Understand ACL Resources on Cisco Nexus 5000 Switches

White Paper

August, 2011

What You Will Learn

This document provides information about how the Cisco Nexus® 5000 Series Switches program access control lists (ACLs) in hardware, and how in certain configurations planning helps prevent exhaustion of the hardware resources relating to ACLs. It explains the ACL merge algorithms and the hardware resources used in Cisco Nexus 5000 Switches to enforce security and apply quality of service (QoS) using router ACLs (RACLs), VLAN ACLs (VACLs), and port ACLs (PACLs). It explains the platform specifics, how to check the resources available, and how to optimize them in order to scale, considering the existing platform architecture limits.

The information discussed in this document is general and does not apply to specific software releases. However, later software releases are likely to include more ACL resource-usage optimizations. It is recommended to read the release notes for your platform before upgrading Cisco Nexus NX-OS software.

Overview

Understanding the ACL capacities when configuring ACLs on the Cisco Nexus 5000 Series Switches helps avoid resource contention and exhaustion Because the platform enforces several types of ACLs in hardware rather than in software, the switch programs hardware lookup tables and various hardware registers in the unified port controller (UPC) so that when a packet arrives the switch can perform a hardware table lookup and perform the appropriate action without affecting performance, while the packets are cut-through switched.

For typical configurations, the main hardware resources are the logical operation units (LOUs) that are combined on this platform between Layer 3 operations (L3Ops) and Layer 4 operations (L4Ops), the ACL-to-switch interface mapping labels, and the available space in the ternary content addressable memory (TCAM). TCAM entries, LOUs, and ACL labels are fixed resources. Therefore, depending on your ACL configuration, you may need to optimize your ACEs to optimize the use of available resources. The available hardware resources differ on Cisco Nexus 5000 Series and Cisco Nexus 5500 Series Switches.

The resources shown in Table 1 are discussed in detail later in this document.

Table 1. Overview of the Cisco Nexus 5000 Series Resources

Nexus Model	TCAM	LOU
5000 series **5010/5020**	1024 VACLs 576 PACLs (ifACL)/UPC 192 QoS 64 SPAN 30 COPP reserved	L4Ops: 8 UDP and 8 TCP L3OPs: as large as the TCAM size
5500 series **5548P** **5548UP** **5596UP**	2048 VACLs 1152 PACLs (ifACL)/UPC 448 QoS 64 SPAN 30 COPP reserved With Layer 3 module 1664 ingress RACLs 2048 egress RACLs	L4Ops: 8 UDP and 8 TCP L3OPs: as large as the TCAM size

Definitions

An **access control list (ACL)** is a means of classifying data.

A **router ACL (RACL)** is an ACL that is applied to an interface that has a Layer 3 address assigned to it. It can be applied to any port that has an IP address: routed interfaces, loopback interfaces, VLAN interfaces, etc. The security boundary is to permit or deny traffic moving between subnets or networks. The RACL is supported in hardware, and has no effect on the performance.

An example follows:

```
ip access-list extended racl
  permit ip host 10.1.1.2 host 192.168.1.14
interface e1/1
  no switchport
  ip address 192.168.1.1 255.255.255.0
  ip access-group racl in
```

Verification:

```
5548-1-LA# show ip access-lists summary

IP access list pacl
        Total ACEs Configured:1

        Configured on interfaces:
                                Ethernet1/1 - ingress (Router ACL)

        Active on interfaces:
                                Ethernet1/1 - ingress (Router ACL)
```

A **VLAN access control list (VACL)** is an ACL that is applied to a VLAN. It can be applied only to a VLAN - no other type of interface. The security boundary is to permit or deny moving traffic between VLANs and permit or deny traffic within a VLAN. The VLAN ACL is supported in hardware, and has no effect on the performance.

An example follows:

```
ip access-list extended vacl
  permit ip any 10.1.1.0 0.0.0.255
vlan access-map myvacl 10
  match ip address vacl
  action forward
vlan filter myvacl vlan-list 10-15
```

A **port access control list (PACL)** entry is an ACL applied to a Layer 2 switchport interface. It cannot be applied on any other type of interfaces, and it works only in the ingress direction. The security boundary is to permit or deny traffic within a VLAN. The PACL is supported in hardware and has no effect on the performance.

An example follows:

```
ip access-list extended pacl
  permit tcp any any
interface Ethernet 1/1
  switchport
  ip port access-group pacl in
```

Verification:

```
5548-1-LA# show ip access-lists summary

IP access list pacl
        Total ACEs Configured:1

        Configured on interfaces:
                                Ethernet1/1 - ingress (Port ACL)

        Active on interfaces:
                                Ethernet1/1 - ingress (Port ACL)
```

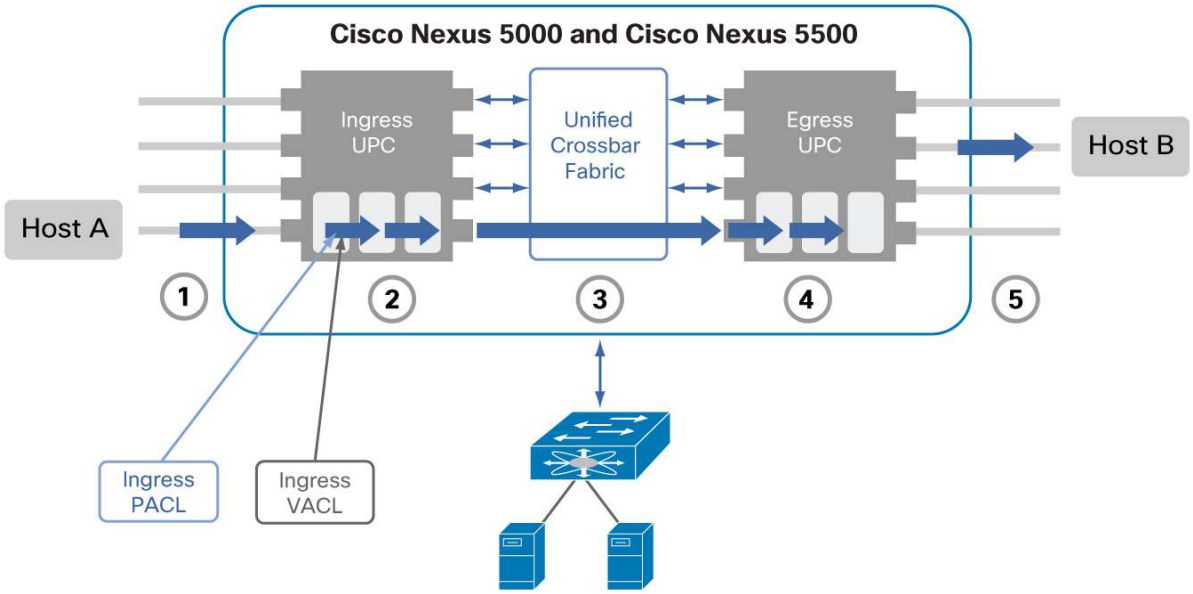
Note: The PACL resources are called ifacl in the commands covered in the section “Monitoring TCAM and Hardware Resource Usage”.

Note: A RACL and VACL cannot coexist on a switch virtual interface (SVI). The first one preempts the configuration and an error message will be displayed if preemption is attempted: “ERROR: RACL and VACL together is not supported on this platform”.

Packet Forwarding Overview

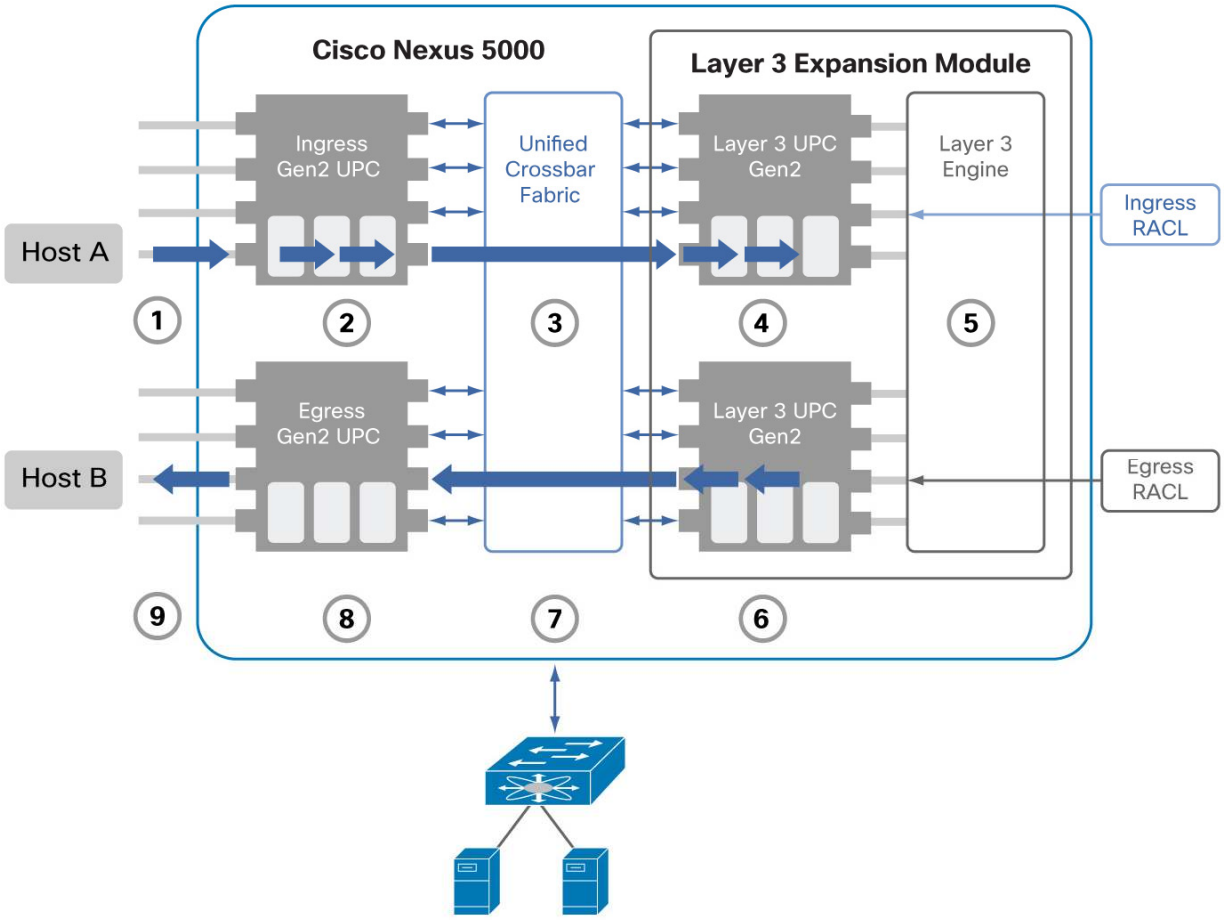
This section explains the packet forwarding across the Cisco Nexus 5000 and Nexus 5500 Series for Layer 2 (Figure 1) and Layer 3 packet forwarding occurring in the Cisco Nexus 5500 Series equipped with the Layer 3 daughter card or expansion module (Figure 2). Understanding the packet forwarding on this platform can help in understanding where the access lists are applied and the related hardware forwarding engines to look at, as covered in the section “Monitoring TCAM and Hardware Resource Usage” later in this document.

Figure 1. Cisco Nexus 5500 and Nexus 5000 Layer 2 Packet Forwarding Overview



1. Packet arrives ingress from host A: MAC address is decoded and bytes are synchronized.
2. Ingress is processed at the UPC.
   a. Ingress forwarding logic: The frame is first parsed, and then the forwarding logic is performed as well as filtering searches and learning.
   b. Ingress buffer (virtual output queue) role is to: queue frames, request service of fabric, de-queue frames to fabric, and monitor queue usage to trigger congestion control.
   c. PACL and VACL are applied at the ingress UPC in this respective order.
3. Unified crossbar fabric processing: Scheduler determines fairness of access to the fabric and determines when frame is de-queued across the fabric.
4. Egress processing at the UPC.
   a. Egress buffers: These buffers are the landing spots for frames in flight when egress is paused.
   b. Egress forwarding logic: this logic applies parsing, fields extraction then performs learning and filtering searches, perform learning, and finally convert to desired egress format.
5. Packet leaves egress to host B: MAC address is encoded, bytes are synchronized, and the packet is transmitted.

Figure 2. Cisco Nexus 5500 Layer 3 Packet Forwarding Overview



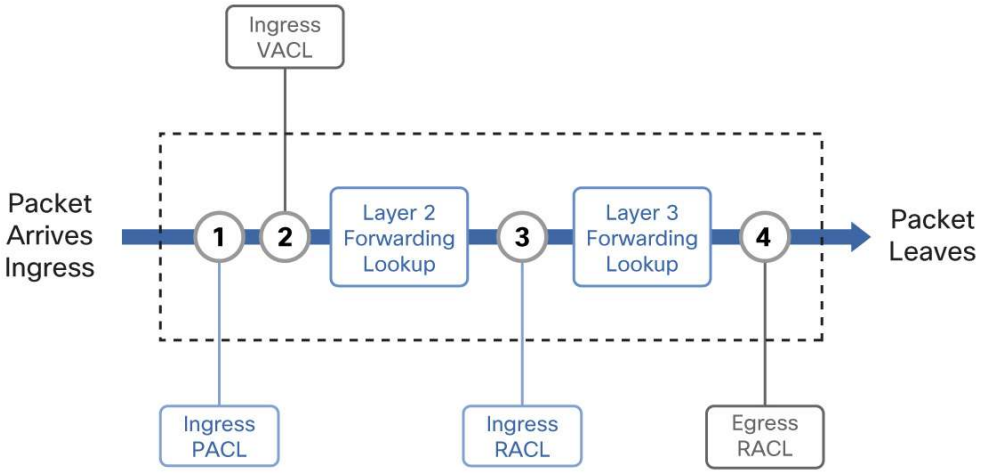
The Layer 3 forwarding engine is connected to the unified crossbar fabric through two Layer 3 UPCs, providing a throughput capacity of the Layer 3 engine up to 160 Gbps. It is an optional two-stage forwarding, where only the Layer 3 traffic to be processed is sent.

1. Packet arrives ingress from host A: MAC address is decoded and bytes are synchronized.
2. Ingress processing at the UPC.
    a. Ingress forwarding logic: the frame is first parsed, and then the forwarding logic is performed as well as filtering searches and learning.
    b. Ingress buffer (virtual output queue) role is to: queue frames, request service of fabric, de-queue frames to fabric, and monitor queue usage to trigger congestion control.
3. Unified crossbar fabric processing: Scheduler determines fairness of access to the fabric and determines when frame is de-queued across the fabric.
4. When MAC address is the router MAC address (example: Hot Standby Router Protocol [HSRP] virtual MAC), the packet is forwarded across fabric to Layer 3 UPC. Ingress RACLs are applied.
5. The Layer 3 lookup occurs and routing is performed in the Layer 3 engine.

6. The packet arrives at the egress Layer 3 UPC. Egress RACL is applied, and the packet is queued to the unified crossbar scheduler.
7. Unified crossbar fabric processing: Scheduler determines fairness of access to the fabric and determines when frame is de-queued across the fabric.
8. Egress processing at the UPC.
    a. Egress buffers: These buffers are the landing spots for frames in flight when egress is paused.
    b. Egress forwarding logic: the frame is first parsed, the fields extracted, then perform learning and filtering searches are performed, and finally convert to desired egress format.
9. Packet leaves egress to host B: MAC address is encoded, bytes are synchronized, and the packet is transmitted.

The ingress UPC parses the packet and applies the policy enforcement. The ACL TCAM space for each UPC resides in the policy enforcement module. The processing order of the ACLs types is as follows: first the PACLs check is applied (ingress), and then the VACL check is applied, followed by the QoS ACL. The RACLs are applied either on the ingress Layer 3 UPC or at the egress Layer 3 UPC. The processing order and the ACLs are summarized in Figure 3.

Figure 3. ACL Processing Order



ACL TCAM Overview

The TCAM is a specialized piece of memory designed for rapid table lookups, based on packets passing through the switch, performed by the ACL engine on the UPC. The result of the ACL engine lookup into the TCAM determines how the switch handles a packet. For example, the packet might be permitted or denied.

The TCAM has a limited number of entries that are populated with pattern values and mask values, each with an associated result. These fields are known as value, mask, result entries (VMRs). The term VMR simply refers to the format in which access control entries (ACEs) are represented in the UPC TCAM. A typical VMR (represented as hexadecimal digits) follows:

```
[2043]> K:IP (3/0) IN v4 L-[V-3f/3 ]
    [2043]  SA:ffffff00/01010100 DA:ffffff00/02020200
    [2043]-> prio:7 PERMIT
```

The “value” in VMR refers to the pattern that is to be matched (such as IP addresses, protocol ports, and so on). The “mask” refers to the mask bits associated with the pattern.

A similar concept as the Cisco Catalyst® VMR applies on the Cisco Nexus 5000 and Nexus 5500 Series NX-OS. A difference is that there are as many mask resources as value resources: each source address (SA) and destination address (DA) is stored with its own mask. The “result” refers to the result or action that occurs in the case where a lookup returns a hit for the pattern and mask; it is as simple as PERMIT or DROP, as shown in the example VMR.

Figure 4. Cisco Nexus 5000 ACL Resources Overview (Cisco Nexus 5010 and Nexus 5020 Switches)

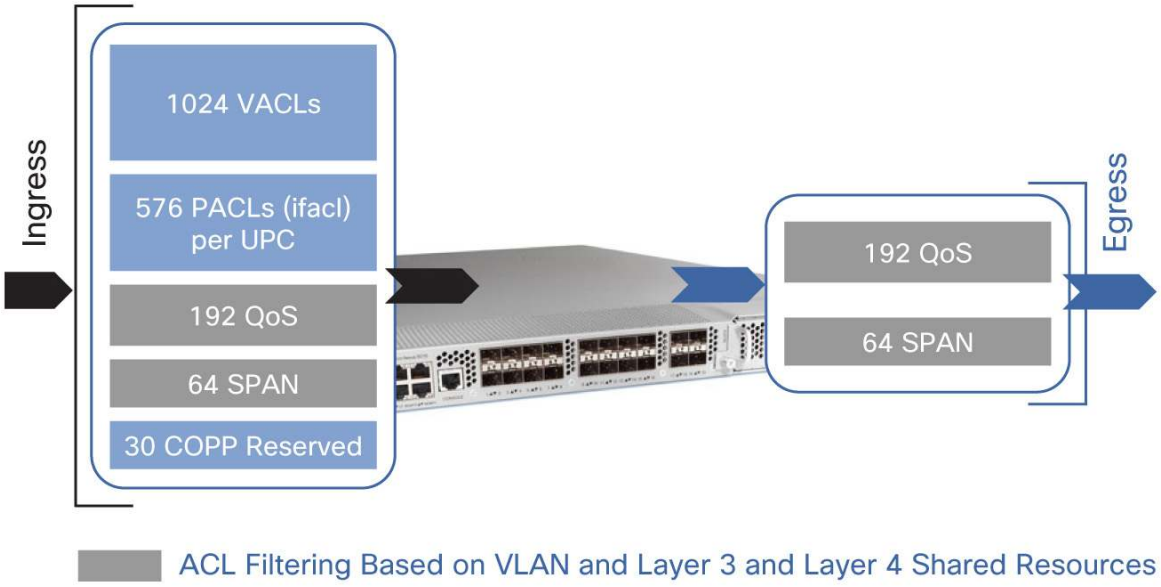
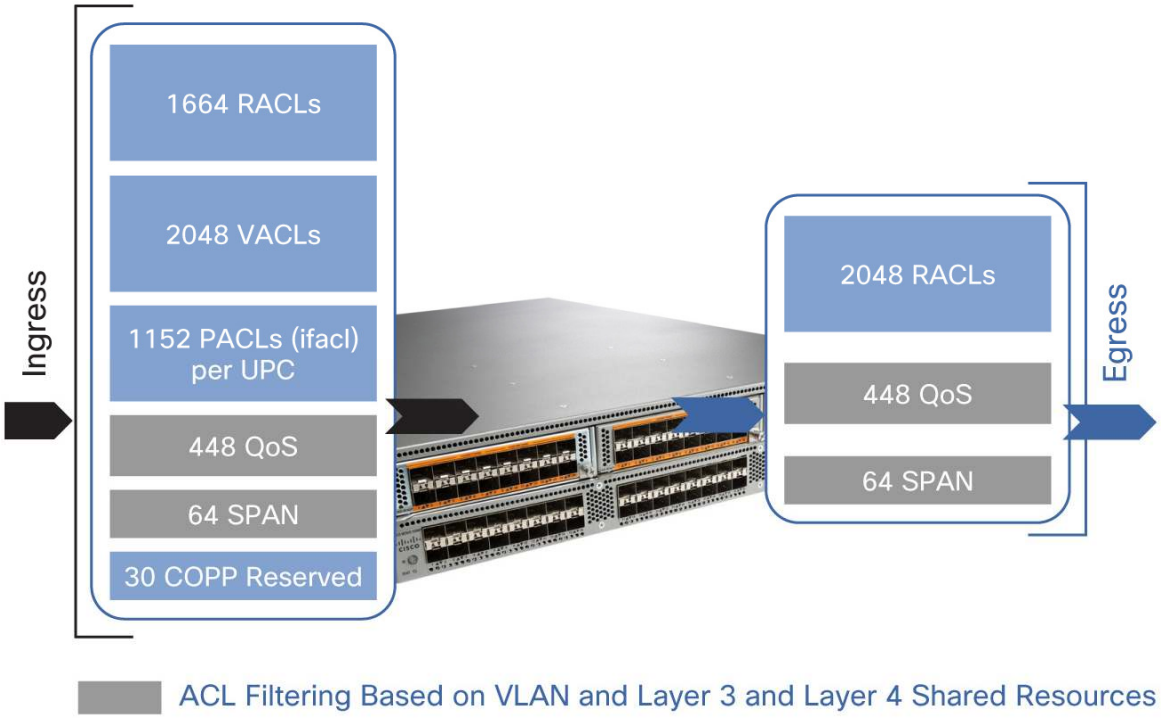


Figure 5. Cisco Nexus 5500 ACL Resources Overview



Monitoring TCAM and Hardware Resource Usage

Layer 2 Resources Monitoring on Cisco Nexus 5000 and Nexus 5500 Series

The Layer 2 resource monitoring is available on both the Cisco Nexus 5000 Series and the Cisco Nexus 5500 Series platforms. When monitoring VACL (VLAN access list), a resource is used on each UPC of the Cisco Nexus 5000 and Nexus 5500 Series Switches. PACL (ifacl), the ACL for a specific interface, is specific to the UPC, and therefore it is important to know what the UPC number is in order to use the appropriate command.

You can use the following command to check the Layer 2 type of ACL TCAM usage: **show platform afm info tcam <UPC#> region [vacl|ifacl|qos]**.

An example follows:

```
5548-1-LA# show platform afm info tcam 0 region vacl
vacl tcam TCAM configuration for asic id 0:
[ vacl tcam]: range     0 - 2047 *
[ifacl tcam]: range  2048 - 3199
[  qos tcam]: range  3200 - 3647
[rbacl tcam]: range  3648 - 3775
[ span tcam]: range  3776 - 3839
[  sup tcam]: range  3840 - 3967

    TCAM [vacl tcam]: [v:1, size:2048, start:0 end:2047]
    In use tcam entries: 11
        0-9,2047
```

```
5548-1-LA# show platform afm info tcam 0 region ifacl
ifacl tcam TCAM configuration for asic id 0:
[ vacl tcam]: range     0 - 2047
[ifacl tcam]: range  2048 - 3199 *
[  qos tcam]: range  3200 - 3647
[rbacl tcam]: range  3648 - 3775
[ span tcam]: range  3776 - 3839
[  sup tcam]: range  3840 - 3967

    TCAM [ifacl tcam]: [v:1, size:1152, start:2048 end:3199]
    In use tcam entries: 24
        2048-2049,2051,2054-2055,2058-2071,3195-3199
```

The steps bellow explain how to find the mapping between the UPC application-specific integrated circuit (ASIC) number and the physical port being configured.

The port-to-UPC number mapping can be retrieved with the command **show hardware internal gatos all-ports** or **show hardware internal carmel all-ports**, for the Cisco Nexus 5000 or Nexus 5500 Series, respectively.

Cisco Nexus 5000 port mapping follows:

On the Cisco Nexus 5000 Series, each UPC ASIC has four 10 Gigabit Ethernet server-facing ports. To look at the mapping of physical ports to UPC ASIC ports, use the command **show hardware internal gatos all-ports**.

An example follows:

```
5020# show hardware internal gatos all-ports
```

Gatos Port Info:

```
name   |log|gat|mac|flag|adm|opr|c:m:s:l|ipt|fab|xgat|xpt|if_index|diag
-------+---+---+---+----+---+---+-------+---+---+----+---+--------+----
xgb1/8 |7  |0  |0  |b7  |en |up |0:0:0:f|0  |55 |0   |2  |1d007000|pass
xgb1/7 |6  |0  |1  |b7  |dis|dn |0:1:1:f|1  |54 |0   |0  |1d006000|pass
xgb1/3 |2  |0  |2  |b7  |en |up |1:2:2:f|2  |56 |0   |4  |1d002000|pass
xgb1/4 |3  |0  |3  |b7  |en |up |1:3:3:f|3  |57 |0   |6  |1d003000|pass
xgb1/16|15 |1  |0  |b7  |en |up |0:0:0:f|0  |50 |1   |2  |1d00f000|pass
xgb1/15|14 |1  |1  |b7  |en |up |0:1:1:f|1  |51 |1   |0  |1d00e000|pass
xgb1/11|10 |1  |2  |b7  |en |up |1:2:2:f|2  |53 |1   |4  |1d00a000|pass
xgb1/12|11 |1  |3  |b7  |en |up |1:3:3:f|3  |52 |1   |6  |1d00b000|pass

<snip>
```

The left column identifies the Ethernet interface identifier xgb1/8 = e1/8. Columns thee and four reflect the UPC (Gatos) ASIC and port that are tied to the physical Ethernet port.

An example follows:

Gatos Port Info:

```
name   |log|gat|mac|flag|adm|opr|c:m:s:l|ipt|fab|xgat|xpt|if_index|diag
-------+---+---+---+----+---+---+-------+---+---+----+---+--------+----
xgb1/8 |7  |0  |0  |b7  |en |up |0:0:0:f|0  |55 |0   |2  |1d007000|pass
```

```
xgb1/8 gat 0 mac 0 would tell you that interface e1/8 maps to gatos 0 port 0 and
so on.
```

Cisco Nexus 5500 port mapping follows:

On the Cisco Nexus 5000 Series, each UPC ASIC has eight 10 Gigabit Ethernet server-facing ports. To look at the mapping of physical ports to UPC ASIC ports, use the command **show hardware internal carmel all-ports.**

An example follows:

```
5548# sh hardware internal carmel all-ports
```

Carmel Port Info:

```
name   |log|car|mac|flag|adm|opr|m:s:l|ipt|fab|xcar|xpt|if_index|diag|ucVer
-------+---+---+---+----+---+---+-----+---+---+----+---+--------+----+-----
xgb1/2 |1  |0  |0 -|b7  |en |up |0:0:f|0  |92 |0   |0  |1a001000|pass| 4.0b
xgb1/1 |0  |0  |1 -|b7  |en |up |1:1:f|1  |88 |0   |0  |1a000000|pass| 4.0b
xgb1/4 |3  |0  |2 -|b7  |en |up |2:2:f|2  |93 |0   |0  |1a003000|pass| 4.0b
xgb1/3 |2  |0  |3 -|b7  |en |up |3:3:f|3  |89 |0   |0  |1a002000|pass| 4.0b
xgb1/6 |5  |0  |4 -|b7  |dis|dn |4:4:f|4  |90 |0   |0  |1a005000|pass| 4.0b
xgb1/5 |4  |0  |5 -|b7  |en |dn |5:5:f|5  |94 |0   |0  |1a004000|pass| 4.0b
xgb1/8 |7  |0  |6 -|b7  |dis|dn |6:6:f|6  |95 |0   |0  |1a007000|pass| 4.0b
xgb1/7 |6  |0  |7 -|b7  |dis|dn |7:7:f|7  |91 |0   |0  |1a006000|pass| 4.0b
xgb1/10|9  |1  |0 -|b7  |en |up |0:0:f|0  |80 |0   |0  |1a009000|pass| 4.0b
xgb1/9 |8  |1  |1 -|b7  |dis|dn |1:1:f|1  |87 |0   |0  |1a008000|pass| 4.0b
xgb1/12|11 |1  |2 -|b7  |en |up |2:2:f|2  |81 |0   |0  |1a00b000|pass| 4.0b
xgb1/11|10 |1  |3 -|b7  |en |up |3:3:f|3  |86 |0   |0  |1a00a000|pass| 4.0b
xgb1/14|13 |1  |4 -|b7  |en |up |4:4:f|4  |82 |0   |0  |1a00d000|pass| 4.0b
xgb1/13|12 |1  |5 -|b7  |en |up |5:5:f|5  |85 |0   |0  |1a00c000|pass| 4.0b
xgb1/16|15 |1  |6 -|b7  |en |up |6:6:f|6  |83 |0   |0  |1a00f000|pass| 4.0b
xgb1/15|14 |1  |7 -|b7  |en |up |7:7:f|7  |84 |0   |0  |1a00e000|pass| 4.0b
xgb1/18|17 |2  |0 -|b7  |en |dn |0:0:f|0  |75 |0   |0  |1a011000|pass| 4.0b
xgb1/17|16 |2  |1 -|b7  |en |up |1:1:f|1  |76 |0   |0  |1a010000|pass| 4.0b
xgb1/20|19 |2  |2 -|b7  |en |up |2:2:f|2  |74 |0   |0  |1a013000|pass| 4.0b
xgb1/19|18 |2  |3 -|b7  |en |up |3:3:f|3  |77 |0   |0  |1a012000|pass| 4.0b
xgb1/22|21 |2  |4 -|b7  |en |up |4:4:f|4  |78 |0   |0  |1a015000|pass| 4.0b
xgb1/21|20 |2  |5 -|b7  |en |up |5:5:f|5  |73 |0   |0  |1a014000|pass| 4.0b
xgb1/24|23 |2  |6 -|b7  |dis|dn |6:6:f|6  |72 |0   |0  |1a017000|pass| 4.0b
xgb1/23|22 |2  |7 -|b7  |dis|dn |7:7:f|7  |79 |0   |0  |1a016000|pass| 4.0b
```

```
xgb1/25|24 |3  |0 -|b7  |dis|dn |0:0:f|0  |71 |0   |0  |1a018000|pass| 4.0b
xgb1/26|25 |3  |1 -|b7  |dis|dn |1:1:f|1  |64 |0   |0  |1a019000|pass| 4.0b
xgb1/27|26 |3  |2 -|b7  |dis|dn |2:2:f|2  |65 |0   |0  |1a01a000|pass| 4.0b
xgb1/28|27 |3  |3 -|b7  |dis|dn |3:3:f|3  |70 |0   |0  |1a01b000|pass| 4.0b
xgb1/29|28 |3  |4 -|b7  |dis|dn |4:4:f|4  |66 |0   |0  |1a01c000|pass| 4.0b
xgb1/30|29 |3  |5 -|b7  |dis|dn |5:5:f|5  |69 |0   |0  |1a01d000|pass| 4.0b
xgb1/31|30 |3  |6 -|b7  |en |up |6:6:f|6  |67 |0   |0  |1a01e000|pass| 4.0b
[snip]
```

Layer 3 Resources Monitoring on Cisco Nexus 5500 Series Switches

The Layer 3 resource monitoring is available for the Cisco Nexus 5500 Series platform equipped with the Layer 3 daughter card or expansion module. The ASIC number to use is 8. This ASIC number represents the Layer 3 UPC as described in the “Packet Forwarding Overview” section. Use the following command:

show platform afm info tcam 8 region [vacl|rbacl]

```
Vacl= egress RACL
Rbacl = ingress RACL
```

Example 1: How to check the size of the egress RACL TCAM:

```
5548UP# sh platform afm info tcam 9 region vacl
vacl tcam TCAM configuration for asic id 9:
[ vacl tcam]: range     0 - 2047 *
[ifacl tcam]: range  2048 - 2111
[  qos tcam]: range  2112 - 2175
[rbacl tcam]: range  2176 - 3839
[ span tcam]: range  3840 - 3903
[  sup tcam]: range  3904 - 3967

    TCAM [vacl tcam]: [v:1, size:2048, start:0 end:2047]
    In use tcam entries: 11
        0-9,2047
```

As shown, the TCAM size is exactly 2048, and in this example 11 entries are used.

Example 2: How to check the size of the ingress RACL TCAM:

```
5548-1-LA# sh platform afm info tcam 8 region rbacl
rbacl tcam TCAM configuration for asic id 8:
[ vacl tcam]: range     0 - 2047
[ifacl tcam]: range  2048 - 2111
[  qos tcam]: range  2112 - 2175
[rbacl tcam]: range  2176 - 3839 *
[ span tcam]: range  3840 - 3903
[  sup tcam]: range  3904 - 3967

 TCAM [rbacl tcam]: [v:1, size:1664, start:2176 end:3839]
    In use tcam entries: 5
        2176-2177,3837-3839
```

As shown, the TCAM size is 1664, and in this example five entries are used, with indexes of 2176, 2177, 3837, 3838, and 3839. The indexes can be used to read the TCAM entries with the command:

show platform afm info tcam 8 car-entry <tcam_first_index tcam_last_index>

```
5548UP#show platform afm info tcam 8 car-entry 2176 2829
TCAM entries in the range of 2176 and 2829 for asic id 8:
    K-keyType, L-label, B-bindcheck, DH-L2DA, CT-cdceTrnst
    L(IF-ifacl V-vacl Q-qos R-rbacl)

    [2176]> K:ALL (0/0) IN-OUT L-[R-fff/1]
    [2176]-> prio:0 DROP

    [2177]> K:ALL (0/0) IN L-[IF-ffff/3 ]
    [2177]-> prio:7 PERMIT
```

Note: Because of the resource sharing for the Layer 3 UPC, the egress RACL is using the VACL TCAM resource and the ingress RACL is using the role-based ACL (RBACL) resource.

Logical Operation Units and Layer 4 Operations

LOUs are hardware registers used to store {operator, operand} tuples for TCP and User Datagram Protocol (UDP) port numbers specified in an IP extended ACL, VACL, or QoS ACL. As stated previously, these tuples are called Layer 4 Operations, or L4Ops.

The operator portion of an L4Op is one of the **lt, gt, neq**, and **range** operators. The operand is the source or destination TCP or UDP port number. Therefore, source or destination matched traffic uses a different operator.

L4Ops consume LOU registers. For example, the L4Op {gt, 1023} consumes one register (one-half) of an LOU. The same rule applies for the **lt** and **neq** operators. An L4Op using the **range** operator requires two LOU registers (the entire LOU). Tuples using the **eq** operator (for example, {eq, 5000}) do not consume LOU registers.

Note: In addition to the **lt, gt, neq**, and **range** operators, an ACE matching on a port, UDP or TCP the **eq** operator does not consumes an L4Op for the ACL. Indeed, these ACEs do not consume LOU registers. Instead, the TCAM space is utilized. In the Cisco Nexus 5000 and Cisco Nexus 5500 series, this LOU is called L3 Operations (L3Ops). However, because most ACLs do not match on only on **eq**, exhausting this hardware resource is uncommon. This will be leveraged in this documentation to optimize and resolve resource exhaustion that may occur. The important thing to note is, only one of the L4Ops for the ACL is consumed if one or more ACEs matches on the TCP “established” flag, using the concept of label sharing.

An example follows:

Consider the following ACL:

```
access-list 101 permit tcp host 10.1.1.1 host 10.2.2.2 gt 1023
access-list 101 permit tcp host 10.3.3.3 lt 1023 host 10.4.4.4
access-list 101 permit tcp host 10.5.5.5 host 10.6.6.6 gt 5000
access-list 101 permit tcp host 10.7.7.7 host 10.8.8.8 neq 2000
access-list 101 permit tcp host 10.9.9.9 lt 1023 host 10.10.10.10 gt 1023
```

This ACL has four different L4Ops: gt 1023 lt 1023 gt 5000 neq 2000.

L4Ops are different if the operation (for example, **gt, lt**, etc.) is different, or the operand (the TCP or UDP port number) is different, or both. Therefore, {gt, 1000} and {gt 1001} are two different L4Ops.

An example follows:

```
access-list 101 permit tcp host 10.3.3.3 lt 1023 host 10.4.4.4
access-list 101 permit tcp host 10.5.5.5 host 10.6.6.6 lt 1023
```

If the same {operator, operand} tuple is applied to a source TCP or UDP port and then later to a destination TCP or UDP port, it counts as a different L4Op. In this case, the two {lt, 1023} tuples count as two different L4Ops because one applies to an IP source address and one applies to an IP destination address.

Now consider the following ACL:

```
access-list 101 permit tcp host 10.1.1.1 host 10.2.2.2 gt 1023
access-list 101 deny tcp host 10.3.3.3 gt 1023 host 10.4.4.4
access-list 101 deny tcp host 10.5.5.5 neq 4000 host 10.6.6.6
access-list 101 permit tcp host 10.7.7.7 host 10.8.8.8 range 5000 6000
```

This ACL has four L4Ops.

Monitoring Layer 4 Operations Usage

When monitoring Layer 4 Operations usage (LOU), it is important to know what UPC number it relates to; it depends on the port where the RACL, for example, is applied. The mapping can be retrieved as explained in the section “Monitoring TCAM and Hardware Resource Usage”.

Use the command **show platform afm info lop [all | asic #]**

An example follows:

```
5548-1-LA# show platform afm info lop 0
Logical operators configuration for asic id 0:
    L2 operators:
         0- HDR_PRST [v:1 ref_cnt:1] op: OR {macsec }
         1- HDR_PRST [v:1 ref_cnt:1] op: OR {dot1q dot1p }
        16- DA_CMP [v:1 ref_cnt:1] addr: 00:05:73:d8:60:c8, exclude_len: 5
        17- DA_CMP [v:1 ref_cnt:1] addr: 00:00:00:00:00:00, exclude_len: 0
        18- DA_CMP [v:1 ref_cnt:1] addr: 00:00:00:00:00:00, exclude_len: 0
        19- DA_CMP [v:1 ref_cnt:1] addr: 00:00:00:00:00:00, exclude_len: 0
        20- DA_CMP [v:1 ref_cnt:1] addr: 00:05:73:d8:60:c4, exclude_len: 2
    L2op selectors:
        0- <ref_cnt:1>
            [i: 0 r:   1] [i: 1 r:   1] [i: 0 r:   1] [i: 1 r:   1]
            [i: 2 r:   1] [           ] [           ] [           ]
        1- <ref_cnt:1>
            [i: 0 r:   1] [i: 1 r:   1] [i: 0 r:   1] [i: 1 r:   1]
            [i: 2 r:   1] [           ] [           ] [           ]
        2- <ref_cnt:1>
            [i: 0 r:   1] [i: 1 r:   1] [i: 0 r:   1] [i: 1 r:   1]
            [i: 2 r:   1] [           ] [           ] [           ]
        3- <ref_cnt:1>
```

```
        [i: 0 r:   1] [i: 1 r:   1] [i: 0 r:   1] [i: 1 r:   1]
        [i: 2 r:   1] [           ] [           ] [           ]

v4 L3op selector:
        [cond:9 r:   1] [           ] [           ] [           ]
        [           ] [           ] [           ] [           ]

v6 L3op selector:
        [           ] [           ] [           ] [           ]
        [           ] [           ] [           ] [           ]

UDP l4op selector:
        [i:16 r:   1] [i:17 r:   1] [i:18 r:   1] [i:19 r:   1]
        [i:20 r:   1] [i:21 r:   1] [i:22 r:   1] [i:23 r:   1]
UDP operators:
    16- SRC-PORT [v:1 ref_cnt:1] lo: 49072, hi: 50175 op: RANGE
    17- DST-PORT [v:1 ref_cnt:1] lo: 49072, hi: 50175 op: RANGE
    18- SRC-PORT [v:1 ref_cnt:1] lo: 1023, hi: 0 op: GT
    19- DST-PORT [v:1 ref_cnt:1] lo: 39500, hi: 39551 op: RANGE
    20- DST-PORT [v:1 ref_cnt:1] lo: 39552, hi: 39999 op: RANGE
    21- SRC-PORT [v:1 ref_cnt:1] lo: 39500, hi: 39551 op: RANGE
    22- DST-PORT [v:1 ref_cnt:1] lo: 1023, hi: 0 op: GT
    23- SRC-PORT [v:1 ref_cnt:1] lo: 39552, hi: 39999 op: RANGE

TCP l4op selector:
        [i: 8 r:   1] [i: 9 r:   1] [           ] [           ]
        [           ] [           ] [           ] [           ]
TCP PORT operators:
     8- SRC-PORT [v:1 ref_cnt:1] lo: 1023, hi: 0 op: GT
     9- DST-PORT [v:1 ref_cnt:1] lo: 1023, hi: 0 op: GT
TCP FLAG operators:
```

In this example eight UDP resources and two TCP resources are used. If you use more ACLs for UDP with a different range of ports, then this configuration needs to be optimized. The next section will cover how to optimize the configuration in this case.

Note: When using the same port range in a different ACL statement, the label resource sharing allows using the same Layer 4 Operation resource.

Cisco Nexus 5000 Best Practices and Optimization

- Before implementing a new ACL and ACE, it is good practice to check the TCAM size and the number of operators available.
- Regroup all the ranges as much as possible in fewer range statements; create larger range statements whenever possible.
- For a smaller number of ports, for example fewer than 50 ports, define the ACE with the eq statement instead of the range - it creates more ACE lines in the access-list definition, but does not use a Layer 4

operator and it permits you to preserve the range resource for a larger subset of ports. The value 50 is an example. Depending on the environment, this value may be lower or higher. The limitation of the size of eq statement ACEs is the size of the TCAM for the type of ACL used (VACL, RACL etc.).

Optimization examples follow:

Example 1: Reduce the number of Layer 4 operations resources while conserving the ACE.

In this example ACLs and ACEs are already configured. The example shows how to reduce the number of Layer 4 operators used while conserving the same ACL rules.

1. Check the number of Layer 4 operators:

```
5548-1-LA# show platform afm info lop 0
[snip]
TCP PORT operators:
8- DST-PORT [v:1 ref_cnt:1] lo: 17101, hi: 17150 op: RANGE
9- DST-PORT [v:1 ref_cnt:1] lo: 9888, hi: 9889 op: RANGE
10- DST-PORT [v:1 ref_cnt:4] lo: 19001, hi: 19010 op: RANGE
11- DST-PORT [v:1 ref_cnt:1] lo: 18023, hi: 18025 op: RANGE
12- DST-PORT [v:1 ref_cnt:1] lo: 1500, hi: 1510 op: RANGE
13- DST-PORT [v:1 ref_cnt:1] lo: 19001, hi: 19004 op: RANGE
14- DST-PORT [v:1 ref_cnt:1] lo: 19001, hi: 19030 op: RANGE
[snip]
```

Currently eight Layer 4 TCP Ops are used, the maximum limit possible on the platform.

When looking at the RANGE statements used, line number 8 is using 150 ports in a range statement and line number 14 is using 30 ports. The other lines, in bold, are using a RANGE statement between 2 and 11 ports. All the statements can be optimized. In this example lines 8 and 14 will not be optimized; in other scenarios when larger RANGE statements are needed, they may be optimized in the same manner.

2. Find the ACL with the range statement.

Checking the running configuration or using the **show access-lists** command provides the following:

```
[snip]
ip access-list racl_1_in
150 permit tcp 10.10.1.135/32  172.10.73.15/32 range 19001 19010
160 permit tcp 10.10.1.135/32 172.10.73.15/32 range 9888 9889

ip access-list racl_2_in
150 permit tcp 10.10.1.135/32 204.1.50.15/32 range 18023 18025
160 permit tcp 10.10.1.135/32 204.1.50.15/32 range 19001 19004
170 permit tcp 10.10.1.135/32 204.1.50.15/32 range 1500 1510
[snip]
```

Note: racl_1_in and racl_2_in are the only ACLs using the previous port RANGE values. If other ACLs are using the same RANGE statements, they need to be found and modified, because a Layer 4 Ops is a shared resource among the different ACLs using the same RANGE values using the LABEL sharing method.

3. Change the ACL to reduce the Layer 4 Ops resources.

```
ip access-list racl_1_in
360 permit tcp 10.10.1.135/32 172.10.73.15/32 eq 19001
370 permit tcp 10.10.1.135/32 172.10.73.15/32 eq 19002
... [etc]
460 permit tcp 10.10.1.135/32 172.10.73.15/32 eq 19010
470 permit tcp 10.10.1.135/32 172.10.73.15/32 eq 9888
480 permit tcp 10.10.1.135/32 172.10.73.15/32 eq 9889

ip access-list racl_2_in
150 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 18023
151 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 18024
152 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 18025
160 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 19001
161 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 19002
162 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 19003
163 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 19004
170 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 1500
171 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 1501
172 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 1502
...
180 permit tcp 10.10.1.135/32 204.1.50.15/32 eq 1510
```

Note: The example does not show how to edit an access list. The instructions are available in the configuration guide.

4. Check the Layer 4 Ops resources.

```
5548-1-LA# show platform afm info lop 0
[snip]
TCP PORT operators:
8- DST-PORT [v:1 ref_cnt:1] lo: 17101, hi: 17150 op: RANGE
9- DST-PORT [v:1 ref_cnt:1] lo: 19001, hi: 19030 op: RANGE
[snip]
```

The ACL modification has reduced the Layer 4 Ops resources from 8 to 2. There is now margin for other ACEs or ACLs with larger RANGE statements. The ACL has been optimized using the best practices.

Example 2: Error while applying a RACL to an interface

Methodology to follow: When you receive an error message, check the ACL configuration in the running configuration and the available resources; then optimize the ACL, reapply, and finally check the Layer 4 resources.

1. Error while applying ACL

```
5548-1-LA(config-if)# ip access-group example in
```

ERROR: L4 logical operator allocation error

2. Check ACL configuration and TCAM and L4Ops available resources.

```
5548-1-LA# sh ip access-lists example
```

IP access list example

```
        10 permit tcp 1.1.1.0/24 11.11.11.0/24 range 32001 32010
        20 permit tcp 2.2.2.0/24 22.22.22.0/24 range 32200 32250
        30 permit tcp 3.3.3.0/24 33.33.33.0/24 range 32300 32350
        40 permit tcp 4.4.4.0/24 44.44.44.0/24 range 32400 32450
        50 permit tcp 5.5.5.0/24 55.55.55.0/24 range 32500 32550
        60 permit tcp 6.6.6.0/24 66.66.66.0/24 range 32600 32650
        70 permit tcp 7.7.7.0/24 77.77.77.0/24 range 32700 32750
        80 permit tcp 8.8.8.0/24 88.88.88.0/24 range 32800 32850
        90 permit tcp 9.9.9.0/24 99.99.99.0/24 range 32900 32950
```

```
5548UP-1-LA# show platform afm info lop asic 0
[snip]
    v4 L3op selector:
            [          ] [          ] [          ] [          ]
            [          ] [          ] [          ] [          ]

    v6 L3op selector:
            [          ] [          ] [          ] [          ]
            [          ] [          ] [          ] [          ]

    UDP l4op selector:
            [          ] [          ] [          ] [          ]
            [          ] [          ] [          ] [          ]
    UDP operators:


    TCP l4op selector:
            [          ] [          ] [          ] [          ]
            [          ] [          ] [          ] [          ]
    TCP PORT operators:
    TCP FLAG operators:
[snip]
```

In this case, no resources are yet used from the L4Ops. Therefore, the limit is eight L4Ops for the ACL to be accepted. The ACL is using nine L4Ops and therefore is rejected. It needs to be optimized.

3. Optimize ACL.

IP access list example

```
10 permit tcp 1.1.1.0/24 11.11.11.0/24 eq 32001
11 permit tcp 1.1.1.0/24 11.11.11.0/24 eq 32002
12 permit tcp 1.1.1.0/24 11.11.11.0/24 eq 32003
13 permit tcp 1.1.1.0/24 11.11.11.0/24 eq 32004
14 permit tcp 1.1.1.0/24 11.11.11.0/24 eq 32005
15 permit tcp 1.1.1.0/24 11.11.11.0/24 eq 32006
16 permit tcp 1.1.1.0/24 11.11.11.0/24 eq 32007
17 permit tcp 1.1.1.0/24 11.11.11.0/24 eq 32008
18 permit tcp 1.1.1.0/24 11.11.11.0/24 eq 32009
19 permit tcp 1.1.1.0/24 11.11.11.0/24 eq 32010
20 permit tcp 2.2.2.0/24 22.22.22.0/24 range 32200 32250
30 permit tcp 3.3.3.0/24 33.33.33.0/24 range 32300 32350
40 permit tcp 4.4.4.0/24 44.44.44.0/24 range 32400 32450
50 permit tcp 5.5.5.0/24 55.55.55.0/24 range 32500 32550
60 permit tcp 6.6.6.0/24 66.66.66.0/24 range 32600 32650
70 permit tcp 7.7.7.0/24 77.77.77.0/24 range 32700 32750
80 permit tcp 8.8.8.0/24 88.88.88.0/24 range 32800 32850
90 permit tcp 9.9.9.0/24 99.99.99.0/24 range 32900 32950
```

4. Reapply ACL.

```
5548-1-LA (config-if)# ip access-group example in
```

5. Check Layer 4 Ops resources.

```
5548-1-LA# show platform afm info lop asic 0
[snip]
    TCP l4op selector:
            [i: 8 r:   1] [i: 9 r:   1] [i:10 r:   1] [i:11 r:   1]
            [i:12 r:   1] [i:13 r:   1] [i:14 r:   1] [i:15 r:   1]
    TCP PORT operators:
         8- DST-PORT [v:1 ref_cnt:1] lo: 32200, hi: 32250 op: RANGE
         9 DST-PORT [v:1 ref_cnt:1] lo: 32300, hi: 32350 op: RANGE
        10- DST-PORT [v:1 ref_cnt:1] lo: 32400, hi: 32450 op: RANGE
        11- DST-PORT [v:1 ref_cnt:1] lo: 32500, hi: 32550 op: RANGE
        12- DST-PORT [v:1 ref_cnt:1] lo: 32600, hi: 32650 op: RANGE
        13- DST-PORT [v:1 ref_cnt:1] lo: 32700, hi: 32750 op: RANGE
        14- DST-PORT [v:1 ref_cnt:1] lo: 32800, hi: 32850 op: RANGE
        15- DST-PORT [v:1 ref_cnt:1] lo: 32800, hi: 32850 op: RANGE
[snip]
```

The access list was applied successfully and was optimized to use eight resources. Ten resources from the TCAM were used to reduce one L4Op. Further optimization is possible; in this example each line optimized can save one L4Op and use 51 TCAM entries (the ranges have 51 ports).

Conclusion

The ACL TCAM resources and Layer 4 Ops differ from platform to platform - and even between Cisco Nexus 5000 and Nexus 5500 Series Switches. This document explained the ACL TCAM resources available per type of ACL and the Layer 4 Ops capacity. When implementing ACL at scale on the platform, it is important to understand how to monitor the resources available: the TCAM usage as well as the Layer 4 Operations usage. When new ACLs and ACEs need to be implemented, the best practices explain how to optimize the Layer 4 Operators to scale without encountering errors, or what actions to take when there are not enough remaining free hardware resources.

For More Information

http://www.cisco.com/go/nexus5000

Glossary

ACE - Access Control Entry (one line of an Access Control List)

ACL - Access Control List

COPP - Control Plane Policing

L3Op - Layer 3 Operation (the {operator, operand} tuples used to perform one operation - the L3Ops are eq)

L4Op - Layer 4 Operation (the {operator, operand} tuples used to perform Layer 4 protocol operations - the L4Ops are lt, gt, neq, and range)

LOU - Logical Operation Unit (hardware register used to store L4Op information)

QoS ACL - QoS Access Control List

RACL - Router Access Control List (an access-list applied to an interface in IOS, using the ip access-group {in|out} command)

TCAM - Ternary Content Addressable Memory (specialized piece of memory for storing complex tabular data and supporting very rapid parallel lookups)

UPC: Unified Port Controller, this is the port hardware asic.

uRPF check - Unicast Reverse Path Forwarding check (method of verifying the reachability of source IP addresses in incoming packets)

VACL - VLAN Access Control List (a security ACL mapped to a VLAN on a Cisco Catalyst OS switch, using the set security acl commands)

VMR - Value, Mask, Result (format, used to represent entries in the TCAM, that consists of a pattern value, the associated mask value, and a result for lookups returning a hit for the entry)

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C11-682225-00 08/11