                   Experience with the BGP-4 Protocol

Status of This Memo

Copyright Notice

Abstract

   The purpose of this memo is to document how the requirements for
   publication of a routing protocol as an Internet Draft Standard have
   been satisfied by Border Gateway Protocol version 4 (BGP-4).

   This report satisfies the requirement for "the second report", as
   described in Section 6.0 of RFC 1264. In order to fulfill the
   requirement, this report augments RFC 1773 and describes additional
   knowledge and understanding gained in the time between when the
   protocol was made a Draft Standard and when it was submitted for
   Standard.

Table of Contents

1. Introduction

   The purpose of this memo is to document how the requirements for
   publication of a routing protocol as an Internet Draft Standard have
   been satisfied by Border Gateway Protocol version 4 (BGP-4).

   This report satisfies the requirement for "the second report", as
   described in Section 6.0 of [RFC1264]. In order to fulfill the
   requirement, this report augments [RFC1773] and describes additional
   knowledge and understanding gained in the time between when the
   protocol was made a Draft Standard and when it was submitted for
   Standard.

2. BGP-4 Overview

   BGP is an inter-autonomous system routing protocol designed for
   TCP/IP internets. The primary function of a BGP speaking system is
   to exchange network reachability information with other BGP systems.
   This network reachability information includes information on the
   list of Autonomous Systems (ASes) that reachability information
   traverses. This information is sufficient to construct a graph of AS
   connectivity for this reachability, from which routing loops may be
   pruned and some policy decisions, at the AS level, may be enforced.

   The initial version of the BGP protocol was published in [RFC1105].
   Since then, BGP Versions 2, 3, and 4 have been developed and are
   specified in [RFC1163], [RFC1267], and [RFC1771], respectively.
   Changes to BGP-4 after it went to Draft Standard [RFC1771] are listed
   in Appendix N of [RFC4271].

2.1. A Border Gateway Protocol

   The initial version of the BGP protocol was published in [RFC1105].
   BGP version 2 is defined in [RFC1163]. BGP version 3 is defined in
   [RFC1267]. BGP version 4 is defined in [RFC1771] and [RFC4271].
   Appendices A, B, C, and D of [RFC4271] provide summaries of the
   changes between each iteration of the BGP specification.

3. Management Information Base (MIB)

   The BGP-4 Management Information Base (MIB) has been published
   [BGP-MIB]. The MIB was updated from previous versions, which are
   documented in [RFC1657] and [RFC1269], respectively.

   Apart from a few system variables, the BGP MIB is broken into two
   tables: the BGP Peer Table and the BGP Received Path Attribute Table.

The Peer Table reflects information about BGP peer connections, such
as their state and current activity. The Received Path Attribute
Table contains all attributes received from all peers before local
routing policy has been applied. The actual attributes used in
determining a route are a subset of the received attribute table.

4. Implementation Information

   There are numerous independent interoperable implementations of BGP
   currently available. Although the previous version of this report
   provided an overview of the implementations currently used in the
   operational Internet, at that time it has been suggested that a
   separate BGP Implementation Report [RFC4276] be generated.

   It should be noted that implementation experience with Cisco's BGP-4
   implementation was documented as part of [RFC1656].

   For all additional implementation information please reference
   [RFC4276].

5. Operational Experience

   This section discusses operational experience with BGP and BGP-4.

   BGP has been used in the production environment since 1989; BGP-4 has
   been used since 1993. Production use of BGP includes utilization of
   all significant features of the protocol. The present production
   environment, where BGP is used as the inter-autonomous system routing
   protocol, is highly heterogeneous. In terms of link bandwidth, it
   varies from 56 Kbps to 10 Gbps. In terms of the actual routers that
   run BGP, they range from relatively slow performance, general purpose
   CPUs to very high performance RISC network processors, and include
   both special purpose routers and the general purpose workstations
   that run various UNIX derivatives and other operating systems.

   In terms of the actual topologies, it varies from very sparse to
   quite dense. The requirement for full-mesh IBGP topologies has been
   largely remedied by BGP Route Reflection, Autonomous System
   Confederations for BGP, and often some mix of the two. BGP Route
   Reflection was initially defined in [RFC1966] and was updated in
   [RFC2796]. Autonomous System Confederations for BGP were initially
   defined in [RFC1965] and were updated in [RFC3065].

   At the time of this writing, BGP-4 is used as an inter-autonomous
   system routing protocol between all Internet-attached autonomous
   systems, with nearly 21k active autonomous systems in the global
   Internet routing table.

BGP is used both for the exchange of routing information between a
transit and a stub autonomous system, and for the exchange of routing
information between multiple transit autonomous systems. There is no
protocol distinction between sites historically considered
"backbones" versus "regional" or "edge" networks.

The full set of exterior routes carried by BGP is well over 170,000
aggregate entries, representing several times that number of
connected networks. The number of active paths in some service
provider core routers exceeds 2.5 million. Native AS path lengths
are as long as 10 for some routes, and "padded" path lengths of 25 or
more autonomous systems exist.

6. TCP Awareness

BGP employs TCP [RFC793] as it's Transport Layer protocol. As such,
all characteristics inherent to TCP are inherited by BGP.

For example, due to TCP's behavior, bandwidth capabilities may not be
realized because of TCP's slow start algorithms and slow-start
restarts of connections, etc.

7. Metrics

This section discusses different metrics used within the BGP
protocol. BGP has a separate metric parameter for IBGP and EBGP.
This allows policy-based metrics to overwrite the distance-based
metrics; this allows each autonomous system to define its independent
policies in Intra-AS, as well as Inter-AS. BGP Multi Exit
Discriminator (MED) is used as a metric by EBGP peers (i.e., inter-
domain), while Local Preference (LOCAL_PREF) is used by IBGP peers
(i.e., intra-domain).

7.1. MULTI_EXIT_DISC (MED)

BGP version 4 re-defined the old INTER-AS metric as a MULTI_EXIT_DISC
(MED). This value may be used in the tie-breaking process when
selecting a preferred path to a given address space, and provides BGP
speakers with the capability of conveying the optimal entry point
into the local AS to a peer AS.

Although the MED was meant to only be used when comparing paths
received from different external peers in the same AS, many
implementations provide the capability to compare MEDs between
different autonomous systems.

Though this may seem a fine idea for some configurations, care must
be taken when comparing MEDs of different autonomous systems. BGP

speakers often derive MED values by obtaining the IGP metric
associated with reaching a given BGP NEXT_HOP within the local AS.
This allows MEDs to reasonably reflect IGP topologies when
advertising routes to peers. While this is fine when comparing MEDs
of multiple paths learned from a single adjacent AS, it can result in
potentially bad decisions when comparing MEDs of different autonomous
systems. This is most typically the case when the autonomous systems
use different mechanisms to derive IGP metrics, BGP MEDs, or perhaps
even use different IGP protocols with vastly contrasting metric
spaces.

Another MED deployment consideration involves the impact of the
aggregation of BGP routing information on MEDs. Aggregates are often
generated from multiple locations in an AS to accommodate stability,
redundancy, and other network design goals. When MEDs are derived
from IGP metrics associated with said aggregates, the MED value
advertised to peers can result in very suboptimal routing.

The MED was purposely designed to be a "weak" metric that would only
be used late in the best-path decision process. The BGP working
group was concerned that any metric specified by a remote operator
would only affect routing in a local AS if no other preference was
specified. A paramount goal of the design of the MED was to ensure
that peers could not "shed" or "absorb" traffic for networks they
advertise.

7.1.1. MEDs and Potatoes

Where traffic flows between a pair of destinations, each is connected
to two transit networks, each of the transit networks has the choice
of sending the traffic to the peering closest to another transit
provider or passing traffic to the peering that advertises the least
cost through the other provider. The former method is called "hot
potato routing" because, like a hot potato held in bare hands,
whoever has it tries to get rid of it quickly. Hot potato routing is
accomplished by not passing the EBGP-learned MED into the IBGP. This
minimizes transit traffic for the provider routing the traffic. Far
less common is "cold potato routing", where the transit provider uses
its own transit capacity to get the traffic to the point in the
adjacent transit provider advertised as being closest to the
destination. Cold potato routing is accomplished by passing the
EBGP-learned MED into IBGP.

If one transit provider uses hot potato routing and another uses cold
potato routing, traffic between the two tends to be symmetric.
Depending on the business relationships, if one provider has more
capacity or a significantly less congested transit network, then that
provider may use cold potato routing. The NSF-funded NSFNET backbone

and NSF-funded regional networks are examples of widespread use of
cold potato routing in the mid 1990s.

In some cases, a provider may use hot potato routing for some
destinations for a given peer AS, and cold potato routing for others.
The different treatment of commercial and research traffic in the
NSFNET in the mid 1990s is an example of this. However, this might
best be described as 'mashed potato routing', a term that reflects
the complexity of router configurations in use at the time.

Seemingly more intuitive references, which fall outside the vegetable
kingdom, refer to cold potato routing as "best exit routing", and hot
potato routing as "closest exit routing".

7.1.2. Sending MEDs to BGP Peers

   [RFC4271] allows MEDs received from any EBGP peers by a BGP speaker
   to be passed to its IBGP peers. Although advertising MEDs to IBGP
   peers is not a required behavior, it is a common default. MEDs
   received from EBGP peers by a BGP speaker SHOULD NOT be sent to other
   EBGP peers.

   Note that many implementations provide a mechanism to derive MED
   values from IGP metrics to allow BGP MED information to reflect the
   IGP topologies and metrics of the network when propagating
   information to adjacent autonomous systems.

7.1.3. MED of Zero Versus No MED

   [RFC4271] requires an implementation to provide a mechanism that
   allows MED to be removed. Previously, implementations did not
   consider a missing MED value the same as a MED of zero. [RFC4271]
   now requires that no MED value be equal to zero.

   Note that many implementations provide a mechanism to explicitly
   define a missing MED value as "worst", or less preferable than zero
   or larger values.

7.1.4. MEDs and Temporal Route Selection

   Some implementations have hooks to apply temporal behavior in MED-
   based best path selection. That is, all things being equal up to MED
   consideration, preference would be applied to the "oldest" path,
   without preference for the lower MED value. The reasoning for this
   is that "older" paths are presumably more stable, and thus
   preferable. However, temporal behavior in route selection results in
   non-deterministic behavior, and as such, may often be undesirable.

8. Local Preference

   The LOCAL_PREF attribute was added to enable a network operator to
   easily configure a policy that overrides the standard best path
   determination mechanism without independently configuring local
   preference policy on each router.

   One shortcoming in the BGP-4 specification was the suggestion that a
   default value of LOCAL_PREF be assumed if none was provided.
   Defaults of zero or the maximum value each have range limitations, so
   a common default would aid in the interoperation of multi-vendor
   routers in the same AS (since LOCAL_PREF is a local administration
   attribute, there is no interoperability drawback across AS
   boundaries).

   [RFC4271] requires that LOCAL_PREF be sent to IBGP Peers and not to
   EBGP Peers. Although no default value for LOCAL_PREF is defined, the
   common default value is 100.

   Another area where exploration is required is a method whereby an
   originating AS may influence the best path selection process. For
   example, a dual-connected site may select one AS as a primary transit
   service provider and have one as a backup.

```
                    /---- transit B ----\
       end-customer                        transit A----
                    /---- transit C ----\
```

   In a topology where the two transit service providers connect to a
   third provider, the real decision is performed by the third provider.
   There is no mechanism to indicate a preference should the third
   provider wish to respect that preference.

   A general purpose suggestion has been the possibility of carrying an
   optional vector, corresponding to the AS_PATH, where each transit AS
   may indicate a preference value for a given route. Cooperating
   autonomous systems may then choose traffic based upon comparison of
   "interesting" portions of this vector, according to routing policy.

   While protecting a given autonomous systems routing policy is of
   paramount concern, avoiding extensive hand configuration of routing
   policies needs to be examined more carefully in future BGP-like
   protocols.

9.   Internal BGP In Large Autonomous Systems

   While not strictly a protocol issue, another concern has been raised
   by network operators who need to maintain autonomous systems with a
   large number of peers. Each speaker peering with an external router
   is responsible for propagating reachability and path information to
   all other transit and border routers within that AS. This is
   typically done by establishing internal BGP connections to all
   transit and border routers in the local AS.

   Note that the number of BGP peers that can be fully meshed depends on
   a number of factors, including the number of prefixes in the routing
   system, the number of unique paths, stability of the system, and,
   perhaps most importantly, implementation efficiency. As a result,
   although it's difficult to define "a large number of peers", there is
   always some practical limit.

   In a large AS, this leads to a full mesh of TCP connections
   (n * (n-1)) and some method of configuring and maintaining those
   connections. BGP does not specify how this information is to be
   propagated. Therefore, alternatives, such as injecting BGP routing
   information into the local IGP, have been attempted, but turned out
   to be non-practical alternatives (to say the least).

   To alleviate the need for "full mesh" IBGP, several alternatives have
   been defined, including BGP Route Reflection [RFC2796] and AS
   Confederations for BGP [RFC3065].

10. Internet Dynamics

   As discussed in [RFC4274], the driving force in CPU and bandwidth
   utilization is the dynamic nature of routing in the Internet. As the
   Internet has grown, the frequency of route changes per second has
   increased.

   We automatically get some level of damping when more specific NLRI is
   aggregated into larger blocks; however, this is not sufficient. In
   Appendix F of [RFC4271], there are descriptions of damping techniques
   that should be applied to advertisements. In future specifications
   of BGP-like protocols, damping methods should be considered for
   mandatory inclusion in compliant implementations.

   BGP Route Flap Damping is defined in [RFC2439]. BGP Route Flap
   Damping defines a mechanism to help reduce the amount of routing
   information passed between BGP peers, which reduces the load on these
   peers without adversely affecting route convergence time for
   relatively stable routes.

None of the current implementations of BGP Route Flap Damping store
route history by unique NRLI or AS Path, although RFC 2439 lists this
as mandatory. A potential result of failure to consider each AS Path
separately is an overly aggressive suppression of destinations in a
densely meshed network, with the most severe consequence being
suppression of a destination after a single failure. Because the top
tier autonomous systems in the Internet are densely meshed, these
adverse consequences are observed.

Route changes are announced using BGP UPDATE messages. The greatest
overhead in advertising UPDATE messages happens whenever route
changes to be announced are inefficiently packed. Announcing routing
changes that share common attributes in a single BGP UPDATE message
helps save considerable bandwidth and reduces processing overhead, as
discussed in Section 12, Update Packing.

Persistent BGP errors may cause BGP peers to flap persistently if
peer dampening is not implemented, resulting in significant CPU
utilization. Implementors may find it useful to implement peer
dampening to avoid such persistent peer flapping [RFC4271].

11. BGP Routing Information Bases (RIBs)

[RFC4271] states "Any local policy which results in routes being
added to an Adj-RIB-Out without also being added to the local BGP
speaker's forwarding table, is outside the scope of this document".

However, several well-known implementations do not confirm that
Loc-RIB entries were used to populate the forwarding table before
installing them in the Adj-RIB-Out. The most common occurrence of
this is when routes for a given prefix are presented by more than one
protocol, and the preferences for the BGP-learned route is lower than
that of another protocol. As such, the route learned via the other
protocol is used to populate the forwarding table.

It may be desirable for an implementation to provide a knob that
permits advertisement of "inactive" BGP routes.

It may be also desirable for an implementation to provide a knob that
allows a BGP speaker to advertise BGP routes that were not selected
in the decision process.

12. Update Packing

Multiple unfeasible routes can be advertised in a single BGP Update
message. In addition, one or more feasible routes can be advertised
in a single Update message, as long as all prefixes share a common
attribute set.

The BGP4 protocol permits advertisement of multiple prefixes with a
common set of path attributes in a single update message, which is
commonly referred to as "update packing". When possible, update
packing is recommended, as it provides a mechanism for more efficient
behavior in a number of areas, including:

   o Reduction in system overhead due to generation or receipt of
     fewer Update messages.

   o Reduction in network overhead as a result of less packets and
     lower bandwidth consumption.

   o Reduction in frequency of processing path attributes and looking
     for matching sets in the AS_PATH database (if you have one).
     Consistent ordering of the path attributes allows for ease of
     matching in the database, as different representations of the
     same data do not exist.

The BGP protocol suggests that withdrawal information should be
packed in the beginning of an Update message, followed by information
about reachable routes in a single UPDATE message. This helps
alleviate excessive route flapping in BGP.

13. Limit Rate Updates

The BGP protocol defines different mechanisms to rate limit Update
advertisement. The BGP protocol defines a
MinRouteAdvertisementInterval parameter that determines the minimum
time that must elapse between the advertisement of routes to a
particular destination from a single BGP speaker. This value is set
on a per-BGP-peer basis.

Because BGP relies on TCP as the Transport protocol, TCP can prevent
transmission of data due to empty windows. As a result, multiple
updates may be spaced closer together than was originally queued.
Although it is not common, implementations should be aware of this
occurrence.

13.1. Consideration of TCP Characteristics

If either a TCP receiver is processing input more slowly than the
sender, or if the TCP connection rate is the limiting factor, a form
of backpressure is observed by the TCP sending application. When the
TCP buffer fills, the sending application will either block on the
write or receive an error on the write. In early implementations or
naive new implementations, setting options to block on the write or
setting options for non-blocking writes are common errors. Such
implementations treat full buffer related errors as fatal.

Having recognized that full write buffers are to be expected,
additional implementation pitfalls exist. The application should not
attempt to store the TCP stream within the application itself. If
the receiver or the TCP connection is persistently slow, then the
buffer can grow until memory is exhausted. A BGP implementation is
required to send changes to all peers for which the TCP connection is
not blocked, and is required to send those changes to the remaining
peers when the connection becomes unblocked.

If the preferred route for a given NLRI changes multiple times while
writes to one or more peers are blocked, only the most recent best
route needs to be sent. In this way, BGP is work conserving
[RFC4274]. In cases of extremely high route change, a higher volume
of route change is sent to those peers that are able to process it
more quickly; a lower volume of route change is sent to those peers
that are not able to process the changes as quickly.

For implementations that handle differing peer capacities to absorb
route change well, if the majority of route change is contributed by
a subset of unstable NRLI, the only impact on relatively stable NRLI
that makes an isolated route change is a slower convergence, for
which convergence time remains bounded, regardless of the amount of
instability.

14. Ordering of Path Attributes

The BGP protocol suggests that BGP speakers sending multiple prefixes
per an UPDATE message sort and order path attributes according to
Type Codes. This would help their peers quickly identify sets of
attributes from different update messages that are semantically
different.

Implementers may find it useful to order path attributes according to
Type Code, such that sets of attributes with identical semantics can
be more quickly identified.

15. AS_SET Sorting

AS_SETs are commonly used in BGP route aggregation. They reduce the
size of AS_PATH information by listing AS numbers only once,
regardless of the number of times it might appear in the process of
aggregation. AS_SETs are usually sorted in increasing order to
facilitate efficient lookups of AS numbers within them. This
optimization is optional.

16. Control Over Version Negotiation

   Because pre-BGP-4 route aggregation can't be supported by earlier
   versions of BGP, an implementation that supports versions in addition
   to BGP-4 should provide the version support on a per-peer basis. At
   the time of this writing, all BGP speakers on the Internet are
   thought to be running BGP version 4.

17. Security Considerations

   BGP provides a flexible and extendable mechanism for authentication
   and security. The mechanism allows support for schemes with various
   degrees of complexity. BGP sessions are authenticated based on the
   IP address of a peer. In addition, all BGP sessions are
   authenticated based on the autonomous system number advertised by a
   peer.

   Because BGP runs over TCP and IP, BGP's authentication scheme may be
   augmented by any authentication or security mechanism provided by
   either TCP or IP.

17.1. TCP MD5 Signature Option

   [RFC2385] defines a way in which the TCP MD5 signature option can be
   used to validate information transmitted between two peers. This
   method prevents a third party from injecting information (e.g., a TCP
   Reset) into the datastream, or modifying the routing information
   carried between two BGP peers.

   At the moment, TCP MD5 is not ubiquitously deployed, especially in
   inter-domain scenarios, largely because of key distribution issues.
   Most key distribution mechanisms are considered to be too "heavy" at
   this point.

   Many have naively assumed that an attacker must correctly guess the
   exact TCP sequence number (along with the source and destination
   ports and IP addresses) to inject a data segment or reset a TCP
   transport connection between two BGP peers. However, recent
   observation and open discussion show that the malicious data only
   needs to fall within the TCP receive window, which may be quite
   large, thereby significantly lowering the complexity of such an
   attack.

   As such, it is recommended that the MD5 TCP Signature Option be
   employed to protect BGP from session resets and malicious data
   injection.

17.2. BGP Over IPsec

   BGP can run over IPsec, either in a tunnel or in transport mode,
   where the TCP portion of the IP packet is encrypted. This not only
   prevents random insertion of information into the data stream between
   two BGP peers, but also prevents an attacker from learning the data
   being exchanged between the peers.

   However, IPsec does offer several options for exchanging session
   keys, which may be useful on inter-domain configurations. These
   options are being explored in many deployments, although no
   definitive solution has been reached on the issue of key exchange for
   BGP in IPsec.

   Because BGP runs over TCP and IP, it should be noted that BGP is
   vulnerable to the same denial of service and authentication attacks
   that are present in any TCP based protocol.

17.3. Miscellaneous

   Another routing protocol issue is providing evidence of the validity
   and authority of routing information carried within the routing
   system. This is currently the focus of several efforts, including
   efforts to define threats that can be used against this routing
   information in BGP [BGPATTACK], and efforts to develop a means of
   providing validation and authority for routing information carried
   within BGP [SBGP] [soBGP].

   In addition, the Routing Protocol Security Requirements (RPSEC)
   working group has been chartered, within the Routing Area of the
   IETF, to discuss and assist in addressing issues surrounding routing
   protocol security. Within RPSEC, this work is intended to result in
   feedback to BGP4 and future protocol enhancements.

18. PTOMAINE and GROW

   The Prefix Taxonomy (PTOMAINE) working group, recently replaced by
   the Global Routing Operations (GROW) working group, is chartered to
   consider and measure the problem of routing table growth, the effects
   of the interactions between interior and exterior routing protocols,
   and the effect of address allocation policies and practices on the
   global routing system. Finally, where appropriate, GROW will also
   document the operational aspects of measurement, policy, security,
   and VPN infrastructures.

   GROW is currently studying the effects of route aggregation, and also
   the inability to aggregate over multiple provider boundaries due to
   inadequate provider coordination.

Within GROW, this work is intended to result in feedback to BGPv4 and
future protocol enhancements.

19. Internet Routing Registries (IRRs)

Many organizations register their routing policy and prefix
origination in the various distributed databases of the Internet
Routing Registry. These databases provide access to information
using the RPSL language, as defined in [RFC2622]. While registered
information may be maintained and correct for certain providers, the
lack of timely or correct data in the various IRR databases has
prevented wide spread use of this resource.

20. Regional Internet Registries (RIRs) and IRRs, A Bit of History

The NSFNET program used EGP, and then BGP, to provide external
routing information. It was the NSF policy of offering different
prices and providing different levels of support to the Research and
Education (RE) and the Commercial (CO) networks that led to BGP's
initial policy requirements. In addition to being charged more, CO
networks were not able to use the NSFNET backbone to reach other CO
networks. The rationale for higher prices was that commercial users
of the NSFNET within the business and research entities should
subsidize the RE community. Recognition that the Internet was
evolving away from a hierarchical network to a mesh of peers led to
changes away from EGP and BGP-1 that eliminated any assumptions of
hierarchy.

Enforcement of NSF policy was accomplished through maintenance of the
NSF Policy Routing Database (PRDB). The PRDB not only contained each
networks designation as CO or RE, but also contained a list of the
preferred exit points to the NSFNET to reach each network. This was
the basis for setting what would later be called BGP LOCAL_PREF on
the NSFNET. Tools provided with the PRDB generated complete router
configurations for the NSFNET.

Use of the PRDB had the fortunate consequence of greatly improving
reliability of the NSFNET, relative to peer networks of the time.
PRDB offered more optimal routing for those networks that were
sufficiently knowledgeable and willing to keep their entries current.

With the decommission of the NSFNET Backbone Network Service in 1995,
it was recognized that the PRDB should be made less single provider
centric, and its legacy contents, plus any further updates, should be
made available to any provider willing to make use of it. The
European networking community had long seen the PRDB as too US-
centric. Through Reseaux IP Europeens (RIPE), the Europeans created
an open format in RIPE-181 and maintained an open database used for

address and AS registry more than policy. The initial conversion of
the PRDB was to RIPE-181 format, and tools were converted to make use
of this format. The collection of databases was termed the Internet
Routing Registry (IRR), with the RIPE database and US NSF-funded
Routing Arbitrator (RA) being the initial components of the IRR.

A need to extend RIPE-181 was recognized and RIPE agreed to allow the
extensions to be defined within the IETF in the RPS WG, resulting in
the RPSL language. Other work products of the RPS WG provided an
authentication framework and a means to widely distribute the
database in a controlled manner and synchronize the many
repositories. Freely available tools were provided, primarily by
RIPE, Merit, and ISI, the most comprehensive set from ISI. The
efforts of the IRR participants has been severely hampered by
providers unwilling to keep information in the IRR up to date. The
larger of these providers have been vocal, claiming that the database
entry, simple as it may be, is an administrative burden, and some
acknowledge that doing so provides an advantage to competitors that
use the IRR. The result has been an erosion of the usefulness of the
IRR and an increase in vulnerability of the Internet to routing based
attacks or accidental injection of faulty routing information.

There have been a number of cases in which accidental disruption of
Internet routing was avoided by providers using the IRR, but this was
highly detrimental to non-users. Filters have been forced to provide
less complete coverage because of the erosion of the IRR; these types
of disruptions continue to occur infrequently, but have an
increasingly widespread impact.

21. Acknowledgements

We would like to thank Paul Traina and Yakov Rekhter for authoring
previous versions of this document and providing valuable input on
this update. We would also like to acknowledge Curtis Villamizar for
providing both text and thorough reviews. Thanks to Russ White,
Jeffrey Haas, Sean Mentzer, Mitchell Erblich, and Jude Ballard for
supplying their usual keen eyes.

Finally, we'd like to think the IDR WG for general and specific input
that contributed to this document.

22. References

22.1. Normative References

   [RFC1966]   Bates, T. and R. Chandra, "BGP Route Reflection An
                alternative to full mesh IBGP", RFC 1966, June 1996.

   [RFC2385]   Heffernan, A., "Protection of BGP Sessions via the TCP
                MD5 Signature Option", RFC 2385, August 1998.

   [RFC2439]   Villamizar, C., Chandra, R., and R. Govindan, "BGP Route
                Flap Damping", RFC 2439, November 1998.

   [RFC2796]   Bates, T., Chandra, R., and E. Chen, "BGP Route
                Reflection - An Alternative to Full Mesh IBGP", RFC 2796,
                April 2000.

   [RFC3065]   Traina, P., McPherson, D., and J. Scudder, "Autonomous
                System Confederations for BGP", RFC 3065, February 2001.

   [RFC4274]   Meyer, D. and K. Patel, "BGP-4 Protocol Analysis", RFC
                4274, January 2006.

   [RFC4276]   Hares, S. and A. Retana, "BGP 4 Implementation Report",
                RFC 4276, January 2006.

   [RFC4271]   Rekhter, Y., Li, T., and S. Hares, Eds., "A Border
                Gateway Protocol 4 (BGP-4)", RFC 4271, January 2006.

   [RFC1657]   Willis, S., Burruss, J., Chu, J., "Definitions of Managed
                Objects for the Fourth Version of the Border Gateway
                Protocol (BGP-4) using SMIv2", RFC 1657, July 1994.

   [RFC793]   Postel, J., "Transmission Control Protocol", STD 7, RFC
                793, September 1981.

22.2. Informative References

   [RFC1105]   Lougheed, K. and Y. Rekhter, "Border Gateway Protocol
                (BGP)", RFC 1105, June 1989.

   [RFC1163]   Lougheed, K. and Y. Rekhter, "Border Gateway Protocol
                (BGP)", RFC 1163, June 1990.

   [RFC1264]   Hinden, R., "Internet Engineering Task Force Internet
                Routing Protocol Standardization Criteria", RFC 1264,
                October 1991.

   [RFC1267]   Lougheed, K. and Y. Rekhter, "Border Gateway Protocol 3
                (BGP-3)", RFC 1267, October 1991.

   [RFC1269]   Willis, S. and J. Burruss, "Definitions of Managed
                Objects for the Border Gateway Protocol: Version 3", RFC
                1269, October 1991.

   [RFC1656]   Traina, P., "BGP-4 Protocol Document Roadmap and
                Implementation Experience", RFC 1656, July 1994.

   [RFC1771]   Rekhter, Y. and T. Li, "A Border Gateway Protocol 4
                (BGP-4)", RFC 1771, March 1995.

   [RFC1773]   Traina, P., "Experience with the BGP-4 protocol", RFC
                1773, March 1995.

   [RFC1965]   Traina, P., "Autonomous System Confederations for BGP",
                RFC 1965, June 1996.

   [RFC2622]   Alaettinoglu, C., Villamizar, C., Gerich, E., Kessens,
                D., Meyer, D., Bates, T., Karrenberg, D., and M.
                Terpstra, "Routing Policy Specification Language (RPSL)",
                RFC 2622, June 1999.

   [BGPATTACK] Convery, C., "An Attack Tree for the Border Gateway
                Protocol", Work in Progress.

   [SBGP]   "Secure BGP", Work in Progress.

   [soBGP]   "Secure Origin BGP", Work in Progress.

Authors' Addresses

   Danny McPherson
   Arbor Networks

   EMail: danny@arbor.net

   Keyur Patel
   Cisco Systems

   EMail: keyupate@cisco.com