



Bringing HyperScale Computing to the Enterprise

The need for Enterprises to overhaul their IT systems

About MSys



Established In:
2007



Self-funded,
profitable



Over 350
employees



Global
Presence

Product Engineering Services

- Product Engineering / Development
- Test Engineering
- Rapid Prototyping
- Maintenance & Support
- UI Engineering

Technology

- Cloud Computing
- Storage
- Big Data
- Analytics
- DevOps
- Infrastructure Tech

Storage Services

- Filesystem Development
- Kernel Development
- Backup & DR
- Cloud Storage
- Storage QA
- Storage Certifications



Global
Presence

MSys
Advantages

- End-to-end Storage Expertise
- 75% revenue from storage companies
- Dedicated team with focused skills

About the Speaker



Chirag Jog

VP, Engineering at Msys(Clogeny)



Pune, India



Architect with 10+ years of experience.



chirag@clogeny.com



Built cutting edge products in Storage, Cloud and Big data space



@chirag_jog



Products: Cloud Mobility, Cloud Backup Products



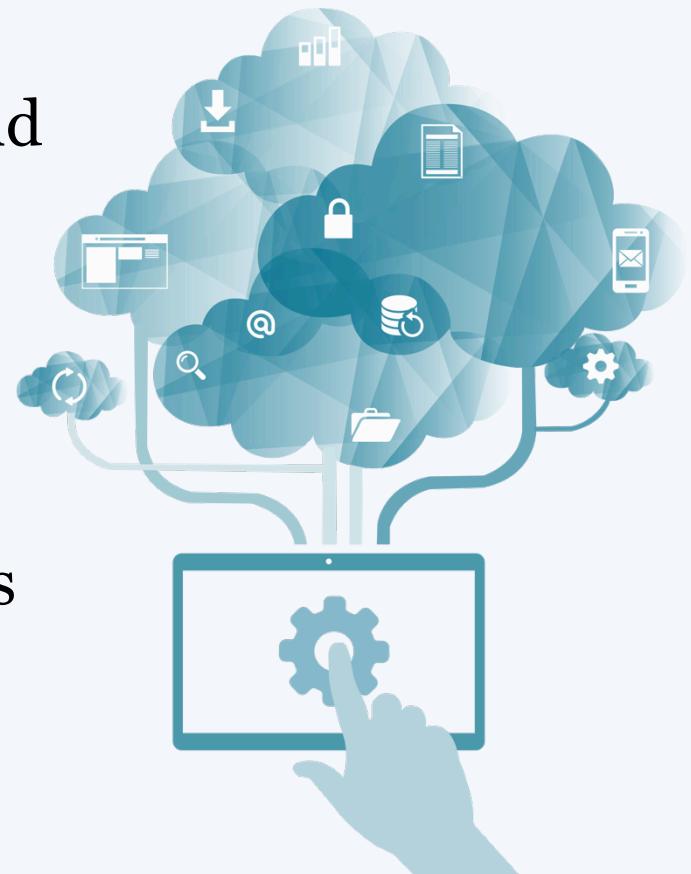
<http://in.linkedin.com/in/chiragjog>



Open source contributor

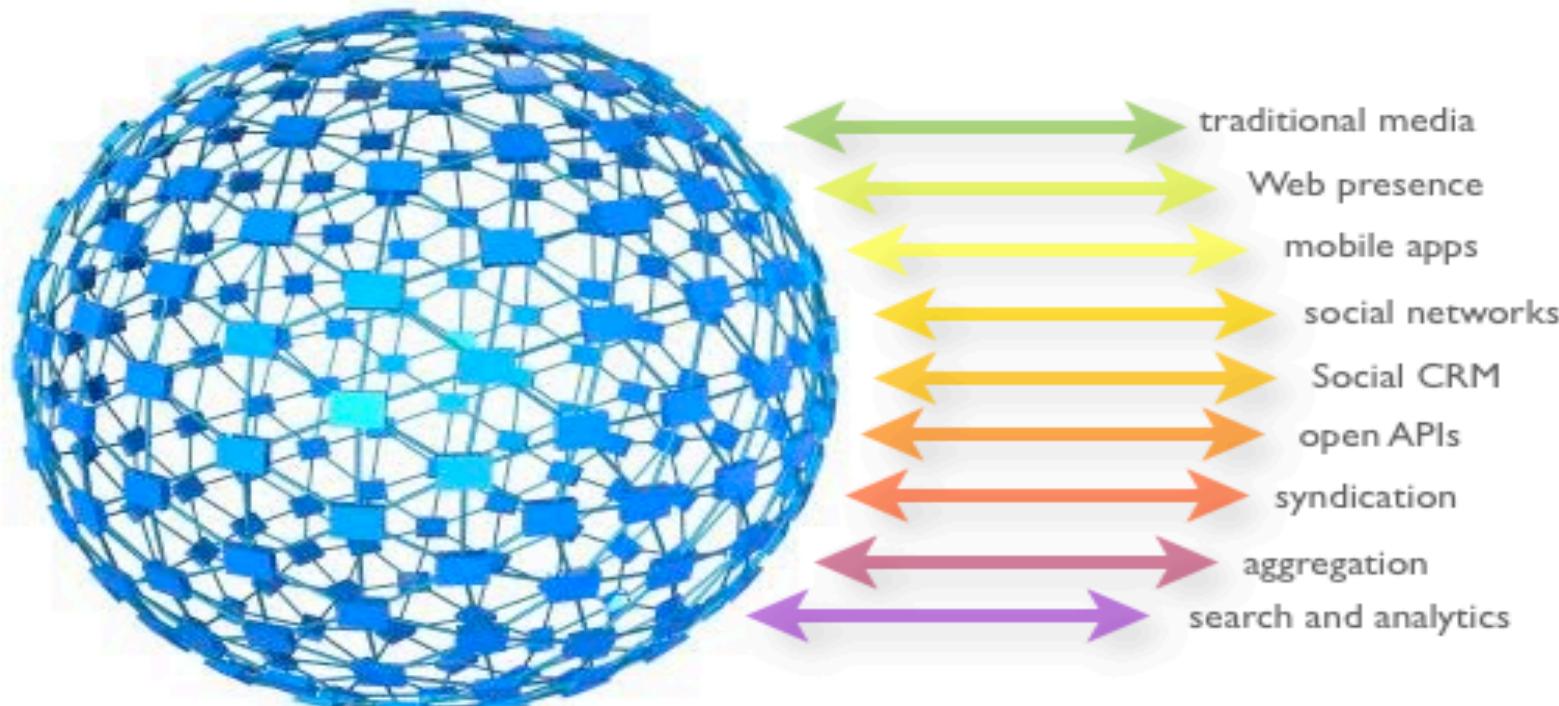
Agenda

- Changing nature of Applications and Data in Enterprise
- Time to re-think the Approach to Infrastructure
- Current Challenges and Limitations
- Trends in Software and Hardware



Next-Gen Enterprise Ecosystem

points of engagement + data + people



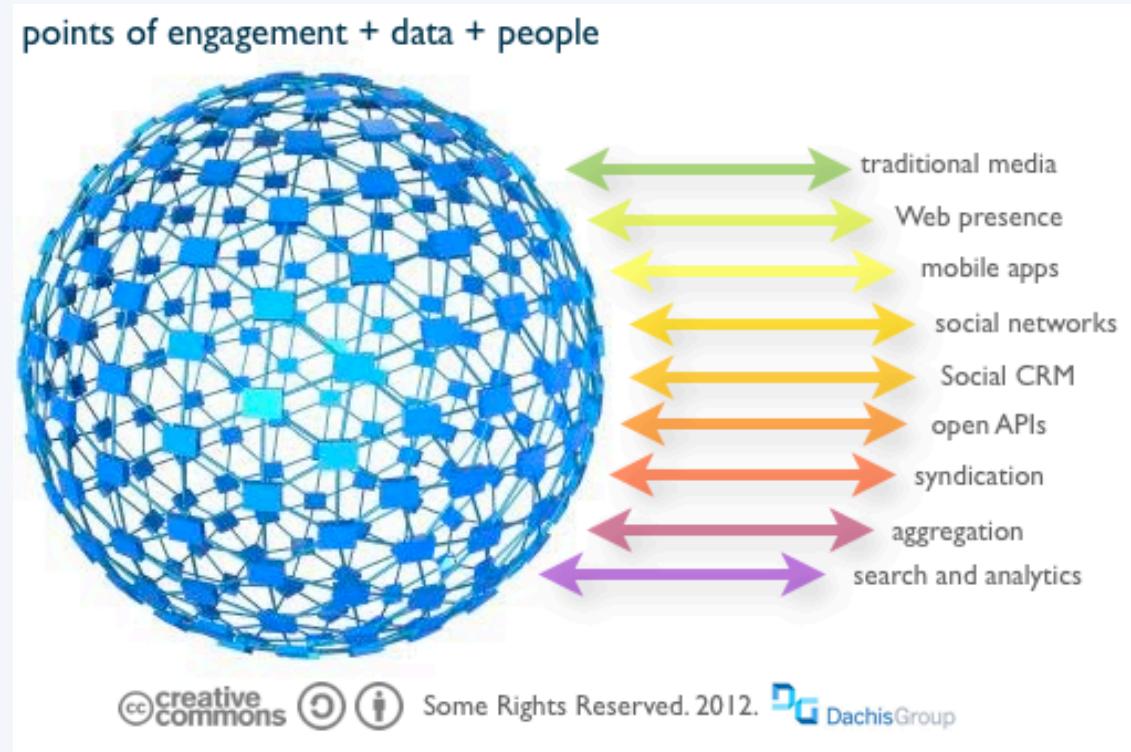
Some Rights Reserved. 2012.



DachisGroup

Next-gen Application and Data

- Web-Scale IT



Next-gen Application and Data

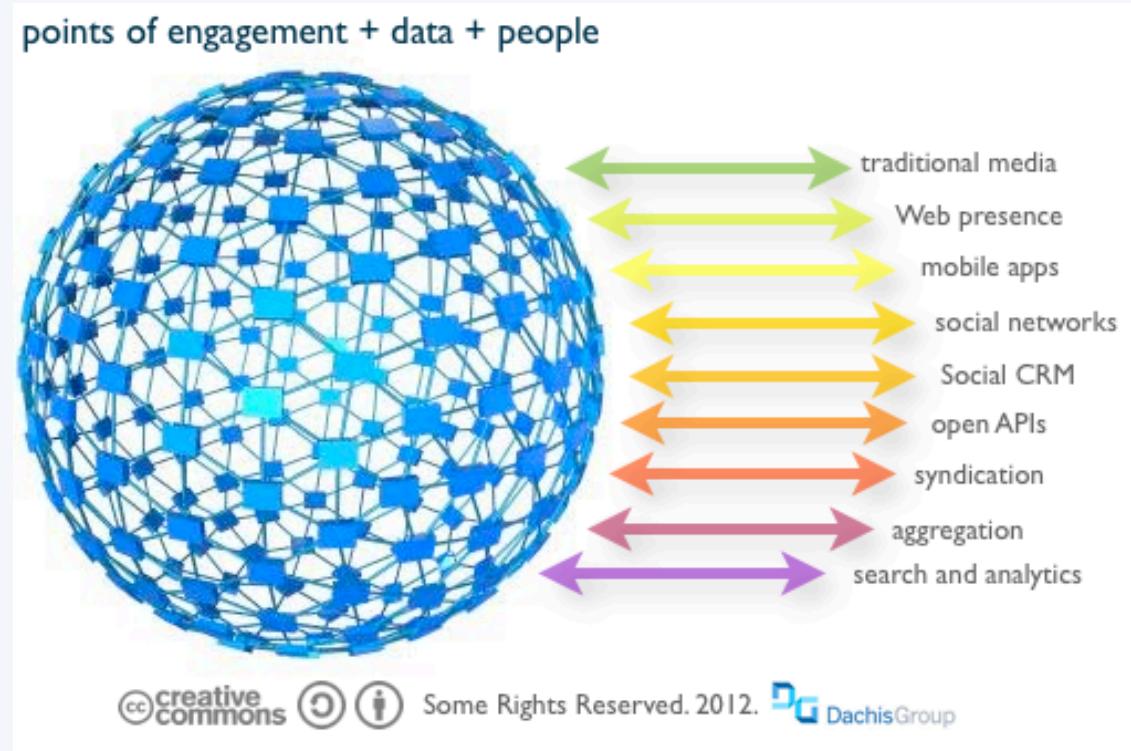
- **Web-Scale IT**
 - Available as-a-Service
 - Scale on demand
 - Always on operation
 - Automated Provisioning
 - Agility, Flexibility



Next-gen Application and Data



- Web-Scale IT
- Real-time Data streams



Next-gen Application and Data

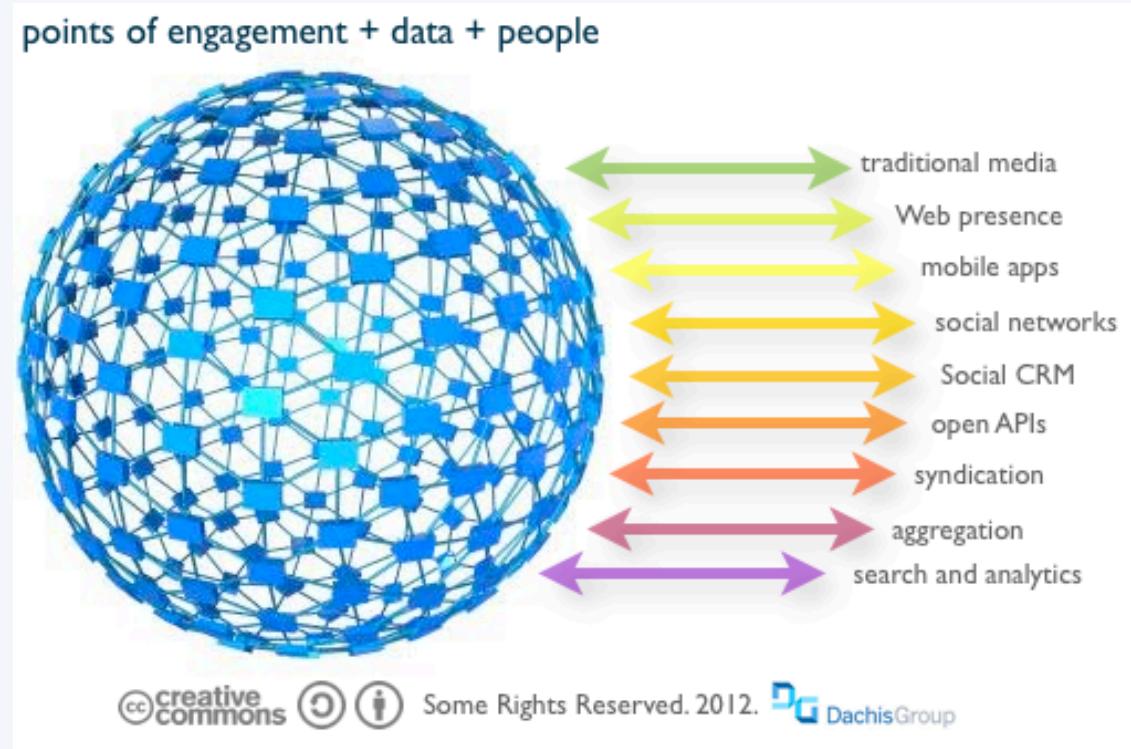


- **Real-time Data streams**
 - customers
 - partners
 - supply chains
 - applications
 - internet of things



Next-gen Application and Data

- Web-Scale IT
- Real-time Data streams
- Near real-time recommendations



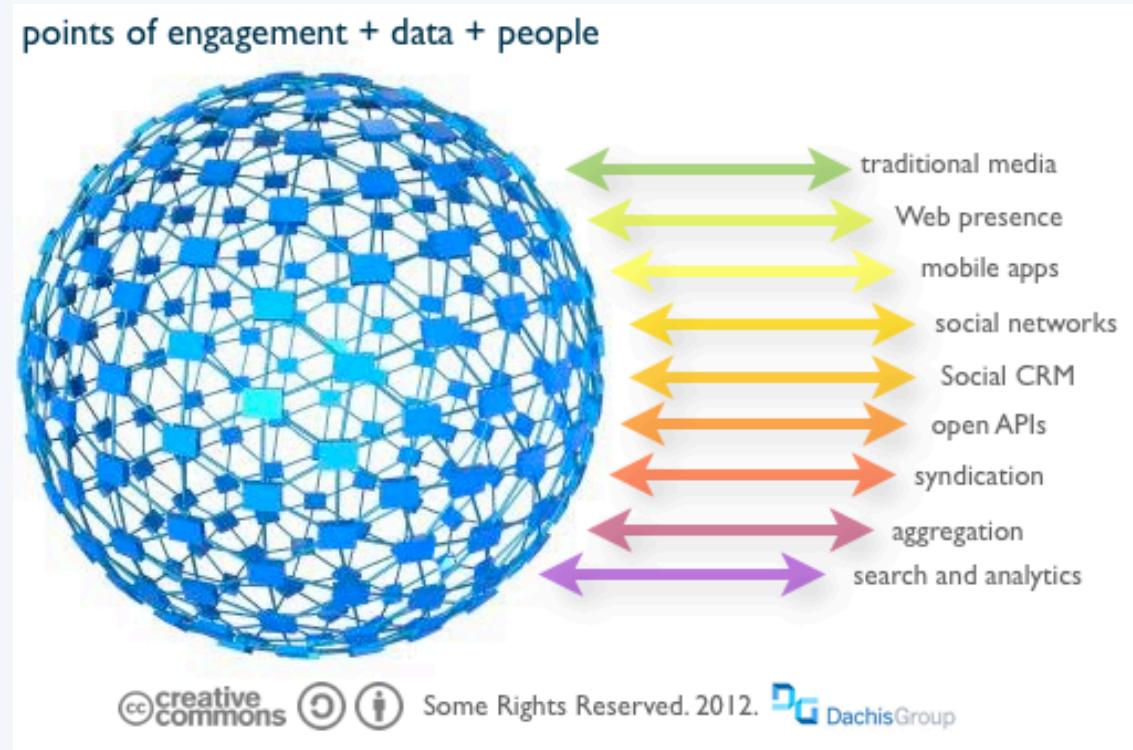
Next-gen Application and Data

- Near real-time recommendations
- real-time transactional and analytical systems

WHEN
GETTING IT
FAST IS TOO
SLOW,
**GET IT IN
REALTIME**

Next-gen Application and Data

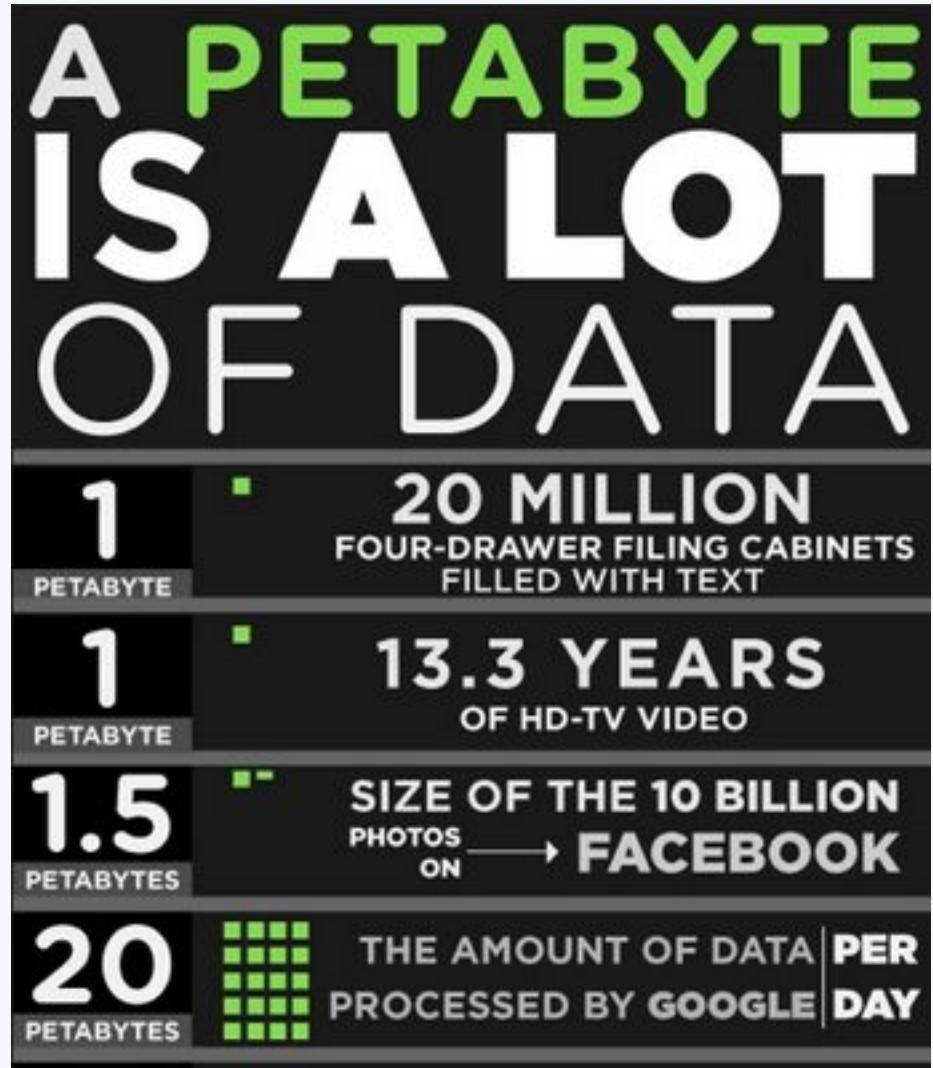
- Web-Scale IT
- Real-time Data streams
- Near real-time recommendations
- **Massive amounts of data**



Next-gen Application and Data



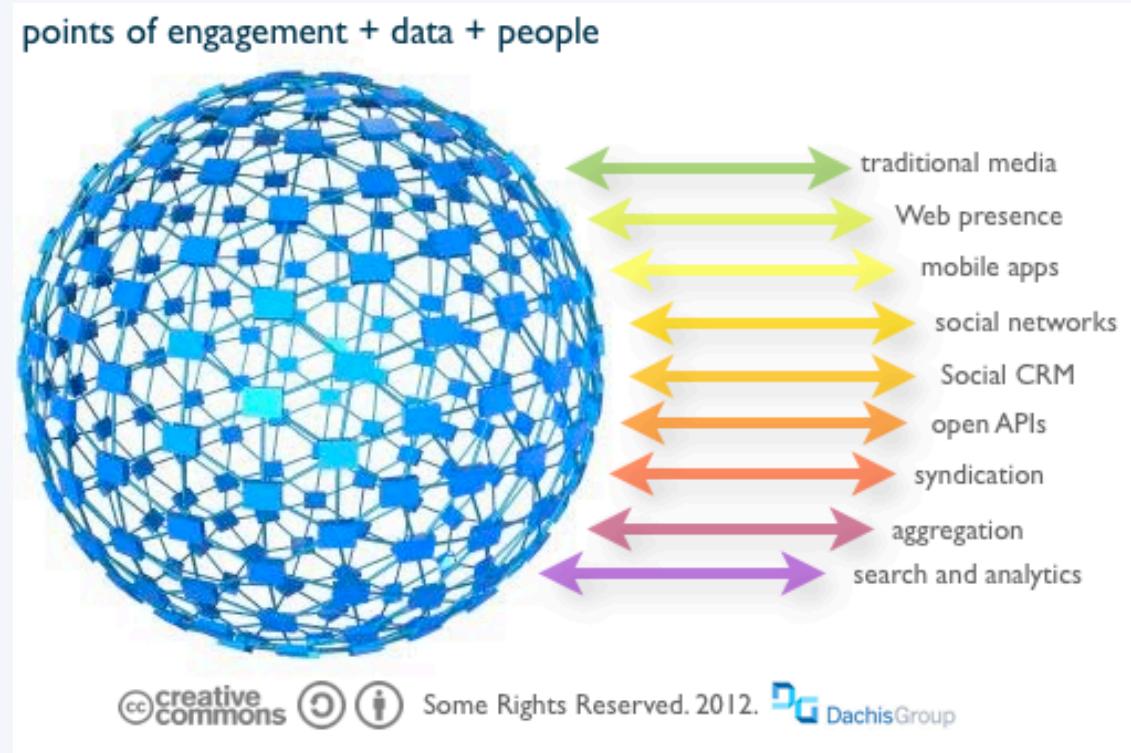
- Massive amounts of data
 - structured, unstructured
 - archive
 - readily available for analytics, compliance, outages



Next-gen Application and Data



- Web-Scale IT
- Real-time Data streams
- Near real-time recommendations
- Massive amounts of data
- **Cost effective solutions**



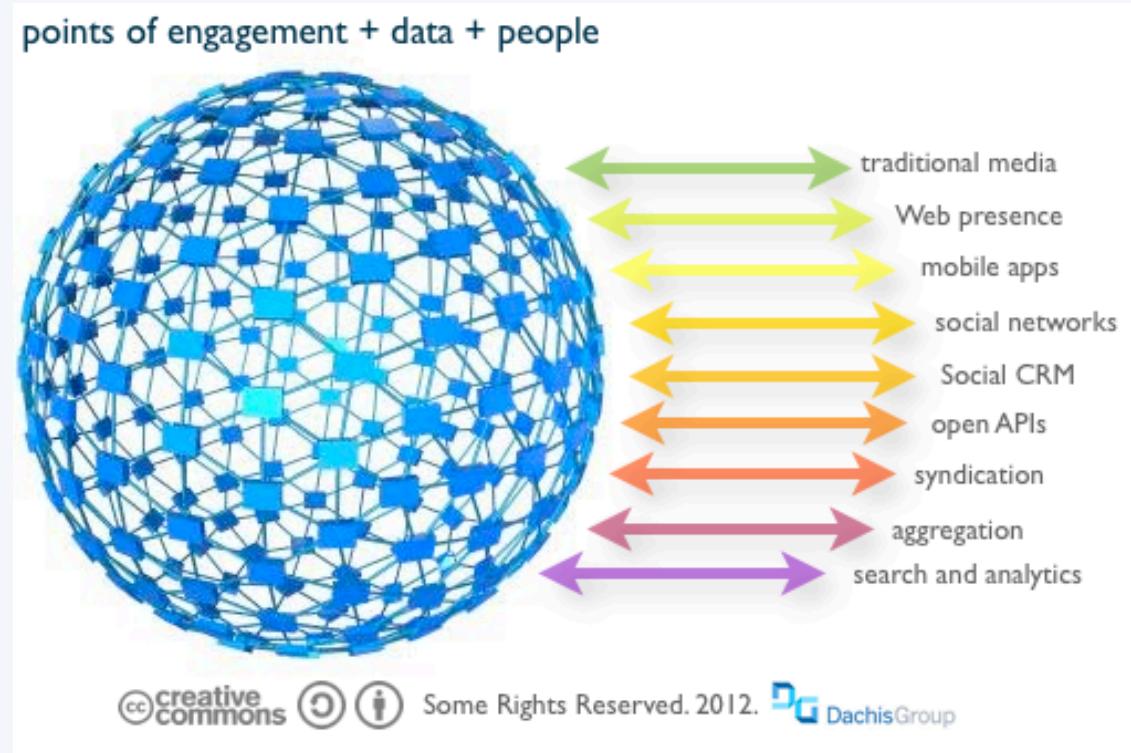
- **Cost effective solutions**
 - cost-effect performance
 - Minimize cost of Human resources
 - Minimize costs of infrastructure



Next-gen Application and Data



- Web-Scale IT
- Real-time Data streams
- Near real-time recommendations
- Massive amounts of data
- Cost effective solutions



Current Constraints- Storage

- Multiple Storage Array Platforms and Arrays

Application

← Conventional SCSI IO →

Simple
Block

Network
File

Proprietary Storage OS

Storage Media



Storage Services

- Replication
- Snapshots & Clones
- Hypervisor integration
- Compression/De-duplication
- Encryption
- Vendor Maintenance

Current Constraints- Storage

- Multiple Storage Array Platforms and Arrays
- Slow Access Speed and throughput of Disk Drives

Application

← Conventional SCSI IO →

Simple Block

Network File

Proprietary Storage OS

Storage Media



Storage Services

- Replication
- Snapshots & Clones
- Hypervisor integration
- Compression/De-duplication
- Encryption
- Vendor Maintenance

Current Constraints- Storage

- Multiple Storage Array Platforms and Arrays
- Slow Access Speed and throughput of Disk Drives
- Difficulty of migrating current management of data on DAS and SAN to more flexible topologies

Application

← Conventional SCSI IO →

Simple Block

Network File

Proprietary Storage OS

Storage Media



Storage Services

- Replication
- Snapshots & Clones
- Hypervisor integration
- Compression/De-duplication
- Encryption
- Vendor Maintenance

Current Constraints- Storage

- Multiple Storage Array Platforms and Arrays
- Slow Access Speed and throughput of Disk Drives
- Difficulty of migrating current management of data on DAS and SAN to more flexible topologies
- Difficulty in managing the volume, growth and complexity of unstructured data

Application

← Conventional SCSI IO →

Simple Block

Network File

Proprietary Storage OS

Storage Media



Storage Services

- Replication
- Snapshots & Clones
- Hypervisor integration
- Compression/De-duplication
- Encryption
- Vendor Maintenance

Current Constraints- Storage

- Multiple Storage Array Platforms and Arrays
- Slow Access Speed and throughput of Disk Drives
- Difficulty of migrating current management of data on DAS and SAN to more flexible topologies
- Difficulty in managing the volume, growth and complexity of unstructured data
- A fragmented and often fragile Data backup process

Application

← Conventional SCSI IO →

Simple Block

Network File

Proprietary Storage OS

Storage Media



Storage Services

- Replication
- Snapshots & Clones
- Hypervisor integration
- Compression/De-duplication
- Encryption
- Vendor Maintenance

Current Limitations - Applications

- Interconnected and Inter-dependent - Apps, Infra, Physical DC
- Apps in a silo - Optimized Infrastructure
- Rigid and easy to knock down
- 1 admin per 300-700 servers.
- Managed infrastructure by hand - updating software, rearranging configuration, SLAs
- Inflexible to innovate - decisions are made for 15-20 years

Million Dollar Question or Exabyte Question



How are exa-bytes of data going to be:

- backed up?
- restored?
- accessed?



How did “they” do it?



Google

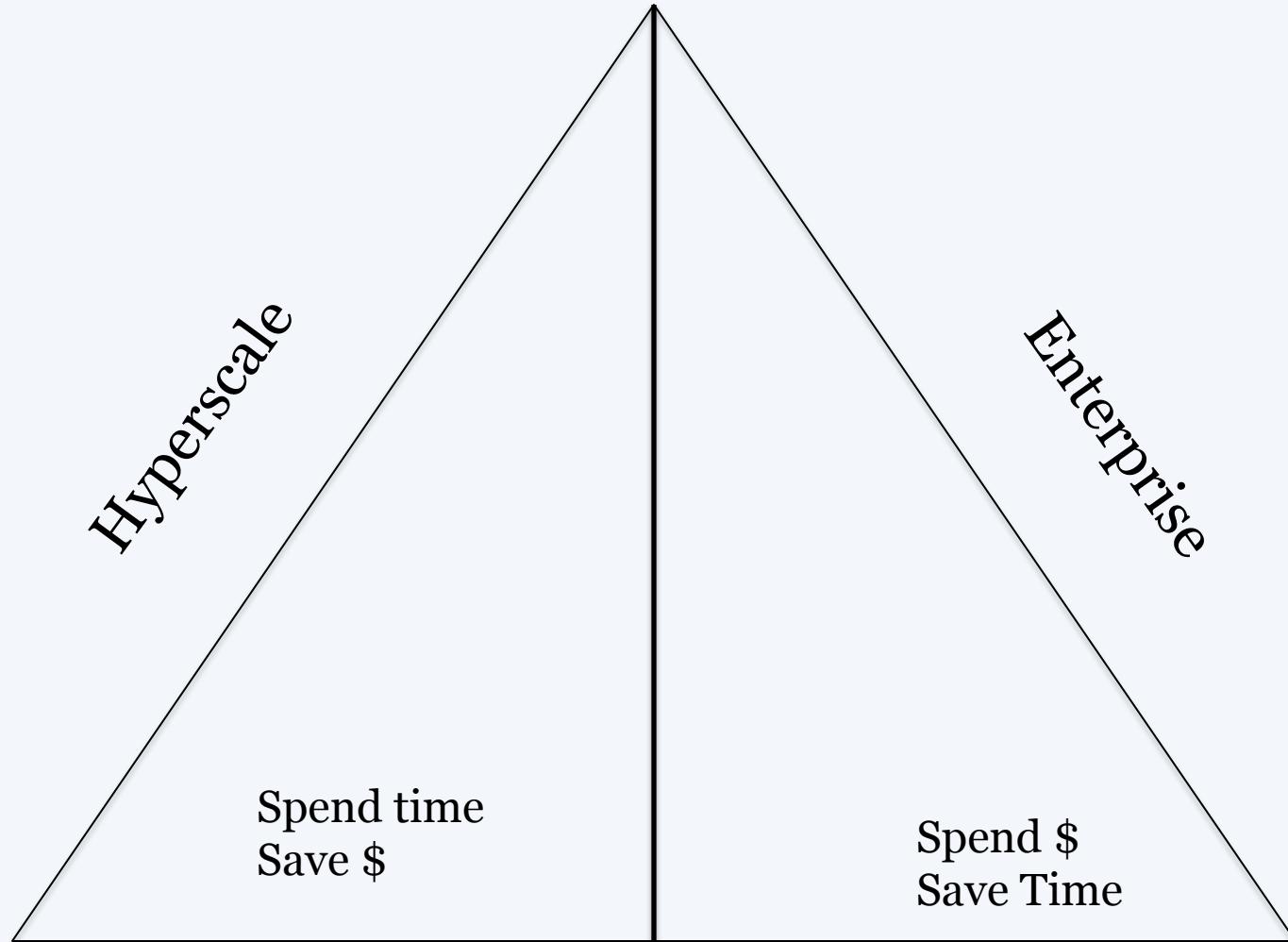
facebook

amazon.com®

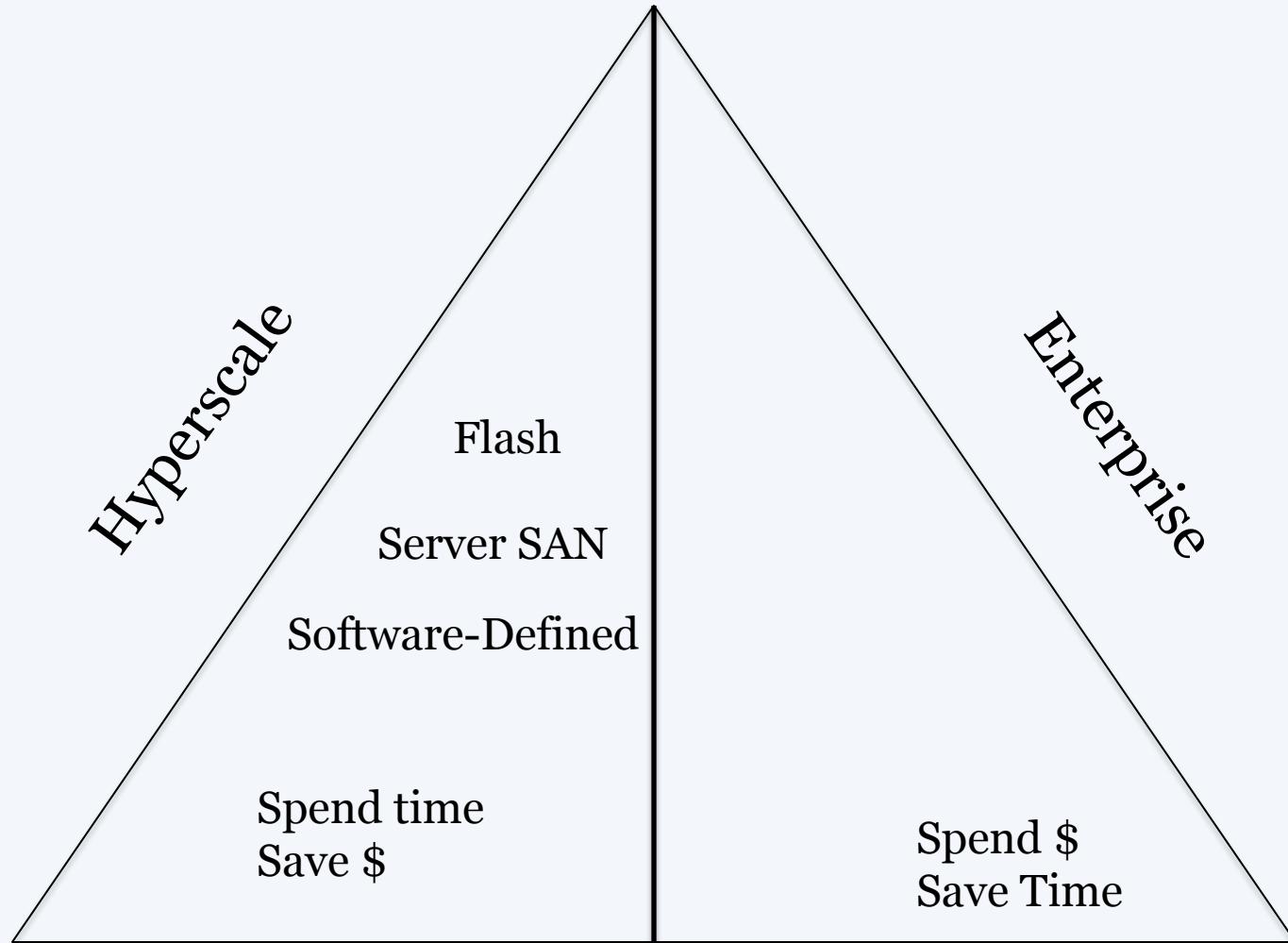


ebay™

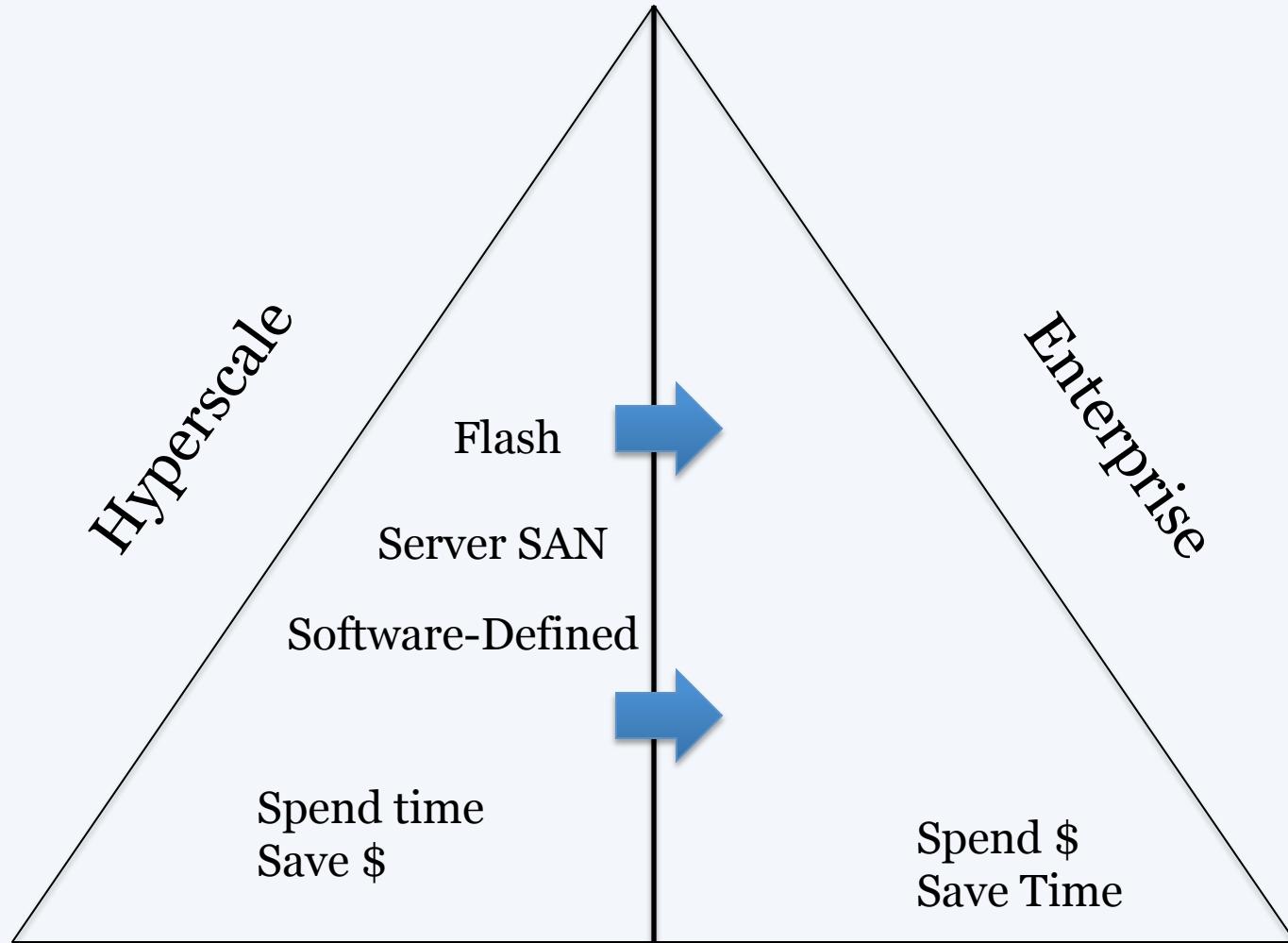
Enterprise learning from Web-Scale



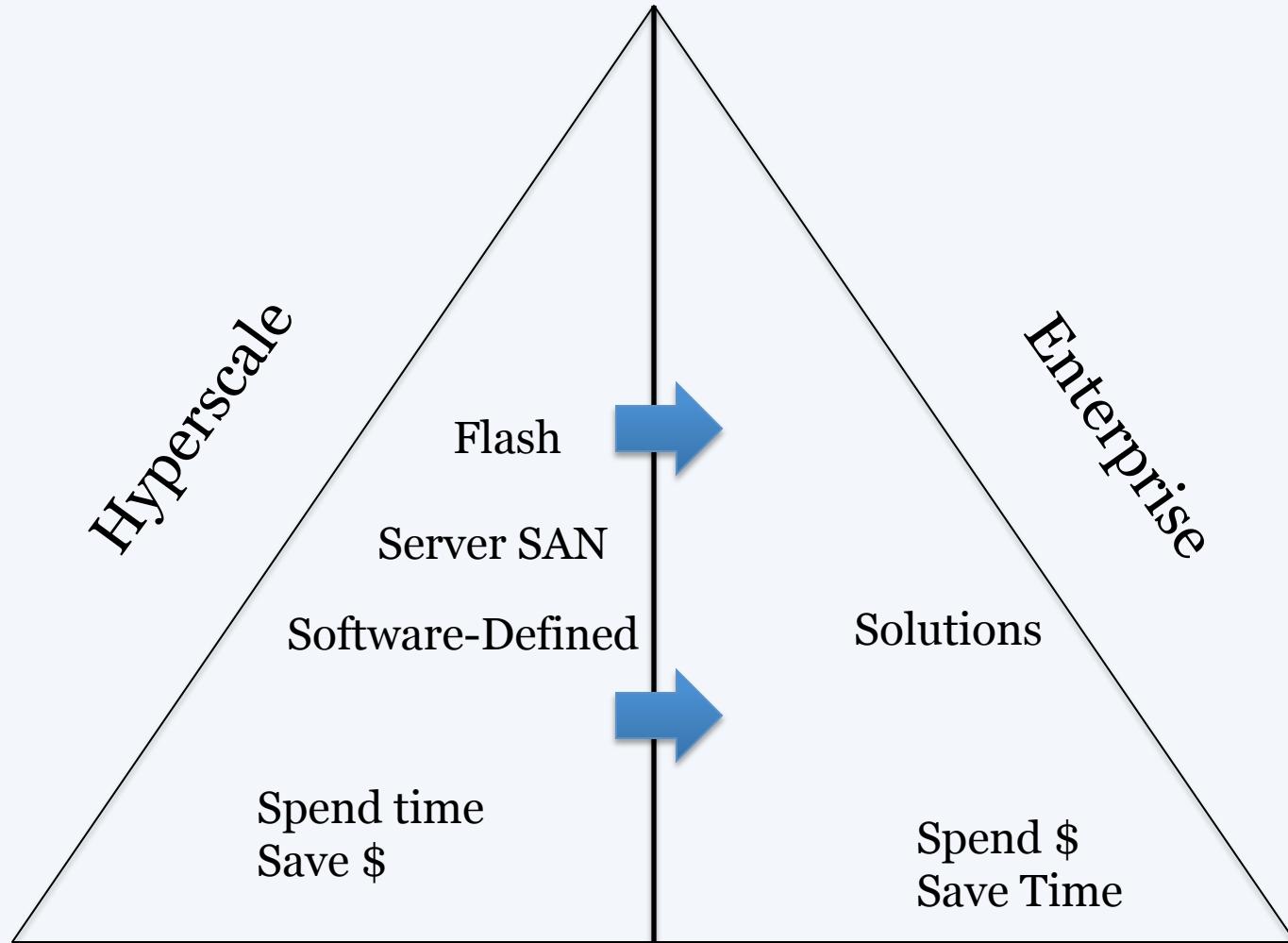
Enterprise learning from Web-Scale



Enterprise learning from Web-Scale



Enterprise learning from Web-Scale



Software Led Infrastructure



- Dynamic Infrastructure - Scale Out

Software Led Infrastructure



- Dynamic Infrastructure - Scale Out
- Automation - Software led solutions

- Dynamic Infrastructure - Scale Out
- Automation - Software led solutions
- Simplification - Rack Level convergence

- Dynamic Infrastructure - Scale Out
- Automation – Software led solutions
- Simplification – Rack Level convergence
- New Design/Architecture
 - Flash
 - Cloud
 - Network
 - Distributed Systems

What did “they” do ?

- Commodity Hardware
- Software Defined Software
(infrastructure)
- Avoid RAID
- Hyper-convergence
- Leverage Flash for Cache/Hot data

Enterprise Goals !

- All Computer/Storage hardware is compatible across the data center
- Rack-level configurations are not dependent on any one piece of code.
- 1 administrator for every 20,000 servers.
- Deploy hardware/software very quickly
- Invest in technical innovation
- Avoid siloes through transparency
- Use Automation



Micro-
services

Containers
Internet Scale
Applications

Internet of
things

Security

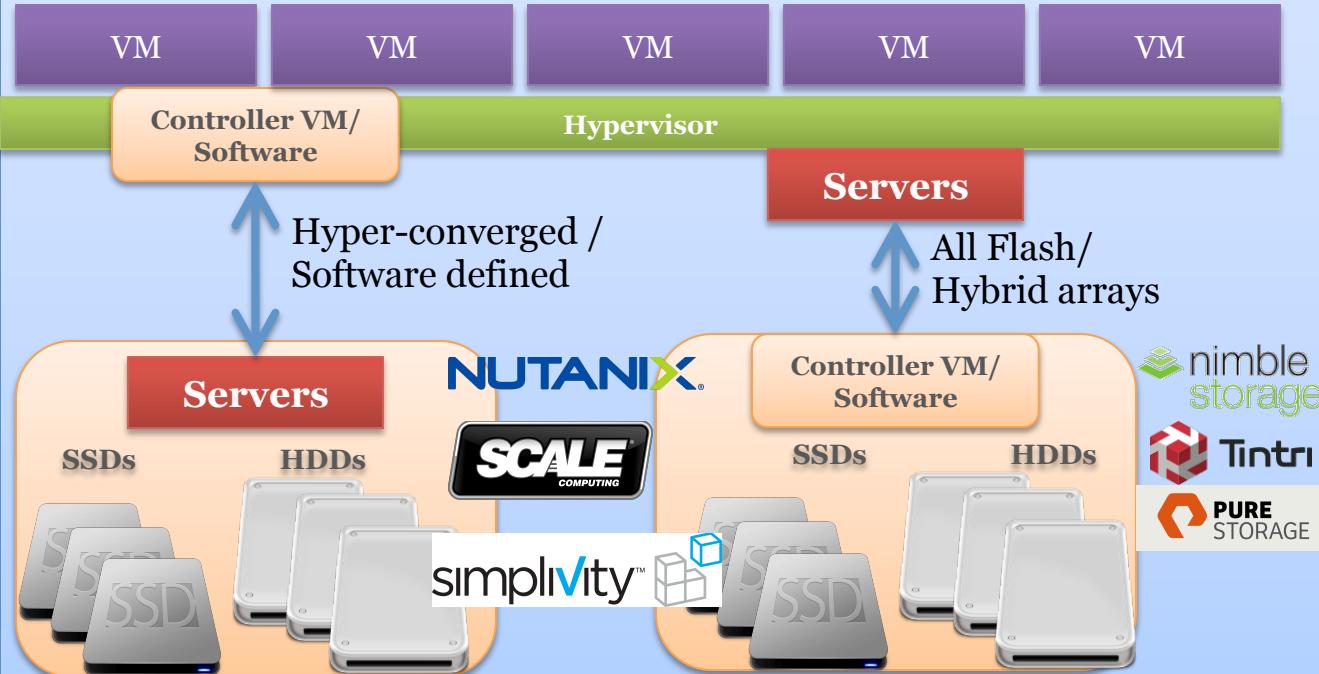
Real Time streams of data

Supply
Chains

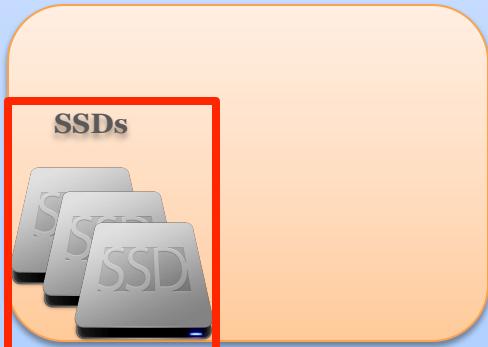
Click
streams

Real-time Analytics and
Batch processing

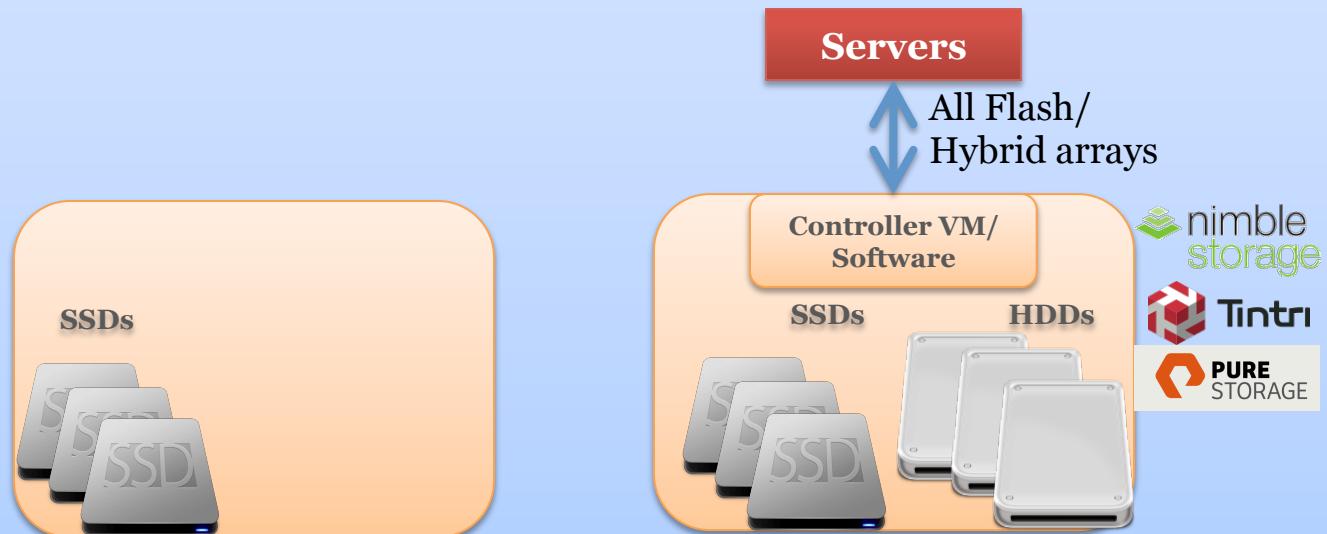




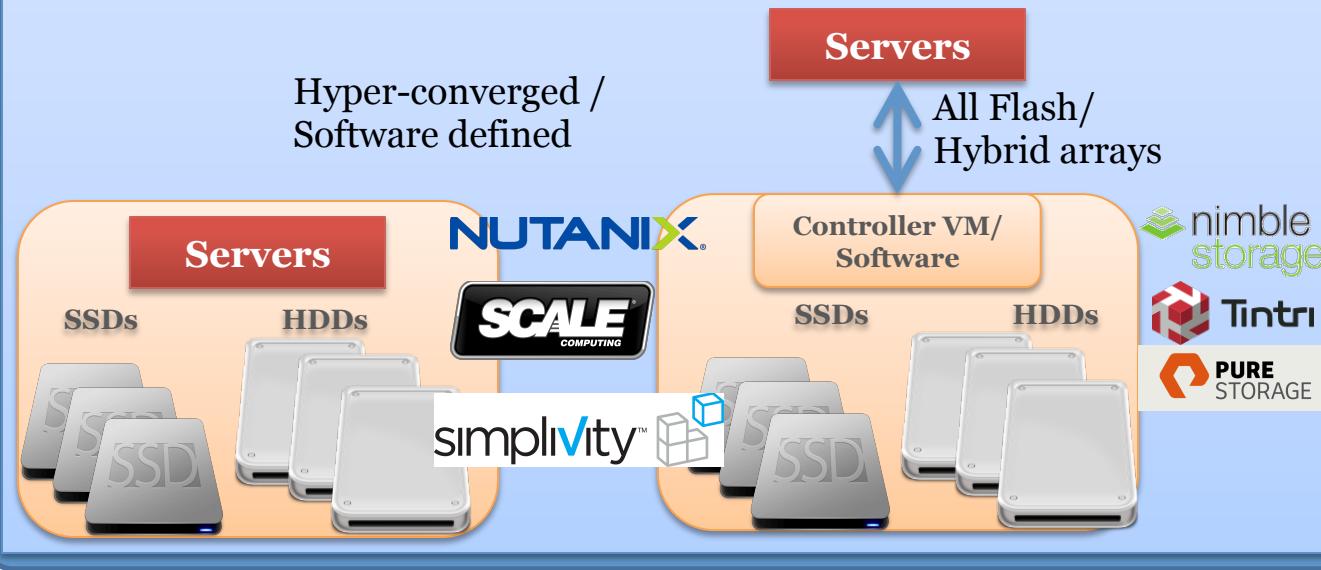
- Flash Storage and Solid State Drives



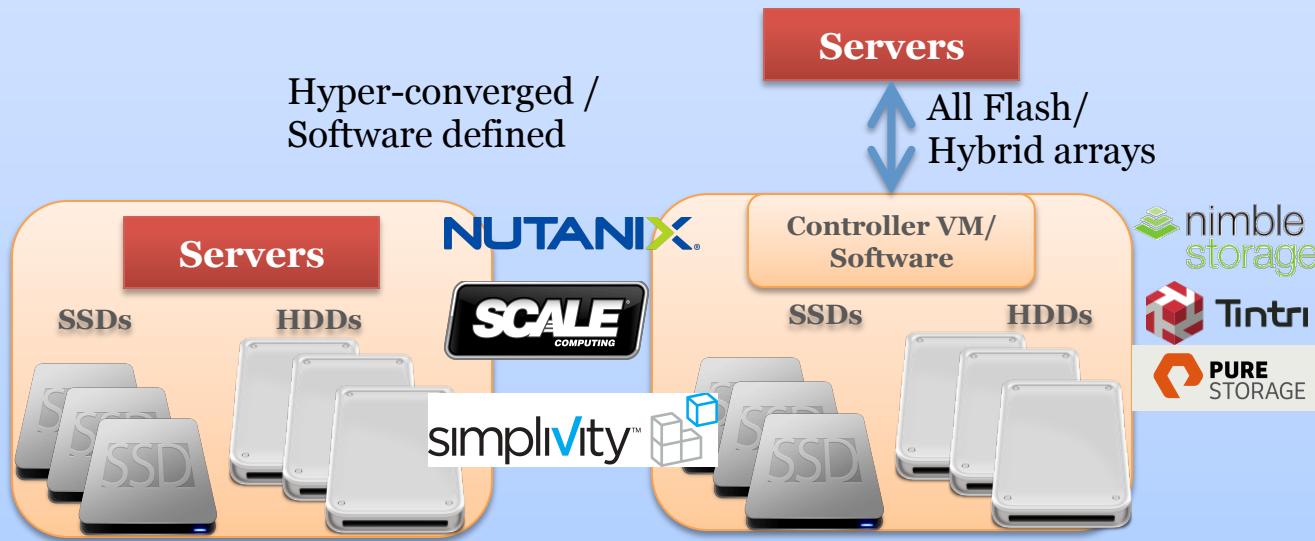
- Flash Storage and Solid State Drives
- All Flash or Hybrid Flash Arrays



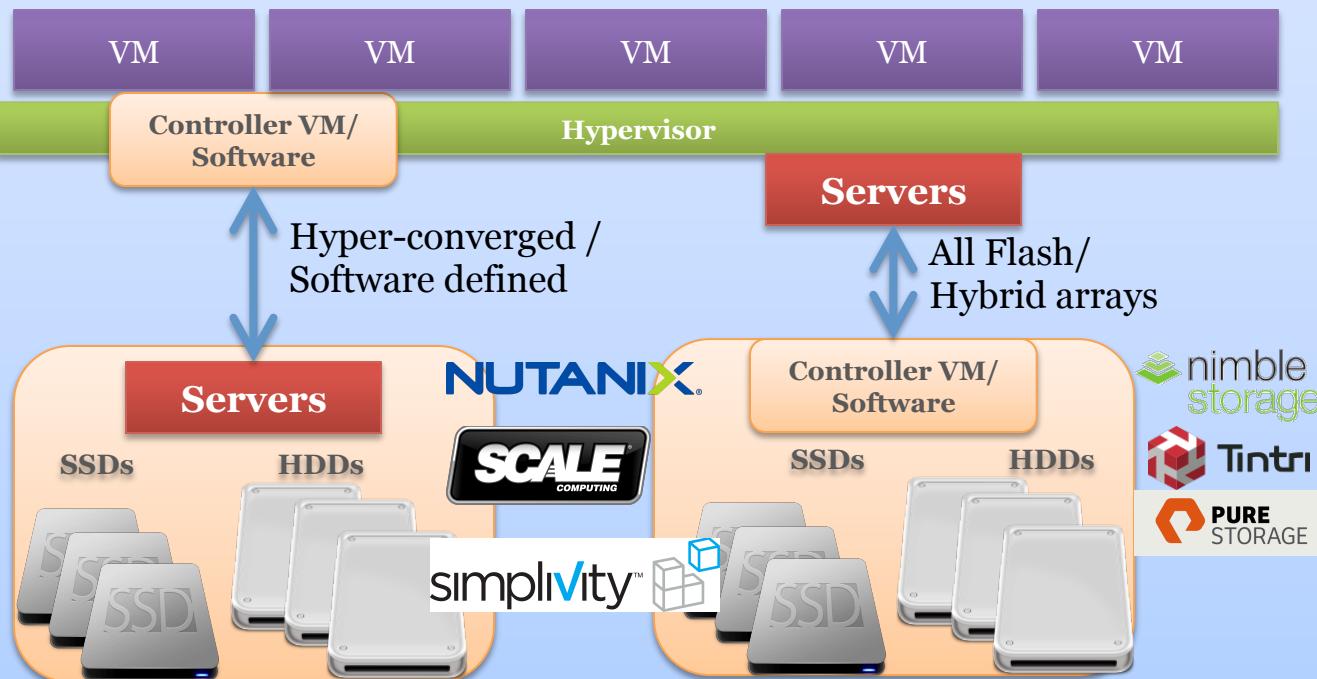
- Flash Storage and Solid State Drives
- All Flash or Hybrid Flash Arrays
- Software Defined Software

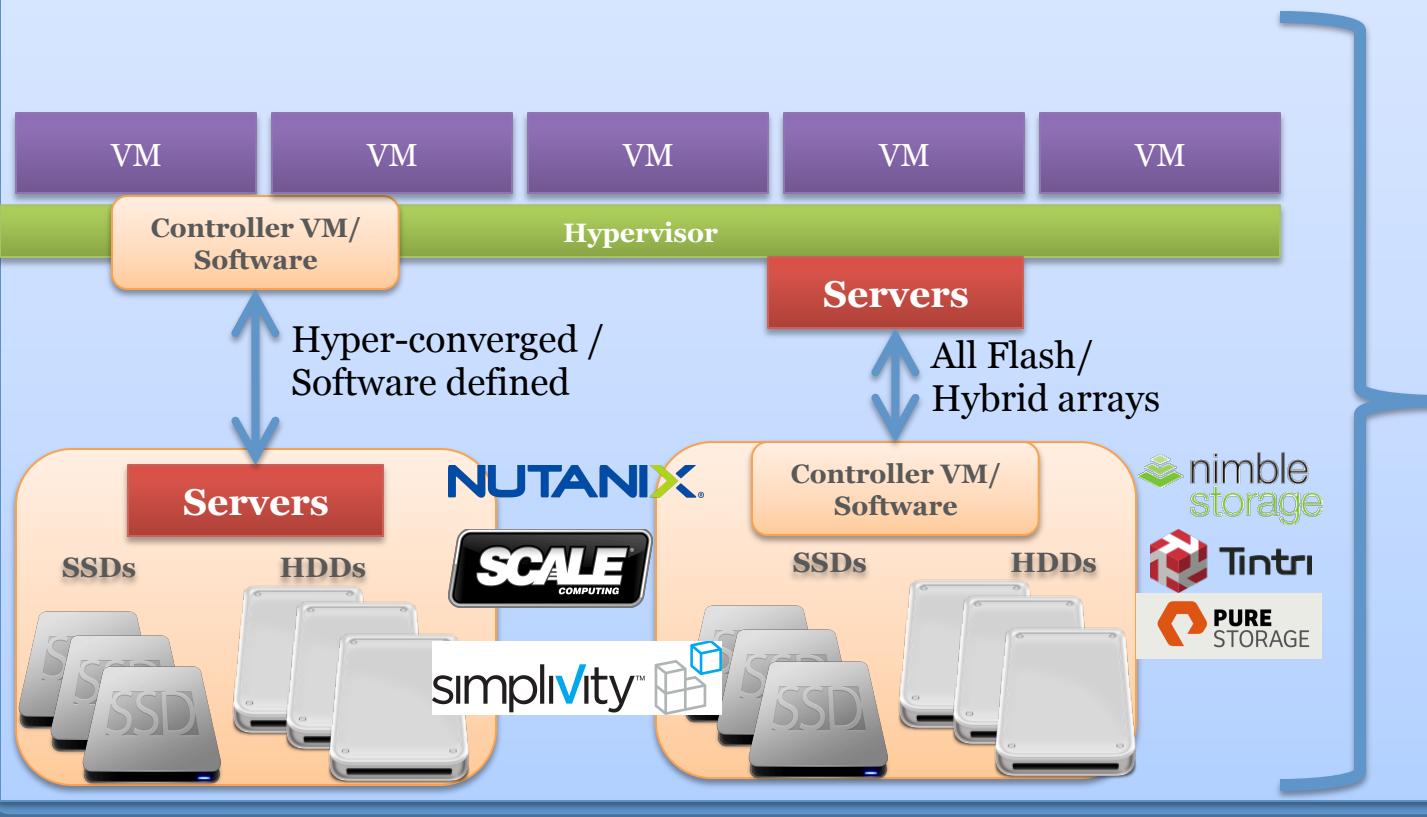


- Flash Storage and Solid State Drives
- All Flash or Hybrid Flash Arrays
- Software Defined Software
- Hyper converged Infrastructure

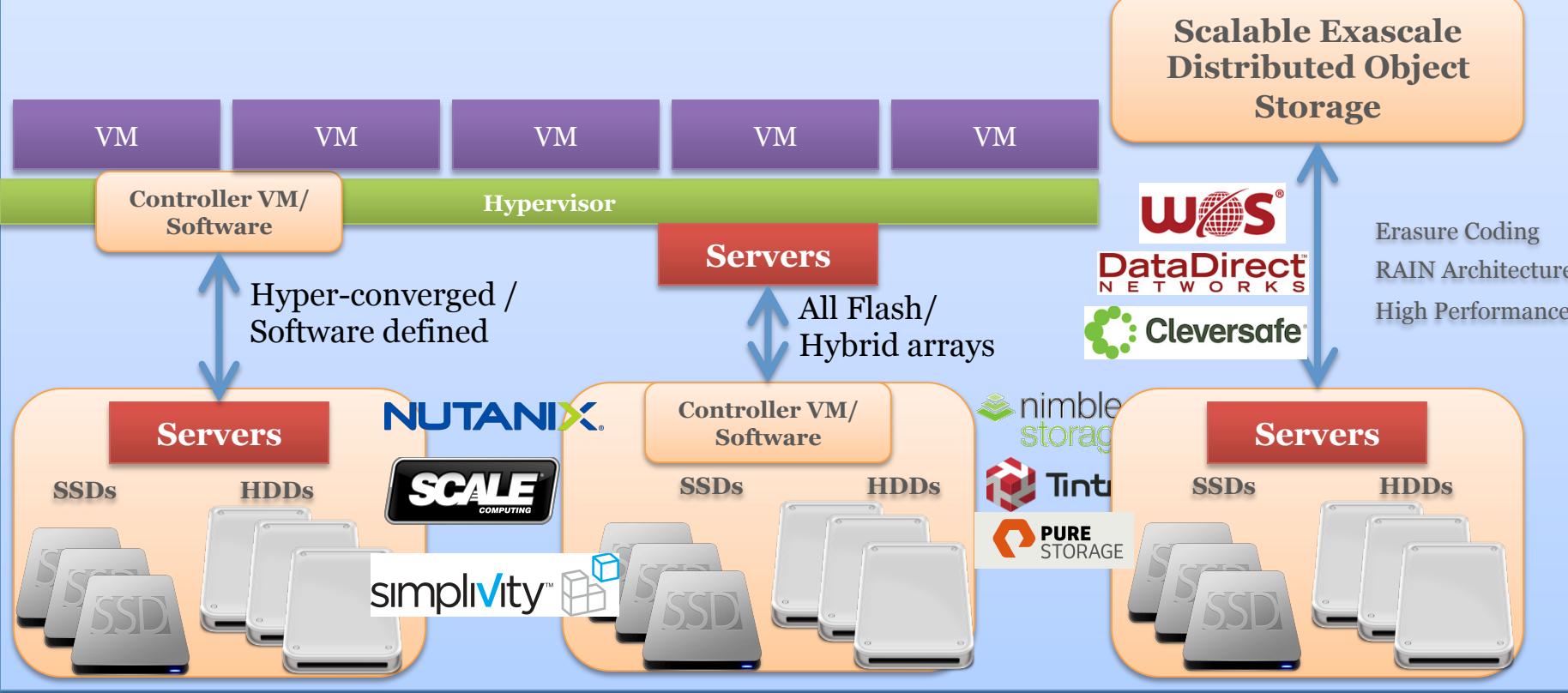


- Flash Storage and Solid State Drives
- All Flash or Hybrid Flash Arrays
- Software Defined Software
- Hyper converged Infrastructure
- IT-as-a- Service

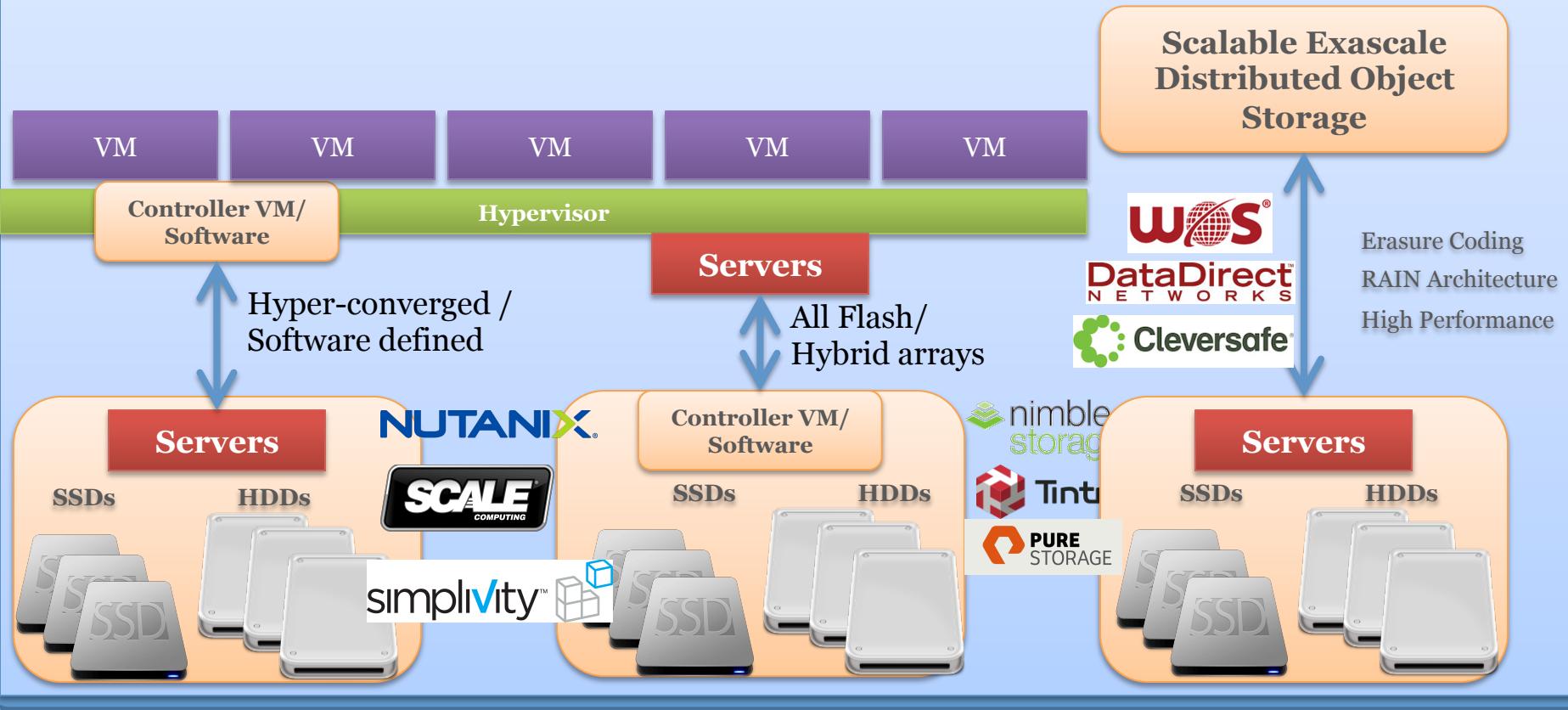




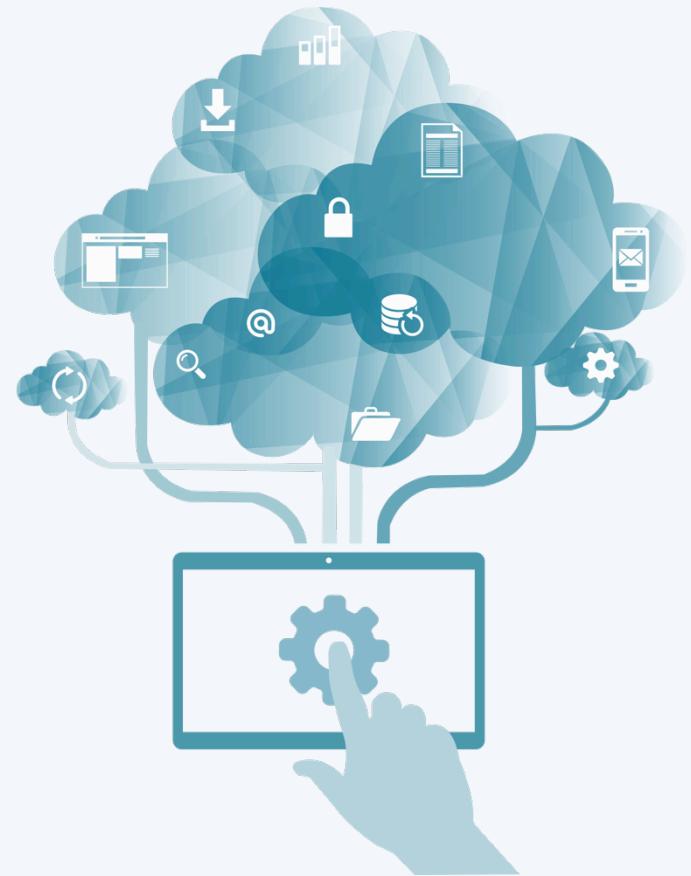
Virtual Private Computing Platform

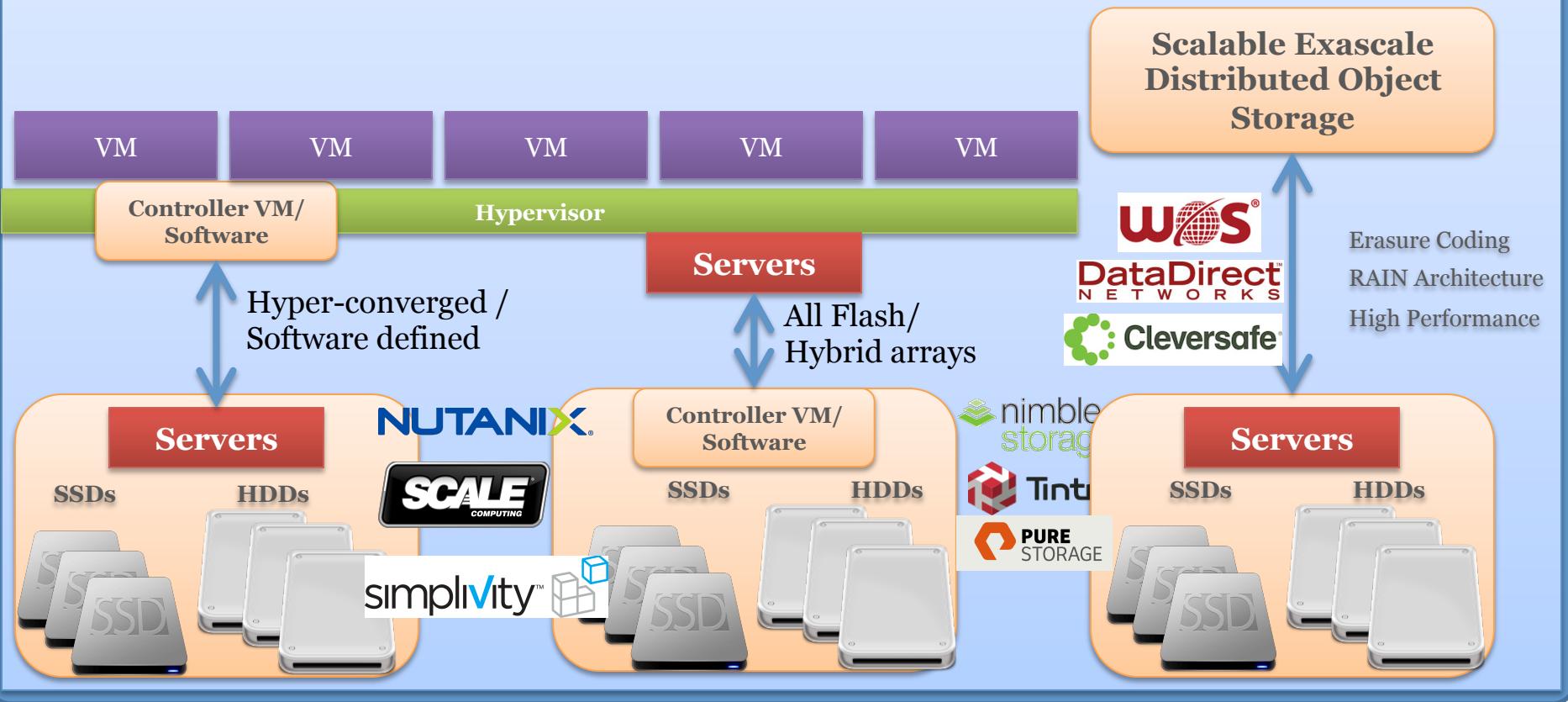


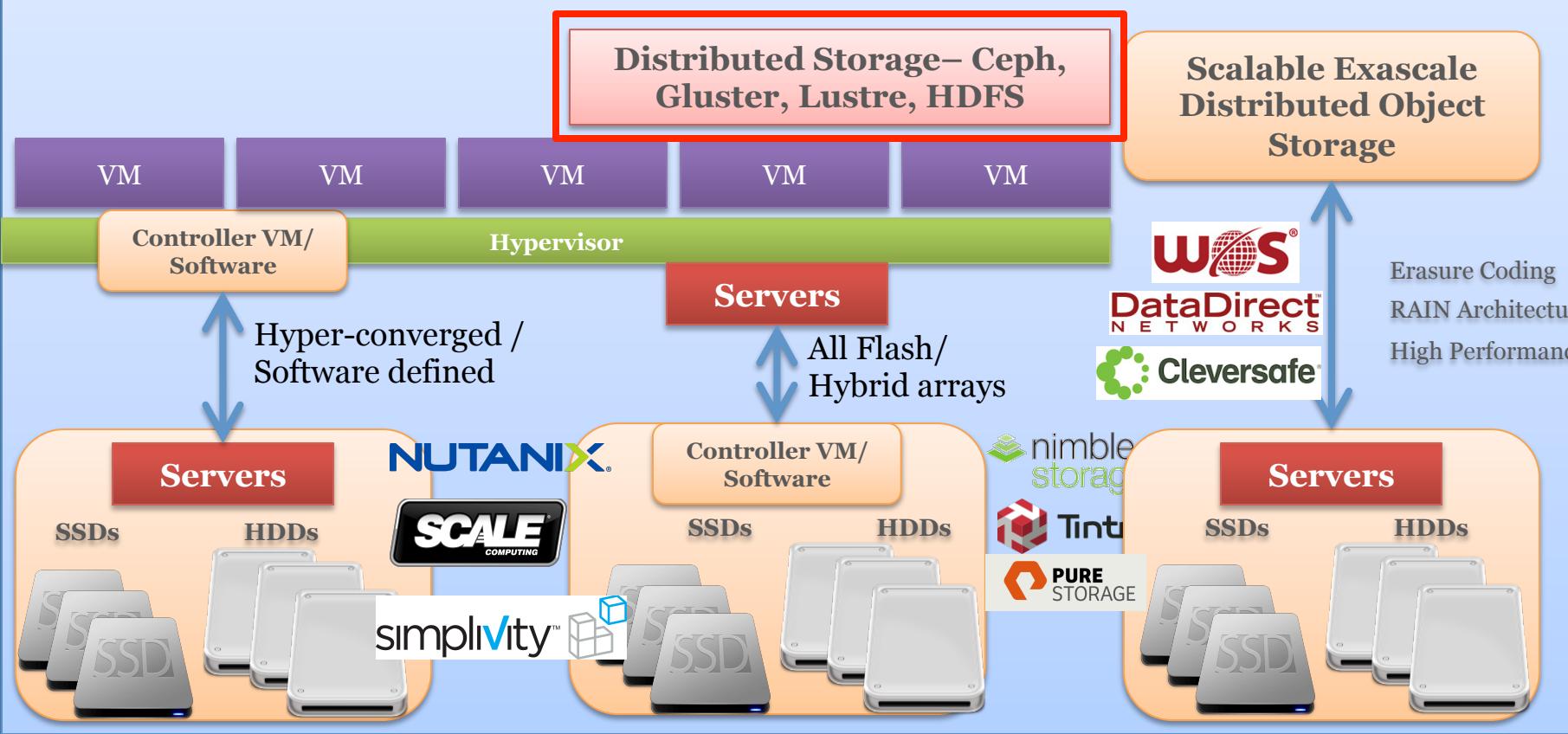
- Object Storage
- Erasure Coding
- Information Dispersal Algorithms (IDA)
- No RAID – but RAIN !



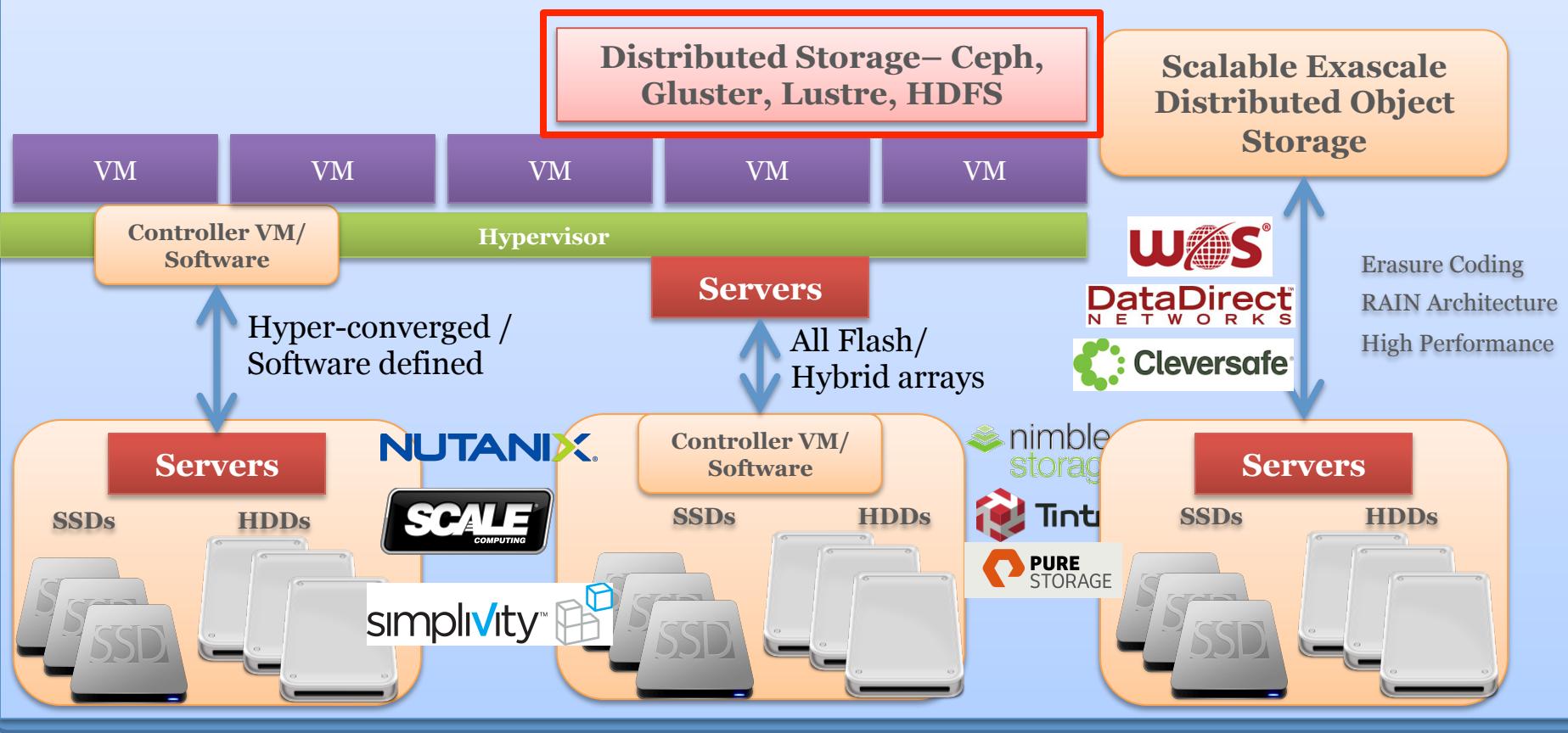
What about the Software to use ?







- Distributed Parallel Filesystems
- None of complexities of RAID
- Replication, Zero Downtime and Self-healing
- Commodity hardware with lowest cost



Distributed Systems – Mesos, Zookeeper

Distributed Storage– Ceph, Gluster, Lustre, HDFS

Scalable Exascale Distributed Object Storage

VM

VM

VM

VM

VM

Controller VM/
Software

Hypervisor

Servers

All Flash/
Hybrid arrays

Servers

SSDs

HDDs

NUTANIX

SCALE
COMPUTING

Controller VM/
Software

SSDs

HDDs

simpliVITY™

WOS®
DataDirect™
NETWORKS

Cleversafe®

Erasure Coding
RAIN Architecture
High Performance

Servers

SSDs

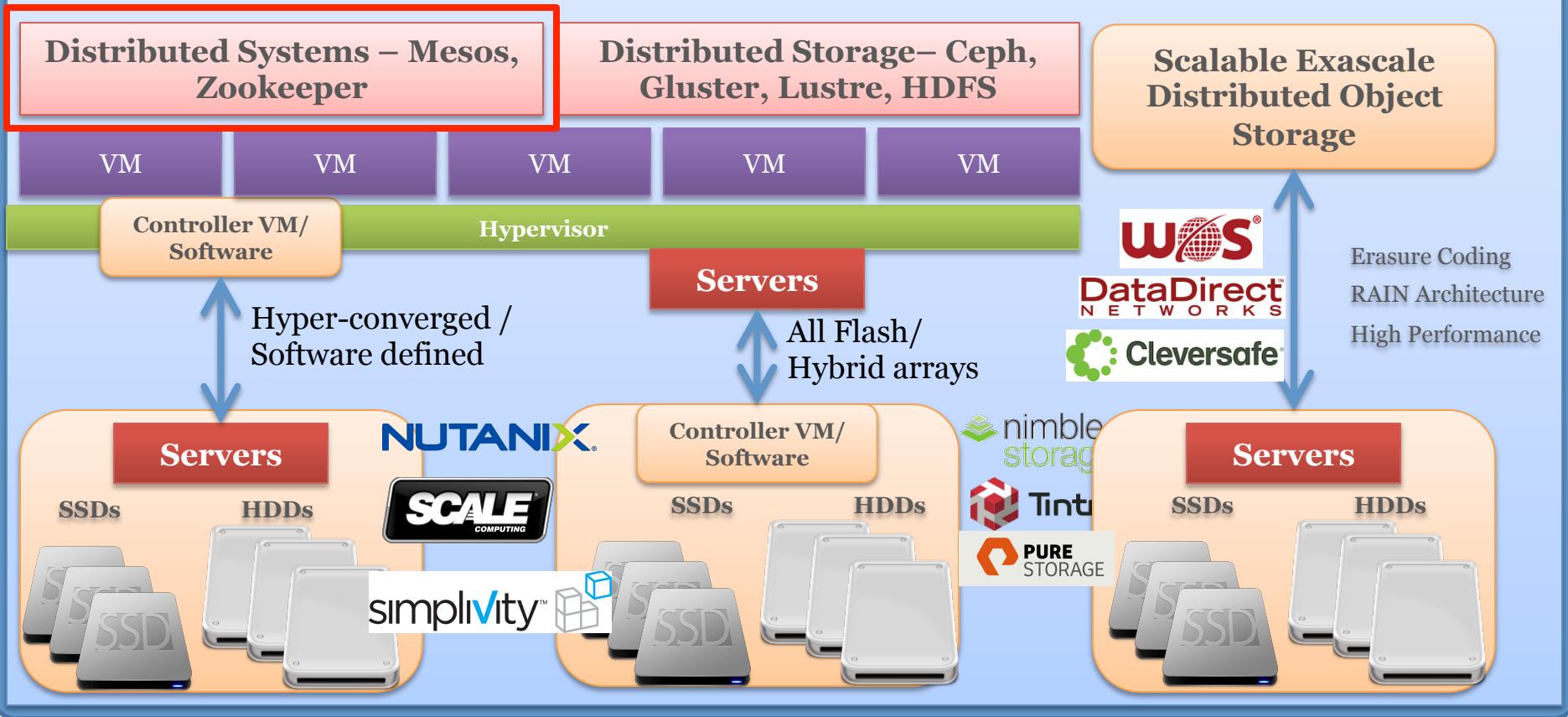
HDDs

nimble
storage

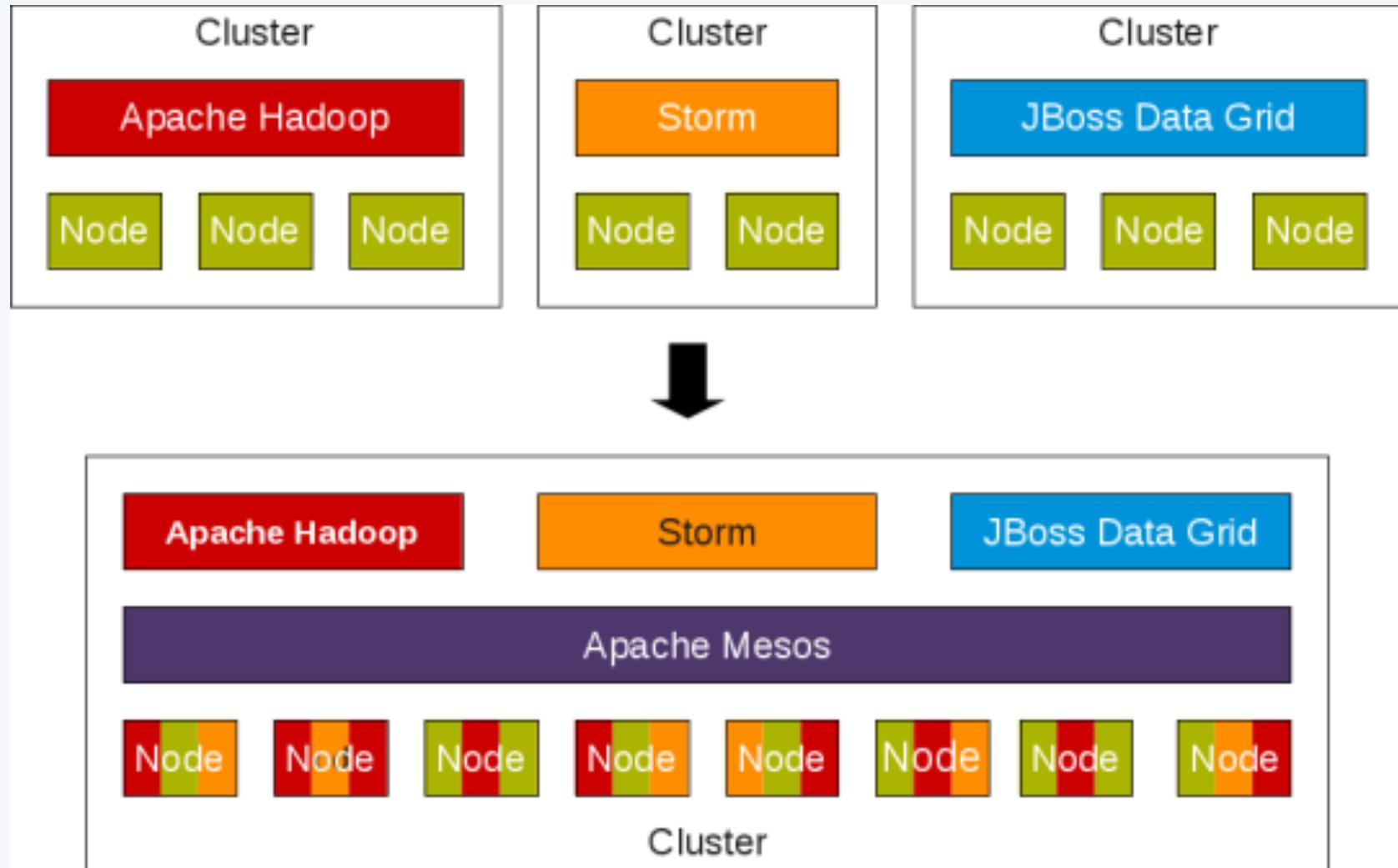
Tinti

PURE
STORAGE

- Apache Mesos - Operating system for the Datacenter
- Apache Zookeeper – Distributed Co-ordination Service
- Consul, Eureka – Service Discovery



Mesos – Server Abstraction



Distributed Systems – Mesos, Zookeeper

Distributed Storage– Ceph, Gluster, Lustre, HDFS

Scalable Exascale Distributed Object Storage

VM

VM

VM

VM

VM

Controller VM/
Software

Hypervisor

Servers

Hyper-converged /
Software defined

Servers

SSDs

HDDs

NUTANIX

SCALE
COMPUTING

simpliVity™

Controller VM/
Software

SSDs

HDDs

WOS®

DataDirect™
NETWORKS

Cleversafe®

Erasure Coding
RAIN Architecture
High Performance

Servers

SSDs

HDDs

nimble
storage

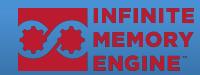
Tinti

PURE
STORAGE

In-memory Computing



Burst Buffer Solutions



Distributed Systems – Mesos, Zookeeper

VM

VM

VM

VM

VM

Controller VM/
Software

Hypervisor

Hyper-converged /
Software defined

Distributed Storage– Ceph, Gluster, Lustre, HDFS

Burst Buffer Solutions

Burst Buffer Solutions

Servers

All Flash/
Hybrid arrays

Scalable Exascale Distributed Object Storage



Erasure Coding
RAIN Architecture
High Performance

Servers

Controller VM/
Software

SSDs

HDDs



NUTANIX



simpliVity™

SSDs

HDDs



nimble storage



PURE STORAGE

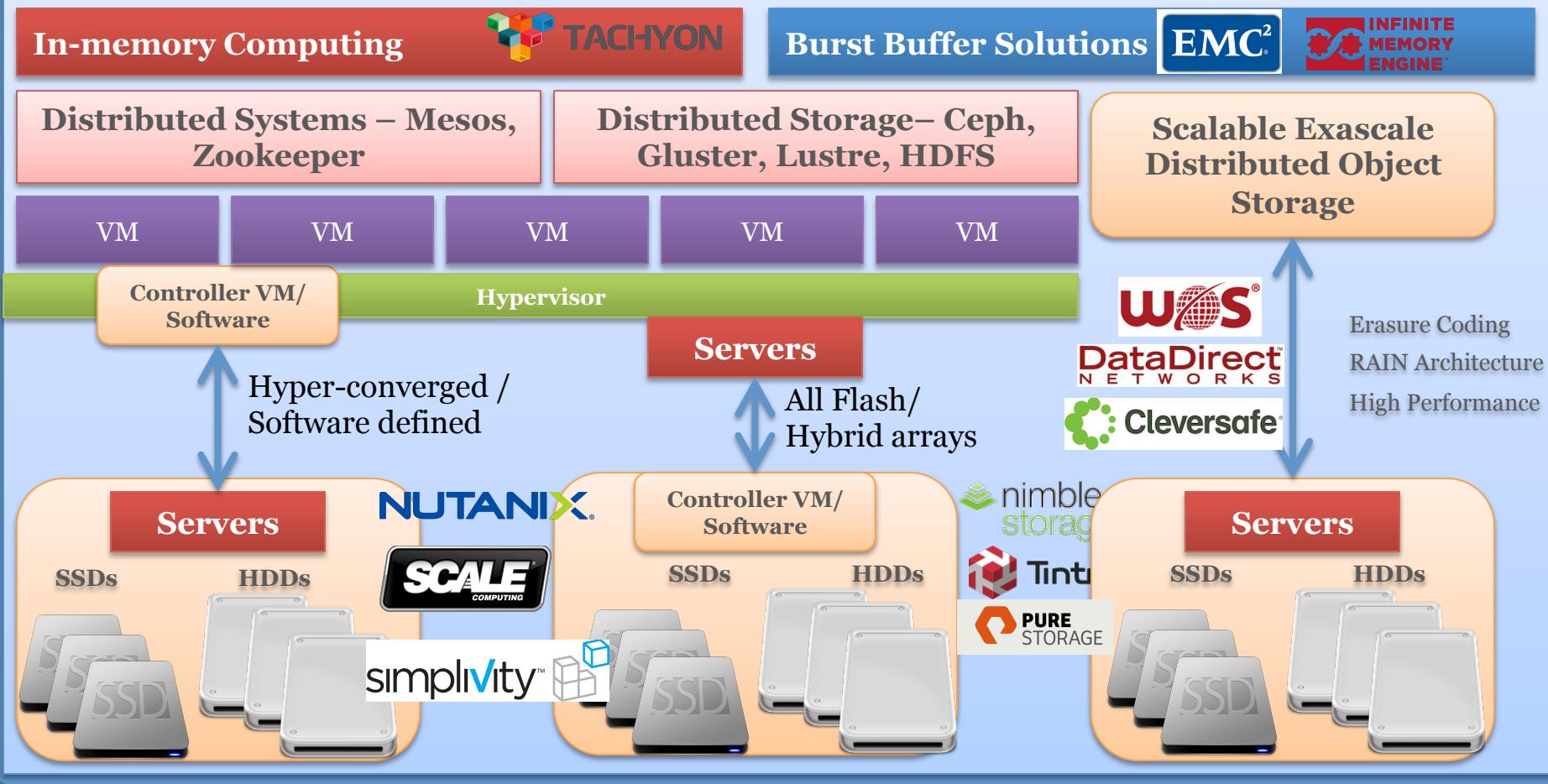
Servers

SSDs

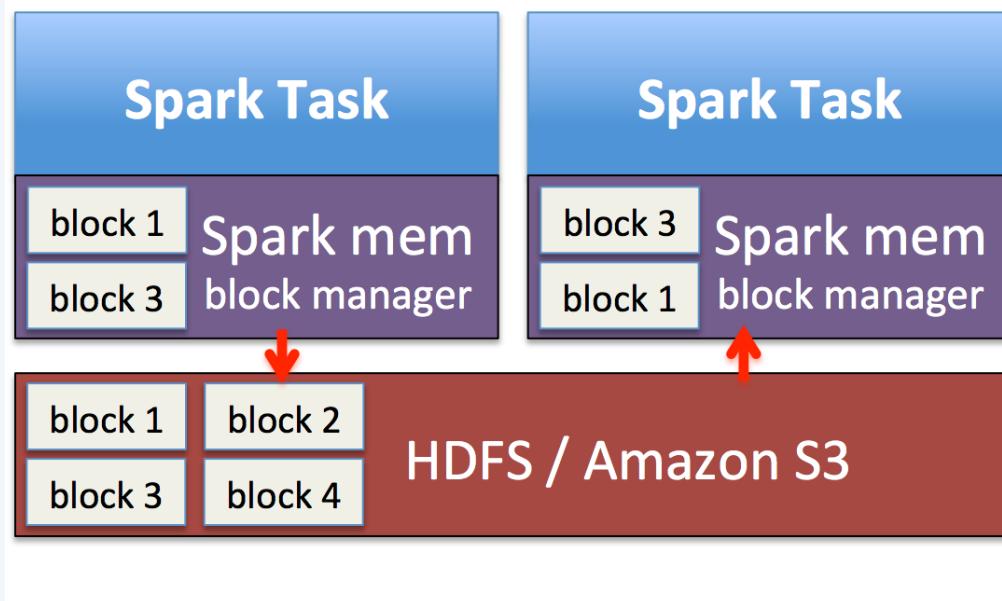
HDDs



- Tachyon – Reliable data sharing at memory-speed across clusters



Tachyon - Current issues



- Data sharing is a bottleneck due to slow writes to disk
- Cache loss when process crashes
- In-memory data duplication and Java Garbage Collection

Tachyon - memory centric storage



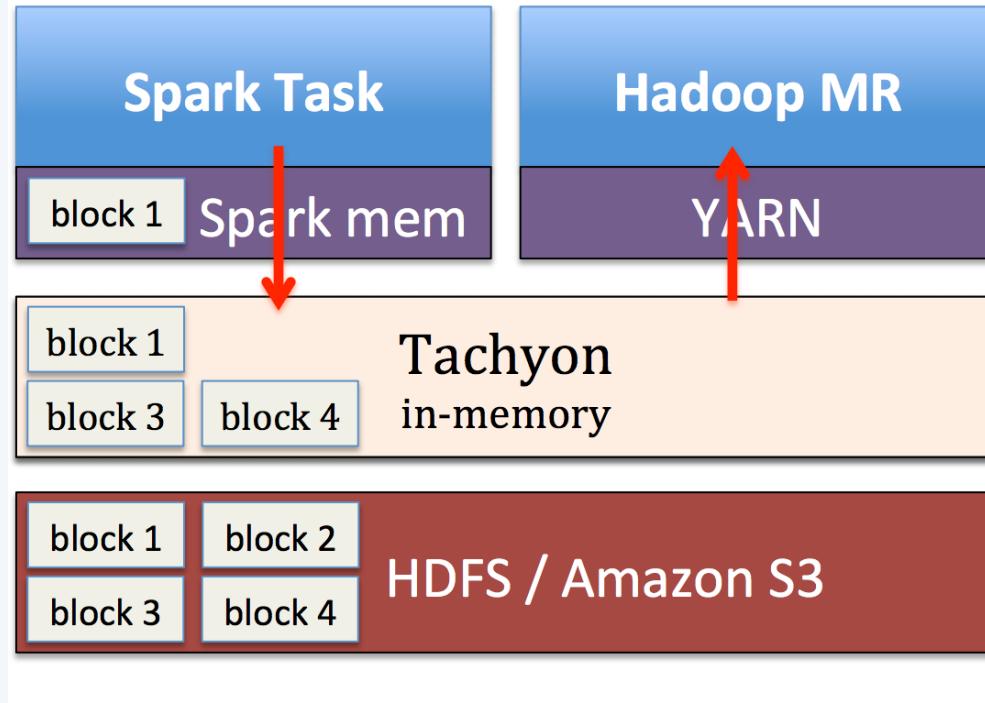
Reliable data sharing at memory-speed
within and across cluster frameworks/
jobs

Computation Frameworks
(Spark, MapReduce, Impala, H2O, ...)

Tachyon

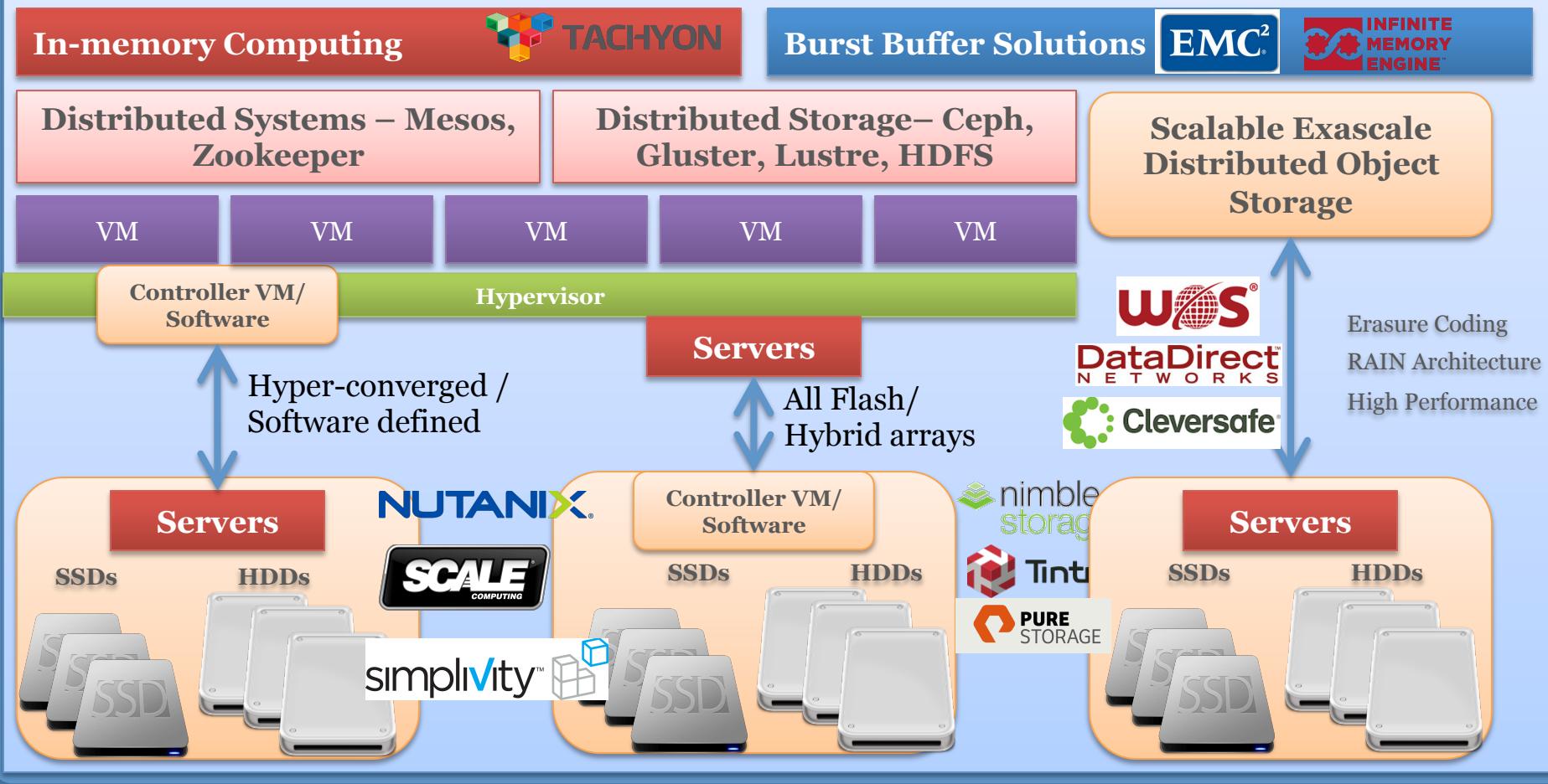
Existing Storage Systems
(HDFS, S3, GlusterFS, ...)

Tachyon - memory centric storage

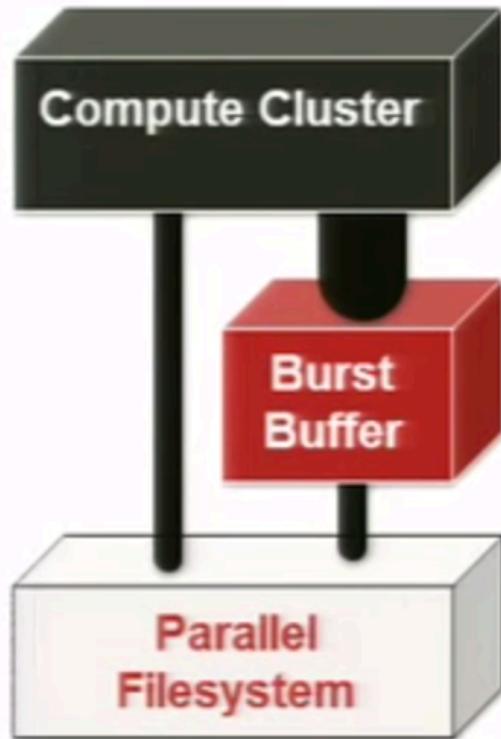


- Memory-speed data sharing among jobs in different frameworks
- Keep in-memory data safe, even when a job crashes.
- No in-memory data duplication, much less GC

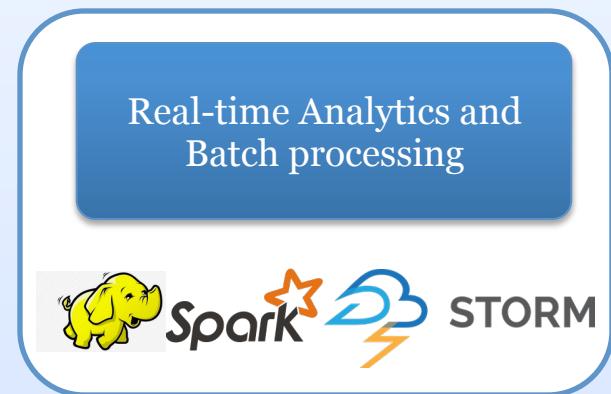
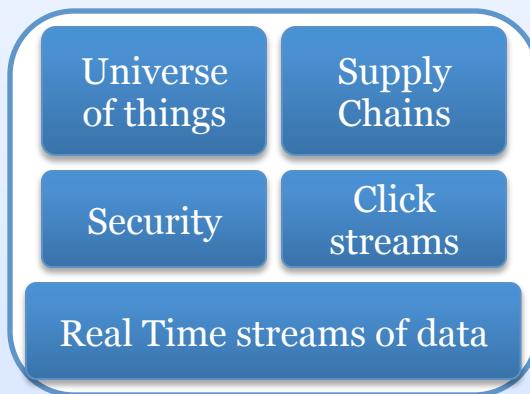
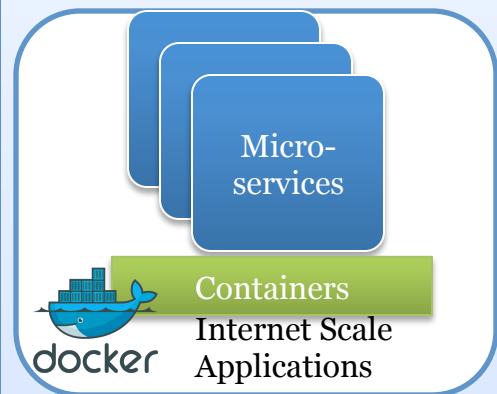
- Tachyon – Reliable data sharing at memory-speed across clusters
- Burst Buffers - Latency reduction, greater bandwidth and high IOPS performance



Burst Buffer Designs



- Introduce fast buffer layer
- Layer between memory and persistent storage
 - pre-stage application data
 - Buffer writes from memory to fast devices
 - Store immediate application data
- Still a “mount point” – POSIX Compliance



In-memory Computing



TACHYON

Burst Buffer Solutions



EMC²
INFINITE MEMORY ENGINE

Distributed Systems – Mesos, Zookeeper



Controller VM/ Software

Hyper-converged / Software defined

Distributed Storage– Ceph, Gluster, Lustre, HDFS



Scalable Exascale Distributed Object Storage



DataDirectTM
NETWORKS



Erasure Coding
RAIN Architecture
High Performance

Servers

All Flash/
Hybrid arrays

Servers

SSDs HDDs



NUTANIX



simpliVityTM

Controller VM/ Software

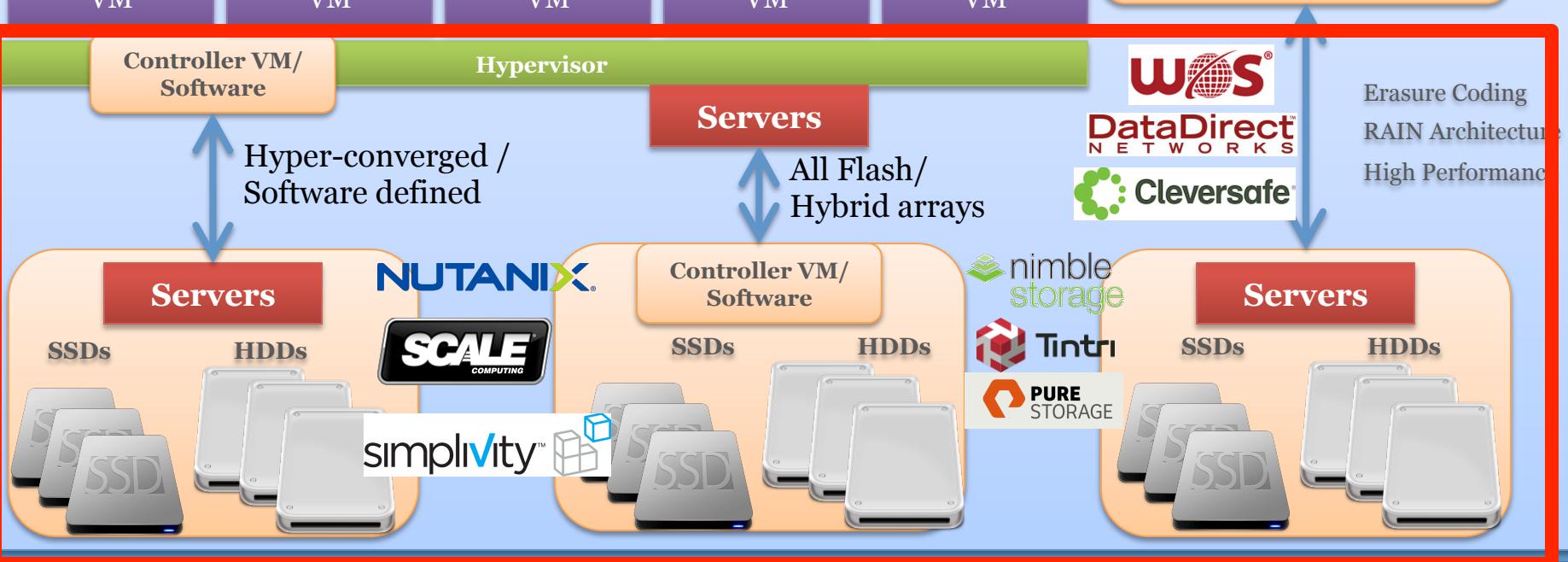
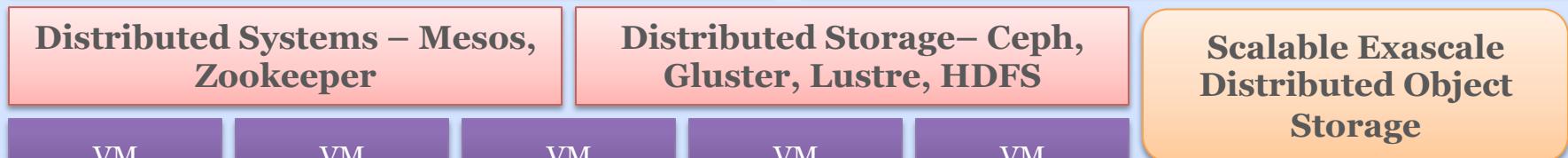
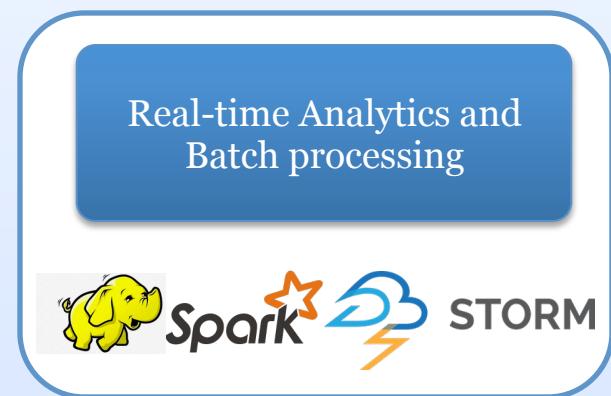
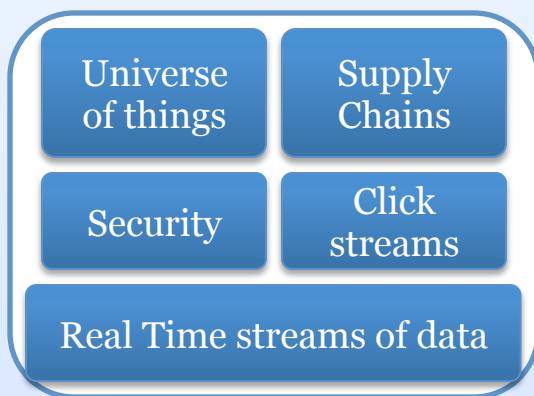
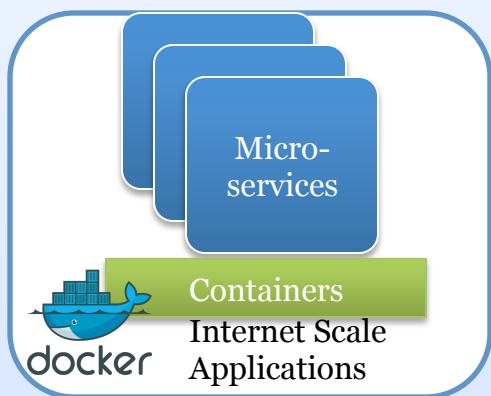
SSDs HDDs

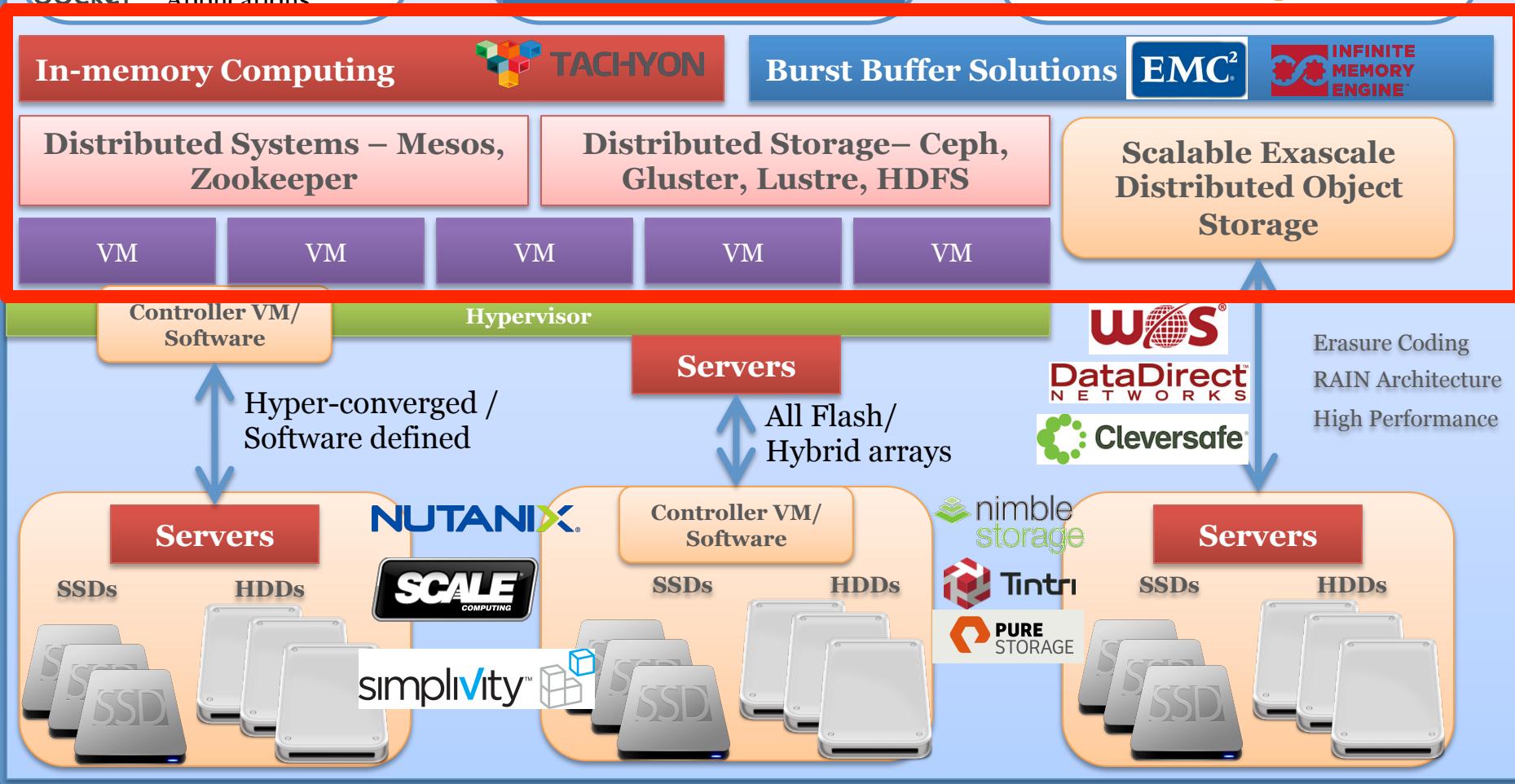
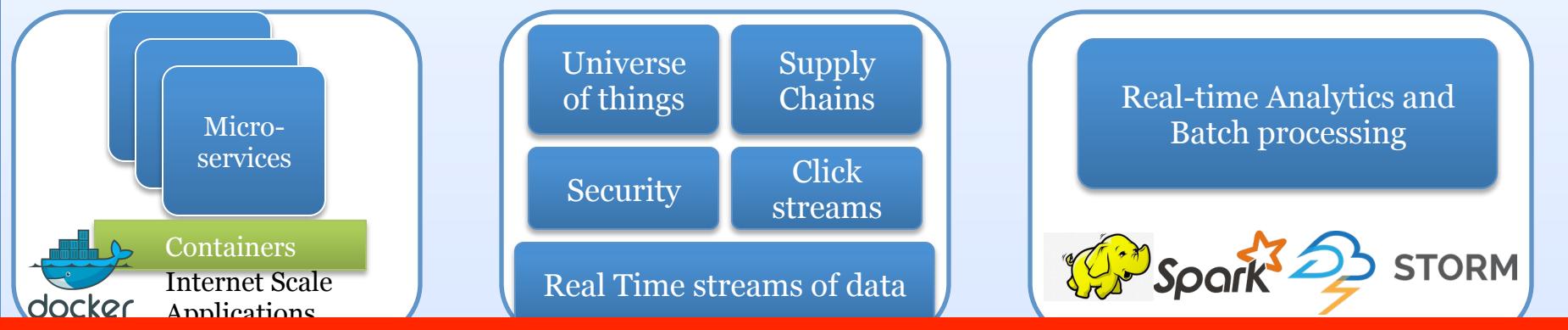


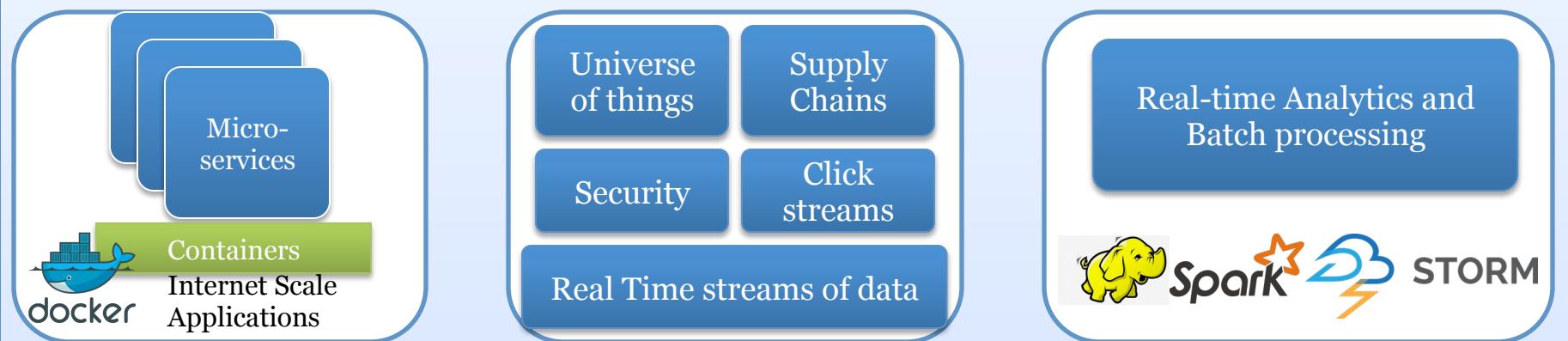
Servers

SSDs HDDs









In-memory Computing



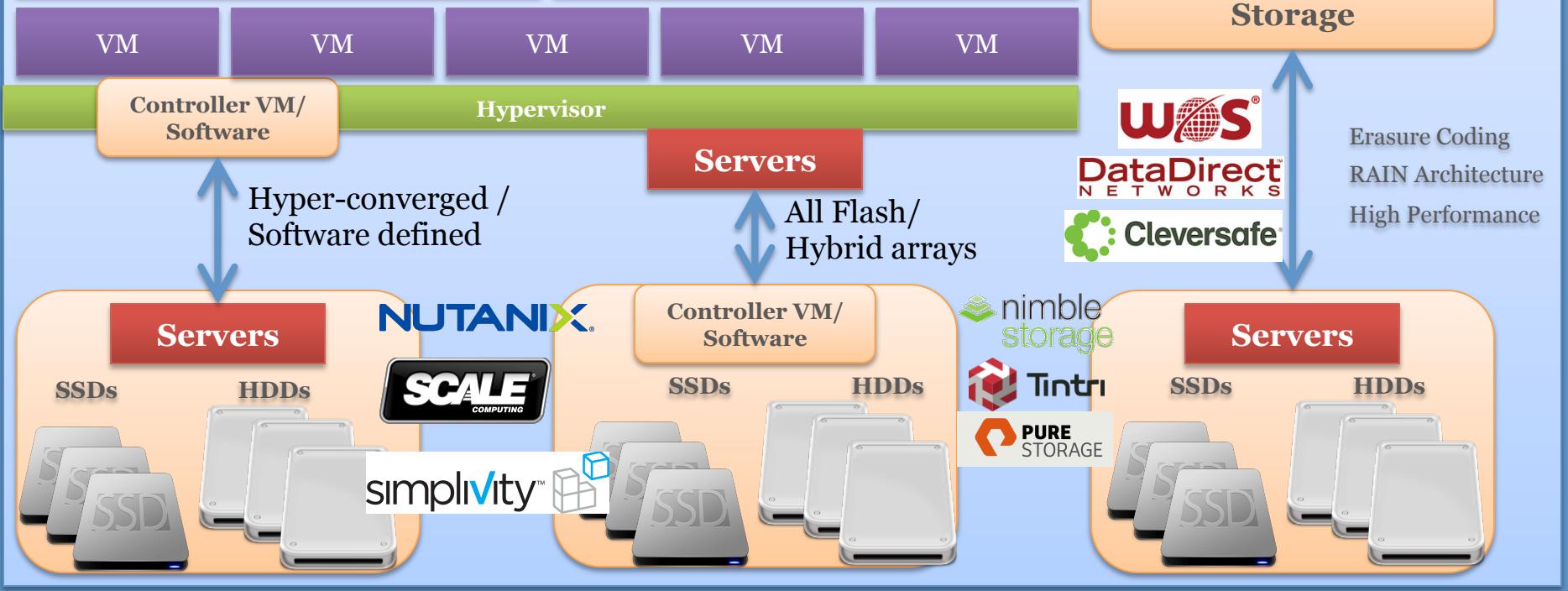
Burst Buffer Solutions



Distributed Systems – Mesos, Zookeeper

Distributed Storage– Ceph, Gluster, Lustre, HDFS

Scalable Exascale Distributed Object Storage



Questions & Queries!





Thank You!

MSys Georgia
4385 Kimball Bridge
Rd,
Suite 203, Johns Creek,
Georgia - 30022

Ph. +1 770-809-3217

MSys Chennai
Bristol IT Park, 4th Floor,
Plot No. 10, South Phase,
Thiru Vi Ka Industrial
Estate,
Guindy, Chennai 600032

Ph. +91-44-39167015

MSys Bangalore
No: 56/3, Ground
Floor,
Vakil Square,
Bannerghatta Road,
Bangalore -560 029

Ph. +91-80-41158363

Clogeny Pune
Plot no. 34/2, Rajiv
Gandhi
Infotech Park – Phase 1,
Hinjewadi, Pune - 411 057

Ph. +91 20 661 43 482
US Ph. +1 408 556 9645

E: info@msys-tech.com
W: www.msys-tech.com

f [/msystech](https://www.facebook.com/msystech)
t [@msys_tech](https://twitter.com/msys_tech)