

IBM Almaden Research Center - Storage Systems



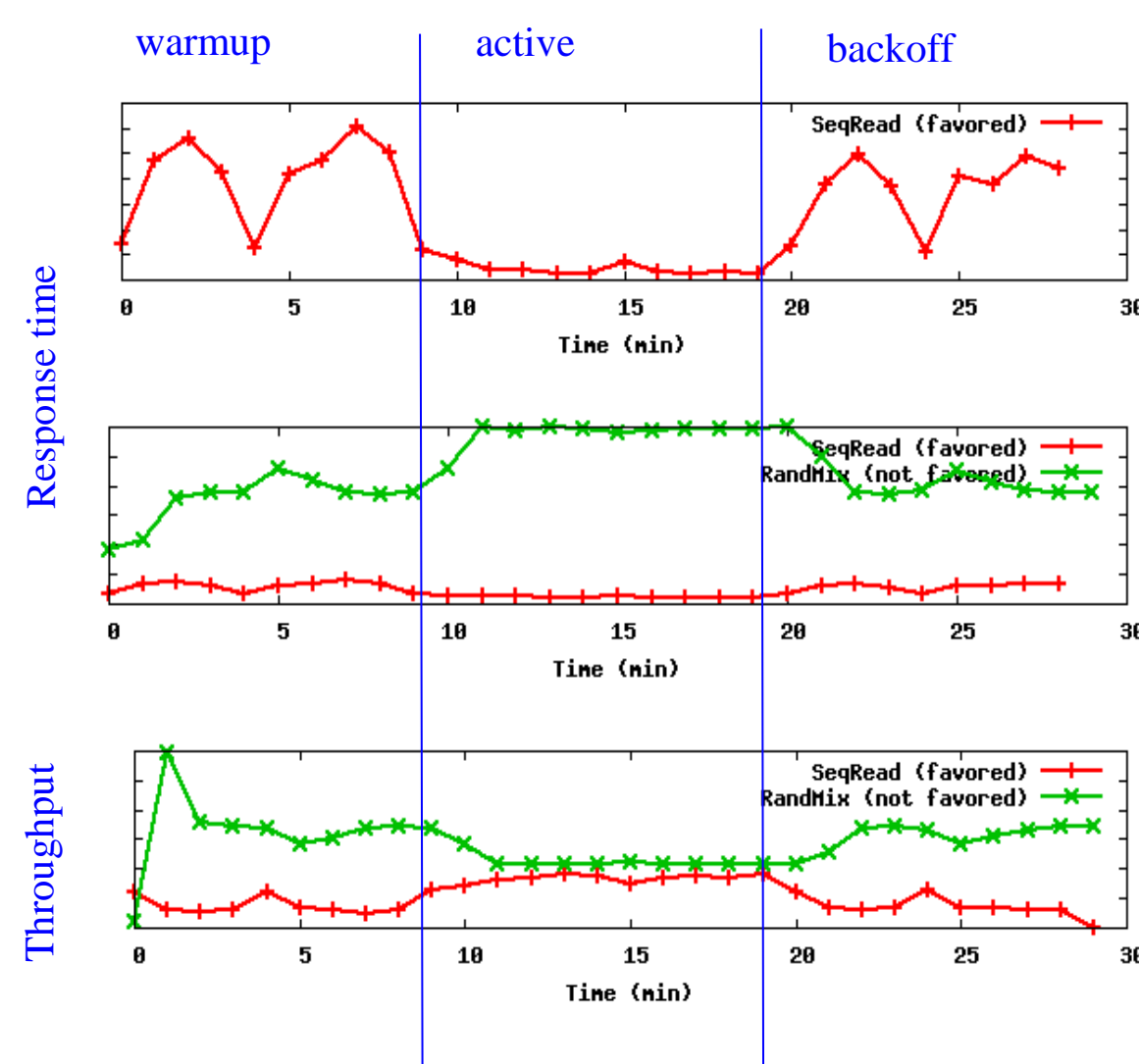
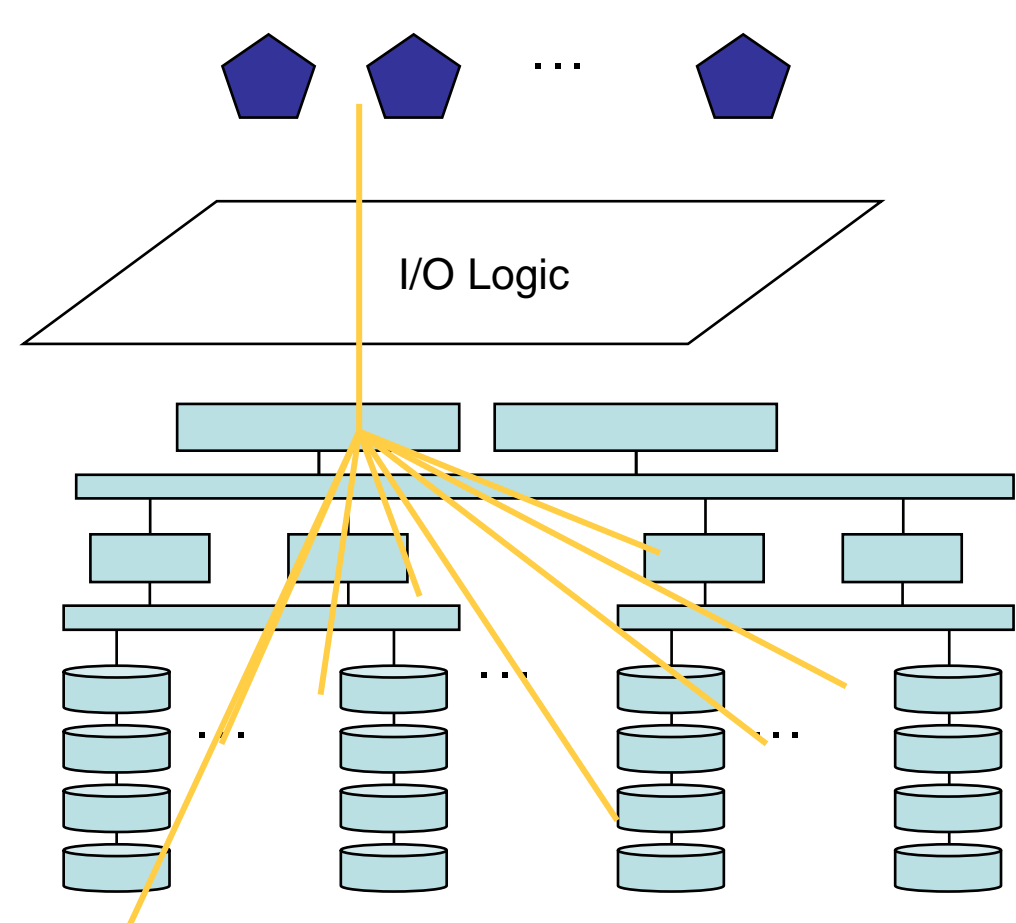
Active QoS Management for Storage

Motivation:

- High-performance storage is valuable, critical infrastructure
- Challenge—performance stress
 - Transient high load in secondary application can swamp performance for critical apps
- Manageability in the face of
 - Heterogeneity
 - Variable workloads

Core Technology

- Rate limits per I/O stream



Goals:

- Dynamically favor most important work
- Use resources fully

Principle

- Control resource utilizations for QoS
- Use rate limits to manage utilization

Challenges

- Smart I/O systems hide utilization from I/O response
 - Prefetch, write-back, ...
 - Inscrutable I/O logic

Virtualization Management

The "Virtual Data Center" Vision

- A dynamic and responsive data center that allows non-disruptive re-allocation of physical resources for application computation, storage and network connectivity.

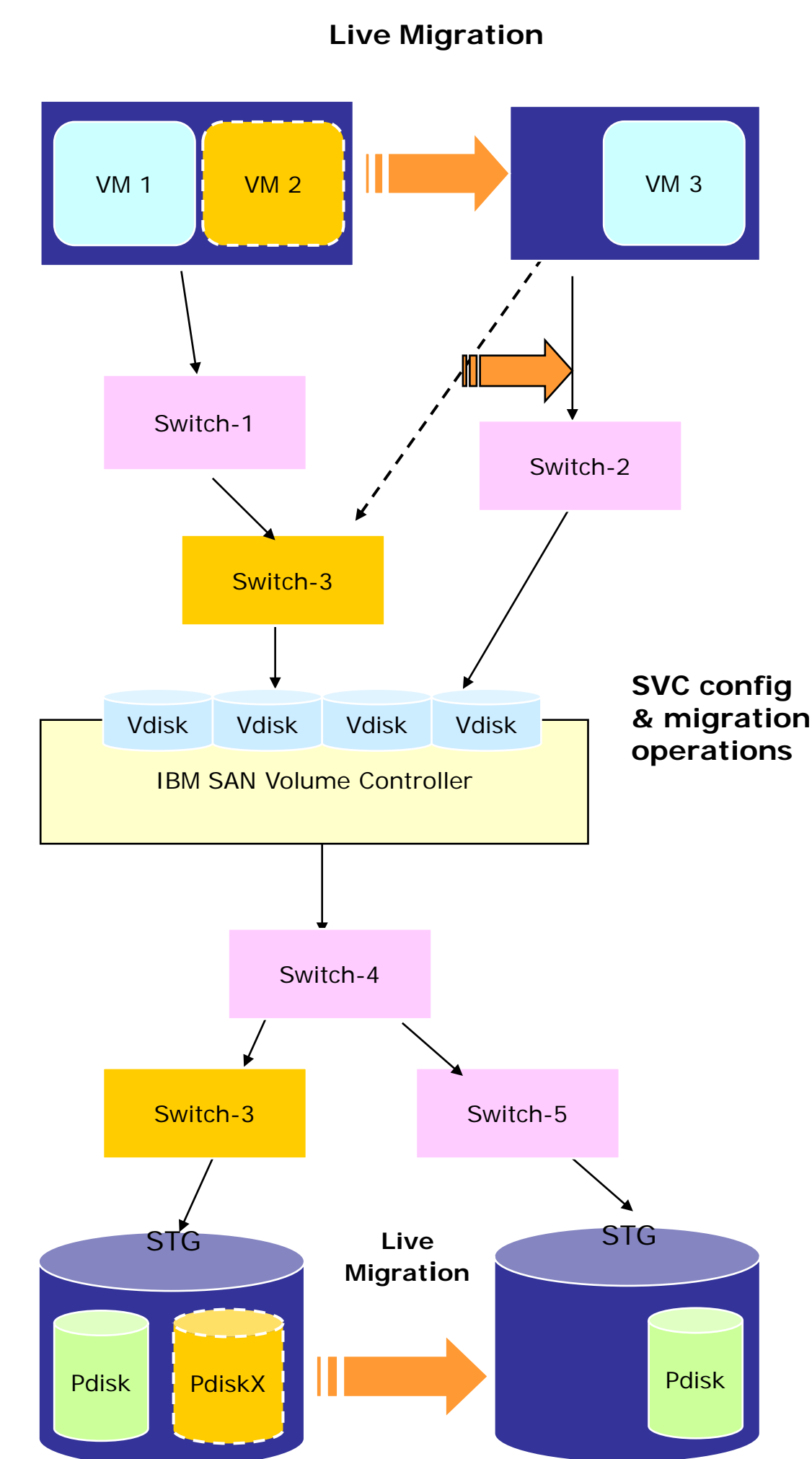
Server Storage Network Integration is key

Pieces in realizing the vision:

- Optimized provisioning of virtual resources (VM, virtual storage) on physical resources
 - Integrated provisioning for server and storage, accounting for network proximity
- SPARK continuous optimizations
 - Changing the virtual to physical resource mapping non-disruptively, instrumented through a variety of mechanisms (VM migration, SVC based migrations, path re-assignments...)
- Support for both block and file storage (incl. NAS)

Representative Publications

- Server-Storage Virtualization: Integration and Load Balancing in Data Centers, SC 2008
- Coupled Placement in Modern Data Centers, IPDPS 2009
- Shares and Utilities based Power Consolidation in Virtualized Server Environments, IM 2009



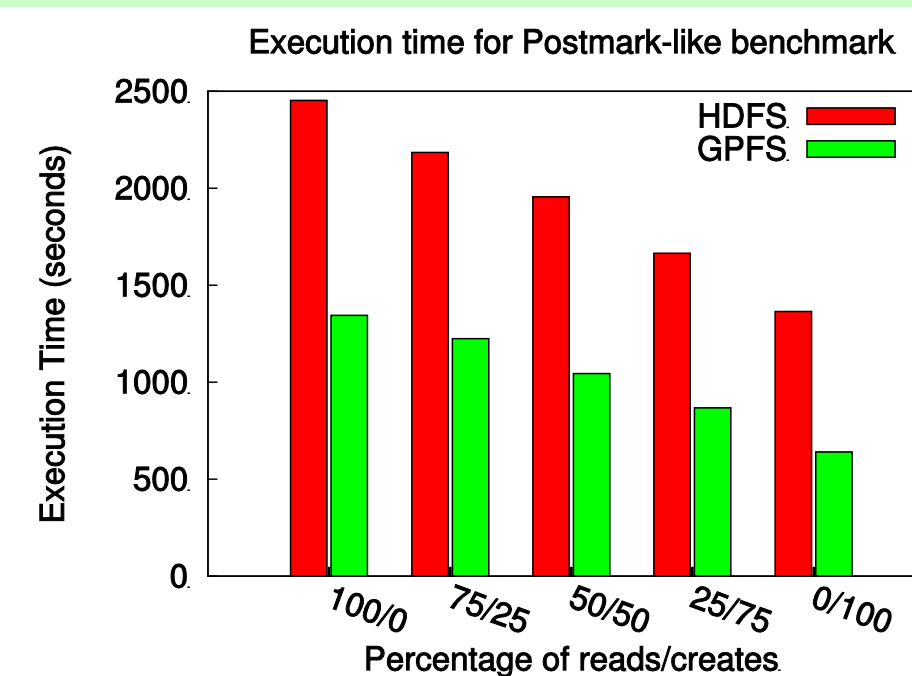
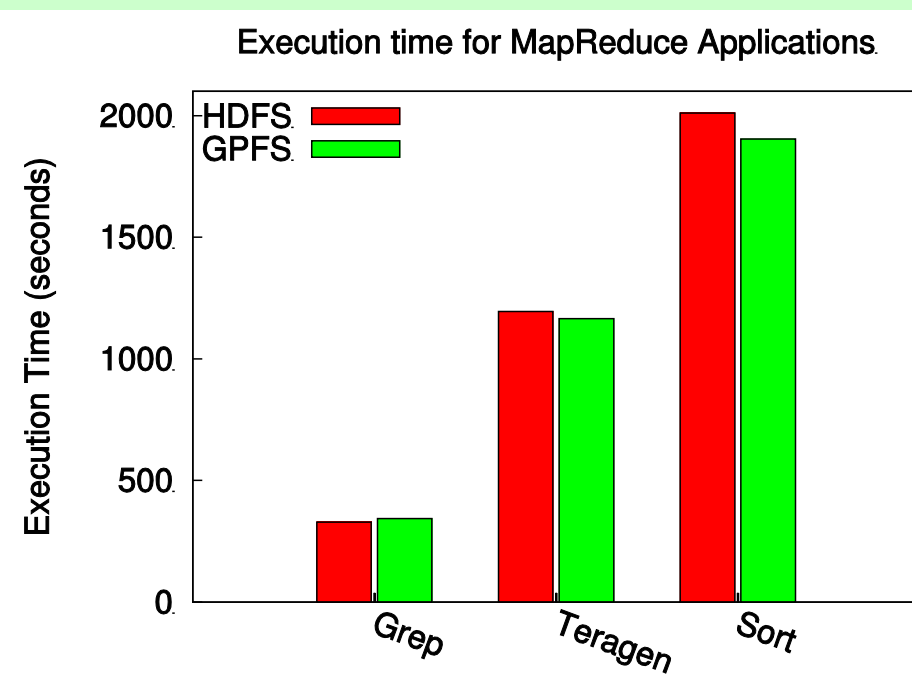
Orion: A low-latency high-throughput Internet scale file-system

Motivation:

- Internet objects are large and small
 - 8.3 billion video streams of 2.5 minutes (2007)
 - 21.3 trillion emails @ 10 KB (2006)
- Re-examine design principles of GFS
 - "High sustained bandwidth is more important than latency"
- Design for:
 - Both high throughput and low latency
 - Commodity components
 - Scale

Innovations:

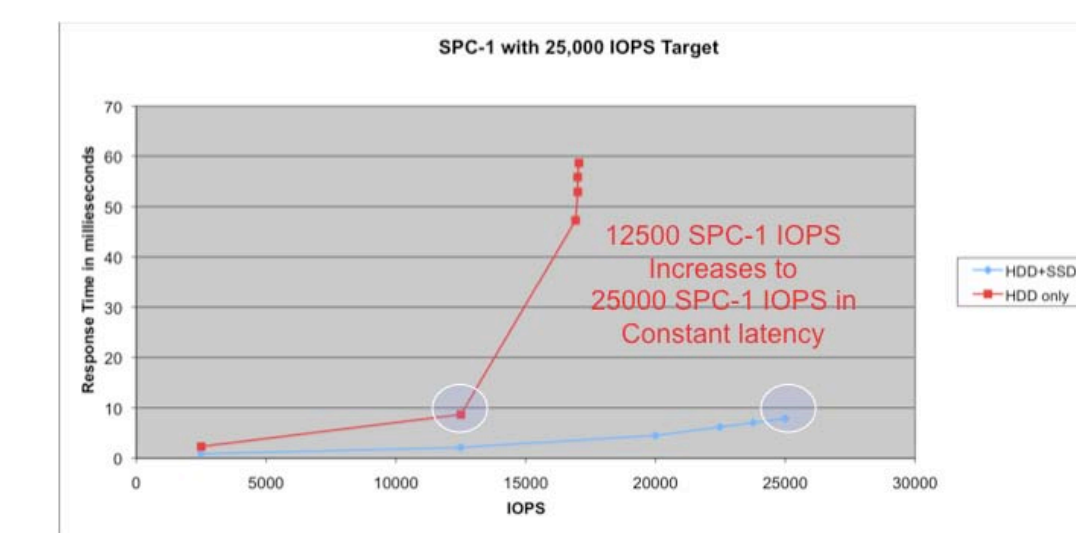
- Start with GPFS, then add...
 - Metablocks: small block size with layout control
 - Small blocks for low-latency apps
 - Large metablocks for high-throughput apps
 - Adaptive Multi-pipelined Replication:
 - Congestion-aware replication
 - Better utilization of available bandwidth
 - Finally, validate that distributed metadata has:
 - Better scalability
 - Low performance overhead



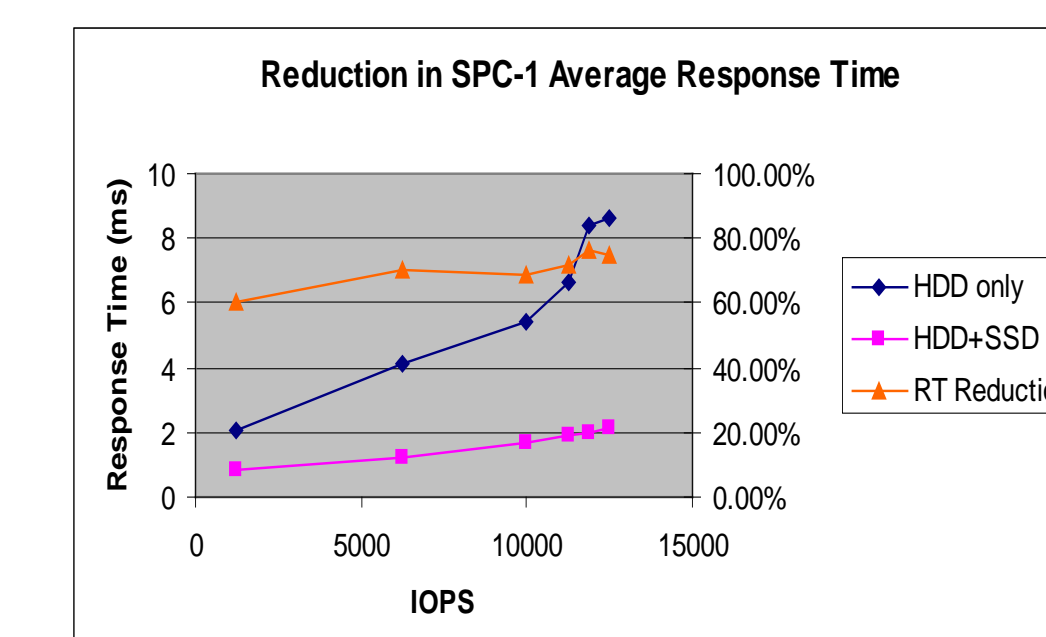
Automated Data Relocation with Solid State Disks on Multiple Tiered Enterprise Storage System

- Dynamically identify hot spot of workload in storage system and automatically place/migrate the right location in multiple storage tiers SSDs, Fiber Channel Disks, SATA Disk.

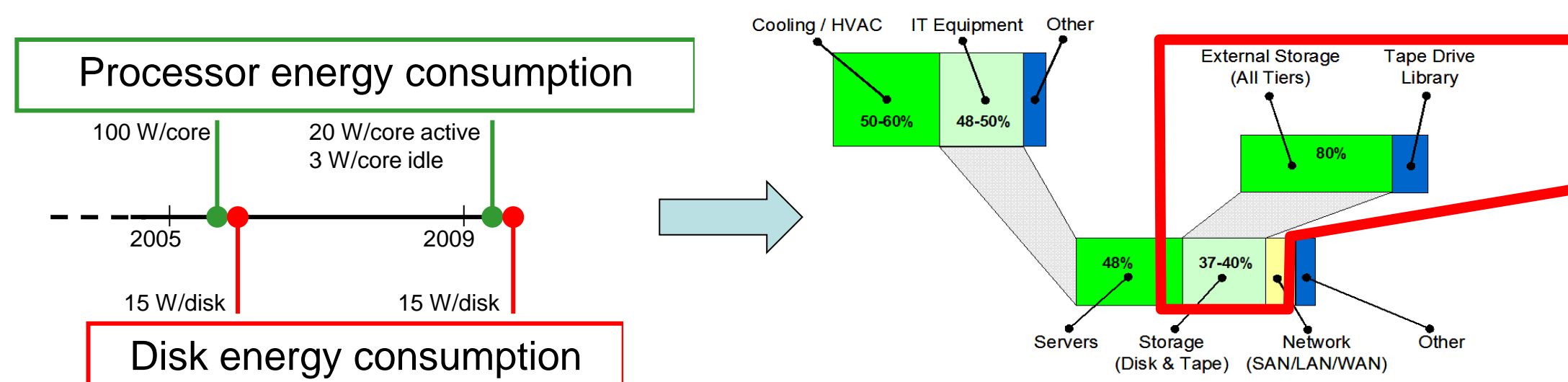
- Develop new insights of infrastructure optimization for both customers and system architects.



SPC-1 benchmark shows fine grain reorganization of 5% of data to SSD to gain 75% reduction of average response time or to double (200%) throughput.

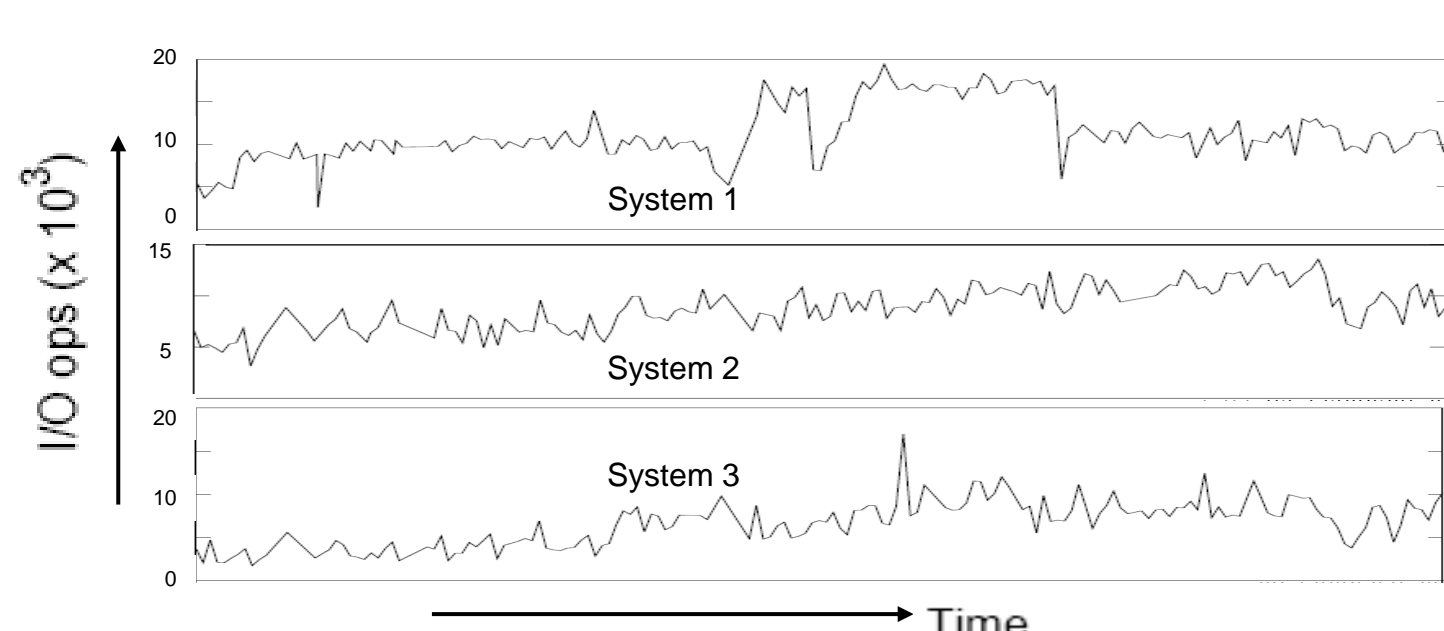


Energy Proportionality for Storage Systems



Storage uses a significant percentage of data center energy → Energy efficiency of storage systems is necessary, but not sufficient

Case for proportionality: Highly variable utilization, requires energy to vary with performance



Financial data center analysis

Framework for choosing energy proportionality techniques

Sensitivity to Avg. Resp. Time	Sensitivity to Peak Resp. Time	Stability of Workload	Techniques
			C T S W A H D
Yes	Yes	No	✓ ✓ ✓ ✓ ✓ ✓ ✓
Yes	No	No	✓ ✓ ✓ ✓ ✓ ✓ ✓
No	No	No	✓ ✓ ✓ ✓ ✓ ✓ ✓
No	No	Yes	✓ ✓ ✓ ✓ ✓ ✓ ✓

C: Consolidation, T: Tiering/Migration, S: Opportunistic Spin-down/MAID, W: Write Off-loading, A: Adaptive Seek Speeds, H: Workload Shaping, D: Dedup/Compression

Ongoing work: Tiering/consolidation for achieving energy proportionality, by leveraging different storage tiers (SSD, SAS, SATA)
Key challenges: (a) Continuous optimization of storage resources, (b) Driving online storage resources to be fully utilized