



Open Source @ IBM

MILANO BICOCCA - **LINUX DAY** 2018

Lorenzo Laderchi

PowerAI Technical Sales
IBM Italia

mailto: laderchi@ibm.com
phone: +39 3479116760



Agenda:

- **POWER** – an introduction on the architecture
- **Opensource** @ IBM
- **OpenPower** Foundation – not only open software
- **PowerAI** – how the open innovation unlocks the workloads of the future

Architettura **POWER**

Performance Optimization With Enhanced RISC



OS supportati:



redhat



ubuntu

fedora



debian



CentOS

Caratteristiche HW:

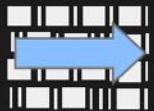
Massive IO bandwidth



1,000+ Concurrent Queries



Large-scale memory processing



Continuous data load



Parallel processing



PowerPC goes to.. Mars



2004
“Opportunity”



2012
“Curiosity”

The environment

- Extreme temperature range:
between **-140 °C** and **20 °C**
- Cosmic radiation

The **architecture**
of the computers running
on the rovers?

IBM PowerPC

Open Source is the engine of innovation

2000

IBM announced that it was adopting Linux and would support it with:

SERVERS

SOFTWARE

SERVICES

A major supporter of Linux, contributing broad resources to the Community to help
make Linux better good enough for Enterprise-level

2001

IBM announces 1 billion US\$ investment on Linux.

2004



2007



2011



2012



2013



Open Source is the engine of innovation

Code (Source)

1991: **10.000** lines of code
2017: **25.000.000** lines of code

"The kernel community added nearly 15 files and 7.500 lines of code. Every day."

Development model

New major kernel release
every 9-10 weeks



Community

Contributions since 2005:
• Developers: **15.600**
• Companies: **1.400**

Top Contributors

The **top 10 contributors**, including the groups “unknown” and “none,” make up nearly **54 %** of the total contributions to the kernel.



Top companies contributing to the Linux kernel, 4.8–4.13

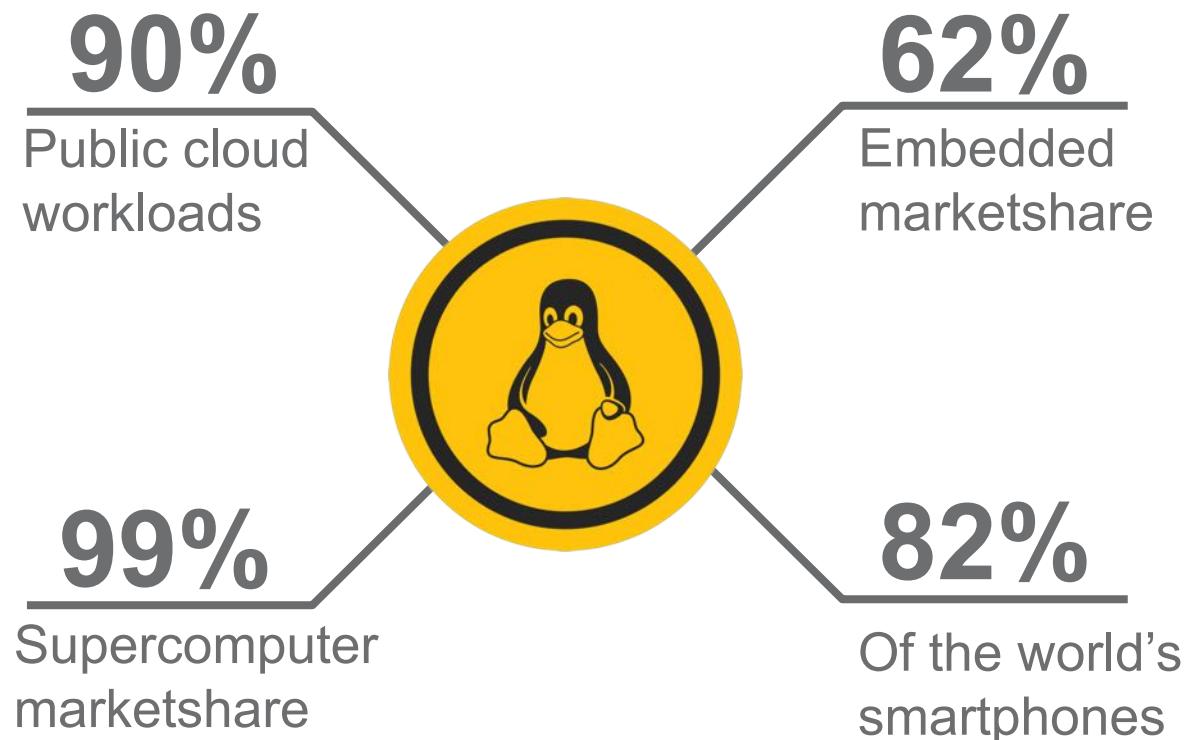
Company	Changes	%
Intel	10,833	13.1%
none	6,819	8.2%
Red Hat	5,965	7.2%
Linaro	4,636	5.6%
unknown	3,408	4.1%
IBM	3,359	4.1%
consultants	2,743	3.3%
Samsung	2,633	3.2%
SUSE	2,481	3.0%
Google	2,477	3.0%
AMD	2,215	2.7%
Renesas Electronics	1,680	2.0%
Mellanox	1,649	2.0%
Oracle	1,402	1.7%

Open Source is the engine of innovation

“The kernel which forms the core of the Linux system is the result of one of the largest cooperative software projects ever attempted.”



The result? An amazing success



IBM commitment to Open Source

IBM is one of the largest single contributor to Open Source Community Projects overall.

Open Innovation

Open Standard

Open Source

Open Governance

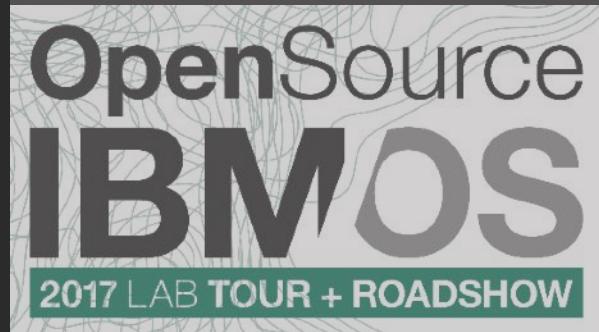
Contributes to more than 150 Open Source Projects



More than 700 Developer dedicated to
Open Source Projects

More than 500 patents donated to
Open Source Communities

The IBM «OpenSource Way»



Learn

Open Source @ IBM
Program touches

78,000
IBMer annually



Consume

Virtually all
IBM products
contain some
open source
• 40,363 pkgs
Per Year

IBM developerWorks

Open Source @ IBM

Learn

Open Source @ IBM
Your one-stop shop for open source development @ IBM
Find everything you need, from the tools, people, and steps to start contributing to open source projects.

Jeffrey Borek
@jrborek

Great turn out for #OpenSource @IBM Lab Tour event today, w Jochen talking about when #Standards & OS collide! 😊 IBM SW Labs, Germany #Cibud

Michael Pöck
@mfpock

Excited to talk to a full house about The Next Chapter of Open Source at IBM with @Kevin Nowka, @fmoore_1, and @bradtopol



Contribute

- **>62K OS Certs per year**
- **~10K IBM commits per month**



Connect

> 1000 active IBM Contributors

Working in key OS projects

Opensource @ IBM

ibm.biz/open_source

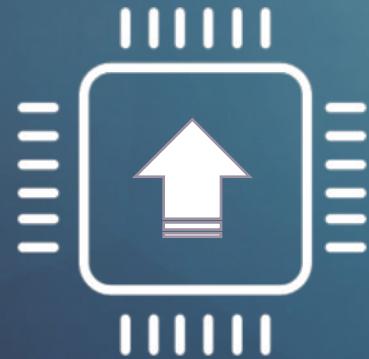


Why OpenPOWER™

Market Shifts



Moore's Law



Workload
Demands Up



Numerous IT
consumption models



Mature Open
software ecosystem

Why OpenPOWER™

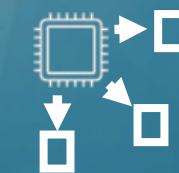
The Processor is No Longer the Only Source of IT Innovation



Moore's Law
Not Holding Up

Why OpenPOWER™

The Processor is No Longer the Only Source of IT Innovation



Accelerators



Storage



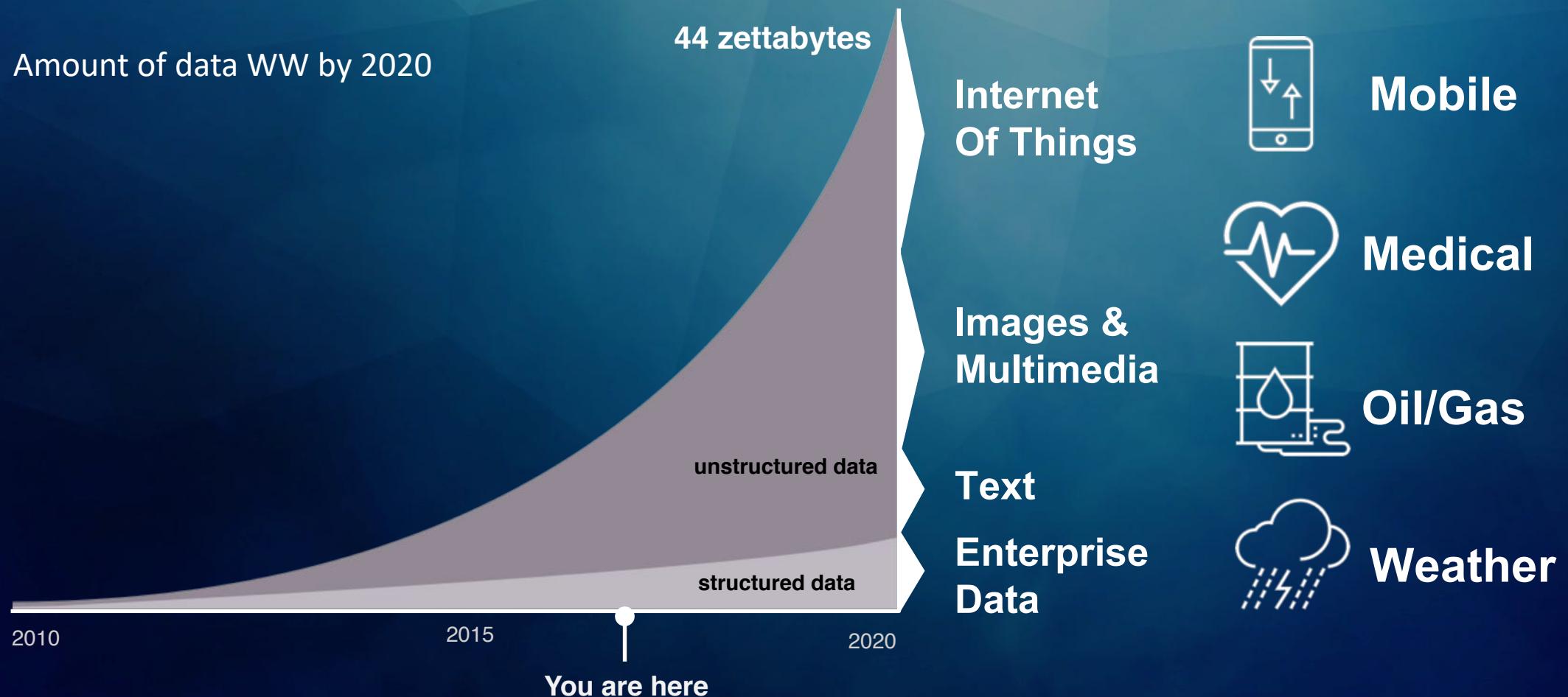
I/O Attach



Memory

Why OpenPOWER™

Data holds competitive value





OpenPOWER™

Ecosystem

IBM



TYAN



Google



**Founding
Members**
2013



Implementation / HPC / Research



System / Integration



System / Integration



I/O / Storage / Acceleration



Boards / Systems



Chip / SOC



335+

Members

33

Countries

70+

ISVs

FREE

Membership for
Universities

Active Membership
**From All Layers
of the Stack**

100k+

Linux Applications
Running on Power

2300 ISVs

Written Code
on Linux

**Partners
Bring
Systems
to Market**

150+

OpenPOWER Ready
Certified Products

20+

Systems Manufacturers

40+

POWER-based systems
shipping or in development

100+

Collaborative innovations
under way

OpenPOWER

Risultati

Tencent

*«Tencent recently purchased a number of OpenPOWER-based systems to add to its growing enterprise data center. With adoption of OpenPOWER technology, Tencent's overall efficiency has improved by more than 30%, and with savings of **30%** on rack and server resources.»*

PayPal

*«PayPal used IBM's OpenPOWER Systems and PowerAI to accelerate **deeplearning** research for fraud prevention by unlocking the computation power on extra large datasets with the Power architecture.»*

Google

*«Google announced that their IBM POWER9-based server, **Zaius**, is deployed and in the process of scaling up in their Data Center. Google's Maire Mahoney declared Zaius “Google Strong” and they are actively adding new production workloads onto Zaius and POWER9.»*

Summit

#1° of 500 - The World's Most Powerful and World's Smartest Supercomputer for Science



9.216

IBM POWER9
CPUs

25

gigabytes per
second between
nodes

27.648

NVIDIA Tesla
GPUs

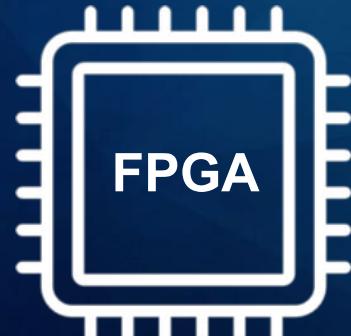
200

quadrillion
calculation per
second

250

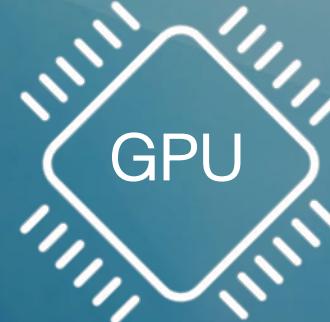
petabytes storage
capacity

Workload Accelerators + POWER9



CAPI

Reconfigurable hardware
Task customized
Low latency & power



NVIDIA NVLink

1000s of simple cores
High bandwidth, floating point, and parallelism

Uses:

- Compression
- Encryption
- high speed streaming search
- Monte Carlo simulations

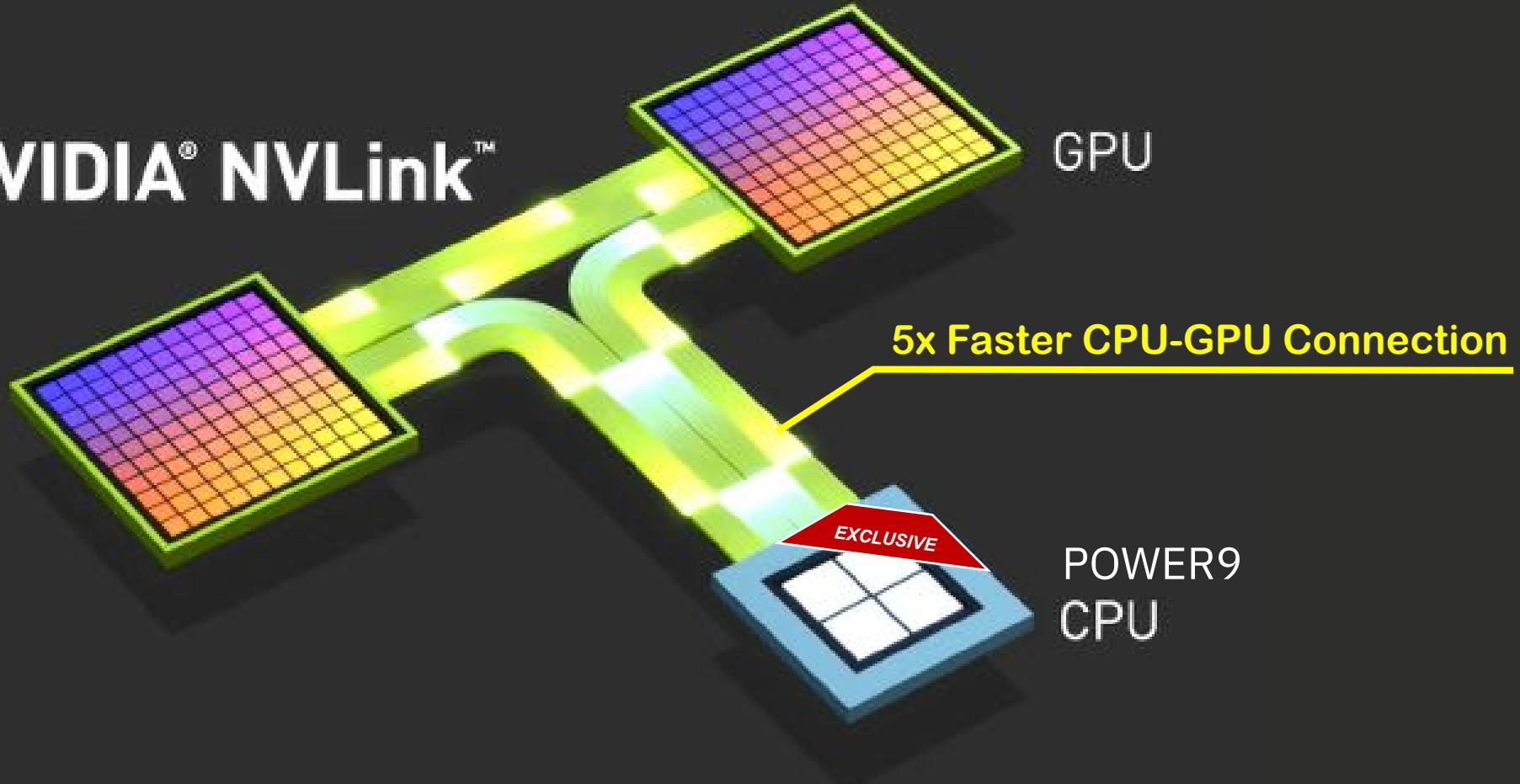
Uses:

- Deep Neural Networks
- Speech Recognition
- Chemistry Simulations
- JAVA
- Hadoop
- Graphics



IBM

NVIDIA® NVLink™

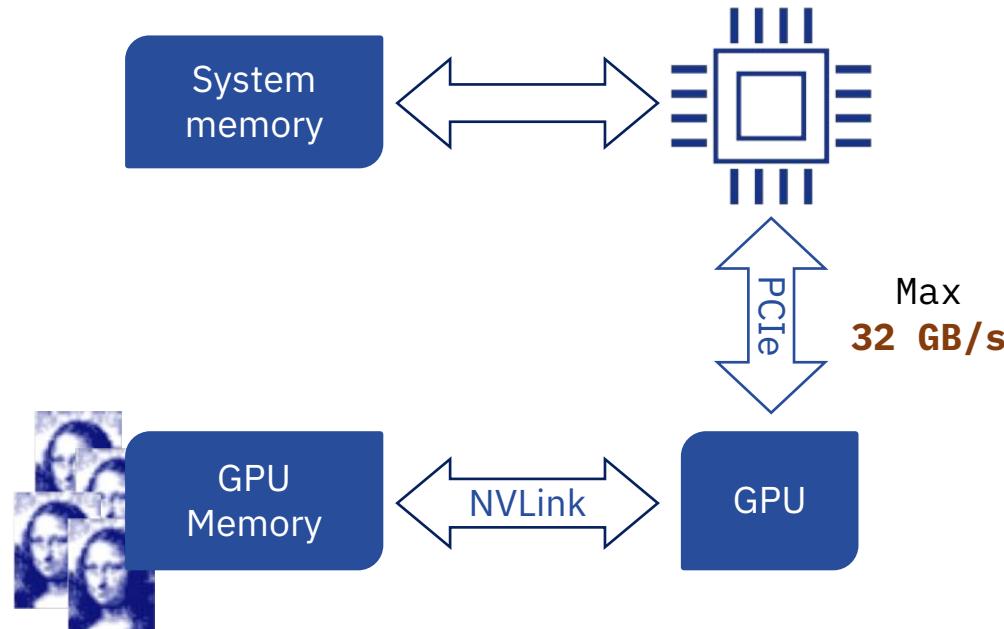


Large Model Support

Train larger, more complex models

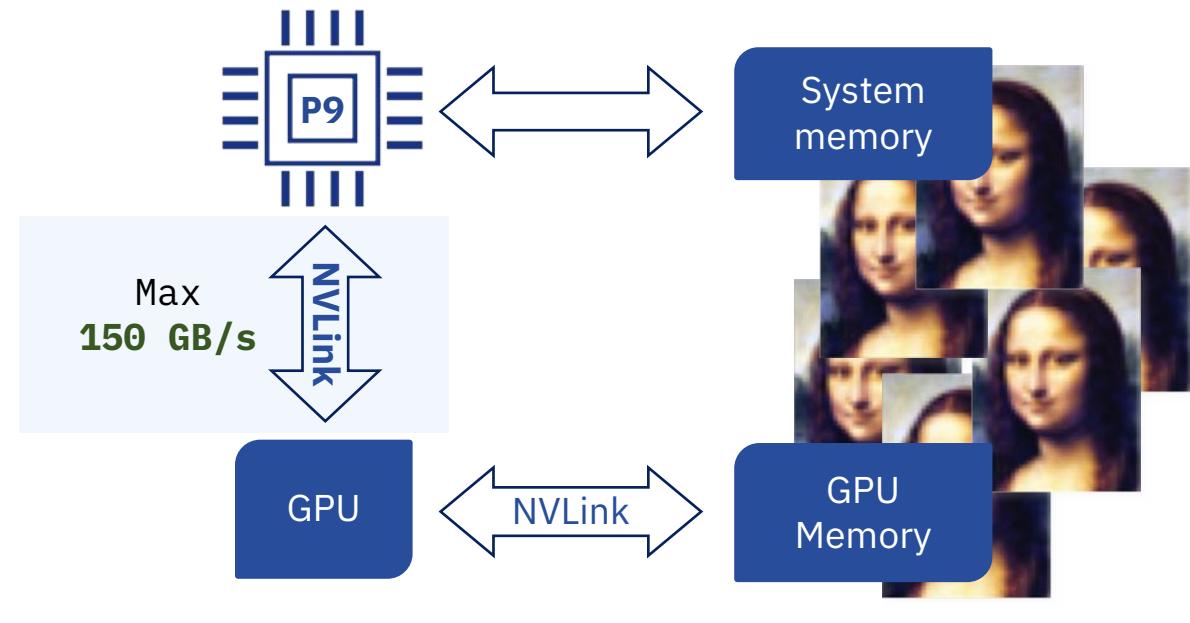
Traditional Model Support

Limited memory on GPU - up to 32 GB on the latest NVIDIA Tesla GPU - forces trade-off in model size / data resolution; PCIe can't provide memory coherence across the system (CPU + GPU) due to the limited bandwidth.



Large Model Support – only on PowerAI

Only available on Power Systems, the NVLink connection between CPU and GPUs makes possible to use the system memory (up to 2TB) for GPU's training tasks. The result? You can use larger datasets with high resolution data, to fully exploit the GPUs capabilities.

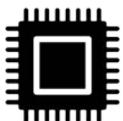


Large Model Support

Train larger, more complex models

Large AI Models Train
~4 Times Faster

Test specifications:



POWER9 (NVlink) **vs** Intel (PCIe)

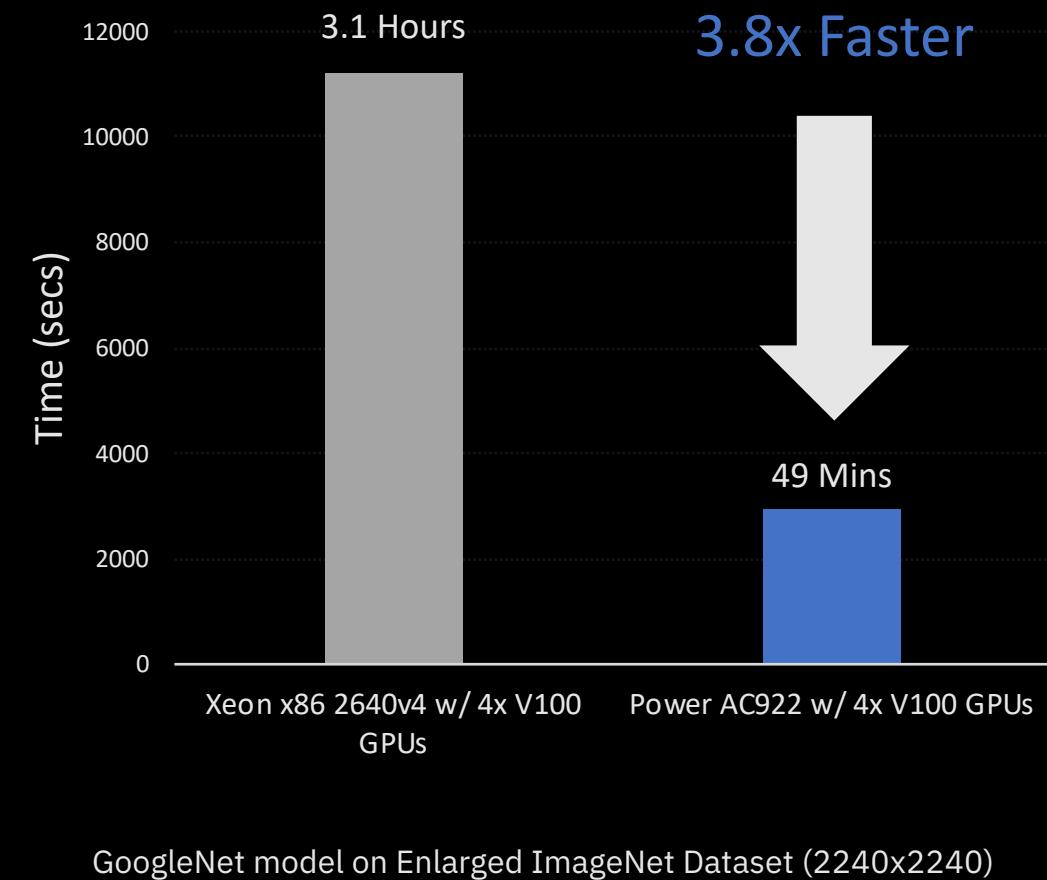


4x NVIDIA Tesla V100



CAFFE & LMS library

Caffe with LMS (Large Model Support) Runtime of 1000 Iterations

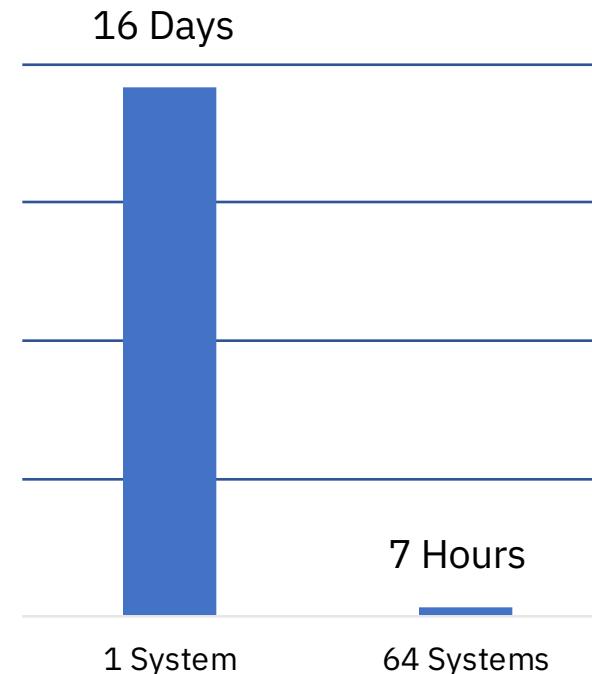


Distributed Deep Learning

Faster training and inferencing

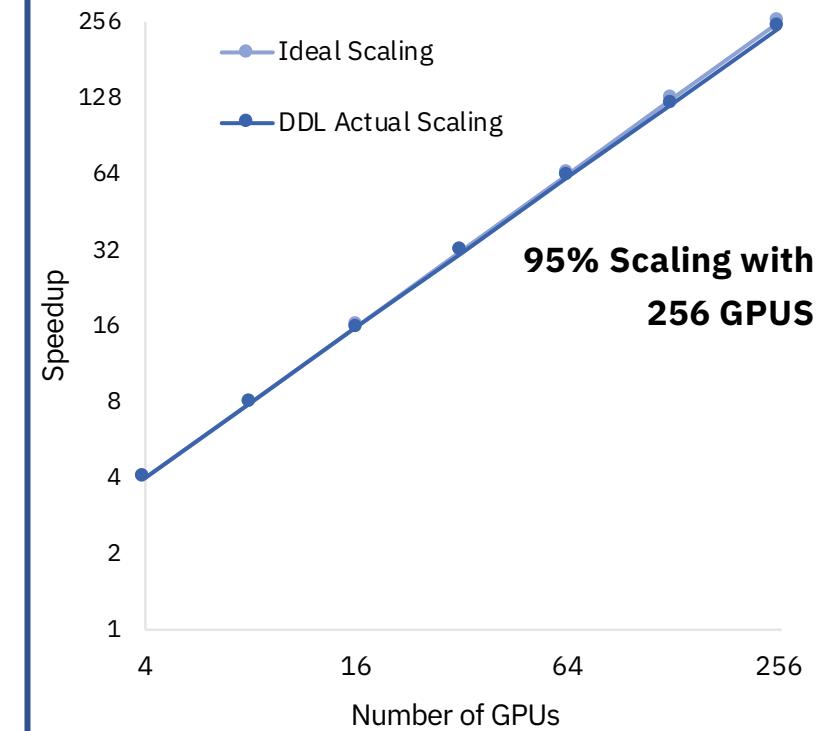
- Deep learning training on large datasets takes **days to weeks**
- On x86 servers scaling on multiple servers is limited due to connection speeds bottleneck
- PowerAI with DDL enables **scaling to 100s of servers**, with near-linear efficiency

16 Days Down to 7 Hours
58x Faster



ResNet-101, ImageNet-22K

Near Ideal Scaling to 256 GPUs



ResNet-50, ImageNet-1K

POWER 9

“The only processor specifically designed for the AI era.”

At a glance:

- 4x** more threads for high performance vs. x86
- 9.5x** more I/O bandwidth than x86
- 2.6x** more RAM possible vs. x86 (up to 2 TB)
- 1st** first CPU to deliver PCIe Gen 4

IBM Cognitive Systems



optimized.



PowerAI

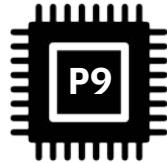
“IBM PowerAI: the complete environment for data science as a service”

Features:

- TF** Optimize the most popular DL opensource frameworks for this specific hardware
- Keras**
- Caffe**
- DDL** Distributed Deep Learning: scale training across multiple nodes
- LMS** Large Model Support: exploit the system memory (up to 2 TB) with GPUs
- SnapML** IBM library that enables GPU acceleration and scaling for ML algorithms

IBM Power Systems AC922

Technical specifications



POWER 9 processor

- ▷ **14nm** lithography
- ▷ **20cores** (x2 sockets)
- ▷ **2.00 GHz** (2.87 GHz Turbo)
- ▷ Up to **2TB RAM**



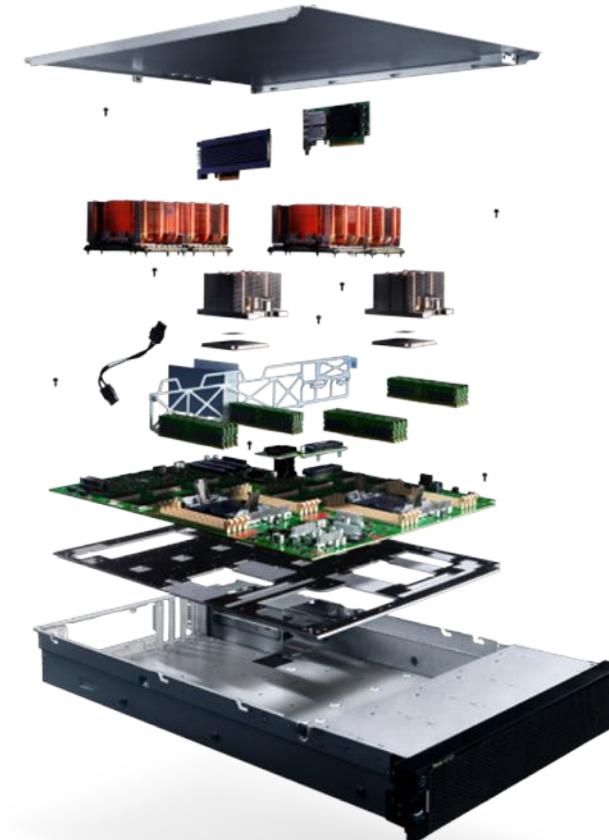
GPU Accelerated

- ▷ Up to 6x NVIDIA Tesla V100 (32gb) per server



I/O protocols

- ▷ OpenCAPI
- ▷ NVLink 2.0
- ▷ PCIe Gen4



PowerAI

Open-Source Based
Enterprise AI Platform

- Integrated & Supported AI Platform
- 3-4x Speedup for AI Training
- Ease of Use Tools for Data Scientists

Developer Ease-of-Use Tools

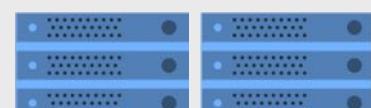
Open Source Frameworks: Supported Distribution



Caffe

SnapML

Faster Training Times via HW & SW Performance Optimizations



GPU-Accelerated
Power Servers



Storage

Come In
WE'RE
OPEN

