President's Information Technology Advisory Committee (PITAC)

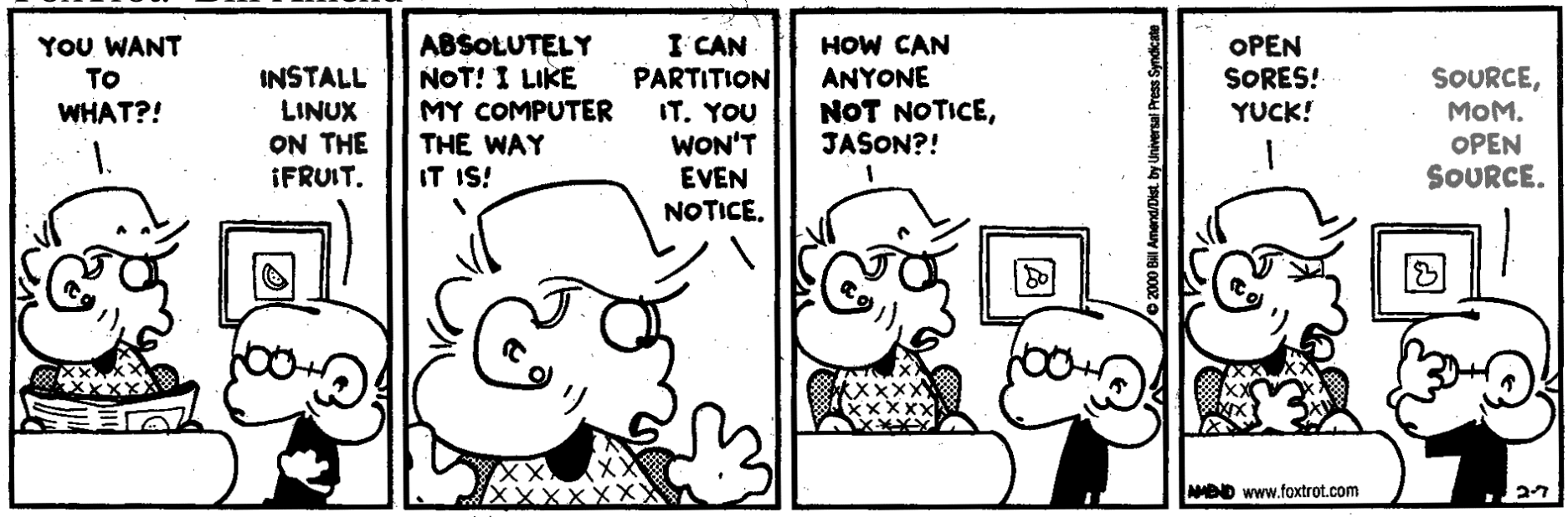
Panel on Open Source Software
for High End Computing

Larry Smarr, Chair

President's Information Technology Advisory Committee

The Time Is Now!

FoxTrot: Bill Amend



Linux and Open Source have Passed the "Huh?" Test in Washington DC

The Open Source Phenomenon

- Working Group on Libre Software, December 1999
  - "The impact of open source technology is expected to be quite noticeable in the software industry, and in society as a whole. It allows for novel development models, which have already been demonstrated to be especially well suited to efficiently take advantage of the work of developers spread across all corners of the planet. It also enables completely new business models, which are shaping a network of groups and companies based on open source software development. And it has, in general, a very positive impact as an enabler for the creation of new markets and business opportunities."

- Examples
  - NCSA HTTPd→Apache — Web Server Software
  - Perl — Interpretive Scripting Language and Libraries
  - BIND — Software Implementation of DNS
  - Sendmail — Email SMTP
  - GNU, Emacs, gcc — Programming Tools, Editor, compiler
  - Linux — Operating System

What Defines Open Source?

- Open Source has become the generic term for describing software that is freely distributable without restriction. (Beckman/Oldehoeft, Feb, 2000)
- The main features that characterize free (open source) software is the freedom that each user has to: (Working Group on Libre Software, Dec, 99)
  - Use the software as one wishes, for whatever one wishes, on as many computers as one wishes, in any technically appropriate situation.
  - Have the software at ones disposal to fit it to ones needs. Of course, this includes improving it, fixing its bugs, augmenting its functionality, and studying its operation.
  - Redistribute the software to other users, who could themselves use it according to their needs. This redistribution can be done for free, or at a charge, not fixed beforehand.
  - The users of a piece of software must have access to its source code.
- Nature of Licenses (GPL vs BSD) and Reuse
- Radical Decentralization of Authorship
- Cathedral vs. Bazaar (Both have been used for OS)

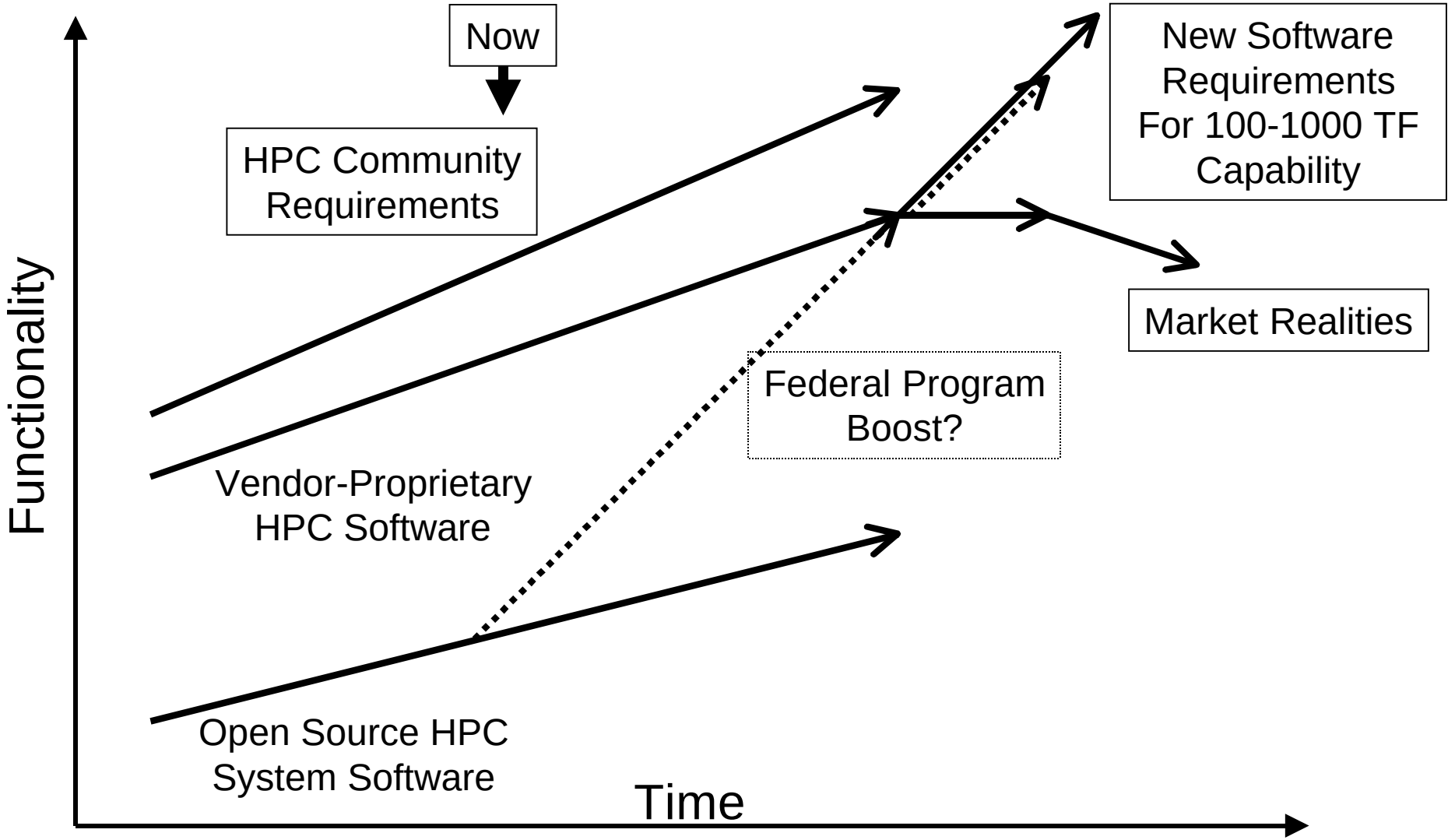
HPC-OS Meeting History

Santa Monica	Livermore	Oak Ridge	Santa Fe	Chicago
August 2-3 1999	August 31 1999	October 5-6 1999	December 13-14 1999	February 17-18 2000

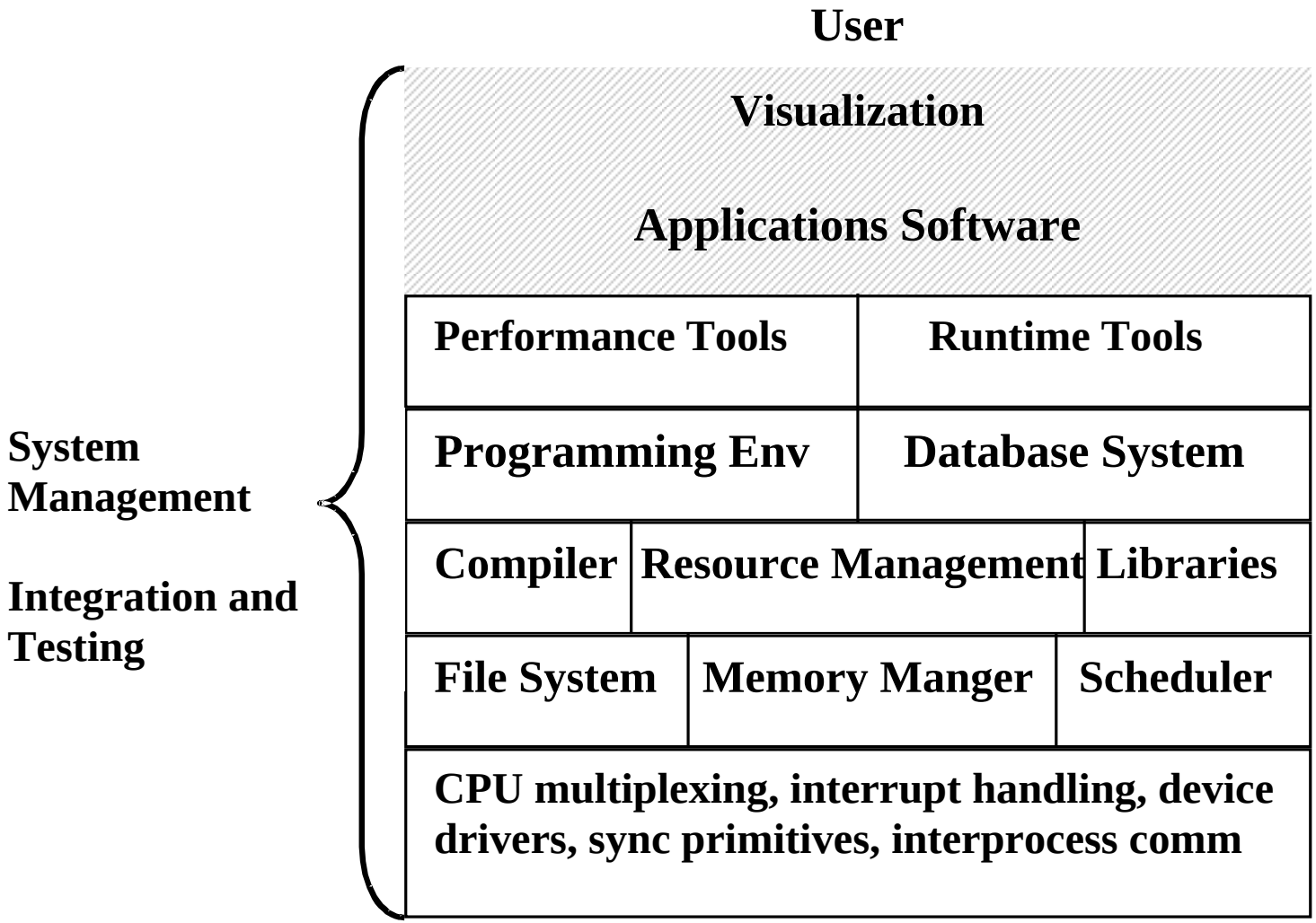
Motivation for the Santa Monica Meeting

- Looming Crisis in High Performance Computing System Software
  - HPC Software Is Complex and Requires Great Expertise to Develop
  - The Market for HPC System Software Is Not Economically Viable
  - Many HPC Requirements Are Not Being Met, and Vendors Can Not Feasibly Be Expected to Meet Them
- Growth of Open Source
  - Demonstrated Success of Open Source Development Model
  - Rise in Visibility of Linux, Apache, and Others
- Thus the Motivating Question:
  - Can We Build a Completely Open Source System Software for HPC Systems, and in So Doing, Improve the World for Both HPC Users and HPC Vendors?

Building Open Source to Meet HPC Needs



The Software Stack Needed for HPC



Adapted from: Distributed Operating System & Algorithms, Chow/Johnson

Source: Jose Munoz, DOE

Many Models For Working with Vendors

- The “Stone Soup” Model
  - Everyone Contributes to Community-Owned Software Base
- The “Reference Implementation” Model
- The SIA Model
  - Government Does the Research, Industry Does the Implementation
- Explicit Partnerships With Industry
- LANL ACL Software Carpentry Design & Implementation
  - $860,000 Over Two Years Paid to People to Develop Software
  - Red Hat, O’Reilly, etc. Have Joined In

Potential Upsides to OSS for HPC Approach

- Brings HPC Community Together to:
  - Unify and Leverage Multi-Agency HPC Investments
  - Define Needed Software Stressing HPC Unique Requirements
  - Improve Coordination, Design, Integration, and Testing of Software
  - Bridge the Gap Between Research and Production Software
  - Create Cooperation Between Industrial and Research Institutions
- Modifications and Extensions to Delivered Software Can Be:
  - Authored Locally
  - Shared Globally
- Drive Cost Effective and Rapid Technical Development
- Code Can be Continued After Company Bankruptcies
- HPC Software From Cottage Industry to Global Enterprise?

Questions About an OSS for HPC Approach

- Is There Enough Lead Time for Community to Catch Up to Hardware Deployments?
- Can Open Source Approach Work with Time Deadlines?
- Are There Successful Precedents for Multi-Agency Multi-Lab Coordinated Development of Complex Software?
- Open Source Can Produce Code Chunks, but Can It Guarantee Layered Integration?
- How to Move From Research Quality Code to Production Quality Code?
  - Coordination and Accountability Must Be Part of Solution

What is the Federal Government's Role?

- Private Sector Open Source is Rapidly Growing
  - Red Hat $10 Billion Market Cap
  - VA Linux $4 Billion Market Cap
- However, Private Sector is Based on Large Installed Base Systems
  - PCs
  - Server Farms
- High Performance Systems Have Small Installed Base
  - "Top 500" Compared to Hundreds of Millions of PCs
  - Federal Government Has Critical Need of HPC
    - National Defense
    - Science and Engineering Research
    - Crisis Management

Areas to Be Explored by Panel:

- Education of Agencies, Universities, Labs, &Researchers About:
  - Advantages of OSS
  - Details of Open Source Software & Licensing
  - How OSS Can Help HPCers Achieve Mission Goals
- Possible Changes in Federal Policy:
  - Ways for Agencies to Encourage Open Source Development
  - Agencies Require OSS in HPC Procurements?
- Focus Federal Funding on :
  - Competitive Grants for OSS Creation and Support
  - Virtual Teams for Specific OSS Goals
  - CS Research on Improvements to OSS

Open Source Software for High End Computing Panel Membership

Chair
Larry Smarr

Overall Architecture
Steve Wallach-PITAC
Rick Stevens-ANL
Bo Ewald-PITAC
Dave Cooper-PITAC
Bill Carlson-NSA/IDA
Tom Sterling-Caltech/JPL
*Kai Li-Princeton

Industry
Dan Frye-IBM
Tom Gibbs-Intel
Greg Lindahl-High Performance Technology Inc
*Tim O'Reilly-O'Reilly & Associates
Todd Needham-Microsoft
Greg Chesson-SGI
* Michael Tieman - Redhat

Systems
Peter Beckman-LANL
Dennis Gannon-IU
Ian Foster-ANL
Sue Graham-PITAC
*Garth Gibson- CMU
Ken Kennedy-PITAC
Irving Wladawsky-Berger-PITAC
John Reynders-LANL
Jim Gray-PITAC

PACI
John Toole-Alliance
Philip Papadopoulos-NPACI

Intellectual Property
John Perry Barlow-EFF

* Pending Confirmation

Application, Algorithms, Libraies
Bob Sugar-UCSB
Bob Lucas-LBL
Jack Dongarra-Univ of Tenn

Government Liason
Bob Borchers-NSF/HCCWG
Paul Messina-DOE
Bill Feiereisen-NASA
Jose Munoz-DOE/HCCWG
Dan Hitchcock-DOE/ER
Bill Nitzberg-NASA
Fred Johnson-DOE/ER

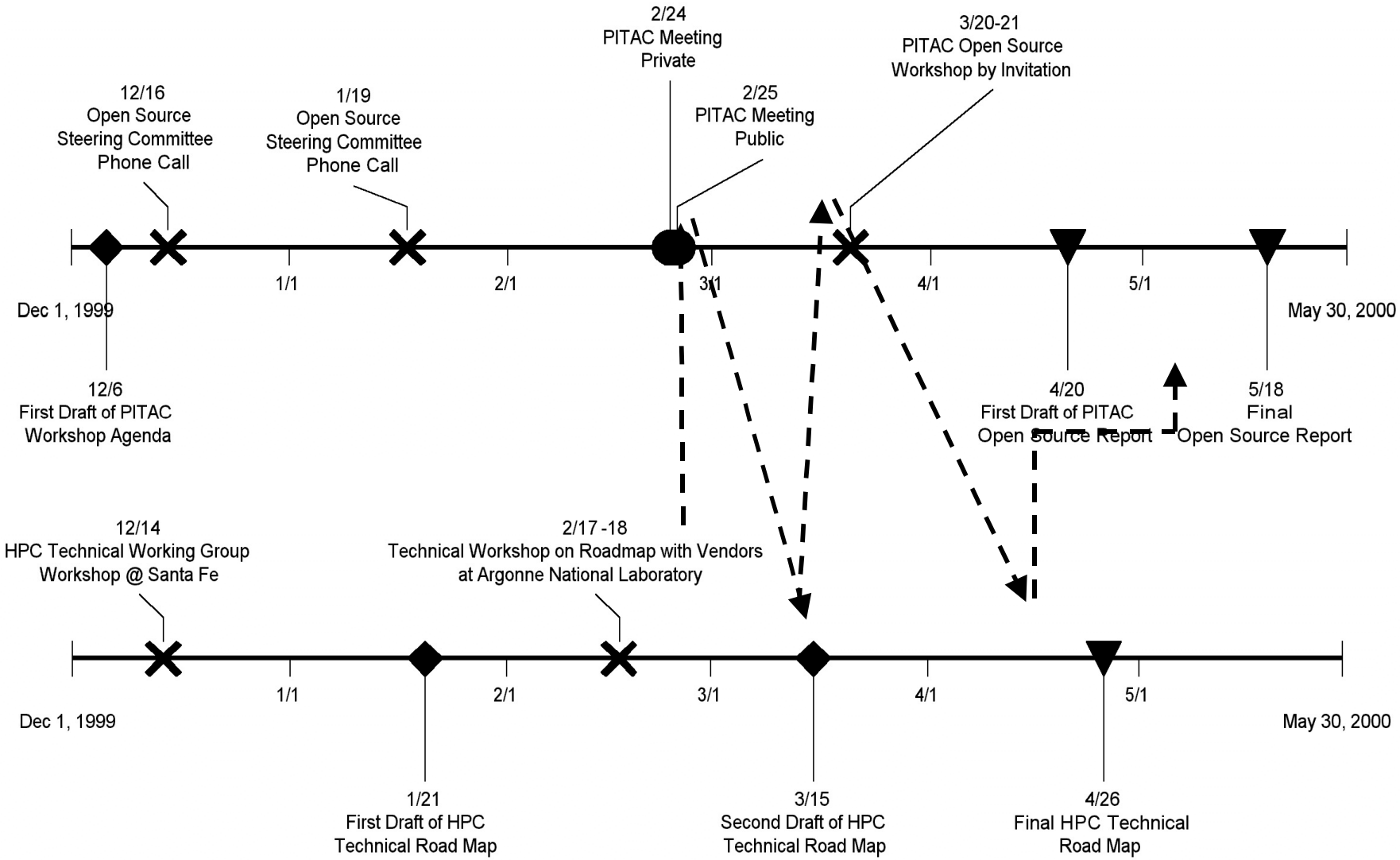
Charter for Panel on OSS for HPC

- Objectives:
  - Develop a Vision of How the Federal Government Can Support the Developing Open Source Software for HPC Activities
  - Define a Policy Framework for Accomplishing This
    - Identify Policy, Legal, or Administrative Barriers to Widespread Adoption of Open Source Software Efforts
    - Define Potential Roles for Public Institutions in Open Source Software Economic Models
  - Encourage Broad Technical Working Groups Drawn From Labs, Academia, Agencies, and Industry to Draw up Technical Roadmaps for Creating Open Source Software for HPC

Charter for Panel on OSS for HPC

- Approach Methodology:
  - Analyze Pertinent Existing Government, Academic, and Private Sector Open Source Efforts
  - Gather and Assess Recent Studies on the Advantage/Disadvantages of Open Source Software Development Efforts
  - Build on the Work of Members of the Community
  - Hold Workshop to Explore the Issues
  - Issue Report

PITAC Panel on OSS for HPC Timeline



PITAC Panel on
HPC Open Source Software

President's Information Technology Advisory Committee