IBM HPC Topics
IBM Imperatives*

- Transform industries and professions with data
- Remake enterprise IT for the cloud
- Reimagine work through mobile and social technologies
- Rethink the challenge of security
- Create new infrastructure for a new era

* IBM's 2014 Annual report
Transform Industries and professions with data
IBM provides solutions that integrate all data sources, structured and unstructured and support descriptive, predictive, prescriptive analytics and is adding cognitive computing with IBM Watson

- IBM reported over $17B in revenue from business analytics in 2014
- In 2014 announced $1B investment plan to accelerate the commercialization of Watson

April 13, 2015  Announcement of Watson Health – a new line of business with headquarters in Boston

- Applying Big Data, cognitive computing, security and cloud computing capabilities to help improve the quality of healthcare.

- Establishment of **Watson Health Cloud** bringing together a de-identified, HIPPA enabled data repository with Watson Health Insights services providing cognitive and advanced analytics

- Announced new partnerships with Apple Health, Johnson and Johnson and Medtronics

- Acquired two healthcare technology companies - Cleveland-based **Explorys** and Dallas-based **Phytel**.

Watson Text Analytics in Biomedical Research

Watson rapidly identifies and normalizes user-defined scientific and clinical concepts within millions of pages of unstructured biomedical information.

Omics Information
Variant Annotations

-exonic NOD2 16 … a frameshift …
-SNP… exonic GJB2 13 … associated with hearing loss … exonic CRYL1,GJB6 13 … a 342kb deletion encompassing GJB6, associated with hearing loss …

Clinical Information
Patient Histories, Clinical Notes, Diagnostic Reports, Discharge Summaries

…was in good health until 2-3 months ago when she gradually developed fatigue and intermittent epigastric pain, … most recent colonoscopy was within normal limits…

Clinical & Scientific Literature

-Peer-Reviewed Articles, Clinical Guidelines, Textbooks, Patents
NHGRI, a branch of NIH, has defined 5 steps for genomic medicine. (source: E. Green et al., Nature 470, 204–213)

Next Generation Sequencing
the focus is on very large data generation, mainly from $1000 whole genome sequencing, and the data processing and reduction includes human, plant, animal, and microbiome genomics

Translational Research
the focus is on data integration including genomic data, and the analytics required to identify biomarkers, understand disease mechanisms, and to identify new medical treatments

Personalized Healthcare
the focus is on delivering genomic medicine to patients to improve outcomes by associating patients with known genomic specific treatments
Significant Computing Challenges with Genomic Medicine:

- **The Omics or Other Data Pipeline**—completing it in a timely fashion and managing the data (long term) that was generated. Research genomics pipelines are now Clinical Genomics Pipelines, adding complexity.

- **The Large Volume of Unstructured Data**—creating “structure” and understanding from clinical notes, journal articles, and other sources.

- **The Integration of Data from Multiple Sources** to allow researchers across the organization to make new discoveries which will ultimately drive changes in patient care. Creating:
  - data schema to support a wide array of use cases
  - metadata query, provenance and security
  - extraction tools and systems to allow further analysis

*COMPLETING THE ANALYSIS WORKFLOW and COMPLETING IT NOW!!!!!*
IBM PowerGene Architecture

PowerGene is a foundation for providing flexible and scalable SOLUTIONS to our customers based on their specific analytical requirements.
Partners Using IBM PowerGene Architecture—Building the Ecosystem

PARTNERS
- Databiology – using IBM SDI as a foundation for clinical needs
- LAB7 – using IBM SDI and POWER as a base for clinical needs
- Broad Institute–GATK application on POWER
- BGI-- SOAP3 acceleration on POWER
- Wash U Genomics Center– IBM Software Defined Environment
- ZATO Health – data federation with GPFS and POWER
- CLC bio (Qiagen)– scaling very large genomics and translational platforms in the IBM HPC Cloud
- BROAD INSTITUTE– GATK acceleration on POWER

Work in process
- Ingenuity (Qiagen)– curated full text (5M) for translational
- TranSMART– data warehouse and Watson supported with SDI/POWER/BigInsights
- Ayasdi– using Symphony (SDI) for application scheduling acceleration
- IRODS tightly integrated with GPFS
PowerGene is the first IBM industry-focused reference architecture for software-defined infrastructure. It is developed for and focuses on life sciences and healthcare industry

- PowerGene.net
- Solution Brief
- Redbook (soon)
Key IT Capabilities Accelerating Scientific Breakthrough

Leading biomedical research organizations are asking for technology solutions that will give them a low-cost competitive advantage in therapeutic discovery

✓ Flexible, scalable, and low-cost high-performance compute and storage solutions capable of efficiently processing rapidly growing quantities of genomic and other types of complex life science data

✓ Seamless integration of complex life science data types

✓ Rapid extraction and analysis of unstructured semantic concepts from very large volumes of clinical and scientific documents

✓ Metadata collection capabilities providing detailed audit trails as source data are transformed into analytical results

✓ Tools for scientific collaboration that enable workload sharing to cross organizations and geographic boundaries in a secure environment appropriate for Protected Health Information
Remake Enterprise IT for the cloud
IBM Cloud revenue in 2014 over $7B. Up 60% from 2013.

Investment of $1.2B to expand IBM Softlayer Cloud Hosting Centers
Investment of $1B to create Bluemix providing platform as a service capability for software developers

Establishment of an HPC Cloud offering

- Providing private bare metal servers and VLANS
- HPC technologies
  - Multi-core servers, Infiniband, technical services, GPFS, Platform LSF and Symphony, Hadoop
  - Hybrid implementations leveraging local and hosted implementations

Support for OpenStack and other multi-tenant cloud based orchestration tools
Create new infrastructure for a new era
IBM Software Defined Infrastructure portfolio

Increase business agility by transform static infrastructure into **dynamic private, hybrid and public cloud** resources

– Accelerate technical computing workloads up to 150X
– Quickly address resource constraints by bursting workloads to the SoftLayer cloud
– Reduce cloud infrastructure deployment time by 76%

Accelerate time to insight and reduced costs by Deploying, running and managing **big data, analytics and technical computing** apps on a shared infrastructure

– Optimize long-running services with lifecycle management for born-on-the-cloud apps
– Gain up to 6X greater performance and 3X increased scalability
– Maximize analytics, file serving and object storage for hundreds of petabytes and gigabytes per second throughput.
Multitude of scale out applications in today’s data centers

- **High Performance Computing**
  - (Batch, Serial, MPI, Workflow)

- **Hadoop / Big Data**

- **High Performance Analytics**
  - (Low Latency Parallel)

- **Application Frameworks**
  - (Long Running Services)

**Example Applications**

**Homegrown**
IBM Software Defined Infrastructure

High Performance Analytics
(Low Latency Parallel)

Hadoop / Big Data

High Performance Computing
(Batch, Serial, MPI, Workflow)

Application Frameworks
(Long Running Services)

Example Applications

Homegrown

Platform Symphony

Platform Symphony
(MapReduce)

Platform LSF

Platform Application Service Controller

Workload Engines

Resource Management

Scheduling & Acceleration With Infrastructure Sharing
IBM Software Defined Infrastructure

Data Management
Combined with Compute
File, Object, Block

IBM Spectrum Storage

IBM Platform Cluster Manager
(Bare Metal Provisioning)

IBM Platform Cluster Manager
(Virtual Machine Provisioning)

On-premise, On-cloud, Hybrid Infrastructure
(heterogeneous distributed computing and storage environment)
### IBM Spectrum Control
- Analytics-driven data management to reduce costs by up to 50 percent
- Virtual Storage Center

### IBM Spectrum Protect
- Optimized data protection to reduce backup costs by up to 38 percent
- Tivoli Storage Manager

### IBM Spectrum Archive
- Fast data retention that reduces TCO for archive data by up to 90%
- LTFS

### IBM Spectrum Virtualize
- Virtualization of mixed environments stores up to 5x more data
- SAN Volume Controller

### IBM Spectrum Accelerate
- Enterprise storage for cloud deployed in minutes instead of months
- XIV Software

### IBM Spectrum Scale
- High-performance, highly scalable storage for unstructured data
- Elastic Storage / GPFS
HPC Hardware Strategy

• High-performance computer and high-performance analytics drive common platform design

• Over $3B committed over the next five years to develop next generation chip technologies

• Embrace standards as foundation for innovation

• Servers will be predominately 2-socket designs

• Developing deeper relationships with technology partners – especially OpenPOWER members

• Majority of floating-point performance will come from GPUs
  • FPGA accelerators also available

• Utilize Industry-standard compliant racks and electronics enclosures
  • Air and water cooling options
Driving industry innovation

The goal of the OpenPOWER Foundation is to create an open ecosystem, using the POWER Architecture to share expertise, investment, and server-class intellectual property to serve the evolving needs of customers.

– Opening the architecture to give the industry the ability to innovate across the full Hardware and Software stack
  • Simplify system design with alternative architecture
  • Includes SOC design, Bus Specifications, Reference Designs, FW OS and Open Source Hypervisor
  • Little Endian Linux to ease the migration of software to POWER
– Driving an expansion of enterprise class Hardware and Software stack for the data center
– Building a complete ecosystem to provide customers with the flexibility to build servers best suited to the Power architecture
Building collaboration and innovation at all levels

Implementation / HPC / Research

System / Software / Services

I/O / Storage / Acceleration

Boards / Systems

Chip / SOC

59 Members as of 10/1/2014: Complete member list at www.openpowerfoundation.org
Accelerated Technical Computing Vision - GPUs

Next-Gen IBM Supercomputers and Enterprise Servers

Long term roadmap integration

POWER CPU + NVIDIA GPU
Power Processor Technology Roadmap

POWER5/5+
130/90 nm
- Dual Core
- Enhanced Scaling
- SMT
- Distributed Switch +
- Core Parallelism +
- FP Performance +
- Memory Bandwidth +
- Virtualization

POWER6/6+
65/65 nm
- Dual Core
- High Frequencies
- Virtualization +
- Memory Subsystem +
- Altivec
- Instruction Retry
- Dynamic Energy Mgmt
- SMT +
- Protection Keys

POWER7/7+
45/32 nm
- Eight Cores
- On-Chip eDRAM
- Power-Optimized Cores
- Memory Subsystem ++
- SMT++
- Reliability +
- VSM & VSX
- Protection Keys+

POWER8
- Twelve Cores
- SMT++
- Reliability ++
- FPGA Support
- Transactional Memory
- CAPI
- Encryption
- PCIe Acceleration

POWER9
- Extreme Analytics Optimization
- Extreme Big Data Optimization
- On-chip accelerators

Future
- On-chip accelerators
Data Centric Systems (DCS) Designed to Minimize Data Motion

The DCS architecture will appeal to segments experiencing an explosion of data

**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**
Next-generation supercomputer coming to Lawrence Livermore National Lab

Lawrence Livermore National Laboratory (LLNL) today announced a contract with IBM to deliver a next-generation supercomputer in 2017. The system, to be called Sierra, will serve the National Nuclear Security Administration’s (NNSA) Advanced Simulation and Computing (ASC) program. Exascale supercomputers, expected in the next decade, will be about 1,000 times more powerful than today’s petaflops (quadrillions of operations per second) systems, with exascale operating at a quintillion operations per second. Under the contracts, Livermore and Oak Ridge will work with IBM, NVIDIA and Mellanox to deploy systems of about 150 petaflops to advance science and ensure national security.

Oak Ridge to acquire next generation supercomputer

The U.S. Department of Energy’s (DOE) Oak Ridge Leadership Computing Facility (OLCF) has signed a contract with IBM to bring a next-generation supercomputer to Oak Ridge National Laboratory (ORNL). The OLCF’s new hybrid CPU/GPU computing system, Summit, will be delivered in 2017.

The system’s vendor, IBM, and major component suppliers, NVIDIA and Mellanox, are all participating in an open architecture technology collaboration known as the OpenPOWER Foundation. Summit will feature more than 3,400 nodes, each with multiple IBM POWER9 processors and multiple NVIDIA Volta GPUs CPUs and GPUs completely connected with high speed NVLink.

The file system will be a GPFS Storage Server system with 1TB/s I/O bandwidth and 120 PB of disk capacity. IBM HPC software including Linux, Platform Computing LSF scheduler, resource manager, system management, and GPFS parallel file system.

Except from OLCF and LLNL Press release material