Specifications for the EuroFlash Cluster

November 7, 2025

Contents

1 Introduction **2**

2 EuroFlash Cluster Overview **3**

3 The EFC Offer **5**

4 Acceptance Procedure and Compliance Tests **8**

5 Requirements Conformity Tables **11**

6 Award Criteria **23**

7 Definitions and Conventions **24**

List of Tables

1 Applicability of Compliance Tests to Nodes . . . 10
2 Rack Requirements . . . 11
3 Common Requirements for All Node Types . . . 12
4 Additional Requirements for Compute and GPU System Nodes . . . 13
5 Specific Requirements for Compute Head Node . . . 15
6 Specific Requirements for Compute Production Nodes . . . 16
7 Specific Requirements for GPU Nodes . . . 17
8 Additional Requirements for Storage System Nodes . . . 18
9 Specific Requirements for Storage Head Node . . . 20
10 Specific Requirements for Storage Data Nodes . . . 21
11 General Terms and Conditions . . . 22
12 Award Criteria . . . 23

List of Figures

1 An overview of the EuroFlash Cluster layout. . . . 4
2 Compute Head Node overview . . . 5
3 Compute Production Node overview . . . 6
4 GPU Node overview . . . 6
5 Storage Head Node overview . . . 7
6 Storage Data Node overview . . . 7

1 Introduction

We are tendering for compute and storage infrastructure to analyse data in real time from the Low-Frequency Array (LOFAR[1]) radio telescope. LOFAR receives streams of data from radio antennas spread across the Netherlands and Europe. It is one of the largest and most scientifically capable radio telescopes ever built. Close to 1,000 scientific papers have been written using data from LOFAR, including discoveries that have been published in the world's leading scientific journals and which have appeared in major news outlets internationally! With the new "EuroFlash Cluster" (EFC) we are tendering here, we will sift through many petabytes of data to discover rare but scientifically interesting flashes of radio light that point the way to the most extreme stars in the Universe.

These flashes are called "Fast Radio Bursts" (FRBs), and they pose one of the most intriguing mysteries in modern astronomy: what produces these brilliant, millisecond-long flashes of radio waves that travel billions of light-years before reaching Earth? Whatever their origin, FRBs require extreme cosmic explosions capable of generating laser-like emission. While recent evidence suggests that FRBs are created by the ultra-magnetised stars known as "magnetars", other evidence suggests that this isn't the whole story. Maybe FRBs come from evaporating black holes or other exotic phenomena. If so, astronomers may be facing not one but several unsolved cosmic puzzles.

Studying these signals remains challenging: they are unpredictable and typically last only milliseconds. Expanding our search capability increases the chances of uncovering new types of signals and understanding where they come from.

The EuroFlash[2] project is funded by an Advanced Grant from the European Research Council (ERC) under the European Union's Horizon Europe research and innovation programme (grant agreement No. 101098079; Principal Investigator: Prof. Dr. Jason Hessels). The project aims to conduct a world-class search for cosmic signals, and is led by scientists at ASTRON[3] (the Netherlands Institute for Radio Astronomy) and Anton Pannekoek Institute for Astronomy[4] at the University of Amsterdam[5]. To support the EuroFlash project's ambitious goals, a dedicated EuroFlash Cluster (EFC) will be procured. This cluster[6] will be owned by ASTRON, and physically housed in the CIT Smitsborg data centre at the University of Groningen[7]. Managed by the EuroFlash team, the cluster will operate 24/7, enabling compute-intensive, high-throughput data processing tasks, including real-time operations that require substantial network and processing capabilities.

With EuroFlash, we will be pushing the boundaries of astrophysics, and we expect that this will lead to surprising discoveries. In addition to publishing these results in scientific journals, our team is experienced in disseminating our research to the broader public via press releases and associated multi-media visualisations and descriptions. Likewise, we look forward to promoting and publicizing the installation of the EuroFlash cluster and its "first light" observations, in partnership with the chosen vendor.

---

[1] `https://www.lofar.eu/`

[2] `https://astroflash-frb.github.io`

[3] `https://www.astron.nl`

[4] `https://api.uva.nl`

[5] `https://www.uva.nl`

[6] **Throughout this document, the term "cluster" refers to the EuroFlash Cluster, unless explicitly stated otherwise.**

[7] `https://www.rug.nl`

2 EuroFlash Cluster Overview

An overview of the cluster layout and its external connectivity is shown in Figure 1. All the components inside the purple dashed rectangle in the figure (excluding only the network cables) are included in this call for procurement.

The cluster consists of three main systems:

a) **"The Compute System":** The compute system is a CPU cluster consisting of (i) one compute head node; and (ii) several compute production nodes (as offered by the vendor), as illustrated in Figure 1. The nodes use the dual-port 100 GbE network to exchange the necessary data. All compute system nodes must be PCIe Gen 5 based, single CPU socket machines.

   The **compute head node** (Figure 2) will serve as the login (user) node, manage cluster job scheduling, and host the home and software directories.

   The **compute production nodes** (Figure 3) will act as the main work horse to execute the necessary computations using a custom, in-house developed software pipeline for astronomical data processing.

b) **"The GPU System":** The GPU system is a CPU/GPU cluster consisting of six identical nodes, as shown in Figure 1. Each of these nodes also hosts two NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs, as shown in Figure 4. These nodes use the dual-port 100 GbE network to exchange the necessary data. The nodes in the GPU system serve as the primary workhorses for performing real-time stream computations using custom, in-house developed correlation software that leverages both CPUs and GPUs. All GPU nodes must be PCIe Gen 5 based single CPU socket machines.

c) **"The Storage System":** The storage system consists of one storage head node and three storage data nodes, as shown in Figure 1. The nodes use the 200 GbE network to exchange the necessary data. All nodes in the storage system must be PCIe Gen 4 based (or higher) single CPU socket machines. It is important to note that the storage system must be equipped exclusively with high-endurance NVMe SSDs for data storage purposes, as the data will be overwritten multiple times each day as observations continue to stream in. The storage system is expected to handle an average daily data ingest of approximately 2 PB (including the overwrites). Boot drives are exempt from the high-endurance SSD requirement.

   The **storage head node** (Figure 5) will handle the file structure and function as a file server.

   The **storage data nodes** (Figure 6) will be used to hold bulk data.

It must be noted that a few components specified in this document, such as particular models of NVIDIA GPU and AMD CPUs, are required to be of a specific make and model. This requirement is critical to ensuring compatibility and optimal performance with our in-house developed, resource-intensive astronomical data processing pipelines developed over a period of several years. These software suites include real-time stream processing capabilities designed for CUDA-enabled GPU architectures. Furthermore, the cluster will be deployed within the existing HPC infrastructure of the LOFAR telescope. To ensure seamless integration and interoperability within this established environment, the specified hardware components are essential. Any deviation from these models is likely to introduce compatibility issues with key software dependencies and overall architecture, leading to substantial re-engineering, testing, benchmarking, and validation efforts.

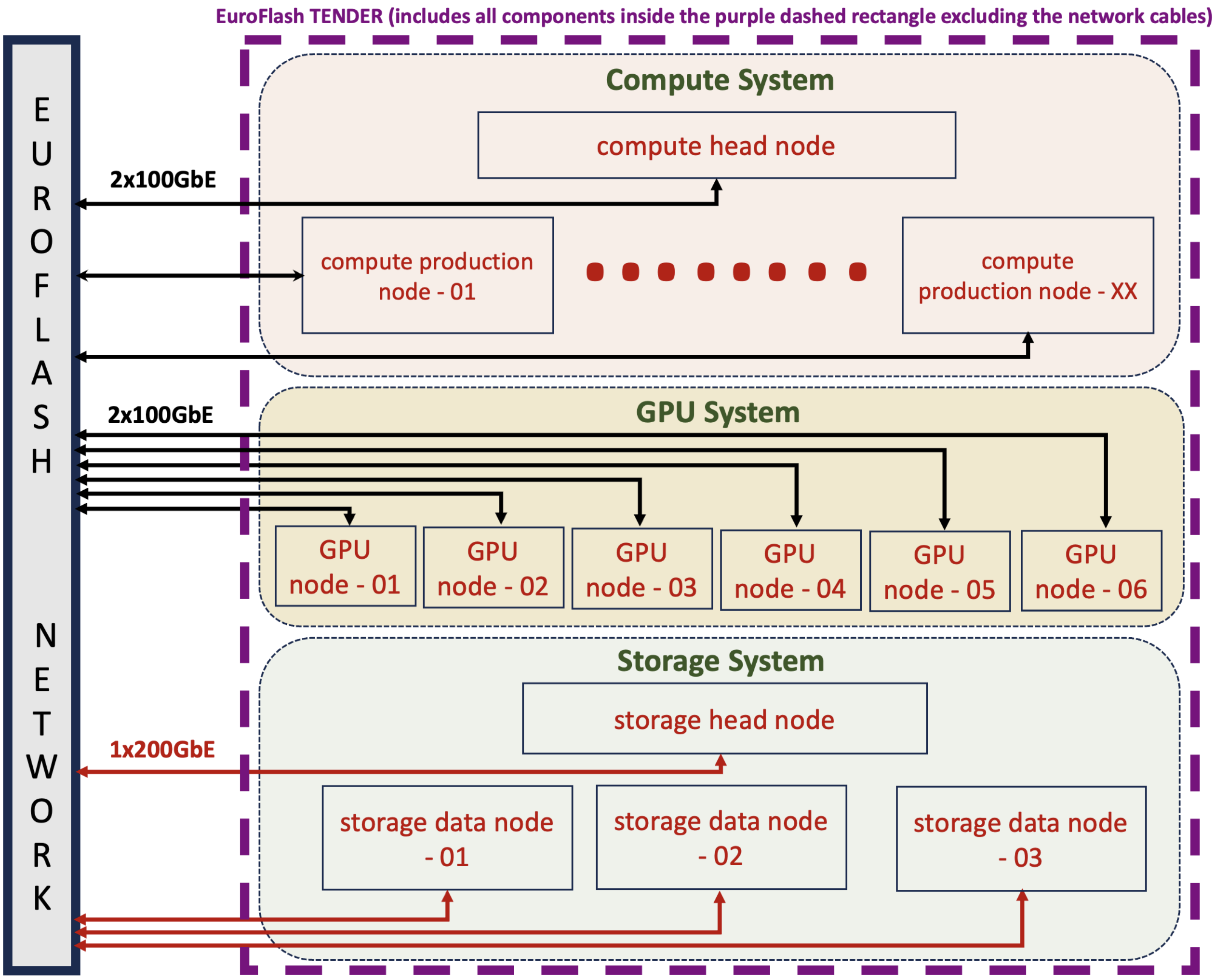

Figure 1: **An overview of the EuroFlash Cluster layout**. All components enclosed within the purple dashed rectangle (excluding network cables) are included in this procurement. Network elements shown outside the dashed box are for context only and are not part of this tender. The EuroFlash cluster consists of three systems, namely: (i) the CPU-based **Compute System**, consisting of one compute head node and several compute production nodes; (ii) the **GPU System**, composed of six identical GPU nodes; and (iii) the **Storage System**, which includes one head node and three data nodes. All nodes are interconnected via high-speed 100 GbE or 200 GbE networks, illustrated using different colors. All nodes must include at least one dedicated server-grade 1 Gbps (or higher) OoBM interface supporting remote management via IPMI, Redfish, or an equivalent protocol; for clarity, this connection is not depicted in the figure, though it is required for all nodes. Note: Although not shown here, the offer must include: (i) two spare SSDs compatible with the compute and GPU systems (DWPD$\geq$1); and (ii) two spare high-endurance SSDs compatible with the storage system (DWPD$\geq$3). This total of four spare SSDs (for the entire cluster and not per node) will be stored on-site to allow immediate replacement in the event of SSD failures within the cluster.

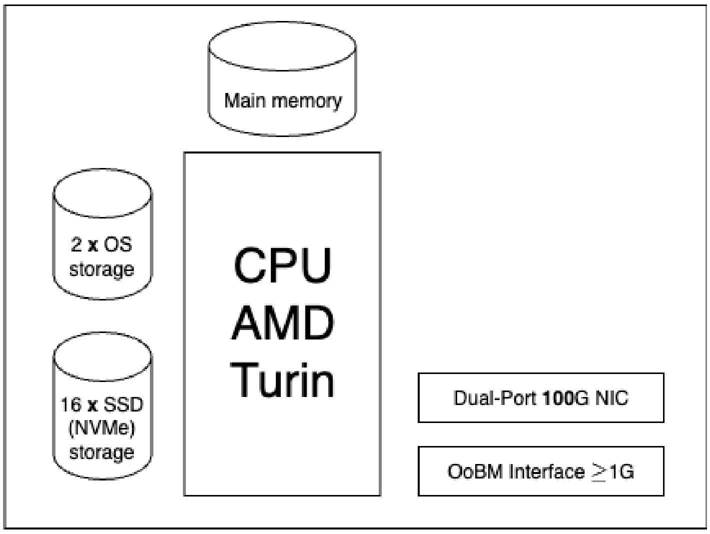

Figure 2: **Compute Head Node overview**. The compute head node is a PCIe Gen 5 based single CPU socket server having one physical CPU. It must include at least two enterprise-grade, flash-based storage devices suitable for use as boot drives in a RAID-1 configuration. Additionally, the node must be equipped with 16×NVMe SSDs for storage. The node must include a single general-purpose, server-grade, 100 Gbps full-duplex dual-port (2×100 GbE) Ethernet interface. The installed NIC must actually deliver sustained throughput of at least 90% of full line rate on both ports simultaneously, as integrated within the node. It must also include at least one dedicated server-grade 1 Gbps (or higher) OoBM interface supporting remote management via IPMI or Redfish (or equivalent).

In accordance with the principle of proportionality, these hardware specifications have been limited to only those components where deviation would result in disproportionate performance, technical, and financial burdens. The additional cost, time, human resources, and risk associated with adapting to alternative hardware falls outside the scope and budget of this procurement. Therefore, all vendors must adhere to the stated hardware specifications and submit proposals that incorporate the explicitly designated components wherever specified in this document.

3 The EFC Offer

The offer from the vendor must include :

1. One compute head node.
2. The offered number of compute production nodes. All the compute production nodes must be identical to each other.
3. Six GPU nodes. All the GPU nodes must be identical to each other.
4. One storage head node.
5. Three storage data nodes. All the storage data nodes must be identical to each other.
6. The offer must include: (i) Two spare SSDs compatible with the Compute and GPU Systems (CG.16); and (ii) Two spare high-endurance SSDs compatible with the Storage System (S.17). This total of four spare SSDs (for the entire cluster and not per node) will be stored on-site to allow immediate replacement in the event of SSD failures within the

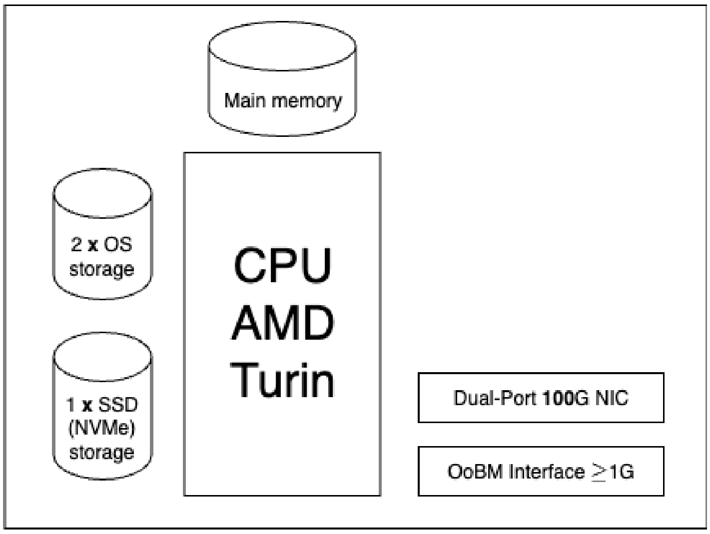

Figure 3: **Compute Production Node overview**. A compute production node is a PCIe Gen 5 based single CPU socket server having one physical CPU. It must include at least two enterprise-grade, flash-based storage devices suitable for use as boot drives in a RAID-1 configuration. Additionally, the node must be equipped with 1×NVMe SSD for storage. The node must include a single general-purpose, server-grade, 100 Gbps full-duplex dual-port (2×100 GbE) Ethernet interface. The installed NIC must actually deliver sustained throughput of at least 90% of full line rate on both ports simultaneously, as integrated within the node. It must also include at least one dedicated server-grade 1 Gbps (or higher) OoBM interface supporting remote management via IPMI or Redfish (or equivalent).

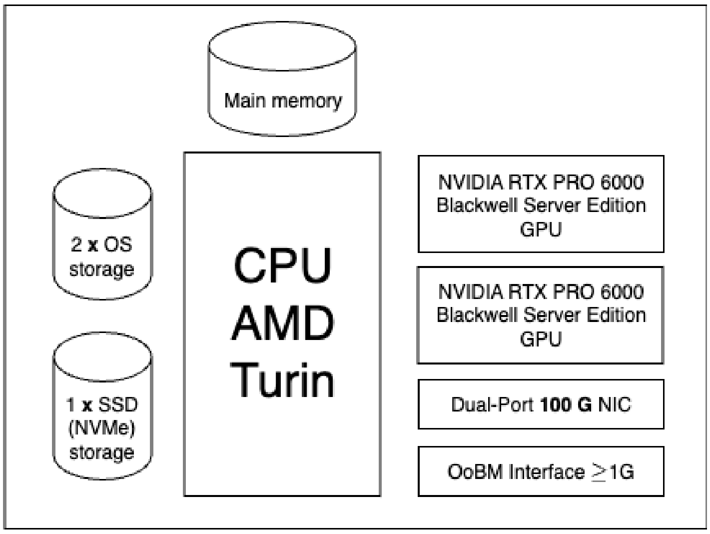

Figure 4: **GPU Node overview**. A GPU node is a PCIe Gen 5 based single CPU socket machine having one physical CPU. It must include at least two enterprise-grade, flash-based storage devices suitable for use as boot drives in a RAID-1 configuration. Additionally, the node must be equipped with 1×NVMe SSD for storage. The node must also include 2×NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs. The PCIe connection between CPU socket and any GPU must consist of at least 16×PCIe Gen 5 lanes. The connections between CPU socket and any GPU cannot be shared with other GPUs, NICs or peripherals. The node must include a single general-purpose, server-grade, 100 Gbps full-duplex dual-port (2×100 GbE) Ethernet interface. The installed NIC must actually deliver sustained throughput of at least 90% of full line rate on both ports simultaneously, as integrated within the node. It must also include at least one dedicated server-grade 1 Gbps (or higher) OoBM interface supporting remote management via IPMI or Redfish (or equivalent).

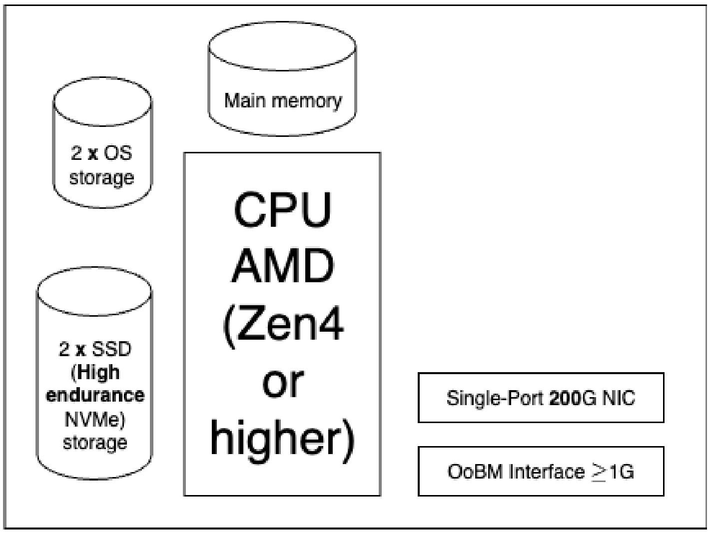

Figure 5: **Storage Head Node overview**. The storage head node is a PCIe Gen 4 (or higher) based single CPU socket machine having one physical CPU. It must include at least two enterprise-grade, flash-based storage devices suitable for use as boot drives in a RAID-1 configuration. Additionally, the node must be equipped with 2× high endurance NVMe SSDs for storage. The node must include a general-purpose, server-grade, single port 200 Gbps full-duplex Ethernet NIC (1×200 GbE). The installed NIC must actually deliver sustained throughput of at least 90% of full line rate, as integrated within the node. It must also include at least one dedicated server-grade 1 Gbps (or higher) OoBM interface supporting remote management via IPMI or Redfish (or equivalent).

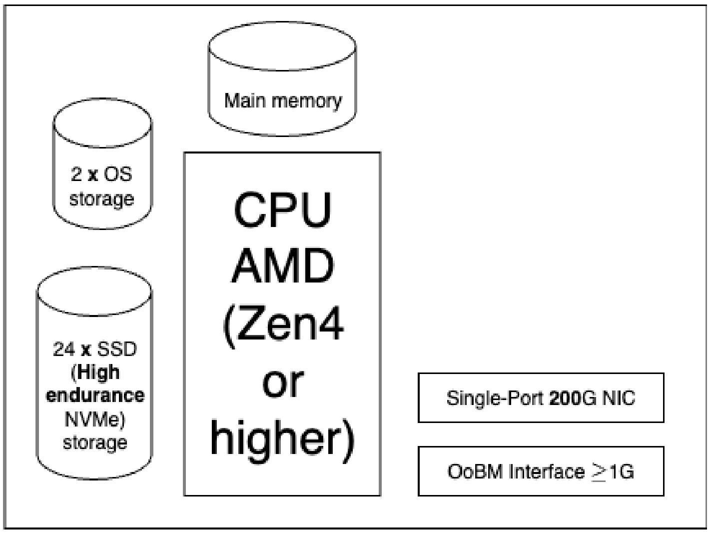

Figure 6: **Storage Data Node overview**. A storage data node is a PCIe Gen 4 (or higher) based single CPU socket machine having one physical CPU. It must include at least two enterprise-grade, flash-based storage devices suitable for use as boot drives in a RAID-1 configuration. Additionally, the node must be equipped with 24× high endurance NVMe SSDs for storage. The node must include a general-purpose, server-grade, single port 200 Gbps full-duplex Ethernet NIC (1×200 GbE). The installed NIC must actually deliver sustained throughput of at least 90% of full line rate, as integrated within the node. It must also include at least one dedicated server-grade 1 Gbps (or higher) OoBM interface supporting remote management via IPMI or Redfish (or equivalent).

cluster. This measure is intended to ensure uninterrupted and optimal cluster operation while replacement drives are arranged by the vendor.

7. The vendor is responsible for installing the supplied equipment in the racks.

8. A support contract on the entire EFC must be provided for a minimum period of 3 years including free replacement of any failed/non-compliant hardware. The support must be "within five working days on site" to diagnose any malfunctioning equipment and "next business day" for any failed hardware replacement.

9. The vendor must make a complete offer that also includes the detailed PCIe configuration of the motherboard layout of each type of node in the cluster. Vendors must submit a detailed Bill of Materials (BoM) listing all proposed components, including part numbers, descriptions, manufacturers, models, and quantities.

10. The vendor must provide the anticipated delivery date of the complete cluster hardware in their proposal for planning purposes. However, the delivery must be completed within ten (10) weeks from the date of execution of the agreement. Timely delivery is essential to the project. Delays not attributable to *force majeure* or beyond the vendor's control will result in financial penalties, as defined in the accompanying document "Beschrijvend document Astron EFC.PDF".

11. It must be possible for the nodes to be installed and run with a non-modified x86-64 bit version of Debian 13 (Trixie) running Linux kernel 6.12.x (or newer) without the need for any vendor-supplied device drivers, boot-time installed firmware, driver disks, or custom kernels. In case any licenses are required, they must be included within the offer and must be perpetual. Trusted Platform Module (TPM) functionality is not required, and the nodes must not depend on TPM for booting, driver compatibility, or operation.

12. The vendor must agree with the Acceptance Procedure and Compliance Tests detailed in Section 4.

13. The offer must comply with all the requirements specified in this document including the Requirements Conformity Tables in Section 5.

4 Acceptance Procedure and Compliance Tests

The to-be-delivered EuroFlash cluster must be able to meet the specifications quoted throughout this entire document. The Cluster must pass initial checks during the following four stages before it can be considered as suitable for work by the EuroFlash team.

(1) Before Delivery: Before delivery, the vendor must prove — using documentation from the components in the nodes — that the specifications are met.

(2) Delivery and Inspection: Within the delivery deadline, the vendor must deliver all the offered hardware to the Center for Information Technology (CIT) - Smitsborg, University of Groningen, Nettelbosje 1, 9747 AJ, Groningen, The Netherlands. This needs to be proven by the vendor, by identification and counting of all hardware, and confirmed on a checklist signed by the EuroFlash team.

(3) Installation: The vendor (or their sub-contractor) will have a period of one week following the delivery deadline to physically install all the supplied equipment in the rack.

(4) Compliance Tests: Immediately following hardware delivery and installation, the EuroFlash team will have a two-week period dedicated to network installation and setting up the EFC. This includes all necessary configuration and tuning to prepare the nodes for compliance testing. In the event of delays caused by the EuroFlash team, this period may be extended accordingly.

The compliance tests will consist of a series of standardized benchmark procedures designed to verify that the supplied hardware meets the specified requirements. These tests aim to produce robust and consistent results, enabling the contracting authority to confirm compliance with the quoted specifications.

The EuroFlash team (client) is responsible for conducting the compliance tests and preparing the corresponding test results report. Vendors are also encouraged to provide their own test methods and/or test software if so desired to further demonstrate that the delivered products meet the stated performance and functionality criteria.

All testing will be conducted using the non-modified x86-64 bit version of Debian 13 (Trixie) running Linux kernel 6.12.x (or newer). In principle hardware RAID is not necessary. During testing, cluster nodes will not utilize hardware RAID controllers for non-boot drives (in case they are present).

The specific types of tests to be performed are listed in Table 1, with brief descriptions provided below.

cpu-perf: This test verifies that there are no CPU related faults or performance bottlenecks in each unique node configuration. Each node shall undergo a sustained CPU stress test using *stress-ng* (`https://github.com/ColinIanKing/stress-ng`). The test shall fully load all CPU cores for a period of not less than several hours, sufficient to confirm stable and reliable operation under continuous full-load conditions. The CPU must operate within its manufacturer-specified frequency range during the test. ~~CPU performance shall be validated using standard library based DGEMM benchmarks, executed on: a single core, and all available cores. The results should reflect expected performance for the given CPU model and configuration.~~

mem-test: This test ensures that no bottle-necks or faults exist in the memory configuration and performance. The absence of memory errors shall be verified using *Memtest86+* (`https://www.memtest.org`). Memory speed is validated based on the frequency and channel configuration reported by the operating system via *dmidecode*. Measured memory bandwidth using *PassMark PerformanceTest* (`https://www.passmark.com/products/pt_linux`) is compared to the theoretical bandwidth calculated from the configured memory frequency and channel configuration. The measured bandwidth must reach at least 90% of the theoretical peak. ~~For each unique node configuration, the memory read and write speed shall be tested, e.g., with the *STREAM benchmark*~~[8].

gpu-copy: This test ensures that no bottle-necks or faults exists in the GPU PCIe configuration and performance. For each GPU in the nodes it shall be tested that the PCIe connection from CPU to GPU and GPU to CPU can reach the specified performance. Measured PCIe bandwidth must reach at least 80% of the theoretical peak.

gpu-perf: This test verifies that each type of GPU in the system is correctly configured and performs as expected, without hardware faults or performance bottlenecks. For each GPU present in the GPU nodes, the following synthetic performance metrics shall be measured at the GPU base clock frequency: Single precision FLOPS (FP32); Double precision FLOPS (FP64); Tensor Core performance (FP16). GPU memory read and write bandwidth. The measured performance must reach at least 90% of the theoretical peak for each metric, based on the manufacturer specifications. In the unlikely scenario that the GPU can not sustain its base clock frequency due to power capping, the actual clock frequency during the benchmark will be used to determine the theoretical peak performance. ~~Compliance with this peak performance~~

[8] ~~`https://github.com/jeffhammond/STREAM.git`~~

Table 1: **Applicability of compliance tests to nodes:** The table below outlines the compliance tests that must be satisfied by the nodes of each system in the cluster.

Compliance test	Compute system	GPU system	Storage system
cpu-perf	yes	yes	yes
mem-test	yes	yes	yes
eth-test	yes	yes	yes
nvme-test	yes	yes	yes
gpu-copy	no	yes	no
gpu-perf	no	yes	no
bar1-test	no	yes	no

~~is mandatory, and there shall be no power capping on the GPUs, which must be allowed to operate up to their design power (TDP) as specified by the manufacturer.~~

eth-test: This test ensures that no bottle-necks or faults exist in the NIC configuration and performance. For each type of NIC delivered, these tests shall verify the performance of the NICs between two nodes, both for the 100 GbE and the 200 GbE networks. Both normal network operations (TCP and UDP communication) as well as RDMA operations with RoCEv2 and DPDK shall be tested. For example, *iperf* and *qperf* might be used to verify normal network operations and RDMA operations.

nvme-test: This test ensures that no bottle-necks or faults exist in the local data storage configuration and performance. These tests shall verify both the read and write capabilities of the local data storage in all different node types. The NVME/storage performance will be validated using the *FIO benchmark* (`https://fio.readthedocs.io/en/latest/fio_doc.html`).

bar1-test: This test ensures the correct configuration and operation of Direct Memory Access to the GPU memory. It shall be verified that the full memory size of the GPU in the GPU node is accessible on the PCIe BAR1 memory and that data can be written to the memory with a DMA transaction using a nvidia-smi query. ~~For example, with *gdrcopy*~~[9].

Non-Compliance: If the equipment proves to be non-compliant, the vendor has a single period of maximum four weeks to meet compliance through, e.g., an exchange, extension or reconfiguration of the nodes. All the compliance tests specified in this section must be successfully met within the already awarded period of four weeks.

[9] ~~`https://github.com/NVIDIA/gdrcopy`~~

5 Requirements Conformity Tables

The following tables list the cluster requirements in addition to any other requirements specified elsewhere in this document. Vendors must add an “X” in the Fully Compliant (FC) column, for each requirement row, to confirm adherence. Where necessary, vendors may provide appropriate remarks in the Remarks column to explain why they consider the requirement fully compliant. Vendors are also encouraged to include any additional explanations or comments for each requirement on separate pages. Only those offers that are fully compliant with all the requirements specified in this document, including those in this section, will be considered for the award of the tender. The vendors must also fill the appropriate values as per their offer in Table 12.

Table 2: **Rack Requirements**
The following general requirements are applicable for the entire cluster.

RID	Requirement	FC	Remarks
RK. 1	The cluster must be mountable in 19 inch racks (with doors) that are already available at the site location. Racks are 800 mm×1200 mm (width×depth) of which only 40 U (height) is available per rack to the vendor. Any additional components necessary for integration into these racks must be supplied.		
RK. 2	The cluster must fit physically in at most 2 adjacent racks. This corresponds to a total of 80 U of usable rack height across two racks.		
RK. 3	The weight of the supplied equipment must not exceed 1500 kg per rack.		
RK. 4	The vendor is responsible for installing the supplied equipment into the existing racks at the site location. All necessary rack-mounting materials, rack mount kits with sliding rails compatible with the available racks, must be included in the offer. Cable management arms are not required and the use of sliding rails is mandatory.		
RK. 5	The cluster must operate using 50 Hz AC and 220/230 V.		
RK. 6	The maximum power consumed must be less than 25 kW per rack. (Two UPS feeds are available; in the event of a power outage, a single feed can supply up to 25 kW.)		
RK. 7	The power inlet on all the supplied equipment must be IEC C14, C16, or C20. Power cords are not required as they are available at the installation site. Cords with matching connectors (C13, C15, and C19) will be provided by the EuroFlash team in data center-specific colors.		
RK. 8	All nodes must have 80 Plus Platinum (or better) Hot-plug, Redundant Power Supply. One feed power outage must not affect their availability.		
RK. 9	All nodes must dissipate their heat directly into the air (i.e., no liquid cooling solutions are allowed).		
RK. 10	All nodes must be able to operate continuously to nominal specifications (under full load) up to an intake air temperature between 18−28°C. The standard intake air temperature in the data center will usually be about 24°C.		

Table 3: **Common Requirements for All Node Types**
The following list of common requirements applies to all the nodes in the cluster. Thus they are applicable to the Compute Head Node, Compute Production Nodes, GPU nodes, Storage Head Node, and Storage Data Nodes.

RID	Requirement	FC	Remarks
C. 1	Each node including all its components must be designed for continuous 24×7 use.		
C. 2	Each node must have only a single CPU socket with one CPU installed.		
C. 3	Each node shall allow the CPU to operate up to its full rated TDP, without any power capping or restrictions that limit performance below the manufacturer-specified levels. Configurable TDP (as per AMD) must be fully supported and exposed in the BIOS/firmware.		
C. 4	All nodes and any PCIe expansion cards must support IOMMU.		
C. 5	Boot drives: Each node must be equipped with at least two enterprise-grade, flash-based storage devices suitable for use as boot drives in a RAID-1 configuration. Each drive must have a total gross capacity between 900 GB and 1300 GB. Only high-performance interfaces such as NVMe or SAS SSDs are acceptable; SATA SSDs are not permitted.		
C. 6	The nodes must allow the operating system to monitor drive health metrics (e.g., SMART data, temperature, endurance) without requiring vendor-specific tools.		
C. 7	Each node must include a dedicated, server-grade OoBM interface with a minimum of 1 Gbps bandwidth, supporting remote hardware-level management protocols such as IPMI or Redfish (or equivalent). The interface must be enterprise-grade (e.g., iDRAC, iLO, or LOM) and capable of operating independently of the host OS.		
C. 8	All software, drivers, licenses related to the OoBM solution and the remote/virtual KVM over IP setup (including any required vendor specific tools and licenses) must be provided as part of the offer and must not have an expiry date.		
C. 9	On each node, remote configuration of BIOS or EFI settings (device boot order, etc.), and access to “magic” SysRq functionality via either the IPMI SoL or rKVM must be possible.		
C. 10	All nodes must be delivered with a uniform username and password for OoBM access; if not, the access credentials must be securely provided electronically.		
C. 11	For each node access must be provided (e.g., BIOS/UEFI passwords) for full control over Firmware, including its upgrade.		
C. 12	The MAC addresses of all Ethernet interfaces must be electronically delivered to the EuroFlash team no later than when the hardware is physically delivered.		
C. 13	Each node must support the remote installation, booting, and operation of an unmodified 64-bit (x86-64) version of Debian 13 (Trixie) with Linux kernel 6.12.x or newer. All proposed hardware components must be fully compatible with this operating system. Any additional drivers, firmware, software, licenses required for full functionality must be included in the offer for perpetual use.		

Table 4: **Additional Requirements for Compute and GPU System Nodes**
The following list of requirements applies to all the nodes in the Compute and the GPU systems. These include the compute head node, all the compute production nodes, and all the GPU nodes. The nodes must also meet the common requirements for all node types specified in Table 3. Additional requirements specific to each node type are provided in separate tables.

RID	Requirement	FC	Remarks
CG. 1	The node must also meet the common requirements for all node types specified in Table 3.		
CG. 2	The node must be a PCIe Gen 5 based single CPU socket machine with one CPU installed.		
CG. 3	The node must fully support AMD Turin CPUs, including providing a minimum of 128 PCIe Gen 5 lanes without performance limitations or bottlenecks.		
CG. 4	The installed CPU must be suitable for a single-socket server configuration and must be selected from the following list of **5^{th} Gen AMD Epyc (Turin/Turin Dense) processors: (i) AMD Epyc 9655 or 9655P (96-core) (ii) AMD Epyc 9745 (128-core), (iii) AMD Epyc 9645 (96-core)**. The same selected CPU model must be installed across all the nodes in the compute and GPU systems.		
CG. 5	All CPUs installed across both the Compute and GPU systems must be identical. Mixing of CPU models is not allowed across the compute head node, compute production nodes and GPU nodes.		
CG. 6	The minimum number of physical CPU cores per socket must be 96.		
CG. 7	The maximum number of physical CPU cores per socket must be 128.		
CG. 8	The base clock frequency of the CPU cores must be at least 2.3 GHz.		
CG. 9	Each CPU Socket must support a minimum of 12 memory channels.		
CG. 10	All main memory modules must be DDR5 6400 MHz RDIMMs, have ECC and run at at least 6000 MHz (Main memory that actually runs at 6400 MHz in the nodes is highly preferred).		
CG. 11	12 memory channels per CPU socket must be fully populated with memory modules. It is acceptable if the system has more than 12 DIMM slots per socket, but only those necessary to achieve the balanced 12 channel configuration must be populated.		
CG. 12	All main memory modules in the node must be of identical make, model and specification.		
CG. 13	The total main memory in the node must be at least 1152 GiB.		
CG. 14	The total main memory in the node must not exceed 1536 GiB.		
CG. 15	All nodes in the Compute and GPU systems must have the same total main memory size; nodes with differing memory capacities are not permitted.		

table continued ...

...table continued

Table 4: Additional Requirements for Compute and GPU System Nodes

RID	Requirement	FC	Remarks
CG. 16	In addition to dedicated boot drives (C.5), each node must also be equipped with one or more SSDs for various purposes (scratch, home directories, etc.). The number of required SSDs for each specific kind of system is further described in the specific requirements tables later. All SSDs (not used for boot) in the compute and GPU systems must meet the following requirements. Usage of external devices (enclosure) to host SSDs is not allowed. SSD requirements: • All SSDs must be NVMe-based, PCIeGen 4 based (or higher). ***and*** • Each SSD must have a gross capacity between 14 TB and 20 TB. ***and*** • Each SSD must have a DWPD of at least 1 based on a 5 year usage period. • Each SSD: Random 4KiB performance : Read > 900 KIOPS; Write > 175 KIOPS (not Simultaneous) ***and*** • Each SSD must support sequential 128 KiB transfers with a minimum performance of over 6000 MB/s for reads and over 3500 MB/s for writes (not simultaneously). The performance levels must be sustained when the SSDs are installed and operating within the node ***and*** • SSDs must be capable of operating as individual disks as well as with software RAID levels 0, 1, 5, 6, and JBOD ***and*** • SSDs across the entire Compute and GPU systems must be identical, coming from the same manufacturer and model.		
CG. 17	The offer must include two (2) spare SSDs as specified in CG.16. These two SSDs will be stored on-site to allow immediate replacement in the event of SSD failures within the Compute and GPU systems. This measure is intended to ensure uninterrupted and optimal cluster operation while replacement drives are arranged by the vendor.		

Table 5: **Specific Requirements for Compute Head Node**
The following list has the requirements for the compute head node. The compute head node must also meet the common requirements for all node types specified in Table 3 and additional requirements for the Compute and GPU systems specified in Table 4.

RID	Requirement	FC	Remarks
CH. 1	The offer must contain one compute head node.		
CH. 2	The compute head node must meet the common requirements for all node types specified in Table 3 and additional requirements for compute and GPU systems specified in Table 4.		
CH. 3	The compute head node must include a single general-purpose, server-grade 100 Gbps full-duplex dual-port Ethernet network interface card (NIC) (2×100 GbE). The installed NIC must actually deliver sustained throughput of at least 90% of full line rate on both ports simultaneously, as integrated within the node. It must support QSFP28 transceivers and be fully compatible with the RoCEv2 protocol. The NIC must be connected via a dedicated PCIe interface with at least 16 lanes. These lanes must be PCIe Gen 5-capable at the platform level, even if the NIC operates at PCIe Gen 4.		
CH. 4	The compute head node must be equipped with 16×hot-swappable SSDs to host home and software directories. SSD requirements are specified in CG.16.		

Table 6: **Specific Requirements for Compute Production Nodes**
The following list has the requirements for the compute production nodes. The compute production nodes must also meet the common requirements for all node types specified in Table 3 and additional requirements for the Compute and GPU systems specified in Table 4.

RID	Requirement	FC	Remarks
CP. 1	The minimum number of offered compute production nodes must be 8 (eight).		
CP. 2	Each compute production node must meet the common requirements for all node types specified in Table 3 and additional requirements for compute and GPU systems specified in Table 4.		
CP. 3	All compute production nodes must be from the same server manufacturer and must be identical to each other including all the installed components and firmware revision levels. Third-party components (e.g., memory, NICs) must be of the same brand and model across all nodes, and not substituted with equivalents from different manufacturers.		
CP. 4	Each compute production node must include a single general-purpose, server-grade 100 Gbps full-duplex dual-port Ethernet network interface card (NIC) (2×100 GbE). The installed NIC must actually deliver sustained throughput of at least 90% of full line rate on both ports simultaneously, as integrated within the node. It must support QSFP28 transceivers and be fully compatible with the RoCEv2 protocol. The NIC must be connected via a dedicated PCIe interface with at least 16 lanes. These lanes must be PCIe Gen 5-capable at the platform level, even if the NIC operates at PCIe Gen 4.		
CP. 5	Each compute production node must be equipped with one hot-swappable SSD to be used as scratch disk. SSD requirements are specified in CG.16.		

Table 7: **Specific Requirements for GPU Nodes**
The following list has the requirements for the GPU nodes. Each GPU node must also meet the common requirements for all node types specified in Table 3 and additional requirements for the Compute and GPU systems specified in Table 4.

RID	**Requirement**	**FC**	**Remarks**
G. 1	The offer must contain six GPU nodes.		
G. 2	Each GPU node must meet the common requirements for all node types specified in Table 3 and additional requirements for compute and GPU systems specified in Table 4.		
G. 3	All GPU nodes must be from the same server manufacturer and must be identical to each other including all the installed components and firmware revision levels. Third-party components (e.g., memory, NICs) must be of the same brand and model across all nodes, and not substituted with equivalents from different manufacturers.		
G. 4	Each GPU node must have 2×NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs.		
G. 5	~~In each GPU node, there shall be no power capping on the GPUs; both GPUs in the node must be allowed to operate up to their full design power (TDP) as specified by the manufacturer.~~ Nodes hosting the GPUs with power-capping are allowed, provided that each node still supports sustained operation of each GPU at at least 450W. (Note: Support for operation at higher power levels, although desired, is not mandatory and has no bearing on the award criteria.)		
G. 6	In each GPU node, the PCIe connection between CPU socket and any GPU must consist of at least 16 PCIe Gen 5 lanes. The connections between CPU socket and any GPU cannot be shared with other GPUs, NICs or peripherals.		
G. 7	In each GPU node, the connection to the GPU must be simultaneously able to sustain a full duplex PCIe Gen 5×16 transfer rate, transferring data from motherboard DRAM to GPU as well as GPU to motherboard DRAM.		
G. 8	Each GPU node must include a single general-purpose, server-grade 100 Gbps full-duplex dual-port Ethernet NIC (2×100 GbE). The installed NIC must actually deliver sustained throughput of at least 90% of full line rate on both ports simultaneously, as integrated within the node. It must support QSFP28 transceivers, the RoCEv2 protocol, NVIDIA GPUDirect RDMA (PeerDirect), and DPDK gpudev for GPU-aware packet processing. The NIC must be connected via a dedicated PCIe interface with at least 16 lanes. These lanes must be PCIe Gen 5-capable at the platform level, even if the NIC operates at PCIe Gen 4.		
G. 9	Each GPU node must be equipped with one hot-swappable SSD to be used as scratch disk. SSD requirements are specified in CG.16		

Table 8: **Additional Requirements for Storage System Nodes**
The following list of requirements applies to all the nodes in the Storage System. These include the storage head node and all the storage data nodes. The nodes must also meet the common requirements for all node types specified in Table 3. It is further followed by additional requirements specific to each type of node.

RID	Requirement	FC	Remarks
S. 1	The node must also meet the common requirements for all node types specified in Table 3.		
S. 2	The node must be a PCIe Gen 4 (or higher) based single CPU socket machine with one CPU installed.		
S. 3	The installed CPU in the node must be suitable for a single-socket server configuration and must be from the 4th generation (or newer) AMD EPYC processor family. The same CPU model must be used across all nodes of the storage system. Thus mixing of CPU models is not allowed across the storage head nodes and storage data nodes.		
S. 4	The node must fully support the installed CPU without performance limitations or bottlenecks.		
S. 5	The minimum number of physical CPU cores per socket must be 48.		
S. 6	The maximum number of physical CPU cores per socket must be 64.		
S. 7	The base clock frequency of the CPU cores must be at least 2.3 GHz.		
S. 8	The CPU must support a minimum of 12 memory channels.		
S. 9	Each CPU Socket must support a minimum of 12 memory channels.		
S. 10	All main memory modules must be DDR5 4800 MHz RDIMMs (or higher), have ECC, and run at at least 4800 MHz.		
S. 11	12 memory channels per CPU socket must be fully populated with memory modules. It is acceptable if the system has more than 12 DIMM slots per socket, but only those necessary to achieve the balanced 12 channel configuration must be populated.		
S. 12	All main memory modules in the node must be of identical make, model and specification.		
S. 13	The total main memory in the node must be at least 576 GiB.		
S. 14	The total main memory in the node must not exceed 768 GiB.		
S. 15	All nodes in the Storage systems must have the same total main memory size; nodes with differing memory capacities are not permitted. This includes both storage head nodes and storage data nodes.		

table continued . . .

. . . table continued

Table 8: Additional Requirements for Storage System Nodes

RID	Requirement	FC	Remarks
S. 16	The node must include a single general-purpose, server-grade 200 Gbps full-duplex, single-port Ethernet network interface card (NIC) (1×200 GbE). The installed NIC must actually deliver sustained throughput of at least 90% of full line rate, as integrated within the node. It must support QSFP56 transceivers and be compatible with the RoCEv2 protocol. The NIC must be connected via a dedicated PCIe interface with at least 16×PCIe Gen 4 (or higher) lanes.		
S. 17	In addition to dedicated boot drives (C.5), each node must also be equipped with one or more high-endurance SSDs for bulk data storage. The number of required SSDs for specific kind of system is further described in the specific requirements tables later. All SSDs (not used for boot) in the storage systems must meet the following requirements. Usage of external devices (enclosure) to host SSDs is not allowed. **High-endurance SSD requirements:** • All high-endurance SSDs must be NVMe-based ***and*** • Each high-endurance SSD must have a gross capacity between 12 TB and 26 TB ***and*** • Each high-endurance SSDs must have a DWPD of at least 3 based on a 5-year usage period and specified for mixed-use workloads (e.g., JEDEC JESD219 or similar) ***and*** • Each high-endurance SSD: Random 4KiB performance : Read > 900 KIOPS; Write > 175 KIOPS (Not Simultaneous) ***and*** • Each high-endurance SSD must support sequential 128 KiB transfers with a minimum performance of over 6000 MB/s for reads and over 3500 MB/s for writes (not simultaneously). The performance levels must be sustained when the SSDs are installed and operating within the node ***and*** • High-endurance SSDs must be capable of operating as individual disks as well as with software RAID levels 0, 1, 5, 6, and JBOD ***and*** • High-endurance SSDs across the entire storage system must be identical, coming from the same manufacturer and model.		
S. 18	The offer must include two (2) spare high-endurance SSDs as specified in S.17. These two high-endurance SSDs will be stored on-site to allow immediate replacement in the event of drive failures within the storage system. This measure is intended to ensure uninterrupted and optimal cluster operation while replacement drives are arranged by the vendor.		

Table 9: **Specific Requirements for Storage Head Node**

The following list has the requirements for the storage head node. The storage head node must also meet the common requirements for all node types specified in Table 3 and additional requirements for the storage systems specified in Table 8.

RID	Requirement	FC	Remarks
SH. 1	The offer must contain one storage head node.		
SH. 2	The storage head node must meet the common requirements for all node types specified in Table 3 and additional requirements for storage systems specified in Table 8.		
SH. 3	The storage head node must be equipped with 2×High endurance SSDs (hot-swappable). High-endurance SSD requirements are specified in S.17.		

Table 10: **Specific Requirements for Storage Data Nodes**

The following list has the requirements for the storage data nodes. The storage data nodes must also meet the common requirements for all node types specified in Table 3 and additional requirements for the storage systems specified in Table 8. The storage data node will be used for storage of bulk data.

RID	Requirement	FC	Remarks
SD. 1	The offer must contain three storage data nodes.		
SD. 2	The storage data node must meet the common requirements for all node types specified in Table 3 and additional requirements for storage systems specified in Table 8.		
SD. 3	All storage data nodes must be from the same server manufacturer and must be identical to each other including all the installed components and firmware revision levels. Third-party components (e.g., memory, NICs) must be of the same brand and model across all nodes, and not substituted with equivalents from different manufacturers.		
SD. 4	Each storage data node must be equipped with 24×High-endurance SSDs (hot-swappable). High-endurance SSD requirements are specified in S.17.		
SD. 5	Each storage data node must be able to sustain a simultaneous aggregate sequential read throughput of at least 60000 MB/s and an aggregate sequential write throughput of at least 45000 MB/s across all 24 high-endurance SSDs.		

Table 11: **General Terms and Conditions**

RID	Requirement	FC	Remarks
GTC. 1	The vendor must agree to comply with the provisions of the document "Bijlage 3 ARIV_2018.PDF".		
GTC. 2	All components supplied must be brand new and unused.		
GTC. 3	The vendor must make a complete offer that also includes detailed PCIe configuration of the motherboard layout of each type of node in the cluster. Vendors must submit a detailed Bill of Materials (BoM) listing all proposed components, including part numbers, descriptions, manufacturers, models, and quantities.		
GTC. 4	The delivery of the complete cluster hardware must be completed within ten (10) weeks from the date of execution of the agreement. Delays not attributable to *force majeure* or beyond the vendor's control will result in financial penalties, as defined in the accompanying document "Beschrijvend document Astron EFC.PDF".		
GTC. 5	The Vendor agrees to adhere to the Acceptance Procedure and Compliance Tests outlined in Section 4. The Vendor shall be considered to have successfully fulfilled the obligations of this Tender only upon successful completion of the acceptance procedure and compliance tests. Accordingly, any payments due shall be released only after such completion and formal acceptance by the Client.		
GTC. 6	A support contract on the entire EFC must be provided for a minimum period of 3 years including free replacement of any failed/non-compliant hardware. The support must be "within five working days on site" to diagnose any malfunctioning equipment and "next business day" for any failed hardware replacement. None of the components may be withdrawn from support for the entire support period.		
GTC. 7	In case the vendor uses a third party to provide any services, this must be clearly described in the offer.		
GTC. 8	Replacement of components such as hot-swappable drives and NICs may be performed by the EuroFlash team without affecting the validity or terms of the Vendor's support obligations.		
GTC. 9	The latest stable BIOS/firmware available at the time of shipment must be installed on the nodes.		

6 Award Criteria

The two aspects, namely the price and the quality would form the basis for evaluation of the Offers which also meet all the eligibility requirements. Tenderers must accurately fill in the 'Value' column of Table 12 truthfully. The methodology for evaluating and selecting the successful offer is defined in the accompanying document "Beschrijvend document Astron EFC.PDF".

Table 12: **Award Criteria : Quality**

Subquality/Subaward criterion description	**Value**
Total number of offered compute production nodes minus *7*	
Sustained Daily Write Endurance (SDWE) of the Storage System[a].	(TB/day)
Total Main memory (DRAM) size in a single node of the Compute System (in GiB).	(GiB)
Guaranteed[b] *operational main memory (DRAM) speed in the nodes of the Compute System (in MHz)* minus *5600.*	(MHz)
Pipeline Performance Factor (PPF) value[c] for the CPU model offered in the Compute and GPU systems.	
Guaranteed[b] *operational main memory (DRAM) speed in the nodes of the GPU system (in MHz)* minus *5600.*	(MHz)
Support contract duration (in years) minus *2.*	(years)
Aggregated power supply efficiency value[d].	
Aggregated SSD NVMe gross storage capacity over all the nodes of the Compute and GPU systems (in TB).	(TB)
Total Main memory (DRAM) size in a single node of the Storage system (in GiB).	(GiB)

[a] Sustained Daily Write Endurance (SDWE) of the Storage System is defined as the product of *"DWPD of a high-endurance drive"* **and** *"Aggregated high-endurance SSD NVMe gross storage capacity over all the nodes installed inside the Storage systems (in TB)"*. DWPD specification is based on a 5-year usage period and specified for mixed-use workloads. The two spare high-endurance drives and the boot drives must NOT be included in this calculation.

[b] Guaranteed means that the main memory must demonstrably operate at the specified frequency (MHz) when installed and integrated within the node running the unmodified 64-bit (x86-64) version of Debian 13 (Trixie) with Linux kernel 6.12.x or newer.

[c] The Pipeline Performance Factor (PPF) values for each of the allowed AMD Epyc Turin/Turin-dense CPU models are: (i) for AMD Epyc 9655 or 9655P (96-core): PPF=1.7; (ii) for AMD Epyc 9745 (128-core): PPF=1.2; and (iii) for AMD Epyc 9645 (96-core): PPF=1.

[d] The aggregated power supply efficiency value is calculated by summing the Power Supply Efficiency Factor (PSEF) of each node in the cluster. The PSEF for a node is assigned as per the following: (i) A node equipped with 80 Plus Ruby power supplies is assigned PSEF=3; (ii) A node equipped with 80 Plus Titanium power supplies is assigned PSEF=2; and (iii) A node equipped with 80 Plus Platinum power supplies is assigned PSEF=1.

7 Definitions and Conventions

Throughout this document the following definitions and conventions are used.

1 Gb = 10^9 bits
1 Gbps = 10^9 bits/second
1 GT/s = 10^9 transfers/second
1 MHz = 10^6 Hertz
1 GHz = 10^9 Hertz (used in context of CPU clock frequency)
1 MB = 10^6 Bytes
1 GB = 10^9 Bytes
1 GiB = 2^{30} Bytes = 1073741824 Bytes (used in context of RAM memory)
1 TB = 10^{12} Bytes
1 TFLOPS = 1 Tflops = 10^{12} FLOPS
GbE is Giga-bit Ethernet.
Single-precision refers to 32-bit floating-point numbers (IEEE 754).
Half-precision refers to 16-bit floating-point numbers (IEEE 754).
FLOPS is Floating point operations per second.

List of some common Abbreviations

AI Artificial Intelligence
BFC Beam Formed Correlator
BIOS Basic Input Output system
BOM Bill Of Materials
CIT Center for Information Technology (University of Groningen), The Netherlands
CPU Central Processing Unit
DGEMM Double-precision GEneral Matrix Multiply
DRAM Dynamic Random Access Memory
DWPD Drive Writes Per Day
ECC Error Correcting Code memory
EFC EuroFlash Cluster
EFI Extensible Firmware Interface
EF EuroFlash
ERC European Research Council
FC Fully Compliant
FHFL Full Height Full Length
FHHL Full Height Half Length
FRBs Fast Radio Bursts
GPU Graphics Processing Unit
HHHL Half Height Half Length
HPC High Performance Computing
iDRAC Integrated Dell Remote Access Controller
IEEE 754 Institute of Electrical and Electronics Engineers Standard for Floating-Point Arithmetic
IGMP Internet Group Management Protocol
ILO Integrated Lights Out
IOMMU Input-Output Memory Management Unit
IPMI Intelligent Platform Management Interface
IPoIB IP-over-InfiniBand
JBOD Just a Bunch Of Drives
KIOPS Kilo Input/output Operations Per Second

KVM Keyboard Video Mouse
LACP Link Aggregation Control Protocol
LDAP Lightweight Directory Access Protocol
LOFAR ERIC LOw Frequency ARray European Research Infrastructure Consortium (https://www.lofar.eu)
LOFAR LOw Frequency ARray (https://www.lofar.eu)
LOM Lights Out Management
MLD Multicast Listener Discovery (In context of snooping)
MTU Maximum Transmission Unit
NC Not Compliant
NIC Network Interface Controller
NVMe Non-Volatile Memory express
OoBM Out of Band Management
OS Operating System
PCIe Peripheral Component Interconnect Express
PXE Preboot Execution Environment
RAID Redundant Array of Independent Disks
RAM Random Access Memory
RID Requirement IDentifier
RoCEv2 RDMA over Converged Ethernet version 2
RU Rack Unit
SKA Square Kilometre Array
SNMP Simple Network Management Protocol
SSD Solid State Drive(Disk)
TDP Thermal Design Power
TPM Trusted Platform Module
UEFI Unified Extensible Firmware Interface
UPI Ultra Path Interconnect
UTP Unshielded Twisted Pair
VLAN Virtual Local Area Network
VPI Virtual Protocol Interconnect