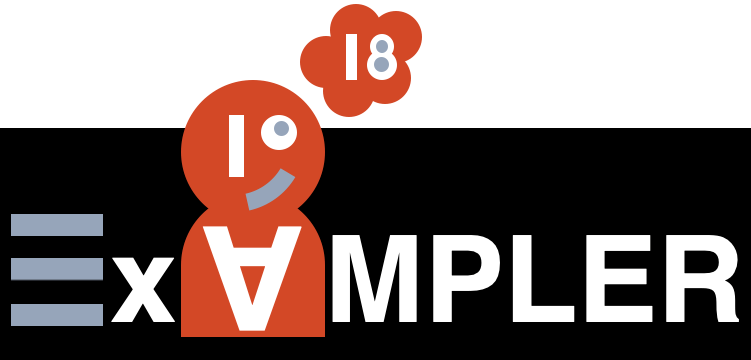
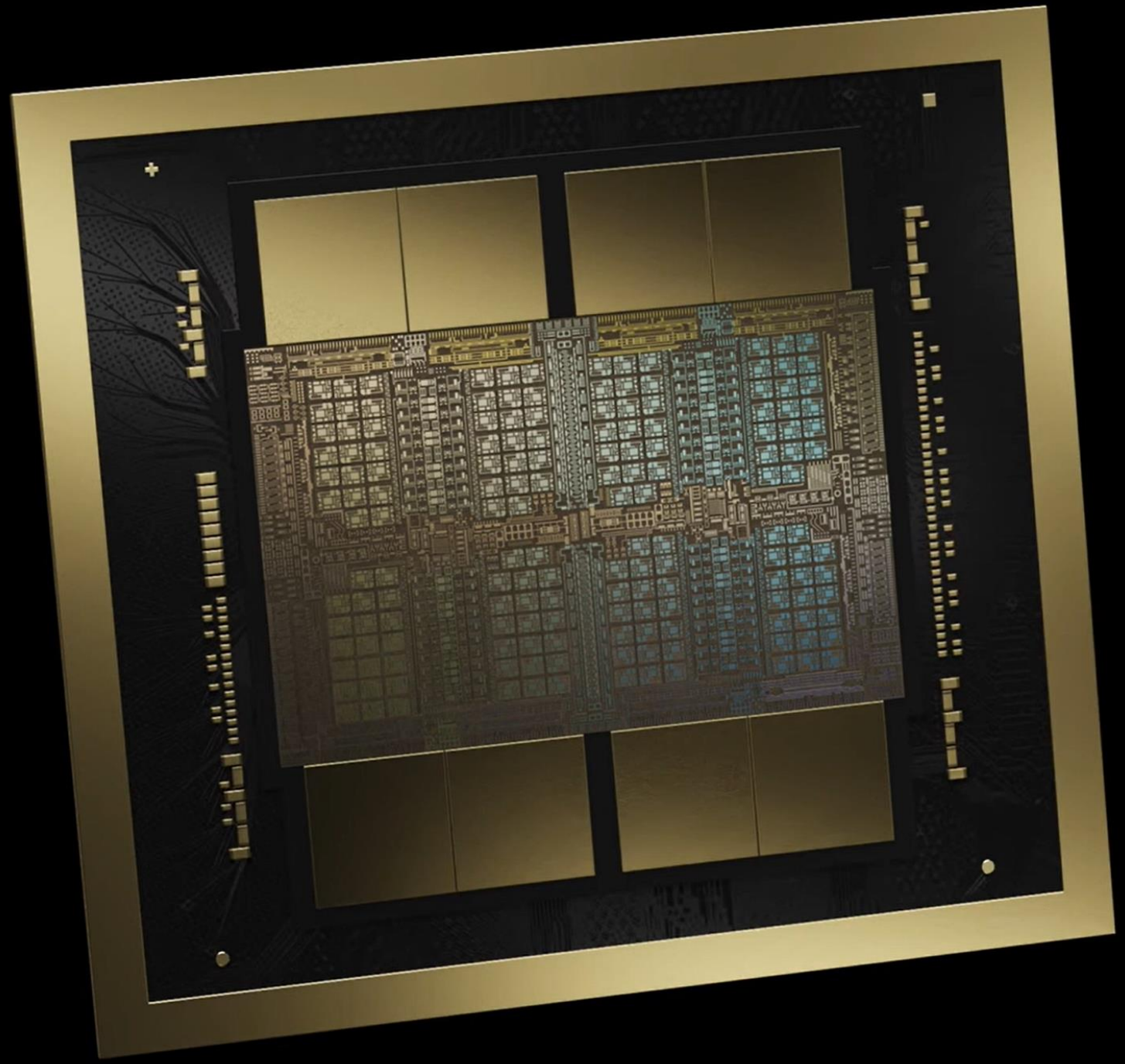
Identifying the Crucial Challenges for ABSS at Exascale

Dr. Richard Milton and Prof. Michael Batty





BLACKWELL

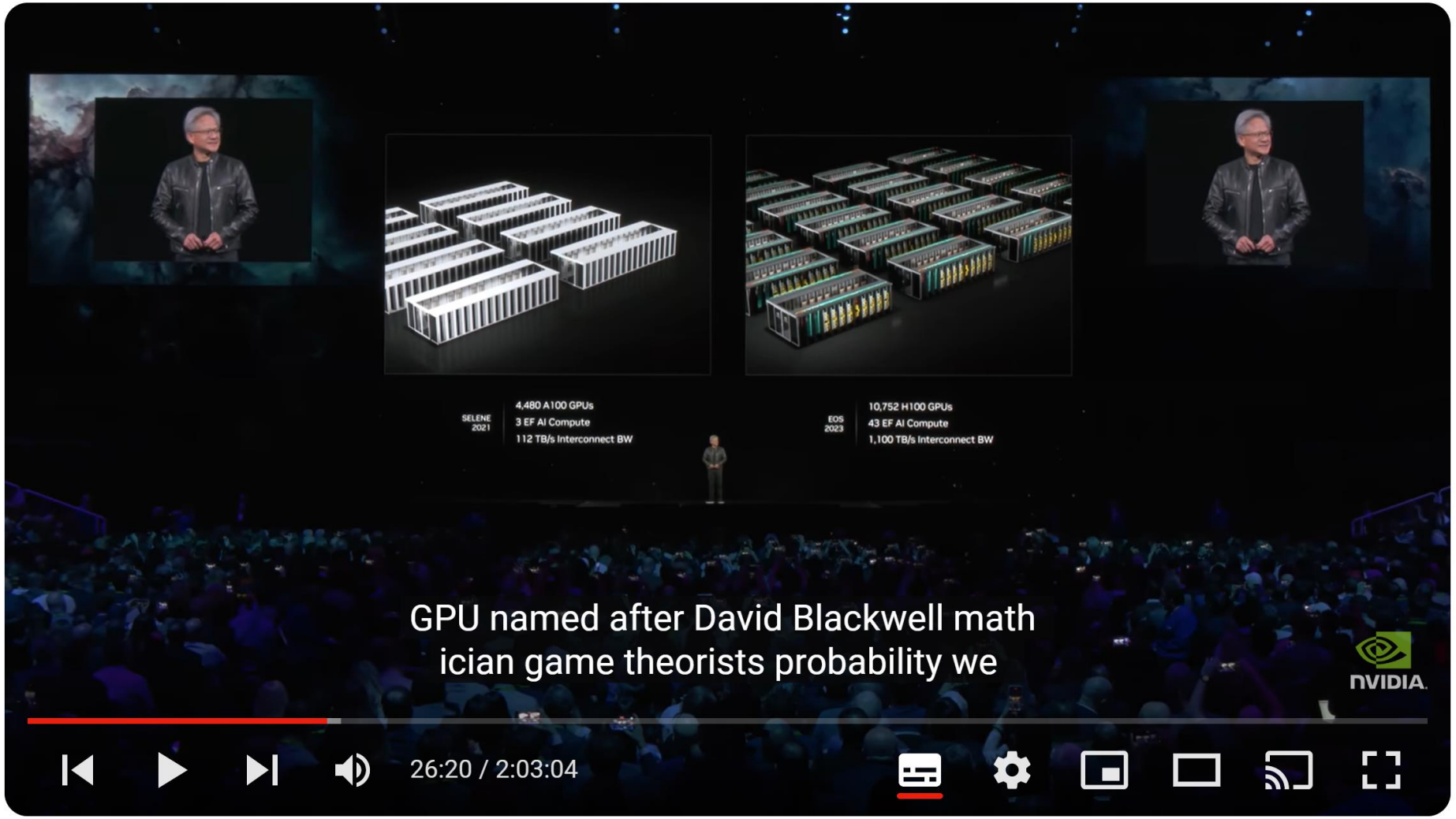
THE ENGINE OF THE NEW INDUSTRIAL REVOLUTION

20 petaFLOPS of AI performance

192GB of HBM3e

8TB/s of memory bandwidth

Full stack, CUDA enabled



GTC March 2024 Keynote with NVIDIA CEO Jensen Huang

https://www.youtube.com/live/Y2F8yisiS6E?si=M5tp8hCGaWwYL5Bs&t=1564 (26m48s to 29m12s)

FULL DATA CENTER WITH 32,000 GPUs

AI FACTORY FOR THE NEW INDUSTRIAL REVOLUTION

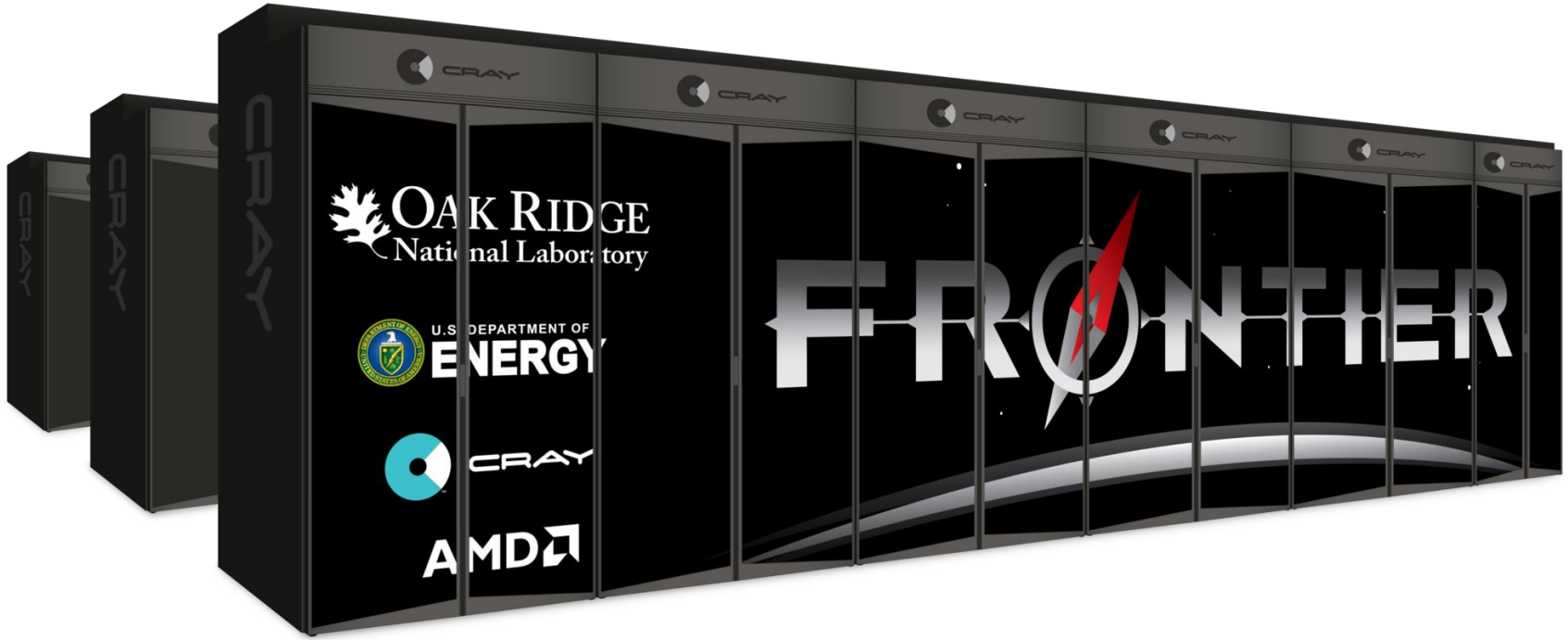
645 exaFLOPS of AI performance

13PB of fast memory

58PB/s of aggregate NVLink bandwidth

16.4 petaFLOPS of In-Network Computing

FRONTIER, Oak Ridge, Tennessee, US
~£476M, Operational May 2022

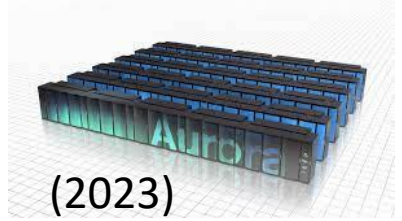


archer2 Archer 2 is Frontier without the accelerators, x61

And the others…

United Kingdom
(2025?)

United States



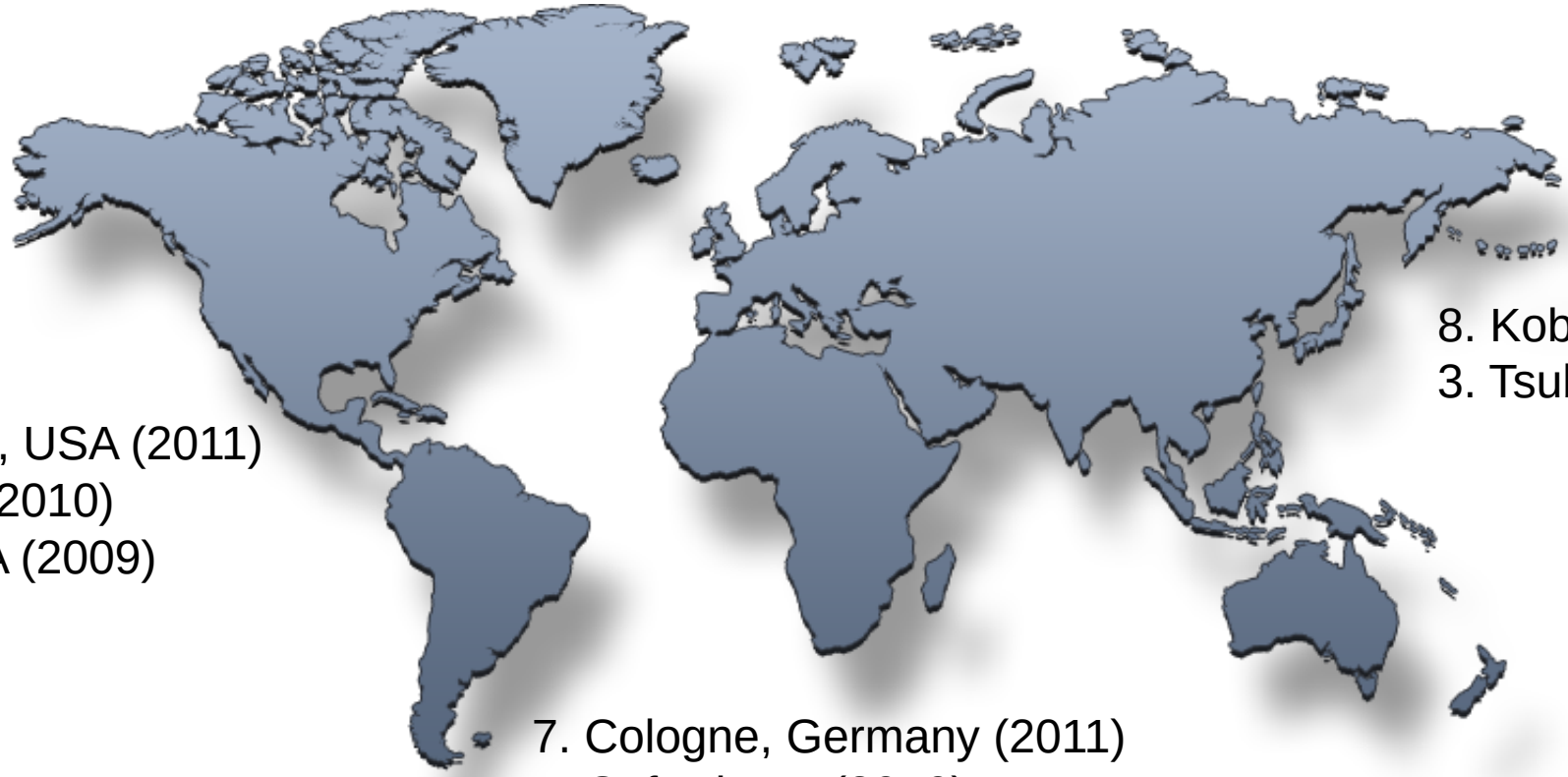
(2023)

(2024)

Europe
Jupiter (2024)

China
Sunway Oceanlite
Tianhe-3

IESP Roadmap Meetings



6. San Francisco, USA (2011)
5. Maui, Hawaii (2010)
1. Santa Fe, USA (2009)

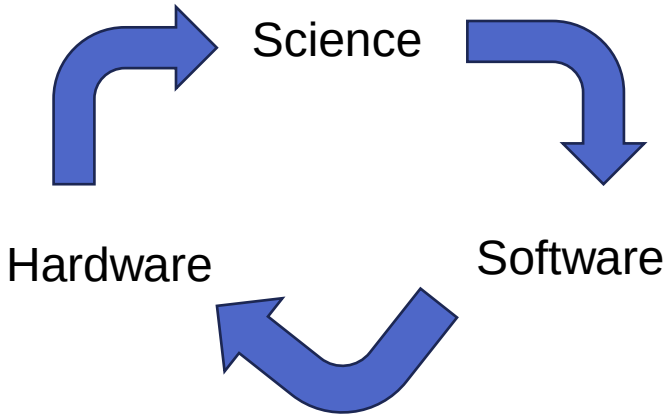
8. Kobe, Japan (2012)
3. Tsubuka, Japan (2009)

7. Cologne, Germany (2011)
4. Oxford, UK (2010)
2. Paris, France (2009)

https://exascale.org/iesp/IESP:Documents.html

CoDesign Vehicles and Proxy Apps

https://proxyapps.exascaleproject.org/app/



A "proxy app" differs from a benchmark in that it is designed to be a simplified example of a real world algorithm

"At the IESP's recent meeting in San Francisco, the ideas of "hardware/software co-design" and the "co-design process" took center stage as organizing concepts for the community's exascale effort."
Jack Dongarra et. al, April 2011 IESP SF Summary Report

mlperf-cosmoflow | Python, TensorFlow

The CosmoFlow training application benchmark from the MLPerf HPC v0.5 benchmark suite. It involves training a 3D convolutional neural network on N-body cosmology simulation data to predict physical parameters of the universe.

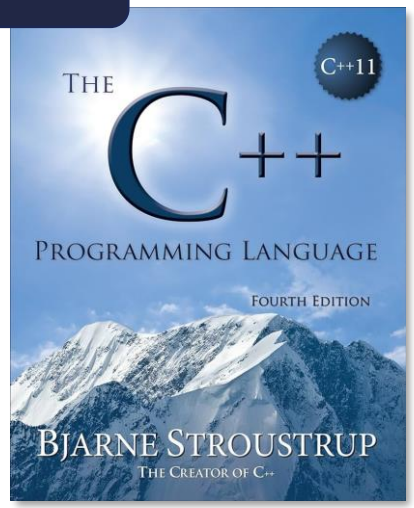
miniSMAC2D | Fortran

Solves the finite-differenced 2D incompressible Navier-Stokes equations with Spalart-Allmaras one-equation turbulence model on a structured body conforming grid

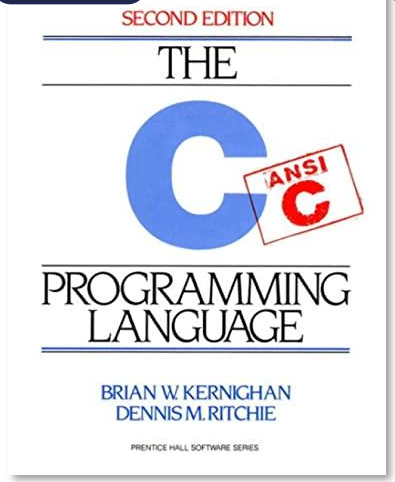
What Languages Achieve an Exaflop?

Summary of ECP Proxy Apps https://proxyapps.exascaleproject.org/

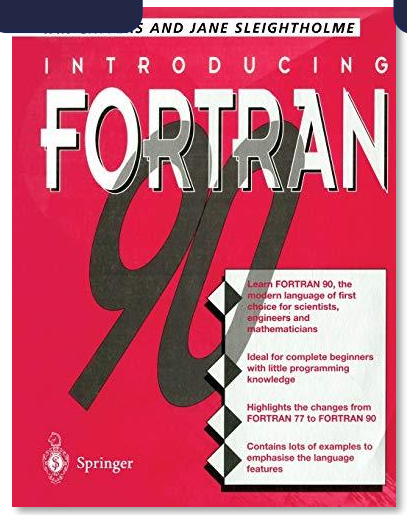
44%



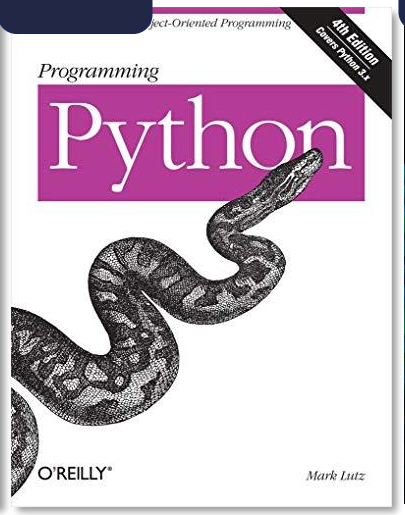
24%



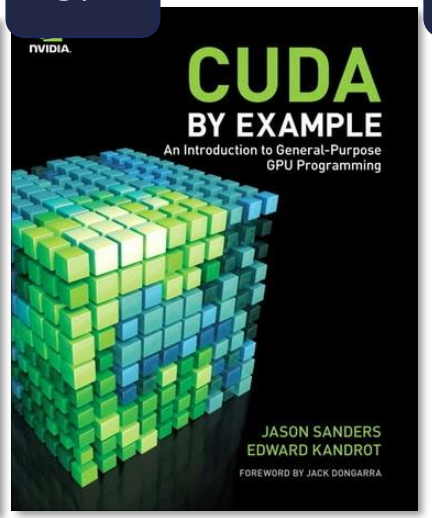
17%



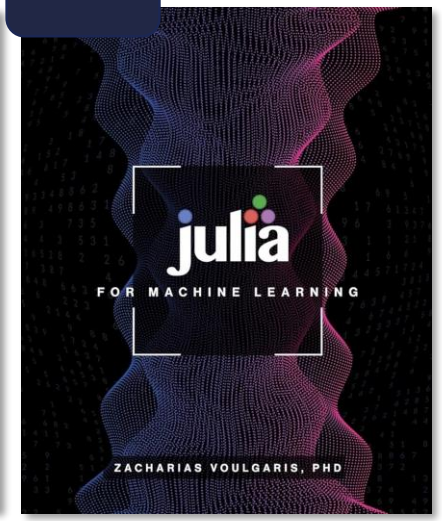
10%



5%



1%



Bernd Mohr: All 25 sites had C, C++ and Fortran. Python 84%, Java 72%

https://exascale.org/mediawiki/images/e/ef/Talk02-Bernd.pdf

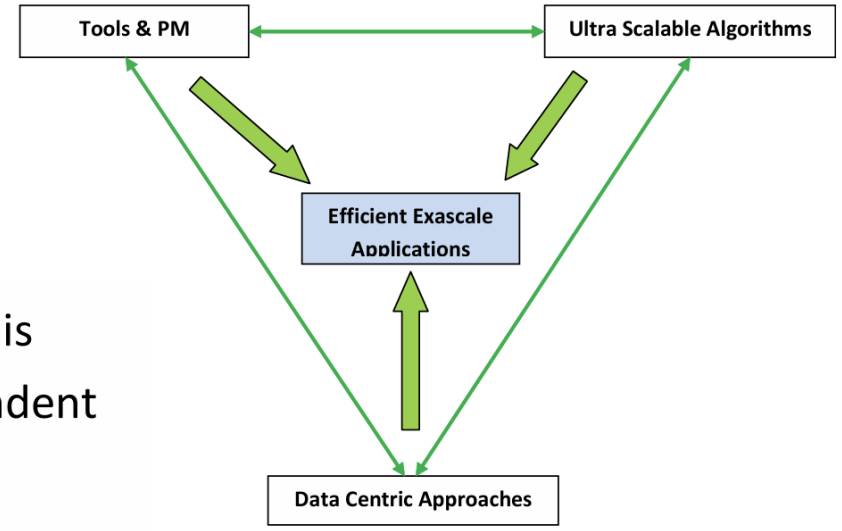
Survey of System Software Stacks in the IESP Community, Bernd Mohr, Jülich Supercomputing Centre (IESP San Francisco)

…and in Europe, the European Exascale Software Initiative (ESSI)

- http://www.eesi-project.eu/vision/roadmap-foundation/
- ESSI2 final meeting and roadmap in 2015 – Three Pillars

The roadmap towards the implementation of efficient Exascale applications and the consecutive recommendations are gathered in three large pillars:

- Tools & Programming Models
- Ultra Scalable Algorithms
- Data Centric Approaches



Note that the Data Centric vision is very new in Europe but is essential for approaching the ultra complex and interdependent challenges of Extreme Computing and Extreme Data.

As already advised by EESI2 (and by US DOE), Exascale requires a new and different approach compared to classical HPC. There is an urgent need for specific and disruptive R&D programs targeting Exascale software.

The UK Situation

Delivery date? 2025?

"Edinburgh has been selected to host a next-gen supercomputer fuelling economic growth, building on the success of a Bristol-based AI supercomputer, creating high-skilled jobs."
– *UK Government Press Statement*

Isambard 3 (Bristol), ARM
NVIDIA Grace, 55,000 cores

- ❑ ARCHER2, Tier-1 national supercomputer service
- ❑ CDS3 Tier-2 HPC service
- ❑ Kelvin-2 (NI-HPC) Tier-2 HPC service
- ❑ Cirrus Tier-2 HPC service
- ❑ Bede (NICE) Tier-2 HPC service
- ❑ Baskerville Tier-2 HPC service
- ❑ Sulis Tier-2 HPC service
- ❑ Isambard
- ❑ JADE
- ❑ MMM Hub

EXTREME SCALE COMPUTING

Revolution or Evolution?

What is in the “X Stack” (for ABSS) ?

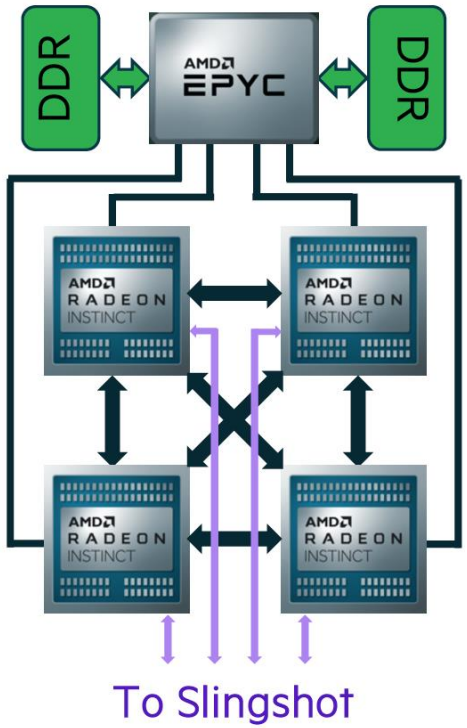
Technology

High numbers of agents

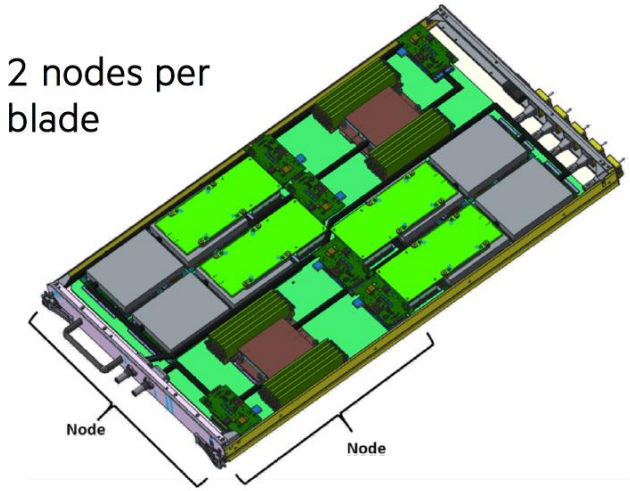
High communication between agents

High compute per agent

If you're not using the accelerators, then you're only doing petascale



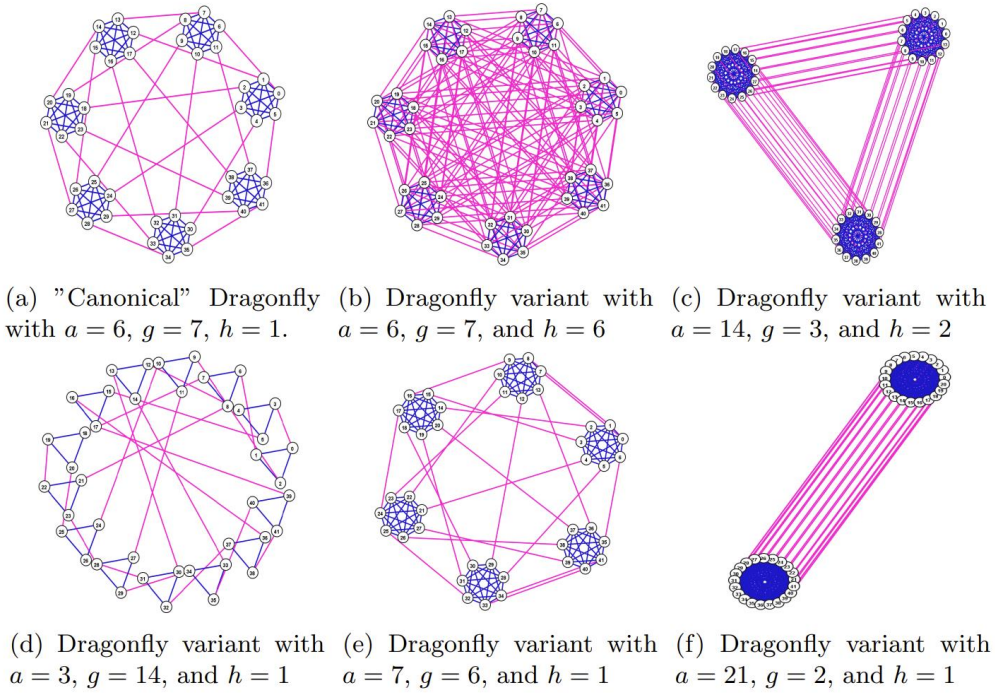
AMD GPU (ORNL)



9,472 nodes, each with 1 AMD EPYC 64 core CPUs and 4 Radeon Instinct MI250X GPUs
606,208 CPU cores and 8,335,360 GPU cores

- What is the reference architecture or step up?
- One big simulation, or massive ensembles?
- Simulation waits, time steps
- Dependent interacting agents

Connections



(a) "Canonical" Dragonfly with $a = 6$, $g = 7$, $h = 1$.

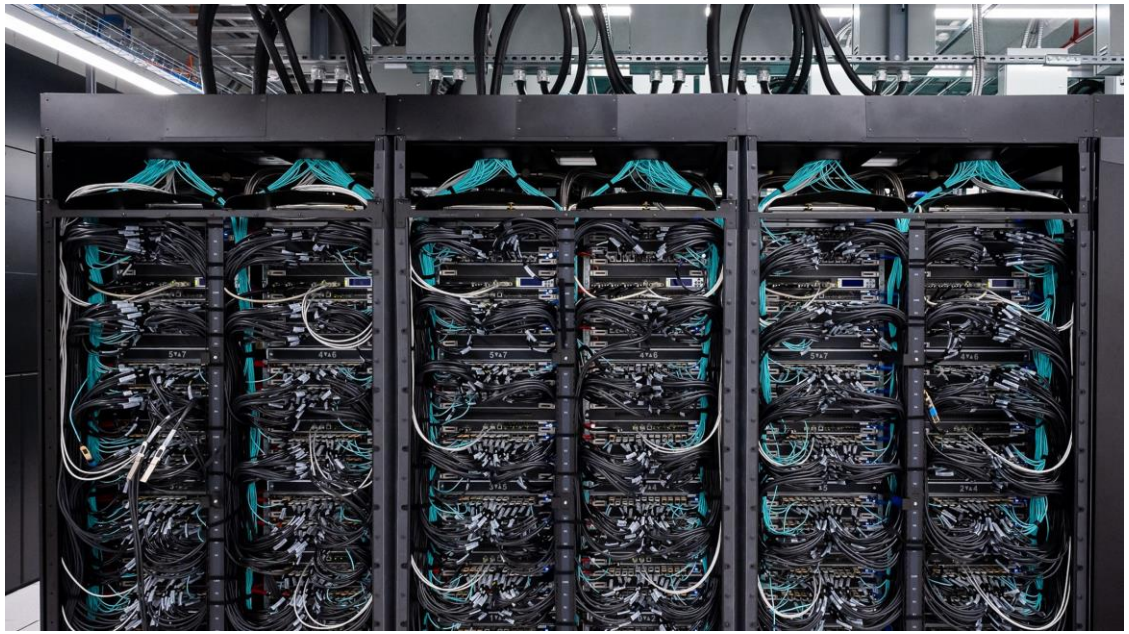
(b) Dragonfly variant with $a = 6$, $g = 7$, and $h = 6$

(c) Dragonfly variant with $a = 14$, $g = 3$, and $h = 2$

(d) Dragonfly variant with $a = 3$, $g = 14$, and $h = 1$

(e) Dragonfly variant with $a = 7$, $g = 6$, and $h = 1$

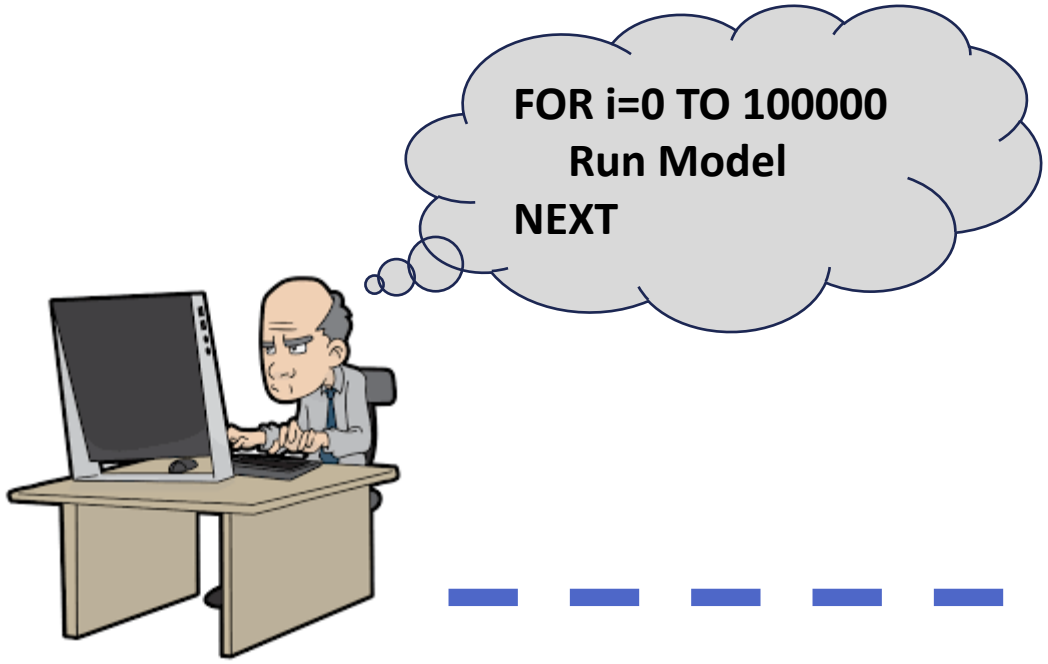
(f) Dragonfly variant with $a = 21$, $g = 2$, and $h = 1$



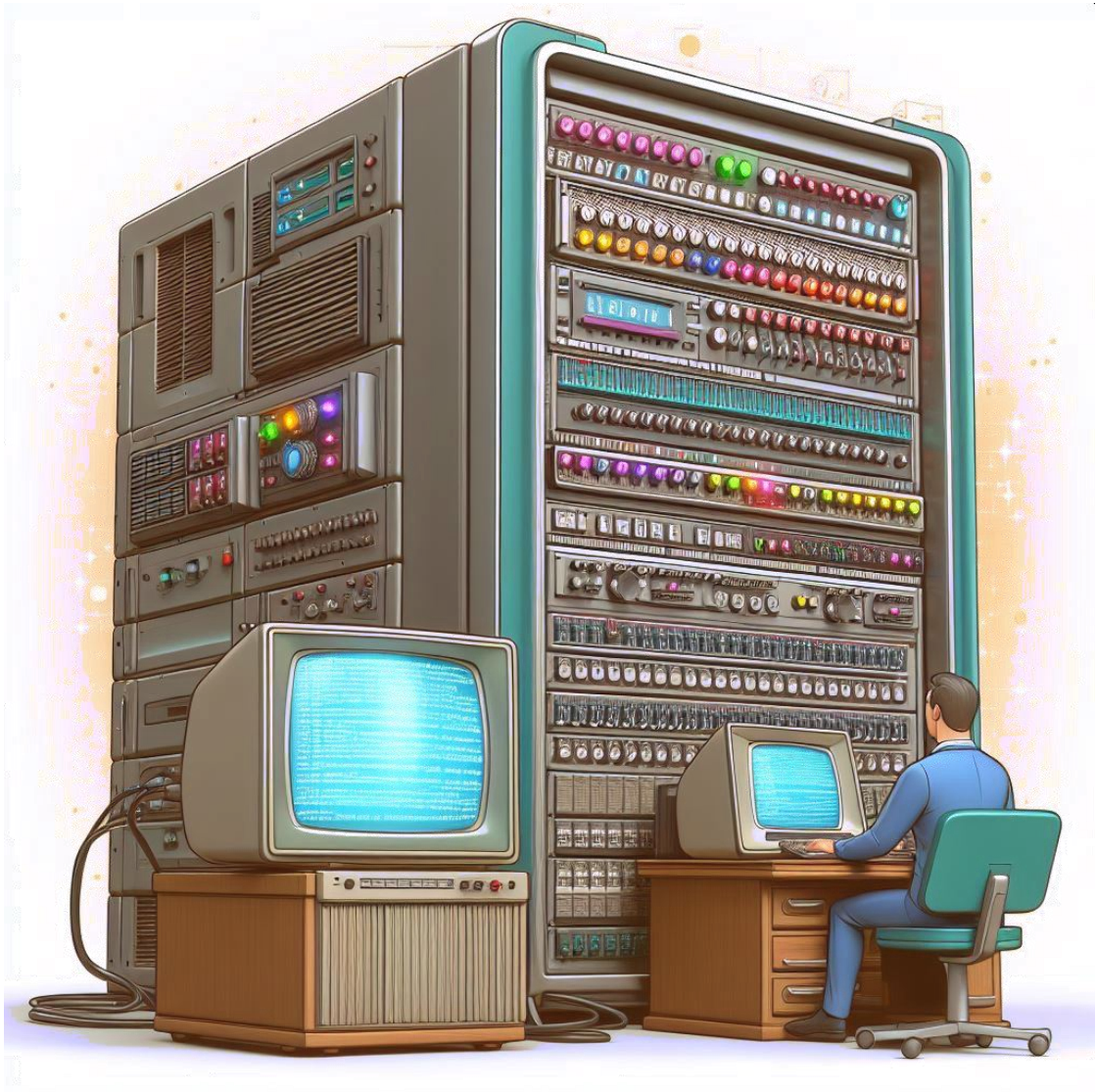
a=number of routers in group, g=number of groups in system, h=number of channels within each router used to connect to other groups.

Policy Testing

- Cost of running an exascale computer
- Emulator models



45 Watt Computer



45 Megawatt Computer

Bounded Worlds

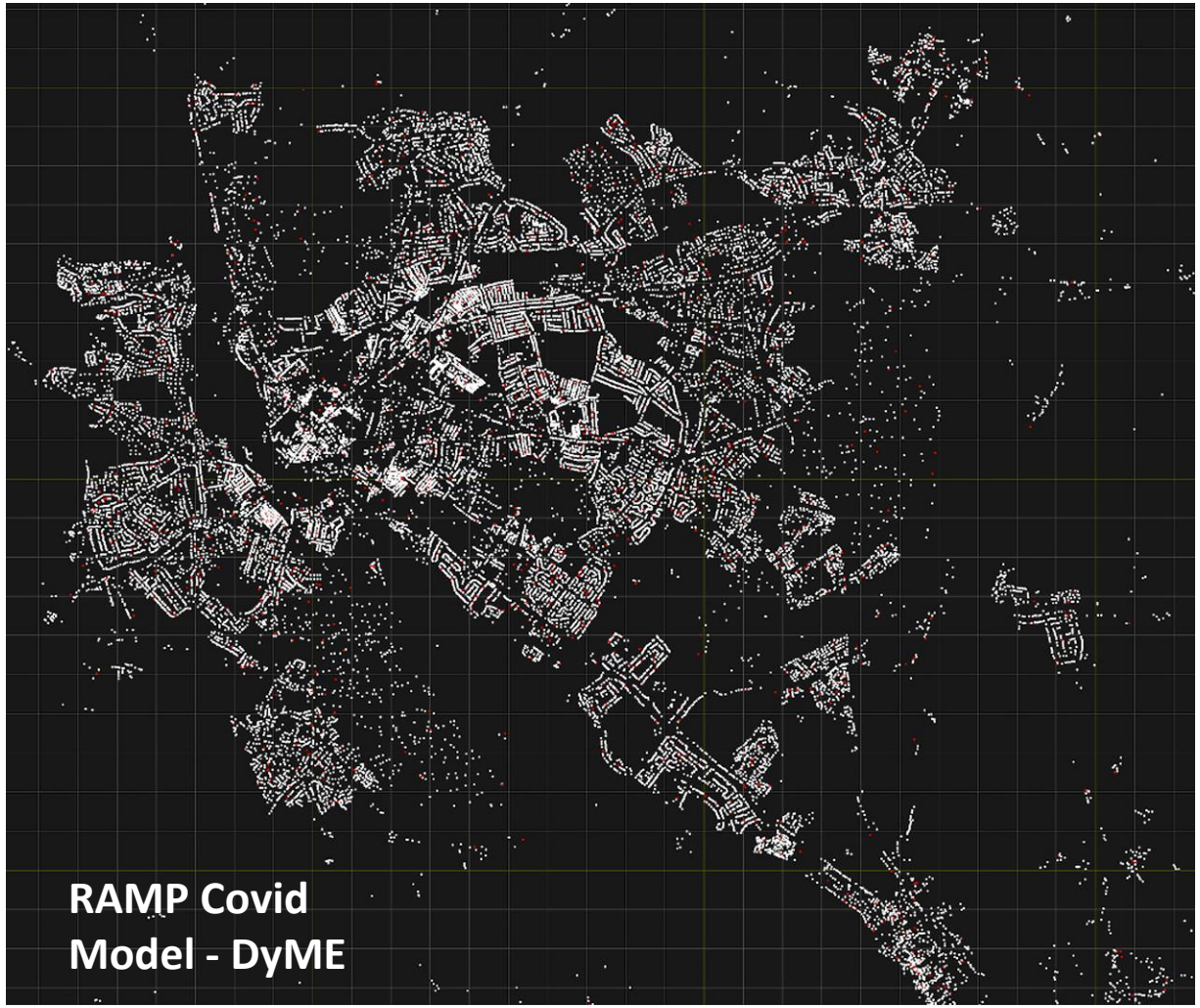
"Agent Based Architecture, A Quantitative Approach"

For example, in a COVID 19 simulation of Devon, there are 796,012 people.

1 exaflop/796,012 = 1,256,262,468,405 FP operations per second per person. That's 1 teraflop each.

Frontier has 9472 nodes, each with 4 GPU chips, giving 37,888 GPUs.

That's 21 people per GPU.



The Science Case

what are other people doing with $900M supercomputers?

The ACM Gordon Bell Prize: https://awards.acm.org/bell

2023	"Large-Scale Materials Modeling at Quantum Accuracy: *Ab Initio* Simulations of Quasicrystals and Interacting Extended Defects in Metallic Alloys." (Frontier, Crusher, Perlmutter, NOTE: 659.7PF)
2022	"Pushing the Frontier in the Design of Laser-Based Electron Accelerators With Groundbreaking Mesh-Refined Particle-In-Cell Simulations on Exascale-Class Supercomputers." (Frontier, Fugaku, Summit)
2021	"Closing the "Quantum Supremacy" Gap: Achieving Real-Time Simulation of a Random Quantum Circuit Using a New Sunway Supercomputer." (Sunway)
2020	"Pushing the limit of molecular dynamics with *ab initio* accuracy to 100 million atoms with machine learning." (Summit)
2019	"A Data-Centric Approach to Extreme-Scale Ab initio Dissipative Quantum Transport Simulations." (Summit)
2018	"**Attacking the Opioid Epidemic**: Determining the Epistatic and Pleiotropic Genetic Architectures for Chronic Pain and Opioid Addiction" and "Exascale Deep Learning for Climate Analytics" (ORNL?)

Agenda, Monday

11.00 – Welcome and Introduction to the ExAMPLER project

11.30 – Talks I: Experts' experiences in high-scale computing for Agent Based Social Simulation

12.30 – *Lunch*

13.30 – Talks II: Experts' experiences in high-scale computing for Agent Based Social Simulation

14.30 – *Coffee, networking*

14.45 – Workshops, discussion on Exascale ABSS.

Determine topics for the group work on the following day.

16.45 – Wrap Up (finish 17.00)

18.15 – Workshop Dinner at Grosvenor Hotel

Agenda, Tuesday

09.00 – Arrival and coffee

09.30 – Group Work

How to get from here to there – migration of tools, development of new ones, processes for ABSS modelling?

Evolution of existing methodology into exascale ABM. Revolutionary ideas for changing how ABM is achieved at exascale. What are the science drivers for exascale ABM?

11.30 – *Coffee*

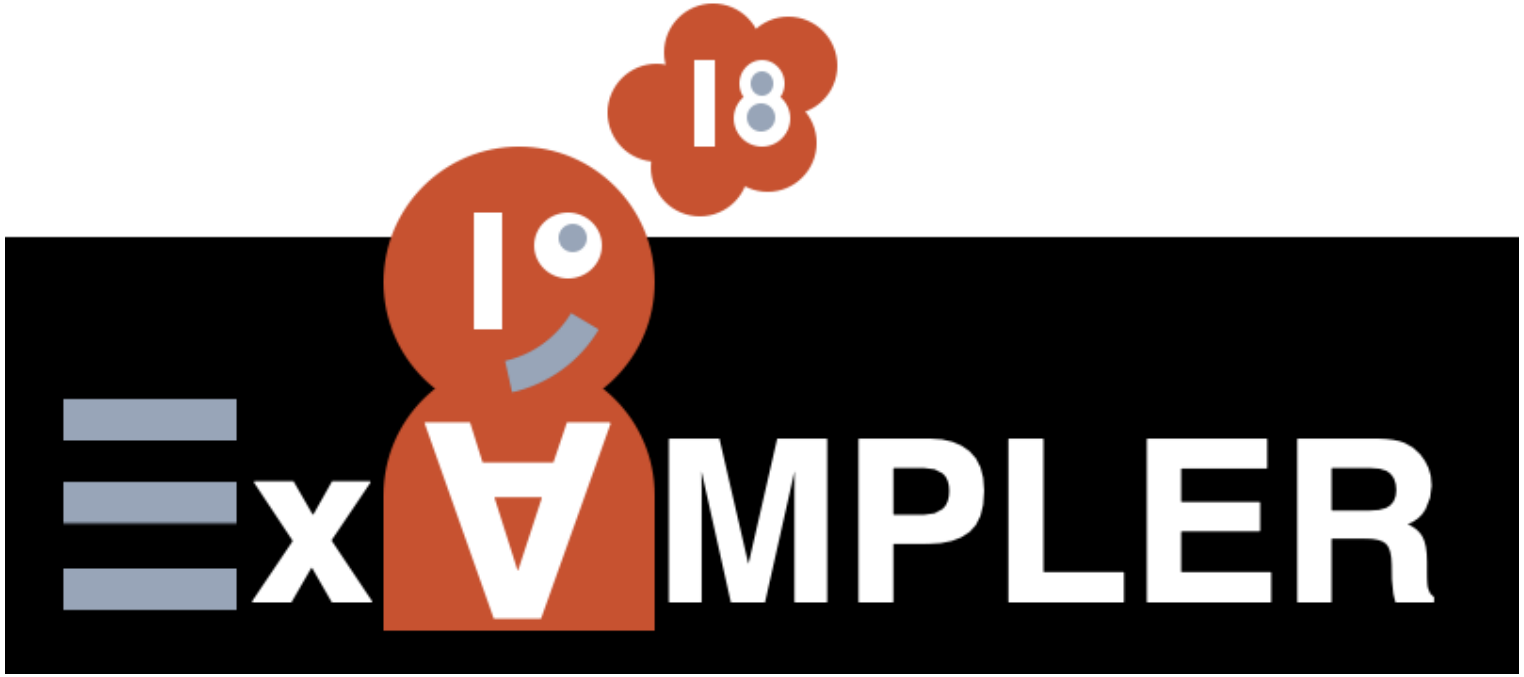
12.00 – Presentations of group work findings

13.00 – *Lunch*

14.00 – Plenary discussion and *Next steps.*

15.00 – *Finish*

Any Questions?



1,000,000,000,000,000,000 FP per sec