Solving the IO Bottleneck in NextGen DataCenters & Cloud Computing

Anil Vasudeva
President & Chief Analyst
imex@imexresearch.com
408-268-0800

IMEX
RESEARCH.COM

Abstract

IMEX RESEARCH.COM

Solving the I/O Bottleneck in NextGen Data Centers & Cloud Computing

Virtualization brings tremendous advantages to data centers of all sizes – large and small. But the rapid proliferation of virtual machines (VMs) per physical server, in its wake, creates a highly randomized I/O problem, raising performance bottlenecks. To address this I/O issue, the IT industry and storage administrators in particular, have begun the adoption of a slew of newer technology solutions - ranging from faster and larger memories in cache, SSDs for higher IOPS cost effectively, high bandwidth (10GbE) networks using NPIVs for virtualized networks between VMs and shared storage systems along with embedded intelligence software to optimize various VM workloads – all in order to meet various SLA metrics of performance, availability, cost etc..

Are you aware of the side-effects that get created from Server Virtualization and prepared to ask pertinent questions of your suppliers of IT Infrastructure Equipment, Storage Virtualization and Data Storage Management software as well as in their implementation to achieve targeted results in performance, availability, scalability, interoperability and data management in their virtualized data centers

This presentation provides an illustrative view of the impact of Server Virtualization on existing storage I/O solutions and best practices. It delineates the roles, capabilities and cost effectiveness of emerging technologies in mitigating the I/O bottlenecks so the IT infrastructure implementers can achieve their targeted performance under various workloads, from their storage systems in virtualized data centers.

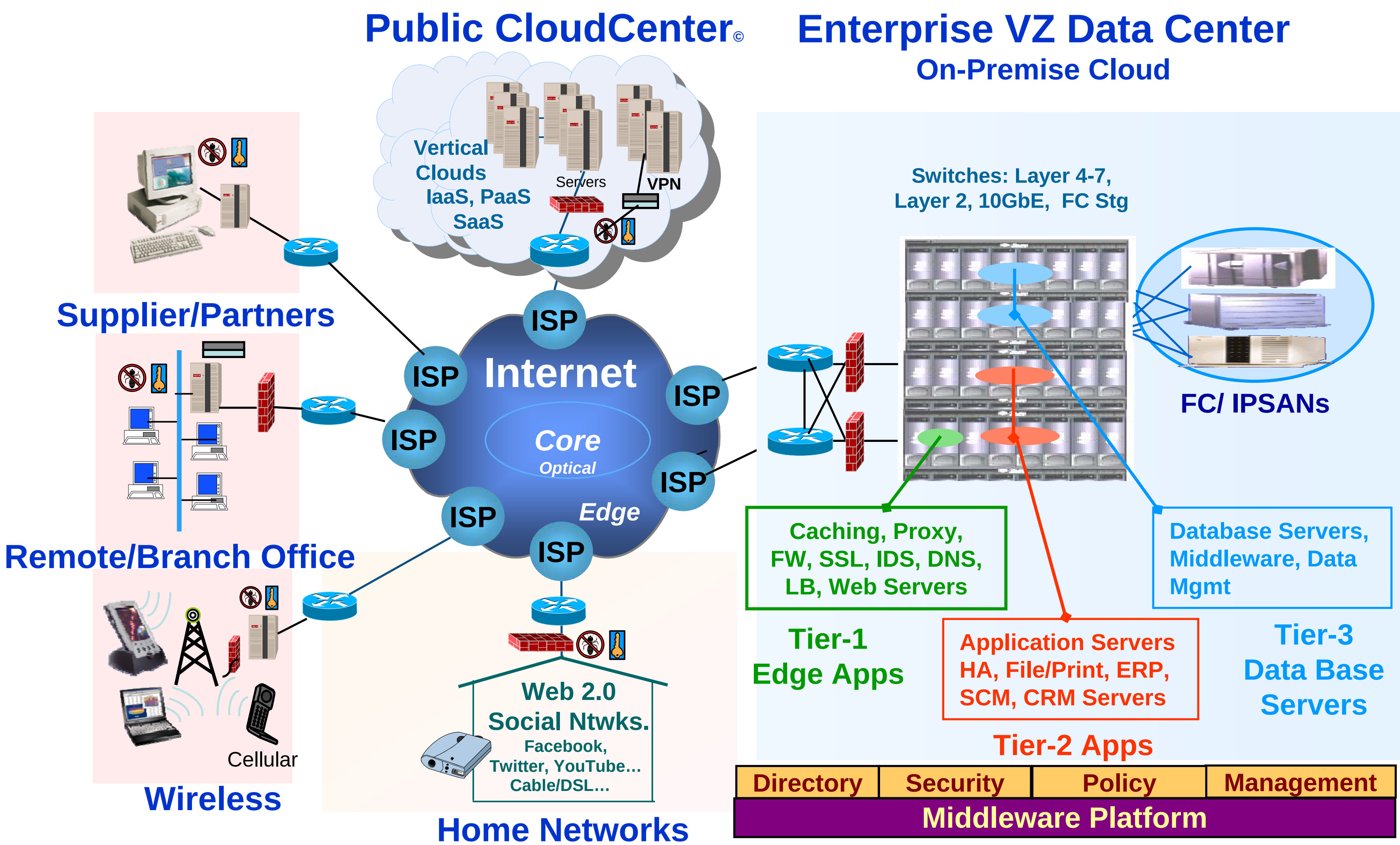
Agenda

IMEX RESEARCH.COM

- Data Centers & Cloud Infrastructure
- Cloud Computing Architecture
- Performance Metrics by Workload
- Anatomy of Data Access
- Data Center Performance Bottlenecks
- Improving Query Response Time in OLTP
- Role of SSD in Improving I/O Perf. Gap
- SCM: A New Storage Class Memory SSDs
- Price Erosion & IOPS/GB
- Choosing SSD vs. Memory to Improve TPS
- New Storage Usage Hierarchy in NGDC & Clouds
- IO Bottleneck Mitigation in Virtualized Servers
- I/O Forensics for Auto Storage-Tiering
- Apps Benefitting from Improved I/O
- Key Takeaways
- Acknowledgements

Data Centers & Cloud Infrastructure

IMEX RESEARCH.COM



Source:: IMEX Research - Cloud Infrastructure Report ©2009-11

IT Industry's Journey - Roadmap

IMEX
RESEARCH.COM

SIVAC
®IMEX

Cloudization

On-Premises > Private Clouds > Public Clouds
DC to Cloud-Aware Infrast. & Apps. Cascade migration to SPs/Public Clouds.

Automation

Automatically Maintains Application SLAs
(Self-Configuration, Self-Healing©IMEX, Self-Acctg. Charges etc)

Virtualization

Pools Resources. Provisions, Optimizes, Monitors
Shuffles Resources to optimize Delivery of various Business Services

Integration/Consolidation

Integrate Physical Infrast./Blades to meet CAPSIMS ®IMEX
Cost, Availability, Performance, Scalability, Inter-operability, Manageability & Security

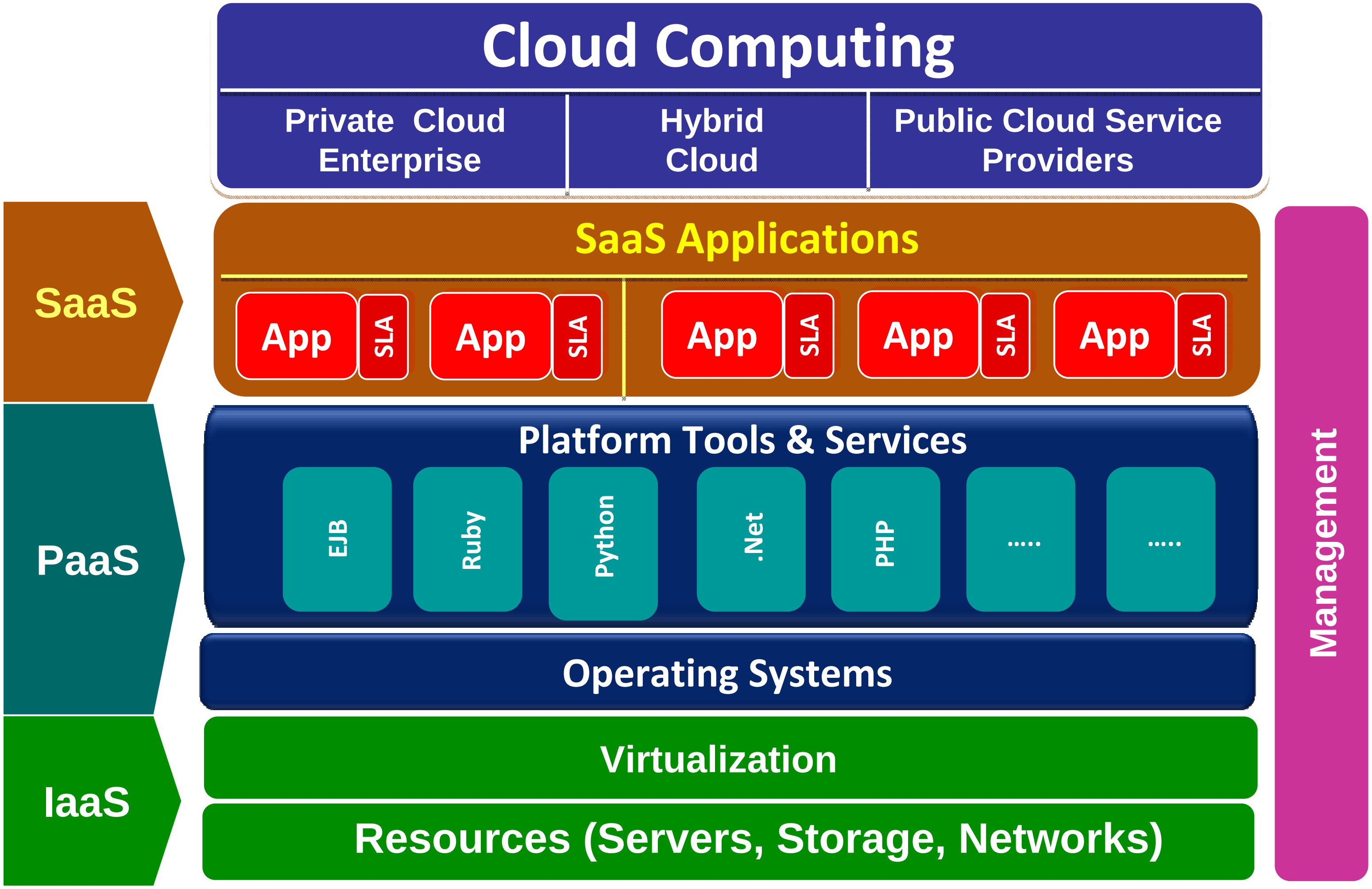
Standardization

Standard IT Infrastructure- Volume Economics HW/Syst SW
(Servers, Storage, Networking Devices, System Software (OS, MW & Data Mgmt SW)

Source:: IMEX Research - Cloud Infrastructure Report ©2009-11

Cloud Computing Architecture

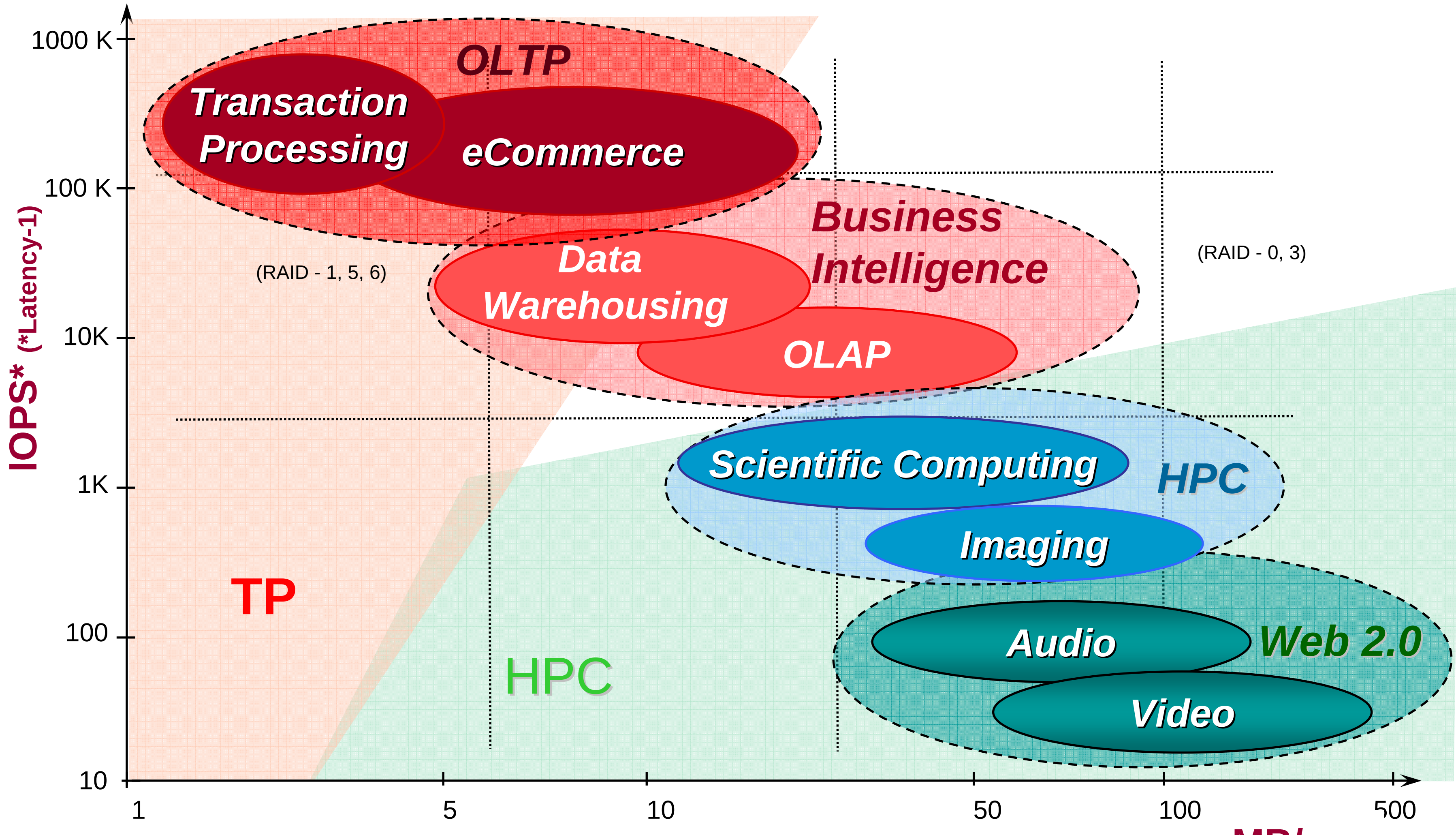
IMEX RESEARCH.COM



Application's SLA dictates the Resources Required to meet specific requirements of Availability, Performance, Cost, Security, Manageability etc.

Performance Metrics by Workload

IMEX RESEARCH.COM

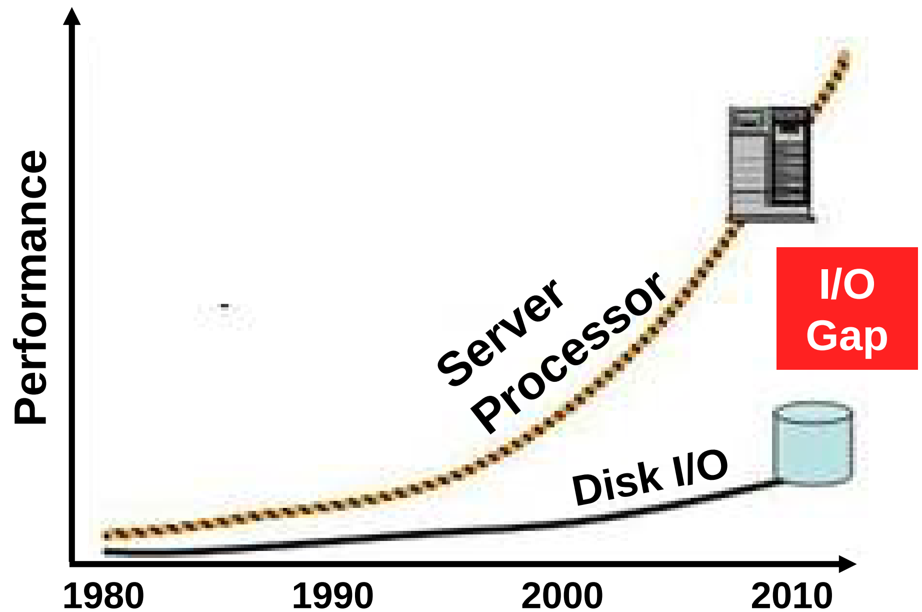


*IOPS for a required response time (ms)
*=(#Channels*Latency-1)

Source:: IMEX Research - Cloud Infrastructure Report ©2009-11

Anatomy of Data Access

IMEX RESEARCH.COM



For the time it takes to do each Disk Operation:

- Millions of CPU Operations can be done
- Hundreds of Thousands of Memory Operations can be accomplished

Anatomy of Data Access

Time taken by CPU, Memory, Network, Disk for a typical I/O Operation during a Data Access

L1 cache reference	0.5 ns
Branch mispredict	5 ns
L2 cache reference	7 ns
Mutex lock/unlock	25 ns
Main memory reference	100 ns
Compress 1K bytes with Zippy	3,000 ns
Send 2K bytes over 1 Gbps network	20,000 ns
Read 1 MB sequentially from memory	250,000 ns
Round trip within same datacenter	500,000 ns
Disk seek	10,000,000 ns
Read 1 MB sequentially from disk	20,000,000 ns
Send packet CA->Netherlands->CA	150,000,000 ns

A 7.2K/15k rpm HDD can do 100/140 IOPS

Data Center Performance Bottlenecks

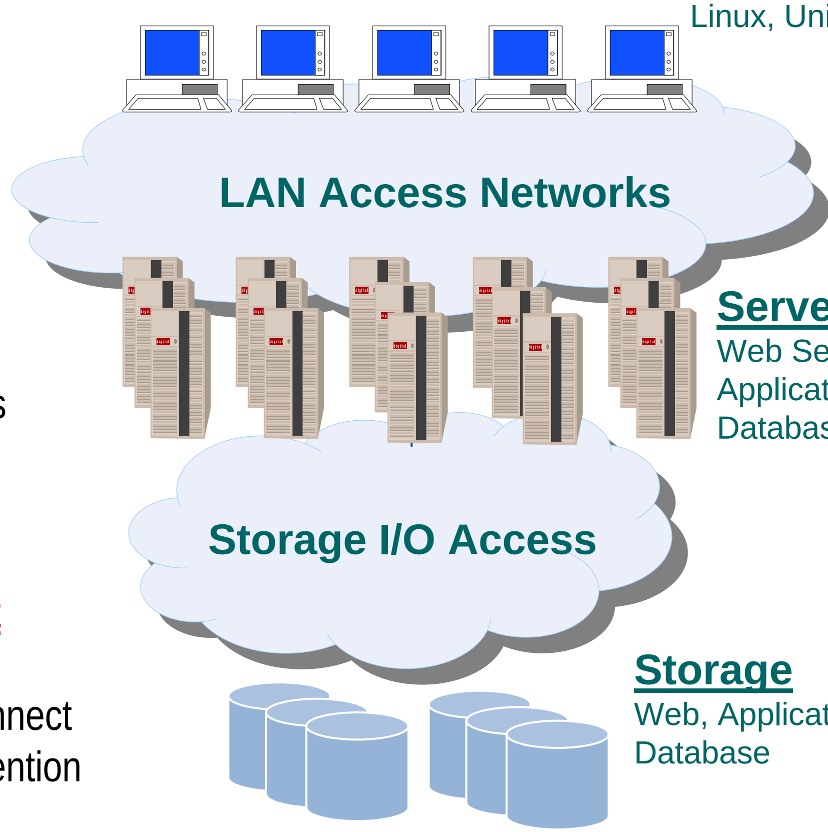
IMEX RESEARCH.COM

Applications
Excessive Locking
Data Contention
I/O Delays/Errors

Network I/O
Network Congestion
Dropped packets
Data Retransmissions
Timeouts
Component Failures

Storage I/O Connect
Lack of Bandwidth
Overloaded PCIe Connect
Storage Device Contention

Clients
Windows,
Linux, Unix



Servers
Web Servers
Application Servers
Database Servers

Storage
Web, Application,
Database

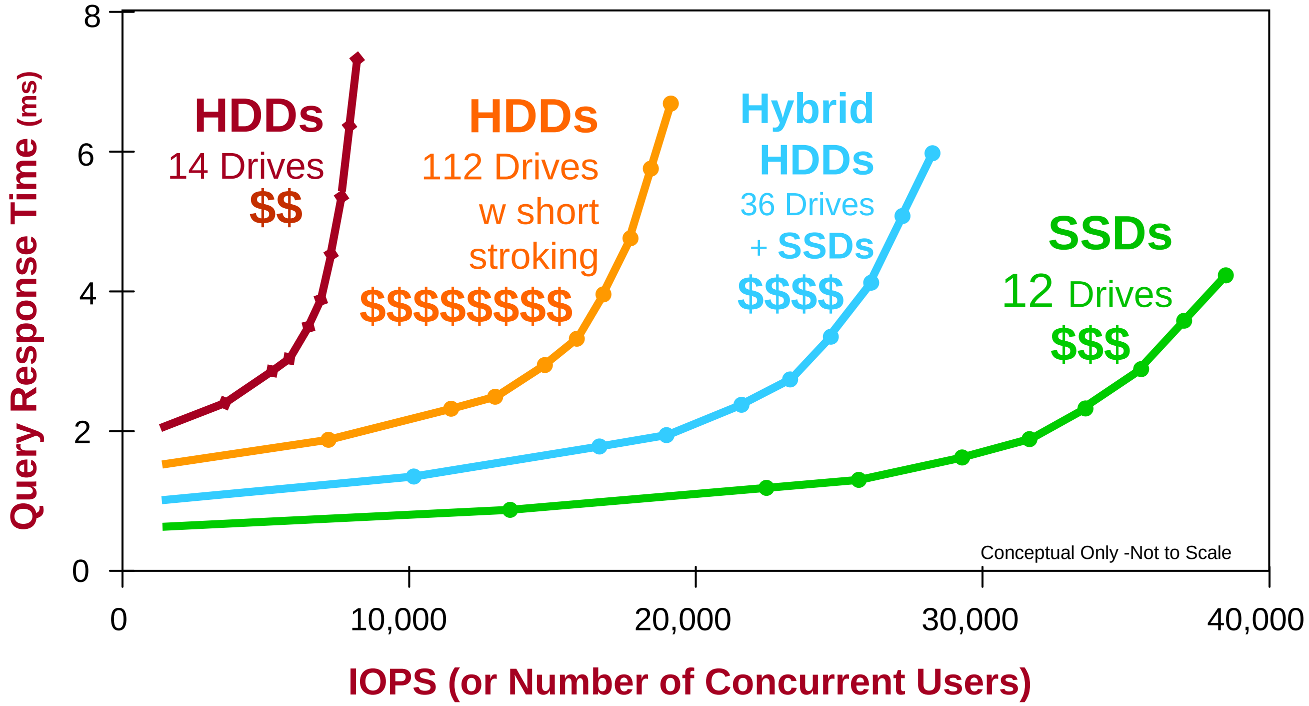
User Bottlenecks
Connectivity Timeouts,
Workload Surges

Server Bottlenecks
Lack of Srvr Power
IO Wait & Queuing CPU
Overhead I/O Timeouts

Device Bottlenecks
Device I/O Hotspots
Cache Flush
Lack of Storage Capacity

Improving Query Response Time in OLTP

IMEX RESEARCH.COM

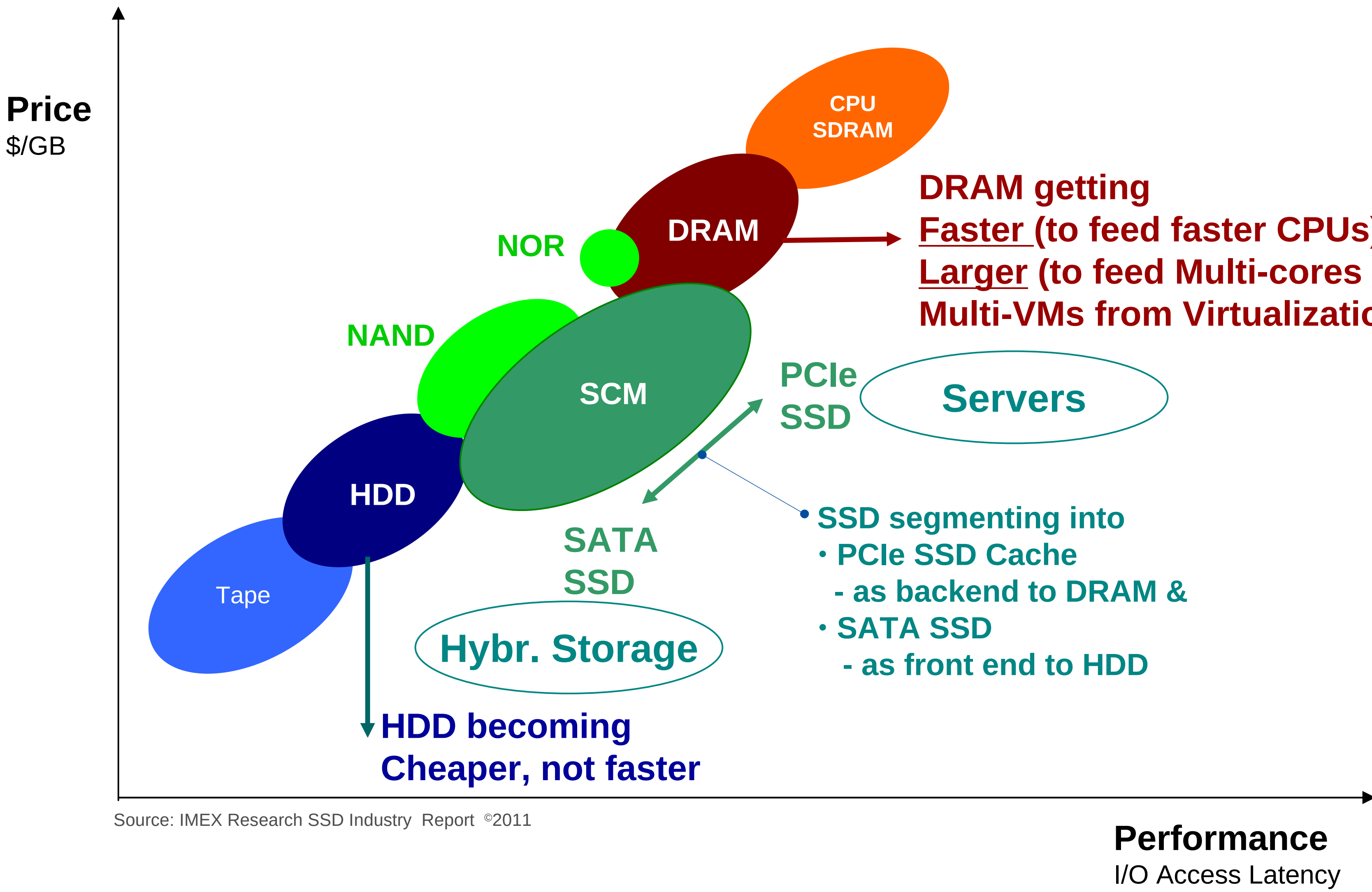


- **Improving Query Response Time**
  - Cost effective way to improve Query response time for a given number of users or servicing an increased number of users at a given response time is best served with use of SSDs or Hybrid (SSD + HDDs) approach, particularly for Database and Online Transaction Applications

Source: IMEX Research SSD Industry Report ©2011

Role of SSD in Improving I/O Perf. Gap

IMEX RESEARCH.COM



Source: IMEX Research SSD Industry Report ©2011

SCM: A New Storage Class Memory

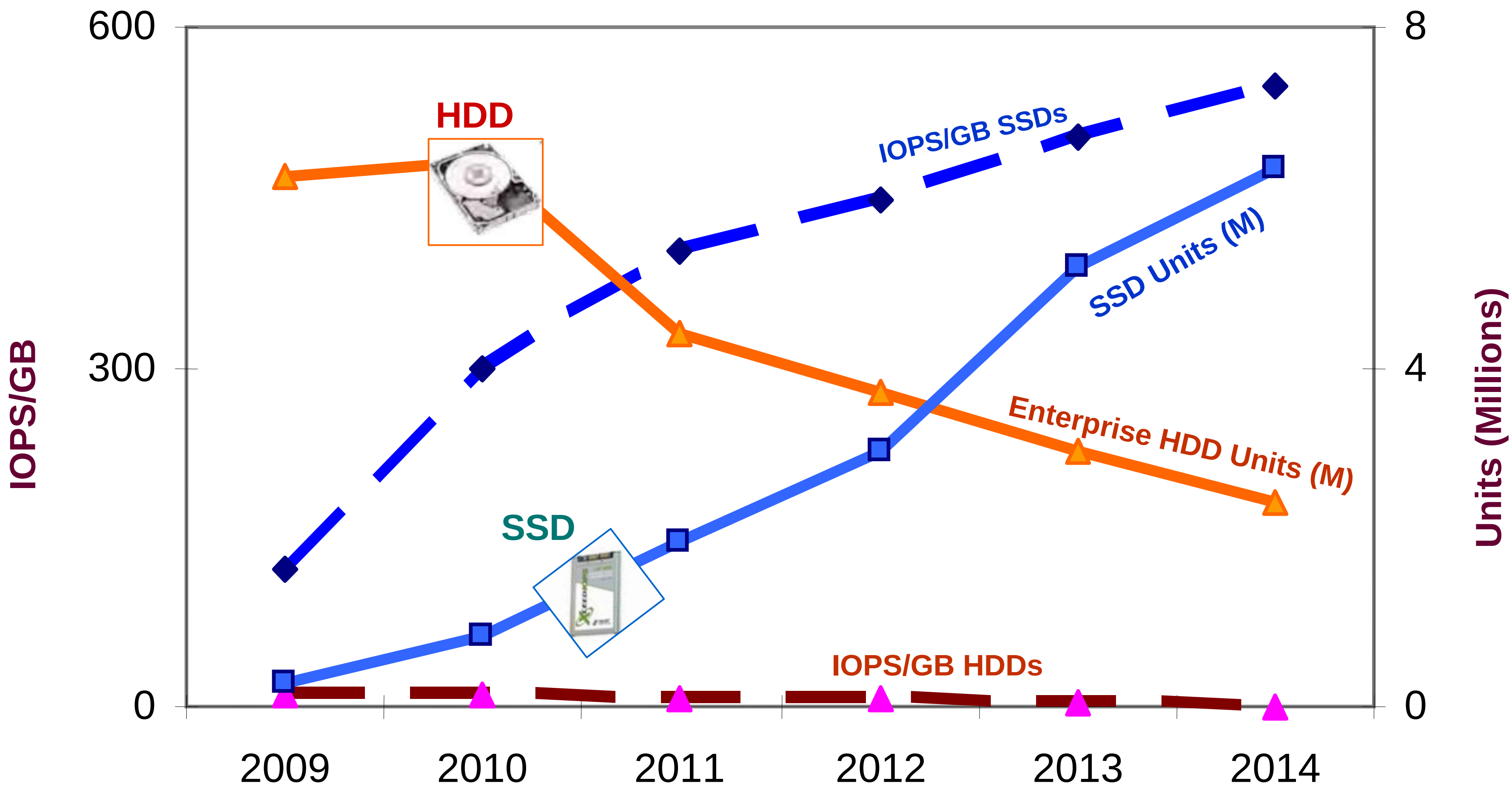
IMEX RESEARCH.COM

- **SCM (Storage Class Memory)**
  Solid State Memory filling the gap between DRAMs & HDDs
  Marketplace segmenting SCMs into SATA and PCIe based SSDs

- **Key Metrics Required of Storage Class Memories**
  Device - Capacity (GB), Cost ($/GB),
  Performance - Latency (Random/Block RW Access-ms); Bandwidth BW(R/W- GB/sec)
  Data Integrity - BER (Better than 1 in 10^17)
  Reliability - Write Endurance (30K PE Cycles No. of writes before death);
  - Data Retention (5 Years); MTBF (2 millions of Hrs),
  Environment - Power Consumption (Watts);
  Volumetric Density (TB/cu.in.); Power On/Off Time (sec),
  Resistance - Shock/Vibration (g-force); Temp./Voltage Extremes 4-Corner (oC,V); Radiation (Rad)

SSDs - Price Erosion & IOPS/GB

IMEX RESEARCH.COM

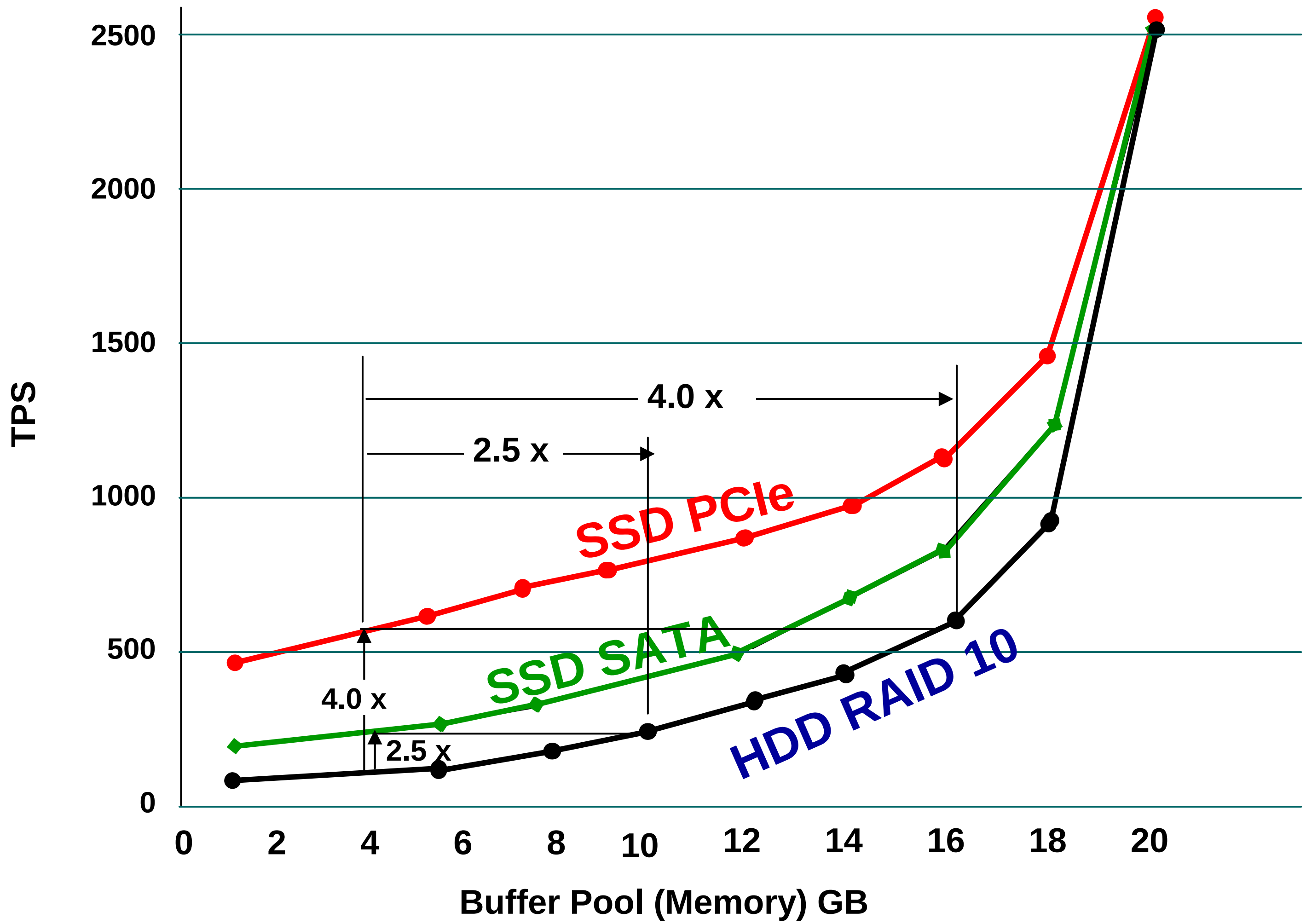


Note: 2U storage rack, • 2.5" HDD max cap = 400GB / 24 HDDs, de-stroked to 20%, • 2.5" SSD max cap = 800GB / 36 SSDs

Source: IMEX Research SSD Industry Report ©2011

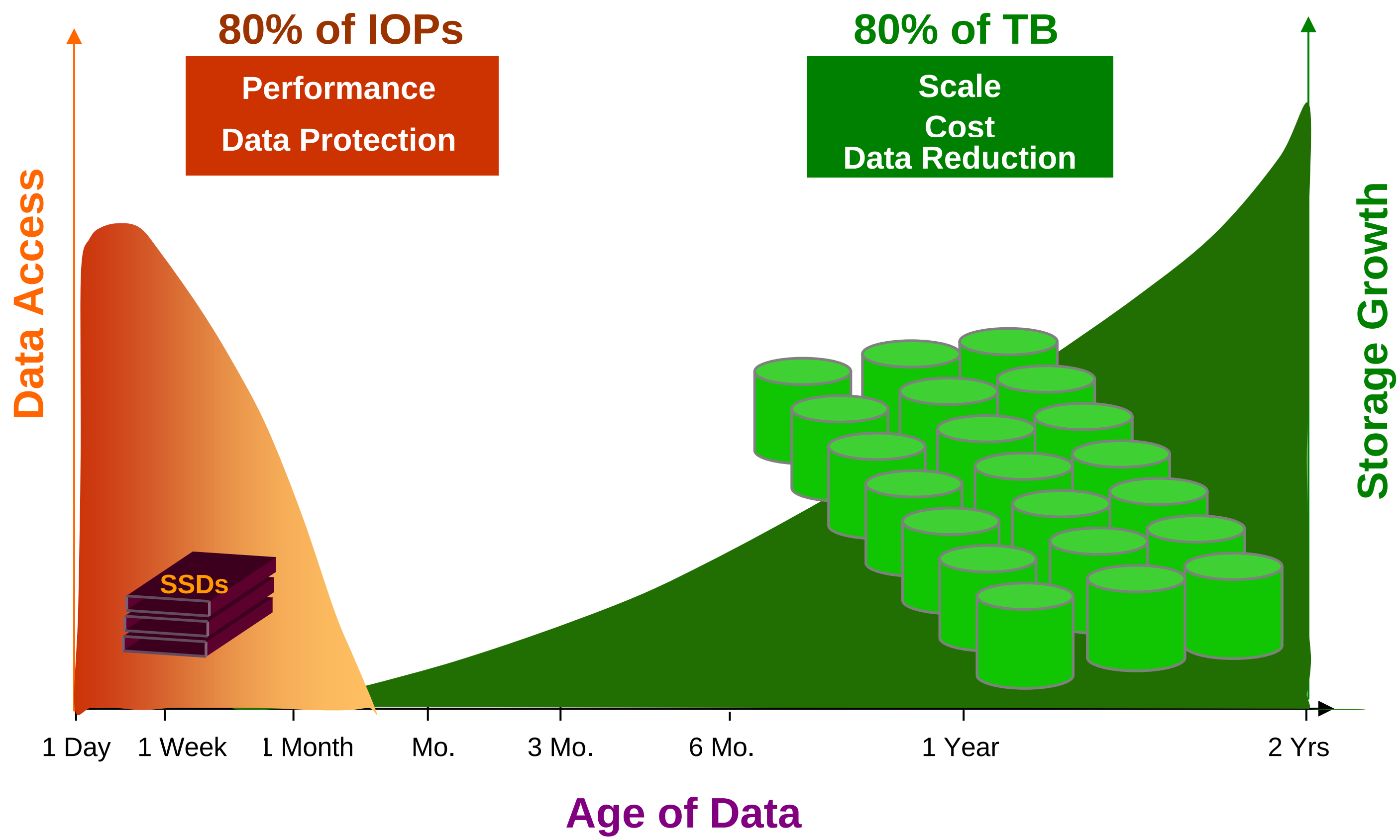
Choosing SSD vs. Memory to Improve TPS

IMEX RESEARCH.COM



Data Storage Usage Patterns – Data Access vs. Age of Data

IMEX RESEARCH.COM

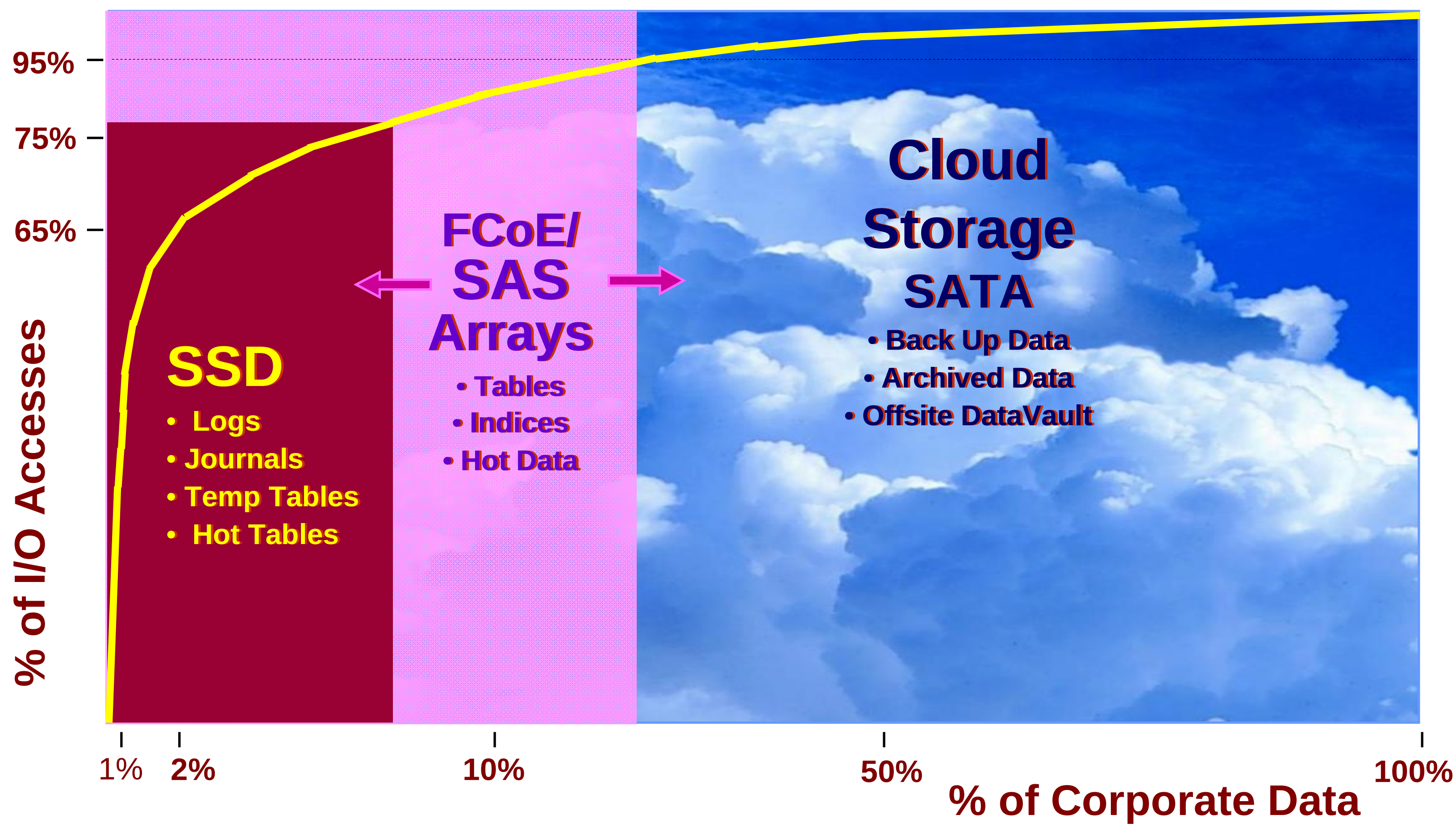


Source:: IMEX Research - Cloud Infrastructure Report ©2009-11

New Storage Hierarchy in NGDC & Clouds

IMEX RESEARCH.COM

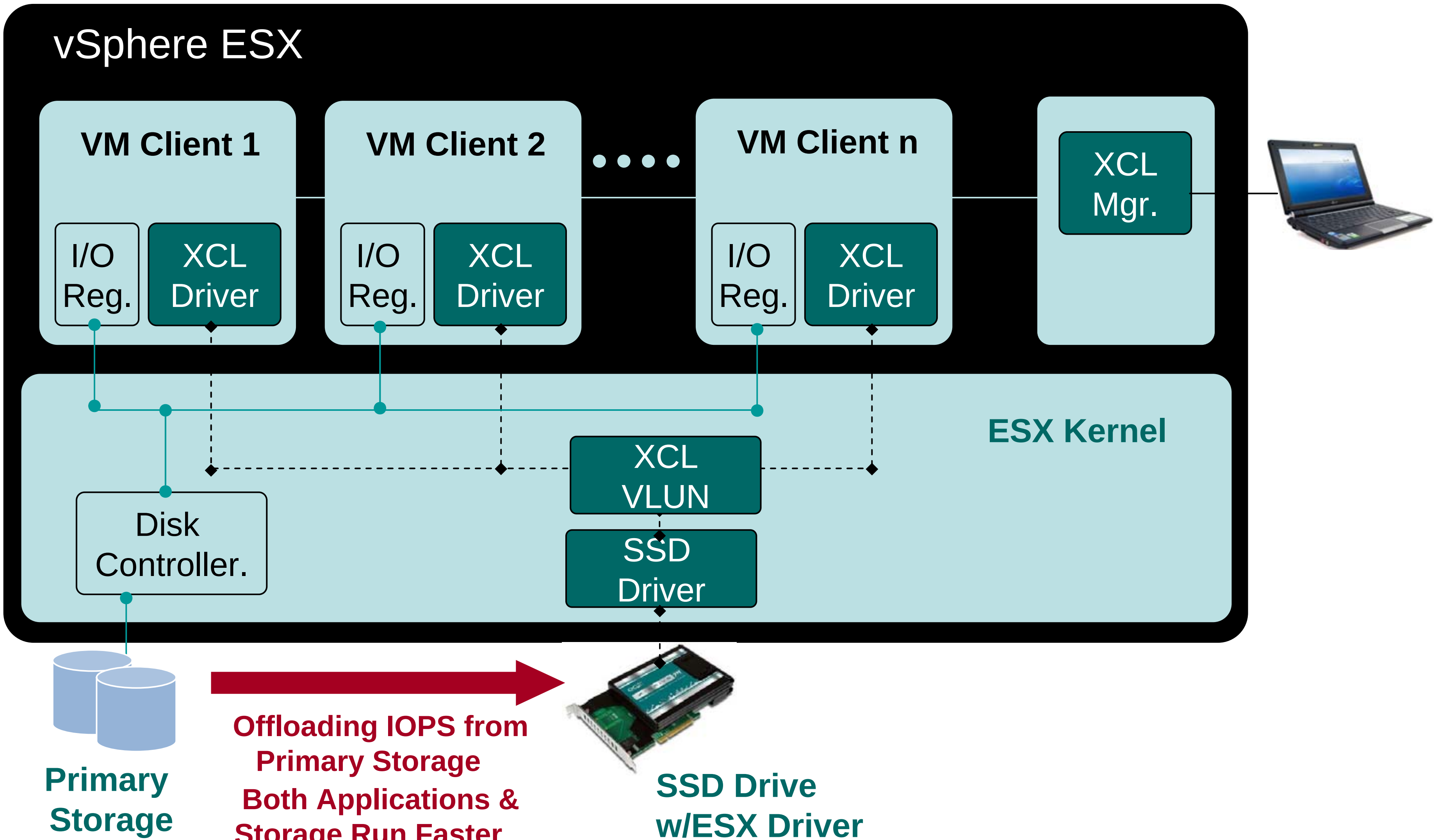
I/O Access Frequency vs. Percent of Corporate Data



Source:: IMEX Research - Cloud Infrastructure Report ©2009-11

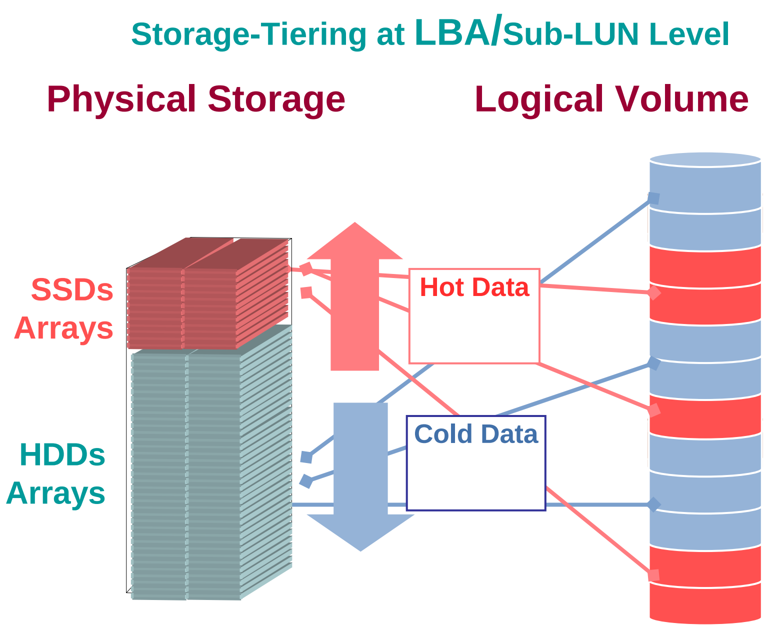
IO Bottleneck Mitigation in Virtualized Servers

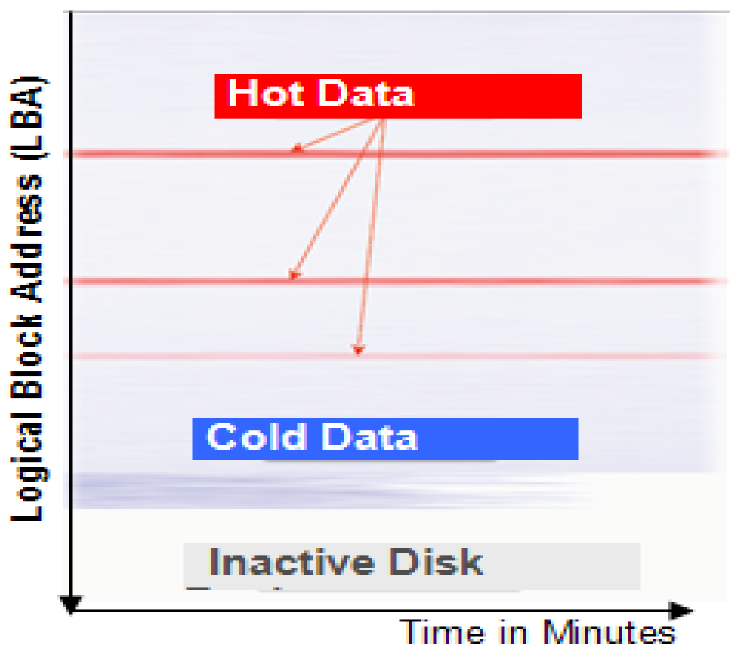
IMEX RESEARCH.COM



IMEX RESEARCH.COM

Storage-Tiered Virtualization





LBA Monitoring and Tiered Placement

- Every workload has unique I/O access signature
- Historical performance data for a LUN can identify performance skews & hot data regions by LBAs

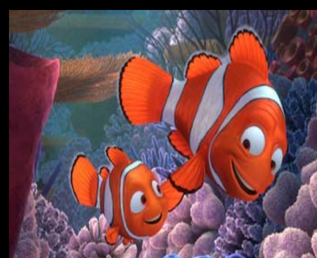
Source: IBM & IMEX Research SSD Industry Report 2011 ©IMEX 2010-11

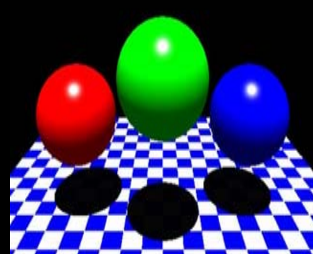
Apps Benefitting from Improved I/O

IMEX RESEARCH.COM

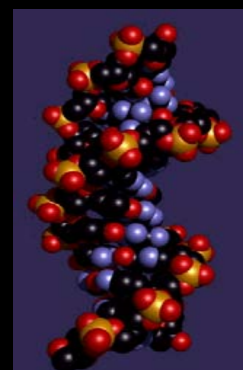
Smart Mobile Devices

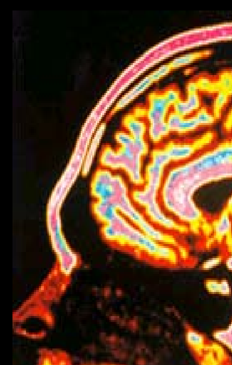
Commercial Visualization





Bioinformatics & Diagnostics





Decision Support Bus. Intelligence

Entertainment-VoD / U-Tube

Data: IMEX Research & Panasas

Instant On Boot Ups
Rugged, Low Power
1GB/s, __ms

Rendering (Texture & Polygons)
Very Read Intensive, Small Block I/O
10 GB/s, __ms

Data Warehousing
Random IO, High OLTPM
1GB/s, __ms

Most Accessed Videos
Very Read Intensive
4 GB/s, __ms

Key Takeaways

IMEX RESEARCH.COM

- **Solving I/O Problems**
  - I/O Bottlenecks occur at multiple places in the Compute Stack, the largest being at Storage I/O
  - SSD comes out cheaper/IOP for IO Intensive Apps
  - To get of Reads – Improve Indexing, archive out old data
  - Minimize the impact of writes – Get rid of temp tables/filesorts on slow disks.
  - Compress big varchar/text/blobs
- **Data Forensics and Tiered Placement**
  - Every workload has unique I/O access signature
  - Historical performance data for a LUN can identify performance skews & hot data regions by LBAs
  - Use Smart Tiering to identify hot LBA regions and non-disruptively migrate hot data from HDD to SSDs.
  - Typically 4-8% of data becomes a candidate and when migrated to SSDs can provide response time reduction of ~65% at peak loads