Increasing Performance of Existing Oracle RAC up to 10X

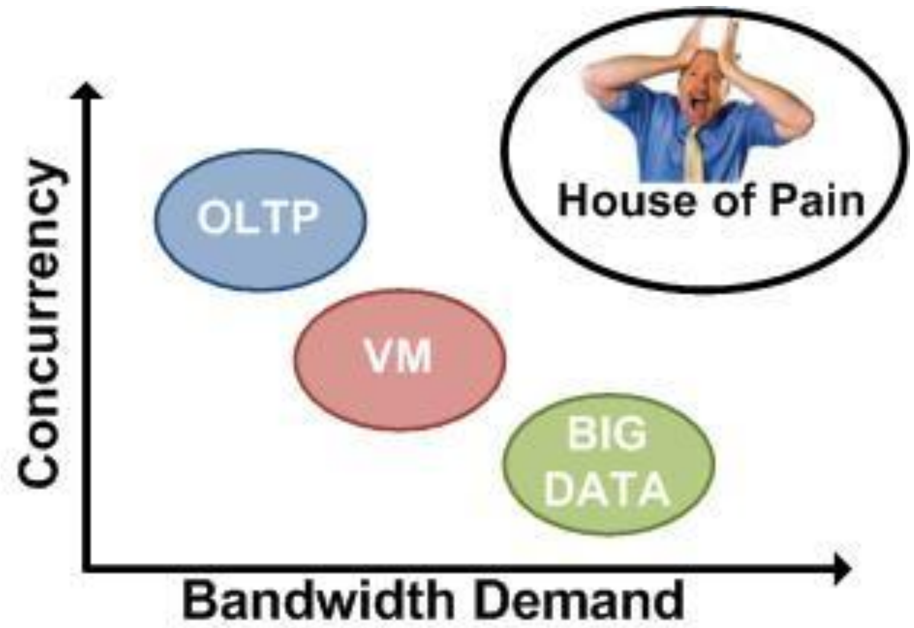
Erik de la Iglesia

www.gridironsystems.com

December 13, 2011

The Problem – Data can be both Big and Fast

- **Processing large datasets creates high bandwidth demand**
  - Rapid ingest through scans
  - Spills and reread of temp
  - Burst demand many times average
- **Concurrent queries and threads access the same data**
  - Data layout for bandwidth may not be concurrency friendly
  - Hot spots on disks or stripes
  - Write demand can stall reads



Databases with demand for bandwidth and concurrency run into a Storage Performance Wall

Oracle RAC with ASM

- **Stripe Tables over many LUNs and Distribute Processing**
  - Prevent multiple servers hitting same LUNs
  - Allow scans to utilize combined server IO
  - Failover of down server
- **Storage Performance limits effectiveness and scaling**
  - Storage system controllers outmatched by processing and IO capability of servers (the IO Performance Gap)
  - Storage architectures not designed for linear performance scaling (designed for capacity)
  - Virtualized layout can still cause physical disk or stripe hot spots

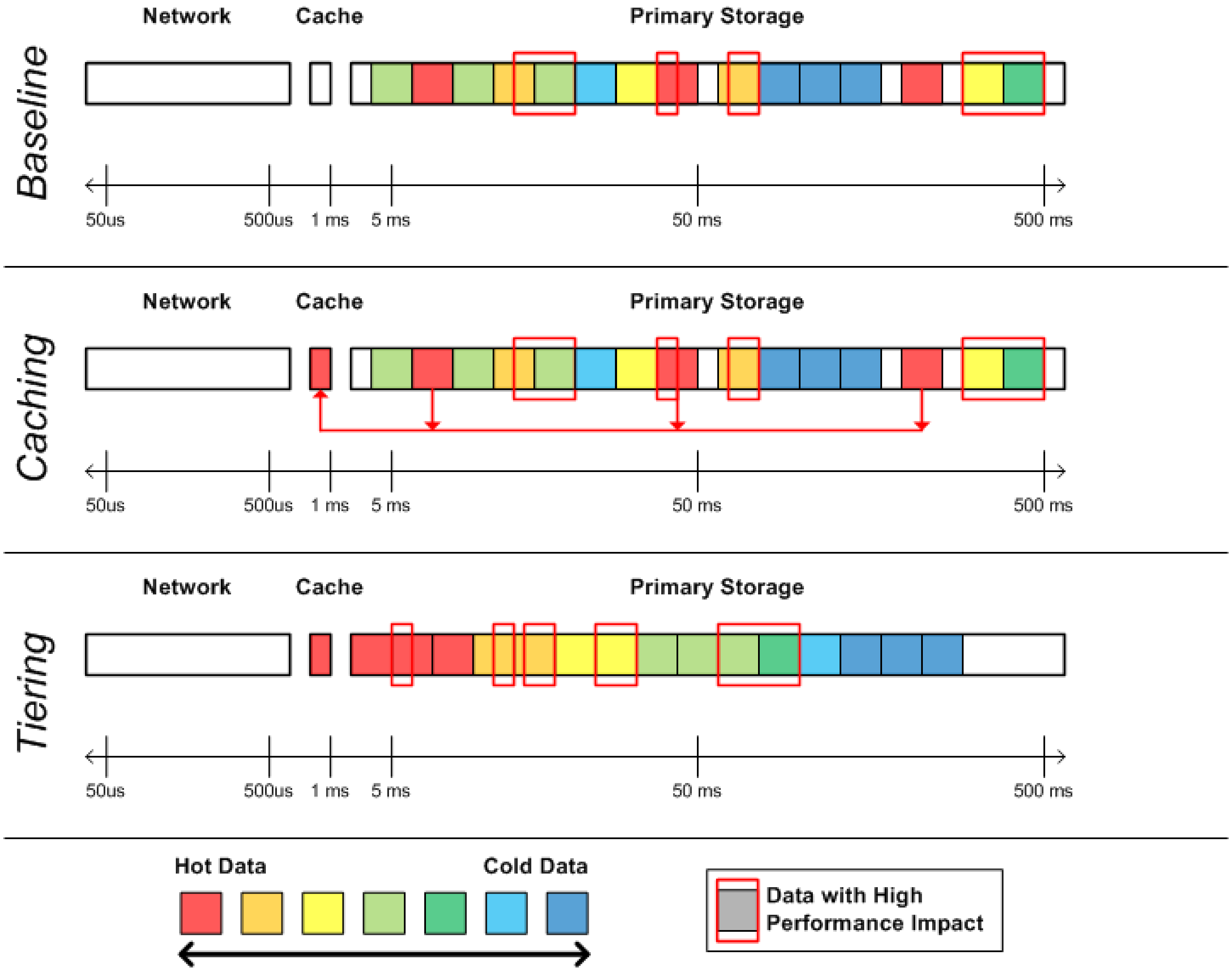
Every storage management function and feature requires resources taken from application IO processing

Flash to the Rescue - Maybe

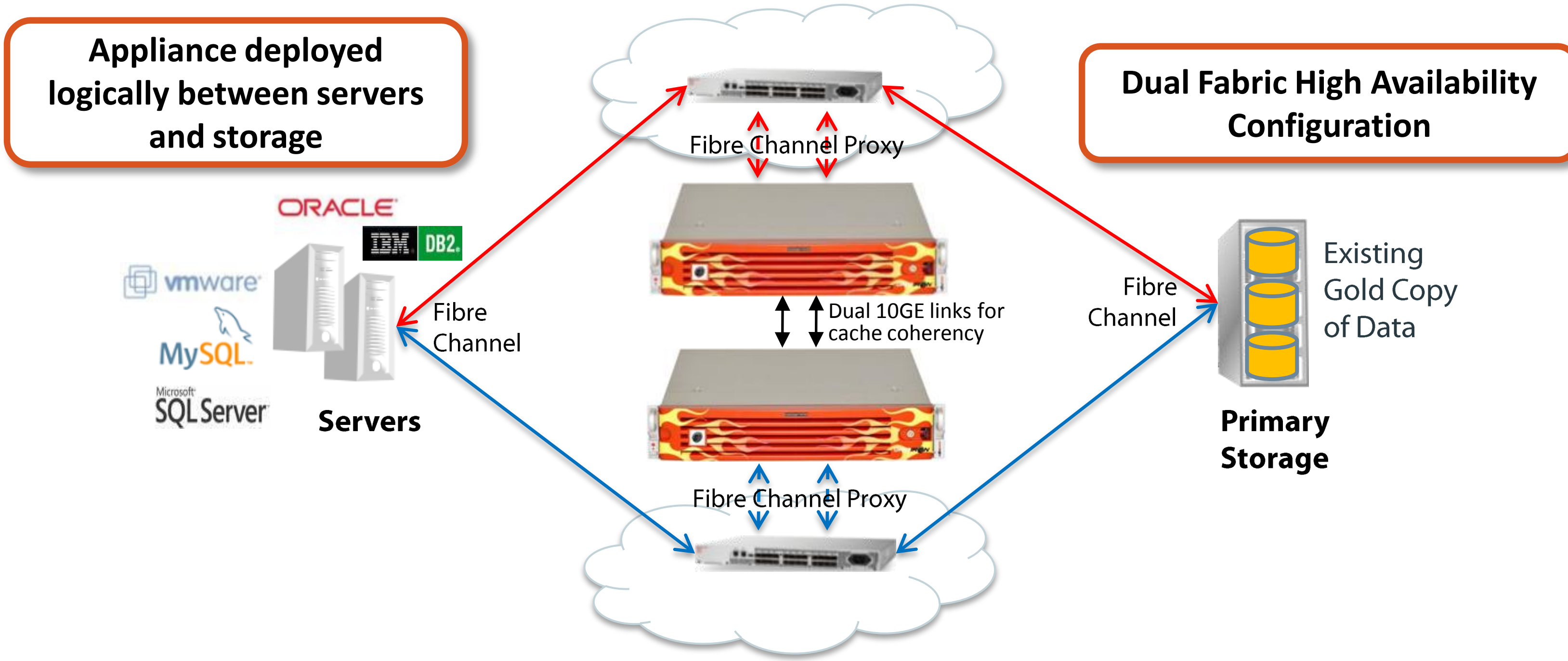
- **Flash in the Servers**
  - Changes to Software
  - How is HA handled?
  - No sharing among servers
  - Large CPU overhead
  - Capacity and effectiveness?
- **Flash in the Storage Array**
  - Architectures not designed for Flash
  - Sequential performance no better than spinning disk
  - Effectiveness of caching and tiering
  - Controllers can still be bottleneck to scaling
- **Full Flash Storage Array**
  - Sized for entire physical disk capacity and growth
  - Not economical even with compression and deduplication
  - Forklift upgrade of existing datacenter – changes to applications and processes

What about deploying in the network?

Caching and Tiering in Database Storage



Network-Based Flash for Database Acceleration



Transparently accelerate data access in the SAN

Solid State Performance with no change to:
Software - Databases - Servers - Storage - Processes

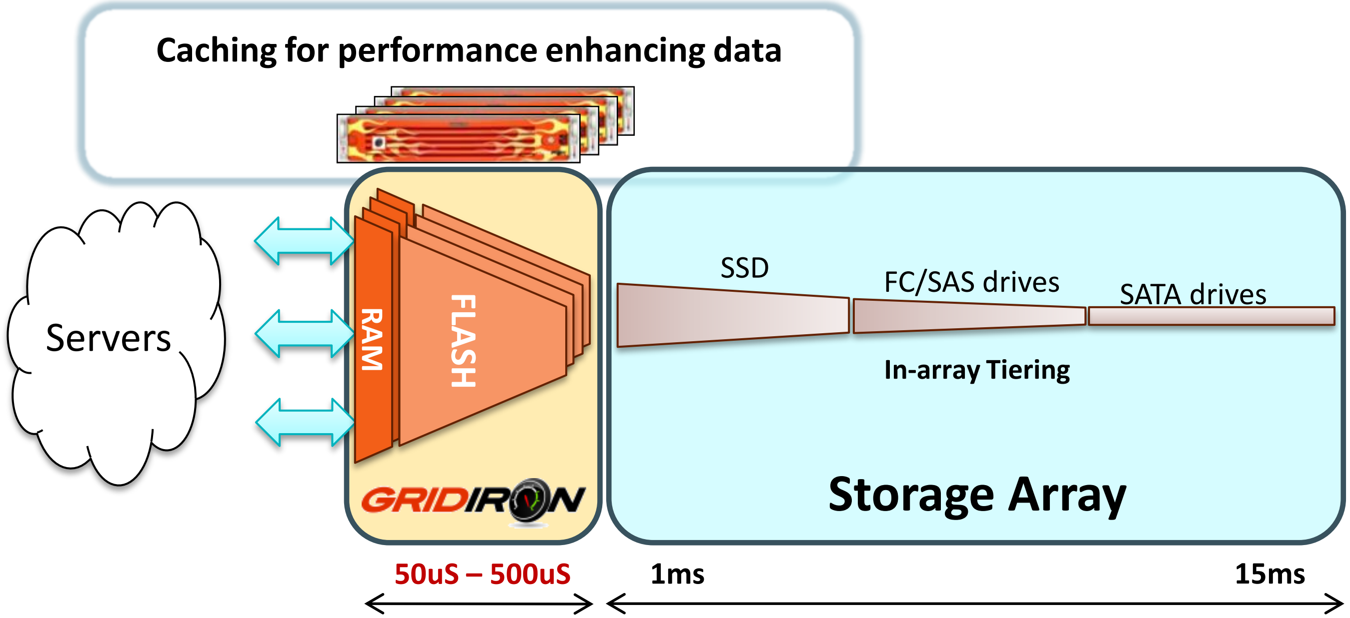
Real-Time Tiering Enables High Concurrent Bandwidth

Acceleration in the Network

- Higher concurrent IO bandwidth
- Higher IOPS
- Low latency multi-level cache

The Learning Process

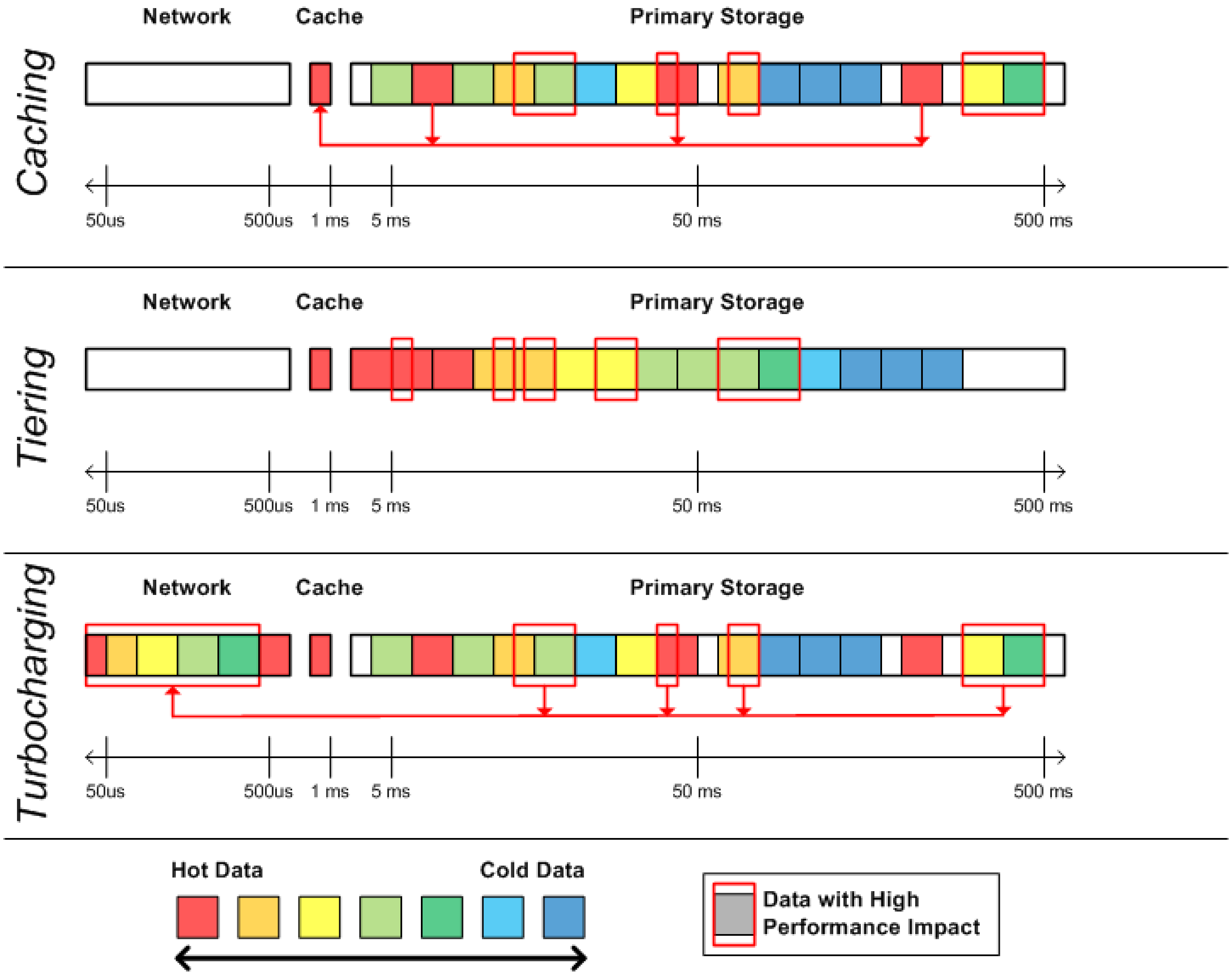
- Learn data access graphs in real time
- Use patterns to manage caching
- Use feedback to continuously refine performance



Results

- **Applications 2-10x faster**
- **Read latency 10-100x shorter**
- **Flash life a non-issue**

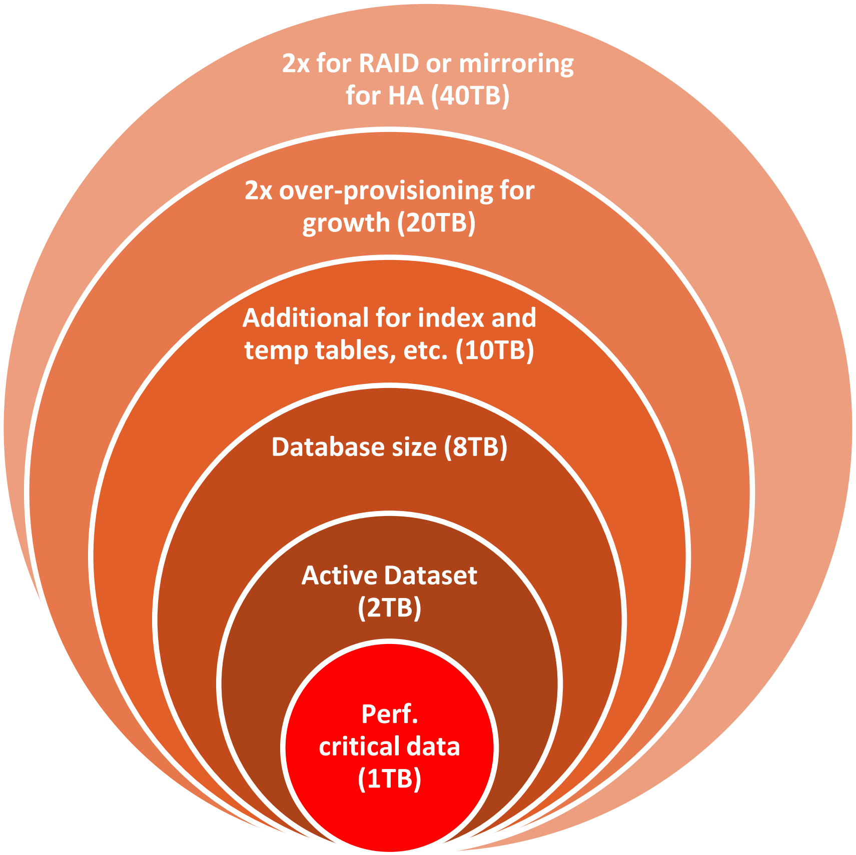
TurboCharging Database Storage



GRIDIRON

Network Solution Provisioning vs. Dataset (8TB Example)

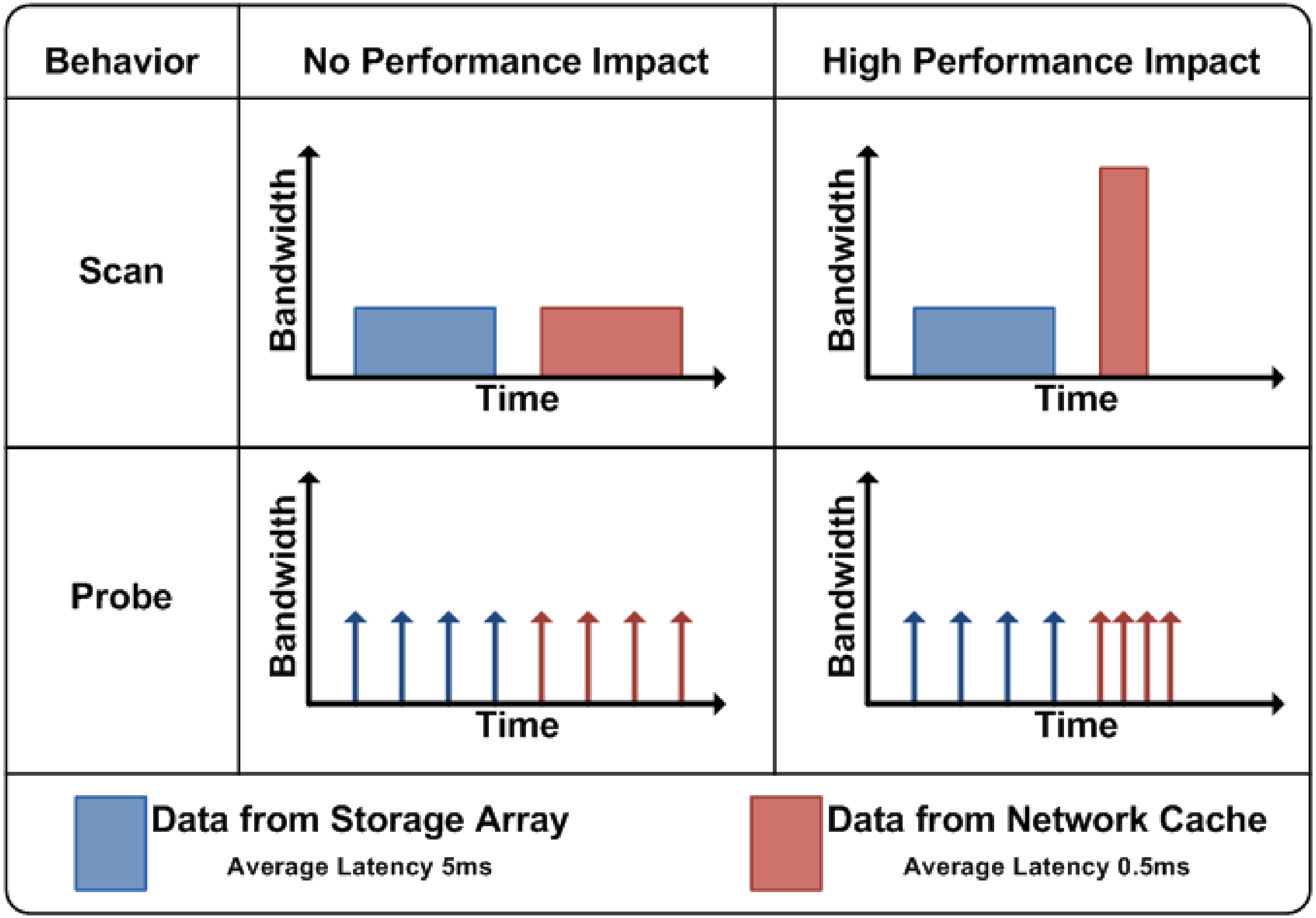
- **Smart Flash in the Network**
  - Sized for a fraction of dataset
  - Adapts in real-time to changes in usage and scale
  - Is shareable among servers, applications and arrays
  - Is always coherent with backend storage state
  - Requires no changes to applications or data management processes
- **Overcomes physical limitations of storage architecture**
  - Highest concurrency access to performance critical data
  - Scale bandwidth and IOPS without regard for architecture of storage system
  - Separate data access from data retention
  - Leverage and extend existing storage investment



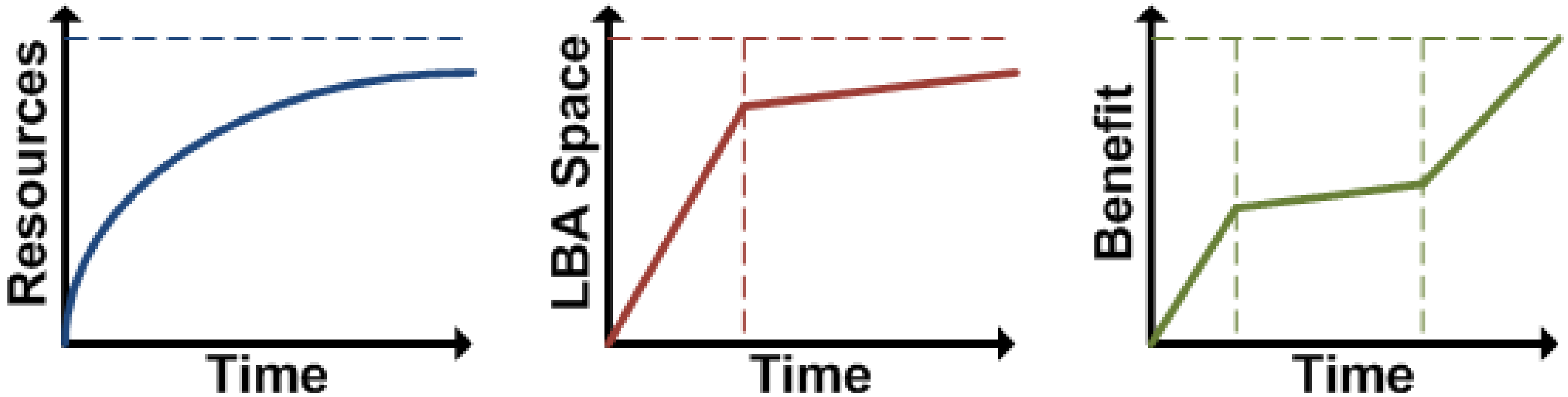
Flash Control and Effectiveness

- **Network Cache in not Primary Storage**
  - Can use RAM for high churn data and critical blocks
  - Learns what not to cache (no capacity churn)
  - Flash not subject to write patterns of application
  - Uses large, aligned and contiguous writes
  - No over-provisioning, RAID or rebuilds
  - Can achieve stripe width far beyond arrays
- **Use profitability as eviction scheme**
  - Collect statistics over entire storage space
  - Set Rank pixelates storage map
  - Use application behavior to dynamically adjust chunk size
  - Perform cost-benefit analysis of each caching decision
  - Reinforce or punish behaviors based on application reaction

Selected Profitability Examples for Database Operations



Dynamic Profitability in Caches



- **Benefit Analysis based on real measurements**
  - Increase in scan bandwidths or peak bursts
  - Spacing of probes or dependent reads
  - Work product of application (write bandwidth)
  - Relative improvement vs. other sets
- **Cost based on resources and state**
  - Amount of flash media used for set
  - Current resource pressure (new vs. speculative sets)
  - Convergence of data space vs. time
  - Set Rank velocity and mean benefit

Case Study: 40 TB DWH For Online Comparison Shopping Site

Challenges

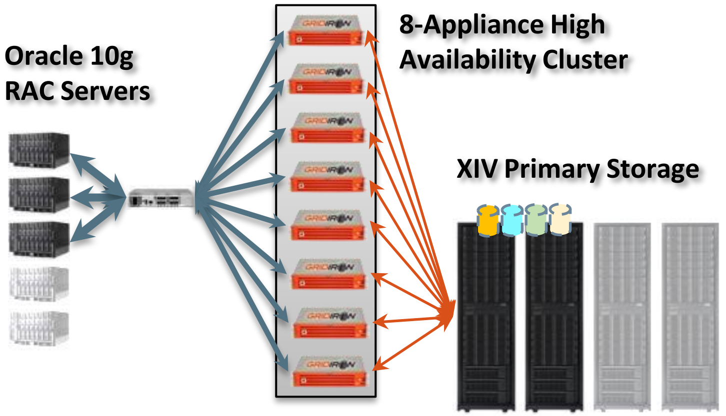
- Customer behavior analytics cycle taking too long (six hours) directly impacting revenue optimization
- Lost revenue from delays in fixing anomalies in customer-facing infrastructure
- Prohibitive storage acquisition and management costs from rapid data growth

Environment

- **Storage:** IBM XIV Storage Systems
- **Servers:** Dell 2950 server nodes (16GB DRAM) with dual QLogic 8Gbps FC HBAs
- **FC Fabric:** QLogic SANbox 9000 FC switches
- **GridIron:** **Eight GT-1100 TurboChargers in a striped configuration**

Benefits

- Business-intelligence reports' run time reduced from 6 hours to 30 minutes
- Near real-time decision-making to optimize operations and maximize revenue
- CapEx savings of over $2M compared to alternatives
- Ability to support more online products
- Ability to handle peak holiday loads without degradation in performance



"Online data analytics is at the heart of what we do as a company. We live and die by our data!"

Burzin Engineer, VP of Infrastructure Services, Shopzilla

Case Study: Acceleration of Software Builds Under VMware VDI

Challenges

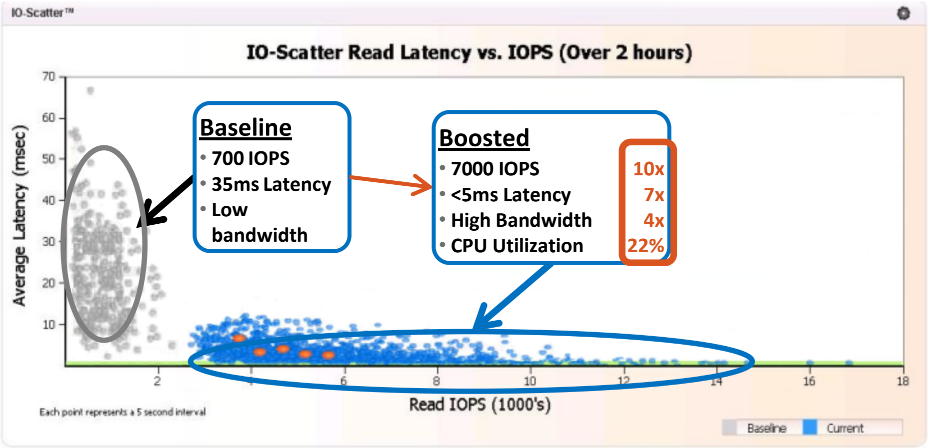
- Revenue and competitiveness impacted by increasingly long software build times
- Virtualized architecture not scalable due to I/O limitations of SAN storage
- Sub-optimal productivity of developers

Environment

- **Storage:** NetApp FAS3270 Storage
- **Servers:** HP c3000 BladeSystems
- **FC Fabric:** Brocade 300 SAN switches
- **VMware:** vSphere Hypervisor™, Virtual Desktop Infrastructure (VDI) 4.0
- **GridIron:** **One GT-1100A front-ending storage of VMware cluster**

Benefits

- Software build times reduced from 70 minutes to eight minutes
- Increase in sales from games with more features
- Savings of over $800,000 from avoiding performance upgrades to storage arrays
- Cost-effective scaling of the virtualized build environment



"Our developers' expectations of the game build environment have risen dramatically! They can now add more features without impacting release dates."

Dan Mulkiewicz, IT Director, High Moon Studios

Turbocharging a RAC deployment with Network Cache

- **Change the bandwidth physics**
  - Partition cache to match peak server demand
  - Storage system primarily used for writes
  - No data layout optimization or management required
  - Scale in situ with server growth
- **Leave the environment untouched**
  - Transparent for servers, applications, storage and processes
  - HA maintained via ASM and old fabric zones
  - Same LUNs with same data
- **Score significant performance wins**
  - Increase concurrent bandwidth
  - Decrease latency where it matters
  - Reserve storage processing for writes and data management

Questions?