DeiC
DANISH eINFRASTRUCTURE COOPERATION

Bilag 1
National HPC Center Machine Types

Specification and Requirements

2. April 2020

Content

1 Introduction

This appendix is based on input from “Arbejdsgruppen for national HPC” as shown in Appendix 3.

The appendix is adapted from the board of DeiC internal discussions that provide a framework for the universities to obtain prices and numbers for building the HPC facilities.

In addition to technical descriptions of the HPC type facility, the offer must also state price and service for the solution. This include price quotations that are customary for the type of facility. For example, it can be node-time price/hour and TB/month for type 2 facilities or virtual machine price/hour for the various virtual machine options for type 1 facilities.

The intention is that the offers can be compared with other similar services. This include both on the commercial market and at other research institutions at home and abroad.

2 Type 1 – Interactive

Type 1 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. A set (e.g. class set) of resources can be issued for to a limited time for educational or training. Type 1 is expected to 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.

In cases where users developed program that depends on compilers this must be done offsite and the program can be transferred to the site.

Demand 2.1 In case the user wants to run the VMs at the home university the host must provide a way to export the VM in a standard format like OVF or similar.

Option 2.1 Provided the user have some very special needs for VMs that cannot be met by the host the host should provide a way for the user to import the VM in a standard format like OVF or similar at an extra cost.

2.1 Access Models

Demand 2.1.1 Provide graphical User Interface (UI) for access to the services. This should be done via one of the common remote desktop access tools.

2.2 Hardware

- ***Demand 2.2.1*** *The hardware must be based on x86 compliant CPUs.*
- ***Demand 2.2.2*** *The service must offer a high degree of flexibility and scalability in terms of hardware capacity.*
- ***Demand 2.2.3*** *The service must offer at least three predefined hardware profiles for virtual machines (e.g. small, medium and large).*
- ***Demand 2.2.4*** *The service must be able to offer a virtual computing system with a hardware profile according to the users' demands.*
- ***Demand 2.2.5*** *A user must be able to request the service with minimum:*
- ***Demand 2.2.6*** *CPU cores: 4 (small) \| 8 (medium) \| 16 (large)*
- ***Option 2.2.1*** *64 (XLarge).*
- ***Demand 2.2.7*** *Memory (GB): 16 (small) \| 32 (medium) \| 64 (large)*
- ***Option 2.2.2*** *256 GB (XLarge).*
- ***Option 2.2.3*** *Storage should be large enough to hold data.*
- ***Option 2.2.4*** *The host may also consider offering different storage types such as a capacity or performance oriented type.*
- ***Option 2.2.5*** *The service should also offer access to at least one type of GPU accelerated profile, possibly more.*
- ***Demand 2.2.8*** *The service must support at least two virtualization technologies, like traditional VM and container systems.*

2.3 Software

- ***Demand 2.3.1*** *The user must be able to install own software with sufficient rights in a protected environment.*

- ***Demand 2.3.2*** *The service must as a minimum offer Virtual Machines (VMs) with pre-installed Linux operating system offered with the latest security patch.*
- ***Demand 2.3.3*** *The service must as a minimum offer Virtual Machines (VMs) with a pre-installed windows operating system. The host must provide license for VMs OS.*
- ***Demand 2.3.4*** *The service offers installing of open source software packages**.***
- ***Option 2.3.1*** *Possibility of using licensed software (bring your own license or by invoice).*
- ***Demand 2.3.5*** *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).*

2.4 Storage

- ***Demand 2.4.1*** *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.*
- ***Option 2.4.1*** *The service may provide backup, at additional cost, of data including software and computer code.*
- ***Demand 2.4.2*** *The service must offer safe and secure erasure of data.*
- ***Demand 2.4.3*** *The service must offer a possibility of capacity scaling.*
- ***Demand 2.4.4*** *The user of the service must be able to transfer data to and from the host facility.*
- ***Demand 2.4.5*** *The facility must interface to the future national DM landscape e.g. for data ingestion and publication.*

3 Type 2 – Throughput

Type 2 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.

3.1 Access Models

Beside the central national HPC platform for login there can be a need to provide login to front-end machine via SSH with suitable safety settings. If a whitelist is utilized, Danish universities, public IP-addresses must be added to avoid barriers for external users.

3.2 Hardware

- ***Option 3.2.1*** *Many x86 cores not necessarily with high clock frequency but there should also be access to cores with high(er) clock frequency.*
- ***Demand 3.2.1*** *High memory bandwidth per core.*
- ***Demand 3.2.2*** *Minimum 8 GB RAM/core, but preferably 12 GB RAM/core.*
- ***Demand 3.2.3*** *Fast network, for example 25 Gb/s Ethernet or min. 56 Gb/s Infiniband, with low latency. Min. 100 Gb/s Infiniband can be prioritized.*
- ***Demand 3.2.4*** *Scratch space (local disk storage on each node), preferably 1 TB or more.*

3.3 Software

- ***Demand 3.3.1*** *A minimal and modern Linux installation, for example CentOS 8. Common development utilities (compilers, libraries, git, cmake etc.) is installed and is easily made available for the users.*
- ***Demand 3.3.2*** *The users must be able to install software in their own home or project directories, for example via Conda, easybuild or spack.*
- ***Demand 3.3.3*** *The users must be able to install commercial software packages, if they have a license*
- ***Demand 3.3.4*** *Access to compute via a queue system.*

3.4 Storage

- ***Demand 3.4.1*** *A fast parallel filesystem, for example Lustre, BeeGFS or CephFS.*
- ***Option 3.4.1*** *Must support quotas on home and project directories.*
- ***Option 3.4.2*** *Quota on home directories should aim for a capacity of at least 100 GB*
- ***Option 3.4.3*** *There should not necessarily be a quota on a project directory but it should be possible.*
- ***Option 3.4.4*** *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*

4 Type 3 – Large Memory

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.

4.1 Hardware

- ***Demand 4.1.1*** *At least 32 cores/node, preferably more for OpenMP or similar and threading performance, not necessarily high clock frequency.*
- ***Demand 4.1.2*** *At least 2TB RAM/socket, preferably 4TB or more.*
- ***Demand 4.1.3*** *At least two sockets/node, preferably 4 sockets/node or more.*
- ***Demand 4.1.4*** *Fast network with low latency, min. 56 Gb/s.*
- ***Demand 4.1.5*** *Minimum 8TB scratch space pr. Node, which is fast and has low latency.*

4.2 Software

- ***Demand 4.2.1*** *Minimal, modern Linux installation, for example CentOS 8.*
- ***Demand 4.2.2*** *Development tools (compilers, performance optimized libraries, git, cmake …) are installed and can be activated with for example the module program.*
- ***Demand 4.2.3*** *The users can install software in their own home directories or project directories for example with Conda.*
- ***Option 4.2.1*** *Sought after scientific applications can be installed and activated for example with module to lower the barrier of entrance.*
- ***Demand 4.2.4*** *Access to compute via a queue system.*

4.3 Storage

- ***Demand 4.3.1*** *Hardware to run a fast and cheap parallel file system, for example Lustre, BeeGFS or CephFS.*
- ***Option 4.3.1*** *Quota on home directories, for example 100 GB.*
- ***Option 4.3.2*** *Quotas on project directories.*
- ***Option 4.3.3*** *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*

5 Type 4 – Accelerated HPC

Type 4 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 as well as more energy efficient HPC for both storage and compute units.

5.1 Hardware

- ***Demand 5.1.1*** *Accelerators: Must support an as broad as possible portfolio of accelerators, AI, FPGA, advanced GPU integration, etc.*
- ***Demand 5.1.2*** *CPUs: CPUs shall primarily function as a control unit for accelerators. CPUs, which optimizes the usage of the accelerators, will be preferred.*

5.2 Software

- ***Demand 5.2.1*** *Linux environment in the variant, which best supports the hardware.*
- ***Demand 5.2.2*** *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).*
- ***Demand 5.2.3*** *The host must choose technology, which allows the users to be as self-driven as possible.*
- ***Demand 5.2.4*** *The host maintains a set of standard development tool, which are required to use the accelerators.*
- ***Demand 5.2.5*** *Access is granted through a queuing system.*

5.3 Storage

- ***Demand 5.3.1*** *Must include a high-speed central storage solution, which allows access directly from accelerators.*
- ***Option 5.3.1*** *Media: Persistent memory with a NVMe capacity tier or plain NVMe. Potentially medias with in-storage compute.*
- ***Option 5.3.2*** *Minimal shared file system with home directories based on conventional technology. Users can have quotas.*

5.4 Interconnect

- ***Demand 5.4.1*** *The interconnect hardware 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.*
- ***Option 5.4.1*** *Technologies, which exclude CPUs in the data plan between accelerators and storage, will be preferred.*

5.5 Development Support

- ***Demand 5.5.1*** *Must offer online documentation.*
- ***Demand 5.5.2*** *Must offer introductory courses/workshops with focus on the hardware – with national availability.*
- ***Demand 5.5.3*** *Sparring on mail, possibly phone/Skype/f2f by appointment.*
- ***Demand 5.5.4*** *Technical sparring anchored in a research group, which uses a Type 4 in research themselves.*
- ***Demand 5.5.5*** *Support user community with focus on national knowledge sharing.*