



Advanced Services

UCL / UKERNA

Cisco 12000 QoS Testing - Bedfont Lakes

Version 1.0

Corporate Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
<http://www.cisco.com>
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100



Contents

Contents	2
Document Control.....	4
History	4
Review	4
Testing Objectives.....	5
Method.....	6
Tests	6
Test A - 2 x Engine 2 Input, 1 x Engine 2 Output.....	6
Test B - 2 x Engine 2 Input, 1 x Engine 3 Output.....	6
Test C - 1 x Engine 2, 1 x Engine 3 Input, 1 x Engine 2 Output.....	7
Test D - Police and Drop Ingress on Engine 3	7
Software.....	7
Traffic Characteristics.....	8
QoS configuration	8
Supported Classes	8
Queuing + Congestion Avoidance.....	8
IOS QoS Configuration Templates.....	9
From Fabric QoS - MDRR / WRED (Engine 2 Linecards).....	9
From Fabric QoS - MDRR / WRED (Engine 3 Linecards).....	10
To Fabric QoS - MDRR / WRED	10
Results.....	11
Overview of Tests Performed.....	11
Test Results.....	12
Conclusion	14



Appendix 1 - Useful show commands.....16

Appendix 2 - Router configuration.....17



Document Control

Author: Nick Carter
Advanced Services
Change Authority: Cisco Systems Advanced Services
Reference Number: KBMS stored document

History

Table 1 *Revision History*

Version No.	Issue Date	Status	Reason for Change
1.0	19/06/03	First Version	First Version

Review

Table 2 *Revision Review*

Reviewer's Details	Version No.	Date

Change Forecast: High

This document will be kept under revision control.



Testing Objectives

This testing was performed as part of the MB-NG (Managed Bandwidth - Next Generation) project. The objective of this testing was to derive a QoS configuration for the 12000 routers that can be used in the MB-NG network. By testing a single 12000 in the lab, we can have confidence that we understand the QoS characteristics of the 12000 routers in the core of the MB-NG network.

4 tests were originally planned:

- Test A - 2 x Engine 2 Input, 1 x Engine 2 Output
- Test B - 2 x Engine 2 Input, 1 x Engine 3 Output.
- Test C - 1 x Engine 2, 1 x Engine 3 Input, 1 x Engine 2 Output.
- Test D - Police and Drop Ingress on Engine 3

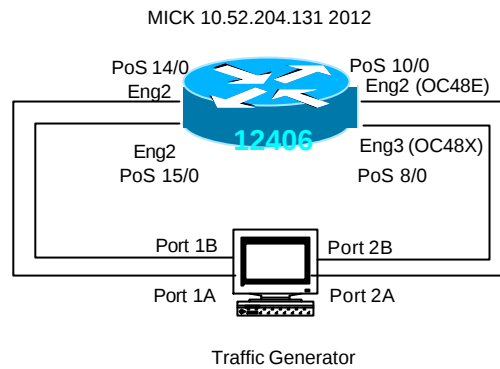
Tests A and B were successfully completed. Test C requires a very new IOS image, version 12.0(23)SZ. This is an engineering "special" image and is not officially supported by Cisco. Test C will also work on a future supported version of IOS - 12.0(26)S (available August 2003). It is planned to run tests C and D, later this year, when this supported image is available.

Method

Tests

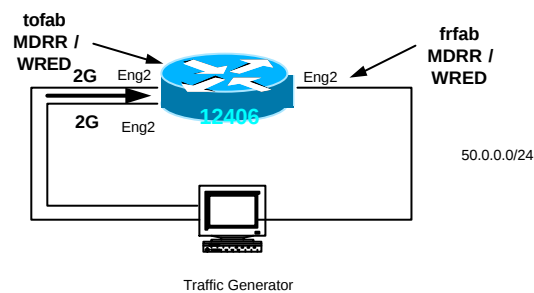
Physical

All links POS STM-16
2488.32Mb/s (payload 2405.38)
IP addressing
Router 223.1.slot.1
Traffic Gen 223.1.slot.2



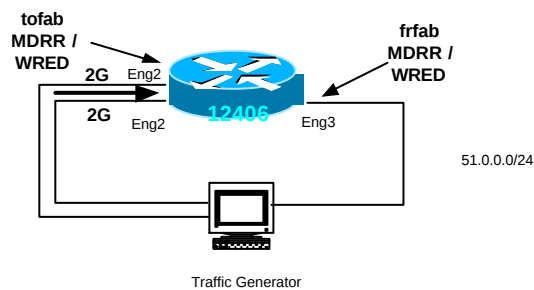
Test A - 2 x Engine 2 Input, 1 x Engine 2 Output

Test A - 2 x E2 Input. 1 x E2 output



Test B - 2 x Engine 2 Input, 1 x Engine 3 Output.

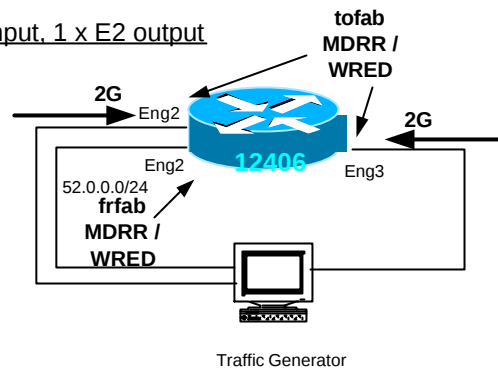
Test B - 2 x E2 Input. 1 x E3 output



Test C - 1 x Engine 2, 1 x Engine 3 Input, 1 x Engine 2 Output.

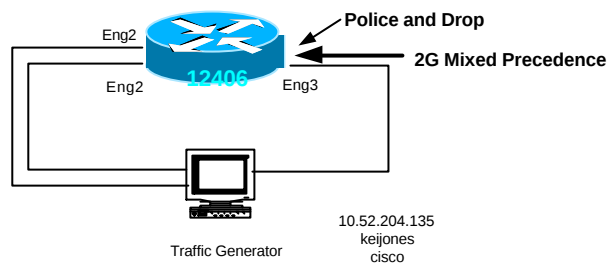
Future Test

Test C - 1 x E3 input, 1 x E2 Input, 1 x E2 output



Test D - Police and Drop Ingress on Engine 3

Test 4 - Police and Drop ingress on Engine 3



Software

IOS 12.0(21)S7 Service Provider (file name: gsr-p-mz.120-21.S7.bin)

Traffic Characteristics

Each Port on the traffic generator produced 2G of traffic, with the following characteristics, destined for 50.0.0.1. Tests were performed using an IP streams "blasted" at the unit under test. There was no host TCP/IP stack running on the traffic generator. Packets were sent regardless of what was being received.

Each packet had a length of 1000 bytes.

The traffic classes were based on the classes proposed by UKERNA as part of the JANET QoS development project. http://www.ja.net/development/qos/qos_dev.html The Premium Class is designed for applications that require a low delay and a low jitter. The IP Plus class is a premium data service with low packet loss and a minimum guaranteed throughput. The Best Effort class is the standard data class which may experience significant packet loss. The Less Than BE class is used for applications that do not have tight network requirements, such as bulk file transfer. The BE class has a small lossless minimum bandwidth to ensure long running applications do not time out through loss of keepalives.

Class	IP Precedence	% of 2G	Mb/s
Network Updates	7 6	0 3	0 60
Premium	5	5	100
<i>Reserved for future use</i>	4	0	0
IP Plus	3	5	100
<i>Reserved for future use</i>	2	0	0
Best Effort	1	86	1,720
Less Than BE	0	1	20

QoS configuration

Supported Classes

Class Name	DSCP Value
Network Updates	6,7
Premium	5
IP Plus	3
Best Effort	1
Less Than BE	0

Queuing + Congestion Avoidance

Queue Number	Class	IP Precedence Values	%	Weight	Quantum
0	Network Updates	6, 7	5	37	23040
1	Premium	5	10	PQ	PQ
2	IP Plus, Best Effort	4,3,2 1	84	748	387072
3	Less Than BE	0	1	1	4608

IOS QoS Configuration Templates

From Fabric QoS - MDRR / WRED (Engine 2 Linecards)

```
cos-queue-group stm16-qos
! Assign class to queues
precedence 0 queue 3
precedence 1 queue 2
precedence 2 queue 2
precedence 3 queue 2
precedence 4 queue 2
precedence 5 queue low-latency
precedence 6 queue 0
precedence 7 queue 0

! Define Queues (MDRR Bandwidths)
queue 0 37
queue 2 748
queue 3 1
queue low-latency strict-priority

! Define WRED profiles
random-detect-label 1 1 2 1
random-detect-label 2 2000 2001 1

! Assign WRED profiles to each class
precedence 0 random-detect-label 2
precedence 1 random-detect-label 1
precedence 2 random-detect-label 2
precedence 3 random-detect-label 2
precedence 4 random-detect-label 2
precedence 5 random-detect-label 2
precedence 6 random-detect-label 2
precedence 7 random-detect-label 2

! Apply from fabric MDRR / WRED
int pos x/y
tx-cos stm16-qos
```

From Fabric QoS - MDRR / WRED (Engine 3 Linecards)

```
! Engine 3 Linecards must be configured using MQC

! Assign class to queues
class-map match-any queue-0
  match ip precedence 6
  match ip precedence 7
class-map match-any queue-1
  match ip precedence 5
class-map match-any queue-2
  match ip precedence 4
  match ip precedence 3
  match ip precedence 2
  match ip precedence 1
class-map match-any queue-3
  match ip precedence 0

! Define Queues and Assign WRED profiles to each class
policy-map e3-stm16-qos
  class queue-0
    bandwidth percent 5
    random-detect
    random-detect precedence 6 2000 2001 1
    random-detect precedence 7 2000 2001 1
  class queue-1
    priority
  ! Need to police here
  police 200000000 2000 2000 conform-action transmit exceed-action drop
  class queue-2
    bandwidth percent 84
    random-detect
    random-detect precedence 1 1 2 1
    random-detect precedence 2 2000 2001 1
    random-detect precedence 3 2000 2001 1
    random-detect precedence 4 2000 2001 1
  class queue-3
    bandwidth percent 1
    random-detect
    random-detect precedence 0 2000 2001 1

! Apply from fabric QoS
int pos x/y
  service-policy output e3-stm16-qos
```

To Fabric QoS - MDRR / WRED

```
! RX QoS on Engine 2 or Engine 3 (E3 requires 12.0(23)SZ)
! 12.0(26)S will be first supported CCO release (due in August)

slot-table-cos rx-table
! Apply "stm16-qos" cos-queue-group as defined earlier to all slots
  destination-slot all stm16-qos

! Apply to fabric QoS to slots 14, 15, 16
rx-cos-slot 14 rx-table
rx-cos-slot 15 rx-table
rx-cos-slot 8 rx-table
```

Results

Overview of Tests Performed

16th - 17th June 2003: Bedfont Lakes

Each test ran for a duration of 20 minutes.

Test A - 2 x Engine 2 Input, 1 x Engine 2 Output

Test B - 2 x Engine 2 Input, 1 x Engine 3 Output.

Future tests:

Test C - 1 x Engine 2, 1 x Engine 3 Input, 1 x Engine 2 Output.

Test D - Police and Drop Ingress on Engine 3

N.B. When loading a new test configuration into the traffic generator we must wait before starting the test. The traffic generator ports are shutdown then re-enabled, which causes the router ports to go down, then come back up. The traffic generator is ready to transmit traffic even before the router has cleared the POS alarms from the router ports. We must wait for the POS alarms to clear on the router ports before starting each test.

Test Results

Test	Description	Src Port	Dest. Port	Prec	Tx Rate ¹ (Mb/s)	Rx Rate (Mb/s)	Tx (Frames)	Rx (Frames)	Lost (Frames)	Ave. Latency (uS)	Max Latency (uS)	Min Latency (uS)
A	2 x Engine 2 Input, 1 x Engine 2 Output	1A	2A	All	2,000.00		300,000,000					
	Config: Test-A.xml			7	0	0	0	0	0	0	0	0
	Results: Test-A-20mins.csv			6	60.00	60.00	9,000,000	9,000,000	0	927.81	1866.46	34.44
				5	100.00	100.00	15,000,000	15,000,000	0	80.55	90.62	37.54
				4	0	0	0	0	0	0	0	0
				3	100.00	100.00	15,000,000	15,000,000	0	926.21	1871.88	23.78
				2	0	0	0	0	0	0	0	0
				1	1720.00	907.27	258,000,000	136,094,074	121,905,926	928.00	1873.14	28.26
				0	20.00	20.00	3,000,000	3,000,000	0	924.46	1863.38	30.86
		1B	2A	All	2,000.00		300,000,000					
				7	0	0	0	0	0	0	0	0
				6	60.00	60.00	9,000,000	9,000,000	0	927.21	1866.44	31.70
				5	100.00	100.00	15,000,000	15,000,000	0	80.47	90.62	34.78
				4	0	0	0	0	0	0	0	0
				3	100.00	100.00	15,000,000	15,000,000	0	925.63	1867.48	20.52
				2	0	0	0	0	0	0	0	0
				1	1720.00	907.31	258,000,000	136,094,074	121,906,519	927.44	1872.22	25.02
				0	20.00	20.00	3,000,000	3,000,000	0	923.95	1861.00	28.10
		1A+1B	2A		4,000.00	2,374.58	600,000,000	356,187,555	243,812,445			

¹ Throughput readings taken whilst test was running

Test	Description	Src Port	Dest. Port	Prec	Tx Rate ¹ (Mb/s)	Rx Rate (Mb/s)	Tx (Frames)	Rx (Frames)	Lost (Frames)	Ave. Latency (uS)	Max Latency (uS)	Min Latency (uS)
B	2 x Engine 2 Input, 1 x Engine 3 Output	1A	2B	All	2,000.00		300,000,000					
	Config: Test-B.xml			7	0	0	0	0	0	0	0	0
	Results: Test-B-20mins.csv			6	60.00	60.00	9,000,000	9,000,000	0	631.38	1281.70	28.38
				5	100.00	100.00	15,000,000	14,999,898	102	88.36	98.48	28.88
				4	0	0	0	0	0	0	0	0
				3	100.00	100.00	15,000,000	15,000,000	0	877.05	1862.78	38.54
				2	0	0	0	0	0	0	0	0
				1	1720.00	907.27	258,000,000	136,088,770	121,911,230	878.87	1866.22	42.14
				0	20.00	20.00	3,000,000	3,000,000	0	1172.29	5036.16	32.36
		1B	2A	All	2,000.00		300,000,001					
				7	0	0	0	0	0	0	0	0
				6	60.00	60.00	9,000,001	9,000,001	0	631.45	1280.60	25.10
				5	100.00	100.00	15,000,000	14,999,898	102	88.37	98.44	28.90
				4	0	0	0	0	0	0	0	0
				3	100.00	100.00	15,000,000	15,000,000	0	877.06	1859.34	35.80
				2	0	0	0	0	0	0	0	0
				1	1720.00	907.33	258,000,000	136,097,426	121,902,574	878.95	1866.96	37.34
				0	20.00	20.00	3,000,000	3,000,000	0	1172.68	5211.94	29.10
		1A+1B	2A		4,000.00	2,372.60	600,000,001	356,185,993	243,814,008			
C	1 x Engine 3 input, 1 x Engine 2 input, 1 x Engine 2 Output											
	<i>Future Test</i>											
D	Police and Drop Ingress on Engine 3											
	<i>Future Test</i>											

Conclusion

The Cisco 12000 router under test, performed well in both Test A and Test B

In test A we saw no packet loss in any of the classes, except the expected loss in the BE class.

The overall throughput achieved (2,374.58Mbps) was 99% of the payload of an STM-16. [The HDLC headers added to the packets would not have been included in this throughput figure.] The throughput of each non-BE class was as exactly as requested.

The average latency, in any class did not even reach 1 ms. A latency of <1 ms, per multiple hops, will support even the most demanding application delay budget i.e. voice. The maximum latency experienced in the Premium queue was only 90.62us, a very low figure, well within the requirements of the most demanding mainstream applications around today.

In test B we only saw packet loss in the Premium Class and the BE class. The packet loss in the Premium class was unexpected. During the 20 minute test we lost 204 premium class packets. This is a very small percentage of the total premium class packets sent, around $6.8e-6$ %. However we would expect to see no packet loss in the premium class. A further test running for 90 minutes showed a $9.3e-6$ % premium class packet loss. It is possible that this packet loss is increasing over time, which would be a worrying trend. It is worth noting that in a real network a device would not be under this continuous overload for this length of time. Further investigation is required, although with such a low loss rate it may be hard to determine exactly what is causing the packet loss.

The overall throughput achieved (2,372.60Mbps) was 99% of the payload of an STM-16. The throughput of each non-BE class was as exactly as requested.

Again the latency figures were very good and no class had an average latency greater than 1.2ms. The maximum latency experienced in the Premium queue was only 98.48us.

In both Test A and Test B we expected to drop some LBE traffic as we were sending 40M and only allocating 24M (1%). But we did not drop any LBE. This could be caused by a rounding error or the minimum granularity supported by the queuing scheduler. Further testing showed it was possible to send up to 86M in the LBE class before any traffic was dropped. This is around 3.6% of the throughput of the link. Further investigation is required to determine the cause of this variation between configuration and observed performance.

Whilst the tests were running both the "from fabric" and the "to fabric" queues were filling. We could observe this using the "exec slot n sh controllers tofab queue" and "exec slot n sh controllers frfab queue" commands. The queuing in the "from fabric" queues was not expected, as we had approximately the same bandwidth in to the queues - 2.5G (Switch fabric to "from fabric" queues) as out of the queues - 2.5G (fabric queues to Port). The switch fabric to "from fabric" queue bandwidth must be higher, for the from fabric queues to build.

Since all the "to fabric" queues share the same memory address space, it is possible for a single queue to consume all the memory available and cause all the other queues to fail. If the WRED profile for a traffic class is not aggressive enough, then the traffic in this single traffic class can affect all other classes. We saw this during our testing as we were refining our WRED profiles. **To properly isolate one traffic class from affecting another traffic class it is very important to apply a correct WRED profile to every traffic class.**

The QoS configuration of the router was optimised to cope with a "packet blast" test. For operation on a production network carrying real traffic flows the routers congestion avoidance configuration should be softened. For example the WRED profile applied to the Best Effort Class was very aggressive and under a real network load would start discarding traffic too soon and too fast, to the detriment of the end hosts.

The configuration of the router, and the router IOS version, used during tests A and B, achieved a consistent and reliable QoS performance. It is recommended that this is the initial configuration and IOS version to be deployed in the MB-NG 12000 routers.

Appendix 1 - Useful show commands

```
! Overview of cos
sh cos

! tofab queuing
exec slot 15 sh controllers tofab queue
exec slot 15 sh controllers tofab queue 10

! frfab queuing (engine 2 and engine 3)
exec slot 15 sh controllers frfab queue

! WRED frfab
sh int pos 10/0 random-detect

! Input / Output tx / rx drops
sh int pos 10/0
```

Appendix 2 - Router configuration

```
Mick16-5#sh run
Building configuration...

Current configuration : 5976 bytes
!
version 12.0
no service pad
service timestamps debug datetime msec
service timestamps log uptime
service password-encryption
!
hostname Mick16-5
!
boot system flash slot1:gsr-p-mz.120-21.S7.bin
enable secret 5 $1$NB0s$DVgesRDik8vDDxSAb1cA9/
!
username guest privilege 15 password 7 030752180500
!
hw-module slot 11 shutdown
hw-module slot 13 shutdown
!
!
hw-module slot 12 srp
!
!
!
!
ip subnet-zero
no ip domain-lookup
!
class-map match-any queue-0
  match ip precedence 6
  match ip precedence 7
class-map match-any queue-1
  match ip precedence 5
class-map match-any queue-2
  match ip precedence 4
  match ip precedence 3
  match ip precedence 2
  match ip precedence 1
class-map match-any queue-3
  match ip precedence 0
!
!
policy-map nicktemp
policy-map e3-stm16-qos
  class queue-0
    bandwidth percent 5
  random-detect
```

```

random-detect precedence 6 2000 2001 1
random-detect precedence 7 2000 2001 1
  class queue-1
  priority
    police 200000000 2000 2000 conform-action transmit exceed-action
drop
  class queue-2
  bandwidth percent 84
  random-detect
  random-detect precedence 1 1 2 1
  random-detect precedence 2 2000 2001 1
  random-detect precedence 3 2000 2001 1
  random-detect precedence 4 2000 2001 1
  class queue-3
  bandwidth percent 1
  random-detect
  random-detect precedence 0 2000 2001 1
!
!
controller SYSCLOCK 3
!
controller SONET 6/0
  clock source line
  framing sonet
!
  sts-1 1
    mode ct3
!
  sts-1 2
    mode ct3
!
  sts-1 3
    mode ct3
!
controller SONET 6/1
  clock source line
  framing sonet
!
  sts-1 1
    mode ct3
!
  sts-1 2
    mode ct3
!
  sts-1 3
    mode ct3
!
!
interface Loopback0
  ip address 223.2.5.1 255.255.255.255
  no ip directed-broadcast
  no ip route-cache
  no ip mroute-cache
!
interface POS0/0
  no ip address
  no ip directed-broadcast
  shutdown

```

```
crc 32
!
interface POS0/1
  no ip address
  no ip directed-broadcast
  shutdown
  crc 32
!
interface POS0/2
  no ip address
  no ip directed-broadcast
  shutdown
  crc 32
!
interface POS0/3
  no ip address
  no ip directed-broadcast
  shutdown
  crc 32
!
interface POS1/0
  description OC-192 Noam pos 1/0
  ip address 223.45.1.1 255.255.255.254
  no ip directed-broadcast
  no keepalive
  shutdown
  crc 32
  clock source internal
  pos ais-shut
!
interface POS3/0
  no ip address
  no ip directed-broadcast
  shutdown
  crc 32
!
interface POS3/1
  no ip address
  no ip directed-broadcast
  shutdown
  crc 32
!
interface POS3/2
  no ip address
  no ip directed-broadcast
  shutdown
  crc 32
!
interface POS3/3
  no ip address
  no ip directed-broadcast
  shutdown
  crc 32
!
interface POS5/0
  description OC-12 Noam pos 5/0
  ip address 223.45.5.1 255.255.255.254
  no ip directed-broadcast
```

```
no keepalive
shutdown
crc 32
clock source internal
pos ais-shut
pos scramble-atm
!
interface POS5/1
description OC-12 Noam pos 5/1
ip address 223.45.5.3 255.255.255.254
no ip directed-broadcast
no keepalive
shutdown
crc 32
clock source internal
pos ais-shut
pos scramble-atm
!
interface POS5/2
description OC-12 Noam pos 5/2
ip address 223.45.5.5 255.255.255.254
no ip directed-broadcast
no keepalive
shutdown
crc 32
clock source internal
pos ais-shut
pos scramble-atm
!
interface POS5/3
description OC-12 Noam pos 5/3
ip address 223.45.5.7 255.255.255.254
no ip directed-broadcast
no keepalive
shutdown
crc 32
clock source internal
pos ais-shut
pos scramble-atm
!
interface POS7/0
description OC-3 Noam pos 7/0
ip address 223.45.7.1 255.255.255.254
no ip directed-broadcast
no keepalive
shutdown
crc 32
clock source internal
pos ais-shut
pos scramble-atm
!
interface POS7/1
description OC-3 Noam pos 7/1
ip address 223.45.7.3 255.255.255.254
no ip directed-broadcast
no keepalive
shutdown
crc 32
```

```
clock source internal
pos ais-shut
pos scramble-atm
!
interface POS7/2
description OC-3 Noam pos 7/2
ip address 223.45.7.5 255.255.255.254
no ip directed-broadcast
no keepalive
shutdown
crc 32
clock source internal
pos ais-shut
pos scramble-atm
!
interface POS7/3
description OC-3 Noam pos 7/3
ip address 223.45.7.7 255.255.255.254
no ip directed-broadcast
no keepalive
shutdown
crc 32
clock source internal
pos ais-shut
pos scramble-atm
!
interface POS8/0
ip address 223.1.8.1 255.255.255.0
no ip directed-broadcast
no keepalive
service-policy output e3-stm16-qos
crc 32
clock source internal
pos ais-shut
pos scramble-atm
no cdp enable
!
interface POS10/0
description OC-48 Till pos 1/3
ip address 223.1.10.1 255.255.255.0
no ip directed-broadcast
no keepalive
tx-queue-limit 1000
crc 32
clock source internal
pos ais-shut
pos scramble-atm
tx-cos stm16-qos
!
interface SRP12/0
no ip address
no ip directed-broadcast
shutdown
!
interface POS14/0
ip address 223.1.14.1 255.255.255.0
no ip directed-broadcast
no keepalive
```

```

crc 32
clock source internal
pos ais-shut
pos scramble-atm
!
interface POS15/0
ip address 223.1.15.1 255.255.255.0
no ip directed-broadcast
no keepalive
crc 32
clock source internal
pos ais-shut
pos scramble-atm
tx-cos stm16-qos
!
interface Ethernet0
ip address 10.52.204.144 255.255.255.224
no ip directed-broadcast
no ip route-cache
no ip mroute-cache
!
ip classless
ip route 50.0.0.0 255.255.255.0 223.1.10.2
ip route 51.0.0.0 255.255.255.0 223.1.8.2
ip route 52.0.0.0 255.255.255.0 223.1.15.2
!
rx-cos-slot 8 rx-table
rx-cos-slot 14 rx-table
rx-cos-slot 15 rx-table
!
slot-table-cos rx-table
destination-slot all stm16-qos
!
cos-queue-group stm16-qos
precedence 0 random-detect-label 2
precedence 1 random-detect-label 1
precedence 2 random-detect-label 2
precedence 3 random-detect-label 2
precedence 4 random-detect-label 2
precedence 5 queue low-latency
precedence 5 random-detect-label 2
precedence 6 random-detect-label 2
precedence 7 random-detect-label 2
random-detect-label 1 1 2 1
random-detect-label 2 2000 2001 1
queue 0 37
queue 2 748
queue 3 1
queue low-latency strict-priority
!
!
line con 0
exec-timeout 0 0
login local
line aux 0
line vty 0 4
login local
!

```

end