Per Packet Load Balancing

Feature history for per packet load balancing

This table describes the developments of this feature, by release.

Table 1: Feature history

Feature name	Release information	Description
Per packet load balancing	Cisco IOS XE Catalyst SD-WAN Release 26.1.x Cisco Catalyst SD-WAN Manager Release 26.1.x	With this feature, Cisco SD-WAN sends packets from a single flow across multiple WAN links, maximizing bandwidth and maintaining performance by reordering packets at the destination.

Per packet load balancing

Per-packet load balancing (PPL) or adaptive single flow distribution is a bandwidth aggregation feature that

- distributes packets from a single flow across multiple WAN links,
- maximizes the use of all available paths, and
- maintains data integrity by reordering packets at the destination.

Traditional load balancing vs PPL

Traditional load balancing uses one network link for each data stream, which can limit overall bandwidth use. PPL shares traffic across multiple available links, allowing you to get the most out of your network and improve performance.

Reorder packets

When a device distributes packets over multiple links, packets may arrive at the destination out of order. The PPL feature automatically reorders these packets on a best-effort basis and releases any buffered packets based on either a time threshold or a memory threshold.

Key concepts for understanding PPL

Explains how PPL efficiently distributes traffic across multiple WAN tunnels by dynamically selecting the best paths based on latency, flowlets, and network policies.

Sender and receiver

The sender is the WAN edge device that sends PPL traffic. The receiver is the WAN edge device that receives this traffic and may reorder packets if they arrive out of order.

Traffic selection

Defines which specific traffic from the service VPNs use PPL, enabling detailed control through network policies. You can configure this using data policies in the "from-service" direction, and any traffic matched in the policy’s match clause becomes eligible for PPL.

Inter-tunnel latency measurement

PPL measures the delay across all available tunnels and selects the tunnel with the lowest latency as a reference point. It compares other tunnels to this reference and includes only those within a set latency range for traffic distribution. This method is a Cisco proprietary algorithm. It ensures efficient and reliable path selection for your network traffic.

Flow splitting

Instead of splitting traffic into individual packets, PPL groups packets into flowlets, which are small sets of packets. You create flowlets either after sending a set number of packets or when you detect natural breaks (packet gaps) in the traffic.

Packet gap

Packet gap is a method used in PPL to detect natural breaks in a flow of traffic. When there is an idle period between bursts of packets that exceeds the maximum latency difference among available tunnels, a new flowlet is started and sent on a different tunnel.

Candidate tunnel group (CTG)

The group of tunnels available for PPL, selected first by “local colors” and then filtered using metrics like latency and loss. A maximum of eight tunnels can be part of the CTG.

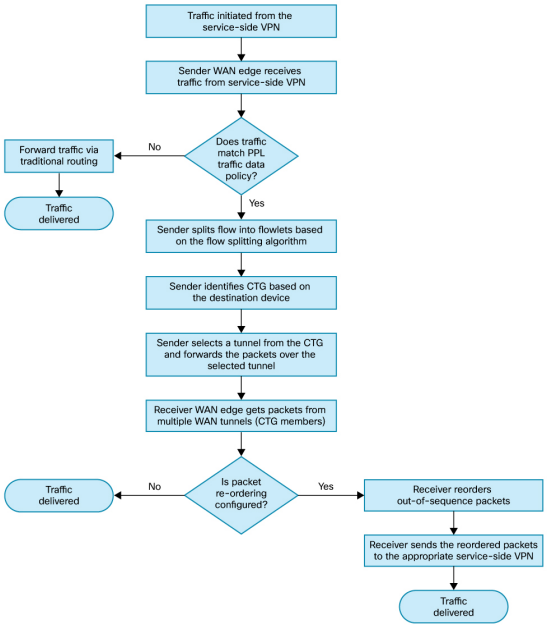
Path selection

Determines how each flowlet is sent across the reasonably available tunnels in the CTG, using round-robin distribution.

Sequence of events for PPL

The sender device receives traffic from service-side VPNs and checks it against the traffic data policy for PPL rules. If matched, the sender splits the flow into flowlets and forwards them over reasonably available tunnels in the Candidate Tunnel Group. The receiver device collects and re-orders the flowlets as needed, then forwards the reordered flows into the appropriate service VPNs.

See the illustration to understand the sequence of events.



Benefits of PPL in Cisco SD-WAN

Benefits of PPL:

- Adaptibility: Adjusts quickly to changes in network conditions to keep performance steady.
- Bandwidth and resource utilization: Combines the capacity of multiple WAN links for higher data transfer rates.

  Uses all available network paths to make the optimal usage of bandwidth.
- Performance: Increases speed and efficiency for elephant flows.

  Supports applications that need more bandwidth than a single link can provide.

Use cases for PPL

Scenario 1: Managing AI traffic between offices and data centers

AI tasks often involve sending large amounts of information from local offices to a central data center for processing. At the same time, users need quick, real-time responses from AI tools. This traffic is unpredictable

and comes in sudden bursts. When all this data is forced through a single network path, it can create network congestion, leading to slow AI performance and delayed data uploads.

How does PPL help?

PPL breaks these large data streams into smaller, manageable pieces and sends them across all available network connections at once. This prevents any single connection from becoming overloaded.

Scenario 2: High-volume workload transfers and backups in enterprise environments

Enterprise environments routinely perform large-scale data transfers for activities such as workload upgrades, Windows operating system or application updates, data center backups, and disaster recovery operations. These scenarios often require moving significant amounts of data between endpoints. Such transfers can easily exceed the capacity of individual WAN circuits, especially when using multiple mid-sized links.

How does PPL help?

PPL splits large, single flows into multiple flowlets and distributes them across all available WAN circuits. This approach maximizes total available bandwidth, prevents any single link from being overwhelmed, and ensures that even deprioritized or non-critical traffic, like updates and backups, completes efficiently without impacting business-critical applications.

Supported platforms for PPL

These tables outline the supported hardware models and their corresponding DRAM requirements for Receiver and Sender PPL roles.

Table 2: Receiver PPL supported platforms

Model	Minimum DRAM
C8500-12X4QC	32 GB
C8500-20X6C	64 GB
C8570-G2	32 GB
C8300-2N2S-6T	16 GB
C8300-2N2S-4T2X	16 GB

Table 3: Sender PPL supported platforms

Model	Minimum DRAM
C8300-2N2S-4T2X	16 GB
C8300-2N2S-6T	16 GB
C8500L-8S4X	15 GB
C8475-G2	32 GB
C8455-G2	32 GB
C8500-12X4QC	16 GB

Model	Minimum DRAM
C8500-12X	16 GB
C8550-G2	32 GB

Note Receiver platforms can act as senders, but sender platforms cannot act as receivers.

Restrictions for PPL

Restrictions for PPL:

- Maximum senders: A receiver can only support up to 256 senders.
- Maximum receivers: A sender can only support up to 512 receivers.
- PPL data policy: You can configure PPL data policy action only in the ‘from-service’ direction.

  Do not use packet duplication (packet dup) or forward error correction (FEC) actions with PPL.
- SLA classes: Limit SLA classes to 14 to reserve one for PPL. If 15 SLA classes are configured, remove one and reload the device before configuring PPL.
- Multicast traffic: Multicast traffic is not supported with PPL.
- Service insertion: Service insertion is not supported with PPL.
- Parallel routing configurations: Sender cannot be configured with any parallel routing decisions such as next-hop, and remote-TLOC through data policy when PPL is enabled.

Configure PPL

Use one of these procedures to configure PPL.

- Configure PPL using configuration groups
- Configure a data policy with load balancing using policy groups
- Configure maximum packet value in PPL using CLI commands
- Configure latency offset value in PPL using CLI commands

Configure PPL using configuration groups

Use these steps to configure PPL using configuration groups.

Procedure

Step 1 From the Cisco SD-WAN Manager menu, choose **Configuration** > **Configuration groups**.

Step 2 Create or edit a System profile.

Step 3 Navigate to Basic Settings and select **Per Packet Load Balancing**.

Step 4 Enable **PPL** on the Sender and configure its parameters mentioned below.

Step 5 Enable packet reordering on the receiver.

Table 4: PPL fields

Field	Description
Local colors	Specify the WAN links (such as MPLS, Internet, and LTE) used to distribute traffic within the network.
Flow splitting	Specify how to divide traffic among the available tunnels. **Packets:** Sends a set number of packets (default: 1024) through one tunnel before switching to the next flowlet. This cycle repeats continuously to balance the network load. To set a maximum packet value, refer toConfigure maximum packet value in PPL using CLI commands **Packet gap:** Switches tunnels based on the time interval between packets. If the time gap between bursts exceeds the maximum latency difference among tunnels, a new flowlet is sent on a different tunnel. This helps distribute traffic more effectively during natural pauses in data transmission.
Candidate tunnel group	Dynamically defines the group of tunnels used in a PPL network. All tunnels: All available tunnels become part of PPL network. Latency offset: Includes only those tunnels whose latency is within a configurable threshold of the lowest-latency tunnel. To set latency offset value, refer to Configure latency offset value in PPL using CLI commands. Adaptive cluster: Groups tunnels dynamically into clusters based on real-time loss and latency measurements. Tunnels with similar loss (primary factor) and latency (secondary factor) are grouped together for packet distribution.
Reorder	Enables packet reordering at the receiving end to reduce out-of-order delivery. If you disable this option, ensure that your application can independently handle packet reordering.

Configure a data policy with load balancing using policy groups

Use these steps to configure data policy with load balancing.

Procedure

Step 1 From the Cisco SD-WAN Manager, choose **Configuration** > **Policy Groups** > **Application Priority and SLA**.

Step 2 Create or edit an existing policy.

Step 3 Add or create a new Traffic Policy within the policy.

Step 4 Set the direction to **From Service**.

Step 5 In the Match Clause, specify the traffic to be matched, such as the DSCP source.

Step 6 In the Action section, select the **radio button** to enable Load Balancing .

Step 7 From the **Load Balancing Algorithm** drop-down list, select **per packet**.

Step 8 Select **Save**.

Configure maximum packet value in PPL using CLI commands

Use these steps to configure the maximum packet value in PPL.

Procedure

Step 1 Enter configuration mode.

Example:

```
config-transaction
```

Step 2 Set the maximum number of packets in a flowlet.

Example:

```
sdwan ppl flow-splitting max-pkts 2048
```

Enter a value between 1024 and 10240 to define the maximum packet count for a flowlet.

Here's the complete configuration example for configuring maximum packet value in PPL.

```
sdwan
 ppl
  enable
  flow-splitting
   max-pkts 2048
   !
  !
 !
```

Configure latency offset value in PPL using CLI commands

Use these steps to configure the latency offset value in PPL.

Procedure

Step 1 Enter configuration mode.

Example:

```
config-transaction
```

Step 2 Configure the latency offset value for the CTG algorithm.

Example:

```
sdwan ppl ctg algo latency-offset 15
```

Enter the desired latency offset value to be used by the PPL CTG algorithm.

Here's the complete configuration example for configuring maximum packet value in PPL.

```
sdwan
 ppl
  ctg
    algo
     latency-offset 15
     !
    !
  !
!
```

Monitor PPL

Use these show commands to monitor PPL.

To verify the configured flow splitting and Candidate Tunnel Group (CTG) algorithms on a sender, use the **show platform software sdwan ftmd ppl cfg** command.

```
sender# show platform software sdwan ftmd ppl cfg
PPL Config
 Enable             : true
 Local Color List   : all-colors
 Path Selection     : round-robin
 OWL
   Interval : 20 sec
   DSCP     : 48 (0x30)
 Flow Splitting
   Algo     : pkt-gap adaptive

 CTG
   Algo     : adaptive-cluster
```

To verify if the PPL configuration flag is enabled for specific sites on a device, use the **show sdwan ftm site-db** command.

```
sender# show sdwan ftm site-db


  Site ID: 400, OD: No, status: Active, sys_cnt: 2, status_change: cnt
 2time 04:14:19:532, last_stats 0, is_pending_tloc_re
play=0, replay_cnt=1 Idle Timer: Not inited

  Remote    system: 172.16.255.41, OD: No, status: Active, tloc_cnt: 4,
 tun_cnt: 16, glean_ipc_cnt: 0, is_pending_tloc_repl
ay=0, replay_cnt=103, pkt_ro_ctx_id: 1, ppl_ctx_id: 1 ppl_cfg: True
ppl_adaptive_flowlet_gap: 1000
        TLOC   :: 172.16.255.41 : bronze : ipsec, nh_index: 32772, OD: No,
 status: Inactive
        TLOC   :: 172.16.255.41 : private2 : ipsec, nh_index: 32776, OD: No,
 status: Inactive
        TLOC   :: 172.16.255.41 : mpls : ipsec, nh_index: 32782, OD: No,
 status: Inactive
        TLOC   :: 172.16.255.41 : biz-internet : ipsec, nh_index: 32784, OD: No, status:
Inactive
<snip>
  Remote    system: 172.16.255.40, OD: No, status: Active, tloc_cnt: 5,
 tun_cnt: 20, glean_ipc_cnt: 0, is_pending_tloc_repl
ay=0, replay_cnt=177, pkt_ro_ctx_id: 2, ppl_ctx_id: 2 ppl_cfg: True
 ppl_adaptive_flowlet_gap: 1000
        TLOC   :: 172.16.255.40 : mpls : ipsec, nh_index: 32773, OD: No,
 status: Inactive
        TLOC   :: 172.16.255.40 : private1 : ipsec, nh_index: 32775, OD: No,
 status: Inactive
        TLOC   :: 172.16.255.40 : biz-internet : ipsec, nh_index: 32777, OD: No,
 status: Inactive
        TLOC   :: 172.16.255.40 : private2 : ipsec, nh_index: 32783, OD: No,
 status: Inactive
        TLOC   :: 172.16.255.40 : bronze : ipsec, nh_index: 32785, OD: No,
 status: Inactive
        Tunnel : 10.1.61.40/12346->10.1.41.40/12346 proto 0x32, idx 11,
sys 172.16.255.40, nh_index 65546
```

To verify Candidate Tunnel Group (CTG) selection for a destination site, use the **show sdwan bfd sessions alt** command.

```
sender# show sdwan bfd sessions alt
*Sus = Suspend
*GREinUDP = GREinUDP encap
*EAAR  = Enhanced Application-Aware Routing
*PPL = Per Packet Loadbalancing
*NA  = Flag Not Set
                                      SOURCE TLOC      REMOTE TLOC
              DST PUBLIC              DST PUBLIC
SYSTEM IP        SITE ID     STATE       COLOR            COLOR            SOURCE IP
              IP                       PORT        ENCAP  BFD-LD      FLAGS
UPTIME
------------------------------------------------------------------------------------------------------------------------
172.16.255.40    400         up          mpls             biz-internet     10.1.61.40
              10.1.40.40               12346       ipsec  20019       PPL
13:01:01:32
172.16.255.40    400         up          mpls             mpls             10.1.61.40
              10.1.41.40               12346       ipsec  20022       NA
13:01:01:31
172.16.255.40    400         up          mpls             private1         10.1.61.40
```

```
              10.1.42.40              12346       ipsec  20020       PPL
13:01:01:31
172.16.255.40    400         up          mpls             bronze          10.1.61.40
              10.1.43.40              12346       ipsec  20021       PPL
13:01:01:32
172.16.255.40    400         up          biz-internet     biz-internet    20.1.60.40
              10.1.40.40              12346       ipsec  20001       PPL
13:01:01:45
172.16.255.40    400         up          biz-internet     mpls            20.1.60.40
              10.1.41.40              12346       ipsec  20004       PPL
13:01:01:45
172.16.255.40    400         up          biz-internet     private1        20.1.60.40
              10.1.42.40              12346       ipsec  20002       PPL
13:01:01:45
172.16.255.40    400         up          biz-internet     bronze          20.1.60.40
              10.1.43.40              12346       ipsec  20003       PPL
13:01:01:45
172.16.255.40    400         up          biz-internet     private2        20.1.60.40
              10.1.44.40              12346       ipsec  20006       PPL
13:01:01:45
```

To verify absolute Candidate Tunnel Group (CTG) data and counters, use the **show platform software sdwan ftmd ppl ctg** command.

```
sender# show platform software sdwan ftmd ppl ctg
OWL-REQ RX TS: Timestamp at which the remote device received the OWL request. In brackets:
 Latency relative to tunnel with lowest OWL.
--------------------------------------------------------------------------------------------------------------------------
LOCAL COLOR     SYSTEM-IP       REMOTE COLOR    BFD-LD    PPL-ID    PKTRO-ID    EPOCH-ID
  OWL-REQ RX TS                PKT-TX      BYTES-TX    PKT-RX      BYTES-RX
--------------------------------------------------------------------------------------------------------------------------
biz-internet    172.16.255.40   biz-internet    20001     1         65535       55905
  1119300231(0)                 0           0           0         0
biz-internet    172.16.255.40   private1        20002     1         65535       55905
  1119300231(0)                 0           0           0         0
biz-internet    172.16.255.40   bronze          20003     1         65535       55905
  1119300231(0)                 0           0           0         0
biz-internet    172.16.255.40   mpls            20004     1         65535       55905
  1119300231(0)                 0           0           0         0
biz-internet    172.16.255.40   private2        20006     1         65535       55905
  1119300231(0)                 0           0           0         0
mpls            172.16.255.40   biz-internet    20019     1         65535       55905
  1119300231(0)                 0           0           0         0
mpls            172.16.255.40   private1        20020     1         65535       55905
  1119300231(0)                 0           0           0         0
mpls            172.16.255.40   bronze          20021     1         65535       55905
  1119300231(0)                 0           0           0         0
biz-internet    172.16.255.41   bronze          20005     4         65535       55905
  1119139948(0)                 0           0           0         0
biz-internet    172.16.255.41   biz-internet    20007     4         65535       55905
  1119139948(0)                 0           0           0         0
biz-internet    172.16.255.41   private2        20008     4         65535       55905
  1119139948(0)                 0           0           0         0
biz-internet    172.16.255.41   mpls            20009     4         65535       55905
  1119139948(0)                 0           0           0         0
mpls            172.16.255.41   bronze          20023     4         65535       55905
  1119139948(0)                 0           0           0         0
mpls            172.16.255.41   biz-internet    20025     4         65535       55905
  1119139948(0)                 0           0           0         0
mpls            172.16.255.41   private2        20026     4         65535       55905
  1119139948(0)                 0           0           0         0
```

```
mpls            172.16.255.41   mpls            20027     4         65535       55905
  1119139948(0)                 0           0           0           0
```

To verify the cumulative tunnel statistics, use the **show sdwan tunnel statistics table** command.

```
sender# show sdwan tunnel statistics ppl table
TUNNEL                          SOURCE  DEST
 PROTOCOL  SOURCE IP   DEST IP     PORT    PORT   ppl-tx-ipv4-pkts  ppl-tx-ipv6-pkts
ppl-tx-ipv4-octets  ppl-tx-ipv6-octets  ppl-rx-ipv4-pkts  ppl-rx-ipv6-pkts  ppl-rx-ipv4-octets
  ppl-rx-ipv6-octets
ipsec     10.1.61.40  10.1.40.40  12346   12346  1024             0                 1224704
             0                   0                  0                 0                   0
ipsec     10.1.61.40  10.1.42.40  12346   12346  1024             0                 1224704
             0                   0                  0                 0                   0
ipsec     10.1.61.40  10.1.43.40  12346   12346  1024             0                 1224704
              0                   0                  0                 0
0
ipsec     20.1.60.40  10.1.40.40  12346   12346  1024             0                 1224704
              0                   0                  0                 0
0
ipsec     20.1.60.40  10.1.41.40  12346   12346  1024             0                 1224704
              0                   0                  0                 0
0
ipsec     20.1.60.40  10.1.42.40  12346   12346  1024             0                 1224704
             0                   0                  0                 0                   0
ipsec     20.1.60.40  10.1.43.40  12346   12346  1024             0                 1224704
             0                   0                  0                 0                   0
ipsec     20.1.60.40  10.1.44.40  12346   12346  1024             0                 1224704
             0                   0                  0                 0                   0
```

To verify the configured packet reordering parameters on a receiver, use the **show platform software pkt-reorder cfg** command.

```
receiver# show platform software pkt-reorder cfg
Enable                : true
Max-allowed-memory    : 40%
Packets-per-context   : 8192
Out-of-order-tolerance : 2048
Dropped-packet-timeout : 100 ms
Idle-timeout          : 1000 ms
Dispact-size          : 2048
```

To check if the packet re-order context ID is correctly assigned for remote senders on a receiver device, use the **show sdwan ftm site-db** command.

```
receiver# show sdwan ftm site-db 300
OD Tunnel Idle Time   : 10 min
 Site ID: 300, OD: Yes, status: Inactive, sys_cnt: 1, status_change:
cnt ltime 08:59:05:953, last_stats 0, is_pending_tloc
_replay=0, replay_cnt=2 Idle Timer: Not running, time_to_expiry: 0 sec
 Remote   system: 172.16.255.80, OD: Yes, status: Inactive,
tloc_cnt: 4, tun_cnt: 20, glean_ipc_cnt: 0, is_pending_tloc_r
eplay=0, replay_cnt=1010, pkt_ro_ctx_id: 7, ppl_ctx_id: 5 ppl_cfg:
True ppl_adaptive_flowlet_gap: 1000
      TLOC  :: 172.16.255.80 : privatel : ipsec, nh_index: 32800,
OD: Yes, status: Inactive
      TLOC  :: 172.16.255.80 : bronze : ipsec, nh_index: 32801, OD:
Yes, status: Inactive
      TLOC  :: 172.16.255.80 : mpls : ipsec, nh_index: 32803, OD:
Yes, status: Inactive
      TLOC  :: 172.16.255.80 : biz-internet : ipsec, nh_index:
32804, OD: Yes, status: Inactive
      Tunnel : 10.1.41.40/12346->10.1.62.40/12346 proto 0x32, idx
960, sys 172.16.255.80, nh_index 66495
```

To view the reordering statistics of a PPL network, use the **show platform software pkt-reorder ctx sdwan** command.

```
receiver# show platform software pkt-reorder ctx sdwan
Context ID Module IP Address Total Packet Time Exceeded Too Old In Window Wrap DM Distance
 Mem Usage Total Mem Usage
1 sdwan 172.16.255.80 0 0 0 0 0 0 0 0
2 sdwan 172.16.255.81 132,213,316 403 130,239,435 829,845 1 43,955,369,702,422,670 0 0
3 sdwan 172.16.255.83 0 0 0 0 0 0 0 0
4 sdwan 172.16.255.82 0 0 0 0 0 0 0 0
```