

Subscriber Load Balancing for 802.3ad Link Aggregation Groups Overview

To configure subscriber load balancing, you configure a scheduler hierarchy with nodes and queues for IP, VLANs, and S-VLANs. The system distributes those nodes and queues in the scheduler hierarchy over the ports within the LAG using a technique called *partitioning*.

Ethernet queues used for hashed load balancing are always present in the scheduler hierarchy.

To ensure that QoS is symmetrically applied to all the links, the router periodically rebalances the load within the LAG using a hash algorithm. You can control the loadbalancing parameters and configure the system to dynamically rebalance. Partitioning the Scheduler Hierarchy

The system then partitions the scheduler hierarchy by binding the IP, VLAN, L2TP session, and MPLS resources for each subscriber to a selected link within the LAG at the time the subscriber interface is configured.

S-VLANs and Subscriber Load Balancing

The system *clones* S-VLAN nodes and queues on each link in the LAG. The system clone S-VLANs so it can allocate subscribers that share a common S-VLAN ID to different links within the LAG. S-VLAN nodes and queues are the only resources that are cloned; the system always allocates nodes and queues for other interface types to a single selected link.

Cloning S-VLAN nodes enables fine-grained load balancing within the LAG because VLANs within the S-VLAN can be allocated to the link with the least traffic. However, cloned S-VLANs can introduce anomalous scheduling behavior. A shaped S-VLAN node within the LAG shapes traffic on a per-link basis. Shaping a LAG S-VLAN node to 2 Mbps on a LAG with 2 links can enable up to 4 Mbps of traffic (2 Mbps per link).

Shared shaping on an S-VLAN within a LAG has the same behavior; the LAG S-VLAN that is shared shaped to 10 Mbps on a LAG with 2 ports allows up to 20Mbps of traffic; 10 Mbps for each link.

PPPoE over VLANs and Subscriber Load Balancing

The system binds PPPoE subscribers stacked over a common VLAN to the same link within the LAG. Because the underlying VLAN node is allocated to a single link, the system allocates all traffic over that VLAN to that link.

PPPoE over Ethernet (No VLANs) and Subscriber Load Balancing

The system allocates subscribers to each link independently. There are no S-VLAN nodes to clone, and no related VLAN nodes that require allocation on the same link.

MPLS over LAG and Subscriber Load Balancing

For QoS purposes, the system considers base tunnels as logical interfaces, but does not consider stacked tunnels. The system assigns MPLS base tunnels stacked over VLANs to the link to which the VLAN is assigned.

Sample Scheduler Hierarchy for Subscriber Load Balancing

Figure 1 displays the scheduler hierarchy for the Gigabit Ethernet interface in slot 3, port 0. Figure 2 displays the scheduler hierarchy for the Gigabit Ethernet interface in slot 3, port 1.

The Ethernet queues are shown in gray; they are not bound to a link in the LAG and are replicated for each link in the LAG. These queues are used for subscribers with QoS profiles that indicate Ethernet queues, and for traffic classes other than best-effort, traffic class 1, and traffic class 2.

When partitioning the scheduler hierarchy that includes 1000 VLAN subinterfaces, the system binds 500 of the subinterfaces to port 0, and binds another 500 to port 1. The binding for a given VLAN subinterface is arbitrary.

The scheduler nodes for a given VLAN subinterface are always allocated to the same port within the LAG. In this example, the scheduler nodes for VLAN 0,0 are all allocated to Gigabit Ethernet interface in slot 3, port 0.

S-VLAN nodes and queues are cloned for each link in the LAG. S-VLAN nodes in each traffic-class group are shown identically on both ports.

Figure 1: Subscriber LoadBalanced Scheduler Hierarchy for Port 0

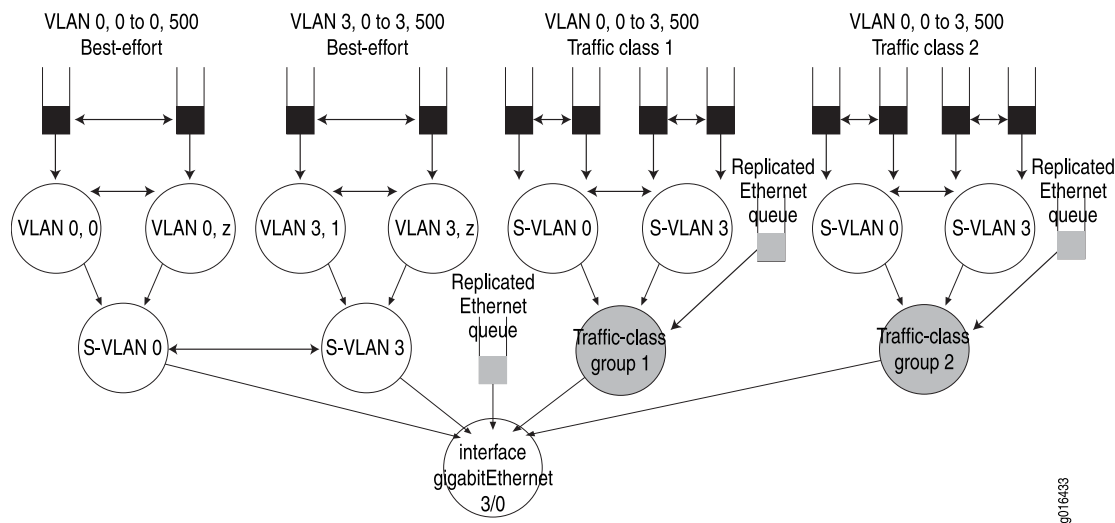
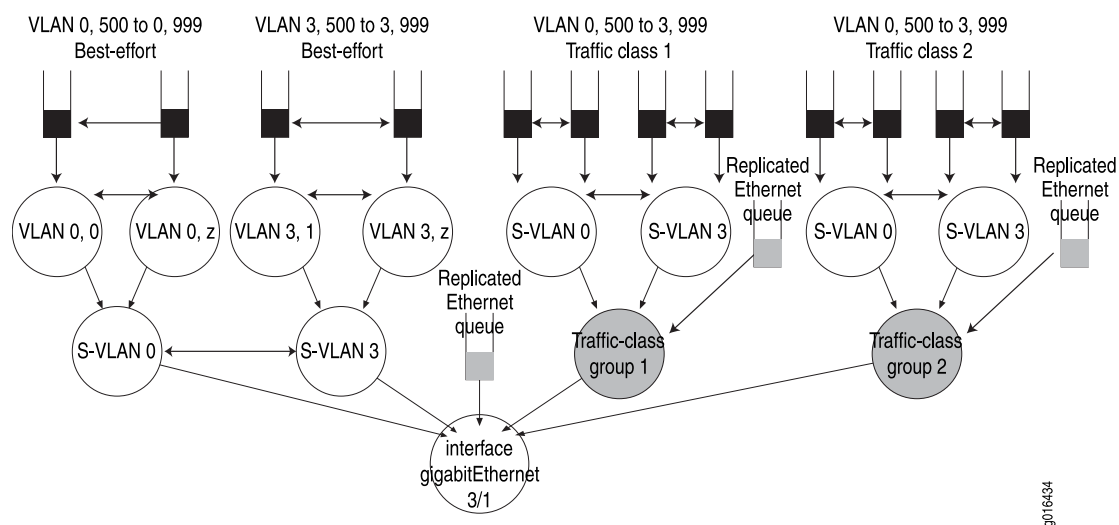


Figure 2: Subscriber LoadBalanced Scheduler Hierarchy for Port 1



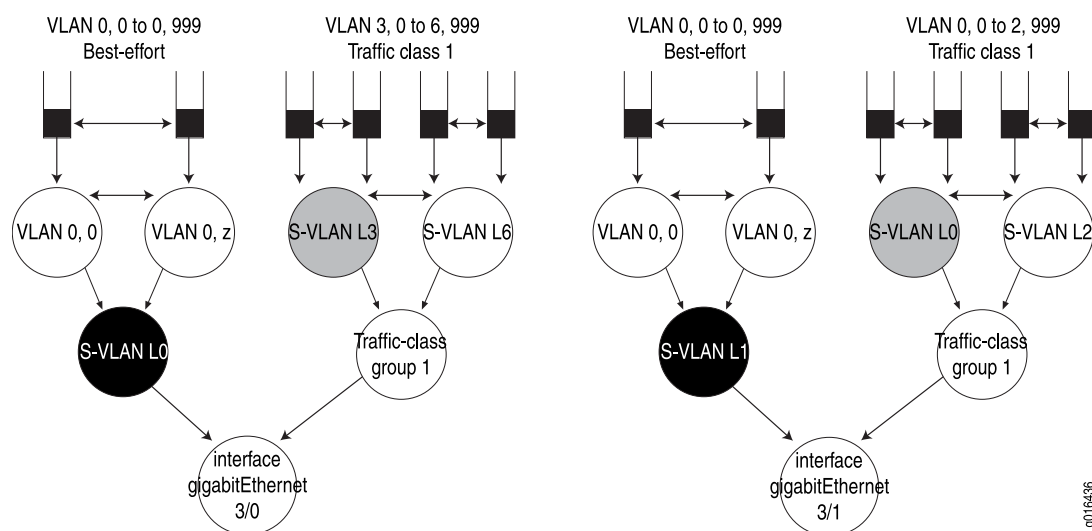
Subscriber Allocation in 802.3ad Link Aggregation Groups

You can configure upper-layer subinterfaces over a LAG interface, including VLANs, PPPoE, and MPLS.

The system balances any upper-layer subinterfaces so that each active link in the LAG carries an equal number of upper-layer subinterfaces. For this purpose, the system counts each upper-layer interface as a single subscriber, regardless of the number of forwarding interfaces stacked above it.

Figure 3 displays a sample allocation of subscribers.

Figure 3: Subscriber Allocation and Load Balancing



In an ideal QoS configuration, queues and nodes are stacked over a single port that corresponds to a LAG, with the port bandwidth equal to the sum of the overall port bandwidth.

However, the actual LAG behavior is different. No level 1 node or queue can exceed the bandwidth of a link. The relative weighting of queues and nodes results in proportional bandwidth allocation only within a link, but not across the entire LAG. Actual traffic might not be evenly balanced across links in the LAG, resulting in latency and loss on one link, while another link may be lightly loaded.

Even though relative weighting is different on a LAG, shaping and shared shaping in the partitioned scheduler hierarchy operate in the same way as a typical Ethernet configuration.

Related Topics ■ For more information about load rebalancing, see [Configuring Load Rebalancing for 802.3ad Link Aggregation Groups](#)

Published: 2010-03-24