Technical Guide

Route Selection

Feature Overview and Configuration Guide

Introduction

This guide describes the route selection process used by the AlliedWare Plus™ Operating System. Understanding the route selection process helps in analyzing and troubleshooting route-related problems.

The process of routing packets consists of selectively forwarding data packets from one network to another. Your device must determine which network to send each packet to, and over which interface to send the packet in order to reach the desired network. This information is contained in your device's routes. For each packet, your device chooses the best route it has for that packet and uses that route to forward the packet. In addition, you can define filters to restrict the way packets are sent.

Products and software version that apply to this guide

This guide applies to AlliedWare Plus products that support routing and/or Layer 3 switching, running version **5.4.4** or later. To see whether your product supports routing and/or Layer 3 switching, see the following documents:

- The product's Datasheet
- The product's Command Reference

These documents are available from the above links on our website at alliedtelesis.com.

Most features described in this document are supported from AlliedWare Plus 5.4.4 or later. These features are available in later releases:

- Version 5.4.5-2.1 and later supports per-flow ECMP routing on AR-series firewalls
- Version 5.4.7-0.1 and later supports extended hardware switching on x310 Series switches
- Version 5.4.7-1.1 and later supports ECMP routing on interfaces with dynamically-assigned IP addresses, on AR-series firewalls.

- Version 5.5.1-2.1 and later supports weighted lottery mode on AR-series firewalls.
- Version 5.5.1-2.1 and later supports descriptions on static routes.

Content

Types of Routes

Your device learns routes from static information entered as part of the configuration process and by listening to any configured routing protocols. The following types of routes are available on your device:

Connected routes

When an IP address (and mask) is applied to an interface, this provides two pieces of information:

- The IP address of the interface.
- The subnet to which the interface is connected.

Your device then knows that the route to this connected subnet is via that interface. Hence, it enters the route into your device's route table as a connected route.This route tells your device to send packets over that interface when the packets are addressed to the interface’s subnet.

Static routes

You can manually enter routes, which are then called static routes. Static routes are used in cases where a routing protocol is not being used, or the route learnt by a routing protocol needs to be over-ridden by a different route, or an alternative backup route needs to be put in place that will take over when a dynamically learnt route is lost.

To create a static route, use the command:

For IPv4:

```
awplus(config)# ip route [vrf <vrf-name>] <subnet&mask> {<gateway-ip>|
<interface>} [<distance>] [weight <1-255>] [description <description>]
```

For IPv6:

```
awplus(config)# ipv6 route <dest-prefix/length> {<gateway-ip>|<interface>}
[<distance>] [description <description>]
```

For information about the route distance, see "Administrative distance" on page 8.

For information about the weight, see "Weighted lottery mode" on page 11.

The description lets you record the route’s purpose. It can be up to 80 printable ASCII characters long, including spaces. The description does not affect routing or forwarding decisions made by the device.

To remove the description, re-enter the route without entering the **description** parameter.

Dynamic routes

Your device learns dynamic routes from one or more routing protocols such as RIP, BGP, or OSPF. The routing protocol updates these routes as the network topology changes.

In all but the most simple networks, we recommend that you configure at least one dynamic routing protocol. Routing protocols enable your device to learn routes from other routers and switches on the network, and to respond automatically to changes in network topology.

Routing protocols use different metrics to calculate the best path for a destination. However, when two or more paths have an equal cost/metric and Equal Cost Multipath (ECMP) is enabled on a system, AlliedWare Plus may utilize two paths or more from the same protocol.

- Routing Information Protocol (RIP)

  Routing Information Protocol (RIP) is a simple distance vector IPv4 routing protocol. It determines the number of hops between the destination and your device, where one hop is one link. Given a choice of routes, RIP uses the route that takes the lowest number of hops. If multiple routes have the same hop count, RIP chooses the first route it finds.

  See the RIP Feature Overview and Configuration Guide for further information about RIP Configuration.

- Routing Information Protocol next generation (RIPng)

  Routing Information Protocol next generation (RIPng) is a simple distance vector IPv6 routing protocol. It is the IPv6 equivalent of RIP. It determines the number of hops between the destination and your device, where one hop is one link. Given a choice of routes, RIPng uses the route that takes the lowest number of hops. If multiple routes have the same hop count, RIPng chooses the first route it finds.

  RIPng (Routing Information Protocol next generation) is an extension of RIPv2 to support IPv6. RFC 2080 specifies RIPng. The main differences between RIPv2 and RIPng are:

  - RIPng uses innate IPv6 authentication to provide RIP updates authentication
  - RIPng does not allow the attachment of arbitrary tags to routes
  - RIPng packets can associate a single next-hop address with multiple routes

  See the RIPng Feature Overview and Configuration Guide for further information about RIPng Configuration.

- Open Shortest Path First (OSPF)

  The Open Shortest Path First (OSPF) protocol is documented in RFC 1247. It has a number of significant benefits over RIP, including:

  - OSPF supports the concept of areas to allow networks to be administratively partitioned as they grow in size.
  - Faster response to topology changes
  - Less protocol traffic
  - Highly configurable route summarization

  AlliedWare Plus also supports OSPFv3, which is the adaptation of OSPF to IPv6, as defined in RFC5340.

See the OSPF Feature Overview and Configuration Guide for further information about OSPF Configuration.

Similarly, the OSPFv3 Feature Overview and Configuration Guide provides further information on the configuration of OSPFv3.

- Border Gateway Protocol (BGP)

  The Border Gateway Protocol (BGP) allows routers in different routing domains to exchange routing information. This facilitates the forwarding of data across the borders of the routing domains. BGP4 is based on distance vector (DV) protocol algorithms.

  See the BGP and BGP4+ Feature Overview and Configuration Guide for further information about BGP Configuration.

RIB and FIB Routing Tables

Your device maintains its routing information in routing tables that tell your device how to find a remote network or host. Each route is uniquely identified in a table by its IP address, network mask, next hop, interface, protocol, and policy. There are two routing tables populated by your device: the **Routing Information Base** (RIB) and the **Forwarding Information Base** (FIB).

Routing Information Base

The RIB records **all** the routes that your device has learned. Your device uses the RIB to advertise routes to its neighbor devices and to populate the FIB. It adds routes to this table when:

- you add a static route using the **ip route** command
- one or more routing protocols, such as RIP or OSPF, learn routing information with other routers or hosts
- your device creates a route for a connected interface
- your device gathers route information from an ICMP redirect message or DHCP message

Understanding the Routing Information Base (RIB)

Viewing table entries

To view the routes in the RIB, use the command:

For IPv4:

```
awplus(config)# show ip route [vrf <vrf-name>|global] database
[bgp|connected|ospf|rip|static]
```

For IPv6:

```
awplus(config)# show ipv6 route database [bgp|connected|ospf|rip|static]
```

The RIB in AlliedWare Plus contains all the routes that the routing protocols have chosen as their best routes to given destinations, plus all the static and connected routes. Because multiple routing protocols may have learnt a route to a given destination, the RIB can contain multiple routes to the same destination. The AlliedWare Plus software must chose between these multiple routes, to determine the overall best route to the destination.

The angle bracket **>** character in show output indicates the routes that have been selected as the best route to their destination. The best routes are installed in the Forwarding Information Base (FIB). Routes which have been installed in the software FIB are marked with a star ***** symbol in show output.

The Administrative Distance and the Metric are seen in the square brackets with **AD** on the left of the backslash and **Metric** to the right of **AD**, so this is shown as: **[AD / Metric]**

See the below list of other information displayed in the RIB:

- Route type
- Prefix and Prefix Length
- Administrative Distance
- Metric
- Next-Hop
- Exit interface
- Uptime

Example RIB Output

See the sample output below for example RIB output, and note that all routes (including the non-best routes) are displayed in the RIB, but note only the best routes are selected for inclusion into the FIB:

Figure 1: Example RIB output after entering the show ipv6 route database command

```
awplus#show ipv6 route database
IPv6 Routing Table
Codes: C - connected, S - static, R - RIP, O - OSPF, B - BGP, D - DHCP
IA - OSPF inter area E1 - OSPF ext. type 1, E2 - OSPF ext. type 2
> - selected route, * - FIB route, p - stale info
Timers: Uptime
R *> 2001:db8:10::/64 [120/2] via fe80::eecd:6dff:fe20:c26b, vlan20, 20:42:47
R 2001:db8:20::/64 [0/1] via ::, vlan20, 21:18:42
C *> 2001:db8:20::/64 via ::, vlan20, 21:18:42
O 2001:db8:40::/64 [110/1] via ::, vlan40, 21:18:21
C *> 2001:db8:40::/64 via ::, vlan40, 21:18:22
O *> 2001:db8:50::/64 [110/2] via fe80::eecd:6dff:fe20:c073, vlan40, 21:17:29
O *> 2001:db8:60::/64 [110/2] via fe80::eecd:6dff:fe20:c073, vlan40, 20:31:06
```

Forwarding Information Base

The RIB populates the **Forwarding Information Base** (FIB) with the best route to each destination. When your device receives an IP packet, and no filters are active that would exclude the packet, it uses the FIB to find the most specific route to the destination. If your device does not find a direct route to the destination, and no default route exists, it discards the packet and sends an ICMP message to that effect back to the source.

A route is only deleted in the FIB if the corresponding route is removed from the RIB or a better route to the same destination is learnt, and displaces the existing entry in the FIB. Changes to the software FIB are propagated to the hardware FIB, so the software and hardware FIB tables mirror each other. The number of FIB table entries also mirror the best selected routes in the RIB.

Viewing table entries

To view the routes in the FIB, use the command:

For IPv4:

```
awplus# show ip route [bgp|connected|ospf|rip|static|<ip-addr>|
<ip-addr/prefix-length>]
```

For IPv6:

```
awplus# show ipv6 route [connected|database|ospf|rip|static|summary|
<ipv6-address>|<ipv6-addr/prefix-length>)]
```

Example FIB Output

See the sample below for example FIB output, and note that only the best selected routes from the RIB are installed in the FIB.

Figure 2: Example FIB output after entering the show ipv6 route command

```
awplus#show ipv6 route
IPv6 Routing Table
Codes: C - connected, S - static, R - RIP, O - OSPF, B - BGP, D - DHCP
IA - OSPF inter area E1 - OSPF ext. type 1, E2 - OSPF ext. type 2
Timers: Uptime
R 2001:db8:10::/64 [120/2] via fe80::eecd:6dff:fe20:c26b, vlan20, 21:28:22
C 2001:db8:20::/64 via ::, vlan20, 22:04:17
C 2001:db8:40::/64 via ::, vlan40, 22:03:57
O 2001:db8:50::/64 [110/2] via fe80::eecd:6dff:fe20:c073, vlan40, 22:03:04
O 2001:db8:60::/64 [110/2] via fe80::eecd:6dff:fe20:c073, vlan40, 21:16:41
```

Administrative distance

When multiple routes are available for the same prefix, the AlliedWare Plus Operating System adds the routes with the lowest **administrative distance** to the FIB. The administrative distance is a rank given to a route based on the protocol that the route was received from. The lower the administrative distance, the higher the route preference. For example, if the RIB has these routes then the AlliedWare Plus Operating System adds routes 1, 2, and 4 to the FIB. It does not add route 3, as this has a higher administrative distance than a route with the same prefix.

Table 1: Adding routes to the FIB based on administrative distance

ROUTE	PREFIX	PROTOCOL	DISTANCE
1	192.168.1.0/24	Static	1
2	192.168.2.0/24	eBGP	20
3	192.168.2.0/24	OSPF	110
4	192.168.3.0/24	OSPF	110

Note: Administrative distance indicates a level of trustworthiness of a route where the lower the administrative distance the higher the trustworthiness of a route.

The following table lists the default administrative distances for routing protocols.

Table 2: Default administrative distances

PROTOCOLS	DISTANCE	PREFERENCE
Connected Routes directly connected to an interface.	-	1 (highest)
Static Routes added using the **ip route** command or learned through DHCP options on interfaces using DHCP to obtain an IP address.	1	2
eBGP Routes learned from BGP that are external to your network.	20	3
OSPF Routes learned from OSPF.	110	4
RIP Routes learned from RIP.	120	5
iBGP Routes learned from BGP that are internal to your network.	200	6 (lowest)
Unknown No traffic will be passed to neighbors via this route.	255	(route is not advertised to neighbors)

Static routes

For static routes, it is possible to specify a non-default distance when adding the route, using the optional **distance** parameter on the command:

```
awplus(config)# ip route [vrf <vrf-name>] <subnet&mask> {<gateway-ip>|
<interface>} [<distance>] [weight <1-255>] [description <description>]
```

```
awplus(config)# ipv6 route <dest-prefix/length> {<gateway-ip>|<interface>}
[<distance>] [description <description>]
```

BGP For eBGP and iBGP routes, enter the Router Configuration mode and use the command to enter a separate administrative distance value for each BGP route type:

```
awplus(config-router)# distance bgp <ebgp> <ibgp> <local>
```

To set the administrative distance for a specific BGP route, use the command:

```
awplus(config-router)# distance <1-255> <ip-address/mask> [<listname>]
```

OSPF To enter a separate administrative distance value for each OSPF route type, enter the Router Configuration mode and use the command:

```
awplus(config-router)# distance ospf {external <1-255>|inter-area <1-255>|
intra-area <1-255>}
awplus(config-router)#distance ospfv3 {external <1-254>|inter-area <1-
254>|intra-area <1-254>}
```

To set the same value for all OSPF or OSPFv3 route types, use the command:

```
awplus(config-router)# distance <1-255>
```

RIP For RIP routes, enter the Router Configuration mode, and use the command:

```
awplus(config-router)# distance <1-255> [<ip-addr/prefix-length> [<access-
list>]]
```

This sets the administrative distance for all RIP routes.

You cannot set an administrative distance for connected routes.

Note: AlliedWare Plus does not populate routes with an administrative distance of 255 in the FIB (Forwarding Information Base). But AlliedWare Plus does populate routes with an administrative distance of 255 in the RIB (Routing Information Base). See the following examples showing the behavior of a static route with an administrative distance of 255, which is only added to the RIB, as seen from the show output.

Output Figure 3: Static route with an administrative distance of 255 that is added to the RIB

```
awplus(config)#ip route 100.0.0.0/24 192.168.1.100 255
awplus(config)#end
awplus#show ip route database

Codes: C - connected, S - static, R - RIP, B - BGP
       O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       > - selected route, * - FIB route, p - stale info

S       100.0.0.0/24 [255/0] via 192.168.1.100, vlan1
C    *> 192.168.1.0/24 is directly connected, vlan1
```

Figure 4: The static route with an administrative distance of 255 is not added to the FIB

```
awplus(config)#ip route 100.0.0.0/24 192.168.1.100 255
awplus(config)#end
awplus#show ip route

Codes: C - connected, S - static, R - RIP, B - BGP
       O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       * - candidate default

C       192.168.1.0/24 is directly connected, vlan1
```

Metric

The Metric is used to find the best route from one routing source. In the routing table it is used as the next tie breaker if multiple routes to a given destination all have the same Administrative Distance (AD).

Metrics used by AlliedWare Plus routing protocols

The following metrics are used by AlliedWare Plus IPv4 and IPv6 routing protocols:

Table 3: AlliedWare Plus IPv4 and IPv6 routing protocol metrics

IPV4 AND IPV6 ROUTING PROTOCOL	METRIC
RIP for IPv4 / RIPng for IPv6	Hop-Count
OSPFv2 for IPv4 / OSPFv3 for IPv6	Cost
BGP for IPv4 / BGP4+ for IPv6	MULTI_EXIT-DISC / IGMP metric to Next-Hop

Equal Cost Multipath Routing

When multiple routes are available for the same prefix within the FIB, then your device uses Equal Cost Multipath Routing (ECMP) to determine how to forward packets.

ECMP allows the AlliedWare Plus Operating System to distribute traffic over multiple equal-cost routes to a destination. The software determines that two or more routes are equal cost if they have the same destination IP address and mask, Administrative Distance, Metric, and protocol. When the software learns such multiple routes, it puts them in an ECMP route group. When it sends traffic to that destination, it distributes the traffic across all routes in the group.

On AR-series firewalls, the AlliedWare Plus Operating System distributes traffic over the routes one flow at a time, so all packets in a session take the same route.

On AlliedWare Plus switches, ECMP calculations are done on a per-packet basis. This means that the packets in a flow may take different paths.

Interfaces with dynamically-assigned IP addresses

On AR-series firewalls, version 5.4.7-1.1 adds support for static Equal Cost Multi-path (ECMP) Routing for interfaces whose IP address is typically dynamically assigned, such as PPP interfaces and virtual tunnel interfaces.

This means that if you create two static routes to the same remote destination network, the egress interfaces can be specified as the next hop for the routes, instead of the next hop ip address. The AR-series firewall will balance traffic flows via the two routing paths. For example, to use ECMP routing via tunnel1 and tunnel2, use commands like the following:

```
awplus(config)#ip route 192.168.2.0/24 tunnel1
awplus(config)#ip route 192.168.2.0/24 tunnel2
```

Number of routes

On most AlliedWare Plus devices, each equal-cost route group can contain up to eight individual routes. ECMP is only used to select between routes already in the FIB.

By default, each equal-cost route group can contain four routes. You can change this setting by using the command:

```
awplus(config)# maximum-paths <1-8>
```

The maximum path setting determines how many routes with the same prefix value and the same administrative distance the FIB can contain for one destination. Once an equal-cost route group has the maximum number of routes, then the RIB cannot add any further routes to the route group. The device only adds to the group if an existing route in the ECMP group is deleted from the FIB.

To disable ECMP, set the maximum paths value to one.

Weighted lottery mode

Weighted lottery mode enables a router to distribute traffic between static routes. This feature is available from version 5.5.1-2.1 onwards on some AlliedWare Plus firewalls and VPN routers, and from 5.5.2-2.1 onwards on all of them.

Weighted lottery mode can be used when you have two or more static routes with the same destination. Using the **ip route** command with the **weight** parameter, you can set a weight for each static route. AlliedWare Plus distributes the work load based on the number of sessions that are connected through the interfaces. It uses the weight that you assign to each interface to calculate a percentage of the total sessions that are allowed to connect through each interface. It then distributes the number of sessions between the interfaces accordingly.

For example, you could configure two static routes, using the following commands:

```
awplus# configure terminal
awplus(config)# ip route 192.0.2.0/24 203.0.113.1 weight 6
awplus(config)# ip route 192.0.2.0/24 198.51.0.1 weight 14
```

The weights are added together, and a percentage allocated to each. In this example, the total weight is 20. For the first route, the weight of 6 is 30% of the total weight of 20, and is allocated 30%. The second route, with a weight of 14 of the total of 20, is allocated 70%.

When a new connection is started, these percentages are used to randomly assign a nexthop. The selection is weighted according to the percentages. So, in the above example, 30% of new connections will be assigned to the first route, and 70% to the second.

Note: The weight parameter is only supported for nexthop static routes, and not interfaces.

If the **weight** parameter has been set using the **ip route** command, it can be seen by using the **show ip route** command:

```
S       10.10.37.0/24 [110/11] via 10.10.31.16, eth1 weight 5
                               via 10.10.31.32, eth1 weight 1
S       10.10.37.0/24 [110/11] via 10.10.31.16, vlan2 weight 5
                               via 10.10.31.32, vlan2 weight 1
```

How AlliedWare Plus deletes routes

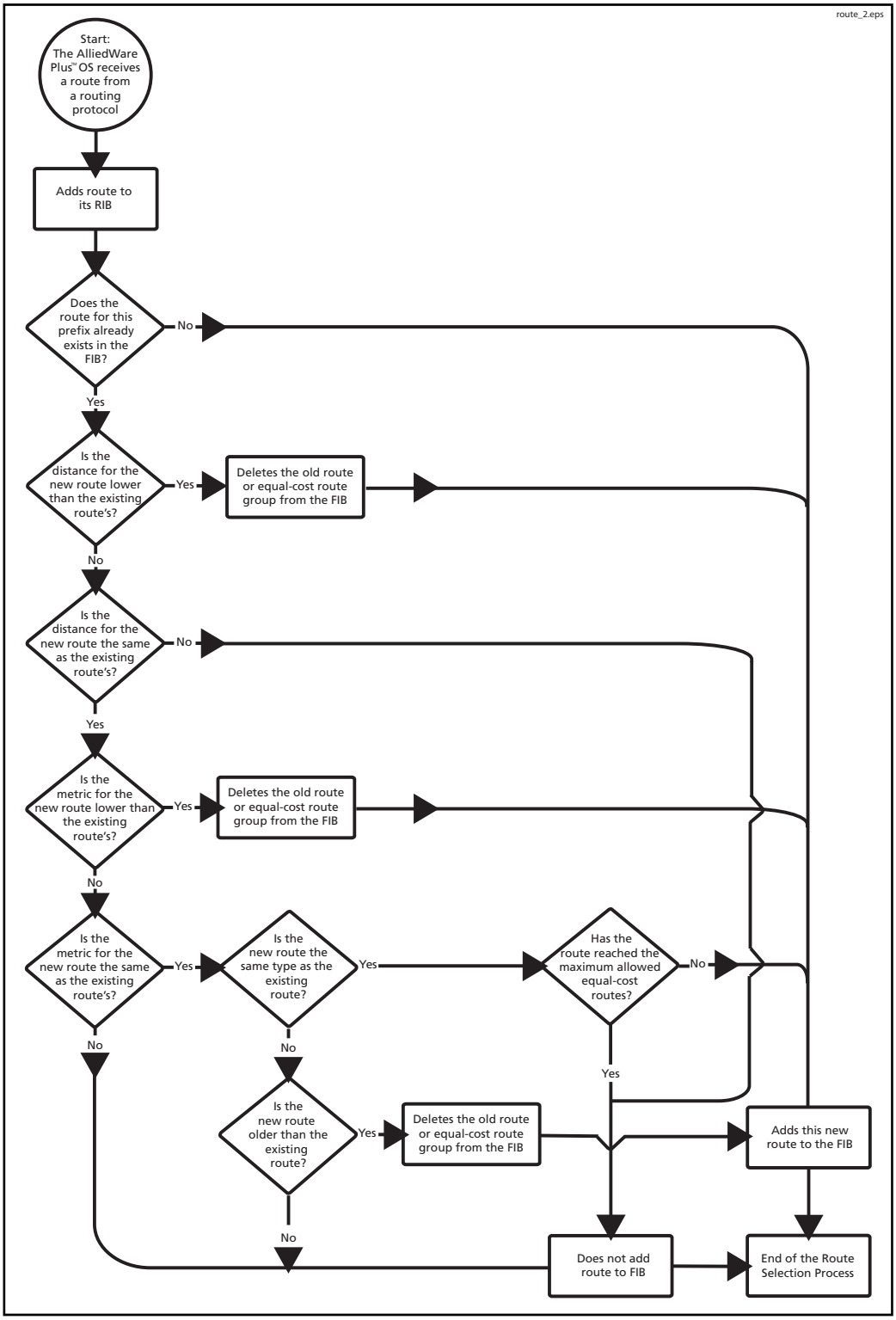
When the AlliedWare Plus Operating System receives a route delete request from a routing protocol, it first deletes the specified route from its RIB. Then it checks if the specified route is in the FIB.

- If the route is in the FIB, it deletes it from the FIB and checks if another route is available in its database for the same prefix.
- If there is another route in the database, the software installs this route in the FIB. When multiple such routes exist, the software uses the route selection mechanism to choose the best route before adding it to the FIB.

How AlliedWare Plus adds routes

The following flow chart shows how the software adds a route to the FIB.

Figure 5: How AlliedWare Plus adds a route to the FIB



Extended hardware switching on x310 Series switches

From version 5.4.7-0.1 onwards, x310 Series switches can hardware-switch traffic to individual hosts in remote networks, if those remote networks are not covered by any routes in the hardware route table. This means the switch can hardware-switch to additional remote hosts when the hardware route table is full.

If the extended hardware switching is disabled (and in versions before 5.4.7-0.1), the switch processes traffic for each such host via the CPU every time it has to send traffic to the host. If the extended hardware switching is enabled, the switch only processes traffic via the CPU the first time it has to send traffic to the host. Then it copies the host into the switch's hardware host table and hardware-switches future traffic to it.

To enable the extended hardware switching, use the following command:

```
awplus(config)#fib cache-remote-host
```

Troubleshooting routes not installed in the RIB

Possible reasons why a route is not installed in the RIB are:

- The Layer 3 interface is not in the Up state.
- Route oscillation (route flap) is occurring with the route being added and removed frequently.
- The routing process from which the route is learned, has deleted the route.
- A routing protocol has learned the maximum number of routes allowed by the license, so the routes are not installed to the RIB.

See the **max-static-routes** command in the product's Command Reference for detailed command description and command example information, where static routes are applied before adding routes to the RIB. The Command Reference is available on our website at alliedtelesis.com.

Troubleshooting routes not installed in the FIB

Possible reasons why a route is not installed in the FIB are:

- The maximum-paths limit may have been reached (currently supports up to eight equal cost paths being installed).
- The maximum-paths command may be set to a lower value preventing more paths being selected as best.
- The desired route type has a higher AD over another route entry in the RIB, so is not preferred.
- The **max-fib-routes** command is configured and the maximum number of installed software FIB routes has been reached.

See the **max-fib-routes** command in the product's Command Reference for detailed command description and command example information to control the maximum number of FIB routes configured. The Command Reference is available on our website at alliedtelesis.com.

C613-22029-00 REV F

Allied Telesis NETWORK SMARTER

North America Headquarters | 19800 North Creek Parkway | Suite 100 | Bothell | WA 98011 | USA | T: +1 800 424 4284 | F: +1 425 481 3895
Asia-Pacific Headquarters | 11 Tai Seng Link | Singapore | 534182 | T: +65 6383 3832 | F: +65 6383 3830
EMEA & CSA Operations | Incheonweg 7 | 1437 EK Rozenburg | The Netherlands | T: +31 20 7950020 | F: +31 20 7950021

alliedtelesis.com