Network Working Group
Internet Draft
Expires August 2001

H.Berkowitz
A.Retana
S.Hares
P.Krishnaswamy

draft-berkowitz-bgpcon-01.txt

March 2000

Benchmarking Methodology for Exterior Routing Convergence

Status of this Memo

Abstract

This is an update of an individual contribution that has been accepted as a work item by the Benchmarking Methodology Working Group, and will split into two BMWG documents. It is being posted for information. T This document defines a specific set of tests that router implementers can use to measure and report the convergence performance of BGP-4 processes. It doe not consider the forwarding performance of such routers once they have converged, or the convergence characteristics of the global routing system.

Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [2].

Table of Contents

1. Introduction

This document describes a specific set of tests aimed at characterizing the convergence performance of BGP-4 processes in routers or other boxes that incorporate BGP functionality. A key objective is to propose methodology that will facilitate the conducting and reporting of convergence-related measurements in a standard fashion. Although both convergence and forwarding are essential to basic router operation, this document does not consider the forwarding performance,if applicable, in the Device Under Test (DUT),for two reasons. Forwarding performance is the primary focus in [1] and it is expected that it will be dealt with in work that

ensues from [1]; further, as convergence characterization is a complex process, we would deliberately like to restrict the initial focus in this document to specifying how to take basic measurements towards this objective.

Subsequent drafts will explore the more intricate aspects of convergence measurement, e.g. in the presence of policy processing and other realistic performance modifiers such as simultaneous traffic on the control and data paths within the DUT. Convergence in Interior Gateway Protocols will also be dealt with in separate drafts.

1.1 Overview and Roadmap

In general, measurements of routing protocol convergence can be classified either as æinternalÆ, with time-stamped tables indicating the time of completion of convergence (such as those described in [4], or æexternalÆ. In an external measurement, a process in the Device Under Test (DUT) is inferred to have converged after a downstream measurement device indicates the corresponding advertisement has been received by it. An alternative type of external measurement is to test for data forwarded to the downstream device that relies upon the route that the Device Under Test just converged upon. The external technique is more readily applicable than the internal technique at present since the requisite NTP timestamp hooks may not yet be in products. However, the external technique is less accurate as it also includes the time to advertise the new route downstream and transmission times for the advertisement. If data forwarding were to feature in the measurement methodology it too would include some extraneous latency- that of the forwarding lookup process in the DUT at the minimum. This document deals only with external measurements limited to route propagation.

A characterization of the BGP convergence performance of a device must take into account, if not also time, all distinct stages and aspects of BGP functionality. This requires that the relevant terms and metrics be as specific as possible. Consequently the first step taken here towards detailing measurements of convergence performance will be to define all the relevant terms and concepts.

The necessary definitions are classified into two separate categories:

. Descriptions of the constituent elements of a network or a router that is undergoing convergence

. Descriptions of factors that impact convergence processes which will influence measurements on convergence.

1.2 Definition Format

The definition format is the equivalent to that defined in [12], and is repeated here for convenience:

X.x Term to be defined. (e.g., Latency)

Definition:
The specific definition for the term.

Discussion:
A brief discussion about the term, its application and any restrictions on measurement procedures.

Measurement units:
The units used to report measurements of this term, if applicable.

Issues:

List of issues or conditions that affect this term.

See Also:
List of other terms that are relevant to the discussion of this term.

2. Definitions of convergence-related router and network states or components

Many terms included in this list of definitions were described originally in previous standards or papers. They are included here because of their pertinence to this discussion. Where relevant, reference is made to these sources. An effort has been made to keep

this list complete with regard to the necessary concepts without overdefinition.

2.1 BGP Peer

Definition:
A BGP peer is another BGP process to which the DUT has established a TCP connection over which a BGP session is active. Peers send BGP advertisements to the DUT and receive DUT-originated advertisements.

Discussion:
This is a protocol-specific definition, not to be confused with another frequent usage, which refers to the business/economic definition for the exchange of routes without financial compensation.

Measurement units:

Issues:

See Also:

2.2 The routing information Base (RIB) and its constituents Adj-Rib-In, Adj-Rib-Out, Loc-RIB

Definition:

These terms were defined in [10]. The RIB contains all destination prefixes to which the router may forward, and one or more currently reachable next hop addresses for them. Routes included in this table potentially have been selected from several sources of information, including hardware status, interior routing protocols, and exterior routing protocols. RFC 1812 [12] contains a basic set of route selection criteria relevant in an all-source context. Many implementations impose additional criteria. A common implementation-specific criterion is the preference given to different routing information sources.

The Forwarding Information Base (see next item) is generated from the RIB.The Loc-RIB contains the set of best routes selected from the various Adj-RIBs,after applying local policies and the BGP route selection algorithm.Adj-RIB-In and Adj-RIB-Out are "views" of routing information from the perspective of individual peer routers. The Adj-RIB-In contains information advertised to the DUT by a specific peer. The Adj-RIB-Out contains the information the DUT will advertise to the peer.

Discussion:

The separation implied between the various RIBs is logical. It does not necessarily follow that these RIBs are distinct and separate entities in any given implementation.

Measurement Units:
Number of route instances

Issues:
Specifying the RIB is important because the types and relative proportions of routes in it can affect the convergence efficiency. Types of routes can include internal BGP, external BGP, interface and IGP routes.

See Also: Route, BGP Route, Route Instance

2.3 The Forwarding Information Base or FIB

Definition: The FIB is referred to in [10] as well as [12] but not defined in either. For the purposes of this document, the FIB is the last lookup on the router data path, based on which a next hop is selected for forwarding each packet.

Discussion: Most current implementations have full, non-cached FIBs per router interface. All the route computation and convergence occurs before a route is downloaded into a FIB.

Measurement Units: N.A.

Issues:

See Also: Route

2.4 Default Free routing tables

Definition:
The size of routing tables in the default free zone of the Internet.

Discussion:
The term originates from the concept that routers at the core or top tier of the Internet will not be configured with a default route (Notation 0.0.0.0/0). Thus they will forward every prefix to a specific nexthop based upon the longest match.
Default free routing table size is commonly used as an indicator of the magnitude of reachable Internet address space. However,default free routing tables may also include routes internal to the infrastructural net that a router is part of.

Measurement Units: number of routes

Issues:
See Also: Routes,Route Instances, Default Route

2.5 Prefix

Definition: A destination address in CIDR format.
Expressed as prefix/length. The definition in [12] is "A
network prefix is..a contiguous set of bits at the more significant
end of the address that defines a set of systems;host numbers select
among those systems."

Discussion: A prefix is expressed as a portion of an IP address
followed by the associated mask such as 10/8.

Measurement Units: N.A.

Issues:

See Also:

2.6 Route

Definition: In general, a æroute Æ is the tuple <prefix, nexthop>. If
MPLS is supported the tuple may include <fec,prefix,nexthop,label>

Discussion: This term refers to the concept of a route common to all
routing protocols.

Measurement Units: N.A.

Issues: None.

See Also: BGP route

2.7 BGP Route

Definition: The tuple <prefix, path attributes> [10]

Discussion: Attributes are mentioned in [10], and are by inference,
qualifying data that accompanies a prefix in a BGP route." For
purposes of this protocol a route is defined as a unit of
information that pairs a destination with the attributes of a path
to
that destination... A variable length sequence of path attributes is

present in every UPDATE. Each path attribute is a triple <attribute type,attribute length, attribute value> of variable length." Nexthop is one type of attribute.

Measurement Units:N.A.

Issues:

See Also: Route, prefix.

2.8 Default Route

A Default Route is a route entry that can match any prefix. If a router does not have a route for a particular packet's destination address, it forwards this packet to the next hop in the default route entry if its FIB contains one. The notation for a default route is 0.0.0.0/0

Discussion: Core routers do not contain default routes. Access and edge routers are likely to have default route entries.

Measurement units: N.A.

Issues:

See Also: default free routing table, route, route instance

2.9 Route Instance

This term is used in the context of a BGP Adj RIB In.

Definition:

Single occurrence of route sent by BGP Peer for a particular prefix. When a router has multiple peers from which it accepts routes, routes to the same prefix may recur in the various Adj-Ribs-In. This is then a case of multiple route instances.

Discussion

Route instances may not be selected by the BGP selection algorithm due to local policy.

Measurement Units:number of instances

Issues: the number of route instances in the Adj-Rib-in bases will vary based on the function to be performed by a router. A core router will likely receive more route instances than an access

router. A core router is situated in the default-free zone.

See Also:

2.10 Unique Route

Definition: A unique route is a prefix for which there is just one route instance.
Discussion:
Measurement Units:N.A.
Issues:
See Also: route, route instance

2.11 Route Views

Definition:

Route views must be further specified as incoming or outgoing. An incoming route view is AFI/SAFI and peer specific and is the Adj-Rib-In for that peer and AFI/SAFI. An outgoing route view is also peer and AFi/SAFI specific and is the Adj-Rib-Out for that peer, for a given AFI/SAFI combination.

Discussion:

Measurement Units: N.A.

Issues:

See Also:

2.12 Policy

Definition:
Policy is "the ability to define conditions for accepting, rejecting, and modifying routes received in advertisements"[16] Policy processing is the set of actions performed by the BGP route selection algorithm that influences route selection in the presence of attributes in the route updates received from peers, or policy actions configured to influence outbound BGP route advertisements.

Discussion:RFC 1771 [10] further defines policy constraints in the hop-by-hop routing paradigm.

Measurement Units:

Issues: Policy is implemented using filters .

See Also: Policy Information Base.

2.13 Policy Information Base

Definition:

A policy information base is the set of incoming and outgoing policies. All references to the phase of the BGP selection process here are made with respect to RFC 1771 [10] definition of these phases.
Incoming policies are applied in Phase 1 of the BGP selection process [10]to the Adj-Rib-In routes to set the metric for the Phase 2 decision process. Outgoing Policies are applied in Phase 3 of the BGP process to the Adj-Rib-Out routes to allow route (prefix and path attribute tuple) to be announced out to a specific peer.

Discussion:

Policies in the Policy information base often instantiated as "route maps" and filter/access lists. The "route maps" often operate on or use the "path attribute" portion of the BGP route. On incoming policy, these "route maps" may set a metric to be compared in Phase 2 of the BGP process.[10] On the outgoing policy, the "route maps" may also set outgoing path attributes to the route sent to the peer.

The filter lists/access lists track the route prefixes.

The amount of policy processing (both in terms of route maps and filter/access lists) will impact the convergence time of the BGP algorithm. The amount of policy processing may vary from a simple policy which accepts all routes and sends all routes to complex policy with a substantial fraction of the prefixes being filtered by filter/access lists.

For this first round of tests for BGP convergence, we recommend that the tests be run under the simple policy of "accept all routes and and send all routes."

2.14 Route Flap

Definition:

RFC 2439 [13]refers to route flapping as

"An excessive rate of update to the advertised reachability of a subset of Internet prefixes.."

We would like to refine this description for the purpose of benchmark specification to be:

"Repeated excessive updates to route instances in the Adj-Rib-In on the DUT."

Discussion:

These repeated updates can be either

a) Implicit replaces of routes [10] categorized in [4] as: either AADiff or AAdup.

b) Explicit replaces of routes [10] categorized by [4] as either: WADiff, WAdup,

c)Erroneous Duplicate Withdrawals for the same route as categorized in [4] as WWDup.

The threshold that can be declared excessive by RFC 2439 [13] is configured by each network on the basis of:

"cutoff threshold (cut)

This value is expressed as a number of route withdrawals. It is the value above which a route advertisement will be suppressed.

reuse threshold (reuse)

This value is expressed as a number of route withdrawals. It is the value below which a suppressed route will now be used again. "

Measurement units

Flapping events per unit time.

Specific Flap events are:

1) AADiff
2) AADup
3) WADiff
4) WADup
5) WWDup

The Flapping event sequence can be characterized as mixture of these events with a percentage per type. An example of this would be:

20% AADiff, 40% AAdup, 30% WADiff 10% WWdup at 100 flap

events per second.

2.15 Convergence

Definition: A router is said to have converged onto a route advertised to it, given that route is the best route instance for a prefix, (if multiple choices exist for that prefix) when this route is advertised to its downstream peers.

Discussion: The best route instance should be set so as to be unambiguous during test setup/definition. This document does not consider forwarding-dependent illustrations of convergence.

Measurement Units: N.A.

Issues:

See Also:

3. Factors that impact the performance of the convergence process

Some of these factors will not be incorporated into the tests in this document. This is because, as mentioned earlier, specifying characterization methodology will be undertaken in stages according to complexity starting with the more baseline tests.

3.1 Number of peers

As the number of peers increases, the BGP route selection algorithm is increasingly exercised. The phasing and frequency of updates from the various peers will have a marked effect on the convergence process on a router.

3.2 Number of routes per peer

The number of routes per BGP peer is an obvious stressor to the convergence process. The number, and relative proportion, of multiple route instances and distinct routes being added or withdrawn by each peer will affect the convergence process. So will the mix of overlapping route instances, and IGP routes.

3.3 Policy processing/reconfiguration

The number of routes and attributes being filtered for, and set,as a fraction of the target route table size is another parameter that will affect BGP convergence.

The two extremes are:

Minimal Policy
Extensive policy--. For example, upto 80 % of the total routes

must have applicable
policy.

3.4 Forwarded traffic

The presence of actual traffic in the router may stress the control path in some fashion if both the offered load due to data and the

control traffic (FIB updates and downloads as a consequence of flaps)are excessive. This is implementation dependent. This condition is a more accurate reflection of realistic operating scenarios than if no data traffic is present.

3.5 Flap dampening

Flap Damping occurs in response to frequent alterations in the route instances input to the DUT. If this is in effect, it requires that the router keep additional state to carry out the damping, which has a direct impact on the control plane due to increased processing.

3.6 Authentication

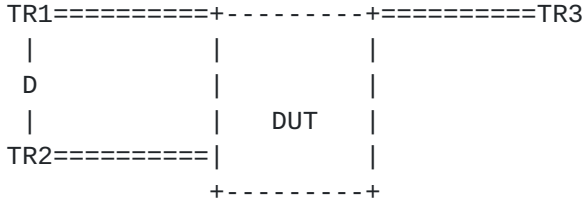
Authentication in BGP is currently done using the TCP MD5 Signature Option [14]. The processing of the MD5 hash, specially in routers with a large number of BGP peers and a large ammount of update traffic may have an impact on the control plane of the router.

3.7 MBGP Processing

Multiprotocol extension for BGP are defined in [15], giving BGP the ability to carry routing information for multiple address families (not only IPv4 unicast). Processing of different protocol information encoded using these multiprotocol extensions may have an impact on the convergence of any one protocol. The tests presented in this document may be applicable to any specific address family.

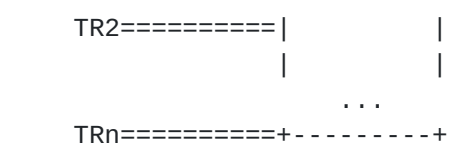
4. Test Configuration

Figure 1 illustrates the single peer test case:



D is a prefix reachable by both TR1 and TR2. It is assumed that neither TR1 or TR2 is the AS of origin for the announcement of D. For all test routers and the DUT, all routes fed in as part of this test process are EBGP routes.

More complex peering arrangements will involve up to n Test Routers, as shown in Figure 2. It is recommended that the Figure 1 configuration always be tested as a baseline, and then additional reports made that show the effect on performance of increasing the number of peers. All tests defined in this document use the topology shown unless explicitly noted.



Interface speeds must be specified as part of the test report. At least a 100 Mbps speed link and a full duplex MAC layer between all connected devices are recommended.
In the absence of other route selection criteria, TR1 shall have an IP address that makes it most preferred.

5. Test setup and methodology

'Test routers' will be providing the test traffic to the Device Under Test and collecting the evidence of convergence from it, if any. The only traffic in the cases described here is route updates/withdrawals. The requisite TCP sessions will have to be established between all test routers and the DUT. Any other equipment required to trace the flow of BGP messages between the devices actually participating in the test will need to be transparent to these sessions. It is also desirable that the 'Test routers' be able to generate protocol message sequences at settable rates.

5.1.1. Stages of convergence and events triggering reconvergence

5.1.1.1 Full Initialization

The DUT establishes a TCP connection, then a BGP session, with a peer, and accepts routes from it. Full initialization of this sort is expected to be relatively infrequent compared with incremental convergence.

5.1.1.2 Incremental Convergence

There are several distinct operations which could be categorized as incremental convergence.
A taxonomy characterising routing information changes seen in operational networks is described in [3] as well as [4]. These papers describe BGP protocol-centric events, or event sequences in the course of an analysis of network behavior. The terminology in the two papers addresses similar but slightly different events. The former refers to Tup,T ,T , and Tlong indicating the
down short ,
occurrence of a route first coming up, being withdrawn,and routes with shorter or
longer ASPATHs respectively. The first two denote explicit events. The last two refer to implicit re-announces of a shorter or longer route.
In [4],the notation used was WADiff (explicit), WADup, AADiff,which is implicit and AADup, also implicit.

With regard to the benchmarking methodology under discussion, we would like to apply the foregoing taxonomies to categorise the tests

under definition where possible, because these tests must tie in to phenomena that arise in actual networks. We avail of, or extend, this terminology as necessary for this purpose. In this document, the meaning of Tup and Tdown are preserved and extended from [3]. The notation Tup(TRx) stands for a Tup event advertised to the router being tested (i.e., DUT). We also introduce Tupinit to indicate the initial announcement of a route to a unique prefix.

{is this used?}The sense of the Tshort and Tlong events is also preserved, but the basic criterion for selecting a "better" route is the final tiebreaker defined in RFC1771, the router ID. As a consequence, this memorandum uses the events Tbetter, Tworse, and Tbest. They are defined as:
Tbest -- The current best path.
Tbetter -- Advertise a path that is better than Tbest.
T -- Advertise a path that is worst than Tbest.
worst

Categories of incremental convergence:
These tests list basic operations that occur on a single router in response to route updates / withdrawals typical of network instabilities. Only the fundamental operations are selected because they form the basis of all more intricate responses. Longer sequences of protocol updates require a compounding of the responses listed here. In addition the arrival rate as well as pattern of route updates/withdrawals is an important factor in the stress

testing of a router's convergence.

--Add single route (Tupinit)
--Delete single route (Tdown)
--Add/deletes of multiple routes in increments until the full table is advertised or withdrawn at once . This could include repetitions of the basic operations of Tupinit

,
WAdiff,WAdup,AAdup,AAdiff
--Delete Peer/Readd
This causes a full convergence type of operation. The test router terminates the TCP connection and BGP session with the peer, then reestablishes the BGP session. When the session is reestablished, routing information must be exchanged again.
--Delete multiple peers and readd.
When multiple peers are sending or receiving routes from the DUT, the percentage of route instances, unique routes, and the total number of routes from or to each peer.
--Failover to an existing less preferred route on withdrawal of preferred route (Wf)

6. Tests measuring Full Initial convergence with a single peer

Procedure:
Initialize the test scenario by establishing an eBGP session between the DUT and TR3. No routing information is exchanged. Initialize TR1 with a predetermined number of prefixes. Suggested fractions are

10%,20%,50% and 100% of the full routing table. The physical link between TR1 and the DUT should also be active at this time.

Establish an eBGP session between TR1 and the DUT; all the prefixes in TR1 should be advertised at this time to the DUT.

The convergence time measurement should start when the first OPEN message is exchanged between TR1 and the DUT. The end of the convergence period is marked when TR3 receives the last UPDATE from the DUT.

It is expected that the DUT will install the routes in its FIB. However, this test will neither check for, nor verify this.

7. Incremental Reconvergence

This set of tests measures the convergence after the initial full BGP table has been transmitted to and processed by the DUT.The test procedures are based on the cases described in section 4.

7.1 Route Announcements

7.1.1. Explicit announce of single new route (Tupinit)[3,4]

This test measures the time required to add a route newly advertised by a peer (Tup(TRx)). Such a route does not exist in the DUT's RIB, and will not displace a route in the RIB.

Procedure :

Initialize the test scenario by establishing an eBGP session between the DUT and TR1 and between the DUT and TR3. TR1 should advertise a predetermined number of routes to the DUT, which in turn should advertise it to TR3.

-Advertise a route originated in TR1; Tup(TR1,D).

--The reconvergence time measurement should start when TR1 sends the UPDATE containing the route D. The end of the reconvergence period is marked when TR3 has received the UPDATE containing D.

7.1.2. Implicit withdraw of single route and replace by new announced route (AAdiff)

This test measures the time required to replace an existing route with one that is preferred (Tbetter(TRx)). Such a route exists in the DUT's RIB, and will be replaced by the new advertisement.

Procedure :

Initialize the test scenario by establishing an eBGP session between the DUT and TR1 and between the DUT and TR3. TR1 should advertise a predetermined number of routes to the DUT, which in turn should

advertise it to TR3. The set of routes advertised by TR1 should contain the test route D.

-Advertise a replacement route for D from TR1; Tbetter(TR1,D).

This route should have LOCAL_PREF value that is preferred over the original advertisement for D.

--The reconvergence time measurement should start when TR1 sends the UPDATE containing the replacement route. The end of the reconvergence period is marked when TR3 has received the new UPDATE containing the replacement.

Variations to this test may consist in selecting other attributes to replace in a consecutive update.The attribute used should be indicated

in the results and no filters should be used.

7.1.3. Duplicate announcments (AAdup)

From [4], this type of event occurs and may be caused by policy changes or flaps within the "MinRouteAdvertisementInterval" of 30 seconds.

7.2 Route withdrawal

7.2.1. Explicit withdraw of single route (Tdown)

This test measures the time required to withdraw a route advertised by a peer (Tdown(TRx)). Such a route exists in the DUT's RIB, and will be removed.

Procedure

Initialize the test scenario by establishing an eBGP session between the DUT and TR1 and between the DUT and TR3. TR1 should advertise a predetermined number of routes to the DUT, which in turn should advertise them to TR3.

Withdraw a route previously originated in TR1; Tdown(TR1,D).

The reconvergence time measurement should start then TR1 sends the withdraw message containing the route D. The end of the reconvergence period is marked when TR3 has received the corresponding withdraw message.

7.2.2. Explicit Withdrawal followed by an reannounce (WAdup)

This test combines 6.2 and 6.3.1.and measures the time required to withdraw ((Tdown(TRx)) and reinstall (Tup(TRx)) a route advertised by a peer. Such a route initially exists in the DUT's RIB, it will be removed and then reinstalled.

Procedure:

Initialize the test scenario by establishing an eBGP session between the DUT and TR1 and between the DUT and TR3. TR1 should advertise a predetermined number of routes to the DUT, which in turn should advertise them to TR3.

Withdraw a route previously advertised by TR1; Tdown(TR1).

After a predetermined ammount of time, TR1 readvertises the same withdrawn route (Tup(TR1)) to the DUT.

The reconvergence time measurement should start then TR1 sends the withdraw message containing the route D. The end of the reconvergence period is marked when TR3 has received the UPDATE containing D.

7.2.3. Failover to existing Alternate Path after Explicit Withdrawal (no announce WF)

This test measures the time to replace a path with an existing alternate after an explicit withdraw (Tdown(TRx)) of the current best path (Tbest).

Procedure:

Initialize TR1 and TR2 with a predetermined number of routes. These routes should be for the same prefixes.

Initialize the test scenario by establishing an eBGP session between the DUT and TR1, TR2 and TR3. TR1 and TR2 should advertise their routes and the DUT should advertise the best path to TR3.

The routes advertised by TR1 and TR2 should be such that the DUT selects the path through TR1 as the best. The decision should be made by comparing the LOCAL_PREF between the two available paths. At this point the DUT should have a path from both TR1 and TR2 for every prefix.

TR1 sends a withdraw for a specific route (D); Tdown (TR1,D).

The reconvergence time measurement should start when TR1 sends the withdraw message for D. The end of the reconvergence period is marked when TR3 receives the new UPDATE containing the path through TR2.

This test may also be executed by increasing the number of routes withdrawn by TR1 or by increasing the number of alternate paths available(increase the test routers up to TRn).

7.2.4. Explicit withdraw of an existing route followed by announce of a different route (WAdiff)

7.3 Repetitive route updates (flaps)

Once the basic protocol update responses have been calibrated,longer event sequences must be tested for. These sequences may look like AwAdiffwadupAAdup..and occur at, eg, 300 per second.Announces will be more overhead intensive than withdraws.

8. Multiple Peers

8.1 Initial Convergence

This test is similar to the single peer initial convergence time, but the number of external peers should increase. All peers are expected to advertise the same number of routes to the DUT.

A ratio of n paths per prefix may be considered such that the first n neighbors must advertise the exact same prefixes (only the AS_PATH should be different). If the number of eBGP peers tested goes beyond n, then the routes should be distributed among all the peers so that the ratio is maintained and all advertise the same number of routes.

Procedure:

Initialize the test scenario by establishing an eBGP session between the DUT and TR3. No routing information is exchanged.

Initialize TR1 and up to TRn with a predetermined number of prefixes such that such that the ratio is maintained. The physical link between TR1 thru TRn and the DUT should also be active at this time.

Establish an eBGP session between TR1 thru TRn and the DUT. Each TR router should belong to a different AS. All the prefixes in TR1 thru TRn should be advertised at this time to the DUT.

The convergence time measurement should start when the first OPEN message is exchanged between any TR and the DUT. The end of the convergence period is marked when all the TR routers have advertised all the paths to the DUT, and TR3 has received the last UPDATE.

The number of test routers should be increased in equal intervals until the maximum number under test is reached.

It is expected that the DUT will install the routes in its FIB. However, this test will not test for this.

9. References

1 Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

2 Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

3 "An Experimental Study of Delayed Internet Routing Convergence." Abha Ahuja, Farnam Jahanian, Abhijit Bose, Craig Labovitz, RIPE 37 - Routing WG.

4 "Origins of Internet Routing Instability," Infocom 99

Craig Labovitz, G. Robert Malan, Farnam Jahanian],
5 "BGP Route Flap Damping" C. Villamizar, R.Chandra, R. Govindan, RFC 2539 November 1998.
6 "Benchmarking Methodology for Network Interconnect Devices",RFC 2544, S. Bradner, J. McQuaid. March 1999.
7 Routing Policy Specification Language (RPSL), RFC 2622, " C.Alaettinoglu, C. Villamizar, E. Gerich, D. Kessens, D. Meyer, T.Bates, D. Karrenberg, M. Terpstra. June 1999.
8 "Route Refresh Capability for BGP-4", RFC 2928, E. Chen.
9 "Terminology for Forwarding Information Based (FIB)based Router Performance Benchmarking", Work in Progress, IETF,draft-ietf-bmwg-fib-term-00.txt
10 "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, Y. Rekhter, T. Li. March 1995.
11 "Benchmarking Terminology for Network Interconnection Devices",RFC 1242, S. Bradner. July 1991.
12 "Requirements for IP Version 4 Routers", RFC 1812, F. Baker. June 1995.
13 "BGP Route Flap Damping", RFC 2439, C. Villamizar, R. Chandra, R. Govindan. November 1998.
14 "Protection of BGP Sessions via the TCP MD5 Signature Option", RFC 2385, A. Heffernan. August 1998.
15 "Multiprotocol Extensions for BGP-4", RFC 2858, T. Bates,Y. Rekhter, R. Chandra, D. Katz. June 2000.
16 Junos 4.2 Software Routing Guide
17 RIPE 178, "RIPE Routing-WG Recommendation for coordinated route-flap damping parameters, Tony Barber, Sean Doran, Daniel Karrenberg, Christian Panigl, Joachim Schmitz

10. Acknowledgments

Thanks to Francis Ovenden for review and Abha Ahuja for encouragement.Much appreciation to Jeff Haas, Matt Richardson, and Shane Wright at Nexthop for comments and input.

9 Author's Addresses

Howard Berkowitz
Nortel Networks
5012 S. 25th St
PO Box 6897
Arlington VA 22206

Phone: +1 703 998-5819 (ESN 451-5819)
Fax: +1 703 998-5058
EMail: hberkowi@nortelnetworks.com
hcb@clark.net

Alvaro Retana
Cisco Systems, Inc.

7025 Kit Creek Rd.

Research Triangle Park, NC 27709
Email: aretana@cisco.com

Susan Hares
Nexthop Technologies
517 W. William
Ann Arbor, Mi 48103
Phone:
Email: skh@nexthop.com

Padma Krishnaswamy
Nexthop Technologies
517 W William
Ann Arbor, Mi 48103
Phone:
Email: kri@nexthop.com

----- End forwarded message -----