CITRIX

# Deployment Modes

## Contents



 

## Customizing the Ethernet ports

October 26, 2018

A typical appliance has four Ethernet ports: two accelerated bridged ports, called *accelerated pair A* (apA.1 and apA.2), with a bypass (fail-to-wire) relay, and two unaccelerated motherboard ports, called Primary and Aux1. The bridged ports provide acceleration, while the motherboard ports are sometimes used for secondary purposes. Most installations use only the bridged ports.

Some SD-WAN appliances have only the motherboard ports. In this case, the two motherboard ports are bridged.

The appliance's user interface can be accessed by a VLAN or non-VLAN network. You can assign a VLAN to any of the appliance's bridged ports or motherboard ports for management purposes.

Figure 1. Ethernet Ports

### Port List

The ports are named as follows:

**Table 1. Ethernet port names**

| Ethernet Port | Name |
|---|---|
| Motherboard port 1 | Primary (or apA.1 if no bypass card is present) |
| Motherboard port 2 | Auxiliary1 or Aux1 (or apA.2 if no bypass card is present) |
| Bridge #1 | Accelerated Pair A (apA, with ports apA.1 and apA.2) |
| Bridge #2 | Accelerated Pair B (apB, with ports apB.1 and apB.2) |
| Bridge #3 | Accelerated Pair C (apC, with ports apC.1 and apC.2) |

## Port Parameters

August 20, 2018

Each bridge and motherboard port can be:

- Enabled or disabled
- Assigned an IP address and subnet mask
- Assigned a default gateway
- Assigned to a VLAN
- Set to 1000 Mbps, 100 Mbps, or 10 Mbps
- Set to full duplex, half-duplex, or auto (on SD-WAN 4000/5000 WANOP and SD-WAN 4000 SE/5100 SE appliances, some ports can be set to 10 Gbps)

All of these parameters except the speed/duplex setting are set on the Configuration: IP Address page. The speed/duplex settings are set on the Configuration: Interface page.

> **Note**s about parameters:

- Disabled ports do not respond to any traffic.
- The browser-based UI can be enabled or disabled independently on all ports.
- To secure the UI on ports with IP addresses, select HTTPS instead of HTTP on the Configuration: Administrator Interface: Web Access page.
- Inline mode works even if a bridge has no IP address. All other modes require that an IP address be assigned to the port.
- Traffic is not routed between interfaces. For example, a connection on bridge apA does not cross over to the Primary or Aux1 ports, but remains on bridge apA. All routing issues are left to your routers.

## Accelerated Bridges (apA, apB, and apC)

October 25, 2018

Every appliance has at least one pair of Ethernet ports that function as an accelerated bridge, called *apA* (for *accelerated pair A*). SD-WAN 410-SE appliance has three pairs of ethernet ports (apA, apB, and apC). A bridge can act in inline mode functioning as a transparent bridge, as if it were an Ethernet switch. Packets flow in one port and out the other. Bridges can also act in one arm mode, in which packets flow in one port and back out the same port.

An appliance that has a bypass card maintains network continuity of a bridge or appliance malfunctions.

 

Some units have more than one accelerated pair, and these additional accelerated pairs are named apB, apC, and so on.

### Bypass card

If the appliance loses power or fails in some other way, an internal relay closes and the two bridged ports are electrically connected. This connection maintains network continuity but makes the bridge ports inaccessible. Therefore you might want to use one of the motherboard ports for management access.

Caution: Do not enable the Primary port if it is not connected to your network. Otherwise, you cannot access the appliance, as explained in Ethernet Bypass and Link-Down Propagation

Bypass cards are standard on some models and optional on others. Citrix recommends that you purchase appliances with bypass cards for all inline deployments.

The bypass feature is wired as if a cross-over cable connected the two ports, which are the correct behavior in properly wired installations.

Important: Bypass installations must be tested - Improper cabling might work in normal operation but not in bypass mode. The Ethernet ports are tolerant of improper cabling and often silently adjust to it. Bypass mode is hard-wired and has no such adaptability. Test inline installations with the appliance turned off to verify that the cabling is correct for bypass mode.

### Using multiple bridges

If the appliance is equipped with two accelerated bridges, they can be used to accelerate two different links. These links can either be fully independent or they can be redundant links connecting to the same site. Redundant links can be either load-balanced or used as a main link and a failover link.

Figure 1. Using dual bridges

When it is time for the appliance to send a packet for a given connection, the packet is sent over the same bridge from which the appliance received the most recent input packet for that connection. Thus, the appliance honors whatever link decisions are made by the router, and automatically tracks the prevailing load-balancing or main-link/failover-link algorithm in real time. For non-load-balanced links, the latter algorithm also ensures that packets always use the correct bridge.

### WCCP (WANOP) and virtual inline modes

Multiple bridges are supported in both WCCP mode (WANOP) and virtual inline mode. Usage is the same as in the single-bridge case, except that WCCP (WANOP) has the additional limitation that all traffic for a given WCCP (WANOP) service group must arrive on the same bridge.

 

### High availability with multiple bridges

Two units with multiple bridges can be used in a high-availability pair. Simply match up the bridges so that all links pass through both appliances.

## Motherboard Ports

August 17, 2018

Although the Ethernet ports on a bypass card are inaccessible when the bypass relay is closed, the motherboard ports remain active. You can sometimes access a failed appliance through the motherboard ports if the bridged ports are inaccessible.

### The Primary Port

If the Primary port is enabled and has an IP address assigned to it, the appliance uses that IP address to identify itself to other acceleration units. This address is used internally for a variety of purposes, and is most visible to users as the Partner Unit field on the Monitoring: Optimization: Connections page. If no motherboard port is enabled, the appliance uses the IP address of Accelerated Pair A.

The Primary port is used for:

- Administration through the web based UI
- A back channel for group mode
- A back channel for high-availability mode

## VLAN Support

July 19, 2018

A virtual local area network (VLAN) uses part of the Ethernet header to indicate which virtual network a given Ethernet frame belongs to. SD-WAN appliances support VLAN trunking in all forwarding modes (inline, WCCP (WANOP), virtual inline, and group mode). Traffic with any combination of VLAN tags is handled and accelerated correctly.

For example, if one traffic stream passing through the accelerated bridge is addressed to 10.0.0.1, VLAN 100, and another is addressed to 10.0.0.1, VLAN 111, the appliance knows that these are two distinct destinations, even though the two VLANs have the same IP address.

 

You can assign a VLAN to all, some, or none of the appliance's Ethernet ports. If a VLAN is assigned to a port, the management interfaces (GUI and CLI) listen only to traffic on that VLAN. If no VLAN is assigned, the management interfaces listen only to traffic without a VLAN. This selection is made on the Configuration: Appliance Settings: Network Adapters: IP Addresses tab.

## Inline Mode

August 20, 2018

In inline mode, traffic passes into one of the appliance's Ethernet ports and out of the other. When two sites with inline appliances communicate, every TCP connection passing between them is accelerated. All other traffic is passed through transparently, as if the appliance were not there.

Figure 1. Inline mode, Accelerating All Traffic on a WAN

> **Note**: Any TCP-based traffic passing through both units is accelerated. No address translation, proxying, or per-site setup is required. Inline mode is auto-detecting and auto-configuring.

Configuration is minimized with inline mode, because your WAN router need not be aware of the appliance's existence.

Depending on your configuration, inline mode's link-down propagation can affect management access to the appliance if a link goes down.

Inline mode is most effective when applied to all traffic flowing into and out of a site, but it can be used for only some of the site's traffic.

## Ethernet Bypass and Link-Down Propagation

October 27, 2018

Most appliance models include a "fail-to-wire" (Ethernet bypass) feature for inline mode. If power fails, a relay closes and the input and output ports become electrically connected, allowing the Ethernet signal to pass through from one port to the other as if the appliance were not there. In fail-to-wire mode, the appliance looks like a cross-over cable connecting the two ports.

Any failure of the appliance hardware or software also closes the relay. When the appliance is restarted, the bypass relay remains closed until the appliance is fully initialized, maintaining network continuity at all times. This feature is automatic and requires no user configuration.

When the bypass relay is closed, the appliance's bridge ports are inaccessible.

 

If carrier is lost on one of the bridge ports, the carrier is dropped on the other bridge port to ensure that the link-down condition is propagated to the device on the other side of the appliance. Units that monitor link state (such as routers) are thus notified of conditions on the other side of the bridge.

Link-down propagation has two operating modes:

- If the Primary port is not enabled, the link-down state on one bridge port is mirrored briefly on the other bridge port, and then the port is re-enabled. This allows the appliance to be reached through the still-connected port for management, high availability heartbeat, and other tasks.
- If the Primary port is enabled, the appliance assumes (without checking) that the Primary port is used for management, high availability heartbeat, and other tasks. The link-down condition on one bridge port is mirrored persistently on the other port, until carrier is restored or the unit is rebooted. This is true even if the Primary port is enabled in the GUI but not connected to a network, so the Primary port should be disabled (the default) when not in use.

## Accelerating an Entire Site

July 19, 2018

Inline mode, Accelerating All Traffic on a WAN shows a typical configuration for inline mode. For both sites, the appliances are placed between the LAN and the WAN, so all WAN traffic that can be accelerated is accelerated. This is the simplest method for implementing acceleration, and it should be used when practical.

Because all the link traffic is flowing through the appliances, the benefits of fair queuing and flow control prevent the link from being overrun.

In IP networks, the bottleneck gateway determines the queuing behavior for the entire link. By becoming the bottleneck gateway, the appliance gains control of the link and can manage it intelligently. This is done by setting the bandwidth limit slightly lower than the link speed. When this is done, link performance is ideal, with minimal latency and loss even at full link utilization.

## Partial-Site Acceleration

August 17, 2018

To reserve the appliance's accelerated bandwidth for a particular group of systems, such as remote backup servers, you can install the appliance on a branch network that includes only those systems. This is shown in the following figure.

 

Figure 1. Inline Mode, Accelerating Selected Systems Only

SD-WAN traffic shaping relies on controlling the entire link, so traffic shaping is not effective with this topology, because the appliance sees only a portion of link traffic. Latency control is up to the bottleneck gateway, and interactive responsiveness can suffer.

## Configuring and Troubleshooting Inline Mode

July 19, 2018

Inline mode requires only basic configuration, because it is applied automatically to any packets passing through the accelerated bridge.

## Virtual Inline Mode

August 20, 2018

> **Note**: Use virtual inline mode only when both inline mode and WCCP mode are impractical. Do not mix inline and virtual inline modes within the same appliance. However, you can mix virtual inline and WCCP modes within the same appliance. Citrix does not recommend virtual inline mode with routers that do not support health monitoring.

In virtual inline mode, the router uses policy based routing (PBR) rules to redirect incoming and outgoing WAN traffic to the appliance for acceleration, and the appliance forwards the processed packets back to the router. Almost all of the configuration tasks are performed on the router. The only thing to be configured on the appliance is the forwarding method, and the default method is recommended.

Like WCCP, Virtual inline deployment requires no rewiring and no downtime, and it provides a solution for asymmetric routing issues faced in a deployment with two or more WAN links. Unlike WCCP, it contains no built-in status monitoring or health checking, making troubleshooting difficult. WCCP is thus the recommended mode, and virtual inline is recommended only when inline and WCCP modes are both impractical.

### Example

The following figure shows a simple network in which all traffic destined for or received from the remote site is redirected to the appliance. In this example, both the local site and remote site use virtual inline mode.

 

Figure 1. Virtual Inline Example

Following are some configuration details for the network in this example:

- Endpoint systems have their gateways set to the local router (which is not unique to virtual inline mode).
- Each router is configured to redirect both incoming and outgoing WAN traffic to the local appliance.
- Each appliance processes the traffic received from its local router and forwards it back to the router.
- PBR rules configured on the router prevent routing loops by allowing packets to make only one trip to and from the appliance. The packets that the appliance forwards back to the router are sent to their original (local or remote) destination.
- Each appliance has its default gateway set to the address of the local router, as usual (on the **Configuration: Network Adapters** page). The options for forwarding packets back to the router are Return to Ethernet Sender and Send to Gateway.

## Configuring Packet Forwarding on the Appliance

July 19, 2018

Virtual inline mode offers two packet-forwarding options:

**Return to Ethernet Sender (default)**—This mode allows multiple routers to share an appliance. The appliance forwards virtual inline output packets back to where they came from, as indicated by the Ethernet address of the incoming packet. If two routers share a single appliance, each gets its own traffic back, but not the traffic from the other router. This mode also works with a single router.

**Send to Gateway (not recommended)**—In this mode, virtual inline output packets are forwarded to the default gateway for delivery, even if they are destined for hosts on the local subnet. This option is usually less desirable than the Return to Ethernet Sender option, because it adds an easily forgotten element of complexity to the routing structure.

**To specify the packet-forwarding option**—On the Configuration: Optimization Rules: Tuning page, next to Virtual Inline, select Return to Ethernet Sender or Send to Gateway.

## Router Configuration

August 20, 2018

The router has three tasks when supporting virtual inline mode:

 

1. It must forward both incoming and outgoing WAN traffic to the SD-WAN appliance.
2. It must forward SD-WAN traffic to its destination (WAN or LAN).
3. It must monitor the health of the SD-WAN appliance so that the appliance can be bypassed if it fails.

### Policy-Based Rules

In virtual inline mode, the packet forwarding methods can create routing loops if the routing rules do not distinguish between a packet that has been forwarded by the appliance and one that has not. You can use any method that makes that distinction.

A typical method involves dedicating one of the router's Ethernet ports to the appliance and creating routing rules that are based on the Ethernet port on which packets arrive. Packets that arrive on the interface dedicated to the appliance are never forwarded back to the appliance, but packets arriving on any other interface can be.

The basic routing algorithm is:

- Do not forward packets from the appliance back to the appliance.
- If the packet arrives from the WAN, forward it to the appliance.
- If packet is destined for the WAN, forward to the appliance.
- Do not forward LAN-to-LAN traffic to the appliance.
- Traffic shaping is not effective unless all WAN traffic passes through the appliance.

> **Note**: When considering routing options, keep in mind that returning data, not just outgoing data, must flow through the appliance. For example, placing the appliance on the local subnet and designating it as the default router for local systems does not work in a virtual inline deployment. Outgoing data would flow through the appliance, but incoming data would bypass it. To force data through the appliance without router reconfiguration, use inline mode.

### Health Monitoring

If the appliance fails, data should not be routed to it. By default, Cisco policy based routing does no health monitoring. To enable health monitoring, define a rule to monitor the appliance's availability, and specify the "verify-availability" option for the "set ip next-hop" command. With this configuration, if the appliance is not available, the route is not applied, and the appliance is bypassed.

> **Important**: Citrix recommends virtual inline mode only when used with health monitoring. Many routers that support policy-based routing do not support health-checking. The health-monitoring feature is relatively new. It became available in Cisco IOS release 12.3(4)T.

Following is an example of a rule for monitoring the availability of the appliance:

 

```
pre codeblock !- Use a ping (ICMP echo)to see if appliance is connected
 track 123 rtr 1 reachabilit y ! rtr 1 type echo protocol IpIcmpecho
192.168.1.200 schedule 1 life forever start-time now
```

This rule pings the appliance at 192.168.1.200 periodically. You can test against 123 to see if the unit is up.

## Monitoring and Troubleshooting

July 19, 2018

In virtual inline mode, unlike WCCP mode (WANOP), the appliance provides no virtual inline-specific monitoring. To troubleshoot a virtual inline deployment, log into the appliance and use the Dashboard page to verify that traffic is flowing into and out of the appliance. Traffic forwarding failures are typically caused by errors in router configuration.

If the Monitoring: Usage or Monitoring: Connections pages show that traffic is being forwarded but no acceleration is taking place (assuming that an appliance is already installed on the other end of the WAN link), check to make sure that both incoming WAN traffic and outgoing WAN traffic are being forwarded to the appliance. If only one direction is forwarded, acceleration cannot take place.

To test health-checking, power down the appliance. The router should stop forwarding traffic after the health-checking algorithm times out.

## High-Availability Mode

July 19, 2018

Two identical appliances on the same subnet can be combined as a *high-availability pair.* The appliances each monitor the other's status by using the standard *Virtual Router Redundancy Protocol (VRRP)* heartbeat mechanism. The pair has a common virtual IP address for management, in addition to each appliance's management IP address. If the primary appliance fails, the secondary appliance takes over. Failover takes approximately five seconds.

High availability mode is a standard feature.

 

## Cabling Requirements

August 17, 2018

The two appliances in the high availability pair are installed onto the same subnet in either a parallel arrangement or a one-arm arrangement, both of which are shown in the following figure. In a one-arm arrangement, use the apA.2 port (and, optionally, the apB.2 port), not the apA.1 port. Some models require a separate management LAN, whether deployed in inline or one-armed mode. This is depicted only in the middle diagram.

Figure 1. Cabling for High-Availability Pairs

Do not break the above topology with additional switches. Random switch arrangements are not supported. Each of the switches must be either a single, monolithic switch, a single logical switch, or part of the same chassis.

If the spanning-tree protocol (STP) is enabled on the router or switch ports attached to the appliances, failover will work, but the failover time may increase to roughly thirty seconds. Without STP, failover time is roughly five seconds. Thus, to achieve the briefest possible failover interval, disable STP on the ports connecting to the appliances.

Figure 2. Ethernet Port Locations (Older Models)

## Other Requirements

October 27, 2018

Both appliances in an high availability pair must meet the following criteria:

- Have identical hardware, as shown by on the System Hardware entry on the Dashboard page.
- Run exactly the same software release.
- Be equipped with Ethernet bypass cards. To determine what is installed in your appliances, see the Dashboard page.

Appliances that do not support high availability display a warning on the Configuration: High Availability page.

## Updating Software on a High-Availability Pair

October 27, 2018

 

Updating the SD-WAN software on an high availability pair causes a failover at one point during the update.

Note: Clicking the Update button terminates all open TCP connections.

### To update the software on an high availability pair

1. Log on to both appliances.
2. On the secondary appliance, update the software and reboot. After the reboot, the appliance is still the secondary. Verify that the installation succeeded. The primary appliance should show that the secondary appliance exists but that automatic parameter synchronization is not working, due to a version mismatch.
3. On the primary appliance, update the software, and then reboot. The reboot causes a failover, and the secondary appliance becomes the primary. When the reboot is completed, high availability should become fully established, because both appliances are running the same software.

 

 