IBM Connect 2015

Seize the Moment

06/02/15

The OpenPOWER Foundation

How IBM, Google and 120 others are reinventing the datacenter

Franz Bourlet
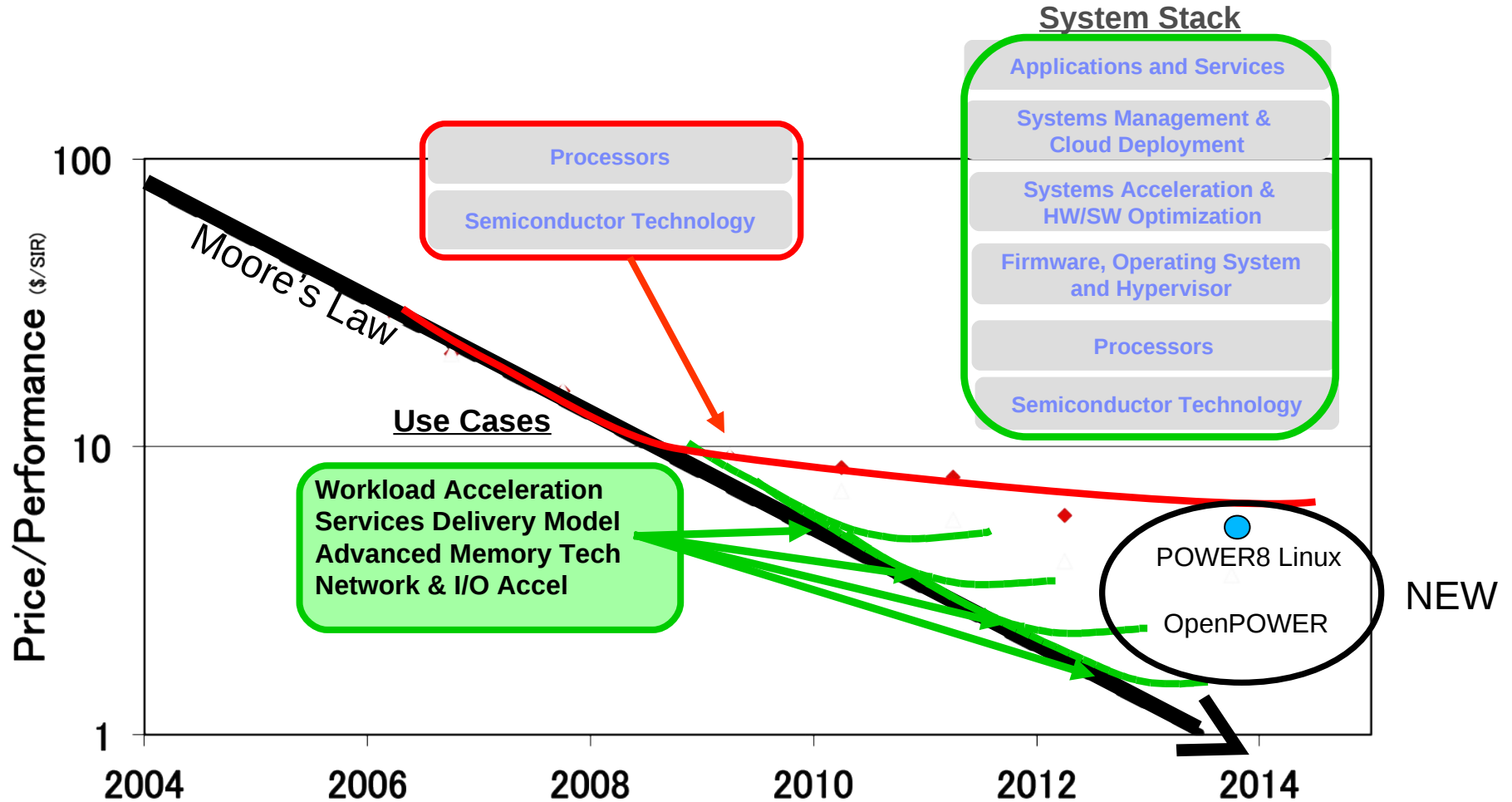
Power Systems Technical Sales

IBM Belgium & Luxembourg

New ways of acquiring servers

The number of companies designing and building servers is increasing :

- Traditionally there have been few companies designing systems: HP, IBM, SUN, Dell, etc.

- Today there are many more: Google, Microsoft, Facebook, Rackspace, Huawei, Sugon, Inspur, etc.

- A fairly mature ecosystem including the Taiwanese ODMs is a key enabler of this trend

Industry trends drive innovation beyond the chip

Microprocessors alone no longer drive sufficient Cost / Performance improvements

System Stack

- **Applications and Services**
- **Systems Management & Cloud Deployment**
- **Systems Acceleration & HW/SW Optimization**
- **Firmware, Operating System and Hypervisor**
- **Processors**
- **Semiconductor Technology**

Processors

Semiconductor Technology

Use Cases

**Workload Acceleration
Services Delivery Model
Advanced Memory Tech
Network & I/O Accel**

Moore's Law

Price/Performance ($/SIR)

100

10

1

2004   2006   2008   2010   2012   2014

POWER8 Linux

OpenPOWER
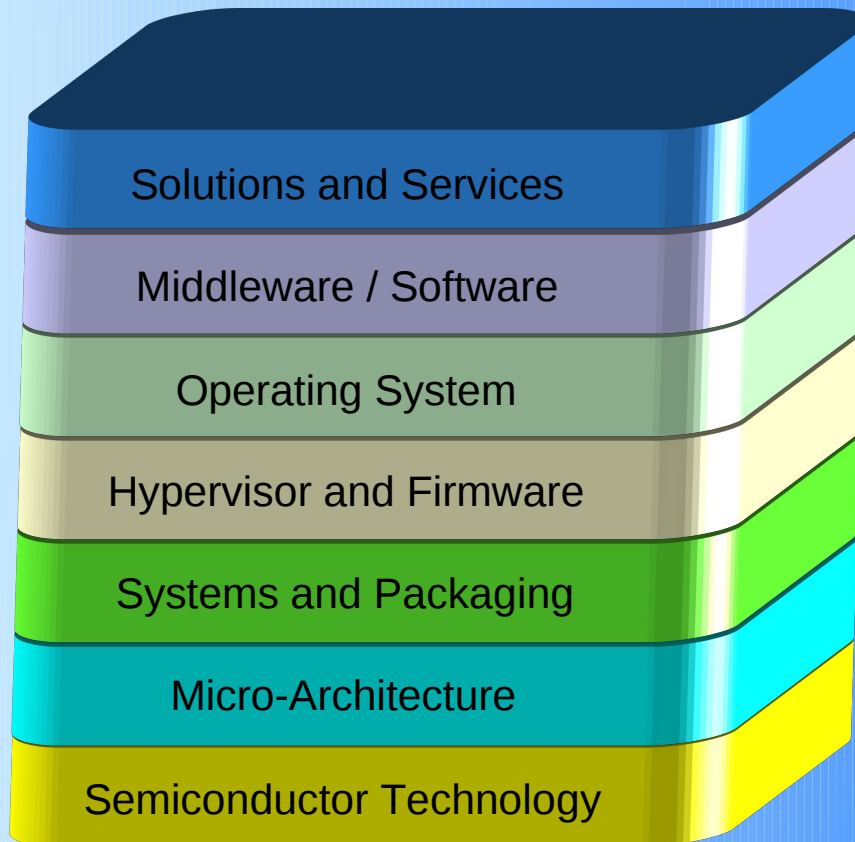
NEW

System stack innovations are required to drive Cost/Performance

ARM Ltd : another economic model

- Founded in November 1990 – www.arm.com

  – Spun out of Acorn Computers

- Runs 95% of mobile phones and tabs

- Arm designs a range of RISC processor cores but does not fabricate silicon itself

- Licenses ARM core designs to semiconductor partners who fabricate and sell to their customers

  – Examples include Apple, Samsung, Nvidia, Qualcomm,...

- Also develops technologies to assist with the design-in of the ARM architecture

  – Software tools, boards, debug hardware, application software, bus architectures, peripherals, etc

POWER has evolved into the ARM of the datacenter : from this...

IBM Research: Innovation Strengthens the Stack

Solutions and Services

Middleware / Software

Operating System

Hypervisor and Firmware

Systems and Packaging

Micro-Architecture

Semiconductor Technology

IBM is Unique in the World
- No other company plays
  in all the layers. IBM is
  world-class in all layers

- Those that play in multiple
  layers are world-class only
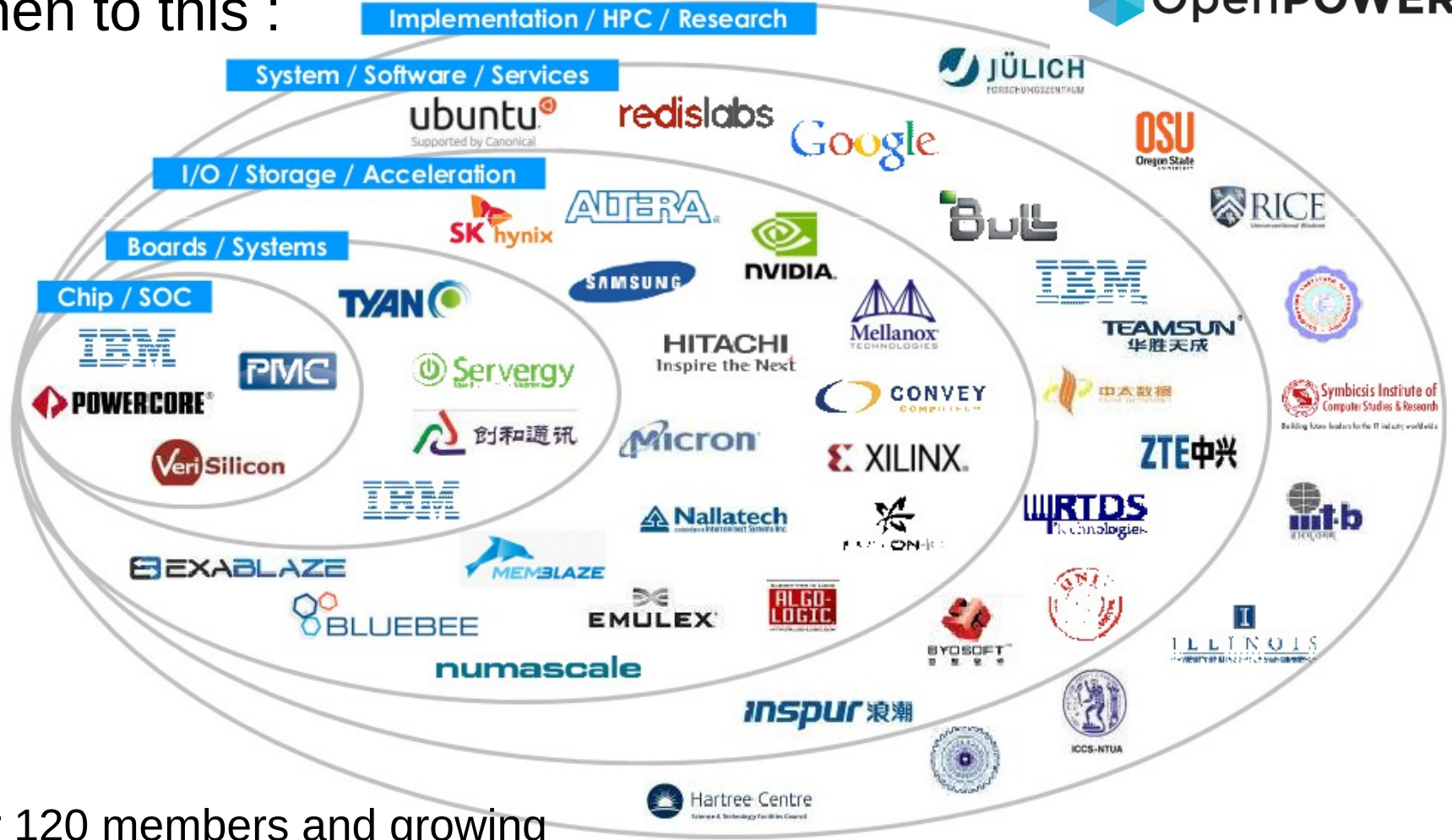 in few.

IBM R&D investment
- $6 Billion annually
- Human capital and culture
- 3000+ Researchers
- 6478 U.S. patents in 2012
- Patent leader past 20 years

As CMOS scaling slows, Value comes from Innovating across the Stack
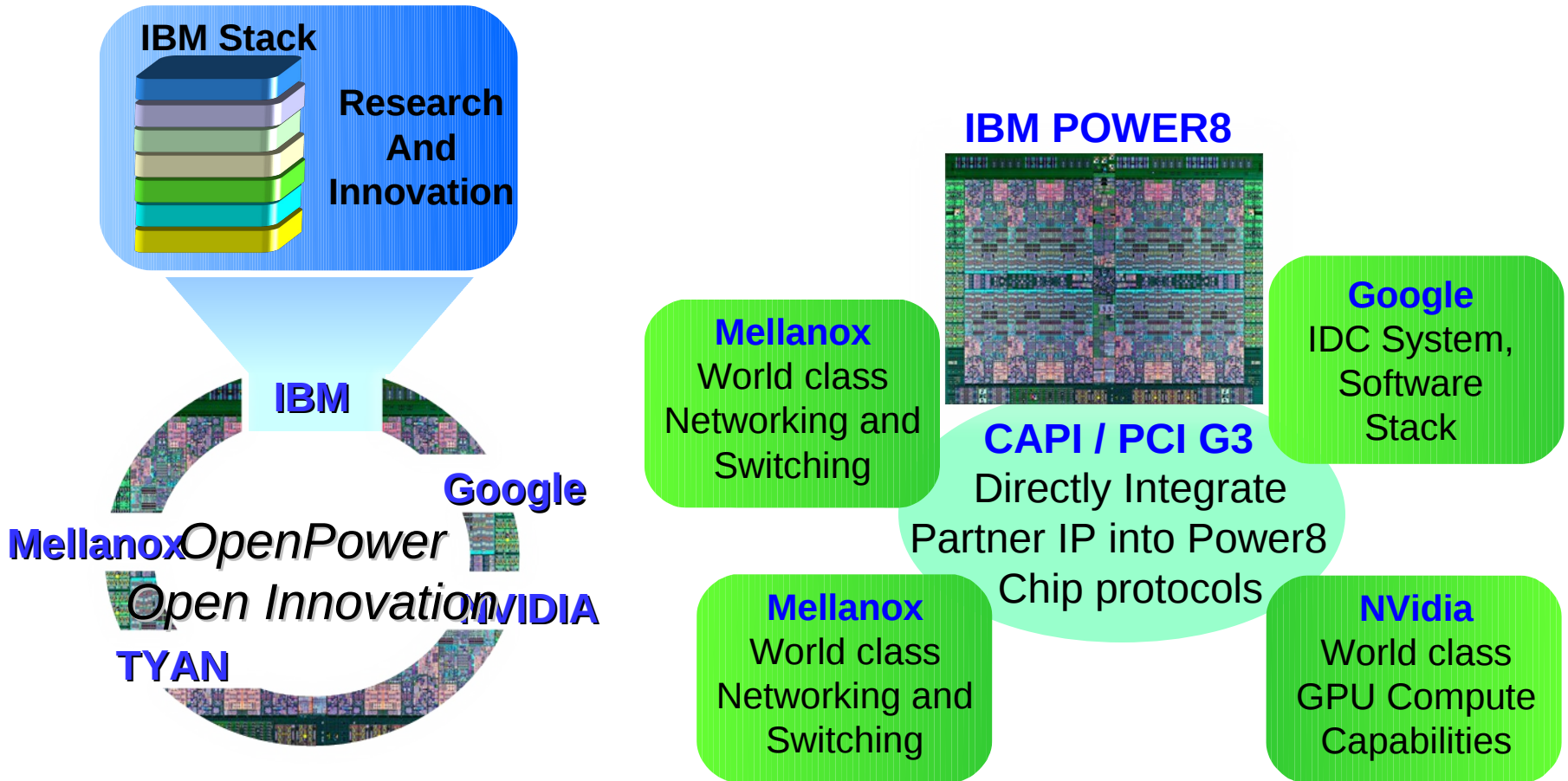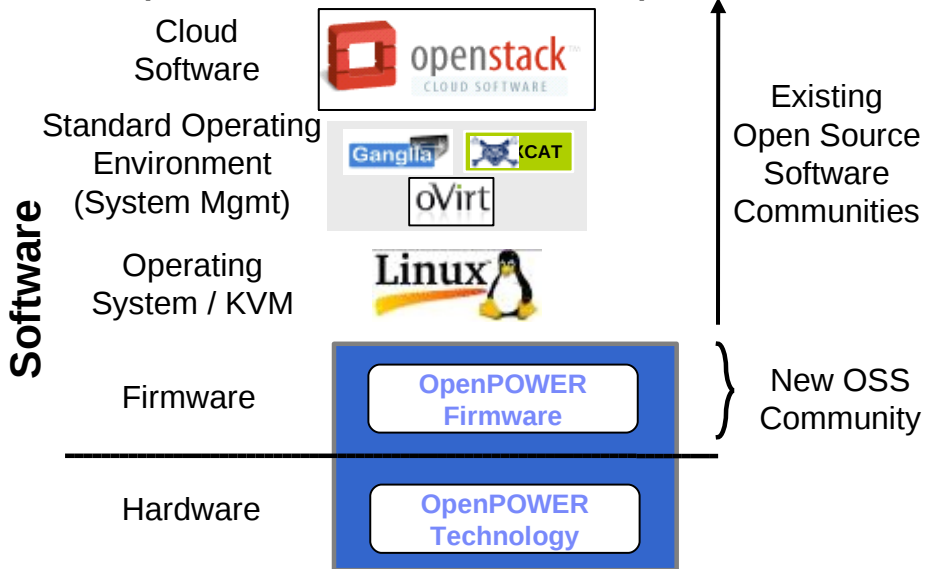
Then to this :

Over 120 members and growing

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.

The most open server architecture in the industry
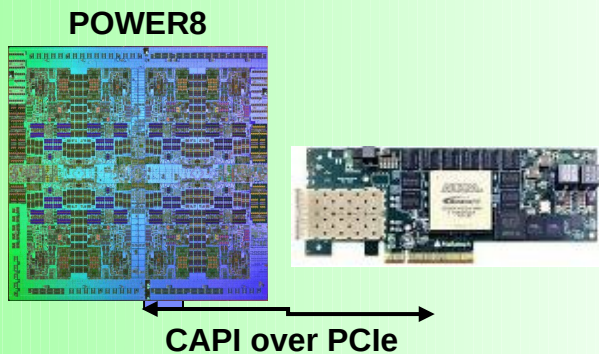
Power Open Source Software Stack Components

Software

Layer	
Cloud Software	openstack CLOUD SOFTWARE
Standard Operating Environment (System Mgmt)	Ganglia, xCAT, oVirt
Operating System / KVM	Linux
Firmware	**OpenPOWER Firmware**
Hardware	**OpenPOWER Technology**

Existing Open Source Software Communities

} New OSS Community

System Operating Environment Software Stack

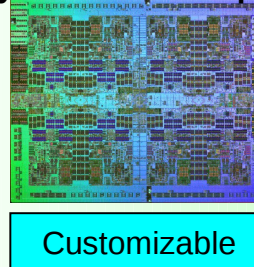A modern development environment is emerging based on this type of tools and services

Bootstrap, node, SOCKET IO, redis, Cassandra, RabbitMQ, NGiNX, mongoDB, MariaDB, HAPROXY

Multiple Options to Design with POWER Technology Within OpenPOWER

Hardware

POWER8

CAPI over PCIe

Framework to Integrate System IP on Chip

Customizable

Industry IP License Model

"Standard POWER Products" – 2014

"Custom POWER SoC" – Future

POWER8 vs Intel Haswell

	Sandy Bridge EP E5-x6xx	Ivy Bridge EP E5-26xx v2	Ivy Bridge EX E7-88xx v2	Haswell EP E5-26xx v3	POWER 7+ Systems	POWER8
Clock rates (GHz)	1.8–3.6	1.7-3.7	1.9-3.4	1.6-3.5	3.1–4.4 GHz	3.0-4.1 GHz
SMT options	1,2*	1, 2*	1, 2*	1, 2*	1, 2, 4	1, 2, 4, 8
Max Threads / sock	16	24	30	36	32	96
Max L1 Cache	32KB	32KB*	32KB*	64 KB	32KB	64KB
Max L2 Cache	256 KB	256 KB	256 KB	256KB	256 KB	512 KB
Max L3 Cache	20 MB	30 MB	37.5 MB	45 MB	80 MB	96 MB
Max L4 Cache	0	0	0	0	0	128 MB
Memory Bandwidth	31.4-51.2 GB/s	42.6-59.7 GB/s	68-85** GB/s	51-68 GB/s	100 – 180 GB/sec	230 - 410 GB/sec

1 = The Micro-architecture features above contributed to premium performance for the POWER8 systems compared to Ivy Bridge systems
 * Intel calls this Hyper-Threading Technology (No HT and with HT)
 32KB running in "Non-RAS mode" 16KB results in better RAS
 85GB running in "Non-RAS mode" and dual-device error NOT supported

POWER8 vs Intel Haswell

	Sandy Bridge EP E5-x6xx	**Ivy Bridge EP** E5-26xx v2	**Ivy Bridge EX** E7-88xx v2	**Haswell EP** E5-26xx v3	**POWER 7+ Systems**	**POWER8**
Clock rates (GHz)	1.8–3.6	1.7-3.7	1.9-3.4	1.6-3.5	3.1–4.4 GHz	3.0-4.1 GHz
SMT options	1,2*	1, 2*	1, 2*	1, 2*	1, 2, 4	1, 2, 4, 8
Max Threads / sock	16	24	30		32	96
Max L1 Cache	32KB	32			32KB	64KB
Max L2 Cache	256 KB				256 KB	512 KB
Max L3 Cache	20 MB				80 MB	96 MB
Max L4 Cache	0				0	128 MB
Memory Bandwidth	31.4-51.2 GB/s	42.6-59.7 GB/s	GB/s	51-68 GB/s	100 – 180 GB/sec	230 - 410 GB/sec

How fast the engine

1 = The Micro-architecture features above contributed to premium performance for the POWER8 systems compared to Ivy Bridge systems
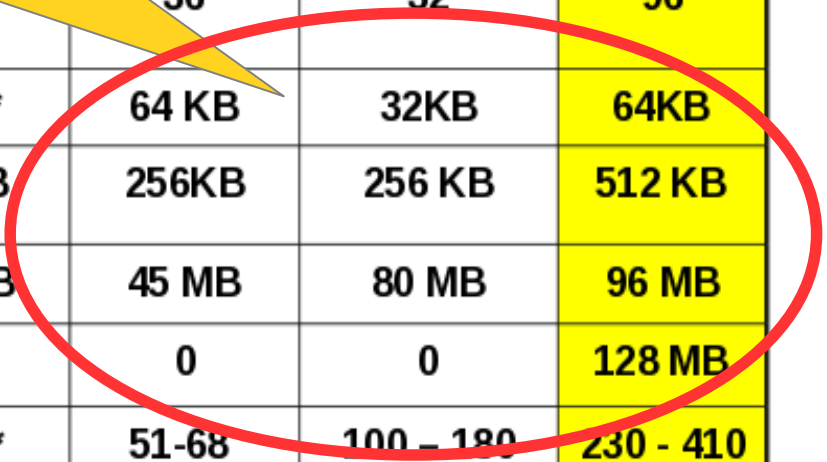 * Intel calls this Hyper-Threading Technology (No HT and with HT)
 ***32KB running in "Non-RAS mode" 16KB results in better RAS**
 85GB running in "Non-RAS mode" and dual-device error NOT supported

POWER8 vs Intel Haswell

	Sandy Bridge EP E5-x6xx	Ivy Bridge EP E5-26xx v2	Ivy Bridge EX E7-88xx v2	Haswell EP E5-26xx v3	POWER 7+ Systems	POWER8
Clock rates (GHz)	1.8–3.6	1.7-3.7	1.9-3.4	1.6-3.5	3.1–4.4 GHz	3.0-4.1 GHz
SMT options	1,2*	1, 2*	1, 2*	1, 2*	1, 2, 4	1, 2, 4, 8
Max Threads / sock	16	24	20	36	32	96
Max L1 Cache	32KB	32			32KB	64KB
Max L2 Cache	256 KB				256 KB	512 KB
Max L3 Cache	20 MB				80 MB	96 MB
Max L4 Cache	0				0	128 MB
Memory Bandwidth	31.4-51.2 GB/s	42.6-59.7 GB/s	GB/s	51-68 GB/s	100 – 180 GB/sec	230 - 410 GB/sec

How large the team

1 = The Micro-architecture features above contributed to premium performance for the POWER8 systems compared to Ivy Bridge systems
 * Intel calls this Hyper-Threading Technology (No HT and with HT)
 32KB running in "Non-RAS mode" 16KB results in better RAS
 **85GB running in "Non-RAS mode" and dual-device error NOT supported

POWER8 vs Intel Haswell

	Sandy Bridge		Bridge	Haswell EP E5-26xx v3	POWER 7+ Systems	POWER8
Clock rates (GHz)				1.6-3.5	3.1–4.4 GHz	3.0-4.1 GHz
SMT options				1, 2*	1, 2, 4	1, 2, 4, 8
Max Threads / sock	16			36	32	96
Max L1 Cache	32KB	32KB*	32KB*	64 KB	32KB	64KB
Max L2 Cache	256 KB	256 KB	256 KB	256KB	256 KB	512 KB
Max L3 Cache	20 MB	30 MB	37.5 MB	45 MB	80 MB	96 MB
Max L4 Cache	0	0	0	0	0	128 MB
Memory Bandwidth	31.4-51.2 GB/s	42.6-59.7 GB/s	68-85** GB/s	51-68 GB/s	100 – 180 GB/sec	230 - 410 GB/sec

How abundant the food

1 = The Micro-architecture features above contributed to premium performance for the POWER8 systems compared to Ivy Bridge systems
 * Intel calls this Hyper-Threading Technology (No HT and with HT)
32KB running in "Non-RAS mode" 16KB results in better RAS
85GB running in "Non-RAS mode" and dual-device error NOT supported

POWER8 vs Intel Haswell

	Sandy Bridge EP E5-x6xx	Ivy Bridge EP E5-26xx v2	Ivy Bridge EX E7-88xx v2	Haswell EP E5-26xx v3	POWER 7+ Systems	POWER8
Clock rates (GHz)	1.8–3.6	1.7-3.7	1.9-3.4	1.6-3.5	3.1–4.4 GHz	3.0-4.1 GHz
SMT options	1,2*		1, 2*	1, 2*	1, 2, 4	1, 2, 4, 8
Max Threads / sock				36	32	96
Max L1 Cache				64 KB	32KB	64KB
Max L2 Cache				256KB	256 KB	512 KB
Max L3 Cache	20			45 MB	80 MB	96 MB
Max L4 Cache	0		0		0	128 MB
Memory Bandwidth	31.4-51.2 GB/s	42.6-59.7 GB/s	68-85** GB/s	51-68 GB/s	100 – 180 GB/sec	230 - 410 GB/sec

How fast the data

1 = The Micro-architecture features above contributed to premium performance for the POWER8 systems compared to Ivy Bridge systems
 * Intel calls this Hyper-Threading Technology (No HT and with HT)
 **32KB running in "Non-RAS mode" 16KB results in better RAS
 **85GB running in "Non-RAS mode" and dual-device error NOT supported

In a nutshell

IBM SWG : 70 PVU on all POWER systems running Linux

- Applies to SWG PVU licensed software products for **ALL** Power cores **running Linux**

- 70 PVU pricing previously only available for 1-2 socket servers
  - PowerLinux 7R1/7R2, p260/p270 and Power 710/720/730/740
  - Intel SandyBridge/IvyBridge servers

- 70 PVU pricing includes:
  - Power IFLs
  - Power S8xx
 - Power E8xx
  - Flex System

- Applies to POWER8 and POWER7

PVU Table per Core (section 1 of 2 - RISC and System z)

Processor Technologies										
	Processor Brand				Processor Type					
Processor Vendor	Processor Name	Server model numbers	Maximum number of sockets per server	Cores per socket					IFL Engine	Proc. Model Number	PVUs per Core
				(1)	(2)	(4)	(6)	(8)	(12)		
	POWER Systems cores running Linux OS	7R1, 7R2, 7R4 POWER IFL, p24L Any POWER System core running Linux	All			■	■	■	■	All	70
POWER7 [4]		770, 780, 795	>4			■	■	■		All	120
		750, 755, 760, 775, PS704, p460, Power ESE	4			■	■	■		All	100
		PS700-703, 710-740, p260, p270	2			■	■	■		All	70
	POWER6	550, 560, 570, 575, 595	All	■						All	120
IBM		520, JS12, JS22, JS23, JS43	All	■						All	80

PVU Table link: http://www-01.ibm.com/software/lotus/passportadvantage/pvu_licensing_for_customers.html

What's in it for me?

Consider a mix of workloads requiring 250 cores on x86 :
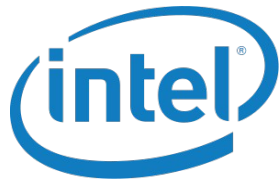
intel

Power™

- <u>Scale out approach (2-socket)</u> :   - <u>Scale out approach (2-socket)</u> :

70 PVU X 250 =   **17.500**   70 PVU X 100 =   **7.000**

What's in it for me?

Consider a mix of workloads requiring 250 cores on x86 :

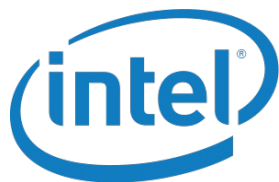- Scale out approach (2-socket) :

70 PVU X 250 =   **17.500**

- Optimized scale out (4-socket) :

100 PVU X 250 =   **25.000**

- Scale out approach (2-socket) :

70 PVU X 100 =   **7.000**

- Optimized scale out (4-socket) :

70 PVU X 100 =   **7.000**

What's in it for me?
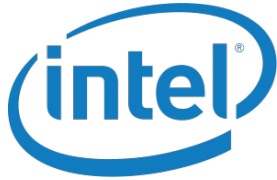
Consider a mix of workloads requiring 250 cores on x86 :

- Scale out approach (2-socket) :

70 PVU X 250 =   **17.500**

- Optimized scale out (4-socket) :

100 PVU X 250 =   **25.000**

- Scale up approach (> 4-socket) :

120 PVU X 250 =   **30.000**

- Scale out approach (2-socket) :

70 PVU X 100 =   **7.000**

- Optimized scale out (4-socket) :

70 PVU X 100 =   **7.000**

- Scale up approach (> 4-socket) :

70 PVU X 100 =   **7.000**

POWER8 combined with business accelerators

Altera FPGA acceleration and IBM CAPI

Monte Carlo 250x faster than POWER8 core alone, reduced C code 40x over non-CAPI FPGA

ALTERA. IBM

Data Engine for NoSQL 24:1 server consolidation, 3x lower cost per user, 40TB CAPI-attached flash

ubuntu® redislabs

CAPI dev kit with FPGA card from Nallatech

Nallatech   OpenCL

US Dept of Energy $325M super computing contract awarded to IBM, Mellanox, and NVIDIA

DoE systems for science and stockpile stewardship

Sierra and Summit systems to be >100 PF, 2 GB/core main memory, local NVRAM, and science performance 4x-8x Titan or Sequoia

IBM
NVIDIA
Mellanox TECHNOLOGIES

OAK RIDGE National Laboratory   Lawrence Livermore National Laboratory

NVIDIA acceleration built into IBM Power S824L

8x faster than x86 Ivy Bridge on pattern extraction

82x faster for Cognos BI and DB2 BLU

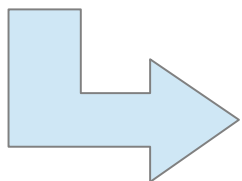NVIDIA

IBM

Tyan OpenPOWER Customer Reference System

TYAN'S OPENPOWER
CUSTOMER REFERENCE SYSTEM
INNOVATIVE, COLLABORATIVE AND OPEN

• Following the OpenPOWER Foundation's concept with hardware opened to community
• Based on POWER 8 Architecture
• First commercialized customer reference system provided by OpenPOWER ecosystem

OpenPOWER

Coherent attachment of accelerator cards

Power Processor

CAPI

FPGA

IBM Supplied POWER Service Layer

Function 0 | Function 1 | Function 2 | Function n

CAPP | **PCIe**

Typical I/O Model Flow

| DD Call | → | Copy or Pin Source Data | → | MMIO Notify Accelerator | → | Acceleration | → | Poll / Int Completion | → | Copy or Unpin Result Data | → | Ret. From DD Completion |
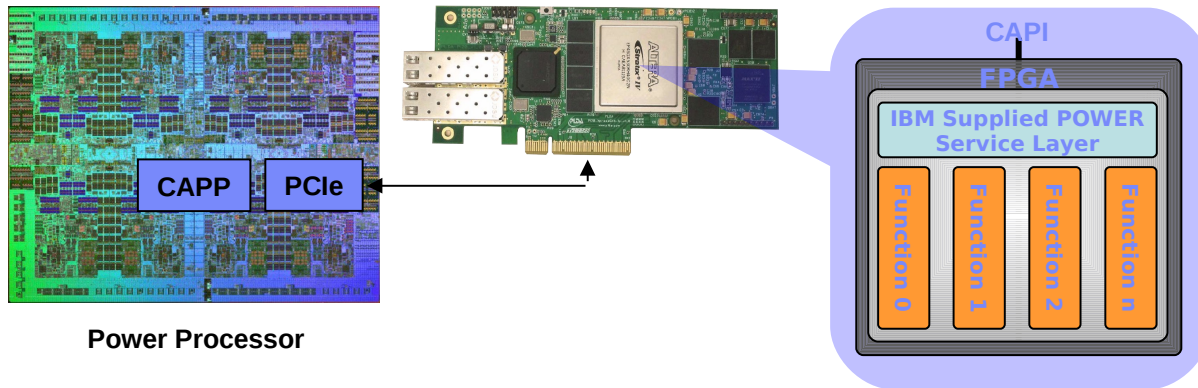
Flow with a Coherent Model

| Shared Mem. Notify Accelerator | → | Acceleration | → | Shared Memory Completion |

Faster application development

Faster application execution (40 to 1)

Power Systems

IBM

Comparable TCA

Linux on Intel Ivy Bridge + VMware

Vs.

Linux on POWER8 + with PowerVM

	Dell PowerEdge R720	HP ProLiant DL380 G8	IBM Power 822L
	$28,366	**$29,829**	**$29,264**
	redhat (intel) vmware' Ivy Bridge	redhat (intel) vmware' Ivy Bridge	redhat PowerVM IBM POWER8
Server list price* -3-year warranty, on-site	**$12,605**	**$14,068**	**$14,895**
Virtualization - OTC + 3yr. 9x5 SWMA	**$10,064** VMware vSphere Enterprise 5.1	**$ 10,064** VMware vSphere Enterprise 5.1	**$9,880** PowerVM for IBM PowerLinux
Linux OS list price - RHEL, 2 sockets, unlimited guests, 9x5, 3 yr. sub./ supp.	**$5,697** Red Hat subscription and Red Hat support	**$5,697** Red Hat subscription and Red Hat support	**$4,489** Red Hat subscription and IBM support
Total list price: (Total cost of acquisition)	$28,366	$29,829	$29,264
Server model	Dell R720	HP Proliant DL380p G8	IBM Power 822L
Processor / cores	**Two 2.7 GHz , E5-2697, Ivy Bridge, 12-core processors**		**Two 3.4 GHz POWER8, 10-core**
Configuration	64 GB memory, 2 x 300GB 15k HDD, 10 Gb two port		Same memory, HDD, NIC

* Based on US pricing for Power S822L announcing on April 28, 2014 matching configuration table above. Source: hp.com, dell.com, vmware.com

New POWER economical model

BLACK BOX

Customer requirements

ODM
Design
Assistance

IBM Developed
IBM HW & FW
AIX, I, & LINUX
POWERVM
New IP

Core Set of Building Blocks

IBM Bring Up
And Verification

IBM Test

IBM Logo system
Sold to IBM's customer

IBM Customer

IBM Logo
Target - Power existing customers
AIX, I, & LINUX

GREY BOX

Customer Defined

OpenPOWER™

IBM specifications

Partner Developed
Partner HW
BMC
Open Source Code
KVM & LINUX

Reference
Material

IBM
Enablement

IBM Test

IBM Logo system
Sold to IBM's customer

IBM Customer

IBM Logo
Target – HPC & MSP
OPAL / KVM / LINUX

WHITE BOX

Customer Defined

Partner Developed
Partner HW
BMC
Open Source Code
KVM & LINUX

Partner Logo system
Sold to Partner's customer

Partner's Customer

Partner Logo
Target – Warehouse Data Centers
OPAL / KVM / LINUX

Linux support for POWER

IBM

redhat / fedora

- RHEL 7
  - POWER8 (native mode) and POWER 7/7+
  - Available June 2014
- RHEL 6
  - POWER8 supported with U5 (P7-compatibility mode)
  - Full support of POWER6 and POWER7 (native mode)
- Fedora
  - Fedora 16 was first release to re-launch POWER
  - Fedora 20 has POWER8 support
- Supported add-ons
  - JBoss
  - High Performance Network Add-on

SUSE / openSUSE

- SLES 12
  - Anticipated to support POWER8 (native mode) and POWER 7/7+
  - Available October 2014
- SLES 11
  - POWER8 with SP3 (P7-compatibility mode)
  - POWER7+ encryption, RNG accelerators with SP3
  - Full support of POWER7 (native mode)
- openSUSE
  - openSUSE 12.2 re-launched for IBM POWER
  - openSUSE 13.2 includes POWER8 support
- Supported add-ons
  - SUSE Linux Enterprise High Availability Extension

ubuntu® / debian

- Ubuntu 14.10
  - Continued support for POWER8
  - Anticipate 4Q14 availability
- Ubuntu 14.04
  - POWER8 enabled (native mode)
  - No official support for POWER7+ and older systems
  - No support for 32-bit applications. 64-bit only.
  - Supported in KVM only at this time
  - LTS
- Supported add-ons
  - JuJu Charms
  - MaaS (Metal as a Service)
  - Landscape
- Debian
  - Debian community now supports Power as of

Built from the same source as x86
Delivered on the same schedule as x86
Supported at the same time as x86

Integrating POWER into the datacenter

PowerVM
PowerKVM
RHEV
Docker

Manage

Provision

Deploy

Procure

Platform Management
- Custom management
- Cloud based

OpenStack
Custom Scripts

Compute   Networking   Storage

Standard Hardware

PXE
IPMI

H/W Management
- Custom scripting
- Heavily Automated
- Open technologies

xCat
Puppet
Chef
Custom Scripts
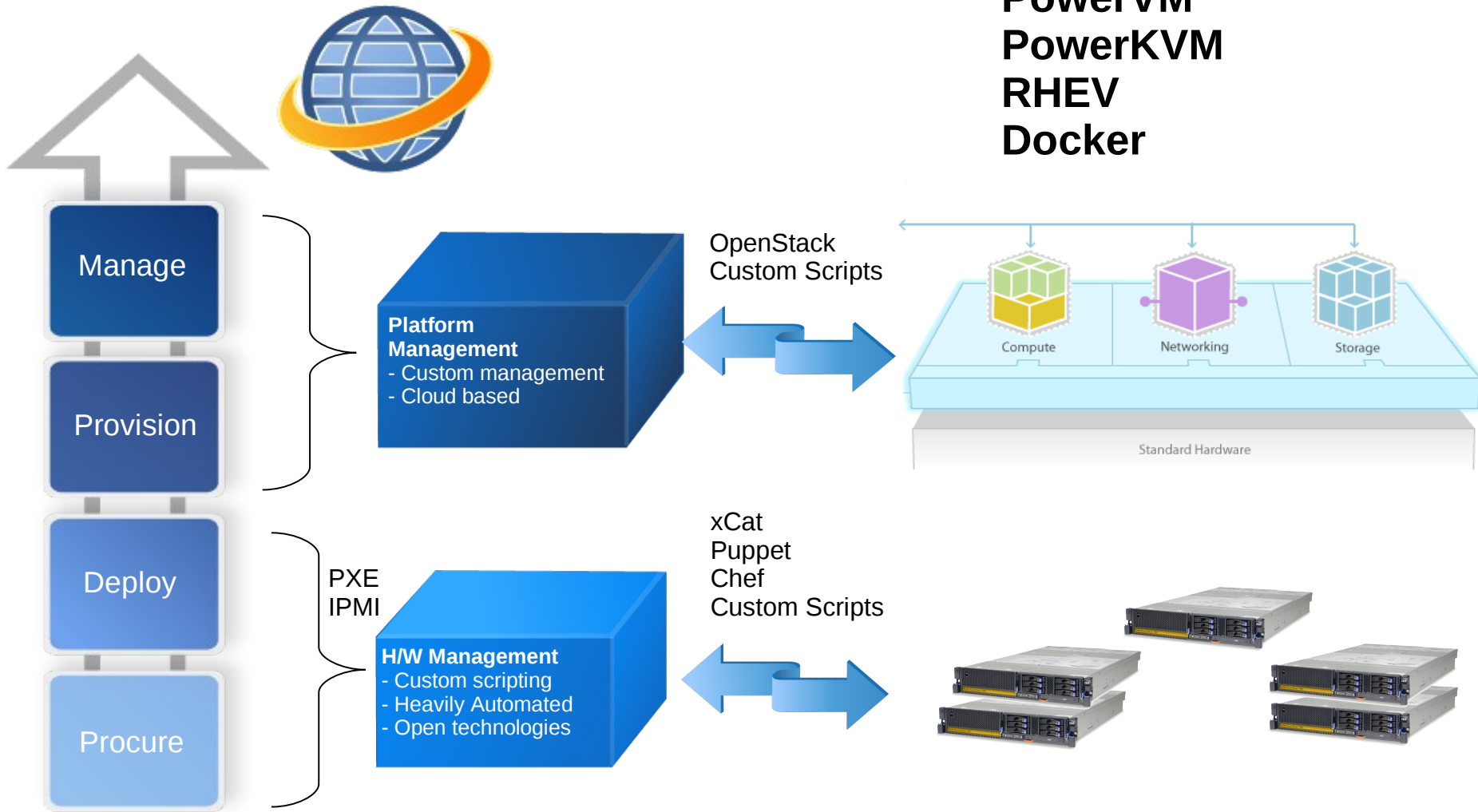
Hot off the press – OVH MSP to offer POWER8 platform

"OVH has become the first commercial partner outside of IBM to formally announce the deployment of POWER8 servers inside their data centre. These will deployed as bare metal servers for customers who have big data and analytics problems and want the most powerful solution that they can get access to. As part of this announcement, OVH will become an OpenPOWER Foundation member."

- OVH announcement
- Google announcement
- Softlayer announcement
- Rackspace announcement

Linux momentum is growing on Power Systems

An open ecosystem, using the POWER Architecture, to **share expertise, investments, and server-class intellectual property** to serve customers' evolving needs.

Linux Centers. Offering access to hardware technical support, porting assistance, demos, toolkits, hands-on labs

Beijing | Austin | New York | Montpellier | Tokyo

500+
ISVs enabled to run on Linux on Power

Global Linux ISV Development

NEW. Power Development Cloud gives Linux developers who want to prototype applications access to Power

NEW. 50,000 new developers from Ubuntu development community

Regional Ecosystem Initiative

Recruiting key solutions:
Open Source Tools
Middleware
Industry Solutions

$1B

in new Linux and open source technologies for IBM Power Systems.

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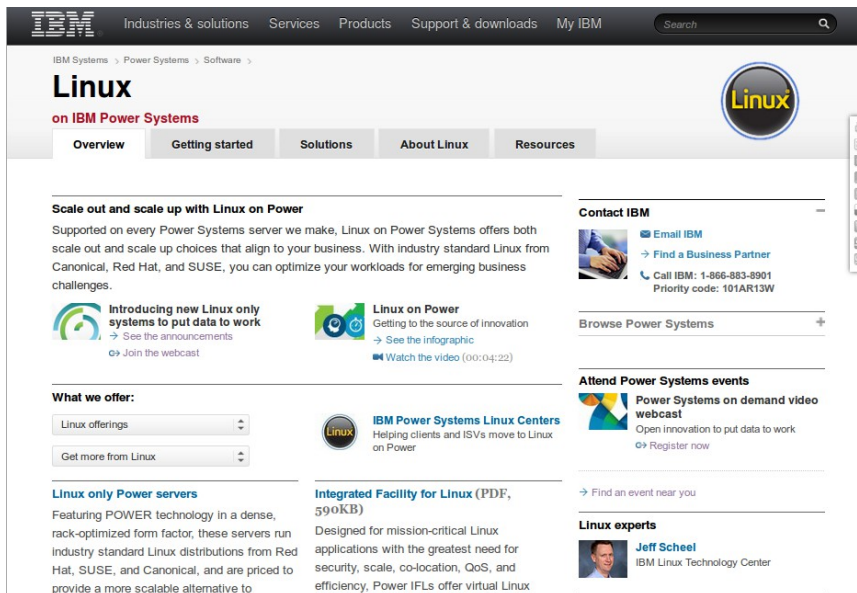
The goal of the OpenPOWER Foundation is
to create an open ecosystem,
using the POWER Architecture to share expertise,
investment, and
server-class intellectual property
to serve the evolving needs of customers.

http://openpowerfoundation.org/

Where to find more information?
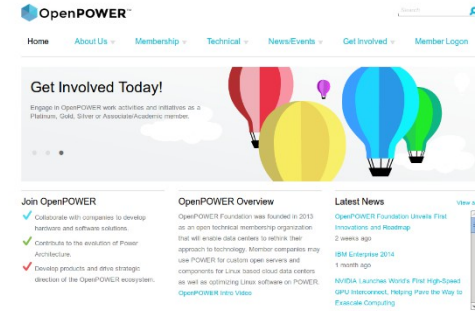
Power Systems Linux Portal

(Product Information)

The OpenPOWER Foundation

The Linux on Power Community

franz_bourlet@be.ibm.com