Results of QoS trial tests in which UKERNA participated

JANET QoS Development Project Report

Version 1

Authors:

Victor Olifer, Rina Samani

With contributions from: Anthony Ryan, Michael Robson and Steve Williams

April 2005

Table of Contents

Introduction

This document describes the tests that were carried out by UKERNA in cooperation with Swansea University and the University of Manchester in March 2004 as a part of the JANET QoS Development Project trial. It complements the descriptions of the tests made by other JANET QoS Development Project partners: Lancaster University, The University of Manchester (with other than UKERNA partners), Southampton University and Imperial College, which can be found in the partners' final deliverables.[1] The document contains examples of the QoS configurations of the UKERNA site's routers and BARs (Backbone Access Routers) on the JANET backbone. The network diagrams of the University of Manchester and Swansea University testbeds also are given.

The tests across JANET were an important part of the Project as they aimed to verify the applicability and possible benefits of the DiffServ[2] (Differentiated Services) model for certain popular traffic types in the multi-domain JANET environment. Three kinds of application were chosen for testing: VoIP, videoconferencing and bulky GRID applications. The tests were conducted according to the Test Plan,[3] which includes a detailed description of the tests' scenarios, generic network diagrams, and measurement and monitoring tools.

The JANET Backbone and Participants' Sites Configuration

The JANET backbone

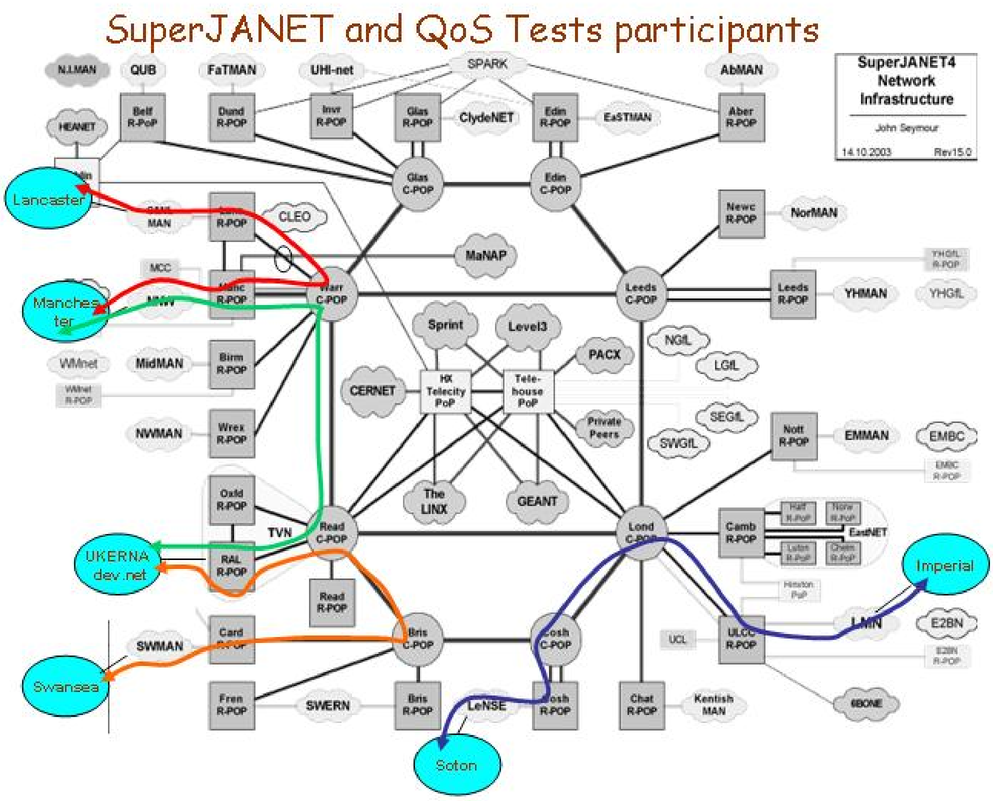

Figure1. The JANET backbone and participants' sites.

The JANET backbone (shown simplified in Figure1) was configured to support three QoS services: IP Premium (also known as Expedited Forwarding, EF), Best Effort (BE) and Less-than-Best Effort (LBE). The standard values of DSCPs (Differentiated Service Code Points) were used for marking traffic classes, so that IP Premium traffic was marked as DSCP 46, Best Effort traffic as DSCP 0, and LBE traffic as DSCP 8.

Participants were paired during each test so that there were bidirectional traffic flows between pairs of corresponding sites.

This document describes the tests between UKERNA and the University of Manchester (flows marked in green in Figure 1) and between UKERNA and Swansea University (flows marked in amber in Figure 1).

Although not all JANET backbone routers were involved in QoS testing, all of them – both core routers installed at C-PoPs (Core Points of Presence) and BARs installed at R-PoPs (Regional Points of Presence) – were configured to support QoS. However, only BARs serving project participants' sites allowed non-Best Effort IP packets (i.e. packets with non-zero DSCPs) to be transferred through the JANET backbone with unchanged DSCPs, and hence to be served as an appropriate QoS class. Those BARs were:

- RAL BAR;
- Manchester BAR;
- Lancaster BAR;
- Cardiff BAR;
- Cosham BAR;
- ULCC BAR.

All the other BARs re-marked IP packets coming in with non-zero DSCPs to DSCP 0, thus marking them for the Best Effort service.

Sample router configurations

Most BARs and all core routers on the JANET backbone are Cisco GSRs. A few BARS are Cisco 7507s. One of the BARs involved in the Project Lancaster BAR was a Cisco 7507 while the others were GSRs.

GSR BAR sample configuration

GSR has a specific distributed architecture affecting QoS configuration. Figure 2 shows a simplified structure. Each line card is shown twice: its receiving part (Rx) is on the left side of the figure and its transmitting part (Tx) is on the right. It features two kinds of queues which can be configured for queuing and/or scheduling incoming traffic:

- Cross-fabric queues (also known as virtual queues or ToFab queues). These store packets going from line cards through a cross-fabric to other line cards. Each Rx line card (slot) supports a separate set of queues for each Tx line card (slot). This set consists of queues for every class of traffic. (In our case we used three kinds of queues – for IP Premium, Best Effort and LBE traffic.)
- Output queues in the Tx part of a line card, known as FromFab queues. These contain packets going to the Tx interface of a line card.

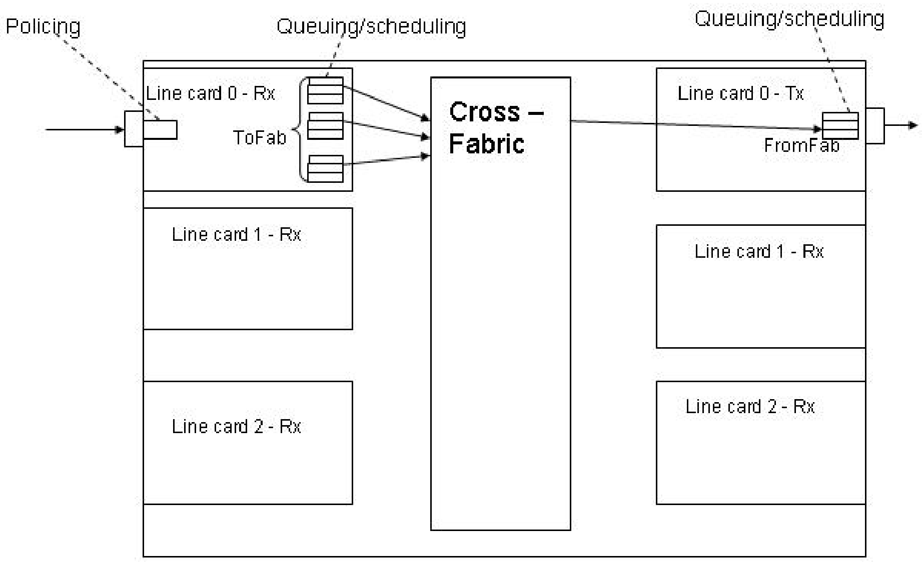

Figure 2. GSR queues.

Policing could be applied to the Rx interface of a line card.

Different types of line cards known as Engine 0, 1, 2 or 3 support different sets of QoS features and different commands for QoS configuring.

An example given below is a typical JANET BAR's QoS configuration, which uses legacy QoS CLI (command line interface) for ToFab and FromFab queuing, and Modular QoS CLI (MQC) for policing. It is also possible to use MQC for FromFab queuing, while for ToFab, legacy CLI is the only way of configuring.

1. Policing

```
: class-map match-any premiumplus
:   match ip precedence 5
: class-map match-any nearpremium
:   match ip dscp 40  41  42  43  44  45  47
: class-map match-any lbe
:   match ip dscp 8
: class-map match-any premium
:   match ip dscp 46
: !
: policy-map police-premium-125M
:   class premium
:     police cir 125000000 bc 4470 be 4470 conform-action transmit
exceed-action drop
:   class nearpremium
:    set ip precedence 0
: !
```

Policy-map “police-premium-125M” was applied to all BAR 2.5Gbit/s input interfaces leading to the Regional Networks.

2. ToFab queuing

(A) A schedule for ToFab queues was defined by cos-queue-groups, which may look like this:

```
: cos-queue-group eng2-rx
:  precedence 0 random-detect-label 0
:  precedence 1 queue 1
:  precedence 1 random-detect-label 1
:  precedence 2 random-detect-label 0
:  precedence 3 random-detect-label 0
:  precedence 4 random-detect-label 0
:  precedence 5 queue low-latency
:  random-detect-label 0 1365 5461 1
:  random-detect-label 1 16 80 1
:  queue 0 99
:  queue 1 1
```

This cos-queue-group says that traffic will go to one of three queues: “0” (default), “1” or “low-latency”. Traffic with precedence 5 (i.e. after policing it only corresponds to IP Premium traffic with DSCP 46) goes to a low-latency queue which is the equivalent of a priority queue. LBE traffic (precedence 1) goes to queue 1 and all other traffic goes to a default queue 0.

The last two lines, “queue 0 99” and “queue 1 1”, implicitly define a bandwidth share for these queues during congestion. The numbers 1 and 99 are not weights but they impact on weights so that LBE traffic receives about 3% according to this configuration example.

(B) A table of different cos-queue-groups specifies a queuing style for every Tx slot.

For example, the following table specifies schedules for a router with two slots, 0 and 11:

```
: slot-table-cos rxcostable     ! Define slot table 'rxcostable'
:  destination-slot 0 eng2-rx   !  All traffic to slot 2 uses profile
:  destination-slot 11 eng3-rx  !  'eng2-rx' while slot 11 uses
                                   'eng3-rx' where eng3-rx is another
                                    cos-queue-group
```

(C) Finally, slot-table-cos tables should be applied. These specify a detailed scheduling of traffic going to a fabric. JANET BARs were configured so that all slots used the same slot-table-cos. To do so, only one command was needed:

```
: rx-cos-slot all rxcostable
```

During tests, BARs were configured with the same queue scheduling parameters for all outgoing slots (e.g. eng2-rx and eng3-rx are identical in the example given above), but it is possible to create several such tables and apply individual scheduling for each Rx-slot/Tx-slot pair.

3. FrFab queuing

A scheduling for FromFab queues used the same set of legacy CLI commands:

```
: cos-queue-group stm16-tx
:  precedence 0 random-detect-label 0
:  precedence 1 queue 1
:  precedence 1 random-detect-label 1
:  precedence 2 random-detect-label 0
:  precedence 3 random-detect-label 0
:  precedence 4 random-detect-label 0
:  precedence 5 queue low-latency
: random-detect-label 0 1365 5461 1
: random-detect-label 1 16 80 1
: queue 0 99
: queue 1 1
```

There are no matrix applications because these are standard output queues of line cards, one queue per Tx card. So,

```
tx-cos  stm16-tx
```

is an example of a command which activates output scheduling for an interface.

Peculiarities found

The link between RAL BAR and the UKERNA development network was by Fast Ethernet. RAL BAR was a Cisco GSR with old Fast Ethernet line cards which were not designed to support scheduling in the output direction. This impacted heavily on the tests in which UKERNA participated and meant that the QoS services could only be tested in one direction, from UKERNA to the sites at Swansea University and the University of Manchester.

UKERNA

UKERNA used its development network for the QoS tests, connected directly to RAL BAR as shown in Figure3. Figure 3 shows only those elements of the network which are relevant to the QoS tests conducted. All but the Cisco 805 interfaces were 100 Mbit/s Fast Ethernet. As the Cisco 805 model only supports 10 Mbit/s Ethernet interfaces, they were connected to the corresponding interfaces of the Cisco switches.

Each Linux box had IPERF[4] and RUDE/CRUDE[5] open-source applications installed and the Windows box had only IPERF. IPERF and RUDE/CRUDE applications were used as traffic generation and measurement tools. Total traffic rate through the output interface of the UKERNA access router was measured by GETIF[6] running on the Windows box.

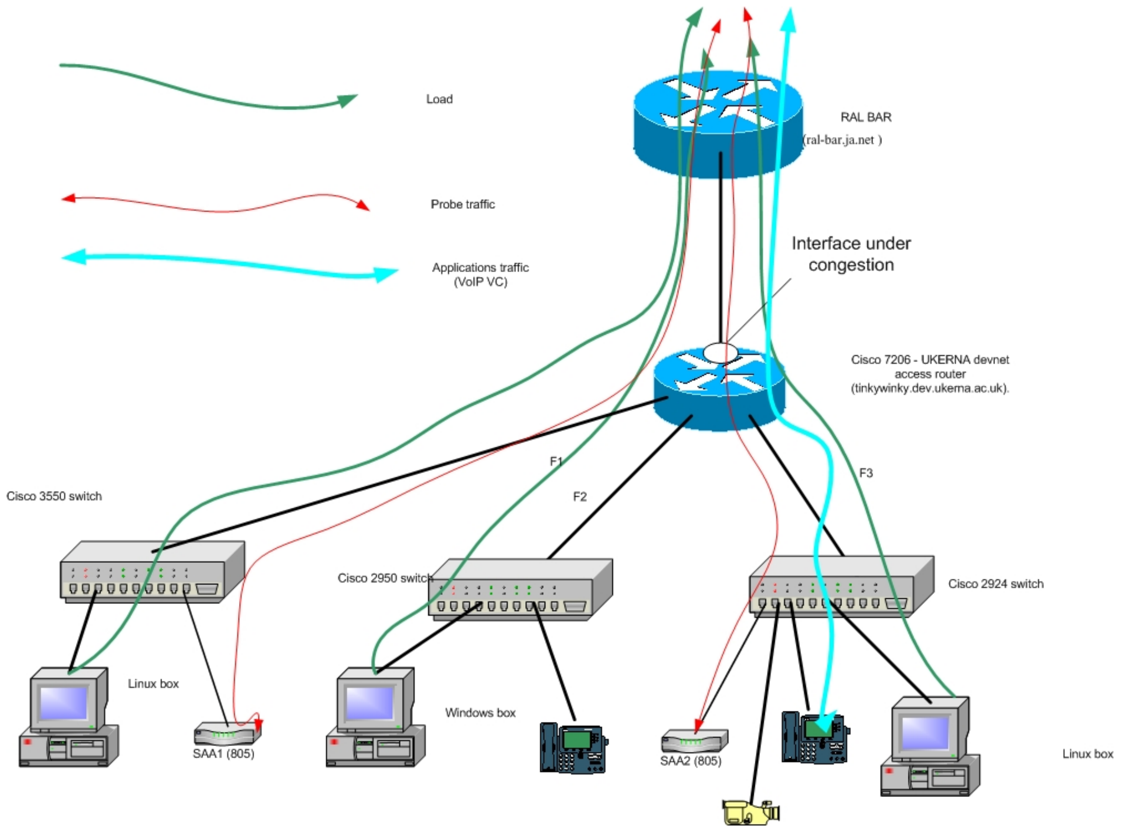

Figure 3. UKERNA development network.

Configuration of the UKERNA development network access router:

Class maps:

```
 class-map match-all lbe
  match ip precedence 1
 class-map match-all premium
  match ip precedence 5
```

Policy maps:

```
policy-map queuing
  class premium
   priority percent 5
  class lbe
bandwidth percent 1
   random-detect
  class class-default
   bandwidth percent 69
   queue-limit 512
```

The “**queuing”** policy map was applied to all UKERNA access router interfaces in the output direction:

```
service-policy output queuing
```

Swansea University

The topology of the Swansea University site network used during QoS tests is shown in Figure 4.

As South Wells MAN is built on Riverstone routers, the testbed consisted of a Riverstone access router connected by switches to several traffic generators and monitoring tools, installed on SUN Solaris platforms. The Site Access Router marked VC traffic as IP Premium and processed it in the priority queue.

A more detailed description of the testbed can be found in the Swansea University final deliverable.[7]

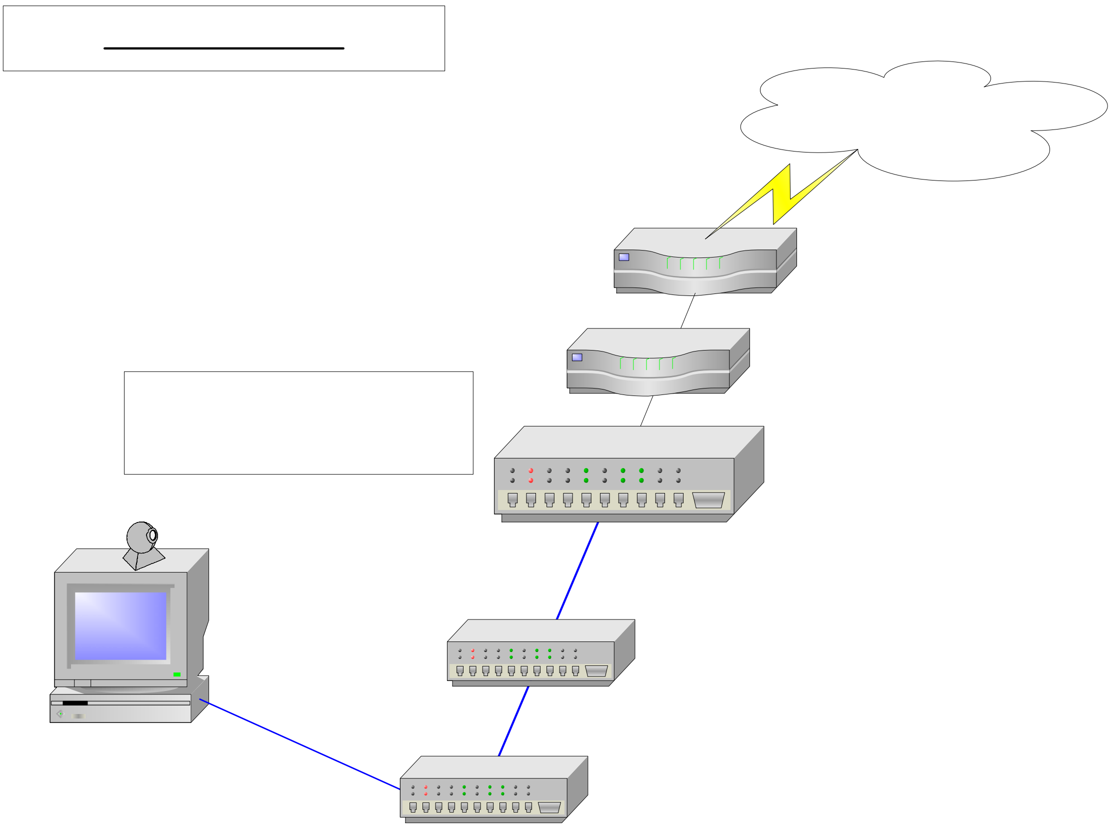

Figure 4. Testbed topology of the Swansea University site

University of Manchester

The University of Manchester used the University campus core network and the Research and Development (R&D) network which was built specifically to check QoS configurations and features before using the University production network. Simplified diagrams of these networks are shown below in Figure 5, together with the relevant part of the Net North West (NNW) Regional Network.

A detailed description of router and switch configuration used at the University of Manchester site can be found in the University final deliverable.[8]

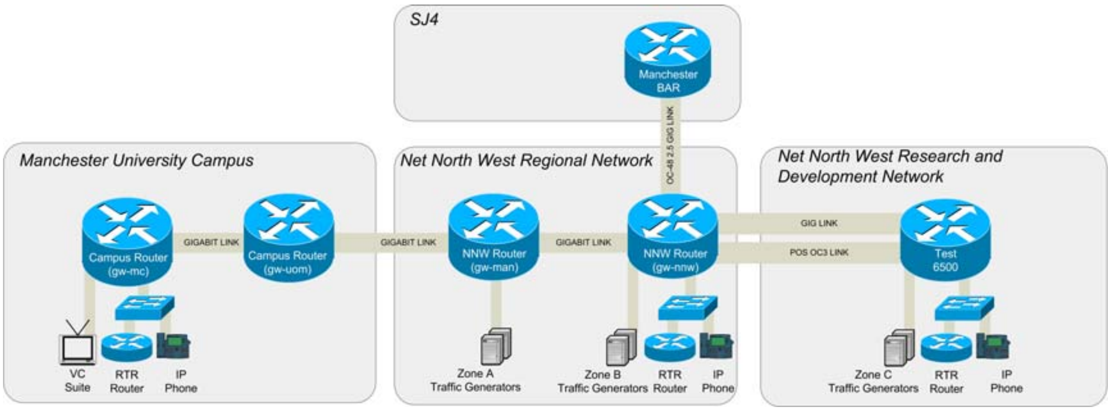

Figure 5. Testbed topology of the University of Manchester site.

Test Descriptions

Introduction

One-way UKERNA tests

During preliminary testing of the UKERNA site in February 2004 it was discovered that the RAL BAR, to which the UKERNA development network was connected, only supported QoS in one direction, namely from UKERNA to the JANET backbone. All traffic from the backbone to UKERNA was treated by the RAL BAR (Cisco GSR) as one class – Best Effort – despite its correct DiffServ marking. JANET NOSC explained that this one-way effect was the result of using an old version of the Fast Ethernet line card in the RAL BAR.

An upgrade of this card was not an option at that time. Therefore UKERNA decided to measure QoS metrics for outgoing traffic only, though traffic itself went in both directions. This also meant that during all tests described in this document the interface of the UKERNA access router (Cisco 7206, Figure 3), connected to the RAL BAR, was only congested for outgoing traffic.

SAA graphs

QoS metrics such as delay, jitter and loss were measured during tests by the centralised QoS monitoring system based on the Cisco SAAs (Service Assurance

Agents). The QoS monitoring infrastructure was established by the project partners; the activity was led by Steve Williams from Swansea University.

The QoS monitoring system is described in detail in the Swansea University final deliverable[9] and in the JANET QoS Development Project report.[10]

The QoS monitoring system produced real time graphs available through a web interface. These graphs are used in a description of each test, given below.

The SAA graphs use the following labels:

- **EF (**or **BE**, or **LBE) AvgRTT**: measured average Round Trip Time in milliseconds for IP Premium/EF (or Best Effort, or LBE) traffic.
- **EF** (or **BE**, or **LBE**) **Packet loss SD** (or **DS**): measured amount of packet loss for IP Premium/EF (or Best Effort, or LBE) traffic from source to destination (SD) or destination to source (DS). As the SAA probe traffic was generated by an SAA agent at Reading C-PoP and reflected by the UKERNA SAA agent, the relevant data are those marked "DS" on the graphs (taking into account the one-way nature of UKERNA tests). The vertical axis of the SAA graphs shows the number of lost packets out of 50, as the SAA agent generates probes of 50 packets. To obtain packet loss as a percentage, double the result shown.

The SAA graphs do not show one-way delays (OWD) and jitter (IPDV, IP delay variance). These parameters were measured during the tests but the data obtained were not comprehensive. OWD could be due to a lack of synchronisation between SAA probes and reflectors – they were only synchronized by the JANET NTTP servers and not by a more precise time source. Non-comprehensiveness of jitter could be due to a big packet loss during our tests which usually distorts the results of jitter measurement. Both effects need further investigation.

Test: UKERNA – Swansea, 2nd March

Test aims

- To check QoS configurations of routers and traffic generators at the UKERNA and Swansea University sites (this was the first test across JANET for these sites);
- To check the functionality of the SAA-based monitoring system;
- To check the degree of protection for IP Premium VoIP traffic against Best Effort bursts.

SAA graph

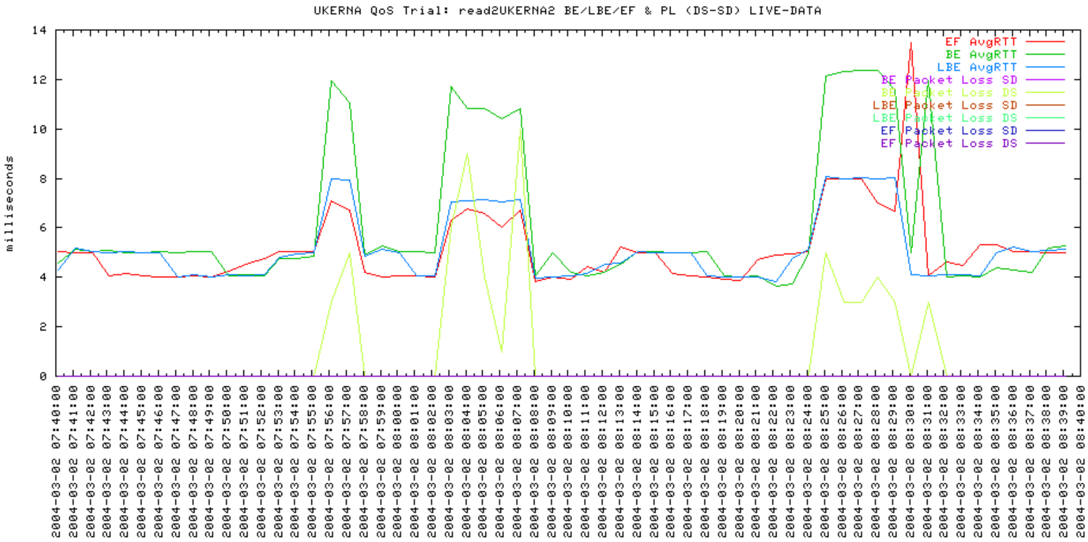

Figure 6. RTT and loss during UKERNA-Swansea test of 2 March 2004.

Figure 6 reflects three distinct periods of the test:

- *8:02 – 8:08.* A VoIP session was established between UKERNA and Swansea using Cisco VoIP phones. Voice traffic was marked by the phones as IP Premium (with DSCP 46).

  Two artificial Best Effort flows created by IPERF were sent towards Southampton University to congest the 100 Mbit/s output interface of the UKERNA access router. The Southampton University site was used as a traffic sink as Swansea University did not have an IPERF server installed (an IPERF session needs both a client and a server participating).

  During this period two IPERF flows, with an average rate of 45 and 60 Mbit/s respectively, congested the output interface of the UKERNA access router up to 100% of its capacity. This resulted in a sharp increase in delays (measured as RTTs) and loss of Best Effort traffic.

  However, the quality of voice, which was served as IP Premium traffic, had not changed and remained very good in both directions.

- *8:24 - 8:30.* The same load was applied but this time the voice packets were re-marked by the input interface of the UKERNA access router from IP Premium to Best Effort (i.e. DSCP 46s were replaced by DSCP 0s). The quality of voice became quite poor in the UKERNA to Swansea direction, e.g. significant crackles were heard.

- *8:30 - 8:33.* Swansea applied an artificial load towards UKERNA. Quality of voice became very poor despite voice packets towards UKERNA having IP Premium marking. This happened because the RAL BAR did not support QoS towards UKERNA.

Test: UKERNA – Swansea, 9th March

Test aims

- To check QoS configurations of routers and traffic generators at the UKERNA and Swansea University sites;
- To check the degree of protection for IP Premium videoconferencing traffic against Best Effort bursts.

SAA graph

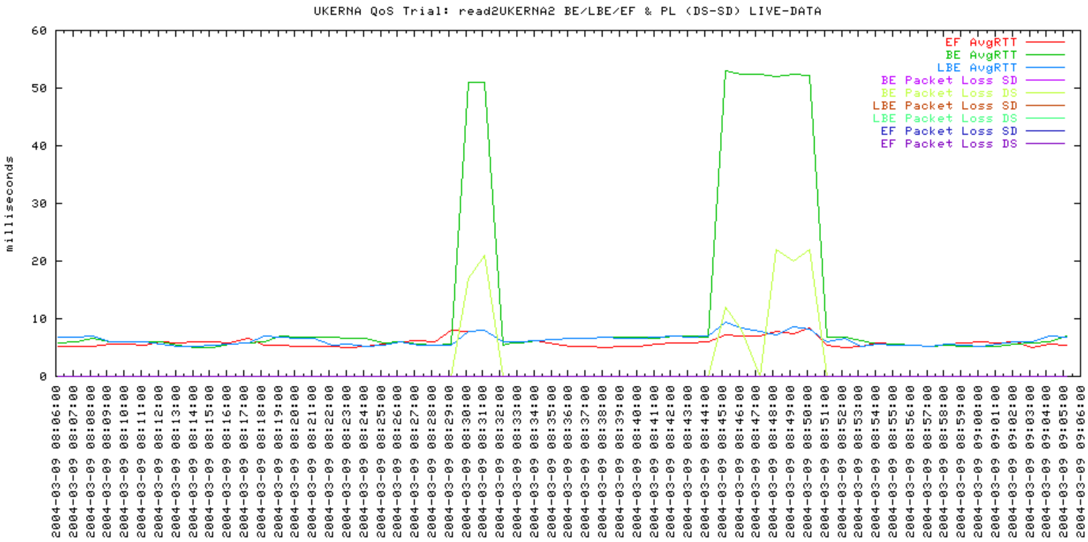

Figure 7. RTT and loss during UKERNA-Swansea test of 9 March 2004.

This test was very similar to the 2 March test except that videoconferencing traffic was examined. The SAA graph in Figure 7 shows a good degree of IP Premium (EF) traffic protection against Best Effort bursts.

The combined artificial load, consisting of two RUDE UDP flows, reached about 96 Mbit/s, which was very close to the total capacity of the UKERNA access router output interface.

From 8:28-8.32 am, videoconferencing traffic was treated as IP Premium. No distortions of video images were observed at the Swansea site during this period.

At 8:43 am the output interface of the UKERNA access router was congested by an artificial load as described above. After this the Swansea University site observed significant video and voice distortions which resulted in automatic cancellation of the session.

Test: UKERNA – Manchester, 16th March

Test aim

- To check the benefits of IP Premium for VoIP and videoconference traffic;
- To check the protection of the IP Premium service against Best Effort bursts.

Path descriptions

The path between the sending box 212.219.210.132 at UKERNA and receiving box 194.66.26.50 at Manchester consisted of 7 intermediate routers, all configured for IP Premium, Best Effort and LBE QoS support.

Tracing route to beacon2.netnw.net.uk [194.66.26.50]:

1	15 ms	<10 ms	<10 ms	cisco.ukerna.ac.uk [193.62.83.59]
2	<10 ms	<10 ms	<10 ms	ral-bar.ja.net [146.97.40.25]
3	<10 ms	<10 ms	16 ms	po10-0.read-scr.ja.net [146.97.35.157]
4	<10 ms	<10 ms	16 ms	po3-0.warr-scr.ja.net [146.97.33.54]
5	<10 ms	16 ms	<10 ms	po1-0.manchester-bar.ja.net [146.97.35.166]
6	<10 ms	<10 ms	16 ms	gw-nnw.core.netnw.net.uk [146.97.40.202]
7	*	*	*	Tested router with private IP address .
8	<10 ms	<10 ms	16 ms	beacon2.netnw.net.uk [194.66.26.50]

Load description

Three artificial flows were used to congest the interface in question: F1, F2 and F3. All flows were Best Effort and generated by RUDE scripts. The flows lasted for 300 sec.

F1:

- generates UDP packets with max (1472 bytes) payload;
- starts with 35.3 Mbit/s UDP load (46 Mbit/s raw Ethernet load)
- adds about 1.1 Mbit/s UDP every 10 sec (first step was about 12 Mbit/s)
- finishes with 82 Mbit/s UDP (84.4 raw load)

To create burstiness in a generated flow, every 10 sec period ended about 300-500 msecs early so that the next step started to create some burstiness, which can be seen on the screenshot.

A graph of F1 recorded by GETIF is shown in Figure 8; the vertical axis is in bytes per second.

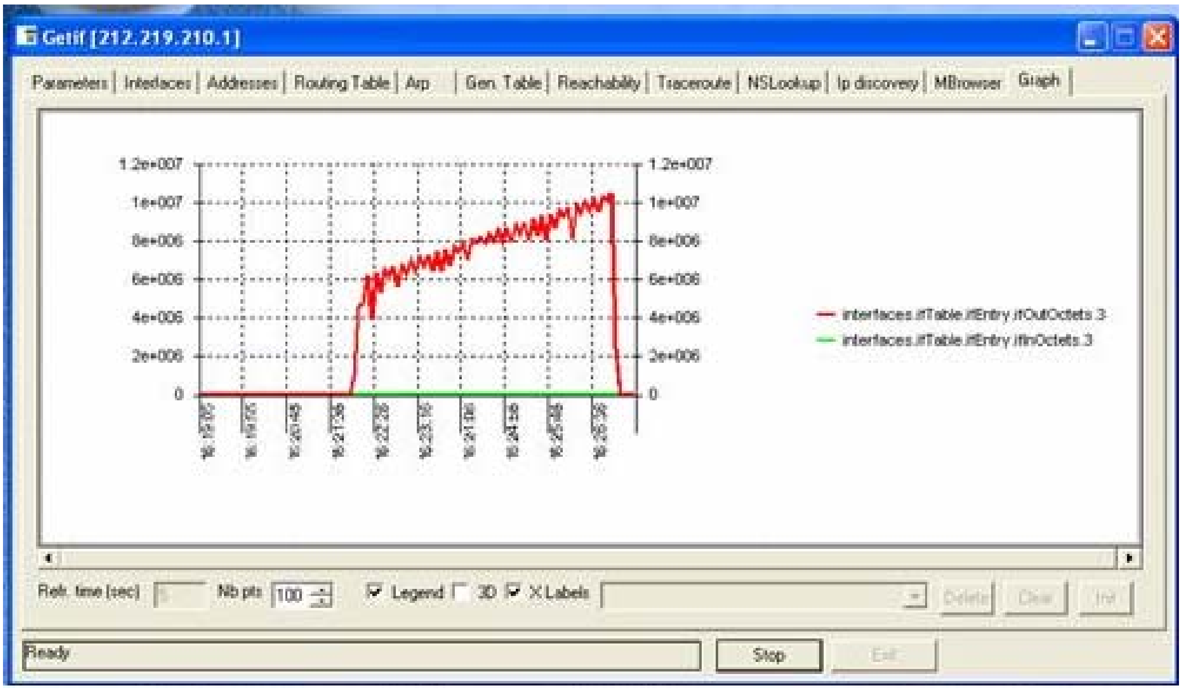

Figure 8. F1 flow generated by RUDE and recorded by GETIF.

F2:

- generates UDP packets with 200 bytes payload;
- starts with 8Mbit/s UDP load (10.64 raw Ethernet load);
- adds about 160 K UDP load every 10 sec with the same mode of pausing between steps as in F1;
- finishes with 12.64 Mbit/s UDP load (16.8 raw Ethernet).

F3:

- generates UDP packets with 500 bytes payload;
- starts with 28Mbit/s UDP load (31.7 raw Ethernet load);
- adds about 400K UDP load every 10 sec with the same mode of pausing between steps as in F1;
- finishes with 39.6 Mbit/s UDP load (44.8 raw Ethernet).

Total Best Effort offered load ramped up from 78 UDP / 88 Ethernet to 134 UDP (146 Ethernet) Mbit/s for 300 sec. The load reached 100% of the link capacity after about 60 sec of congestion.

SAA Graph

Figure 9 shows the SAA graph of 16 March tests together with an artificial load applied.

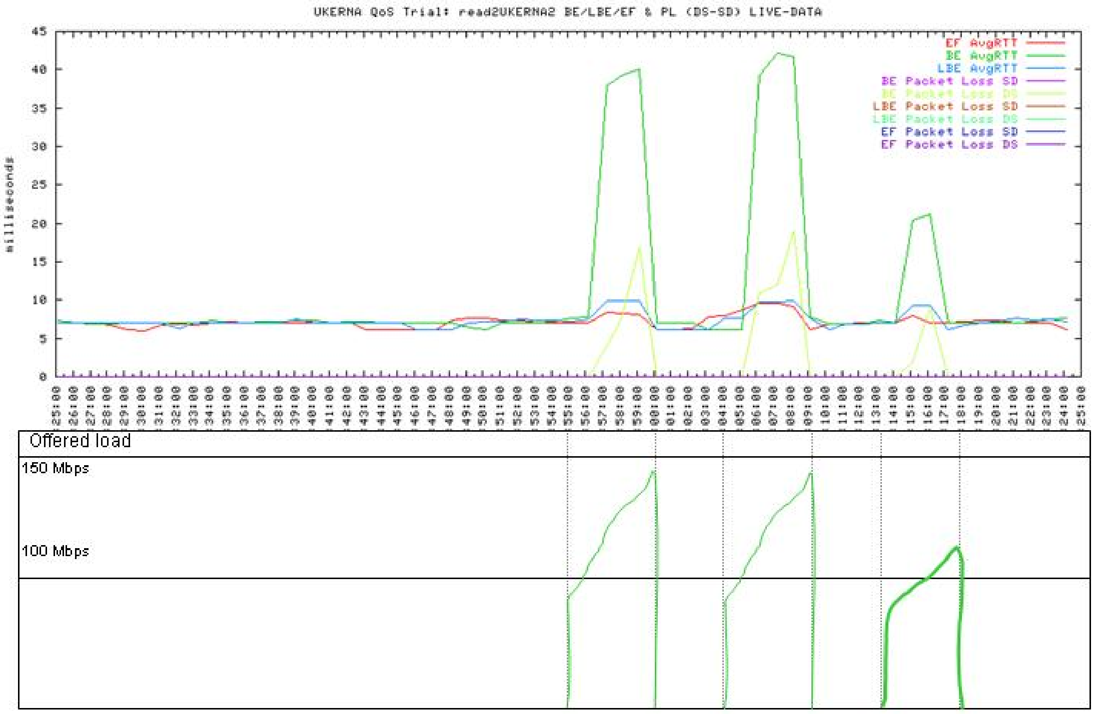

Figure 9. RTT and loss during the UKERNA-Manchester test of 16th March 2004

There are three interesting areas in Figure 9:

- *07:55 – 08:00*. Load was applied for testing the readiness of RUDE generators and other equipment. No VoIP and video applications were running. There were no changes in probe traffic RTTs and loss until 1 minute after the start of load, when total load reached about 100% of link capacity. After the second minute of the load we can see significant increase of RTTs (from 6-7 msec to about 40 msec) and loss of Best Effort traffic(from 0 to 30-35%). IP Premium (EF) and LBE flows remained almost non-affected: a slight increase can be explained as the router's scheduler 'distraction' on Best Effort processing.

- *08:04 – 08:09*. The same load and the same behaviour of probe traffic. Just before this period a videoconference session between UKERNA and Manchester was established and its video and voice flows were processed as Best Effort. After one minute from start of load, several negative effects have started to appear: a slowing down of video picture updating and some voice distortions. After two minutes a video picture froze and voice became unrecognisable. After that, communication between UKERNA and

Manchester was only possible by VoIP because VoIP traffic from IP Phones was served as IP Premium. No VoIP traffic distortions were registered.

- *8:13 – 8:18*. Load was slightly decreased by using only two RUDE flows, F1 and F3. Videoconferencing and VoIP sessions were established during this phase and their traffic was served as IP Premium. No distortions in either videoconferencing or VoIP sessions were observed.

Conclusions

The tests in which UKERNA participated showed clear benefits for videoconferencing and VoIP traffic during periods of severe and long-term Best Effort congestions.

Despite an almost 150% Best Effort load the Cisco 7206 router on the UKERNA development network successfully served IP Premium traffic by means of a priority queue. LBE probe traffic was also processed almost without drops and delays as its rate was far lower than the UKERNA access router output interface's guaranteed 1% (1Mbit/s) of bandwidth.

The benefits of QoS services during periods of moderate congestion need further investigation. It is very likely that such investigation requires more sophisticated traffic generators able to create a very bursty artificial load. More precise measurement tools for one-way delay and jitter registration will be necessary as well.

References

[1] http://www.ja.net/development/qos/partners-deliverables

[2] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and W. Weiss, "An Architecture for Differentiated Services", RFC2475, December 1998.

[3] http://www.ja.net/development/qos/QoS-Testplan-Final.pdf

[4] http://dast.nlanr.net/Projects/Iperf/

[5] http://rude.sourceforge.net/

[6] http://www.wtcs.org/snmp4tpc/getif.htm

[7] http://www.ja.net/development/qos/Swansea.doc

[8] http://www.ja.net/development/qos/Manchester.doc

[9] http://www.ja.net/development/qos/Swansea.doc

[10] http://www.ja.net/development/qos/QoS-final-report.pdf