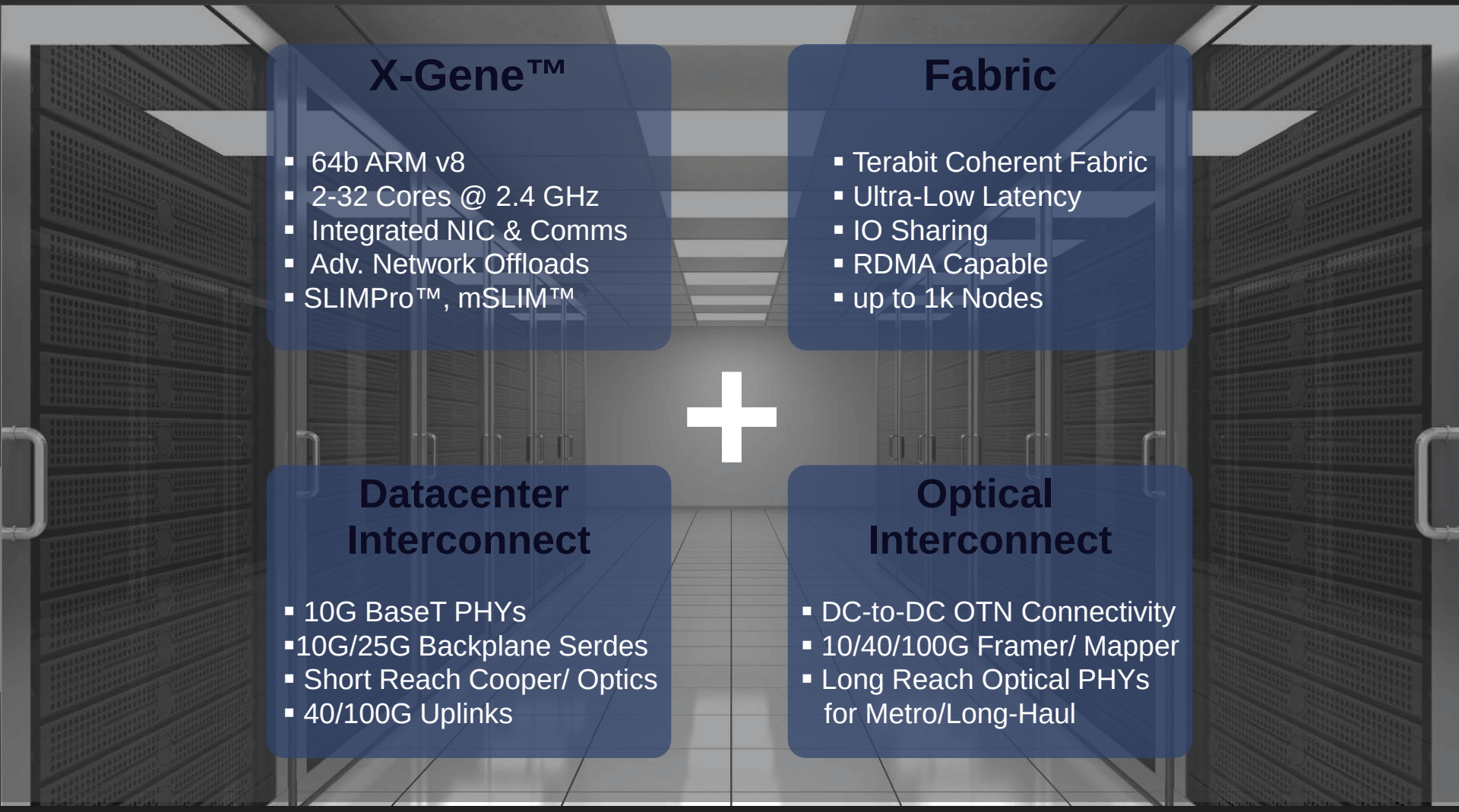
applied micro

Energy-Efficient Computing – Applications in Big Data and HPC

Parag Jain

Sr. Director of Marketing, Server Solutions

AppliedMicro's Aligned IP Portfolio Strengths



X-Gene™

- 64b ARM v8
- 2-32 Cores @ 2.4 GHz
- Integrated NIC & Comms
- Adv. Network Offloads
- SLIMPro™, mSLIM™

Fabric

- Terabit Coherent Fabric
- Ultra-Low Latency
- IO Sharing
- RDMA Capable
- up to 1k Nodes

+

Datacenter Interconnect

- 10G BaseT PHYs
- 10G/25G Backplane Serdes
- Short Reach Cooper/ Optics
- 40/100G Uplinks

Optical Interconnect

- DC-to-DC OTN Connectivity
- 10/40/100G Framer/ Mapper
- Long Reach Optical PHYs for Metro/Long-Haul

Data Explosion Fuels Datacenter Growth

Device Drivers

Content Drivers

Google

Baidu 百度

twitter

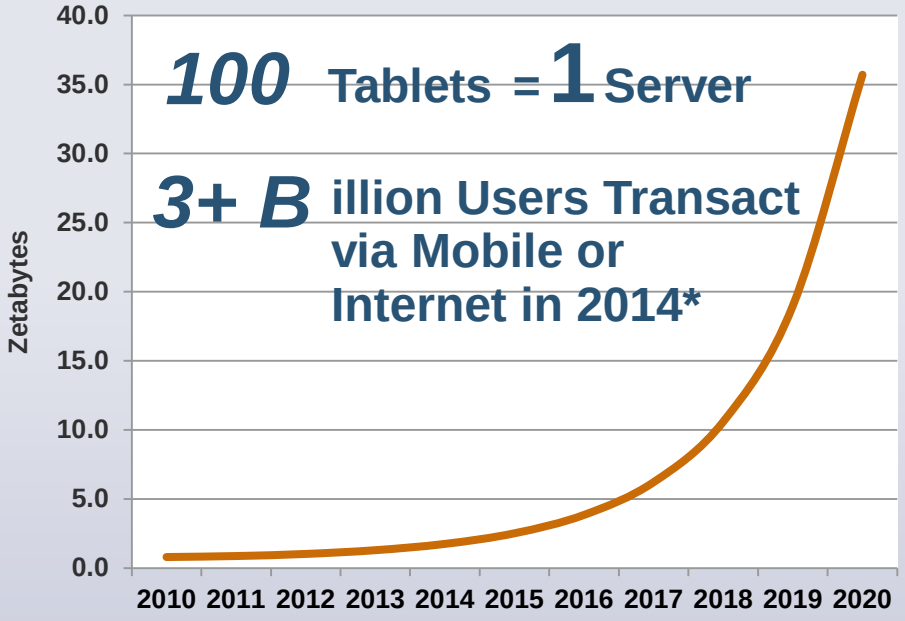
NETFLIX

PetaBytes Today. ZetaBytes Tomorrow.

500 Smartphones = **1** Server

100 Tablets = **1** Server

3+ Billion Users Transact via Mobile or Internet in 2014*



*Source: IDC, Gartner**

Datacenter = Information Super Highway

applied micro™

Datacenters Undergoing Fundamental Shifts

Corporate Data

Consumer Data

Commercial Software

Open-Source Software

ORACLE

Structured Database

Unstructured Database

hadoop

Managed Data Sets

Fragmented Data Sets

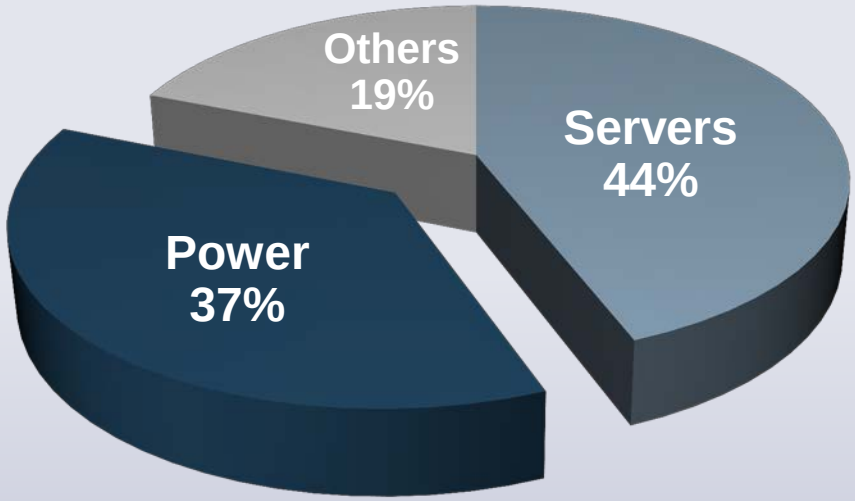
Centralized Systems

Distributed Systems

Why Low Power Servers?

Large Power Penalty → **Cost of Power = Cost of Server HW**

Being Taxed on Power Inefficiencies



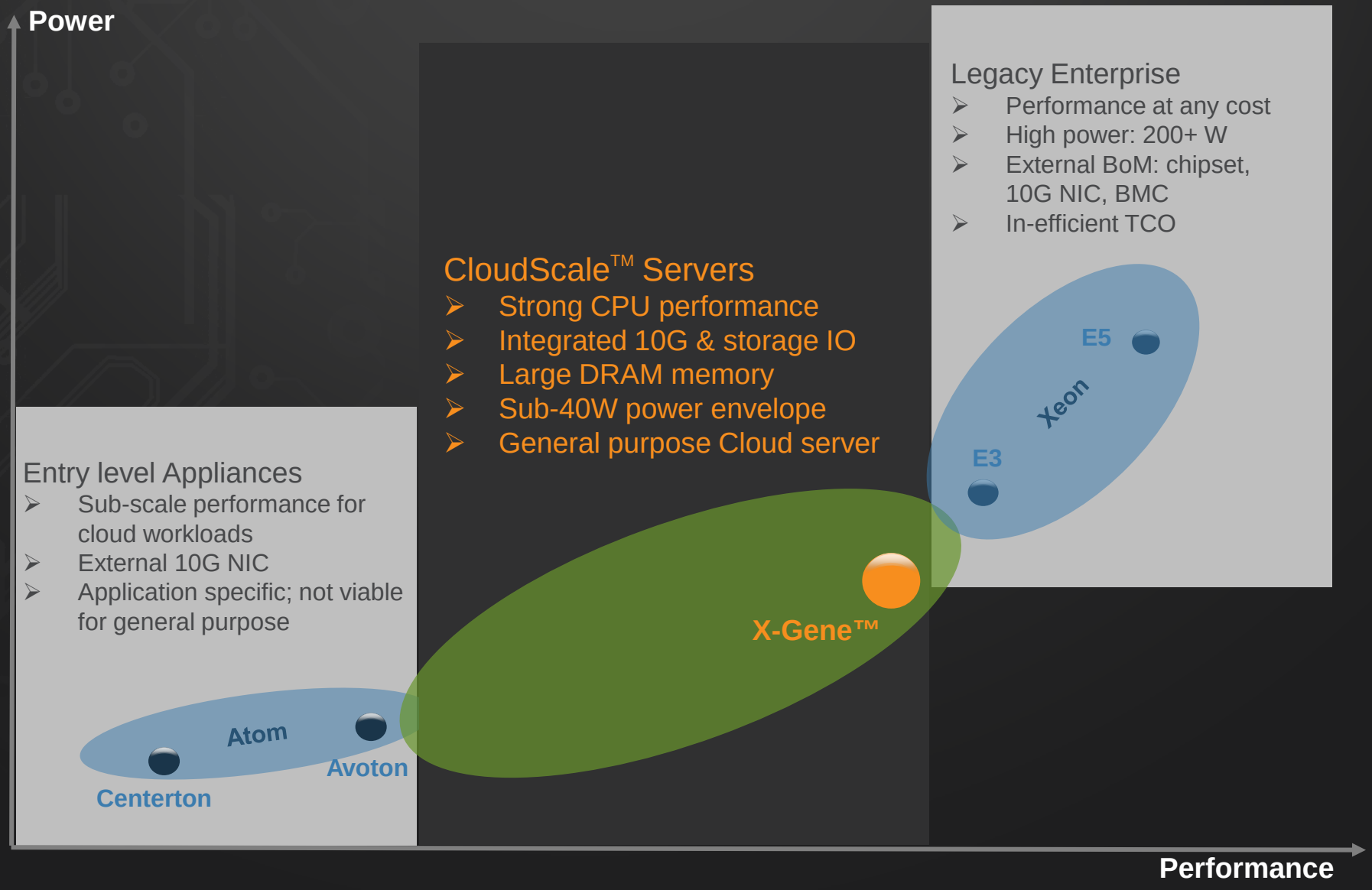
Resultant of Large Compute Intensive Cores

Source: James Hamilton, Uptime Institute, AppliedMicro

Right-sized solution generates significant TCO savings over life of Datacenter

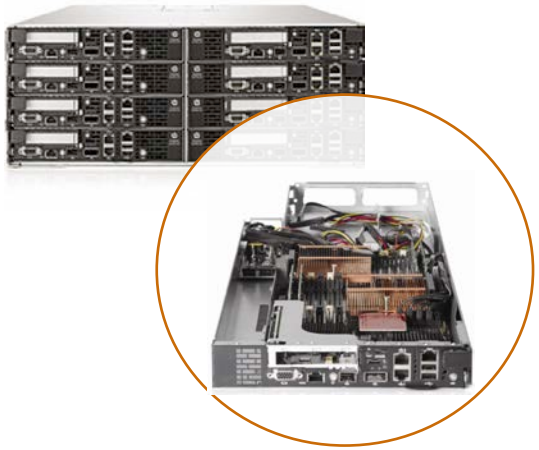
applied micro™

ARM64b Addresses Mainstream Cloud Workloads



Server Platform Evolution

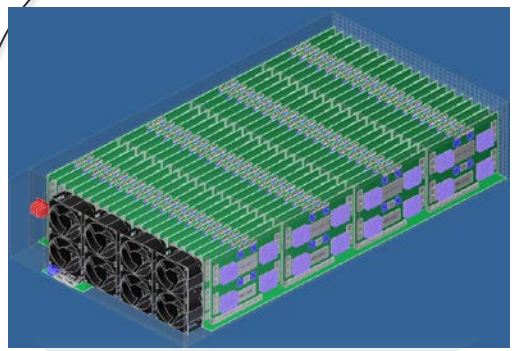
Traditional servers

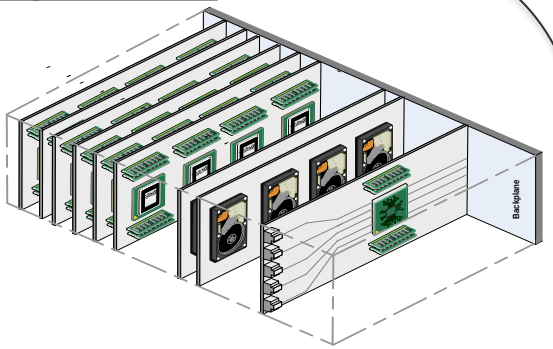


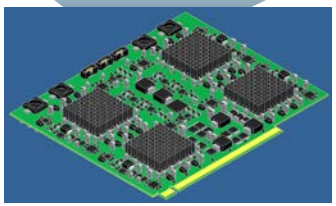
4U Chassis

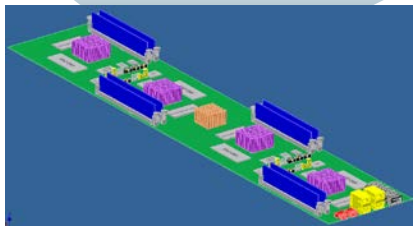
- 8 servers → 2GHz 16 cores per server
- Shared power supply & cooling
- 2.4KW power supply

High Density servers









4U Chassis

- 180 servers → 2GHz 4 cores per server

4U Chassis

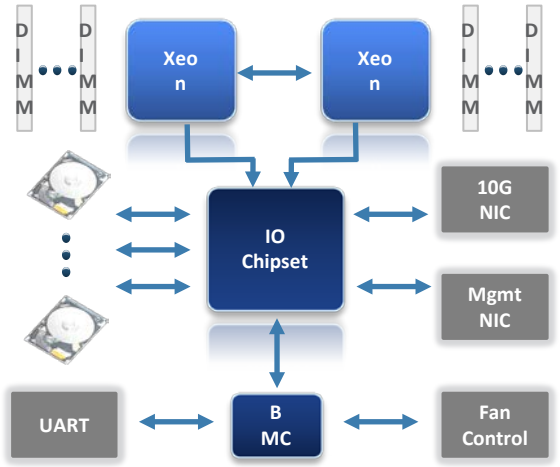
- 40 servers → 2GHz 16 cores per server

- Shared power supply and cooling
- Disaggregation of storage and networking blades
- Common management framework

4 – 6x the density of cores compared to traditional Xeon servers

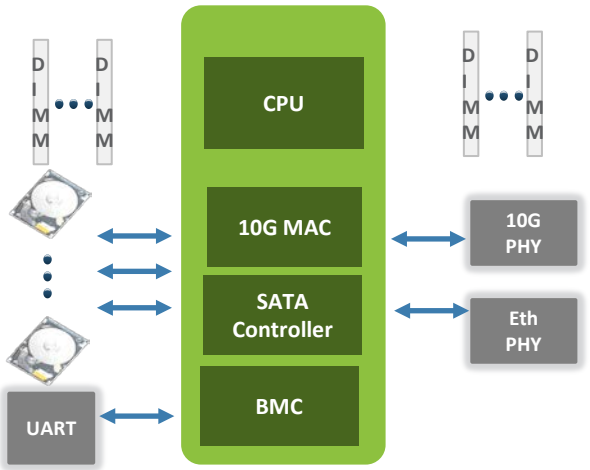
Server Technology Evolution

Traditional servers



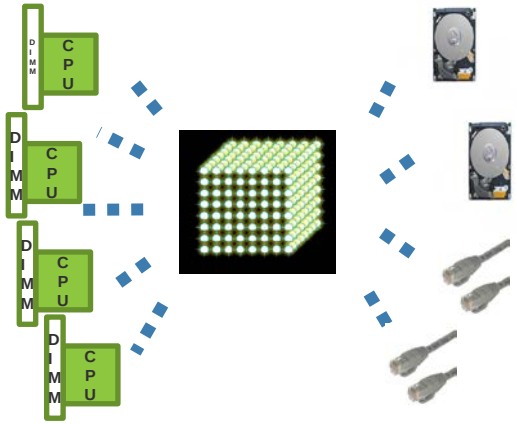
- 2 socket Xeon → 8C @95W TDP
- External chipset → PCIe, SATA connectivity
- External 10G NIC

Low Power servers



- Power efficient CPU cores → 8-16C @ 20-35W
- SoC approach → integrated SATA, PCIe …
- Integrated 10G → significantly reduce latency

High Density Compute Clusters



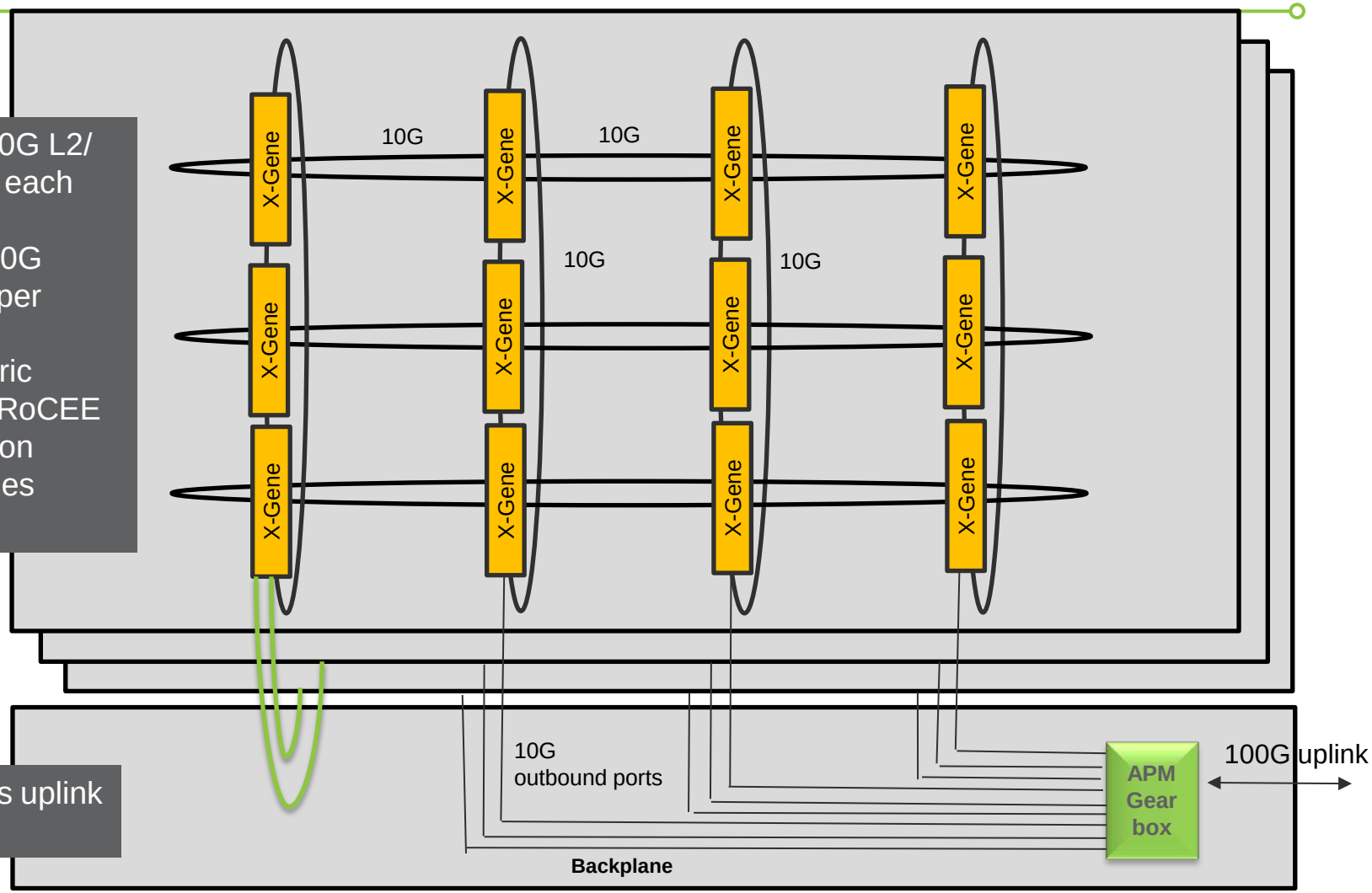
- IO sharing within the rack → networking, storage, PCIe devices
- Flexible topology → scales to 1000's of nodes
- Optimized IPC – RDMA over 10GE (RoCEE)

Optimized compute for datacenter workloads

Scale Out Fabric

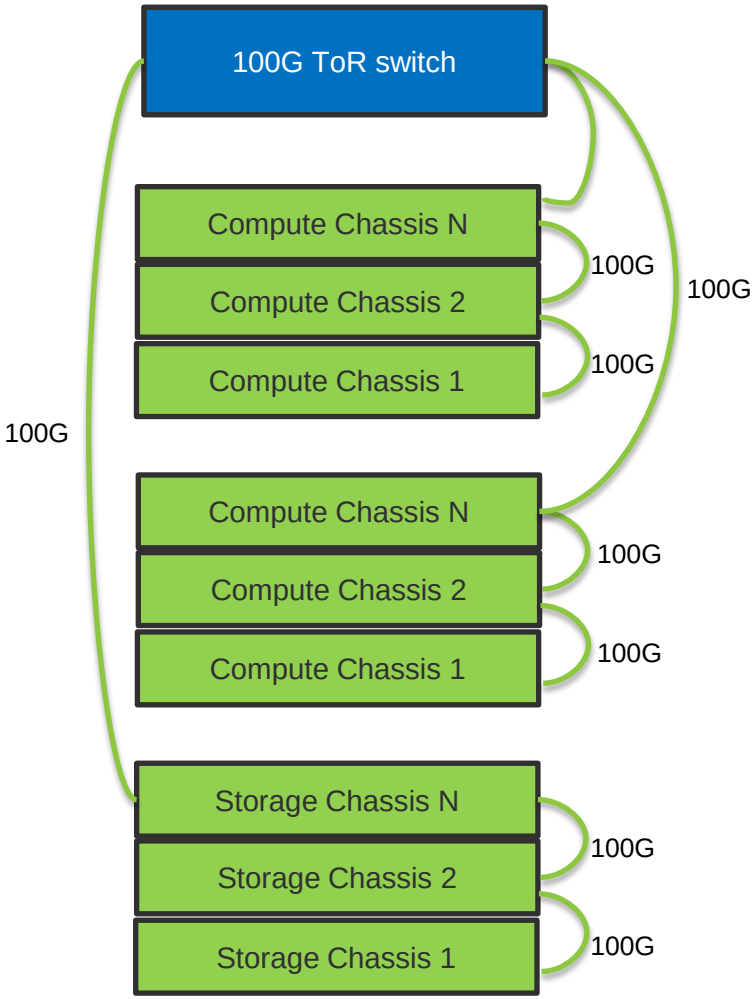
Next-Gen Server Concepts

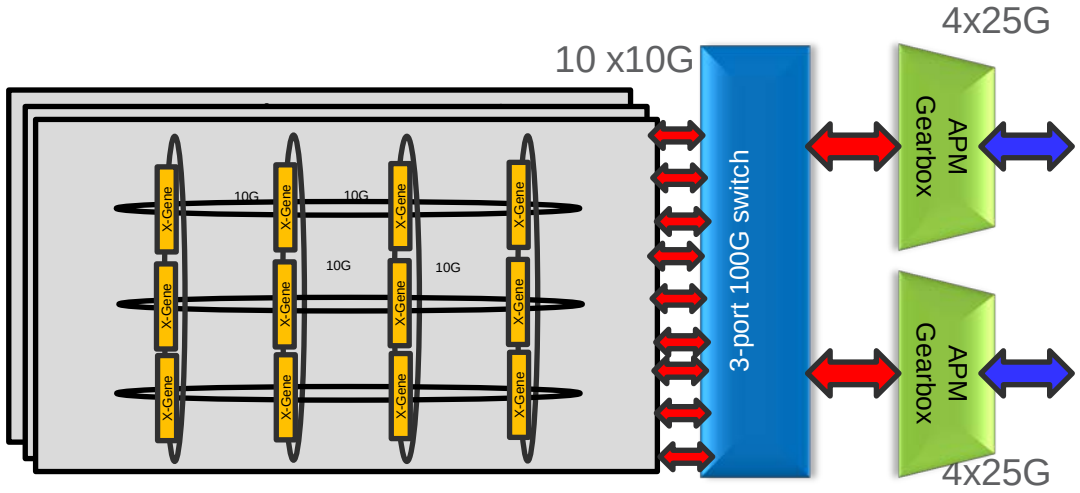
- Distributed 10G L2/ L3 switch on each server
- 6x10G + 2x10G switch ports per server
- 3D Torus fabric
- Low latency RoCEE communication between nodes



- 100G chassis uplink

Next Gen Intra-Rack Concepts

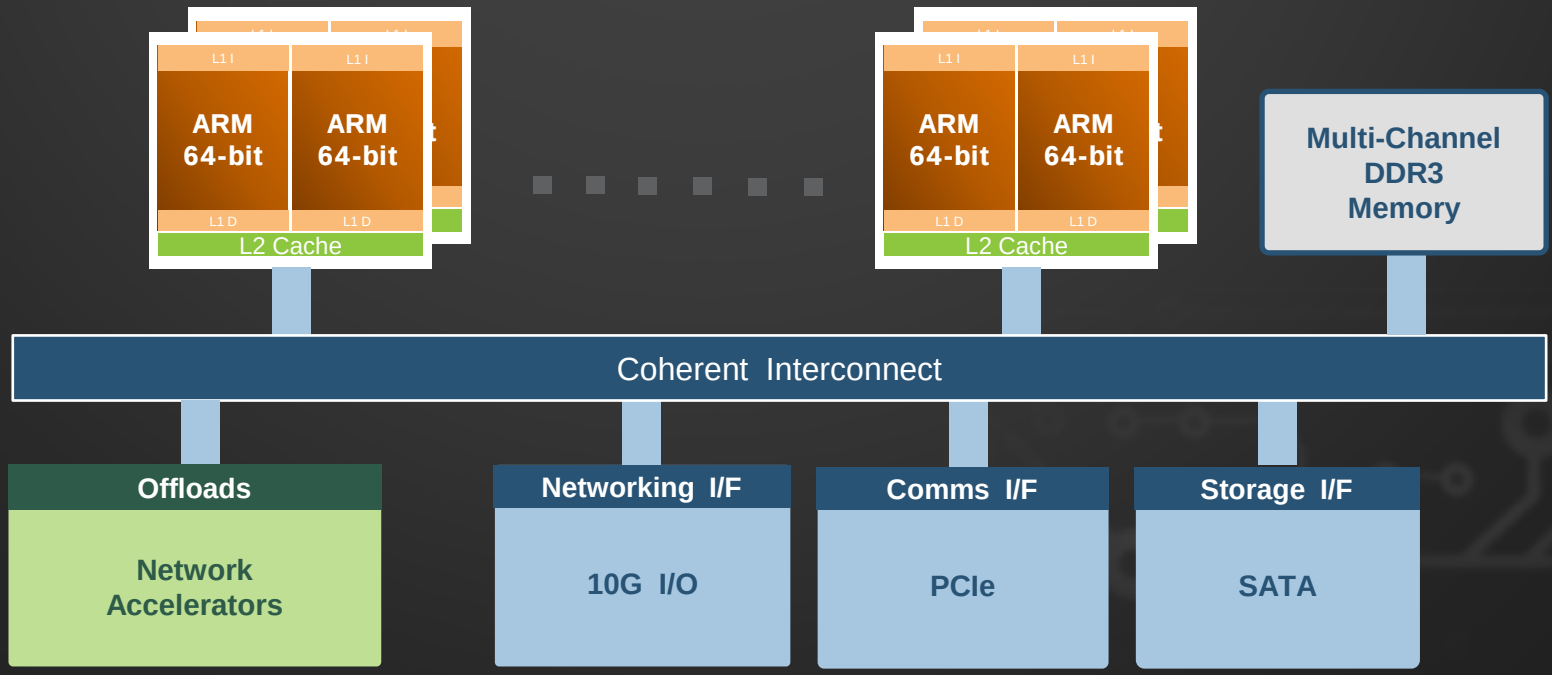




GroupHug Chassis Switch

- Multiple chassis daisy-chained at 100G; one 100G to ToR
- 3port 100G switch → chassis-to-chassis and chassis-to-ToR communication at 100G
- Low latency 100G path to Flash/ SSD storage

X-Gene™: Right-Sized for the Cloud



Server Class 64bit Architecture

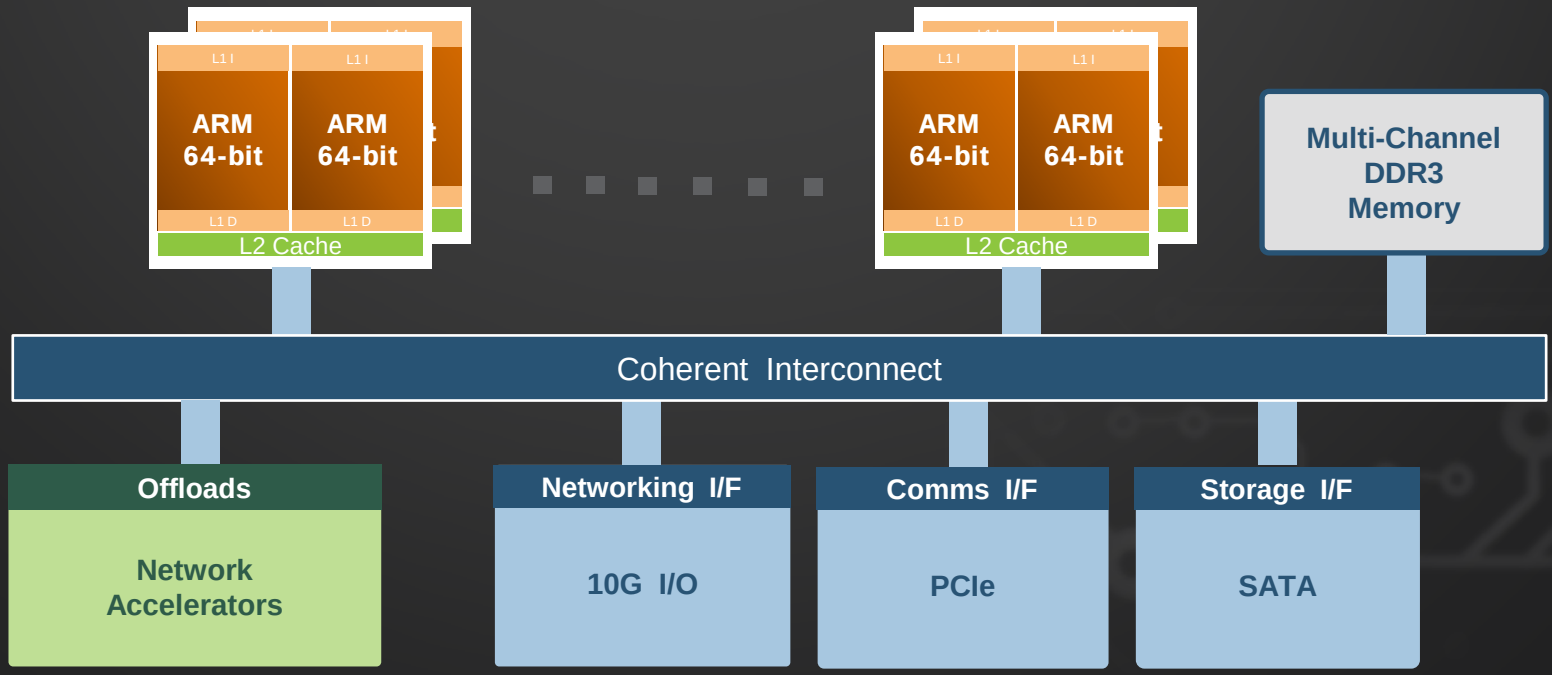
High Single Thread Performance

Server on a Chip™ Integration

Unique Mixed-Signal IP: 10/40/100G

First to Market

X-Gene™: Right-Sized for the Cloud



Samples Availability: Q1’13.

Server Availability: 2H’13.


apm
applied
micro