# Forum TERATEC 23

## Unlock the future

**31 MAI & 1er JUIN 2023 • Au Parc Floral, Paris**

*Un événement organisé par* infoprodigital Teratec

# *Emergence of new technologies, new constraints and different deployment environments: an opportunity for modular and/or disaggregated systems*

June 1st, 2023 - Forum TERATEC

Xavier LE VAILLANT
xlv@2crsi.com

*Disclaimer: The views and opinions expressed in this presentation are those of the speaker and do not necessarily reflect the views or positions of any entities he represent.*

# Insatiable demand for computing power

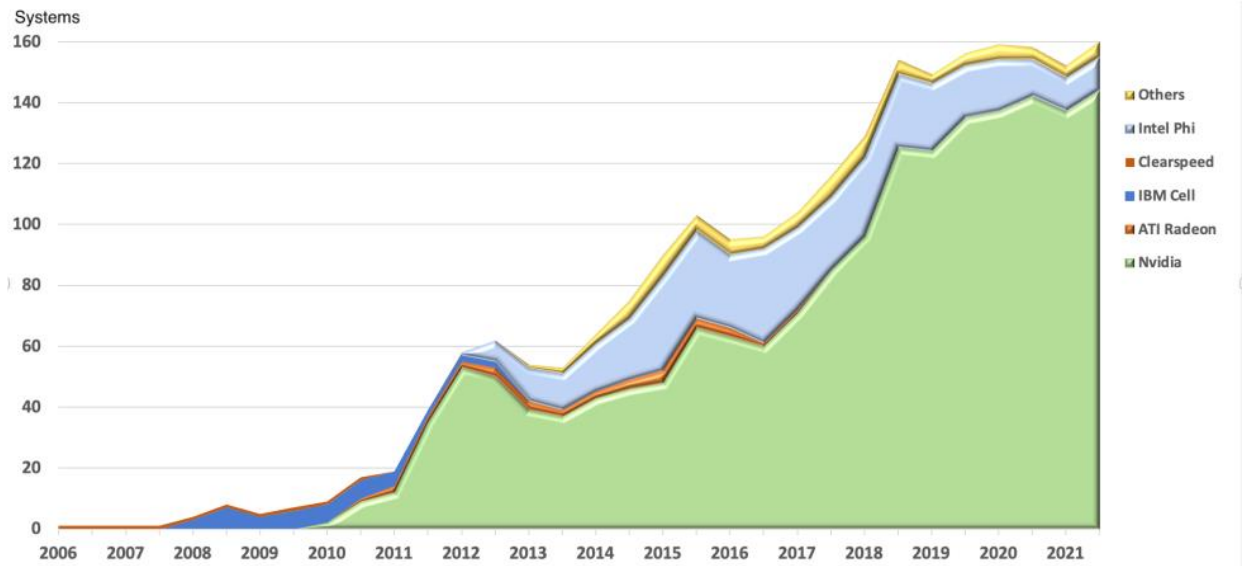
(Getty Images/iStockphoto)

*"Powering the services and products essentials to our daily lives"*

- SMARTER CLIENT DEVICES & EDGE
- AI & ANALYTICS EVERYWHERE
- 5G & COMMS INFRASTRUCTURE
- ADAPTABLE & INTELLIGENT SYSTEMS
- GAMING, SIMULATION & VISUALIZATION
- CLOUD, NETWORK, HYPERSCALE & SUPERCOMPUTER

# More computing power… with Accelerators

## Systems Using Accelerators on the TOP500

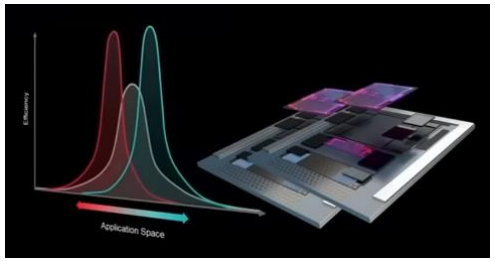



Reinventing High Performance Computing: Challenges and Opportunities, Daniel Reed & al., 2022

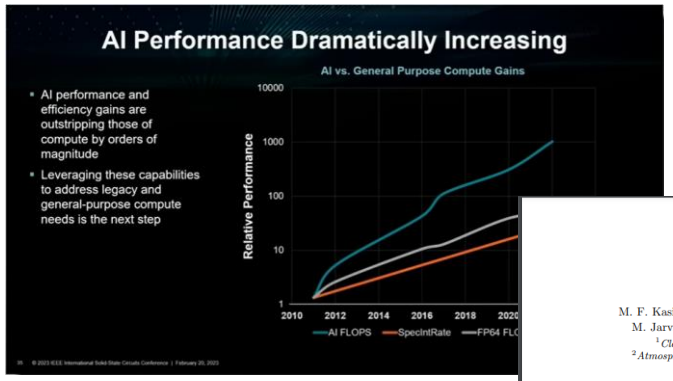
## Age of Accelerators

- GPUs, FPGAs, Vector engines, DPUs
- Accelerators needed to get to value
- Different accelerators process data differently

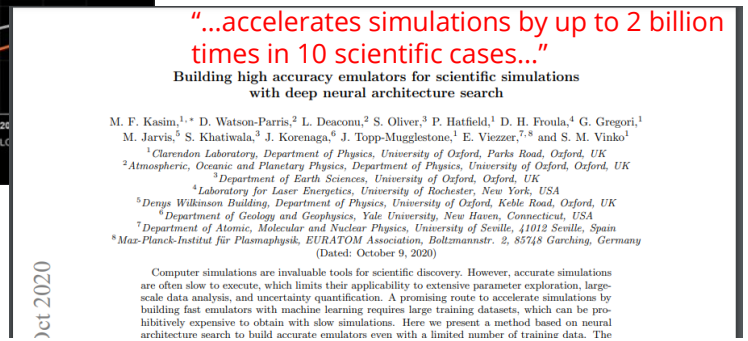

Domain-specific computation



# (AI-)Accelerated HPC




ISSCC 2023 Plenary - Lisa Su: Innovation For the Next Decade of Compute Efficiency



"...accelerates simulations by up to 2 billion times in 10 scientific cases..."

**Building high accuracy emulators for scientific simulations with deep neural architecture search**

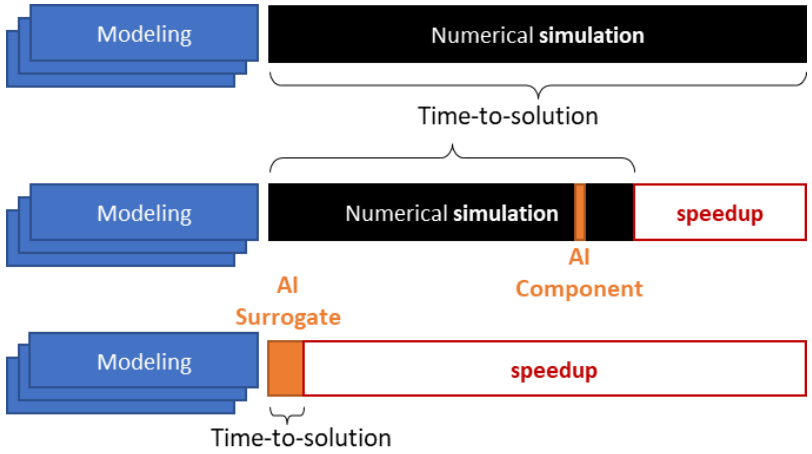
M. F. Kasim,[1,*] D. Watson-Parris,[2] L. Deaconu,[2] S. Oliver,[3] P. Hatfield,[1] D. H. Froula,[4] G. Gregori,[1] M. Jarvis,[5] S. Khatiwala,[3] J. Korenaga,[6] J. Topp-Mugglestone,[1] E. Viezzer,[7,8] and S. M. Vinko[1]

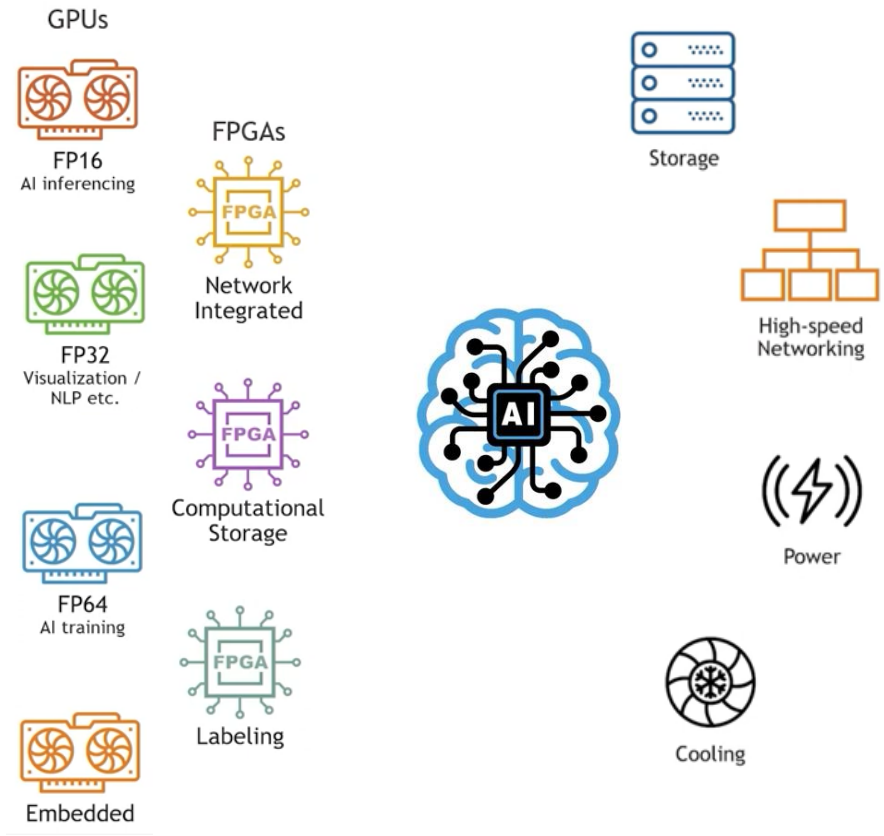
[1] *Clarendon Laboratory, Department of Physics, University of Oxford, Parks Road, Oxford, UK*
[2] *Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, UK*
[3] *Department of Earth Sciences, University of Oxford, Oxford, UK*
[4] *Laboratory for Laser Energetics, University of Rochester, New York, USA*
[5] *Denys Wilkinson Building, Department of Physics, University of Oxford, Keble Road, Oxford, UK*
[6] *Department of Geology and Geophysics, Yale University, New Haven, Connecticut, USA*
[7] *Department of Atomic, Molecular and Nuclear Physics, University of Seville, 41012 Seville, Spain*
[8] *Max-Planck-Institut für Plasmaphysik, EURATOM Association, Boltzmannstr. 2, 85748 Garching, Germany*

(Dated: October 9, 2020)

Computer simulations are invaluable tools for scientific discovery. However, accurate simulations are often slow to execute, which limits their applicability to extensive parameter exploration, large-scale data analysis, and uncertainty quantification. A promising route to accelerate simulations by building fast emulators with machine learning requires large training datasets, which can be prohibitively expensive to obtain with slow simulations. Here we present a method based on neural architecture search to build accurate emulators even with a limited number of training data. The

https://arxiv.org/abs/2001.08055




**New workloads are driving new levels of HPC complexity... and applying visible pressure to systems architecture and Datacenters**

# THE EUPILOT

Pilot using Independent Local & Open Technologies

## Paving the way for the first 100% EU Exascale supercomputer

- Extend open source to include open source hardware for HPC
- Software/hardware co-design for improved application performance and system energy efficiency.

- System integration innovations
- Stimulate European collaboration
- Combine industry standard methodology and cutting-edge research to accelerate exploitation

The European PILOT project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No.101034126. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Spain, Italy, Switzerland, Germany, France, Greece, Sweden, Croatia and Turkey.

www.eupilot.eu

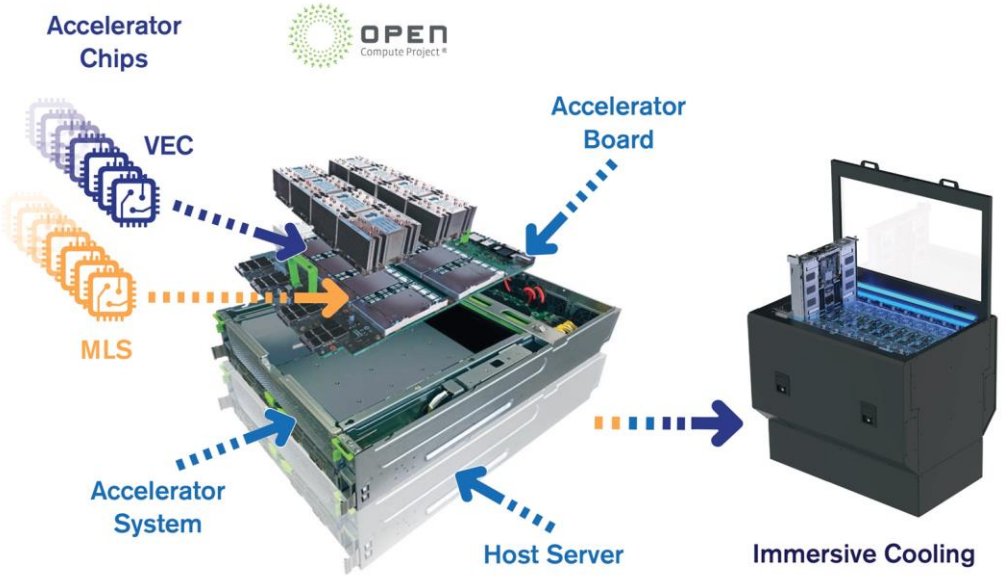
# THE EUPILOT

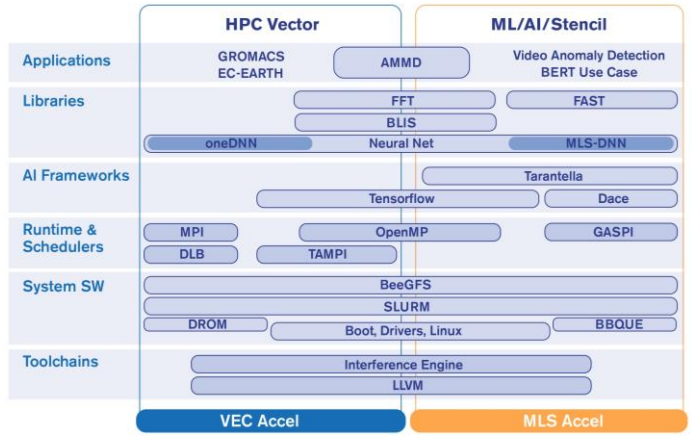
Pilot using Independent Local & Open Technologies

## Paving the way for the first 100% EU Exascale supercomputer

- Extend open source to include opensource hardware for HPC
- Software/hardware co-design for improved application performance and system energy efficiency.
- System integration innovations
- Stimulate European collaboration
- Combine industry standard methodology and cutting-edge research to accelerate exploitation

## Co-design & Open standards




The European PILOT project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No.101034126. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Spain, Italy, Switzerland, Germany, France, Greece, Sweden, Croatia and Turkey.

www.eupilot.eu

# Composable Disaggregated Infrastructure : What and why now?

**Composable Disaggregated Infrastructure (CDI) brings the agility, savings and efficient resource-sharing of the cloud to the management of on-premises equipment**.

Using orchestration and high-bandwidth, low-latency fabrics, shared resources can be combined on-demand for shifting workloads. The goal is to get the right ratio for a specific AI training or inference job, change configurations as the workload pipeline changes, and free up expensive GPUs and other accelerators for additional work.

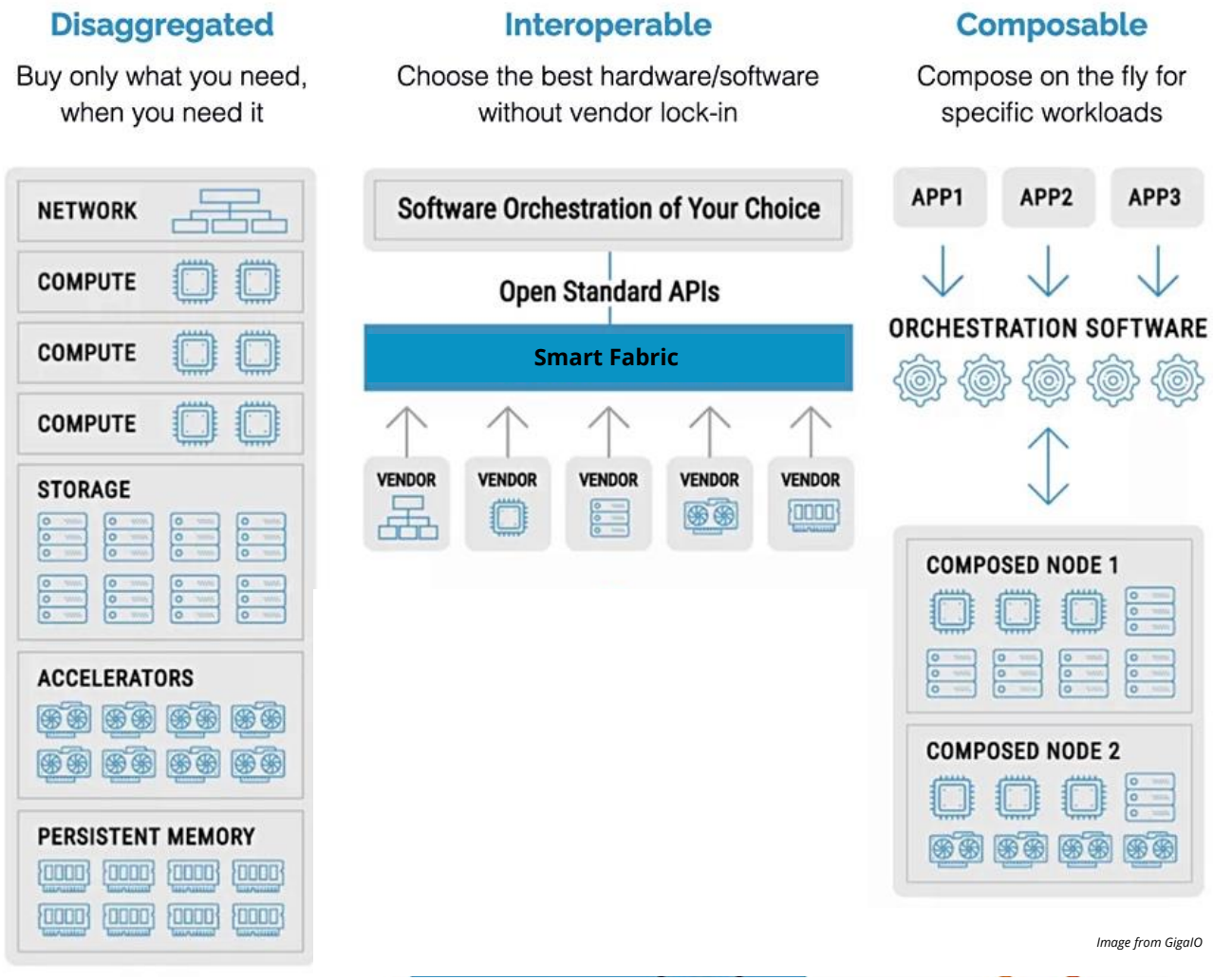
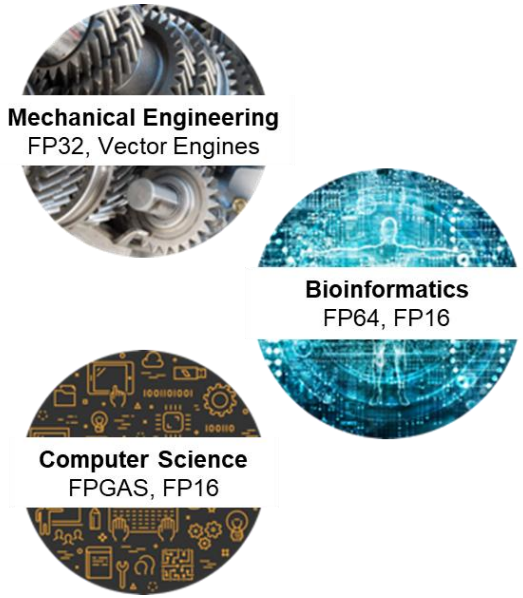


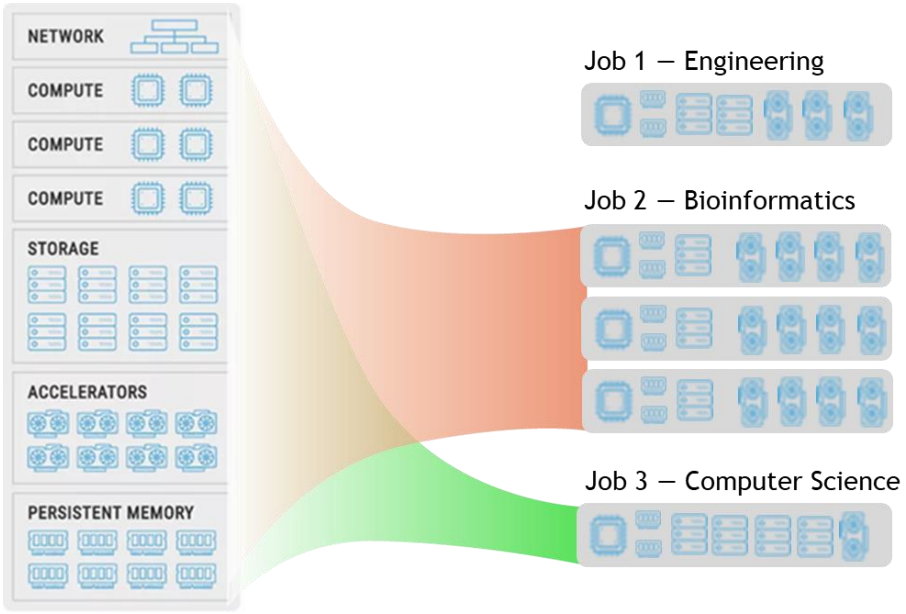

*Image from GigaIO*

# Composability: one-size-fits-all architectures are outdated

e.g., different research activities




Software defines hardware uniquely for each workload




**The End of Stranded Resources**
- The right resources in the right place, on demand

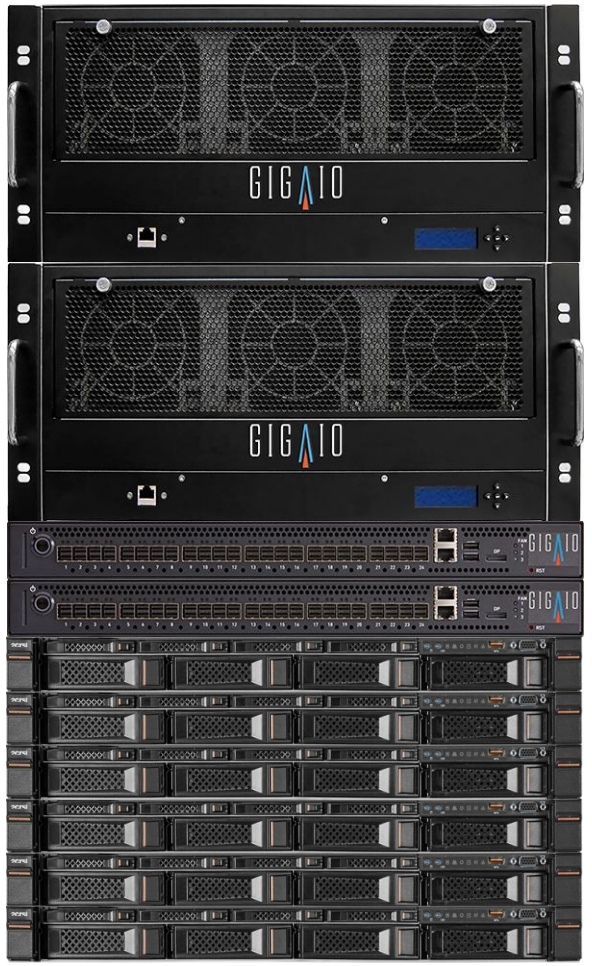
**Scale-Up and Scale-Out as You Grow**
- In-place scaling and selection of servers and accelerators as requirements evolve

**True Heterogeneity**
- The right GPUs, servers and accelerators for the job

# Composability: e.g., GigaPod™ - preconfigured for easy deployment

## A perfect entry point and easily expanded

- **1 GigaPod kit** — includes FabreX™ composable switches, network adapter cards, and cables
- **Up to 2 GigaIO Accelerator Pooling Appliances** — each up to 8 mix-and-match GPUs
- **Up to 6 servers** — *storage* to 300TB each or *application* server - dual AMD EPYC™ 7713, 75F3, or 7543 — each with up to 1TB memory and 128 cores
- **Bright Cluster Manager** — supporting High Availability (HA) storage and Bright for Data Science

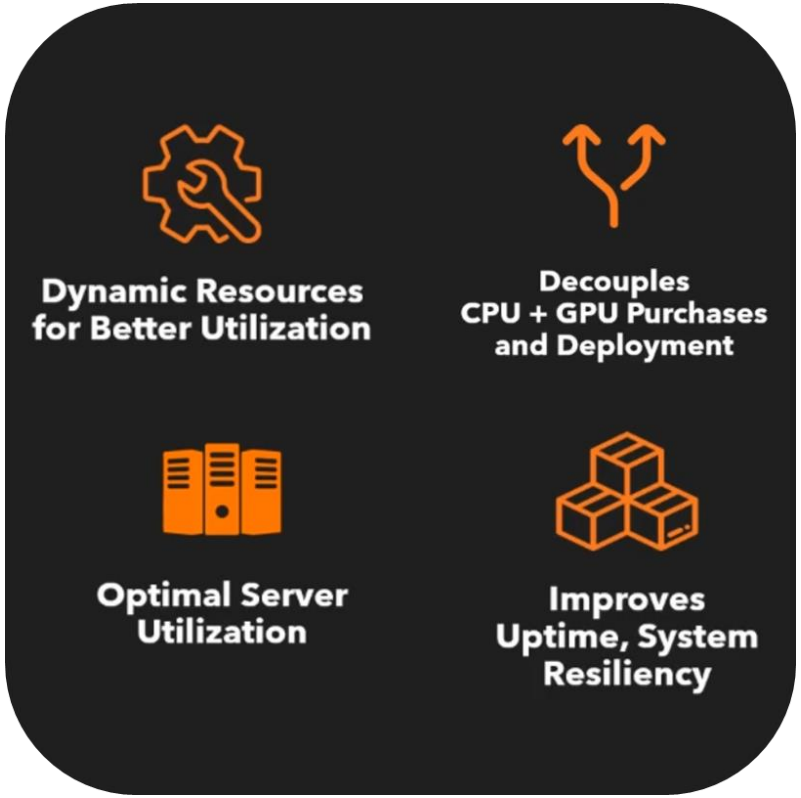



**Supports up to:**

| 640 CPU Cores | 4.6TB Memory | 307 TB Storage | 16 Accelerators |
|---|---|---|---|

# Composability: Benefits

- **Improved system utilization** by more fully leveraging expensive on-premises assets.
- **Flexible hardware profiles** – create the *Impossible server*
- **Pay as you Grow** - Simplified system expansion and reduced system  costs via modular resource-specific nodes
- **Reduced Power & Cooling** (Sustainability)
- **Better Managed Life Cycles**



# Composability: Challenges

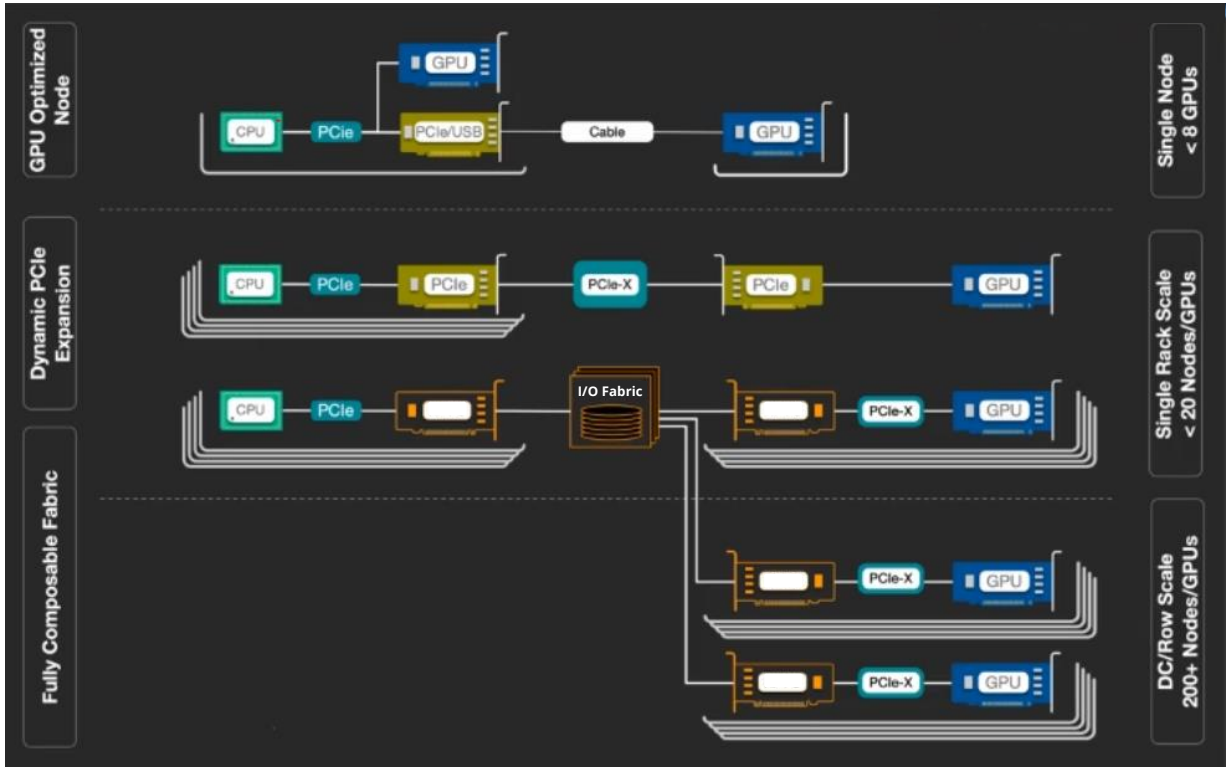
- **Usage and operational impacts**
  - Which workloads are most suitable for composability?
- **Resource impacts**
  - Changes (if any) in application code  to support composability?
  - Will it increase or reduce support requirements?
- **Performance impacts**
  - What about the latency to manage, provision, monitor, and re-claim system resources between jobs?
  - Will increased physical distance also add latency?
  - Scaling? How far?
- **Cost impacts**
  - Additional network (MPI, I/O, and now PCIe…)?

# Composability: I/O Fabric (one of) the biggest challenge

**Traditional System Model**
- Inefficient architecture
- Low ratio of GPUs to CPUs
- Stranded resources

**PCIe Expansion**
- Single rack scale
- Low radix, non-resililant
- Large blast radius

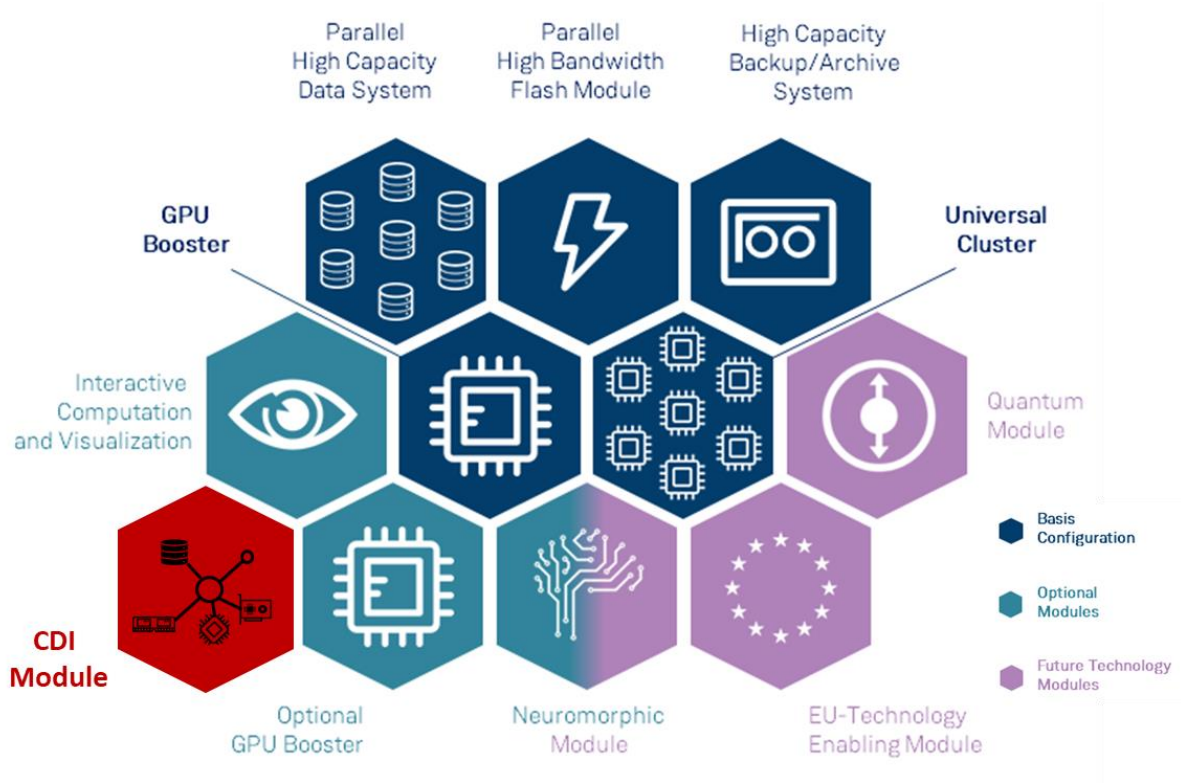
**Fully Composable Fabric**
- Data center scale
- Highly resilient
- No single point of failure
- Very high radix
- Fully dynamic path selection
- Accelerates CLX memory scalability

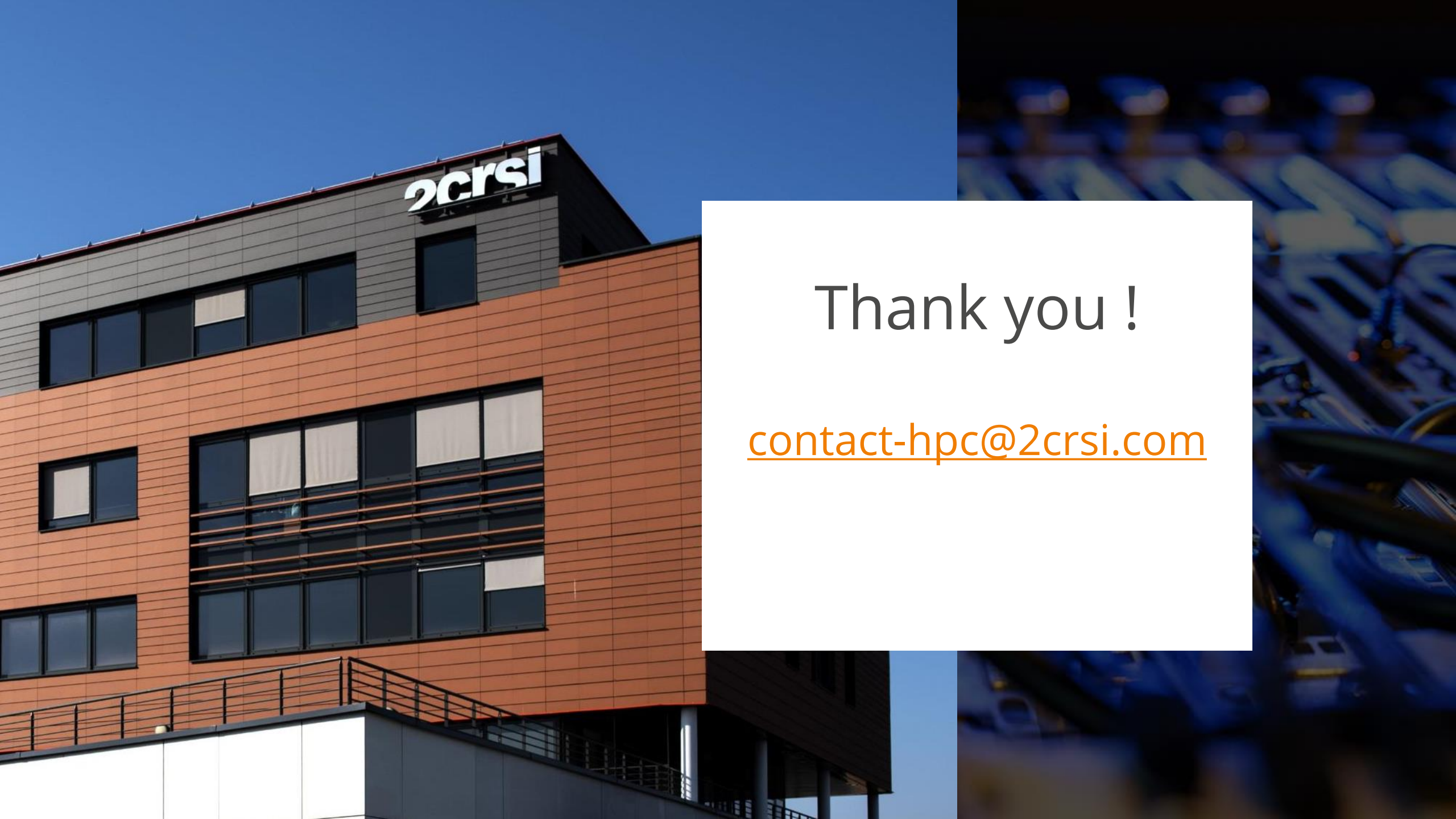


# Closing thoughts - Heterogenous system & workloads

- Complex, heterogeneous modern workloads will continue to stress existing system architectures
- Increasing interdependence between complexities of new workloads (e.g., AI, Quantum), access to resources at scale, and user demands for accelerating time to results
- CDI is becoming a way to solve these problems

Proposal of a *revised* first EU Exascale system…






# Thank you !

contact-hpc@2crsi.com