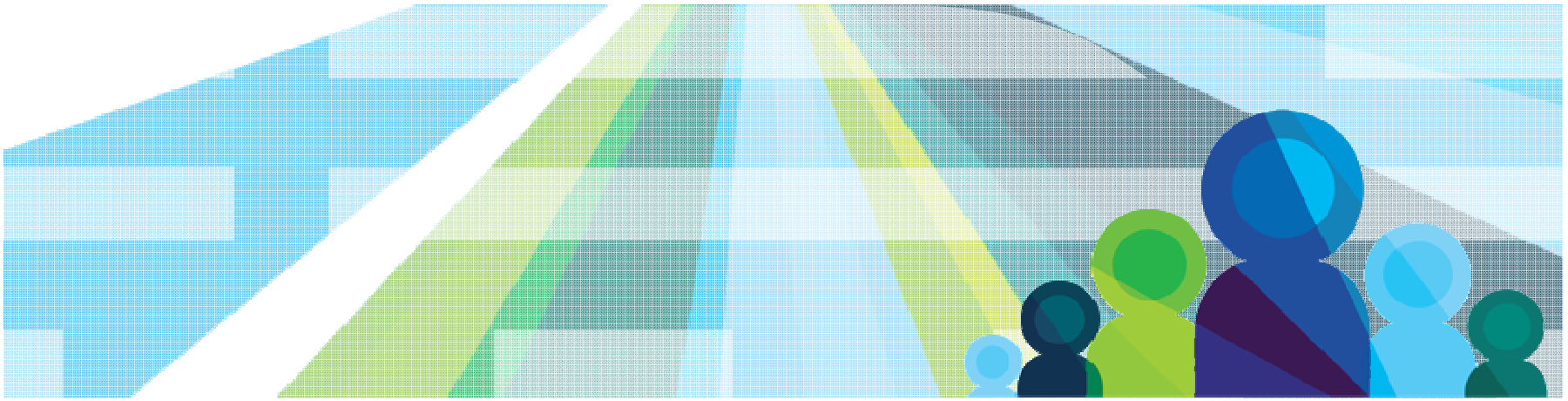

# Pourquoi Linux on Power ?

*Luca Comparini – EcoD Europe, Linux on Power Growth Initiative Leader*

*@lucacomparini*

*www.linkedin.com/in/lucacomparini*



# POWER platform: best known for...

**Design point is "purposeful"**
for new and existing Enterprise workloads (Analytics, Big Data, OLTP, ERP, SAP, etc.)

**Design Goals**
Increase the performance per core while reducing the cost per workload

**Optimization**
Integrate processor micro-architecture with hypervisor, OS, IBM middleware and storage

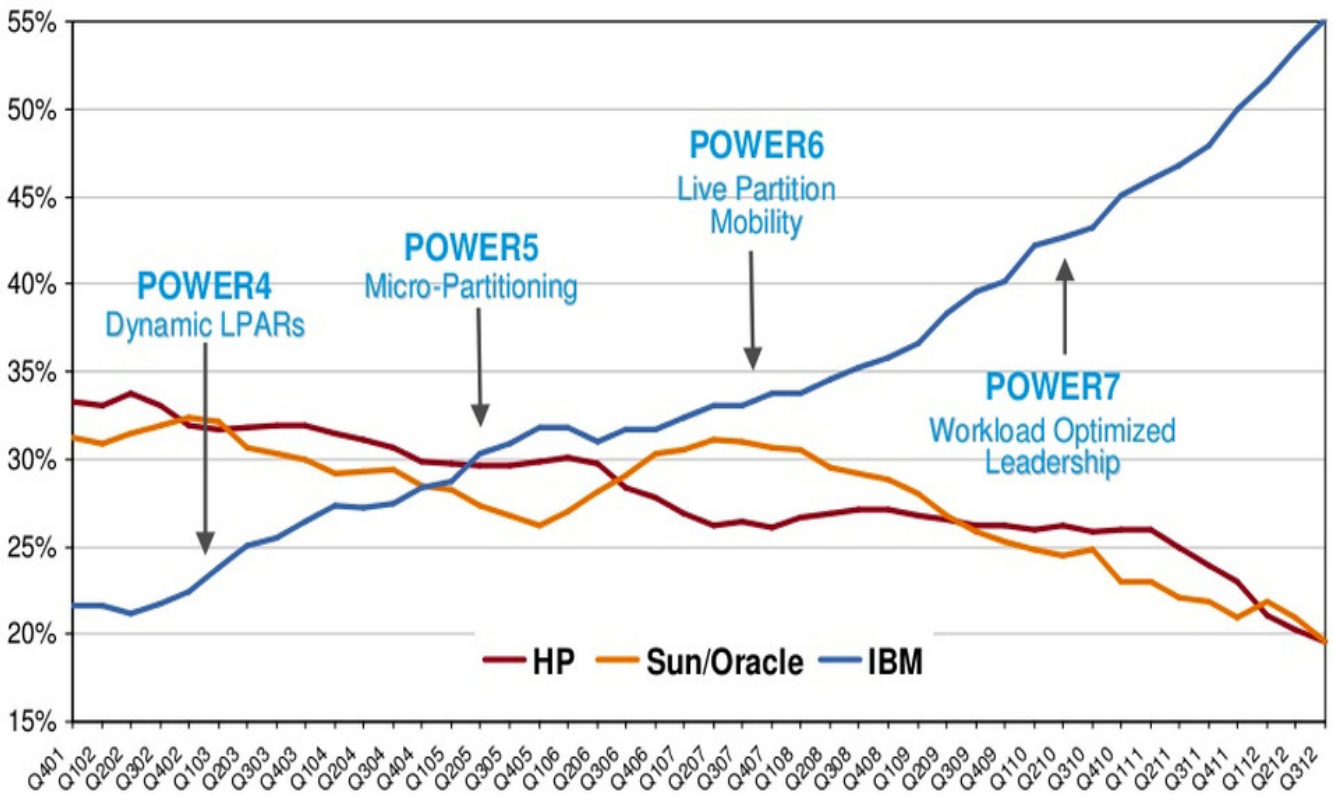
***Designed for enterprise reliability***

→ *Higher Performance per Core = less software licenses needed to accomplish the same amount of work*
→ *Less software licenses = lower cost per workload*

*Traditional, back-office applications and workloads. We call these "Systems of Record." If you are a retailer, think of your warehouse inventory records; if you are a healthcare provider, think of your patient records; if you are an insurer, your claims records, etc. In addition, there are traditional IT applications for customer relationship management, and finance and accounting systems, to name a few.*

Systems of Record

# IBM Power in the Unix market (market share)




http://www.idc.com/getdoc.jsp?containerId=IDC_P348

# Power is not best known for...

## Facts on Linux

Over 90% of world fastest supercomputers run on Linux

8 of the world's top 10 website run on Linux (Google, YouTube, Yahoo, Facebook, Twitter)

80% of all Stock Exchanges rely on Linux

US Department of Defense is the “single biggest install base for Red Hat Linux”

95% of the servers used by Hollywood animation films run on Linux

Systems of Engagement

*“Systems of Engagement.” For online retailers, customers can browse what is in their stores, recommend items and complete transactions via their phones; insurers have new mobile apps where customers can file accident claims via their mobile devices; and so on. These new mobile and social apps require an infrastructure where they can be quickly developed, deployed, modified, and redeployed. And then of course, you want to turn the data generated from these applications into insights that can be acted on.*

# Power8 – innovation for today

## More Cores

*Industry Best Practice*

12 processor cores per socket (50% more than before) that deliver better per core performance

**What this means**
Enjoy better scale up performance, and more throughput per scale out server node.

## More Threads

*Industry Leading*

**SMT8** – 8 dynamic threads per core, supporting SMT1, 2, 4, & 8 modes dynamically across VMs

**What this means**
You choose – Deploy VM's in the optimal SMT mode based on application needs.

## More Cache

*Industry Leading*

At 100MB, 3X the on-chip cache as POWER7 – plus 128MB of new off-chip cache as well

**What this means**
Memory-intensive applications (like database) will perform better as memory latency is enhanced.

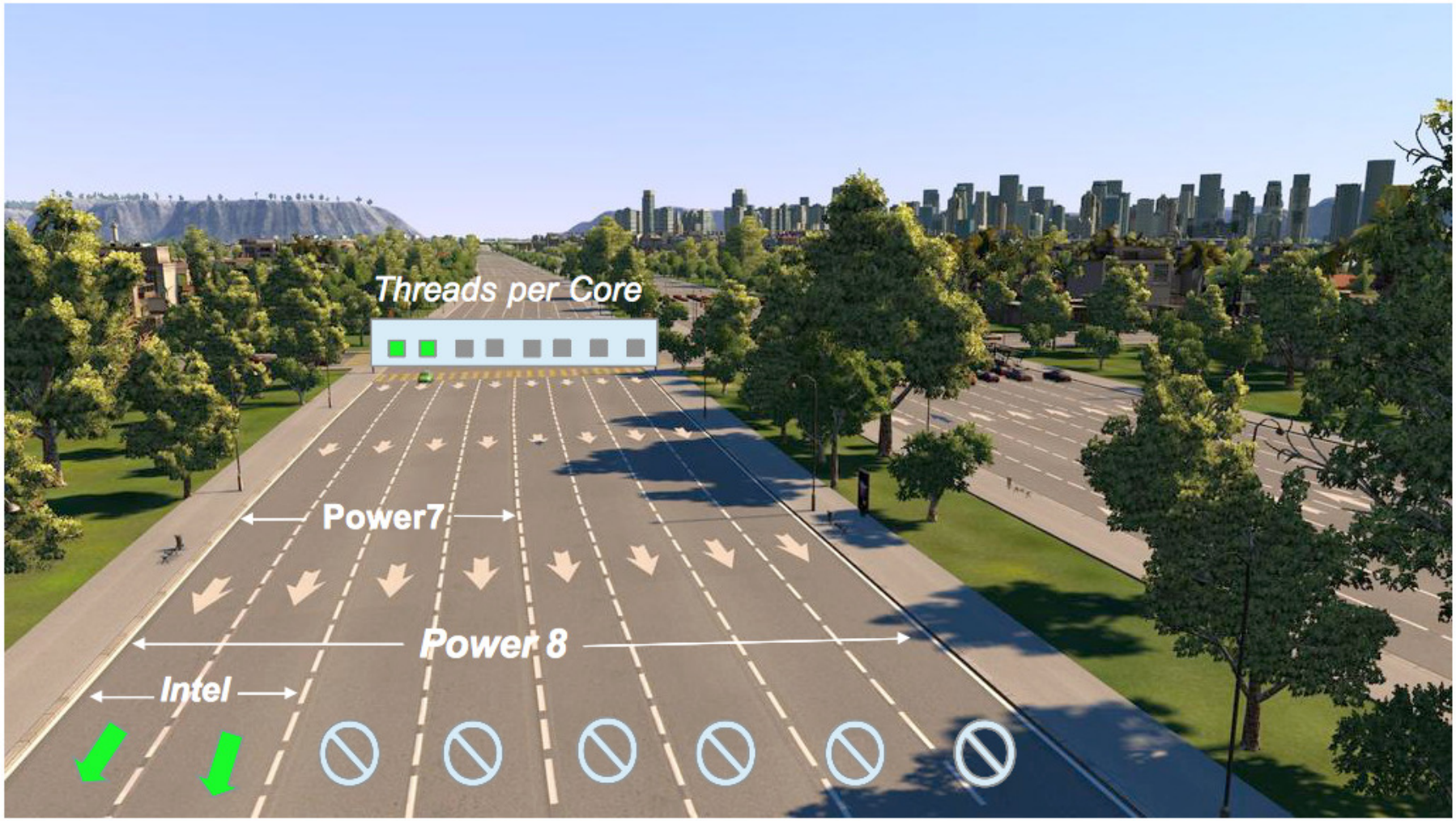
## More Bandwidth

*Industry Leading*

2.3X our prior gen (and 4.6X the competition) to memory, and 2.4X our prior gen to I/O.

**What this means**
Data-hungry applications (like big data & analytics) will respond twice as fast and scale more efficiently.

# Understanding the impact of “more threads”




Intel HyperThreading: 2 x Threads per Core
Power Simultaneous Multi-Threading: 1 / 2 / 4 / 8 x Threads per Core
+ 1.6 times single thread performance of Power7+

# Understanding the impact of “more bandwidth”

*Data is here (Disks)*


*Processors are here*



96 MB L3 cache
128 MB L4 cache
230GB/s sustained memory bandwidth

BUILT ON Power

# Unusual transports

*Unfortunately moving data is a different story*


# Unusual transports

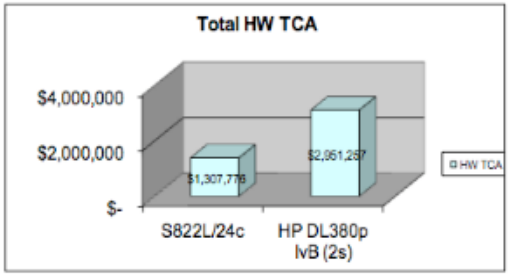
*Unfortunately moving data is a different story*

# Power8 vs Intel: geek comparison

| | Sandy Bridge EP | Ivy Bridge EP E5-26xx v2 | Ivy Bridge EX E7-88xx v2 | POWER 7+ Systems | POWER8 |
|---|---|---|---|---|---|
| Clock rates | 1.8–3.6GHz | 1.7-3.7GHz | 1.9-3.4 GHz | 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 | 16 | 24 | 30 | 32 | 96 |
| Max L1 Data Cache | 32KB | 32KB* | 32KB* | 32KB | 64KB |
| Max L2 Cache | 256 KB | 256 KB | 256 KB | 256 KB | 512 KB |
| Max L3 Cache | 20 MB | 30 MB | 37.5 MB | 80 MB | 96 MB |
| Max L4 Cache | 0 | 0 | 0 | 0 | 128 MB |
| Memory Bandwidth | 31.4-51.2 GB/s | 42.6-59.7 GB/s | 68-85** GB/s | 100 – 180 GB/sec | 230 - 410 GB/sec |

# The equation between performance and $$$

## *Do the same with less*




| Power S822L | HP DL 380 G8 |
|---|---|
| **TCA/TCO is for 34 servers, 816 cores** | **TCA/TCO is for 100 servers, 2400 cores** |
| **3x** better virtualized throughput vs. an HP 2 socket **Ivy Bridge** | 100 HP servers needed for ~ equal virtualized throughput of 34 Power S822L |
| 2S, 24 cores each | 2S, 24 cores each |
| POWER8, 3.0GHz | **Ivy Bridge**, 2.7GHz |
| PowerVM | VMware vSphere Ent |

***66% Less Systems and Cores***

*Lower SW Licenses Fees*
*Reduced Management Costs*
*Reduced Floor Space*

***58% Lower HW TCA***

*Vs Ivy Bridge with VMWare*

## *Do more with the same*

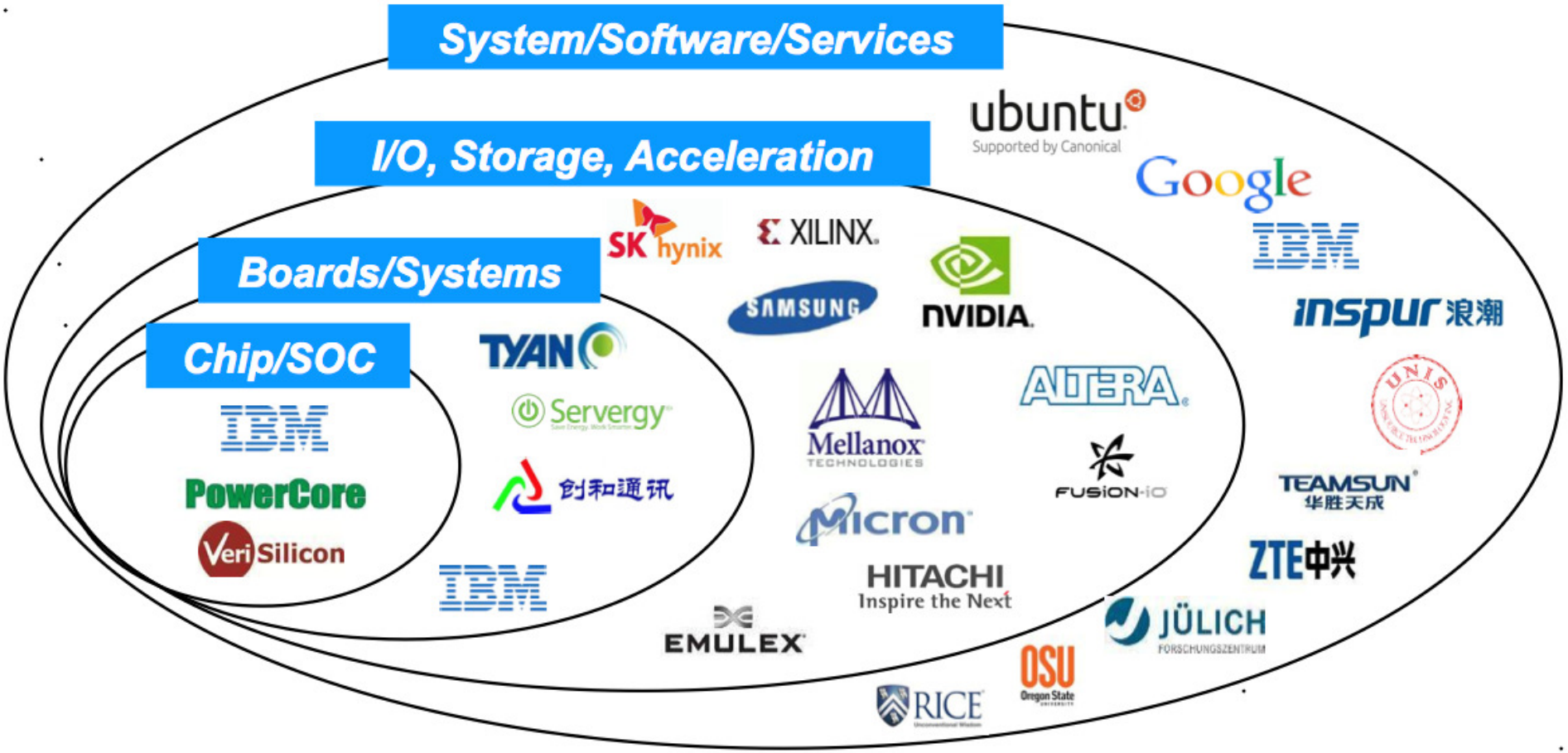
***Power8 delivers insights 82x faster***

*Running Cognos BI reports, with DB2 BLU acceleration vs Ivy Bridge with a traditional database*

- DB2 with BLU Acceleration on POWER8 for Cognos BI is 'Fast on Fast on Fast!'
- First processer designed for Big Data to run more concurrent queries in parallel faster, across multiple cores with more threads per core
  - **18x more throughput for simple reports**
  - **40x more throughput for intermediate reports**
  - **747x better throughput for complex reports**

Additional information available in The Edison Group paper at https://www.ibm.com/services/forms/signup.do?source=stg-web&S_PKG=ov23147

OpenPOWER™



# Non IBM Power8 products

The Tyan reference (ATX) board, SP010, measures 12" by 9.6"

- one single-chip module (SCM)
- four DDR3 memory slots
- four 6 Gb/sec SATA peripheral connectors
- two USB 3.0 ports
- two Gigabit Ethernet network interfaces
- keyboard and video
- intended for developers

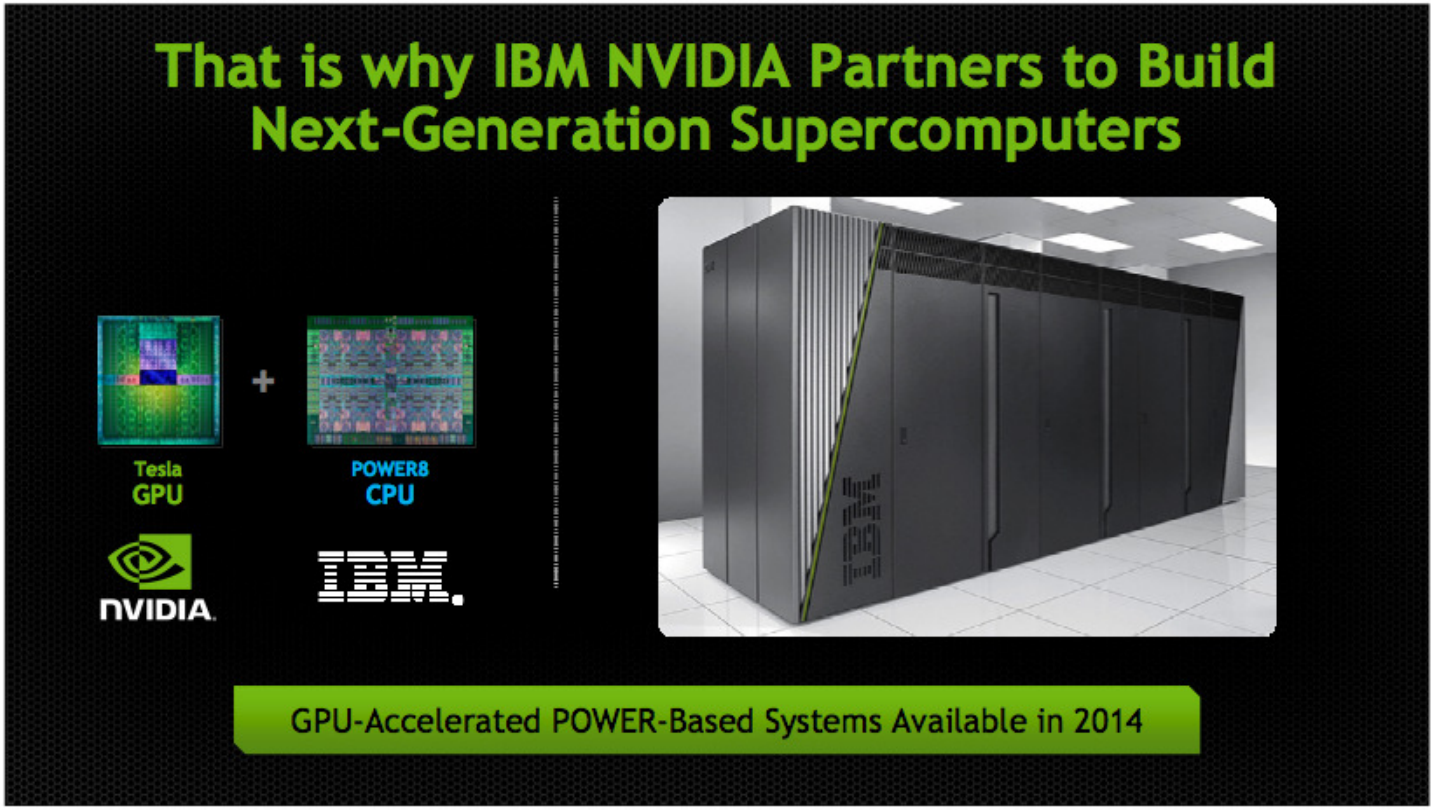
The Google reference board

- two single-chip module (SCM)
- four modified SATA ports
- Google use only

OpenPOWER™

NVIDIA

*http://www.computerworld.com/s/article/9241409/IBM_opens_up_Power_chip_design_partners_with_Google*




*One of the more interesting aspects of OpenPower is IBM and Google partnering with Nvidia, which is tuning its CUDA parallel programming toolkit for Power processors. Both could help resolve the issue around the software and make Power another good alternative, McGregor said.*


## Power 8 – CPU or Accelerator ?

| | Intel Xeon (E5-2697V2) | IBM Power 8 | NVIDIA K40 |
|---|---|---|---|
| Cores | 12 | 12 | 15 |
| Frequency (GHz) | 2.7 (3.5) | 4.116 | 0.745 |
| GFLOPS (SP/DP) | 259.2 / 129.6 | 790.3 / 395.1 | 4.29 / 1.43 |
| Threads | 24 | 96 | 240 |
| GB/s (peak) | 59.7 | 192 | 288 |
| *Vector unit (SP/DP)* | *8 / 4* | *4 / 2* | *32 / 32* |
| *Contexts (SP/DP)* | *24 / 24* | *96 / 96* | *1440 / 480* |

- Bandwidth of a GPU : entering the 200 GB/s range
- Flexibility of a CPU : small vector units
- Great FLOPS : 3x on the E5-V2.

Altimesh

How does it translate in real-world applications ?

## FINANCIAL SERVICES INDUSTRY USE-CASE

- Fixed cash flows pricer – accumulate discounts of cash flows with linear interpolation on the interest rate

$$\pi = \sum_{cash\ flows} N * e^{-T * r(T)}$$

$$r(T) = \frac{T - T_-}{T_+ - T_-} * r_+ + \frac{T_+ - T}{T_+ - T_-} * r_-$$

- Implementations: Java, Default (C++), optimized with FMA, optimized without FMA. All implementations have same algorithmic optimizations (precalculated lookups and interpolations).

Default implementation of exp seems to be the biggest performance blocker

# Growing the ecosystem

*Success factors*

- **Better TCA**
- **Better TCO** (thanks to “do even more with less”)
- **Even better TCO** for IBM Power Linux solutions (for aggressive PVU)

- **Choice of Hypervisor**: PowerKVM attracts MSP using OpenStack
  (Ubuntu in > 50% OpenStack deployment; Ubuntu > 50% Amazon EC2 Instances)
- **OpenPower:** open innovation
- **Bi-Endian**: choice granted by Linux distribution

# Power virtualization options

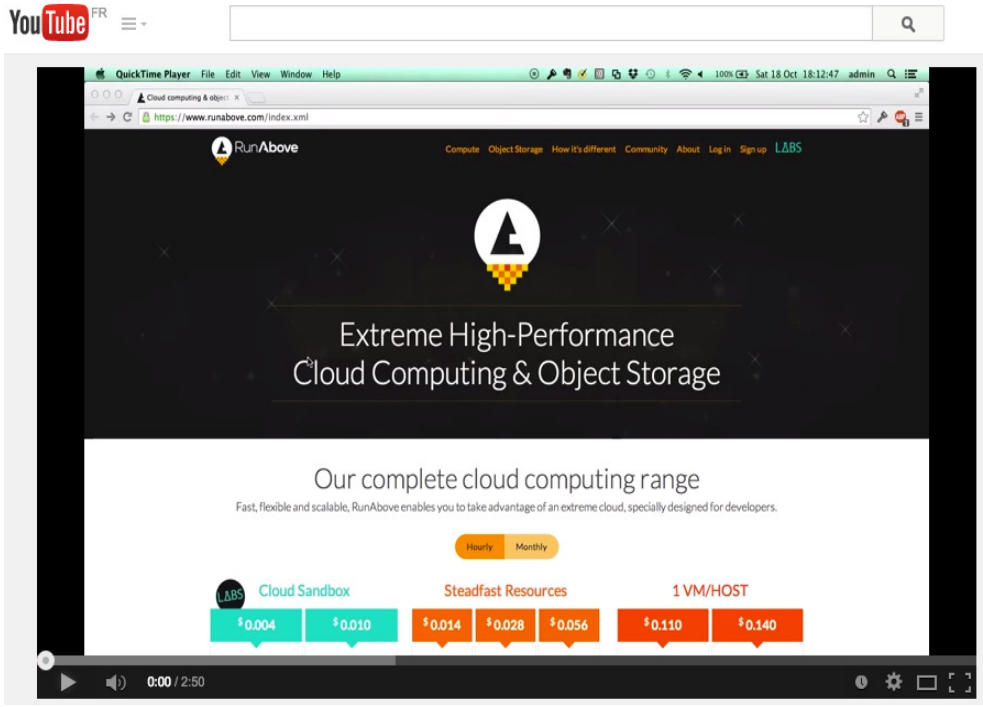
*Reflect the customer segmentation: enterprise vs born on web/cloud*

**PowerVM**: Provides virtualization of Processors, Memory, Storage, & Networking for AIX, IBM i, and Linux environments on Power Systems.

| | PowerVM | PowerKVM |
|---|---|---|
| GA Availability | **2004** | **Q2 2014** |
| Supported Hardware | **All P6, P7, P7+, P8 Systems** | **PowerLinux P8 Systems S812L & S822L** |
| Supported OS | **AIX, IBM i & Linux** | **Linux** |
| Workload Mobility | **Supports AIX, IBM i & Linux** | **Linux** |
| Basic Virtualization Management | **IVM / HMC / FSM** | **Virtman/libvirt** |
| Advanced Virtualization Management | **PowerVC/VMControl** | **PowerVC, Vanilla OpenStack** |
| Admin Type | **Power Centric** | **Linux/x86 Centric** |
| Established Security Track Record on Power | **Yes** | **No** |
| Open Source Hypervisor | **No** | **Yes** |

**PowerKVM**: Open Source option for virtualization on Power Systems for Linux workloads.

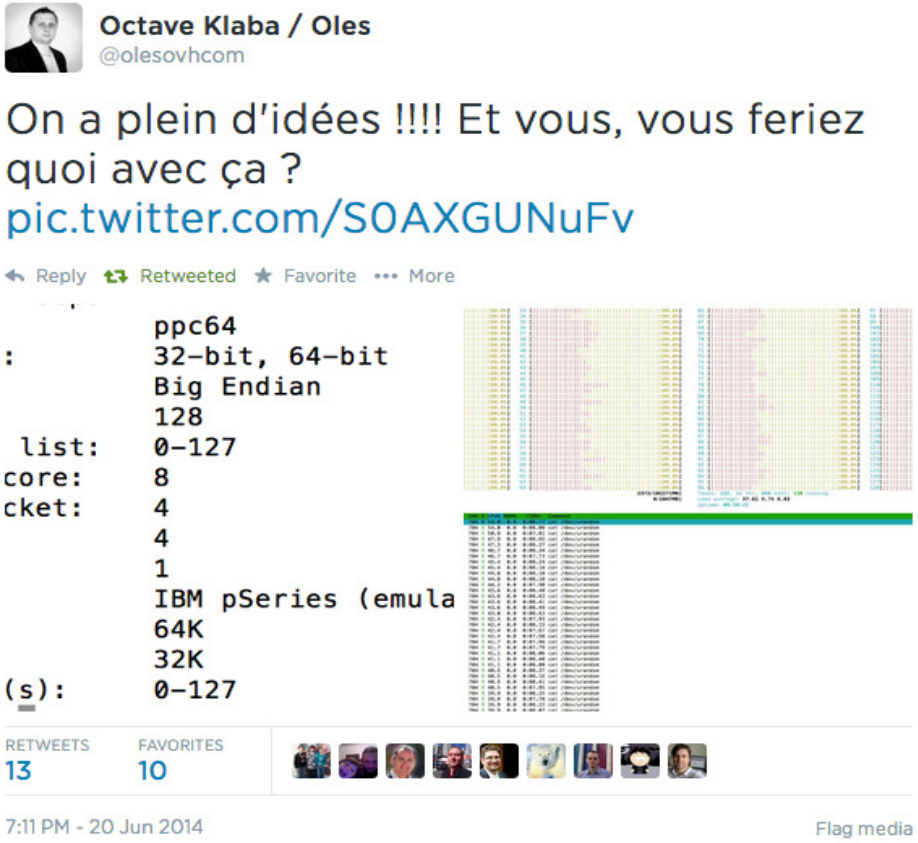
# Runabove.com

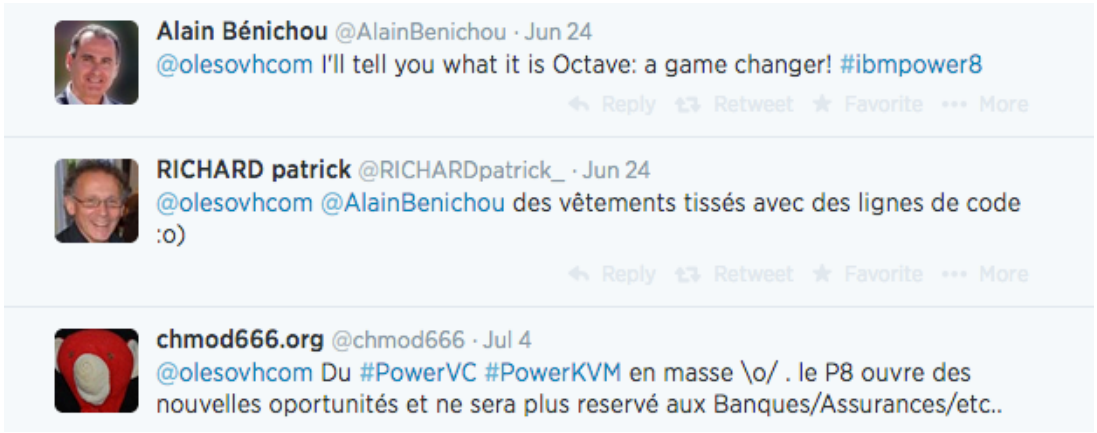



*https://www.youtube.com/watch?v=kE4cjPW2q14*

# Runabove.com – history of OVH and Power8

# OVH and Power8



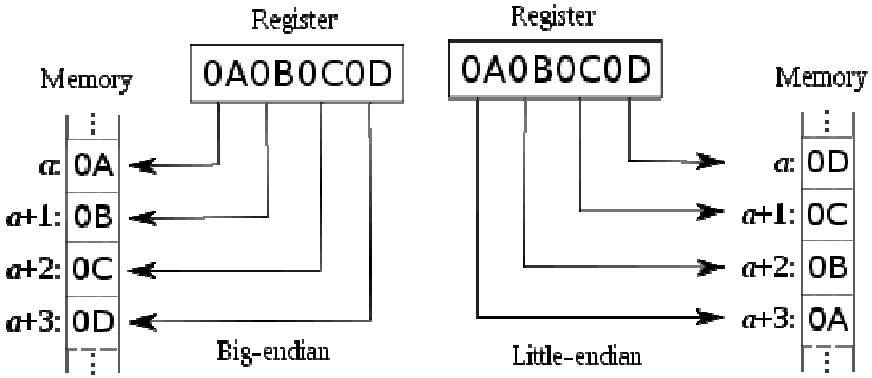
# IBM SWG licensing (Power vs Intel)

| Processor Technologies | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Processor Brand | | | Processor Type | | | | | | | | | |
| Processor Vendor | Processor Name | Server model numbers | Maximum number of sockets per server | Cores per socket (1) | (2) | (4) | (6) | (8) | (12) | IFL Engine | Proc. Model Number | PVUs per Core |
| IBM | POWER Systems cores running Linux OS | 7R1, 7R2, 7R4 POWER IFL, p24L, S812L, S822L<br>Any POWER System core running Linux | All | | | ■ | ■ | ■ | ■ | | All | 70 |
| | POWER8 | S814, S822, S824 | 2 | | | | ■ | ■ | ■ | | 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 |
| | | 520, JS12, JS22, JS23, JS43 | All | | ■ | | | | | | All | 80 |

| Processor Vendor | Processor Name | Maximum number of sockets per server | Cores per socket (1) | (2) | (4) | (6) | (8) | (10) | (12) | (16) | Proc. Model Number[1] | PVUs per Core |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Intel® | Xeon® (Sandy Bridge EP) | 4 | | | ■ | ■ | ■ | | | | E5-4600 to 4699 | 100 |
| | | 2 | | ■ | ■ | ■ | ■ | | | | E5-1400 to 1499<br>E5-1600 to 1699<br>E5-1600V2 to 1699V2<br>E5-2400 to 2499<br>E5-2600 to 2699<br>E5-2600V2 to 2699V2 | 70 |
| | Xeon® (Nehalem EX)[2] | > 4 | | ■ | ■ | ■ | ■ | ■ | | | 6500 to 6599<br>7500 to 7599<br>E3-1200 to 1299<br>E3-1200 V2 to 1299V2<br>E7-2800 to 2899<br>E7-4800 to 4899<br>E7-8800 to 8899 | 120 |
| | | 4 | | ■ | ■ | ■ | ■ | ■ | | | | 100 |
| | | 2 | | ■ | ■ | ■ | ■ | | | | | 70 |
| | Xeon® (Nehalem EP) | 2 | | ■ | ■ | ■ | | | | | 3400 to 3699<br>5500 to 5699 | 70 |
| | Xeon® (pre-Nehalem) | All | | ■ | ■ | ■ | | | | | 3000 to 3399<br>5000 to 5499<br>7000 to 7499 | 50 |
| AMD | Opteron | All | | ■ | ■ | ■ | ■ | | ■ | ■ | All | 50 |
| Any | Any single-core | All | ■ | | | | | | | | All | 100 |

# What is endianness

*In computing, memory commonly stores binary data by organizing it into 8-bit units called bytes. When reading or writing a data word consisting of multiple such units, the order of the bytes stored in memory determines the interpretation of the data word.*



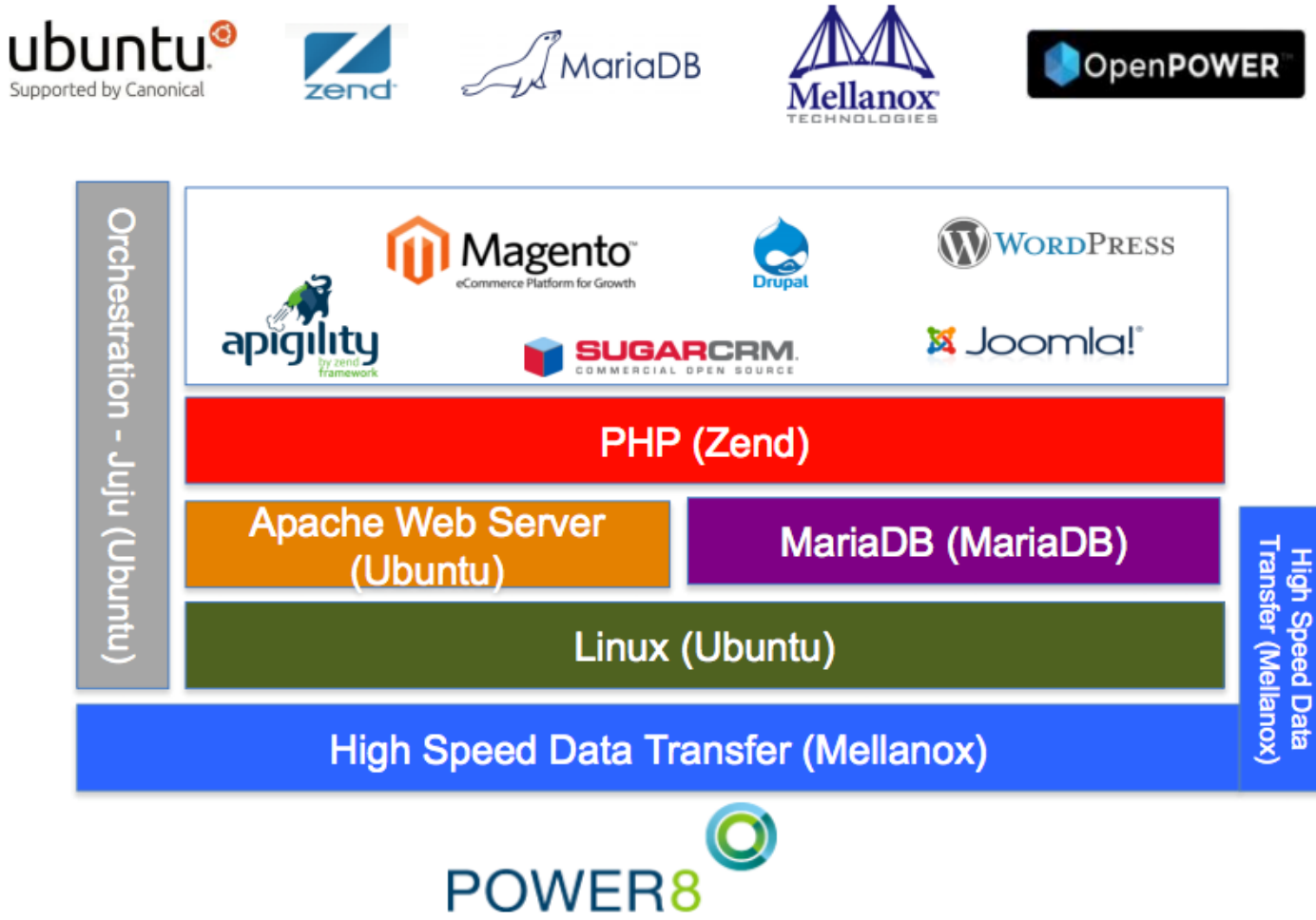

*Each memory storage location is associated with an index, called its address, which indicates its position. Bytes of a single data word are generally stored in consecutive memory addresses.*

***Big-endian** systems store the most significant byte of a word in the smallest address and the least significant byte is stored in the largest address.*

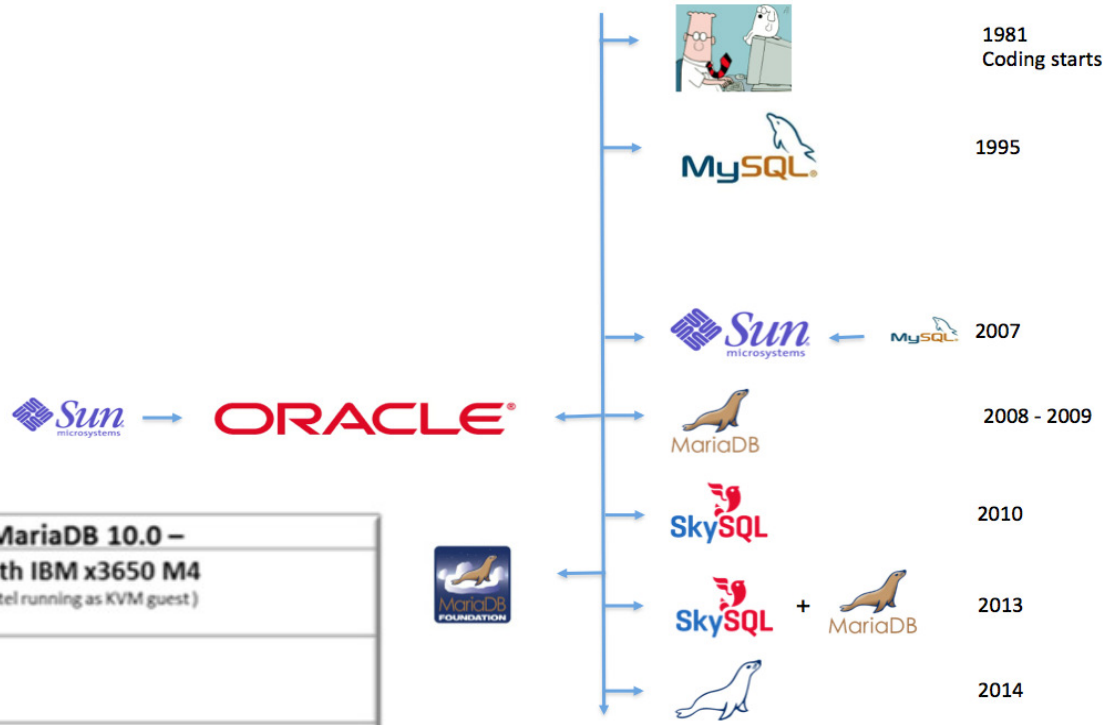
***Little-endian** systems, in contrast, store the least significant byte in the smallest address.*

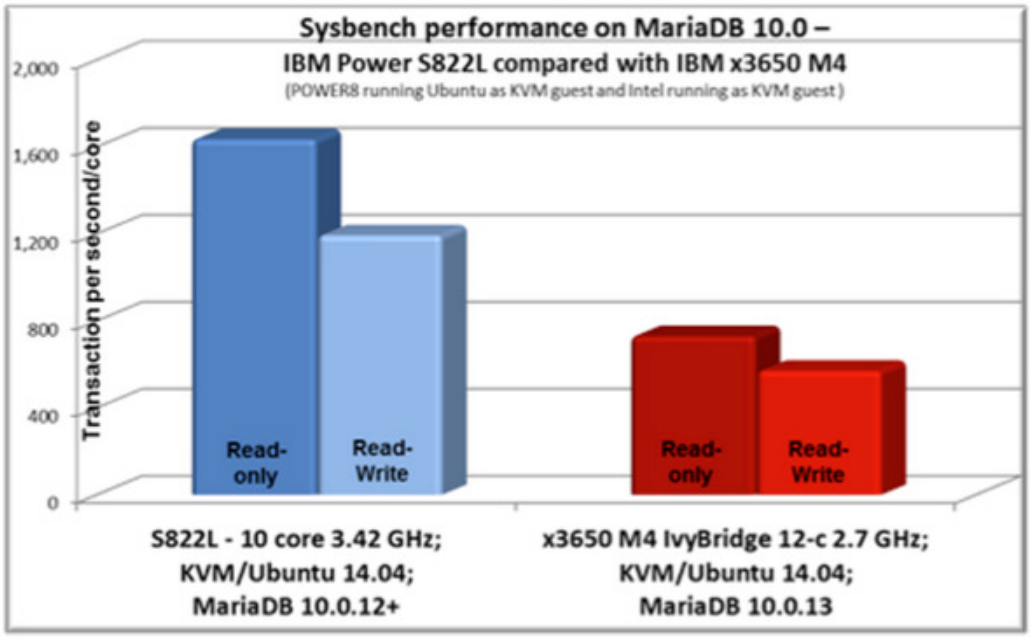
*Source: http://en.wikipedia.org/wiki/Endianness*

# Open Power for Open Source partners



MariaDB




Preliminary benchmarking results shows that MariaDB performance on POWER8 per core is 2.2x more than on x86. Source: MariaDB


## Benefits

- Oracle Compatibility
- Enhanced Performance Features
- Advanced Security
- Postgres Enterprise Manager
- Virtual Private Database
- Index Advisor
- MS SQL Server Replication
- All Benefits from EDB Tools

| License Fee Per Core | Oracle Enterprise Edition *(4 sockets/32 cores) (Power processor)* | EDB Postgres Plus Enterprise Edisiton *(4 sockets/32 cores) (Power processor)* |
|---|---|---|
| Database | $47,500 | included in subscription |
| Partitioning | $11,500 | Included |
| Data Guard | $11,500 | Included |
| Diagnostics | $5,000 | Included |
| Total License Fee per Core | $75,500 | included in subscription |
| Total License Fee per Server (CapEx) | **$2,416,000** | **$0** |
| Annual support/subscription cost per core | 22% of License Fee | $6,900 per socket |
| Annual Support/Maintenance per Server (OpEx) | **$531,520** | **$27,600** |
| **Total 3 Year License and Support Cost** | **$4,010,560** | **$82,800** |

No CAPEX • Annual OPEX reduction 95% • 3 YR TCO cost savings 98%

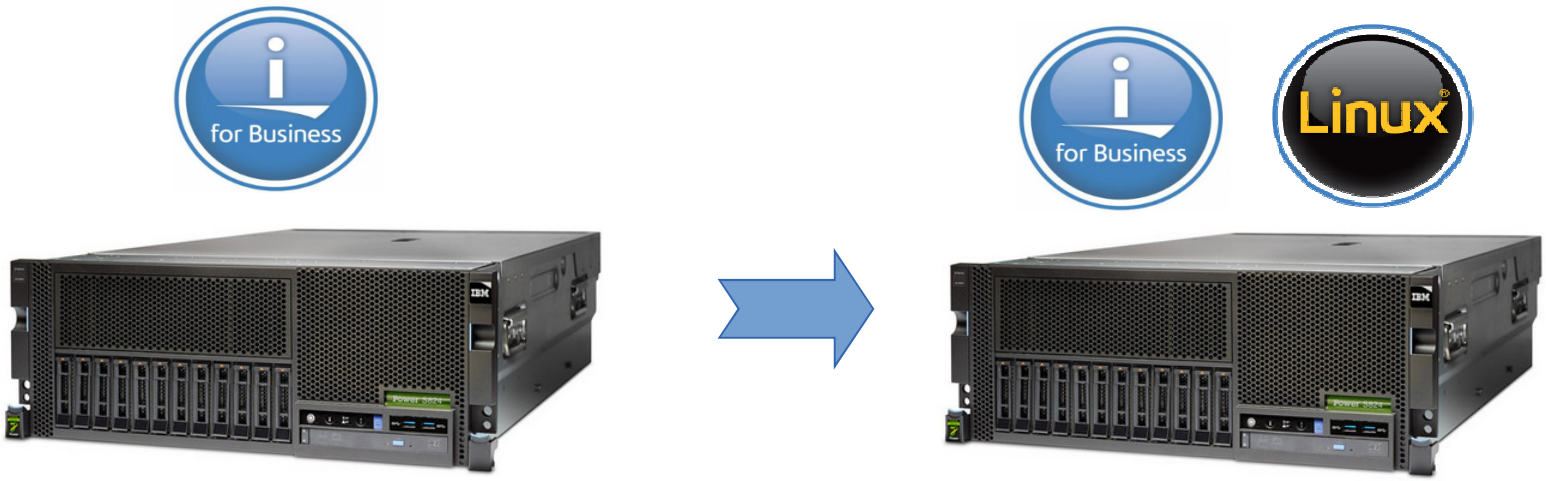




# What's the easiest way to try it, for IBMi clients?

- Primary OS could be IBMi: Linux is a secondary OS
- Enabler: IBMi promo
  - Power standard comes with a Linux license almost for free

- Value proposition:
  - Business Intelligence (existing Power workload being the Data Source)
  - Infrastructural add-on
  - Front-ends (web-applications)



# Previous architecture

Jgalilleo Client

## Clients

*Rich Client is resource consuming (CPU, Memory, Network). Presentation speed depends on Client.*

## ISV view

*Rich Client has to be maintained and supported on too many OS.*

## Other environments

*Each component required a server or virtual server, leading to additional costs for hardware or virtualization licenses. Complexity increasing exponentially*

| | |
|---|---|
| CRM | Workflow |
| Qualità | B2B / B2C |
| App | Configuratore |
| ... | Project Management |

?

## Power

*Only a part of the solution on the enterprise level platform. Lot of unused resources*

- Jgalileo Business Logic (RPG-CLLE-Java)
- Jgalileo DB (db2 for i)
- Archiviazione Documentale (GD)
- Spedizione Documenti (SD)

# New architecture

## Client

*Thanks to x2go the front-end is executed on the server, with the logic of terminal application (citrix-like).*
*From a maintainability perspective everything is on the server.*

Jgalilleo Client

CRM
Qualità
App
Workflow
B2B / B2C
Configuratore
Project Management

## Power

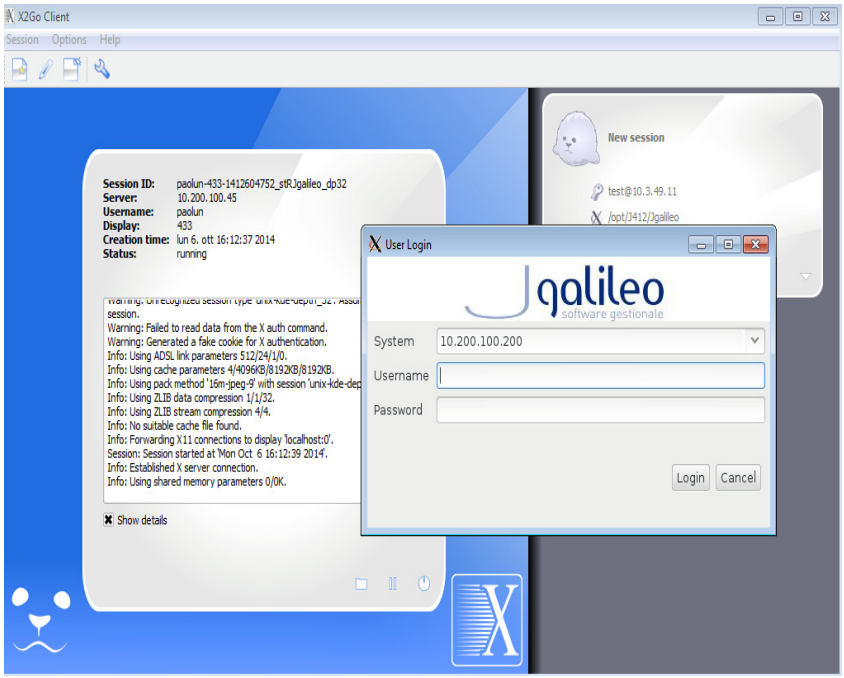
*Consolidated design eases High Availability and Disaster Recovery*

Jgalileo Business Logic (RPG-CLLE-Java)
Jgalileo DB (db2 for i)
rchiviazione Documentale (GD
Spedizione Documenti (SD)

## Linux

*Re-use the free resources for running "other" workloads, instead of allocating additional servers for this.*

# X2go for presentation logic




## *X2GO Client*

*Installed on the user's laptop/desktop.*
*Allows the presentation logic within a terminal application window.*
*Supports Windows, MacOS, Linux.*

## *X2GO Server*

*Installed on a Linux on Power partition.*
*Sends the visual session executed on Linux on Power to the x2go client.*

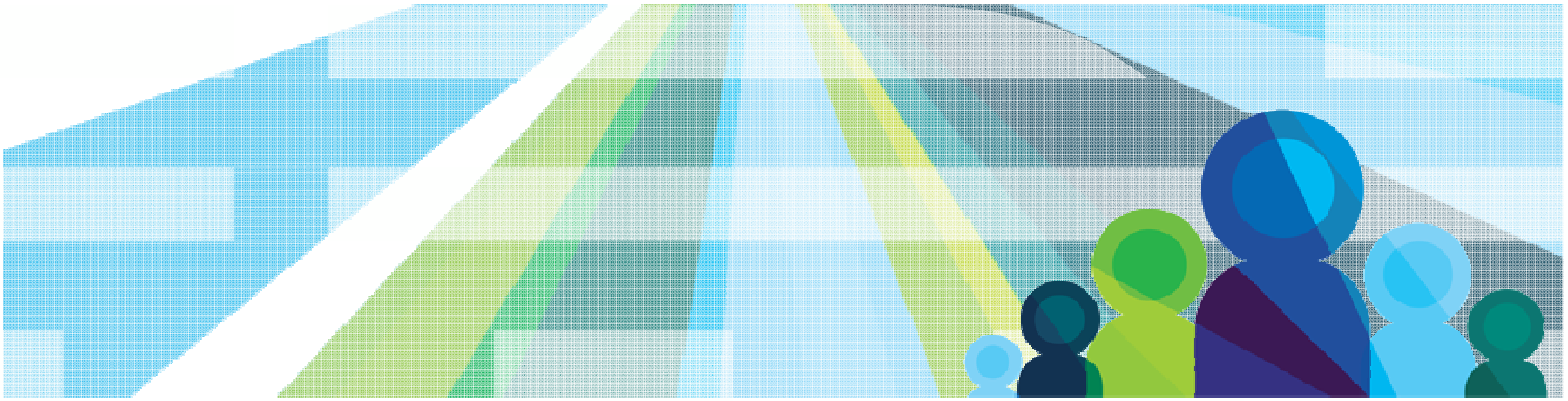
*Communication is optimized (compressed) and is designed to work also when network bandwidth is poor.*
*Supports encryption.*

# Few notes from the benchmark


***~2x improvement from Power7+ to Power8***
***…*** *Marketing does not lie!*

***~3x improvement with JVM, OS and Hypervisor tuning***


# Resources for ISVs



# Tools and Assets

- Porting/Migration tools:
  - Chiphopper (migration/validation support with marketing benefits)
  - Rapid Port Program (migration support)
  - Migration Factory
- Testing/Performance optimization:
  - **21** Innovation Centers in Europe, in **19** countries (3 x **ESP POWER8**)
    - All IIC capable of performing POWER engagements
  - Montpellier PowerLinux Center
  - STG Lab Services
- Hardware Platform (other then IIC and MOP):
  - Power Development Platform
  - Hardware Malls (hardware on premises via leases and discounts)
- Co-marketing:
  - IBM/ISV joint marketing events
- Sales Kit
  - Eagle TCO

# IBM Systems Application Advantage (Chiphopper)

If you answer yes to these questions, you are an excellent candidate for IBM Chiphopper

- Is your application written in C/C++, Java, COBOL, SmallTalk, PHP, Perl, Python or Shell?
- Does your application run on a Linux 32-bit or 64-bit x86 processor and does not have hardware-specific code or use your own kernel modules?
- Are you interested in porting your application to Red Hat Enterprise Linux or SUSE Linux Enterprise Server on IBM Power Systems, System z or both?

If you are still not sure if Chiphopper is right for you, send an email to chiphop@us.ibm.com. We can help you find the right IBM resources to assist you with your multi-platform Linux work.

## Entry criteria

- IBM Partnerworld Member
- ISV nominate application for Chiphopper online: ibm.com/isv/go/chiphopper
- Leverage PDP or IIC for testing

# IBM Power Development Cloud

The Power Cloud that enables developers offers no-charge remote access to IBM hardware, including IBM POWER8™, IBM POWER7+™ and IBM POWER7™ processor-based servers on the IBM AIX®, IBM i and Linux® operating systems.

PDP is intended for Development, Porting and Functional testing only.

User systems include virtualized CPU, disk and networking. Users have root access to their systems but cannot access the Hardware Management Console (HMC) or the Virtual I/O Server (VIOS). Reservation duration is up to 14 days with configuration defaults. System defaults can be adjusted using a promotional code.

Access to IBM Middleware.

https://www.ibm.com/partnerworld/pdp

| | POWER8™ with AIX & Linux | PowerKVM™ with Linux | POWER8™ with IBM i |
|---|---|---|---|
| **Server specification** | IBM Power System S824 | IBM Power System S822L | IBM Power System S824 |
| **OS supported** | AIX 7.1, AIX 6.1, RHEL 6.5, SLES 11.3 | RHEL7, Ubuntu 14.04 | IBM i7.2 |
| **vCPU/Memory** | 1 to 2 vCPU / 2 to 4 GB RAM | 1 to 2 vCPU / 2 to 4 GB RAM | 1 to 2 vCPU / 2 to 4 GB |
| **Storage space** | 35 GB storage on IBM System Storage® DS8000® | 15 to 35 GB storage on IBM Storwize V7000® | 100 GB storage on IBM System Storage® DS8000® |
| **% of real CPU per Virtual CPU** | 20% | 20% | 20% |

# IBM Hardware Mall

As a member of PartnerWorld, you may qualify to lease or purchase equipment at very attractive terms to support your development needs.

Qualified members can get cost effective and timely access to the latest technologies for any of the following business-related purposes:

- Product development, testing, support or maintenance
- Product demonstrations
- Use in customer demonstration or briefing centers

ibm.com/partnerworld/page/isv/hardwaremall

## Entry criteria

- Countries: Europe, North America
- Membership: All levels
- Cost: Variable

# IBM Migration Support

The IBM Migration Factory makes is easier than you think with resources to help you move your solutions to the IBM POWER platform and help you differentiate your solutions from others. Read about some of the thousands of customers that have gained the POWER platform advantages and how they were able to provide exceptional client experience.

- Port to multiplatform Linux
- Solaris to Linux
- HP-UX to Linux
- Windows to Linux
- Skills migrations and re-platforming

http://www-03.ibm.com/systems/power/migratetoibm/path.html

# IBM Innovation Centers (IIC)

The IBM Innovation Centers offer the resources you need including technical expertise, marketing and sales facilities and skills, and IBM hardware and software. Collaborate with IBM and other Business Partners through our worldwide centers.

**21** Innovation Centers in Europe, in **19** countries (3 x **ESP POWER8**)

- All IIC capable of performing POWER engagements
- 3 x Power8 Early Shipment Programs

https://www.ibm.com/partnerworld/iic

## Value

- Offer technical skills and hardware access across multiple platforms, local or remote, at little or no cost.
- Shorten development and testing time with hands-on assistance and leading edge technology.
- Qualify for "Ready for" marks to broaden your marketing reach
- Attend workshops, seminars, and more conducted by trained subject matter experts.
- Support Proof-of-concept or prototyping, Validation testing, Porting and enabling on IBM platforms, IBM operating systems (also specific Linux distributions) and IBM middleware, migration, integration testing, and performance and scalability testing

# IBM Power Systems Linux Centers

The Centers leverage existing capabilities within IBM to deliver services to Clients, ISVs, SIs and MSPs.

**Center Locations**

Beijing, China; Austin, TX; New York, NY, Montpellier, France; Tokyo, Japan

**Capabilities**

Briefing, demonstrations, porting and migration assistance, optimization and benchmarking, training

www.ibm.com/systems/power/software/linux/centers/

## Value

- Tailored customer briefings to learn about using Power Systems to deploy world-class Linux applications. Linux training workshops for how to program, port and optimize applications using Red Hat and SUSE environments on Power Systems.
- Hands-on assistance to show developers how to take advantage of Power parallel processing and advanced virtualization capabilities.

# IBM Systems Lab Services

IBM Systems Lab Services helps BPs and ISVs to drive faster adoption of new technologies. IBM Lab Services has over 50 experienced Linux on Power technical consultants (world-wide).

IBM Lab Services teams with IBM Business Partners and ISVs to assist with technically complex on client site situations including proof of concepts, application performance optimization, virtualization optimization and management.

The majority of pre-sales Lab Services engagements with BPs and ISVs are funded through IBM Power Systems brand programs that are based on Power Systems HW/SW deal opportunity. Post sales services, include Power Linux within PowerCare.

# Support for ISVs

- Power Systems Linux web site
  www.ibm.com/systems/power/software/linux/
- Power Linux Community
  www.ibm.com/developerworks/group/tpl/
- Power Linux on Google+
  http://plus.google.com/communities/100156952249293416679
- Power Linux on Twitter
  @ibmpowerlinux
- Applications from IBM using IBM Software Products Reports
  http://www-969.ibm.com/software/reports/compatibility/clarity/index.jsp
- Applications from ISVs using the ParterWorld Global Solutions Directory
  www.ibm.com/partnerworld/gsd