Network Working Group	R. White
Internet-Draft	Cisco Systems
Expires: September 13, 2008	J. Burns
	Wachovia
	March 12, 2008

Common Practices in Routing Protocols Deployment
draft-white-rppract-01

Status of this Memo

Abstract

This document discusses common practices used in deploying routing protocols in both public and private networks. The focus is not to describe how routing protocols should be deployed, but rather how they are generally deployed, to provide those working on specifications which impact the operation of routing protocols with guidance in what will likely be deployed, or what will likely not be deployed. The focus in thie document will be ionterdomain routing, but it will cover aspects of intradomain routing, as well.

1. Background

When considering new extensions to existing routing protocols, it's useful to consider them in the context of existing usage of these protocols. Various questions come to mind, such as:

- *Common Underlying Principles of Network Designs
- *Common Practices in Route Origination
- *Common Practices in Routing Database Management
- *Common Practices in Aggregation
- *Common Practices in Peering
- *Common Practices in Security
- *....

Each of these topics will be covered in a separate section below.

2. Common Underlying Principles of Network Designs

There are a number of underlying principles most network designers take into account when designing networks which don't fit neatly into any single category below; these are covered in this section. Most of these principles apply across multiple layers, and with many different protocols, so while examples are given, these principles can be found in many parts of any given network design, in ways that may not be immediately obvious or apparent.

2.1. Network Growth

Networks always grow, through organic growth and through mergers and aquisitions. To counter this, network design is often focused on the removal of information from the routing database. There are three types of information which are commonly removed from the routing database:

- *Fine grained reachability information is often replaced with more gross levels of reachability information.
- *State changes are often hidden at various points within the network.

- *Topology information is often reduced from a fine grain view of the network to a single point of reachability.

One mechanism used to remove these types of information from the routing database is the aggregation of routing data. Common techniques used to aggregate routing data are covered in greater detail in a later section. Removing information from the routing database through aggregation virtually always causes suboptimal routing. The corollary to this is that for finer grained traffic control, more state is always required, hence traffic engineering must always be balanced with stability in network design.

Another mechanism commonly used to remove information from the routing database is virtualization, or splitting the network into two pieces vertically, throughout the entire physical topology. One instance of this is using BGP to carry external routes while carrying an IGP to carry internal routes. This splits the database into two pieces, based on the characteristics of the information, and carries them separately. While the traffic being routed is carried along the same topologies, the control plane data is split, or virtualized.

2.2. Deterministic Behaviour

Network designs often favor deterministic behaviour in the face of failures or changes over non-deterministic behaviour. This is generally supported by the observation that the Mean Time To Repair is virtually always a larger component of network downtime than the Mean Time Between Failures. Deterministic behaviour is a tremendous aid in troubleshooting, which can decrease the Mean Time To Repair dramatically. Some examples of designing for deterministic behaviour include:

- *Link metrics are normally manually engineered to select a primary and alternate path through the network for any given source/destination pair, rather than allowing the routing protocol to naturally process the paths, and build paths which might fail over in non-deterministic ways.

- *Trees for routing multicast routing may be manually configured throughout a network, to control the paths and backup paths available to certain classes of traffic

---

2.3. Convergence Verses Network Stability

Newer classes of traffic place a great deal of load on network convergence. At one time, a convergence time of 3 to 9 minutes was considered acceptable, as witnessed by the default timers and operation of early distance-vector protocols. Networks now must contend with very high speed links, across which loops with durations in the 100s of milliseconds can lead to a total failure of sections of the network. Networks must also contend with applications which cannot accept any loss of connectivity above the 100s of milliseconds, and some applications which cannot tolerate any packet loss.
The primary problem with these sorts of requirements is that extermely high network convergence speeds allow no time for dampening rapid changes in the network, and, in fact, can amplify rapid network changes, reducing network stability, sometimes to the point where the network fails to converge. Network design thus must be built around converging quickly while maintaining stability, a sometimes difficult balance to achieve. Some of the techniques designers use to balance between stability and convergence speed include:

- Pushing detection as close to the hardware as possible. For instance, point-to-point links are used where possible, so the physical media state is tied directly to the logical media state.

- When logical state doesn't track physical state directly, using layer 2 mechanisms where possible to detect circuit outages.

- Using exponential backoff and other dampening mechanisms to prevent a positive feedback loop from forming, adversely impacting network performance.

3. Common Practices in Route Origination

Interior and exterior gateway protocols have a number of ways in which they classify routing information, the primary of which is the way in which destinations have been injected into the protocol.

3.1. Interior Gateway Protocols

For interior gateway protocols, routing information is normally classified as originating either from within the protocol, or from a source which is external to the protocol. Destinations which are learned of through a direct connection, such as a connection to a

subnet on a router running the protocol, are called internal routes. Destinations which are learned of through some other means, outside the protocol, are called external routes.
Virtually all routing information is injected into interior gateway protocols as internal routing information, unless there is a specific reason for injecting external information into the IGP routing domain. Some specific reasons might include:

> When multiple routing protocols are being used in the same network. Generally, this occurs when two networks are merged, or when a part of the network runs a different routing protocol for policy or design reasons.
>
> When interaction is occuring with a network not under the local administrator's control. Generally, injecting external live routing information between interior gateway protocols between routing domains is not encouraged, but there are instances when this occurs.
>
> To inject manually configured reachability information into the protocol. This generally occurs along the edges of a network, to provide reachability to destinations not within the network itself.
>
> To provide reachability across some form of layer 3 virtual private network, when no mechanism is deployed or supported to provide the transport of native routing information across the VPN.

Generally, injection of external routing information is avoided where possible in network designs, unless there is a specific policy or design related reason to do so.

3.2. Exterior Gateway Protocols

For exterior gateway protocols, the distinction between internal and external routing information is blurred, as all information is considered to be external. There is an indicator of where a specific piece of routing information originated, but this information is used very low on the decision process, and so it's generally not considered a factor in route choice.
However, there is another aspect of route origination which is a common concern in exterior gateway protocols, such as [BGP]--how routing information is locally originated on a given router. In all implementations of [BGP], routing information can either be originated from the local routing table, or it can be originated from a local manually configured route. Generally, to improve network stability, routes are injected into BGP by manually configuring a local static route, and injecting the manually configured route into the protocol, rather than by pulling information from the dynamic routing table.

4. Common Practices in Routing Database Management

When managing policies and filters in the routing database, explicit and obvious mechanisms are generally preferred over implicit, or less obvious, mechanisms. Some examples of this include:

- When redistribution between routing protocols, route tags are preferred over lists of redistributed routes to prevent routing loops from forming.

- When filtering at an AS boundary in [BGP], filtering based on the AS Path length is generally preferred over filtering on communities, or other attributes, because the AS Path is obvious and well known, while a lot of network engineers will not examine other attributes.

5. Common Practices in Aggregation

Aggregation of reachability information in a network occurs both in the IGP and the the EGP, and there are different common practices for each one. The two section below discuss these practices. In a third section, the common practice of allowing longer prefixes matches through an aggregation point is discussed.

5.1. Aggregation Practices in IGPs

Normally, aggregation in IGP is performed through manual configuration, and the aggregate route information is pulled from the local RIB. Quite often, the metric of the resulting aggregate route is forced to remain constant (which prevents state changes in one part of the network from impacting other parts of the network) through the use of a virtual interface, or a manually configured metrics attached to the aggregation configuration.

5.2. Aggregation Practices in EGPs

While aggregation commands are available in most implementations of [BGP], and there are extensive rules covering how to aggregation the various attributes of a set of aggregated routes, aggregation is not used in most BGP deployments. Instead, it is much more common for a manually configured route to originated into BGP to advertise an aggregate. Filters are normally used in conjunction with these manually originated routes to prevent components of the aggregate from being leaked to peering routers.

5.3. Allowing Components Through Aggregation

It is common to allow components to be advertised along with aggregated routing information to provide optimal routing to specific destinations. To provide an example:

```
 +----[B]---10.1.2.0/24
 |     |
[A]    +----10.1.3.0/24
 |     |
 +----[C]---10.1.4.0/24
```

In this network, the network designer might want to reduce the amount of routing data and state flowing to A. In order to do this, manual summaries can be configured at B and C, so only a shorter prefix covering all the reachable destinations is advertised. However, as noted earlier, the consequence of configuring this manual aggregation of routing information would be the introduction of suboptimal routing in the network, from A, towards 10.1.2.0/24 and 10.1.4.0/24. To counter this, the network engineer might opt to leak these two specific routes through the aggregate.
What is seen from the outside as a "multihoming" problem is, then, actually a traffic engineering problem. Most often providing two alternate paths in any network will result in the desire to optimally route traffic through those paths, whether they are equal cost or not. In most cases, leaking more specific reachability information is the quickest and most obvious way to reach the right balance of routing information verses optimal routing.

6. Common Practices In Peering

Many network design problems need to be taken into account when setting up peering, both for IGPs and for [BGP]. Common practices in this area include:

> eBGP peers are normally set up for fast down detection where possible, which is generally only possible with sessions over point-to-point links.
>
> eBGP sessions are generally manually configured not to accept a TCP keepalive timer less than 10 or 15 seconds, to prevent the peering router from negotiating very low TCP keepalive timers, which consumes processor.
>
> [OSPF] designated routers and [IS-IS] Designated Intermediate Systems are normally chosen through manual configuration. Deterministic behaviour is the goal in all cases where one router within a set is chosen for a role or a specific set of processing.

7. Common Practices in Security

Security practices generally center around preventing state changes and false routing information from entering the network, and preventing access to infrastructure devices, including routers, within the network. Some commonly used techniques in this area include:

- Filtering reachability information at network edges so infrastructure devices are not reachable outside the network.

- Configuring packet filters at network edges to directly prevent infrastructure devices from being reached from outside the network.

- Filtering reachability information at network edges to prevent the injection of private routes, bogus routes, or routes used for internal infrastructure.

- Route count limiters at the network edge where live routing data is accepted from an outside network, to prevent overflowing local routing tables.

Cryptographic secuirity mechanisms, such as MD5, are not generally configured for various reasons, including:

*Processing requirements cryptographic mechanisms are generally high, which can produce generally undesirable side effects.

*Key management for cryptographic mechanisms is generally difficult to imeplement and manage.

8. Acknowledgements

....

9. Informative References

[BGP-MD5]	Heffernan, A., “Protection of BGP Sessions via the TCP MD5 Signature Option,” RFC 2385, August 1998 (TXT).
[RFC4271]	Rekhter, Y., Li, T., and S. Hares, “A Border Gateway Protocol 4 (BGP-4),” RFC 4271, January 2006 (TXT).

Authors' Addresses

	Russ White
	Cisco Systems
Phone:	
Fax:	
Email:	riw@cisco.com
URI:	
	John Burns
	Wachovia
Phone:	
Fax:	
Email:	john.burns1@wachovia.com
URI:	

Full Copyright Statement

This document and the information contained herein are provided on an “AS IS” basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.