

# Grid High-Performance Networking Research Group

## draft-ggf-ghpn-netissues-1

Volker Sander (Editor), Forschungszentrum Jülich

William Allcock, ANL

Pham CongDuc, Ecole Normale Supérieure Lyon

Jon Crowcroft, University of Cambridge

Doan B. Hoang, University of Technology, Sydney

Mark Gaynor, Boston University

Inder Monga, Nortel Networks Labs

Pradeep Padala, University of Florida

Marco Tana, University of Lecce

Franco Travostino, Nortel Networks Labs



# Motivation

- **GHPN focuses on the relationship between network research and Grid application and infrastructure development**
  - Document by Jon Crowcroft: “Top Ten Things Network Engineers wish Grid Programmers knew”
    - Overview of ongoing networking efforts of interest for the Grid community
  - Document draft-ggf-ghpn-netissues-0
    - The ambitious nature of Grids stress the capabilities of the networking infrastructure
    - List experienced problems and activities
    - Identify open issues that should be addressed by a community effort



## Since GGF8 ...

- Merged the two documents into one reflecting the particular importance of the network for the Grid community
- New sections on
  - Sensors and sensor-nets
  - Traffic considerations, traffic engineering
  - Discussion on IPv6 pros/cons
- Issues raised @GGF8 and still requiring an action
  - Multiple virtual systems running on single host
  - Over-provisioning is about bandwidth only
  - Transactional traffic is not represented



# Document Outline

- **Introduction**
- **Scope and Background**
- **End-Systems**
- **IPv6**
- **Routing**
- **Access Domains**
- **Transport Domains**
- **General Issues**
- **Macroscopic Traffic and System Considerations**
- **Traffic Engineering**
- **Security Considerations**



# Sensors and SensorNets

- Allow integration of real-time data into grid applications
- Some sensors will be complex, and wired directly into the grid, using standard grid protocols
- Other sensors will be embedded devices without access to wired power, and wired Internet service
- For the latter, today application/sensor binding typically occurs in an ad-hoc fashion
- Plenty of wireless technologies out:
  - WiFi, Cellular, Bluetooth, passive (i.e. Rfid tags) free space lasers, etc
- Power management to extend battery life is a prime consideration
- IP at the network layer is often too inefficient
  - 20 bytes of the IP header requires too much power and too much bandwidth
  - non-IP networking layer is a major complication
- Routing within sensor nets is hard
  - limited power
  - dynamic nature
  - non-unique addressing across sensor realms



# Macroscopic Traffic and System Considerations / TE

- Self similarity does not really matter
- Traffic phase effects do matter
  - Flash crowds
- Beware of asymmetry
  - e.g., ADSL
  - Client-Server
  - Master-Slave
  - NAT-Boxes
  - BGP-Paths
- Traffic Engineering
  - The existing Interior Gateway Protocols (IGP) are not adequate for TE
  - Overlay could build virtual topologies on top of the physical networks
  - MPLS could be one of this overlay model



# IPv6

- **Initially designed to solve the problem of a narrow address space**
- **However it has several other benefits:**
  - Facilitates the allocation for multicast, and re-allocate for mobile, which makes dynamic addressing and group communication much more viable
  - Obviates the need for NATs too which improves the chances of pure end-to-end connectivity without extra NAT-traversal stages
  - Features a flow identifier field, which could be used to speed-up the identification (and possible grouping) of packets into flows (and aggregates) for special purpose forwarding treatment
  - Operationally, the zero-knowledge configurability of IPv6 is potentially very useful in large site management.
- **The big missing piece is a stable router deployment in cores. This does not obstruct IPv6s usefulness in the edge (e.g. wireless access), but does undermine the use of flow id for forward performance where it matters (core) or multicast.**
- **IPv6 WG has been formed, that's where these issues belong**



# Open Discussion

- **What else is missing?**
- **Can we derive a set of actions to be performed by the Grid community?**
- **Is there any explicit relation to the optical draft?**

# Backup



# End Systems

- **Communication Protocols and their Implementations**
  - Problems caused by limitations of existing APIs
  - More API-related efforts needed
- **Operating System Capabilities and Configuration Issues**
  - Service impact caused by tuning parameters
  - Tuning guides (and tools) needed
- **OS and System-Level optimizations**
  - Lists potential solutions to improve the performance
    - OS-bypass, Super-NIC,...
  - Bridge to Jon's document
- **Multi-Stream File Transfers**
  - Quantitative analysis on what is achievable
  - The Grid wants multi-stream file transfers



# Routing

- **Fast forwarding is there**
  - Clusters can do 10Gbps
  - Firewalls can do better than they currently do
- **Faster Convergence**
  - Current Convergence of routing protocols is not convenient
- **Theory and practice**
  - it takes a good programmer about 3 months to implement OSPF
  - It then takes around 3 years to put in all the defenses
- **Better (multi-path, multi-metric) routing**
  - Multipath in limited cases appears to work quite well
  - Multimetric relies on good understanding of traffic engineering and economics, and to date, hasn't seen the light of day
- **MPLS**
  - Pros: Protection capabilities, CoS-support
  - Cons: Multicast not really supported
- **BPG**
  - Policies are hard!
    - Influence what class of packets gets routed where to
    - Getting the global view is tough



# Access Domains

- **Firewalls**

- Do have an impact on the level of service
- Jon's document states that firewalls can do better
- Connectivity
- Do we need a "grid-ftp"-aware firewall

- **Network Address Translators**

- Problems caused by callbacks (exposing the IP-address)
- More work input here

- **Middleboxes with L4-7 impact**

- Firewalls, intrusion detectors, SSL accelerators, Traffic Shaping
- Resistance to use other L4-7 protocol: TCP remains
- How to control (discover and signal) those Boxes?

- **VPNs**

- Convenient, but not integrated into Grid Security
- Error handling problems due to isolation of VPN console
- GSS-API style interface needed?



# Transport Domains

- **Service Level Agreements**
  - **Grids and SLS**
    - Knowledge about the expectable capabilities important for Grid resource allocation strategies
    - End-to-End QoS
  - **SLS assurance**
    - Monitoring network SLS parameters by applications or RM
    - Relates to general service assurance efforts in Grids
  - **On-demand SLS**
- **Overprovisioned Networks**
  - Nice to have, but are they deployed end-to-end in the foreseen future?
  - Is the onion model still valid (10GigE)?
  - Meshing of multiple fibers might come up
  - Relates to Jon's document



# General Issues

- **Service Orientation and Specification**
  - The Grid is about the composition of services
  - Required services by applications
  - Relation OIF's UNI
  - Advance Reservations
- **Programming Models**
  - GFD-E.5 (Advance Reservation API) lists uniform operations for interfacing to SLAs
- **Support for Overlay Structures**
  - Offer interesting features to improve effectiveness of VOs
  - Relation to Jon's document
- **Multicast**
  - Well-designed to suit the Grid requirements
  - Reliable multicast is the best candidate for providing an efficient multipoint communication support, but scalability is a challenge



# Security Considerations

- **Security Gateways**

- IPsec devices often enforced by policies
- No differentiation between application layer secured communication and unsecure communication
- GSI can be relaxed if it would be aware of the capabilities of the Security Gateways

- **Authentication and Authorization Issues**

- AAA servers are deployed for network services
- Required to be HA

- **Policy Issues**

- The standards for application-level security policies and network-level security policies are evolving independently, while the respective implementations may actually end up sharing infrastructure elements
- Will there ever be a closure with tools verifying adequacy, equivalence, and congruence of security policy selections