International Telecommunication Union

ITU-T

Make way for Grids

Franco Travostino

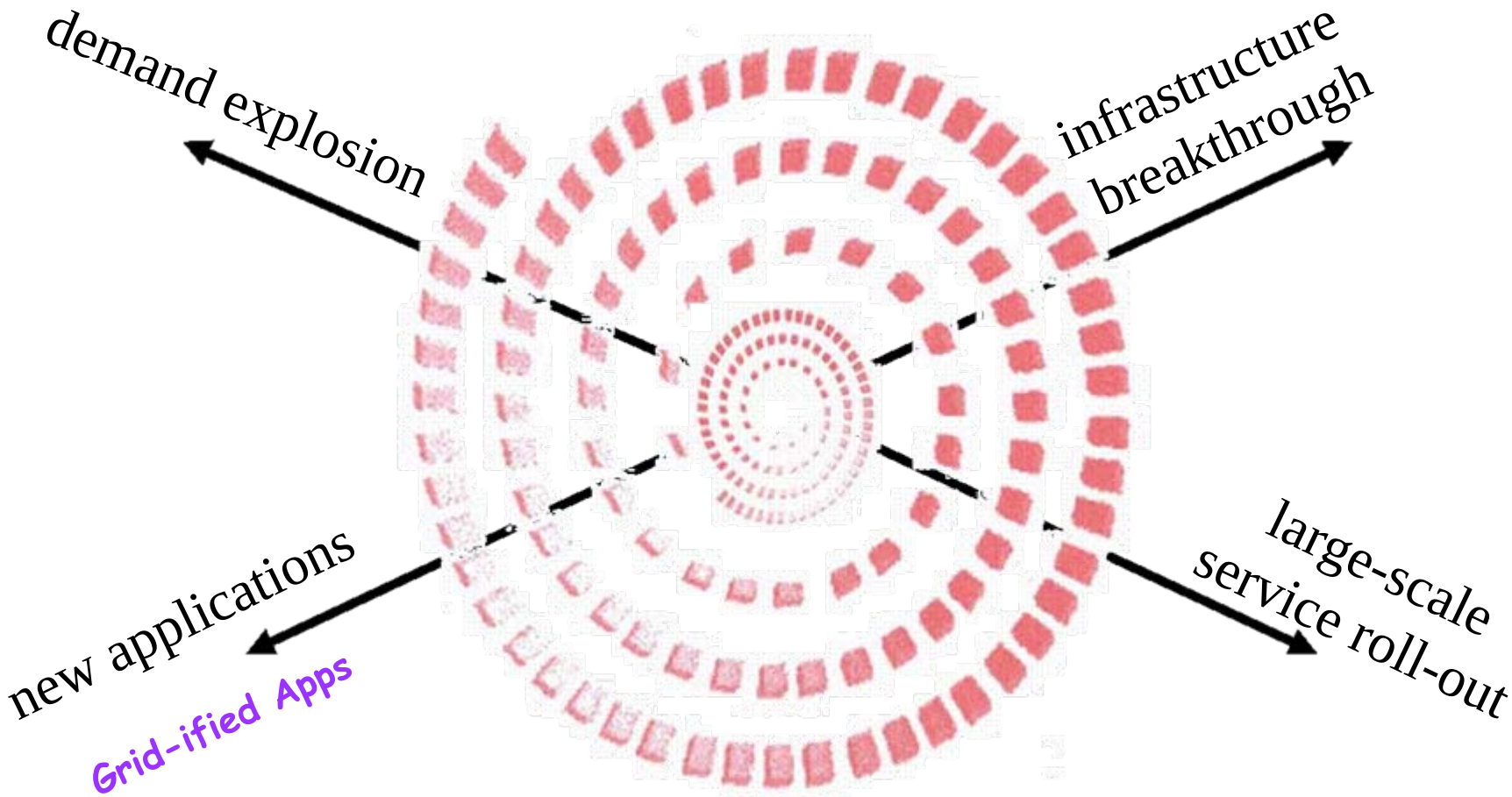
Director of Advanced Technology and Research, Nortel

Area Director for Infrastructure, Open Grid Forum

OpenGridForum

ITU-T/OGF Workshop on Next Generation Networks and Grids
Geneva, 23-24 October 2006

Placing Grids in the Spiral of Goodness



- Fine-tuning iterations required, applications ↔ services
- High stakes: Low price of admission vs. Arrested development

Recurring traits of Grid traffic

> When it rains, It ***pours***: The better the pipe, the more traffic wants to go in
  - "better pipe" = lower latency or higher throughput, depending on Grid type
  - parallel streams make it less likely that traffic will be storage- or I/O-bound

> Geographically-significant extents
  - from intranets quickly onto extranets, often all the way to global scale
  - A Metcalfe law at play, substitute users w/ distant high-capacity sites

> A node pushes data into the network to one or more recipients
  - often, a whole "herd of elephants" has to move across relatively few locations

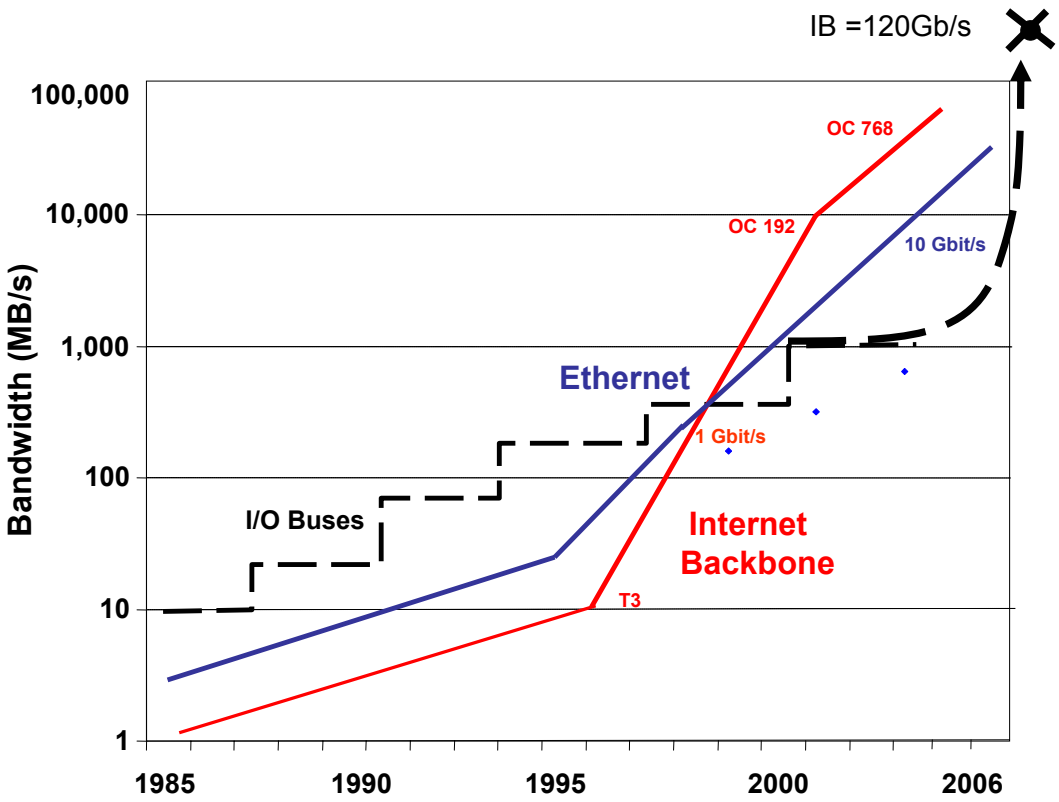
> Workflow-oriented more so than periodic
  - ***someone*** knows what's the next step after current step finishes
  - network requirements, physical span to expand or contract accordingly

Over…

> Wireline
- IP services
- Ethernet services
- L1/L0 services

> Wireless
- w/ Cellular, WiFi, WiMax transport
- e.g., for data acquisition from sensors
- e.g., for untethered access and control of a Grid

System I/O is the first hop in the network

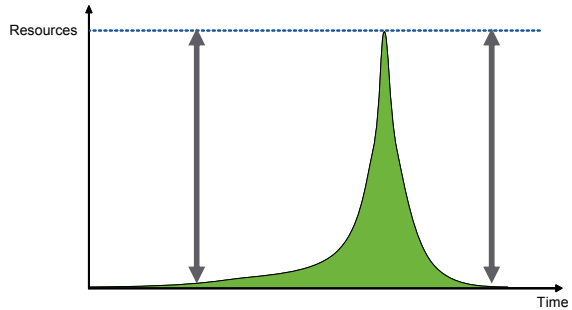


> MAN/WAN rates had surpassed legacy I/O rates (e.g., PCI)

> Infiniband™ et al. restore I/O lead and re-open traffic flood gates

Problems with unknowable hardwired networks

> Peak-provisioning… steep or elusive peak
- "Over-provisioning is the only QoS that you can really depend on"
- Especially across multiple domains

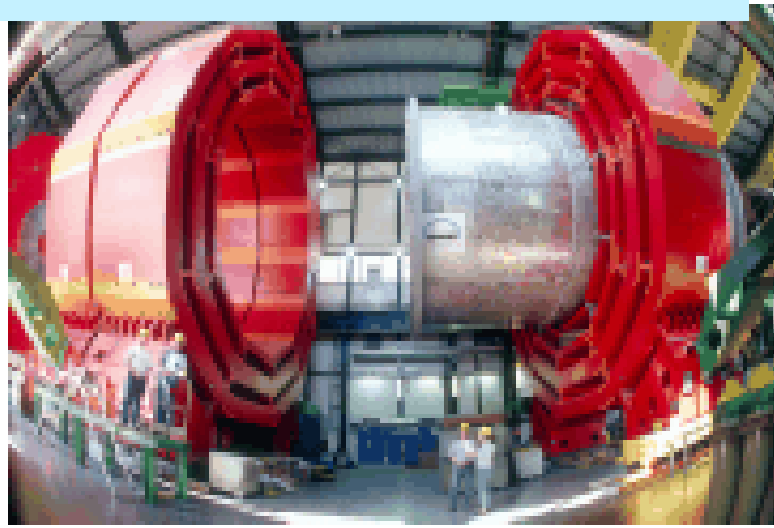


> CoS tiered service
- Only mainstream SLAs admitted
- Lag before the supply side reflects new requirements in a SLA

> New apps urged to go global, yet without any perf compromise

> How likely that *everyone* plays by the rules
- New apps are plain un-tested or outright buggy
- Exploits: from TCP Daytona to (D)DoS

Contrasts



> Soon, Cern's LHC will post 10+ PB new data per year to researchers worldwide

vs.

> One size fits all networks

" *a standard TCP connection with 1500-byte packets and a 100 ms round-trip time, achieving a steady-state throughput of 10 Gbps would require an average congestion window of 83,333 segments and a packet drop rate of, at most, one congestion event every 5,000,000,000 packets*" (Floyd, '03)

Requirements posed to the network

- High performance
  - in bit blasting and/or low latency
- Virtualization
  - abstracted network resources
- Finesse
  - fine granularity of control (e.g., app hints service allocation)
- Site autonomy
  - SLA-based federation
- Advance reservation of service
  - Footholds that I can rely on
- Resource bundling
  - network AND {cpu | data | etc.}
- Determinism
  - can predict completion
- Collision-free addressing
  - hitless VO joins
- Security
  - non-interference, can work w/ firewalls
- Multicast, anycast
  - efficient data displacement, lazy-evaluated binding
- Automation-friendly software control
  - no GUIs, no Operators
- Feedback loops
  - between network and higher-order software

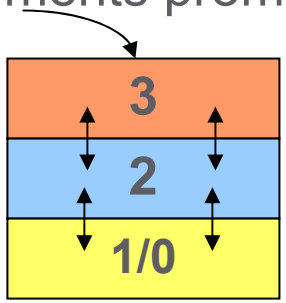
Solution Domains

> IP, premium or best-effort

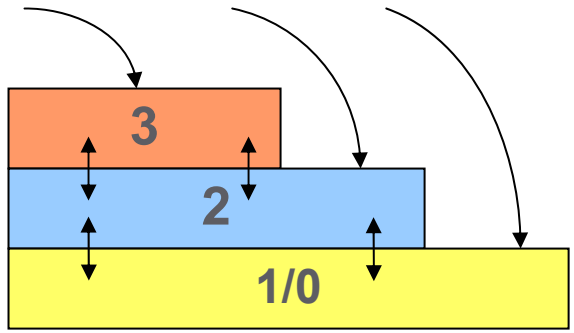
> Provider-class Ethernet
- E-line and E-LAN w/ hierarchical scaling, virtualization, and isolation

> "Wavelength" services
- TDM/WDM circuits consuming a wavelength or fraction/multiple thereof

Grid requirements prompt for:



a)
cross-layer coordination



b)
services directly-matched to appl.

What's wrong with an IP CBR service?

> Low-latency applications will suffer queuing delay & jitter
  - **Most noticeable over short distances**

> High-throughput applications demand high rates
  - **Cost, cost, cost**

> Non-negligible packet loss and restoration times

> Questionable isolation
  - **Against buggy or outright malicious applications**

What's wrong with a wavelength service?

> Physical constraints to connectivity matrix

> Set-up time

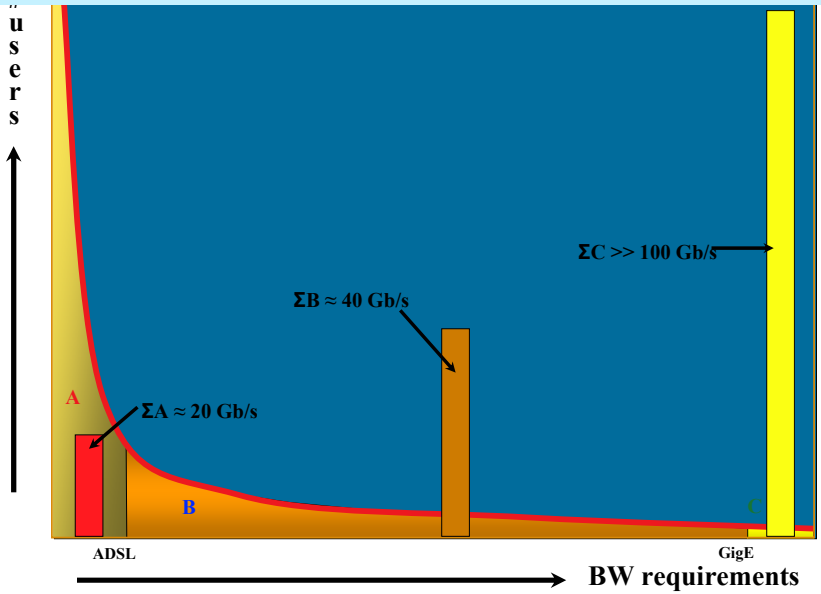
> Challenges in overall utilization

> Blocking probability exposed

} Can be mitigated via software

The rise of hybrid packet optical solutions

- SURFnet6 (2005 – …)
  - 1st proof point of country scale
  - 1-10 Gb/s dynamic wavelength svcs
  - ref. Kees Neggers' keynote



deLaat's Curve (2003)

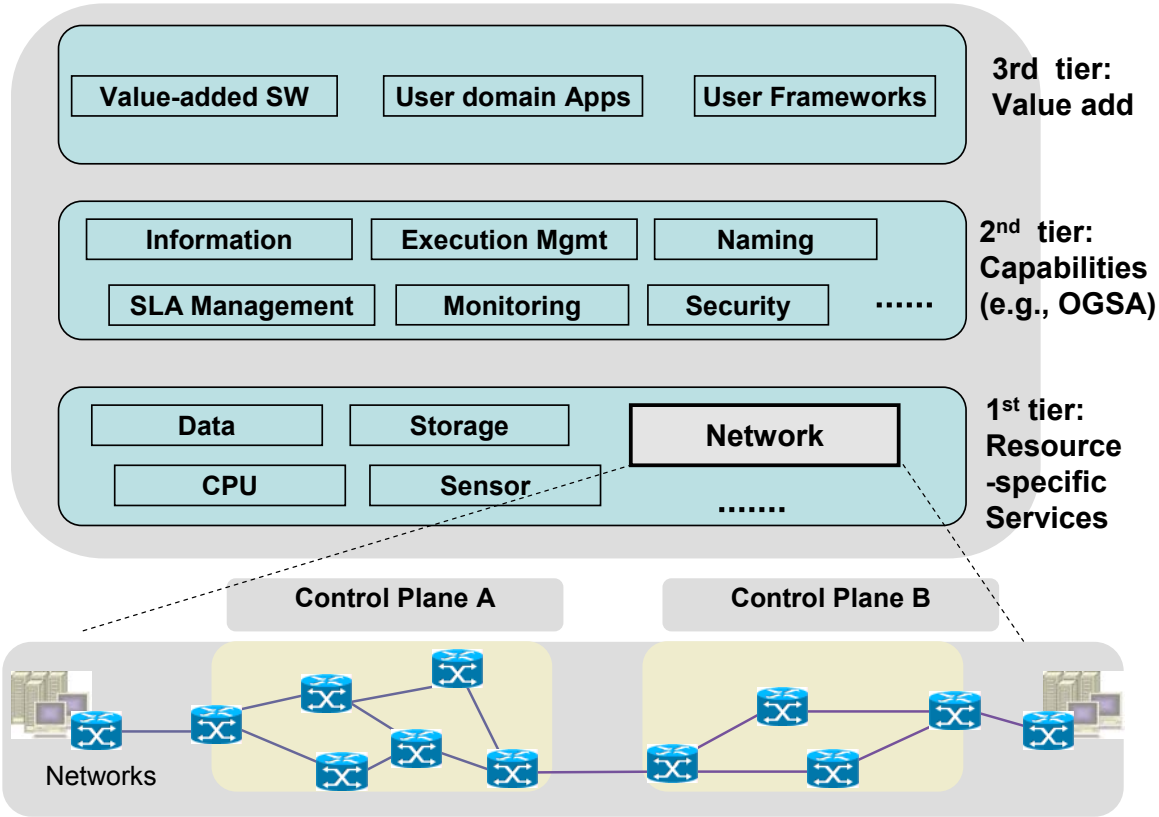
- DARPA's Coronet (2010 - …) with worldwide reach
  - Traffic breakdown: 75% IP, 25% wavelength svcs
  - 100 Tb/s aggregate demand
  - 100 Gb/s bit rate per wavelength
  - Setup and restoration time as low as 50+(2*RTT) msec
  - Commercial carriers to build it and operate it

Cost and performance arguments indicate that hybrid solutions have a sustainable edge & are future-proof to new apps

Bringing it all together 1 / 2
How does the Grid engage the network

> The network as a 1st-class Grid-managed resource akin to CPU and Data

> Can possibly be scheduled by a Grid community scheduler



Grid Network Services

sampler

- UCLP
- DRAC
- VIOLA
- NARB
- NRMS
- ...

In OGF, the Infrastructure Area advances this vision and the Grid/Network seams

Bringing it all together 2 / 2
OGF groups that worked / work on network themes

> Grid High Perf Networking RG
- grid network services architecture and interoperability

> Data Transport RG
- transport protocols: evolutions and new ideas

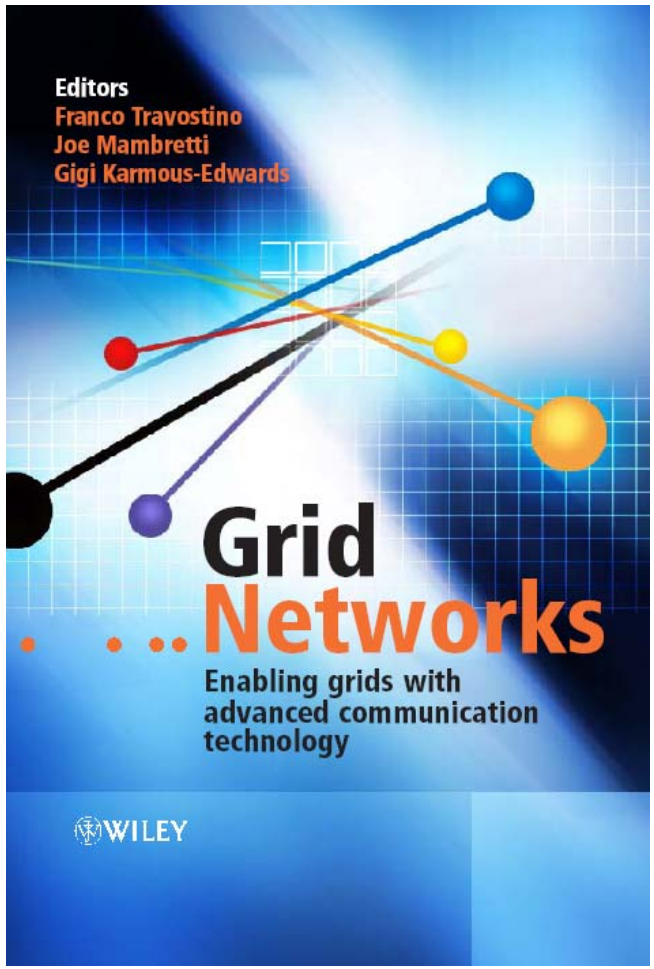
> IPv6 WG
- to weed out gratuitous IPv4 dependencies from OGF specs

> Network Monitoring WG
- standardizes ways to publish network metrics to the Grid

> GridFTP WG
- standardizes extensions to FTP

In OGF, the Infrastructure Area advances this vision and the Grid/Network seams

Additional Info



Available Fall 2006