Available online at www.prace-ri.eu

Partnership for Advanced Computing in Europe

HPC Infrastructures Workshop 2021

Gert Svensson[a] *, Hayk Shoukourian[b], Volker Weinberg[b], Huub Stoffers[c], Radosław Januszewski[d], Mirko Cestari[e], Ezhilmathi Krishnasamy[f], Andreas Johansson[g], François Robin[h], Javier Bartolome[i], Rafael Meideros[j], Nicolas Eyraud[k]

[a] *PDC Center for High Performance Computing, KTH Royal Institute of Technology, Sweden*
[b] *Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities (BADW-LRZ), Germany*
[c] *SURF, Netherlands*
[d] *Poznan Supercomputing and Networking Center (PSNC), Poland*
[e] *CINECA, Spain*
[f] *University of Luxembourg, Luxembourg*
[g] *National Supercomputer Centre (NSC) at Linköping University, Sweden*
[h] *Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), France*
[i] *Barcelona Supercomputing Center, Spain*
[j] *Institute for Development and Resources in Intensive Scientific Computing (IDRIS), France*
[k] *Paris-Saclay District Heating and Cooling Network, France*

Abstract

The annual series of European Workshops on HPC Infrastructures (EWHPC) aims to bring together specialists in High-Performance Computing (HPC) centre design and operation from around the world to discuss the latest trends in HPC infrastructure and supporting technologies for supercomputing centres.

Due to the COVID-19 pandemic, the 11th EWHPC was held as an online event. It attracted more than 110 registrations from 53 different HPC sites and provided a general overview on the EuroHPC Joint Undertaking (JU) as well as updates on the following three EuroHPC JU petascale sites and their systems:

- MeluXina, LuxProvide, Luxembourg,
- VEGA, IZUM, Slovenia, and
- Karolina, IT4Innovations National Supercomputing Center, Czech Republic.

The 11th EWHPC event was held on the 18th and 19th of May 2021 and was followed, as usual, by a PRACE session on HPC infrastructures. The session featured technical presentations from PRACE sites and was held on the 20th of May 2021.

This document summarises the presentations and discussions held during the online EWHPC event and the PRACE session on HPC infrastructures.

* Corresponding author email address: gert@pdc.kth.se

Table of Contents

List of Figures

List of Tables

1 Introduction

The 11th European Workshop on HPC Infrastructures (EWHPC) had initially been planned to be held at the Leibniz Supercomputing Centre (LRZ), Bavarian Academy of Sciences and Humanities, from the 25th to 27th May 2020, however that was not feasible due to the COVID-19 pandemic. Instead, a shorter online version of the event was offered on the 14th October 2020. The 11th workshop was instead postponed until 2021. It was held as an online event on the 18th and 19th of May 2021 and was followed on the 20th of May by a PRACE session on HPC infrastructures (which traditionally follows the EWHPC workshop each year). The PRACE session was also held online and featured technical presentations from PRACE sites. The 11th European Workshop on HPC Infrastructures is summarised in this document, along with the associated PRACE session.

The program committee for the online EWHPC event consisted of the following site representatives:

- François Robin and Jean-Philippe Nominé (CEA),
- Ladina Gilly and Tiziano Belotti (CSCS),
- Herbert Huber and Michael Ott (LRZ),
- Javier Bartolomé (BSC),
- Norbert Meyer (PSNC), and
- Gert Svensson (KTH).

In addition, Evangelos Floros (EC) and Oscar Diez (EC) were invited to attend the meetings of the programme committee.

As with previous events in the EWHPC series, attendance at the 2021 workshop was by invitation only. The event brought together more than 100 participants from 53 different supercomputing sites. It provided a unique platform for sharing the latest infrastructure trends and issues, as well as lessons learned at HPC sites facing similar challenges and requirements.

The agenda of the EWHPC 2021 workshop featured six sessions, starting with an overview presentation on the EuroHPC Joint Undertaking (JU) followed by updates on the following three EuroHPC JU petascale systems:

- MeluXina, LuxProvide, Luxembourg,
- VEGA, IZUM, Slovenia, and
- Karolina, IT4Innovations National Supercomputing Center, Czech Republic.

A talk concerning planning for next year's 12th EWHPC (which will be held in Kajaani, Finland, from the 8th to the 10th of June 2022) concluded this year's workshop. For up-to-date information about future workshops, please refer to the EWHPC official website: https://www.euhpcinfrastructureworkshop.org/.

As mentioned earlier, the EWHPC event in 2021 was followed by the traditional PRACE session on HPC infrastructures (which included technical presentations from PRACE sites) on the 20th of May 2021. The PRACE session on HPC infrastructures included attendees from PRACE Tier-0 and Tier-1 sites and featured talks given by five PRACE partners (LRZ, SURF, KTH, CINECA and BSC). These talks were followed by a presentation on requirements for exascale system installation based on a survey of PRACE sites that was undertaken by LRZ. The session finished with a talk about heat reuse with the IDRIS supercomputer, Jean Zay, in Paris. The PRACE session also provided further opportunities for the experts in attendance to exchange best practices.

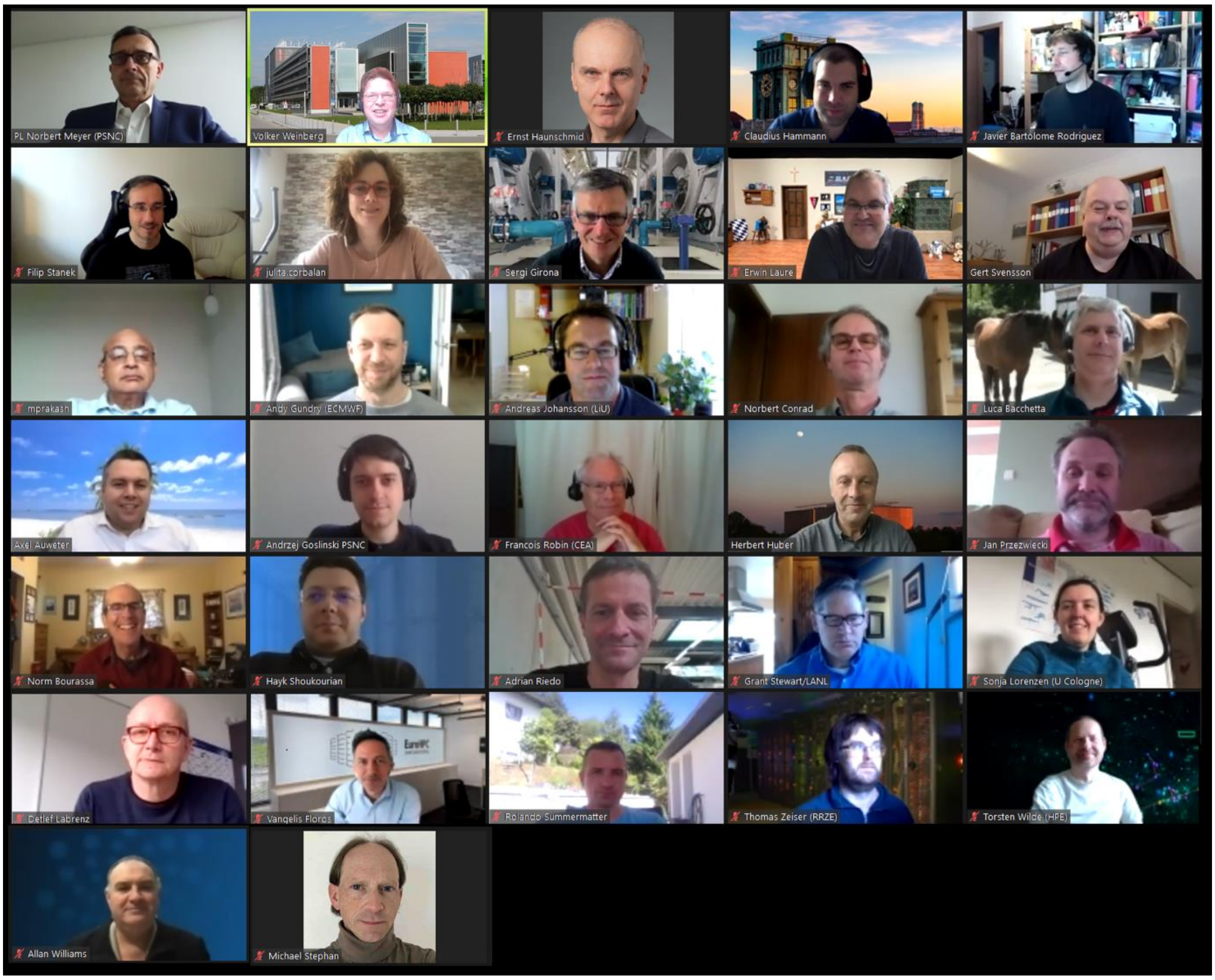

Table 1: Attendees at one session during the workshop

This document provides a summary of the facts, insights, and best practices that were discussed during the EWHPC 2021 online event and the following PRACE session on HPC infrastructures so that information can be more broadly disseminated amongst interested parties and experts involved in HPC facility management. The proceedings of the previous series of EWHPC events can be found via: https://prace-ri.eu/infrastructure-support/european-workshops-on-hpc-infrastructures/.

2 Session I – Chaired by Herbert Huber (LRZ)

2.1 High-Density In-Rack Power Distributions for HPC – "Why every connection counts"

Adrian Riedo, Riedo Networks

Adrian Riedo is employed by a company founded in 2005 that is usually referred to as "RNX", which is short for "Riedo Networks". RNX provides systems for power distribution and monitoring to data centres, including HPC data centres. The company, based in Switzerland, manufactures its own hardware and software solutions, mainly in Switzerland and the European Union.

The power density per rack has increased steadily in HPC facilities and is expected to continue doing so. The first part of Adrian Riedo's presentation summarised the basics of power distribution in a data centre, which will help people to more easily grasp the contemporary challenges in this field of engineering that were addressed in the second part of the presentation, namely safe and efficient in-rack power distribution for high-density racks. The third part of the presentation was a small case study about a high-density rack (which is part of the "Alpha Centauri" computer produced for the University of Dresden) that elucidates the balancing of requirements and making of trade-offs that are necessary when engineering a power provisioning solution.

2.1.1 Power distribution basics

The electrical power distribution in a data centre can be compared to a tree. The various levels of branching each have their own circuit breakers that match the magnitude of the currents that must be handled. Besides safety, "selectivity" (i.e., strictly containing any power overdraw or short circuit events, so they trip only the single, nearest circuit breaker, and avoid any cascading effects on circuit breakers higher up in the power chain) is of utmost importance for the design of the power distribution tree.

At the root of the tree is the connection between the data centre and the local electrical grid. Typical currents at this level are in the order of one to several tens of thousands of amperes, representing power levels of five megawatts and beyond. In order to distribute currents of this magnitude in a safe manner, ample physical space is needed: the distribution panels built for this purpose can take up several rooms.

After the first segmentation at the root, branches go into different areas of the data centre building, using very thick cables or bus bars. The currents in these branches typically support currents ranging from several hundred amperes up to one thousand amperes. Most installations have yet another intermediate level of power distribution panels, especially when power is delivered to the racks via cables rather than via bus bars with tap-off boxes.

Finally, at the rack level – the leaf level of the tree for a data centre - a three-phase Alternating Current (AC) is supplied, in accordance with common European standards, with a nominal voltage of 230 volts, phased to neutral in a star-wiring topology.

2.1.2 In-rack power distribution

The in-rack power distribution constitutes a tree in itself: power from the rack Power Distribution Unit's (PDU's) inlet is distributed further to multiple outlets. In most cases, especially for PDUs used in HPC systems, another set of circuit breakers limits the current of those outlets.

The following three classes for rack PDUs are in common use in European data centres.

- **11 kW (3 × 16 A feed)**

  A rack density of 11 kW for HPC systems is largely a thing of the past. However, these PDUs are still very popular for non-HPC applications or some non-HPC parts of a larger HPC facility because they do not require additional circuit breakers within the rack. 3 × 16 amperes are simply fed into the outlets. An overloaded power draw on the collective outlets, or a short circuit, will trip the circuit breakers that are installed in the panel board or in the tap off-box outside the rack.

- **22 kW (3 × 32 A feed)**

  These are often used in current HPC systems. The PDUs have the same (55 mm wide) form factor as the 11 kW PDUs but are a bit longer in order to accommodate six

miniature circuit breakers. The 22 kW of the inlet is split into six modules that are each limited to about 3.6 kW. Each module has eight outlets, 48 sockets in total, but the total power per module is limited to 3.6 kW. If there is an overdraw on the group of connected sockets, the miniature circuit breaker for the corresponding module will trip.
Most HPC racks do not require power redundancy. It is common practice to install two of these PDUs in a single HPC rack, either fed from the same bus bar or from two different bus bars. Thus, these PDUs are used in a manner that enables a rack density in the 40 kW range.
If redundancy is required with this rack density, obviously, four of these PDUs would be necessary. This configuration is very rare, as it is pretty challenging in terms of the physical placement of the PDUs in the rack.

- **44 kW (3 × 63 A feed)**

  To provide 44 kW per PDU, 12 miniature circuit breakers are needed. The RNX product of this type, as well as the products for the other implementations, still maintain the de facto standard width of 55 mm, but the total number of outlets is reduced to 36. The more circuit breakers are required, the less space is left to accommodate outlets – unless the format were to be changed, but that would lead to a number of other practical inconveniences.

Adrian Riedo considers that this type of PDU could become the de facto standard for most HPC installations. Still, some cases require racks with more power density and consequently higher currents fed into a single PDU. The following engineering question arises from such cases.

2.1.3 High-current challenges

How can more power – specifically 66 kW / 3 × 100 A – be packed into a PDU of almost the same volume while minimising heat loss?
As the current rises, so do the losses of power to heat. There are three aspects to take into account.

- **Resistance of wired connections**

  The following formula expresses a simple inescapable truth of physics that applies to wiring losses: $P = R \times I^2$, where P is (loss of) power, R is resistance, and I is the current. So, it is essential to minimise the resistance of every wired connection.

- **Circuit breakers**

  With 12 or more circuit breakers in a PDU, the choice of circuit breaker model is another important aspect. At full load, the selection of an ideal circuit breaker can reduce power losses to heat up to 50%.

- **Outlet connectors (and cables)**

  Every outlet will generate heat as well, of which the largest part is transferred to the cables – luckily – and so does not stay within the PDU. The better the connection and cables, the lower the resistance. RNX recommends cables that have a copper cross-section that can easily carry a bit more than 16 amperes.

RNX did research on all three aspects to establish their relative importance and to assess the relative merits of alternatives.

2.1.3.1 Wired Connections inside the PDU

Clamps to connect wires that are excellent connectors exist and are a convenient product. They have spring contacts that compensate for temperature differences. Nevertheless, clamps will never be better than bare copper connections. RNX compared the power losses in clamps with power losses in soldered joints and ultrasonic welding joints. RNX found that ultrasonic welding produces the most compact solution with the lowest power losses to heat. In a high-power PDU, where hundreds of connections have to be made where there is barely any space to do so, ultrasonic welding, therefore, is the preferred technology. The heat loss reduction with these joints is 20 times lower than can be achieved with clamps. The joint in ultrasonic welding actually contains more copper than the wires on either side of the joint and is of such high quality that the power losses in the joint are actually a bit lower than with pure wire.

Figure 1: Clamps (left) Ultrasonic weld (right)

In a comparison performed in an experiment with currents of 25 amperes, the difference between clamps and ultrasonic welding joints in degrees Celsius for a single connection was only about 6 degrees. This may not seem much, but there are hundreds of such connections, and literally every connection counts.

2.1.3.2 Miniature Circuit Breakers

Miniature circuit breakers in high current PDUs are mandatory. There are two types of technologies commonly in use: hydraulic-magnetic and thermal-magnetic circuit breakers. Both types use a coil to generate a magnetic field that moves a piston. In the case of a short circuit, the current in the coil rapidly increases and makes the piston open a contact. The circuit breakers must also trip in overload situations, where the rise of the current is much slower. In this respect, the two types are fundamentally different. The thermal-magnetic device has a bimetal that slowly deforms. If the load is high enough, the amount of deformation will trigger the opening mechanism. The hydraulic-magnetic device uses the very same piston as is used for short circuits. It is put in a fluid to achieve a time delay – hence the name “hydraulic”.

The thermal variant is generally cheaper but generates more heat loss, especially when operating close to maximum current. In comparison experiments performed by RNX with currents of 16 amperes, the thermal-magnetic device generated about 50% more heat. The hydraulic variant trips sooner and is more sensitive to load changes near the maximum load.

Hydraulic-magnetic devices are commonly found in US data centres as a technology to implement circuit breakers at various levels of the power provisioning tree. In Europe, we mostly find thermal-magnetic technology used at the facility level. This has consequences for the type of circuit breaker model that can be used in the rack PDUs. In principle, the hydraulic variant is recommended in high-density racks. However, in many situations, selectivity of circuit breakers is required, but more difficult to guarantee if a mixture of different types is used at different levels of the power chain.

2.1.3.3 Outlets and Cables

So-called C19 and C13 connectors (see Figure 2), as specified by the IEC 60320 standard, are commonly used. In line with expectations, the C19 connectors have less power loss to heat than the C13 connectors. The loss of the C19 connector in an experiment with a current of 20 amperes was even less than the loss of a C13 connector with a 15 A current. However, with the currents well below 16 A, the choice of the connector quickly loses relevance.

Figure 2: C13 connector (left), C19 connector (right)

2.1.4 High-density PDU for "Alpha Centauri"

Alpha Centauri is the name of a high-density installation at the University of Dresden. The cluster consists of 34 high-density direct liquid-cooled nodes, provided by NEC Deutschland GmbH, that each contains the following components.

- 8 × NVIDIA A100-SXM4 GPUs
- 2 × AMD Epyc 7352 CPUs (24 cores @ 2.3 GHz)
- 1 TiB main Random-Access Memory (RAM)
- 3.5 TB Non-Volatile Memory express (NVMe) device

With that hardware installed, the maximum power usage per node is about 4.8 kW. However, the chassis of the nodes is equipped with four power supplies of 3 kW each. The trend is clearly to go up to densities up to 10 kW in the same chassis space. One critical requirement decision to be made is how many power supplies go to a single circuit breaker. Ideally, from a power provisioning perspective, it would be just one, but there are other design criteria, such as saving costs. After considering the various criteria, the following requirements were settled on for the rack.

- 1 or 2 PDU per rack
- At least 44 kW, possibly 60 kW+, per rack
- At least 36 high power 16 A connections per rack
- At least two C19 power supply connections available per circuit breaker

The final product was a 66 kW (3 x 100 A) PDU, 2 metres long, with the following characteristics.

- Eighteen magnetic-hydraulic miniature circuit breakers

- All internal wires are made of 3 x 10 mm^2 copper
- All joints of wires are made by ultrasonic welding
- Power inlet: 3 x 100 A, polyurethane-isolated cabling

Each circuit breaker can provide full power to a single C19 connector, but a second connector is in place which can obviously be used if the maximum power draw of the hardware that is actually connected warrants it. Because the circuit breakers are of fixed length, there is room for a third connector – possibly of another type.

The power distribution problem per se, in this case, could also have been solved with four traditional PDUs, which then would require connections to four tap-off boxes. Now, a rack can be run with a single PDU. Besides the minimisation of losses to heat, the reduction of complexity and the saving of costs *outside* of the rack is an essential factor that has been taken into account. The higher density PDU reduces complexity and costs outside of the rack by reducing the number of tap-off boxes and feeds involved. This reduced the number of circuit breakers outside the rack that are involved. In addition, it reduces the number of residual current monitoring sensors that are mandatory for many installations nowadays.

2.2 Modular and Monolithic UPS designs – A Comparison

Stephania Ceccarini, Borri/Legrand

Stefania Ceccarini is a product specialist at Borri UPS. She gives technical advice about data centre level Uninterruptible Power Supply (UPS) products. Her presentation was an overview of different UPS types and different UPS modes of operations that each have their advantages and disadvantages. The presentation can serve as a small reference handbook for assessing which UPS type and mode of operation are most suitable for a particular data centre.

2.2.1 Data centre typology

Data centres vary in size and in terms of their importance for their owners and users. Therefore, the features that are required from a UPS and the relative weight of key factors in selecting an appropriate UPS depend to a large extent on the type of data centre where the UPS will be installed. Data centres can be classified as follows based on the usage of their IT equipment.

- **In-house** data centres in common, not IT-related, businesses: IT equipment in these types of data centres is often not required to be available 24/7. Such data centres are usually relatively small. The impact of downtime for IT equipment is relatively low. Maintenance that requires downtime can be scheduled relatively easy.
- **Enterprise** data centres are self-owned data centres whose core business is IT or highly IT-dependent. Most IT equipment needs to be in production 24/7. Downtime impact is high.
- **Colocation** data centres are owned by IT space rental companies. Their core business is to provide data centre services, housing and hosting, 24/7. Downtime impact is very high. In principle, all data centre maintenance must be performed in such a way that it is transparent to the data centre's customers.

2.2.2 Key factors in UPS selection

There are two main classes of UPS topology – monolithic and modular – that are addressed in this subsection. Irrespective of the UPS topology, however, there are three drivers in UPS selection that are invariably important in the selection of a UPS.

- **Power efficiency** is the desire to minimise power usage by the UPS itself. The power efficiency of a UPS is directly related to the data centre's electricity bill.
- **Scalability** refers to the ability to change – usually to increase – the amount of power that can be delivered to the data centre, preferably in a cost-effective way and preferably also in a way that is not disruptive to the activities going on in the data centre.
- **Availability** is the capability of the electrical infrastructure to sustain the load required by the data centre without downtime – especially for enterprise and colocation data centres, ideally continuously, 24/7, without any incidents or changes in the infrastructure having a disruptive effect on any part of the IT equipment in the data centre.

2.2.3 Power efficiency and UPS modes

Optimising the power efficiency of a UPS is inextricably linked to selecting the mode of operation, and it is a trade-off game without any free rides. Improving the power efficiency is achieved by not using some of the protective features that a UPS could offer. Choosing a particular mode requires a risk assessment in which the compatibility of the IT load and the stability of the local grid are key considerations.

Current UPS systems offer three distinct operating modes that are colloquially called "online double conversion", "line interactive", and "eco mode". In the formal UPS standards, these modes are designated respectively as "VFI", "VI", "VFD". V stands for voltage, F is frequency. D means dependent, and I means independent. So, VFI implies that the voltage and frequency of the output are independent of the voltage and frequency of the input. Two further variations on these modes are presented as well.

2.2.3.1 Online Double Conversion Mode (VFI)

This is the classical mode of operation, in which a UPS is continuously online, and all load is fed through the UPS. The input is continuously fed through surge protection, and the UPS is constantly regenerating the output voltage and waveform that the load needs. This is the most protective mode of operation that a UPS can offer. It is also the least power-efficient mode.

2.2.3.2 Automatic Power Adaptation for Online Double Conversion Mode (VFI)

A relatively recent development is that the VFI mode of operation can be combined with automatic adaptation of the power capacity of the UPS by switching on and off particular components in response to the load. The power efficiency of a UPS is highest when it is operated at the high end of its capacity. The effect of this improvement is shown in Figure 3, which compares the performance of a modern UPS with a "legacy" UPS of a decade ago, which lacks the adaptive capability of modern UPSes.

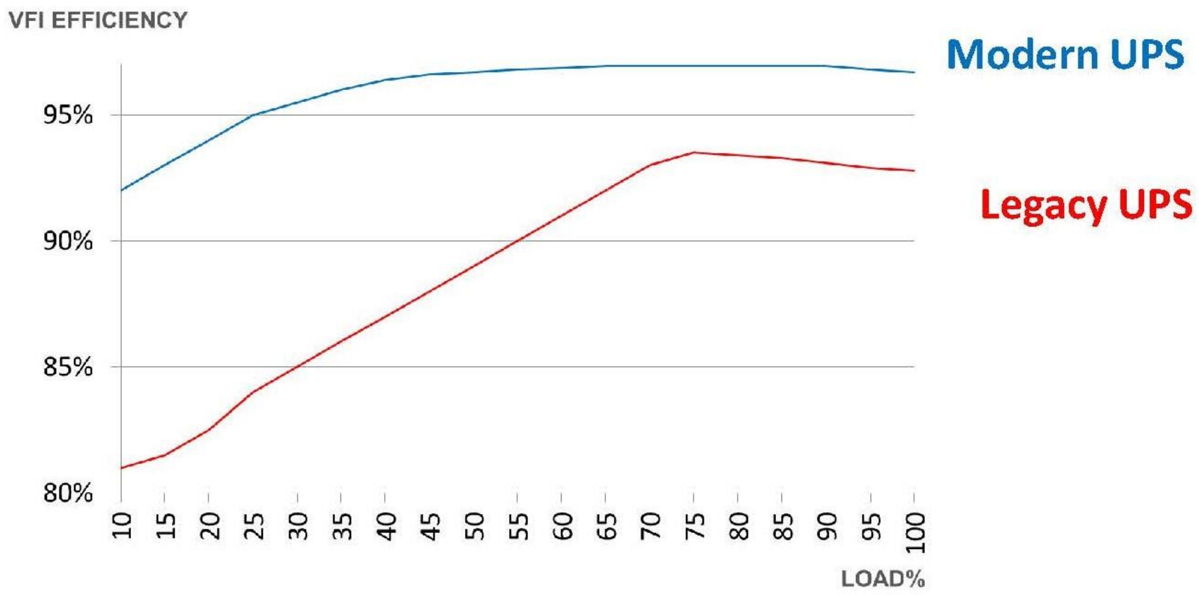

Figure 3: Power efficiency of a legacy UPS compared to that of a modern UPS

There is a trade-off between adaptive capability for the VFI mode and low failure rate. The internal complexity of UPSes has increased. The modern UPS uses a much more tightly coupled approach which means that its mean time between failures (MTBF) is lower. For most data centres, it is pretty common to run at load percentages between 25% to 50% most of the time, so the modern UPS is a considerable cost saver. In this case, efficiency usually takes priority over MTBF. The following factors should also be taken into account when selecting the UPS mode.

- Different modes of operation are available that offer additional protection features.
- Higher UPS power efficiency modes of operation imply lower protection.
- Choosing between modes of operation requires a thorough risk assessment.
- Significant aspects to be taken into account in the risk assessment are the characteristics of the load and its fluctuations and the stability of the main power grid feed.

2.2.3.3 Line-interactive Mode (VI)

This mode offers a power efficiency of 97-98%. It, of course, protects against a sudden power outage by invoking a battery immediately whenever there is a micro-interruption in the feed of the grid. When there is a normal feed from the grid, most of the load is fed directly via a bypass. The UPS is not idle, however. Part of the load is fed through the UPS that is now acting as a voltage stabiliser and power factor corrector. Figure 4 illustrates this mode schematically.

VI mode does not protect against rapid frequency swings and other types of noise that might occur on the feed from the grid. It may also negatively affect battery life. Whether VI mode is useful depends in large part on the stability of the grid. If the grid is too unstable, the mode effectively changes too often to double conversion.

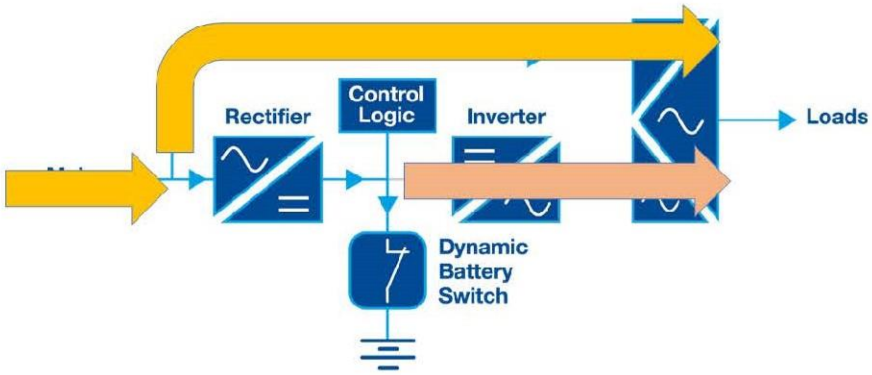

Figure 4: UPS line-interactive mode, with the UPS acting as a contributing interactive stabiliser rather than as a regenerator

2.2.3.4 Eco-mode (VFD)

This mode offers a power efficiency of 98-99%. It provides protection against a power outage, including micro interruptions. It does not offer voltage stabilisation or power factor correction. When there is a feed from the grid, it is fed to the load completely and unmodified via a bypass. The modules of the UPS are on hot standby, of course, but otherwise, they are idle. This differs from VI mode in that there is no impact on battery life. However, as with VI mode, it has no benefits if the grid is too unstable. This mode of operation is illustrated schematically in Figure 5.

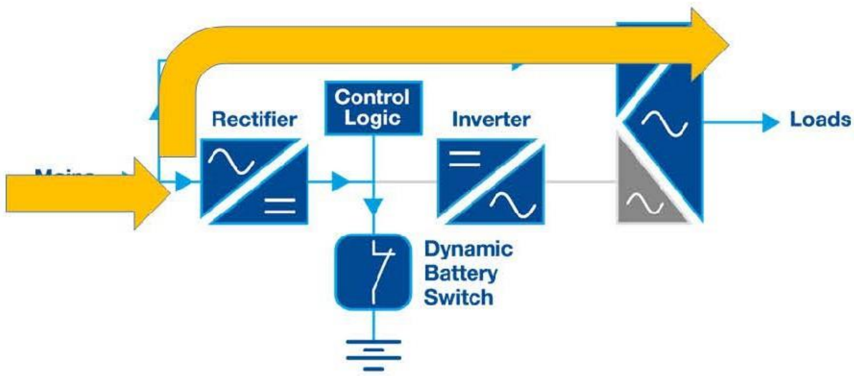

Figure 5: UPS eco mode, with the UPS just on hot standby when there is a feed from the grid.

2.2.3.5 Offline Mode (VFD)

Some manufacturers, including Borri, offer yet another mode in which the UPS is offline or in a “cold standby” mode. It is a variant of eco mode (VFD). In this mode, the power efficiency is 99-99.5%. The UPS awakens at intervals (every 14 days) to check and reload the battery. It also wakes if there is a long enough gap in the feed from the grid. So, the UPS does not protect against micro-interruptions. This implies that there are virtually no losses, but also that the invocation of the path through the UPS takes time, viz. the time of the gap plus the start-up time of the rectifier and inverter. The total delay is in the order of 10 milliseconds. Whether this delay is acceptable depends on the compatibility of the IT load with this behaviour. So, this mode may or may not offer protection against a power outage from the grid, depending on the tolerance of the IT load for the delay. But it does not provide any other type of protection. Figure 6 illustrates the offline mode schematically.

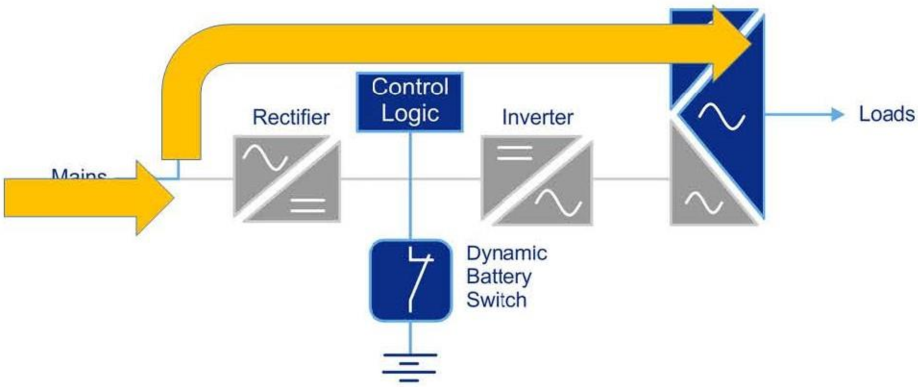

Figure 6: UPS offline mode, "cold standby" with a start-up delay.

2.2.3.6 UPS Power Efficiency of Monolithic Versus Modular

The topology of a UPS also has an effect on its efficiency. A modular UPS is generally slightly less efficient than a monolithic one. A modular UPS is, in fact, multiple UPSes operated in parallel. The more building blocks used in parallel, the smaller and less efficient the individual building blocks will be. Although the efficiency curve for the VFI mode has been flattened, a monolithic UPS still performs better relatively compared to a modular one when operated with loads at the lower ranges of its capacity (that is, up to 60%).

2.2.4 Scalability

Modular UPSes were developed mainly to overcome scalability problems with monolithic UPSes. With a Monolithic UPS, there is simply no way to expand the capacity of that very same UPS. Indeed, another monolithic UPS can be added to operate in parallel, but that cannot be done without disruptive maintenance to the data centre's electrical infrastructure. To accommodate growth scenarios for a data centre, both capital and operating expenditure are higher with a monolithic UPS. With the monolithic options, the data centre has to invest in a higher capacity than is effectively needed in the earlier stages of operation. Subsequently, the UPS will handle a load at the lower end of its capacity – which is less efficient - for an extended period of time.

Modular UPSes can be extended by adding power modules at any time, and the extension can be performed with the UPS in full operation. A modular UPS' capital and operating expenditure are lower whenever the power required at the start is less than 25% of the capacity required in the final stage of a growth scenario. Moreover, a modular UPS has the ability to implement N+1 redundancy of its power modules as a cost-effective feature.

For colocation data centres, modular UPSes are definitely the best choice. Whether they also are suitable for in-house and enterprise data centres is very site- and situation-specific. For example, if there is no scenario of foreseeable capacity growth, a monolithic UPS for any given capacity is substantially cheaper than a modular one.

2.2.5 Availability

Availability has several aspects. The availability of a single device is a function of its MTBF and its Mean Time To Repair (MTTR). The usual definition is expressed as follows:

$$Availablility = \frac{MTBF}{(MTBF\ +\ MTTR)}$$

This implies that complete availability is achieved if either the MTBF is infinite (the device never fails) or if the MTTR equals 0 (all repair is instantaneous). Neither infallible UPSes nor instantaneous maintenance exists. However, the capability to "hot repair" a UPS while it is in production and protecting the IT load is a quality that heightens the effective availability. Furthermore, a redundancy scheme with more than a single device and failover capability can be deployed to improve the effective availability of the service provided by a device.

On the one hand, the MTBF of a modular UPS is generally lower due to the increased complexity and number of components involved. On the other hand, while not all component failures in a modular UPS can be fixed with hot repairs, many can, especially if the failure is in a power module and the modular UPS was implemented with an N + m redundancy scheme of its power modules. In addition, the MTTR in modular UPSes is generally substantially lower than in monolithic ones. Table 2 lists MTBF and MTTR figures common for contemporary modular and monolithic UPSes. Thus, the dramatically lower repair times more than compensate for their lower MTBF.

	Modular UPS	Monolithic UPS
MTBF	190,000 h	385,000 h
MTTR	0.5 h	8.0 h
Availability	99.99974 %	99.99792 %

Table 2: MTBF and MTTR figures common for contemporary modular and monolithic UPSes

2.3 Cyber Threats - Welcome to a New World!

Stéphane Perez, CEA

The presentation by Stéphane Perez heightened the awareness of the audience to the multitude of cyber threats that exist nowadays. This was particularly important given that, at present, such threats are frequently not perceived well enough, and hence are not taken into account sufficiently.

HPC architectures, especially those in industry, can typically be represented by a multi-layer model such as the one presented in Figure 7. At the top level (or office report level), there is an enterprise network on which several systems exist that can have access – mediated, controlled, and logged by a firewall - to a network for Supervisory Control and Data Acquisition (SCADA) of the facility. Remote access from the outside world to the enterprise network itself is usually also enabled and is either controlled by a firewall or by the combination of a firewall and a virtual private network.

SCADA servers, upgrade and deployment control stations, etc., with an interface on the SCADA network are in turn connected to one or more networks to which remotely manageable components are connected at the Programmable Logic Controller (PLC) level.

In Figure 7, the following two activities or situations that deviate from the hierarchical structure (and therefore need special care and attention) are added to the layered model.

- Access to remotely manageable and configurable components at the programmable logic controller level by means of Wi-Fi connections
- Human actors having direct physical access to all sorts of equipment at various levels in the course of maintenance procedures

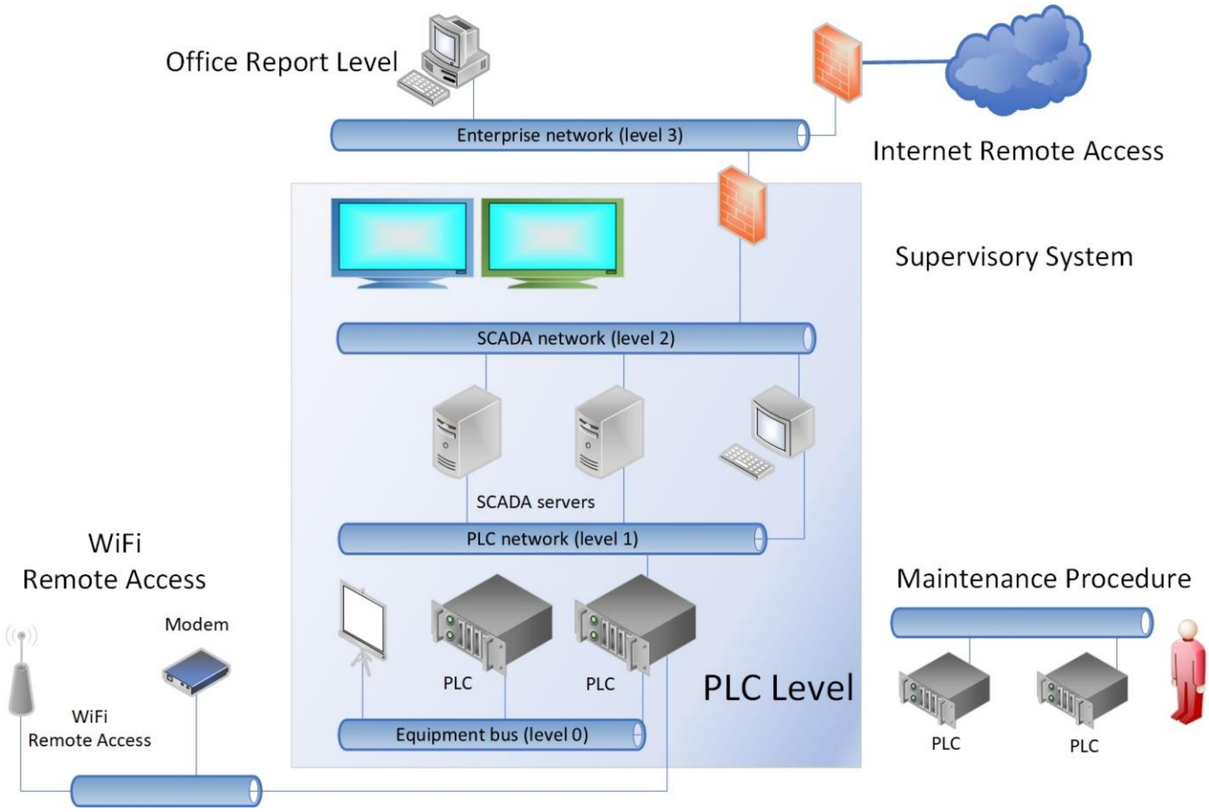

Figure 7: Multi-layer model of HPC architecture

Typically, HPC architectures are protected mainly, or even exclusively, at the top level. The actual equipment is in a data centre, or a computer room, into which only properly authenticated authorised personnel are allowed. Other than that, there are often no or few additional procedures and precautions in place that specifically address protection at the PLC and SCADA levels.

The perception is that the most severe threats come from the outside via the network. However, threats and infected or compromised components can exist on every level:

- office desktops compromised through downloading of malware,
- inadequate or incomplete firewall rules,
- infected USB keys,
- non-secure Wi-Fi applications,
- infected maintenance PCs,
- infected devices brought into the PLC and/or SCADA network by authorised personnel, possibly from hired third parties that move their own specialised equipment from one environment to another, and
- a compromised or insufficiently tested software supply chain.

Of course, there has to be adequate protection on the "front door" and "official main entrance" as well. But recent quantitative studies that categorise vulnerabilities convincingly demonstrate that this in itself is grossly insufficient. Most vulnerabilities are found to occur in fact on the SCADA and PLC level, with more than 50% in embedded web servers that provide monitoring and control interfaces to remotely configurable and manageable components. Moreover, there seems to be little progress in this area in that the total number of vulnerabilities discovered each year in these areas so far has not decreased but has remained constant.

The presentation contained a relatively large body of high impact incidents that convincingly exemplified that the above issues do occur and that the Achilles heel was the lack of a more in-depth organisation of defence. The following examples illustrate just a few of these attacks.

- In 2008 an attack occurred on the BP Baku-Tbilisi-Ceyhan pipeline. An attack through non-secure Wi-Fi networks was combined with a physical attack. Surveillance cameras and security alarms could be disconnected, and a lot of equipment was destroyed.
- In 2008 a nuclear plant in Georgia had to do an emergency shutdown after a software upgrade. Technically this was not a security threat with a malevolent attacker, but the problems were clearly caused by insufficient control and testing in the software supply chain.
- In 2010 "Stuxnet" (a computer worm or malware program) sabotaged Iranian nuclear facilities. The hack was very sophisticated and exploited unknown "zero-day" vulnerabilities in Windows and in Siemens Step 7 software. "Stuxnet" entered the target environment on infected USB keys and could spread because of a lack of controlled separation of networks.

The above list can be expanded. In 2021, in France, there has been an increase of so-called ransomware attacks that were successful from the point of view of the attackers, mainly because there was an insufficient in-depth organisation of defence.

An additional problem is that the hacker community has become very diverse, and so have the reasons for becoming a target. Hackers may be individuals, "lone wolves" or "script kiddies", but increasingly they are organised collectives with considerable expertise and resources at their disposal. They might be interested in abusing your resources for various purposes, but the motive may also be espionage – getting access to essential data and programs residing on your systems. Hackers can be hired by interested parties or directed and supported by state actors with access to specialised expertise and intelligence. The goal could also be to obstruct, delay, or sabotage particular research that is being conducted. The ransomware attacks demonstrate that any way to effectively obstruct primary production processes will do for some hackers, as long as the victim is willing to pay to alleviate the obstruction.

Perez argues that organisations should adopt strict protocols for USB keys that enter their systems operations as well as their office reporting environments. A study, conducted on campus by the University of Illinois in 2015 with 300 USB keys "dropped" on campus, showed that a large number of the keys were plugged in by people who found them to inspect what was on them. This suggests that the risk awareness of many individuals in this respect is dangerously low.

Organisations should have cheap test devices, such as Raspberry Pi systems, and make it mandatory to test every USB key before it is accepted into any segment of the organisation's environment. On the test device, the required data must be transferred to an already trusted USB device belonging to the company. The external USB key should never get past the test device.

In a similar vein, outsourcing management and maintenance tasks to third parties must be accompanied by a protocol that ensures that portable devices used by the maintenance operator do not travel from one environment to another. The best procedure is to have dedicated devices for particular types of maintenance that are handed out to operators and then collected after the maintenance and thus never leave the company.

Another recommendation is to avoid applying patches without prior testing in a site-specific test environment. This measure protects against bad patches in general and reduces the risk of falling victim to a compromised software supply chain.

Perez also recommends that companies thoroughly organise the segregation of networks at the lower levels of their architecture. A security operations centre with resources for detecting non-authorised computers, detecting IP spoofing, and performing general log file analysis is also highly recommended. Several open-source software initiatives provide a framework for analysing networked environments. As a member of CEA, Perez recommends the IVRE analysis toolkit for this purpose. IVRE relies on other well-known open-source tools, such as NMAP. It is free and available on GitHub.

3 Session II – Chaired by Michael Ott (LRZ)

3.1 Facility Power Management for the Upcoming Perlmutter HPC System

Melissa Romanus, Abdelbaky, NERSC

The presentation began with an introduction to the National Energy Research Scientific Computing Center (NERSC) as an institution. NERSC is an HPC centre in the USA that offers its computing services to a wide spectrum of users, which includes thousands of researchers with multiple codes.

NERSC has been planning to expand its computing power by introducing a new machine, Perlmutter, that will offer many times more computing power than the current system. The new system imposes new requirements on the data centre facility. Therefore, the first stage of preparing for the new machine is to prepare the supporting infrastructure. The most basic requirement is to ensure that enough electricity is available and can be delivered efficiently to the new system.

The data centre used by NERSC is a compressor-free facility that does not use chillers. Hence the global Power Usage Effectiveness (PUE) for the whole facility is 1.08. However, having a compressor-free facility, without any air-conditioning units, means that with global warming, there is a notable decrease in the efficiency of the infrastructure as the outdoor temperature increases.

The current power cap is 12.5 MW and is being increased to 25 MW to be able to host the new supercomputer. Perlmutter will be a Cray Shasta supercomputer, so it will be an almost 100% Direct Liquid Cooling (DLC) cooled system consisting of a significant CPU-only partition based on AMD Milan CPUs supplemented by 6,000 NVIDIA A100 accelerators. In addition to the extensive HPC system, dedicated storage will be installed, namely 35 PB of Lustre storage with 5+ TB/s bandwidth. All components of the machine will be connected using Slingshot fabric.

Usually, when installing a new cluster, one expects it to have a larger footprint and offer more compute capabilities. While the latter is true for Perlmutter, the footprint of the new machine is smaller than that of the old one due to the increased power density of the new system. It is estimated that the CPU-only cabinets will consume 400 kW while the cabinets equipped with GPUs will consume 300 kW. The estimated power draw of the entire system under normal load is around 6 MW.

Due to seismic activity in the area where the data centre is located, both power and cooling installations must be connected to the building via flexible connectors. Otherwise, the movement of the raised floor might damage the infrastructure.

Because of the expansion of the power supply system, a series of tests were conducted with the goal of detecting any potential issues before the real, expensive system is delivered and installed. This task proved to be a complex one as several hundreds of circuits needed to be tested. Infra-red imaging was done while testing to figure out potential hot spots.

The validation and test procedures covered not only the wires, fuses and other physical infrastructure components, but also the validation of the metering and monitoring system. The monitoring system collects data every few seconds because of the slow Modbus interface. Then data is gathered via open-source message-broker software RabbitMQ and sent to what is called the Operations Monitoring and Notification Infrastructure (OMNI) for collecting and managing data. Recordings are stored in Elasticsearch to provide easy and flexible access to the data, as well as facilitating analysis and visualisation using external tools such as Kibana, Grafana or other modules written in Python, Go, and so forth. Due to the size and complexity of the power infrastructure, 90,000 metrics per second are being collected from 6,000 sensors in the data centre.

In order to emulate real-world loads, the tests were performed using 12 load banks (model Mosebach X100LR). All crucial components were tested: compute racks connections, cooling distribution units, and auxiliary equipment racks. The analysis and failure detection were performed online in semi real-time using Grafana dashboards. Data from the meters was correlated with the settings on the load banks.

A direct result of the testing procedure was the detection of a series of issues that might cause problems for the production system. The main issues were: bad circuits, bad fuses reporting 0 from phase, incorrect wiring, and misreadings on some of the meters. Testing with load banks revealed a bad bus kit and resulted in a melted bus kit on the main breaker. A side effect of the tests was the optimisation of the collection of the data from the meters. There was a lot of missed data in the beginning, so the test procedure also resulted in software fixes. Infrared imaging revealed potential hot spots for further investigation.

The test did not verify whether the infrastructure can handle rapid ramp-up of computing power to full load. During the procedure, the power draw was increased in a steady, gradual way.

Even though it was time- and resource-consuming, the test procedure was extremely useful and revealed a series of issues that, if unfixed, might have escalated into big problems if the supercomputer had been installed before fixing the problems.

3.2 The Wild Ride To the First Production Cray EX 3000

Jim Rogers, Oak Ridge National Laboratory

Oak Ridge National Laboratory (ORNL) signed a partnership with the US Air Force Weather (AFW) enterprise for hosting a computing system. Weather forecasting systems, unlike most HPC systems, have strict reliability and availability requirements, as both global and regional weather forecasts must be delivered on a timely basis. This leads to different requirements when designing such systems. Jim Rogers presented key tenets of the ORNL design for the AFW system.

3.2.1 Infrastructure requirements for reliability

3.2.1.1 Electrical Reliability

Electrical reliability at ORNL is achieved via redundancy and diversity. The total demand is approximately 2 MW. UPS systems support only a small fraction of the mission-critical systems, and those are mostly storage systems. All critical systems are dual-fed, with only one side protected by UPSes. The ORNL campus is powered by 161 kV transmission lines from three physically diverse routes. The power is delivered from multiple large fuel sources, including nuclear, hydro/dam, natural gas, coal, wind, and solar power. Thanks to this diversity, ORNL has a highly reliable power supply and can deliver electricity to the campus facilities for about 0.11 $USD per kW/hour.

The electrical system is designed to be redundant on several levels as follows.

- High voltage 161 kV lines: There are three connections; a drop on the power source (power plant) results in voltage sag, but no interruption as all three 161 kV lines end up on a single bus at the substation.
- Internal for the facility, medium voltage: If one of the 13.8 kV feeders dies, one of the HPC systems may be impacted, but the other will continue operations.
- Compute service level: Mission-critical calculations are performed on two systems simultaneously to ensure results are on time (e.g., global weather modelling).

The redundancy of the 161 kV infrastructure was inadvertently tested in January 2021 when a 161 kV line was destroyed in a vehicle accident that destroyed a transmission tower, so the line to the Bull Run coal power plant was damaged. As predicted, a voltage sag was recorded – but it was within acceptable boundaries – so there were no issues for the ORNL campus as a result of the accident.

3.2.1.2 Network Connections

The design of the ORNL network connections follows similar principles to those for its power infrastructure.

- The Wide Area Network (WAN) is based on multiple redundant paths via different providers to ensure stable and robust connections to all systems.
- Local area network is high speed and redundant so high loads and failures are handled. All systems are installed within an enclave that is compliant with the US Federal Information Processing Standard (FIPS) Publication 199 at the "moderate" level for both integrity and confidentiality.

3.2.1.3 Infrastructure

The ORNL facility has a single large, medium temperature cooling loop (20 °C) with a total flow of around 19,000 l/min, which uses cooling towers to dissipate the heat. The annual PUE of the data centre is 1.05.

The ORNL facility offers a virtual tour with detailed information about the operations and infrastructure. It can be accessed here: https://my.matterport.com/show/?m=iBfbj7ET4LT.

3.2.1.4 Specifications of the AFW System

The new AFW computer is a CRAY EX 3000 system consisting of two computers in a redundant setup. Each computer has four compute cabinets, and 800 nodes. In total, there are 50 compute blades; each blade hosts dual compute nodes, each with 2 × AMD Rome 2.2 GHz microprocessors with 256 GB of memory. Even if the machine is 100% water-cooled, one must assume that 3-5% of the heat will radiate to the room and be dealt with using traditional air-conditioning.

One of the immediate lessons learned during the preparation phase was that the Machine Unit Specification (MUS) usually lists much higher power consumption figures than the actual consumption. Therefore, one has to take these values with a grain of salt in order not to overinvest in expensive power delivery infrastructure. One should pay attention to how cooling and power equipment and empty space are aligned with each other to avoid stranded capacity.

For the AFW system, the estimated power capacity, based on MUS specifications, is 592 kW, however the actual values differed significantly from the estimated number. For the High-Performance LINPACK (HPL) test, the power consumed was close to 500 kW, which was an extreme case. While executing performance and scaling tests, power consumption peaked at 450 kW, and during the stability tests, power consumption was even lower at 400 kW. The values provided by MUS are over the worst case (HPL) by 20%, 30% above the pre-selected scaling tests, and almost 50% over the real-life usage pattern.

The machine's power factor depended on the amount of power consumed and varied from 0.993 for the HPL test to 0.98 for real-life workloads.

3.2.1.5 Satisfying Certificate of Readiness Requirements

The project experienced various challenges related to meeting the Certificate of Readiness for AFW on schedule.

3.2.1.6 Hardware Installation

Hardware installation was a relatively easy task and went smoothly; software stack readiness was the biggest obstacle. After three months of struggle with the Kubernetes-based Shasta software stack, the team decided to abandon this stack. The core services (for booting the system) were implemented using traditional Linux mechanisms on the external services cluster. Finally, the HPE Performance Cluster Manager (HPCM) was used as it proved to be a much more stable solution along with the Slingshot fabric manager.

3.2.1.7 COVID-related Issues

The maturity of all components is crucial for stable and controlled operations. The machine was delivered in October 2019 and in early 2020, but the monitoring software was not ready. Cray filled both systems with pre-production water chemistry, but the telemetry interface was not ready. Due to the COVID pandemic, all non-essential personnel were sent home, so the progress of software development was stalled. The system was running without remote supervision. After a few months, it turned out that the temperatures were higher than expected, and the flows were not at the required levels. The source of the anomaly was biological contamination of the OEM-provided coolant, and the steady degradation of the cooling loop parameters went unnoticed because the system was not monitored remotely, and someone had to be on-site to figure out that something was wrong with the machine.

This resulted in a week-long clean-flush cycle to make sure it was cleaned. The long-term water chemistry state needs to be monitored and is currently being monitored constantly.

As with most HPC systems, this system produces a mass of data that humans cannot efficiently process. A machine learning-based mechanism is needed to detect and predict failures; there are, however, no off-the-shelf solutions. Solutions need to be developed in-house.

3.3 AIOPS – Using AI to Improve Operations

Avi Purkayastha, NREL (National Renewable Energy Laboratory)

The presentation began with a brief introduction to Artificial Intelligence for IT Operations (and the role of the National Renewable Energy Laboratory (NREL) in the AIOps project.

AIOps is a 3-year collaboration project between NREL and HPE with the goal of leveraging Artificial Intelligence and Machine Learning (AI/ML) and real-time streaming data to improve data centre operational efficiency through enhanced monitoring, prognostics and anomaly detection.

The NREL efforts focus on data gathering, analysis and deployment, while HPE focuses on anomaly detection. As a result of the collaboration, a streaming data platform for gathering and storing metrics was created (see Figure 8).

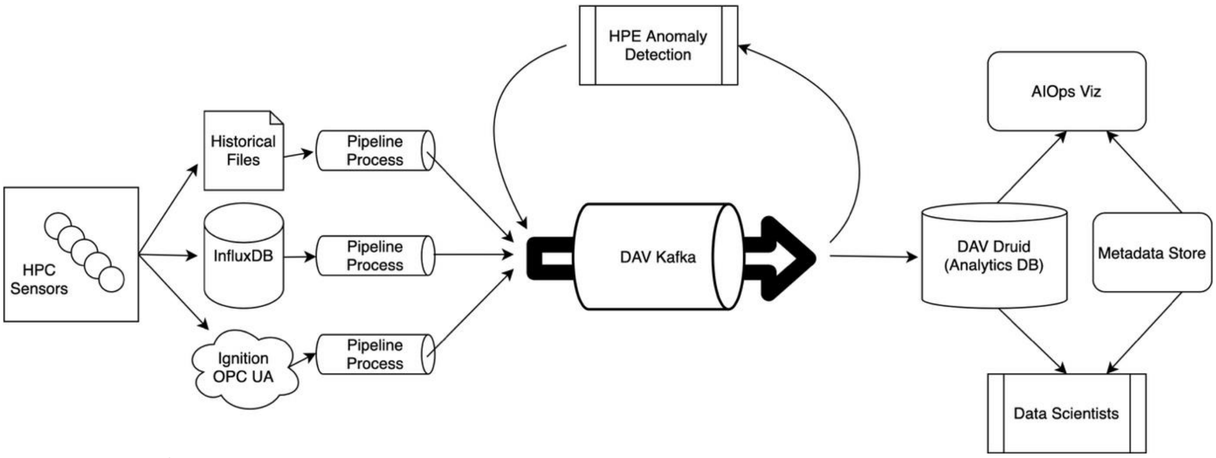

Figure 8: Data platform components

The next part of the presentation focused on a description of the data centre where all research and development was done.

The facility hosts a mixture of Direct Liquid Cooled (DLC) systems and air-cooled systems with hot aisle containment.

Both the facility and the supercomputer infrastructure are instrumented. The monitoring system collects about 4 thousand data points per minute for the facility, while the supercomputer monitoring system collects 1 million data points per minute. The collected metrics include cluster power, cooling data, job metadata, and power readings from meters (including intelligent breakers). In addition, building, cooling tower and thermosyphon data are also gathered at intervals of one minute.

A dashboard presenting current readings from the data centre can be visited here: https://www.nrel.gov/hpc/cool.html.

Currently, the monitoring system shows PUE and Energy Reuse Effectiveness (ERE). While these two metrics are useful on their own, the goal of the collaboration activity is, based on PEU and ERE models, to develop an AI-based prediction model from historical data to look at possible usage or utility patterns in future.

Another application that the team is working on is an Energy Balance Model (Figure 9) that shows how the energy consumed by the data centre is split between different equipment groups, how much is dissipated to the air and how much is captured by DLC.

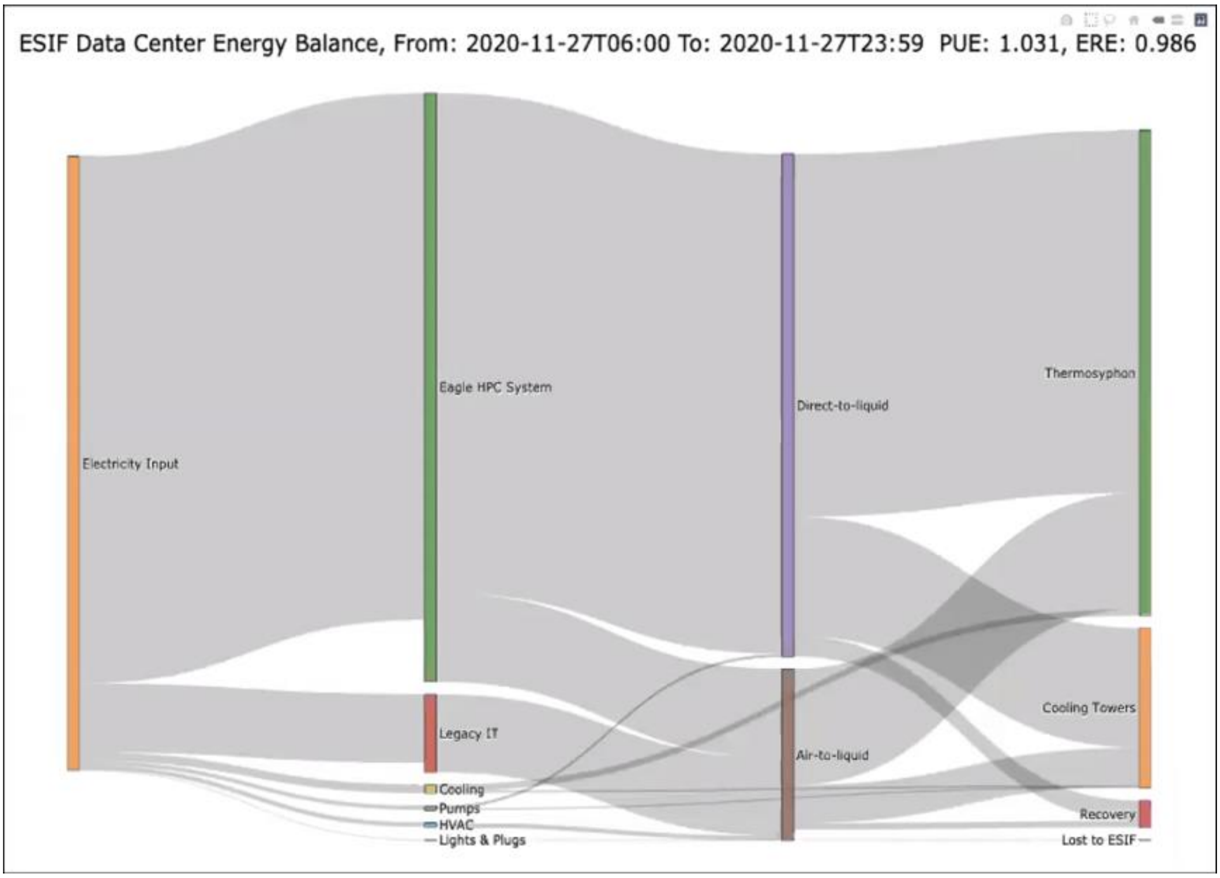

Figure 9: Energy balance model of ESIF data centre

This is an early warning system for energy leaks in the data centre. It also helps with optimisation procedures because it can show how the split changes with environmental changes, such as coolant temperature.

Predictive cooling is a part of the project that relies on the availability of power profiles for HPC jobs that run on the HPC system. The profiles, along with statistical data about power characteristics of the applications, makes it possible to predict the required cooling capacity for a given job and, consequently, adjust the cooling system settings to the optimal setting before the job is launched. This way, it is possible to, e.g., speed up the fans or provide additional water to the cooling towers the moment the job enters execution to avoid rapid handling of the power spikes that would occur if the cooling system just reacted to the coolant temperature. This makes it possible to prevent rapid changes in cooling system settings and thus saves energy.

4 Session III: EuroHPC Petascale Systems – Chaired by Evangelos Floros (EuroHPC)

4.1 Update on EuroHPC

Evangelos Floros, EuroHPC JU

The European High-Performance Computing (EuroHPC) Joint Undertaking is a legal and funding agency created by the European Commission in 2018 with the aim of fostering the development of HPC technologies and supporting infrastructures in Europe. This mission is implemented via three main activities/pillars: HPC ecosystems, Infrastructure and Operations, and Research and Innovation (R&I).

EuroHPC is a consortium of 33 member states, along with two private members – the European Technology Platform for High Performance Computing (ETP4HPC) and the Big Data Value Association (BDVA) – with a budget of 1.5 billion euro. The EuroHPC Joint Undertaking (EuroHPC JU) became autonomous in September 2020, and its headquarters in Luxembourg were inaugurated in May 2021.

EuroHPC JU has selected the following eight entities for hosting its supercomputers.

- Three pre-exascale systems, each above 150 petaflops, will be located in Finland (CSC), Italy (CINECA), and Spain (BSC).
- Five petascale systems will be located in Portugal (University of Minho), Luxembourg (LuxProvide), the Czech Republic (IT4I), Slovenia (IZUM), and Bulgaria (Sofia Tech Park).

At the time of writing, seven contracts had already been signed and were under implementation. The contract for the MareNostrum 5 system to be deployed at BSC was still pending.

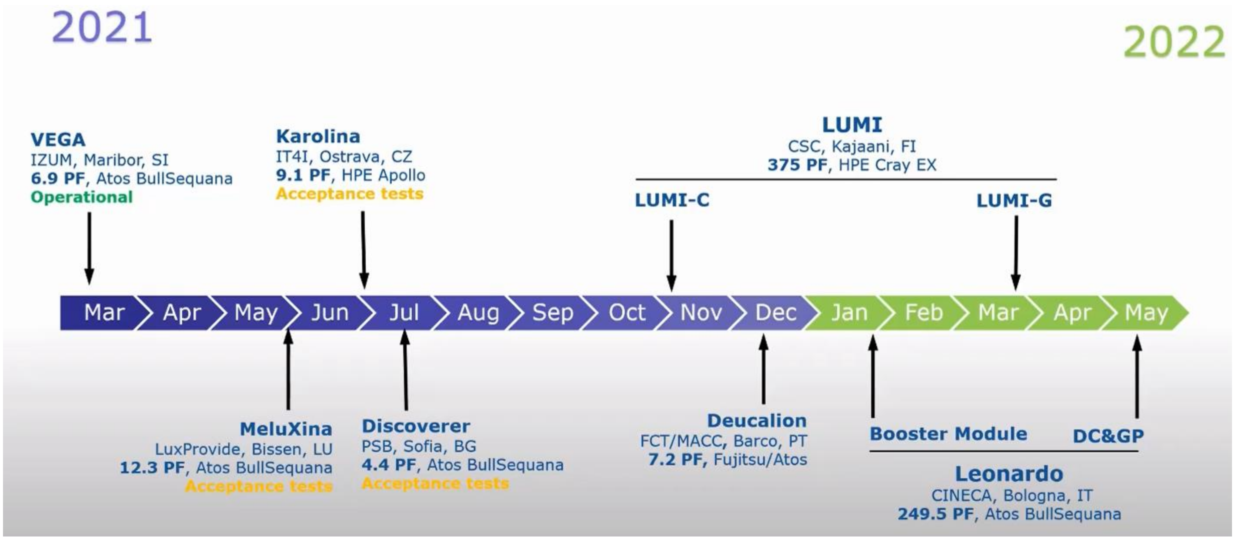

Figure 10: Deployment timeline of EuroHPC systems

Figure 10 shows the availability timeline of the EuroHPC systems, i.e., the dates when it is estimated that the acceptance process for the systems will be finished. As can be seen, the petascale systems are the first in the queue, with the VEGA system already being operational. The last of the planned petascale systems will be Deucalion, located in Portugal, with an expected performance of 7.2 petaflops and built by Fujitsu and Atos. All the compute performance numbers indicated in Figure 10 represent sustained performance values.

As can be seen, the deployment phases of the two large pre-exascale systems are split into two phases. For LUMI (hosted by CSC, Finland), the first (CPU-only) partition of the system is scheduled to arrive at the beginning of November 2021, while the deployment of the GPU partition is foreseen for the beginning of March 2022. The LUMI system is being built by HPE Cray, and it is estimated that it will deliver a performance of 375 petaflops.

The next pre-exascale system, Leonardo, which will be deployed at CINECA (Bologna, Italy), is expected to be completely operational towards May 2022 and to deliver a computational performance of 249.5 petaflops.

One of the recent major milestones for EuroHPC was the adoption in March 2021 of an access policy, which defines several access modes for offering resources on a periodic and continuously open call basis. More specifically, these access modes are as follows.

- **Extreme-scale**: mainly focusing on pre-exascale systems and offering a large amount of resources (similar to PRACE Tier-0 calls)
- **Regular**: intended to be used by medium to large-scale applications, utilising the EuroHPC petascale systems (a new type of call)
- **Development**: limited access (up to 1 year) across all EuroHPC systems (similar to PRACE Type B calls)
- **Benchmark**: limited access (up to 3 months) across all EuroHPC systems (similar to PRACE Type A calls)
- **Fast track for industry & academia**: aims to provide quick access to the EuroHPC infrastructure for stable applications requiring further computational time (a new type of call)

EuroHPC will rely on the existing knowledge, experience, and know-how of PRACE for the effective and efficient implementation of this access policy. These types of calls are already open on the PRACE website: https://prace-ri.eu/hpc-access/eurohpc-access/eurohpc-ju-benchmark-and-development-access-calls/; however, currently only the VEGA system is being offered due to the previously mentioned deployment timelines (see Figure 10). Because of this timeline, no extreme scale or fast track calls will be available during 2021.

From 2019 to 2021, EuroHPC has successfully launched several research and innovation calls, aiming to boost HPC uptake and further support research and development, as well as the use of HPC technologies, across Europe. So far, four calls have been launched, closed and evaluated, which provided grants to 26 projects that are being implemented at the time of writing.

The new JU regulation, which is planned to be adopted by the EC during the summer of 2021, aims to extend the mandate of EuroHPC to ensure a world-leading federated, secure, hyper-connected supercomputing and quantum computing service and data infrastructure ecosystem in Europe.

4.2 Supercomputer MeluXina Hosted in a Green Data Centre

Roger Lambach, LuxProvide

LuxConnect is a hosting entity/infrastructure provider that does not offer any IT services. It was founded in 2006 by the Luxembourg government and is 100% owned by the Luxembourg state. However, the company acts, and is regulated, as a private company. LuxConnect currently has 25 employees and had a turnover of 27 million euro in 2020.

LuxConnect has deployed 1,500 km of dark fibre cables (up till now) while not offering any active services. LuxConnect's customers in that domain, namely telecom operators, are renting cables from LuxConnect and putting in their own active equipment and offering services to end customers. It should be noted that not all parties are eligible to get this type of service from LuxConnect (e.g. banks are not), as LuxConnect has no intention of interfering in the market of the telecom operators.

At the time of writing, LuxConnect operates four multi-tier data centres in Luxembourg: three located in Bettembourg which offer 10,000 m^2 of IT floor space; and one in Bissen with

5,000 m^2 of IT floor space. Bissen is the data centre where the EuroHPC MeluXina system is installed. All these sites are interconnected by different fibre routes as this allows for redundancy, i.e. each location can be controlled from another. (The optical cable length between these two areas is 35 km). Customers are offered a further redundancy by being allowed to install their equipment at more than one site.

Figure 11: ICT campus Bettembourg

Figure 11 shows the ICT Campus in Bettembourg, with three data centres (i.e. DC1.1, DC1.2, and DC1.3) and the office building in front of DC1.2. DC1.1 was the first to be completed, and it opened in June 2009 and is a Tier IV data centre. DC1.2 has been operational since 2011 and will house the tape library of the MeluXina system. DC1.3 is the latest data centre and was opened in 2015. DC1.3 is double the size of the other two data centres; however, only half of it (the left part) has been fully built at this stage. In principle, DC1.3 makes it possible to have two data centres that are completely separated internally. Given the ongoing rapid changes in technology, particularly in the area of the cooling infrastructures supporting data centres (e.g. higher inlet temperatures are becoming acceptable to many hardware vendors), it was decided to wait before completing the second half of the DC1.3 building, so that any necessary adaptations of the supporting infrastructure could be executed effectively and efficiently.

LuxConnect had decided to ensure 100% green electricity usage from its first day, with the power mainly coming from hydroelectric sources. LuxConnect uses free cooling as long as outside temperatures are below 12 °C. However, LuxConnect is constantly working to raise this temperature threshold. They adapt the cooling infrastructure to the new temperature ranges allowed by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE)[a]. LuxConnect is working continuously to improve energy efficiency, supervised by the government agency MyEnergy, and the company needs to demonstrate a reduction of energy consumption every year without compromising any service level agreements (SLAs) on the Tier IV level. The latest building constructed in Bettembourg, DC1.3, has a Power Usage Effectiveness (PUE) indicator of 1.3 (DC1.1 has a PUE of 1.6, DC1.2 has a PUE of 1.45, and the building at Bissen was designed to have a PUE of 1.35). Furthermore, the wasted heat from the servers is reused to heat office spaces and diesel generators.

DC2 (Figure 12), where MeluXina is located, and DC1.3 are Tier IV data centres as certified by the Uptime Institute.

---

[a] https://www.ashrae.org/

Figure 12: LuxConnect's green data centre DC2 in Bissen

As can be seen from Figure 12, on the left side, there is space for expansion. Two extensions are already being planned. The first one is an office building, which would also be used for training. The other extension is expected to be used to enlarge the existing IT equipment floor space to accommodate a larger number of high-end IT systems.

On the right side of Figure 12, the building labelled Kiowatt is a part of the Kiowatt installation, which supplies DC2 with heat converted into the required cooling for the data centre. However, it should be noted that, since DC2 is a Tier IV data centre, it is independent of this installation, meaning that in case of maintenance of the Kiowatt plant, DC2 is still capable of supplying the required cooling.

Figure 13 illustrates the process of generating cooling from heat at Kiowatt. As can be seen, waste wood is burned in a steam boiler to produce steam. This steam is injected into a steam turbine, and the electricity that comes out is fed into the public grid as green energy.

Most of the rest of the heat that results from this process is used for cooling the data centre. The remainder of the heat, as can be seen from Figure 13, is used for producing BADGER wood pellets (for heating), as well as for district heating during the wintertime. This, in its turn, means the data centre can rely on free cooling during the wintertime, resulting in a PUE for DC2 of between 1.2 to 1.25, which is very low for a Tier IV data centre.

LuxProvide is the operator of HPC, which takes care of all IT services. It was founded in June 2019 as a daughter company of LuxConnect (100% subsidiary of LuxConnect) and is headquartered in Bissen. Currently, it has 15 employees. The mission of LuxProvide is to operate the leading-edge national supercomputing and data infrastructure, provide HPC, HPDA, Big Data, as well as AI commercial services, and in general to empower the Luxembourg digital ecosystem.

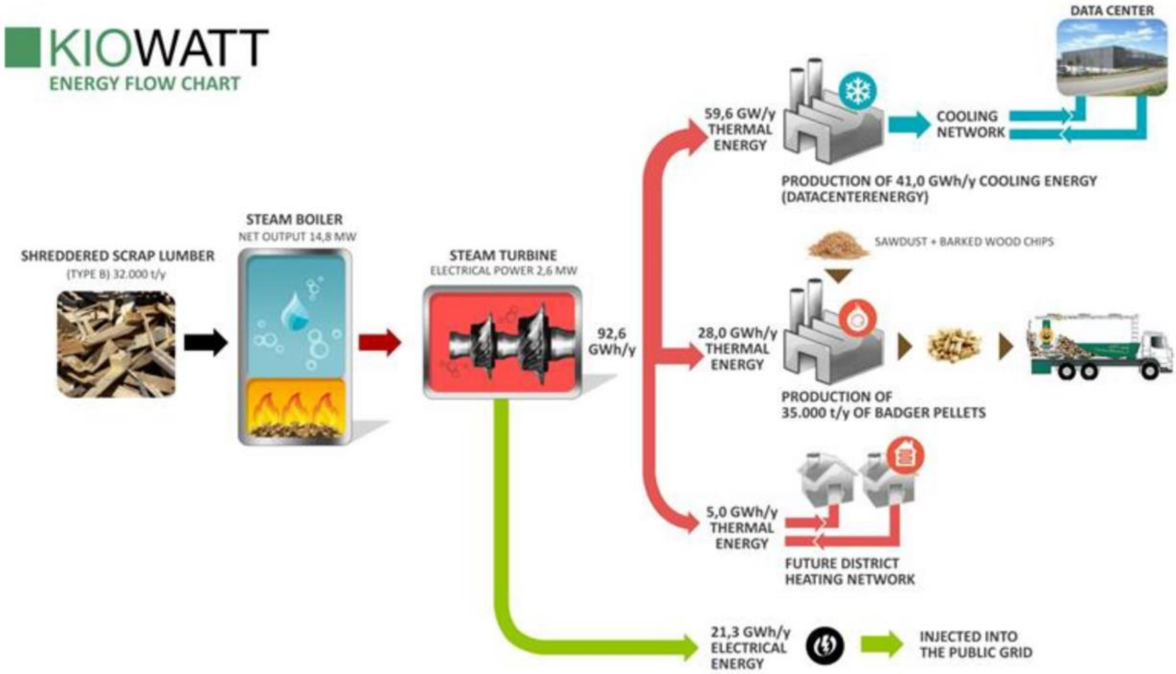

Figure 13: Kiowatt: generating cooling energy for data centres

The EuroHPC MeluXina system, operated by LuxProvide, is a 10.5 petaflops machine[b] featuring the latest AMD EPYC processors and NVIDIA A100 GPUs[c]. The system was built by Atos and had a total acquisition cost of 30.4 million EUR. The expected lifetime of the MeluXina system is five years. MeluXina will serve a large variety of complex, data-driven computational workloads.

Given that MeluXina is a EuroHPC machine, 35% of its capacity will serve EuroHPC projects, with the remaining 65% dedicated to industry, research and public administration.

Figure 14 shows the architecture of the MeluXina modular system. The main cluster module, a CPU-only module, is connected via an InfiniBand 200 Gb/s interconnect to the storage module, large memory module, and the accelerator module. This accelerator module features two types of hybrid nodes: 200 CPU-GPU hybrid nodes and 20 CPU-FPGA hybrid nodes. The large memory module has a total capacity of 80 TB RAM with 20 large memory nodes to support big data applications.

The MeluXina system also features a system module with 2N customer access nodes and 2N service nodes for enabling user access to the system. As can be seen from Figure 14, all these modules are connected to the cloud module, which is mainly used for user portals. The cloud module is connected to the Ethernet network. However, it is also connected to the InfiniBand network for redundancy.

Figure 15 provides a performance overview regarding MeluXina's compute, storage, and interconnect installations. As can be seen, the main storage is divided into four tiers:

- Tier-0 with lower capacity, but with higher speed;
- Tier-1 is the largest, with a storage capacity of 12.5 PB intended for general-purpose usage;
- Tier-2 is intended for backup with 7 PB storage capacity; and

[b] High-Performance LINPACK (HPL) performance
[c] According to recent June 2021 Green500 rankings, the MeluXina system has been named as the greenest supercomputer in the European Union and the fourth greenest in the world.

- Tier-3 is the tape library located in Bettembourg, which is intended for long-term archiving, and currently has a capacity of 5 PB. However, it can be enlarged based on needs.

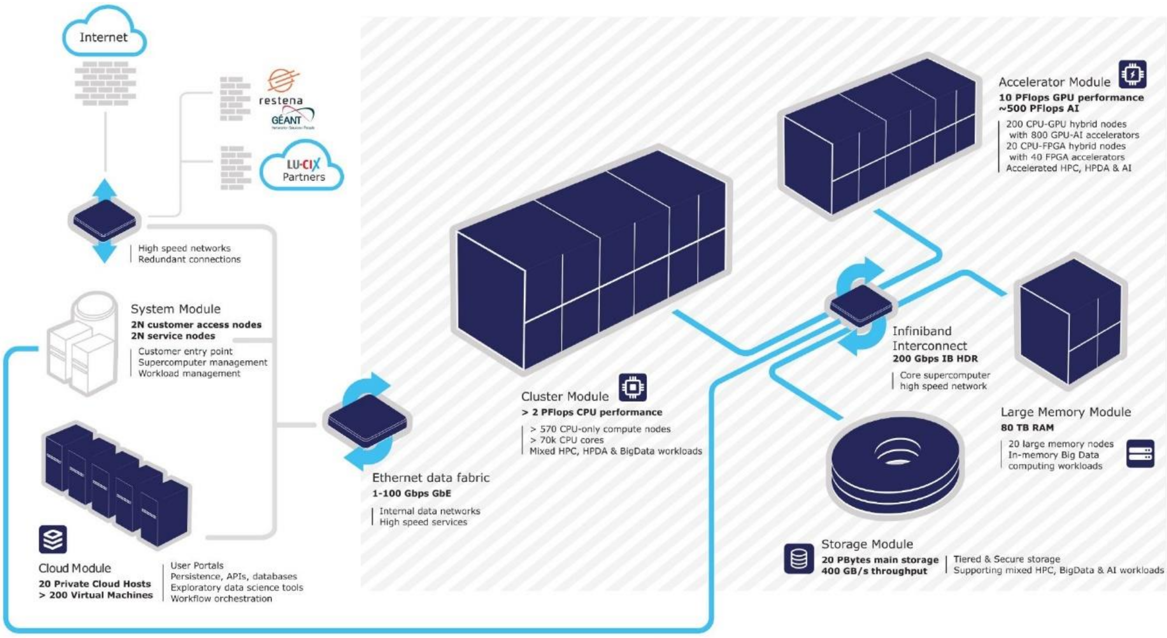

Figure 14: The architecture of the MeluXina system

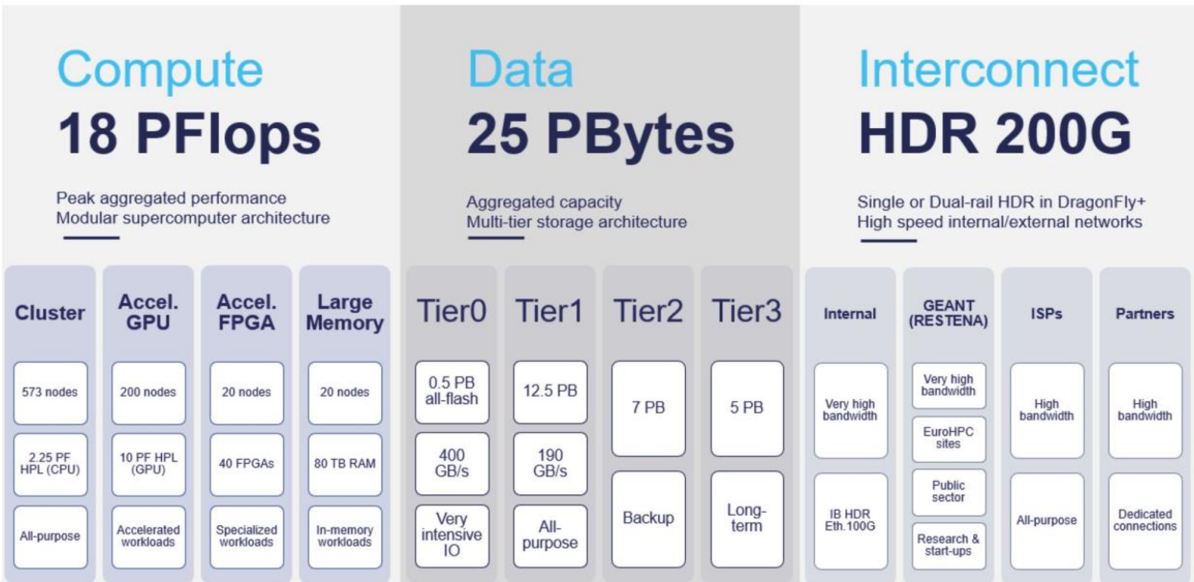

Figure 15: An overview of MeluXina's compute, storage, and interconnect performance

4.3 Vega

Dejan Valh, IZUM

Vega, built by Atos, is the first operational EuroHPC petascale system. It is hosted at the Institute of Information Science (IZUM), Maribor, Slovenia. IZUM is a public institution with approximately 120 employees.

Figure 16 provides a brief performance overview for the Vega system.

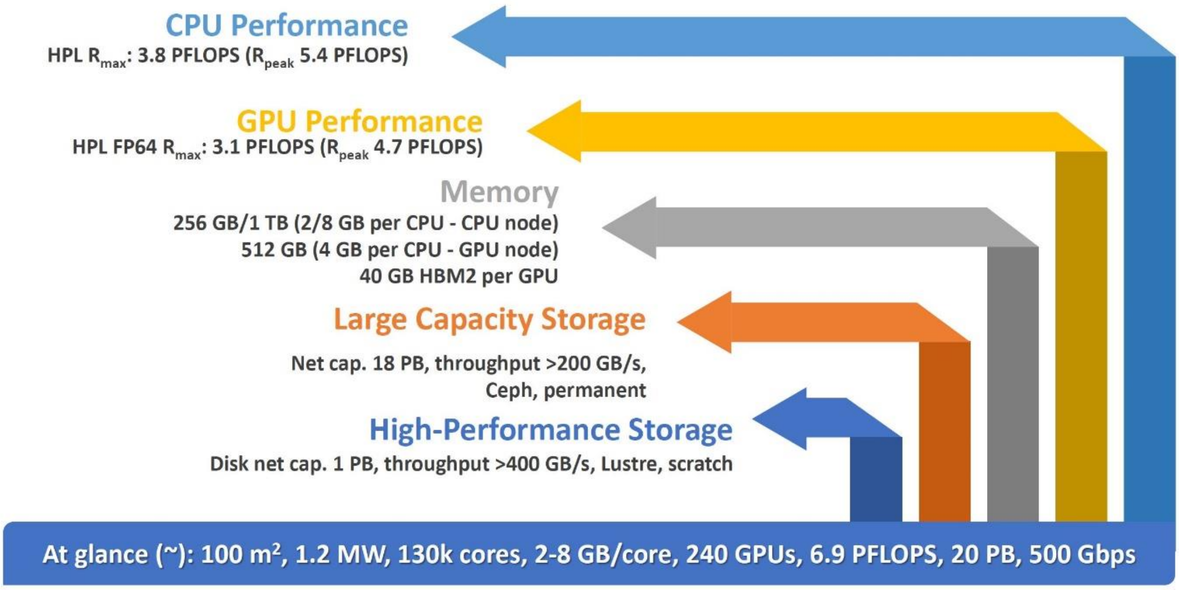

Figure 16: Vega performance overview

The system has two partitions – one is CPU-only and the other is accelerated with GPUs – with an aggregated HPL performance of 6.9 petaflops. These partitions are interconnected via a Mellanox InfiniBand HDR100 using Dragonfly+ topology.

Figure 17 provides the architectural overview of the system. As can be seen, the compute nodes of both partitions are water-cooled and are located in the RIVR1 room (Figure 19, Figure 20). The left side of Figure 18 illustrates the directly liquid-cooled CPU blades. Each blade is equipped with 3 CPU nodes, with each node featuring 2 CPUs which are AMD EPYC 7H12 64-core processors with a Thermal Design Power (TDP) of 280 W. The right side of Figure 18 illustrates the GPU blade with 2 CPUs and 4 GPUs (NVIDIA A100). Overall, there are 320 blades in the CPU partition (i.e. 960 nodes) and 60 blades in the GPU partition of the Vega system. A total of 1.2 MW of power consumption was recorded for the system during testing periods.

The system has two storage tiers. One is based on Ceph and is used for permanent storage, with a net capacity of 18 PB and a throughput of more than 200 GB/s. This storage is based on 61 Supermicro SuperStorage nodes, each node featuring 24 × 16 TB 3.5" SATA3 7.2K RPM data disks.

The other storage tier is used for scratch storage and is based on Lustre, featuring 1 PB of disk capacity with more than 400 GB/s throughput. The storage hardware is based on 10 DDN Exascaler blocks, each with 23 × 6.4 TB NVMe devices.

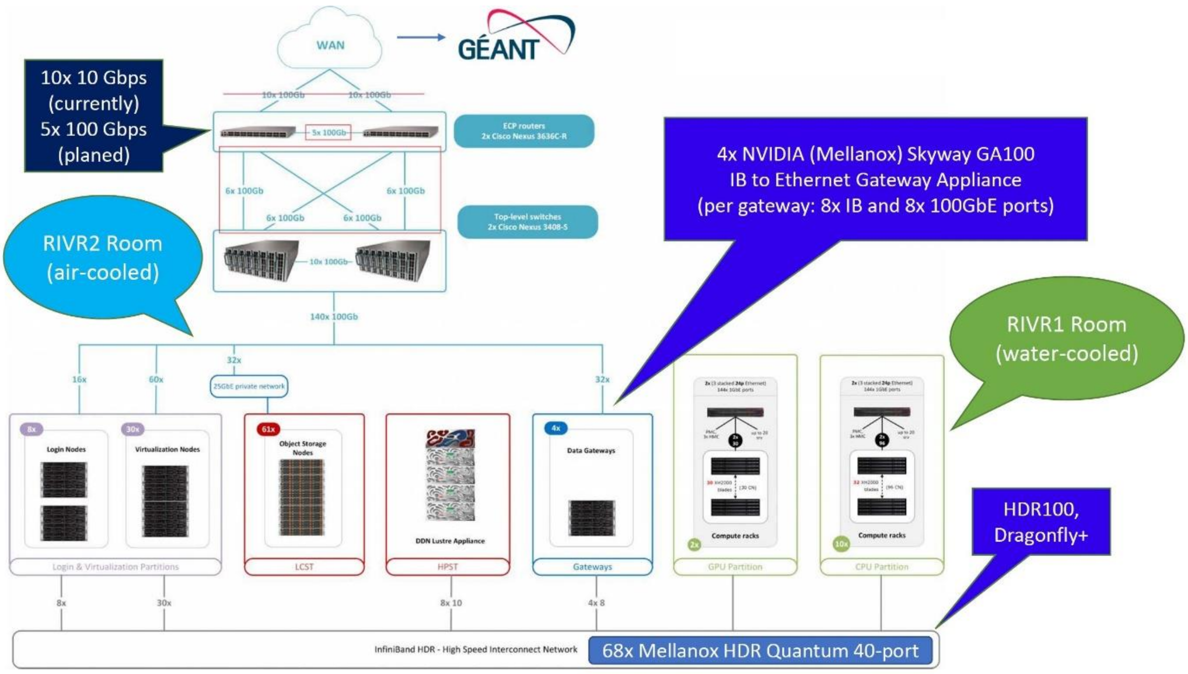

Figure 17: Architectural overview of the Vega system

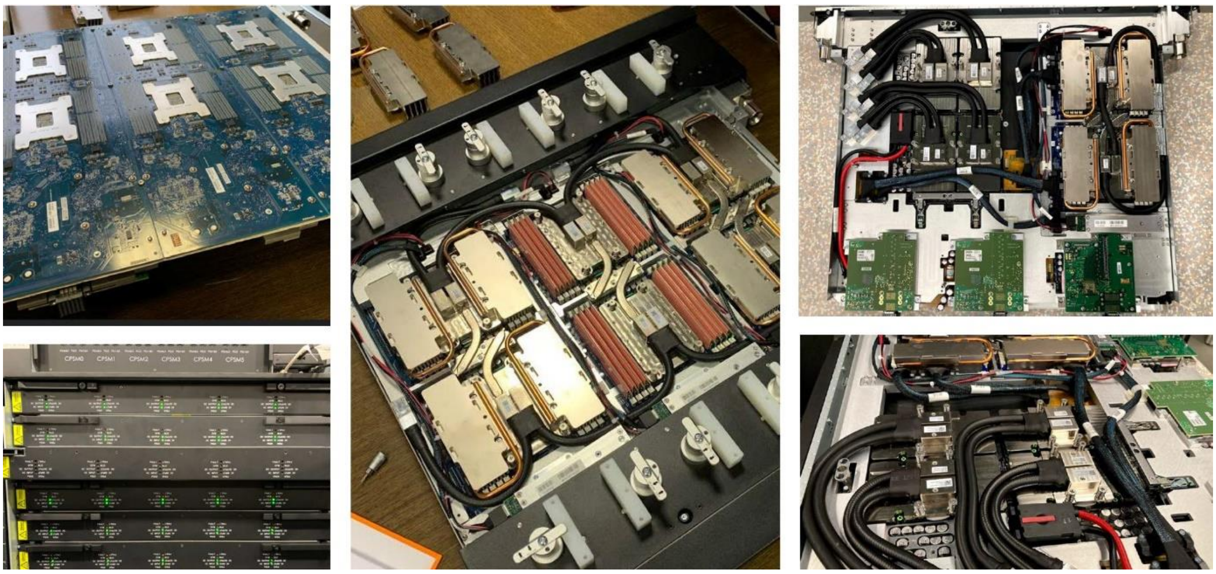

Figure 18: Overview of the Vega compute hardware

Figure 19 illustrates the layout of the RIVR1 room where the compute segment of the Vega system is located. The existing data centre for the production of the library automation system is depicted on the right side of the figure. The left side part of Figure 19 is the floor segment dedicated to the Vega system. There are 12 Atos BullSequana XH2000 Direct Liquid cooling racks, with 10 serving the CPU-only partition of the system. Figure 20 illustrates the current view of the RIVR1 room.

Figure 21 illustrates the layout of the RIVR2 room, which in particular hosts the system login nodes, virtualisation nodes, and data gateways. The power consumption of this system area is limited to 110 kW. The Vega system has 4 CPU login nodes and 4 GPU login nodes. These nodes have the same configuration as the CPU and GPU partitions – this was done to allow users to test their applications using a similar hardware infrastructure[d]. The virtualisation and

[d] However, the GPU login nodes, in contrast to the compute ones, feature only one NVIDIA A100 GPU.

service nodes are intended for some monitoring activities as well as for the provision of certain cloud services to Vega users.

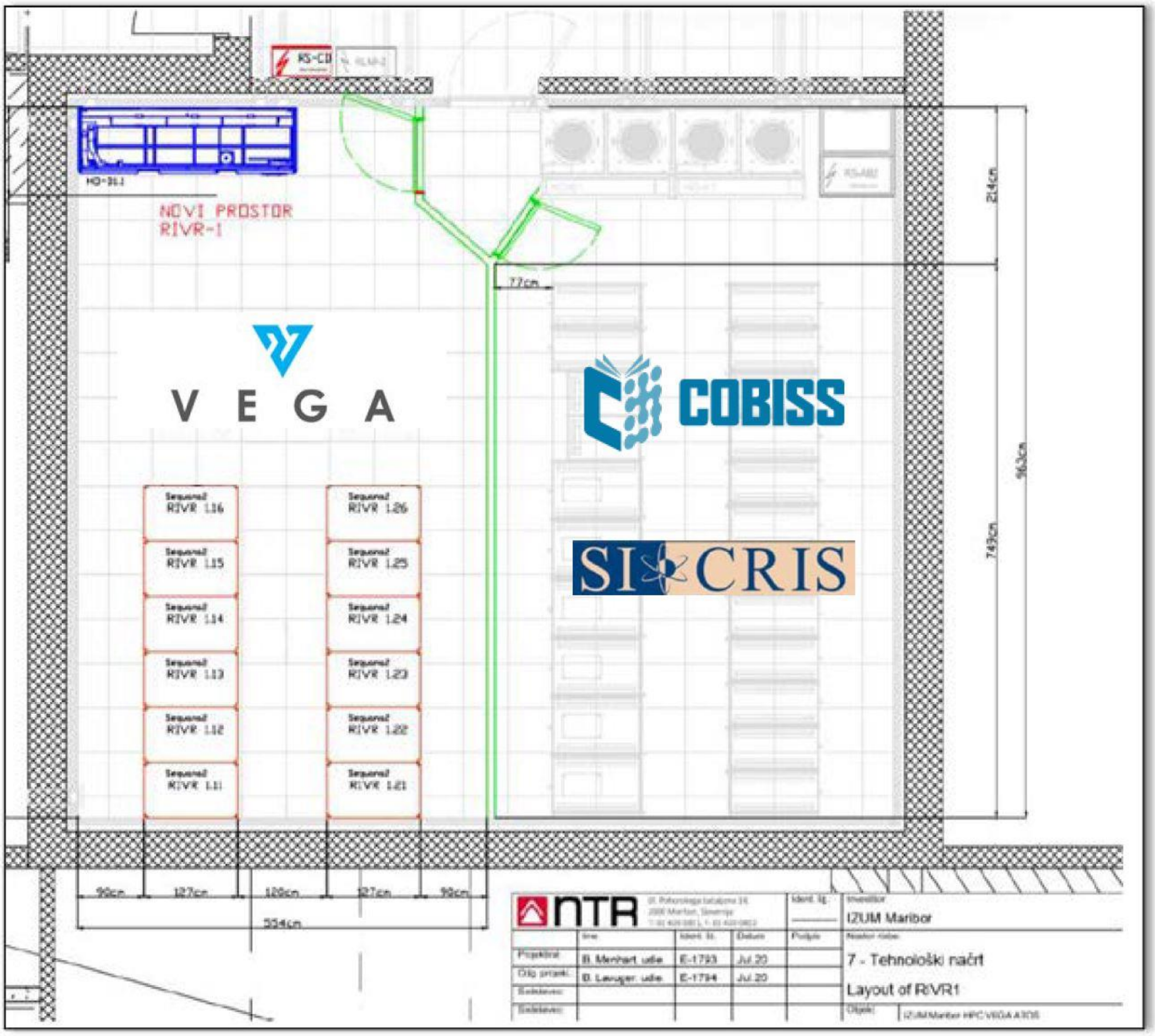

Figure 19: The layout of the IZUM RIVR1 room

Figure 22 provides an overview of the power supply infrastructure. The whole facility is backed up by a diesel generator (left side of Figure 22). In case of a power outage, the storage in the RIVR2 room is backed up with the help of a UPS (UPS C), while the power for the compute part is backed up by UPS D (Figure 22).

Figure 23 provides the schematic overview of the cooling infrastructure. Two adiabatic dry coolers are utilised to provide warm-water cooling delivering an inlet water temperature of 35 °C and taking back water with an outlet temperature of up to 50 °C. This inlet temperature water is then delivered to the compute racks in the RIVR1 room. Furthermore, as can be seen from Figure 23, the cooling infrastructure also features chillers that are aimed at delivering cold water to the Computer Room Air-Conditioning (CRAC) unit and in-row coolers in the RIVR1 and RIVR2 rooms, respectively.

As this is a EuroHPC system, 35% of its capacity will be dedicated to user projects approved by EuroHPC, with the remaining 65% allocated to national-scale initiatives. It is possible to virtually visit some parts of the centre where the system is located using the following link: https://www.izum.si/virtualni-sprehod/VEGA.html.

Figure 20: Current view of IZUM RIVR1 room

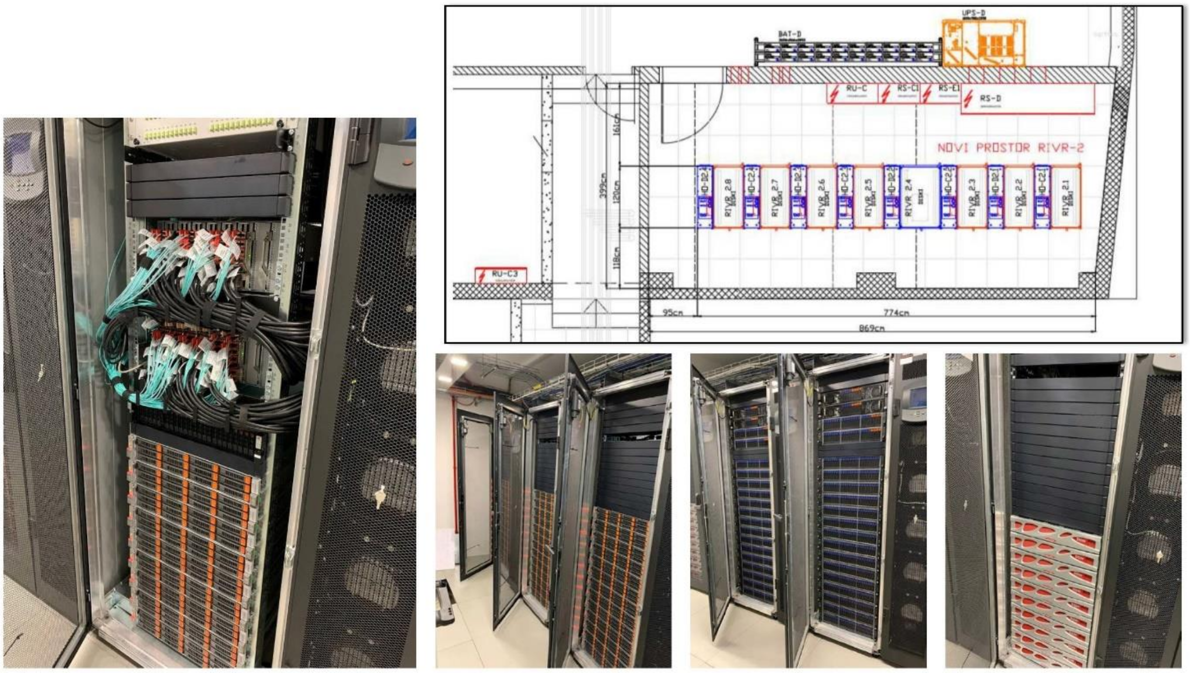

Figure 21: The layout of the IZUM RIVR2 room

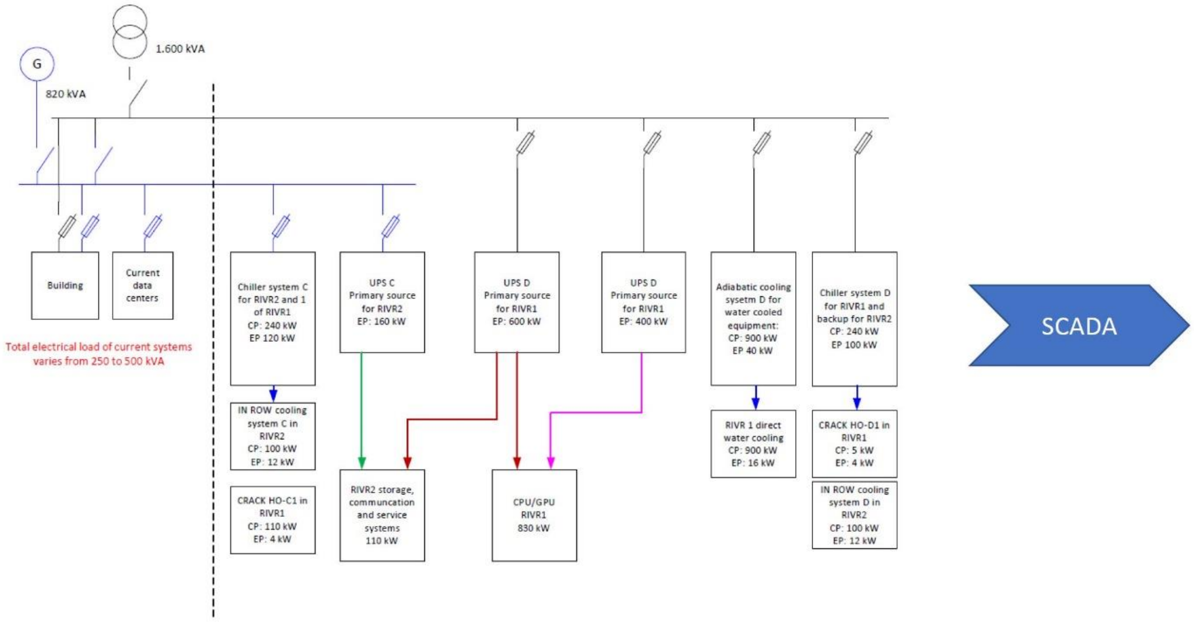

Figure 22: IZUM power supply infrastructure scheme

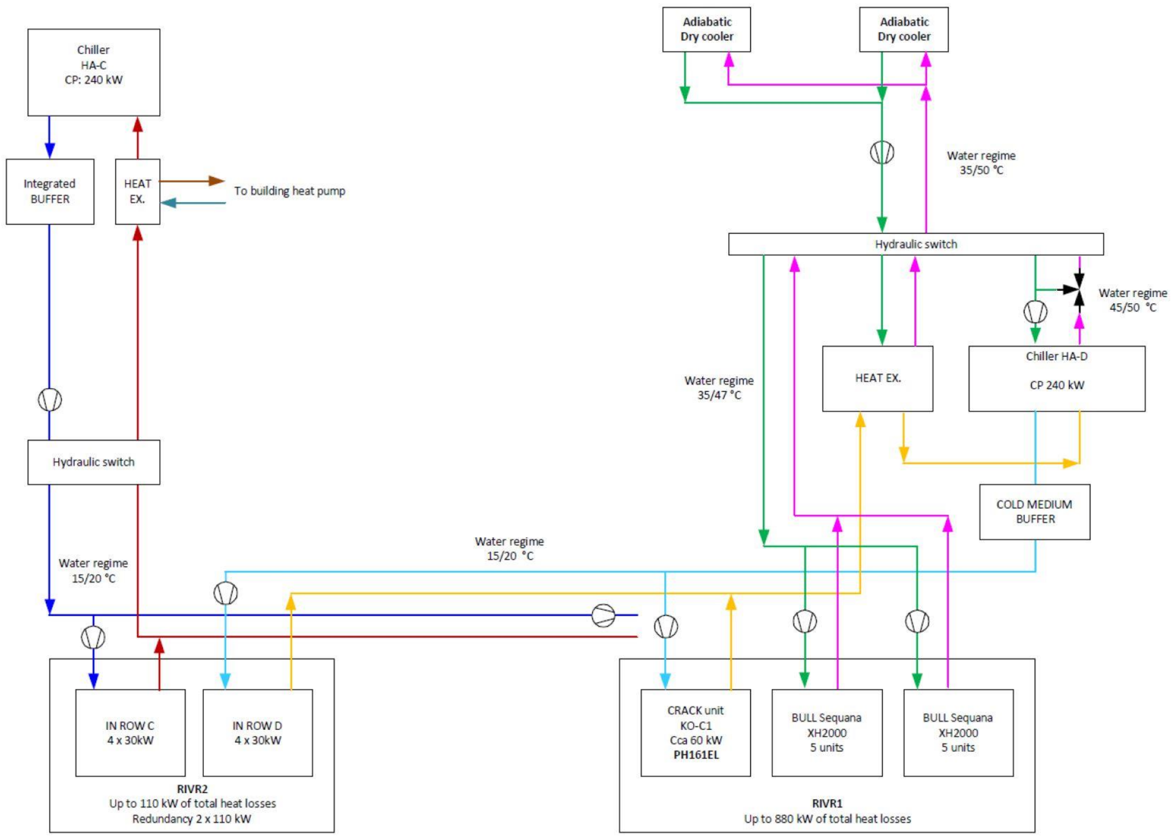

Figure 23: IZUM overview of the cooling infrastructure

5 SESSION IV: EuroHPC Petascale Systems – Chaired by Evangelos Floros (EuroHPC)

5.1 Karolina

Branislav Jansik, IT4Innovations National Supercomputing Center

5.1.1 Introduction

The Karolina supercomputer is operated and hosted by the IT4I National Supercomputing Centre, which was established in 2011 in Ostrava (Czech Republic) and is part of a more extensive computing service environment. In fact, IT4I is a unit of the Technical University of Ostrava (VSB) and is also a member of e-INFRA CZ, which operates the national academic network, storage resources and computing grid infrastructure in the Czech Republic.

Currently, IT4I is operating three supercomputers: Salomon, Barbora, and DGX-2, while Karolina is under installation. IT4I is organised as five research laboratories and supercomputer services units. It is very active in EU HPC activities, being a member and hosting entity of EuroHPC, a member of PRACE, and participating in EUDAT, ETP4HPC, BDVA, and a number of other HPC activities. IT4I has a strong international collaboration that includes 14 Horizon2020 projects, as well as multiple cooperation projects with industry. Training and education activities are also a very important aspect for IT4I; the aim is to increase the uptake of HPC in the Czech Republic and Europe – in fact, the IT4I training program is now part of the University of Ostrava programme.

5.1.2 Infrastructure

All the IT4I supercomputers – Barbora, Salomon (about to be decommissioned) and DGX-2 – share a common network infrastructure (100 Gbit/s technology), a storage infrastructure named “project storage” (which holds the data required during the lifetime of a project), and other common infrastructure such as virtualisation infrastructure to host ancillary services.

5.1.3 DGX-2 and Barbora supercomputers

DGX-2 is a fascinating little system equipped with 16 NVIDIA V100 cards, interconnected via NVSwitch, which makes it possible to exploit unified address memory space across the memories of the GPU cards. The whole system features a performance of over 130 teraflops in double precision. This system was bought as a precursor to the Karolina supercomputer in order to test some key features such as flash storage and the unified address memory space.

The Barbora supercomputer is currently operated by IT4I, and the production phase of the system started in 2019. It is composed of 192 compute nodes, one Symmetric MultiProcessing (SMP) node, and eight accelerated nodes with NVIDIA GPUs. Barbora features an interesting solution for scratch storage. It is a burst buffer solution based on NVMe. In addition, NVMe storage units are allocatable by the compute nodes via the NVMe over fabric technology. The system features over 840 teraflops in double-precision performance. The main purpose of the Barbora supercomputer was to bridge the gap between IT4I’s main supercomputer Salomon, which is about to be decommissioned, and the newly procured Karolina supercomputer.

5.1.4 Project storage

The supercomputers are connected to the centralised "project storage", which was designed to be independent, extendable, scalable and redundant in order to allow continuous operation (even during upgrades and updates) by simply removing and adding new elements of the storage, without impacting operations. The purpose of the storage is to hold project data that is produced during the lifetime of the projects. Currently, the storage features three units giving a total capacity of 15 PB and 40 GB/s of throughput. The storage can be accessed by the individual computing nodes of the supercomputers, but this model of use is recommended only for low I/O intensity applications. For demanding I/O applications, dedicated scratch storage is available on the supercomputers.

5.1.5 KAROLINA supercomputer

The Karolina architecture is composed of the following four main partitions.

- Universal partition – composed of 720 CPU nodes giving a total of over 92,000 cores
- Accelerated partition – composed of 72 accelerated nodes, each featuring 8 NVIDIA A100 GPUs
- Data analytics partition – based on a single SMP node with 24 TB of RAM and 768 compute cores
- Cloud partition – based on 36 nodes with the same architecture as the universal partition

All partitions are interconnected via an InfiniBand High Data Rate (HDR) network with a full fat-tree topology. Besides these computing partitions, there are servers as the front-end and to host auxiliary services.

An important feature of the system is the scratch storage, which is designed to be extremely powerful in throughput performance (over 1 TB/s) while still providing a capacity of over 1 PB.

All the compute components of Karolina can access the "project storage" via dedicated data gateways. The system is capable of an overall peak performance of 15.2 petaflops, with two-thirds of the performance coming from the accelerated partition. From the Top500 standpoint, running a hybrid LINPACK benchmark poses a challenge. By submitting separated sustained values for the main partitions, the system was ranked in the 69th position.

5.1.5.1 Universal Partition

The universal partition is based on 720 HPE Proliant servers equipped with:

- dual socket AMD CPUs providing a total of 128 cores per node,
- 256 GB of DDR4 memory per node, and
- an NVIDIA Mellanox HDR 100 Gbit/s interconnection link.

The whole partition provides 3.8 petaflops of peak performance in double precision.

5.1.5.2 Cloud Partition

The cloud partition is based on 36 HPE Proliant servers equipped with:

- dual socket CPUs providing a total of 128 cores per node,
- 256 GB of DDR4 memory per node,
- an NVIDIA Mellanox HDR 100 Gbit/s interconnection link,

- dedicated NVMe storage, and
- a 10 Gbit/s ethernet card.

5.1.5.3 Accelerated Partition

The accelerated partition is based on 70 HPE Apollo 6500 servers equipped with:

- dual socket AMD EPYC CPUs providing a total of 64 cores per node,
- 512 GB of DDR4 memory per node,
- 4 × NVIDIA Mellanox HDR 200 Gbit/s links, and
- 8 NVIDIA A100 GPU accelerator cards, all interconnected via an NVswitch.

The accelerated partition configuration was upgraded after this presentation was given to contain 72 HPE Apollo 6500 servers, 1024 GB of DDR4 per node and a dual socket AMD EPYC CPU totalling 128 cores per node.

The whole partition provides over 11 petaflops of peak double-precision performance. The NVswitch allows communication between GPU cards without impacting other interconnected cards so that multiple communications across the accelerator cards can run concurrently. Tests have shown that this architecture provides more performance than architectures featuring peer-to-peer interconnection. A100 GPU cards are very powerful, each providing 9.7 teraflops of double-precision peak performance, or 19.5 teraflops if A100 tensor cores are used for algorithms that involve matrix-matrix multiplication. In reduced precision (FP32-FP16), A100 cards can provide way higher performance. IT4I tested and verified these performance figures. Enhanced connectivity towards the InfiniBand interconnection via 4 × HDR 200 Gbit/s is expected to play a critical role in machine learning workloads.

5.1.5.4 Data Analytics Partition

The data analytics partition is based on HPE Superdome flex, and it contains 32 processors interconnected via ASICs in an all-to-all topology through a NUMA network. This node features over 20 TB of main memory and enhanced connectivity towards the InfiniBand interconnection via 2 × HDR 200 Gbit/s to support data analytics workloads.

5.1.5.5 Network Interconnection

The Karolina supercomputer features a full fat-tree non-blocking topology based on an HDR 200 technology, employing split cables to serve a bandwidth of 100 Gb/s or more per compute node depending on the partitions described in the preceding subsections. Tests carried out by IT4I have shown that typically over 92% of the maximum link throughput bandwidth is achieved (>11.5 GB/s).

5.1.5.6 Scratch Storage

The scratch storage is essentially an attached burst buffer storage – completely full flash-based – on the ClusterStor E1000 unit. It is composed of 24 independent storage units essentially made of storage nodes with multiple NVMe devices attached via PCIe4 bus. This storage infrastructure runs a POSIX Lustre file system. The storage offers a capacity of 1.3 PB and a throughput of over 1.2 TB/s. It is designed to match the computing power capacity of the accelerated partition specifically to address machine learning workloads that need to feed data to the GPU cards at a high rate.

5.1.5.7 Installation Plan

At the time of writing, the Karolina supercomputer was in the process of being installed. The floor of the computing room was opened, and the delivery of the racks started in January 2021. Then, the cabinets were connected to the power and cooling infrastructure in February. A partial acceptance of the CPU partition was carried out in April. Currently, the system – complete with all the partitions – is in place and is undergoing the acceptance procedure, even if achieving the committed LINPACK value is taking longer than expected.

5.1.6 Data centre installation

The IT4I cooling infrastructure consists of five water cooling loops: two hot water groups with dry cooling towers on the roof, and three cold water-cooling loops cooled by chillers. Since three supercomputers were in production, the cold cooling loops were in full use during the Karolina installation. This caused an incident when the valves were opened for the Karolina installation, which led to an outage of the cooling system and a partial outage of the computer room. IT4I has now gained sufficient experience to be able to install a new supercomputing system while maintaining other systems in production.

6 Session V – Chaired by Javier Bartholome (BSC)

6.1 Quantum Computers and their Infrastructure Implications

Kristel Michielsen, JUNIQ, JSC

Prof. Kristel Michielsen, leader of the Quantum Information Processing research group and head of the Jülich Unified Infrastructure for Quantum Computing (JUNIQ) at the Jülich Supercomputing Centre (JSC), talked about the current state of, and latest advancements in, quantum computing research at JSC. The following section summarises her presentation.

JUNIQ was founded to provide researchers with access to state-of-the-art quantum technology. Right now, researchers can access the D-Wave 2000Q to conduct reviewed research projects. Additionally, the JUNIQ fleet of accessible Quantum Computer (QC) devices is going to expand over the next few years and will include, amongst other things, the D-Wave Advantage system, commercial and experimental NISQ devices and the Pasqal Quantum Simulator. Currently, JUNIQ's infrastructure is being expanded by installing a D-Wave Advantage system, creating a unified access platform, developing algorithms and prototype applications for QC, performing training and establishing user support structures, and embedding QC into the high-performance computing environment of JSC.

Two of these systems, namely the D-Wave Advantage and the Pasqal Quantum Simulator, will be hosted in the new JUNIQ building, which provides perfect conditions for operating QC. Vibration-free platforms and permanent air temperature control are only two of the various attributes that have been taken into account to create a noise-free environment for such sensitive devices.

Here are details of the recent and planned development at JUNIQ.

- 2019: installation of Atos QLM
- 2019: cloud access to D-Wave 2000Q; development of NISQ device within OpenSuperQ; starting to embed QC devices into the HPC infrastructure of JSC for hybrid quantum computing

- 2021: finishing construction of the building that will be hosting D-Wave Advantage and Pasqal Quantum Simulator; installing and operating D-Wave Advantage System
- 2022: installation of Pasqal Quantum Simulator
- 2023 – 2029: integration of several QC devices of different maturity

The access to stand-alone QC as well as to hybrid systems embedded in the Jülich Modular Supercomputer will provide researchers and industrial partners a fruitful ground for high impact research activities.

6.2 EAR Power Cap for Large Scale Heterogeneous Clusters

Julita Corbalan, BSC

The EAR project was started by the Barcelona Supercomputer Centre (BSC) in Spain in collaboration with Lenovo in 2016. The main focus of the project is an energy management framework that provides services in energy accounting, energy and power control, and energy optimisation. These three services are provided by five components of the EAR software: the EAR Library (EARL), EAR submission plug-in, EAR daemon (EARD), EAR database manager (EARDBD), and EAR Global Manager (EARGM). This section presents the EAR proposal for managing the power cap for nodes and clusters.

EAR is an open-source project and, since August 2019, it has been used in the SuperMUC NG supercomputer at LRZ.

The main objectives of the EAR project are that the software should be easy to use and transparent, optimise the energy at run time, be easy to configure and support large scale systems. Figure 24 gives an overview of the EAR components, and Figure 25 shows the EAR functionality.

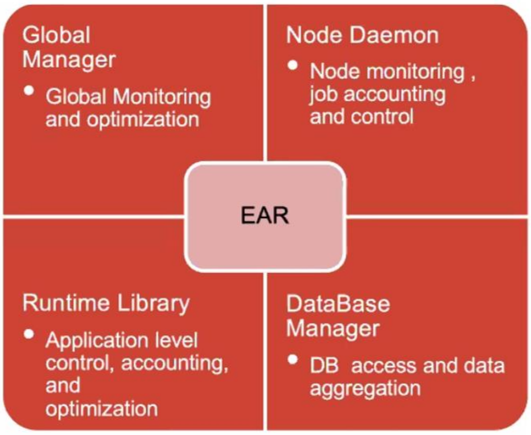

Figure 24: Components in EAR

An overview of the EAR power cap functionality for clusters and the components of the power cap software follows.

EAR power cap requirements: The EAR power cap software should be architecture-independent, support large-scale systems, target heterogeneous systems with CPUs and GPUs, be easily configurable, and the software should vary the CPU/GPU frequencies to optimise the energy consumption. In addition, the combined power cap software should have a minimal impact on application performance. Figure 26 shows the hierarchical design of the EAR power cap software; Meta-EAR GM, EAR GM, and the Node Manager are responsible for the power cap control, power balance, and power cap API, respectively.

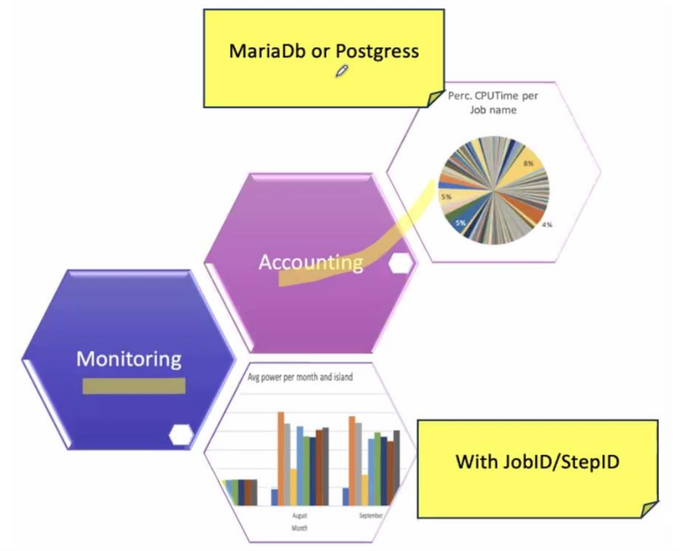

Figure 25: Functionality of the EAR software

Cluster power cap [Meta-]EARGM: This is a cluster-wide component of the EAR software that continuously measures the energy and power from a cluster. The EARGM can be configured for manual operation; in that case, if there is an energy/power warning, it will only inform the system administrator, and no automatic actions will be taken. It can also be configured for automatic operation, in which case it will automatically modify the node energy/power settings in case of any warnings. In general, EARGM will provide three levels of warnings, with each level of warning being dedicated to a different problem. Figure 26 gives an overview of the Meta-EARGM functionally in a cluster environment.

Moreover, Meta-EARGM provides three functionalities for power control. Firstly, EAR daemons will not consume more than the allocated amount of power. Secondly, power is released and then re-allocated. Thirdly, global messages are optimised to be efficient, and they are not used in the critical path of the power management process.

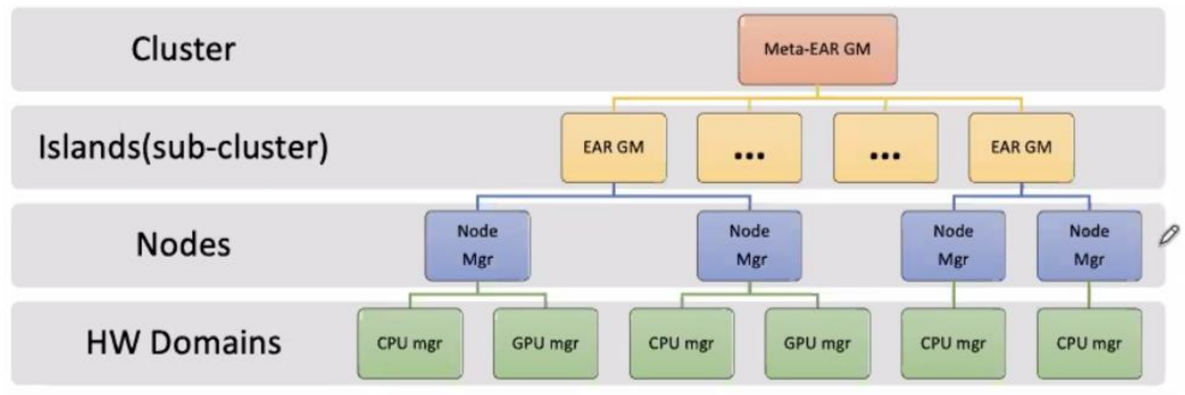

Figure 26: Overview of EAR power cap hierarchical design

Node power cap Manager: The power cap manager for the nodes is implemented as an EARD thread, and it controls both the power control and power balance for the nodes in the cluster. In terms of power control, the node power cap manager measures DC power and verifies the power allocated to the hardware (HW) domains. In terms of the power balance, the node power cap manager sets the default power cap for the HW domains. The power allocated to the node is re-distributed among the different domains (CPU and GPU) based on the power cap status of the

HW domains. At each new job, the power cap distributions for the HW domains are set using the default settings. The node power cap manager supports the following events: a new job, the end of a job, a new power cap, run to idle, idle to run. Figure 28 shows the schematic overview of the node power cap manager in the compute node.

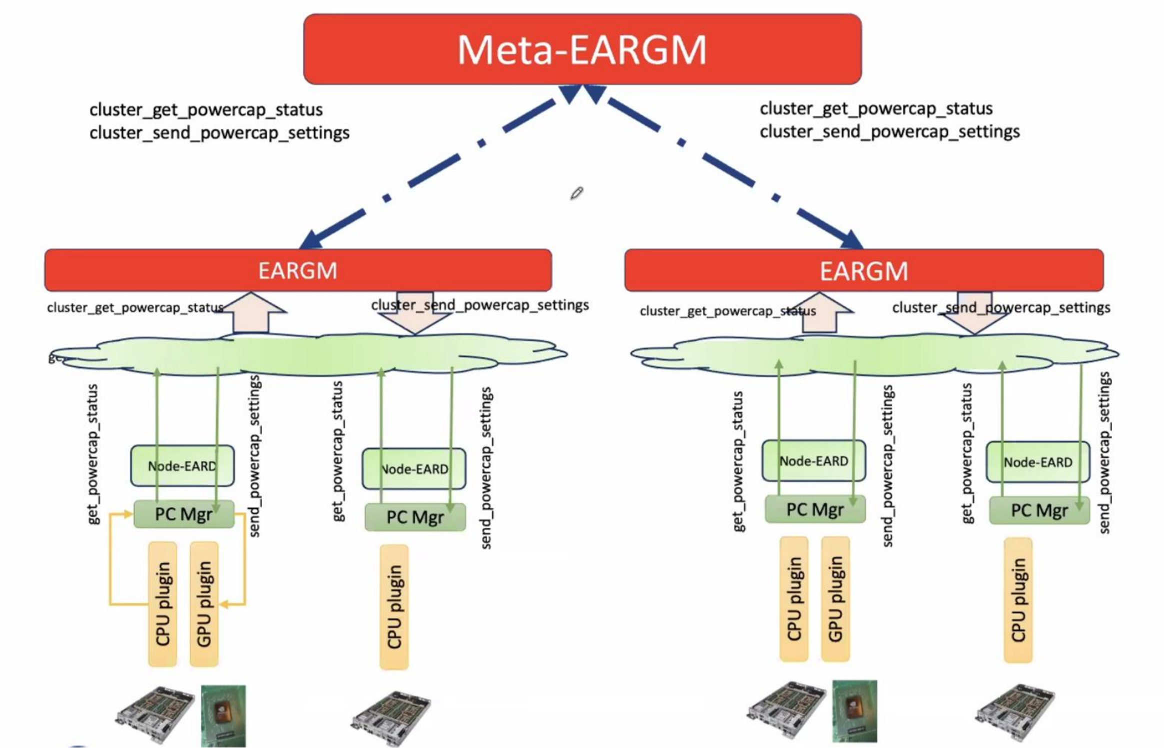

Figure 27: Overview of Meta-EARGM functionality

HW domains: EAR provides plug-ins to support heterogeneous systems, and these plug-ins are implemented in the power cap API. Four HW domains are considered, although only two HW domain plug-ins have been implemented so far. They are CPU (to control the CPUs) and GPU (to control the GPUs). For CPUs, the main plug-ins are the Intel Node Manager (INM) and Dynamic Voltage and Frequency Scaling (DVFS) plug-ins. For GPUs, there are the NVIDIA Management Library (NVML) power limit and NVML frequency scaling plug-ins. All these plug-ins help to measure and control the power in the cluster.

EAR messaging capabilities: EAR supports collective operations, that is, data gathering, broadcasting, and reduction operations for power/energy measurements and monitoring. For example, in SuperMUC-NG with 12 nodes per range, the cluster power status cost is estimated in between 30 to 300 ms.

Node power cap evaluation: A simple test case has been experimented with on a heterogeneous compute node to show the capabilities of the EAR framework. For example, two Intel Xeon Gold 6148@2.4 GHz CPUs consume (TDP) 150 W, and two NVIDIA V100 16 GB GPUs consume (TDP) 250 W; in total, the full compute node consumes 800 W (= 300 W + 500 W).

Finally, the talk concluded that the EAR4.0 status would be evaluated in terms of the Node Power cap, including the plug-ins, and the EARGM Cluster, at the end of Q2, 2021. Meta-EARGM is also under development and will be included in EAR4.0. In the context of the external collaboration, the EAR software will be used in the Snellius system at SURF. The Snellius system is heterogeneous, consisting of AMD and Intel CPUs plus NVIDIA A100

GPUs. Furthermore, the EAR software will be tested against more test cases, especially focusing on I/O, and used in BSC spin-off initiatives (www.eas4dc.com).

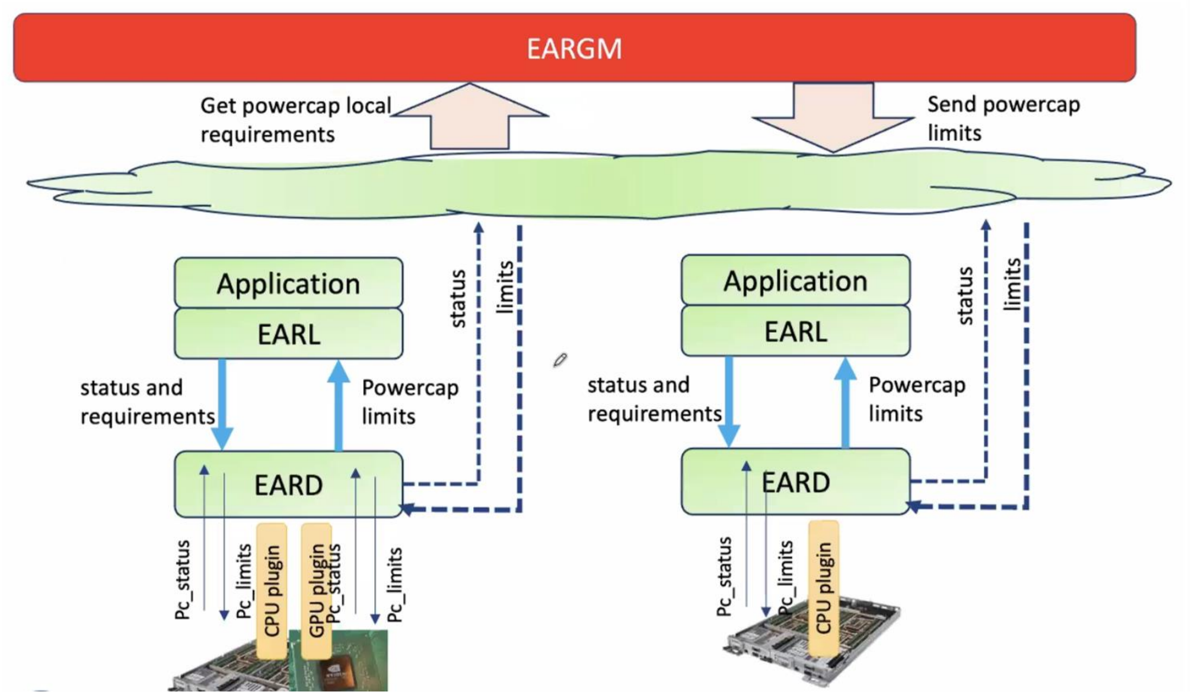

Figure 28: EAR node power cap in the compute node

6.3 Efficient Preventive Fire Protection: Official Rules vs. Performance-Based

Claudius Hammann, Garching Fire Brigade

This presentation was given by Dr Claudius Hammann and is based on his own research. The presentation mainly focuses on basic fire protection strategies at HPC centres and how fire prevention can be addressed or solved. In order to limit the effect of accidental fires, there are a few steps that should be practised to prevent a potential fire or to stop an actual fire and save lives. They are: prevent the fire, halt the spread of the fire and smoke, use efficient fire extinguishers and, finally, save human beings and animals. Two strategies can be adopted during this process: fire safety concepts according to rules, and fire safety concepts according to a performance-based code.

The fire safety concept according to rules covers people and animals, as well as environmental and cultural artefacts. In contrast, the fire safety concept according to a performance-based code focuses on non-cash assets. The main challenge for the performance-based code is protecting the HPC building (non-cash assets) from fire. This is usually considered more challenging than protecting an ordinary building from fire.

Several algorithms are being used for the performance-based code. They are Fault Tree Analysis (FTA), Machine Learning (ML), Classical Statistic, Fuzzy Logic, Bayes Theory and Network Science. Dr Claudius Hammann's thesis is focused on FTA, and the following section gives a brief overview of FTA.

Fault-Tree-Analysis (FTA): The basic concepts of the FTA approach are to describe the failure rate of a complex technical system using Boolean algebra, by splitting the system into a top event and lower-level events and to describe the system with a diagram, as shown in Figure 29.

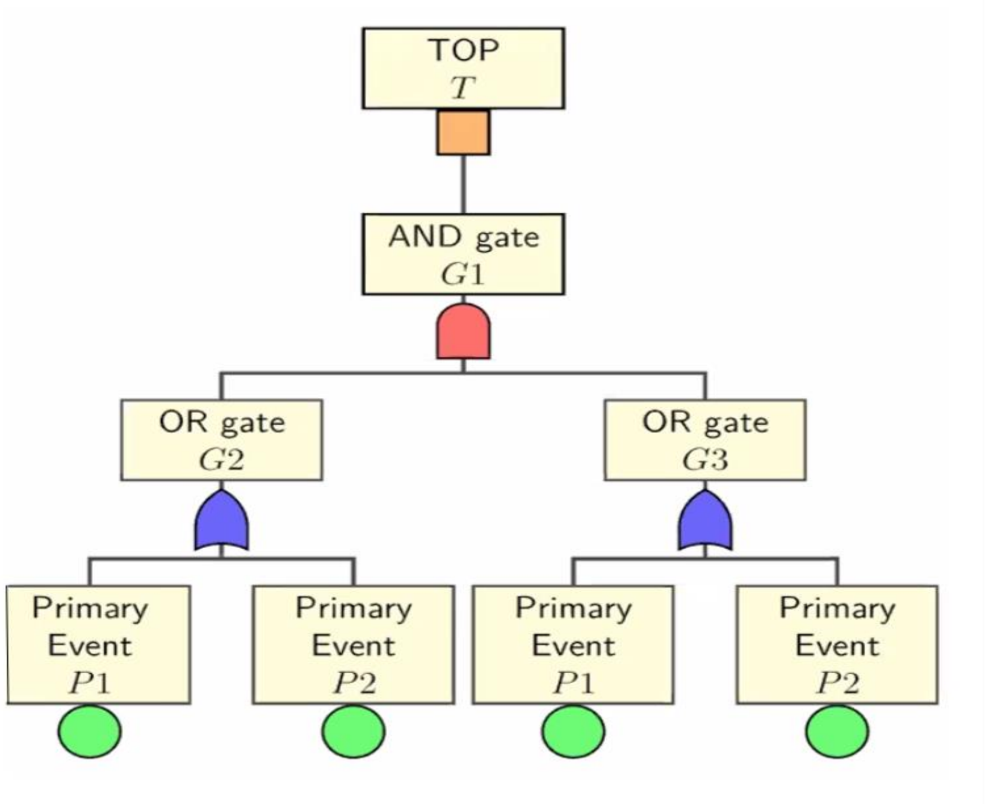

Figure 29: Basic flow structure of FTA analysis

This approach is regarded as simple but, at the same time, it is an extremely effective and useful tool for analysing complex system failures. In an FTA analysis, the whole system in question is represented by Boolean systems which are integrated with a series of lower-level events. There are three important symbols that are used: event symbols, gate symbols and transfer symbols. Each symbol has its own purpose and meaning to support the FTA. For example, Table 3 shows the basics of the Boolean symbols and their descriptions. The FTA does qualitative and quantitative analysis to trace back the fault in the system.

FTA analysis involves five steps:

1. Define the undesired event to study
   Defining the undesired event can be difficult, although some of the potential events are very easy to define as they can be clearly observed. An engineer with a wide knowledge of the design of the system is the best person to help define and number the possible undesired events. Undesired events are used then to make FTAs. Each FTA is limited to one undesired event.
2. Obtain an understanding of the system
   Once an undesired event is selected, all causes with probabilities of affecting the undesired event of 0 or more are studied and analysed. Getting exact numbers for the probabilities leading to the event is usually impossible as it may be very costly and time-consuming to do so. System analysts can help with understanding the overall system. System designers have full knowledge of the system and this knowledge is

very important so as to avoid missing any cause that could affect the undesired event. For the selected event, all possible causes are then numbered and put in sequence in the likely order of occurrence. They are then used for the next step which is drawing or constructing the fault tree.

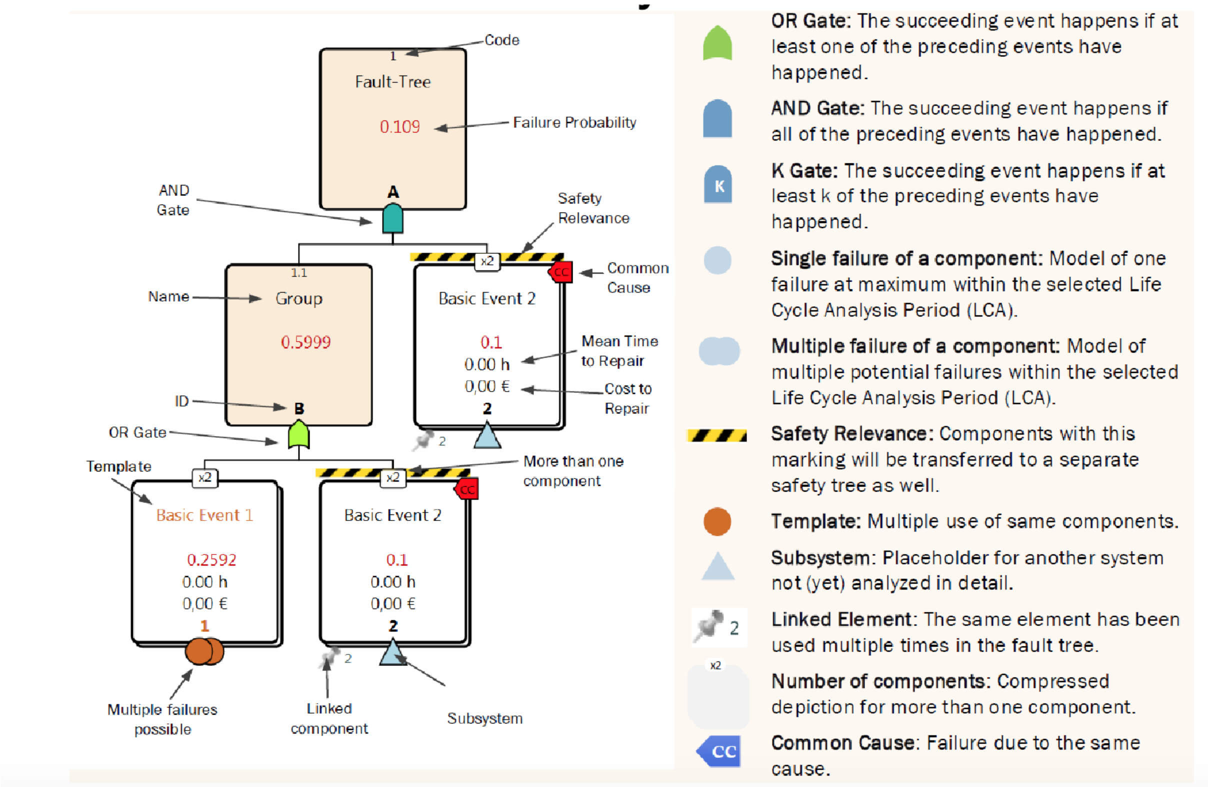

Table 3: Basics of fault tree analysis - Boolean symbol name and its description

3. Construct the fault tree
   After selecting the undesired event and having analysed the system so that we know all the causative effects (and if possible, their probabilities) we can now construct the fault tree. The fault tree is based on AND and OR gates which define the major characteristics of the fault tree.
4. Evaluate the fault tree
   After the fault tree has been assembled for a specific undesired event, it is evaluated and analysed for any possible improvement or, in other words, to study the risk management and find ways to improve the system. A wide range of qualitative and quantitative analysis methods can be applied. This step is a preliminary for the final step which will be to control the hazards that have been identified. In short, in this step we identify all possible hazards affecting the system in a direct or indirect way.
5. Control the hazards identified
   This step is very specific and differs largely from one system to another, but the main point will always be that, after identifying the hazards, all possible methods are pursued to decrease the probability of them occurring.

One of the main purposes of the FTA is to identify risks before they occur. FTA is also widely used in many industries and disciplines, for example, the aerospace, nuclear, chemical, pharmaceutical, and petrochemical industries. FTA is very good at showing how resistant a system is to single or multiple initiating faults. It is not good at finding all possible initiating faults.

7 Session VI – Chaired by Gert Svensson (PDC)

7.1 Update from the Energy Efficient HPC Working Group

Natalie Bates, EE HPC WG

The Energy Efficient HPC Working Group [1] (EE HPC WG) consists of 800+ members from the global scientific and engineering community. In this group 25+ countries are represented with a mix of staff from HPC sites (50% of members), vendor representatives (30%) and researchers (20%). A goal of the group is to bring facility managers and HPC system administrators together, and this is reflected in the group members working at HPC sites. Vendors represented include component, cooling, facility and control system providers as well as system integrators.

Setting standards is out of the scope of the group's work, which is mainly concerned with solving concrete problems. When such a problem exists, a task force is formed to work on that particular issue, and when the work is completed, the task force is dissolved. If the result of the task force is something that should be developed into a standard, the result is turned over to a standardisation organisation as the basis for creating a formal standard. In other cases, the output of a task force is published on the EE HPC WG web page, for example, as a white paper.

In many cases, the area explored by a team continues to be relevant and, after handing over the results of the initial work, the team will reform for further work. Currently, the EE HPC WG consists of the following seven active teams:

- Cooling Controls,
- Operational Data Analytics,
- Procurement,
- Website Migration,
- Liquid Cooling Specification,
- Power API, and
- Electrical Commissioning.

Here are examples of the work being done in these different groups.

Cooling Controls deals with the design process for control systems of mechanical cooling systems. The current work is focused on how to manage upgrades where a new HPC resource is installed in an existing data centre. Deploying a totally new system in a purpose-built data centre usually works well, but the issues involved in the design for this are often overlooked later in the data centre lifetime. This team's current deliverable is a white paper about producing documents known as "Owner's Project Requirements" (OPRs). These documents are used for designing facilities for mechanical cooling. The outline of the white paper matches the OPR structure and includes real-life examples for each section.

Operational Data Analytics (ODA) collects information from sites exploring the use of machine learning and ODA. One of this team's deliverables consists of presentations from participating sites and the discussions that were held afterwards – this deliverable facilitates information sharing about ongoing work between those active in the field rather than being a final report. Another of this team's goals is to spread the word about the relevant technology and develop a model for HPC ODA, so its use becomes more widespread in the HPC community. This model is intended to combine the existing four-pillar model developed at LRZ with a heuristic four-stage ODA model.

Procurement is the team that is least like a task force since the need for procurement will never go away. It was originally created in 2012 and the team's purpose is to encourage sites to consider energy efficiency and power management requirements in procurement processes. The team's current focus is on use cases for power and energy data. The baseline has moved and some originally innovative requirements are now seen as being basic requirements. For example, newer and more innovative requirements deal with automated use of exposed data to optimise application energy consumption.

Website Migration is a team tasked with updating the infrastructure behind the EE HPC WG website, which is currently hosted by the Lawrence Livermore National Laboratory (LLNL). LLNL is migrating to a new website back end, so changes are required regardless of whether the site stays at LLNL or not. As part of this effort, the team is also considering how information is structured and who the audience for different data sets is. Currently, data used by the team is mainly spread among Google drives, and some is stored on GitHub, with the website being used for publications. Navigating and finding data spread out over multiple drives is problematic, and one of this team's goals is to improve this situation.

Liquid Cooling Specification has as its purpose to support the development of an open specification for liquid-cooled system racks. Its deliverable is a draft specification for the fluids used in the cooling systems, for example, the different qualities of water that are available.

Power API promotes the use of the community-supported PowerAPI toolkit. This team will produce a showcase about how PowerAPI is used in the Fugaku system.

Electrical Commissioning deals with the challenges of delivering power to Exascale HPC systems with their high power densities and very dynamic loads. As their deliverable, this team will publish a guide on how to design data centres and commission the delivery of electricity.

7.2 Dealing with Power Swings of Future Exascale Systems

Grant Stewart, LANL

Many presentations at the workshop are from the computer system perspective and are about how to minimise the impact of HPC installations on the power grid. Here is the view from the other side about how the local power grid can accommodate dynamic HPC loads.

Los Alamos National Laboratory (LANL) [2] has a varied mission and employs more than 14,000 people, so hosting HPC data centres is just a part of the power load at the site. Many aspects of the work at LANL are classified, limiting the level of detail that can be provided about installations. The largest HPC systems at the facility are approaching Exascale class.

Three buildings are used for large system installations using both chilled water and direct warm water cooling. Current systems have a dynamic power load which affects the local power grid, and the magnitude of these fluctuations is expected to grow with expansions in compute capacity.

The local site power grid is operated by LANL itself using both on-site power generation (>25% of current peak load) and feeds from the outside. From external sources, three 115 kV transmission lines enter the facility, providing the capacity of importing 200 MVA. Four large substations in the 112 MVA category transform the electricity from 115 kV to 13.8 kV for site distribution. The transmission system is connected in a "ring bus" with inter-tie circuits on the distribution side. Being a local utility means that, from a federal perspective, LANL is deemed to be a utility transmission owner-operator regulated by the North American Electric Reliability Corporation (NERC).

As a utility, LANL also operates a merchant desk to be able to buy energy on the open spot market. For the baseline load, long-term power and transmission service contracts are used. Gas for the local power generator is purchased on a long-term contract in the winter, but gas prices are less volatile during the summer. Power transmission and local power generation are scheduled in advance, and indirectly LANL takes part in the energy imbalance market of loads and power generation for the western USA. Market requirements have made it essential to forecast power requirements in five and fifteen-minute intervals instead of the earlier hourly schedules.

Several other dynamic loads of more than 10 MW exist on the site, including multiple linear accelerators and pulse power magnet systems, so the HPC loads are not unique in their requirements. Having only internal customers frees the local utility from some customer constraints imposed on external utility companies. Power fluctuations are more tolerated by internal LANL users.

A simplified example of the grid design from the regional level down to one HPC resource on the site is shown in Figure 30 below.

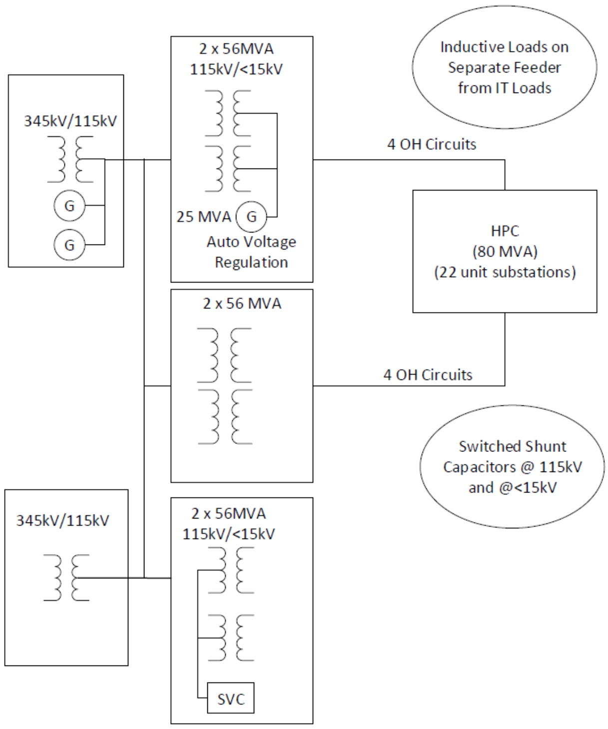

Figure 30: Example of the grid design from the regional level down to one HPC resource

Two 345 kV to 115 kV substations (owned by an external utility) with three feeds connect to the LANL main 115 kV substations (2 x 56 MVA each). Local power generation with 25 MVA capacity is connected on the low voltage side to a 115 kV substation, which helps to deal with power fluctuations and has contingency lines available to help maintain synchronicity. Four overhead lines, each from two different substations, connect to the HPC data centre, where 22 smaller substations distribute the power locally. The data centre facility equipment is capable of handling 80 MVA in total. Inductive loads are separated from IT equipment load feeds.

Installations at LANL to stabilise the local grid and improve the power factor include the following.

- Switched shunt capacitors with varying capacity (2 MVAR, 12.5 MVAR and 25 MVAR) are installed at both the 115 kV and smaller substations.
- A large Static Var Compensator (SVC) with a 100 MVAR capacity is installed centrally at a 115 kV substation.
- A combustion gas turbine generator used for power generation provides large rotational inertia close to the HPC resources and has automatic voltage regulation installed.

The heat generated in transmission lines has been an issue that was resolved by changing from cables in buried ducts to overhead lines. Power factor correction is seen as crucial to having a good relationship with upstream transmission providers.

Long-term planning and developments on the transmission side have been required to prepare for these expansions in the dynamic load. It remains challenging to forecast the amount of power that compute jobs will use, which matches the experience of other sites.

Two different computer models are used for contingency planning and the analysis of transient events that are observed in relation to dynamic loads. Faults and load changes are simulated to ensure N-1 redundancy exists. To aid in the planning, specific criteria have been defined to draw the line between normal variations and scenarios that cause problems.

- The GE Positive Sequence Load Flow (PSLF) model [3] was developed by GE Energy Consulting and is used for the main transmission lines. The following performance criteria are used at LANL:
  - load not exceeding equipment or transmission limits,
  - operating temperature of transmission conductors not exceeding 90 °C,
  - solution converges, and
  - power factor of at least 0.95
- The Electrical Transient Analyzer Program (ETAP) [4] is used for smaller parts with the following criteria (based on 60 Hz line frequency):
  - voltage dips not exceeding 10% for more than 20 cycles,
  - voltage swells not exceeding 10% for more than 1 cycle,
  - no system voltage collapse, and
  - circuit breakers/relays clearing time is less than 6 cycles.

A large HPC application suddenly ramping down can be compared to a fault event from a power grid perspective, and the same type of modelling is used to plan for it.

In the event of a fault, one of the linear accelerators is expected to trip before any HPC resource due to how the electrical installations have been built over time.

This is a summary of the features required on the utility level for dynamic loads:

- power system stabiliser and rotational inertia from a local generator,
- contingency circuits for maintaining synchronicity of generator,
- grid stiffness,
- power transformers with automatic tap changers,
- transmission system “ring bus” with distribution inter-tie circuits,
- voltage support from SVC,
- automatically switched shunt capacitor banks,
- crucial to have a power factor of 0.95 or better on distribution circuits, and
- over-current protection that can discern, and not trip on, dynamic load changes.

A key takeaway from the work done at LANL is that it is the sum of all parts that is important. There is not one single equipment investment that can be done which will fix everything. When asked by the audience to identify the most essential part, Grant mentioned the local power generator.

7.3 ASHRAE TC 9.9 White Paper on Liquid-Cooling

Mani Prakash, Intel

The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) is an organisation that, among other activities, publishes standards for data centre cooling. The Technical Committee (TC) 9.9 [5] is their committee working on data centres.

This white paper [6] on liquid cooling has been a work in progress for a couple of years with many authors, several of whom have changed affiliation during the time period. Content-wise the paper is intended to be a high-level introduction to **why** you want to use liquid cooling, not a detailed investigation into how you should implement it.

Server vendors have benefited in the last decade from performance improvements in the components that have allowed them to keep increasing system performance with only minor increases in system power consumption. This ended in 2018 when performance kept improving but also started to lead to corresponding increases in system power.

Shrinking process sizes in the manufacturing of chips and the use of lower voltages have also led to demands that the chip packaging temperature (known as Tcase) is decreased to reduce the internal junction temperature to levels required by the silicon. Denser chips also concentrate more power into these small packages. To be able to meet this demand and transport all the heat away from components fast enough, liquid cooling is becoming a requirement in more cases. This is not only the case for CPU/GPU components but also, for example, Dynamic Random-Access Memory (DRAM) as it increases in bandwidth and density.

ASHRAE, in general, has changed the classes used in its data centre guidelines for the fifth edition of the document. In terms of air cooling, a more constrained H1 class has been introduced for systems not meeting the full requirements of the A1/A2 classes. Major changes have been made to the water cooling classes where the W1-5 classes have been replaced with six new classes (W17, W27, W32, W40, W45 and W+) with names based on the water temperature at the facility in question.

Until the early 2000s, vendors shipped single-core products, and power increases were moderate, but CPU frequencies slowly approached the level where liquid cooling was required. As CPU vendors started shipping multi-core products, the single hot spot on the chip packaging was split into multiple smaller hot spots and more easily air-cooled. This kept system cooling manageable until the mid-2010s when the number of cores per package and reduced voltage levels in combination led to power densities requiring more cooling. Some of the current top-end CPUs even require liquid cooling. More variants of both high-power CPUs and GPUs are expected in the future, which is likely to change the cooling requirements.

Thermal resistance (between case and ambient temperature) is used as a vendor-independent metric for cooling system requirements. In the last 20 years, this has dropped from 1.40 to 0.20 °C/W, which is a considerable decrease. The inverse of this can be seen as the difficulty of cooling a component.

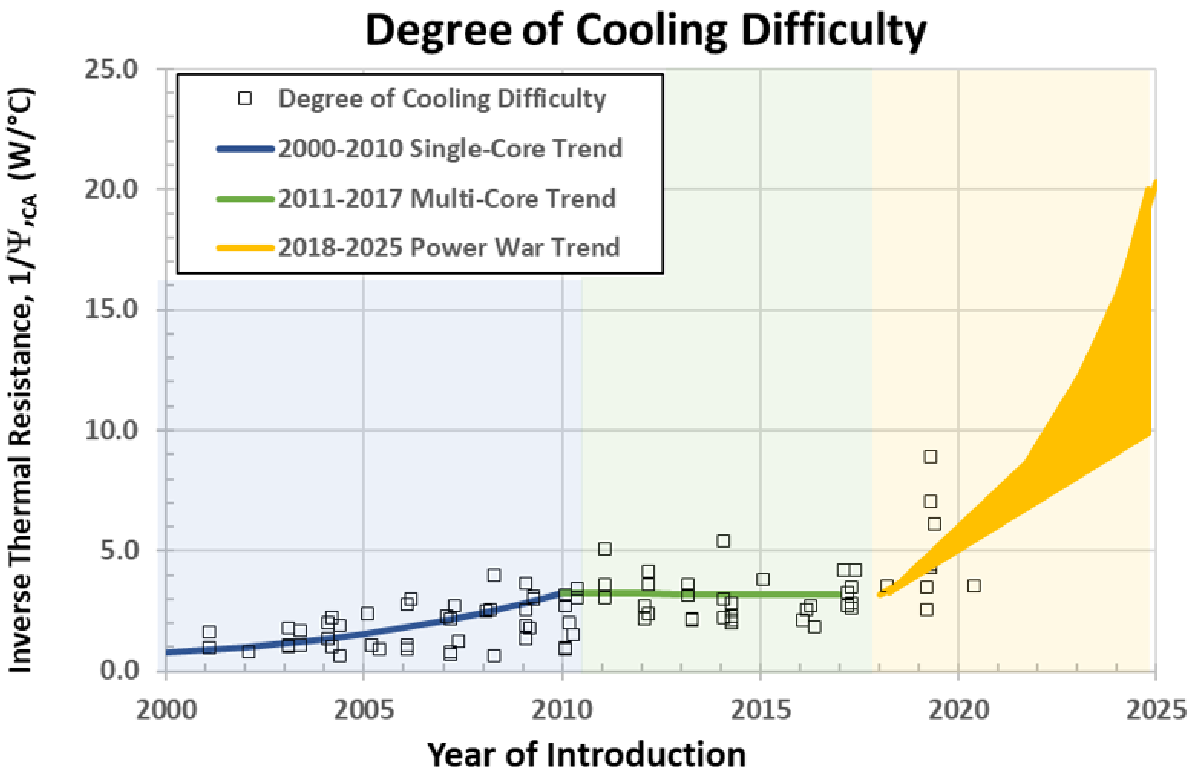

Figure 31: Degree of cooling difficulty

Staying on air cooling requires increased airflow for higher densities, and this also leads to a higher cost for both electricity to drive the fans and possibly more air handling units. As the power density of systems increases, the capability of the air handlers can be exceeded, which means that parts of racks need to be left empty for the air cooling to be adequate.

Acoustics has historically been an issue, especially for mechanical hard drives. Increasing the airflow and the number of fans will make this an even bigger problem.

The memory subsystems are increasingly having cooling issues. Dual In-line Memory Module (DIMM) capacities and the number of channels per DIMM is increasing, as is the number of DIMM slots per CPU socket. A single 256 GB DDR5 DIMM can use 20W at the maximum frequency as an example. System layout requirements mean this becomes a significant contributor to the heat generated around a CPU socket. Cooling the memory with liquid is therefore also becoming a requirement for larger system nodes.

Node sizes constrain the possibility of air cooling; the denser a system becomes, the harder it gets to push air through the system and remove heat. Current HPC trends of very dense racks make liquid cooling more attractive.

Physical limitations constrain the possible values of Tcase, which means there will be trade-offs needed for CPU SKUs with a high Tcase. Waste heat re-use benefits from a high Tcase such that inlet temperatures can be kept high for liquid cooling, but this may then require lower performance on the processor side for stable operation. Developments in cold-plate technologies, such as two-phase and multi-zone cold plates, may be able to offset some of the requirements for a lower Tcase. Audience questions made it clear that current political trends are pushing hard for more re-use of waste heat from data centres, so this may become a large issue in the near future. CPU vendors are aware of this, but fewer products will be able to be used in the higher liquid cooling classes in the ASHRAE guidelines.

Some additional considerations that are not strictly about liquid cooling but affect how it is deployed are covered in an appendix to the white paper. Examples of these topics are TCO, waste heat re-use and power delivery to racks.

8 PRACE Session on Infrastructures – Chaired by François Robin (CEA) and Volker Weinberg (LRZ)

8.1 "Snellius" – The new Dutch National Supercomputer Hosted by SURF

Huub Stoffers, SURF

SURF will replace the current Dutch Tier-1 system, Cartesius, an Atos/Bull system of which the first phase went into production in 2013, with a new system from Lenovo. SURF tendered for a new system, making use of a competitive dialogue procedure. There were two rounds of dialogue with potential vendors in which they could ask questions and provide input on concept requirement specification documents (produced by SURF) for the new system. The final Invitation To Tender document (ITT) was published in July 2020.

The ITT envisaged that the new system would go into production in July 2021, but this target will not be met. Some of the delays that have arisen have been Covid-related and arose from measures taken to battle the pandemic. For example, quarantines had a substantial negative effect on the lead times for getting people and materials on location at the right time. Other delays have been to do with problems that occurred in the process of preparing the data centre for the new system and are further elucidated below. The – still challenging – target is to go into production in September 2021.

8.1.1 Heterogeneous architecture

SURF wants the new system to be an HPC Tier-1 capacity system that can efficiently run highly parallelised applications over a large number of cores, but the system must be able to accommodate a wide variety of other workloads as well. While this might be achievable with a cluster of uniform, very resource-rich nodes, this is not very cost effective. The system requested in the ITT, therefore, is heterogeneous with respect to compute and storage resources.

The ITT requires the cluster to contain CPU-only worker nodes, GPU-enhanced worker nodes, a limited number of "fat" nodes with more memory per core than the standard "thin" nodes, and a limited number of "service" nodes that, among other things, can be used by end-users to run automated data transfers in and out of the cluster in batch jobs.

All node types must be on the same high bandwidth low latency interconnect and have access to all shared file systems: scratch, user-managed scratch or "project" file systems, and home file systems. Some node types are required to have more network bandwidth to the outside world. Some node types must have a truly node-local scratch file system implemented on a fast NVMe device.

For scratch and project shared parallel file systems, the sustained aggregate bandwidth for writing must be at least 24 MiB/s per TiB of net storage capacity, which amounts to a minimum aggregate bandwidth of 48 GiB/s for an individual project's file system of 2 PiB, and to 240 GiB/s aggregate bandwidth for the collective of all of the project space file systems together. An additional shared parallel file system, which is envisaged to be dedicated to jobs that have a very meta-data intensive profile, must be implemented entirely using fast NVMe technology only. The sustained aggregate bandwidth for writing concurrently to this file system must be at least 40 MiB/s per TiB, amounting to at least 8 GiB/s for its initial capacity of 200 TiB. The sustained aggregate bandwidth for reading from this filesystem must be at least 60 MiB/s per TiB, amounting to at least 12 GiB/s for a 200 TiB file system.

The ITT also requested a delivery scheme with a 3-phased expansion of the system to its final capacity to take better advantage of ongoing developments in the market. In the later stages of the system, this will result in increased heterogeneity. In addition to nodes optimised for distinct functions, there will be two generations of nodes with equivalent functions.

8.1.2 Data centre requirements and restrictions formulated in the ITT

The new system is to be installed in the same computer room that houses the current system. The main compute racks of the current system are all DLC, using water with an inlet temperature of approximately 30 °C and returning with a temperature of 40 °C (or higher). The cooling capacity of this loop can be upgraded to 2,400 kW. The room has a fairly limited capacity for heat dissipation by air of 600 kW, and SURF has been clear that it has no intention to change that. The ITT explicitly specifies that the available air-cooling capacity should be reserved for the auxiliary equipment (which is complementary to the compute part of the system) as it tends to be less dense and have a smaller number of nodes and may not be available in a liquid-cooled package. Examples of such auxiliary equipment include: networking equipment, storage arrays and/or storage servers, and possibly nodes for the management and deployment of the system. Furthermore, the new system is expected to do better than the current one in this respect and use no more than 370 kW of the air-cooling capacity. All compute-racks must use a water-cooled solution that is almost (95%+) room neutral.

Based on the requirements and restrictions formulated in the ITT, as well as on the responses and product roadmaps shown in the rounds of dialogue, we had expected the offers from all the vendors to invariably propose a "hot water" DLC cooling solution for the compute racks. The final ITT document formulates very restrictive constraints on the use of air-cooling. However, it does not specifically prescribe DLC as an obligatory mode of cooling. Things turned out otherwise though, with the winning tender featuring a hybrid cooling solution out of necessity (as explained later).

The Lenovo offers won the tender fair, square, and convincingly, given the predetermined criteria for scoring the offer. But the Lenovo design team came to the conclusion during the development of their proposal, given the combination of technical and timeline requirements, that SURF would be best served by a solution using AMD CPUs for most of the requested node types. However, at that time, the development of compute node chassis by the company had mainly been in the context of using Intel CPUs.

8.1.3 Lenovo offer: CPU-only and GPU-enhanced compute nodes

In line with the requirements of the tender, Lenovo's offer specified an initial phase and two extension phases, where the last extension left some options to be decided on later. The nodes delivered in a particular phase are connected in a non-blocking two-level fat tree. The distinct phases will be interconnected by means of a pruned tree, using InfiniBand NDR switches.

The summarised specifications of the compute nodes that were offered are as follows.

Phase 1 (Q3 2021)

CPU-only nodes:

- All nodes: 2 × 64-core 2.6 GHz AMD 7H12 ("Rome") CPUs
  - Thin nodes: 504 × Lenovo ThinkSystem SR645, 256 GB

- Fat nodes: 72 × Lenovo ThinkSystem SR645, 1 TB, 6.4 TB NVMe
  - High memory nodes: 2 × Lenovo ThinkSystem SR665, 4 TB
  - High memory nodes: 2 × Lenovo ThinkSystem SR665, 8 TB
  - Mellanox ConnectX-6 HDR100 (total node interconnect bandwidth 100 Gb/s)
- Aggregate peak performance of all CPU-only nodes: 3.1 PFLOPS

GPU-enhanced nodes:

- 36 × Lenovo ThinkSystem SD650-N V2
- 2 × 36-core 2.4 GHz Intel Xeon 8360Y (Ice Lake) CPUs
- 1 × NVIDIA HGX A100 4-GPU (4 × A100 40 GB**,** Ampere, Redstone)
- 512 GB
- GPU nodes phase 1: 2 x Mellanox ConnectX-6 HDR (total node interconnect bandwidth: 400 Gb/s)
- Aggregate peak performance of all GPU-enhanced nodes: 3 PFLOPS

Phase 1 thus constitutes a compute capacity that is about 3.4 times the capacity of the current systems. The average energy consumption of the system in this phase is expected to be around 620 kW, which is roughly equal to 70% of the energy consumption of the current system.

Phase 2 (Mid 2022)

- **An extension of uniform “thin” CPU-only compute nodes**
- Two socket Lenovo nodes, using future generation AMD EPYC processors, colloquially named “Genoa”, of which the final specifications are still under a Non-Disclosure Agreement (NDA)
- 2 GB memory per core
- Mellanox ConnectX-6 HDR100 (total node interconnect bandwidth 100 Gb/s)
- Aggregate peak performance: 5.1 PFLOPS (or better)
- The nodes will be cooled using DLC

Phase 1 plus Phase 2 thus constitute a compute capacity that is about 6.1 times the capacity of the current system. The average energy consumption of the system in this phase is expected to be about 1,200 kW, which is about 140% of the energy consumption of the current system.

Phase 3 (Mid 2023)

For this phase, there is a limited number of options. The final choice will be made 1.5 years after the start of production of Phase 1 and will be based on actual usage/demand.

- An extension with more CPU-only thin nodes of the same type as the Phase 2 extension, with an aggregate peak performance of 2.4 PFLOPS
- An extension with GPU-enhanced nodes, using a future generation of NVIDIA GPUs of which the final specifications are still under NDA, with an aggregate peak performance of 10.3 PFLOPS
- The nodes in this option must be equipped with 2 × Mellanox Next Data Rate (NDR) (total node interconnect bandwidth: 800 Gb/s)
- A still to be determined amount of additional storage, which could also be more nodes with truly node-local or node-dedicated NVMe devices

Assuming a CPU-based Phase 3, the total peak performance of Snellius will reach some 13.6 petaflops, which amounts to about 7.6 times the capacity of the current system. Assuming a GPU-based Phase 3, the total peak performance will reach some 21.5 petaflops, which amounts to about 11.9 times the capacity of the current system.

The average energy consumption of the system in the final phase, assuming the GPU option, is expected to be about 1,430 kW, which is about 160% of the energy consumption of the current system.

8.1.4 Cooling solution

Figure 32 shows the inside of the Phase 1 compute nodes. Clearly, the GPU-enhanced compute node, using Intel CPUs and NVIDIA GPUs, is on a board with DLC-cooling, whereas the CPU-only node, using AMD CPUs, is not. The CPU-only nodes of Phase 1 use Rear-Door Heat eXchanger (RDHX) as a mode of cooling in an – almost – room-neutral way, and, in compliance with the requirements of the ITT, leave the limited capacity for air-cooling for management, storage and network components.

The amount of effort it takes for a system integrator to develop a DLC board is substantially higher than for an air-cooled board. And boards developed for Intel Xeon CPUs cannot simply be used with AMD EPYC CPUs. Lenovo did not plan, and does not have, a production-ready liquid-cooled chassis for AMD Rome (or Milan) CPUs. It was not practically feasible to get it production-ready and produced in time for this system. Moreover, given the AMD roadmap and the size of the order(s), the amount of development that would be needed to produce such a board would not be commercially feasible for the vendor, particularly as the development of the next-generation AMD EPYC (Genoa) CPU is well underway. For the upcoming AMD EPIC CPU, and hence for the phase 2 extension of the system, Lenovo plans to have a DLC chassis. Thus, in phase 1, even more heterogeneity than was planned (pertaining to cooling solutions) is introduced into the system. Since phases 2 and 3 are extension phases – the nodes of phase 1 are planned to remain operational until the end of the life of the system – the hybrid cooling solution will be a permanent characteristic of the system.

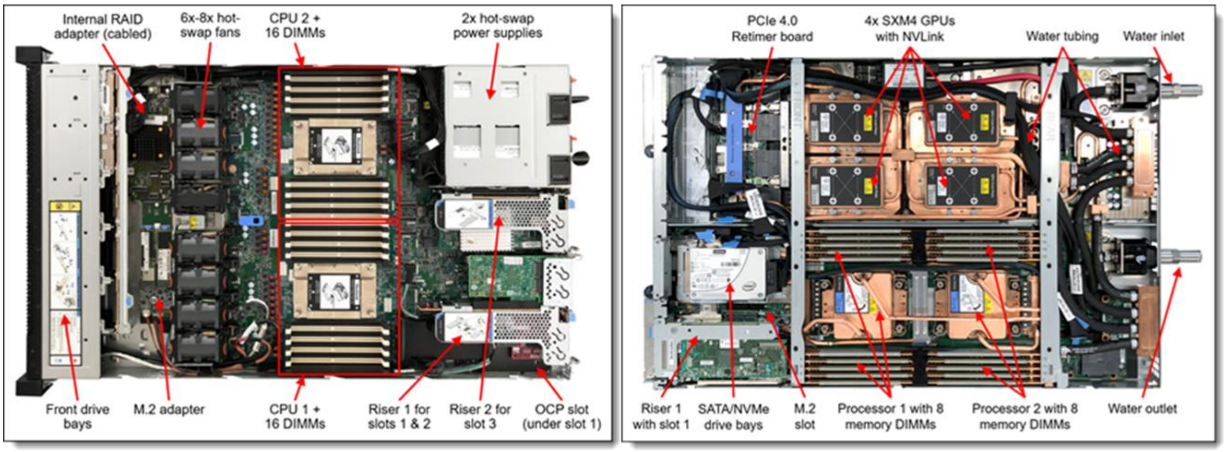

Figure 32: Snellius phase 1 nodes: on the left a CPU-only node; on the right a GPU-enhanced node

8.1.5 Data centre adaptation

The computer room, which has been housing the Dutch National Supercomputer since 2016, is leased in a data centre that is operated by a commercial colocation data centre provider. SURF leases two floors of the centre, one of which is used for the Dutch Tier-1 system. When SURF was moving to the data centre (that was built in 2015 and 2016), the heat exchanger and piping infrastructure for providing a hot water loop to the supercomputer room were built in.

The racks with RDHX cooling need an inlet temperature that is at least 10 °C lower than the inlet temperature for the DLC racks. The original solution proposed by Lenovo was to use the existing loop, with a lowered inlet temperature of about 20-22 °C. The RDHX racks “pre-heat” water for the DLC racks that would finally return the water to the data centre side at a temperature of about 40 °C. With the information initially provided by the data centre provider, this seemed a feasible option. The alternative was to add a second loop with lower inlet and outlet temperatures and leave the existing loop intact.

In the end, the second solution was implemented, with a new dedicated loop feeding a Lenovo cooling distribution unit. It proved cumbersome and time-consuming to decide on and engineer a satisfactory solution in the tripartite setting involving SURF, the data centre provider and Lenovo. Some essential aspects of data centre infrastructure information turned out to be incorrect or outdated, e.g. on materials used for piping (affecting water quality) and on higher pressure levels to take into account (having an effect on safety).

It did not help that the data centre changed ownership after less than a year of operation. The original company merged with another data centre providing company and, because of market share regulations aiming to prevent de facto monopoly situations, the resulting larger company was obliged to sell some of its data centres in the Amsterdam area to competitors. With the obligatory transfer of the data centre to a different owner, a significant amount of the specific local knowledge acquired with designing and implementing the power and cooling accommodation for Cartesius in 2016 was no longer present within the company currently owning the data centre. A lot of the engineering work had to go out to subcontractors of the data centre provider. After several weeks of engineering work, a – non-satisfactory – solution was developed by one of the data centre provider’s subcontractors, and the leading role was transferred from the data centre provider to Lenovo. The data centre provider company, a multi-national company, has also involved several non-local experts now but in a more reviewing role: checking whether the data centre engineering done by Lenovo is compliant with the standards and regulations of the data centre provider.

8.2 Dardel – A new system at PDC, KTH

Gert Svensson, PDC

Gert Svensson presented the new general-purpose HPC system at the PDC Center for High Performance Computing, KTH Royal Institute of Technology, Stockholm, Sweden. The new system, Dardel, replaces PDC’s current flagship system, Beskow, which is a Cray XC40.

The process of procuring and installing the new system at PDC started with a request from the Swedish National Infrastructure for Computing (SNIC) for proposals to install and operate a new general-purpose HPC system for academic research. The idea was to concentrate hardware resources from several Swedish sites into one site. PDC and KTH submitted a project proposal that was approved in the autumn of 2019. The project application proposed a system that would consist of a partition based on CPUs for running traditional HPC jobs and a more forward-looking GPUs partition for codes that have been, or will be, adopted to upcoming exascale systems with GPUs.

The work on the procurement process started immediately after the project proposal was accepted. A decision was made early on to use a negotiated procedure with prior announcement as the formal type of procurement process. A Total Cost of Ownership (TCO) approach was used to select the winning bid. That means that the offer with the highest score for a given amount of funding (for preparing the computer centre, as well as installing the system and operating it for five years) would win. The scoring was allocated using the weights and criteria in Table 4.

Weight	Criteria
55%	Aggregated performance of the benchmarks tests (petaflop-days)
15%	Quality of solution to technical and maintenance requirements
10%	Evaluation of commercial requirements
10%	Quality and extent of collaboration projects with the vendor
10%	Price and quality of options

Table 4: Procurement scoring for Dardel

The benchmark cases for the performance of applications on the potential new system covered GROMACS, NEK5000, VASP, PowerFLOW, PyFR and CP2K tests. Since KTH expected increased competition in the GPU market in the near future, the vendors were allowed to put in bids that involved delivering the GPU part of the new system after the CPU part. To take into account that the vendors could propose different delivery dates for the GPU part of the system, the performance in the benchmark tests was multiplied by the number of days that it was expected the GPUs would be operational (given the different delivery dates and assuming the same fixed end date for all vendors), so the results of the benchmark tests were given as petaflop-days.

It became evident during the spring of 2021 that the winning bid was a proposal from Hewlett Packard Enterprise (HPE) for an HPE Cray EX system. The contract with HPE included necessary adaptations to the PDC computer hall, like connecting the new system to the KTH power network and to the KTH district cooling system (which would enable the heat generated by the system to be re-used) and strengthening the computer hall floor. The contract also included several options for a substantial expansion of the system. It was decided to execute some of these options so that the system would include additional nodes for industrial research, especially with the heavy vehicle manufacturer, Scania. That resulted in a system that will consist, in total, of 858 CPU nodes (with dual AMD Epyc 2.25 GHz CPUs, each with 64 cores) and 56 GPU nodes (each configured with four GPUs that will be a future version of the AMD Instinct GPU). The GPUs are expected to be delivered at the end of 2021. The resulting system will be similar to the much larger EuroHPC Lumi system in Finland.

The new system at PDC has been named Dardel [7] in honour of the Swedish author, Thora Dardel, and her first husband, Nils Dardel, who was a Swedish painter. The front panels of the Dardel system feature "The Dying Dandy" oil painting (which is one of Nils' most famous works), and a portrait of Thora that also was painted by Nils, along with text in the background of the blue panels from Thora's book about Nils.

Figure 33: Rendering of KTH's HPE Cray EX system Dardel

8.3 Hosting Leonardo on Bologna Big Data Technopole

Mirko Cestari (CINECA)

8.3.1 Introduction

CINECA will host and manage one of the three recently procured EuroHPC precursors to exascale systems. The Leonardo supercomputing system [8] will be installed at the newly built data centre located in the Bologna Big Data Technopole (henceforth denoted as Technopole).

The Ministry of University and Research, along with the Emilia-Romagna regional government, supported CINECA's candidacy to act as the hosting entity for the pre-exascale system by providing part of an old tobacco manufacturing building that eventually became the new data centre. This action was deemed consistent with the plan to develop a big data thematic technopole that would also include the hosting of the Italian Institute for Nuclear Physics (INFN) main data centre, a Tier-1 centre known as CNAF, in Bologna. The Emilia-Romagna region and the Ministry of University and Research had already established a collaboration in order to promote the Technopole project at national and international levels. This collaboration was successful in obtaining a decision from the European Centre for Medium-Range Weather Forecasts (ECMWF) council to relocate its own data centre to the Bologna Technopole. Therefore, by virtue of hosting the ECMWF, CINECA and INFN data centres, Bologna Technopole has become one of the leading European hubs for computing and data processing.

CINECA addressed all the technical aspects of the procurement procedure and the design of the supercomputing system architecture, significantly promoting the competition between the candidates, which resulted in final offers outbidding the tender minimum requirements. In fact, the result of the procurement procedure is a system capable of nearly 250 petaflops and equipped with over 100 petabytes of storage capacity. It will be based on Atos Bull Sequana XH2000 technology, with over 13,000 GPUs based on the NVIDIA Ampere architecture and an NVIDIA Mellanox HDR InfiniBand interconnect. The system will provide 10 to 20 times the computing power of the current CINECA flagship system, Marconi-100.

8.3.2 Leonardo supercomputing system

8.3.2.1 System architecture

The overall architecture is composed of two main modules.

- A Booster Module, whose purpose is to maximise computational capacity, is designed to satisfy the most computationally-demanding requirements in terms of *time-to-solution* while optimising the *energy-to-solution*. This result is achieved with 3456 nodes, each equipped with four NVIDIA Ampere-based GPUs (labelled "A100-64" from now on) and with two 32-core Intel Ice Lake CPUs, for a total computational performance of 240.5 petaflops of sustained performance.
- A General Purpose/Data-Centric Module (labelled GP-DC from now on) aims to satisfy a broader range of applications; its 1536 nodes are equipped with two 26-core Intel Sapphire Rapids CPUs per node in order to reach over 8 petaflops of sustained performance.
- These modules are coupled with the following equipment in order to build the whole infrastructure:
  - a front-end interface composed of 16 login nodes with two Ice Lake CPUs (32 cores each), 512 GB RAM and 6 TB disks in RAID-1 configuration (In addition,

16 additional nodes are equipped with 6.4 TB NVMe disks and two NVIDIA Quadro RTX8000 48GB to be used as visualisation nodes.);

- a fast storage area, with a read performance of 744 GB/s and write performance of 620 GB/s on the top of 106 PB of net capacity; and
- a 200 Gbit/s InfiniBand network.

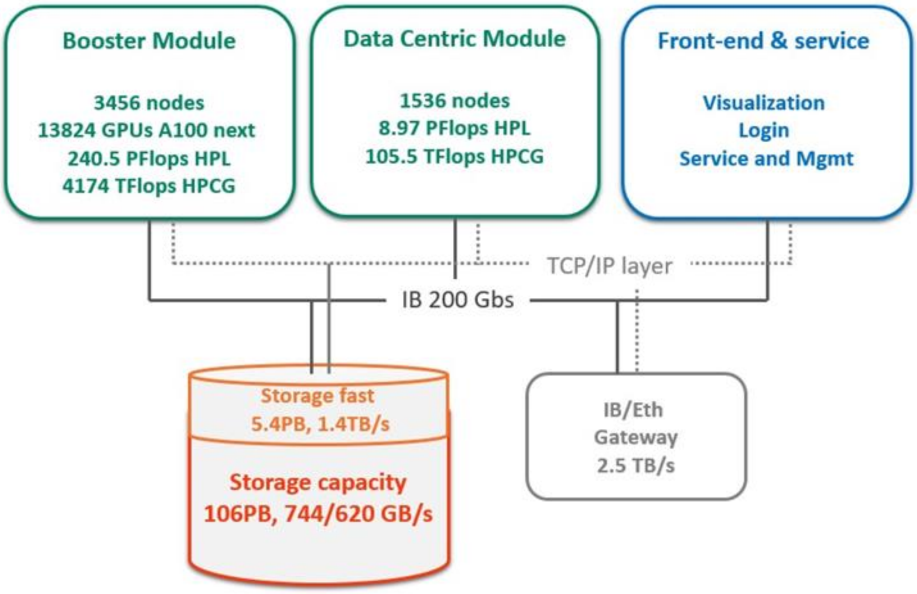

Figure 34: Overview of Leonardo system

8.3.3 High speed interconnect

The low latency high bandwidth interconnect is based on a Mellanox HDR200 solution and designed as a Dragonfly+ topology. This is a relatively new topology for InfiniBand-based networks that makes it possible to interconnect a vast number of nodes with a moderate number of switches while keeping the network diameter very small[3]. In comparison to non-blocking fat-tree topologies, costs can be reduced, and scaling out to a larger number of nodes becomes feasible. In comparison to the 2:1 blocking fat tree, close to 100% network throughput can be achieved for arbitrary traffic. Dragonfly+ topology features a fat-tree intra-group interconnection, with two layers of switches, L1 and L2, and an all-to-all inter-group interconnection with a third layer of switches, L3.

The solution comes with improved adaptive routing support, which is crucial for facilitating high bisectional bandwidth through non-minimal routing. In fact, intra-group routing and inter-group routing need to be balanced to provide a low hop count and high network throughput. This is obtained with routing decisions being evaluated in every router on the packet's path and allows a minimum network throughput of ~ 50%.

8.3.4 Storage architecture

The storage system features a capacity tier and a fast tier. This architecture allows incredible flexibility and the ability to address even the most demanding IO use cases in terms of bandwidth and Input/output Operations Per Second (IOPS). The storage architecture, in conjunction with the booster compute node design and its GPUDirect capability, makes it possible – in principle – to improve IO bandwidth and reduce IO latency towards the GPUs, thereby improving the performance for a significant number of use cases.

The Fast tier (5.4 PB) is based on DDN Exascaler5 and will act as a high-performance tier specifically designed to support high IOPS workloads. This storage tier is completely full flash and based on NVMe and SSD disks, and therefore provides high metadata performance, which is especially critical for AI workloads and, in general, when many files creation are required. A wide set of options are available in order to integrate the fast and capacity tiers to make them available to end-users.

The capacity tier (106 PB) will provide the parallel file system, which will be based on Lustre. Some key security features that will be evaluated are i) Lustre multi-tenancy, available from version 2.10, aiming to improve security and isolation so that only authenticated users can access a selected portion of the storage namespace; and ii) Lustre encryption at rest, available from version 2.14, based on cryptofs, transparent to the parallel filesystem and able to handle file multiple access (multiple clients).

	Storage Fast Tier	**Capacity Tier**
Net capacity	5.4 PB	106 PB
Disk technology	Full flash (NVMe and SSD)	NVMe and HDD
Bandwidth:	Aggregated: 1400 GB/s r/w io500 score: 676 GiB/s	Aggregated: read performance of 744 GB/s and write performance of 620 GB/s io500 score: 197 GiB/s

Table 5: Leonardo storage tiers

8.3.5 Software ecosystem

All the nodes are equipped with Red Hat Enterprise Linux 8; the scheduler provided is SLURM. In addition, the vendor provides the system with some preinstalled software.

- Several compiler suites are preinstalled: Intel Parallel Studio Cluster Edition to optimise the performance of CPU-based applications, and NVIDIA HPC SDK to take advantage of all the GPU features on the Booster Module.
- Debugger and profilers are included: in particular, the Arm Forge Ultimate Suite, which includes Arm DDT for CPU and GPU debugging, will be installed, as well as Arm MAP and Arm Performance Reports for profiling, and the Intel Parallel Studio includes the Intel Debugger too. As for the NVIDIA tools, the graphical tool Nsight System is also provided for profiling.
- Some of the most critical optimised numerical libraries are provided through the Intel MKL library.
- Containerisation is supported through several different tools.
  - The Atos Containers Framework enables the management of containers via Singularity.
  - The NVIDIA Container Framework enhances the creation and automated deployment of containers.
  - SLURM integration is improved via Pyxis.
  - ParTec Parastation also supports the execution of containerised applications, improving the flexibility of a pure Singularity approach.
- Monitoring is enabled via the Atos SMC xScale suite, based on Prometheus, and using Grafana as the front end.
- Detection and tracking of issues are performed by the Parastation HealthChecker.

CINECA staff will provide additional software to satisfy requirements from scientific and industrial communities running programs on the machine.

8.3.6 Energy efficiency

Atos provides two different software tools which enable dynamical adjustment of power consumption. The Bull Energy Optimiser keeps track of energy and temperature profiles via IPMI and SNMP protocols. Such tools can interact with the SLURM scheduler in order to tune some of its specific features, like a selection of the jobs based (also) on the expected power consumption or a dynamical capping of the CPUs' frequencies based on the overall consumption.

This dynamical tuning procedure is enhanced by a second tool called the Bull Dynamic Power Optimiser, which monitors consumptions core by core in order to cap frequencies to the value that leads to the optimal balance between energy saving and time degradation for the application that is running.

When it comes to the GPU consumption, the NVIDIA Data Centre GPU Manager is provided. It throttles the GPU's clock when a custom threshold is exceeded. Atos also claims it to provides a CPU energy saving of ~17% with a related time degradation of ~2.8%.

8.3.7 Technopole data centre

The newly built data centre is located close to the city centre of Bologna, and it uses the buildings of an old tobacco manufacturing plant. The buildings were designed by Pietro Nervi and are part of Italian cultural heritage. Therefore, the original buildings were maintained and carefully re-adapted for their new scope. Figure 35 shows the site hosting the new Technopole data centre.

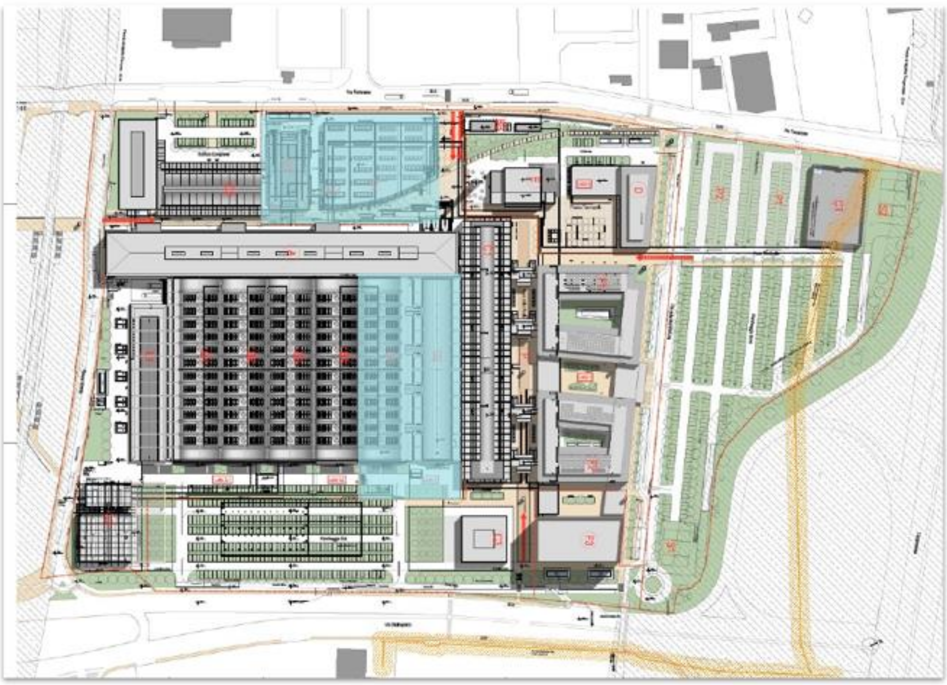

Figure 35: The site hosting the Technopole data centre. Highlighted In light blue the data hall (bottom) and the building hosting the mechanical and electrical infrastructure (top)

CINECA and the INFN new data centre will follow a two-stage evolution plan. In the first stage (2021-2025), the data centre will feature 10 MW of IT load, 1240 m^2 of computing room floor space, 900 m^2 of ancillary space, a direct liquid cooling capacity of 8 MW, chilled water (18° - 23° C) cooling capacity of 6 + 2 MW and power capacity of 3 + 1 MW (no-break) and 9 + 3 MW (short-break).

	Phase 1 (2021-2026)	Phase 2a (2026-2031) *Liquid Cooling Expansion*	Phase 2b (2026-2031) *Air Cooling Expansion*
IT Loads	~10 MW	~20 MW	~20 MW
Computing Room	1,240 m^2	3,840 m^2	3,840 m^2
Ancillary Spaces	900 m^2	900 m^2	900 m^2
Cooling Capacity (Tempered water ~37-47 °C) *	8 MW	16 MW	8 MW
Cooling Capacity (Chilled Water 18-23 °C) **	6+2 MW	6+2 MW	12+4 MW
No-Break Power Capacity (UPS) **	6 + 2 MW	12 + 4 MW	12+4 MW
Short-Break Power Capacity (GEN) **	7.5 + 2.5 MW	15 + 5 MW	15 + 5 MW

Table 6: CINECA infrastructure in different phases

* No redundancy
** 3 + 1 redundancy

The second stage (2025-2030) will see an increase to 20 MW IT load and an additional 2,600 m^2 of computing room space floor will be available while having the mechanical and electrical infrastructure able to comply with two different expansion strategies: stage 2a: *Liquid Cooling Expansion* (16 MW direct liquid-cooled + 4 MW air-cooled), or stage 2b: *Air Cooling Expansion* (8 MW direct liquid-cooled + 12 MW air-cooled).

In designing the data centre, particular care was devoted to limiting the PUE that, for Leonardo, is estimated below 1.1. Simulation of the PUE was based on the expected operation power conditions, assumed at 65% of the peak power conditions[e] on loads and losses calculated for all systems (IT, mechanical and electrical) and on Bologna's historical external conditions, following a strategy compliant with Level 3 Green Grid/ASHRAE. Therefore, the yearly average PUE was projected with the following calculation.

$$PUE = \frac{IT\ Loads + Cooling\ Loads\ (dry\ cool., pumps, chillers) + Losses\ (transformers, UPS, distrib.)}{It\ Loads} < 1.1$$

[e] Peak Power Condition (PPC) is only reached when high-performance LINPACK is run over all the system. A value of 65% of the PPC is typical during production operational conditions.

8.4 MareNostrum 5

Javier Bartolomé, BSC

The Barcelona Supercomputing Center is a public research centre located in Barcelona, Spain, providing supercomputing services as well as performing R&D in various disciplines. BSC is a consortium based on the Spanish government (60%), Catalan Government (30%), and the Technical University of Catalonia (10%). Currently, there are more than 700 employees at BSC facilities conducting research and supporting the HPC services.

Figure 36 represents the history of the MareNostrum system, different installation phases of which have been deployed inside a disused chapel since 2004.

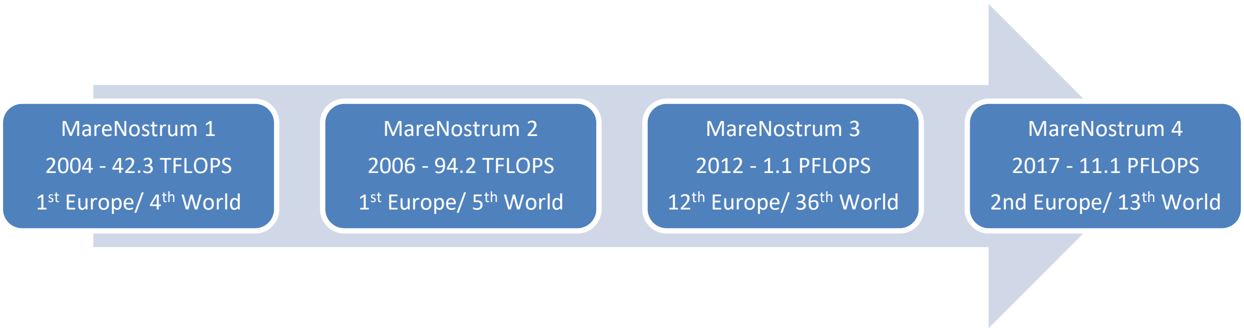

Figure 36: History of MareNostrum

In 2004, the data centre had a power capacity of 3 MVA, which was upgraded to 5 MVA in 2012. The power consumption of the system increased throughout the subsequent installation phases of MareNostrum, which reached the limit of the cooling capacity, despite continuous efforts aimed at increasing the cooling capacity. Table 7 outlines the details of the chapel's evolving infrastructure.

System	Year	Total power	IT (KW)	Cooling Capacity (KW)	IT Cooling	DC Cooling
MN1	2004	3x1 MVA	650	4x Chillers 235 kW	Air	CRAHs: 10 x 75,5 kW
MN2	2006	3x1 MVA	750	5x Chillers 235 kW	Air	CRAHs: 8x 75,5 kW CRAHs: 2x 146,2 kW
MN3	2012	2x2 MVA 1x1 MVA	1080	5x Chillers 235 kW 2x Chillers 513,8 kW	RDHX, Air	CRAHs: 6x 75,5 kW CRAHs: 2x 146,2 kW HxB: 2x 1400 kW (1+1)
MN4	2017	2x2 MVA 1x1 MVA	1300	5x Chillers 235 kW 2x Chillers 513,8 kW	RDHX, Air	CRAHs: 6x 75,5 kW CRAHs: 2x 146,2 kW HxB: 2x 1400 kW (1+1)

Table 7: MareNostrum - Chapel infrastructure

This cooling capacity limit eventually meant that the existing building infrastructure was not adequate. At the time of the presentation, BSC was in the process of moving to new headquarters, the construction of which had going to be finalised by the end of 2020, but was delayed to the first quarter of 2021 due to the pandemic. Figure 37 shows BSC's new building, which was built next to the chapel (see Figure 38) and will host the MareNostrum 5 Pre-Exascale EuroHPC system.

Figure 37: BSC new headquarters

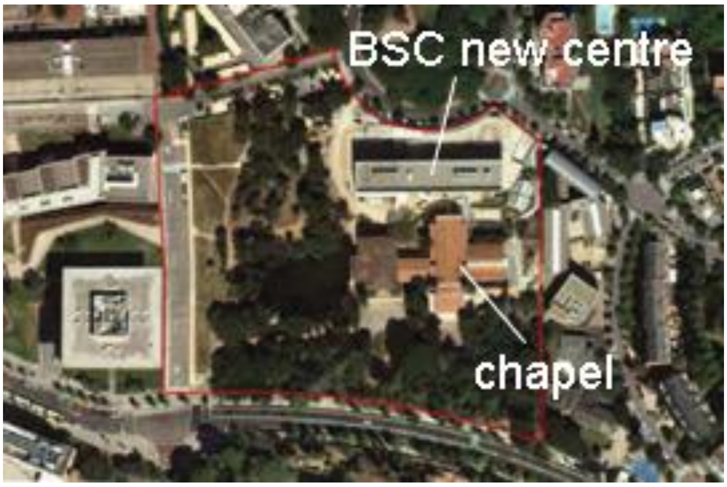

Figure 38: Location of BSC new headquarters

8.4.1 MareNostrum 5: A European pre-exascale supercomputer

MareNostrum 5 is a EuroHPC Joint Undertaking consortium joined by Spain (hosting entity), Portugal, Croatia, and Turkey, with a total investment of 217 million Euro: 150 million Euro for capital investment, with the rest being allocated to operational costs and additional projects[f].

The peak performance goal for the MareNostrum 5 system is set to 200 petaflops, which comes from the EuroHPC objective for the system.

The system will have a General-Purpose (GP) compute partition. Based on the feedback received from the user community, there is still a need to support researchers with general-purpose processors. This, on the other hand, makes it rather challenging to achieve the objectives set by the EuroHPC JU for having a system with a high peak performance. For that reason, an accelerated compute partition was added, thus making the system heterogeneous. Certain maximum and minimum performance limits were set for each of the partitions. For example, for accelerated partitions, it was decided to have a maximum sustained performance of 110 petaflops when running High-Performance LINPACK tests. In addition, the system will feature emerging compute technologies (e.g., new general-purpose technologies and new accelerated technologies) together with high-performance storage and a very high bandwidth interconnect. The end-goal objective is to ensure that the system is compatible with existing

[f] The acquisition and operation of the system will be funded jointly by EuroHPC JU, through EU's Connecting Europe Facility and the Horizon 2020 research and innovation programme, as well as participating states Spain, Portugal, Croatia, and Turkey.

applications and to provide assistance to efficiently solve problems arising from various domains, ranging from the earth and life sciences through engineering to AI and AI-driven applications.

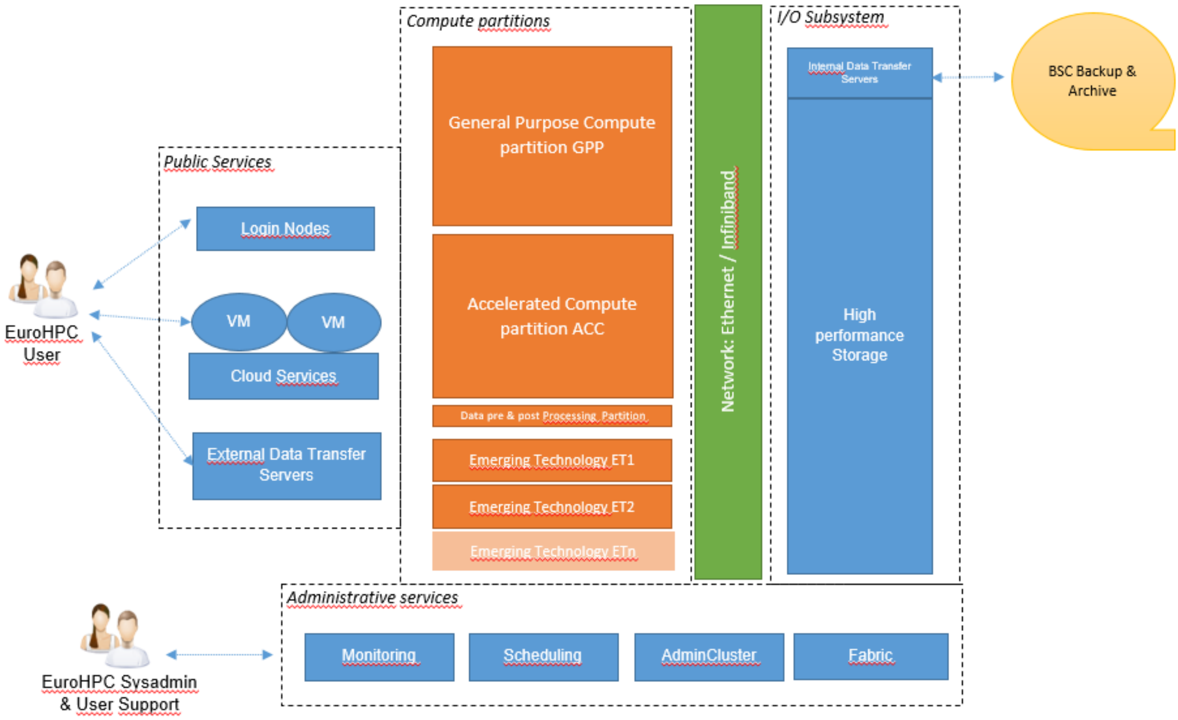

Figure 39: MareNostrum 5 system overview

For achieving this, BSC needed to upgrade its power supply feed and substation with a total capacity of 110 kV and 2 × 25 MVA, which could be expanded, if necessary, to 2 × 40 MVA. The expected initial power consumption was 20 MW. The first installation phase included installing a power line of 1,145 m connecting BSC to the general power grid. This underground connection is expected to be ready by the time the system is installed. At the time of writing, it was estimated that the system would be running in full production towards the end of 2021. The newly built substation will have a total capacity of 4,288 m^2 with an area of 42 m × 26 m and a depth of 18 m. The objective of the substation is to transform 110 kV to 25 kV. There will also be an emergency line of 5 MVA installed. One of the requirements set in the tender was to ensure that all energy provided would, by contract, be green energy. The list below outlines further tender requirements set for the MareNostrum 5 infrastructure.

- The power consumption should not exceed 12 MW when running HPL.
- The PUE should be below 1.08.
- Certain requirements per rack were also indicated, such as power dissipation, weight and dimensions, cabling, visibility, etc. One of the essential requirements was to ensure that each rack dissipates a minimum of 95% of the heat that is generated.
- Cold-water loop 18 °C, up to 1 MW
- Warm-water loop 35 °C, up to 12 MW
- Have an exhibition centre for MareNostrum 5, i.e. have a facility that is visitable

Under these conditions, a public tender was conducted. The tender was awarded on the 1st of August 2019 and formalised on the 26th of November 2019. The tender was awarded to Climava SL[8] with an award price of 12,557,990 EUR (excluding VAT). The expected date for the

[8] https://www.climava.com/

delivery of the site was initially set for September 2020. Due to the impact of COVID-19, that was later shifted to April 2021, meaning that no HPC system would be installed at the BSC facilities before that date. Table 8 indicates the available space for hosting the MareNostrum 5 system.

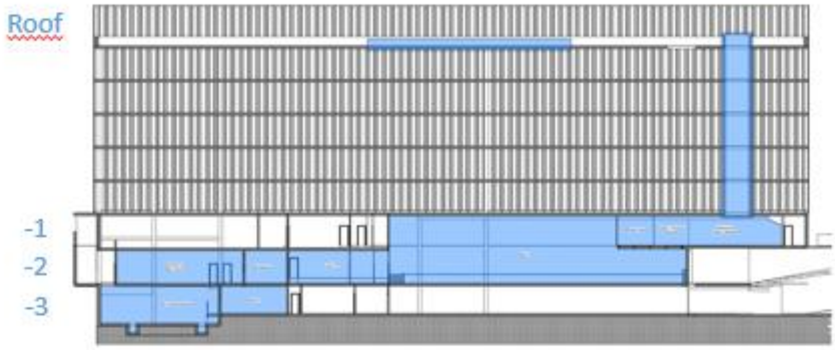

Figure 41: Space available for MareNostrum 5

As BSC intends to keep this facility operational for at least 15 years (as was previously the case with the chapel), the electrical loads were designed accordingly.

Figure 42 illustrates the supporting infrastructure. As can be seen, there are five transformers shown in the bottom part of Figure 42 that transfer the electrical energy to the main distribution panel.

Floor		m²	Total
Roof		320	320
-1	Chillers & Pumps room	466	711
	Riser / *"PATIO"*	9	
	Visitors area	236	
-2	Compute Room	847	1374
	Access to compute room	46	
	Batteries room	73	
	Low voltage room	408	
-3	Transformers	426	470
	Fire extinction	49	
Total		rounded	2875

Table 8: Overview of MareNostrum 5 floor space occupation

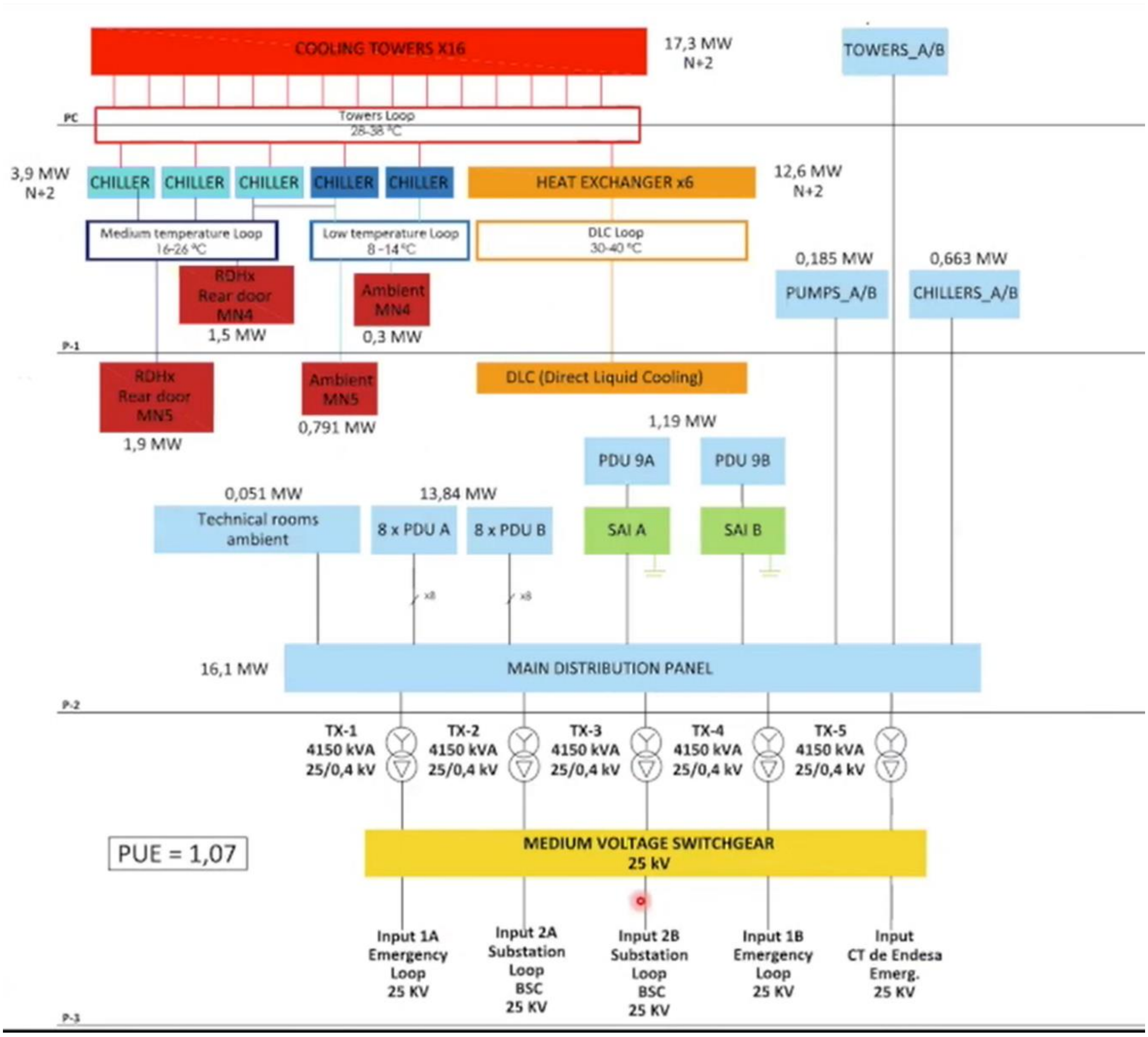

Figure 42: Mare Nostrum 5 infrastructure diagram

The main power distribution panel, in its turn, feeds the PDUs and UPSes[h]. From the upper part of Figure 42, it can be seen that the cooling towers (14 + 2) are operating with a temperature of 38 °C for the inlet and 28 °C for the outlet. As illustrated in the diagram, the water then flows via heat exchangers to the compute servers. Figure 43, Figure 44 and Figure 45 show different areas of the building and the progress on the construction of the different areas at the time of writing.

The plan was to decommission the current cooling infrastructure of MareNostrum 4 in July 2021 and start using the cooling facilities of the new building for that purpose. It was also expected that the construction of the computer room would be finished at the end of the same month. In addition, there are two other projects planned which will focus on reducing water usage in the infrastructure (based on the use of groundwater instead of an industrial water supply) and an osmosis plant to maximise water re-use in the cooling towers.

[h] Labelled SAI (Spanish word for UPS) in Figure 42.

Figure 43: Roof of MareNostrum 5

Figure 44: MareNostrum 5 heat exchanger, chiller and pumps

Figure 45: MareNostrum 5 computer room

8.5 Installation Requirements and Best Practices for Hosting Exascale Systems

Hayk Shoukourian (LRZ)

This presentation was based on the PRACE 6IP/WP5 deliverable D5.1.

Supercomputing centres continuously aim to improve their resources to facilitate better service provision to the HPC community. In terms of this objective, HPC data centres position themselves at the forefront of computing technologies, with regular system hardware and software updates making it possible to ensure research can be undertaken continuously, thereby enabling further scientific discoveries.

Accommodating these updates on a regular basis is, however, not a trivial task, as planning and commissioning a data centre with a lifespan of several years that is capable of hosting diverse supercomputers belonging to different generations is quite a challenge.

This presentation discusses installation requirements and best practices for hosting next-generation HPC systems (i.e. pre-exascale and exascale), with a specific focus on the current European ecosystem. More specifically, the presentation outlines foreseeable major challenges related to a wide spectrum of aspects relevant to the building infrastructure of data centres: ranging from the hosting facility, through power availability and usage of renewable energy sources to cooling solutions and waste heat reuse; best practices for addressing those issues are also presented. For more details, please, see the deliverable [9].

8.6 Reuse by the Paris-Saclay District Heating and Cooling Network of the Waste Heat from the IDRIS Supercomputer

Rafael Meideros, IDRIS, and Nicolas Eyraud, EPAPS

The Institute for Development and Resources in Intensive Scientific Computing (IDRIS) is the main computing centre of the French National Centre for Scientific Research (CNRS). It was founded in 1993, and since 2007 it has hosted the French Tier-1 systems funded by Grand équipement National de Calcul Intensif (GENCI). It is located in the Science University Paris-Saclay (~20 km south-west of Paris).

The operation of the current Tier-1 system, Jean Zay, started in January 2020. This machine, with an initial peak performance of 16 petaflops, was upgraded to 28 petaflops in mid-2020. Jean Zay is an HPE SGI 8600 supercomputer composed of a partition containing scalar nodes and a partition with accelerated nodes using NVIDIA Tesla V100 GPUs. The nodes are connected to an Omni-Path Architecture (OPA) network. The power consumption measured for HPL is 2.1 MW. It should be noted that the electrical line feeding the site has a capacity of 10 MVA and the capacity of the transformers is 6.2 MVA while the UPSes are able to provide 2.5 MW.

In terms of cooling, Jean Zay uses the following two types of cooling.

- Direct warm-water cooling for the CPU and GPU: The inlet temperature is in the range 30-32 °C, and the outlet temperature is in the range 35-41 °C.
- Air-cooling (with rear door heat exchangers) for memory network and storage: The inlet temperature is 12 °C, and the outlet temperature is 18 °C.

At IDRIS, heat-reuse is implemented at the following two levels (see Figure 46).

- At the level of the chillers for the cold-water loop used for air cooling, which can provide up to 300 kW with a temperature of 52 °C: This heat is used for heating IDRIS itself and a set of nearby buildings.
- At the level of the adiabatic heat exchangers for the warm-water loop, which can provide up to 1,200 kW at a temperature of around 35 °C: This heat is transferred to the Paris-Saclay district heating and cooling network.

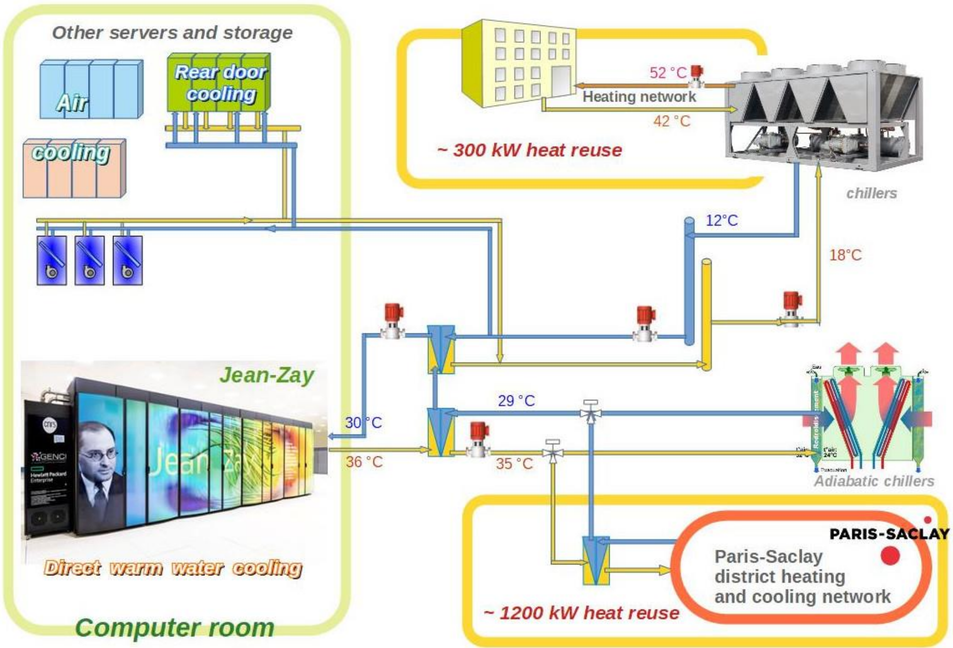

Figure 46: Cooling and heat reuse of the Jean Zay supercomputer

As explained earlier, IDRIS is located at the Paris-Saclay campus, a world-class cluster for innovation in the Metropolis of Greater Paris, or Grand Paris (see https://www.epaps.fr/). The development of this campus is very intensive; the goal is to gather 15% of French research there with 265,000 employees, 65,000 students and 430,000 inhabitants.

In order to provide heating and cooling for offices, university buildings, student housing, and residential housing, it was decided to create a District Heating and Cooling (DHC) network able to reuse waste heat from industry, laboratories and data centres and to take advantage of an aquifer with a temperature of 30 °C (700 m deep). An ambient temperature network (30 °C) was chosen to reduce losses and be compatible with the temperature of the available heat sources.

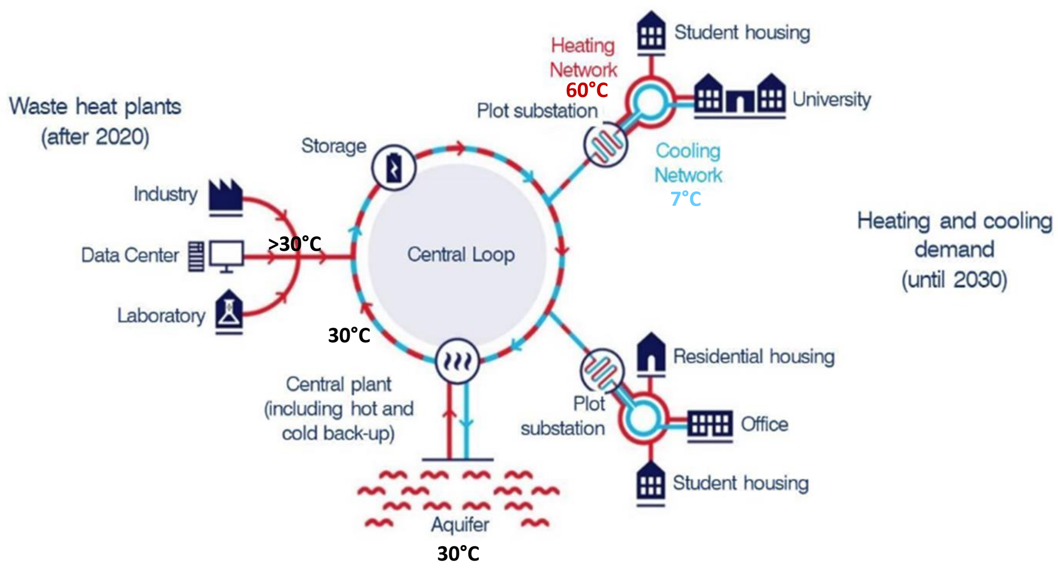

Figure 47: The Paris-Saclay DHC network

Heat pumps are used to raise the temperature if needed, and boilers provide backup when needed. The expected DHC energy mix is 60% renewable and waste heat recovery, 36% electricity and 4% from gas boilers.

At this stage, due to the large size of the Paris-Saclay campus, two networks are in operation: one for the Moulon area, and one for the École Polytechnique area. At a later stage, it could be interesting to connect the two networks. IDRIS will be connected to the Moulon area network in early 2022. A substation of the DHC network will be built for this purpose in which a heat exchanger will transfer heat from IDRIS (Figure 47) to the DHC network. With this substation, 4,000 MWh of heat per year (which is equivalent to the annual needs of approximately 1,000 households) is expected to be transferred to the DHC network. The cost of this substation will likely be amortised in 20 years. The cost of the substation is 665 thousand euros, while the overall cost of the DHC network (2015-2022) is 55 million euros. (It is expected that this infrastructure will be expanded further with the amount invested in it doubling by 2030.)

The following are several of the key factors contributing to the success in reusing the waste heat from IDRIS in the Paris-Saclay DHC network.

- Favourable heat transfer conditions:
  - The ambient temperature of the DHC network (30 °C) is compatible with the outlet temperature of the warm water loop cooling the supercomputer.
  - Waste heat sources are available all year long (so there will be no or rare maintenance shutdowns).
- Limited investments:
  - The location of IDRIS is very close to the DHC network (~500 m).
  - Space is available for a new substation in the IDRIS enclosure.
- Strong and public project support:
  - As these are public investments (EPA Paris-Saclay and CNRS), there are lower investment return requirements.
  - The ADEME national heat fund supports heat reuse projects.
  - Both CNRS and EPA Paris-Saclay have low-carbon strategies.
  - The IDRIS/CNRS and EPA Paris-Saclay teams are both strongly involved.

9 Conclusions

The presentations that were given during the 11th European Workshop on HPC Infrastructures and the plenary discussions following these presentations revealed some important trends in developments of technologies for cooling, heat-re-use, power provisioning and power management, and tooling for integral data centre management.

9.1 Situation in Europe

The EuroHPC Joint Undertaking in Europe has provided an enormous boost to HPC research and development in Europe catalysed by the new large HPC resources that have been made available to academic and business researchers. This was clearly demonstrated by many of the talks at this workshop. However, some of the installations were delayed by the pandemic situation all over the globe. This expansion in HPC resources means that there are many infrastructure problems on a level that has not been seen before in Europe, and it was beneficial to discuss those issues to see how they are being handled by different data centres.

9.2 Trends with Respect to Power

Despite the progress that has been made in terms of flops per watt in modern CPUs and GPUs, most centres envisage that they will need more power in the coming years to supply a level of HPC capacity that meets the demands of researchers from both academia and industry. Since the larger data centres have a power draw of several tens of megawatts, the potentially large impact of their machines on the regional electricity grids still figures on the agenda. Maximising energy efficiency may be constrained by the need to operate in a more grid-friendly way, or measures must be taken to improve “grid-stiffness”. There may be financial benefits in scheduling jobs in a more predictable grid-friendly way by keeping the power consumption within a specific window. In last year’s HPC Infrastructures workshop, people had already concluded that the pricing of electricity is often such that it is cost-optimal to strive for a predictable power draw that fluctuates between predetermined lower and upper bounds. In this year’s workshop, it was demonstrated that sites – with the help of their monitoring and subsequent analyses of application runs – can predict the power usage of the applications that are sitting in the queue waiting to be scheduled and can take this into account. Thus, they can actively schedule to stay in the “safe” area of power usage between the upper and lower bounds. Many centres today operate with the compute part of the equipment connected directly to the local electricity grid. When it comes to storage resources, it is still necessary to have a battery power backup supply provided by UPS systems. With the introduction of more power coming from solar and wind sources in national and regional power grids, more fluctuations in the amount of power supplied by the grid are to be expected. This workshop showed that developments in UPS technology and selecting the right UPS for the task can reduce the power loss in the UPS system to a minimum.

9.3 Trends with Respect to Cooling and Heat Re-use

Air-cooled solutions for the compute part of the current state of the art HPC systems are generally deemed infeasible. Direct liquid cooling solutions are predominant for compute systems in larger installations. while air-cooling is still dominating for storage. Most centres, including the largest centres in Europe and elsewhere, make direct liquid cooling mandatory in their procurements for new large HPC systems. Re-use of waste heat has been employed in northern Europe for a long time due to the cold climate and the existing district heating (and in some cases also district cooling) networks. Now, several new HPC installations are taking advantage of heat re-use, and in several instances, they use district heating and cooling networks. New district heating networks are also being set up. Several new installations using adsorption chillers that can transform waste heat into cooling for buildings or other computers were also presented in this workshop.

9.4 Monitoring and Control

The predominance of warm-water-cooled high-density systems has led to tighter coupling between the management of the actual facility and the management of the HPC systems. Most facility management teams have come to the conclusion that broad instrumentation of all components of the data centre and the HPC system and its operating system and job scheduling software is a basic requirement for improving energy efficiency. This, however, leads to large data streams from multiple sources that must somehow be integrated and aggregated into meaningful and manageable overviews. Many sites have invested in tooling that they can customise for their data centre and HPC systems. There is no adequate out-of-the-box commercial product available. Most centres build their own integrated data centre management tool, relying heavily on open-source projects – such Apache Kafka, Apache Cassandra, Grafana, and OpenDCIM – to provide core blocks of functionality. Sites are also trying to go beyond mere monitoring. Broad instrumentation, along with integrated high-frequency data capturing and analyses, open prospects of automating control, including machine learning for feeding back decisions into facility control systems and/or the job scheduling environment.

9.5 Final Remarks

The 11th HPC Infrastructures workshop was very successful in bringing together experts from various disciplines and who have various stakeholder roles working on HPC site infrastructures. Figure 48 shows that the annual workshop, which is by invitation only, has become an institution that is capable of consistently attracting a stable number of experts in the field. The next workshop in the series will be hosted by CSC in Finland and is planned to take place in June 2022. These workshops are important for sharing best practices, and they contribute to sustaining a high level of expertise in the operation and management of HPC data centres and also to spreading advanced technologies more widely amongst European HPC centres. The EuroHPC Joint Undertaking programmes are leading to large increases in the size and complexity of HPC systems in Europe, which means it is becoming more and more important to discuss infrastructure issues and share solutions. The presentations at this year’s workshop confirmed that HPC centres in Europe are well able to host and operate large supercomputers in reliable and energy-efficient ways. The workshop activities are critical for making sure that EuroHPC can indeed deploy large systems in Europe and for identifying key technologies that may benefit from the research and infrastructure funded by EuroHPC.

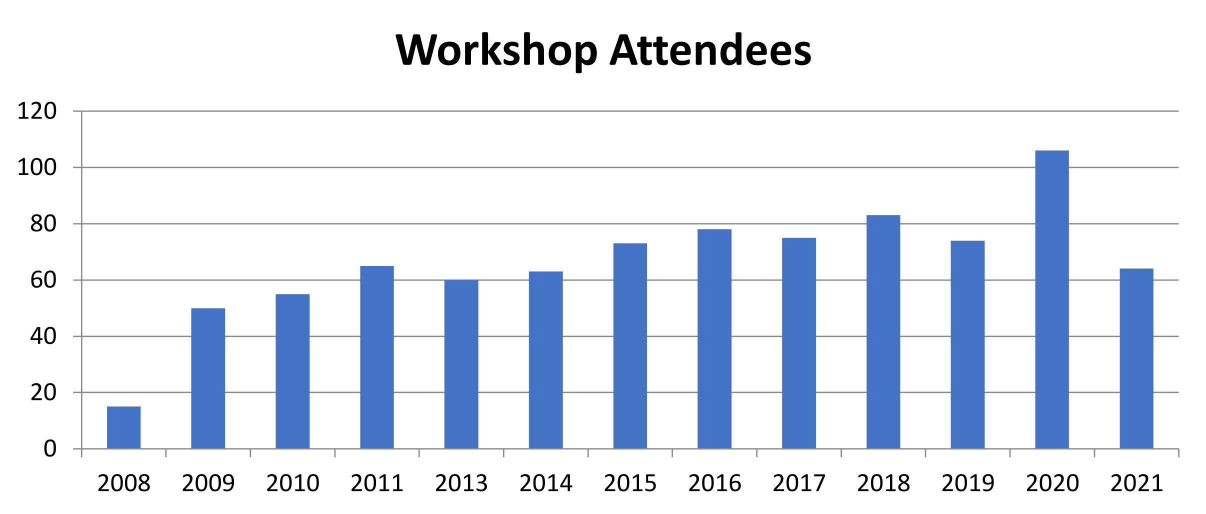

Figure 48: Number of attendees at the annual workshops on HPC Infrastructures

Acknowledgements

This work was financially supported by the PRACE project funded in part by the EU's Horizon 2020 Research and Innovation programme (2014-2020) under grant agreement 823767.

References

[1] EE HPC WG, "EE HPC WG home page," [Online]. Available: https://eehpcwg.llnl.gov/.

[2] LANL, "Los Alamos National Laboratory home page," [Online]. Available: https://www.lanl.gov/.

[3] GE Energy Consulting, "GE PSLF," [Online]. Available: https://www.geenergyconsulting.com/practice-area/software-products/pslf.

[4] ETAP, "ETAP home page," [Online]. Available: https://etap.com/.

[5] ASHRAE, "ASHRAE Technical Committee 9.9 home page," [Online]. Available: https://tc0909.ashraetcs.org/.

[6] ASHRAE TC 9.9, "Emergence and Expansion of Liquid Cooling in Mainstream Data Centers," [Online]. Available: https://tc0909.ashraetcs.org/documents/ASHRAE_TC0909_Emergence_and_Expansion_of_Liquid_Cooling_in_Mainstream_Data Centers_5_May_2021.pdf.

[7] PDC, KTH, "Press Release: New supercomputer coming to KTH!," [Online]. Available: https://www.pdc.kth.se/about/pdc-news/press-release-new-supercomputer-coming-to-kth-1.1053705.

[8] CINECA, "Leonardo Press Release: The World's Fastest European Technology-based Supercomputer for CINECA and Europe," 18 November 2020. [Online]. Available: https://www.cineca.it/area-stampa/comunicati-stampa/worlds-fastest-european-technology-based-supercomputer-cineca-and.

[9] PRACE, “PRACE IP Projects Public Deliverables,” [Online]. Available: https://prace-ri.eu/about/ip-projects/public-deliverables/#PRACE6IP.

Abbreviations

AC	Alternating Current
AFW	US Air Force Weather
AI/ML	Artificial Intelligence and Machine Learning
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BDVA	Big Data Value Association
BSC	Barcelona Supercomputer Centre
CEA	Commissariat à l’Energie Atomique et aux Energies Alternatives
CNRS	French National Centre for Scientific Research
DHC	District Heating and Cooling
DIMM	Dual In-line Memory Module
DRAM	Dynamic Random-Access Memory
DVFS	Dynamic Voltage and Frequency Scaling
ECMWF	European Centre for Medium-Range Weather Forecasts
EEHPCWG	Energy Efficient HPC Working Group
ERE	Energy Reuse Effectiveness
ETAP	Electrical Transient Analyzer Program
ETP4HPC	European Technology Platform for High Performance Computing
EuroHPC	JU EuroHPC Joint Undertaking
EWHPC	European Workshop on HPC Infrastructures
FIPS	Federal Information Processing Standard
FTA	Fault Tree Analysis
GENCI	Grand Équipement National de Calcul Intensif
GP	General-Purpose
HPC	High-Performance Computing
HPCM	HPE Performance Cluster Manager
HPE	Hewlett Packard Enterprise
HPL	High-Performance LINPACK
IDRIS	Institute for Development and Resources in Intensive Scientific Computing
INFN	Italian Institute for Nuclear Physics
INM	Intel Node Manager
IOPS	Input/Output Operations Per Second

ITT	Invitation To Tender
JUNIQ	Jülich Unified Infrastructure for Quantum Computing
KTH	Kungliga Tekniska Högskolan (Royal Institute of Technology)
LANL	Los Alamos National Laboratory
LCA	Life Cycle Analysis
LLNL	Lawrence Livermore National Laboratory
LRZ	Leibniz Supercomputing Centre
ML	Machine Learning
MTBF	mean time between failures
MTTR	Mean Time To Repair
MUS	Machine Unit Specification
NDA	Non-Disclosure Agreement
NDR	Next Data Rate
NERC	North American Electric Reliability Corporation
NERSC	National Energy Research Scientific Computing Center
NSC	National Supercomputer Centre, Sweden
NVMe	Non-Volatile Memory express
NVML	NVIDIA Management Library
ODA	Operational Data Analytics
ORNL	Oak Ridge National Laboratory
OMNI	Operations Monitoring and Notification Infrastructure
OPA	Omni-Path Architecture
OPR	Owner's Project Requirement
PDC	Center for High Performance Computing
PDU	Power Distribution Unit
PLC	Programmable Logic Controller
PPC	Peak Power Condition
PSLF	Positive Sequence Load Flow
PSNC	Poznan Supercomputing and Networking Center
PUE	Power Usage Effectiveness
QC	Quantum Computer
RAM	Random-Access Memory
RDHX	Rear-Door Heat eXchanger
RNX	Riedo Networks
SCADA	Supervisory Control and Data Acquisition

SMP	Symmetric MultiProcessing
SNIC	Swedish National Infrastructure for Computing
SVC	Static Var Compensator
TC	Technical Committee
Tcase	chip packaging temperature
TCO	Total Cost of Ownership
TDP	Thermal Design Power
UPS	Uninterruptible Power Supply
WAN	Wide Area Network