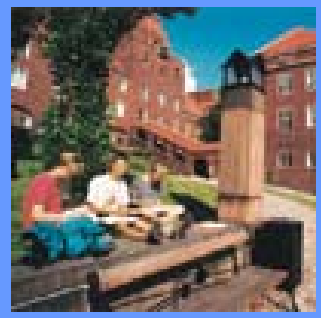
Nordic Grid Consortium

CSC
Helsinki

Parallab
Bergen



PDC
Stockholm

Petter Bjørstad, Parallab

Lennart Johnsson, PDC

Jari Järvinen, CSC

Why Grids

- Science has no bounds
- Science is increasingly based on (international) teams
- Sharing of unique (and scarce) instruments
- Sharing of data
- Sharing of human resources and expertise

Fiberoptic Communication

In 2010

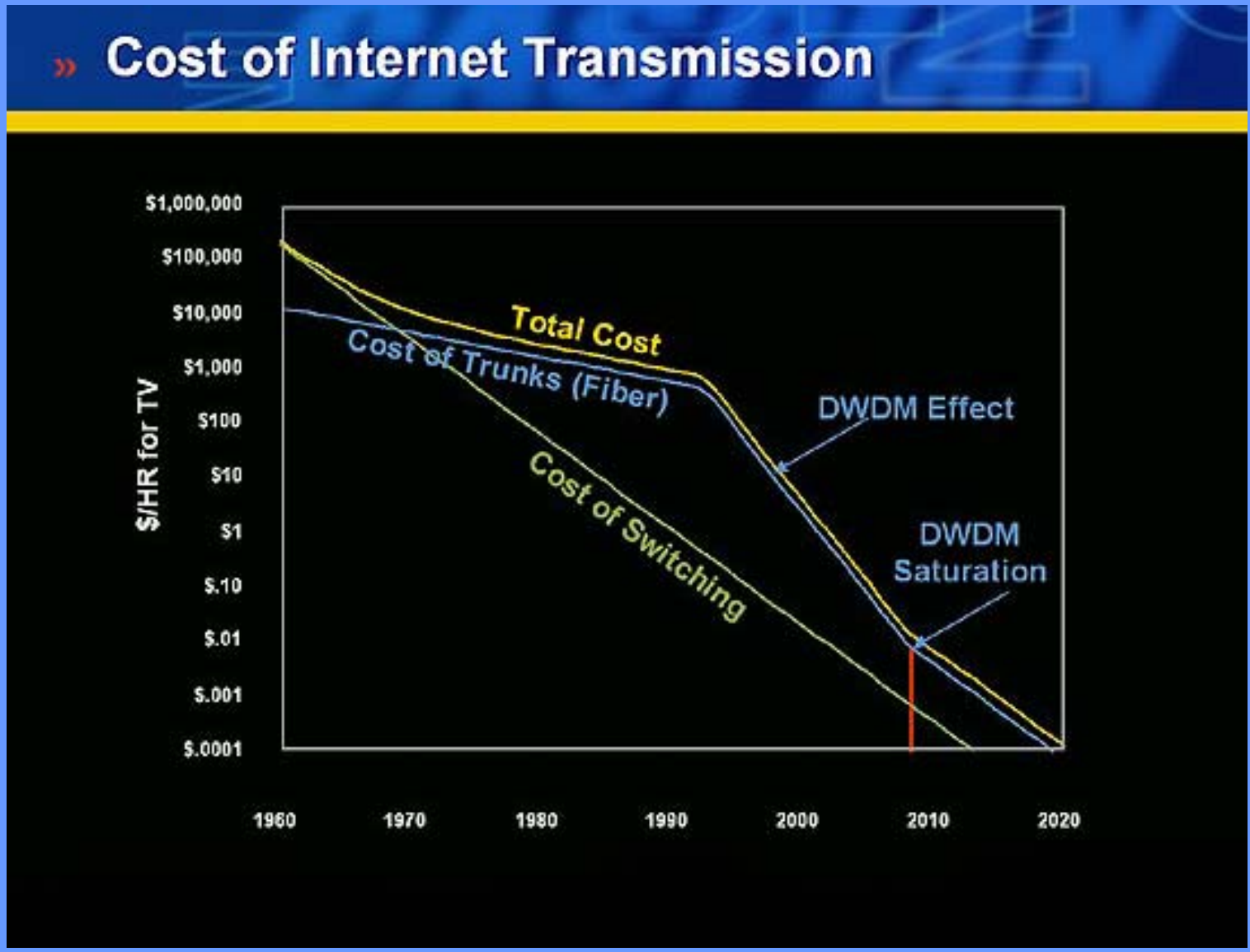
a million books can be sent across the Atlantic for 1$ in 8 seconds

all books in the American Research Libraries can be sent across the Atlantic in about 1 hr for $500

Fiberoptic Communication Milestones

- First Laser 1960
- First room temperature laser, ~1970
- Continuous mode commercial lasers, ~1980
- Tunable lasers, ~1990
- Commercial fiberoptic WANs, 1985
- 10 Tbps/strand demonstrated in 2000 (10% of fiber peak capacity). (10 Tbps is enough bandwidth to transmit a million high-definition resolution movies simultaneously, or over 100 million phone calls).
- WAN fiberoptic cables often have 384 strands of fiber and would have a capacity of 2 Pbps. Several such cables are typically deployed in the same conduit/right-of-way

Nordic Grid Consortium



L Roberts
Caspian
Networks

Nordic Grid Consortium

Cost of Fiberoptic Communication (2000)

- Fast Ethernet transceivers, 80 km, $700
- Gigabit Ethernet transceivers, 60 km, $2000
- 10 Gigabit Ethernet transceivers, 40 km, sampled.
- Ethernet typical price curve: factor of two/year
- MAN/WAN fiber cost: $0.5 – 6/strand meter, expected to drop to $0.07 – 0.10.
  - Installation: $3 – 200/m
  - Fiber: $0.05 – 0.15/strand meter
  - Termination: Several hundred dollars plus a few dollars/strand

Fiberoptic Communication Cost

Example: US – Europe

Installation cost: $1billion
Capacity: 384 strands operating at 1 Tbps peak
Utilization: 10 %
Amortization: 10 yr
Cost/sec: 1 billion/(10x8760x3600)= $3/sec
Effective capacity: (384/2)x(1 Tbps)x0.1 = 19 Tbps
Cost/bit: ~$0.15/Tb or about a dollar/Tbyte.
(or ~$100/Tbyte at 10 Gbps/strand)

The Battle of the Atlantic

Capacity coming online	Gbps*	RFS
Level 3/Global Crossing (Project Yellow)	3Q00	1,280
TAT-14 (Club)	640	4Q00
FLAG Atlantic-1 (FLAG/GTS)	2,560**	2Q01
Hibernia (360networks, Inc.)	1,920	2Q01
Atlantic Crossing -2 (Global Crossing)	2,560***	1Q01
TyCom Global Network	2,560	4Q01
Oxygen	No Go!	-------
Total	8,960	

Does not include C&W Apollo cable (RFS 2003)

* = Design capacity
** = Teleglobe buying 2 fibers
*** = Cancelled, AC-2 joining Level 3

The World Wide Computer

- Internet Provides Connectivity
- Web Provides Hyperlinked File System
- Distributed Storage Moving from SAN to NAS
- Peer-to-Peer Computing Provides Vast CPU Power
- Result--The Distributed Global Computer
  - Storage everywhere
  - Scalable computing
  - Wireless Interfaces Greatly Outnumber PC Interfaces

"When the Network is as fast as the computer's internal links, the machine disintegrates across the Net into a set of special purpose appliances"

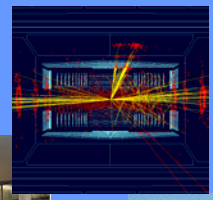
-Gilder Technology Report June 2000

Courtesy Smarr/Stevens

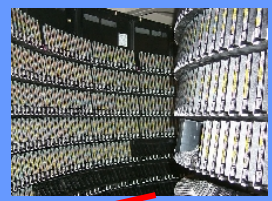
Nordic Grid Consortium

E-Science: Data Gathering, Analysis, Simulation, and Collaboration





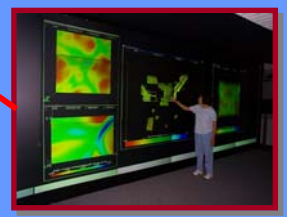




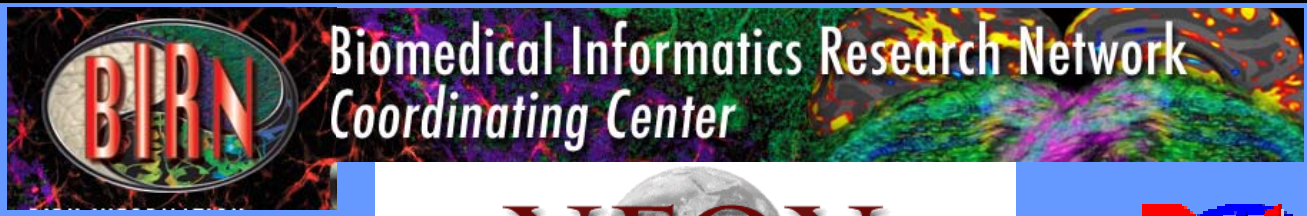




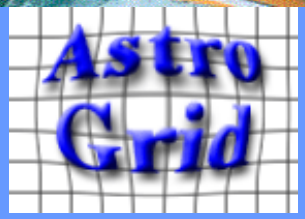


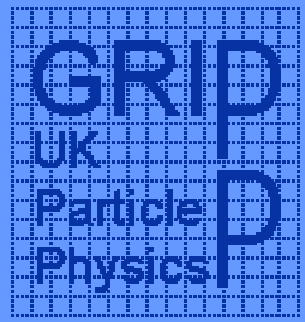


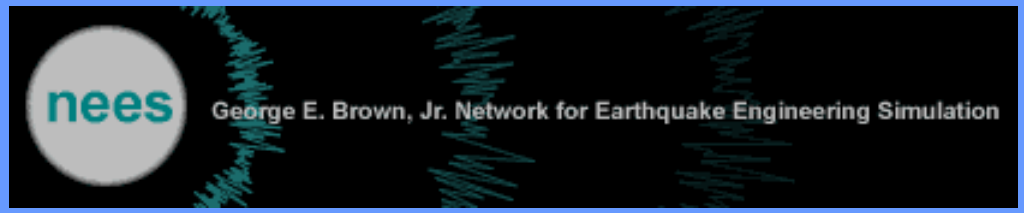
Nordic Grid Consortium



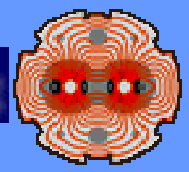
PAMELA





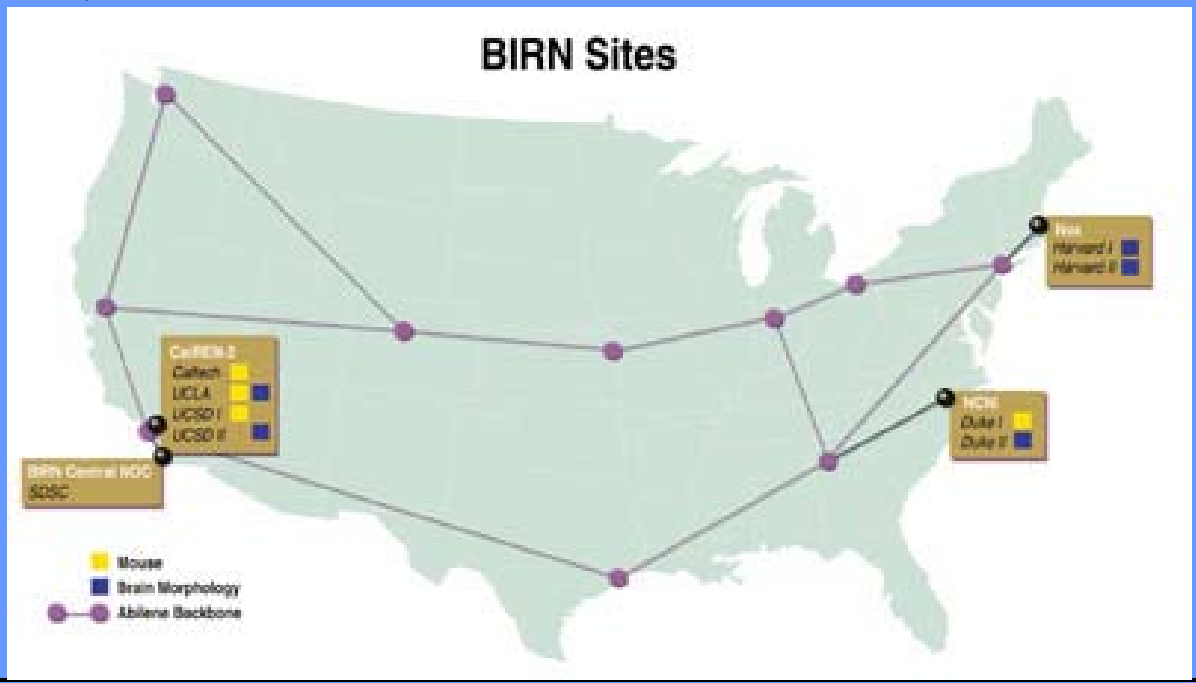


LHC The Large Hadron Collider Project



BIOMEDICAL INFORMATICS RESEARCH NETWORK (BIRN)

The BIRN is an NCRR initiative aimed at creating a testbed to address biomedical researchers' need to access and analyze data at a variety of levels of aggregation located at diverse sites throughout the country. The BIRN testbed will bring together hardware and develop software necessary for a scalable network of databases and computational resources. Issues of user authentication, data integrity, security, and data ownership will also be addressed. $20M.



Nordic Grid Consortium

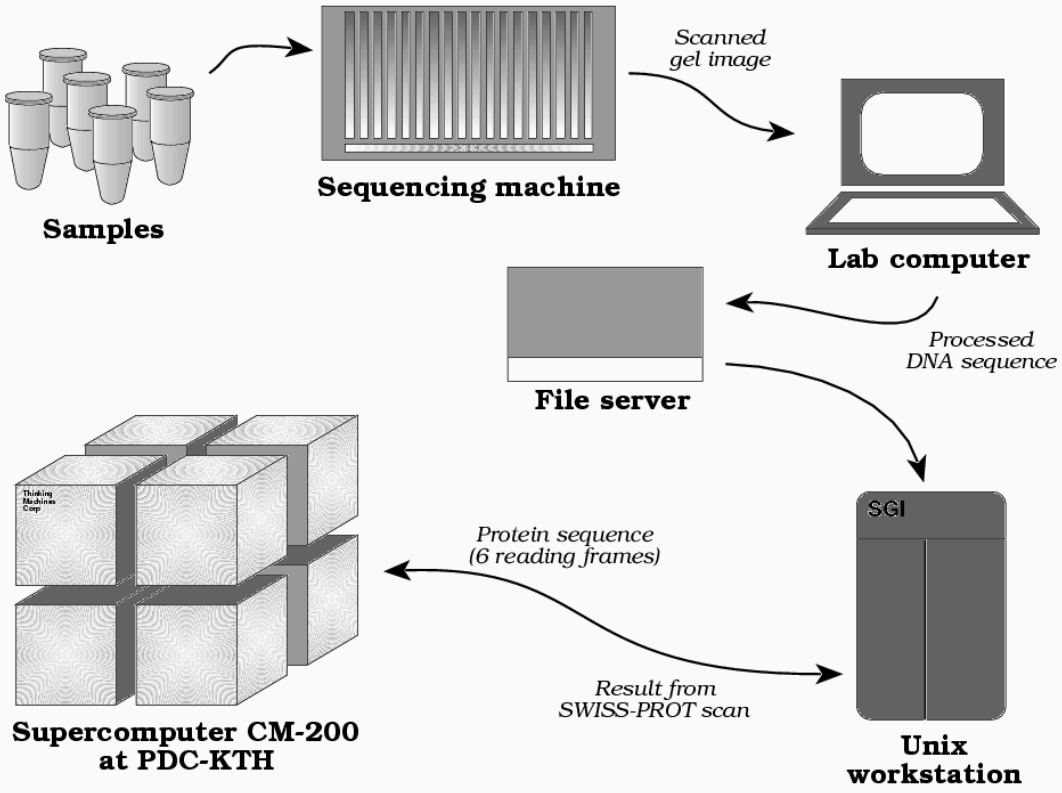
A NATIONAL DIGITAL MAMMOGRAPHY ARCHIVE

Digital Mammography

- About 40 million mammograms/yr (USA) (estimates 32 – 48 million)
- About 250,000 new breast cancer cases detected each year
- Over 10,000 units (analogue)
- Resolution: up to about 25 microns/pixel
- Image size: up to about 4k x 6k (example: 4096 x 5624)
- Dynamic range: 12-bits
- Image size: about 48 Mbytes
- Images per patient: 4
- Data set size per patient: about 200 Mbytes
- Data set per year: about 10 Pbytes
- Data set per unit, if digital: 1 Tbytes/yr, on average
- Data rates/unit: 4 Gbytes/operating day, or 0.5 Gbytes/hr, or 1 Mbps
- Computation: 100 ops/pixel = 10 Mflops/unit, 100 Gflops total; 1000 ops/pixel = 1 Tflops total

Nordic Grid Consortium

Automated Massive DNA Sequencing

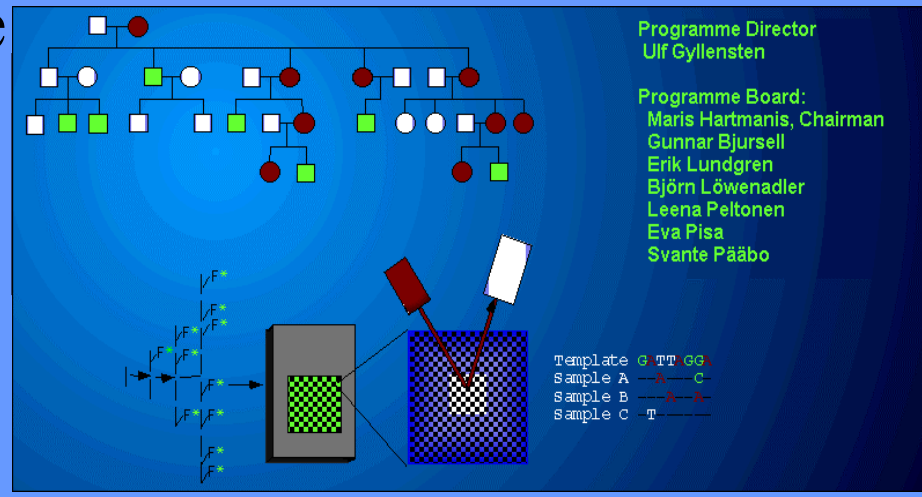


PDC in collaboration with Center for Structural Biochemistry, KI 1992

Nordic Grid Consortium

UU/SSF National Large-Scale Genotyping laboratory, since 1998

- Data repository and long term storage at PDC



Genotyping Machine

HSM

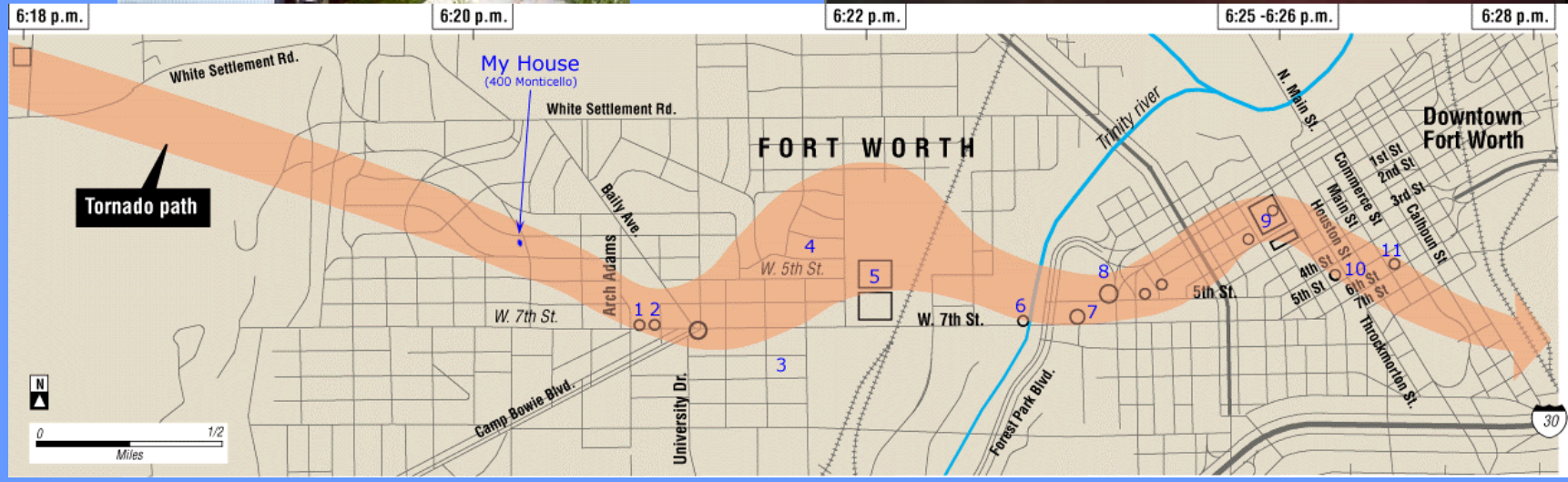
toad.pdc.kth.se

betula.genpat.uu.se

http://genpat.pdc.kth.se:8001/

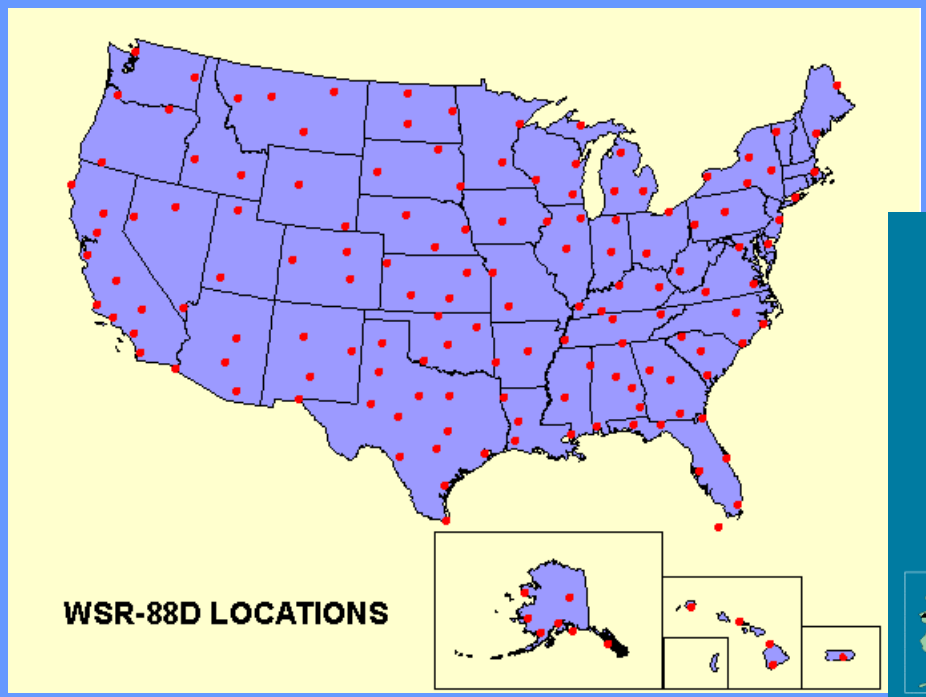
Nordic Grid Consortium

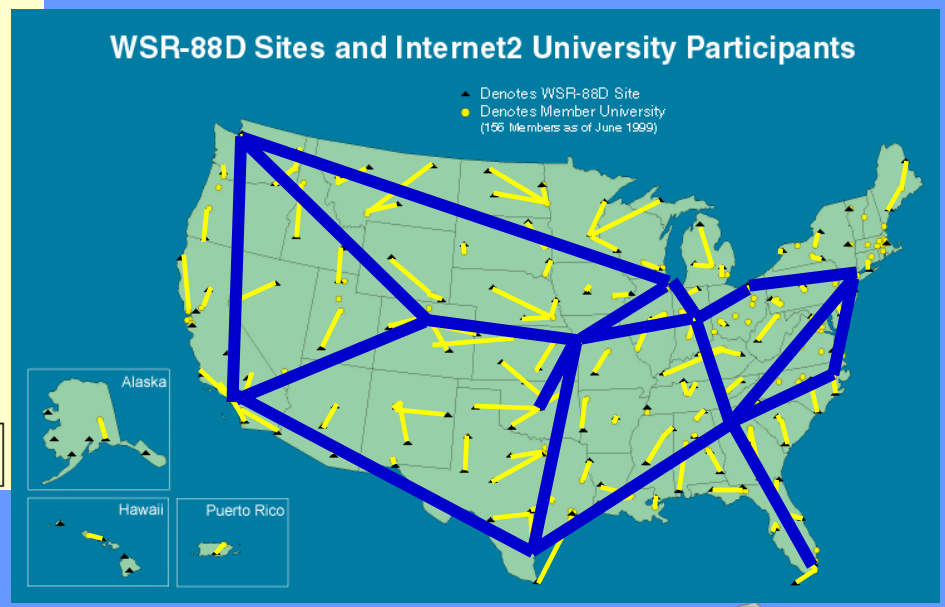
March 28, 2000 Fort Worth Tornado



Nordic Grid Consortium

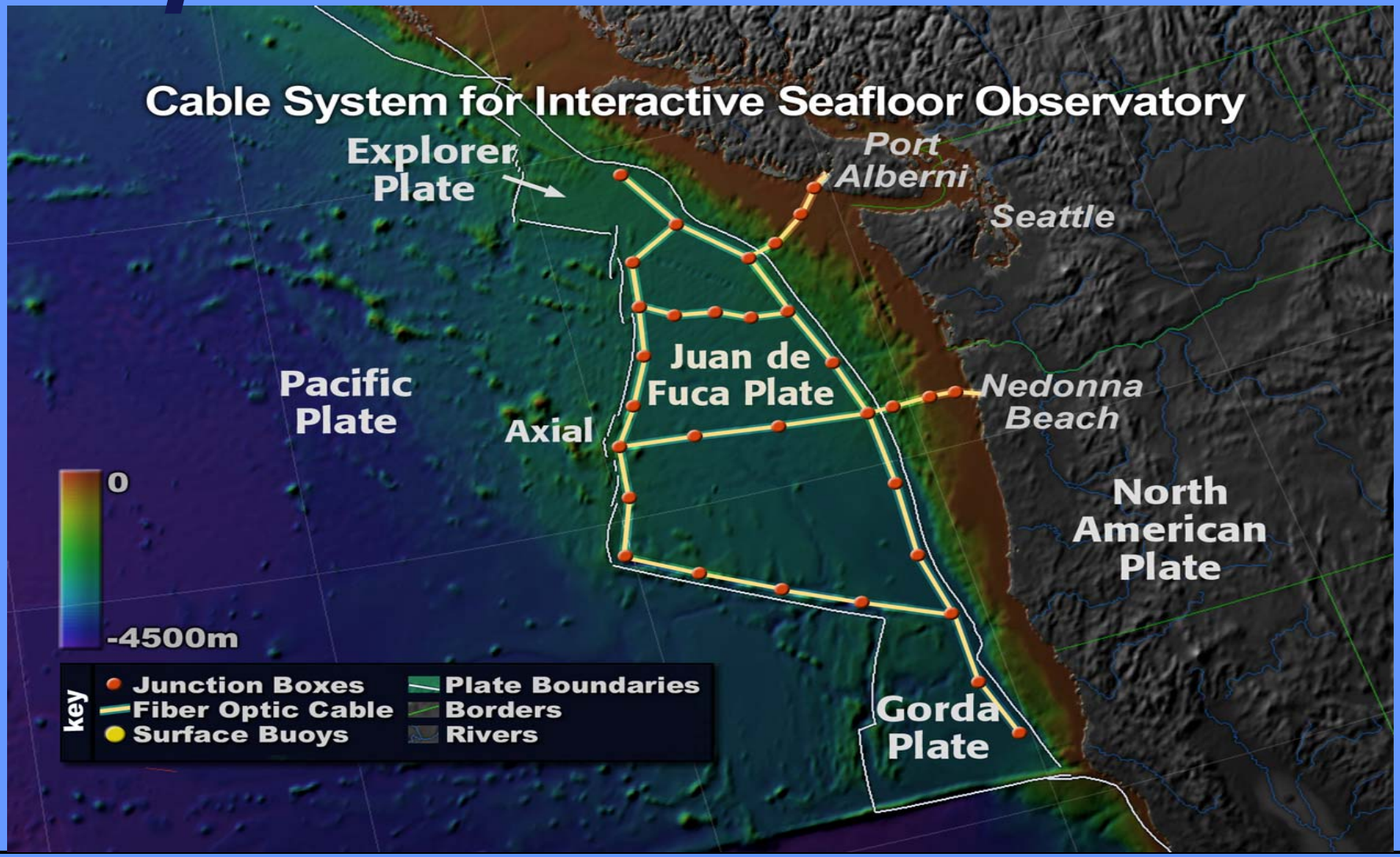
In 1988 … NEXRAD Was Becoming a Reality





Nordic Grid Consortium

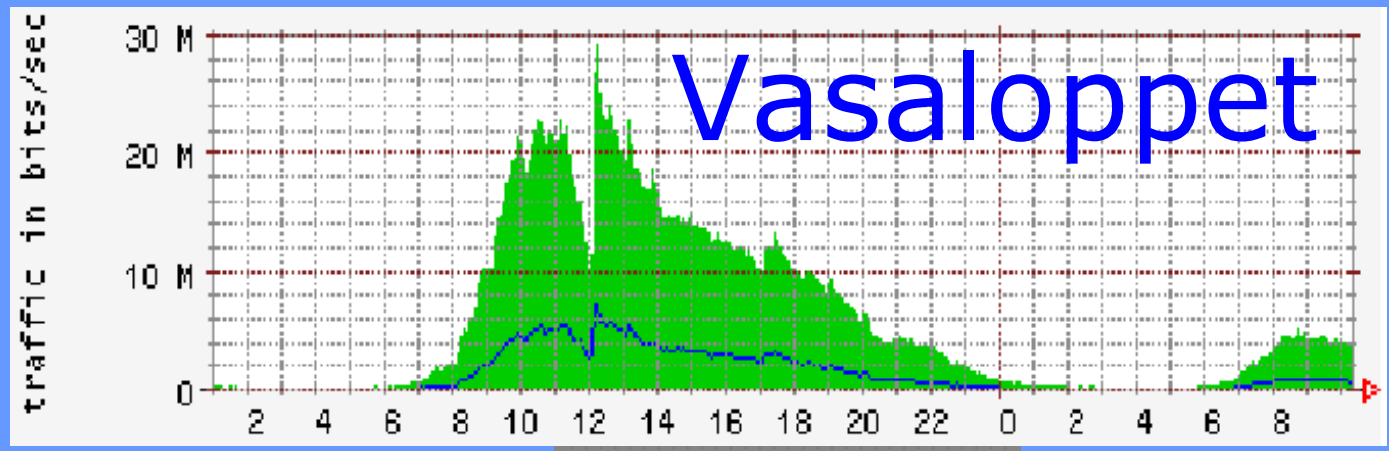
Neptune – Undersea Grid



Neptune
Fiberoptic Telescope to Inner Space

- NEPTUNE's 3000 kilometers of fiber-optic cable will provide power and communications to scientific instruments.
- The system will provide real-time flow of data and imagery to shore-based Internet sites. It will permit interactive control over robotic vehicles on site and will provide power to the instruments and vehicles. NEPTUNE may also serve as a unique testbed for sensor and robotic systems designed to explore other oceans in the solar system.
- The NEPTUNE network is expected to be operational by 2006 and will cost approximately $250 million to develop, install, and operate through the first five years.
- The NEPTUNE system must have the following characteristics to meet the scientific requirements:
  - Plate scale (covering the full Juan de Fuca tectonic plate)
  - Power (order of tens of kW)
  - Bandwidth (order of many Gbits/sec)
  - Real-time data return and robotic control capability
  - Robust design for high reliability
  - Precision timing at all instruments
  - Available for nominal 20-30 years.

Nordic Grid Consortium



~68.4 milljon hits during the race
~37 thousand simultaneous connections peak
peak rates: 29 Mbps out, 7 Mbps in
peak IP addresses, 165,915
peak one day hits 32,155,188
peak 5 minutes 446,695 hits
160,632 GBytes

http://www.vasaloppet.se

Nordic Grid Consortium

Objectives

- Provide the best possible environment and resources for data and compute intensive research and education
- Provide training and user support across a broad spectrum of applications, computing platforms, data resources, and visualization systems
- Provide convenient, transparent, secure access to eligible users
- Research and development of software systems for enhanced user services and efficient use of Grid resources
- Be a resourceful partner in European and intercontinental endeavors for research, development and deployment of integrated large-scale data, computing and visualization resources and systems serving academia and its partners

Nordic Grid Consortium

Founding Partners

- CSC, the Finnish Ministry of Education's National Center for High-Performance Computing and Networks. http://www.csc.fi

- Parallab, University of Bergen's Computational Science and High-Performance Computing Laboratory. http://www.parallab.uib.no

- PDC (ParallalDatorCentrum), Royal Institute of Technology (KTH), Stockholm, Sweden, the lead national center serving Swedish academic research and higher education funded in part by the Swedish Research Council. http://www.pdc.kth.se

Resources

- Hardware
- Software
- Data
- Staff

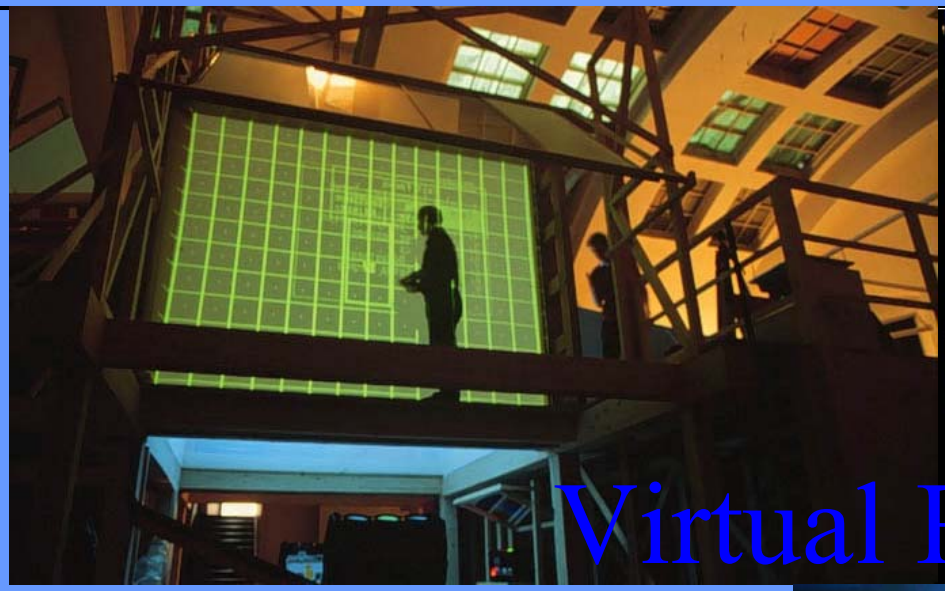
Nordic Grid Consortium

Type of System	GFlop/s / GByte			
	CSC	Parallab	PDC	Total
DM	523 / 90		84 / 36	607 / 127
SMP	2445 / 590	499 / 197	187 / 106	3131 / 892
DSM	79 / 168	25 / 12	4 / 4	108 / 184
V/P			7 / 6	7 / 6
PC-Cluster		81 / 66	273 / 82	354 / 147
Total	3047 / 848	605 / 274	554 / 234	4207 / 1356

Type of System	GByte Disk (global/local)			
	CSC	Parallab	PDC	Total
DM and SMP	2,100	5,800	1 584 / 704	9 484 / 704
DSM	560	35	200 / 594	795 / 594
V/P			128	128
PC Cluster			- / 3 288	- / 3 288
Total	2,660	5,835	1 912 / 4 586	10 407 / 4 586

Type of System	TByte on Tape			
	CSC	Parallab	PDC	Total
Archive	10	32	5	47
Backup	18		20	38
Total	28	32	25	85

Nordic Grid Consortium



Virtual Reality

Nordic Grid Consortium

Nordic Grid Consortium

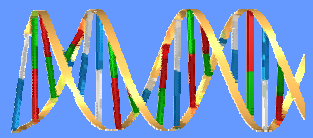
Engineering Software	CSC	Parallab	PDC
ABAQUS/Standard	■		■
ABAQUS/Explicit	■		■
ADAMS	■		
Ansys	■		
CFX	■		
Elmer	■		
FIDAP	■		
Fluent	■		■
Linflow	■		
Opera-3d	■		

Geoscience Software	CSC	Parallab	PDC
Arpege		■	
BOM		■	
GrADS	■		
HYCOM		■	
Melts	■		
MICOM		■	
SU		■	
UNCERT	■		

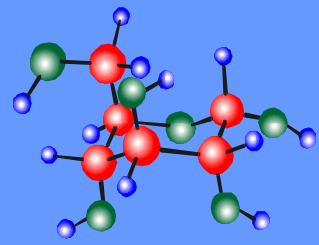
Numerical Libraries	CSC	Parallab	PDC
ARPACK	■		■
BLACS	■	■	■
BLAS	■	■	■
DXML	■		
ESSL	■	■	■
FFTW			■
IMSL	■		
ITPACKV	■		
LAPACK	■		■
MASS		■	■
METIS			■
NAG	■		■
NAG FL90plus	■		■
netCFD			■
NSPCG	■		
OSL		■	
ParMETIS			■
P_ARPACK	■		■
PBLAS	■	■	■
PDE2D	■		
PESSL	■	■	■
PETSc			■
PSPARSLIB			■
PWSSMP	■	■	■
ScaLAPACK	■		■
SCSL	■		
SPARSKIT			■
SPASPAK	■		
SSL II/VP			■
SSL II/VPP			■
WSSMP	■	■	■

Nordic Grid Consortium

Data Bases



- **Bio/Life-sciences**
  - databases for sequences of DNA and amino acids
  - database of macromolecule construction
  - gene maps
  - Coming up:
    - genomes
    - 3D-databases
    - gene expression databases
    - brain image databases

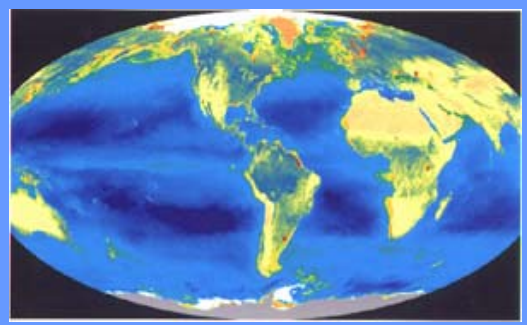


- **Chemistry**
  - MDL Cross-Fire
    - database of chemistry and material sciences
    - the world's most extensive electronic database of organic chemistry
  - Cambridge Structural Database System
    - databases of compound constructions
  - SpecInfo
    - theworld's largest electronic spectra database
  - MDL Databse
    - for drug development
    - for preparation of chemicals

Nordic Grid Consortium

Data Bases

- Geosciences
  - topography models and databases (e.g. GLOBE, Terrainbase
  - characteristics databases (e.g.data on flora, hydrology, glacier, soil, and bedrock)
  - image databases (e.g. satellite images)
  - map databases (basic geographic maps of Finland)



- Language Bank
  - Corpus bank of Finnish texts
    - research material of nearly 180 million words.words of Contemporary Finnish
  - Corpus of Middle French
  - Corpus of contemporary English

Nordic Grid Consortium

Staff

CSC

Systems - 40

Support - 40

Development – 20

Admin/Doc - 20

Total - 120

PDC

Systems - 9

Support - 7

Development - 4

Admin/Doc - 3

Total - 23

Parallab

Systems - 2

Support - 8

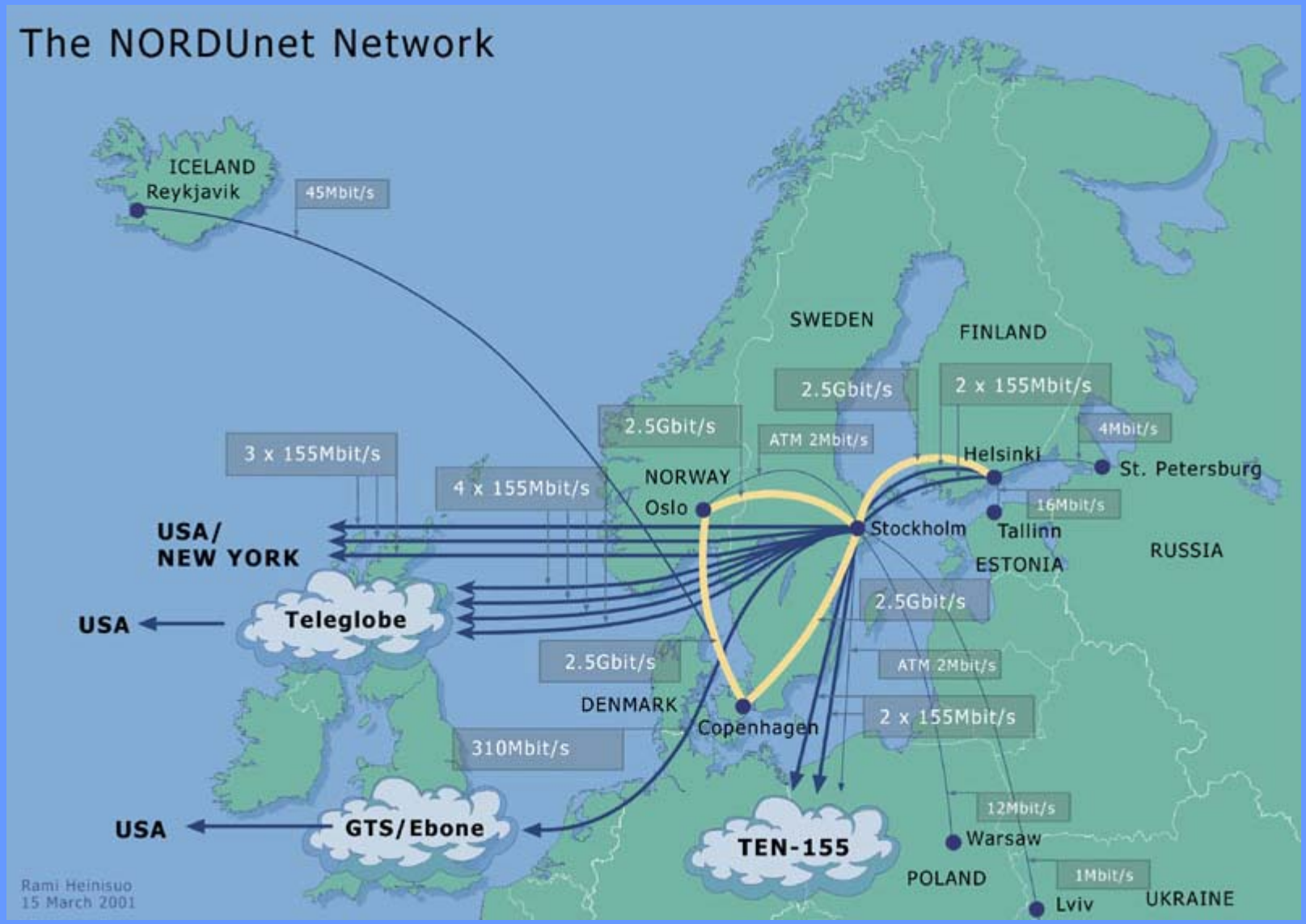
Development - 8

Admin/Doc - 2

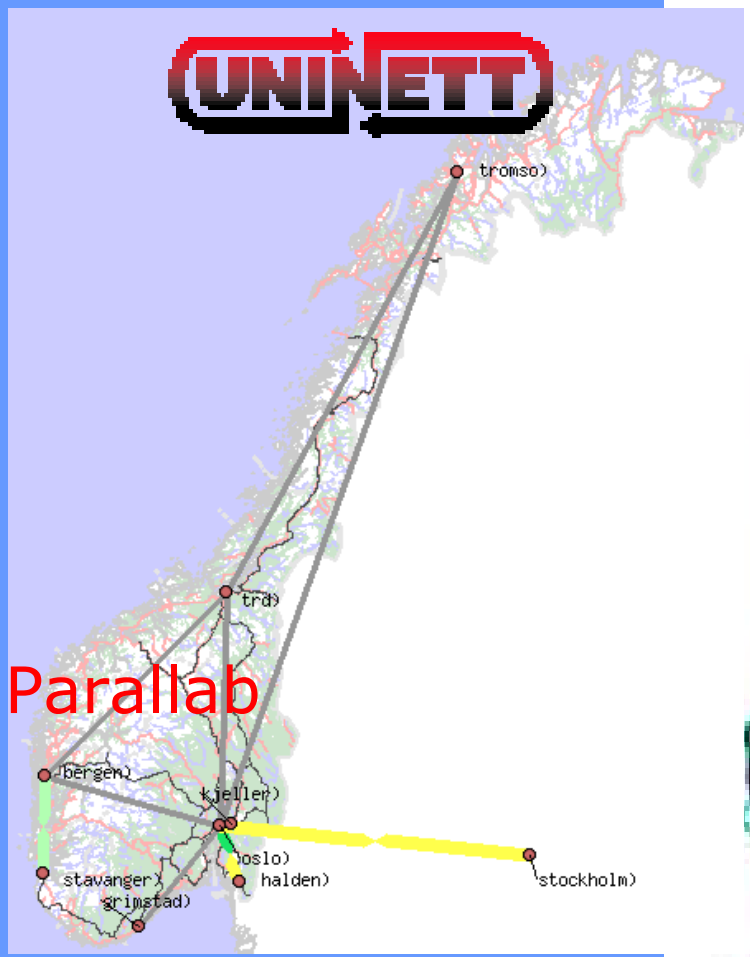
Total - 20

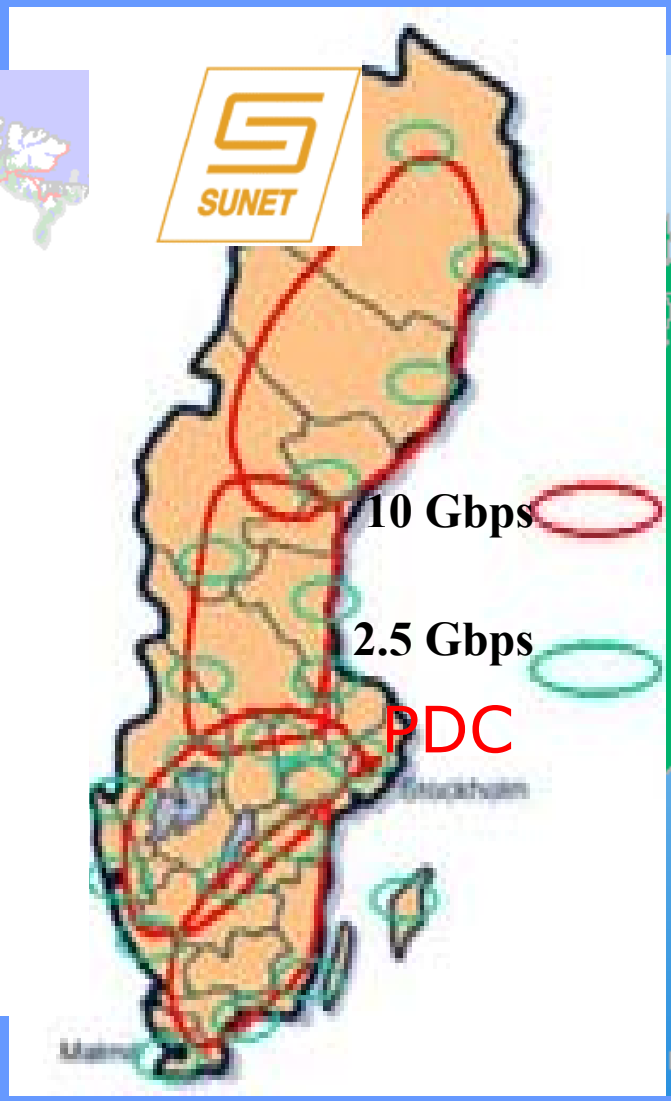
Network environments

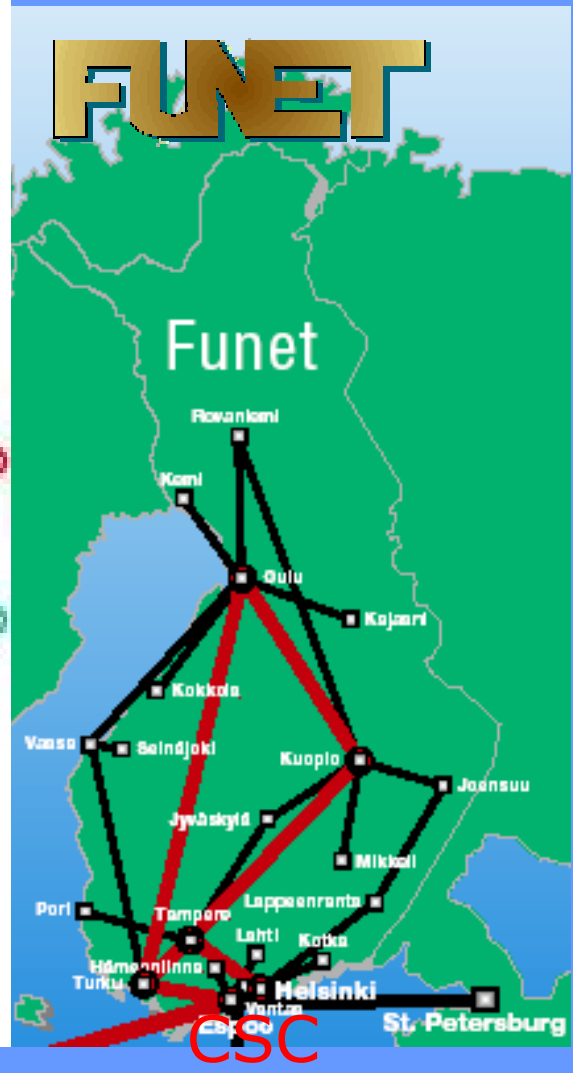
Nordic Grid Consortium



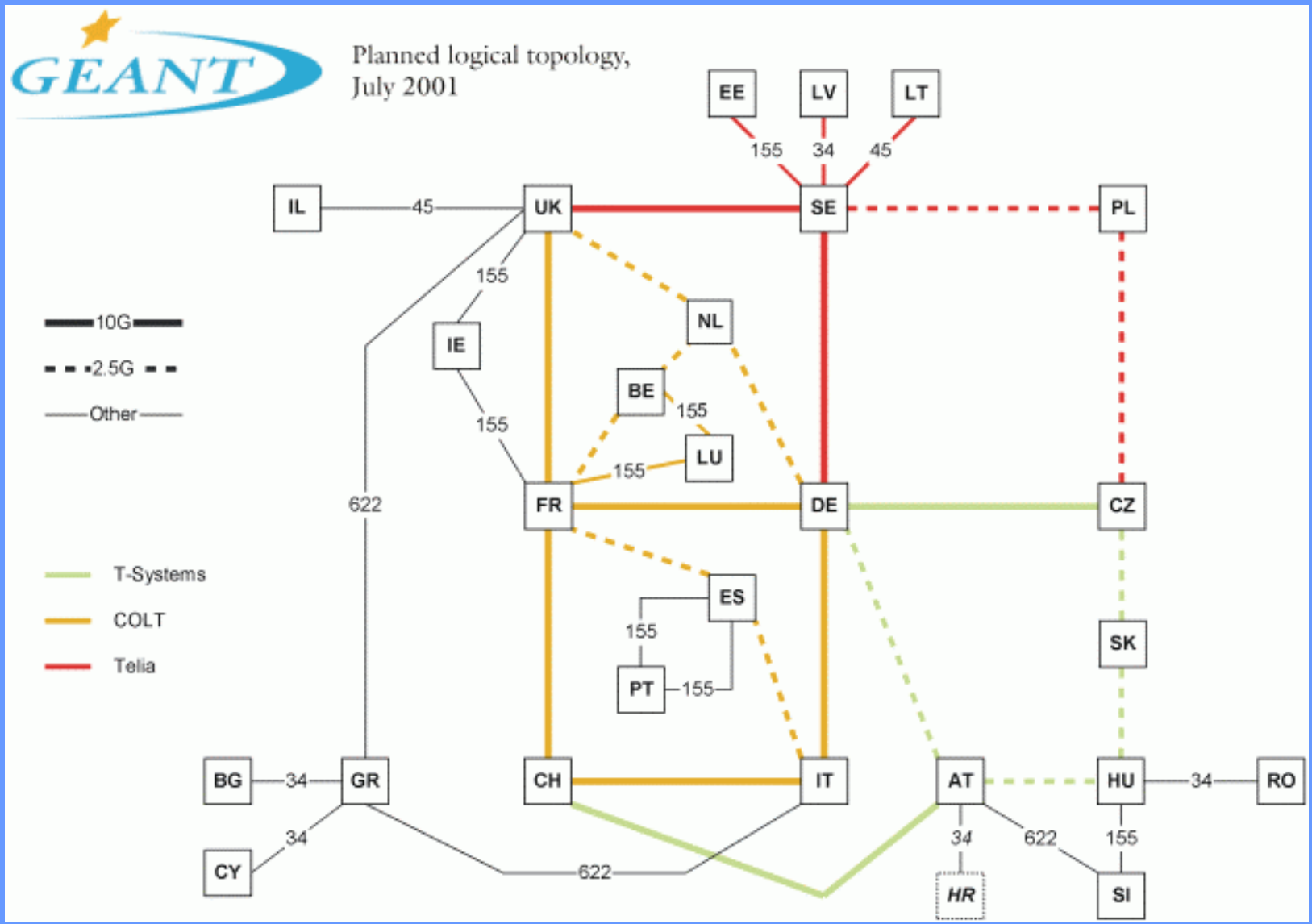
Nordic Grid Consortium





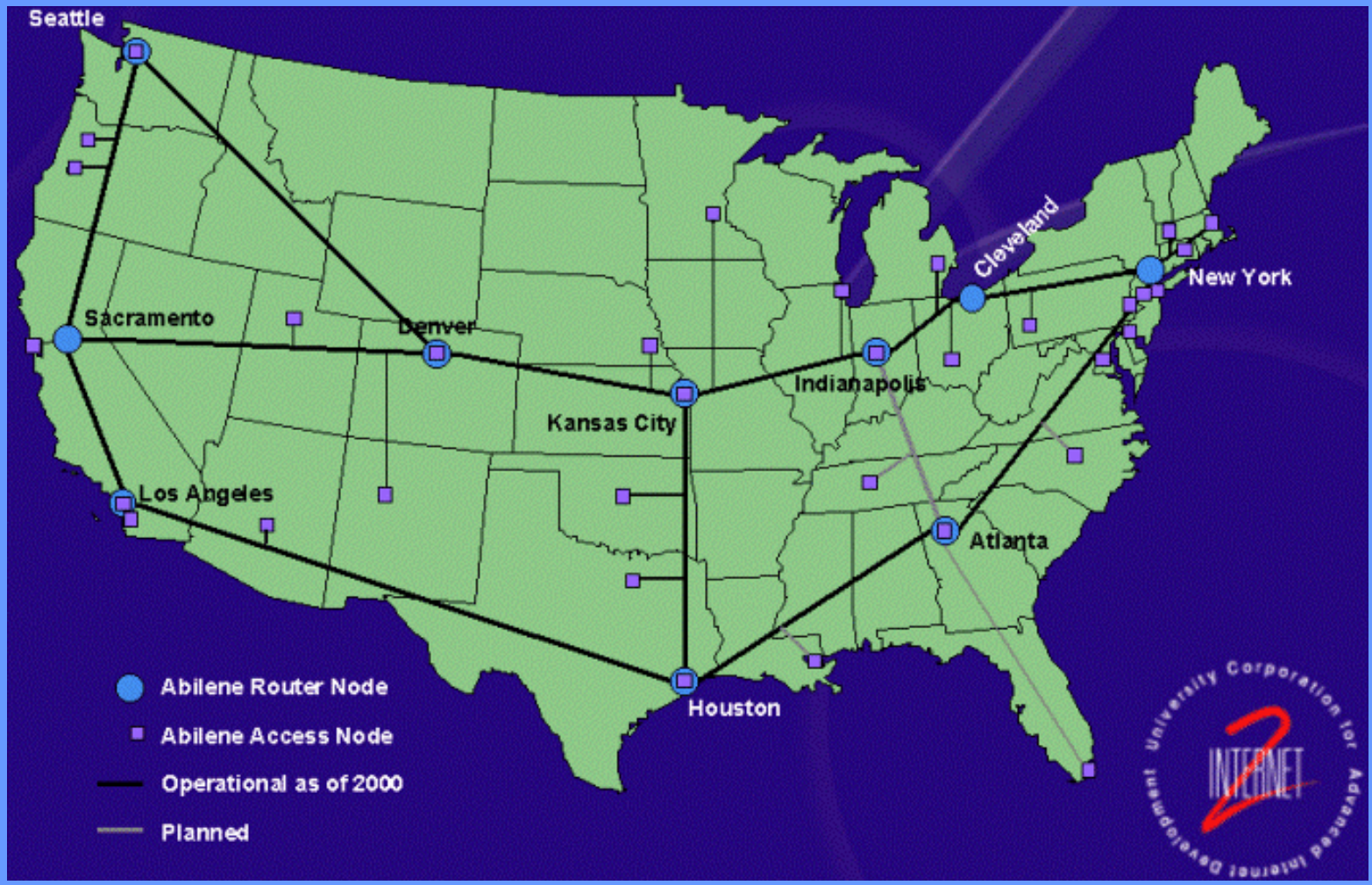


Nordic Grid Consortium

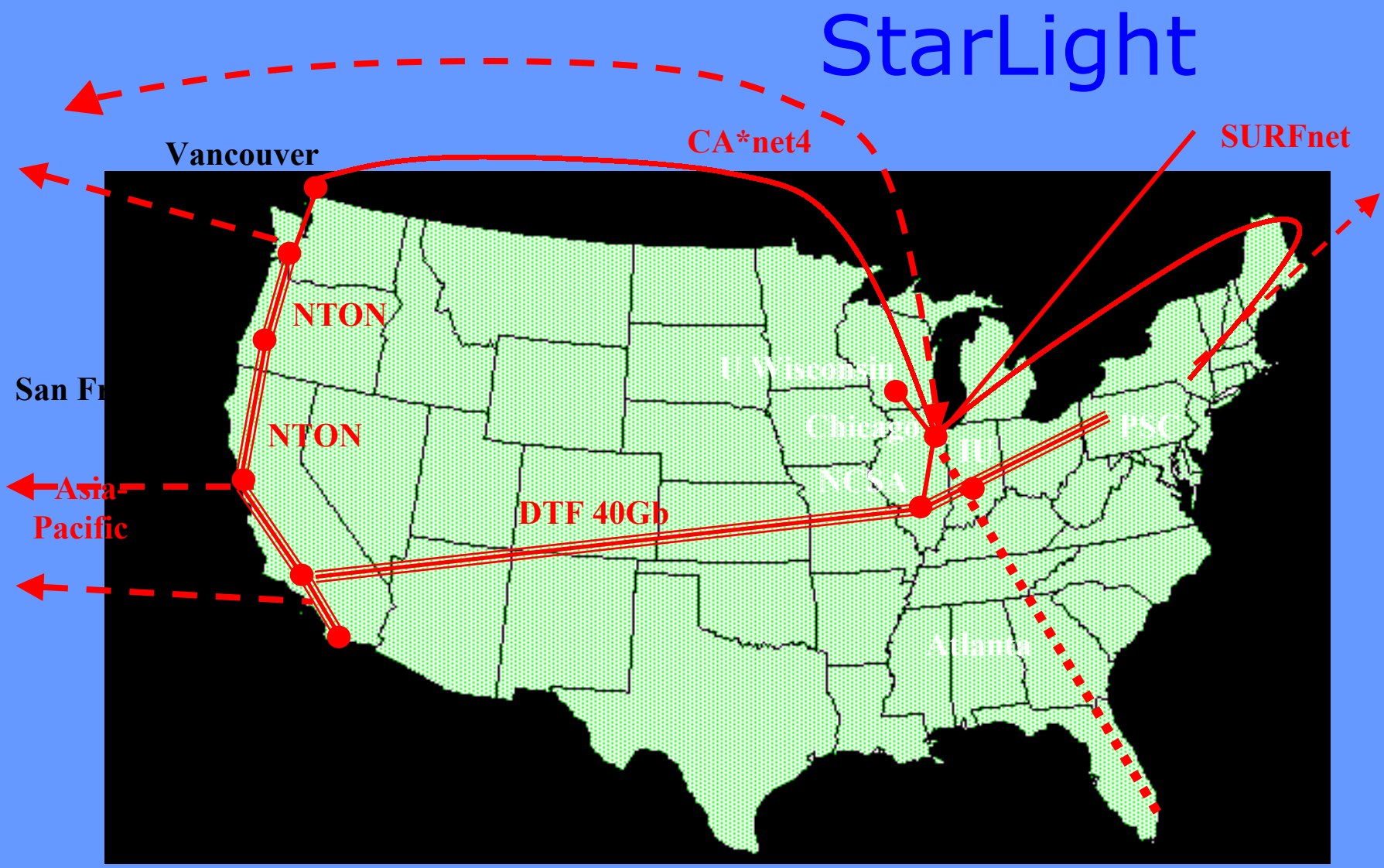


Nordic Grid Consortium

Internet2 - Abilene

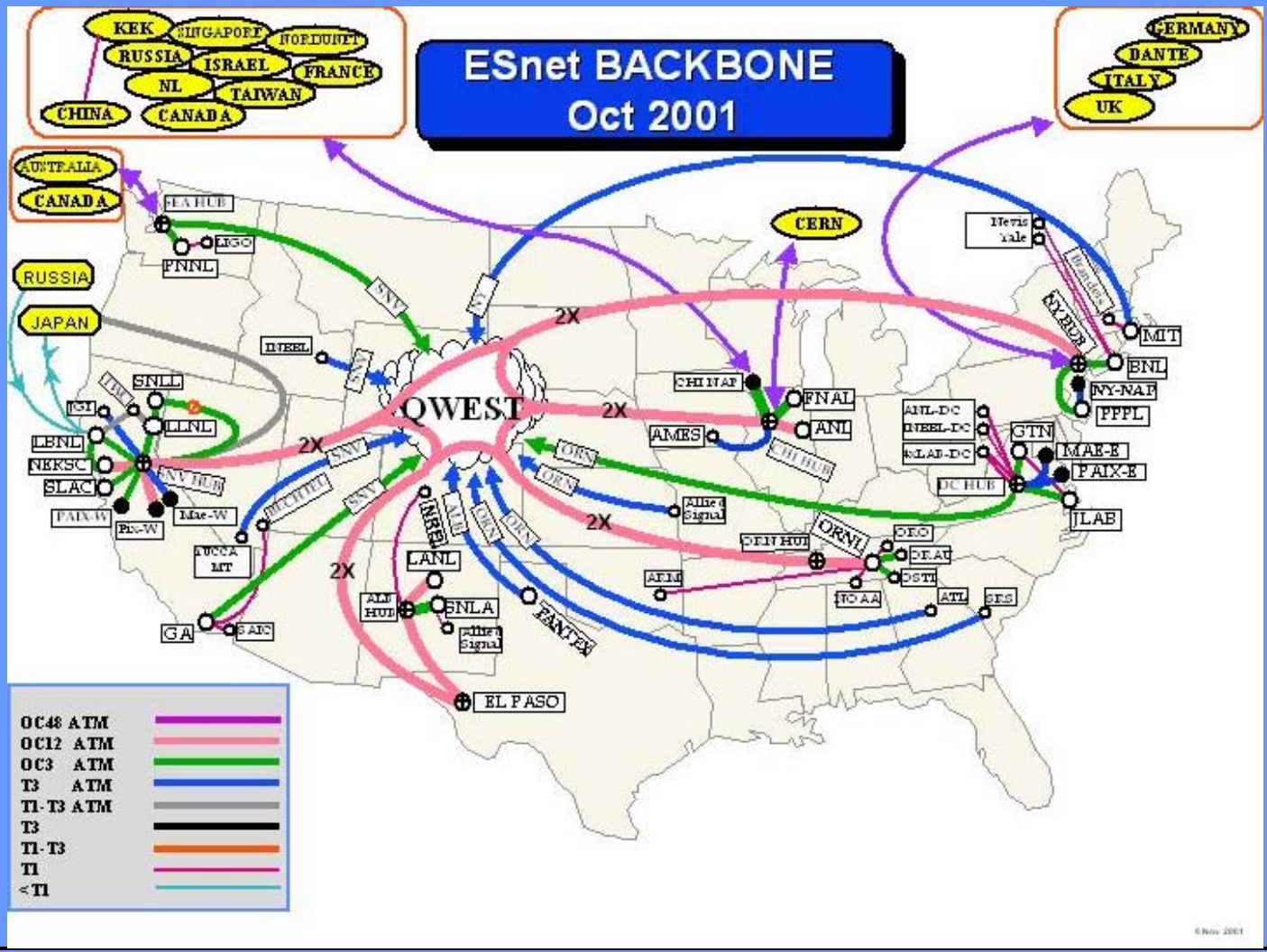


Nordic Grid Consortium

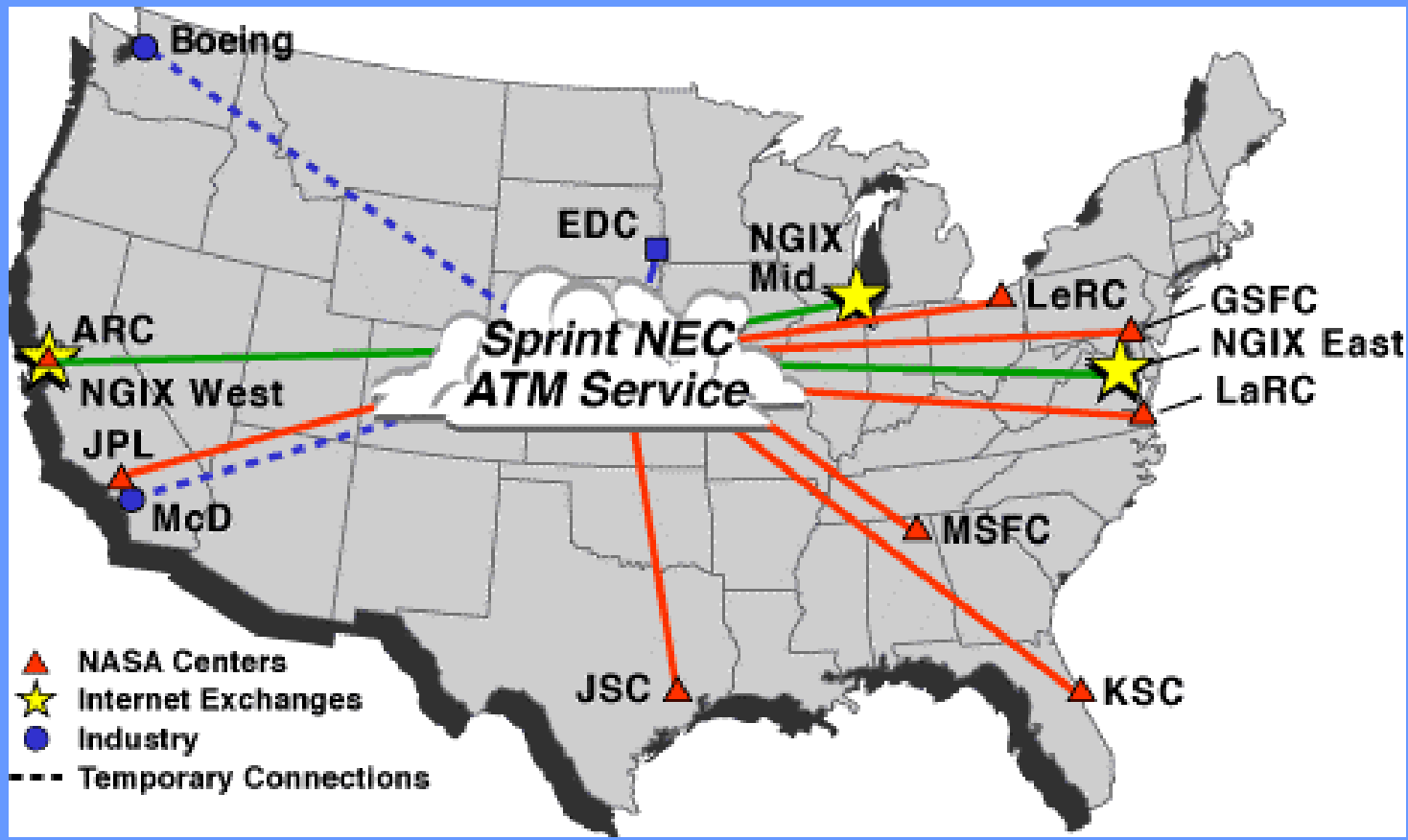


Nordic Grid Consortium

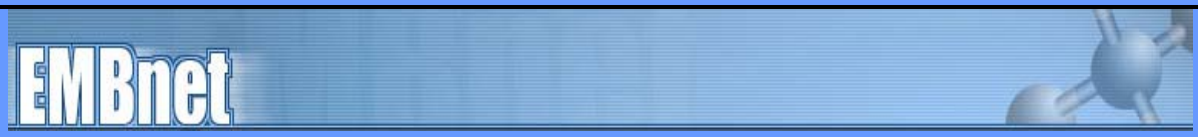
DoE Energy Sciences Network



NASA Research and Education Network (NREN)



Nordic Grid Consortium



GrADS
Grid Application Development Software Project

6NET

EUROMAP
LANGUAGE TECHNOLOGIES

HPCN TTN
network

N'S³
ENScube

JaCo3

Sharing the future
nedlib

enacts

DataTAG

NPACI

CAVERN
互

EUROGRID

NeuroGenerator NG
- a databasegenerator for the neuroimaging community

Data GRID

Nordic Testbed for Wide Area Computing and Data Handling

ALLIANCE

FINNOVA
TECHNOLOGY TRANSFER NODE AT CSC

VIRTUAL FIRES

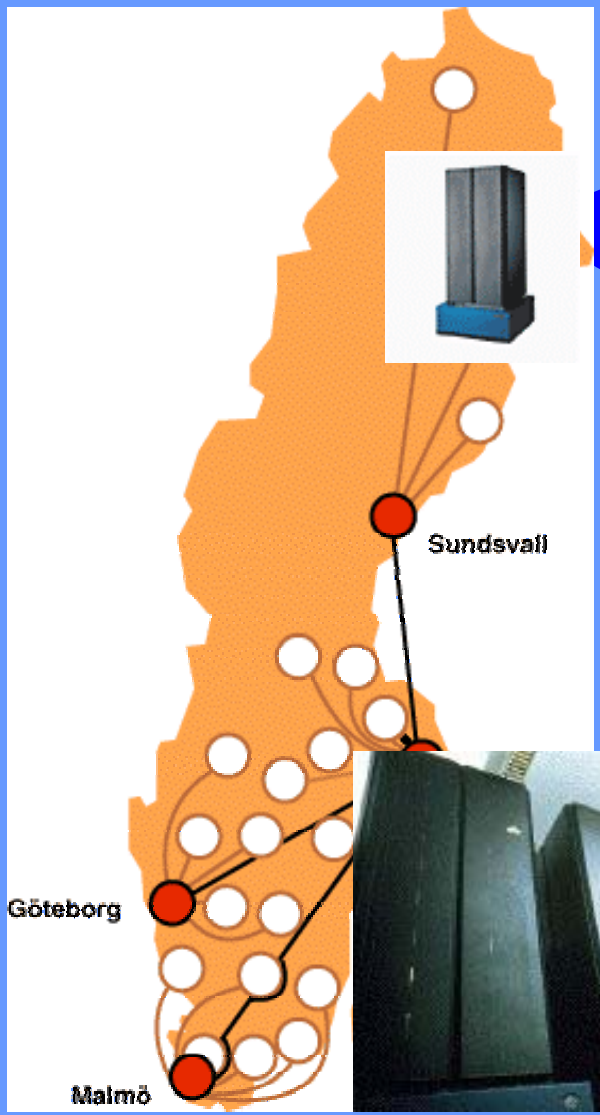
Nordic Grid Consortium

Plans

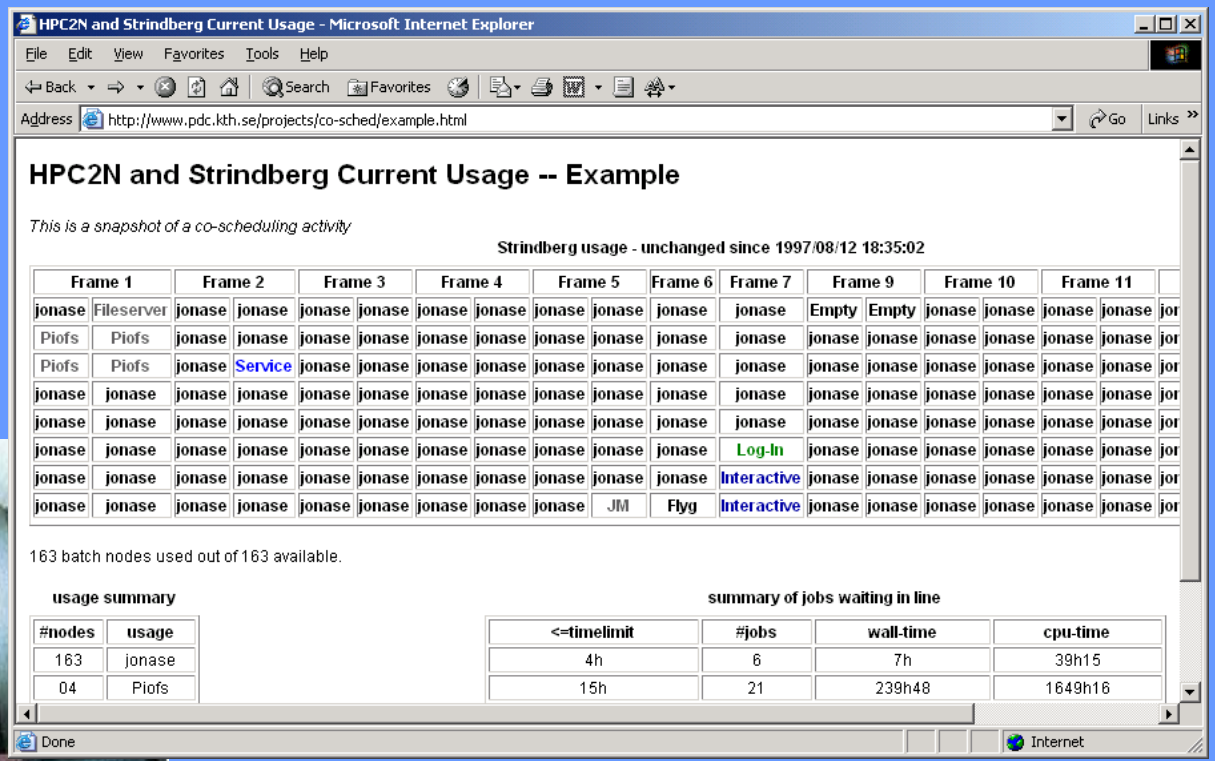
- Work with sponsoring agencies to facilitate resource sharing
- Set up a NGC portal
- First edition shared software environment by end of April (Globus)
- “Friendly” users in summer of 2002
- Account and user access managment
- Limited service in accordance with sponsor policies by the end of 2002

Acknowledgements

Nordic Grid Consortium

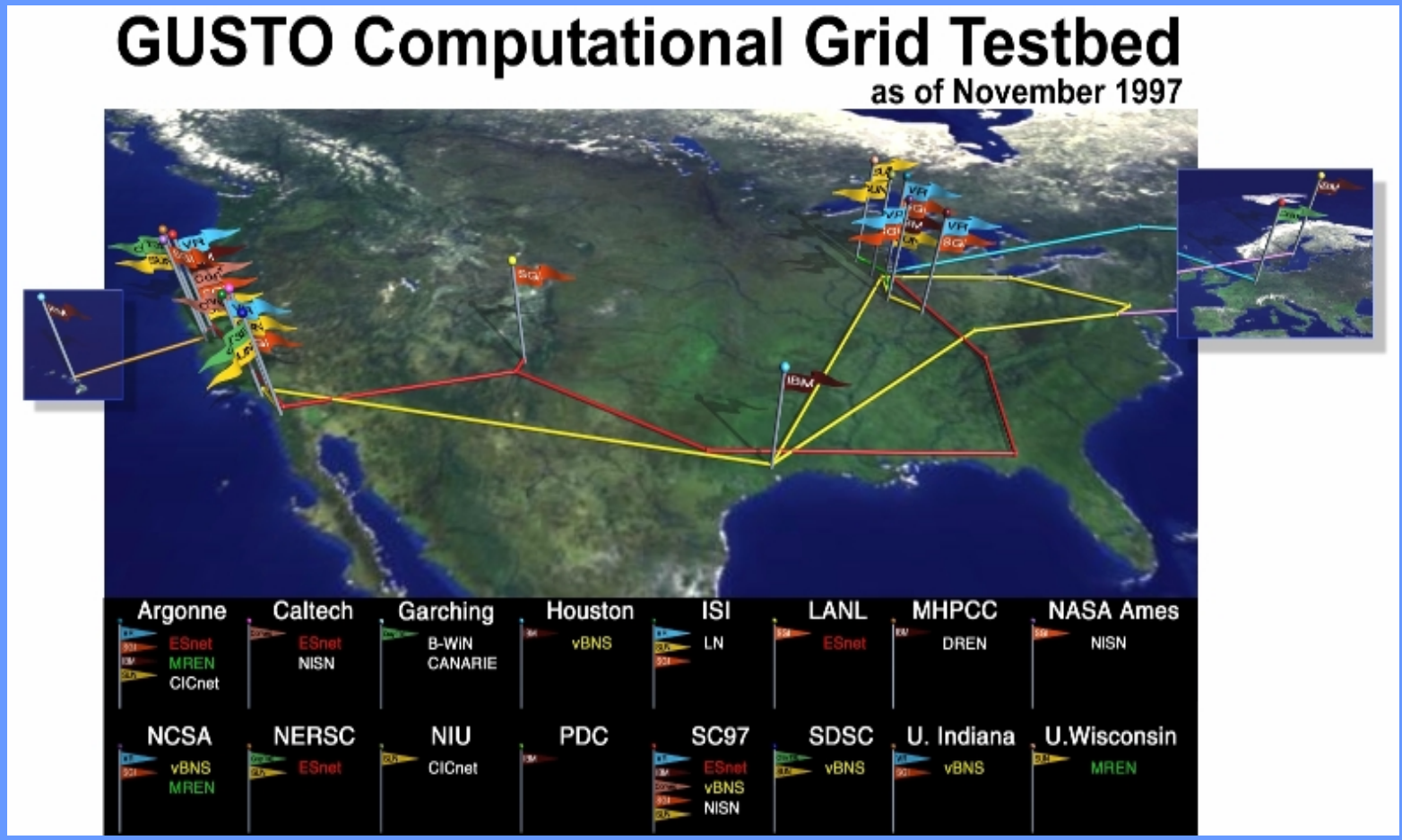


Co-Scheduling of jobs between HPC2N and PDC



http://www.pdc.kth.se/projects/co-sched/

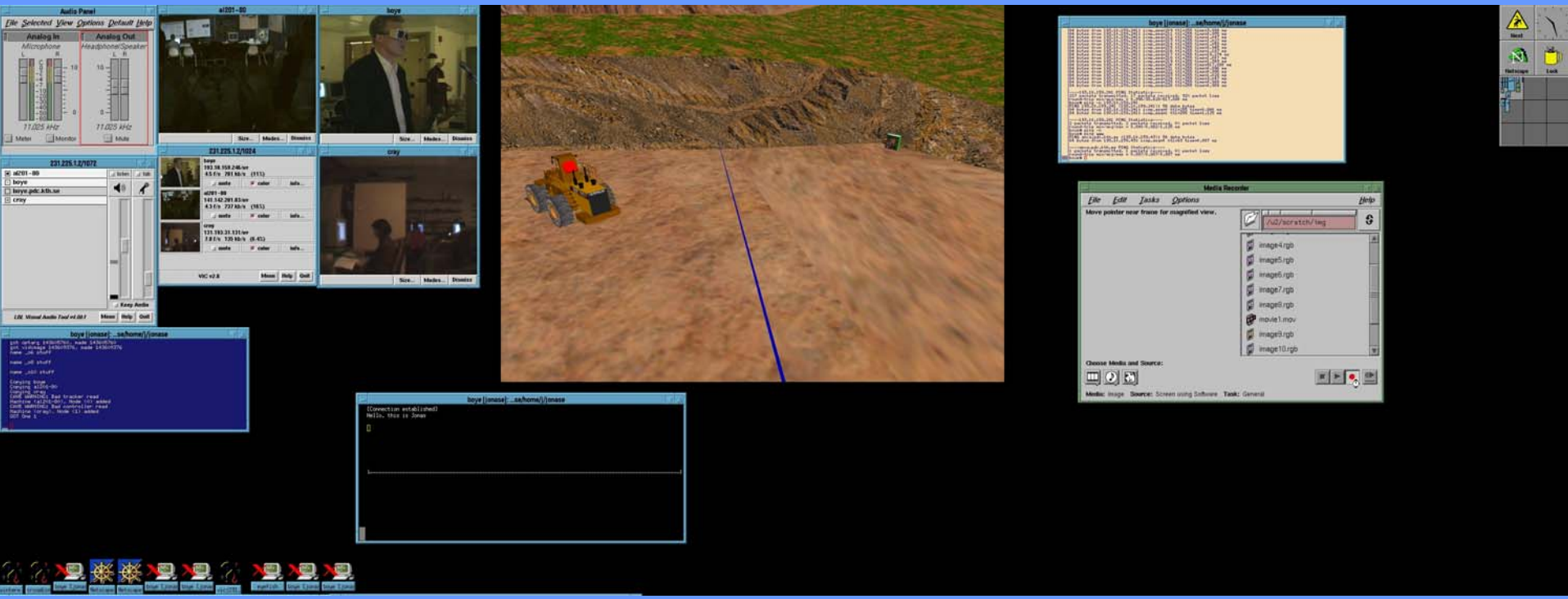
Nordic Grid Consortium



GUSTO tesbed for Grid applications demonstrated at Supercomputing97 exhibition

Nordic Grid Consortium

Alliance98

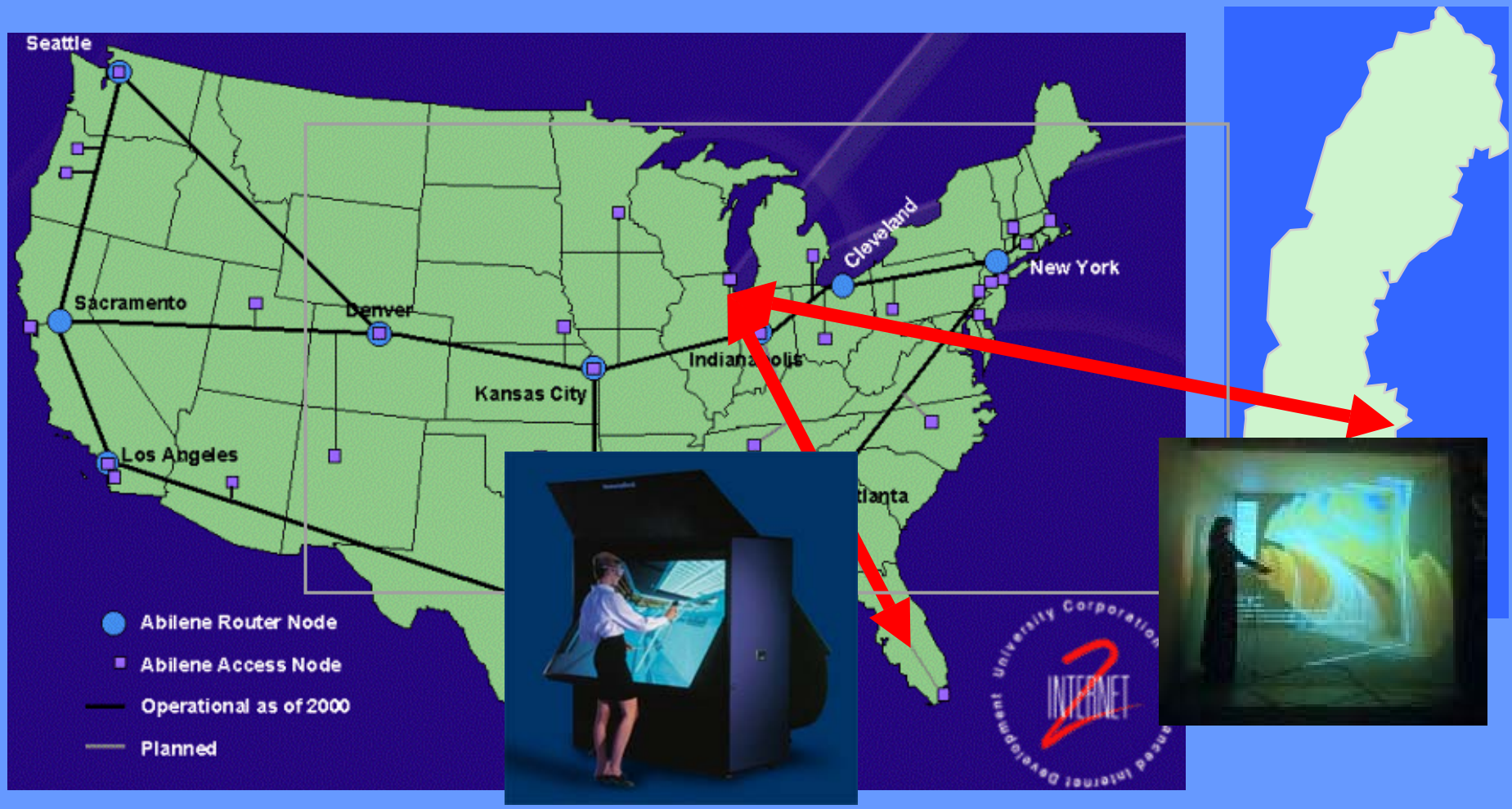


Interactive Collaborative Virtual Environment

http://www.pdc.kth.se/projects/alliance98/

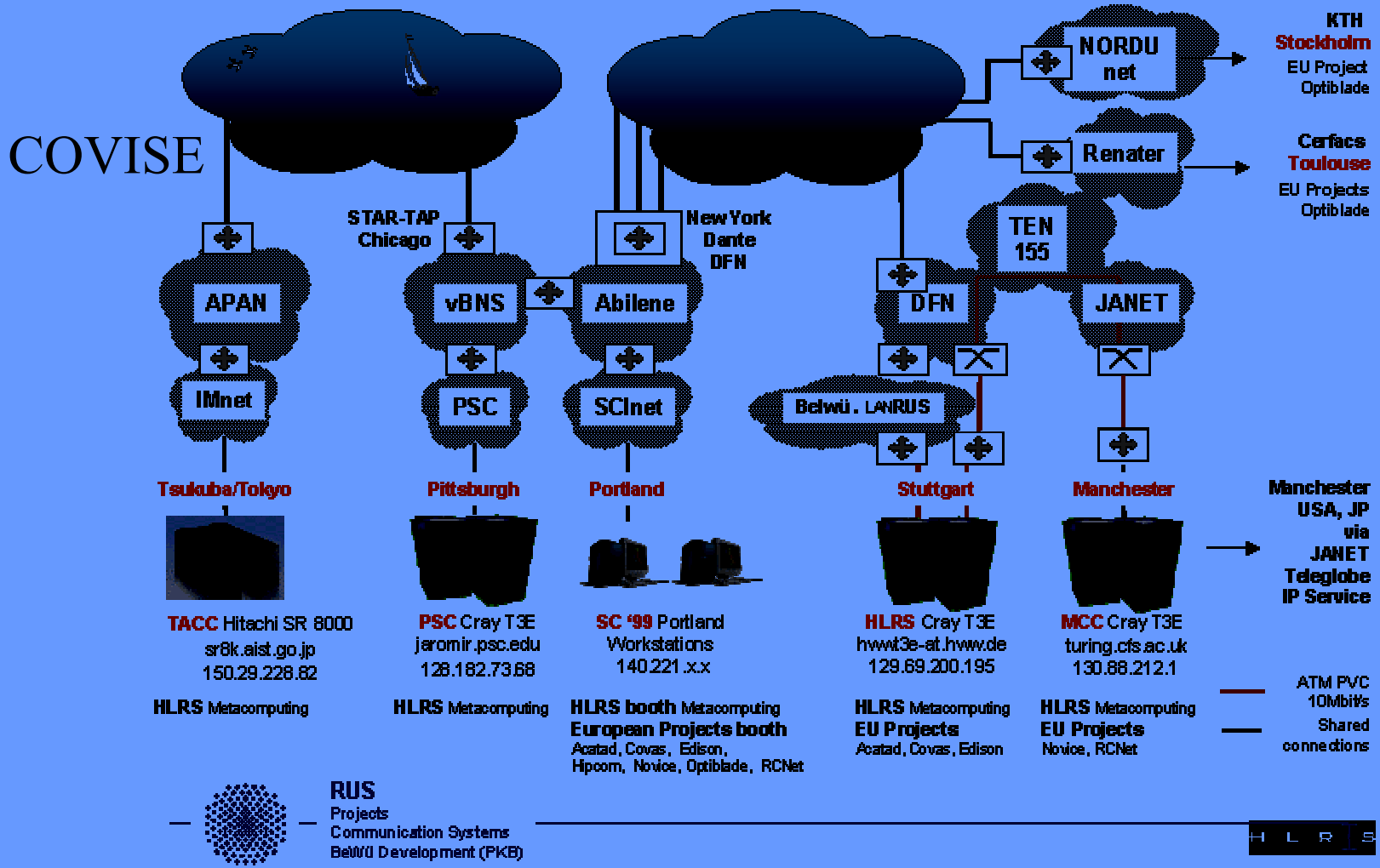
Nordic Grid Consortium

EnVis@SC98



Nordic Grid Consortium

SC '99



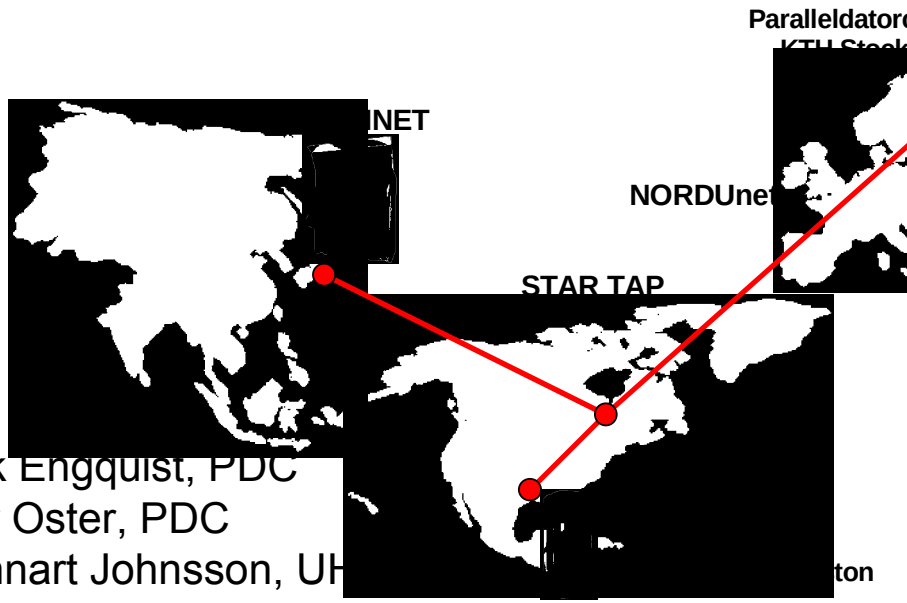
Nordic Grid Consortium

Computational Steering

the globus project
www.globus.org

GEMSviz at iGRID 2000

NORDUnet



SWEDISH UNIVERSITY COMPUTER NETWORK
SUNET • Karta • Sunetten • Universitet och högskolor
Webbkatalogen • Filarkivet • Filsökning • E-postkatalog
SUNET

Erik Engquist, PDC
Per Oster, PDC
Lennart Johnsson, UH

iGrid2000
Empowering Global Research Community Networking
www.startap.net/igrid2000
University of Illinois at Chicago and Indiana University
in collaboration with Tokyo University and Keio University

APAN Asia-Pacific Advanced Network

STAR TAP SM

The END