CSC

ICT Solutions for Brilliant Minds

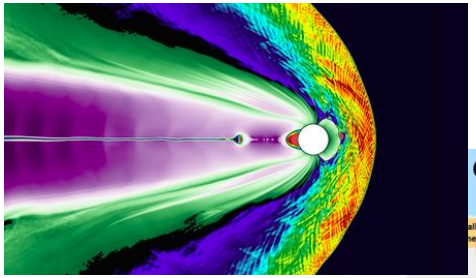
Linux in Supercomputers

Dr. Sebastian von Alfthan

Supercomputers

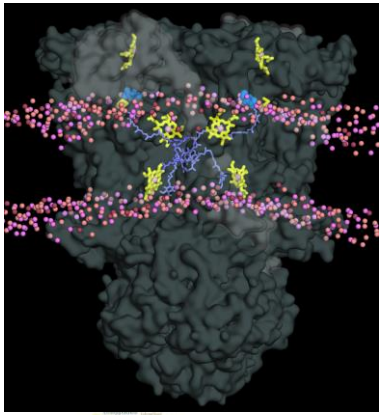
Supercomputers provide huge computational performance and storage capacity for wide range of scientific fields

Supercomputers



Large scale simulations

- For example climate change, space weather, fusion reactors, astronomical phenomena, particle physics



Mid-scale simulations

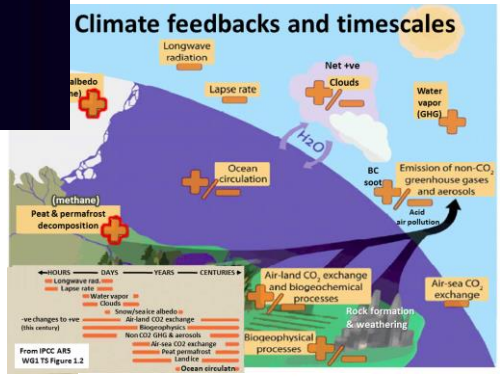
- For example materials science, energy technology, GIS

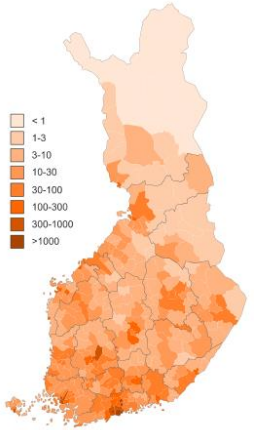
Data-intensive computing

- For example computational econometrics, bioinformatics, language research
- Also developing solutions for sensitive data

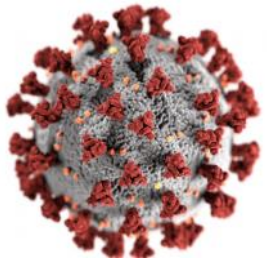
Artificial intelligence

- For example natural language research, business applications, computer vision





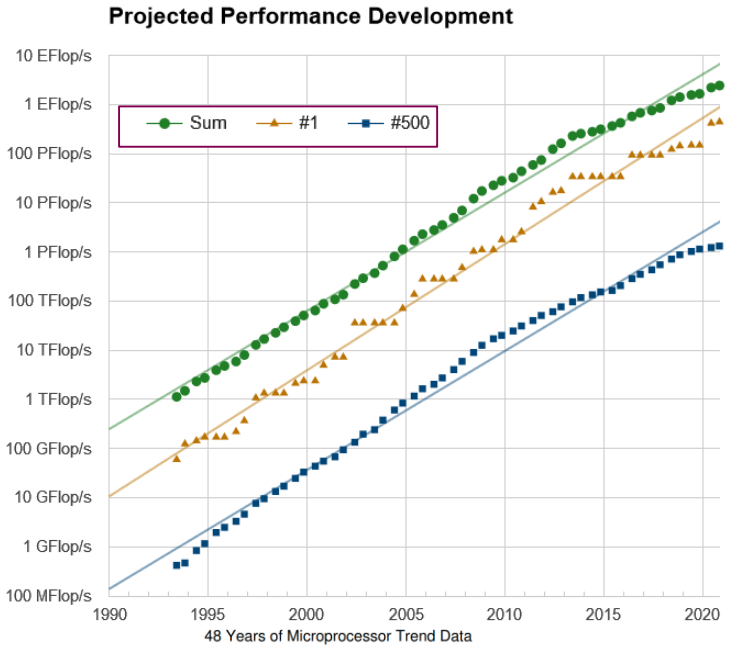
Real-world examples – covid research



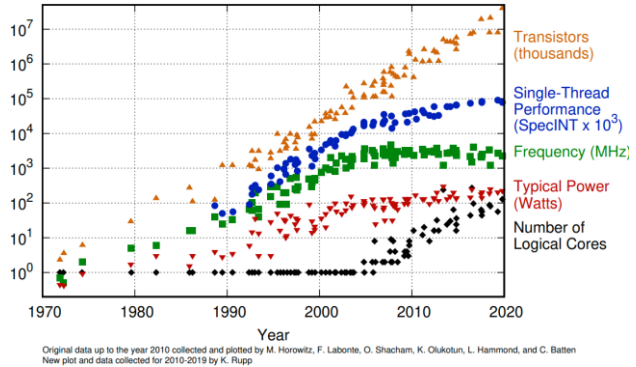
- Large computational resources were a vital tool in the fight against COVID
- CSC and other computing center prioritized COVID research and provided fast track access
- Spreading of aerosol particles in air
  - Computational fluid dynamics modelling of airborne transmission of coronavirus
  - Medium scale simulations with few hundred CPU cores
  - PI Ville Vuorinen, Aalto University

Supercomputer development

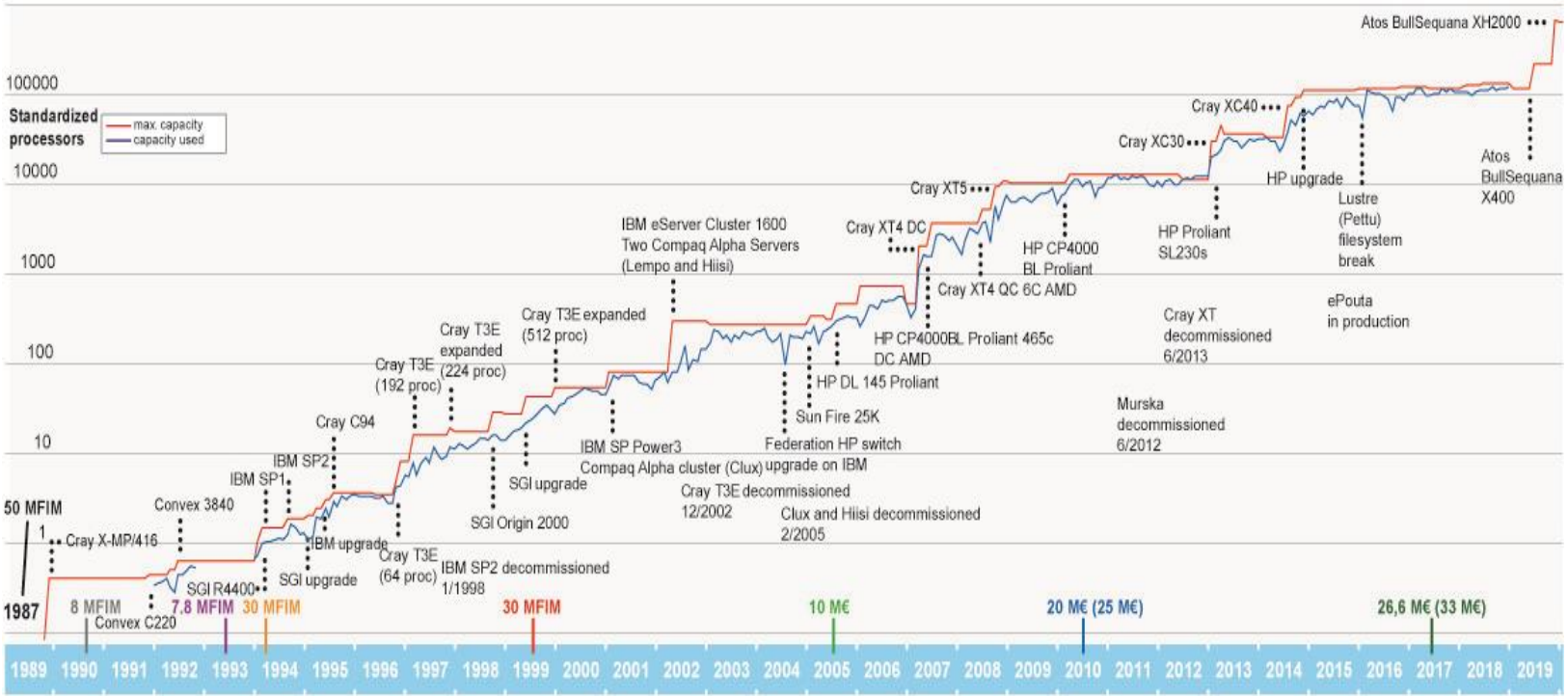
- The one constant theme has been an exponential increase in performance
- Top500 list has tracked the performance of 500 fastest systems since early 90's
- Over the years the architecture has changed significantly – finding performance wherever possible
  - Pipelines – multiprocessing – vector processors – massively parallel systems - accelerators



https://top500.org/



Supercomputer development in Finland



CSC and Linux release

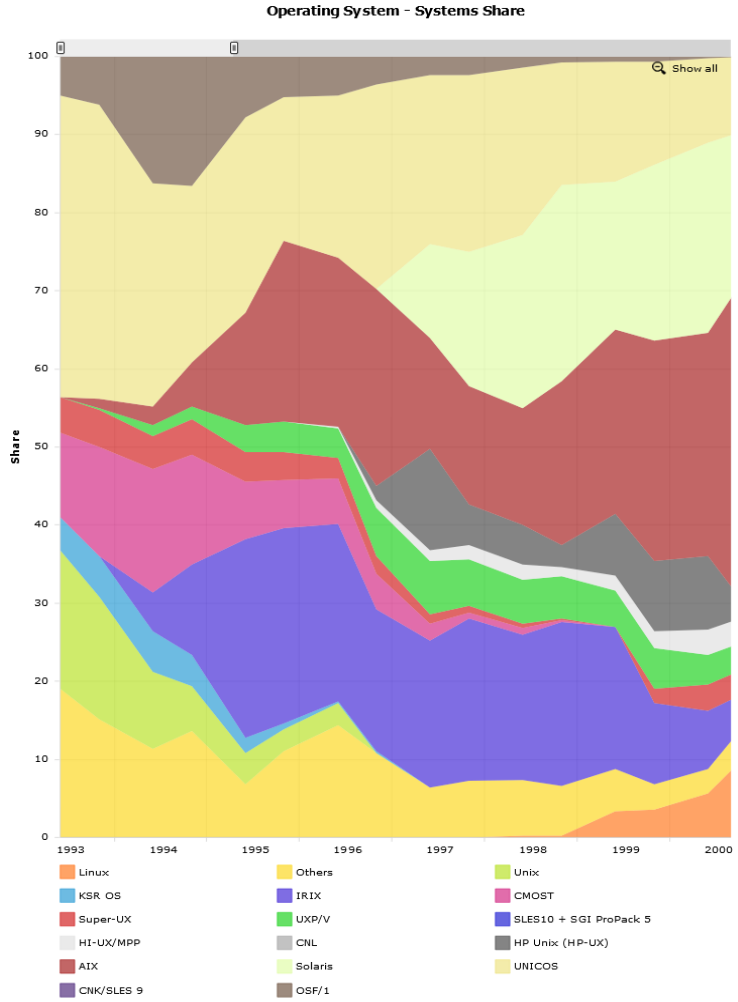
- In 1990 CSC set up one of the premiere file sharing servers of its time – nic.funet.fi also known as ftp.funet.fi
- In 1991 a student from Helsinki University wanted to host his new minix compatitable os at ftp.funet.fi
  - Ari Lemmke who was managing /pub/OS/ did not like original name Freax – since then known as Linux
  - **Linus Torvalds** releases 17.9.1991 version 0.0.1 of Linux at ftp.funet.fi
  - And still there: https://ftp.funet.fi/pub/Linux/historical/kernel/old-versions/

FUNET

by CSC

Linux and supercomputing

- Since the beginning supercomputers had been specialized systems
  - Custom CPU
  - Custom system architecture
  - Custom OS
- Slowly things were changing
  - Custom CPUs being replaced by workstation CPUs when transitioning from vector CPUs to first massively parallel supercomputers
  - More standardized parallel programming models - MPI standard released in 1994
  - UNIX dominant OS in 1990s



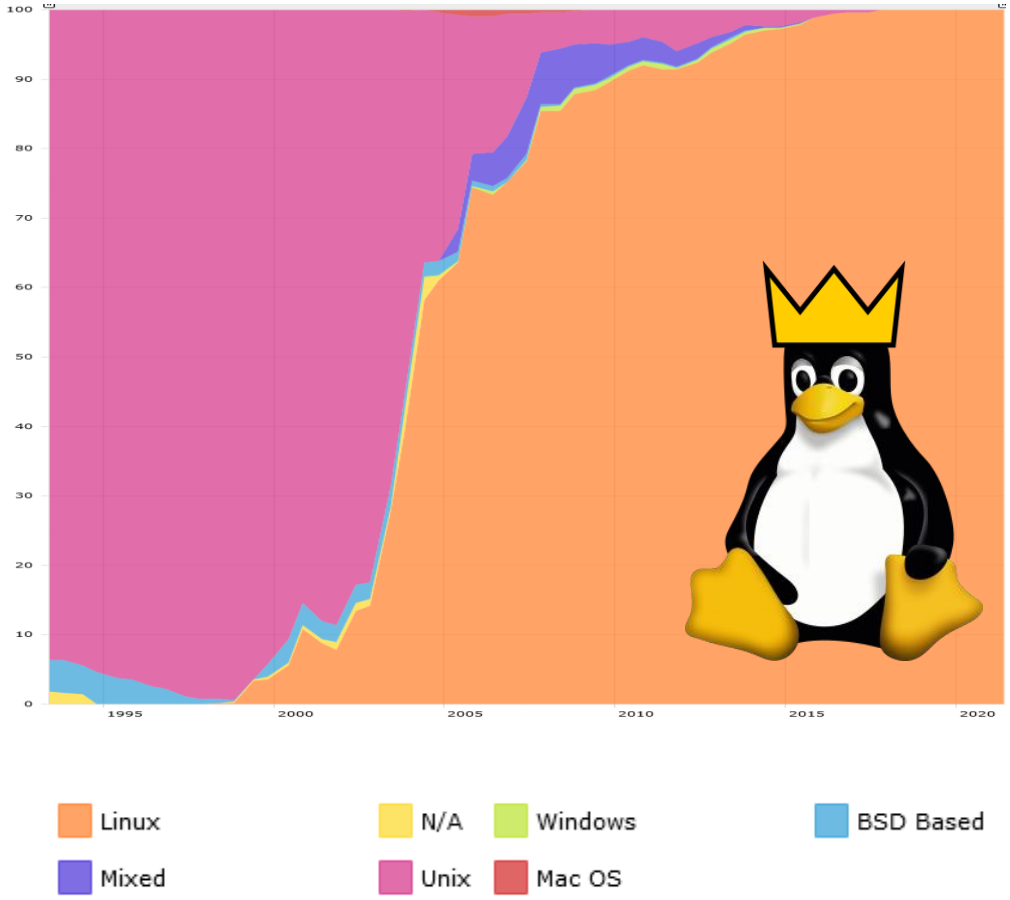
Beowulf – the next revolution

- In early 1994 the Beowulf project was initiated at NASA
  - commodity-based cluster system designed as a cost-effective alternative to large supercomputers
- Many similar projects initiated, also at CSC – cost much lower than traditional Supercomputers
- Enablers
  - Large PC market and commodity off the shelf components available (processors, network cards, …)
  - Full stack of open source software available: GPU compilers, MPI library, and …. **Linux**
- Needed an open source OS to enable the model

Avalon

Linux breakthrough

- Also big supercomputer vendors picked up on Beowulf and Linux
- 2008 Linux took over #1 position (Roadrunner)
- Since 2017 all 500 fastest supercomputers run linux
- Overall supercomputing software became to rely almost completely on open source software



Linux breakthrough

- First Linux based supercomputers at CSC in 2005-2007
  - Sepeli Cluster 2005 – RHEL 4
  - Murska Cluster
  - Louhi Cray XT4
- Overall supercomputing software became to rely almost completely on open source software

Why linux?

User experience

- Linux felt largely familiar to UNIX users
- Linux is de facto standard in supercomputing and cloud
- You can run a supercomputer OS in your laptop

Performance

- Vendors can modify and tweak open source Linux kernel to achieve best performance – OS still very performant
- Vendors can develop kernel support for prototype hardware – best initial support in Linux

Why linux?

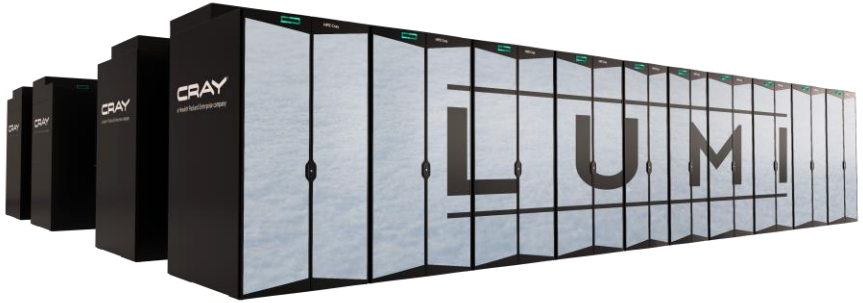
Cost

- Licensing costs lower
- OS development costs lower

Flexibility

- Can combine different hardware – no vendor lock in
- Recruitment - easier to find Linux experts
- Can build even from scratch open source based services
  - CSC's cPouta Cloud, Taito supercomputer, …

CSC's current systems – All running Linux



1 system

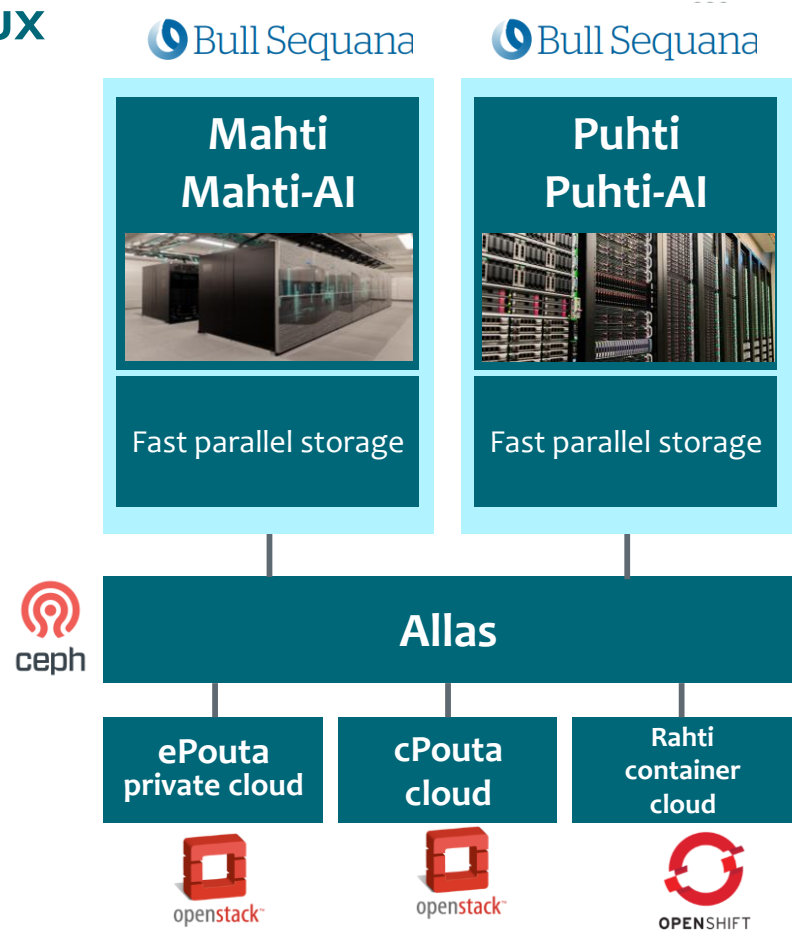
550+ PFLOP/s

Peak Performance

Computing power equivalent to

1 500 000

Modern laptop computers



Sebastian von Alfthan (Ph.d)

Development Manager,
Computing environments,
CSC

Keilaranta 14 (PL 405),
02101 Espoo
+358 40 5888 688
Sebastian.von.alfthan@csc.fi

facebook.com/CSCfi

twitter.com/CSCfi

youtube.com/CSCfi

linkedin.com/company/csc---it-center-for-science

github.com/CSCfi