Grids: an Overview of Vision, Standards, and Early Field Experimentation

Franco Travostino

travos@nortelnetworks.com

Director, Advanced Technology, Nortel Networks
Co-chair of the Hi-Perf Networking Group, Global Grid Forum

Outline

- 1. What's a Grid?
- 2. What's the GGF?
- 3. "Net-heads" anywhere in the GGF?
- 4. What's the GHPN about?
- 5. A Mainstream Use Case for the GHPN: eScience
- 6. Invitation to 2 GHPN Drafts

[The slides in this Unit reflect my own insights and material found in literature (with provenance shown whenever possible). They do NOT carry any official endorsement by the GGF]

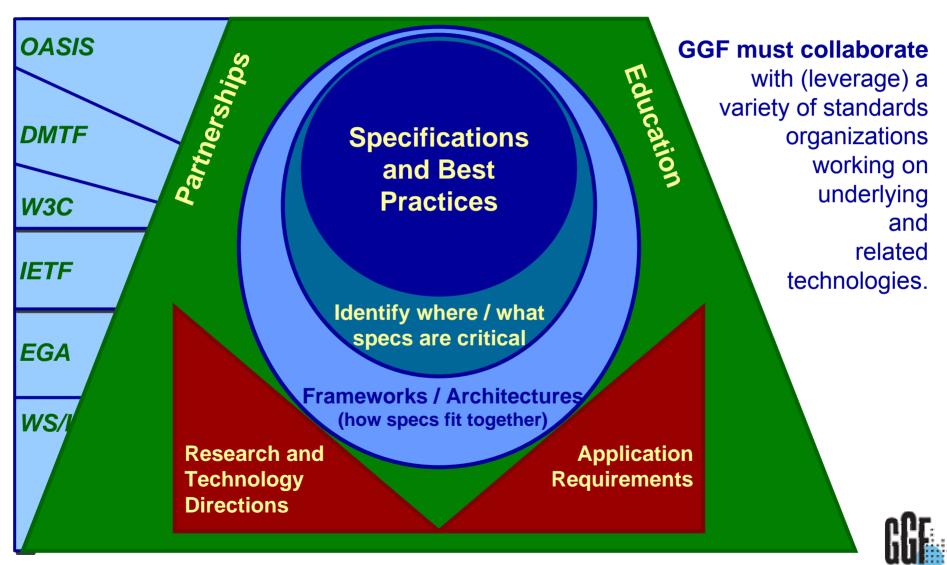
Outline

- 1. What's a Grid?
- 2. What's the GGF?
- 3. "Net-heads" anywhere in the GGF?
- 4. What's the GHPN about?
- 5. A Mainstream Use Case for the GHPN: eScience
- 6. Invitation to 2 GHPN Drafts

What is GGF?

- An Open Process for Development of Agreements and Specifications
 - Grid "Recommendations" process
 - modeled after Internet Standards Process (IETF, RFC 2026)
 - Persistent, Reviewed Document Series
 - similar to RFC series; initiated October 2001
 - Intellectual Property policies (RAND) modeled after IETF policies
- A Forum for Information Exchange and Collaboration
 - Comparing notes on experiences, approaches
 - Research Groups, Workshops, Online Communities
 - Spanning vertical layers, including people who are
 - Doing Grid Research
 - Designing and Building Grid software
 - Deploying Grids
 - Using Grids
 - Spanning technology areas- from directory services to scheduling to data handling to security to collaboration environments to...

Adoption and Relevance



Liaisons (as of October 2004)

- Internet Engineering Task Force (IETF)
 - IETF: Brian Carpenter (IBM)
 - GGF: Cees de Laat (Univ Amsterdam), David Martin (IBM)
- Distributed Management Task Force (DMTF)
 - DMTF: Winston Bumpus (Dell, DMTF President)
 - GGF: John Tollefsrud (Sun)
- OASIS
 - OASIS: Karl Best (OASIS VP), James Bryce Clark (Mgr, Tech Stds)
 - GGF: Dave Snelling (Fujitsu), John Tollefsrud (Sun)
- WS/I
 - WS/I: Michael Weiner (IBM, WS/I President)
 - GGF: Dane Skow (FNAL)
- W3C
 - WC3: Steve Bratt (COO)
 - GGF: Mark Linesch (GGF Chair), Steve Crumb (GGF)
- EGA
 - EGA: Don Deutsch (Oracle, EGA President
 - GGF: Mark Linesch (GGF Chair)



Organized into Seven Areas

- Applications and Programming Environments
 - Satoshi Matsuoka, Craig Lee
- Architecture
 - Andrew Grimshaw, David Snelling
- Data
 - Peter Clarke, David Martin
- Information Systems and Performance
 - Geoffrey Fox, John Tollefsrud
- Peer-to-Peer: Desktop Grids
 - Cees DeLaat, David DeRoure
- Scheduling and Resource Management
 - Bill Nitzberg, Stephen Pickles
- Security
 - Olle Mulmo, Dane Skow



57 GGF Groups as of October 2004

Applications and Programming Models Environments (APME) - Grid Checkpoint Recovery (GridCPR-WG) • Grid Remote Procedure Call (GridRPC-WG) • Advanced Collaborative Environments (ACE-RG) • Applications and Test Beds (APPS-RG) Astronomical Grid Community (Astro-RG) • Grid Computing Environments (GCE-RG) • Grid User Services (GUS-RG) • Life Sciences Grid (LSG-RG) Particle and Nuclear Physics Applications (PNPA-RG) Preservation Environments (PE-RG) • Production Grid Management (PGM-RG) Simple API for Grid Applications (SAGA-RG) • User Program Development Tools for the Grid (UPDT-RG)

<u>Architecture (ARCH)</u> - New Productivity Initiative (NPI-WG)• Open Grid Service Common Management Model (CMM-WG)• Open Grid Services Architecture (OGSA-WG) • Open Grid Services Infrastructure (OGSI-WG)• Grid Policy Architecture (Policy-RG)• Grid Protocol Architecture (GPA-RG)• Semantic Grid (SEM-RG)• Service Management Frameworks (SMF-RG)

- •<u>Data (DATA)</u> Data Access and Integration Services (DAIS-WG)• Data Format Description Language (DFDL-WG) GridFTP-WG IPv6 (IPv6-WG) Grid File System (GFS-WG) Grid Storage Management (GSM-WG) Information Dissemination (INFOD-WG)
- OGSA Data Replication Services (OREP-WG) Data Transport (DT-RG)• Grid High-Performance Networking (GHPN-RG)• Persistent Archives (PA-RG) Transaction Management (TM-RG)

<u>Grid Security (GRID SEC)</u> - Authorization Frameworks and Mechanisms (AuthZ-WG)• CA Ops (CAOPs-WG)• Grid Security Infrastructure (GSI-WG)• Open Grid Service Architecture Authorization (OGSA AUTHZ-WG)• Open Grid Service Architecture Security (OGSA-SEC-WG)• Authority Recognition (ARRG-RG)• Site Authentication, Authorization, and Accounting Requirements (SAAA-RG)

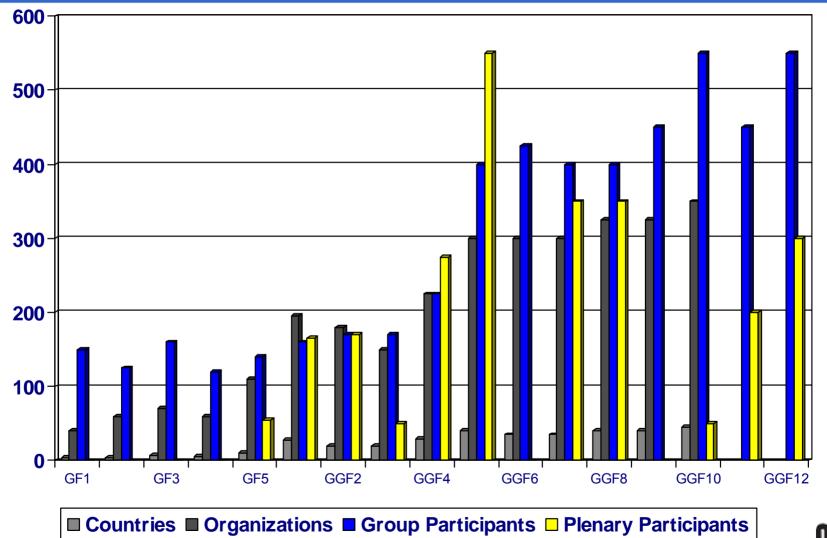
<u>Information Systems and Performance (ISP)</u> - CIM based Grid Schema (CGS-WG)• Discovery and Monitoring Event Description (DAMED-WG)• Network Measurement (NM-WG)• Grid Information Retrieval (GIR-WG)• Grid Benchmarking (GB-RG)• Network Measurements for Applications (NMA-RG) Relational Grid Information Services (RGIS-RG)

Peer-to-Peer (P2P) - Appliance Aggregation (APPAGG-RG) OGSA-P2P-Security (OGSAP2P-RG)

<u>Scheduling and Resource Management (SRM)</u> – Configuration Description, Deployment, and Lifestyle Management (CDDLM-WG) Distributed Resource Management Application API (DRMAA-WG)• Grid Economic Services Architecture (GESA-WG)• Grid Resource Allocation Agreement Protocol (GRAAP-WG)• Job Submission Description Language (JSDL-WG)• OGSA Resource Usage Service (RUS-WG)• Usage Record (UR-WG) Grid Scheduling Architecture (GSA-RG) Workflow Management (WFM-RG)

Working Groups: Developing specifications, API's, frameworks **Research Groups**: Developing Use Cases, Surveys, Technical reports...

GGF Meeting Participation







GGF Participation Demographics

- General Participants (circa early 2004)
 - 60% from universities and Federal laboratories
 - 40% from (over 100 companies)
- Leadership
 - ~100 working group and research group chairs
 - ~30 each from {.com, .edu, .gov } (I.e. 1/3 from commercial)
 - 21 steering group members
 - 7 from industry (IBM, Sun, Fujitsu, The Aerospace Corp, Avaki, Altair) (I.e. 1/3 from commercial)
 - 18 advisory committee members
 - 6 from industry (Intel, HP, Sun, IBM, Boeing, Platform)
- Example: OGSA, DAIS and OGSI Working Groups
 - Chairs from Fujitsu, Oracle, IBM, Univ of Manchester, Univ of Chicago, Argonne National Laboratory
 - 60+ companies, 50+ labs and universities involved



GGF Document Series

- Modeled after RFC Series (Internet Standards Process)
 - Minor differences from RFC 2026, with benefit of excellent advice from IETF participants (Bradner, Carpenter, etc.)
- Types of Documents
 - Informational
 - Workshop reports, general technical information
 - Experimental
 - Rapid prototype of specifications, API's
 - Community Practice
 - Generally accepted practice (process or technical)
 - Recommendations Track
 - Technical specifications, API's



30 Documents as of October 2004

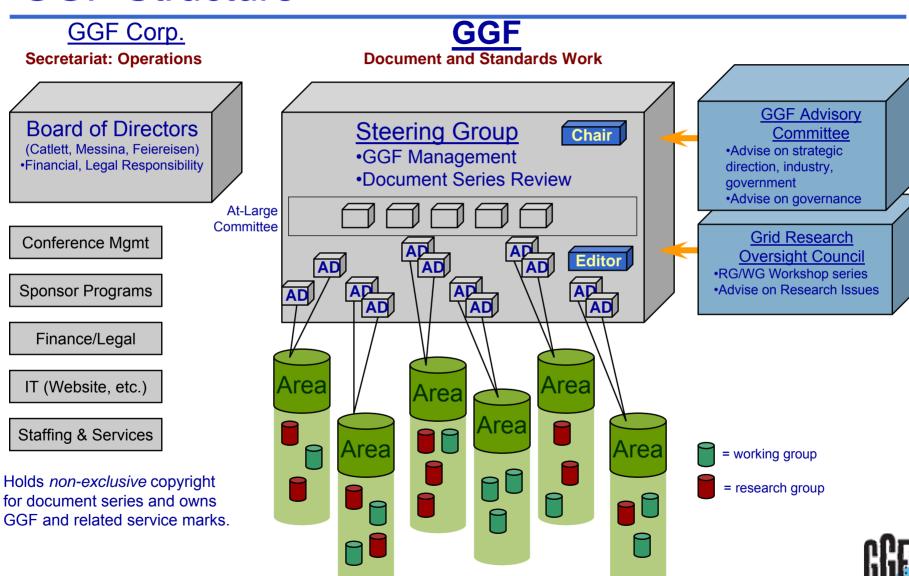
GFD.1	GGF Document Series	GFD.15	Open Grid Services Infrastructure V.1
GFD.2	GGF Structure	GFD.16	Global Grid Forum Certificate Policy Model
GFD.3	GGF Management	GFD.17	CA-based Trust Issues for Grid Authentication and Identity Delegation
GFD.4	Ten Actions When Superscheduling		
GFD.5	Advanced Reservation API	GFD.18	An Analysis of the UNICORE Security Model
GFD.6	Attributes for Communication between Scheduling Instances	GFD.19	Job Description for GGF Steering Group Members (GFSG)
GFD.7	A Grid Monitoring Architecture	050.00	GridFTP: Protocol Extensions to FTP for the Grid
GFD.8	A Simple Case Study of a Grid Performance System	GFD.20	
GFD.9	Overview of Grid Computing Environments	GFD.21	GridFTP Protocol Improvements
GFD.10	Grid User Services Common Practices	GFD.22	Distributed Resource Management
GFD.11	Grid Scheduling Dictionary of Terms and Keywords	GFD.23	Application API A Hierarchy of Network Performance Characteristics for Crid Applications
GFD.12	Security Implications of Typical Grid Computing Usage Scenarios	GFD.24	Characteristics for Grid Applications GSS-API Extensions
GFD.13	Grid Database Access and Integration: Requirements and Functionalities	GFD.25	An Analysis of "Top N" Event Descriptions
GFD.14	Services for Data Access and Data Processing on Grids	GFD.26	Persistent Archive Concepts
		GFD.27	Grid Information Retrieval Concepts

30 Documents as of October 2004 (cont)

GFD.28	Job Submission Information Model
GFD.29	Open Grid Services Architecture Use Cases
GFD.30	Reserved for OGSA v.1
GFD.31	Open Grid Services Infrastructure Primer



GGF Structure



www.qqf.orq

GGF Steering Group (GFSG)

- GGF Steering Group
 - Mark Linesch (HP) [Chair]
 - John Tollefsrud (Sun)
 - Peter Clarke (UCL/UK)
 - David Martin (IBM)
 - Cees De Laat (UVA/NL)
 - Dane Skow (FNAL)
 - Craig Lee (The Aerospace Corp.)
 - Bill Nitzberg (Altair)
 - Andrew Grimshaw (Avaki and UVa)
 - David Snelling (Fujitsu)
 - Sataoshi Matsouka, (Tokyo Inst. Of Tech)
 - David DeRoure (U of Southhampton)
 - Ole Mulmo (KTH)
 - Stephen Pickles (U of Manchester)
 - Geoffrey Fox (IU)

- "At-Large" GFSG Subcommittee
 - Dennis Gannon (IU)
 - Alan Blatecky (SDSC)
 - Malcolm Atkinson, (UK eScience)
 - Ken Klingenstein (Internet2)
 - Hiro Kishimoto (Fujitsu)
 - Charlie Catlett (ANL, UofC)
- GFSG Structure
 - Two area directors per area
 - At-Large Subcommittee
 - Appeals Process
 - Oversee liaisons with Other Groups
 - IETF, Internet2, W3C, DMTF, OASIS, IPv6 Forum

- GFSG Role
 - Operational Management and policy
 - Document series review
 - · Chartering of new groups
 - · Group oversight and review

- GFSG Participation
 - Individuals, not representatives
 - Selected by Nomcom



Grid Research Oversight Council (GROC)

Co-Chairs

- Satoshi Matsuoka (Tokyo Inst. of Technology)
- Dennis Gannon (Indiana University)

GROC Members

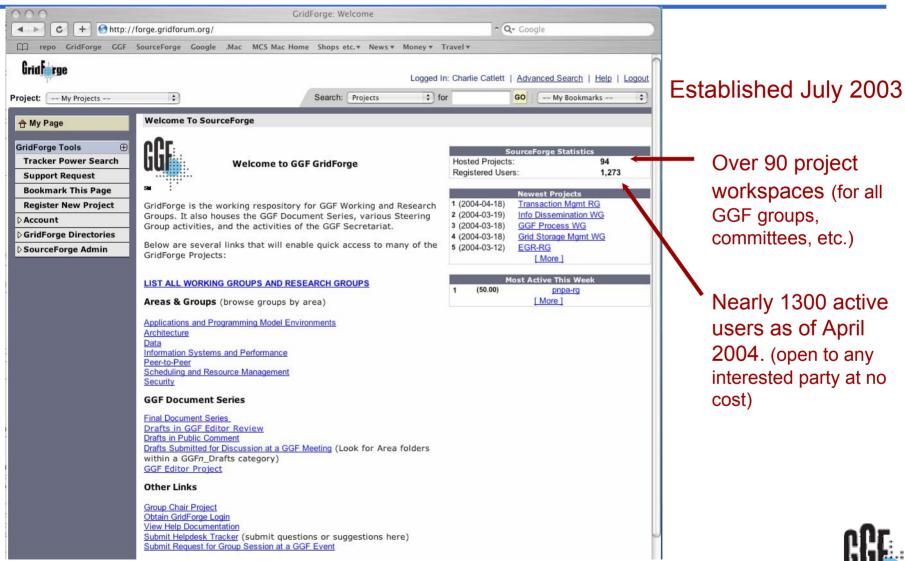
- Satoshi Sekiguchi
- Domenico Laforenza
- Malcolm Atkinson
- Fran Berman
- Henri Bal
- David Hutchinson
- Jack Dongarra
- Alexander Reinefeld
- Carl Kesselman

Charter

- Create and manage a highquality workshop series in conjunction with GGF tri-annual meetings.
 - Proposals accepted exclusively from GGF working and research groups
 - Solicited topics (from GROC, GFSG, GFAC, GGF Chair)
 - Unsolicited topics (from group chairs)
 - Arrange for publication of workshop proceedings as appropriate.
- Provide input to GFSG on approval of new Research Groups
- Provide peer-review on selected GGF documents



GridForge Community Portal



Outline

- 1. What's a Grid?
- 2. What's the GGF?
- 3. "Net-heads" anywhere in the GGF?
- 4. What's the GHPN about?
- 5. A Mainstream Use Case for the GHPN: eScience
- 6. Invitation to 2 GHPN Drafts

Groups with primary net focus

- Grid High Perf Networking RG
 - level set on Grid/Net pain points; grids on Optical Nets and R&E networks; network service architecture(s)
- Data Transport RG
 - transport protocols: evolutions and new ideas
- IPv6 WG
 - · to weed out IPv4 dependencies ahead of IPv6 roll-out
- Network Monitoring WG
 - standardizes ways to publish network metrics to the Grid
- Liaison functions into IETF, DMTF, ...
 - · Cross RG/WG scope; liaisons are reciprocated



Having close encounters with the net

- CIM-based Grid Schema WG
 - · Defines CIM-schemas not yet tackled by DMTF
- OGSA WG
 - On CIM still
- Grid Resource Allocation Agreement WG
 - Standardizes a common resource management protocol, inclusive of advance reservation
- Appliance Aggregation Architecture RG
 - Aggregation of client-side resources (appliances), and their interfacing to the Grid



Special Mention: Security !!!

- Everybody's problem
 - Challenge to Grid-heads and Net-heads @GGF
 - All Documents must have a security section



Outline

- 1. What's a Grid?
- 2. What's the GGF?
- 3. "Net-heads" anywhere in the GGF?
- 4. What's the GHPN about?
- 5. A Mainstream Use Case for the GHPN: eScience
- 6. Invitation to 2 GHPN Drafts

GHPN-RG

Chairs:

Jon Crowcroft, University of Cambridge, UK Jon.Crowcroft@cl.cam.ac.uk

Franco Travostino, Nortel Labs, Boston USA travos@nortelnetworks.com

Secretaries:

Volker Sander Forschungszentrum Jülich, Germany <u>v.sander@fz-juelich.de</u>

Nagi Rao, ORNL, USA raons@ornl.gov



GHPN-RG?

- To bridge gaps between Networking and Grid advanced communities
 - E.g., Grid requirements not yet understood by 'Netheads'
 - E.g., Net rules/features ignored by 'Gridheads'
- while countering any premature ossification around Grid/Net dominant scenarios
 - E.g., Advocacy of Grids using non-traditional mixes of end-systems and net infrastructures
- Outcomes: Informational or BCP Drafts, new RGs/WGs, liaisons with other Groups

GHPN Charter—the long version

The Grid High-Performance Networking Research Group focuses on the

relationship between network research and Grid application and infrastructure development. The objective of GHPNRG is to bridge the gap between the networking and grid research communities. It accomplishes its goal by serving as a forum for information exchange on advances and requirements in both fields, as well as by providing a focal point for liaison activities between the GGF and the various networking standards bodies. Specific topics of interest include, but are not limited to:

- End-to-end performance
- High-performance transport protocols
- Emerging network technologies
- The interface between Grid applications and network services
- Deployment of new technologies on the Internet and overlay networks



GHPN Charter—the long version 2

- The GHPNRG provides a forum for such topics until sufficient maturity and interest to both communities is reached that naturally results in the formation of a separate WG or RG to pursue them further. The discussions of GHPN are carried out both during the meetings and on the GHPN mailing list.
- Two specific goals of the GHPN-RG are identifying:
 - grid application requirements and implementations that are not supported or understood by the networking community and
 - advanced networking features that are not being utilized by grid applications.



"High-Performance" Networking RG?

It does NOT necessarily imply:

Gigabit/s or Terabit/s !!!

- It's short for: "Highly effective OR efficient networking, AND qualitatively different from established practice (if any)"
 - effective, adj: producing a decided, decisive, or desired effect
 - efficient, adj: productive of desired effects, especially: productive without waste



These can be in GHPN's scope

- · Grids with wireless systems
 - · Traits: Mobility, resource constrains, etc.
- Grids exploiting Optical Networks
 - Traits: Circuits, scheduled connection services, diversity routing
- Grids encompassing sensor-nets
 - CERN LHC and Berkeley's Motes are two kinds of sensors



But:

- No tutorials, primers, 101
 - Suggestion: Use glossary + references to level set quickly
- · No "miniature-GGF" in a Draft
- Bumpy & tortuous road is somewhat expected due to material's novelty
- OGSA must be light at the end of tunnel
 - If there is no intercept with present/planned GGF work, with any WG/RGs, then there is no value in pursuing it here further
 - · E.g., a conference, a journal may then be the proper outlets



A bit about our constituency

Active GHPN participants come from

Research

 ANL, University of Essex, CANARIE, University of Tennessee, Universiteit Van Amsterdam, University of Technology Sydney, University of Illinois at Chicago, Northwestern University, Forschungszentrum Jülich GmbH, Ecole Normale Superieure Lyon, Univ. of Cambridge, Boston University, University of Florida, University of Lecce

Industry Labs and industry

 Alcatel CIT Research, AT&T, Cisco, MCNC Institute, Nortel Networks



2 Drafts in PC

- draft-ggf-ghpn-opticalnets-2
 - · Convergence of optical nets and Grids
 - · Greenfield-charting realm
 - · 30 days PC finished on 10/9/04, soon to become a GFD
- · draft-ggf-ghpn-netissues-4
 - · Pain points in attaching Grids to the network
 - · Level-set realm
 - 30 days PC finishes on 11/08/04



3 More Drafts in the making

- draft-ggf-ghpn-transportsurvey-1
 - · A survey of L4 protocols for bulk data transfer other than TCP
- draft-ggf-ghpn-netservices-usecases-2-3
 - Black-box use cases of Grids setups exploiting network behaviors, and quantitative analysis of the same
- · draft-ggf-ghpn-netservices-1
 - Scoping of Grid Network Services that elevate the network resource to being a Grid Managed Resource akin to processing and storage



Outline

- 1. What's a Grid?
- 2. What's the GGF?
- 3. "Net-heads" anywhere in the GGF?
- 4. What's the GHPN about?
- 5. A Dominant Use Case for the GHPN: eScience
- 6. Invitation to 2 GHPN Drafts

[The slides in this Unit reflect my own insights and material found in literature (with provenance shown whenever possible). They do NOT carry any official endorsement by the GGF/GHPN]

Outline

- 1. What's a Grid?
- 2. What's the GGF?
- 3. "Net-heads" anywhere in the GGF?
- 4. What's the GHPN about?
- 5. A Dominant Use Case for the GHPN: eScience
- 6. Invitation to 2 GHPN Drafts

Opticalnets Draft-1

Focuses on deployment of photonic network infrastructure for Grid applications

Motivation:

- Predictions for deployment of data-intensive Grid applications that will require transfers of Terabytes or even Petabytes of data
- These applications will require a high bandwidth network environment where bandwidth will be allocated on demand or by user/application driven scheduled reservation

This draft aims:

- To suggest solutions towards an efficient and intelligent network infrastructure for Grid taking advantage of recent developments in optical technologies
- A solution to support high-demand application with potential to support all types of Grid application

Opticalnets Draft-2

Grid applications

- · Mainly for data-intensive and/or long-lived applications
- No widespread applications requiring Grid over Optics
- Today's applications with large BW requirements (i.e. high energy physic centers, radiotelescopes) belong to well defined communities of users and destinations with typically long lived persistent relationships
- Additionally, sampled users mentioned emerging application scenarios such as high bandwidth interactive applications, data visualisation applications, applications that require bandwidth to reduce latency



Opticalnets Draft-3

Styles of operation

- Users have IRUs to dark fiber and control the network
- Users access OVPN, L1VPN services from providers operating commercial DWDM systems
- In either case, software layers abstract the network as a Grid resource, within the OGSA architecture

Transport and switching technology

- · Wavelength switching
- Hybrid IP router/wavelength switching
- Optical burst switching



Opticalnets Draft-4

- Testbeds show two Control Plane approaches
 - A fixed optical mesh between users with slow "automated fiber patch panel" switching (OBGP)
 - A shared optical "cloud" with rapid λ switching between users (GMPLS, ASON, OBS)
 - Grid User Network Interface
 - GUNI is seen building upon and extending OIF UNI
 - Scheduled connectivity



Netissues -> TOC

- Scope and Background
- Use Case Documents
- High-Throughput
- Performance controllability
- Dynamic Network resource allocation and reservation
- High Availability
- Security Controllability
- Multicast
- Sensor Networks and Wireless
- Grid Traffic



Netissues -> Use Case Documents

The network must provide:

- High performance transport for bulk data transfer (over 1Gb/s per flow)
- Performance controllability to provide ad hoc quality of service and traffic isolation.
- Dynamic Network resource allocation and reservation
- Security controllability to provide a trusty and efficient communication environment when required
- High availability when expensive computing or visualization resources have been reserved
- Multicast to efficiently distribute data to group of resources.

Other issues have been pointed out by the networking community:

- What is the grid traffic impact on infrastructures and other traffics
- How to integrate wireless network and sensor networks in Grid environment



Netissues -> High-Throughput

Requirements:	 High average throughput Advanced protocol capabilities available and usable the end-systems Lack of use of QoS parameters 			
Current issues	1)Low average throughput 2)semantical gap between socket buffer interface and protocol capabilities of TCP	the		
Analyzed reasons	<pre>1a) End system bottleneck, 1b) Protocol misconfigured, 1c) Inefficient Protocol 1d) Mixing of congestion control and error recovery 2a) TCP connection Set up: Blocking operations vs asynchronous 2b)Window scale option not accessible through API</pre>			
Available solutions	1a) Multiple TCP sessions 1b) Larger MTU 1c) ECN			
Proposed alternatives:	1)Alternatives to TCP (see DT-RG survey document) 2)OS by-pass and protocol off-loading 3)Overlays 4)End to end optical paths	oOr.		

Netissues -> Performance Controllability

Requirement:	1) Traffic protection 2) QoS-aware networking infrastructure
Current issue	1) API: Form of SLA with measurable parameters constituting a SLS
Available solutions	1) Overprovisionning 2) DiffServ 3) MPLS-TE with DiffServ Scheduling
Proposed alternatives:	BoD, lambdaoD Overlays



Netissues -> Dynamic Network Resource Allocation and Reservation

Requirements:	1) Advance reservation capabilities
Current issue	OIF UNI 1.0 does not provide appropriate functionality yet
Analyzed reasons	
Available solutions	GFD-E.5 Advance Reservation API
Proposed alternatives:	



Netissues -> High Availability

Requirements:	1)Network reliability 2)Efficient Routing
Current issue	IP-Restoration QoS-Routing
Analyzed reasons	
Proposed alternatives:	MPLS-TE Multipath OSPF Overlays and P2P Ipv6



Netissues -> Security Controllability

Requirements:	1)Site and Network security 2)On demand security
Current issue	1) that there isn't a "one size fits all" site and network security solution 2) computing overhead, packet header overhead, high-availability, and policy 3) Firewalls/NATs and Grids
Analyzed reasons	
Available solutions	Middleboxes with L4-7 impact but lead to ossification around a L4 protocol called TCP VPN with use L2TP protocol in conjunction with the IPsec protocol and the Internet Key Exchange (IKE) protocol



Netissues -> Multicast

Requirements:	1)Reliability 2)Low recovery latency 3)Efficient congestion control
Current issue	Limited deployment No real standard no 'one solution fits all'
Analyzed reasons	
Available solutions	
Proposed alternatives:	End-system/end-host multicast Overlays, P2P



Netissues -> Sensor Networks and Wireless Technologies

Requirements:	Integration of wireless sensors networks
Current issue	power limited very dynamic topology unreliable communications IP overhead TCP limits over wireless networks
Available solutions	
Proposed alternatives	



Netissues -> Grid Traffic

Requirements:	Understand the nature of the Grid traffic and how it may impact the network and the protocols
Current issue	Internet Traffic is self similar, what's about Grid traffic? Traffic Phase Effects Flash Crowds
Available solutions	Network monitoring
Proposed alternatives	



GGF

Path-oriented Use Cases

SII	Banwidth	Latency	Jitter	Packet Loss	Mcast	Reliability	Co-Allocation	Data security
Visuali- zation	. Compute/data intensive . distributed data-sets: TBy/PBy . Depending on image resolution: 68/680 Mb/s; 6.8/68 Gb/s	< 200 ms	< 15 ms (if lat= 10 ms)	Yes	Yes	Yes TCP/UDP with forward error correction	Computing/ Storage/ network	No
HEP file replica	Several PBy/year .Tier-0 to Tier-1 sites (raw data, 2 PBy) -re-processed data, from Tier-1 to other Tier-1 - Local copies	No(but dependenc y of TCP on RTT)	No	Yes	Yes	Yes	Computing/ Storage/ Bandwidth, loss	No
EMT	Small data-sets from wireless sensors to PDAs and vice- versa; network capacity used depends on sensor sampling rate	< 1 s	No	0 – 20%	No	Yes	Storage (disconnected operation); bw and latency	Yes

www.ggf.org

Email über alles

Drafts in the making, new Drafts, comments, ideas, why and why-nots, technical news ...

Home page: http://forge.gridforum.org/projects/ghpn-rg/

