Data storage and services, data centers and networks

Pirjo-Leena Forsström

PhD, Director

CSC

Index

- European Data Infrastructure
- CSC and Finland
- Data Center
- Networking
- Data Services

EUDAT

Towards a European Collaborative Data Infrastructure

CSC – Tieteen tietotekniikan keskus Oy
CSC – IT Center for Science Ltd.

Sharing and federating Scientific Data

e-Infrastructures Vision

empower research communities through ubiquitous, trusted and easy access to services for data, computation, communication and collaborative work

Source: European Commission

History of the EUDAT concept

EUDAT has its origins in the work of the **PARADE** (Partnership for Accessing Data in Europe) initiative

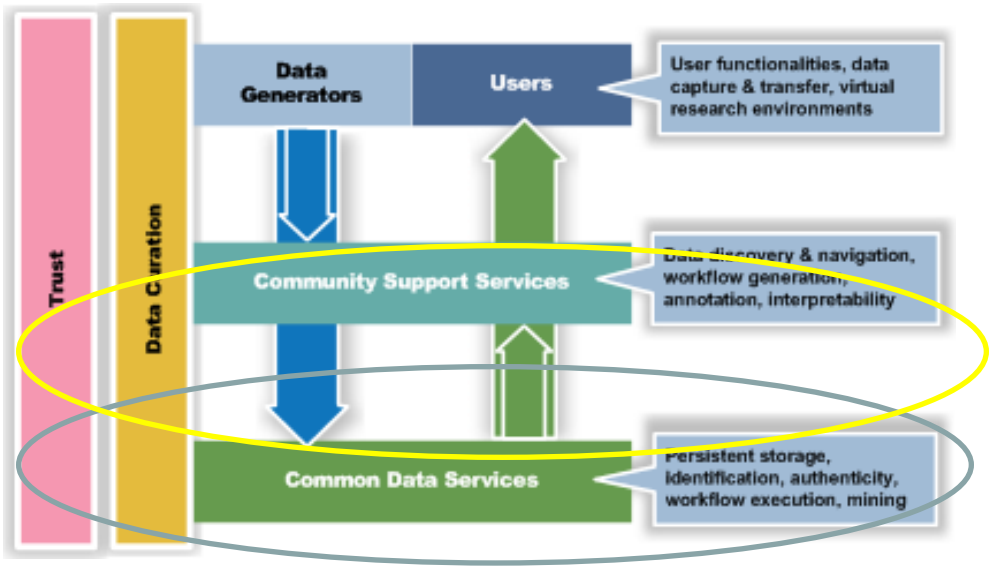
- Early consortium of 13 European partners (key service providers with expertise in HPC, grid computing, networking and data storage)
- PARADE White Paper (October 2009) defining a Strategy for a European Data Infrastructure that should be persistent, multidisciplinary, and based on the need of user communities

The concept of a shared pan-European infrastructure was then supported and further elaborated by a number of policy and experts bodies:

- **e-IRG** and **ESFRI**: e-IRG Blue Paper (September 2010) recommends ”to identify and promote common (long term) data related services across different RI”
- **High Level Expert Group** (**HLEG**) report on Scientific Data: ”Riding the Wave: How Europe can gain from the rising tide of scientific data” (October 2010) calls for a ”Collaborative Data Infrastructure” for scientific data, that supports seamless access, use, reuse, and trust of data.

EUDAT will build on these different initiatives to set up this CDI.

Towards a Collaborative Data Infrastructure



Source: HLEG report, p. 31

Data generators and users gather, capture, transfer and process data in virtual research environment.

They draw upon support services in their specific scientific communities (community infrastructure layer)

These communities draw on a broad set of common data services that cut across the global system (generic data infrastructure layer)

- EUDAT will focus on building this generic data infrastructure layer and offer a trusted domain for long term data preservation accompanied with related services to store, identify, authenticate and mine these data.

- Community Support Services can be incorporated into the common data service infrastructure when they are of use to other communities.

EUDAT core services

Core services are building blocks of EUDAT's Common Data Infrastructure

mainly included on bottom layer of data services

Fundamental Core Services

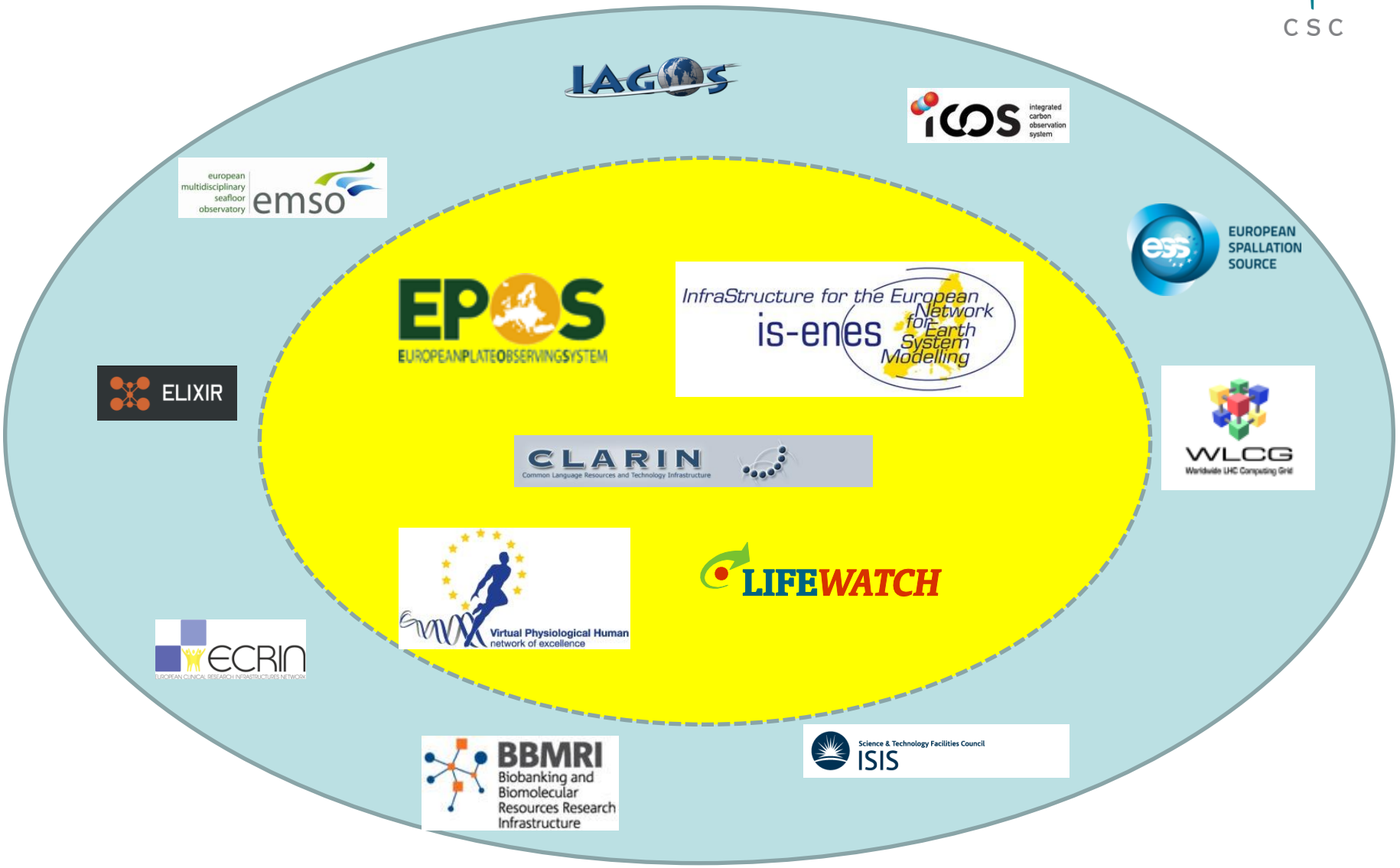
- Single Sign On (federated AAI)
- Data access and upload
- Long-term preservation
- Persistent identifier service
- Workspaces
- Web execution and workflow services
- Network services
- Monitoring and accounting services

Extended Core Services (community-supported)

- Joint meta data service
- Joint data mining service

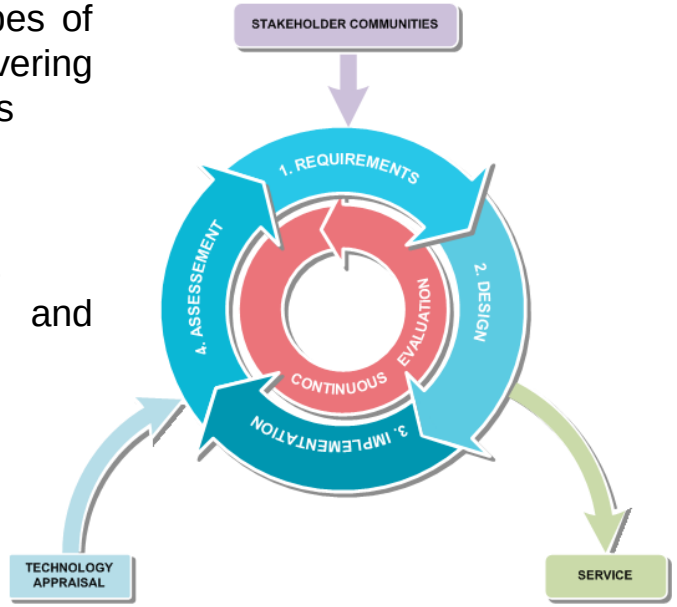
The EUDAT Consortium

The EUDAT Communities



EUDAT Services Activities – Iterative Design

EUDAT's Services activity is concerned with identification of the types of data services needed by the European research communities, delivering them through a federated data infrastructure and supporting their users



1. Capturing Communities Requirements (WP4)

- Services to be deployed must be based on user communities needs
- Strong engagement and collaboration with user communities and innovative ways to capture requirements

2. Building the services (WP5)

- User requirements must be matched with available technologies
- Need to identify:
  - available technologies and tools to develop the required services (technology appraisal)
  - gaps and market failures that should be addressed by EUDAT research activities
- Services must be designed, built and tested in a pre-production test bed environment and made available to WP4 for evaluation by their users

3. Deploying the services and operating the federated infrastructure (WP6)

- Services must be deployed on the EUDAT infrastructure and made available to users, with interfaces for cross-site, cross-community operation
- Reliability, 24h/7d availability and accessibility of the shared services, with operational security, data integrity and compliance with stakeholder requirements and policies.

EUDAT Service Strategy

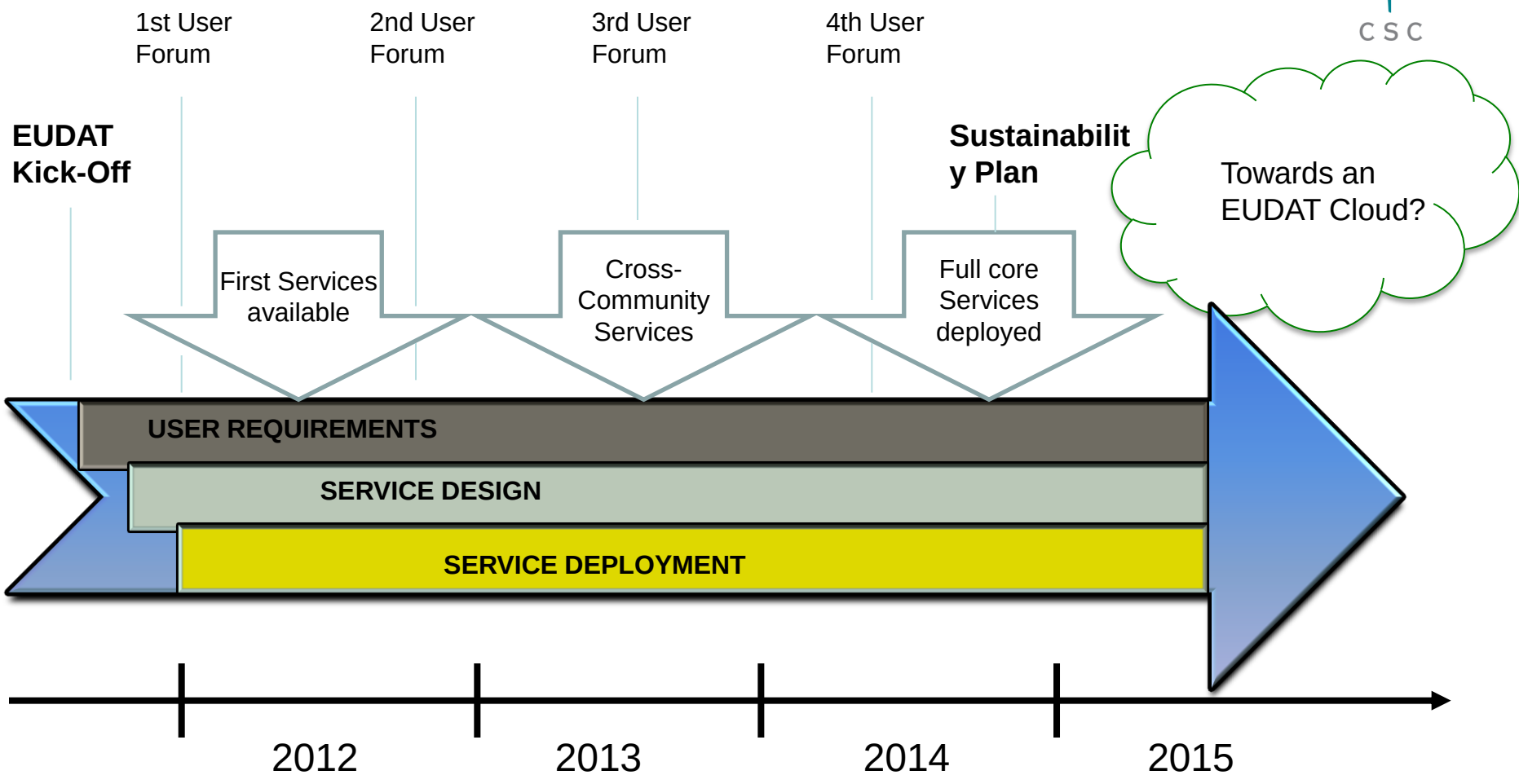
Target: "To deliver a Collaborative Data Infrastructure (CDI) with the capacity and capability for meeting researchers' needs in a flexible and sustainable way, across geographical and disciplinary boundaries."

- The infrastructure must be **Collaborative**
- The infrastructure must be driven by **researchers' needs**
- The infrastructure must be **sustainable** yet **flexible**
- The infrastructure must be **pan-European**
- The infrastructure must be **multi-disciplinary**

Key benefits

- Enabling innovative **multi-disciplinary** data intensive research
  - Development of common services supporting multiple research communities
  - Inter-disciplinary collaboration and exploitation of synergies between communities
- Ensuring wide access to and **preservation** of data in a **sustainable** way
  - A robust generic infrastructure capable of handling the scale and complexity of data that will be generated over the next 10-20 years
  - Put Europe in a competitive position for important data repositories of world-wide relevance
- Economies of scale and **cost-efficiency**
  - Shared resources and work are less costly

EUDAT Timeline



CSC – Finnish IT Center for Science

- Founded in 1971 - Operates on a non-profit principle (shares owned by the Ministry of Education and Culture
- Staff 200 – Turnover 21,9 M€
- CSC provides high performance computing (HPC) and data services for more than 3000 researchers in Finland and national organisations
  - Computing services
  - Data services for Science and Culture (Long term data preservation, data management, etc.)
  - Application Services (Scientific software and databases)
  - Information Management Services
  - Network Services (FUNET – Finnish Research and Education Network)
- Sustainable ICT, data research environments and cloud computing are strategic goals

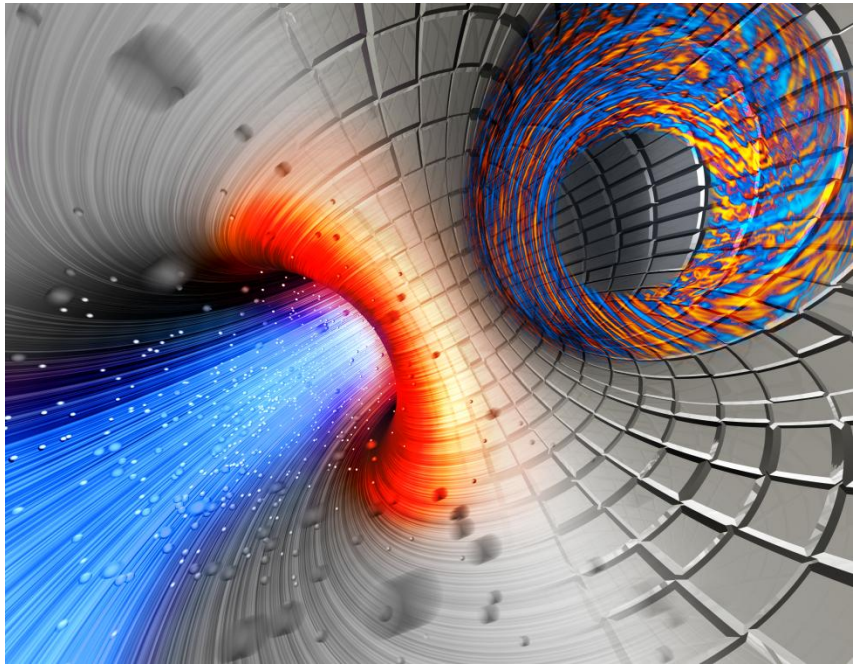
CSC's new Green Flagship Data Center

CSC – Kajaani Data Center

- Built in an old paper mill by the river Kajaani
- Three Hydropower stations on site (30MW)
- Free cooling (air + water) all year
- Space: 8000m^2 (4000m^2 + 4000m^2)
- 25M€ Investment
  - 20 M€ Computing (new Supercomputer + Cluster)
  - 5 M€ Storage and data solutions
- Installation phase: 2011-2012

Mission

- CSC, as part of the Finnish national research structure, develops and offers high quality information technology services



Customers

- 3000 researchers use CSC's computing capacity
- Funet connects about 80 organizations to the global research networking infrastructure
  - universities
  - polytechnics
  - 35 industrial clients and research institutions
  - Total of 350 000 end users

CSC's services

Funet Services

Computing Services

Application Services

Data Services for Science and Culture

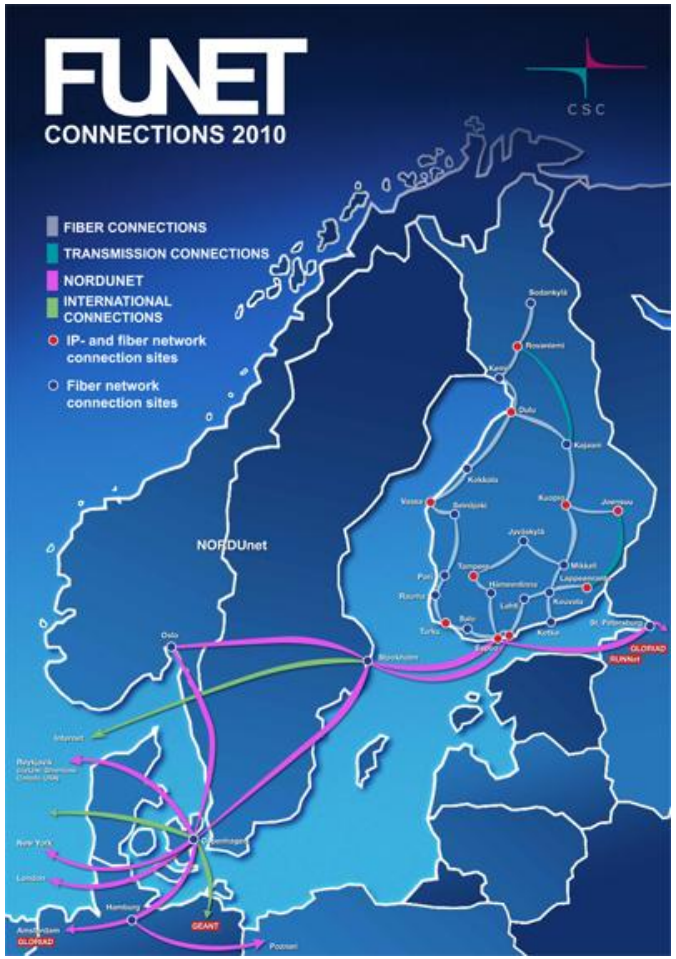
Information Management Services

CSC's CrayXT4/XT5

CRAY XT4/XT5 alias Louhi

- 2356 AMD Quad Opteron 2,3 GHz CPUs
- 10864 cores
- Memory ~ 11,7 TB
- Theoretical computing power 100 teraflop/s

Funet Backbone Network



- Funet backbone provides reliable and high-capacity connections for all Funet member organizations in Finland. Funet is connected to international academic networks via NORDUnet.
- Funet backbone supports advanced services like IPv6 and IP multicast. Link speeds range up to 10 Gbps.
- Since spring 2009, light paths, dedicated high-capacity links for special applications and users, have been available in many locations.

International collaboration

- Computing centers
- International research network organizations:
  - NORDUnet, TERENA, Internet2, Dante (Géant2)
- International science network organizations:
  - European Molecular Biology Network (EMBnet), EMBRACE
- Nordic and European HPC projects and GRID-organizations:
  - Nordic Data Grid Facility, NorduGrid, DEISA2, EGEE-III, NEG, ESO, Sirene, PRACE, EGI
- CSC chairing: TERENA, E-IRG, EGI, NORDUnet, PRACE (vice-chair)

THE FINNISH GREEN FLAGSHIP DATA CENTER

CSC — IT Center for Science Ltd. Is building one of the most eco-efficient datacenters in the world.

- Boreal climate is an advantage
- Natural source of cooling:
  - Cool climate
  - Abundance of cool water
  - Abundance of hydro power, 3 plants on site

Green Computing: a ”hot” topic

Energy efficiency is at the heart of the EU’s effort to tackle the problems of climate change (global warming) (EC Com 2009)

Europe’s targets for 2020:

- saving 20% of primary energy consumption;
- reducing greenhouse gas emissions by 20%;
- raising the share of renewable energy to 20%.

- ICT sector is responsible for 2% of global carbon emissions (Gartner report, 2007)

Examples of Policies & Tools

2005
Green Paper on Energy Efficiency: Set the goal of reducing 20% of energy consumption by 2020.

2006
April: EC DIRECTIVE on *Energy end-use efficiency and energy services*: Calls for "National Energy efficiency Action plans from Member States"

2008
May: EC Communication on *Addressing the challenge of energy efficiency through ICTs*: Recognition of the potential of ICTs to provide a cost-effective means of improving energy efficiency across industry and broader civil society.

2009
March: EC Communication on *Mobilizing ICTs to facilitate the transition to an energy-efficient, low-carbon economy:*

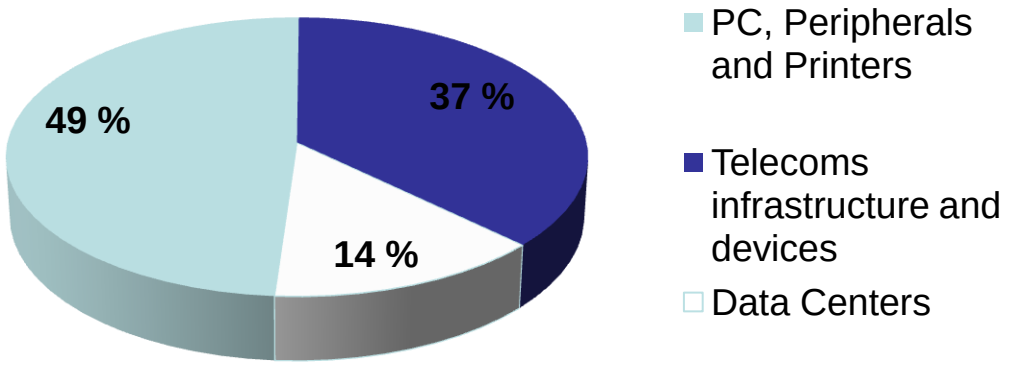
- ICT can enable energy efficiency improvements by reducing the amount of energy required to deliver a given service (by monitoring and directly managing energy consumption; by providing the tools for more energy-efficient business models, working practices and lifestyles; by delivering innovative technologies, etc.)
- ICT can provide the quantitative basis on which energy-efficient strategies can be devised, implemented and evaluated (quantity energy consumption, measuring energy performance at a system level, etc.)
- ICT sector must measure and quantify the energy performance of its own process
- ICT sector must reduce the energy consumption of its own process (including operations, manufacturing, service delivery and the supply chain)

October: EC recommendation on *Mobilizing ICTs to facilitate the transition to an energy-efficient, low-carbon economy*. Sets a few recommendations for the ICT sector and MS to facilitate the transition to an energy efficient and low-carbon economy

- FP7: ICT Work Programme (ICT for a low carbon economy) & Research Infrastructures WP (data infrastructures)
- Cohesion Policy 2007-2013: Funds to support the development of ICT solutions that improve energy performance

EU's European cloud strategy (Kroes) March 27, 2011
define global standards in cloud computing, "aim high" and help Brussels forge a truly European Cloud computing strategy. Data protection, interoperability and standardization are three of the core objectives

Source of carbon emissions by ICT subsectors (%)



Source: Smart 2020 report

• **PCs** (workstations, desktops and laptops) will continue to remain the largest pollutants in the future, despite technological improvements, due to the increasing demand in the developing world.

• **Telecoms** infrastructure and devices demand will also increase significantly, but their relative contribution to the global footprint is expected to decrease due to power consumption reduction from smart chargers and stand-by modes.

• Despite first-generation virtualization and other efficiency measures, **data centers share will grow faster than any other ICT technology, driven by the need for storage, computing and other IT services.**

The case for greener data centers

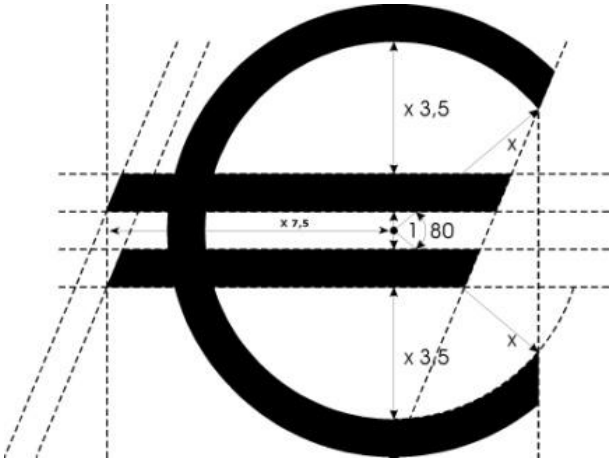
• Due to their enormous energy consumption, data centers have a large and growing carbon footprint;

• There is much room for energy efficiency improvements in data centers that can reduce this footprint;

• Virtualization and cloud computing are major trends driving down the global footprint but ultimately rely on big data centers;

•Data centers are an area where improved environmental performance can result in significant costs savings.

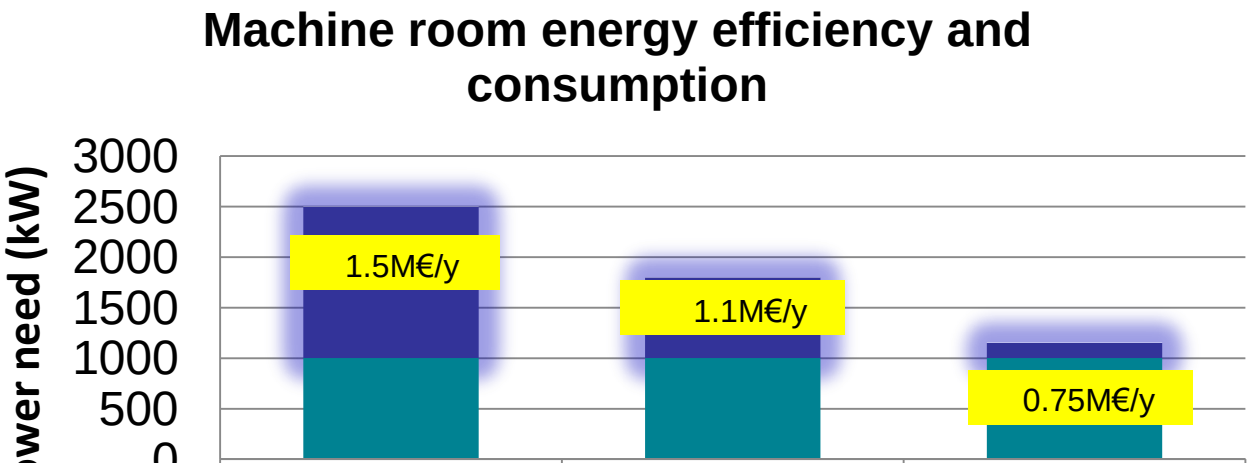
Datacenter customers want cost savings AND quality, green a plus



- Strong preference over cost efficiency in governmental and commercial customers
  - Total costs: energy, data communications, rent, services to be optimized

- Green IT is a plus but not the most important purchase criteria for many customers
- Luckily, energy efficiency improvements lead to savings

Focus on improving energy efficiency => saving costs



	90's (PUE 2.5)	2000-luku (PUE 1.8)	2010 (PUE 1.15)
PUE	2,5	1,8	1,15
Waste	1500	800	150
IT	1000	1000	1000

$$PUE = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

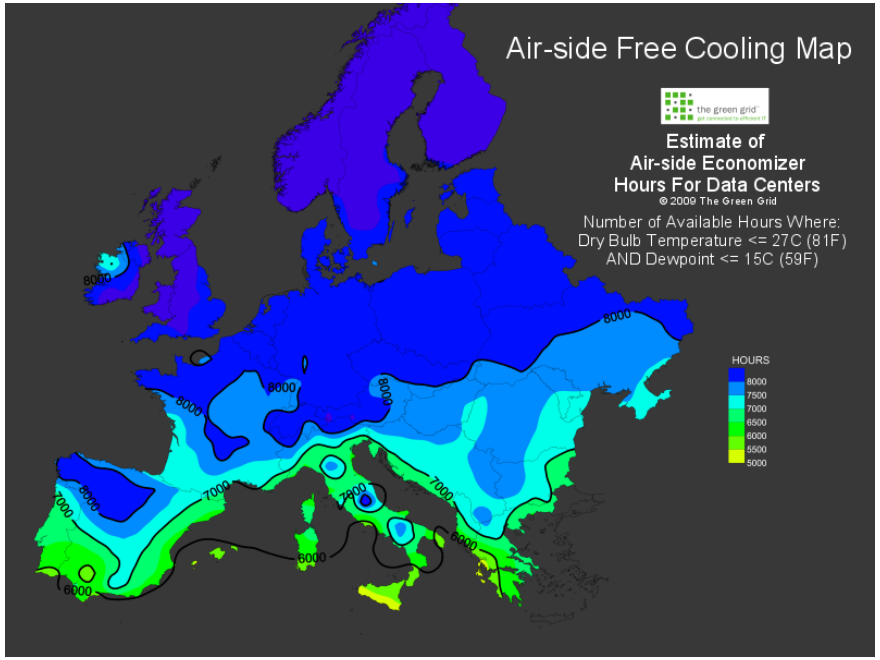
Power Usage Effectiveness (PUE)

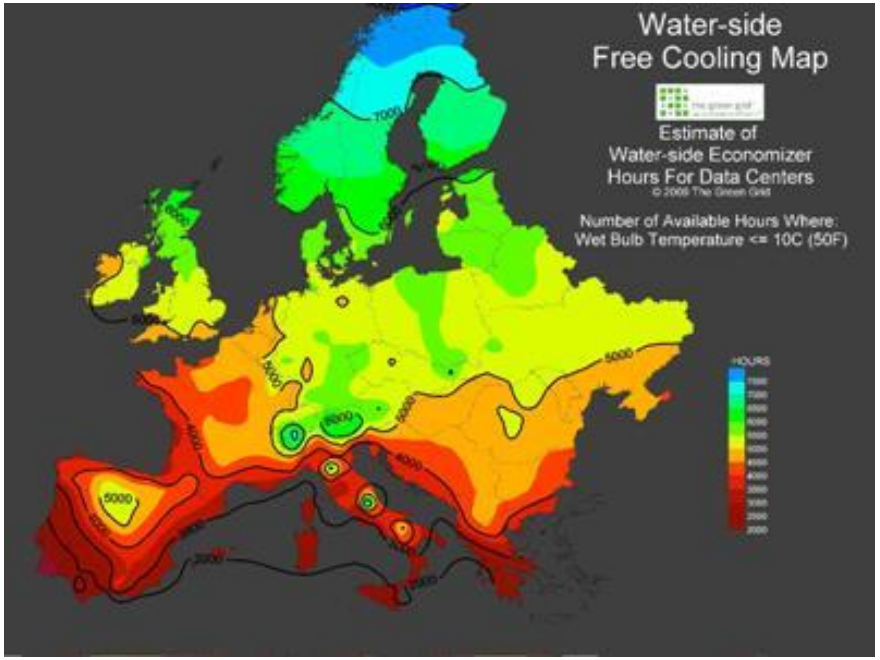
Average energy consumption and costs of running operations for 1MW IT equipment according to different PUE (based on Finnish Electricity costs for Industry)

Case FINLAND

- Safe geological location
- Lot of cooling water available
- Optimal climate: Free cooling potential up to 365 days per year
- Extremely stable electricity distribution grid

How to improve energy efficiency? Take advantage of the climate using external free cooling

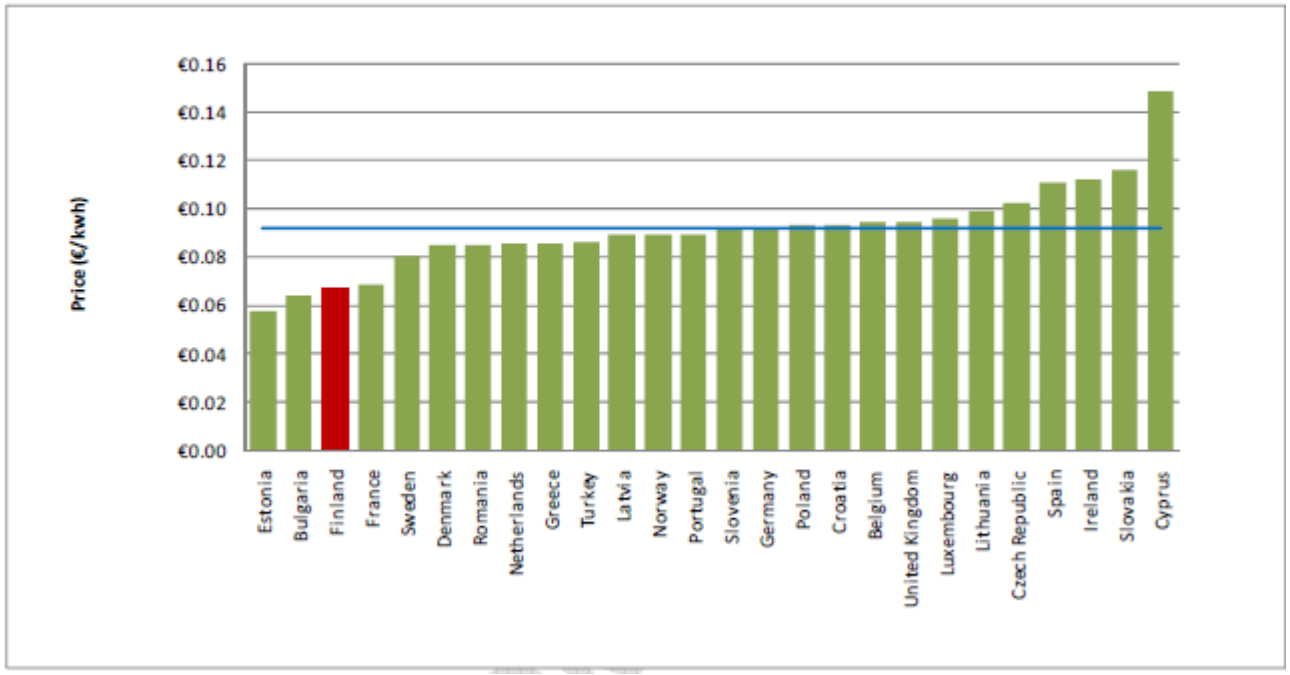




Air vs Water cooling Maps by Green Grid 2009

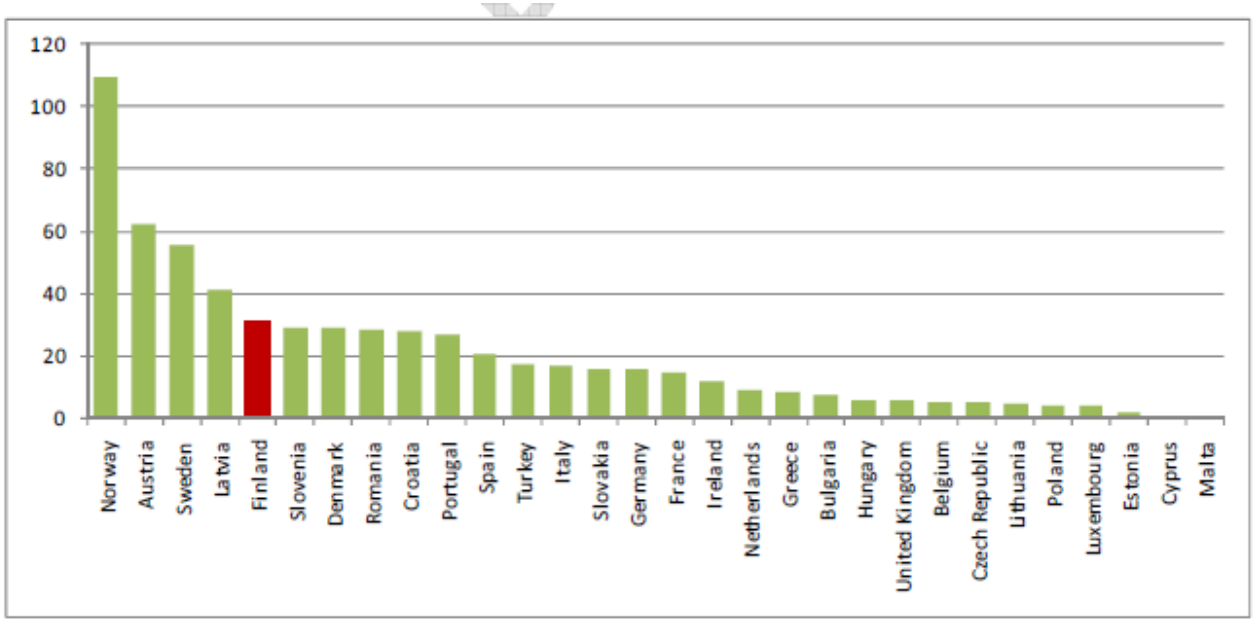
Finnish energy prices

Eurostat (2010 Prices)



This indicator presents electricity prices charged to final consumers. Electricity prices for industrial consumers are defined as follows: Average national price in Euro per kWh without taxes applicable for the first semester of each year for medium size industrial consumers (Consumption Band Ic with annual consumption between 500 and 2000 MWh). Data extracted April 2011.

Percentage of electricity generated from renewable sources in Finland

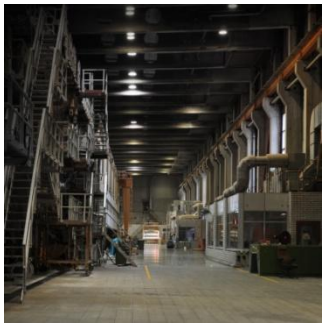


Source: Eurostat (April 2011)

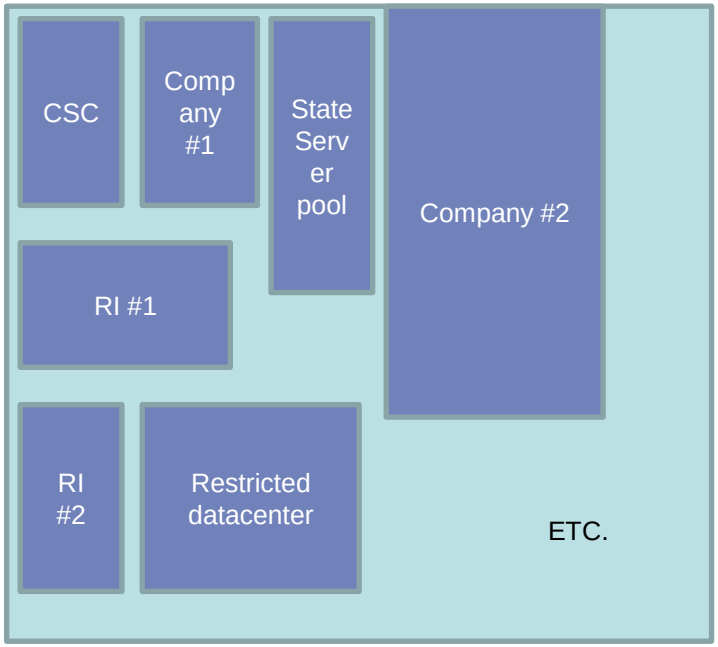
It measures the contribution of electricity produced from renewable energy sources to the national electricity consumption. Electricity produced from renewable energy sources comprises the electricity generation from hydro plants (excluding pumping), wind, solar, geothermal and electricity from biomass/wastes. Gross national electricity consumption comprises the total gross national electricity generation from all fuels (including autoproduction), plus electricity imports, minus exports.

CHOSEN LOCATION IN MORE DETAIL

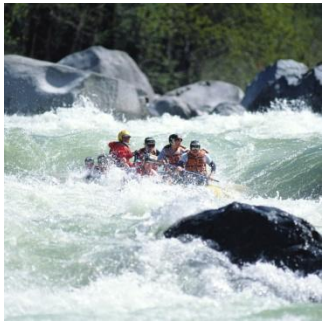
UNIQUE SYNERGY BENEFITS



- The infrastructure is already in place: The Data Center is being built in an old paper mill by the River Kajaani
- Space: 8000m^2 in total, of which CSC will first occupy 4000m^2 with an option for another 4000m^2 for future growth



SUSTAINABLE AND GREEN ENERGY SUPPLY



- Three hydropower stations on site: Maximum capacity of the power plants is 30 MW
- Additional bio-energy plant of 80 MW
- Short energy supply chain
- Massive water stations; pumps, pipelines and filtering systems ensuring sufficient cooling capacity
- Raw water from the River Kajaani used for cooling: Water temperature between 0-21°c

Network connections

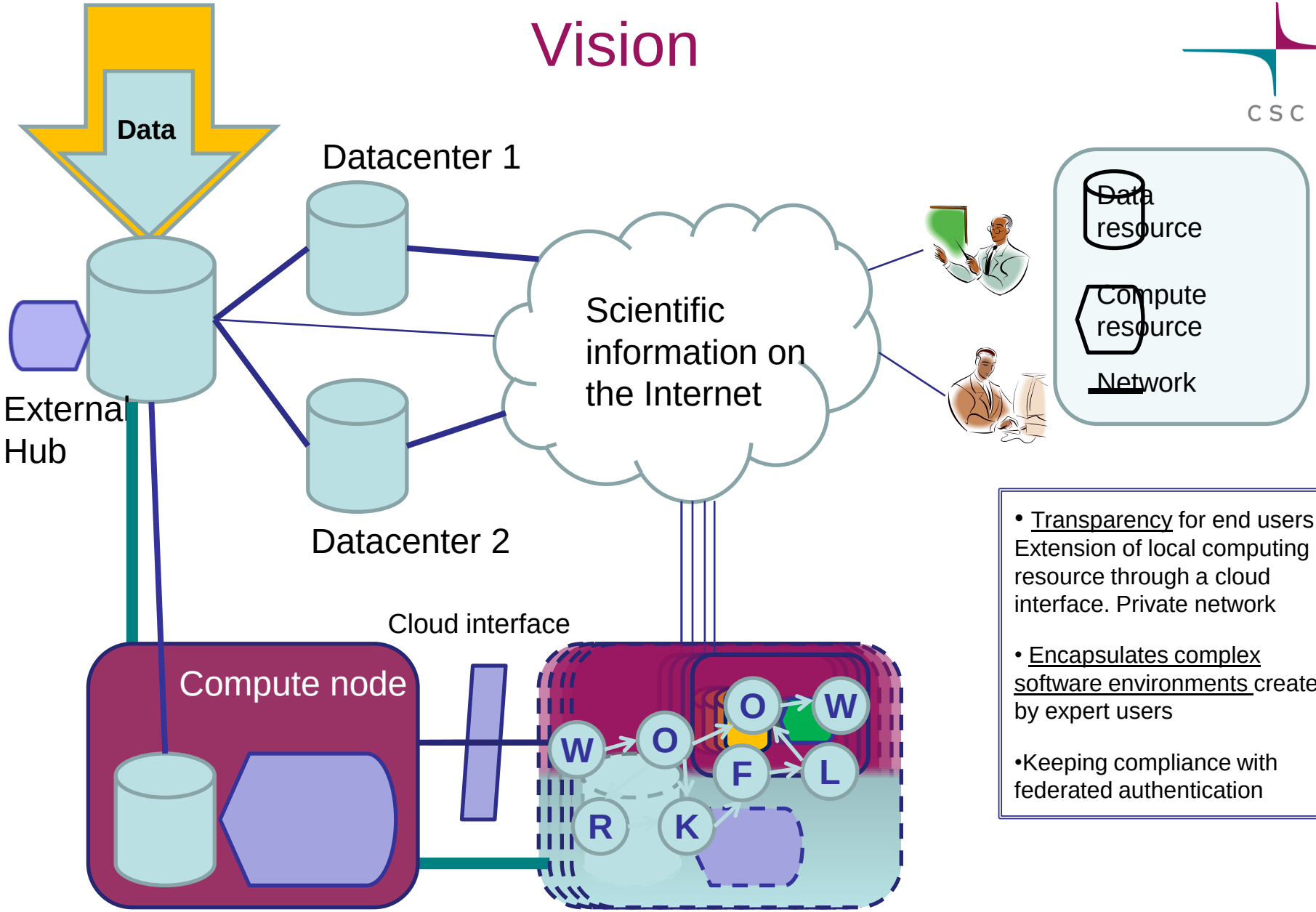
- The Funet hybrid backbone network provides reliable, high-capacity IP services and lightpaths for all *Funet member organizations* in Finland.
- Funet is connected to international academic networks via *NORDUnet* (Nordic academic network collaboration).
- Link speeds towards member organizations - up to 10 Gbps (already in use), capacity of up to 40 /100 Gbps enabled / coming
- Funet optical backbone network is based on leased dark fiber between and inside major Finnish university locations, including **Kajaani**

Case Kajaani

- Kajaani
- Won a competition of several sites, based on hard facts
- The aim is to

1. Start off a new industry in Finland: massive eco-efficient DataCenters
2. Kick off an ecosystem around the DC: EDC Park
3. Build a DC for CSC
4. Also, build eco-efficient DC space for a Finnish commercial operator

Vision



- Transparency for end users. Extension of local computing resource through a cloud interface. Private network
- Encapsulates complex software environments created by expert users
- Keeping compliance with federated authentication

Data services

Core data services are building blocks of Data Infrastructure

mainly included on bottom layer of data services

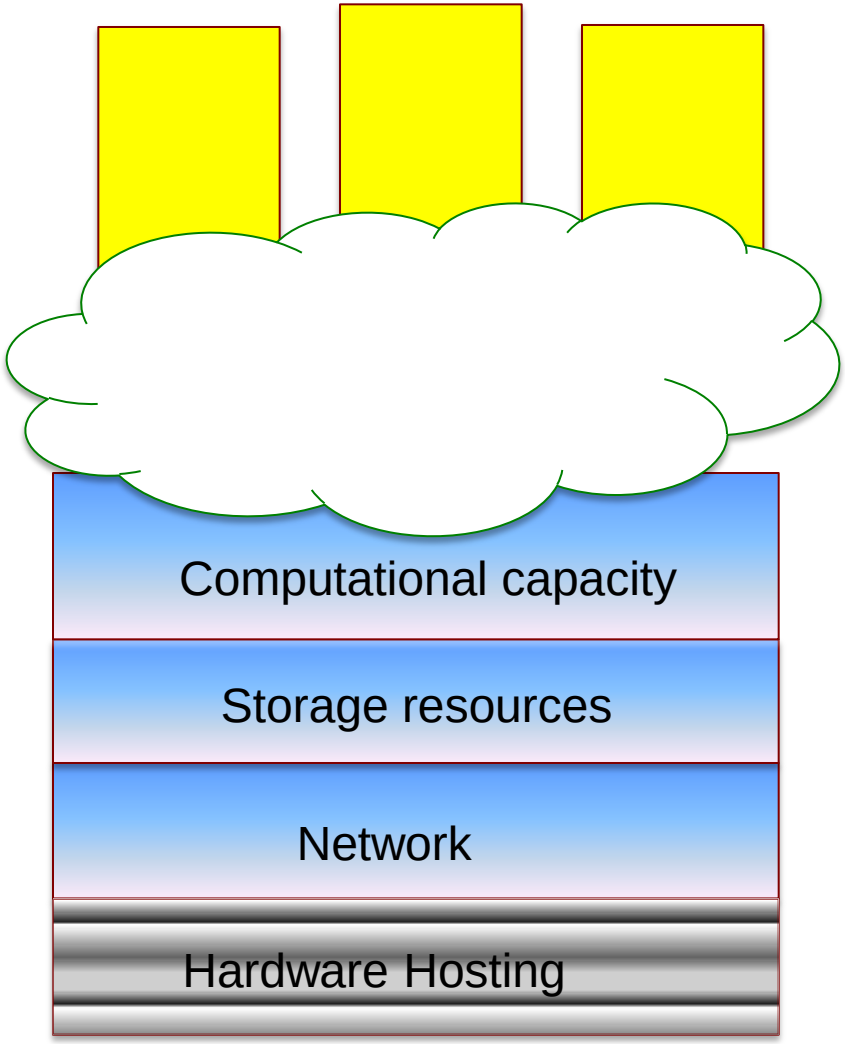
Fundamental Core Services

- Single Sign On (federated AAI)
- Data access and upload
- Long-term preservation
- Persistent identifier service
- Workspaces
- Web execution and workflow services
- Network services
- Monitoring and accounting services

Extended Core Services (community-supported)

- Joint meta data service
- Joint data mining service

Scalable Virtual Research Environment (SVRE)

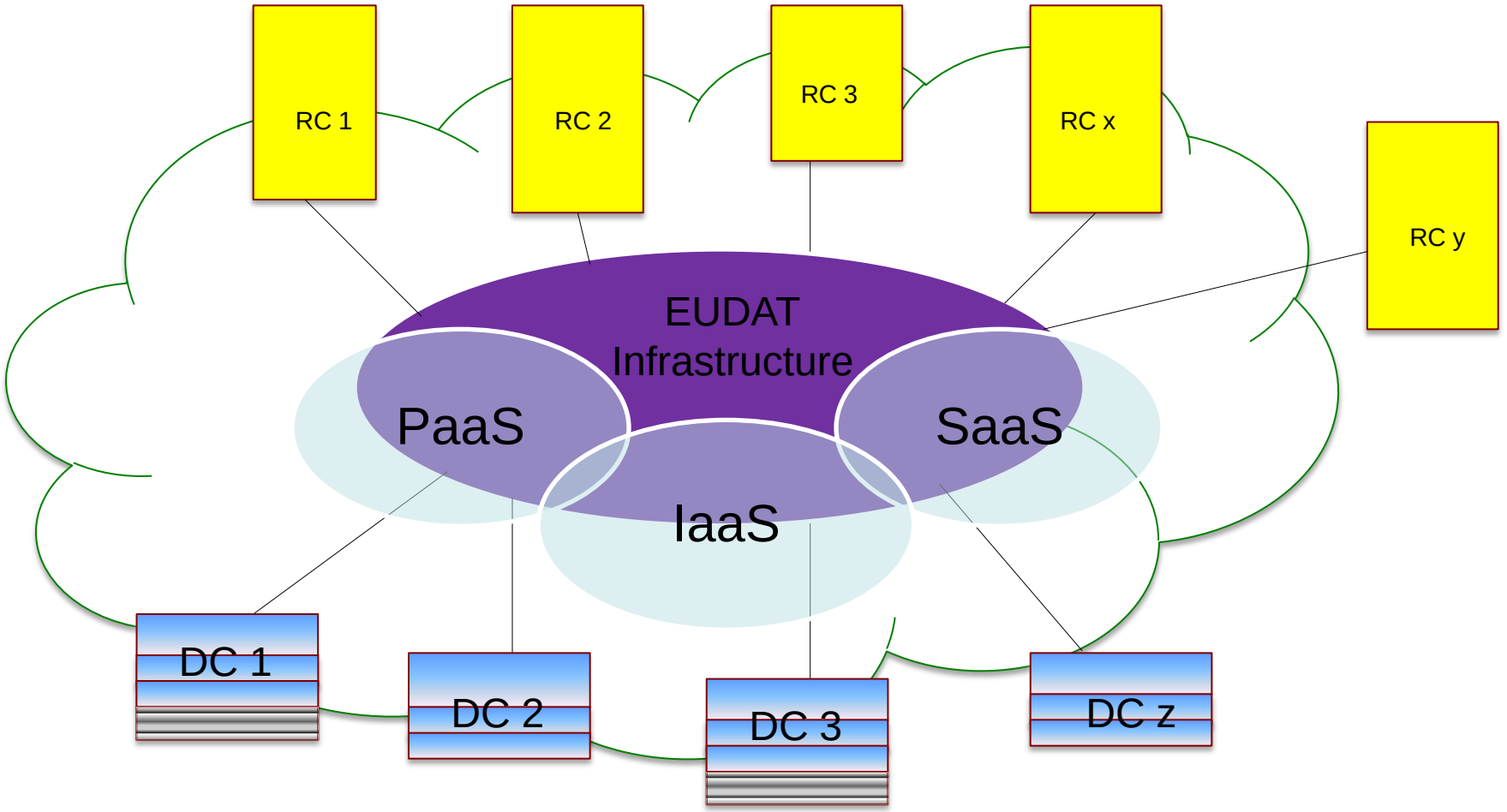


Thematic resources and processes

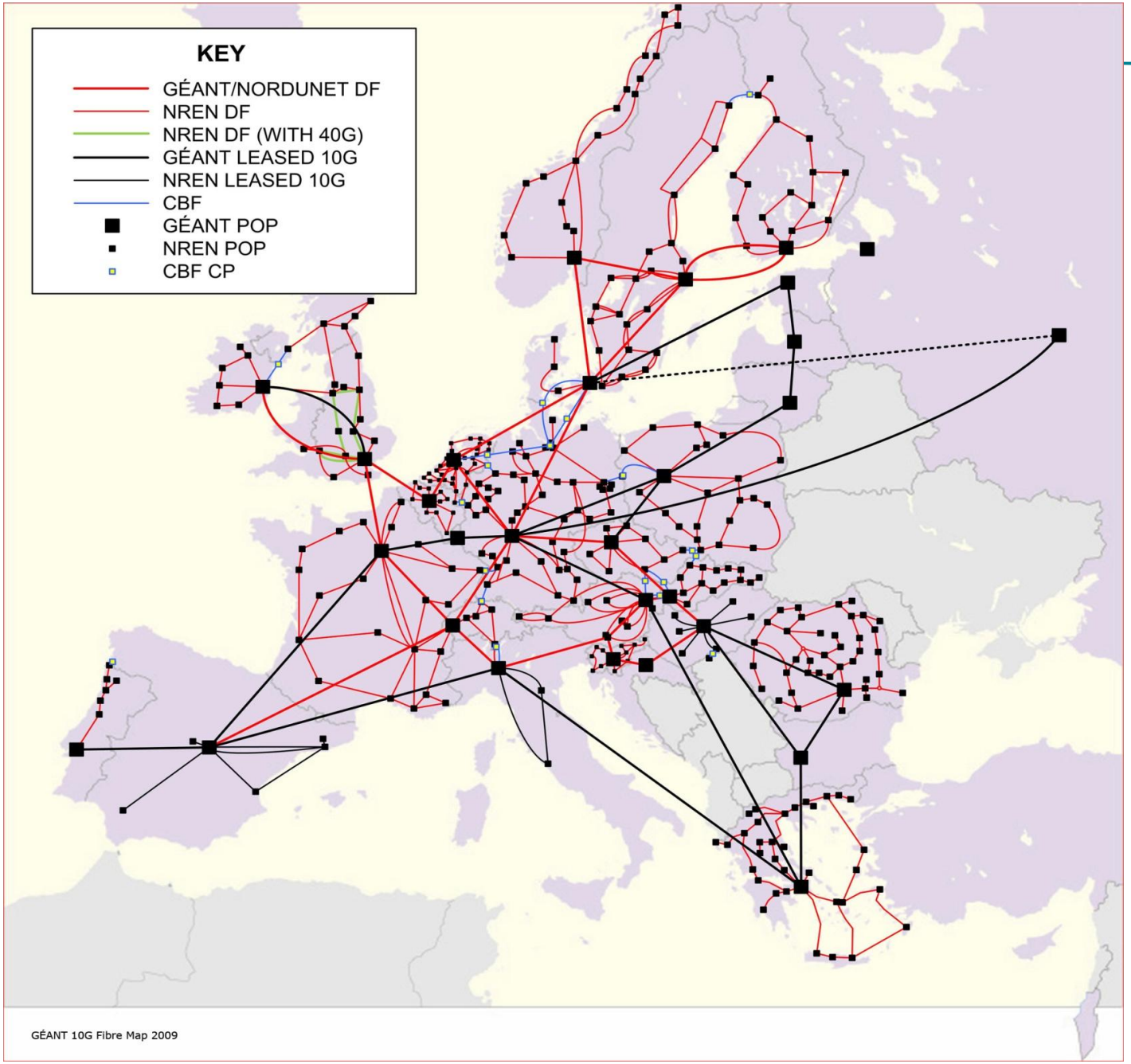
Interoperability layer
AAI, Orchestration,
Secure data exchange/management

Scalable platforms

How could an EUDAT Cloud look like?




FUNET
YHTEYDET 2010
CSC
KUITUVERKKO
SIIRTOYHTEYDET
NORDUNET
ULKOYHTEYDET
IP- ja kuituverkon liityntäpisteet
Kuituverkon liityntäpisteet
NORDUnet
Sodankylä
Rovaniemi
Kemi
Oulu
Kajaani
Kokkola
Vaasa
Seinäjoki
Kuopio
Joensuu
Jyväskylä
Tampere
Mikkeli
Pori
Lappeenranta
Hämeenlinna
Rauma
Lahti
Kouvola
Salo
Turku
Kotka
Espoo
Helsinki
Pietari
RUNNet
Oslo
Tukholma
Internet
Reykjavik
(IceLink: Grönlanti, Kanada, USA)
Internet
New York
Kööpenhamina
Lontoo
Hampuri
Amsterdam
GÉANT
Poznań


KEY
GÉANT/NORDUNET DF
NREN DF
NREN DF (WITH 40G)
GÉANT LEASED 10G
NREN LEASED 10G
CBF
GÉANT POP
NREN POP
CBF CP
GÉANT 10G Fibre Map 2009

CSC

Connectivity - Capacity

- **Funet: Kajaani - Helsinki**
  - x * 10 Gbps in production, 40 Gbps in pilot 2011, 100 Gbps on roadmap
  - fully redundant ring network, reliability 100 %, 24/7 monitoring
- **NORDUnet: Helsinki - Copenhagen**
  - x * 10 Gbps and 40 Gbps in production,
    100 Gps on roadmap and in testing
  - fully redundant, separate fiber paths also undersea, reliability 100%
  - added resiliency possible through northern connection to Sweden
- **GÉANT: Copenhagen - CERN**
  - x * 10 Gbps , 40 Gbps ready, 100 Gps on roadmap and in testing
  - fully redundant connectivity available

Connectivity – Latency

- Kajaani – Helsinki : 8 ms
- Kajaani – Copenhagen : 25 ms
- Kajaani – CERN : 47 ms

- backup connectivity roughly <= 50% more

  (round trip latency)

Funet / NORDUnet references

- NDGF LHC OPN 10 Gbps
  (Helsinki – Copenhagen - CERN)
- DEISA OPN 10 Gbps
  (Helsinki - Frankfurt)
- Runnet-Funet-NORDUnet-GÉANT
  n * 10 Gbps
  (St. Petersburg-Helsinki-Stockholm-Copenhagen-Amsterdam)
- about 65 Lightpaths and other private connections in production nationally


75t


IDEARK

How do we measure?

- Everywhere:
  - Power generation and distribution
  - UPS, Cooling systems, power distribution
  - Hot and cold air mixing, leaks, bad layout
  - Server fans, power supplies
  - CPU/memory, switches, storage
  - Inefficient code
  - Low server/storage utilisation, bad capacity planning, legacy applications

BREEAM

PUE

EU code of conduct for data centers

Partnerships

- Datacenter initiatives require multifaceted competencies and joint efforts
- Kajaani datacenter is a “Megadatacenter”. The more tenants there are the more beneficial it will be for the parties
- CSC is the anchor tenant of this datacenter. CSC is willing to share information on green datacenters / welcome other research organizations and HPC´s to Kajaani.
- Should we establish the open research datacenter? Like Facebook´s open compute project http://opencompute.org/
- Colocation of multiple research organizations could lead to new partnerships and cooperation possibilities
- In any case, Kajaani will be a place for new kinds of partnerships, bringing together
  - Traditional paper and pulp industry, telecom and IT
  - Large global companies, small local organizations
  - Governmental / research and commercial companies

Next Steps

- Kajaani datacenter preplanning completed
- Expected production summer 2012
- Procurement for supercomputing and data management started
- Datacenter related activities boosted by different Finnish stakeholders
  - TEKES, TIVIT, Teknologiateollisuus, SITRA, CSC, ministries, Invest in Finland, …

Thank you, let´s stay in touch!

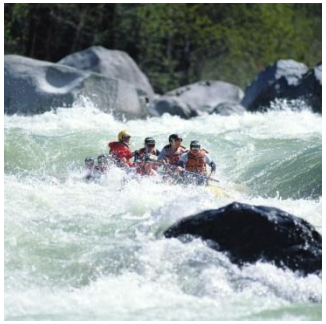
Pirjo-leena.forsstrom@csc.fi

Impact and benefits

CSC and others

1. Scalable cost- and energy efficient datacenter solution for CSC
2. Possibility to consolidate datacenters here, for example for the state and governmental organizations
3. Possibility to attract foreign industry to establish their activities in Finland
4. Opportunities in hosting European research infrastructures? Trying to find the optimal solution to resource utilization, sharing best practices.
5. Employment and revitalization of rural areas
6. Industrial opportunities in environmental technology and energy
7. Innovative new business models. Reuse of old factories and the built infrastructure around them
8. Building a reputation in advanced technology and sustainable development
9. Find successful ways of building trust between research and ICT service providers

SUSTAINABLE AND GREEN ENERGY SUPPLY



- The Data Center aims for a near zero carbon footprint during its operational lifetime
- Three hydropower stations on site: Maximum capacity of the power plants is 30 MW
- Additional bio-energy plant of 80 MW
- Short energy supply chain
- Massive water stations; pumps, pipelines and filtering systems ensuring sufficient cooling capacity
- Raw water from the River Kajaani used for cooling: Water temperature between 0-21°c

Greener DataCenters – why?

- Datacenters use 2 % of electricity = more than all of Sweden
- IT industry carbon emission equals that of aviation
- The energy consumed by data center servers and related infrastructure equipment worldwide doubled between 2000 and 2005, and doubled again 2005 – 2010
- Something has to be done, this is our mission

CSC´s datacenter target setting => practical considerations

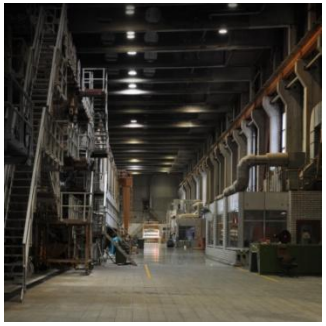
TARGETS

- **Scalability**
  - Long term commitment from CSC
  - Enablement of consolidation & synergy benefits
- **Green datacenter as a target**
  - Minimal carbon footprint
  - Brownfield, with a major part of the infrastructure already existing
  - BREEAM bespoke Outstanding (#1 in the world?)
  - Top class efficiency (PUE Cat.3 < 1.2)
  - 100% renewable energy contracts and ”delivery”
  - Heat re-use by ourselves and others
- **Cost efficiency**
  - Energy, networks, total investment cost
  - No diesel generators or tanks – our generator is a dedicated hydroplant
- **Reliability**
  - Up to 99.997 availability (cooling & power)
- **Policies and guidelines to be followed**
  - EU Code of conduct for Datacenters

PRACTICAL CONSIDERATIONS

- Finding the optimal location
- Solving the energy and cost efficiency requirements. Self build or rent?
- Ensuring connectivity
- Budgetary frame

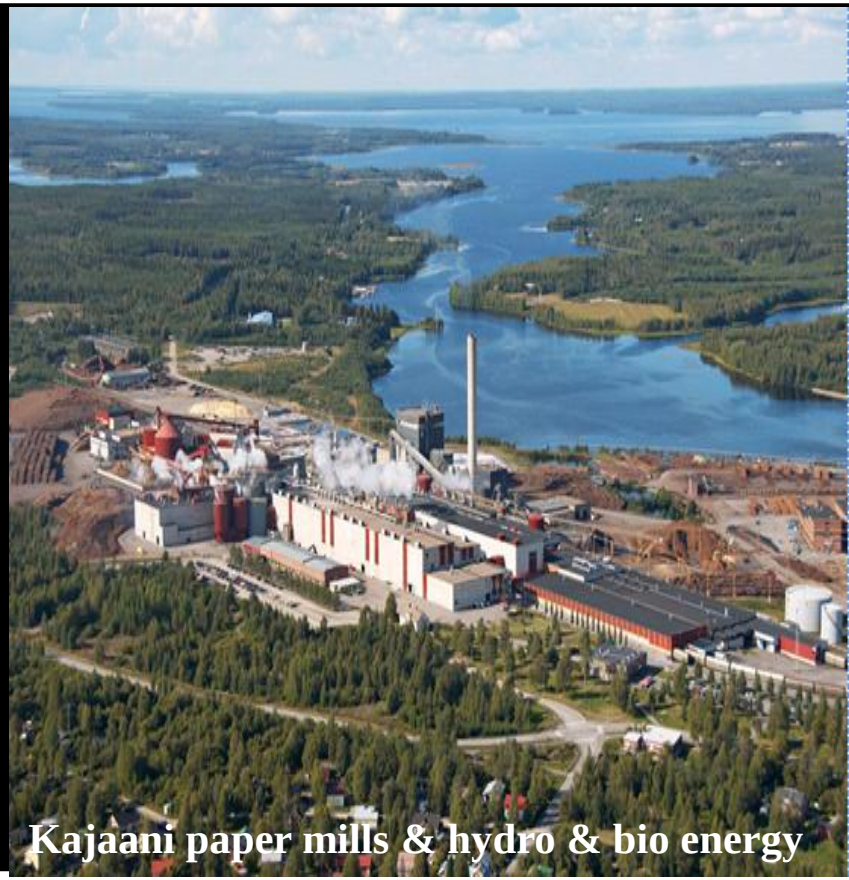
UNIQUE SYNERGY BENEFITS



- The infrastructure is already in place:
  The Data Center is being built in an old paper mill by the River Kajaani
- Space: 16 000m^2 in total, 8000 for CSC, of which CSC will first occupy 4000m^2 with an option for another 4000m^2 for future growth
- Project started in 2010 and estimated to be up and running by summer 2012

Finland as a datacenter location 1/2

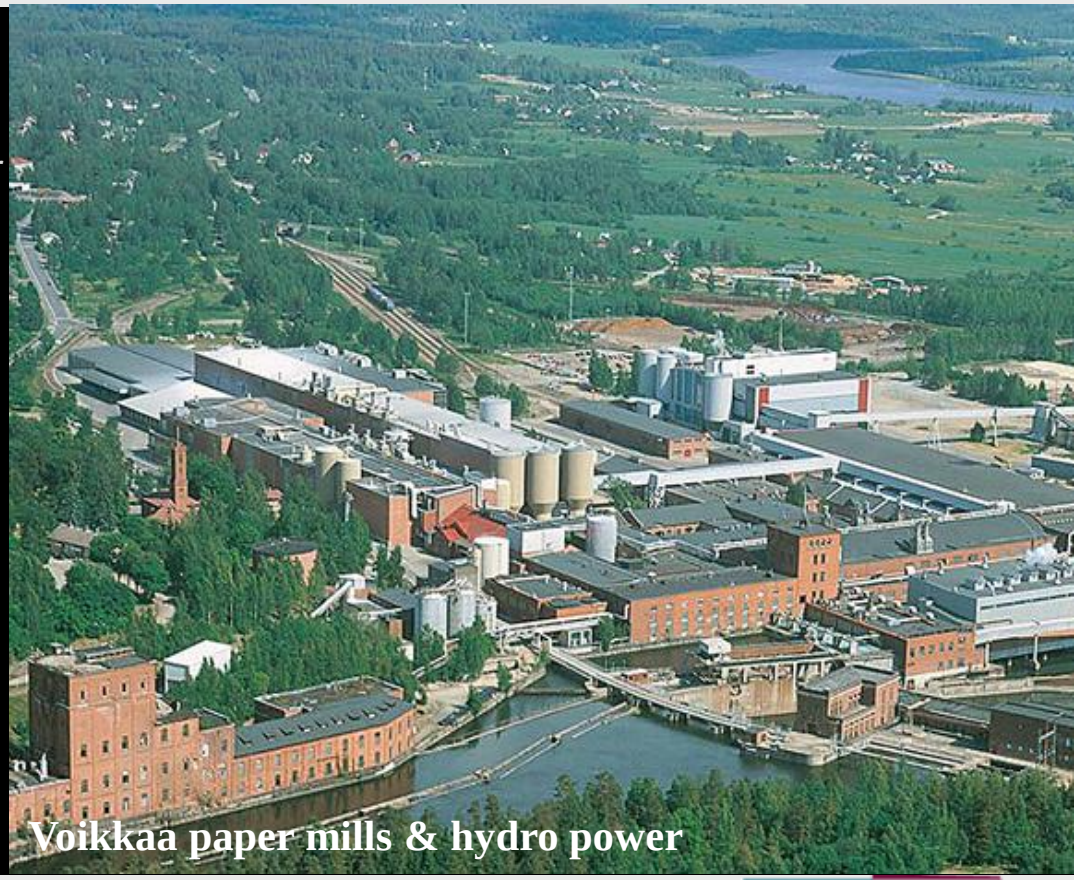
- *Modern and reliable infrastructure (national power grid, roads, airline connections, data networks)*
- *World class education system and competences on ICT & energy*
- *Steady economical and political conditions*
- *Cheap energy (www.energy.eu) and strongly increasing CO_2 –free capacity*
- *Cool climate and water resources*
- *No major earthquakes (4.1 biggest ever)*
- *No major storms or other dangerous natural phenomena*
- *European Gateway to Russia*



Kajaani paper mills & hydro & bio energy

Finland as a datacenter location 2/2

- *Massive paper and pulp industry facilities with local green energy production capabilities and raw water intakes*
- *Heavy duty energy grid (local and national)*
- *Existing, large scale land / building mass*
- *Existing facility services onsite*



Voikkaa paper mills & hydro power

CSC

Storage

- **File system sizing**
- Currently biggest HPC clusters have around 10PB of storage capacity but common size is approximately 1PB in top 10. Our expectations for needed storage capacity are in 500TB - 2PB range. Constantly growing storage capacity causes huge problems in backups and long recovery times may be challenging when non-planned breaks occur so it's not very wise to just grow capacity without thinking actual needs for the storage systems.
- **Performance and capacity**
- File system performance and its responsiveness is important factor because that affects directly how customers encounter our systems. If the file operations are continuously lagging, the environment will be unpleasant to work in. Steady growth of hard disk capacities has been a common trend in last decades but IO-bandwidth and IO-operations per second (IOPS) have not improved significantly. Hard disk performance deteriorates when it's utilization rate increases and files are added and removed. Solid-state drives (SSD) are finally solving storage performance problems. First generation SSDs suffered various problems, e.g. asymmetric read vs. write performance, flash cell wear out limit was low, performance dropped drastically after being used some time and drives had a quite low capacity (3).

Parallel filesystems

- **Lustre** is a opensource (GPL) parallel file system generally used for large scale cluster computing. Lustre key features are scalability, performance, POSIX compliance and high-availability. Lustre file system can be build up from different kind of hardware, there is no vendor lock in. Using solid state disks Lustre is capable of sustaining over one terabyte per second (1TB/s) of aggregate bandwidth with access latencies under 50 microseconds (4). Lustre has still some performance problems, e.g. it doesn't get any substantial benefit of using SSDs on metadata servers which are usually performance bottlenecks. There are also some preliminary support for Lustre HSM integration via additional HSM engine.
- **IBM General Parallel File System (GPFS)** is an enterprise class scalable, highly-available, high performance file system optimized for multi-petabyte storage management. GPFS supports full POSIX filesystem semantics and Hierarchical storage management (HSM) is integrated.
- **Parallel NFS (pNFS)** is an extension to NFSv4 that separates metadata and data paths, and allows clients to access storage devices directly and in parallel. pNFS has been proposed to eliminate the single server bottleneck associated with the current NFS servers while keep the ease of management and interoperability features of NFS. pNFS client and server implementations are not matured yet, Linux distribution makers are planning to ship their operating systems with pNFS support at next year (2010) (5).

Parallel filesystems (cont.)

- **PVFS** is designed to provide high performance for parallel applications, where concurrent, large IO and many file accesses are common. PVFS provides dynamic distribution of IO and metadata, avoiding single points of contention, and allowing for scaling to high-end terascale and petascale systems. Missing features: hardlinks, shared mmapping of files and there is no free space balancing on servers.
- **IBRIX Fusion** is a scalable parallel file system which HP acquired recently. IBRIX has all the enterprise level file system features and tools for it's management. IBRIX can scale to up to 16 petabytes of capacity in a single namespace. IBRIX supports only NFS and CIFS protocols and it's not clear that does HP sell it separately or only with HP hardware.
- **SUN QFS** is yet another enterprise grade parallel file system which has also tight integration to their SAM (HSM) software. QFS future doesn't look very healthy, technical specifications says that maximum QFS volume size is 4PB and they have not released client drivers to RHEL5 or SLES11. Sun released parts of the SAM-QFS source code to the OpenSolaris? project in March 2008 but at least public development list is more or less dead (6)

Filesystem	Product vitality	License	Price	Performance + scalability	Reliability	Management	Future	Other Comments
Lustre	yes	GPL	Free	Good performance and scales well	Clustered metadata not yet available, otherwise redundant	Requires quite lot of knowledge	Depends on SUN / Oracle fusion	So many moving parts, maintenance and tuning needs an enthusiastic attitude
GPFS	yes	Proprietary, complex licensing model	Expensive	Good performance and scales well	High reliability and availability	Management tools included in product		
pNFS	yes	GPL/ Opensolaris binary license/ Proprietary	Free - Expensive	Young product, should perform and scale well	Designed to be reliable	N/A	Multiple big SAN vendors funds the development	Who wants to be the guinea pig...
PVFS	yes	GPL	Free	PVFS fits best for a systems where all applications can be tested with it	HA by using standard OSS tools	Manual work	Rarely used in top500 clusters	Doesn't suite well for a general purpose file systems
IBRIX	yes	Proprietary	Expensive	Modular design, scales up to 16PB	High availability and dynamic load balancing	Management tools included in product	HP has just bought IBRIX	Only NFS and CIFS protocols are supported
QFS	yes/no?	Proprietary	Expensive	4PB volume size limit	yes	Management tools included in product		

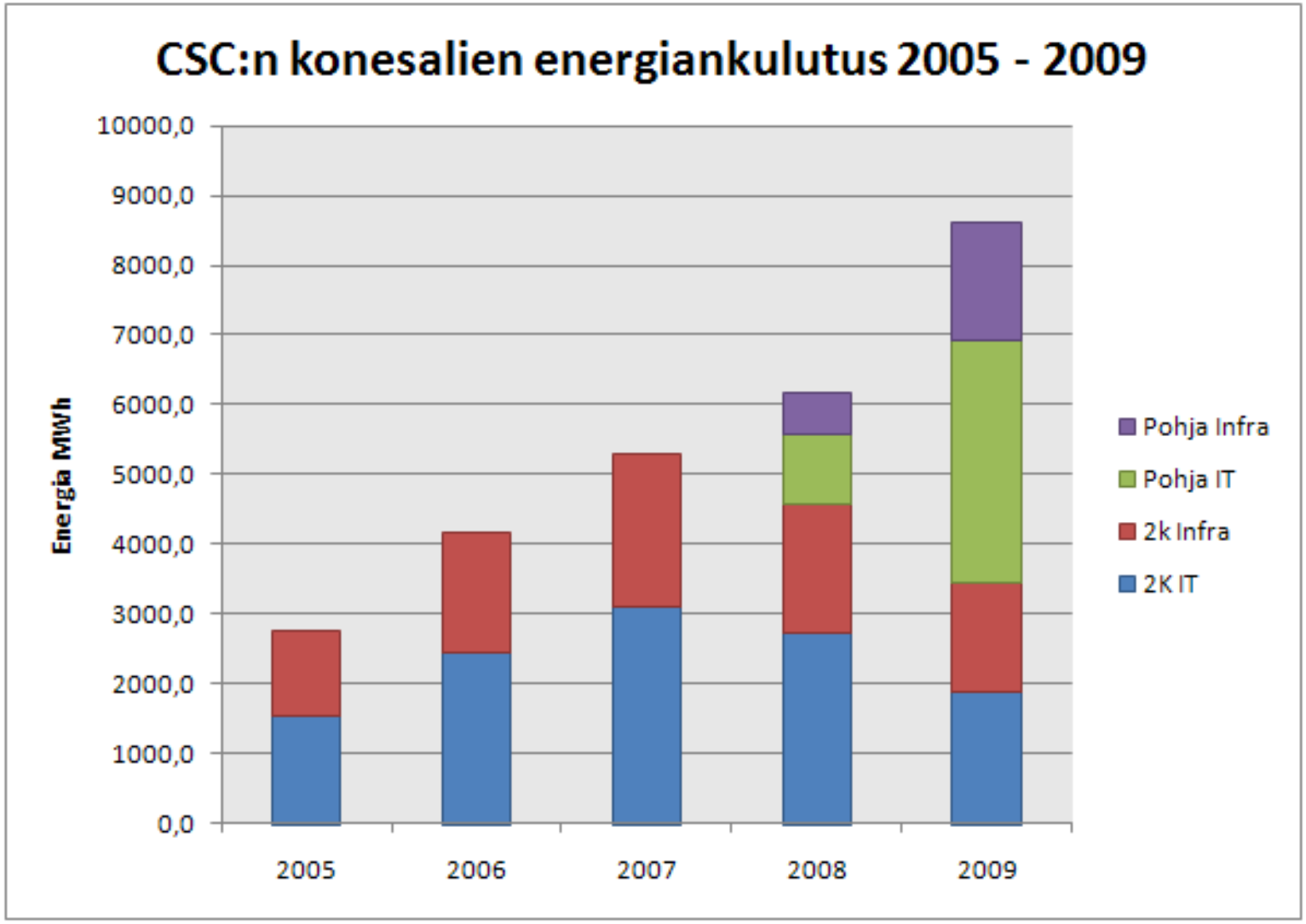
Disk and bus technologies

- File server IO buses are often very heavily loaded and can be described as a performance bottleneck. PCI express version 3.0 promises to double the interconnect performance rate compared to PCIe version 2.0. Currently one internal SSD card can saturate whole PCI Express x16 bus which should be shared between Infiniband, 10GbE and FC controller cards on file servers (4).
- SAS currently supports point data transfer speeds up to 6 Gbit/s, but is expected to reach 12 Gbit/s by the year 2012. SATA bus speed was just lifted to 6Gbit/s in Serial ATA Revision 3.0 specs (7). These buses cannot deliver enough bandwidth to the fastest SSD accelerators, but on the other way traditional disks cannot reach the bus bandwidth.

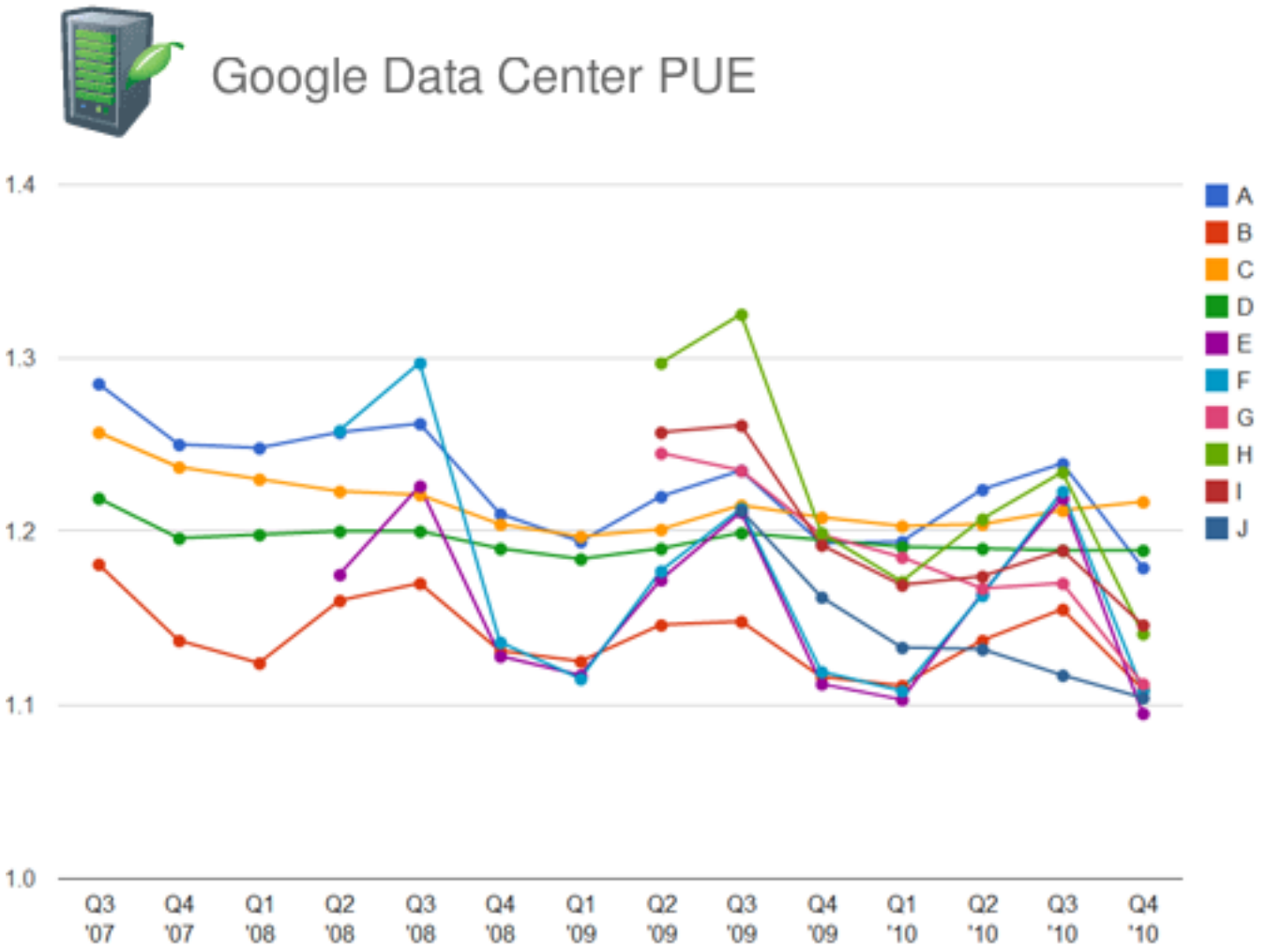
SSD/memory caches

Multiple vendors have a different approaches for speeding up disk operations. External SSD caches can be attached to servers with FC, IB or 10GbE and can be used as a standard SAN device. Internal accelerators can give out more performance than external ones but a redundant configuration reduces performance advantage. One possible scenario for SSD accelerators could be a very fast scratch work space where IO-intensive applications could store data temporally. This kind of scratch disks doesn't need to be redundant, data loss is some times acceptable. Scratch disk would be cleaned actively and format on every service break. Final results should be moved on more reliable storage. SSD accelerators breakthrough will happen in a near future, time will show which implementation will be the winner.

Does energy efficiency matter?



Google Data Centers



http://www.google.com/intl/en/corporate/datacenter/

Kajaani project goals

- Environmentally friendly datacenter
  - "the most eco-efficient datacenter in the world“
  - The first “outstanding” BREEAM Data Center
  - "World record efficiency (PUE 1.1 – 1.2)"
  - Minimal carbon footprint
  - Use Kajaani river for cooling, U
  - Aiming to top of Green500
- Cost effective to operate
  - Low energy costs
  - Low energy use through efficiency gains

$$PUE = \frac{Total\ Facility\ Power}{IT\ Equipment\ Power}$$

Key aspects

- Free cooling almost 365/24
- Water-based, but optimised use of ambient air as well (water cooling is more efficient than air cooling)
- No diesel generators
- Smart metering and management
- Virtualisation
- Over time, better SW
- Final goal: as many GFLOPS/W as possible
- Current record at Green 500: 1,7 GFLOPS/W

Green500

The inaugural Green500 list was announced on November 15, 2008 at SC|08. As a complement to the TOP500, the unveiling of the Green500 ushered in a new era where supercomputers can be compared by performance-per-watt.

Ranking:		MFLOPS/W
1.	IBM Thomas J. Watson Research Center	1684.20
2.	National Astronomical Observatory of Japan	1448.03
3.	GSIC Center, Tokyo Institute of Technology	958.35
4.	The National Center … (NCSA), USA	933.06
5.	RIKEN institute, Japan	828.67
6.	Universitaet Wuppertal	773.38
7.	Universitaet Regensburg	773.38
8.	…	
9.	…	
10.	…	

Really, a DataCenter

- Just a huge heater with integrated logic
- A 2 MW DC wastes about € 1 million per year (of unused heat)
- So, one additional goal is to find a way to capture and re-use the heat