QoS deployment in site networks

In this section we discuss the issues facing a site that is considering the deployment of QoS in its network. We first run through some of the general considerations for a site before looking at some of the specifics of certain technologies that may impact QoS deployment, along with the user and application viewpoint for QoS. Finally we discuss how these issues may vary for large (university campus and satellite) and small (single site college) networks.

5.1 General Considerations

The first question to ask when considering a QoS deployment is the rationale behind such a deployment. Do you have a specific application requirement, such as a university-run VoIP system, or is it a reaction to an observation of general network congestion on parts of the network? It is obviously prudent to have a clear idea of the problem you are trying to address when deploying QoS support. Whatever the reason, there are some general points to consider before exploring the details of a deployment. These include:

- **How is network traffic currently monitored?** Do you have the capability to determine which network links are potentially congested, from the core down to individual switch edge ports? If you cannot determine where congestion is occurring, over the course of a day, with views of historical data, it may be prudent to facilitate that before creating a solution to a problem that you may not have the full picture for. It may also be desirable to identify the top data flows in the network, and how these may change over the course of a regular day or week.
- **Where is the congestion happening?** If you have a clear idea of where congestion occurs and the QoS deployment is in reaction to that congestion, then you should be able to determine the appropriate action to apply on the congestion point(s). This may most typically be at the border between the site and the Regional Network Operator. It is perfectly possible to take the approach of combining over-provisioning with targeted QoS methods in reaction to observed problems.
- **Which applications require prioritised QoS treatment?** When deploying QoS, perhaps pervasively, to prioritise certain application traffic on the network, you should have a clear specification of the QoS requirements of the application(s), in terms of traditional QoS metrics that can both be applied and measured for compliance. The most commonly cited applications for elevated QoS treatment are VoIP, video streaming and videoconferencing.
- **Which applications could be deprioritised?** There may also be certain applications for which traffic can be deprioritised which may, for example, include non time-critical data backups or transfers. By doing so, regular traffic may be less affected by periodic heavy use of the network by bandwidth-hungry, but non-time-sensitive, applications. This may also facilitate the use of such applications during working hours where such operations have not been considered previously due to its potential impact on ‘regular’ traffic. In UNIX terms, deprioritising traffic is the equivalent of using ‘nice’ to use only spare CPU cycles in long running computations.

- **Are there any specific high-bandwidth applications?** Are there any high bandwidth applications, either within the site or running to/from other sites? These should be identified with a view to determining whether QoS methods should be applied. Typical intra-site usage may be for distributed backups, while remote traffic may include specific grid or e-Science applications. We describe JANET's support for managed bandwidth between sites below, but ultimately an end-to-end QoS arrangement will involve all participating sites, their Regional Network Operators and JANET.
- **Are there any specific media constraints?** There may be specific media for which you wish to improve performance. With gigabit and now ten gigabit Ethernet links available, the most common 'problematic' media is likely to be 802.11 wireless networks, e.g. (theoretical) 11Mbit/s 802.11b, or 54Mbit/s 802.11a/g. Here, options may include configurations to help protect such media from unwanted broadcast or multicast traffic.
- **Do you currently use IP multicast?** Multicast is a technology that can reduce bandwidth demands where users are sending/receiving identical concurrent streams of data; rather than multiple unicast streams, a single multicast stream is used, saving bandwidth. Where it is deployed, QoS may also be applied to the multicast stream(s), e.g. where multicast is used for AccessGrid® conferencing or reception of live video material.
- **Is the existing network equipment QoS capable?** It is important to inventory the existing network components to assess their QoS capabilities. For layer 3 routing equipment, the ability to classify, prioritise and queue accordingly IP traffic is required, to support the operation of a Differentiated Services (DiffServ) domain. It is relatively unlikely that any Integrated Services (IntServ) support will be required. For layer 2 switching equipment, the ability to configure port prioritisation is required (typically using IEEE 802.1p). In addition, various network components may be used that should support QoS if an end-toend service guarantee is required, e.g. IP firewalls.
- **Does the current procurement specification include QoS requirements?** It is prudent that any tenders issued should specify the required of support for QoS methods, in particular support for DiffServ and 802.1p. This should allow the procurement of equipment with a common baseline of QoS capability, even if QoS configurations are not applied pervasively in the network.
- **Are there existing network usage policies?** Assuming there are existing policies, these may help to map out how QoS services may be defined and be expected to be used. Typically, certain traffic may be prioritised or deprioritised in accordance with those policies. It is also important that the users understand these, so application behaviour is (to them) more predictable. A QoS deployment in principle is an implementation of policy, so it is important to review those policies prior to any deployment.
- **Are QoS management tools available?** When deploying QoS, one should have a set of management and monitoring tools that can be used to enforce and check the implementation. The network equipment vendor may have a specific package that enables this, or it may be part of a general network management package. The key aspects are the ability to configure devices (including admission control), to police (in particular) prioritised QoS-marked traffic, and to be able to account for that traffic.

- **Is there an SLA between the site and RNO?** If so, it may be necessary to negotiate QoS elements for that SLA. This might include the maximum percentage of Premium IP that can be sent and what happens to out of profile traffic (e.g. is excess Premium traffic dropped or remarked to Best Effort?). The SLA should apply to both inbound and outbound traffic.

These topics provide some initial areas to consider and investigate.

5.2 Technology Considerations

In this section we look at specific network elements that may require ‘special’ handling in a QoS deployment, either because they are potential locations for network congestion or they may have implications (complications) for a QoS deployment. The degree of complication depends on whether you are trying to deploy QoS to address specific network bottlenecks or to set up an end-to-end path with service guarantees. In the latter case, devices on the path that do not fully support QoS may need attention, or may need to be bypassed in some way for a specific traffic flow or set of flows. These issues have been discussed in general in a previous section; here we discuss examples of deployment-oriented problems.

5.2.1 Proxies, Firewalls and Middleboxes

There are various types of these devices.

It is now very common for sites to have one or more firewalls in their network. Each such firewall may add latency to traffic that passes through it, depending on the overall volume of traffic and number of firewall rules being applied. In some cases, a combined router/firewall device will have QoS support and can be configured to queue traffic in addition to filtering it. In the general case a firewall will be equal in the way it handles all traffic. Where end-to-end QoS is required, it is advisable to check with the firewall vendor for specific methods to support QoS on the firewall device.

In smaller sites, network proxies and traffic controllers are becoming increasingly common. These may be application-specific proxies or they may be ‘smart’ devices that shape traffic according to some defined policy for the site. Such shapers are generally not viewed as classic QoS devices, in that they may not support or honour DiffServ QoS, but they may be configurable to provide an equivalent that solves the same problem. Again, it is advisable to check with the vendor on DiffServ capabilities of such devices.

In both cases, if a device does not support DSCP-based DiffServ prioritisation, it is useful also to check that DSCP values are passed unaltered through the devices so that some QoS handling can be applied elsewhere on the path.

5.2.2 Network Address Translation (NAT)

NAT appears to be used quite commonly in smaller college networks, particularly where network support is outsourced. The principle problem that NAT adds from the QoS perspective is that the IP identities of nodes may be ‘hidden’ by the NATing process. A number of hosts may appear to be the same node in the eyes of a QoS classifier because a NAT box has mapped a number of private internal IP addresses to a single global IPv4 address. The simplest solution here is not to perform NAT on devices that need specific QoS

treatment.

Note that while the availability of new IPv4 address space is rapidly approaching exhaustion (predictions currently lie around 2009-2011), there is currently no problem obtaining public IPv4 address space from JANET, where justification can be provided.

5.2.3 Site Border: Access Technologies

The most likely congestion chokepoint on a network is its connection point to the upstream provider, typically a campus-to-RNO connection. A large university site may enjoy a 1Gbit/s uplink while a smaller college will typically have less, perhaps only some tens of Mbit/s.

The site border is probably the single most effective point at which to begin applying traffic (de)prioritisation. It might be tempting to then only configure QoS on that edge device; however that may not be prudent. The border devices need not perform classification; that function can be applied on other trusted internal devices. It is often desirable to apply DiffServ classification as close to the source of the traffic as possible, since the traffic volumes are likely to be less there, and thus classification is less of a load on the device. This approach does require trust of those internal devices, but these are usually in a single management domain in a typical campus.

The same principle may apply to ‘borders’ to remote sites within a single administratively managed campus, where leased line or wireless point-to-point links are used to remote campuses. These will typically be the first points to apply ‘reactive’ QoS solutions to network congestion.

5.2.4 Internal Site Media: Ethernet / Wireless

A large campus may today be deploying multiple trunked 1Gbit/s Ethernet links in its core, or possibly 10Gbit/s Ethernet. As with the JANET core, the campus core is unlikely to be congested. Congestion is more likely where lower bandwidth links or media are present, which today is probably where 802.11a/b/g wireless networks are deployed within the site.

Support for QoS on wireless access points (APs) tends currently to be somewhat vendorspecific. While some support QoS via 802.1p (much as they support VLAN tagging/handling), capabilities are still generally in their infancy.

One approach to this problem is to try to protect the wireless networks from excessive traffic, e.g. restricting where multicast traffic (via IPv4 IGMP or IPv6 MLD snooping) and broadcast traffic (via subnet sizing) can flow. In many cases, reception of multicast on wireless LANs is desirable (e.g. listening to audio content while on the move), but at the same time any significant multicast traffic on a wireless LAN can significantly impact unicast performance. Many wireless APs ship with a default multicast rate limit configured.

5.2.5 Remote Access Considerations

Many organisations (sites) require remote users to access their site services via a VPN, which places the user within the site from a network security perspective. As tunnelled connections, over arbitrary networks, VPN connections make it very difficult to deploy QoS in support of

remote users.

Some sites choose to enforce their wireless users to connect to a local VPN server to access internal services. Due to the likely higher bandwidth demands of such users, such VPN connections are often rate limited. Again, this makes general QoS support for applications run by VPN users problematic to deploy.

5.3 User / Application Requirements

Another important issue to consider when assessing the deployment of QoS within a site is the requirements for the QoS-enabled applications that will be used and the types of users that will be supported. A more thorough investigation of these issues will be provided in section 5.5.5 (Policy) and section 6 (Applications of QoS) but a summary of the issues is provided below.

For large sites, the application requirements for deploying QoS will be similar to those discussed in section 3.1 and 4.2 in that a range of applications may require some level of service and must be classified appropriately. In addition to bulk transfer applications (which may be given a Less than Best Effort service) and web traffic (which will typically be classified Best Effort), a number of applications may be selected for prioritisation. In summary, we identify five groups of QoS application that might be classified in this way within a site:

- VoIP
- Videoconferencing
- Multimedia Streaming
- E Science / Grid
- Control Traffic

Each of these applications has specific requirements and tolerances and so could/should be handled differently within a QoS-enabled site if they are not all simply classified as Premium IP traffic. These are discussed at greater length in Section 6 but on a broad level it is possible to classify these applications based on their tolerances for bandwidth availability, packet loss delay and jitter. As such, some classes (e.g. control traffic) can be identified as more critical and so may require a dedicated ‘expedited’ traffic class of their own.

The user requirements for QoS deployment in a site can also conceivably be applied by grouping users and applications based on the level of service they expect, or are allocated, from the network. On an abstract level we could classify a set of ‘gold’ users (staff for example) who are eligible to receive preferential treatment; a set of ‘silver’ users, such as students or employees, who will receive a normal Best Effort service; and ‘bronze’ users that will receive a Less than Best Effort service and which could be assigned to public or non-authenticated users. These classifications can then be mapped to specific QoS classes (on a per-application basis) which determine how sending hosts can mark their traffic.

5.4 Site Management Implications

The size of a site is likely to affect the network equipment deployed and how it is managed. This will also have an impact on how QoS support is deployed.

5.4.1 Large Campus

A large university campus has the potential to be multi-site and supported by a team of network engineers. There is probably a significant router infrastructure, with an overprovisioned core network.

Some campus departments may be large enough and have the skills to manage their own equipment, typically the computer science departments. Here, QoS deployment needs collaboration like any other service. The central computing service can choose to trust the QoS classifiers applied by the department, or may choose to reclassify traffic received from the department. Where specific QoS support is required to the port level, shared management of the layer 2 switch equipment may be required, e.g. where specific ports participate in a campus-managed VoIP deployment.

5.4.2 Small Site Specifics

In smaller sites, e.g. schools or small colleges, there may just be a single site, with just one or even no dedicated network engineers. While such a site is likely to be under single management, that management may be outsourced.

Access bandwidth in such sites is more likely to be limited and ‘smart’ border devices (e.g. single box traffic shapers and proxies) are more likely to be deployed. NAT is also more likely to be in use, with the complications for QoS that that brings.

5.5 QoS Implementation

In this section we look in more detail at QoS implementation in a site network.

Before proceeding, it is important to have considered the general considerations we outlined above, and in particular to know what is to be achieved (pervasive deployment, or deployment targeted at congestion points), to have reviewed the policies to be implemented and to have an inventory of the capabilities of the deployed equipment.

The general approach for a site is to apply QoS at layer 3 via DiffServ, just as described in the Regional Network section above. However, a site also may have a requirement to deliver QoS down to the physical port level in rooms, and thus complementary layer 2 implementations may be required, typically via IEEE 802.1p. Thus in this section we cover both DiffServ and 802.1p configuration.

5.5.1 Over-Provisioning

As discussed previously for JANET and Regional Networks, one may choose to overprovision parts of the network and elect not to deploy QoS methods on devices there. With trunked 1Gbit/s and now 10Gbit/s available, bandwidth in the core may be more than adequate.

Over-provisioning is discussed in Section 2.4 above. The most important aspect of this is to monitor the usage of links considered to be over-provisioned to gauge whether congestion does occur at any time, and if so to form a plan of action if required. In considering

overprovisioning capacity, it is possible you should tune network capacity to that of the Regional Network Operator (and, implicitly via the Regional Network Operator's relationship with JANET(UK), to the JANET core).

Congestion probability tends to increase towards the edges of networks, and that may be where to focus QoS deployment efforts.

5.5.2 Layer 2 QoS Configuration

Generally, classic QoS methods are considered at layer 3, e.g. DiffServ. However, in a site network one also must consider layer 2 because end nodes are rarely deployed one per subnet, with a dedicated router (though this has been done in some videoconferencing deployments at JANET sites, for example).

There are different methods available to help assure improved layer 2 QoS, e.g.:

- use of IEEE 802.1p, which prioritises traffic at layer 2 on a switch device (on IOS this can be done with the switchport priority command for an interface)
- IGMP/MLD snooping, which can prevent unwanted IPv4/IPv6 multicast traffic being seen by a device on a given switch port
- limiting subnet sizes, such that broadcast traffic is reduced.

Perhaps the most difficult layer 2 issues arise when handling wireless LANs. Here, many devices in effect share one medium. Some vendors do support QoS methods on their access points, but solutions are vendor-specific.

5.5.3 Layer 3 QoS Configuration

In this section we consider DiffServ and assume IntServ is very unlikely to be used. The details of a DiffServ deployment within a site will depend on where the support is being added, whether it is at points of congestion (typically the site border, or links to satellite campuses) or pervasively. In either case, configurations today are static rather than dynamic and while the ability to configure QoS on demand may emerge in the future, static provisioning and configuration is the current common practice.

There are some specifics to plan into the configuration, including:

- which DiffServ classes are required to be supported, and which DSCP values are used
- how and where packets are classified in the network, and (by policy) which applications or sources are given which class of service
- where admission control needs to be configured, such that the DSCP values of packets traversing the device are honoured
- provisioning for Premium IP, where used, i.e. the maximum percentage of elevated traffic allowed at given points on the network
- how and where to police the traffic, with particular consideration for excess Premium IP traffic, and what treatment out of profile traffic is given (i.e. remarked as BE or dropped)
  • how unsupported DiffServ values are treated. These could be allowed to pass transparently (unaltered) or could again be remarked to BE.

In the work done to date in the JANET QoS project, only three DiffServ classes have been

used, those being Premium IP (DSCP 46), Best Effort (BE) (DSCP 0) and Less than Best Effort (LBE) (DSCP 8). There has been discussion of IP Plus for certain applications but we would expect that sites deploying QoS today would just use Premium, BE and LBE. The exception may be where specific vendor solutions use certain DSCP values, perhaps for VoIP solutions. In this case, we advise consulting the recommended DSCP values for JANET [JANETDSCP]. These recommendations are made with a view to future QoS evolution on JANET such that sites can deploy DSCP marking with the best likelihood of future interoperability.

The common approach on JANET is to not remark DSCP values, to allow communicating sites to use DiffServ without capability being removed by the core. It is thus probable that this is a sensible site policy, but given that sites are not (generally) providing transit, a policy to remark certain DSCP values on entry/exit to a site network is unlikely to impact third parties.

As a site, one should seek to negotiate a QoS SLA where possible with your RNO, which will define the various QoS classes that can be used, and the treatment associated with those classes. The RNO would most likely police all traffic sent upstream, so it is important to understand how out of profile (e.g. excess Premium IP) traffic will be handled. It is not uncommon for excess traffic simply to be discarded.

In most cases, one would not expect to see more than 10-20% of capacity reserved for Premium IP, because the more an RNO offers to each connecting site, the worse its own worst case scenario of provisioning becomes. For LBE, the QoS project work suggested a lower bound of 5% on LBE traffic (such that LBE would always have a certain minimum percentage of all traffic).

In the general case, there are various capabilities that are needed in the routing equipment to facilitate QoS:

- marking or classifying traffic
- queuing based on DSCP values on a per hop bases
- policing traffic (and handling out of profile traffic).

Each site will need to decide where each function is required and how it is implemented. Common practice is to classify traffic as close to the source as possible, but this will depend on router capabilities. Additionally, one may choose to only configure DSCP handling at points of congestion. The site should also consider policing (and/or possibly reclassifying) traffic at the network borders, e.g. with an internal department that runs its own network, or more commonly at the upstream point of attachment.

On Cisco® IOS, there are various ways to mark traffic. Marking can be performed using Access Control Lists (ACLs), such that traffic matching a named ACL can be marked with a given DSCP value, or class-based marking can also be used. The specifics of implementing per hop DSCP handling will be vendor-specific. Here is an example of configuring Premium IP on IOS:

class-map EF

match ip dscp 46

!

policy-map TEST

class EF

bandwidth percent 99

!

interface GigabitEthernet0/1

service-policy output TEST

In IOS there is a single command that can be used to police traffic and take a given action based on the observed behaviour, e.g.

policy-map TEST

class premium-aggregate-1

police 1000000 10000 10000 conform-action transmit exceed-action drop

By dropping excess Premium IP traffic rather than remarking to be BE, it ensures that the Premium IP service either works as intended, or fails. To ship packets as BE that are believed to be handled as Premium IP by the source will only cause problems for the application users. It is usually better to reconsider the application usage, or to change the provisioning, than to pass Premium traffic on as BE.

5.5.4 Monitoring and Measurement

It is important to establish a network monitoring capability such that the site can assess where the potential bottlenecks are in the network, which flows may be particularly demanding, and whether the QoS implementation is having the desired effect.

The site may find that the upstream RNO has tools that can be used to view their perspective of the traffic. For example, JANET has supported the Netsight [Netsight] system for a number of years.

In the JANET QoS project, Cisco® IP-SLA tools were used, in particular the Service Assurance Agent (SAA). This gave a very detailed view of behaviour of different DSCP class traffic across the networks.

Locally, there are many tools available that should be considered, for example:

- rtg or mrtg for per-port usage (rtg appears to offer much higher performance than mrtg for this task)
- rude and crude for capacity and performance testing
- netflow to detect and inspect the highest individual or aggregate flows (this will need a netflow collector running on a system)
- multicast beacons (e.g. dbeacon) to view multicast performance.

One should also investigate available tools to catch reports of traffic policing actions, to determine where out of profile traffic may be being seen.

5.5.5 Policy

The policy enforced by a site will be important in determining how QoS will be deployed and supported. The aim of this policy is to act as a management tool to define how QoS resources are allocated within a site in an unambiguous manner and to act as a guideline for long-term deployment and usage. The QoS policy can also be influenced by external entities such as standards body recommendations (e.g. IETF), existing provider policy from the RNO/JANET, and de facto best practice. This policy will be specific to the site in question and may vary in purpose, granularity and details but, in summary, the QoS policy defined by the site should be as detailed as possible and should include: the traffic marking conventions (DSCP values) followed for application classes; the resource allocation on QoS enabled links (and their boundaries); and the user classifications.

As described above, the convention is to specify three DiffServ classes to represent aggregate traffic classed as shown in the table below:

Traffic Class	**DSCP**	**CoS**	**Applicatio**
Premium IP	46	5	VoIP, vide multimedia
Best Effort	0	0	Web traffic transfers
Less than Best Effort	8	1	Batch oper transfers

The classifications presented here are for illustrative purposes to demonstrate how this aspect of the QoS Policy could be structured. Moreover, certain classes of QoS application (such as control traffic) may not be included in the above list but still need to be represented in some way.

Based on this, the policy will define the admission/marking/forwarding behaviour for routers

within the site to determine how QoS-enabled traffic is handled. This policy should also define how incoming/outgoing traffic is handled at the boundary between peers or the upstream provider, be it another larger site or the Regional Network. In the remainder of this section we will focus on examining the issues related to creating these two key aspects of the policy.

5.5.6 Managed Bandwidth

Sites can be connected by a managed bandwidth connection provided by the JANET Lightpath service. This is a point-to-point connection which directly connects two subnets or just two end nodes of two organisations. These organisations might both be Janet-connected, or one of them may be a JANET-connected organisation while the other may be connected to some external education and research network. It is up to sites to determine how to use the connection provided as it is a raw bandwidth service and the provider does not control or monitor a usage of the connection. Sites should consider several aspects of the organisation's use of a managed bandwidth connection:

- What subnets or what end nodes/applications are allowed to use this connection? Will the subnet(s) or nodes that are authorised to use the connection be physically separated from other subnets of the sites or not? If the latter is the case, how will access control for this connection be organised to prevent unauthorised use of the connection?

The complete physical separation of computers that will use the managed bandwidth connection from other computers and servers of the site is the simplest case in terms of access organisation as no additional tools to prevent unauthorised access are needed. However, it means that users of the computers on the isolated subnet will have no access to other organisation's data resources (such as databases and file stores) and to the Internet. If it is not acceptable, a second internal connection for the isolated subnet is needed which should be firewalled to filter unauthorised access to the managed bandwidth connection from the outside of the isolated subnet. Of course, in such a case the isolated subnets are really only semi-isolated but they will be called 'isolated' in the following paragraphs for simplicity.

- At what layer will isolated subnets of the connected sites communicate: at layer 2 or layer 3? Both ways have some advantages and disadvantages. Layer 2 connectivity is simpler to organise as it doesn't require having routers within subnets, only Ethernet switches. However, in the layer 2 case two isolated subnets might become too dependent on each other as they should belong to the same IP network. For subnets that belong to different organisations this might be inconvenient as it requires the assignment of IP addresses to the new nodes in some coordinated way. Layer 3 connectivity gives sites' administrators more freedom in organising and managing their isolated or semi-isolated subnets. This is especially important in the case of semi-isolated subnets as coordinated tuning of remote firewalls on both ends of the distributed IP network may be complex.
- How much traffic is each node or application allowed to direct into the connection? Shall access control (traffic classification) and policing be deployed to enforce the limits on the amount of traffic allowed to use this connection?

This is a typical problem of the QoS area as it is about providing the appropriate balance between the level of connection utilisation and the required characteristics of traffic latency and loss. All considerations of section 2 and 3 of this Technical Guide are relevant to these problems. The difference here is that it is the site administrators (not Regional Networks

administrators) who have responsibility for this issue.
