Ontario

Government of Ontario

Information & Technology Standards

Government of Ontario IT Standard (GO-ITS)

GO-ITS Number 57.1

Network Quality of Service Standard

Version #: 1.0

Status: Approved

Prepared for the Information Technology Standards Council (ITSC) under the delegated authority of the Management Board of Cabinet

Last Review Date: 2010-09-16

Copyright & Disclaimer

Template Info

Template Name	Template #	Template Version No.	Template Author	Template Completion Date
GO-ITS Template	09.03.25	1.0	Design: PMCoE Boilerplate: TAB/OCCTO	2009-03-26

Document History (including ITSC and ARB approval dates)

Date	Summary
2009-08-24	**Created**: Network Quality of Service Standard, initial draft version 0.1
2010-08-18	**Endorsed**: IT Standards Council endorsement
2010-09-16	**Approved**: Architecture Review Board approval • Approved version number set to 1.0

Table of Contents

1. Foreword

Government of Ontario Information Technology Standards (GO-ITS) are the official publications on the guidelines, preferred practices, standards and technical reports adopted by the Information Technology Standards Council (ITSC) under delegated authority of the Management Board of Cabinet (MBC). These publications support the responsibilities of the Ministry of Government Services (MGS) for coordinating standardization of Information & Information Technology (I&IT) in the Government of Ontario. Publications that set new or revised standards provide enterprise architecture guidance, policy guidance and administrative information for their implementation. In particular, GO-ITS describe where the application of a standard is mandatory and specify any qualifications governing the implementation of standards.

2. Introduction

2.1. Background and Purpose

Quality of Service (QoS) is the process of marking flows of packets or traffic, classifying them for a class (level) of service, and then processing them based on those markings. The ultimate goal is to guarantee a certain level of performance for a data flow during times of network congestion. The level of services is based on a priority hierarchy, from high to low priority and data flows are classified based on which are most susceptible to jitter, delay, and packet loss while the network is congested.

Packets marked as in a high priority service category will get "preferential" treatment while travelling across the network while lower priority marked packets get less preference through the network. When a mixture of high and low priority marked packets encounter congestion on the network, QoS helps to manage the congestion by employing different queuing and scheduling features across network components.

As an example, packets not highly impacted by jitter, latency, and packet loss such as general web browsing would be placed in a lower priority queue while the network is congested. While a data flow that is highly impacted like voice or video calls would be prioritized in order to guarantee performance.

In recent years, networks and network based applications have evolved and a significant amount emphasis is being placed on performance management through the implementation of QoS. At the present time, the Ontario Government's network adequately supports hundreds of network applications without such services enabled. The key business driver for the need to standardize QoS within the Ontario Government is related to the development and deployment of future enterprise video, voice, and Unified Communications (UC) services which are accompanied by strict network performance requirements.

With the implementation of new network intensive services on the horizon, this standard shall define a set of guidelines, roles and responsibilities and a traffic prioritization scheme to guarantee performance during times of congestion.

This document is intended to define the requirements for service classes and design criteria when deploying QoS.

2.2. Scope

2.2.1. In Scope

- All internal data communications networks including, Data Centre infrastructure, Local Area Networks (LAN), and Wide Area Network (WAN)
- All data, voice, and video networks used within the OPS
- Unified Communications services including, call control, telephony, instant messaging, presence, speech recognition, unified messaging and video conferencing
- Network infrastructure devices including, routers, switches, and firewalls

2.2.2. Out of Scope

- Broader Public Sector (BPS) and other arms length organizations including, academic institutions, health care providers, major transfer payment recipients, municipalities, and school boards
- Infrastructure devices including, servers, storage, backup devices, and load balancers.
- Implementation of QoS

2.3. Objectives of GO-ITS 57.1

This standard outlines:

- Mandatory requirements and design criteria for deploying Quality of Service on the Government of Ontario Network.
- A defined, structured set of Network Traffic Queues

2.4. Network Objectives

The Ontario Government Network Traffic Queues ***(Appendix D)*** defines a framework for how current and future network traffic must be classified. The main focus of any QoS deployment should be to:

- Define a set of traffic prioritization rules for times of network congestion
- Ensure that time-sensitive and mission-critical applications have the network resources required, while allowing other applications to coexist and access the network.
- Increase application, network, and infrastructure efficiency, resiliency, and performance of network intensive services.
- Govern control over network resources and allows for effective management of the network from a business, rather than a technical perspective.
- Accommodate new applications and technologies (Unified Communications, VoIP, Video Conferencing)
- Support continued infrastructure growth

2.5. Applicability Statements

2.5.1. Organization

Government of Ontario IT Standards and Enterprise Solutions and Services apply (are mandatory) for use by all ministries/clusters and to all former Schedule I and IV provincial government agencies under their present classification (Advisory, Regulatory, Adjudicative, Operational Service, Operational Enterprise, Trust or Crown Foundation) according to the current agency classification system.

Additionally, this applies to any other new or existing agencies designated by Management Board of Cabinet as being subject to such publications, i.e. the GO-ITS publications and enterprise solutions and services - and particularly applies to Advisory, Regulatory, and Adjudicative Agencies (see also procurement link, OPS paragraph). Further included is any agency which, under the terms of its Memorandum of Understanding with its responsible Minister, is required to satisfy the mandatory requirements set out in any of the Management Board of Cabinet Directives (cf. Operational Service, Operational Enterprise, Trust, or Crown Foundation Agencies).

As new GO-IT standards are approved, they are deemed mandatory on a go-forward basis (Go-forward basis means at the next available project development or procurement opportunity).

When implementing or adopting any Government of Ontario IT standards or IT standards updates, ministries and I&IT Cluster must follow their organization's pre-approved policies and practices for ensuring that adequate change control, change management and risk mitigation mechanisms are in place and employed.

2.6. Requirements Levels

Within this document, certain wording conventions are followed. There are precise requirements and obligations associated with the following terms:

Must	This word, or the terms "REQUIRED" or "SHALL", means that the statement is an absolute requirement.
Should	This word, or the adjective "RECOMMENDED", means that there may exist valid reasons in particular circumstances to ignore the recommendation, but the full implications (e.g., business functionality, security, cost) must be understood and carefully weighed before choosing a different course.

2.7. Contact Information

2.7.1. Roles and Responsibilities

Accountable Role Definition

The individual ultimately accountable for the process of developing this standard. There must be exactly one accountable role identified. The accountable person also signs off as the initial approver of the proposed standard before it is submitted for formal approval to ITSC and ARB. (Note: in the OPS this role is at a CIO/Chief or other senior executive level)

Accountable Role:
Title: Director, Telecommunications Services Branch
Ministry/Cluster: Ministry of Government Services
Division: Infrastructure Technology Services

Responsible Role Definition
The organization responsible for the development of this standard, There may be more than one responsible organization identified if it is a partnership/joint effort. (Note: the responsible organization provides the resource(s) to develop the standard)

Responsible Organization:
Ministry/Cluster: Ministry of Government Services
Division: Infrastructure Technology Services
Branch: Telecommunications Services Branch

Support Role Definition
The support role is the resource(s) to whom the responsibility for actually completing the work and developing the standard has been assigned. There may be more than one support role identified. If there is more than one support role identified, the following contact information must be provided for each of them. If there is more than one support role, the first role identified should be that of the editor – the resource responsible for coordinating the overall effort.

Support Role (Editor):
Ministry/Cluster: Ministry of Government Services
Division: Infrastructure Technology Services
Branch: Telecommunications Services Branch
Section: Technical Services
Job Title: Senior Manager, Technical Services
Name: Michael Hipwell

Phone: 416-325-1869
Email: Michael.Hipwell@ontario.ca

The above individual will be contacted by the Standards Section once a year, or as required, to discuss and determine potential changes and/or updates to the standard (including version upgrades and/or whether the standard is still relevant and current).

2nd Support Role (if applicable):
Section: Enterprise Strategic Planning, Services & Architecture Section
Job Title: Manager, Enterprise Architecture Office
Name: Christopher Chua
Phone: 416-327-4415
Email: christopher.chua@ontario.ca

3rd Support Role (if applicable):
Section: Enterprise Strategic Planning, Service & Architecture Section
Job Title: Technical Architect
Name: David Lin
Phone: 416-212-3133
Email: david.lin@ontario.ca

Consulted

Please indicate who was consulted as part of the development of this standard. Include individuals (by role and organization) and committees, councils and/or working groups. (Note: consulted means those whose opinions are sought, generally characterized by two-way communications such as workshops):

Organization Consulted (Ministry/Cluster)	Division	Branch	Date
Ministry of Government Services	Infrastructure Technology Services	Data Centre Operations	March/May 2009/2010
Ministry of Government Services	Infrastructure Technology Services	ESPDS Branch	March/May 2009/2010
Ministry of Government Services	Infrastructure Technology Services	Guelph Data Centre Project Team	May 2010
Ministry of Government Services	Corporate Security Branch		March/May 2009
Ministry of Government Services	OCCTO	Corporate Architecture Branch	March/May 2009
Ministry of Government Services	OCCIO	Corporate Architecture and Standards Branch	March/May 2009

Committee/Working Group Consulted	Date
Technology Architecture Domain Working Group	May 25, 2010
Security Architecture Domain Working Group	June 9, 2010
ITS Architecture Core Team	June 11, 2010
Solutions Delivery Leadership Committee (SDLC)	July 8, 2010
Information Technology Standards Council	July 21, 2010
IT Service Management Leads Forum	July 22, 2010

2.8. Recommended Versioning and/or Change Management

Changes (i.e. all revisions, updates, versioning) to the standard require authorization from the "responsible" organization.

Once a determination has been made by the responsible organization to proceed with changes, the Standards Section, Technology Adoption Branch, OCCTO, will coordinate and provide assistance with respect to the approvals process.

The approval process for changes to standards will be determined based on the degree and impact of the change. The degree and impact of changes fall into one of two categories:

> ***Minor changes*** *- requiring communication to stakeholders. No presentations required. No ITSC or ARB approvals required. Changes are noted in the "Document History" section of the standard;*
>
> ***Major changes*** *- requiring a presentation to ITSC for approval and ARB for approval (Note: ARB reserves the right to delegate their approval to ITSC)*

Below are guidelines for differentiating between minor and major changes:

Major:

- represents a major version change to one or more specifications
- impacts procurement
- requires configuration changes to current solutions
- impacts other standards
- responds to legislative, policy or procurement changes

Minor:

- represents incremental version changes to one or more specifications
- does not impact procurement (other than informational)
- does not require configuration changes to current solutions
- does not impact other standards
- is not related to legislative, policy, or procurement changes

2.9. Publication Details

All approved Government of Ontario IT Standards (GO-ITS) are published on the ITSC Intranet web site. Please indicate with a checkmark below if this standard is also to be published on the public, GO-ITS Internet Site.

Standard to be published on both the OPS Intranet and the GO-ITS Internet web site (available to the public, vendors, etc.)	☒

3. Compliance Requirements

The ITS and the network vendor(s) must comply with the QoS Service Classes as defined in Section 4.3 of this document. Strict enforcement of these Service Classes is essential in order to guarantee service levels and meet contractual commitments for time-sensitive services.

Technical specifications and version control information will be published on an on-going basis as required.

4. Technical Specification

This section presents the technical definition for QoS Internet Standards Based Architecture, key service concepts for IntServ and DiffServ, mandatory requirements, architecture implementation guidelines and traffic marking rules.

4.1. QoS Based Architectures

This QoS standard covers the definitions of QoS, traffic and service specifications, IntServ and DiffServ frameworks.

4.1.1. Definition of QoS

QoS is defined as service differentiation and performance assurance for data flows running over the LAN or WAN. Service differentiation provides different services to different applications according to their requirements. Performance assurance addresses bandwidth, loss, delay and delay variation. The specifications of traffic and its desired service can be given on a per-flow basis or in service level agreement (SLA) from the perspective of packet forwarding treatment an application receives.

There are two major frameworks, Integrated Services (IntServ) and Differentiated services (DiffServ) recognized as the principal architectures for providing QoS.

4.1.2. Integrated Services (IntServ)

An Integrated Services (IntServ) approach utilizes Resource Reservation Protocol (RSVP) to reserve capacity and capability across a network; this makes it ideal for end-to-end delay and jitter-sensitive applications, such as Voice and Video. With RSVP capacity is reserved and prioritized across the network on a per flow basis.

This type of prioritization, because of the reservation of resources, does not scale well in large network environment with a multitude of applications contending for a finite amount of resources. Although RSVP may be ideal in small well behaved networks, in most large networks today, it is impractical to implement RSVP as a QoS mechanism due to the number of flows and reservations that would have to be managed and the amount of bandwidth and CPU cycles that would be required to maintain the environment.

4.1.3. Differentiated Services (DiffServ)

Differentiated Services (Diffserv) utilizes a different approach to prioritize traffic. Instead of looking at individual flows, Diffserv marks each individual packet. By marking groups of packets from applications with similar requirements with the same identifiers, the network elements can recognize the service levels required by each packet and ensure the proper treatment across the network. Network prioritization is assessed on a hop by hop (Per Hop Behaviour) basis and each node is responsible to manage congestion based on the Diffserv priorities.

In general, Diffserv implementation follows the following steps.

- Each Packet is categorized and a series of bits are set in a field in the IP header upon network entry or at a network boundary using legacy IP precedence or the more scalable Differentiated Services Code Points (DSCP).
- Using this field's bit marking the nodes inside the network queues and forwards the packets based on the markings.
- In the event of congestion, these bit markings are used to decide what traffic will be prioritized and what traffic will be dropped.
- Individual packets are marked for a certain class of service and are handled consistently at any point in the network to ensure a consistent level of service. As an example, HTTP packets will be marked for a certain Per Hop Behaviour (PHB) versus a voice packet which will have a "higher" marking PHB and will be handled at a higher level of service. As apposed to IntServ Flows, DiffServ is granular at the packet level and classifications can be coarse to fine as needed. With DiffServ the bits called Differentiated Services Code Points or (DSCP) can be scaled to 64 different markings. This means a packet may be marked in 64 different ways, although the vast majority of networks only implement 3-8 different priorities.
- The major advantage of this approach is scalability and performance since the packets are only examined for classification and marking at the edge of the network and then the traffic is prioritized based on the DSCP markings.

4.2. DiffServ Key Service Concepts (Reference RFC 4594)

Differentiated Services is a general architecture that may be used to implement a variety of network intensive services (e.g., voice and video). Three fundamental forwarding behaviours have been defined and characterized for general use. These are basic Default Forwarding (DF) behaviour for elastic traffic, the Assured Forwarding (AF) behaviour, and the Expedited Forwarding (EF) behaviour for real-time (inelastic) traffic. The fact that four code points are recommended for AF and that one code point is recommended for EF are arbitrary choices, and the architecture allows any reasonable number of AF and EF classes simultaneously. The choice of four AF classes and one EF class in the current document is also arbitrary, and network operators MAY choose to operate more or fewer of either.

4.2.1. Assured Forwarding (AF)

Assured Forwarding behaviour is explicitly modeled on Frame Relay's Discard Eligible (DE) flag or ATM's Cell Loss Priority (CLP) capability. It is intended for networks that offer average-rate Service Level Agreements (SLAs) (as FR and ATM networks do). This is an enhanced best-effort service; traffic is expected to be "elastic" in nature. The receiver will detect loss or variation in delay in the network and provide feedback such that the sender adjusts its transmission rate to approximate available capacity as with most TCP based data applications.

For such behaviours, multiple DSCP values are provided to identify the traffic; a common queue is used to store the aggregate and active queue management is used to protect the network and to limit delays. Traffic is metered and marked as it enters the network depending on the arrival rate of the aggregate. The premise is that it is normal for users occasionally to use more capacity than their contract stipulates (burst), perhaps up to some bound. However, if traffic should be marked or lost to manage the queue, this excess traffic will be marked or lost first

4.2.2. Expedited Forwarding (EF)

The intent of Expedited Forwarding Per-hop Behaviour (PHB)[1] is to provide a building block for low-loss, low-delay, and low-jitter services. It may be used to build an enhanced best-effort service: traffic remains subject to loss due to line errors and reordering during routing changes. However, using queuing techniques, the probability of delay or variation in delay is minimized. For this reason, it is generally used to carry voice and for transport of data information that requires "wire like" behaviour through the IP network. Voice is an inelastic "real-time" application that sends packets at the rate the codec produces them, regardless of availability of capacity. As such, this service has the potential to disrupt or congest a network if not controlled.

To protect the network from abuse, at minimum traffic should be policed at various points to ensure that the design of a queue is not overrun, and then the traffic should be given a low-delay queue (often using priority, although it is asserted that a rate-based queue can do this) to ensure that variation in delay is not an issue, to meet application needs.

4.2.3. Class Selector (CS)

The Class Selector provides support for historical code point definitions and PHB requirements. The Class Selector DiffServ field provides a limited backward compatibility with legacy (pre DiffServ) practice, as described in *RFC2474, Section 4*. Backward compatibility is addressed in two ways. First, there are per-hop behaviours that are already in widespread use (e.g., those satisfying the IPv4 precedence queuing requirements specified in *RFC1812*, and we wish to permit their continued use in DS-compliant networks. In addition, there are some codepoints that correspond to historical use of the IP Precedence field, and we reserve these codepoints to map to PHBs that meet the general requirements specified in *RFC2474.*

A DiffServ-compliant network can be deployed with a set of one or more Class Selector-compliant PHB groups. Also, a network administrator may configure the network nodes to map codepoints to PHBs, irrespective of bits 3-5 of the DSCP field, to yield a network that is compatible with historical IP Precedence use. Thus, for example, codepoint '011000' would map to the same PHB as codepoint '011010'.

In the majority of implementations, CS and AF are used simply as identifiers with each implementation applying its own data priority and queuing mechanism to control data prioritization and queuing behaviour.

4.3. Mandatory Requirements

4.3.1. Differentiated Services QoS Implementation

Due to the scalability and the adoption rate of Diffserv by Canadian Carriers for MPLS WAN implementations, a Diffserv approach will be followed for QoS implementations. The current guidelines for Diffserv implementation are detailed in *RFC 4594*.

4.3.1.1. Service Classes

Traffic flowing through a network may be classified in many different ways. For the purposes of this standard, service classes will be divided into the following two groupings:

4.3.1.1.1. Network Control Group

The network control traffic group is further divided into two service classes:

[1] RFC3246 - http://tools.ietf.org/html/rfc3246

1. **Network Control Service Class (EF) -** for routing and network control functions.
2. **Operations, Administration, and Management (OAM) Service Class (AF31) -** for configuration and management functions.

4.3.1.1.2. User/Subscriber Traffic Group

The user/subscriber traffic group is broken down into ten service classes to provide service differentiation for all the different types of applications/services:

1. **Telephony Service Class (EF)** - is best suited for applications that require very low delay variation and are of constant rate, such as IP telephony (VoIP) and circuit emulation over IP applications.
2. **Signalling Service Class (CS4/AF42)** - is best suited for peer-to-peer and client-server signalling and control functions using protocols such as SIP, SIP-T, H.323, H.248, and Media Gateway Control Protocol (MGCP).
3. **Multimedia Conferencing Service Class (AF41) -** is best suited for applications that require very low delay and have the ability to change encoding rate (rate adaptive), such as H.323/V2 and later video conferencing service.
4. **Real-Time Interactive Service Class (AF41)** - is intended for interactive variable rate inelastic applications that require low jitter and loss and very low delay, such as video conferencing applications that do not have the ability to change encoding rates or to mark packets with different importance indications.
5. **Multimedia Streaming Service Class (CS4/AF42)** - is best suited for variable rate elastic streaming media applications where a human is waiting for output and where the application has the capability to react to packet loss by reducing its transmission rate, such as streaming video and audio and webcast.
6. **Broadcast Video Service Class (CS3)** - is best suited for inelastic streaming media applications that may be of constant or variable rate, requiring low jitter and very low packet loss, such as broadcast TV and live events, video surveillance, and security.
7. **Low-Latency Data Service Class (AF31)** - is best suited for data processing applications where a human is waiting for output, such as web-based ordering or an Enterprise Resource Planning (ERP) application.
8. **High-Throughput Data Service Class (AF21)** - is best suited for store and forward applications such as file transfer protocol (FTP) and billing record transfer.
9. **Standard Service Class (CS0)** - is for traffic that has not been identified as requiring differentiated treatment and is normally referred to as best effort.
10. **Low-Priority Data Service Class (CS1)** - is intended for packet flows where bandwidth assurance is not required.

Note: It is understood that due to equipment limitations and the design requirements of specific QoS implementations, the service classes above may not be applicable and may be combined with other classes by the OPS or the network vendor(s).

See **Appendix B**, for a comparison table of each Service Class' tolerance to delay, jitter, and loss.

See **Appendix C**, for Differentiated Service Code Point (DSCP) Markings and Service Classes

4.4. QoS Design Criteria

As defined earlier there are a significant number of service classes and differentiators related to implementing QoS and network intensive services. At the present time, within the industry the majority of organizations have yet to differentiate application traffic to this level of granularity. This is typically due to the fact that organizations have not yet deployed or do not yet have a clear business requirement for implementing all service classes and differentiators.

For example if an organization is not deploying Voice over IP (VoIP) or video conferencing there may not be a requirement to deploy or allocate these service classes. However, if there are plans to use that technology in the future, it would be wise to reserve these service classes for a future implementation and not allocate them to applications un-necessarily.

As the Ontario Government network evolves, this standard may be revised from time to time to as business requirements dictate changes to the defined service classes.

See **Appendix D** for the Ontario Government Network Traffic Queues

4.4.1. QoS Principles

The following principles have been developed and are mandatory for design and administration of data communication services which may require QoS.

Principle #1: General QoS Design

Do not enable QoS features simply because they exist. Instead, start to design from a high level and clearly define the organization objectives.

- **Rationale**
  - Efficiency in controlling network characteristics and assessing traffic data
  - Accuracy of use of classes of service
  - Ability to measure and justify application requirements to services enabled
- **Implications**
  - The OPS Network (subsequently referred to as « GONET ») should be run as a centralized service to ensure the QoS enabled service is as one centralized service

Principle #2: Classification and Marking

Classify and mark applications as close to their sources as technically and administratively feasible and use DSCP markings whenever possible

- **Rationale**
  - Promote end-to-end Differentiated Services and per-hop behaviours (PHBs)
  - Avoid allowing users to mark their own traffic.
  - Promote DSCP markings whenever possible to ensure interoperability and future expansion.
- **Implications**
  - Proper analysis and design is needed to ensure accurate markings

Principle #3: Policing and Markdown

Police traffic flows as close to their sources as possible.

- **Rationale**
  - This principle applies to legitimate flows also because denial-of-service (DoS) attack and worm-generated traffic might be masquerading under legitimate, well-known TCP and UDP ports, causing extreme amounts of traffic to be poured onto the network infrastructure.
  - Whenever supported, markdown should be done according to open standards-based rules, such as *RFC 2597* ("Assured Forwarding PHB Group").

- **Implications**
  - Congestion-management policies, such as DSCP-based WRED, should be configured

4.5. Traffic Marking Rules

The following must be used to correctly correlate traffic to the appropriate classes.

4.5.1. LAN EF

- All traffic in the IP range that matches the Ports for Voice bearer traffic.
- Any routing protocol traffic (OSPF, BGP or RIP)

4.5.2. LAN AF41

- All traffic in the IP range, this should be considered urgent service and used for the new Video Conferencing Service
- Any Internal Government address space that matches the Ports used for Video in the current implementation
- Any special application that would require time sensitive treatment.

4.5.3. LAN CS4 or AF42

- System Network Architecture (SNA) Data-link switching (DLSw) or Telnet3270 Traffic
- All other IP ranges for Call Setup
- Any internal traffic with ports matching WEBCast type traffic.

4.5.4. LAN AF21

- Any traffic to and from Exchange servers.
- Any Priority Bulk Transfers (Specifically identified)

4.5.5. LAN CS0

- Any HP Client Configuration Management (Radia) traffic
- Any HP Software Server Automation traffic
- Any SMS traffic

4.5.6. LAN CS1

- Any Traffic to or from a non Internal IP Address (Web or other protocols)

4.5.7. LAN AF31 (default) *

- All other Data (Default Class)
  - Data Base
  - Office applications
  - OPS Intranets
  - OAM
  - Desktop Video Conferencing
  - Unclassified Data

For sites subscribing to a WAN service, but without a LAN, classification and marking will be dealt with on a case by case basis by the Network Architecture and Security Committee. In order to mitigate potential risks, applications that would normally fall into a particular category but are found to behave poorly and have an adverse effect on the overall network may be re-prioritized into a lower service class to protect the general performance of the network. Before being introduced into the network, new applications should be assessed for QoS requirements as part of Architecture Review or by consultation with the ITS Telecommunications Services Network Team.

5. Related Standards

5.1. Impacts to Existing Standards

Identify any Standards that reference or are referenced by this Standard and describe the impact.

GO-IT Standard	Impact	Recommended Action
*No impact to existing standards		

5.2. Impacts to Existing Environment

Impacted Infrastructure	Impact	Recommended Action
*No impact to existing environment at present	When Network QoS is enabled	Develop QoS Implementation Plan and follow this standard

6. Appendix A: References

RFC 2474: Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers

K. Nichols, S. Blake, F. Baker, D. Black. "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers". Dec. 1998

Description: This document defines the IP header field, called the DS (for differentiated services) field. In Ipv4, it defines the layout of the TOS octet; in Ipv6, the Traffic Class octet. In addition, a base set of packet forwarding treatments, or per-hop behaviors, is defined.

Remote Link: http://tools.ietf.org/html/rfc2474

RFC 1812: Requirements for IP Version 4 Routers

F. Baker, "Requirements for IP Version 4 Routers". June 1995

Description: This document defines and discusses requirements for devices that perform the network layer forwarding function of the Internet protocol suite. The Internet community usually refers to such devices as IP routers or simply routers; The OSI community refers to such devices as intermediate systems. Many older Internet documents refer to these devices as gateways, a name which more recently has largely passed out of favor to avoid confusion with application gateways.

Remote Link: http://tools.ietf.org/html/rfc1812

RFC 3246: An Expedited Forwarding PHB (Per-Hop Behavior)

B. Davie, A. Charny, J.C.R. Bennett, K. Benson, J.Y. Le Boudec, W. Courtney, S. Davari, V. Firoiu, D. Stiliadis. "An Expedited Forwarding PHB (Per-Hop Behavior)". March 2002

Description: This document defines a PHB (per-hop behavior) called Expedited Forwarding (EF). The PHB is a basic building block in the Differentiated Services architecture. EF is intended to provide a building block for low delay, low jitter and low loss services by ensuring that the EF aggregate is served at a certain configured rate. This document obsoletes RFC 2598.

Remote Link: http://tools.ietf.org/html/rfc3246

RFC 4594 - Configuration Guidelines for DiffServ Service Classes

J. Babiarz, F. Baker, K. Chan**. "**Configuration Guidelines for DiffServ Service Classes" August 2006

Description: This document describes service classes configured with Diffserv and recommends how they can be used and how to construct them using Differentiated Services Code Points (DSCPs), traffic conditioners, Per-Hop Behaviors (PHBs), and Active Queue Management (AQM) mechanisms. There is no intrinsic requirement that particular DSCPs, traffic conditioners, PHBs, and AQM be used for a certain service class, but as a policy and for interoperability it is useful to apply them consistently.

Remote Link: http://www.ietf.org/mail-archive/web/tsvwg/current/msg06771.html

Cisco Enterprise QoS Solution Reference Network Design Guide, Nov 2005

Description: This document provides design considerations and guidelines for implementing Cisco Quality of Service within an enterprise environment.

Remote Link: http://www.cisco.com/univercd/cc/td/doc/solution/esm/qossrnd.pdf

7. Appendix B: Network Service Class Comparison Table

	Service Class Name	Traffic Characteristics	Tolerance to:		
			Loss	Delay	Jitter
Network Control Group	Network Control	Variable Size packets, Mostly Inelastic short messages, some bursting (Routing Protocols)	Low	Low	High
	OAM	Variable Size Packets, Elastic and Inelastic Flows	Low	Med – High	High
User/Subscriber Traffic Group	Telephony	Fixed size small packets, constant rate, inelastic	Very Low	Very Low	Very Low
	Signalling	Variable Size packets, bursty short lived flow	Low	Low	High
	Multimedia / Video Conferencing	Variable size packers, constant transmit rate, adaptive, can adapt to loss	Low – med	Very Low	Low
	Real-Time Interactive	RTP/UDP streams, inelastic, variable rate	Low	Very Low	Low
	Multimedia Streaming	Variable size packets, Elastic with variable rate	Low – Med	Med	High
	Broadcast Video	Constant and variable rate, inelastic, non-bursty	Very low	Med	Low
	Low Latency Data	Variable rate, bursty short lived elastic flows	Low	Low – Med	High
	High-Throughput Data	Variable rate, bursty Long lived elastic flows	Low	Med – High	High
	Standard	A bit of everything	Unspecified		
	Low Priority Data	Non Real Time and elastic	High	High	High

A "High" in the jitter-tolerant column implies that data is buffered in the endpoint and that a moderate level of network-induced variation in delay will not affect the application. Applications that use TCP as a transport are generally good examples. Routing protocols and peer-to-peer signaling also fall in this class; although loss can create problems in setting up calls, a moderate level of jitter merely makes call placement a little less predictable in duration.

8. Appendix C: Differentiated Services Code Point Markings & Service Classes

Service Class Name	DSCP Name	DSCP Value	Application Examples
Network Control	CS6	110000	Network Routing Protocols
Telephony	EF	101110	IP Telephony Bearer
Signalling	CS5	101110	IP Telephony Signalling
Multimedia / Video Conferencing	AF41 AF42 AF43	100010 100100 100110	H.323 video Conferencing (Adaptive)
Real-Time Interactive	CS4	100000	Video Conferencing and Interactive applications (Gaming, SNA etc)
Multimedia Streaming	AF31 AF32 AF33	011010 011100 011110	Streaming Video and Audio on Demand
Broadcast Video	CS3	011000	Broadcast TV & live events
Low Latency Data	AF21 AF22 AF23	010010 010100 010110	Client Server transactions, WEB based Applications
OAM	CS2	010000	OAM&P
High Throughput Data	AF11 AF12 AF13	001010 001100 001110	Store and Forward Type applications
Standard	DF (CS0)	000000	Undifferentiated Applications
Low-Priority Data	CS1	001000	Any Flow that has no BW assurance (Scavenger)

(Reference RFC 4594)

9. Appendix D: Ontario Government Network Traffic Queues

Data Type	LAN Queue	WAN QOS	Bandwidth
Priority Data (EF) ○ VoIP (Bearer Traffic) ○ Routing Protocols*	Strict Priority Queue (EF)	Strict Priority (EF)	As Required for Number of VoIP calls (No more than 33% of total Bandwidth)
Interactive Video (AF41) ○ Video Conferencing ○ Telepresence ○ Real Time applications	Priority Data (Threshold 1) 80% of Queue	Priority Data (AF 3)	As Required for types and number of Video + 10 - 40% of Non EF Bandwidth
Streaming Video (CS4, AF42) ○ Online training ○ WEB Cast ○ SNA ○ Call Setup	Priority Data (Threshold 2) 20% of Queue		
Normal Data (Default) (AF31)[2] ○ Data Base ○ Office applications ○ OPS Intranets ○ OAM ○ Desktop Video Conferencing ○ Unclassified Data	Normal Data (Threshold 1) 80% of Queue	Normal Data (AF2)	55-80% of Non EF Bandwidth Depending on amount of AF3
Priority transfer (AF21) ○ E-Mail ○ Other Bulk Data Transfer	Normal Data (Threshold 2) 20% of Queue		
File Transfer (DF or CS0) ○ HP Client Configuration Management (Radia) traffic ○ HP Software Server Automation traffic ○ SMS ○ Readily identifiable, low priority Data Backups	Low Priority Data (Threshold 1) 80% of Queue	Low Priority Data (AF1)	5% of Non EF Bandwidth
Low priority Data (CS1) ○ Readily identifiable, low priority Internet WEB Browsing ○ Readily identifiable, low priority access to Non-Ontario Government address spaces.	Low Priority Data (Threshold 2) 20% of Queue		

Note: The Managed Network Service Providers must provide mechanisms for ensuring agreed upon bandwidth limits are not exceeded during normal operations.

[2] Applications that would normally fall into this category but are found to be poorly behaved and have an adverse effect on overall network performance may get re-prioritized into a lower service class to protect the overall performance of the network.