Por que la infraestructura importa?

Manejando la explosión de datos

Pablo Pereira
Ineramericana de Cómputos

- 1997 - 2014
  - Administrador de sistemas Unix/AIX
  - Especialista en sistemas IBM Power
- Rol actual
  - Administrador y soporte de sistemas Unix/ AIX
  - Especialista Técnico en sistemas IBM Power
- Jiap 2012

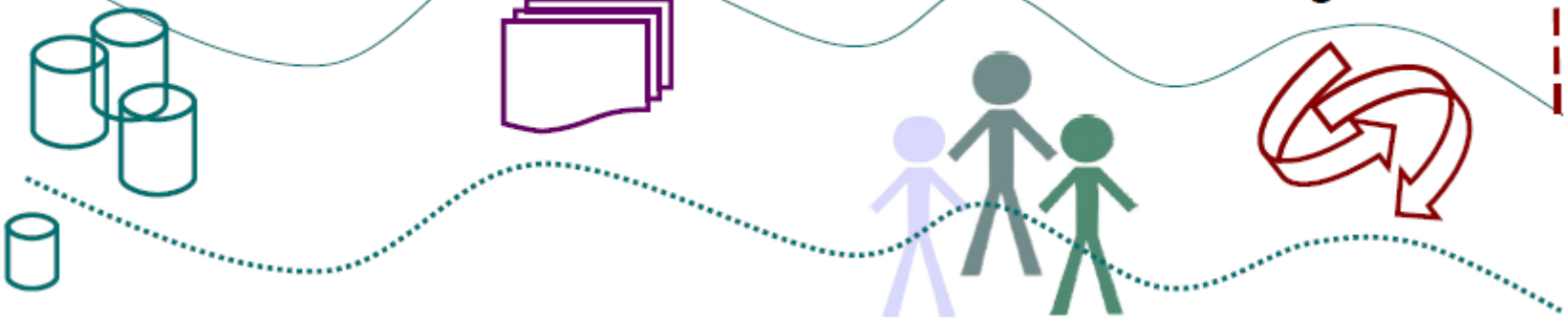

The New York Times Magazine
How your shopping habits reveal even the most personal information. By Charles Duhigg

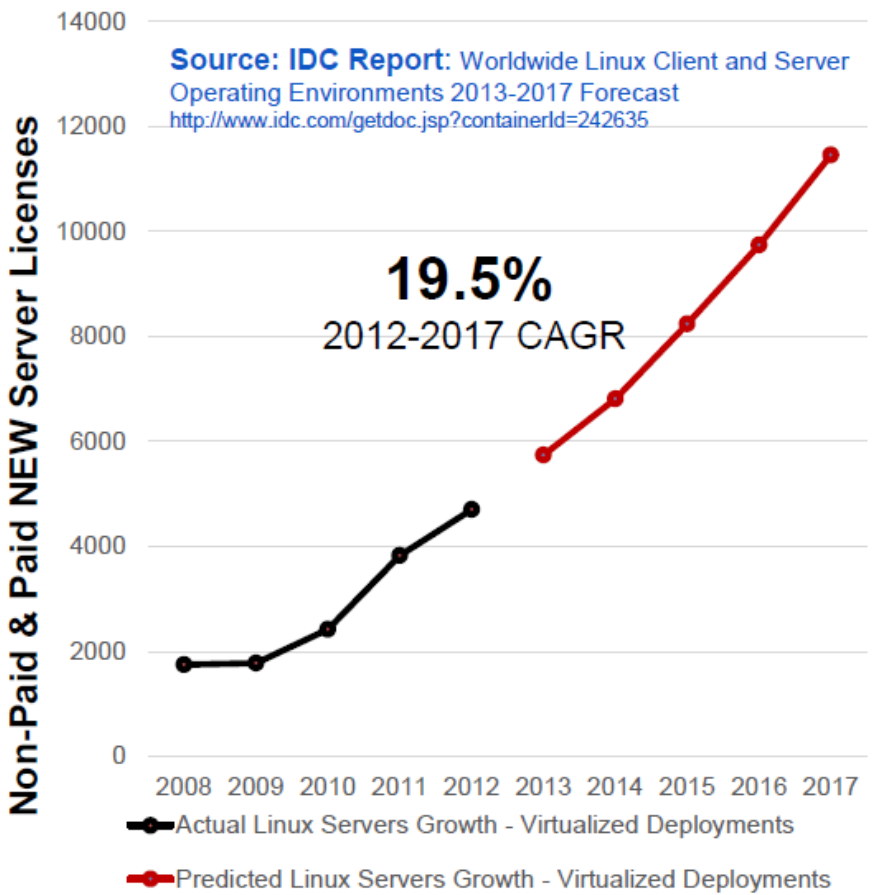
Businesses are “dying of thirst in an ocean of data”

90%
of the world’s data was created in the last two years

80%
of the world’s data today is unstructured

20%
is the amount of available data traditional systems leverages





FIVE YEAR PLANS FOR INCREASED OS INVESTMENTS

Increasing Use of Linux

80%

Increasing Use of Windows

20%

LINUX IS CORE TO THE CLOUD

Maintaining or Increasing Linux to Support Cloud

74%

Decreasing Linux to Support Cloud

1%

Source: **2013 Enterprise End User Report**
http://www.linuxfoundation.org/publications/linux-foundation/linux-adoption-trends-end-user-report-2013


MYTHBUSTERS

1990 POWER1/SP1 | 1993 POWER2/SP2 | 1996 P2SC | 1998 POWER3 | 2001 POWER4 | 2004 POWER5 | 2007 POWER6 | 2010 POWER7 | 2014 POWER8

IBM is Heavily Investing in the Success of Linux and the Ecosystem to Support it

$1 (2001) + $1 (2013) Billion USD

- **100%** of our systems
- **400+** software products
- **500** patents donated
- **600+** developers

IBM Watson	Power Development Cloud	Power Linux Centers Beijing, New York, Tokyo, Austin, Montpellier
OpenStack, KVM	OpenPOWER	PartnerWorld & Developer Ecosystem

IBM Innovation Center Locations

North America
- Austin
- Chicago
- Dallas
- Silicon Valley
- Toronto
- Cambridge

Latin America
- Mexico City
- Sao Paulo

NE Europe
- Amsterdam
- Barcelona
- Copenhagen
- Dublin
- Ehningen
- Hursley
- La Gaude
- Milan
- Paris
- Southbank
- Stockholm
- Tel Aviv
- Zurich

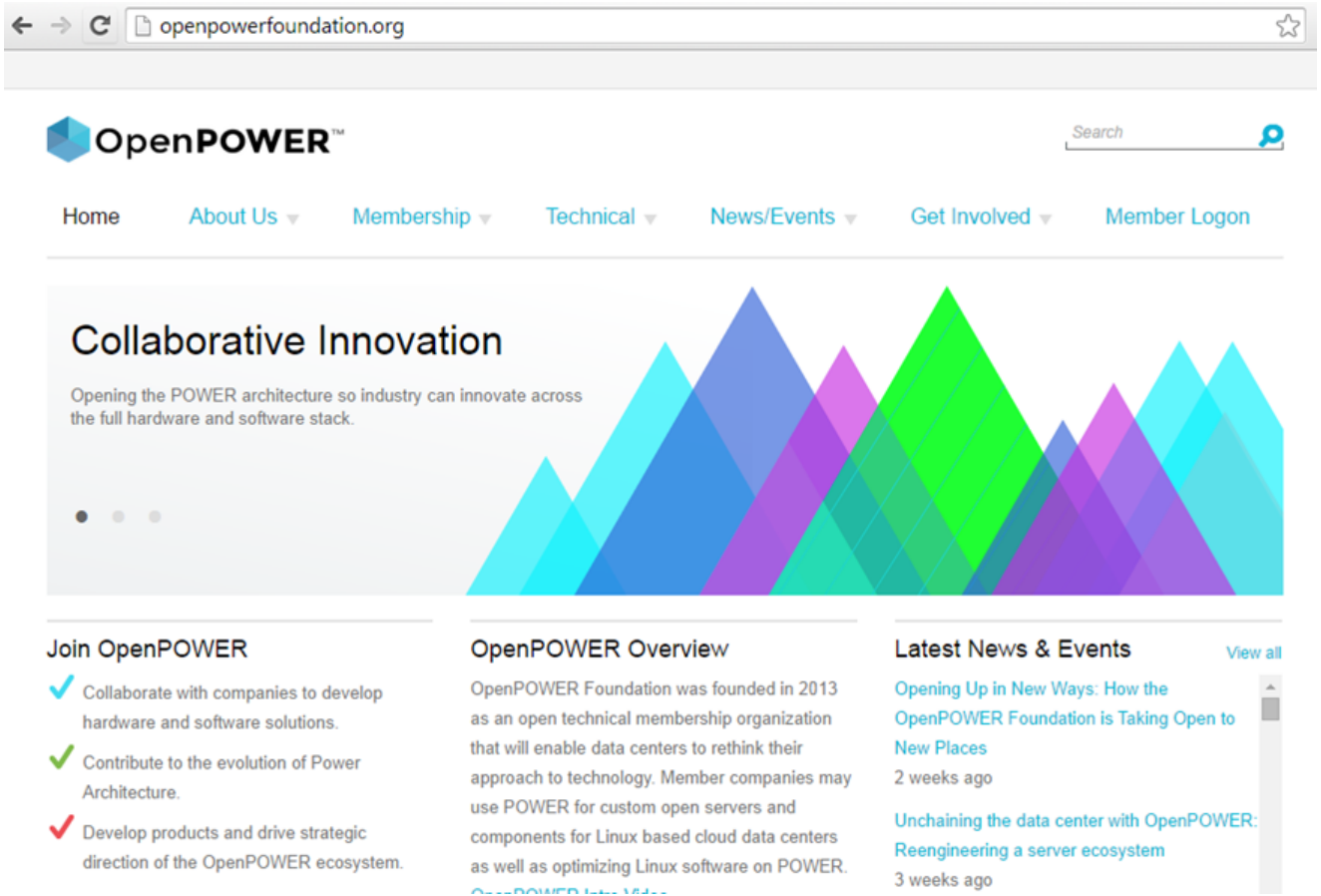
CEE
- Bratislava
- Bucharest
- Budapest
- Istanbul
- Kiev
- Ljubljana
- Moscow
- Prague
- Warsaw

MEA
- Casablanca
- Johannesburg

Asia Pacific
- Bangalore
- Ho Chi Minh City
- Kuala Lumpur
- Manila
- Seoul
- Shanghai
- Sydney

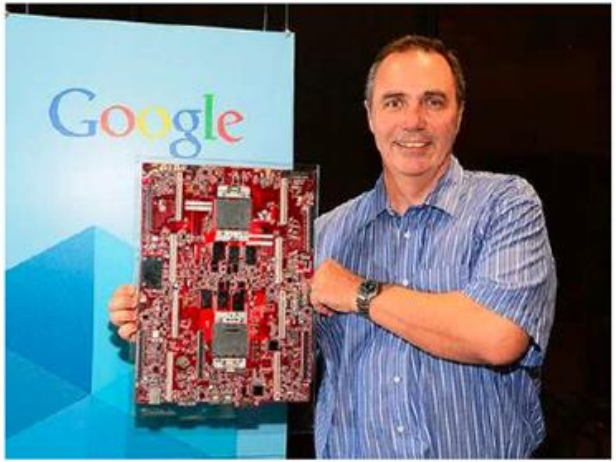
Japan
- Tokyo

Blue = site with PureFlex Systems


openpowerfoundation.org
OpenPOWER™
Search
Home
About Us
Membership
Technical
News/Events
Get Involved
Member Logon
Collaborative Innovation
Opening the POWER architecture so industry can innovate across the full hardware and software stack.
Join OpenPOWER
Collaborate with companies to develop hardware and software solutions.
Contribute to the evolution of Power Architecture.
Develop products and drive strategic direction of the OpenPOWER ecosystem.
OpenPOWER Overview
OpenPOWER Foundation was founded in 2013 as an open technical membership organization that will enable data centers to rethink their approach to technology. Member companies may use POWER for custom open servers and components for Linux based cloud data centers as well as optimizing Linux software on POWER.
Latest News & Events
View all
Opening Up in New Ways: How the OpenPOWER Foundation is Taking Open to New Places
2 weeks ago
Unchaining the data center with OpenPOWER: Reengineering a server ecosystem
3 weeks ago

What OpenPOWER Means for our Customers

- OpenPOWER members will bring Power Servers with specialized Power processors, subsystems and components with big performance gains
- Customers will have a choice of vendors to buy Power Servers and/or components from



Google's Gordon MacKean with IBM POWER8 server motherboard

Tyan's motherboard for POWER8

Linux on Power – Runs The Same Commands as Linux on x86

- **Supports Red Hat, SUSE Enterprise, Ubuntu Linux versions consistent with x86_64**
  - POWER support available simultaneously with other platforms
  - List of packages nearly identical (except minor differences like bootloader)
  - Packages at same version/level – including kernel and device drivers

- **Leverage same opens source system tools**
  - SDK: Same Free Eclipse-based development environment
    http://www14.software.ibm.com/webapp/set2/sas/f/lopdiags/sdklop.html
  - Advance Toolchain: Same Open Source tools (GNU), IBM tested and supported on Power




NEW RAS Features Add in POWER8

Increasing POWER RAS for each generation is a mandate

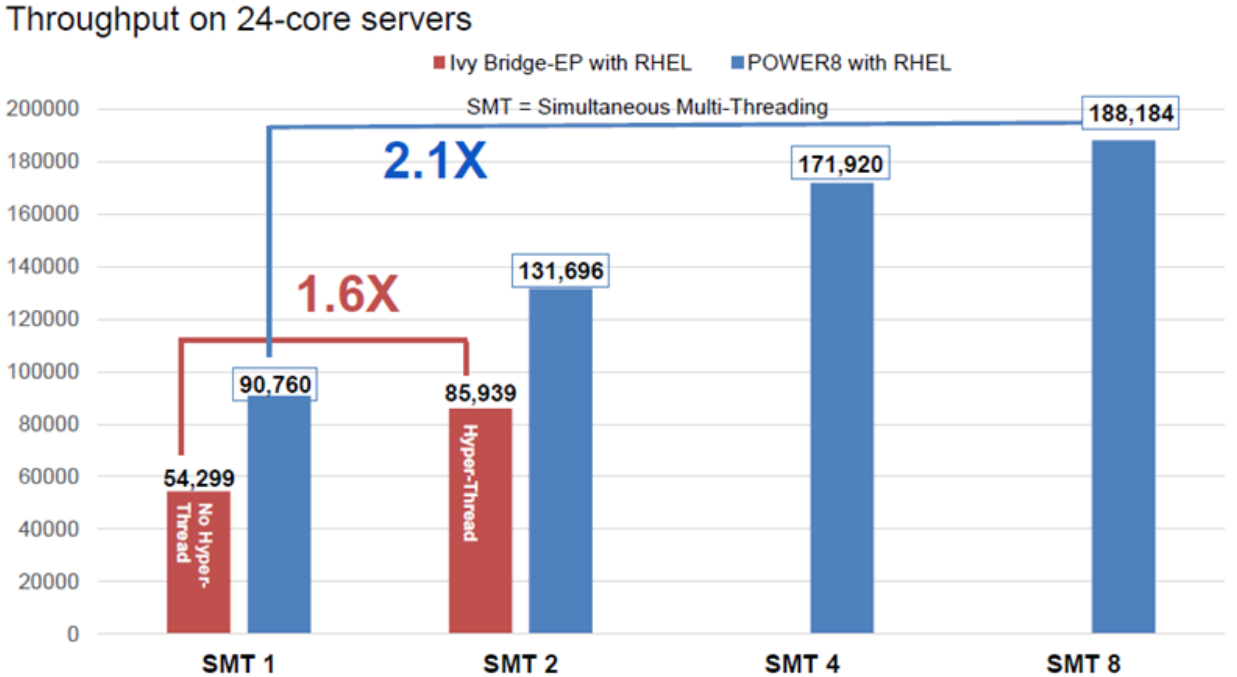
1. "Enterprise Memory" methodology added to POWER8
- Previously only on Enterprise Power Servers

2. Integrated Power cooling and monitor on the POWER8 processor
- Improved predictive failure
- Enhances the ability to "overclock" without sacrificing RAS

3. PCIe controller integrated in processor eliminating external I/O Hub controller (Both performance and RAS enhancement)

4. PCIe HotPlug added for serviceability

Linux on POWER8 Throughput Compared to Linux on Ivy Bridge



This is an IBM internal study designed to replicate a typical IBM customer workload usage in the marketplace. It consists of a POWER8 S824 with 24 cores, 256GB Memory, 3.52 GHz, RHEL 7.0, WAS 8.5.5.2, DB2 9.7, JDK 7.0 FP1 compared to an Ivy Bridge EP 24 cores, 256GB Memory, 2.7 GHz, RHEL 6.5, WAS 8.5.5.1, DB2 9.7, JDK 7.0 FP1. The results were obtained under laboratory conditions, and not in an actual customer environment. IBM's internal workload studies are not benchmark applications, nor are they based on any benchmark standard. As such, customer applications, differences in the stack deployed, and other systems variations or testing conditions may produce different results and may vary based on actual configuration, applications, specific queries and other variables in a production environment. Prices, where applicable, are based on published US list prices for both IBM and competitor, and the Total Cost of Acquisition (TCA) includes the list HW and SW prices and 3 years of service & support which is then divided by the number of transactions to get $ per user interaction per second.

IBM FlashSystem

JIAP 2014

Leandro Naya
Interamericana de Cómputos

- **2011 – 2014**
  - **System X Technical & Sales Specialist**
  - **Storage Technical & Sales Specialist**
- **Mi Rol Actual**
  - **Power Systems Technical & Sales Specialist**
  - **Storage Technical & Sales Specialist**
- **Jiap 2013**

Leandro Naya
Fuera del trabajo…

Why Flash…

In the last 10 years:

CPU Performance **8 - 10x** increase

DRAM Speed **7- 9x**

Network Speed **100x**

Bus Speed **20x**

Disk Speed **1.2**

…and everything waits

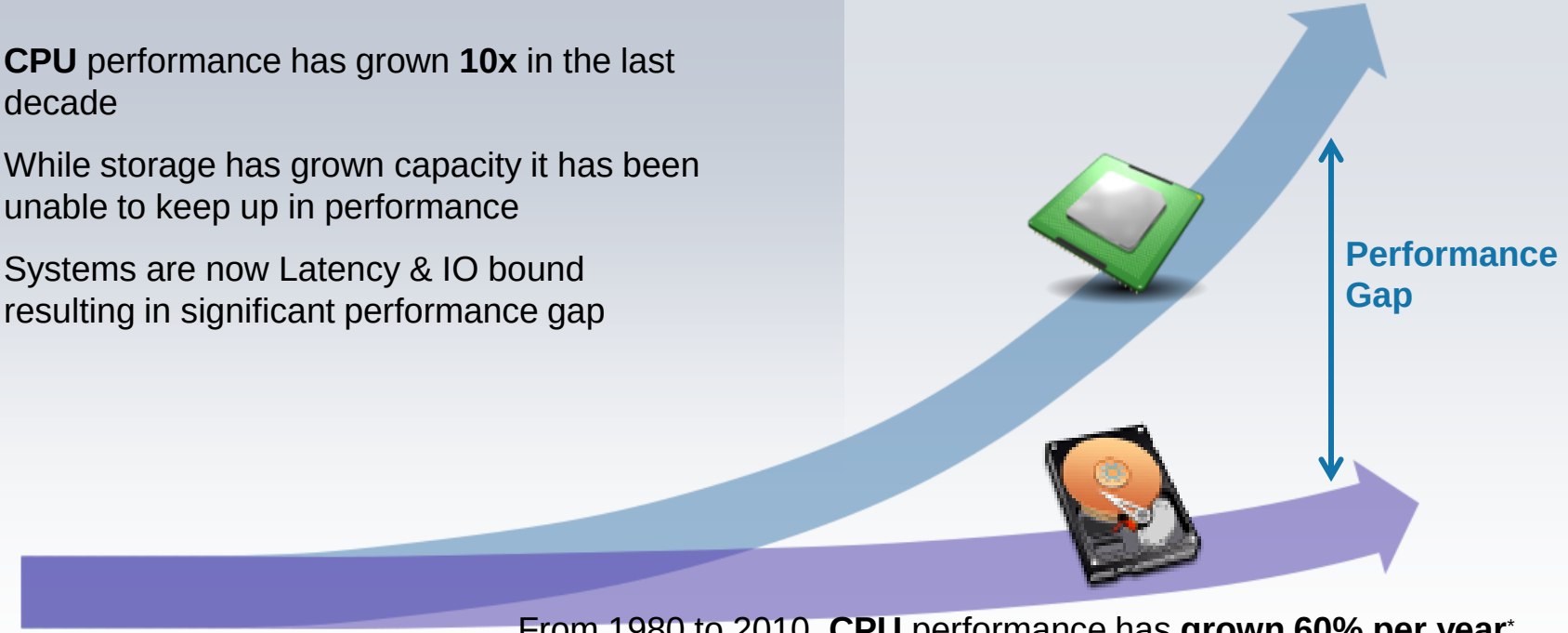
Source: IBM and Industry Estimates

IT infrastructure challenges

CPU performance has grown **10x** in the last decade

While storage has grown capacity it has been unable to keep up in performance

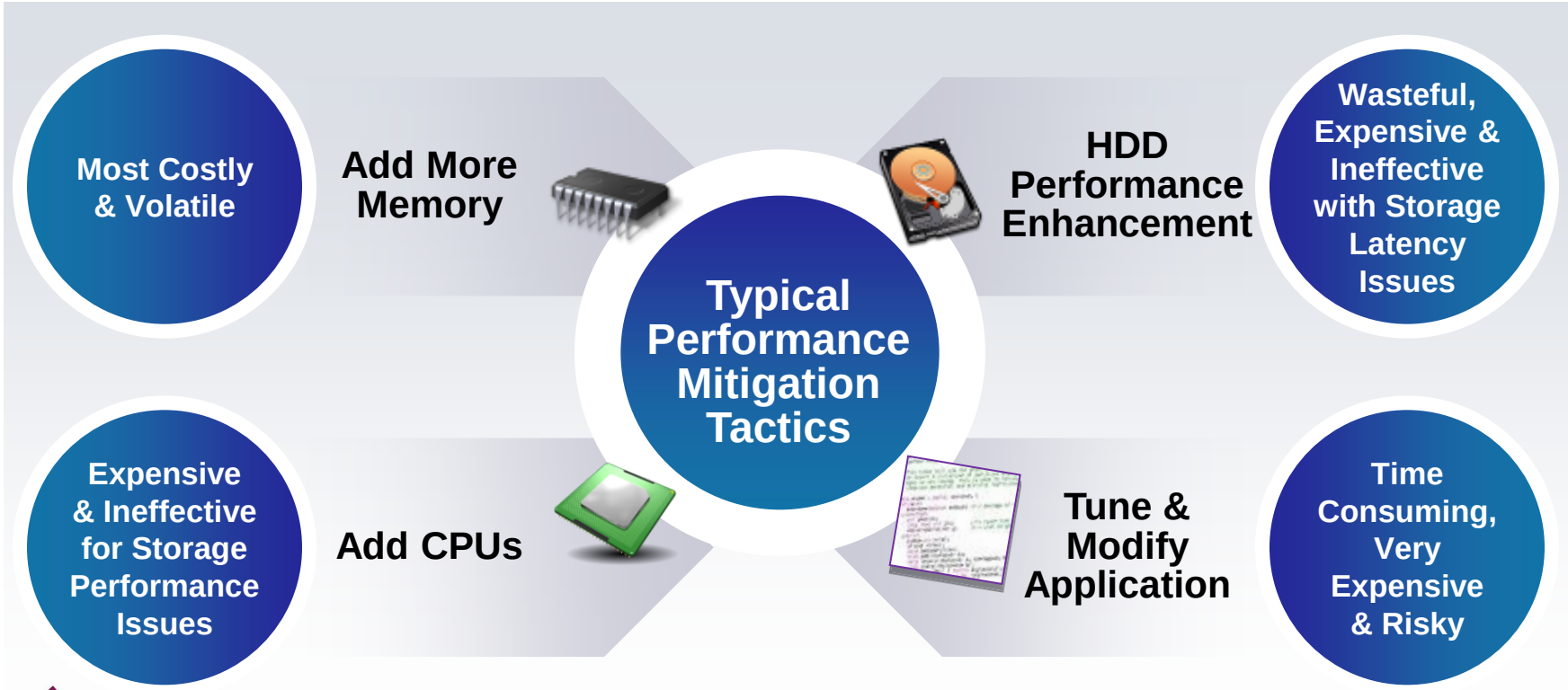
Systems are now Latency & IO bound resulting in significant performance gap



From 1980 to 2010, **CPU** performance has **grown 60% per year***

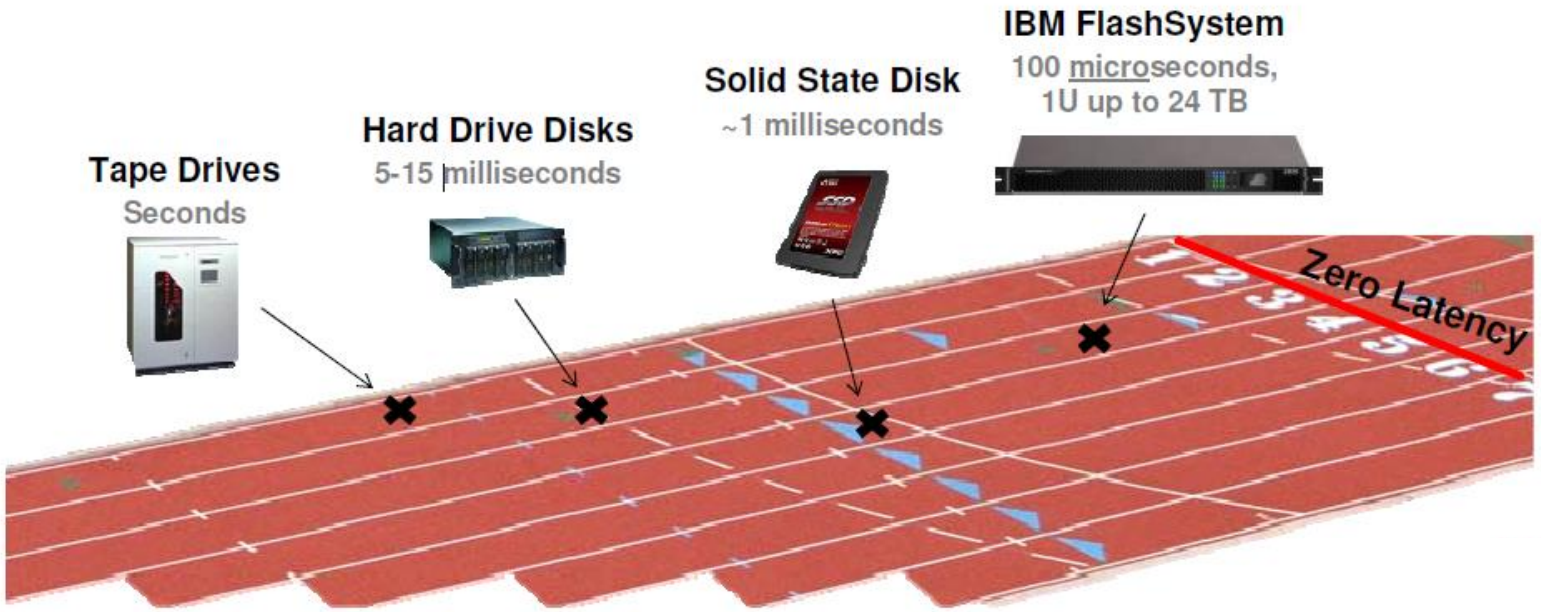
…and yet, **disk** performance has **grown ~5% per year** during that same period**

Client responses to performance gap



Race to Zero Latency

With each new gen. of storage, comes performance gains in order of magnitude



How are SSDs and IBM FlashSystem™ Different?

Think of SSD as a fast car running on a dirt road. The design (or road) is intended for disk, not for Flash.

SSD: Flash inside a disk enclosure **designed for Disk, not Flash**

These factors add latency:

- Array Operation System
- Array Controllers
- Disk Data Protection (RAID)
- SAS controllers and shared data path/bus
- Tiering and Variable Performance

FlashSystem: Purpose-built (**FPGA**), **hardware-only data path**

These factors ELIMINATE latency:

- No software layers, No Operating System
- HW-Optimized and Designed for Flash
- Massive internal parallelism

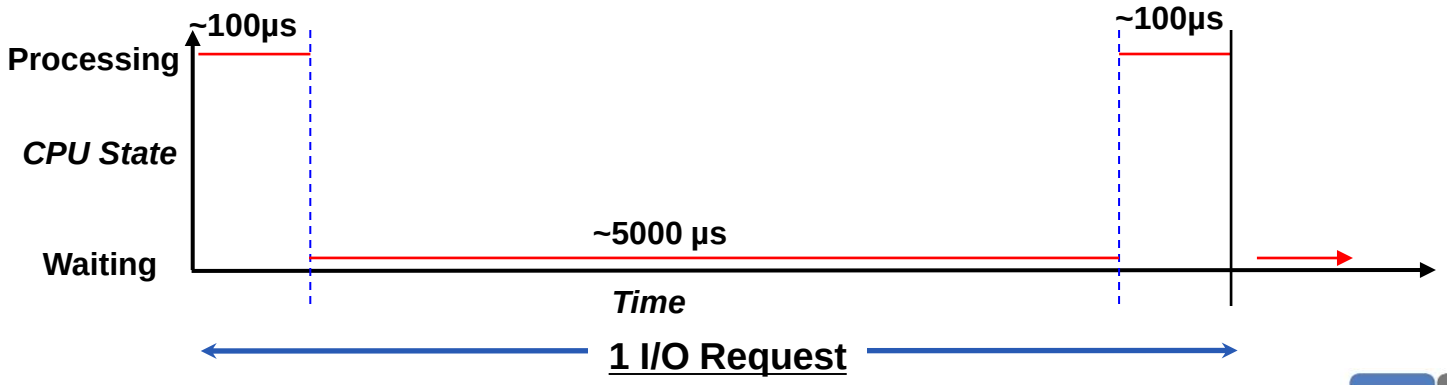
Maximum application efficiency: Application, Server and CPU
Compelling reduction of TCO and shorter ROI

Storage disk latency is driving lower CPU utilization/efficiency

I/O Serviced by Disk

1. Issue I/O request ~ 100 µs
2. Wait for I/O to be serviced ~ 5,000 µs
3. Process I/O ~ 100 µs

- Time to process 1 I/O request = 200 µs + 5,000 µs = 5,200 µs
- CPU Utilization = Wait time / Processing time = 200 / 5,200 = ~4%



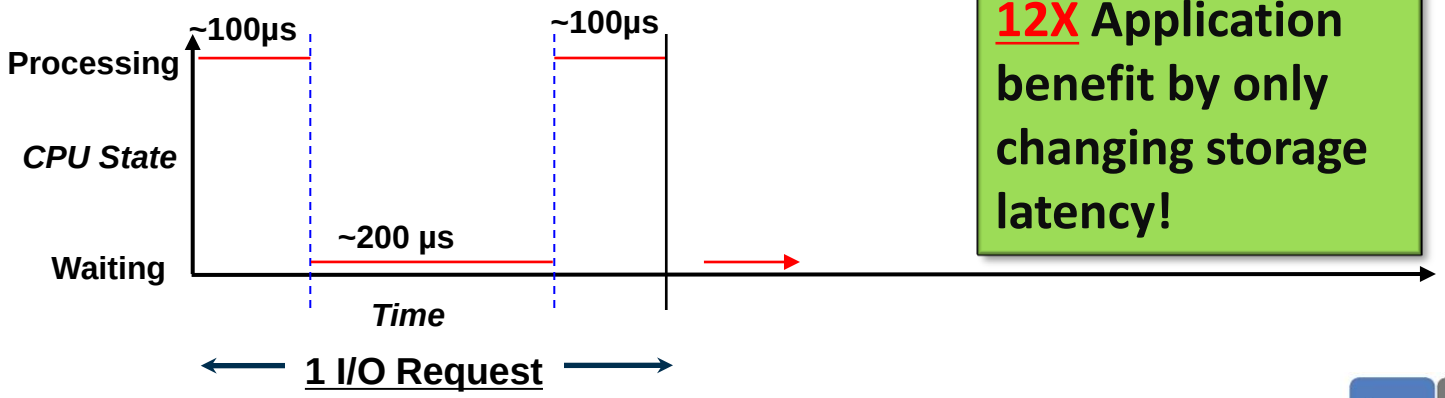
IBM FlashSystem's micro-second latency helps drive CPU utilization higher

I/O Serviced by IBM FlashSystem

1. Issue I/O request ~ 100 µs
2. Wait for I/O to be serviced ~ 200 µs
3. Process I/O ~ 100 µs

UP from ~4%

- Time to process 1 I/O request = 200 µs + 200 µs = 400 µs
- CPU Utilization = Wait time / Processing time = 200 / 400 = 50%



12X Application benefit by only changing storage latency!

Introducing the solution, the FlashSystem 840 replacing performance HDD in your data center

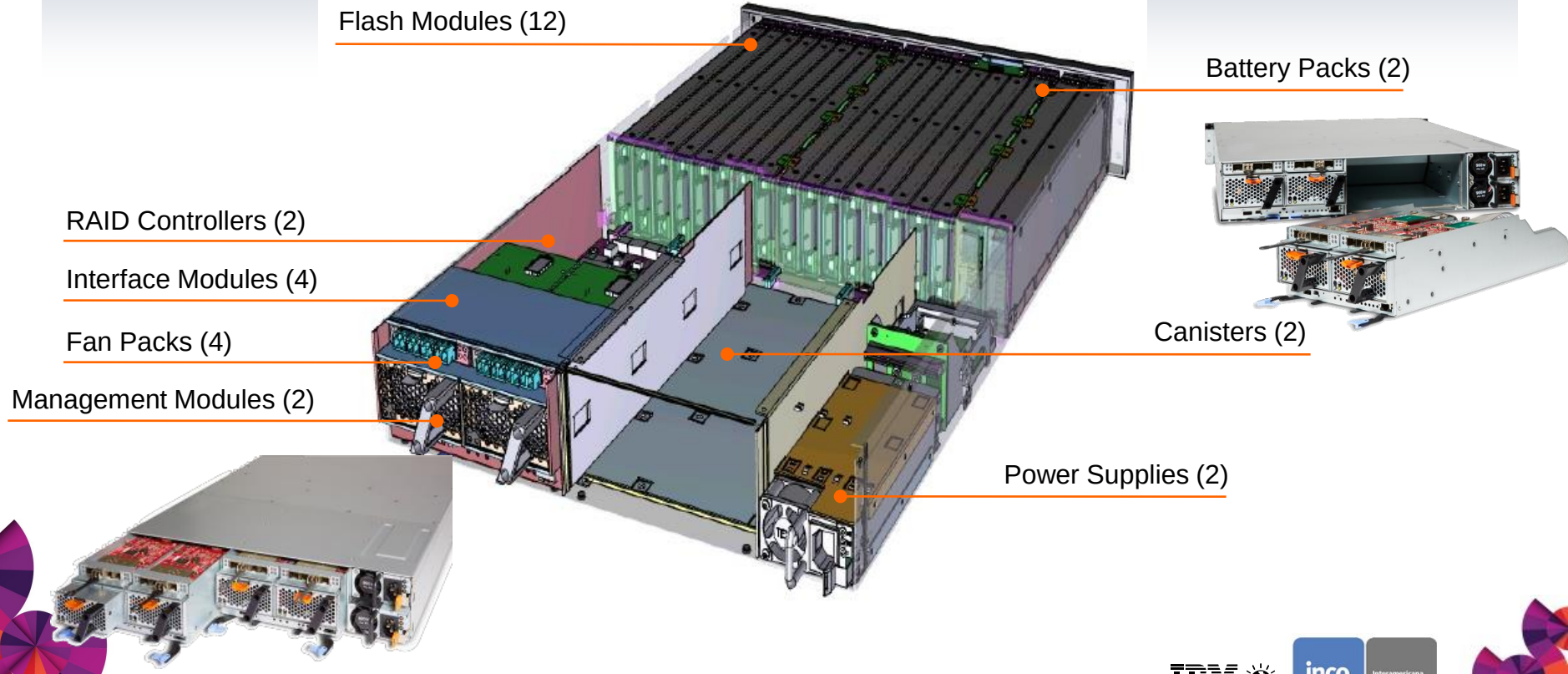
Data center optimized to deliver extreme performance, flexible capacity and total system protection

Performance at-a-glance

Performance criteria[ab]	Maximum capacity (12 flash modules)	Middle capacity (8 flash modules)	Minimum capacity (4 flash modules)
100% Read IOPS	1.1 M	1.1M	1.0 M
100% Write IOPS	600 K	400 K	225 K
70/30 IOPS	750 K	500 K	225 K
100% large block sequential read	8 GB/sec	8 GB/sec	4 GB/sec
100% large block sequential write	4 GB/sec	2.5 GB/sec	1 GB/sec
Read latency	135 µs	135 µs	135 µs
Write latency	90 µs	90 µs	90 µs

a. Data gathered using an Oakgate storage test appliance and an FC protocol analyzer.
b. All measurements are made in a RAID 5 configuration, 4 TB cards, and 90% of usable capacity.

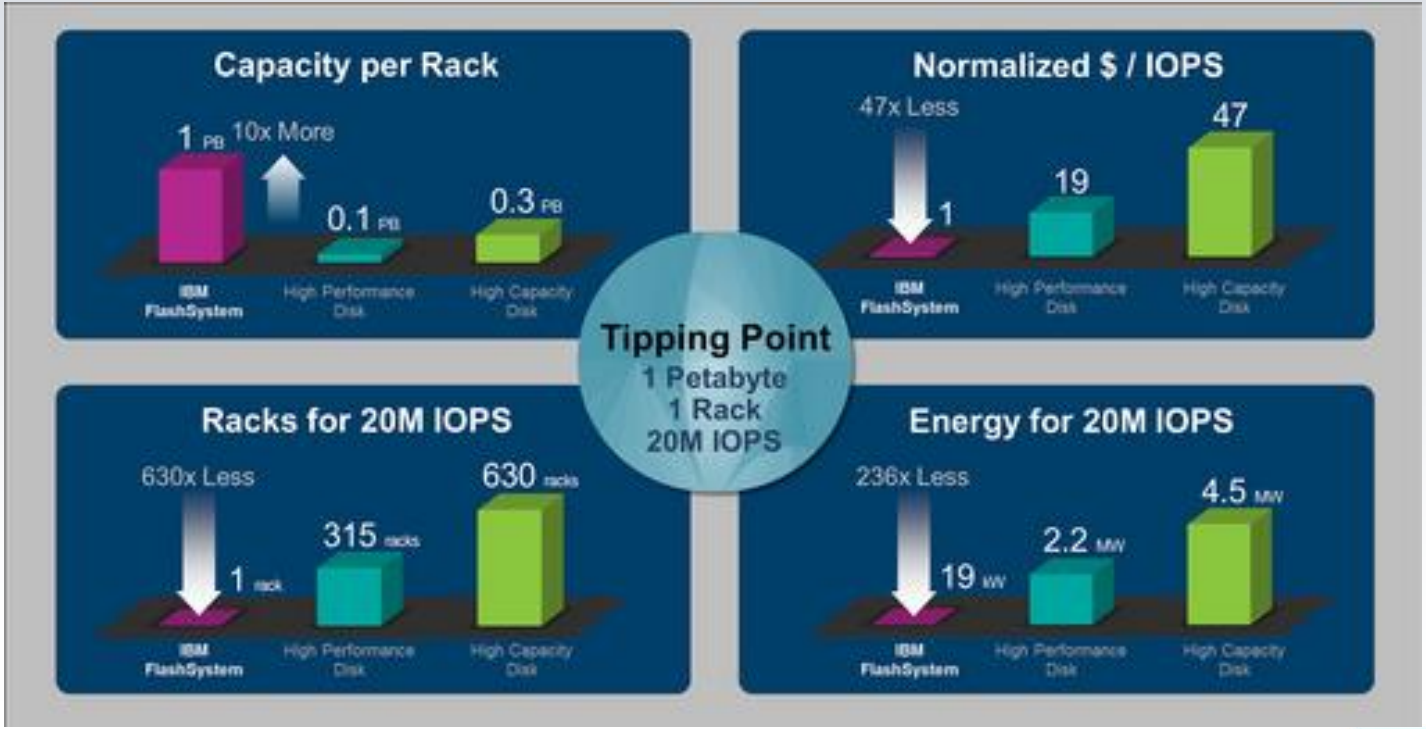
IBM FlashSystem 840 Components



Example from IBM Research: 20M IOPS

http://ibmresearchnews.blogspot.fr/2013/04/mission-impossible-not-for-these-ibm.html

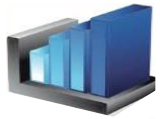
2 Racks: 10 p730 + DB2 + IBM FlashSystem



Application Sweet Spots: *Do More, Do it Faster!*

OLTP Databases

- Financial, gaming, real-time billing, trading, real-time monitoring, query acceleration (DB2/Oracle), etc.

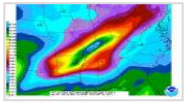


Analytical applications (OLAP)

- Business intelligence, batch processing, ERP systems, reporting, massive data feeds, etc.

Virtual Infrastructures

- VDI, Consolidated virtual infrastructures, user profiles, etc.



HPC/Computational Applications

- Simulation, modeling, rendering, FS metadata, scratch space, video on demand, thread efficiency, etc.

Cloud-scale Infrastructures

- On-demand computing, content distribution, web, caching, metadata, GPFS, active file management, etc.

Financial | Government | E-Commerce | HPC | Telecom

IBM inco Interamericana de Computos S.A.

A Few IBM FlashSystem Customers…

Thank You