Understand how to Carve Nexus 9000 TCAM Space

Contents

Introduction

Background Information

- Terminology
- ACL TCAM Regions
- Nexus 9500 Series TCAM Allocation
- Nexus 9300 Series TCAM Allocation

Configuration

- Example Scenario
- Verification Commands

Errors and Solutions

Design Guidelines and Limitations

Related Information

Introduction

This document describes how to carve Nexus 9000 ternary content-addressable memory (TCAM).

Background Information

This document is not intended to be an exhaustive list of the many TCAM combinations. The purpose of this document is to help users understand how the TCAM allocation works so that they can determine valid configurations that meet their needs. It covers the current and most common concepts, configuration, and error messages.

To use a non-default feature for Nexus 9000 Series switches, one must manually carve out TCAM space for the features. By default, all TCAM space is allocated.

NOTE:
The device on which this document is based is not Cloud-Scale or Cisco Silicon One. Therefore, there may be variations in region names, default values, and other related elements.

Terminology

- Feature Width - There are single-width and double-width features. A single-width feature requires a minimum of one slice. A double-width feature at minimum requires two slices.

For both single-and double-width features, the total size, if greater than 256, must be a multiple of 512. A slice can be allocated to one region only.

For example, you cannot use a 512-size slice in order to configure two features of size 256 each, nor can you use a 512-size slice in order to configure a single double-width feature.

- Slice - A unit of memory allocation. Slices can be of size 256 or of size 512, measured in bytes.
- TCAM - Ternary Content Addressable Memory. This is the space in hardware where access-lists (ACLs) are stored. This is a specialized piece of memory that stores complex tabular data and supports very rapid parallel lookups.

ACL TCAM Regions

You can change the size of the ACL TCAM regions in the hardware. The egress TCAM size is 1K, divided into four 256 entries. The ingress TCAM size is 4K, divided into eight 256 slices and four 512 slices.

The IPv4 TCAM regions are single wide. The IPv6, quality of service (QoS), MAC, control-plane policing (CoPP), and system TCAM regions are double wide and consume double the physical TCAM entries.

For example, a logical region size of 256 entries actually consumes 512 physical TCAM entries.

You can create IPv6, port ACLs (PACLs), VLAN ACLs (VACLs), and router ACLs (RACLs), and you can match IPv6 and MAC addresses for QoS. However, Cisco NX-OS cannot support all of them simultaneously.

You must remove or reduce the size of the current TCAM regions in order to enable the IPv6 and MAC TCAM regions. For every TCAM region configuration command, the system evaluates if the new change can fit in the TCAM.

If not, it reports an error, and the command is rejected. You must remove or reduce the size of current TCAM regions to make room for new requirements.

ACL TCAM region sizes have these guidelines and limitations:

- On Cisco Nexus 9500 Series switches, the default ingress TCAM region configuration has one free 256-entry slice in Cisco NX-OS Release 6.1(2)I1(1).

This slice is allocated to the switch port analyzer (SPAN) region in Cisco NX-OS Release 6.1(2)I2(1). Similarly, the RACL region is reduced from 2K to 1.5K in Cisco NX-OS Release 6.1(2)I2(1) in order to make room for the virtual port-channel (vPC) convergence region with 512 entries.

- On Cisco Nexus 9300 Series switches, the Application Centric Infrastructure (ACI) leaf line card is used in order to enforce the QoS classification policies applied on 40G ports. It has 768 TCAM entries available to carve in 256-entry granularity. These region names are prefixed with ns-.
- For the ACI leaf line card on Cisco Nexus 9300 Series switches, only the IPv6 TCAM regions consume double-wide entries. The rest of the TCAM regions consume single-wide entries.
- When a VACL region is configured, it is configured with the same size in both the ingress and egress directions. If the region size cannot fit in either direction, the configuration is rejected.

Both the Nexus 9300 and 9500 Series switches have four slices of size 512 bytes and eight slices of size 256 bytes. By default, all slices and all space are used, though the default allocation is different between the Nexus 9300 series and the 9500 series.

 **Note**: The Nexus 9332PQ uses the same default allocation as the Nexus 9500.

Nexus 9500 Series TCAM Allocation

The Nexus 9500 Series switches have this TCAM allocation by default:

<#root>

Nexus9500#

show system internal access-list globals

slot 1
=======

  Atomic Update : ENABLED
  Default ACL   : DENY
  Bank Chaining : DISABLED
  Fabric path DNL : DISABLED
  NS Buffer Profile: Mesh optimized
  Min Buffer Profile: all
  EOQ Class Stats: qos-group-0
  NS MCQ3 Alias: qos-group-3
  Ing PG Share: ENABLED

  LOU Threshold Value : 5

INSTANCE 0 TCAM Region Information:

Ingress:

Region	GID	Base	Size	Width
IPV4 PACL [ifacl]	3	0	0	1
IPV6 PACL [ipv6-ifacl]	4	0	0	2
MAC PACL [mac-ifacl]	5	0	0	2
IPV4 Port QoS [qos]	6	0	0	2
IPV6 Port QoS [ipv6-qos]	7	0	0	2
MAC Port QoS [mac-qos]	8	0	0	2
FEX IPV4 PACL [fex-ifacl]	9	0	0	1
FEX IPV6 PACL [fex-ipv6-ifacl]	10	0	0	2
FEX MAC PACL [fex-mac-ifacl]	11	0	0	2
FEX IPV4 Port QoS [fex-qos]	12	0	0	2
FEX IPV6 Port QoS [fex-ipv6-qos]	13	0	0	2
FEX MAC Port QoS [fex-mac-qos]	14	0	0	2
IPV4 VACL [vacl]	15	0	0	1
IPV6 VACL [ipv6-vacl]	16	0	0	2
MAC VACL [mac-vacl]	17	0	0	2
IPV4 VLAN QoS [vqos]	18	0	0	2
IPV6 VLAN QoS [ipv6-vqos]	19	0	0	2
MAC VLAN QoS [mac-vqos]	20	0	0	2
IPV4 RACL [racl]	**21**	**0**	**1536**	**1**
IPV6 RACL [ipv6-racl]	22	0	0	2
IPV4 Port QoS Lite [qos-lite]	61	0	0	1
FEX IPV4 Port QoS Lite [fex-qos-lite]	62	0	0	1
IPV4 VLAN QoS Lite [vqos-lite]	63	0	0	1
IPV4 L3 QoS Lite [l3qos-lite]	64	0	0	1
IPV4 L3 QoS [l3qos]	**37**	**3072**	**256**	**2**
IPV6 L3 QoS [ipv6-l3qos]	38	0	0	2

```
              MAC L3 QoS [mac-l3qos]     39        0        0        2

   Ingress System      1     2048      256        2

                         SPAN [span]      2     4096      256        1

                  Ingress COPP [copp]     40     2560      256        2

         Ingress Flow Counters [flow]     43        0        0        1
           Ingress SVI Counters [svi]     45        0        0        1

                  Redirect [redirect]     46     3840      256        1

            NS IPV4 Port QoS [ns-qos]     47        0        0        1
       NS IPV6 Port QoS [ns-ipv6-qos]     48        0        0        2
         NS MAC Port QoS [ns-mac-qos]     49        0        0        1
           NS IPV4 VLAN QoS [ns-vqos]     50        0        0        1
      NS IPV6 VLAN QoS [ns-ipv6-vqos]     51        0        0        2
        NS MAC VLAN QoS [ns-mac-vqos]     52        0        0        1
            NS IPV4 L3 QoS [ns-l3qos]     53        0        0        1
       NS IPV6 L3 QoS [ns-ipv6-l3qos]     54        0        0        2
         NS MAC L3 QoS [ns-mac-l3qos]     55        0        0        1

    VPC Convergence [vpc-convergence]     57     1536      512        1

--------------------------------------------------------------------------
* - allocated 512 entry slice due to unavailability of 256 entry slices
--------------------------------------------------------------------------
Total: 4096

--------------------------------------------------------------------------
Egress
----------
                         Region            GID     Base     Size     Width
--------------------------------------------------------------------------
               Egress IPV4 VACL [vacl]     31        0        0        1
          Egress IPV6 VACL [ipv6-vacl]     32        0        0        2
            Egress MAC VACL [mac-vacl]     33        0        0        2

             Egress IPV4 RACL [e-racl]     34     4352      768        1

        Egress IPV6 RACL [e-ipv6-racl]     35        0        0        2

                        Egress System     24     3584      256        1

         Egress Flow Counters [e-flow]     44        0        0        1
--------------------------------------------------------------------------
Total: 1024

--------------------------------------------------------------------------
```

The slice allocation is as shown for ingress:

Slice 1 (512): RACL

Slice 2 (512): RACL

Slice 3 (512): RACL

Slice 4 (512): VPC Convergence

Slice 5 (256): Layer 3 QOS

Slice 6 (256): Layer 3 QOS

Slice 7 (256): SPAN

Slice 8 (256): REDIRECT

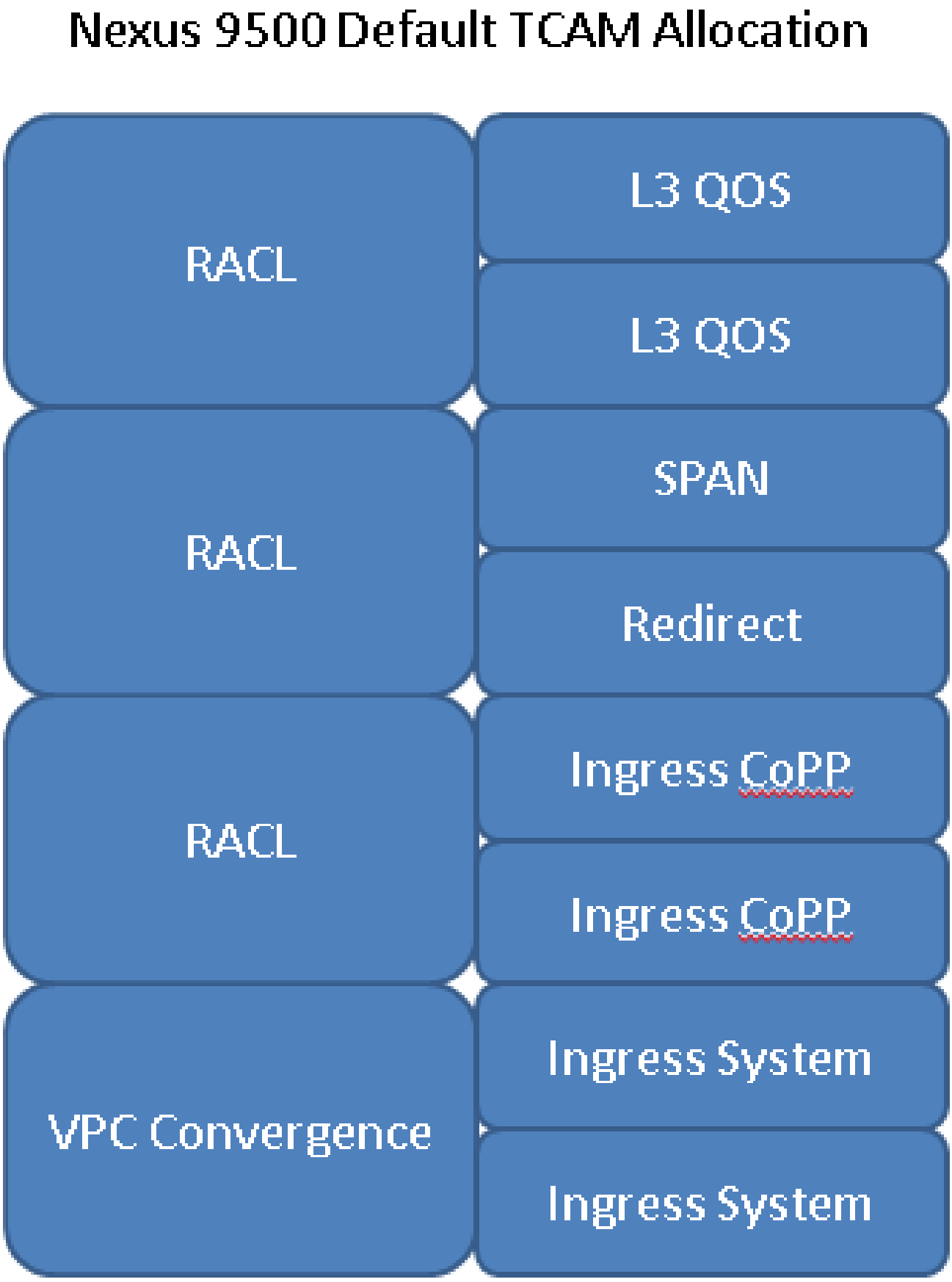
Slice 9 (256): Ingress CoPP

Slice 10 (256): Ingress CoPP

Slice 11 (256): Ingress System

Slice 12 (256): Ingress System

Ingress Utilization Conceptualized:



Nexus 9300 Series TCAM Allocation

The Nexus 9300 Series switches have this TCAM allocation by default:

<#root>

Nexus9300#

show system internal access-list globals

slot 1
=======

      Atomic Update : ENABLED
      Default ACL   : DENY
      Bank Chaining : DISABLED
      Fabric path DNL : DISABLED
      NS Buffer Profile: Burst optimized
      Min Buffer Profile: all
      EOQ Class Stats: qos-group-0
      NS MCQ3 Alias: qos-group-3
      Ing PG Share: ENABLED

      LOU Threshold Value : 5

```
------------------------------------------------------------------
              INSTANCE 0 TCAM Region Information:
------------------------------------------------------------------
Ingress:
----------
                        Region                          GID     Base     Size     Width
------------------------------------------------------------------

        IPV4 PACL [ifacl](   1)       3        0      512        1

                  IPV6 PACL [ipv6-ifacl](   2)       4        0        0        2
                    MAC PACL [mac-ifacl](   3)       5        0        0        2

                 IPV4 Port QoS [qos](   4)       6     3072      256        2

                IPV6 Port QoS [ipv6-qos](   5)       7        0        0        2
                  MAC Port QoS [mac-qos](   6)       8        0        0        2
              FEX IPV4 PACL [fex-ifacl](   7)       9        0        0        1
          FEX IPV6 PACL [fex-ipv6-ifacl](   8)      10        0        0        2
            FEX MAC PACL [fex-mac-ifacl](   9)      11        0        0        2
             FEX IPV4 Port QoS [fex-qos](  10)      12        0        0        2
         FEX IPV6 Port QoS [fex-ipv6-qos](  11)      13        0        0        2
           FEX MAC Port QoS [fex-mac-qos](  12)      14        0        0        2

                    IPV4 VACL [vacl](  13)      15      512      512        1

                  IPV6 VACL [ipv6-vacl](  14)      16        0        0        2
                    MAC VACL [mac-vacl](  15)      17        0        0        2
                   IPV4 VLAN QoS [vqos](  16)      18        0        0        2
                IPV6 VLAN QoS [ipv6-vqos](  17)      19        0        0        2
                  MAC VLAN QoS [mac-vqos](  18)      20        0        0        2

                    IPV4 RACL [racl](  19)      21     1024      512        1

                  IPV6 RACL [ipv6-racl](  20)      22        0        0        2
```

```
          IPV4 Port QoS Lite [qos-lite](  21)      63        0        0        1
    FEX IPV4 Port QoS Lite [fex-qos-lite](  22)      64        0        0        1
         IPV4 VLAN QoS Lite [vqos-lite](  23)      65        0        0        1
          IPV4 L3 QoS Lite [l3qos-lite](  24)      66        0        0        1
                    IPV4 L3 QoS [l3qos](  34)      37        0        0        2
                IPV6 L3 QoS [ipv6-l3qos](  35)      38        0        0        2
                  MAC L3 QoS [mac-l3qos](  36)      39        0        0        2

                     Ingress System(  37)       1     2048      256        2
                            SPAN [span](  39)       2     3584      256        1
                     Ingress COPP [copp](  40)      40     2560      256        2

            Ingress Flow Counters [flow](  41)      43        0        0        1
              Ingress SVI Counters [svi](  43)      45        0        0        1

                 Redirect [redirect](  44)      46     1536      512        1

               NS IPV4 Port QoS [ns-qos](  45)      47        0        0        1
          NS IPV6 Port QoS [ns-ipv6-qos](  46)      48        0        0        2
            NS MAC Port QoS [ns-mac-qos](  47)      49        0        0        1
              NS IPV4 VLAN QoS [ns-vqos](  48)      50        0        0        1
         NS IPV6 VLAN QoS [ns-ipv6-vqos](  49)      51        0        0        2
           NS MAC VLAN QoS [ns-mac-vqos](  50)      52        0        0        1
               NS IPV4 L3 QoS [ns-l3qos](  51)      53        0        0        1
          NS IPV6 L3 QoS [ns-ipv6-l3qos](  52)      54        0        0        2
            NS MAC L3 QoS [ns-mac-l3qos](  53)      55        0        0        1

    VPC Convergence [vpc-convergence](  54)      57     4096      256        1

           IPSG SMAC-IP bind table [ipsg](  55)      59        0        0        1
        Ingress ARP-Ether ACL [arp-ether](  56)      62        0        0        1
------------------------------------------------------------------------
* - allocated 512 entry slice due to unavailability of 256 entry slices
------------------------------------------------------------------------
Total: 4096

------------------------------------------------------------------
Egress
----------
                            Region                  GID     Base     Size     Width
------------------------------------------------------------------
                 Egress IPV4 QoS [e-qos](  25)      28        0        0        2
             Egress IPV6 QoS [e-ipv6-qos](  26)      29        0        0        2
               Egress MAC QoS [e-mac-qos](  27)      30        0        0        2

              Egress IPV4 VACL [vacl](  28)      31     4352      512        1

             Egress IPV6 VACL [ipv6-vacl](  29)      32        0        0        2
               Egress MAC VACL [mac-vacl](  30)      33        0        0        2

            Egress IPV4 RACL [e-racl](  31)      34     4864      256        1

           Egress IPV6 RACL [e-ipv6-racl](  32)      35        0        0        2
        Egress IPV4 QoS Lite [e-qos-lite](  33)      36        0        0        1
```

```
                        Egress System(  38)      24     3840      256        1

            Egress Flow Counters [e-flow](  42)      44        0        0        1
------------------------------------------------------------------------------
Total: 1024

------------------------------------------------------------------
```

Slice 1 (512): IPv4 PACL

Slice 2 (512): VACL

Slice 3 (512): RACL

Slice 4 (512): Redirect

Slice 5 (256): Port QOS

Slice 6 (256): Port QOS

Slice 7 (256): SPAN

Slice 8 (256): VPC Convergence

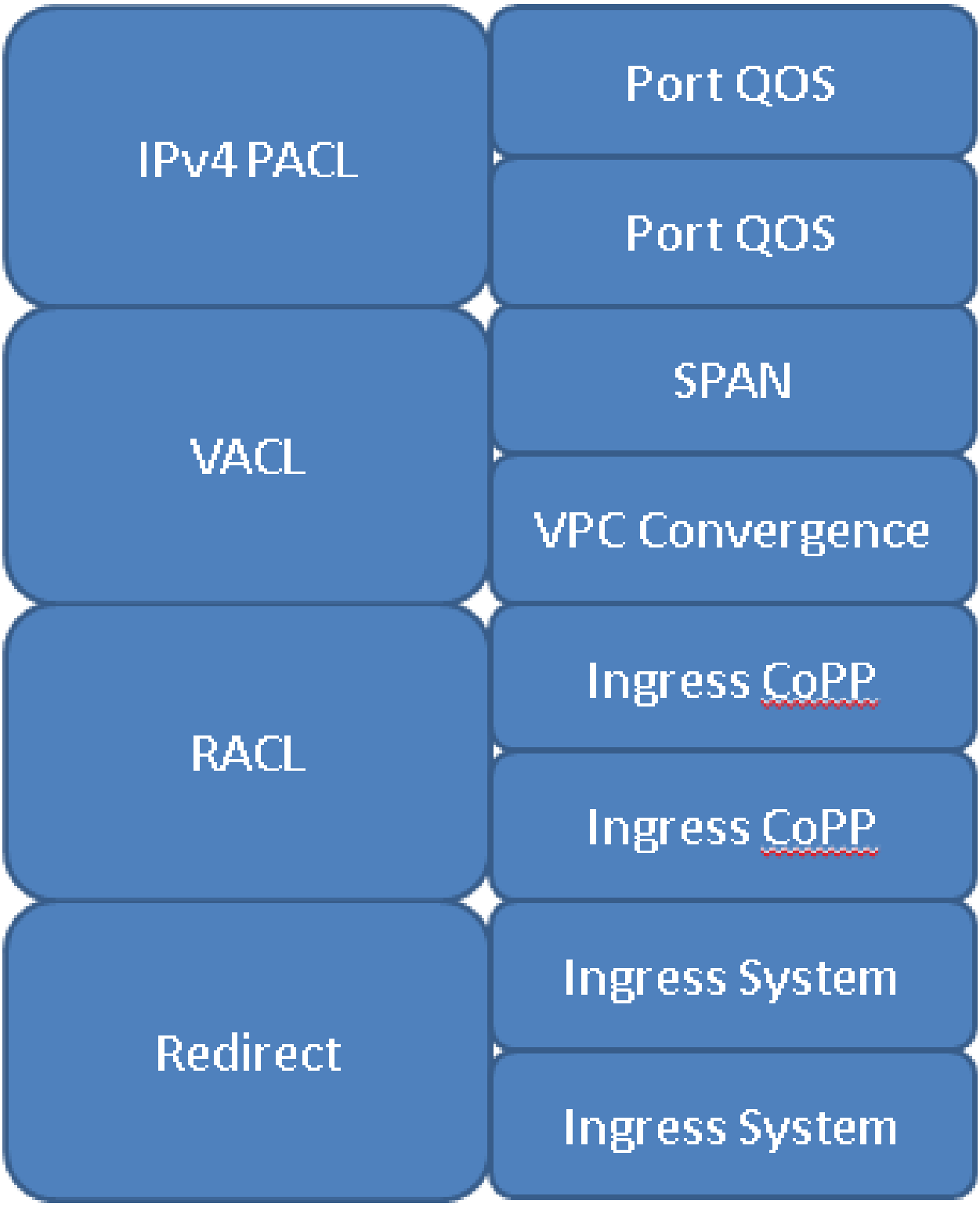
Slice 9 (256): Ingress CoPP

Slice 10 (256): Ingress CoPP

Slice 11 (256): Ingress System

Slice 12 (256): Ingress System

Ingress Utilization Conceptualized:



Configuration

In order to reconfigure a TCAM region, use the **hardware access-list tcam region <*feature_name*> <*feature_size*>** command in the configuration terminal. Once you have changed the regions to be the intended sizes, you must reload the device.

Example Scenario

You have a Nexus 9300 and want to allocate the TCAM space in order to best fit your needs. You need to free 512 bytes of TCAM. This allows you to add more to IPv4 PACL.

However, you decide that you do not need 512 VACL or 512 RACL, but need some of both so you decide to unallocate 256 bytes from VACL and RACL. This frees up 512 space as these commands show:

```
<#root>

Nexus9300(config)#

hardware access-list tcam region vacl 256


Warning: Please save config and reload the system for the configuration to take effect

Nexus9300(config)#

hardware access-list tcam region racl 256


Warning: Please save config and reload the system for the configuration to take effect
```

With 512 bytes free, you try to allocate an additional 512 to IPv4 PACL, but see this output:

```
<#root>

Nexus9300(config)#

hardware access-list tcam region ifacl 1024


ERROR: Aggregate TCAM region configuration exceeded the available Ingress TCAM slices.
Please re-configure.
```

Even though 512 bytes were freed up, both the VACL and RACL space, which 256 were pulled from, were size 512 blocks. As such, the previous commands unallocated space, but did not unallocate any slices. In order to increase the IPv4 PACL size to 1024, you need to take 512 bytes from a single feature which frees up both a slice and space:

```
<#root>

Nexus9300(config)#

hardware access-list tcam region vacl 512


Warning: Please save config and reload the system for the configuration to take effect

Nexus9300(config)#

hardware access-list tcam region racl 0


Warning: Please save config and reload the system for the configuration to take effect
```

```
Nexus9300(config)#
```

```
hardware access-list tcam region ifacl 1024
```

```
Warning: Please save config and reload the system for the configuration to take effect
```

Verification Commands

- **show hardware access-list tcam region** - Verifies the current software configuration.
- **show system internal access-list globals** - Verifies the current hardware configuration.
- **show system internal access-list input entries detail** - Shows the specific ACLs configured for each instance.
- **show hardware access-list resource utilization** - Shows the current utilization of each configured TCAM region.
- **show hardware access-list resource entries** - Shows the number of ACL entries configured for each instance

Errors and Solutions

These are the common errors seen during TCAM configuration:

```
ERROR: Aggregate TCAM region configuration exceeded the available
Ingress TCAM slices. Please re-configure.
```

This error occurs when you try to configure a valid amount of TCAM space with regards to the 4k limit, but your allocation consumes more slices than are available.

The only solution for this error is to revise your overall TCAM design in order to free up slices.

This error is more common when you try to configure a new double-width feature, as they require at least two slices of 256 or 512.

```
ERROR: Aggregate TCAM region configuration exceeded the available
Ingress TCAM space. Please re-configure.
```

Similar to the slice error, the solution is to reconfigure the allocated space not to exceed the total limit. This error message is only seen when all TCAM slices have already been allocated and you try to allocate more space.

```
ERROR: TCAM regions with size more than 256, ... have size
in multiple of 512 entries
```

Due to hardware limitations, TCAM sizes greater than 256 cannot be combined in any manner that combines an odd number of 256 blocks with a 512 block. For this reason, when you configure a TCAM region that is larger than 512, the only valid sizes are multiples of 512.

Design Guidelines and Limitations

TCAM space is limited. The choice for what is best for you depends entirely on the specific use case. By default, all TCAM space is already allocated, so you need to decide where you want to allocate TCAM space in order to allocate elsewhere.

- In the case of ingress, four of the eight available size-256 slices cannot be unallocated (used by CoPP and ingress system).
- One 256 slice is used by SPAN. If you borrow from this, it removes the ability to use the SPAN and Packet-Tracer features completely (not recommended to remove for troubleshoot purposes).
- A size 256 or 512 slice is used for vPC on the Nexus 9300 and 9500 platforms respectively. A reallocation removes the ability to use vPC
- A size 512 or 256 slice is used for Redirect on the Nexus 9300 and 9500 platforms respectively. If you borrow from this, it removes the ability to use DHCPv4, DHCPv6, or BFD.
- If atomic updates are enabled, and you are over 50% utilization for one TCAM feature, you cannot remove a line from any ACL due to lack of space.
- By default QoS policy applied on multiple interfaces does not share the label since statistics are enabled by default. In order to share the label for the same QoS policy applied on multiple interfaces, you have to configure the QoS policy with the no-stats option as this example shows:
  <#root>

  ```
  (config-if)#

  service-policy type qos input my-policy no-stats
  ```

- Where possible, use the basic version of features. With the basic versions, the switch uses half of the TCAM space for that feature. This causes a double-width feature to be single-width. The cost is that the feature does not keep track of violated policer statistics; it only keeps track of conformed policer statistics. This is usually the better option as it saves TCAM space.

- Users cannot reduce the default amount of Ingress System and CoPP TCAM. These are already at the minimum value and cannot be reduced.
- All QoS features are double-width.
- SVI policy-maps are not supported.

Related Information

- Nexus 9000 TCAM Carving Configuration Guide
- Nexus 9000 ACL TCAM Regions
- Cisco Technical Support & Downloads