

TUM Infrastructure for SLICES-DE

Overview

- ▶ **IT Operations** (ITO) of the School of Computation, Information and Technology operates...
 - **Network infrastructure** for the Mathematics/Informatics (MI) building in Garching
 - **Server infrastructure** for research, teaching, and studying
 - an **Autonomous System** with own IP prefixes
- ▶ Virtual Machine Cluster
 - 12 Hypervisors running VMware® ESXi™
 - Highly redundant storage backend
 - Currently around 1500 VMs running on the cluster
- ▶ Tightly and redundantly interconnected network
 - 2 backbone routers and 2 central switches
 - 2×200 Gbit/s interconnect
 - Many downstream switches in server rooms and offices
- ▶ Up to dual 100 Gbit/s to each server possible
- ▶ $2 \times 2 \times 40$ Gbit/s uplink to the LRZ and through them the DFN

Resilient Architecture



- ▶ Every server is connected to two **UPSes** and two independent switches.
- ▶ Critical infrastructure is mirrored in two rooms for **redundancy**.

Special requirements

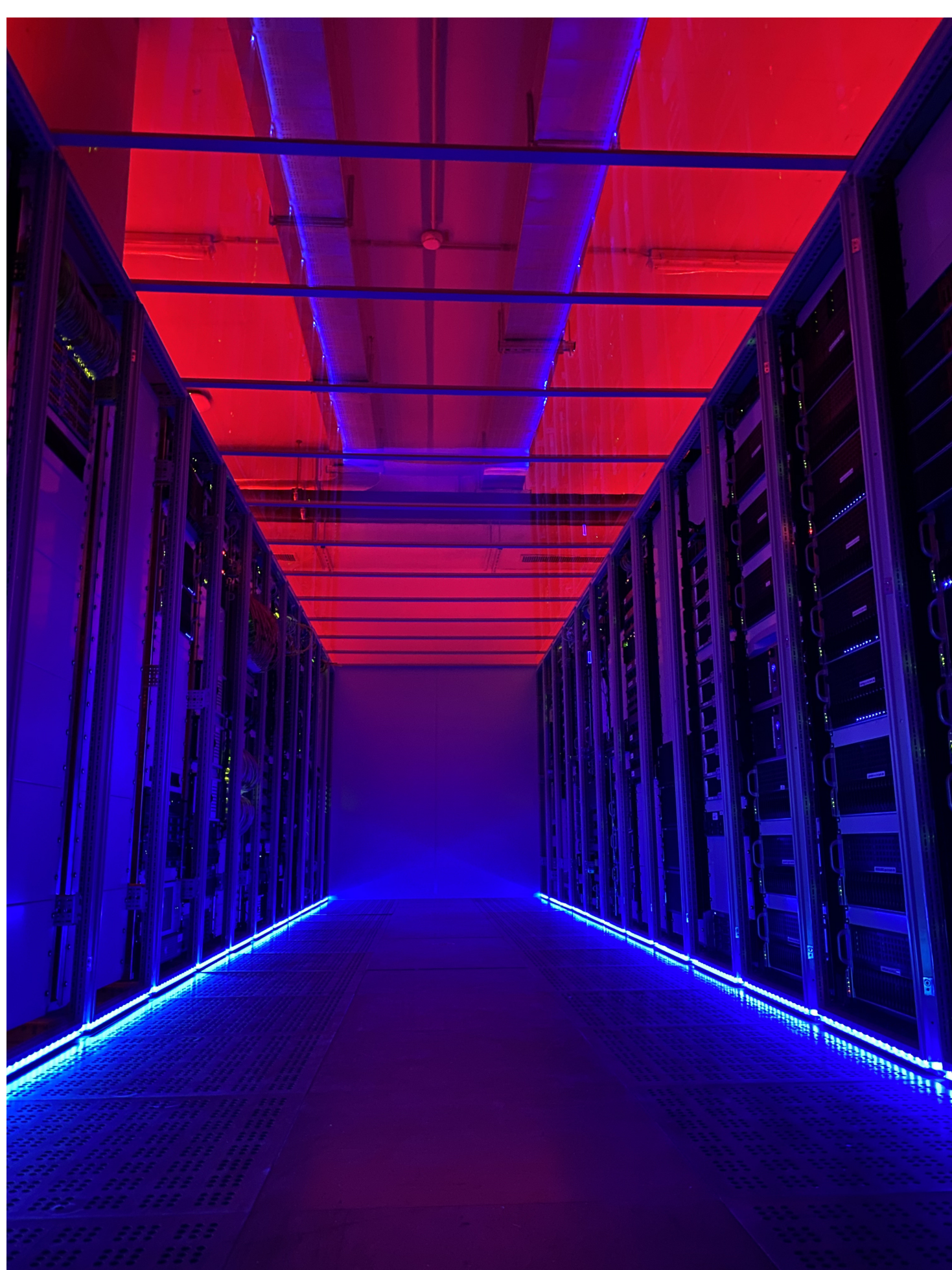
- ▶ As the local IT operator working for and with the chairs the main strength lies in flexibility.
- ▶ Most innovation and changes come from *client* requirements or wishes.
- ▶ Special service: consultation for and procurement (in conformation with the law) and inventORIZATION of hardware
- ▶ Adaptable to new or extraordinary requirements:
 - A chair uploads a lot of data regularly into the Ceph cluster from an adjacent building
 - arranged to provide them a dedicated 10 Gbit/s line for faster uploads
 - Physical rewiring of building-wide fiber-optic cables for 5G Campus-network antennas
 - Restricted access rooms for services with high security requirements, e.g., SAP UCC, TUM TrustCenter, TUMexam, ...

Storage Cluster

- ▶ Distributed Storage System: CEPH [1]
- ▶ Total 10 PB raw capacity
 - 9 PB HDD suitable for cold data
 - 1.32 PB SSD suitable for warm data
 - Erasure Coding renders $\frac{4}{7}$ usable
 - Currently composed of over 102 servers and over 1200 OSDs (drives)
 - Chairs contribute servers and can use the respective (net) quota
- ▶ In simple terms: a distributed key-value store
- ▶ Multiple ways to **access data**:
 - Rados Block Device (RBD): a **block device** to mount and format
 - CephFS: a distributed **filesystem** mountable on multiple hosts
 - Rados Gateway: **S3-compatible** interface to access Ceph data blocks



Server rooms



- ▶ 6 large rooms in MI building
 - 4 for **Co-location** for chairs
 - Both fiber-optic and RJ45 switches with redundant uplinks
 - Up to 100 Gbit/s to the servers possible
 - Central building-wide cooling system set up in a **hot/cold-isle** configuration
- ▶ Smaller, chair-individual rooms in each floor
- ▶ More rooms in the SAP and EI buildings (connected to LRZ network)
- ▶ **Unattended Housing**, i.e., the chair members can access and modify their servers any time they need; only rackspace, power and networking is provided.
 - Experimental setups are possible and encouraged.



- ▶ New building planned:
 - Large server rooms with a total of 1200 kW electrical capacity
 - **In-Row Cooling System** for more efficiency (cooling closer to the heat source)

[1] Sage Weil, Scott Brandt, Ethan Miller, and Darrell Long. Ceph: A scalable, high-performance distributed file system. OSDI '06: Proceedings of the 7th symposium on Operating systems design and implementation, USA, 2006. USENIX Association.