

HUNT Cloud Services 1.5

We believe it should be a simple thing for researchers to get flexible, elegant and secure computer environments.



What you will learn

This eBook introduces scientific cloud services provided by HUNT Cloud.

The primary audience of the eBook is scientific research leaders in need of digital lab environments.

Contents

Lab environments	4
Compute	5
Store	6
Transfer	7
Secure	8
Comply	9
Participate	10
Start	11

Lab environments

HUNT Cloud provide digital labs for analysis and storage of sensitive data - including research unrelated to HUNT. Our primary goal is to give researchers more precious time, power and possibilities to extract knowledge for a better world!

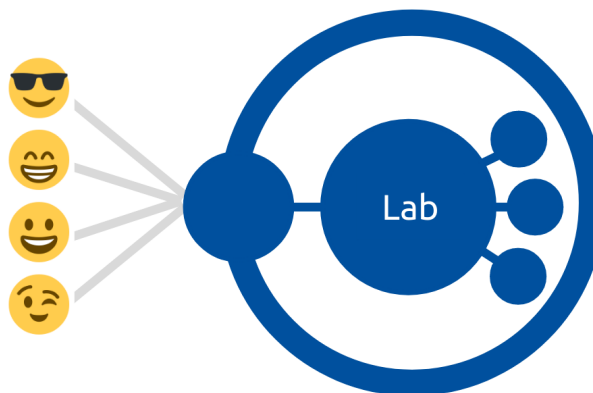
Managed environments

We offer managed digital labs as turn-key ready research environments.

We provide compute and storage resources that are tailored to your research group's workload. We include common analytical tools that allow your lab users to start their data analysis right away. You identify and authorize your own lab users whom can be located around the globe.

We manage a comprehensive security and compliance framework that allow your lab to process sensitive research data out of the box.

We organize your lab as a «gated community». This means that we carefully control and monitor the outer fence so you get the freedom to scale resources and install software on the inside in accordance with your need.



Unmanaged machines

You may increase the power of your lab by adding new virtual machines within your lab environment.

We provide these as unmanaged machines with a clean operating system. Your lab get full administrative privileges to install, configure and maintain these machines in accordance with your need.

Unmanaged machines are also great for short-lived analysis that allow you to scale your lab to quickly analyze large amounts of data, such as batch processing of biomedical data.

Compute

We believe that information should be analyzed to the widest possible extent to generate the widest possible set of knowledge!

Flexibility

You get access to the flexibility of cloud computing. You may quickly add or adjust the amount of compute power and memory as your analytical needs develop over the life-cycle of your projects.

This flexibility do also allow you to explore new tools and analytical approaches as soon as they emerge. All to add a competitive advantage in your exciting race towards more knowledge for a better world.

Machines types

We provide a range of machine types that can be deployed within your lab:

CPU optimized machine types. We provide three series of machines types optimized for computation with 2GB, 4GB or 8GB of memory per virtual CPU. Sizes range from 2 vCPU to 48 vCPU per machine.

GPU optimized machine types. We provide a limited series of machine types intended for machine learning. These machine types are fitted with Nvidia Tesla P100 GPU accelerators.

[Click here to see our machine types](#)

Software

We provide you with the opportunity to install your own software in addition to those managed by us. This may include software such as R, Stata, SPSS and Matlab, container virtualizations such as Docker and Singularity, workflow managers like Snakemake, and big software deployments such as Kubernetes, Slurm and Spark etc.

We do also provide **graphical interface** for commonly used software such as:

STATA®



MATLAB

R Studio®

IBM
SPSS

Store

We believe that knowledge should be derived from the widest possible set of information.

Scalability

We have designed our storage system so you can store and retrieve any amount of data in a simple and secure manner.

We recognize that your ability to collect information evolve faster than regular storage technologies. We are therefore continuously optimizing and developing our storage services towards reduced cost and increased performance.

Storage types

Block storage. This is our main storage resource for your lab. We deploy a predefined set of volume types to guide your lab users towards improved performance and cost-effective big data management. You choose the size of your storage volumes and quickly upgrade these as your data grows.

Tape storage. We provide encrypted tape storage for data that need special protection. This services is also well-suited for cost-effective storage of data with long-term retention requirements, such as requested by regional ethical committees. Tapes may be stored offsite and offline for increased protection.

[Click here to see our volume types](#)

Restore

We provide restore services to protect your data against accidental deletions and accidental overwrites. We have already activated nightly restore functionality for some volume types. You may include other volume types and adjust the backup frequency in accordance with your restore needs.

Your data is automatically replicated three times across three physical storage machines to protect against equipment failures. Your data is stored on equipment in Trondheim, Norway, that is owned and managed by HUNT data center, NTNU.

Transfer

Internal transfers

Lab users may transfer data between your lab environment and their local computers. We protect this transfer in encrypted communication channels.

You may request transfers to other lab environments within HUNT Cloud. We protect these transfers so you can quickly share data with named collaborators from other lab environments in a secure and compliant manner.

External transfers

You can authorize transfers between your lab environment and external collaborators. We provide several secure solutions for such transfers. The choice of solution depends on the data types, expected duration of the transfer and the technical capability of the external party.

We do also provide secure solutions for permanent transfers into lab environments from data capture equipments, such as sequencers and image machines.

Secure

We believe that secure systems are easy to use.

Design principle

We strive to provide elegant and easy-to-use security controls for optimal user compliance.

Our role model is «drinking water» for which you may enjoy instant access without thinking too much about the purification process. We do therefore allow for an initial installation of «water pipes» as long as you may enjoy subsequent instant access without thinking too much about the security controls.

We ensure that your data is protected by an extensive list of 121 information security controls, including key controls, such as:

Key security controls

- Lab users are required to enter an authentication code in addition to their regular credentials to log in.
- Each lab user has unique credentials to ensure one person per login.
- Access from the outside are only allowed in encrypted tunnels.
- Access are restricted to communication that are agreed.
- Labs are logically isolated from each others network communication.
- Research projects are logically isolated from each others data.
- All data are located in Norway.

Monitoring and logging

We monitor an extensive list of system activities, including access attempts and network activity.

Individual logs are carefully preserved in accordance with their collection purpose to allow for retrospective audit capabilities.

Comply

We believe that trust from data donors is the single most valuable asset in biomedical research.

Laws and regulations

We enable your research to be compliant with Norwegian acts and regulations that regulate research, such as the Personal Data Act (Personopplysningsloven), the Health Research Act (Helseforskningsloven), the Health Registry Act (Helseregisterloven), the Health Personnel Act (Helsepersonelloven), and the Data Protection Directive from EU (GDPR).

ISO 9001

Our compliance with the international standard in quality management «ISO 9001» is certified by Kiwa, Norway.

[Click here to see our «ISO 9001:2015» certificate and scoping document.](#)

ISO 27001

«ISO 27001» is one of the most widely recognized and accepted independent security standards. Our compliance is certified by Nemko, Norway.

This international standard specifies requirements for establishing, implementing, maintaining and continually improving an information security management system. An independent auditor examines our data center, infrastructure and operations.

This standard helps you, data controllers, and regulators, to confirm that our operations meet strict security and compliance requirements.

[Click here to see our «ISO 27001:2013» certificate and scoping document.](#)

Self assessments

We do maintain self-assessed statements of applicability for «ISO 27017» that cover cloud security and «ISO 27018» that cover personal identifiable information in public clouds.

We do also maintain self-assessed statements of applicability for the Norwegian «Norm for informasjonssikkerhet» and Subpart C of «HIPAA».

Audits

We are proud to work with data controllers that expect strict independent compliance verifications. We welcome audits from data controllers and research leaders to ensure compliance with their expectations.

Participate

We are proud of the warm and helpful tone among our users!

Cloud Community

We invite all lab users to our «cloud community». This enables lab users with a virtual space to meet users in similar situations from other labs.

The primary goal of our community is for users to learn from each other when needed and provide assistance to others when they can.

We do also utilize the collective intelligence from our community for continuous user feedback and to quickly communicate new features and services across labs.

Lab forum

We invite all lab owners to participate in our quarterly «Lab forums». We use this forum to get feedback on our services and discuss strategic aspects of our system development.

We are humbled by this opportunity for structured assistance from such an experienced tier of lab owners.

Partner program

We partner with impact centers and early adopters to ensure continuous competitive development of our cloud services.

We collaborate closely with each partner in our program to tailor solutions that elevate their scientific goals.

New features that gets developed in the program are made available to all users within our «cloud community» as soon as the quality gets approved by the partners.

Start

Give us a shout if you believe your research may be elevated by cloud services.

Preparation

We usually start with a «virtual coffee» to learn more about your research and discuss how we may optimize your lab environment towards your scientific goals.

We will need a «Data Processor Agreement» for labs that are affiliated with data controllers outside NTNU, as well as a «Project Agreement» that define your lab's resource allocations and maintenance responsibilities.

We will deploy your lab as soon as the paper work are in place.

Onboarding

We recommend that you identify one or two lab users that can act as «technical coordinators» for your lab.

We will work closely with your coordinators to ensure that the lab environment meet your strict expectations before we onboard your remaining lab users. Lab users will follow our «onboarding guide» for their first login.

[Click here to see our onboarding guide.](#)

Support

We aim to empower your technical coordinators to provide smooth «first-level» support for your lab users.

We will provide «second-level» support for questions that are not resolved within your lab or our cloud community.

We have partnered with good friends at several core facilities to ensure access to expert help for your lab users when it comes to scientific questions.

About HUNT data center

HUNT Cloud is delivered by HUNT data center at NTNU, Norway.

The HUNT study in Norway has controlled digital sensitive health information since 1984.

HUNT data center. HUNT data center is part of the HUNT Research Centre at NTNU. HUNT data center was established in 2013 to enable collection, accessibility and analysis of large-scale biomedical data. The center is driven by three core missions: (i) to elevate research, innovation and health services, (ii) to transfer the excitement and expertise of big data processing to the next generation of biomedical scientists, and (iii) to organize competitive digital infrastructure to facilitate the evaluational and educational missions.

HUNT Cloud. HUNT Cloud is the rental part of HUNT data center that deliver flexible, simple

and secure computer environments to a wider biomedical community, including project unrelated to HUNT.

Culture. HUNT data center is continuously developed by a small dedicated team that wholeheartedly believe in information as a source for knowledge! The team is organized around site-reliability engineering principles doing everything from core infrastructure and unboxing bare metal to guiding researchers towards workflow magic. The team is located at the K.G. Jebsen Center in Genetic Epidemiology in Trondheim. HUNT data center offer internship positions.

Technology. HUNT Cloud is designed on open-source software and open infrastructure prin-

ciples utilizing OpenStack cloud orchestration and Ceph distributed storage in symphony with MAAS, LXD and Juju. Additional software and tools include Ansible, Python, R, Docker, Singularity, Prometheus and Grafana. Code and software are continuously tested in dedicated environments.

Responsibility. HUNT data center is part of the HUNT Research Centre, Department of Public Health and Nursing, Faculty of Medicine and Health Sciences, NTNU, Norway.

Image courtesy. Front page photo by Anne-Line Bakken, ISM/NTNU.

eBook version. Rev 1.5 from 2019-07-14.

Contact information. cloud@hunt.ntnu.no

www.ntnu.edu/mh/huntcloud