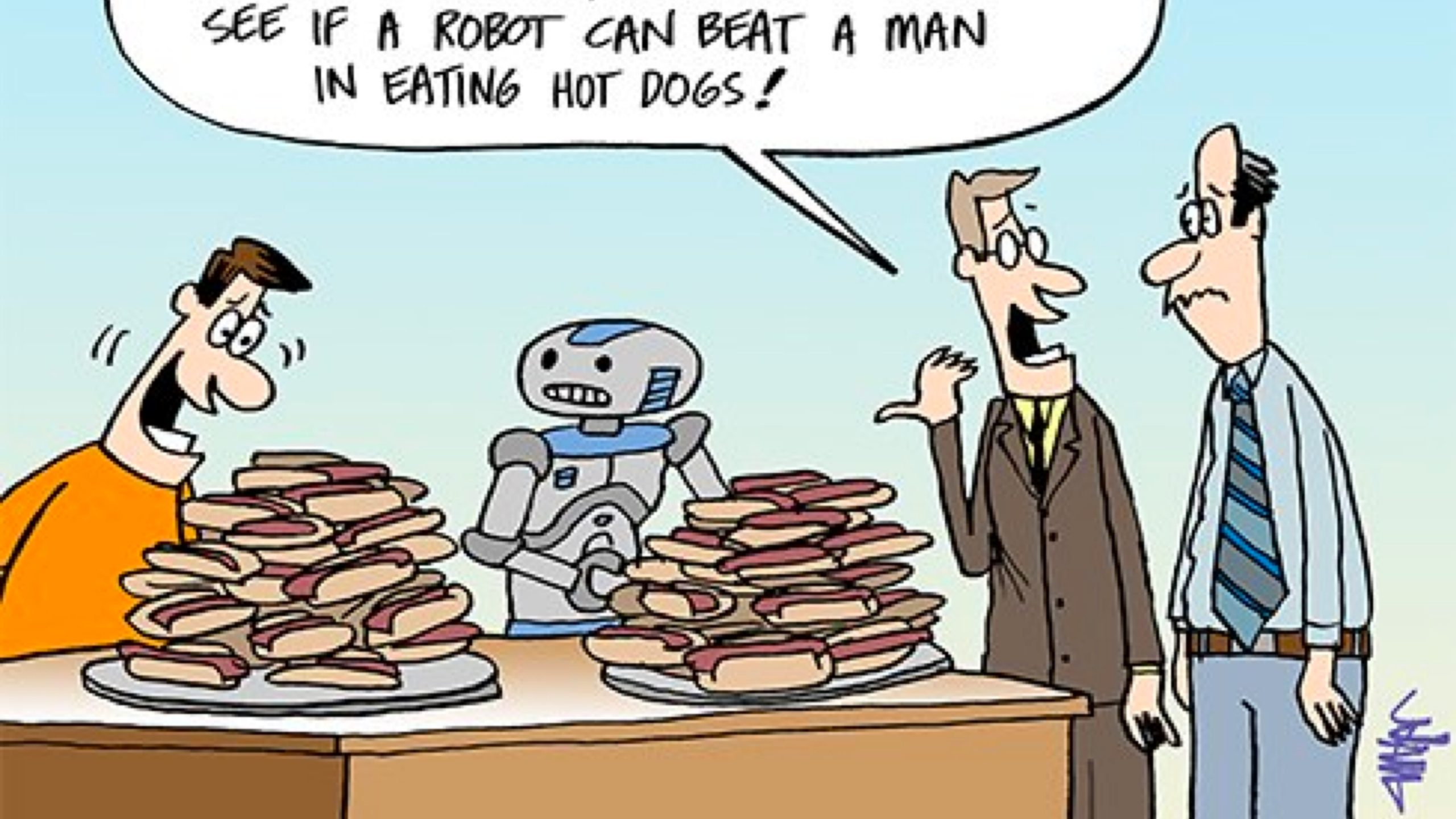
OpenPOWER™ & AI Workshop
Lund University

Ing. Florin Manaila
Senior Architect and Inventor
Cognitive Systems (HPC and Deep Learning)
IBM Systems Hardware Europe
Member of the IBM Academy of Technology (AoT)

Lund, March 21, 2019

By OpenPOWER Academia


SEE IF A ROBOT CAN BEAT A MAN IN EATING HOT DOGS!

AGENDA

OpenPOWER™

10:00 - 10:45:	**OpenPOWER and POWER9 features**
10:45 - 11:15:	**Accelerated Interconnect - Mellanox**
11:15 – 11:30:	Break
11:30 – 12:00:	**Accelerated Inferencing – Xilinix**
12:00 – 13:00:	Lunch
13:00 – 13:30:	**IBM PowerAI and AC922 deep learning system**
13:30 – 14:00:	**IBM PowerAI Large Model Support (LMS) and Distributed Deep Learning (DDL)**
14:00 – 15:15:	**Hands on Exercises** (PowerAI Vision, PowerAI – TF, Pytorch, SnapML, Rapids cuDF, Rapids cuML, BYOC)
15:15 – 15:30:	Break
15:30 – 17:00:	**Hands on Exercises** (PowerAI Vision, PowerAI – TF, Pytorch, SnapML, Rapids cuDF, Rapids cuML, BYOC)

OpenPOWER Foundation

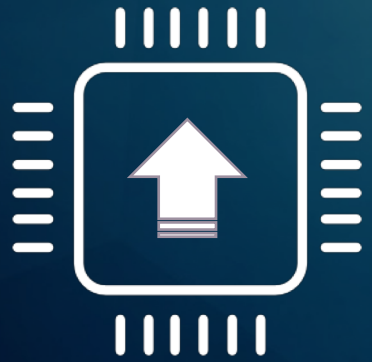
The OpenPOWER Foundation is an ***open technical community*** *enabling* ***collaborative development*** *on top of the POWER architecture, providing opportunity for* ***member differentiation*** *and industry growth.*

Why  OpenPOWER™

Why OpenPOWER™

Market Shifts

Moore's Law



Workload Demands Up

Numerous IT consumption models

Mature Open software ecosystem

Why OpenPOWER™

Market Shifts

Moore's Law

Workload Demands Up

Numerous IT consumption models

Mature Open software ecosystem

Strategy

Vibrant ecosystem through open development

Accelerated innovation through collaboration of partners

Amplified capabilities driving industry performance leadership

Why OpenPOWER™

Market Shifts

Moore's Law

Workload Demands Up

Numerous IT consumption models

Mature Open software ecosystem

Strategy

Vibrant ecosystem through open development

Accelerated innovation through collaboration of partners

Amplified capabilities driving industry performance leadership

Industry Adoption + Open Choice

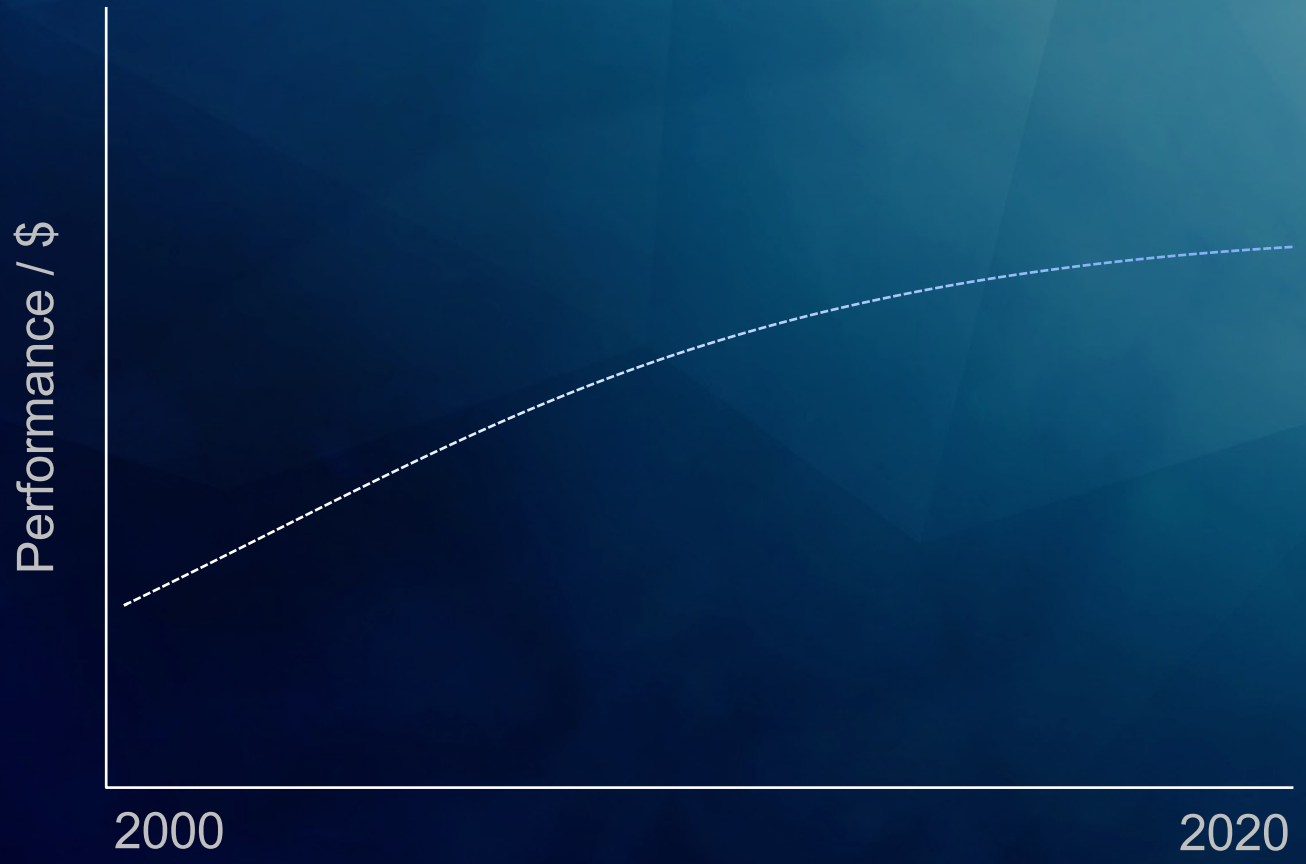
Cloud Computing Hyperscale & Large scale Datacenters

High Performance Computing & Analytics

Domestic IT Agendas

Why OpenPOWER™

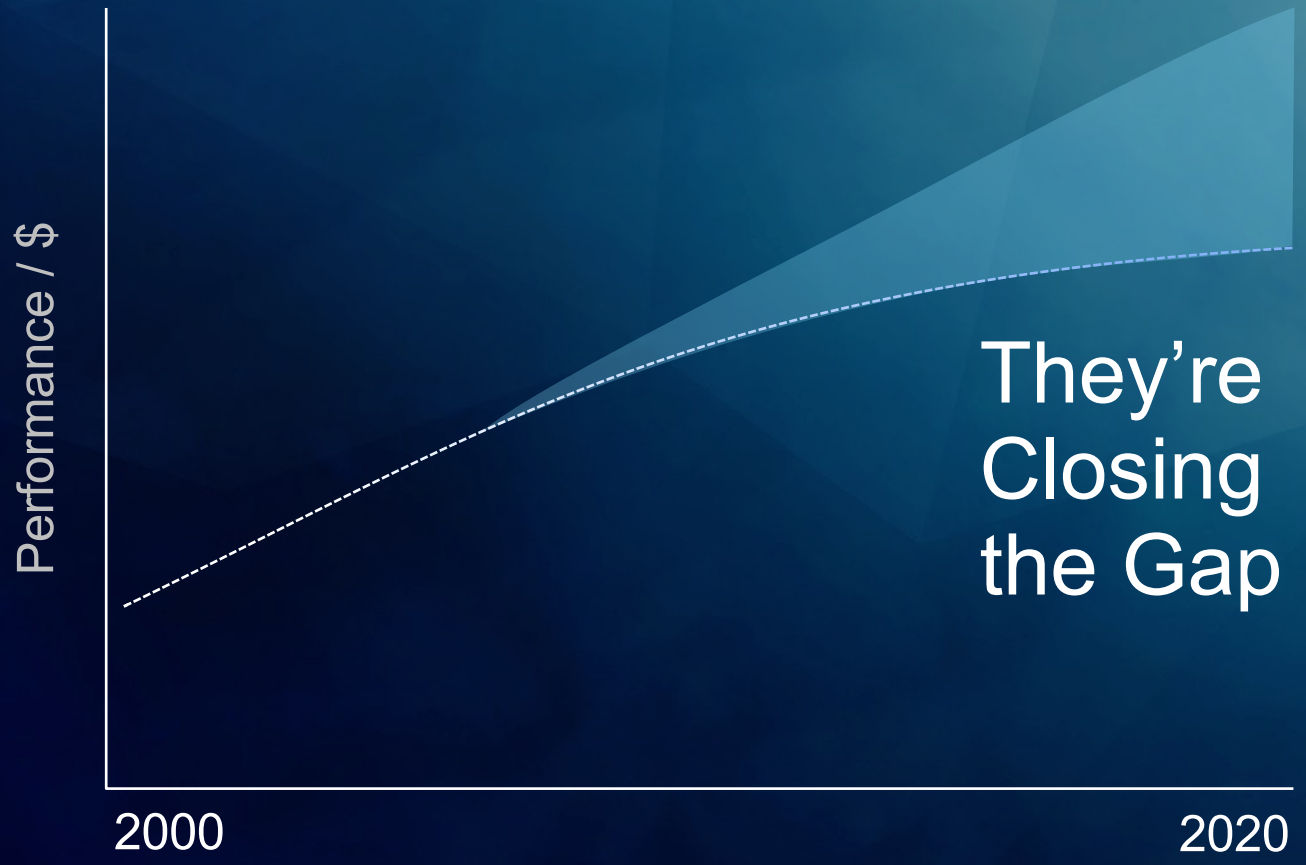
The Processor is No Longer the Only Source of IT Innovation



Moore's Law
Not Holding Up

Why OpenPOWER™

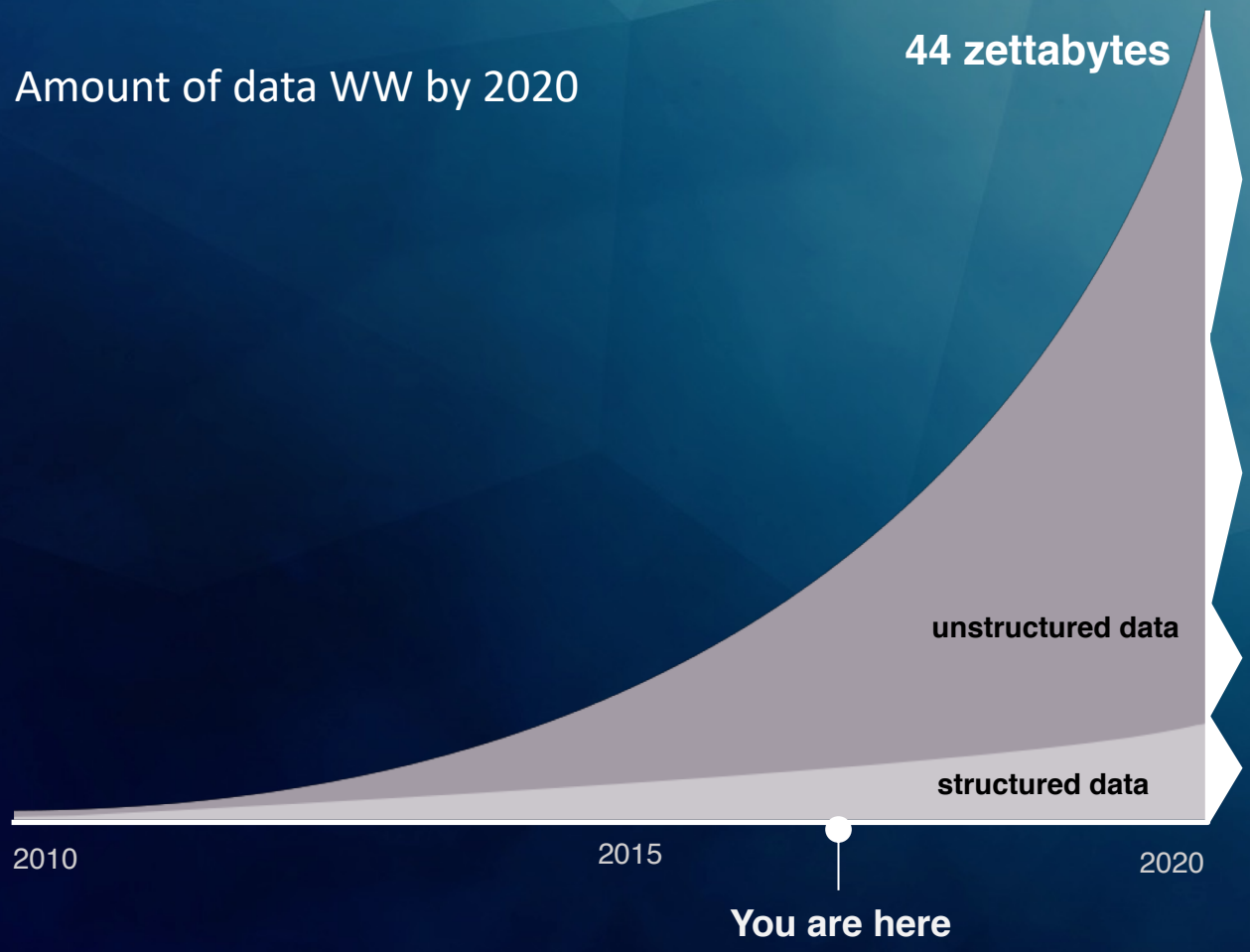
Next Gen Data Centers Depends on **Performance per $ Innovation from:**



- **Accelerators**
- **Storage**
- **I/O Attach**
- **Memory**

Why OpenPOWER™

Data holds competitive value



Internet Of Things

Images & Multimedia

Text

Enterprise Data

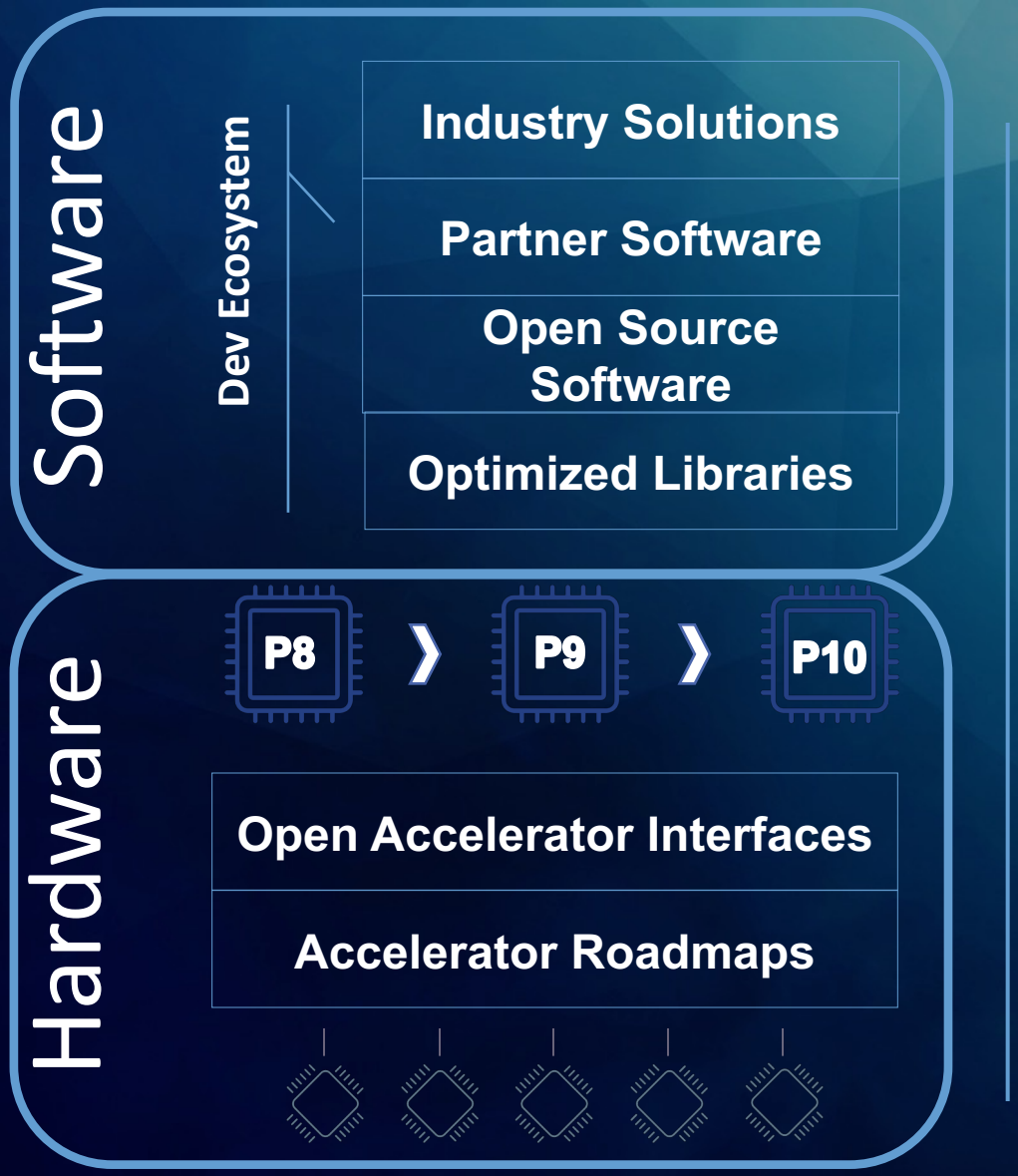
Mobile

Medical

Oil/Gas

Weather

Optimization for the Cognitive Era



3 Strategic Tenets

Software and Hardware Co-optimization
for the entire Data journey of AI workloads

Built on industry leading accelerated technologies

Deployed and delivered via a cloud operational model

Founding Members

2013

Vision › Execution › Adoption

Incorporated – Dec 2013

OpenPOWER Summit – March 2015

300+ Members – July 2017

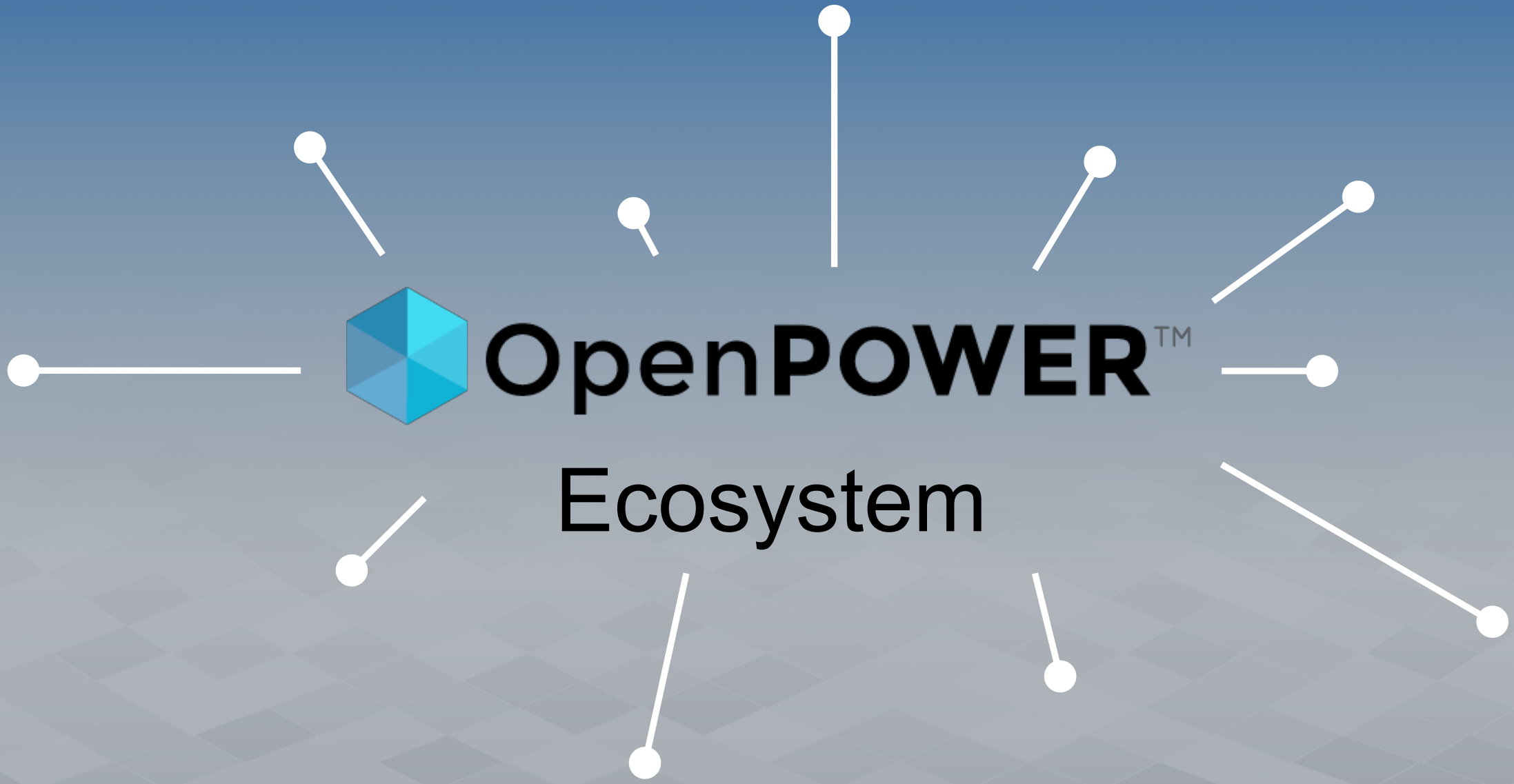
NVIDIA

TYAN

IBM

Mellanox TECHNOLOGIES

Google


OpenPOWER™
Ecosystem

This is What A Revolution Looks Like

Implementation / HPC / Research

Software

System / Integration

I/O / Storage / Acceleration

Boards / Systems

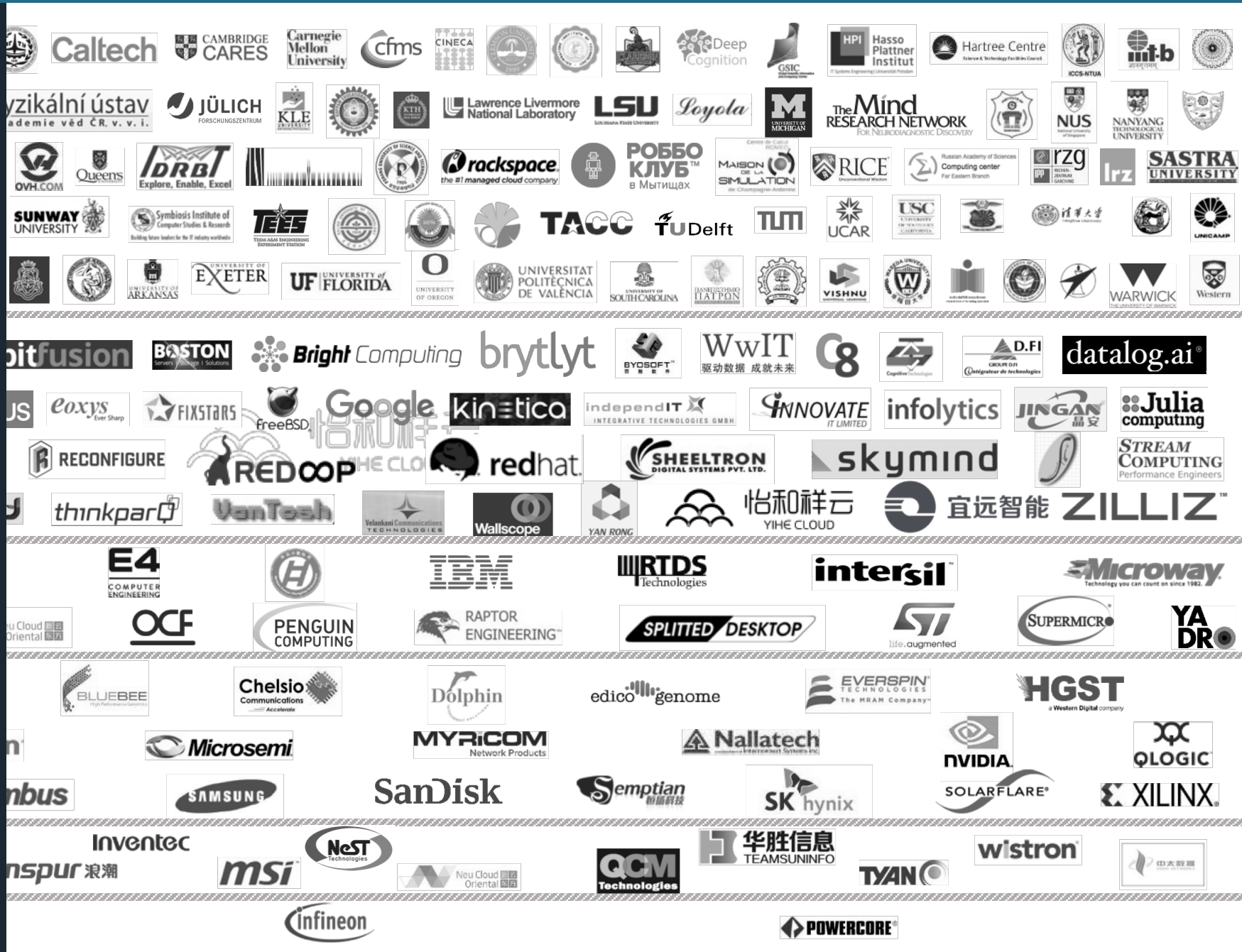
Chip / SOC

This is What A Revolution Looks Like

343+
Members

34
Countries

70+
ISVs



This is What A Revolution Looks Like

343+
Members

34
Countries

70+
ISVs

Active Membership
From All Layers of the Stack

100k+ Linux Applications Running on Power

2300 ISVs Written Code on Linux

This is What A Revolution Looks Like

343+
Members

34
Countries

70+
ISVs

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

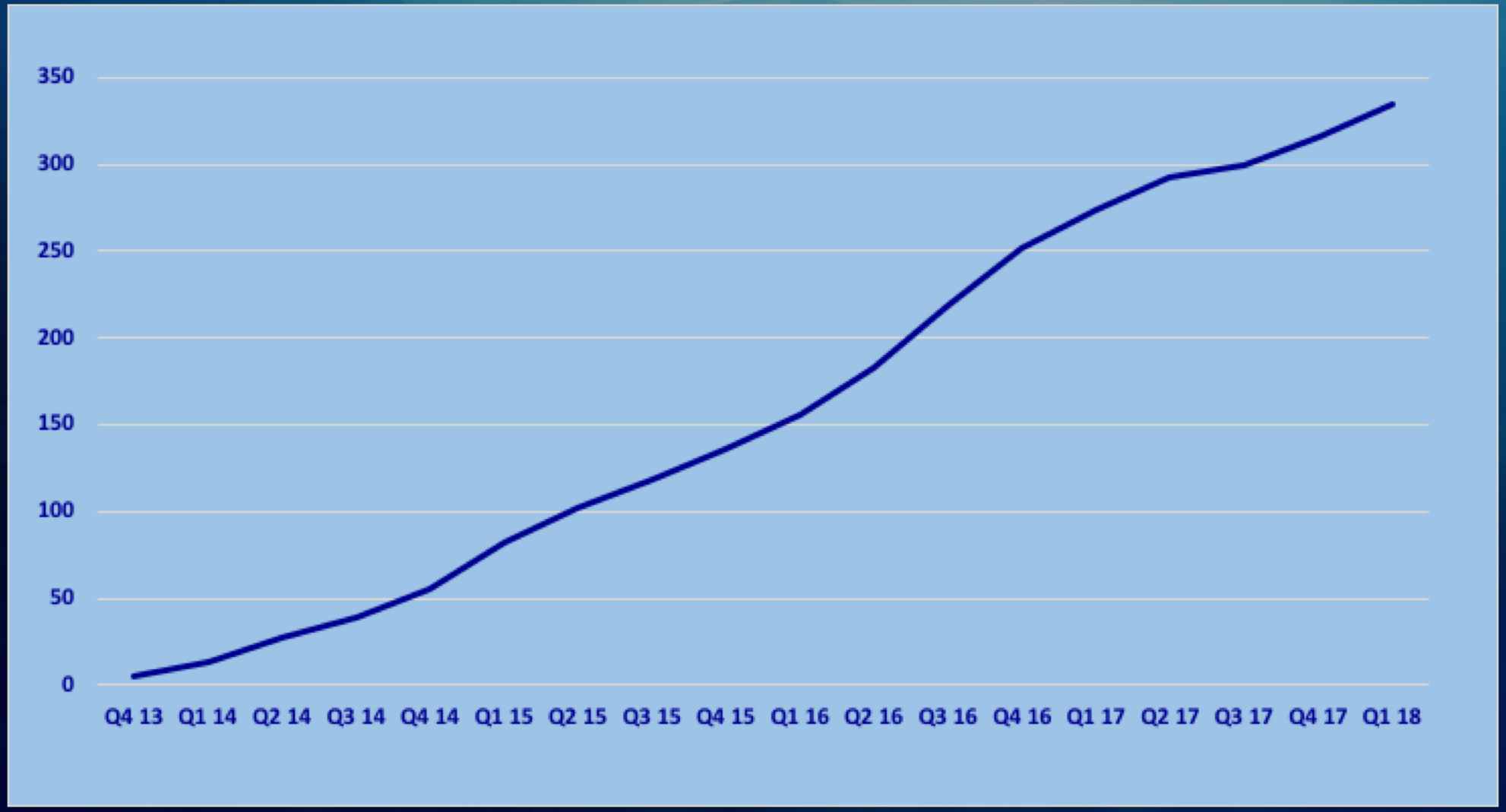
150+ OpenPOWER Ready Certified Products

20+ Systems Manufacturers

40+ POWER-based systems shipping or in development

100+ Collaborative innovations under way

OpenPOWER Foundation Membership Trend



This Ecosystem of Innovators

Creates True Differentiation in Performance and Cost

Growing Ecosystem Of **OpenPOWER Servers**

Micron, Rambus, BROADCOM, HGST, neo4j, IBM, POWERCORE, redhat, ubuntu Supported by Canonical, kinetica, 红旗 Linux, ZILLIZ, RED OOP, Semptian, 阿里云, 微云数聚, EasyStack, BONC 东方国信 让数据改变工作与生活, 娃步众成, FLYSLICE, CTACCEL, Think Force, 宜远智能, Nallatech, SAMSUNG, 怡和祥云, 灵雀云, redislabs, Microsemi, XILINX, ALPHA DATA, NVIDIA, Mellanox TECHNOLOGIES, NEC

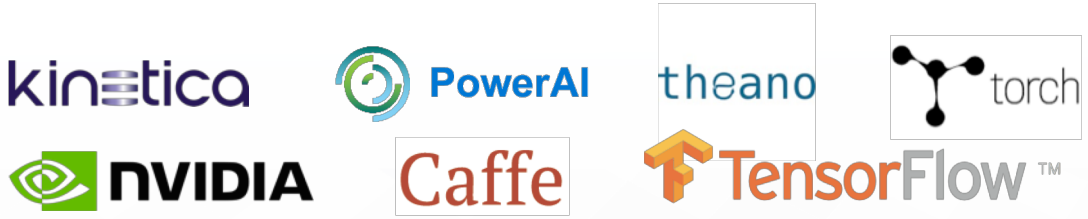
Growing Ecosystem Of **OpenPOWER Innovation**

Cirrascale, 中太服务器 ZOOM SERVER, NIMBIX, Neu Cloud 新云 Oriental 东方, ingrasys, GIGABYTE, inspur 浪潮, HITACHI Inspire the Next, TL 英特力, TEAMSUN, PENGUIN COMPUTING, YADRO, E4 COMPUTER ENGINEERING, msi, SUPERMICRO, IBM, Mark III Systems, RTDS Technologies, PET, AEWIN, Inventec, wistron

Ecosystem Driven **Customer Choice**

Enables OpenPOWER to meet the market demands by…

… **enabling AI Enterprises**
with offerings optimized for AI

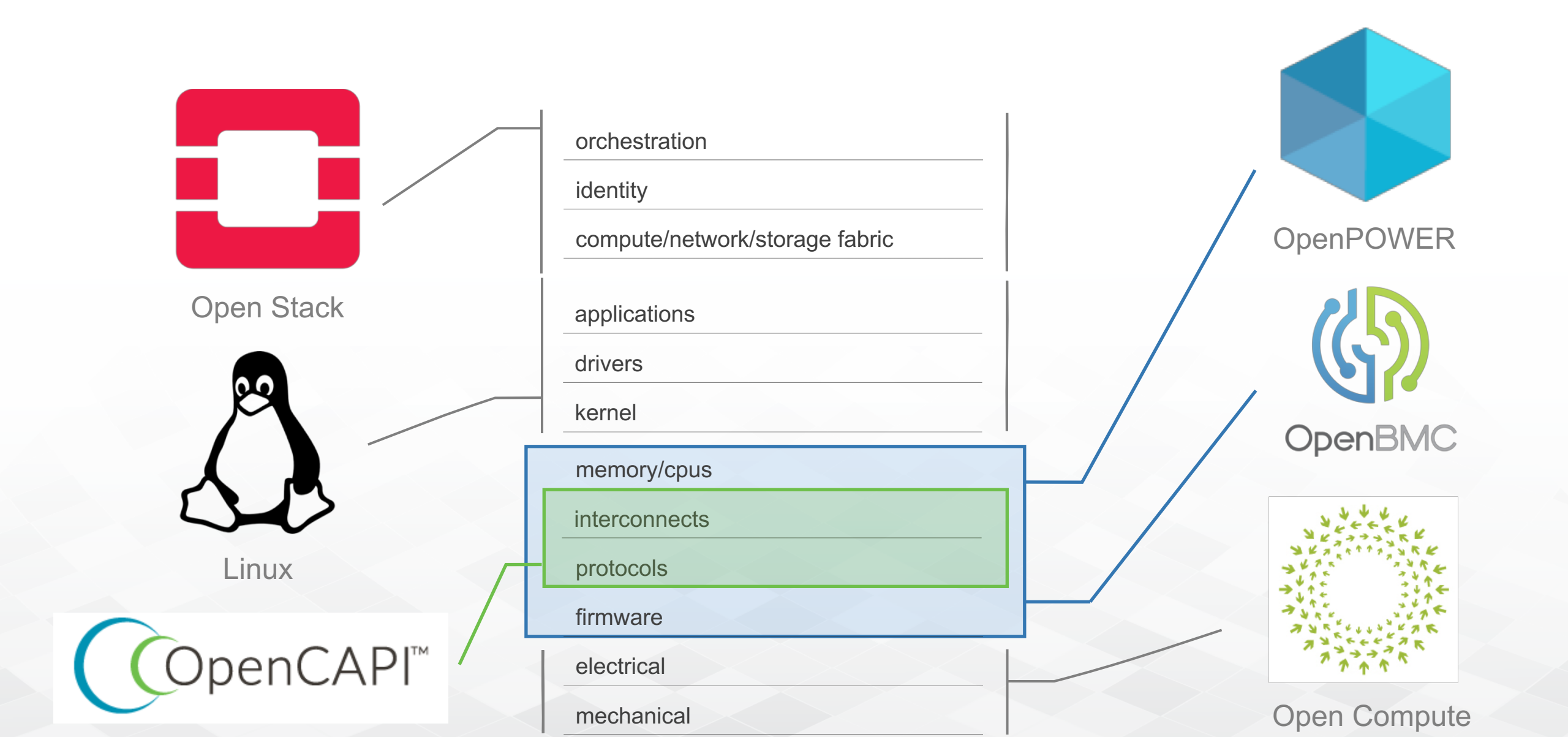


… **collaborating in Building**
Hyperscale Datacenters
with technology partners

… **delivering Price-Performance**
with cloud and software partners

Cross community

Collaboration is essential



…as is the option of an entirely Open software, firmware and hardware stack

Linux support for POWER

- ✓Same source and distribution release schedules as x86
- ✓Simplified x86 application migration with **little endian** distributions
- ✓Enterprise support for all three from IBM or distributors

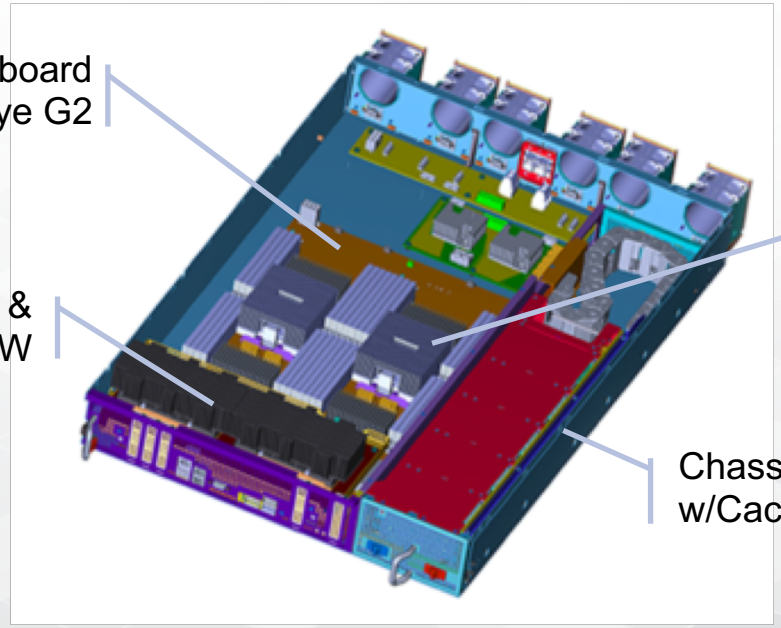
rackspace HOSTING + Google + OPEN Compute Project

Altering the status quo in our industry:

Novel + Community-Driven
Solutions to Very Difficult Problems

Open All The Way Down:

Hardware + Software + Developers
Community Feedback



Read more in the Rackspace Blog

What's Special about OpenPOWER?

POWER Processor Roadmap

	POWER7 Architecture		POWER8 Architecture		POWER9 Architecture			POWER10
	2010 POWER7 8 cores 45nm New Micro-Architecture New Process Technology	2012 POWER7+ 8 cores 32nm Enhanced Micro-Architecture New Process Technology	2014 POWER8 12 cores 22nm New Micro-Architecture New Process Technology	2016 POWER8 w/ NVLink 12 cores 22nm Enhanced Micro-Architecture With NVLink	2017 P9 SO 24 cores 14nm New Micro-Architecture Direct attach memory New Process Technology	2018 P9 SU 24 cores 14nm Enhanced Micro-Architecture Buffered Memory	2019 P9 w/ Adv. I/O 24 cores 14nm Enhanced Micro-Architecture New Memory Subsystem	2020+ P10 TBD cores New Micro-Architecture New Technology
Sustained Memory Bandwidth	Up To 65 GB/s	Up To 65 GB/s	Up To 210 GB/s	Up To 210 GB/s	Up To 150 GB/s	Up To 210 GB/s	Up To 350 GB/s	Up To 435 GB/s
Standard I/O Interconnect	PCIe Gen2	PCIe Gen2	PCIe Gen3	PCIe Gen3	PCIe Gen4 x48	PCIe Gen4 x48	PCIe Gen4 x48	PCIe Gen5
Advanced I/O Signaling	N/A	N/A	N/A	20 GT/s 160GB/s	25 GT/s 300GB/s	25 GT/s 300GB/s	25 GT/s 300GB/s	32 & 50 GT/s
Advanced I/O Architecture	N/A	N/A	CAPI 1.0	CAPI 1.0 , NVLink 1.0	CAPI 2.0, OpenCAPI3.0, NVLink2.0	CAPI 2.0, OpenCAPI3.0, NVLink2.0	CAPI 2.0, OpenCAPI4.0, NVLink	TBD

Statement of Direction, Subject to Change

POWER9 – Data Capacity & Throughput

Big Caches for Massively Parallel Compute and Heterogeneous Interaction

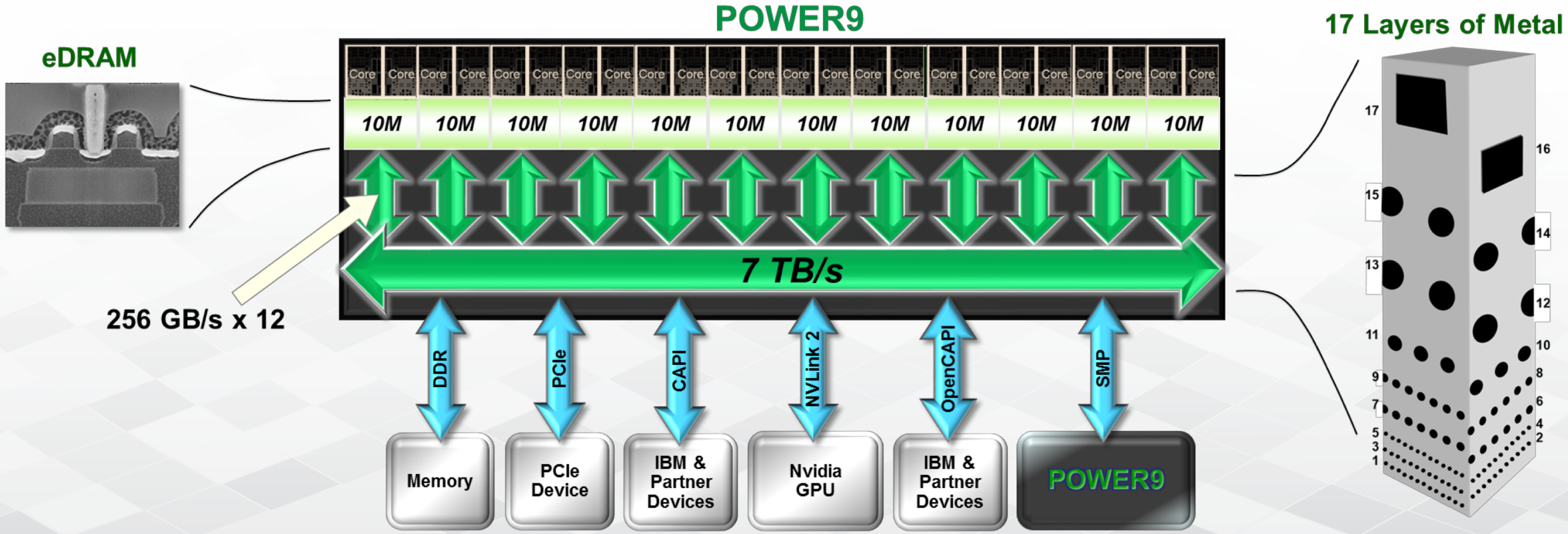
L3 Cache: 120 MB Shared Capacity NUCA Cache

- 10 MB Capacity + 512k L2 per SMT8 Core
- Enhanced Replacement with Reuse & Data-Type Awareness
  12 x 20 way associativity

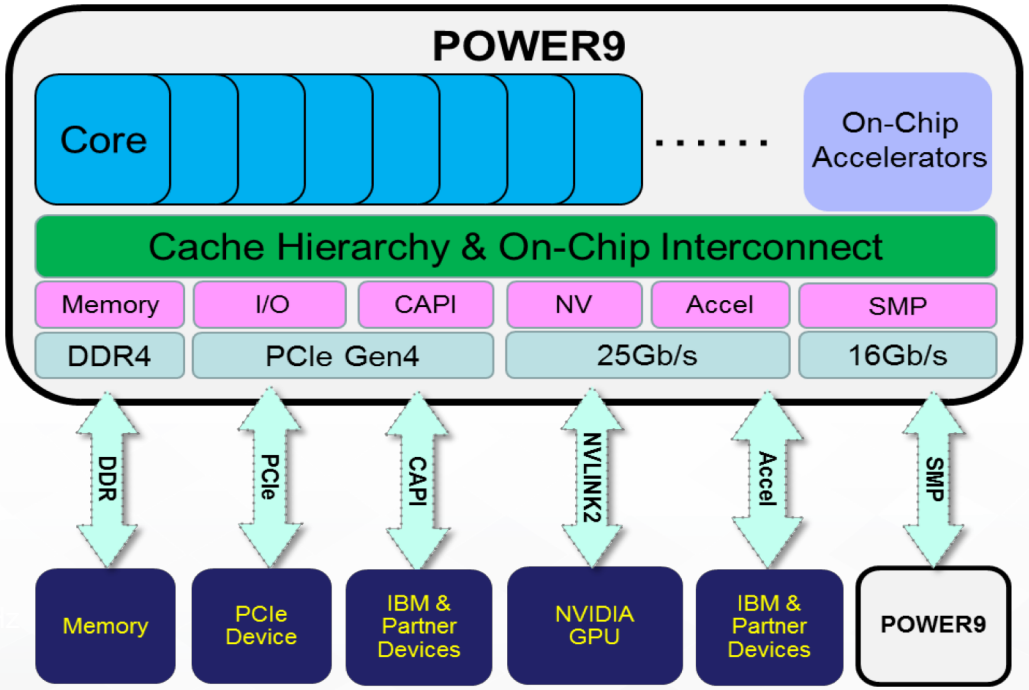
Extreme Switching Bandwidth for the Most Demanding Compute and Accelerated Workloads

High-Throughput On-Chip Fabric

- Over 7 TB/s On-chip Switch
- Move Data in/out at 256 GB/s per SMT8 Core



POWER9 SO with Advanced Accelerator Attach



24 newly designed POWER9 cores

- Leveraging execution slices for improved performance on cognitive, analytic, and big-data applications

Large, low-latency, eDRAM cache for big datasets

Global Foundries 14HP FinFET technology with eDRAM: 8B transistors

Cloud-focused innovation in Energy Efficiency, Security, and Quality of Service

State-of-the-art IO PCIe Gen 4: 48 lanes

Leadership platforms for hardware acceleration

- High bandwidth, GPU interconnect (NV link2.0)
- Next-generation CAPI2.0 interface for coherent accelerator and storage attach
- On-chip compression & cryptography accelerators
- New 25Gb/s advanced accelerator attach bus

1st chip in POWER9 family

- Optimized for 2 socket scale out servers & hyperscale datacenters
- DDR4 direct attach: 8 memory channels, >120 GB/s Sustained

Full POWER9 family will address a broad range of scale out & enterprise servers

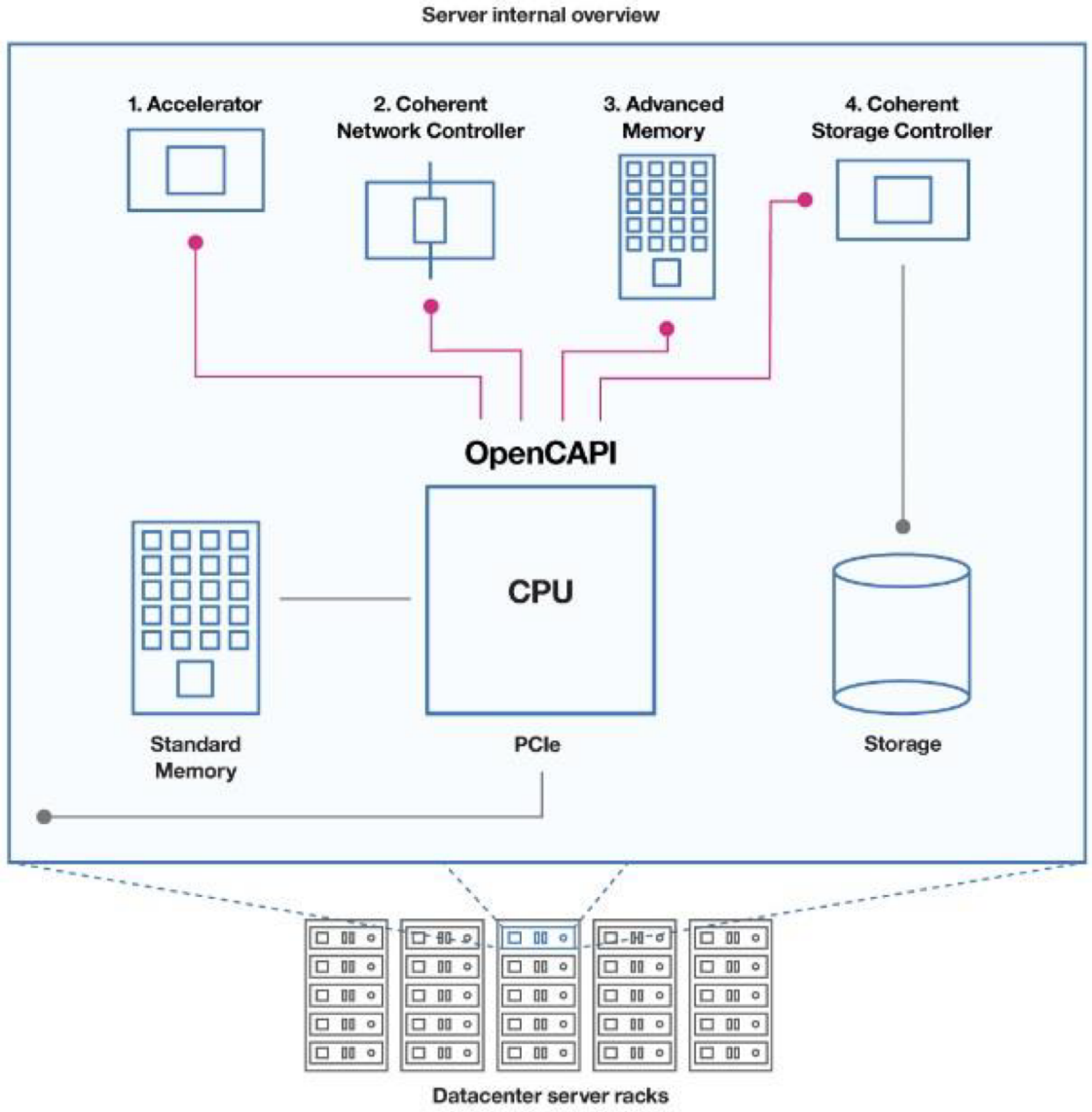
OpenCAPI

1. **Accelerators:** The performance, virtual addressing and coherence capabilities allow FPGA and ASIC accelerators to behave as if they were integrated into a custom microprocessor.

2. **Coherent Network Controller:** OpenCAPI provides the bandwidth that will be needed to support rapidly increasing network speeds Network controllers based on virtual addressing can eliminate software overhead without the programming complexity usually associated with user-level networking protocols.

3. **Advanced Memory:** OpenCAPI allows system designers to take full advantage of emerging memory technologies to change the economics of the datacenter.

4. **Coherent Storage Controller:** OpenCAPI allows storage controllers to bypass kernel software overhead, enabling extreme IOPS performance without wasting valuable CPU cycles.



Workload Accelerators + POWER9

CAPI

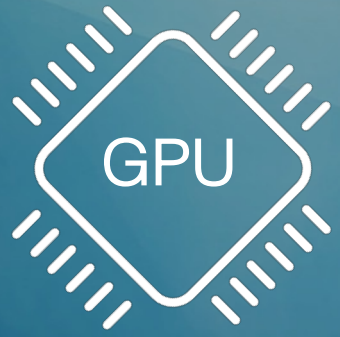
Reconfigurable hardware

Task customized

Low latency & power

Uses:

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



NVIDIA NVLink

1000s of simple cores

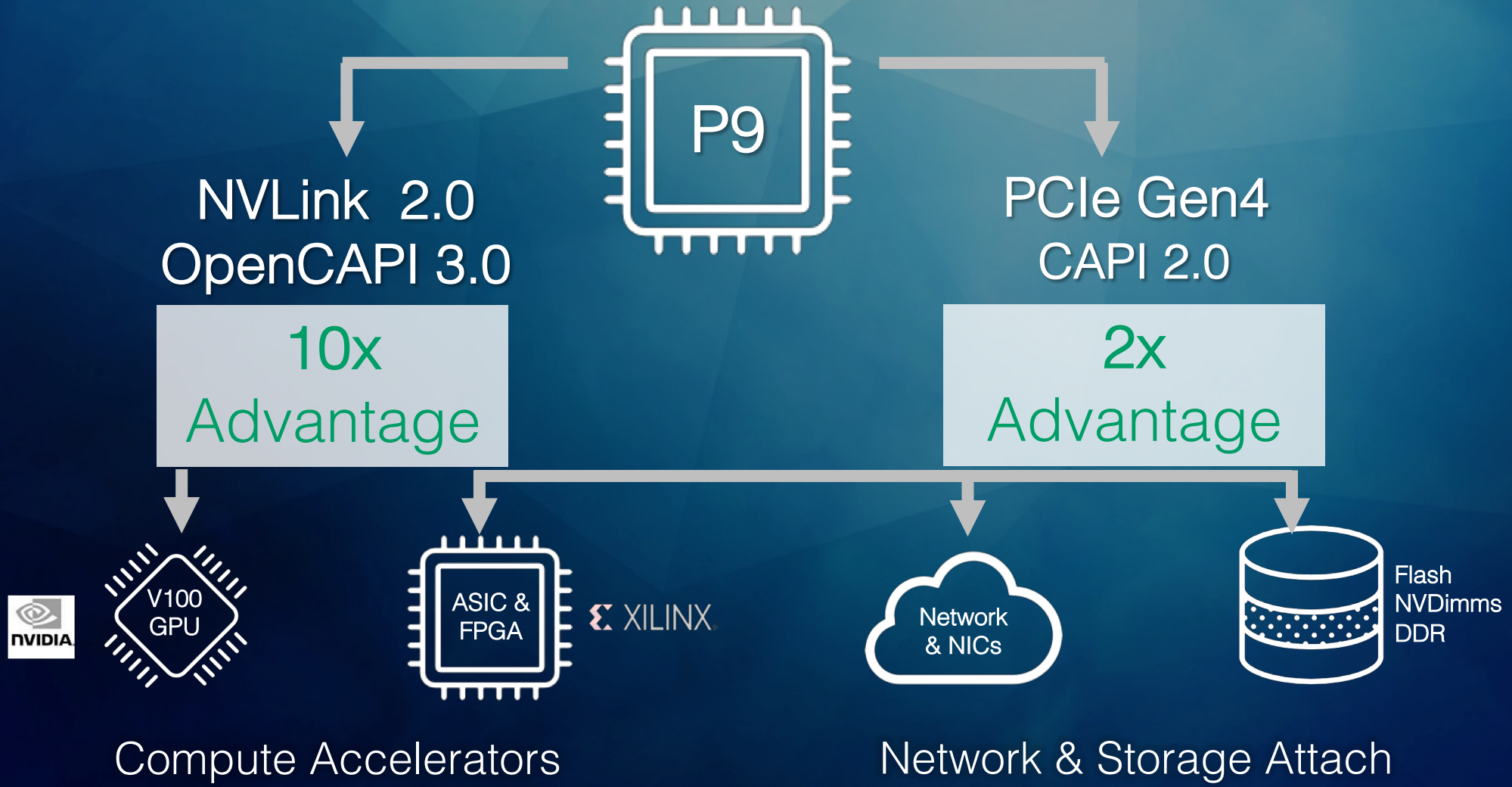
High bandwidth, floating point, and parallelism

Uses:

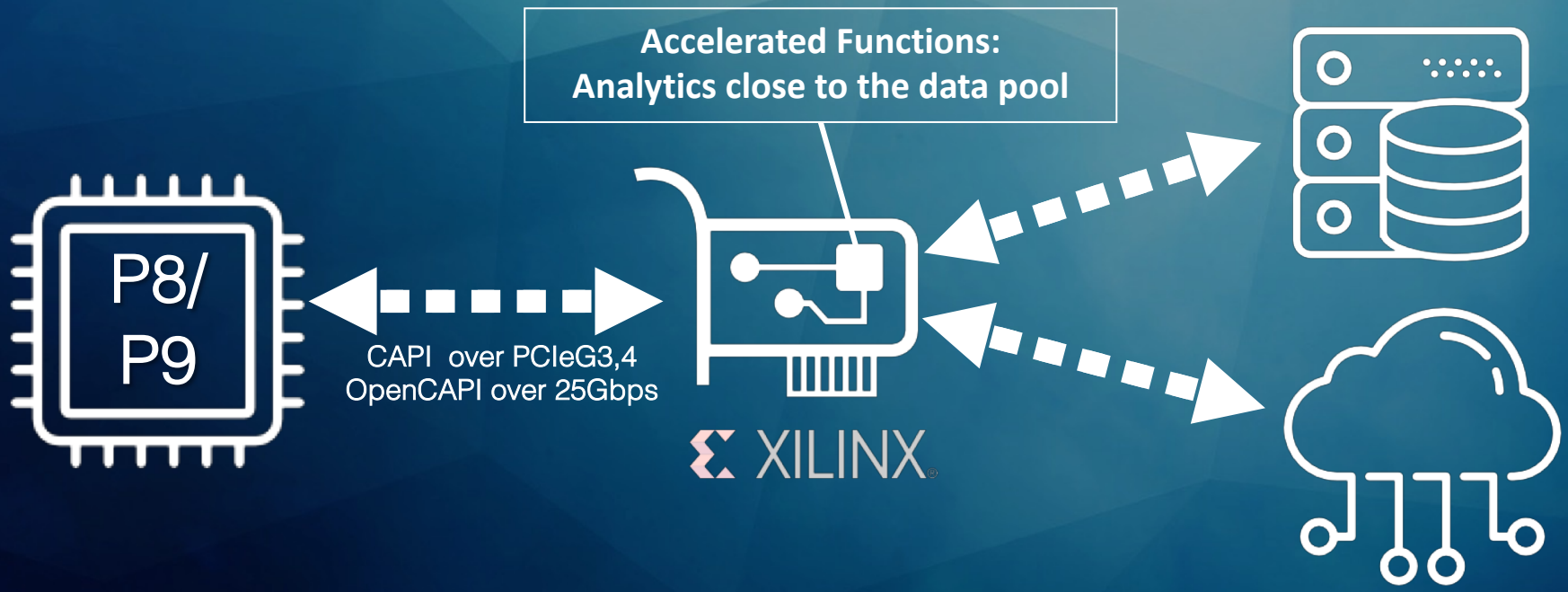
- Deep Neural Networks
- Speech Recognition
- Chemistry Simulations
- JAVA
- Hadoop
- Graphics

Power9, the Next Generation

CPU Designed for Accelerated Computing

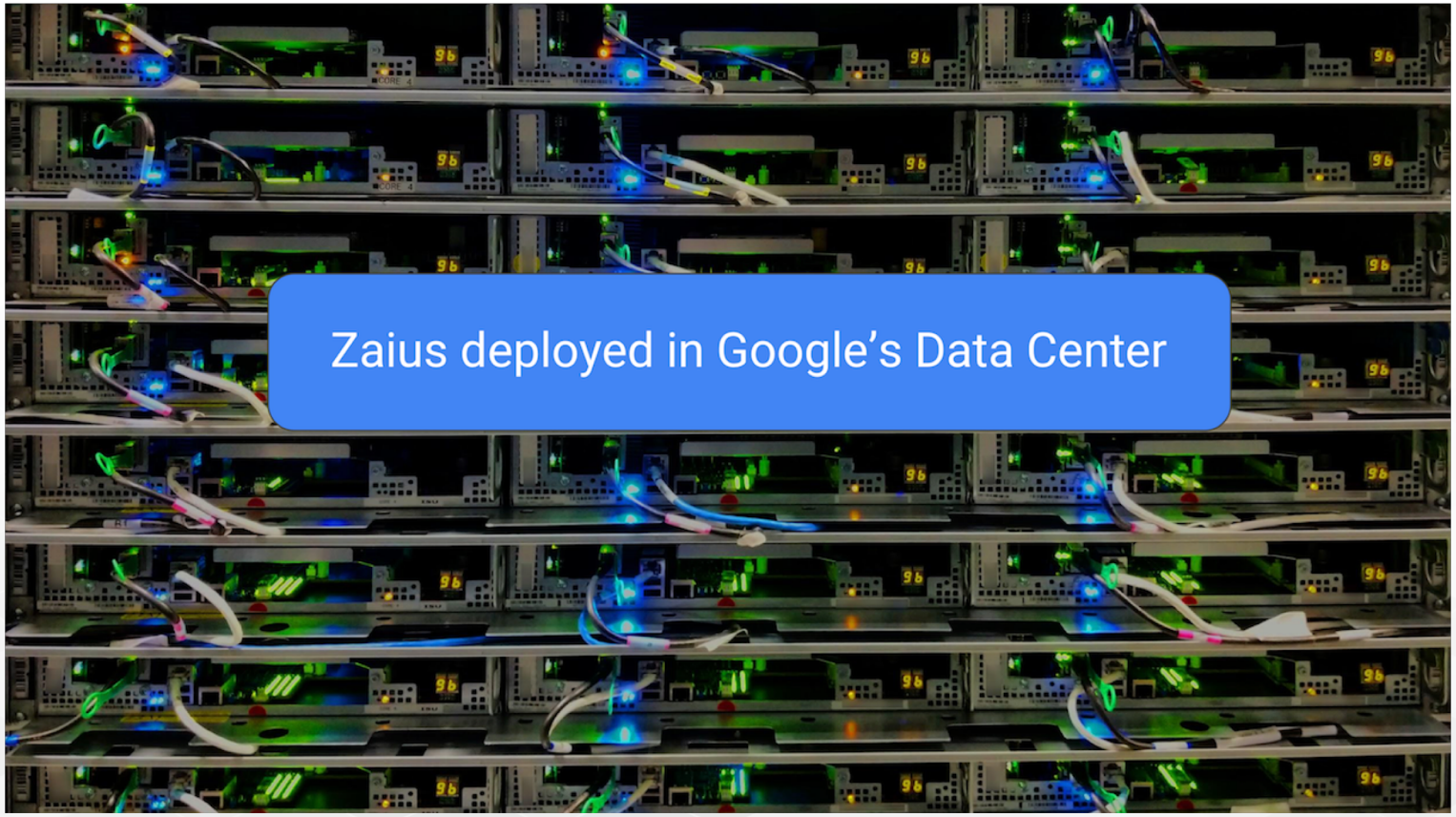


CAPI Coherent Accelerator Processor Interface



- Secure, trusted, and virtualized
- Greater bandwidth & access to your data
- Enables applications not possible on I/O
- Removes device driver and it's code stack

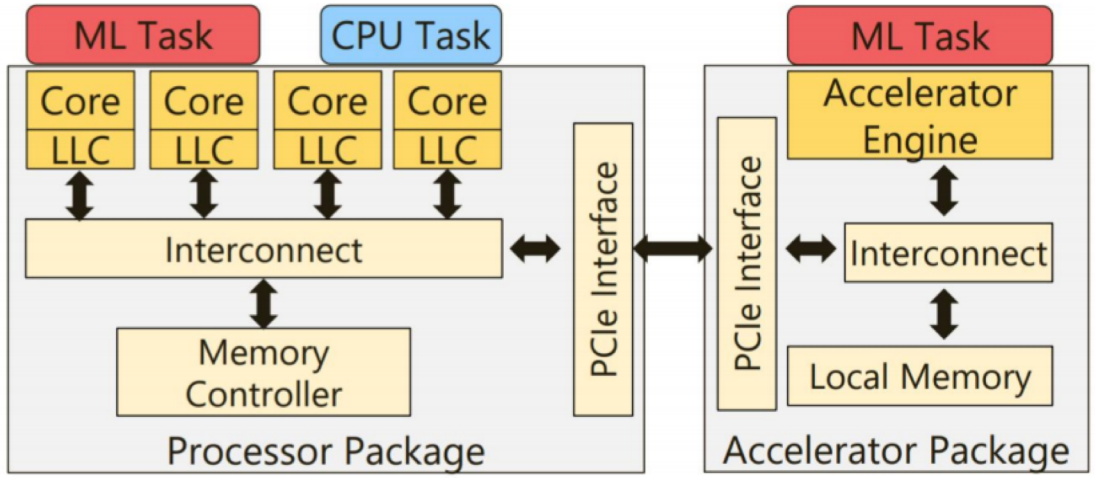
Who is using OpenPOWER and How?

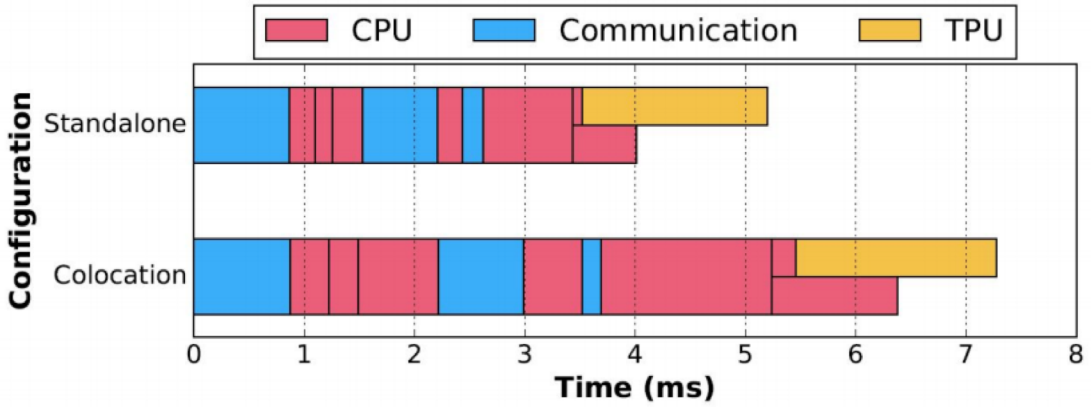


"We're declaring this Zaius platform to be Google Strong"
– Máire Mahony, Google

Source: https://openpowerfoundation.org/wp-content/uploads/2018/04/Maire-Mahony.pdf

More Memory Bandwidth

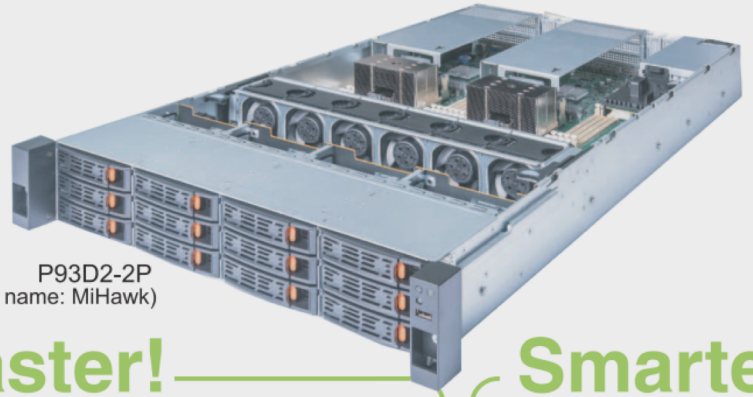




DRAM Bandwidth Contention
Impacts RNN server execution

Kelp QoS for Accelerated ML Systems - ISCA 2018 Submission Source:Haishan Zhu, David Low

OpenPOWER™ wistron

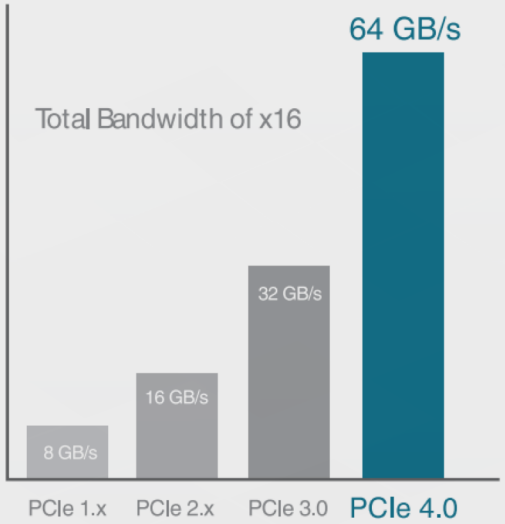


P93D2-2P
(Code name: MiHawk)

Built For The Cognitive Era

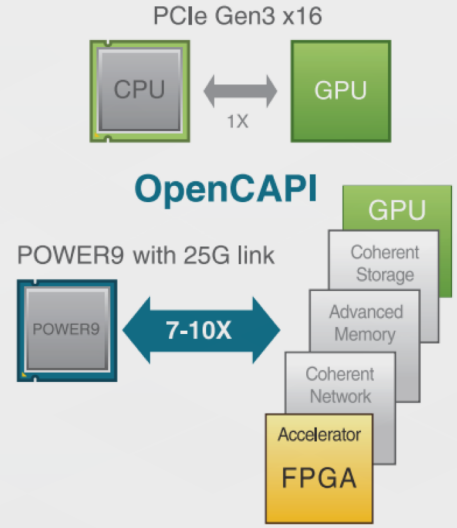
Faster!

PCIe 4 WW 1st PCIe Gen4 to enhance the I/O data rate.



Smarter!

OpenCAPI™ Brand new design for AI, Big Data and Cloud.



Better!

POWER9 Extreme technology for next decade cognitive era.

Feature	Intel Xeon 61xx Processor	IBM POWER9 (LaGrange)
# of Threads	2 per Core Up to 44	**4 per Core Up to 88**
L3 Cache	Up to 30.25 MB	**Up to 120 MB**
Max # of Memory Channels	6	**8**
Max Memory Size	768 GB	**4TB**
Open Innovation Interface	None	**OpenCAPI by 25G link**

https://openpowerfoundation.org/wp-content/uploads/2018/06/MiHawk-down.pdf

MiHawk can install up to twenty-four NVMe U.2 drives and could reach up to 80GB/s on optimal IO bandwidth

Who is using OpenPOWER and How?

Tencent 腾讯

“With adoption of OpenPOWER the overall efficiency has improved by more than 30% more performance @ 30% less rack and server resources.”

3x
vs x86
4
World Records

Spark Terasort

In recent results running former x86 infrastructure, with 2/3rd fewer servers.

512x SuperMicro POWER8 servers

Alibaba Cloud

“OpenPOWER servers are available on the Ali X-Dragon Cloud platform for pilot programs.”

PayPal

"IBM Power System and PowerAI helped PayPal to accelerate the deep learning research on Fraud Prevention problems by unlocking the computation power on extra large datasets with the Power architecture." - PayPal Data Science Team

Summit Overview

OpenPOWER™

Compute System

10.2 PB Total Memory
256 compute racks
4,608 compute nodes
Mellanox EDR IB fabric
200 PFLOPS
~13 MW

Compute Rack

18 Compute Servers
Warm water (70°F direct-cooled components)
RDHX for air-cooled components

39.7 TB Memory/rack
55 KW max power/rack

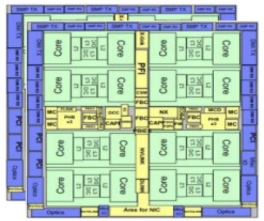
Compute Node

2 x POWER9
6 x NVIDIA GV100
NVMe-compatible PCIe 1600 GB SSD

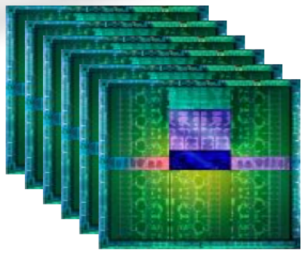
25 GB/s EDR IB- (2 ports)
512 GB DRAM- (DDR4)
96 GB HBM- (3D Stacked)
Coherent Shared Memory

Components

IBM POWER9
- 22 Cores
- 4 Threads/core
- NVLink

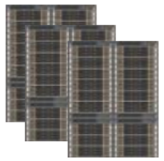


NVIDIA GV100
- 7 TF
- 16 GB @ 0.9 TB/s
- NVLink



GPFS File System

250 PB storage
2.5 TB/s read, 2.5 TB/s write



TOP 500 The List.

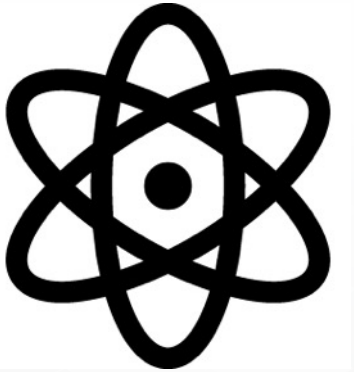
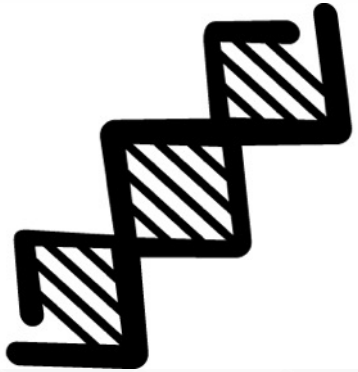
#1 JUNE 2018

https://openpowerfoundation.org/wp-content/uploads/2018/04/Jack-Wells.pdf

OpenPOWER™

Get Involved

Open Allows You To Create What You Need


The Foundation Workgroups Focused on a Range of Areas

- Hardware Specific
- Compliance
- Industries
- …and more

Accelerates Infrastructure Standards

13 OpenPOWER Work Groups	Collaboration	Work Product
Application and Domain Focused	**Unique needs for specific applications / Solutions**	Solution frameworks and optimization guidance
Integrated Solutions WG		
Personalized Medicine WG		
WG for Physical Sciences		
Machine Learning WG		
Interoperability and Inclusion	**Ensure ecosystems solutions work together**	
OpenPOWER Ready WG		Ready Definition and Criteria
Compliance WG		Compliance Specs
System Interfaces for Innovation	**Defining standards for developing and integrating innovative hardware subsystems**	
Memory WG		OpenPOWER Memory Bus Spec
Accelerator WG		PSL/AFU Interface Spec, CAPI SNAP
FRU Service Interface Spec WG		FSI Specification
Input/Output WG		Porting guide, testing guides, etc.
25GIO Interoperability Mode WG		25GIO Interoperability Spec
Fundamental System Architecture	**HW and SW Stack for System Architecture and Interoperability**	
System Software WG		ELFv2 ABI, LoPAPR
HW Architecture WG		ISA Profile, IODA, CAIA

SP010 – Tyan OpenPOWER Customer Reference System
CAPI – Coherent Accelerator Processor Interface
AFU – Accelerator Function Unit
FSI – Field Replaceable Unit (FRU) Service Interface
OPMB – OpenPOWER Memory Bus
ABI – Application Binary Interface
SDK – Software Developer Kit