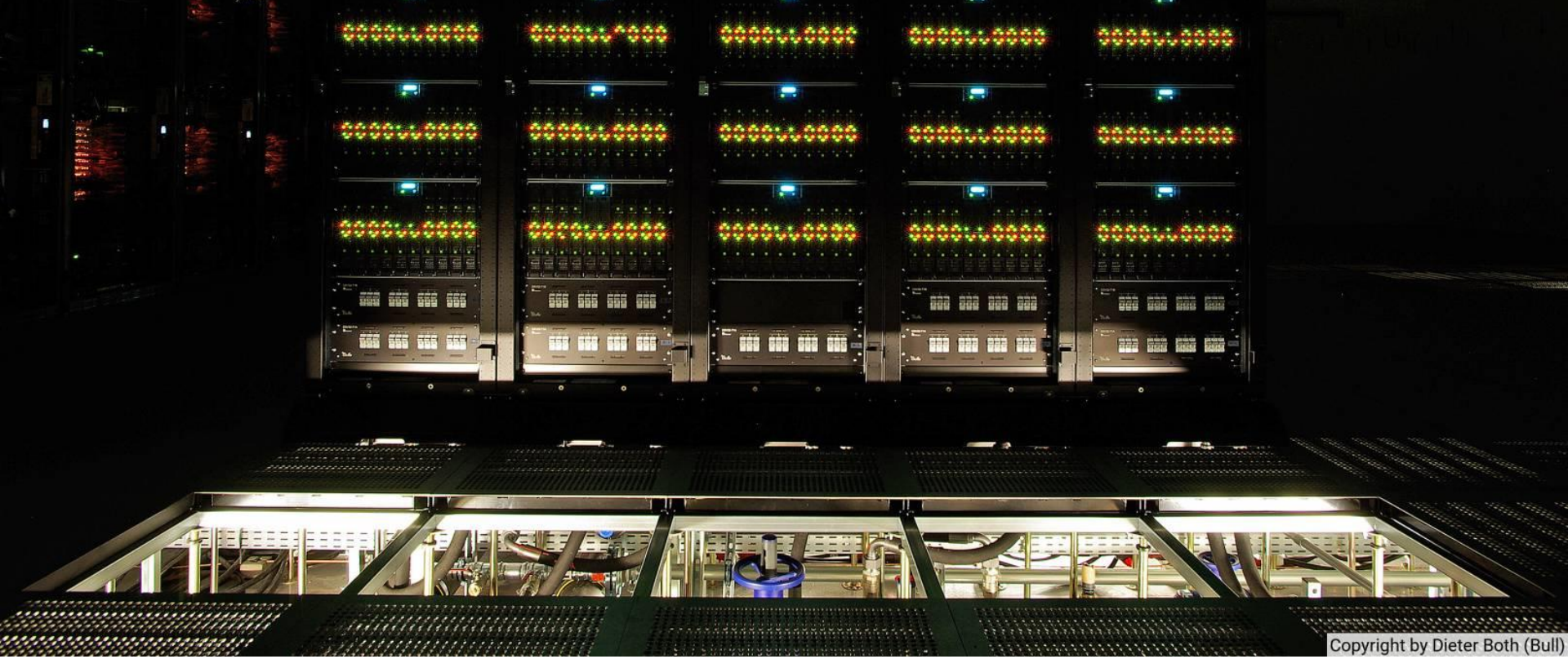


HPC Resources and Computing Projects

Dieter an Mey
IT Center, RWTH Aachen University

HPC Intro, March 3, 2020

IT Center | RWTH AACHEN UNIVERSITY

Overview

- **Access**
- **Core hours**
- **HPC Resources: Computers, Storage, Software**
- **Computing Projects**
- **Links + Upcoming Events**

Access

Who can use the RWTH Compute Cluster?

Access to the RWTH Compute Cluster

- **Members of RWTH Aachen University**
  can use the RWTH Compute Cluster
  for education and research without any fees
  up to certain default limits (compute and storage resources)
- **Employees of RWTH and of Research Center Jülich (FZJ)**
  for education and publicly funded research projects without any fees.
  an application for a computing project has to be submitted for more resources.
  FZJ: one-time registration in the RWTH Identity Management
- **Students of RWTH Aachen University**
  lecturers can apply for resources for participating students.
  Theses advisers can apply for resources for the candidates.
- **Scientists of other German institutions in education and research can submit proposals for large computing projects**
- **Anyone can take part in a computing project as a project member under the responsibility of a PI (principal investigator of an existing computing project)**

https://doc.itc.rwth-aachen.de/display/CC/Access

Access to other HPC Resources

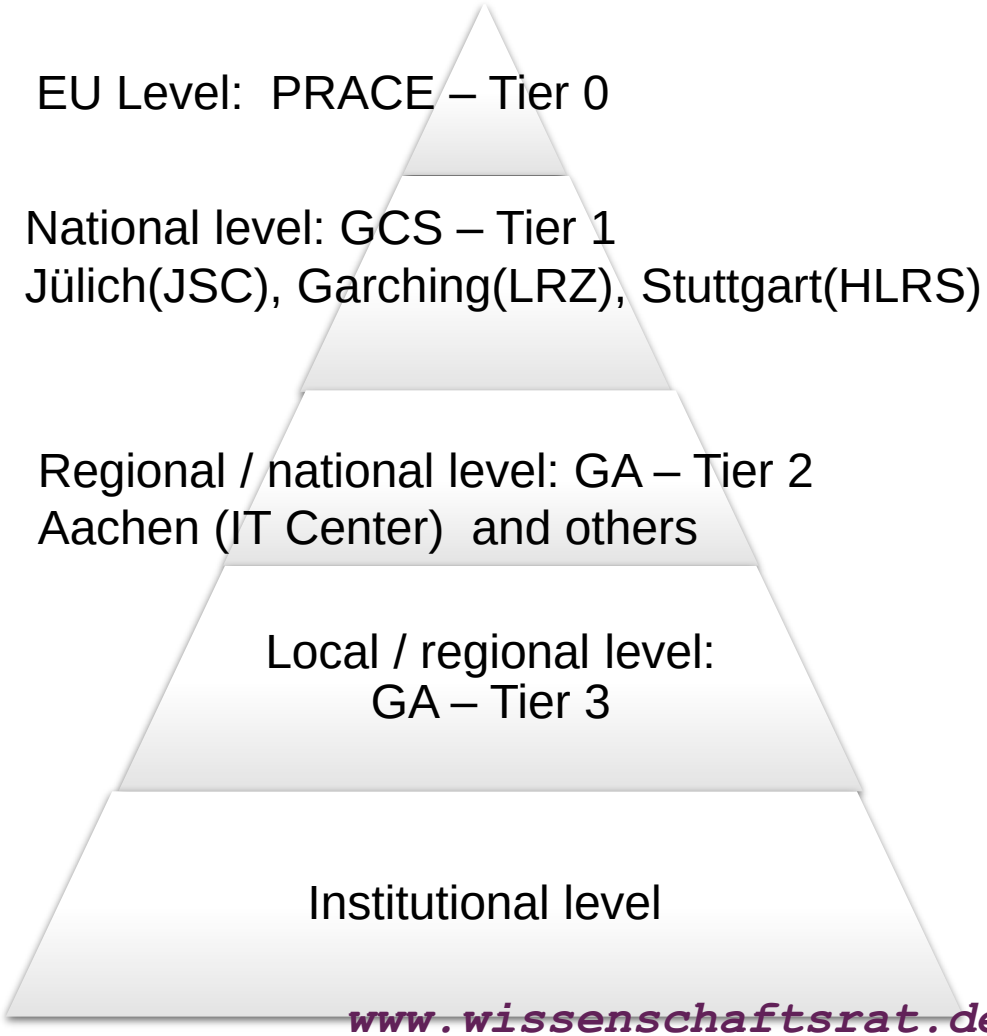
Apart from using local compute resources scientists at RWTH can apply for resources on

- **machines of Research Center Jülich (mainly an Intel Xeon Phi based Booster system) in the context of the JARA-Partition.**
- **national HPC facilities organized in the Gauss Centre for Supercomputing**
- **European facilities in the context of the Partnership for Advanced Computing in Europe (PRACE).**
- **Such an application is carefully reviewed taking technical and scientific aspects into account.**

Citizens of countries that are subject to the export control policy of the German Federal Government may need additional authorization from the German Federal Office for Economic Affairs and Export Control (BAFA) before they are allowed to use the RWTH Compute Cluster.

https://doc.itc.rwth-aachen.de/display/CC/Access

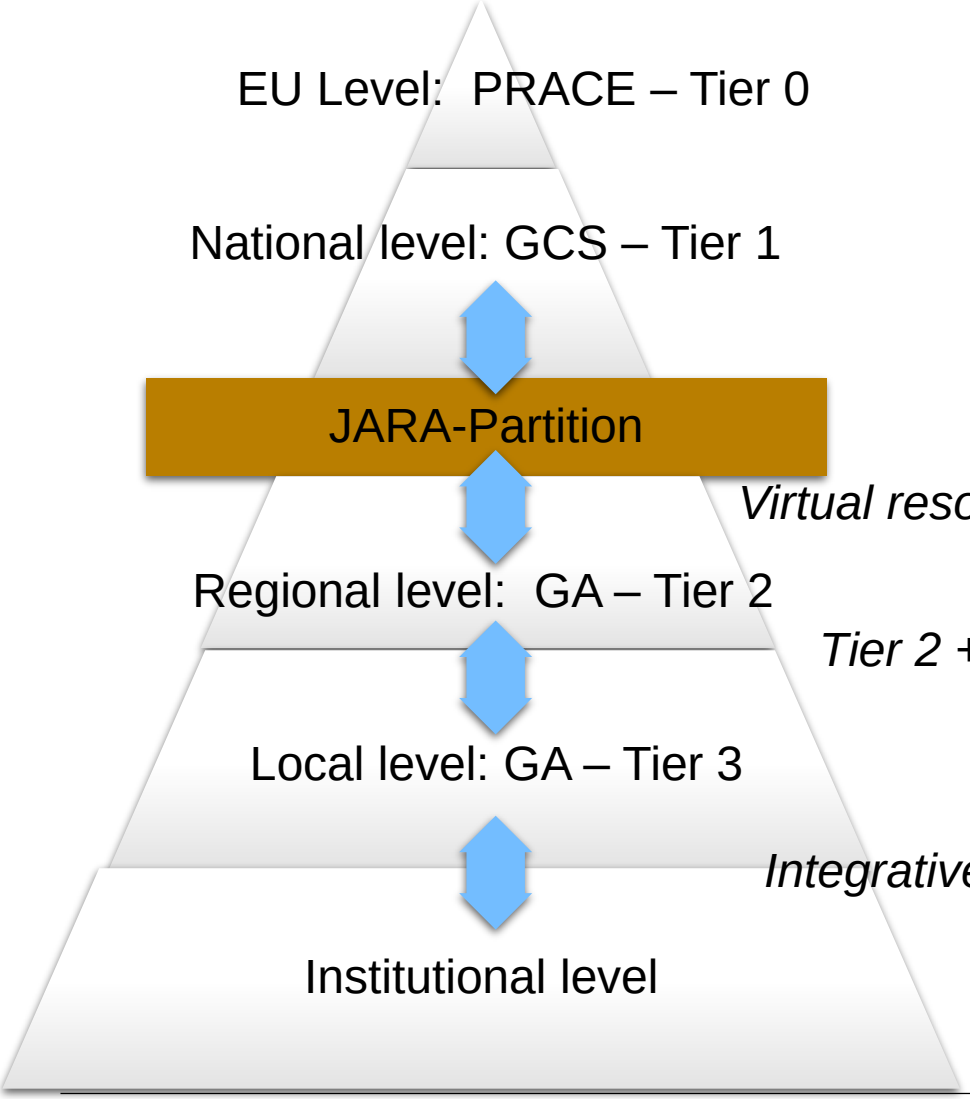
HPC Pyramid – Structure of the Supply of Computing Resources



- Tier 0: European Level: Partnership for Advanced Computing in Europe (PRACE)
- Tier 1: National Level: Gauss Centre for Supercomputing
- Tier 2: Regional-National Level: Gauß Allianz (GA)
- Tier 3: Local (university) - Regional Level: Gauß Allianz (GA)
- Institutional Level
- To learn about future trends: Nationales Hoch und Höchstleistungsrechnen (NHR)

www.wissenschaftsrat.de/download/archiv/4488-15.pdf

Refined RWTH/FZJ HPC Pyramid



- Tier 1.5: JARA-HPC (since 2012)
  Resources for researchers of FZJ and RWTH
- Making transitions easy: integrating resources from different levels

Virtual resource consisting of Tier 1 + 2 resources

Tier 2 + Tier 3 within a single cluster

Integrative hosting of institute resources

Access to HPC Resources

- File applications for PRACE high-end projects on resources currently located at France, Germany, Italy, Spain and Switzerland
  via ***www.prace-ri.eu***

- File applications for GCS large-scale projects (> 35 Mio Coreh) on resources at Jülich, Garching or Stuttgart twice a year directly at one of the 3 national centers, see ***www.gauss-centre.eu***

- File applications for GCS regular projects (< 35 Mio Coreh) at any time (Garching, Stuttgart) resp. twice a year (Jülich) directly at one of the 3 national centers, see ***www.gauss-centre.eu***

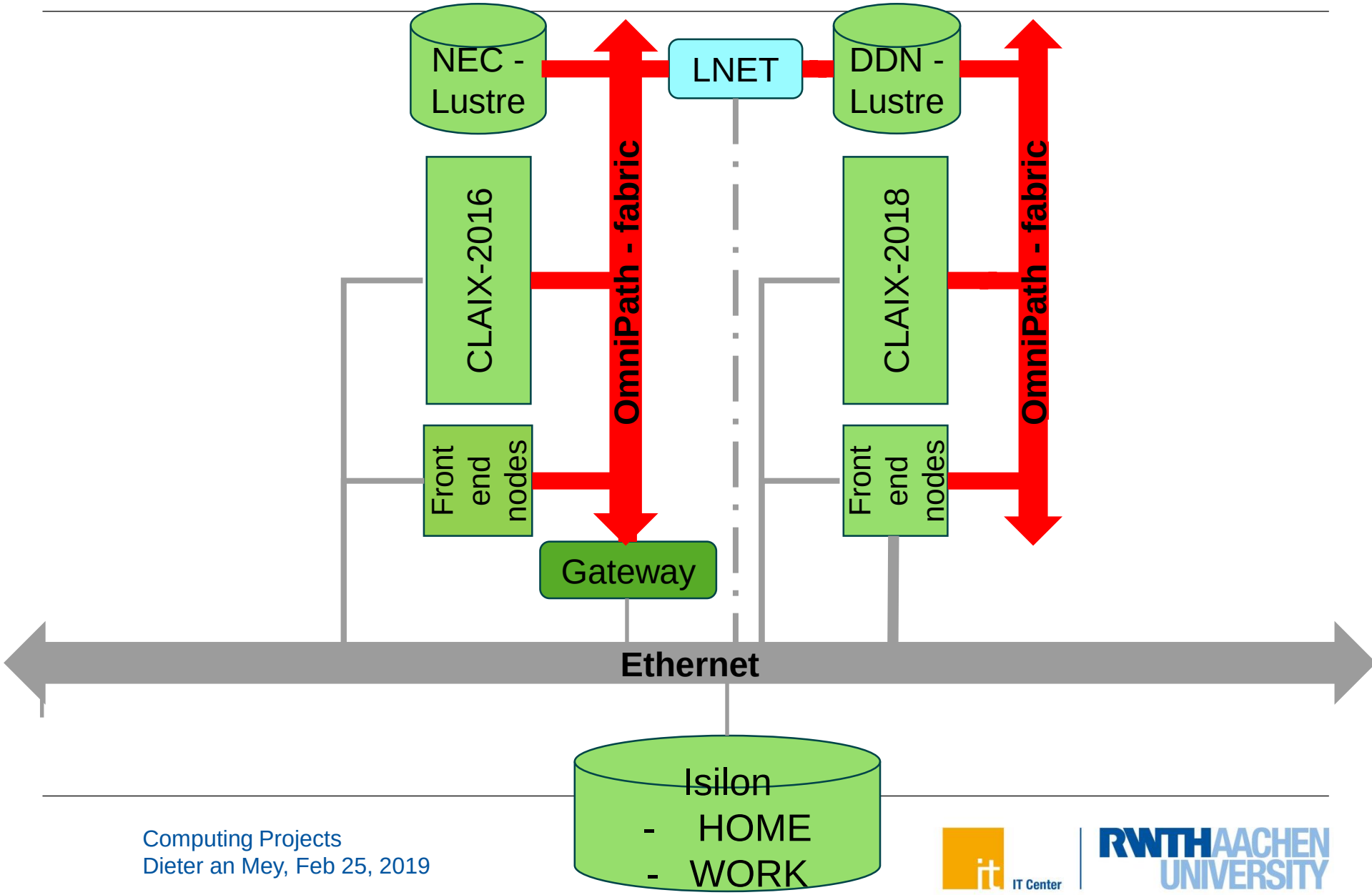
- File applications for JARA-HPC resources twice a year (researchers of FZJ and RWTH only) (>2.4 Mio Coreh)
  see ***https://www.jara.org/de/654***

- File applications for the RWTH Compute Cluster any time
  see ***www.itc.rwth-aachen.de/hpc-projects***

HPC Resources

Computing – Storage – Software (– Networks)

RWTH Compute Cluster Configuration Overview



CLAIX-2016 (Tier 2) versus CLAIX-2018 (Tier 2 + Tier 3)

- **609 MPI nodes**

2-socket Intel Xeon E5-2650v4 (Broadwell-EP)
12 cores each (24 cores total per node)
2.2 GHz, 2.9 GHz turbo mode
128 GB main memory
Local SSD: 120 GB

- **10 MPI-GPU nodes**

MPI nodes w/ 2 NVIDIA P100 SXM2 GPUs
NVLink GPU-GPU interconnect:

- **8 SMP nodes**

8-socket Intel Xeon E7-8860v4 (Broadwell-EX)
18 cores each (144 cores total per node)
2.2 GHz, 3.2 GHz turbo mode
1024 GB main memory

- **Frontend and service nodes**

- **1032 + 216 MPI nodes**

2-socket Intel Xeon Platinum 8160 (Skylake)
24 cores each (48 cores total per node)
2.1 GHz, 3.7 GHz turbo mode
192 GB main memory
Local SSD: 480 GB

- **48 + 6 MPI-GPU nodes**

MPI nodes w/ 2 NVIDIA V100 SXM2 GPUs
NVLink GPU-GPU interconnect

- **Frontend and service nodes**

- **Vendor:** NEC

Computing Resources

CLAIX-2018 - NEC

3.5 PFlop/s
= 3.5 * 10^15 Flop/s

seit Dezember 2018

Ebene 2 HPC-System | GPU-Knoten | MPP-System | SSD Speicher

1.307 Knoten
Omnipath

251 TB
Hauptspeicher

62.736 CPU-Kerne
Intel

96 GPGPUs
Nvidia

CLAIX-2016 - NEC HPC1812-Rg-2

seit Dezember 2016 - Top500.org

Ebene 2 HPC-System | GPU-Knoten | MPP-System | SMP-Knoten

678 TFlop/s
Gesamtleistung

633 Knoten
Omnipath

88 TB
Hauptspeicher

16.152 CPU-Kerne
Intel

20 GPGPUs
Nvidia

https://gauss-allianz.de/de/profile/IT%20Center%20der%20RWTH%20Aachen

(Mio) Core hours – the Metric for Computing Resources

- **Core hour = usage / reservation of one core for one hour**

- **Compute resources are available up to a certain limit (measured in Core-Hours per year)**
- **Technically everyone with an account on the cluster has a certain budget**

- **8760 hours per year / 720 hours per month**
- **48 cores per compute node (CLAIX-2018-MPI)**

=> 0.42 Mio Coreh per node per year / 35.000 Coreh per node per month

Personal default quota on the RWTH Compute Cluster

- **Monthly quota for a RWTH scientist: 2000 hours (nearly 1 PC / laptop continuously)**
- **Monthly quota for a RWTH student: 500 hours (nearly 1 core continuously)**

Inquiring user‘s default quota

```
r_wlm_usage -q
User:                                          user1234
Status of user:                     RWTH-Mitarbeiter
Quota monthly (core-h):                          2000
Remaining core-h of prev. month:                   53
Consumed core-h current month:                    818
Consumable core-h (%):                             62
Consumable core-h:                               3235
```

Sliding Window (3 months)

53 (remainder from previous month)
+ 2000 (for the current month)
+ 2000 (for next month)
– 818 (consumed this month)
= 3235 Core-h left over to be consumed this month at most !
or 1235 Core-h left over for this month = 62% * 2000

Inquiring project quota

```
r_wlm_usage -q -p proj1234
Group:                                          proj1234
Start of Accounting Period:                   01.02.2018
End of Accounting Period:                     01.02.2019
State of project:                                 active
Quota monthly (core-h):                            75000
Total quota (core-h):                           0.900 Mio
Remaining core-h of prev. month:                   15053
Consumed core-h current month:                     21452
Consumable core-h (%):                                91
Consumable core-h:                                143602
```

15053
+75000
+75000
-21452
= 143602

```
r_wlm_usage -p proj1234 -m 3  # display 3 months

| Project  | Queue | User     |   current |  Dec 2018 |  Nov 2018 |      total |
| -------- | ----- | -------- | --------- | --------- | --------- | ---------- |
| proj1234 | rwth  | user1234 |           |           | 50,012.28 |  50,012.28 |
| proj1234 | rwth  | user1235 |           |           |    409.16 |     409.16 |
| proj1234 | rwth  | user1236 |           |  6,725.58 |    534.81 |   7,260.38 |
| proj1234 | rwth  | user1237 | 21,451.77 | 53,220.94 |  3,631.62 |  78,304.33 |
| proj1234 | total |          | 21,451.77 | 59,946.52 | 54,587.84 | 135,986.12 |
```

GigaBytes – The Metric for Storage Resources

- **Where to store your data?**
  How much data?
  Where is your data located?

1 Byte = 8 bits = space for 1 character
1 GigaByte (MB) = 10^6 Bytes
1 GigaByte (GB) = 10^9 Bytes
1 GigaByte (TB) = 10^12 Bytes
1 GigaByte (PB) = 10^15 Bytes

- **Data to be computed has to reside in the computer's main memory:**
  CLAIX-2016: 128 GigaBytes per 24 cores
  CLAIX-2018: 192 GigaBytes per 48 cores

- **Scratch files can be put into the compute node's scratch (/tmp) file system for the life time of a batch job**
  CLAIX-2016: 120 GigaByte of fast SSD
  CLAIX-2018: 480 GigaByte of fast SSD

- **Permanent files reside on the WORK-Fileserver (no back-up, limited life time) or on the HOME-Fileserver (Back-up)**

- **Lustre parallel Filesystem (HPCWORK)**

Storage Hierarchy

Storage	Life time of data	Total Size	Default quota	Peak Bandwidth	remark	usage
memory	program	128 GB / 192 GB		40 GB / 140 GB/s	volatile	CLAIX-2016 / CLAIX-2018
NVME	*job*	*2 TB*		*> 1 GB/s*	*Automatically cleaned*	*only few CLAIX-2016 nodes*
SSD (solid state disk)	job	120 GB / 480 GB		~ 1 GB/s	local disk for each node, automatically cleaned	CLAIX-2016 / CLAIX-2018 $TMP
BEEOND	job	#nodes * 480 GB		#nodes * 1 GB/s	temporary parallel scratch file system for a parallel job running on #nodes nodes	CLAIX-2018 (mid 2020)
Lustre parallel filesystem	Job / short term medium storage	3 PB / 10 PB	default 1TB max 30 TB	30 / 100 GB/s	shared between all cluster nodes, no backup suitable for large blocks of IO	CLAIX-2016 / CLAIX-2018 $HPCWORK
WORK filesystem	medium term storage	< 1 PB	default 250 GB max 1 TB	<1 GB/s	shared between all cluster nodes, no backup	$WORK
HOME filesystem	permanent storage	< 1 PB	default 150 GB max 500 GB	<1 GB/s	shared between all cluster nodes, backup	$HOME
ARCHIVE	long term storage	> 1 PB		< 1 GB/s	shared between all users of RWTH Research data has to be stored 10y (DFG requirement)	

https://doc.itc.rwth-aachen.de/display/CC/Available+File+Systems

(Application) Software

The bad news:

- **Users are responsible for providing licensies for software they need**
  - unless the software usage is free
  - or the software has been licensed centrally
- **Unfortunately, regulations vary considerably between different software packages.**
- **Users from external institutions may need to provide licences**

The good news:

- **The operating system – CentOS Linux**
- **the batch system – Slurm**
- **the development environment – compilers, tools, MPI library**
- **and many other software packages**

are free or are provided by the IT Center

- **see https://doc.itc.rwth-aachen.de/display/CC/Installed+Software**
- **and https://doc.itc.rwth-aachen.de/display/CC/Usage+of+Software**

Software availability

no star - everyone can use this software.

★ software is restricted to members the RWTH University, external users - like members of the FH Aachen, the Research Center Jülich, etc. are NOT allowed to use this software.

☆ software is restricted to employees and / or institutes, which means students are NOT allowed to use this software.

★ this software needs an additional user registration or a special license.

TECHNICS

- ☆ abaqus
- ☆ adams
- ★ ansys
- ★ cfx
- ★ comsol
- ☆ dymola
- ☆ dytran
- fds
- ★ fluent
- ☆ hyperworks
- ★ icem
- ☆ lsdyna
- ☆ lsopt
- lsprepost
- ☆ marc
- ☆ nastran (MSC)
- openfoam
- ★ petromod
- ☆ starcd
- ☆ starccm
- superforge

CHEMISTRY

- abinit
- cp2k
- ★ cpmd
- dalton
- ★ gamess
- ★★ gaussian
- ★ gaussview
- gromacs, VMD
- lammps, liggghts
- meep
- ★ molcas
- openbabel
- openmx
- qe
- siesta
- ★ turbomole[-mpi|-smp]
- ★ vasp

MISC

- freerdp
- gnuplot
- ★ matlab
- opnet
- revolve
- tensorflow

GRAPHICS

- ☆ ensight
- paraview
- ★ tecplot
- tecplotchorus
- visit
- wave

MATH

- gurobi
- ☆ maple
- ☆ mathematica

Check for updates !

https://doc.itc.rwth-aachen.de/display/CC/Installed+software

Computing Projects

Management of Computing (and Storage) Resources

What is a Computing Project ?

Organisationally

- **A measure to manage large amounts of compute cycles to fullfill the demands of the financing agencies (DFG, WR) for scientifically soundness**
- **the more cycles, the higher the quality requirements**

Technically

- **A linux group and associated file space**
- **A list of users (Linux user names) in that group**
- **A category for (Slurm) batch jobs for these users having higher limits In addition to your personal default resources (concerning file space and compute cycles)**
- **A bunch of (sysadmins‘) scripts for book keeping**
- **Data stored in the online submission and reviewing system JARDS and related software packages**

Slurm

- **It‘s all about batch jobs …**
- **Interactive usage is not accounted for - it is limited anyhow**
- **You are <u>not</u> supposed to comsume a large amount of compute resources interactively**
- **The sysadmins are taking measures to prevent you from doing this.**

- **Since May 1, 2019 all machines of CLAIX-2016 and CLAIX-2018 are operated with the Slurm batch system**

https://doc.itc.rwth-aachen.de/display/CC/Using+the+SLURM+Batch+System

Filing a proposal

Check out *www.itc.rwth-aachen.de/hpc-projects*

- **What is project-based resource management?**
- **Who is (not) affected?**
- **Support for research and education**
- **Roles and tasks related to any computing project**
- **Getting an HPC account for the RWTH Compute Cluster**
- **Filing a compute time application**
  1. Provide a description of the project
  2. fill out the JARDS online form
     a) A PI from a trusted domain (RWTH, FZJ, UKA) can finalize the JARDS online form himself (paperless submission)
     b) The PI has to sign the generated pdf-file, and scan it and send it by email to ***hpc-projects@itc.rwth-aachen.de***
- **Add Acknowledgements to related publications**
- **Compute Project Categories including Links to JARDS**
- **More information (technical details, reviewing)**

Roles and tasks related to a computing project

- **Principal investigator (PI)**
  Leading scientist, senior researcher
  (for projects related to education: Student's Advisor / Lecturer)
  Takes any responsibility for the computing project
  has to "sign" the project application
  is responsible for reporting about the project
  takes part in reviewing other scientists' applications
  (or nominates a replacement reviewer)
- **Person to contact (PC**) (previously called technical contact, TC)
  the IT Center's contact person for any questions
  technical management of the project
  adding project members
  PI = PC is OK.
- **Project members (PM)**
  Anyone whom the PI trusts
  performs the computations / submits the project's batch jobs
  PI and PC can act as project members as well.

Project Categories

Applicants		Project Category	Submission and Reviewing procedures	monthly cpu Mio Core-h		annual cpu time quota Mio Core-h per year	
				min	max	min	max
	Exter-nals	BUND	A detailed project description is required. Scientific and technical reviews	>=0,05	0,5	>=0,6	**6 (100 total per year)**
RWTH FZJ UKA		PREP	Short project description, technical reviews only		0,02		0,24
		NOVA	Short project description, technical reviews only		0,02		0,24
		JARA-HPC = RWTH Large	Submission twice a year following. A detailed project description is required. Technical and external scientific reviews	>0,2		**>2,4**	**(300 total per year)**
		RWTH Medium	A project description is required. Scientific and technical review	>0,02	0,2	>0,24	**2,4**
		RWTH Small	Short project description, technical reviews only	>0,002	0,02	>0,02	0,24
		RWTH lecture	No project description, technical reviews only				0,048 per project
		RWTH thesis	No project description, technical reviews only				
		RWTH open (scientists)	no application, no reviewing		0,002		0,024
		RWTH open (students)	no application, no reviewing		0,0005		0,006

https://doc.itc.rwth-aachen.de/display/CC/Overview+of+Compute+Project+Categories+and+Links+to+Submission+Forms

Number of Submitted Projects by Project Category

Project Category	2016	2017	2018	2019
bund	1	2	1	2
jara	45	37	42	36
nova		12	14	18
prep		7	10	6
rwth-m	26	25	26	28
rwth-s	90	93	116	127
thes	127	152	168	201

Links

Upcoming Events

Weblinks

- www.itc.rwth-aachen.de/hpc
- www.itc.rwth-aachen.de/hpc-projects (Computing Projects)
- https://doc.itc.rwth-aachen.de/display/CC/Home (RWTH Compute Cluster documentation)
- https://doc.itc.rwth-aachen.de/display/CCP/Home (Parallel Programming) is going to be replaced by
- https://hpc-wiki.itc.rwth-aachen.de/hpc/HPC_Wiki (the new HPC Wiki)

- www.itc.rwth-aachen.de/ppces (the annual weekly HPC seminar in Mid-March)
- http://www.itc.rwth-aachen.de/aixcelerate (the annual tuning workshop)
- https://doc.itc.rwth-aachen.de/display/VE/Introduction+to+High-Performance+Computing+2020 (this event)

- https://doc.itc.rwth-aachen.de/display/CC/Installed+Software (installed software)
- https://doc.itc.rwth-aachen.de/display/CC/Usage+of+Software (Software Usage)

Upcoming Events: PPCES March 16-20, 2020 (dates to be confirmed)

The first two days will focus on performance programming of single compute nodes mainly using Shared Memory Programming w/ OpenMP.

Part I: Basic OpenMP Programming - Monday, March 16, 9:00 - 17:30

Part II : Advanced OpenMP Topics - Tuesday, March 17, 9:00 - 17:30

Days 3 and 4 will focus on performance programming of multiple compute nodes using Message Passing with MPI

Part I: Basic MPI Programming – Wednesday, March 18, 9:00 - 17:30

Part II: Further MPI Concepts - Thursday, March 19, 9:00 - 17:30

The last day will focus on accelerator programming, particularly on programming NVIDIA GPGPUs.

Part I: Basic GPGPU Programming Principles with OpenACC - Fri, March 20, 9:00 - 12:30

Part II: Advanced Concepts and Tuning of GPGPUs with OpenACC - 14:00 - 17:30

Upcoming Events: aiXcelerate Nov 30 – Dec 3, 2020 (dates to be confirmed)

Mon + Tue: Tutorial on performance engineering techniques and tools

Wed + Thu: Tuning Workshop with an opportunity to work on own codes under guidance of experts from Intel, RRZE and RWTH

aIXcelerate

Thank you for your attention

Any open questions?

Please, provide your feedback!

https://sharepoint.ecampus.rwth-aachen.de/units/rz/HPC/public/Lists/HPCIntro%20Feedback/overview.aspx