

New Internet and Networking Technologies for Grids and High-Performance Computing



Stanford, USA
August 25th, 2006

C. Pham

<http://www.univ-pau.fr/~cpham>
University of Pau, France
LIUPPA laboratory

1

Computational Sciences

- Use of computers to solve complex problems
 - Modeling techniques
 - Simulation techniques
 - Analytic & Mathematic methods
 - ...
- Large problems require huge amount of processing power: supercomputers, high-performance clusters, etc.

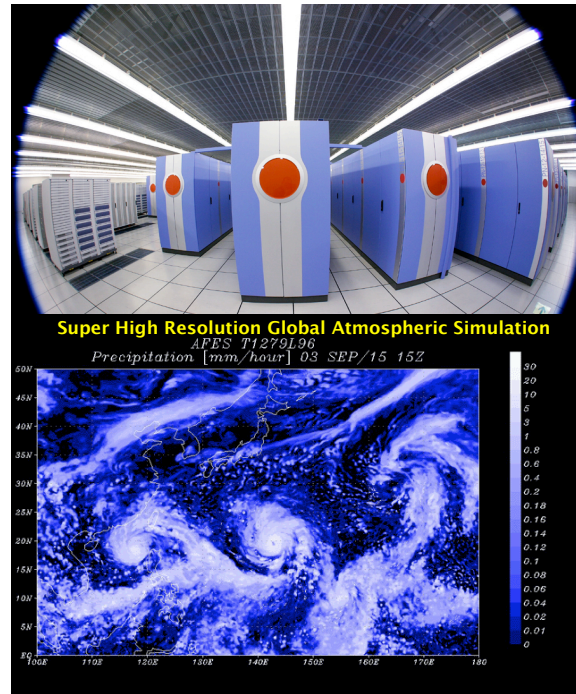
2

Earth Simulator

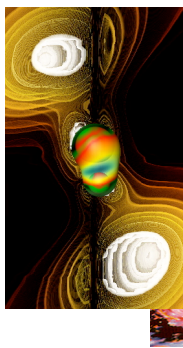
©JAMSTEC

No. 1 position for five consecutive TOP500

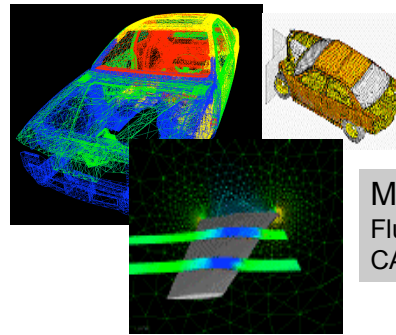
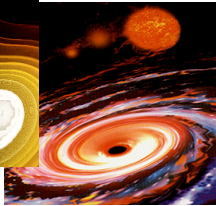
- Intensive numerical simulations
- Ex: Super High Resolution Global Atmospheric Simulation



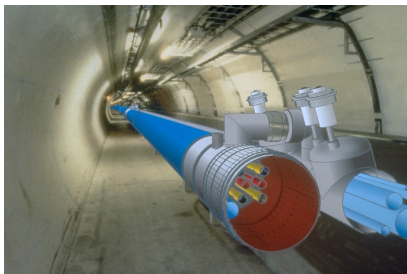
A large variety of applications



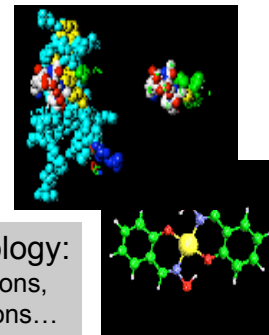
Astrophysics:
Black holes,
neutron stars,
Supernovae...



Mechanics:
Fluid dynamic,
CAD, simulation.



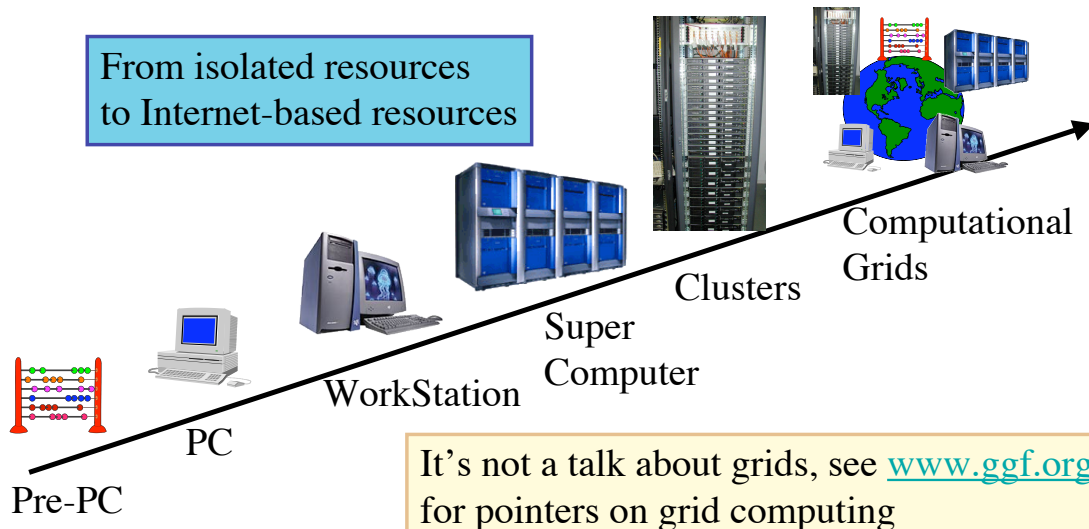
High-Energy Physics:
Fundamental particles of matter,
Mass studies...



Chemistry&biology:
Molecular simulations,
Genomic simulations...

This talk is about...

- How the Internet revolution could be beneficial to computational sciences



Purpose of this tutorial

- Audience
 - Scientists/students from parallel, distributed, computer or grid and computational sciences
- Purpose
 - Provides a comprehensive survey of advanced networking technologies
- Expected results
 - Understanding of why the network is important in a grid infrastructure
 - Knowledge of current advanced technologies for decision making processes

Layout explanation

N
E
W

C
H
A
P
T
E
R

□ Body text

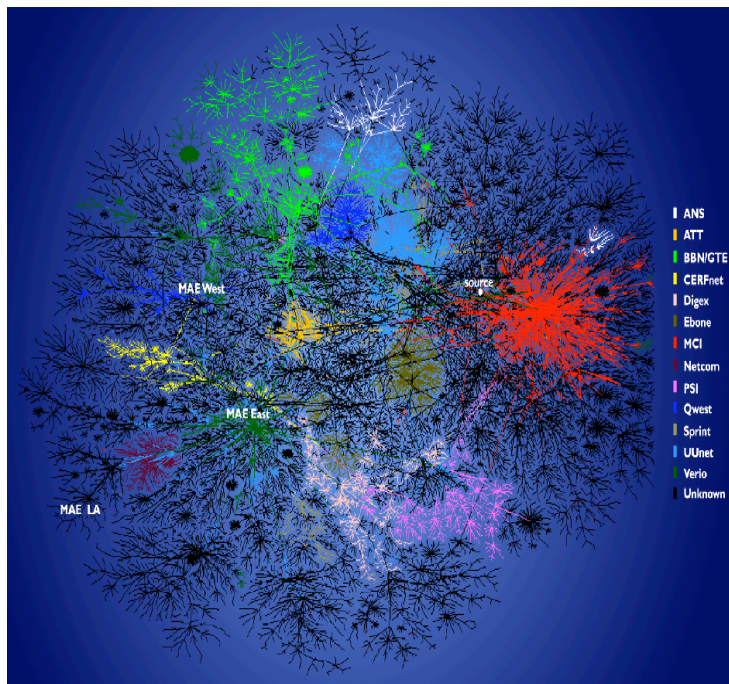
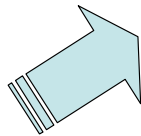
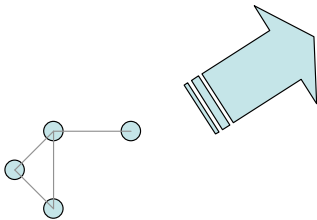
← Indicates a new chapter

↘ Indicates which chapter the slide belongs to

Introduction

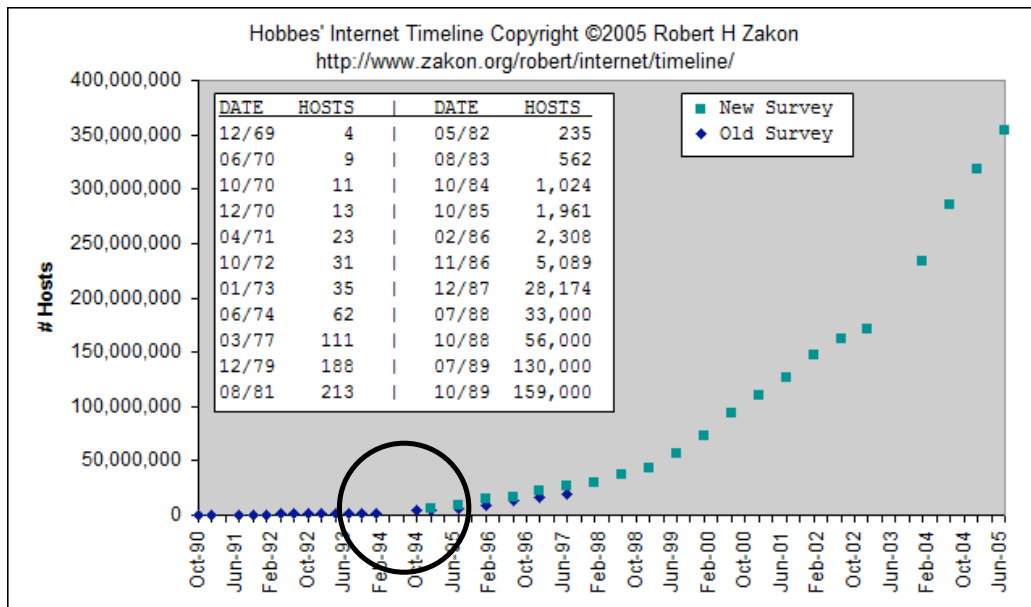
7

The big-bang of the Internet

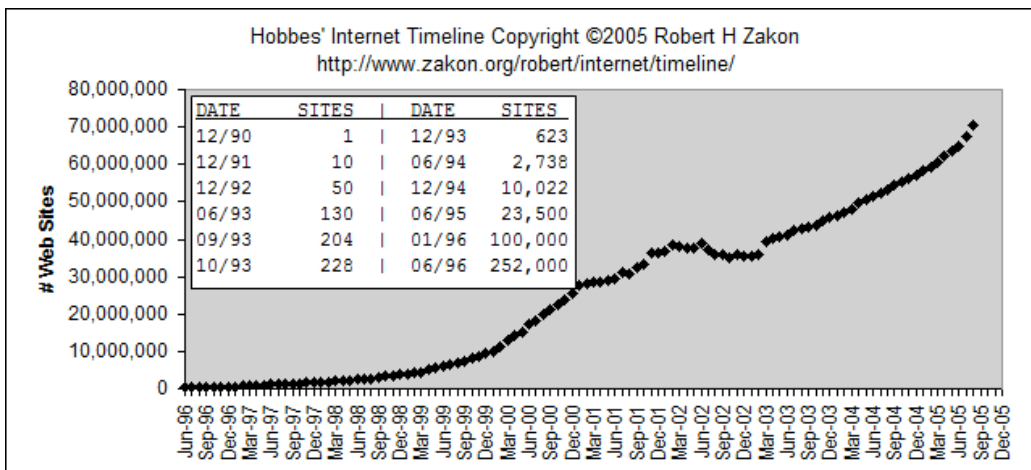


Introduction

Internet host

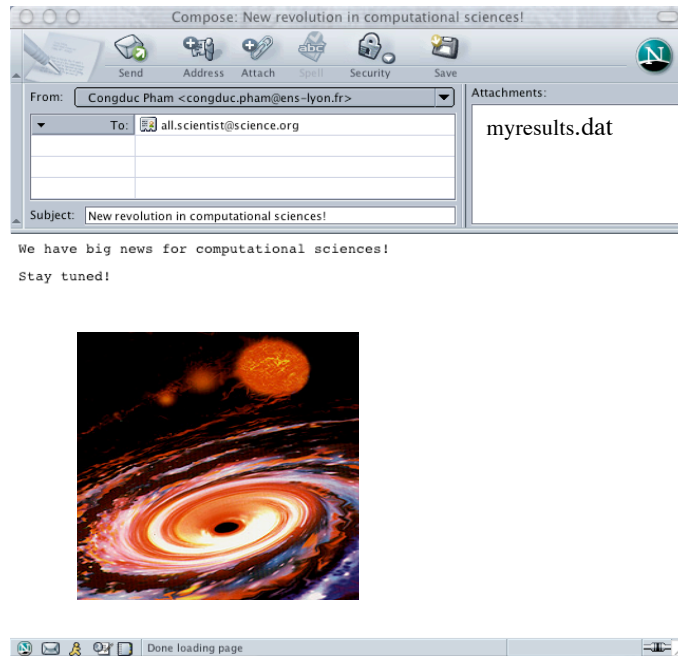


www.web-the-big-bang.org



Internet usage: e-mail...

- ❑ Convenient way to communicate in an informal manner
- ❑ Attachments as a easy way to exchange data files, images...



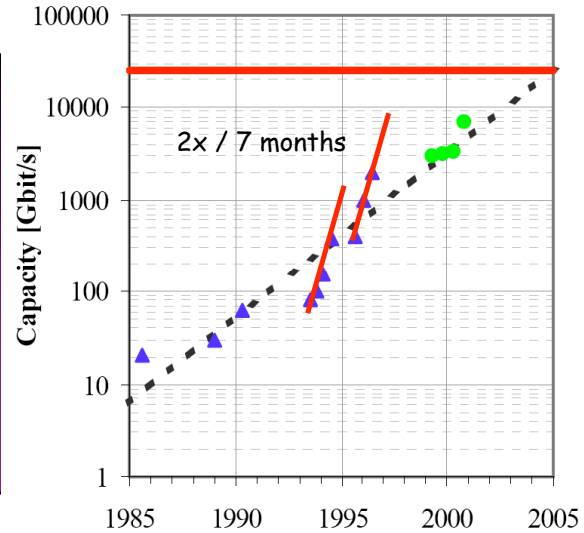
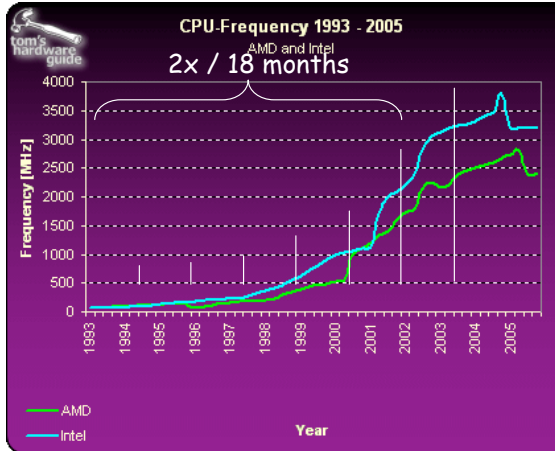
...and surfing the web

- ❑ A true revolution for rapid access to information
- ❑ Increasing number of apps:
 - ❑ e-science,
 - ❑ e-commerce, B2B, B2C,
 - ❑ e-training, e-learning,
 - ❑ e-tourism
 - ❑ ...





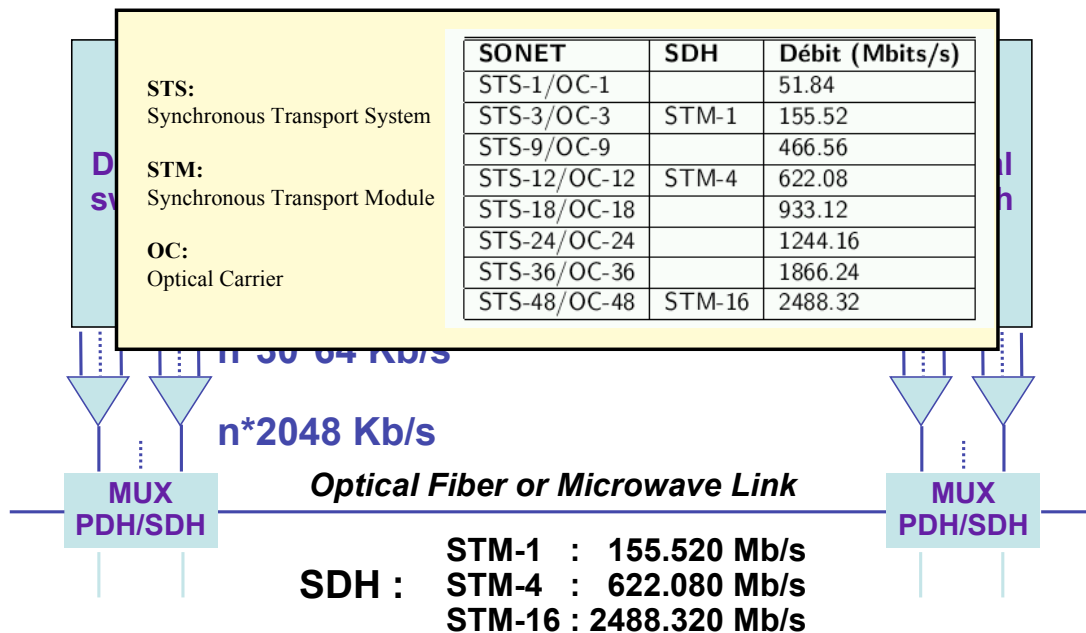
The optical revolution!



Source « Optical fibers for Ultra-Large Capacity Transmission » by J. Grochocinski

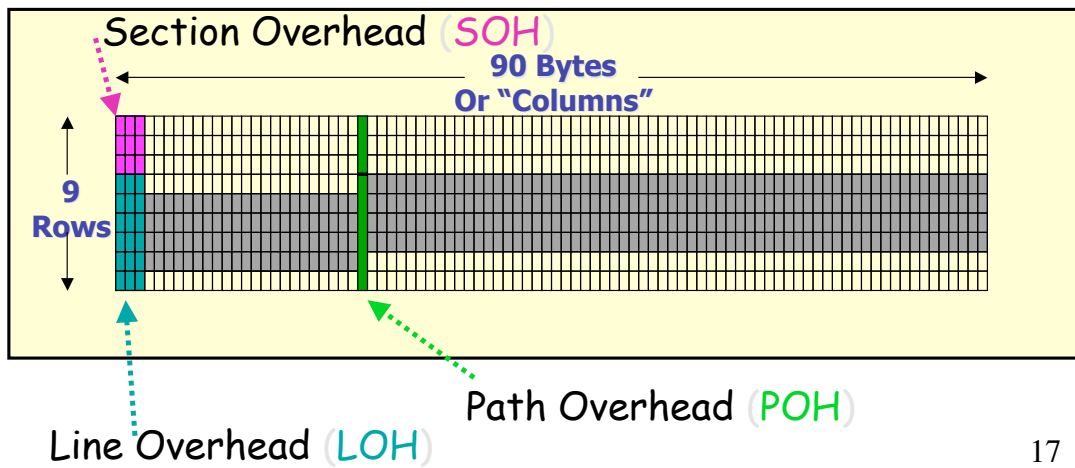
SONET/SDH in the core

95% of exploited OF use SONET/SDH



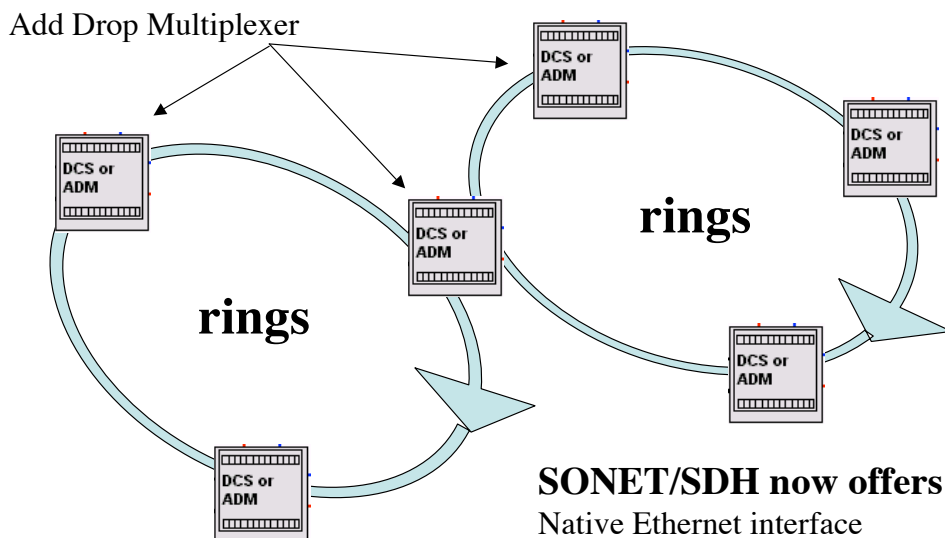
The SONET frame

- Basic frame length is 810 bytes (TDM)
 - Sent every 125us, raw throughput of 51.84 Mbits/s (STS-1)
 - Better seen as a block with 90 columns and 9 lines



17

SONET/SDH transport network infrastructure



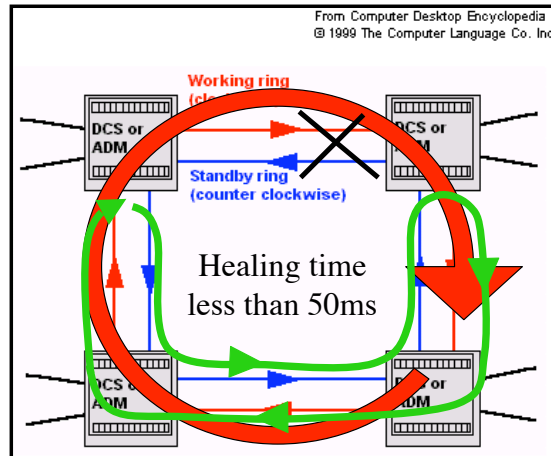
SONET/SDH now offers
 Native Ethernet interface
 Generic Framing Procedure
 Virtual Concatenation

18

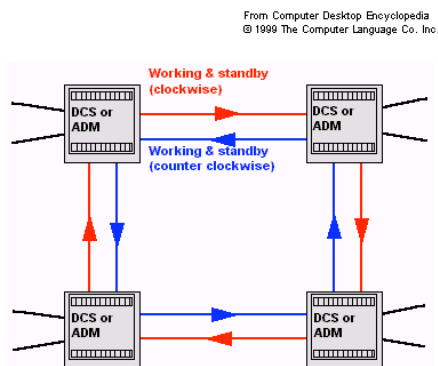
SONET/SDH and resiliency

- SONET/SDH has built-in fault-tolerant features with multiple rings
- Ex: simple case

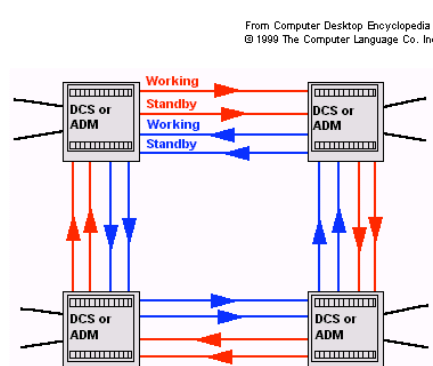
DCS
(Digital Cross-Connects)



High availability in SONET/SDH networks



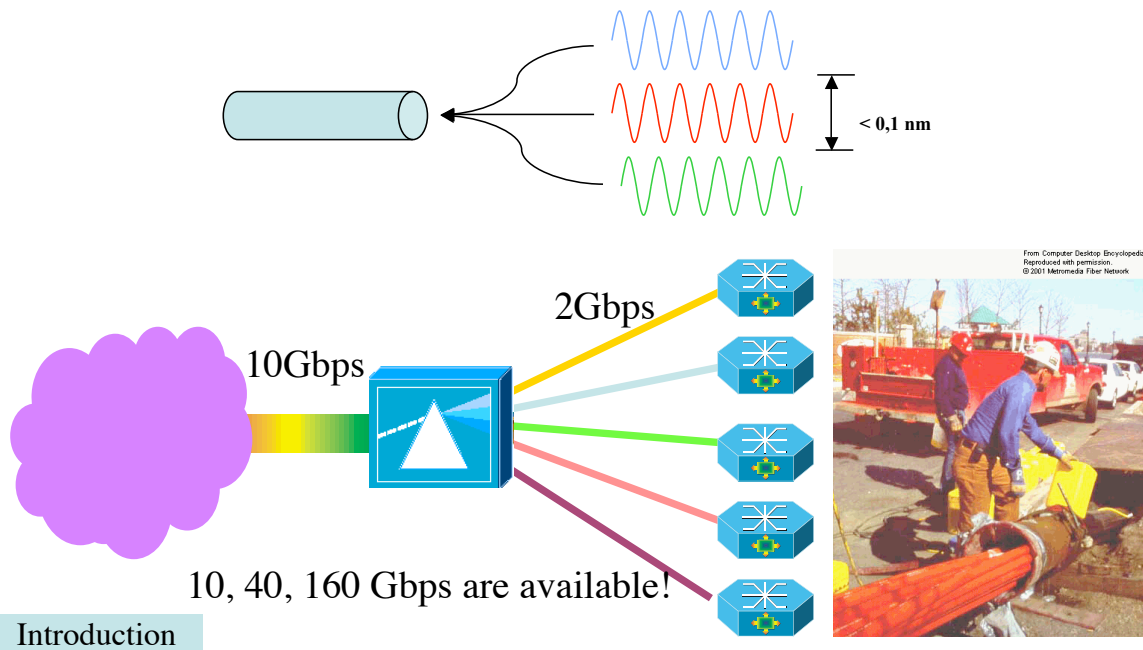
bi-directional



Found in most operators' networks

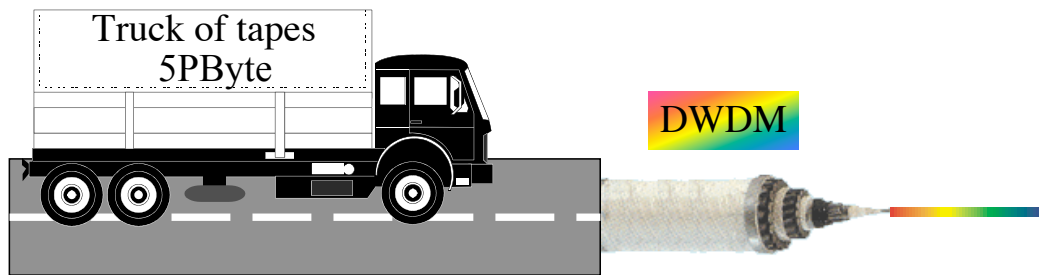
DWDM, bandwidth for free?

DWDM: Dense Wavelength Division Multiplexing



Introduction

The information highways



NEWS of Dec 15th, 2004

3 A throughput of 1.28 Tbits/s has been achieved on a 430kms regular monomode fiber between France Telecom and Deutsch Telecom using 8 DWDM channels (EU project TOPRATE)

Revisiting the truck of tapes (18 of 18)

Consider one fiber

- Current technology allows for 320 λ in one of the frequency bands
- Each λ has a bandwidth of 40 Gbit/s
- Transport: $320 * 40 * 10^9 / 8 = 1600 \text{ GByte/sec}$
- Take a 10 metric ton truck
- One tape contains 50 Gbyte, weights 100 gr
- Truck contains $(10000 / 0.1) * 50 \text{ Gbyte} = 5 \text{ PByte}$
- **Truck / fiber = 5 PByte / 1600 GByte/sec = 3125 s \approx one hour**
- For distances further away than a truck drives in one hour (50 km) minus loading and handling 100000 tapes **the fiber wins!!!**

Introduction

Fibers everywhere?

NEWS of Dec 15th, 2004

Verizon and SBC are deploying large optical fiber infrastructures in the US using FTTC or FTTP scenario

NEWS of May 31st, 2005

US Fiber-to-the-home (FTTH) installations have grown 83% since October 2004, now reaching 398 communities in 43 states

Verizon is on track to pass three million homes with fiber by the end of 2005

campus 2.5Gbps Network Provider 10Gbps 1Gbps GigaEth residential FTTH FTTC

Introduction 23

High Performance Routers

©Cisco

©Nortel Networks

©Alcatel

©Lucent

PRO/8812

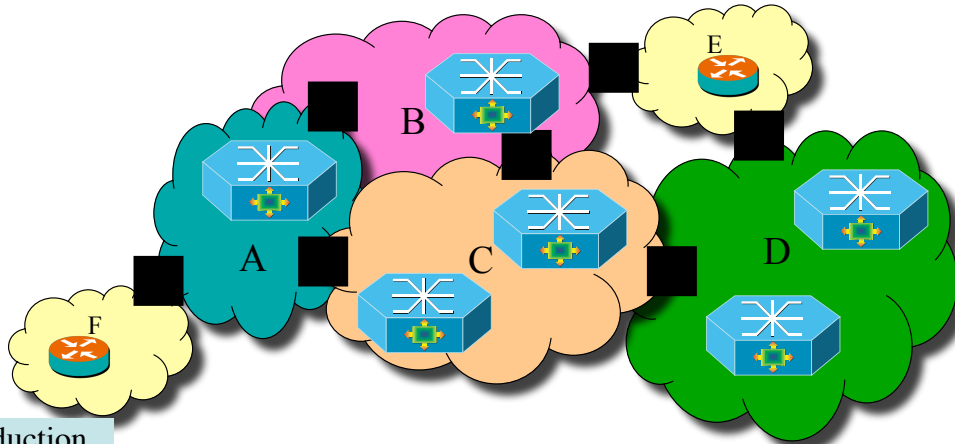
©Procket Networks

and more...

Introduction 24

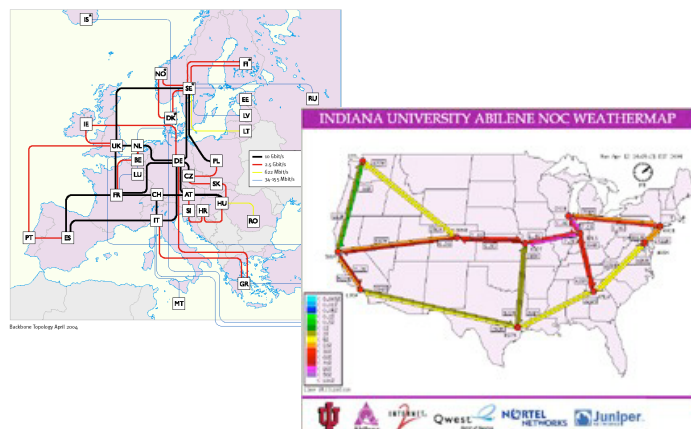
Operator's infrastructure

- ❑ Backbones are optical: OC48 (2.5Gbps), OC192 (10Gbps), OC768 (40Gbps), OC3072(160Gbps)
- ❑ New technologies deployed by operators, POPs available worldwide

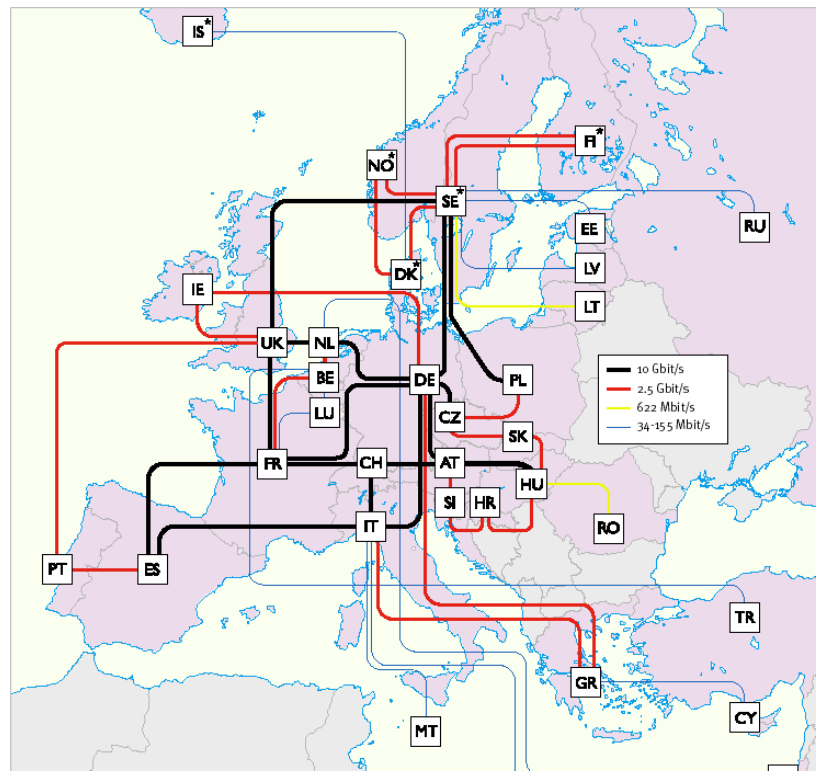


The new networks

- ❑ vBNS
- ❑ Abilene
- ❑ SUPERNET
- ❑ DREN
- ❑ CA*NET
- ❑ GEANT
- ❑ DATATAG
- ❑ ...much more to come!



GEANT



Backbone Topology April 2004

Introduction

New applications on the information highways

Think about...

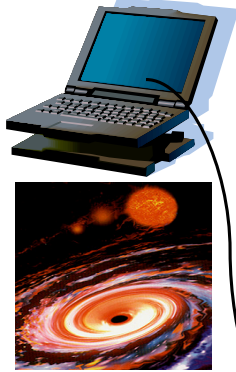
- video-conferencing
- video-on-demand
- interactive TV programs
- remote archival systems
- tele-medicine
- virtual reality, immersion systems
- high-performance computing, grids
- distributed interactive simulations



Introduction

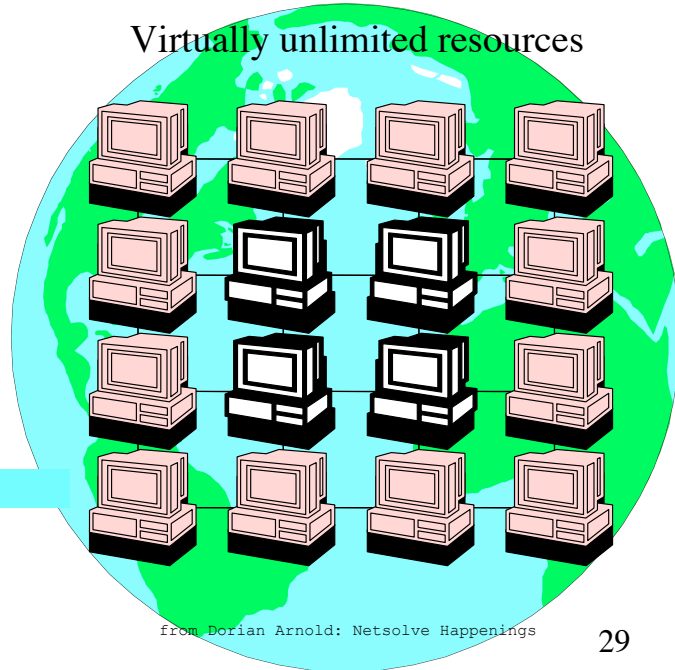
Computational grids

user application



1PFlops

Virtually unlimited resources



from Dorian Arnold: Netsolve Happenings

29

Introduction

One Grid definition

The Grid

“A Grid is a collection of distributed computing resources over **network** that appear to an user or an application as **one large virtual computing system**”

SECURITY

DATA TRANSFER

MANAGEMENT

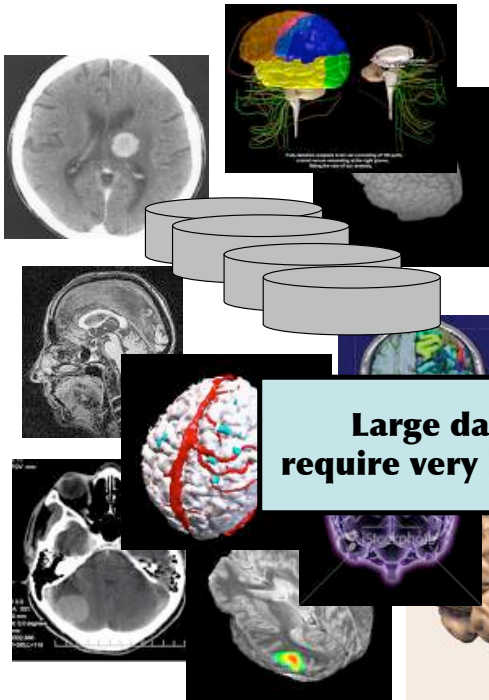
RESOURCES ALLOCATION

MONITORING

Introduction

30

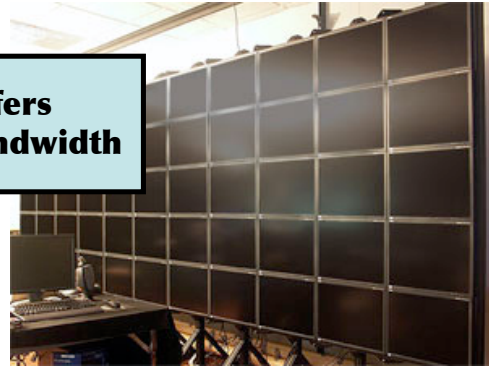
Real-time interactive large-scale scientific collaborations



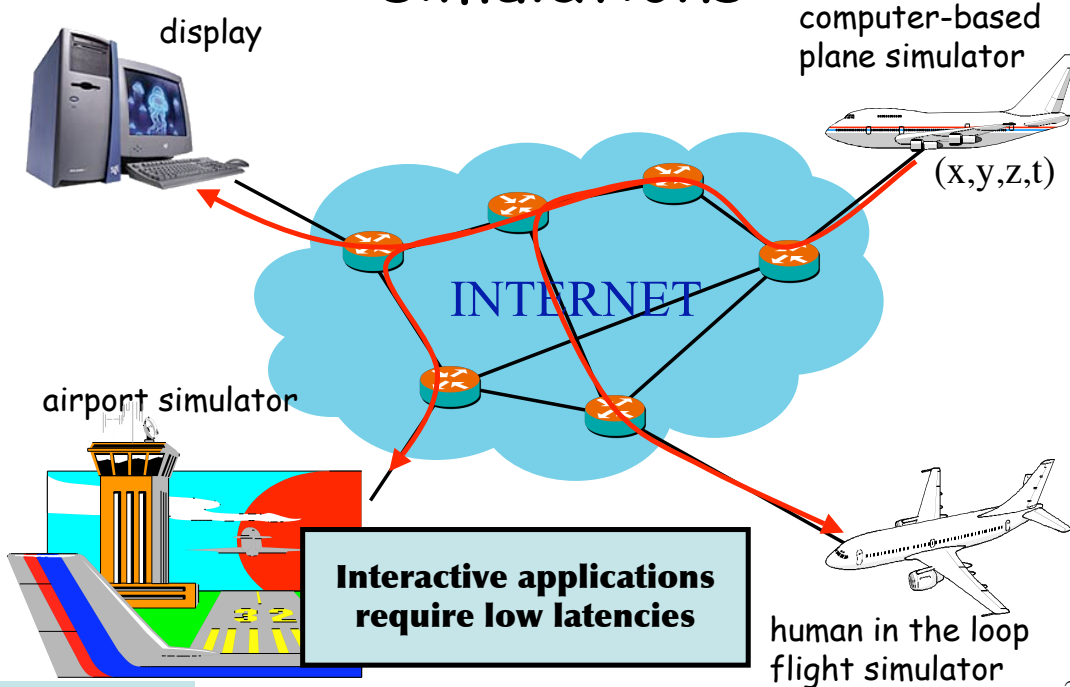
Multimodality brain mapping
require the ability to process, share,
and interactively visualize multiple
100Gbytes datasets!

Today, to visualize and explore eight
3D images require 64Gb/s !

**Large data transfers
require very high bandwidth**



Wide-area interactive simulations



New technologies for Nobel Prize

Resource Broker:
LANL is best match...
but down for the moment

OK! Resource Estimator
says need 5TB, 2TF.
Where can I do this?

From Director@earth.org

Congratulations, you have done a great job, it's the discovery of the century!!

The phenomenon was short but we manage to react quickly. This would have not been possible without efficient multicast facilities to enable quick reaction and fast distribution of data.

Nobel Prize is on the way :-)

This scenario requires
High bandwidth
Low latencies
Multicast facilities

Introduction

In search for the perfect grid

For me, as a user, a computational grid should be:

Easy to use

Fast & Performant

Reliable

Transparent



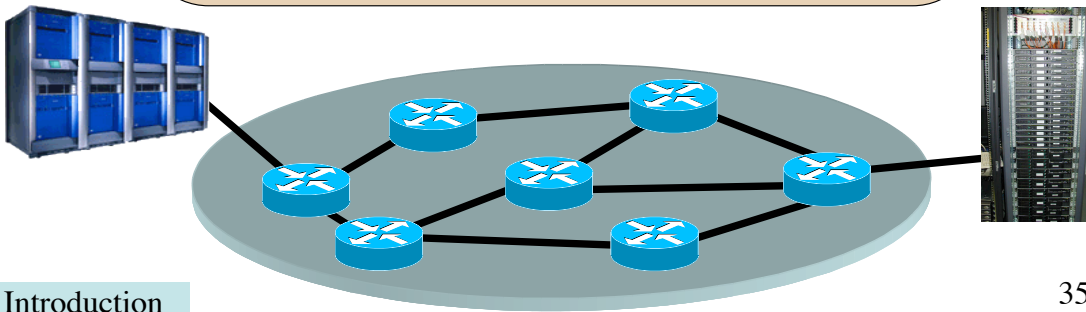
Introduction

Networking issues in the Grid

Manages network resources (link, routers, bandwidth) to offer reliability and guaranteed/predictable performances

Optimizes communication protocols to offer full/optimal utilization of network resources

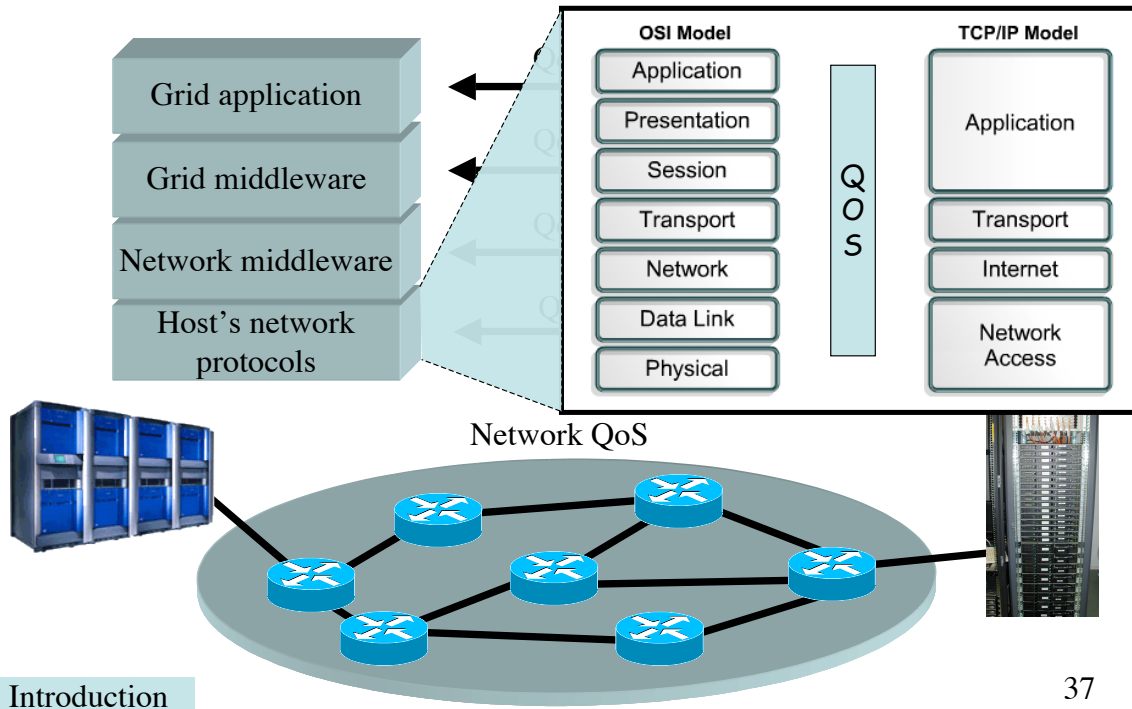
Deploys new technologies to offer new value-added/efficient communication features



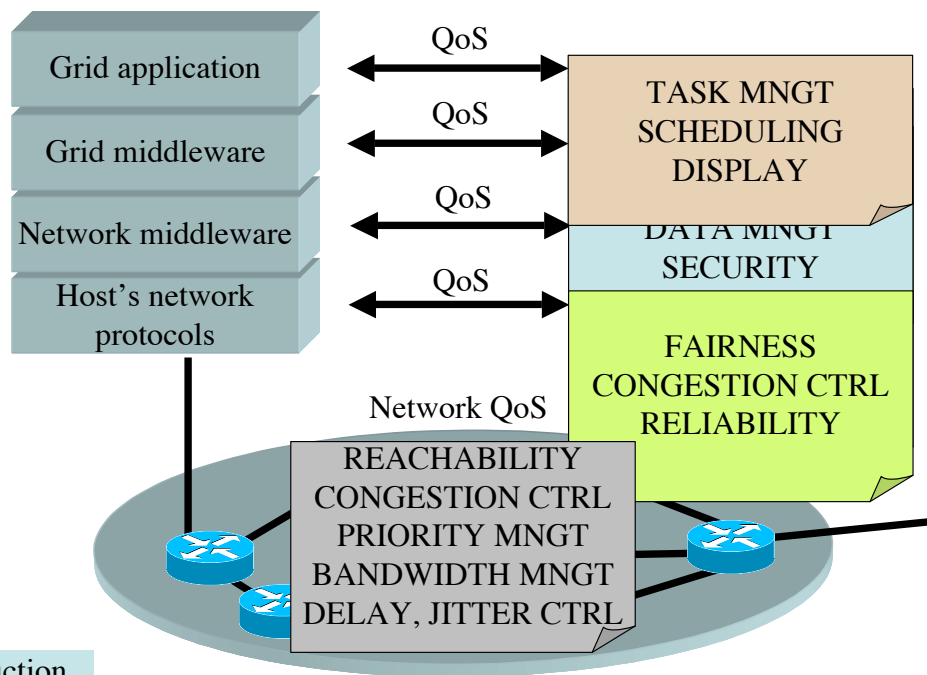
Grid's Quality of Service

- ❑ QoS: « the collective effect of service performance which determines the degree of satisfaction of a user of the service » (ITU)
- ❑ Very broad scope!
- ❑ Unfortunately, difficult to achieve
 - ❑ Spans several communication layers
 - ❑ End-to-end QoS needs coherence between the multiple intervening parties (operators, providers, users)
 - ❑ Lack of a solid, well-established business model

QoS: a difficult task!



QoS: different meanings at different layers



Limitations of the current Internet

❑ Bandwidth

- Raw bandwidth is not a problem: DWDM
- Provisioning bandwidth on demand is more problematic

❑ Latency

- Mean latencies on Internet is about 80-160ms
- Bounding latencies or ensuring lower latencies is a problem

❑ End-to-end performances

- Links are getting faster and faster!
- Why my FTP is still going so slow?

❑ Communication models

- Only unicast communications are well-defined
- Multi-parties communication models are still lacking

Application people come from Venus, Networking people come from Mars

Application guys

The network is a cloud.

Only see TCP, IP and sometimes routing protocols

Will use what is available and working!



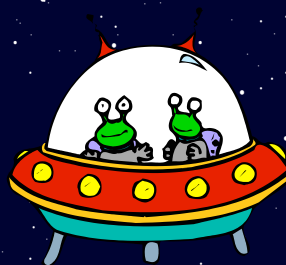
Networking guys

Don't care about applications!

If any applications then must be mainly FTP and web traffic!



Middleware guys



GGF GHPN

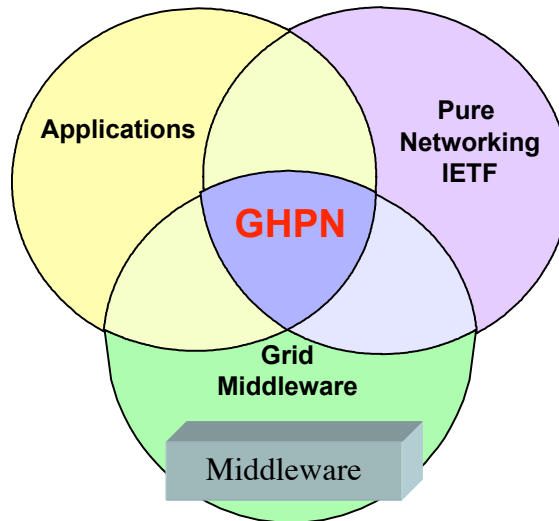
will make a bridge between 3 communities?

Application guys

The network is a cloud.

Only see TCP, IP and sometimes routing protocols

Will use what is available and working!



Networking guys

Don't care about applications!

If any applications then must be mainly FTP and web traffic!

Only problems!

Introduction

41

Purpose of this tutorial

(in more details)

- ❑ **THE ACRONYM JUNGLE!**
 - ❑ GRAM, GARA, GridFTP
 - ❑ SONET/SDH, OXC
 - ❑ MPLS, LSR, GMPLS, VPLS, ASON
 - ❑ DiffServ, DSCP
 - ❑ TCP, HSTCP, STCP, XCP...
- ❑ Explain the basic concepts behind these technologies for Grid/HPC users

42

New technologies addressed in this talk

- ❑ More Quality of Service: **Differentiated Services**, who pays more gets more!
- ❑ **Bandwidth provisioning**: MPLS for virtual circuit in the core networks
- ❑ **Beyond TCP**: fast transport protocols for very high-speed networks
- ❑ **Multicast**: enhancing the communication model