Increasing Performance of Existing Oracle RAC up to 10X

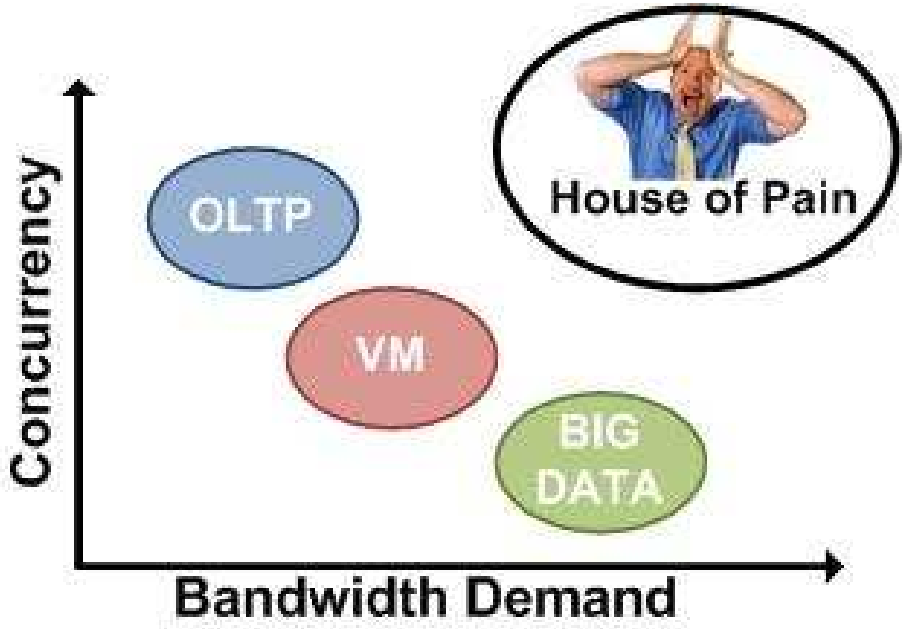
Prasad Pammidimukkala

www.gridironsystems.com

February 23, 2012

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 – Potential Limited by Storage

- **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

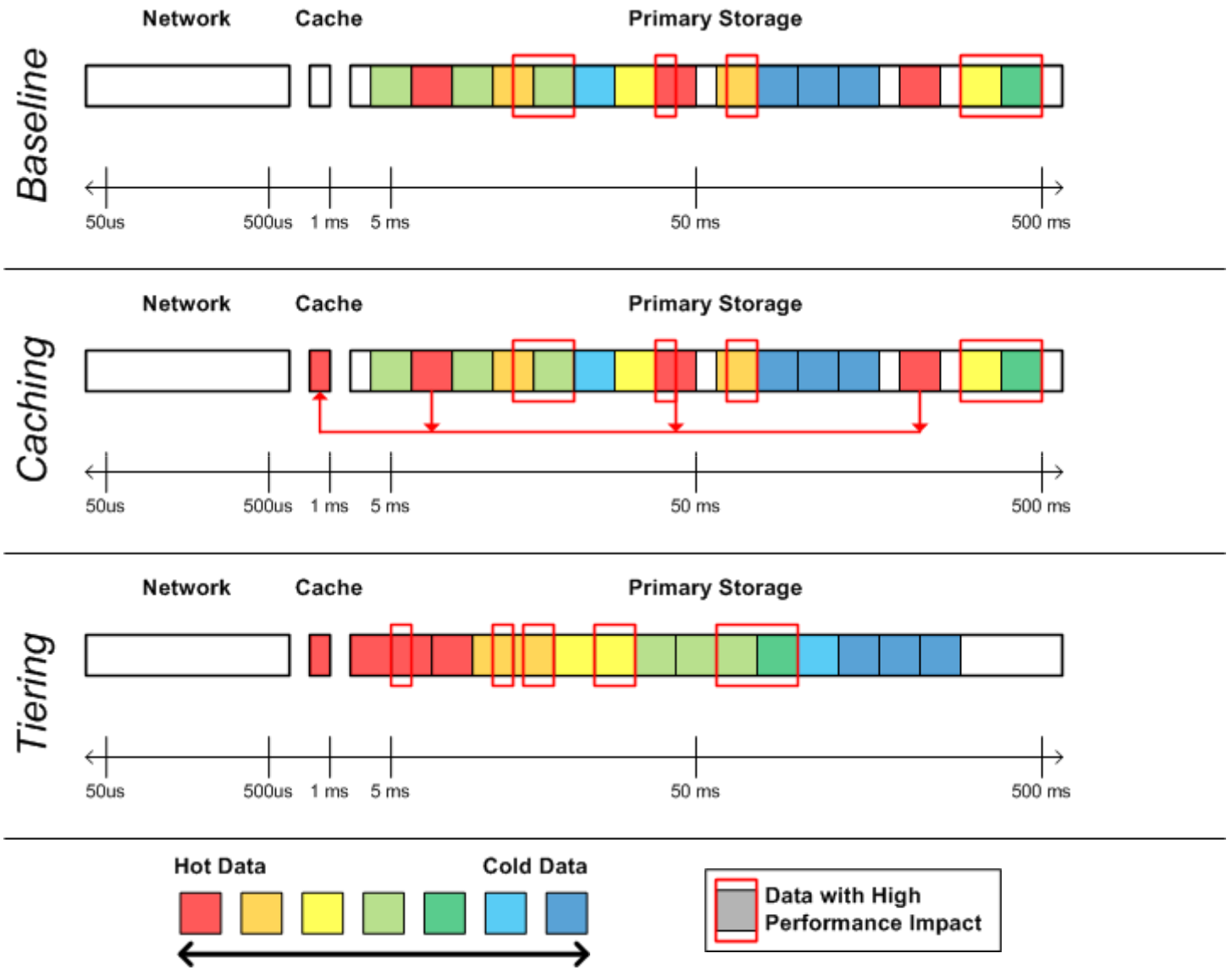
Every storage management function and feature requires resources taken from application IO processing

Flash to the Rescue - Maybe

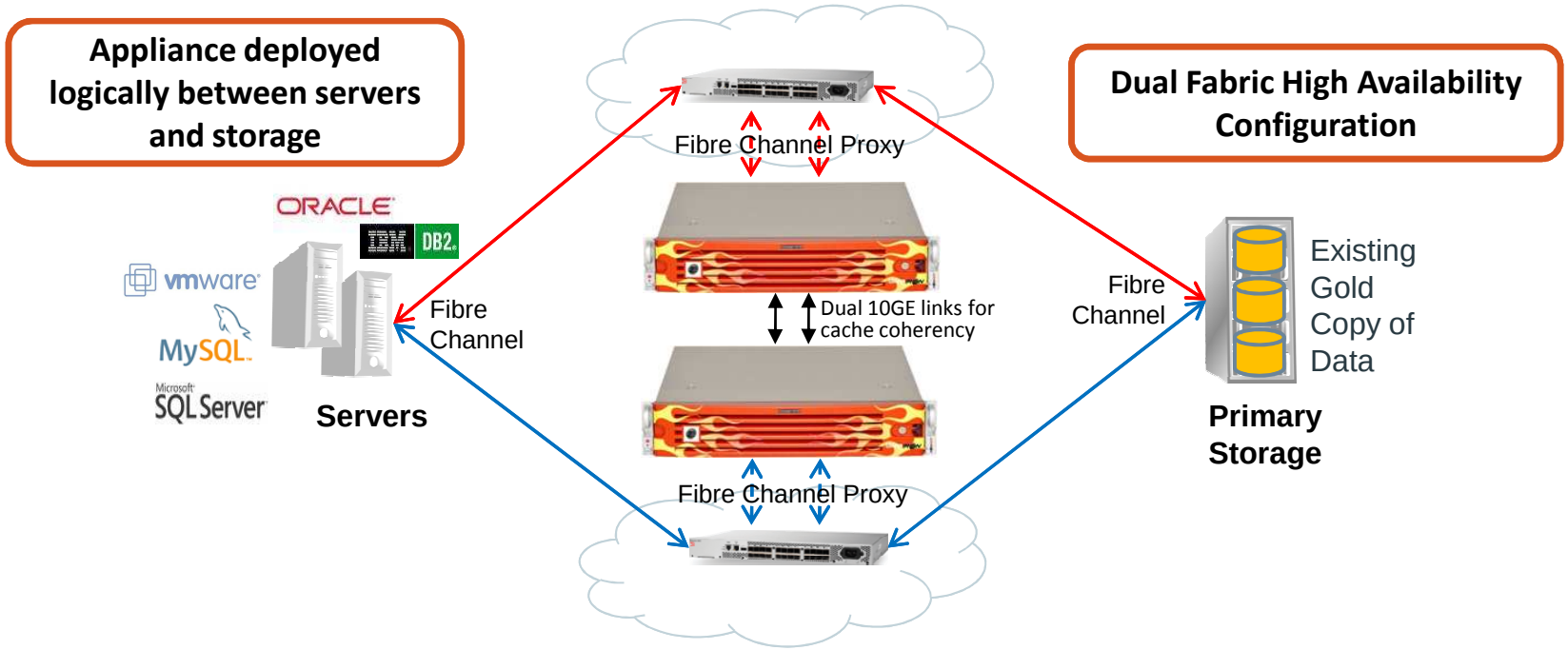
- **Flash in the Servers**
  - Changes to Software
  - How is HA handled?
  - No sharing among servers
  - Large CPU overhead
  - Capacity and effectiveness?
- **Flash in the Storage Array**
  - Architectures not designed for Flash
  - Sequential performance no better than spinning disk
  - Effectiveness 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

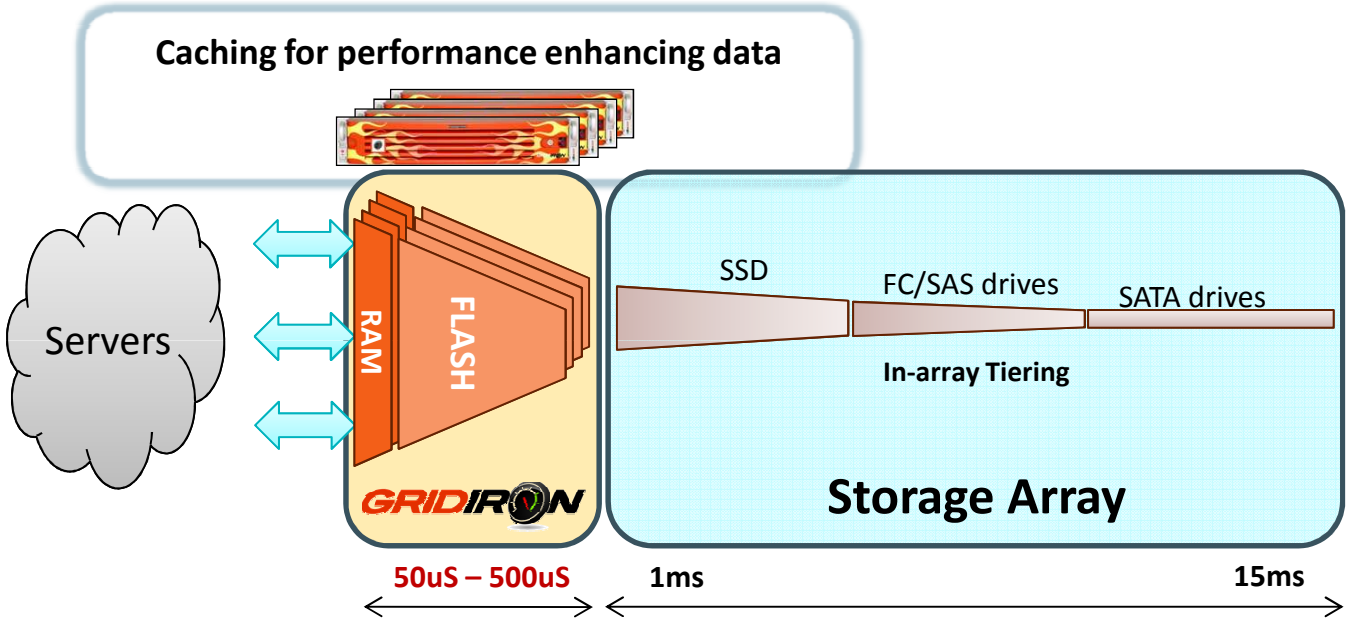
Real-Time Tiering Enables High Concurrent Bandwidth

Acceleration in the Network

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

The Learning Process

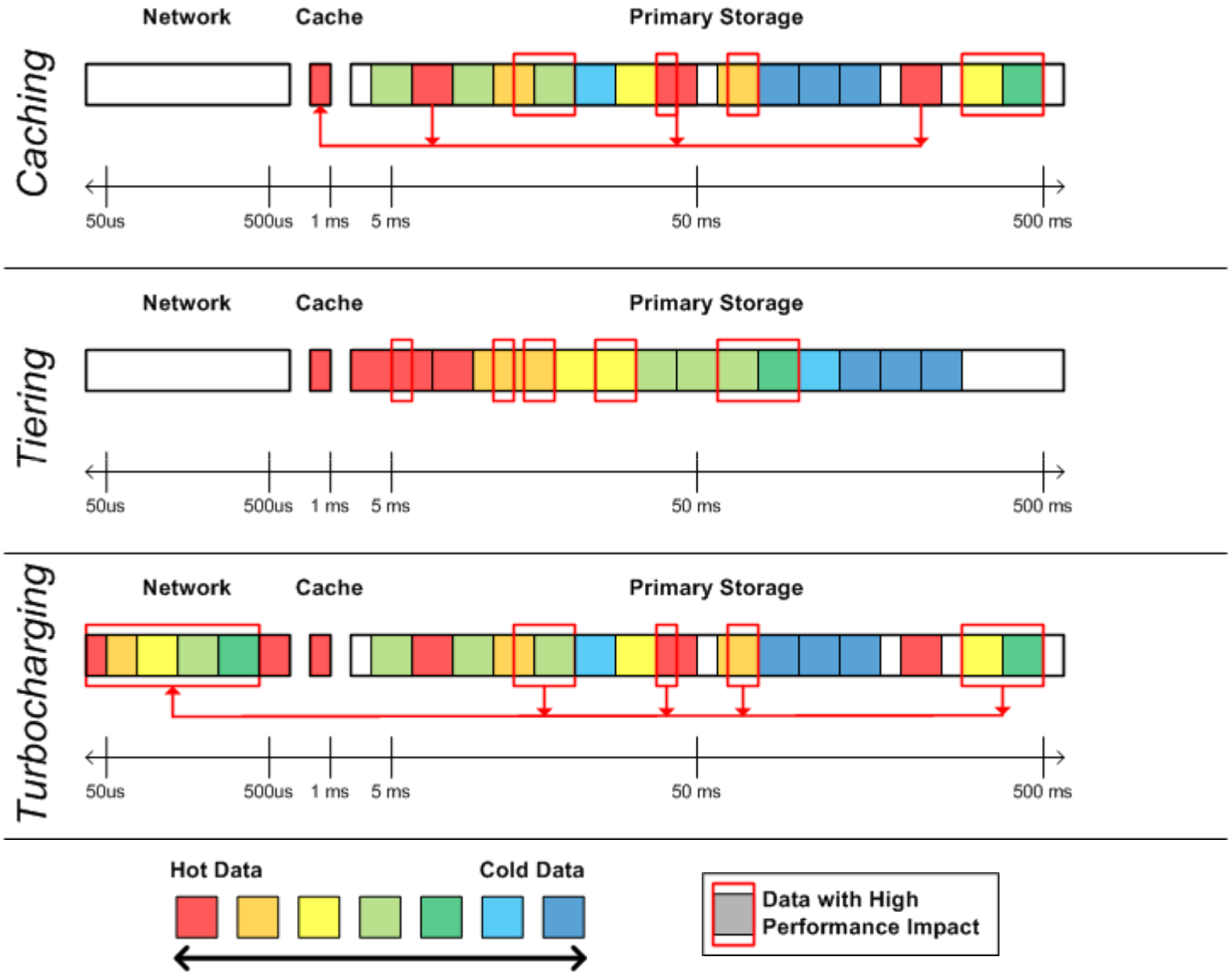
- Learn data access graphs in real time
- Use patterns to manage caching
- Use feedback to continuously refine performance



Results

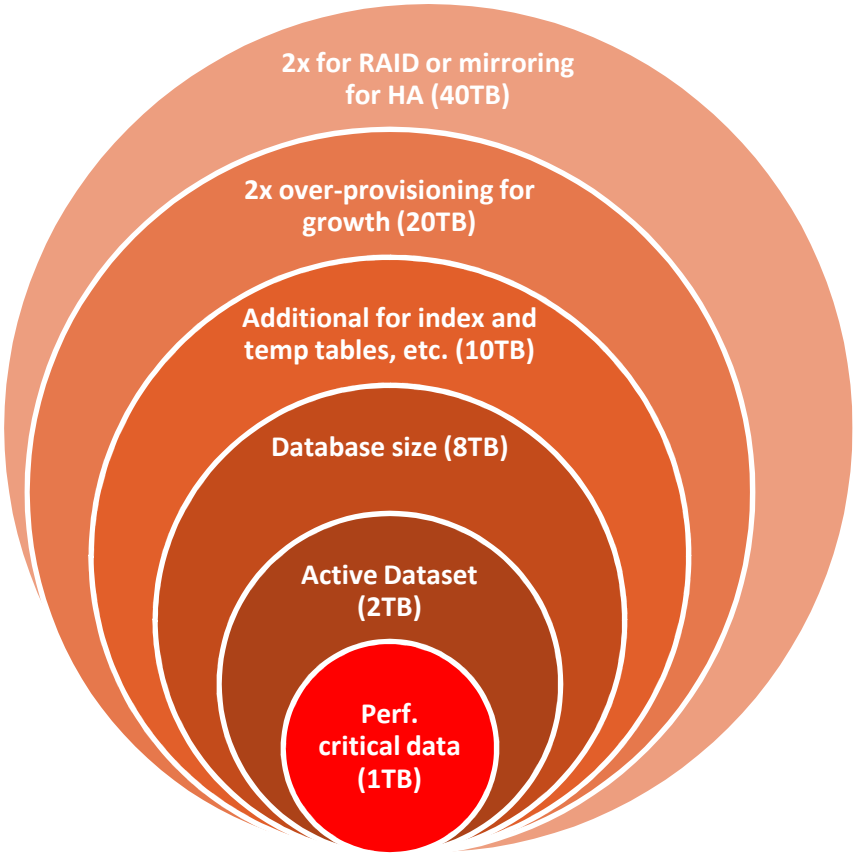
- **Applications 2-10x faster**
- **Read latency 10-100x shorter**
- **Flash life a non-issue**

TurboCharging Database Storage



Network Solution Provisioning vs. Dataset (8TB Example)

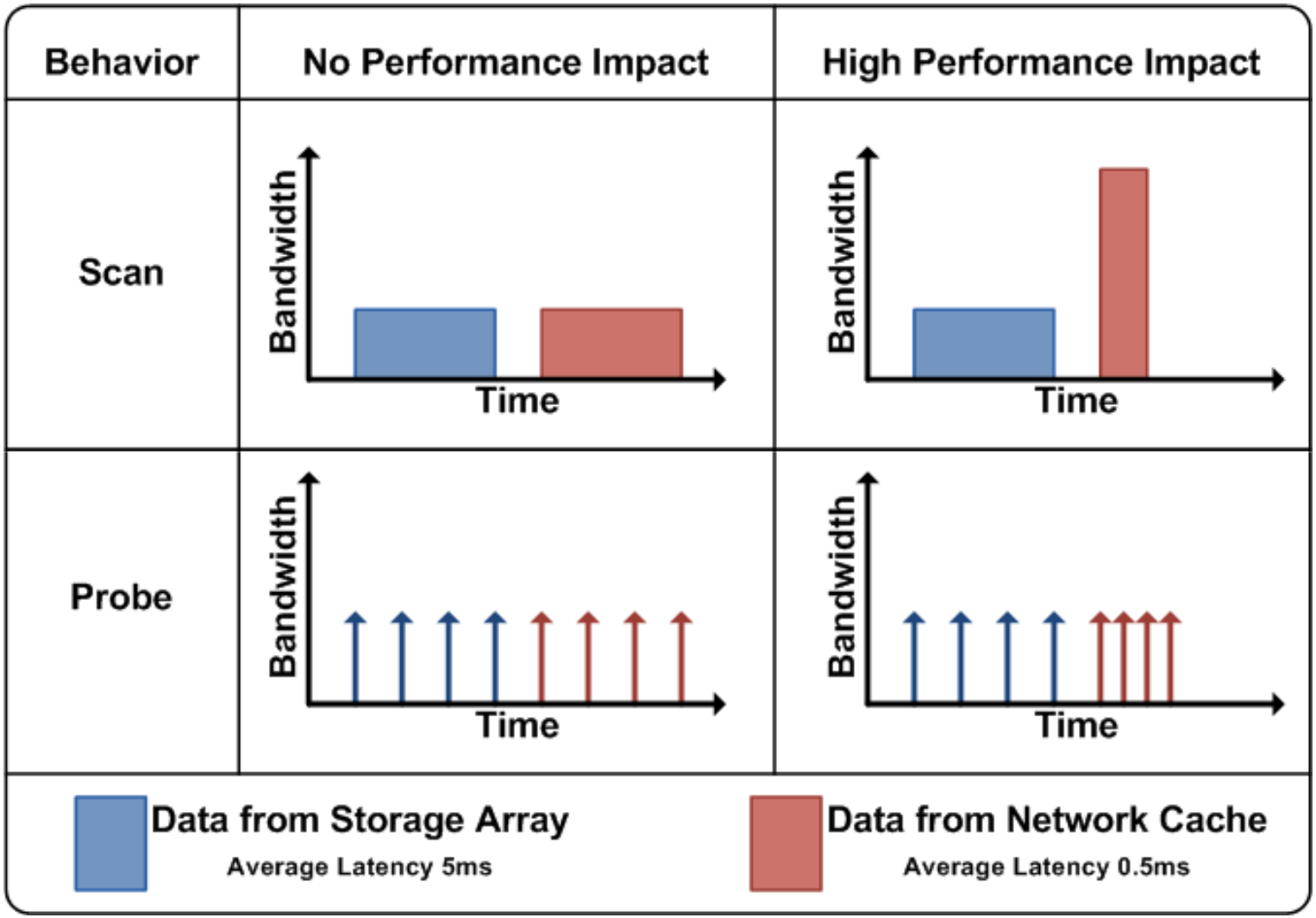
- **Smart Flash in the Network**
  - Sized for a fraction of dataset
  - Adapts in real-time to changes in usage and scale
  - Is shareable among servers, applications and arrays
  - Is always coherent with backend storage state
  - Requires no changes to applications or data management processes
- **Overcomes 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



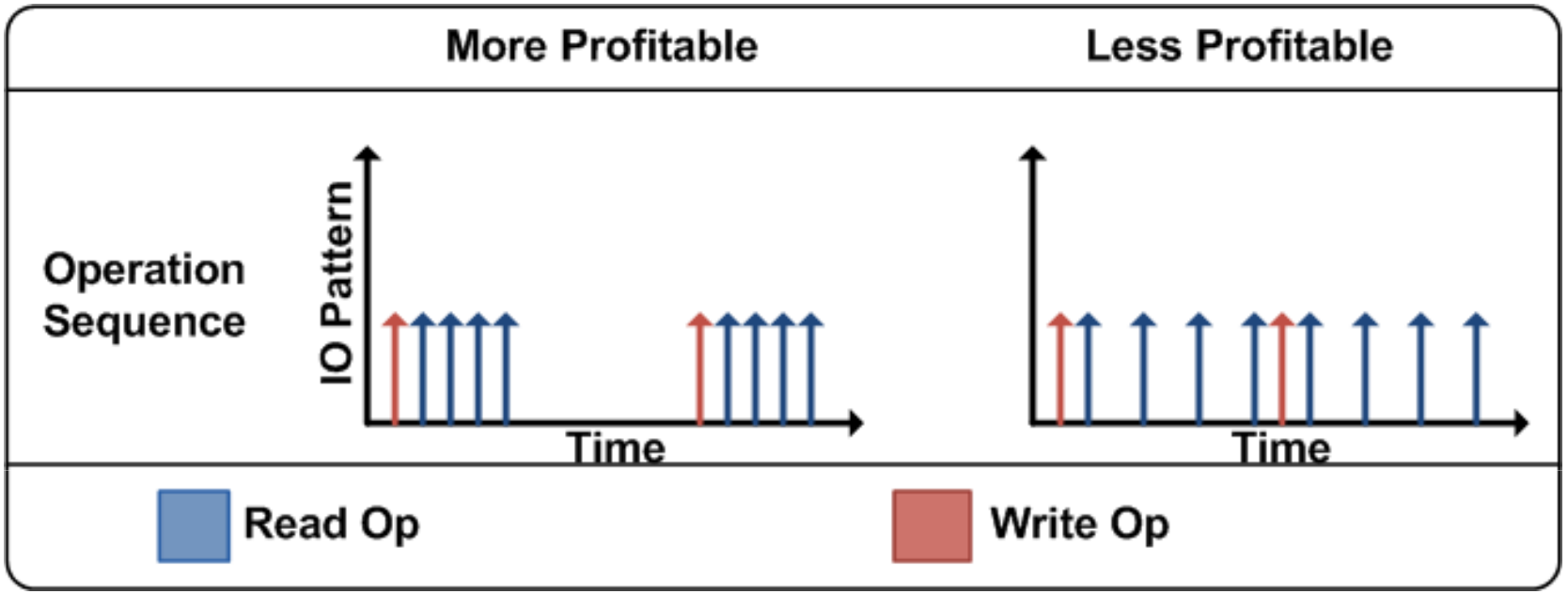
Flash Control and Effectiveness

- **Network Cache is 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

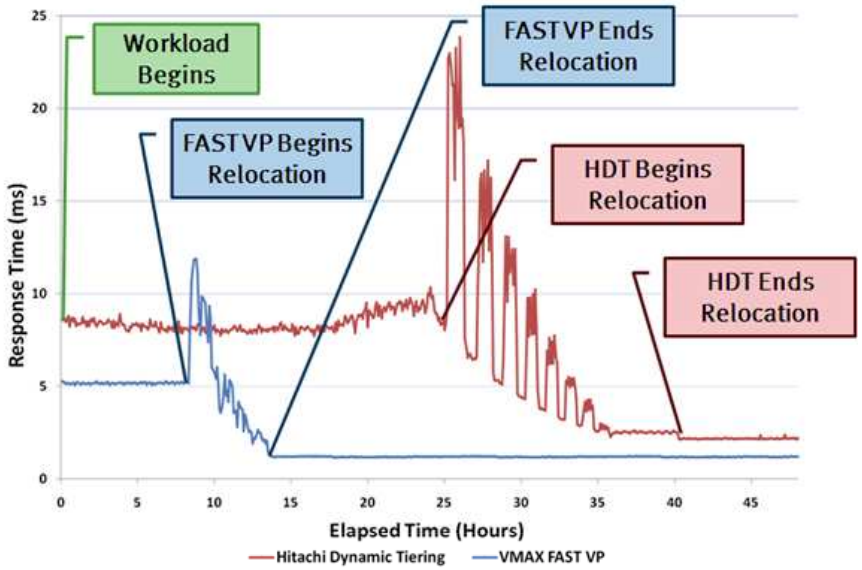


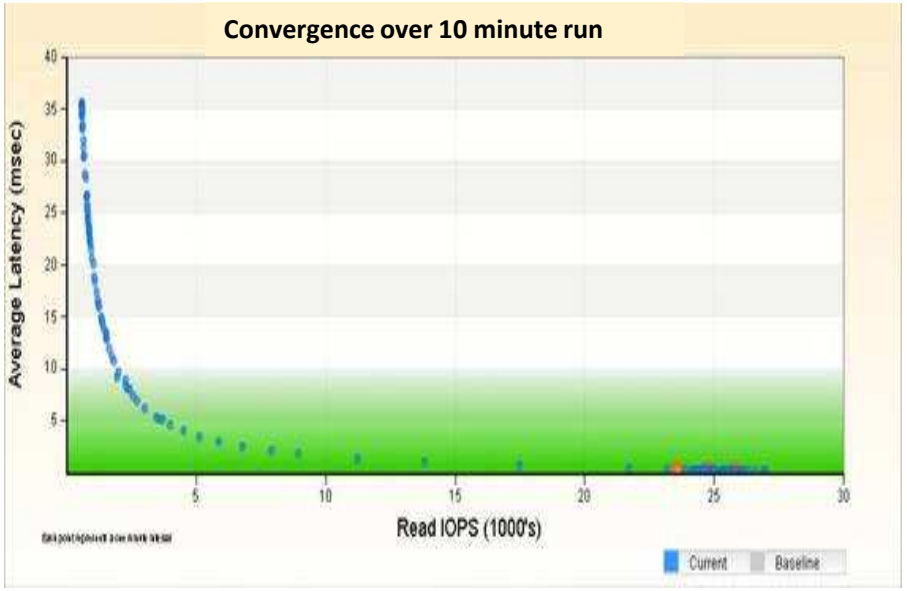
Selected Profitability Examples for Database Operations



- **Read count is a poor metric**
  - Need to consider write count and read-read delta time
  - Every cache eviction is lost storage performance
  - Opportunity cost of waiting for read hits can be high

Extending In-Array Tiering to Real-Time





1. **Above graph is a comparison between EMC FAST and Hitachi Dynamic Tiering**
2. **After 9 hours, FAST started relocating data and completed in14 hours. Response time improved from 5ms to 1ms. During data relocation, latencies more than doubled**
3. **After 1 minute, GridIron started improving both response times & IOPS and completed in 10 minutes. GridIron took latencies & IOPS from 35ms and 200 IOPS and improved them to latencies of 0.1ms and 25,000 IOPS**
4. **Note that GridIron improves latency as well as throughput beyond the physical capabilities of the storage array**

Source: http://thestorageanarchist.typepad.com/weblog/2011/04/4001-when-you-say-tiering-do-you-mean-degradation.html

Successful Proof Points in Multiple Real World Applications

Customer	Application / Business Problem	Big Data Characteristics	GridIron Performance Improvement	CapEx Savings With GridIron
shopzilla	**Oracle Data Warehouse** • "Real time" reports taking 6 hours • Lost revenue from delays • Over-provisioning storage for performance	• Bandwidth: >10GB/s • Concurrency: 25+ users • Data set size: 40TB DWH • Data turnover: continuous ETL	• Critical reports **6 hrs -> 30 mins.**	• $2M from storage and server consolidation
REGULUS A 3i Infotech Company 3i Infotech Innovation • Insight • Integrity	**Oracle Data Warehouse** • User complaints due to missed SLAs • Storage struggling to service complex queries • Massive applications overwhelming storage systems resulting in poor performance	• Concurrency: Multiple applications sharing storage	• **4x improvement** in IOPS • **5x reduction** in latency	• $1M from storage life extension and use of lower cost SATA drives for capacity expansion
COCC Financial Technology Solutions	Hosted financial services apps based on **MS SQL** • Slow DataMart transaction analytics reports • Meeting SLAs with hosted clients • Cost-prohibitive to dedicate infrastructure per hosted client	• Concurrency: Several applications and hosted customers interacting with each other	• **3x improvement** in Data Mart response times • **2x increase** in hosting capacity	• Savings of $1,2M
GILEAD				
ACTIVISION HIGHMOON STUDIOS	Large eDiscovery - **MS SQL under VMware** • Multi-hour query times affecting productivity • Need to support concurrent users • Serialized system impacting business	• Bandwidth: >2 GB/s • Concurrency: 4 users	• Query Times reduced **by >50%** • Increased query capacity **by 6x**	• Saved $775K on a storage upgrade (only 2x)
	Video game **software builds** under **VMware** • Builds taking 70 minutes to complete • Game quality impacted by long build time • Virtualized architecture not scalable	• IOPS: 40,000 Random • Concurrency: 24 users with parallel builds	• Build time reduced from **70 -> 8 mins.**	• $800K vs. alternatives

Case Study: 40 TB DWH For Online Comparison Shopping S

Challenges

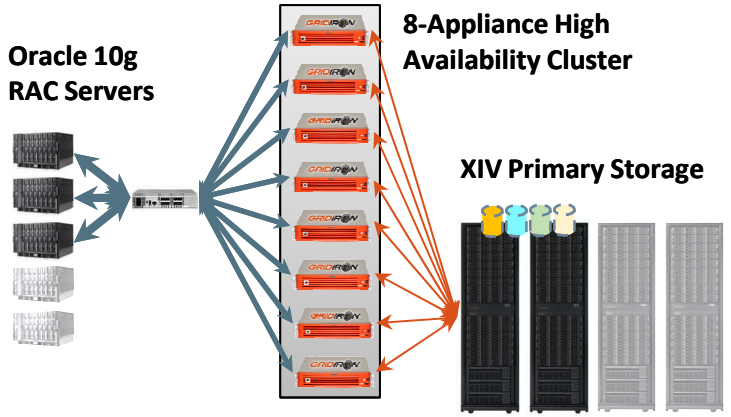
- Customer behavior analytics cycle taking too long (six hours) directly impacting revenue optimization
- Lost revenue from delays in fixing anomalies in customer-facing infrastructure
- Prohibitive storage acquisition and management costs from rapid data growth

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: DWH on Microsoft SQL in a Hosted Environmer

Challenges

- Slow response times of DataMart transaction analytics reports
- Meet SLAs with hosted clients using shared infrastructure
- Cost-prohibitive to dedicate infrastructure to hosted clients

Benefits

- 3x improvement in DataMart response times
- Exceeded SLAs with hosted clients
- 2x increase in hosting capacity
- Savings of $1,275,000
- Happy clients from better user experience

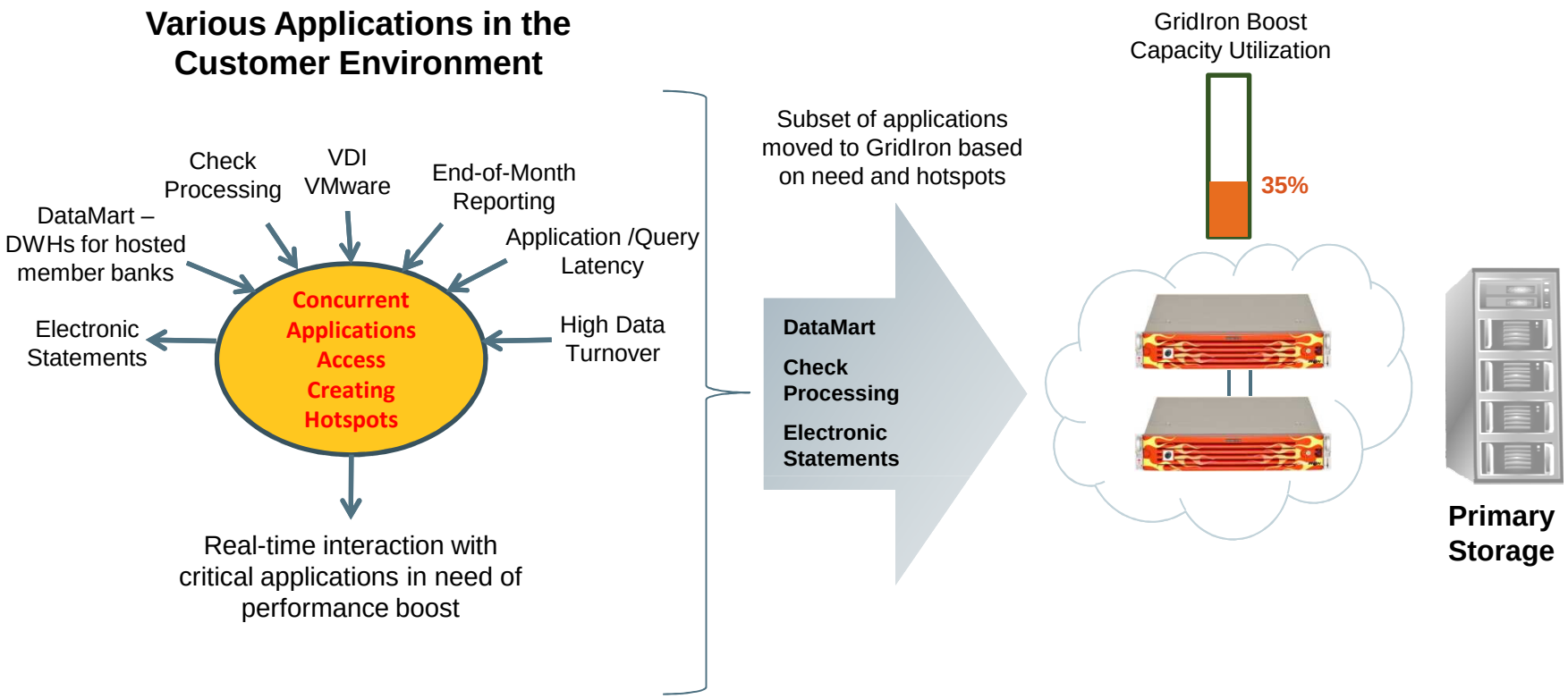
Environment

- **Storage:** Sun Storage 6540 Array
- **Servers:** Dell 1950 and 1850 servers with 4 processors
- **Application:** DataMart Transaction Analytics Reporting with Microsoft SQL
- **FC Fabric:** Brocade 48000
- **GridIron:** **Two GT-1100A TurboChargers in an active-active high-availability cluster**

"GridIron enabled us to exceed the SLAs with our hosted clients without any upgrades to our hosting infrastructure."

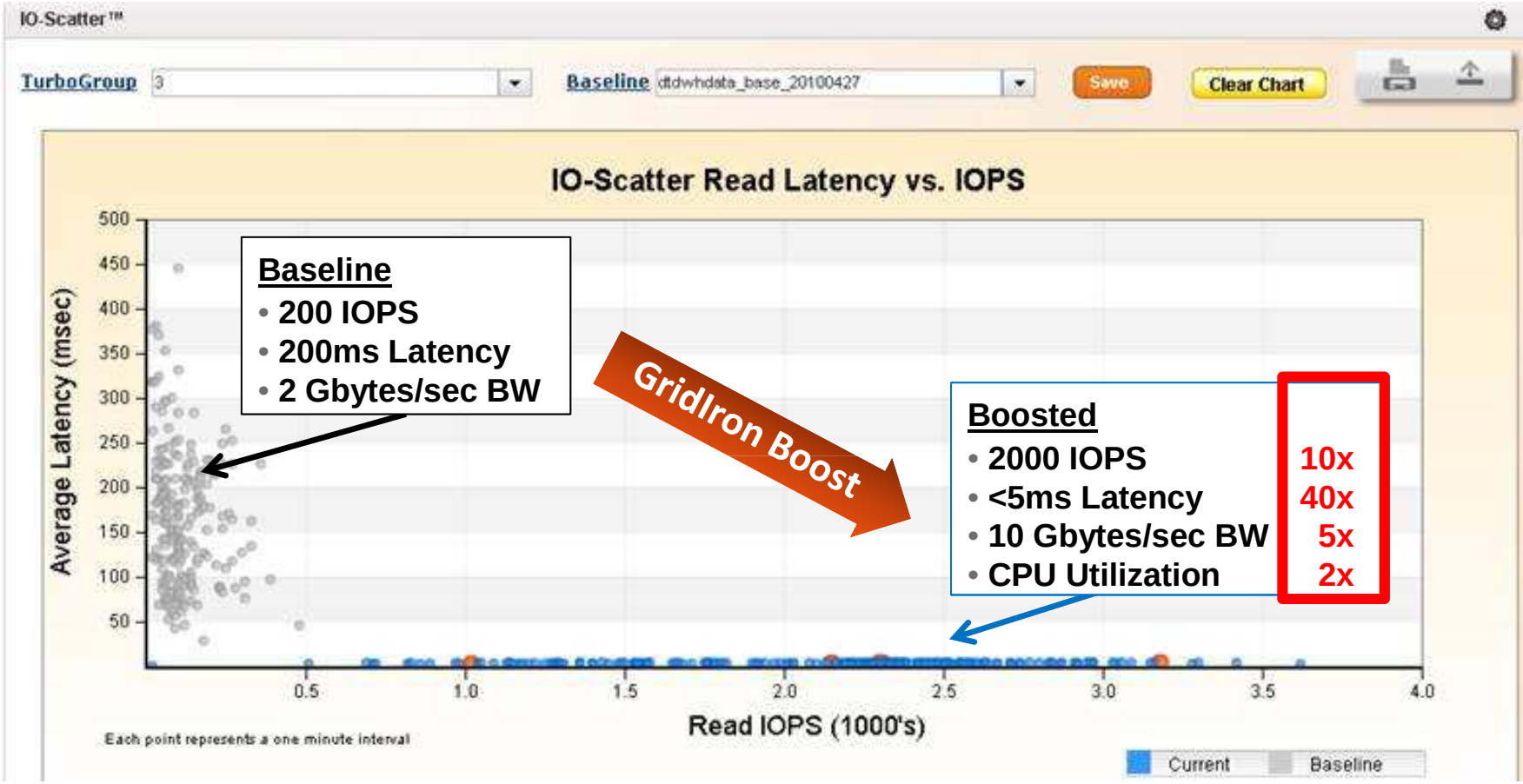
Mary Sokolowski, Storage Architect, COCC

Leverage Investment Across Multiple Applications



- Add performance where you need it, when you need it
- Deliver concurrent, sustained performance across multiple applications
- Fix performance hotspots in minutes, without changing apps or infrastructure

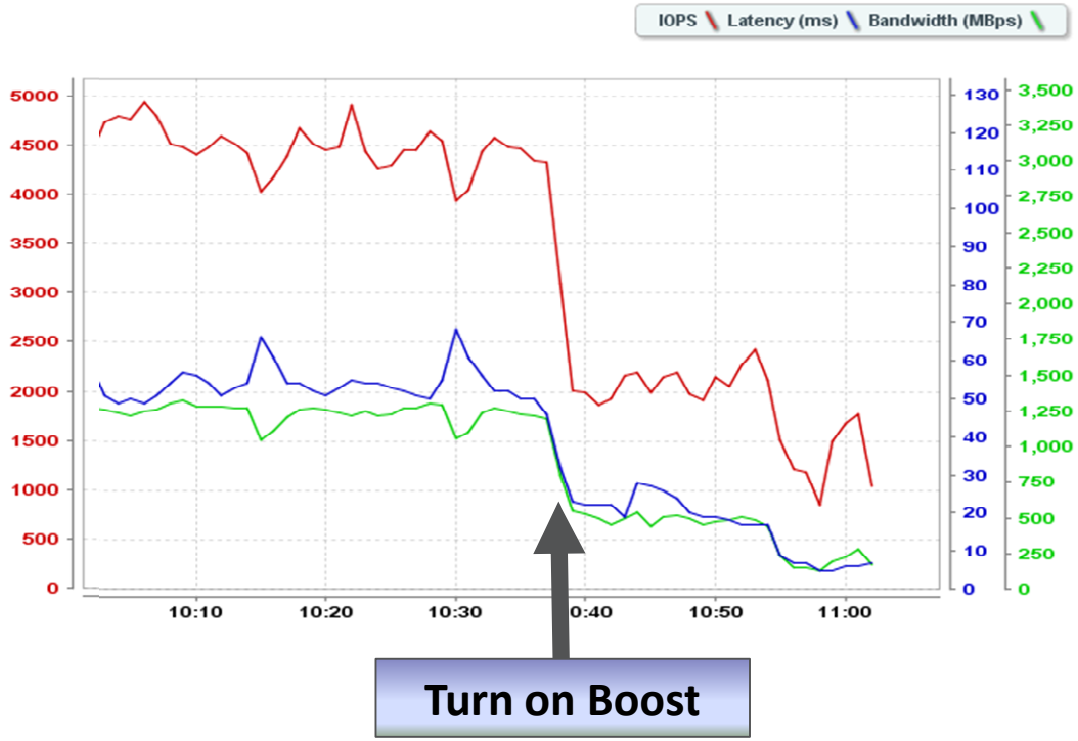
Improve Overall System Performance



Multiple applications can concurrently access the same array without interference

Dramatically Decreases Load on Back-end Storage

- Backend storage array performance improves dramatically with GridIron



- 4 x lower IOPS
- 7x lower latency
- 6x lower bandwidth

Storage controller has more bandwidth for writes and other tasks

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

Realize the true performance potential of Oracle and Oracle RAC by eliminating the IO bottleneck

Questions?