Network Working Group   B. Jamoussi
Request for Comments: 2340   D. Jamieson
Category: Informational   D. Williston
                                                            S. Gabe
                                        Nortel (Northern Telecom) Ltd.
                                                           May 1998

Nortel's Virtual Network Switching (VNS) Overview

Status of this Memo

Copyright Notice

Abstract

   This document provides an overview of Virtual Network Switching
   (VNS).

   VNS is a multi-protocol switching architecture that provides COS-
   sensitive packet switching, reduces the complexity of operating
   protocols like PPP and frame relay, provides logical networks and
   traffic segregation for Virtual Private Networks (VPNs), security and
   traffic engineering, enables efficient WAN broadcasting and
   multicasting, and reduces address space requirements. VNS reduces the
   number of routing hops over the WAN by switching packets based on
   labels.

   VNS has been proven in production networks for several years.

Table of Contents

1. Introduction

   There are several key problem areas with today's wide area backbone
   networks that carry LAN traffic: scalability, service
   differentiation, redundancy, administration, and traffic containment.

   First, scalability is becoming a major concern because of the rapid
   growth in bandwidth demand and geographical reach. As the size of the
   WAN network grows traditional point-to-point and NBMA topologies or
   network models lose their performance.

   Second, the need to provide several Classes of Service (CoS) has
   never been greater. The days of a single "best effort" service are
   over and service providers demand ways to differentiate the quality
   of the service offered to their clients based on several policies.

   Third, the WAN is often carrying mission-critical traffic and loss of
   service is not acceptable. So far, path redundancy has been addressed
   inefficiently by requiring additional links or VCs.

   Fourth, network operators demand easy and simplified network
   administration. Large NBMA topologies require extensive PVC
   provisioning until SVC deployment becomes more ubiquitous. For
   Point-to-point models, IP address space may be used inefficiently and
   non-trivial network schemas are required to contain reserved address
   space.

   Finally, proper segregation of traffic is becoming a must. This
   requirement is being addressed today by adding leased lines or VCs
   used to separate traffic flows based on regions or interest or
   protocol.

   Nortel's Virtual Network Switching (VNS) is a technology that
   provides efficient solutions to these challenges.

Section 2 provides an overview of VNS. The VNS header is specified in
Section 3. Section 4 describes the VNS label distribution mechanism.
Section 5 defines how a VNS network can be partitioned into Logical
Networks (LN). Section 6 outlines VNS routing. Section 7 defines both
unicast and multicast forwarding. Section 8 describes the mechanisms
used to engineer the traffic. Section 9 defines the COS based
switching of VNS. Section 10 provides network migration scenarios
using VNS. A summary of VNS is provided in Section 11.

2. What is VNS?

Virtual Network Switching (VNS) is a CoS-sensitive multi-protocol
label switching architecture that reduces or eliminates the number of
layer 3 hops over the WAN by switching traffic based on labels.

VNS makes a network of point to point links appear to be a single
LAN (broadcast, multiple access) media. The network used by a
particular instance of VNS is called a Logical Network (LN) which is
described in more detail in Section 5.

In reference to the ISO Network Layering Model, the Data Link Layer
is expanded to include VNS network layer. To the ISO Network Layer,
(e.g., IP), VNS is treated as a Data Link Layer.

```
        -----------------------
       | Application           |
        -----------------------
       | Presentation          |
        -----------------------
       | Session               |
        -----------------------
       | Transport             |
        ----------------------   ------------------------
       | Network (e.g., IP)   |  / Network VNS          |
        ---------------------- /                        |
       | Data Link            |------------------------|
        ----------------------                          |
       | Physical             |  \ data link (e.g., ATM) |
        -----------------------   ------------------------
```
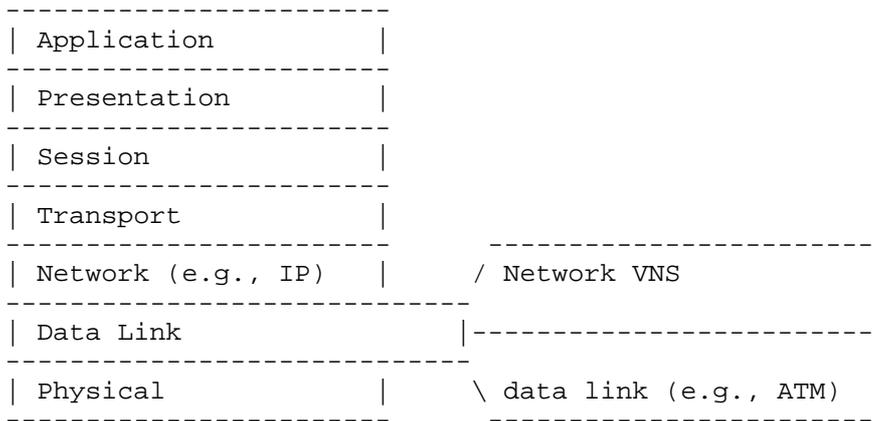
               Figure 1. ISO Network Layering Model for VNS

In a VNS Network, three separate nodal functions are defined. An
ingress node, an egress node, and a tandem node. The ingress and
egress nodes define the boundary between an IP network and the VNS
network. Therefore, these nodes run both IP routing and VNS routing.
However, tandem nodes need only run VNS routing.

A LAN packet is encapsulated in a VNS header as it enters the LN. The
label in the header is used to switch the packet across the LN. The
encapsulation header contains the identifier of the last node (or
egress node) that processes the packet as it traverses the LN. It is
the first node (or ingress node) that decides to which egress node
the packet is sent. All nodes between the ingress and egress nodes
(known as tandem nodes) decide independently the best packet
forwarding route to the egress node identified in the packet.

The network layer protocols view VNS as a shared broadcast media,
where the speed to reach any node on the media is the same for all
nodes. VNS ensures that traffic destined to other nodes is forwarded
optimally. This transparent view of the VNS means that all the
details of the network (for example, topology and link states) can be
hidden from the Upper Layer Protocols (e.g. Layer 3 routing
protocols) and their applications. VNS also ensures that changes to
topology and link state are hidden.

The network layer protocol on the ingress node views the network
layer protocol on the egress node as its logical and directly
connected neighbor. This is significant because the network layer
protocols always decide which directly connected neighbor should
receive a forwarded packet. The details of the actual topology
supporting the connectionless network are managed entirely by the
Virtual Network Switching and are hidden from the network layer
protocols. To the network layer, VNS simply appears to be another
Data Link Layer (or media), even though VNS is a network layer itself
running on top of the actual Data Link Layer (for example, ATM
trunks).

For the ingress node to choose the egress node that provides the best
path to the packet's final destination, it must have knowledge of the
following:

   - the nodes that can be reached in the network
   - the topology of the network that is using the VNS services for
     transport across the network (but not necessarily the topology
     of the full network)

This knowledge is obtained through the network layer routing
mechanisms such as, IP's Open Shortest Path First (OSPF) and Address
Resolution Protocol (ARP).

Once the network layer protocol on the ingress node has decided which
neighbor to transmit the packet to, it is the responsibility of VNS
forwarding, a part of VNS, to deliver the packet to that node. Once
the packet arrives at the egress node, the packet is delivered to the
network layer protocol, which then forwards it to its ultimate

destination.

Tandem nodes have no interaction with the network layer protocols.
They only require knowledge of the VNS network topology. They make
their packet forwarding decision on the egress node identifier and
LN identifier carried in the VNS header of the packet.

3. VNS Header

VNS defines a unicast header shown in Figure 2 and a multicast header
shown in Figure 3.

```
    3                   2                   1                   0
  1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |      TTL      |       LNN         |x|LS-Key |x|DP | CmnHdr   |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | Protocol Type |        Destination Node Identifier          |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |  COS  |x x x x|        Source Node Identifier               |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |               Network Layer Header (e.g. IP)                 |
 /                                                              /
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                         Data                                 |
 /                                                              /
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
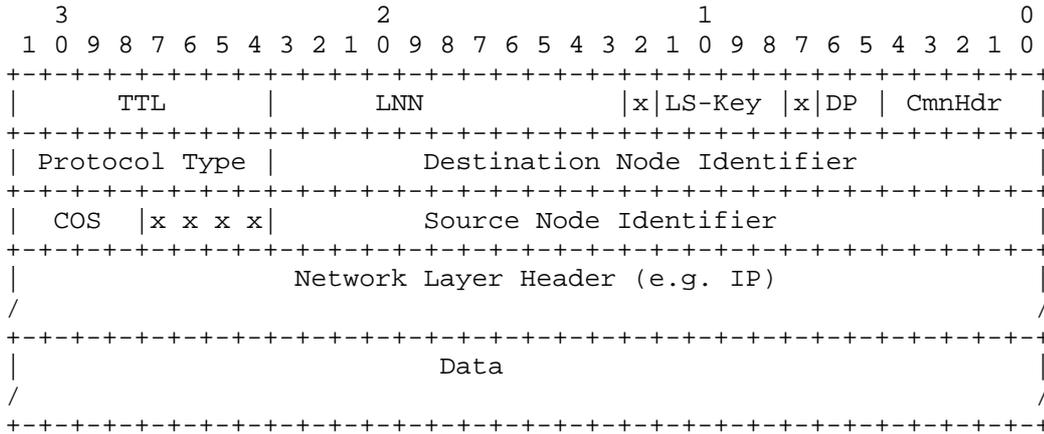
Figure 2. Unicast VNS Header

The unicast header includes the following fields:

- Common Header (CmnHdr): The common header identifies the packet to
be a VNS encapsulated packet.

- Discard Priority: Indicates the level of congestion at which the
packet should be discarded. The value of this field is assigned on
the originating node based on policy information (see Section 9).

- Load Spreading Key: indicates the stream to which the packet
belongs for the purposes of equal cost multipath and trunk load
spreading (see Section 8).

- LNN: The Logical Network Number defines the logical network the
packet belongs to. This field in is used in conjunction with the
destination node identifier as the VNS switching label (see Section
5).

   - TTL: The Time To Live field is used to detect and discard packets
   caught in temporary routing loops.

   - Destination Node Identifier: This field contains an ID which
   uniquely identifies the destination node. This ID is unique to the
   physical network not just the LN. In conjunction with the LNN, this
   forms a global VNS switching label.

   - Protocol Type: indicates the type of Network layer protocol being
   carried in the packet. Examples include IP, IPX, and Bridging. If the
   packet is a multicast packet then this is indicated in this field.

   - Source Node Identifier: This field contains an ID which uniquely
   identifies the source node (ingress node).

   - CoS: The Class of Service field is used to provide routing class of
   service. The COS field also affects the Emission Priority of the
   packet in the scheduler (see Section 9).

   - Reserved Fields: All the fields marked with "x" are Reserved.

```
    3                   2                   1                   0
  1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |      TTL       |         LNN           |x|LS-Key |x|DP | CmnHdr |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | PT = Multicast|         Destination Node Identifier           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |  COS  |x x x x|         Source Node Identifier                |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | Protocol Type |x x x x x x x x|    Multicast Group            |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |               Network Layer Header (e.g. IP)                  |
 /                                                              /
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 /                          Data                                /
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
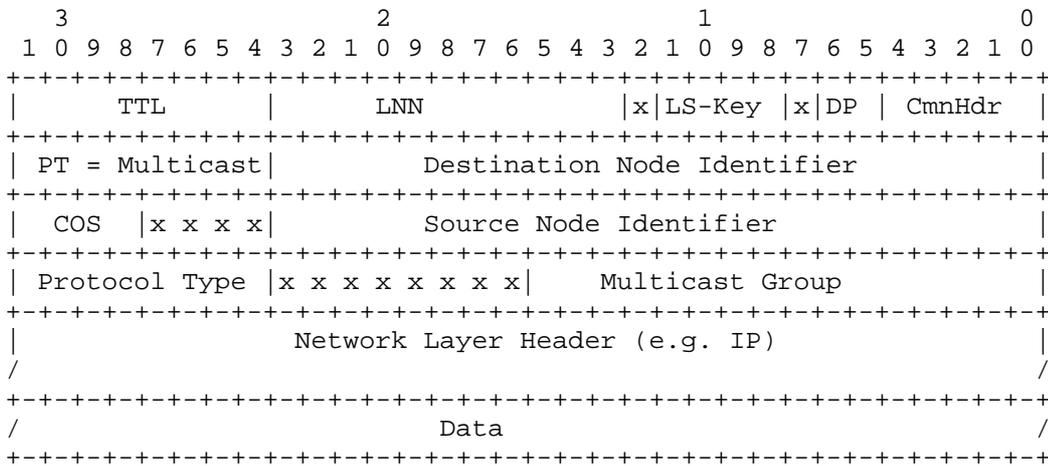
                     Figure 3. Multicast VNS Header

   The multicast header shown in Figure 3, includes all the fields of
   the unicast header. In addition, the multicast header includes the
   following fields:

   - Multicast Group: this field is used to identify a sub-group within
   the logical network that receives the multicast packets.

   - Protocol Type: indicates the type of Network layer protocol being
   carried in the packet. Examples include IP, IPX, and Bridging.

4. VNS Label Distribution

   Label distribution in VNS is based on a distributed serverless
   topology driven approach. Standard ARP or address gleaning is used to
   distribute and map network layer addresses to VNS addresses.

   A VNS Label is an 6 byte encoding of the LNN and the node ID. VNS
   Labels are treated as MAC addresses by the network layer. This means
   that labels are distributed by the same means network layers use to
   distribute MAC addresses. Thus, VNS leverages existing L2/L3 mapping
   techniques and doesn't require a separate Label Distribution
   Protocol.

5. Logical Networks (LNs)

   A logical network consists of a subset of the nodes in a network
   together with a subset of the trunking facilities that link those
   nodes. Logical networks partition the network into subnetworks that
   serve a subset of the overall topology.

   Each of the logical networks supported on any given node has a
   separate routing and forwarding table (built by VNS). Therefore,
   routing decisions are based on the resources available to the logical
   network, not the entire network.

   Each instance of VNS will discover all the trunks which are connected
   to neighbors which support a matching LNN. This provides a huge
   administrative saving, since VNS provisioning is on a per-node basis,
   not on a per-link basis. VNS provisioning requires only a unique
   node ID and an LNN. Discovery of which trunks support which LNNs is
   done at run time, relieving administrative effort, and allowing the
   LN to dynamically adapt to topology changes.

   Multiple Logical Networks provide the following benefits to the
   network system:

      - Logical networks allow service providers to service multiple
      private networks or (Virtual Private Internets) easily over one
      network.

      - Logical networks can be used to limit the impact of one network
      layer protocol on the others. This is particularly true for
      protocols that broadcast or multicast a large percentage of either
      their control or data packets. This increases the effective
      bandwidth of the trunks and allows the overall network to scale

   better.

      - Logical networks allow for the configuration of the network to
      meet individual community of interest and geographical
      subnetworking needs.

      - Routing control traffic has significance only in the local
      subnetwork that is isolated to that subnetwork.

      - Logical networks allow different instances of the same protocol
      to share trunk facilities.

6. VNS Routing

   VNS routing is a link state routing system which uses many concepts
   similar to OSPF and PNNI. One of the most significant departures from
   the others is its ability to calculate shortest path trees for
   routing unicast traffic and spanning trees for routing multicast
   traffic within a Logical Network.

   There is only one type of interface that VNS routing supports and
   this is known as a VNS link. A link is a set of trunks that join two
   VNS neighbor nodes. Each node in a VNS network maintains information
   about the state of locally attached links. This information is
   flooded throughout the network whenever there is a significant change
   to the link's state or attributes (i.e. up/down, speed change,
   available bandwidth change).

   Each node stores and forwards the link state information received
   from all other nodes. This allows each node to have the same view of
   all of the nodes in the network together with all of their link state
   information. This data is used to compute both the shortest path to
   reach each node in the Logical Network and a spanning tree for the
   Logical Network.

   Logical networks are not bound to a particular trunk or link. They
   are configured on a node. By default, a link will support a specific
   logical network if the two nodes which it connects both are
   configured to support the logical network number. This provides a
   significant savings in operations over having to configure logical
   networks on links or trunks.

   When a link first comes into service, a protocol is run which allows
   the two neighboring nodes to exchange information about the logical
   networks they support. This allows the two nodes to determine if the
   links are to be considered as a locally attached link for a logical
   network.

7. VNS Forwarding

   VNS supports two types of forwarding: unicasting and multicasting. In
   the first type, the data packet arrives on the ingress node and
   unicasting forwards the data packet to a single destination (egress
   node). In the second type, the data packet arrives on the ingress
   node and multicasting forwards the data packet to all other nodes in
   the logical network.

7.1 Unicast

   When a packet first enters the LAN internetwork, the network layer
   routing protocol determines the next hop of the best route for the
   packet to reach its final destination. If the best route is through a
   VNS Logical Network, the network layer routing protocol relies on VNS
   forwarding to get the packet to the egress node. A VNS packet header
   containing the node ID (the unique ID assigned to each node) of the
   egress node is added to the front of the packet and VNS forwarding is
   invoked to deliver the packet. The network layer routing protocol
   learns the egress node ID through an Address Resolution Protocol
   (ARP) for IP and Source Address learning for bridging.

   As the packet traverses the LN, routing decisions are made to
   determine the next hop in the route to reach the destination node ID
   specified in the VNS header. A forwarding table is built on each node
   that assists in making the routing decision.

   Each VNS instance on each node builds and maintains a forwarding
   table for its LN. Each forwarding table has an entry for every node
   that is a member of the logical network.

7.2 Multicast

   In addition to the unicast forwarding function, VNS also supports a
   multicast forwarding service for traffic within an LN at the VNS
   layer. Multicast packets are delivered to all nodes supporting the
   logical network to which the multicast packet belongs. The packets
   are sent along the branches of a spanning tree that is built by each
   node supporting the logical network and is based on a common root
   node (so that each node's view of the tree is the same as other
   nodes). In other words, multicast packets are sent intelligently,
   consuming a minimum of network bandwidth. If the network topology is
   stable, each node receives each multicast packet only once.

   Multicast packets received at any node are not acknowledged. They are
   simply forwarded to the specified network layer interface and sent to
   any other neighbor nodes on the spanning tree.

8. Traffic Engineering

   VNS forwarding supports two types of traffic engineering mechanisms:
   equal cost multipaths and trunk load spreading.

   Equal cost multipaths allows different streams (unique network layer
   source and destination address pairings) to be load spread between
   multiple relatively equal cost paths, through the Logical Network to
   the egress node.

   Trunk load spreading between two neighbors can take place when
   multiple VNS trunks are defined between neighbors. Again, the load
   spreading is based on network layer streams.

8.1 Equal Cost Multipaths

   From any point in a logical network, there may be multiple paths to
   reach a specific egress node. If VNS routing determines that more
   than one of these paths are of equal cost, VNS packets will be load
   spread between two of them.

   Equal cost multipath forwarding is supported not only on ingress
   nodes but on tandem nodes as well. Each packet on an ingress node is
   tagged with an equal cost multipath key. This key is acted upon at
   the ingress node and stored in the VNS header to be used on tandem
   nodes.

   The equal cost multipath key is calculated by running an algorithm
   over the source and destination network layer addresses. This means
   that, in a stable network, any given stream will always take the same
   path through a Logical Network avoiding the problems that misordering
   would otherwise cause.

8.2 Trunk Load Spreading Between Neighbors

   VNS allows multiple trunks to be configured between neighboring VNS
   nodes. VNS routing considers the aggregate bandwidth of those trunks
   to determine the metric between the nodes. Also, VNS load spreads its
   traffic amongst those trunks.

   As is the case with equal cost multipaths, the trunk load spreading
   key is calculated on the ingress node from an algorithm run over the
   source and destination network layer addresses. The key is then
   stored in the VNS header to be used on all tandem nodes through the
   Logical Network.

9. Class of Service

   At the ingress to a VNS Network, packets are classified according to
   the Class of Service (Cos) policy settings. The CoS differentiation
   is achieved through different Emission and Discard priorities. The
   semantics of the classification is carried in the VNS label (DP and
   COS Fields described in Section 3) to be used at the ingress node as
   well as all tandem points in the VNS network to affect queuing and
   scheduling decisions.

10. VNS Migration Strategies

   VNS supports several upper layer protocols such as IP, IPX, and
   Bridging. Therefore, it is a multiprotocol label switching
   architecture. In addition, VNS is not tied to a particular L2
   technology. It runs on cell (e.g., ATM) trunks, frame trunks, or a
   mixture of both.

   VNS can be gradually introduced in a network. It can be implemented
   between switching elements interconnected by point to point links.
   Each of the switching nodes can run layer 3 routing simultaneously
   with packet switching. VNS also allows for the interconnection of VNS
   clouds through an ATM VC.

   Since VNS can run on a mixture of Frame and Cell trunks, it allows
   for the graceful migration of the frame links to ATM without
   requiring a complete immediate overhaul.

11. Summary

   VNS addresses scalability problems in several ways:

      1. By a generally distributed design which doesn't
         require a Label Distribution Protocol, or servers of any kind.
      2. By providing an efficient, distributed multicast mechanism.
      3. By allowing administrators to control the size of a
         Logical Network, limiting traffic to a subset of the physical
         topology.
      4. By reducing layer 3 address space/subnet requirements in the
         WAN which reduces the routing table size.

   VNS provides redundancy transparent to the network layer protocol by
   managing the network of trunks independently of the network layer.
   VNS will automatically discover any topology changes and re-route
   traffic accordingly.

VNS eases network administration by dynamically keeping track of
which trunks are available for each LNN. Network administrators
don't have to configure VNS or network layer addresses on a per link
basis. Network layer addresses only have to be assigned on a per
Logical Network basis. For nodes which will only be tandem VNS
nodes, network layer addresses aren't required at all.

Since VNS traffic is constrained within an LNN, administrators have
control of where VNS traffic is allowed to flow.

Finally, VNS supports switching of several Upper Layer Protocols and
supports several media (cell and Frame) or a mixture thereof.
Switching in the core of the WAN removes the need for routers and
improves the performance due to a reduction in the number of fields
that need to processed.

12. Security Considerations

Logical networks provide a means of restricting traffic flow for
security purposes. VNS also relies on the inherent security of the L2
media such as an ATM Virtual Circuit.

13. Acknowledgments

The authors would like to acknowledge the valuable comments of Terry
Boland, Pierre Cousineau, Robert Eros, Robert Tomkins, and John
Whatman.

14. Authors' Addresses

    Bilel Jamoussi
    Nortel (Northern Telecom), Ltd.
    PO Box 3511 Station C
    Ottawa ON K1Y 4H7
    Canada

    EMail: jamoussi@Nortel.ca

    Dwight Jamieson
    Nortel (Northern Telecom), Ltd.
    PO Box 3511 Station C
    Ottawa ON K1Y 4H7
    Canada

    EMail: djamies@Nortel.ca

    Dan Williston
    Nortel (Northern Telecom), Ltd.
    PO Box 3511 Station C
    Ottawa ON K1Y 4H7
    Canada

    EMail: danwil@Nortel.ca

    Stephen Gabe
    Nortel (Northern Telecom), Ltd.
    PO Box 3511 Station C
    Ottawa ON K1Y 4H7
    Canada

    EMail: spgabe@Nortel.ca

15. Full Copyright Statement