



High Performance Computing with System x

Luigi Brochard, IBM Distinguished Engineer
Director, Technical Computing IBM System x

**16th ECMWF Workshop,
October 27-31 2014**

High Performance Computing Challenges

1. Unceasing demand for Compute + Data

- Increased performance leading to higher complexity and scalability
- Faster access to growing data

2. Convergence of Compute + Data

- Rise of Big Data clusters driving need to support any/all distributed applications (compute + data) in a common infrastructure

3. Clusters → Grids → HPC Clouds

- From single purpose clusters of static resources to workload-driven dynamic hybrid HPC resource clouds

4. Integrated Solutions

- Deliver comprehensive integrated systems – easy-to-use, optimized for workloads, and ready-to-run applications



High Performance Computing Challenges



1. Unceasing demand for compute + data



- Delivering innovative, energy-efficient, high-performance systems
 - NeXtScale systems for compute intensive
 - NeXtScale and X6 systems for data intensive
- Delivering innovative, reliable, high-performance storage
 - GPFS and GPFS Server Storage (GSS)
- Intelligent system/job management
 - xCAT, Platform Cluster Manager, Platform LSF
 - Integration of Energy Aware features in xCAT and LSF



IBM NeXtScale System™

Modular, high-performance system for scale-out computing

Chassis



Compute



Storage



Acceleration



Primary Workloads



- A architecture for now and the future
- Better data center density and flexibility
- Compatible with standard racks
- Optimized for Top of Rack Switching
- Software Defined Networking (SDN) ready
- Optimized for Energy Efficiency
- Best of iDataPlex extended beyond HPC

Standard Rack



IBM NeXtScale System continues to evolve



Compute
IBM NeXtScale nx360 M5



Now Shipping ✓

Storage
nx360 M5 + Storage NeX
Up to 8x 3.5" HDDs



Now Shipping ✓

Acceleration
nx360 M5 + PCIe NeX
Up to 2x GPUs



Now Shipping ✓

NeXtScale M5 installations under way



516 nx360M5 compute nodes
380 nodes with Xeon Phi 7120P
GPFS, IB FDR14
12 x NSD servers



3096 nx360m5 compute nodes
Direct Water Cooled
GPFS, IB FDR14
10 GSS26, 5 PB -100 GB/s

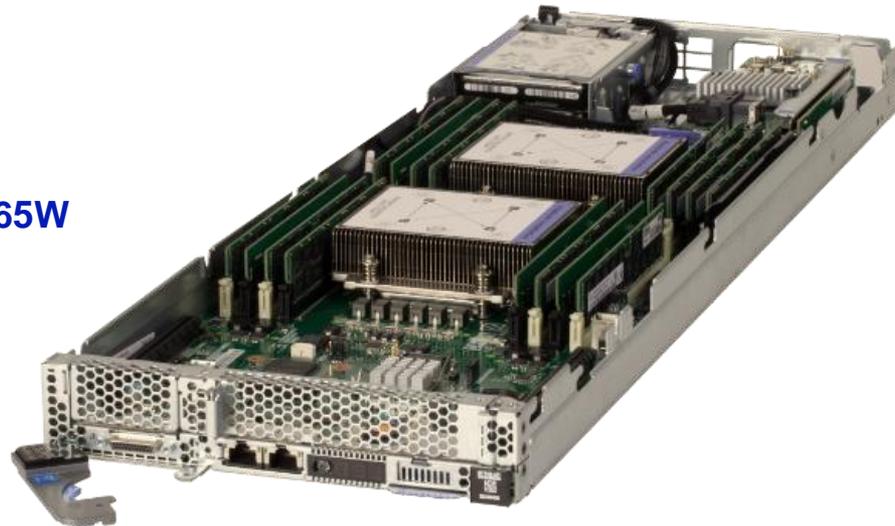


516 nx360M5 nodes
GPFS, IB FDR14
24 x NSD servers
4 SFA12K, 3PB, 120 GB/s



IBM NeXtScale nx360 M5 Improvements

- Processing
 - TOP BIN Haswell and Broadwell
 - 8, 10, 12, and **14, 16 and 18 cores up to 165W**
- Memory
 - **16 DDR4 DIMMs**
 - **2 DIMMs/channel at 2133Mhz**
 - **256GB sweet spot capacity**
- RAID
 - **Extra embedded PCI slot – allows mezz, PCI card, and RAID on system**
 - **Embedded RAID connects to mini SAS external port**
- Disk
 - Flexibility : Diskless or 1x 3.5” HDD
 - **Option for 2 x 2.5” HDD / SSD Hot Swap or additional standard PCI slot**
 - IO Throughput - **4 x 2.5” Simple Swap**
 - Extreme IO Throughput - **4 x 1.8” SSD**

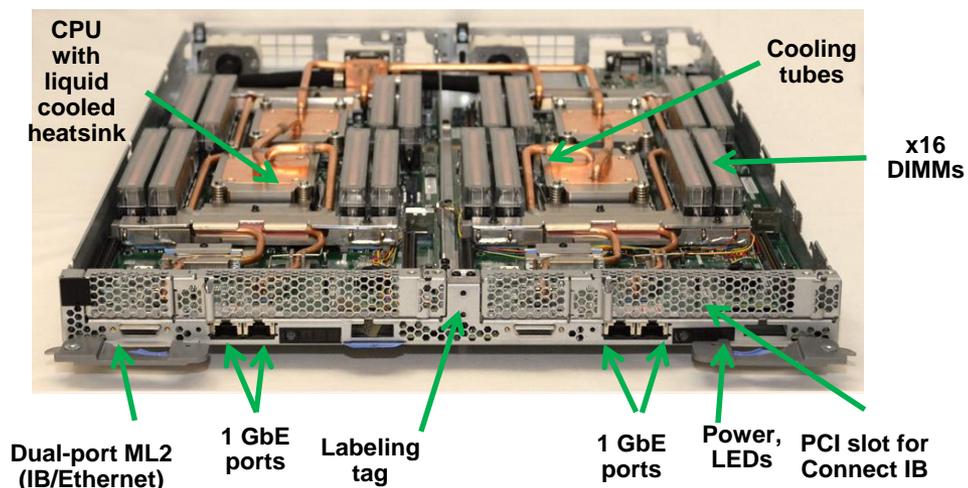


IBM NeXtScale System with Direct Water Cooling Technology (DWC)

Water Cool Node & Chassis

- Reduce power consumption
 - **6% savings**
- Inlet water temperature from 18° C to 45° C
 - **No Chillers required**
- Faster Processor
 - **2.8 GHz 16 core Intel E5-2698a v3 CPUs**

nx360 M5 WCT Compute Tray (2 nodes)

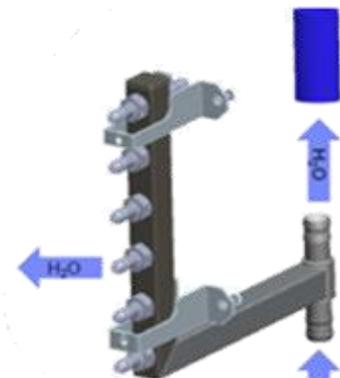


n1200 WCT Enclosure



6 full wide bays
12 compute nodes

n1200 WCT Manifold



Power efficiency designed into HW, SW and management

Efficient Hardware

- 80 Plus Platinum power supplies at over 94% efficiency – **saves 3-10%**
- Extreme efficiency voltage regulation – **saves 2-3%**
- Larger, more efficient heat sinks require less air – **saves 1-2%**
- Smart sharing of fans and power supplies reduce power consumption – **saves 2-3%**
- Underutilized power supplies can be placed into a low power standby mode.
- Energy efficient turbo mode
- Energy Star Version 2⁽¹⁾

Control beyond the server

- Choice of air or water cooling
- No fans or auxiliary cooling required for water cooled version, saving power cost
- Pre-set operating modes - tune for efficiency, performance, or minimum power
- Chassis level power metering and control
- Power optimally designed for 1-phase or 3-phase power feeds
- Optional intelligent and highly efficient PDUs for monitoring and control

Powerful Energy Management

- xCAT APIs allow for monitoring/controlling power at chassis and node level
- Powerful sleep state⁽²⁾ control reduces power and latency
- LSF Energy Aware features allows for automatic energy tuning – **saves 5 - 10%**
- Platform software can target low-bin CPU applications to lower power on CPUs in mixed environments
- Platform Cluster Manager Adv. Edition can completely shut off nodes that are not in use
- Open Source monitoring tool friendly allows of utilization reporting

(1) Pending announcement of product ; (2) On select IBM configurations

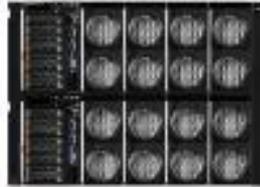


System x rack servers for data intensive

Scale-Up



x3950 X6



**X6 Platform with
Maximum Scalability
and Performance**

x3850 X6



**X6 Platform with
Fastest Performance
and Mainframe
Reliability**

**Mission
Critical**

Scale-Up for Maximum Performance and Capacity

- 120 cores SMP
- 12TB of memory – the largest in-memory databases
- 22 PCI slots – maximum IO bandwidth and flexibility
- Partitioning of a single 8-way into two 4-ways and back with FlexNode
- 50% more flash capacity leveraging eXFlash memory-channel storage
- Ability to upgrade processor and memory technology to next generation within the same chassis

Performance and Reliability for Mission Critical Applications

- 60 cores SMP
- 6TB of memory – business analytics, in-memory applications
- 11 PCI slots – maximum IO bandwidth and flexibility
- 2S and 4S Modularity – modular book design that allows clients to create the most cost effective 2-Socket EX or 4-Socket EX.
- 50% more flash capacity leveraging eXFlash memory-channel storage
- Ability to upgrade processor and memory technology to next generation within the same chassis



3850 X6 Data Intensive Capacity



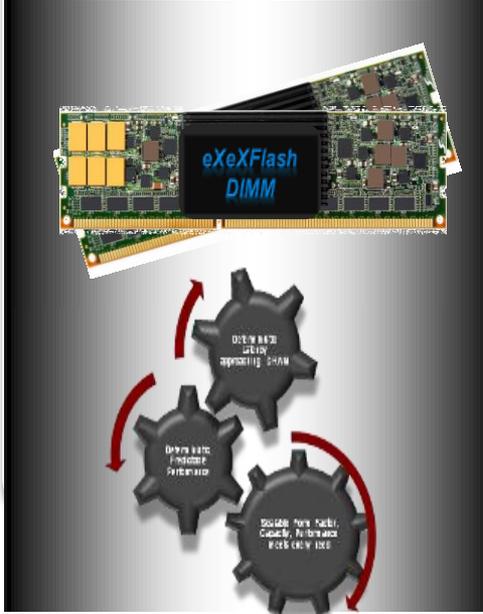
x3850 X6 *without* eXFlash DIMMs can support up to 25.6TB of Flash storage using 16 x eXFlash 1.8” 400GB SSDs and 8 x 2.4TB MLC Duo Adapters.

x3850 X6 *with* eXFlash DIMMs can support up to an additional 12.8TB Flash storage leveraging 32 x 400GB eXFlash DIMMs for a total of 38.4TB of flash storage in the server.



eXFlash DIMMs provides Highest Performance Storage

IBM eXFlash DIMM (IBM System x)



- **Fastest flash response time**
 - Lowest write latency of any Flash offering, as low as 5 microseconds
 - up to 51% lower latency than PCIe* and 86% lower latency than SSD**

- **Highest IOPs**
 - >150K IOPS per eXFlash DIMM†, scaling linearly to up to 4.8M IOPS per server without negatively impacting latency.

This is a much higher than :

- Data Direct Network: 1.7M IOPS
- Cacheio: 1M IOPS
- Fusion-IO: 285K IOPS
- EMC ExtremIO: 250K IOPS

*Fusion IOScale @ 19µs write latency **Intel S3500/S3700 SSD @ 65µs write latency



GPFS Server Storage

1H14

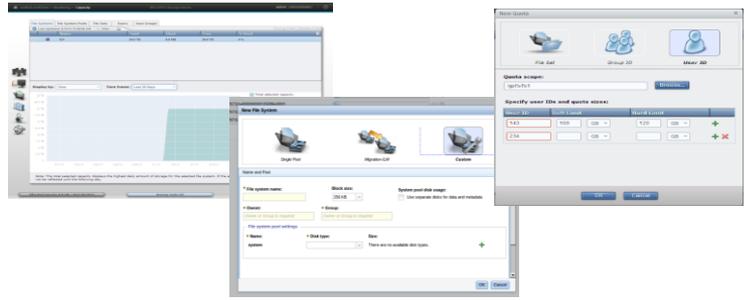
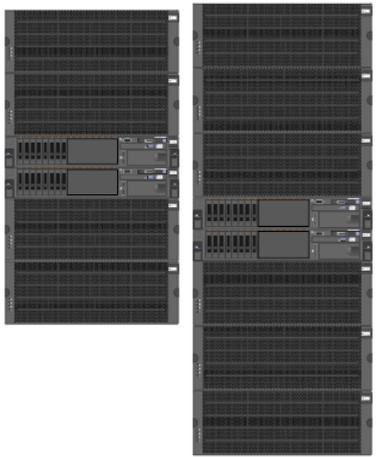
2H14

2015

GPFS Storage Server
 GSS 24 and GSS26
 2 x x3650M4
 4 or 6 60HDD 4U JBOD
 SAS connected
 10Gb or IB based

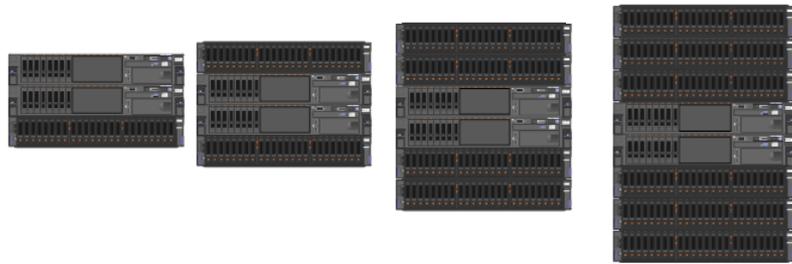
GPFS Storage Graphical User Interface (GUI)
 System Monitoring
 System Maintenance
 User Configuration

GPFS Storage Server
 GSS 24 and GSS26
 Haswell refresh



GPFS Storage Small (S) Solutions
 GSS21S, 22S, 24S, 26S

GSS 21S – 24 SSD model
GSS 22S – 48 2.5" SAS HDD or SSD
GSS 24S – 96 2.5" SAS HDD or SSD
GSS 26S – 144 2.5" SAS HDD



A decorative graphic in the top-left corner consisting of several overlapping, semi-transparent circular segments in shades of purple, pink, and red.

Lenovo, System x and HPC

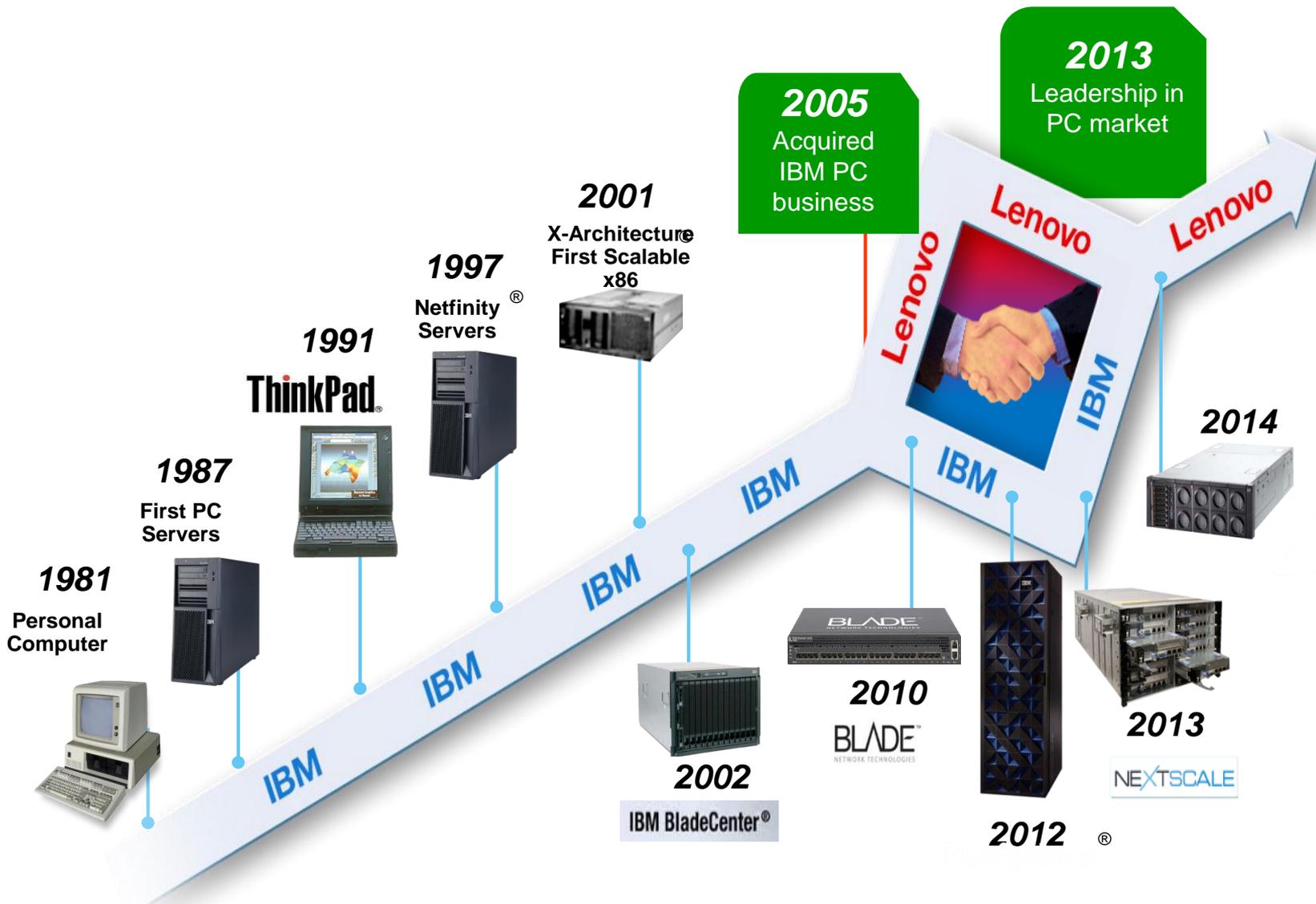


Announcement Highlights

- On October 1st, Lenovo acquired IBM's x86 server portfolio and related resources and operations including the following:
 - System x, x86 BladeCenter, Flex System blade servers, Flex System based integrated infrastructure systems, NeXtScale, iDataPlex, related blade switching and solutions, selected systems software
 - Development, sales and marketing, finance, legal, integrated supply chain, operations, IT and manufacturing
 - Service and support (maintenance)
- IBM will retain its enterprise systems portfolio, including System z, Power Systems, Storage Systems, Power-based PureFlex servers, PureApplication and PureData appliances
- Lenovo and IBM have entered into a strategic relationship which will include a global OEM and reseller agreement for sales of IBM's industry-leading entry and midrange:
 - Storwize disk storage systems, tape storage systems,
 - General Parallel File System software,
 - SmartCloud Entry offering
 - elements of IBM's system software portfolio, including Platform Computing solutions



Deep History of Innovation and Collaboration in the x86 Industry



World Class Service and Support Continues

IBM is a recognized leader in service and support

- Access to the full IBM systems and support capabilities worldwide
- 209 countries/nations, speaking ~127 languages
- >500 part stocking locations with more than 9M shipments annually



Lenovo's Service Commitment

“After the deal closes, IBM will continue to provide maintenance delivery on Lenovo's behalf for an extended period pursuant to the terms of a five-year maintenance service agreement with IBM. Customers who originated contracts with IBM should not see a change in their maintenance support for the duration of the customer's contract.”

SOURCE: <http://shop.lenovo.com/us/en/news/ibm-server>



Lenovo's - Commitment to the HPC Marketplace

- **Post close - Lenovo intends to fully support all current client commitments**
- **Post close - Lenovo will support all commitments that are made by IBM pre-close**
- **Lenovo will use the products and skills it has acquiring from IBM to continue to build out a powerful portfolio of open standard systems and technologies for their clients**
- **Lenovo will also use the engineering skills to continue to differentiate in the x86 space**
 - Large scale systems expertise
 - Water Cooling technologies
 - Graphics and co-processing knowledge
 - Power management and control
 - Performance optimization
- **Lenovo is not buying to continue the current market position of System x**
 - goal is to grow the x86 server business just as they have the PC business.



Thank-You !

