HPC Planning

(strategic management, procurement, funding)

CRAIG TULL (LBNL)
MARK GARY (LLNL)

Session goal

Survey/discuss planning best practices <u>from an HPC perspective</u>

Answer the following questions

1. *What are best practices and tools? Inside and outside HPC*
2. *What are the top challenges?*
3. *What new technologies are needed?*

Getting started

- Introductions (name & affiliation)
- Other volunteers welcome (out-brief, report…)
- Out-brief and survey questions

SW Planning session outline

- First, what is unique about HPC planning?
- Existing HPC planning best practices
- How is project cost planned, managed, monitored?
- Sustainment funding
- Interagency and international cooperation

<Break>

- Strategic choices in software licenses
- Top challenge - Forecasting requirements
- Other top HPC planning challenges
- New **technologies** needed for HPC planning?

Practices

Challenges

First, what is unique about HPC Planning?

How does HPC planning differ from industry, data collection environments, high-end analytics,…?

- HPC SW planning involves both integrating extant products and green-field products.
- Lifecycle of software overlaps use and development.
- Recurrent challenges (eg. scaling) must be resolved.
- When does research become development become production?
- There **are** many similarities with industry. Many industries (eg Google, Sun) do SW research.

What is unique? (cont.)

- Funding source and reward model are different than industry. However, we are expected to leave behind used systems if we expect sustained funding.
  - Quantitative ($) measure of success in industry.
- Requirements from stakeholder are sometimes unrealistic. Scientists don't necessarily see out-year utility of speculative CS research.
- HPC centers must support a wide spectrum of SW.
  - From hardened, guaranteed to someone else supports.
- Planning must encompass entire lifecycle through sustainability.

What rules and resources exist for HPC project planning?

- What mandates exist?
  - Mandates - DOE Order 414.1C
    - Not too much attention paid to this order.
  - ITIL - ASCR has considered, but not mandated
    - more for operation than SW
  - ASCR requires OSS license for government-funded products (slight preference for BSD). And encourage copywriting.
  - Risk Grading is important
    - Simulations lead to real-world policy decisions.

Rules & Resources (cont)

- What applicable standards exist?
  - IEEE 1058 Software project management
  - IEEE 1074 Software life cycle processes

What are your best practices?

- **Early investment**
  - Laboratory LDRD & seed funding - process differs between labs.
    - Even at this point, high expectation of success is required.
    - Something small and speculative is needed to explore new ideas.
  - PathForward, Alliances, …
    - These proved very productive. A very clear, over-arching vision helped.
    - The well-structured evaluation process informed agency decisions.
    - Substantial funding made these substantive.
    - Close management was essential.
- Investment choices
  - Test infrastructures
  - Balanced infrastructures
  - Investment in software and people are a critical investment choice.

Best Practices (cont)

- Planning vehicles
  - Yearly blueprint process to synchronize.
  - ASCR does an annual operational review of each site.
  - Lehman review process is typically reserved for facilities. The SW supported is a (small) part of that.
  - HW evolution and SW evolution are tightly coupled.
  - Agile/Extreme programming is very effective on the short term. Hard feedback from tightly-coupled users invaluable.
    - Fly your own airplane.
- Tactical procedures required to keep facility operations on track. (eg. daily/weekly meetings)
  - Software developers need to be involved.

Best Practices (cont)

- Measured innovation
  - Cost benefit analysis and prioritization.
- Proposals submitted within an organization/program.
  - Formal peer-review process.
- Vendor/Industry Relationships - different models work in different circumstances.
  - Funded Investigation - reputations (only) at risk
  - Fully Defined Purchase - financial risk on vendor
  - Design & Development - specification of goals and requirements
  - Codesign & Codevelopment (eg HPSS) - close partnership
  - Base + Value Added - additional reward for additional value
- **Tracking and metrics to determine use and value inform quantitative planning/investment decisions by developers, facilities, agencies.**
  - eg. NERSC module loads, spyware-like tools to monitor run-time use, download statistics for OSS, filename pattern matching, …

Current Management Schemes

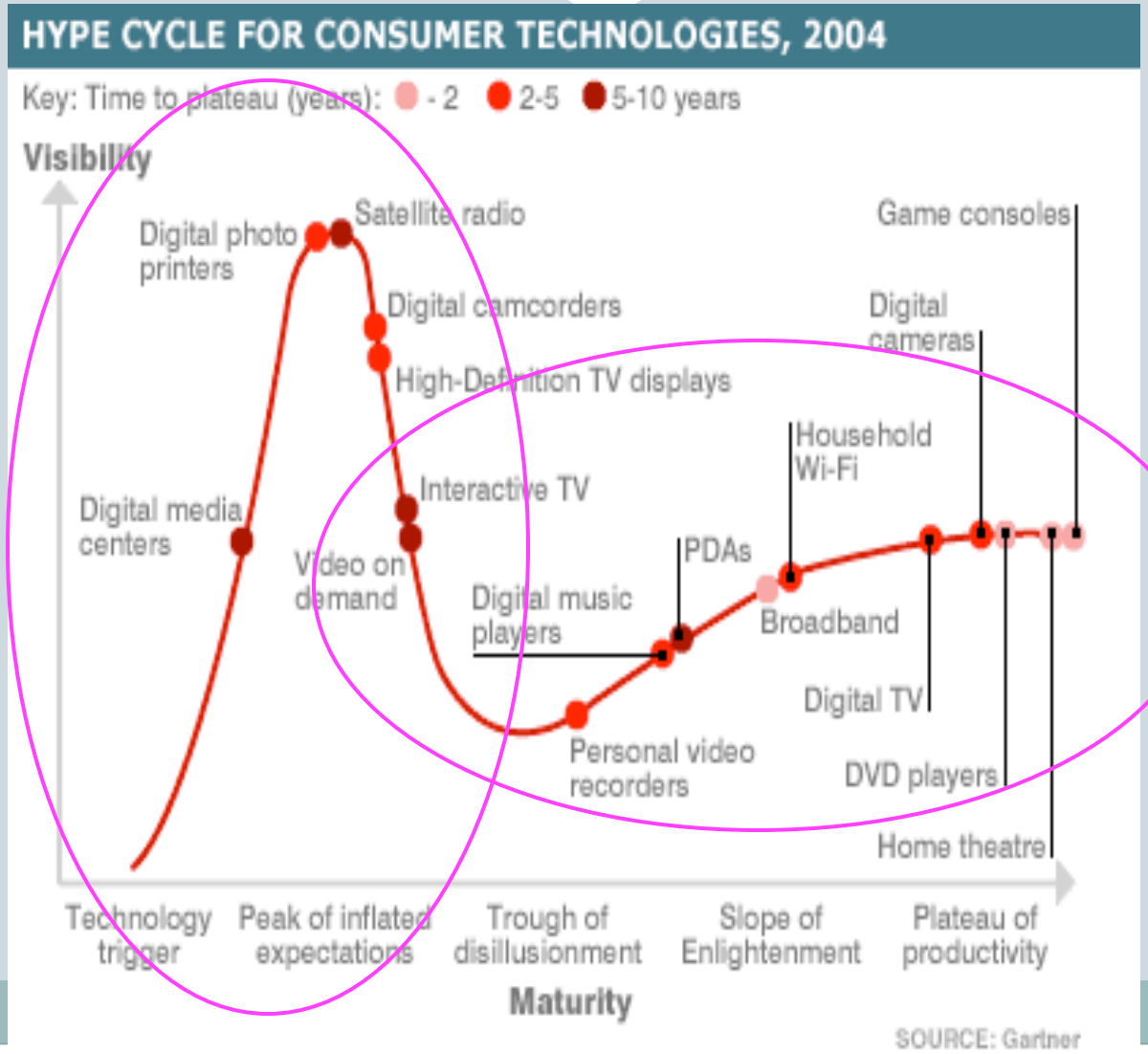
- Most/All sites need improvement to facilitate planning.
- Commercial SW is more strictly managed because costs are concrete.

How is project cost managed/monitored

- EVMS typically applied to procurement & installation of HW. Pure SW projects glean no recognizable value added. Too rigid for SW technology innovation.
- Limited lifetimes applied to some products.
- Two page explanation of goal of project.
- Continual feedback needed to keep SW relevant.
- Project-related CE standards and practices increase productivity/decrease costs.

Gardner's Hype Cycle

Easy to get agency R & D funding here.



Here is where useful SW sees most of it's life.

When software development funding ends, what are your strategies to maintain it?

- We don't have pay-for-use model.
  - Agencies must be willing to sustain critical SW and fund the projectizing necessary.
- Are the R&D and the maintenance developers the same people? Often, no.
- Rewards for good people following the lifecycle must be available. (eg. salary, reputation, employment)
- Many small items can add up.
- **Is a SW sustainability center called for?**
  - First line of user support for multiple SW.

Strategies for interagency and international cooperation

- Examples
  - HPSS, Lustre, OSG, OpenMPI, …
  - Tri-Lab software stack - enable users to migrate between machines.
- Some observations:
  - Big things are easy, but small things lead to redundancy.
  - Exascale power costs will lead to fewer large HW facilities.
    - Most centers will be invested HW outside the facility.
- Collaboration challenges:
  - Foreign nationals/Export control
  - Competition/Proprietary property
  - Legal/Cybersecurity/Licensing/Copyright
  - Trust relationships take a long time to build
    - reliance on long-term availability.

Top challenge - Forecasting requirements

- How do you gather requirements and how accurate are they?
  - Science visits, conferences, requirements database, high-level validation, embedded developers, constraintless thinking paired with allocation process, education & negotiation to mitigate unrealizable expectations, exploration of real/underlying requirement, trouble ticket feature requests (don't always work), face-time and fielding complaints, user-centric documentation & implementation plans, use cases
- When and how often do you do this? **Continually**
- Do your requirements include software performance?
- Do you use metrics to validate requirement extrapolations?
  - Measurability is a best practice, but…
- Do you follow up to verify the accuracy of the requirements?
  - Alpha-user feedback and real-world use are the best measure of accuracy.

Other top challenges surrounding HPC planning

- Vendor and technology unknowns
- **Funding gap between R&D and applied use by domain scientists**
- **Measuring impact on delivered science**
  - Measuring use is easier, but non-trivial
- Interoperability and compatibility and platform coverage.
- **Planning occurs at the component level without cross-cutting requirements and design principles.**

What new technologies are needed to help with HPC planning?

- Collaboration technologies?
- Standards (focus on science rather than mechanism)
  - eg. Portable data formats, software bus architectures

What else?

Next steps

- Survey questions
- Out-brief summary
- Report

HPC Planning
(strategic management, procurement, funding)

CRAIG TULL (LBNL)
MARK GARY (LLNL)

Planning

- Measuring use and forecasting needs
- Assessing criticality and establishing support priorities
- Identifying and managing cost
- Strategies for interagency and international cooperation
- Strategic choices in software licenses

Planning

- Strategic management, procurement, funding HPC software products often have lives that span multiple decades while serving many generations of machines and operating environments. Careful project planning is the foundation upon which these projects are built. From requirements gathering and cost estimation to collaboration and team building, deliberate and realistic planning is the key to product usefulness and longevity. But how do HPC software projects differ from typical software development projects? Do HPC requirements or the HPC community introduce impediments to successful planning? Successful collaboration? Are we in the HPC community successfully leveraging non-HPC methodologies? This session will address these questions, investigate the facets of good software planning, and explore alternative planning approaches.

What are best practices and tools?

- *Continual and conscious seeking of requirements.*
- *Measurement and tracking of use and value.*

What are the top challenges?

What new technologies are needed?

Findings and Recommendations

Strategic choices in software licenses

- Open source and its impact
  - Risk mitigation
  - Influencing HPC requirement implementation
  - Gatekeeper choice
- Proprietary licenses
  - Charging models and change
  - Vendor stability