White Paper
Evaluating High Availability Mechanisms
Introduction

The Need for High Availability

Improving network availability is critically linked to the business success of any service provider. A December 2004 survey of European and North American service providers by research group Infonetics revealed that their top three business priorities were revenue growth, profitability and competition.  

Revenue growth will come largely from deploying value add services such as Triple Play real time services (IPTV, VoD and VoIP), mission critical business applications and VPN. These services place stringent availability demands on the network to meet the user service expectation and work as well as the mechanisms they replace. For example the PSTN service is renowned for its “five nines” reliability corresponding to less than five minutes of outage per year.

Expense management (capital and operating) is crucial to business profitability. Service providers are moving aggressively to rationalize network services onto a converged IP/MPLS infrastructure, eliminating the need for multiple networks and redundant infrastructure. As a result, the converged infrastructure will be responsible for an ever increasing revenue load. To protect against highly critical service outages that greatly impact the bottom line, it must be reliable and support additional capabilities such as hitless in-service upgrades.

The service provider market is a highly competitive landscape. Each provider’s competitive strategy is unique with varying business models that rely on various services and technologies to realize a successful business. Successfully competing requires continuous network availability that meets customer expectations and builds a solid reputation. Network availability can be exploited as a competitive differentiator and in a world where metrics are critical, availability can be relatively easily tracked and audited.

Network Reliability and Service Availability

Improved network reliability and rapid recovery or repair times are a prerequisite for high network service availability. Reliability in the context of a communications network identifies the probability that the network will continue operating as it is intended to for a specified period of time. Network availability is a measure of the network users being able to access services and can be defined using a formula based on two variables.

\[
\text{Network Availability} = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}}
\]

Mean time before failure (MTBF) is a measure of network reliability. Mean time to repair (MTTR) measures the time it takes for the network to recover and return to full operational status in the event of a failure. This formula shows that increasing the time between failures and reducing the time to recover from failures will have a positive impact upon network availability. High availability techniques need to address both network reliability through improved fault tolerance and network recovery time.

1. Infonetics. “Service Provider Plans for IP, MPLS, and ATM: North America & Europe”. Dec 2004
2. 99.999% availability with a 1 day period, equates to less than 5 minutes of downtime.
The Causes of Network Outages

Network service outages can be attributed to many issues. A 2002 study undertaken by Network Strategy Partners LLC (NSP), entitled Reliable IP Nodes: A Prerequisite to Profitable IP Services categorized and quantified this based on research undertaken by the University of Michigan and interviews with Tier 1 carriers. Outages were loosely organised into planned and unplanned outages. Planned outages were those resulting from hardware and software maintenance, and upgrades to network nodes. An unplanned outage was attributed to an unexpected failure whilst in operation such as software, hardware, power and link failures.

Network Failure Types

![Network Failure Types Diagram]

Planned Maintenance & Upgrades

Performing planned maintenance and upgrades to deployed routing nodes is a high risk activity. Network upgrades are among the most common source of outages because new defects are easily introduced to a network during software updates or when hardware replacement procedures are followed incorrectly. Many vendors now support the ability to upgrade router software while the device is in service, without downtime or dropped packets. This feature is often called “hitless upgrade” or “in-service upgrade” and relies upon one or more high availability mechanisms such as redundant control planes, non-stop forwarding and graceful restart.

Unplanned Node Failures

According to the NSP report, unplanned node failures account for approximately 39% of network outages and service disruption. Of these around one-third are hardware related and two-thirds software related.

Hardware

Hardware failures are those caused by the failure of a physical system component such as a route controller, line card, chassis or power supply. Routers and switches typically offer redundancy for the various system hardware components that are activated in the event of a failure of the primary system. Unfortunately the switchover time cannot be assumed to be negligible and may result in minutes of service disruption to network users.

Software

Software failures are those caused by the failure of a system software process or set of processes. Software failures may either affect the forwarding plane, control plane or other auxiliary system such as the management system of the node. Many of the high availability techniques explored in this paper, such as non-stop forwarding, graceful restart and non-stop routing attempt to mitigate the impact of software failures.

Link Failures

Unplanned link failures are attributed to a failure in the underlying transport layer of an IP/MPLS network and include fibre or cable cuts and transport equipment failures. Whilst not a direct failure of the routing nodes in the network they do lead to service outages being experienced by network subscribers and as such must be dealt with. High availability techniques such as MPLS fast-reroute and SONET APS are critical to ensuring the impact of link failures is minimised.

The Costs of Network Outages

Many of the costs to service providers of outages and downtime are the tangible, easily measurable costs such as direct costs of service restoration and equipment upgrades, loss of revenue during outages and penalties associated with customer SLAs violations. However, the “real” costs include losses that are harder to quantify but may be far greater. For example, consider the consequences of lost revenue from disgruntled customers moving or taking new business to competitors, the cost of a tarnished image and the lessened ability to credibly market future “premium” differentiated services and position them against competitors.

The following table is taken from an analysis presented in the NSP report mentioned earlier. It attempts to put some numbers to the cost of network outages. Of course obtaining exact data for these costs is always challenging given the corporate sensitivity, however the use of accurate estimates and analysis coupled with appropriate interpretation allows it to serve as a very useful data point. Three classes of outage are defined, each with an unplanned downtime cost per minute. An estimate is made of the duration of downtime and the number of outages expected for a traditional dual router configuration. Using statistical analysis a figure is derived for the number outages per year when High Availability mechanisms are deployed. From all of this the costs are calculated and compared.

<table>
<thead>
<tr>
<th>Severity Level</th>
<th>Downtime Cost $ (per min)</th>
<th>Down X (hours)</th>
<th>Dual Router Configuration</th>
<th>Single Reliable IP Node</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Outages</td>
<td>Down X Cost $</td>
</tr>
<tr>
<td>1</td>
<td>1,000,000</td>
<td>2</td>
<td>4.00</td>
<td>8,000,000</td>
</tr>
<tr>
<td>2</td>
<td>100,000</td>
<td>2</td>
<td>1.00</td>
<td>200,000</td>
</tr>
<tr>
<td>3</td>
<td>10,000</td>
<td>2</td>
<td>285</td>
<td>5,700,000</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>290</td>
<td>13,900,000</td>
<td>20.09</td>
</tr>
</tbody>
</table>

Figure 2: Unplanned network outage costs (per year)

Many conclusions can be drawn from this table however the most striking and relevant to this discussion is that high availability techniques deployed in reliable IP nodes can reduce the cost of network outages by over 90%. This reduction supports a very strong business case for deploying high availability techniques without considering other potential capital and operational expense advantages achieved through of reducing the number of routing nodes in the network.

**Traditional Network Resiliency**

**SONET SDH**

SONET/SDH Automatic Protection Switching (APS) is a layer-1 mechanism for the protection of links against fibre cuts and excessive bit error rates. Originally designed for voice services, APS offers 50-millisecond restoration using diverse network paths. For 100% restoration, “1+1” protection is employed; traffic is simultaneously sent on working and standby paths, and the receiver selects the path to use.

An alternative scheme that offers partial restoration but uses fewer resources (and is therefore cheaper) is called “1:N” protection. A single standby path, carrying only an idle signal, protects multiple working paths. Failover and restoration can be signalled using a protocol carried on the SONET/SDH K1 and K2 overhead bytes. Carrier-class routers now integrate APS into optical line interface cards to eliminate the need for additional SONET network elements and to add protection against line card failures. There are also transport-layer mesh restoration schemes based on APS that have already been deployed.

**RPR**

Resilient Packet Ring (RPR) combines the benefits of SONET/SDH reliability and 50ms restoration with Ethernet’s efficiency and low cost. Standardized in IEEE 802.17, RPR is based on technology familiar to carriers and increases bandwidth utilization of regional and metro fibre rings, whilst retaining resiliency achieved with duel rings. Many network device vendors already provide RPR or RPR-like products, including Packet Ring interfaces for metro aggregation routers that enable carriers to optimize their existing SONET/SDH backbones for new data services.

**Hot Standby Control Cards**

Today’s carrier-class routers offer redundancy for all hardware components. Power supplies, fans, switch fabrics and line cards can all be backed up by “hot standby” technology. Control cards (also known as route processors or routing engines), which process routing protocol information and maintain routing tables, can also be backed up. The purpose of backing up hardware components is to minimise service disruption in the event of a hardware failure.

Putting this into perspective, the objective of a hot-standby control card is to prevent the need to physically swap out a card in the event of a failure. It is not intended to provide 50 ms switchover time. Typically, the backup control card is idle during normal operation of the primary control card. When the primary card fails all of the session and state information is lost as the backup card goes through a lengthy boot procedure. This usually involves layer-2 configuration and link reestablishment, TCP and routing session reestablishment with adjacent routers, and reconvergence. This process can take in the order of five to ten minutes.
Protocol Reconvergence

IP Routing protocols include reconvergence algorithms to deal with node or link failures. Unfortunately, these mechanisms can take up to tens of minutes to reach a consistent view of the network’s topology after a fault because of routing advertisement oscillations (route flaps) during the complicated path selection process. Depending on the age and type of the router and its software, during this time forwarding can be halted. In a large network, the storm of routing protocol messages and processing overhead can bring a network to its knees. Faster route processors and routing enhancements (such as route flap “damping”\(^5\) to reduce route oscillation) attempt to speed reconvergence but at the same time, networks (and therefore route tables) are continually becoming larger and the need for no service disruption critical.

During reconvergence, the router relearns the routing topology and rebuilds its routing table (also known as the Routing Information Base or RIB). This procedure involves the exchange of potentially thousands of route update messages using a protocol such as OSPF or IS-IS (for routers within the same Autonomous System) or BGP4 (for routers in neighbouring domains). Once the routing table is synchronized, the router can calculate the forwarding table (also known as the Forwarding Information Base or FIB) for its line cards by calculating the best path to each destination. Packet forwarding can then resume. In reality, the FIB may be recalculated many times during reconvergence and partial routing states may be re-advertised. This entire process can take from five to 15 minutes\(^6\) or much longer if there is significant route flapping or if manual intervention is required. Clearly, this is inadequate for a carrier-class router. “Five-9s” availability (99.999%) requires five minutes or less of downtime per year!

High Availability Technologies

The problem with traditional IP/MPLS networks is that planned or unplanned service disruptions result not only in a single node outage, but also the propagation of a disruptive ripple through the entire network as routes flap, forwarding is delayed and packets are dropped. Many minutes of service disruption, performance degradation and lost revenue transpire before the network reconverges around the failure. These consequences and the associated recovery time are unacceptable in a network environment delivering converged services where the IP/MPLS core is shouldering a far greater revenue load.

To solve this problem, HA technologies are being deployed to minimise the impact of outages. This involves two key components; improving fault tolerance and reliability within the network and reducing the time required to restore the network to full operational status after an outage occurs. HA technologies typically decouple the control and data planes of the router so that in the event of a control plane outage data forwarding can continue and revenue loss is avoided, thus improving fault tolerance. Graceful Restart and Non-stop Routing techniques for routing protocols work to eliminate the reconvergence ripple and minimise the recovery time. However, these HA mechanisms must in themselves be bullet proof so that they provide a net improvement to network reliability and this is where testing becomes critical.

Non-stop Forwarding

Most core routers have taken a significant leap forward by maintaining the router’s forwarding table and link states and continuing packet forwarding during route processor failure and restart. During control module failover, layer-2 link state information is maintained for ATM, Frame Relay and Ethernet links, and the router continues to forward packets over routes that were available on the last-known state of the network. Failover is nearly instant and there should be little or no packet loss.

Regardless of other resiliency mechanisms used, Non-stop Forwarding offers a big advantage because it prevents or reduces packet loss during reconvergence. The failed router must still relearn the network topology and recalculate its routing and forwarding tables, but while this is occurring, it continues to forward packets according to its existing forwarding table.

Critics of Non-stop Forwarding call it “headless forwarding” and claim that during failover, other network changes or failures could cause the failing router to ignore route update messages. The router would continue to forward packets according to its “stale” forwarding table, potentially resulting in forwarding loops, routing “black holes” and misdirected or discarded packets. Whether this is a likely or serious scenario depends on the frequency and magnitude of network route changes.

**Graceful Restart**

Activity within the IETF has resulted in the standardisation of protocol extensions to BGP, OSPF, IS-IS and MPLS that build upon “plain” Non-stop Forwarding and address some of its limitations. Known as “Graceful Restart”, these enhancements enable a router to stay on the forwarding path even using its last-known, “frozen” forwarding table as its routing software restarts.

Without Graceful Restart, the neighbours of a restarting router remove the restarting router from their forwarding paths and reconverge on alternative paths, potentially causing a storm of route update messages and route flapping throughout the network. However, if the restarting router and all its immediate neighbours use the Graceful Restart procedures, the restarting router can rediscover its neighbours and relearn its routing table from them without triggering them to start reconvergence.

Graceful Restart relies on the cooperation of neighbour routers, which must implement the same protocol extensions as the restarting router. If just one neighbour does not know these extensions, then the fallback is to use the “regular” (slow) restart procedure. Therefore, service providers using routers from multiple vendors or interfacing with other provider networks will be very keen to test router interoperability and provider inter-working.

Route reconvergence is accelerated during Graceful Restart because route updates are communicated in a single block to the restarting router without interruption. There are no repeated recalculation of the FIB; the restarting router calculates the new FIB only once, after all updates have been communicated and acknowledged. Another advantage is that route updates are localized or isolated within the restarting router and its immediate neighbours. This prevents route flapping and network-wide instability.

Graceful Restart is unfortunately not a perfect solution. If the network topology changes (for example, if there is another network failure) during the restart procedure, the restarting router may not be able to change its forwarding table quickly enough (or at all during restart), causing its entries to become “stale”. To avoid this, a router employing Graceful Restart techniques will typically revert to “regular” restart procedures if a significant topology change is detected by neighbours before the restarting router is ready. This may be preferable but is obviously still not ideal. Because both scenarios are undesirable and could affect availability, service providers and equipment manufacturers will want to test the likely impact on their networks.

Carrier routers will typically need Graceful Restart for multiple routing protocols (BGP, OSPF and IS-IS) and MPLS (LDP and RSVP). It has been suggested that there could be undesirable interactions or dependencies between different routing protocols as they attempt to restart. Specifically, if the Interior Gateway Protocol (OSPF or IS-IS) is unstable or flaps during restart, it may trigger BGP to recommence its restart or to fallback to a “regular” restart. Service providers may wish to simulate and test this in their own evaluation laboratories.
Non-stop Routing

Routing devices are now being built with fully-redundant, “pre-booted” backup control cards. The backup control card maintains a complete copy of routing state information carried by the primary, including the routing table and routing sessions.

During failover, the router continues communication with its neighbours and persists in maintaining up-to-date routing and forwarding tables. There is little or no disruption to routing protocol interactions with the network. There is no need for a lengthy route reconvergence, no “flooding” of the network with update messages, and no packet loss. This approach minimizes the need for duplicate hardware configurations – that is, only one router is needed in each POP.

Non-stop Routing enables new routes to be learned and FIBs updated immediately following failover, avoiding the possibility of stale forwarding table entries, potential routing loops and “black holes”. Control module restarts (following a failure or a software upgrade) are a major cause of router downtime, so Non-stop Routing should significantly reduce network outages.

There are a couple of ways to implement a non-stop routing solution. The first uses mirrored routing cards, wherein the backup card runs the same processes as the primary card in lockstep – similar to telephony switches. The advantage of this approach is simplicity and instant or near-instant failover, avoiding the possibility of stale forwarding table entries, potential routing loops and “black holes”. Control module restarts (following a failure or a software upgrade) are a major cause of router downtime, so Non-stop Routing should significantly reduce network outages.

MPLS

Fast Reroute

MPLS Fast Reroute (FRR) is a technology that can protect against both link and router failures by providing a given label-switched path (LSP) with a pre-established backup LSP. Generally speaking, when a failure is detected, the router immediately upstream from the failure quickly switches all traffic to the backup tunnel. At the same time, it notifies the head-end router at the start of the LSP so that a new, optimal path can be established.

Fast Reroute provides a temporary detour for traffic to circumvent lengthy outages and potential forwarding loops during reconvergence. Manufacturers are currently claiming LSP failover times ranging from 200 milliseconds down to as little as 5 milliseconds. A common “standard” is 50 or 60 milliseconds, in line with SONET/SDH restoration times. In contrast, IP shortest path first (SPF) rerouting can take in the order of several seconds or much longer in large networks.

Fast Reroute is not a complete solution for node protection. It cannot help if a network’s ingress or egress router fails. A further challenge is introduced when an LSP crosses an inter-carrier boundary, making it difficult to provide backup LSP tunnels across carrier borders. Fast Reroute does not prevent the need for route reconvergence following a failure (plus a second reconvergence after node recovery), or the associated storm of route update messages and potential route flapping. Furthermore, some implementations may include proprietary behaviour and that present interoperability concerns in a multi-vendor network or between peer networks for carrier interconnect.

Make Before Break

Make Before Break (MBB) is a component part of Fast Reroute. Specifically, in the event of a topology change due to a node or link failure, a head-end LSR will recalculate the optimum path for an LSP based on updated traffic engineering parameters advertised by the IGP. Once this path is calculated the LSP will be signalled and once tunnels are established the LSR will switch traffic to these paths.
High Availability Deployment

Service Providers

The deployment of high availability mechanisms in carrier networks is occurring now and will continue to grow as the technologies mature and the fundamental market drivers of network convergence and triple play services reach full potential. Direct carrier feedback gathered by Infonetics revealed significant growth in the deployment of HA mechanisms from 2003 to 2004 and looking ahead to 2005.\(^7\)

Figure 5: Infonetics HA deployment survey

The graph above shows the percentage of carriers who indicated they have, or plan to have, the given HA technology deployed within their network. The two interesting conclusions that can be derived from the graph are firstly that we are seeing a move away from traditional resiliency techniques towards newer high availability mechanisms and secondly that a mix of HA technologies will be deployed to achieve five nines reliability within IP/MPLS networks.

From 2004 to 2005 there will be a marked increase in the deployment of all major HA technologies (Non-stop routing, graceful restart, MPLS fast re-route and hitless software upgrades). As these technologies mature and their implementations stabilise, their widespread deployment is inevitable as service providers look for the most efficient way to build reliability into the network. A mix of high availability mechanisms will be needed to ensure that the various causes of failure are covered, whether they are planned or unplanned, hardware or software.

What becomes critical in any case is the need to test. It is only through testing and measuring that it is possible to know that HA technologies are stable and solve the problem they were intended to. When the complexity of a multi-service, multi-protocol converged network, running various HA mechanisms is considered, it becomes abundantly clear that testing is of paramount importance. Simulation and testing must be performed to identify the optimum network configuration and verify that no adverse interactions are caused by the various HA mechanisms at play.

Enterprise

These trends are further reinforced by a more recent survey of the purchasing drivers for new routers within large North American enterprises covering verticals including finance, healthcare, manufacturing and government (all are responsible for managing WANs).

Figure 6: Purchasing Drivers for new routing equipment.\(^8\)

Over 71% of respondents indicated that directly decreasing network downtime is a key driver for purchasing new routing equipment. Looking further at the remaining drivers they too can also be extrapolated in some cases to having an impact on network availability and satisfying customer SLAs.
High Availability Verification

The test methodology used in the HA verification process will greatly affect the quality of the solution deployed in the network and the experience network users encounter. A poorly or inadequately tested network is likely to fail to meet recovery expectations at some point in its deployed life, however thorough testing based on sound process will provide a complete picture of understanding and evidence that the HA solution scales and performs robustly.

Non-Stop Forwarding

Verification of Non-stop Forwarding requires test equipment that tightly integrates IP packet generation and performance measurement with the simulation of large networks through emulation of multiple routing protocols.

To verify a router’s ability to maintain packet forwarding during failover, first advertise a set of routes (using OSPF, BGP4 or IS-IS) from the test equipment, send traffic on those routes through the System Under Test (SUT), establish that traffic is being forwarded with zero or negligible loss, and baseline the latency and latency variation. Failover can then be initiated by extracting a primary control card or by forcing the routing engine or routing processes to restart. Continue to test forwarding performance until recovery is complete. Measure the packet loss and time-to-recovery of packet forwarding; in theory, there should be no packet loss. Compare the latency and latency variation measured during the test against baseline performance to determine the potential impact on SLAs.

A similar method can be used to evaluate “hitless upgrade”. Instead of restarting or extracting a control card, perform a “hot” upgrade of the router software during the test and measure performance degradation during the upgrade.

Figure 7: Testing Non-stop Forwarding

Graceful Restart

Graceful Restart (GR) requires both the restarting router and its immediate neighbours implement Graceful Restart procedures and routing protocol extensions. Therefore, it is important to test a “complete system” – including both a restarting router and two or more cooperating neighbours (sometimes known as receiving routers).

Graceful Restart can only be comprehensively verified and measured using test equipment that supports Graceful Restart extensions. GR extensions allow the tester to connect directly to the SUT and act as cooperating nodes without the need for intervening routing nodes. This allows for the most accurate measurement of restart time using the direct monitoring of protocol session states or timestamp analysis of protocol exchange packets. It also means that test bed size, cost and complexity can be minimised.

Graceful Restart Scenario 1 – Continuous Forwarding

To test continuous forwarding during Graceful Restart, configure a test bed with the restarting router (DUT) connected to three test ports TP1, TP2 and TP3. The test ports are configured with Graceful Restart enabled in Cooperating Neighbour mode.

Simulate a network behind destination ports TP2 and TP3 by advertising routes using OSPF, BGP4 or IS-IS. Setup link costs (weights) such that the default route to the simulated network is via the low cost path TP1-DUT-TP2 and the alternative route TP1-CN1-TP3 offers a higher-cost path to the same simulated network.

As in the Non-stop Forwarding test, advertise a large number of routes through the SUT, generate test traffic from TP1, establish traffic forwarding, and baseline the packet performance (loss, latency and latency variation) to TP2. Force a control card failure or restart to occur in the DUT (or perform a live upgrade of the router software). Ensure continuity of forwarding to TP2 and measure IP performance both during the restart procedure and at the end of restart (when the FIB is recalculated).

As a final check, ensure that no traffic was rerouted at any time via the alternative high-cost path (i.e., to TP3).

Figure 8: Testing Graceful Restart – Continuous Forwarding
**Graceful Restart Scenario 2 – Delayed Reroute & Restart Duration**

A similar configuration is used to test the impact of an external topology change on Graceful Restart and to measure restart duration.

Begin by following the same steps as for testing continuous forwarding. Simulate a network behind test ports TP2 and TP3 and advertise the high and low cost routes. Generate IP test traffic from TP1 to TP2 and measure the baseline forwarding performance.

Following failure or upgrade of the restarting router, immediately withdraw some routes (on the low cost path) at TP2. While the system is busy “gracefully restarting”, this topology change should have no effect on forwarded traffic. The restarting router will use its “stale” forwarding table to continue to forward all traffic to TP2. The restarting router will not yet have knowledge of the withdrawn routes or if it does, it will not use this knowledge to recalculate its FIB until Graceful Restart is complete.

When restart is complete, the DUT will learn the topology change from TP2. It will then rebuild its FIB and re-advertise the topology change to its neighbours TP1 and TP3. Traffic on the withdrawn routes will now be forwarded by the restarting router to the high-cost path via TP3. Check that traffic on those routes is being forwarded to TP3 and compare the IP packet performance to the baseline.

Restart duration, can be measured by monitoring the protocol state, noting the time at which the restart completes and the peer returns to regular operating state. This measurement can also be made by observing the time taken for the first packet to arrive at TP3 following commencement of restart. “Effective” packet loss can also be calculated by counting the packets that were arriving at TP2 following the withdrawal of their routes.

---

![Figure 9: Testing Graceful Restart – Delayed Reroute](image-url)
**Graceful Restart Scenario 3 – Graceful Restart Abort**

This test is used to verify that the SUT will revert to a regular (no-graceful) protocol restart in the event that graceful restart is aborted due to timer expiry, cooperating neighbour failure or other trigger.

Configure the test in the same way as the Non-stop Forwarding test case, connecting three test instrument ports to the router. Simulate a network behind destination test ports TP2 and TP3 by advertising a large number of routes using OSPF, IS-IS or BGP4 from both test ports. Configure the routes through TP3 as higher-cost paths so that by default, all traffic will be routed to TP2. Load each route with IP test packets from TP1 and establish baseline packet performance to TP2 (see Figure 12). Following failure or upgrade of the restarting router, force a failure with a cooperating node. This should cause the restarting router to revert to a regular restart.

![Diagram](image-url)
**Graceful Restart Scenario 4 – Multi-protocol Scaled Graceful Restart**

This test verifies forwarding plane continuity and control plane recovery of a SUT during multi-protocol, graceful restart, or fully redundant (hot) switchover. The SUT is typically a PE/P restarting router connected to many test ports that simulate both edge and core networks. The test is scaled across multiple ports, and multiple protocols are tested concurrently using mixed IP/MPLS traffic to accurately reflect a realistic carrier network. For example, a typical PE router may be peering with many hundreds or perhaps thousands of routers. Similarly, a P router may be acting as a route reflector and also have many peering relationships. During a restart, this will present these devices with an extremely large number of restart messages that need to be processed. It is critical to verify that this load can be reliably handled to ensure a successful restart that minimises service disruption.

Routing protocols (OSPF, ISIS, and BGP4) are used to advertise simulated networks on either edge and core sides of SUT while MPLS protocols (RSVP and LDP) are restricted to the core side networks. IP traffic is sent between multiple OSPF, ISIS, and BGP-4 edge/core networks and MPLS traffic is sent between RSVP and LDP core networks. After taking baseline performance measurements, a failover situation is initiated in the SUT. Performance measurements are continued during the SUT restart period, recovery period, and post-dwell period. The test passes if restart successfully completes within the required period and packet loss for each protocol is less than the minimum acceptable packet loss, MPLS labels remain unchanged, and the SUT completes the restart process within the user specified times.

Figure 11: Multi-protocol scaled Graceful Restart
Graceful Restart Scenario 5 - Conformance Test

Most, if not all, live networks are multi-vendor environments that utilise several different protocol implementations. It is even possible to find multiples device from a given vendor using different protocol implementations. To reduce the total time required to deploy graceful restart and to minimise the likelihood of interoperability issues it is critical to test the conformance of Graceful Restart functionality to the relevant IETF specification or draft using an independent reference.

Conformance testing subjects a DUT to a multitude of functional test cases that carry pass or fail results. Typically the DUT is connected to one or more test ports (depending on the conformance test case) and a series of test cases are executed. Each test case is a stimulus and response message exchange sequence. The tester will send a message to the DUT then wait for the response. The response is analysed and verified to be as expected.

Figure 12: Graceful restart conformance.
Non-stop Routing

Of all the High Availability technologies described in this paper, Non-stop Routing is arguably the most difficult to measure. If it works perfectly, the control plane failover is almost instant and undetectable, and there is nothing to measure. However any problems with the operation of a non-stop routing implementation will show up in a corrupted forwarding plane.

To verify expected operation and performance of Non-stop Routing, configure the test topology in the same way as the Non-stop Forwarding test case, connecting three test instrument ports to the router. Simulate a network behind destination test ports TP2 and TP3 by advertising a large number of routes using OSPF, IS-IS or BGP4 from both test ports. Configure the routes through TP3 as higher-cost paths so that by default, all traffic will be routed to TP2. Load each route with IP test packets from TP1 and establish baseline packet performance to TP2 (see Figure 14).

The next parts of this test will require some automation within the test instrument. Start withdrawing routes from TP2 at a constant rate (see figure 14A). As each route is withdrawn, after a short “rerouting delay”, the router will forward traffic to the alternative, higher-cost path (TP3). At the same time, measure packet performance at TP3. In particular, graph the bit rate or bandwidth of packets arriving at TP3. You should begin to see a straight line, perhaps with some small “bumps” caused by variations in the rerouting speed.

Force control card failover by extracting the router’s primary control card, restarting its routing processes or initiating software upgrade. Continue to withdraw routes from TP2 and continue to measure packet performance at both ports. If the router’s failover to the backup control card is almost instantaneous (sub-second) and seamless, rerouting performance will be unaffected and the graph of packet bandwidth will continue along a relatively straight line (see dashed line in figure 14B). If the router takes more than a few seconds to failover and to establish routing processes on the backup control card, then a plateau will be observed in the second graph during the routing engine’s “downtime” (see solid line in figure 14B). The duration of this plateau is a good estimate of the time taken for the router to failover to the backup control card. When routing processes are fully re-established, traffic on the withdrawn routes should be quickly redirected to TP3 and the packet bandwidth at TP3 should “catch up” to the expected rate. Ensure that this “catch up” does occur – that is, ensure that all route withdrawals were processed and none were missed.

Although this test scenario was designed for testing Non-stop Routing, it can be used to test the impact of other resiliency technologies on routing performance.

Figure 13: Testing Non Stop Routing

Figure 14A: Routes to TP2

Figure 14B: Packets arriving at TP3

Figure 14A: Figure 14B:
Evaluating High Availability Mechanisms

MPLS

MPLS Test Scenario 1 - Fast Reroute

Testing MPLS Fast Reroute is a complex procedure that requires a test instrument that tightly integrates simulation of network topologies (using OSPF or IS-IS), emulation of MPLS signalling protocols, and generation and performance measurement of labelled test traffic. For test scenarios such as this that require multiple protocols, an automated application is generally preferred.

To measure the time it takes to redirect traffic to a backup LSP tunnel after the primary LSP has been dropped, three router test ports are needed: one source port, a destination port for the primary LSP, and a second destination port for the backup LSP.

Begin by simulating a common network topology on the two destination ports of the router under test using OSPF or IS-IS. Establish the LSP tunnels and generate traffic on the source port through the router. Physically sever the link between the SUT and the primary destination port to force failover from the primary to the backup LSP tunnel. Finally, measure the switchover time to reroute traffic to the backup LSP and destination port.

When MPLS reroutes traffic to a backup LSP tunnel, the backup LSP should respect the same Class of Service (CoS) priorities configured on the primary LSP, critical for maintaining customer SLAs. Configure the SUT and traffic source with several different streams and repeat the above test, confirming that the backup LSP maintains the relative stream priorities.

Figure 14: Verifying MPLS Fast Reroute

- Laser off (Link and Node Protection)
- Bring down PPP (Link and Node Protection)
- Withdraw LSA (Node Protection Only)
- SONET Alarm (Line and Node Protection)
**MPLS Test Scenario 2 - Make Before Break**

This test measures the time it takes a system under test (SUT) to converge and redirect traffic to a more preferential Label Switched Path (LSP). Although possible to do with 3 ports, this test is best completed with 4 ports and the SUT acting as an LER, because it is the headend router that performs path re-optimisation. The source sends IPv4 traffic (non-MPLS) into the network and the SUT acts as head-end. Outgoing ports are the primary path, backup path, and re-optimised path. When the primary path is no longer available the SUT will fast re-route to the backup, then after using CSPF algorithms calculate and switch to the re-optimised path.

Begin by simulating a common network topology on the two destination ports of the router under test using OSPF or IS-IS. OSPF/ISIS topology needs to make primary least costly, backup most costly. Establish the LSP tunnels and generate traffic (IPv4, non-MPLS) on the source port through the router. Physically sever the link between the SUT and the primary destination port to force failover from the primary to the backup LSP tunnel. After the fast re-route the SUT should perform a CSPF calculation to determine if the backup tunnel is optimal. Based on the IGP settings it will find a new path. The SUT should then establish a new LSP tunnel along this optimal path and redirect traffic from the primary LSP to this new, more preferential LSP. Convergence time measurements are then taken and a check for no packet loss.

![Figure 15: Make before break](image-url)
Service Disruption Testing

This test extends the Multi-Protocol Scaled Graceful Restart or the MPLS Scaled Fast Reroute test to include L3 VPN service delivery via BGP and LDP. The objective is to verify that VPN performance is maintained and traffic is not lost during a failover situation; a critical component of end user SLAs.

Begin by simulating multiple VPN sites on the various test ports using BGP and LDP. IP traffic should be sent between VPN sites. After taking baseline performance measurements, a failover situation is initiated in the SUT. Performance measurements are continued during the SUT restart period to verify that VPN services are not disrupted. The test can be extended to cover other network services including L2 and multicast VPNs.

Figure 16: L3 VPN Service Impact Test
Testing With Agilent N2X

Agilent provides the industry’s most complete High Availability solution with scalable emulation software, a dedicated productivity application automating complex scenarios and a Graceful Restart Conformance Test suite. Equipment manufacturers and services providers can save many months of test engineering effort through the rich set of N2X automation tools. Agilent N2X ensures test realism, scalability and faster time to insight.

Protocol Emulation

Agilent N2X is the only solution to provide comprehensive coverage of all routing and signalling Graceful Restart protocol extensions, MPLS fast reroute and make before break. N2X provides the industry’s most scalable and flexible multi-protocol emulation that enables realistic test simulation of converged networks. Coupled with accurate data plane convergence measurements N2X users can verify device control plane restart performance and scalability with concurrent, multi-protocol graceful restart emulation on a single port. N2X can also characterise the impact of high availability mechanisms on service delivery. By simultaneously emulating L2/3 services per port.

Conformance Test

Agilent N2X is the only solution to provide graceful restart conformance test suites for IPv4 routing protocols (BGP-4, IS-IS and OSPFv2). Conformance testing will accelerate the deployment of graceful restart protocol extensions and minimise interoperability issues by ensuring the vendor implementation supports all features specified. The N2X Graceful Restart protocol conformance suite operates within a unique Test Manager with industry leading usability and reporting features. This allows for rapid result diagnosis and problem isolation saving engineers time and effort involved in developing and validating protocol implementations. The N2X Test manager also has the flexibility to be easily integrated with customer’s regression strategy.

High Availability Productivity Application

The High Availability Productivity Application was developed by Agilent to specifically verify High Availability mechanisms. Building upon many of the unique capabilities offered in N2X, it delivers a rich set of automated test cases designed to maximize testing productivity. The N2X High Availability Productivity Application simplifies and accelerates testing of scaled multi-protocol router and network resiliency scenarios. Agilent has already studied and interpreted multiple IETF documents to develop commonly used test scenarios thereby saving many valuable test engineering person months of time. Furthermore, the application provides rapid test execution and dynamic graphical reporting for quick analysis on highly scaled scenarios.

Conclusion

Driven predominantly by network convergence and triple play services a single IP/MPLS infrastructure will carry multiple (if not all) services and applications. As a result even small periods of network downtime will result in potentially large amounts of lost revenue, diminished customer satisfaction and reduced brand equity. Traditional IP routing solutions delivering recovery times in the order of minutes are no longer acceptable. Network deployment using HA mechanisms is critical for building the necessary reliability into the network.

The deployment of these features and technologies is accelerating amongst service providers and adding further complexity to already complex and highly scaled multi-service networks. Quantifying of the robustness of these HA mechanisms is essential. Test equipment with integrated support for HA technologies such as Agilent’s N2X will give confidence that HA mechanisms are performing correctly.
References

- Network Strategy Partners LLC. “Reliable IP Nodes: A Prerequisite to Profitable IP Services”; 2002.
This page intentionally left blank.