Wireless Layer 3 Roaming Best Practices

Large WLAN networks (for example, those found on large campuses) may require IP session roaming at layer 3 to enable application and session persistence while a mobile client roams across multiple VLANs. For example, when a user on a VoIP call roams between APs on different VLANs without layer 3 roaming, the user's session will be interrupted as the external server must re-establish communication with the client's new IP address. During this time, a VoIP call will noticeably drop for several seconds, providing a degraded user experience. In smaller networks, it may be possible to configure a flat network by placing all APs on the same VLAN.

However, on large networks filled with thousands of devices, configuring a flat architecture with a single native VLAN may be an undesirable network topology from a best practices perspective; it may also be challenging to configure legacy setups to conform to this architecture. A turnkey solution designed to enable seamless roaming across VLANs is therefore highly desirable when configuring a complex campus topology. Using Meraki's secure auto-tunneling technology, layer 3 roaming can be enabled using a mobility concentrator, allowing for bridging across multiple VLANs in a seamless and scalable fashion.

Typical Campus Architecture

Large campuses are often designed with a multi-VLAN architecture to segment broadcast traffic. Typically, network best practices dictate a one-to-one mapping of an IP subnet to a VLAN, e.g., client devices joining VLAN 10 will be assigned an IP address out of the subnet range 10.0.10.0/24. In this design, clients in different VLANs will receive IP addresses in different subnets via a DHCP server. Multi-VLAN architectures can vary to include multiple subnets within a building (e.g., one for each floor/area), or multiple subnets across a large site (e.g., one for each building/region in a large campus or enterprise environment).

As seen in the diagram below, the typical campus architecture has the core L3 switch connected to multiple L3 distribution switches (one per site), with each distribution switch then branching off to L2 access switches configured on different VLANs. In this fashion, each site is assigned a different VLAN to segregate traffic from different sites. Without an L3 roaming service, a client connected to an L2 access switch at Site A will not be able to seamlessly roam to an L2 access switch connected to Site B. Upon associating with an AP on Site B, the client would obtain a new IP address from the DHCP service running on the Site B scope. In addition, a particular route configuration or router NAT may also prevent clients from roaming, even if they do retain their original IP address.
With layer 3 roaming, a client device must have a consistent IP address and subnet scope as it roams across multiple APs on different VLANs/subnets. Meraki’s auto-tunnelling technology achieves this by creating a persistent tunnel between the L3 enabled APs and depending on the architecture, a mobility concentrator. The two layer 3 roaming architectures are discussed in detail below.

**Distributed Layer 3 Roaming**

Distributed layer 3 roaming maintains layer 3 connections for end devices as they roam across layer 3 boundaries without a concentrator. The first access point that a device connects to will become the anchor access point. The anchor access point informs all of the other Meraki access points within the network that it is the anchor for a particular client. Every subsequent roam to another access point will place the device/user on the VLAN defined by the anchor AP.

Distributed layer 3 roaming is very scalable because the access points are establishing connections with each other without the need for a concentrator. The target access point will look up in the shared user database and contact the anchor access point. This communication does not traverse the Meraki Cloud and is a proprietary protocol for secure access point to access point communication. UDP port 9358 is used for this communication between the APs.
As you can see in the above diagram, Anchor AP is the AP where the client gets connected the first time. An AP to which the client is associated is called a hosting AP, it does not connect with the broadcast domain of the client. Hosting AP will create a tunnel with the Anchor AP to maintain the IP address of the client.

In case the hosting AP has direct access to the broadcast domain of the client, then the hosting AP will become the Anchor AP for that client.

A client's anchor AP will timeout after the client has left the network for 30 seconds.

### Broadcast Domain Mapping

Each Meraki Access point sends layer 2 broadcast probes over the Ethernet uplink to discover broadcast domain boundaries on each VLAN that a client could be associated with when connected. This is done for multiple reasons. One reason is because there can be instances where AP1 is connected to an access port (no VLAN tag) and AP2 is connected to a trunk port where the same VLAN is used, but the VLAN ID is present and tagged on the uplink. These broadcast frames are of type 0x0a89 sent every 150 seconds.
There could also be situations where the same VLAN ID is used in different buildings (representing different broadcast domains), so it’s important to ensure exactly which APs and VLAN IDs can be found on which broadcast domains. Apart from tunnel load balancing and resiliency, the broadcast domain mapping and discovery process also allows for anchor APs and hosting APs to have a real-time view into which VLANs are shared between the two APs. This allows for efficient decision making when it comes to layer 2 vs layer 3 roaming for a client, as described in the “VLAN Testing and Dynamic Configuration” section below. This is important so that anchor APs for clients can be dynamically switched for load balancing reasons, or in failover situations where the original anchor AP is no longer available.

Meraki APs will send out probes to discover the following broadcast domains:

- The AP’s native VLAN
- Any VLAN that is configured for the SSID on the AP
- Any VLAN that is dynamically learned via a client policy
- Any VLAN that an AP has recently received a broadcast probe on from another Meraki AP in Dashboard network

The power of the broadcast domain mapping is that this will discover broadcast domains agnostic of VLAN IDs configured on an AP. As a result of this methodology, each AP on a broadcast domain will eventually gather exactly the AP/VLAN ID pairs that currently constitute the domain. Whenever a client connects to another SSID the Anchor AP for that client is updated.

**Broadcast Domain Discovery**

The following steps establish the AP/VLAN ID (VID) pair that correspond to a broadcast domain:

1. APs periodically broadcast a BCD announcement packet that contains the AP’s VLAN ID for that broadcast domain, giving a {sender AP,VID} pair on each broadcast domain the AP interacts with.
2. Create equivalence classes based on AP/VID pairs recently observed in BCD announcement packets on the same broadcast domain.

Additional notes:

- Each AP on a broadcast domain will eventually gather exactly the AP/VID pairs that currently constitute the domain.
- In principle, any AP/VID pair can be used to refer to a broadcast domain. Given AP1/VID1, as long as you know the full list of pairs for that broadcast domain, you can tell whether some other AP2/VID2 refers to the same domain or not.

An AP could theoretically broadcast BCD announcement packets to all 4095 potentially attached VLANs, however it will limit itself to the VLANs outlined above.

**Roaming with Broadcast Domains**

The Meraki MRs leverage a distributed client database to allow for efficient storage of clients seen in the network and to easily scale for large networks where thousands of clients may be connecting. The client distributed database is accessed by APs in real-time to determine if a connecting client has been seen previously elsewhere in the network. This requires that the APs in the Meraki network have layer 3 IP connectivity with one another, communicating over UDP port 9358. Leveraging the Meraki Dashboard, the APs are able to dynamically learn about the other APs in the network (including those located on different management VLANs) to know whom they should communicate with to look up clients in the distributed client database.

The following process describes how client roaming operates with distributed layer 3 roaming

1. Anchor APs have a full set of AP/VLAN ID pairs for each attached broadcast domain as described above.
2. On client association, the hosting AP retrieves the client data from the distributed store.
   - If the hosting AP does not find an entry in the store:
     - The hosting AP then becomes the anchor AP for the client. It stores the client in the distributed database, adding a candidate anchor AP set.
The candidate anchor set consists of the AP's own AP/VLAN ID pair plus two randomly chosen pairs from the same anchor broadcast domain.

- If the hosting AP does find an entry in the store:
  - It checks to see if the client’s VLAN is available locally, from the previous broadcast domain discovery process outlined above. If the associated VLAN ID is available, the hosting AP will become the anchor AP and the VLAN for that client will dynamically be provisioned for the client. See the section “VLAN Testing and Dynamic Configuration” below.
  - Otherwise, the hosting AP sets up an anchor AP for the client (picking a random pair from the candidate anchor set).

3. As long as the hosting AP continues to host the client, it periodically receives updates to the candidate anchor set from the anchor AP. The anchor AP replaces any AP/VLAN ID pair in the candidate anchor set that disappears with another randomly chosen AP/VLAN ID pair for that broadcast domain. The hosting AP updates the distributed store’s client entry with changes to the candidate

### VLAN Testing and Dynamic Configuration

The anchor access point runs a test to the target access point to determine if there is a shared layer 2 broadcast domain for every client serving VLAN. If there is a VLAN match on both access points, the target access point will configure the device for the VLAN without establishing a tunnel to the anchor. This test will dynamically configure the VLAN for the roaming device despite the VLAN that is configured for the target access point and the clients served by it. If the VLAN is not found on the target AP either because it is pruned on the upstream switchport or the Access Point is in a completely separated layer 3 network, the Tunneling method described below will be used.

- **Local VLAN testing and dynamic configuration** is one method used to prevent all clients from tunneling to a single anchor AP. To prevent excess tunneling the layer 3 roaming algorithm determines that it is able to place the user on the same VLAN that the client was using on the anchor AP. The client in this case does a layer 2 roam as it would in bridge mode.

### Tunneling

If necessary, the target access point will establish a tunnel to the anchor access point. Tunnels are established using Meraki-proprietary access point to access point communication. To load balance multiple tunnels amongst multiple APs, the tunneling selector will choose a random AP that has access to the original broadcast domain the client is roaming from. If the target AP detects a connectivity failure to the currently selected anchor AP, as a failover mechanism the target AP will choose a new anchor AP. Hosting AP will ping Anchor AP every second to ensure that the Anchor AP has not failed. This ping is integrated as a part of the L3 communication on UDP port 9358.

All APs must be able to communicate with each other via IP. This is required both for client data tunneling and for the distributed database. If a target access point is unable to communicate with the anchor access point the layer 3 roam will time out and the end device will be required to DHCP on the new VLAN. Data packets are not encrypted between two Anchor APs on the wired side but control and management frames are encrypted.

- **Fast roaming protocols** such as OKC and 802.11r are not currently supported with distributed layer 3 roaming. The best roaming performance will be using Layer 2 roaming with 802.11r.

### Design Example

Let’s walk through an example of the distributed layer 3 roaming architecture from start to finish. In this example network, we’ll use the following configuration:

1. 5x MRs on management VLAN 10 tagged with ‘Group A’
2. 5x MRs on management VLAN 20 tagged with ‘Group B’
3. SSID: Corporate
4. Client IP Assignment: layer 3 roaming
5. VLAN ID:
We will assume that the total of 10 APs are online and connected to Dashboard, and have IP connectivity with one another.

Client A associates with a ‘Group A’ AP on management VLAN 10, and receives an IP address in VLAN 15 as expected. This AP becomes the anchor AP & hosting AP for the client. The APs in the Meraki network have built out the broadcast domain mapping pairs (AP/VLAN ID) and are exchanging periodic updates. Client A roams to a ‘Group B’ AP on management VLAN 20, client VLAN 25. The ‘Group B’ AP is now considered the hosting AP and reads the distributed client database to see if the client has connected previously. It finds an entry for the client and checks locally to see if the client’s broadcast domain is available on the switchport. The broadcast domain is not available, and the hosting AP will now pick an anchor AP out of the candidate anchor set (supplied from the distributed client database check) which will be any AP that has advertised itself to the distributed client database as having access to client VLAN 15. Once the anchor AP is selected, along with two candidate anchors for resiliency, the tunnel is established and the hosting AP updates the distributed client database with this information.

The hosting AP will periodically refresh the anchor AP and distributed database. The anchor AP’s entry for a client has an expiration time of 30 seconds. If the client disconnects from the network for 45 seconds, as an example, it may connect back to a new anchor AP on the same broadcast domain associated with the client. The distributed database expiration timer for a client is the DHCP lease time. This effectively determines how long a client’s broadcast domain binding is remembered in the distributed database. If a client disconnects from the network, and then reconnects before the DHCP lease time has expired, then the client will still be bound to its original broadcast domain.

In another scenario, let’s imagine a large enterprise campus with 10 floors. Following common enterprise campus design, the customer has segmented one VLAN per floor for the users. To accommodate for client mobility and seamless roaming throughout the campus building, the customer wishes to leverage distributed layer 3 roaming. Using AP tags, the configuration will specify a VLAN ID assignment for a given SSID based on the tag. In this case, the following configuration will be used:

- SSID: Corporate
- Client IP Assignment: layer 3 Roaming
- VLAN ID:
  - Floor 1 - VLAN 11
  - Floor 2 - VLAN 12
  - Floor 3 - VLAN 13
  - Floor 4 - VLAN 14
  - Floor 5 - VLAN 15
  - Floor 6 - VLAN 16
  - Floor 7 - VLAN 17
  - Floor 8 - VLAN 18
  - Floor 9 - VLAN 19
  - Floor 10 - VLAN 20

The switchports which the MRs will be connecting to will be configured as trunk ports. Switches on floors 1-5 will allow VLANs 11,12,13,14,15. Switches on floors 6-10 will allow VLANs 16,17,18,19,20. With this configuration, a user who associates on floor 1 will receive an IP address on VLAN 11. As they roam throughout the building, changing floors, the roams will be layer 2 only with no tunneling required.
Only when the client roams to the upper half of the building (or vise versa) will a tunnel be formed to keep the client in its original broadcast domain. Keep in mind that even if the client originally received IP addressing on VLAN 11, since AP’s on Floor 5 have access to that broadcast domain (discovered via the Broadcast Domain Mapping & Discovery mechanism), then that client will maintain it’s VLAN 11 IP addressing information and will simply use the AP on floor 5 as it’s new anchor.

This type of design allows for maximum flexibility by allowing for traditional layer 2 roams for users who spend the majority of their time in a specific section of the building, and allowing for continued seamless roaming for the most mobile clients.

Repeaters don’t have their own IP address, so they cannot be anchor APs. When a client connects to a repeater, the repeater becomes the client’s hosting AP, and the repeater assigns its gateway as the client’s anchor AP.

### Concentrator-Based Layer 3 Roaming

Any client that is connected to a layer 3 roaming enabled SSID is automatically bridged to the Meraki Mobility Concentrator. The Mobility Concentrator acts as a focal point to which all client traffic will be tunneled and anchored when the client moves between VLANs. In this fashion, any communication data directed towards a client by third party clients or servers will appear to originate at this central anchor. Any Meraki MX can act as a Concentrator, please refer to the MX sizing guides to determine the appropriate MX appliance for the expected users and traffic.
The diagram below shows the traffic flow for a particular flow within a campus environment using the layer 3 roaming with concentrator.