FIB Scaling on Merchant Silicon

Vince Schuele

ipinfusion™

ufiSpace

www.epsglobal.com

AGENDA

- Current State of the IRT
- What is RIB/FIB
- Broadcom DNX
- Scaling characteristics

Current Internet Routing Table status

- Various route servers show approximately 970k IPv4 and 205k IPv6
- First questions in BGP applications are always around scale

Table History

Date	Prefixes	CIDR Aggregated
07-03-25	1003061	564880
08-03-25	1002475	565493
09-03-25	1002624	565619
10-03-25	1002500	565748
11-03-25	1002712	566629
12-03-25	1003357	566623
13-03-25	1002960	566828
14-03-25	1003335	566921

Plot: BGP Table Size

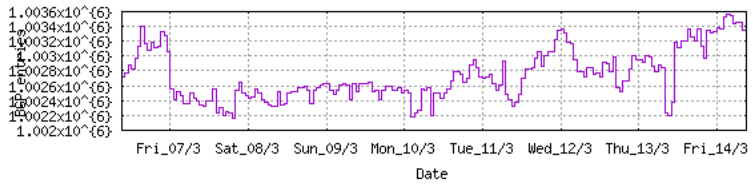
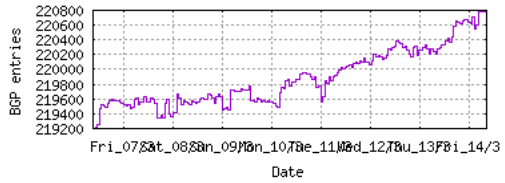


Table History

Date	Prefixes	CIDR Aggregated
07-03-25	219524	112353
08-03-25	219409	112438
09-03-25	219443	112446
10-03-25	219551	112520
11-03-25	219556	112909
12-03-25	220055	113072
13-03-25	220254	113308
14-03-25	220639	113308

Plot: BGP Table Size

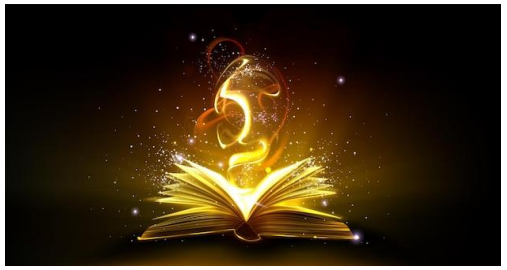


https://www.cidr-report.org/as2.0/

FIB vs RIB what's the difference

Forwarding Information Base

- Hardware tables, TCAM space, and encapsulations
- Mostly magic



Routing Information Base

- Various parts:
  - Adj-RIB-In
  - Adj-RIB-Out
  - Local RIB
- Scales with respect to memory space

BORING!

Broadcom Silicon – DNX Family

Q2a

800 Gbps

Q2c

2.4 Tbps

J2c+

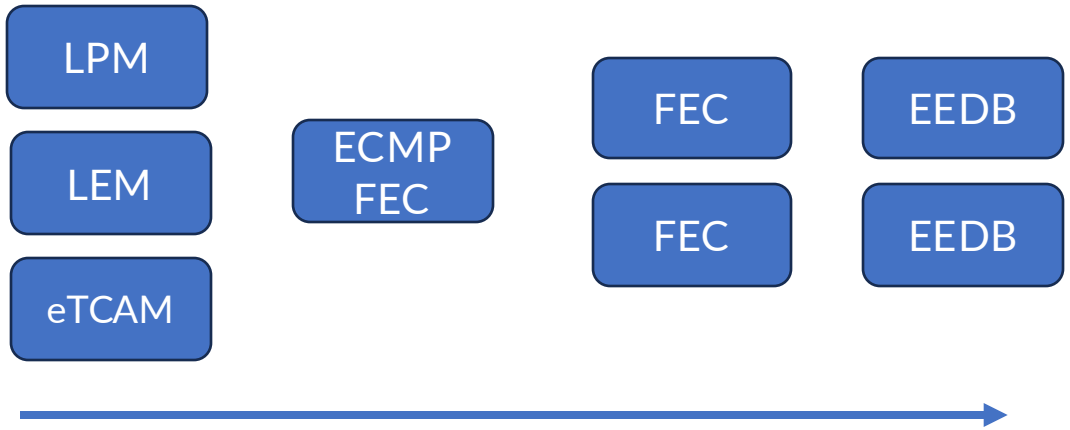
14.4 Tbps

What tables are we working with?

Acronym	Meaning	Purpose
LPM	Longest Prefix Match	Prefixes shorter than /32 /128
LEM	Largest Exact Match	/32s /128s MPLS labels MACs
eTCAM	External TCAM	Prefix shorter than /32 /128
ECMP FEC	Equal cost multipath Forward Equivalence Class	Next hop info
FEC	Forward Equivalence Class	Next hop info
EEDB	Egress Encapsulation Database	Egress encapsulation

How do they work together

IP Routing



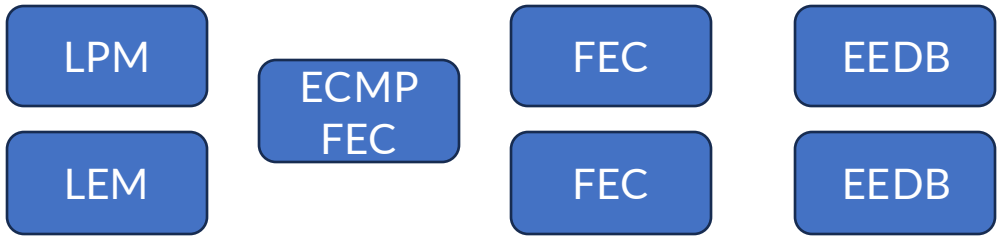
1x ECMP FEC prefix set

FEC per next-hop

Egress object per next hop

LPM remains constant

MPLS Forwarding



Outlabel is unique so the DB entry is unique

Extra utilization based on the number of paths

Scalability Q2A

S9510-28DC
24 x 1/10/25G SFP28, 2 x 40/100G QSFP28, 2 x 100/400G QSDP-DD supporting 400G ZR+ (2)

```
R2#show forwarding profile limit
(s): public and private tables are shared

Active in HW: balanced (cfg1)
-------------------------------------------------------------------------
|                     Forwarding Profile Table Size                     |
-------------------------------------------------------------------------
|cfg  |Entry    |Balanced  |Balanced  |l2-xl     |l3-xl     |Balanced  |
|     |         |          |p-and-p   |          |          |Extended  |
-------------------------------------------------------------------------
|     MAC       355.0k     372.0k     590.0k     77.5k      253.0k     |
|                                                                       |
|     SRv6-SID  30.0k      30.0k      30.0k      37.0k      30.0k      |
|                                                                       |
|     LPM                                                               |
|1(s) Public    1048.0k    917.0k     131.0k     1441.0k    917.0k     |
|     Private   1048.0k    917.0k     131.0k     1441.0k    917.0k     |
-------------------------------------------------------------------------
```

Current Internet IRT:

IPv4: 980K
IPv6: 225k

LPM usage:

shared space between IPv4 and IPv6

IPv4 ~ 1 entry
IPv6 ~ 2 entries

10-12% loss of space due to hashing collisions

IPv6 is assumed to be shorter than /64

IPv6 with prefix length longer than /64 will use more than 2 LPM Entries

Result:

Q2A and 7946-30XB can **not** hold a full IPv4 and IPv6 IRT simultaneously

Filtering is required. Filtering out /24s removes ~600k IPv4 routes

Scalability Q2c and OP2

```
Software version: UFI_S9600-32X-OcNOS-SP-MPLS-6.3.2-MR 08/29/2023 03:03:29
 Copyright (C) 2023 IP Infusion. All rights reserved

 Software Product: OcNOS-SP, Version: 6.3.2
 Build Number: 47
 Release: MR
 Hardware Model: Ufi Space S9600-32X-B
 Software Feature Code: MPLS
 Software Baseline Version: 6.0.117

Installation Information:
 Image Filename: OcNOS-SP-MPLS-Q2-6.3.2-47-MR-installer
 ONIE-SysInfo: x86_64-ufispace_s9600_32x-r0

SRH-Qumran#show forwarding profile limit
(s): public and private tables are shared

Active in HW: balanced (cfg1)
-------------------------------------------------------------------------
|                    Forwarding Profile Table Size                      |
-------------------------------------------------------------------------
|cfg  |Entry    |Balanced  |Balanced  |l2-xl     |l3-xl     |Balanced  |
|     |         |          |p-and-p   |          |          |Extended  |
-------------------------------------------------------------------------
|     MAC       710.0k     744.0k     1180.0k    155.0k     506.0k     |
|                                                                       |
|     LPM                                                               |
|1(s) Public    2097.0k    1832.0k    131.0k     2883.0k    1832.0k    |
|     Private   2097.0k    1832.0k    131.0k     2883.0k    1832.0k    |
-------------------------------------------------------------------------
```

```
ok-core-0#show forwarding profile limit
(s): public and private tables are shared

Active in HW: balanced (cfg1)
-------------------------------------------------------------------------
|                    Forwarding Profile Table Size                      |
-------------------------------------------------------------------------
|cfg  |Entry    |Balanced  |Balanced  |l2-xl     |l3-xl     |Balanced  |
|     |         |          |p-and-p   |          |          |Extended  |
-------------------------------------------------------------------------
|     MAC       710.0k     744.0k     1180.0k    155.0k     506.0k     |
|                                                                       |
|     LPM                                                               |
|1(s) Public    20480k     20480k     20480k     20480k     20480k     |
|     Private   20480k     20480k     20480k     20480k     20480k     |
-------------------------------------------------------------------------
ok-core-0#
ok-core-0#sh ver
Software version: UFI_S9600-72XC-OcNOS-SP-MPLS-6.5.3-MR 02/28/2025 15:56:16
 Copyright (C) 2025 IP Infusion. All rights reserved
```

Q2c can support 2.8m LPM with l3-xl enabled

OP2 can support 20m LPM by default

Forwarding profile can be changed with a configuration and reload

```
R3PLAB(config)#forwarding mdb-profile ?
  balanced           select mdb profile balanced
  balanced-extended  select mdb profile balanced-extended
  balanced-p-and-p   select mdb profile balanced-p-and-p
  l2-xl              select mdb profile l2-xl
  l3-xl              select mdb profile l3-xl
```

- UfiSpace S9600-32X
- UfiSpace S9600-28DX

→ **Q2c**

- UfiSpace S9600-72XC

→ **Q2c w/eTCAM**

- UfiSpace S9600-56DX
- UfiSpace S9600-64X

→ **2xQ2c**

FIB scale tested to 8M IPv4 and 4M IPv6 with OP2 (20M LPM) and will continue to scale with IRT growth

ECMP only routes only count as one LPM entry with multiple next hops meaning FIB install of current IRT is ~1.6M entries

Scalability J2c+

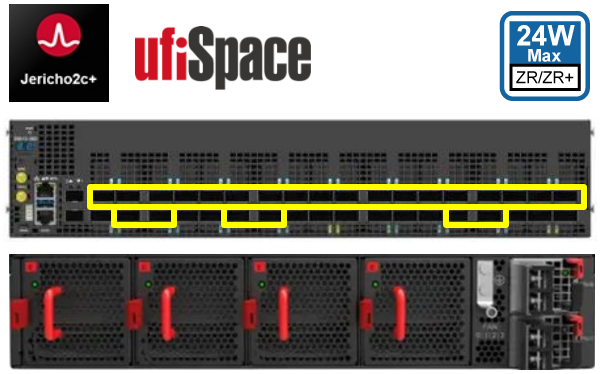
OcNOS 6.5

```
7011#show forwarding profile limit
(s): public and private tables are shared

Active in HW: balanced (cfg1)
-------------------------------------------------------------------------
|                     Forwarding Profile Table Size                     |
-------------------------------------------------------------------------
|cfg  |Entry    |Balanced  |Balanced  |l2-xl     |l3-xl     |Balanced  |
|     |         |          |p-and-p   |          |          |Extended  |
-------------------------------------------------------------------------
|      MAC       713.0k     744.0k     1150.0k    155.0k     500.0k    |
|                                                                       |
|      SRv6-SID  60.0k      60.0k      60.0k      75.0k      60.0k     |
|                                                                       |
|      LPM                                                              |
|1(s)  Public    2097.0k    1572.0k    131.0k     3145.0k    1835.0k   |
|      Private   2097.0k    1572.0k    131.0k     3145.0k    1835.0k   |
-------------------------------------------------------------------------
```

```
OcNOS#show forwarding profile limit
(s): public and private tables are shared

Active in HW: balanced (cfg1)
-------------------------------------------------------------------------
|                     Forwarding Profile Table Size                     |
-------------------------------------------------------------------------
|cfg  |Entry    |Balanced  |Balanced  |l2-xl     |l3-xl     |Balanced  |
|     |         |          |p-and-p   |          |          |Extended  |
-------------------------------------------------------------------------
|      MAC       713.0k     744.0k     1150.0k    155.0k     500.0k    |
|                                                                       |
|      SRv6-SID  60.0k      60.0k      60.0k      75.0k      60.0k     |
|                                                                       |
|      LPM                                                              |
|1(s)  Public    20480k     20480k     20480k     20480k     20480k    |
|      Private   20480k     20480k     20480k     20480k     20480k    |
-------------------------------------------------------------------------
```

Forwarding profile can be changed with a configuration and reload

Changing to utilize TCAM reduces port usage by 4 ports



S9610-36D
36 x 40/100/200/400G QSFP-DD
supporting 400G ZR+ (24)

FIB scale tested to 8M IPv4 and 4M IPv6 with OP2 (20M LPM) and will continue to scale with IRT growth

ECMP only routes only count as one LPM entry with multiple next hops meaning FIB install of current IRT is ~1.6M entries

Simulated full table – Q2a

```
PE4#show ip bgp sum
BGP router identifier 172.16.10.250, local AS number 65000
BGP table version is 20
1 BGP AS-PATH entries
0 BGP community entries

Neighbor          V    AS     MsgRcv     MsgSen  TblVer    InQ   OutQ   Up/Down  State/PfxRcd   Desc
192.85.1.3        4 65100       2469         22      20      0      0  00:03:54        970000

Total number of neighbors 1

Total number of Established sessions 1
PE4#show ipv6 bgp sum
BGP router identifier 172.16.10.250, local AS number 65000
BGP table version is 17
1 BGP AS-PATH entries
0 BGP community entries

Neighbor          V    AS     MsgRcv     MsgSen  TblVer    InQ   OutQ   Up/Down  State/PfxRcd   Desc
2001::2           4 65100       3300         27      17      0      0  00:04:14        250000

Total number of neighbors 1

Total number of Established sessions 1
```

TCAM is full

```
PE4#show ip route summary

-------------------------------------------------
IP routing table name is Default-IP-Routing-Table(0)
-------------------------------------------------
IP routing table maximum-paths   : 8
Total number of IPv4 routes      : 868439
Total number of IPv4 paths       : 868439
Pending routes (due to route max reached): 101570
Route Source    Networks
connected       6
bgp             868427
isis            6
Total           868439
FIB             868439

ECMP statistics (active in ASIC):
 Total number of IPv4 ECMP routes : 0
 Total number of IPv4 ECMP paths  : 0
PE4#show ipv6 route summary

-------------------------------------------------
IPv6 routing table name is Default-IPv6-Routing-Table(0)
-------------------------------------------------
IPv6 routing table maximum-paths : 8
Total number of IPv6 routes      : 179237
Total number of IPv6 paths       : 179237
Pending routes (due to route max reached): 70766
Route Source    Networks
connected       3
bgp             179234
Total           179237
FIB             179237

ECMP statistics (active in ASIC):
 Total number of IPv6 ECMP routes   : 0
 Total number of IPv6 ECMP paths    : 0
```

What happens when the hardware is full?

Challenges:

- Routes don’t get installed
- These routes won’t be reachable from your device

Solutions:

- Selective route download
- Filtering
  - filtering /24 removes ~600k routes
  - still must reach those /24
- Static default
- Do routers with one egress really need the full table?

Simulated full table – multi-path – Q2a

```
PE4#show ip bgp sum
BGP router identifier 172.16.10.250, local AS number 65000
BGP table version is 21
2 BGP AS-PATH entries
0 BGP community entries

Neighbor          V    AS     MsgRcv     MsgSen  TblVer    InQ   OutQ   Up/Down  State/PfxRcd   Desc
192.85.1.3        4 65100       2488         44      21      0      0  00:13:31        970000
192.86.1.3        4 65101       2555        963      21      0      0  00:01:31        960000

Total number of neighbors 2

Total number of Established sessions 2
PE4#show ipv6 bgp summary
BGP router identifier 172.16.10.250, local AS number 65000
BGP table version is 18
2 BGP AS-PATH entries
0 BGP community entries

Neighbor          V    AS     MsgRcv     MsgSen  TblVer    InQ   OutQ   Up/Down  State/PfxRcd   Desc
2001::2           4 65100       3319         50      18      0      0  00:13:50        250000
2001:0:0:1::2     4 65101       2555        587      18      0      0  00:01:32        250000

Total number of neighbors 2

Total number of Established sessions 2
```

TCAM is full

```
PE4#show ip route summary

-------------------------------------------------
IP routing table name is Default-IP-Routing-Table(0)
-------------------------------------------------
IP routing table maximum-paths   : 8
Total number of IPv4 routes      : 965985
Total number of IPv4 paths       : 1921963
Pending routes (due to route max reached): 4025
Route Source    Networks
connected       7
bgp             965978
Total           965985
FIB             965985

ECMP statistics (active in ASIC):
 Total number of IPv4 ECMP routes : 955978
 Total number of IPv4 ECMP paths  : 1911956
 Number of routes with  2 ECMP paths: 955978
PE4#show ipv6 route summary

-------------------------------------------------
IPv6 routing table name is Default-IPv6-Routing-Table(0)
-------------------------------------------------
IPv6 routing table maximum-paths : 8
Total number of IPv6 routes      : 81205
Total number of IPv6 paths       : 162407
Pending routes (due to route max reached): 168799
Route Source    Networks
connected       3
bgp             81202
Total           81205
FIB             81205

ECMP statistics (active in ASIC):
 Total number of IPv6 ECMP routes   : 81202
 Total number of IPv6 ECMP paths    : 162404
 Number of routes with  2 ECMP paths: 81202
```

Simulated full table – Q2a

```
PE4#show forwarding profile limit
(s): public and private tables are shared

Active in HW: l3-xl (cfg1)
-------------------------------------------------------------------------
|                      Forwarding Profile Table Size                    |
-------------------------------------------------------------------------
|cfg  |Entry    |Balanced  |Balanced  |l2-xl     |l3-xl     |Balanced  |
|     |         |          |p-and-p   |          |          |Extended  |
-------------------------------------------------------------------------
|      MAC       355.0k     372.0k     590.0k     77.5k      253.0k    |
|                                                                       |
|      SRv6-SID  30.0k      30.0k      30.0k      37.0k      30.0k     |
|                                                                       |
|      LPM                                                              |
|1(s)  Public    1048.0k    917.0k     131.0k     1441.0k    917.0k    |
|      Private   1048.0k    917.0k     131.0k     1441.0k    917.0k    |
-------------------------------------------------------------------------
```

NOT RECOMMENDED for full IRT – even though it is fits in the lab the device is at capacity and will be overloaded in the field

```
PE4#show ip route summary

-------------------------------------------------
IP routing table name is Default-IP-Routing-Table(0)
-------------------------------------------------
IP routing table maximum-paths   : 8
Total number of IPv4 routes      : 970003
Total number of IPv4 paths       : 970003
Pending routes (due to route max reached): 0
Route Source    Networks
connected       6
bgp             969997
Total           970003
FIB             970003

ECMP statistics (active in ASIC):
 Total number of IPv4 ECMP routes : 0
 Total number of IPv4 ECMP paths  : 0
PE4#show ipv6 route summary

-------------------------------------------------
IPv6 routing table name is Default-IPv6-Routing-Table(0)
-------------------------------------------------
IPv6 routing table maximum-paths : 8
Total number of IPv6 routes      : 250003
Total number of IPv6 paths       : 250003
Pending routes (due to route max reached): 0
Route Source    Networks
connected       3
bgp             250000
Total           250003
FIB             250003

ECMP statistics (active in ASIC):
 Total number of IPv6 ECMP routes   : 0
 Total number of IPv6 ECMP paths    : 0
```

Simulated full table – Q2a with GRT and VRF

```
PE4#show ip route summary

-----------------------------------------------
IP routing table name is Default-IP-Routing-Table(0)
-----------------------------------------------
IP routing table maximum-paths   : 8
Total number of IPv4 routes      : 970003
Total number of IPv4 paths       : 970003
Pending routes (due to route max reached): 0
Route Source    Networks
connected       6
bgp             969997
Total           970003
FIB             970003

ECMP statistics (active in ASIC):
 Total number of IPv4 ECMP routes : 0
 Total number of IPv4 ECMP paths  : 0
```

```
PE4#show ip route vrf TEST summary

-----------------------------------------------
IP routing table name is TEST(2)
-----------------------------------------------
IP routing table maximum-paths   : 8
Total number of IPv4 routes      : 274336
Total number of IPv4 paths       : 274336
Pending routes (due to route max reached): 160812
Route Source    Networks
connected       2
bgp             274334
Total           274336
FIB             274336

ECMP statistics (active in ASIC):
 Total number of IPv4 ECMP routes : 0
 Total number of IPv4 ECMP paths  : 0
```

Doubles the number of entries used ~ 1.25m prefixes installed

Each vrf utilizes it's own hardware resources i.e. "separate" tables

What did we learn

- FIB and RIB sizes are independent of each other – except when they're not (VRFs)
- Routing tables need to be managed
- Various options to manage the routing table without scaling up hardware
- With the current generation of silicon there is runway for growth
- 10m RIB != 10m FIB

Set up a Free Consultation /

Apply for a Proof of Concept

www.epsglobal.com

Thank You

FIB Scaling on Merchant Silicon

Vince Schuele