TS Open Day
F2 F2E card input discard troubleshooting

Rong Cheng
Data Center/Lan Switching Team TAC CSE

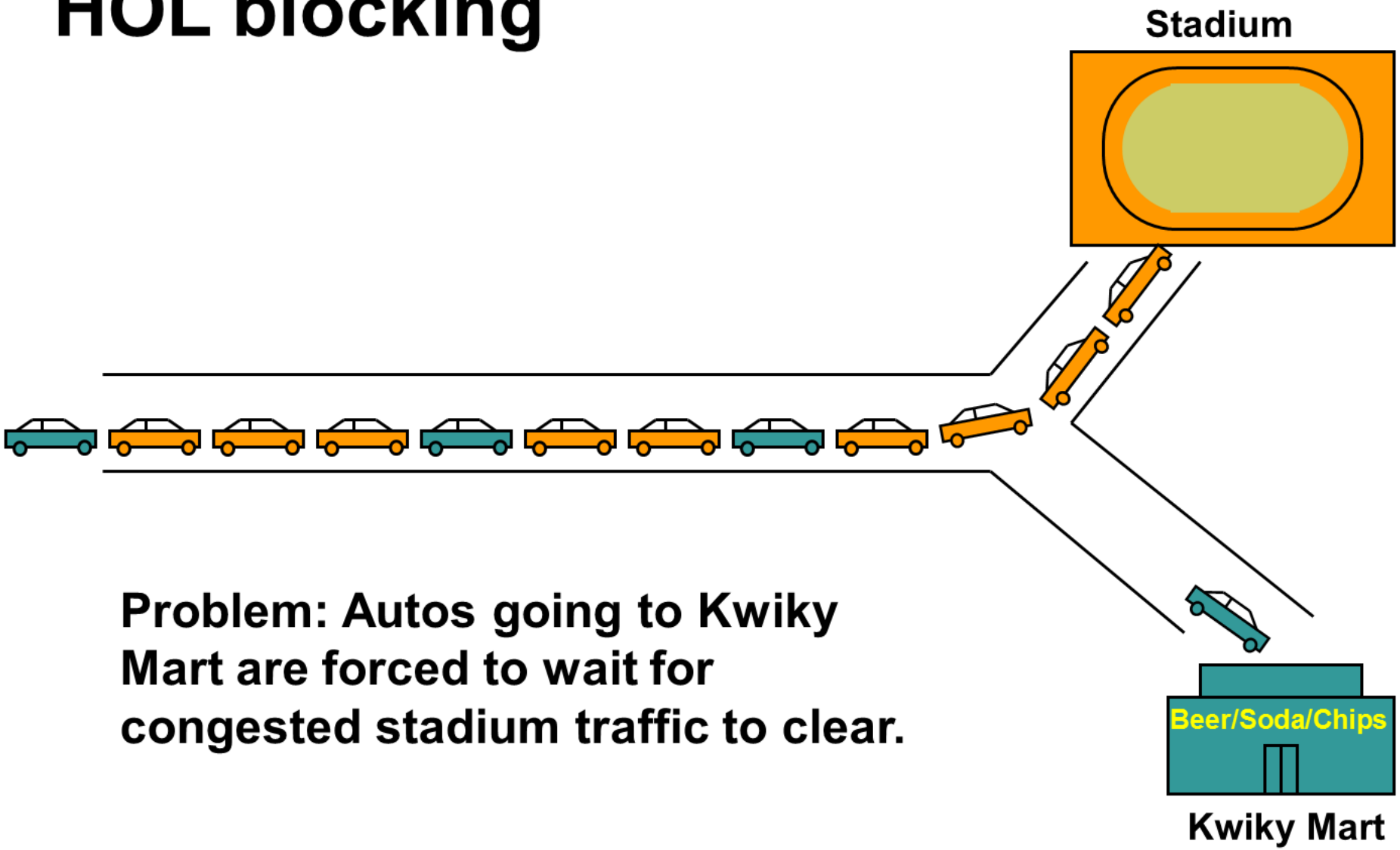
Dec-18-2015

Agenda

- Virtual Output Queues
- Input Discarding Troubleshooting -F2 Card
- Input Discarding Troubleshooting -F2E Card
- LAB Test

Virtual Output Queues (VOQs)

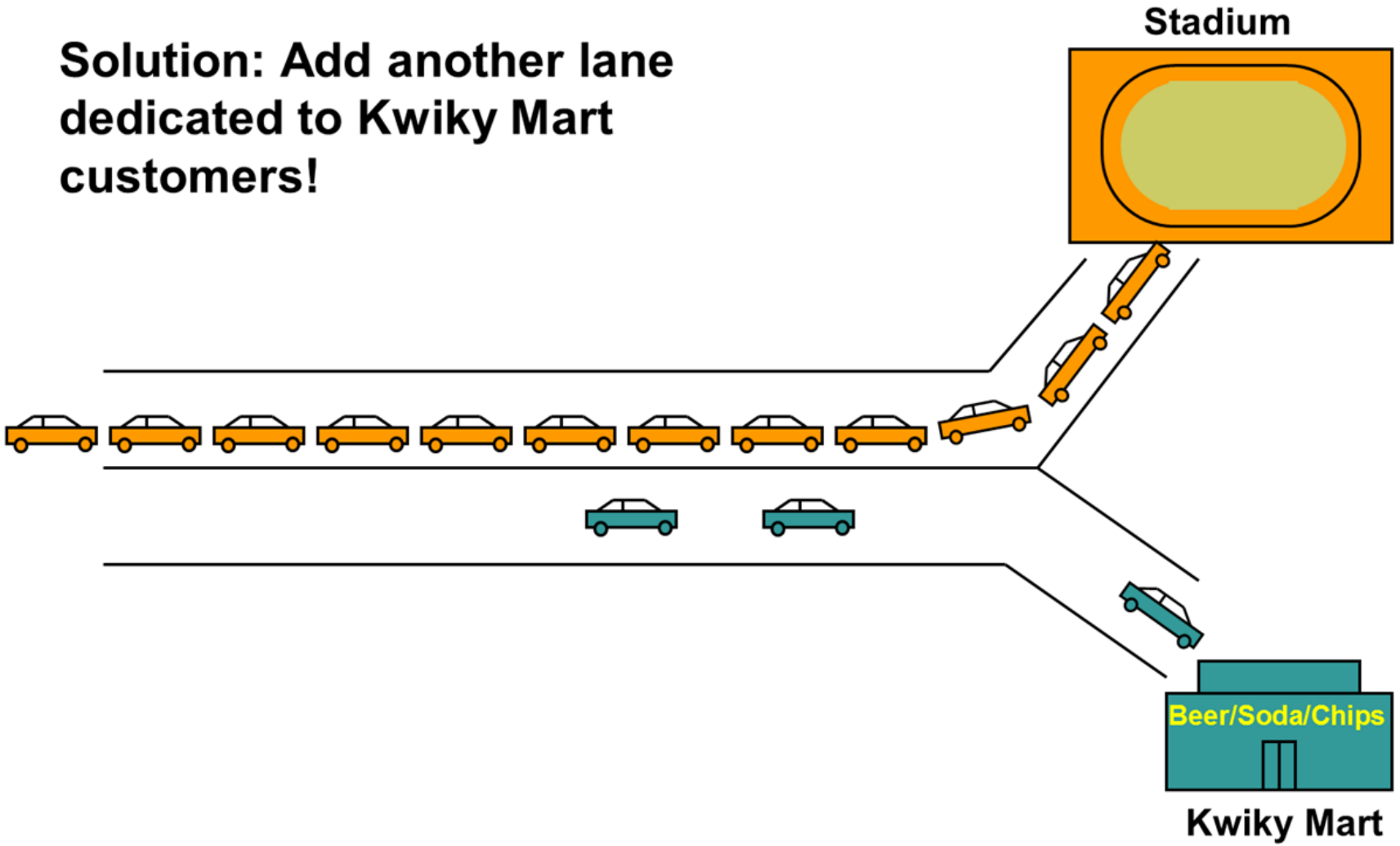
HOL blocking



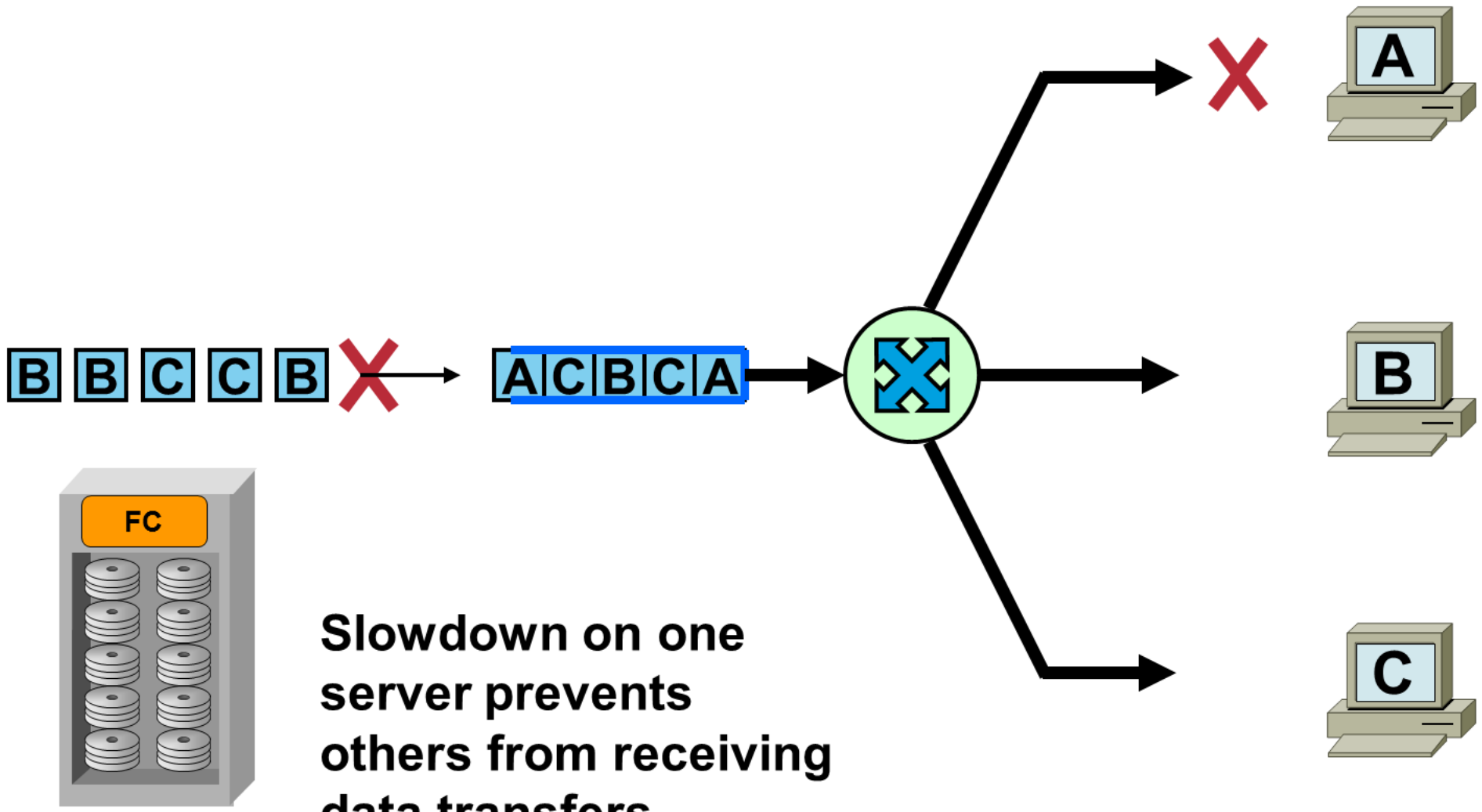
Problem: Autos going to Kwiky Mart are forced to wait for congested stadium traffic to clear.

VoQ concept

Solution: Add another lane dedicated to Kwiky Mart customers!

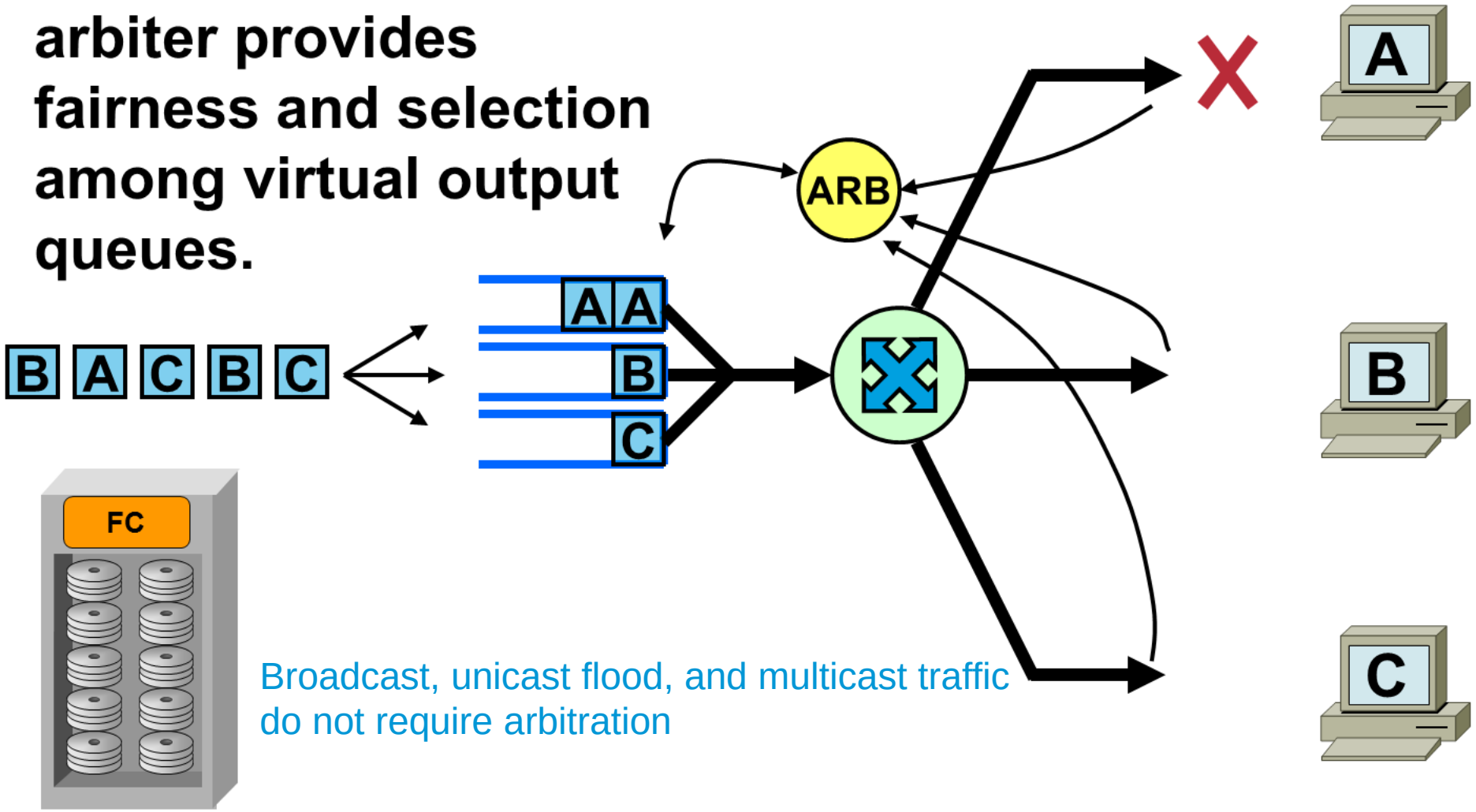


HoL Blocking



VoQ

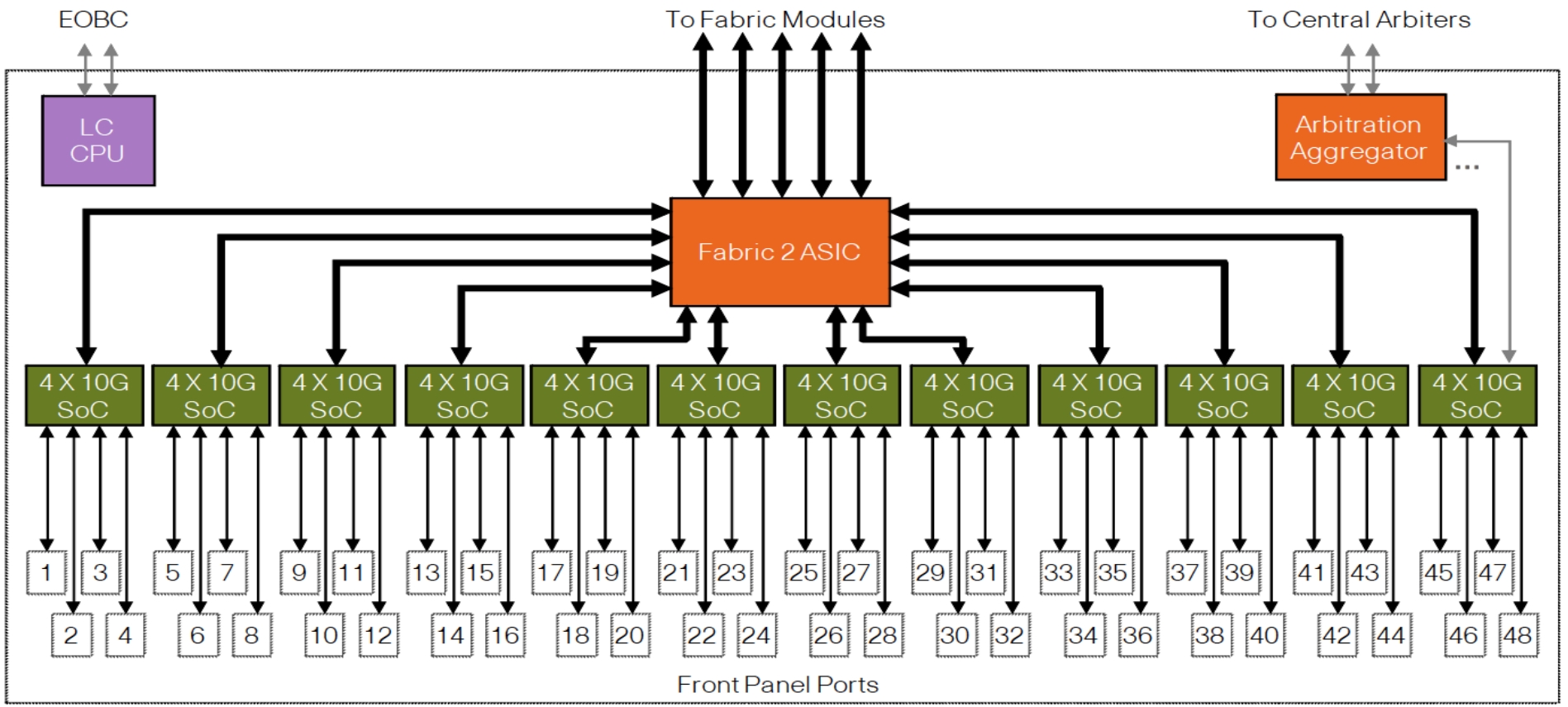
Sophisticated central arbiter provides fairness and selection among virtual output queues.



Broadcast, unicast flood, and multicast traffic do not require arbitration

Input Discarding Troubleshooting -F2 Card

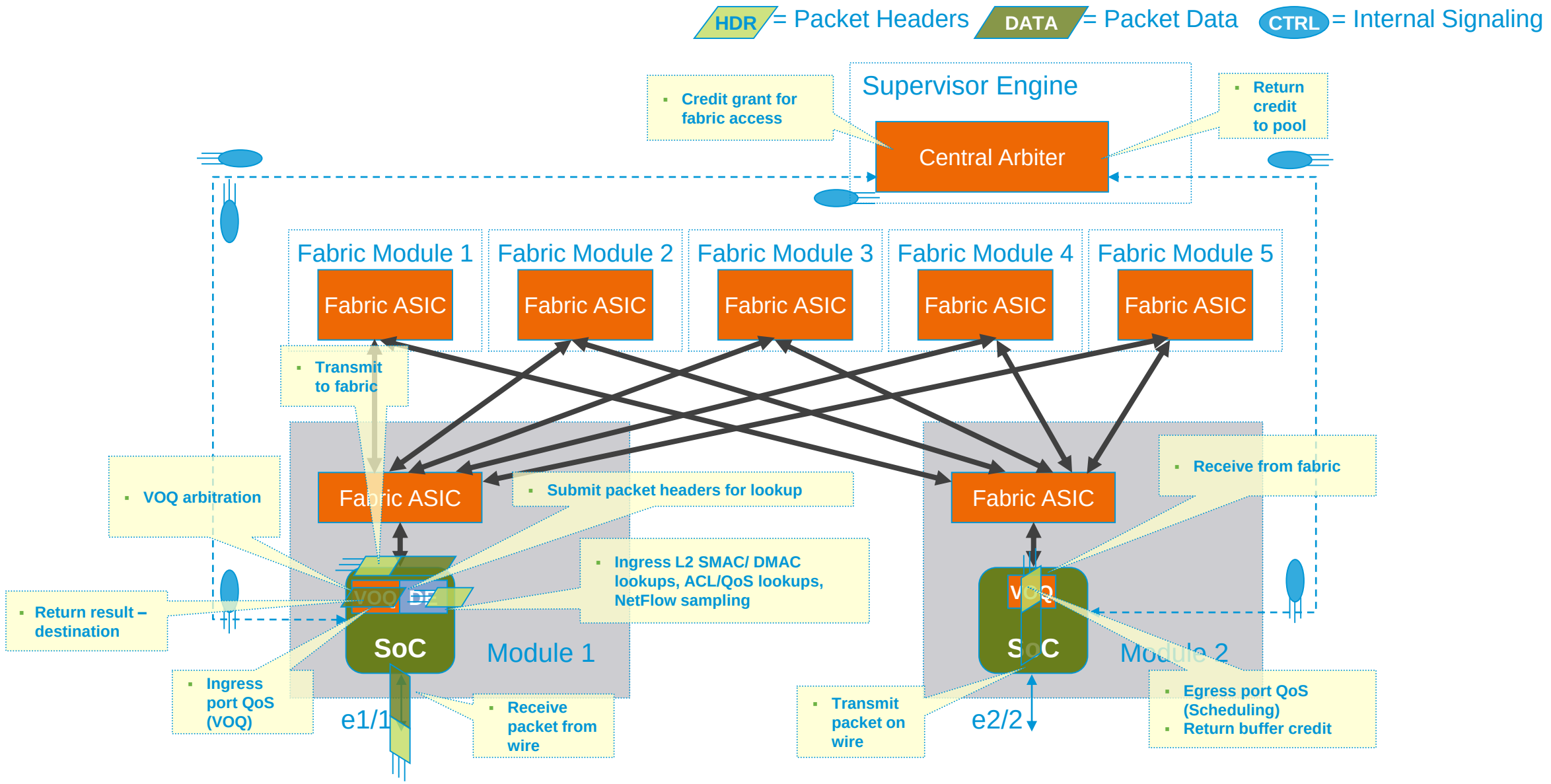
F2 Architecture Overview



With regard to data forwarding, each SoC has two dedicated connections:

1. To the arbitration aggregator
2. To a bundled connection to the 1st stage fabric, local to the module.

F2 Packet Flow



F2 Input discarding Troubleshooting

N7K1# **show int eth2/5 >>>module 2 is F2 card**
Ethernet2/5 is up
admin state is up, Dedicated Interface
---
  full-duplex, 1000 Mb/s

……
0 input with dribble 11590977 input discard

N7K1# **attach module 2**
module-2# **show hardware internal qengine inst 1 voq-status non-empty**

VQI:CCOS	BYTE_CNT	PKT_CNT	TAIL	HEAD	THR	
00**36**:3	6154	3077	6804	14168	1	<------- VQI is 36 here

F2 Input discarding Troubleshooting

module-2# **show hardware internal qengine vqi-map | egrep "VQI|36"**

```
VQI   SUP  SLOT  LDI  EQI  FPOE  NUM  XBAR   IN   ASIC  ASIC  SV  FEA_
NUM   VQI  NUM   NUM  NUM  BASE  DLS  MASK   ORD  TYPE  IDX   ID  TURE
----  ---  ----  ---  ---  ----  ---  -----  ---  ----  ----  --  ----
--snip
36        no     1        0        0       8        1    0x155       0    CLP      0     0   0x81
--snip
```

Mapping function

port = 4 * floor(LDI / 4) + r(LDI % 4)
r(0) = 2
r(1) = 1
r(2) = 3
r(3) = 4
Module Number = SLOT NUM + 1 (zero-based)

port = 4 * floor (0 / 4) + r (0 % 4)
port = 4 * 0 + r (0)
port = 4 * 0 + 2
port = 2

Module Number = 1 + 1 = 2

Physical Port = Eth 2/2

Input Discarding Troubleshooting -F2E Card

Advanced Feature in F2E

```
N7K-1# show hardware queuing drops egress
VQ Drops
-------------------------------------------------------------------------------
-
|     Output    |      VQ Drops      |   VQ Congestion    | Src | Src   |     Input
|
|   Interface   |                    |                    | Mod | Inst  |   Interface
|
-------------------------------------------------------------------------------
-
|      Eth1/1 | 0000000000393210 | 0000000000000000 |   1 |     7 |  Eth1/29-32|
-------------------------------------------------------------------------------
-
Egress Buffer Drops
-----------------------------------
|     Output    |      EB Drops      |
|   Interface   |                    |
|-----------------------------------|
|  Eth3/21-24 | 0000000000001200 |
```

VOQ drops are due to unicast /credited traffic

EB Drops are due to mcast / uncredited traffic

Advanced Feature in F2E

2. For F2E card, we have two options (burst-optimized & mesh-optimized) to control queues drop behavior

Burst-optimized – (default) All thresholds set to the maximum, used for better burst absorption when ports burst traffic to another port.

Mesh-optimized – VOQ thresholds set to smaller values so that packets in VOQs building faster are dropped first, used to manage congestion when steady traffic destined to multiple ports.

We can use following network-qos template to switch between this two modes.

```
N7K-7004-4(config)# policy-map type network-qos nq-8eone
N7K-7004-4(config-pmap-nqos)# class type network-qos c-nq-8e
N7K-7004-4(config-pmap-nqos-c)# congestion-control tail-drop threshold mesh-optimized
N7K-7004-4(config-pmap-nqos-c)# system qos
N7K-7004-4(config-sys-qos)# service-policy type network-qos nq-8eone
N7K-7004-4(config-sys-qos)# end
```

Lab Test 1

Step 1. N7K-3(config)# show int ex/y | in “input discard”

```
N7K-3# show int e4/23 | in "input discard"
    0 input with dribble  1254603 input discard
N7K-3# show int e4/23 | in "input discard"
    0 input with dribble  3895748 input discard

N7K-3# show int e4/27 | in "input discard"
    0 input with dribble  0 input discard

N7K-3# show int e4/35 | in discard
     0 input with dribble  0 input discard
    0 lost carrier  0 no carrier  0 babble  0 output discard
```

Common Causes of Input Discards

SPAN

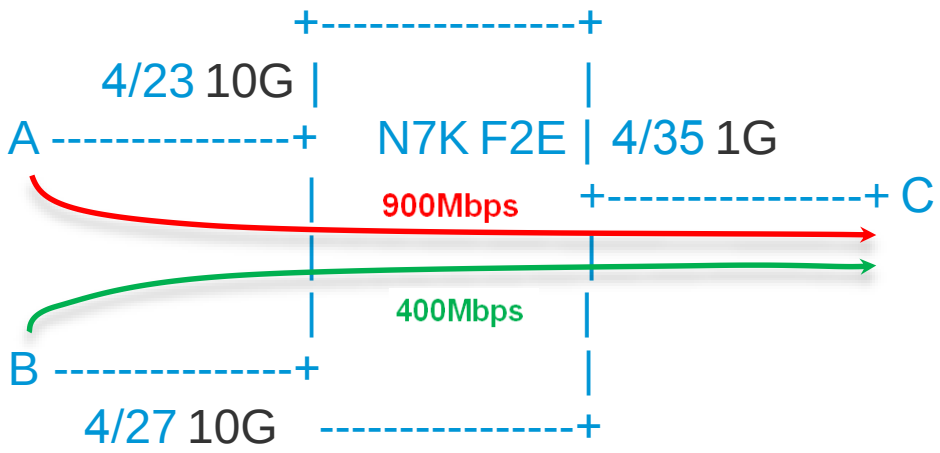
Inappropriate design

Burst Traffic

Lab Test

Lab Test 1

Topology:



we first begin by sending **900 Mbps** of traffic from host A to host C.
Then we add an additional **400 Mbps** traffic from host B sent to host C.
So now host C's 1 Gbps port is oversubscribed, and we begin seeing input discards on host A 's port.

Lab Test 1

Step 2. find the egress congestion port

```
N7K-3# attach module 4
Attaching to module 4 ...
To exit type 'exit', to abort type '$.'
module-4# show hardware internal qengine inst 5 voq-status non-empty
VQI:CCOS   BYTE_CNT   PKT_CNT   TAIL   HEAD   THR
--------   --------   -------   ----   ----   ---
0046:3     16880      1055      6035   6109   1
module-4# show hardware internal qengine inst 5 voq-status non-empty
VQI:CCOS   BYTE_CNT   PKT_CNT   TAIL   HEAD   THR
--------   --------   -------   ----   ----   ---
0046:3     16624      1039      11752  10473  1
```

```
module-4# show hardware internal qengine vqi-map | egrep "VQI|46"
VQI    SUP   SLOT   LDI   EQI   FPOE   NUM   XBAR    IN    ASIC   ASIC   SV   FEA_
NUM    VQI   NUM    NUM   NUM   BASE   DLS   MASK    ORD   TYPE   IDX    ID   TURE
46     no    3      34    2     170    1     0x155   0     CLP    8      0    0x80
```

Lab Test 1

Step 3. find the egress congestion port

Mapping function

Module Number = SLOT NUM + 1 (zero-based)

port = 4 * floor(LDI / 4) + r(LDI % 4)

r(0) = 2

r(1) = 1

r(2) = 3

r(3) = 4

SLOT NUM ==3

LDI NUM == 34

Module Number = 3 + 1 =4

port = 4 * floor (34 / 4) + r (34 % 4)

port = 4 * 8 + r (2)

port = 4 * 8 + 3

port = 35

Physical Port = Eth 4/35

Lab Test 1

Step 4. Then I make some changes, change the F2E to mesh-optimized mode.

```
N7K-3(config)# policy-map type network-qos nq-8eone
N7K-3(config-pmap-nqos)# class type network-qos c-nq-8e
N7K-3(config-pmap-nqos-c)# congestion-control tail-drop threshold mesh-optimized
N7K-3(config-pmap-nqos-c)# system qos
N7K-3(config-sys-qos)# service-policy type network-qos nq-8eone
N7K-3(config-sys-qos)# end
```

Step 5. After that, input discard was not found. The output discard will arise.

```
N7K-3# clear counters
N7K-3# show int e4/23 | in "input discard"
    0 input with dribble  0 input discard
N7K-3# show int e4/27 | in "input discard"
    0 input with dribble  0 input discard
N7K-3# show int e4/35 | in discard
     0 input with dribble  0 input discard
    0 lost carrier  0 no carrier  0 babble  655350 output discard
```

Lab Test 1

Step 6.

N7K-3(config)# clear statistics module 4 device all

N7K-3(config)# show hardware internal errors module 4

```
|-----------------------------------------------------------------------------|
| Device:Clipper XBAR                Role:QUE                        Mod: 4   |
| Last cleared @ Mon Dec 26 11:30:30 2016                                     |
| Device Statistics Category :: CONGESTION                                    |
|-----------------------------------------------------------------------------|
Instance:5
Cntr  Name                                          Value              Ports
----- ----                                          -----              -----
  131 VQ buffer tail drop count                     0000000000001166   21-24 -
  132 VQ credited pkt replica VOQ tail drops        0000000002309698   21-24 -
  137 VQ credited pkt replica drop count            0000000002309701   21-24 -
  752 VQ buf tail drop count 19                     0000000000001166   21-24 -
 8338 VQ VQI 46 CCOS 3 drop count                   0000000001883280   21-24 -
```

Lab Test 1

Step 7. Find the egress congested port.

Module-4#show hardware internal qengine vqi-map | egrep “VQI|46”

```
module-4# show hardware internal qengine vqi-map | egrep "VQI|46"
VQI    SUP   SLOT   LDI   EQI   FPOE   NUM   XBAR    IN    ASIC   ASIC   SV   FEA_
NUM    VQI   NUM    NUM   NUM   BASE   DLS   MASK    ORD   TYPE   IDX    ID   TURE
46     no    3      34    2     170    1     0x155   0     CLP    8      0    0x80
```

Lab Test 1

Step 8. Double check the egress congested port.

```
N7K-3(config)# show hardware queuing drops egress

VQ Drops
-------------------------------------------------------------------------------
|     Output    |      VQ Drops      |  VQ Congestion     | Src | Src  |    Input     |
|  Interface    |                    |                    | Mod | Inst |  Interface   |
-------------------------------------------------------------------------------
|      Eth4/35 | 00000000002604165 | 0000000000000000 |   4 |    5 |  Eth4/21-24 |
-------------------------------------------------------------------------------

Egress Buffer Drops
-----------------------------------
|     Output    |      EB Drops     |
|  Interface    |                   |
-----------------------------------
-----------------------------------
```

Thank you.