

LBE/Scavenger

TF-NGN 8

Berlin, July 1st 2002

Overview

- LBE deliverable D9.9 for GEANT due Aug 31
- Need to decide how to press ahead with workplan; had some discussion already
 - Technical issues presented at TF-NGN 7 in April
 - Also held BoF at Limerick (TNC2002) in May
- Who is interested?
 - What are the application areas?
- What needs to be done on GEANT core?
 - To enable experiments between NRENs and end sites

Interest from...?

- To date we have interest from:
 - Arnes (Slovenia)
 - CESNET (Czech)
 - INFN-CNAF (Italy)
 - RedIris (Spain)
 - UKERNA/UoS (QoS pilots being formulated)
 - UNINETT (Norway)
- Are there others?

Original ideas

- From Tiziana's TF-NGN 7 presentations:
 - Work plan definition
 - Ongoing – some agreement at Limerick meeting
 - Review the Scavenger configuration
 - Done – though GEANT routers different
 - Co-existence with IP Premium
 - Not seen as a problem – GEANT spec underway
 - Identify test cases: e.g. GRID data?
 - Ongoing – candidate set discussed

Original work plan

- Phases:
 - 0: gather NREN requirements and define the service specification
 - 1: baseline tests and service feasibility
 - 2: performance analysis in user-based scenario (e.g. DataGRID) – can be transparent LBE on core
 - 3: implement user and NREN-based scenario, then trial on production networks
 - 4: joint US/Internet2/Abilene experiments (Scavenger)
 - Some revision after Limerick meeting...

Original (TCP) application areas

- Including but probably not limited to:
 - Mirroring (Uninett)
 - Usenet news and mail distribution (HEAnet/Surfnet)
 - Limiting point to point traffic not entitled to use NRN resources (RedIris)
 - Isolate production traffic from test bulk traffic, e.g. DataTAG (INFN-CNAF)
 - Non TCP-compliant programs (INFN-CNAF)
 - Student dormitory traffic (UoS)
 - Certain types of GRID applications (UoS)
 - Others?

Internet 2

- Scavenger running for over a year
 - Approx 2% of traffic is LBE-marked
- Abilene core is not congested
 - Core transports DSCP 8 (Scavenger code point) transparently (does not alter value)
 - Scavenger applied at edges where congestion occurs.
 - At least four universities using Scavenger for dormitory traffic
- Abilene migrating to Juniper Q3 2002
 - Using T-640 routers (more output queues than current GEANT routers)

GEANT status

- Packets currently retagged to DSCP 0 on core
 - Can be checked by NANOG traceroute
- DANTE seeking method for DSCP transparency
- GEANT-US link not heavily loaded
 - 20% as of May 2002, probably more now e.g. due to UKERNA's loss of transatlantic carrier.
- Some implementation examples for M-series
 - http://www.dante.net/staff/nicolas/Cos/pip_lbe_backbone.html
- (Thanks to Nicolas.Simar@dante.org.uk)

Checking DSCP propagation...

- Useful to be able to test where DSCP is being cleared if LBE DSCP not transported end to end
- Can use NANOG traceroute (latest version):
 - <ftp://ftp.login.com/pub/software/traceroute/beta>
 - (thanks Simon! 😊)
 - Run with `-t` option for TOS value.
 - Run from UoS shows TOS cleared at entry to GEANT core. Implies DSCP passes from UoS through MAN and JANET OK. Likewise for SWITCH.

Limerick BoF conclusions

- Four were noted:
 - Adopt a common DSCP (8).
 - Ensure DSCP is allowed to pass unaltered for end to end transparency
 - Let the NRENs and end users decide on implementation of LBE
 - Approach the IETF with suggestion to recommend global use of some DSCP values (may be tricky!)
- Minutes will be available from our LBE site

Some points to note...

- GEANT is not congested in the core
 - Usefulness currently in applying at the edges
- Packets can travel 90% to destination and then still be dropped.
 - But will still cause TCP backoff at source, easing congestion
- Server has to set DSCP, not client
 - Stig's FTP server patch shows this can be easy (one socket option argument).

Stig's FTP server patch

- Patch for troll-ftpd
- Command “site dscp value”
 - Works for *sverresborg.uninett.no*
 - Sets DSCP for data transfer (not control)
 - E.g. site dscp 32
 - NB. The server must be accessed in passive mode
- Very simple patch.

GEANT implementation issues

- Needs own hardware queue, GEANT Juniper routers have four hardware queues (BE, Premium IP, routing protocols, 1 free!)
 - DANTE seems happy to use all four queues on GEANT core to enable LBE.
 - Candidate specification drafted by Nicolas @ DANTE
- May be some problems in passing DSCP unaltered over whole GEANT core.
 - Requirement to alter/reset core routers?

Application enabling...

- Example of FTP server
 - Option for client to set DSCP
 - Equivalent of Unix “renice” on network
- Need to identify other applications
 - Any need to standardise DSCP options in applications?
 - If so, how?
- Note many FTP clients set DSCP too
 - Need to check and understand this – Stig suggested study of DSCP behaviour from applications...?

Do we need a testbed?

- Traffic generation/loading
 - Smartbits – appropriate? SmartTCP?
 - Other TCP tests?
- Measure the aggregate LBE throughput that applications can get out of a LBE queue with and without congestion.
 - Detail level of tests?
- If GEANT passes DSCP unaltered, we can run tests between NRENs/end sites...

Where next?

- LBE can be deployed incrementally:
 - Transparent transport of DSCP 001000
 - Implementation of queues
 - See Tiziana's TF-NGN 7 slides
 - Initially only at customer edge (requires transparent transport)
 - Implies no need for testbed?
 - Begin implementation in edge routers, where congestion is observed.
 - Low cost to deploy

Where next: applications?

- Are we happy with original scenarios and usages?
- Enabling LBE (Scavenger) on link between GEANT and Abilene.
 - Enable European-US end-to-end experiments, even if we don't know exactly what they are now.
- Is there a clear user requirement?
 - If not, can we encourage one?
 - With Scavenger, applications can be nice without looking at MRTG graphs or other sources to predict network load.

Other related areas to consider?

- Some areas to perhaps consider:
 - LBE through firewalls
 - Remote host signalling request for LBE
 - Any uses for LBE with UDP? (no backoff)
 - UDP usually used for media streaming?
 - Enforced LBE, e.g. for student dorms?
 - Application behaviour with LBE
 - Should be “predictable” with TCP
 - Min spec router to implement LBE drops

Other DiffServ services? ABE

- Alternative Best Effort (ABE)
 - Aimed at real-time applications like voice
 - We're not considering this yet
 - Some interest in Internet 2 community
 - Lack of implementations?
 - See: www.abeservice.com

Info

- Tiziana's LBE site:
 - <http://www.cnaf.infn.it/~ferrari/tfngn/lbe/>
- Scavenger:
 - <http://qbone.internet2.edu/qbss/>
 - Includes configuration examples
- Need input for August 31st deliverable!
 - Over to Tiziana for GEANT implementation...