White Papers for Five-Year Strategic Plan (FYSP) for Federal NITRD Program

(August 2008)

INPUT FOR NITRD FIVE-YEAR PLAN AUGUST 2008

Cyber-Physical Systems

This paper could apply to any area of software practice, but its suggestions are especially important for Cyber-Physical Systems (CPS). The area deserves the effort. By their embedded nature, CPS are low profile, and the area has been under-served by CS/SE research. However, CPS are critical to many US industries, they are knitted into the modern lifestyle, and defective cyber-physical systems can be deadly. This paper focuses on exploiting the potential of interagency projects with the cooperation of industry partners in ways that could ameliorate long-standing problems that are especially troublesome to CPS.

This paper suggests strategies to address two complaints:

1. Researchers argue that their ideas improve the software development process, but their research doesn't provide convincing evidence. They ask us to "bet the company" on new ideas supported by anecdotes and arguments. If they have experimental evidence, it is often based on experience with a few students working on small projects. Almost all embedded software developers are deeply conservative about the way they build software.They need convincing reassurance before they will adopt new technologies.

2. Either researchers don't attack the problems that bother practitioners, they have addressed those problems and the results are not implemented in commercially available software and well-documented methodologies, or wonderful things have been offered to us in attractive packages and we've ignored them. The latter is unfortunately likely. (See the first complaint.)

Validate Research Results Using Real Software Development

Any large software project could be an opportunity to test and compare development tools and methodologies. Some large software projects have been used to test new ideas, but I don't know of a case where a real software project was used to compare competing ideas (except, maybe, the control software for the space shuttle). The data generated by a comparison is more useful than the data from a test. To take an example that would be of special interest to me, a test can show that the RTSJ is useful for a large real-time software project, but a comparison would show its strengths and weaknesses relative to alternative technologies.

A software development project can be used to generate comparative data by duplicating the appropriate part of the project in at least two separate and equivalent sets of groups using different tools/methodologies/whatever. To be statistically valid, there should be several groups in each set. [1] This approach can be used to get experimental data on many questions that trouble software engineers, but it will be expensive. A single instance of a "real" software project may cost

[1] Maybe statistical trickery can let one experiment answer several questions, giving better value for these experiments.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

millions of dollars. Replicating it *n* times with suitable experimental measurement and isolation could cost more than *n* times the cost of a single instance.

Interagency cooperation, and cooperation with industry, may make suitable projects available for this kind of research. DoD, automotive companies, NASA, aerospace companies, DOE, and FAA have large, demanding software projects that would be excellent candidates for experimentation. Other problems include: finding worthwhile questions whose answers are measurable (statistics only helps when there's something meaningful to measure), and convincing the owners of the projects to let them be used this way.

This type of experiment seems likely to be most useful in two scenarios: evaluating broad research directions early in the R&D process, and demonstrating the usefulness of ideas that seem powerful to the research community but are not immediately accepted by practitioners.

Build Research and Development Communities

From a pragmatic viewpoint, research doesn't matter unless it improves real software products, or makes software production more efficient. From this viewpoint some aspects of operating system software, compiler software, software project management, data base technology, and networking get attention. Other important areas are mainly neglected.

I spend much of my life testing and debugging software, but over the past 25 years few breakthroughs have appeared in the testing/debugging software that I use. I wish those products were advancing faster. If CS/SE R&D were proportioned according to my workload, about half its effort would be directed at improving testing and debugging technology.

It might help focus researchers' efforts if CS/SE research groups were tightly coupled to users. Tying the groups together would make it easy for the researchers to understand the problems facing developers, and create a tight feedback loop that might engender useful new research and uncover ideas that have been composting in research libraries. A "real" software project might resist being asked to use tools with less-than-professional polish. That problem can be lessened by including a professional software company in the community.

I envision a systems group from CMU dedicated to providing improved systems software for the Mars lander team at JPL, with the JPL team having a matching commitment to depend on systems software from the CMU team, and a software product team from IBM putting a professional polish on the CMU software and providing first rate documentation and support. It might be best if the three groups were located together.

The teams should collaborate through several generations of the product.

Not all important work is done by large groups. It would be hard, but useful, to build communities to focus researchers' attention on tools and methodologies for development teams with fewer than six members.

Networking and Information Technology Research and Development
NITRD - A Testing Perspective

New applications over computer networks appear every day ranging from simple message systems to VOIP and IPTV. Also, a multitude of new protocols for ad hoc wireless network have been proposed. Despite all the improvements on network communication systems, little has been done to advance the way these new applications/protocols are tested. Most existing work on network testing is restricted to protocol conformance and many criteria have been proposed to assess the quality of a test suite. However, those criteria focus on the node level and not on the network level. For example, state coverage is used to assess how many states of a protocol have been exercised by a determined test suite. The same applies to transition coverage. While these criteria are important, they do not provide an overall indication about the quality of a test suite at the network level. Some criteria have been proposed to account for all possible interactions between a communication protocol for end-to-end nodes. While improving over state and transition coverage, they do not provide assessment criteria at the network level. If testing an isolated node or a couple of nodes was sufficient, then there would be no need for large test beds such as DETER, PlanetLab and simulation systems such as ns2, where experiments with thousands of nodes can be conducted.

One question to be answered when testing any system is “How adequate is the test suite?”, that is “How can researchers qualify their network experiments?” If the adequacy of a test suite cannot be properly measured, then it is very unlikely that someone will have high confidence in the quality of the product to be released. The same applies for network testing where a network experiment or set of experiments is analogous to a test suite. Basically, three measurements have been used to determine the adequacy of a test suite for network testing: scalability; execution time; and confidence interval. A myriad of research papers and industrial experiments use these factors as the main aspects to evaluate their test suites. These factors form an important subpart but do not constitute the entire picture of test suite adequacy. For example, suppose an experiment is conducted using a thousand nodes executing over a period of five days (120 hours) using some appropriate random traffic generator. Would such an experiment be “adequate” for exercising the network? In general, there can be no definitive answer. Further, the same events could be generated over and over again and many possible scenarios could be left untouched by the experiment. The experiment may be repeated many times and a 99% confidence interval is computed. Even in this case, there is no guarantee that the experiments are not biased. This is a clear indication of the neccessity of better testing tools and techniques for networking experiments. The identification of multi-layer or cross-layer assessment criteria along with automatic instrumentation of source code with respect to the identified criterion would bring the current practise in networking testing to a much higher level. Although one cannot validate other's experimental results the availability of a proper assessment criteria gives an indication of how much one can trust the produced results.

Also, not many experiments are conducted using test beds such as DETER or PlanetLab. One of the main reasons is the effort require to deploy and execute an experiment on these test beds. There is a clear need for tools that can help on the (re)deployment of network experiments. The availability of such tools would encourage the testing of proposed networking solutions using more real life environments and would consequently improve their quality. The tool should also facilitate the reconfiguration of the environment to make possible the testing of more diverse scenarios.

Another problem to be addressed is to avoid unnecessary resource usage. Many experiments are programmed to run for many days or even weeks but they may reach a saturation point (no progress in terms of testing new scenarios; rate of change of coverage is zero) earlier in this period. In this case, all the rest of the time and resources are being wasted as no new scenarios are tested. Therefore techniques to identify the "saturation point" as well as to dynamically tune the experiments to regain the testing of new scenarios would result in substantial savings in terms of time (no need to abort and re-start the experiment) and resource usage. This goal can be achieved by: (i) dynamically modeling the behavior of the networking experiment; (ii) automatically identifying the inputs and variables affecting the resource under observation; and (iii) tuning these inputs to regain increase in coverage.

In summary, the following itens are needed to improve the current status of network testing:

- a tool to facilitate deployment of experiments on easily reconfigurable test beds.
- a cross-layer or multi-layer network assessment criteria and correspondent coverage analysis tools.
- a technique to identify the saturation points during execution of experiments and dynamically tune the experiments.

A multi-agency collaboration is keen to realize the proper development of the testing tools/techniques listed above. First, deployment tools may need to follow a specific standard or language. The individual development of them by each agency would lead to a myriad of distinct languages and would defy the goal of easy deployment in any given scenario. Second, assessment criteria defined according to specific needs of an agency would be weaker than a more comprehensive set of criteria defined in a global manner. Third, some agencies may have a focus on some specific networking needs and the combined effort would lead to the development of tools with a broader scope and consequently more effective testing tools. Finally, testing tools are general purpose and the multi-agency development would incur substantial savings.

INPUT TO THE FIVE-YEAR STRATEGIC PLAN FOR THE FEDERAL NETWORKING AND INFORMATION TECHNOLOGY RESEARCH AND DEVELOPMENT PROGRAM

August 15, 2008

Networking and Information Technology (NIT) is a dynamic field that transforms society, science and industry. It is not surprise that the center of gravity in this dynamic field continue shifting driven by changing societal challenges and emerging technological opportunities. I believe that we are in the right position now to identify, articulate and analyze the consequences of a new realignment in the field driven by the appearance of Cyber Physical Systems (CPS). In my input to the Federal Strategic Planning process, I intend to expose some elements of this new area and argue that the structure of the Federal NITRD investment need to be modified.

1 Future needs for NIT capabilities

NIT, once a provider of tools for the sciences and engineering, has become a uniquely interdisciplinary and pervasive field at the very core of the scientific and industrial innovation in the USA. We are surrounded by the results of a massive multidisciplinary R&D activity in such applications as analysis, and visualization capabilities for medical diagnosis; robotic exploration of Mars; computerized control in cars; international air-traffic control systems; computerized monitoring and control of industrial process. The "NIT explosion" is an auto-catalytic process, which simultaneously proceeds in many direction. The effects of this process are profound:

- The most exciting developments occur at the intersection of NIT with other areas. There are many examples for new science and technology fields, such as bioinformatics, smart materials, distributed robotics that turn into new disciplines and have the potential to jumpstart new industries.
- IT is becoming so pervasive that the classical structure of IT research and industry is changing drastically. For example, the tight integration of physical and information processes in embedded systems requires the development of a new systems science, which is simultaneously computational and physical. This ultimately leads to a new education and project management structure, which is very different from the existing models.

Arguably, the late eighties and early nineties were the "Big Bang" of IT. Triggered by the Strategic Computing Initiative of DARPA, Federal NIT investment created research breakthroughs in parallel and high performance computing. During the nineties, the center of transformational activity of NIT shifted to networking and the Internet. Starting from the mid-nineties, a new, less obvious, nevertheless perhaps the most pervasive expansion of NIT started emerging; the fusion of information processing with physical processes – called Cyber Physical Systems. CPS literally changes the physical world around us. From electric shavers to airplanes and from cars to factory robots, computers monitor and control our physical environment. NIT is rapidly taking over the role of being the universal system integrator for physical systems of all size. This trend is based on a fundamental technical reason: NIT is uniquely suited for implementing and controlling complex interactions among physical system components.

I believe that the deep integration of NIT with our physical environment will be the center of gravity of NIT in the next decade. This is a profound revolution that transforms entire industrial sectors into producers of cyber-physical systems. The deep integration is more than adding computing and communication equipment to conventional products where both sides maintain separate identities. The result is about creating new capabilities that fundamentally changes product capabilities and quality. CPS has extraordinary significance for the future of the U.S. industry. There is much more at stake than extending our leadership in NIT to an exploding new market segment. Falling behind in the foundations of CPS may render our scientific and technological infrastructure obsolete, leading to rapid loss in our competitiveness in major industrial segments including automotive, aerospace, defense, industrial automation, health/medical equipment, critical infrastructure and defense. We are in the midst of a pervasive, profound shift in the way humans engineer physical systems and manage their physical environment.

2 The Role of the NITRD Program

Federal investment in NITRD is particularly critical in periods of major transformations in technology. Federal investment has major role in stimulating basic research, attracting attention to solving fundamental challenges and helping the transformation of the research infrastructure.

I fully agree with the PCAST Report that the current prioritization of Federal NITRD does not reflect the fundamental shift in the center of gravity of NIT transformational activities. Roughly half of the federal dollars are invested in high performance computing (HEC I&A and HEC R&D), while CPS is not even represented as a crosscut.

I recommend changing this and - as recommended by PCAST - place CPS as first priority and Software as second priority for NITRD.

In order to increase the relevance of Federal NITRD in improving US Competitiveness, I have the following specific recommendations:

1. *CPS needs to be identified as a separate crosscut.* For example, this may happen by changing HCSS to CPS with a drastically extended role covering the full scope of CPS research. This is justified, since high-confidence system and software is essential and inherent part of the CPS research needs. While I was at DARPA between 1999 and 2002, I served as co-chair of SDP and saw clearly the significance of *clear and explicit* articulation of priorities. Without this change, NCO will not be able to coordinate CPS investment along a well defined research agenda.

2. The CPS research portfolio need to include cyber security and networking but driven by the specific needs and circumstances of cyber-physical integration.
3. The closest (although much narrower) existing crosscut, HCSS, receives 3.7% of the NITRD investment. This seems to be quite misaligned with the proposed priorities. Just to get in rough parity with the EU programs would require about $1B/ year investment in the area.
4. Software, which is currently included in the SDP crosscut, receives only 2.1% of the NITRD investment. The lack of investment in software research is a huge problem that impacts not only CPS but all NIT application areas.

3 Key Challenges for a CPS crosscut NITRD

As part of the community working on the formulation of a CPS research agenda, I concur with the main research challenges captured in the CPS Executive Summary Document[1]. The following table shows my perception about the relevance of the different areas to NITRD Agencies:

	New abstraction layers for design	**Semantic foundations for composing models**	**Composition platforms for heterogeneous systems**	**Predictability under limited compositionality**	**Foundation for system integration**	**Compositional certification**	**Agile design automation**	**Open Architectures**	**Reliable systems from unreliable components**	**Resiliency to cyber attacks**
NSF	■	■	■	■	■	■	■	■	■	■
DARPA	■	■	■	■	■	■	■	■	■	■
OSD and Services	■	■	■	■	■	■	■	■	■	■
DOE	■	■	■	■	■	○	○	■	■	■
NSA	□	□	■	■	□	□	□	■	■	■
NASA	■	■	■	■	■	■	■	■	■	□
NIST	■	■	□	□	□	■	■	■	■	■
AHRQ	○	○	□	○	○	■	○	■	○	■
DOE/NNSA	■	■	■	■	■	○	○	■	■	■
NOAA	○	○	□	○	○	○	○	□	□	○
EPA	○	○	□	○	□	○	○	□	□	○
NARA	○	○	○	○	○	○	○	○	○	○

(■: very important; □: important; ○: no information)

The strong overlap in relevance shows the importance of establishing interagency coordination.

[1] CPS Steering Group: Cyber-Physical Systems Executive Summary. March 6, 2008

4 Role of International Collaboration

NITRD is a worldwide enterprise and US is part of the international efforts. The chance for establishing substantial and mutually beneficial international collaboration and leverage large depends on the significance of the different areas on industrial competitiveness and national security. Not surprisingly, those NIT areas that are highly precompetitive and serve more the general scientific progress have a better chance for leveraging international efforts.

Based on this, my assessment for the areas where international collaboration is highly feasible and beneficial are: HEC I&A, HEC R&D, HCI & IM, and SEW.

The areas where some international collaboration in foundations are possible but direct relevance to competitiveness and national security decreases opportunity for leveraging investments are: HCSS, CSIA, LSN and ASDP.

CPS is clearly an area where the international competition is extremely high. For example, a declared goal of the Artemis program is to increase EU leadership in this area, therefore the research projects are tightly controlled and international collaboration is much harder. In certain basic research areas, such as Agile Design Automation or Resiliency to Cyber Attacks the mutual interests can be aligned for establishing collaborative programs, but the precondition for this is parity in investment.

5 Industry/Government/Academy Partnership and Technology Transitioning

A unique aspect of CPS is that the US Systems Industry (aerospace, automotive, process, automation, health energy, defense) recognizes the huge importance of the area in future competitiveness and expressed willingness to co-invest and create new partnership constructs. It is essential to ensure that a much tighter collaboration is created between industry and academia, where industry challenges more directly inspire research and academic solution are more easily validated in industrial strength testbeds. Willingness of industry and academia to create new forms of collaboration need to be exploited.

The Institute for Information Infrastructure Protection is pleased to have this opportunity to share our thoughts with the Federal NITRD program on the Five-Year Strategic Plan. As a consortium of academic institutions, non-profit, and national laboratories, the I3P provides a unique perspective on the Networking and IT R&D agenda.

The N&IT challenges facing the government mirror many of those faced by industry. Both have identity and authentication, intrusion detection, and vulnerability analysis challenges, to name a few. However, the Federal agencies in the NITRD program face other challenges that are unique to government agencies, or have different or additional characteristics when compared to industry. We will briefly discuss three of these in the limited space allowed in this RFI response.

Attribution in cyber space

Attribution, in the N&IT arena, was defined by an IDA report (IDA Paper P-3792, October 2003) as "determining the identity or location of an attacker or an attacker's intermediary." To that we add "even when the intermediary may or may not be a willing participant in the attack" in our I3P white paper on "The role of cyber attack attribution" (see http://www.thei3p.org/docs/publications/whitepaper-attribution.pdf). This is a particularly vexing problem, as no real path towards reliable attribution was included in the various communications protocols used on the Internet. National defense concerns and capabilities aside, the ability to reliably attribute arbitrary Internet traffic to its origin, or some small set of origins, is of increasing interest to the industry as well as to the member agencies of NITRD. However, this capability must be harmonized with the privacy interests of the public from the start. Furthermore, this cannot be left as just a law enforcement challenge since the private sector has interest in it for preventing fraud and the law enforcement agencies around the world cannot possibly handle all of the demand.

While we agree that full attribution may not always be possible or even desirable, we believe that R&D towards the capability of attribution in cyber space should be a high priority of the NITRD. We do not advocate the position of a specific technological or policy solution to address this challenge. Rather, we see the future of attribution in cyber space as a gradual co-evolution of technology, socialization, policy, and law. This multi-disciplinary challenge is one at which both the NITRD and the I3P excel.

Protection of medical information

By necessity of our mobile population, longer life expectancies, and great leaps in medical technology, the medical records and ongoing medical treatments of citizens are increasingly dependent upon cyber space. In the past, this discussion has focused on standardizing medical record formats and the confidentiality of their exchange. However, in the next few years that focus will likely be replaced by the concerns for the veracity of medical information as well as its protection from unauthorized update. These new concerns are not limited to the traditional medical data at rest or during exchange, but will include such situations as quickly providing personal medical information during a disaster response or the gathering of live medical data via telemetry outside of a medical institution. This live information has the same privacy and accuracy concerns as the traditional medical data but, by its very nature, telemetry data is more vulnerable to interception, disruption, or other interference. Both the NIH and the AHRQ have taken some steps towards these concerns, and we believe that this growing concern deserves much further study from the technology, policy, education, and legal perspectives.

People: the ultimate critical infrastructure

The public and private sector have taken many steps to ensure the reliability and security of everyday activities on both private intranets and the open Internet. The development and application of hundreds of technologies, books of policies and procedures, and many hours of user education have achieved today's level of reliability and security. This limited success is due, in part, to a failure to realize that efforts towards increasing the reliability and security of any system must be focused upon the users of that system – users of systems make the final decision on how to use them. In the current case, this missing realization is that the people using our computer systems and networks are indeed the ultimate critical infrastructure of all public and private organizations. Whether talking about a malicious insider or an under-educated innocent, the everyday decisions made by individuals around the world clearly affect the reliability and security of the systems we all depend on.

This is indeed a very hard problem to address, requiring the interplay of a broad collection of disciplines. However, the actions taken to date have failed to provide the level of reliability and security that will be needed as the world economy grows even more dependent upon cyber space. Well-known private sector concerns about process control systems (including public utilities), supply chains, and the availability and safeguarding of medical information, combined with the innumerable other critical systems inside of private and public institutions, truly makes the case that the people using these systems should be the focus of a strong R&D agenda across all of the member agencies of the NITRD as well as the private sector.

The Institute for Information Infrastructure Protection (The I3P) is a consortium of leading universities, national laboratories and nonprofit institutions dedicated to strengthening the cyber infrastructure of the United States. See www.i3p.org for more information.

Autonomic Management of Next Generation Wireless Networks for Ubiquitous Services

(A White Paper submitted to NITRD Program, Aug 25, 2008)

1 INTRODUCTION

1.1 Background. Recent advances in the areas of Web services, mobile wireless communication and networking, smart devices, sensors and embedded systems promise interactions and services that have never been experienced before. As the advances in information and communications technology are riding Moore's Law, the ability to share services and resources has also witnessed exponential growth. In near future, such advances will lead to *ubiquitous service spaces* based on mobile and pervasive computing paradigm, that encompass data, computational resources, and services distributed over heterogeneous (wireline and wireless), pervasive network infrastructures. This will offer tremendous opportunities in a variety of novel and attractive application domains such as environmental monitoring and control, advanced automotive systems, critical infrastructure control, reliable healthcare systems, pervasive security, community based computing, social networking, cloud computing, and so on.

In these applications, typically a large-scale wireless networking system consisting of massive numbers of connected elements (e.g., devices, processors, sensors, and actuators) is designated to play a crucial role. The high complexity of such heterogeneous, dynamic systems raises new challenges in the science of design. These systems must be (i) *context-aware* and *self-adaptive* to uncertain internal and external environments to achieve reliable and robust performance, (ii) *self-managing* to prevent the escalation of maintenance cost, (iii) capable of *self-optimizing* the performance in different application scenarios, and (iv) capable of *self-organization* in order to enable rapid deployment and reconfiguration of the network infrastructure. This will enable the systems to respond to changes in the environment as well as the needs of the network operators, service providers and subscribers (application users).

1.2 Motivation. Current computing systems are rarely transparent and adaptive. They rely on human understanding of different computing platforms and technologies and often extensive manual programming to compose and aggregate services. This is clearly unrealistic in large-scale, highly unpredictable and dynamically evolving complex systems.

Uncertainty is the defining characteristic of mobile and pervasive computing networking systems and poses stiff challenges to the seamless functionality of ubiquitous services and applications. The uncertainty appears in various facets, for example at the physical layer (uncertainty in stochastic wireless channels and scarce spectrum), at the network layer (uncertainty in topology due to user mobility and wireless bandwidth availability) as well as at the application layer (uncertainty in traffic load and resource demands, application profiles and quality of service),. Today's mobile devices come with a number of embedded communication technologies such as 2G/3G/4G cellular networks, Bluetooth, Wi-Fi, WiMAX, cognitive radio, etc. As an example, in a disaster management scenario, rescue and public safety teams are likely to carry smart devices to communicate with each other. Some of the desirable requirements are the ability to transmit real-time voice, still pictures or low-quality videos, or even remotely actuate rescue equipment. Reliable communication and connectivity among the members of the rescue team play a significant role in the success or failure of the rescue operation and thus directly impact the lives that are in jeopardy. It is relatively easy to install a mesh or a peer-to-peer network to provide a wireless backbone. However, when a disruption occurs, the available wireless devices are unlikely to maintain a well-connected and reliable network. Instead, fractions or *clouds* of devices and network elements will be sporadically connected to each other, and possibly, to the surviving part of the infrastructure. In the extreme case, a single disconnected user may be a degenerate version of a cloud. Due to user mobility and spectrum scarcity, these clouds will be extremely dynamic. As a result, traditional networking approaches will fail to preserve dependable communication since many of them will require continuous end-to-end paths between communicating endpoints, and critical services based on multi-hop routing are unlikely to live up to the expectations. *Cognitive radio* permits dynamic spectrum assignment, and potentially alleviates part of this problem by opportunistically grabbing unused spectra and putting them to use at the time of crisis. Such systems must be spontaneously deployable, able to self-organize (so that nodes can seamlessly join or leave without the need of global control), reach consensus about maintaining the best-quality path in the face of conflicting requirements, and be able to self-heal following node or link or software failures. They must also be able to survive in spite of disruptive incidents and/or attacks by opportunistically exploiting and dynamically reconfiguring the available network components and resources. In other words, these systems must be *autonomic* whose holy grail is *self-management*.

Dynamic distributed systems go through periods of change and stability. An *adaptive system* reacts to changes in the *environment*, by encapsulating the needs of the users. Besides mobility and resource (e.g., bandwidth) requirement, the cognition and opportunistic self-allocation of unused spectra is an important part of the environment, which has interesting consequences: dynamically allocated parts of a spectrum may have to be released immediately after a designated user of that band appears in the scenario. This will trigger a disruption in communication, and a *self-healing* communication system has to restore functionality. Another significant feature is the need for *self-optimization*: a subset of nodes engaged in communication via a resource-constrained path must be able to spontaneously improve the communication quality in case resources like a specific band of the spectrum capable of providing better functionality becomes available, or a mobile node appears in the neighborhood promising to lower the latency. The scale of the application will have additional impacts on the functioning of such systems. A small-scale system deployed in a tornado-affected neighborhood will find it easier to abide by a set of policies established by a single administrator of the equipments. However, large-scale applications overseeing relief work in an earthquake-hit metropolitan area will have additional challenges: for example, the latency of some tasks may be unacceptable due to large number of routing hops. Also, the selfish behaviors of nodes under different administrative controls have the potential to adversely affect the overall performance in certain areas, as demonstrated by game theorists.

2 THE RESEARCH CHALLENGES

The diverse nature of the activities enabled by the next generation wireless networks and ubiquitous service requirements defines a grand challenge in computing whose investigations will require meaningful collaboration between research in various disciplines including mobile, and pervasive computing, autonomic computing, and adaptive service computing. Below, we highlight five fundamental challenges of next generation ubiquitous service spaces.

2.1 Challenge 1: Uncertainty Management

Uncertainty is the major driving force and perhaps the over-arching principle guiding an autonomic framework. Indeed, the uncertainty associated with the system produces unique challenges to achieving survivable communication services providing acceptable end-to-end quality of service (QoS). This calls for the design of new adaptive protocols to tame the uncertainty at various levels. Currently, uncertainties in wireless mobile networks are perceived mostly at the physical communication layer (due to rapidly varying wireless link qualities, or variable points of network attachment for mobile users). How to identify various uncertainties and their impact on the networking layers, as well as on the application goals? How to provide fault-tolerance and survivability when disconnections, disruptive incidents or attacks occur in an uncertain service support environment?

2.2 Challenge 2: Context-Awareness and Situation Modeling

Uncertainty management requires awareness of the contexts of the participating entities (such as devices or other system components), as well as the context of the entire environment or situation. A *context* is any relevant attribute of a device that provides information about its interaction with other devices and/or its surrounding environment at any instant of time. A sequence of device/entity contexts with the underlying interpretation (semantics) defines a *situation*. The understanding and analysis of the behavior of a system is paramount to capturing contexts unambiguously in the presence of uncertain (noise) and incomplete information. Past profiles help the system components anticipate future disruption and choose the best recovery action. How to generate *context awareness* and make the system *situation aware*?

2.3 Challenge 3: Autonomic Service Discovery

A crucial facet of self-management in mobile and ubiquitous computing environments is *service discovery*. Without this, nodes in infrastructure-less and dynamic environments such as mobile ad hoc networks (MANET) or vehicular ad hoc networks (VANET), will be confined to using only their own services and resources (or those that are pre-configured statically by system administrators). Even though many service discovery protocols and architectures have been proposed for such volatile environments, most are far from being considered autonomic. How to perform autonomic service discovery to cater to the needs of a volatile environment? How to utilize the promise of the newer tools like cognitive radios for efficient and expeditious autonomic discovery and recovery?

2.4 Challenge 4: Autonomic Service and Resource Management

Flexible and seamless configuration and delivery of services in large, autonomous, and complex evolving service spaces is another major goal. To achieve this objective, two strongly coupled aspects deserve attention: (i) composite services engineering, and (ii) resource provisioning. How to efficiently and dynamically manage the spectrum (bands) in cognitive radios? How to improve the QoS by preventing the over-provisioning of scarce resources (bandwidth) and using learning algorithms to profile and anticipate future resource usage, so that the resulting system performance is optimal or near optimal? How to facilitate efficient on-demand composite services configuration and protect users with guaranteed QoS despite variations in user activities, services spaces, and network infrastructures?

2.5 Challenge 5: Self-*Algorithms

The driving force behind successful self-management is the design of specific algorithms for different management tasks. How to design efficient algorithms for self-healing, self-optimization, self-stabilization, self-protection, fault-containment and graceful degradation to improve the availability and maximize the functionality of the applications?

While generic algorithmic tools for some of these are available, their applicability varies from one scenario to another. Accordingly, there is a pressing need to redesign them to suit the specific application needs. Two components are likely to improve the effectiveness of these algorithms. The first is the static data about uncertainly profiles collected from typical applications of the same class – this will help shape the ground rules or policies. The second is the learning mechanism using which the system becomes smarter with time: thus if the same fault repeats, or the same attack is launched several times, then ideally the system should be able to recover faster or protect itself more expeditiously with every new episode. How to add learning components to the various algorithms that are the cornerstones of autonomic computing?

3 CONCLUSION

Autonomic behavior of complex network systems consists of an assortment of *self-** (self-star) properties that guarantee dependability, availability and continued functionality. The proposed agenda will spawn further research not only in mobile/ pervasive and autonomic computing, but also in distributed algorithms, fault tolerance, security, game theory, and information theory. Eventual immunity to malicious actions is an issue that will continue to challenge the best minds as crooks invent new forms of abuse. The proposed framework must integrate different technologies (e.g., sensors, mobile devices, wireless mesh) prevailing in the (next generation) pervasive networking infrastructures. It will also be a crucial test bed for cognitive radio based applications that rely on the opportunistic harnessing of spectra. Ubiquitous (groupware) communication services will address real-time/interactive streaming (voice), interactive applications, mobile instant messaging, off-line streaming, chat, e-mail, data sharing, environment monitoring, and provide the assured quality of service that will strike a balance between security / privacy and usability.

Response to NITRD RFI
White Paper to Inform the Five-Year Strategic Plan
Collective Intelligence

Internet search technology has created an environment in which any user has access to nearly unbounded amounts of information about any selected topic. Two important aspects of search are that (1) the search engine should consider the largest possible collection of relevant information, and (2) the search engine should select the best match of points in this information space with the search target. Secondarily, the search recommendations should also be of “high quality” (meaning it is significant, verifiable, factual, and correct content). Generally speaking, after searching for information each user has to vet the information recommended by a search engine to determine its *relevance* with respect to then intent of the search, and to evaluate the *veracity* of the information.

The overhead work of vetting the information – both with respect to its relevance and veracity – can be significant. Any such vetting work can be abstracted and saved as *persistent state* in the user’s search environment, e.g., bookmarks of locations are persistent state representing the location of pertinent & reliable information. Bookmarks effectively store this state in a simple hierarchical name space; bookmarks can be saved in a root node, or in child subdirectories (folders). For example, the user might be a university teacher, a software researcher, a member of the Chandler family, a tourist, a chess player, a John Mayer fan, and a luthier: the user can create a folder for each of these interests in which to store topic-specific bookmarks.

While it is not commonly done, it would be possible for the system to exploit this state, including its structure, to improve the search function. First, assume that the folder hierarchy is organized so that each folder represents one of the user’s *perspectives* of the web (the persistent *substate* associated with each folder represents information about a particular view of the web as, say, a luthier). This substate information can be exploited by having the user select a perspective prior to searching; the overall search process (the search engine, possibly with supplementary software) then uses this filter to locate and discriminate among information normally returned by a search engine.

Bookmarks are only one example of the information that might make up a perspective state, and search is only one example of the NIT capability that uses such information. Any kind of information that the user/system can collect, encode, organize, and save can be part of a perspective state. Such generalization enables the system to use the state as the basis for broader tasks. For example, information stored in cookies could be part of a perspective state; social network relationships can be part of the state; and information extracted from RSS feeds (e.g., photos and videos that friends favor) could be saved in the state. The logical extent of the state is bounded by the nature of the data collection tools, the data organization methods, and the nature of the data analysis algorithms.

Some of these perspectives are unique to the user, but others are likely to be valuable to other users that have similar interests. That is, users with similar perspectives can leverage their effort to refine their perspective by collaboratively constructing and sharing state. The exemplar user might share some state with other luthiers, other state with other *guitar* luthiers, other state with other *acoustic* guitar luthiers, etc. Social network technology provides manual means (friends, groups, and networks) for creating, maintaining, elementary sharing among people with like interests (i.e., with like perspectives). In an attempt to choose terminology that generalizes such relationships, we characterize such shared perspectives as defining *groups*, *teams*, *communities* or *tribes* (all of these terms have been used by people building social network software). I imagine a NIT capability that maintains perspectives with programmable state for each user, that supports networks (or at least hierarchies) of perspectives, and that has a well-conceived state sharing model. Such a system requires a carefully designed and widely accepted structure (e.g., a standard) for representing state, a means of managing communities and their shared states, and a refinement in the way application tools (such as search) use the shared state.

This input focuses on an area of study in which users work in groups to accrue shared value – the implicit knowledge in the shared state – from which each team member derives benefit. This approach correlates with that of *collective intelligence*;[1] in this RFI "intelligence" is described as an evolved state models that incorporates potentially sophisticated data structures to represent structured (possibly hierarchical) and shared knowledge.

I imagine a new set of information processing capabilities in future NIT that facilitates the identification and preservation of group knowledge that represents the group's collective intelligence. I also expect that research in the area of *harvesting* collective intelligence will become an increasingly important component of the new NIT environment. The technology will help to: form groups (as in a social network); to structure and archive information of high interest to all/most group members; assist group members in using the knowledge implied by the collected information, and enable others to provide tailored services and goods to groups. There is already a clear commercial trend in this general direction (e.g., see FriendFeed.com), thus an NITRD effort is needed to coordinate the research frontiers of the work at the same time that it encourages technology transfer and commercial development of the capability.

An ominous aspect of this work relates to cybersecurity: such technologies are also likely to enable (antagonistic) third parties to observe the group information. Because of the implications regarding invasion of privacy, identity theft, etc., it seems clear that while harvesting information can enable highly refined information harvesting and group support, it is likely to be important to have technology, standards, policies, and legislative actions that address security related to the technology. For example, it is important to understand how to create NIT that enables harvesting of collective intelligence in the aggregate without using/storing information about individuals in a group. This aspect of collective intelligence is likely to affect many different NITRD agencies. The evolution/revolution of collective intelligence is deeply intertwined with cybersecurity – what is the correct balance between the ability to exploit collective intelligence without violating social, customary, and legal barriers? I imagine that to the extent collective intelligence emerges as an important NIT, it will necessarily be associated with a carefully thought out cybersecurity strategy.

I believe that the idea of collective intelligence is fundamental in contemporary and evolving NIT; as a consequence, it seems important that part of the NITRD charter should encourage, monitor, influence, and support the evolution of the area as it evolves. Information retrieval for the masses is an important multi-disciplinary application for the NIT infrastructure. Work in this area spans networks, systems, security, software, web programming technology, AI, anthropology, ethnography, and more. Significant research and technology improvements are needed to exploit the notion of collective intelligence.

[1] See http://cci.mit.edu/ describing the multi disciplinary M.I.T. Center for Collective Intelligence, formed in 2006 to study the emerging concept of collective intelligence.

August 25, 2008

To Whom It May Concern,

Computing and communications are becoming inextricably linked to our physical world. New consumer and industrial products include some form of computing to enhance functionality and to better control processes. For instance, modern cars would not be able to satisfy emission control requirements, would not be able to travel safely on slippery roads, and would not be able to inform the drivers about traffic congestion and alternative routes without embedded and networked computing devices. These *cyber-physical systems* that integrate the physical with the computational to provide new and improved functions are clearly at the forefront of technological development, and their quality, safety, and dependability is of utmost importance.

The electronics industry keeps coming up with new components (processors, networks, interfaces), but these devices are only enablers of new functions. To implement new functions (like stability control on cars) domain engineers from different disciplines, including mechanical engineering, electrical engineering, software engineering, and industrial engineering have to work together – in a truly interdisciplinary manner -- to build these systems. New aspects, like cybersecurity, human interfaces and cognition, and total energy management are being recognized as important factors for quality and dependability as well.

It seems that competitiveness in any industry that manufactures physical products (vehicles, consumer devices, healthcare equipment, etc.) depends on how well it can take advantage of the computational ('cyber') technology and can incorporate that into its products. However, the cyber-physical systems change the design paradigm: systems have to be created with deep understanding of the interdependencies between the physical and the computational. Designers have to be trained to follow a 'systems' thinking and understand the subtleties of cyber-physical systems, where computing and physics (and biology and chemistry and psychology and economy and so on) interact. We need new foundations for the science and engineering of cyber-physical systems that is applicable across many industries.

NITRD could play an essential coordinating role to make this happen. By providing coordination across government agencies, it is in a unique position to ensure that the science and engineering of cyber-physical systems becomes a reality, and that foundations developed are used in real, potentially societal-scale systems. The table below attempts to provide a partial list of that could play a specific role in CPS.

Agency	Potential topics the agency could support in a CPS effort
NSF	Fundamental science of cyber-physical systems Engineering foundations for systems integration; Education
NASA	Cyber-physical systems in the aerospace area
FAA	High-consequence and autonomous systems
NSA	Cybersecurity issues in CPS
DARPA/DoD	Defense applications; extreme-scale CPS
DOE	Energy systems as CPS
NIST	Joint academia-industry projects with engineering applications
NIH, FDA	Medical devices as CPS

One can envision multi-agency programs, where a specific *‘grand challenge’ problem* is defined, and teams that include participants from academia, industry, and government work on the problem in a competitive manner, are funded by various agencies, and solve basic science and engineering, as well as applied engineering problems.

An organization like NITRD could also foster *academia/industry partnerships*. For example, on a joint, academia/industry project the academic part could be funded by NSF, and the industrial part could be funded by NIST. The results that include academic results, but also prototype designs and implementations should be made ‘open source’, for public use.

NITRD could also help the coordination by providing a ‘*clearinghouse*’ of CPS research results and application examples. Along the lines of the ESCHER effort a community of CPS researchers could be created that allows the rapid dissemination of results and tools.

Publicly funded R&D activities should be done in an open environment, with industry participation in a pre-competitive phase. Organizations that assist with the dissemination and rapid transitioning of the results of government investment should be supported. Past examples include: BSD Unix and ESCHER.

INTRODUCTION – This white paper is in response to the Request for Input (RFI) posted by the National Science Foundation (NSF) on 21 July 2008 to inform the five-year strategic plan for the Federal Networking and Information Technology Research & Development (NITRD) program.

The NITRD research agenda, as evidenced by the existing strategic plan, does span most of the foundational research areas in the NITRD domain. However, we feel that the relative emphasis given to those areas should be adjusted, and in particular, the cross cutting area of Cyber Physical Systems (CPS) needs to be given explicit and increased attention and budget transparency. The existing strategic plan does not identify CPS as an area, though aspects of CPS research could find a home in the existing areas, and more critically, the cross cutting challenges presented by CPS can not be adequately addressed without a holistic approach involving aspects of several of the current areas. For example, by their nature CPS tend to be, or have elements that are, safety critical, and similarly have significant security requirements. Accordingly, foundational CPS research needs to include integral high confidence, assurance, and security dimensions, rather than developing solutions and then attempting to add security, safety and assurance after the fact.

NITRD PRIORITIES – From our perspective the R&D objectives, as indicated by funding levels, are not optimally prioritized. Nearly 50% of the FY 2008 and 2009 NITRD budgets ($1.5B out of $3.3B in FY 08) are allocated to High End Computing (HEC), including Architecture, Infrastructure, and R&D. HEC is not at present an area where we feel US competitiveness is at stake. High levels of HEC funding appear to be institutional priorities of a past era. Expenditures for Human Computer Interaction and Information Management ($0.8B) also appear out of proportion relative to the need and potential gains in research and competitiveness to be attained. The Presidential Committee of Advisors on Science and Technology (PCAST) correctly pointed out the need to substantially increase the level of spending on CPS – which is not even explicitly mentioned among the programs in NITRD budget documents.

We are also concerned about the isolation of Cyber Security and Information Assurance (CSIA) from the systems domains of Human-Computer Interaction and Information Management (HCI&IM); Large Scale Networking (LSN); High Confidence Software and Systems, Social, Economic and Workforce Implications of IT (SEW); and Software Design and Productivity (SDP). CPS must include an essential CSIA program element because of the unique vulnerabilities and consequences associated with the target industries. What we need is CPS focused R&D in CSIA, tightly integrated with all other research challenges.

INDUSTRY / ACADEMIC / GOVERNMENT PARTNERSHIP – We believe that a public-private research partnership to advance the capabilities of cyber-physical systems, analogous to the European Union's ARTEMIS Embedded Computing Systems Initiative, is one way of addressing the CPS research challenge, and this could be achieved by creating Industry / University Consortia to perform pre-competitive research on industry-provided test beds. The "industrial strength" fidelity of the test beds is critical to ensuring that the research focuses on the highest payback elements of the problem space of cyber-physical systems. Consortia focused on more applied levels have been highly successful and include USCAR (U.S. Council for Automotive Research) and AVSI (Aerospace Vehicle Systems Institute). Funding for the consortia could be assembled from: 1) Industry with Internal Research & Development investment; 2) Academia through Government funding; 3) Test bed development through Government funding.

We propose a model based upon joint work of integrated projects as opposed to loose / spontaneous collaborations. While the latter model can sometimes produce important benefits, we believe the focus needs to be the synergistic development of fundamental science directly motivated and evaluated on realistic challenge problems from industry. In this rapidly evolving field where time and resources are limited, this is the most effective way to build a core technology base. Knowledge and technology is best transitioned by people working on well defined problems using industrial strength test beds.

Industry has had a very limited voice in influencing research priorities of NITRD program. Organizations like PCAST have influence at a strategic level but they have little influence in implementation. We believe that proactive industrial participation in shaping NITRD priorities and participation in the research agenda is key to achieving breakthroughs required.

The CPS research agenda is cross cutting and spans multiple industries. Much of the research required is of a pre-competitive nature – where industry sponsored research dollars are inherently limited. The current approach of federal government sponsored research in this area has, to date, been ineffective in both addressing “industrial strength” real-world challenge problems, and creating transition pathways outside of the academic world. Greater industrial participation in executing the research agenda is critical to success and will spur the focused industrial-academic collaboration needed for significant progress. We believe that the Grand Challenge Application approach from the existing strategy has merit, but should have been made more concrete in the form of NITRD sponsored challenge problems and test beds to bring together government, industry, and academia, to provide a means of exploring the cross cutting nature of domains such as CPS and to foster cross fertilization between fundamental research and emerging problems.

An example candidate would be a CPS challenge focused on Autonomous Aerial Vehicles in the Next Generation National Airspace. This challenge problem would exercise all of the elements of CPS, including massive distribution, high assurance and certification, and security. Elements of mixed-criticality functions operating in a common compute platform and challenges associated with migration onto multi-core compute substrates are also of high interest. Such challenges would provide a fertile ground for research grounded in a critical problem whose resolution is essential to the future U.S. national security and economic prosperity.

August 25, 2008

Topic: Bosch response to the Subcommittee on Networking and Information Technology Research and Development (NITRD) Request for Input (RFI)

Bosch has had the opportunity to attend several recent Cyber-Physical Systems (CPS) workshops including the *Workshop on High-Confidence Automotive CPS* (April 07), the *Embedded Systems to Cyber-Physical Systems Workshop* during CPS Week (April 07), and the *Joint Workshop on High Confidence Medical Devices, Software, and Systems (HCMDSS)* and *Medical Device Plug-and-Play (MD PnP) Interoperability* (June 07).

The quality and diversity of research ideas as well as the distinguished list of participants were found to be quite impressive. From these meetings, we have the following observations and suggestions:

Development and execution of multi-agency/multi-disciplinary programs

- Coordination: Since CPS is a broad, multi-disciplinary topic, continued coordination between the complex set of government, academic and industry participants is especially important.
- Discussion: Experts from many fields will have useful contributions to make to CPS. The workshops and break-out sessions help facilitate constructive discussion and goal-setting.
- Awareness: It is important to keep industry aware of the latest CPS developments and funding opportunities. The CPS workshops have addressed this need well so far.
- US/EU Cooperation: To the extent possible, international cooperation is beneficial as many corporations are multi-national. Industry is reluctant to 'pay twice' for the same research in different countries, so a lack of coordination can mean that one country loses out entirely.

Strategic goals, key challenges, opportunities and research priorities

- Good fit to Bosch business: CPS domains are strongly related to current and future products and services sold by Bosch in three business sectors[1]:
    - Automotive Technology (€27.2B / $6.2B)
    - Industrial Technology (€5.9B / $1.3B)
    - Consumer Goods and Building Technology (€11.7B / $2.0B)

  [1]Note: figures are expressed as sales worldwide in billion EUR / North America in billion USD for 2007

- Research challenges: Many research challenges were discussed at the workshops. We found the following challenges relevant and compelling:

  - Predictable system integration through compositionality
  - Architectures and tools for reliable and resilient CPS systems
  - Smart environments and scalable digital services
  - Verification and Validation (V&V) of model-based design
  - Efficient manufacturing and logistics
  - Achieving high confidence and interoperability in medical devices, software, and systems

Transition R&D results; Government, commercial, academic interactions

- Funding opportunities: Bosch engages in many research collaborations with individual universities and multi-university research centers. If government funding is available for the university partner, this helps lower the barrier to entry to setting up collaborations.
- Identifying partners: Ultimately, transfer of technology occurs between individual partners. The CPS meetings facilitate good partnerships.

Examples that illustrate the impact of realizing the vision, achieving the proposed goals, and meeting the identified challenges

- CPS visionaries: Since CPS is a new topic, there are not many completed projects yet. However, the workshops featured many forward-thinking speakers with insightful observations:
  - Dr. Jeannette Wing, National Science Foundation: Cyber-Physical Systems is the #1 priority. Two important classes of systems are coverging: Embedded and Real-Time with Pervasive and Mobile.
  - Dr. Alan Taub, General Motors: Automotive represents the intersection of cyber-physical systems problems in other areas.
  - Dr. Helen Gill, National Science Foundation: In the next five years, embedded computing components are expected to account for a significant percentage of product value; up to 40% in some areas.

In summary, future NIT capabilities are needed to address the business areas and research challenges mentioned above. With respect to roles, the NITRD Program offers value by organizing stakeholder meetings, facilitating networking and providing funding opportunities. The academic community can provide a sound scientific foundation for the advancement of CPS. Since CPS challenges are shared globally, coordination of international agencies is important to avoid potentially wasteful double work on pre-competitive topics.

We look forward to the opportunity to participate in future activities regarding Networking and Information Technology and Cyber-Physical Systems.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

Dear Madam/Sir:

It is a pleasure to share my thoughts on the Five-Year Strategic Plan for the Federal NITRD Program as per your Request for Input of June 30. Let me introduce some of my views by an excerpt from the abstract to a presentation of mine entitled “Engineering cyber-physical systems: a strategic outlook”.

Computation has established itself as the technology of choice to implement system functionality of otherwise unattainable capabilities. The exponential increase in computational power, however, has outpaced advances in design approaches. To bridge this productivity gap, the level of abstraction in design must be raised by exploiting automation. Where automatic code generation has made great strides for Von Neumann architectures, it has yet to scale the concurrency, and, moreover, heterogeneity of emerging electronics platforms necessary to keep pace with Moore’s Law. In other developments, heterogeneous platforms are coming about because of the increasing demand for networked low power mobile devices as well as high-power manipulation such as increasingly autonomous robots. All this is giving rise to a new breed of *cyber-physical systems*, systems that integrate physical with computation and network capabilities.

Our society is quickly and increasingly relying on cyber-physical systems for its technological needs. Six general drivers for networked information systems in general and cyber-physical systems in particular can be identified:

(i) the need to reduce *power* consumption and to provide high power density to implement functionality;

(ii) the rise of *heterogeneity* in computation, in systems, and in design methods;

(iii) the increasing exploitation of *computation* in scientific discovery, in (automated) design, and in product differentiation;

(iv) the omnipresence of *communication* connecting everybody and everything;

(v) the conversion and extension of information terminals to devices for *physical interaction* such as sensors, actuators, and manipulators;

(vi) the emerging sense of *autonomy*, especially in consumer and defense products.

Examples of future systems that rely on these drivers are no-charge mobile devices, smart infrastructure (e.g., power grid and transportation), cyber government, and automated domestic help and health care (where automated could be teleoperated as a stop-gap towards truly autonomous machines). These future systems will be self-adaptive in order to support, for example, self-reconfiguration, self-healing, and self-(re)design. Research challenges to successfully produce such systems can be identified as shown in Table 1. These areas are invariably of a multi-disciplinary nature, involving Computer Science, Electrical Engineering, and Mechanical Engineering at the core with some profound psychological, ethical, and social elements involved. As Table 1 conveys, many different areas within each of these disciplines are involved.

To address the agenda in Table 1, the NITRD Program should provide a strategic framework that outlines the key research challenges. Expertise in the respective areas can then be developed in academic institutes and amalgamated into a selection of comprehensive demonstrator programs. Early multi-disciplinary and cross-institute collaboration should be encouraged.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

Table 1: Cyber-physical systems research and development agenda

Area	Research Goals		Technical Agenda	
Control	·	Computation and communication in control	·	Technology to represent implementation effects
	·	Collaborative control methods	·	Mixed-initiative control technology
	·	Self-adaptive control theory	·	High confidence control synthesis
Mechatronics	·	10 times increase in power density	·	Integrative system technology
	·	Fail-safe and reliable mobile machines	·	Mixed-signal computational technologies
	·	Machine form factor and dexterity on par with humans	·	Methods for approximate performance
Machine intelligence	· · ·	Modalities for human/machine interaction Certifiable learning and programming Robust operation in open environments	· · ·	Symbol extraction from massive correlated data sets Situational awareness and ambient intelligence Rigorous testing and verification technology
Design automation	· · ·	Theory and methodology for multi-view, multi-abstraction models Novel design languages for heterogeneity Eliminate system integration pain	· · · ·	Performance estimation technologies Model relation and transformation technology Technologies for semantic definition Component and platform based system design
Network	·	Secure and trusted connections	·	Secure transmission and authentication
	· ·	Always available Mixed protocol systems	· ·	technologies Technologies for safe co-protocol design Technologies for data intensive applications

NITRD should inventory and streamline the vested interest of all stakeholder agencies. In particular, with regards to cyber-physical systems, DARPA, NSF, and NASA have direct interests. Less pronounced is the relation with NIH research, where automated health care is a prime consideration.

Furthermore, NITRD could attempt to help start, for example, a series of workshops of a specific multi-disciplinary nature to provide acclaim within the academic merit system. In addition, an infrastructure to support an ecosystem of computational models could be an effective approach to increase the utility of computation in cyber-physical system design. This could be chartered under NIST. Farther removed could be exploiting multi-view modeling for weather forecast as a NOAA responsibility. Finally, overlapping activity at the Department of Homeland Security, such as the efforts on securing our critical infrastructure, may be identified.

The commercial sector should help provide the means for the academic programs, including realistic scenarios, data sets, and problems to enforce the multi-disciplinary character that cyber-physical systems inherently impose. Furthermore, internships should be actively installed as a means to transfer technology.

Strategically selected contacts should be established with international programs, which is best done at the academic and research institute level. For example, collaboration on architecture and performance research with European institutes and robotic research with Japanese and Korean institutes could be valuable. This will allow the U.S. to play the role of integrator and to define general-purpose approaches of a comprehensive nature.

If you wish to discuss the matter presented here in more detailed, I very much look forward to your comments, questions, or suggestions. I hope to have provided some valuable feedback.

White Paper

Five-Year Strategic Plan (FYSP) for Federal Networking and Information Technology Research & Development Program

The FYSP must include enhanced countermeasures for insider threat detection and mitigation as a priority.

Awareness of information value has literally exploded in recent years. It is widely-known there is a large and growing international black market for personally identifiable information to facilitate identity theft. Evidence of this can be found in the near daily announcements of information loss or theft that has reached epidemic proportions.

Technology vendors have responded to the seemingly endless stream of information loss disclosures by providing Data Loss Prevention (DLP) tools in an effort to stem the hemorrhaging of information from enterprise networks.

However, these first generation DLP tools were designed to detect accidental loss of information through carelessness or negligence on the part of insiders with authorized access to the information—they were not designed with malicious users in mind.

The increasing value of information coupled with continuing deployments of DLP tools will, in effect, drive nefarious insiders to find and use more technically sophisticated ways to steal information.

One way insiders can steal information is through the use of digital steganography. By using any of the more than 1,000 steganography applications available as freeware or shareware on the Internet, insiders can steal sensitive information with no risk of detection. There is no risk of detection because neither the current generation of DLP tools nor any other current generation network security appliance can detect the use of digital steganography by insiders.

In addition to insider theft of sensitive information, steganography applications can also be used to steal intellectual property. In his remarks at the inauguration of the National Intellectual Property Rights Coordination Center, DHS Secretary Chertoff said “Our national assets and our productive resources in many ways are concentrated in our intellectual property.”[1]

Accordingly, protecting intellectual property from loss or theft is an Economic, National, and Homeland Security imperative.
Steganography is also being used by criminals to conceal evidence of criminal activity such as distribution of child pornography and drug trafficking, for example, and is being used by terrorists for covert communication.

In terms of importance, the Federal Plan for Cyber Security and Information Assurance Research and Development[2] stated that steganographic technologies “… pose a potential threat to U.S. national

[1] Remarks by Homeland Security Secretary Michael Chertoff at the Inauguration of the National Intellectual Property Rights Coordination Center, July 10, 2008, http://www.dhs.gov/xnews/releases/pr_1215736423265.shtm

[2] Federal Plan for Cyber Security and Information Assurance Research and Development, Report by the Interagency Working Group on Cyber Security and Information Assurance, National Science and Technology Council, April 2006, http://www.nitrd.gov/pubs/csia/csia_federal_plan.pdf

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

security." and that "The threat posed by steganography has been documented in numerous intelligence reports."

In terms of capability gaps, the Plan states that "Resources to evaluate, integrate, and deploy the numerous basic research advances are limited and should be enhanced."

The role for the NITRD Program is to prioritize the expansion of research on existing analytical detection capabilities highly enough to be competitive for funding. The role for the commercial sector is to transition the expanded research results into state-of-the-art digital forensics tools for use by examiners in federal, state, and local law enforcement agencies; the intelligence community; and the private sector and state-of-the-art network security tools for real-time detection of insider use of steganography.

Research into advanced steganography countermeasures is inherently a multi-agency effort because every agency has sensitive information that could be stolen by insiders using steganography. Accordingly all agencies would benefit from enhanced countermeasures to detect insider use of steganography.

A key strategic goal would be to establish a national repository of steganography applications, fingerprints, and signatures that could be used to develop world-class steganalysis tools.

A key challenge will be to provide resources sufficient to establish and maintain the repository and perform the technical research on every steganography application in the repository in order to discover the digital signature of the application, if one exists, and the corresponding algorithm to extract information hidden with the application that can then be integrated into automated steganalysis tools.

The impact of prioritizing research on steganography applications and subsequent development of world-class steganalysis tools will be the detection of attempts to use steganography to steal sensitive information, to include intellectual property, along with detection of attempts to use steganography to conceal evidence of criminal activity that would have otherwise gone undetected.

Improved steganalysis tools will enhance efforts to combat cyber crime and will enhance National and Homeland Security by improving insider threat detection and mitigation capability.

White Paper: Response to NITRD RFI on the Five-Year Strategic Plan for the Federal NITRD Program

Our vision for the NITRD program revolves around four major thrusts, namely network science, evolvable networking architectures, innovative networking paradigms and cyber security. Next, we discuss the thrusts in the context of the two questions posed by the RFI:

- ***What do you imagine as the future in terms of desired NIT capabilities?***

Establishing a new field of network science is of paramount importance to the society which depends on a diversity of complex networks, e.g. transportation, biological, social, military. It is a multi-disciplinary area that requires multiple agencies joining forces to coordinate incoherent research investments and identify dualities and common concepts across domains. The objective is to develop rigorous mathematical basis and measurement tools to model, at sufficient fidelity and with sufficient scale, design, analyze, predict and control the behavior of networked entities.

The emergence of diverse networking applications, e.g. embedded networked systems, automation, networked rovers, and delay-tolerant applications, clearly demonstrates that a manually-configured static networking stack (e.g. TCP/IP) cannot fit all. An efficient and effective substitute for stove-piped systems is to build *evolvable networking software* that can evolve to synthesize, at run-time, the optimal networking architecture and seamlessly morph across stacks to match the diverse objectives and constraints of vastly different applications.

Geo-Networking (GeoNET) is a strong candidate for *networking paradigms beyond IP* due to optimally exploiting location information to simplify routing, address translation and mobility management. Furthermore, GeoNET capitalizes on location-based services projected to dominate over the coming decade. This requires new approaches to geo-based addressing, geo-MAC for interference avoidance, geo-routing, geo-based constructs, e.g. geo-cast and geo-gather.

Rate/Reliability/Latency Network Utility Maximization (R/R/L NUM) extends rate-based NUM to dynamically balance the fundamental rate-reliability-latency trade-off in wireless networks. It is a key enabler of clean-slate networking architectures, founded on network and information theory formulations. This requires mathematical tools to develop an overarching NUM formulation, dynamic multi-objective utility functions, and optimization decompositions.

Finally, cyber security is not merely about security, its more about *Self-resilient Trustworthy systems* where new models, logics, and theories for analyzing and reasoning about security, reliability, privacy and usability come into play. Cyber security is an inherently multi-disciplinary area where technology, policy and business need to converge to support foundational research.

- ***What roles do you imagine for the NITRD Program and for the academic, commercial, international, and other domains in achieving that future?***

The NITRD program should foster fundamental and applied research through creating programs along one, or more, of the four areas identified above. Academics should lead the network science and R/R/L NUM initiatives with strong support form the industry to apply developed theories and feedback their insights. Both academia and industry should co-lead the cyber security initiative due to its complex nature that goes beyond technology to policy and business aspects that need strong industry involvement. The evolvable networking architecture and GeoNET initiatives would also require strong partnership between industry and academia due to their systems focus. The international community should streamline the process of know-how dissemination, information sharing across borders through multi-disciplinary conferences, committees and student exchange.

Next, we address the following six challenges highlighted in the RFI:

- ***Development and execution of multi-agency and multi-disciplinary programs***

Network Science Program: NSF-NIH-DARPA-DOD-DOE/SC.
Evolvable Networking Architecture Program: NASA-DARPA-DOD-NSF.
GeoNET Program: NASA-DHS-DARPA-DOD-NIST-NSF.
R/R/L NUM Program: NSF-DARPA-DOD-DOE.
Cyber security and Information Assurance Program: DHS-NIST-NSA-NSF-DARPA.

Form a core multi-disciplinary team that manages and coordinates the work of multiple sub-teams, each focused on a subset of disciplines with highest synergy.

- ***Determination of strategic goals, key challenges, opportunities, and research priorities***

Strategic Goals: National security, prosperity and welfare of the people, US IT leadership, Information and Decision superiority, US energy independence, and US dominance in space.
Key Challenges: collaboration, communication and reaching the best abstractions (APIs) across disciplines, bridging the gap between theory and practice, optimal transition milestones, unconventional & innovative types of cyberattacks, and diverse types of distributed applications.
Opportunities: capitalize on H/W trends attributed to Moore's law, advances in sensing, MEMS and embedded systems, plethora of emerging wireless technologies, strong market pull for new applications, convergence of control, comm. and computing, and seeds for network theory.
Research priorities: *Theory:* highest priority is R/R/L NUM followed by long-term network science and cyber security. *Systems:* highest priority is evolvable networking architectures, followed by GeoNET and cyber security.

- ***Examples that illustrate the impact of realizing the vision, achieving the proposed goals, and meeting the identified challenges***

Establishing the new field of network science should lead to new findings and theories with direct impact on operating transportation, financial and communication networks as well as better understanding the immensely complex biological and social networks.

Availability of evolvable networking architectures significantly reduces the resources needed to build a "tailored" networking infrastructure for diverse applications that exist nowadays or will emerge sometime in the future. This will surely have a significant societal and economic positive impact due to deploying cost-effective services for the public much faster.

GeoNET simplifies packet routing, enables scalable networked systems, e.g. border control, facilitates NASA exploration and science missions, enables next generation GIG and DHS pre-dominantly location-based applications.

Successful deployment of a reliable cyber security infrastructure facilitates and streamlines the process of information sharing among trusted parties for civilian applications, with prevalent impact on the global economy (e.g. e-commerce), as well as national security.

- ***Transition of R&D results into practice***

Leverage COTS technologies, as much as possible, shortens the transition path and strengthens risk mitigation. Maintain close collaboration among government, industry and academia in multi-phase programs. Promote proof-of-concept prototypes and demonstrations early-on in the program serves dual-role: proves/disproves the soundness of theory and underlying assumptions early-on and provides key insights to basic research about challenges and limitations.

- ***Role of the U.S. in the international NIT arena***

As it has always been, the US should lead NIT innovations encouraging the international partners to plan an aggressive R&D portfolio, share research findings and exchange experiences.

- ***Interactions among government, commercial, academic, and international sectors***

Industry should work closely with academics to define problems and mature their research to sub-system and prototype levels. The US government agencies need to coordinate with international counterparts in order to layout a coherent R&D portfolio that is effective and efficient.

Formally Verifiable Architecture Patterns for Safe Medical Device Plug and Play (MD PnP)

NSF CPS Research Planning Workshops have identified that compositional system modeling, analysis, synthesis, and integration are at the frontier of engineering sciences. At the center of this development is the current transition from a reductionist approach to a compositional approach in engineering and science. The science of system composition has clearly emerged as one of the grand themes driving many of our research questions in networking and distributed systems. By *system composition* we mean that the QoS properties and functional correctness of the system can be derived from the system architecture structure, and the functional and QoS properties of constituent components.

Medical systems are an important class of network-controlled reconfigurable CPS systems with a high degree of complexity. Each year thousands of patient injuries and near-misses are caused by improper or unsafe medical device-device and/or device-human interactions. The Institute of Medicine (IOM) report[1] revealed that at least 98,000 annual hospital deaths are attributable to medical errors.

Networked medical devices systems have both loosely coupled and tightly coupled subsystems, including infusion pumps, respirators, robot-assisted surgery devices, monitoring and diagnostic devices, as well as medical information systems, and medical personnel. The have different degrees of safety, reliability and real time requirements. For example, after major surgery, a patient is allowed to "operate" an infusion pump with potentially lethal pain killers (patient controlled analgesia (PCA)). When pain is severe, the patient can push a button to get more pain medication. The use of PCA devices is an integral part of modern post-operative care. This is an example of a safety critical device controlled by a non-safety critical and error-prone operator (patient). In spite of operator errors, the operation of PCA system must be verifiably safe. Furthermore, the role of each device and their QoS requirements will change at different stage of a medical procedure. This raises many system composition challenges.

At the heart of compositional research is the development of formally specified and verified architecture patterns that can be deployed at scale. An architectural design is mature only if, under the proposed architecture, system requirements will be met without significant changes to component definitions, to interfaces and to the rules of interactions (protocols) during the development process. That is, lower level designs and implementations can unfold as planned. Matured software architectures are characterized by explicit stated assumptions and properties that we can predict and formally verify. In this regard, most building architectures are mature, while most software developments are not supported by matured architecture design patterns.

However, when there are no architecture patterns to follow, even in a matured engineering discipline, the system development is facing serious problems, such as the Tacoma bridge collapse[2] and the Sydney Opera House cost and schedule overrun[3].

It is vital for our nation to develop the scientific and technological foundation upon which formally specified and verified MD PnP architecture patterns can be developed and deployed at

[1] Kohn LT, Corrigan JM, and Donaldson MS, eds. To Err Is Human: Building a Safer Health System, The National Academies Press, 2000.

[2] http://www.wsdot.wa.gov/tnbhistory/Connections/connections3.htm

[3]The adjusted cost of the building increased almost fourfold, from 15 to 55 million Australian dollars - 20 million before Utzon's departure and 20 million after. http://www.gsd.harvard.edu/research/publications/hdm/current/21_tombesi.html

scale. To this end, we need a national research initiative that includes medical community, computer science community, engineering community and the FDA to jointly develop:

- Domain models for dynamic composition (MD PnP), which capture the structure and dynamics of requirements and their changes during medical procedure workflow[4].
- Advanced technologies to integrate the operational context with low level alarms to provide timely “situational awareness” for medical personnel, because operational contexts are different at different stages of a complex medical procedure.
- A coherent suite of formally verified QoS protocols for each domain model. We need to develop cross domain QoS protocol interferences model and then formally verify that candidate protocols for different QoS dimensions will not interfere with each others.
- Formally specified and verified architecture patterns that can ensure system level safety in spite of the faults and failures in non-safety critical components, because complex and unverifiable components, e.g., human operator and certainly highly complex (legacy) software components are difficult to avoid.
- Evidence based certification that utilizes formally specified and verified architecture patterns.
- Computer-aided configuration time verification tools to ensure that dynamically configured medical device networks is compliant with formally verified and certified architecture patterns.

Input to NITRD Program's Strategic Plan for Cyber-Physical Systems

In response to the RFI for the NITRD Five Year Strategic Plan, I am attaching a recent white paper we submitted to the workshop on Automotive Cyber-Physical Systems. The research community at large has responded with overwhelming support for this critical domain of cyber-physical systems, recognizing its importance not only for its strategic economic impact, but also for its criticality in the safety and reliability of embedded and ubiquitous computing elements within the larger social context of our everyday lives. Thus while the attached white paper may appear to focus only on automotive applications, the underlying themes of safety, dependability, security and reliability are common to a wide range of CPS applications. I urge the planning activity to put the CPS agenda on top of their strategic plans for the next five years.

Position Paper: Guaranteed End-to-end QoS in Automotive Cyber-Physical Systems

Next-generation Automotive Cyber-Physical Systems (ACPS) face increased demand for safety-critical functionalities, such as electronic drive assist applications, X-by-wire systems, as well as increased demand for non-critical functionalities, such as multimedia applications, Internet, and ad-hoc networking with other cars and stationary objects. While automotive systems have always been extremely cost-sensitive with a focus on safety and reliability, the increasing embedded software/hardware content in these systems raises new issues for guaranteeing Quality of Service (QoS) – which we broadly interpret to include timing, safety, dependability, security, accuracy and other aspects of interest.

Limitations

Fundamental limitations for Automotive Cyber-Physical Systems (ACPS) include:

- Lack of good formal representations and tools capable of expressing and integrating multiple viewpoints and multiple aspects. This includes lack of robust formal models of multiple abstraction layers from physical processes through various layers of the information processing hierarchy; and their cross-layer analyses.
- Lack of strategies to cleanly separate safety-critical and non-safety-critical functionality, as well as for safe composition of their functionality during human-in-the-loop operation.
- Ability to reason about, and tradeoff between physical constraints (e.g., battery capacity, wiring harness complexity, etc.) and QoS of the ACPS.

Challenges

- *Modeling and verifying end-to-end timing behaviors for emerging ACPS platforms* is challenging due to the extensive amount of in-car networking that involves end-to-end interactions among multiple layers (application, middleware, network, OS, hardware architecture) in a distributed environment. A holistic approach to understanding timing in these distributed multi-layer systems is both essential and of significant benefit because (a) ACPS applications often need to meet end-to-end hard or soft real-time needs, (b) existing techniques for timing analysis of ACPS applications do not account for the spectrum of timing granularities in a cross-layer system, which can vary by orders of magnitude, and (c) knowledge of timing parameters at the different levels can dramatically improve the utility (e.g., QoS, energy, etc.) and performance of ACPS applications that often execute in constrained environments where CPU, memory, network and device energy may be limited. Furthermore, data integrity is critical in the context of ACPS. In particular, how can we guarantee the integrity of data in the context of end-to-end delivery across multiple abstraction levels (from sensors, across multiple distributed networking and software abstractions, down to embedded hardware platforms, and all the way back to actuators? The relationship between the traditional ERTS constraints – timing, resources, energy, etc. – and the quality of data needs to be modeled and characterized to enable reliable monitoring and manipulation of the ACPS: both within the vehicle, as well as remotely.

- *Integration of Time-Triggered and Event-Driven Behaviors.* A fundamental challenge that needs to be resolved to improve the interaction between safety-critical and non-critical functionalities is the composition of time-triggered and event-driven systems. Time-triggered systems were developed as an effort to improve predictability, and are commonly utilized in automotive safety-critical subsystems. Time-triggered systems are based on a mathematical model that allows the scalable schedulability analysis of hard real-time properties, such as execution times and latencies. Event-driven systems are a popular choice for non-critical subsystems, as they are less costly, simpler to implement due to the lack of global synchronization, and often provide better average utilization, performance, and throughput than time-triggered systems. In order to facilitate more efficient communication between safety-critical and non-critical subsystems, we need to develop standard methods for the build-by-composition design of mixed time- and event-driven systems.

- *Providing guarantees on functionality and real-time properties for ACPS.* Due to the complexity of the safety-critical subsystems and their interactions, we need to enhance our capability to formally analyze and guarantee functional and real-time properties in ACPS. In the past decades significant advances were achieved in static schedulability analysis, and we have scalable methods for the analysis of time-triggered systems. Simulations and test-beds are widely used in the industry for the functional validation of safety-critical subsystems. Model-checking provides alternative methods for the real-time analysis of event-driven systems. Control theory provides the foundations of runtime monitoring and control based on mathematical foundations. No method is likely sufficient for the analysis of mixed time- and event-driven systems. We need to find ways to combine various analysis methods to support the build-by-composition design of ACPS.

Innovations

Promising innovations for achieving ACPS include:

- Integration of formal models and analyses with simulation, testing, monitoring of deployed systems in a mutually synergistic manner.
- Formal models that address both cross-layer and end-to-end considerations.
- Methods and standards for build-by-composition design of mixed time-triggered and event-driven systems.
- Formal guarantees for critical functionality and QoS in ACPS that will in turn enable formally-based certification processes.

Biographical Sketches

Nikil Dutt a Chancellor's Professor of CS and EECS at UC Irvine and is affiliated with the Center for Embedded Computer Systems (CECS) UC Irvine. He received a PhD in Computer Science from the University of Illinois at Urbana-Champaign (1989). Dutt leads research projects in several aspects of networked embedded system design and automation, including issues in software/hardware codesign, storage and memory requirements, and validation/verification of complex embedded systems. Email: dutt@ics.uci.edu, Phone: 949-824-7219.

Gabor Madl is a Ph.D. candidate and graduate student researcher at the University of California, Irvine, and is a member of Professor Nikil Dutt's research group. His research focuses on the combination of formal methods and simulations for the model-based analysis and evaluation of embedded systems. He received his M.S. in computer science from Vanderbilt University and in computer engineering from the Budapest University of Technology and Economics. Email: gabe@ics.uci.edu.

High Confidence Medical Device Software Systems

The area of Cyber Physical Systems (CPS) – that is, NIT systems connecting with the physical world needs to be the highest priority among the eight technical priority areas recommend by the PCAST report. This is necessary to maintain American Competitiveness, as CPS is where other countries EU and Asian countries are aggressively investing in their R&D programs. The application domains of CPS include healthcare, transportation, process control and energy distribution, large-scale physical infrastructures, and defense systems.

Given the shortage of caregivers and the growth of an aging US population, the future of US healthcare quality does not look promising and definitely is unlikely to be cheaper. Advances in healthcare technology and health information systems offer a tremendous opportunity for improving the quality of our healthcare while reducing healthcare costs [1].

The development and production of medical device software and systems is a crucial issue, both for the US economy and for ensuring safe advances in healthcare delivery. As devices become increasingly smaller in physical terms, but larger in software terms, the design, testing, and eventual Food and Drug Administration (FDA) device approval is becoming much more expensive for medical device manufacturers both in terms of time and cost. Furthermore, the number of devices that have recently been recalled due to software and hardware problems is increasing at an alarming rate. As medical devices are becoming increasingly networked, ensuring even the same level of health safety seems a challenge [2].

The cross-cutting nature of medical device design—transcending the informational, physical, and medical worlds—along with the possibility of a nationwide networked medical system that actively monitors and regulates the health of our nation's citizens, raises immense scientific and technological R&D challenges for the IT, medical, and regulatory communities. The challenges envisioned for the next five to ten years include the following:

- System integration. As we embrace a plug-and-play vision of medical device networks in future digital hospitals and digital homes, we must collectively facilitate the development of medical device systems and coordinate them with the development of standards for the architecture and communication of interoperable plug-and-play device networks. Achieving these goals while establishing quality-of-service levels that ensure system and patient safety on the one hand, and patient security and privacy on the other, is a great challenge.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

- Critical infrastructure. As we move toward an environment in which all patients are constantly monitored and actively plugged into a nationwide medical information network, we are creating a new critical infrastructure that will literally monitor the nation's health. We need new methods to ensure the safety and security of that network, particularly methods involving the active use of information for medical purposes. In the presence of abnormal conditions or attacks, the system's performance must degrade gracefully and safely, and the system must identify, contain, and, if possible, repair faults while providing timely notification to human operators.

- Embedded real-time systems design. Medical devices are embedded not only inside information networks but also inside human patients, whose critical life-functions they monitor and regulate. The design of medical devices is therefore more than an NIT issue; it must also include the device's interaction with the patient and the environment and the context in which they coexist. Thus, we need a fundamental rethinking of medical device design—toward a holistic approach that integrates functional, computational, and communication designs in the presence of highly uncertain patient models in both normal and abnormal conditions.

- Verification, validation and certification. Current design practice makes certification and verification an afterthought, taking place at the end of the design cycle, when it is frequently too late to change design choices. As medical devices become more complex and more interconnected, it is becoming increasingly evident that certification should be incorporated in early design stages. Furthermore, certification and design frameworks are currently not component-based, resulting in time-consuming and expensive certification of large integrated systems. This drawback makes the current approach inefficient for certification of incremental or evolutionary designs, and creates difficulties in maintaining or upgrading legacy systems.

To address these challenging R&D issues, there must be coordination among funding agencies (NSF, NIH, DoD), the regulating and standardization agencies (FDA, NIST). Such coordination activities should promote collaborations between academic researchers and medical device manufacturers. There then lies the potential to create a new scientific community and a new generation of scientists and engineers that integrate computer science, engineering, and medicine.

References

[1] Insup Lee, George Pappas, Rance Cleaveland, John Hatcliff, Bruce Krogh, Peter Lee, Harvey Rubin, Lui Sha, "High-Confidence Medical Device Software and Systems," IEEE Computer, vol 39, no 4, April 2006. pp. 33-38.

[2] Proceedings of Joint HCMDSS/MD PnP workshop, Julian Goldman, Insup Lee, Oleg Sokolsky, Susan Whitehead, Eds., IEEE Press, June 25-27, 2006. (www.cis.upenn.edu/hcmdss07/)

A “cyber-physical” system takes on different definitions depending on the application and the underlying technology. As used here and as is prominent across UTC applications the issues that are central to cyberphysical systems are (a) scale and (b) interdependency.

Scale refers to the sheer size of the emerging *meta*-systems, composed of different physical systems and interconnected by diverse sets of communication networks. Building systems are one UTC application where scale could refer to the number of occupants in a building — new high rise buildings and campuses contain upwards of 30,000 participants and have thousands of points of sensing and controls for air quality and fire protection. Avionics systems are another example of scale — aircraft electric power systems are increasing in power density by factors of 3–5X in each new generation of aircraft and are connected by multiple networks with hundreds of thousands of switching scenarios.

Interdependency refers to coupling of physical systems through communication networks. Physical devices that were locally operated and controlled are now coupled by local and wide area communication networks with significantly enhanced computation distributed across the network – increasingly the networks are wireless and increasingly controlled by end users or consumers through PDA, cell phones or other mobile devices. An effect of these complex communications networks is that information can be obtained and used for control in ways that create new behaviors -both desired and so called “emergent” -that make the design and assurance of robust operation increasingly problematic.

Developing the underlying scientific and engineering knowledge base that allows full understanding and exploitation of these kinds of modern cyber-physical systems is critical. Elements that appear in UTC applications that today characterize the area and that form the basis of recommendations to NITRD include:

Scale. The systems that UTC sees emerging in building systems and in aerospace systems are of increasing scale -indeed, one could say “societal scale” -orders of magnitude different in qualitative and quantitative size to systems currently deployed today. The coupling of the diverse physical systems in the design and operation of products enables new functionality to be provided and especially tailored to customer needs and desires; however, the tools for engineering systems at societal scales do not currently exist.

Heterogeneity and Interconnected Systems. The systems that UTC sees emerging in building systems and in aerospace systems are diverse and the coupling provided by modem communication networks assembles cyber-physical systems where new coupling occurs.

Uncertainty. Uncertainty provides the motivation for a large degree of coupling in cyber-physical systems -the motivation is to react to environmental issues or to customer needs and desires. Uncertainty also is a key ingredient -with a high degree of interconnectivity -of “emergent” behavior.

Why does the area matter? There are three key issues that are represented by UTC applications and that are clear in other sectors (for example in transportation and health care).

Functionality. Products are being required to have higher levels of performance. An excellent example is in the area of buildings which are heading to “net zero energy” where over the course of a year the energy consumed is equal to the energy produced, requiring energy efficiency gains coupled to renewable energy sources. Computation and communication is essential to delivering these new levels of performance and the coupling of building functionality -HVAC, lighting, elevators together with thermal storage and renewable energy sources is exactly a “cyber-physical systems. The drivers of sustainability and reactions to global warming as well as energy security are accelerating the design and implementation of such functionality.

Delivering functionality at acceptable cost and risk. Coupling subsystems through communication networks makes information available globally and enables new performance. Coupling of subsystems however increases the cost of designing such systems — bringing different groups together — and increases the risk that coupling can cause undesirable behavior. The classic example of coupling on power grids and collapse during blackouts of entire regions is an excellent example of the cost and risk that “cyber-physical systems” bring forward. Risk also includes the development of dependable systems which focuses on the safety critical nature of cyber-physical systems and the role of dependable software.

National competitiveness. Increased energy efficiency is an example of the benefits that “cyberphysical systems” can bring to society. The product offerings will increase the value of both the physical and the “cyber” or information technology portions. The design and manufacturing base as well as the talent base is critical for national competitiveness.

Recommendations:

Public-private partnerships for R&D. There needs to be Federal and State investments in the basic science of “systems of systems” and technology that addresses key cyber-physical issues of scale, heterogeneity and uncertainty. These investments must produce new methodology, tools and talent and must involve academia, National Laboratories and industry working together.

Industrial involvement and involvement of mission oriented agencies. There must be investments in the mission oriented agencies including DOD and DOE to focus on industrial situations of national importance and that are targeted to identify and mature technology that can be effectively commercialized.

Standards. There must be investments at NIST and related Federal and State agencies to identify and put in place standards for communication and control that enable interoperability and effective commercialization of cyber-physical systems.

August 25, 2008

~~Introduction~~

The Council on Competitiveness is the only group of corporate CEOs, university presidents and labor leaders committed to the future prosperity of all Americans and enhanced U.S. competitiveness in the global economy through the creation of high-value economic activity in the United States. A nonpartisan, nongovernmental organization in Washington, D.C., the Council shapes the debate on competitiveness by bringing together business, labor, academic and government leaders to evaluate economic challenges and opportunities.

The Council greatly appreciates this opportunity to provide input during the development of the Networking and Information Technology Research and Development (NITRD) program's next Five-Year Strategic Plan. As we have seen so often in the past, sustained and coordinated federal government support for fundamental research in networking and information technologies (NITs) frequently seeds the innovations and innovation platforms (e.g., the Internet) of tomorrow. Government programs and partnerships with industry are also key competitiveness mechanisms for allowing U.S. companies to maintain leadership positions in NIT.

Building on Valuable Past Work

Given the competitiveness value of a number of the research priorities from the program's most recent five-year plan,[1] the Council hopes that the following areas will continue as part of the next iteration of the plan currently under development:

- **Compact, teraops-scale supercomputing systems**. Advanced computing capabilities are a central component to U.S. innovation, especially as that computing power becomes more mainstream (or "trickles down") and can be accessed by U.S. companies either through partnership programs with universities and government labs or directly by purchasing from vendors.

[1] See http://www.nitrd.gov/pubs/strategic_plans/2002_2006_NITRD_Strategic_Plan.pdf.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

- **Reliable, universally accessible high-speed networks**. Just as access to high-speed networks have enabled many of the technological innovations seen in the last decade, even faster, next-generation networks will no doubt enable tomorrow's innovations, making leadership in networking a key U.S. competitive advantage.

- **Social, economic, and workforce implications of IT**. The more we understand about technology's role in supporting the strength of the U.S. economy and workforce, the better able U.S. companies will be in the future to take advantage of new technologies to enhance our overall global competitiveness.

New Areas of Focus

In addition to furthering the existing priorities mentioned above, the Council also feels that continued, coordinated, multi-agency research and development efforts in the following focus areas would also be most beneficial for U.S. economic competitiveness:

- **Government/industry partnerships**. The Council sees great value in new and continued support for programs (like the Department of Energy's INCITE program) that allow partnerships between government and industry—partnerships which Council research shows have been successful and have added to the global competitiveness of U.S. companies.

- **Advanced software research**. As computing hardware performance continues to improve and the shift toward multicore architectures continues, significant advances will be needed in software tools, programming languages, parallel processing, and related education and training to allow U.S. developers and companies to leverage this growing computing power for innovation and economic growth.

- **Computational expertise**. America clearly has the technological edge today—indeed, the most powerful computing systems in the world are in the United States. But America lacks sufficient numbers of computational scientists and engineers to fully exploit this leadership position. According to Council surveys, the biggest single constraint on the deployment of advanced computation tools is the lack of computational scientists. For this reason, NITRD support for a multi-agency program to increase the number of U.S. computational scientists would be a sound investment for the nation.

Information technology continues to play a critical role in our future prosperity and in maintaining the strength of our nation's economy. Thank you again for this opportunity to contribute to the development of NITRD's next five-year strategic plan. We hope that our expertise will help you guide the coordinated information technology R&D efforts of participating agencies.

Position Paper on Federal R&D Investment in Cyber-Physical Systems

Cyber-physical systems (CPS) consist of the interconnection of numerous physical and computational processes. Numerous examples of CPS may be found throughout the national civil infrastructure. This position paper discusses a recent CPS project called CSOnet that attempts to address national problems concerning wastewater runoff. The objective of this document is to use CSOnet to highlight the broad interdisciplinary nature of CPS, to comment on the importance of encouraging industrial and academic collaborations, and to petition for a greater federal support of such projects.

CSOnet [1] is a sensor-actuator network that is being built in South Bend Indiana to address the problem of combined-sewer-overflow (CSO) events. A combined sewer system combines sanitary and storm water flows into a single system. CSO events occur when storm water flows exceed the capacity of the existing system, thereby causing operators to dump the untreated water directly into a river or stream. As these waters are highly impacted by biological and chemical contaminants, the occurrence of CSO events has a significant negative impact on public health and water quality.

CSO events are a problem of national concern. Combined sewers are found in nearly 800 municipalities, centered primarily in the Midwestern and Northeastern United States. In addition to this, however, is the fact that the US environmental protection agency (EPA) has begun levying fines against local municipalities for each CSO event, as part of the clean water act. These fines are significant and can be construed as an unfunded federal mandate requiring municipalities to reduce their CSO events.

In 2004, under the leadership of the University of Notre Dame, a small group of academic (University of Notre Dame, Purdue University), public sector (city of South Bend, Indiana), and private (EmNET LLC) sector partners began developing the concept of using a sensor-actuator network to control the occurrence of CSO events. At present, the project has deployed a 150+ sensor network in the city of South Bend for monitoring the occurrence of CSO events. The actuation (control) part of the system will control CSO events. This component will be in place by summer 2009. While started at South Bend, the project is starting to expand to other Indiana cities though the small business (EmNET LLC) that was created as part of the original project. CSOnet can therefore be viewed as an example of a successful CPS engineering project.

Sensor-actuator networks such as CSOnet differ greatly from traditional data networks in that the data transmitted across the network is used to control the external environment. Networks having such feedback paths exhibit extremely complex behaviors that can be difficult to predict. Understanding and controlling this complexity requires developers that are comfortable in a wide range of engineering and scientific disciplines. Not only do these developers need to be aware of traditional communication and computer networking practices, they must also be knowledgeable about the physical processes that these networks are being used to control. In the case of CSOnet, we drew on a wide range of engineering expertise that included environmental engineering, fluid dynamics, network middleware, adaptive antenna design, and real-time systems. Identifying a group with such a diverse set of talents and integrating that group into a team was essential for the success of the CSOnet project.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

The very breadth of the CSOnet team made it difficult to secure federal funding for such a project. Agencies such as the National Science Foundation (NSF) support fundamental research in very narrow technical disciplines. Technical reviewing panels are often drawn from researchers specializing in one area. It can be difficult for such panels to appreciate the depth of innovation present in projects such as CSOnet because these projects require expertise in such a broad range of specialties. As a result, federal funding for CSOnet has been very limited; $150,000 in the last couple years of the project. This is a small fraction of what was required in the first 3 years of the project. In contrast the related European WIDE project (INFSO-ICT-224168) is being supported (2008-2011) at a level of 2.7 million euros by the European Union. The WIDE project is building a sensor-actuator network similar to CSOnet for the city of Barcelona.

Funding for CSOnet was obtained from state and city governments. The state of Indiana's 21st Century Technology Fund (CTF) supported CSOnet project with a one million dollar grant between 2004-2007. This was used to build a prototype system. An additional one million dollars was obtained from the 21st CTF (2007-2009) to scale up to the entire South Bend metropolitan area. During this period, the City of South Bend contributed over $150,000 in services and funds to the project.

State programs such as the 21st CTF, however, are primarily economic development programs. Their primary interest is in stimulating local businesses in order to grow the state economy. There is little interest in funding the fundamental research required to ensure the success of CSOnet. To ensure adequate freedom for academic partners in this project, we created a small start-up company (EmNET LLC) to serve as a buffer between the actual government stakeholders (City of South Bend) and the academic units (Notre Dame and Purdue). This worked very effectively and it essentially forced the academic units to act as research arms for the local company.

CSOnet is a large-scale CPS system for controlling CSO events; a problem of national importance. The success of the project required assembling a broad team of academics, private, and public sector stakeholders. It required the creation of a small business entity to serve as the technology transfer bridge. It required significant funding (on the level of 2-3 million for 3-5 years) to stimulate interdisciplinary research in a broad range of engineering areas as well as funding for the more mundane job of product development and testing. Increased federal funding that encourages fundamental CPS research while also supporting the participation of small business would be a great help in promoting projects such as CSOnet that have the potential of maintaining and securing our national civil infrastructure.

[1] L. Montestruque and M.D. Lemmon (2008), CSOnet: a metropolitan scale wireless sensor-actuator network, *International Workshop on Mobile Device and Urban Sensing (MODUS)*, 2008

AISec: Leveraging Artificial Intelligence for Personalized Security and Privacy

A white paper in support of the NITRD RFI on the Five-Year Strategic Plan for the Federal Networking and Information Technology Research & Development Program

There is a long tradition of using artificial intelligence (AI) to tackle security problems. A prevalent research method is to collect data capturing a particular malicious activity (e.g. network intrusions, spam) and using AI techniques such as machine learning, train a detector for future malicious activity of the same type. While this approach clearly yields powerful security technologies it is largely an after-the-fact approach to security in that data of security breaches is needed first, before the protection mechanism can be developed. The growth of the Web and efficient AI-based techniques for mining large corpora mean that we can now *anticipate* the adversary to a degree not possible previously. While this is a significant advance, we argue that even more transformative technologies are possible through continued collaboration between AI and security. In particular, we call for research in a new subdiscipline called "AISec" that leverages and extends AI advances in predictive modeling to achieve *personalized* security technology. We envision security technology that is personalized to the user in terms of their security vulnerabilities and privacy preferences, thus achieving high usability while providing strong security and privacy.

We highlight three problem areas which have recently gained a proactive security advantage through leveraging AI before discussing the AISec agenda in more detail.

Password reset. Online service providers commonly require users to not only select "challenge questions" that they may be asked in the event that they forget their password and are unable to login. Common challenge questions include, "What's your mother's maiden name?" and "What's your favorite pet's name?". Recently, Jakobsson et al demonstrated that publicly available records can be mined to determine the answers to many such questions (see, for example, [GJ05]) . Subsequently, they developed a more secure approach based on preferences [JSWY08].

CAPTCHAs. The tests humans are asked to perform when registering for a Web site (typically, typing in a distorted word) are called CAPTCHAs. The state of the art in CAPTCHAs is constantly evolving as advances in computer algorithms frequently ruin the effectiveness of a particular CAPTCHA at distinguishing between humans and computers. Despite this "arms race", CAPTCHAs are routinely presented and put into use with little, if any, testing, often with bad results. For example, in [G08], Golle demonstrated that standard machine learning techniques can be used to break a CAPTCHA introduced at one the security community's most competitive conferences, thus underscoring, that AI techniques should be routinely applied to proactively identify weaknesses.

Data Privacy. Documents are commonly redacted prior to release to protect sensitive content when responding to Freedom of Information Act (FOIA) requests or legal subpoena. While the act of redaction can be time-consuming and tedious, the process of determining what to redact is even more challenging and error-prone (examples of redaction failures are discussed in [SGZ07]). In [CGS08] a fast data mining approach was demonstrated that allows the keywords closely associated with a sensitive topic to be quickly identified. The approach leverages the Web to model the adversary's knowledge and provide proactive privacy protection against inferences.

An Opportunity: Personalized Security.

While the above technologies are novel and effective, we believe even more transformative technologies are possible through a closer collaboration between the AI and security and privacy communities. The data mining community has demonstrated the power of predictive models for advertising. Even seemingly generic data like browser history or search terms has proven to be strongly indicative of demographic attributes (see, for example [adLabs]) and even identity [NYT]. We propose that the AI and security communities work together to design analogous models to transform the way users experience security and privacy today. With a sharing of data between industrial partners, the key features to predictive models of a user's security habits and privacy preferences can be identified. Such models will enable the user's security experience to be tailored to them, that is, their security vulnerabilities and privacy preferences, thus increasing usability while providing better security and privacy.

Consider for example, a model based on the types of software applications installed, browser habits and history, keyword searches and network connection patterns. The model might suggest that someone who frequently clicks on links in emails, connects to many unknown wireless networks and uses short passwords is a risk-taker and so needs a more stringent security policy. In particular, risk-takers might experience stricter browser requirements around certificate acceptance, and less flexible security posture requirements for connection to their employer's internal network. In contrast, users with good security practices, might enjoy more leeway with installing security patches and fewer hurdles to corporate intranet access.

Similarly, a user who engages in limited online social networking and contributes anonymously when they do, might be predicted to have strong privacy concerns around demographic information. For such a user, relevant parts of a Web site's privacy policy could be highlighted for them before they register, and they could be warned about sites that are likely to violate their privacy preferences. This automatic prediction of privacy preferences would be especially powerful in light of the well-documented difficulty users have in articulating their true privacy concerns [LHDL04].

Building on work done in the AI community on social influence (see, for example, [CCHS08]) the model might also predict the user's vulnerability to social engineering attacks. A vulnerable user could be supported with a more stringent warning system or automated protections. For example, an automated protection mechanism might bounce suspicious emails that appear to come from friends and send re-send requests to friends using email addresses from the recipient's contact list.

In conclusion, it is well-documented that one-size-fits-all security mechanisms frequently frustrate users and as result are often simply turned off, resulting in no security at all (see, for example, [WSC06]). We believe a research agenda in *personalized* security would translate into less corporate and government data leaks, as it targets security policy for the user and incentivizes good behavior. The interdisciplinary nature of this research agenda and the requirement for collaboration between multiple industrial partners and academia makes support from an influential government body like the NITRD crucial.

[adLabs] Microsoft's adCenter Labs. http://adlab.msn.com/

[CGS] *Detecting Privacy Leaks Using Corpus-Based Association Rules*. R. Chow, P. Golle and J. Staddon. KDD 2008.

[CCHS08] *Feedback Effects between Similarity and Social Influence in Online Communities*. David Crandall, Dan Cosley, Daniel Huttenlocher, Jon Kleinberg, Siddharth Suri. KDD 2008.

[G08] *Machine Learning Attacks against the Asirra CAPTCHA*. P. Golle. ACM CCS 2008.

[GS05] *Messin' with Texas, Deriving Mother's Maiden Names Using Public Records*. V. Griffith and M. Jakobsson. ACNS '05.

[LHDL04] *Personal Privacy through Understanding and Action: Five Pitfalls for Designers*. Lederer, S., J.I. Hong, A. Dey, and J.A. Landay. Personal and Ubiquitous Computing 2004. 8(6): p. 440 - 454.

[JSWY08] *Love and Authentication*. M. Jakobsson, E. Stolterman, S. Wetzel, L. Yang. (Notes) ACM Computer/Human Interaction Conference (CHI), 2008.

[NYT] *A Face Is Exposed for AOL Searcher No. 4417749*. M. Barbaro and T. Zeller. New York Times, August 9, 2006.

[SGZ] *Web-Based Inference Detection*. J. Staddon, P. Golle and B. Zimny. USENIX Security 2007.

[WSC06] *User experiences with sharing and access control*. T. Whalen, D. K. Smetters. E. Churchill. CHI Extended Abstracts 2006: 1517-1522

Mixed Critical Systems

Background

Testing, validation, verification, and certification of unmanned avionic systems, with components of different criticality are critical barriers to rapid insertion, growth and effective utilization of unmanned assets within the military and civilian communities. Increasing software complexity multiplies the cost of its testing and validation to a degree that becomes prohibitive. The adoption and advancement of new and emerging technologies has been hindered by lack of significant evolution of standard certification processes. This is especially true for mixed criticality architectures, which are increasingly common on unmanned aircraft capable of flying and managing missions autonomously. Characterization and the development of acceptable and cost effective methods of software, component, system, and network certification remain a stumbling block to reach the desired NIT goals in the area of High Confidence Software and Systems (HCSS).

Mixed Criticality

Mixed criticality is the concept of allowing applications at different levels of criticality to interact and co-exist on the same computational platform. Unfortunately, certification of such systems is more difficult, because it requires that even the components of less criticality be certified at the highest criticality level. One approach to achieve mixed criticality without increased certification expense and effort is to use ARNC653 time and space partitioning. The ARNC653 approach allocates a predefined faction of CPU time and memory of the whole system to each partition. By defining and restricting available time and space resources by partition; this prevents faults and failures in one partition from corrupting another partition leading to system failures. Each partition can be certified to a different known level of criticality.

An equivalent construct in the security arena is that of Multiple Independent Levels of Security (MILS). MILS is a departure from present operating system architectures that were designed prior to the Internet, when there was little threat of network attacks. As a result, these early systems did not incorporate security as a design requirement and use “fail first, patch later approach to inevitable failures and intrusions. MILS allows the co-execution of various application images, at differing levels of security, with dedicated hardware with strongly controlled data flow between them. A “separation kernel” provides the MILS policy enforcement layer.

Both approaches demand significant investments in time and effort to effectively meet the appropriate certification levels and thus constitute a significant barrier to the adoption of unmanned systems.

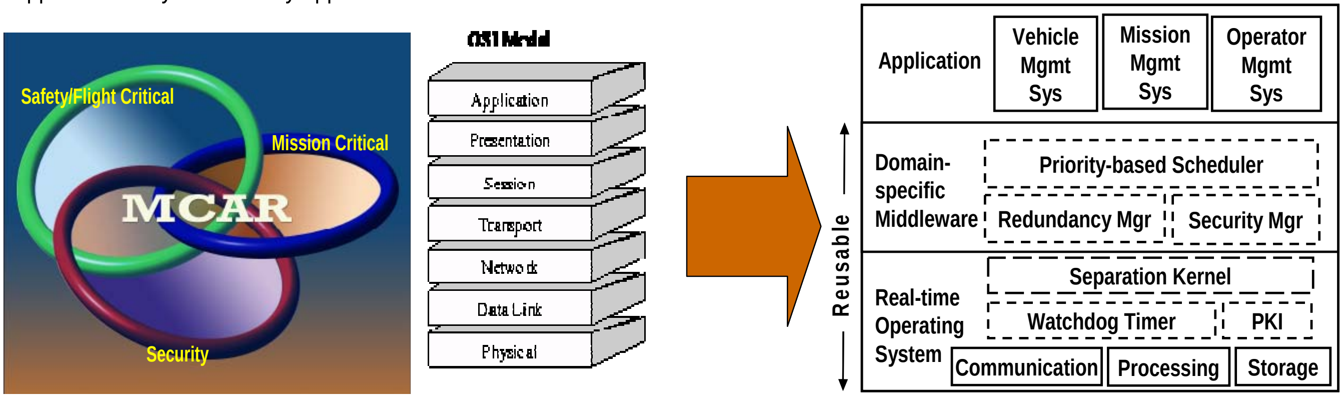
Mixed Criticality Challenges

- Certification of Compostable components: Future systems are likely to be constructed from compostable components with known levels of certifiability. The challenge is to identify, develop and implement both a certification process and a compostability framework that will support compostable and incremental certification.
- Certification by Design: A means of identifying and assessing the “certifiability” of a component, even before it is implemented, is a necessary condition, in order to support both compostable and incremental certification. This includes both formal and analytical methods that seek to define a design process that embeds attributes of certifiability with it.
- Reconfigurable Systems: It is inherently difficult to bind the decision space of an avionics system that has the ability to reconfigure itself under certain failures or contingencies. This poses a challenge to certifying reconfigurable systems form a cost, effort and complexity standpoint.
- Fault Tolerant Systems: Modern mixed critical systems need to exhibit a high degree of fault-tolerance. Canonical or generic approaches to implementing fault tolerance in avionics systems are extremely hard to come by. Consequently, it is even harder to define an approach to certifying the behavior of such systems. Nevertheless, a more flexible process than the one currently available is needed urgently.

Ongoing Efforts

The MCAR program is example of a multi-agency and multi-disciplinary program combining the expertise of government, industry, and academia to address a common issue of mixed criticality certification. The Air Force Research Laboratory (AFRL) joined with the National Science Foundation (NSF) created the MCAR, or Mixed Criticality Architecture Requirements project, which aims to build industry consensus on the problems, challenges and potential solutions to the certification of mixed critical systems. With a strong active commitment of all stakeholders, whether industry, government, or academic; the MCAR project combines the experience and expertise of three major industry players (Boeing, Lockheed, and Northrop Grumman), industry representative experts in real-time operating systems and middleware frameworks, as well as academic experts addressing future mixed criticality architectures with a focus on certification and security.

MCAR aims at trying to solve the problems of certification of today's software by unifying embedded system development with enhanced middleware components development to expand the system certification process. MCAR places an emphasis on component and subsystem reuse by migrating common requirements from the individual applications into a common component-based middleware layer. Combined with upgradeable hardware building blocks, the development of a reusable middleware level provides a path to cost effective certification. The MCAR Program layered software approach consists of a high confidence Real-Time Operating System (HC-RTOS) and Domain-specific Management System specific Middleware. The HCRTOS provides the foundation for the architecture, providing the low-level fault tolerance and separation support for safety and security applications.



Desired NIT Capabilities

In order to address the challenges discussed above, we believe NITRD represents a force for change that can help focus attention, direct funding and establish priorities for future infrastructure. Some of the areas of relevance are:

1. Infrastructure to support Enhanced Certification Process for the certification of emerging mixed-critical avionics systems
2. Cross-Spectrum Technology Push program to mature and transition research and technology in the area middleware frameworks for Mixed Critical systems
3. Technology push programs to develop tools and frameworks to support Cross-Vendor Assessment and evaluations of Real Time Operating Systems

NITRD Program Roles and Functions

Given the broad spectrum of players comprising NITRD, there is little doubt that with the right set of problems and focusing of effort across academia, government and industry will yield positive results. Therefore, the choices of program elements and goals, players as well as mechanisms to implement the goals are critical to its success. Some of these elements are noted below:

1. Development and execution of multi-agency and multi-disciplinary programs to mature the research in the challenge areas of Mixed Criticality
2. A concerted effort to accelerate the transition of research and technology in Cyber Physical Systems into Mixed Critical implementations, tools and frameworks into application realistic environments through a cross-agency effort. Cert.Auth + AFRL + NSF + Industry + Academia = Mixed Criticality Success
3. Targeted Steering of strategic goals, key challenges, opportunities, and research priorities; could include steering available research funding for Academia and small businesses to address Mixed Criticality challenges, as well as providing clearinghouse / dissemination mechanisms across academia, government and industry.
4. Providing a strategic goal to the Mixed Critical Systems community of building an open source test bed and framework for the assessment and evaluation of emerging best-of-breed solutions within realistic application bindings

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

White Paper: Inform the Five Year Strategic Plan for the Federal Networking and Information Technology Research & Development

Mobility in modern society comes at a high cost. Traffic injuries and deaths, congestion-induced delays, and pollution incur enormous costs worldwide. For example, in the U.S. traffic congestion resulted in an estimated 3.7 billion hours of delay and 2.3 billion gallons of wasted fuel in 85 major urban areas in 2003, yielding an estimated cost of $65 billion [1]. Over 6 million crashes occur each year resulting in over 40,000 fatalities and an estimated $150 billion in economic loss [2]. Vehicle emissions are the leading cause of health-related air pollution. Travel demand is expected to rise by 50% in the U.S. to reach more than 4 trillion vehicle-miles traveled by 2020 [3]. Natural and man-made disasters can create gridlock, impeding emergency vehicles and personnel and potentially costing lives. Similarly, transportation has created major societal problems in Europe and developing countries such as China and India. Even modest improvements of only a few percent can result in billions of dollars in savings each year.

Intelligent Transportation Systems (ITS) exploit networking and information technology to alleviate these concerns, and are now widespread. However, ITS deployments have reached an inflection point, and are entering a new era. Costs associated with deploying and maintaining infrastructure, technological advances in sensing and wireless communications, and the lure of increased capability, coverage, and efficiencies are driving ITS toward systems that feature widespread exploitation of in-vehicle sensing, vehicle-to-roadside and potentially vehicle-to-vehicle communication [4-9]. For example, the USDOT's Vehicle-Infrastructure Integration (VII) initiative is a public-private partnership where wireless communication devices are being installed in the nation's vehicle fleet and roadside infrastructure [9-13]. It is estimated that 10% of the nation's vehicle fleet may be instrumented within two years of the commitment to deploy the system [11], resulting in a major shift in the operation of transportation systems in a relatively short period of time. Networking and information technology deployments utilizing in-vehicle and roadside computing and communications will need to be integrated with more traditional infrastructures such as adaptive control devices, embedded sensors, and traffic management.

As such deployments become widespread travelers will depend upon networking and information technology more and more in every day use. This technology will only become visible to travelers when it fails to operate properly. The consequences of such failures will be high, resulting in severe economic loss, and in conditions such as emergencies, failures could result in loss of life. Such deployment must have several key attributes:

- *Resiliency* - Future transportation systems must be resilient, reliable, and robust to failures; their supporting networking and information technology built over large-scale transportation cyberinfrastructures must share these features, particularly under communication failures resulting from natural disasters and attacks.
- *Adaptive and transparent* - The system must be able to automatically adapt to unexpected events such as crashes and roadway incidents. It must be transparent and provide services to travelers without creating driver distractions.
- *Scalability* - such systems must be implementable in rural, suburban, and urban settings, with population ranges from a few thousand to millions. Large wireless networks must provide effective communication services to hundreds of thousands of network nodes.

The realization of effective systems presents enormous technical challenges. Over a large urban area, such a system will involve the interaction of potentially millions of nodes, including vehicles, control and detection devices, and management services. Under emergency scenarios, such a system can be envisioned to reach the individual person level, requiring near-instant communication interaction with

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

hundreds of thousands to millions of individuals in real time. In the future, the ability to manage the transportation system under normal and emergency conditions will not be constrained by the availability of data. The true limitations concern data management, processing, and networking. Deployed systems will be composed of a heterogeneous collection of in-vehicle, roadside, and traditional (e.g., TMC/EMC-based servers) computation and sensor nodes that must analyze current system states, predict future states, and rapidly adapt to unexpected disruptive events on short time scales.

Federal Networking and Information Technology Research and Development must support the study and development of such systems as well as rapid commercialization if they are to achieve their fullest potential. Creating and sustaining this new paradigm presents many major challenges. For instance, what will be the form of the system architecture? Can a resilient system be ensured? How will system recovery be implemented? Can the system have sufficient redundancies while remaining efficient operation? Will the system continue to function under the added stress of emergency events (e.g. natural or manmade)? What vulnerabilities are introduced into the transportation system with increasing automation? Can a communications system be designed that achieves the reliability required to operate such a system? How should the massive amounts of real-time data produced in such a system be processed and mined to extract useful information in a timely fashion? Can approaches to information sharing and data anonymization be found that maintain acceptable levels of privacy among travelers while still providing adequate information to effectively manage the transportation infrastructure? What business models are appropriate in realizing public-private partnerships that meet the requirements of government, industry, and the public at large?

The Federal Networking and Information Technology Research and Development program must support the fundamental research necessary to test and model such systems. For instance, simulation models that capture *both* the transportation and communication infrastructures and their interdependencies are essential to answer questions such as these. These models must be able to capture emerging behaviors of such deployments on a large-scale, e.g., a large metropolitan area. So-called “urban canyons” create major challenges for wireless communication, and cross-layer protocol interactions must be captured to accurately characterize Quality of Service (QoS) properties of the network infrastructure, which ultimately affect the overall behavior of the deployment. This necessitates detailed modeling of the network, including accurate capture of the effects of physical terrain. Such simulations must incorporate detailed models of individual vehicle movement and behaviors, operating in conjunction with detailed, accurate models of the communication infrastructure.

These research challenges are inherently multidisciplinary. Technology development efforts must combine research from transportation, wireless communication, computer science, statistics, operations research, and modeling and simulation, among others. Environmental engineering expertise is needed to help ensure the creation of sustainable infrastructures. Research in the social sciences is required to address issues such privacy, and public policy issues similarly plays a critical role that must be considered in the deployment of such systems.

Effective research and development programs fundamentally require successful multi-agency collaboration. Both in day-to-day and under emergency conditions federal, state, local, and private agencies must be seamlessly networked. Different aspects of the transportation network fall under the purview of each agency, often with several agencies having some shared responsibility over a portion of the system. Another crucial feature of the Federal Networking and Information Technology Research & Development program should be meeting the challenges of integrating the information technology systems of the disparate agencies responsible for operating the transportation system and developing a means to foster and maintain this integration over time.

References

1. Schrank, D. and T. Lomax, *The 2005 Urban Mobility Report*. 2005, Texas Transportation Institute, The Texas A&M University System.
2. Intelligent Transportation Society of America, *National Intelligent Transportation Systems Program Plan: A Ten-Year Vision*. 2002, ITS America in cooperation with the U.S. Department of Transportation.
3. Department of Energy. *http://www.eia.doe.gov/oiaf/aeo/aeotab_7.htm*. 2001 [cited.
4. Tian, J. and K. Rothermel, *Building Large Peer-to-Peer Systems in Highly Mobile Ad Hoc Networks: New Challenges?*, in *Technical Report 2002/05*. 2002, University of Stuttgart.
5. Morsink, P., et al. *Design of an application for communication-based longitudinal control in the CarTALK project*. in *IT Solutions for Safety and Security in Intelligent Transport (e-Safety)*. 2002.
6. Xu, Q., R. Sengupta, and D. Jiang. *Design and Analysis of Highway Safety Communication protocol in 5.9 GHz Dedicated Short Range Communication Spectrum*. in *IEEE VTC'03*. 2003.
7. Ziliaskopoulos, A.K., *An Internet Based Geographic Information System that Integrates Data, Models and Users for Transportation Applications*, in *Transportation Research, Part C*. p. 427-444.
8. Ziliaskopoulos, A.K. and J. Zhang. *A Zero Public Infrastructure Vehicle Based Traffic Information System*. in *TRB 2003 Annual Meeting*. 2003.
9. Bechler, M., W.J. Franz, and L. Wolf. *Mobile Internet Access in FleetNet*. in *KiVS 2003*. 2003.
10. Werner, J. *Details of the VII Initiative 'Work in Progress' Provided at Public Meeting*. 2005 [cited 2005 April 15]; Available from: http://www.ntoctalks.com/icdn/vii_pubmtg_v1.php.
11. Werner, J. *USDOT Outlines the New VII Initiative at the 2004 TRB Annual Meeting*. 2004 [cited 2005 April 12]; Available from: http://www.nawgits.com/icdn/vii_trb04.html.
12. Wilson, C. *The Merits of Vehicle-to-Infrasturcture Communication for Intersection Safety*. Newsletter of the ITS Cooperative Deployment Network 2004 [cited 2005 April 12]; Available from: http://www.nawgits.com/icdn/wilson_itespring04.html.
13. Werner, J. *More Details Emerge about the VII Effort*. Newsletter of the ITS Cooperative Deployment Network 2004 [cited 2005 April 12]; Available from: http://www.ntoctalks.com/icdn/vii_details_itsa04.html.

ATTN: Networking and Information Technology Research and Development (NITRD) program representatives

Thank you for the opportunity to respond to your Request for Information (RFI) for the Five-Year Strategic Plan for the NITRD program. My responses for the RFI include:

- **Multi-core Software Development Tools –** Currently there are few debugging tools for multi-core software development. For instance, improved software debugging methods for multi-core chips is needed. For debugging a software error, the engineer must uncover the processing interactions between the multiple cores that caused the error. Today's current debugging tools can provide insight into an individual core's processing, but these tools are incapable for providing insight to every core's processing simultaneously. Also, additional communication protocols and tools are required for improved inter-core communication for multi-core chips. Many current communication protocols like TCP/IP are designed for very specific environments. The inter-core communication environment requires very fast communication speeds over short distances of less than a few inches over silicon.

  The desired result is for a suite of software engineering tools that allow for easy and convenient multi-core software development. The multi-core software debugging tools should allow software developers to simultaneously debug all threads operating on the multiple cores with each core processing synchronously with the other cores on the micro-processor. One or more communication protocols and tools would be developed for very fast inter-core communication for multi-core chips. Multiple protocols and tool suites may be required for differing chipset design.

  To develop these tools, there should be a multi-disciplinary approach involving academia, software vendors, and hardware vendors. The National Sciences Foundation (NSF) and similar research and development organizations such as the Air Force Research Laboratory (AFRL) should provide program leadership and research direction.
- **Validation/Verification and Certification of Multi-core chips –** Certifying systems comprised of uni-core hardware architectures is very difficult. Now, new certification methods must be devised for multi-core chips. Multi-core chips have the following certification issues:
  - The multiple cores on a microprocessor share L2 cache memory. To certify multi-core chips, techniques and methodologies must be developed to ensure that no core is able to over-write another core's L2 cache memory.
  - In a real-time operating system (RTOS) executing on a multi-core system, a thread is scheduled to execute. Only core 1 is available for executing the thread, so the thread executes on core 1. The next time the thread is scheduled to execute several cores are currently not utilized, which core will the thread execute on? Techniques and procedures must be developed whereby thread execution is predictable and consistent for the system's overall execution.

  The certification, validation, and verification of embedded systems has become extremely expensive as our computer systems become more complex. In 2007, 40% of all processors that Intel shipped contained multiple cores. By 2011, it is expected that 95% of all processors that Intel ships will contain multiple cores. Similar multi-core migration trends are being seen with other microprocessor manufacturers. The sooner that research

organizations begin investigating validation/verification and certification for multi-core chips, the better off the computer industry will be.
To develop these certification techniques and methodologies, there should be a multi-disciplinary approach involving academia, software vendors, hardware vendors, and the National Security Agency (NSA). The NSA, NSF, and similar research and development organizations such as the Air Force Research Laboratory (AFRL) should provide direction and program leadership.

- **Architecture and Implementation Security Vulnerability Discovery –** To reduce the effects of viruses and mal-ware, many network devices execute Intrusion Detection/Prevention software. Many Intrusion Detection/Prevention packages are very good but they require frequent updates to stop the latest viruses and mal-ware. Furthermore, these tools eat up valuable processing time that the software applications they are protecting should be using. It would be beneficial to generate tools that identify security vulnerabilities while developing a system's architecture and code. The earlier these security vulnerabilities are discovered and solutions are developed to fix the vulnerabilities, the less expensive the overall system will be.
To develop these Architecture and Implementation Security Vulnerability Discovery techniques, methodologies, and tools, there needs to be a multi-disciplinary approach involving academia, software middleware vendors, software operating system vendors, hardware vendors and the NSA. The NSA, NSF, and similar research and development organizations such as the Air Force Research Laboratory (AFRL) should provide program leadership and direction.

Multi-Framework Multi-Technology Software Development and Certification of Dynamic Behavior

Introduction

The reality of information technology today is that software systems in the foreseeable future will a) be more dynamic and flexible, b) have to utilize an increasing number of technological advances in both hardware and software, c) be intertwined with increasingly more essential services and critical functions, and d) be composed of COTS components of diverse origin and pedigree. There is a challenge to ensure that the increase in capabilities and flexibility needed by future software systems does not come with a corresponding increase in complexity and unpredictability that limits our ability to reliably build, certify, and deploy them. There remain significant gaps in currently available technical capabilities that prevent system owners, operators and integrators from properly developing and certifying systems with the flexibility they need in order to be deployed in the dynamic environments of today and as envisioned for future. Any 5 to 10 year research plan in the area of information technology must attempt to bridge this capability gap. In particular, two topic areas stand out that need immediate attention: i) complexity of multi-framework and multi-modal system development, ii) certification of dynamic behavior.

Multi-framework Multi-technology System Development

The current trend of Service Oriented Architectures (SOA) and Software as a Service (SaaS) will continue in the near future, implying that new IT capabilities will be *composed* rather than developed as collections of homogeneous components. Individual components are likely to be developed (and tested) using a wide range of technology elements including different programming languages, development environments, supporting libraries and even programming paradigms (e.g., publish-subscribe vs. point to point). It will not be uncommon in a system where C code embedded in VB scripts associated with cells in an Excel spreadsheet that is populated by scanning online data sources be visualized with Java/Google maps and delivered to your browser via web services. Driven by forces in the globalized economy, it would be hard to ascertain the level of testing each of the components underwent or the environments in which each component was tested—it may even be impossible to ascertain the origin of the components. While technology is facilitating cooperation and coexistence of these kinds of multi-paradigm multi-technology components that cross from one application domain to another, human stakeholders face a tremendous challenge—the learning curve for these different paradigms and technologies remain steep, and it is not easy for an expert to adapt to new software domains. A software engineer who is an expert in web services will not be able to easily grasp the C/VB intricacies that might be quite natural to an Excel jockey. With every new wave of technology a significant wealth of acquired human expertise becomes obsolete and needs to be reacquired. Either human stakeholders need to be brought up to speed (through training or hiring), which incurs cost and delay or the organization makes a conscious decision to remain with older generations of technology (which may prove costly from another perspective where newer technologies may offer more efficient solutions). A third option entrusts outsiders with critical responsibilities which incurs both cost and additional risks. It is time to invest in new research that focuses on *reducing* *the learning curve associated with change.* The changes to be investigated by the multi-framework and multi-modal system development are at various levels ranging from programming language and software libraries, to software engineering and software domain. Recent advances in semantic metadata and mark up offers a potential solution for conveying abstract program structure, data flow and semantics. It is foreseeable that source or intermediate code (such as Java byte code and MS CIL) can be annotated to carry structural, dataflow and even semantic information described in terms of information technology ontology. Automated tools can then be used on the annotated source or byte code to extract such descriptions and present it to the stakeholder or to check compliance of vendor/stake holder specified invariants about structure, flow or semantic. Code-generation tools can generate code from constructs involving these high level structures and flow in the desired programming language, for a specific software development environment or with respect to specific libraries for run time analyses and checking. The goal is not to verify units of functionality or generate the code for entire functional components, which distinguishes this research from prior automated program verification or code-generation work.

Certification of Dynamic Behavior

There has long been a need to *certify* software for deployment in certain environments, such as those that deal with life-critical, safety-critical, financial, space, and military environments. Current focus on certification/accreditation of software primarily focuses on process—how the software is constructed—and testing. Many of the techniques used for certification of software are essentially risk assessment approaches, convincing a certification authority to accept the risk that the software is correct by documenting the process by which it was built and the thoroughness of its testing. This has led to two trends for software needing certification: (1) building *stovepiped* software from scratch because reusing open-source, off-the-shelf, or third-party software presents difficulties in certifying the resulting system; and (2) developing static, inflexible software because of the increased complexity in certifying dynamic, flexible systems. However, looking forward to future systems for the next 5-10 years and beyond, these trends cannot continue. As stated above, it is becoming increasingly critical to build new systems by composing existing components and services. Before too long, it will simply not be cost effective or feasible to build new systems from scratch (and may already not be). Furthermore, certified systems are no longer deployed only in the closed, embedded environments of yesterday. Systems are increasingly internetworked, deployed in unpredictable environments, and deployed for longer lifetimes, which means that they must be flexible to handle dynamic interconnections, inputs, environmental conditions, and unforeseen conditions and interactions. Systems of the future simply will not be able to certified using a process that requires documentation of the creation of every line of code, or the enumeration of and testing of every possible state it can enter. Research In new approaches to software certification—approaches more suited to composed and dynamic systems—needs to be included in the 5 to 10 year information technology research plan. Promising areas of research include looking at automated annotation of software components and services with metadata describing their performance, QoS, provenance, security, and other characteristics to aid in their reuse; *safe composition* techniques that allow software to be reused, but limits its ability to *misbehave* in the integrated system; proof carrying code; open-source certified real-time operating systems; software isolated processes; and composition languages, frameworks, and tools.

New Paradigms for Verification and Certification of Systems of Cooperating Medical Devices

Over the last two decades, rapid advances in computing power, personal computing platforms, computer networking, and interoperability standards and middleware for system integration have had a dramatic impact on many sectors of the national and international economy. These technological advances are now poised to revolutionize health care systems.

- Pervasive networking will enable integration of national networks, regional health care centers, local hospitals and clinics, primary care physician offices, home care systems, and body-area networks.
- Health-care information technology infrastructures will be oriented toward "systems of systems" architectures built upon middleware information backplanes that integrate and blend monitoring and treatment devices with other information producers and consumers in the extended healthcare network. Device data streamed into medical records will be automatically mined to extract knowledge that can drive a host of activities such as automated treatment, dosing, trend analysis geared toward health prediction, and large-population assessments of human health and treatment effectiveness.
- With information technology as a catalyst, health care systems will increasingly exhibit collective intelligence built upon the intelligence of individual devices and data mining and knowledge gathering components.
- Operating rooms and other diagnosis and treatment contexts will shift from the use of a collection of fixed monolithic devices to plug-and-play components that enable flexible and rapid re-configuration of diagnostic, recording, and treatment systems
- Precision robotics and high-speed networks will hasten advances in telemedicine and robotic surgery.
- Portable healthcare devices will support multiple care contexts, and boundaries of these component-based systems and their extended information environments will be difficult to define.
- As generations of technology-savvy healthcare consumers enter retirement, these consumers will embrace and even demand sophisticated home healthcare monitoring, treatment, and records systems integrated with national information databases (e.g., prescription drug information systems) and local hospital and primary care systems.

Considering the huge costs associated with transitioning increasing numbers of aging citizens into private and government-sponsored medical insurance plans, it is especially important that these advances provide quality medical care at a reasonable cost. The U.S. government must therefore facilitate technology innovations as well as innovations in medical processes, workflows, and regulatory policies that reduce the cost of medical care but still provide the highest levels of safety, security, and overall quality.

Unfortunately, industry and government find themselves in a rather shocking situation – the technology exists to *assemble* many of the types of medical systems described above, but the technology to *guarantee the safety and security* of these systems is lacking. Moreover, many government agencies such as the Food and Drug Administration currently *do not have the regulatory regimes in place* to provide oversight to ensure the safety and effectiveness of these systems.

While many modern medical devices have some form of connectivity that may be used to upload information to an electronic health record or provide computer-based audits, they are still monolithic in nature. Likewise, current Verification and Validation (V&V) techniques used in industry primarily target single monolithic systems. Moreover, FDA's regulatory regimes are designed to approve single stand-alone devices – there are no guidelines in place for how the industry might bring to market a collection of cooperating medical device components from different vendors, where each component goes through a separate pre-market approval process. Today, V & V activities typically account for as much as 50% of the overall cost in bringing devices to market. The increasing amount of software in modern medical monolithic devices is already pushing current V&V technology and regulatory regimes to their limits.

Moving to highly integrated, componentized, distributed systems of medical devices such as those envisioned above renders current V&V and regulatory regimes obsolete. This situation will increasingly result in significant safety risks. In the best case, these emerging systems will be "shoe-horned" into existing V&V and regulatory frameworks that are unable to cope with their complexity. In the worst case, as health-care organization are under significant pressure to reduce costs, clinical technicians will begin to "home roll" their own unapproved device integrations at extreme risks to patients.

Assessment

To meet these challenges, a paradigm shift is needed in V&V technology, training, and regulatory paradigms. Bringing about this shift will require significant coordination and cooperation between academic researchers (to develop alternate V&V technologies), government regulators (to propose new regulatory regimes that recognize those technologies and can accommodate componentized systems of medical devices), and industry (to evaluate proposed technologies in realistic environments, and to jointly develop standards for integration). This cannot be achieved by incremental expansion of existing research programs at NSF and other agencies. Instead, a high-level coordinated inter-agency effort with new research programs targeted to this problem will be necessary. It should also be emphasized that this is not just a technology problem, but also a regulatory regime problem and culture problem. Significant funds must be devoted to developing and evaluating new regulatory regimes along with appropriate education and workforce training resources for those regimes.

Specific Recommendations

- Develop new research programs at agencies such as NSF, NIH, and NIST that target V&V technology, new regulatory paradigms, and integration standards that can support complex "systems of systems" of medical devices. These programs should prioritize research programs that involve collaborations with industry and government agencies.

- Within existing software and embedded system research programs, emphasize research on technologies that can lead to higher confidence in large-scale safety/security critical systems. This includes new approaches to certification such as assurance cases as well as techniques that lead to the production of independently auditable evidence of correctness (e.g., as recommended in the National Academies study "Software for Dependable Systems: Sufficient Evidence?")

- Provide infrastructure grants to develop test beds for medical device integration and coordination research that can be released as resources for academic, industry, and government research.

- Provide grants that support research internships at FDA that target medical device integration technology.

- Continue sponsorship of workshops and other meetings that bring together government regulators, clinicians, medical device developers, along with technical domain experts (e.g., in interoperability, verification/validation, and certification).

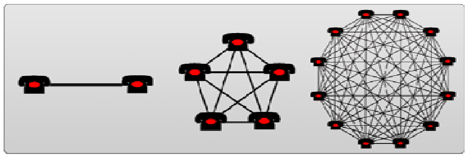
NITRD Collaboration 2.0
Response to
RFI to inform the NITRD Five-Year Strategic Plan
August 25, 2008

NITRD Future State Vision

The Collabworks/Touchstone Consulting Consortium represents the best of innovative process & methodology to leverage Web2.0, collaborative project technology, and proven government agency transformation & innovation. We have chosen to focus our response on a key element of NITRD's success; ***Collaboration 2.0***. Our combined thought leaders view this time-value infrastructure-like capability as a critical element in the adoption of true "**network-effec**t" driven organic innovation and rapid problem solving with an open channel for continuous adoption of commercial leading practices. Over the five year horizon NITRD will be more integrated with needs, more nimble to competitive threats, and far more leveraged with R&D efforts that matter both domestically and among international partners. This Collaboration 2.0 model will shift away from the traditional top down doctrine that causes many critical initiatives to perish from sheer inertia.

Metcalfe's Law is the core value proposition of Collaboration 2.0

Collaboration 2.0 is not about tools, meetings, or events. It about identifying shared problems and issues within a network of users, leveraging solutions among users and forming trusted relationships with suppliers addressing a network of users. Metcalfe's law **(now called "the network effect")** states that the value of a telecommunications network is proportional to the number of users of the systems. , if only one person had a phone, the system wouldn't be very useful, but as more phones are added, the value of the network increases dramatically. In other words, the more that participate in the network the greater the value to all participants.



The Internet itself and several social sites, including Wikipedia, eBay and Skype are clear examples of network effects – the more that participate the more value for all the participants. Several important books have been written in the past three that describe the evolution of society, consumerism, and business as a result of the network effect. Included among them are: The Long Tail, Open Business Models, Wikinomics, The Wisdom of Crowds, and The Starfish and the Spider. CollabWorks is the catalyst that enables the network effect for entities. NITRD can now hardness a collaborative infrastructure to transform events, forums, and meetings into problem solving networks that grows organically as the network effect creates more value for each participant.

The Collaboration 2.0 challenge

- Maintain pace innovation pace and relevancy with industry evolution driven by global competitive forces
- Traditional trickle down R&D from government to commercial will matter less. Foreign states, particularly rough states will leverage commercial evolution rather than internal R&D
- Therefore, NITRD strategies, dialog, action need to flow from an commercial-government network

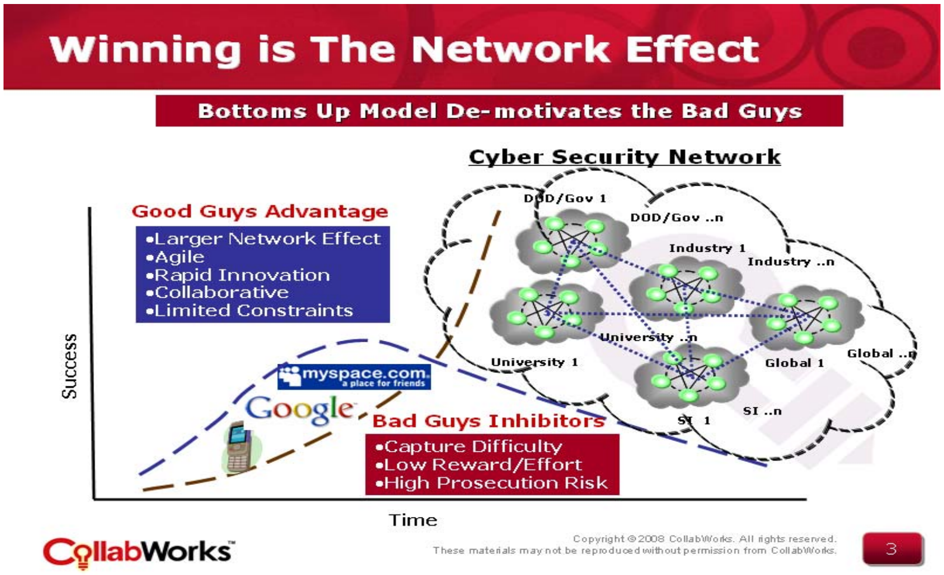
NITRD Development & Execution

- Focus on the macro, act on the micro; don't allow "business-as-usual to stifle collaboration & innovation
- Organic, not top down; 1-5 year horizon demands asymmetric problem solving capability. Our adversaries use this advantage, beating top down approaches with less cycle time and far fewer resources
- User driven demand pull via collaboration – meaning the solution starts with the user not the supplier
- Collaborative services infrastructure with tool, processes, knowledge capture and retrieval, expert networks - all driven by collaborative processes shaped by the participants
- Project solution model which combine social networking, sourcing, and business process management

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

Vision & Realization

- IT operations will be global, decentralized, with plug and play architectures to leverage virtual computing services (internal and outsourced) where the concept of the perimeter will give way to the concept of a meta management layer with system network, manageability, optimization and security built in. Knowledge and solutions will flow among a participating network of entities as shown below.
- Adversaries will find capturing valued information increasingly difficult and far less rewarding.
- 1000’s of entities contributing transparently in an organically growing network.
- 1000’s of rapid innovation projects that are driven by users that include both government and industry where suppliers are invited into the process (demand pull, not supplier push0
- Global framework and government standards driven by participating networks not top down standards groups



Winning is The Network Effect

Bottoms Up Model De-motivates the Bad Guys

Good Guys Advantage

•Larger Network Effect
•Agile
•Rapid Innovation
•Collaborative
•Limited Constraints

Bad Guys Inhibitors

•Capture Difficulty
•Low Reward/Effort
•High Prosecution Risk

Copyright © 2008 CollabWorks. All rights reserved.
These materials may not be reproduced without permission from CollabWorks.

About Us
CollabWorks™ is bringing the network effect to a network of entities

Our mission: Simplify the processes for effective Entity to Entity Collaboration™.

CollabWorks™ has the expertise, processes, and legal and collaborative infrastructure to help your enterprise build trusted relationships, solve important problems, and share them among a growing network of enterprise participants. These companies leverage their respective strengths and typically save 3-4x over typical outsourcing strategies. By fostering customer collaboration networks, suppliers benefit from consolidated customer requirements and more aligned value chains. Domain experts and thought leaders benefit by extending their skills across a network of enterprise participants.

The Touchstone Consulting Group is a leading provider of specialized strategic management consulting services tailored to the specific needs of government organizations. Since its founding, Touchstone has focused on enabling leaders in both the public and private sectors to realize results on a national-scale.

We collaborate with our government customers to deliver measurable results today and into the future. Specifically, Touchstone’s solutions are designed to help government agencies manage complex, technological, and transformational initiatives within tight budgetary constraints and with limited human resources.

Orchestrating Computation and Physical Dynamics: Response to RFI for NITRD

August 21, 2008

This response considers the orchestration of computing with physical processes. It argues that to realize its full potential, the core abstractions of computing need to be rethought to incorporate essential properties of physical dynamics. All branches of computing, networking, and systems theory are affected, resulting in a research agenda that can only be effectively pursued as a national or international effort.

Most microprocessors today are embedded in systems that are not first-and-foremost computers. They are cars, medical devices, instruments, communication systems, industrial robots, toys, games, etc. Key to these microprocessors is their interaction with physical processes through sensors and actuators. Such microprocessors increasingly resemble general-purpose computers. They are becoming networked and intelligent, often at the cost of reliability. At the same time, general-purpose computers are increasingly being asked to perform complex interactions with physical processes. They integrate media such as video and audio, and through the migration to handheld platforms and pervasive computing, sense physical dynamics and control physical devices.

The foundations of computing do not support this mission well. In the Turing-Church abstraction, computation is about the transformation of data, not about physical dynamics. This abstraction deliberately omits the passage of time, an essential property of physical dynamics.

Computer scientists and engineers have heavily exploited this omission. The central fact is that "correct" execution of nearly any computer program has nothing to do with how long it takes to do anything. Engineers have to step outside the programming abstractions to specify timing properties. Timing needs to become a correctness property rather than a quality of service measure. This requires profound changes in computing and networking.

Computers have become so fast that surely the passage time in most physical processes should be able to be handled without special effort. But then why is the latency of audio signals in modern PCs as large as it was 20 years ago? Audio processes are quite slow by physical standards, and a latency of a large fraction of a second is enormous. To achieve good audio performance in a computer (e.g. in a set-top box, which is required to have good audio performance), engineers are forced to discard many of the innovations of the last 30 years of computing. They often work without an operating system, without virtual memory, without high-level programming languages, and without memory management, and they use microprocessors without caches, dynamic dispatch, or speculative execution. Those innovations are built on the key premise that time is irrelevant to correctness. By contrast, what these systems need is not faster computing, but physical actions taken at the right time. It needs to be a semantic property, not a quality factor.

But surely the "right time" is expecting too much, the reader may object. The physical world is neither precise nor reliable, so why should we demand this of computing systems? Instead, we must make the systems robust and adaptive, building reliable systems out of unreliable components. Clearly systems need to be designed to be robust, but we should not blithely discard the reliability we have. Electronics technology is astonishingly precise and reliable, more

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

than any other human invention. We routinely deliver circuits that will perform a logical function essentially perfectly, on time, billions of times per second, for years. Shouldn't we exploit this remarkable achievement?

This problem is going to get worse. As embedded systems become more networked and intelligent, their character fundamentally changes. They are no longer black boxes, designed once and immutable in the field. Instead, they are pieces of a larger system, a dance of electronics, networking, and physical processes. An emerging buzzword (that few seem particularly fond of) for such systems is *cyber-physical systems* (CPS). Such systems will unquestionably have an enormous impact on technical dominance.

Applications of CPS have the potential to rival the 20-th century IT revolution. They include high confidence medical devices and systems, assisted living, traffic control and safety, automotive systems, process control, energy conservation, environmental control, avionics, instrumentation, critical infrastructure control (electric power, water resources, and communications systems for example), distributed robotics (telepresence, telemedicine), defense systems, manufacturing, and smart structures. It is easy to envision new capabilities that are technically well within striking distance, but that would be extremely difficult to deploy using today's methods. Consider, for example, a city with no traffic lights, where each car provides the driver with adaptive information on speed limits and clearance to pass through intersections. We have in hand all the technical pieces for such a system, but achieving the requisite level of confidence in the technology seems decades off.

The challenge of integrating computing and physical processes has been recognized for some time, having motivated for example the emergence of hybrid systems theories. These theories blend the dynamical systems models of electrical engineers with automata models of computer scientists. But the effort needs to extend beyond systems theories into the abstraction stack on which engineers build applications. Today's computing and networking technologies *unnecessarily* impede progress towards CPS applications, and dynamical systems theories unnecessarily omit software and network behavior.

The solution pervades the abstraction stack. Beginning bottom-up, computer architects have gone overboard exploiting the irrelevance of timing to achieve better performance. Multi-level caches, dynamic dispatch, speculative execution, and bus architectures are all notable culprits. The research challenge is to achieve comparable performance with predictable and repeatable timing.

Continuing up the stack, there is a long history of attempts to insert timing features into programming languages. These are generally done, however, without fundamental changes in the semantics of the languages, particularly with regard to concurrency. Domain-specific languages with temporal semantics (e.g. Simulink, LabVIEW) have firmly taken hold in some areas, showing that there are alternatives. These are radically different from imperative and functional languages that dominate the programming language community. But they remain outside the mainstream of software engineering, are not well integrated into software engineering processes and tools, and have not benefited from many innovations such as data abstraction and strong type systems.

An attractive alternative to new programming languages is notations that work at the level of tasks or components. These are attractive because they can exploit experience with conventional imperative or functional languages, which can be used to specify detail functionality, and they can embrace models of physical systems. Various innovations in coordination languages and actor models look promising, but require adaptation to express temporal dynamics. I envision an innovation like what C++ did to C, which provided language support for object-oriented design. For example, an actor model with temporal dynamics could be supported by a C++++ that introduced notations for expressing concurrency and timing.

Similar research will be needed in systems theory, semantics, verification, security, operating systems, and networking. The emphasis needs to be on predictable repeatable dynamics rather than on performance optimization. This requires more than incremental improvements, which will, of course, continue to help. But effective orchestration of computing and physical processes requires semantic models that reflect properties of interest in both.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

Performance, Robustness, and Cyber Security of Critical Infrastructure Systems – A Cyber-Physical Systems Research Theme

Background

Critical infrastructures are complex physical and cyber-based systems that form the lifeline of modern society, and their reliable and secure operation is of paramount importance to national security and economic vitality. The US President's Commission on Critical Infrastructure Protection (CCIP) [1] has identified telecommunications, electric power systems, natural gas and oil, banking and finance, transportation, water supply systems, government services, and emergency services, as the eight critical infrastructure systems. These infrastructures are not only complex and intertwined with electric power and cyber systems, but are highly interdependent among themselves and hence a disruption in one infrastructure will have cascading effects on others [1, 2]. The disruption could be due to natural events, such as hurricanes, earth quakes, and wild fires or due to man-made malicious events, such as physical destructions or electronic intrusions into infrastructure systems. Identifying, understanding, and analyzing such interdependencies among infrastructure systems pose significant challenges [2-4]. These challenges are greatly magnified by the geographical expanse and complexity of individual infrastructures and the nature of coupling among them. The rest of the discussion focuses on electric power infrastructure.

The electric power grid, as of today, is a highly automated network. A variety of communication networks are interconnected to the electric grid for the purpose of sensing, monitoring, and control. These communication networks are closely associated with the supervisory control and data acquisition (SCADA) systems in the network. The data provided by the SCADA system is utilized in the energy management systems (EMS) of the power grid, for a wide range of system operation functions and real-time control of the power grid. Currently, the electric power infrastructure does not have adequate measures to guarantee protection against many forms of natural and malicious physical events on the infrastructure, which makes it highly vulnerable [3-6]. One of the primary concerns has been the issue of large-scale fault events and their impact on the overall performance and stability of the electric grid. Various incidents in the recent past [3-5] have indicated the extent to which the electric grid is vulnerable and the urgent need to protect them against physical and electronic faults and intrusions.

Theme 1: Performance and Robustness of Critical Infrastructure Systems

Research: There is need for fundamental research harnessing the enabling power of sensor networks for real-time monitoring and control of complex dynamical critical infrastructure systems. The goal should be to significantly improve the robustness and performance of the electric energy grid through innovative embedded sensor network design and associated data aggregation, fault diagnosis, and decision algorithms [6].

The overarching goal should be to lay a strong foundation for design and analysis of embedded sensor networks whose optimization/constraints are governed by the underlying dynamics of the physical system. One approach is to deploy sensors in critical and vulnerable locations of the power systems to sense mechanical properties of its various components and transmit the sensed data through a suitable

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

wireless network to the central control center, and fuse the information with existing data for the electrical quantities in the system to arrive at an ideal preventive or corrective control decision..

Theme 2: Cyber Security of Critical Infrastructure Systems

Three modes of malicious attacks on critical infrastructure are generally envisioned: 1) Attacks upon the system - The system itself is the primary target with ripple effects throughout society, 2) Attacks by the system - The population is the actual target, using parts of the system as a weapon, 3) Attacks through the system - The system provides a conduit for attacks on other critical infrastructures. In some sense, the cyber system forms the backbone of nation's critical infrastructures, which means that a major cyber security incident could have significant impacts on the safe operations of the physical systems that rely on it.

Security threats against utility assets have been recognized for decades [3-5]. Insecure computer systems may lead to catastrophic disruptions, disclosure of sensitive information, and frauds. Cyber threats result from exploitation of cyber system vulnerabilities by users with unauthorized access. A potential cyber threat to supervisory control and data acquisition (SCADA) systems, ranging from computer system to power system aspects, is recognized. It is shown that an attack can be executed within an hour once the computer system security is compromised. The ever increasing power of the Internet facilitates simultaneous attacks from multiple locations. The highest impact of an attack is when an intruder gains access to the supervisory control access of a SCADA system and launches control actions that may cause catastrophic damages. Another primary concern has been the possibility of massive denial of service (DoS) attacks on the SCADA control system and the resulting impacts on the overall performance and stability of the electric power systems.

Research: The overarching research should be to develop a comprehensive cyber security framework for critical infrastructure systems integrating the dynamics of the physical system as well as the dynamics of the cyber-based control network. The integration of cyber-physical attack/defense modeling with physical system simulation capabilities must be developed to quantify the potential damage a cyber attack can cause on the physical system in terms of capacity/load loss, equipment damage, or economic loss in other forms [7]. The integrated model should provide a foundation to design and evaluate effective countermeasures, such as mitigation and resilience algorithms against large scale cyber-based attacks.

Conclusion

These challenging problems call for inter-disciplinary research collaboration between physical as well as cyber systems researchers, and call for laying strong foundation for inter-disciplinary educational program on cyber-enabled critical infrastructure systems focusing on protection, security, resiliency, and sustainability.

References

[1] Presidents Commission on Critical Infrastructure Protection, Critical Foundations: Protecting Americas Infrastructures (1997). [Online]. Available at: http://www.ciao.gov/.

[2] M. Rinaldi, James P. Peerenboom, and Terrence K. Kelly, "Identifying, understanding, and analyzing critical infrastructure dependencies, *IEEE Control Systems Magazine,* pp. 11-25, Dec. 2001.

[3] Control Systems Security Program, US-CERT, Department of Homeland Security. http://www.us-cert.gov/control_systems/

[4] M. Amin, "Security challenges for the electricity infrastructure," *IEEE Security and Privacy Mag.*,

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

vol. 35, no. 4, pp. 8–10, Apr. 2002.

[5] G. Ericsson, “Toward a framework for managing information security for an electric power utility - CIGR´E experiences,” *IEEE Trans. on Power Delivery,* vol. 22, no. 3, pp. 1461–1469, Jul. 2007.

[6] R. A. Le´on, V. Vittal, and G. Manimaran, “Application of sensor network for secure electric energy infrastructure,” *IEEE Trans. on Power Delivery,* vol. 22, no. 2, pp. 1021–1028, Apr. 2007.

[7] C.-W. Ten, C.-C. Liu, and G. Manimaran, "Vulnerability assessment of cyber security for SCADA systems," *IEEE Trans. on Power Systems*, to appear, 2008.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

NITRD Strategic Plan
White Paper Submission

This note will argue that a major focus of NITRD strategy over the next five years should be on **Principles of Embedded Software**.

Software controls an enormous range of devices, ranging from safety-critical systems such as anti-lock braking systems in automobiles, cardiac pacemakers, and aviation flight-control; to systems with national defense and security implications such as missile guidance and command and control; to consumer devices such as cell phones and microwave ovens. The importance of such embedded software truly amounts to a silent revolution in device design; new features are routinely implemented using off-the-shelf sensors, actuators and microprocessors, with the specific desired behavior captured in source code. This new control-engineering paradigm has opened tremendous opportunities for great improvements in existing devices such as cars; indeed, widely quoted estimates from General Motors executives estimate that 90% of new feature content, and consequently profits, in automobiles will be software-driven. Embedded software has also enabled the development of devices, such as cell phones, that would not have been possible otherwise.

It is widely reported in the trade press that over 98% of new microprocessors are used in embedded applications: all of these microprocessors run embedded software. Already, however, shortages in the talent needed to develop this software can be observed, and are indeed driving the use of off-shoring as a stop-gap measure. Without revolutionary advances in the efficiency with which such software is developed, verified, certified and deployed, the pace of innovation in devices will necessarily slow, and our country's competiveness and security even threatened.

The European Union faces similar pressures in its device-oriented industries and has mounted a high-profile series of research initiatives, variously called Artemis / Artist, with a total 10-year budget of over €1.5bn, aimed at improving embedded-system design practices. The initiatives combine basic and applied research projects conducted by conglomerations of university, government and industry researchers. It is vitally important that the US also invest in the area of embedded systems and software in order to give its industrial base the scientific advances it needs to continue to innovate and compete.

By its very nature, embedded software is multi-disciplinary. Developing it requires input not only from software experts but also from control and system engineers with deep knowledge of the device domain (medical, automotive, aerospace, etc.), with their differing environmental assumptions and regulatory regimes. Advancing the state of the art in embedded software development practices will therefore require effective multi-agency collaborations within US Federal research and regulatory agencies. For example, medical devices represent an important societal good as well as a significant market niche for US companies; of special-interest are high value-added products such as radiological equipment, implanted cardiac

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

devices, and infusion pumps. Software currently accounts for a substantial portion of the development and certification costs of such equipment. New technologies for software design and verification would help, but would require support not only from basic-science agencies such as the NSF and NIH but also others such as the FDA. Similarly, in the aerospace arena, regulatory agencies such as the FAA should be involved in the support of embedded-software research that would also be funded by NASA, DARPA and the Air Force. An effective model would involve the shouldering of research costs by the research agencies, with the regulatory agencies providing capabilities for evaluation and input to technology development. This model is being used with some success by Fraunhofer via an NSF-funded project on medical-device software being conducted with input from the FDA, and expanding this model appropriately would offer significant avenues not only for research advances but for evaluation and future commercialization.

Input to the NITRD request for its' Strategic Plan

August 25th 2008

The Open Science Grid (OSG) serves high throughput collaborative science by providing a national distributed infrastructure in the US, federating with other national and international cyber-infrastructures to meet the needs of broad communities of science and research at all scales. The OSG vision is of a sustained, growing infrastructure which transcends administrative boundaries and scientific domains, and through practical application of advanced IT technologies and methodologies contributes to the ubiquitous and democratic use of a global, shared, distributed computational eco-system for scholarship, disseminated to society and industry at large.

The technical directions and principles of the OSG are driven by the needs of the stakeholders: the Large Hadron Collider (LHC) about to start a two decade research program at CERN, the Laser Interferometer Gravitational Wave Observatory (LIGO) which is on the edge of being upgraded to LIGO providing more than 10 times the sensitivity, and the chemists, biologists, climate and other applications groups that are collaborating and contributing to the OSG Consortium.

Towards these goals we input the following components to the NITRD strategic plan:

1) *R&D in the manageability and integration of advanced networks and middleware* providing distributed compute and data management technologies to provide more effective and usable heterogeneous end-to-end systems for researchers – from small groups of ad-hoc collaborators to large multi-national scientific initiatives: Improving the intelligence of routing and bridging components to enable seamless use across administrative and technological (performance, capabilities) domains.

2) *Cross-agency development and implementation projects which address the full life-cycle computational needs of the scientific research community.* A particular need is the integration of the nation's high performance and leadership class facilities at DOE and NSF with the campus infrastructures being deployed at 100s of the nation's universities. The research life-cycle increasingly relies on the full spectrum of computational scales.

3) *R&D into interoperation of Federated cyber-infrastructures at all scales - local to worldwide.* Interfacing/Routing/ Bridging between autonomous self-managed infrastructures nationally and internationally. (cf network bridges/ routers - at the next "transport" or service layer up the end-to-end stack). These approaches will increase our national competitiveness in the global research economy.

4) R&D addressing the sociology, governance, practice and technical needs of Collaborative Research Communities at all scales. This include leading the way in understanding how to conduct international treaties and collaborations for open shared cyber-infrastructures. We have existing examples in the production systems of the large physics and astrophysics collaborations that can be built on for other scientific domains.

5) *Internationally co-sponsored partnerships of small targeted research projects* targeted specifically to augment and prove the R&D and enable ad-hoc, dynamic, innovation and discovery. These would stimulate focused activities for independent researchers globally to find each other and collaborate democratically on shared hypotheses. This builds on the experience and successes of NITRD in working across multiple agencies in the national arena.

Response to RFI for Five-Year Strategic Plan for the Federal NITRD Program

Q. What do you imagine as the future in terms of desired NIT capabilities?

While NIT advances over the past two decades have brought about tremendous improvements in individual productivity and organizational efficiency, their impact on the societal infrastructure – for transportation, health-care, electricity/water distribution – has been marginal. There exists a tremendous potential for substantial reengineering of the societal infrastructure – related to movement, distribution, awareness and participation – that can be made possible by the emerging NIT capabilities. The interdependencies between these systems and their adaptabilities provide opportunities for resource conservation, better quality of life and improved societal dialogue & conduct. Crude examples based on extrapolation of what we see today include intelligent highway systems (versus single, independently-operated vehicles), dynamically configured and integrated intensive care or emergency transport units (versus separately-operated sensors or support systems), and intelligent home and industrial environments (versus traditional heating, lighting, and air conditioning subsystems). Much more is possible when the new technological underpinnings have been devised.

But these advances will not happen as a matter of entirely proliferation of commercial technologies. To be sure there are fundamental challenges: scientific, societal and economic. Scientifically, a whole new engineering discipline is needed to deal with phenomena at the intersection of logical and physical processes, societal needs for policy, equity and fairness (especially as infrastructure investments involve use of public funds.)

Q. What roles do you imagine for the NITRD Program and for the academic, commercial, international, and other domains in achieving that future?

The funding for S&T has undergone dramatic changes due to changes in the agency priorities. Even as the changes in individual funding sources have been gradual their cumulative impact is beginning to challenge the fundamental structure and assumptions underlying the federal (and increasingly state) outlay for supporting S&T research. One of the major pillars that has come under increased stress and questioning is the NSF directed research as rooted in the NRC act, and going back to its genesis in NDRC of 1940. The original concept leading up to formation of "Research University" was based on policy pillars spelled out in the act. The most important among these was the role of NSF in directing *unsolicited* research.

Current realities make such unsolicited research as one of the least attractive options for individual researchers; in many subject areas this is not even an option. In this environment, NSF as a research funding institution faces dramatic challenges in creating programs that are forward looking and truly reflective of the original spirit behind the NRC act that directly lead to the successful "research university" in the post world-war era. This success was in no small measure attributed to the leveraging effect of capability-seeking S&T 6.2 funding by DARPA as well as successful consortia funding such as Sematech and SRC. All of these institutions have undergone changes with the primary effect being the removal of their amplifying effect on the NSF/DOE sponsored research. NSF now faces the real danger of creating programs that are likely to mislead – rather than benefit from – the tremendous diversity of S&T talent in the nation; by creating artificial priorities and programs that do not reflect the true state of knowledge which is increasingly tied to the practice.

NITRD has a special role. NITRD has the capability to take a broader view of the state of knowledge and technology, and thus advise individual agencies of not only the opportunities that they should pursue, but also the moral imperatives emerging from such a broader view. As it does so, it must strive to overcome the barriers and mindsets regarding knowledge created across institutions, across industry and academia. In particular, it is no longer true that universities or a group of academic working directly in areas related to IT accurately represent the state of the art in knowledge; or that they know the best opportunities for advancing the state of knowledge. Indeed, even as industrial labs have changed, the state of the art – and visibility into such knowledge – has shifted to small-scale enterprise and their backers as in the leading edge venture capital community. This visibility into VC community is important because as a group they would not support investments necessary to advance societal infrastructure, and yet they provide – through their investments – important technology pieces that dramatically reduce the costs in retooling infrastructure through commercialization. Further and specifically in the context of advancing the state of the knowledge in *cyber-physical systems (CPS),* the programs should also seek to meaningfully engage public-works agencies, such as state DOTs, water resources management. These will be crucial to technology transition efforts into practice.

Q. Technical challenges that should be addressed

There are many scientific challenges underlying the theory and engineering needed to build CPS systems. Many of these have been spelled out in related reports by NSF organized workshops. To be specific, I will focus on one such challenge. Current CPS systems lack the ability to capture spatial information – information related to the location of actions as well as the use of location information in defining actions. While geographical location information can be 'stored' in various forms and at various levels, semantic support to use this information at various levels of the system implementation is severely lacking. When building CPS, this limitation manifests in many ways: from inadequately specified CPS functionalities and their validation to a lack of any guarantees related to availability or unavailability of computational resources as a function of location. Cyber-physical systems are at one level embedded sensor systems: they react to and manipulate spatiotemporal sensory information. *Yet these lack models and methods to capture such information, validate their behavior, performance against timing and spatial requirements.* This presents a fundamental barrier to the scaling and use of cyber-physical systems to societal-scale applications because a whole host of constraints, from energy, power, bandwidth, processing to resource availability, simply rule out the use of all the sensors at all times. To achieve the goal of semantic support for location and time at all levels, we need to address the following technical problems: (a) How do we capture location (and timing) information into CPS models that allows for validation of the logical properties of a program against the constraints imposed by its physical (sensor) interactions? (b) What are useful models for capturing faults and disconnections within the coupled physical-computational systems? How can we reason with these models to define the notion of system availability? (c) What kind of properties that can be verified, and assertions that can be ensured in applications that make use of both physical (real) time as well as location information? Do these propositions require direct algebraic support for location? How best these location and timing aware assertions can be validated? (d) What programming model is best suited for CPS applications utilizing dynamic behaviors? Are there any specific operating system or 'middleware' services that can ease the task of building such applications, and doing so reliably? (e) What are the metrics to measure effectiveness of physically-coupled embedded systems? How do we characterize operational efficiency with measures that take into account spatial information?

To answer these questions, our shared research challenges span the choice of abstractions (models and methods), programming models that use location information, to infrastructural support for location (virtualization, location determination/validation support ser-vices, etc). Promising projects to address these challenges will span programming, formal methods, distributed and embedded real-time systems, and whole host of disciplines that come together in building sensors and sensor network applications.

Regulation-relevant network science and engineering

Both the Internet revolution and the Wireless revolution have moved beyond a purely technical domain. They are core parts of the national infrastructure and billions of dollars are spent on them every year. However, these are interesting times in that the Internet and Wireless communication are subject to *both* radical *technological advancements* and radical *regulatory developments*. Unfortunately, the current structure of federal institutions is not able to deal with radical developments in both simultaneously. Policy makers are asked to write and implement forward-looking regulations that reflect the collective will of the citizenry and are likely to be effective in the real world. Political processes can help understand the will of the public but they must be complemented with sound science and an understanding of the engineering trade-offs.

Consider the Internet first. It has revolutionary possibilities to tap the creativity of the citizenry and enable new forms of community engagement and collective action. At the same time, economic incentives play a major role. However, the network protocols were originally designed without attention to such incentives. Consequently, the Internet is faced with severe market failures and inefficiencies that can and should be corrected. Yet, the network neutrality debate is politically charged. Both the two extremes (full neutrality and zero neutrality) are undesirable for economic or public-good reasons. Research is essential to understand the trade-offs involved so that middle-ground approaches can be found. Similarly, the poor level of security in the Internet results in large part from the absence of proper incentives for users to protect their computers. The trade-offs of mechanisms (e.g. certification agencies, insurance schemes, due care or liability policies, rebates on access fees for protected computers) to provide such incentives must be understood before good policies can be formulated.

The issues in the wireless domain are even more stark. For decades now, the paradigm has been that users required a license to operate a wireless system and that any piece of wireless equipment was restricted to a few spectrum bands wherein they obeyed fixed protocols. This was how interference was precluded. Advancements in circuits driven by Moore's law has brought us to the brink of radio equipment that are frequency-agile and "cognitive" in the protocols that they follow. This promises to have a revolutionary impact that dovetails with the impact of the Internet and personal computing. It now technically conceivable that two people in their garage will be able to roll out their own wireless service and have a reasonable chance of growing organically as they grow their user base without disrupting other wireless systems. This will enable rapid and creative innovation without holding it hostage to either standards committees or the business plans of the incumbents. Wireless spectrum policy must be reformulated to take this into account and the underlying trade-offs need to be understood in order to do this with a suitably light hand of regulation. The core question is to determine the fundamental limits on the overhead required for regulation the same way that we have such guidance from Shannon's theory for spectral efficiency.

For both the Internet and Wireless systems, the main federal agency relevant to policy-making is the FCC. How can it get access to the trade-offs relevant to making policy decisions? The traditional approach is to just call together a panel from the National Academy of Sciences and the National Academy of Engineering and asking them for a report of our current understanding. The problem is that we do not have the body of research results required to provide that understanding. So where is that basic research going to come from? Who is going to pay for it?

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

The challenge is that both the Internet and wireless revolutions will have their impact in part due to enabling "creative destruction" as new business models and new ways of doing things emerge. Thus, existing industry cannot be trusted to take the lead in doing this sort of research since incumbent players will be accused of taking a side in favor of their own economic interests. A similar criticism applies to startup companies. Venture capital firms are less likely to invest in startups pursuing technologies that can only be deployed if the regulatory environment changes. Even if they do, the current patent system in the United States would encourage them to try to get key patents that would then lead incumbents to viscerally oppose any such regulatory change since they would be locked-out of benefiting from it.

Government research labs could in principle take the lead in developing the relevant understanding. The problem comes in finding people --- the past structure of science and engineering education in the United States has not produced many people who are simultaneously technically trained and have an appreciation for the kinds of policy trade-offs involved. Therefore, these research problems are best handled within Academia where commercial interests are not driving the choice of questions and students can be trained in a way to appreciate these new kinds of trade-offs.

The question is then which agency or agencies should be funding the work that needs to be done. In the past, DARPA has done a good job of taking the lead in exploring focused new areas. However, the DARPA culture tends to be deliverable-oriented and has rarely funded the kind of basic theoretical research required here. Other agencies might also be uncomfortable with having regulation-oriented basic research coming out of DARPA --- would it be perceived to favor military interests over civilian ones? The National Science Foundation has an excellent track record of funding unbiased theoretical research. The challenge is that the existing peer-review culture is quite conservative in today's Malthusian funding climate. This could be remedied by giving certain new pots of money in the name of federal agencies like the FCC. This money could be accessed by NSF program managers only to fund projects that help provide the basic scientific and engineering foundations for the regulatory questions likely to be relevant to that agency in the future.

Science Cartography:
Communicating, Navigating, Managing, and Utilizing our Collective Scholarly Knowledge Across Disciplinary, Cultural, and Geospatial Boundaries

I believe it would be highly beneficial if 13 Federal agencies would collaborate on federating the existing publication, patent, funding datasets with the goal to make this vast amount of data/knowledge and expertise easy to understand, navigate, manage, and utilize. The resulting dataset can also be used to analyze, model, and (visually) communicate the structure and dynamics of science and technology in support of science policy and economic decision making. A more detailed argumentation and discussion of benefits for different user groups follows below.

The number of currently active researchers exceeds the number of researchers ever alive. Researchers publish or perish. Some areas of science produce more than 40,000 papers a month. We are expected to know more works than one can possibly read in a lifetime. We receive many more emails per day than can be processed in 24 hours. We are supposed to be intimately familiar with datasets, tools, and techniques that are continuously changing and increasing in number and complexity. All this while being reachable twenty four hours, seven days a week.

Not only library buildings and storage facilities, but also databases are filling up quicker than they can be built. In addition, there are scientific datasets, algorithms, and tools that need to be mastered in order to advance science. The center figure depicts just how much information exists. No single man or machine can process and make sense of this enormous stream of data, information, knowledge, and expertise.

All this leads to a quickly increasing specialization of researchers, practitioners, and other knowledge workers; a disconcerting fragmentation of science; a world of missed opportunities for collaboration; and a nightmarish feeling that we are doomed to ‘reinvent the wheel’ forever.

Over the last several hundred years, our collective knowledge has been preserved and communicated via scholarly works such as papers or books. Works might report a novel algorithm or approach, report experimental results, or review one area of research among others. Some report several small results others one large result like the Human Genome. Some confirm results while others disprove results. Some are filled with formulas while others feature mostly artistic imagery. Different areas of science have very different publishing formats and quality standards. The description of one and the same result, e.g., a novel algorithm, will look very different if published in a computer science, biology, or physics journal.

Unfortunately, our tools to access, manage, and utilize our collective knowledge are rather primitive. In fact, our main means of accessing everything we collectively know are search engines. Search engine providers try to make us believe that we can live in ‘flatland’. That is, no directory structures are necessary and data objects can have any name. Superior retrieval software will find whatever we need. However, the usage of search engines resembles charging a needle with a search query and sticking it into a haystack of unknown size and consistency. Upon pulling the needle out, one checks the linearly sorted items that got stuck on it. This seems to work well for fact-finding. However, it keeps us always at the bottom of confirmed and unconfirmed records in the haystack of our collective knowledge. Of course, we can explore local neighborhoods of retrieved records via Web links or citation links. However, there is no ‘up’ button that provides us with a more global view of what we collectively know and how everything is interlinked. Search engines are a bad choice when it comes to identifying patterns, trends, outliers, or the context in which a piece of knowledge was created or can be used. Without context, intelligent data selection, prioritization, and quality judgments become extremely difficult.

Consequently, even the smartest people on this planet can neither keep up with the amount nor the accelerating speed of knowledge production. This becomes a major concern if scientific results are needed to enable all human beings to live a healthy, productive, and fulfilling life. We simply need better tools to keep track, access, manage, and utilize our collective scholarly knowledge and expertise.

Maps of science that guide our scholarly endeavors promise to make a difference. They aim to serve today’s explorers navigating the world of science. These maps are generated through scientific analysis of large-scale scholarly datasets in an effort to connect and make sense of the bits and pieces of knowledge they contain. They can be used to objectively identify major research areas, experts, institutions, collections, grants, papers, journals, and ideas in a domain of interest. Local maps provide overviews of a specific area: its homogeneity, import-export factors, and relative speed. They allow one to track the emergence, evolution, and disappearance of topics and help to identify the most promising areas of research. Global maps show the overall structure and evolution of our collective scholarly knowledge.

Science maps have been designed for diverse users and information needs as discussed subsequently.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

Science Maps as Visual Interfaces to Scholarly Knowledge

The number and variety of available databases in existence today is overwhelming. Databases differ considerably in their temporal, geospatial, and topical coverage. Database quality reaches from ‘downloaded from the Web’ to ‘manually curated by experts’. Visual interfaces to digital libraries provide an overview of a library’s or publisher’s holdings as well as an index into its records. They apply powerful data analysis and information visualization techniques to generate visualizations of large document sets. The visualizations are intended to help humans mentally organize, electronically access, and manage large, complex information spaces and can be seen as a value-adding service to digital libraries.

Mapping Intellectual Landscapes for Economic Decision Making

A deep understanding of technology, governmental decisions, and societal forces is required to make informed economic decisions that ensure survival in highly competitive markets. Discontinuities caused by disruptive technologies have to be determined and relevant innovations need to be detected, deeply understood, and exploited. Companies need to look beyond technical feasibility to identify the value of new technologies, to predict diffusion and adoption patterns, and to discover new market opportunities as well as threats.

The absorptive capacity of a company, i.e., its ability to attract the best brains and ‘to play with the best’ has a major impact on its survival. The importance of social networking tools and network visualizations increases with the demand to understand the "big picture" in a rapidly changing global environment.

Competitive technological intelligence analysis, technology foresight studies, and technology road mapping are used to master these tasks. Easy access to major results, data, tools, expertise is a key to success. But exactly what is the competition doing, who makes what deals, and what intellectual property rights are claimed by whom?

Last but not least, companies need to communicate their image and goals to a diverse set of stakeholders -- to promote their products, to hire and cultivate the best experts, and to attract venture capital.

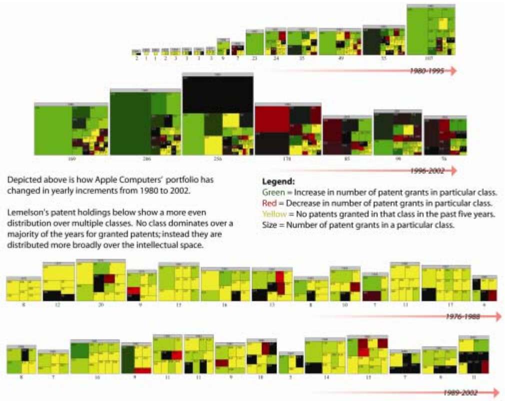


Figure 1 Claiming Intellectual Property Rights via Patents
The evolving patent portfolios of Apple Computer (1980-2002) and Jerome Lemelson (1976-2002) are shown here. The number of patents granted per year matches the size of the square. Each square is further subdivided into patent classes which are color coded in green if the number of patents increased, in red if it decreased, and yellow if no patent was granted in this class in the last five years. While Apple Computer claims more and more space in the same classes, Lemelson’s patent holdings are distributed more broadly over the intellectual space.

Science of Science Policy (Maps) for Government Agencies

Increasing demands for accountability require decision makers to assess outcomes and impacts of science and technology policy. There is an urgent need to evaluate the impact of funding/research on scientific progress: to monitor (long-term) money flow and research developments, to evaluate funding strategies for different programs, and to determine project durations and funding patterns.

Professional science managers are interested to identify areas for future development, to stimulate new research areas, and to increase the flow of ideas into products. Hence, they need to identify emerging research areas; understand how scientific areas are linked to each other; examine what areas are multi-disciplinary, measure collaborations and knowledge flows at the personal, institutional, national, and global level; identify and compare core competencies of economically competing institutions and countries; identify and fund central and not peripheral research centers, etc.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

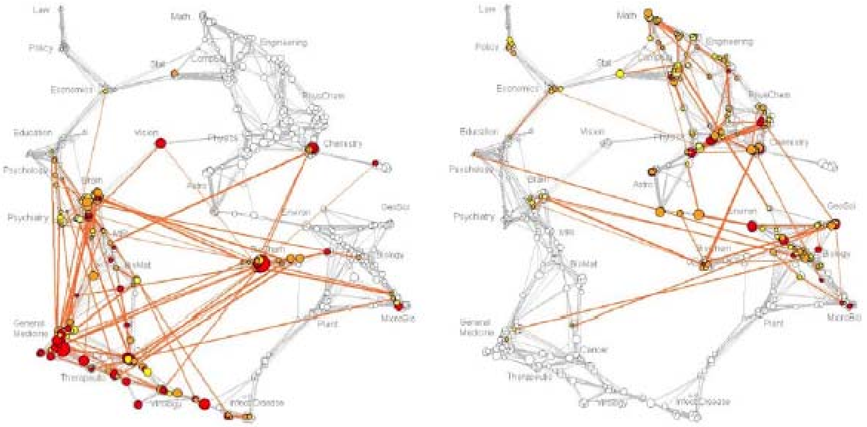


Figure 2: Funding Profiles of NIH (left) and NSF (right)
Using a base map of science, the core competency of institutes, agencies, or countries can be mapped and visually compared. Shown here are funding profiles of the National Institutes of Health (NIH) and the National Science Foundation (NSF). As the base map represents papers, funding was linked by matching the first author of a paper and the principal investigator using last name and institution information. A time lag of three years between funding of the grant and publication of the paper was assumed. While NIH mostly funds biomedical research, NSF focuses on math, physics, engineering, computer science, environmental and geo sciences, and education. Overlaps exist in chemistry, neuroscience, and brain research.

Professional Knowledge Management Tools for Scholars

Most researchers wear multiple hats: They are researchers, authors, editors, reviewers, teachers, mentors, and often also science administrators.

As researchers and authors, they need to strategically exploit their expertise and resources to achieve a maximum increase of their reputation. Expertise refers to the knowledge they already have or can obtain in a given time frame but also expertise that can be acquired via collaborations. Resources refer to datasets, software, and tools but also to people supervised or paid.

They need to keep up with novel research results; examine potential collaborators, competitors, related projects; weave a strong network of collaborations; ensure access to high quality resources; and monitor funding programs and their success rates. Last but not least, they need to review and incorporate findings and produce and diffuse superior research results in exchange of citation counts, download activity, press, etc.

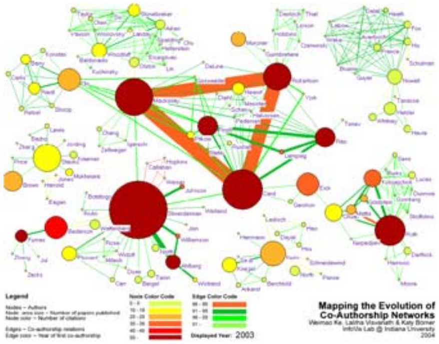

Legend
Nodes – Authors
Node area size – Number of papers published
Node color – Number of citations
Edges – Co-authorship relations
Edge color – Year of first co-authorship
Node Color Code
Edge Color Code
Displayed Year: 2003
Mapping the Evolution of
Co-Authorship Networks
Weimao Ke, Lalitha Visvanath & Katy Borner
InfoVis Lab @ Indiana University
2004

Figure 3: Mapping the Evolution of Co-Authorship Networks
This is the last frame of an animation sequence that shows the evolution of authors (nodes) and their co-authorship relations (links) in the domain of information visualization. Node area size coded reflects the number of papers an author published, its color denoted the number of citations these papers received.
Link width equals the number of co-authorships and link color denotes the year of the first collaboration. Large, dark nodes are preferable. Large, light nodes indicate many papers that have not (yet) been cited. Shneiderman -- working in a student dominated academic setting – has a very different co-author environment than Card, Mackinley, and Robertson which are at Xerox Parc for most of the time captured here.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

As editors and reviewers, researchers act as gatekeepers of science. They need detailed expertise of their domain as well as related domains of research to ensure that only the most valuable and unique works become manifested in the eternal pile mountain of scholarly works.

As teachers and mentors, they provide students with a deep understanding of the structure and evolution but also the peculiarities of a domain of research and practice. They might give an overview of the teaching material to be covered first; then highlight major people, important papers, and key events; and provide a red thread or pathway and help students discover, understand, and interrelate details.

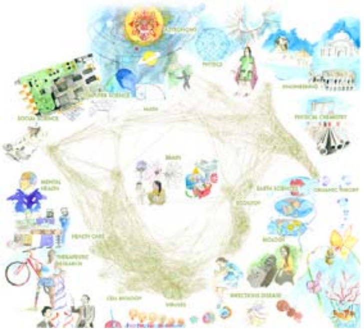
As science administrators, they are responsible for decisions regarding hiring and retention; promotion and tenure; internal funding allocations; budget allocation; outreach. Here a global overview of major entities and processes and their temporal, spatial, and topical dynamics is important.

Science Maps for Kids

In school, children learn about many different sciences. However, they never get to see science ‘from above’. Hence, they have a very limited understanding regarding the sizes of different sciences or their complex interrelations.

How will they be able to answer questions such as: What intellectual travels did major inventors undertake to succeed? Why is mathematics necessary to succeed in almost all sciences? How do the different sciences build on each others work? Or how would they find their place in science and where would it be?

Imagine a map of our collective scholarly knowledge would hang next to the map of the world in each class room. Imagine students could not only travel our planet online but also explore the web of knowledge. How would this change the way we learn, understand, and create?



See also http://scimaps.org/kids/Web_topic_big.jpg - different base map as funding overlay

Figure 4: Hands-on Science Maps for Kids invite children to see, explore, and understand science from above. This map shows our world and the places where science is practiced or researched. A complementary map shows major areas of science and their complex interrelationships. Children and adults alike are invited to help solve the puzzle by placing major scientists, inventors, and inventions at their proper places. Look for the many hints hidden in the drawings to find the perfect place for each puzzle piece. What other inventors and inventions do you know? Where would your favorite science teachers and science experiments go? What area of science do you want to explore next?

References

Chen, C. (2002) *Mapping Scientific Frontiers.* Springer-Verlag, London.

Börner, K., C. Chen, and Boyack, K. W. (2003) Visualizing Knowledge Domains, in *Annual Review of Information Science & Technology*, B. Cronin, Editor. Information Today, Inc./American Society for Information Science and Technology: Medford, NJ, pp. 179-255.

Shiffrin, R. M. and Börner, K. (eds.) (2004) *Mapping Knowledge Domains*. PNAS. Vol. 101 (Suppl. 1). PNAS.

Boyack, K.W., R. Klavans, and Börner, K. (2005) Mapping the Backbone of Science. *Scientometrics*, **64** (3), 351-374.

Börner, Katy, Dall'Asta, Luca, Ke, Weimao & Vespignani, Alessandro. (2005). Studying the Emerging Global Brain: Analyzing and Visualizing the Impact of Co-Authorship Teams. *Complexity*. Vol. 10*(4)*, 58-67.

Börner, Katy. (2009). Atlas of Science: Guiding the Navigation and Management of Scholarly Knowledge. ESRI Press.

Mapping Science Exhibit online at http://scimaps.org and on display at the National Science Foundation, 4201 Wilson Boulevard, Arlington, VA.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

Response to NITRD RFI
White Paper to Inform the Five-Year Strategic Plan
Soft Real-time Computing in Cyber-Physical Systems

It is common for cyber-physical systems to be designed to support real-time processing, i.e., the system is required to respond to external conditions within a bounded amount of time. Historically such systems are designed to incorporate strict *hard real-time* (HRT) scheduling policies that *guarantee* the system will respond to each external event before a prescribed deadline elapses. In essence, failure to meet the guarantee means that the system has failed, usually resulting in the system being restarted (all ongoing tasks managed by the cyber-physical system are halted, then restarted). In certain cases, (e.g., mission-critical task processing) HRT policies are required to avoid loss of human life or failure of a large, expensive system (e.g., an unmanned space mission). A slightly relaxed variant uses the same admission tests and technology as strict HRT systems, but simply kills a job that overruns its deadline, and then continues operation (without restarting the system).[1] It is quite difficult to design pure HRT systems, since systems every possible execution path of the code must have bounded latency; the presence of multiple tasks can cause unexpected delays due to secondary activities of other tasks (e.g., I/O) or synchronization.

By 1990, researchers realized that many real-time applications would be acceptable even if they occasionally failed to meet the traditional HRT assurances – people began to build real-time system using *soft real-time* (SRT) policies. There are many approaches to implementing SRT policies, e.g., many depend on the way that a job is managed if it misses its deadline, i.e., there is no generally accepted policy for what the system should do in this case. SRT is widely used for tasks such as audio/video streaming tasks (including application domains such as security or observation video and digital voice communication over a network), reading sensor data (such as temperature monitoring), controlling actuators (such as redirecting a tracking telescope). Depending on the exact nature of the application software, SRT systems can support high confidence cyber-physical systems at far less cost; in some cases satisfactory SRT systems can be built when it is infeasible to build a pure HRT system. In SRT systems, different facets of HRT execution are relaxed; for example: jobs in a task may be permitted to overrun their deadline by a fixed amount of time; pre specified percentage of jobs in a task may be permitted to miss their deadlines, such that the system minimizes the amount of time that the collection of software tasks misses their respective deadlines; job may be scheduled so that on *average*, a software tasks does not exceed a bound on the resources it uses (although in any given short period of time it may exceed that bound); jobs may be delayed for an entire task phase; etc.

SRT methodology is greatly influenced by HRT methodology, typically sharing the same assumptions, early principles, and proven approaches for HRT systems. For example, early HRT schedulers often used rate monotonic (RM) scheduling policies since they satisfy well-known bounds for admission and scheduling that were published in 1973, and since RM was widely used in HRT system in the next two decades. The earliest deadline first (EDF) scheduling approach was an early competitor with RM, and has certain desirable properties in SRT systems (e.g., it can have better processor utilization than RM); however in early kernel scheduler technology, it was more difficult to implement than RM. EDF schedulers have been shown to have certain

[1] Real-time workload is typically specified as a collection of *tasks*, each of which is executed as a series of *jobs*. Thus a task is often periodic, with one job executing in each period of the task. Each job is specified with an a priori service time and deadline.

advantages in SRT system, yet it is sometimes difficult to publish SRT research, or to obtain funding for a grant proposal, if it depends on EDF scheduling.

Along the same lines, SRT system applications are typically designed under the assumption that all software tasks are implemented as well-informed, well-intentioned HRT style programs that, e.g., would not attempt to make one statement of resource need to the system, but then either naively or covertly *far* exceed that stated need.[2] For example, in SRT systems that use slack time scheduling techniques, the scheduler is likely to ignore the possibility that a task could effectively launch a denial of service attack on the real-time scheduler simply by greatly understating the service time (earning high admission and scheduling priority, but using much greater amounts of resources if they are available). This greatly influences the breadth, complexity, and availability of application software for SRT systems; each SRT system typically uses its own specialized set of applications rather than drawing from an open pool of application programs.

The essence of this input statement is that SRT systems are an engineering reality for most cyber-physical systems, but that SRT systems and software technology could provide far more useful support to practical cyber-physical systems than is currently possible using the existing hybrid theoretic foundations derived from classic HRT systems.

I imagine a design environment in which SRT software is required to state its resource requirements *a priori*, but in which cyber-physical system applications can be built using modern software techniques and constraints. Designers should be able to easily choose a SRT (or hybrid HRT/SRT) system policy based on pure SRT criteria rather than as softening of HRT; if the particular cyber-physical system requires strict HRT management, then it is should be built using the classic HRT principles. But if the cyber-physical system can be built with favorable cost-reliability tradeoffs or alternatively infrastructure (such as EDF scheduling), the design environment should support a spectrum of policies that enable the designer to relax certain policies while maintaining other, and have a clear understanding of the implications of such policy changes.

I imagine a real-time based design environment in which the cyber-physical system designer can incorporate a much more diverse set of applications from diverse software providers while still be able to make the selected assurances about the real-time behavior of the system. Success in this area would greatly increase the amount and quality of software that could be used in cyber-physical systems.

It is very difficult to “reset” an entire discipline, such as soft real-time computing, by reexamining basic principles, since many results in the area rely on assumptions built into the logical design environment. However, modern cyber-physical systems are being held back by the aging set of assumptions that were developed for hard real-time systems, and by the creation of sound SRT methodology through the “sprawl” of HRT methodology. In other cases, the technology is limited by principles that are simply accepted as “the way to do things,” even though technology has evolved (or could evolve) to accommodate better approaches. An effort to reexamine and update the base assumptions cannot possibly be done by any single researcher, or even a single funding agency; it is a movement that can only be by an agency such as NITRD.

[2] HRT systems rely on each job’s service time being specified as its worst case execution time, which causes the scheduler to operate using the most conservative admission and scheduling approach. SRT systems typically depend on the applications to make more aggressive service estimates, but ones that approximate the WCET, or perhaps the average job execution time for all jobs in a task.

White Paper to Inform the Five-Year Strategic Plan for Federal networking and Information Technology Research & Development Program

1. Previous Five-year Strategic Plan for NITRD's IT-R&D is unstructured. It looks like F&A-like address or comments. The part-to-part and section-to-section are loosely coupled. It looks not professional. Fortunately, the assessment report titled "Leadership under Challenge: Information technology R&D in Competitive World" can be used to serves as the background review and current-status of the IT R&D in USA. It provides useful information to design a state-infrastructure with a clear roadmap for the next IT- R&D. The strategic plan should be addressed clear to guide the accomplishment of measurable and tangible goals and aims. Such strategic plan must be engaged to "American Competitiveness Initiative" protected by American Competitiveness Act Bill, and referenced by recent summary titled "Rising above the Gathering Storm," from NAS, NAE, and Institute of Medicine. Therefore, the vision may be

   > "Construct an undefeatable US networking and information technology (NIT) for the nation towards the 21st. century development.

   Mission May be

   > "Assist the US government to coordinate with major federal agencies, to federate national resources, to strengthen national IT capacities, to initiate concrete programs, and to enhance the national IT-R&D for sustaining the global leadership in networking and information technology (NIT) for the nation's competitiveness (including nation's defense and security, economics blooming, health and environmental improvement, education fostering).

2. The NIT-R&D programs may be framed in to the following groups base don the priorities (none of the following challenges can be faced and solved without IT R&D efforts)
   a. National security and defense
   b. National healthcare
   c. Global environmental crisis
   d. Energy crisis
   e. National economy through information-driven intelligence and decision making in global economics and marketing
   f. Scientific discovery and high-end computing
   g. Education networking
   h. …

3. Technical implementation goals must be tangible and measurable.
   a. Multi-agency cooperative projects. NITRD services as a national coordinator and assessor.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

b. Develop multiple layers networks (today's networking is a single layer and still regional based). It is extremely important to national security and homeland defense

c. Develop dedicated healthcare IT networking for cyber-medicine and cyber-health information systems (currently is organizational based). It impacts US healthcare system, such as tele-radiology, tele-medicine, virtual hospitals/clinics etc. The future health information systems in hospitals and clinics are absolutely based on the new IT infrastructure.

d. Develop region-to-region, state-to-state, and country-to-country networking and information to access environmental data, forecasting, disaster responding, and human life and property protections.

e. Crisis networking systems to handle any disasters (earthquake, tornado, global warming, flooding, etc), in case there is not power supplies, how can the information still networks for communications

f. Develop government sponsored public-private partnership networks for example, such systems allow provide intelligent mechanism embedded in the products for global economics. Such systems or products are not only IT software driven, but also networked. In way, it establishes networks between customers and enterprise for future decision making and knowledge, while interactively support customer needs without through regular computing systems. (That is can be called embedded computing in IT networks)

g. Continue to provide NIT R&D for unprecedented computing resources (High-end computing, and data intensive computing networks for scientific and engineering models and simulations and data intensive computing for discovery and innovative CAD designs, to sustain and grow US global competitiveness.

h. Promotion the network for enterprise software that should be highly emphasized. For example, how to oversee the data transfer in specific region or state or country? How to monitor the data transfer like traffic transportation or air flight monitor and management?

i. Develop educational networks, such as virtual schools for K12 education, college education, STEM promotion, cyber virtual museum and educational games.

j. Develop programs to conduct R&D results to practical technical transfers

k. In order to understand the global competitiveness, IT-R&D must have a project to conduct serious study and launch international collaborations with other countries' networks and protocol standards, global cooperation or international cooperative research without lose our strengths for competitiveness.

Response to NITRD Request for Input.

Networking and Information Technology is among the most pervasive and profound technologies affecting all aspects of scientific, societal, as well as economical advances. Establishment of a national/international strategic plan is timely and much needed. We believe that NIT's impact is no longer confined in first generation applications such as voice/data/video transmission. An NIT infrastructure to support the next generation application will be most beneficial and significant in the decades to follow.

Description of Information Sought:

What do you imagine as the future in terms of desired NIT capabilities?

- NIT is expected to provide ubiquitous and secure information communication capabilities for real-time applications in a massive scale. First generation applications such as voice, data, and video transmissions have already impacted global economy in an unprecedented scale. The use of NIT in network control is just emerging but will have significant consequence in our way of living. Second generation applications include real time network control of factories, land/sea/air vehicles, hospitals, farm, entertainment etc. In order to support such a generational leap of advanced network infrastructure a theoretical development and commercialization advancement must be processed in coherent and synergistic manor This is possible only with significant national investment and leadership. Furthermore, the future technological development must include broad based market impact, to avoid any digital-divide type of deficiencies.

What roles do you imagine for the NITRD Program and for the academic, commercial, international, and other domains in achieving that future?

- NITRD program should establish cross agencies, international collaborations among university, industries, and federal research laboratories for several high priority applications. For such a complex process, the strategic vision must include a roadmap jointly developed by government, academia, and industry. Furthermore, program development should be an international effort with inputs from all stakeholders to maximize the benefit and minimize replication. The highest priority as we see it is the definition of standards. While NIT is evolving and maturing, a shift from spontaneity to guided development should occur. At present, there are numerous hardware, software, and implementation protocols. The transition into a new generation of technology is not a smooth process and, thus, the lack of standards can be very detrimental.

In addressing these questions, submitters are challenged to present views and input on one or more of the following subjects, in relation to NIT:

Development and execution of multi-agency and multi-disciplinary programs

- It is envisioned that NITRD acts as a broker and a facilitator to arrange workshops, symposium, portals, funding, initiatives in order to successfully fulfill the vision. Besides offering programmatic initiatives, NITRD should encourage industry–led projects and multi-national projects. The latter is especially significant as presently, multi-national research projects are difficult to fund, set-up, and execute.

Determination of strategic goals, key challenges, opportunities, and research priorities

- NITRD needs a strong industry component. However, industry participation is a very difficult factor to quantify; as the experience in NIST Advanced Technology Projects had demonstrated. Nevertheless, a mechanism must be set up to encourage industry to participate whole heartedly by transforming industry from a technology receiver to an active member. The key metrics for evaluation of success and the traditional proposal evaluation process must be broadened.

Examples that illustrate the impact of realizing the vision, achieving the proposed goals, and meeting the identified challenges

- Example 1: Integrate global resource challenges to solve high priority issues (e.g., pollution, energy crisis) in the world that cannot be solve otherwise.
- Example 2: Establish a massive ubiquitous and secure information communication network for real-time transportation applications. Every single vehicle on the road could interact with their neighbor vehicles, as well as connect to local base stations. The local base stations are connected to the central station. The central station will coordinate with base stations to perform real-time analyses on the traffic conditions, perform real-time optimization using distributed computing, and inform each vehicle the optimal route (e.g., minimal traveling, minimal energy, minimal CO_2 emission, and improved safety) they should select based on the information provided from each vehicles' on bord electronics, e.g., GPS. Many of these concepts have already been used and some are commercially available for individual cars. This individual optimization certainly will not lead to globally optimal results. However, by tasking all of the vehicles and infrastructure to create a real-time and globally optimal solution would be relatively basic; yet the payback would also be substantial. Combining this technology with congestion charging infrastructures / cities will provide opportunities for consideration research and environmental improvements.

Transition of R&D results into practice

- Success of R&D transition depends on a number of factors: industry involvement in basic research phase, academic involvement in the deployment phase, and government involvement in all phases. It should be noted that the transition process must be done with a global mindset.

Role of the U.S. in the international NIT arena

- The NIT arena is increasingly competitive in the first generation applications. To maintain leadership, focused innovation is the key. With its prowess in basic research, creativity, and infrastructure, the United States should take on the leadership role. We should learn and actively involve other partners in terms of research, development, and manufacturing. NITRD is encouraged to generate centralized information clearing house and an office to assist the assembly of such information.

Interactions among government, commercial, academic, and international sectors

- Traditionally, government financial support has been a strong catalyst for interactions. For the scale of NIT, many other issues such Intellectual Properties (domestic and international), international regulations/arbitration, etc. must be resolved. Multi-national corporations' experience can be a helpful input to this process.

Coalition for
Academic
Scientific
Computation

CASC

August 25, 2008

These comments are submitted on behalf of the Coalition for Academic Scientific Computation (CASC), an educational 501 (c) (3) non-profit organization with 54 member institutions representing many of the nation's universities and computing centers. Our members have been involved with the NITRD from its creation and they support the goals and objectives of the NITRD.

The NITRD legislation, from its inception in 1991 through the amendments provided in the Next Generation Internet Research Act of 1998 (Public Law 105-305) and more recently the America Competes Act of 2007 (Public Law 110-69), continues to serve as a viable vehicle for research and development needed to maintain and advance U.S. competitiveness.

The current NITRD Program Strategic Goals, as authorized, are as relevant today as when they were promulgated and the interagency committee structure of the NITRD, with the involvement of the Office of Science and Technology Policy, and the support of the Office of Management and Budget, has recognized that the program requires vigorous and active attention to evolving technological, social and economic advances. This internal examination has been expanded by PCAST through their review and discussed in the report of the Networking and Information Technology subcommittee.

This current activity, the development of a new NITRD strategic plan, is perhaps the most important. Unfortunately the success of any interagency plan involving some 13 Federal agencies depends on steadily increasing funding, with long term commitments, and not the cyclical increases and decreases that we have recently faced. No recommendations that we can offer to the NCO can solve this overriding reality. Long-term steady funding for the NITRD program must be an integral part of the Federal budgeting process. The recommendation for long term funding was stated in the 1999 PITAC report "Information Technology Research: Investing in our Future" and is as valid today as it was then.

Additionally, it is important that the NITRD "do more to exploit existing technology transfer mechanisms," exploring opportunities to expand interactions with the private sector as well as with the academic community. The increasing pace of technological change mandates that the tripartite approach – government, industry and academia - collaborate to ensure that the investment in the NITRD is an investment in U.S. innovation and competitiveness.

We acknowledge the importance of the role and function of the National Coordination Office for NITRD. We support the wide-ranging agenda of the NITRD and also the need to enhance the resources of the NCO to guide the coordination and collaboration efforts of the participating agencies.

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

Recommendations in the PCAST report "Leadership Under Challenge: Information Technology R&D in a Competitive World," are of particular interest to the higher education community and are directly relevant to this request for comment.

These are the PCAST recommendations which we call to your attention:

Recommendation #4 – The OSTP Director should call on senior officials from Federal agencies with large academic networking and information technology R&D budgets to meet with senior officials from the Nation's major research universities to address how to better conduct large-scale , long-term , multidisciplinary academic research in the development and application of networking and information technology important to the Nation. (The NITRD NCO should be directed to support this effort.)

Recommendation #1 – To provide a solid basis for subsequent action, the NITRD Subcommittee should charge the NITRD NCO to commission one or more fast-track studies on the current state of and future requirements for networking and information technology undergraduate and graduate education.

Recommendation # 7 – The NITRD subcommittee should facilitate efforts by leaders from academia, industry and government to identify the critical issues in software design and development and help guide the NITRD planning on software R&D.

Recommendation #10 – The NITRD subcommittee should develop, implement and maintain strategic plan for Federal investments in HEC R&D, infrastructure, applications and education and training. Based on the strategic plan, the NITRD Subcommittee should involve experts from academia and industry to develop and maintain a HEC R&D roadmap.

Recommendation #16 – The NITRD Subcommittee, with support from the NITRD NCO should develop a set of metrics and other indicators of progress for the NITRD program and use them to assess the NITRD program progress.

We appreciate this opportunity to comment and stand ready to participate, providing advice and comment, in continuing discussions regarding the evolving five-year NITRD strategic plan.

Trustworthy Cyber-Physical Systems, Integrated Systems, and Networked Software

This paper presents some of the key areas of our interest that are desired networking and information technology (NIT) capabilities for critical aviation infrastructures as well as US defense and intelligence.

Desired Future NIT Capabilities and Strategic Goals:
Air transportation around the world is evolving to overcome major challenges such as limited capacity, increased air traffic, and environmental pollution. Within the next two decades, advanced NIT enabled applications and processes will be introduced for enhanced aircraft operation, control and maintenance. Consequently, the next-generation airplanes and air traffic management are modeled as safety- and security-critical *cyber-physical systems* on which human lives and well-being depend. The emerging wireless, ad hoc, and sensor networks for aviation, US Air Force, and border control are also cyber-physical systems that are trusted to provide information that can be used in real-time, reliable, safe and useful decisions. Furthermore, recent NIT developments in aviation include *integration of systems* across diverse domains to facilitate network applications, and *open source software* based solutions for reduced manufacturing and maintenance costs of networked systems.

Therefore, to ensure safety, security, reliability, efficiency, usability and yield of future air transportation as well as battlefields, the strategic goal is to support, progress fundamental research and collaboration in the trustworthy design, development, verification and validation of cyber-physical systems, integrated systems, as well as networked software.

Key Challenges and Research Priorities:
The following will require coordination between two or more agencies in the NITRD.

- Framework to formalize the relationship between security and safety.
    - Integration of the mainly discrete methods of security analysis into the quantitative, probabilistic approaches of safety analysis.
    - Combining security analysis which refers to non-functional properties, with the functional software correctness analysis to achieve an overall system safety level.
- Assessment of security technologies in the context of safety certification.
    - Evaluation of avionics software/systems for encryption and authentication.
- Long-term security mechanisms for airplane information assets.
    - Protection of information assets throughout airplane lifecycle.
- Effective formal methods.
    - Visual representations of FM with transparent analysis to facilitate the communication of FM benefits to business management.
    - Specification language that is accessible to software architects/developers without substantial training, and easy for customers to understand so they can contribute to the formal specification.
- High assurance of networked system-of-systems.
    - Interoperable domain standards and policies for NIT and security.
    - Design of scalable pervasive public key infrastructure for establishing trust in large-scale, multi-stakeholder aviation applications such as airplane software distribution.
    - Inexpensive, efficient, scalable, user-friendly methodologies for end-to-end assurance assessment.
    - Security models for multi-disciplinary global collaboration.
- Security of integrated modular systems

    - Evaluation of potential impact of loadable software at different safety levels residing on the same platform
- Security assessment tools for cyber-physical systems.
    - Evaluation of trustworthiness of information received from the physical world.
    - Evaluation of the impact of attacks on the security and performance of networks.
- Evaluation of potential attacks on next-generation networked aviation
    - Security assessment of communications, navigation and surveillance technologies, such as Asynchronous Dependent Surveillance, before use for US air traffic management.
    - Securing operation and control of unmanned aerial systems for civilian applications
    - Secure design of onboard wireless networked control systems
- Early demonstration or evaluation of emergent security technologies.
    - Quantum cryptography.
- Human-computer interaction for airplane operators.
    - Introduction of onboard networks and security technologies will warrant data representation and network monitoring tools to ease the cognitive load of pilots, aircraft maintenance and traffic control personnel. Due to the safety-critical aspects of such software-based tools and the global operation of airlines, high confidence design and assessment is needed.
- Wireless-enabled environmental monitoring and control.
    - Wireless sensors deployed on the aircraft can provide feedback on pollution-related factors of legacy, current and future airplanes. However, key enablers for this include the design of efficient and non-interfering wireless sensor network architecture and the mitigation of security concerns.
- Trustworthiness and privacy of ubiquitous computing in aviation.
    - Addressing the lack of centralized authorities.
    - Assessment of user privacy concerns, such as traceability of RFID systems.
- Security of open source software.
    - Effective assessment methodologies for real-world open source software practices.
    - Metrics for determining assurance levels of open source products.

Opportunities:

- Common solutions for commercial and defense platforms. This approach has many benefits:
    - competitive advantages for airframe manufacturers from technology reuse.
    - employment of foreign nationals in the US for developing commercial platforms, derivatives of which are then offered to the US military.
- Common solutions for cyber-physical systems in aviation and other transportation sectors.
    - Recent advances in vehicular networks can benefit the aviation industry with the networking and security of e-enabled airplanes.

Achieving Desired NIT Capabilities:

NITRD should continue to support forums that allows industries to be forward looking in terms of identifying important technology areas, needs and challenges, and collaborate with universities to bring focused research and expertise in these areas. However, in aviation, diverse organizations have diverse interests and spans of control, sometimes overlapping and sometimes mutually contradictory. Therefore, the desired end-state must be a coherent set of regulatory laws, practices, processes that work together to satisfy needs of the public weal, requirements for compliance, and business interests of manufacturers, maintainers, and operators of transportation infrastructure and products.

The US High Energy Physics activities are inherently international and inter agency.

Many of these activities require a worldwide-distributed computing system and dedicated data transfers. The largest of these activities currently involve approximately 100 sites. These sites include NSF and DOE funded US University groups and DOE facilities. These activities also include NASA and non-US sites experimental facilities, for example the LHC in Switzerland and Daya Bay in the PRC.

US scientists participate in these activities in the context of international collaborations, with equal data access for all collaboration members/ therefore, the distributed infrastructures used in the US must interoperate with those created by other significant international partners.

The US supports a program of research work allowing the construction of these kinds of systems. The program has included both networks and grid technology. Projects such as these have advanced the state of the art for constructing usable, large-scale federated systems.

1) Given future demand, the level of expertise required to form, build and operate these sorts of systems needs to be lowered, continued work is warranted, both on technology and operational methods.

2) The infrastructure developed by differing major international partner's needs to interoperate. Examples can be drawn from the Open Science Grid, which functions as part of the World Wide LHC Computing Grid; and Dice, which provides a forum for coordinating advanced network technology in the US, Canada and Europe.

3) Research on the business and social aspects of these kinds of federation is needed, for example, coordination of operations, and the level of organization of Scientific Virtual Organizations, for example Voss.

4) Operational frameworks and operational work are needed to provide practical, reliable inter-operating systems that crosscut across specialties. Informative examples include grid operating Organizations; efforts to connect campus networks To high speed wide area infrastructures; and resolve connectivity problems at end sites in an efficiently way

5) There is a need to make ensure that community written software is increasingly well crafted and well maintained. Informative existing work is the VDT work within the Open Science Grid. This area requires more development and expanded work. Similar efforts need to extend to the conception and generation of system notions and software.

August 25, 2008

RE: Submissions to Inform the Five-Year Strategic Plan for NITRD Program

This responds to the July 25, 2008 Federal Register notice (73 F.R. 43477) inviting submissions in support of the strategic plan of the Federal Networking and Information Technology Research and Development Program (NITRD). Thank you very much for providing the opportunity to comment on this critically important program.

I write on behalf of the more than 350 member companies of the Information Technology Association of America. The Information Technology Association of America (ITAA) is the premier IT and electronics industry association working to maintain America's role as the world's innovation headquarters. Following its April 1, 2008 merger with the Government Electronics and Information Technology Association (GEIA), ITAA provides leadership in market research, standards development, business development, networking and public policy advocacy to some 350 corporate members doing business in the public and commercial sector markets. These members range from the smallest start-ups to industry leaders offering Internet, software, services and hardware solutions. ITAA offers the industry's only grassroots-to-global network, carrying the voice of IT to companies, markets and governments at the local, state, national and international levels to facilitate growth and advocacy. The Association maintains a formal alliance with more than 40 regional groups in the U.S. and Canada, representing 16,000 technology-related companies through the Technology Councils of North America. ITAA is also the U.S. member of the World Information Technology and Services Alliance, a network of nearly 70 industry associations from around the world. For more information, visit www.itaa.org.

Today, use of information technology – hardware, software, applications, networking, services – is becoming pervasive in all sectors of the economy, and by every conceivable actor: organizations (internally and externally), whether business or government or any other imaginable organization; individuals, whether as students, consumers, citizens, educators, business owners and employees, or farmers and fishermen. IT is having a profound effect on the productivity and growth of every economic sector in which it is used. At base, IT is not just another albeit critically important industry among those listed in the ranks of farming, manufacturing, services or other enterprises. IT increases the productivity of every industry it touches. It enables innovation in every economic sector. It allows companies to improve efficiencies and lower the cost of doing business, which in turn frees up capital for other investments and lowers the prices of consumer goods. IT enables the creation of altogether new technologies and branches of science, such as nanotech and growing, detailed knowledge of the human genome. There is a deep integration of IT into other engineering systems in today's world. Information Technology also improves the lives of every citizen who can access it by facilitating communications and commerce, increasing the transparency of government and improving access to government services. At the same time, developing a stronger IT workforce for the challenges ahead has never been more important and is starting to become a real problem unless serious steps are taken as soon as possible.

All of this needs to be reflected in the strategic plans of the NITRD. ITAA is a strong supporter of this program and looks forward to continuing to engage in conversations with the NCO about its strategic plans for the future. For our introductory comments, we rely on last year's PCAST report, "Leadership Under Challenge, Information Technology R&D in a Competitive World", http://www.ostp.gov/pdf/nitrd_review.pdf ("PCAST Report") We generally endorse its findings, and are pleased to see that many of the priorities for research listed in that report (see generally Chapter 4) appear to be reflected in the list of priorities for research on the NITRD web site, www.nitrd.gov Of particular interest to our organization is the discussion of R&D on the subject of Cyber Security and Information Assurance. This must be a critical national priority, as a matter of national security and to assure that there is adequate consumer and enterprise assurance and trust in the medium of the Internet to continue to allow it to continue to grow and develop.

We also strongly endorse the PCAST recommendation that "Federal agencies should rebalance their networking and information technology R&D funding portfolios by increasing: (1) support for important networking and information technology problems that require larger-scale, longer-term, multidisciplinary R&D and using existing or new mechanisms to address those problems and (2) emphasis on innovative and therefore higher-risk but potentially higher-payoff explorations." (PCAST Report at 26). As the Report notes, this is arguably the "most critical need" of the NITRD program, and has many supplemental benefits including providing a mechanism for increasing the number of researchers and students in the field, by making R&D more relevant to recognizable problems and issues in the modern economy, or as the report puts it, may attract students and researchers to work on "more challenging problems relevant to the future." (PCAST Report at 27). Given the critical nature of the IT sector, a large scale, longer term, multidisciplinary approach to the basic federal program for funding R&D in this sector is equally critical.

Thank you again for the opportunity to comment. We look forward to continuing the conversation begun by this submission. If we can provide additional information, please feel free to contact me at
pbond@itaa.org

Vision for Interagency Network R&D Needs Over the Next Five Years

Background and Context

Over last five years, it has become clear that researchers and educators are demanding two research tools from research internetworks: virtual lans (VLANS) and optical switching and the tools to make connections easy to manage, end-to-end. However, the value of these internetworks is only as good as their interconnections – both within the US and internationally.

In order to support the research platforms of the US government, with its various open science and international research use of high performance networks, it is apparent that serious basic and applied research needs to be directed to the network of the future. Terabits per second networks will be essential components of distributed Petascale science and advanced analysis. These terabit networks will likely be based on advanced optical transport infrastructure, ultra high-speed protocols and dynamically reconfigurable services. Drafting a concrete and robust roadmap to develop, test and implement these advanced capabilities will be critical to agency missions in the next decade.

Key Federal requirements for Petascale analysis and science will emerge years before they are required by the commercial communities. Such requirements can be met only through agency network R&D directly targeted at addressing key needs for internetwork infrastructure and services. Agencies cannot compromise mission-critical programs by assuming commercial communities will address their needs in a timely manner. It is clear that no other R&D community is positioned to address these needs within the required time frame and only an organization like NLR can readily step up to the developing needs for infrastructure that can and will support the required research efforts.

Networking R&D Requirements for Agency Petascale Science and Analysis

The federal government has a broad spectrum of network requirements ranging from routine production IP services to very specialized high-throughput capabilities that are required to support emerging distributed Petascale and Exascale science and analysis applications. With DOD, DOE, NASA, and NOAA all moving towards Petascale and Exascale scientific discovery, it becomes imperative to interconnect these distributed facilities for the advancement of science. In terms of end-to-end networking, Petascale and Exascale applications will require both *capacities* and *capabilities* unprecedented in currently envisaged network infrastructures and associated support technologies*:*

Network Capacity utilizing Optics: To support Petabyte and Exabyte-scale data distribution and other applications, Terabit-capable networks will be needed. This implies protocols and services that can efficiently operate at ultra high speeds. The agencies' supercomputing, storage, visualization and experimental facilities will be required to have sufficient capacity to handle Petabyte or larger data sets. The one hundred Gbps/lambda circuit technologies, currently emerging from development laboratories, barely meet this requirement. It still takes 24 hours to transfer a Petabyte of data at such rates. Nonetheless, this serves as a starting point in planning for Petascale- and Exascale analysis, since current experience shows that conducting large-scale data transfers using soft aggregation of circuits or data streams is difficult and support-intensive. While planning 100-1000 Gbps capacity requirements for agency cores, it is critical that the needed network capacity be provisioned end-to-end: including metro, campus, edge and host. Technologies capable of providing connections with such capacities are necessarily disruptive and therefore require theoretical and experimental research. Hence, capacity solutions based on such technologies must be fostered and developed through highly focused efforts on experimental networks and systems, since the advanced technologies may prove to be too disruptive for production network infrastructures.

End-to-End Capabilities*:* In addition to the transport network's path capacities, an extremely important and vital part of the solution consists of optimized systems of software and edge/host technologies that will enable users to achieve throughputs commensurate with the provisioned capacities. As we already see today, research networks will include multiple vendors' hardware and software, multiple protocols and encompass multiple service-providers.

The challenges of developing the needed capacities and capabilities for Petascale and Exascale applications are multi-fold, spanning the wide-area connections, edge and hosts systems, systems and application software and middleware tools. The development is very complex and often requires non-traditional solutions to achieve the needed quantum leaps in the capabilities. Examples might include special interconnects to supercomputers and direct wide-area InfiniBand interfaces to storage systems. Such developments would require specific combinations of specialized technologies in this area and would be extremely unlikely to become available as incidental byproducts of other projects.

Research Recommendations

We see the need for R&D in five core areas. We believe a strong, coordinated interagency approach will be needed to deliver robust working solutions in all five areas in a five-year time frame

- Transport protocols, technologies and data distribution services to enable networks to move the large datasets required
- End-to-end federated network measurement based on Petascale-derived performance characteristics
- Multi-layer federated network provisioning to enable regional, national and international network use on an end-to-end basis
- High-performance end-system middleware that augments a multiprotocol, multiprovider infrastructure
- Experimental network testbeds to support the breakable and configurable research of network scientists

Networking and Information Technology Research & Development: Input for Strategic Planning

Futurists, product planners, and government agencies envision new semi-autonomous vehicle, vehicle-to-vehicle, and vehicle-to-infrastructure capabilities that reduce fuel consumption, reduce congestion, improve air quality, and improve safety. Given this future, the automotive industry must have access to new engineering tools and methods that yield the productive development of these complex Cyber-Physical Systems at unquestioned levels of quality.

Our goal is to partner with NITRD and other collaborators to conduct research in robust methods for Cyber-Physical Systems development. Our particular experience is in automotive engine control and the following paragraphs expose critical aspects Cyber-Physical Systems development in the engine control context. Of course, the complexity and difficulty grow exponentially for the products and systems outlined above.

To reduce fuel consumption and emissions, improve vehicle performance, and meet higher customer expectations, the powertrain control system is becoming more and more complex. As a result of the increased complexity, the software has a structure composed of thousands of variables and tuning parameters with highly interacting functions. The software is becoming very difficult to develop and test and the cost for securing the highest quality is increasing exponentially. **Model-Based Development** (MBD) has been promoted [1], [2], [3] for several years to alleviate this scenario.

We envision an engineering framework to enable the introduction of MBD to production processes. Challenges related to three aspects of MBD are introduced in this position paper:

1) **Plant Modeling.**

2) **Verification and Validation (V&V)**.

3) **Visualization.**

Challenges and Research Needs

Plant Modeling: If we are to be successful, we must recognize that plant models are prerequisite for model-based development. These plant models must a) capture the "appropriate" physical phenomena and dynamics, b) be "sufficiently" accurate, and c) be "available" when dictated by the production process schedule. Given these requirements, we believe there is significant research and development opportunity to improve plant modeling processes, methodologies, and tools so that plant modeling can be considered a predictable and dependable production process.

In fact, toward this end, we are participating in the formation of a Plant Modeling Consortium. Our initial goal is to create a community to promote plant modeling research and development in the above context. Potential topics for consortium consideration include: a) plant modeling process definition and standardization, b) domain-neutral, conservative modeling methods and tools, c) statistical and machine learning modeling methods and tools, d) model simplification methods and tools, e) improved parameter identification methods and tools, f) integration of physical and statistical modeling methods, g) improved solvers for a-causal model formulations, and h) model confirmation & validation methods.

From our control domain perspective, we consider the behavioral / functional specification

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

and verification of the embedded powertrain software to be a fundamental challenge. It is essential that we develop robust modeling methods that yield ‘sufficient’ plant models to support this control algorithm specification and verification.

Verification and Validation (V&V): Conventional V&V processes of powertrain control software rely on exhaustive experimental testing. However, continuously increasing complexity of modern powertrain control systems poses great challenges to the conventional processes. We list a few of the challenges in the MBD context:

Defining levels of abstractions and corresponding specification languages are needed for being able to develop more complex control systems and efficiently perform V&V [5]. For example, a Simulink model visually provides hierarchies, however it has all the details of a C-code. A higher level architecture abstraction that captures properties such as component interactions, real-time communication schedules and dependency structure are needed.

Ensuring semantics-preservation between architecture abstraction layers during the development process is a key challenge to develop high confidence control software. As we move down the abstraction layers, there are more details added to the design. The less abstract implementation must preserve the semantic properties of the high level model. If it is not preserved as in between simulation and experiment, the gap and its effect on robustness need to be analyzed.

There are many V&V methods, tools and platforms available or being developed. V&V processes need to be developed in every step of the control software development process including requirements management, control logic design, software implementation and integration and calibration. The process should efficiently use a comprehensive end-to-end V&V tool chain with confidence.

Visualization: It is rare to develop powertrain control software from scratch. This means that we have to cope with legacy software which has been incrementally expanded over time to add new control functions. Drastic changes are avoided as far as possible, since reliability of legacy software has been proven through market use. Thus legacy powertrain software tends not to be well structured as it increases with complexity and size. It is difficult to understand how the system works and how partial changes will affect the whole system. As control software keep growing, this is becoming one of the key issues in the design process and demand for visualization is increasing in this light.

Conventionally, understanding legacy control software starts with code reading to capture software structure, and then we infer system behavior. To help structural understanding, there already exist commercially available visualization tools. The difficulty we have in using these tools is that the graph representations given by these tools are still too complex and their interpretation is as painful as code reading. A mechanism for selecting what to show and what not to show is required.

Simulation is often used to understand dynamic behavior of the software. However, stimuli should be provided to excite important system behavior and this requires deep knowledge on the software and the system itself which we don’t have without large investment.

Understanding control system from legacy code is time-consuming and human-dependent activity. Methodologies to capture and extract representative system properties while hiding unimportant details seem to be the key. We believe that tools that can support this activity will

The opinions and positions in the white papers and comments posted on this web site are those of the submitters only and do not necessarily represent those of the Federal government, the NITRD program and its participating agencies, or the National Coordination Office.

improve the productivity of development processes significantly.

References

[1] T. Ueda, A. Ohata, "Trends of Future Powertrain Development and the Evolution of Powertrain Control Systems," in Proc. 2004 SAE–Convergence Conference, October, 2004.

[2] A. Ohata, K. Butts, *"Towards a Concurrent Engine System Design Methodology,"* in Proc. 2005 American Control Conference, June, 2005.

[3] A. Ohata, *"Model-Based Development: Plant Modeling Environment,"* Vision of the Future Plenary Presentation, Fifth IFAC Symposium on Advances in Automotive Control, August, 2007.

[4] Plant Modeling Consortium Website: http://www.maplesoft.com/consortium/

[5] http://aadl.info/

Cover Note

This document was developed and used as an internal, working guide for Federal NITRD program participants. As a working guide, the document does not represent official Administration policy. This text is being released now (July, 2008) to provide background information for those interested in contributing to the development of a new, NITRD program strategic plan. It should be used for historical and informational purposes and should not be considered a model for a new strategic plan.

National Coordination Office

INFORMATION TECHNOLOGY RESEARCH AND DEVELOPMENT (IT R&D) PROGRAMS

FIVE-YEAR STRATEGIC PLAN FOR FY 2002-FY 2006

Part I: Overview

Today, at the beginning of a new century and a new millennium, information technology (IT) is transforming our world, generating unprecedented U.S. prosperity and building revolutionary new infrastructures for commerce, communication, human development, and national security. In this remarkable period of transformation, the United States stands preeminent as the world's information technology pioneer, its research leader, and its foremost developer and deployer of cutting-edge computing, high-speed telecommunications, and IT systems.

"When historians look back a decade or so hence," Federal Reserve Board Chairman Alan Greenspan said in a March 6, 2000, address on the New Economy, "I suspect they will conclude we are now living through a pivotal period in American economic history. ... It is the growing use of information technology throughout the economy that makes the current period unique."

The New Economy arose from the Nation's immense industrial and entrepreneurial enterprise. But it was Federal investment in fundamental, long-term IT research and development (R&D) that launched the digital revolution, and that investment continues to play a critical role in generating the technological breakthroughs the country needs to meet vital national objectives and achieve the full promise of information technology in such public benefits as:

- Immediate on-site medical care, in the home and at remote locations
- Reliable, failure-resistant systems for such mission-critical applications as air-traffic control, defense, financial transactions, life support, and power supply
- Reduction of battlefield risk for military personnel

- Industrial process and product modeling, visualization, and analytical capabilities, such as in aircraft design and production, automotive efficiency and safety, and molecular synthesis of new drugs
- Expanded e-commerce with assured security and privacy of information
- On-demand universal access to education and knowledge resources
- Advanced computing capabilities that underpin the Nation's leadership in science and technology and the success of critical civilian and national security missions of the Federal government
- More accurate weather forecasting and improved environmental analysis
- High performance networking and information systems for emergency and disaster management
- Access to information anytime, anywhere, with any device

As the President's Information Technology Advisory Committee (PITAC) noted in its 1999 report on the status of U.S. information technology research, bipartisan support of Federal investment in fundamental IT R&D over the last several decades has produced "spectacular" returns for the economy and for society generally. But the Committee, made up of prominent IT industry leaders and university researchers, warned that Federal funding for IT R&D is now seriously inadequate, far outpaced by the explosive growth in societal needs for advanced IT capabilities. Measured in constant dollars, the Committee found, Federal research investment in critical IT areas – such as networking, software, and high performance computing – had been flat to declining for a decade, at the same time that IT-related business activity had grown to represent virtually a third of GDP growth. The PITAC concluded that the existing level of Federal IT R&D investment represented a significant strategic threat to the Nation's economic future.

The U.S. House of Representatives reached a similar conclusion in its findings on the Networking and Information Technology Research and Development Act (H.R. 2086) passed in 2000, saying: "Current Federal programs and support for fundamental research in information technology are inadequate if we are to maintain the Nation's global leadership in information technology."

There is also a growing nationwide need for trained IT scientists, engineers, educators, and technical workers. "Fundamental research in information technology has contributed to the creation of new industries and new, high-paying jobs," argued the House bill. "Scientific and engineering research and the availability of a skilled workforce are critical to continued economic growth driven by information technology."

The national stakes for the U.S. are high. In the past, the benefits of any single area of scientific research might be limited in scope – enabling, for example, development of one weapon or treatment for one disease. But information technology is by its nature pervasive, providing systems, tools, and capabilities that daily touch hundreds of millions of citizens. A balanced, diversified Federal IT R&D portfolio not only advances vital Federal missions but helps the Government support overarching public goals in education, environmental protection, health care, law enforcement, productivity, transportation safety, and many other dimensions of our national life.

The National Academy of Sciences, in a recent study of the Nation's IT research needs, noted that amid the deluge of new IT devices, "the critical role of the first half of the R&D process is overlooked: the research that uncovers underlying principles, fundamental knowledge, and key concepts that fuel development of numerous products, processes, and services." The Council concluded: "The Federal government should boost funding levels for fundamental IT research, commensurate with the growing scope of research challenges."

Corporate IT leaders strongly support a significant expansion of the Federal IT R&D enterprise on the ground that such fundamental, long-term work is critical to the national interest but will not be carried out by industry. Citing contemporary pressures on industry to focus on short-term product-related research, they argue that Federal IT R&D fills a gaping void in the Nation's IT research portfolio, complementing private-sector efforts with a focus on pre-competitive, long-term studies in such critical IT areas as advanced scientific computing, large-scale high-speed networks, and fundamental enabling information technologies. These areas, not addressed by industry, support essential government functions and effectively drive overall national innovation.

This Strategic Plan represents the collaborative research framework of the 12 Federal agencies whose critical missions require advanced IT R&D. These agencies participate in the Federal government's interagency Information Technology Research and Development Programs – successor to the High Performance Computing and Communications Program established by Congress in 1991. The Federal IT R&D agencies have established a 10-year record of highly successful collaborative accomplishments in multiagency projects and in partnerships with industry and academic researchers. The multiagency approach leverages the expertise and perspectives of scientists and technology users from many agencies who are working on a broad range of IT research questions across the spectrum of human uses of information technology.

Pursuing the Strategic Plan will enable the Federal IT R&D agencies to address the most significant scientific and technical challenges standing between today's networking and computing capabilities and the affordable advanced technologies and tools that both the Federal government and the Nation need. As Chairman Greenspan contended in his speech on IT, "We should ... persevere in policies that enlarge the scope for competition and innovation and thereby foster greater opportunities for everyone."

Key national IT research priorities proposed under this Strategic Plan are summarized in the table on the following page. The sections that follow describe the imperatives for Federal leadership in fundamental IT R&D; the technical elements of the Strategic Plan's five-year research agenda; examples of cutting-edge National Grand Challenge Applications that the IT R&D agencies propose to pilot, test, and/or demonstrate as part of the Strategic Plan; and the structure of the multiagency IT R&D Programs' coordinated administration and management.

Research Priorities in Federal IT R&D Programs' Five-Year Strategic Plan

Research Area	Research Goals	Technical Agenda
Compact, teraops-scale supercomputing systems	• Platforms capable of a trillion or more operations per second – with large global memory bandwidth, high-speed interconnects/switching, low latency • Long-range scalability to petaflops computational speeds and 1,000 petabyte level storage	• Systems software and tools for terascale and beyond-terascale platforms • Systems architectures, including configuration, new device and storage technologies, software for managing highly parallel computations, and hierarchical programming environments • Advanced nonclassical computing concepts
New processor technologies	• Increase processor speeds to 1,000 or more times current levels • Quantum and DNA computers for embedded systems, rapid design of target and therapeutics for pathogen combatants	• Innovative computational structures, 3-D architectures, hybrid technologies • Computation at the molecular scale, microprocessor fabrics, quantum structures, biological substrate computing
Reliable, universally accessible high-speed networks	• Networks 1,000 or more times as fast as today's Internet • Secure transmission, with guaranteed quality of service, of HDTV, sound, and terabytes of data • Ubiquitous access from wireless, embedded, and wired devices • New classes of applications • Testbeds and supporting infrastructure	• Fundamental network research (optical and wireless networks, resource and network management, increased bandwidth requirements) • Technologies for scalability and ubiquitous access • Technologies for distributed data-intensive computing, collaboration, and computational steering, distance visualization and instrumentation, workflow management, management of large-scale distributed systems
High-confidence software and systems	• Fail-safe, secure, and reliable software and systems for mission-critical applications • Technologies and tools for rigorous testing, validation, and certification	• Theoretical, scientific, and technological principles for high-confidence design • Encryption, secure transmission, user I.D. systems • Formal methods for specification and validation • Technologies and tools to reduce the time and cost of assuring quality and security
New paradigms for software design	• Software governed by scientific and engineering principles • Autonomous, self-programming software for embedded systems	• Computer languages, compilers • Interoperability of applications • Tools for integrated software and system design • Reusable middleware for embedded systems • Automation of software development • Empirical testing of models and methods
Large-scale information systems and data sets	• Next-generation technologies for management and use of massive data sets and information archives • Interoperability, accessibility, usability of data sets and data management systems	• Tools for collection, synthesis, curation, indexing, mapping, provisioning, and fusion • Protocols for data compatibility, conversion, interoperability, interpretation • Methods of preservation, archiving • Interoperable interfaces, metadata systems • Ultra-scale storage, management, and data-mining technologies
Human-computer interaction	• Improved integration of humans and computers in complex task environments • Augmented human capabilities • Universal accessibility to digital systems	• Technologies for language translation, speech-based interfaces, content extraction, multimodal systems • Intelligent systems, such as "smart spaces," for ubiquitous computing with multiple interactions • Universal designs integrating device-independence, usability, and accessibility considerations • Ultra-rapid machine translation, prosodic processing, auditory modeling
Social, economic, and workforce implications of IT	• New knowledge about the co-evolution of IT and society, economic systems, education, the workforce, and workforce development • Innovative applications of IT in education and training • Infrastructure for SEW research	• Social and technical systems • Internet governance and "citizenship" • Intellectual property and information privacy • Collaboration and learning • Workforce development • Universal participation

Part II: Why we need IT research

The promise of information technology

No longer just a provider of tools for the sciences and engineering, information technology today is the uniquely interdisciplinary field at the very core of American innovation in every sector.

IT begins with fundamental research in the sciences and in engineering and stretches across the applied scientific and engineering knowledge it takes to design, construct, and maintain computing and telecommunications equipment. IT encompasses the mathematics and computer science expertise that goes into writing the complex sets of instructions – the software – that enable digital devices to do what people want them to do. IT also engages the thinking and imagination of scholars, students, government and business officials, and ordinary computer users in virtually every field who help figure out how to harness computing and communications capabilities to human needs, interests, and aspirations. All these scientific and technical skills and knowledge bases working together produce the complex digital systems that we have quickly come to rely on in our day-to-day lives.

Whether we are aware of it or not, we are surrounded by the results of this multidisciplinary R&D activity, in such applications as precision instrumentation and visualization capabilities for medical diagnosis and treatment; inventory-management systems for agile, just-in-time manufacturing; the Mars rover and astronomical images from the far reaches of the universe; monitoring and management of large-scale financial systems; standardized transmission protocols for electronic mail and audio, video, and sound files; international air-traffic communication and control systems; and weather forecasts based on collection and analysis of data from real-time observations of wind, water, and other environmental systems.

IT's rich mix of basic and applied science was showcased in October 2000, when Nobel Prizes in physics and chemistry were awarded for the first time in the program's history to six researchers, including four Americans, whose discoveries – the integrated circuit, the hetero semiconductor structure, and conductivity in plastics – involved information technology. The prizes also highlighted a key underpinning of U.S. scientific achievements that is not well understood. Three of the four American prizewinners are university-based scholars who received early and continuing research support from the Federal government.

The Federal role in the IT revolution

The Nation's computer and telecommunications industry leaders are the first to acknowledge that Federal investments in fundamental IT R&D produced both the knowledge base and the technical workforce that are powering the New Economy and U.S. leadership in advanced computing and networking. As the corporate leaders on the President's Information Technology Advisory Committee (PITAC) put it in the Committee's 1999 report:

> "Federal funding has seeded high-risk IT research and yielded an impressive list of billion-dollar industries. Federally funded university research has trained most of our leading IT researchers. Information technology industries provide hundreds of thousands of jobs and much of the Nation's recent economic growth. The Federal investment to date has had tremendous benefits for our government, our Nation, and our economy."

Following is a sampling of the Federally sponsored IT R&D that has fueled the Information Age and dynamic business opportunities throughout the private sector:

- **The first operational, electronic stored-program computer (SEAC)** was developed for the U.S. Air Force by the predecessor of NIST; a similar machine, SWAC, built by the agency the same year for the U.S. Navy, was the fastest in the world at the time.
- **The Internet** grew out of ARPANET, the network invented by DARPA-funded researchers in the 1960's.

- **The first graphical Web "browser"** was developed by university-based researchers supported by NSF**;** Web search engines grew out of initial research investments by DARPA and NSF.
- **Java,** the programming language that supports interoperability across networks, is based on concepts first explored by Federally funded researchers.
- **The mouse and the graphical user interface (GUI),** now standard to desktop computers, stem from DARPA-funded research in the late 1960's.
- **High-speed optical networks**. The Federal government's Next Generation Internet (NGI) Initiative has produced the world's first prototype optical networks with end-to-end transmission speeds and carrying capacity a thousand times those of the current Internet. With $276 million in funding over three years, this program has stimulated development of new private-sector companies with a combined value of more than $50 billion – a near-2,000 percent return on the Federal investment.
- **The world's first and largest public medical database,** integrating research findings and medical-journal citations, was developed and is managed by NIH's National Library of Medicine.
- **Speech and spoken dialogue technologies** funded over decades by DARPA have led to new customer call center concepts and more efficient service for industry worldwide.
- **Parallel computing** concepts explored by Federally supported researchers for two decades laid the groundwork for the development of commercial high-end computing platforms in the late 1980's and 1990's.
- **Relational databases** – the industrial-strength software systems needed to store and manage large quantities of information, such as financial records, census data, and business inventories – were pioneered by university researchers funded by NSF in the 1970's.
- **Reduced instruction set computing (RISC) technology**, the basis for many of today's fastest microprocessors, was advanced by DARPA-funded research in the 1970's and early 1980's.
- **Machine learning research,** sponsored by DOE and NSF, was employed in decoding the human genome and also spawned the data-mining industry.
- **Numerical linear algebra libraries research sponsored by DOE, DARPA, and a number of other Federal agencies has produced high performance libraries of numerical linear algebra software that are used by thousands of researchers worldwide. These libraries have become a critical part of the world's scientific computing infrastructure.**

Why Federal investment is key

Federal IT R&D supports critical agency missions and national needs

All of the projects cited above were funded in support of critical Federal missions, including national defense and national security, critical infrastructure protection, energy systems, aerospace engineering, weather and climate forecasting, and advanced biomedical and other scientific research. National defense and national security needs alone require advanced IT research efforts on a continuing basis to equip the military with cutting-edge weapons technologies and secure communications systems and to accurately model and design these advanced systems. Federal research responds to a basic reality of the interdisciplinary IT field: What can be accomplished using IT is determined by the weakest or the slowest technology, not by the strongest or the fastest. For that reason, Federal IT R&D pursues a balanced, diversified portfolio of research interests, reflecting the wide range of enabling technologies required for agency missions.

The talent "pipeline" we need to continue making technological advances is in jeopardy

The irony of the Nation's remarkable IT growth is that the commoditization of technologies in widely available, low-cost products has had a perverse effect on the fundamental-research engine of innovation, hollowing out its substance and thinning the ranks of the basic researchers. The flight of IT researchers into more lucrative jobs in industry – an acute and worsening problem today in both universities and the Federal government – combined with the private sector focus on product development, results in a strategic threat to the Nation's continuing leadership in advanced computing and communications. Judith Estrin, CEO of Packet Design and former CTO of Cisco Systems, Inc., made this point in a recent address, saying: "As a result, there is an architectural vacuum. Who will do the longer-term thinking for our industry?"

The Federal government plays a special role in our society as the primary supporter of research in many fields that generate innovations critical to the national interest and that expand the pool of highly trained scientists, engineers, and technicians needed to work on national challenges. A high-priority Federal IT research program will provide the opportunity to reverse the "brain drain," reinvigorating and repopulating the IT research community with fresh generations of talented people working on the most profound and challenging science and engineering problems of our era.

Federal IT R&D produces broadly useful technologies and tools that spur innovation across the U.S. economy

The coordination of Federal IT R&D investments across many agencies and private-sector partnerships leverages mission-related research, producing general-purpose, broadly useful, and interoperable technologies, tools, and applications. Federal IT R&D has thus been a powerful engine of technology transfer, the direct result of its focus on widely applicable solutions to basic IT problems and its mechanisms of funding R&D in universities, research institutes, and corporations. The large number of Federally funded breakthroughs subsequently commercialized in the private sector – often by graduates of U.S. research universities whose education was funded through the IT R&D Programs – leverage the Federal investments even further.

No other sector does this fundamental scientific discovery work

Many people incorrectly assume that the most significant IT R&D comes from private industry and that such research is best handled there. While it is true that total private IT R&D spending exceeds that of the Government, the lion's share of private-sector research is actually product-related development. In fact, private sector leaders are among the most ardent champions of Federal investment in long-term, fundamental IT research – precisely because the U.S. government is the only real source of support for that kind of work. IT executives emphasize that today's economic pressures – including international competition, stock market fluctuations, and short product lifecycles – push industry to concentrate resources on short-term, product-focused research rather than high-risk investigations with uncertain payoffs. Companies simply will not take on the responsibility for mid- and long-term fundamental research, particularly when the results are unlikely to produce an immediate proprietary economic benefit to their bottom line but rather one that accrues only to whole (sometimes new) industries.

Yet it is fundamental research that has driven the digital revolution. The Federally funded projects cited above explored core technical problems in IT that had to be solved to advance the capabilities of computers, networks, and information systems generally. These projects were not designed to result in commercial products within six months. They achieved results over years of experimentation and revisions that spread across the research community, enabling many scientists to join in the problem-solving. It is this ongoing foundational research process that has generated scientific, technical, and engineering breakthroughs that benefit us all.

Venture capitalists do not finance fundamental research

Some think IT R&D should be funded by private venture capital investment, which reportedly has risen over the last 10 years from $5 billion to $100 billion annually. But as the corporate representatives on the PITAC noted, venture capital flows only *after* "ideas freely flow from universities and national labs to existing and new companies." The leaders added, "The basic feedstock for these investments has been Government support of basic IT research. … If this feedstock is allowed to deplete, the economic growth engine could slow or disappear." David Morgenthaler, former president of the National Venture Capital Association, underscores the significance of this problem, noting that venture capitalists rarely invest to develop enabling technologies – all those bits and pieces of science and engineering that make the hardware and software innards of IT systems but do not stand to make a fortune by themselves. In addition, venture capital investments are usually made with the expectation of recovering the investment and profit within three years, not the longer timescale of high-risk R&D.

Private-sector R&D bypasses the research most significant to Federal and national needs

Private-sector investment strategies therefore bypass key technology areas that are the most critical to Federal government missions and that help support the continuing superiority of the U.S. IT industry. These areas include high-end computing, mass storage, optical networking, interoperable systems and applications, security, privacy, new generations of embedded and large scale systems, improved processes for developing new software, and effective human uses of IT. In fundamental IT R&D, the research time horizons are much longer and there is no guarantee of the success of any one research path.

The value of the coordinated multiagency approach

The breadth and diversity of Federal missions and activities uniquely equip the Government to lead the high-profile effort proposed in this Strategic Plan to assure U.S. IT leadership and economic competitiveness in the new century. The active interplay among computer science, engineering, mathematics, physical and biological sciences, social sciences, and technology users that IT research requires is rare in the private research community but a characteristic strength in Federal IT R&D activities. The Government's multiagency approach to IT R&D leverages the expertise and perspectives of scientists, engineers, and technology users from many agencies who are working on a broad range of IT research questions across the spectrum of human uses of information technology. Moreover, the research enterprise proceeds through multidisciplinary collaborations among university, government, and private-sector researchers, creating dynamic interactions among parties working on IT problems. Industry leaders concur with research scientists that this is an unparalleled resource for fundamental IT R&D.

The multiagency Federal effort proposed in the Strategic Plan will be coordinated by the Interagency Working Group on IT R&D, which is composed of representatives of the participating IT R&D agencies: AHRQ, DARPA, DOE/NNSA (National Nuclear Security Agency), DOE Office of Science, EPA, NASA, NIH, NIST, NOAA, NSA, NSF, and ODUSD/URI.

Part III: What Federal IT R&D can accomplish

In the following sections on their strategic agenda, the IT R&D agencies describe the highest-priority IT research and development issues that must be addressed now to meet pressing government and national needs.

1. Developing the next generations of computing and data storage technologies

Critical Federal missions and industry needs both call for new scientific and technical paradigms that significantly raise high-end computational speeds, provide adaptable and reconfigurable computing environments, and reduce the size, cost, and power requirements of high performance computing and data storage equipment. For example, the world's fastest computing platform today is DOE's "Option White" system at the Lawrence Livermore National Laboratory. A massively parallel system made up of 512 IBM multiprocessor nodes, it requires 13,000 square feet of floor space and more than 3.2 megawatts of electricity for power, cooling, and mechanical equipment. Option White is capable of 12.3 teraops (trillions of operations per second) in processing speed. But even such a system is not adequate for the massive computational requirements of the most complex scientific problems.

At the same time, the Nation's high-end computing sector – the companies that produce computing platforms that are much faster than the standard desktop computer – is a shrinking fraction of the U.S. marketplace. Business purchasers of high-end machines prefer less costly and physically demanding mid-range machines. As a result, the technical challenges of developing technologies that break through today's upper-end barriers in computing speed, storage capacity, and equipment are left orphaned.

Currently, the Government supports several dozen high-end computing platforms at academic computing centers and national laboratories, along with a number of mid-range machines, that are used by both academic and Federal researchers. But these are not nearly enough to support the high-end research and applications needs of university-based or government scientists. Nor do they offer a viable model for scalability to the processing speeds and storage capacity that future advanced applications will demand. Today's Option White platform, for example, has 160 terabytes of storage space spread over 7,000 disk drives. This amount of storage space represents about six times the contents of the entire Library of Congress, but it is only a small fraction of the scientific data that future research will call for.

Finding cost-effective solutions will require computational science research in disciplines such as physics, chemistry, materials science, and electrical engineering, as well as innovations in computer science and applied mathematics. Next-generation supercomputing architectures, systems software, and middleware must also address interoperability needs of Federal agencies. These technological breakthroughs will also aid U.S. competitiveness.

Under this Strategic Plan, the IT R&D Programs will sponsor research to increase the delivered performance of computing systems by 1,000 times the speeds of today's fastest systems, while reducing cost, energy consumption, and footprint, and to develop interoperable systems software and tools that will:

- Improve sustained application performance, ease of use, manageability, and high-speed network connectivity of teraops-scale systems
- Be scalable (expandable) to petaops-scale systems (petaops systems perform a thousand trillion calculations per second)
- Provide a unifying environment for high end scientific computing

Technical research agenda: Next-generation computing and data storage

- Systems software technologies (including operating systems, programming languages, compilers, memory hierarchies, I/O, and performance tools)
- Systems architectures that integrate device and component technology, systems software, and programming environments (including configuration, hardware, node functionalities, device technologies, software for managing highly parallel computations, and hierarchical programming), and network connectivity
- Effective software component technologies for high performance computing
- Advanced computing concepts (including nonconventional architectures, components, and algorithms)

2. Surmounting the silicon CMOS barrier

The complementary metal oxide semiconductor (CMOS) chip, the two-dimensional miniature electronic map on a silicon wafer that is the standard building block of computing systems, is fast approaching its physical limits. That is, the electronic signals that can be routed through its pathways are finite in quantity and speed. Even the complex technical amassing of chips that produced Option White demonstrates these limitations in the machine's enormous size and power requirements. The IT R&D agencies propose to find new materials and methods to create wholly new designs for processors in computing devices. Both Federal missions and private-sector IT innovation require mid-term incremental improvements in computational speeds and long-term breakthroughs to radically new processor architectures capable of teraops and petaops speeds.

This Strategic Plan will fund research at the theoretical and empirical intersection of biology, information science, and micromechanics [Bio:Info:Micro]. Advances in photonics, micro-electromechanical systems (MEMS), sensors, actuators, opto-electronics, digital, analog, and mixed signal processing, and new fabrication technologies make it possible, for example, to conceive of integrated designs in 3-D on a chip with billions of transistors. This work will focus on design of new, modular hybrid architectures that include fault-tolerance, programmability (including novel approaches such as amorphous computing methods), and security features needed in embedded systems for defense.

A related research area is biological substrate computing, the potential in organic molecules – such as DNA, RNA, and proteins – to provide vast storage and processing capacities. For example, one gram of DNA contains 10^{21} DNA bases, which is equal to 10^{8} terabytes of information storage. Breakthroughs in this area could result in:

- High-volume, content-addressable storage
- Solutions to computationally hard problems that now are not solvable
- Self-assembly of nanostructures using DNA/RNA tiling. The nanostructures in turn could be used for nanoscience such as molectronics (described below)

This Strategic Plan will also support long-range research in quantum physics to explore the potential of atomic matter – such as quanta of light or molecular nuclei – to serve as high-speed processing mechanisms. This area holds great promise as a future means of providing:

- Ultrasecure communications over optical backbone networks
- Orders of magnitude increases in the speed of algorithms such as for searching unsorted databases or factoring large numbers
- Quantum computers that can give detailed and faithful simulations of molecular processes and phenomena in physics

Technical research agenda: Surmounting the CMOS barrier

- Innovative computational structures, 3-D architectures, hybrid technologies
- Reconfigurable systems on a chip, adaptive and polymorphous computing
- Processor in memory (PIM) and other efforts to provide memory performance commensurate with processor performance
- New computational substrates
    - Quantum computing
    - Biological substrate computing
    - "Smart fabric." Using technology for interweaving battery, fiberoptic cable, and metal connectors, scientists can produce fabric that can be embossed with enough processors to provide on-person processing on the order of tens of teraops (the size of today's larger supercomputers)
    - Molectronics: Computation at the molecular scale, which holds the potential of providing extremely fast, high-density processing power for the next generation of strategic computing for the military

3. Building versatile, scalable, secure networks for the 21st century

The Internet is at the heart of the IT revolution, and Federally sponsored networking research plays a pivotal role in generating the technological advances critical to the Net's growth and evolution. Though the focus of Federal work is exploration of technologies and tools to support critical agency missions, the Federal research emphases on long-range needs – such as scalability, level and quality of service, reliability, security, interoperability, and flexible access – not surprisingly turn out to be the core research problems that must be solved to transition the Internet into a secure, reliable, expandable, very-high-speed communications system for the Nation in the new century.

The U.S. urgently needs a new generation of basic enabling technologies to "modernize" the Internet for rapidly growing traffic volumes, expanded e-commerce, and the advanced networked applications that will be possible only when next-generation networks are widely available. The Strategic Plan therefore proposes an ambitious research program to press ahead immediately to:

- Understand how to extend the network infrastructure so that it is available anytime and anywhere, and so that it includes ubiquitous networking that extends the network to millions and potentially billions of new devices and chips embedded in larger devices such as appliances, automobiles, and other elements of our transportation system
- Provide needed network services, such as management, reliability, security, and high-speed transmission rates

Technical research agenda: Building versatile, scalable, secure networks

- Fundamental network research
  - Optical networking (flow, burst, and packet switching; access technology; gigabit per second interfaces; simplified protocol layering)
  - Network dynamics and simulation (automated management, automated resource recovery, network modeling)
  - Fault tolerance and autonomous management
  - Resource management (discovery and brokering, advance reservation, co-scheduling, policy-driven allocation mechanisms)
  - Wireless technologies (technical standards for discovery, co-existence, configuration)
  - Assuring an increasing capability to support bandwidth requirements
  - Enhancing and scaling networks to improve robustness and handling of transient interactions among billions of devices
  - Enhancing and scaling networks (maximizing access from the "edges of the network" such as methods for ubiquitous broadband access, tether-free networking, network security and privacy, network management)
  - Understanding global-scale networks and information infrastructure (end-to-end performance, backbone structures, applications)
- Enabling new classes of applications (distributed data-intensive computing; collaboration; computational steering of scientific simulations; distance visualization; operation of remote instruments; large-scale, distributed systems)
- Testbeds and infrastructure (extend the reach of high-performance networks, improve access technologies)

4. Exploiting advanced IT to sustain U.S. leadership in science and engineering

The Federal investments proposed in this Strategic Plan include development and demonstration of the world's most advanced computational science and IT systems in the sciences and engineering. These high-performance IT capabilities, like all important innovations, will drive new waves of exploration and discovery at the leading edge of scientific and engineering knowledge, where the U.S. must remain in the years ahead. High-end scientific computation and visualization technologies and tools will enable researchers to "see," interact with, and analyze the structures and behaviors of organic and inorganic matter more precisely than previously possible – from the tiniest building blocks of the universe, to the tolerances of manufacturing designs, to the properties and interactions of the biosphere's large-scale phenomena. These exciting explorations will leverage Federal IT R&D investments to attract talented researchers into engineering and science and strengthen our national leadership in these fields.

Examples of the advanced science and engineering activities that will be possible with high performance IT systems include:

- Modeling and simulation in the biological, chemical, environmental, material, and physical sciences, such as:
  - A dynamic integrated model of the Earth's atmosphere, oceans, and soil at scales ranging from kilometers to meters
  - A model of the human body and its components at scales ranging from atoms to molecules, cells, and organs, to the whole body
  - Complete engine simulation, including combustion, chemical mixing, and multiphase flow
  - Simulation for controlled fusion to optimize the design of future nuclear reactors
  - Design of new chemical compounds for biological and manufacturing applications

  - Models of chemical, manufacturing, and assembly plants for optimization
  - Simulations of automobile crash tests for different spatial configurations that reliably substitute for real tests
- Data assimilation, fusion, visualization, manipulation for modeling and simulation
- Modeling and simulation in IT
  - High-end computing systems
  - Network dynamics
  - Large-scale IT systems such as embedded systems and distributed heterogeneous applications
- Collaboration technologies in clinical medicine, scientific research, and professional education and training

5. Ensuring reliable operation of critical systems, with protection against failures and attacks

The 1999 PITAC report correctly argued that fundamental research in software development should be an absolute national priority in Federal IT R&D. The Committee highlighted a reality of the Information Age: The software running today's computer systems and networks is a vast patchwork of often idiosyncratically designed, insecure, and non-interoperable code whose fragility manifests itself daily in unreliability, security breaches, performance lapses, errors, and difficulties in upgrading. Unlike the design of bridges and airplanes, for example, there exists today no framework of formal scientific and engineering principles governing software development. At the same time, the demand for software currently exceeds our capacity to produce it, and the software that is developed is very costly and increasingly complex, with many programs running to millions of lines of code – far too many to be closely validated or made secure from attack with today's technology.

If all that were at stake were the frustrations of home computer users, perhaps we could leave software development as a cottage craft rather than a formal scientific discipline. But with software already managing large-scale and mission-critical systems as aircraft and air traffic, medical devices including life-support systems, electrical power grids, international financial networks, and advanced weaponry, we must act immediately to turn software development into a science-based discipline. We must conduct the research necessary to develop software governed by formal principles and methods and structured so that its security and reliability can be assured through automated testing and validation. Mission-critical systems must be able to withstand hacker, criminal, and enemy attacks as well as unanticipated system interactions, "self-healing" so they can continue to function after an attack or system failure, and designed to guarantee predictably high levels of data integrity and security.

The Strategic Plan proposes to fund research to develop and demonstrate revolutionary high confidence software and systems development and assurance capabilities that balance risk, cost, and effort to achieve systems that behave in predictable and robust ways. The goals of this research effort are to:

- Provide a sound theoretical, scientific, and technological basis for assured construction of safe, secure systems
- Develop hardware, software, and system engineering tools that incorporate ubiquitous, application-based, domain-based, and risk-based assurance
- Reduce the time, effort, and cost of assurance and certification processes
- Provide a technological base of public domain, advanced prototype implementations of high confidence technologies to enable rapid adoption

Technical research agenda: Ensuring reliable operation of critical systems

- Foundations of assurance and composition
  - Rigorous modeling and reasoning about high confidence properties
  - Intereoperable methods and tools

  - System composition and decomposition
  - Specification
  - Safety and security foundations
- Scalable fault prevention, detection, analysis, and recovery
  - Robust system architectures
  - Monitoring, detection, and adaptive response
- Correct-by-construction software technologies
  - Programming languages, tools, and environments
  - Systems software, middleware, and networking, including reusable middleware services such as efficient, predictable, scalable, dependable protocols for timing, consensus, synchronization, and replication for large-scale distributed embedded applications and domain-specific services
- Evidence technologies for verification and validation
- Experimentation and reference implementations
  - Assured reference implementations and assurance cases, such as Public Key Infrastructure (PKI) for advanced networks, software control of physical systems, and mobile networked devices
  - Domain-specific certifications technologies, such as technologies for cost-effective verification and validation and verified hardware/software co-design technologies

6. Making software for the real world

The PITAC contended that the longstanding “crisis” in software quality and productivity threatens U.S. security and economic viability. A conspicuous example of the enormous opportunities and challenges before us is embedded software – that is, software operating with and controlling the physical world. Embedded software is extremely hard to build because its design cannot be based on an idealized model of the real world. While the primary stakeholder is DoD (embedded software is the main reason for significant time and cost overruns in major weapon programs and presents a profound technical challenge for developers), embedded software has tremendous commercial significance. Examples of this may be found in automotive electronics (where it is predicted that the cost of the embedded computers and software will exceed that of the drive train and body by early 2003), consumer electronics such as personal digital assistants (PDA’s,) cell phones, television sets, and other household devices, and industrial process control.

Given the staggering impact of the software industry on both the private sector — where personnel costs have reached $400 billion a year — and the Federal government, the Strategic Plan will sponsor fundamental research that will lead to more cost-efficient, productive software development methods. This will result in higher-quality software with predictable characteristics, as well as support the construction of advanced applications that stress and evaluate current and evolving best practices.

The National Research Council, in its new report "Making IT Better," argues that the single greatest challenge in IT research today is presented by large-scale systems, which now power society's most complex and critical infrastructures but which have not been the IT research community's primary focus to date. Citing the growing complexity, heterogeneity, distribution, and integration of these vast interconnected systems, the report urges that research to improve their design, development, and operation be made a national priority.

In large-scale systems, the validity of theoretical approaches is drastically challenged by scalability pressures and by the inherent heterogeneity of components. We cannot achieve improvements without evaluating the practical applicability of methods and techniques and actually testing them in large-scale application platforms. Therefore, this Strategic Plan proposes an aggressive research program that not only addresses the scientific foundations of software design, but also investigates the related engineering process and conducts substantial experimental evaluations.

Technical research agenda: Making software for the real world

- Science of software and system design
  - Languages and compilers —for example, domain-specific languages to make software specification and development easy for end users and languages that are easier to use and harder to abuse
  - Effective methods for composing software and systems – better techniques for composing, analyzing, and verifying complex systems, and making them interoperable on widely distributed heterogeneous systems
  - Foundations for advanced frameworks and middleware – adaptive and reflexive components, composition frameworks and middleware, theoretical basis for the construction of scalable distributed software systems
- Automating the engineering process
  - Methods for putting together software “components” to reduce development time and increase reliability, including technologies for developing distributed, autonomous and/or embedded software; automation of software and systems development
  - Integrated software and systems development process, including methods for specifying, analyzing, testing, and verifying software and physical systems
  - Interoperability of applications running concurrently across wide area networks
  - Integrated configurable tool environments that enable rapid compostion and customization of integrated domain-specific development environments
- Pilot applications and empirical evaluation
  - Technologies for embedded software applications and other complex applications
  - Empirical studies of software and systems development projects

7. Expanding human capabilities and supporting universal human development

Information technology holds the potential to help *all people* enhance their individual capacities and skills. The Federal research agenda outlined in this Strategic Plan aggressively pursues technical innovations that bring us closer to universal access to and usability of computing and communications systems. First, it focuses on investigation of ways to integrate advanced functionalities – i.e., computing technologies that input and output speech, translate languages, are activated by sensory data or remote instruction, and the like – so that they best support people performing multiple tasks in varying configurations within complex work environments. Second, it supports end user-focused research to re-invent such IT components as interfaces, search engines, and communications capabilities from the standpoint of expanding the user's capabilities and ease of use. Third, it supports development of technologies, tools, and devices that enable all individuals to live full and independent lives, whatever their age or physical capacities.

Technical research agenda: Expanding human capabilities

- Development of advanced functionalities
  - Language-engineering technologies (including translation between languages and between spoken and written languages, and spoken-language query systems)
  - Spoken, aural, and multimodal interfaces – for hands-free and untethered computing in military and advanced aerospace applications and for computer access for the blind
  - Sensor technologies for use in such settings as health care, national defense, and emergency management, and for the severely physically disabled
  - Real-time interaction with databases to accelerate decisionmaking (for example, Web query response times of less than one second)

- Integration of advanced functionalities
    - Intelligent systems, such as "smart spaces," for ubiquitous computing with multiple interactions; collaborative mobile agents
    - Remote collaboration, visualization, and virtual-reality environments
    - Computer-assisted prosthetics for motion, sight, and hearing; monitoring systems; and remote consultation technologies to increase the independence of the elderly and disabled
    - Technologies and methods for modeling and sharing expertise; models and metrics for collaborative performance of complex tasks

8. Managing and enabling worlds of knowledge – Rx4

Small Federal IT investments to date have pioneered development and implementation of digital repositories of information and such basic enabling technologies as search engines, record management systems, and linkages among distributed archives. Creating digital libraries across the range of human knowledge and developing the technologies and tools to make that knowledge universally available on demand is a core challenge in information technology whose advances benefit every profession, every academic discipline, every learner, and every citizen. Digital libraries form the basis of the Nation's 21st century knowledge network, but we also need advanced IT technologies for working with the information, from visualization, data fusion, and analysis capabilities to remote collaboration and metadata notation schemes, to advanced interoperable systems. This Strategic Plan proposes an expanded research program in this area to build on early Federal successes to reach next-generation technologies in archiving; data access, manipulation, management, and analysis; interoperability; remote collaboration; preservation; and information security.

This research will enable IT to provide the **right** information to the **right** people in the **right** form at the **right** time, a goal the IT R&D Programs call "**Rx4. "**

Technical research agenda: Managing and enabling worlds of knowledge

- Data storage and management technologies
    - Tools for collection, indexing, synthesis, and archiving
    - Protocols for data compatibility, conversion, interoperability, interpretation
    - Technologies and tools for fusion of databases, such as molecules and macromolecular structures in biology or disparate real-time weather observations, with remote access and analysis capabilities
    - Component technologies and integration of dynamic, scalable, flexible information environments
    - Digital representation, preservation, and storage of multimedia collections
    - Protocols and tools to address legal issues such as copyright protection, privacy, and intellectual property management

- Usability of large-scale data sets
    - Intelligent search agents, improved abstracting and summarization techniques, and advanced interfaces
    - Digital classification frameworks and interoperable search architectures
    - Metadata technologies and tools for distributed multimedia archives
    - Ultra-scale data-mining technologies
    - Testbeds for prototyping and evaluating media integration, software functionality, and large-scale applications

9. Supporting development of a world-class IT workforce

U.S. employers in every sector as well as a variety of studies identify the growing shortage of skilled IT workers as the single greatest threat to U.S. competitiveness over the next 10 years. The Strategic Plan proposes to accelerate and expand research on issues in IT education and workforce development, with a focus on barriers and impediments to IT careers among women, minorities, and other underrepresented groups. This research will also foster promising IT applications for classroom and work-related learning by establishing research centers devoted to exploring and developing IT learning technologies.

As researchers representing many different disciplines, participants in the Federal IT R&D Programs know firsthand that the shortage of IT researchers is already jeopardizing their ability to carry out the research agenda that is crucial for the Nation's future. To address this problem, the Strategic Plan also proposes that the Federal research community strive to double the number of new IT researchers over the next five years and increase the support levels for existing faculty.

Technical research agenda: Development of a world-class IT workforce

- Models of cognitive development
- Effects of IT systems on learning
- Software for self-instruction and collaborative learning
- Integration of technologies in learning environments
- Technology-based workforce development and training

10. Understanding the effects of IT to maximize its benefits

New modes of learning, research, communication, commerce, and human services are proliferating so rapidly that we as a society have hardly paused to contemplate the changes or analyze their effects on people and institutions. Most IT research investments to date understandably have centered on development of the new technologies themselves. The PITAC report brought into focus, however, the need for "investment in research to identify, understand, anticipate, and address" the unintended consequences of the increasing pace of technology transformation. The Strategic Plan proposes a vigorous interdisciplinary research program to look much more closely at the nature and dynamics of the interactions between IT and social systems. It will do this by developing both baseline empirical maps of the landscape of social change and new theories and models to describe the complex process of adaptation and interchange between humans and large-scale technical systems.

The research agenda will address a major initial challenge: Development of an intellectual architecture for this new interdisciplinary research area. Researchers currently working on IT-related studies are scattered across many different disciplines without either a "magnetic" center to draw them together or a multidisciplinary communications network oriented to their work. Federal leadership will enable us to create research centers, build a national infrastructure for social, economic, and workforce-related (SEW) research, and attract additional developing scholars to the work to be done. This capacity-building effort will provide policymakers, for the first time, with current, research-based findings about IT's societal effects.

Technical research agenda: Understanding the effects of IT

- Universal participation in a digital society, including such topics as the digital economy, modes of work, Internet governance and Internet citizenship, "digital divide" issues, and cybercrime and law enforcement

- Fundamental theoretic and legal analyses and empirical studies of intellectual property and privacy issues in the digital age
- Using large-scale social technologies for collaboration and learning in science, education, and the workplace
- Ethical principles in IT socio-technical designs
- IT for learning and research in education

Part IV: Putting Federal research into action – National Grand Challenge Applications

Most of the IT enabling technologies the Nation needs – and that constitute the Federal research agenda proposed in this Strategic Plan – are invisible to the public. It is the combination of component technologies in far-reaching applications that marks the visible ultimate goal and crowning achievement of fundamental IT research. Many people think such applications *are the main focus* of IT research. But as this paper explains, applications are effectively *the final step* in an R&D process that begins with methodical, multidisciplinary investigations across a variety of basic and applied sciences.

The bulk of the budget for this five-year Strategic Plan will support fundamental research in enabling technologies. But the IT R&D Programs also propose to test and validate these technologies in prototypes and demonstrations of advanced IT applications. Representative National Grand Challenge Applications are described in the following sections. Several sections are devoted to specific IT applications. Others point to key areas of the national interest in which many advanced applications are needed and several Grand Challenge Applications will be prototyped.

Next-generation national defense and national security systems capabilities

The R&D called for in this Strategic Plan will provide the base technologies to ensure that the U.S. maintains its dominant position in the application of information technologies to critical national defense and national security needs. The investments that are called for in this Plan will provide the national defense and national security communities with the advanced information technologies needed to support weapons programs, military and intelligence operations, and adverse information warfare environments. Systems with these capabilities will be able to perform the computationally intensive fine grain simulation of new aircraft and smart weapons, and will permit full maintenance and reliability simulation of the Nation’s nuclear weapons stockpile. This R&D will enable the efficient design and development of robust and reliable software with the high fault tolerance and high levels of security assurance and intrusion resistance that are vital to the Defense command, control, communications, and intelligence infrastructure. R&D in both microsensors and embedded and autonomous devices will enable the modeling and the management of huge battlespaces involving hundreds of thousands of objects in dynamic combat, support, and intelligence operations. As a result, it will be possible to link autonomous sensor, surveillance, and combat weapon systems to battle management and cyber warfare systems in order to support both defensive and offensive operations with minimum risk of casualties.

The IT R&D Programs will develop and demonstrate new generations of highly secure, fault tolerant computing, networking, and storage technologies, including high end computing systems and distributed autonomous and embedded devices and systems, needed in weapons systems, battlespace, and national security applications.

Improved health care systems for all citizens

Secure, high-speed networks and software that is reliable, interoperable, and safe from intrusion will enable basic improvements in the national health care infrastructure, such as high-confidence software for medical devices including life-support systems; management and usability of patient information; interactions between patients and health care providers; timely analysis of provider and institutional quality; and hospital systems, inventory, and procurement management.

More dramatic will be the extension of monitoring, diagnosis, care, emergency treatment, and even surgery to citizens in remote locations, or unable to reach the hospital, or housebound. Experimentation with telemedicine is showing the enormous promise in combining high-speed networking, two-way real-time video, embedded and robotic devices, and remote visualization and instrumentation to get needed care to citizens immediately wherever they are located. These capabilities will also make it possible to help maintain the independence of aging citizens and of citizens with physical limitations. In addition, this set of technologies will enable a whole new generation of techniques and practices in medical training and physician and health care professional continuing education.

The IT R&D Programs will prototype and demonstrate high-confidence medical devices; multimodal systems for remote and emergency on-site patient care; advanced home devices and services for individuals with physical limitations; and advanced, distributed multimedia capabilities for medical education, biomedical research, and clinical practice.

Creating scientifically accurate, 3-D functional models of the human body

Advances in computational speeds, visualization software, and data storage capacities are bringing us closer to being able to generate large-scale 3-D models and simulations of enormously complex phenomena such as the human body. To suggest how computationally challenging such models are: It is taking the world's fastest computing platforms in the Federal government's national research laboratories to begin to create quantitatively accurate visualizations of the Nation's nuclear weapons stockpile. It will take substantially more computational capacity to generate a precise 3-D visual model of the human body, starting from atoms, molecules, and cells, through organs and the vertebrate and circulatory systems. Federally funded researchers are working today on visualizing the neuronal structure of the brain. The scale of this problem alone is exemplified by the fact that one cubic millimeter of cerebral cortex may contain on the order of five billion interdigitated synapses of different shapes and sizes and a wide variety of subcellular chemical signaling pathways. Being able to visualize, manipulate, and test representations of structures and processes at this level of matter will mark an invaluable innovation for both scientific research and education.

The IT R&D Programs propose to harness IT advances to create a complete, functional digital model of the human body.

IT tools for environmental modeling and monitoring

Advanced IT modeling, simulation, visualization, and analysis tools will also improve our ability to study and understand such complex phenomena as global warming, food shortages, energy depletion, drought, natural disasters, and human/environment interactions. More accurate measurement and analysis of such phenomena will provide better information for decision making in both the private and public sectors.

Developing a next-generation environmental monitoring, modeling, and prediction system will require real-time monitoring and observations above the Earth, on the Earth, and under ground. Because these real-time observations will be global in scale, the system will require high-speed digital connectivity and high-end computing platforms. The data must then be integrated with timely contextual knowledge in geophysics, biology, chemistry, and atmospheric and oceanic sciences. A key challenge in developing this application is the great complexity of assimilating observational data with models. Scientists will need new methods of visualization to understand the complexity and the spatial and time evolution of the underlying processes. Integration and synthesis of multidisciplinary data with advanced, high-resolution models will require coordination of component technologies, specialized languages for scientific software, storage strategies with very large capacities and good access characteristics, and metadata and search capabilities that include environmental semantics, data fusion, and data mining and/or automated pattern recognition.

The IT R&D Programs propose to prototype and demonstrate environmental monitoring and modeling systems to improve forecasting of hazardous weather, evolution of hazardous spills, response of ecosystems to environmental change, and impacts of earthquakes.

Integrated IT systems for crises management

In a major natural or human-caused disaster, there is a great need for an instantaneous common communication system and a common capability for real-time distribution of precise information, disaster guidance and directives, situational updates and analyses, and instructions for distributed disaster workers. To date, we have not put development of such a coordinated crises management system on the national agenda. It is time to put the mobile wireless, nomadic, and satellite communications technologies now available together with scalable wireline networking capabilities, advanced microsensing technologies, data analysis and system-management software, and with our extensive multidisciplinary experience in crisis management (for example, public health, emergency response, medical triage, fire, and policing) to create a state-of-the-art crisis coordination and management system that can be deployed immediately and effectively in any kind of catastrophic situation.

The IT R&D Programs propose to support creation of a collaborative, interdisciplinary Enabling IT Center for Crises Management to develop and demonstrate this comprehensive technologies framework. Federal agencies, with state and local government and private-sector partners, have the technologies, the personnel, and the broad experience in major environmental and other disasters to successfully build this much-needed grand challenge application.

IT-enabled integrated intermodal transportation system

The current national transportation system is made up of three main modes -- land (*e.g.,* pedestrian, highways, transit, motor carriers, rail, pipelines), water (maritime, waterways), and air and space. These modes are connected “intermodally” to provide for the transport of people, natural resources, and goods. But intermodal connections are not optimal today. Public emphasis on the value of time (and therefore, doorstep-to-destination speed) will take on increased importance early in the 21st century and will result in demand for faster, more efficient, and affordable transportation services. In addition, with the information revolution, a new mode of transport is emerging – the information or virtual mode. This “fourth” mode will provide global connectivity anywhere, anytime. While some physical transport needs will be reduced by the information revolution, throughout history increased communication ability has also resulted in increased demand for physical transportation. Information will not only make individual transportation modes more efficient, but could fundamentally change the ways in which we can conceive of totally integrated intermodal transportation systems in the 21st century.

The long-term goals for this IT investment are:

- An integrated national transportation system that can move anyone and anything, anywhere, anytime, on time at an affordable price
- A transportation system without fatalities and injuries resulting from system or operator error, or terrorist intervention
- A transportation system without transportation-related environmental impacts (e.g., noxious emissions, greenhouse gases, noise) and not dependent upon foreign energy

Integrated, advanced aviation system technology

Aviation safety and capacity are national issues that are reported in the newspaper daily. The air transportation system is on the verge of gridlock, with delays and cancelled flights last summer reaching all-time highs and passenger rage skyrocketing. As demand for air transportation continues to increase, fueled by a strong economy and the package-delivery demands of e-commerce, the capacity of the air traffic control system needed to accommodate the anticipated growth is falling farther behind. It has become painfully apparent that the present air traffic control system cannot continue to be scaled up to provide the capacity increases required in the next 15 to 25 years. We need a fundamental change in the management of the aviation system and information technology is the key.

High-performance computational and networking,technologies, in combination with advanced applications in visualization, modeling, simulation, and distributed instrumentation make it possible now to design a fully integrated, large-scale aviation system in the air and on the ground. Such a next-generation IT infrastructure would vastly increase the capacity of the air transport system to move people and cargo through integrated airspace operations. This integrated system would enable real-time sharing of information from distributed sources such as weather stations, automated air-traffic management systems, flight controllers, passenger managers, and other transport-related nodes. IT challenges are to develop:

- The critical core component technologies to meet the requirements of the air transportation system
- A virtual airspace transportation environment for simulating the air traffic components at the system level with the requisite degree of fidelity
- Evaluation of candidate system-level concepts and architectures making use of the “virtual air transportation environment”

Creating a world-class infrastructure for lifelong learning

Lifelong education, training, and development have become necessities of the Information Age. With human knowledge doubling every two years and dynamic work environments calling for continuous skills development and adaptability to new information, the ability to keep learning is perhaps this era's core requirement for successful employment and career development. We currently have, or will soon, the enabling technologies in high-speed networks, software for information management, real-time collaboration, 3-D visualization, and the like to create multifaceted learning environments and experiences for learners of every age with every kind of academic, vocational, or personal learning focus. IT can provide ubiquitous access to structured knowledge (systematic course work, laboratory activities, and rich digital libraries) as well as immersive environments for experiencing scientific phenomena and different cultures and environments. IT interfaces and experiences can be tailored to individual learning styles, ages, physical and mental capacities, and interests, with automated feedback systems to guide progress.

The IT R&D Programs propose to demonstrate prototypes of advanced learning systems for education, training, and development across age groups and needs.

Part V: Realizing the IT promise

The imperative for Federal leadership

As the developers of this five-year Strategic Plan, the Federal IT R&D Programs strongly urge the new Administration to build on the interagency effort's demonstrated 10-year record of success with this bold investment to make fundamental IT R&D a high-visibility national priority. Federal IT research not only is essential for critical Federal missions but also helps fulfill the promise of information technology in innovations and applications that benefit all Americans.

For all the reasons noted in the Strategic Plan, this vital fundamental research will not be undertaken by the private sector. But without it the Nation cannot achieve the scientific and technical breakthroughs urgently needed in mission-critical national defense and national security applications, advanced scientific research, manufacturing and services that lie at the heart of our economic competitiveness, and biomedical research and health care. All of these vital sectors now depend on effective communications networks, reliable computing systems, and information-rich digital archives.

The Strategic Plan details these enormous national benefits, and it sets out the technical challenges that Federal research priorities must address to drive substantial IT advances. The Federal IT R&D agencies agree with the President's Information Technology Advisory Committee, however, that it will take aggressive Federal leadership to press the IT research agenda that will maintain and strengthen the Nation's competitive edge in the global economy and assure that our national defense and national security systems are second to none.

Management of the Federal IT R&D Programs' Strategic Plan

The Interagency Working Group on Information Technology R&D (IWG/IT R&D, successor to the Subcommittee on Computing, Information, and Communications and before that the Subcommittee on High Performance Computing, Communications, and Information Technology) serves as the coordinating leadership body for the IT R&D Programs. Made up of representatives of the 12 Federal agencies in the Programs, plus representatives from the Office of Science and Technology Policy and the Office of Management and Budget (OMB) in the Executive Office of the President, the IWG provides IT R&D policy, program, and budget guidance for the Executive Branch and coordinates multiagency IT R&D activities. The IWG also works with other Federal agencies that need advanced IT to identify their requirements and accelerate development of appropriate technologies.

Six Coordinating Groups (CGs), representing the major research emphases of the IT R&D Programs and the program managers in participating agencies, report to the IWG. These groups confer regularly to coordinate the objectives and activities of the multiagency projects in their specialized research domains, called Program Component Areas (PCAs). The PCAs are:

- High End Computing and Computation (HECC), which includes both HEC (High End Computing) R&D and HEC Infrastructure & Applications (I&A)
- Human Computer Interface & Information Management (HCI&IM)
- Large Scale Networking (LSN)
- Software Design and Productivity (SDP)
- High Confidence Software and Systems (HCSS)
- Social, Economic, and Workforce Implications of IT and IT Workforce Development (SEW)

Funding for agency IT R&D activities is implemented through standard budgeting and appropriations processes that involve the participating agencies and departments, OMB, and the Congress. Some activities are funded and managed by individual agencies. Others involve multiagency collaboration, with mutual planning and mutual defense of budgets. For some highly complex, mission-critical R&D efforts, such as the thee-year Next Generation Internet Initiative begun in FY 1998 and the current HEC R&D program, agencies create integrated programs and budgets and detailed management plans. The work of the IWG and the Coordinating Groups, and interactions with OSTP, OMB, and the Congress, are supported by the National Coordination Office for IT R&D.

The 12 agencies that developed the Strategic Plan and their primary IT R&D interests are:

AHRQ – the Agency for Healthcare Research and Quality – focuses on developing state-of-the-art IT for use in health care applications such as computer-based patient records, clinical decision support systems, and standards for patient care data, information access, and telehealth.

DARPA – the Defense Advanced Research Projects Agency – is focused on future generations computing, communications, and software technologies, and human use of information technologies in national defense applications such as battlefield awareness.

DOE/NNSA – Department of Energy National Nuclear Security Agency – was established to develop new means of assessing the performance of nuclear weapon systems, predict their safety and reliability, and certify their functionality through high-fidelity computer simulations and models of weapon systems.

The DOE Office of Science is discovering, developing, and deploying the computational and networking tools that enable researchers in the scientific disciplines to analyze, model, simulate, and predict complex physical, chemical, and biological phenomena important to DOE. The office also provides support for the geographically distributed scientific teams and remote users of experimental facilities that are critical to the Department's missions.

EPA – the Environmental Protection Agency – has the IT R&D research goal of facilitating multidisciplinary ecosystem modeling, risk assessment, and environmental decision-making at the Federal, state, and local levels, and in corporations, through advanced use of computing and other information technologies.

NASA – the National Aeronautics and Space Administration – is extending U.S. technological leadership to benefit the U.S. aeronautics, Earth and space science, and spaceborne research communities.

NIH – the National Institutes of Health – is developing the basic knowledge for the understanding, diagnosis, treatment, and prevention of human disease, including the storage, curation, analysis, and retrieval of biomedical data and information.

NIST – the National Institute of Standards and Technology – is working with industry, educational, and government organizations to make IT systems more useable, secure, scalable, and interoperable; apply IT in specialized areas such as manufacturing and biotechnology; and encourage private-sector companies to accelerate development of IT innovations.

NOAA – the National Oceanic and Atmospheric Administration – is an early adopter of emerging computing technologies for improved climate modeling and weather forecasting, and of emerging communications technologies for disseminating weather warnings, forecasts, and environmental information to users such as policymakers, emergency managers, and the general public.

NSA – the National Security Agency – is addressing some of the most challenging problems in the country in computing, storage, communications, and information assurance in order to help ensure our national security.

NSF – the National Science Foundation – is the lead agency in the IT R&D Programs, with interest in developing new fundamental IT knowledge; applications in the biological, chemical, geophysical, physical, and mathematical sciences and engineering; educating world-class scientists and engineers and a knowledgeable IT workforce; and research infrastructure.

ODUSD/URI – the Office of the Deputy Under Secretary of Defense's University Research Initiative – focuses on IT R&D for Department of Defense applications, research infrastructure, and science and engineering education.

Other Federal agencies participate in information technology research and development, and coordinate with the IT R&D Programs, using funds that are not budgeted under the IT R&D Programs.