Overview Of | How Allied Telesis Routers And Managed Layer 3 Switches Choose The Best IP Route

Introduction

In a simple network, any given router (or layer 3 switch) will typically have only one route to any given destination, but as a network becomes more complex it will often become multiply connected, giving multiple paths from any given point to any other point. When a network reaches this level of complexity, routers will have multiple entries for any given IP route. The router needs criteria by which it can choose from among the multiple routes.

Which products does this document apply to?

This document applies to the following Allied Telesis routers and managed layer 3 switches:

- AR300, AR400, and AR700 series routers
- AT-8600, AT-8700XL, AT-8800, Rapier, and Rapier i series switches
- AT-9800 series switches
- AT-8900, AT-9900, and AT-9900s series switches
- x900 series switches
- SwitchBlade series switches

The criteria

There are a number of criteria available to the router for choosing an IP route. Let us first list them, in no particular order, and then look at how the router uses these criteria in its decision process.

Metric

The long standing standard mechanism for choosing between IP routes is the metric. A metric is basically a measure of the speed of a particular route. Different routing protocols define different methods for calculating metrics. For example, the RIP protocol simply counts hops, i.e. the number of routers that a packet must pass through to get to the destination. The OSPF protocol defines a more elaborate calculation process which takes into account the number of hops and the bandwidths of the links between the routers.

Preference

Unfortunately, there does not really exist any well-defined process for comparing the metrics of routes learnt by different routing protocols. Hence it is necessary to introduce another parameter, known as preference, which the router assigns to different types of route as follows:

- Interface routes are assigned preference 0.
- Routes learnt by OSPF from within the OSPF cloud are assigned preference 10.
- Statically added routes are assigned preference 60.
- Routes learnt by RIP are assigned preference 100.
- Routes learnt by OSPF as external routes are assigned preference 150.

Netmask

The netmask on an IP route provides a measure of the specificity of a route. If there are two different routes in the routing table:

- One route to 204.34.5.0 with netmask 255.255.255.0
- One route to 204.34.5.64 with netmask 255.255.255.224

the latter route covers a smaller subnet than the former. So, if the router were looking for the route to 204.34.5.72, then it would choose the latter route, as it more specifically points to the subnet containing that particular IP address.

Policy

Policy routing provides a means of routing packets based on protocol type, source, destination, etc. The policy to be used when routing a given packet can be set by some external device, which modifies the Type of Service (TOS) field in the packet's header, or by policy filters defined on the router itself. If a packet is to be routed according to a particular policy, the router will only look for routes that have been allocated that particular policy value.

The routing decision process

The router decides upon the best route to a given destination as follows:

1. Eliminate those routes whose destination network does not contain the required destination address.
2. Eliminate those routes whose policy does not match the policy defined for the packet in question (either by the packet's TOS field, or the router's policy filters).

   The default policy value for any packet is 0, and the default policy value for an IP route is 0.
3. If there is more than one route still in contention at this stage, choose the route(s) with the lowest preference value.
4. If there is more than one route still in contention at this stage, choose the route(s) with the lowest metric.

5. If there is more than one route still in contention at this stage, choose the route(s) with the most specific netmask.
6. If there is more than one route still in contention at this stage then:
   - If Equal-Cost-Multipath (ECMP) routing is enabled (it is enabled by default), then share packets equally across the routes.
   - If ECMP routing is disabled, then just choose the first route in the list of contending routes.

Example

In the routing table below, there are several routes to the network 192.168.1.0:

```
IP Routes
-------------------------------------------------------------------------------
Destination       Mask              NextHop             Interface           Age
DLCI/Circ.        Type     Policy   Protocol            Metrics      Preference
-------------------------------------------------------------------------------
192.168.1.0       255.255.255.0     0.0.0.0             eth0-1              395
-                 direct   0        interface           1                     0
192.168.1.0       255.255.255.0     0.0.0.0             eth0-1               63
-                 remote   0        ospf-intra          1                    10
192.168.1.0       255.255.255.0     192.168.5.1         ppp0-3              175
-                 remote   0        static              2                    60
192.168.1.0       255.255.255.0     192.168.2.1         eth0-2              176
-                 remote   0        rip                 2                   100
192.168.1.0       255.255.255.0     192.168.200.1       fr0-0               181
20                remote   0        rip                 2                   100
192.168.203.0     255.255.255.0     192.168.0.1         eth0-0               31
-                 remote   0        ospf-ext            3                   150

-------------------------------------------------------------------------------
```

If the router had to forward a policy 0 packet to 192.168.1.10, it would choose the route at the top of the table, because by step 3 of the decision process above, that would be the only route left in contention.

Potential trap: Static and OSPF-intra routes

Static routes may be ignored if there is a competing OSPF-intra route present. This is because OSPF-intra routes are assigned preference 10, whereas static routes are assigned preference 60.

For example, consider an OSPF-intra route such as:

```
--------------------------------------------------------------------------------
Destination       Mask              NextHop             Interface           Age
DLCI/Circ.        Type     Policy   Protocol            Metrics      Preference
--------------------------------------------------------------------------------
...
...

192.168.1.0       255.255.255.0     0.0.0.0             eth0-1               63
-                 remote   0        ospf-intra          1                    10
...
...

--------------------------------------------------------------------------------
```

Then add a static route to 192.168.1.192 with netmask 255.255.255.248, via some other interface:

```
--------------------------------------------------------------------------------
Destination       Mask              NextHop             Interface           Age
DLCI/Circ.        Type     Policy   Protocol            Metrics      Preference
--------------------------------------------------------------------------------
...
...

192.168.1.0       255.255.255.0     0.0.0.0             eth0-1               63
-                 remote   0        ospf-intra          1                    10
192.168.1.192     255.255.255.248   0.0.0.0             ppp4-1                4
-                 remote   0        static              3                    60
...
...

--------------------------------------------------------------------------------
```

You may initially expect that the router will now use the static route to forward traffic to 192.168.1.194, as the static route has the more specific netmask.

However, the router will actually continue to use the OSPF route, as it has the lower preference.

If you want to force the router to use the static route, then you must specifically alter the preference on that route to be some value less than 10:

```
set ip rou=192.168.1.192 mask=255.255.255.248 int=ppp4-1
 next=0.0.0.0 preference=5
```

This means that the routing table ends up as:

```
-------------------------------------------------------------------------------
Destination       Mask              NextHop           Interface           Age
DLCI/Circ.        Type     Policy   Protocol          Metrics      Preference
-------------------------------------------------------------------------------
...
...

192.168.1.0       255.255.255.0     0.0.0.0           eth0-1               63
-                 remote   0        ospf-intra        1                    10
192.168.1.192     255.255.255.248   0.0.0.0           ppp4-1                4
-                 remote   0        static            3                     5
...
...

-------------------------------------------------------------------------------
```

USA Headquarters | 19800 North Creek Parkway | Suite 200 | Bothell | WA 98011 | USA | T: +1 800 424 4284 | F: +1 425 481 3895
European Headquarters | Via Motta 24 | 6830 Chiasso | Switzerland | T: +41 91 69769.00 | F: +41 91 69769.11
Asia-Pacific Headquarters | 11 Tai Seng Link | Singapore | 534182 | T: +65 6383 3832 | F: +65 6383 3830
www.alliedtelesis.com

C613-16047-00 REV B