Networks for HEP & the Digital Divide

Harvey B. Newman

California Institute of Technology

CHEP06, TIFR Mumbai

February 15, 2006

SCIC in 2005-2006
http://cern.ch/icfa-scic

Three 2006 Reports:
Rapid Progress, Deepening Digital Divide

- ***Main Report: "Networking for HENP"*** **[H. Newman, et al.]**
  - ➔ **Includes Updates on the Digital Divide, World Network Status; Brief updates on Monitoring and Advanced Technologies**
  - ➔ **27 Appendices: A World Network Overview**
    Status and Plans for the Next Few Years of Nat'l & Regional Networks, HEP Labs, & Optical Net Initiatives
- ***Monitoring Working Group Report*** **[L. Cottrell]**

Also See:

- **TERENA (www.terena.nl) *2005 Compendium:* In-depth Annual Survey on R&E Networks in Europe**
- **http://internetworldstats.com: Worldwide Internet Use**
- ***SCIC 2003 Digital Divide Report*** **[A. Santoro et al.]**

ICFA Report 2006 Update: Main Trends and Issues for 2006

- **HEP's role as the Leading User and Co-Developer of network technologies has come into focus in 2004-2006 (e.g. SC2005):**
  - **➔Rapid progress in tools, applications and "system-level" tuning for high speed data transport; global monitoring**
- ***More than Just Bandwidth is Needed***
  - ***➔Grids and the networks at their foundation need to be Reliable and Efficient for large data flows***
  - ***➔.. and Effective in serving a large community, by providing*** **managed fair-sharing of networks (as well as CPU and storage)**
- **A great deal of work remains to integrate the latest network developments with production grids (OSG, EGEE)**
  - **➔To support a large community doing analysis**

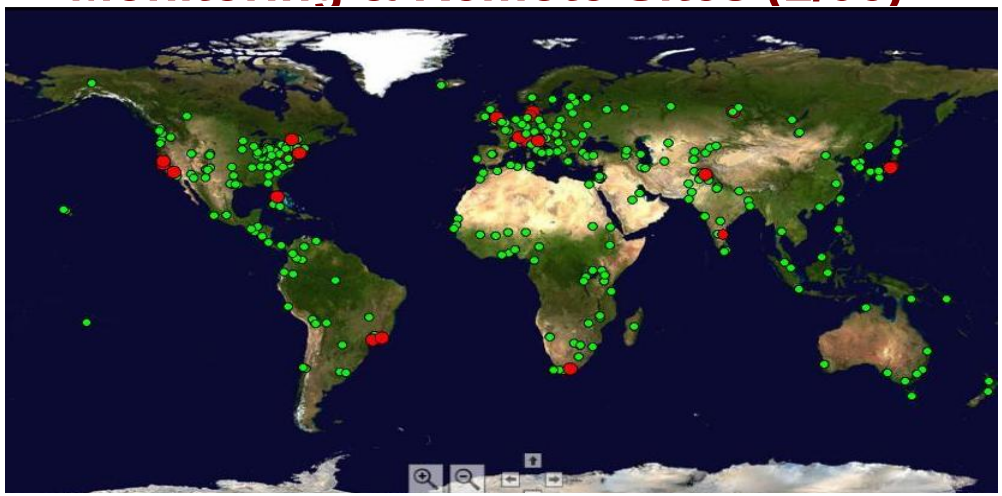
SCIC Main Conclusions for 2006

- ***As we progress we are in danger of leaving the communities in the less-favored regions of the world behind***
  - ➔ ***Scientific Collaboration demands equality***
- ***We must Work to Close the Digital Divide***
  - ➔ ***To make physicists from all world regions full partners in the scientific discoveries***
  - ➔ ***This is essential for the health of our global collaborations, and our field***
- ***We are learning to help do this effectively, in some cases***
  - ➔ ***Brazil and Central Europe***
- ***A great deal of work remains: India, Russia, China, Central and Eastern Europe are focal points for 2006***

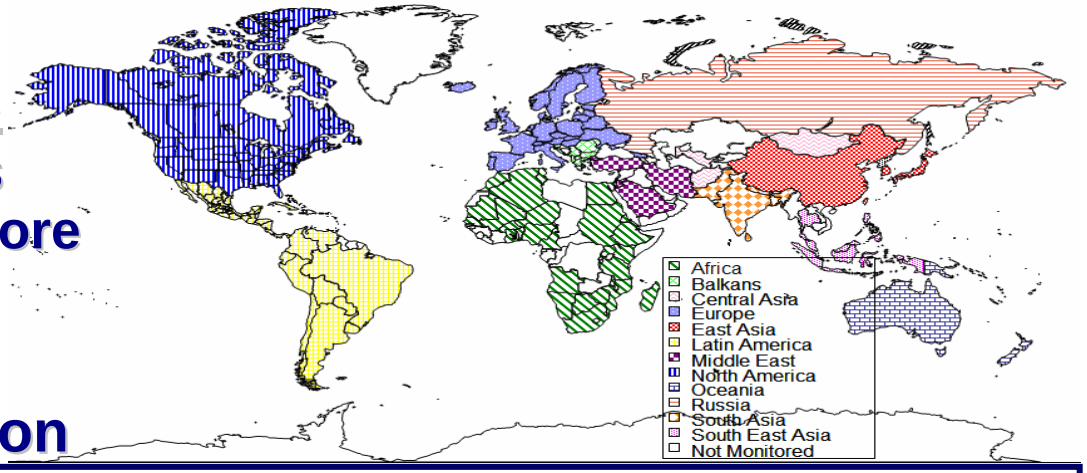
SCIC Monitoring WG
PingER (Also IEPM-BW)

R. Cottrell

Monitoring & Remote Sites (1/06)



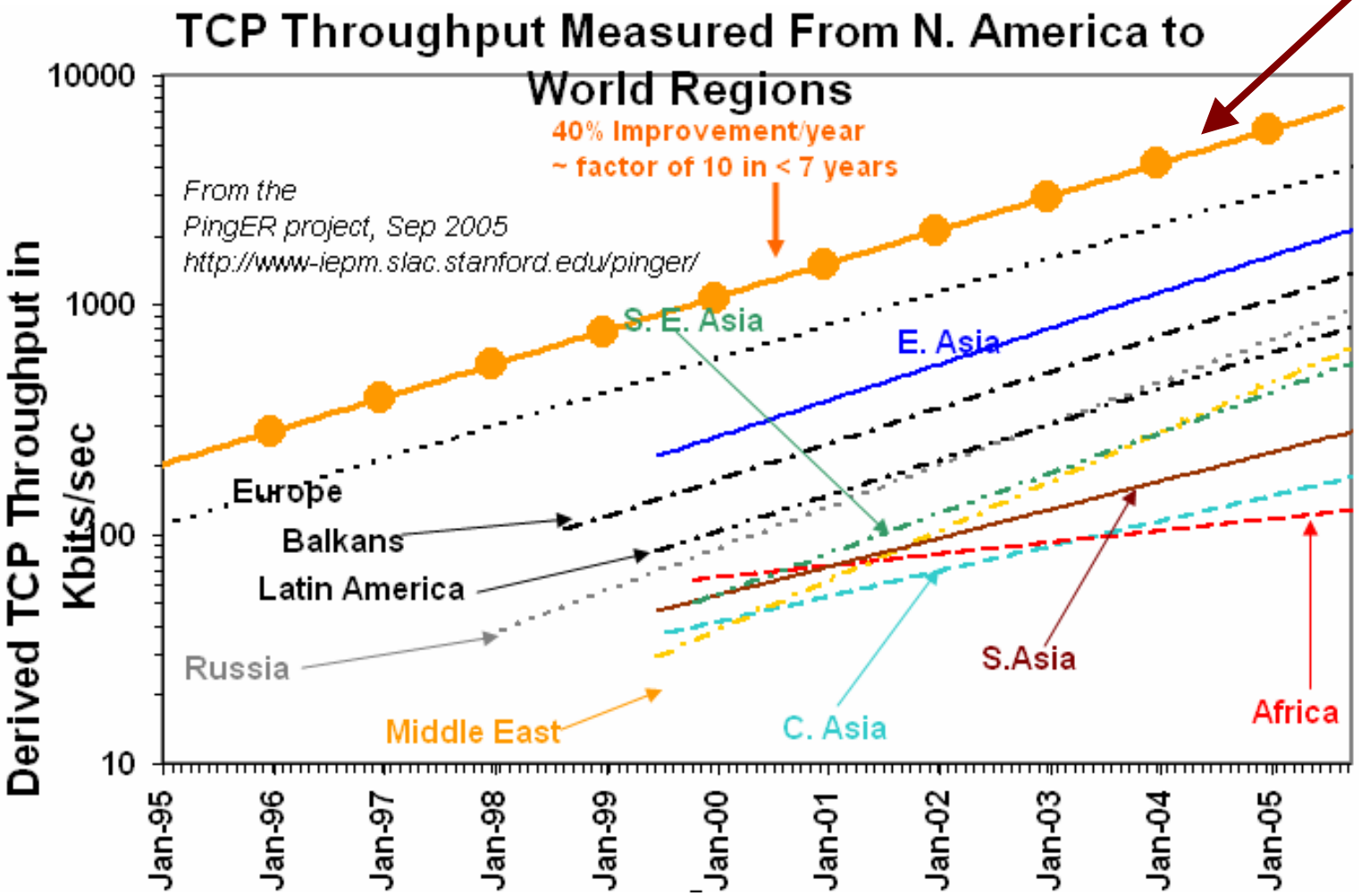
- **Measurements from 1995 On**
  - *Reports link reliability & quality*
- **Countries monitored**
  - → Contain 90% of world population
  - → 99% of Internet users
  - → Seek to Complete Coverage
- **3700 monitor-remote site pairs**
  - → 35 monitors in 14 countries Capetown,Rawalpindi, Bangalore
  - → 1000+ remote sites in 120 Countries
  - → Aggregation by World Region



Countries: N. America (2), Latin America (18), Europe (25), Balkans (9), Africa (31), Mid East (5), Central Asia (4), South Asia (5), East Asia (4), SE Asia (6), Russia includes Belarus & Ukraine (3), China (1) and Oceania (5)

SCIC Monitoring WG - Throughput Improvements 1995-2006

Progress: but Digital Divide is Mostly Maintained



40% annual improvement Factor ~10/7 yrs

Behind Europe

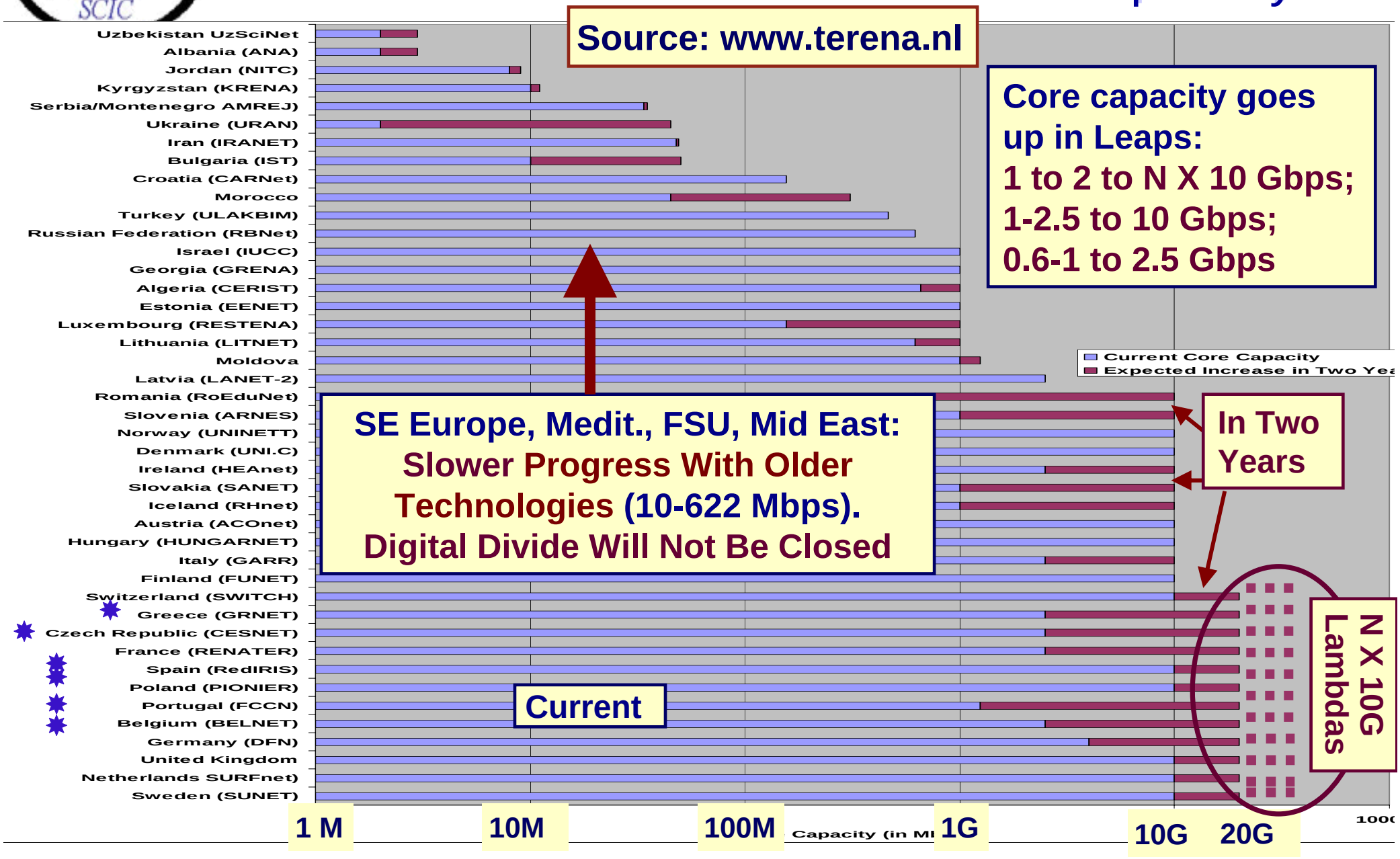
- **6 Yrs: Russia, Latin America**
- **7 Yrs: Mid-East, SE Asia**
- **10 Yrs: South Asia**
- **11 Yrs: Cent. Asia**
- **12 Yrs: Africa**

INDIA, Central Asia, and Africa are in Danger of Falling Even Farther Behind

*Bandwidth of TCP < MSS/(RTT*Sqrt(Loss))*

Matthis et al., Computer Communication Review 27(3), July 1997

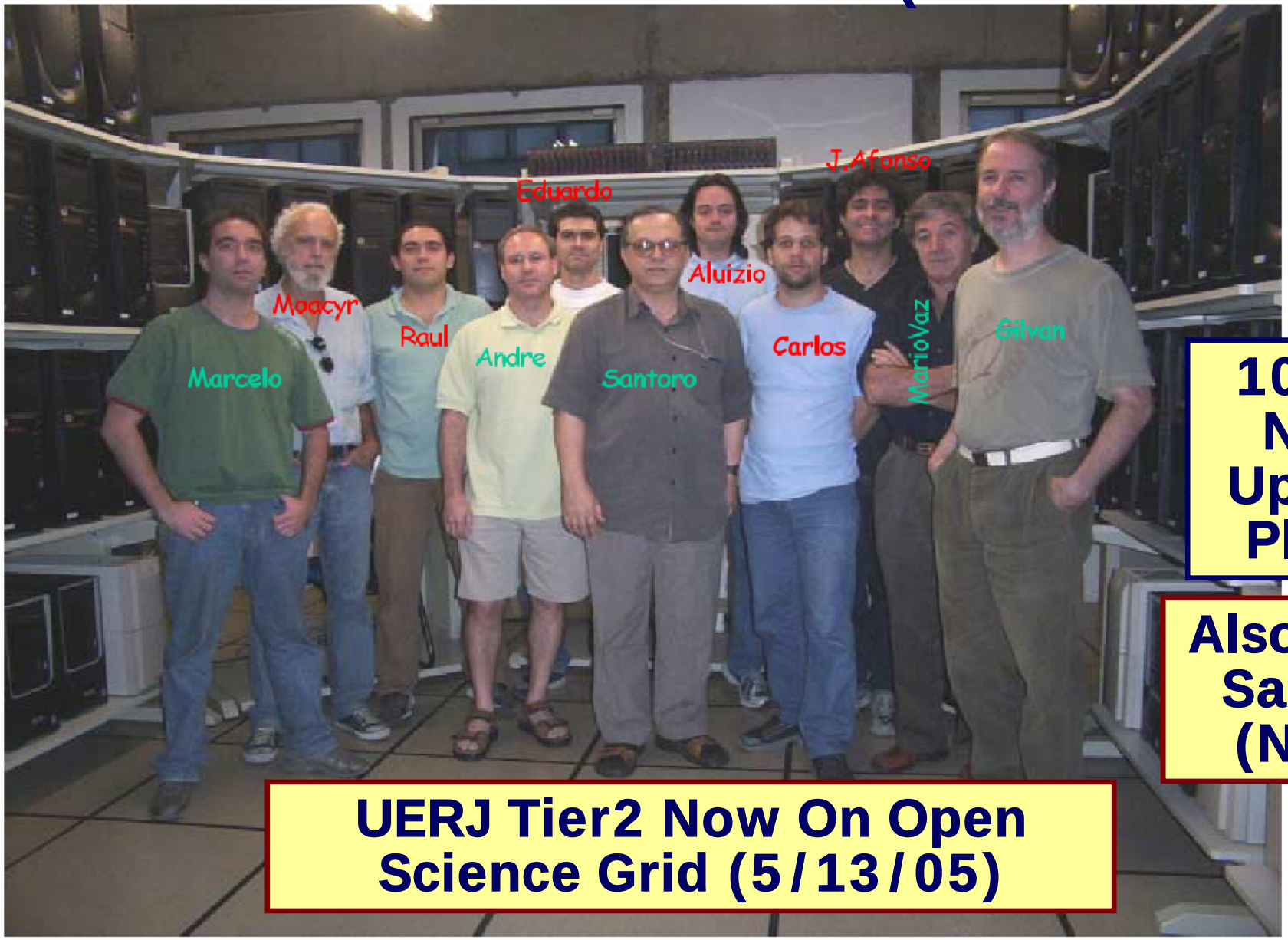
Digital Divide Illustrated by Network Infrastructures: TERENA Core Capacity



Work on the Digital Divide from Several Perspectives

- **Work Within the Community: Locally and Globally**
- **Technical Help with Modernizing the Infrastructure:**
  - **Provide Tools for Effective Use: Data Transport, Monitoring, Collaboration**
  - **Design, Commissioning, Development**
- **Share Information: *Monitoring, BW Progress*; Dark Fiber Projects & Pricing**
  - **Model Cases: Poland, Slovakia, Brazil, Czech Rep., China …**
  - **Encourage Access to Dark Fiber**
- **Encourage, and Work on Inter-Regional Projects**
  - ***INDIA Links to US and Europe, and the World's NRENs***
  - **GLORIAD, Russia-China-Korea-US-Europe Optical Ring**
  - **Latin America: CHEPREO/WHREN (US-Brazil); RedCLARA**
  - **Mediterranean: EUMEDConnect; Asia-Pacific: TEIN2**

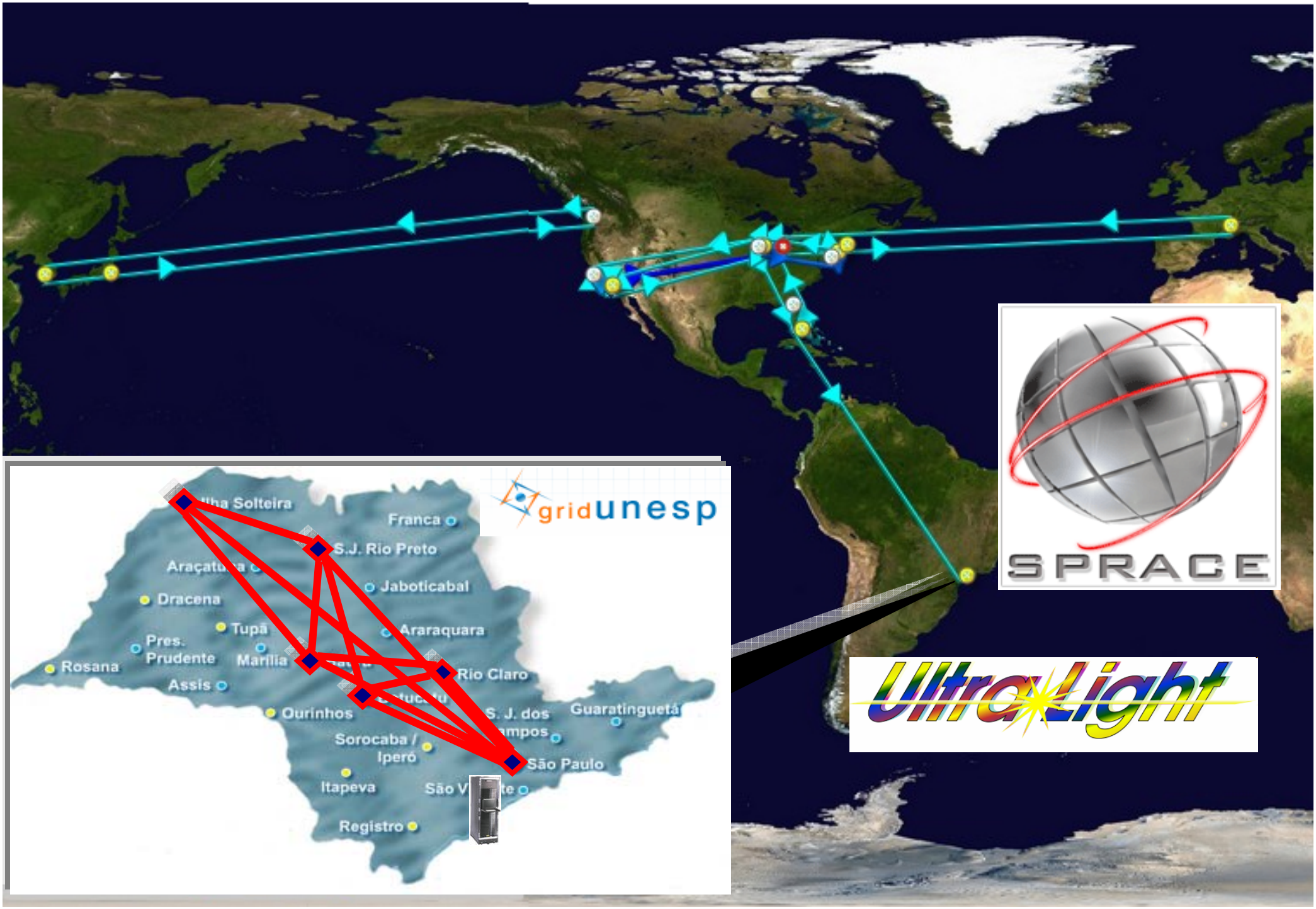
UERJ T2 HEPGRID Inauguration: Dec. 2004: The Team (Santoro et al.)



100 Dual Nodes; Upgrades Planned

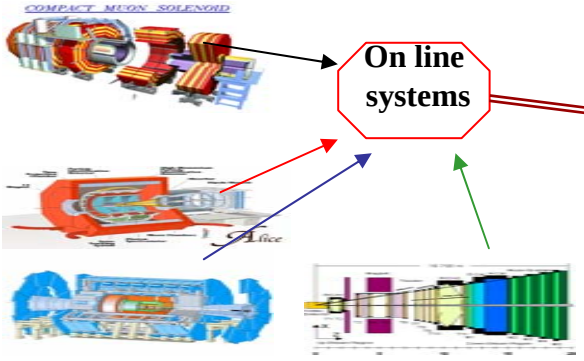
Also Tier3 in Sao Paulo (Novaes)

UERJ Tier2 Now On Open Science Grid (5/13/05)


gridunesp
Franca
S.J. Rio Preto
Jaboticabal
Araraquara
Dracena
Tupã
Pres. Prudente
Marília
Rosana
Assis
Rio Claro
Ourinhos
S. J. dos
Guaratinguetá
Sorocaba /
Iperó
São Paulo
Itapeva
Registro
SPRACE
UltraLight

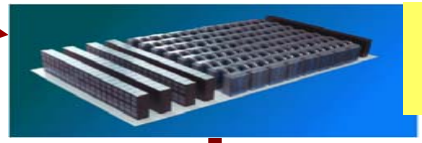
Science-Driven: HEPGRID (CMS) in Brazil

HEPGRID-CMS/BRAZIL is a project to build a Grid that
- **At Regional Level will include CBPF,UFRJ,UFRGS,UFBA, UERJ & UNESP**
- **At Int'l Level will be integrated with CMS Grid based at CERN, FNAL; focal points include OSG and bilateral projects with Caltech Group**



On line systems

Brazilian HEPGRID



T0 +T1

ICFA DD Workshop 2/04; T2 Inauguration + GIGA/RNP Agree 12/04

CERN

2.5 to 10Gbps:No Longer a Dream

T1

France | Germany | Italy | BRAZIL | 622 Mbps | USA

UNESP/USP SPRACE-Working

UERJ Regional Tier2 Ctr

T2 ➔T1

T3 ➔T2

UFRGS

Gigabit

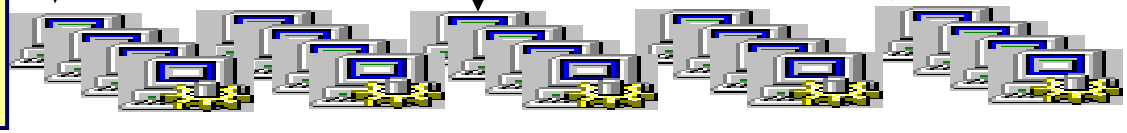
UERJ

CBPF

UFBA

UFRJ

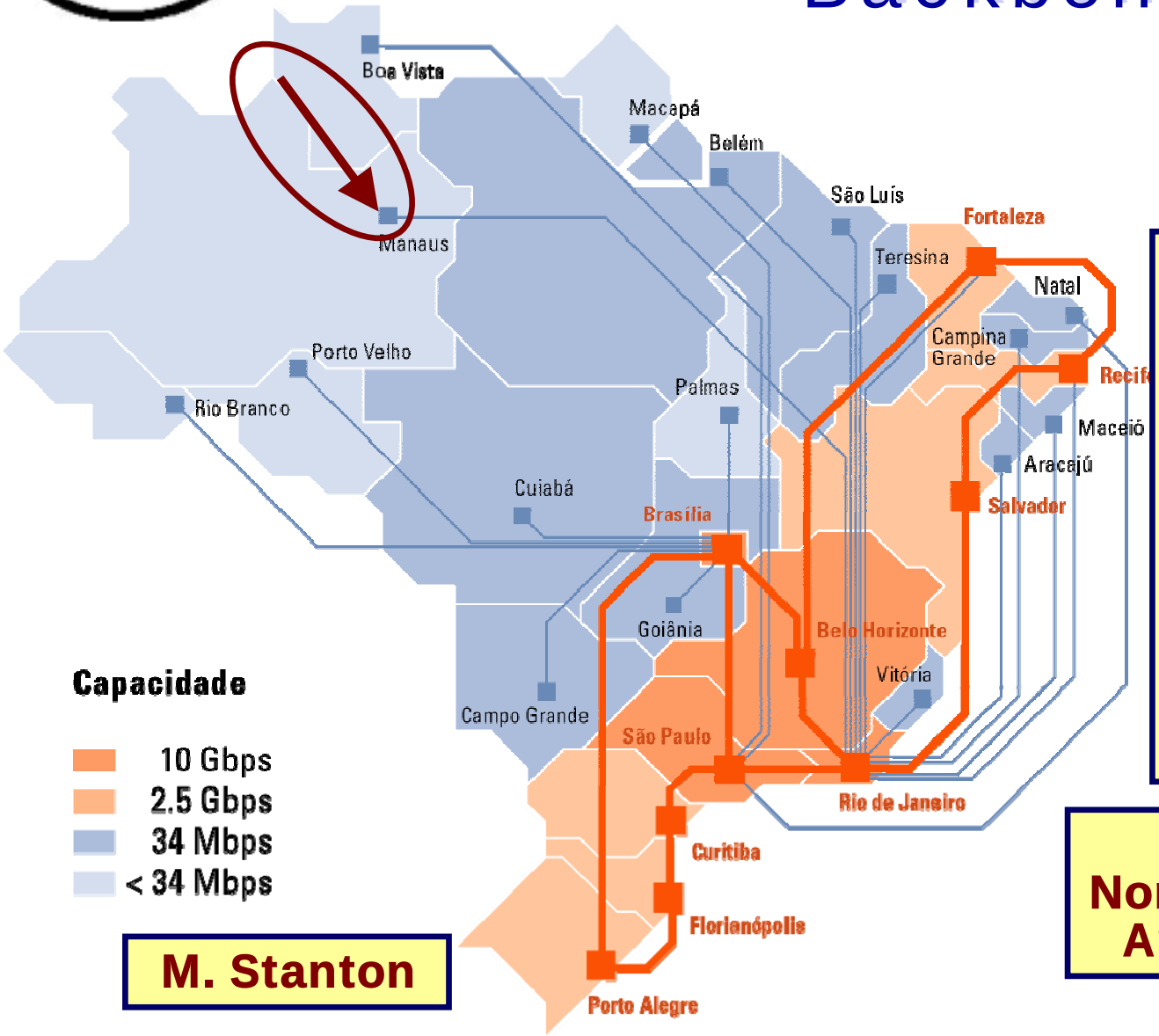
T4



UERJ: T2➔T1, 100➔500 Nodes; Plus T2s to 100 Nodes

Extend to Colombia by 2006

Brazil: RNP2 Next-Generation Backbone



M. Stanton

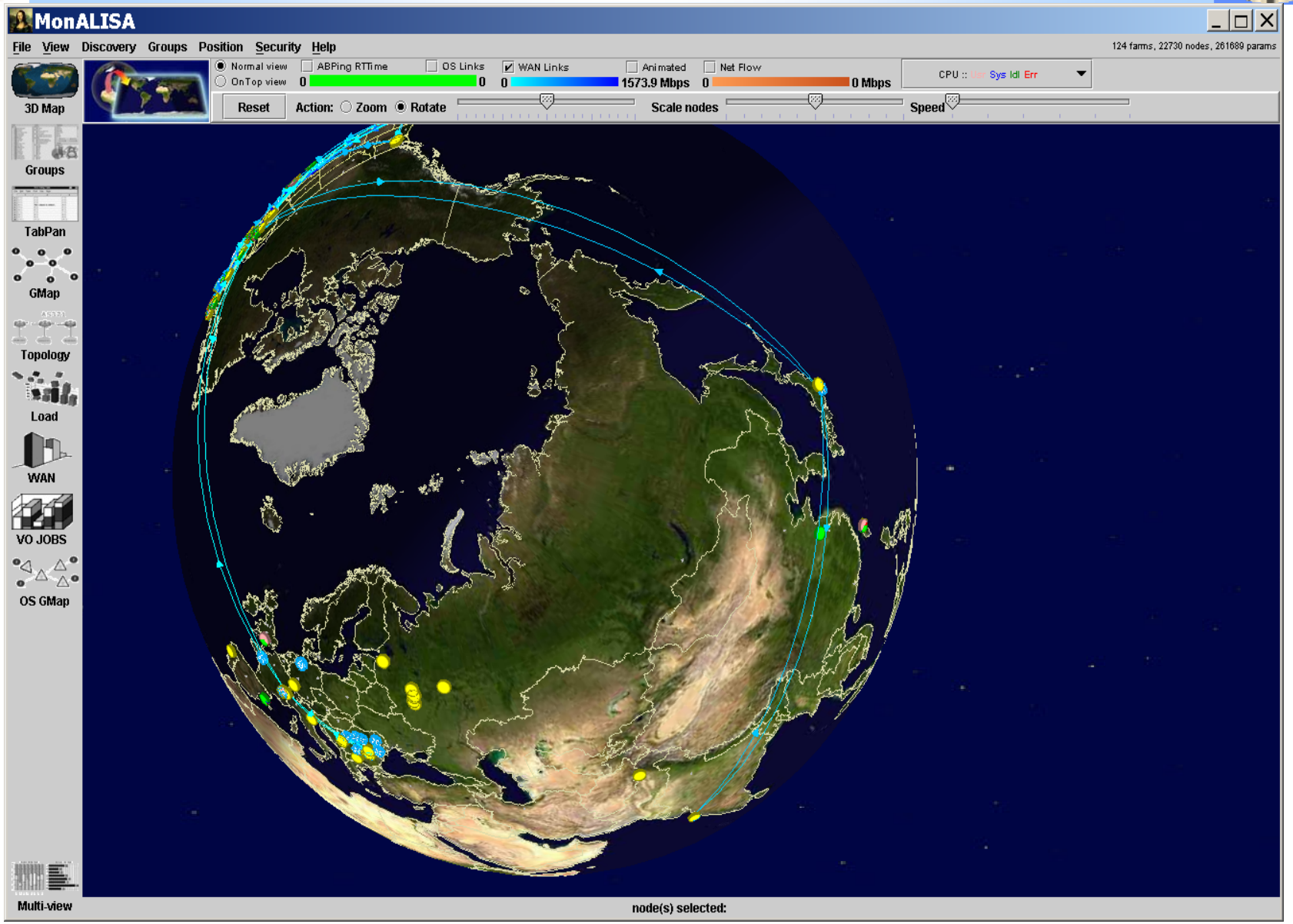
New vs. Old
A factor of 70 to 300 in Bandwidth

2006:

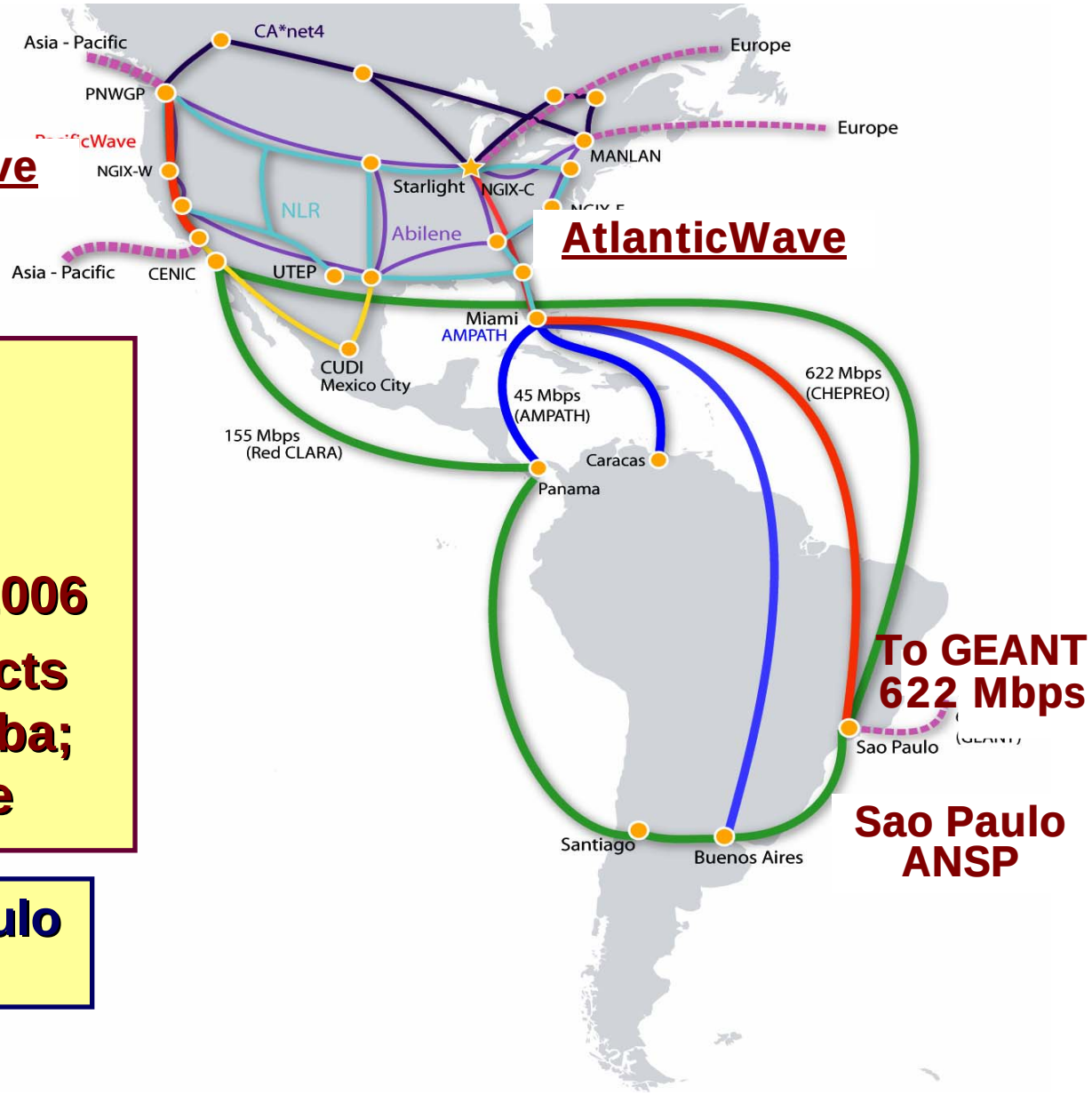
- **Buildout of dark fiber nets in 27 cities with RNP PoPs underway**
- **Expect 200 Institutions Connected at 1 GbE in 2006**
- **2.5G WHREN (NSF) Link to US; 622M Link to GEANT**

Plan: Extend to the Northwest; Dark fiber across Amazon jungle to Manaus

Mumbai-Japan-US Links



Closing the Digital Divide: R&E Networks in / to Latin America



- RNP2 (Rio de Janeiro) and ANSP (Sao Paulo)
- WHREN/LILA (US NSF)
  - ➔1.2 to 2.5 G Ring by 2006
- RedCLARA (EU): Connects 18 Latin Am. NRENs, Cuba; 622 Mbps Link to Europe

iGrid05, SC05: 2.5G Sao Paulo to San Diego (RNP + GC)

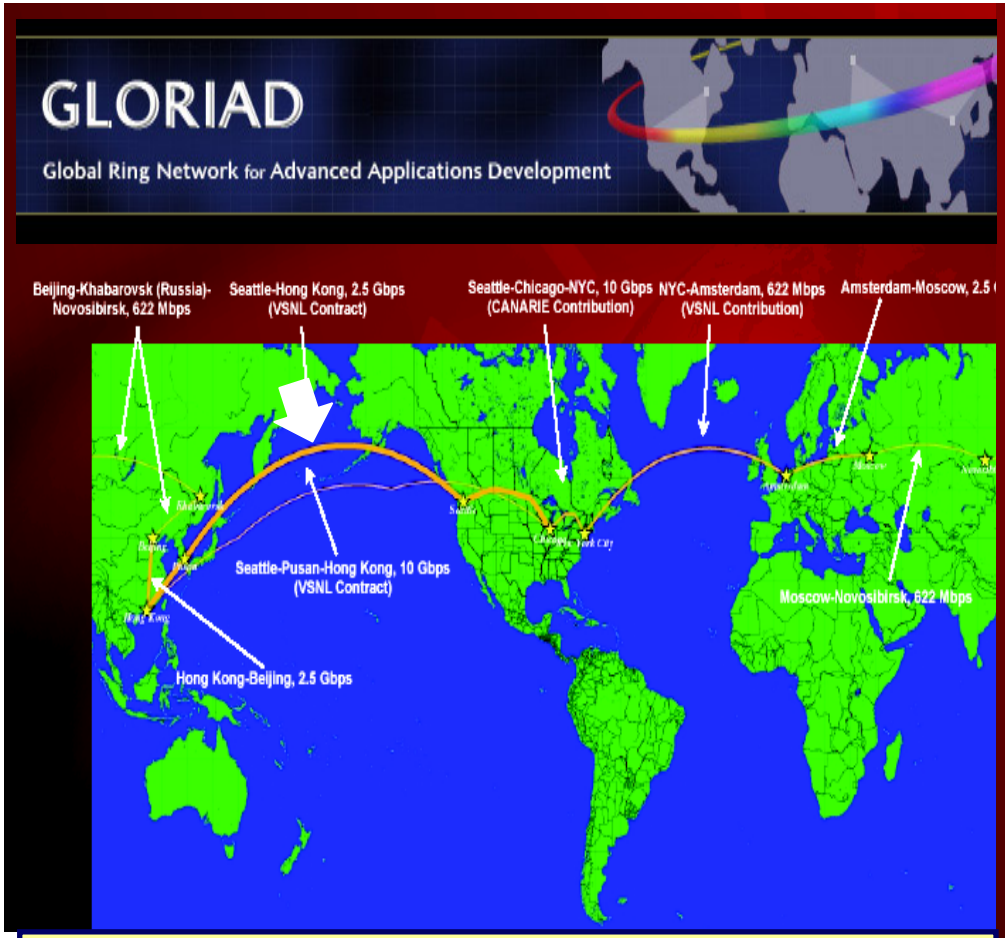
CHINA: CERNET Map January 2006

- Backbone raised to multiples of 10 Gbps
- Regional bandwidth to multiples of 2.5 Gbps
- 2.5 Gbps GLORIAD Link
- Setting up 622 Mbps or 2.5G link to GEANT2 in the CNGI (China Next Generation Internet) Project

From 6 to 78M Internet Users in China from January – July 2004; 111M Users in January 2006

http://www.cnnic.net.cn/en/index/0O/02/index.htm

GLORIAD: 10 Gbps Optical Ring Around the Globe by March 2007



China, Russia, Korea, Japan, US, Netherlands Partnership

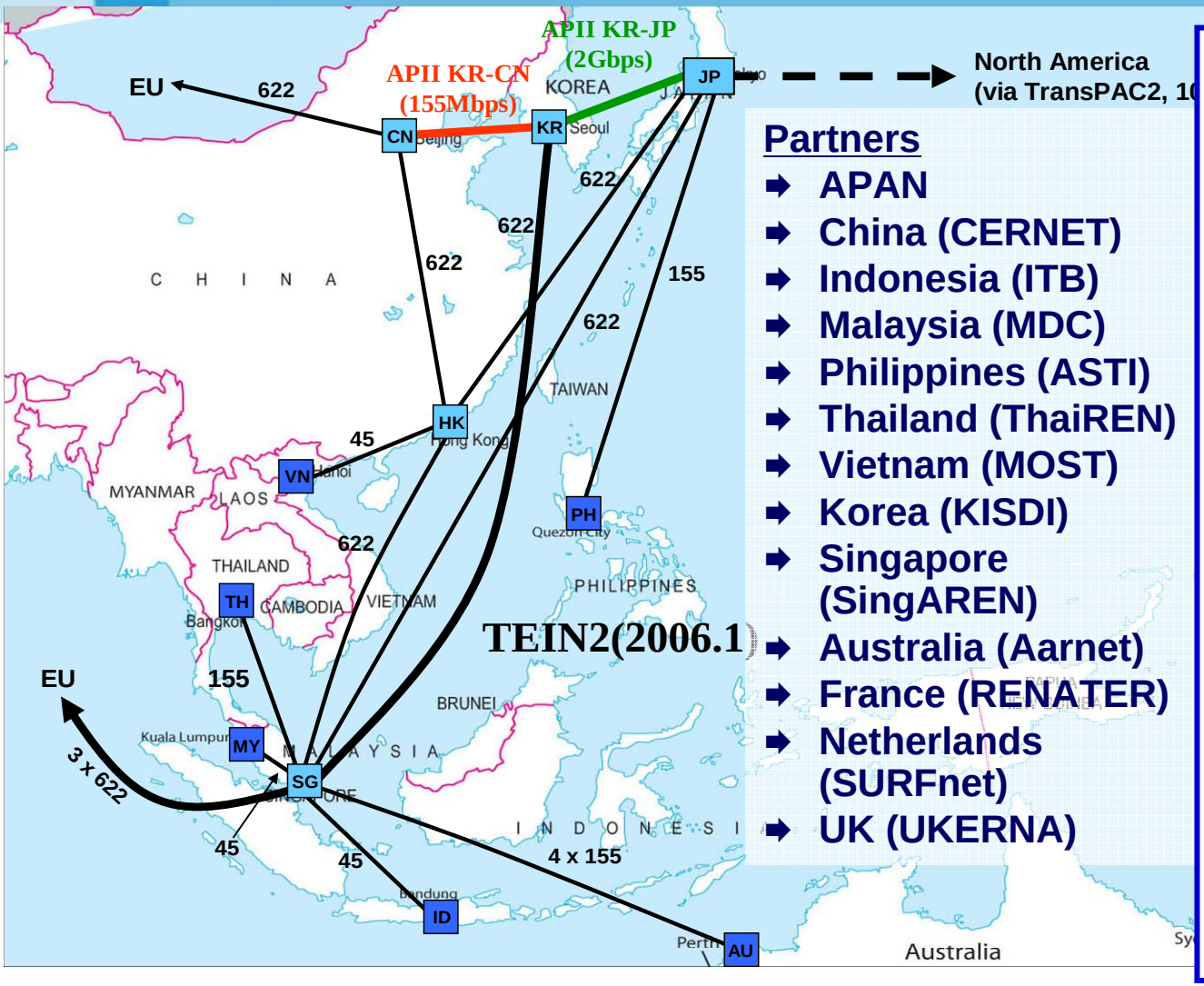
US: NSF IRNC Program

GLORIAD Circuits Today

- 10 Gbps Hong Kong-Daejon-Seattle
- 10 Gbps Seattle-Chicago-NYC (CANARIE contribution to GLORIAD)
- 622 Mbps Moscow-AMS-NYC
- 2.5 Gbps Moscow-AMS
- 155 Mbps Beijing-Khabarovsk-Moscow
- 2.5 Gbps Beijing-Hong Kong
- 1 GbE NYC-Chicago (CANARIE)

G. Cole

TEIN2 (EU and Partner NRENs and Agencies): Improving Connectivity in the Asia-Pacific Region [*]



Partners

- APAN
- China (CERNET)
- Indonesia (ITB)
- Malaysia (MDC)
- Philippines (ASTI)
- Thailand (ThaiREN)
- Vietnam (MOST)
- Korea (KISDI)
- Singapore (SingAREN)
- Australia (Aarnet)
- France (RENATER)
- Netherlands (SURFnet)
- UK (UKERNA)

622 Mbps Core
Tokyo, Hong, Singapore

Spurs

622M: China, Korea
+2G Korea-China

4 X 155M: Australia

155M: Thailand
Philippines
Taiwan

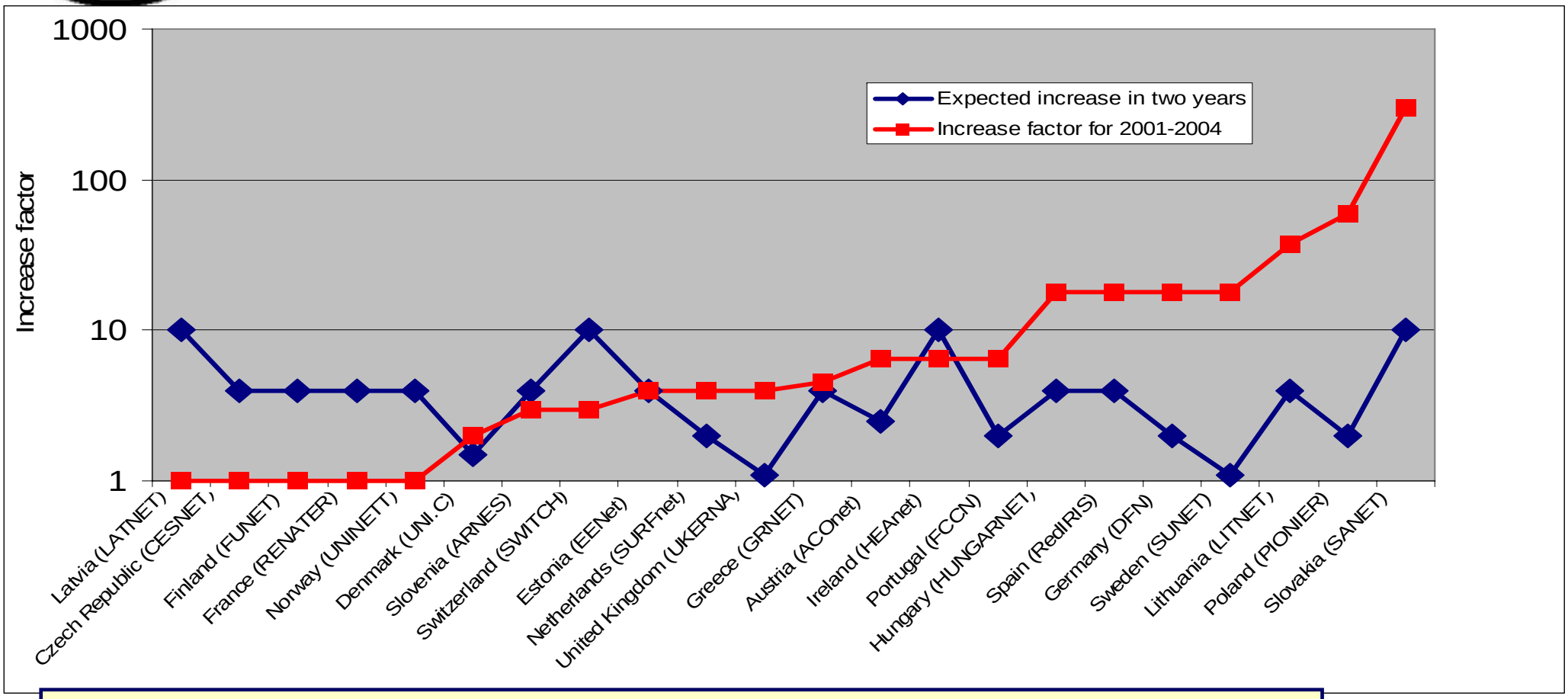
45M: Vietnam
Malaysia
Indonesia

10G to US (TransPAC)

4 X 622 Mbps to GEANT

[*] Before TEIN2 many North-South links were 0.5 -2 Mbps

Core Network Bandwidth Increase for Years 2001-2004 and 2004-2006

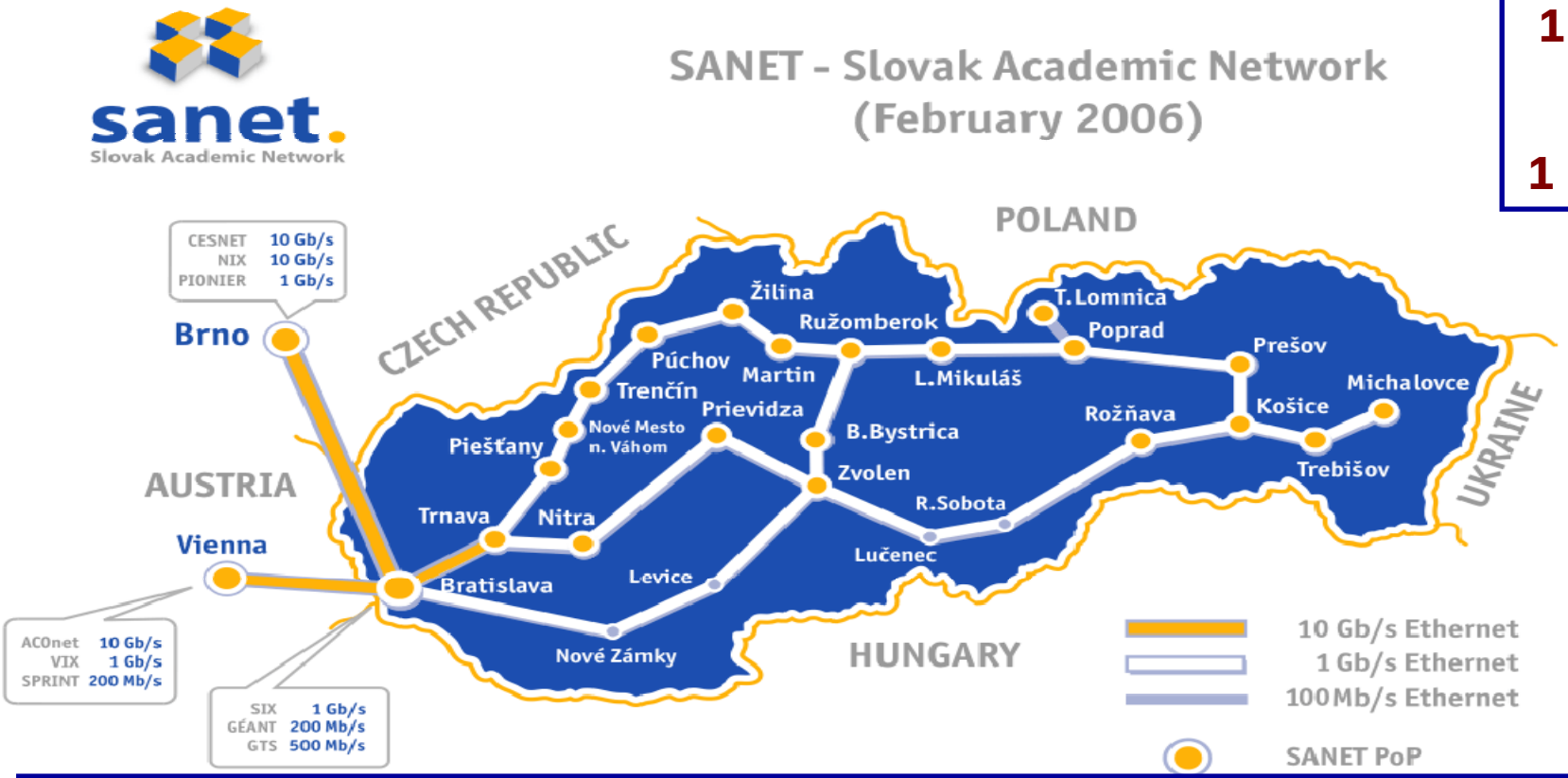


- **Countries With No Increase Already Had 1-2.5G Backbone in 2001**
  - **These are all going to 10G backbones by 2006-7**
- **Countries Showing the Largest Increase Are:**
  - ➔ **PIONIER (Poland) from 155 Mbps to 10 Gbps capacity (64X)**
  - ➔ **SANET (Slovakia) from 4 Mbps to 1 Gbps (250X).**

Source: TERENA

SLOVAK Academic Network
February 2006: All Switched Ethernet

120km CBDF
Cost 4 k€
Per Month
1 GE 2/16/05

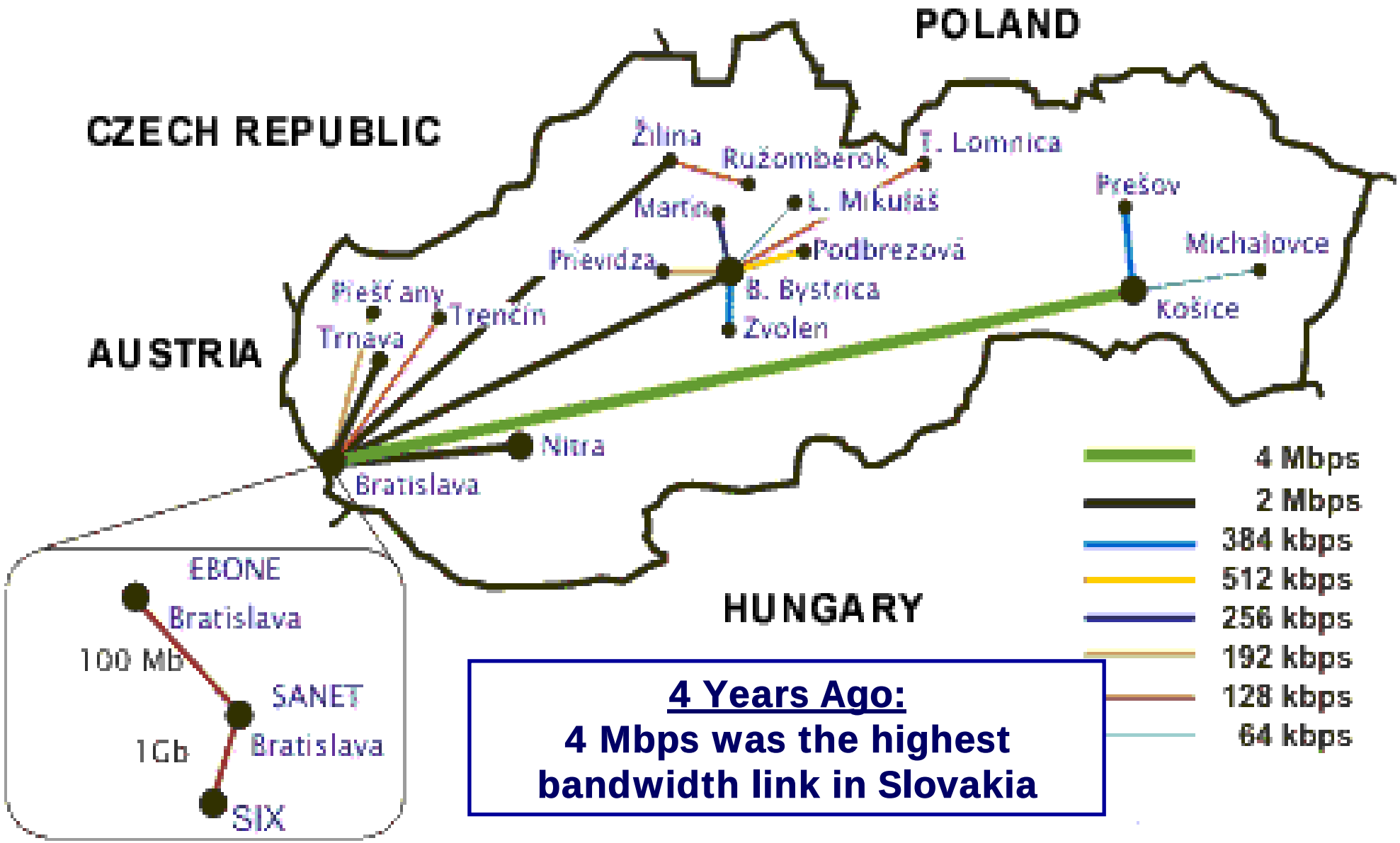


T. Weis

- **1660 km of Dark Fiber CWDM Links, 1 to 4 Gbps (GbE)**
- **August 2002: Dark Fiber Link, to Austria**
- **April 2003: Dark Fiber Link to Czech Republic**
- **2004: Dark Fiber Link to Poland**
- **Planning 10 GbE Backbone; Dark Fiber to Ausria and Czech Republic, and Regional and Metro Nets**

2500x: 2002-2006

SANET - Slovak Akademic Data Network
(January 2002)



4 Years Ago:
4 Mbps was the highest bandwidth link in Slovakia

PIONIER (Poland) Cross Border Dark Fiber Plan Locations

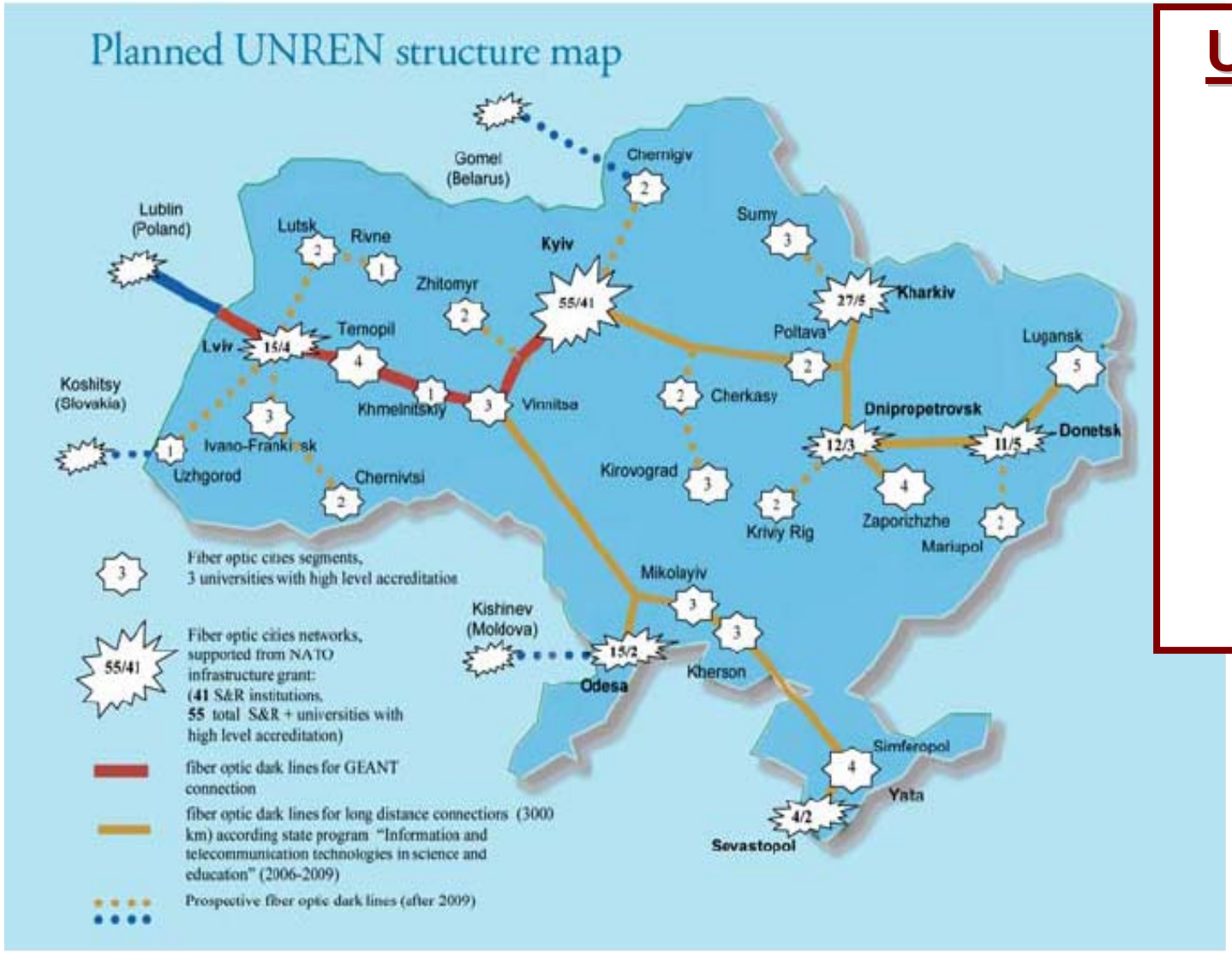
Plan: Single GEANT PoP in Poznan

Key Enabler of Networking in Czech Rep., Slovakia, & the Ukraine

64X Backbone Improvement 2002-5

M. Turala

Ukraine: *UNREN Research and Education Network Plan*



UNREN (2006-9)

3000 km Dark Fiber Infrastructure

Kyiv-Lublin (Poland) Dark Fiber Link to Connect to GEANT2

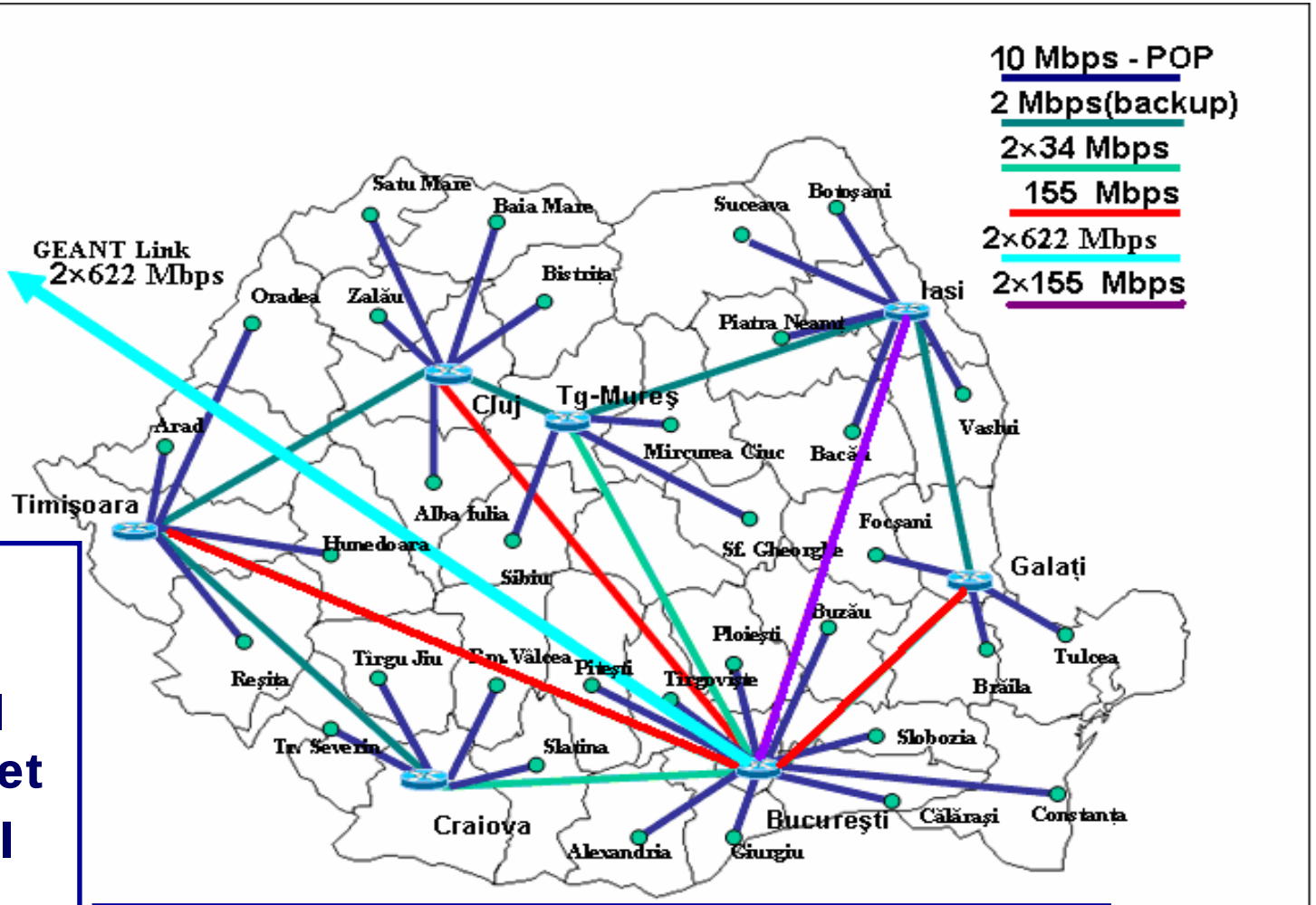
Romania: RoEduNet Topology

Connects 610 Institutions to GEANT:

- 38 Universities
- 32 Research Institutes
- 500 Colleges & High Schools
- 40 Others

RoGrid Plans for 2006

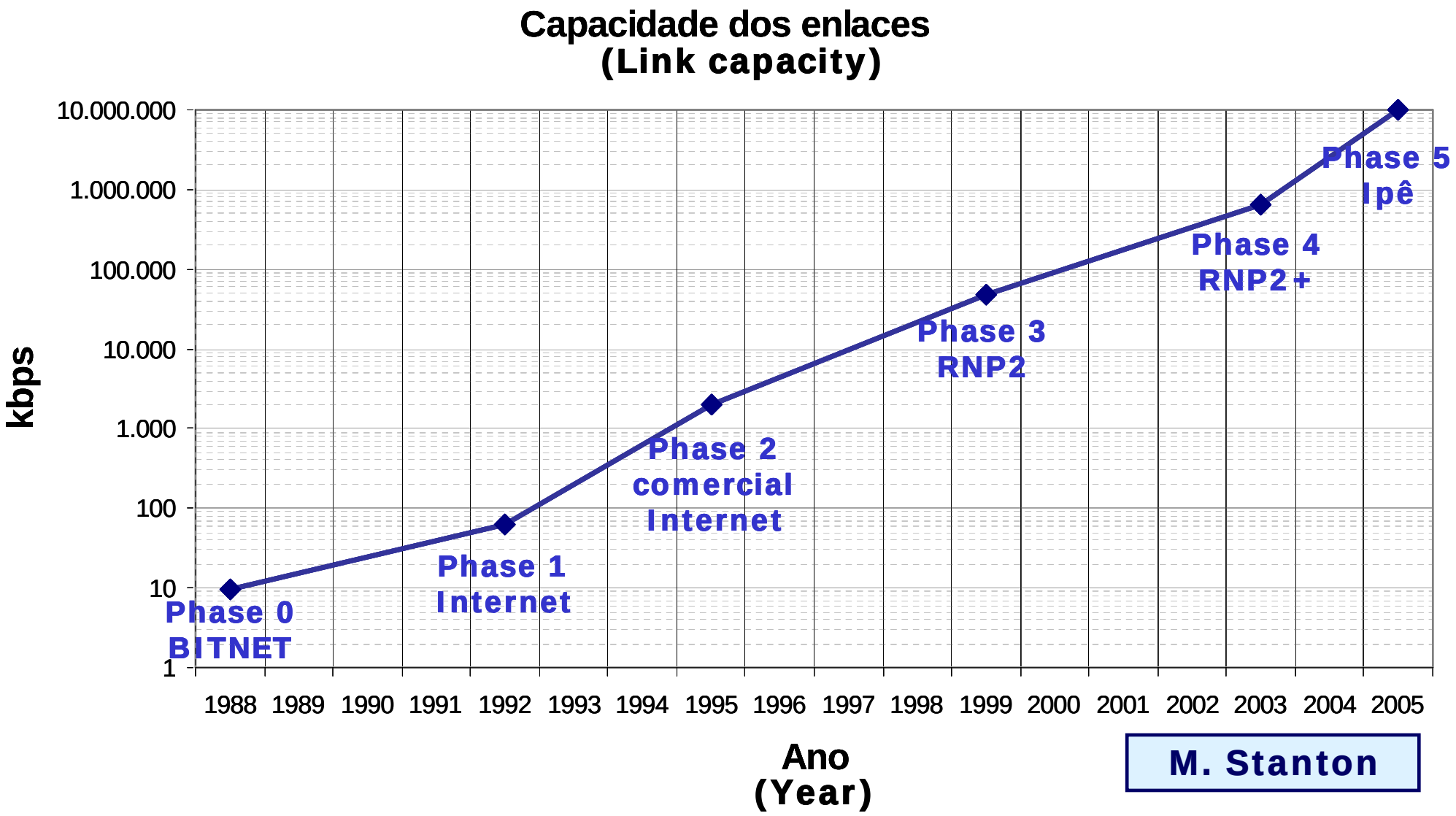
- 10G Experimental Link UPB-RoEdunet
- Upgrade 3-4 Local Centers to 2.5G



Future Plan: Dark Fiber Infrastructure with 10G Light-paths

N. Tapus

The Progress of RNP (Brazil): A factor of 1 Million in 17 Years; A factor of 16 in the Last 2 Years



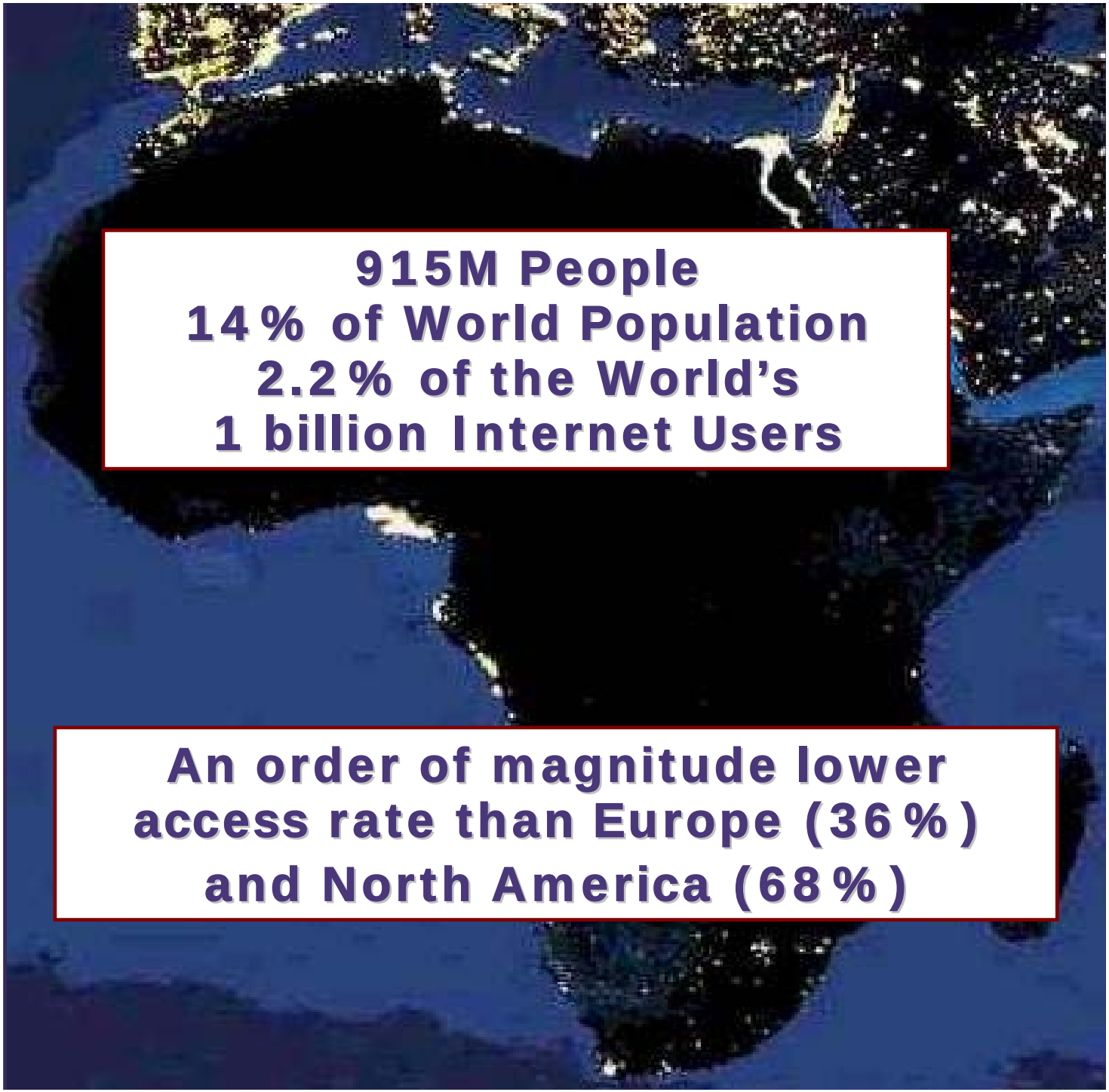
New Focus on AFRICA

Only world region genuinely in decline:

Lack of energy, infrastructure.
Lack of expertise

Problems of
- Disease
- Political unrest
- Pricing
- Import duties & policies
- Corruption

M. Jensen



Internet Users: Africa and the Rest of the World

- **Internet Penetration in Africa is 2.5% (1.4% in 2004): Still more than an order of magnitude less than Europe (36%), and North America (68%)**

WORLD INTERNET USAGE AND POPULATION STATISTICS
Updated December 31, 2005
http://www.internetworldstats.com

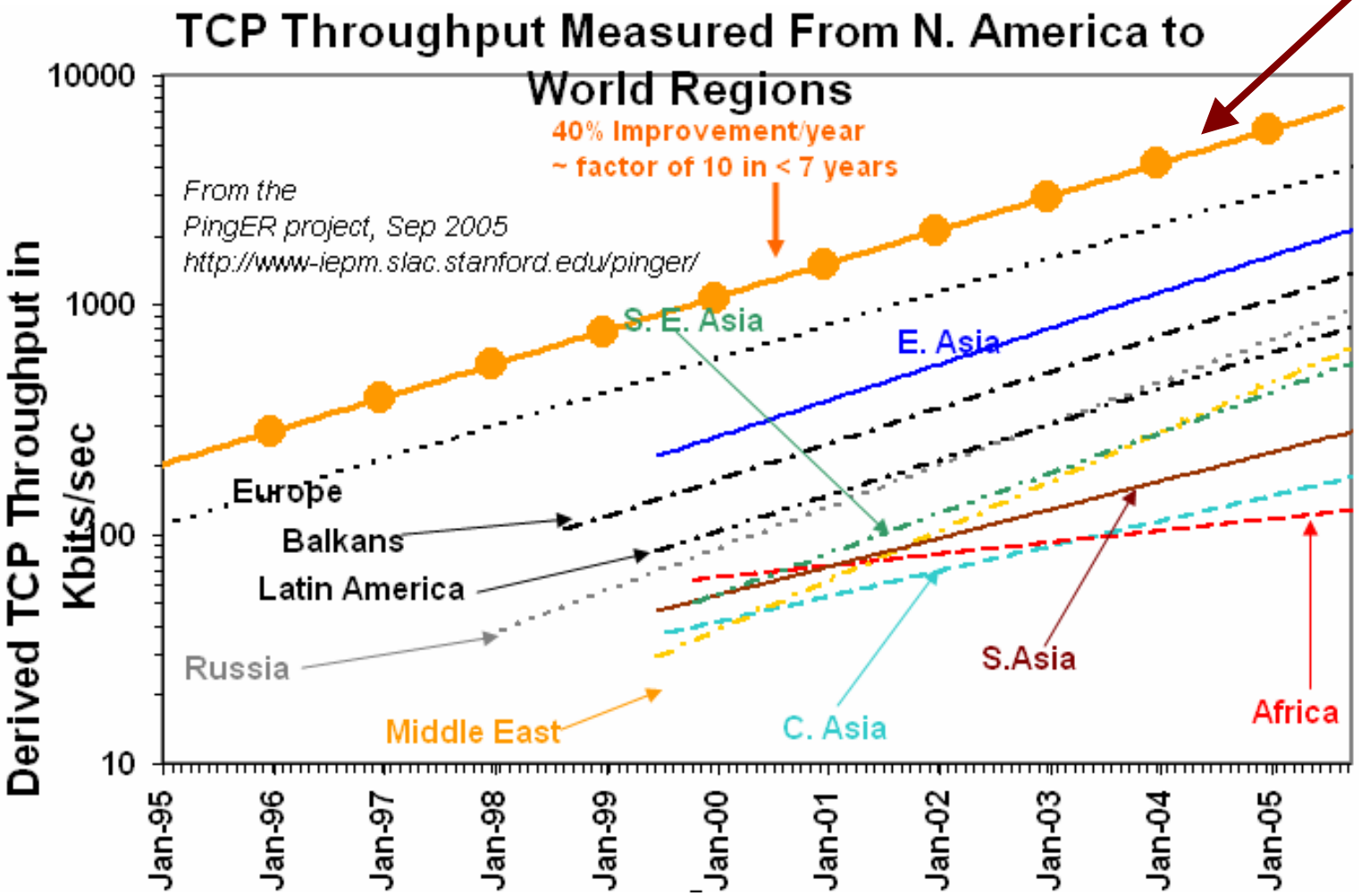
World Regions	Population (2006 Est.)	Population % of World	Internet Usage, Latest Data	% Population Penetration	Usage % of World	Usage Growth 2000-2005
Africa	915,210,928	14.1 %	22,737,500	2.5 %	2.2 %	403.7 %
Asia	3,667,774,066	56.4 %	364,270,713	9.9 %	35.7 %	218.7 %
Europe	807,289,020	12.4 %	290,121,957	35.9 %	28.5 %	176.1 %
Middle East	190,084,161	2.9 %	18,203,500	9.6 %	1.8 %	454.2 %
North America	331,473,276	5.1 %	225,801,428	68.1 %	22.2 %	108.9 %
Latin America/Caribbean	553,908,632	8.5 %	79,033,597	14.3 %	7.8 %	337.4 %
Oceania / Australia	33,956,977	0.5 %	17,690,762	52.9 %	1.8 %	132.2 %
WORLD TOTAL	6,499,697,060	100.0 %	1,018,057,389	15.7 %	100.0 %	182.0 %

INTERNET USAGE AND POPULATION IN ASIA http://internetworldstats.com						
ASIA	Population (2006 Est.)	Internet Users, (Year 2000)	Internet Users, Latest Data	Penetration (% Population)	(%) Users in Asia	Use Growth 2000-2005
China	1,306,724,067	22,500,000	111,000,000	8.5 %	30.5 %	393 %
Hong Kong *	7,054,867	2,283,000	4,878,713	69.2 %	1.3 %	113 %
India	1,112,225,812	5,000,000	50,600,000	4.5 % [*]	13.9 %	912 %
Indonesia	221,900,701	2,000,000	18,000,000	8.1 %	4.9 %	800 %
Japan	128,389,000	47,080,000	86,050,000	67.2 %	23.7 %	83 %
Korea, South	50,633,265	19,040,000	33,900,000	67.0 %	9.3 %	78 %
Malaysia	27,392,442	3,700,000	10,040,000	36.7 %	2.8 %	171 %
Pakistan	163,985,373	133,900	7,500,000	4.6 %	2.1 %	5,501 %
Philippines	85,712,221	2,000,000	7,820,000	9.1 %	2.1 %	291 %
Singapore	3,601,745	1,200,000	2,421,000	67.2 %	0.7 %	102 %
Taiwan	22,896,488	6,260,000	13,800,000	60.3 %	3.8 %	120 %
Vietnam	83,944,402	200,000	5,870,000	7.0 %	1.6 %	2,835 %
TOTAL ASIA	3,667,774,066	114,303,000	364,270,713	9.9 %	100.0 %	219 %

[* Less Than 1M Broadband Users in India]

SCIC Monitoring WG - Throughput Improvements 1995-2006

Progress: but Digital Divide is Mostly Maintained



40% annual improvement Factor ~10/7 yrs

Behind Europe

- **6 Yrs: Russia, Latin America**
- **7 Yrs: Mid-East, SE Asia**
- **10 Yrs: South Asia**
- **11 Yrs: Cent. Asia**
- **12 Yrs: Africa**

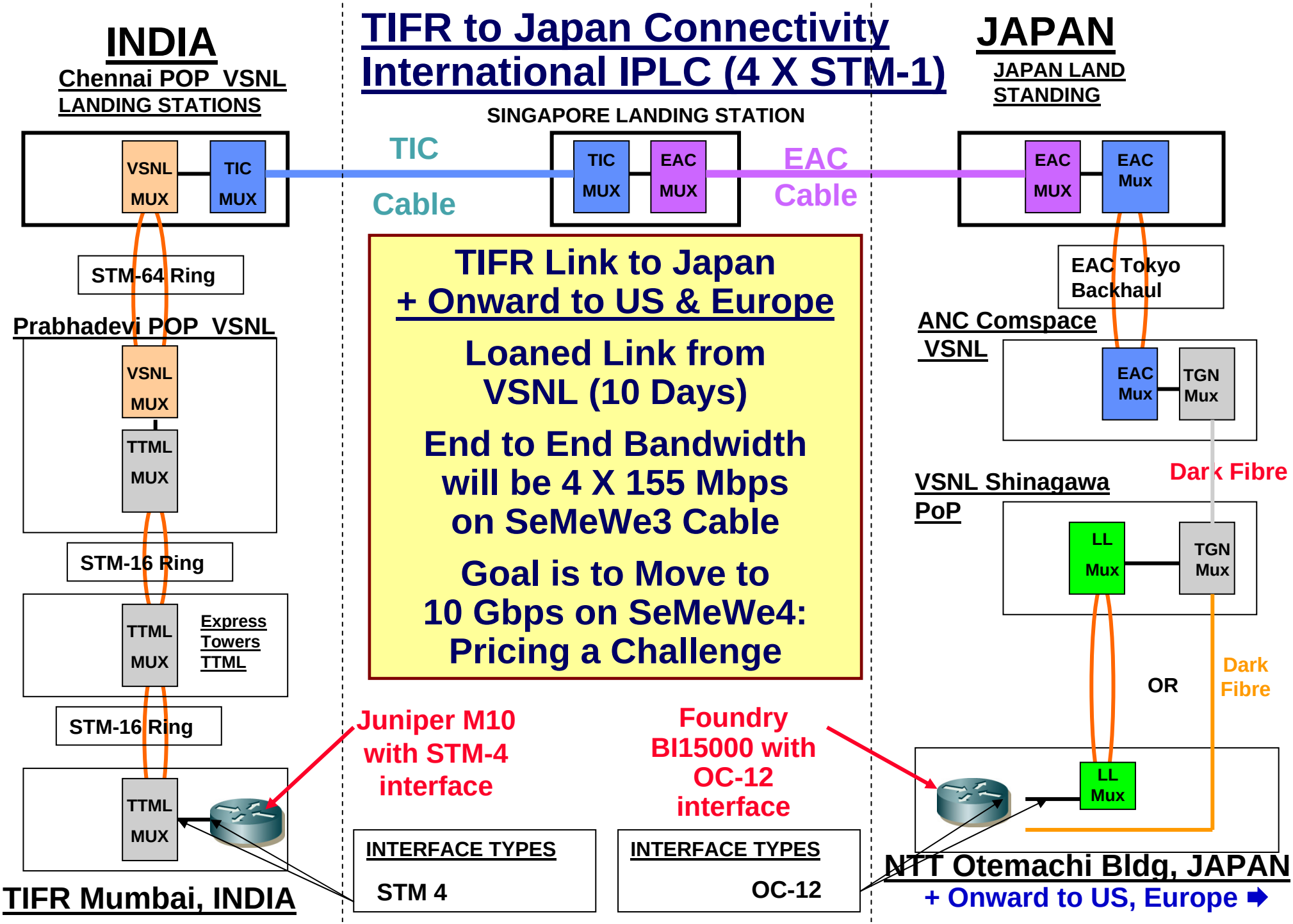
INDIA, Central Asia, and Africa are in Danger of Falling Even Farther Behind

*Bandwidth of TCP < MSS/(RTT*Sqrt(Loss))*

Matthis et al., Computer Communication Review 27(3), July 1997

SCIC Digital Divide Workshops and Panels

- ***2002-2005:***
  An effective way to raise awareness of the problems, and discuss approaches and opportunities for solutions with national and regional communities, and gov't officials
  - ***ICFA Digital Divide Workshops: Rio 2/2004; Daegu 5/2005***
  - ***Workshop on R&E Networking in Africa at CERN: Sept. 2005***
- ***In 2006***
  - ***CHEP06 Mumbai: Digital Divide Panel [SCIC, TIFR, CDAC, Internet2]***
  - ***Workshop on "Moving India into the Global Community Through Advanced Networking", February 18 [4 X 155 Mbps Link Donated by VSNL: Demos]***
  - ***Side Event to ICHEP06 (Moscow), on Networking in Russia***
  - ***SCIC meeting in Central Europe (Poland, Estonia discussed)***


INDIA
Chennai POP VSNL
LANDING STATIONS
TIFR to Japan Connectivity
International IPLC (4 X STM-1)
JAPAN
JAPAN LAND
STANDING
SINGAPORE LANDING STATION
VSNL MUX
TIC MUX
TIC
Cable
TIC MUX
EAC MUX
EAC
Cable
EAC MUX
EAC Mux
STM-64 Ring
EAC Tokyo
Backhaul
Prabhadevi POP VSNL
ANC Comspace
VSNL
VSNL MUX
TTML MUX
EAC Mux
TGN Mux
Dark Fibre
VSNL Shinagawa
PoP
STM-16 Ring
LL Mux
TGN Mux
TTML MUX
Express
Towers
TTML
OR
Dark
Fibre
STM-16 Ring
TTML MUX
LL Mux
TIFR Link to Japan
+ Onward to US & Europe
Loaned Link from
VSNL (10 Days)
End to End Bandwidth
will be 4 X 155 Mbps
on SeMeWe3 Cable
Goal is to Move to
10 Gbps on SeMeWe4:
Pricing a Challenge
Juniper M10
with STM-4
interface
Foundry
BI15000 with
OC-12
interface
INTERFACE TYPES
STM 4
INTERFACE TYPES
OC-12
TIFR Mumbai, INDIA
NTT Otemachi Bldg, JAPAN
+ Onward to US, Europe

Advanced Networks, HEP and the Digital Divide

- **Networks used by HEP and other fields of DIS are moving to the N X 10G; much faster than Moore's Law**
- **Hybrid "Dark Fiber", R&E community owned/operated hybrid networks emerging as the means to rapid progress, in a growing list of nations:**
  - **ca, nl, us, jp, kr; *pl, cz, br, no, cn, pt, ie, gr, sk, si, …***
- **HEP & CS are learning to use long range networks effectively**
  - **7+ Gbps TCP flows over 10-30 kkm; 151 Gbps Record**
- ***We Are Working to Close to Digital Divide: Equality in Collaborations***
  - ***To Allow Scientists in All World Regions to Take Part in Discoveries***
- ***HEP Groups in US, EU, Japan, Korea, Brazil, Russia …; With International R&E Network and Grid Organizations***
  - ***Are here to join as partners in a program of leadership***
  - ***For science. For research and education. For the global community.***

SCIC Work in 2005 – 2006: Partnerships

Continue Digital Divide Focus

- *Work on Specific Improvements, Case by Case:*
  - → *India with TIFR, Internet2 the World Bank, CDAC & VSNL: CHEP06 Side Event*
  - → *Russia and China: With MSU, IHEP Beijing and GLORIAD*
  - → *Pakistan with PERN and NUST*
  - → *Brazil and Latin America, with RNP, ANSP, WHREN, RedCLARA*
  - → *Asia-Pacific with Kyungpook, KEK, Aarnet, and APAN*
  - → *Africa, with IEEAF, CERN, UNU, ICTP Trieste, …*
- *New Initiatives started in 2005; Continuing in 2006: Armenia, Ukraine*
- **Help with Modernizing the Infrastructure:**
  - ❒ **Provide Tools for Effective Use: Data Transport, Monitoring, Collaboration**
  - ❒ **Design, Commissioning, Development**
- *Encourage Creation of New "Culture of Collaboration", for example in the LHC Computing (& Analysis) Models*

Extra Slides Follow

SCIC in the WSIS Stocktaking Report (July 2005): http://www.itu.int/wsis

- **The SCIC is listed in Section 7.8, E-science:**

" The International Committee on Future Accelerators: Standing Committee on Inter-regional Connectivity monitors the world's research and education networks, tracking requirements and deals especially with digital divide issues. Its main goal is to foster global scientific collaboration, so enabling scientists around the world to participate in frontier scientific discoveries."

[No other initiative cited is quoted as working towards *equality* in scientific research. Such equality is a foundation of HEP's collaborations.]

- **ICFA SCIC Brochure Presented at WSIS Phase2 (Tunis) by Hans Hoffmann**

Role of Science in the Information Society; WSIS 2003-2005

- **HEP Active in WSIS I, Geneva**
  - Theme: *"Creating a Sustainable Process of Innovation"*
  - CERN RSIS Event
  - SIS Forum & CERN/Caltech Online Stand at WSIS I (12/03) > 50 Demos: Advanced Nets & Grids, Global VideoConf., Telesurgery, "Music Grids"…
- **Visitors at WSIS I:**
  - Kofi Annan, UN Sec'y Gen'l
  - John H. Marburger, Science Adviser to US President

WSIS II: TUNIS
11/16-11/18/2005
www.itu.int/wsis

World Conference on Physics and Sustainable Development

Durban, South Africa 10/31-11/2/05

"The World Conference will serve as the first global forum to focus the physics community toward development goals and to create new mechanisms of cooperation toward their achievement."

www.saip.org.za/physics2005/WCPSD2005.html

Poland: *PIONIER* 20G + 10G Cross Border Dark Fiber Network (1/2006)

6000 km of Owned Fiber;
Multi-Lambda Core

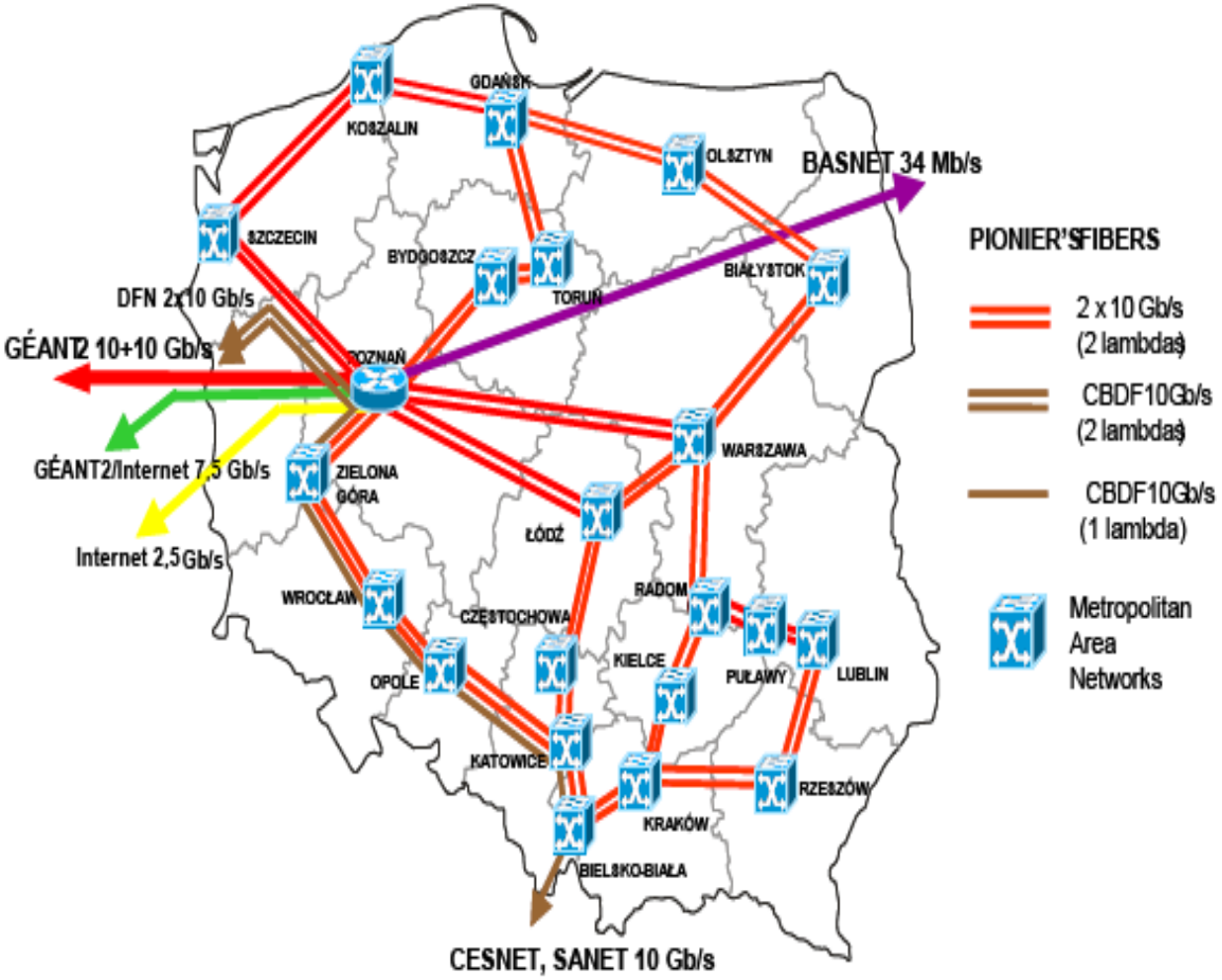
21 Academic MANs

- 19 of them have 20 or 10G backbone
- ➡ Moving to 20G on all major links

Cross Border Dark Fibers

- 20G to Germany
- 10G Each to Cz, Sk
- Move to Connect All Neighbors at 10G in 2006

20G to GEANT;
10G to Internet2



- Supports Grids, Digital Libraries, Interactive TV
- e-Regional Initiatives

M. Przybylski

Derived Throughput (kbps) Between Monitoring Countries Remote Regions

Remote regions ====v	Monitoring Country Top Level Domain ==========>													
	CH	DE	DK	HU	UK	CA	US	RU	JP	BR	IN	PK	ZA	Median
Europe	33985	9529	6504	14288	27683	2385	2214	1143	1151	1371	704		1113	2299
N. America	2200	2298	2081	2134	3089	607932	66557	546	2128	1529	902		963	2131
Russia	4543	3394	3445	2857	3049	1137	1362	2865	910		528	110	861	2109
S.E. Europe	5446	5464	5330	18427	3824	2048	1492	1041	1129		641		1096	2048
E. Asia	767	1165	1060	1159	1330	1675	2277	631	103904	590	1869	115	523	1159
M. East	917	1489	1303	1303	1250	758	964	420	472		260		498	917
Oceania	744			1079		1447	1469	349			802		656	802
S.E. Asia	540						1055							798
L. America	842	617	482	567	594	706	1289	261	496	15980	265	98	375	567
Africa	450	428	596	548	977		423	266	239	784			652	499
C. Asia	275						323							299
S. Asia	371						1957			97	798	203	81	287
Median	804	1894	1692	1303	2190	1561	1415	546	1020	1078	704	113	654	859
	CH	DE	DK	HU	UK	CA	US	RU	JP	BR	IN	PK	ZA	

Good > 5000 kbps; **Acceptable** 1000 to 5000 kbps
Poor 500 to 1000 kbps; **Very Poor** < 500 kbps

- Intra-Continental Europe (Including Russia, SE, Baltics); Intra-US: *Much Improved*
- Latin America (to Europe & Asia); Mid-East, Central and So. Asia, Africa: *Poor to Very Poor, and Far Behind*

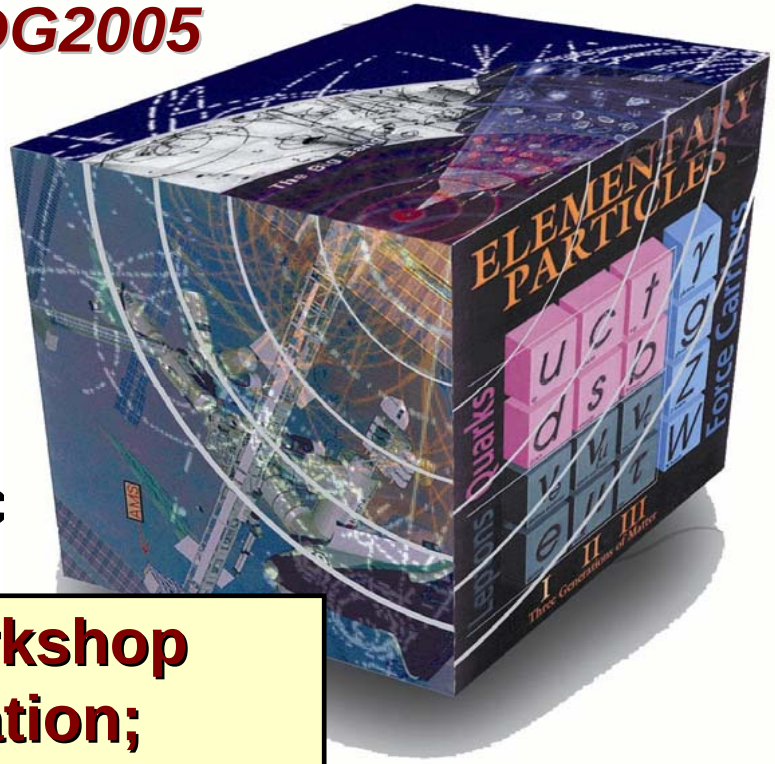
What about lambda switching?

- Two factors argue that it will be a long time before we have lambda switching services in production networks
  1) There will not be enough lambdas available to satisfy the need
     - Just provisioning a single lambda ring around the US (11,000km) is still about $2M, even on R&E networks
     - However, this should drop by a factor of 5-10 over next decade
  2) Even if there were a "lot" of lambdas (hundreds?) there are thousands of large-scale science users
     - Just considering sites (and not scientific groups) there are probably 300 major research science research sites in the US and a comparable number in Europe
- So, lambdas will have to be shared for the foreseeable future
  - Multiple QoS paths per lambda
  - Guaranteed minimum level of service for best effort traffic when utilizing the production IP networks
  - Allocation management
    - There will be hundreds to thousands of contenders with different science priorities

W. Johnston, ESnet

International ICFA Workshop on HEP Networking, Grids, and Digital Divide Issues for Global e-Science

http://chep.knu.ac.kr/HEPDG2005



➡ Focus on Asia-Pacific

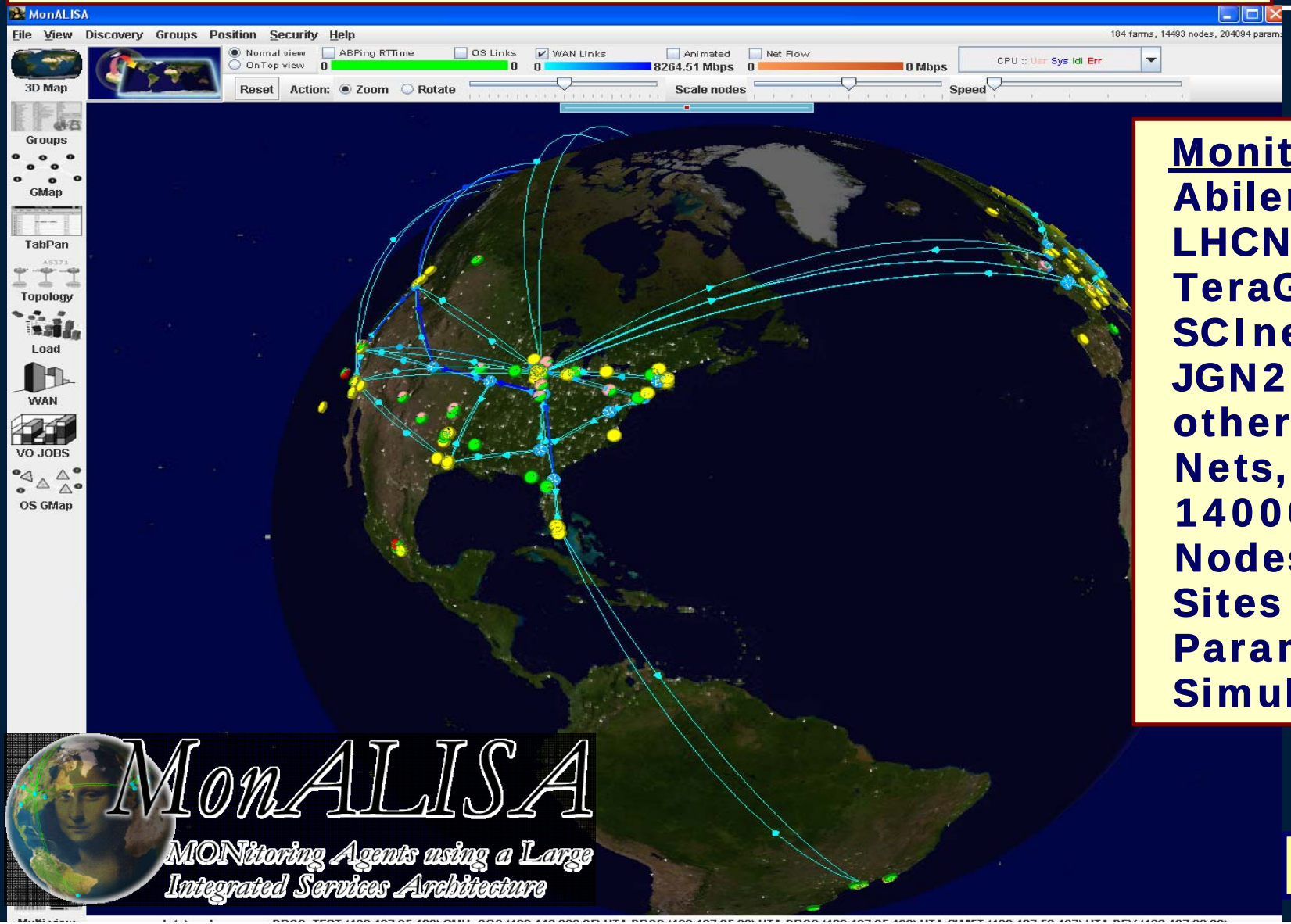
May 23-27, 2005
Daegu, Korea

Dongchul Son
Center for High Energy Physic

➡ Following successful Rio Workshop
➡ Brazil HEPGrid Tier2 Inauguration; joined Open Science Grid
➡ 2006: ICHEP06 Side Event; Workshop in Central Europe

HEP at SC2005

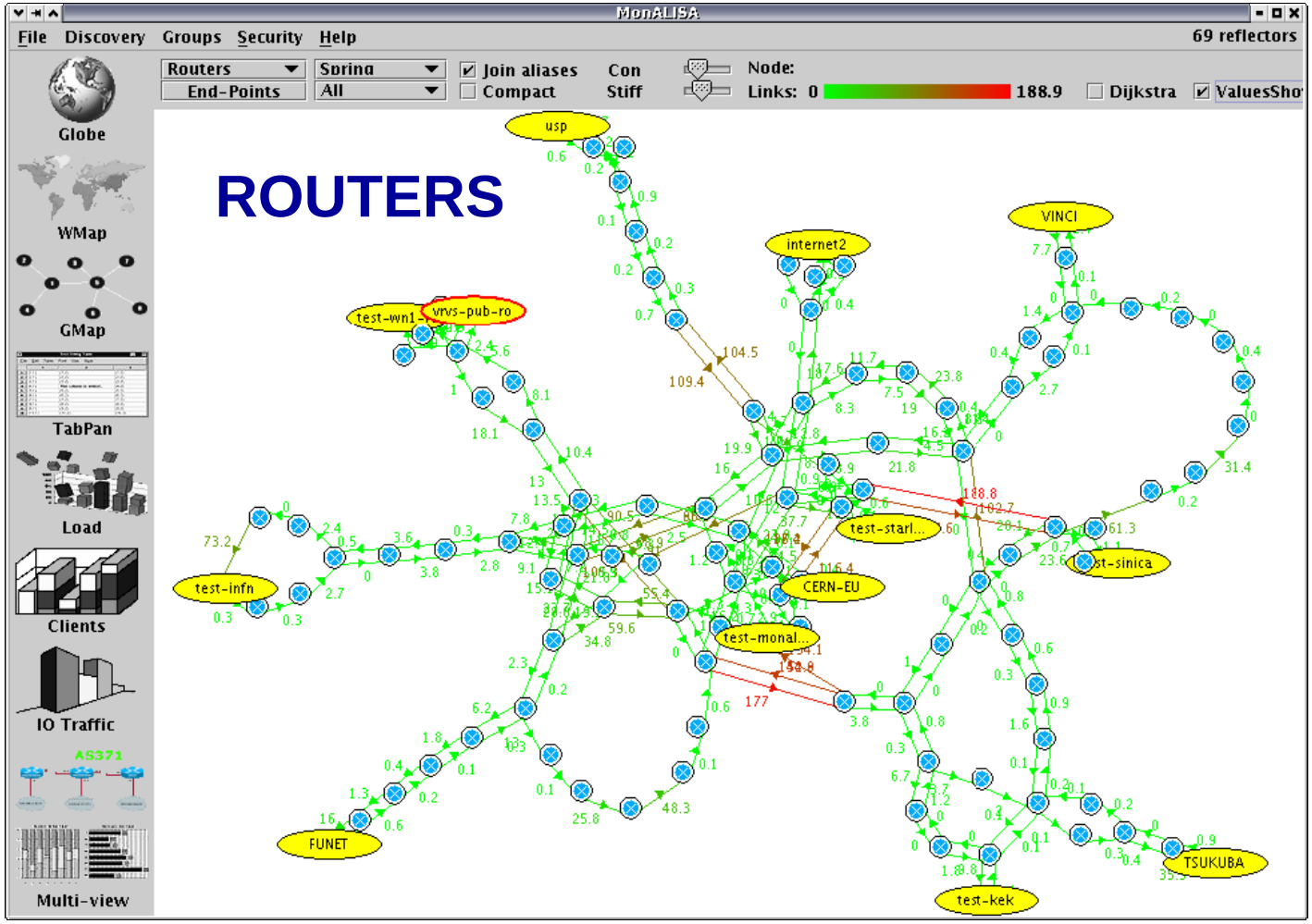
Global Lambdas for Particle Physics



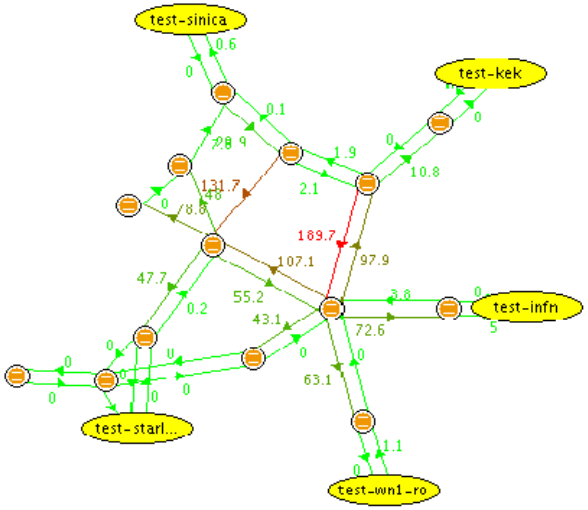
Monitoring NLR, Abilene/HOPI, LHCNet, USNet, TeraGrid, PWave, SCInet, Gloriad, JGN2, WHREN, other Int'l R&E Nets, and 14000+ Grid Nodes at 250 Sites (250k Parameters) Simultaneously

I. Legrand

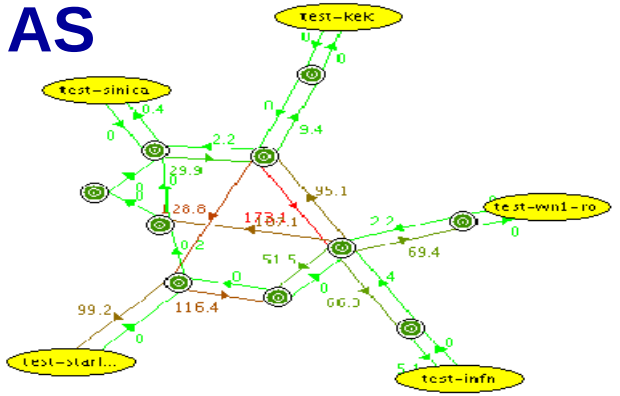
MONALISA Discovering and Monitoring Network Topology, Latency, Performance



NETWORKS

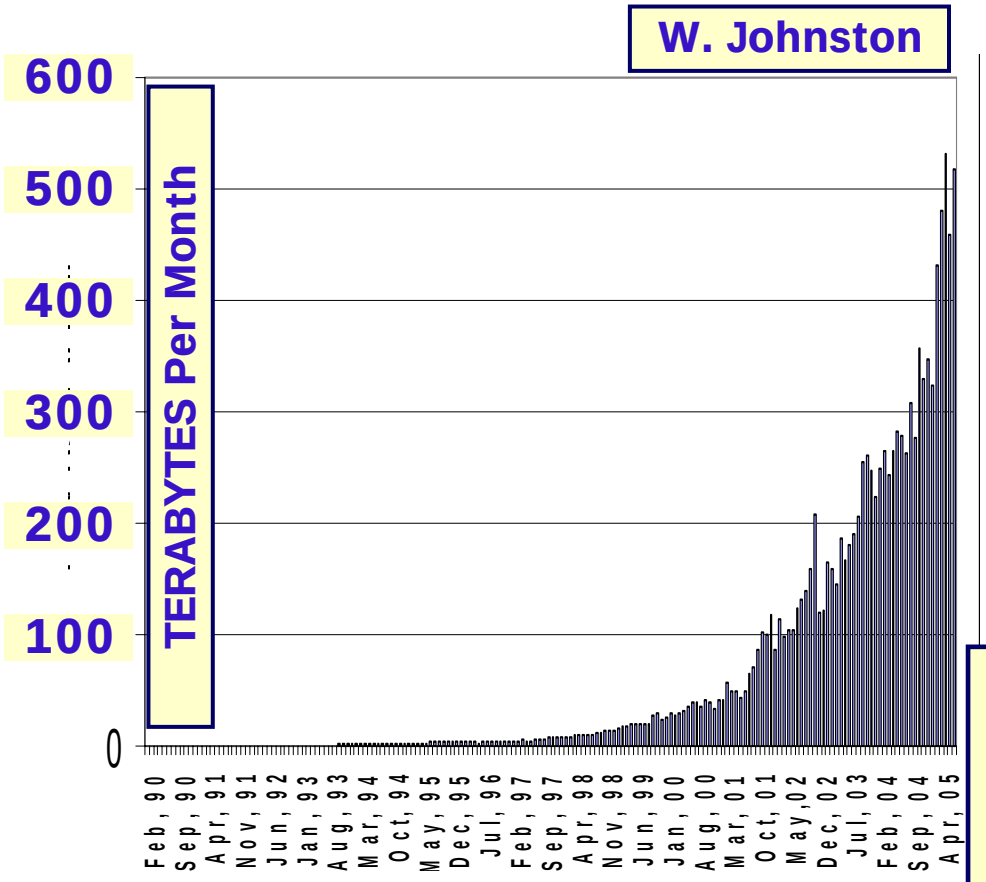


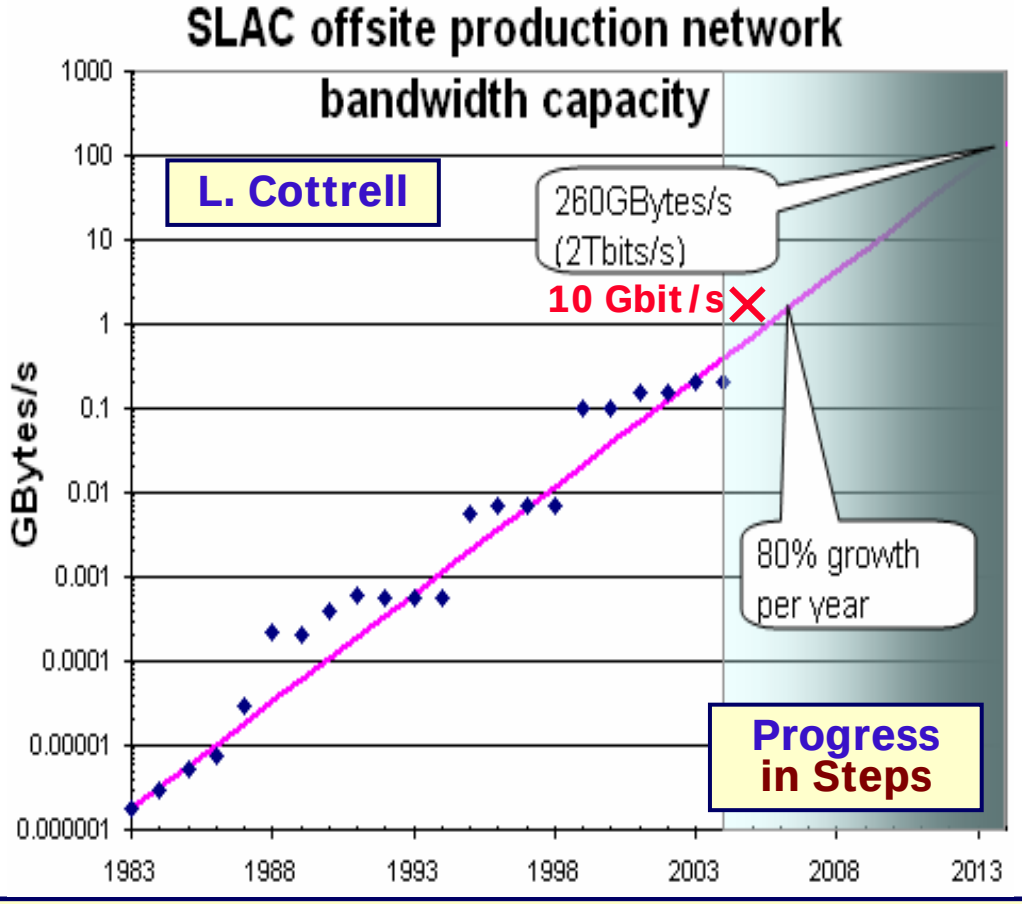
AS



Long Term Trends in Network Traffic Volumes: 300-1000X/10Yrs

ESnet Accepted Traffic 1990 – 2005
Exponential Growth:
Avg. +82%/Year for the Last 15 Years





- **SLAC Traffic ~400 Mbps; Growth in Steps (ESNet Limit): ~ 10X/4 Years.**
- **Summer '05: 2x10 Gbps links: one for production, one for R&D**
- **Projected: ~2 Terabits/s by ~2014**

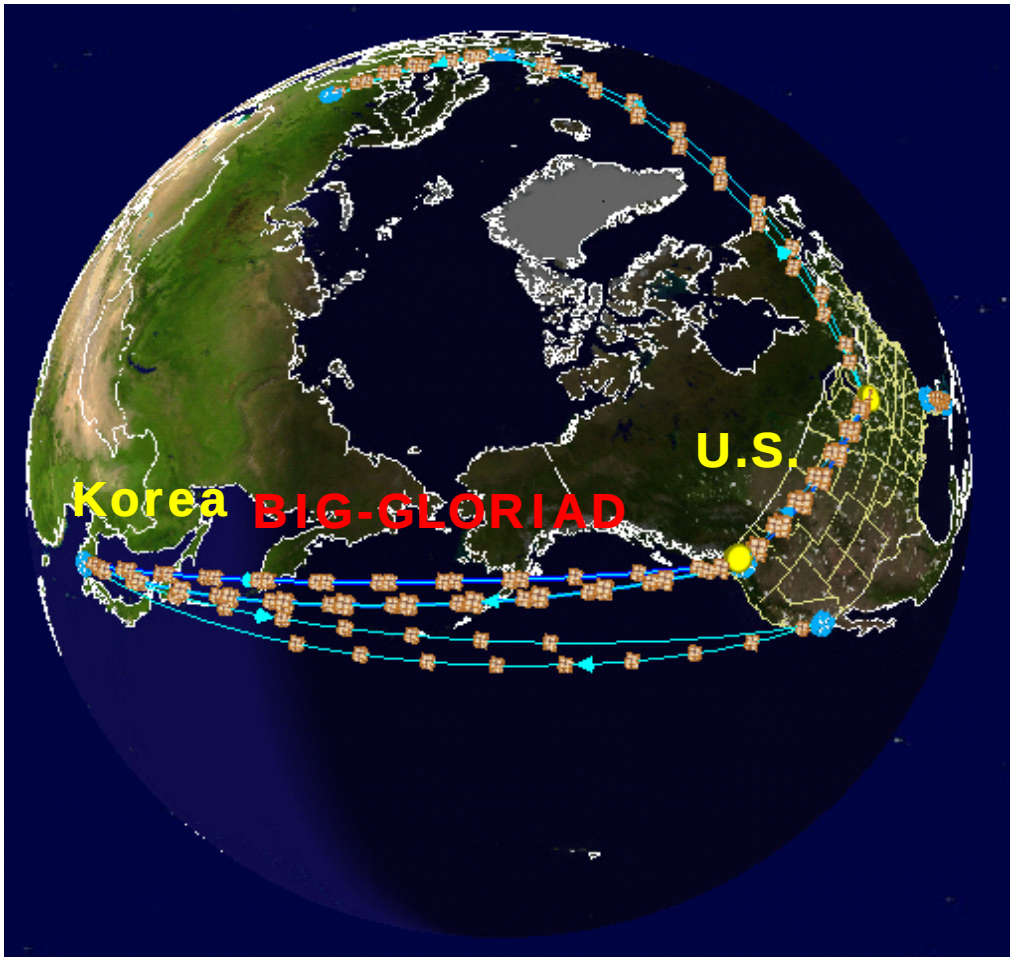
Data Samples and Transport Scenarios

10^7 Event Samples	Data Volume (TBytes)	Transfer Time (hrs) @ 0.9 Gbps	Transfer Time (hrs) @ 3 Gbps	Transfer Time (hrs) @ 8 Gbps
AOD	0.5-1	1.2 – 2.5	0.37-0.74	0.14 – 0.28
RECO	2.5 - 5	6 - 12	1.8 – 3.7	0.69 – 1.4
RAW+RECO	17.5 - 21	43 - 86	13 - 26	4.8 – 9.6
MC	20	98	30	11

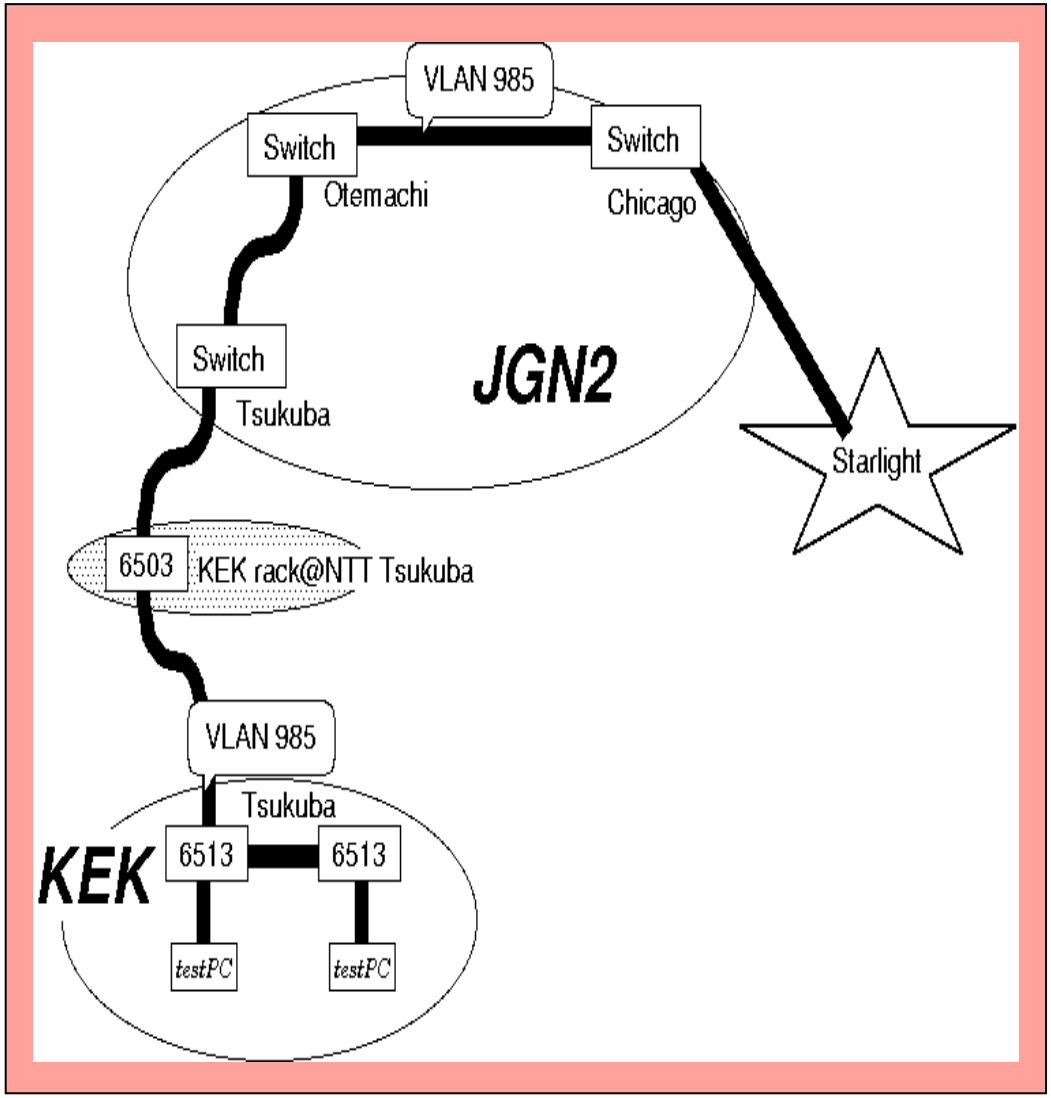
- 10^7 Events is a typical data sample for analysis or reconstruction development [Ref.: MONARC]; equivalent to just ~1 day's running
- Transporting datasets with quantifiable high performance is needed for efficient workflow, and thus efficient use of CPU and storage resources
- ➡ One can only transmit ~2 RAW + REC or MC samples per day on a 10G path
- ✸ *Transport of significant data samples will require one, or multiple 10G links*
- ➡ *Movement of 10^8 event samples (e.g. after re-reconstruction) will take ~1 day (RECO) to ~1 week (RAW, MC) with a 10G link at high occupancy*

KNU (Korea) at SC05

- Uses 10Gbps GLORIAD link from Korea to US, which is called BIG-GLORIAD, also part of UltraLight
- Try to saturate this BIG-GLORIAD link with servers and cluster storages connected with 10Gbps
- Korea is planning to be a Tier-1 site for LHC experiments



KEK (Japan) at SC05
10GE Switches on the
KEK-JGN2-StarLight Path



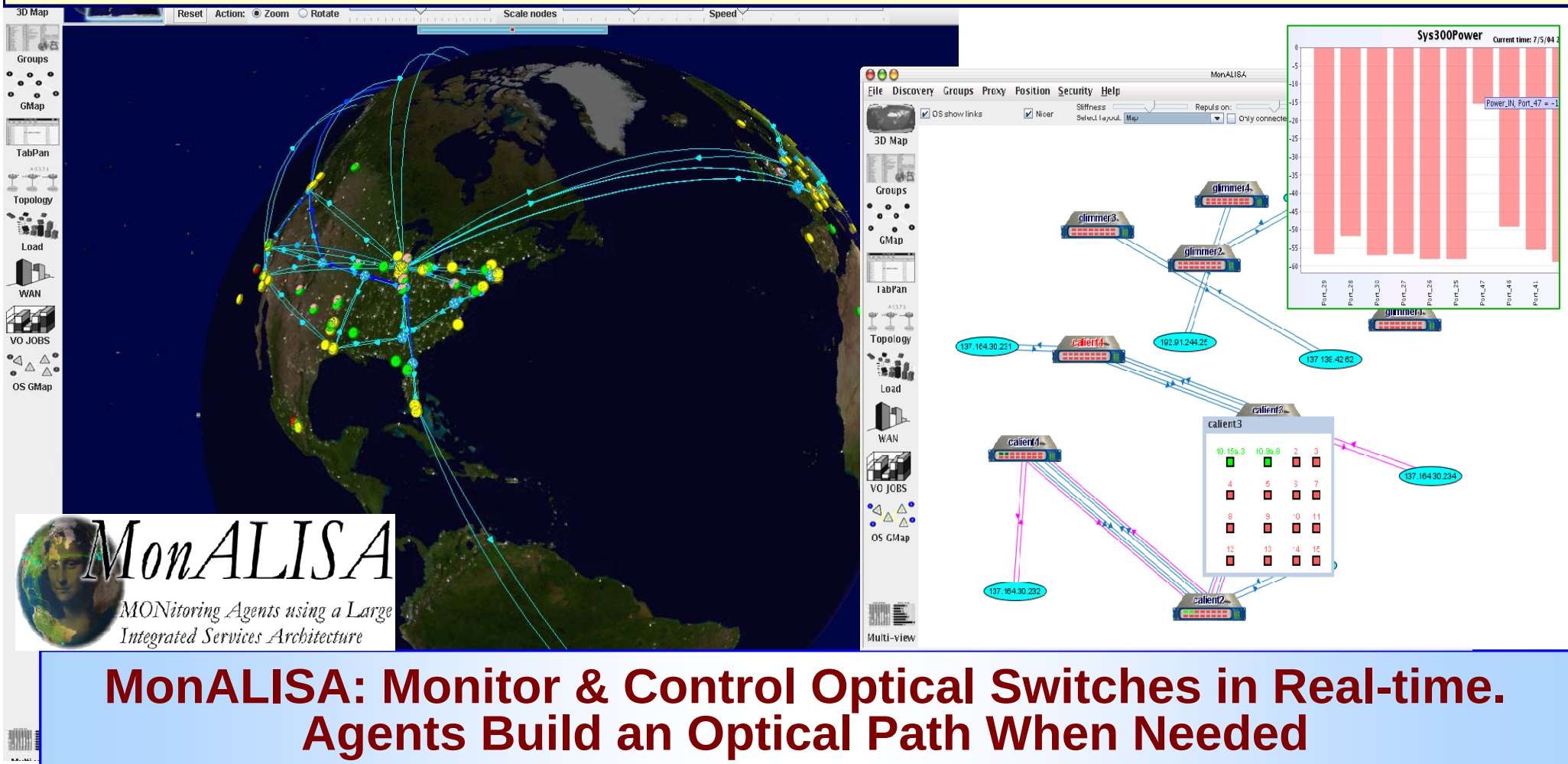
JGN2: 10G Network Research Testbed

- Operational since 4/04
- 10Gbps L2 between Tsukuba and Tokyo Otemachi
- 10Gbps IP to Starlight since August 2004
- 10Gbps L2 to Starlight since September 2005

Otemachi–Chicago OC192 link replaced by 10GE WANPHY in September 2005

A New Network & Grid Services Paradigm

- Grids with networks as an active, managed element
- Global services, dynamic bandwidth provisioning for workflow optimization
  - *Agent-based services spanning all layers*

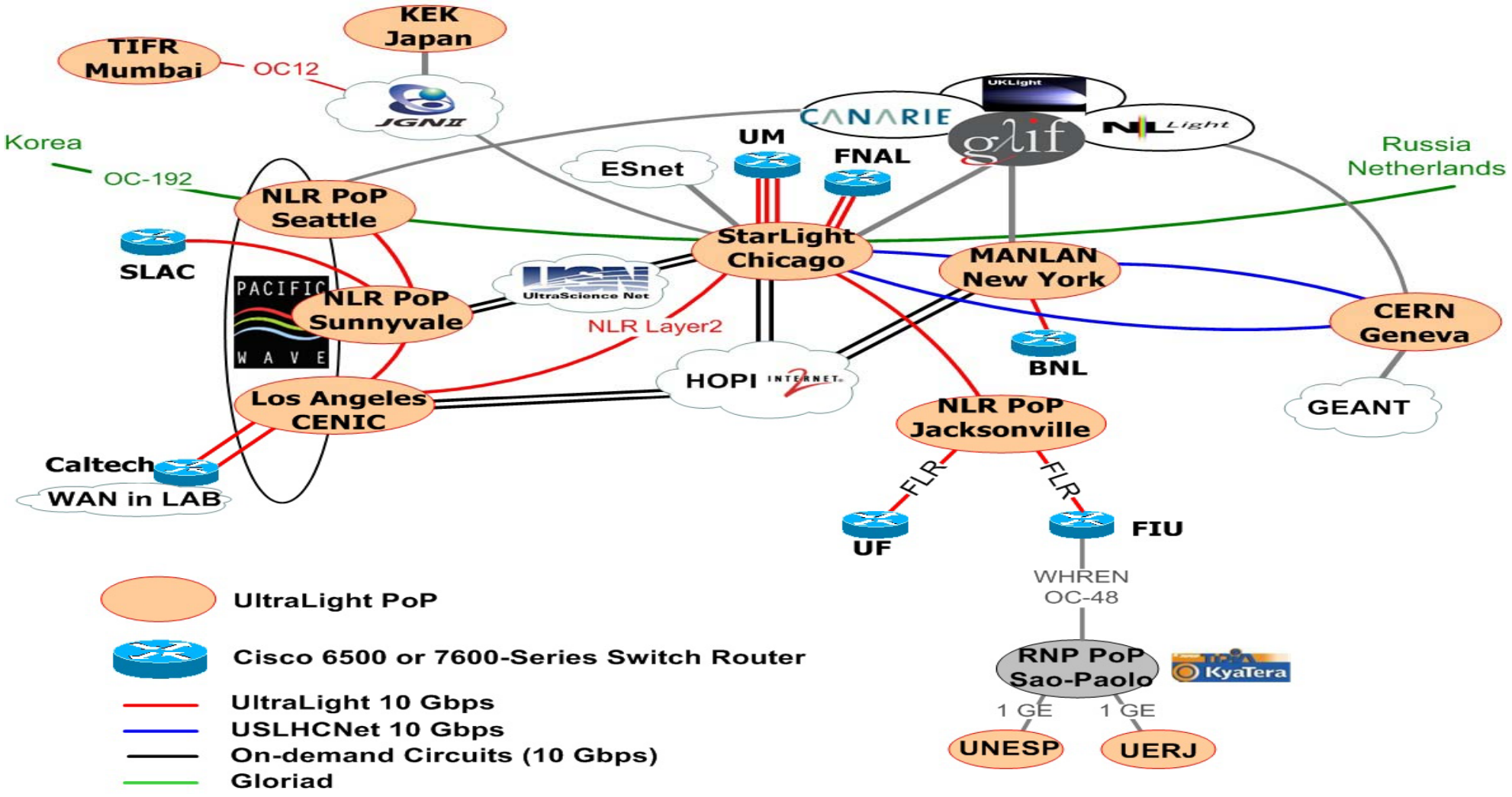


MonALISA: Monitor & Control Optical Switches in Real-time. Agents Build an Optical Path When Needed

SC05 BWC Takeaways & Lessons

- **Substantive take-aways from this Marathon exercise:**
  - **An optimized Linux kernel (2.6.12 + FAST-TCP + NFSv4) for data transport; after 7 full kernel-build cycles in 4 days**
  - **Scaling up SRM/gridftp to near 10 Gbps per wave, using Fermilab's production clusters**
  - **A newly optimized application-level copy program, bbcp, that matches the performance of iperf under some conditions**
  - **Extensions of SLAC's Xrootd, an optimized low-latency file access application for clusters, across the wide area**
  - **Understanding of the limits of 10 Gbps-capable computer systems, network switches and interfaces under stress**

4 Continent Testbed



ICFA Standing Committee on Interregional Connectivity (SCIC)

- **Created in July 1998 in Vancouver ; Following ICFA-NTF**

CHARGE:

- **Make recommendations to ICFA concerning the connectivity between the Americas, Asia and Europe**
- **As part of the process of developing these recommendations, the committee should**
  - **Monitor traffic on the world's networks**
  - **Keep track of technology developments**
  - **Periodically review forecasts of future bandwidth needs, and**
  - **Provide early warning of potential problems**
- **Representatives: Major labs, ECFA, ACFA, North and Latin American Users, Russia, China**

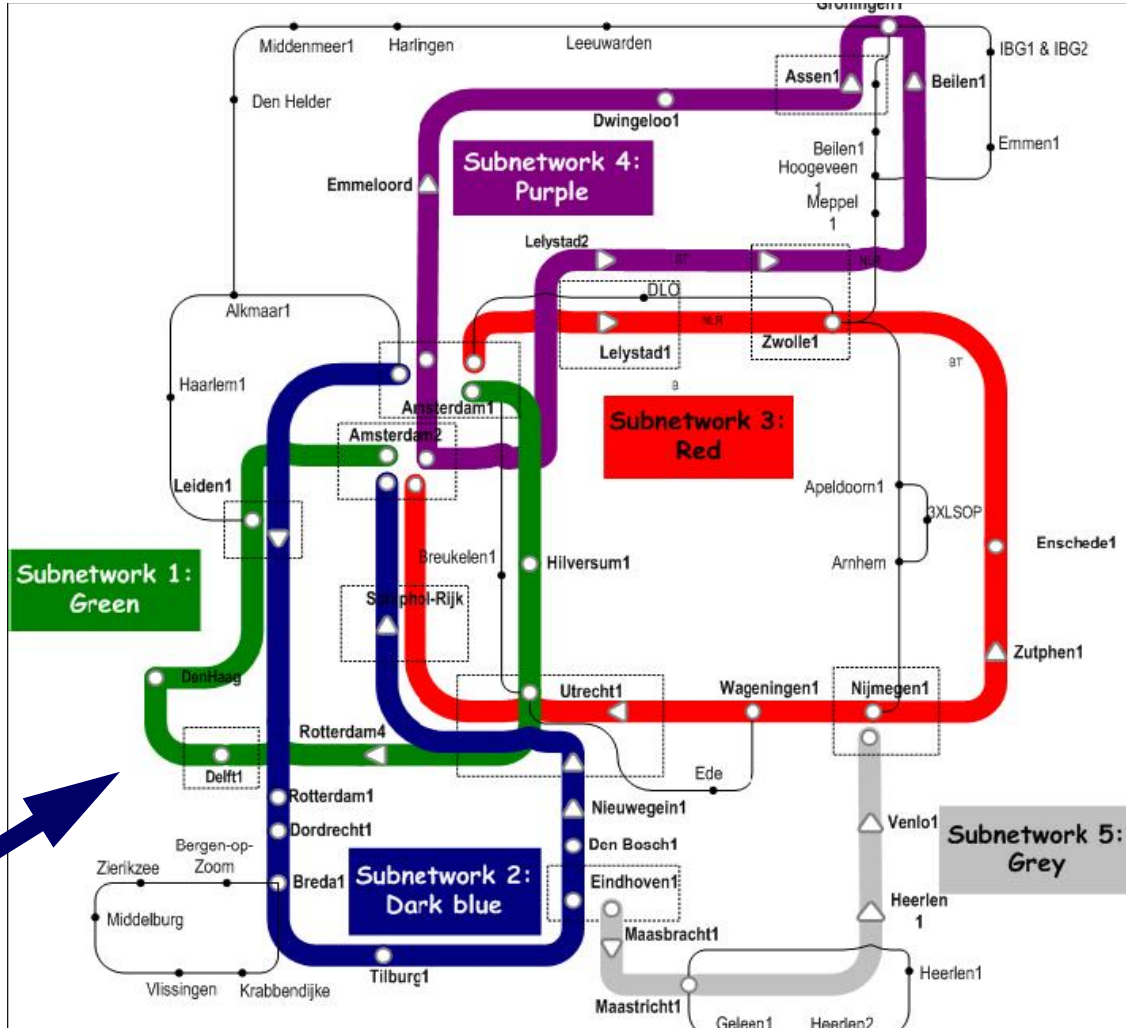
Tier-2s

The Proliferation of Tier2s
➡ LHC Computing will be More Distributed and Dynamic

~100 Identified – Number still growing

J. Knobloch

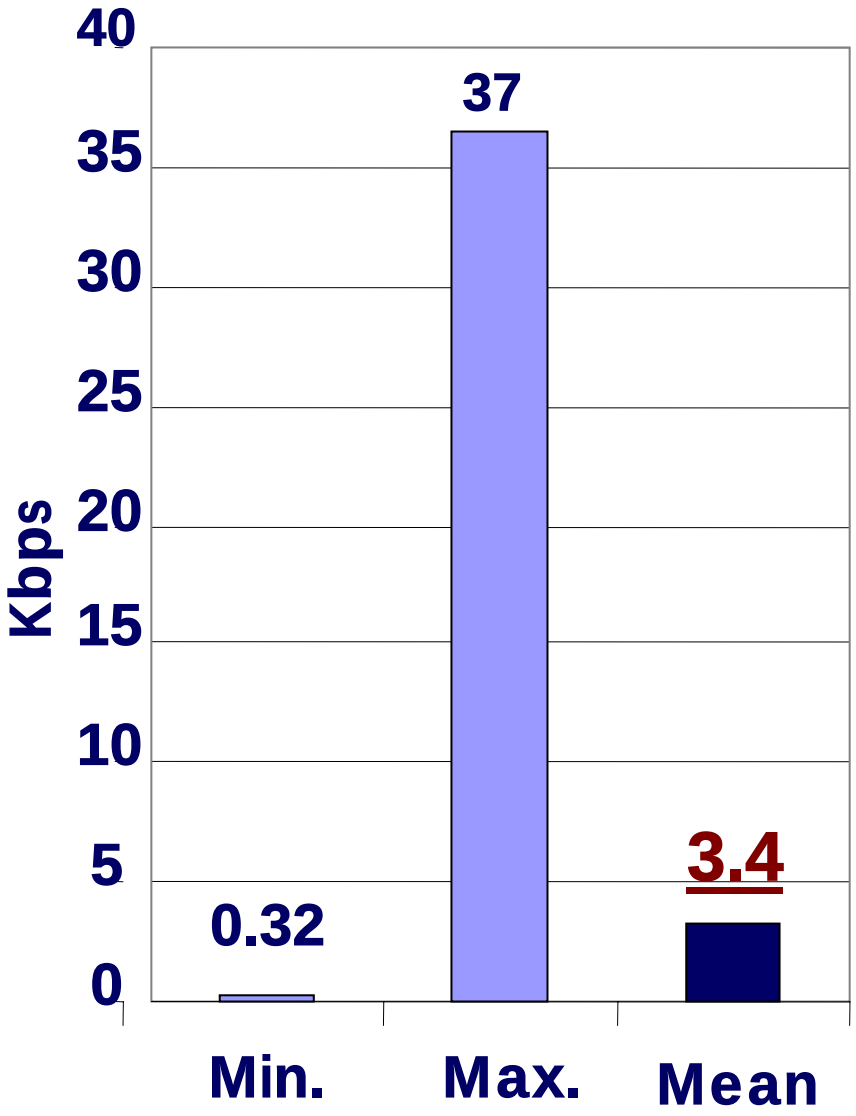
SURFNet6 in the Netherlands
5300 km of Owned Dark Fiber



Optical Layer: 5 Rings
Up to 72 Wavelengths
Support for HEP,
Radioastronomers
Medical Research

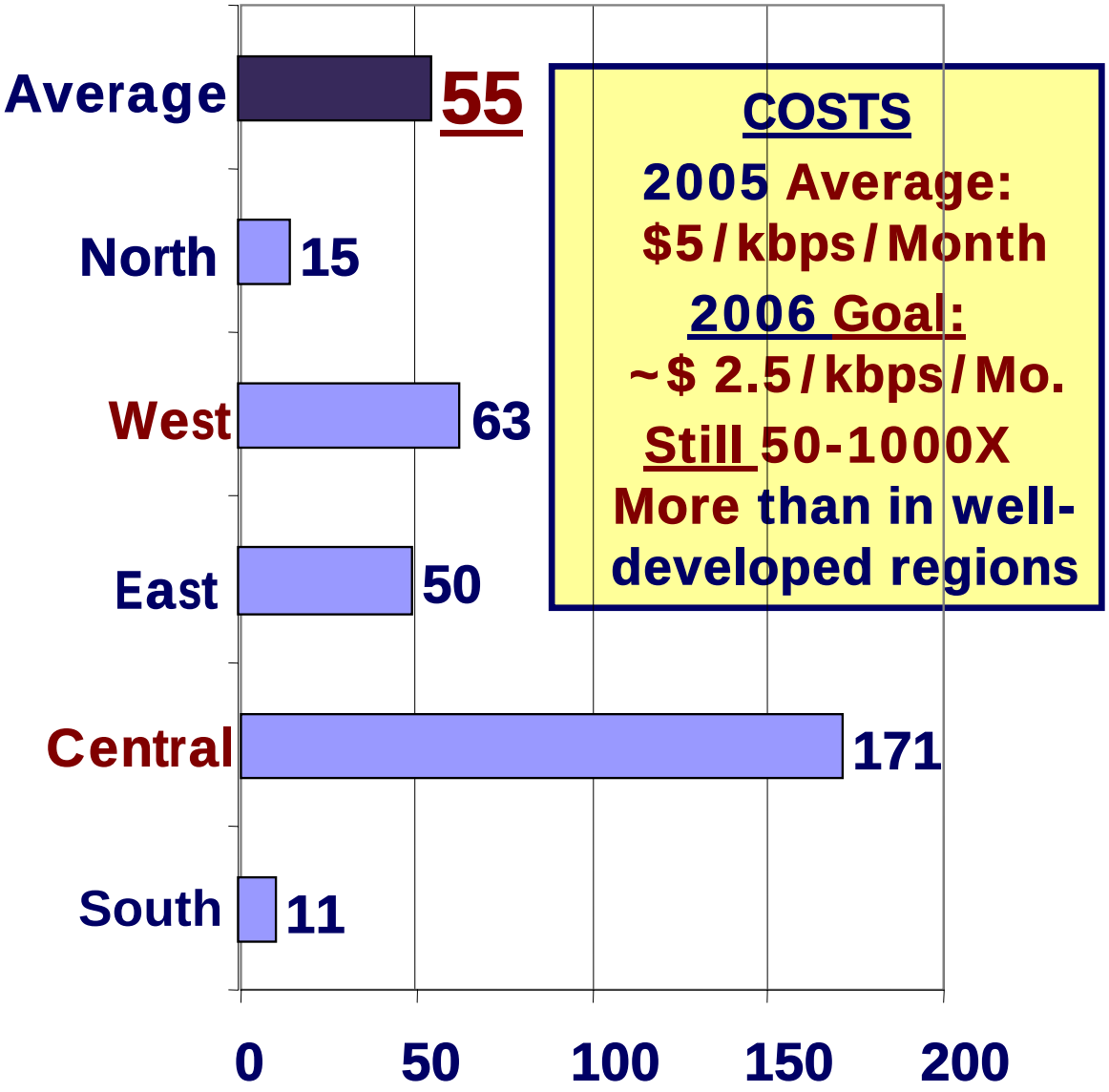
K. Neggers

AFRICA: Bandwidth per networked computer (Kbps)



M. Jensen

Users Per Networked Computer by African Region



COSTS

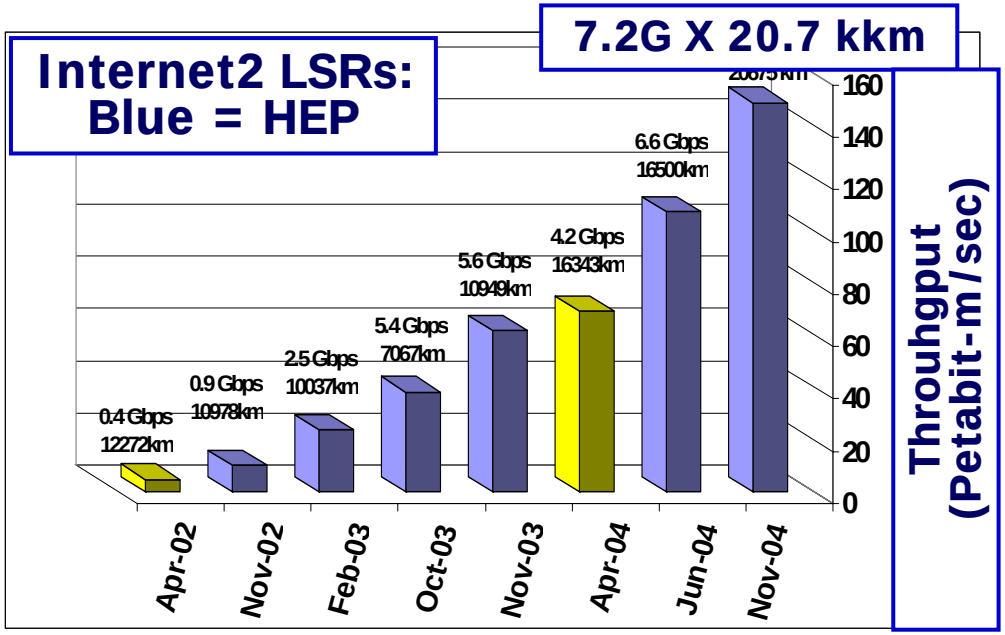
2005 Average: $5/kbps/Month

2006 Goal: ~$ 2.5/kbps/Mo.

Still 50-1000X More than in well-developed regions

Internet2 Land Speed Records & SC2003-2005 BWC Records

- Single Stream 7.5 Gbps X 16 kkm with Linux: July 2004
- IPv4 Multi-stream record with FAST TCP: 6.86 Gbps X 27kkm: Nov 2004
- PCI-X 2.0: 9.3 Gbps Caltech-StarLight
- *Concentrate now on reliable Terabyte-scale file transfers*
  - ➔ *Disk-to-disk Marks: 536 Mbytes/sec (Windows); 500 Mbytes/sec (Linux)*
  - ➔ *System Issues: PCI-X Bus, Network Interfaces, Disk I/O Controllers, Linux Kernel,CPU*



Nov. 2004 Record Network

NB: Computing Manuf.'s Roadmaps for 2006: One Server Pair ~ One 10G Link

SUPERCOMPUTING 2005: 151 Gbps Bandwidth Challenge Record