Bilag 3
Overall requirements for Machine Types 1-4

DeiC National Committee for HPC

Proposal for
Overall Requirements for Machine Types 1-4

by

The DeiC National Committee for HPC

3rd March 2020

Contact: Prof. T. Larsen, Aalborg University,
prodekan-tech-forsk@aau.dk ; +45 2020 6856

Shared Requirements for all Machine Types 1-4

This section includes general demands for all four machine types listed in the report "Fremtidigt Nationalt HPC Landskab, 15. September 2019, Anbefalinger fra Arbejdsgruppen for fremtidigt national HPC landskab".

General abbreviations/definitions are included in Appendix "Abbreviations/Definitions".

The present section includes demands that must be met by all four machine types.

General

- The hardware must be hosted and maintained at a university and be directly connected to the Danish research network.

Access Models

Access models includes the idea of HPC as a service and the host participates with other Danish hosts to develop storage, compute and network infrastructure.

- One common login user authentication for researchers to Danish HPC.
- Multi factor authorization (2FA) must be used when private/public keys are not used for login.
- By default, resources on an individual machine can be shared between users (not over-subscription but node sharing). However, it must be possible for users to allocate a full machine for the single user.
- It must be possible to deliver resource usage reporting (compute, storage) at any time at user and project level.

Security

Security is an ever-continued process and the service provider must acknowledge that this area must be continuously developed – meaning that existing and coming rules and standard on national and international level must be adapted.

- The operating system and system administrated software must be timely updated according to the operating software and application software supplier.
- The entire system must be reinstalled in case of a serious exploit (unintended root access).
- The host must have a written and approved security policy and this must be updated annually and exposed to an audit.
- The host must have a written and implemented security policy that is governing at security breaches.
- The host is responsible for the technical security in the service provided and corresponding components used to provide the service. The user is responsible for data classification and data processing but the host facility may provide support and tools for data management.

User Support

- Operational support via phone, mail or web must be accessible all working days in the time window 08-16.
- Processing of support requests or incidents must be initiated within two working days.

- Must include assistance to install user requested software but not support on how to use specific software programs.
- Optional: Provide support and advice for data management at the host facility – e.g. data management planning considering systems and legal aspects. This potential service may be provided at extra cost.

Type 1 – Interactive HPC

This computing type has focus on interactive computing resources and easy access for new users – HPC-as-a-service is a key aspect here. Furthermore, this type may constitute a platform for educational purpose. It is expected that this type will increase the usage of HPC as an easy entry level for new users – also prototyping and testing may be important use cases, even for more advanced HPC users.

Access Models

- Provide graphical User Interface (UI) for access to the services.

Security

- Regarding handling of personal sensitive S1 may be applied.

Hardware

- The hardware must be based on x86 compliant CPUs.
- The service must offer a high degree of flexibility and scalability in terms of hardware capacity.
- The service must offer at least three predefined hardware profiles for virtual machines (e.g. small, medium and large).
- The service must be able to offer a virtual computing system with a hardware profile according to the service providers demands.
- A user must be able to request the service with minimum:
    - CPU cores: 4 (small) | 8 (medium) | 16 (large) → optional: 64 (XLarge).
    - Memory (GB): 16 (small) | 32 (medium) | 64 (large) → optional: 256 (XLarge).
    - Storage (TB): 1 (small) | 2 (medium) | 4 (large) → optional: 8 (XLarge). The host may also consider to offer different storage types such as a capacity or performance-oriented type.
- The service should also offer access to at least one type of GPU accelerated profile, possibly more.
- The service must support at least two virtualization technologies, like traditional VM and container systems.

Software

- The user must be able to install own software with sufficient rights in a protected environment.
- The service must as a minimum offer Virtual Machines (VMs) with pre-installed Linux operating system offered with the latest security patches.
- The service offers installing of open source software packages and also a possibility of using licensed software (bring your own license or by invoice).
- There should be a SaaS offer based on the user needs/requests (ready to use pre-installed software environments, e.g. Jupyter notebooks, RStudio and Matlab).

Storage

- The service must as a minimum offer storage for data computation and temporary storage space (minimum 1 TB). This storage is not a data archive.
- The service must as a minimum offer ** TB of data and temporary storage for the user.
- The service may provide backup, at additional cost, of data including software and computer code.
- The service must offer safe and secure erasure of data.

- The service must offer a possibility of capacity scaling.
- The user of the service must be able to transfer data to and from the host facility.
- The facility must interface to the future national DM landscape e.g. for data ingestion and publication.

Type 2 – Throughput HPC

This type is characterized by having a large amount of compute kernels which can be a mix between cost-efficient and compute-efficient units with high throughput capacity with focus on high security. The type is ideal for many small and intermediate size jobs which use large amounts of data/files.

Access Models

AM1 as described in the Appendix “Abbreviations/Definitions”.

Security

- Regarding handling of personal sensitive S1 must be applied.

Hardware

- Many x86 cores not necessarily with high clock frequency but there should also be access to cores with high(er) clock frequency. High memory bandwidth per core.
- Minimum 8 GB RAM/core, but preferably 12 GB RAM/core.
- Fast network, for exmaple 25 Gb/s Ethernet or min. 56 Gb/s Infiniband, with low latency. Min. 100 Gb/s Infiniband can be prioritised.
- Scratch space (local disk storage on each node), preferably 1 TB or more.

Software

- A minimal and modern Linux installation, for example CentOS 7 or 8. Common development utilities (compilers, libraries, git, cmake etc.) is installed and is easily made available for the users.
- The users must be able to install software in their own home or project directories, for example via Conda, easybuild or spack.
- The users must be able to install commercial software packages, if they have a license
- Access to compute via a queue system, for example Slurm.

Storage

- A fast parallel filesystem, for example Lustre, BeeGFS or CephFS.
- Must support quotas on home and project directories.
- Quota on home directories should aim for a capacity of at least 100 GB
- There should not necessarily be a quota on a project directory but it should be possible.
- The users choose themselves which data that must be backed up and will only pay for data in backup such that normal storage can be kept as cheap as possible

Type 3 – Large Memory HPC

The focus is here on applications that not easily or efficiently can be distributed between many computer nodes. There is a demand for a large flat memory-space as seen in large matrix problems or other problems with large memory demands and relatively few compute kernels.

Security

- Regarding handling of personal sensitive S1 must be applied.

Hardware

- At least 32 cores/node, preferably more for OpenMP and threading performance, not necessarily high clock frequency.
- At least 2TB RAM/socket, preferably 4TB or more
- At least 2 sockets/node, preferably 4 sockets/node or more
- Fast network with low latency, min. 56 Gb/s
- Minimum 8TB scratch space pr. Node, which is fast and has low latency.

Software

- Minimal, modern Linux installation, for example CentOS 7 or 8.
- Development tools (compilers, performance optimised libraries, git, cmake …) are installed and can be activated with for example the module program.
- The users can install software in their own home directories or project directories for example with Conda.
- Sought after scientific applications can be installed and activated for example with module to lower the barrier of entrance.
- Access to compute via a queue system, for example Slurm.

Storage

- Hardware to run a fast and cheap parallel filesystem, for example Lustre, BeeGFS or CephFS.
- Quota on home directories, for example 100 GB
- Quotas on project directories
- The users choose themselves which data that must be backed up and will only pay for data in backup such that normal storage can be kept as cheap as possible

Type 4 – Accelerated HPC

This is a compute type where the majority of computational resources comes from accelerated devices of different kind. This enables Danish researchers to develop and test their codebase on the next generation of accelerated devices. This may be novel FPGA technologies as well as in-memory and in-storage computing units.

Security

- Regarding handling of personal sensitive S1 may be applied.

Hardware

- Accelerators: Must support an as broad as possible portfolio of accelerators, AI, FPGA, advanced GPU integration, etc.
- CPUs: CPUs shall primarily function as a control unit for accelerators. CPUs, which optimises the usage of the accelerators, will be preferred.

Software

- Linux environment in the variant which best supports the hardware.
- The users shall to the greatest extent possible install software themselves. The host should be helpful installing packages, which the user does not have permissions to install (for example kernel modules).
- The host must choose technology, which allows the users to be as self-driven as possible.
- A set of standard development tool, which are required to use the accelerators, is maintained by the host.
- Access is granted through a queuing system, for example Slurm.

Storage

- Must include a high-speed central storage solution which allows access directly from accelerators.
- Media: Persistent memory with a NVMe capacity tier or plain NVMe. Potentially medias with in-storage compute.
- Minimal shared filesystem with home directories based on conventional technology. Users can have quotas.

Interconnect

- The interconnect must be looking towards the future but should primarily deliver a good platform for near communication between accelerators and if possible direct communication between accelerators and storage.
- Technologies, which exclude CPUs in the data plan between accelerators and storage, will be preferred.

Development Support

- Must offer online documentation.
- Must offer introductory courses/workshops with focus on the hardware – with national availability.
- Sparring on mail, possibly phone/Skype/f2f by appointment.
- Technical sparring anchored in a research group, which uses a Type 4 in research themselves.
- Support user community with focus on national knowledge sharing.

Appendix: Abbreviations/Definitions

- **AM1:**
  Login to front-end machine via SSH with suitable safety settings.
    - If a whitelist is utilized, all Danish universities public IP-addresses must be added to avoid barriers for external users.
- **S1:**
  Handling of personal sensitive data (legal, organizational, technical) that meets demands from the public health authorities, archive data legislation and the legislation for personal data. For example, the host facility is welcomed to meet compliance with GDPR or ISO27001 standards.

    - Logging of certain file operations (open/close) to make it possible to track unauthorized access to user data.
    - Export control that ensures that data cannot leave a secure area without approval.
    - Access via remote desktop or similar that makes it difficult to apply operations such as copy/paste, which may leak data.