



# Data Center Storage Solution - Redundancy in a Box

FTF-SNT-F1205

Madhu Indurthi | Freescale Semiconductor  
Storage Networking Systems Architect

Lance Smith | Servery Inc. President & CEO

Mike McKean | Synapse Design. Senior Director, Engineering

J U N E . 2 0 1 5



External Use

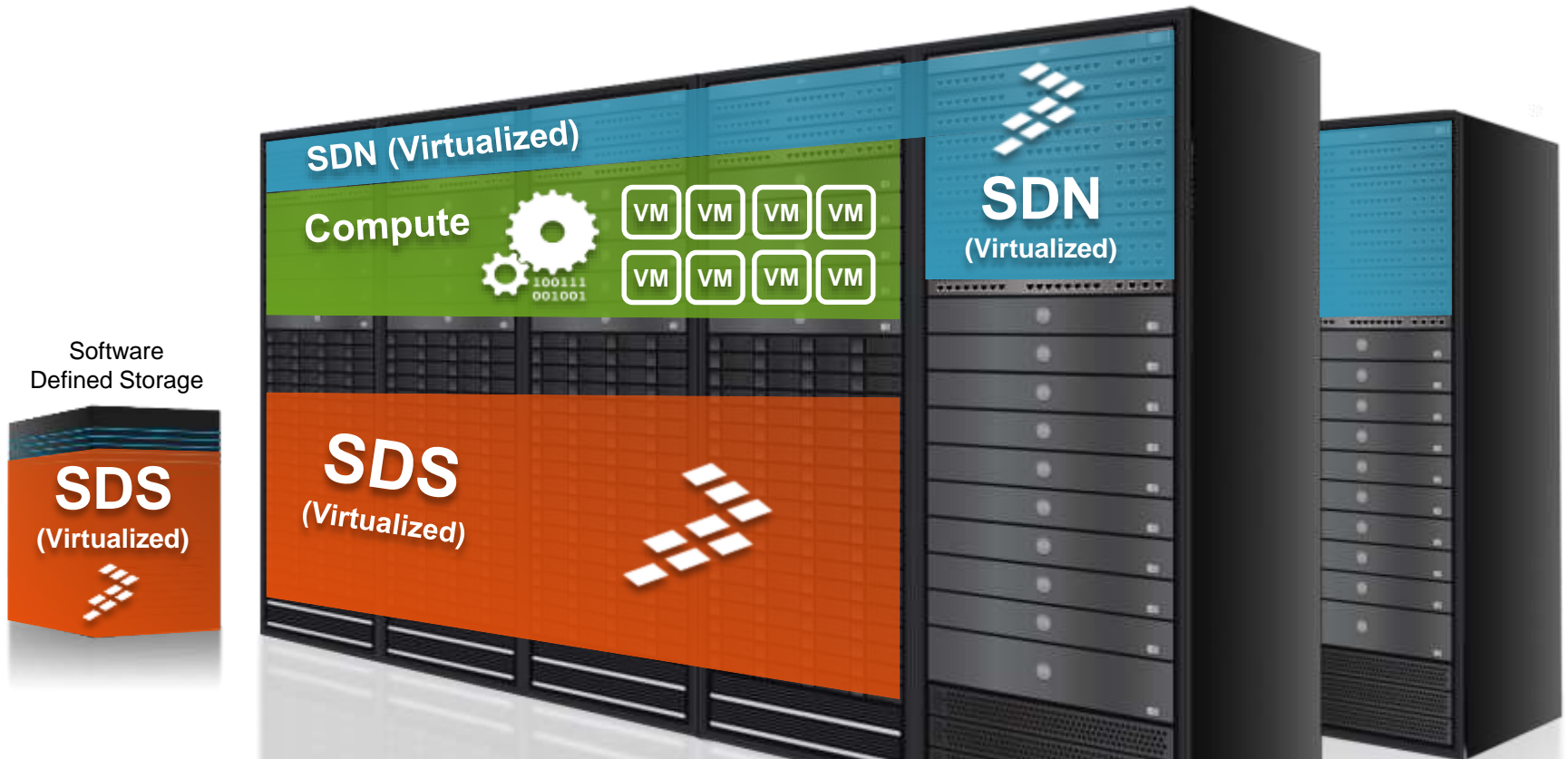


# Agenda

- Trends in the Data Center
- Freescale Storage Solutions – Software Defined Storage (SDS) Platform
- Storage Use Cases
- Storage Solutions for cNAS, Ethernet Drive, Storage Accelerators & iNIC's
- High-end Next Generation Storage Solutions
- Synapse: Data Center Cloud Solutions with Freescale
- Servergy: CYPHRE - SaaS Cloud Platform with Freescale



# Primary Data Center Market Trends

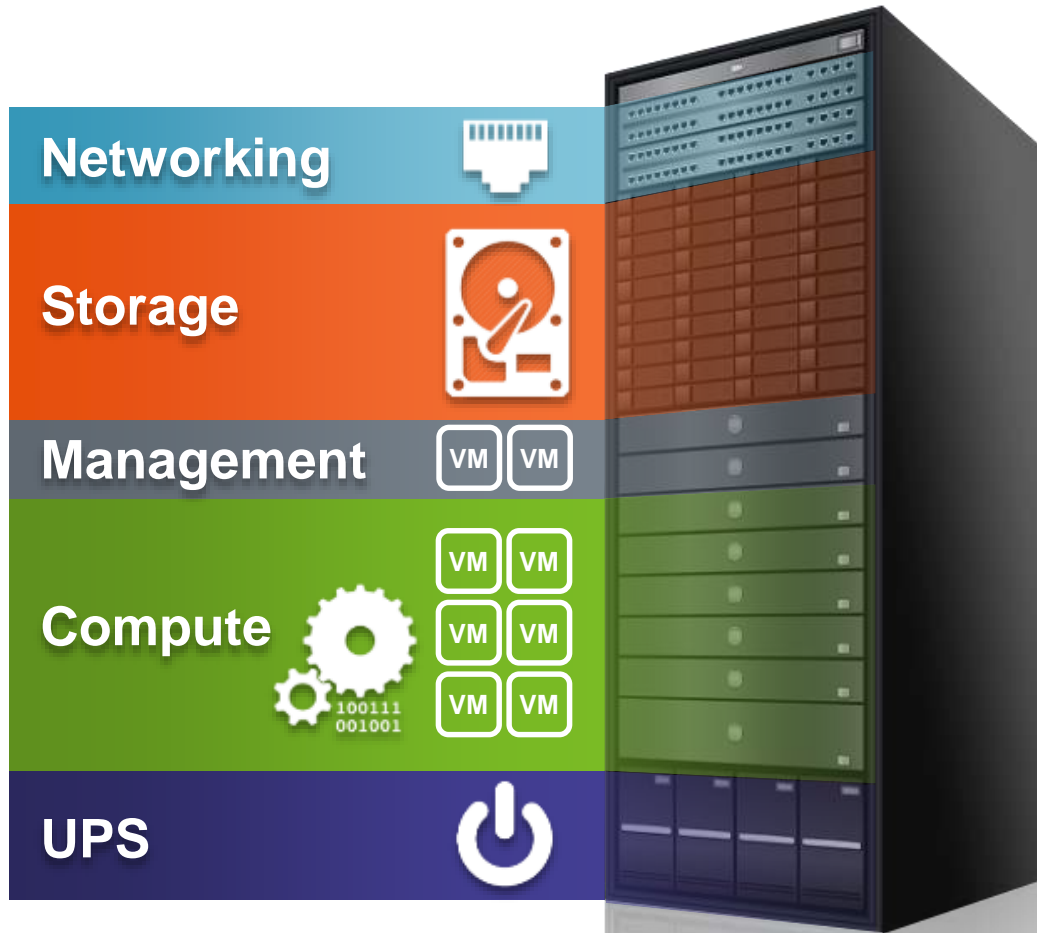


**Compute:** Increased Security & CPU Offload

**Storage:** Software Defined Storage, Cloud Storage

**Networking:** Evolution via Virtualization

# Enterprise Data Center



NSS ToR SDN switch



Software Defined Storage (SDS)



Intelligent Storage Acceleration



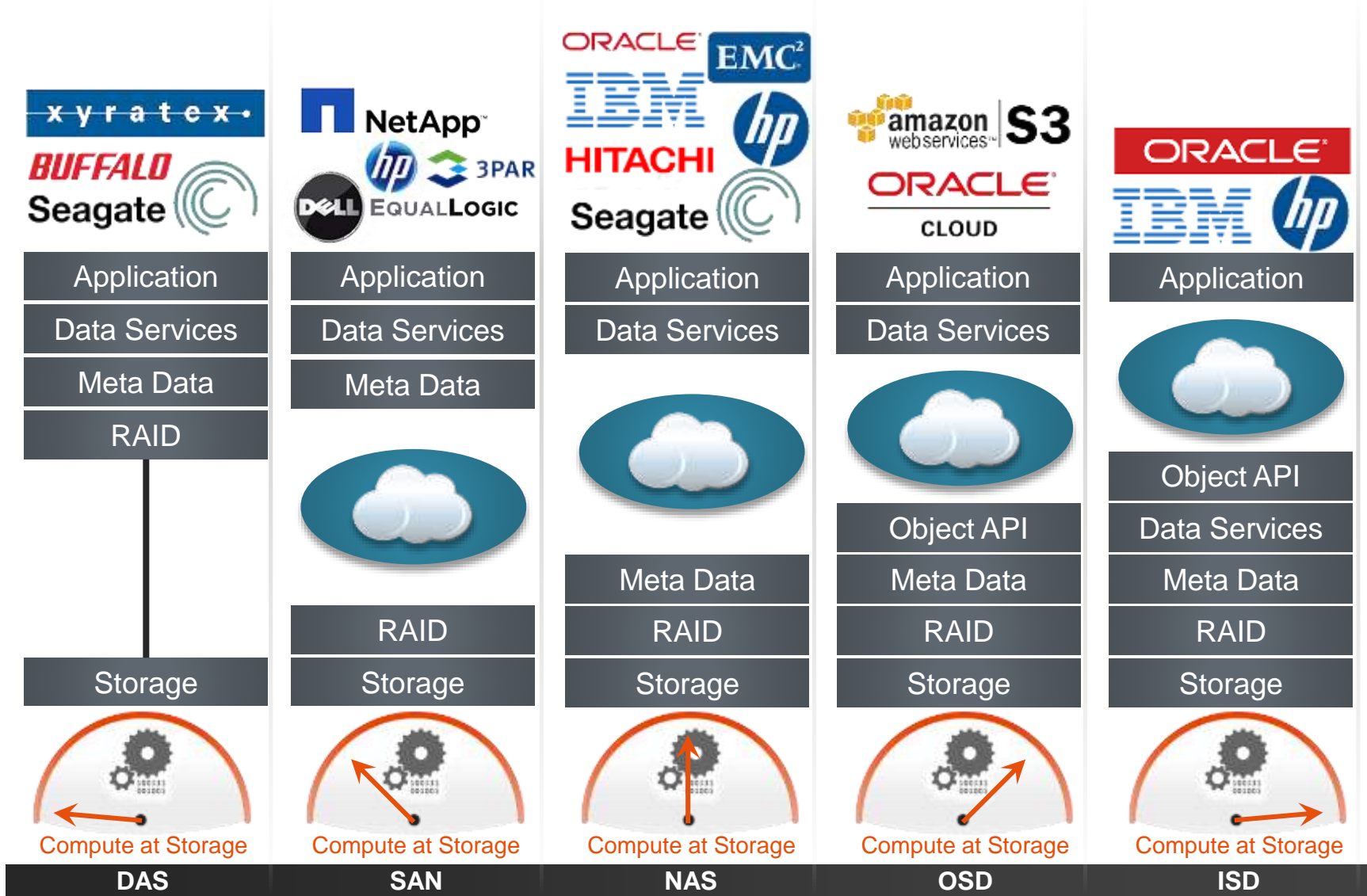
Multi-Functional Appliance Platform



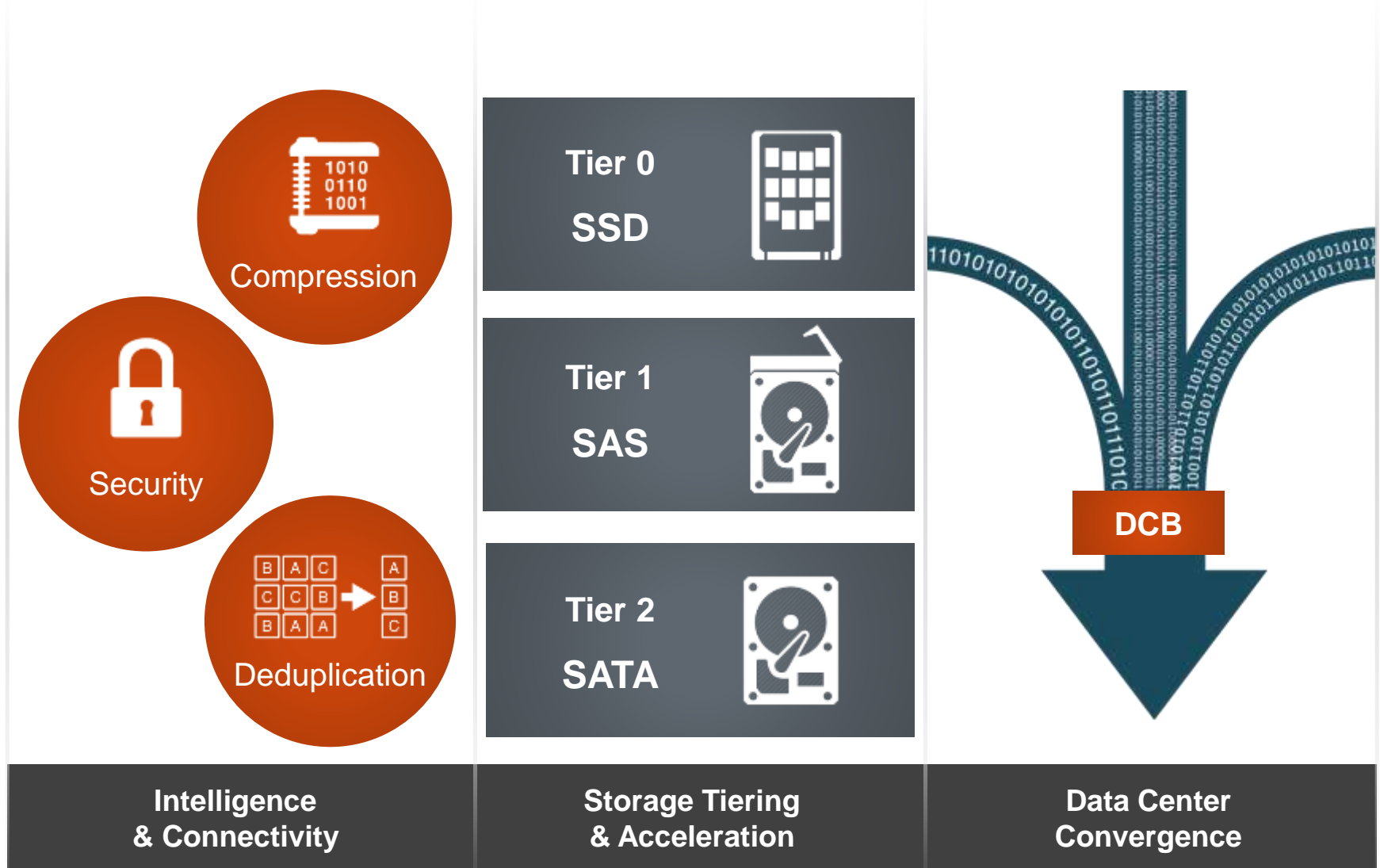
Smart iNICs



# Evolution of Data Storage



# Data Center – Customer Concerns

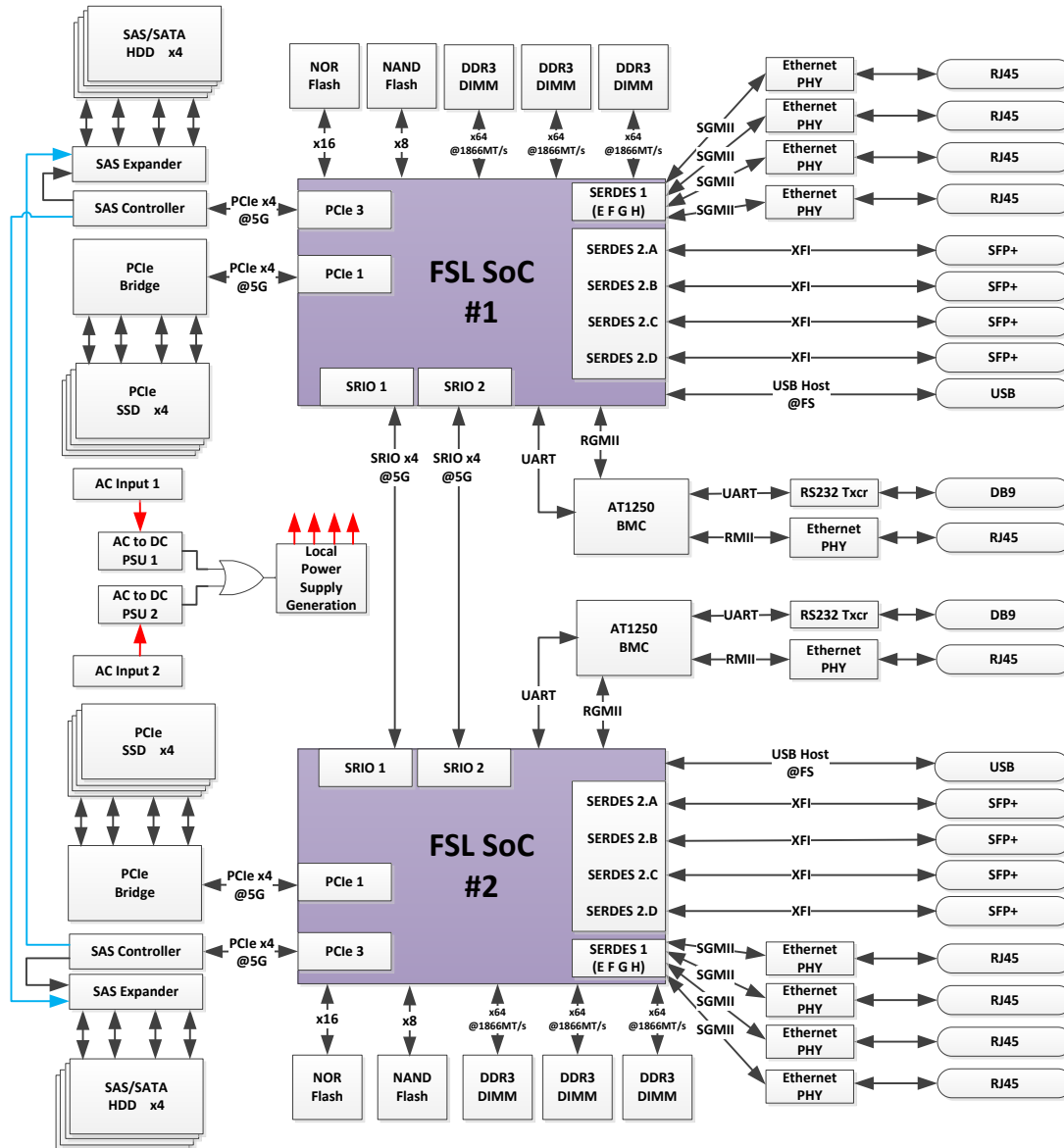




# Freescale Storage Solutions



# Freescale 80G SDS Solution with HDD/SSD



## 2U dual-SoC SDS/Storage Appliance

- ONIE
- NVMe for PCIe connectivity to SSD's
- Dual redundant SAS cards for Highly Available HDD's in the chassis
- RAID5/6
- RDMA over Serial RapidIO
- DPAA across Serial RapidIO for load balancing
- Ceph deployment on the Storage Appliance
- iSCSI offload
- FCoE offload
- Data-at-rest encryption using T4 and/or integrated C29x
- Dual redundant power supplies

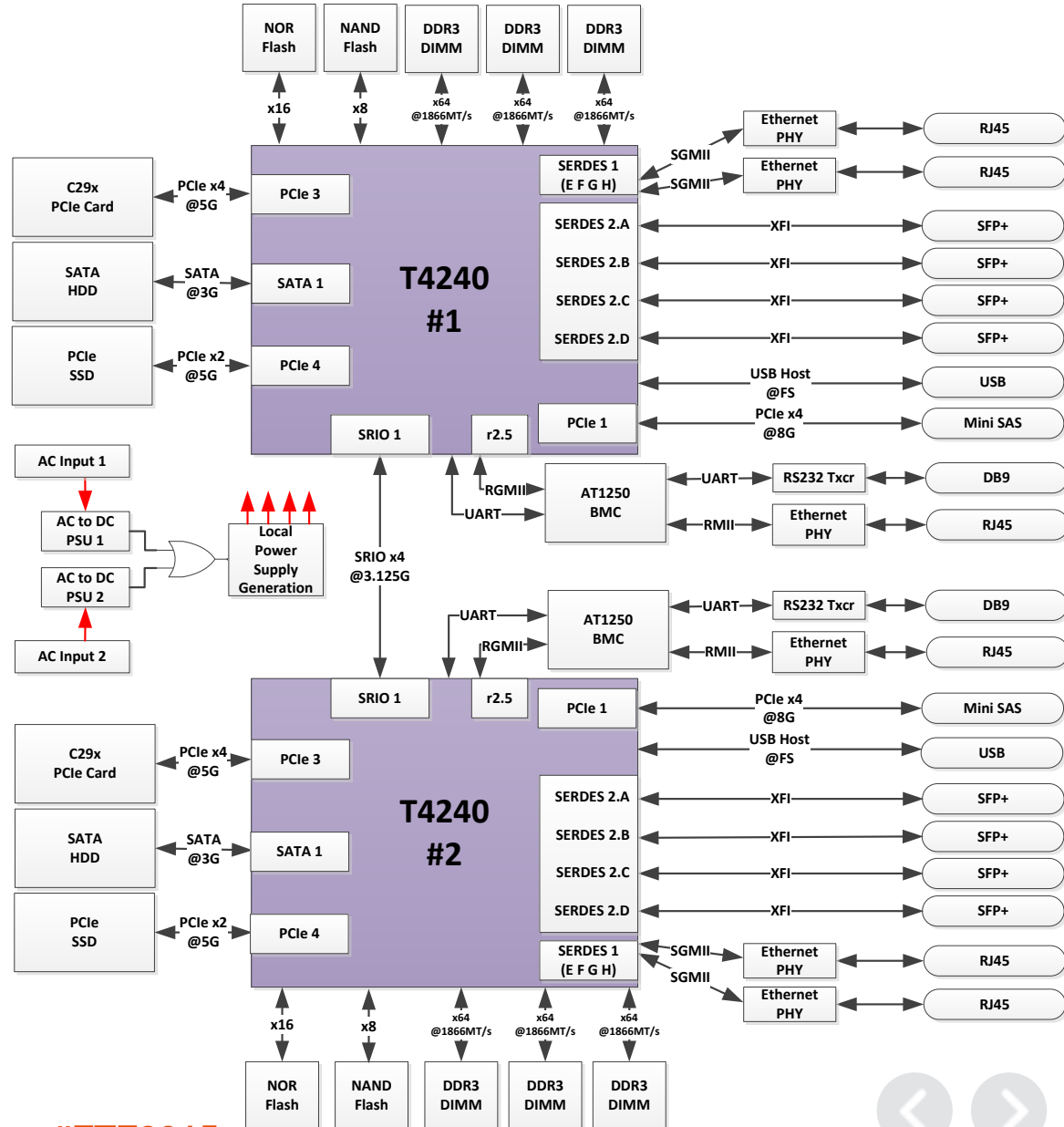




# Freescale Dual QorIQ T4240 Server

## Key Features

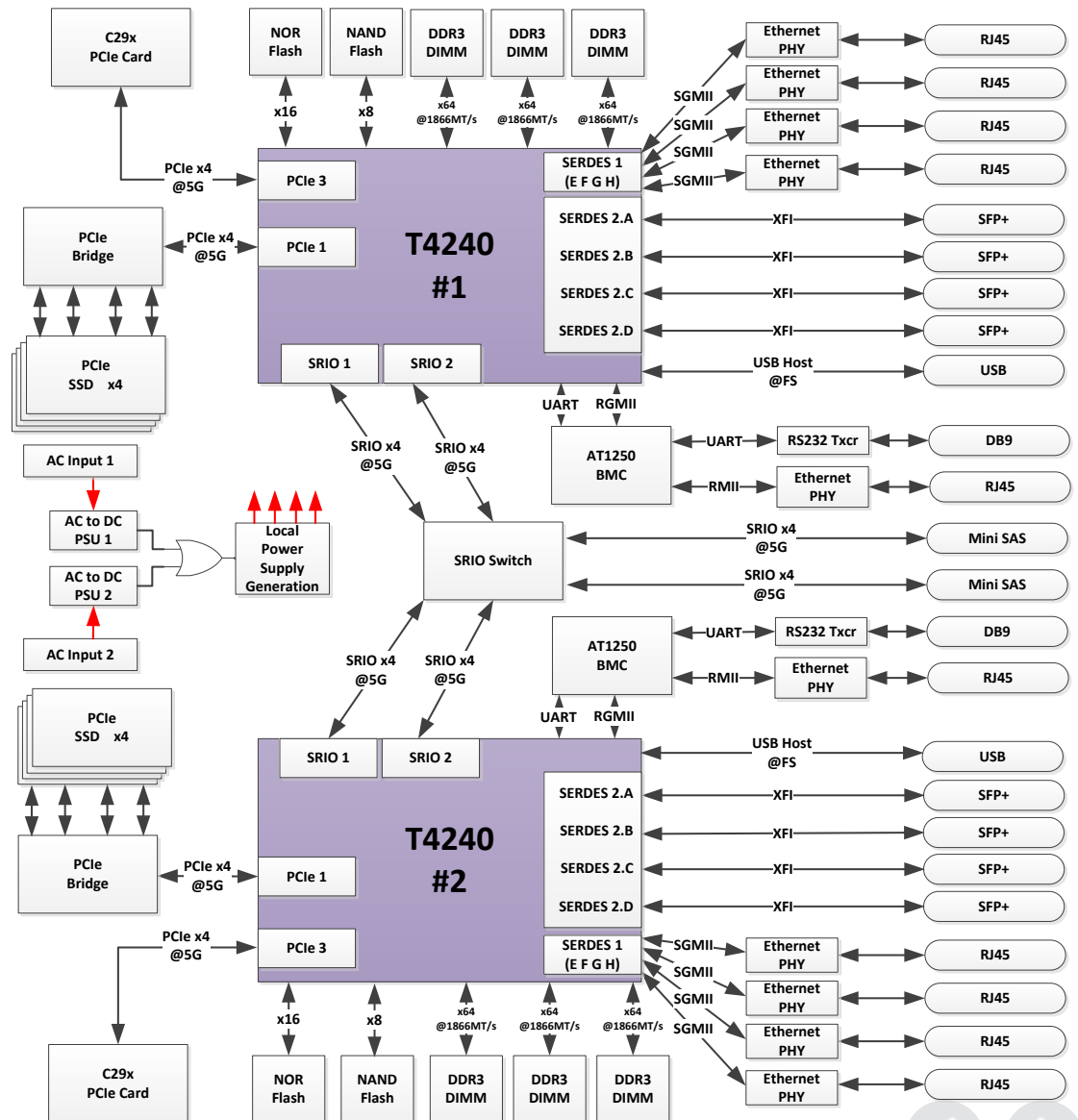
- 1U 24 core, Dual-T4 solution
- Integrated dual C29x for crypto acceleration
- 8 x 10G + 4 x 1G integrated networking with DCB
- NVMe support for PCIe connectivity to SSD's
- RAID implementation on AltiVec
- Load balancing and distribution across T4's
- Redundant power supply
- Hardware acceleration for compression, deduplication & encryption



# Freescale Dual QorIQ T4240 Scale-out Server

## Key Features

- 2U 24 core, scale-out solution
- Scale-out via Serial RapidIO switch for clustered solutions like Hadoop
- Integrated dual C29x for crypto acceleration
- 8 x 10G + 8 x 1G integrated networking with DCB
- NVMe support for PCIe connectivity to SSD's
- RAID implementation on AltiVec
- Load balancing and distribution across T4's
- Redundant power supply
- Hardware acceleration for compression, deduplication & encryption



# Enterprise NAS – 2U Redundancy-in-a-Box

- Dual T4240's with 24 cores arranged in three core clusters sharing a 2MB L2 cache
- Up to 1.8GHz at 1.0V with 64-bit ISA
  - Scale out option to 4 processors
- 2U Storage Appliance option
- Integrated storage(iSCSI, FCoE) protocols
- 75W thermal power (without drives)
- Memory
  - 6x (DDR3 +ECC), 24GB total (expandable)
  - 1x 128Mbytes Flash
- I/O
  - 8 x 10GE (SFP+) + 4 x1G
  - IPMI via UART and 1GE
  - 2x SATA2
  - PCIe expansion slots optionally supporting
    - Dual C293 SSL coprocessor
    - Dual PCIe switch with 4x SSD
    - Dual SAS controllers with up to 24 drives
    - Multiple PCIe SSD (up to 24)

## FRONT VIEW



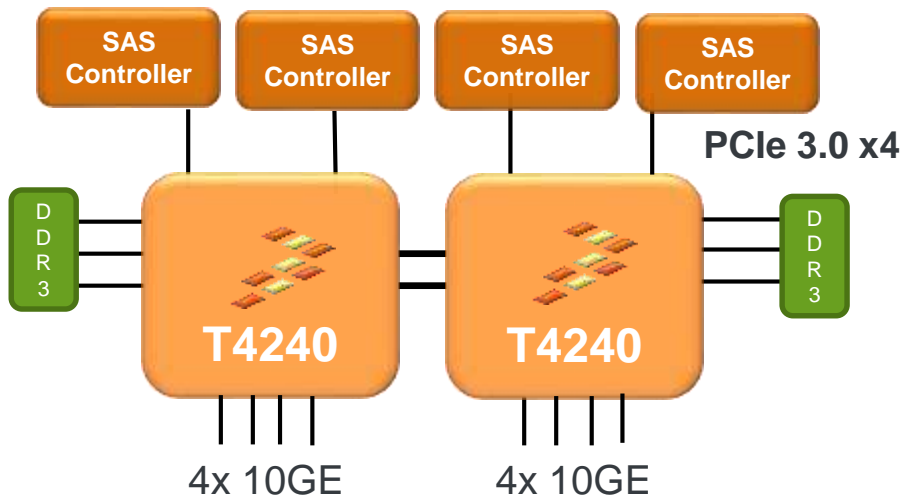
## REAR VIEW



- Dual Chip standard low latency interconnect (Serial RapidIO)
- Dual redundant power supplies
- Dual BMC (Base Board Management Control)
- Load balancing across T4240's



# Comparing a Redundant Storage Solution with 80G

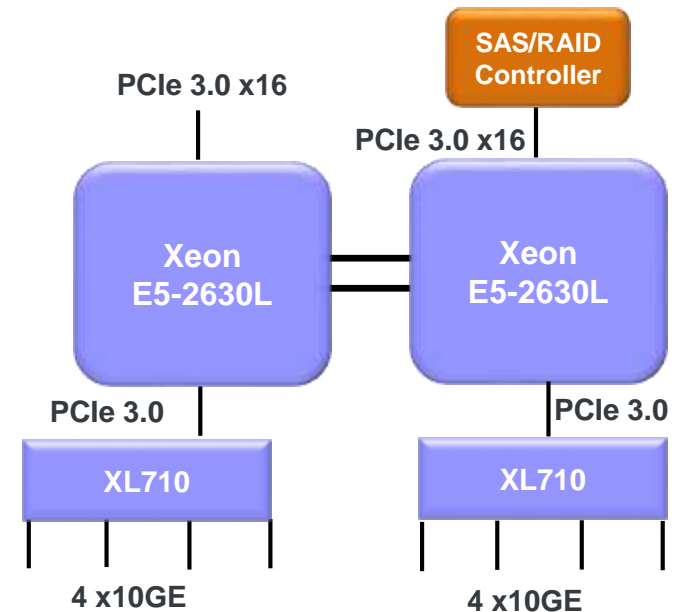


**Freescale: ~44% less price,  
~29% less power,**

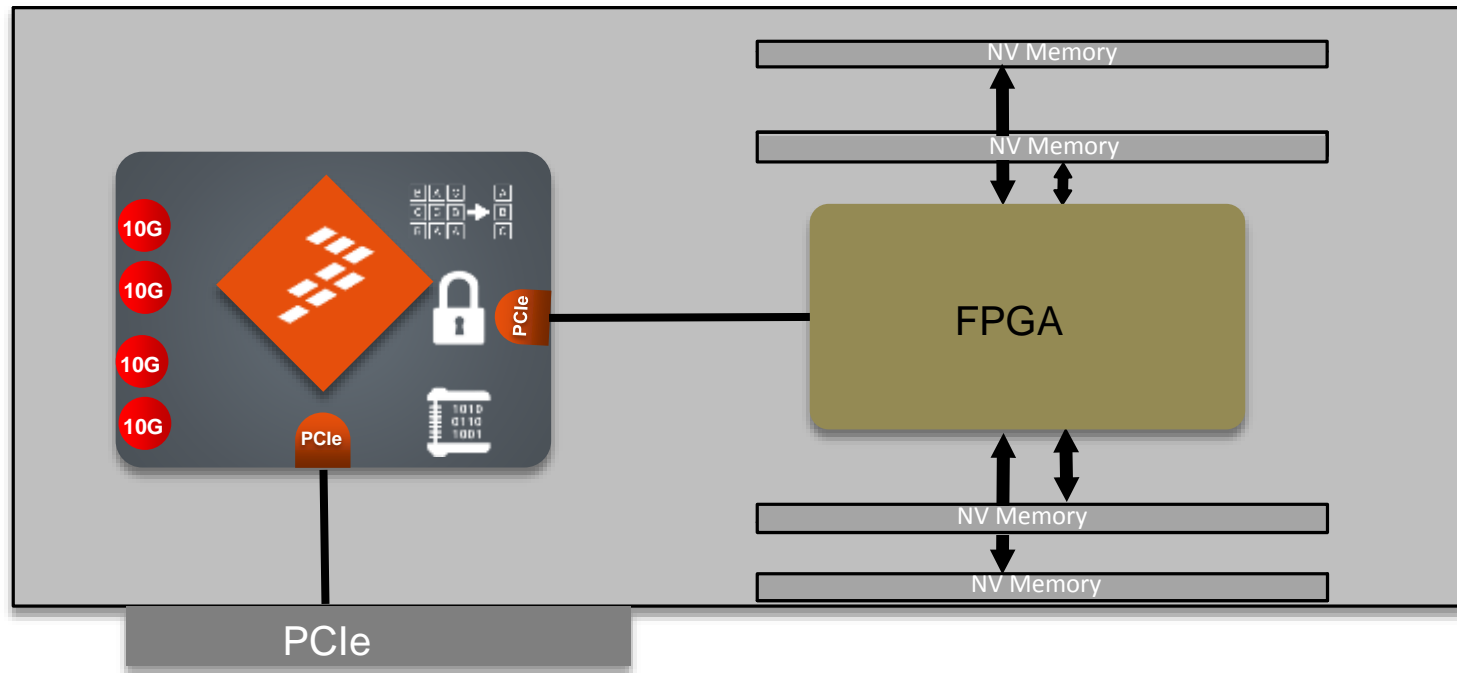
**2 x cores  
AND**

**2 chips instead of 4 for the  
same functionality!**

Parameter	Freescale	Intel
Cores (per processor)	12 @ 1.8 GHz	6 @ 2 GHz
Data Accelerators: PME DCE SEC	10Gb/s 20Gb/s 40Gb/s	Software
Chips needed	2	4
List Price	T4240: 2 x \$570 Total: <b>\$1140</b>	E5-2630L 2 x \$662 XL710 2 x \$160.5 Total <b>\$1645</b>
Power (thermal or TDP)	T4240: 2 x 52W Total: <b>104W</b>	E2630L: 2 x 60 120W XL710 : 2 x 7 14W Total <b>134W</b>

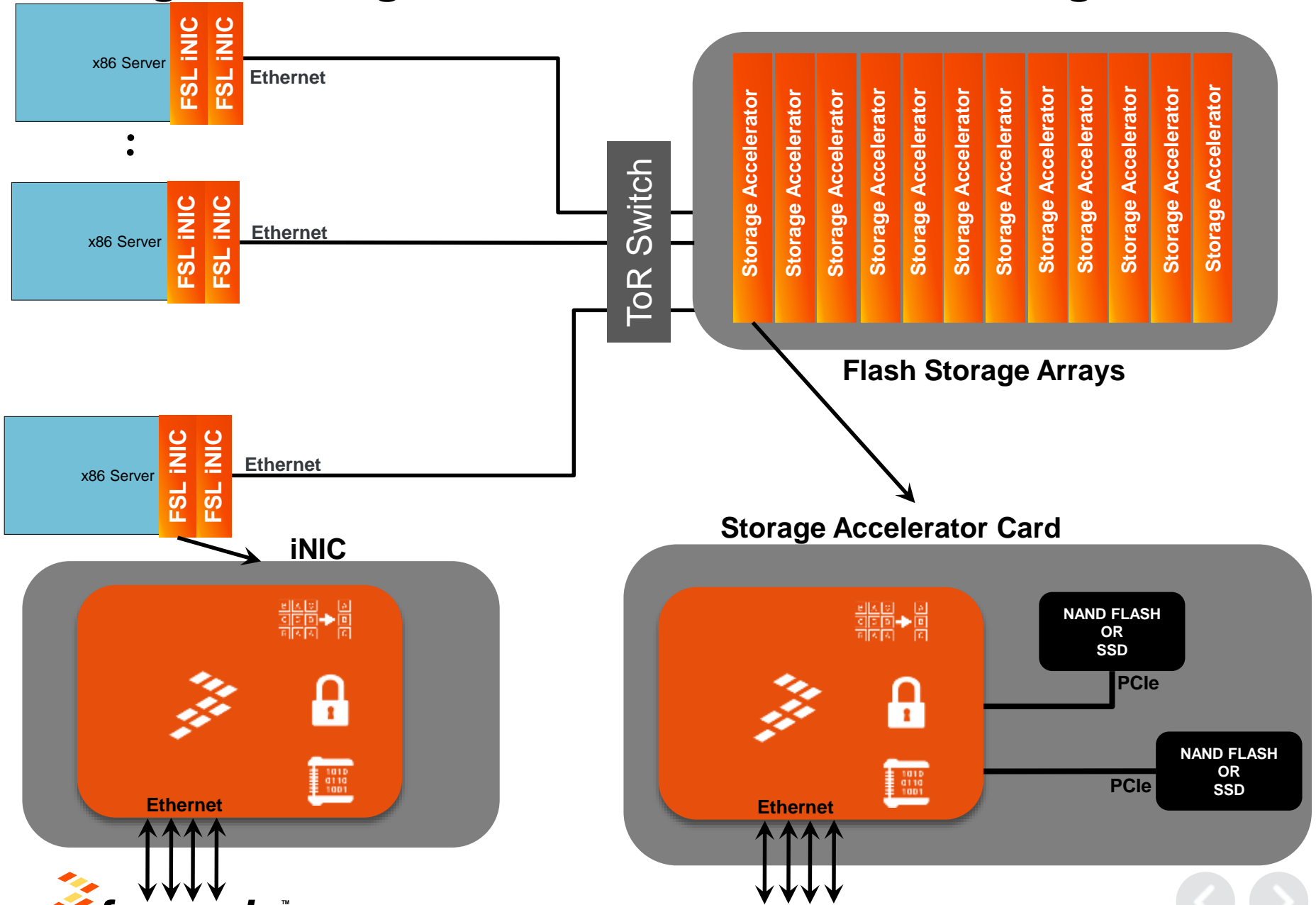


# Freescale Intelligent Storage Accelerator



- RDMA over Converged Ethernet (RoCE)
- NVMe Controller on Freescale SoC
- NVMe over RoCE, hardware offloading
- Built-in compression, Crypto hardware accelerators
- ARM TrustZone® support
- 4 x 10GbE with DCB support
- Two PCI-e Gen 3.0: x4 or x8 controllers
- 16-bit Local Bus interface (FPGA backdoor communication)
- One USB port for control and debugging
- One SPI (Flash)
- SoC will be mounted as a storage device

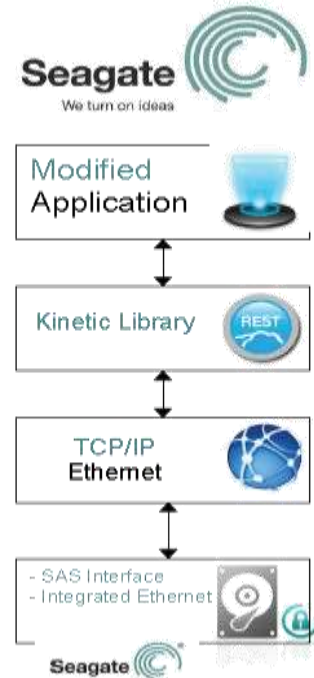
# Intelligent Storage Acceleration on Host & Storage



# Ethernet Drive

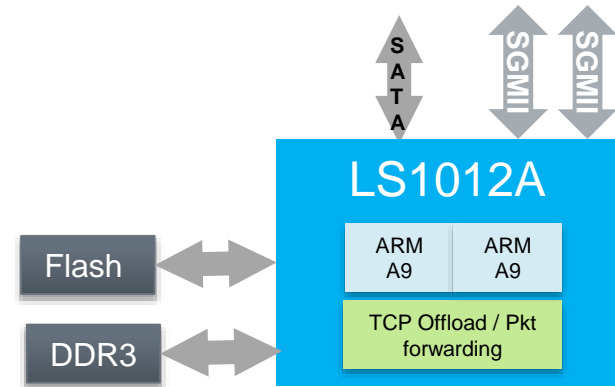
Legacy IT infrastructure is a slow-moving ship. When you have something as new and different as Seagate Kinetic, you're not going to see a massive rush to it all at once. But does it have the potential to change storage architecture in the future? Absolutely."

*Dave Reinsel, Vice President at Research Firm IDC*



## Freescale Differentiation

- PPF's Samba optimizations
- SATA optimizations & performance
- AoE offload using PPF
- Integrated dual Gigabit ethernet & SATA interfaces





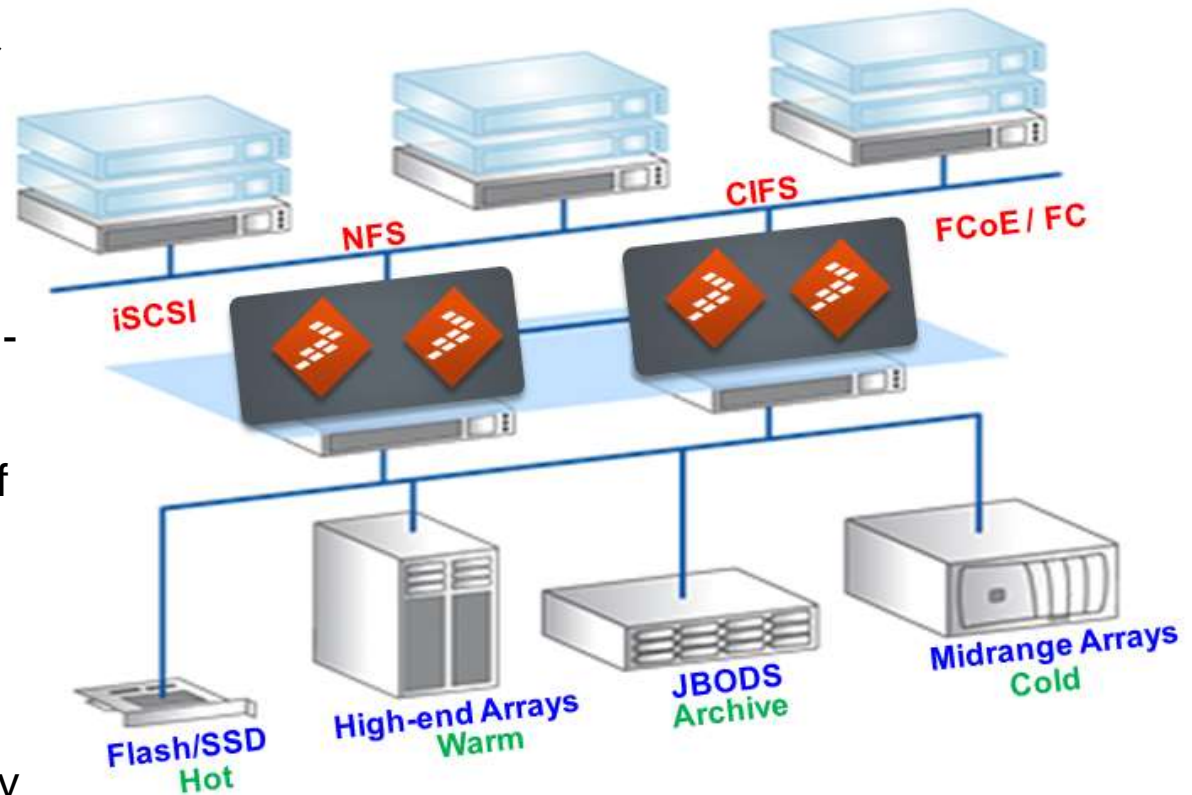
# Freescale Storage Use Cases



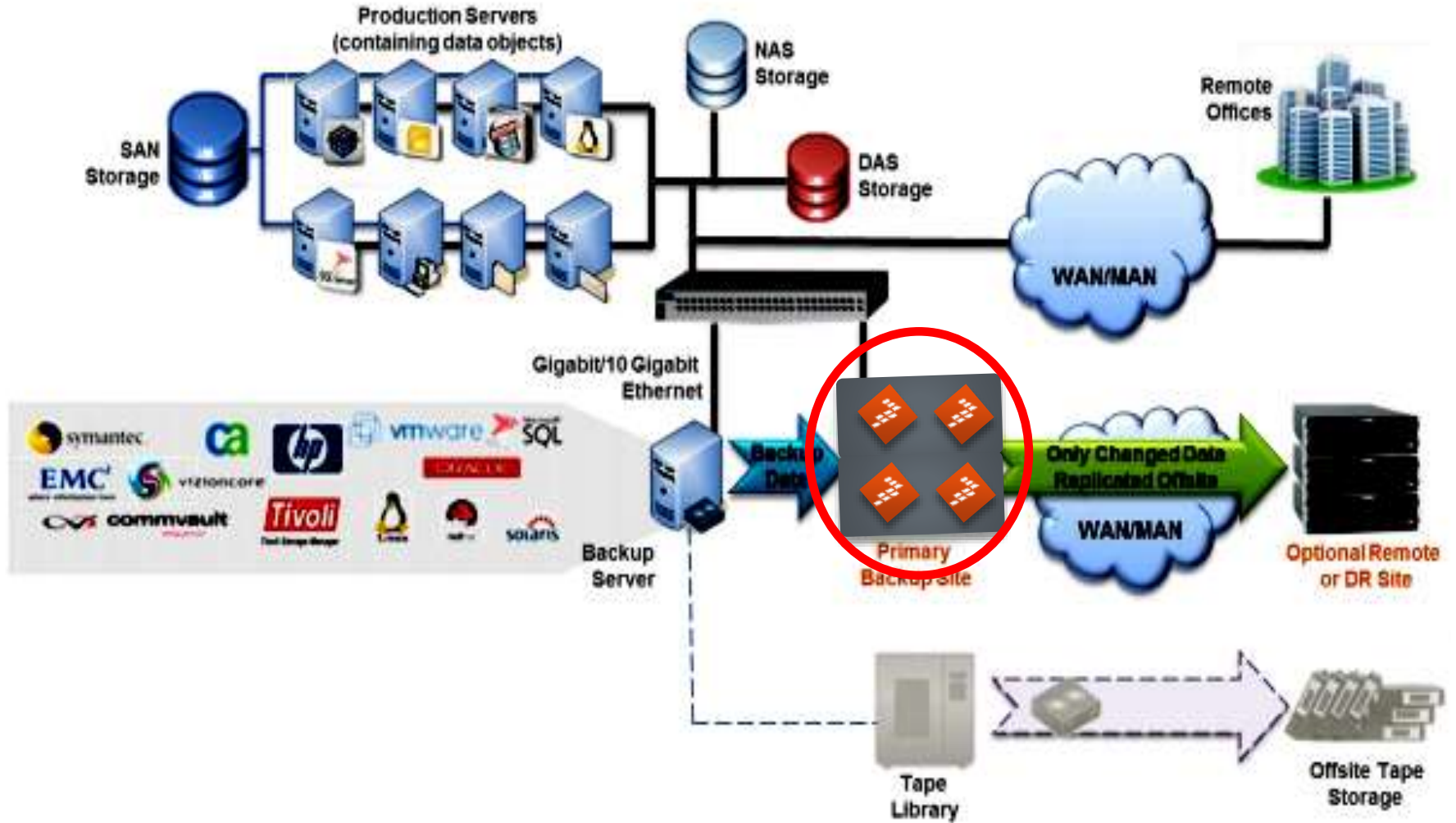
# Software Defined Storage

## Features

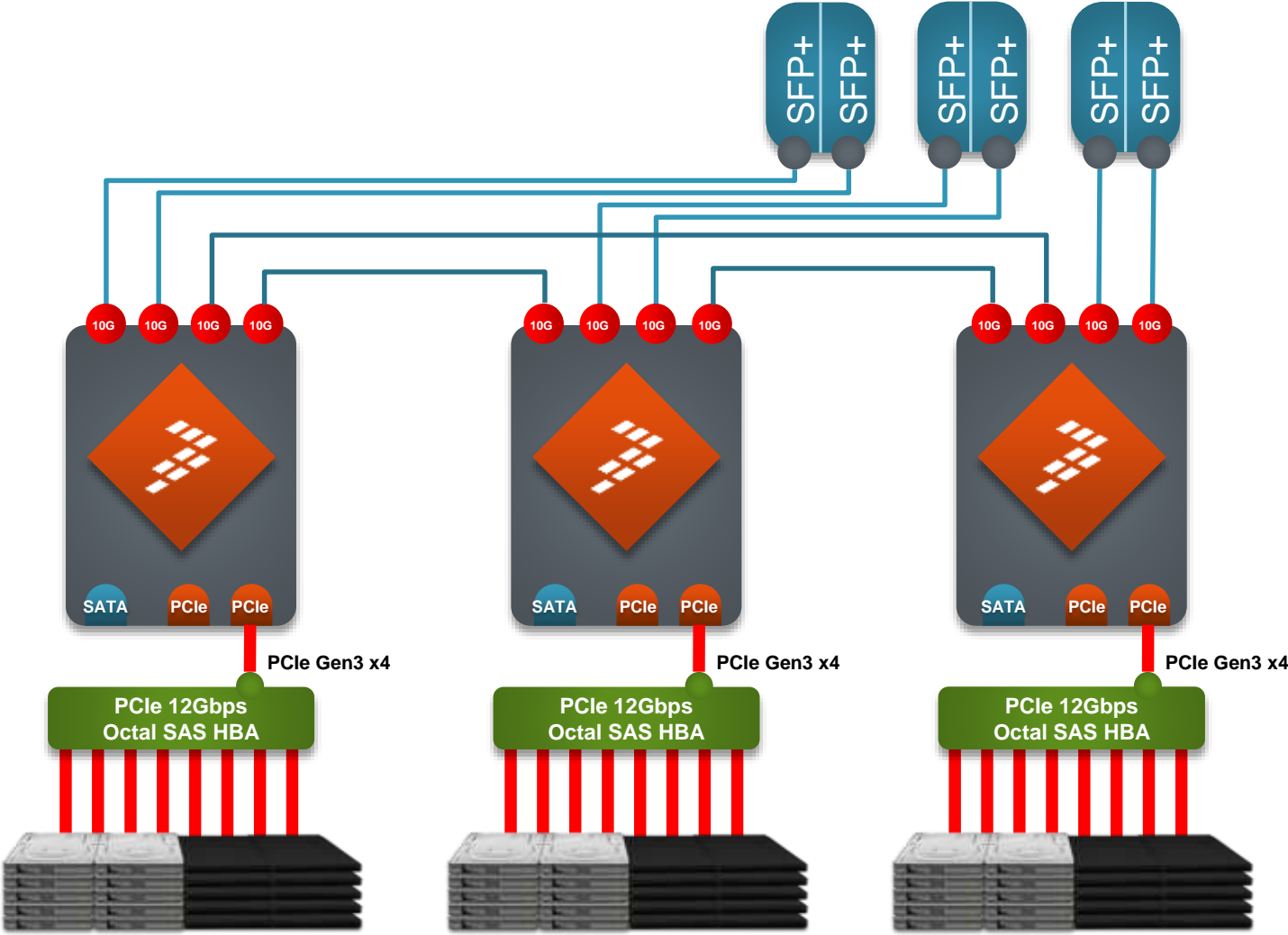
- Flexible & resilient storage
- Enable access to file, block & object storage systems
- Provisioning & management
- Guaranteed QoS, complete redundancy-in-a-box
- Data Services: Dedupe, compression, encryption, thin-provisioning
- Abstraction layer to storage systems with varying levels of capacity & performance
- Unlimited scalability
- Guaranteed SLA's across heterogeneous storage systems
- Significant reduction in energy costs



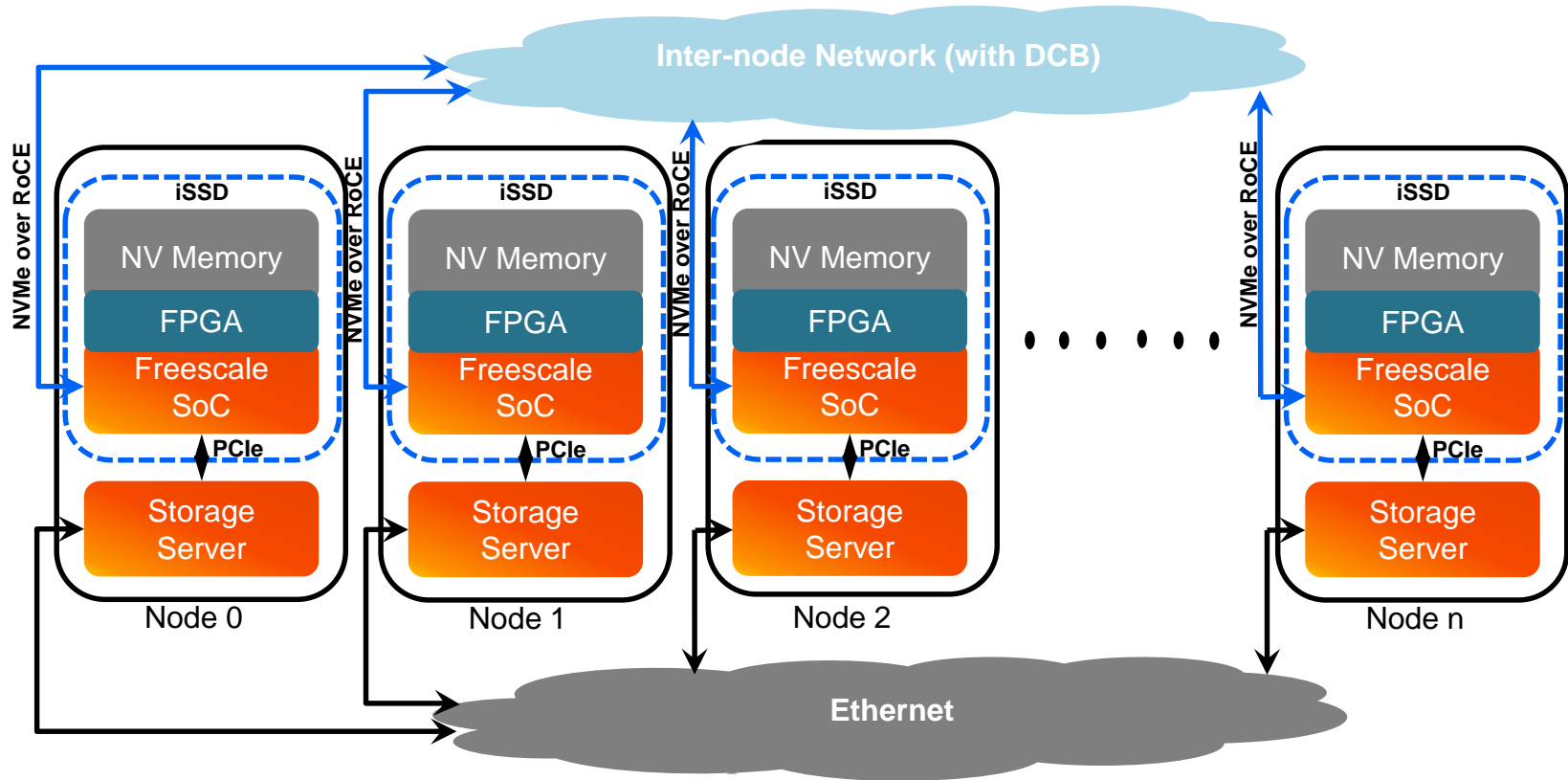
# Backup Appliance: DeDupe, Compress, Encrypt



# Cloud Storage Systems



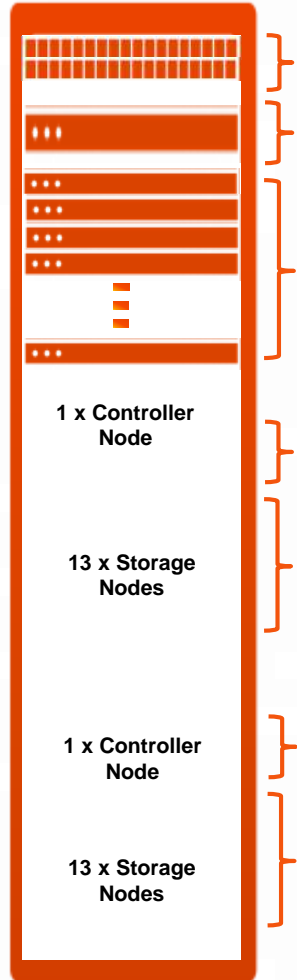
# Scalable Systems Architecture with Freescale Storage Accelerator



- PCIe NVMe SSD's deliver low latency hardware acceleration
- Low latency, high bandwidth inter controller network is achieved with NVMe over RoCE

# Ethernet Drive - Cloud Storage Solution

## Data Center Rack



2 x 48 Port Ethernet Switches

1 x Controller Node

13 x Storage Nodes

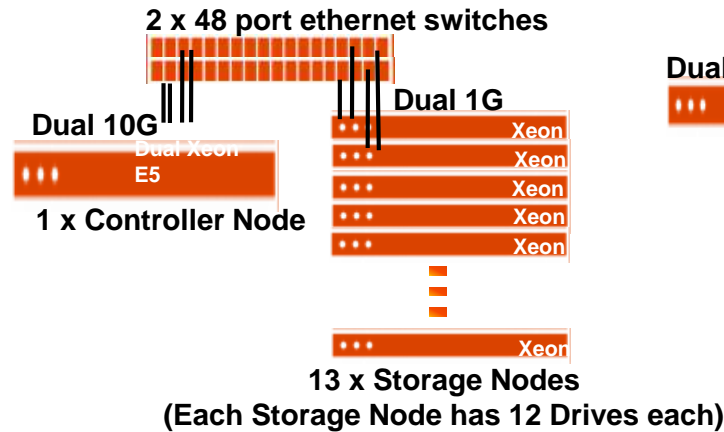
1 x Controller Node

13 x Storage Nodes

1 x Controller Node

13 x Storage Nodes

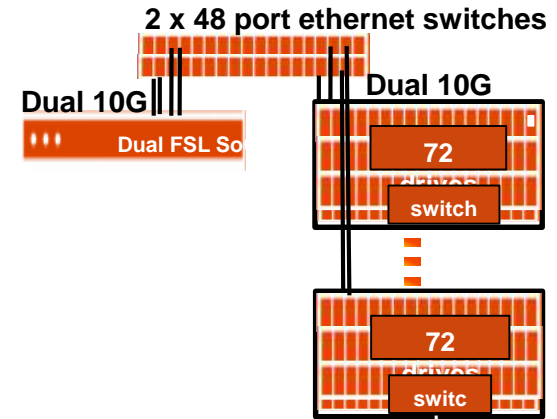
### Current Solution



**TOTAL RACK COST**

$$93210 + 11916 = \$105,126$$

### Ethernet Drive Solution



4 x Storage Nodes

(Each Storage Node has 72 drives)

Plus a 72 x 1G + 2 x 10G embedded switch

**TOTAL RACK COST**

$$87048 + 1140 = \$88,188$$

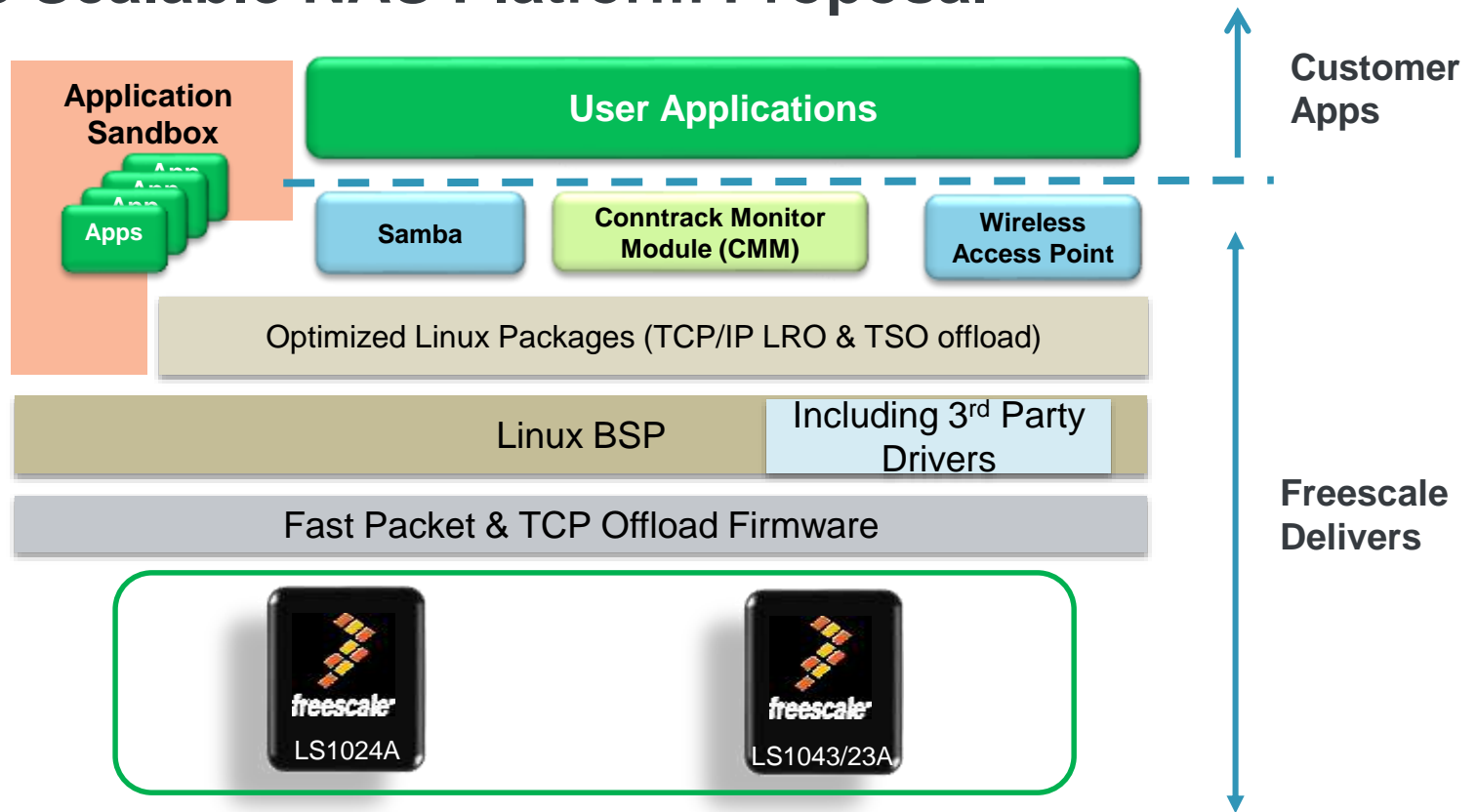
**Ethernet Drive architecture is 53% more space efficient and 16% more cost efficient not including total system power & thermal design**

# Freescale Storage Solutions cNAS, Ethernet Drives, Storage Accelerators & iNIC's





# Freescalable Scalable NAS Platform Proposal



**Freescalable:**  
broadest  
ARM SoC  
solutions  
supplier

- Up to 6,000DMIPS
- Hardware packet engine
- ~140MB/s read/write
- Video transcoding via co-processor
- ~3W typical power

- Up to 14,000DMIPS
- Hardware packet engine
- >200MB/s read/write
- Video transcoding via co-processor
- ~6-8W typical power

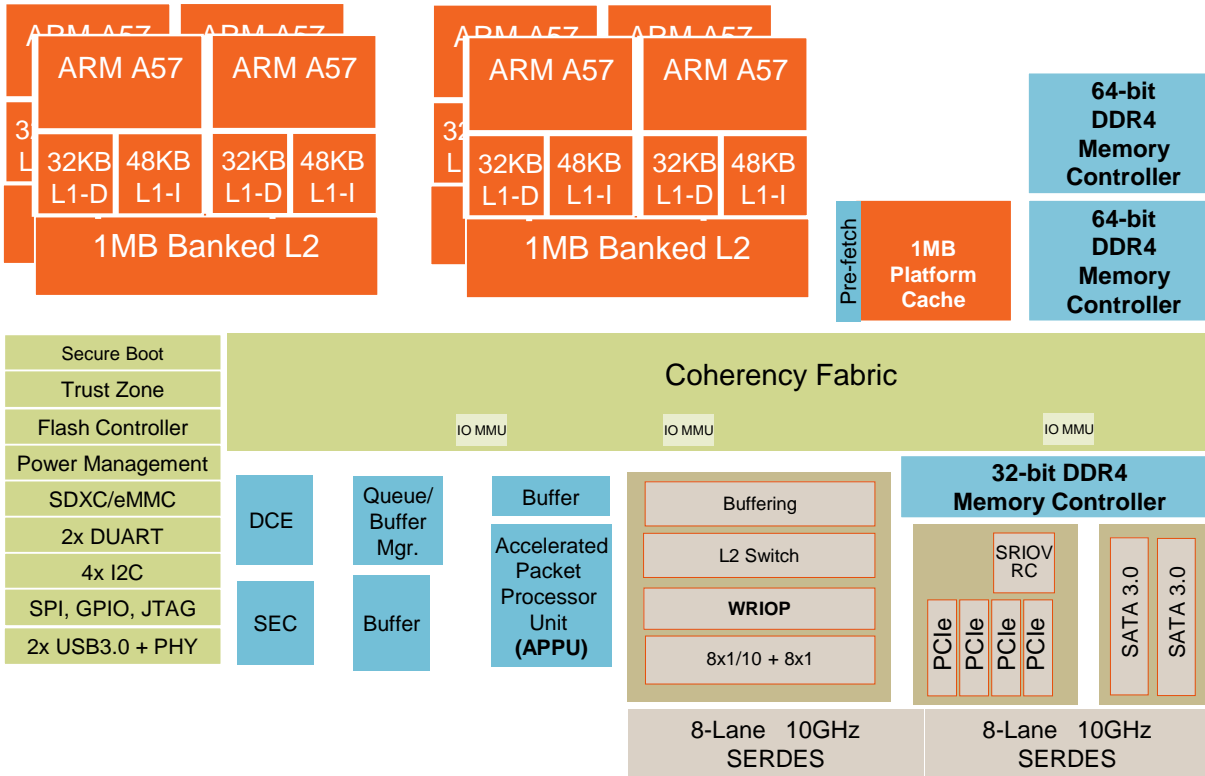
\* Chips not shown to scale

# High-end Storage Solutions Overview and Definition



# Layerscape LS2085A Family

## Industry's Leading 64b ARM Solution



### General Purpose Processing Layer

- 8 x ARM A57 CPUs, 64b, 2.0GHz
  - 1MB L2 cache / cluster
- HW L1 & L2 Prefetch Engines
- Neon SIMD in all CPUs
- 1MB L3 platform cache w/ECC
- 2x64b DDR4 up to 2.4GT/s
- 

### Accelerated Packet Processing

- 40Gbps Packet Processing Engine
- 10Gbps SEC- crypto acceleration
- 10Gbps Pattern Match/RegEx
- 10Gbps Data Compression Engine

### Express Packet IO Layer

- Supports 1x8, 4x4, 4x2, 4x1 PCIe Gen3 controllers
  - SR-IOV support, Root Complex
- 2 x SATA 3.0, 2 x USB 3.0 with PH

### Network IO

- Wire Rate IO Processor:
  - 8x1/10GbE + 8x1G
  - XAUI/XFI/KR and SGMII
  - MACSec on up to 4x 1/10GbE

### Other Parametrics

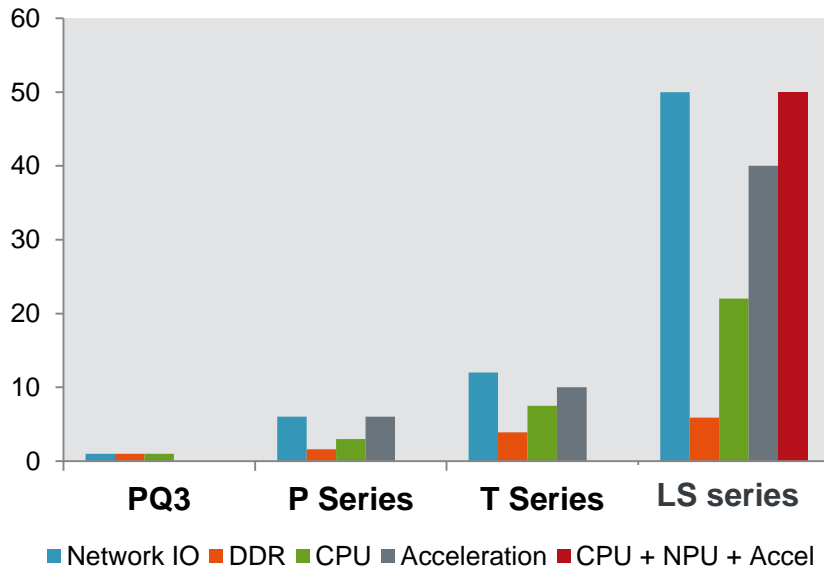
- 37.5x37.5 Flipchip
- 1mm Pitch
- 1292pins
- 45W Typical Power, 60w MAX Power (est)

### Datapath Acceleration

- **SEC**- crypto acceleration
- **DCE** - Data Compression Engine
- **PME** – Pattern Matching Engine



# LS series: A New Architecture for a New Network



Many-core processor approach is not sustainable due to power, software complexity and integration costs



Need to provide right mix of high performance and programmability

## MUST HAVE:

### Advance Packet Processing

- Tightly coupled accelerators called as C functions
- Hardware preloaded task state, headers, stack frame
- Customer programmable
- Run-to-completion model using standard C (C99)



**3-4x Performance**  
over general purpose cores  
in a lower power envelope



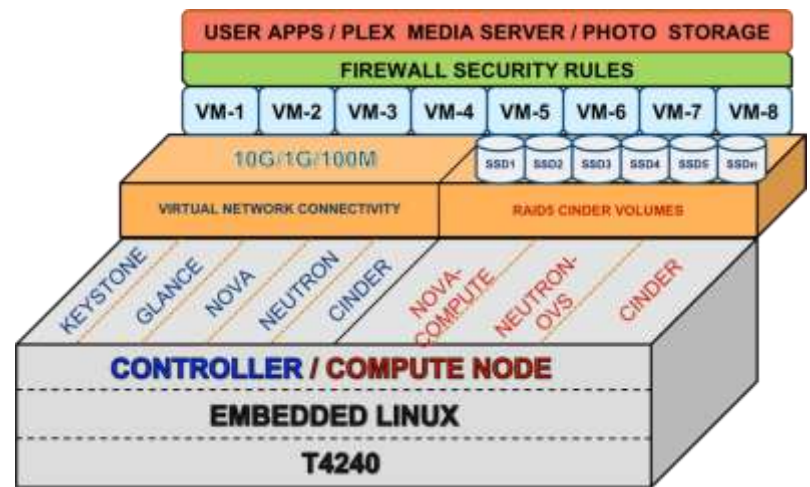
# SYNAPSE Design

Data Center Cloud Solutions with Freescale



# SYNAPSE design

Design Partner for the Fortune 500



Core for core, Synapse Platform has **3x the perf** of an x86 device (CoreMark/watt)

Core for core, Synapse Platform dissipates approx  $\frac{1}{2}$  **the power** of an x86 device

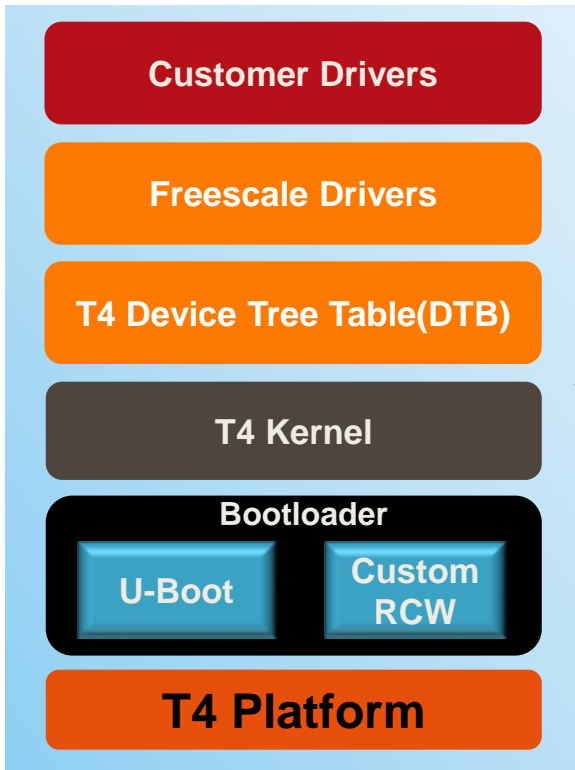
Hardware:

- T4240RDB + Openstack™
- + SAS/SATA HBA
- + SAS/SATA HDD/SSD
- + chassis/ power

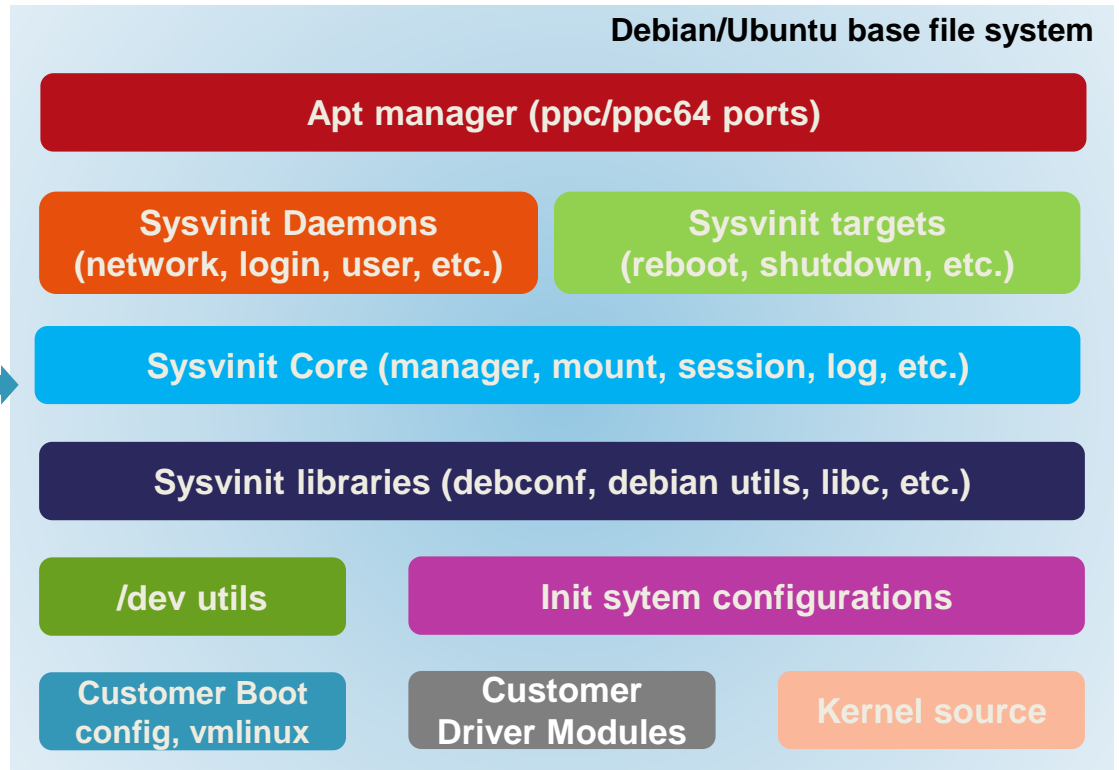


# Synapse Debian/Ubuntu for Freescale T4

## Freescale SDK with Customer Drivers



## Synapse Debian/Ubuntu File System with Customer drivers integrated







# Servergy Inc.

CYPHRE - SaaS Cloud Platform with Freescale





# Storage Platforms

- The QorIQ family and the T4240 are great Storage Chips
  - All the functional pieces needed in one place
  - Great H/W and S/W support from Freescale
- Ceph distributed storage cluster
  - Best in Class capabilities, reliability, and scalability
  - Massive Community support and ongoing development
- Production Deployment in 2014
  - 3 production clusters





# Storage Platforms

- Ceph Deployment Details
  - 3 production clusters (3 to 5 monitor nodes)
    - Cluster performance increased with larger deployments
  - 216TB in 10U, Scales to PB range in a rack
  - Excellent Bandwidth with 10G bonded network
  - Excellent cost (\$/GB)
- Caveats: Ceph is a an early technology, quickly maturing
  - Regular monitoring required
  - Cluster expansion process caused unexpected downtime





# SaaS Cloud Platform

- Beyond Storage... T4240 is an entire SaaS Platform
  - Integrated other features of the QorIQ Platform
  - Freescale has all the pieces of a solid SaaS platform
- Security Engines to protect the Cloud
- Compute power to power the Apps
- Raw bandwidth to drive data to the edge
- We're now in production with our SaaS application
  - Based on and Optimized for T4240
  - The most secure EFSS solution available





# SaaS Cloud Platform

- Our production SaaS stack is designed for Five-9's
  - 2N+1 infrastructure w/ Automatic failover
  - Data backup and replication / DR
  - Each layer is independently upgradeable
- Freescale QorIQ has a dominant role in our stack
  - 2x Load Balancers w/ SSL Termination (NGINX)
  - 3+ Application Servers (LAMP/Apache)
  - Back end Storage Ceph clusters
  - Webserver (LAMP/Apache)
  - Ubuntu Package Repository (Ubuntu 14.04 LTS)



# Freescal Has World Class Support....and MORE

**Global Technical Information Center**  
Design & Support Resource

**Networking Applications Team**  
Depth of Expertise & Knowledge

**Design With Freescal, Freescal Technology Forum**  
Training



## Networking Software and Services Group

- Commercial Solutions
- Engineering Services
- Guaranteed Performance
- Service Level Agreement Support...and MORE

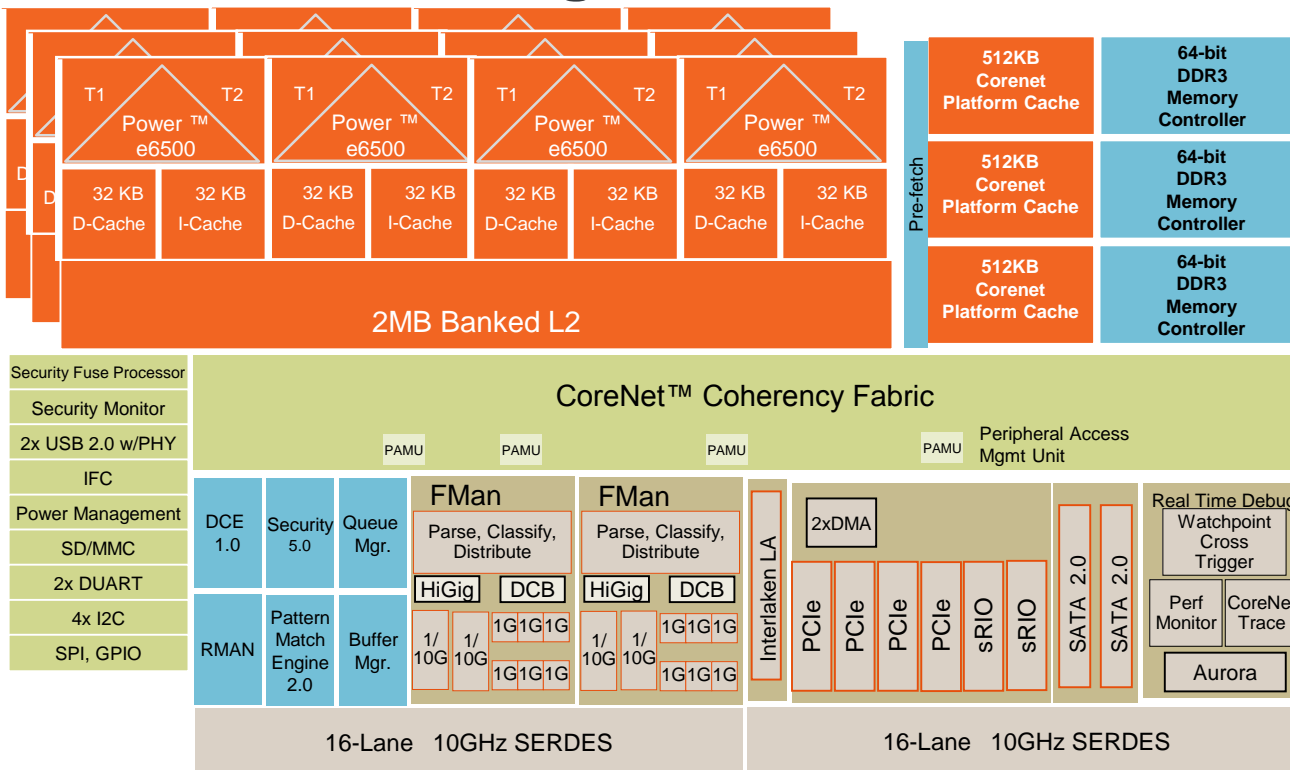
- Visit Pedestal 415 in the Technology Lab

# Backup





# T4240 Block Diagram



## Processor

- 12x e6500, 64b, up to 1.8GHz
- Dual threaded, with 128b Altivec
- Arranged as 3 clusters of 4 CPUs, with 2MB L2 per cluster; 256KB per thread

## Memory SubSystem

- 1.5MB CoreNet Platform Cache w/ECC
- 3x DDR3 Controllers up to 2.1GHz
- Each with up to 1TB addressability (40 bit physical addressing)
- HW Data Prefetching

## CoreNet Switch Fabric

### High Speed Serial IO

- 4 PCIe Controllers, with Gen3
  - SR-IOV support
- 2 sRIO Controllers
  - Type 9 and 11 messaging
  - Interworking to DPAA via Rman
- 1 Interlaken Look-Aside at up to 10GHz
- 2 SATA 2.0 3Gb/s
- 2 USB 2.0 with PHY

### Network IO

- 2 Frame Managers, each with:
  - Up to 25Gbps parse/classify/distribute
  - 2x10GE, 6x1GE
  - HiGig, Data Center Bridging Support
  - SGMII, QSGMII, XAUI, XFI

## Device

- TSMC 28HPM Process
- 1932-pin BGA package
- 42.5x42.5mm, 1.0mm pitch

## Power targets

- ~60W thermal max at 1.8GHz
- ~50W thermal max at 1.5GHz

## Datapath Acceleration

- **SEC**- crypto acceleration 40Gbps
- **PME**- Reg-ex Pattern Matcher 10Gbps
- **DCE**- Data Compression Engine 20Gbps





[www.Freescale.com](http://www.Freescale.com)