A strategic vision of the whole IT operations on Kumpula campus

Summary

This document describes the present state of IT operations on Kumpula campus, as well as improvements needed in the near future.

The shared IT costs at the Faculty of Science at the University of Helsinki are some 2 million euros annually. The largest single expense consists of workstation acquisitions, around 1 million euros. Some of the workstations and other equipment are also purchased with department or project funding. The campus is home to a significant IT infrastructure that has been obtained through separate funding (Ukko2, Kale, Vorna etc). Further, a large part of the computational infrastructure belonging to CSC (Puhti, Mahti, etc) is employed at the campus. The services on campus are produced by local branches of the IT Centre (on-site support, IT for Science), other units of the IT Centre such as the data communication group, CSC, and commercial ISPs.

The current computing and storage infrastructure will become outdated in the near future, but the need for our own infrastructure will remain. The costs for workstations are considerable, but it is vital for on-campus operators to have equipment that fulfils their needs. In future, we may be able to cut back costs for workstations by deploying so-called virtual desktops (VDI) more widely, both when it comes to freshman laptop workstations, computer labs, as well as high-performance computers.

Data- and computing-intensive research has spread into the different disciplines and research groups on campus. Within scientific computing, a critical turning point is under way on a national and European level, as the available computing capacity is increasing exponentially with the emergence of new supercomputers (Mahti, Puhti and LUMI). Further, the event of GPU technologies in scientific computing is creating new opportunities and challenges, such as the need to educate users. Changes accentuate the need for various services on different levels and for resources. The scientific approaches are very varied according to research problems, which leads to the need for diverse tools and infrastructures. Data-intensive research also highlights the significance of data management.

The strategic vision for developing the IT operations in Kumpula is that *Kumpula campus will be a global trailblazer for computational science*.

User-oriented solutions are the main principle for the development of all services and their accessibility – it must be easy to use and access services. The other strategic outlines are categorised under the following headings: Capabilities, cooperation and information flow, work environments, infrastructures and support.

Document information

Version	1.0
Authors	Ville Tenhunen et.al.
Date of writing	22.4.2020
Approver	The IT steering group on Kumpula campus
Approval date	25.3.2020
Organisation	UH/KITO
Publicity status	public

Version history

Version	Creator	Date	Changes
0.1	ville.tenhunen@helsinki.fi	12.1.2020	init()
0.2	ville. tenhunen@helsinki. fi	24.1.2020	Changes made on the basis of committee meetings are marked in *italics*.
0.3	ville. tenhunen@helsinki. fi	3.2.2020	Version corrected according to comments from steering group (above-mentioned italics etc removed)
0.31	hannu. toivonen@helsinki. fi	5.2.2020	Details clarified

0.4	ville. tenhunen@helsinki. fi, sami. maisala@helsinki. fi, natalia. jiganova@helsinki. fi	14.2.2020	Clarifications and additions to several chapters on the basis of committee discussion.
0.41	hannu. toivonen@helsinki. fi	19.2.2020	Clarifications and styling
0.42	hanna. vehkamaki@helsinki. fi	25.2.2020	Comments in red, a few additions on the CSC interface
0.43	hannu. toivonen@helsinki. fi	28.2.2020	Additions to facility information
0.44	natalia. jiganova@helsinki. fi	12.3.2020	Additions on the AV equipment in teaching and meeting facilities
0.45	sami. maisala@helsinki. fi	17.3.2020	Small clarifications and styling
0.46	sami. maisala@helsinki. fi	14.4.2020	Corrections based on steering committee comments.
1.0	hannu. toivonen@helsinki. fi	22.4.2020	final version

Contents

1 Introduction

This document presents a strategic vision of the all-over IT operations on Kumpula campus. On 11 December, 2018, the Dean of the Faculty of Science asked the IT steering committee on the Kumpula campus to create this vision.

The main body of the document describes the current status, changes in the operational environment, and the vision and outlines for IT operations for the next few years.

In the description of the current status, the main focus is on IT services, hardware environment, maintenance responsibilities, management models, and IT costs on the campus.

In the changes to the operational environment, we have attempted to take into account the main factors that affect the outlining: general adjustment factors like the university's strategy, changes in the operational environment of the campus, international development projects, and available resources.

The strategic vision for developing the IT operations in Kumpula is that *Kumpula campus will be a global trailblazer for computational science*.

The outlines are based in some central principles of IT operations. In the vision, we have assessed the IT-related capabilities that we can invest in to promote the advancement of core operations on campus. The other policy drafts are ordered under the headings cooperation and information flow, work environments, infrastructures, and support.

2 Current Status

2.1 IT services on campus

Services offered by the IT Centre

There are two IT Centre teams working on-campus, i.e. on-site support and IT for Science.

The duties of *on-site support* are:

- services to students (freshman laptops, computer stations, VDI kiosks, workstations in the working spaces, etc)
- teaching-related maintenance (workstations in classrooms, lecture halls, exam aquarium)
- maintenance of staff workstations
- maintenance of workstations in joint-use facilities
- maintenance of IT environments in labs (lab networks, workstations for measuring instruments)
- IT acquisitions
- recycling of IT equipment

A total of 6 persons are employed in the on-site team.

The duties of *IT for Science* are:

- maintenance and development of scientific computation infrastructure
- FGCI coordination (Finnish Grid and Cloud Infrastructure)
- development and support of the university's own Linux distribution
- helpdesk and consultation for High Performance Computing (HPC) environments, storage, and software
- development and maintenance of version control system
- maintenance of the infrastructures and applications for geoinformatics
- participation in the maintenance and development of the VDI infrastructure
- maintaining the computer lab infrastructure in Kumpula
- participation in the university's data support
- support to acquisitions when required
- support for education in scientific computation as resources allow

From the beginning of year 2020, there were 9 employees in the IT for Science team; 6 in Kumpula, 2 in Viikki, and 1 in Meilahti.

The on-site support or the IT for Science team in Kumpula do not offer much support in the use and maintenance of applications.

Further, the services of the following IT Centre teams are used on campus:

- the data communications group (on-campus network connections)
- capacity services (centralised storage and server room services (Viikki server room services))
- the Linux team (maintenance of Linux servers)
- the Windows team (Windows servers, AD)
- user administration (centralised access administration)
- IT communications (Unitube)

A rough estimate of the extent of these support functions would be several person years annually.

Other services

Services produced by other agents are also utilised on campus:

CSC

- the customer service of CSC, such as support for scientific computation
- local CSC support: the units on campus can request support from CSC for local operations in the form of support staff, for example

- CSC is a partner in some of the infrastructure projects of the research groups at the faculty (INAR)

Unigrafia

- printing services
- video streaming and recording services

The Facility Services at the University of Helsinki (TILA/HY247)

- AV and presentation technology in e.g. lecture halls and meeting rooms (Facilities administrates the funds, HY247 produces the services)

Others

- the services and infrastructures from commercial cloud service providers have their own user support
- the workstations and applications in the laboratories have varied support services depending on to which and what kind of infrastructure or equipment the workstations or applications are connected. Typically, such support is produced by the provider of the equipment or application.

2.2 Hardware and its lifespan

Computation infrastructure managed by the faculty: Ukko2, Kale, Vorna

The university's computation infrastructures Ukko2, Kale, and Vorna in numbers:

	Vorna	**Ukko2**	**Kale**
Nodes	180	36	60
Memory	c. 12 TB	14 TB	16TB
Cores	c. 6000 HT	2176 HT	1840 HT + 768
I/O nodes	12 I/O nodes	-	-
Topology	336Gb/s	336Gb/s	336Gb/s - 600Gb/s
GPU	-	8	20
Disc system	280TB (~4GB/s)	280TB (~4GB/s)	1PB (~20GB/s)

Locally, these clusters can have an additional fast, Lustre-based disc capacity of around 1 PB.

In future, these three clusters and any possible new additions will form an infrastructure called Turso. The basic principle with Turso is that its parts can be switched out as their lifespan requires, without the service level changing.

Assessment of cluster lifespans

- Ukko2: serviceable until 2023-2024
  - Service agreement until 2021
- Kale: serviceable in stages until 2022-2026
  - The cluster has been extended in four parts, the oldest of which is not included in the service contract
- Vorna: serviceable until 2023-2024
  - No service contract, but we have spare parts for servicing the hardware
  - In production until 2023
  - Used for teaching until 2024

Here, serviceable means the lifespan of the technology. A significant part of the cluster hardware has outlived its warranty period.

New acquisitions for the clusters will mainly be made through the budget for the FGCI project. Other investments are made by university units and projects that participate in the cluster operations. The IT Centre has made minor investments, mainly into communications hardware.

The partners in the current cluster are the following:

- Ukko2: CS
- Vorna: CS
- Kale:
  - the vice-rector (Paula Eerola)
  - HIP (Katri Huitu)
  - HIIT (Petri Myllymäki)
  - The Institute of Biotechnology (Howy Jacobs)
  - Faculty of Science (Kai Nordlund)
  - Department of Languages (Sanna-Kaisa Tanskanen/Ulla Tuomarla)

Physically, the clusters are located both in the server rooms in Exactum (Ukko2 and Vorna) and the IT Centre server room in Viikki (Kale, the new FGCI hardware). The capacity services team of the IT Centre is in charge of the server room technology and maintenance in Viikki, and the IT for Science team in Kumpula.

Departmental computation infrastructures and other significant systems

The departments and groups in Kumpula have their own calculation clusters and other important IT systems. Such systems include

- The Department of Computer Science (CS):

        - Honka (96 cores and 3T of memory)
        - Haapa (96 cores and 1.5T of memory)
        - The Nodes Laboratory (Exactum B124, front space; B125, Faraday cage)
        - Usability research UIX laboratory (Exactum A339)
        - Software Factory for teaching and research in software engineering (Exactum BK113)
- The Department of Geosciences and Geography:
    - the geoscience cluster (Whipp), (33 nodes)
- The Department of Physics and the Helsinki Institute of Physics (HIP):
    - the Alcyone FGI cluster (Chemistry and CS are also partners) (892 cores, 4TB RAM, 14TB discs)
    - PhEDEx, the data-transfer service for the CMS tests, Frontier, the calibration service, and the information system BDII.
- Department of Chemistry:
    - Gold, HP DL585, 64 core AMD Opteron 6386 SE. Acquired in 2013, under (extended) warranty until 2020
    - Silver, Dell PowerEdge R940, 96 core Intel Xeon Platinum 8168. Acquired in 2018, under warranty until 2023
- The Institute for Atmospheric and Earth System Research (INAR):
    - eddy2. atm server: HPE ProLiant DL380 Gen10 16 physical cores,128G of memory, 40T of storage

Some of these can be connected to the above-mentioned Turso complex, but due to the nature of their operations, some of these systems will remain in the use of the departments or laboratories, and are their responsibility.

Furthermore, the following storage systems are located in the campus server rooms:

- a NetApp storage systems (Numenor, Kratti) Service agreement until 2021
- a Lustre disc system (Vakka) No service agreement

Workstations in the charge of the faculty

There are a total of 2156 workstations under centralised maintenance of the IT Centre on the Kumpula campus. More than half of them are used by faulty staff. Further, there are 1700 laptop computers for freshmen, which have been distributed to students (hardware registry information). In addition, there are a significant number of workstations on campus that are not included in the centralised maintenance, such as high-performance workstations or lab equipment.

Computer labs managed by the IT Centre

Physicum D210, 16 workstations

Physicum D211, 13 workstations

Compute labs managed by departments

The departments in the faculty have several computer classrooms and other public-access equipment. The departments stand for the costs of their computer classrooms. All the equipment is maintained by the IT Centre except the Nodes Laboratory in CS.

CS

Exactum B221, 25 workstations, need to renew 15 workstations 06/2020, 10 CUDA computers 12/2023

Exactum BK107, 40 workstations – during 2020, the rate of use of these classrooms will be mapped and their purpose reconsidered

Exactum DK107, 32 computer docks, primarily the workspace for software engineering project work

Exactum DK110, 18 workstations, freely accessible for students, including evenings

Exactum DK108, 3 workstations, freely accessible for students, including evenings

Geo

Physicum A113, A114: 24 workstations, need for renewal 07/2020

Physicum A112: 18 workstations, need for renewal 11/2021

Physicum A111: 8 workstations, need for renewal 12/2023

Physicum, environment for virtual reality, two workstations, need for renewal 02/2024

Chemistry

Chemicum, lab spaces, 10-15 computers in common use, need for renewal 12/2023

Mathematics

Exactum C128, 26 workstations, need for renewal 09/2023

Physics

Physicum, teaching labs A209, three workstations, need for renewal 2 computers 12/2020, 1 computer 12/2023

Physicum, teaching spaces D204 one workstation, and D208 one workstation, need for renewal 12/2023

AV equipment in teaching and meeting rooms managed by departments

The need to renew AV equipment has been mapped in summer 2019, the report is in attachments. The state of the equipment varies, and the digitisation of projectors half-finished.

The faculty has decided to renew the following facilities and their AV equipment

Department	Needs per department
INAR	Physicum B409: a new screen, mobile stand and soundbar (combination if loudspeaker + camera + microphone*)
PHYS	Hall1 (at the accelerator): Renewal of projector
	Seminar room 115 (at the accelerator): renewal of video projector + automatic distributor
	K134F (at the accelerator – renewal of video projector + automatic distributor
CS	Exactum
	C221 and CK110: 2 projectors
	B221: digitalisation + new projector filter
GEO	Physicum C107: 1 projector
	Physicum A111+112: 1 projector
	Physicum A113 and A114: 2 projectors
CHEM	Chemicum A227: 1 projector

Steering machines for measuring instruments in the labs managed by departments

There are a significant number of workstations in the laboratories, which are not included in the centralised maintenance, such as the steering machines for measuring instruments. The installation of operating systems on the workstations is mainly the responsibility of the IT Centre. Most of the equipment is on a closed lab network.

The current state of workstations varies, the management of their lifespan is inadequate.

The infrastructures of the IT Centre in Kumpula

The most significant infrastructure component the IT Centre has on campus are the telecommunications connections. The IT Centre’s communications group is in charge of telecommunications on Kumpula campus and its connection to the backbone network of the university.

The following networks are deployed on campus:

- Landline network
  - closed lab networks
- Wireless network (eduroam)
- Visitors' network (HUPnet)
- Open wireless network in Ubicampus (speed: 1G). Discussions have been ongoing about extending the wireless network to the whole campus.

The university is updating its own backbone network to 100 Gbit/s , which will cut back the need for separate light paths and other similar extraordinary solutions.

Infrastructures of CSC

Kumpula campus is the largest single user of national computation resources.

The primary infrastructures maintained and operated by CSC are

- Mahti the supercomputer
- Puhti the supercomputer
- cPouta and ePouta cloud services
- Datalake Allas
- the Fairdata services (incl. the IDA storage service and the PAS service)
- Support services

Further infrastructures

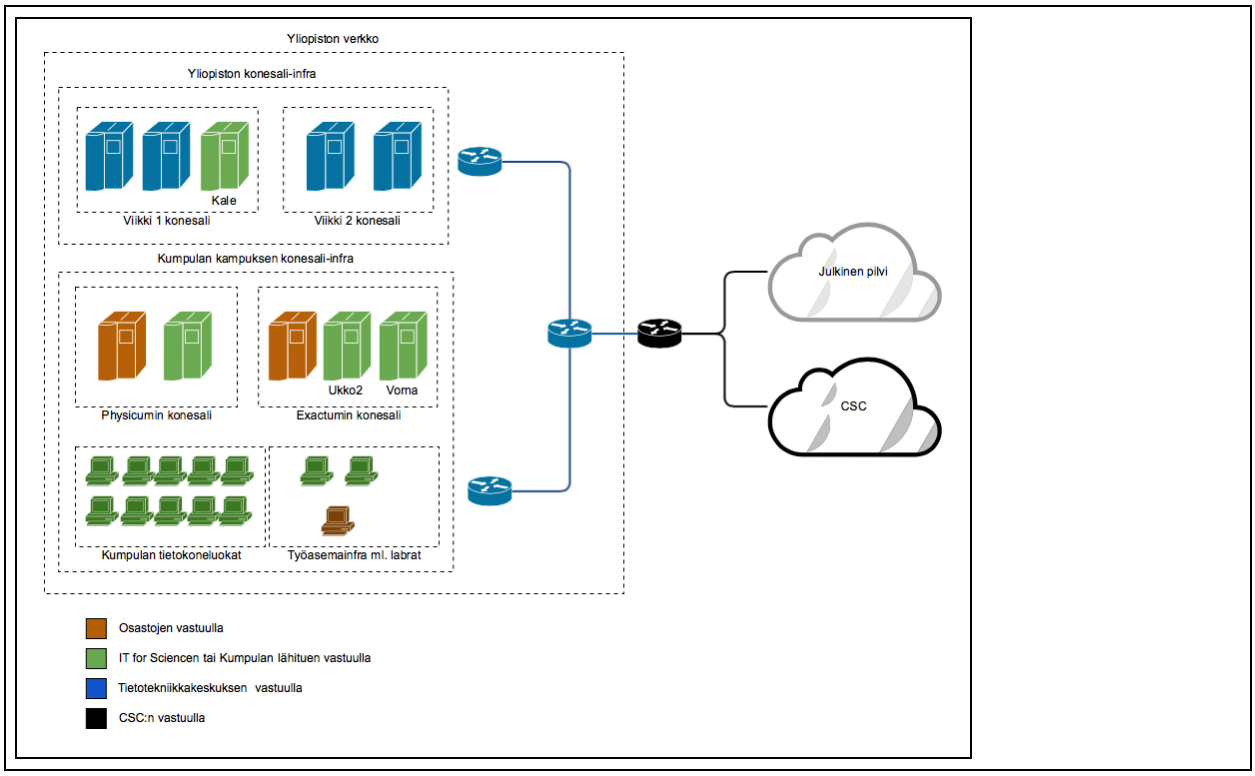
Commercial cloud services such as Amazon, AWS, Microsoft Azure, and Google Cloud are used to a minor extent on Kumpula campus. The total volume annually (2019) is a some tens and thousands of euros. The university has its framework arrangements, with which these services can be acquired, but on-campus, many such cloud services directed at consumers are used, though the user agreements and security may be significantly inferior to those within the framework. We do not have information about the extent of use of such consumer services.

2.3 Responsibilities for upkeep

The on-site support team is in charge of workstations and IT for Science of the server infrastructure when it comes to IT maintenance on Kumpula campus. Online communications are the responsibility of the telecommunications team at the IT Centre.

There are two main server rooms in Kumpula, the ones in Exactum and Physicum. Their upkeep is completely in the charge of IT for Science and departmental administrators. Further, there are clusters in the IT Centre server rooms in Viikki, which IT for Science is responsible for upkeeping.

The following figure shows the main features of IT infrastructures and administration responsibilities on Kumpula campus.



2.4 Management model

The dean has appointed the *Kumpula campus IT steering group* with representation from all the departments on campus. The steering group acts as the steering group for IT for Science at the Faculty of Science. The steering group sets the objectives for IT for Science, processes the reports on how tasks progress, decides on prioritisation of acquisitions for the most significant projects every six months, and when necessary, names the product owner in large projects.

The duties of the steering group also include drawing up a strategic vision of the IT operations on campus, i.e. this document. The steering group promotes IT operations and cooperation between all agents on campus, and supports the faculty leadership in developing the digital campus. It is up to the steering group to identify overall resourcing needs and bring them to the attention of departments, the dean, and the scientific expert and infrastructure group of the faculty (Matias). Vice-Dean Hannu Toivonen is the chairperson of the steering group.

The steering group has also appointed a client panel for IT for Science, and its chairperson (Professor Keijo Heljanko). The client panel has adopted the name *Kumpula IT infrastructure user group*, and 2-3 representatives have been appointed from each department of the faculty to discuss practical issues. The client panel reports to the steering group regularly.

The coordination of infrastructure projects at the faculty, including IT infrastructure, is the remit of the *scientific expert and infrastructure group at the faculty (Matias).* In developing IT infrastructure, the IT steering group is collaborating with Matias.

2.5 Current IT costs on campus

The annual costs for IT at the faculty and departments (except infrastructure projects) were around 1.8-2.0 Million euros[1] in 2018-9. This includes costs that have been allocated to the faculty departments. The numbers do not include the costs of the IT Centre, which mainly consist of salaries.

Half of the costs, i.e. 1 million euros, are used for acquisition of workstations and laptop computers. In 2018 and 2019, some €100,000 have been used for the acquisition of so-called high-performance workstations. In 2019, a total of €340,000 were used on laptops for freshmen, a total of 488 computers.

Acquisitions for infrastructure are not included in this itemisation. They are funded from separate sources (e.g. FGCI).

The costs on the university level are also affected by the expenses caused by the aging infrastructure. A typical example is outdated servers requiring more electricity and more cooling power than state-of-the-art servers. This does not appear in the IT costs for the campus, but in facility costs as e.g. electricity expenses.

In addition to economic costs the campus must pay attention to how its IT solutions affect the environment and sustainability.

3 Changes in operational environment

3.1 General changes to operational environment

The research and teaching on Kumpula campus are very dependent on available information technology. The significance of data- and computation-intensive research is large and keeps growing.

Other trends that are causing a change in operational environments include the growing corporate collaboration within many disciplines. Simultaneously, international cooperation, e.g. over EU borders, is constantly increasing in importance.

The increasing and border-crossing work is also putting pressure on service-providing both on Kumpula campus and at the university more generally. Services – whether they are end-user services or infrastructure services – can no longer be produced with only academic users in mind or a permanent contract with the university.

The changes are bringing forth the need to specify where IT for Science ends and CSC begins, including which CSC services are subject to a charge, which of them are fully or partially funded by the faculty, and which by research group project funding.

Strategic choices at the university

In the preparations[2] of the future strategy of the university for 2021-30, the following viewpoints on primary queries in environment analysis[3] have been listed:

Primary queries for environment analysis – solution for academic civilisation.

- **Teaching and learning:** How will our university foster critical and ethically operating future experts and trailblazers ?
- **Science and the infrastructures of knowledge:** How do we promote science that is open, works long term, keeps a high standard, and is responsible?
- **Science education and communications:** How do we enforce the university's role as an open learning and research environment that can predict social changes and pass on powerful information and expertise, on which to base solutions?
- **Working and studying culture:** How do we strengthen diversity and communal wellbeing at the university while giving everyone a chance to find meaning in their work and studies in order to attain our shared objectives?
- **Cooperation:** How do we increase the impact of research and innovations on a global scale and the competitiveness of the university via strategic cooperations?

The University of Helsinki strategy for 2021-30

The board of the University of Helsinki approved the 2021–2030 strategy for the university on 26 February 2020.

To be a stronger creator of culture and wellbeing, our vision for 2030 is

The power of science for the benefit of the world

Ground-breaking basic research, unprejudiced curiosity, and limitless cooperation are the prerequisites for rising into the global elite. We are building a sustainable future in the spirit of global responsibility by constructing understanding for solving both local and universal problems – for the benefit of the whole world.

The strategic choices for implementing our vision are:

1. everyone has a right to science and learning
2. openness enforces science and cooperation
3. our university is the best place for studying and working
4. Our university is a champion of answerability and sustainability

Data policy

The research data policy approved at the University of Helsinki in 2015 can be sketched with the following headings:

"Good management of data is taught and observed at the University of Helsinki. Each member of academia is responsible for the observance of good data management.

- Good data management is the duty of every university
- Research-data know-how is part of every researcher's education
- The university serves up the infrastructure for research data
- Support in legal issues about data management
- Data-management plans and meta-data
- Research data is shared and open
- The research-data policy pertains to digital material."

3.2 Changes in operational environment on campus

Data-intensive research

The amount of data-intensive research is increasing on campus and at the university, while its methods are developing. This increases the demand for computation and storage infrastructures, whether they are produced by the university, national and international discipline-based networks, or e.g. funded by the EU. Data-intensive research is becoming more common in many disciplines, and also opens up opportunities for inter-disciplinary collaborations. The growth increases the need at all levels of service provision and the challenges cannot be solved with just one single service, infrastructure or concept.

This creates new or increasing requirements for the services offered to students, teachers, and researchers.

The requirements for long-term storage of data are growing and extending to include raw data as well as material or data sets that are included in publications. This means demand for more storage capacity and a longer lifespan for storage. The transfer of data and material will also grow significantly in future, which may pose its own demands on the capacity of available data networks.

The quantitative and qualitative increase of data also means that the meaning of data management will grow, so that ownerships of data, management of meta-data, the discoverability of data, and the reusability of data sets must increasingly be acknowledged for their whole lifespan. A primary and general trend is that the significance of data and data sets at the back of the repeatability of experimenting is growing, i.e. the pressure to open up more and more data sets is increasing. The requirements for open research are causing an ever-growing need for researchers to publish large data sets.

International research collaborations are increasing, which means sharing more and more data and infrastructures with outside (including non-European) research groups and scientists. Such cooperation bodies are not necessarily given access to e.g. EuroHPC. Some parts of data or materials may gain international interest, which increases their use and various related technical requirements, as well as the management or administration of material (access rights, legal issues, funding, etc.). Internationalisation also calls for global standards like the FAIR principles or other customary practices to be applied when it comes to services produced on campus or purchased for the campus. The FAIR principles, or more generally, the openness of science is significant in relation to research funding, such as funding applications from the Academy of Finland or EU's Horizon programmes.

Data intensity has already brought with it new requirements and changes in work procedures. There are new demands at least pertaining to 3D-modelling and Virtual Reality technologies. The significance of visualisation and GPU calculation will increase in future.

Studies and teaching

Studying and teaching are turning digital and moving towards the 24/7 concept. The tools for digitalised learning include massive open online courses (so-called MOOCs) and other learning environments that are independent of place and time. Furthermore, teaching and studying can be supported with new digital tools and algorithms, such as Oodikone that has been developed in Kumpula.

Studying is moving towards a constant-learning model, where some courses are completed distantly and the operational environment is increasingly international. Naturally, this demands suitable tools. We have to analyse whether the freshmen's laptops are the correct and cost-efficient solution, or whether we could provide students with the course environments required through e.g. VDI.

It is important for teaching and especially studying that the working environments for students hold a high standard while being similar to each other. Their similarity enables efficient centralised support for the students, as well as peer-to-peer support among the students, and the interoperability of applications can be assured without extra adjustments. An alternative to ensuring workstation-based interoperability is course-based VDI workstations (Virtual Desktop Infrastructure), which can be used on very different kinds of workstations. In this way, we could ensure that all the participants in a course have the same and correct work environment ready at their hands, including Master's students and continuous learners, who do not have the freshmen's laptop.

Constant learning (and research etc) is apparent to the IT support services in that they receive help requests and error reports any time, regardless of date or time.

The strategy for constant/open and online teaching is being developed now at the Faculty of Science.

Changing the IT service models

The need for IT services change as operations change. On Kumpula campus, new and developing technologies also act to change the needs of the operations. As new opportunities appear, they will be seized and our own operations developed to bring new IT services to the attention and use of ever larger user groups.

The IT service models will develop further in accordance with user needs and demand. The advantage with services such as cloud computing is that their acquirement has been automated and defaulted so that the users can deploy them themselves. Services models are not as well suited for projects or use cases where the available capacity has to be tailored extensively, or where there is a distinct reason to keep the data close to the working environment (e.g. transfer of masses of data etc). In addition to SaaS (Software as a Service), PaaS (Platform as a Service), or IaaS (Infrastructure as a Service), current cloud services are also developing in the direction of services on a higher level of abstraction, such as FaaS (Function as a Service), which means the user does not need to worry about the platform or application, but can run application functions such as ready algorithms. Examples of such platforms include AWS Lambda, Google Cloud Functions, or Microsoft Azure Functions.

In science IT, the cloud model used is the Pouta services produced by CSC, services based on the IaaS model to provide different kinds and sizes of virtual servers. Some amount of similar services are available both from the IT Centre and the platforms locally maintained in Kumpula.

The meaning of usability grows for all IT services as studying increasingly disconnects from time, place and degree, and as research is entwined into research data and its services.

Local vs cloud services

In principle, it is easy to gain capacity from cloud services to a very scalable degree.

We can expect the services of CSC to be used more and more on campus since CSC has a considerable amount of computation capacity for new scientific questions, along with related storage space.

Compared to using CSC's or UH's own service, the use of commercial cloud services requires larger economic investments (e.g. for long-term data storage). Economical reasons can thus limit the use of commercial cloud services as a part of a research project. The services of CSC are typically free for Finnish research projects and the services of UH are usually free for university researchers if they are funded with university funds. It is also easier to come by support for internal services than for external cloud services, since the support services of UH only have little experience with external cloud services.

Adding to current use of external services will also increase the significance of communications as part of the infrastructure.

Local resources and infrastructures hold their own place in the whole when considering e.g. the need for developing education, applications, and algorithms. The services on various levels will continue to complement each other.

Cloud services are not suited for all kinds of research, e.g. because transferring data may be very time-consuming and unfeasible due to data security or protection reasons.

The above-mentioned implementation needs for configuration and application environments can limit the use of cloud services or use of services external to the university in general (including licensing reasons). It makes sense to use computation environments in teaching or algorithm development and testing in local environments before the true production runs are made in larger environments. For storage, the local storage services on campus or e.g. the university's Datacloud service are suited for such use cases, where the researcher does not want to transfer data too far away or it is not connected with e.g. computation.

Technological advances

Technological advances will keep creating new opportunities in future, but also new demands on support services. In short, we could mention quantum computation, which the Department of Chemistry is experimenting with in 2020; CSC is acquiring resources in this area, as well. Another example is object storage that is becoming increasingly common in storage solutions in parallel with traditional file systems. Let us separately discuss two important areas of development: graphical processors (GPU) and virtual desktop interfaces (VDI).

- Container technologies
- We must follow up new technological solutions and react to them.

The deployment of new technologies requires its own support services and processes. A basic knowledge of new technologies is vital for the support services to provide efficient solutions and redirection to the correct service point.

GPU

In the past few years, traditional computation based on CPU use has been joined and partially replaced by computation using GPU (Graphics Processing Unit). GPUs are especially well suited for parallel computation, which significantly benefits many tasks within scientific computing. AI applications are one example of suitable application areas. In new national super-computers, there is a multiple amount of GPU resources compared to before, and thus the importance of GPU technologies and applications is growing considerably. However, the utilisation of GPU resources requires a new kind of programming, e.g. using TensorFlow or CUDA[26]. These new technologies require a new kind of skills from the user and new kinds of support services.

VDI

VDI (Virtual Desktop Infrastructure) services have emerged alongside the traditional workstation infrastructure, which means the hardware and OS infrastructure is for the most part maintained on the servers rather than the workstation equipment. The VDI services are not yet ready and their usability does not reach a level where they could replace all needs in practice.

Data networks

Data communications hold a primary role in current research and teaching infrastructures, and their significance will grow as services are increasingly accessed on the local server level or, e.g., cloud services. Similarly, the use of mobile networks will increase in relation to the traditional fixed network.

The above-mentioned VDI services, for example, demand reliable and functional network connections, though they do work serviceably e.g. on 4G-network speeds. The base network of the university is already being raised to a speed of 100Gb/s, which will help in many ways and cut back on the need for separate light paths.

Sensitive data and regulations

The past years have seen an increase in regulations for the handling of personal data[9] and sensitive data[10] in research with the event of the GDPR (General Data Protection Regulation). The GDPR does not have such a large impact on Kumpula campus as some other campuses (such as Meilahti, Central campus), but within administration and some research projects there is a clear connection.

Furthermore, two new sets of regulations have been enacted recently, the information management act[12] and the accessibility directive[13]. Taking them into account is primarily the university's business, but will be reflected on the campus level at a later point when instructions have been set.

IT support and integration into research groups

As the need for services grows and they become ore varied, we will also need more extensive support to implement and use the services. As the significance of data has grown, new needs have also arisen.

The current IT services at the university do not quite meet the IT needs of research groups. The IT services are mainly concentrated to generic IT services for workstations and server environments, while also providing consultation on the use and configuration of these environments. For applications and specific research problems, the university's IT services offer little support or solutions. There is both demand for these services, and need for them within the support services. Within resources, we would like to recommend that on-site support and IT for Science acquaint themselves with the services of other bodies providing support for scientific computation and data-intensive research, like CSC (https://research.csc.fi), EGI (https://www.egi.eu/services/), EUDAT (https://eudat.eu/services/) or GÉANT (https://www.geant.org./Services), so that they can learn to support researchers on campus in acquiring and deploying them.

The gap between research and IT has been bridged with the help of CSC staff in some university groups.

The primary identified need is provision of data management services. This means comprehensive support for data-intensive research so that researchers and doctoral students are

provided with solutions to problems concerning scientific computation and data management, while the development of more efficient work methods is promoted. This service also offers data-related technical services from infrastructure and applications (platforms, servers, workstations, applications. etc) to issues typical for research groups (algorithms, optimisations, etc). Data management support is its own chapter in the whole, i.e. such as how it would be sensible to organise the data so that it is as usable as possible at different stages of its lifespan. Moreover, we also need more support for existing infrastructures and applications. The role of data steward[23] is part of this whole. The role includes services and function pertaining to data management and quality assurance.

In a wider perspective, it is about implementing a sort of IT porter service, where the users' problems are met as individually as possible, though the solutions may be quite standardised. The same service would include the mapping of how opportunities offered by IT could be utilised as efficiently as possible to support studying, teaching, and research.

One growing need is the support of infrastructures external to the university, since it makes sense to use e.g. CSC's services more widely on campus in future.

The university's data support network[14] will provide support for computing in future, as well. This is a network support service jointly organised by the library of the University of Helsinki, the IT Centre, and research administration. This service offers university staff support and instructions on the FAIR principles, storage, project funding, data management plans, etc, as they pertain to research data.

3.3 International and national development projects

Scientific computation

In future, the significance of international collaboration projects will grow. For the Kumpula campus, the main one is the EuroHPC[15], which will provide Finland and Finnish researchers with a significant amount of calculation capacity for scientific computation in the form of the LUMI super computer[16]. The LUMI super computer is expected to be in use in early 2021.

Before the event of those, the national supercomputers Mahti and Puhti[17] brought in by the DL2021 project have been or will be deployed by CSC. Once all these new resources are in use, the capacity for scientific computation in Finland will be about 15-20 times as high as in January 2019. This is a huge increase that offers completely new opportunities.

DL2021 has also provided a storage solution called Allas.

Data management

In developing techniques, architectures, and procedures for data management and use, the European Open Science Cloud (EOSC)[18] is the main international project. It produces development projects in dedicated areas and funds multinational projects that often have a scientific edge. The objective of EOSC is to provide 1.7 million European researchers and 70

million professionals in different fields with a virtual environment with open and seamless services for storage, management, analysis, and reuse of research data, across borders and scientific fields. EOSC combines existing scientific data infrastructures that are currently distributed within disciplines and EU member states.

In future, EOSC and EuroHPC will be connected to each other in some way.

On a national level, the coordination of national open science and research managed by TSV can be considered the equivalent of EOSC.

3.4 Available resources

There are no changes in sight for the basic funding of the Faculty of Science and the IT Centre. When producing services, the importance of cost efficiency grows, while at the same time available resources must be used as flexibly as possible to develop them. Gaining new competences for support services demands time and education.

When developing the infrastructure, project funding becomes more important. When it comes to computing infrastructures, FCCI (former FGCI[20]) has gained good experiences of that. However, the university's own computing environment has already been planned to be able to incorporate new computing resources from various projects, which means the funds are in joint use on the university level. The same logic has been used when building the university's own research-data infrastructure Datacloud[21].

4 The vision

A strategic vision for developing IT operations in Kumpula:

Kumpula campus is a global trailblazer for computational science.

5 Outlines

Here, we present the outlines that we want to emphasise for the development of IT operation on Kumpula campus.

Principles

- User-oriented solutions are the main principle for the development of all services and their accessibility – it must be easy to use and access services.
- We will develop the skills of researchers to utilise the available computing infrastructure.
- Different resources will be utilised efficiently, joint use of resources will increase
- Everyone has suitable tools available for their duties

- We will invest in our own infrastructures along with national and international infrastructures

Capabilities

The primary capabilities in Kumpula campus IT operations, which will be developed further in the near future are:

- Infrastructure capabilities: use of cloud services, appropriate use of workstation environments, more efficient utilisation of VDI, better and faster network connections
- Skills and education: services for data management (data manager, data engineer, etc) as well as support for scientific computing (e.g. GPUs etc)
- Processes: simplification and better transparency
- Communications: a better representation of what services, resources, or capacities are available (and e.g. for teachers better information about existing services)
- User-centred solutions: making the planning, implementation, and further development of support services and infrastructures more user-centred
- Governance: administration needs to be efficient in supporting the core operations; we need up-to-date information on the lifespan of existing components, and especially their rate of aging, as well as investment needs

Cooperation and information flow

- We will form closer cooperations between research and the IT services, so that the services can understand the needs of research better, and research is aware of existing possibilities
- For present and future users of campus IT services and infrastructure, we will provide an easily accessible place to find information on services available on campus (whichever body is providing the service)
- On-campus researchers, teachers, and students will be informed about IT infrastructures and services via various communication channels and events, such as the PI meetings in cooperation with the faculty, and other similar events
- We will implement education about the methods, resources, and infrastructures of computational science as facilitated by IT for Science, and market suitable courses on it
- Researchers will be activated to give feedback and information about what kind of services are needed, and to provide peer support, e.g. in the form of the Kumpula Campus IT-infrastructure User Group
- In parallel with the Kumpula IT-infrastructure User Group, we will create methods for routinely updating an overview of the needs of research groups
- When it comes to research data, campus operations must aim at increasingly opening up the material and implementing the FAIR principles more systematically
- The faculty will be encouraged to find data agents

Working environments

- Both staff and students are offered appropriate tools for their tasks

- The laptops for freshmen ensure that students have up-to-date tools and a standardised work environment for seamless completion of courses and related support – including peer support. However, it is a considerable investment, so the policy for laptops for freshmen should be revisited regularly to make sure it is expedient.
- The need for computer labs on campus must be checked by department, and we must establish the needs of the education programmes

Infrastructures

- In addition to utilising CSC services and international infrastructures, the campus will continue to invest in its own scientific high performance computation infrastructure
- When further developing infrastructures, we have to consider overlapping services on the faculty level so that they are not implemented as bans, but as conscious and justifiable choices, e.g. by the campus IT steering group finding existing solutions for new projects (this will require a campus-level mechanism that may assess the impact of projects on the infrastructure)
- We will build a management model for the lifespan of hardware and applications, as well as equivalent investment plans to display e.g. the technology debt
- We need a separate overall plan for the current state and development of the management and storage of research data to take into account the local storage infrastructures in Kumpula and their future, as well as the options available for the rest of the university, so that it is possible to handle research data throughout its lifespan in a controlled and efficient way while taking into account the FAIR principles etc.
- We want to extend the use of VDI infrastructures, e.g. for specific project-related workstation needs
- Services produced and mediated by CSC should be adopted more widely as a part of our own service supply
- The computation and storage infrastructures of the university (Turso, Datacloud) should be consolidated through project support
- We will maintain the external network and campus network connections at a sufficient level
- When developing infrastructures or IT operations in general in the future, we also need to take into account sustainability aspects in accordance with university-level recommendations, for example

Support services

- The growing significance of data-intensive research also requires developing support services, e.g. in the roles of data engineer, data manager, or data steward
- The support services will be brought closer to research groups and departments so that the services can be tailored ever more individually
- Locally, we will maintain our know-how to the degree that we can help researchers on campus get started with e.g. the infrastructures maintained by CSC
- When developing the support services, we will cooperate with central service providers like CSC

- The utilisation of the university's data support (a joint project between the university library, IT Centre, and the university's research services) could be made more efficient as a part of developing data management

References

1. Numbers from SAP: IT expenses 2017-2018
2. Strategy preparation: https://flamma.helsinki.fi/fi/group/yliopisto/strategia-2030
3. Main issues in environment analysis: https://flamma.helsinki.fi/fi/group/ajankohtaista/uutinen/-/uutinen/ajattele-2030-strategiatyon-ymparistoanalyysi-on-yliopistoyhteison-osalta-valmis/3263632
4. The strategic choices of committees as the base of final choices:

   https://flamma.helsinki.fi/fi/group/ajankohtaista/uutinen/-/uutinen/kommentoi-ehdotuksia-strategian-2021-2030-valinnoiksi-5-11-mennessa-/6501285

5. Committees: the scientific committee TINE, the steering group of doctoral education, the steering group of corporate collaboration and innovation work, and the infrastructure committee for research
6. The future of the strategy work: https://flamma.helsinki.fi/fi/group/ajankohtaista/uutinen/-/uutinen/rehtori-niemel-c3-a4-kiitos-kampuskohtaamisista-ja-strategiaty-c3-b6h-c3-b6n-osallistumisesta-/7463919
7. Research data policy: https://www.helsinki.fi/fi/tutkimus/tutkimusymparisto/tutkimusaineistot/tutkimusdatapolitiikka
8. FAIR principles: https://www.go-fair.org/fair-principles/
9. Personal data: https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_fi
10. Sensitive personal data: https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en
11. GDPR in Wikipedia: https://en.wikipedia.org/wiki/General_Data_Protection_Regulation
12. The law on public administration information management: https://www.finlex.fi/fi/laki/alkup/2019/20190906
13. The accessibility directive: https://vm.fi/saavutettavuusdirektiivi
14. Data support: https://datasupport.helsinki.fi/
15. EuroHPC: http://eurohpc.eu/
16. LUMI: https://datacenter.csc.fi/wp/about-lumi/
17. Mahti, Puhti, and Allas: https://www.csc.fi/web/blog/post/-/blogs/new-computing-and-data-infrastructure-puhti-mahti-and-all-1
18. European Open Science Cloud: https://www.eosc-portal.eu/
19. Open science and research: https://avointiede.fi/
20. Finnish Cloud and Grid Infrastructure (FGCI): https://research.csc.fi/fgci

21. Datacloud: https://wiki.helsinki.fi/display/Tutkimusdata/Getting+started+with+Datacloud
22. JHS 179: http://www.jhs-suositukset.fi/suomi/jhs179
23. Data steward: https://en.wikipedia.org/wiki/Data_steward
24. Graphics Processing Unit (GPU): https://fi.wikipedia.org/wiki/Grafiikkasuoritin
25. TensorFlow: https://en.wikipedia.org/wiki/TensorFlow
26. CUDA: https://en.wikipedia.org/wiki/CUDA
27. The University of Helsinki strategy for 2021-2030:https://www.helsinki.fi/fi/yliopisto/strategia-2021-2030/helsingin-yliopiston-strategia-2021-2030-tieteen-voimalla-maailman-parhaaksi#section-90808

Appendices

1. AVtilojenuusmistarpeet 2019.xls