ETH zürich

Scientific IT Services

Celebrating 25 years of cluster shareholder model at ETH Zurich

Olivier Byrde

HPC-CH forum on financial aspects of HPC

ZHAW, Wädenswil, 10 Oct 2025

Table of contents

History of HPC at ETH: from vector processors to multi-site clusters

1990
Central supercomputers (Cray X-MP)

2000
Independent clusters (Asgard, Hreidar, Gonzales, etc.)

2007
Large unified cluster (Brutus)

2014
Very large unified cluster (Euler)

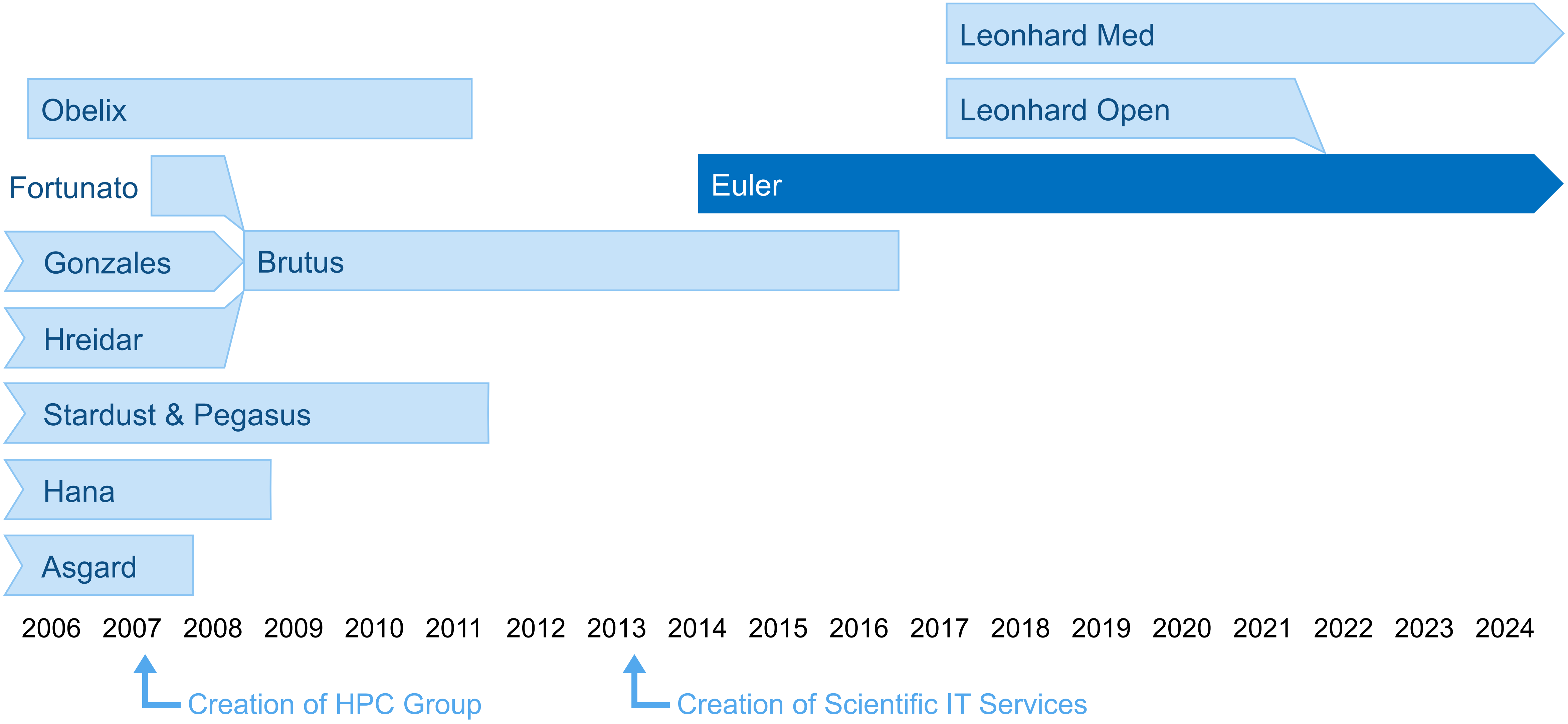
2026+
Next generation HPC

Single processor
1,000 cores
10,000 cores
100,000 cores
1,000,000 cores

History of HPC at ETH: origin of the cluster “shareholder” model

- In the 1980’s and 1990’s, all HPC systems at ETH — Cray X-MP, J90, SV1; HP Superdome, etc. — were paid centrally and made available to all users
  - Access was granted on demand with little scientific oversight; there was no way to make sure that the resources were allocated fairly and used efficiently
  - The size of these systems was limited by the budget of IT Services — and politics!
- In 1994 the first Beowulf cluster opened the door to cheap computing for everyone
  - Many groups at ETH started building their own little cluster
- In 1999, the departments of Physics, Mathematics and Materials of ETH joined forces to procure one large shared cluster called Asgard
  - This cluster, built by the Swiss company Dalco AG, was the largest in Europe
  - The hardware was paid by these departments and operated by IT Services
  - This marked the beginning of the cluster shareholder model at ETH

History of HPC at ETH: standardization & consolidation



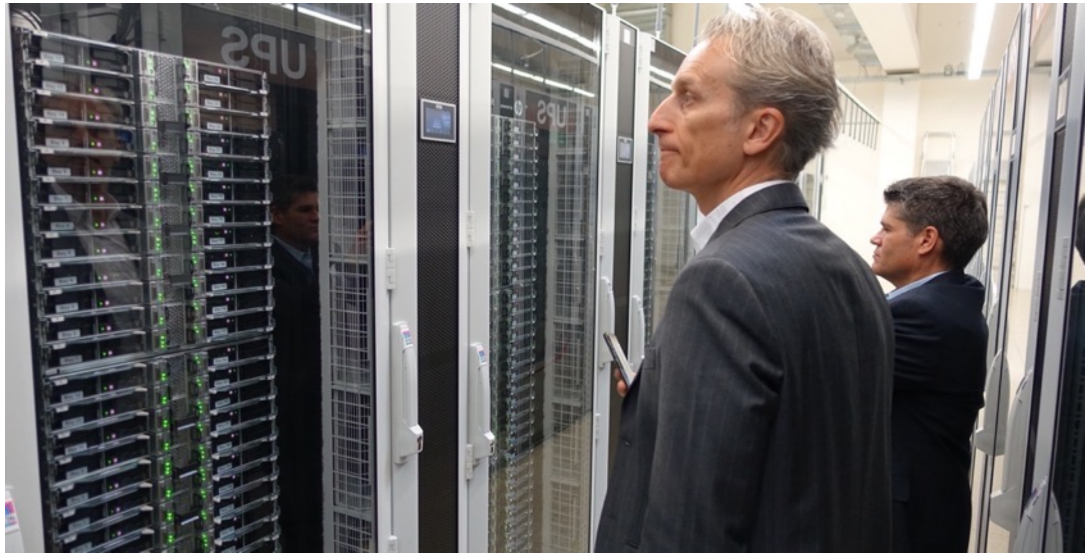
History of HPC at ETH: Euler over the years

Euler I (2014)

Euler II (2015)

Euler III (2017)

Euler IV (2018)



Shareholder model: ownership

- Shareholders are not considered merely as users or customers but as co-owners of the HPC infrastructure
  - To be competitive with local clusters, shareholders only cover the cost of the hardware
  - All operating costs (space, power & cooling, maintenance, system management and user support) are covered by the central IT Services

- In the early clusters, this “ownership” was valid for the lifetime of the underlying hardware
  - This was problematic because different components in a cluster may have different lifetimes, so professors did not know up-front for how long they could use the cluster
  - As you can imagine, they wanted to keep old hardware in operation until it died, even when that was not economically viable

Shareholder model: share validity and hardware life-cycle

- Since the introduction of Euler, a share is no longer tied to hardware but gives shareholders the right to use the corresponding resources anywhere in the cluster
- Advantages:
  - Shares are elastic and allow shareholders to use (far) more resources during peaks
  - Shares can be bought at any time and are valid for 4 years
- As sole owner of the hardware, the IT Services can replace/upgrade it as we see fit
  - We usually expand the cluster once per year to meet the demand of ETH researchers
  - We always buy extra compute nodes to be able to fulfill orders placed by shareholders between major cluster expansions; they are placed in the "public share" of the cluster until they are assigned to a shareholder
  - Likewise, older nodes that are no longer assigned to a shareholder are placed in the public share until they are decommissioned

Shareholder model: standardized cluster service

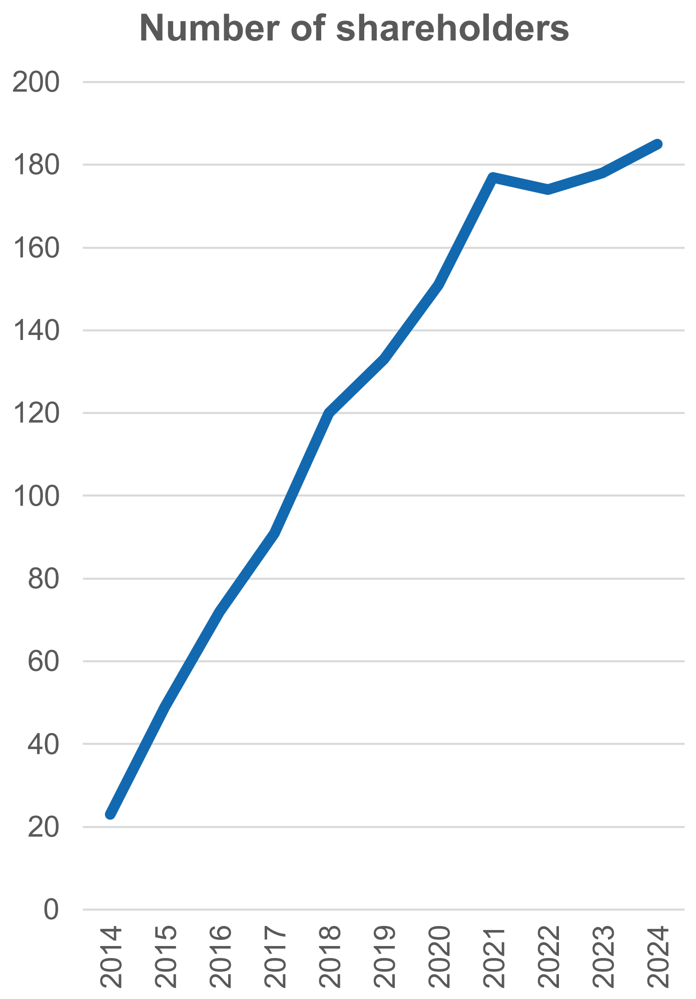
- Professors at ETH have no obligation to buy into Euler
  - They can buy their own cluster, cloud resources, etc. using ETH framework contracts
  - They will only choose Euler if it is competitive in terms of price and service quality
- Likewise, the IT Services have no obligation to fulfill the needs of every professor
  - We do our best to cover the needs of everyone, but we do not have enough manpower to handle special projects (e.g. Windows-based applications)
  - If we cannot satisfy a professor's needs ourselves, we help them find another solution within IT Services (e.g. virtual machines), or at CSCS (Alps) or in the cloud
- To avoid any misunderstanding, all contractual details pertaining to the shareholder model are covered in a binding Service Level Agreement (SLA)
  - This document also contains the specifications and price of all types of compute nodes, GPUs and storage systems that we offer in Euler

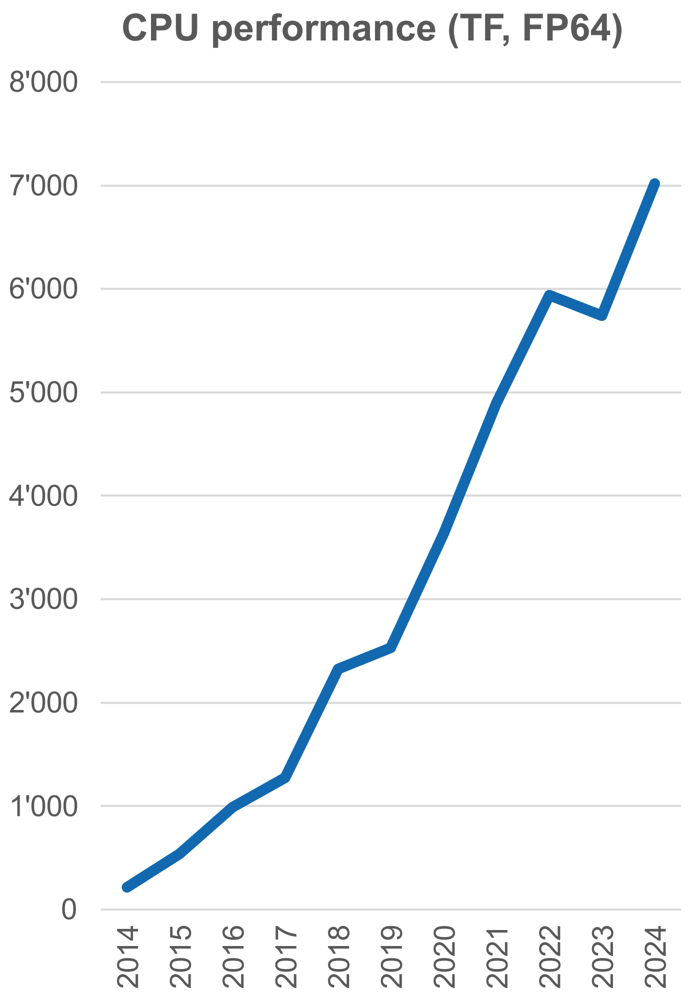
Financial aspects: Euler is all about money

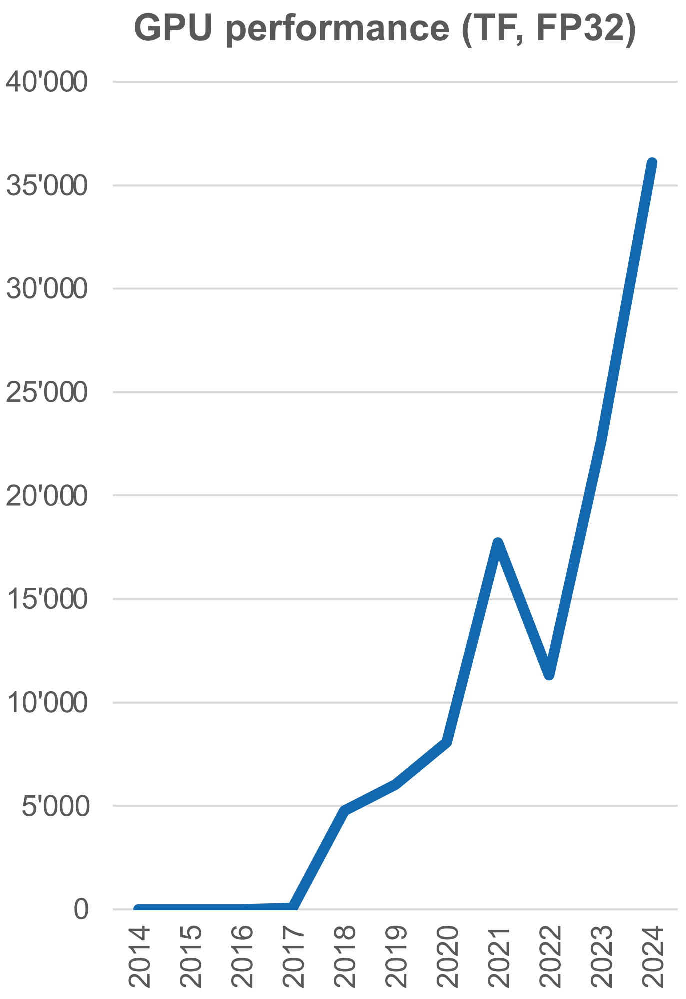
Financial aspects: take the money where it is

- The main limitation of the old central supercomputers was the budget of IT Services
  - The shareholder model removed this limitation by pooling money from researchers
- ETH professors are used to submitting proposals to fund their research
  - Funding agencies (not us!) are responsible for evaluating the scientific value of a proposal
  - Our job is to ensure that every shareholder gets the computing resources they bought
- The standardization of this model (e.g. SLA) led to a rapid increase in the number of shareholders, which in turn fueled the growth of the Brutus cluster, and later, Euler
  - In the last 10 years, shareholders invested over 30 million CHF in the central clusters

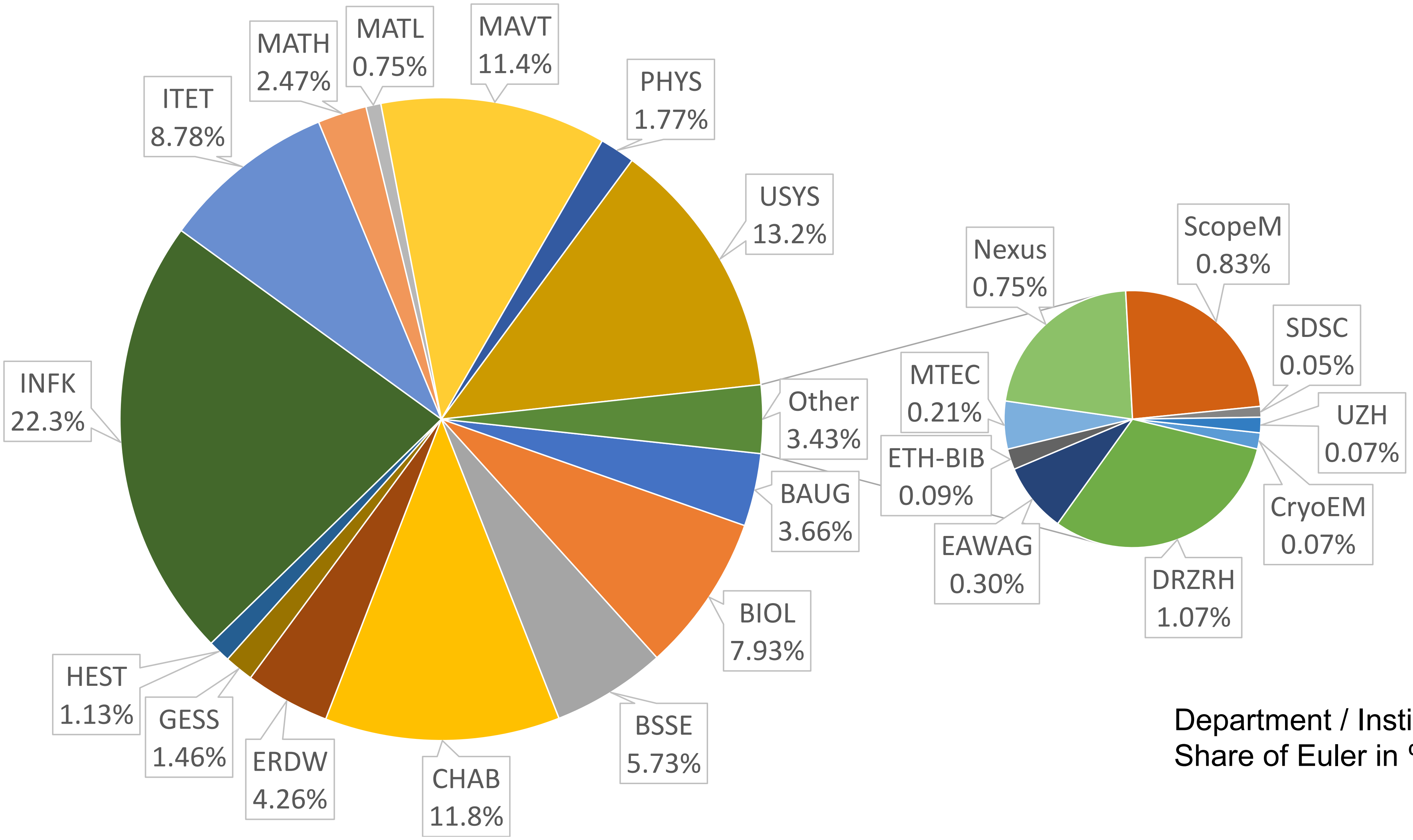
Financial aspects: shareholders fuel cluster growth







Financial aspects: investments by ETH departments (active shares)



Department / Institute
Share of Euler in %

Financial aspects: budget issues

- The financial regulation of ETH does not allow us to transfer money from one year to the next
  - Everything that we buy in a given year must be paid in the same year using money from our own budget and from shareholders
  - Since some hardware has a long (and sometimes unpredictable!) delivery time, we stop accepting orders from shareholders at the end of July, to make sure that we can spend their money before the end of the year
  - If some money cannot be spent on time, e.g. if a delivery is delayed, that money is lost ☹
- Since we need to pay the hardware upon delivery, we also need to charge the full cost to shareholders up-front (and in the same year)
  - We cannot rent resources on a yearly basis, we can only sell resources for 4 years

Financial aspects: procurement process

- According to ETH procurement rules:
  - Orders above 230 kCHF require a public Request For Proposals (RFP) on simap.ch
  - Orders above 500 kCHF require approval by the Executive Board of ETH
- In the “good old days”, we used to issue an RFP for every cluster procurement. That was time-consuming and required a lot of effort ☹
- Since Euler, we do not publish RFPs for individual procurements anymore, but for general framework contracts valid for 5+ years
  - We select the 3-4 vendors who submitted the best proposal for a pre-defined reference configuration, and establish a framework contract with each one
  - Whenever we need to make a large procurement, we invite these vendors to submit an offer (mini-bid)
  - This process prevents vendor lock-in and ensures that we always get competitive prices

Financial aspects: SNSF and ERC grants

- Research groups who get funding from the SNSF or ERC cannot simply invest it into Euler because an Euler share can be categorized neither as equipment nor as consumable

- To solve this problem, the IT Services and the Finance department of ETH have devised a special pay-per-use model for the Euler cluster
  - This model cannot be used to buy into Euler, but allows the owner of an Euler share to reclaim the cost of the CPU and/or GPU time used for a specific project over time
  - The IT Services take care of the usage accounting and billing to the SNSF or ERC

- This model has a high administrative overhead because the CPU and GPU tariffs must be re-calculated yearly based on the effective cost over the previous 3 years, including:
  - Amortization of all eligible hardware in the cluster — for each type of compute node!
  - Resource utilization — also for each type of compute node
  - Power & cooling — ditto

Outlook

- As budgets are getting tight, we expect the growth of Euler — 40% per year — to slow down in the coming years, unless shareholders find new sources of funding
  - The pay-per-use model represents a tiny fraction of the investments in Euler; now that Switzerland has access to ERC funds again, it may become more important in the future
- Given that new CPUs are not getting very much faster/more efficient, we will probably keep older compute nodes in operation longer than in the past
  - This will require more space, power & cooling, so ETH is exploring the option to build a common data center together with PSI in Villigen
- GPUs are getting more expensive each year, and their power requirements are exploding!
  - For now, Alps offers an attractive solution for researchers who need powerful GPUs
  - Will ETH — or CSCS or CH as a whole — be able to afford new GPUs in a few years?

Questions?