FENIX
RESEARCH INFRASTRUCTURE

Fenix Virtual Machine Services Models

The ICEI project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 800858.

Contents

1. Acronyms

AAI	Authentication and Authorization Infrastructure
ACD	Active Data Repositories
ACL	Access Control List
API	Application Programming Interface
ARD	Archival Data Repositories
BSC	Barcelona Supercomputing Center
CapEx	Capital Expenditure
CDP	Co-design Project
CEA	Commissariat à l'énergie atomique et aux énergies alternatives
CINECA	Consorzio Interuniversitario
CLI	Command Line Interface
CSCS	Centro Svizzero di Calcolo Scientifico
DL	Data Location Service
DM	Data Mover Service
DT	Data Transfer Service
FPA	Framework Partnership Agreement
FURMS	Fenix User and Resource Management Services
GoP	Group of Procurers
GUI	Graphical User Interface
HBP	Human Brain Project
HPAC	High Performance Analytics and Computing
HPC	High Performance Computing
HPDA	High Performance Data Analytics

HPST	High-Performance Storage Tier
IaaS	Infrastructure as a Service
IAC	Interactive Computing Services
ICCP	Interactive Computing Cloud Platform
ICEI	Interactive Computing E-Infrastructure for the Human Brain Project
ICN	Interactive Computing Node
IdP	Identity Provider
IPR	Intellectual Property Rights
JP	Joint Platform
JSC	Jülich Supercomputing Centre
LCST	Large-Capacity Storage Tier
MS	Monitoring Services
NDA	Non-Disclosure Agreement
NETE	External Interconnect
NETI	Internal Interconnect
NMC	Neuromorphic Computing
NVM	Non-Volatile Memory
NVRAM	Non-Volatile Random Access Memory
OIDC	OpenID Connect
OpEx	Operational Expenditure
PaaS	Platform as a Service
PCP	Pre-Commercial Procurement
PI	Principal Investigator
PID	Persistent Identifier
PIE	Public Information Event
PRACE	Partnership for Advanced Computing in Europe
Q&A	Questions and Answers
QoS	Quality of Service
R&D	Research & Development
R&I	Research & Innovation
RBAC	Role-Based Access Control

RFI	Request For Information
SCC	Scalable Computing Services
SGA	Specific Grant Agreement
SIB	Science & Infrastructure Board
SLA	Service Level Agreement
SP	Subproject
TCO	Total Cost of Ownership
TGCC	Très Grand Centre de Calcul
UI	User Interface
US	User Support Services
VM	Virtual Machine Services

2. Introduction

This document describes how Fenix VM Models could be defined in a similar manner as AWS1. The choice of Fenix VM Models provided by the different Fenix Resource Providers depends on the locally available hardware resources. The way how hardware resources are allocated is assumed to be done in a similar way at the different sites in order to allow for performance of the Fenix VM Models at the different sites to be reasonably similar. Next, the approach for announcing VM resources in user calls and finally the reporting of Fenix infrastructure resources is discussed.

3. VM Models

The different models differ (potentially) in terms of the following features:

Number of virtual CPUs (vCPU)

Amount of memory (MEM)

Number of virtual GPUs (vGPUs)

Amount of capacity-optimised storage (HDD)

Amount of performance-optimised storage (SSD)

All other differences, e.g. single CPU performance, memory bandwidth etc., are assumed to be negligible or not relevant.[2] It is furthermore assumed that applications can be deployed such that they can run on different processor architectures, i.e. also on non-x86 processors.

[1] See: https://aws.amazon.com/ec2/instance-types/

[2] AWS also considers, e.g. network performance.

Currently, there is no clear need for provisioning different amounts of capacity-optimised storage space (HDD).

Model name	vCPU	MEM [GiByte]	vGPU	SSD [GiByte]
gpp.s	1	3		
gpp.m	2	8		
gpp.l	4	16		
gpp.xl	16	64		
gpp-ssd.l	4	16		800
gpp-ssd.xl	16	64		800
gpu.m	2	8	1	
gpu.l	4	16	1	
gpu.xl	16	64	1	

4. Hardware Resource Allocation

The mapping of VM models onto hardware should follow similar rules at all Fenix sites in order to improve similarity in terms of performance. The following rules are proposed:

Allocate one vCPU per hardware thread.

MEM refers to physically available memory minus space needed for OS as well as system services (including PFS). No overbooking of MEM.

SSD refers to the effectively available storage. No overbooking of SSD.

Moderate overbooking of GPUs, e.g. 2 vGPU, per GPU (to be explored based on the use cases)

Let us consider the consequences of this policy for a specific example. Assume the following hardware resources being available:

Server type	Qty.	Physical resources per server				Virtual resources per server			
		#HW threads	Host memory	SSD [GiByte]	#GPUs	#vCPUs	MEM [GiByte]	SSD [GiByte]	#vGPU

			[GiByte]						
A	8	40	192	0	0	40	176	0	0
B	8	40	192	1600	0	40	176	1600	0
C	4	40	192	0	1	40	176	0	2
Total virtual resources:						800	3520	12800	8

The maximum number of VMs per model would be:

Model name	Server A	Server B	Server C	Total
gpp.s	320	320	160	800
gpp.m	160	160	80	400
gpp.l	80	80	40	200
gpp.xl	16	16	8	40
gpp-ssd.l	0	16	0	16
gpp-ssd.xl	0	16	0	16
gpu.m	0	0	2	8
gpu.l	0	0	2	8
gpu.xl	0	0	2	8

However, different combinations would be possible. One example:
One Server C could host 2 VMs of model gpu.m and 36 VMs of model gpp.s.

5. Resource Announcement

The following information is made available:
The total number of vCPU, MEM, SSD, vGPU.

The maximum number of VMs per model.

The information to be announced for the example presented in section 3 has already been provided in that section.

6. Resource Reporting and KPI Evaluation

The resources are reported as follows:

The total amount of available resources is defined in terms of total number of vCPU, amount of MEM, amount of SSD and number of vGPU.

The total amount of allocated resources is obtained by adding the number of vCPU, amount of MEM, amount of SSD and number of vGPUs for each of the allocated VMs. VMs are assumed to be running 24/7, i.e. the VM execution times are not reported.

The KPI for infrastructure utilisation is evaluated for vCPU, MEM, SSD and vGPU separately.

Let us reuse the example of available hardware resources introduced in section 3 and assume that 64 gpp.s, 8 gpp.xl, 4 gpp-sdd.xl and 4 gpu.m VMs have been allocated. In total, this translates into the following resource consumption:

	Total resources	**Utilisation [%]**
#vCPU	264	33
MEM [GiByte]	992	28
SSD [GiByte]	3200	25
#vGPU	8	100