IBM POWER8 HPC Review

April 12th, 2017

Jamie Syptak, Systems Architect, IBM

Harry Parks HPC|Data Analytics - Solution Sales

THE UNIVERSITY OF UTAH

Topics

- OpenPower … why
- Power8 Platform & Chip Design
- NVidia Tesla P100 GPU Deep Learning
- CAPI & Open CAPI
- Spectrum Scale & Components
- Power8 Solutions with Elastic Storage Sever
- Big Data Analytics - Hadoop
- Open Software Solutions and Application Performance

© 2016 IBM Corporation

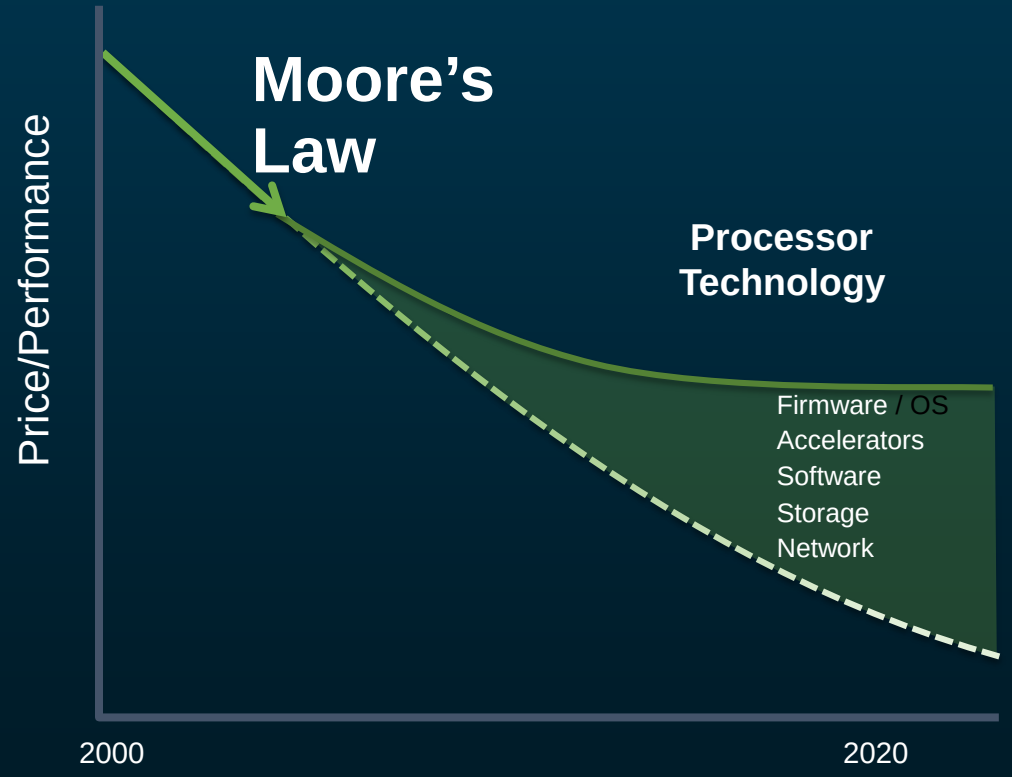
POWER – OpenPower …why

© 2016 IBM Corporation

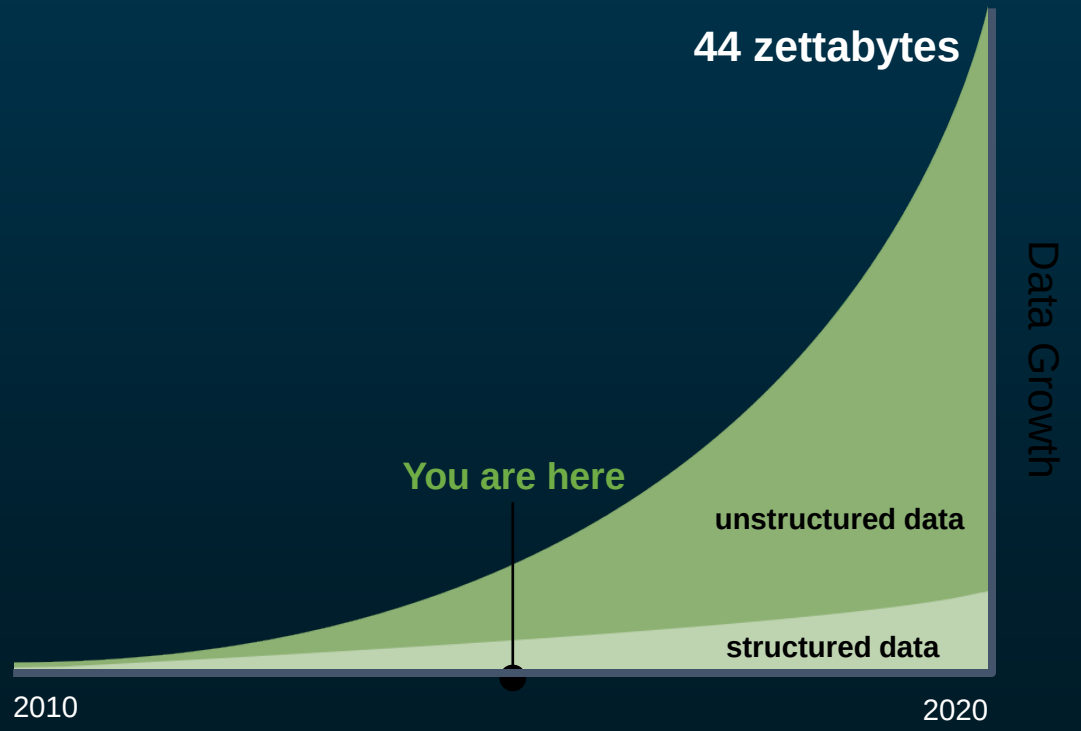
OpenPOWER™

Today's challenges demand innovation

Full system and stack open innovation required



Data holds competitive value



© 2016 OpenPOWER Foundation

OpenPOWER, a catalyst for Open Innovation

Market Shifts

- Moore's law no longer satisfies performance gain
- Growing workload demands
- Numerous IT consumption models
- Mature Open software ecosystem

OpenPOWER Strategy

- Vibrant ecosystem through open development
- Accelerated innovation through collaboration of partners
- Amplified capabilities driving industry performance leadership

Industry adoption, Open choice

Cloud Computing
Hyperscale & Large scal Datacenters

High Performance
Computing & Analytics

Domestic IT Agendas

OpenPOWER is an open development community,
using the POWER Architecture to serve the evolving needs of customers.

This is what a revolution looks like

© 2016 IBM Corporation

POWER – Power8 Platform & Chip Design

© 2016 IBM Corporation

Introducing the LC Portfolio of OpenPOWER servers

HPC

Data Intensive

Compute Intense

S812LC
MTM 8348-21C

- Storage rich single socket system for big data applications
- Memory Intensive workloads

NEW

S822LC For Big Data
MTM 8001-22C

- Ideal for storage-centric and high data through-put workloads
- Brings 2 POWER8 sockets for Big Data workloads
- Big data acceleration with work CAPI and GPUs

NEW

S822LC For High Performance Computing
MTM 8335-GTB

- Introducing CPU-GPU NVLink, delivering >2.5X the bandwidth to GPUs
- POWER8 with NVIDIA NVLink
- Up to 4 integrated NVIDIA "Pascal" GPUs

NEW

S821LC
MTM 8001-22C

- 2 POWER8 sockets in a 1U form factor
- Ideal for environments requiring dense computing

S822LC
MTM 8335-GCA

- Up to 2.2X memory bandwidth of Intel x86 systems
- Memory Intensive workloads

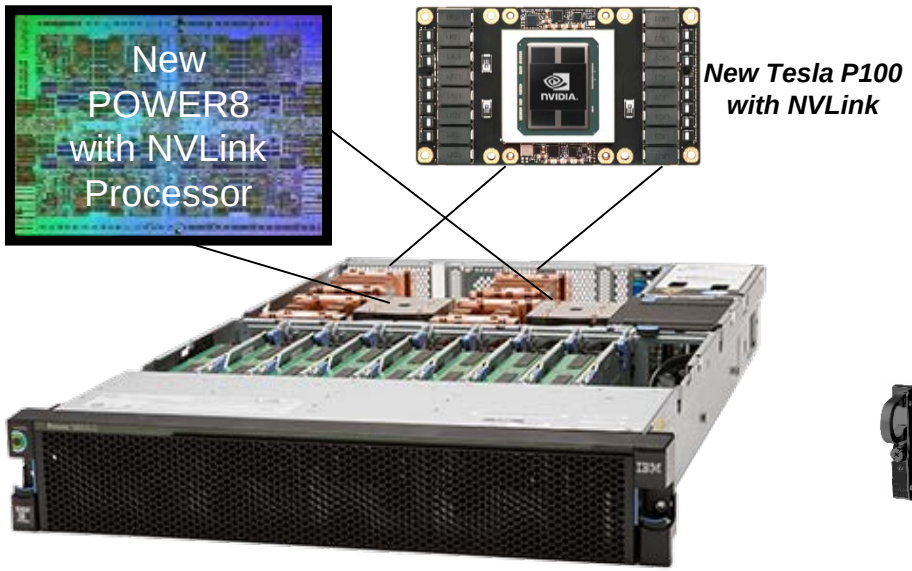
IBM Power Systems LC Portfolio

Systems designed to take **Data Rich** and **High Performance Computing (HPC)** Linux workloads *to the next level*

© 2016 IBM Corporation

Current Power Linux Servers

S822LC for High Performance Computing



System Details

- 2-socket, 2U
- Up to 20 cores (2.86-3.26Ghz)
- 1 TB Memory (32 DIMMs)
- 230GB/sec memory bandwidth
- 2x SFF (HDD/SSD), SATA
- Up to 4 integrated NVIDIA Tesla P100 GPUs
- 3 PCIe slots, 3 CAPI enabled, IB Add-in
- Air or water cooled

S822LC for Big Data

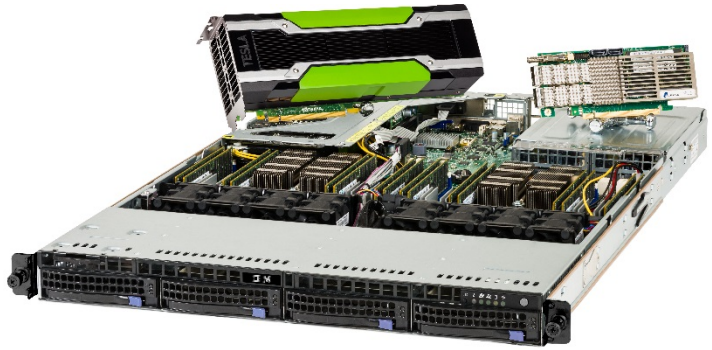
K80 or Tesla P100

System Details

- 2-socket, 2U
- Up to 20 cores (2.9-3.3Ghz)
- 512 GB Memory (16 DIMMs)
- 115GB/sec memory bandwidth
- 12 SFF/LFF (HDD/SSD) 96 TB storage
- 5 PCIe slots, 4 CAPI enabled
- 2 NVIDIA PCIe GPU capable

S821LC

K80 or Tesla P100

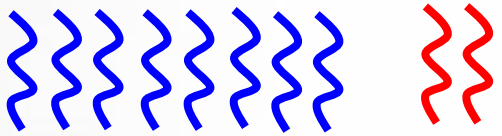


System Details

- 2 socket, 1U
- Up to 20 cores (2.09-2.32Ghz)
- 512 GB Memory (16 DIMMs)
- 115 GB/sec memory bandwidth
- 4 SFF/LFF (HDD/SSD), 32 TB Storage
- 4 PCIe slots, 3 CAPI enabled
- 1 NVIDIA PCIe GPU capable

IBM Innovates with POWER8: Breakthrough performance for YOUR data

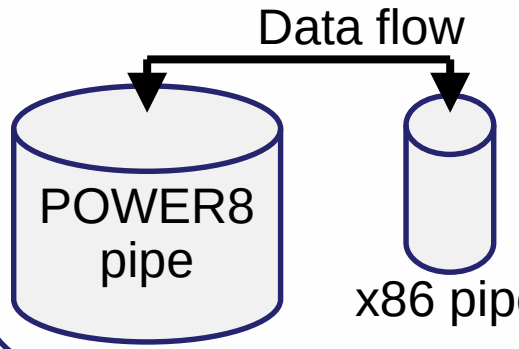
4X
Threads per core*



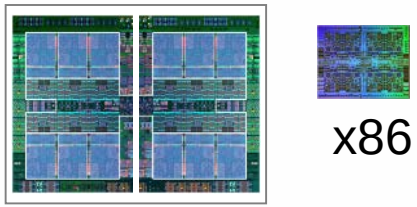
POWER8 SMT8 | x86 Hyperthread

Parallel Processing

4X
Mem. Bandwidth*

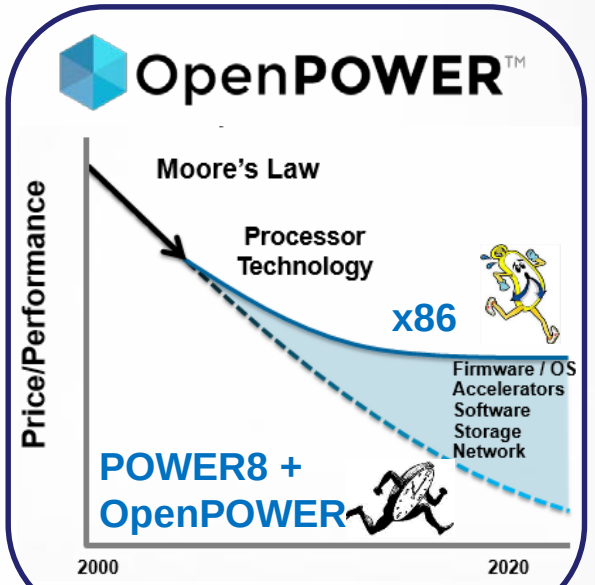


4X
More cache*



POWER8

x86



These design decisions result in best performance for data centric workloads like:

Database, NoSQL, Big Data Analytics, OLTP

SMT=Simultaneous Multi-Threading
OLTP = On-Line Transaction Processing

© 2016 IBM Corporation

Power 8 Internal Speeds –Important for Big Data

Processor Comparisons:	(Best Version) Haswell	POWER8	Comments
SMT / Core (Threads)	2T	8T	4X
L1 DCache / Core	32 KB	64 KB	2X
L2 Cache / Core	256 KB	512 KB	2X
L3 Cache / Processor	16 - 45 MB	80 - 96 MB	2.1X – 6.0X
L4 Cache / System	None	64 MB – 2 GB	Helps Big Data
L2 Cache Bw / Core	116 GB/s	320 GB/s	2.76X
Maximum “Sustained” Mem Bw	53 GB/s	224 GB/s	4.2X
STAC-A2 Greeks / Core	121/ms	38/ms	3.2X
STAC-A2 Assets / Core	3.125	1.667	1.9X
STAC-A2 Paths / Core	1.17	0.22	5.3X
Transactional Memory	No	Yes	Helps Scalability
PCI-Exp Gen3 / Processor	16x (32 GB/s)	48x (96 GB/s)	3X
Native/Custom Engine Accelerator	No	CAPI/FPGA	Helps Analytics
Large “In Memory” Flash	No	CAPI/Flash	Helps Big Data

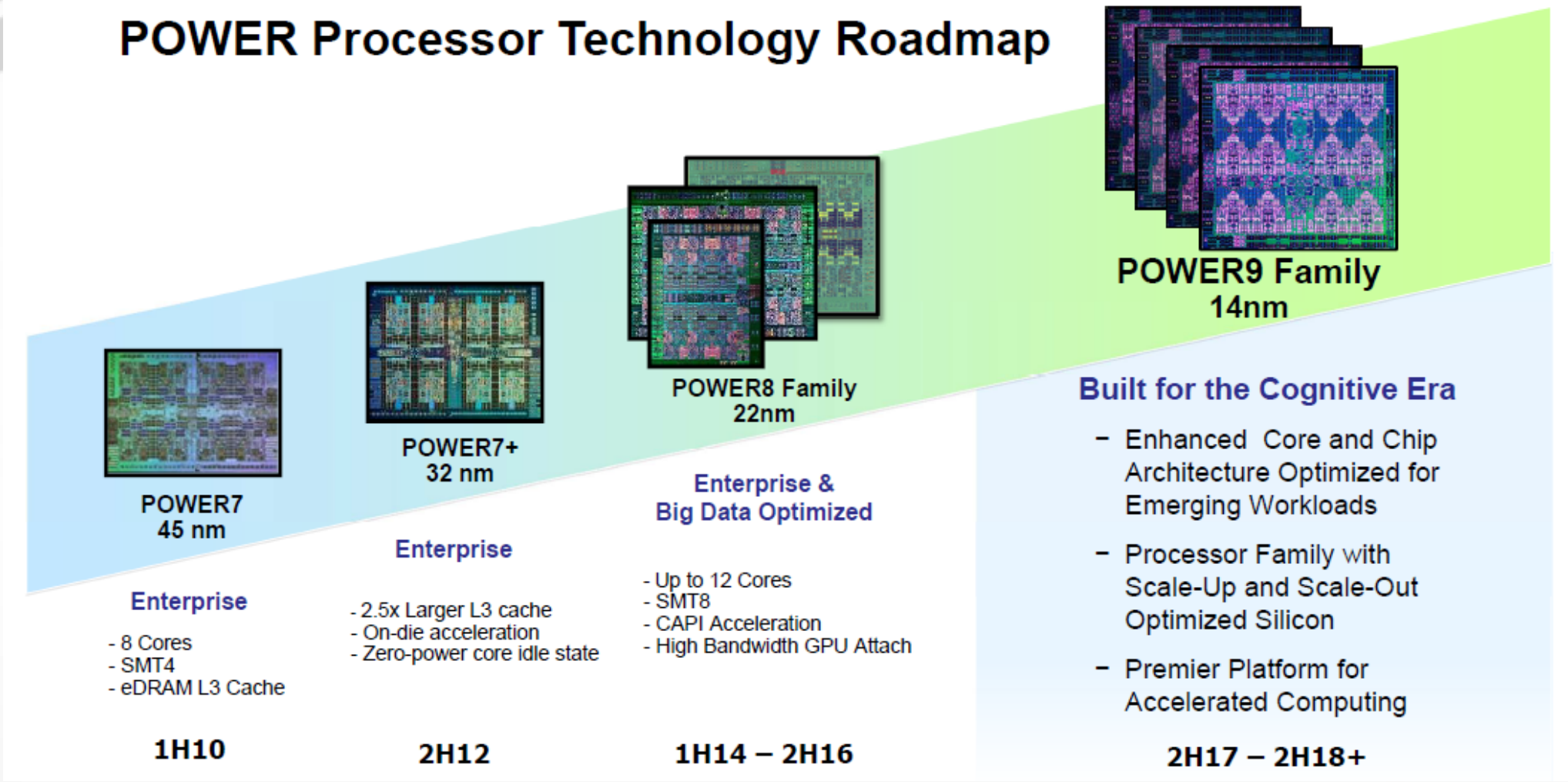
POWER8 Core is ~2X to 3X of a Haswell Core!

POWER8 Provides a “New Class” of Capabilities for Attacking Analytics, Big Data, HPC & Cloud Workloads!

© 2016 IBM Corporation

Return to top

Hot Chips 2017 – Roadmap



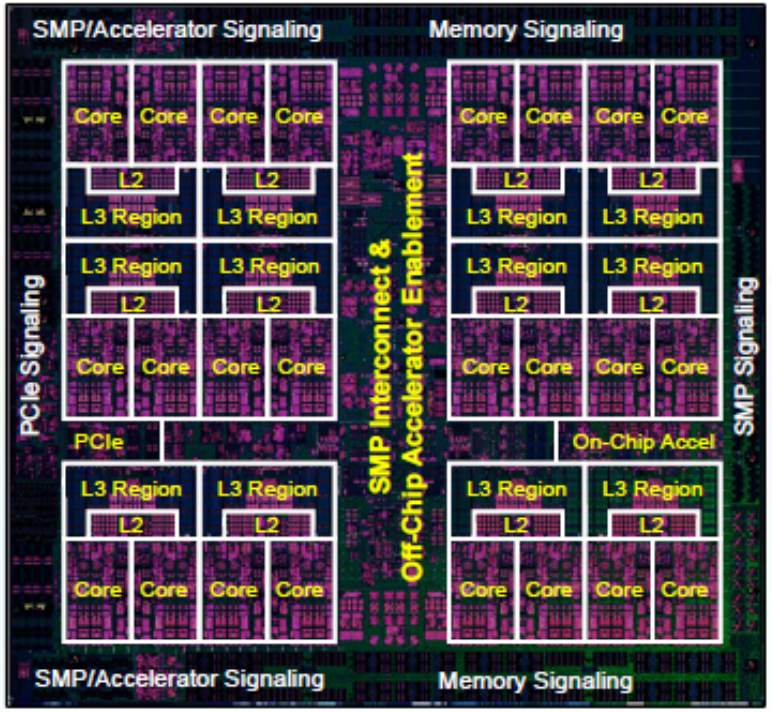
© 2016 IBM Corporation

Hot Chips 2016 – POWER9 Processor Common Features

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
- Workload optimized frequency
- Hardware enforced trusted execution



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, advanced new features
- CAPI 2.0: Coherent accelerator and storage attach (PCIe G4)
- New CAPI: Improved latency and bandwidth, open interface

State of the Art I/O Subsystem
- PCIe Gen4 – 48 lanes

High Bandwidth Signaling Technology
- 16 Gb/s interface
  - Local SMP
- 25 Gb/s interface – 25G Link
  - Accelerator, remote SMP

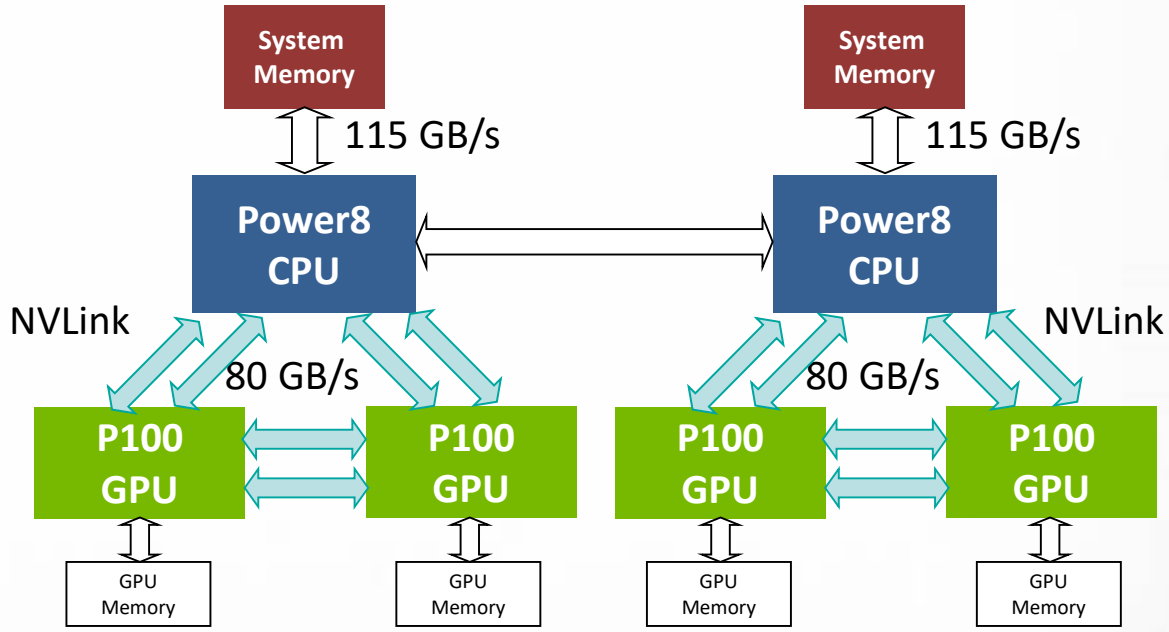
© 2016 IBM Corporation

POWER – NVidia with NVlink Tesla P100 GPU – Deep Learning Machine Learning

© 2016 IBM Corporation

POWER8+P100+NVLink for increases system bandwidth

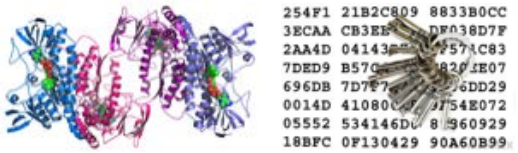
- NVLink between CPUs and GPUs enables fast memory access to large data sets in system memory
- Two NVLink connections between each GPU and CPU-GPU leads to faster data exchange
- First to market: volume shipments starting September, 2016



© 2016 IBM Corporation

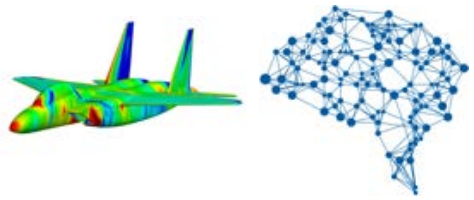
Why it Matters:
Use Cases where NVLink will have the most Impact

Stream Data at Same Rate as Computation



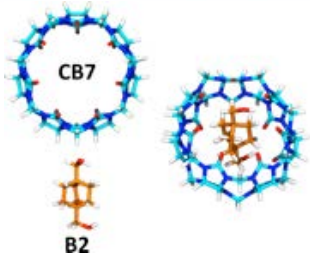
Genomics, Cryptography, Video Processing, etc.

Burst Data at Startup and Teardown



CFD/CAE, Machine Learning, Deep Learning, etc.

Constant Data Transfers between adjacent GPUs



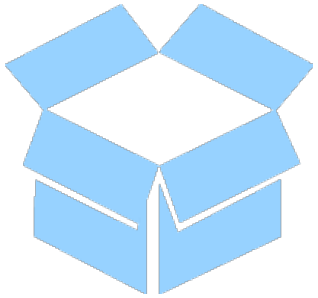
Molecular Dynamics, Amber, etc.

Mask Bus Transfers from Host-Device

Accelerated Databases, Analytics, etc.

Introducing PowerAI: Get Started Fast with Deep Learning

IBM



Package of Pre-Compiled Major Deep Learning Frameworks

Easy to install & get started with Deep Learning with Enterprise-Class Support

Optimized for Performance To Take Advantage of NVLink

Enabled by High Performance Computing Infrastructure

Simplify Access and Installation

- Tested, binary builds of common Deep Learning frameworks for ease of implementation
- Simple, complete installation process documented on IBM OpenPOWER
  - http://openpowerfoundation.org/blogs/ and search Deep Learning
- Future focus on optimizing specific packages for POWER: NVIDIA Caffe, TensorFlow, and Torch

Caffe

TensorFlow

torch

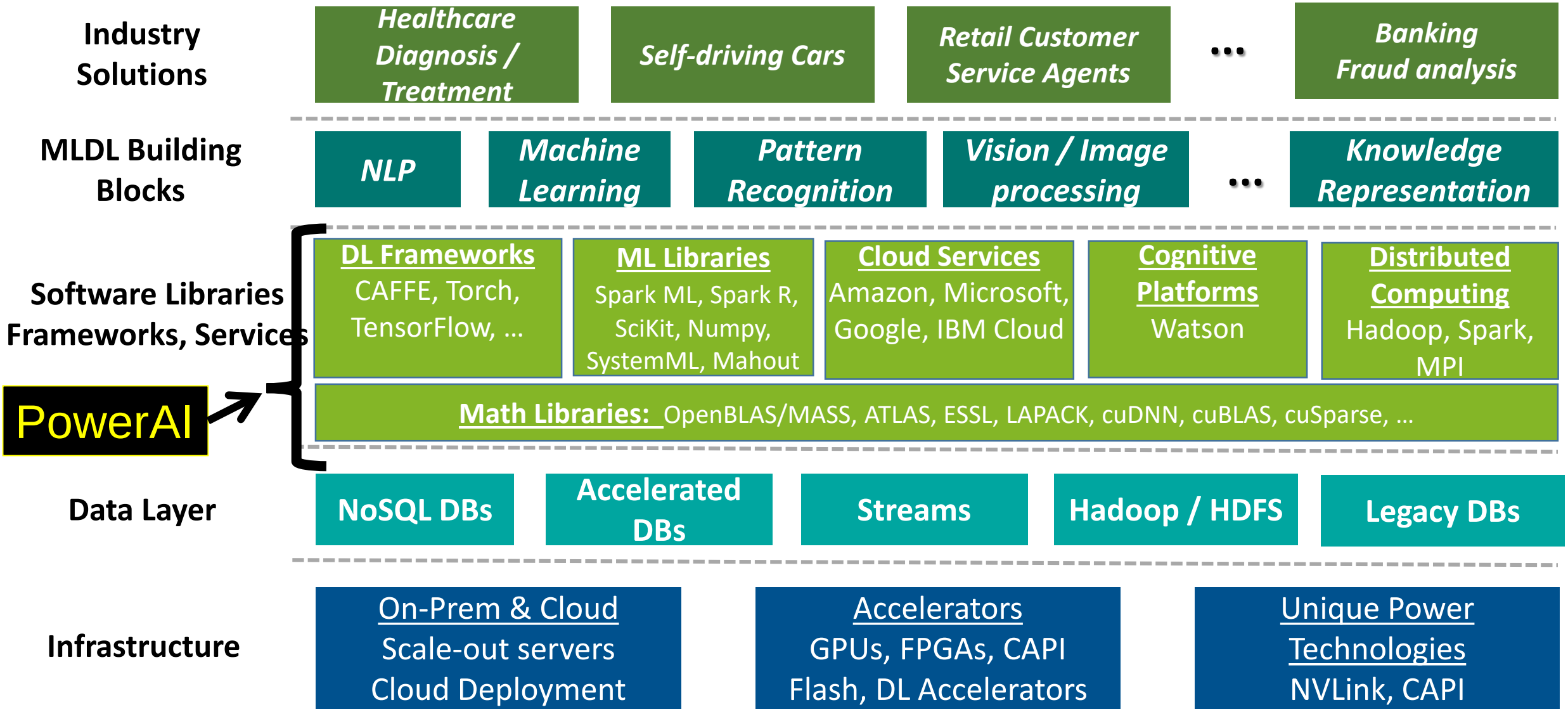
- Additional docs… DIY…

https://www.ibm.com/developerworks/community/blogs/home/search?q=%22Deep+Learning+on+OpenPOWER%22&t=entry&f=all&maxresults=50&sortby=0&order=desc&lang=en

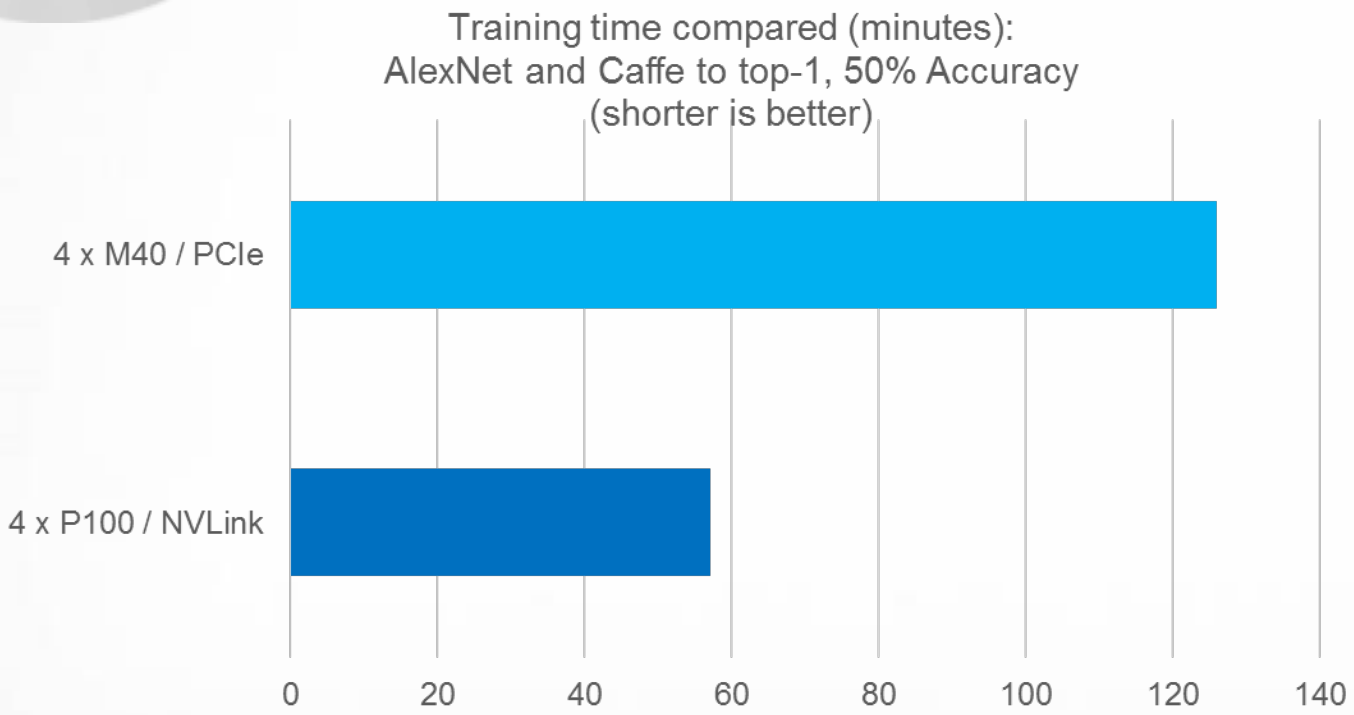
	Already ported	Future focus
OS	Ubuntu 14.04	Ubuntu 16.04
CUDA	7.5	8.0
cuDNN	5.1	5.1
Built w/ MASS	Yes	Yes
OpenBLAS	0.2.18	Optimize
Caffe	1.0 rc3	
NVIDIA Caffe	0.14.5	Optimize
NVIDIA DIGITS	3.2	
Torch	7	Optimize
Theano	0.8.2	
TensorFlow	0.9	Optimize
CNTK	Nov 2015(*)	
DL4J	0.5.0(*)	
Chainer		
GPU	2x K80	4 x P100
Base System	822LC	Minsky

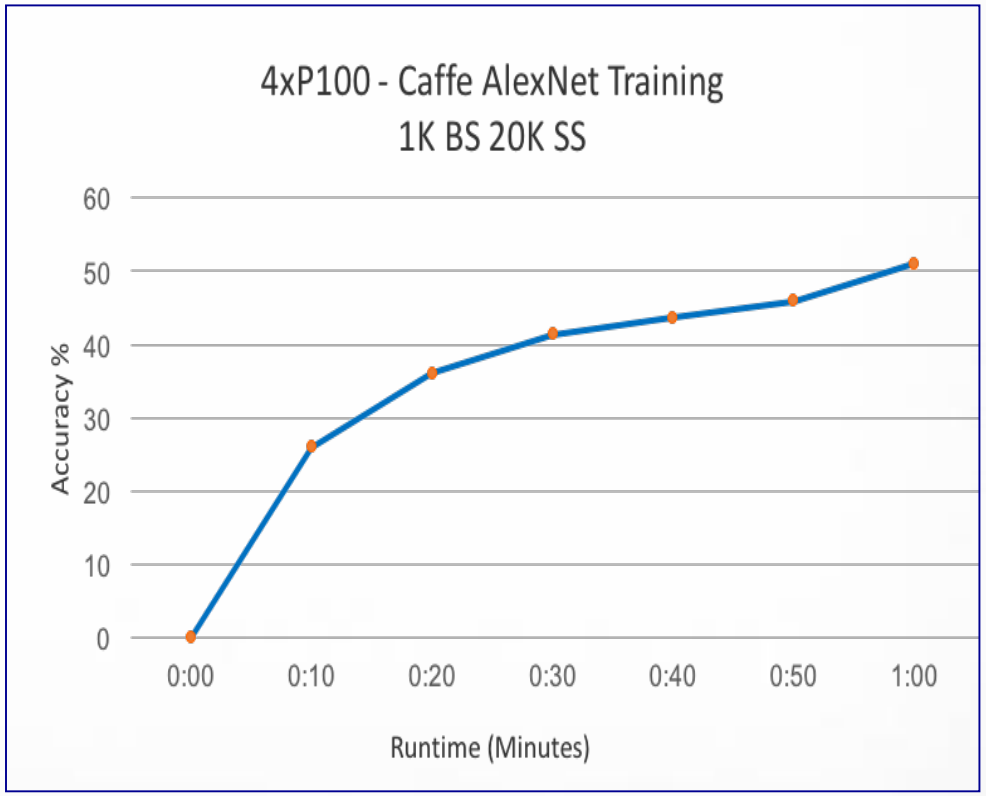
© 2016 IBM Corporation

Machine Learning / Deep Learning Software Stack



Improve performance: 2.2X faster training time





AlexNet Trained in Under 1 Hour (57 mins)

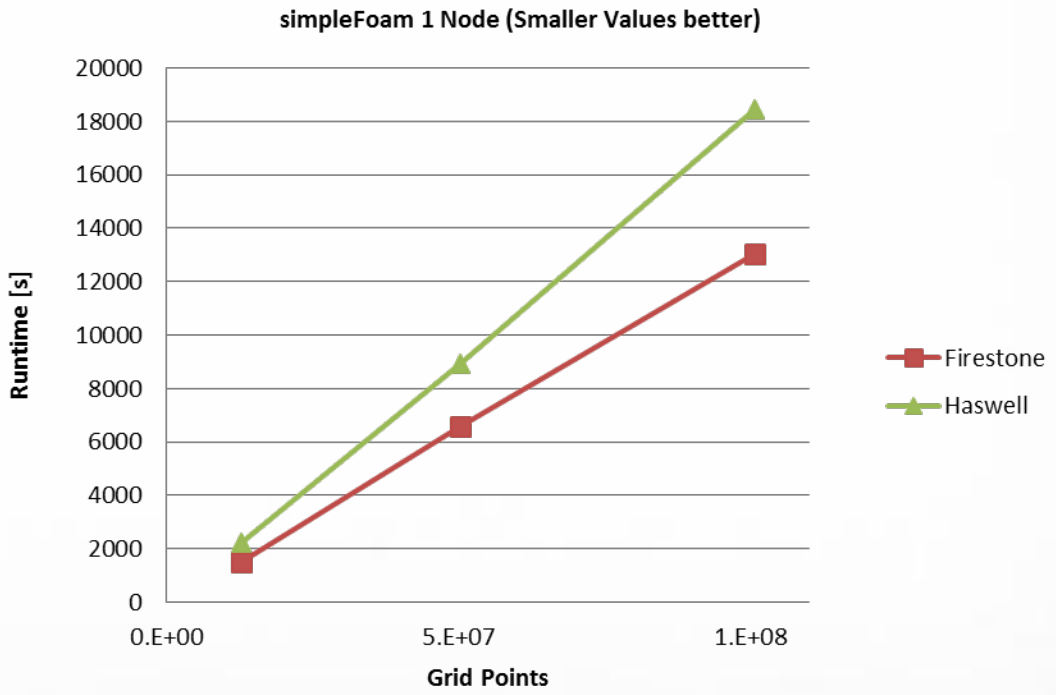
© 2016 IBM Corporation

Realize CFD results 40% faster on OpenFOAM on IBM Power System S822LC compared to Xeon E5-2600v3 Systems

Iterate faster through improved time to solution

- IBM Power S822LC delivers 1.4x the OpenFOAM performance from a superior processor design
  - Higher Memory bandwidth
  - Increased L3 cache
  - SMT (simultaneous multi-threading)

- Moreover, IBM Power S822LC delivers exceptional throughput on the *largest* meshes

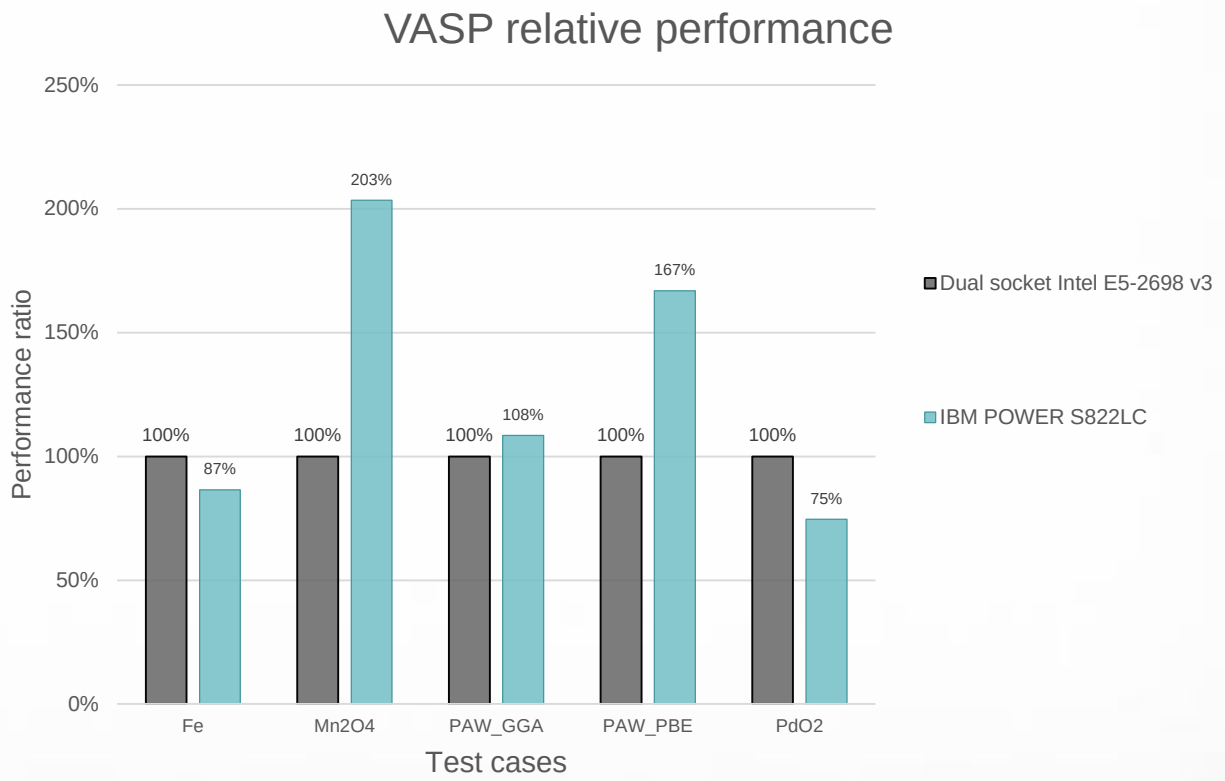


- Results are based on IBM internal testing of systems running OpenFOAM version 2.3.0 code benchmarked on POWER8 systems.. Individual results will vary depending on individual workloads, configurations and conditions.
- IBM Power System S822LC, POWER8; 3.5 GHz, 512 GB memory; 2x 10 core processors/ 4 threads per core
- Bull R424-E4, Intel E5-2680v3, 2.5 GHz, 128 GB memory, 2x 12 core processors / 1 thread per core

© 2016 IBM Corporation

The POWER8 processor is designed for Big Data and HPC, with up to 2X the performance on VASP

- Results based on five standard test cases
  - CPU-only
- Memory bandwidth helps with application performance



- Results are based on IBM & BSC internal testing of systems. Individual results will vary depending on individual workloads, configurations and conditions.
- IBM Power System S822LC; 10 cores, POWER8; 2.93 GHz, 256 GB memory
- Intel Xeon data is based on IBM & BSC internal measurements. 16 cores Intel Xeon E5-2698 v3, 2.3GHz, 256 GB

System	#MPI	#OMP
Dual socket Intel	32	1
IBM Power System S822LC	20	1

© 2016 IBM Corporation
IBM Confidential

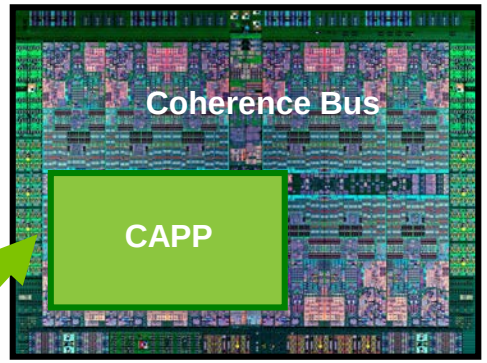
POWER – CAPI & Open CAPI

© 2016 IBM Corporation

Power 8 CAPI – Coherent Accelerator Processor Interface

- Virtual Addressing
  - Accelerator can work with same memory addresses that the processors use
- Hardware Managed Cache Coherence
  - Enables the accelerator to participate in “Locks” as a normal thread Lowers Latency over IO communication model

POWER8



PCIe Gen 3
Transport for encapsulated messages

Processor Service Layer (PSL)

- Present robust, durable interfaces to applications
- Offload complexity / content from CAPP

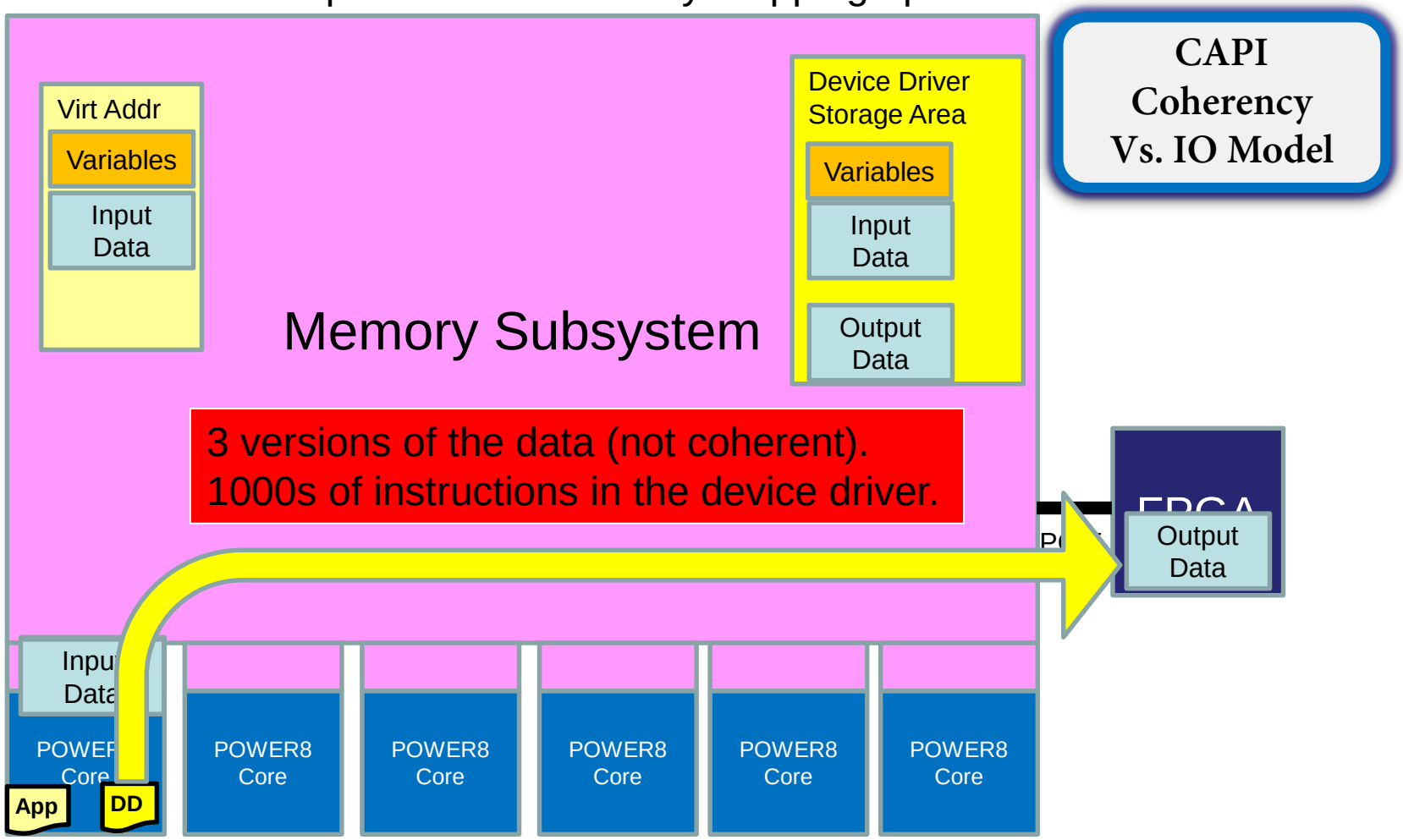
Customizable Hardware Application Accelerator

- Specific system SW, middleware, or user application
- Written to durable interface provided by PSL

© 2016 IBM Corporation

What was done before CAPI?

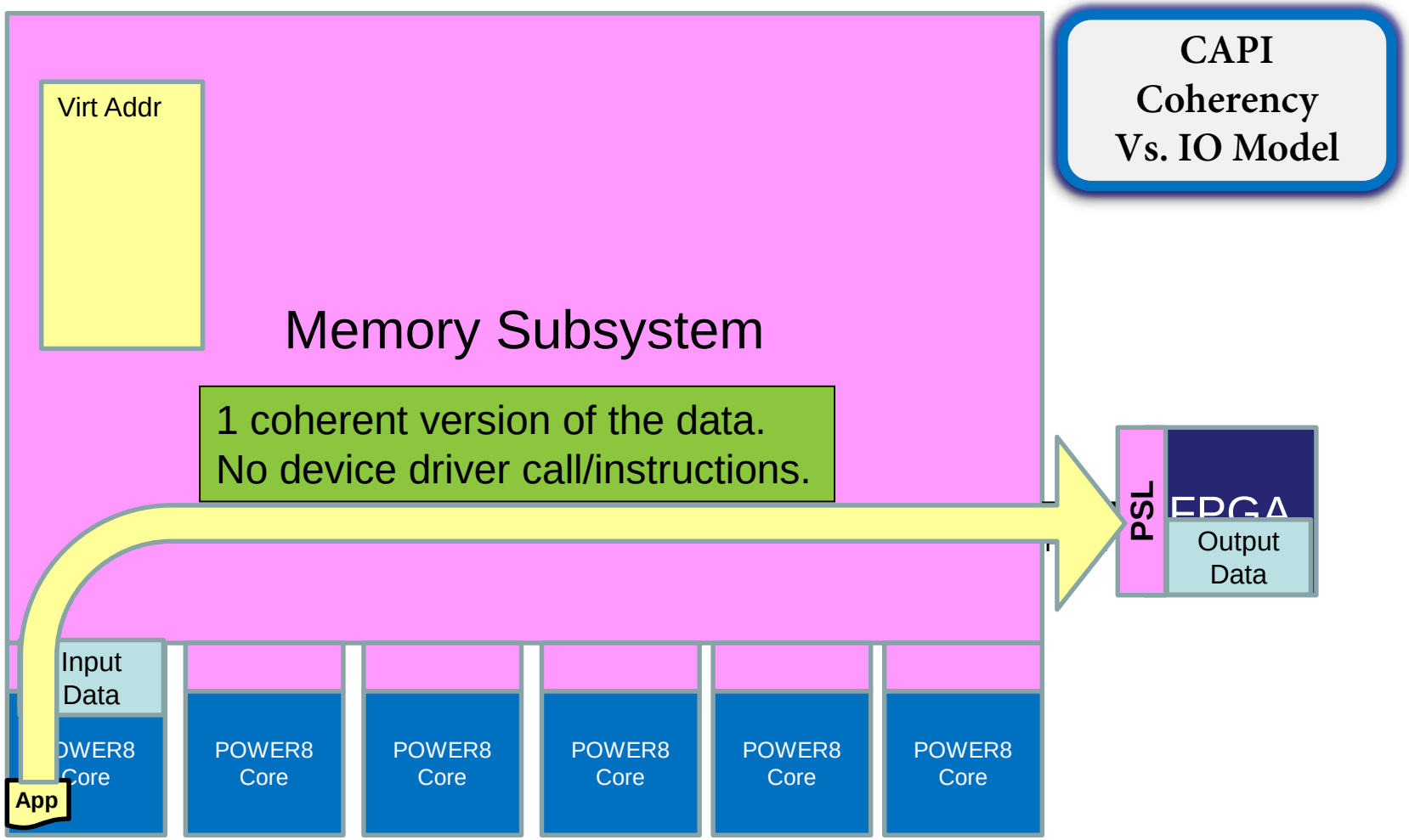
Prior to CAPI, an application called a device driver to utilize an FPGA Accelerator.
The device driver performed a memory mapping operation.



© 2016 IBM Corporation

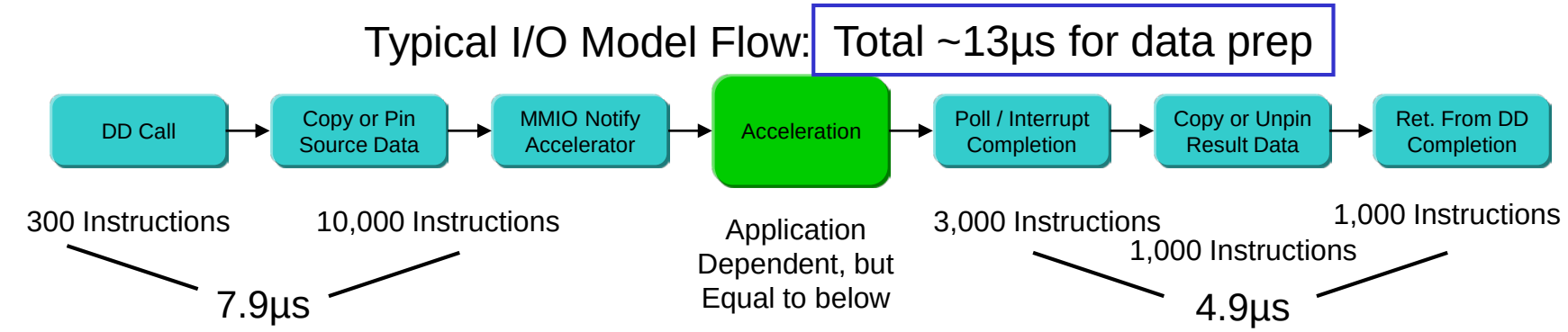
CAPI Coherency

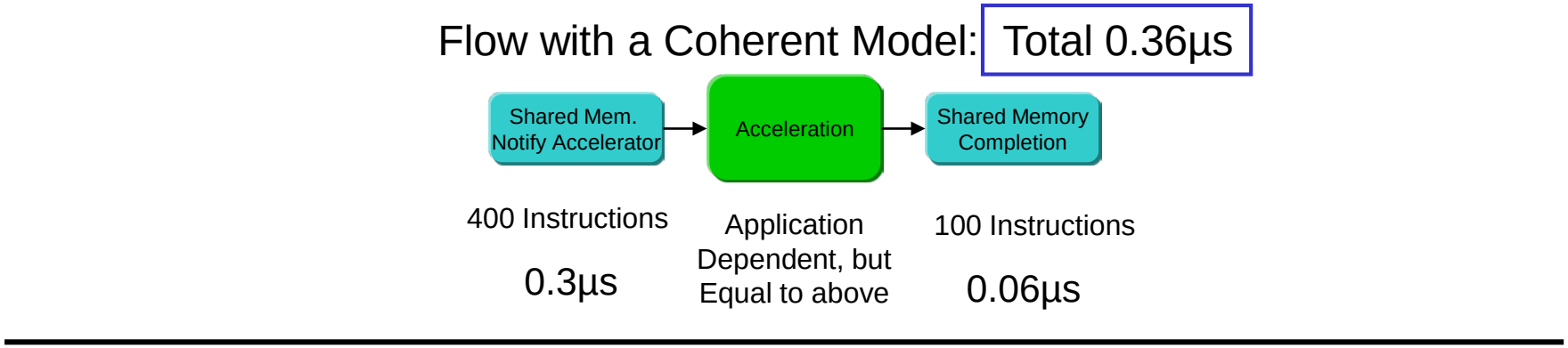
With CAPI, the FPGA shares memory with the cores



© 2016 IBM Corporation

CAPI vs. I/O Device Driver: Data Prep



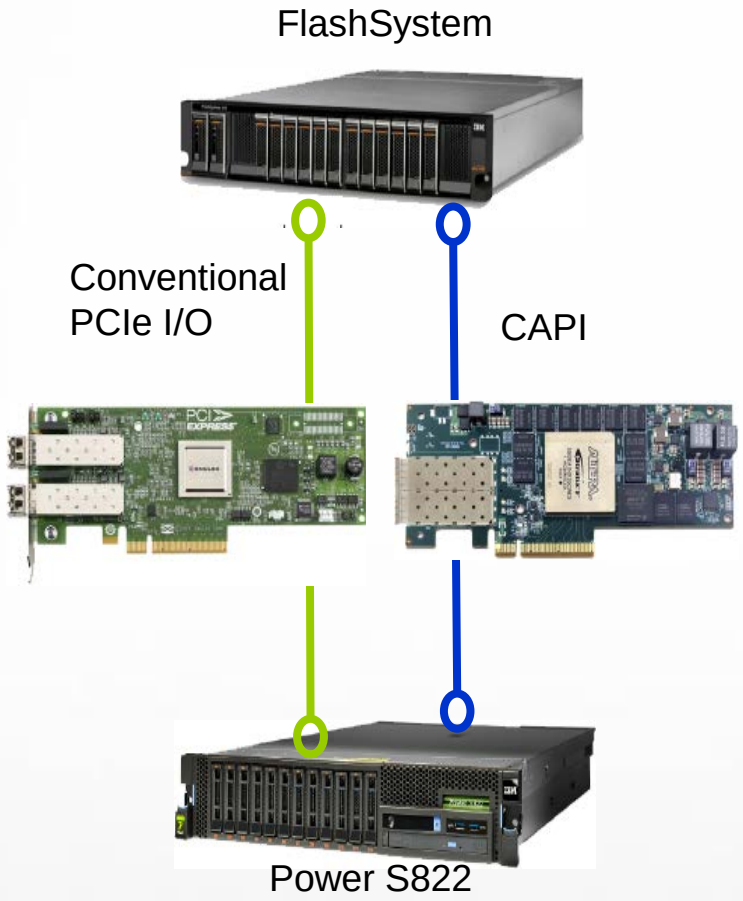


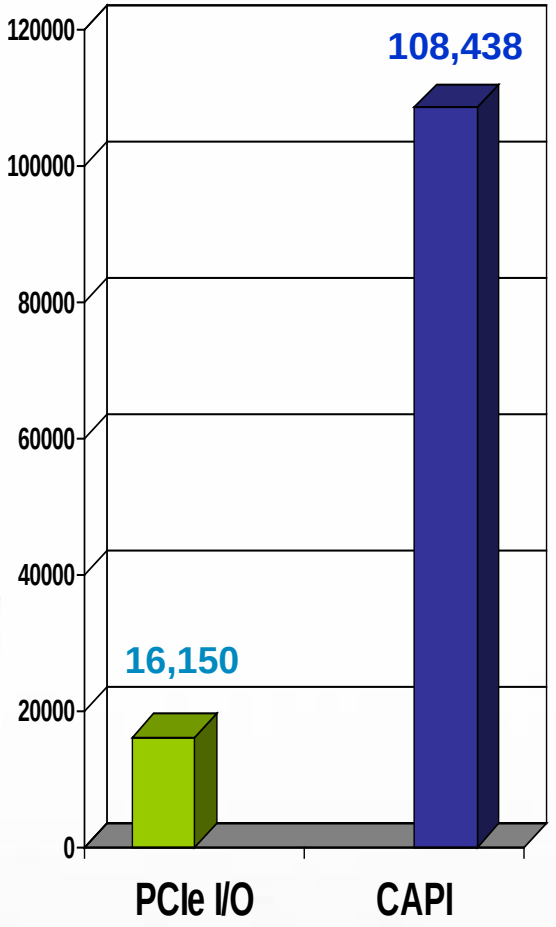
© 2016 IBM Corporation

Storage Access Acceleration

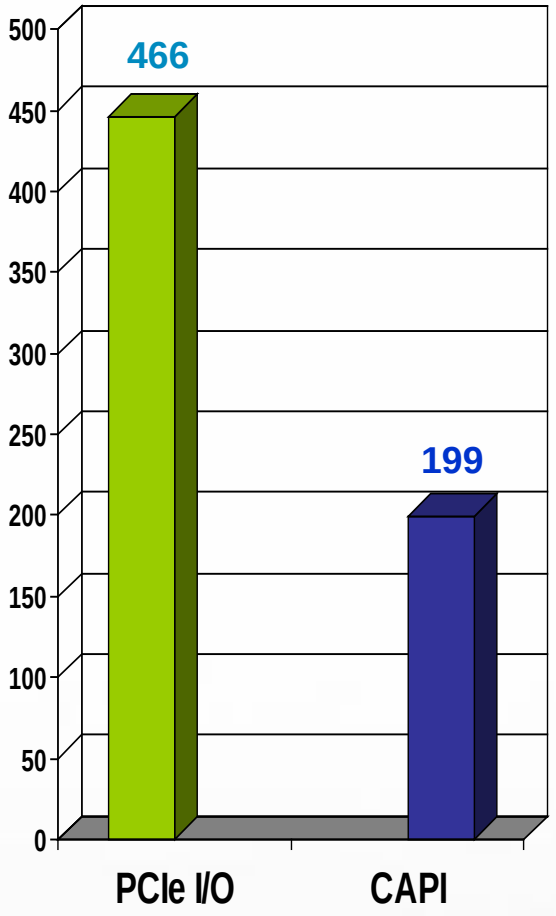
Demonstrating the Value of CAPI Attachment

Identical hardware with 2 different paths to data





IOPs per HW Thread



Latency (us)

© 2016 IBM Corporation

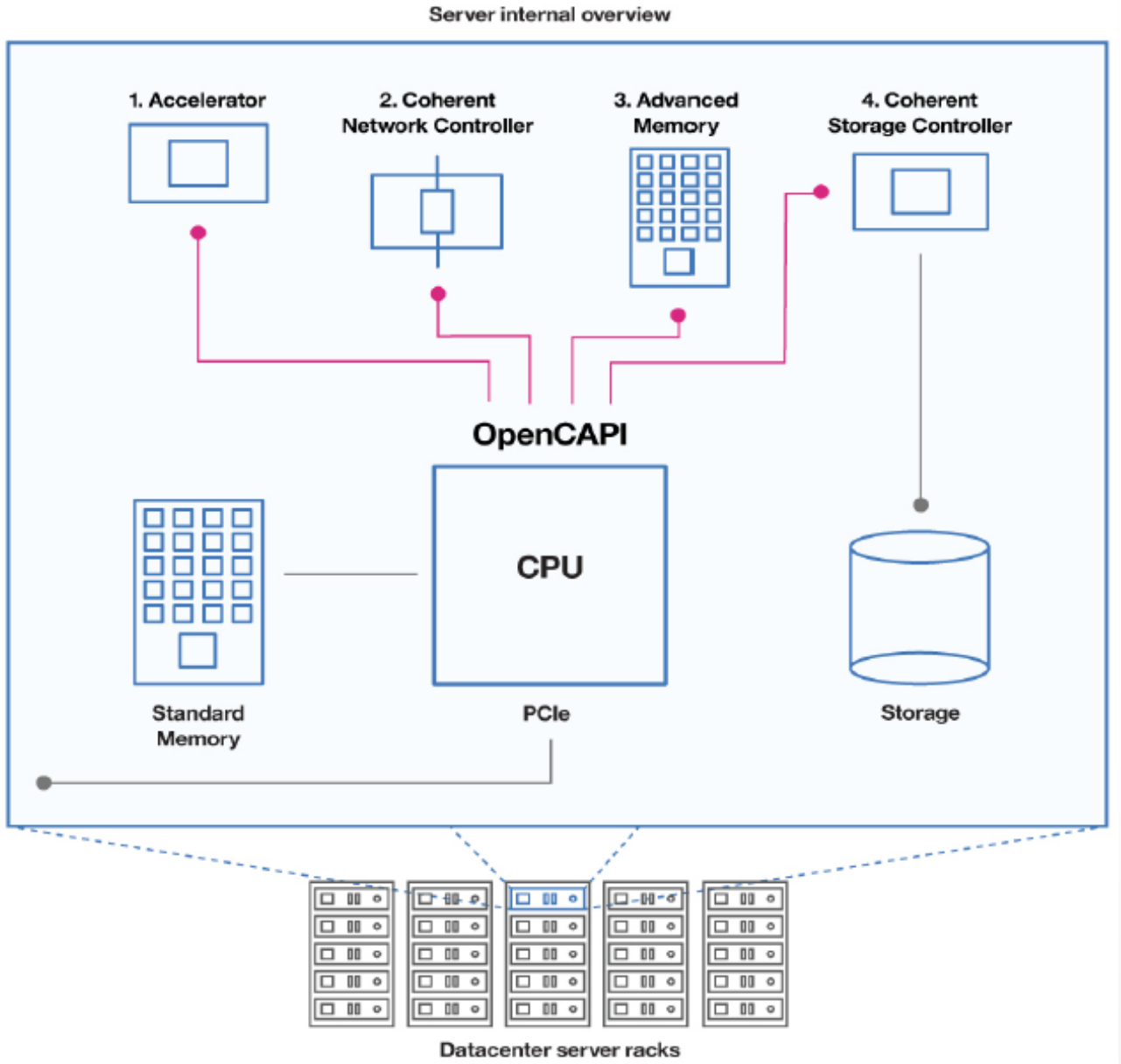
OpenCAPI

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.



© 2016 IBM Corporation

POWER – Spectrum Scale & Components

© 2016 IBM Corporation

IBM Software Defined Infrastructure
Multi-scale Infrastructure for High Performance Computing & Analytics

High Performance Computing
Design / Simulation / Modeling

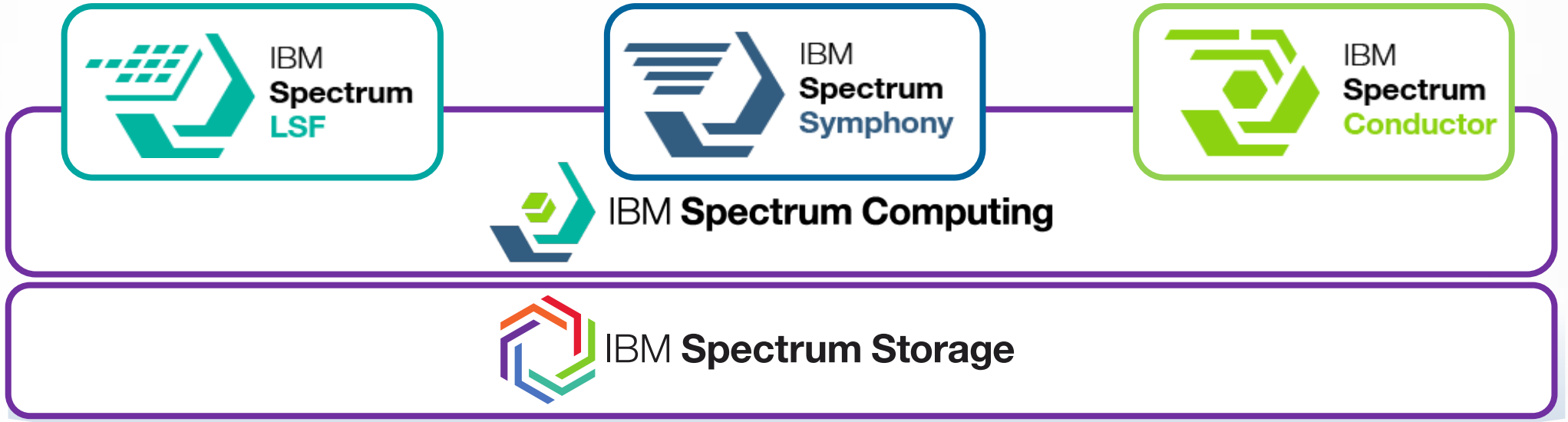
High Performance Analytics
Trade / Risk Analytics

'New-gen Workloads'
Hadoop, Spark, Containers

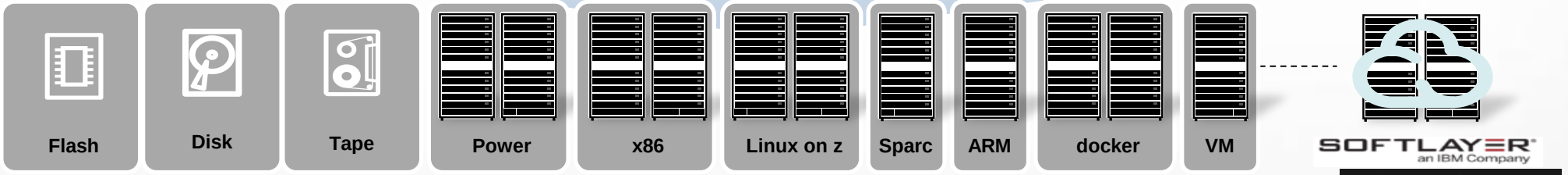
Workload Aware Scheduling

Shared Resource Management

Shared Multi-tier Data Management

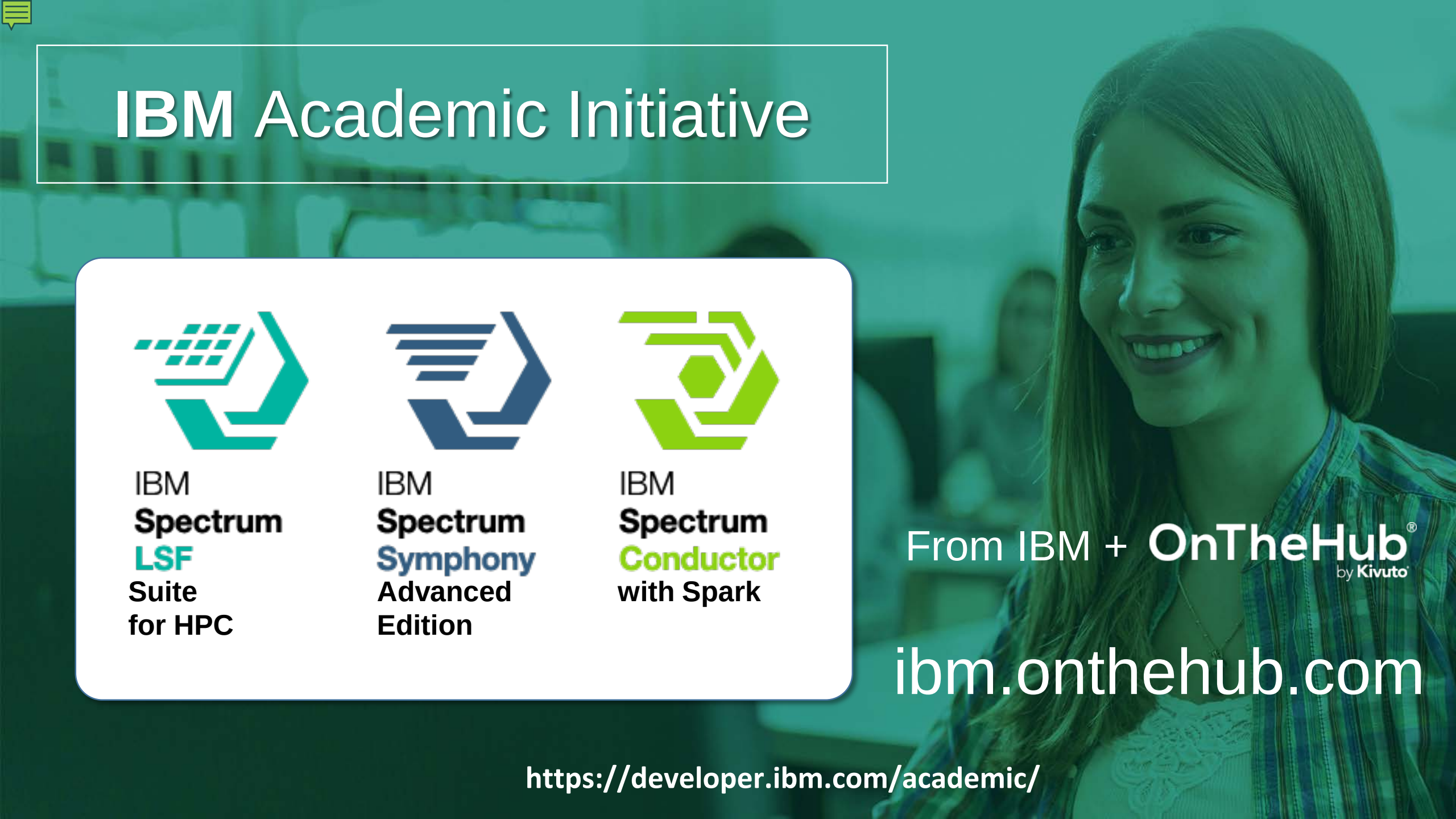


Heterogeneous Servers & Storage



Hybrid Cloud Infrastructure

IBM Academic Initiative



IBM Spectrum LSF Suite for HPC

IBM Spectrum Symphony Advanced Edition

IBM Spectrum Conductor with Spark

From IBM + OnTheHub® by Kivuto

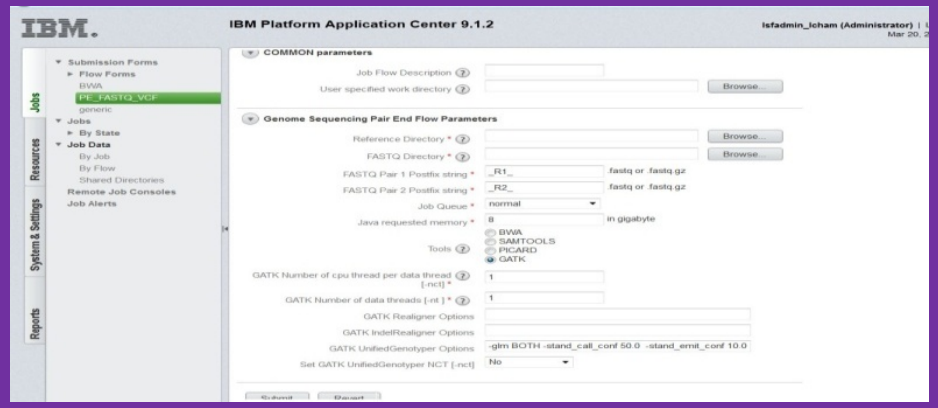
ibm.onthehub.com

https://developer.ibm.com/academic/

IBM Spectrum LSF Application Center, IBM Spectrum LSF Process Manager

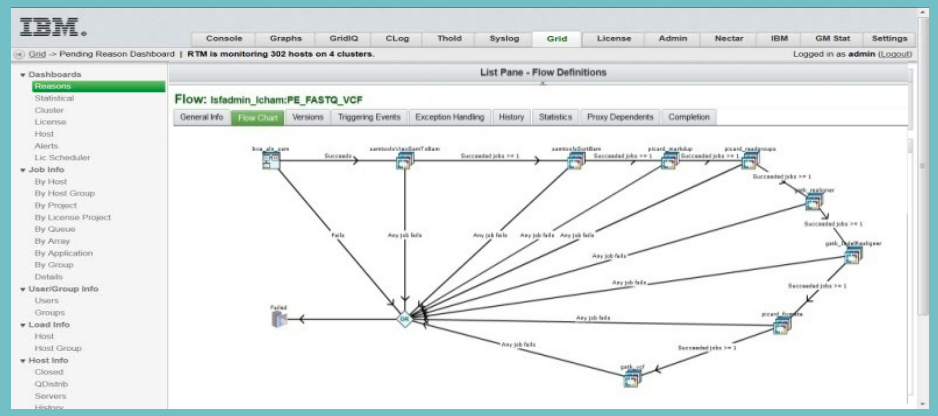
Simplify User Experience and Automate Workflows

IBM Spectrum LSF Application Center



Provides flexible, application-centric interfaces for cluster users and administrators that are easy to use, deploy, manage and support

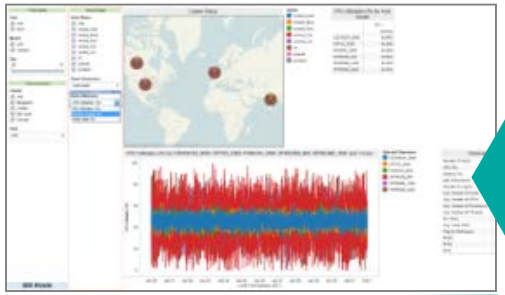
IBM Spectrum LSF Process Manager

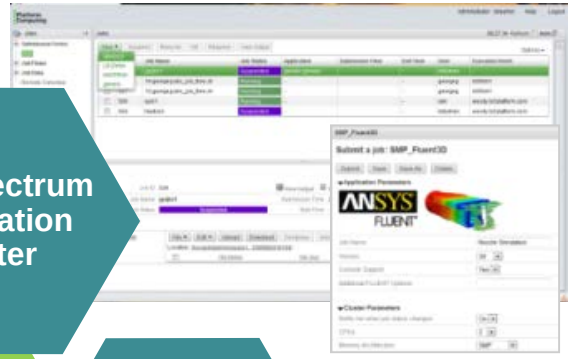


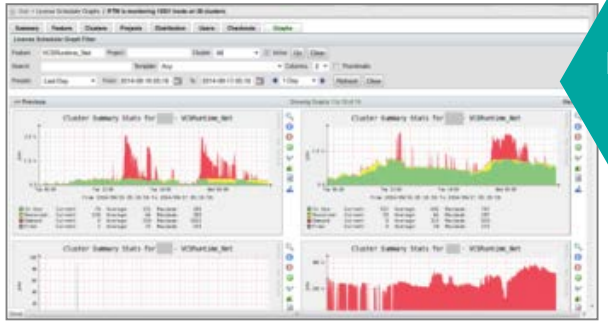
Enables organizations to design and automate computational or analytic processes, making them reliable and capturing repeatable best practices

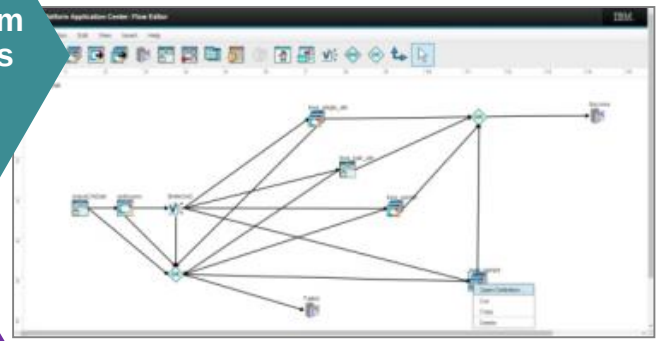
© 2016 IBM Corporation

The IBM Spectrum LSF Family









IBM Spectrum LSF Analytics

IBM Spectrum Application Center

IBM Spectrum LSF RTM

IBM Spectrum LSF Process Manager

IBM Spectrum LSF

IBM Spectrum LSF Data Manager

IBM Spectrum LSF License Scheduler

IBM Spectrum LSF Session Scheduler

Hadoop Connector

Spark Connector

Docker Connector

Resource Connector for OpenStack & Cloud

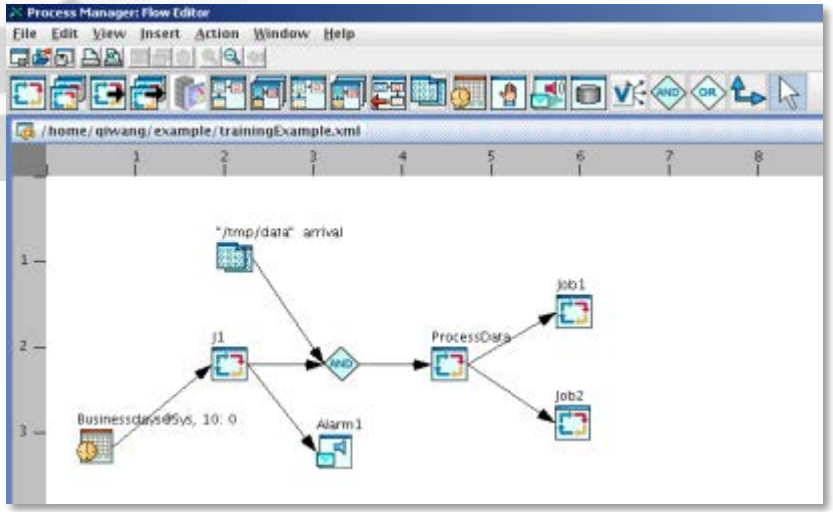
IBM Spectrum MPI

IBM Spectrum Cluster Foundation

© 2017 IBM Corporation

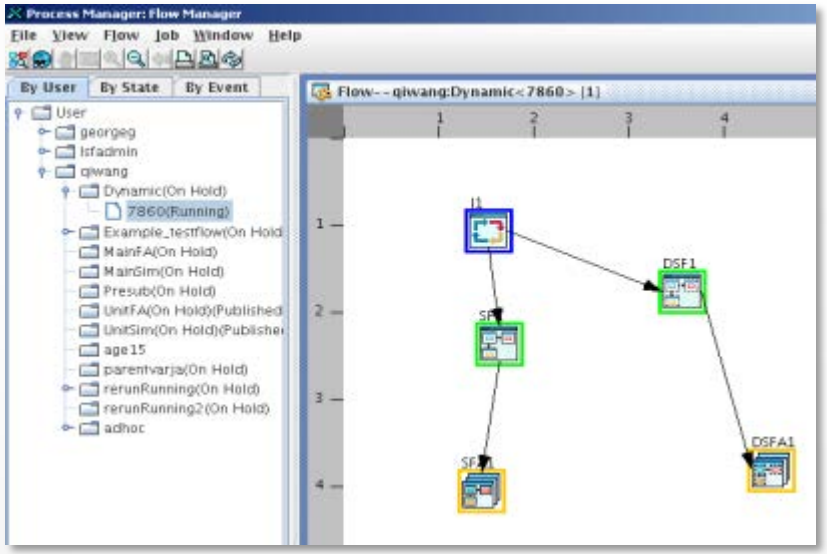
IBM Spectrum LSF Process Manager

Visual Flow Editing, Management



IBM Spectrum LSF Process Manager Flow Editor

- Intuitive drag-and-drop interface
- Creates self-documenting flows
- Support for sub-flows, job arrays
- Rich error-handling / retry capability
- Save workflows in XML format
- Publish flows directly to Flow Manager



IBM Spectrum LSF Process Manager Flow Manager

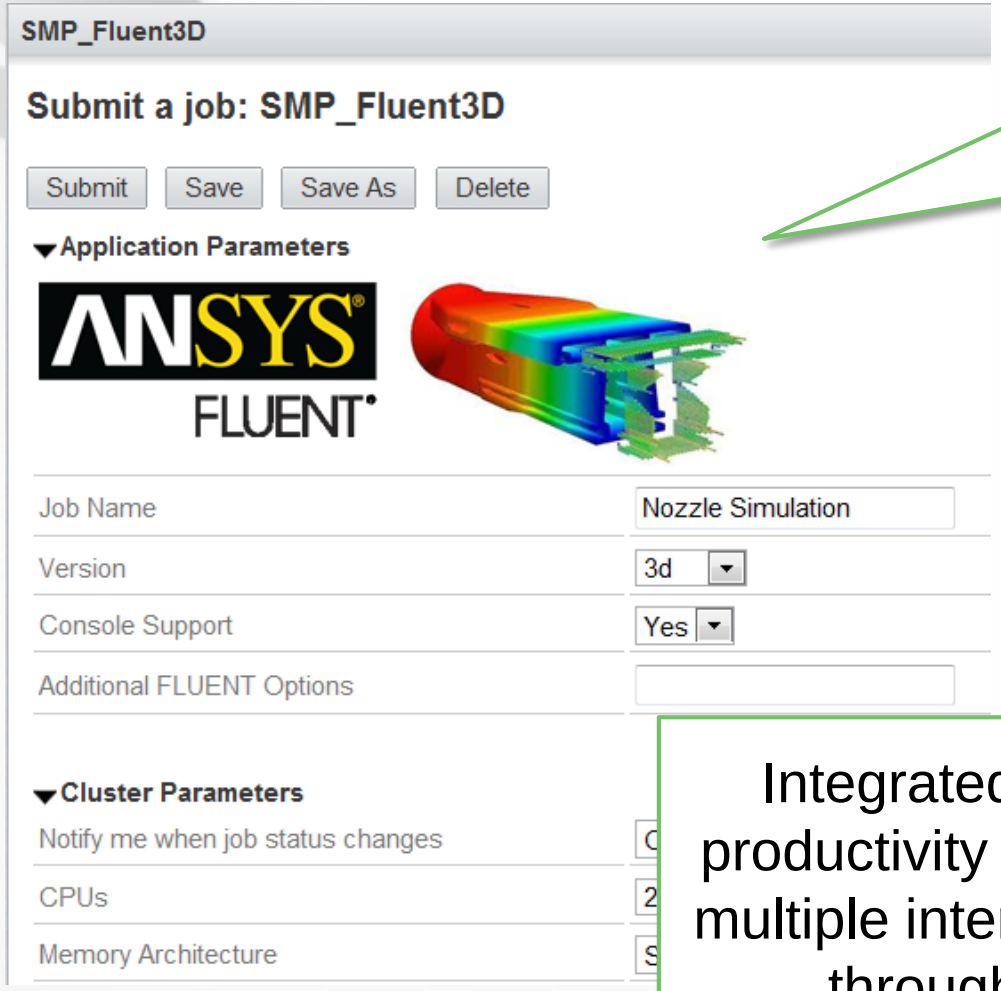
- Manages multiple flows for multiple users and groups simultaneously
- Monitor workflow execution graphically
- Trigger flows automatically through calendar events, the flow manager or the command line.

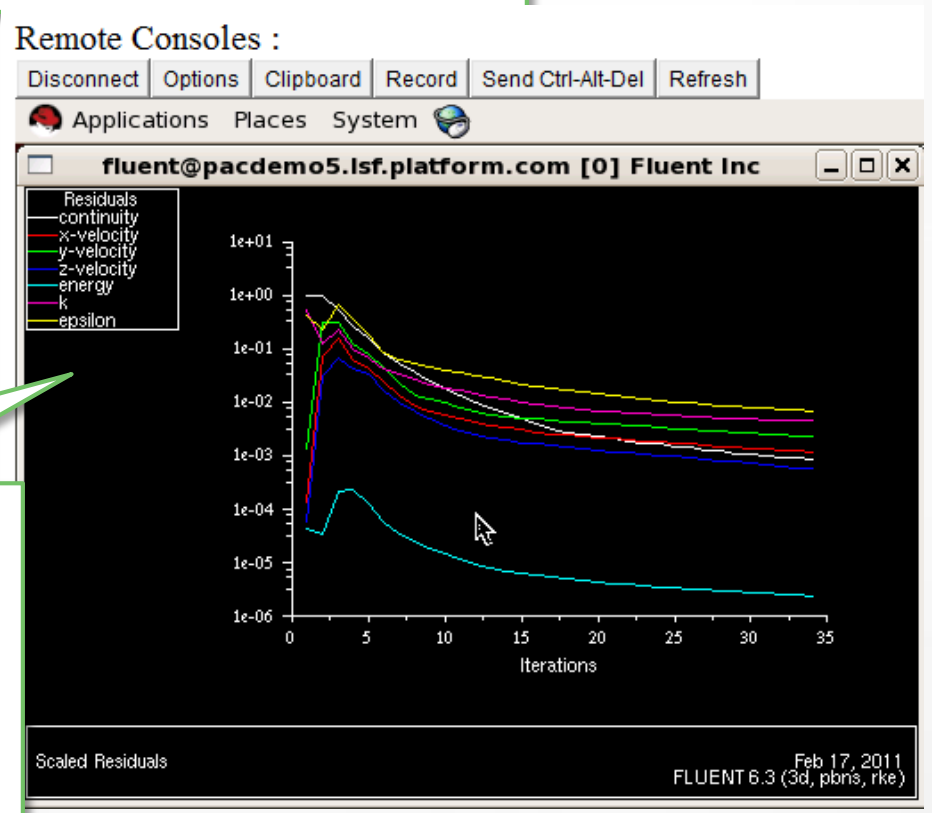
© 2016 IBM Corporation

IBM Spectrum LSF Application Center

Intuitive Application Interfaces

Guided, self-documenting interfaces boost productivity, reduce training and lower support costs.



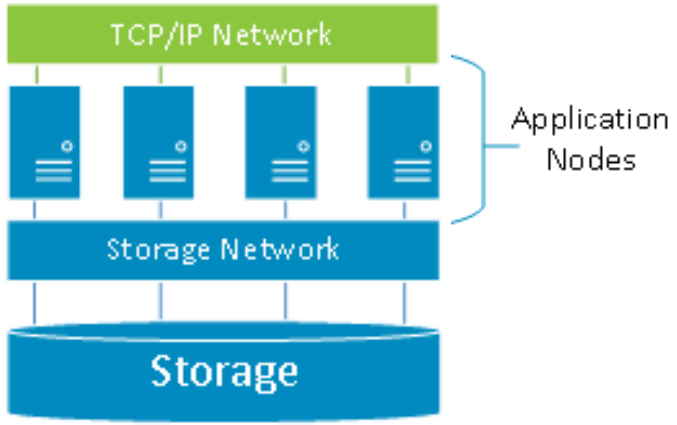


Integrated console boosts productivity enabling access to multiple interactive applications through the browser.

© 2016 IBM Corporation

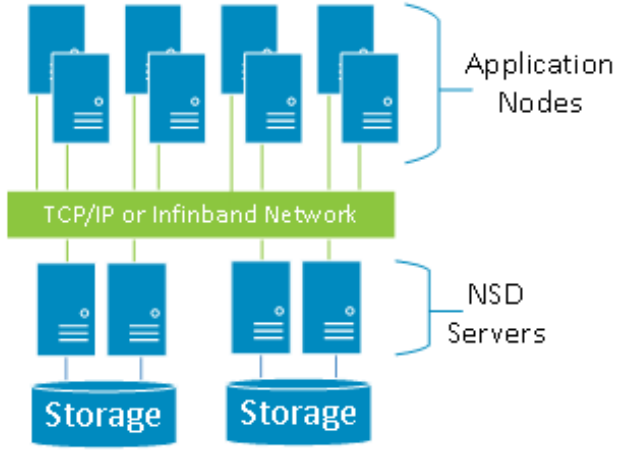
Spectrum Scale deployment models

Enterprise Integrated Model



Unify and parallelize storage silos

Network Shared Disk (NSD) Model



Modular High-Performance Scaling

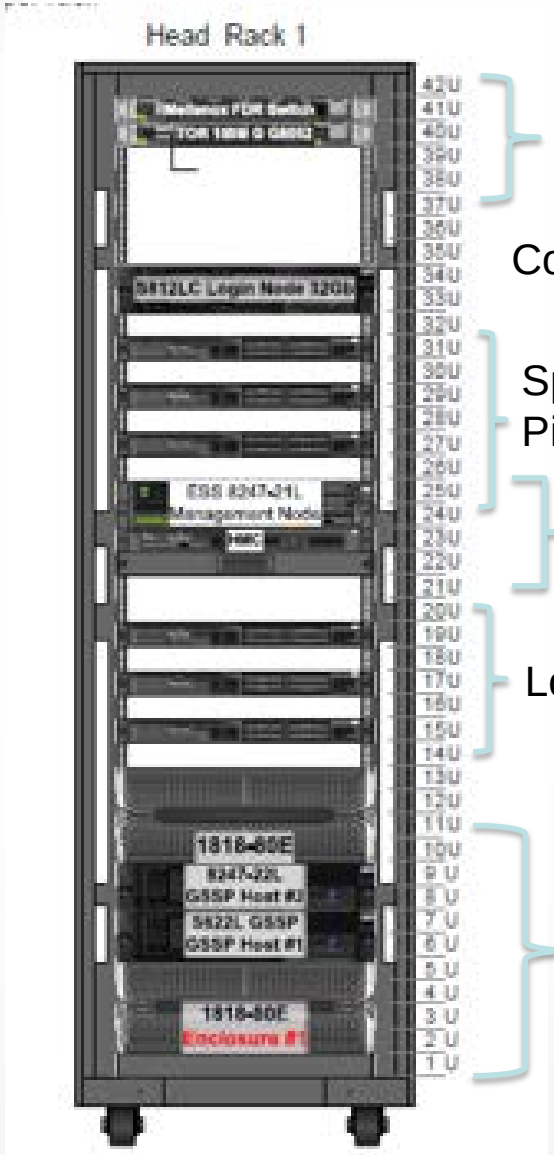
Shared Nothing Cluster (SNC) Model

Span storage rich servers for converged architecture or HDFS deployment

POWER – Power8 HPC Solutions & Elastic Storage Sever

© 2016 IBM Corporation

System Head Rack (Login, Management, Storage)



Top of Rack Mellanox Leaf Switches (Ethernet or Infiniband)

Compute Nodes (FAT)

Spectrum Foundation Pipeline Manager SW.

Spectrum Scale Storage Management Node

Login / SW Mgt Nodes

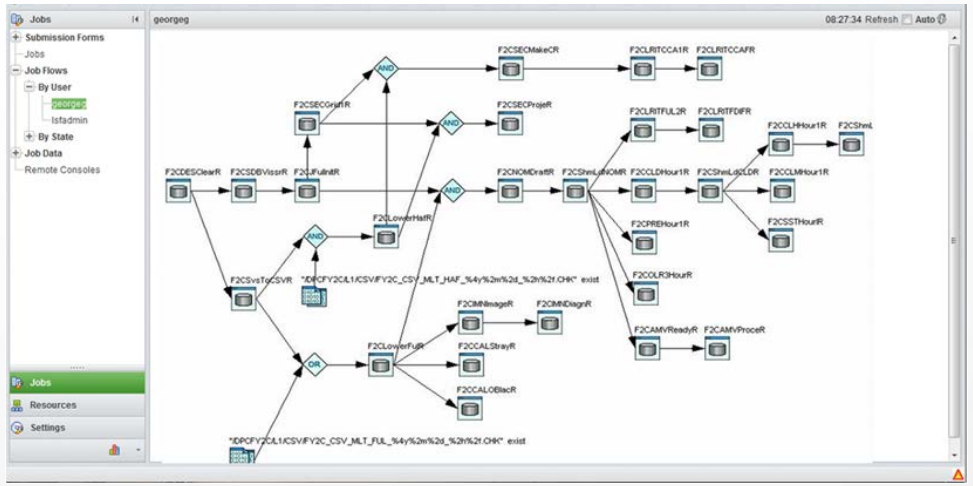
Spectrum Scale
ESS GI2 Storage disc
500 TB Disc
Parallel
6-18 GB/sec

Software Stack

Spectrum Compute Foundation

- Cluster Manager
- LSF Scheduler
- Process Manager
- Application Center
- Workload Manager
- Real Time Manager

End User, Usability



© 2016 IBM Corporation

Sample IBM HPC Dense Compute

- Multiple racks of 1U compute nodes based on Power8+ cpus
- 20 or 24 cores per server.
- 160/192 threads simultaneous
- Infiniband or Ethernet Interconnect
- 22.5KW per rack (no GPU)

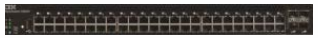
© 2016 IBM Corporation

IBM Delivers Full HPC Solutions

S822LC servers
2 POWER8 CPUs
1U, 2U

NVIDIA GPUs
P100, K80

Mellanox IB, Ethnt switches

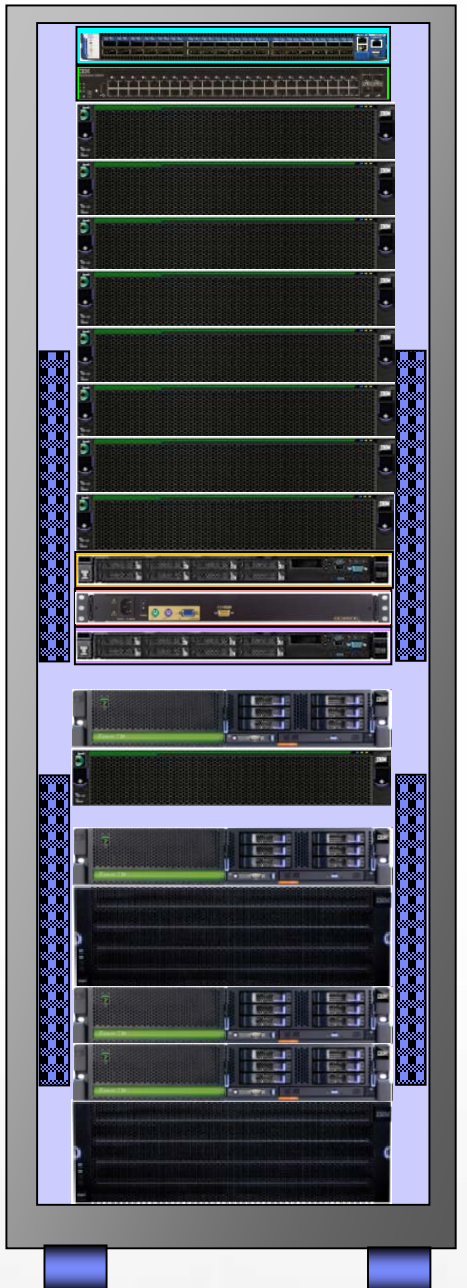


Ethernet switches

Parallel File System
Elastic Storage Server

HPC Software

XL Compilers CUDA ESSL PGI	Spectrum Scale xCAT	Spectrum Compute Fa miily LSF,PAC,PPM	Parallel Env



InfiniBand switch
Ethernet switch

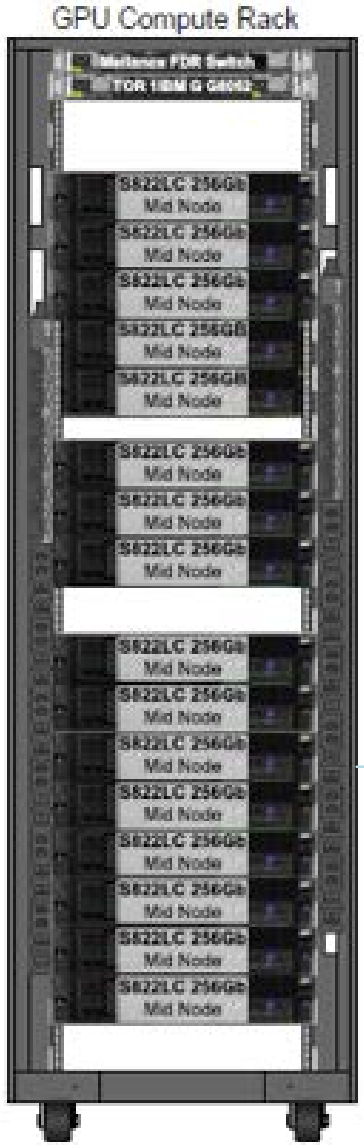
compute nodes

Keyboard/monitor
UFM appliance

Management server
Login node

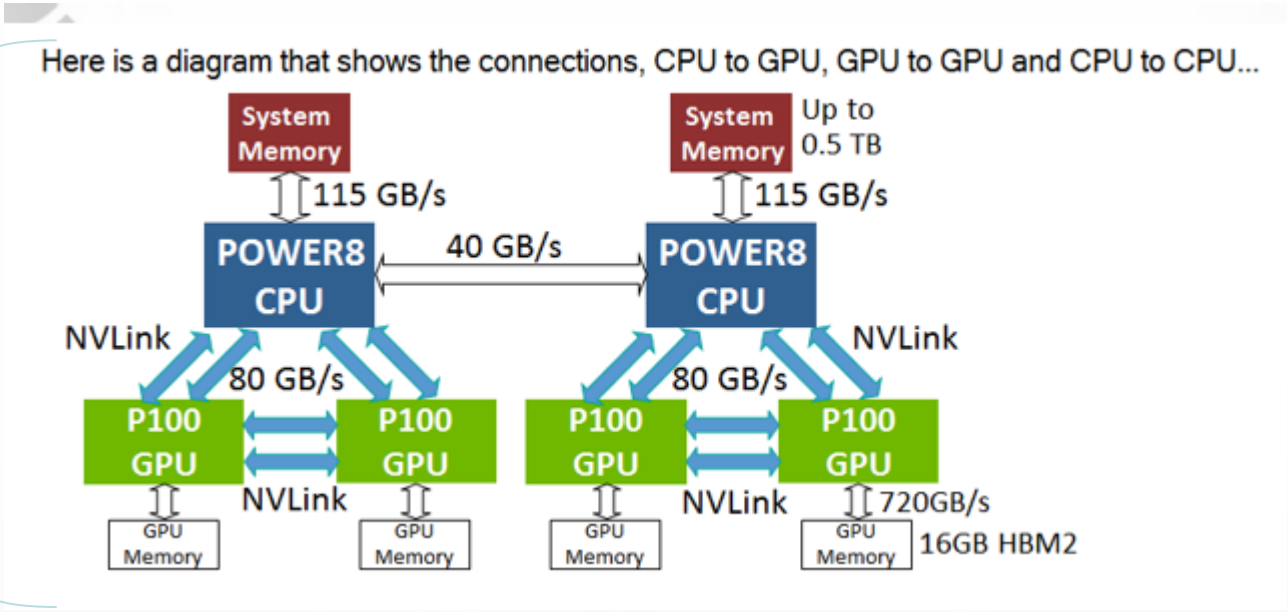
ESS Storage

© 2016 IBM Corporation



Compute Intense FAT nodes for GPU Workloads
Nvidia NVLINK© Technology = IBM OpenPower 822LC

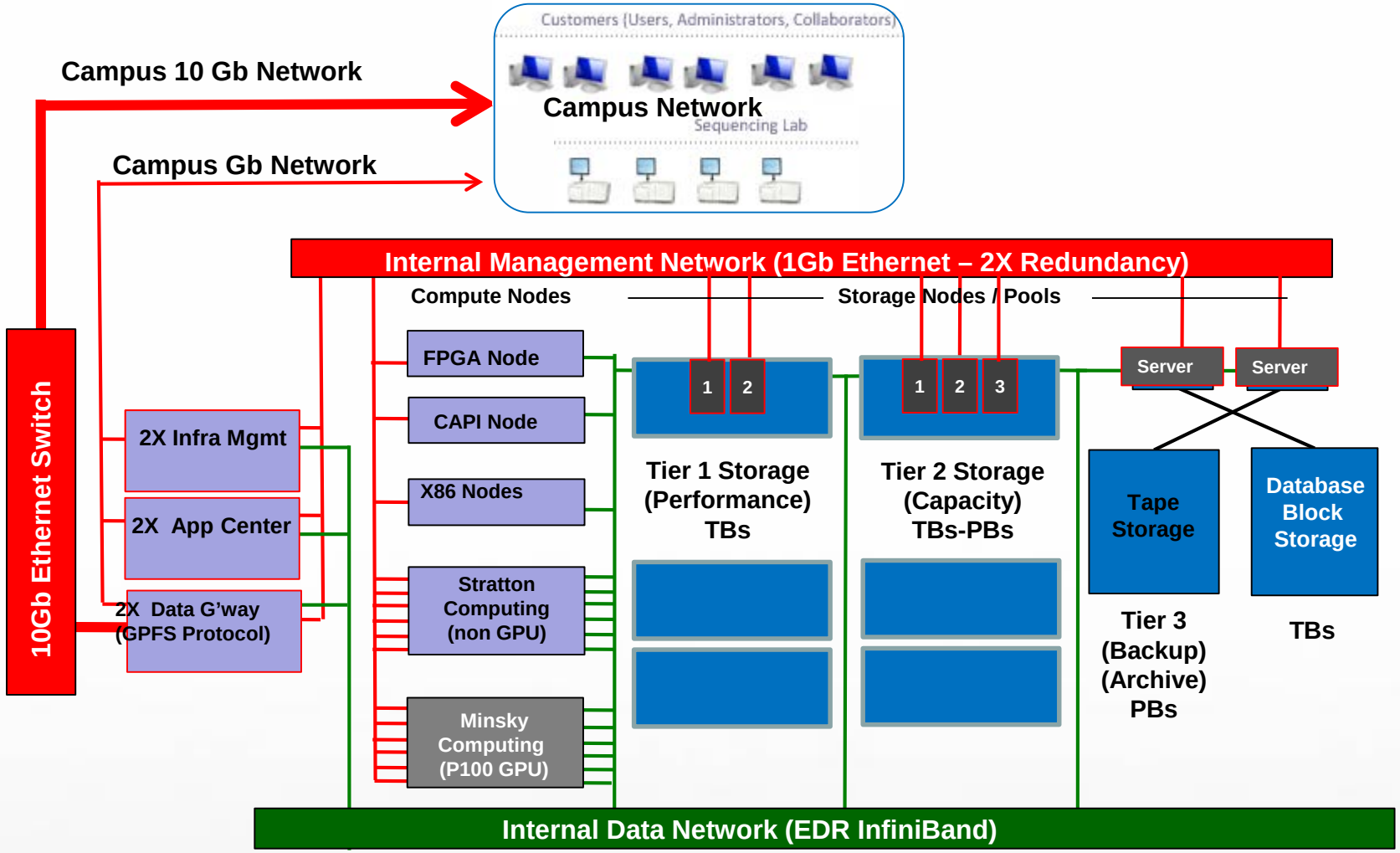
Sample: One or More racks of 2U Servers
16 High Performance IBM Power822LC servers
2 or 4 Nvidia Pascal P100 GPUs , 256 Gb Memory
+ FAT Nodes w 1TB Memory, GPU (or not)



Intense Deep Learning and Bioinformatics Workloads are Seeing 5-10X improvement over Intel x86 PCI based systems

© 2016 IBM Corporation

Sample Solution Architecture



© 2016 IBM Corporation

HPC Storage - Elastic Storage Server

- Elastic Storage Server (ESS) is an integrated file/object based storage solution consisting of:
  - IBM Power System Servers
  - Storage Enclosure and drives
  - Spectrum Scale Software including Spectrum Scale RAID software
  - Networking components
  - **Scalable Parallel File System** ****Built on 15 Years of IBM GPFS**
  - CIFS, NFS, SMB, Posix gateways
  - **Spectrum Object Transparent Cloud Tiering.**

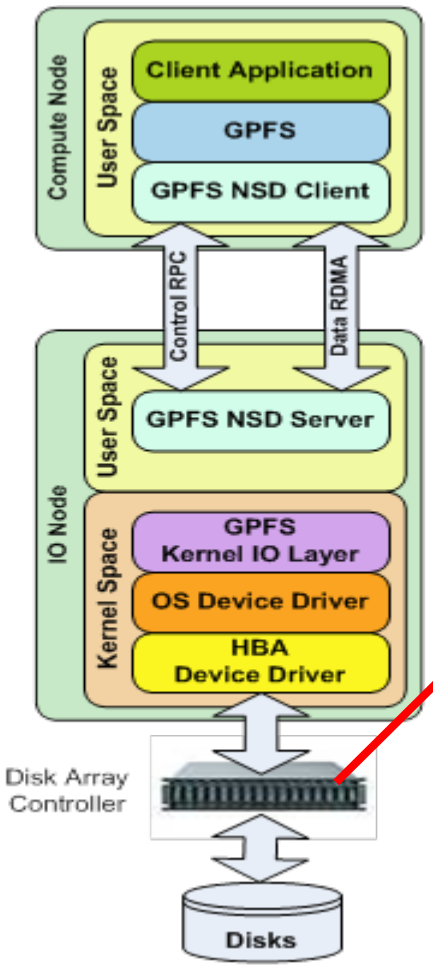
- ESS is a building block solution for Spectrum Scale
  - ESS's can be scaled horizontally to provide massive scaling of:
    - Capacity
    - Throughput **(3-12 GB/second) *****
    - Number of files
    - Number of Objects

- All of the intelligence in ESS lies in Spectrum Scale and Spectrum Scale RAID Software

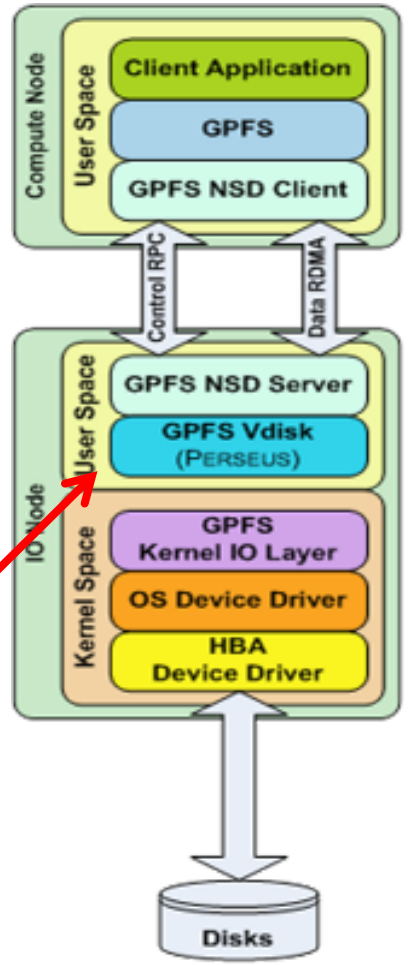
© 2016 IBM Corporation

ESS – *The World's Fastest Spectrum Scale Product*

GPFS „Classic“

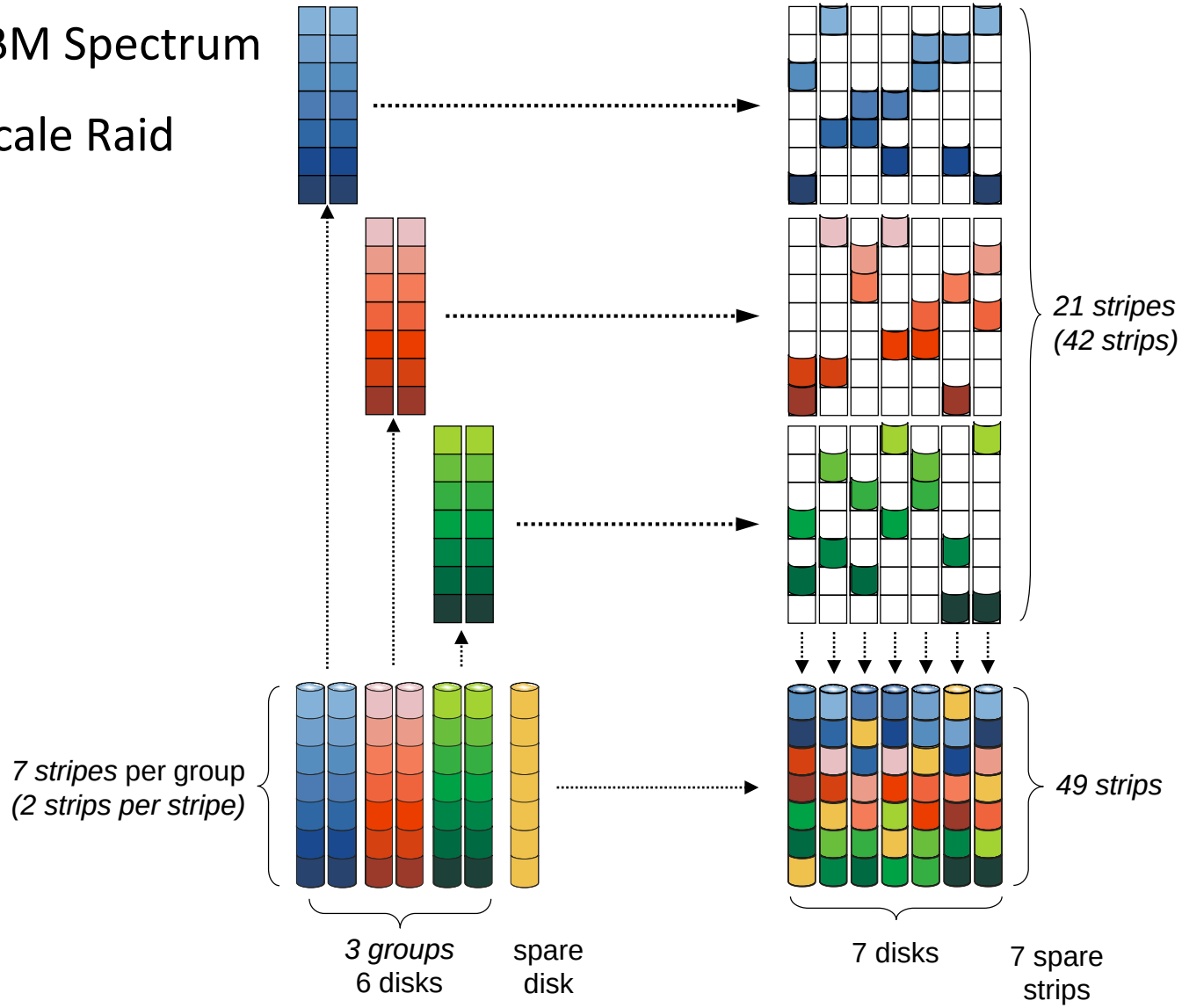


IBM ESS



IBM Spectrum

Scale Raid

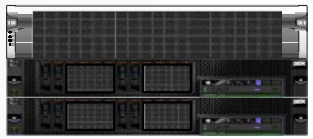


IBM's DeepFlash ESS enhances the Elastic Storage Server family.

GS models use 2U 24x2.5" JBODs or SSDs, GL models use 4U 60x3.5" JBODs,
GF models use 3U 32 JBOF enclosures

Support drives: 1.8TB SAS, 400GB, 800GB, 1.6TB SSD 2.5"; 2TB,4TB,6TB and 8TB NL-SAS 3.5" HDDs

Supported NICs: 10GbE, 40GbE Ethernet and EDR Infiniband

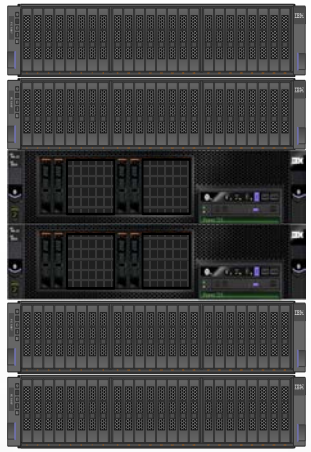


GF1 building block

GF2 building block

GS1 building block

GS2 building block



GS4 building block

GS6 building block

GL2 building block

GL4 building block

GL6 building block

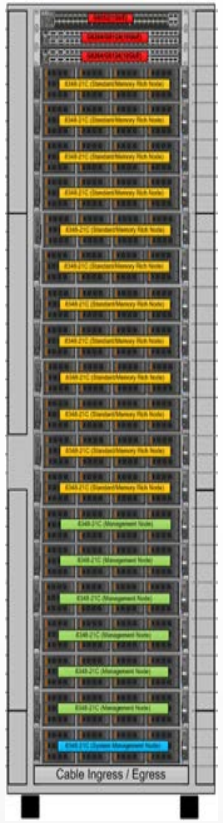
© 2016 IBM Corporation

POWER – IBM Data Engine for Hadoop & Spark

© 2016 IBM Corporation

IBM Data Engine for Hadoop and Spark

OpenPOWER innovation with IBM Open Platform with Apache Hadoop for a high performance, storage dense and fully integrated cluster offering.



- ✓ **Optimized configurations for Hadoop or Spark workloads**
- ✓ **Based on S812LC servers with up to 14*6TB disk drives per server**
- ✓ **Optionally preloaded with IBM BigInsights and IBM Open Platform**
- ✓ **Simplify operations – easy to deploy and manage**
- ✓ **Adapt and scale to your changing analytics needs**

Single vendor support

Up to 2x better price performance for Spark workloads*

Delivered as a **fully integrated cluster** ready to run

OpenPOWER innovation with IBM S821LC servers

- All results are based on IBM Internal Testing of 3 SparkBench benchmarks consisting of SQL RDD Relation, Logistic Regression, SVM

© 2016 IBM Corporation

New Members of LC Server line ideal for Big Data

Storage-rich Offerings
Spark / Hadoop / BD&A

Compute Intensive Offerings
Cloud/HPC Workloads

S822LC *for Big Data*

Data Nodes

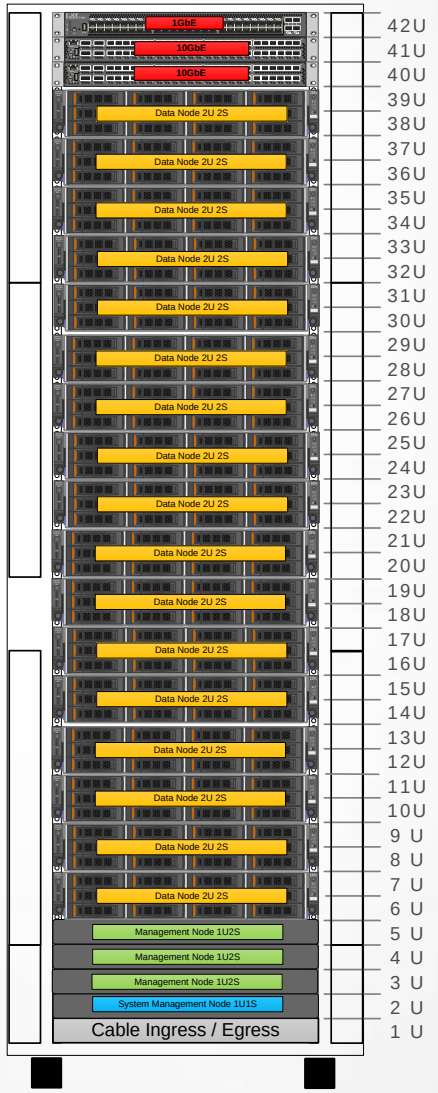
- 2 POWER8 Processors in a 2U
- Up to 2X BETTER price performance on OS Databases
- 40% more ops/s vs x86
- GPU and CAPI enabled
- 96 TB Storage

S821LC

Mgmt Nodes

- 20 POWER8 cores in HALF the space
- GPU and CAPI enabled
- 32 TB of storage

HORTONWORKS®



© 2016 IBM Corporation

Typical IDEHS Configurations

- **Starter/Small (max 1 rack)**
  - 1 sys mgmt, 1 mgmt, 3 data
  - Ideal for dev/test, POC or single use where HA and performance not a consideration
  - Capacity: 72TB raw
- **Medium**
  - 1 sys mgmt., 3-6 mgmt., 7 data (min 3)
  - Redundant 10GB top of rack switches
  - Ideal for small to medium production cluster
  - Capacity: 504TB raw
- **Large (max 2 rack)**
  - 1 sys mgmt., 3-6 mgmt., 12 data
  - Ideal for multiple lines of business in production
  - Capacity: 864TB raw

Data Node Default Options:

- **Standard Analytics Node Default Configuration**
  - 1x POWER8 2.92GHz 10Core + 128GB (16x8GB) DRAM + 12 x 6TB (front drives) + 2 x 1TB HDD (rear drives)
- **‘Memory Rich’ Node Default Configuration:**
  - 1x POWER8 2.92GHz 10Core + 256GB (16x16GB) DRAM + 10 x 6TB HDD + 2 x 960GB SSD + 2 x 1TB HDD(rear drives)

© 2016 IBM Corporation

Actual Use case for Hadoop with Power8 SMT8

Louisiana State University- Genome Assembly with Hadoop

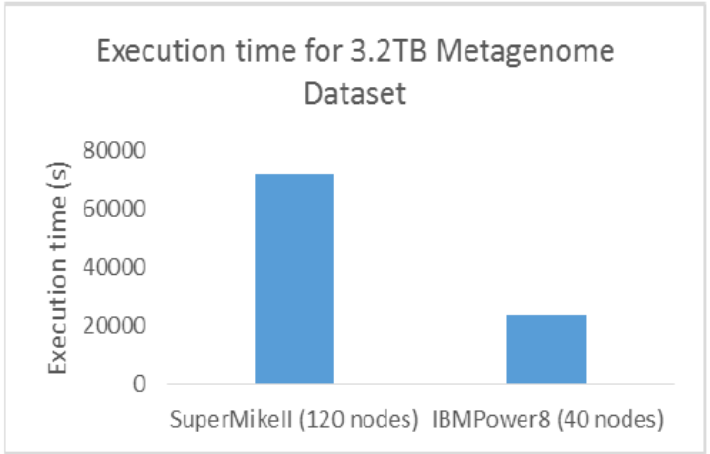
Hadoop workload with SMT8 enabled
+ External GPFS storage (HDFS) connector

- **9x Performance over Intel nodes.**
- 120 Nodes Dell vs 40 Nodes IBM

3x Time performance, 1/3 nodes.

Analyzing Large Size 3.2TB Metagenome

Figure 6: Execution time for 3.2TB



To explore the capability of IBM POWER8 processor with large scale data, we analyzed a dataset size of 3.2TB of metagenome data. The process completed in 6 hours and 22 minutes on the IBM POWER8 cluster using only 40 nodes. This same process takes more than 20 hours to complete on 120 Intel nodes available at LSU. ***That is, for***

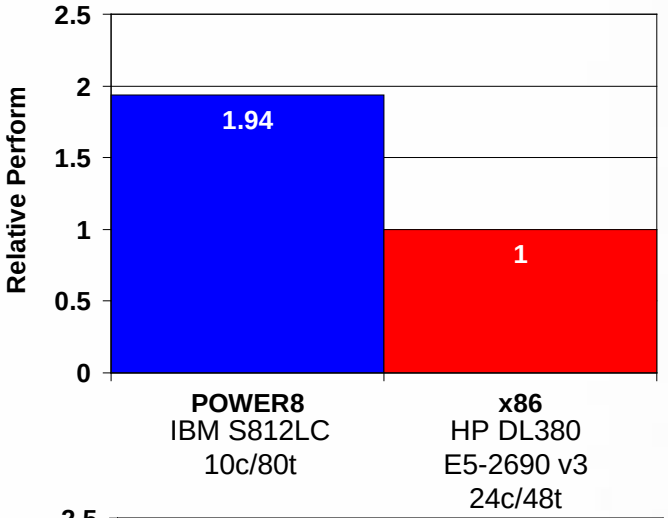
LSU Whitepaper published

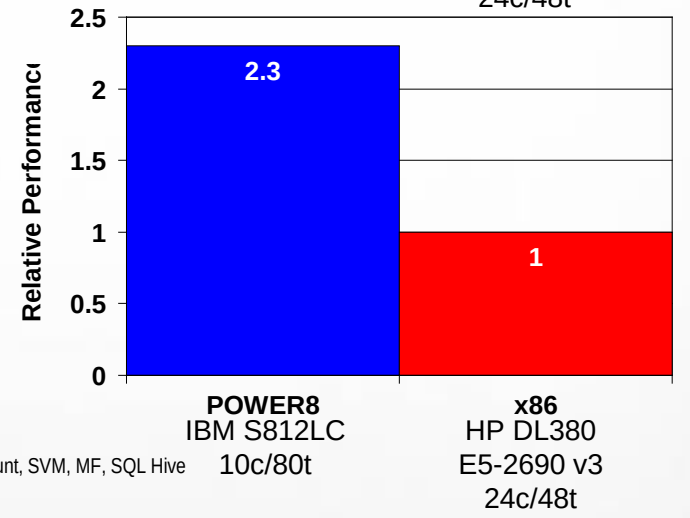
http://www.lsu.edu/mediacenter/docs/LSU-IBM_POWER8_GenomeBenchmark.pdf

© 2016 IBM Corporation

POWER8 Delivers more for Spark

- S812LC delivers optimized Spark price-performance based on an average of 10 SparkBench benchmarks
  - Complete the same Spark workloads for less than ½ the cost of Intel Xeon E5-2690 v3 systems
    - 2.3X BETTER performance per dollar spent
  - 94% more Spark workloads in the same rack space versus Intel Xeon E5-2690 v3 systems
    - 1.94X BETTER performance per system (10 core S812LC vs 24 core DL380)





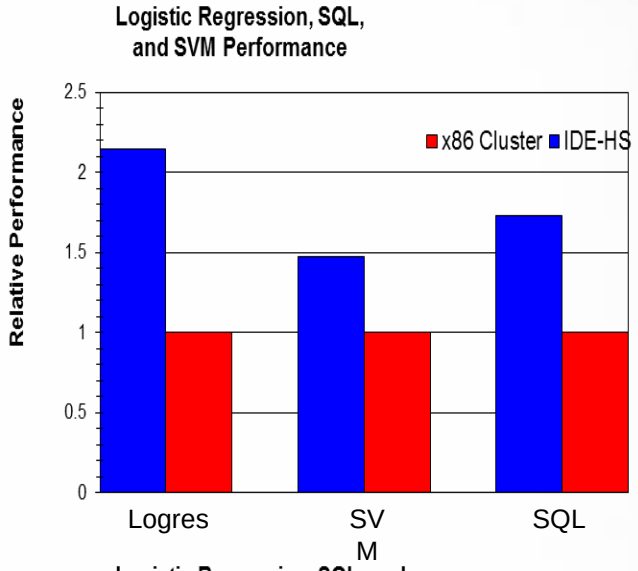
- All results are based on IBM Internal Testing of 10 SparkBench benchmarks consisting of SQL RDD Relation, Twitter, Pageview Streaming, PageRank, Logistic Regression, SVD++, TriangleCount, SVM, MF, SQL Hive
- IBM Power System S812LC 10 cores / 80 threads, POWER8; 2.9GHz, 256 GB memory, Ubuntu 15.04, Spark 1.4, OpenJDK 1.8
- Intel Xeon HP DL380; 24 cores / 48 threads, E5-2690 v3; 2.3GHz , 256 GB memory. Ubuntu 15.04, Spark 1.4, OpenJDK 1.8
- Pricing is based on list prices of HP DL380 and estimated prices of IBM Power S82LC

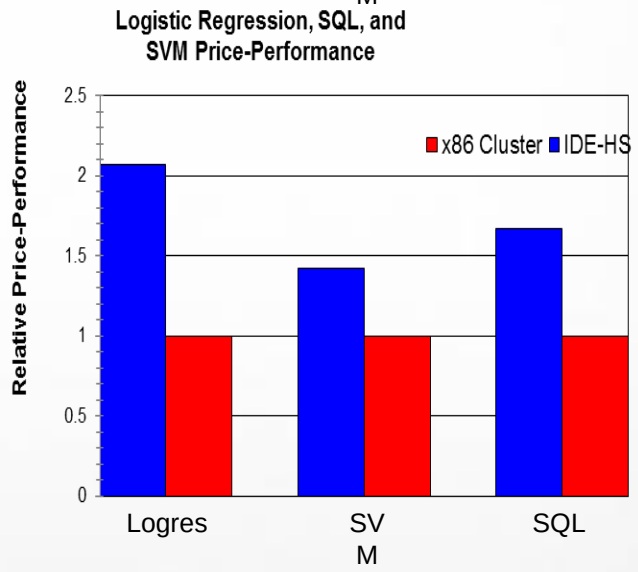
© 2016 IBM Corporation

IBM Data Engine for Hadoop and Spark (IDE-HS) Cluster Performance

Designed for the Cognitive Era to Make Better Decisions even Faster

- IBM Data Engine for Hadoop and Spark infrastructure delivers Spark workload scaling to minimize execution times and reduce batch windows
  - **2.1X** more performance per dollar spent for Spark Logistic Regression based **Machine Learning** used in **model training** by wide variety of lines of business
  - **1.4X** more performance per dollar spent for Support Vector Machine (SVM) – a **Machine Learning** algorithm used in product **Recommender Systems**
  - **1.7X** more performance per dollar spent for **Spark SQL query processing** used widely in Big Data clusters





- All results are based on IBM Internal Testing of 3 SparkBench benchmarks consisting of SQL RDD Relation, Logistic Regression, SVM
- 6 Data Nodes and 1 Management Node. Each node is IBM Power System S812LC 10 cores / 80 threads, POWER8; 2.92GHz, 256 GB memory, RedHat 7.2, Spark 1.5.1, OpenJDK 1.8
- 6 Data Nodes and 1 Management Node. Each node is x86 E5-2620V3 12 cores / 24 threads, E5-2620 V3; 2.4GHz, 256 GB memory, RedHat 7.1, Spark 1.5.1, OpenJDK 1.8
- Pricing is based on web prices of HP DL380 and list prices of IBM Power S812LC

© 2016 IBM Corporation

POWER – Open Software Solutions and Application Performance

© 2016 IBM Corporation

Open Source Ecosystem

Dev. Env /Tools

Available:
Angular.js,Backbone, Bootstrap, buildbot, busyBox, Continuum, CruiseControl, Docker , Eigen lib, Erlang, Ganglia, GCC, gccgo, GDB, GoLang, Gump, Jenkins, Jruby, Lynx, LLVM, logstash, logstash-forwarder, kibana, Maven, Nagios, NGINX, node.js, OCaml, OpenJDK, Phantom.js, PHP, phpMy Admin, Perl, PyCi, Python, Python-Django, Python-Pip ecosystem, PyPy, R, RabitMQ, rsyslog, Ruby, Ruby on Rails (rbenv), Ruby Gems, scala, snappy, Socket.io (npmjs), SpiderMonkey, Strider, Supervisord, SystemTap, Vagrant, V8, Xerces

Optimized:
R

Port In Progress:
Travis-CI

Databases & search engines

Available:
Accumulo (column), Cassandra
CouchDB (document)
Couchbase (noSQL)
Derby, InfluxDB, Firebird
ElasticSearch
EnterpriseDB, LevelDB,
MariaDB (v10.x optimized)
Memcached (KVS)
MongoDB 3.2 (document)
MySQL, Percona Server, PostgreSQL, PrestoDB, RabbitMQ, Redis (KVS), Riak, Sphinx, SQLite, TokyoCabinet
Virtuoso (graph)
Wired Tiger, ZODB

Port In Progress:
GPUDB, Neo4j (graph)

Optimizing:
CouchDB

Evaluating:
InfiniSQL

Big Data & Analytics

Available:
Hadoop Base, Hive, HBase, Accumolo, Ambari, Avro, ElasticSearch Falcon, FluentD Flume, Hue, Kafka, Knox, Lucene-Solr, Mahout, Oozie, Parquet, Phoenix Pig, Ranger, Slider, Spark, Sqoop, Storm, Tez, Zookeeper

Optimizing:
Hadoop

Evaluating:
Clusterpoint

Cloud Managemnt Stack

Available:
Apache Web Server
Apache tomcat
Ceph, Chef Client,
Glassfish, Jetty,
Juju & Juju gui
Landscape client
MAAS, OpenStack, Puppet
Apache Qpid
Thrift
Ceilometer client, Sensu Server & Client

Port In Progress:
Chef Server v12

Other HA, Security etc.

Available:
Alfresco, BTRFS, Bootstrap
CentOS
Chroma-key
Cluster Glue, corosync, CRMshell, DRBD
Evolution data svr, ClusterGlue,
HAProxy, Heartbeat keepalived
Ldirectord, Linux-HA. mesos
OpenSSL
Pacemaker, REAR, resource agents,
Samba, Tophat
WordPress

Port In Progress:

Evaluating:
Cluster-Network
Open Identity Stack (forgerock.com)

Technical Computing

Available:
ALLPATH-LG, bfast, BioConductor, BLAST, BOOST, Bowtie2, BWA, EMBOSS, FASTA, FastQC, HMMER, HTSeq, ISAAC, MUSCLE, PICARD, PLINK, RNAStar, RSEM, SAMTools, SAMTools 1.0, SeqAn, SHRiMP, SOAPAligner, SOAPbuilder, SOAPDenovo, TMAP, TopHat, ABySS, Balsa, Caffe, Code-Saturne, GATK, GMP, Heat3d, IOR, Lattice-Boltzmann, LAPACK, OpenFOAM, waterman ….

Optimized & GPU enabled:
AMBER14, CHARMM, Galaxy, GAMESS, GROMACS, HPC-challenge, KKRnano, LatticeQCD, MAFIA, NAMD. NEST, PEASOUP, PLUTO, Quantum Espresso, QUDA, SOAP3-DP, WRF

Evaluating:
CEGMA, DIALIGN-TX, openQCD, SIFT, STREAM, miRdeep2, ucsctools, ViennaRNA,

Available: Open source application is ported and available on distro (Ubuntu or RHEL or SLES) (black), in community (purple), Lab7 (green) or IBM IOP (blue). Does not mean it is optimized. Does not mean that a commercial ISV version is available.

Evaluating: Needs to be vetted in new business development prioritization process. Some of these are available codes that need optimization to be competitive with x86.

© 2016 International Business Machines Corporation

Available Technical Computing Packages

ABySS, **ALLPATHS-LG**, ALYA, Amber14, ATLAS

BAMtools, Barracuda, bcftools, bedtools, bfast, BioConductor, **BioPerl,** BLAS (libblas3), **BLAST (NCBI)**, Boost, **Bowtie, Bowtie2, BreakDancer, BWA, bzip2**

Caffe, c-ares, CHARMM, **Chimerascan, ClustalW**, Code-Saturne, CoMD (LJForce), Cosmo SVN, CP2k, CPMD, **Cufflinks**

DELLY2

EBSEQ (R), Eigen (eigenlib), **EMBOSS**, Eureka, Eureka-client,
FASTA, FastQC, FASTX-Toolkit, fftw (vectorized), **FreeBayes,**

Galaxy, GAMESS, GATK, **GenomonFisher,** GMP, **gnuplot, Graphviz**, GROMACS

Heat3d, **HMMER**, HOMM-COMM, HPC Challenge (incl. Linpack HPL), HPCG, HPL, **HTSeq, Htslib**

IGV, IOR, **iRODS (beta), iSSAC (Illumina)**

Jurrassic, KKRnano

Lattice-Boltzmann, LatticeQCD (LQCD), LAPACK (liblapack3), LES, LSQR, **LibGD**<partially>, **libpng**, Ludwig

MAFIA, **Matplotlib**, Mdtest, MG2, **Mothur**, MPAS-A, MurmurHash,

NAMD, NEST, nlopt, **nose (Python), NumPy**, NWChem,

Oases, OpenARC, OpenFOAM, OpenQCD

PEASOUP, **PICARD**, PIConGPU, **PLINK**, PLUTO. POPPPerf, **Pysam**. PyReshaper

QuantumEspresso, QUDA

R, regCM. **RNAStar, RSEM**

Sailfish, SAMtools, **Scalpel, SciPy**, SeqAn, SHOC, **SHRiMP, SOAP3-DP, SOAPAligner/SOAP2, SOAPbuilder, SOAPdenovo2, SPlazerS**, spice, SQLite, **SRA-Tools, STAR-fusion,** STREAM

Tabix, tassel, T-Coffee, TMAP, TopHat, Trinity

VASP, variant_tools, Velvet/Oases, native vector support libs, **W**aterman, WRF

© 2016 IBM Corporation

© 2016 International Business Machines Corporation

HPC Software Stack for Power Systems

	Products	Client Benefits
Systems Management	PCM Standard Ed xCAT	• Ease of Use: web portal • Customizable: admin productivity • Faster time to system productivity • Robust monitoring
Application Runtime	IBM MPI runtime ESSL / PESSL CUDA runtime	• Optimized Parallel Runtime • Optimized LAPACK and ScaLPACK libraries • User controlled workflow support
Development Productivity	PE Developer Ed XL Complier suite Totalview/DDT debuggers	• Modern application development environment using Eclipse • Performance analysis tools to help analyze applications • Optimized compiler for Power
Workload Management	Spectrum LSF	• Optimized utilization of resources • Policy and resource aware scheduling • Robust add-on features
Data Management	Spectrum Scale HPSS Spectrum Protect	• Scalable/reliable storage for parallel files system (ESS) • ILM for transparent migration of data from storage to tape and back • Enhance availability with RAID-based ESS and Tape
Application Environment	Spectrum Application Center Spectrum Process Manager Spark,Hadoop Connectors	• Simplify job submission for repeatable workload: customization • Customizable • Workflow management - Faster time to system productivity

© 2016 IBM Corporation

IBM Confidential

Compiler Offerings include Proprietary and Open Source versions of Acceleration Enabled Programing Models

PGI®

CUDA

Key Features:

- Gives direct access to the GPU instruction set
- Supports C, C++ and Fortran
- Generally achieves best leverage of GPUs for best application performance
- PGI/NVIDIA Compiler
- CUDA C/C++ for Power via XL NVCC

OpenACC
Directives for Accelerators

Key Features:

- Designed to simplify Programing of heterogeneous CPU/GPU systems
- Directive based parallelization for accelerator device
- PGI/NVIDIA Compiler
- OpenACC/gcc

OpenMP

Key Features:

- OpenMP 4.0 introduces offloading and support for heterogeneous CPU/GPU
- Leverage existing OpenMP high level directives support
- IBM XL Compiler
- Open Source LLVM OpenMP Compiler

© 2016 IBM Corporation

IBM Confidential

Return to top

Give us your hardest workload!

How can IBM make you and the university successful?

Thank you….

© 2016 IBM Corporation