Network Quality of Service for GRID-enabled Applications and Middleware

Tiziana Ferrari
Istituto Nazionale di Fisica Nucleare – CNAF

Tiziana.Ferrari@cnaf.infn.it
Overview

- Grid-application and middleware categories
- Requirements and services
- QoS today
  - Status
  - Issues
- Future work
Grid-based applications and middleware

- Extremely **wide range of requirements** due to
  - the *variety* of potential GRID-aware applications which can take benefit from distributed computing
  - the **rich set of network-based services** supported by the Grid middleware, the Grid “engine”
    - Resource brokerage
    - Job scheduling
    - Data replication
    - Information service
    - Network-based optimization
    - Grid resource monitoring
    - Advance reservation (network, disk, cpu etc)
Broad application categories

1. Applications handling audio/video/image content
2. Short-lived, reliable data transactions
3. Bulk data transfer
4. Test applications
5. ...
Audio/video/image handling (1)

- Videoconferencing
- remote visualization
- real-time remote analysis of images
- tele-immersion

- Critical parameters:
  - Packet loss frequency
  - one-way delay
  - IPDV

→ Expedited Forwarding
Short-lived, reliable data transactions (2)

- Data-oriented applications accessing large amounts of small data portions (remote file analysis: HEP applications, earth observation, ...)
- remote hardware control
- client/server transactions in GRID middleware

- Critical parameters:
  - Number of completed transactions over time
  - RTT
  - packet loss and packet loss pattern

→ Active queue management: ECN, WRED
→ Assured Forwarding for minimum guaranteed bandwidth with allowed excess traffic to accommodate short rate peaks
Bulk data transfer (3)

- database replication for load balancing and job locality

- transfer of large data collections to one or more sites (see the multi tier data analysis hierarchy in HEP experiments)

- Critical parameters:
  - Guaranteed minimum bandwidth,
  - packet loss minimisation for improved performance at high-speed
Bulk data transfer (3) - cont

→ TCP stack tuning and optimization in case of link capacity under-utilization for increased rate aggressiveness and improved performance at high-speed
→ application adaptation
→ dynamic selection of single stream/parallel streams data transfer
→ active queue management, other ...
Bulk data transfer (3) - cont

Delivery of guarantees to TCP-based applications requires:

- very good understanding of performance existing high-speed network infrastructures:
  
  • Effect of *aggregation* of hundreds of TCP streams in a single traffic class (different depending on the aggregation degree)

  • effect of *systematic errors* like router and OS bugs on packet loss and consequently rate adaptation

  • Typical frequency of both short and long-term congestion
Bulk data transfer (3) - cont

• tolerance of traffic policers to *aggregate burstiness*
  (especially at very high speed:
  – Policing Token bucket depth configuration
  – tuning of queue lengths of schedulers for maximum
tolerance to burstiness and packet loss minimization
  – RED thresholds tuning
  – re-shaping: benefits (controlled burstiness) and penalties
    (additional queueing stages and consequent delay and
    buffer size tuning issues)

Note: TCP services not sufficiently understood and
scarcely addressed by large research network providers
Test applications (4)

- Any grid-based application and middleware component under alpha and beta test

- Critical parameters:
  - isolation between traffic classes in case of congestion or in presence of background legacy traffic

→ Less than Best-Effort Services
QoS today

• Good implementation and support of basic router and switch QoS functional blocks
  – QBone Premium service (Abilene)
  – IP Premium
    • GEANT, alpha-phase testing
    • Some European NRNs (not all of them for the moment)
  – Less than Best-Effort services: QBSS (Abilene)
QoS today (cont)

- Issues:
  - Inter-domain seamless QoS support, for consistent end-to-end services (dynamic SLA set-up, network dimensioning)
  - Control of user access to services (accounting, authorization and authentication is needed, no agreement on inter-domain policy management architectures)
  - Service monitoring (point-to-point and end-to-end)!
  - Feedback to applications of current service level (end-to-end QoS), application adaptation
  - Killing applications
  - Great overall complexity!
Areas for future work

• Network Quality of Service for GRIDs
  – What services (Per-Hop Behaviours) - not currently addressed by ISPs and IETF are particularly requested by GRID-enabled user applications and GRID middleware?

• Advance reservation: linking QoS functionalities and services with GRID middleware and user applications for a more dynamic use of network services
  – Technical issues to be addressed for the design and implementation of a Network Resource Broker in a true inter-domain environment
  – Differentiated service – *permanent configuration*, no dynamic network dimensioning
  – Other complementary approaches: *MPLS LSPs*, optical networking and *lambda switching* in a non-homogeneous inter-domain environment
  – *WEB services* model for service specification and management