LUXEMBOURG
LET'S MAKE IT HAPPEN

Cloud & HPC Resource Allocations for Research Projects and External Partners

High Performance Computing & Big Data Services

hpc.uni.lu
hpc@uni.lu
@ULHPC

LUXEMBOURG
LET'S MAKE IT HAPPEN

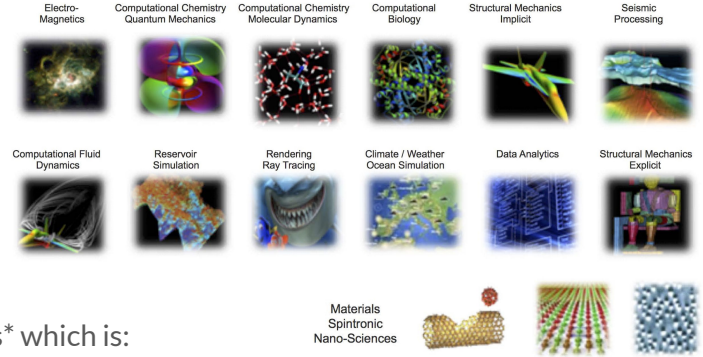
amazon web services™

Microsoft Azure

Google Cloud Platform

UNIVERSITÉ DU LUXEMBOURG

HPC, BD & Cloud Computing



- **HPC**: **H**igh **P**erformance **C**omputing - **BD**: **B**ig **D**ata
- **Cloud Computing:**
  - Network access to a shared pool of configurable computing resources* which is:
    - Ubiquitous, Convenient, On-demand

- All scientific disciplines are becoming **computational** today
  - Modern scientific discovery requires very high computing power and handles huge volumes of data
  - cf. J. Rifkin report: "3rd Industrial Revolution Strategy for the Grand Duchy of Luxembourg"
  - **Research Projects, Industry and SMEs** are increasingly relying on the computing resources…
    - … to invent innovative solutions while reducing cost and decreasing time to market

- HPC, BD & Cloud Computing: Essential tools for **Research**, **Science**, **Society and Industry**


Andy Grant, Head of Big Data and HPC, Atos UK&I

To out-compete you must out-compute

Increasing competition, heightened customer expectations and shortening product development cycles are forcing the pace of acceleration across all industries.

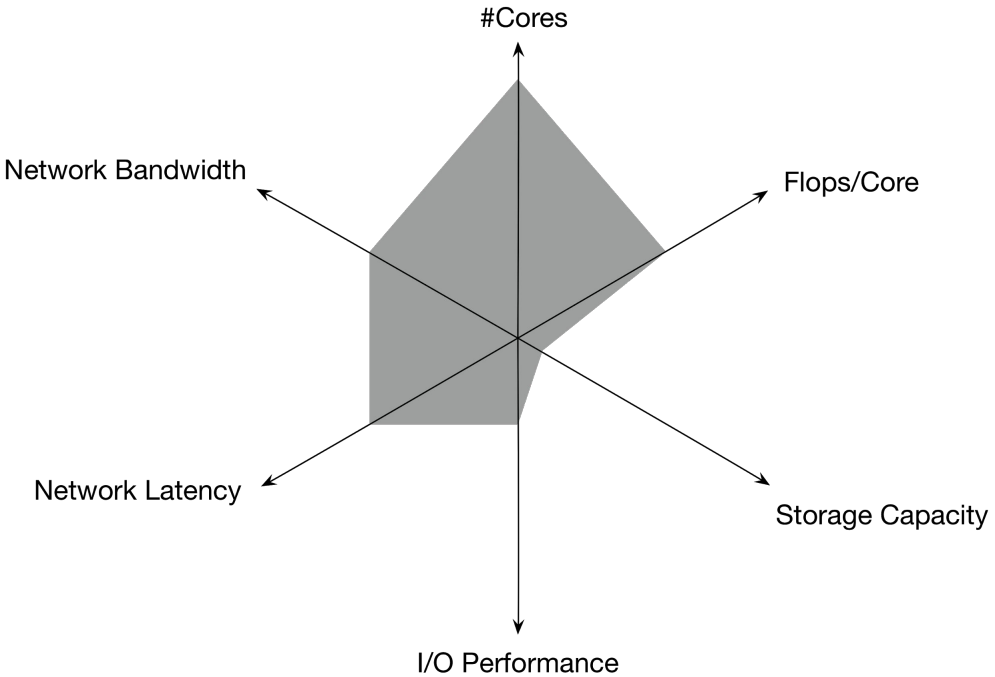

Science & Technology Facilities Council

Prof Tony Hey CBE FREng, STFC Chief Data Scientist & Joint Chair, UK E-infrastructure Leadership Council

Excellent research requires excellent research infrastructure

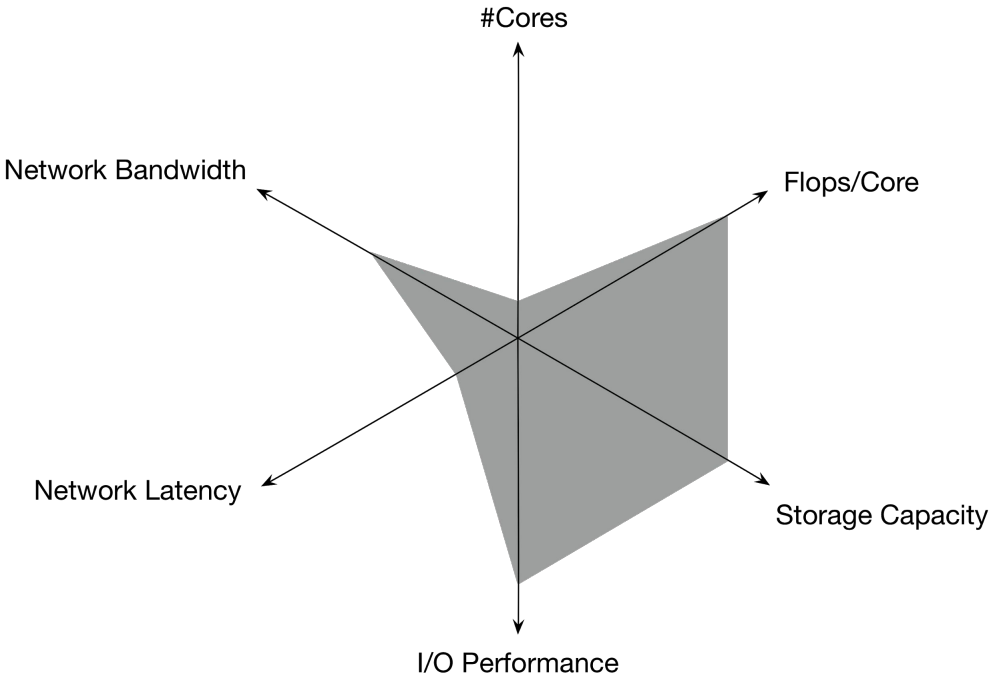
Different Needs for Different Domains

Material Science & Engineering



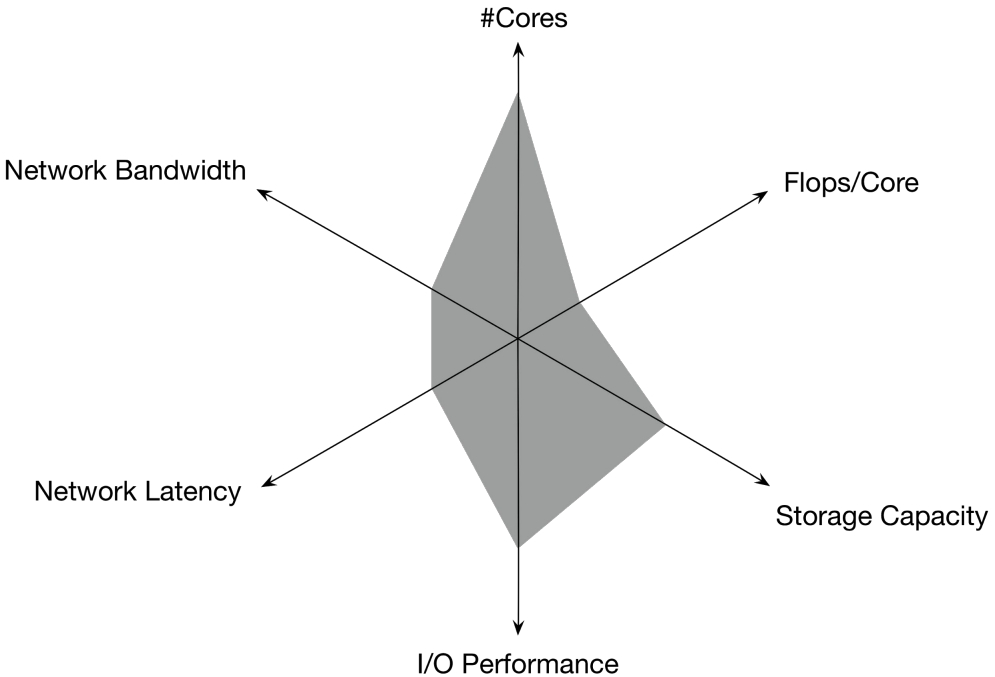
Different Needs for Different Domains

Biomedical Industry / Life Sciences



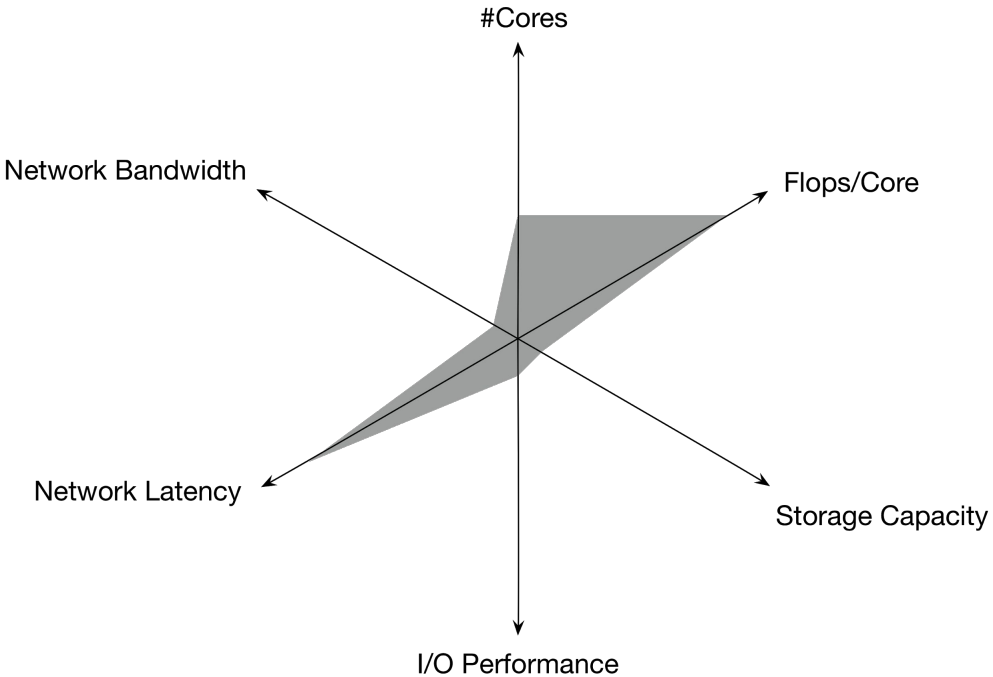
Different Needs for Different Domains

Deep Learning / Cognitive Computing



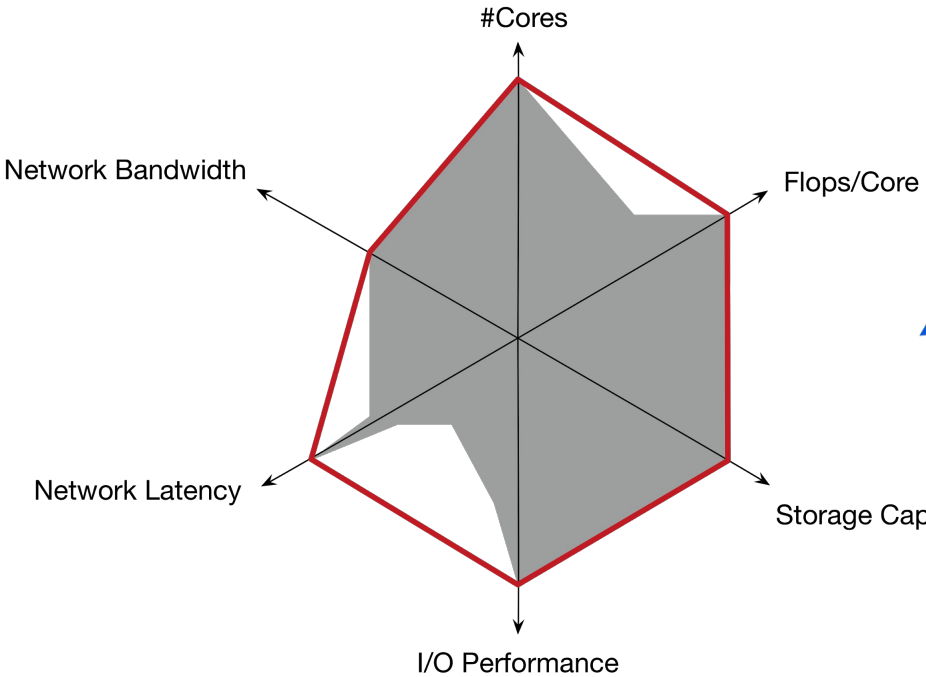
Different Needs for Different Domains

IoT, FinTech



Different Needs for Different Domains

ALL Research Computing Domains



amazon web services™

Microsoft Azure

Google Cloud Platform

High Performance Computing & Big Data Services

hpc.uni.lu

hpc@uni.lu

@ULHPC

LUXEMBOURG
LET'S MAKE IT HAPPEN

Cloud Computing Resource Allocation (Job)

1. Select instance type of computing resource

Cloud Instance Types

Entry Level	A-Series	T2	n1-standard
General Purpose	D-Series	M5	n1-standard
Compute Optimized	F-Series	C5	n1-highcpu
Memory Optimized	E-Series	X1e	n1-highmem
Storage Optimized	L-Series	I3	*Can attach SSDs to VMs
GPU Enabled	N-Series	G3	*Can attach GPUs to VMs

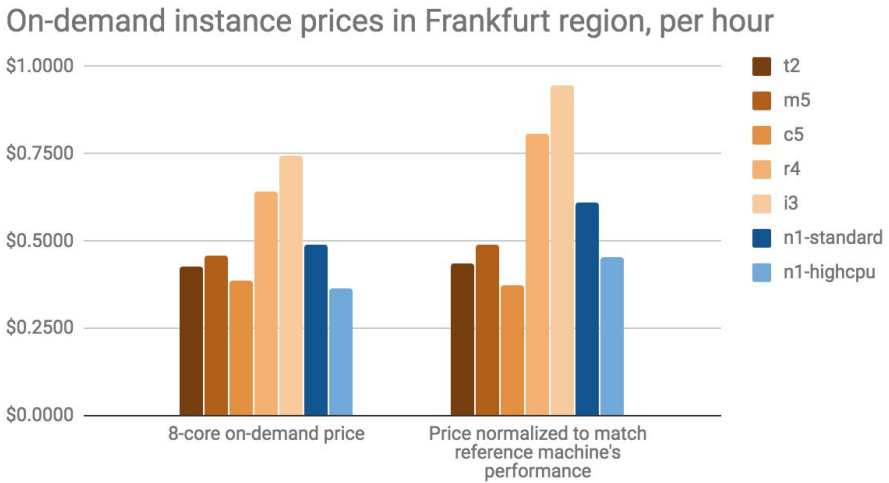
Cloud Computing Resource Allocation (Job)

2. Select number of resources, deploy and use
3. Pay depending on instance type and run time
    - Charged per hour of (vCPU and Memory usage) or (VM uptime) depending on Provider
    - Subject to service availability

ON DEMAND AZURE VS AWS PRICING

LINUX OS (AUGUST 2018)

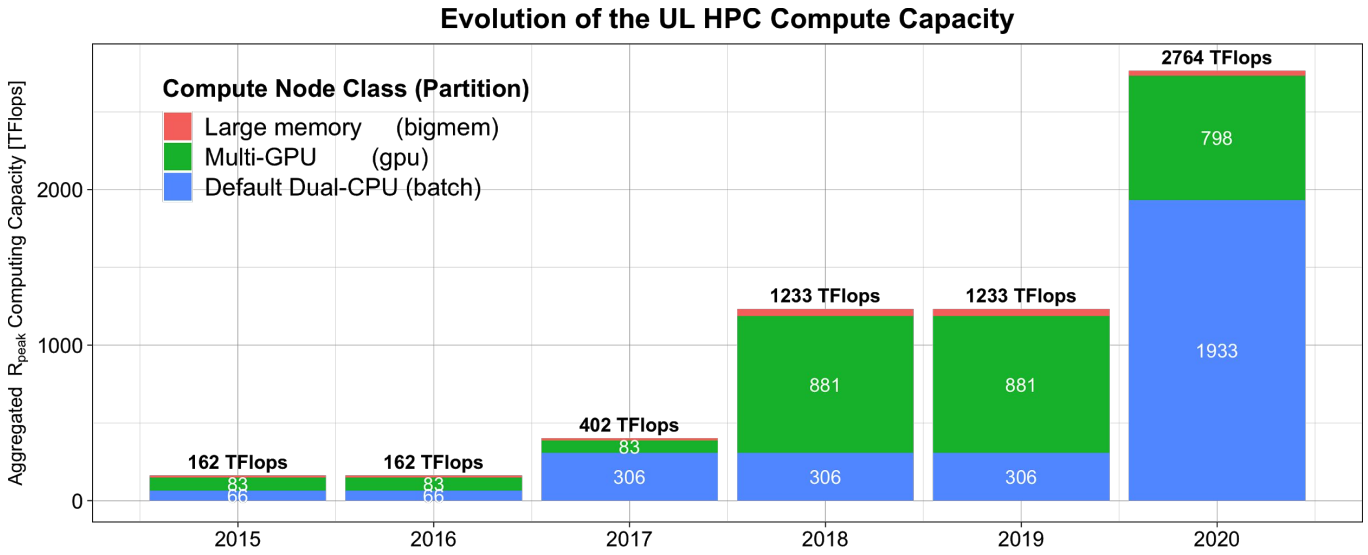
TYPE	vCPU	MEM.	AZURE	AWS
General Purpose	2	8GB	$0.0840	$0.0928
	4	16GB	$0.1670	$0.1856
	8	32GB	$0.3350	$0.3712
Compute Optimized	2	4GB	$0.0850	$0.0850
	4	8GB	$0.1690	$0.1700
	8	16GB	$0.3380	$0.3400
Memory Optimized	2	16GB	$0.1330	$0.1330
	4	32GB	$0.2660	$0.2660
	8	64GB	$0.5320	$0.5320



hpc. uni.lu

HPC Facility @ Uni.lu

- 3 types of computing resources:
  - "*regular*" nodes: Dual CPU, no accelerators, 128 to 256 GB of RAM
  - "*gpu*" nodes: Dual CPU, 4 Nvidia accelerators, 768 GB RAM
  - "*bigmem*" nodes: Quad-CPU, no accelerators, 3072 GB RAM



High Performance Computing & Big Data Services

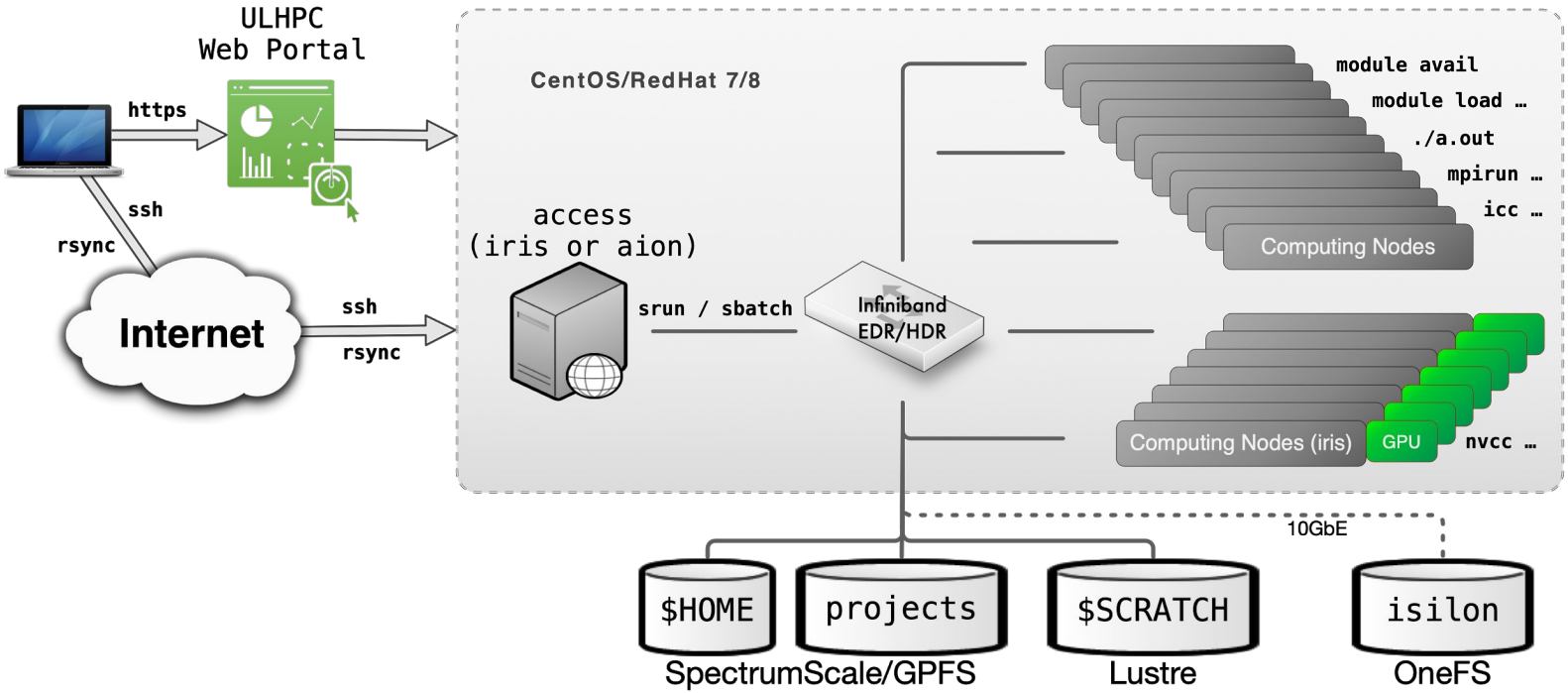
hpc.uni.lu

hpc@uni.lu

@ULHPC

LUXEMBOURG
LET'S MAKE IT HAPPEN

HPC Resource Allocation (Job)



HPC Resource Allocation and Billing (Job)

1. Select instance type of computing resource
2. Select number of resources (computing nodes/cores/memory/GPUs), reserve and use
3. Pay depending on resource & run time: $\texttt{NNodes} \times [\alpha_{cpu} \times \texttt{Ncores} + \alpha_{mem} \times \texttt{Mem} + \alpha_{GPU} * \texttt{Ngpus}] \times T_{exec}$
   a. Charged per hour of (vCPU and Memory and GPU usage)
   b. Subject to service availability

Cluster	Type of node	Partition / Usage	#Cores per node	CPU	α_{cpu}	α_{mem}	α_{GPU}
Iris, Aion	regular	interactive	28/128	n/a	0	0	0
Iris	regular	batch, long	28	broadwell	1.0*	1/4 = 0,25	0
Iris	regular	batch,long	28	skylake	1.0	1/4 =0,25	0
Iris	gpu	GPU	28	skylake	1.0	1/27 ≈ 0,037	50,0
Iris	bigmem	Large-memory	112	skylake	1.0	1/27 ≈ 0,037	0
Aion	regular	batch, long	128	epyc	0,57	1/1.75 ≈ 0,57	0

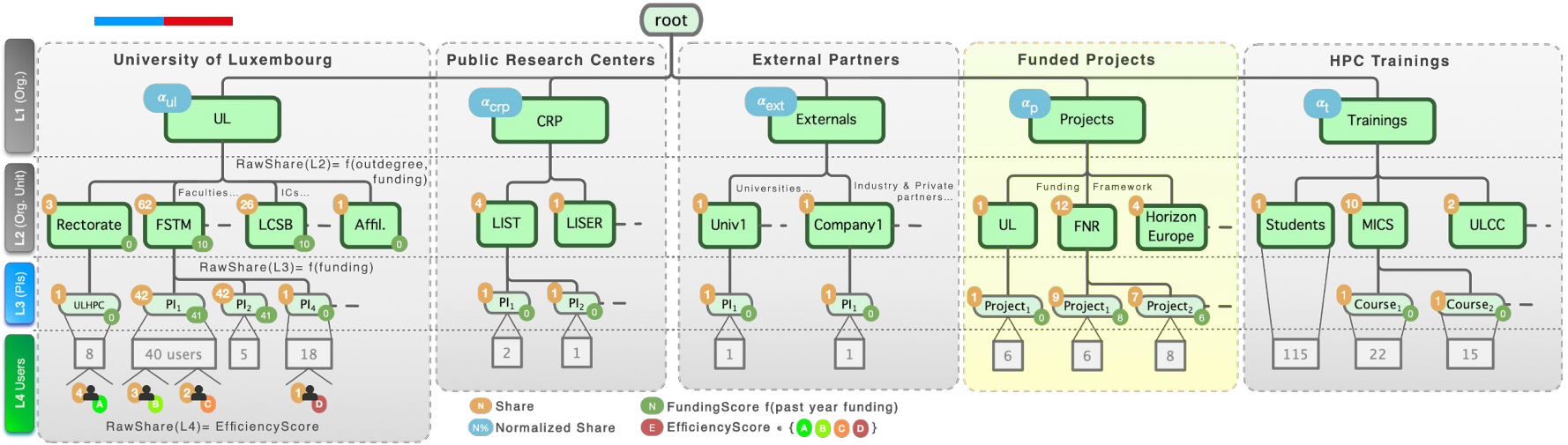
Examples of Billing Usage

- Continuous use of **2 regular skylake nodes (56 cores, 224GB Memory)** on `iris` cluster
  - 28 cores per node, 4 GigaByte RAM per core i.e., 112GB per node
  - For 30 days: 2 nodes * [(α_{cpu} + α_{mem} x 4) x 28 cores] * 30days * 24hours =80640 SU = 2419,2€ VAT excluded
- Continuous use of **2 regular epyc nodes (256 cores, 448GB Memory)** on `aion` cluster
  - 128 cores per node, 1,75 GigaByte RAM per core i.e., 224 GB per node
  - For 30 days: 2 nodes * [(α_{cpu} + α_{mem} x 1,75) x 128 cores] * 30days * 24hours = 289382,4 SU = 8681.47€ HT
- Continuous use of **1 GPU nodes (28 cores, 4 GPUs, 756GB Memory)** on `iris` cluster
  - 28 cores per node, 4 GPUs per nodes, 27 GigaByte RAM per core, 756 GB per node
  - For 30 days: 1 node * [(α_{cpu} + α_{mem}*27) x 28 cores + α_{GPU}*4 GPUs]* 30days * 24h = 184320 SU = 5529.6€ HT

For the provisioning: If the project PI is not able to anticipate the type and amount of resources needed, **we suggest to provision 5529,60€ for every 12 PM of funded personnel** (i.e., 1 month of continuous usage on the most expensive type of resource).
In particular, by default:

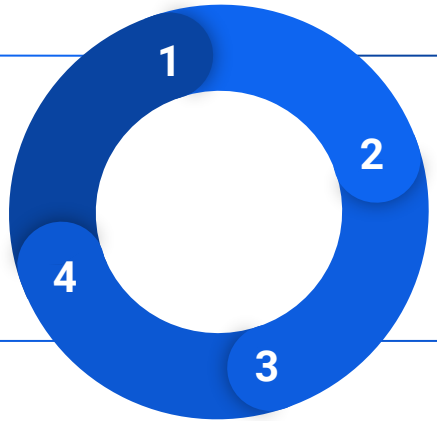
- **Budget for 1 funded PhD student (36+12PM): 22118,4€**
- **Budget for 1 funded PostDoc (24PM): 11059,2€**

Project Allocations, Accounting and Reporting



- UL HPC **compute time** accounted for each allocated resource and instance type
  - Project compute time allocations under the billing model presented previously
  - Easy reporting, per Project, End user; for any time interval with an **exact** usage reported
  - End user task scheduling under fair-share
- UL HPC **data storage** allocated in TB/month
  - Allocation done/increased on request - Default capacity (1TB) for project data free of charge
- UL HPC **expert time** allocated in pools of 4h (Default UL expert rate)

Project Workflow & End User-view



1. Project granted

Contractual agreement signed with funding agency
Provisional HPC budget agreed and setup with accounting dept.
Project details sent to ULHPC team, incl. target capacities

0. Project proposal preparation

Budget plan incl. planned computing/storage costs

2. Initialization

Project groups, data storage spaces, user accounts, project accounts are created.
End users can now access HPC resources.

4. Finalization

Final Project reporting and billing.
Project data migrated from HPC resources.
Project groups, data storage, user accounts closed.

3. Project underway

Project members can run computational jobs, access and transfer project data, request and receive help from HPC experts.
Regular usage report delivered for monitoring/billing

HPC compute time allocation accessible by project members by selecting the project-specific account:
`srun/sbatch --account PROJECT_NAME [...]`

HPC data storage spaces under project-specific path, accessible by project members:
`/work/projects/PROJECT_NAME`

Compute time billing can be reported on at any time and for any time interval:
`sreport -T billing -t hours cluster AccountUtilizationByUser Accounts=PROJECT_NAME Start=2021-01-01 End=2021-12-31`

Data storage space live allocation and utilization can be reported at any time: `df-ulhpc [-i]`