Regis University
College for Professional Studies Graduate Programs
Final Project/Thesis

Disclaimer

Revision Date 09/30/2010

INVESTIGATING BASIC QUALITY OF SERVICE DESIGN POSSIBILITIES FOR REGIS UNIVERSITY ACADEMIC RESEARCH NETWORK EDGE ROUTERS

A THESIS

SUBMITTED ON THE 23RD OF SEPTEMBER, 2010

TO THE DEPARTMENT OF COMPUTER SCIENCE

OF THE SCHOOL OF COMPUTER & INFORMATION SCIENCES

OF REGIS UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF MASTER OF SCIENCE IN SYSTEM ENGINEERING

BY

Robert Zwick

APPROVALS

Paul Vieira, Thesis Advisor

D. M. Likarish

Ranked Faculty Name

Ranked Faculty Name

Abstract

The Regis University Academic Research Network (ARNe) had network resources, such as VoIP, that required preservation their ability to receive near real-time forwarding treatment across the network. Quality of Service (QoS) design ideas were examined from four actual implementations described in research cases. Additionally, research involving surveys from Cisco certified professionals was examined, and Cisco technical literature was examined. Case study methodology, involving the study of multiple cases, was the primary tactic utilized in this research. Examination and triangulation of data from the research indicated that ARNe would benefit from moving forward with a basic QoS design and implementation, integrating concepts identified in the data. Additionally, data supported that a basic QoS design and implementation on ARNe would provide Computer Science and Information Science students an opportunity to more fully appreciate QoS through further research and hands-on experience.

Acknowledgements

I would like to express my appreciation to the Cisco certified professionals who helped me to understand QoS in networks. Without their help, the results of this research effort would have been substantially diminished in value.

I would like to thank Regis University alumni Robert Moon; Robert helped extensively through discussions that put QoS into perspective for ARNe.

I would like to thank Paul Vieira, my thesis advisor, for reviewing, and then making good, constructive criticisms during the thesis process.

I am grateful to my wife, Pam, who supported me and tolerated time spent apart during my hours of writing and rewriting throughout this research.

Table of Contents

List of Figures

List of Tables

Chapter 1 – Introduction

Thesis

Investigating basic Quality of Service (QoS) design possibilities for Regis University edge routers will illuminate configuration possibilities for allowing high priority data, such as Voice over IP (VoIP), to flow through edge routers before all lower priority data, and will provide a foundation for Regis University students to later design, implement and fine tune a very basic QoS configuration on the Regis University Academic Research Network (ARNe).

Problem Definition

ARNe has enabled students and faculty to access several types of applications and data resources in support of a variety of academic research projects and other educational needs. Each instance of a student or faculty member accessing an ARNe network resource requires the allocation of a portion of the network's limited bandwidth. The continually expanding and changing environment of ARNe, such as its VoIP implementation at the DTC and ILB campuses, illustrates the need to ensure that data which relies on near-real time processing, such as VoIP and videoconferencing, receives priority processing through the network. A second problem that this paper addresses is that Regis University Computer Science and Information Science students need education in, and direct experience with, the design and implementation of QoS in a network.

Relevance

Investigating possibilities for a basic QoS design on ARNe is relevant in two ways. First, by collecting and analyzing data through this research, it will aid in, at minimum, in creating a prophylactic QoS design. When the QoS design is implemented, it will assure that identified high priority data with low latency requirements, such as VoIP, receives preferential treatment.

Lower priority data, such as File Transfer Protocol requests, would be processed through the network on a lower priority basis. The relevance of QoS was further explained when Cisco stated that “In order to manage the multitude of applications such as streaming video, Voice over IP (VoIP), e-commerce, Enterprise Resource Planning (ERP), and others, a network requires Quality of Service (QoS) in addition to best-effort service. Different applications have varying needs for delay, delay variation (jitter), bandwidth, packet loss, and availability... The IP network should be designed to provide the requisite QoS to applications.”(Cisco, 2005) The second area of relevance for investigating a QoS design for ARNe involves the educational experience provided for Regis University students.

By providing research data supporting a minimal QoS design on ARNe edge routers, future students will better understand QoS methods that may be used in creating a baseline implementation of QoS for current ARNe requirements; and that baseline implementation will, in turn, facilitate future students to further study, research, and implement additional QoS mechanisms. As Hartpence stated,

> Quality of Service (QoS) techniques and protocols can be a critical part of network operations. Whether attempting to define traffic patterns, improve the performance of real time applications such as VoIP or even as part of a security strategy, QoS can be central to a network implementation. QoS topics include but are not limited to MPLS, DiffServ, IntServ, policies, aggregation, traffic shaping, queuing and good network design. Due to the wide range of concepts and their impact, educational programs focusing on communication should include quality of service mechanisms as topic for all students before they graduate. Courses need not include all of the advanced QoS ideas, but should provide the student with the knowledge and skills necessary to make good decisions when deploying topologies that

> must serve the needs of contemporary users. In addition to getting the lab experiences in their program, students should also have an opportunity to apply their skills in a practical and where possible, an interesting, and maybe even entertaining, environment. (Hartpence, 2009)

Assumptions, Constraints, and Risks

It is assumed that the readers of this document have intermediate-level knowledge of computer networking techniques and terms. The scope of the thesis was specifically restricted to identifying, through research, minimalistic QoS design possibilities for VoIP services at ARNe edge routers. While QoS for VoIP is the focus, the basic designs investigated will provide data which could be applied just as easily to QoS for other data types, such as streaming video and video conferencing. A complete QoS design for a network typically also involves prioritizing other services and involves additional QoS configurations on core routers and switches; these aspects will be mentioned, but not thoroughly researched or analyzed. Beyond the scope of this thesis, it is assumed that any individual or student who later uses this data in creating a QoS design will obtain proper permission from ARNe stakeholders before changing ARNe configurations. Any network changes should involve creating a test plan and a restoral plan to mitigate risks. One example of a risk mitigation technique is to first create a backup of the startup configuration of any router or switch in ARNe before making changes; in the event that a QoS implementation erroneously causes a problem, the original configurations would be restored from the removable storage media device containing the backup configurations.

Chapter 2 – Review of Literature and Research

Introduction

An exploration of available literature and research finds that there are different QoS mechanisms and implementation methods available for use in networks of different sizes. It became quickly evident that Cisco, a major manufacturer of networking equipment, generates extensive, detailed technical literature that describes networking principles, techniques and configuration guides, much of which is relative to QoS. Given that the routers in ARNe are Cisco products, and that Cisco provides extensive technical information specific to QoS and its implementation within its products, much of the technical literature referenced in this paper is from Cisco sources.

With respect to research sources, there is a variety of data presented on this paper from both scholarly sources and from surveys that exclusively included four experienced networking professionals that hold Cisco certifications. Two of the Cisco certified professionals held the Certified Network Associate (CCNA) credential, one held the Cisco Certified Network Professional (CCNP) credential, and one held the Cisco Certified Internetwork Expert (CCIE) credential. The survey information provided by these Cisco professionals, collected from September 2009 – October 2010, is given the same weight in research value as other sources, since ARNe specifically utilizes Cisco routers at its DTC and ILB campuses. Before taking a look at QoS models, we will briefly discuss a realistic limitation impacting QoS.

QoS and the Internet

When considering QoS, we need to take into account that the majority of networks interface with Worldwide Web resources, and interface with each other, via the Internet. For the most part, the Internet was designed to provide “best effort” connectivity, throughput, and delivery. This means that, by default, data is processed on a First In First Out (FIFO) manner and buffered equally at each Internet router. There are instances where Internet Service Providers (ISP) provide increased throughput and service guarantees honoring customer site QoS identification, but that comes at higher monetary cost and that higher service level can only be maintained in sections of the Internet that the ISP has direct control over. Research by Clark and Fang stated

> The current Internet assumes the ‘best-effort’ service model. In this model the network allocates bandwidth among all of the instantaneous users as best it can and attempts to serve all of them without making any explicit commitment as to rate or any other service quality. When congestion occurs, the sources of traffic are expected to detect this and slow down, so that they achieve a collective sending rate equal to the capacity of the congestion point. (Clark & Fang, 1998)

This information about the Internet is brought to the reader’s attention in order that unrealistic expectations are not developed when considering QoS implementations. The bottom line is that, for the most part, it is typically only economically feasible to implement and expect QoS be maintained within the network where the data is created, and, in some cases, in other external network components where the data originator’s QoS desires are honored. These points made, let us look at the prevalent basic QoS models and architectural considerations.

QoS Models

There are two basic QoS models for IP networks defined by the Internet Engineering Task Force (IETF). These models include the Integrated Service Model (IntServ), and the Differentiated Service Model (DiffServ). IntServ was introduced as RFC 1633 by the IETF Network Working Group of Braden, Clark, & Shenker (1994), and DiffServ was introduced as

RFC 2474 by the IETF Network Working Group of Nichols, Blake, Baker, & Black (1998). For an extensive explanation of these models, along with the most recent changes and available extensions, refer to the IETF website for thorough descriptions. The technical descriptions of these models are summarized when Cisco stated

> To facilitate true end-to-end QoS on an IP-network, the Internet Engineering Task Force (IETF) has defined two models: Integrated Services (IntServ) and Differentiated Services (DiffServ). IntServ follows the signaled-QoS model, where the end-hosts signal their QoS needs to the network, while DiffServ works on the provisioned-QoS model, where network elements are set up to service multiple classes of traffic with varying QoS requirements. (Cisco, 2005)

All four of the Cisco certified professionals surveyed during this paper's research similarly expressed that the two prevailing QoS models in use included IntServ and DiffServ; these professionals also unanimously agreed that DiffServ was more prevalent in present corporate networks. Let us take a take a closer look at these two models, beginning with IntServ.

IntServ is a fine-grained approach to QoS. It depends upon local and distant end network devices, and all routing devices in-between, to predetermine and agree upon data priority and service levels on a flow-by-flow basis. This requires the use of the Resource ReSerVation Protocol (RSVP) which sets up the path with the sending, receiving, and intermediary routing devices. Once a path is "reserved", the intermediary routers attempt to honor the path parameter constraints until it is cancelled or times out. Cisco stated that IntServ

> …provides a way to deliver the end-to-end Quality of Service (QoS) that real-time applications require by explicitly managing network resources to provide QoS to specific

user packet streams (flows). It uses "resource reservation" and "admission control" mechanisms as key building blocks to establish and maintain QoS.

IntServ uses Resource Reservation Protocol (RSVP) to explicitly signal the QoS needs of an application's traffic along the devices in the end-to-end path through the network. If every network device along the path can reserve the necessary bandwidth, the originating application can begin transmitting. Besides end-to-end signaling, IntServ requires several functions on routers and switches along the path:

- Admission Control: determine whether a new flow can be granted the requested QoS without impacting existing reservations
- Classification: recognize packets that need particular levels of QoS Policing: take action, including possibly dropping packets, when traffic does not conform to its specified characteristics
- Queuing and Scheduling: forward packets according to those QoS requests that have been granted. Cisco (n.d.)

DiffServ is a coarse grained approach to QoS. DiffServ involves creating traffic classification types, assigning priorities to those classifications, marking traffic to fit into appropriate classification types, and then processing the data flow for each classification type as a whole, in accordance with its prioritization. In this scheme, if two resources are trying to send data, and they fit into different classification priorities, then the IP packets marked with the higher classification will be processed first. According to all four Cisco certified professionals, DiffServ is the predominant model used in networks today, and the general process for a DiffServ QoS implementation involves identifying data types in the network, defining priority classifications, marking priority classification type in the DSCP field of the IP headers at the edge of networks - as close to the source as possible, and then managing and enforcing rules for classified data “flows” through any portions of the network(s) that you have control over. There are several nuances to the mechanisms available under DiffServ. Cisco described DiffServ as follows.

> Differentiated Services is a multiple service model that can satisfy differing QoS requirements. With Differentiated Services, the network tries to deliver a particular kind of service based on the QoS specified by each packet. This specification can occur in different ways, for example, using the 6-bit differentiated services code point (DSCP) setting in IP packets or source and destination addresses. The network uses the QoS specification to classify, mark, shape, and police traffic and to perform intelligent queueing. Differentiated Services is used for several mission-critical applications and for providing end-to-end QoS. Typically, Differentiated Services is appropriate for aggregate flows because it performs a relatively coarse level of traffic classification.
>
> DS Field Definition
>
> A replacement header field, called the DS field, is defined by Differentiated Services. The DS field supersedes the existing definitions of the IP version 4 (IPv4) type of service (ToS) octet (RFC 791) and the IPv6 traffic class octet. Six bits of the DS field are used as the DSCP to select the Per-Hop Behavior (PHB) at each interface. A currently unused 2-bit (CU) field is reserved for explicit congestion notification (ECN). The value of the CU bits is ignored by DS-compliant interfaces when determining the PHB to apply to a received packet.
>
> Per-Hop Behaviors

> RFC 2475 defines PHB as the externally observable forwarding behavior applied at a DiffServ-compliant node to a DiffServ Behavior Aggregate (BA). With the ability of the system to mark packets according to DSCP setting, collections of packets with the same DSCP setting that are sent in a particular direction can be grouped into a BA. Packets from multiple sources or applications can belong to the same BA. In other words, a PHB refers to the packet scheduling, queueing, policing, or shaping behavior of a node on any given packet belonging to a BA, as configured by a service level agreement (SLA) or a policy map. The following sections describe the four available standard PHBs:
>
> - Default PHB
> - Class-Selector PHB (as defined in RFC 2474)
> - Assured Forwarding PHB (as defined in RFC 2597)
> - Expedited Forwarding PHB (as defined in RFC 2598)
>
> For more information about default PHB, see RFC 2474, *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers.* (Cisco, 2007)

More technical details about DiffServ, along with a sample implementation, are located within the referenced Cisco document. Additionally, refer to current IETF guidance for the most recent revisions and updates on both IntServ and DiffServ. Now that we have had a brief look at IntServ and DiffServ, we will explore research that supports the claim that QoS has value in networks, and would benefit ARNe.

Research Regarding QoS Implementation

As we review the following QoS research, we will keep in mind that the purpose of this paper is to find possibilities for a minimal QoS design ARNe that allows VoIP to be processed ahead of other data, and for allowing future students to have a good starting point for understanding QoS and its design and implementation. The research collected is focused on showing that QoS *can* help in ARNe and is unlikely to do any harm if designed and implemented carefully. The resulting research, therefore, focuses on showing that QoS has demonstrated benefits and limited risk.

Davie confirms that QoS, specifically DiffServ, has successfully been deployed, assuring priority treatment for time sensitive applications such as VoIP. He finds no noteworthy risk involved with its implementation and points out that DiffServ has proven both effective and scalable. Davie found that

> … there has been a moderate amount of diffserv deployment in enterprise networks. One primary driver for this is VOIP, with the IP phones and VOIP gateways marking all traffic that they source as EF, and the routers and switches being configured to treat such traffic with the appropriate PHB. … diffserv *is* now deployed in a large number of service provider networks and appears to be meeting a range of needs well. The biggest reason for lack of deployment of diffserv appears to be excess capacity in the ISP backbones, which can hardly be interpreted as a negative statement about diffserv.
>
> If excess capacity existed *everywhere* in the Internet, and was sure to do so forever, it would clearly be time to give up on the idea of deploying any QoS mechanisms in the Internet. When we consider the fact that the Internet includes expensive access circuits from customers to providers and peering points that are not always well provisioned, it seems that that time has not arrived.

> In one of its primary goals - scalability - diffserv has been extremely effective. The fact that so many providers have turned on diffserv features in large networks, even if only for VPN or VOIP applications, indicates that the scalability hurdle, which is a high one, has been successfully cleared. The PHB definitions of diffserv also appear to have been very effective. They are all implemented in a wide range of products and they are all in use in large provider networks today…the rising popularity of VOIP in large corporations has been a clear driver for diffserv in those environments, where it appears to be working well…Thus we conclude that diffserv is, on balance, a successful technology. (Davie, 2003)

While the research presented by Davie provided a generalized affirmation that DiffServ QoS has been effectively implemented in many cases, we now move on to specific examples and results of DiffServ implementations. The first set of data is collected from results of testing in a laboratory environment.

Data presented by Davidson, Peters, Bhatia, Kalidindi, & Mukherjee compared using VoIP across a network first with no QoS and then with QoS. Their QoS laboratory set up is described as follows.

> For the testbed, Cisco used a simple network with two Cisco VoIP gateways and one 56-kbps WAN. It completed two tests:
>
> - Test A consisted of testing with and without QoS enabled while steadily increasing the saturation of the WAN link.
> - Test B consisted of testing with and without QoS enabled while sending traffic across the WAN link in a bursty nature.

> When no QoS was used, FIFO queuing was implemented. When QoS was used, MCML PPP, WFQ, and IP Precedence were utilized. Cisco sent the traffic through two in-house traffic generation tools. It measured latency and voice quality using a voice-quality test tool that utilizes the ITU-T Perceptual Speech Quality (PSQM) recommendation P.861. (Davidson, et al., 2006)

Their first test dataset, the results of utilizing consistent data flow rates, is presented in Figure 2.1. The data showed that as bandwidth saturation increased, the PSQM and delay measurements were consistently better—had lower numbers—with QoS enabled.

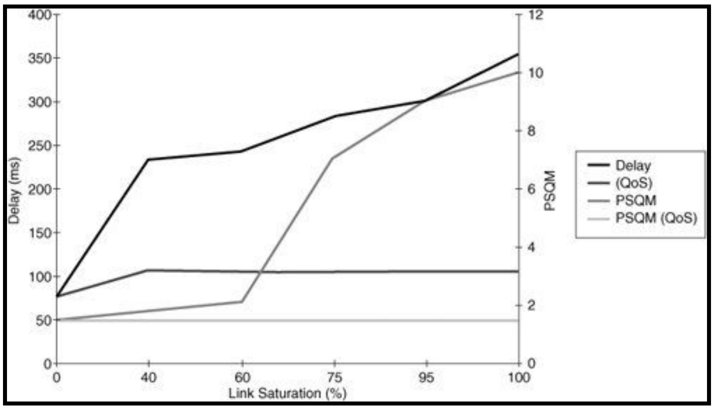


Figure 2.1 Results from Test A (Davidson et al., 2006)

Their second test data, with bursty data rates, is presented as a graph in Figure 2.2. Again, the data set showed that PSQM and delay measurements were consistently better—had lower numbers—with QoS enabled.

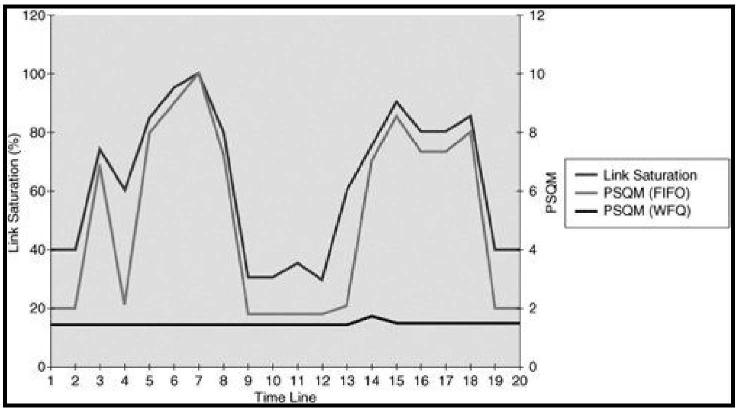


Figure 2.2 Results from Test B (Davidson et al., 2006)

The quantitative data presented by Davidson et al. is now followed by another research example using laboratory testing data. In this case, TCP flows are examined – first with no QoS, and then with DiffServ implemented.

Sander, Foster, Roy, and Winkler conducted quantitative research and testing regarding a DiffServ implementation for high priority TCP flows. They found that their DiffServ implementation did have a positive impact on ensuring high priority data QoS for TCP data with minimal impact on lower priority traffic. Their research study involved tests incorporating Committed Access Rate (CAR) to mark the IP precedence field of higher priority packets and to police the data flows and premium aggregate. The priority markings utilized corresponded to Expedited Forwarding per-hop forwarding behavior, specified in IETF RFC 2598. CAR was implemented at the edge router ingress port. Their test also utilized Weighted Fair Queuing (WFQ) to ensure a defined bandwidth available to flows, based on IP precedence class. WFQ was implemented at the edge router egress port and in interior routers. (Sander et al., 2000) Figure 2.3 illustrates a diagram of the network used for the testing.

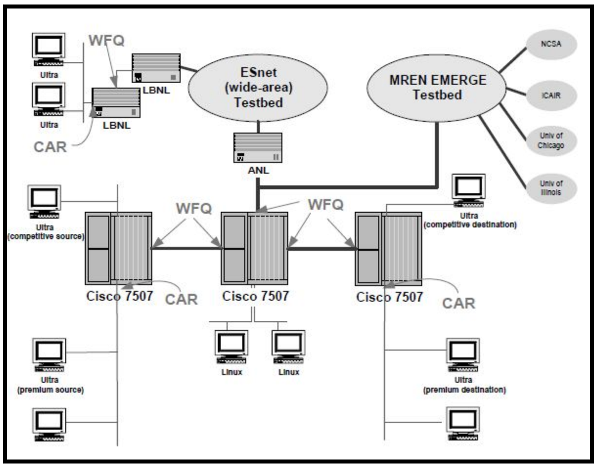


Figure 2.3 Test Network Diagram (Sander, et al., 2000)

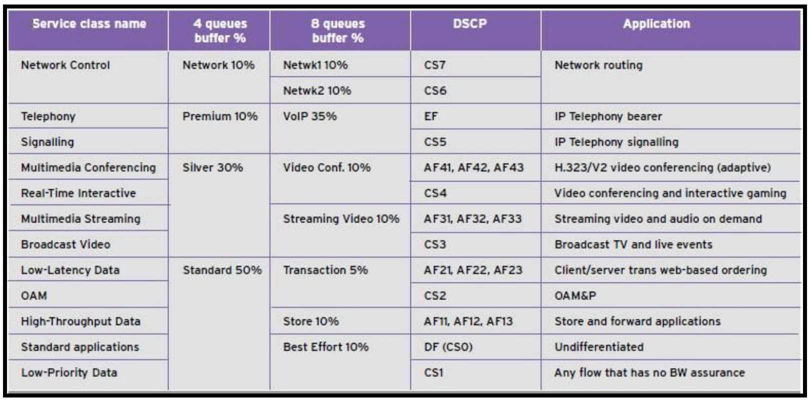
Extensive test data, incorporated with mathematical analysis can be reviewed in their research paper. The summary of their findings, indicating that they tested TCP flows using a DiffServ model, with CAR and WFQ, stated

> We have presented a quantitative evaluation of DS implementation for high-performance TCP-flows and demonstrated that end-to-end QoS can be delivered to such flows if DS mechanisms are configured carefully. The basic challenge is dealing with the burstiness introduced by TCP's sliding window. This must be addressed by appropriate policies at the edge routers; these policies must support bursts correlating to the TCP window size. Burstiness also introduces problems on the interior interfaces, because the available bandwidth might be exceeded by accumulating aggregate bursts. For that reason the implementation of the EF PHB should avoid queuing by overprovisioning the guaranteed amount of bandwidth for the premium class as much as possible.
>
> We have shown that CAR and WFQ can be used to implement a DS architecture. Policing at edge routers is done based on the applied bandwidth and the estimated round-trip

> time. The EF PHB is implemented by guaranteeing 99% of the available bandwidth to the premium class, which minimizes the maximum latency of premium traffic. We have demonstrated that the impact of this configuration on best-effort traffic in the absence of premium traffic is negligible. (Sander et al., 2000)

We have seen tested DiffServ implementations from Davidson et al. and from Sander et al. Another demonstration of a documented improvement in performance for priority network resources after QoS implementation can be seen in a report by Fernandes of Nortel Networks, which is presented next.

Nortel Networks is a well recognized provider of networked telephony systems. They provide access to case studies for viewing some of their best practices. One of their case studies examines how they deployed a DiffServ QoS design to ensure proper prioritization and treatment of several categories of data, including VoIP and multimedia conferencing. For their design and implementation, they utilized a four tier template, with configuration values recommended in IETF RFC 4594. In the study, Fernandes stated that they chose to follow IETF recommendations in order to comply with known, accepted standards. Their template for designing QoS identifies details such as their Differentiated Services Code Point (DSCP)-to-service class mapping, and is illustrated in Figure 2.4



Service class name	4 queues buffer %	8 queues buffer %	DSCP	Application
Network Control	Network 10%	Netwk1 10%	CS7	Network routing
		Netwk2 10%	CS6	
Telephony	Premium 10%	VoIP 35%	EF	IP Telephony bearer
Signalling			CS5	IP Telephony signalling
Multimedia Conferencing	Silver 30%	Video Conf. 10%	AF41, AF42, AF43	H.323/V2 video conferencing (adaptive)
Real-Time Interactive			CS4	Video conferencing and interactive gaming
Multimedia Streaming		Streaming Video 10%	AF31, AF32, AF33	Streaming video and audio on demand
Broadcast Video			CS3	Broadcast TV and live events
Low-Latency Data	Standard 50%	Transaction 5%	AF21, AF22, AF23	Client/server trans web-based ordering
OAM			CS2	OAM&P
High-Throughput Data		Store 10%	AF11, AF12, AF13	Store and forward applications
Standard applications		Best Effort 10%	DF (CS0)	Undifferentiated
Low-Priority Data			CS1	Any flow that has no BW assurance

Figure 2.4 Template with DSCP to service mapping (Fernandes, 2008)

This case study provided some specific testing methods, testing results, and presented lessons learned; as Fernandes stated

> We used a number of different methods to test various aspects of the voice path. For example, we did voice payload and bi-directional trace route testing as well as TCP (Transmission Control Protocol) performance tests. The reason for performing these tests is that the voice payload is User Datagram Protocol (UDP), and yet voice signaling is TCP. We wanted to ensure that the payload was not subjected to high amounts of latency, jitter and discards. The TCP tests were used to ensure the signaling channels would take the correct queue. After each site's first QoS deployment, we also used CLI Manager and other network management tools to proactively monitor the router interfaces. Based on interface queue drops and tags, we adjusted the router configurations.
>
> Lessons Learned
>
> - Classify your applications and agree to the QoS (DSCP) application marking.
> - Put together a standardized switch/router configuration.

- When rolling out configurations across the network, use network tools to simplify the configuration deployment.
- Proactively monitor network links and QoS on router interfaces, and adjust router buffer configurations to address any drops in the voice queue.
- Remember, especially for users on a VPN transport, that the default service offering associated with the Internet is a best-effort service. (Fernandes, 2008)

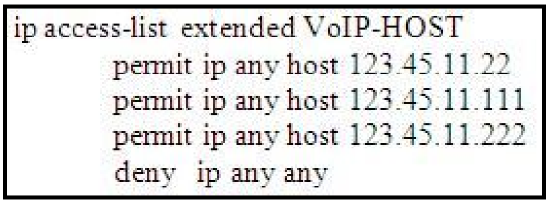
Nortel Network's case study by Fernandes provided an example of a full QoS design and implementation, using DiffServ with features such as DSCP marking, class mapping, and queuing, in accordance with IETF recommendations. Let's take a look at a somewhat smaller scale QoS design implemented by Rivada Networks in a deployable communications system.

Rivada Networks designs, integrates, and operates deployable communication systems for government agencies. These systems require an assurance of quality of service for certain applications due to their use in public safety and security. The Joint Incident Site Communications Capability (JISSC) system illustrates one example of how Rivada Networks designed and implemented a DiffServ QoS solution which provided priority treatment for services such as VoIP and VTC. Their design technique, parallel to the technique used by Nortel Networks, involved common basic practices. The basic steps included the following.

- Identified data traffic types and their priority criteria
- Grouped the data traffic types into prioritized classes
- Defined policy rules for each priority class
- Applied policy rules to appropriate interfaces

After identifying data traffic and priority expectations, Rivada Networks created a DSCP to service mapping table, very similar to the one utilized by the Nortel Networks example in Figure 2.4. With this level of the design plan completed, they then incorporated the information into the JISCC router configurations using IETF and Cisco recommended settings. Examples of selected details from the edge router configuration, with Rivada Networks proprietary details removed or substituted, show criteria such as IP packet DSCP marking, traffic shaping, and queuing. The details are not unique, as they are employed nearly identically in similar IETF and Cisco compliant implementations available through research, such as the Nortel Networks case study. Brief explanations are provided for before each section of the router configuration examples.

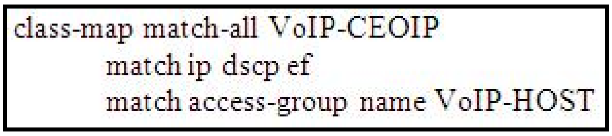
Access-lists and extended access lists permit or deny packets, based on a variety of criteria such as source or destination IP address, IP protocol information, TCP or UDP protocol, and TCP or UDP source or destination port. They associate these packets with the user defined name of the access-list, in this case “VoIP-HOST”. In Figure 2.5, the “permit” and “deny” commands create limits such that only hosts with the IP addresses shown are in agreement with that extended access list.



```
ip access-list  extended VoIP-HOST
          permit ip any host 123.45.11.22
          permit ip any host 123.45.11.111
          permit ip any host 123.45.11.222
          deny  ip any any
```

Figure 2.5 Extended Access List example (Zwick, 2010)

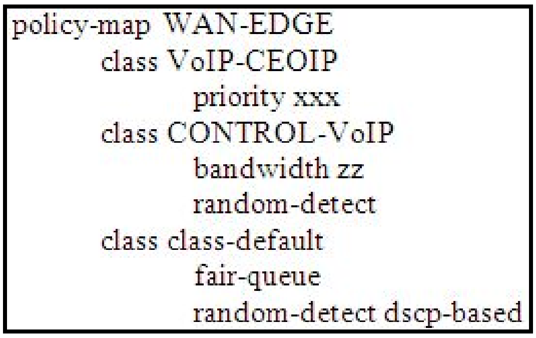
Class maps categorize packets into classes, based on matches to access lists, extended access lists, and other criteria. As an example, in Figure 2.6, the class-map named “class-map match-all VoIP-CEOIP”, a match occurs when an IP packet’s DSCP field is set to the “ef” value (46), *and* it belongs to the access list “VoIP-HOST”. Both criterion must be met for the match since the word “all” instead of “any” is used.



```
class-map match-all VoIP-CEOIP
        match ip dscp ef
        match access-group name VoIP-HOST
```

Figure 2.6 Class Map example (Zwick, 2010)

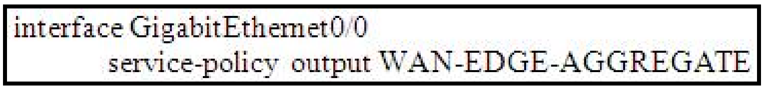
Policy maps take the categorized “flows” of IP packets and assign them rules related to bandwidth and priority, which ingress and egress interfaces will later enforce for each flow. These rules include criteria such as per flow bandwidth allocation, queuing method, and identifying total aggregate speed available. In Figure 2.7, for the policy-map named “policy-map WAN-EDGE”, the flow identified by the class “VoIP-CEOIP” is guaranteed bandwidth of xxx kbps utilizing the “priority” command; the flow identified as “CONTROL-VoIP” is guaranteed bandwidth of zz kbps utilizing the “bandwidth” command . There are nuances to the bandwidth rules and impacts of the priority and bandwidth commands which will not be explained in this research. The queuing mechanisms shown here include First In First Out for VoIP (and other services not shown), with all lower priority services sharing remaining buffer space based on Class-Based Weighted Fair Queuing. An example of the final stage of QoS design and implementation, applying policies to interfaces, is discussed after this policy map excerpt.



```
policy-map WAN-EDGE
        class VoIP-CEOIP
                priority xxx
        class CONTROL-VoIP
                bandwidth zz
                random-detect
        class class-default
                fair-queue
                random-detect dscp-based
```

Figure 2.7 Policy Map example (Zwick, 2010)

Applying a policy to an interface is fairly straight forward. In Figure 2.8, the policy map WAN-EDGE-AGGREGATE (not shown previously) was applied to the egress interface GigabitEthernet0/0. The “service-policy output” command applied the rules which were specified in the WAN-EDGE-AGGREGATE policy map.



```
interface GigabitEthernet0/0
        service-policy  output WAN-EDGE-AGGREGATE
```

Figure 2.8 Applying policy to interface example (Zwick, 2010)

The research data from Rivada Networks was the result of direct involvement by the researcher, as the researcher was an employee of Rivada Networks. Data was collected by the researcher, proprietary information was changed or removed, and then the data was summarized.

We have taken a look at research intended to support the idea that investigating a basic QoS design would benefit ARNe and its students. Let us take a look at research providing a brief overview of the portions of ARNe that would be a likely starting point for considering a basic QoS design. This research identifies primary components and potential basic priority data needs of ARNe at the DTC and ILB campus.

First, in Figure 2.9, we examine the ARNe network diagram which was available from the Regis University SEAD electronic shared document repository on October 7, 2009. While the diagram is not updated due to the current move of the DTC equipment, the router details (models and their basic connectivity), remain essentially unchanged, with the exception that one Cisco IP phone will be connected to the 2821 edge router at DTC and one Cisco IP phone will be connected to the 2821 edge router at ILB. Both 2821 routers will run a version of Cisco CallManager Express, with FXS port connectivity to an analog phone line at each site. The diagram is directly followed by a description of ARNe provided by Regis University Assistant Professor Dan Likarish.

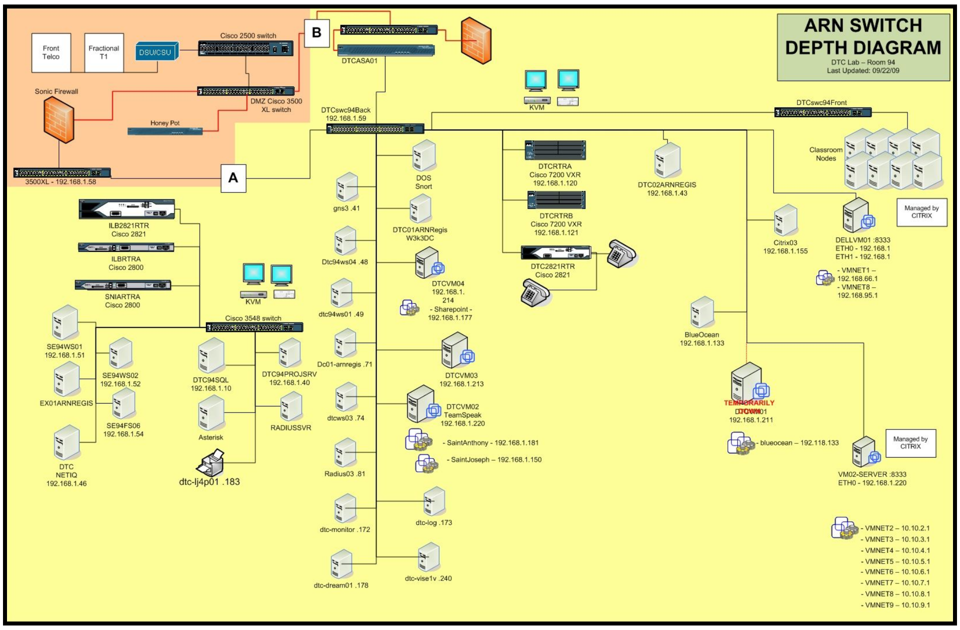


Figure 2.9 ARNe at DTC and ILB (Regis University, 2009)

ARNe (Regis University Academic Research Network) ARNe for short, is the graduate student run and managed intranet that spans our campus locations, Fort Collins to Colorado

> Springs and online. The CIT labs/classrooms are nodes on ARNe that are managed by all our second year graduate students in conjunction with faculty advisors.
>
> Both graduate and undergraduate teaching occurs within the labs during each term of the year and through our Vlabs. Each CIT classroom allows students to gain valuable hands-on experience with system engineering, provides advanced teaching tools for instruction, interact with other MSCIT students and a place for students to complete their professional project…ARNe has a three-fold mission:
>
> - Support the MSCIT teaching curriculum (we also support the undergraduate teaching at CIS/SPS) through hands-on experience with hardware and software
> - Provide a pleasant, stimulating environment for student professional projects
> - Support and encourage the collaboration between Oracle DBAs and the system engineering students. (Likarish, 2004)

A sample of services in ARNe that may benefit from QoS was built from the description by Likarish, from direct observations and interactions with ARNe throughout the SEAD practicum course August to September, 2009, and from interactions with Regis University alumni. The sample of services are identified as: VoIP for the IP phones, access to Adobe Connect meetings for on campus students and staff, Citrix access for both on and off campus students and staff, streaming video for instructional presentations for both on and off campus students and staff, and Virtual Machine lab access for both on and off campus students and staff.

A consensus among the Cisco professionals surveyed indicated that VoIP is typically the service least tolerant of latency and jitter; therefore, VoIP was selected as the service which QoS research would receive slightly more weight in subject coverage. If it is later determined that another ARNe service requires a higher quality of service than VoIP, the same QoS design principles investigated in this paper would similarly apply to designing QoS for that service.

Findings from experiences with SEAD and interactions with alumni indicated that, with regard to VoIP, Cisco CallManager Express was to be configured on the DTC model 2821 Cisco router in the near future. This VoIP implementation would support one IP phone connected to the DTC 2821 Cisco router, and would remotely support one IP phone connected to the 2821 Cisco router at ILB. At the DTC 2821 router, access to the Public Switched Telephone Network (PSTN) would be provided to both locations via an FXS port on the router However, recent design changes were made, and the current plan is to have Cisco CallManager Express, an FXS port, and PSTN line at both DTC and ILB sites. This somewhat lessens, but does not negate the need for VoIP QoS, because calls placed from one site IP phone to the other site IP phone will still travel over the network paths versus the Public Switched Telephone Network (PSTN) path. Additionally, as identified earlier, the results from research into possibilities for a very basic QoS design can just as just easily be applied to another ARNe service, such as streaming video or video conferencing, as it is applied to VoIP. With this background understood, we now move on to recommendations offered by one of the Cisco professionals surveyed; the recommendations included general QoS design steps, and then a recommendation for more specific VoIP QoS design steps in ARNe.

General Recommendation:

Identify traffic and its requirement

- ○ network audit

- ○ service level required
- ○ project audit
  - § determine how each type of traffic is important for projects

Divide traffic into classes

- ○ low latency: priority 5: such as video and voice
- ○ Mission critical: priority 4
- ○ Best-Effort: priority 2
- ○ Scavenger: data packets as priority 0
- ○ Note: classification can be recognized based on:
  - § DSCP
  - § IP precedence
  - § Source address
  - § Destination address

Define policies for each traffic class

- ○ set minimum bandwidth guarantee for priority one = 500kbps
- ○ set maximum bandwidth limits for scavenger = 100kbps
- ○ assign priorities to each class
- ○ manage congestion

Specific VoIP QoS Recommendation:

1. Create ACLs:

```
access-list 100 permit ip any any precedence 5

access-list 100 permit ip any any dscp ef

access-list 101 permit tcp any host a.b.c.d range 2000 2002  ! note ports for Cisco Call
```

```
manager
access-list 101 permit tcp any host a.b.c.d range 11000 11999 ! note ports for MGCP
```

2. Create class-map

```
class-map VOIP-RTP
    match access-group 100
class-map VOIP-Control
    match access-group 101
```

3. Create policy-map

```
policy-map QOS-POLICY
    class VOIP-RTP
        priority 100 ! PQ minimum 100kbps bandwidth
    class VOIP-Control
        bandwidth 8 ! minimum 8kbps bandwidth
    class class-default
        fair-queue ! all other traffic wfq
```

4. Apply it to an interface

```
interface  s1/0
    service-policy output QOS-POLICY
```

A variety of research has been presented, beginning with research from Davie supporting that DiffServ is effective and scalable. Research from Davidson, et al. provided data on VoIP QoS effectiveness, while research from Sander, et al. provided data on TCP QoS effectiveness with a DiffServ implementation. Research data from DiffServ implementations for multiple services at both Nortel Networks and Rivada Networks has also been presented. The final research in this section provided background on ARNe itself, along with the recommendations of a Cisco certified professional familiar with ARNe. With this data obtained, we will now examine the research methodology used in this paper.

Chapter 3 – Methodology

The qualitative methodology was utilized in conducting research and analysis for this paper. This methodology is described and contrasted with quantitative methodology by Leedy and Ormrod (2004, p.94) as follows "…quantitative research is used to answer questions about relationships among measured variables with the purpose of explaining, predicting and controlling phenomena. In contrast, qualitative research is typically used to answer the questions about the complex nature of phenomena, often with the purpose of describing and understanding the phenomena from the participants' point of view." The qualitative methodology was selected with the intent of presenting comparisons and contrasts of empirical data, primarily collected from case studies and technical literature closely related to QoS and its implementation. The category of qualitative methodology that this research and analysis most closely follows, therefore, is case study, referencing multiple case studies. Case study methodology is described by Yin (2008) as he stated

> 1. A case study is an empirical inquiry that
>
> - investigates a contemporary phenomenon in depth and within its real-life context, especially when
> - the boundaries between phenomenon and context are not clearly evident…
>
> 2. The case study inquiry
>
> - copes with the technical distinctive situation in which there will be many more variables of interest than data points, and as one result

> - relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result
> - benefits from the prior development of theoretical propositions to guide data collection and analysis.
>
> In essence, the two-fold definition shows how case study research comprises an all-encompassing method--covering the logic of design, data collection techniques, and specific approaches to data analysis… case studies can be conducted and written with many different motives. These motives vary from simple presentation of individual cases to the desire to arrive at broad generalizations based on case study evidence but without presenting any of the individual case studies separately…Case studies can cover multiple cases and then draw a single set of cross-case conclusions. (p.18-20)

Within the explanation by Yin, we see the term triangulation used. Triangulation is further described by Niglas (2000) as one or more of four types: data triangulation, investigator triangulation, theory triangulation, and methodological triangulation. Niglas (2000) stated that

> Denzin triangulation means more than using multiple measurements of the same phenomenon - in addition to the use of diverse data, it involves combining different methods and theories, as well as perspectives of different investigators. Denzin (as cited in Denzin,1978) has clearly identified four different types of triangulation:
>
> *data triangulation* - the use of variety of data sources and data sets in a study. Data may be both qualitative and quantitative, gathered by different methods or by the same method from different sources or at different times.

> *investigator triangulation* - the use of several different researchers. Here the importance of partnership and teamwork is underlined as the way of bringing in different perspectives.
>
> *theory triangulation* - the use of different theoretical viewpoints for determining competing hypotheses as well as for interpreting the single set of data.
>
> *methodological triangulation* - the use of multiple methods to study a single problem or phenomenon. It may also include the use of the same method on different occasions and situations.

Triangulation was a key component of the research methodology used to discern meaning from the research sources identified in this paper. With an understanding of the qualitative case study methodology and triangulation, we can now move on to analyzing the research data with a goal of appreciating possibilities for a basic QoS implementation in ARNe.

Chapter 4 –Results

In this chapter, several tables were used aid in describing and comparing a summary of data collected from the research discussed earlier in this paper. The purposes of these tables included the following with regard to QoS design and implementation.

- Identify commonalities
- Identify differences
- Use triangulation to suggest conclusions which apply to ARNe

Table 4.1 is the side-by-side comparison of the services provided QoS in the research cases. An analysis follows the table.

Table 4.1 Comparison of services provided QoS

	Service(s) provided with higher levels of QoS
Davidson et al.	VoIP
Sander et al	TCP
Fernandes	VoIP, VTC, Network control/routing, Streaming video
Rivada Networks	VoIP, VTC, Network control/routing

An obvious conclusion from this set of data was that all of these research cases did have one or more services configured to receive higher priority treatment by QoS mechanisms. This result confirms that QoS *can* be designed and implemented on networks with a range of service prioritization needs, from minimal to complex. The next research analysis involved comparing the relative size of the networks.

During the research collection process, it was very difficult to find QoS implementations in networks that were extremely close in size and complexity compared with ARNe. Table 4.2 provided a comparison of network sizes for the research cases that were chosen.

Table 4.2. Comparison of network sizes

	Size of network at QoS implementation sites relative to ARNe
ARNe	Reference standard
Davidson et al.	Somewhat smaller
Sander et al	Somewhat smaller
Fernandes	Various sizes, but generally larger
Rivada Networks	Somewhat smaller

While the usage of the term “size” introduces some subjectivity, however, the results did indicate that QoS can be implemented on a variety of network sizes, ranging from smaller than ARNe to larger than ARNe. Using inductive reasoning, the assumption could be made that a network the exact size of ARNe can have a QoS implementation. Next, we will look at a comparison of the QoS model and mechanisms used in the research examples.

With regard to QoS models, the decision was made early in the research process to focus on the DiffServ model versus the IntServ model, based primarily on concurrence of four Cisco certified experts that DiffServ was the predominant model used in networks, as stated earlier in this paper. Additionally, from a pragmatic point of view, it stood to reason that Regis University will students benefit more from direct experience designing, implementing and fine-tuning a QoS model that is most often used in the networks of potential employers. Table 4.2 illustrates a summary comparison of the QoS implementations.

Table 4.3. Comparison of QoS Implementation Types

	Type of QoS implementation
Davidson et al.	First no QoS with FIFO, then DiffServ & IP Precedence, MCML PPP and WFQ
Sander et al	DiffServ with EF PHB, CAR and WFQ
Fernandes	DiffServ with RFC 4594 recommended settings, using EF PHB,AF PHB, and CBQ
Rivada Networks	DiffServ with RFC 4594 recommended settings, using EF PHB,AF PHB, and WFQ

Other than the fact that all four research sources implemented DiffServ, we can see that they all marked data IP headers in accordance with the Per Hop Behavior (PHB) settings for: IP Precedence, Expedited Forwarding, or Assured Forwarding, as specified in IETF RFC 4594. Using these markings, data was then mapped to an aggregate grouping for any same priority data. The aggregates, or classes, were then allocated bandwidth by a queuing mechanism. Three of the four implementations used Weighted Fair Queuing for at least one implementation tested, while one opted for Class Based Queuing. The final set of data is analyzed next.

The most basic question or problem posed in this paper alluded to the assumption that a basic QoS design *could* be successfully implemented on ARNe. The research data presented in Table 4.4 displays the status of whether each research source had a successful QoS implementation. Success, in this instance, is taken to mean that services designed to receive preferential QoS treatment, actually received preferential treatment in the implementation, while not causing unexpected negative impacts on lower priority data.

Table 4.4. Comparison of QoS Implementation Results

	Did test results show that QoS implementation was successful
Davidson et al.	Yes
Sander et al	Yes
Fernandes	Yes
Rivada Networks	Yes

In all four cases, QoS was implemented and test results showed that the implementations were successful. Through the Comparison of services provided QoS (Figure 4.1), the Comparison of network sizes (Figure 4.2), the Comparison of QoS Implementation types (Figure 4.3), and the Comparison of QoS Implementation results (Figure 4.4), one could conclude that the services existing on ARNe can be provided useful, basic QoS capabilities by creating and implementing a design based on the DiffServ model, utilizing IETF and Cisco recommendations. With the data collected by research analyzed, we now move on to general conclusions based on the results presented.

Chapter 5 – Conclusions

Introduction

This chapter includes general conclusions which could be drawn from the research results. These conclusions also suggest direction(s) for future research on a QoS design and implementation on ARNe, and some lessons learned during the research process. We begin some general conclusions.

General Conclusions

Just as with any system development process or lifecycle, the planning stage is the first, critical element. This research identified specific preliminary planning steps for considering and designing QoS. These same steps can be found throughout numerous Cisco technical sources, and in case studies examining successful QoS design and implementation. One may draw the conclusion from this that planning is essential to QoS design, just as it is to any system design effort. We move on now to a conclusion about DiffServ.

Through the research identified in this paper, and through numerous sources encountered during the background study leading up to the research, it becomes obvious that DiffServ is the QoS model of choice in today's corporate networks, due to its simple design, ease of use, and scalability. For the sake of getting the most value with regards to institutional educational goals, it is logical to conclude that curricula and assets providing exposure to DiffServ design and implementation is a choice for supporting Computer & Information System students. We move on now to a final conclusion regarding areas for future research.

Early in this paper, it was stated that QoS involves additional configurations on core routers and switches in a network such as ARNe. While getting the basic QoS design completed and implemented on ARNe edge routers will speed prioritized data processing through those edge routers, it is just as important to configure QoS in those other network components mentioned. These devices will, often by default, strip the prioritization information from the packet header and then the packet reverts to best-effort treatment for the rest of the trip through the network--unless the devices are configured not to do so. So, with regard to a conclusion about future research specifically following the path of this paper, it seems reasonable that further QoS design exploration for core routers and switches in ARNe is a valid area for continued academic research.

Lessons Learned

While working through the research process in support of exploring the thesis of this paper, several challenges were encountered that may have been mitigated by taking a slightly different course of action. The following are lessons learned regarding this research process, which may help future researchers

- First, discuss the thesis research process with a researcher that had recently successfully completed similar research in the same general area, in the same school of study. This reduces possible confusion over exactly how the process works and what is expected.
- For research involving aspects of specific routing and switching designs, for example QoS, seek one or more accurate, working router and switch configurations of the design under study; then, seek a subject matter expert to explain both the design planning process and to explain, step-by-step, how it was implemented in the configuration(s).
- With regard to research relating to an active network, complete research at an accelerated pace. During the course of the research for this paper, one set of

network equipment moved and there were design changes. Fortunately, the design changes did not significantly impact continued research.

Summary

The research data presented in this paper has served the purpose of exploring basic QoS design possibilities for ARNe edge routers. The data also supported that DiffServ and IntServ are the two QoS models defined by the IETF, with DiffServ being the more prevalent model currently used. Data collected and analyzed from laboratory and real-life QoS implementations and testing demonstrated that QoS was effective at guaranteeing preferred treatment of selected data, with very little negative impact on the remaining data. With the data from this research, and technical guidance from networking professionals and QoS configuration literature, a basic QoS design can be created and implemented on ARNe edge routers.

References

Hartpence, B. H. (2009). Qos content and experiences for IT, networking and security programs. In *Proceedings of the 10th ACM Conference on Sig-information Technology Education* (Fairfax, Virginia, USA, October 22 - 24, 2009). SIGITE '09. ACM, New York, NY, 60-64. DOI= http://doi.acm.org/10.1145/1631728.1631749

Clark, D. and Fang, W. (1998). Explicit allocation of best-effort packet delivery service. IEEE/ACM Trans. Netw. 6, 4 (Aug. 1998), 362-373. DOI= http://dx.doi.org/10.1109/90.720870

Braden, R., Clark, D. and Shenker, S. (1994). “Integrated services in the Internet architecture: An overview,” *Internet Engineering Task Force, RFC 1633*. Retrieved from: http://www.ietf.org/rfc/rfc1633.txt

Nichols, K., Blake, S., Baker, F., and Black, D. (1998). “Definition of the differentiated services field (DS field) in the IPv4 and IPv6 Headers” *Internet Engineering Task Force, RFC 2474*. Retrieved from: http://www.ietf.org/rfc/rfc2474.txt

Cisco. (2005). *DiffServ—The Scalable End-To-End Quality Of Service Model*. [White Paper]. Retrieved from http://www.cisco.com/en/US/technologies/tk543/tk766/technologies_white_paper09186a00800a3e2f.pdf

Cisco. (n.d). *Integrated Services*. Retrieved from http://www.cisco.com/en/US/products/ps6611/products_ios_protocol_group_home.html

Cisco. (2007). *Overview of DiffServ for QoS*. Retrieved from http://www.cisco.com/en/US/docs/ios/12_2/qos/configuration/guide/qcfdfsrv_ps1835_TSD_Products_Configuration_Guide_Chapter.html

Davie, B. 2003. Deployment experience with differentiated services. In *Proceedings of the ACM SIGCOMM Workshop on Revisiting IP Qos: What Have We Learned, Why Do We Care?* (Karlsruhe, Germany, August 25 - 27, 2003). RIPQoS '03. ACM, New York, NY, 131-136. DOI= http://doi.acm.org/10.1145/944592.944598

Davidson, J, Peters, J, Bhatia, M, Kalidindi, S, and Mukherjee, S. (2006) Voice over IP Fundamentals. Cisco Press: Indianapolis IN. Retrieved from http://fengnet.com/book/voip/ch08lev1sec6.html

Sander, V., Foster, I., Roy, A., and Winkler, L. 2000. A differentiated services implementation for high-performance TCP flows. *Comput. Netw*. 34, 6 (Dec. 2000), 915-929. DOI= http://dx.doi.org/10.1016/S1389-1286(00)00162-6

Fernandes, S. (2008). Quality of Service – A Key Attribute for IP Networks. Nortel Networks. Retrieved from http://www.nortel.com/corporate/nortel_on_nortel/collateral/qos_for_voip.pdf

Regis University. (2009). DTC Room 94 Diagram. Retrieved from https://in2.regis.edu/sites/scis/IT/SEAD/Systems/Network%20Diagram%20DTC%20Room94.vsd

Likarish, D. (2004). Practicum FAQ: What is ARNe?. Retrieved from http://www.dlikarish.com/nlp_overview.shtml

Leedy, P. & Ormrod, J. (2005). Practical Research: Planning and Design (8th ed.). Upper Saddle River, NJ: Pearson Education.

Yin, R. K. (2008). Case Study Research: Design and Methods (4th ed.). Thousand Oaks, CA: Sage Publications.

Niglas, K. (2000). Combining quantitative and qualitative approaches. Paper presented at the European Conference on Educational Research, Edinburgh, 20-23 September 2000 http://www.leeds.ac.uk/educol/documents/00001544.htm

Denzin, N.K. (1978) The research act: An introduction to sociological methods. New York: McGraw-Hill