Abstract

- **High Performance Computing facilities face increased pressures to survive and thrive in the next millennium. HPC facilities must combine effective techniques of the past with innovative methods of the future. This tutorial explores requirements and pressures on HPC centers, and presents effective methods being employed and new approaches to employ to overcome these challenges.**

Abstract (cont.)

- **Topics include:**
  - **The current state of HPC computing and projections.**
  - **System management that make possible to achieve CPU utilization greater than 90% for MPPs running many large jobs.**
  - **Archive storage issues of improving transfer bandwidth and practical advice for running Terabyte archives.**
  - **Innovations for client services to ensure the "intellectual resource" is equally valued by clients as the systems.**
  - **Introduce the Effective System Performance Test, a new way to objectively measure and compare not just system performance (e.g. sustained performance of a applications) but also system effectiveness (e.g. how many system resources, especially CPU time can really be used by the workload over time).**
  - **Integrating production with a good research program is critical to maintaining a robust HPC facility. The tutorial will address how to achieve and maintain this delicate balance.**
- **It explores what a facility needs to do to thrive in the new millennium.**

High Performance Computing Facilities for the Next Millennium

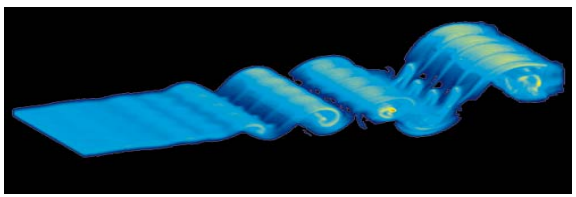
Large Scale Computing Today and in the Future

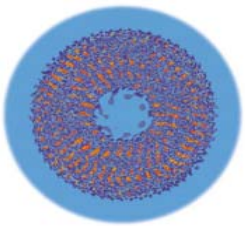
SC 99 Tutorial
November 14, 1999

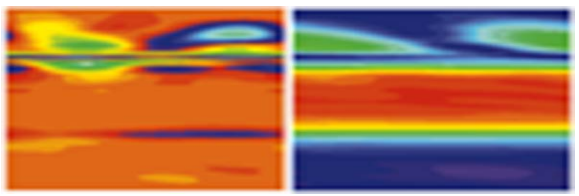
Bill Kramer
NERSC Deputy Division Director
kramer@nersc.gov
510 486-7577

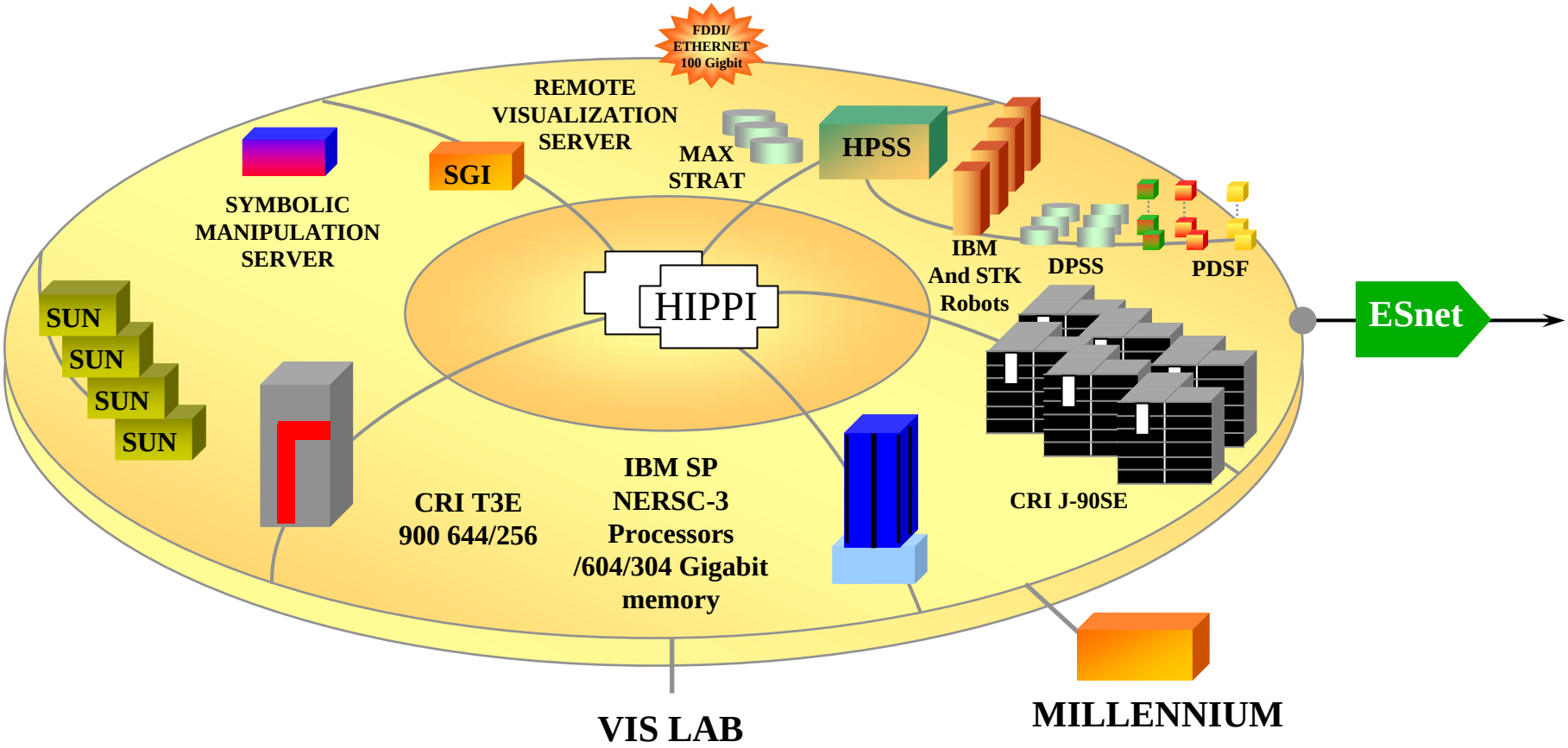
Outline

- Large Scale Computing Today and in the Future - *Bill Kramer*
- Client Services - *Francesca Verdier*
- Mass Storage - *Keith Fitzgerald*
- Computational Systems - *James Craw*
- Dealing with new technology - *Tammy Welcome*
- The NERSC Effective System Performance Test - *Bill Kramer*

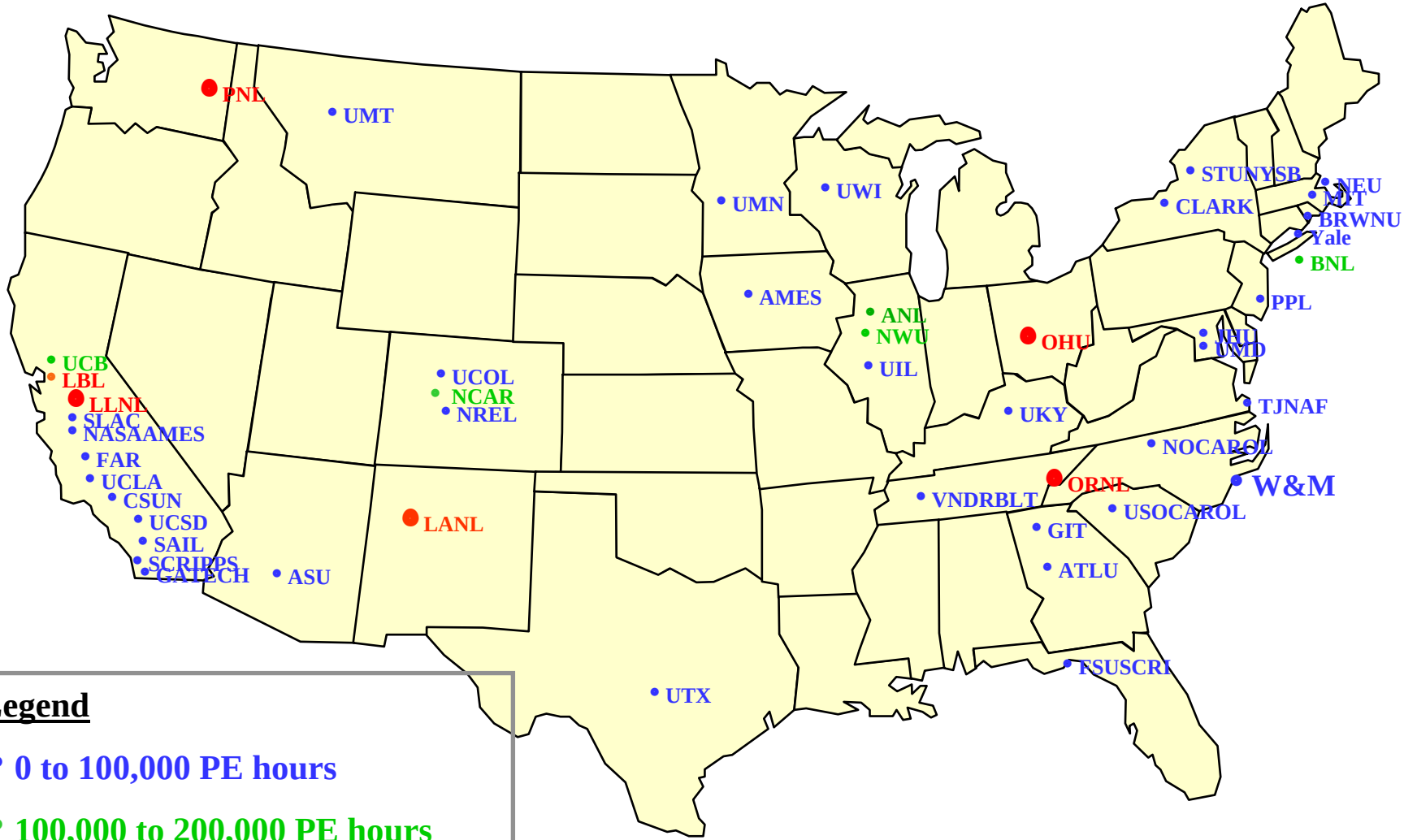







FDDI/
ETHERNET
100 Gigbit
REMOTE
VISUALIZATION
SERVER
SGI
SYMBOLIC
MANIPULATION
SERVER
MAX
STRAT
HPSS
IBM
And STK
Robots
DPSS
PDSF
HIPPI
ESnet
SUN
SUN
SUN
SUN
CRI T3E
900 644/256
IBM SP
NERSC-3
Processors
/604/304 Gigabit
memory
CRI J-90SE
VIS LAB
MILLENNIUM

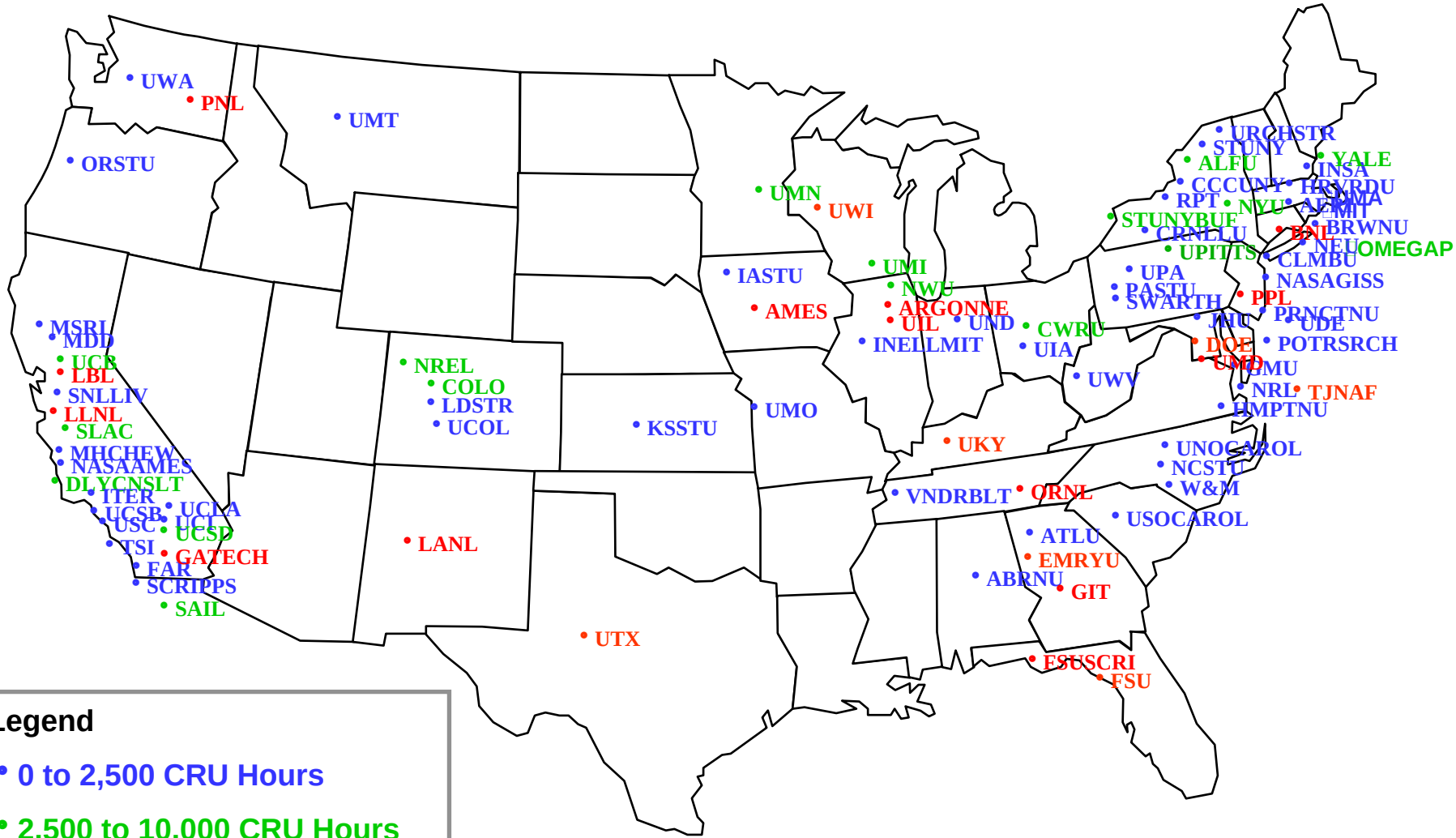
NERSC FY 99 Massive Parallel Processing Allocations by Site (MPP)



Legend

- 0 to 100,000 PE hours
- 100,000 to 200,000 PE hours
- 200,000+ PE hours

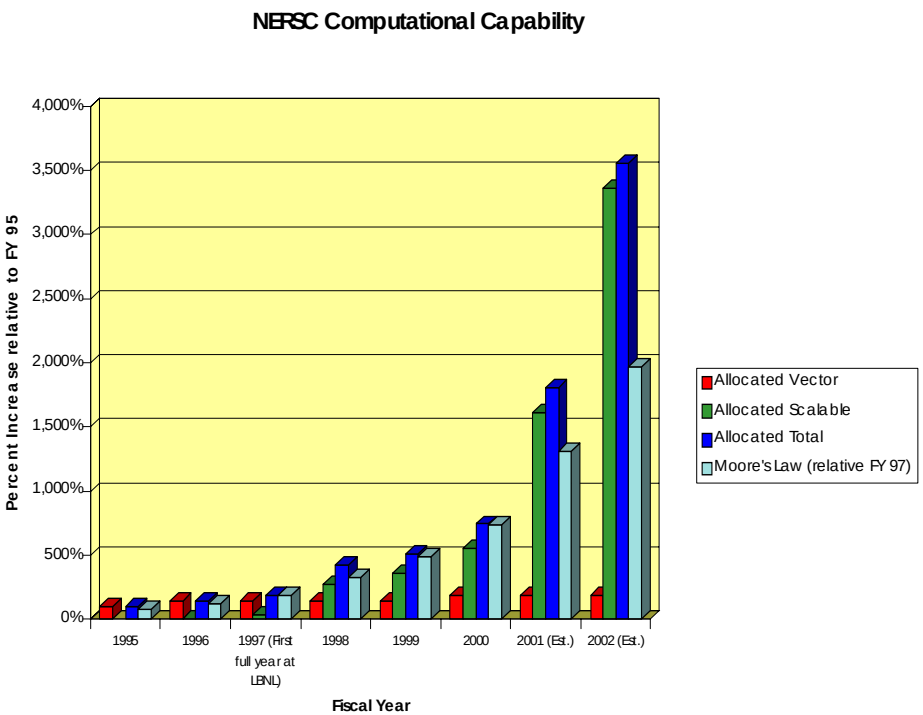
NERSC FY 99 Parallel Vector Systems Allocations by Site (PVP)



Legend

- 0 to 2,500 CRU Hours
- 2,500 to 10,000 CRU Hours
- 10,000 to 100,000 CRU hours

Why we are doing this



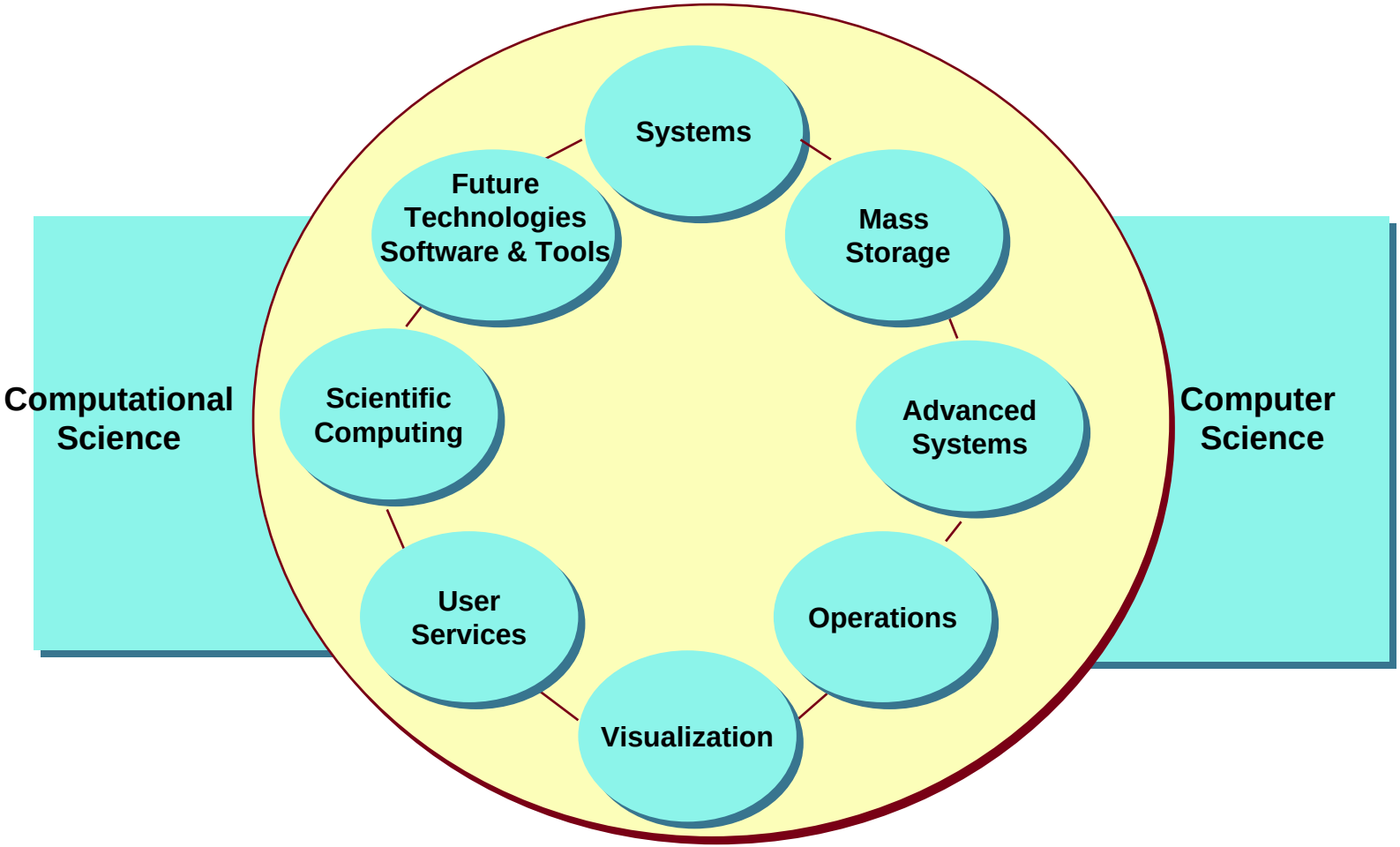
- Over the past 3 years, NERSC has completed a paradigm shift to production Massively Parallel computing
  - 1996 - 100% of production computing resources were vector SMP
  - 1999 - less than 25% of production computing resources are vector SMP
- The shift was completed in FY 00 when all major products had more MPP time than vector time.

To share our experiences
To Discuss the future

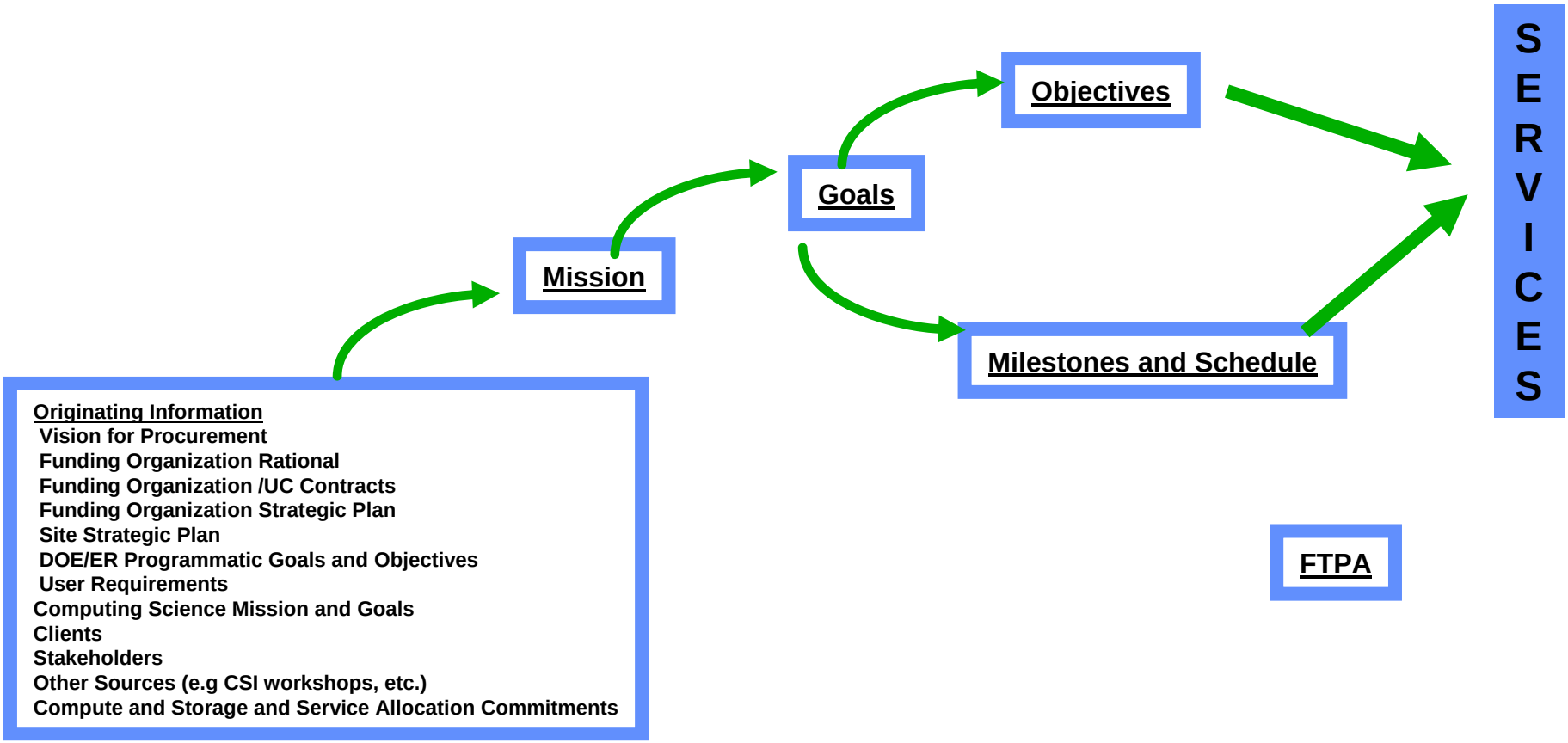
Comments on High Performance Computing Today

Service Architecture

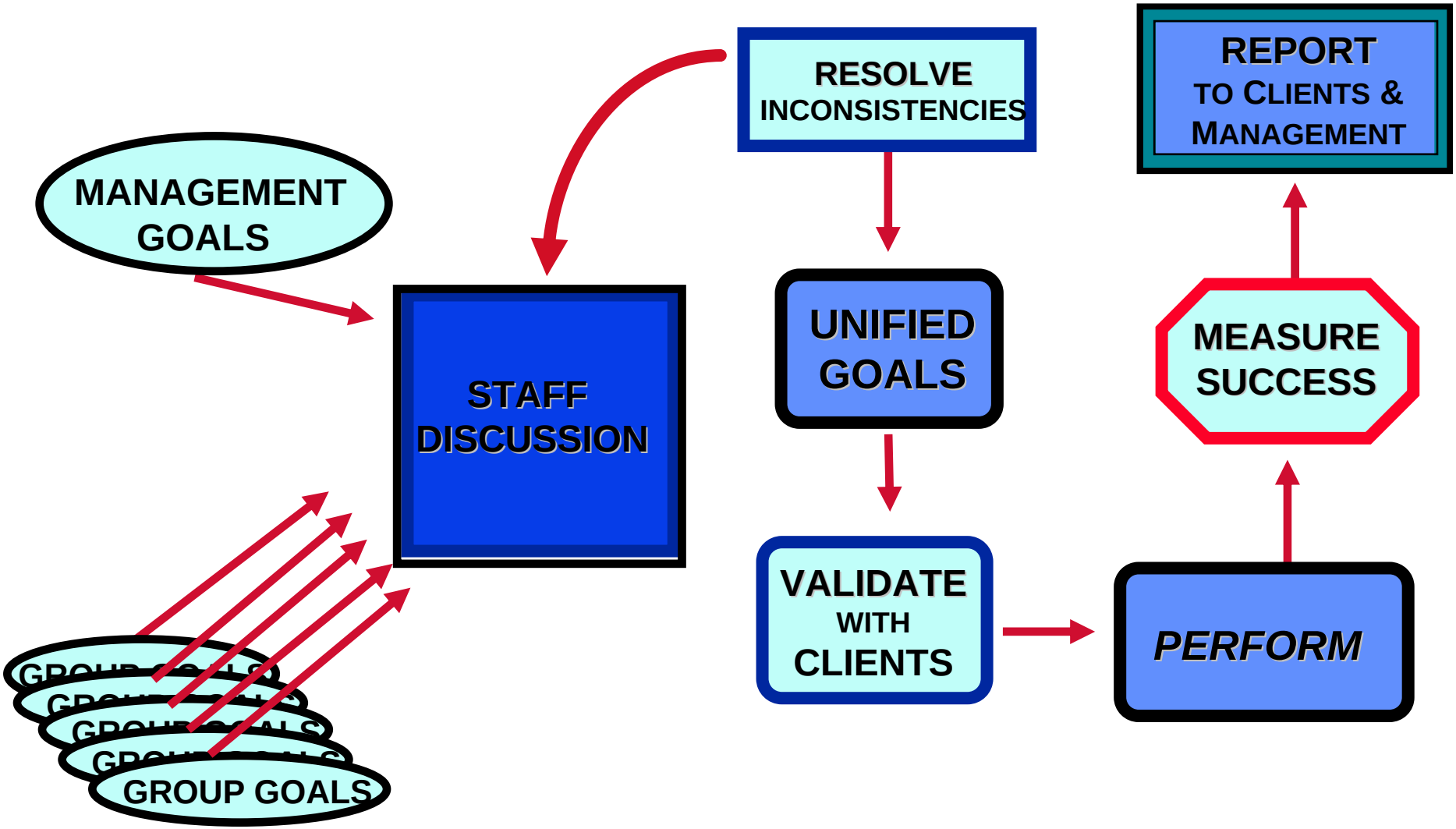
Intellectual Home of Large Scale Computational Facilities



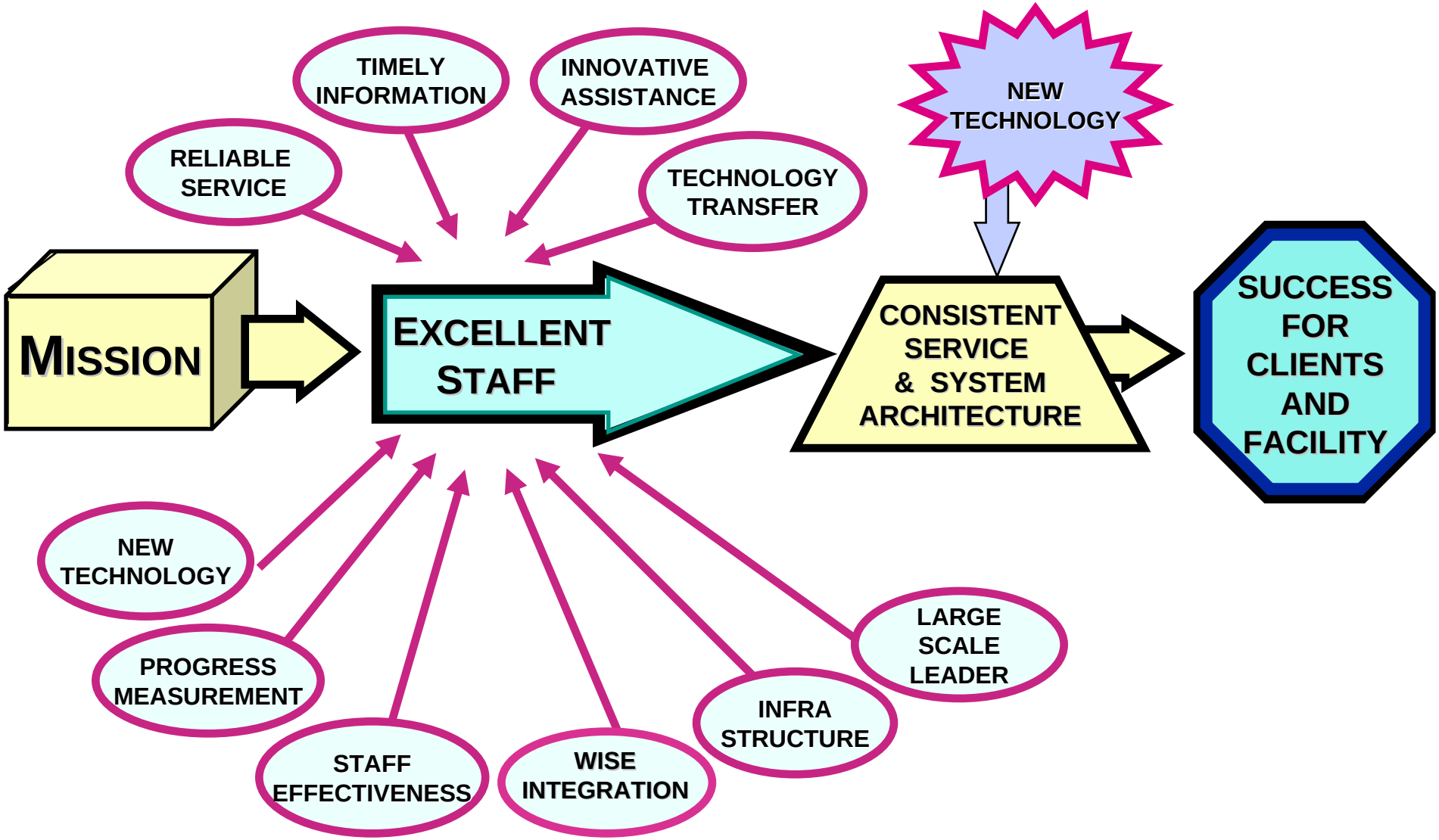
Originating Requirements



Goal Process



Goals


TIMELY INFORMATION
INNOVATIVE ASSISTANCE
RELIABLE SERVICE
TECHNOLOGY TRANSFER
NEW TECHNOLOGY
MISSION
EXCELLENT STAFF
CONSISTENT SERVICE & SYSTEM ARCHITECTURE
SUCCESS FOR CLIENTS AND FACILITY
NEW TECHNOLOGY
PROGRESS MEASUREMENT
STAFF EFFECTIVENESS
WISE INTEGRATION
INFRA STRUCTURE
LARGE SCALE LEADER

Reliable and Timely Service Examples

- <u>System Goals -</u> effective within 6 months after a computational and storage systems going into service
  - A service interruption is any event or failure (hardware, software, human, environment) that disrupts full service to the client base.
    - Any partial degradation (e.g., a node with some but not all CPUs out of service) is treated, for the sake of these goals, as a complete failure.
    - Any shutdown that has less than 24 hours notice is treated as an unscheduled interruption.
    - A service outage is the time from when computational processing halts to the restoration of computation (e.g. not when the system was booted, but rather when user jobs are recovered and restarted).
    - If an outage occurs within 2 hours of the system that does not have checkpoint/restart being restored to service, it is treated as one continuous outage.
    - If an outage occurs within 2 hours of the system being restored to service, it is treated as one continuous outage.

Gross Availability	95%
Schedule Availability	96%
Mean Time between Interruptions (hours)	120
Mean Time to Restoration of service (hours)	4

Reliable and Timely Service Examples

The end measure of a site is how much productive scientific work users accomplish. Sites must assist users in being as productive as possible by providing systems, tools, information, consulting and training. The objective is to understand codes and how they are used, and target bottlenecks for assistance and improvement.

- **A four tier service architecture ensures response to client problems within 4 working hours.**
  - **Resolve at least 95% of problems within two working days.**
  - **Escalate unresolved problems for review within 72 hours.**
  - **Install accounts within one working day (update activities regularly).**
- **Provide timely and accurate electronic information.**
  - **All system outages announced at least 24 hours in advance.**
  - **All planned system changes announced at least 7 days in advance.**
- **Improvements and Enhancements**
  - **Improve system usability by testing, working with vendors and monitoring and tuning performance and system parameters.**

Reliable and Timely Service Examples

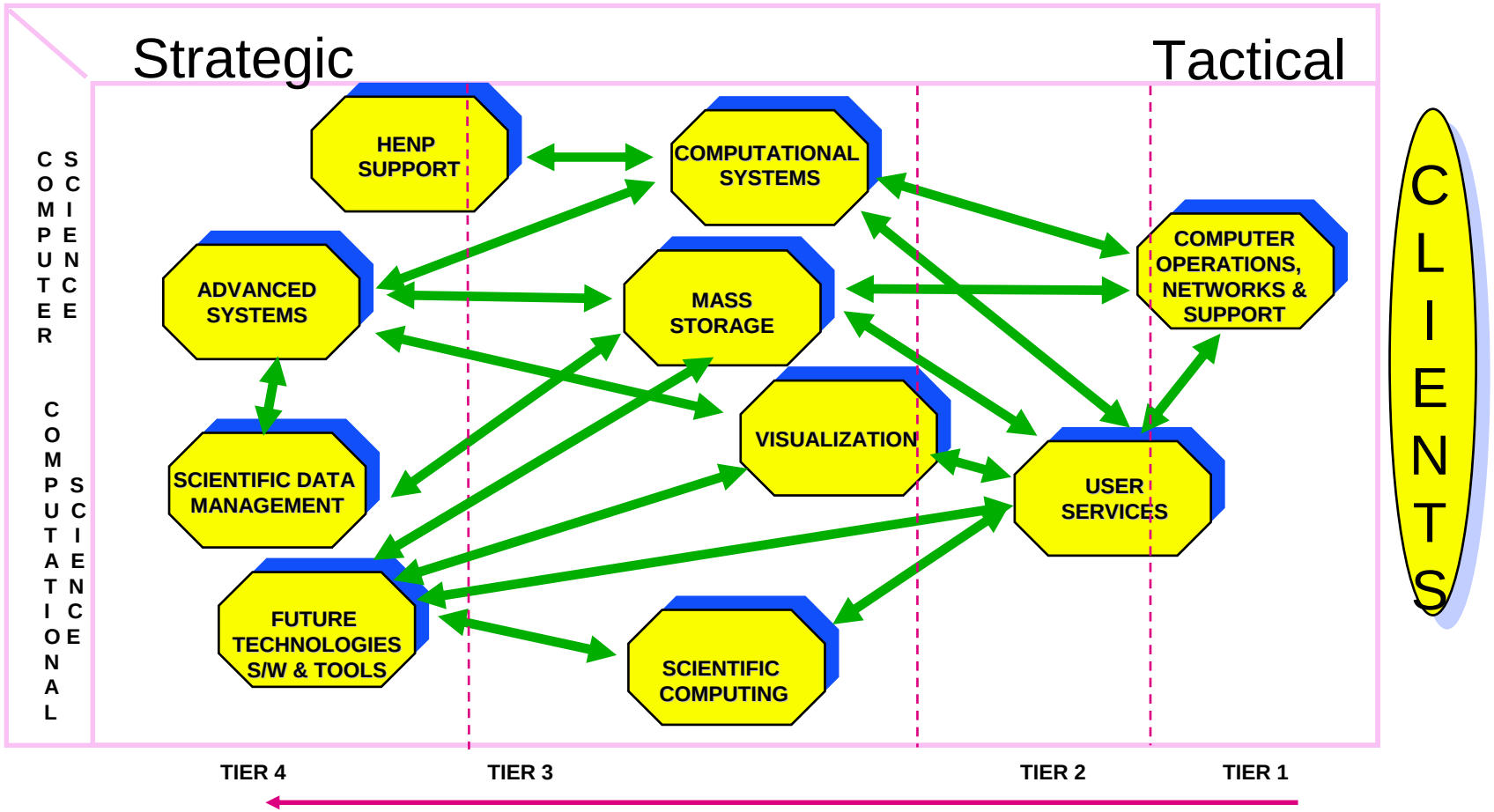
- **Test and Integration**
  - **Ensure all system changes improve (or at least do not diminish) service. To ensure system changes have minimal detrimental impact on users' ability to do work, the site will thoroughly test systems and software, fully inform the user community of the impact of changes/enhancements and create workable solutions for users when they exist. If detrimental impact is unavoidable, ensure the benefits of the change well outweigh the detriments.**
- **Metrics**
  - **Establish baseline metrics ("primary indicators") that measure the quality of service from the user's perspective.**
- **Security**
  - **Maintain a protected and secure environment consistent with open access and integration with remote sites. Assure quick response to problems in order to return systems and functions to service as well as providing the user the tools and functions necessary for them to protect their own information.**

Example: Scalable Computational Science Competency

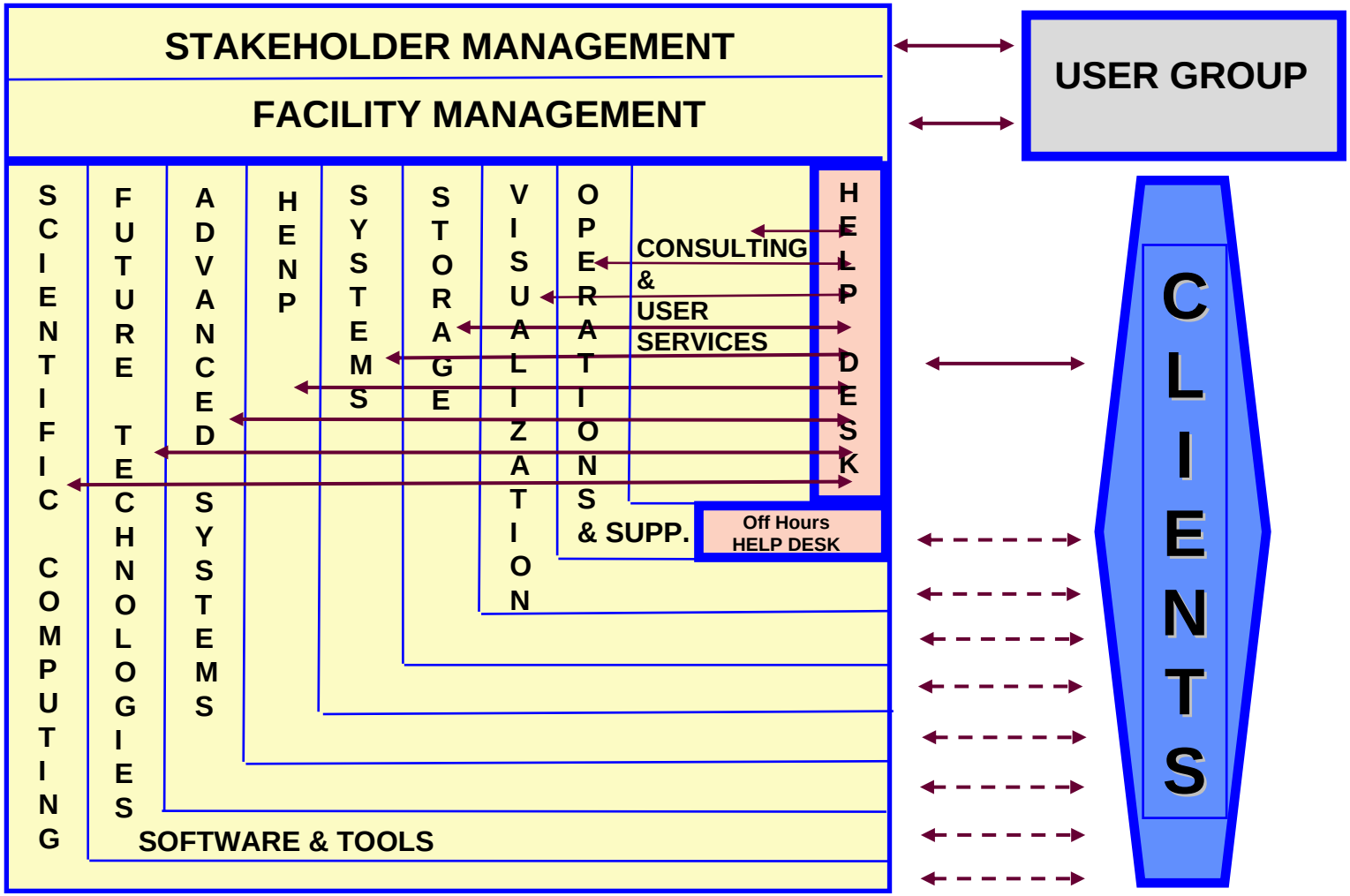
ER	Maps to	Discipline	Maps to	Computational technology
Fusion		magnetic fusion		particle in cell
Basic Sciences		computational chemistry material sciences		local density functional
Earth Sciences		climate research computational biology		partial diff. equations
Physics		QCD accelerator design particle detection simulation		Monte Carlo technique searching, pattern matching
Computer Sciences		combustion applied mathematics		image processing

NERSC has or will build competency in all technological areas of relevance to ER research

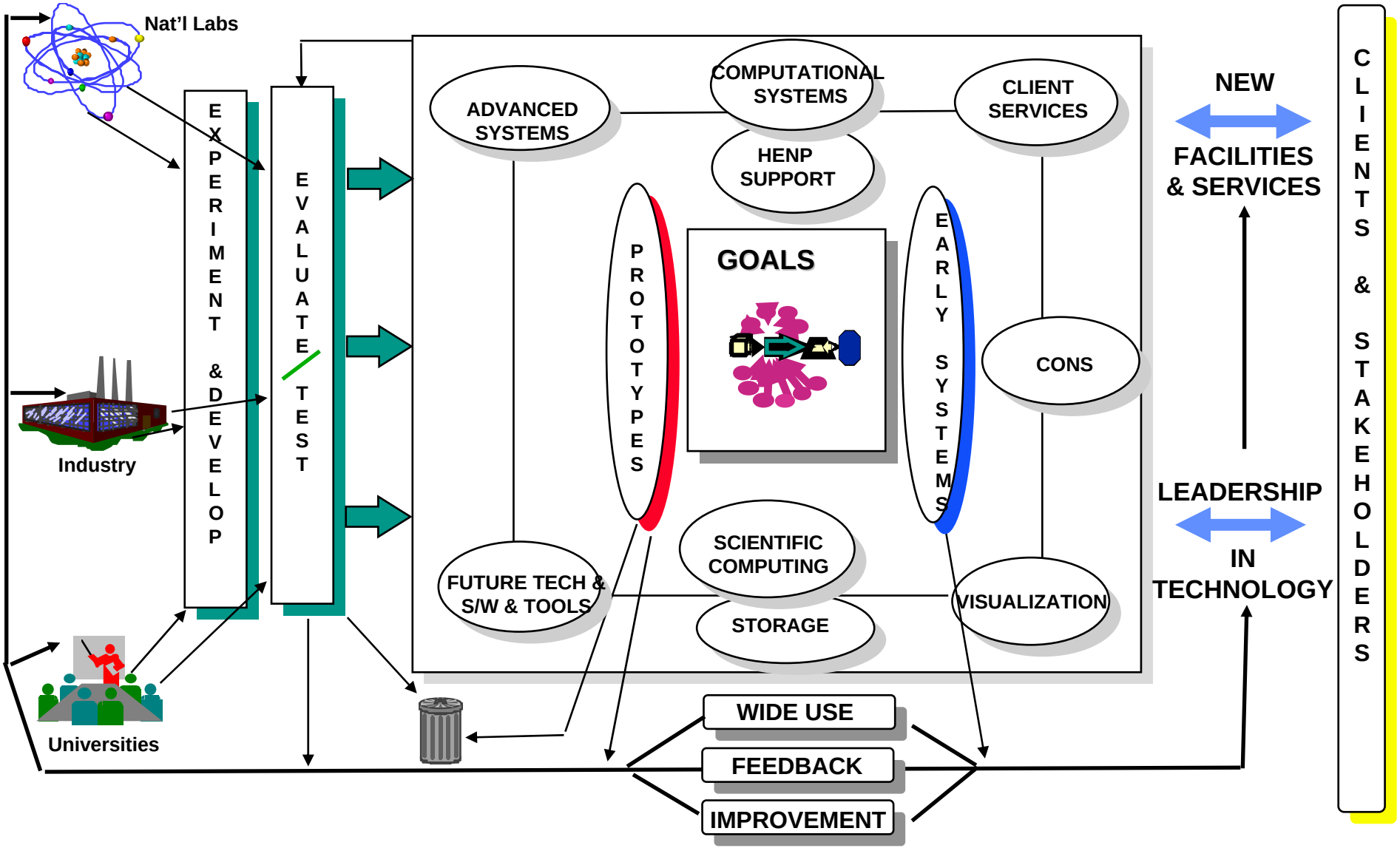
Major Communication Paths & Levels of Support



Service Architecture

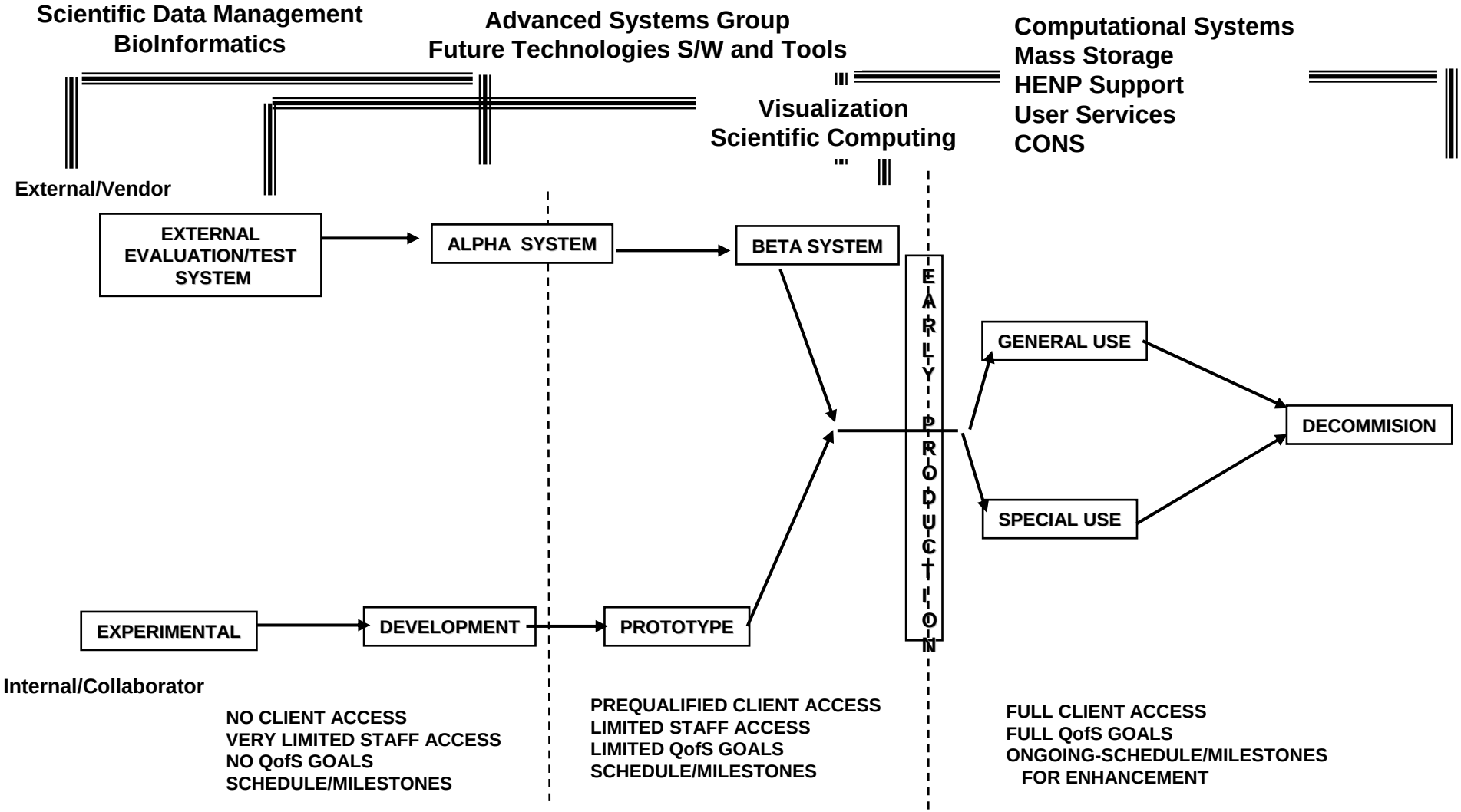


New Technology Integration



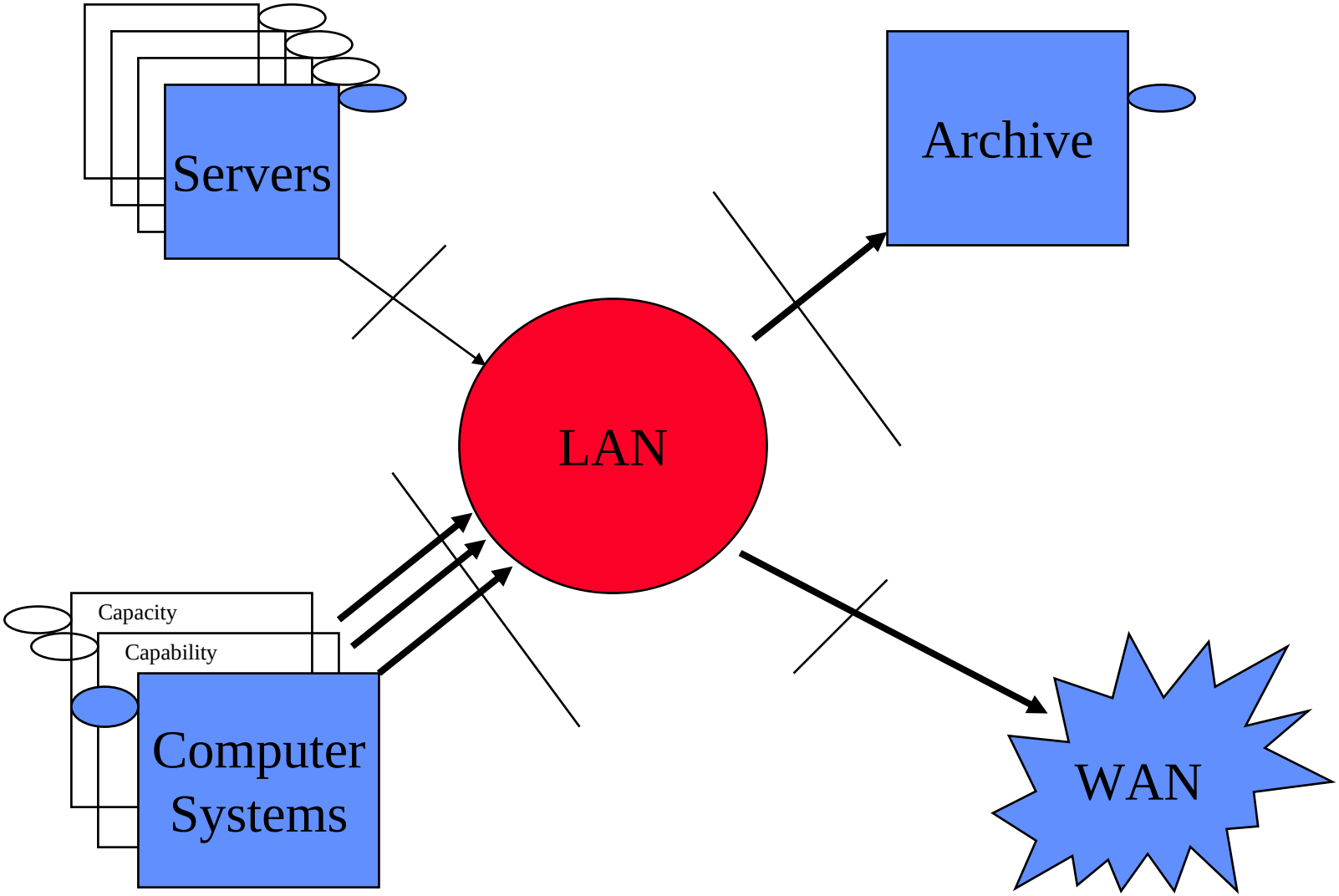
H/W and S/W Life Cycle

Technology can enter at any phase

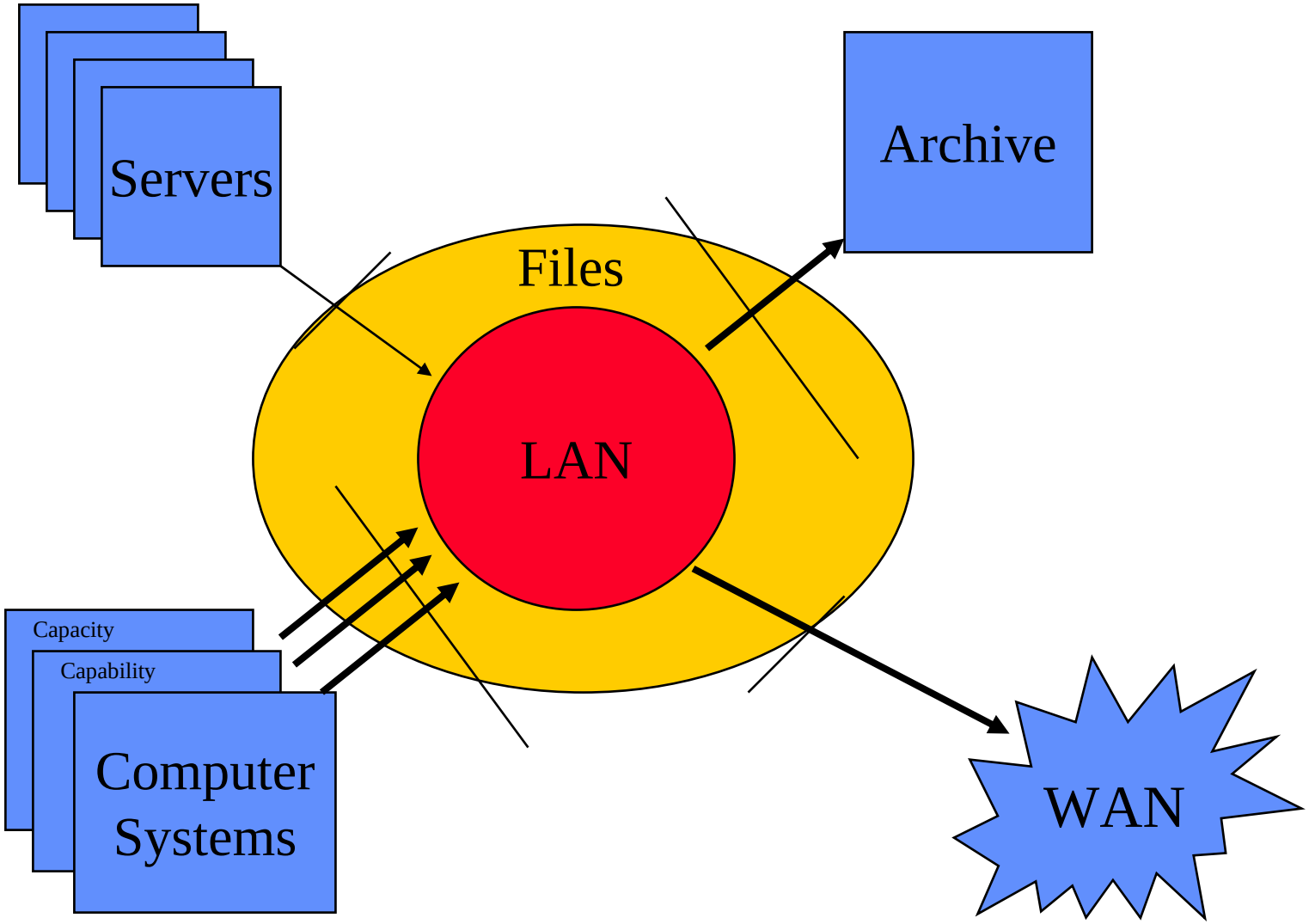


EXPERIMENTAL SYSTEM MATURITY **MATURE**

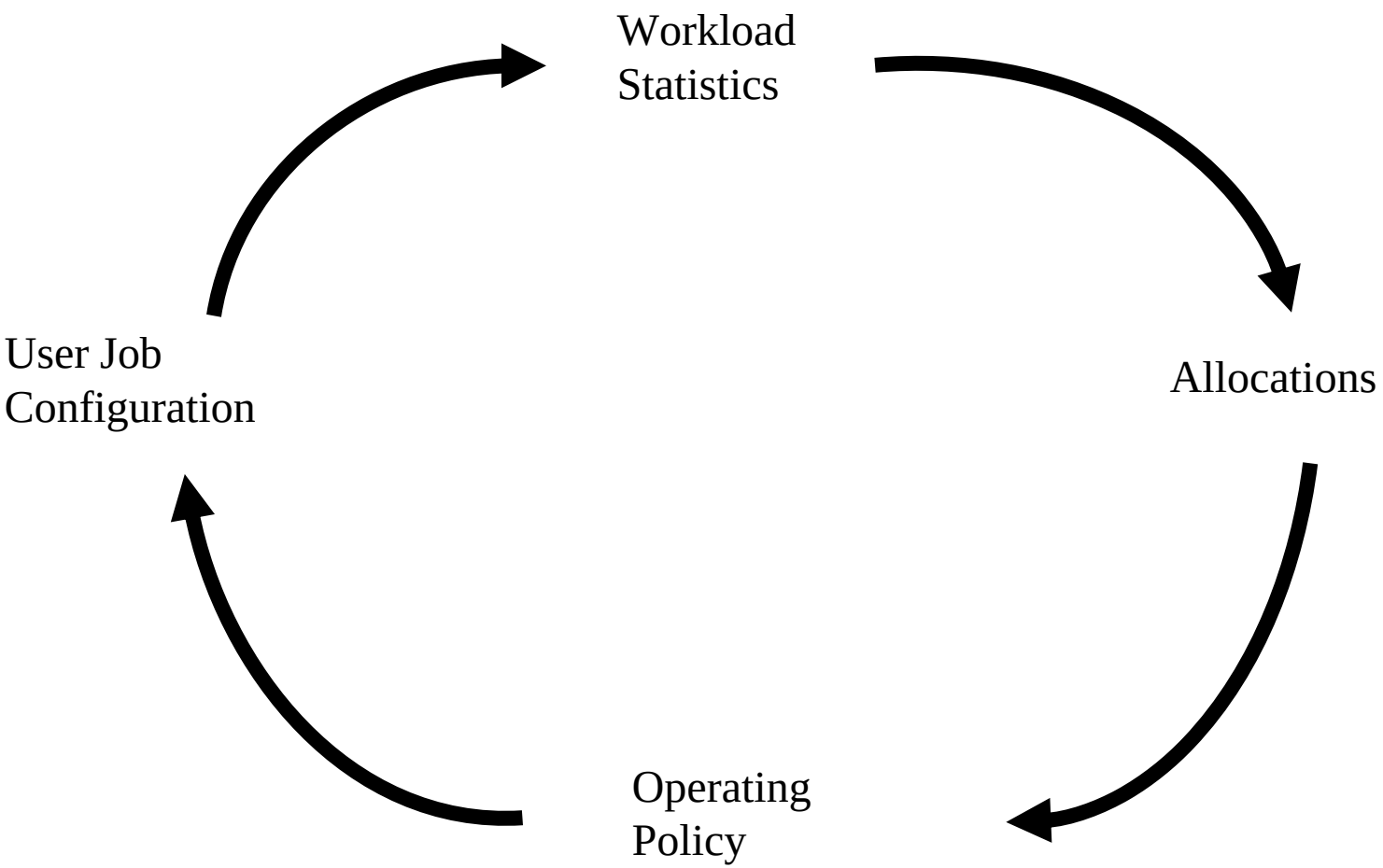
Traditional Architecture



Desired Architecture



Behavior Impacts Workload



Allocations and User Behavior

- Allocations play a key role in how well systems are utilized
  - Types
    - Straight annual allocations
    - Over allocation by fixed percents
    - User selected priority
    - Quarterly expiring allocations
    - Combinations of the above
  - Users behavior impacts utilization
    - e.g. System is able to process jobs but no one submits jobs

Observations

- When there are no limits on individual/group usage
  - Users respond to turnaround shifts by reconfiguring jobs
  - Very difficult to prioritize usage
    - System Managers and users
- When there is very high over-allocation
  - Similar to no limits
  - Adds user disappointment of not getting what was promised
- When allocations are actual or close to actual (<75% over allocation)
  - Users respond to hoarding and gorging behavior
  - User selectable priorities/charging essentially provide over allocation
    - Priorities that impact only CPU scheduling are of limited value
    - The lemming syndrome - a few users raise priority, then all do to maintain turnaround

Comments on High Performance Computing in the Future

Input

- Three major areas of input
  - NERSC-III Procurement
    - FY99/00 addition
    - 4-5 times T3E
    - 1/6-1/3 sustained Tflop-yr average over 3 years
  - Strategic Simulation Program
    - Major initiative
      - *DOE and other agencies*
    - FY 00 at 10 Tflops Peak
    - FY 03/04 at 80 Tflops Peak
    - About $500 M/yr
  - General tracking

System Summary

- There are commodity and custom processors
  - MERCED and follow-ons (Intel)
  - Alpha (Digital)
  - Ultra Sparc (SUN)
  - Power 3 (IBM)
  - R12000 (SGI-O2000)
  - SGI/CRI Proprietary (SN2v)
  - Tera

Vendor Summary

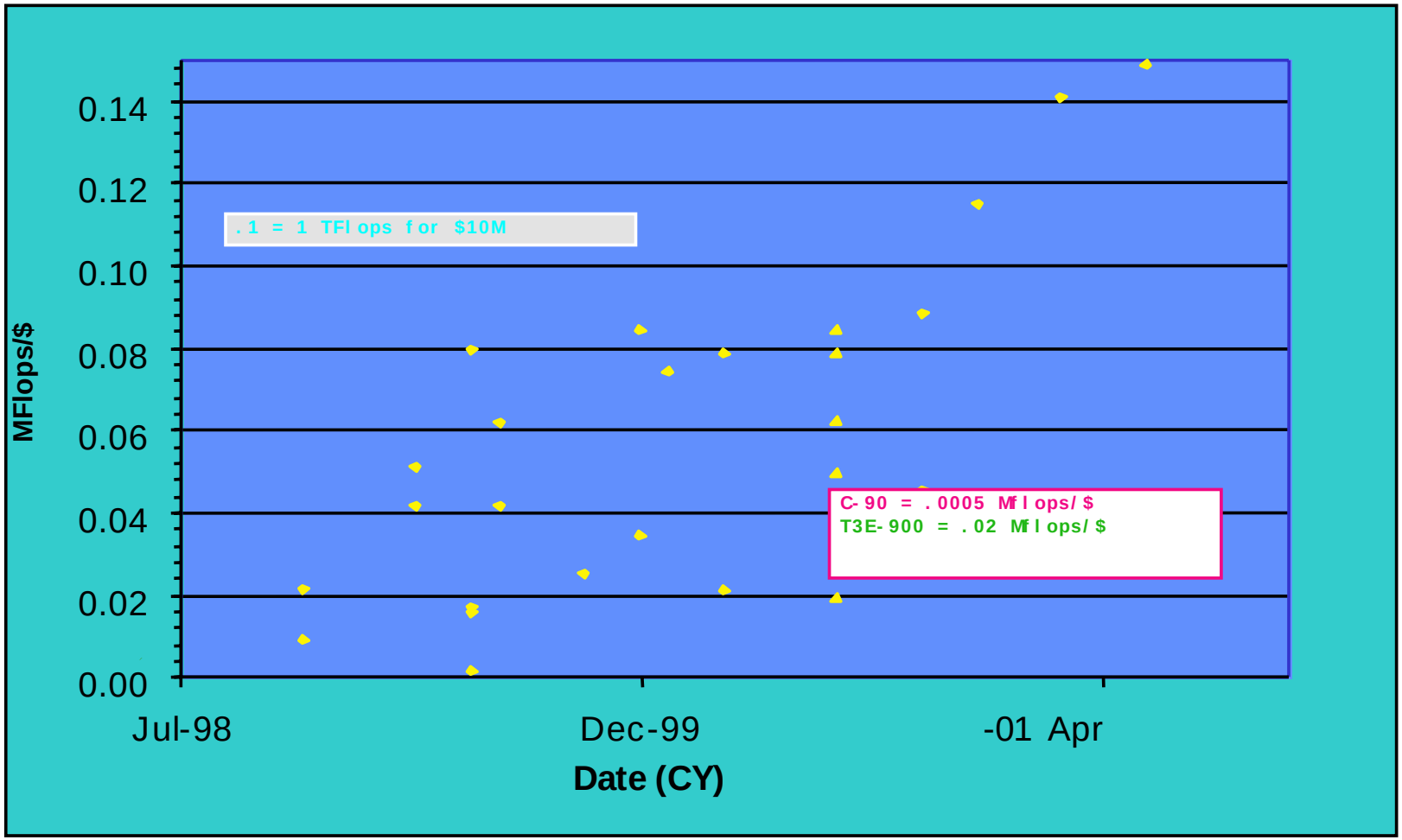
- **Parameters for Computational system only**
- **SS*-1 - Computational System Estimates**
  - **10 TFlops, FY 00**

	Memory Bytes/Flop	Disk Bytes/Flop	Cost/Flop $	Size SF	Power MW
Minimum	**.05**	**5**	**11.80**	**4,096**	**1.1**
Average	**.23**	**9**	**22.14**	**8,264**	**2.5**
Maximum	**1.00**	**15**	**50.00**	**10,716**	**5.3**

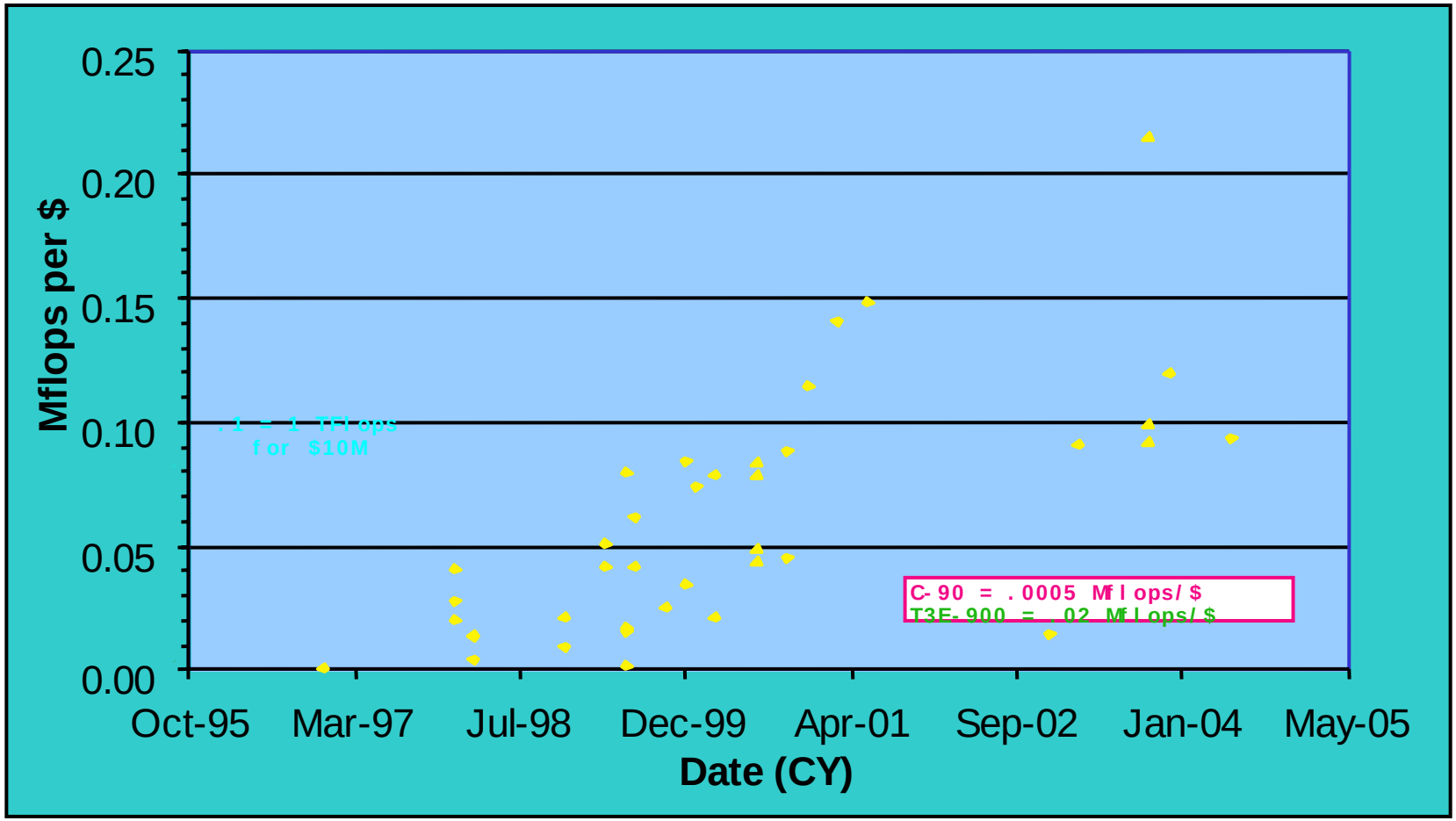
- **SS*-2 - Computational System Estimate**
  - **40 TFlops, FY 04**

	Memory Bytes/Flop	Disk Bytes/Flop	Cost/Flop $	Size SF	Power MW
Minimum	**.10**	**5**	**3.51**	**4,096**	**1.1**
Average	**.26**	**9**	**8.81**	**11,575**	**4.0**
Maximum	**1.03**	**15**	**20.00**	**15,115**	**6.0**

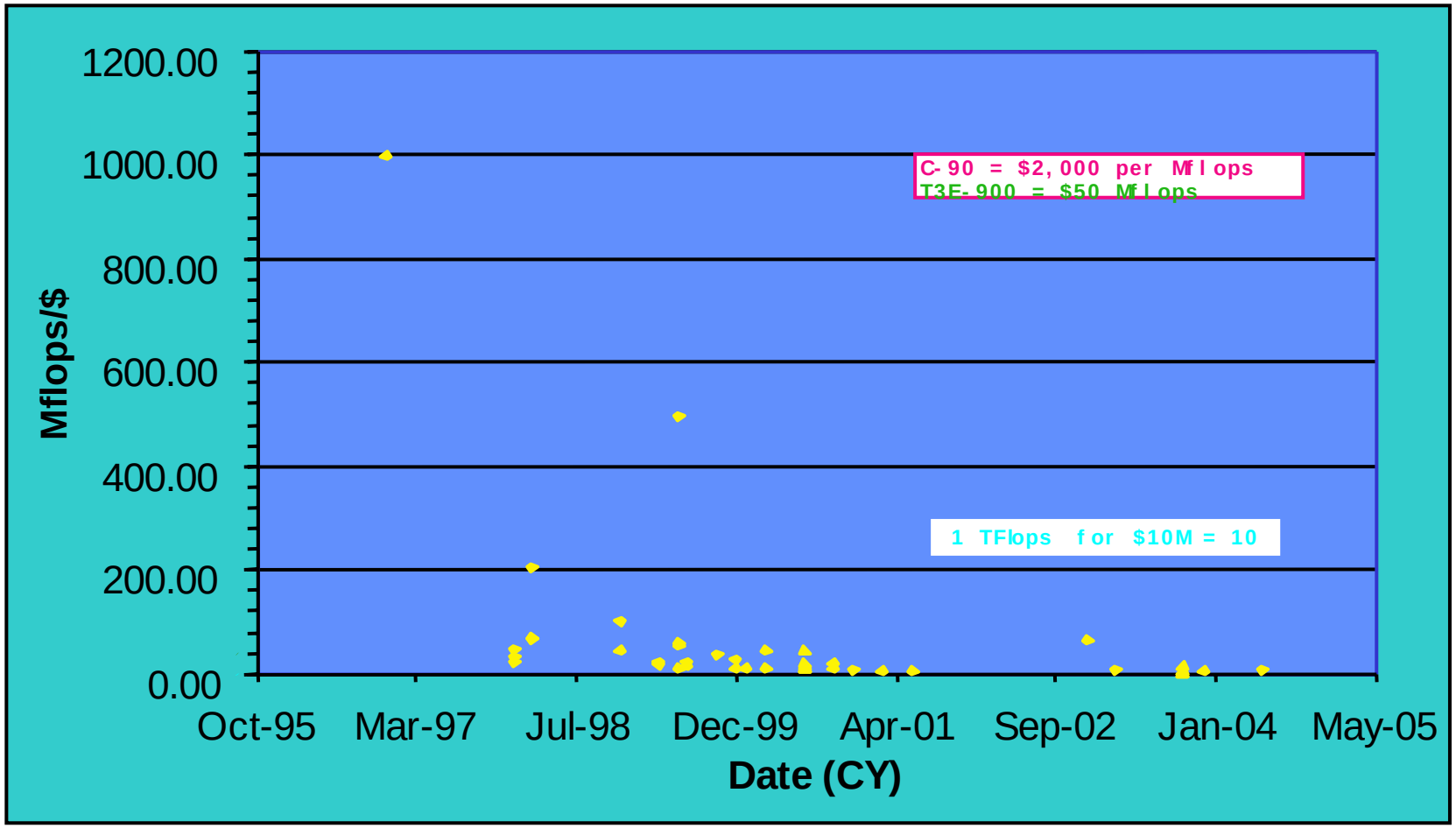
Large Scale System Costs



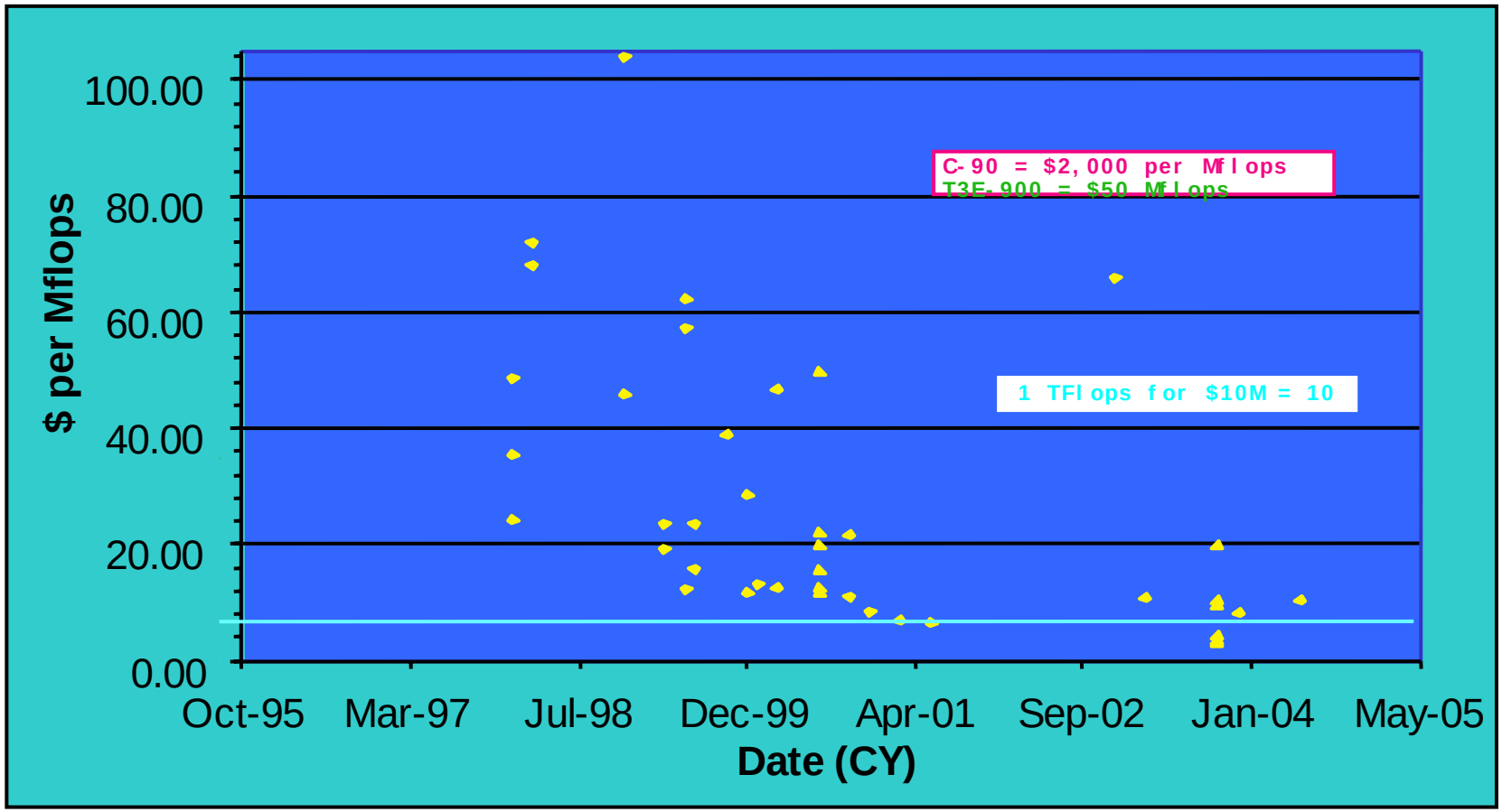
Large Scale System Costs



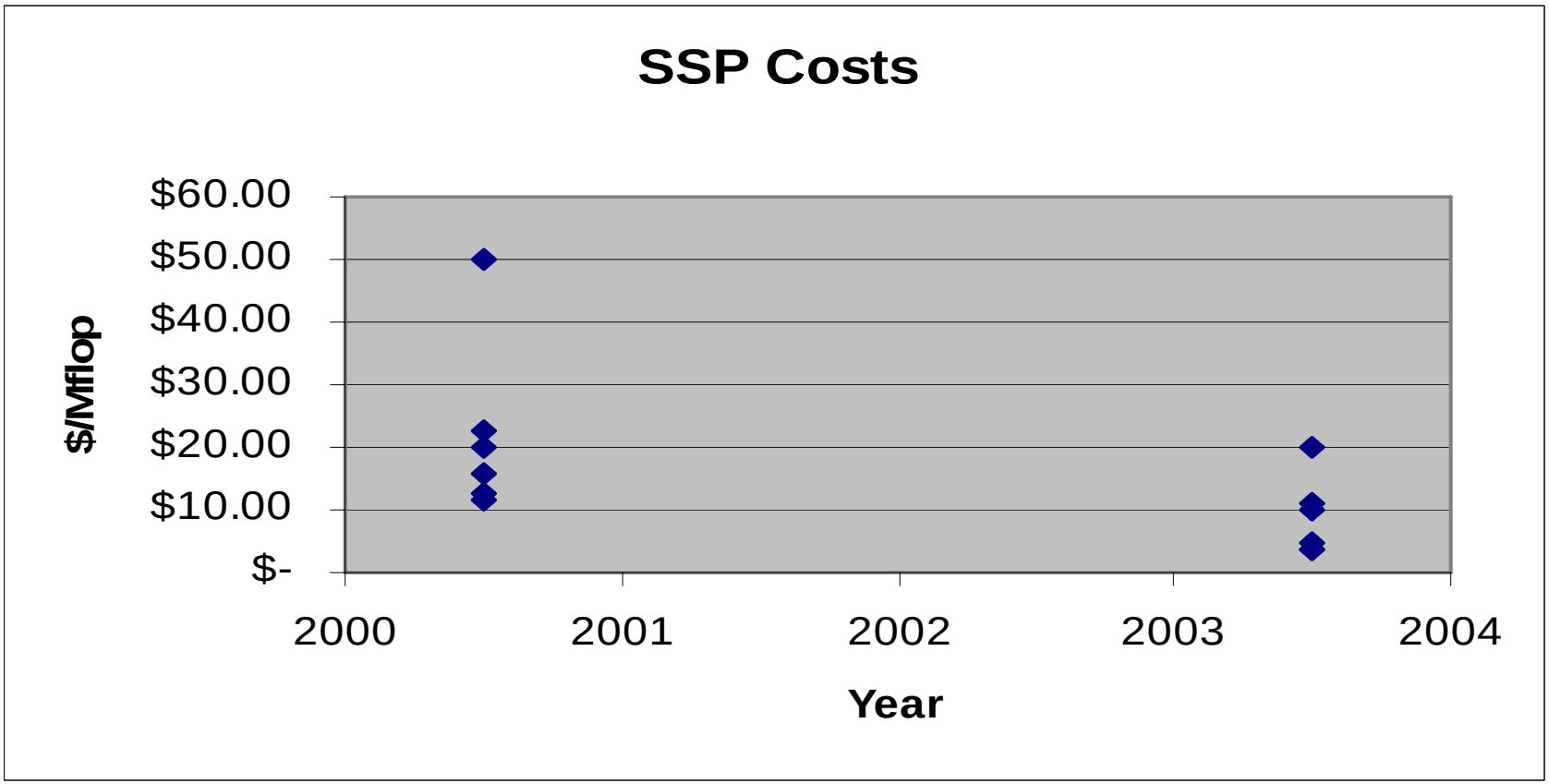
Large Scale System Costs



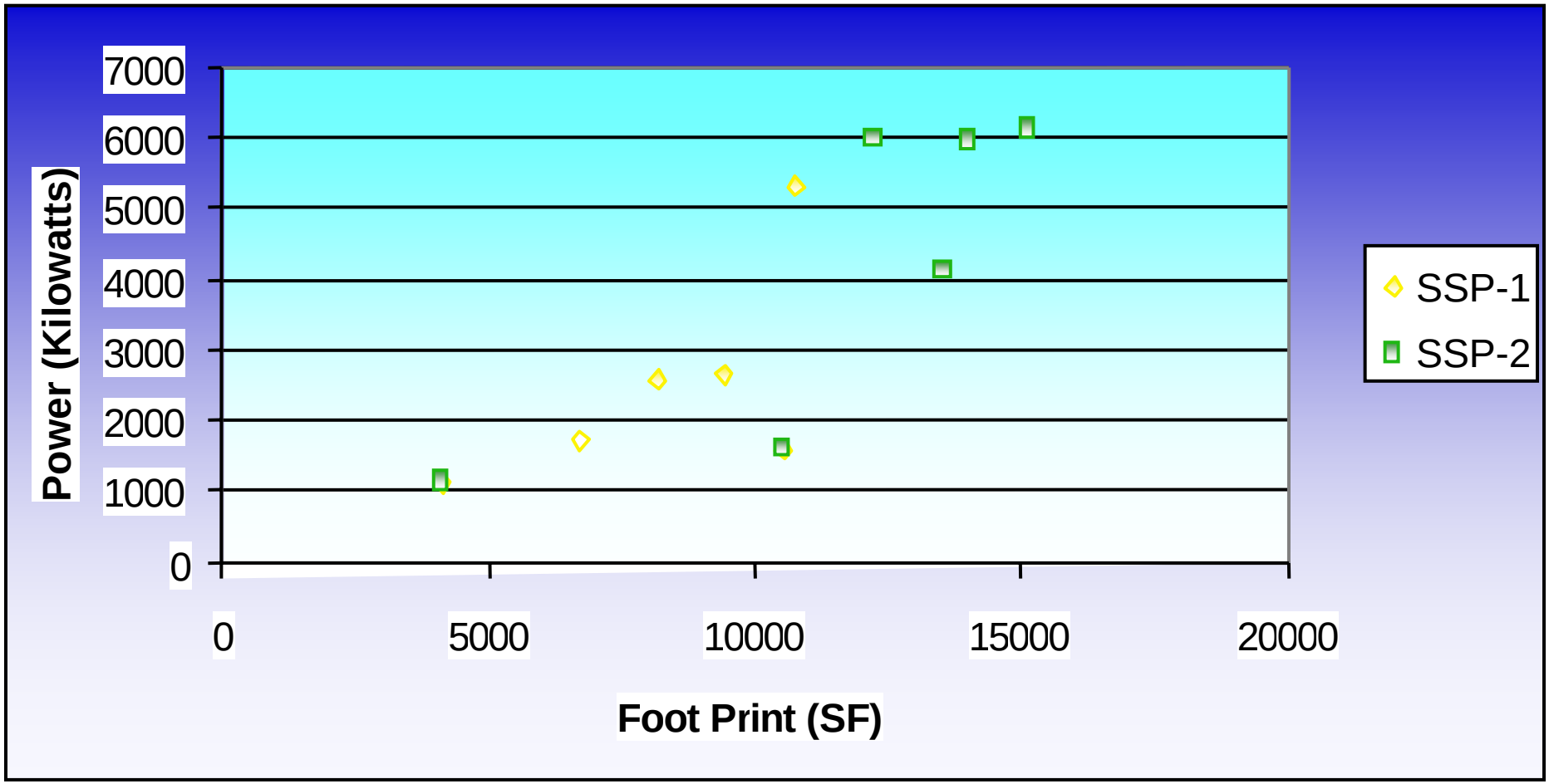
Large Scale System Costs



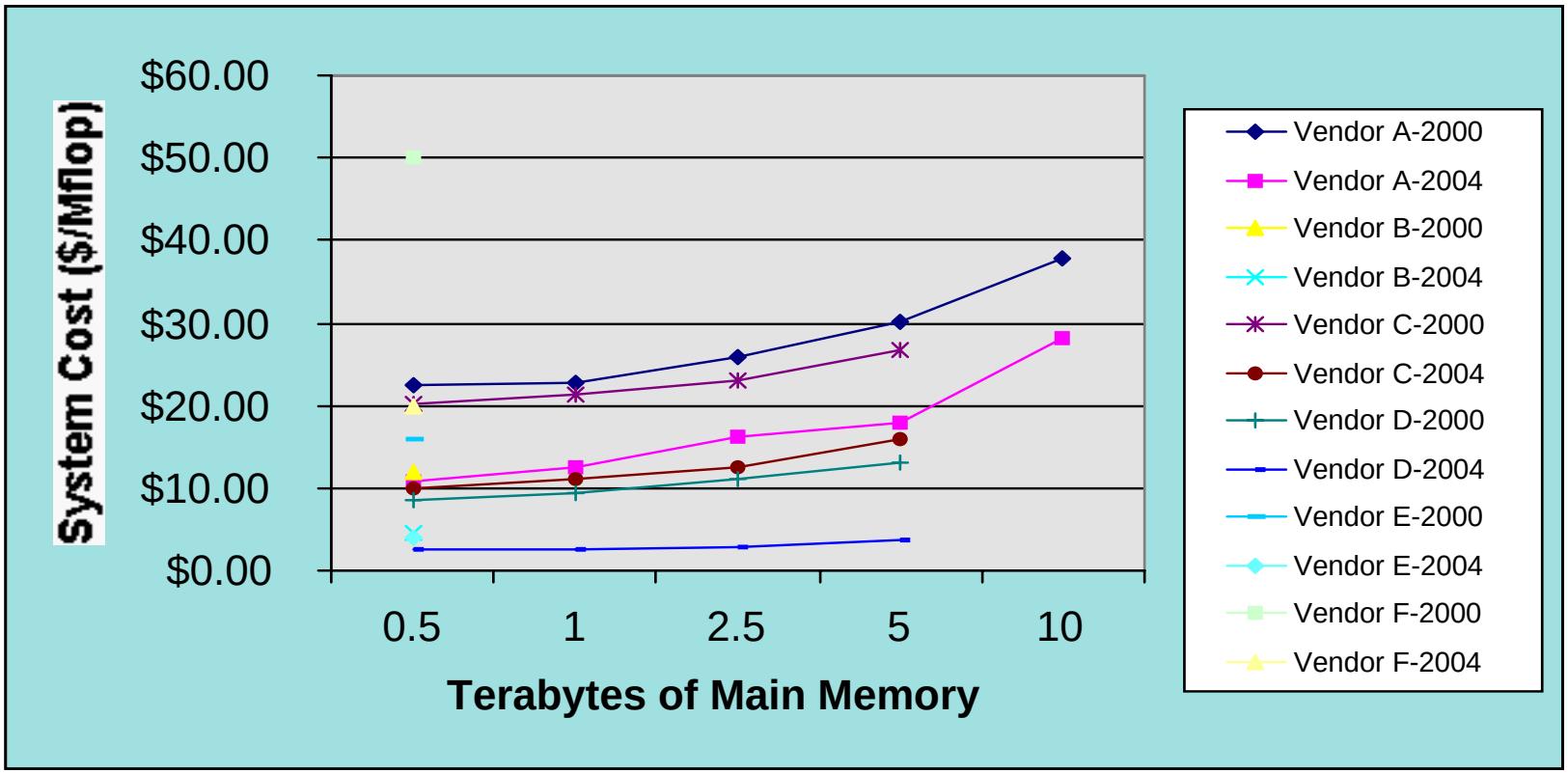
Longer Term Price/Peak Performance



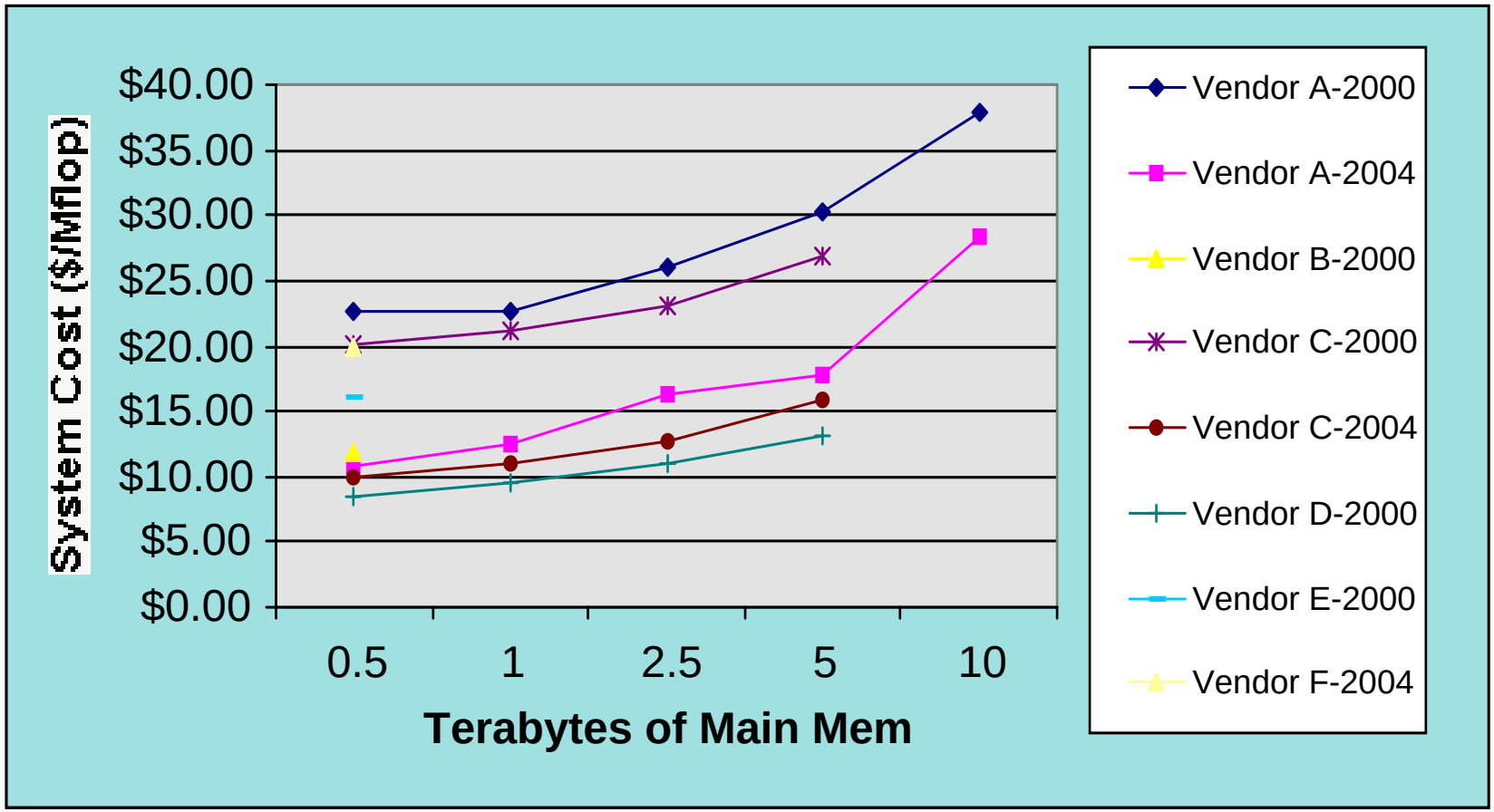
Large Scale System Facilities



Influence of Memory Costs



Influence of Memory Costs



The Bottom Line

- The Good News
  - The various architectures converge into one basic model
    - No more MIMD/SIMD
    - No more Vector vs MPP
    - No more SMP vs Distributed Memory
- The Bad News
  - The range of options in the architecture is so broad, and with so much complexity, we may wish for the old days

General Conclusions

- FY 00
  - $20-25/Mflops
  - $6/Mbyte of Memory
  - $.3-.2/Mbyte of disk (high performance)
- FY 03/04
  - $10-11/Mflop
  - $1.5/Mbyte of Memory
  - $.1-.05/Mbyte of disk (high performance)

Trends

- All systems coherent, non-uniform distributed shared memory
  - SMPs of some type
- To reach Teraflops performance sustained, there needs to be the order of 1,000 processors
  - To reach 10 - 80 Teraflops of performance needs on the order of 10,000 processors
- Large scale systems will be made up of clusters of SMPs
  - Smallest SMP is 16 processors
  - Largest is 1024 processors in hardware
    - Software may only scale to order 100 processors
- All systems will have multiple single system images

Trends

- No system will have a uniform programming environment
- To do large scale problems, applications will have to deal with hierarchical distributed memory programming (MPI, OpenMP, etc.)
- CPU costs are becoming proportionally less
  - Large scale systems may have memory and CPU cost the same
- Complexity of the systems increases dramatically
  - System Software
  - Failure Modes
  - Application Software
  - Interconnections

Trends

- There is no dominant vendor (we hope)
  - High performance simulation market is place too small for all vendors.
  - At least two of the potential vendors are driven by transaction processing, not computation limited applications

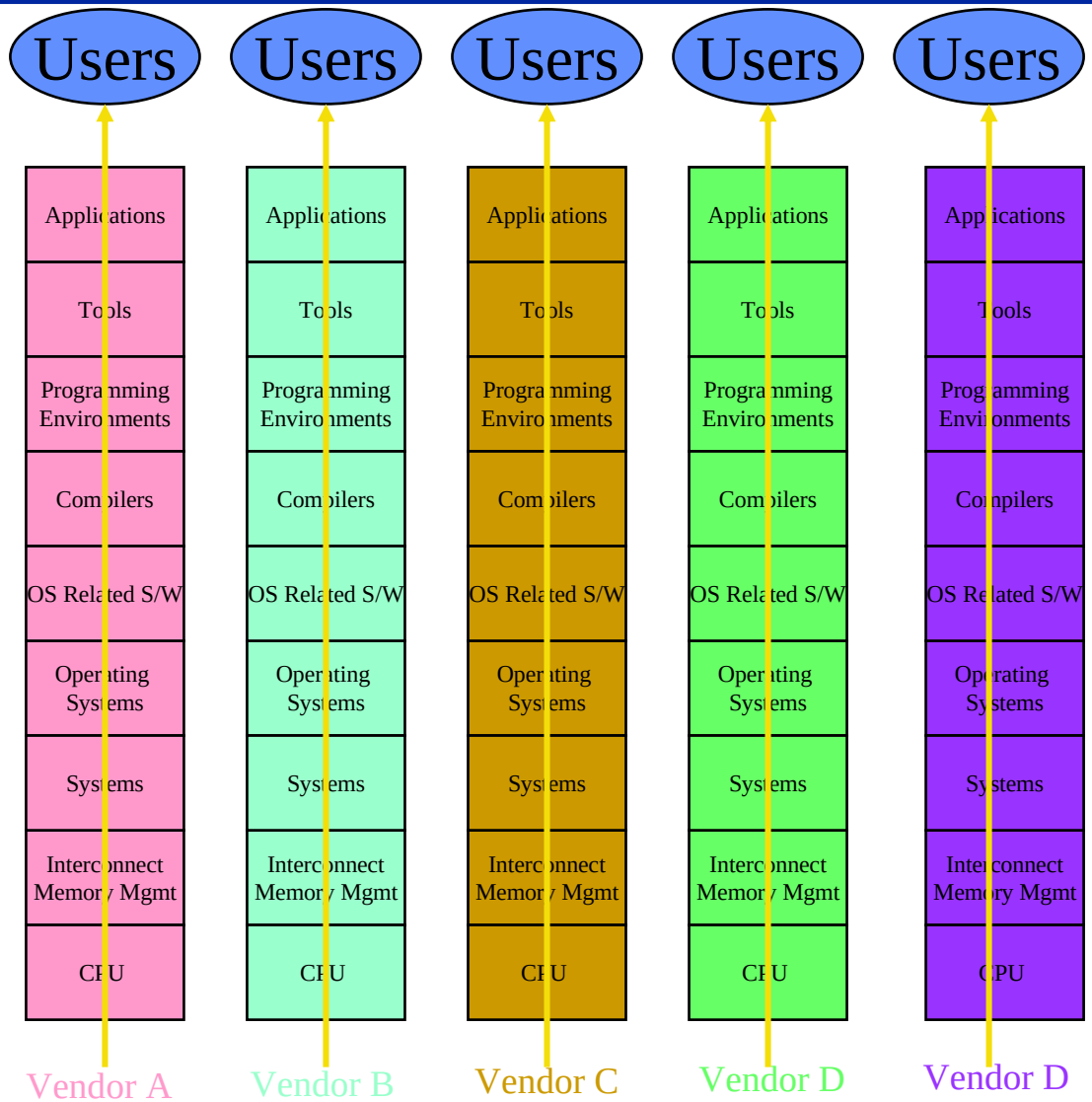
- Vendors will still supply systems for the next 5 years
  - PC clusters have a way to go -
    - scale to 10-1,000 systems in the next 3 years but not beyond
    - Limited to restricted applications
    - After 5 years, may be a major force in large scale computing

Trends

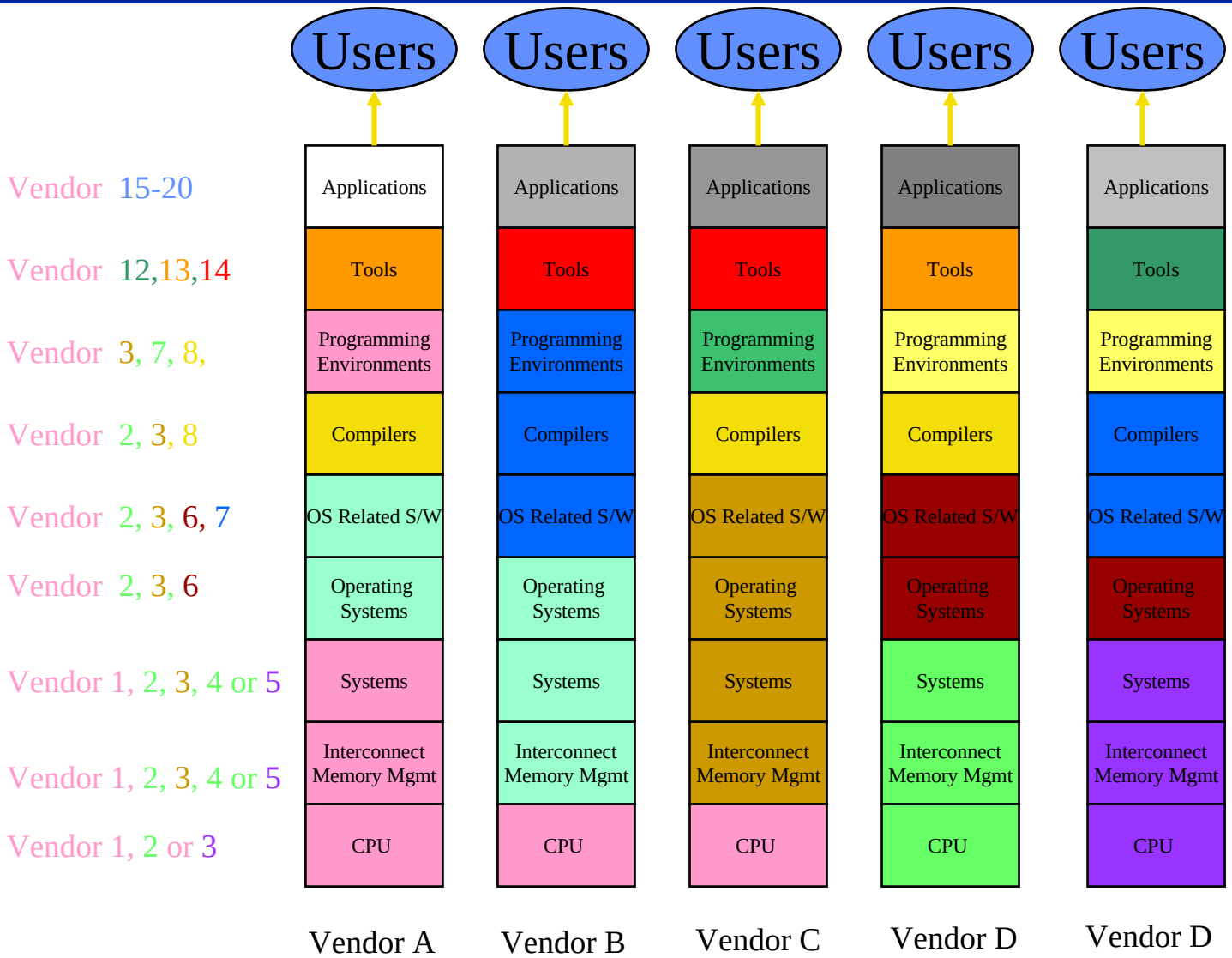
- Industry moving from vertical integration to horizontal integration
  - Some computer vendors will be integrators
    - Commodity CPU and memory
    - Custom interconnect
    - Custom software
  - Layered products may not be as integrated as we would like

- People have to expect more vendor changes
  - 4 generations(?) longest I ever dealt with (Cray-1, Cray-2, YMP, C90)
  - Likely that each new system will be a different vendor/architecture

Vertical Integration



Horizontal Integration



Likely Environment

- All processors are in the 1 to 2 Gigaflops range in 2000/2001
- SMPs clustered into massively parallel systems
  - SMP choices will range from 16 to 1,024 processors
  - Interconnections will range from very tight to moderately loose
- 1,000s of processors - some systems will be 10,000s processors to reach 40 Tflops in 2003
- System and application software complexity a significant challenge
- More tightly integrated with other components
- Resource management beyond CPU cycles will not exist but will be needed *CAN BE DONE - AND* ***DONE WELL***

Archive Storage Trends

- Capacity is manageable
  - Density increasing
  - Format remains the same

- Bandwidth the biggest issues
  - Aggregate to and from the systems
  - Real time peaks (on-demand)
  - Individual drive speeds
    - Mitigation strategies - (e.g. tape striping) are hard
    - Seek times for data on very large tapes
    - Assumption that archives will have only very large files is not valid

Storage Trends

- Web and Data Base not well integrated into archive systems
  - Archive systems oriented to large files
  - Web is asynchronous, immediate access
  - Data Base is record oriented
- Scientific Data management needs to be applied
  - Data Use
  - Data Organization
- Data Intensive computing needs to be enabled

Archive Storage Trends

- Industry will be driven by the commercial market place
- Scientific demands remain larger/more difficult than commercial, but market share small
- Scientific sites must add value to what is commercially available
- Solutions must scale
  - To balance with computation
  - To be cost effective
- Introduction of solutions will span longer time periods
  - Technology introduced with bigger increases in price performance but with longer time periods

What must a facility do to prosper in this environment

Create a System and Service Architecture that endures

- Decide on Focus
- Create synergistic System and Service Architectures
- Attributes of an Architecture
  - Functional Definitions
  - Interface Definitions
  - Processing/Service Rates
  - Measurement
- Design Considerations
  - Meeting Client Requirements
    - Access, performance, usability, assistance
  - Balance
  - Feasibility
  - Cost
  - Rate of Change

Focus

- Single Function vs multi function
- Capability vs Capacity
- Lowest cost vs best value

Challenges

Data Throughput

Data Locality

Complexity

Using incorrect Metrics

Resource Scheduling

Data Throughput

- Example - HiPPI
  - At NERSC
    - 14 HiPPI connections from compute systems
    - 8 HiPPI connections to Max Strat Raid Disk
    - 14 HiPPI connections to HPSS Servers
  - Performance per link
    - Peak 100 MB/s
    - IPI-3 Third Party from C-90 to Disk 60 MB/s
    - TCP/IP - Cray Memory 50-60 MB/s
    - Compute to storage servers (mem) 14 MB/s
    - Compute to storage disk 4-5 MB/s
    - User experience 1-2 MB/s

- DPSS - .75 MB/s workstation to disk

Other Data Throughput Experience

- Good News
  - HPSS improved past bandwidth by 10-20 times
  - Potential exists for more
  - Other technology may impact this further

- Bad News
  - IP bottlenecks
    - IPI-3 Third party not implemented by vendors
      - Not general
      - Example of modular approach vs performance
      - Still working on user level
  - Differentiated service is not likely to impact this much

Data Locality

- Network attached disks still a challenge
  - Max Strat is very expensive and no client software
  - Some solutions for other types of markets
  - Scalability problematic
  - Generality a problem if performance is needed

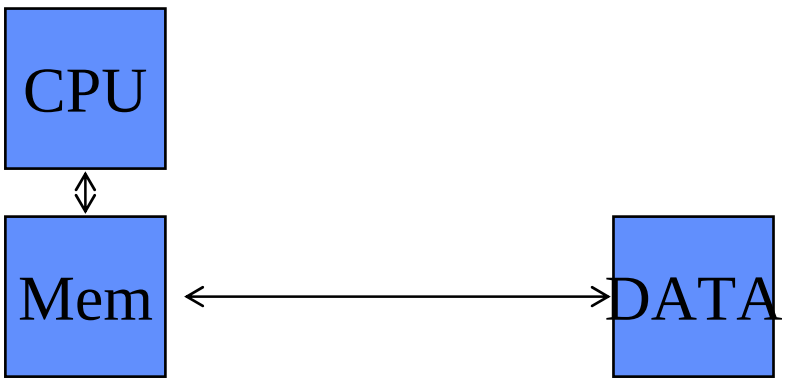
- Distributed file systems can not perform at the rates needed for High End Systems
  - NFS, DFS, AFS
    - General solutions
    - Performance is a serious limitation
    - Scalability is a serious issue

Memory Scheduling

- Last major advance was Virtual Memory
  - Memory scheduling is not integrated with CPU scheduling
  - Memory hierarchy is very complex and will increase in complexity
    - Model of memory scheduling is still essential the simple SMP
  - Not even syntax to describe requirements of a job
    - How memory varies
    - How it needs to be organized

What else is needed

- **Holistic Scheduling**
  - **Integrated scheduling of the CPU, Memory, Network and Data**
    - **For a job**
    - **For a process**
  - **Basic things do not exist**
    - **How to describe needs**
    - **How to exchange priority information**
      - ***CPU Priority vs Network Q of S***



Production and R&D Portfolio

- Computer Science and Computational Science areas
- Sites will become more integrators of technology - like the old times
  - Clusters, cost effective solution scaling, etc
- Environment changes with each system generation
  - Research to support site's users
  - Research to chart future directions for site's users
  - Research to change the state of the art for site and others
    - Discipline Sciences
    - Large Scale computing
- R&D without production verification has little meaning

Summary

- Effective system determination is extremely important in our cost conscious world
  - Both before and after purchase
- It is possible to achieve high utilization with MPP systems
- Improvement will continue
- Storage and Bandwidth increasing important