EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

de.NBI Cloud
German Network for Bioinformatics Infrastructure

Jens Krüger
High Performance and Cloud Computing Group
Eberhard Karls Universität Tübingen

December 05th 2023, Munich

de.NBI
GERMAN NETWORK FOR BIOINFORMATICS INFRASTRUCTURE

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Agenda

- Infrastructure
  - de.NBI Cloud
  - Sustained storage
  - Secure processing environments
  - Information security management
- Legal framework
  - General data protection regulation basics
  - Processing contracts
  - Standard operating procedures
- Research data management
  - Annotation
  - Ontologies
  - National Research Data Infrastructure and beyond

de.NBI Cloud

- The de.NBI Cloud is the compute and storage infrastructure of the German Network for Bioinformatics Infrastructure.
- It is an academic cloud for research in the broader field of life sciences
- For academic users, its usage is available free of charge

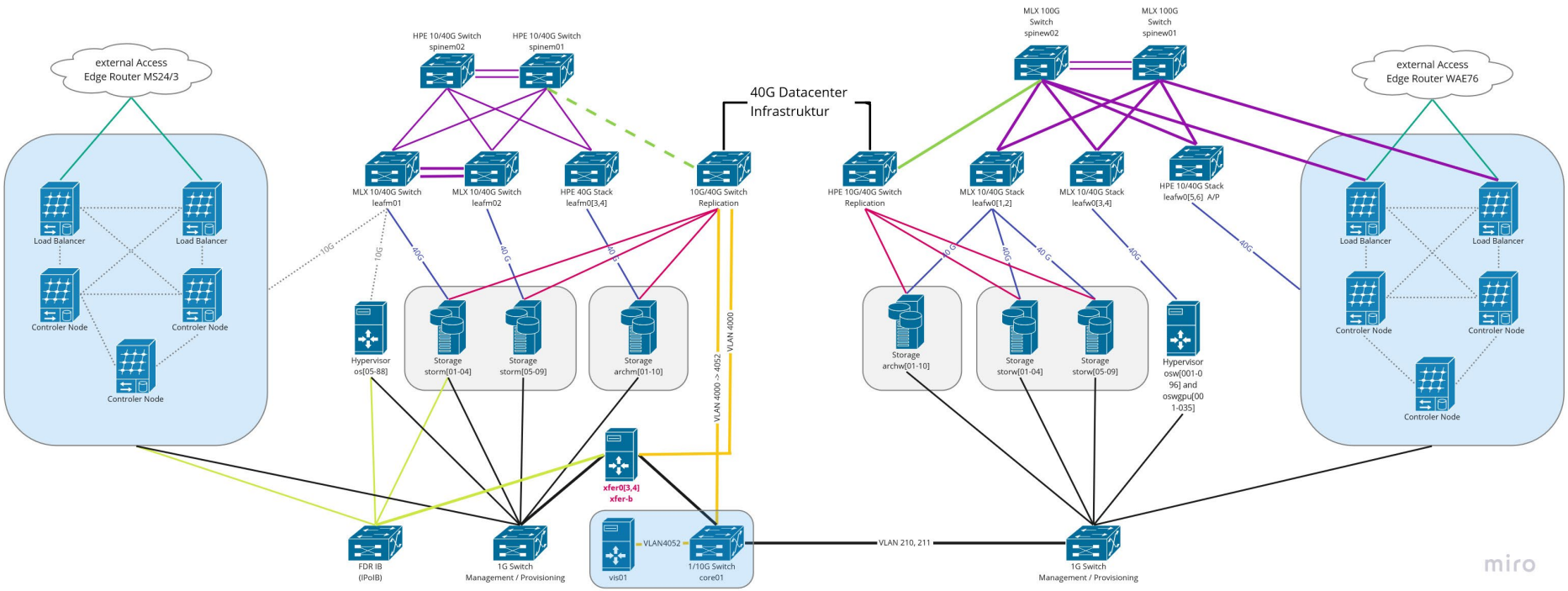
EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

de.NBI Cloud Tübingen

- Tech Specs
  - 6720 CPU cores
  - 118 TByte RAM
  - 180 GPUs
  - 13 PByte Quobyte
  - 14.7 PByte Ceph

- Access policies
  - Open for all German life scientists and beyond
  - ELIXIR AAI / Life Science RI

Topology



EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

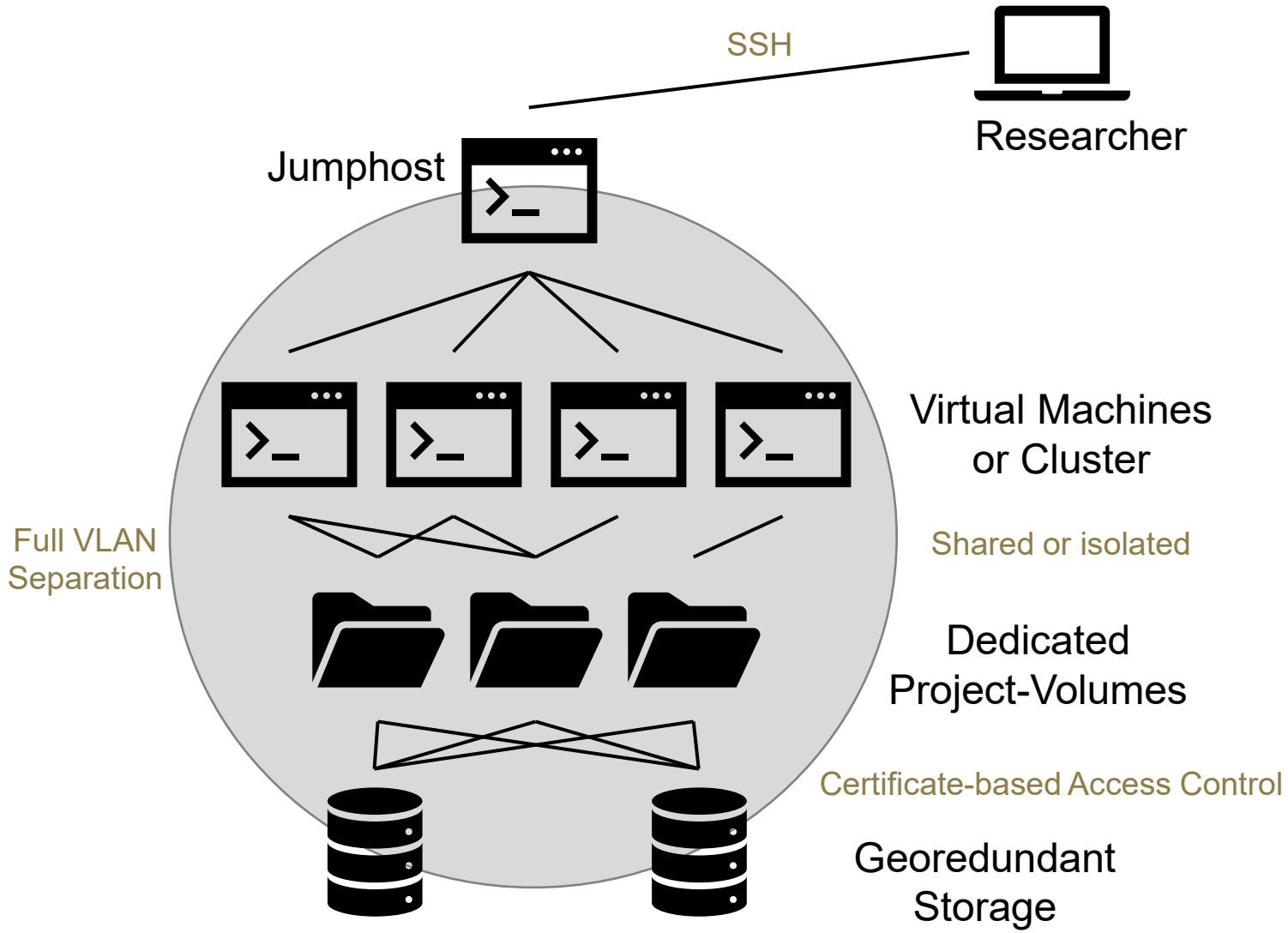
Sustained Storage

- Quobyte and Ceph storage instances
- High level of failure resistance
- Geo-replication over two sites
- Fine grained access control
- Certificate based authentication mechanisms
- Encryption-at-rest

- Long term commitment to support scientific collaborations, research data management projects and associated partners

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Secure processing environments



ISMS and Certification

- The de.NBI Cloud Tübingen has a state-of-the-art Information Security Management System (ISMS)
- A set of security controls and well-trained personnel ensure the confidentiality, availability, and integrity of assets from threats and vulnerabilities

- The cloud is certified according to ISO27001:2017

CERTIFICATE

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

DIN EN ISO/IEC 27001:2017

Technical equivalent to ISO/IEC 27001:2013 including Cor 1:2014 and Cor 2:2015

DEKRA Certification GmbH hereby certifies that the organization

Eberhard Karls Universität Tübingen

Scope of certification:

Operation of the de.NBI Cloud site Tübingen

Certified location:

Wächterstraße 76, 72074 Tübingen, Germany

has successfully implemented the above-mentioned information security management system according to the standard and to the statement of applicability from 18.11.2021 and applies it effectively. The conformity was inspected during the certification audit documented in audit report no. A21091322.

Certificate registration no.: DS-0222009 Certificate valid from: 2022-02-24

Validity of previous certificate: - Certificate valid to: 2025-02-23

Language translation

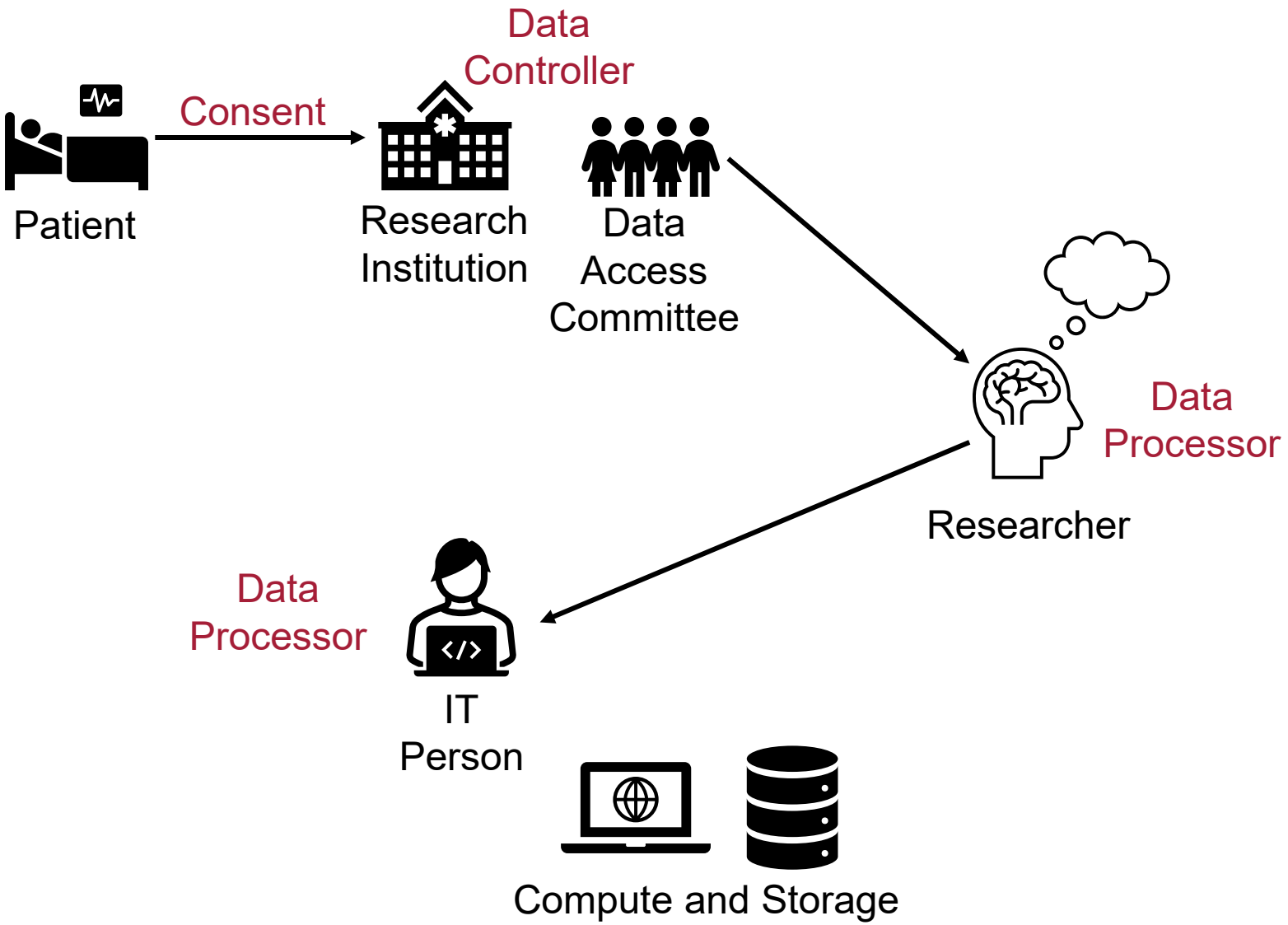
Dr. Gerhard Nagel

DEKRA Certification GmbH, Berlin, 2022-02-24

DEKRA Certification GmbH * Handwerkstraße 15 * D-70565 Stuttgart * www.dekra.de/audits

page 1 of 1

GDPR in a Nutshell



Processing contracts

- The data controller commissions the analysis of human scientific data to a processor. The processor may have further sub-processors.
- If processors or controller belong to different legal entities a processing contract between the institutions has to be made.
- The processing contract contains the purpose of processing e.g. scientific research refereeing to the patient/donator consent.
- The contract regulates the access policies and security requirements.

- Not obeying the law or uncontrolled data leaks may have unpleasant consequences for the institutions involved.

Standard Operating Procedures

- For a given project the handling of its datasets is accompanied by Standard Operating Procedures (SOP) defining the eligible access, the protocols for transfer, the processing workflows, naming conventions and finally decommission.
- These documents are created with the principal investigators involved, making sure that regulations imposed by the data controllers or by law are followed by all people involved.

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Research data management

- Essential for modern data management is the annotation and documentation of research data
- Ideally the whole research data cycle is covered from data collection to secondary reuse
- Research data management should be findability, accessibility, interoperability, and reusability
- Basis for everything is a consistent, consequent and transparent annotation of all data
- Essential for the generation of extra value is the definition of ontologies representing the inherent properties and the relations between them within the data

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Infrastructure, NFDI and beyond

GHGA
- Part of the German National Research Infrastructure
- Archiving and processing of human genome data
- National European Phenome Genome Archive node

Elixir
- European Life Science Infrastructure
- Coordinates national efforts

EOSC-Life
- European Open Science Cloud
- Collaboration on Infrastructure

IRTG 2804 Women's Mental Health
- Swedish/German international research training group
- Broad range of clinic and genome data
- Secure processing within the de.NBI Cloud Tübingen

DataPLANT
- Part of the German National Research Infrastructure
- Archiving and processing of plant genome data
- Federated research infrastructure over multiple sites

Science Data Center BioDATEN
- Research data Management in Baden-Württemberg
- Aiming at diverse research communities from different fields of life sciences

TRR356 Plant-Microbe
- Collaborative research centre
- Federated research infrastructure over multiple sites

… and some more

Acknowledgement
Among many others, support by de.NBI and the High Performance and Cloud Computing Group at the Zentrum für Datenverarbeitung of the Eberhard Karls University of Tübingen is acknowledged.