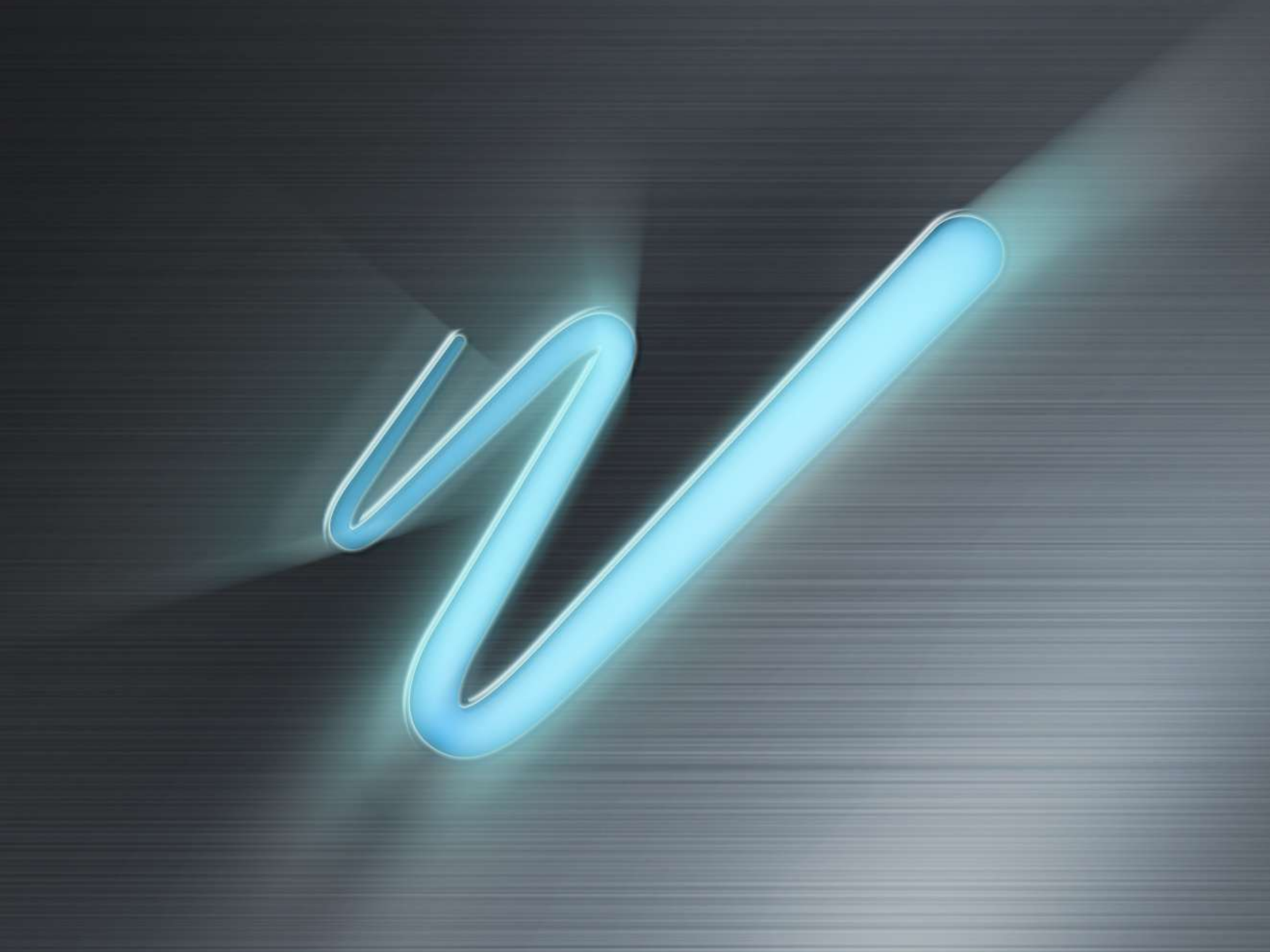


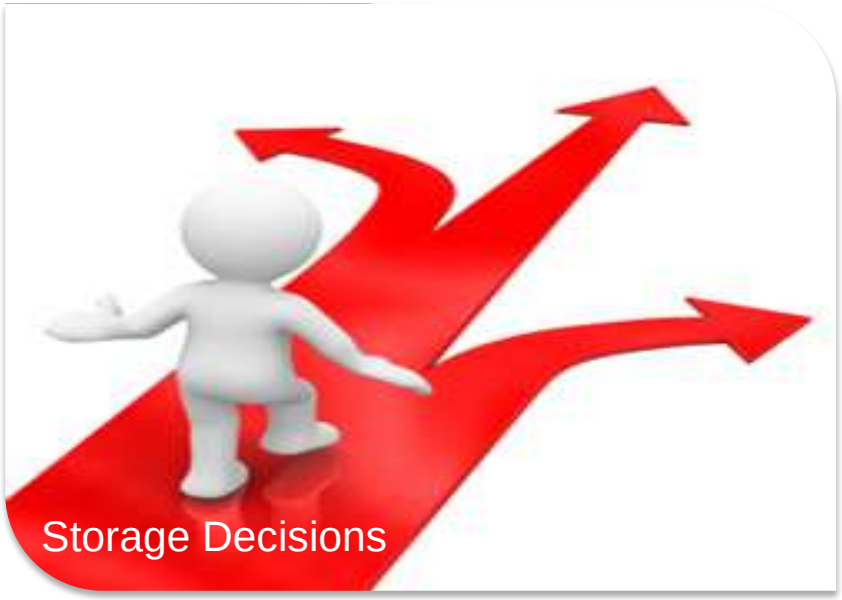
Achieving Extraordinary Workload Performance via Hyperscale Solid State Storage

STORAGE DECISIONS

BUSINESS
THE WORKLOAD CHALLENGE

Accelerate Workloads

Strategic Objectives



IT Costs

ACCELERATING WORKLOADS MEANS ACCELERATING BUSINESS

Analytics & Intelligence

High IOPS

Batch Processing

High IOPS
High Bandwidth

Email

Online Transaction Processing

High IOPS
Low Latency

Video Transcoding

High IOPS & Bandwidth

Virtual Desktops

High Write IOPS

Databases Loads

High IOPS

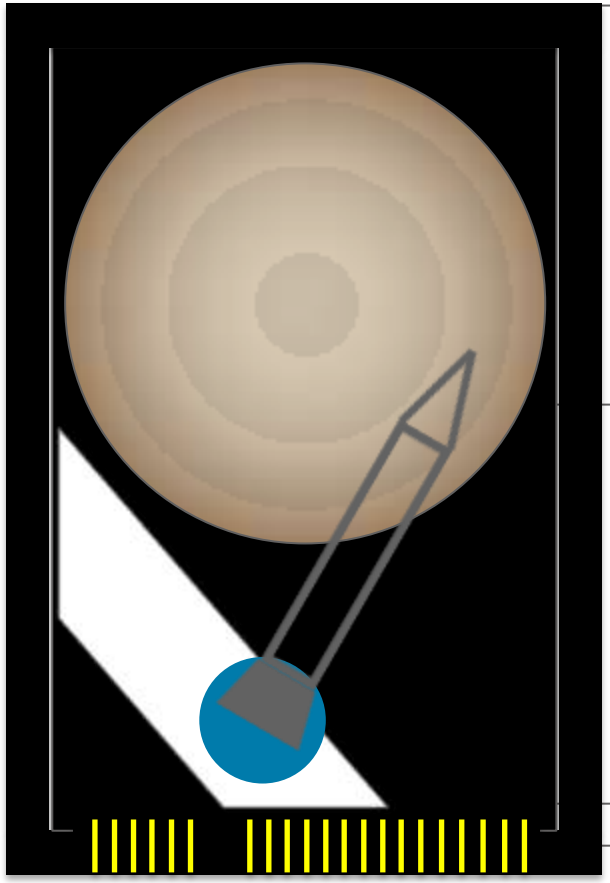
High Performance Computing

Low Latency

WRITE TO THE POINT

THE I/O BALANCING ACT.

HISTORICALLY HDDS DRIVE APPLICATION PERFORMANCE



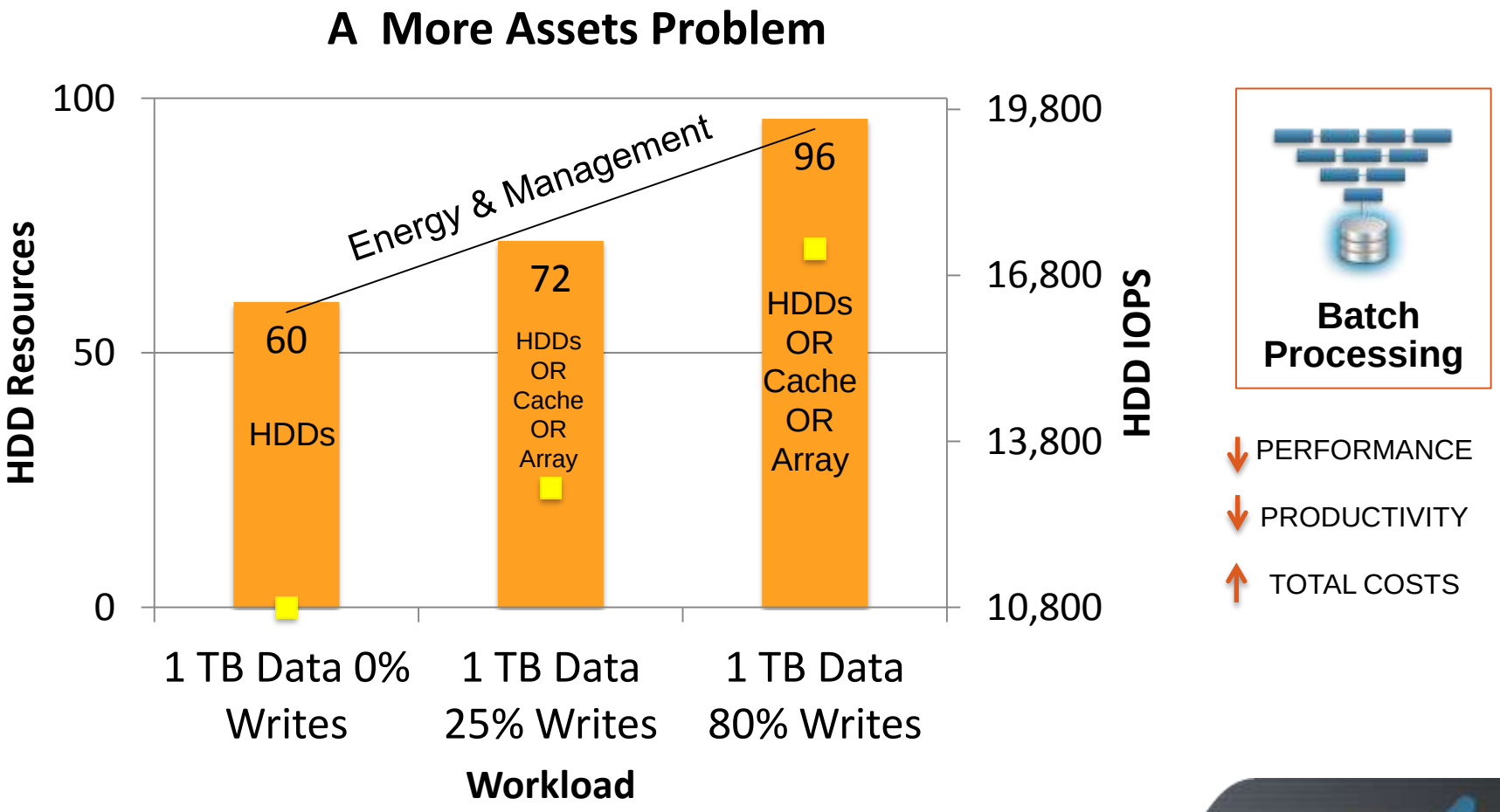
Speed

- 10s of MB/s Data Transfer Rates
- 100s of Write / Read operation per second
- .001s Latency (ms)

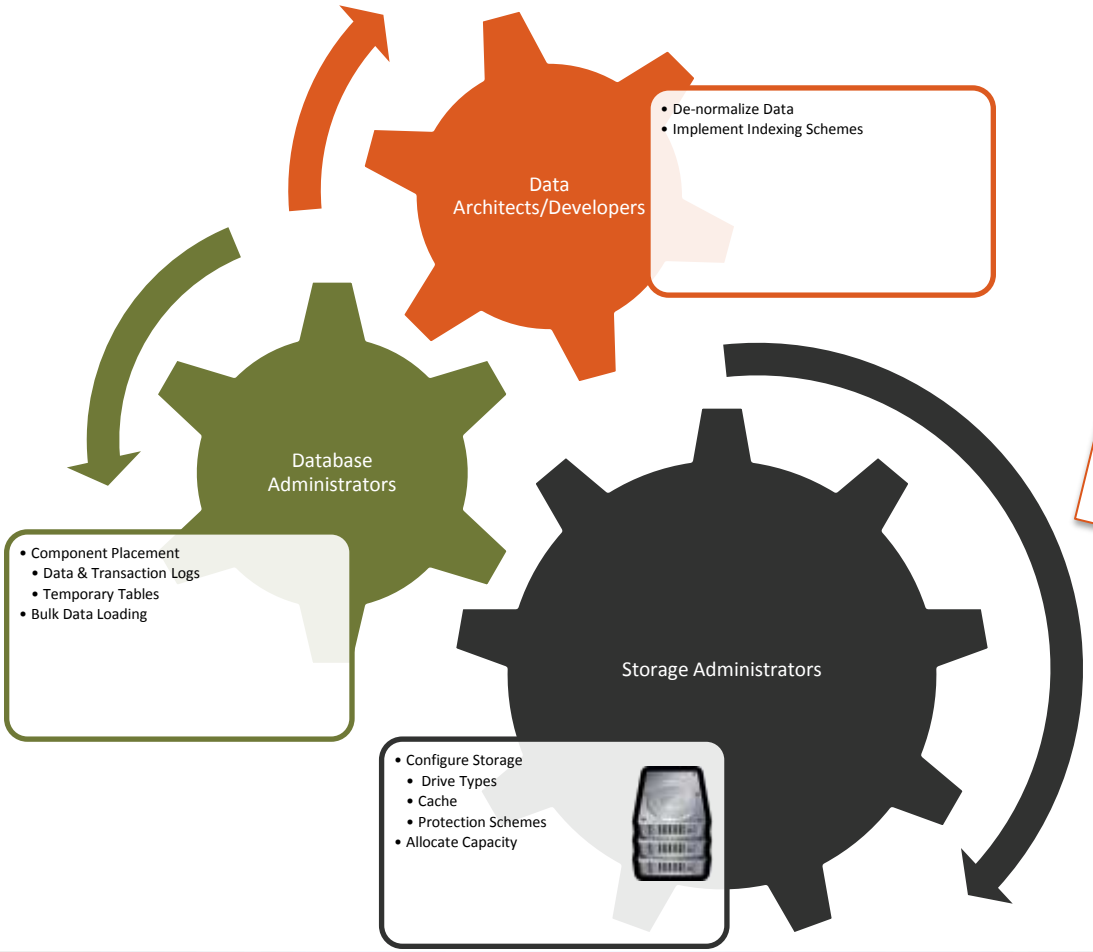
Design

- Motors
- Spindles
- High Energy Consumption

TODAY, WORKLOAD ACCELERATION IS CONSTRAINED BY WRITES



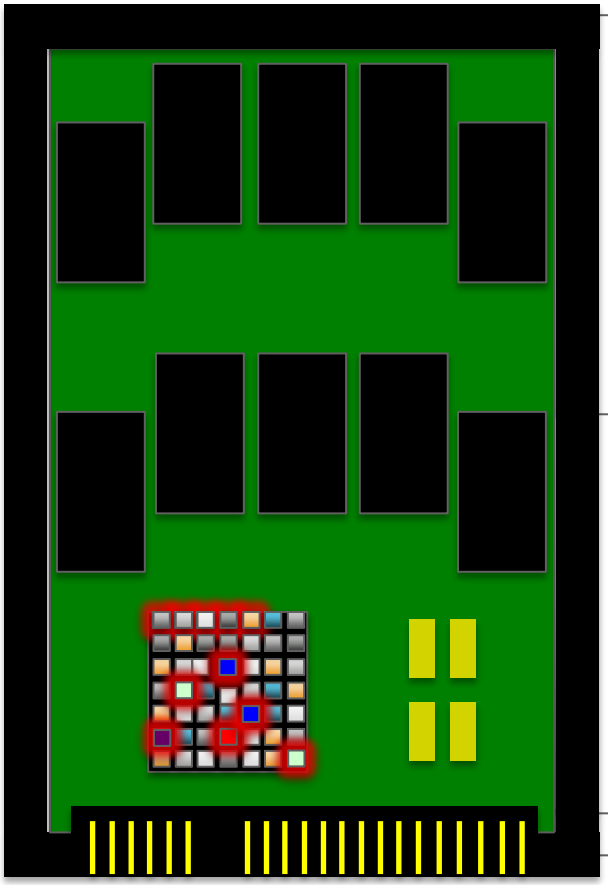
ACCELERATING WORKLOADS DRIVES IT EFFORT



IT is focused on managing Application Performance

DO THE WRITE THING

FLASH IMPROVES WORKLOAD PERFORMANCE



Speed

- 100s of MB/s data transfer rates
- 1000s of Write or Read operations per second
- .000001 Latency (µs)

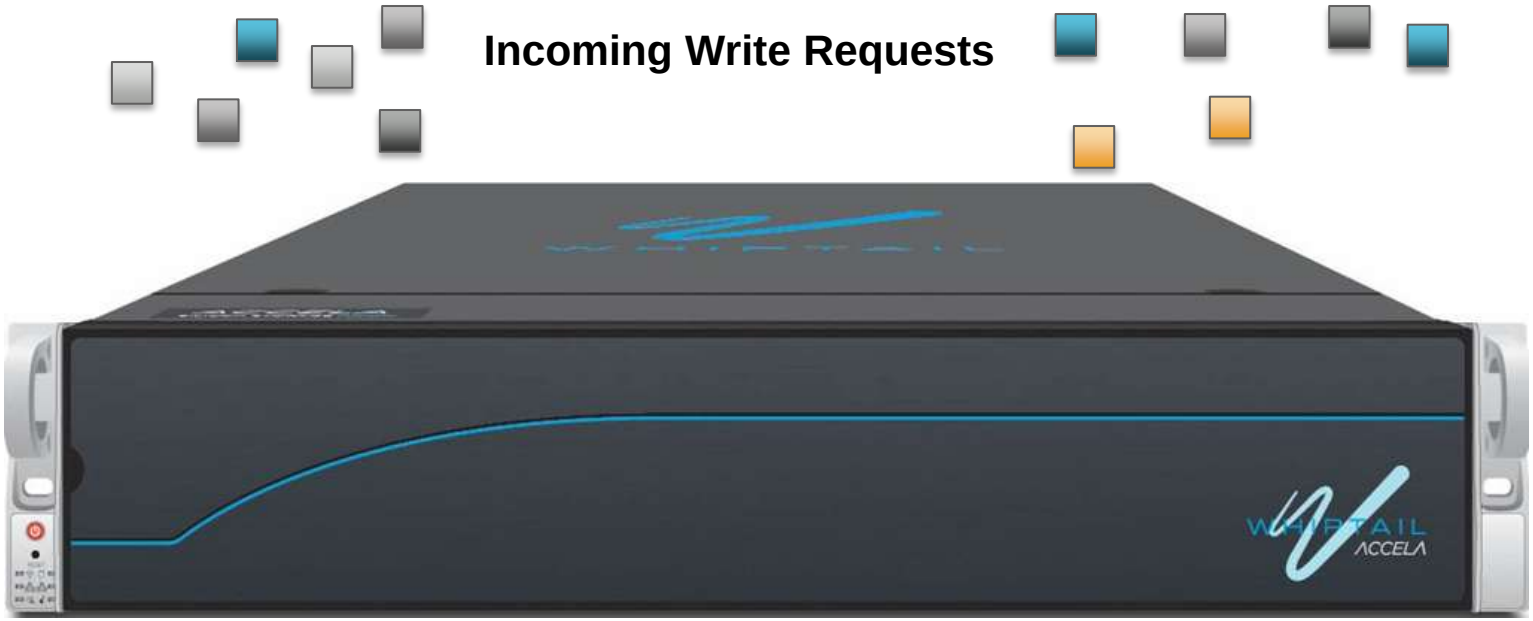
Design

- Silicon
- NAND Flash
- Low energy consumption

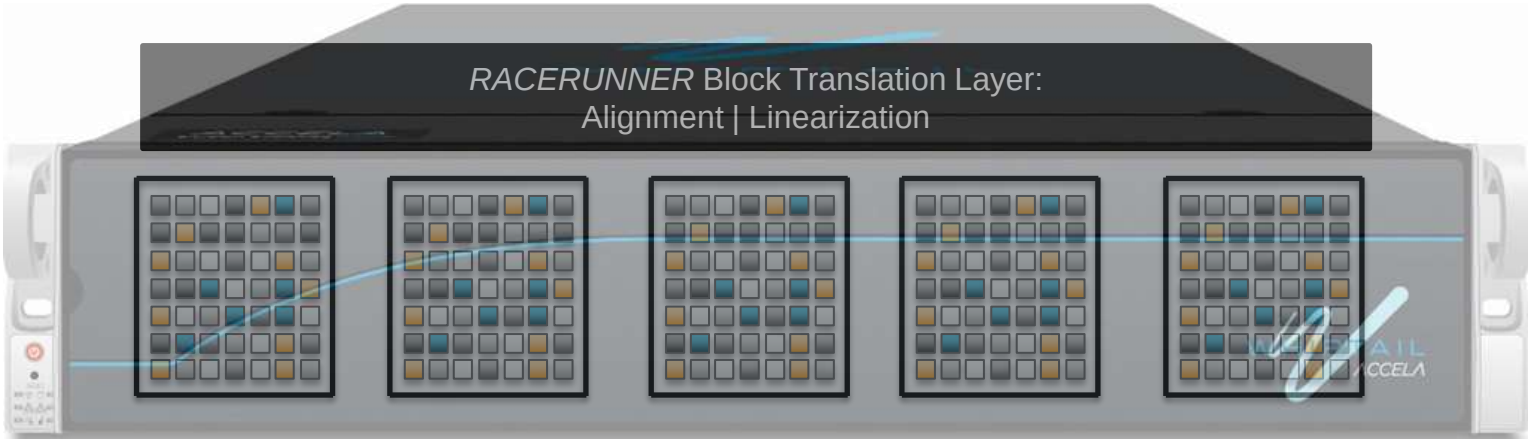
THE BLOCK TRANSLATION LAYER MANAGING FLASH WRITES

- Mask Asymmetric Write Behavior (Page Erase)
- Logical block Addressing to Physical Mapping
- Garbage Collection
- Bad Block Management
- Align & Distribute Writes

RACERUNNER OS: ALIGNS WRITES TO NAND & RAID

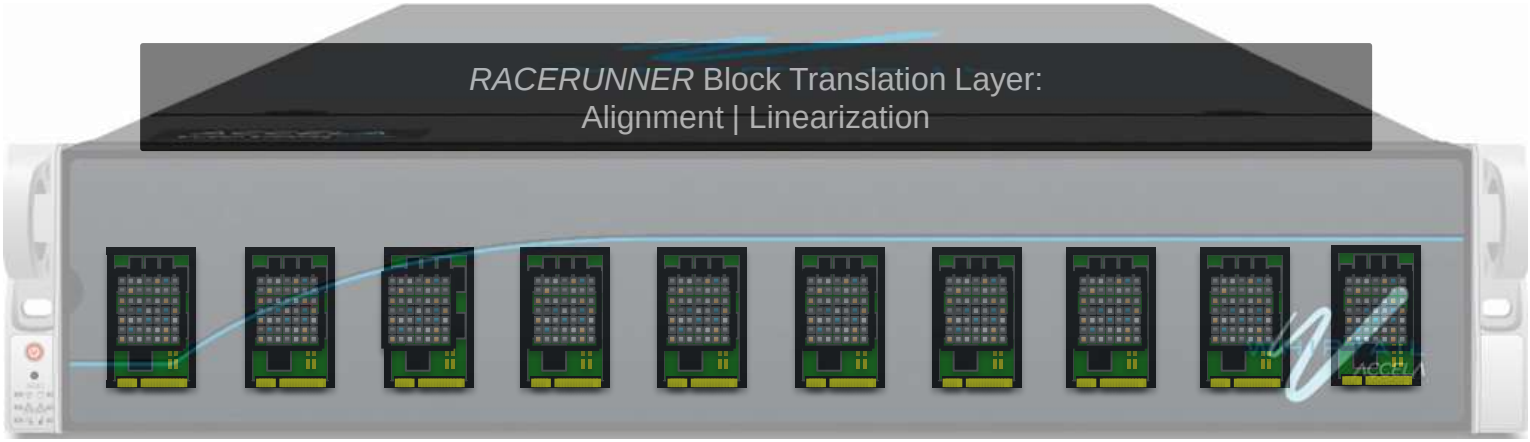


RACERUNNER OS: ALIGNS WRITES TO NAND & RAID



Write requests are aligned in native NAND page size before passed to RAID.

RACERUNNER OS: ALIGNS WRITES TO NAND & RAID



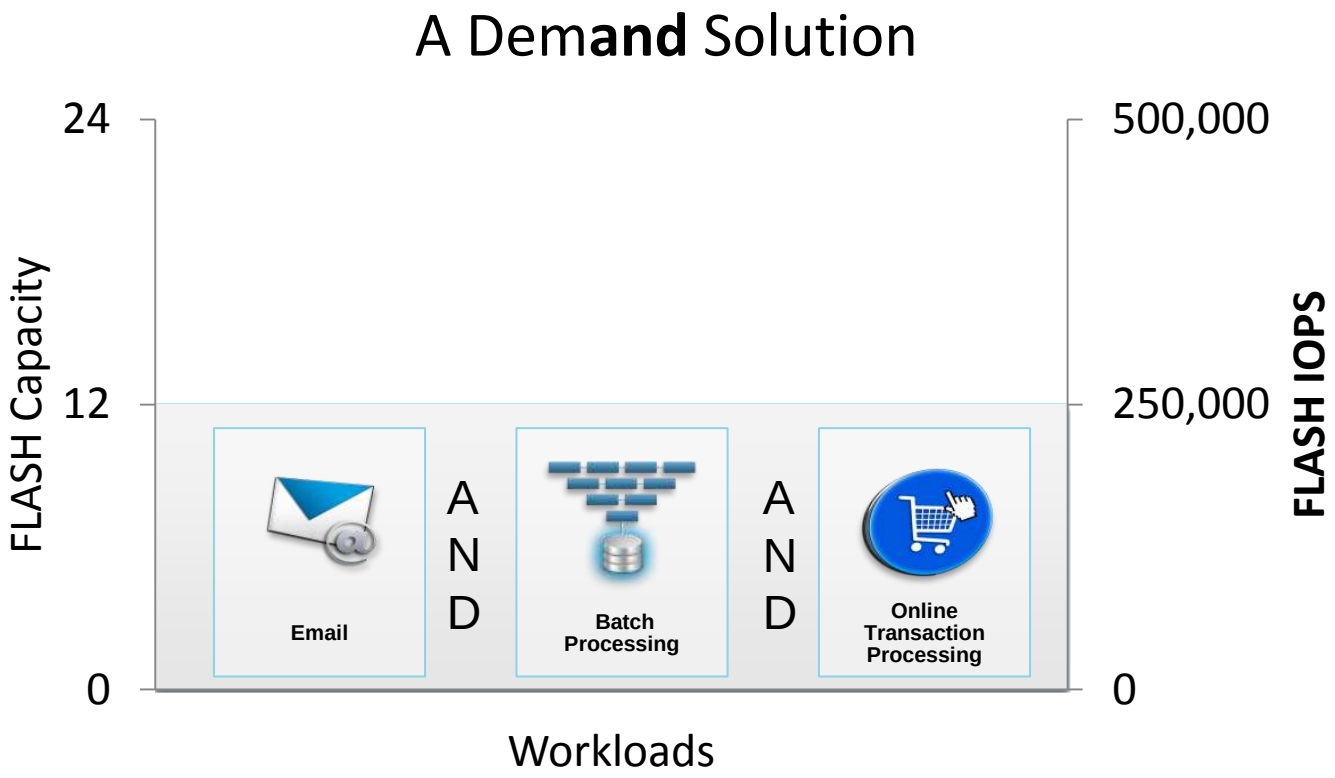
Write requests are flushed to media as full RAID stripes.

REMOVE WRITE CONSTRAINTS AND ACCELERATE WORKLOADS

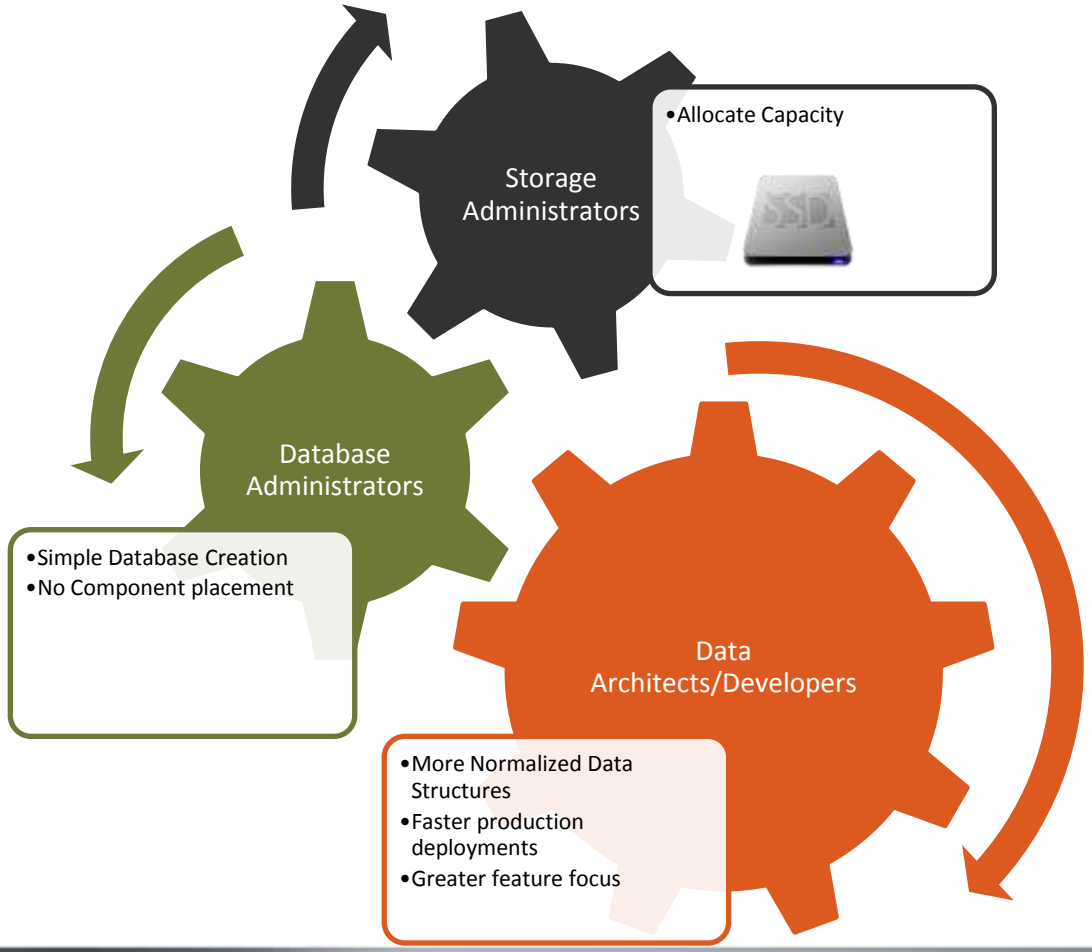
PERFORMANCE ↑

PRODUCTIVITY ↑

TOTAL COSTS ↓



ACCELERATED WORKLOADS DRIVES BUSINESS EFFECTIVENESS



IT is Focused on Application Features

MEASURING PERFORMANCE

HOW DO WE ACCELERATE APPLICATION WORKLOADS?

Increase

- **BANDWIDTH:** Sustained WRITE or READ bit-rate (GB/s or TB/h)
- **IOPS:** Input/Output Operations Per Second or Transactions
- **SYMMETRY** between Writes and Reads
- **WORKLOAD CONSOLIDATION:**

Decrease

- **LATENCY:** Milliseconds to Microseconds
- **ENERGY** : Power and Cooling.
- **CAPACITY**: eliminate overprovisioning
- **WORKLOAD TUNING:**

Velocity

- Near time
- Real time
- Streams
- Batch

RETHINKING WORKLOADS

WHIPTAIL Labs

- Reengineer Workloads for Solid State Storage
- Workload Reference Architectures
- Workload Benchmarks
- Workload Best Practices

Product Engineering

- Solid State Storage platforms
- Data management features

Research & Development

- New advancements in Solid State Storage
- New data management techniques

WORKLOAD TESTING PROCESS

- DVD Store
- TPC
  - E
  - DS
  - VMS
- Virtual Desktops

- SQLIO (SQLServer)
- ORION (Oracle)
- VDIMARK (Desktop Users)
- Yahoo! Cloud Serving Benchmark (Big Data)
- JetStress (EMAIL)

- Load data
- Export data
- Run complex Queries
- Workload Recovery

- Benchmarks
- Reference Architectures
- Best Practices

VALIDATING RESULTS

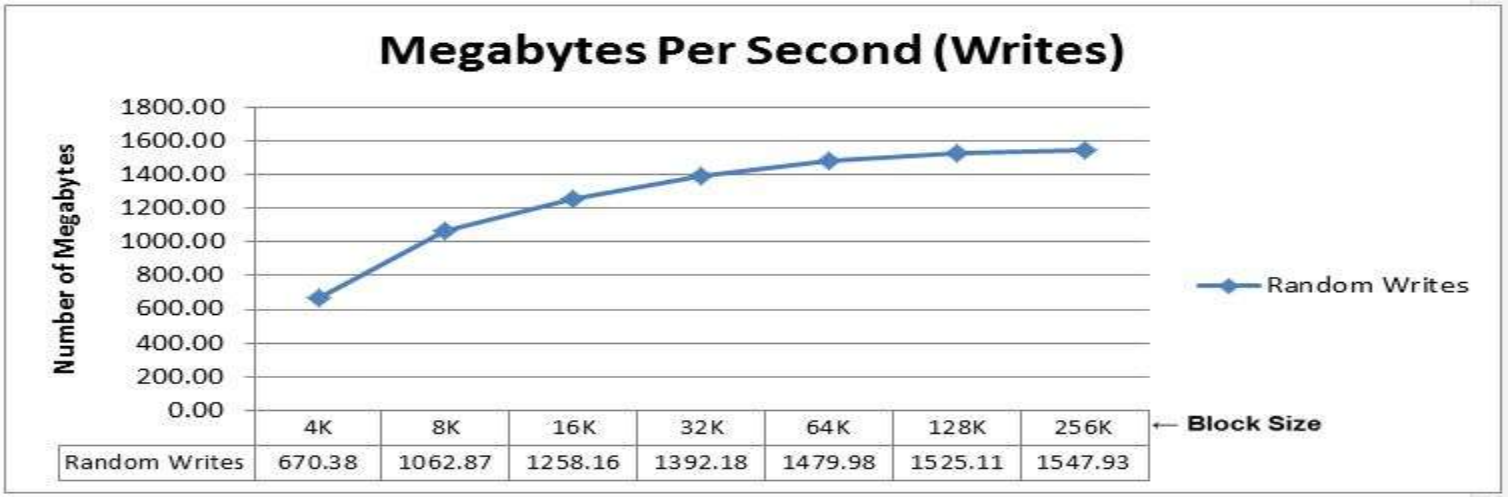
Develop Prototypes	Scale Testing	Validate
		
• Reference Architectures • Benchmarks	• Architectures • Benchmarks • Compare / Contrast	Audit Results

MULTI-WORKLOAD TESTING

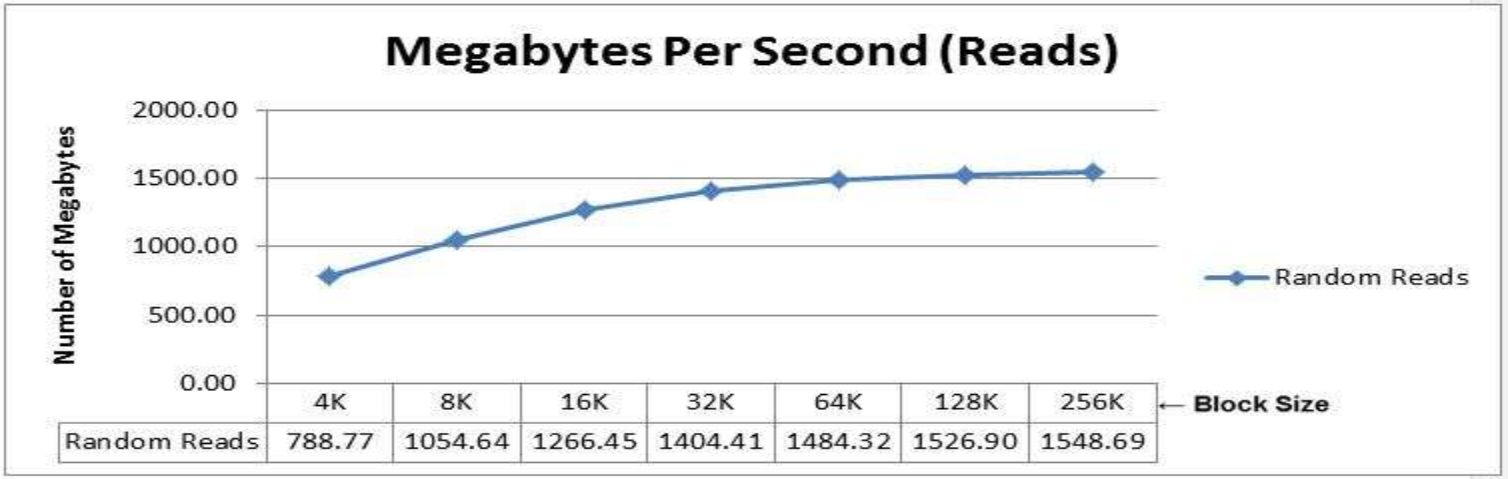
LUNS

OLTP Workload 100% WRITES	SQLIO_01	10.00G	SSRA
	SQLIO_02	10.00G	SSRB
	SQLIO_03	10.00G	SSRA
	SQLIO_04	10.00G	SSRB
Month end Reporting workload 100% READS	SQLIO_05	10.00G	SSRB
	SQLIO_06	10.00G	SSRA
	SQLIO_07	10.00G	SSRB
	SQLIO_08	10.00G	SSRA
	SQLIO_09	10.00G	SSRA
	SQLIO_10	10.00G	SSRB

BANDWIDTH TEST RESULTS
OLTP & REPORTING WORKLOADS

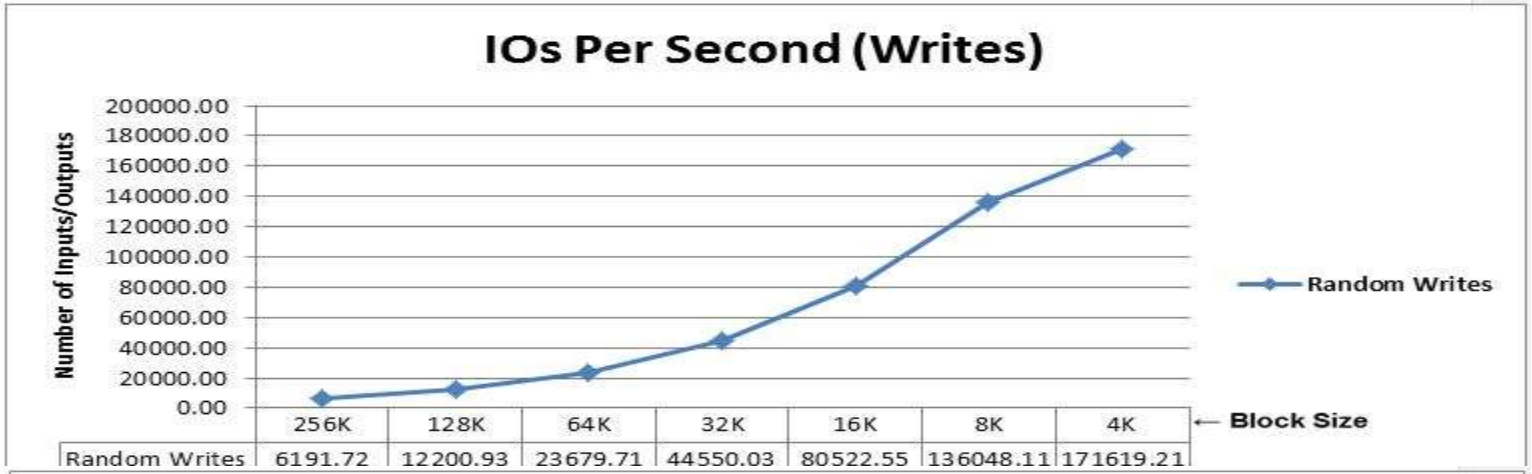


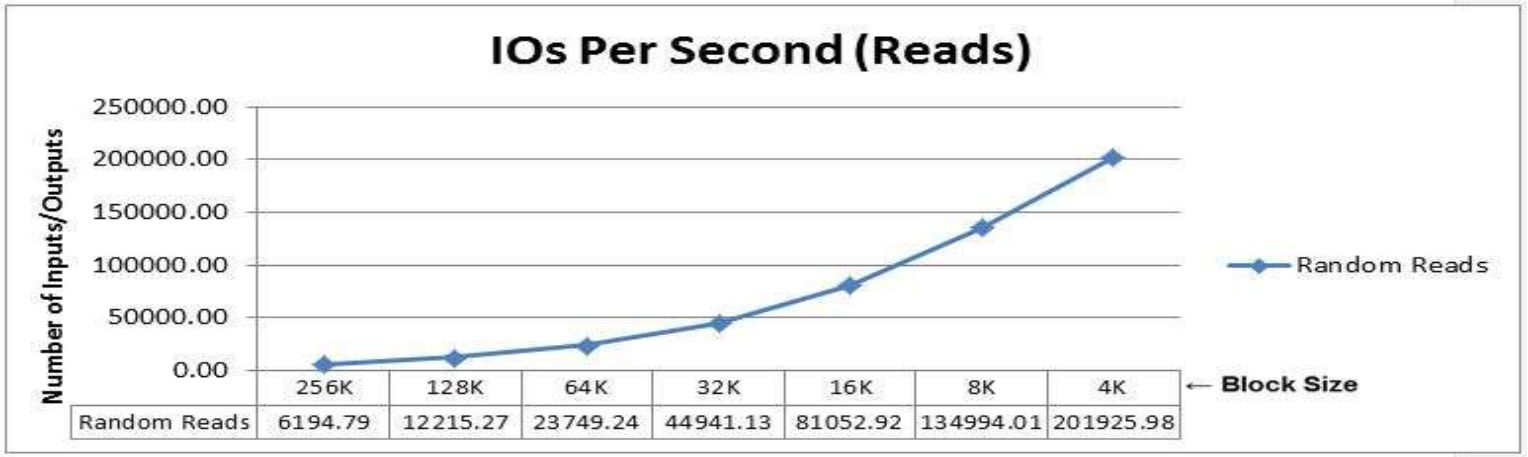
	4K	8K	16K	32K	64K	128K	256K
Random Writes	670.38	1062.87	1258.16	1392.18	1479.98	1525.11	1547.93



	4K	8K	16K	32K	64K	128K	256K
Random Reads	788.77	1054.64	1266.45	1404.41	1484.32	1526.90	1548.69

IO TEST RESULTS
OLTP & REPORTING WORKLOADS





MULTI-WORKLOAD REFERENCE ARCHITECTURE

Mercury

Workload Engines	Workload Type	Workload Demand
Dell DVD Store MS SQL Server	1200 Transactions Per Second (Continuous)	4,000 IOPS .05 GB/s
VMWare View	600 Desktops Boot Storm (2:30)	109,000 IOPS .153 GB/s
SQLIO MS SQL Server	Heavy OLTP Simulation 100% 4K Writes (Continuous)	86,000 IOPS .350 GB/s
	Batch Report Simulation 100% 64K Reads (Continuous)	16,000 IOPS 1 GB/s
		215,000 IOPS 1.553 GB/s

Raid 5 HDD Equivalent = 3,800
RAID 10 HDD Equivalent = 2,000

- *INVICTA*
  - 350,000 IOPS
  - 3.5 GB/s
  - 18 TB
- 8 Servers

In 2012 Mercury traveled to Barcelona, New York, San Francisco, Santa Clara, and Seattle demonstrating the ability to accelerate multiple workloads on to Solid State Storage.

PERFORMANCE BENCHMARK
THE CARL SAGAN RUN

Create 4.225 Billion Records

- Randomly generated census like data
- 10 fields
- 85 bytes per record

Load records into Oracle Database

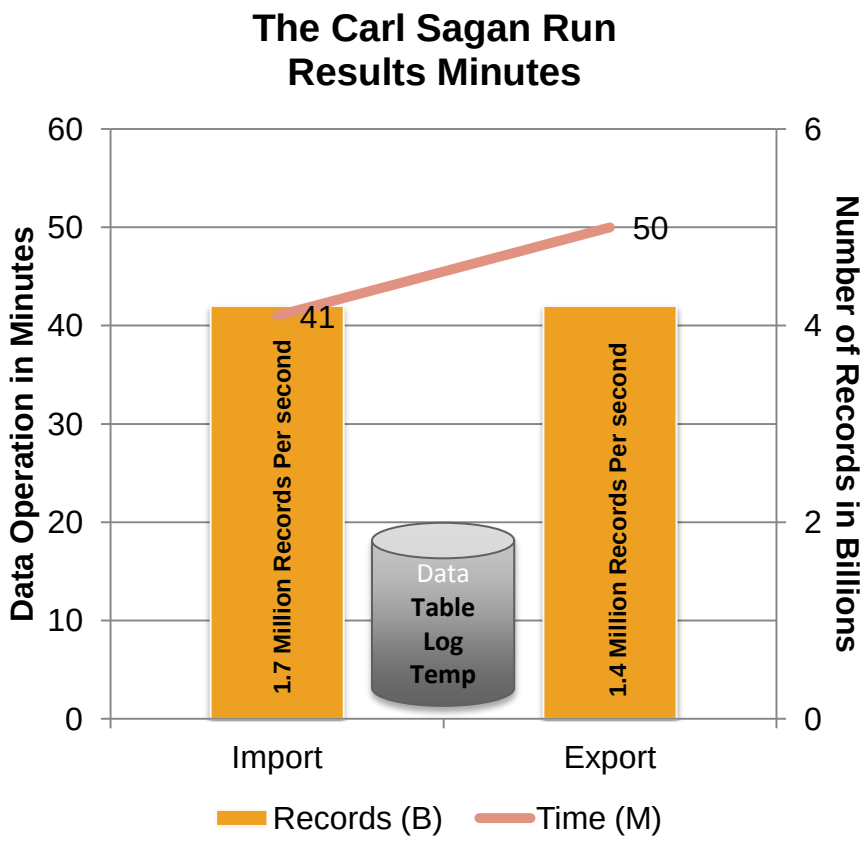
- A single 1.5TB LUN for table, temp and logs
- Tablespace .365 TB
- No Indexes
- 76,000 Writes per second
- 1.2 GB/s

Generate additional loads via ORION

- Run Heavy OLTP Simulation
- 70,000 Writes/Reads per second (40/60)
- 1 GB/s

INVICTA

- 4 Nodes
- 24 TB Capacity
- 425,000 IOPS



WORKLOADS IN THE QUEUE

Document Oriented

Key Value Pairs

Reference Architecture

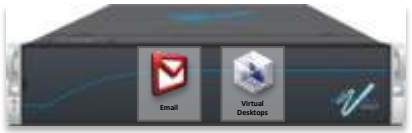
Data Warehouse
Reference Architecture

Q&A


The Evening Herald
EXTRA!

PRODUCT LINEUP
TO *INFINITY* AND BEYOND

1-9 Application Workloads



ACCELA
1.5TB – 12TB
250,000 IOPS
1.9 GB/s Bandwidth

1-60 Application Workloads

INVICTA
2-6 Nodes
6TB-72TB
650,000 IOPS
7GB/s Bandwidth

1-300 Application Workloads

INVICTA – INFINITY (Q1/13)
7-30 Nodes
86TB-360TB
800,000 – 4 Million IOPS
40GB/s Bandwidth

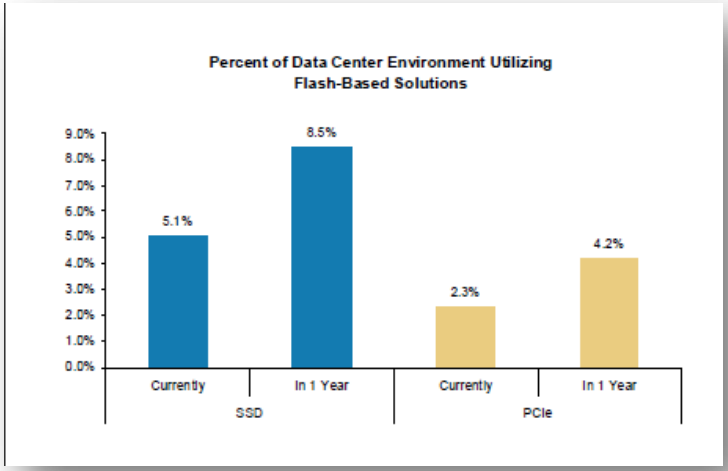
CITRIX ready

vmware READY

Scalability Path

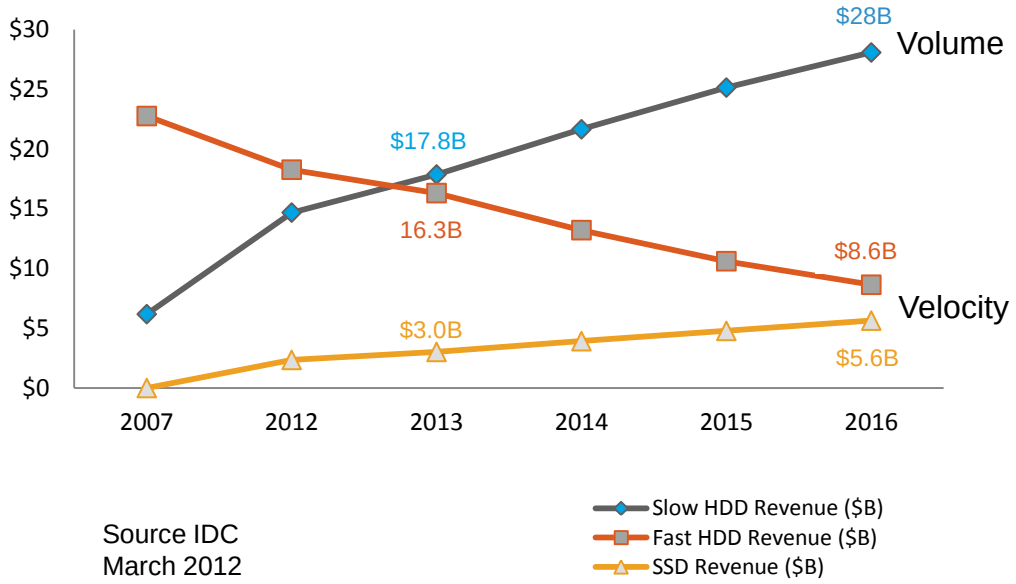
SOLID STATE IS QUICKLY GAINING FAVOR WORKLOAD PERFORMANCE

Data Center Adoption



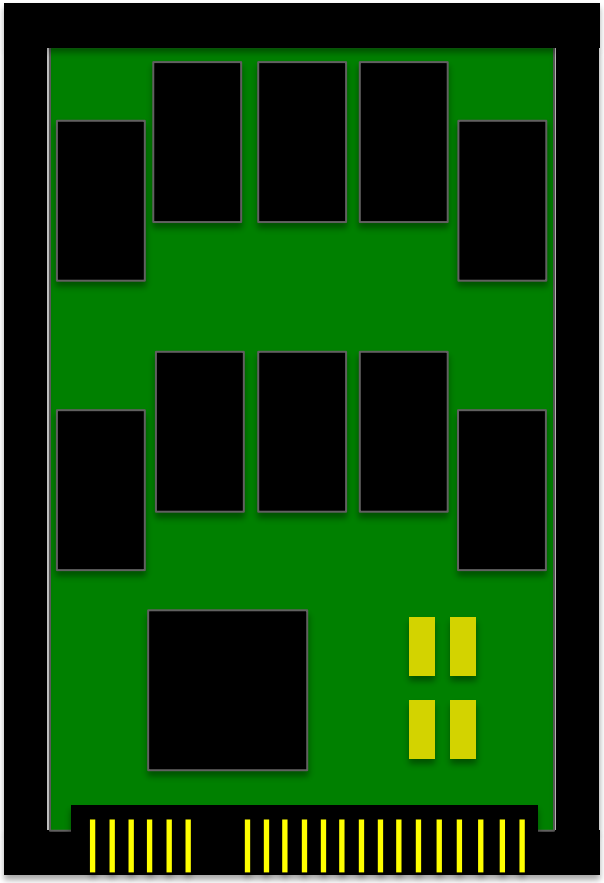
Source Morgan Stanley
April 2012

Storage Systems Revenue ($B)

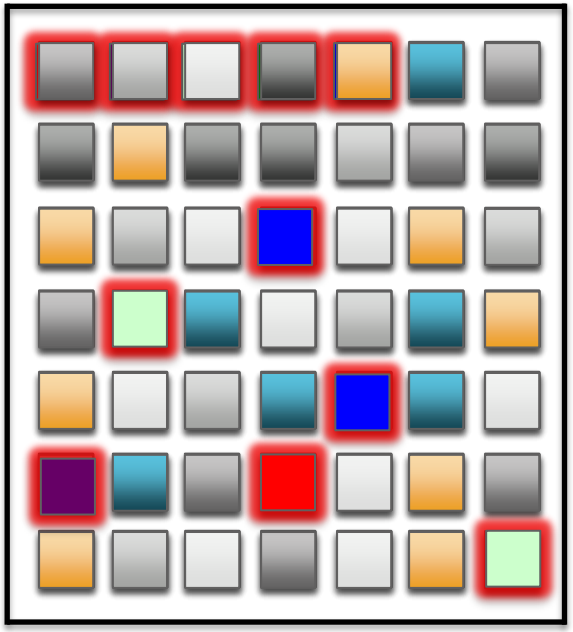


Source IDC
March 2012

NAND FLASH FUNDAMENTALS: *FLASH WRITE PROCESS*



2MB NAND Page



1. NAND Page contents are read to a buffer.

2. NAND Page is erased (aka, “flashed”).

3. Buffer is written back with previous data and any changed or new blocks – including zeroes.

RACERUNNER OS:
ALIGNING WRITES TO NAND