Sustainability in Exascale computing

Crispin Keable

December 2022

1

Atos HPC, AI & Quantum
Our Vision

Help our client achieve **breakthrough science or business benefits** through HPC, AI and Quantum simulation

Design innovative, energy efficient on-premise and cloud based **High Performance Computing** (**HPC**) solutions

Atos decarbonization ambition
To reach "net-zero"

Give access to **Artificial Intelligence** (**AI**) through innovative solutions such as Atos Trusted SuperPODs

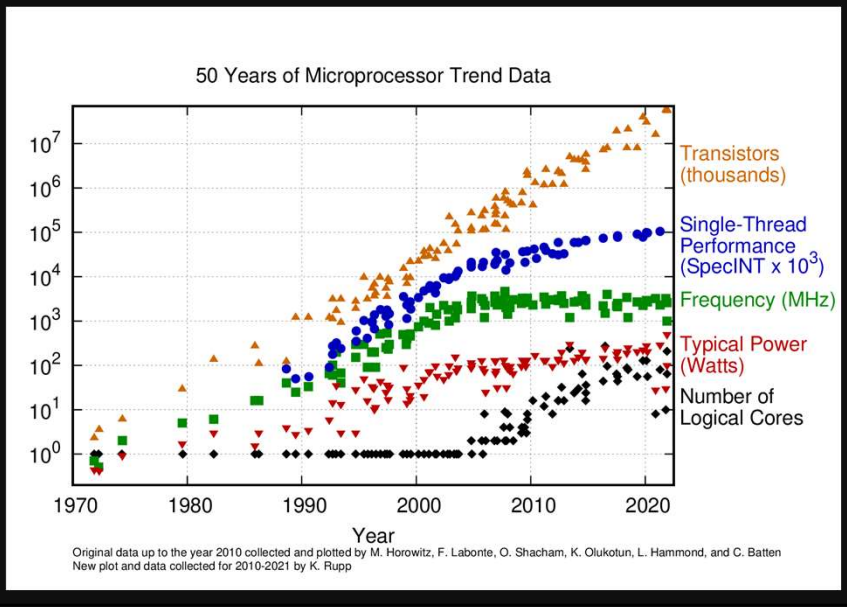
Step into the future with the **Quantum Computing** (**QC**) universal gateway

2

2

CPUs and GPUs are getting more powerful

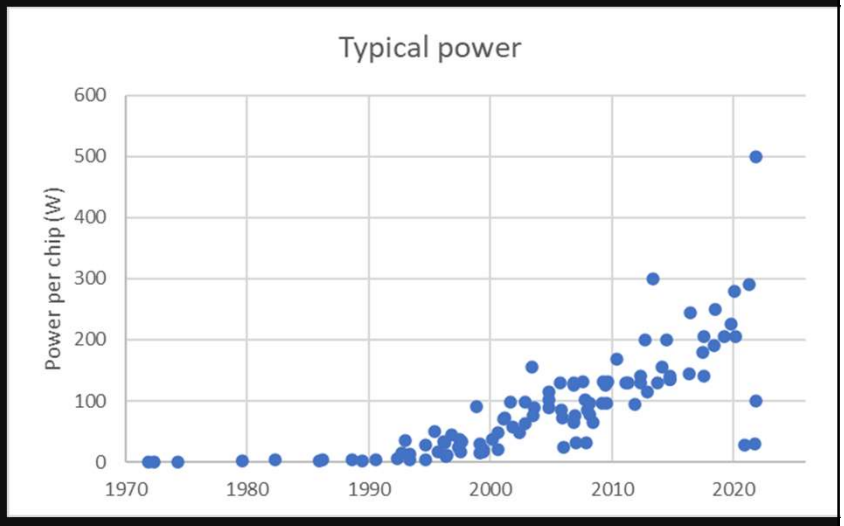
- Log graph shows the continuing Moores law climb
- Log graph
  - Great for showing exponential growth
  - Not so great for upward drift of CPU power used



3

Processor power is growing

- Switching to linear plot for power only, the growth in CPU and GPU power use is clear
- Narrowing the time window makes it even more clear
- Adding data for upcoming processors starts to look worrying!



4

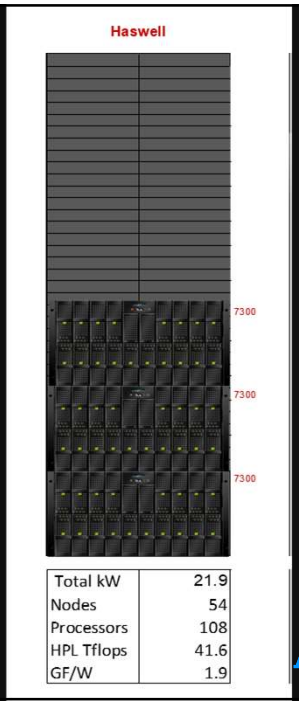
What can you put in an air cooled rack?

- Thinking back, HPC has been air cooled for decades
- Looking at typical HPC technologies in the last years, we consider
  - Haswell, with a typical rack configuration from 2014
  - Haswell, with water cooled doors
  - SkyLake, typical HPC rack (2018)
  - NVIDIA DGX1 – Ampere (2021)
  - NVIDIA DGX1 – Hopper (2023)
- Assume a DC rack limit of 20kW per rack to air
- GF per rack goes up, as does the GF/W
- However, rack density goes way down – ballooning DC costs

5

2014, Haswell ac

- Standard 19" rack, with high density 7U chassis
- Air cooled
- 18 dual socket nodes per chassis (Bull B500)
- Rack is 50% full



Total kW	21.9
Nodes	54
Processors	108
HPL Tflops	41.6
GF/W	1.9

6

2014 – Haswell, WCD

- 19" rack, with water cooled door
- Same Haswell compute chassis/nodes
- Bull B500 again

- Rack is 83% full

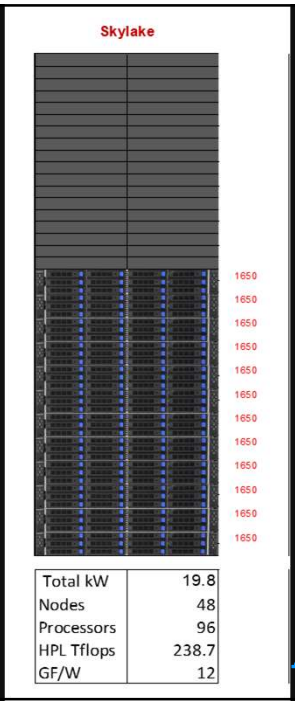
Total kW	36.5
Nodes	90
Processors	180
HPL Tflops	69.3
GF/W	1.9

7

2018, Skylake

- 19" rack, air cooled
- 2U, 4 node dual socket configuration
- Bull X400

- Rack is 57% full



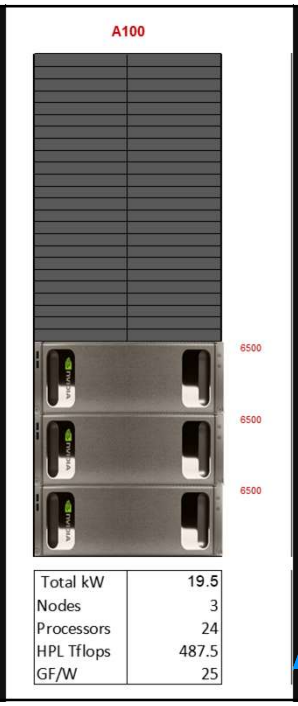
Total kW	19.8
Nodes	48
Processors	96
HPL Tflops	238.7
GF/W	12

8

2021, NVIDIA A100

- 19" rack air cooled
- DGX1 – 8x A100 in 6U per node

- Rack is 43% full



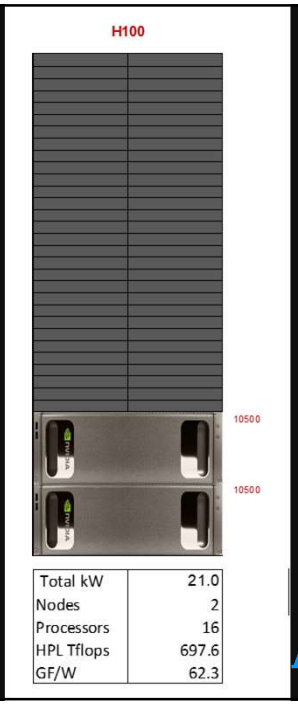
9 | Dec 2022 – CIUK | © Atos

9

2023, NVIDIA next gen

- NVIDIA Hopper based DGX
- 6U, 8 GPUs per node

- Rack is 28% full



10 Dec 2022 – CIUK | © Atos

10

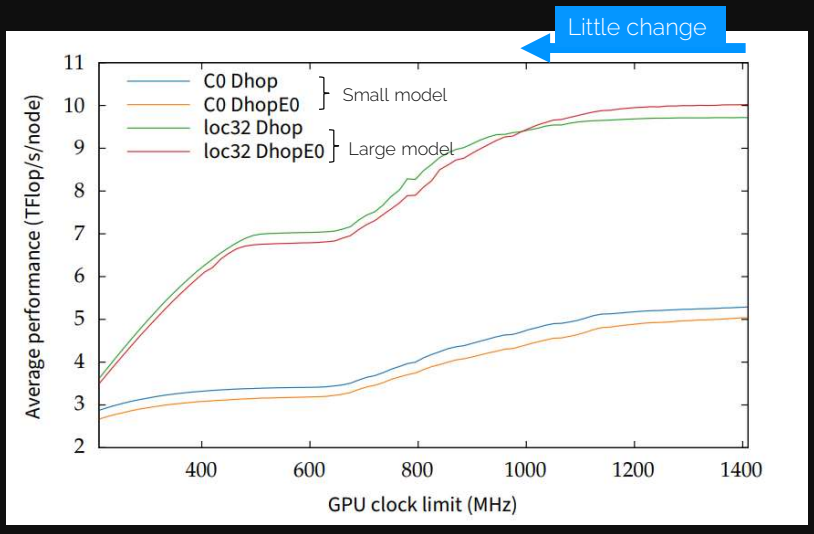
Edinburgh HPC energy use study

- Based on the 'Tursa' system, live in early 2022
- System is a BullSequana XH2000 configured with 112x A100-40 blades.
- Each blade is one node, with
  - 4x NVIDIA A100-40 GPUs
  - 2x AMD Rome
  - 4x NVIDIA InfiniBand 200Gbps networks
- High performance network is unblocked HDR, providing 800Gbps to the node, or 200Gbps to each GPU
- System is designed for maximum applications scalability for GPU enabled codes
- Apart from applications performance, optimising energy use is **critical**

11

Edinburgh energy study

- Using real application code to measure energy use
- This study runs the code, while measuring energy use per job
- Study maps performance against clock rate of A100
- Very little performance gain above 1GHz
- Translating to cost, clocking back A100 saves ~20% energy cost for virtually no loss of science
- Thanks to Antonin Portelli, Edinburgh University



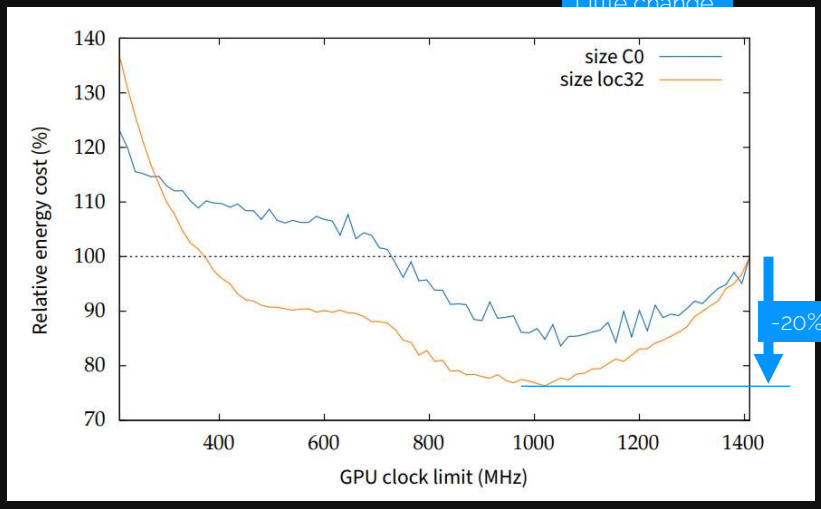
https://zenodo.org/record/7057319#.Y2okWXbP2Um

12

Edinburgh energy study

- Using real application code to measure energy use
- This study runs the code, while measuring energy use per job
- Study maps performance against clock rate of A100
- Very little performance gain above 1GHz
- Translating to cost, clocking back A100 saves ~20% energy cost for virtually no loss of science

- Thanks to Antonin Portelli, Edinburgh University

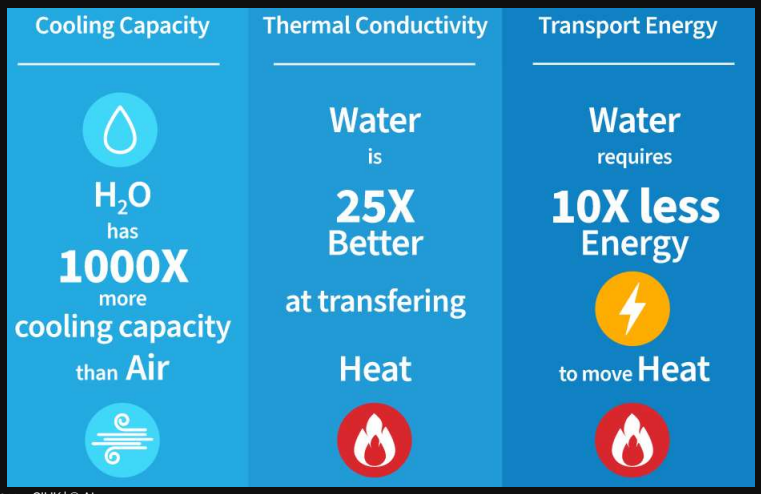


https://zenodo.org/record/7057319#.Y2okWXbP2Um

13

Overview of cooling technologies
Using water instead of air for components cooling

- Atos has offered Direct Liquid Cooling servers since 2011



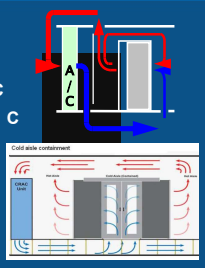
Atos continues to investigate other technologies such as liquid immersion cooling, which may be needed in the future as processor power continues to climb

14

Impact on PUE and DC operating costs (Tier2 assumption)

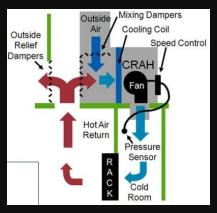
Air-cooled
5-20 kW/rack
Room 20° C
A/C water 7-12° C

PUE 1,6-2



Air-cooled with inrow
10-30 kW/rack
Room 20° C
A/C water 7-12° C

PUE 1,4-1,6

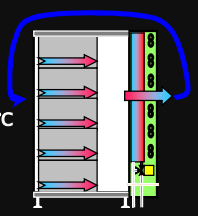


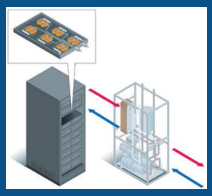
Free cooling air/air
In container or next generation datacenter
Room up to 35°C

PUE < 1.1

Water-cooled doors
40 kW/rack
Room 23°C
Water 12-18°C

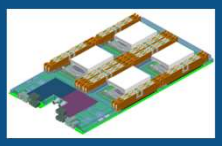
PUE 1.3-1.5





DLC Water Block
80 kW/rack
Room up to 27°C
Water 38-47°C

PUE 1,2-1,3



Full DLC
150 kW/rack
Room up to 27°C
Water 38-47°C

PUE < 1.1

15

5 MW Cost Study comparing Atos HW technology

PUE, energy use, and CO2 are key factors in the cost of a large-scale HPC/AI solution

Based on 4000 X400 or XH3000 nodes with the following assumption :

- Power consumption 1250 W/nodes
- Price 7500 € per node
- X400 CooliT nodes are limited to 60kW 12 double-twin per rack,
- air solution is based on 20kW/rack datacenter
- 4-years TCO

5 MW - 4000 compute nodes	Full DLC system XH3000	DLC water block system X400 CooliT	Air system X400 air
Rack number	42	83	250
PUE	1.1	1.3	2
IT estimated cost	€ 42,416,667	€ 32,500,000	€ 30,750,000
Infrastructure	€ 4,166,667	€ 8,333,333	€ 25,000,000
Cooling	€ 1,000,000	€ 3,000,000	€ 4,000,000
Total CAPEX	€ 47,583,333	€ 43,833,333	€ 59,750,000
CAPEX gbp	£ 41,557,496	£ 38,282,387	£ 52,183,406

	€/KWH	XH3000	X400 CooliT	X400 air
Finland	0.04	€ 55,292,133	€ 52,943,733	€ 73,766,000
France	0.25	€ 95,763,333	€ 100,773,333	€ 147,350,000
Germany	0.4	€ 124,671,333	€ 134,937,333	€ 199,910,000
United Kingdom	0.229	£ 62,335,667	£ 67,468,667	£ 99,955,000

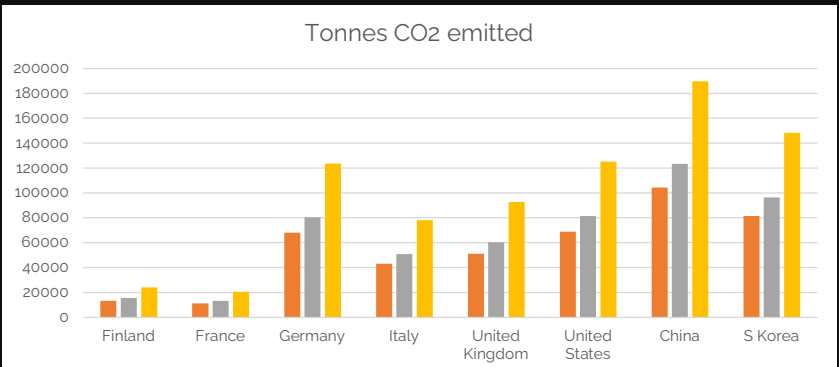
16

5 MW Cost Study comparing Atos HW technology

PUE, energy use, and CO2 are key factors in the cost of a large-scale HPC/AI solution

Based on 4000 X400 or XH3000 nodes with the following assumption :

- Power consumption 1250 W/nodes
- Price 7500 € per node
- X400 CooliT nodes are limited to 60kW 12 double-twin per rack,
- air solution is based on 20kW/rack datacenter
- 4-years TCO



	XH3000		X400 CooliT		X400 air	
	Infra + Power	Carbon	Infra + Power	Carbon	Infra + Power	Carbon
Finland	€ 55,292,133	€ 662,957	€ 52,943,733	€ 783,494	€ 73,766,000	€ 1,205,376
France	€ 95,763,333	€ 562,742	€ 100,773,333	€ 665,059	€ 147,350,000	€ 1,023,168
Germany	€ 124,671,333	€ 3,395,726	€ 134,937,333	€ 4,013,131	€ 199,910,000	€ 6,174,048
United Kingdom	€ 71,374,338	€ 2,548,722	€ 77,251,623	€ 3,012,126	€ 114,448,475	€ 4,634,040

17

BullSequana XH3000

6 Key Value Proposition Points

1 Ready for Accelerated Hybrid computing

- Most powerful computing technologies embedded
- Ready for GPU and QPU acceleration next

intel AMD NVIDIA SIPEARL

2 Any Scale

- "All-In-One" rack design
- Seamless scalability to ExaScale

3 Flexible and dense by design

- Standardized slots to mix and interchange compute and switch blades
- Large choice of computing and interconnect technologies
- Industry-leading density

4 Unrivaled efficiency

- 4th generation DLC technology with 40°C inlet water temperature
- Smart Power Management Suites
- Lowest TCO with free cooling and PUE ~1.0

5 Open & Future-proof

- Standardized design for easy future evolutions
- OpenSequana program to seamlessly integrate 3rd party innovations

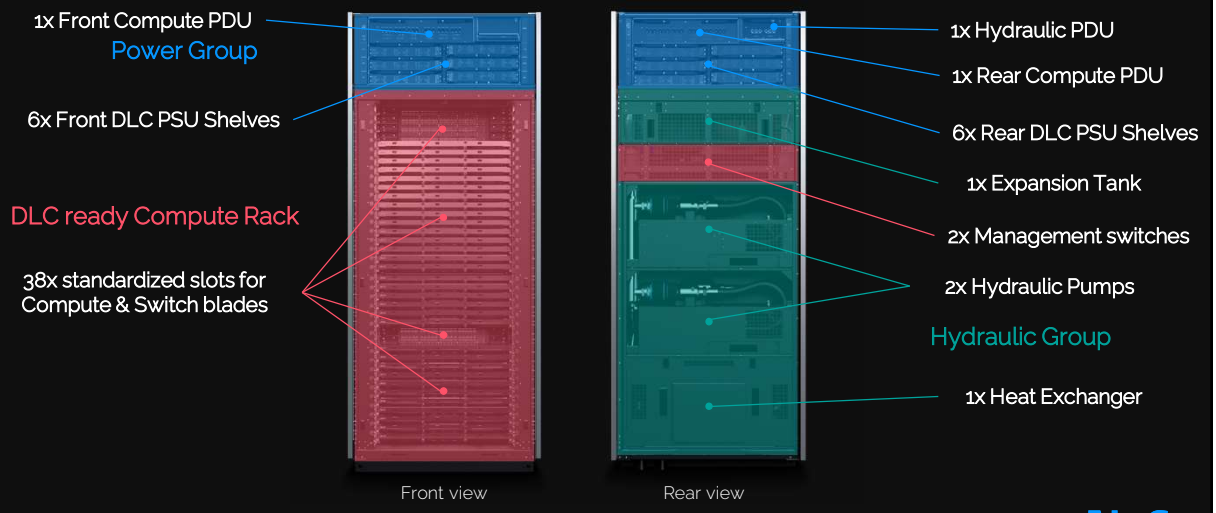
6 Secured by design

- Embedded security with hardware-anchored secure boot
- Vulnerability monitoring and threat anticipation

18

Infrastructure overview

Architecture Overview



19

NetZero[1] | Context

we share the same challenges

1
Economic

40[2] billion tCO_2

Get to Net Zero

2
Social

3
Legal

European Commission

THE EU TAXONOMY

ERP COMPLIANT DIRECTIVE 2009/125/EC ECODESIGN

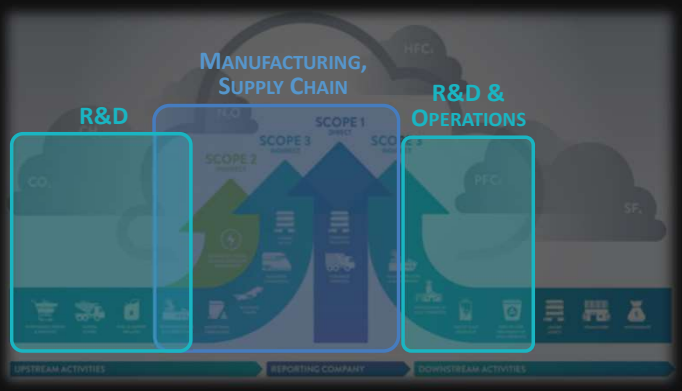
THE GREEN 500 level 2/3 EEHPC WG: Power Measurement Meth.

[1] On 28 October,, the Science Based Targets Initiative (SBTi) released the Net Zero Standard, which represents the first global standard for corporate Net-Zero target-setting in line with science
[2] Energy Transition Commission: Mind the Gap: How Carbon Dioxide Removals Must Complement Deep Decarbonization to Keep 1.5°C Alive (March 2022)

20

Sustainability | A holistic approach

everyone has a role to play



CONSIDERATIONS

Atos impact is possible across **the entire lifecycle:**

1. Eco **DESIGN**.. designing Green(er) IT. In some cases, the Green IT can be stand alone technology that can be leveraged as part of an IT for Green solution for our customers
2. Eco **MANUFACTURING & SUPPLY CHAIN**.. reduce energy consumption and waste in the Angers factory and influence supply chain behavior
3. Eco **OPERATIONS**.. influence and incentivize client behavior with new "business" elements that can in included in our Decarbonization Level Agreements (i.e. reduction on non-green energy use, DC PuE improvements, ecoact offsets...)

21

Frugal HPC

Our DNA, Green HPC Platforms

Sustainability with a holistic approach:

- REQUIREMENTS & DESIGN
- TENDER
- FACTORY
- DELIVERY
- PRODUCTION & EOL

TODAY

1/ Generation of compute... **less energy lost**
Direct Liquid Cooling with 40°C water at the inlet

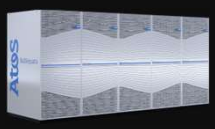
2/ **Understand** compute consumption
Simulate peak and operational consumption phase
Accurate modeling of production performance

3/ Effective use of compute... **less energy used**
Profile and optimize application behavior to minimize its energy footprint by optimizing runtime (MPI. etc.) and I/O flows (FastIO Library, SBB/SBF)

Measure energy and temperature metrics from all components of the supercomputer and link them with job execution information (provide per-job information)

Policy driven control and capping of power consumption. Static rules today, tomorrow "tight coordination" with resource planning

4/ **Ensure Availability** of compute
Predictive maintenance to reduce energy waste by optimizing execution and/or avoiding unproductive jobs



eco friendly design

software solutions to measure & control energy to solution

high availability supported by predictive maintenance

sovereignty via European Centric Supply Chain

22

From frugal to carbon neutral HPC

covering CO_2 footprint measurement, reduction, reporting and offsetting

Sustainability with a holistic approach:

Requirements & Design
Tender
Factory
Delivery
Production & EOL

Today

1/ 3rd generation of DLC compute blades…
Energy efficient design
Direct Liquid Cooling with 40°C water at the inlet

2/ **Understand** supercomputer consumption
Simulate peak and operational consumption phase
Accurate modeling of production performance
Automatic out-of-band data gathering in production

3/ Effective use of compute…
Optimize energy usage
Profiling tools (runtime, I/O, energy, etc.)
Live energy optimization (automated DVFS)
Power capping
Static rules today, tomorrow "tight coordination" with resource manager

4/ **Maximize solution availability**
Out-of-box monitoring (logging, events, etc.)
Predictive maintenance to reduce energy waste by optimizing execution and/or avoiding unproductive jobs

Tomorrow

5/ **Carbon Accounting**
Visibility on the Baseline CO_2 footprint impact on TCO (complete lifecycle supply chain, run, end of life)

European elements in our HPC TCO > 50% means reduction in supply chain CO_2 footprint related to transportation (70% by 2023/2024, 90/100% by 2030)

Provide CO_2 removal credits to offset the incompressible part of the platform carbon footprint

6/ Angers Factory, **carbon neutral** since 2017
Site Energy consumption, Transportation
Plastic free program, Waste generation…

7/ **Decarb Levers**, to reduce the CO_2 baseline
Build Green Co-innovation programs with Atos R&D teams to measure, reduce and report on CO_2 footprint

Leverage a set of existing – or develop new – "decarbonization levers" that unlock potential carbon footprint and energy savings

Provide user-land information (job-related) related to energy consumption and CO_2 footprint (end-of-job report, dynamic power optimizer, etc.)

CO_2 footprint per BOM component

Includes extraction, fabrication, transportation and end of life

Bureau Veritas EIME LCA database

40% to 60% of CO_2 total lifecycle footprint
(depending on energy source + system size)

23

From frugal to carbon neutral HPC

Innovation driving the next steps

Sustainability with a holistic approach:

Requirements & Design
Tender
Factory
Delivery
Production & EOL

Today

1/ 3rd generation of DLC compute blades…
Energy efficient design
Direct Liquid Cooling with 40°C water at the inlet

2/ **Understand** supercomputer consumption
Simulate peak and operational consumption phase
Accurate modeling of production performance
Automatic out-of-band data gathering in production

3/ Effective use of compute…
Optimize energy usage
Profiling tools (runtime, I/O, energy, etc.)
Live energy optimization (automated DVFS)
Power capping
Static rules today, tomorrow "tight coordination" with resource manager

4/ **Maximize solution availability**
Out-of-box monitoring (logging, events, etc.)
Predictive maintenance to reduce energy waste by optimizing execution and/or avoiding unproductive jobs

Tomorrow

5/ **Carbon Accounting**
Visibility on the Baseline CO_2 footprint impact on TCO (complete lifecycle supply chain, run, end of life)

European elements in our HPC TCO > 50% means reduction in supply chain CO_2 footprint related to transportation (70% by 2023/2024, 90/100% by 2030)

Provide CO_2 removal credits to offset the incompressible part of the platform carbon footprint

6/ Angers Factory, **carbon neutral** since 2017
Site Energy consumption, Transportation
Plastic free program, Waste generation…

7/ **Decarb Levers**, to reduce the CO_2 baseline
Build Green Co-innovation programs with Atos R&D teams to measure, reduce and report on CO_2 footprint

Leverage a set of existing – or develop new – "decarbonization levers" that unlock potential carbon footprint and energy savings

Provide user-land information (job-related) related to energy consumption and CO_2 footprint (end-of-job report, dynamic power optimizer, etc.)

Future

8/ **Global topic** supported by multiple groups
Green4BDS, dedicated R&D teams in BDS and at Atos Group level

9/ Increase **"green" energy** consumption
Local Green Energy with H_2 cell designed for HPC

Model energy consumption profiles and optimize usage of smartgrid capabilities to manage peak loads

10/ Modular, extensible, adaptable **architectures**
Optimized heterogeneous scheduling across "sub clusters" with high variety of hardware (including shared memory models)

11/ **Data aware**, data usage efficiency (DuE)
Consider data path and data volume in scheduling policies linked to application profiles

24

Decarbonization Level Agreements | Benefits

Decarb Dashboard
Provide visibility on carbon footprint evolution to support regulatory reporting & internal program needs

Accelerate IT Decarb
An auditable service which is aligned to your sustainability objectives

DLA

Tangible CO_2 Reduction
Binding commitment to reduce the footprint across your entire emission landscape

Carbon Neutral Compute
Driven by an eco friendly design + co-innovation program + innovative business model

Reduced Risk & TCO
Lowers exposure to risk related to the fluctuation in carbon offset prices. Reduction in energy usage lowers TCO

25

Nimbix Supercomputing Suite
Global Elastic Supercomputing as-a-Service

3 Supercomputing "as-a-Service" Consumption Models

Elastic

Pay-as-you-go, on-demand, secure and scalable supercomputing through a single user interface.

Dedicated

Leverage powerful dedicated Bullsequana HPC servers as "Bare Metal as-a-Service" for the best of infrastructure and on-demand scalability, convenience, and agility.

Federated

Federated Supercomputing-as-a-Service offers a unified service console to manage all compute zones and regions in a public or private HPC, AI, and supercomputing federation

26

JARVICE XE software for Public, Private, or Hybrid HPC

Simplified User Experience

Simple point-click-run workflows on any infrastructure

Simplified Administration

Unified SaaS for HPC and Deep Learning

Unified platform for multi-cloud, multi-datacenter deployments

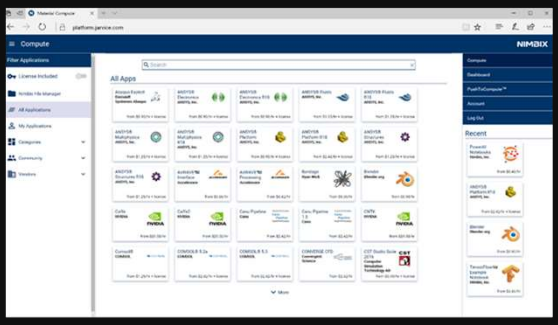
Containerized Application Distribution & Deployment

Platform-as-a-Service (PaaS) continuous integration and deployment for in-house algorithms or customization of commercial applications

Automatic synchronization with HyperHub Application Marketplace

Reduced Infrastructure Complexity

Unified infrastructure layer with Kubernetes



27

Summary

Smart software tools are invaluable

Systems and DCs need to be considered together

Smart hardware design

We need a holistic approach

Sustainability considerations & Technology trends point in the same direction

28

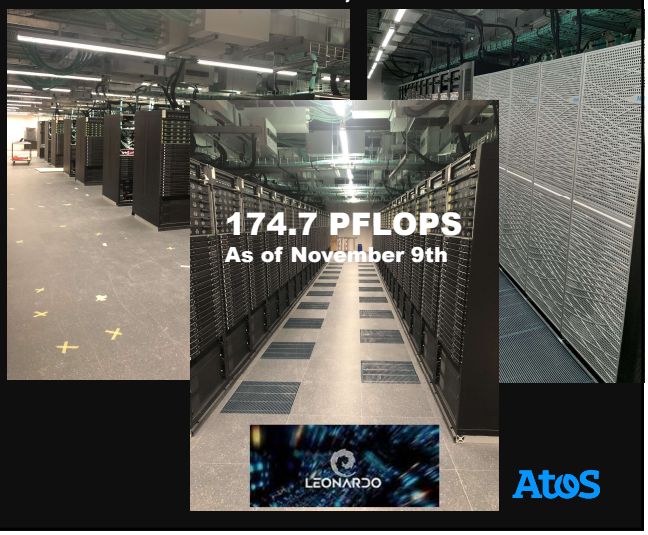
CINECA Installation

Fantastic effort to deliver and Build Leonardo

EuroHPC Joint Undertaking

May 2022

Today



29

30