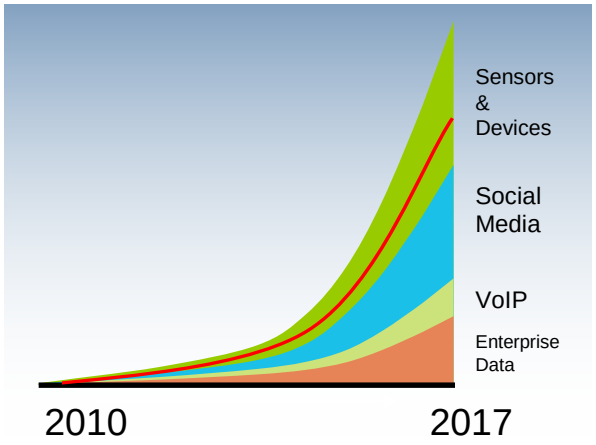
IBM Technical Computing

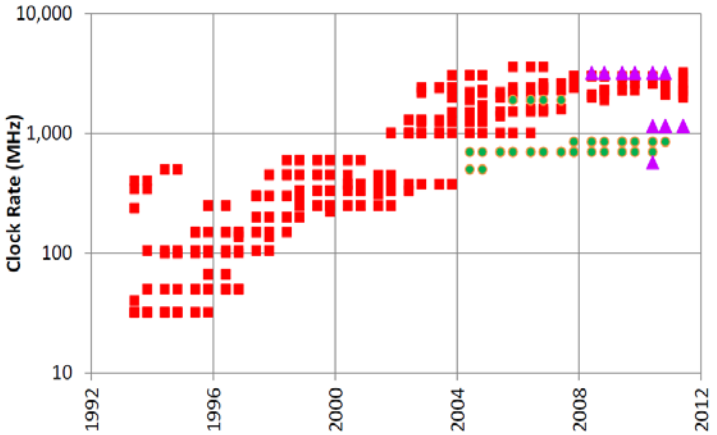
2015 ICT INAF workshop

Data growth outpaces Computing Technology elements

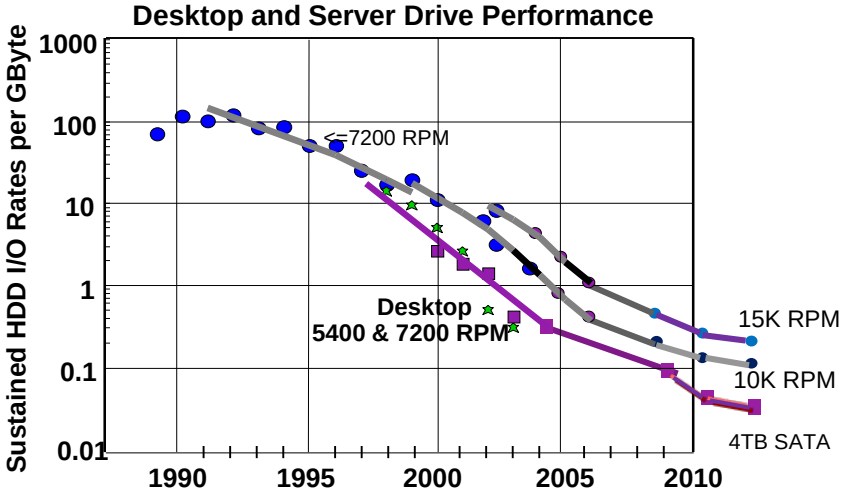
Data volume grows exponentially



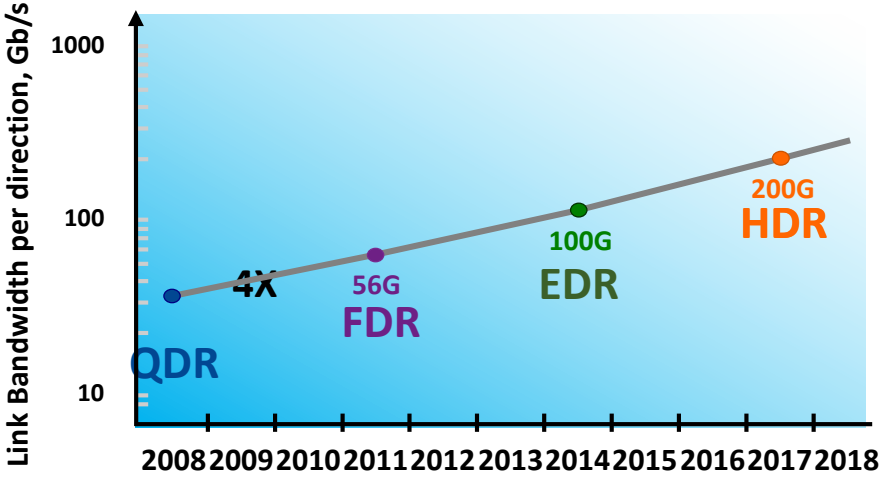
Microprocessor clock rates have stalled …



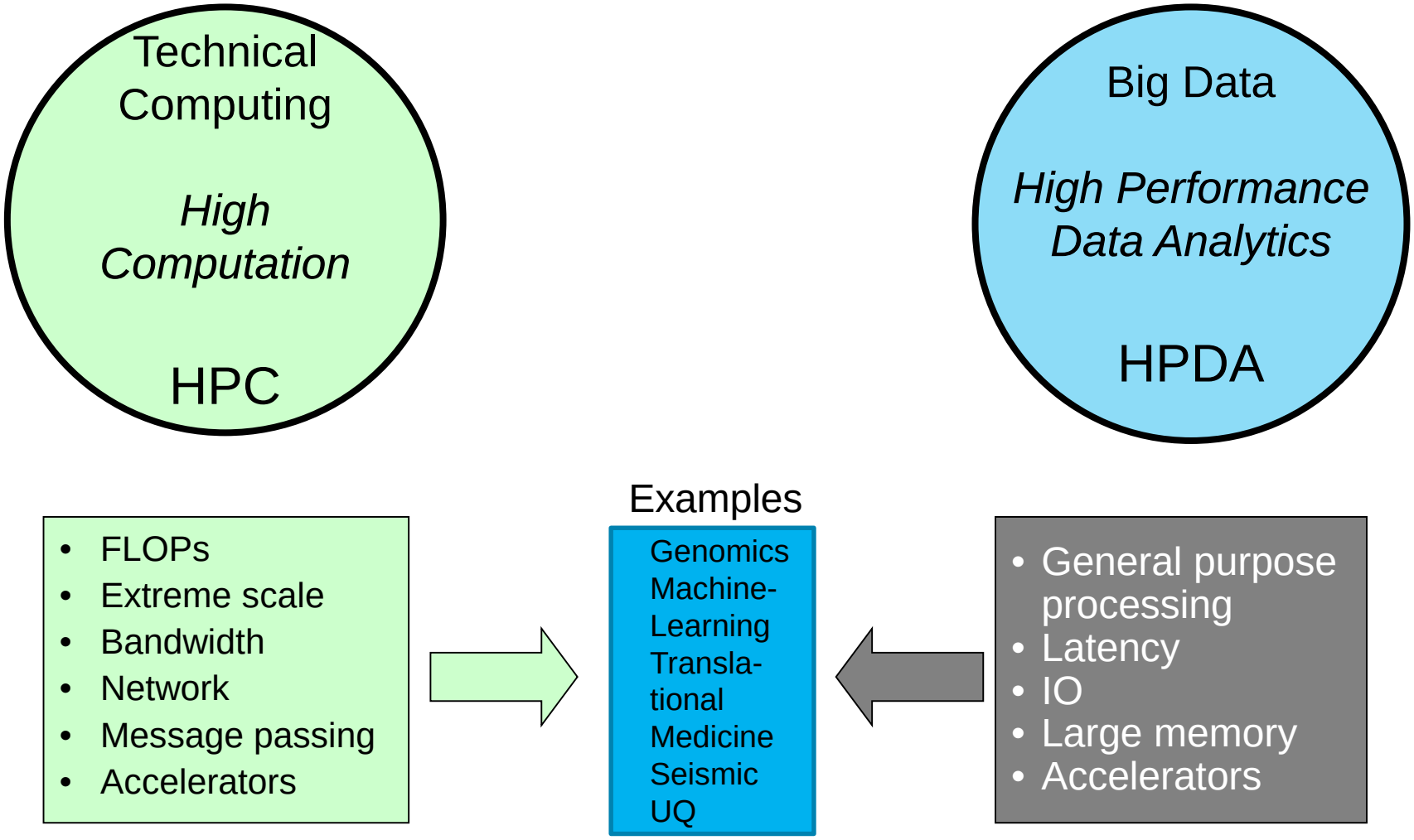
I/O Performance / Capacity loosing ground …



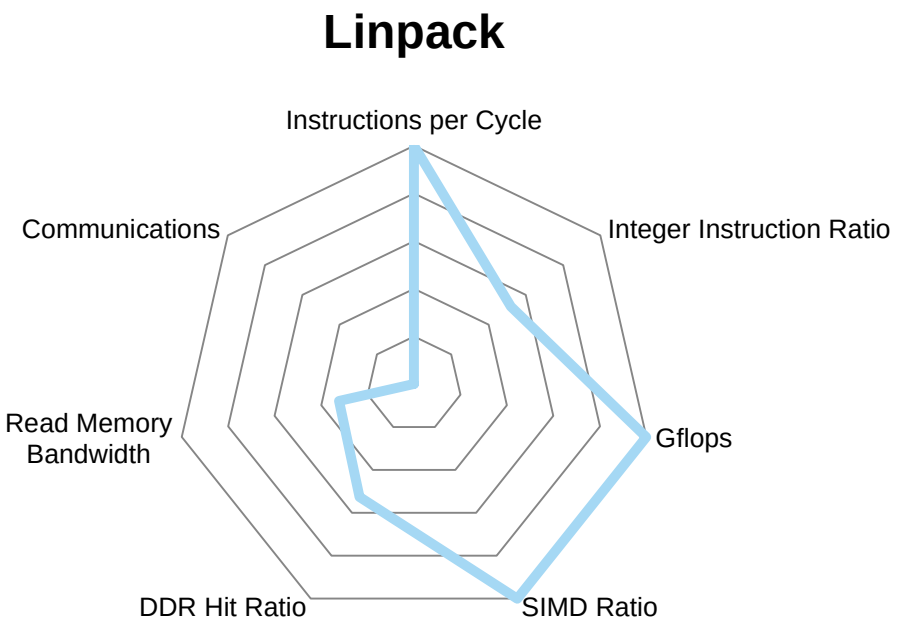
And network bandwidth can't keep up.



The Evolution of Demand

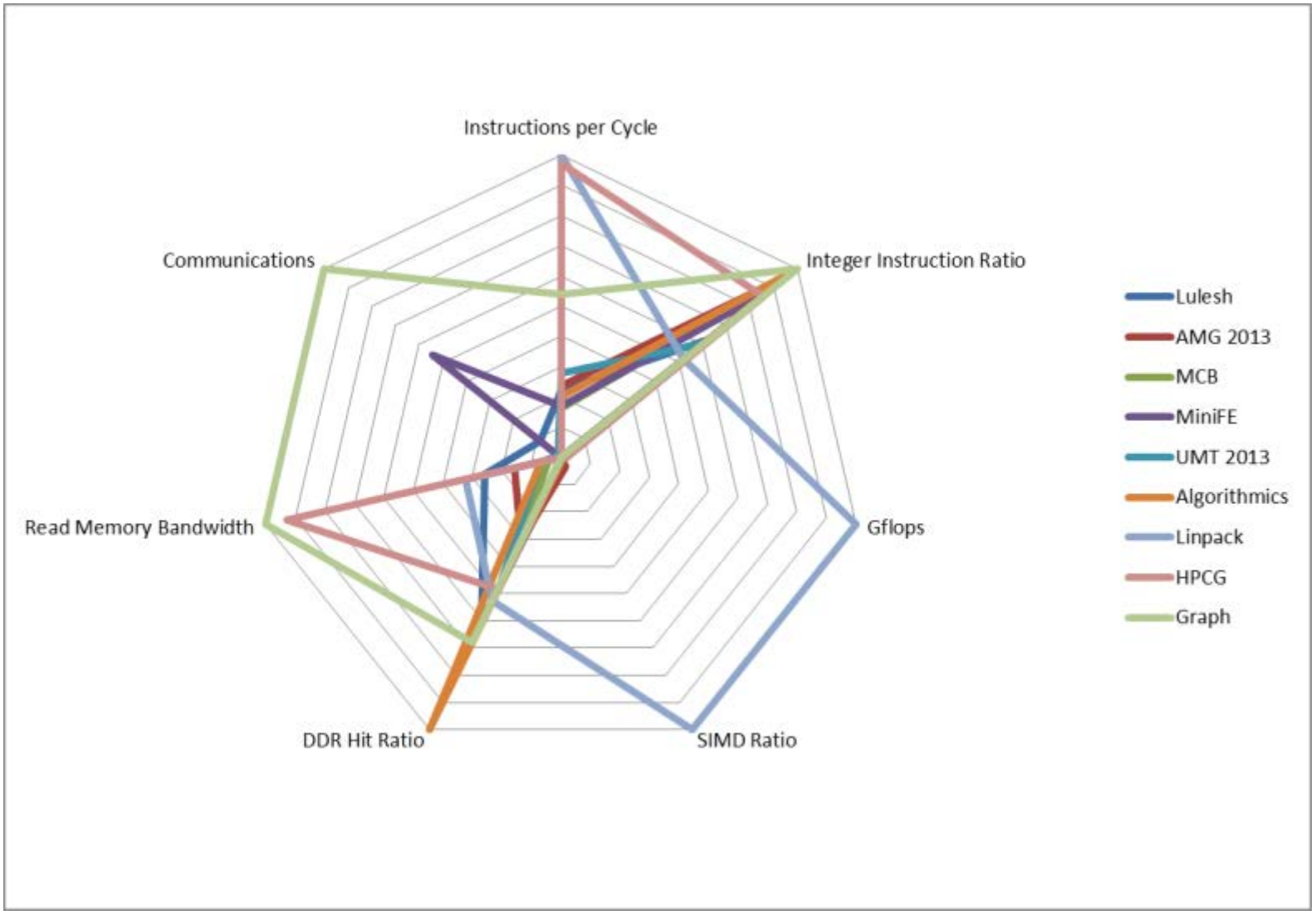


An Application Fingerprint: Benchmark



- High Floating Point
- Very high use of SIMD instructions
- Low use of main memory bandwidth

The Diversity of Architectural Requirements



The IBM Technical Computing strategy is to provide accelerated computing solutions with a data centric focus

The systems will employ heterogeneous (hybrid) architectures combining POWER host processors and closely-coupled accelerators (GPUs and FPGAs) in clusters of scale-out servers, to provide the most cost effective solutions for compute-intensive and data-intensive problems.

IBM Data Centric Design Principles

The onslaught of Big Data requires a composable architecture for big data, complex analytics, modeling and simulation. The DCS architecture will appeal to segments experiencing an explosion of data and the associated computational demands

Principle 1: Minimize data motion
- Data motion is expensive
- Hardware and software to support & enable compute in data
- Allow workloads to run where they run best

Principle 2: Enable compute in all levels of the systems hierarchy
- Introduce "active" system elements, including network, memory, storage, etc.
- HW & SW innovations to support / enable compute in data

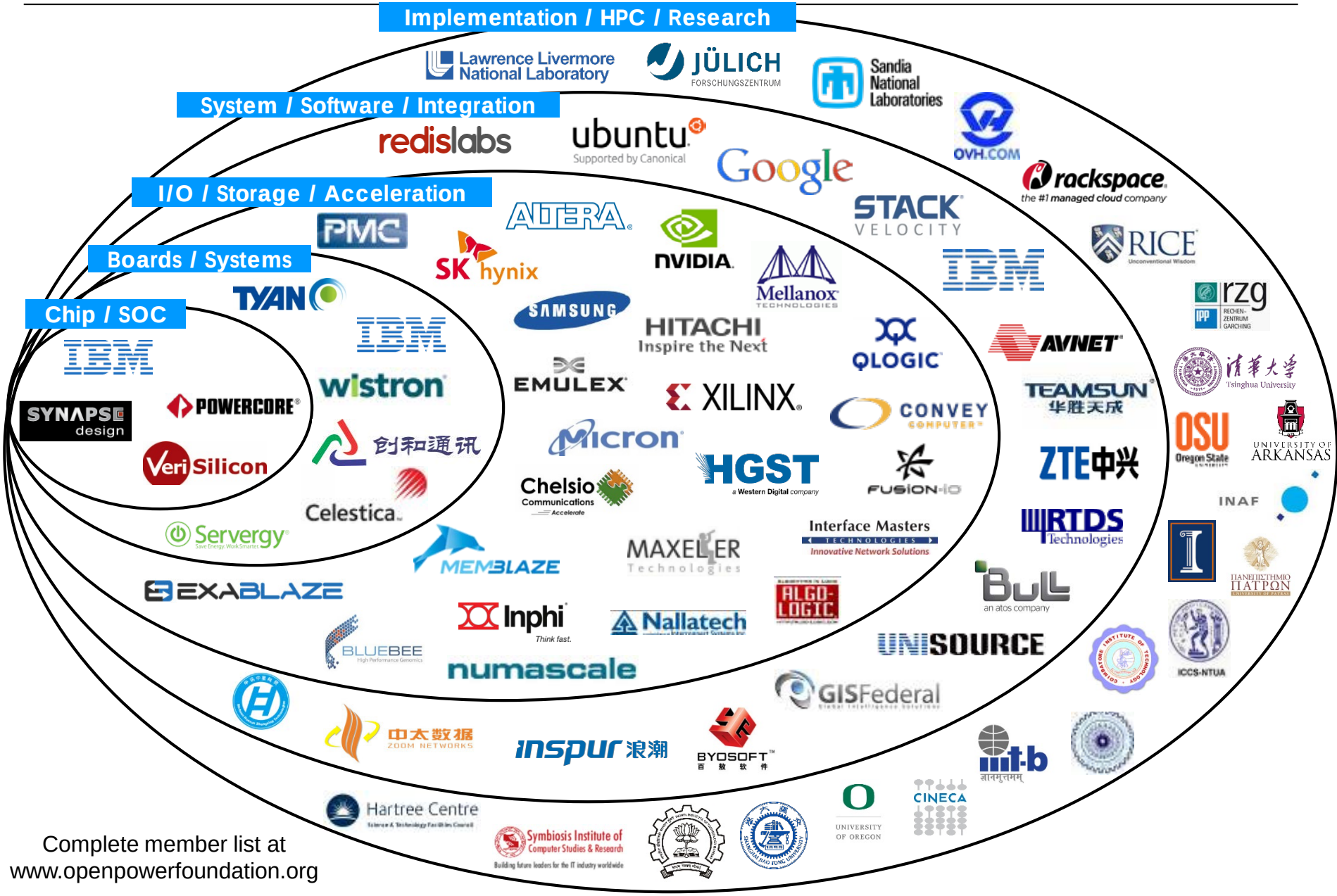
Principle 3: Modularity
- Balanced, composable architecture for Big Data analytics, modeling and simulation
- Modular and upgradeable design, scalable from sub rack to 100's of racks

Principle 4: Application-driven design
- Use real workloads/workflows to drive design points
- Co-design for customer value

Principle 5: Leverage OpenPOWER to accelerate innovation and broaden diversity for clients

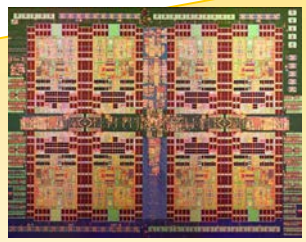
Architectural Elements
- Compute dense nodes for high efficiency
- Integrated compute and storage nodes for working set data and for data intensive work
- Common active network to offload remote computation and data transfer from the cores

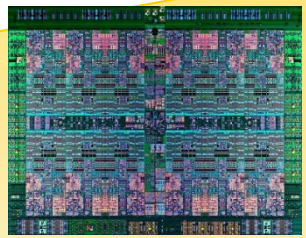
Fueling an Open Development Community

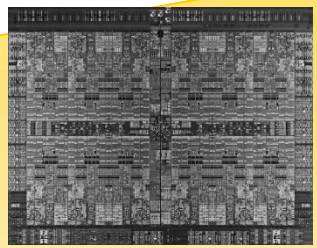


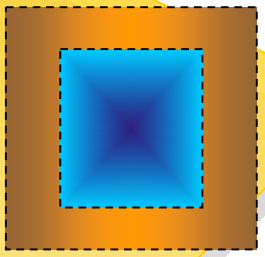
Complete member list at
www.openpowerfoundation.org

Continued Investment in POWER









POWER7+
32 nm

- 8 cores
- Scale-out
- Accelerators
- 2x SPFP
- Power Gating
- CPMs
- Extreme L3 Size

2012

POWER8
22 nm

- 12 Cores
- SMT8
- 2X DPFP
- PCIE Gen 3
- Coprocessor (CAPI)
- Enhanced Prefetch

2014

POWER8+
22 nm

- NVLINK1.0
- 2X CAPI

2016

POWER9
14 nm

- 24 Cores
- New µArchitecture
- Direct-attach DDR4
- Gen4 PCIe
- CAPI 2.0
- NVLINK2.0

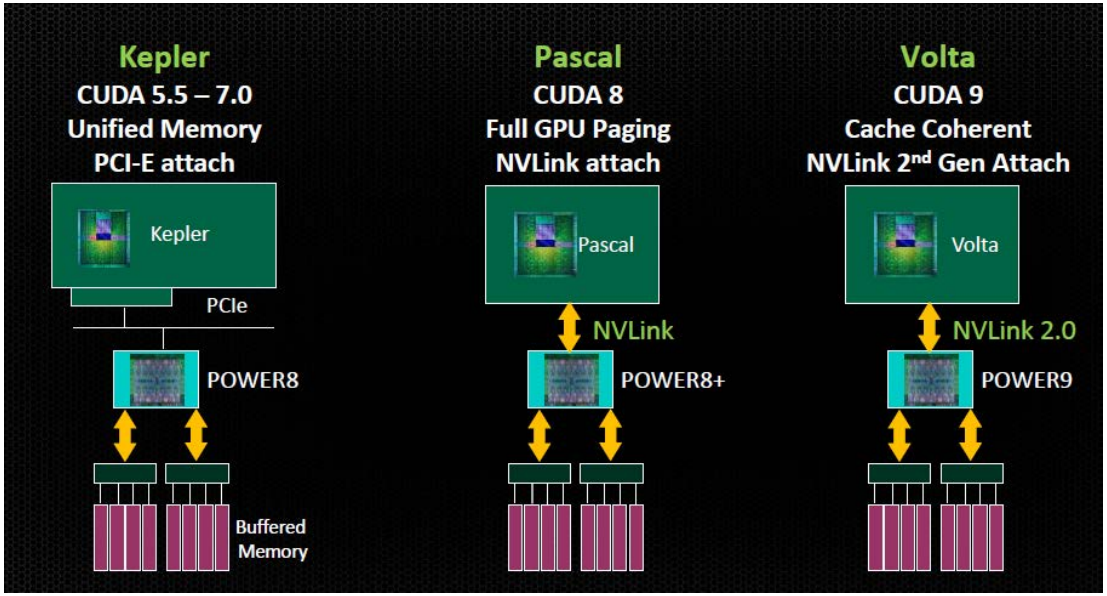
2017

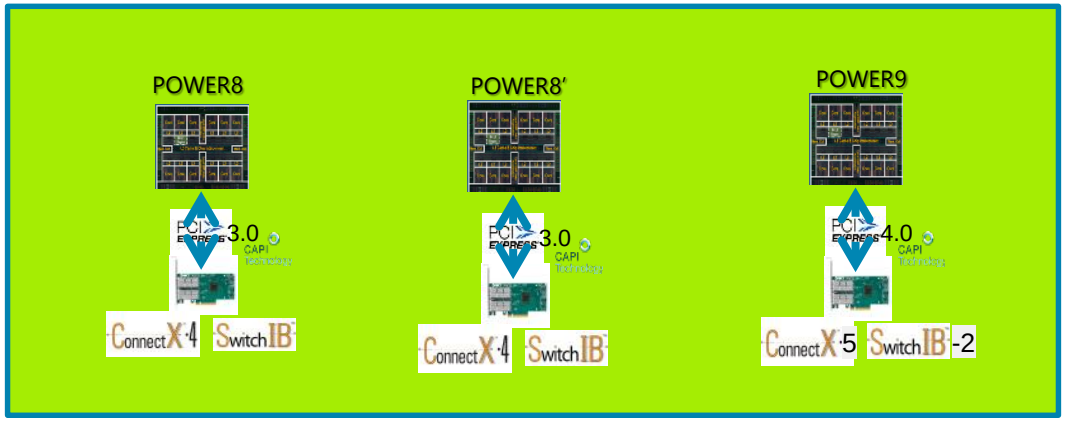
POWER10
10 nm

- >>24 Cores
- New µArchitecture
- Future NVLINK

Future

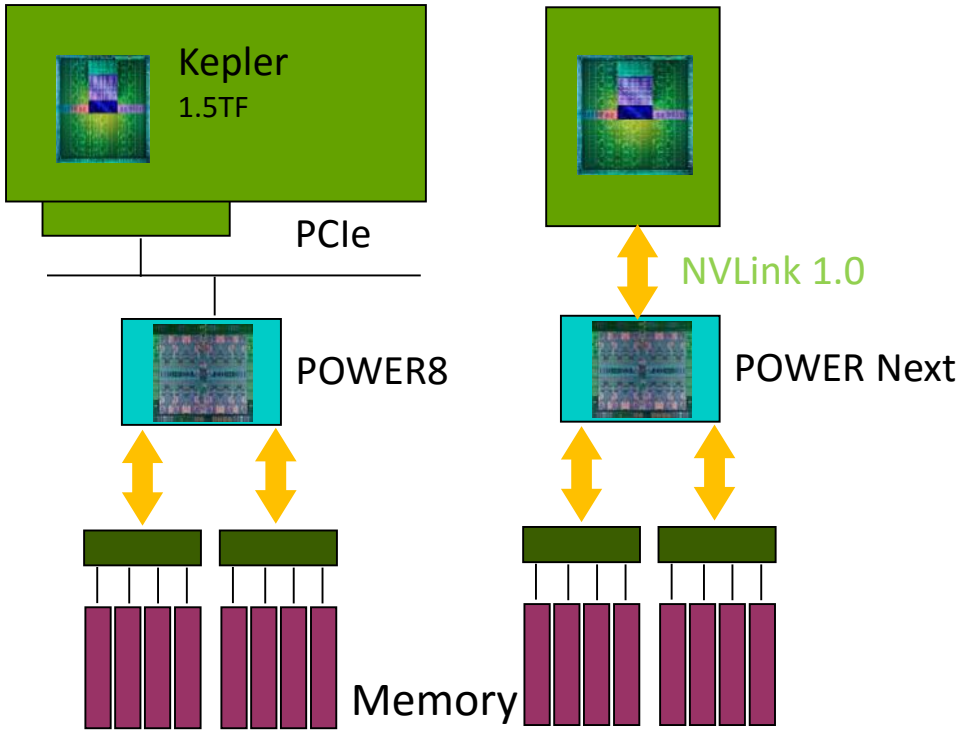
Key Close Partnerships



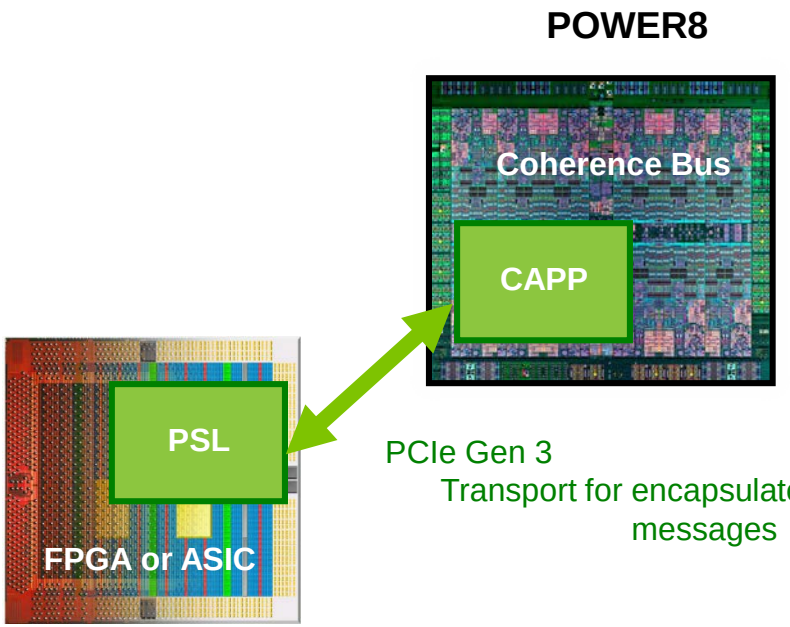


POWER Innovations – tighter integration with accelerators

NVLink – couples CPUs and GPUs closely
5-12x PCI-E Gen3 bandwidth!



CAPI – Coherent Accelerator Processor Interface
Hardware Managed Cache Coherence



- **Example workloads:**
  - Monte Carlo
  - Data compression
  - Streaming compression/decompression

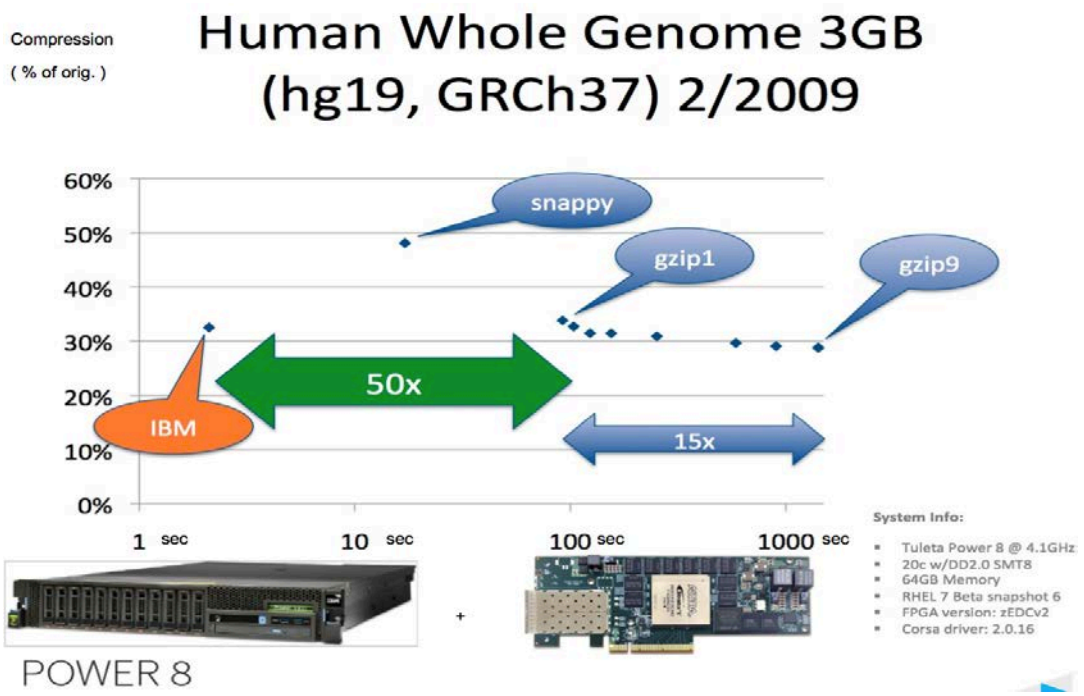
IBM Accelerated GZIP Compression

What it is:

- An FPGA-based low-latency GZIP Compressor & Decompressor with.

Results:

- **Single-thread** througput of ~2GB/s and a compression rate significantly better than low-CPU overhead compressors like snappy



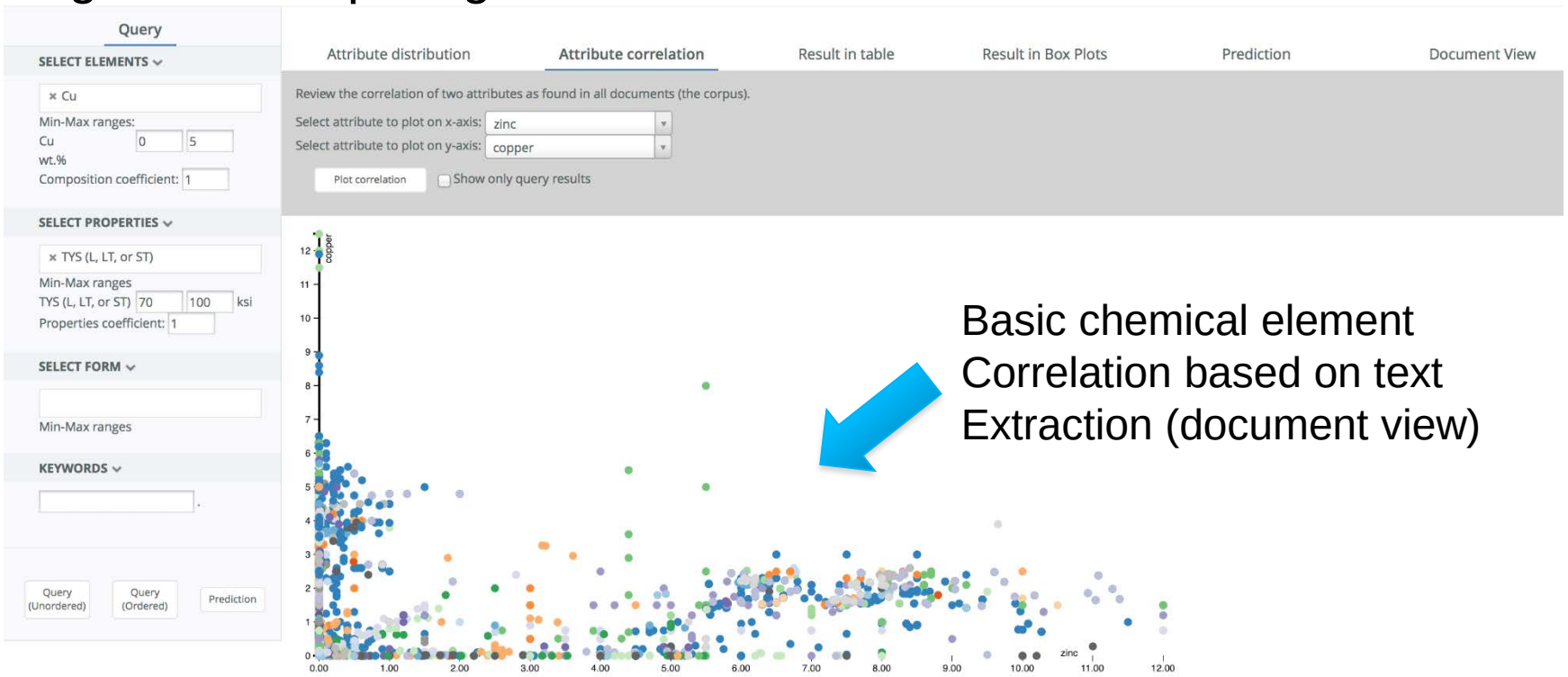
Monte Carlo Simulations

- Monte Carlo methods are used to price complex financial derivatives which in turn allows financial institutions determine the risk (e.g.VaR) of their investment portfolios
  - Required by regulators of risk compliance
  - Risk measures such as VaR are susceptible to underlining risk factors e.g.:
    - Interest rate changes
    - Fluctuating exchange rates
    - Other relevant financial factors
- Joint demo with Altera at Impact & Feedback
- Internally demonstrated results

Running
1 million iterations

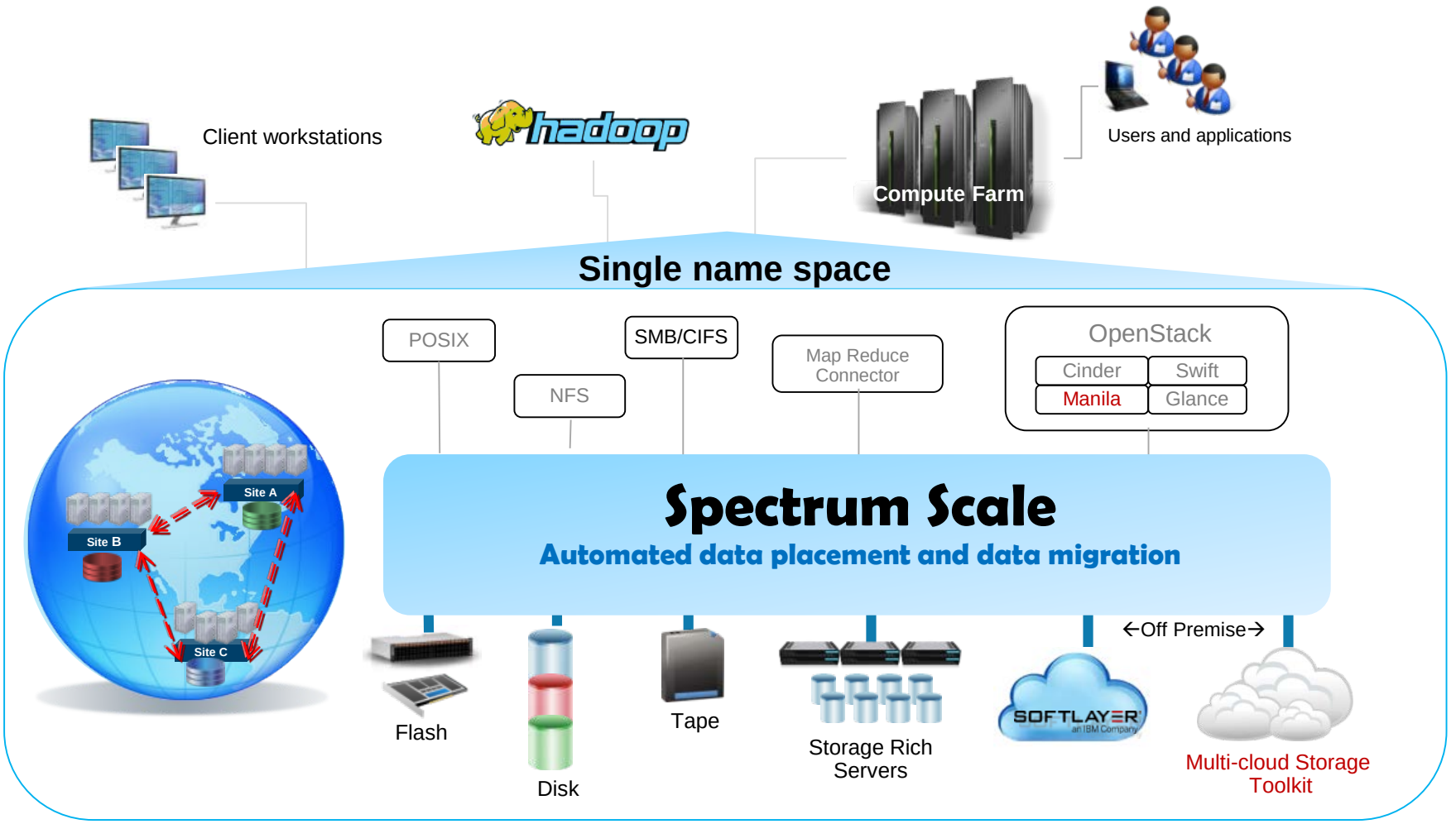
At least
250x Faster
with CAPI FPGA + POWER8 core

Cognitive Computing on Power8 + GPUs



- Cognitive system for the discovery of new Aluminum alloys
- New kind of application that combines Big Data & classic workloads
- Big data solutions (WDA) working with HPC back end to allow discovery
- C++ back end for compute of intensive kernel runs on P8+K80, in harmony with Big Data solutions (Cassandra + Titan Graph DB) on P8: high B/W + compute
- Answers in seconds

Unleash New Storage Economics on a Global Scale

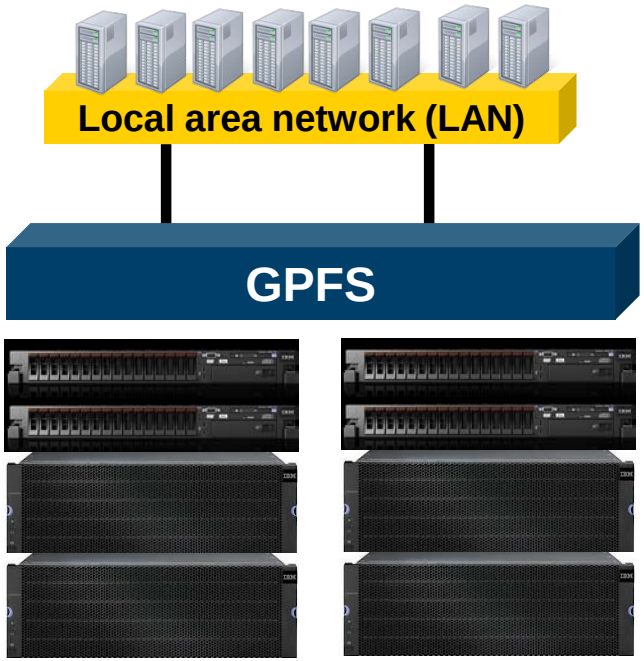


Note: Features market in red are planned for the future.

Elastic Storage Server at a glance

- Software RAID on the I/O Servers
  - SAS attached JBOD
  - Special JBOD storage drawer for very dense drive packing
  - Solid-state drives (SSDs) for metadata storage
  - No “storage controller”
- Deculstered RAID Fast rebuild
  - Much lower rebuild impact
  - Much longer MTDDL

Features
- Auto rebalancing
- Only 2% rebuild performance hit
- Reed Solomon erasure code, “8 data +3 parity”
- ~10^5 year MTDDL for 100-PB file system
- End-to-end, disk-to-GPFS-client data checksums



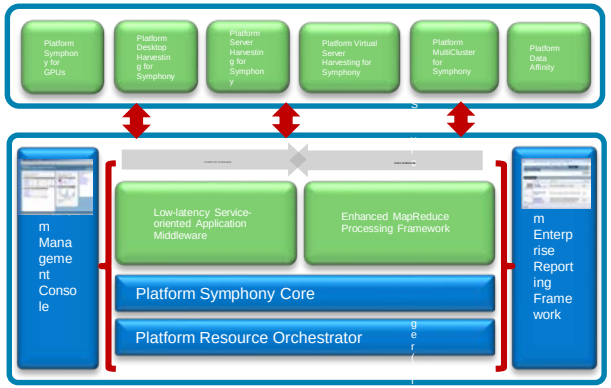
Leadship performance!

FileSystem Block Size	Write GB/Sec	Read GB/Sec
16 MB	18.337	29.481

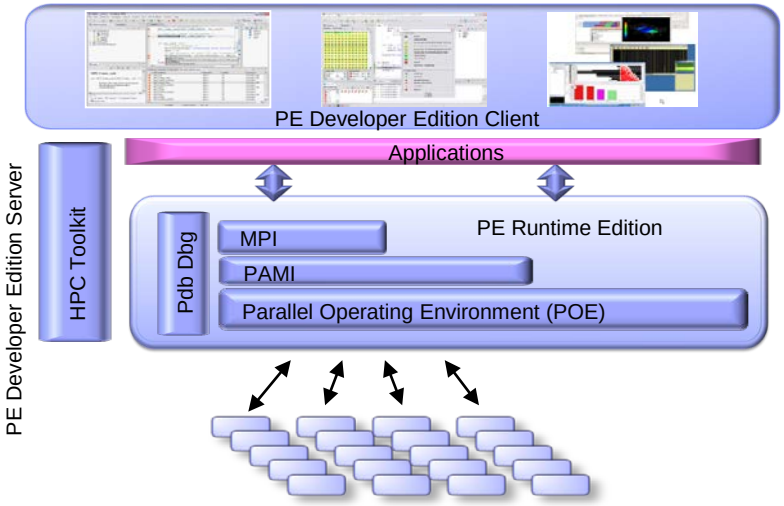
IBM Technical Computing Software Portfolio

Platform LSF
Workload Management

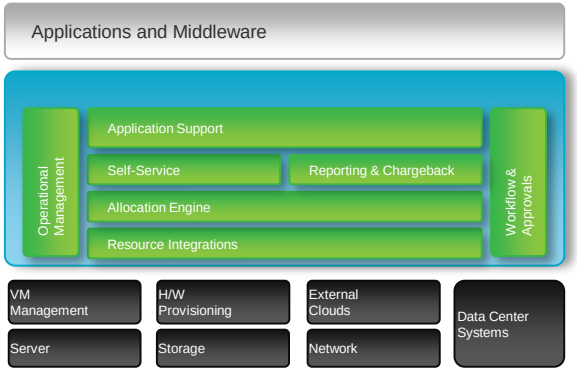
Platform Symphony
HPC Grid Services Management



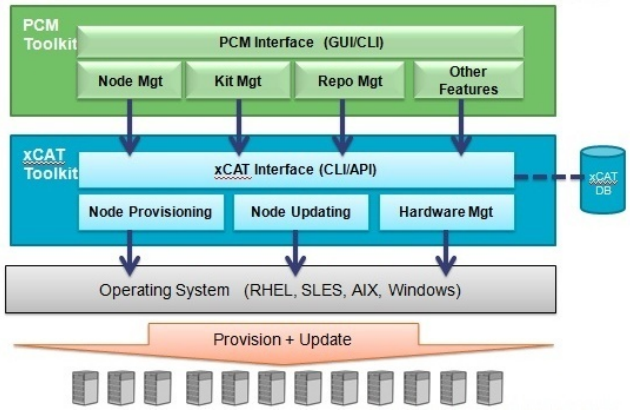
Parallel Environment (PE)



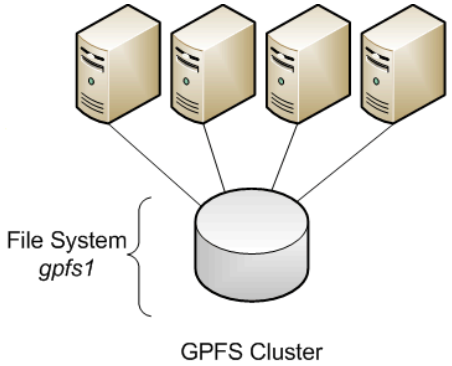
Platform Cluster Manager – Advanced Edition
Dynamic HPC Infrastructure Management



xCAT + Platform Cluster Manager
Systems Management and Provisioning

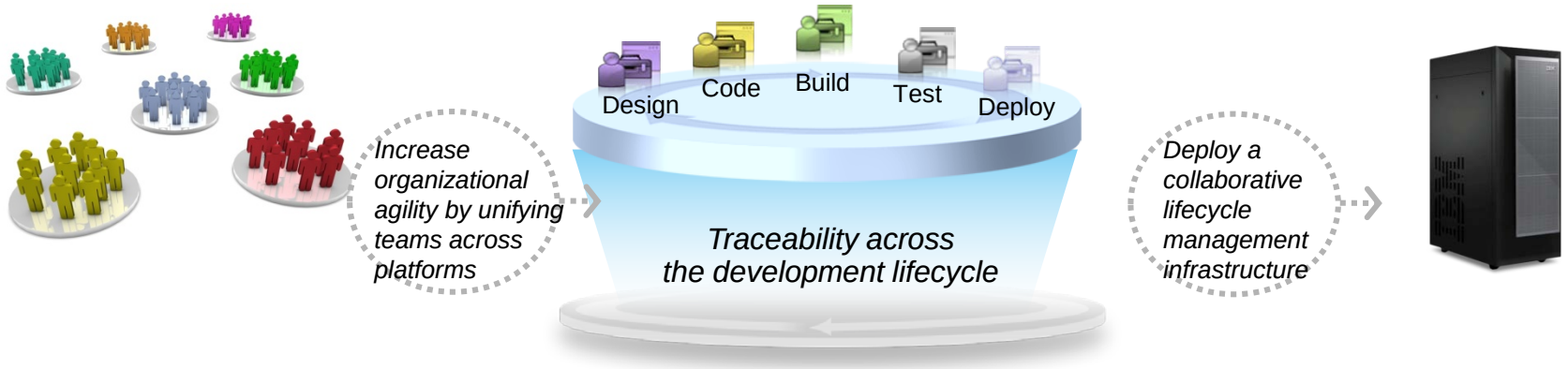


General Parallel File System (GPFS)
File System Management



Parallel Environment Developer Edition

Parallel Environment Developer Edition + Parallel Debugger for Design, Code, Build, Test



Key Features:

- High Function Eclipse Based Application Development Environment
  - C/C++/Fortran
  - Parallel Tools Platform
- Supports IBM PE RTE, Platform LSF
- Customers can incrementally add software capability into their environment

Advanced IBM Development Tools:

- IBM HPC Toolkit for analyzing performance of parallel and serial applications.
- XLC Compiler transformation report viewer
- IBM PAMI assistance tools

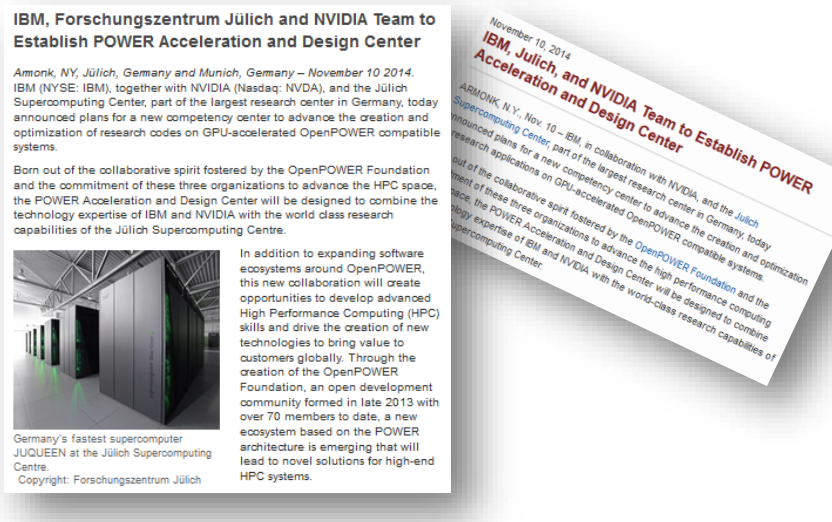
What's New:

- LSF scheduler/resource manager batch plug-in for PE DE Workbench IDE
- Ubuntu 14.04.1 Little Endian (Non Virtualized)
- Power8 Hardware Counters

POWER Acceleration & Design Centers (PADC)

- Germany
  - IBM R&D Labs in Böblingen and Zürich, Research Center Jülich, NVIDIA Europe

  Mission:
  - Suport scientists & engineers to target the grand challenges facing society using OpenPOWER technologies
  - Create competence and knowledge for Application and Technology developers



IBM, Forschungszentrum Jülich and NVIDIA Team to Establish POWER Acceleration and Design Center

Armonk, NY, Jülich, Germany and Munich, Germany – November 10 2014. IBM (NYSE: IBM), together with NVIDIA (Nasdaq: NVDA), and the Jülich Supercomputing Center, part of the largest research center in Germany, today announced plans for a new competency center to advance the creation and optimization of research codes on GPU-accelerated OpenPOWER compatible systems.

Born out of the collaborative spirit fostered by the OpenPOWER Foundation and the commitment of these three organizations to advance the HPC space, the POWER Acceleration and Design Center will be designed to combine the technology expertise of IBM and NVIDIA with the world class research capabilities of the Jülich Supercomputing Centre.

In addition to expanding software ecosystems around OpenPOWER, this new collaboration will create opportunities to develop advanced High Performance Computing (HPC) skills and drive the creation of new technologies to bring value to customers globally. Through the creation of the OpenPOWER Foundation, an open development community formed in late 2013 with over 70 members to date, a new ecosystem based on the POWER architecture is emerging that will lead to novel solutions for high-end HPC systems.

Germany's fastest supercomputer JUQUEEN at the Jülich Supercomputing Centre.
Copyright: Forschungszentrum Jülich

November 10, 2014

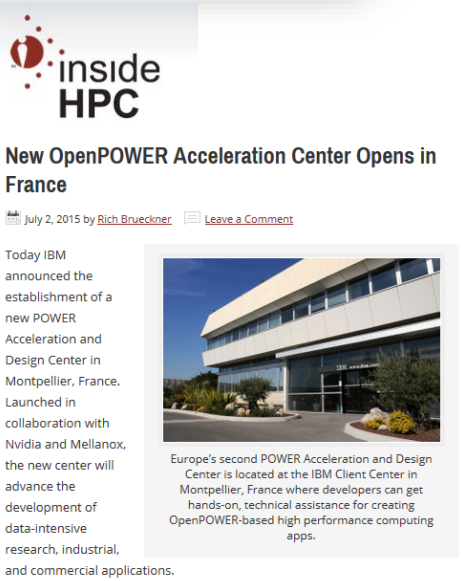
IBM, Julich, and NVIDIA Team to Establish POWER Acceleration and Design Center

ARMONK, N.Y., Nov. 10 – IBM, in collaboration with NVIDIA, and the Julich Supercomputing Center, part of the largest research center in Germany, today announced plans for a new competency center to advance the creation and optimization research applications on GPU-accelerated OpenPOWER compatible systems.

out of the collaborative spirit fostered by the OpenPOWER Foundation and the ment of these three organizations to advance the high performance computing ace, the POWER Acceleration and Design Center will be designed to combine logy expertise of IBM and NVIDIA with the world-class research capabilities of upercomputing Center.

- France
  - IBM Montpellier Client Center, IBM Zürich Research Lab, Mellanox, NVIDIA

  Mission:
  - Expand the software ecosystem around OpenPOWER
  - Increase computational performance and energy efficiency
  - Advance the development of data-intensive research, industrial, and commercial applications
  - Focus on direct customer porting and tuning



inside HPC

New OpenPOWER Acceleration Center Opens in France

July 2, 2015 by Rich Brueckner Leave a Comment

Today IBM announced the establishment of a new POWER Acceleration and Design Center in Montpellier, France. Launched in collaboration with Nvidia and Mellanox, the new center will advance the development of data-intensive research, industrial, and commercial applications.

Europe's second POWER Acceleration and Design Center is located at the IBM Client Center in Montpellier, France where developers can get hands-on, technical assistance for creating OpenPOWER-based high performance computing apps.

IBM Awarded $325M U.S. Department of Energy CORAL contracts

Two super computers for Oak Ridge and Lawrence Livermore Labs in 2017.

Current DOE Leadership Computers

Titan (ORNL) 2012 - 2017

Sequoia (LLNL) 2012 - 2017

Mira (ANL) 2012 - 2017

5X – 10X Higher Application Performance versus Current Systems

>100 PF, >2GB/core main memory, 800 GB/node local NVRAM, ~10MW
120 PB, 1 TB/s GPFS™ File System
Mellanox® Dual-Rail InfiniBand, IBM POWER® CPUs, NVIDIA® Volta™ GPUs

Discussion

ibm.com/technicalcomputing