

QoS in the Janet network

JANET Policy Framework

This section describes the initial JANET QoS Policy Framework, which is based on the DiffServ architecture being implemented at the edge of the network. It is intentionally simple, and may be revised over time as experience is gained in the production deployment of elements of DiffServ.

Transparent Core

The model assumes that the JANET core network is sufficiently over-provisioned for the foreseeable future to remove the requirement for differentiated queuing, and so it behaves like any other standard IP network. As a result, the core will be QoS neutral and DSCP transparent: packets marked with a particular DSCP value will enter and leave the core with the value unchanged.

No access control is performed on traffic entering the core and no shaping takes place on traffic leaving the core. (Ingress policing and shaping may take place at the very edge of the network, on the links in from and out to end sites - see below.) The model requires that the core remain well over-provisioned. With the present JANET backbone this should be easily possible as substantial over-provisioning of the core was an absolute requirement of the SuperJANET5 project.

The model applies to almost all borders with the core - whether to a Regional Network or an external peer. Any end site connections to the core are treated as for any other end site, as described below.

Sites connecting to the Regional Network/JANET

A Regional Network (RN) will ideally also be QoS neutral and DSCP transparent; however, given the variety of architectures in use at the present time, this may not be possible. In these situations, further coordination with the RN concerned is necessary to determine if alternative methods are available - for example, the RN may have already deployed a QoS mechanism.

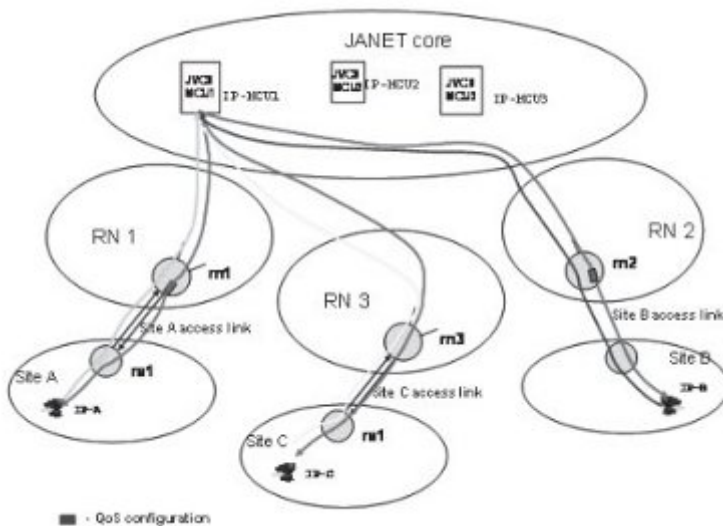
As such, the recommended procedure is that the site may request QoS configuration from its RN, or directly from the JANET NOC if it is directly connected, if it is specifically affected by congestion on its JANET access link. The RN or JANET edge router should then be configured to prioritise traffic accordingly, based on **source and/or destination address** classification of ingress traffic. For example, if a site wishes to have an amount of bandwidth reserved for JVCS (JANET Videoconferencing Service) conferences, the JANET edge router would be configured based on the IP addresses of the JVCS MCUs and/or the address of the

site codec.

It may also be desirable to perform ingress policing, to ensure that the rate of the traffic entering JANET does not exceed the prioritised bandwidth assignment on a site's access link. This will be performed by either the RN or JANET based on the site's point of connection to the network. Using the videoconferencing example, it would be a waste of effort to assign a maximum 1Mbit/s of bandwidth outbound to the site if the other endpoint was sending at 1.5Mbit/s. Not only would the videoconference break down, but JANET would pointlessly carry the extra 0.5Mbit/s of traffic that is inevitably going to be dropped or otherwise rendered useless. Both egress and ingress policing/access control will therefore be required on the JANET/RN side of site access links.

In essence, the JANET QoS model puts the emphasis for QoS configuration at the edge routers, where sites connect to an RN or directly to JANET, as experience has shown that site access links are the most prone to congestion on JANET. However the model does not mean that all these edge routers must be configured for QoS. Only those routers where QoS becomes necessary on an access link need configuration and there is no expectation that, for example, an RN must deploy DiffServ across its whole network. The model is simply concerned with prioritising resources where necessary.

Figure 3-1 illustrates an example of QoS deployment at the edge routers for supporting JVCS traffic. Site A and site B access links were found prone to congestion so that QoS was deployed at the edge routers rn1 and rn3. At the same time Site C access link was found to have sufficient bandwidth to serve all traffic with an acceptably low level of packet delay and loss, making QoS unnecessary for this link. QoS configurations within routers rn1 and rn3 classify ingress traffic using the IP address of the JVCS MCU: IP-MCU1. This restricts priority treatment to JVCS streams only (other traffic is treated normally) and prevents misuse of privileged bandwidth on routers rn1 to which site A and site B are connected.



with Janet Videoconferencing

[1]

3.1.3 QoS Services

The JANET QoS policy provides for the priority treatment of specific traffic on a site access link when it is suffering ongoing congestion. In terms of the DiffServ model, this translates to

providing two classes of service - IP Premium/Expedited Forwarding and Best Effort. This functionality is available to any JANET site on its access link, except where specific technical limitations make it impossible - for example if the router connecting the link to JANET does not support multiple queues, or its implementation may adversely affect other traffic.

The DiffServ architecture provides for many other classes of service, the use of which is not defined within this policy. For example, this policy does not specify the inclusion of Less Than Best Effort service since it is assumed that there is sufficient capacity in the JANET core to make this unnecessary. This, of course, does not preclude their use on a bi-lateral basis between a site and its Regional Network or between two specific sites. These classes are not available to sites connecting directly to the JANET backbone, however.

3.1.4 Use of DSCP Based Classes

Within JANET, the core makes no use of DSCP values as provisioning is based on IP addresses so in theory there are no reserved classes. However, if DSCP values were to be standardised across JANET and used in some consistent way, it might help sites to cooperate when deploying QoS and reduce their efforts on consolidation of approaches. As a result, some work was done within the QoS Development Project to arrive at a list of recommended DSCP values for use within the JANET community and this is available in a separate document [JANETDSCP].

3.1.5 Coordinated Provisioning

This model relies on a manual provisioning system whereby a request for bandwidth is first checked to ensure that implementing the request is feasible. This is currently done on a case by case basis whereby requests are judged on their merits as no unified approach to QoS provisioning is in place. In summary, if the request came from the site sourcing the traffic, the check will also verify that the destination site is aware of and agrees to the priority treatment. All requests should be made to JANET Service Desk (service@ja.net [2]). If all is well, a request will then be sent to the destination RN (or NOC for directly connected sites) to have the appropriate configuration put in place. This configuration may include access control and policing parameters for the edge router to which the destination site is connected.

Provisioning requests might be made for both long and short term flows. For example, a site may wish to have permanent configuration in place to prioritise traffic from the JVCS MCUs. Short term requests for weeks or months might be made to satisfy the needs of research or education projects.

3.1.6 Interoperability with External Domains

There are several QoS initiatives running in both the public and commercial sectors, of which those from the NHS, BECTA and GÉANT2 are the closest to JANET users. Some of the project owners have announced the use of several QoS classes and dedicated DSCP values.

According to the approach described above, a JANET QoS service would accept requests for QoS from an external source of traffic to a JANET-connected end point on the same basis and according to the same procedure as for internal sources. This means that a request for such a service should specify the IP addresses of both ends of a connection and these

attributes will play a key role in access control procedure when real traffic will be directed into JANET. A DSCP value assigned by an external source might be taken into account but only as an additional attribute for a required type of service, not as a main access control attribute.

It is worth stressing that traffic to/from external domains will be treated no differently from internal traffic. It will receive no special treatment within JANET until it reaches the site access link where QoS configuration has been requested and configured.

3.2 Provisioning and Maintenance

The description given in this section assumes the simplest case of QoS configuration on an access link according to the JANET QoS Policy: requested and authorised traffic is placed in a priority queue, effectively providing the IP Premium/Expedited Forwarding DiffServ class of service. All other traffic is placed in a single other queue - DiffServ Best Effort.

Classification of authorised traffic at the edge should be based on the use of IP addresses; DSCP values will only play an auxiliary role for the classification. Having only one priority queue means that all traffic which needs better than Best Effort treatment is placed in that queue. To become authorised, traffic should pass through admission control which means that its requested sustained rate will be checked against available bandwidth of a priority queue at the output interface of an edge router.

3.2.1 Provisioned Parameters

IP Premium provisioning deals with the following parameters [JANETThinkTank]:

- Operational per-flow policing limit. To prevent excess traffic from being admitted to the priority queue on a link, ingress policing is deployed at the JANET side of a site access link. In the JANET model, individual flows are identified based on source and destination IP address, and may include parameters such as TCP or UDP port number.
- Aggregate IP Premium limit. The maximum proportion of IP Premium traffic allowed on a link. For any particular access link, an overall amount of bandwidth reserved for IP Premium traffic is known, which is the aggregate of all traffic that has been agreed to be prioritized. This summary amount of privileged traffic should not exceed the technological limit beyond which IP Premium loses its elevated features as packets start to congest the link themselves. As was stressed above in section 2.4, the precise value of this limit is unknown; however, 20% seems to be a reasonable estimation which guarantees oversubscription mode for IP Premium traffic.

The simplest way of enforcing the aggregate IP Premium limit is manually: a person who makes a decision about accepting or rejecting the next request for IP Premium should add the requested bandwidth to the aggregate amount of IP Premium previously agreed for this link and then check whether or not this new aggregated amount exceeds 20% limit of total bandwidth.

Aggregate IP Premium rate may be policed as an additional measure against human error but it is not mandatory (or possible for some router models) as the possibility to do aggregate policing with individual flow policing depends on a router implementation.

- Traffic shaping. This technique might be deployed, as close to the traffic source as

possible, to smooth out bursts of traffic and prevent packets being dropped by policing. This is done by buffering packets and transmitting them into the network to keep within the flow's bandwidth assignment. If the traffic source is sending data at such a rate that the buffer fills then clearly shaping will not help and an adjustment of the bandwidth assignment is required (assuming that the high traffic rate is not due to a problem).

The ultimate place for applying policing and strict classification is the egress router on the JANET/RN side of a site access link. Ingress policing and classification might also be deployed at the same point, but is not mandatory in the JANET model. Not deploying ingress policing and strict classification places the responsibility on the users of the bandwidth assignment to keep within its bounds, otherwise traffic transits JANET for no reason other than to be dropped.

3.2.2. Dynamic and Static Provisioning

The provisioning of QoS in the JANET model is manual and static as it involves a user request, followed by effort from staff at the site, Regional Network and JANET(UK). All necessary network configurations are specific to the request, and remain until they are removed, manually, whether the QoS requirement is for days, weeks or longer, and whether the requirement is still there or not. Static provisioning also makes sense for established stable services, such as JVCS, where the requirement is likely to be ongoing. Requesting QoS configuration every time a videoconference was to take place (via JVCS or some other system) would clearly make little sense, on staff effort grounds if nothing else.

The drawback of this approach is that the bandwidth assigned from the IP Premium class still figures in the calculation for whether further QoS requests can be accepted on a site's access link. If a site with a lower speed, congested JANET access link made little use of JVCS, it may prefer not to have bandwidth for JVCS permanently assigned. To stress, though, this does not mean that Best Effort traffic cannot make use of any bandwidth unused from a QoS assignment. It simply means that it is taken into consideration when assessing request for further prioritised bandwidth.

Where a time limit is placed on a QoS request, the provisioning might be classed as dynamic. The term 'dynamic' might be confusing as it sometimes refers to the possibility of automatically establishing service by a user or an application. In the context of this section, it simply refers to the duration of a connection. If a connection is established on a shortterm basis, for example, for several weeks or months, it is considered a dynamic connection which needs dynamic provisioning. In this case, while each request still goes through a normal procedure of checking available bandwidth and configuring access control and policing, it treats bandwidth in a more economical way and is more universal. However, it clearly requires more staff effort as requests must be validated and configured manually on an ongoing basis.

3.2.3 Towards Automation of Provisioning

The model described above greatly simplifies provisioning compared to the general case when all interfaces are treated as congestion prone. As a result, the procedures for checking available bandwidth (i.e. fulfilling admission control) and placing QoS configuration for only one egress interface are relatively simple and can be done manually.

However, when a network supports many such interfaces then automated tools may be of use. The first step might involve deploying a standard database of interfaces including their capacity and a record of any current QoS assignments made. Then, with some small level of database or other scripting, this system could provide a simple 'yes/no' answer to assignment requests, doing automated admission control without requiring any detailed knowledge of the network.

The next step towards fuller automation of provisioning might be a reservation system such as the GÉANT2 AMPS system [AMPS]. These types of system maintain a model of the network state and can potentially re-configure equipment automatically or, in most cases, send notifications to operators and engineers for the required configuration.

3.3 Monitoring and Measurement

Monitoring the operation of QoS across a network is an important function, particularly when troubleshooting or reporting on the behaviour of traffic in different classes. Unfortunately very few packages, commercial or open-source, are available that produce this type of information and few pieces of network equipment report statistics on a per-queue basis.

QoS monitoring also poses interesting problems in that when the network is performing as expected, it is likely that the most interesting statistics will be end-to-end, showing that particular requirements are being met such as end-to-end delay. When there is a problem however, the end-to-end figures will likely only be the first indication and finding the location of the problem requires knowledge of how the queuing is working at each hop between the source and destination. As already noted, per-queue statistics are not a common feature in network equipment so it can be hard to determine where and exactly why packets may have been dropped or held in a long queue.

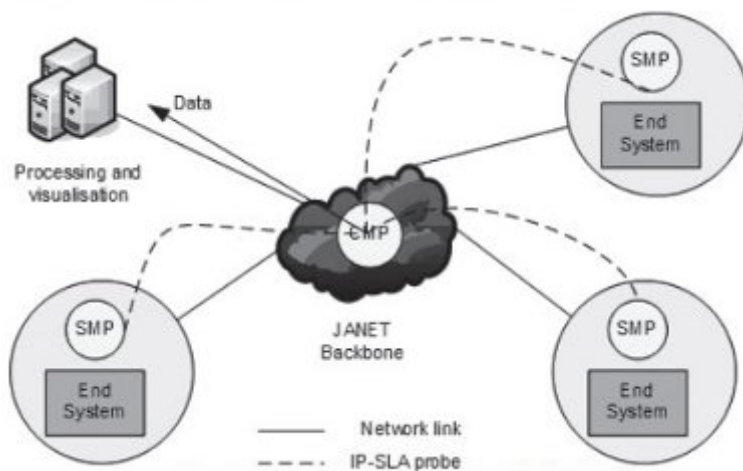
Beyond the existing Netsight [Netsight] system, during the JANET QoS trial, a bespoke system was deployed based on Cisco® IP-SLA, which was formerly known as Cisco® Service Assurance Agent (SAA) and before that Cisco® Response Time Reporter (RTR). The system gathered data on the end-to-end response times and packet loss of the Premium IP, BE and LBE classes used in the QoS experiments, and was primarily used to verify that the QoS deployment behaved as expected under the different (artificially created) types of network congestion conditions. The system did not attempt to address gathering of per-hop statistics. The system used during the trials is documented below as an example of what can be done, and was not designed to be packaged as a deliverable of the JANET QoS project. The software is therefore not available for general distribution, although example software configurations for the IP-SLA probes are provided for information only in Appendix A1.

3.3.1 JANET QoS Trial Monitoring Architecture

The monitoring system was based on a distributed architecture, with a core monitoring point (CMP) and a set of site monitoring points (SMP) placed as close as possible to each end-point/end-system used in the various experiments. A Cisco® 7206VXR/NPE400 was used as the CMP and was located in one of the JANET Core Points of Presence (C-PoPs), so that it was in the centre of the network. Cisco® 805 low-end routers were initially used as the SMP equipment; however Cisco® 1760 routers were used later instead as the power of the 805 was not enough to provide stable real-time calculations.

The CMP ran IP-SLA probes to each SMP, consisting of a burst each minute of 50 packets, sent at 20ms intervals between each packet. One probe was sent for each class of service - BE, LBE and EF - to each SMP. Over time, the results of these probes indicate differences in RTT and packet loss for each traffic class. In some of the experiments, probes were also run directly between SMPs. The IP-SLA feature uses the router's own system clock to timestamp probe data, so care had to be taken to ensure that the clocks on the CMP and all the SMPs were as tightly synchronised as possible.

Note that this system is not taking measurements of actual production traffic as it introduces its own traffic to obtaining an indication of whether the network is behaving as expected. In the JANET QoS experiments, SMPs were placed as closely as possible to the end-points, as these were few and clearly defined. Interpreting data gathered from the nearest practical locations in a full-scale deployment of QoS, where many users are using the same traffic class, may be more difficult. An example configuration of the IP-SLA probe is given in Appendix A1.

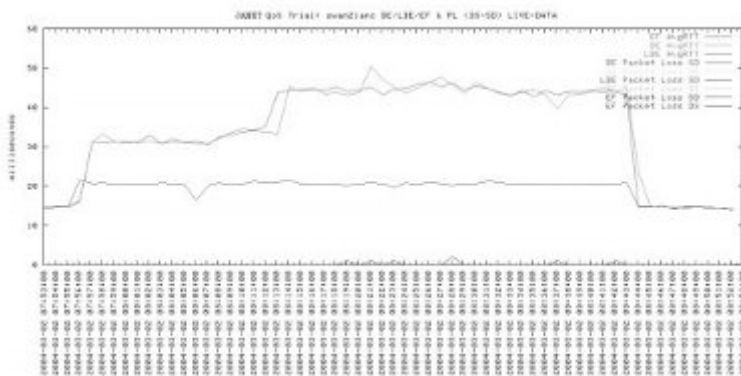


[3]

3.3.2 Data Archive and Display

The IP-SLA process on the CMP stores the results of each probe within the router itself, and the results were not available via SNMP at the time of the trials (this may still be the case). An archiving solution was developed using scripting to retrieve the data via the router CLI, and then store the statistics in a database. The resulting information was then processed every minute to produce graphs of the data, and these and the raw data itself were published on a web server. Extended time lapse graphs were also created, for the past 60 minutes and for the previous 24 hours. The example graph in Figure 3-3 shows data gathered during one of

the experiments. It shows the RTT of EF traffic (the lowest line) as well as BE and LBE traffic from Swansea University to Lancaster University. This test focused on the behaviour of IP Premium/EF under congestion - BE and LBE were treated identically.



[4]

Congestion was artificially introduced into the network and the effect of this traffic can be seen to start affecting performance at 07:55, where the RTT of BE and LBE rises sharply whilst that of the EF class remains low. An increase in RTT of EF traffic of approximately 5ms went unexplained, but the differential treatment of the classes is apparent, thus confirming the correct configuration and operation of QoS across the network.

3.4 Managed Bandwidth Services in JANET

Managed bandwidth based on ATM networks was a feature of SuperJANET, SuperJANET2 and SuperJANET3, although relatively little use of it was made. This was due to a combination of factors, including limited experience, the reach of the ATM network, network capacity and actual realised demand. The development of SuperJANET4 coincided with the period of growing availability of high-speed optical networks, dramatically changing this picture. As such, technologies which had previously been used exclusively by internal telcos became available for end users. Normally, this is not a commodity service like IP; it is rather for 'heavy' demanding users who routinely exchange large amounts of data. Prime examples of such users in this scope are the scientific and research community.

UKLight was the first UK-wide project to target this area. It resulted in the creation of an optical infrastructure (separate from the production JANET IP infrastructure) based on SDH multiplexers. UKLight allowed UK academic organisations to establish end-to-end bandwidth channels between their sites within the UK or to sites of their research partners in Europe and the US. UKLight became a successful service providing more than 20 connections between major research centres with bandwidth from hundreds of Mbit/s to 1Gbit/s.

The introduction of the JANET Lightpath service combined the UKlight facility and the transmission network of the JANET backbone into a single service. This raised the maximum capacity of a lightpath to 10Gbit/s.

JANET Lightpath does not use the packet infrastructure of JANET, but instead provides users with point-to-point links of guaranteed bandwidth - effectively a hard QoS limit. Any more fine grained QoS that might be used on networks built with JANET lightpaths is completely under the control of the body managing the layer 2 or 3 network operated over the lightpaths.

3.5 Experiences with a QoS-Aware SuperJANET4

Since the rollout of the 10Gbit/s backbone under SuperJANET4 in 2002, the network core has always maintained sufficient capacity to avoid congestion. There are, however, areas of the network outside the backbone which may be susceptible to congestion, particularly as traffic levels continue to grow. The first phase of the JANET QoS Development Project [JANETQoSP1] was established to define an open, non-proprietary framework for QoS services that, as far as possible, will not exclude any part of the end-to-end path between users across JANET. The project also aimed to assess the effort required to deploy and manage QoS services on the production multidomain and multivendor networks, and evaluate the effectiveness of QoS services for applications used by the JANET community. More than ten UK universities, in partnership with the Regional Networks to which they are connected, expressed an interest in participating and six universities confirmed their participation and committed resources to the Project.

3.5.1 QoS Services Tested

Three types of QoS services were selected for evaluation: IP Premium, to serve time-sensitive applications; BE for the common Internet applications; and LBE for bulky elastic traffic which could run on the network during periods of low activity from other applications. Despite the differences in QoS implementation on different router platforms, the IP Premium service is usually based on processing traffic in a priority queue, which decreases delays and loss of packets for time-sensitive applications. The LBE service is usually given a very low guaranteed level of a router's resources (usually 1%-3% of the interface capacity), with the possibility of using more resources as they are freed by IP Premium and BE traffic. VoIP and videoconferencing traffic was selected for the IP Premium service as these were considered to be the most important time-sensitive applications for the JANET community. The effect of bulky Grid applications using LBE on BE AccessGrid traffic was investigated for the LBE service trial. All other applications came under the BE class of service.

The main aim of the QoS trial was to assess the behaviour of time-sensitive application traffic on the production network at times of congestion. Traffic behaviour was examined under two conditions: when certain traffic is processed as BE and when the same traffic is processed in the high priority DiffServ class. Comparison of traffic behaviours allowed the project group to assess the impact of enabling QoS for certain applications. A phased approach was taken, to decrease the risk of the IP Premium traffic having an impact on the JANET production service. (LBE traffic cannot distort BE production traffic as BE traffic is always given precedence.)

By the end of 2003, all routers on the core and selected BAR routers on JANET were configured to support the three QoS classes. Access links to the Regional Networks that were participating in the trial were configured appropriately to accept IP Premium and LBE traffic from the participants. All other access links to JANET were configured to mark all IP Premium and LBE packets as BE, in order to process them as normal production traffic. The JANET backbone was configured with the following objectives:

- to give the other participants an example of possible QoS configurations
- to evaluate the efforts necessary for configuring and managing QoS across the backbone
- to check that the Cisco® GSRs deployed on the JANET backbone could support QoS.

By March 2004, all areas of JANET (including the respective parts of the Regional Networks and Campus Networks) that provide the end-to-end paths for the trial traffic had been configured and activated with QoS. During January-February 2004, all participants in the QoS trial were conducting local tests on their campus networks in preparation for the main tests to be conducted across JANET.

3.5.2 Results of the Tests

The main series of tests on JANET were carried out during March-May 2004 by six universities, their Regional Networks, and JANET. The results showed that the IP Premium service gave significant benefits to both VoIP and videoconferencing applications during congestion periods, especially at times of severe congestion. The LBE service also showed positive results, using all the remaining bandwidth available on the network after IP Premium and BE traffic is processed.

During the tests the behaviour of the applications was compared subjectively and objectively. The subjective assessments were made both while QoS was enabled and disabled. During severe congestion periods, participants observed different types of distortions on the quality of BE voice traffic during VoIP sessions, and on the BE traffic for both voice and video during the videoconferencing sessions. These distortions were not observed when VoIP and videoconferencing traffic were served as IP Premium. For the objective assessments during the tests, the participants used the centralised monitoring infrastructure based on the Cisco® SAA as described in section 3.3.1. Open source tools such as IPERF and RUDE/CRUDE were also used for traffic measurement to verify the data obtained from SAA.

The results obtained from the measurement tools showed the positive impact of QoS on the behaviour of the applications, especially the significant reduction in packet loss for IP

Premium traffic against BE traffic, and for BE traffic against LBE traffic under the same network conditions. At the same time, the tests revealed some unexpected router behaviour in terms of QoS support. For example, AccessGrid videoconference traffic, which was served as BE, was suppressed by LBE traffic, which resulted in breaking the established voice and video sessions. A possible reason could be the multicast nature of the AccessGrid traffic, as unicast BE traffic in the other tests was not suppressed by LBE.

The tests demonstrated the clear advantages of enabling QoS on a multi-domain and multivendor network for selected time-sensitive applications during congestion periods. These results were also used as part of the process to define the JANET(UK) approach to QoS on

the present JANET backbone under SuperJANET5. Based on the successful results and experience of running QoS across a selection of JANET, at a meeting held in London on 16 June 2004, the project partners recommended continuing work in the QoS area by establishing Phase 2 of the JANET QoS Development Project.

Source URL: <https://community-stg.jisc.ac.uk/library/janet-services-documentation/qos-janet-network>

Links

- [1] <http://community.ja.net/system/files/images/tg-qos-3.1.jpg>
- [2] <mailto:service@ja.net>
- [3] <http://community.ja.net/system/files/images/tg-qos-3.2.jpg>
- [4] <http://community.ja.net/system/files/images/tg-qos-3.3.jpg>