Describe vPC Control and Data Plane
The vPC peers use the Cisco Fabric Services protocol to synchronize forwarding-plane information and implement necessary configuration checks. The Cisco Fabric Services protocol travels on the peer link and does not require configuration by the user.

Cisco FSoE is used as the primary control plane protocol for vPC. It performs several functions:

vPC peers must synchronize the Layer 2 MAC address table between the vPC peers. If one vPC peer learns a new MAC address on a vPC, that MAC address is also programmed on the Layer 2 forwarding table of the other peer device for that same vPC. This MAC address learning mechanism replaces the regular switch MAC address learning mechanism and prevents traffic from being forwarded across the vPC peer link unnecessarily.

Cisco Fabric Services performs the synchronization of IGMP snooping information. Layer 2 forwarding of multicast traffic with vPC is based on modified IGMP snooping behavior that synchronizes the IGMP entries between the vPC peers. In a vPC implementation, IGMP traffic entering a vPC peer switch through a vPC triggers hardware programming for the multicast entry on the two vPC member devices.

Cisco Fabric Services is also used to communicate essential configuration information to ensure configuration consistency between the peer switches. Similar to regular port channels, vPCs are subject to consistency checks and compatibility checks. During a compatibility check, one vPC peer conveys configuration information to the other vPC peer to verify that vPC member ports can form a port channel. In addition to compatibility checks for the individual vPCs, Cisco Fabric Services is also used to perform consistency checks for a set of switchwide parameters that need to be configured consistently on the two peer switches.

Cisco Fabric Services is used to track the vPC status on the peer. When all vPC member ports on one of the vPC peer switches go down, Cisco Fabric Services is used to notify the vPC peer switch that its ports are now orphan ports and that traffic that is received on the peer link for that vPC should now be forwarded to the vPC.

Layer 3 vPC peers synchronize their respective ARP tables. This feature is transparently enabled and helps ensure faster convergence time on reload of a vPC switch. When two switches are reconnected after a failure, those switches use Cisco Fabric Services to perform bulk synchronization of the ARP table.

Between the pair of vPC peer switches, an election is held to determine a primary and secondary vPC device. This election is nonpreemptive. The vPC primary or secondary role is primarily a control plane role. It determines which of the two switches is responsible for the generation and processing of spanning-tree BPDUs for the vPCs.

The vPC peer switch option can be implemented, which allows the primary and secondary to generate BPDUs for vPCs independently. The two switches use the same spanning-tree bridge ID to ensure that devices connected on a vPC still see the vPC peers as a single logical switch.

The two switches actively participate in traffic forwarding for the vPCs. However, the primary and secondary roles are also important in certain failure scenarios, most notably in a peer-link failure as described later.

vPC Data Plane Operation
vPC is designed to limit the use of the peer link specifically to switch management traffic and the occasional traffic flow from a failed network port.

The peer link does not carry regular traffic for vPCs. It carries only the traffic that needs to be flooded, such as broadcast (including STP BDPUs), multicast, and unknown unicast traffic. It also carries traffic for orphan ports.

One of the most important forwarding rules for vPC is that a frame that enters the vPC peer switch from the peer link

cannot exit the switch from a vPC member port. This principle prevents frames that are received on a vPC from being flooded back onto the same vPC by the other peer switch. The exception to this rule is traffic that is destined for an orphaned vPC member port.

vPC Peer-Link Failure
A vPC peer-keepalive link is used to send keepalive messages between the vPC peer devices. The keepalive messages on the vPC peer-keepalive link determine whether a failure is on the vPC peer-link only or on the vPC peer device. The keepalive messages are used only when all the links in the peer-link fail.

If the vPC peer link fails, the following procedure happens:

The software checks the status of the remote vPC peer device using the peer-keepalive link.

The peer-keepalive link is a link between vPC peer devices that ensures that the two devices are running.

If the vPC peer device is running and peer-link failed, the secondary vPC device disables all vPC ports on its device, to prevent loops and disappearing traffic or flooding traffic. The secondary vPC peer also shuts down all switch virtual interfaces (SVI)s associated with VLANs that are configured as allowed VLANs for the vPC peer link.

Data is forwarded using only the remaining active links of the port channel.

vPC Peer Switch
Although vPCs provide a loop-free Layer 2 topology, STP is still required to provide a fail-safe mechanism to protect against incorrect or defective cabling or possible misconfiguration. When you first bring up a vPC, STP reconverges. STP treats the vPC peer link as a special link and includes the vPC peer link in the STP active topology.

Cisco recommends that you set all the vPC peer-link interfaces to the STP network port type so that bridge assurance is enabled on all vPC peer links. It is also recommended that you do not enable the STP enhancement features on vPC peer links. If the STP enhancements were previously configured, they do not cause problems for the vPC peer links.

The vPC peer switch feature was added to Cisco NX-OS to address performance concerns around STP convergence. This feature allows a pair of Cisco Nexus devices to appear as a single STP root in the Layer 2 topology. This feature eliminates the need to pin the STP root to the vPC primary switch and improves vPC convergence if the vPC primary switch fails.

If the peer switch option is not used, the vPC primary is responsible for generating and processing BPDUs and uses its own bridge ID for the BPDUs. The secondary relays BPDU messages, but does not generate BPDUs itself for the vPCs. When the peer switch option is used, the primary and secondary switches send and process BPDUs. However, they use the same bridge ID to present themselves as a single switch to devices connected on a vPC.

STP Recommendations
Several guidelines apply to using vPC with STP.

The following are the spanning-tree recommendations:

Configure aggregation vPC peers as root and secondary root.

Align the STP primary root and HSRP active router with the vPC primary peer.

If you implement the vPC peer switch, the two vPC peers behave as a single STP root.

Set all the vPC peer-link interfaces to the STP network port type so that Bridge Assurance is automatically enabled on

all vPC peer-links.

Configure the ports on the downstream devices as STP edge ports.

Compatibility Parameters
Many configuration and operational parameters must be identical on all interfaces in the vPC.

After you enable the vPC feature and configure the peer link on both vPC peer devices, Cisco Fabric Services messages provide a copy of the configuration on the local vPC peer device configuration to the remote vPC peer device. The system then determines whether any of the crucial configuration parameters differ on the two devices.

The devices automatically check for compatibility for some of these parameters on the vPC interfaces. The per-interface parameters must be consistent per interface, and the global parameters must be consistent globally.

Configuration Parameters that must be identical:

Port channel mode (on, off, or active)

Link speed, duplex mode, and trunk mode (native VLAN, allowed VLANs) per channel

STP mode and STP region configuration for MST

Enable/disable state per VLAN

STP global settings (Bridge Assurance, port type, Loop Guard)

STP interface settings (port type, Loop Guard, Root Guard)

MTU

The configuration parameters listed above must be configured identically on both devices of the vPC peer link; otherwise, the vPC moves fully or partially into a suspended mode. If any of these parameters are not enabled or defined on either device, the vPC consistency check ignores those parameters.

Note
Enter the show vpc consistency-parameters command to ensure that none of the vPC interfaces are in the suspend mode and to display the configured values on all interfaces in the vPC. The displayed configurations are only those configurations that would limit the vPC peer link and vPC from coming up.

When any of the following parameters are not configured identically on both vPC peer devices, a misconfiguration might cause undesirable behavior in the traffic flow. Configuration parameters that should be identical:

MAC aging timers and static MAC entries

VLAN interface: each device on the end of the vPC peer link must have a VLAN interface configured for the same VLAN on both ends and they must be in the same administrative and operational mode. Those VLANs configured on only one device of the peer link do not pass traffic using the vPC or peer link. You must create all VLANs on both the primary and secondary vPC devices, or the VLAN will be suspended.

All ACL configurations and parameters

QoS configuration and parameters

STP interface settings (BPDU Filter, BPDU Guard, Cost, Link type, Priority, VLANs for Rapid PVST+)

Port security

Cisco Trusted Security (CTS)

DHCP snooping

Dynamic ARP Inspection (DAI)

IP Source Guard

IGMP snooping

HSRP

PIM

All routing protocol configurations

You can configure the graceful consistency check feature, which suspends only the links on the secondary peer device when a mismatch is introduced in a working vPC. This feature is configurable only in the CLI and is enabled by default.

As part of the consistency check of all parameters from the list of parameters that must be identical, the system checks the consistency of all VLANs. The vPC remains operational, and only the inconsistent VLANs are brought down. This per-VLAN consistency check feature cannot be disabled and does not apply to Multiple Spanning Tree (MST) VLANs.

vPC Object Tracking Feature
vPC object tracking is used to prevent traffic black-holing if there is failure of a module where both peer-link and uplinks to the core resides. By tracking the interface, this feature can suspend vPC on affected switch and prevent traffic black-holing.

If you must configure all the vPC peer links and core-facing interfaces on a single module, you should configure, using the CLI, a track object and a track list that is associated with the Layer 3 link to the core and on all vPC peer links on both vPC peer devices. You use this configuration to avoid dropping traffic if that particular module goes down because when all the tracked objects on the track list go down.

Look at the following example. You have one module used for the vPC peer link and the Layer 3 uplinks to the core. In the event where you are to lose the module due to a hardware failure you would lose the vPC peer-link and the Layer 3 uplinks. If this were to happen on the vPC secondary box (B), it would not be a problem as the operational primary peer would take over suspending the vPC port-channels and VLAN interfaces on the operational secondary. The problem is in the case of a hardware failure on the operational primary device (A). If you did not use object tracking, the vPC would suspend all vPC port-channels on B and the VLAN interfaces. The peer link would also be down. You would not have a way to route the Core traffic into your vPC VLANs in this scenario.

Object Tracking gets around this by bringing down vPC on the operational primary so that you don't get into the scenario where vPC brings down the VLAN interfaces and vPC port-channels on the box that has the remaining uplinks to the core.

The system does the following:

Stops the vPC primary peer device sending peer-keepalive messages, which forces the vPC secondary peer device to take over.

Brings down all the downstream vPCs on that vPC peer device, which forces all the traffic to be rerouted in the access switch toward the other vPC peer device.

You should create a track list that contains all the links to the core and all the vPC peer links as its object. Enable tracking for the specified vPC domain for this track list. Apply this same configuration to the other vPC peer device.