UNIVERSITEIT
GENT

ANNUAL REVIEW HPC-UGENT

2020

Department ICT, Infrastructure office

E hpc@ugent.be
T +32 9 264 4716

Campus Sterre S9
Krijgslaan 281, 9000 Gent

https://www.ugent.be/hpc

CONTENTS

1 ABOUT HPC-UGENT

In scientific computing*, computers are used to solve complex problems. (*aka: supercomputing or high-performance computing - HPC)

1.1 Our mission

HPC-UGent provides centralised scientific computing services, training, and support for researchers from Ghent University, industry, and other knowledge institutes.

HPC-UGent is part of the central ICT department of Ghent University, and is a strategic partner of the Flemish Supercomputer Center (VSC - http://vscentrum.be/).

1.2 Our vision

HPC-UGent offers a professional scientific computing environment that is stable, user-friendly, and serves the diverse purposes of researchers from Ghent University, industry and other research institutions.

We provide a structural training curriculum for new and advanced users, and provide supporting course material.

We present a supercomputing portfolio that is well known within Ghent University and beyond, and we establish ourselves in the international community via contributions to centralised solutions, such as EasyBuild (https://easybuilders.github.io/easybuild/).

1.3 Personnel

In 2020, the HPC-UGent team consists of 9 people:
Álvaro Simón García, Andy Georges, Ewald Pauwels, Kenneth Hoste, Kenneth Waegeman, Stijn De Weirdt, Wouter Depypere, Balázs Hajgató and Bart Verheyde.

Tasks include:
- User support
- Training
- Infrastructure installation and upkeep (both hardware and software)
- Outreach and marketing
- Collaboration with other supercomputing centers

Additionally, several consultants (Inuits) are also regularly engaged to help solve software installation requests: Pavel Grochal, Jakub Zárybnický, Denis Kristak

2 INFRASTRUCTURE

2.1 Overview

The Ghent University compute infrastructure (Tier-2) consists of several specialised clusters, jointly called Stevin, hosted in the S10 datacenter.

In 2020, the following changes to the infrastructure were applied:

- All shared filesystems (HOME, DATA, SCRATCH) were migrated to a new, larger storage system, opening up a total storage space of more than 3.75 PB.
- The InfiniBand network, providing a high-speed interconnect between most clusters and the storage system, was drastically rewired and extended in order to keep on accommodating new HPC hardware.
- Cluster *joltik* was taken in production. This cluster introduced GPU computing capabilities in the Tier-2 infrastructure of Ghent University.
- Cluster *kirlia* was taken in production. This is a new big-memory cluster, intended as a replacement for cluster *phanpy*, which is going offline in 2021.
- Fast scratch area *arcanine* was taken in production. This is a temporary, very fast storage space for 'live' data for calculations, recommended for very I/O-intensive jobs. Total storage space is 70 TB on NVME hardware. *Arcanine* is a replacement for the *phanpy* scratch area, which was gradually phased out by the end of 2020.
- Cluster *golett* reached its end of life and was taken offline.
- Better connectivity between HPC storage and DICT shares was enabled using on-demand mountpoints and Globus. See chapters 6.2.5 and 6.2.6 in the manual (https://www.ugent.be/hpc/en/support/documentation.htm) for more details.
- New cluster *doduo* was initialized and introduced for testing by selected pilot users.

2.1.1 Compute clusters

Cluster name	#nodes	CPU / GPU per node	Usable memory per node	Local disk per node	Network interconnect
Phanpy	16	2 x 12-core Intel E5-2680v3 (Haswell-EP @ 2.5 GHz)	480 GiB	3 x 400 GB (SSD, striped)	FDR InfiniBand
Golett	200	2 x 12-core Intel E5-2680v3 (Haswell-EP @ 2.5 GHz)	53 GiB	500 GB	FDR-10 InfiniBand
		Decommissioned in two stages: first part 13 Oct, second part 9 Nov			
Swalot	128	2 x 10-core Intel E5-2660v3 (Haswell-EP @ 2.6 GHz)	116 GiB	1 TB	FDR InfiniBand
Skitty	72	2 x 18-core Intel Xeon Gold 6140 (Skylake @ 2.3 GHz)	177 GiB	1 TB	EDR InfiniBand
Victini	96	2 x 18-core Intel Xeon Gold 6140 (Skylake @ 2.3 GHz)	88 GiB	1 TB	10 Gb ethernet
Joltik	10	2 x 16-core Intel Xeon Gold 6242 (Cascade Lake @ 2.8 GHz) 4x NVIDIA Volta V100 GPUs (32GB GPU memory)	256 GiB	800 GB SSD	Double EDR InfiniBand
Kirlia	16	2 x 18-core Intel Xeon Gold 6240 (Cascade Lake @ 2.6 GHz)	738 GiB	1.6 TB NVME	HDR-100 InfiniBand
Doduo (pilot)	128	2x 48-core AMD EPYC 7552 (Rome @ 2.2 GHz)	250 GiB	180 GB SSD	HDR-100 InfiniBand

2.1.2 Shared storage

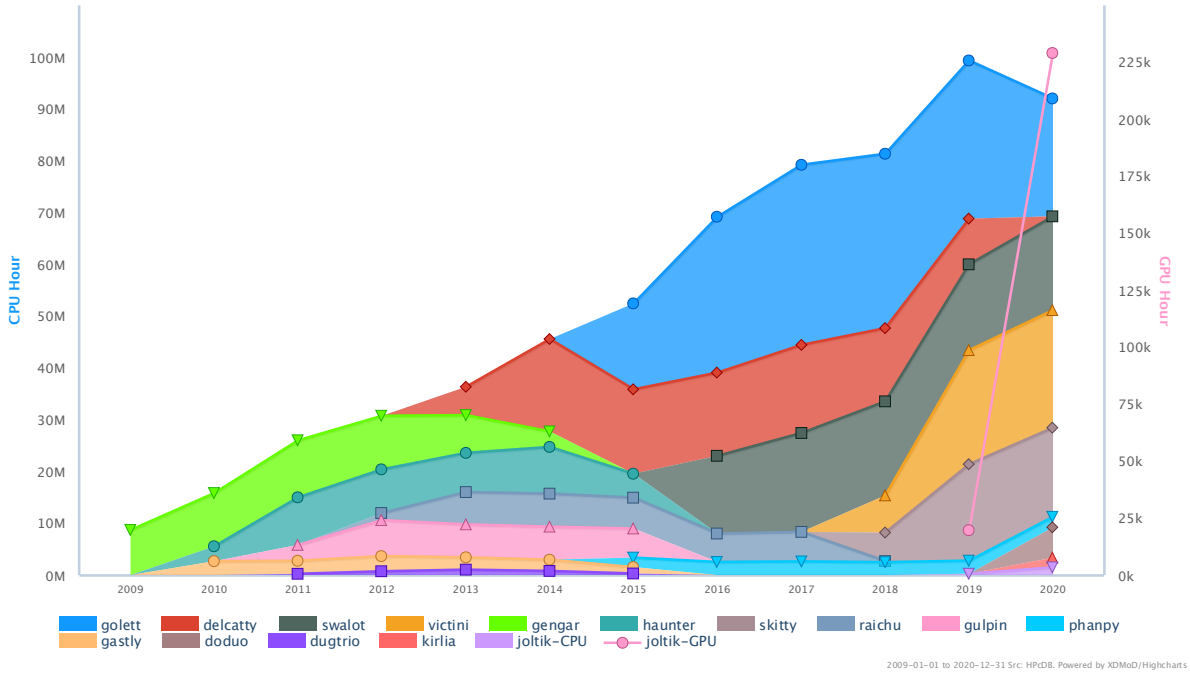
Partition	Size
$VSC_HOME	51 TB
$VSC_DATA	1.8 PB
$VSC_SCRATCH **$VSC_SCRATCH_KYUKON**	1.9 PB
$VSC_SCRATCH_PHANPY	39 TB SSD Phased out Dec
$VSC_SCRATCH_ARCANINE	70 TB NVME

2.2 Usage

In the graphs below, the used or available compute time on the compute clusters is typically expressed in *CPU hours* or *GPU hours*.

- 1 GPU hour corresponds to the work done by a graphical processing unit for one hour of time.
- Modern CPU processors contain many cores that are capable of running a computational task. Since the number of cores varies a lot between processor types and production years, 1 CPU hour is used here to mean the work done by one core in a CPU processor for one hour of time.

2.2.1 Historical perspective on used compute time



This plot gives a historical perspective on the amount of compute time that is used every year, coloured according to the cluster that does the work. All clusters together, since the start in 2009, have now produced more than 606 million CPU core hours (equivalent to 69.196 years of compute work on one core) and 248 thousand GPU hours.

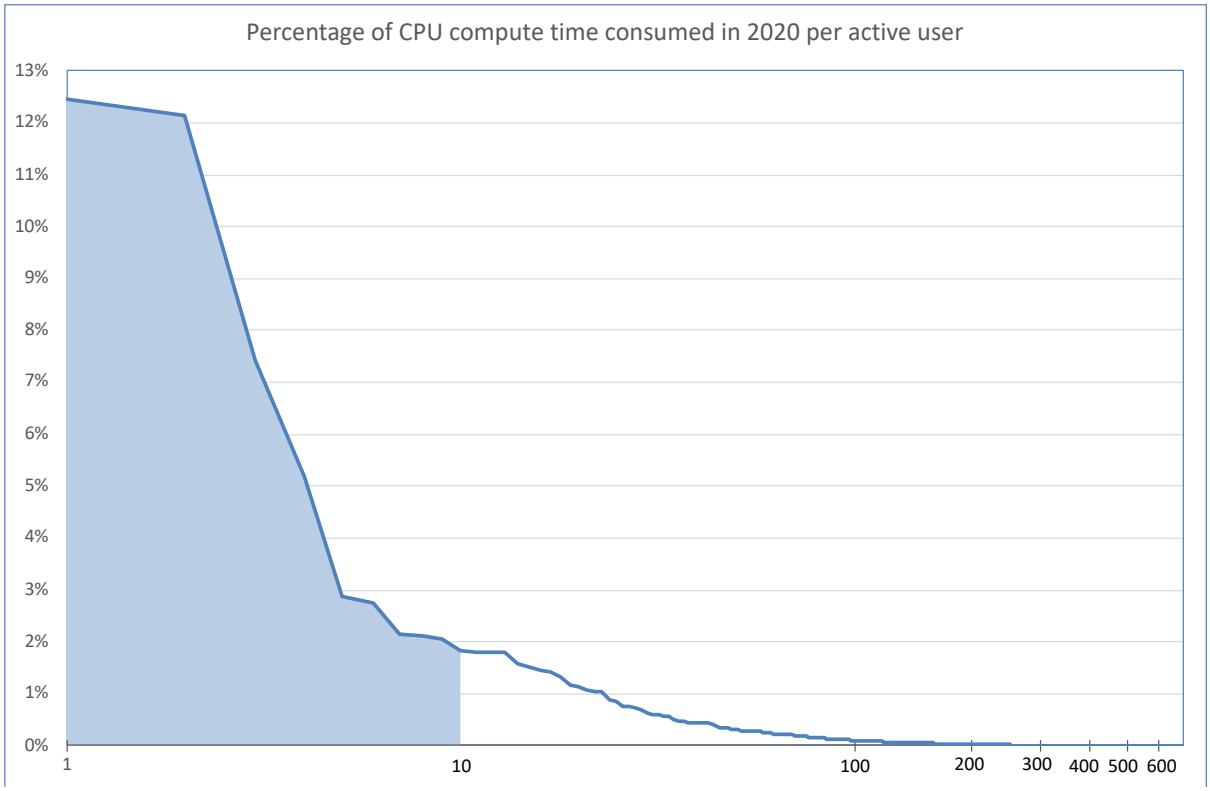
2.2.2 Consumed compute time in 2020

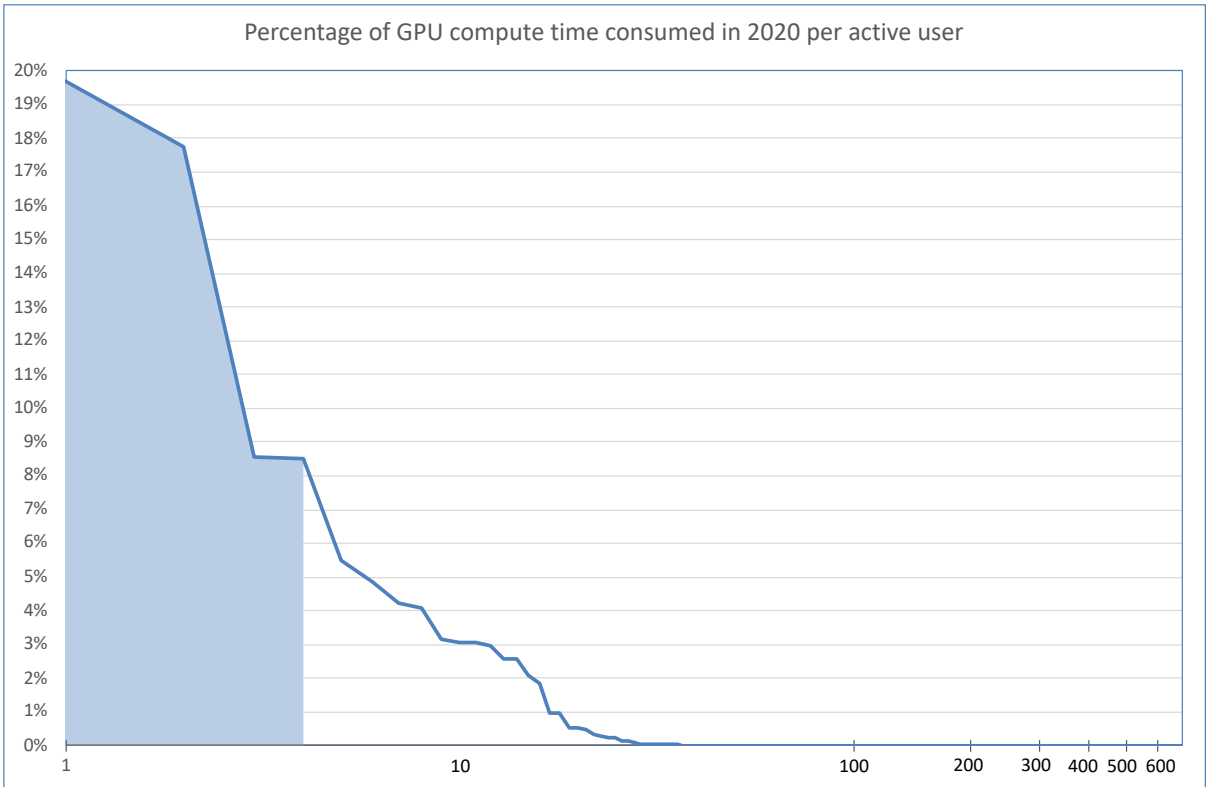
2.2.2.1 Consumed compute time per compute cluster

Cluster name	Compute time consumed		Effective use percentage	
	CPU hours	*GPU hours*	*CPU*	*GPU*
Phanpy	2 024 716		60%	
Golett	23 734 156		69%	
Swalot	18 111 027		81%	
Skitty	17 210 229		76%	
Victini	22 696 689		75%	
Joltik	1 499 605	229 124	53%	65%
Kirlia	1 862 691		42%	
Doduo (pilot)	5 870 828		22%	

In 2020, a total of 93 009 941 CPU core hours (this corresponds to 10 618 core years) and 229 124 GPU hours have been consumed on the Tier-2 compute clusters of Ghent University.
The effective use percentage expresses how much of the theoretically available compute power in one year (#nodes x #cores/node or #GPUs/node) was used. Down-times were not taken into account, so the percentages represent a lower bound.

2.2.2.2 Consumed compute time versus active users





In 2020, 691 persons actively used CPU compute time on the Tier-2 compute clusters of Ghent University, 74 persons actively used GPU compute time.
Both for CPU and GPU, the user base typically contains a limited number of power users in addition to regular users with a lower usage profile. The plots above graph per user (x axis) what percentage of all compute time produced in 2020 this user consumed. The blue area indicates the 50% level of all Tier-2 CPU/GPU compute time consumed.

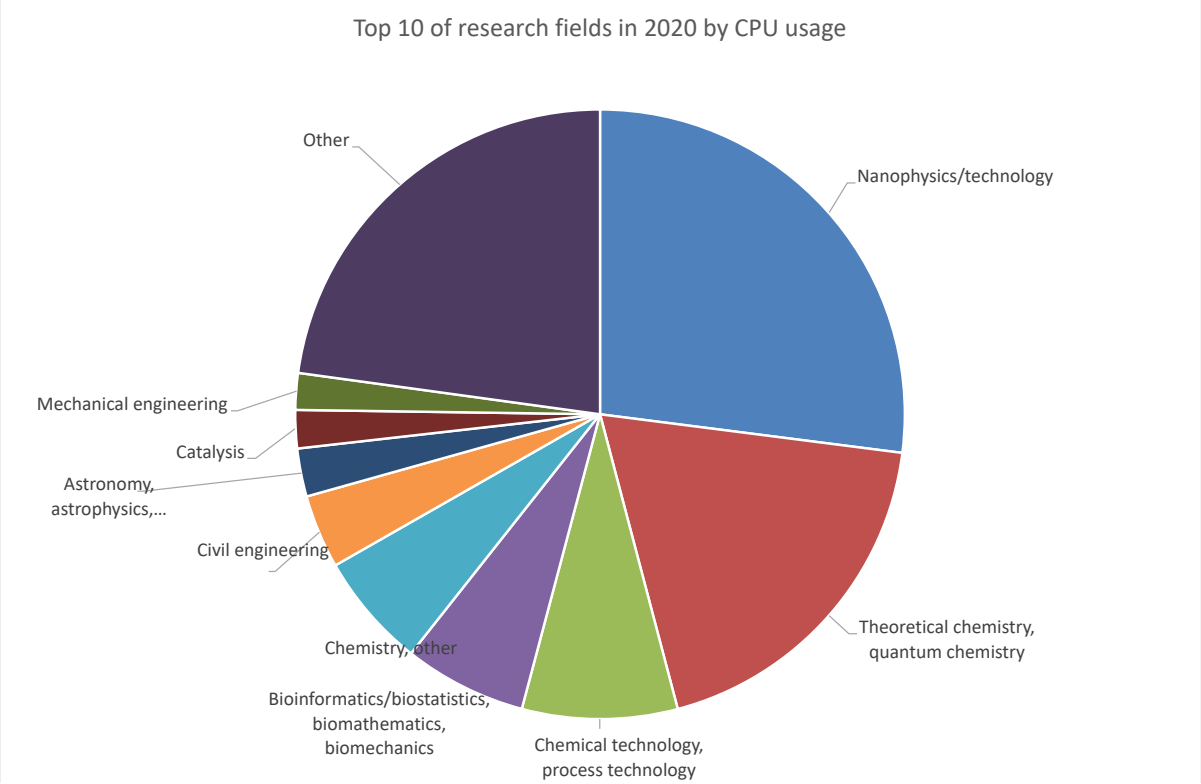
Everyone can get access to the Tier-2 compute clusters of Ghent University. The conditions that apply depend on the affiliation of the researcher. Researchers affiliated with Flemish university associations (constituting the Flemish Supercomputer Center – VSC) get free access. Other Flemish or federal research institutes can get access for their researchers on a contract basis, with 1500 euro worth of compute time given out free of charge per year. Industry can buy compute time in a pay-what-you-use model.

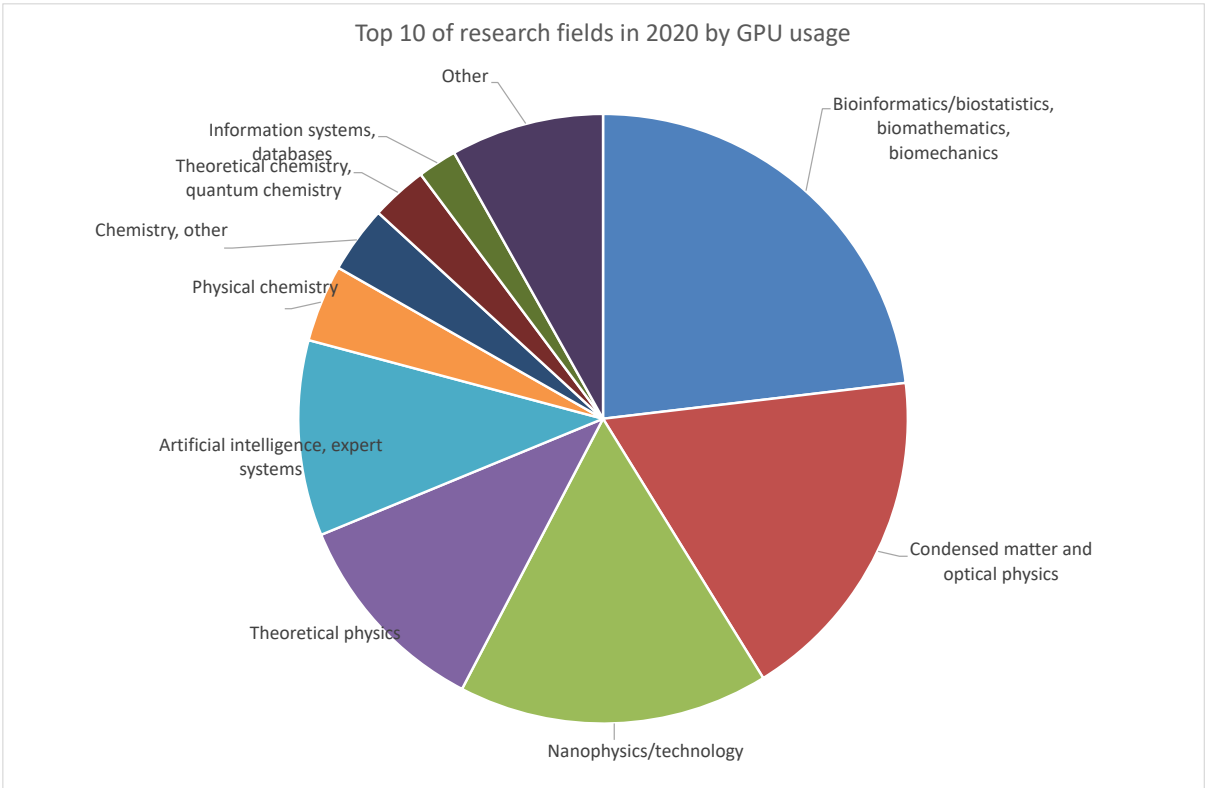
Breakdown of consumed compute time by affiliation	*CPU time*	*GPU time*
UAntwerpen	0.05%	0.00%
VUB	0.16%	5.51%
UGent	99.47%	94.43%
KULeuven / UHasselt	0.01%	0.00%
Other research institutes	0.20%	0.00%
Industry	0.10%	0.06%
Total	**100.00%**	**100.00%**

Master or Bachelor students enrolled in a Flemish university association can also get access to the Tier-2 compute clusters of Ghent University. Several teachers effectively rely on the infrastructure for training purposes, stimulating several students to become a user. Master students often rely on the infrastructure to perform research included in their Master thesis.

Number of students/researchers versus breakdown of consumed compute time by category	*CPU time*	*GPU time*
Ma/Ba students	2.60%	5.32%
Researchers	97.40%	94.68%
Total	**100.00%**	**100.00%**

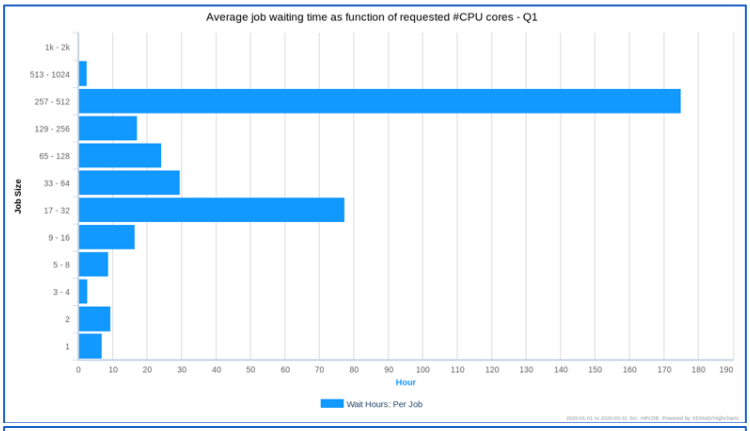
2.2.2.3 Consumed compute time versus research field

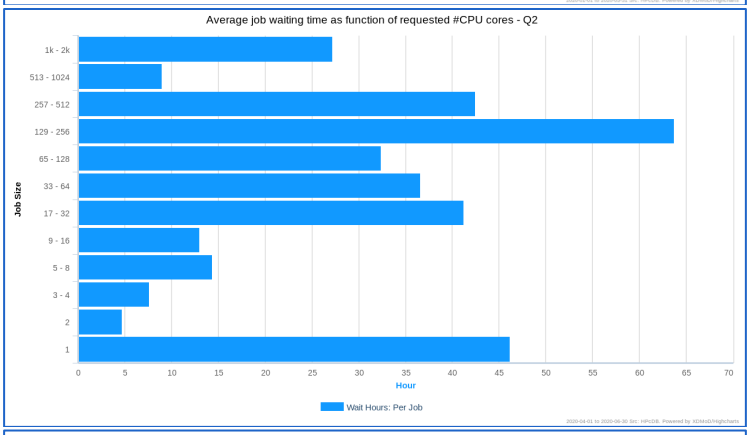


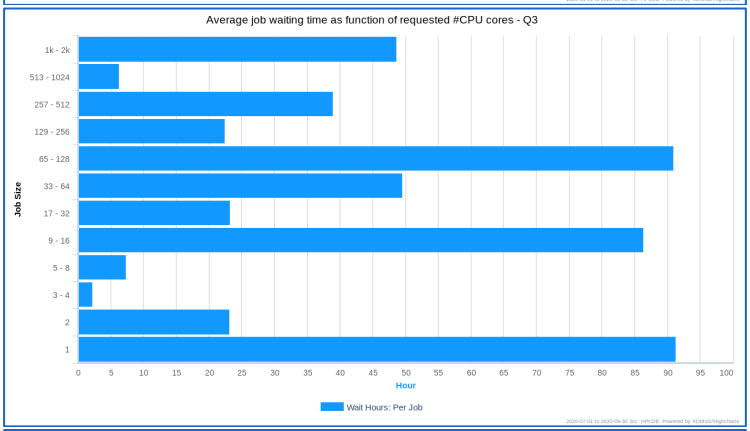


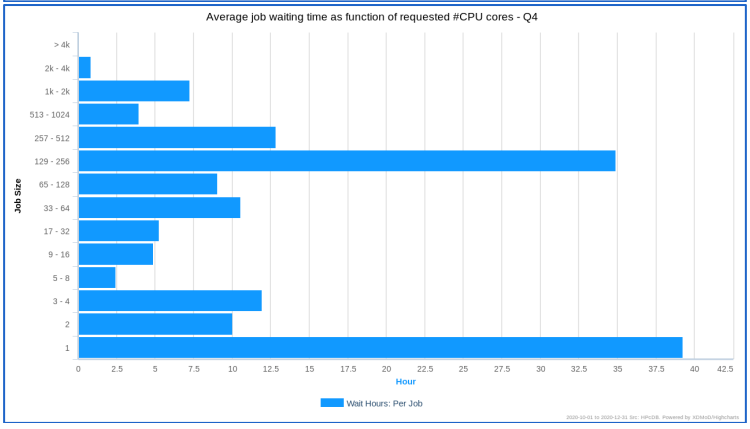
These are the research fields that users enter when requesting their account to access HPC-UGent resources. No doubt, there is overlap between certain fields and some users have not listed a research field at all. As such, the above distributions merely gives an indication of the top research fields that actively use scientific computing, and which ones actively use either CPU or GPU resources.

2.2.2.4 Queue time versus job size on CPU clusters, reported per quarter









2.2.2.5 Queue time versus number of jobs submitted on CPU clusters

	Number of jobs submitted on CPU clusters				Cumulative percentages			
Job Wait Time	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
0 - 1s	97022	23450	20757	89659	10%	3%	4%	18%
1 - 30s	321913	170612	117576	127790	45%	24%	29%	45%
30s - 30min	224662	182515	118117	77047	69%	46%	54%	60%
30 - 60min	41069	67601	38380	23929	74%	55%	62%	65%
1 - 5hr	134661	235205	105501	95402	88%	83%	85%	85%
5 - 10hr	64488	98207	47384	45046	95%	95%	95%	94%
10 - 18hr	39565	30477	24528	25183	100%	99%	100%	99%
18+hr	3482	6202	1406	4449	100%	100%	100%	100%

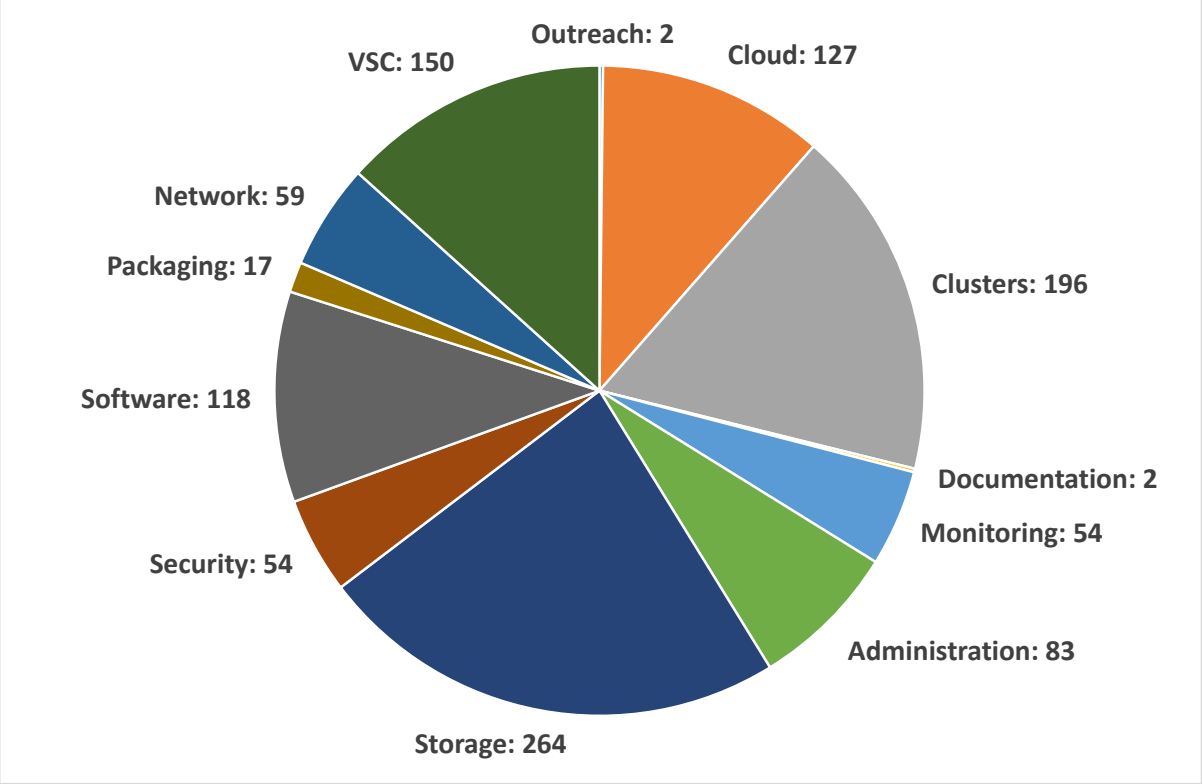
On average, considered over all quarters of 2020:

- 64% of all jobs submitted start within 1 hour
- 95% of all jobs submitted start within 10 hours at most

3 DEVELOPMENT AND MAINTENANCE

To maintain compute infrastructure capabilities, HPC-UGent continuously performs maintenance works. On hardware, such as installing new components or replacing broken ones, and on software, ranging from the operating system to higher-level software that enables services directly available to the end user. To report, plan and follow-up hardware and software bugs, issues or developments, these are managed with the aid of a tracker, JIRA. Larger developments, grouping many individual or interrelated issues are grouped as 'Epics'. These effectively track the main development and maintenance projects that HPC-UGent executes. Github services are used to facilitate software development and Jenkins services for continuous integration. The number of 'pull requests' (PRs) by HPC-UGent staff to the various repositories is indicative of the effort spent over the last year.

3.1 Number of JIRA issues resolved in 2020, grouped per main component



3.2 Projects finalized in 2020

Issue #	Title	Summary
1	Tier-1c tendering procedure	Wrap-up of the purchasing procedure for the next VSC Tier-1 supercomputer
2	Port to python3	Full HPC-UGent internal codebase migration to python3.

3	Migrate VMs to new VMware	Migration of all HPC-UGent virtual machines that are used for admin purposes to new UGent infrastructure.
4	Setup GPU cluster (joltik)	Finalization of pilot phase and start of production.
5	RedHat OpenStack Platform deployment	Milestone in the development of the Tier1-Cloud infrastructure for VSC, setting up a complete new Cloud platform using Redhat OpenStack and Director.
6	Improve automation infrastructure	Several improvements to HPC-UGent admin environment, furthering automation of e.g. updates, tests, …
7	EUdat and RDM IRODS integration	Research and tests into EUdat and IRODS as RDM-ready tools for VSC data component. Trial project and setup finalized, all information transferred to KULeuven who will manage the new Tier-1 Data component.
8	Set up new Globus endpoint for UGent shares	Enable mounting of UGent shares on HPC infrastructure
9	Trial of Burst Buffer on Slurm	Research and tests into Slurm feature for shared high-speed storage resources - abandoned due to time constraints.
10	VSC security data	Engagement and collaboration of HPC-UGent in ISO-27001 project for VSC
11	Tier2 site update 2020Q1	General maintenance of all Tier2 clusters and servers.
12	Arcanine scratch pilot	Finalization of pilot project for the new very fast storage space and start of production.
13	SSH key evaluation and improvements	Changes to accountpage and security response related to the VSC-wide SSH key update and cleanup of 27 May 2020.
14	Migrate HOME/DATA/SCRATCH to GPFS5 on new storage	All tasks related to migration to the new, larger storage system.
15	Cloud Cluster Phase III	Milestone in the development of the Tier1-Cloud infrastructure for VSC, relinquishing OpenNebula.

3.3 Github PRs in 2020 by HPC-UGent staff, per repository

github.ugent.be/hpcugent	
Repository	#PRs
quattor	674
jobcli	75
vsc-accountpage	44
vsc-install	37
vsc-jobs	21
vsc_user_docs	20
vsc-config	19
vsc-base	14
vsc-utils	11
vsc-mympirun	8
vsc-accountpage-clients	7
eb_inuits	7
vsc-administration	6
Lmod-UGent	5
vsc-filesystems	4
slurm	4
slurm-prepilogue	4
vsc-filesystems-quota	4
vsc-ldap	2
gpfsbeat	2
nhc	2
pyslurm	2
vsc-ldap-extension	2
vsc-modules	2
logstash-patterns	1
vsc-zk	1
csub	1
mympingpong	1
django-wayf	1
slurm-spank-talamini	1
slurm-torque-wrappers	1
openstack-dashboard-vsc	1
xpmem	1

github.com/easybuilders	
Repository	#PRs
easybuild-easyconfigs	653
easybuild-easyblocks	114
easybuild-framework	110

github.com/hpcugent	
Repository	#PRs
quattor	286
legacy-rpms	84
clusterbuildrpm-server	23
icinga-checks	21
oncall-service	16
documents	14
vsc-cluster-modules	12
vsc-quattor	11
software-stack	9
vsc-monitoringdb	8
quattor-host-mngmt-tools	8
pbsmon-web	7
vsc-freeipa	7
vsc-ood	7
ssh-attack	7
vsc-cloud	6
vsc-burstbuffer	6
vsc-project	4
android-systems-locator	3
service_sanity	3
vsc-firewall-client	3
generatemotd	2
sync-ldap-collector	2
easyblocks_ugent	2
otrs_stats	2
vsc-profiles	2
flexlm	2
vsc-host-tools	2
slurm-node-resume	2
vsc-reporting	2
pbsmon2php-client	1
vsc-postgres	1
reposnap	1
vsc-backup	1
vsc-testing	1
vsc-rpms	1
rundeck	1
vsc-data	1
vsc-gridftp	1

4 TRAINING AND SUPPORT

4.1 Training overview and evaluations

As our rector has eloquently put it “Corona sucks”. The pandemic has had a considerable impact on trainings that were scheduled for 2020. For some, a remote alternative was developed, albeit without the same 1-to-1 hands-on experience as was considerably appreciated by participants in the previous years.

4.1.1 Introduction to HPC @ Ghent University

Trainer: Kenneth Hoste (HPC-UGent)

Date	#participants
20/01/2020	21

Satisfaction scores[1]	
Training content	100%
Lecturer	100%
Hands-on	100%

Selected suggestions for improvement and remarks:

- I would like to see a little bit more exercises in an introduction course
- Very nice work on the available manual, this helps a lot to get you started!

[1] These satisfaction scores indicate positive response of evaluation respondents on questions “The presentation gave me all the information I wanted” (Yes or Mostly) or “This course helped build up my knowledge of the topic at hand” (Strongly Agree or Agree); “The lecturer presented well” (Strongly Agree or Agree) or “The lecturer(s) had the expertise required to meaningfully elaborate upon the subject of the course” (Strongly Agree or Agree); and "The hands-on session was in accordance with my expectations” (Strongly agree or Agree).

4.1.2 Introduction to HPC @ Ghent University – remote

Trainer: Kenneth Hoste (HPC-UGent)

Date	#participants
18/03/2020	12
15/09/2020	15

Satisfaction scores[1]	
Training content	100%
Lecturer	100%

Most attendants did not find the digital format of this training to have heavily affected the learning experience. Points raised:

- Barrier to ask questions is a little bit higher
- More effective than a physical meeting

Selected suggestions for improvement and remarks:

- Short but clear introduction to the HPC. Afterwards you can start with submitting jobs.
- Thorough explanations, practical session
- Well prepared, effective, reusable material
- Maybe a bit more exercises could be useful, or an extra course about Linux command line

4.1.3 Specialist Workshops in Scientific Computing – Introduction to multithreading and OpenMP

Trainer: Reinhold Bader (Leibniz Supercomputing Centre LRZ, Germany)

This course was originally scheduled for 28-29 May 2020, but was cancelled as a result of the COVID-19 pandemic.

4.1.4 Specialist Workshops in Scientific Computing – Introduction to MPI

Trainer: Jan Fostier (INTEC, IDLab, UGent & imec)

This course was originally scheduled for 3 June 2020, but was cancelled as a result of the COVID-19 pandemic.

4.1.5 Suggestions from participants for new training sessions

- Multi-job submission, automating job submissions
- Using modules
- Mapping and blasting
- Linux command line
- More information on interactive jobs
- Python modules
- Working in R
- Integration with vsc storage
- Containers (docker)

4.2 Lectures and community meetings

4.2.1 HPC-UGent user meeting

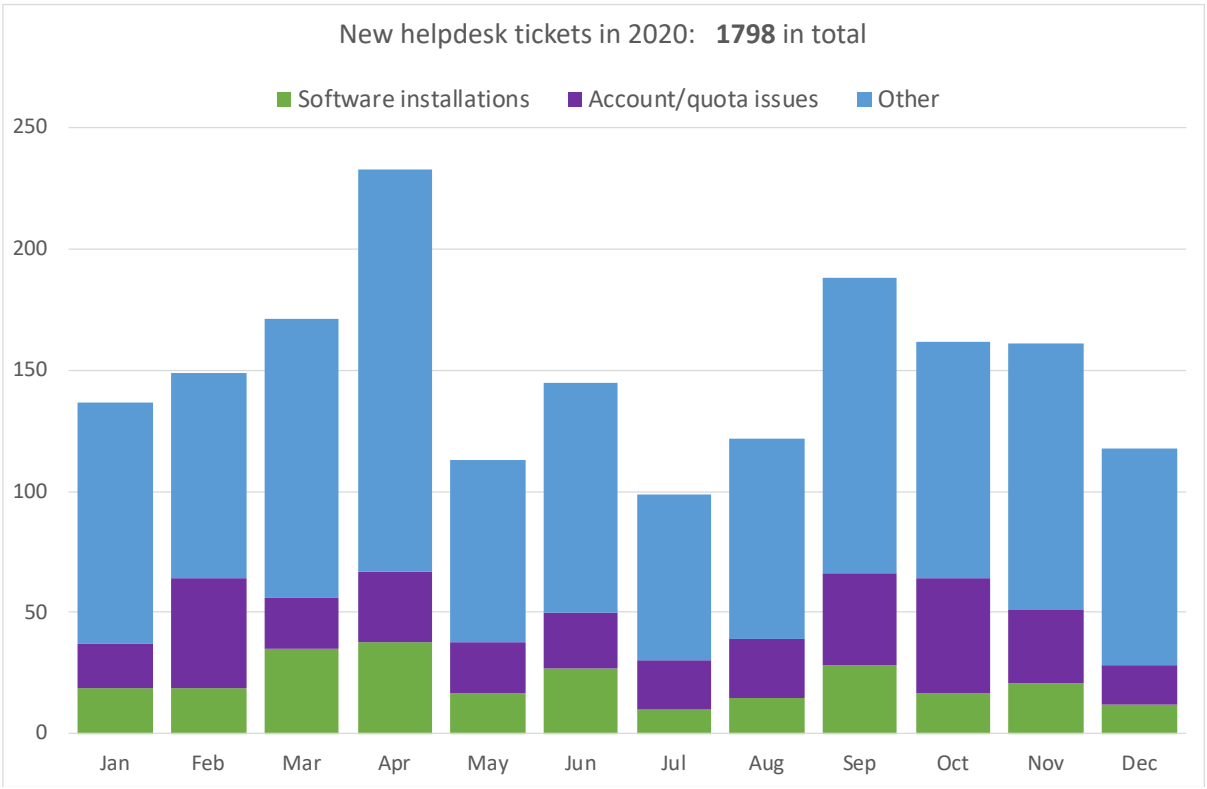
Purpose of this meeting is to inform all users about future plans for the HPC-UGent infrastructure, and to bring together users and the HPC-UGent team.
Featured presentations

- Ewald Pauwels, HPC-UGent
  "Overview of HPC-UGent usage, status of the VSC, future plans, review of user poll results"
- Jonathan Leliaert, DyNaMat
  "GPU-accelerated micromagnetic simulations using mumax3"
- Sebastiaan Theuns, Laboratory of Virology
  "GPU acceleration of revolutionising diagnostics of infectious diseases in veterinary medicine"

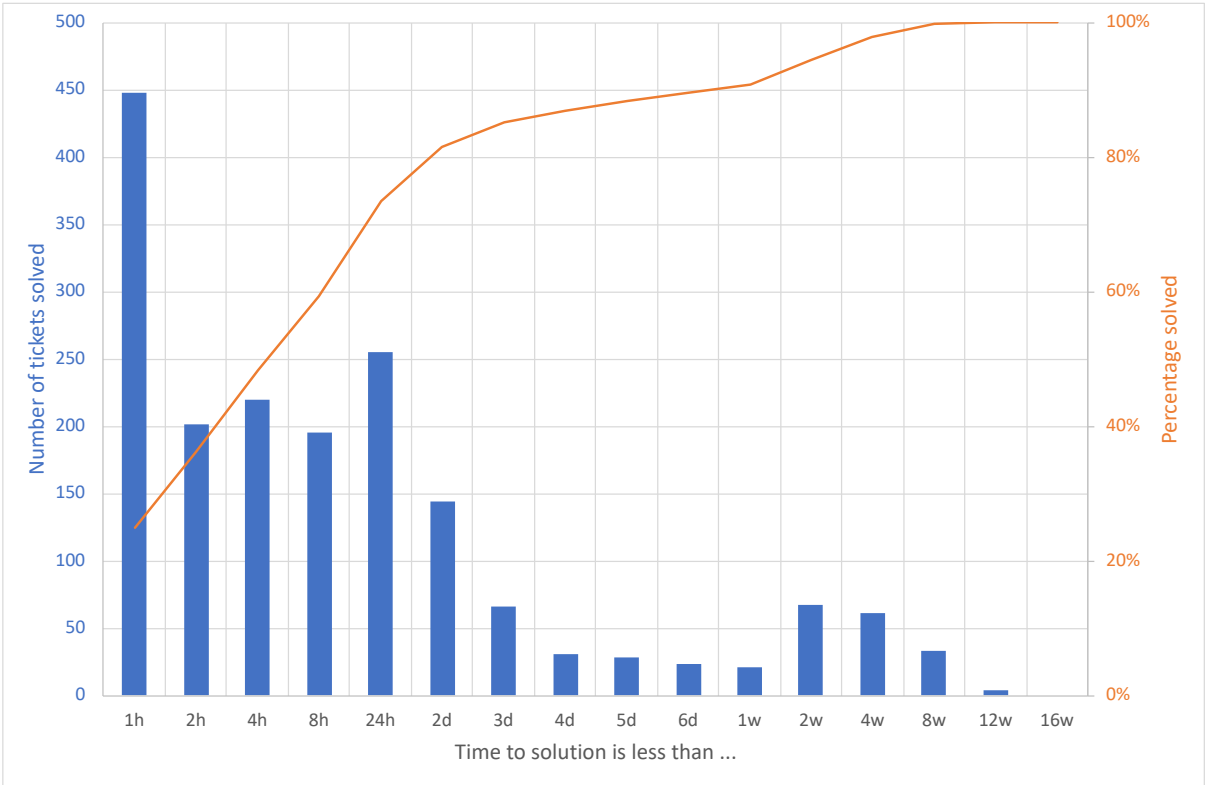
Date	#participants
17/02/2020	30

4.3 Helpdesk

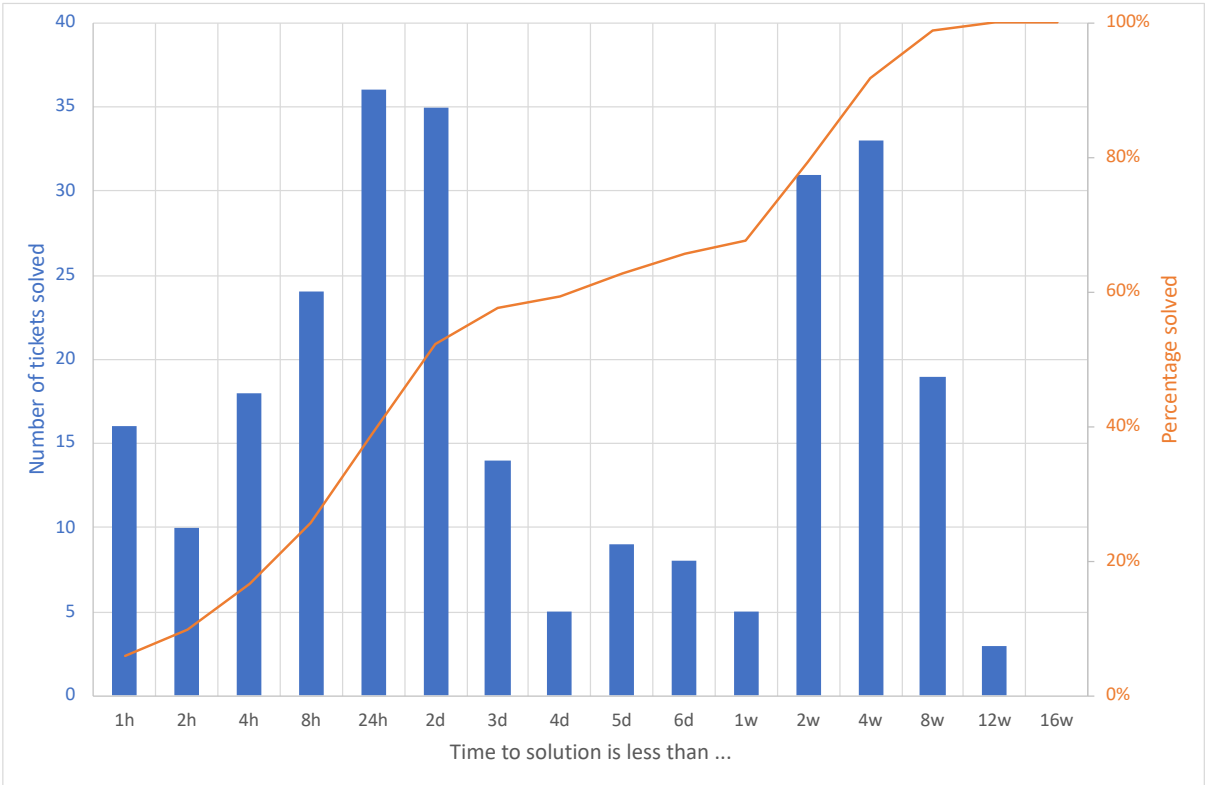
4.3.1 Number of new tickets per month



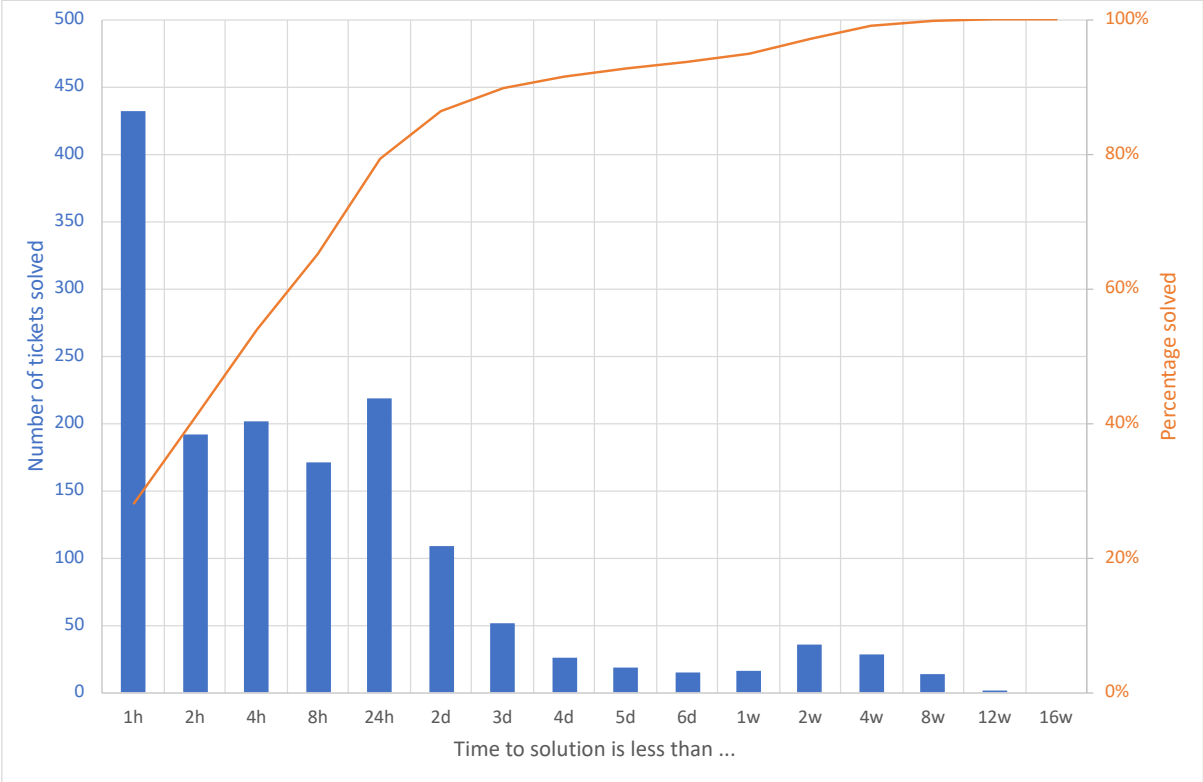
4.3.2 Average time to resolution – overall



4.3.3 Average time to resolution – software installation requests



4.3.4 Average time to resolution – NOT software installation requests



4.3.5 Speeding up software installation requests – consultancy

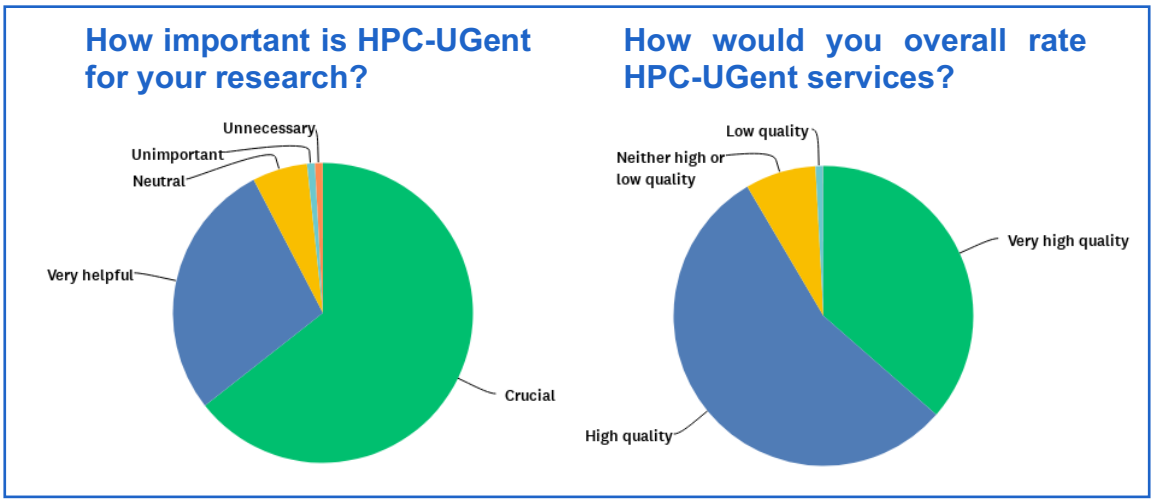
To meet the growing request for central software installations, HPC-UGent regularly engages consultancy from Inuits to help out, in addition to the regular HPC-UGent helpdesk staff. This has enabled us to keep up, as best as possible, the pace of deploying and installing specific scientific software centrally.

On average, every software installation takes about 7 hours of time. This is one order of magnitude larger than the average effort for 'regular' (non-software-install) helpdesk tickets.

4.4 User evaluation

In order to improve HPC-UGent services, all current users were asked to complete a user survey. This survey was anonymous, short (average completion time 3 minutes) and in all 118 users responded. A selection of questions and responses is shown.

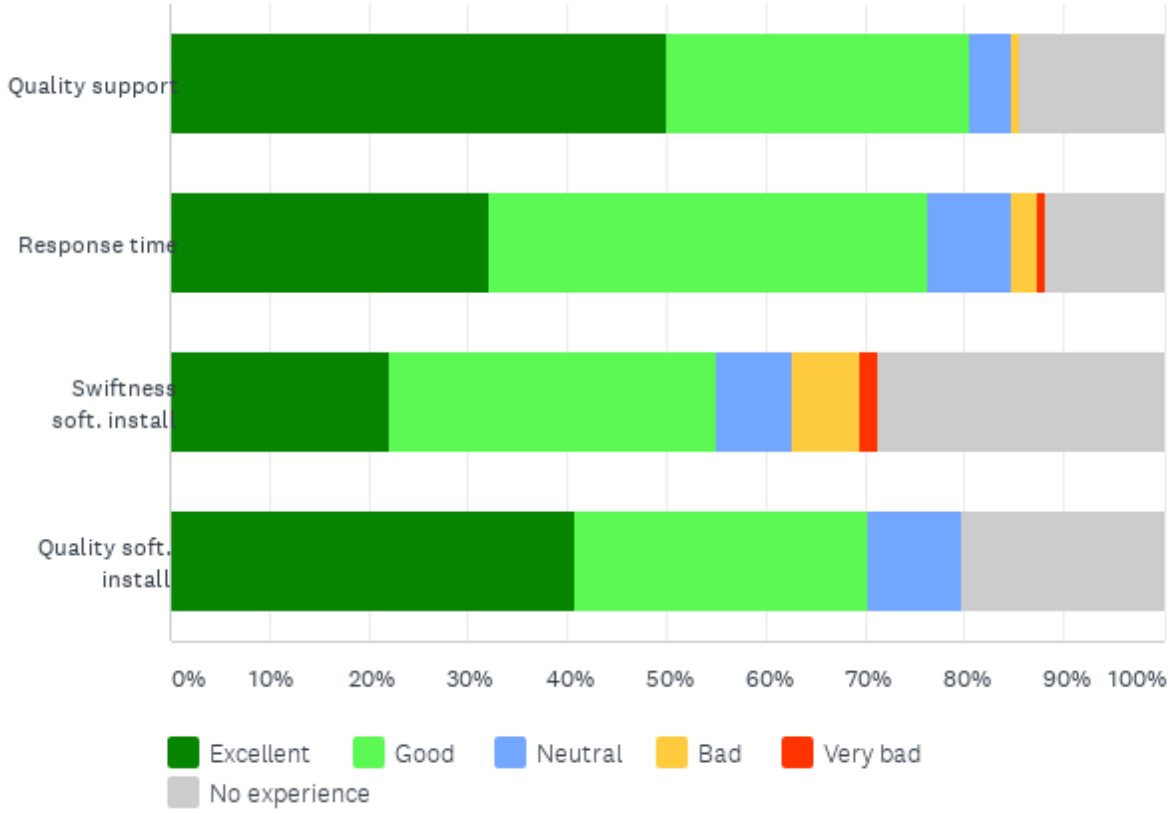
How often do you use the HPC-UGent scientific computing infrastructure?		
	#responses	*%*
On a daily basis	32	27%
(More than) once per week	46	39%
(More than) once per month	27	23%
1-2 times per year or less	13	11%



How important is HPC-UGent for your research?		
	#responses	*%*
Crucial - I can't do my research without	76	64%
Very helpful - It allows me to do my research at a faster pace and at a higher level	33	28%
Neutral	7	6%
Unimportant - I can just as well do my research in another way	1	1%
Unnecessary - I don't need it at all to do my research	1	1%

How would you overall rate the services that HPC-UGent provides?		
	#responses	*%*
Very high quality	43	36%
High quality	65	55%
Neither high or low quality	9	8%
Low quality	1	1%
Very low quality	0	0%

Please rate the following aspects related to HPC-UGent user support



Would your research benefit from specific IT hardware or services that HPC-UGent currently does not provide?		
	#responses	*%*
No – The current compute platform suffices	71	60%
I don't know	39	33%
Yes – I need specific hardware/services	8	7%

- More GPUs
- More high-memory clusters

- Dedicated development environment (queue/server) for GPU porting
- Lower queue times

- Conda support

How could we further improve HPC-UGent services?
Documentation
• Expand documentation with links to external sources • Step-by-step documentation on e.g. pip and conda • Workflows including git/github • Workflows including jupyter notebooks
User experience
• Shorter queue times (for companies) • Way to estimate queue time   ○ See number of jobs/corehours equested per cluster by all users • Project applications for VOs to get compute time (to reduce load) • Automatic distribution of jobs across clusters in terms of occupancy and efficiency • Longer wall-clock time • Better estimate of the impact of maintenance on e.g. license servers
Infrastructure
• More nodes/clusters • More GPUs • Higher file transfer speeds • Fast parallel and accessible communication to SCRATCH on victini • More storage in HOME directory
Training
• Online courses/recordings on how to use HPC efficiently • Specific workshops, targeted to specific research groups • Introduction to parallel programming • How-to lecture on compiling your own complex software model • Using git • Jupyter notebooks on HPC web portal • More mails with suggested courses (e.g. PRACE)
Software
• Easy local module installation • Permit installation of published R packages without separate manual requests • Update all the software on a regular basis
User support
• Speed up software installation • Speed up response time in general • (Better) support for first time users and non-computational science researchers • More staff
Consultancy
• Advanced programing support (e.g. testing and improving code) • Support for compiling complex software models

User encouragements

“Keep up the good work!”

“We <3 HPC-UGent!”

“We are super happy with the HPC-UGent services! Many thanks to a great team for all the services and support!”

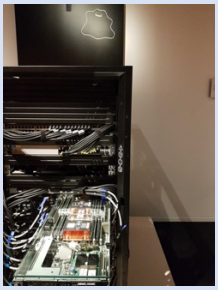
“I really appreciate the great support. It has been extremely helpful to me and my research.”

5 OUTREACH

5.1 Within Ghent University

Date	Event
3/04/2020 8/05/2020 19/06/2020 10/12/2020	The HPC-UGent team regularly meets with the UGent data stewards. The scope here is broader than the typical HPC-UGent community of users, comprising all UGent researchers and all infrastructure and services of the ICT department (DICT).
7/02/2020	"*Overview of DICT storage platforms*" Short Sessions on Hot Topics (SSHT) – Faculty of Economics and Business Administration
9/11/2020	"*DICT storage platforms – Theory*" Short Sessions on Hot Topics (SSHT) – Faculty of Economics and Business Administration
26/11/2020	"Roadmap: compute resources for education" FWET computer committee, Faculty of Sciences
8/12/2020	"*DICT storage platforms – Applied*" Short Sessions on Hot Topics (SSHT) – Faculty of Economics and Business Administration
10/12/2020	"Roadmap: compute resources for education" UGent ICT user committee

5.2 To policy makers, industry and general public

Date	Event
11/02/2020	Derelict cluster-on-wheels ditto has been decommissioned and is now part of the exposition in the new Ghent University Museum.  https://www.gum.gent
14/02/2020	HPC introduction and UGent datacenter tour for Romanian consul to Belgium

5.3 Within international HPC community

Date	Event
29-31/01/2020	5th EasyBuild User Meeting (EUM'20) Barcelona, Spain https://github.com/easybuilders/easybuild/wiki/5th-EasyBuild-User-Meeting Co-organization of meeting Keynote presentation "*EasyBuild State of the Union*" Lecture: "*10 things you did not know yet about EasyBuild*"
1-2/02/2020	FOSDEM'20 Brussels, Belgium Co-organization and support of "*HPC, Big Data and Data Science* devroom" https://archive.fosdem.org/2020/schedule/track/hpc_big_data_and_data_science/
5/03/2020	Scientific Software Repository meeting TU Delft, the Netherlands https://www.eessi-hpc.org
7/05/2020 2/07/2020 6/08/2020 3/09/2020 1/10/2020 5/11/2020 3/12/2020	Monthly *"European Environment for Scientific Software Installations (EESSI)"* update meetings Remote https://www.eessi-hpc.org
11/06/2020	Opennebula Cloud e-tech day, hosted by CSUC (Consorci de Serveis Universitaris de Catalunya) and Cloud Admins Barcelona. Keynote presentation "*HPC Clouds at UGent*" https://www.meetup.com/Cloud-Admins-Barcelona/events/270506256/
16-18/06/2020	6th Annual High Performance Container Workshop Remote http://qnib.org/2020/06/17/isc2020-hpcw/ Presentation "*Build Tools (Spack/Easybuild/...)*"
19-20/06/2020	HPC Knowledge Portal meeting (HPCKP'20) Remote https://hpckp.org/past-edition/hpckp-20/ Presentation "*European Environment for Scientific Software Installations (EESSI)*"
23/06/2020	EasyBuild tutorial Remote https://github.com/easybuilders/easybuild/wiki/EasyBuild-tutorial

	Co-organization and co-presentation This tutorial was accepted for ISC'20 but postponed to ISC'21. As such, it was organized as an online tutorial separate from ISC'20.
24/06/2020 8/07/2020 5/08/2020	Organization of EasyBuild Tech Talk "*The ABCs of Open MPI*" Remote Speakers: Jeff Squyres (Cisco, Open MPI) and Ralph Castain (Intel, Open MPI, PMIx) https://github.com/easybuilders/easybuild/wiki/EasyBuild-Tech-Talks-I%3A-Open-MPI
30/06/2020	Confcall with Vienna supercomputing center (Austria) on Slurm scheduling configuration. Remote
24/09/2020	Extreme-scale Scientific Software Stack (E4S) Forum Remote https://oaciss.uoregon.edu/E4S-Forum20/agenda.html
30/09/2020	Organization of EasyBuild Tech Talk: "*Yes! You Can Run Your Software on Arm.*" Remote Speaker: Chris Edsall (Univ. of Bristol) https://github.com/easybuilders/easybuild/wiki/EasyBuild-Tech-Talks-II:-Arm
24/11/2020	Interview for imakefoss.org Remote https://imakefoss.org/talks/ 1-week curator for @imakefoss Twitter account Contribution to imakefoss YouTube channel
25/11/2020	Remote SORSE event (International Series of Online Research Software Events) Software demo on "*European Environment for Scientific Software Installations (EESSI)*" https://sorse.github.io/programme/software-demos/event-028/
4/12/2020	SESAMENET partner meeting Remote
12/10/2020	EuroCC Introductory session Belgium Remote Presentation on "*WP30.5 – Mapping HPC/BigData/AI Competences*"
15/10/2020	EuroCC-CASTIEL Initial competence mapping meeting Remote
26/10/2020	EuroCC internal meeting for Belgian national competence center Remote

19/11/2020	EuroCC-CASTIEL update on competence mapping, industry engagement and techtransfer Remote
10/12/2020	EuroCC-CASTIEL meeting of competence mapping champions Remote
15/12/2020	Presentation on European Environment for Scientific Software Installations (EESSI) Remote HPC Champions workshop @ UK https://www.archer2.ac.uk/training/courses/201215-champions/

Highlight:

EuroCC Belgium

Within the EuroCC project under the European Union's Horizon 2020 program (H2020), participating countries are tasked with establishing a single National Competence Centre (NCC) in the area of high-performance computing (HPC) and high-performance data analytics (HPDA). Each NCC will coordinate activities in all HPC-related fields at the national level and serve as a contact point for customers from industry, science, (future) HPC experts, and the general public alike.

For Belgium, the NCC is a consortium of 12 entities:

- CECI: the 5 French speaking Universities (http://www.ceci-hpc.be/)
- VSC: the 5 Flemish speaking Universities (https://vscentrum.be/)
  also Tier-1 operator for Flanders
- Cenaero: Research center (http://www.cenaero.be/)
  Tier-1 operator for Wallonia
- Innoviris: Research & Innovation hub (https://innoviris.brussels/)

The objectives of EuroCC Belgium are:

- Support usage of high-performance computing and data analytics
- Set up a full training offer
- Set up a framework to ease access to HPC/HPDA for industry
- Identify available competences at national and European level
- Facilitate access to systems, scientific and technical expertise and knowledge pools
- Set up communication and dissemination actions

Role of HPC-UGent

HPC-UGent represents Ghent University in this consortium and is joint task-leader of WP30.5, set on mapping HPC/Big Data/AI technical competences within Belgium. UGent is further involved as a contributor in WP30.3 focusing on 'Technology Transfer/Business Development' and WP30.4 'Collaboration with Industry'.

Link: https://www.enccb.be

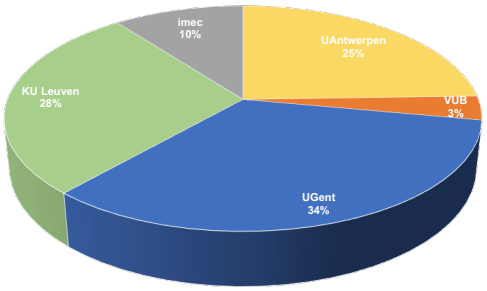
6 BUDGET

Budget line	UGent budget statement	Income
Scientific computing (Tier2)	1.320.000	
Tier1 Flemish supercomputer and VSC cloud	6.525.000	
Central compute infrastructure (wages)	952.000	
Impulse funding Tier1 staff	63.000	
FWO recurrent subsidy for personnel (Tier1+2)		665.000
FWO recurrent subsidy for investment and operations		1.316.977
FWO project subsidy Tier1 compute		6.000.000
FWO project subsidy Tier1 cloud		475.000
FWO project subsidy Tier1 compute and cloud - power		50.000
FWO project subsidy Tier1 operational management team 2020 – staff		63.333,33
FWO project subsidy Tier1 cloud 2020 – staff		190.000
	euro	euro

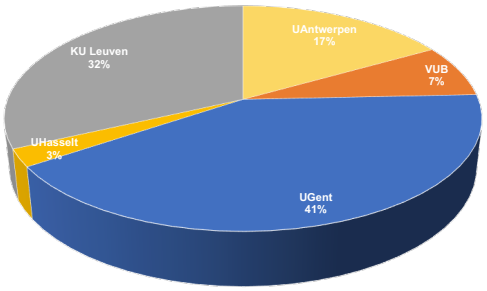
7 TIER-1 USAGE

Every year, three calls are open to apply for access to the Tier-1 system of the Flemish Supercomputer Center (VSC). HPC-UGent stimulates and assists researchers as much as possible to apply for access. The tables and graphs below give insight in the number of proposals (#proj) and the awarded compute time (#nd, expressed in Tier1 nodedays), listed per institute.

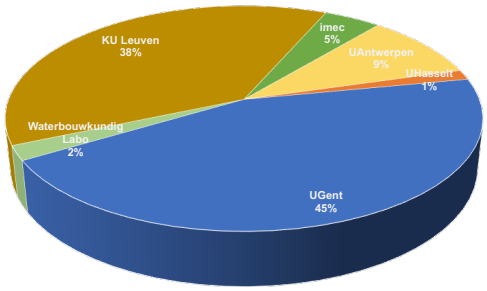
Call 1 3/02/2020	Requested		Granted	
Institute	#proj	#nd	#proj	#nd
UAntwerpen	3	22346	3	22346
VUB	1	3120	1	3120
UGent	**9**	**38448**	**7**	**30648**
KU Leuven	6	53580	4	25604
imec	2	9453	2	9453



Call 2 15/06/2020	Requested		Granted	
Institute	#proj	#nd	#proj	#nd
UAntwerpen	5	20156	5	20116
VUB	4	8978	4	8978
UGent	**14**	**58311**	**12**	**49424**
UHasselt	1	3000	1	3000
KU Leuven	10	59954	8	38213



Call 3 5/10/2020	Requested		Granted	
Institute	#proj	#nd	#proj	#nd
UAntwerpen	2	9912	2	9912
UGent	**12**	**46847**	**12**	**46847**
UHasselt	1	1500	1	1500
KU Leuven	11	56625	10	40025
Waterbouwkundig Labo	1	2209	1	2209
imec	1	4755	1	4766
von Karman institute	1	5000	0	0



Percentage of awarded compute time, listed per institute per call

8 USER IN THE SPOTLIGHT

At the dawn of the COVID-19 crisis, a number of eager researchers at the Department of Data Analysis and Mathematical Modelling at UGent's Bioscience Engineering faculty decided to mobilise their expertise in the battle against SARS-CoV-2. At present, we have joined forces with model developers and analysts at UHasselt, UAntwerpen, UNamur, ULB and VUB, bundled in the RESTORE consortium (https://covid-en-wetenschap.github.io/restore.html, see Figure 1). The ensemble of model outputs provided by this collection of scientists has proven to be indispensable for the Belgian Ministry of Health in its task to quickly and efficiently decide on new social policies. HPC-UGent plays a crucial role in the smooth execution of our mission.

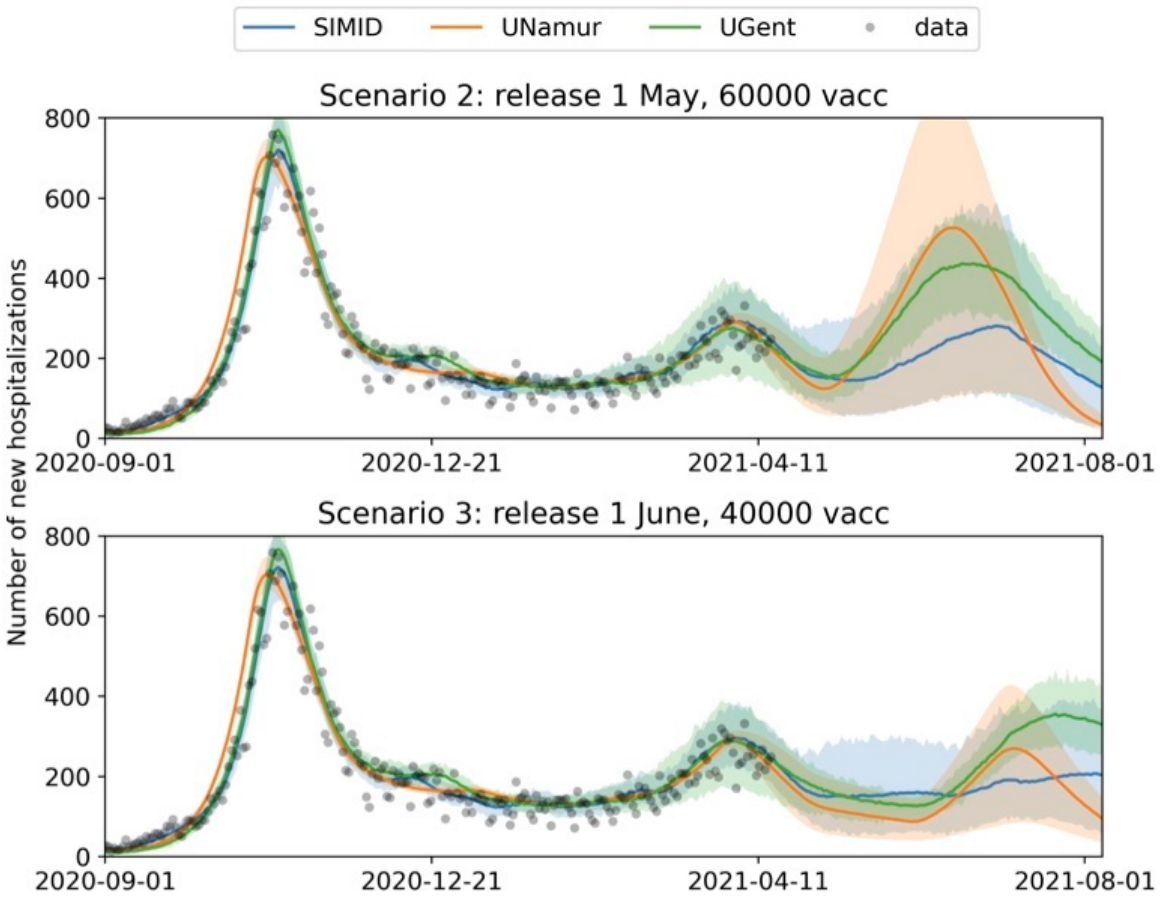


Figure 1. Ensemble of predictions for new hospitalisations based on data (black circles) and various models (blue: UHasselt; orange: UNamur; green: UGent). Different scenarios of the future (timing of relaxations, tempo of vaccinations) cause different predictions. Note that this prediction is over two months old and new calibrations would paint a less worrisome picture.

The framework developed at UGent is based on a tailor-made extension of epidemiological metapopulation models ("SIR models"), embodied by a large number of coupled ordinary differential equations (ODEs). Due to the virus's propensity to manifest differently in distinct ages, a distinction is made between 9 age groups. In addition, the geographical spread of the virus is of interest – we therefore consider a 43-fold geographical stratification. This effectively introduces a 9-times-43 multiplication of the number of coupled ODEs that is to be solved numerically. The solution of this ODE system with particular initial conditions and empirical parameter values allows for the long-term prediction of e.g. the number of hospitalised patients between the ages of 50 and 60 in the arrondissement of Antwerp. Whilst this is already computationally intensive, it is certainly not unmanageable on a local machine. The real need for remote parallel computing on HPC-UGent comes into play when considering the *calibration process* of the developed model, because this requires solving these ODEs for a large number of iterations and comparing every iteration with the data.

The model is supported by a large number of parameters. These include for example the average length of stay in the hospital, or the daily average mobility between two regions. The values of most of these parameters are known rather precisely, informed by large datasets from Sciensano, Proximus, mathematical translations of political policies, and previous COVID-19-related publications. *However*, some parameter values – eight in total – we cannot know beforehand, such as the intrinsic infectivity of SARS-CoV-2. In order to determine their values, and therefore also determine an appropriate model outcome, the parameter space must be explored. For every combination of these parameters, the model predicts a different outcome, which is compared to historical hospitalisation time series until a sufficiently close match is achieved.

Even on a supercomputer, an eight-dimensional parameter space should not be explored exhaustively. In order to efficiently identify the best-fit extremum, the space is sampled with a Markov chain Monte Carlo algorithm, with a large number of parallel chains. We use the ***emcee*** package for MCMC sampling, and the ***multiprocessing*** package for parallel computing, both in Python. For a typical run, we use 36 cores on the skitty cluster for 108 chains, which allows us to achieve convergence in a few tens of thousands iterations within approximately one day. This procedure provides best-fit parameter values, as well as a large sample library used to express model uncertainties.

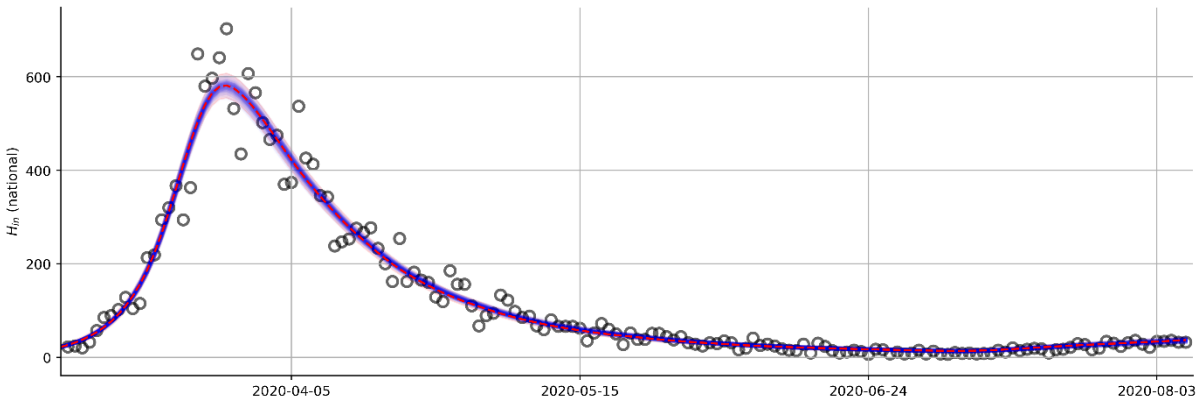


Figure 2. The national number of new hospitalisations in the first COVID-19 wave. This result is calculated in a spatially explicit model, which requires supercomputer capacity. The blue zone represents simulations based on samples made in a calibration process using an Markov chain Monte Carlo algorithm on HPC-UGent. The red dashed line is the median value. The black circles represent the raw daily data from Sciensano. The resulting parameter values allow for the extrapolation of the model into the (near) future, aiding medical personnel and politician in their decision-making process.

The model outputs and the associated uncertainty intervals have played an important role in providing objective information in the political decision-making process for matters related to COVID-19, in an environment where a subjective response inspired by blinding emotion was the dangerous alternative. The physical infrastructure as well as the quick and professional support provided by (the people at) HPC-UGent have been of vital importance in this endeavour.

Tijs Alleman, Michiel Rollier, Jan Baetens, Jenna Vergeynst, Gisele Miranda, Ingmar Nopens, Joris Meys

Contact: jan.baetens@ugent.be
Research unit KERMIT (Knowledge-based Systems)
Research unit BIOMATH
Ghent University