



Flash Memory: Key to Efficient Real-Time Analysis of Big Data

Andy Walls

IBM Fellow and CTO Flash Systems



The Flash Revolution to date

- Consumer flash in phones, tablets, USB sticks, etc has driven down prices of all flash
- Fast HDD price declines have slowed
- Flash has come into the enterprise as a replacement of performance disks



Tier 1 Disk Replacement

- Automatic Tiering gave immediate performance improvement
- All Flash Arrays have been cheaper on \$/IOP basis for a long time
- Data reduction techniques have eroded \$/GB advantage
- Flash vastly simplifies the management of a tier 1 data center
- Consolidation of application data onto one pool of storage now possible



Flash rapidly coming down in cost

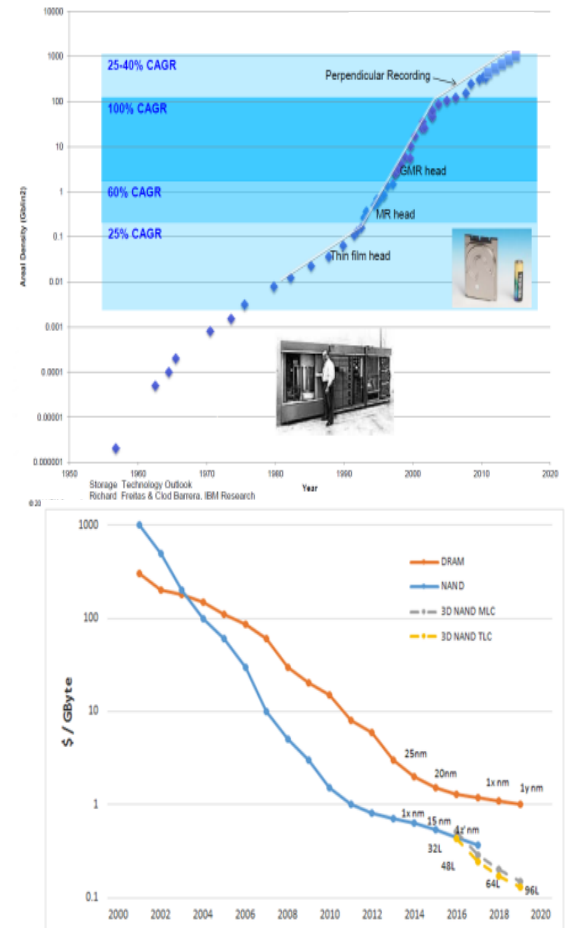
- **Flash media cost improving at ~30% CGR**

- Low latency, High IOPs, High bandwidth (No matter what)
- Fast enough to handle Data Reduction built in
- Endurance that performs well
- Hardware data path and accelerators
- Easy to Use
- Extremely reliable and available.

- **HDD improvement rate is slowing**

- And improvement coming with very large drives like 12TB
- The IOP density continues to shrink

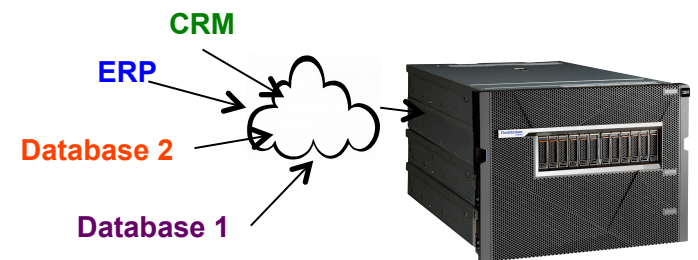
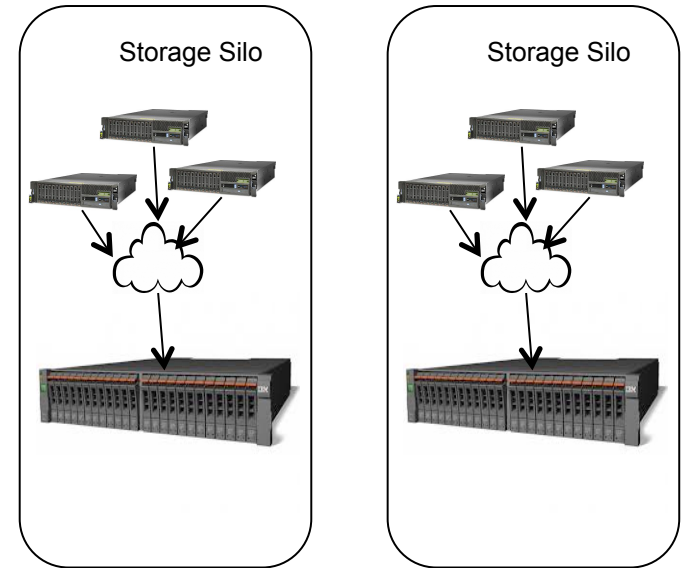
- **Tiering is giving way to All Flash Data Center**





Consolidation

- Servers and Storage often exist in silos
- New applications result in new servers and storage
- Server virtualization allows for allocating new VMs
- But it is hard to consolidate storage onto Spindles!
- Tiering is challenging for consolidation
- All Flash Arrays can finally allow consolidation





Flash Revolution Accelerating

- New workloads are demanding more and different storage
- It is no longer about just **STORING** the data and accessing it when client does something.
- Analyzing all the data becoming paramount
- Flash starting to encroach on 10K HDDs – and even near line drives



Where Flash is going

- Amazing lithography reductions
- 3DMLC will accelerate the density increases
- 3DTLC
- Dies per package increasing
- Research and investment in increasing 3D levels

Soon, you can have a data center in your pocket!



Data is Exploding

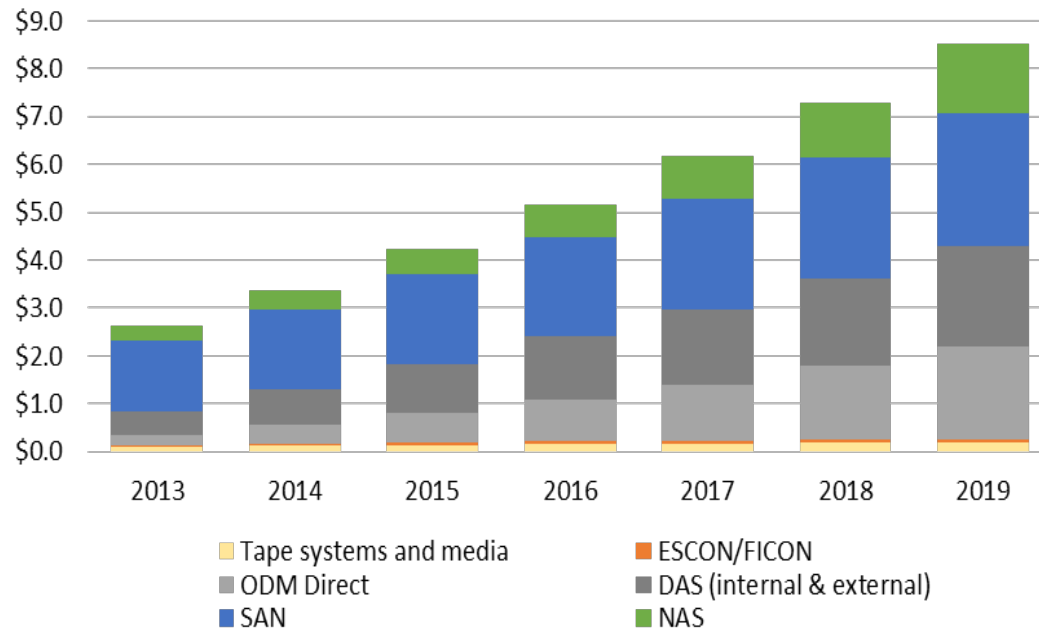
- Unstructured data is growing at an enormous rate
 - At least 40% a year
- Data is no longer a by-product of a company it is becoming the company
- What will produce all the data?
 - For a long time data was created in structured form
 - Then came social media
 - And soon it will be THINGS (The internet of things!)



Big Data Storage Revenue



“Big Data will form the foundation as businesses transform themselves into data-driven entities ...”





Making Sense out of all the Data

- Distributed (scale out) software stacks that provide extreme scalability
- No SQL databases, Hadoop, Spark
- **Providing**
- HPC Workloads – Oil and Gas, Life Science, etc
- Analyzing buying trends, providing recommendations to consumers
- Marketing analysis and optimization
- Medical record analysis





Requirements

- Analytics for data at rest
- Read Intensive in nature
- Requires high bandwidth rather than high IOPs
- Density is key
- High Performance Clustered File systems close to the storage
- Redundancy built into the file system in the form of distributed RAID
- Distributed (scale out) software stacks that provide extreme scalability



How does flash enable this

- For example, new DeepFlash 150 has 500TB in 3U – PBs on a floor tile
- Less than half the power or cooling of HDDs
- Very high throughput JBOF
- Scaleout storage software can provide Exabytes under single namespace
- Only the beginning





Now fast forward to Real Time

- Fraud Detection
- On the fly Ad insertion
- Call Data Record analysis for network optimization
- Real time traffic congestion avoidance
- Security Threats
- Real time analysis of monitors and sensors on babies to improve treatment



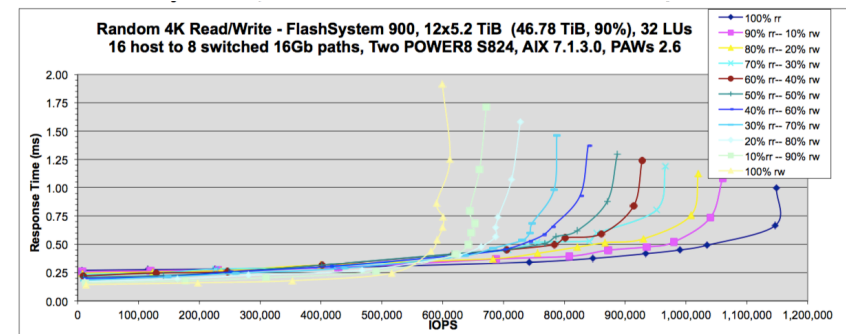
Enabling analytics

- Low latency
- High IOPs
- Read and writes
- In Memory Databases – logically expanded to include flash
- Key Value Store acceleration

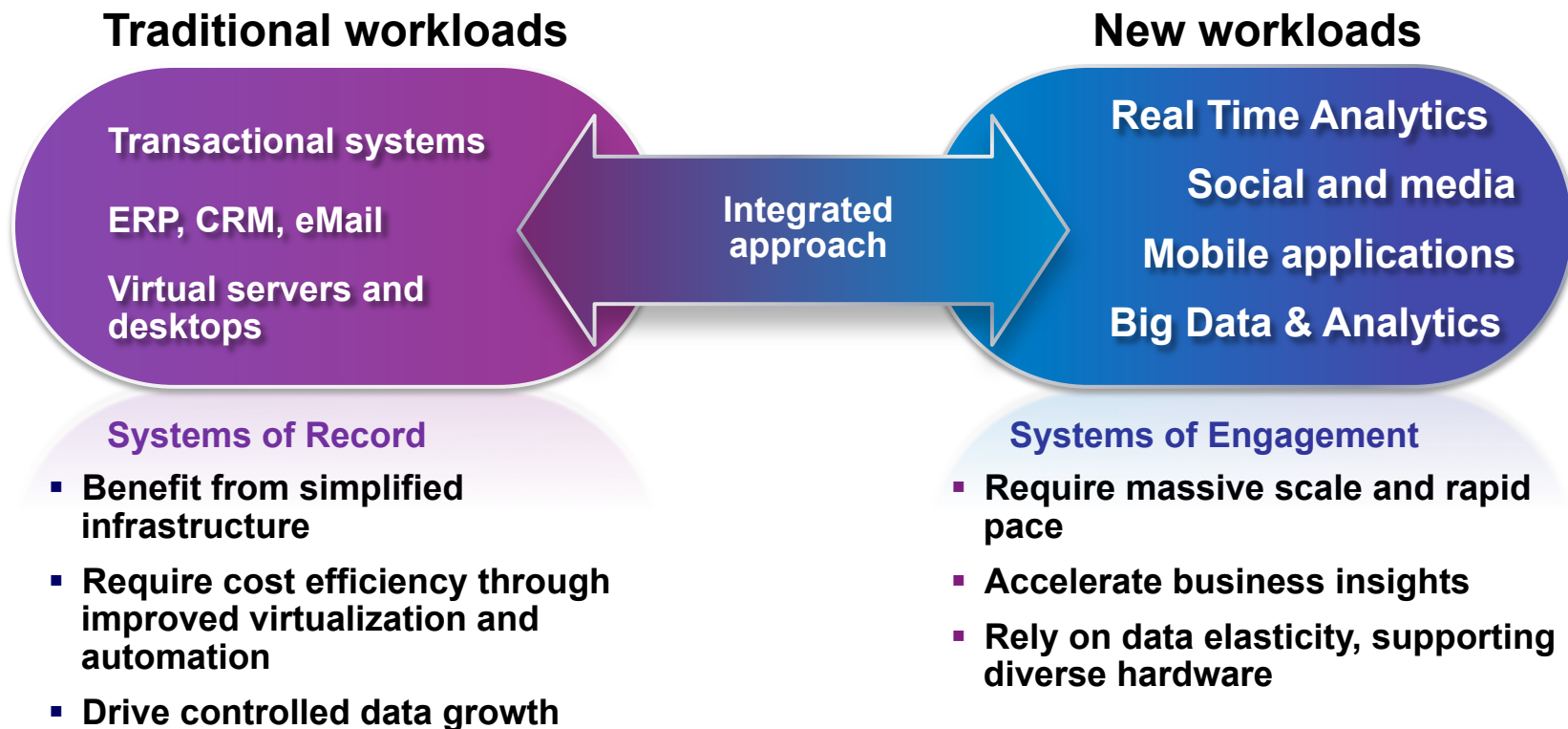


Characteristics of these Flash Systems

- Very low response time
- Must be consistent
- Even during rebuilds and code loads, etc.
- Accelerators and hardware data path
- Essentially expand In memory databases
- Flash becomes the colder tier!



Market shifts change data economics



A Vision of Where the Enterprise is Headed

