



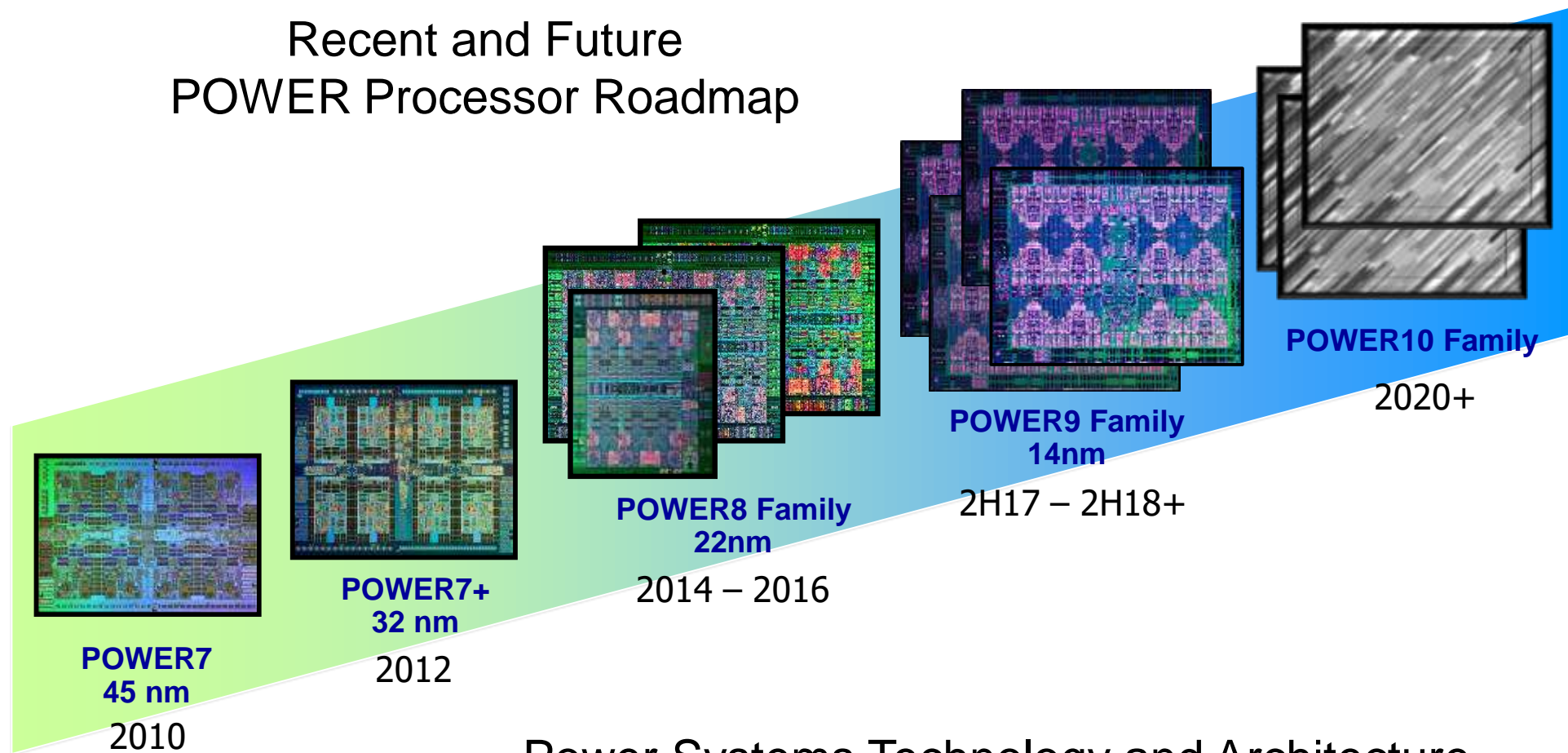
POWER9

Jeff Stuecheli

POWER Systems, IBM Systems



Recent and Future POWER Processor Roadmap



Power Systems Technology and Architecture
Leveraging the economics of the New Era

New Core Microarchitecture

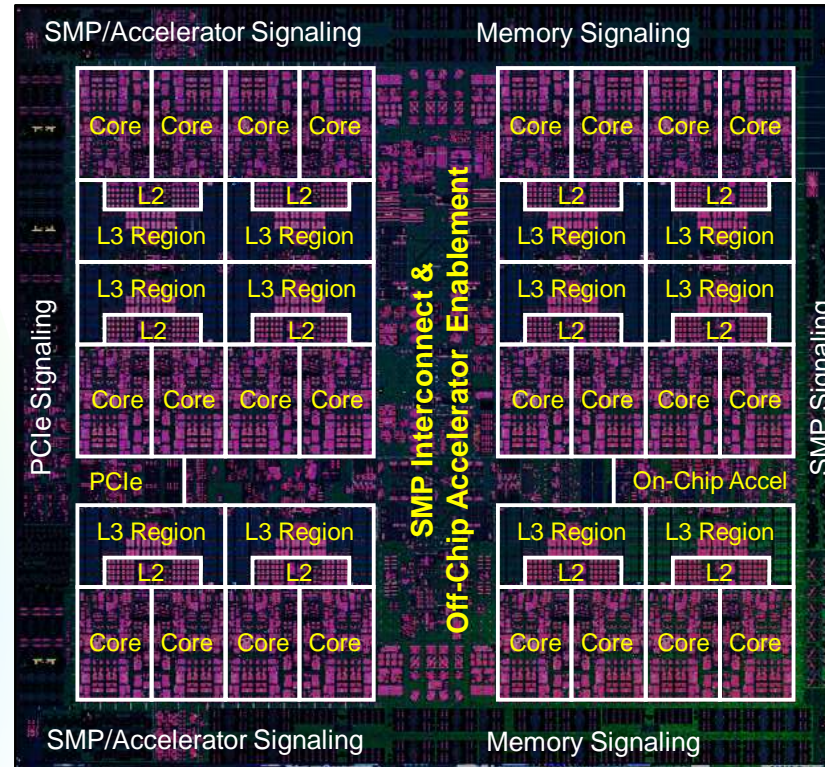
- Stronger thread performance
- Efficient agile pipeline
- POWER ISA v3.0

Enhanced Cache Hierarchy

- 120MB NUCA L3 architecture
- 12 x 20-way associative regions
- Advanced replacement policies
- Fed by 7 TB/s on-chip bandwidth

Cloud + Virtualization Innovation

- Quality of service assists
- New interrupt architecture
- Energy Scale (Workload optimized frequency)



14nm finFET Semiconductor Process

- Improved device performance and reduced energy
- 17 layer metal stack and eDRAM
- 8.0 billion transistors

Leadership

Hardware Acceleration Platform

- Enhanced on-chip acceleration
- Nvidia NVLink 2.0: High bandwidth and advanced new features (25G)
- CAPI 2.0: Coherent accelerator and storage attach (PCIe G4)
- OpenCAPI: Improved latency and bandwidth, open interface (25G)

State of the Art I/O Subsystem

- PCIe Gen4 – 48 lanes

High Bandwidth Signaling Technology

- 16 GT/s interface
 - Local SMP
- 25 GT/s Common Link interface
 - Accelerator, remote SMP

Optimized for Stronger Thread Performance and Efficiency

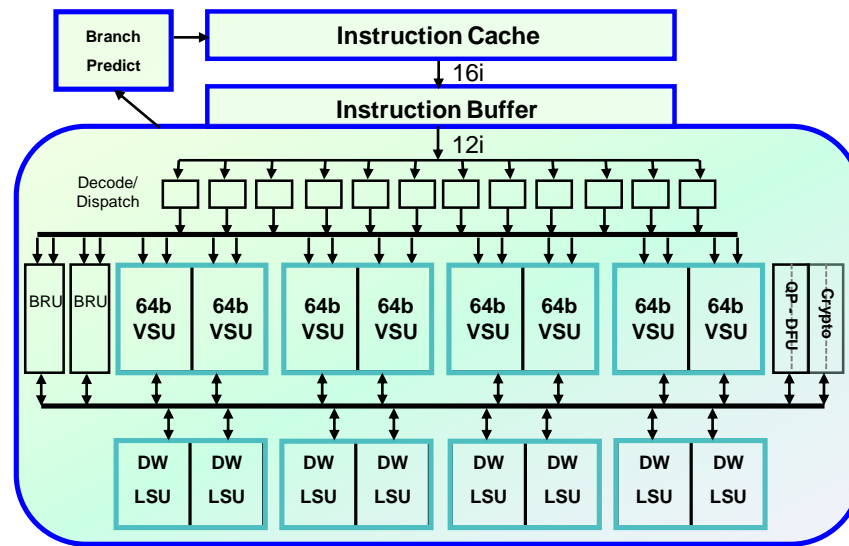
- Increased execution bandwidth efficiency for a range of workloads including commercial, cognitive and analytics
- Sophisticated instruction scheduling and branch prediction for unoptimized applications and interpretive languages
- Adaptive features for improved efficiency and performance especially in lower memory bandwidth systems

Available with SMT8 or SMT4 Cores

8 or 4 threaded core built from modular execution slices

POWER9 SMT8 Core

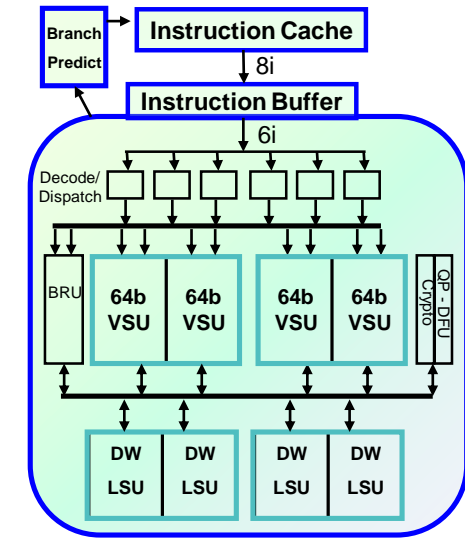
- PowerVM Ecosystem Continuity
- Strongest Thread
- Optimized for Large Partitions



SMT8 Core

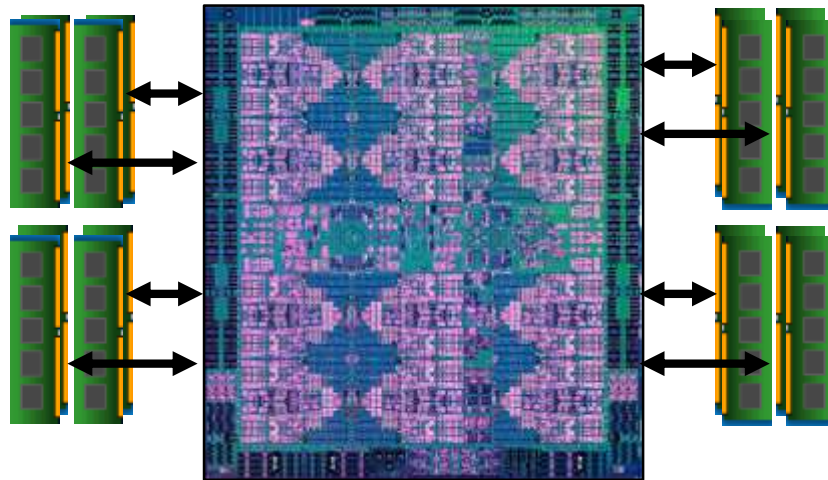
POWER9 SMT4 Core

- Linux Ecosystem Focus
- Core Count / Socket
- Virtualization Granularity



SMT4 Core

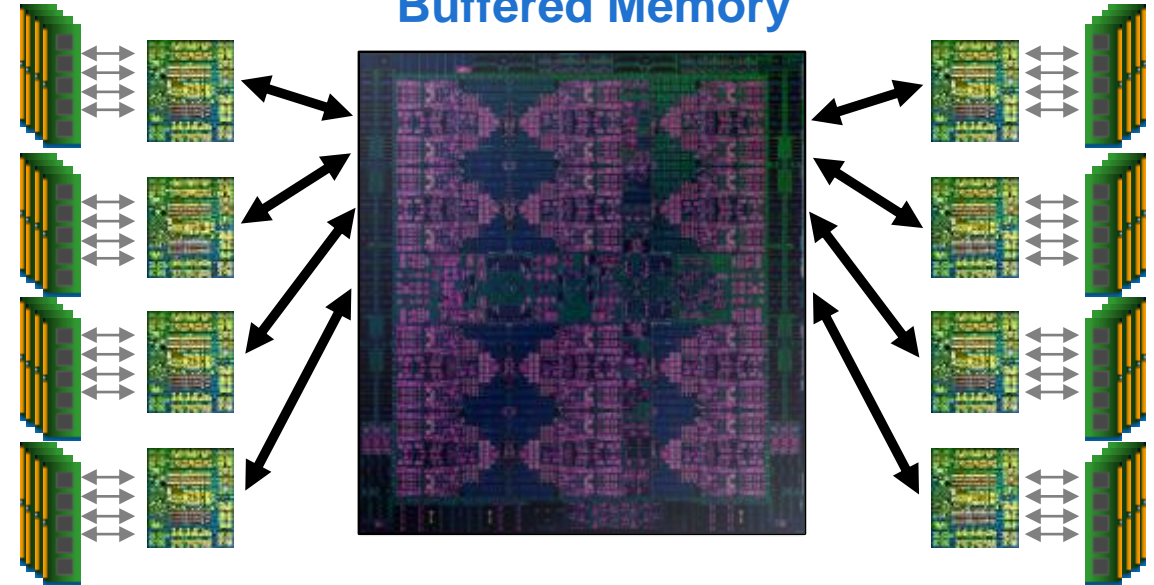
Scale Out Direct Attach Memory



8 Direct DDR4 Ports

- Up to 120 GB/s of sustained bandwidth
- Low latency access
- Commodity packaging form factor
- Adaptive 64B / 128B reads

Scale Up Buffered Memory



8 Buffered Channels

- Up to 230GB/s of sustained bandwidth
- Extreme capacity – up to 8TB / socket
- Superior RAS with chip kill and lane sparing
- Compatible with POWER8 system memory
- Agnostic interface for alternate memory innovations

Four targeted implementations

Core Count / Size

SMP scalability / Memory subsystem

SMT4 Core

24 SMT4 Cores / Chip

Linux Ecosystem Optimized

SMT8 Core

12 SMT8 Cores / Chip

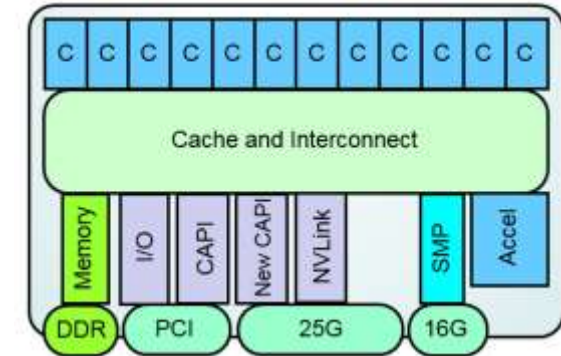
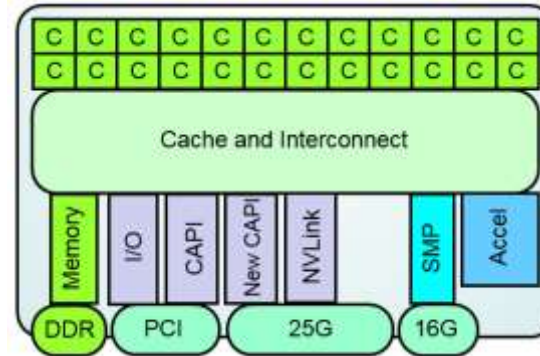
PowerVM Ecosystem Continuity

Scale-Out – 2 Socket Optimized

Robust 2 socket SMP system

Direct Memory Attach

- Up to 8 DDR4 ports
- Commodity packaging form factor



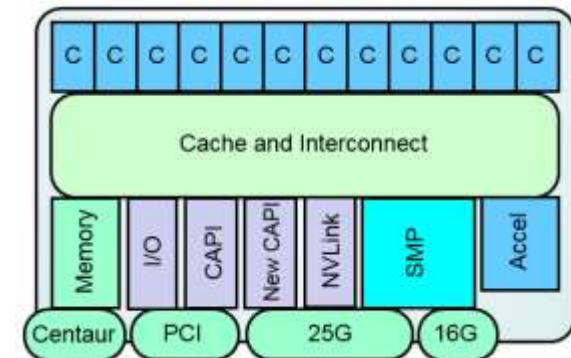
Scale-Up – Multi-Socket Optimized

Scalable System Topology / Capacity

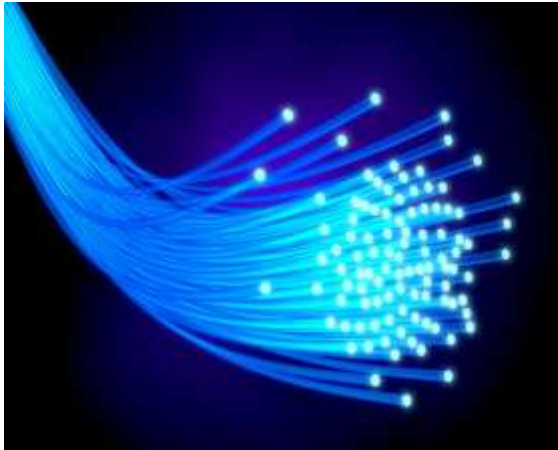
- Large multi-socket

Buffered Memory Attach

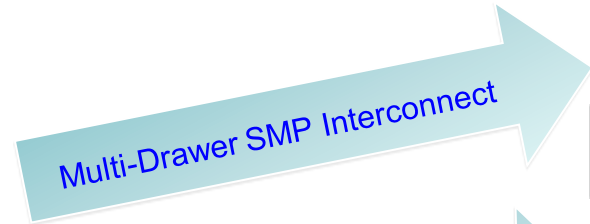
- 8 Buffered channels



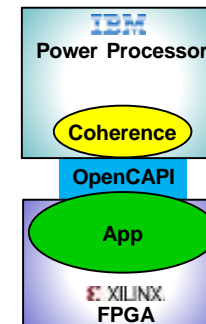
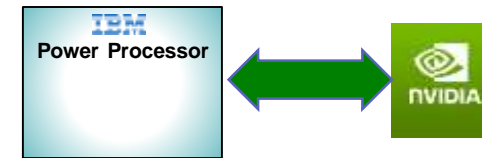
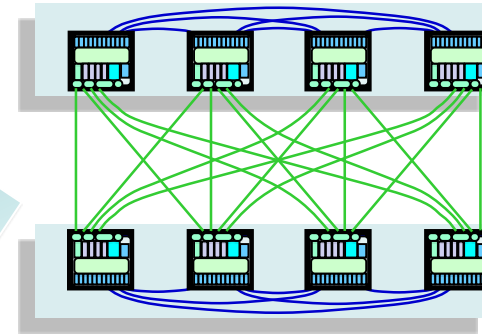
Modular Constructs → High-speed 25 GT/s Signaling



**Utilize Best-of-Breed
25 GT/s Optical-Style
Signaling Technology**



**Flexible & Modular
Packaging
Infrastructure**



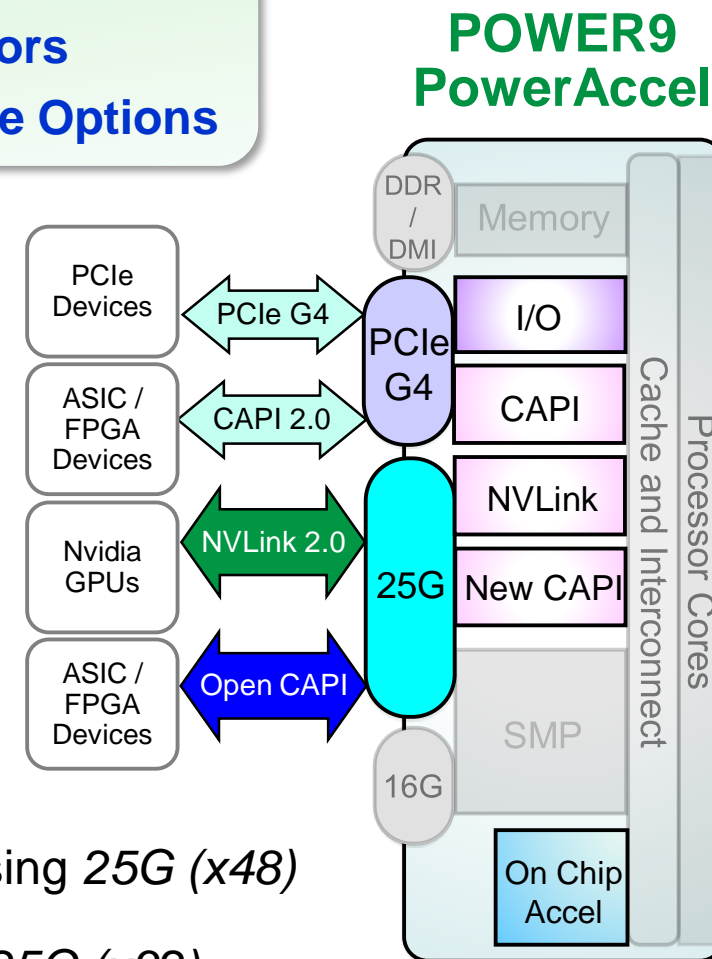
- Extreme Processor / Accelerator Bandwidth and Reduced Latency
- Coherent Memory and Virtual Addressing Capability for all Accelerators
- OpenPOWER Community Enablement – Robust Accelerated Compute Options

• **State of the Art I/O and Acceleration Attachment Signaling**

- PCIe Gen 4 x 48 lanes – 192 GB/s duplex bandwidth
- 25Gb/s Common Link x 48 lanes – 300 GB/s duplex bandwidth

• **Robust Accelerated Compute Options with OPEN standards**

- On-Chip Acceleration – Gzip x1, 842 Compression x2, AES/SHA x2
- CAPI 2.0 – 4x bandwidth of POWER8 using *PCIe Gen 4*
- NVLink 2.0 – Next generation of GPU/CPU bandwidth and integration using 25G (x48)
- Open CAPI 3.0 – High bandwidth, low latency and open interface using 25G (x32)



OpenCAPI – An Open heterogeneous architecture standard

Power Systems

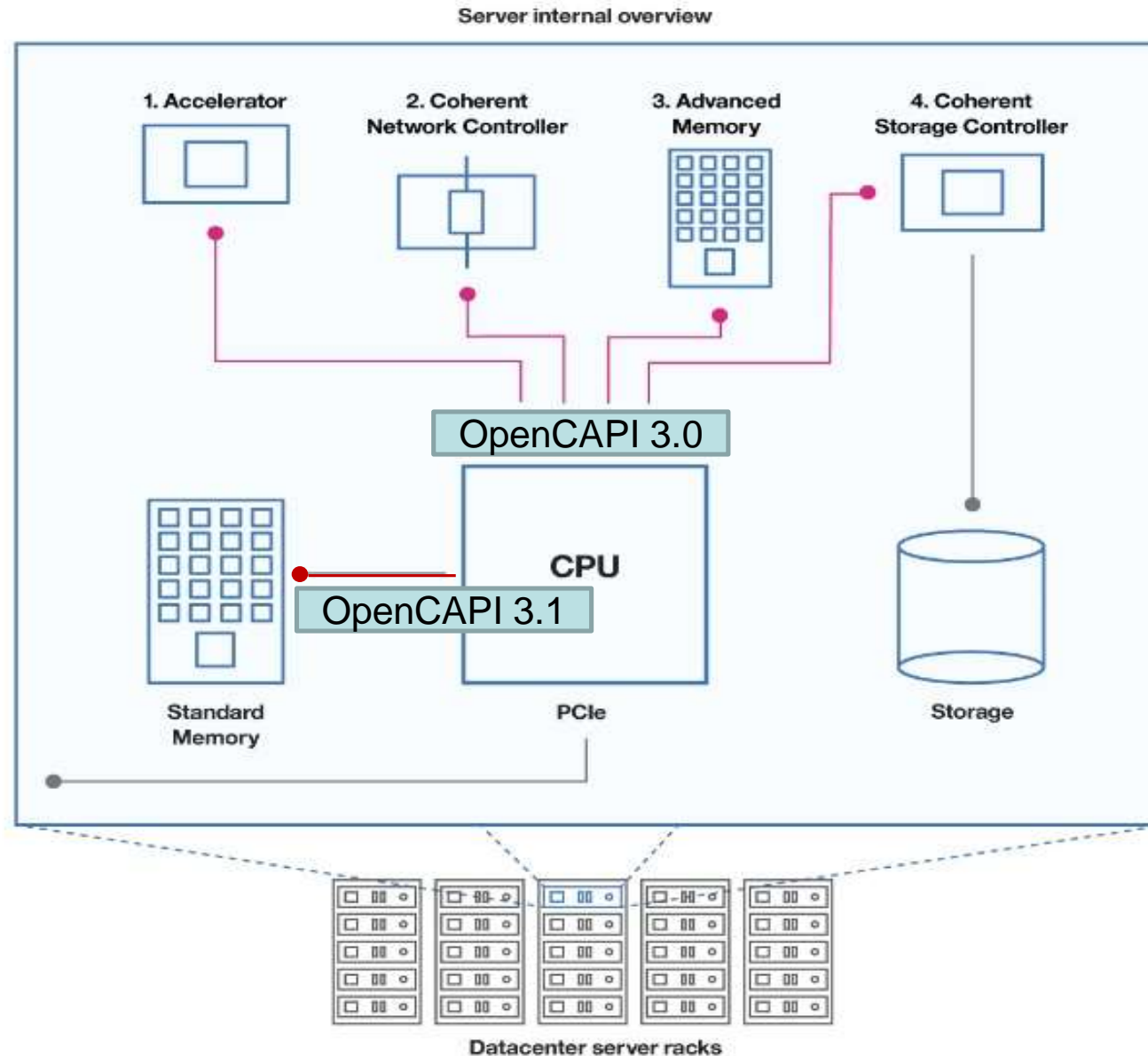


1. Accelerators: The performance, virtual addressing and coherence capabilities allow FPGA and ASIC accelerators to behave as if they were integrated into a custom microprocessor.

2. Coherent Network Controller: OpenCAPI provides the bandwidth that will be needed to support rapidly increasing network speeds. Network controllers based on virtual addressing can eliminate software overhead without the programming complexity usually associated with user-level networking protocols.

3. Advanced Memory: OpenCAPI allows system designers to take full advantage of emerging memory technologies to change the economics of the datacenter.

4. Coherent Storage Controller: OpenCAPI allows storage controllers to bypass kernel software overhead, enabling extreme IOPS performance without wasting valuable CPU cycles.



OpenCAPI specifications are downloadable from the website

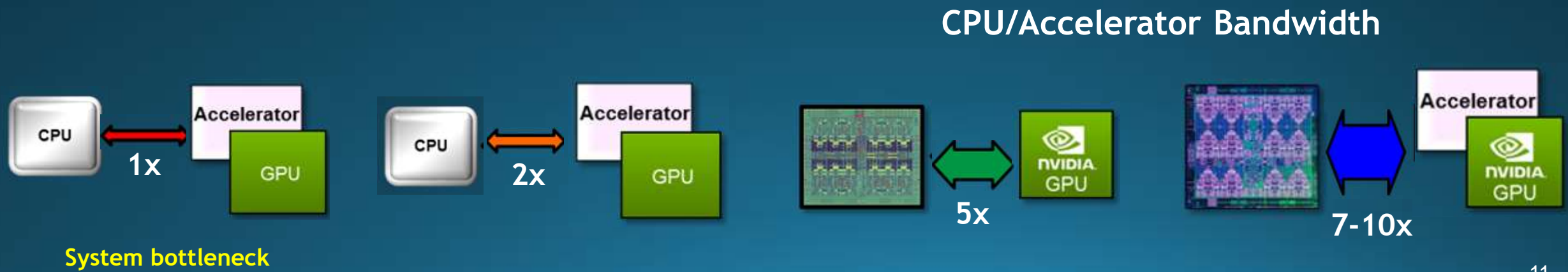
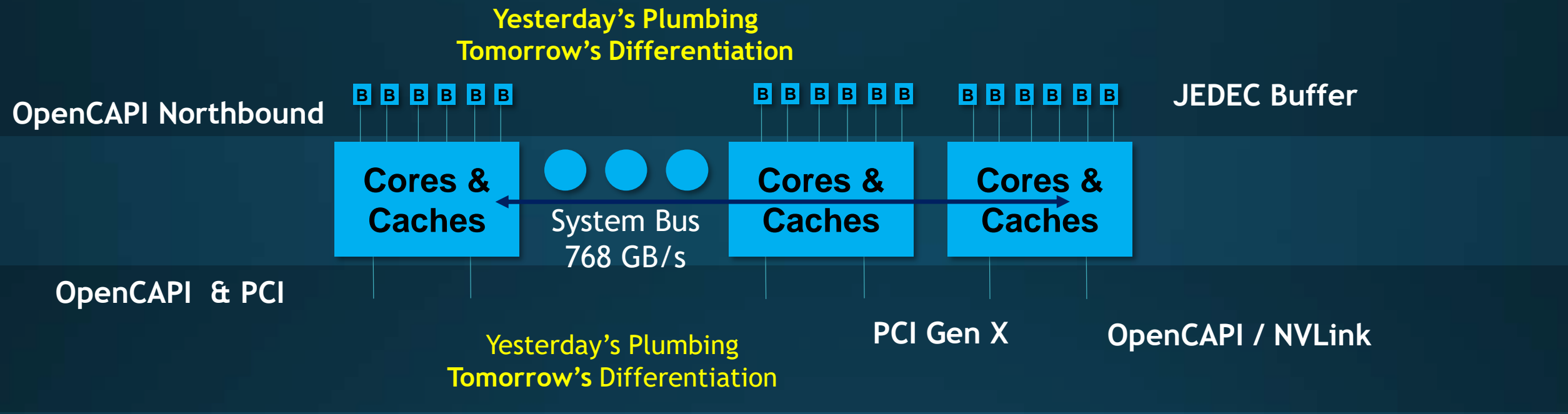
at www.opencapi.org

- Register
- Download

Proposed POWER Processor Technology and I/O Roadmap

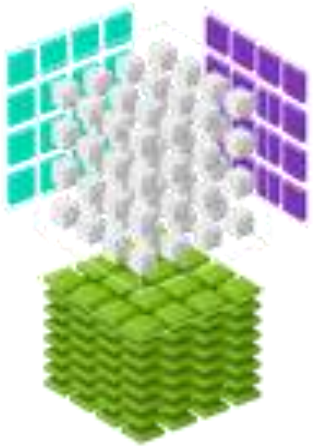
	POWER7 Architecture		POWER8 Architecture		POWER9 Architecture			POWER10
	2010 POWER7 8 cores 45nm New Micro-Architecture New Process Technology	2012 POWER7+ 8 cores 32nm Enhanced Micro-Architecture New Process Technology	2014 POWER8 12 cores 22nm New Micro-Architecture New Process Technology	2016 POWER8 w/ NVLink 12 cores 22nm Enhanced Micro-Architecture With NVLink	2017 P9 SO 12/24 cores 14nm New Micro-Architecture Direct attach memory New Process Technology	2018 P9 SU 12/24 cores 14nm Enhanced Micro-Architecture Buffered Memory	2019 P9 w/ Adv. I/O 12/24 cores 14nm Enhanced Micro-Architecture New Memory Subsystem	2020+ P10 TBD cores New Micro-Architecture New Technology
Sustained Memory Bandwidth	Up To 65 GB/s	Up To 65 GB/s	Up To 210 GB/s	Up To 210 GB/s	Up To 150 GB/s	Up To 210 GB/s	Up To 350 GB/s	Up To 435 GB/s
Standard I/O Interconnect	PCIe Gen2	PCIe Gen2	PCIe Gen3	PCIe Gen3	PCIe Gen4 x48	PCIe Gen4 x48	PCIe Gen4 x48	PCIe Gen5
Advanced I/O Signaling	N/A	N/A	N/A	20 GT/s 160GB/s	25 GT/s 300GB/s	25 GT/s 300GB/s	25 GT/s 300GB/s	32 & 50 GT/s
Advanced I/O Architecture	N/A	N/A	CAPI 1.0	CAPI 1.0 , NVLink 1.0	CAPI 2.0, OpenCAPI3.0, NVLink2.0	CAPI 2.0, OpenCAPI3.0, NVLink2.0	CAPI 2.0, OpenCAPI4.0, NVLink3.0	TBD

Future Evolution of System Architecture



CORAL – IBM Delivers Summit and Sierra

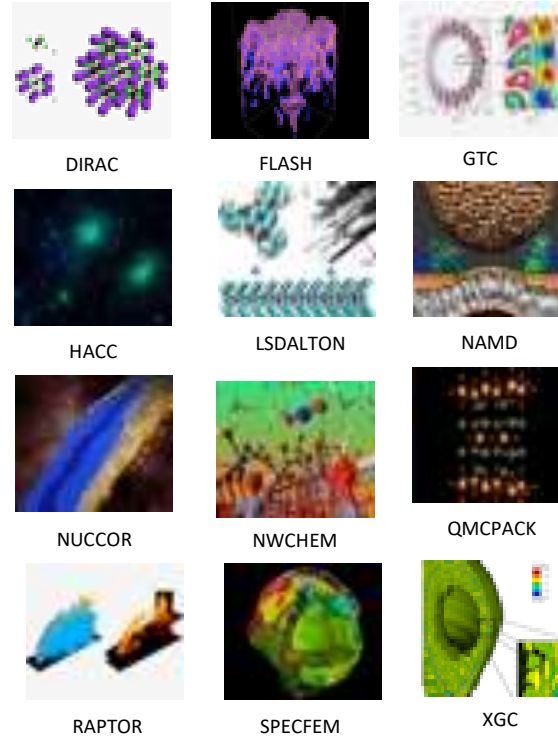
AI Exascale Today



Order of Magnitude Leap in Computational Power



Real, Accelerated Science



3+EFLOPS

Tensor Ops

10X

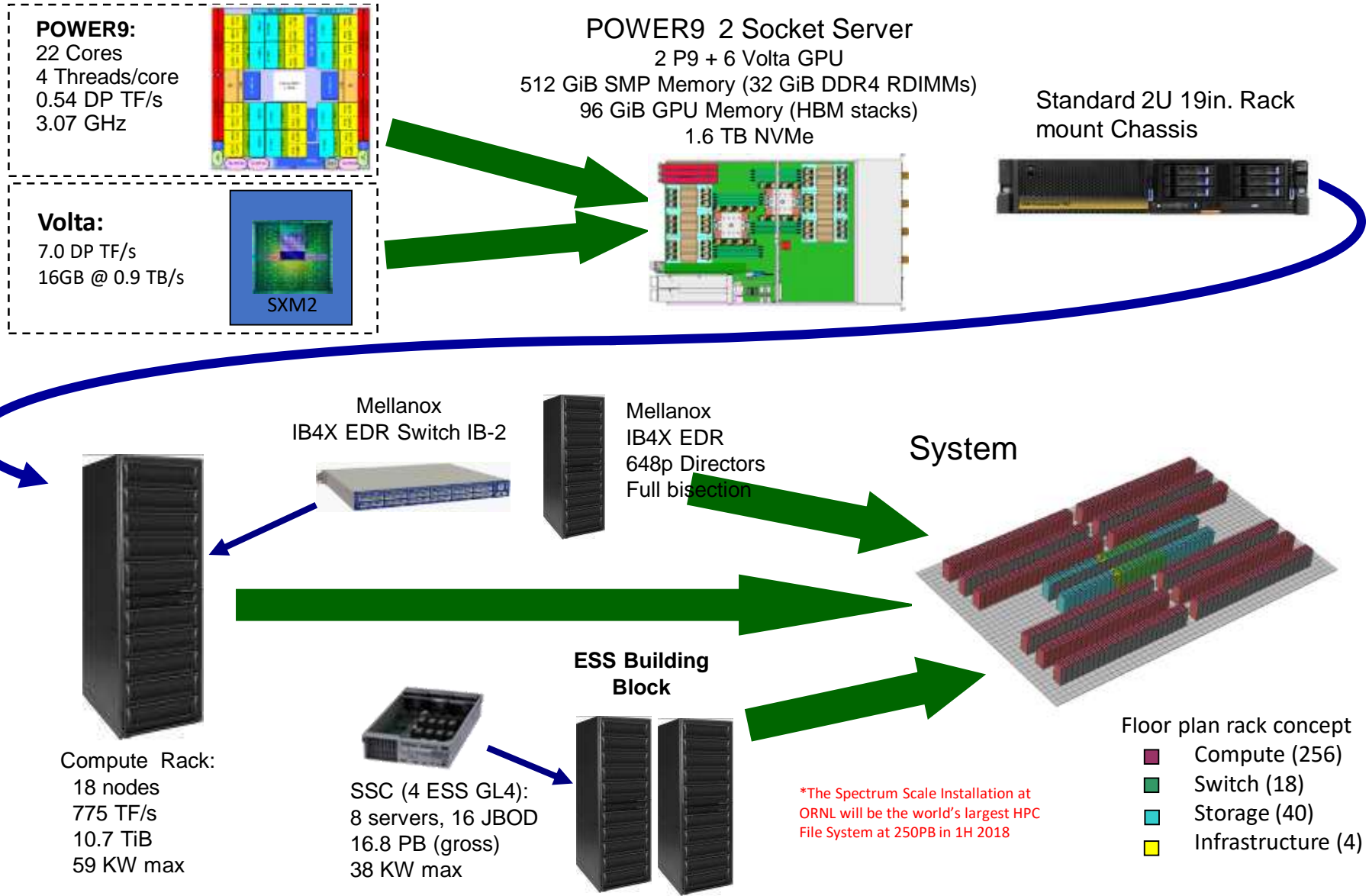
Perf Over Titan

5-10X

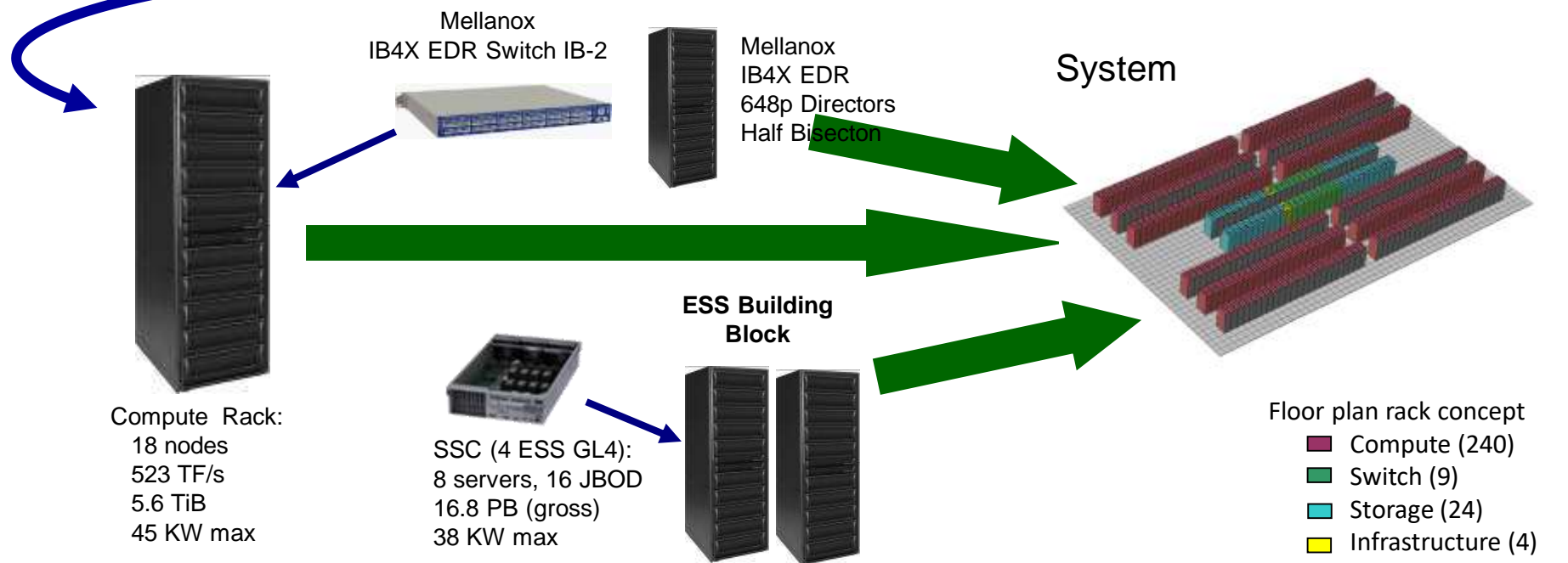
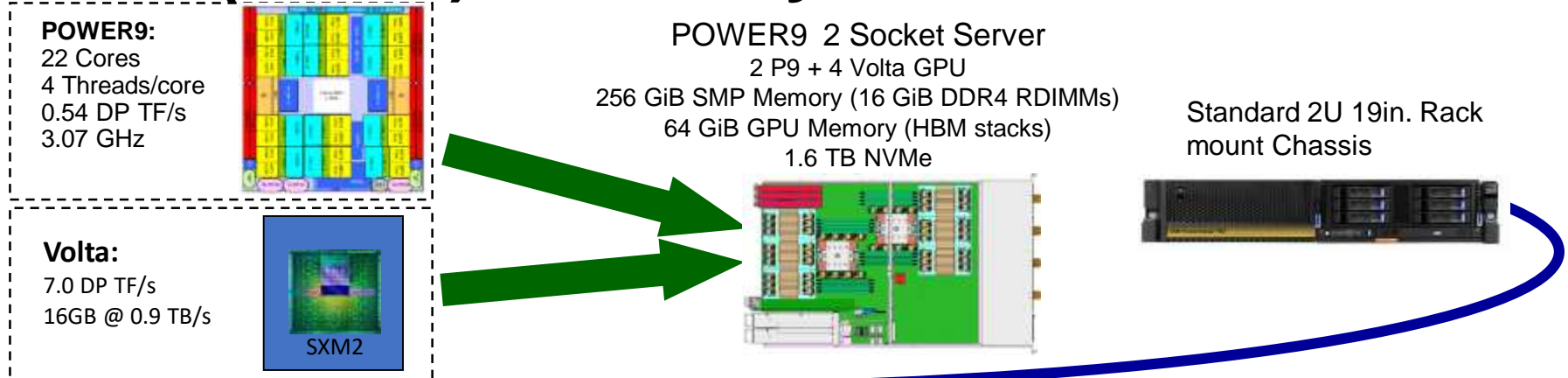
Application Perf Over Titan

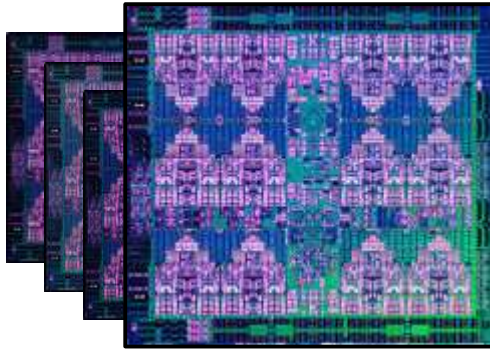
- Deployments beginning with full acceptance in 2018
- Significant application performance over Titan (AMD/NVIDIA)
 - Achieved with 1/4 the servers

CORAL ORNL (Summit) 200PF System



CORAL LLNL (Sierra) 125PF System





Built for the Cognitive Era



Enhanced Core and Chip Architecture for Emerging Workloads

- New Core Optimized for Emerging Algorithms to Interpret and Reason
- Bandwidth, Scale, and Capacity, to Ingest and Analyze

Processor Family with Scale-Out and Scale-Up Optimized Silicon

- Enabling a Range of Platform Optimizations – from HSDC Clusters to Enterprise Class Systems
- Extreme Virtualization Capabilities for the Cloud

Premier Acceleration Platform

- Heterogeneous Compute Options to Enable New Application Paradigms
- State of the Art I/O
- Engineered to be Open

Special notices

This document was developed for IBM offerings in the United States as of the date of publication. IBM may not make these offerings available in other countries, and the information is subject to change without notice. Consult your local IBM business contact for information on the IBM offerings available in your area.

Information in this document concerning non-IBM products was obtained from the suppliers of these products or other public sources. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. Send license inquiries, in writing, to IBM Director of Licensing, IBM Corporation, New Castle Drive, Armonk, NY 10504-1785 USA.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

The information contained in this document has not been submitted to any formal IBM test and is provided "AS IS" with no warranties or guarantees either expressed or implied.

All examples cited or described in this document are presented as illustrations of the manner in which some IBM products can be used and the results that may be achieved. Actual environmental costs and performance characteristics will vary depending on individual client configurations and conditions.

IBM Global Financing offerings are provided through IBM Credit Corporation in the United States and other IBM subsidiaries and divisions worldwide to qualified commercial and government clients. Rates are based on a client's credit rating, financing terms, offering type, equipment type and options, and may vary by country. Other restrictions may apply. Rates and offerings are subject to change, extension or withdrawal without notice.

IBM is not responsible for printing errors in this document that result in pricing or information inaccuracies.

All prices shown are IBM's United States suggested list prices and are subject to change without notice; reseller prices may vary.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

Any performance data contained in this document was determined in a controlled environment. Actual results may vary significantly and are dependent on many factors including system hardware configuration and software design and configuration. Some measurements quoted in this document may have been made on development-level systems. There is no guarantee these measurements will be the same on generally-available systems. Some measurements quoted in this document may have been estimated through extrapolation. Users of this document should verify the applicable data for their specific environment.

Revised September 26, 2006

Special notices (continued)

IBM, the IBM logo, ibm.com AIX, AIX (logo), IBM Watson, DB2 Universal Database, POWER, PowerLinux, PowerVM, PowerVM (logo), PowerHA, Power Architecture, Power Family, POWER Hypervisor, Power Systems, Power Systems (logo), POWER2, POWER3, POWER4, POWER4+, POWER5, POWER5+, POWER6, POWER6+, POWER7, POWER7+, and POWER8 are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries.

A full list of U.S. trademarks owned by IBM may be found at: <http://www.ibm.com/legal/copytrade.shtml>.

NVIDIA, the NVIDIA logo, and NVLink are trademarks or registered trademarks of NVIDIA Corporation in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries or both.

PowerLinux™ uses the registered trademark Linux® pursuant to a sublicense from LMI, the exclusive licensee of Linus Torvalds, owner of the Linux® mark on a world-wide basis.

The Power Architecture and Power.org wordmarks and the Power and Power.org logos and related marks are trademarks and service marks licensed by Power.org.

The OpenPOWER word mark and the OpenPOWER Logo mark, and related marks, are trademarks and service marks licensed by OpenPOWER.

Other company, product and service names may be trademarks or service marks of others.