

IBM's High Performance Computing Strategy

CAS2K5
Don Grice

IBM's High Performance Computing Strategy

Solving Problems More Quickly at Lower Cost

- Aggressively evolve the POWER-based Deep Computing product line
- Develop advanced systems based on loosely coupled clusters
- Deliver supercomputing capability with new access models and financial flexibility
- Research and overcome obstacles to parallelism and other revolutionary approaches to supercomputing

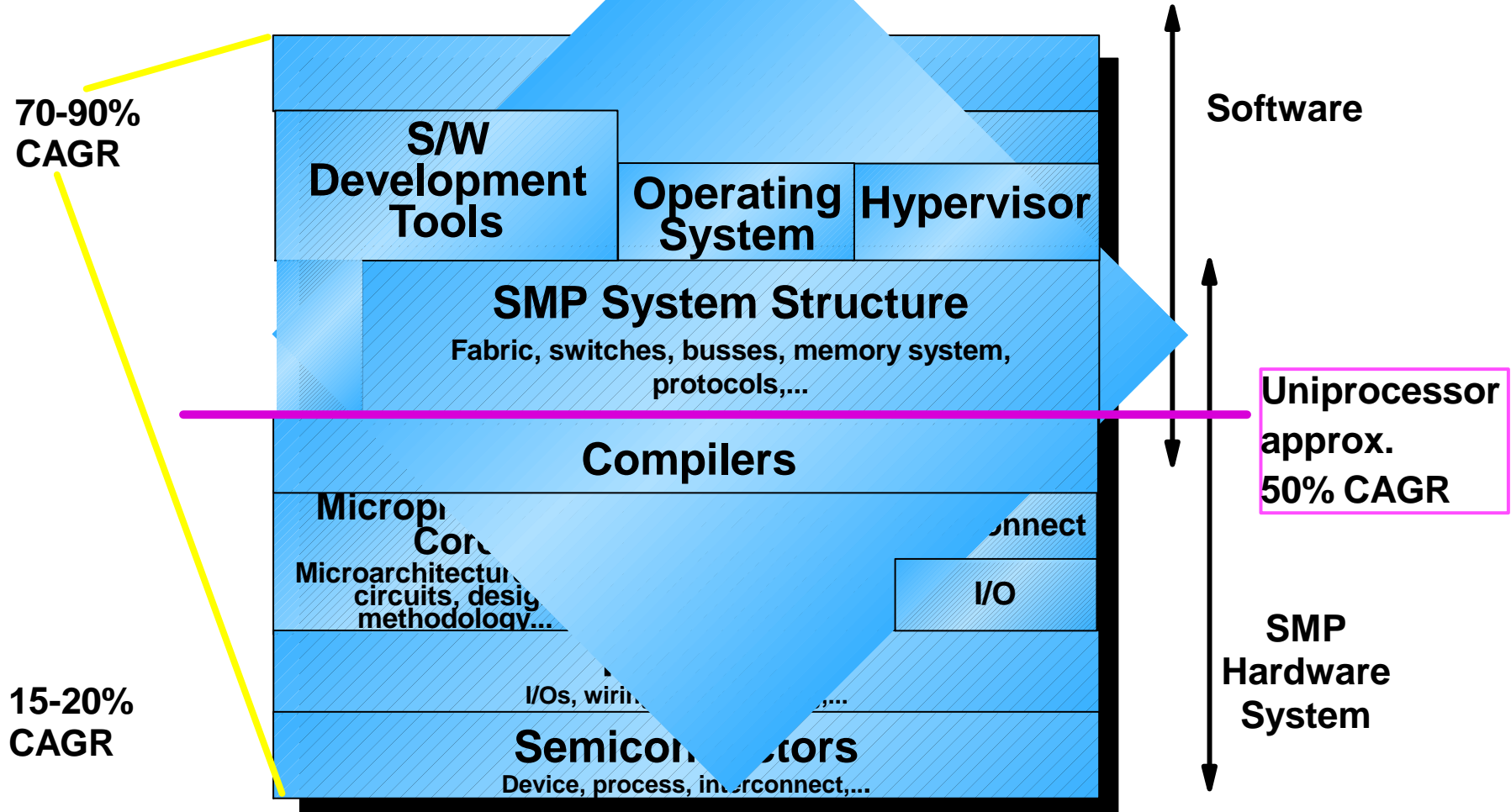


Technology Scaling; A Legacy Strategy

- Rate of Frequency Improvement is Slowing
 - ▶ Application Scaling Will Require More Parallelism
- Circuit Density Continues to Improve
- What is the Best Use of the Circuits?

System performance gains of 70-90% CAGR derive from far more than semiconductor technology alone

Performance improvements will increasingly require system level optimization

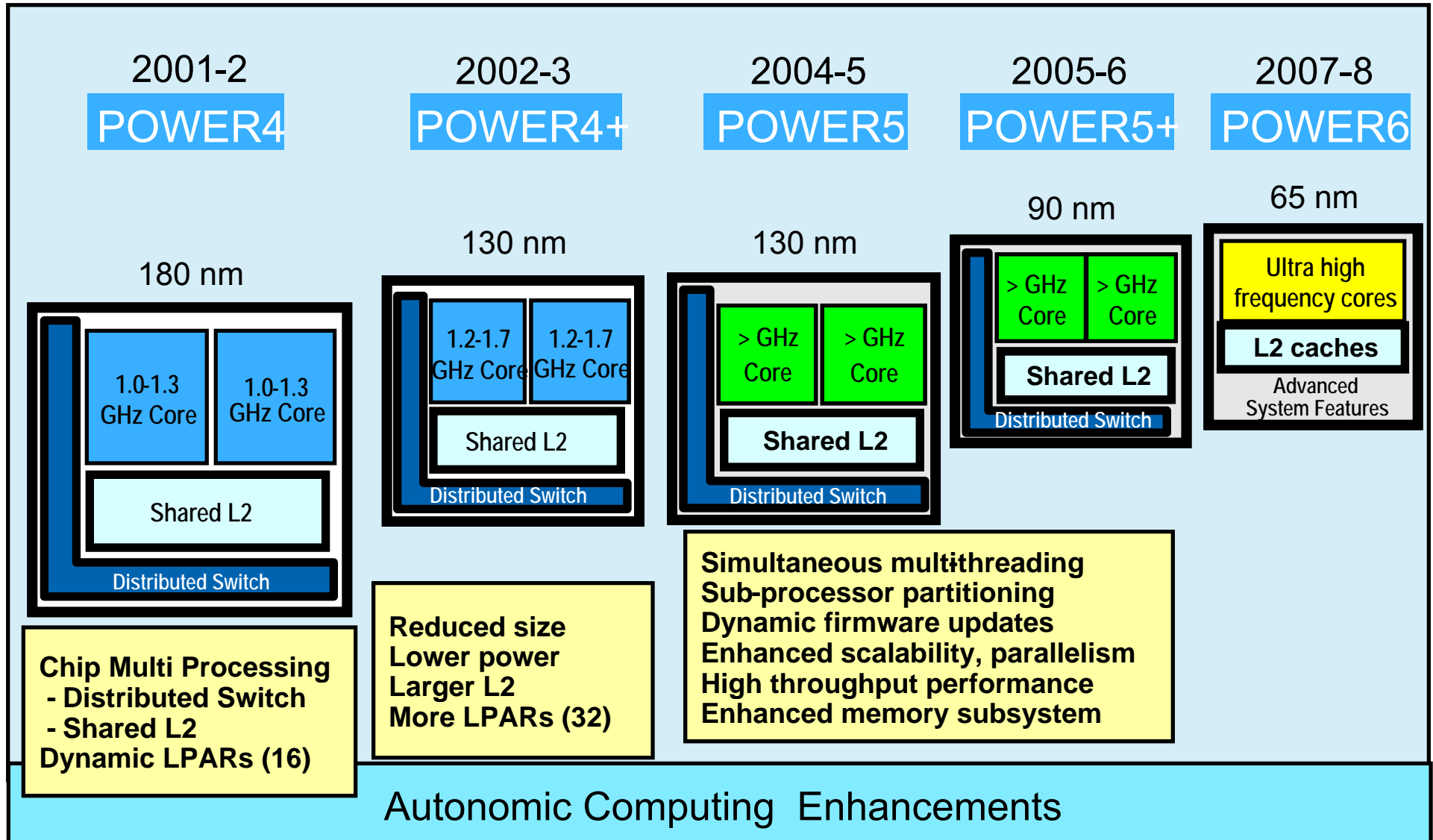


Multi-Core Options

- **Homogeneous Symmetric Multi-Core General Purpose CPUs (Multiple Threads)**
- **Homogeneous Symmetric Multi-Cores with Specialized Instructions**
- **Heterogeneous Cores with Specialized Accelerators**
- **'System on a Chip'**

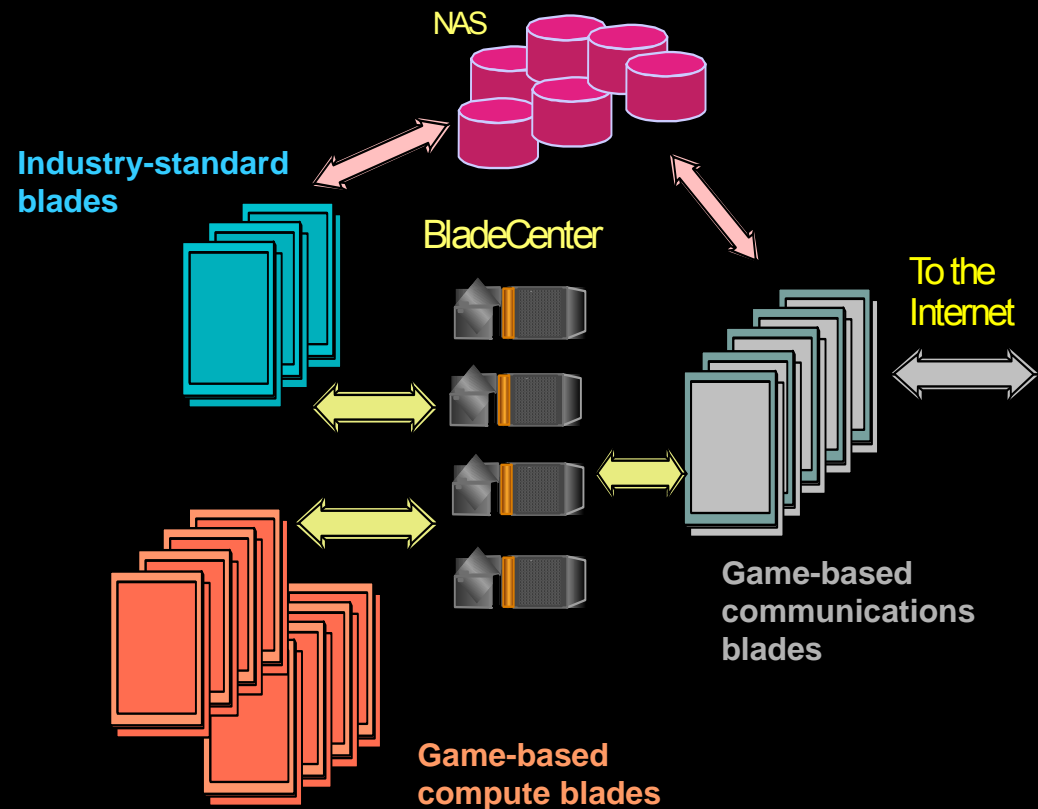
Science Driven Design

Multiple Core Processors and Optimizations



Game Processor Technologies

- The game console market is driving the development of microprocessors with high numeric processing capabilities, high bandwidths, and features to tolerate memory latency
- Research is exploring the use of game-processor technologies to boost performance in areas including:
 - High-performance scientific computing
 - On-line client-server games
 - Video applications including surveillance
 - Secure communications acceleration

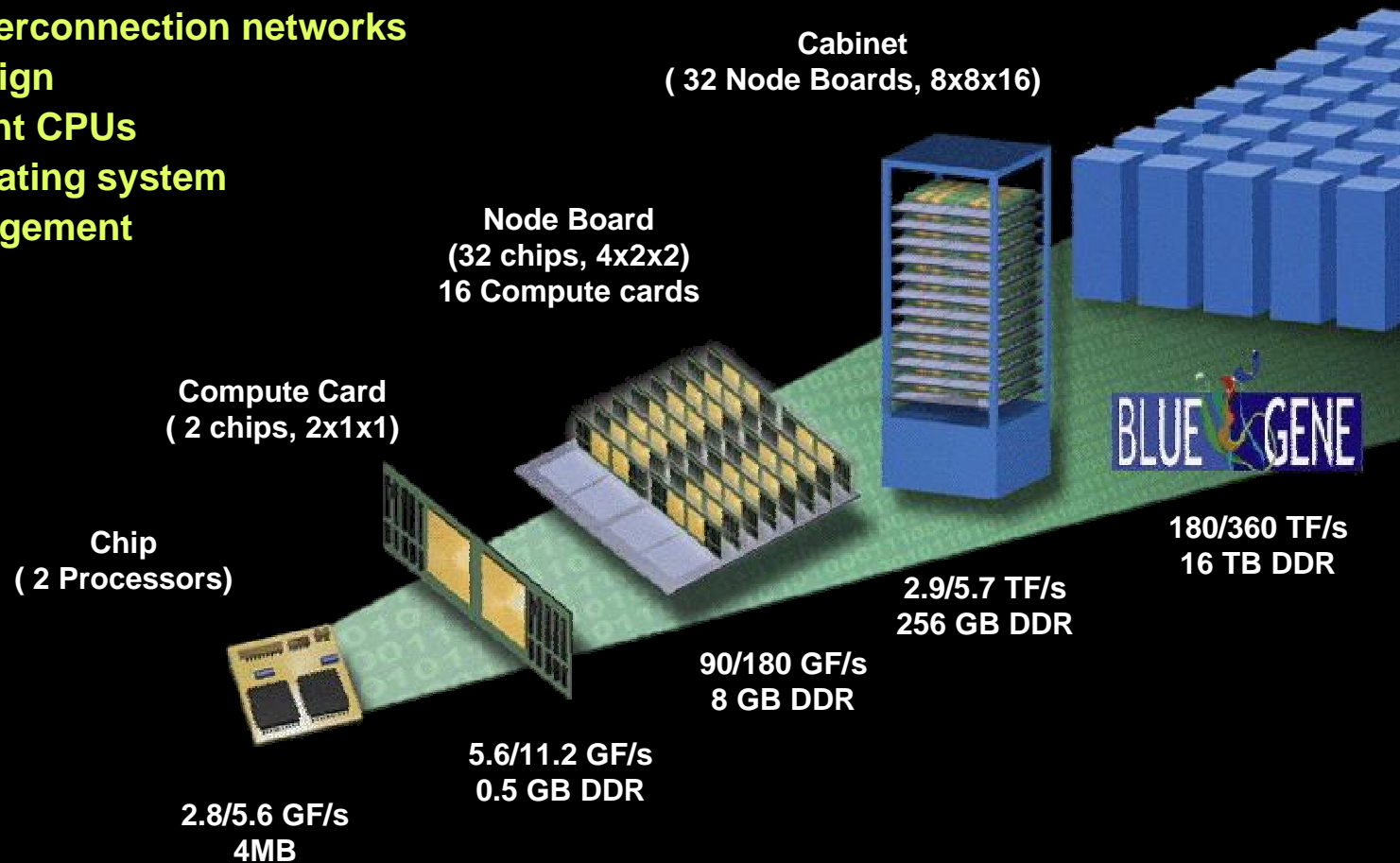


BlueGene/L

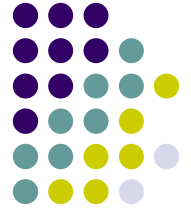
64K CPUs

System
(64 cabinets, 64x32x32)

- **Key Blue Gene Technology Elements**
 - ▶ **System-on-a-chip**
 - ▶ **Integrated interconnection networks**
 - ▶ **Balanced design**
 - ▶ **Power-efficient CPUs**
 - ▶ **Efficient operating system**
 - ▶ **System management**



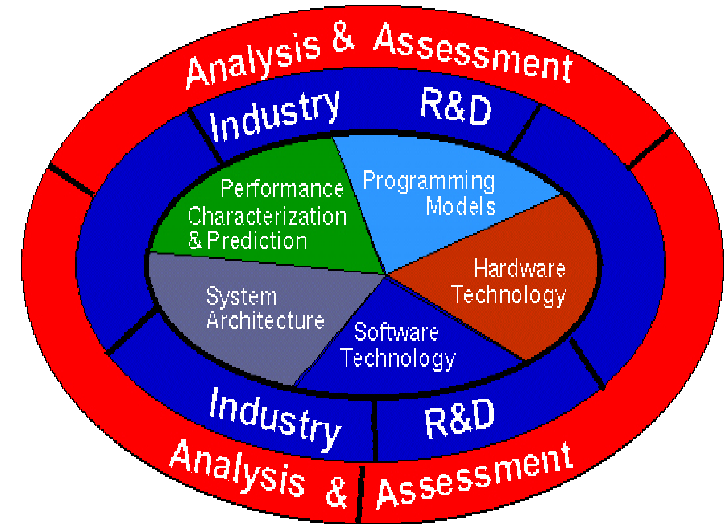
High Productivity Computing Systems



Provide a new generation of **economically viable** high productivity computing systems for the national security and industrial user community)

Goals:

- **Performance** (time-to-solution): speedup critical national security applications by a factor of 10X to 40X
- **Programmability** (time-for-idea-to-first-solution): reduce cost and time of developing application solutions
- **Portability**
- **Robustness** (reliability): includes security in addition to traditional RAS



HPCS Program Focus Areas

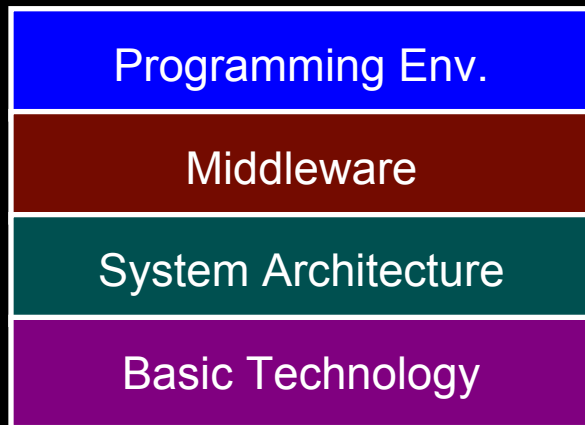
Applications:

- Intelligence/surveillance, reconnaissance, cryptanalysis, weapons analysis, airborne contaminant modeling and biotechnology

Ref: R. Graybill, DARPA 1/27/03

2010: High Productivity Computing Systems Research PERCS (Productive, Easy-to-use, Reliable Computer System)

Balanced attack across all system layers



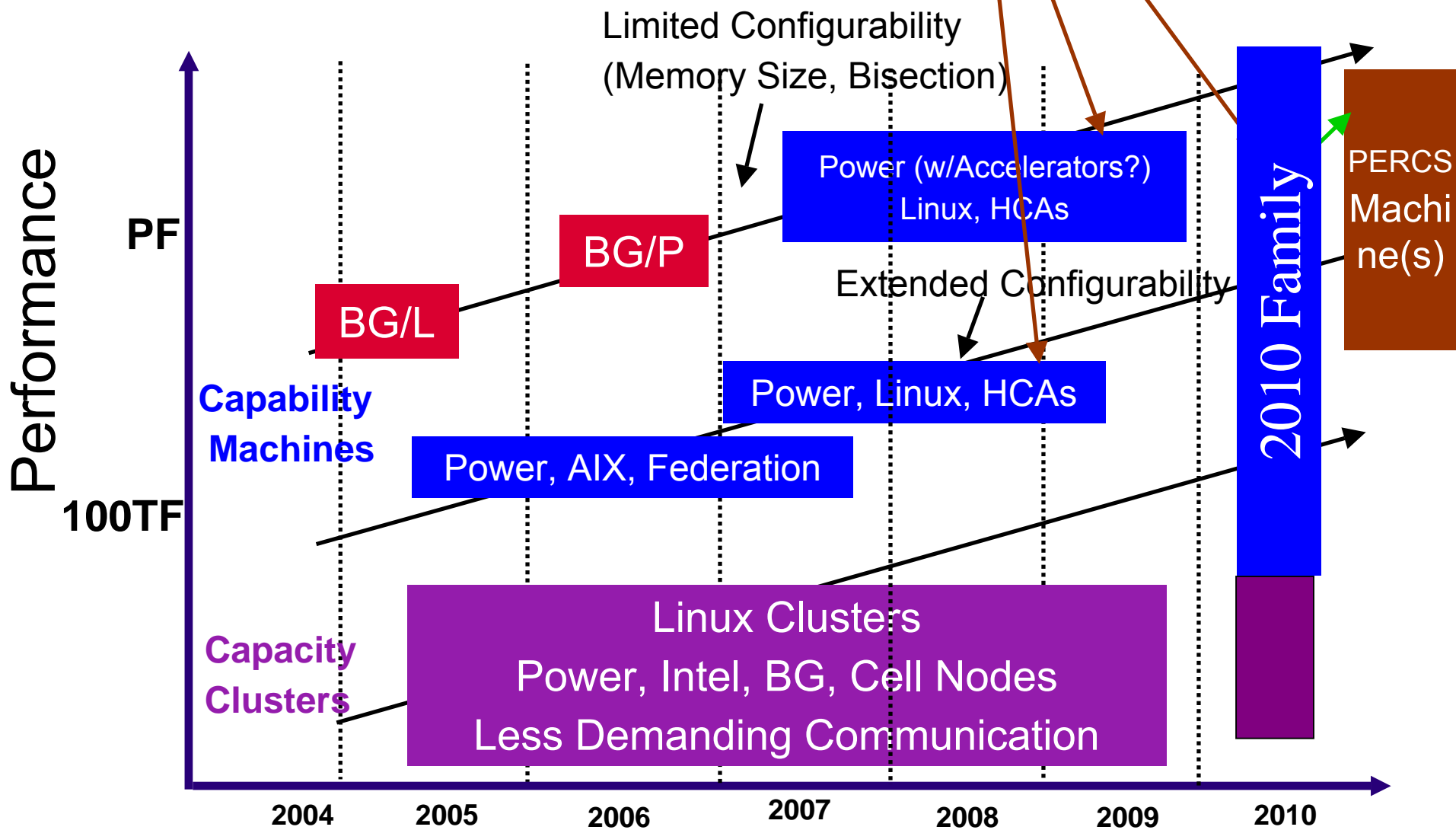
Main theme: A system that adapts to the application, not the other way around

- Continuous program optimization
- System performance evaluation methodology and infrastructure

- Programming environments
 - Focus on simplifying programming tasks and reducing development cycle
- Scalable OS and middleware
 - Support for on-demand computing
- Compilers
 - Tolerating memory latency
- Development of new systems analysis tools
 - An execution-driven evaluation infrastructure
- Systems architecture
 - Application dependent morphing architectures under software control, addressing the memory wall
- Circuits/power/technology
 - High performance, lower power circuits, system level power analysis, advanced packaging

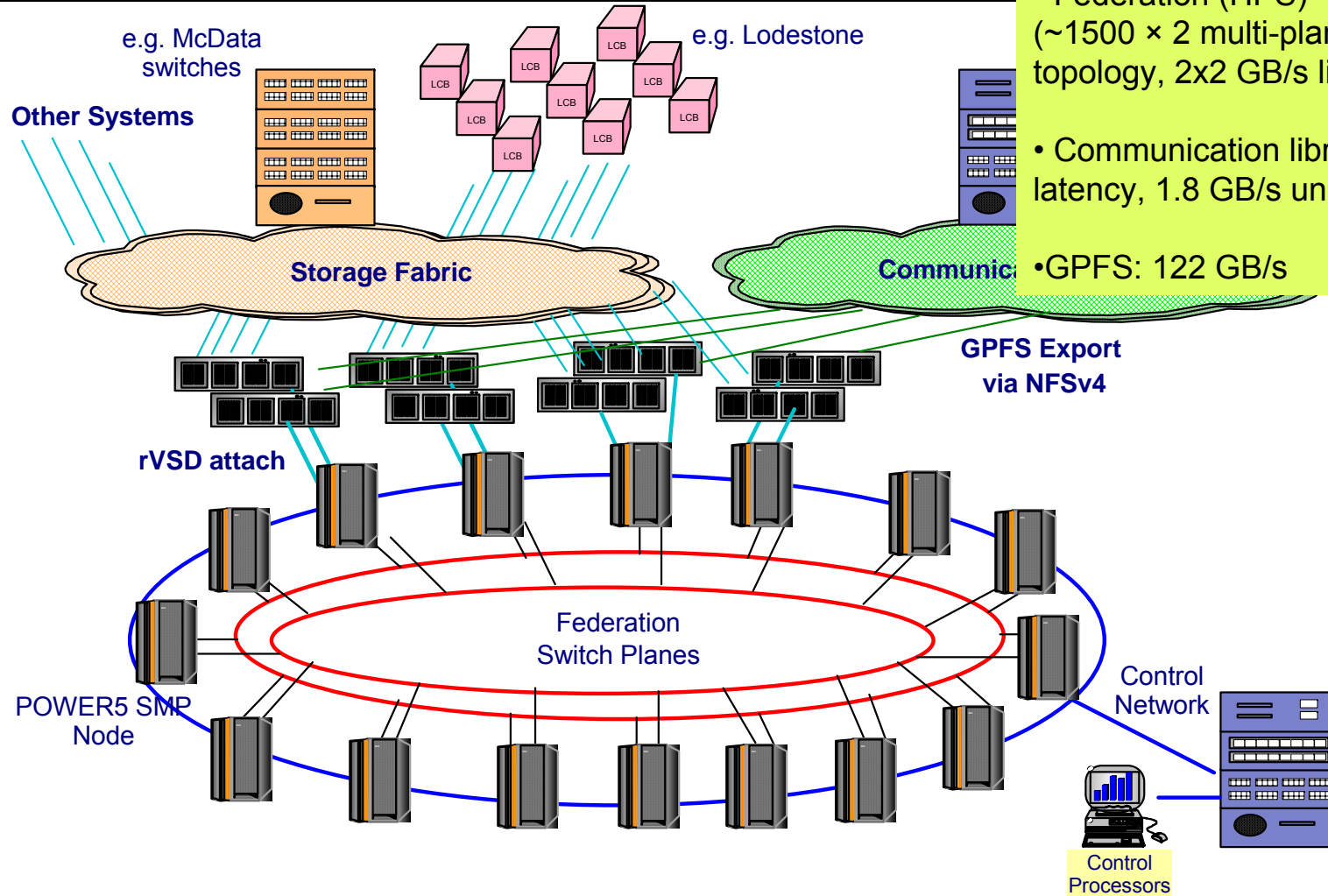
HPC Cluster Directions

PERCS Technologies

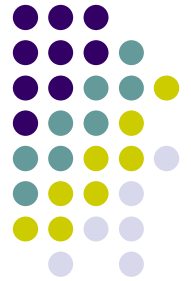


ASCI Purple Architecture

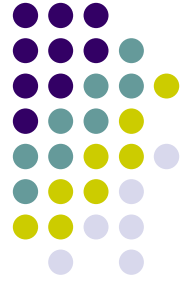
- 100TF Machine
- ~1500 8-way Power5 Nodes
- Federation (HPS) ~12K CPUs (~1500 × 2 multi-plane fat-tree topology, 2x2 GB/s links)
- Communication libraries: < 5 μs latency, 1.8 GB/s uni
- GPFS: 122 GB/s



Pain Points in Peta Scale Computing



- **Productivity:** Application Development and Optimization
 - Time to Solution
- **Application Performance and Scaling**
 - Serial/Node/Single Server Level
 - e.g. Memory Usage/Efficiency
 - Cluster Level Communications
 - Communications Efficiency (Sustained BW and Latency)
 - Collectives Efficiency
- **Data Generation/Storage/Management**
 - File System Size and Life Cycle Management (migration)
 - Multi-Cluster Access to Data
 - Parallel I/O BW



Software Initiatives

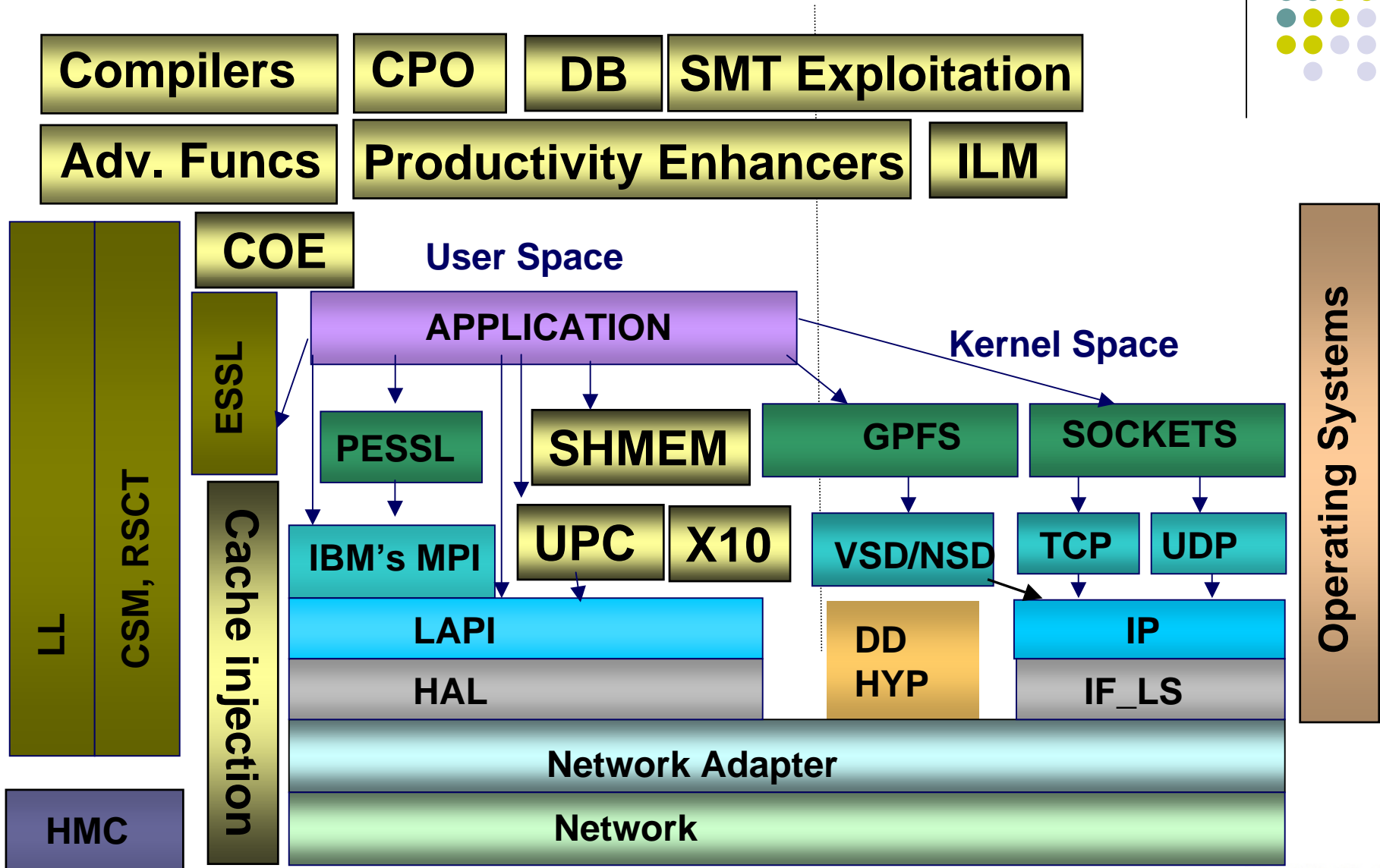
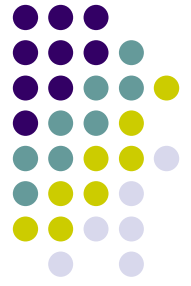
- Productivity enhancers
 - Programmer productivity (tools, libraries, ...)
 - Significant effort on improving sustained performance
 - Compilers
- Programming Model Run Time
 - Collectives, cache miss free latency and other performance critical paths
- Data Management: File systems, storage software
- Resource Management
- Systems Management
- Operating Systems, Hypervisors



Data Management

- Continued enhancement of multi-cluster capability
- Scalable file systems
 - 100PB file systems
 - 100s of millions of files
- Scalable HSM/ILM
 - HPSS or TSM based
- Tighter integration with databases
- Simpler access models (Google for the enterprise)

Software Architecture - 2010: Vision



Thank you

Questions?

