#SHAREorg

SHARE

Buffer-to-Buffer Credits, Exchanges, and Urban Legends

Lou Ricci, EMC
Howard L. Johnson, Brocade

14 August 2013 (8:00am – 9:00am)
Session 14281

QR Code

SHARE in Boston

Legal Stuff

- Notice
  - IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing to: *IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.*
  - Any references in this information to non-IBM Web sites are provied for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.
- Trademarks
  - The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both: FICON® IBM® Redbooks™ System z10™ z/OS® zSeries® z10™
  - Other Company, product, or service names may be trademarks or service marks of others.

Abstract

- Performance in a FICON network is influenced by the underlying flow control mechanisms of Fibre Channel. In this session, we examine how Buffer-to-Buffer credits flow from the channel to the control unit. We also look at how exchanges are used in FICON applications and how they change with the introduction of zHPF. During both examinations, we explore the role of the FICON Director in managing Buffer-to-Buffer credits and exchanges over a cascaded network. Throughout the session, we debunk the various FICON “Urban Legends” featuring credits and exchanges. Take the opportunity to learn from two of the FICON industry’s leading experts in channel and fabric development and join our session.

Agenda

- Buffer Credits
  - What are they and how do they work?
    - How do you fill the pipe?
    - What if you can't fill the pipe?
    - What's wrong with multiple senders and one receiver?
    - What happens when the pipes are different sizes?
    - What's it like in the real world?
- Exchanges
  - What are they and how do they work?
    - What's an Exchange?
    - How many exchanges are needed?
    - Can they be "reused?"
    - Can you have too many exchanges?

What are they and how do they work?

BUFFER CREDITS

What is Buffer-to-Buffer Credit?

- The greater the BB Credit….
  - A. The faster frames can be sent
  - B. The farther apart the two ports can be
  - C. The larger the frames can be
  - D. None of the above

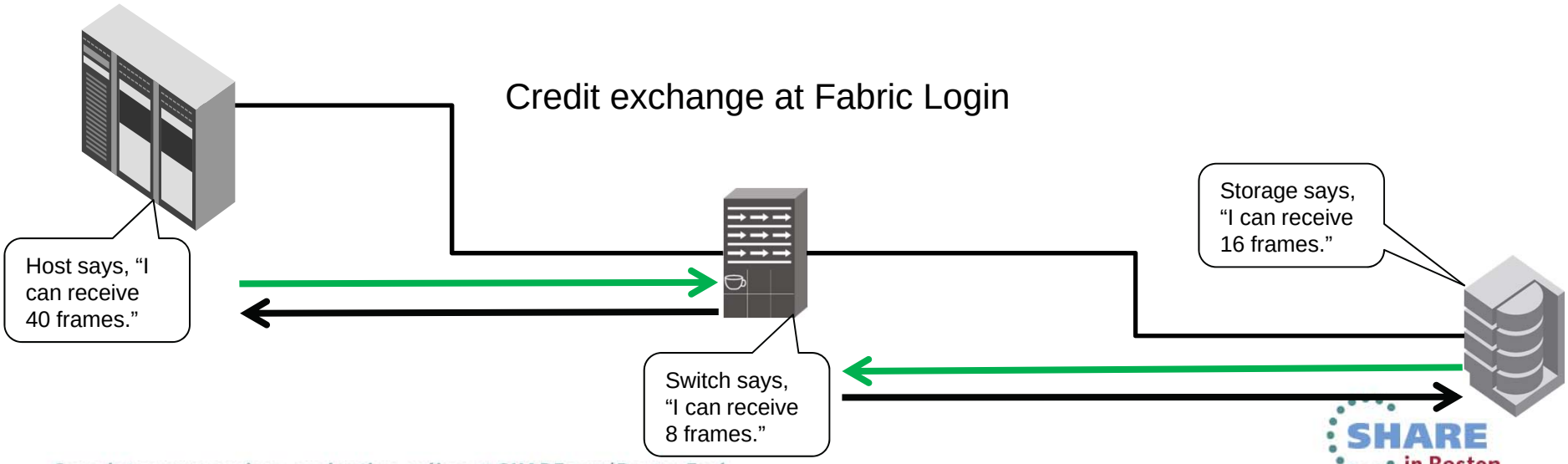
What is Buffer-to-Buffer Credit?

- The greater the BB Credit….
  - A.
  - B. The farther apart the two ports can be
  - C.
  - D.

Flow Control

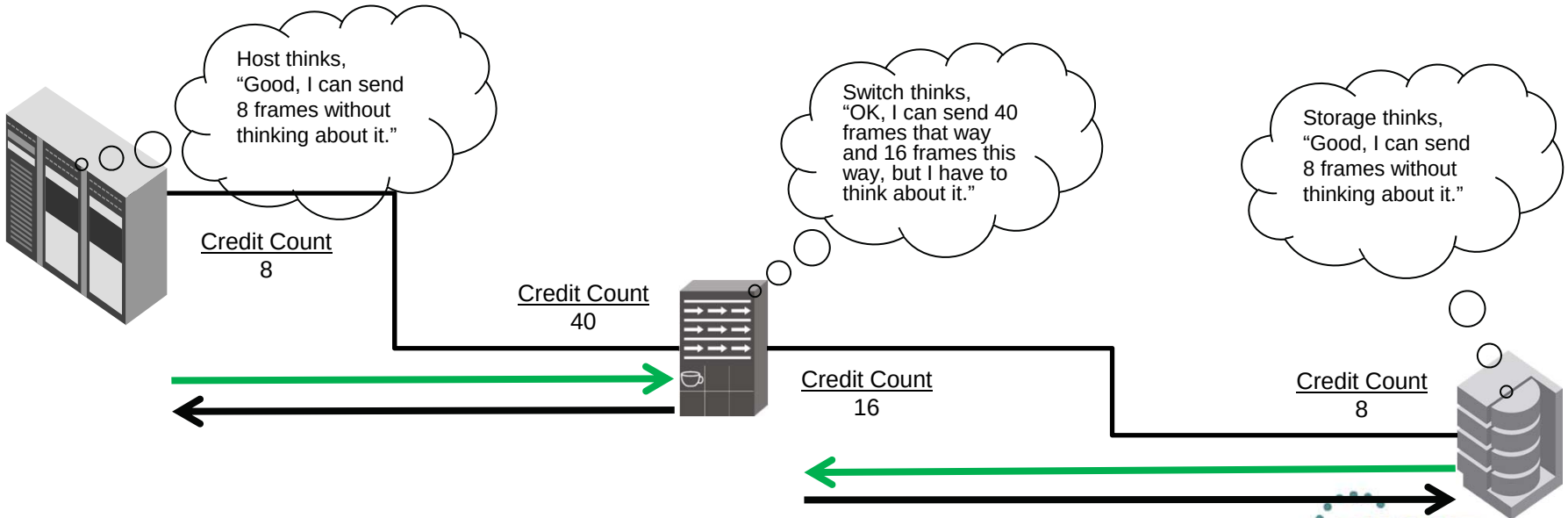
- Related to the devices' ability to receive and process frames
- Manages when frames are coming faster than they can be processed
- Dropped frames occur when frames are arriving too fast to be processed

- Frames can only be transmitted when the receiver is ready
- Credit establishment communicates the number of frames a device can receive at a time
- The credit value is exchanged at login
- Transmission stops when credit runs out
- The receiver indicates when it is ready to receive more frames



Buffer Credit

- At initialization, the two ports establish credit
  - Each buffer credit corresponds to a frame (regardless of size)
- Each side can support different values
  - Credit Count
- If a port doesn't have credit, it can't send a frame
  - Credit Count has reached zero
- Mechanism limits frame drops

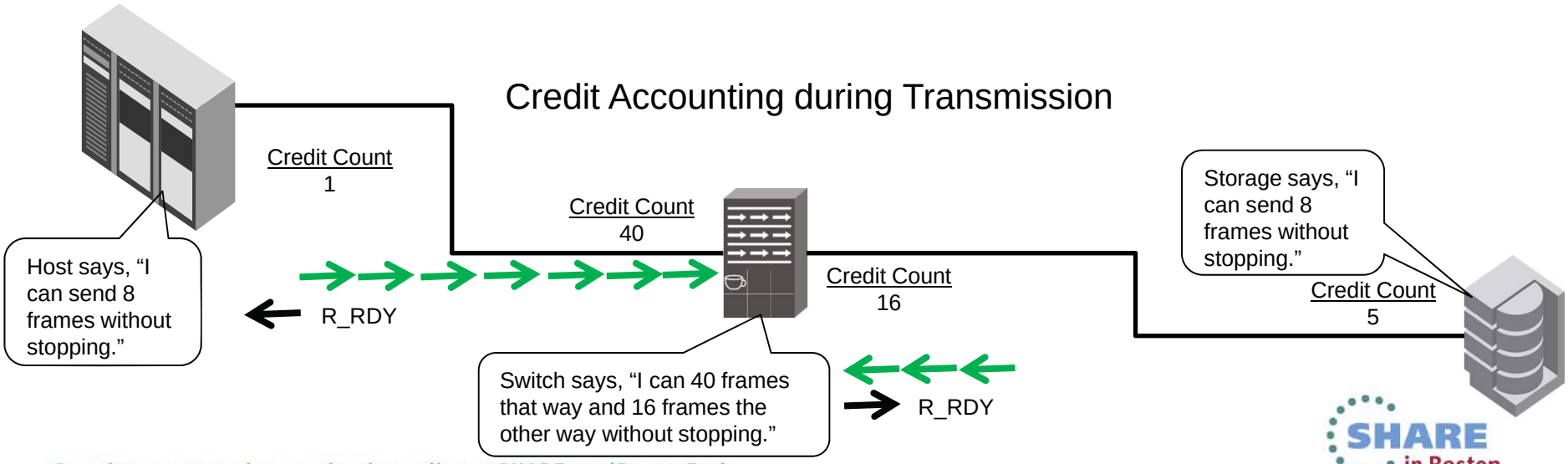


Credit accounting after Fabric Login

Receiver Ready (R_RDY)

- R_RDY
  - Used for link level flow control
  - Called buffer-to-buffer credit (BB Credit)
- R_RDY is not a frame
  - It is a “primitive” so it doesn’t consume a buffer

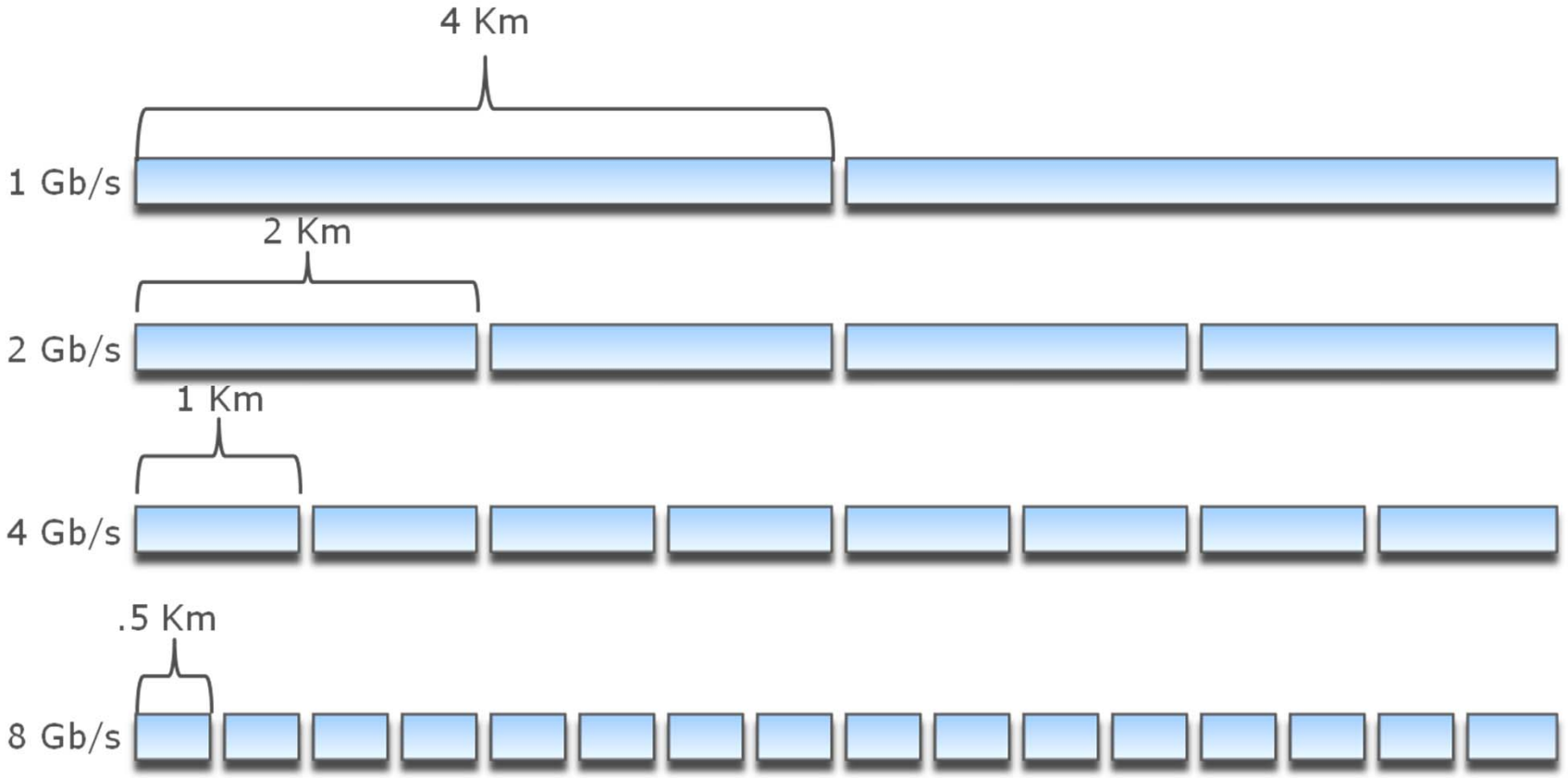
- Frame transmission
  - BB Credit is decremented
    - Once for each frame transmitted
  - When BB Credit = 0
    - Transmission stops
- Frame received
  - R_RDY is sent
  - Causes transmitter to increment BB Credit



Urban Legend:
Buffer Credits at Zero are a Problem

- Buffer credit determines DISTANCE
  - The distance two nodes can be apart and still maintain full link frame rate
- Buffer credit is the number of FRAME buffers
  - A port provides for it's NEAREST neighbor for RECEIVING frames
  - Does NOT have to be symmetrical
- Buffer credit is a FRAME count
  - Not a data SIZE
  - A 1 byte frame consumes 1 buffer credit
  - A 2K byte frame consumes 1 buffer credit
- Number of credits needed is determined by:
  - Raw Link Speed
  - Speed of light thru a fiber
  - Distance between two adjacent nodes

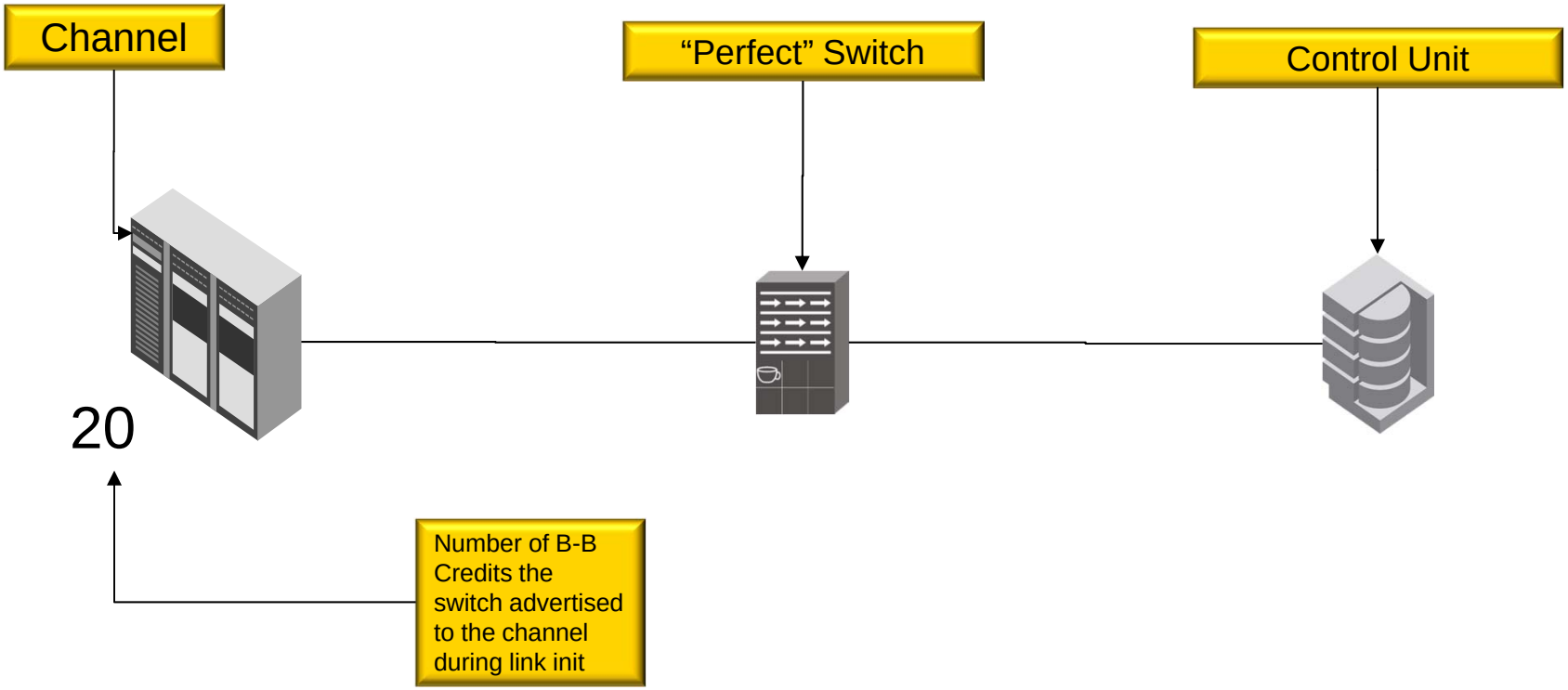
How Long is a Fibre Channel Frame?



Example: A full pipe

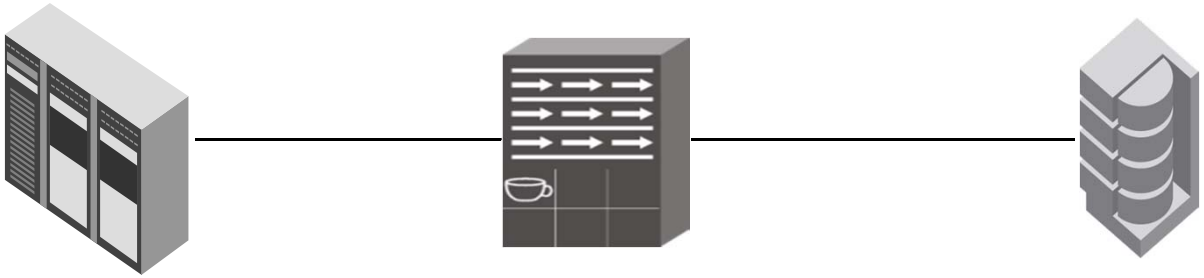
BUFFER CREDITS

Initial Conditions

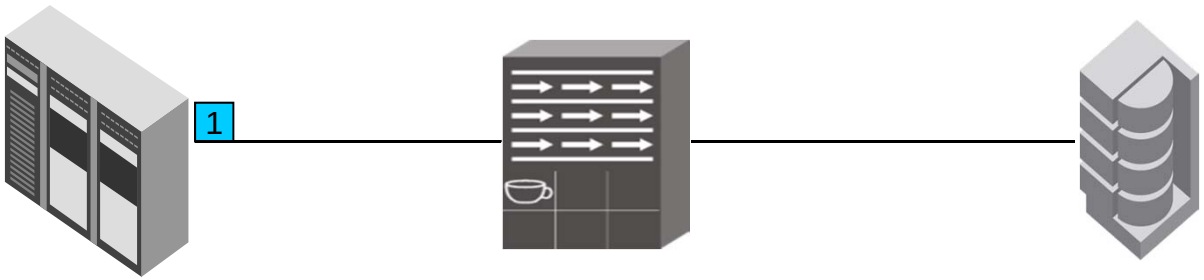


NOTE: In these animations, both the frames and R_RDY's are numbered. This is for illustrative purposes only. In reality, neither the frames nor the R_RDY's are numbered. The arrival of an R_RDY only informs the receiver that A frame has been forwarded, not WHICH frame has been forwarded.

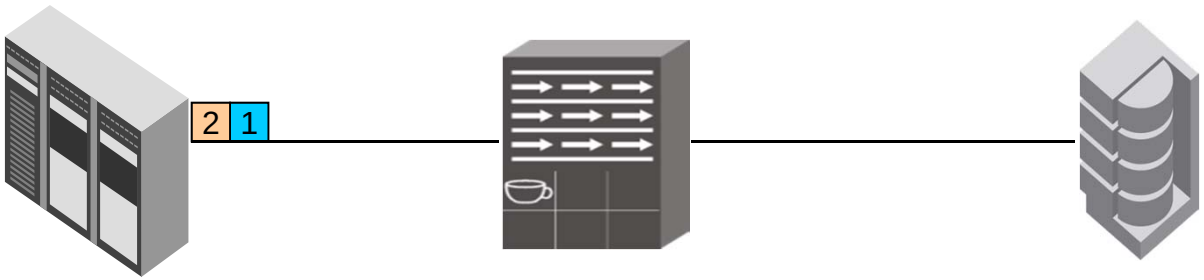
20

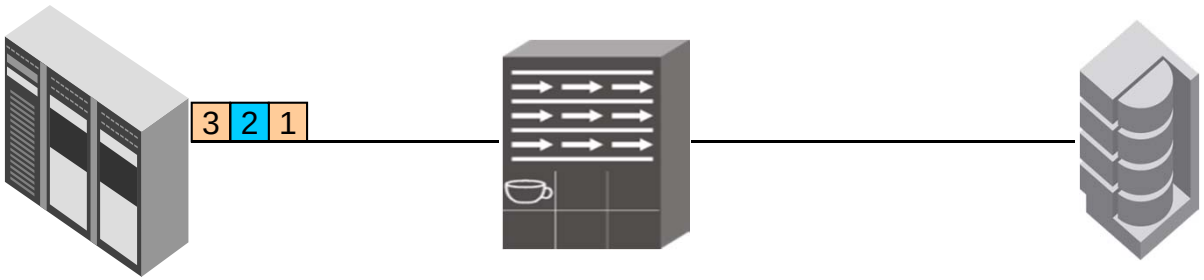


19

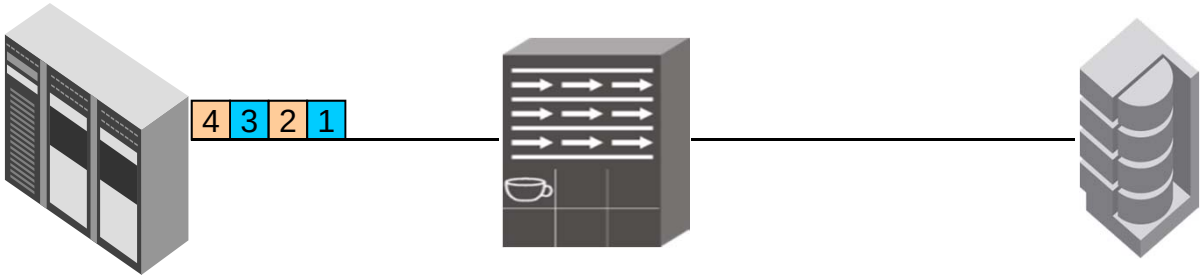


18

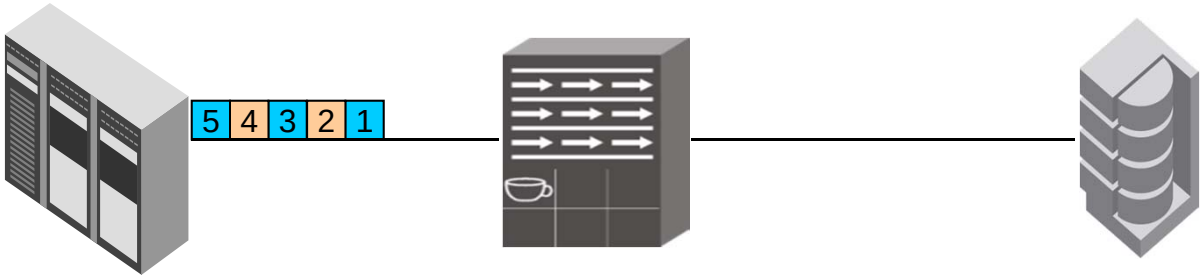



3
2
1

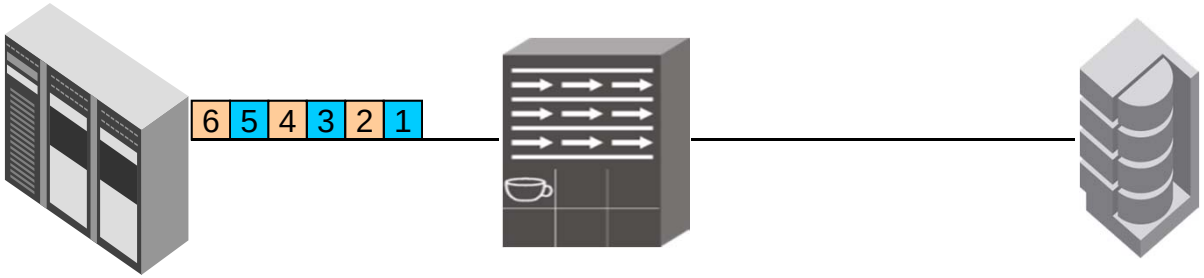
16



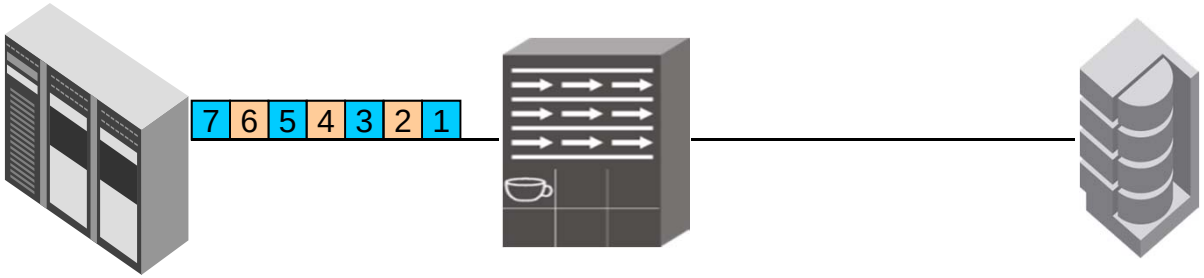
15



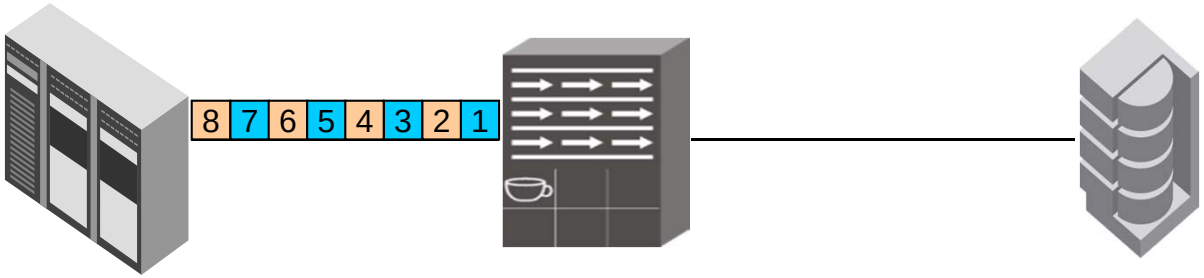
14

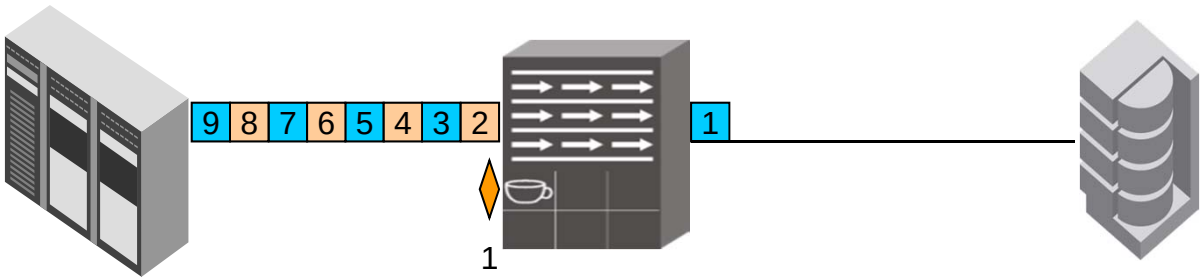


13

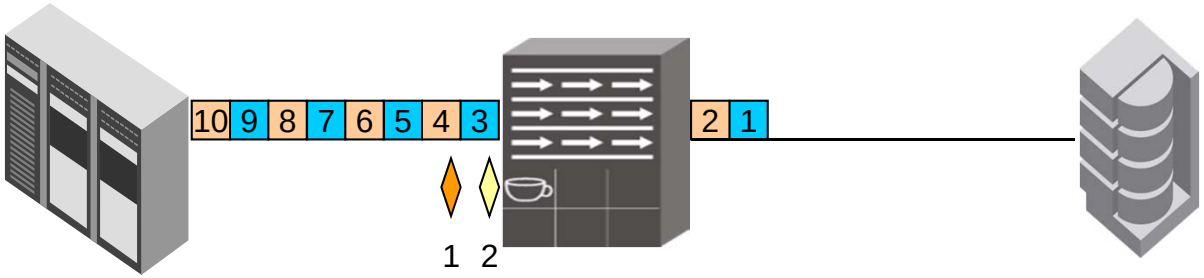


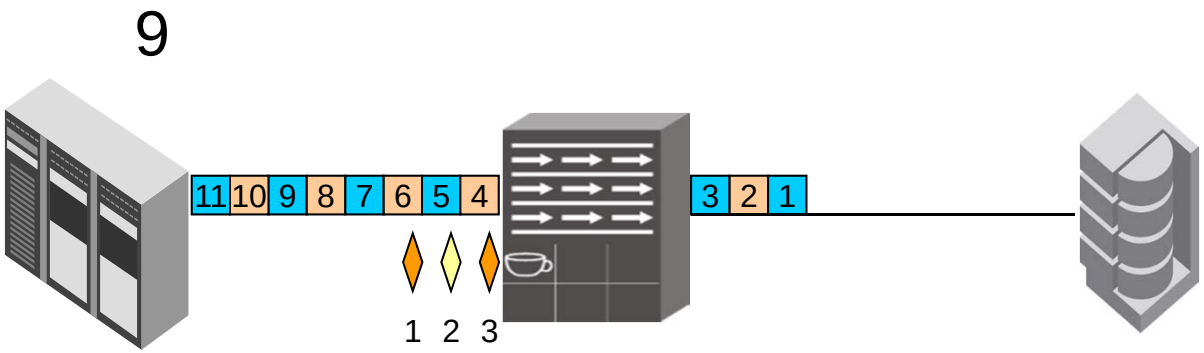
12



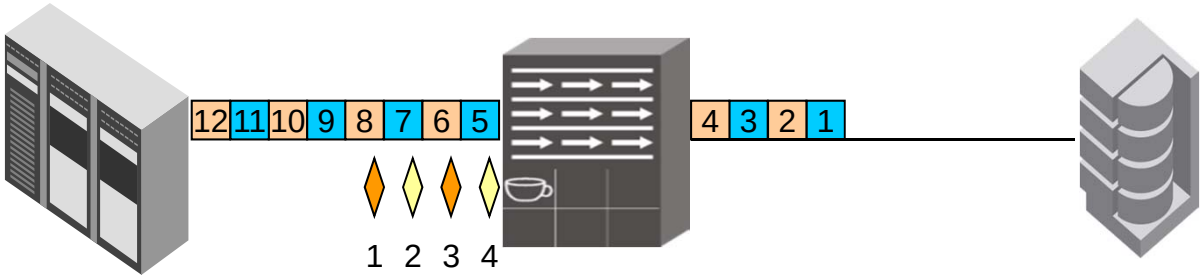

9
8
7
6
5
4
3
2
1
1

10

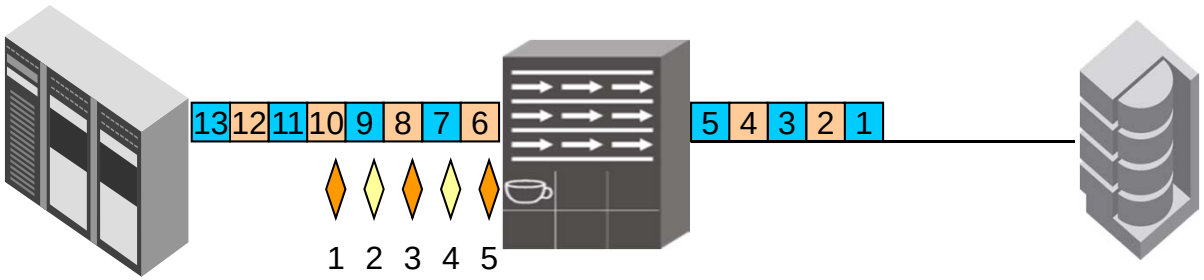



9
11 10 9 8 7 6 5 4
3 2 1
1 2 3

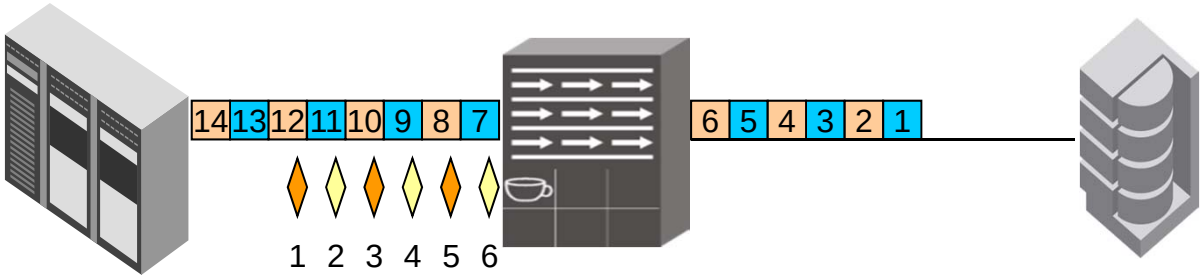
8



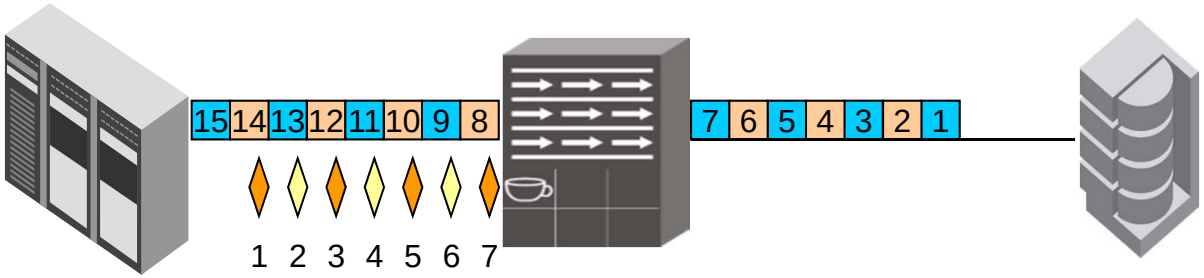
7



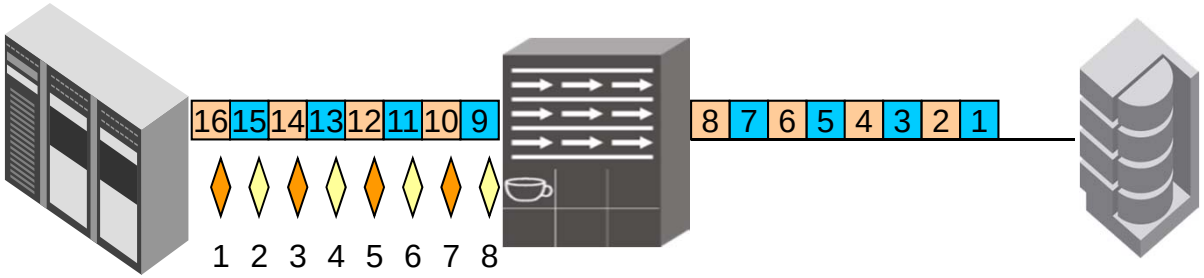
6



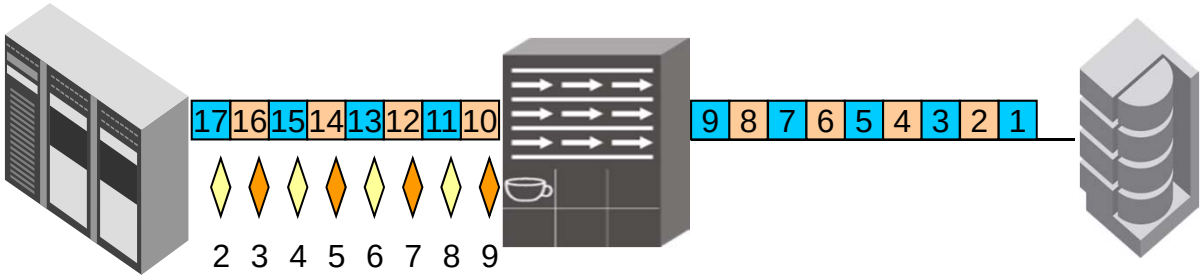
5



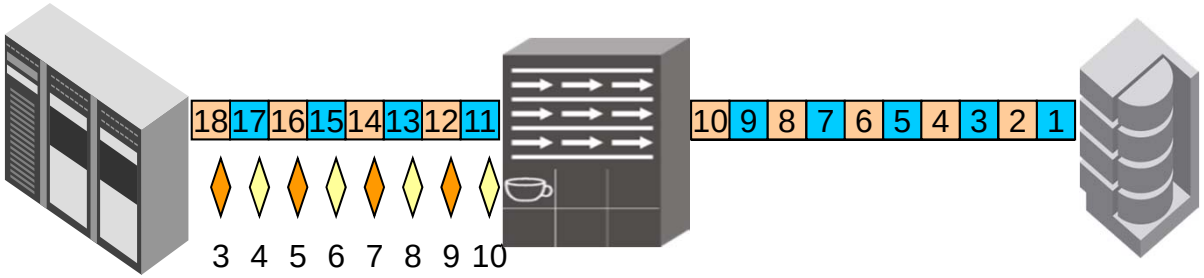
4



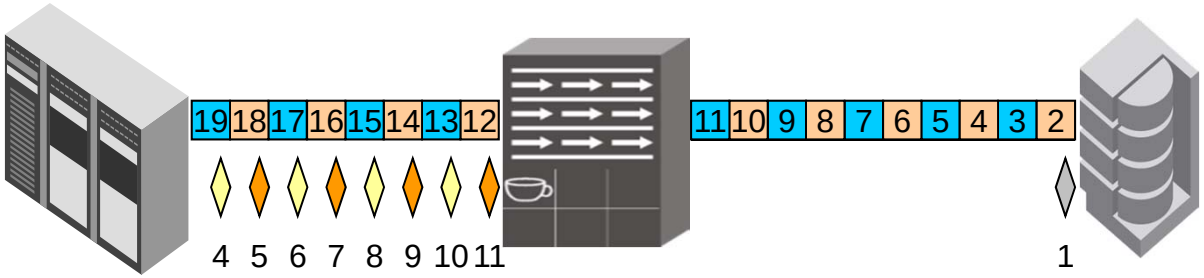
4



4



4



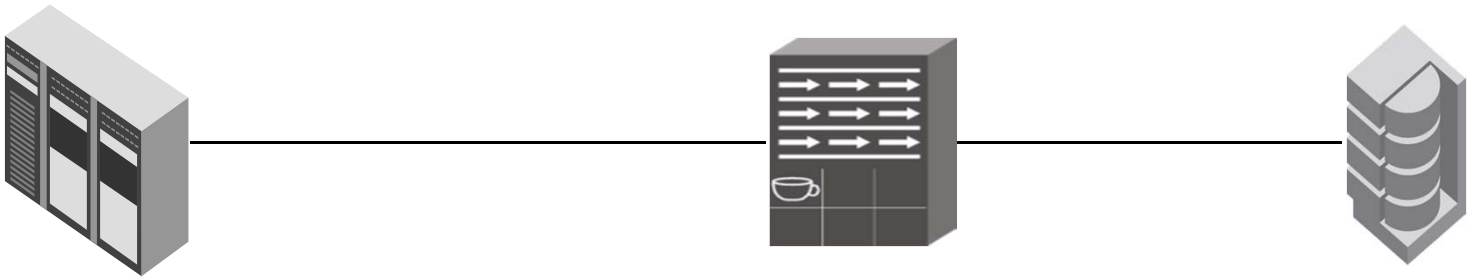
THIS PAGE INTENTIONALLY LEFT BLANK

Example: A not so full pipe

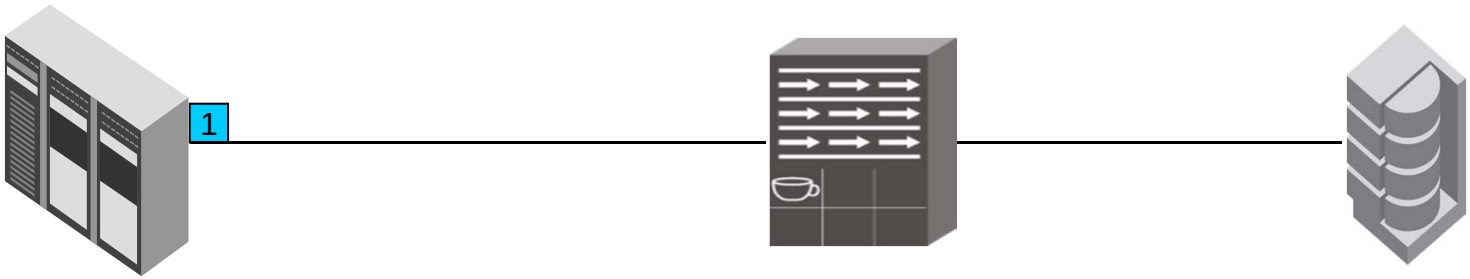
BUFFER CREDITS

Suppose the switch is too far away from the channel for the B-B credit it advertised to the channel

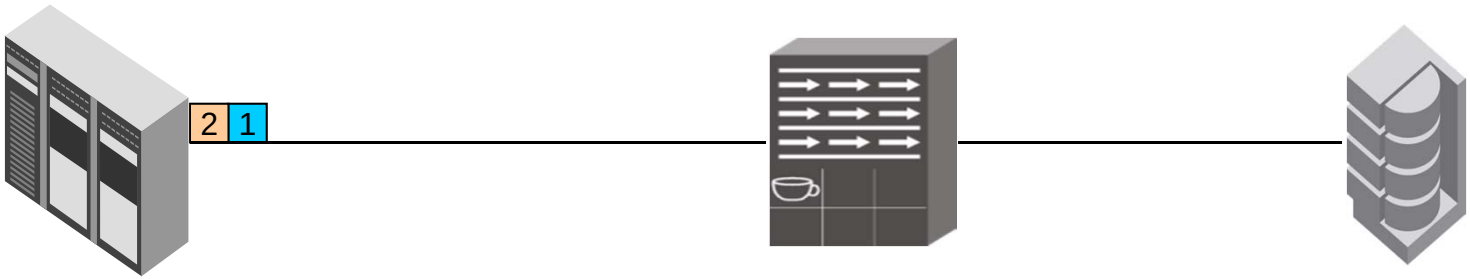
20

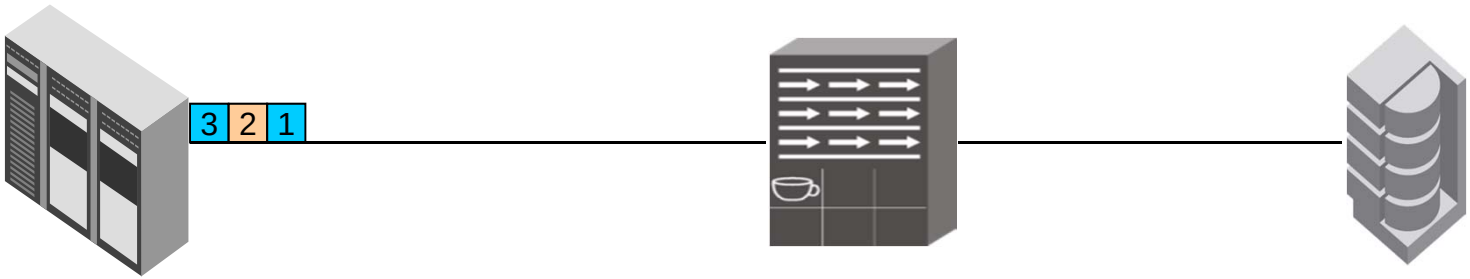


19

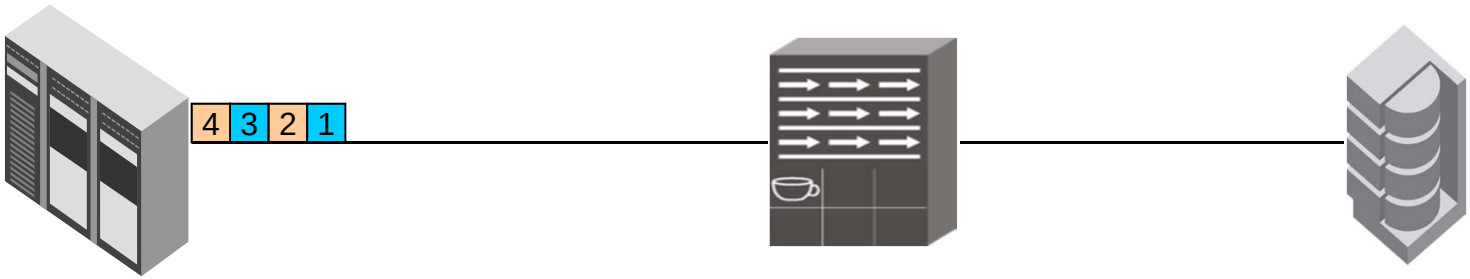


18

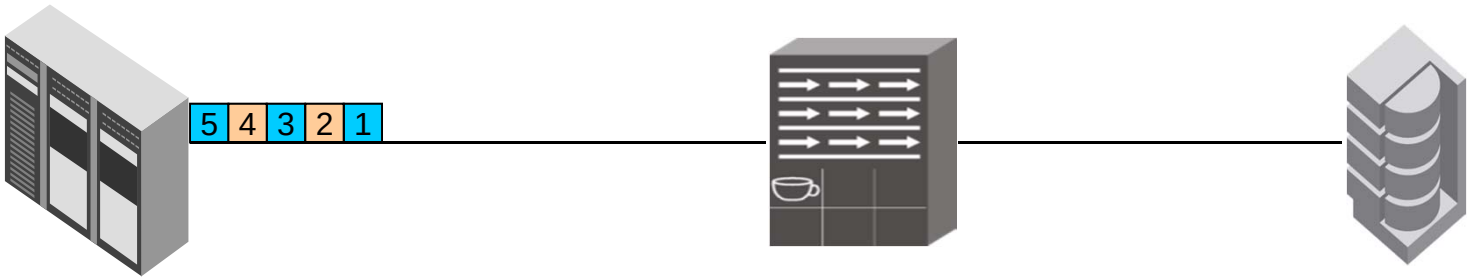



3
2
1

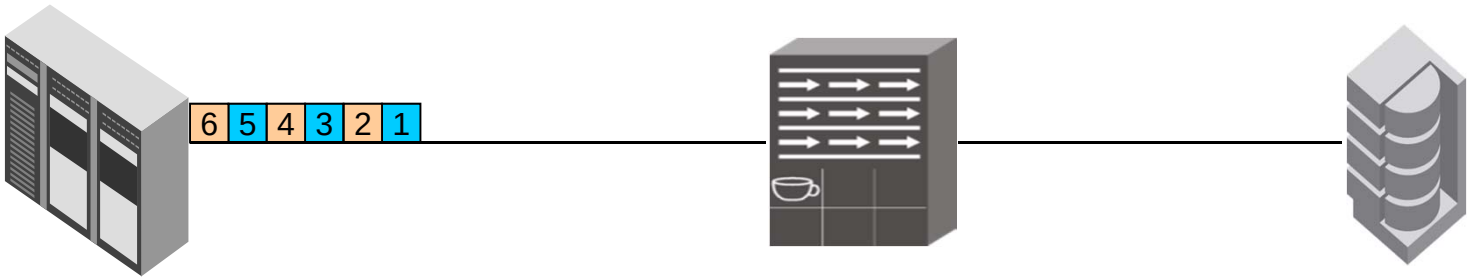
16



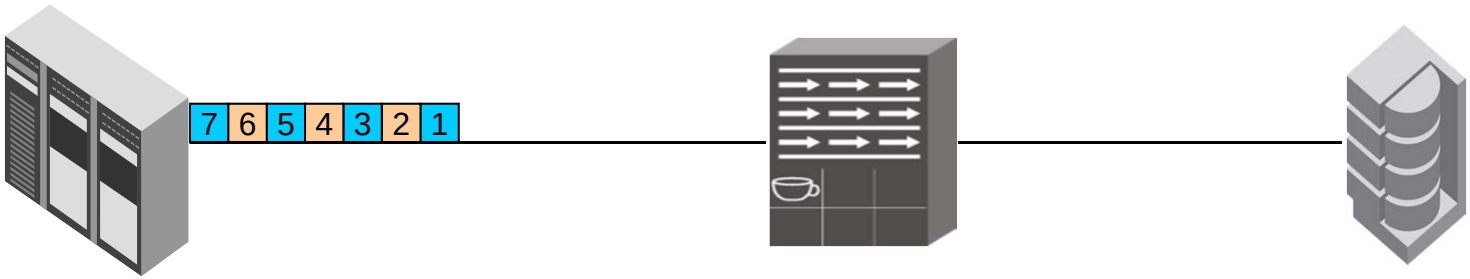
15



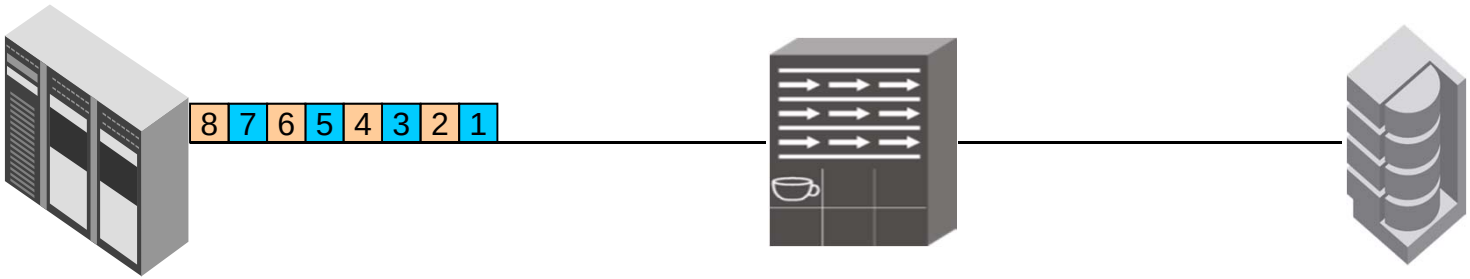
14



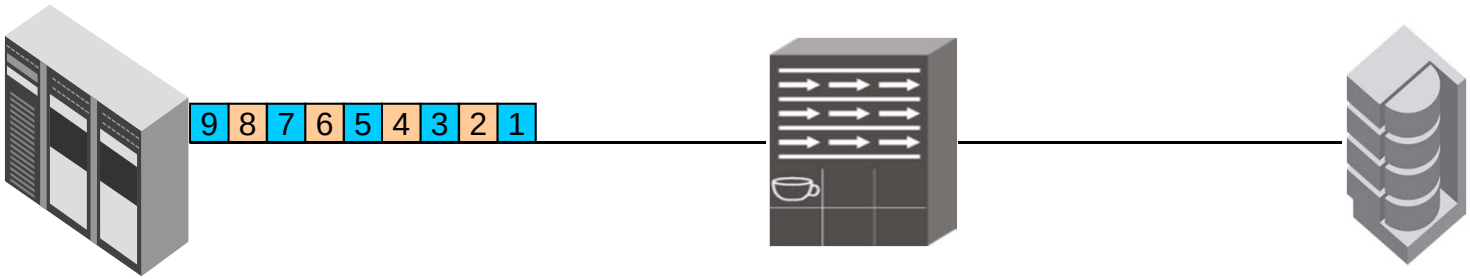
13



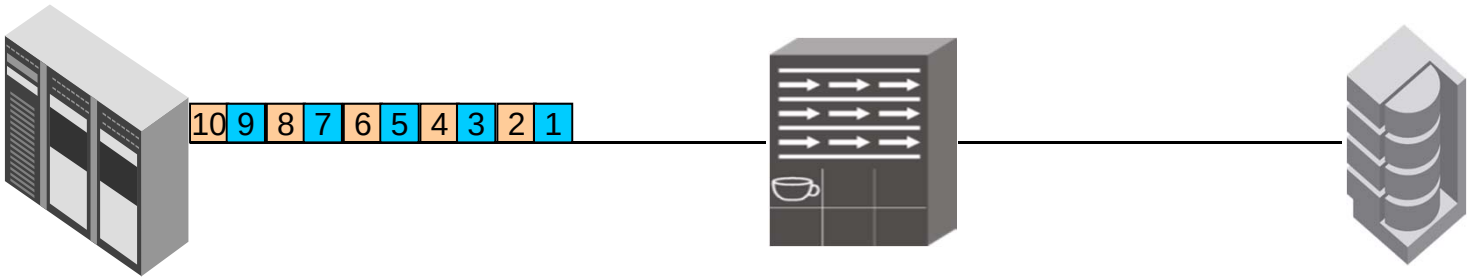
12



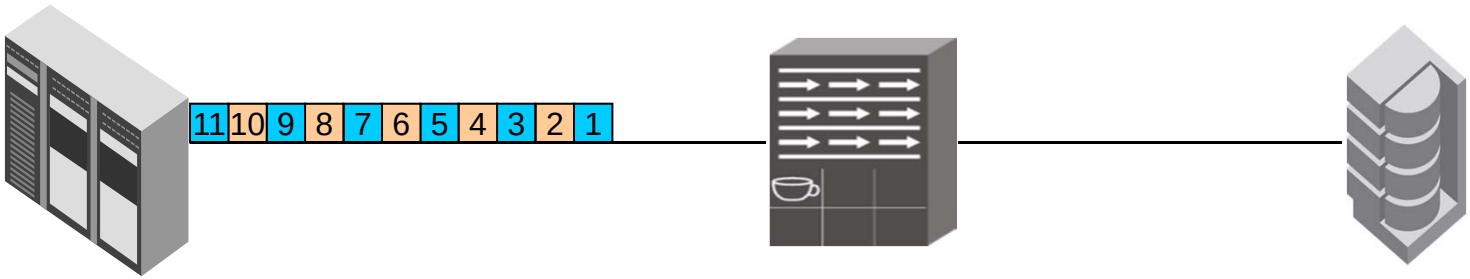
11



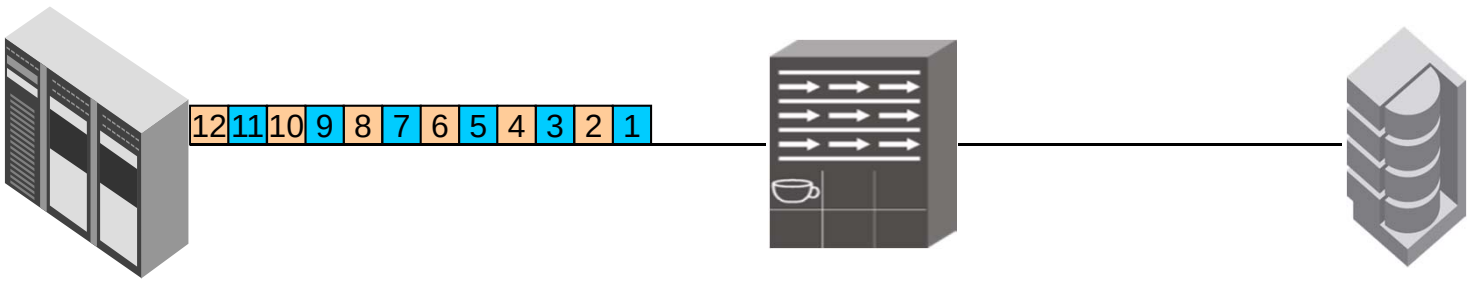
10

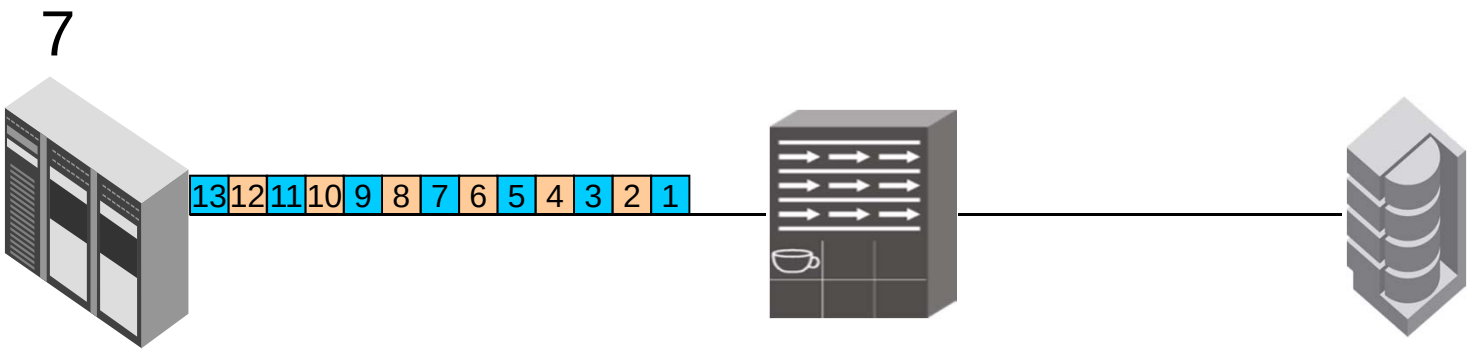


9

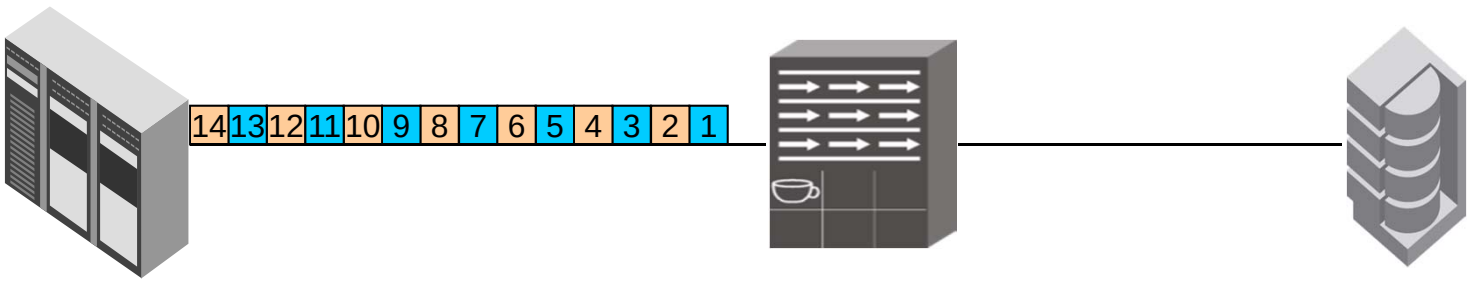


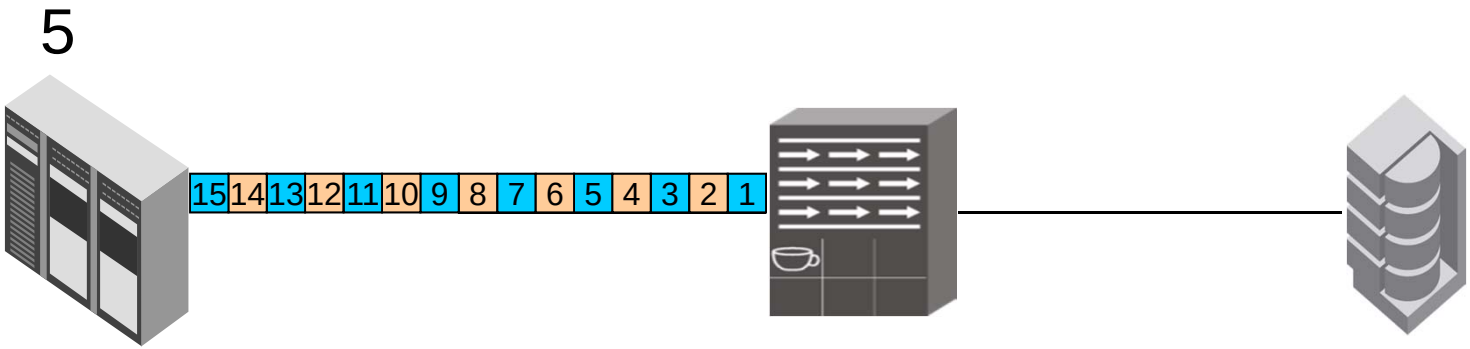
8

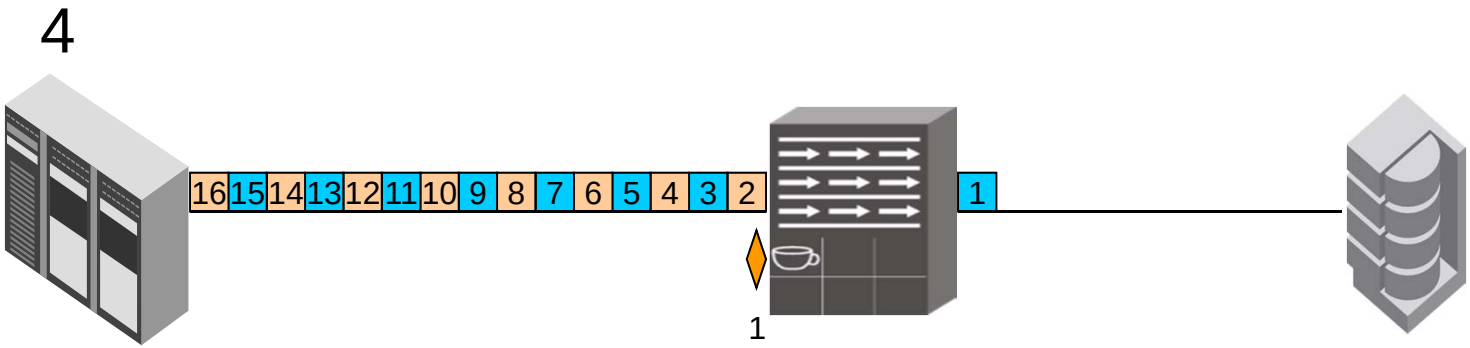


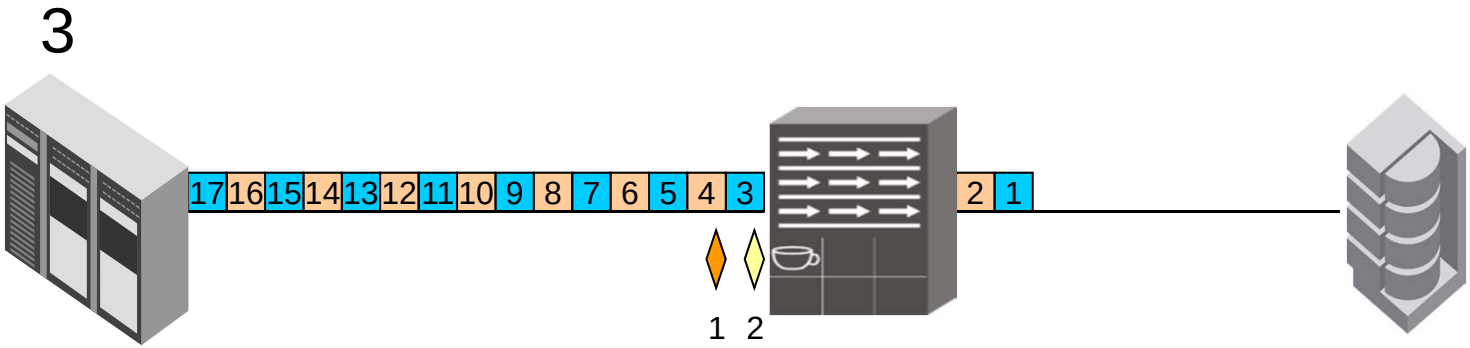

7
13 12 11 10 9 8 7 6 5 4 3 2 1

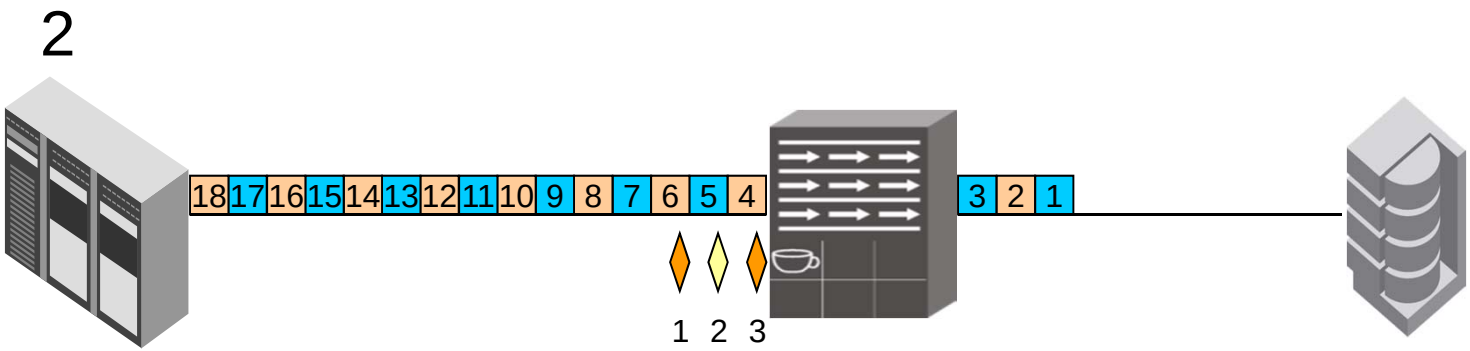
6

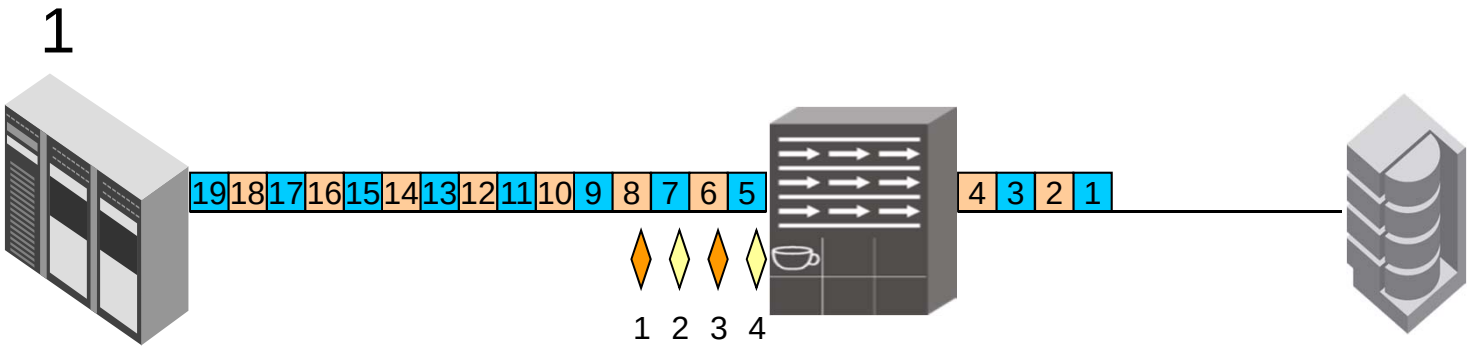


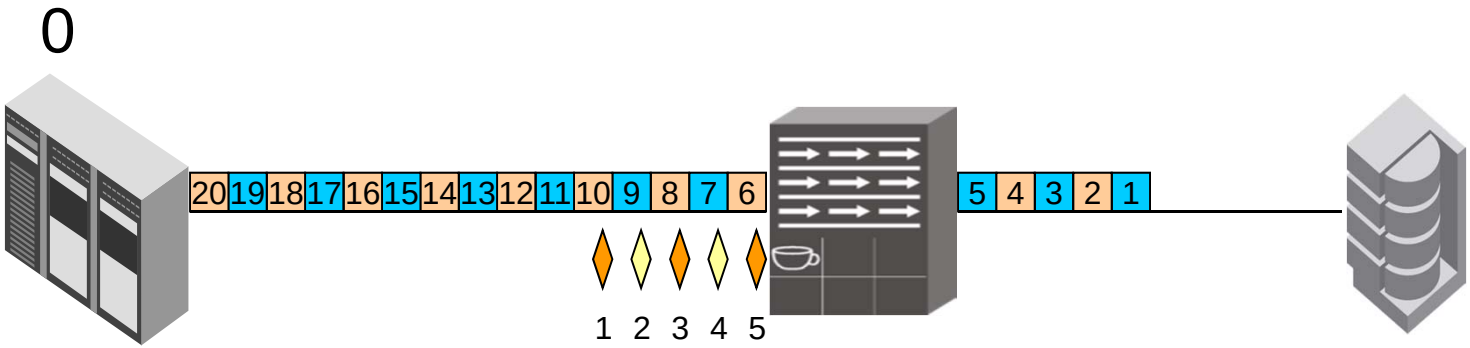

5
15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

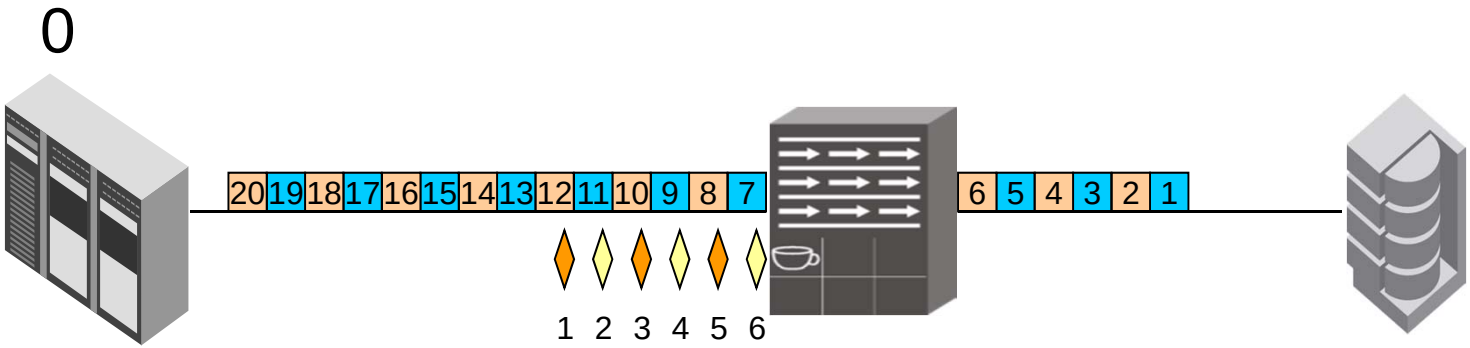

4
16 15 14 13 12 11 10 9 8 7 6 5 4 3 2
1
1

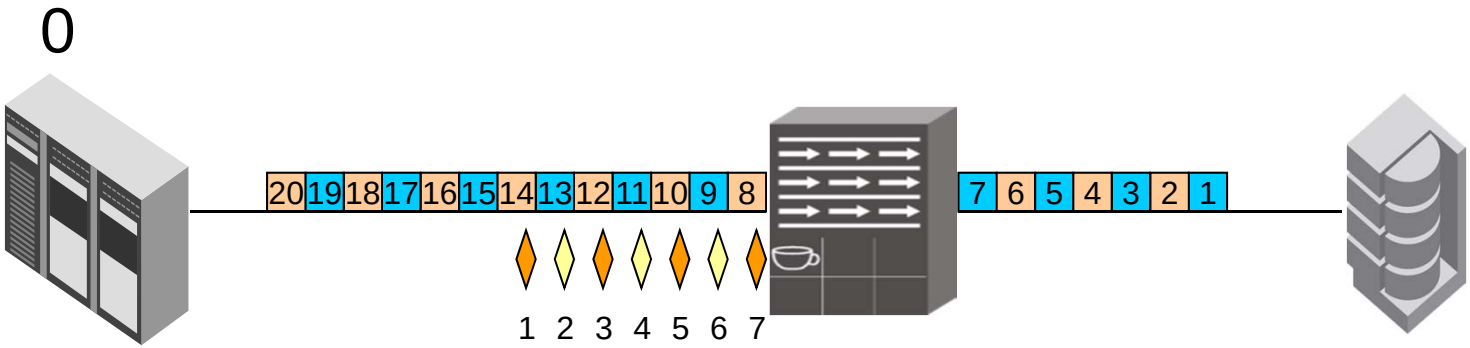

3
17 16 15 14 13 12 11 10 9 8 7 6 5 4 3
2 1
1 2

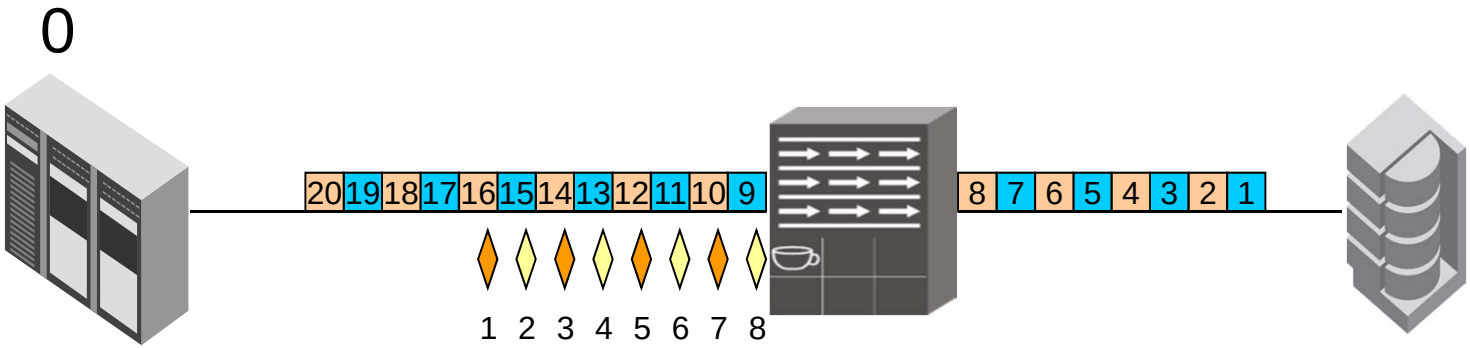

2
18 17 16 15 14 13 12 11 10 9 8 7 6 5 4
3 2 1
1 2 3

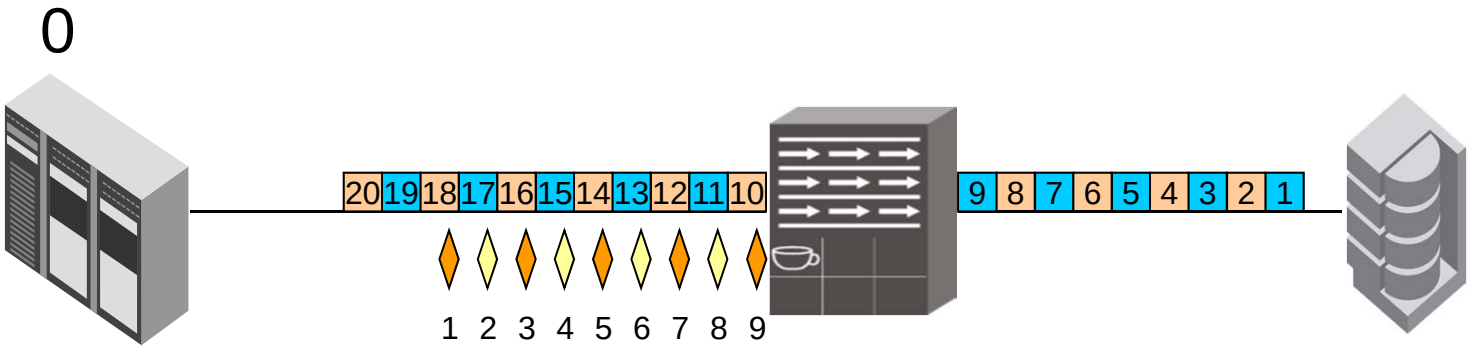

1
19 18 17 16 15 14 13 12 11 10 9 8 7 6 5
4 3 2 1
1 2 3 4

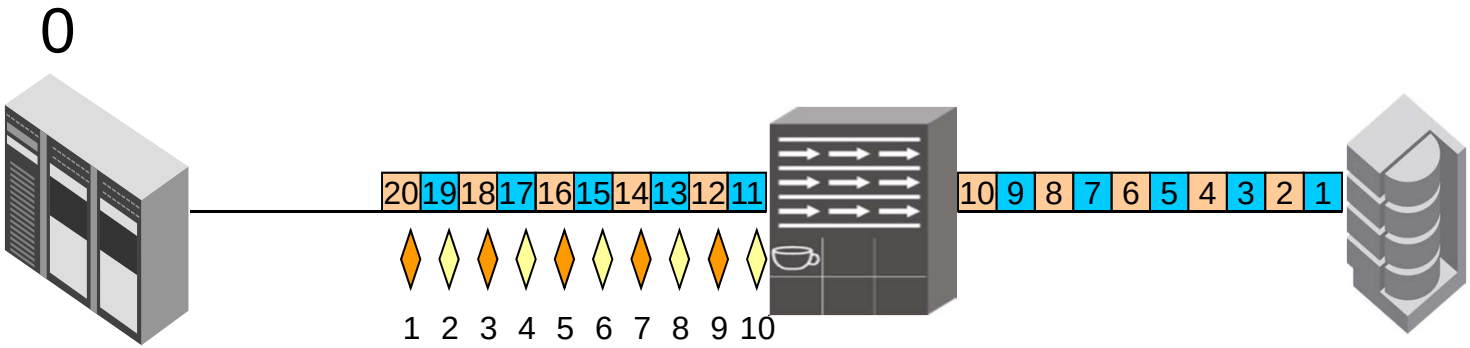

0
20 19 18 17 16 15 14 13 12 11 10 9 8 7 6
5 4 3 2 1
1 2 3 4 5

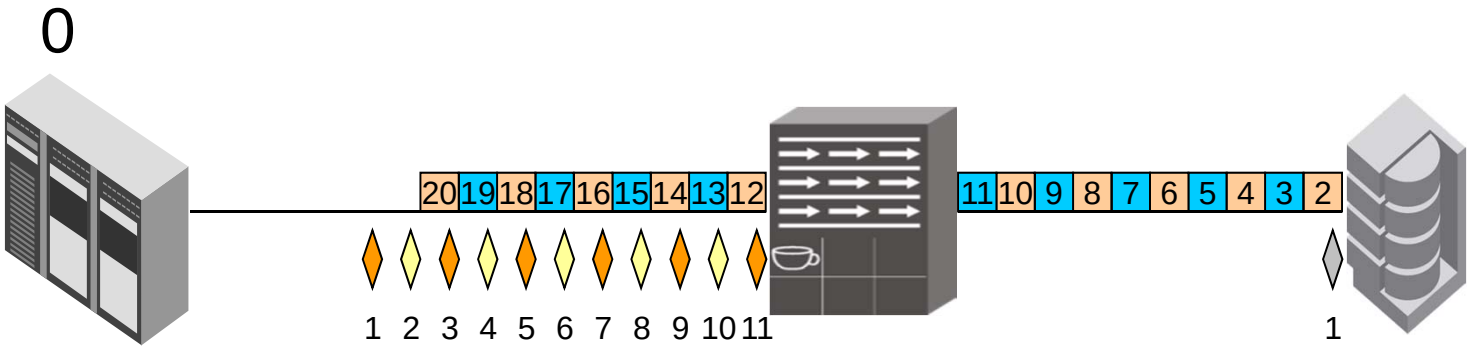

0
20 19 18 17 16 15 14 13 12 11 10 9 8 7
6 5 4 3 2 1
1 2 3 4 5 6

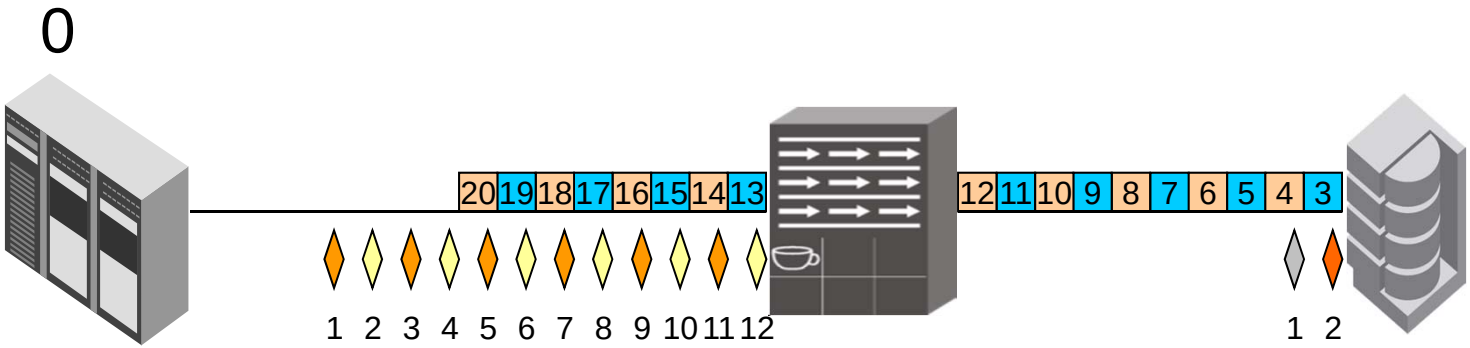

0
20 19 18 17 16 15 14 13 12 11 10 9 8
7 6 5 4 3 2 1
1 2 3 4 5 6 7

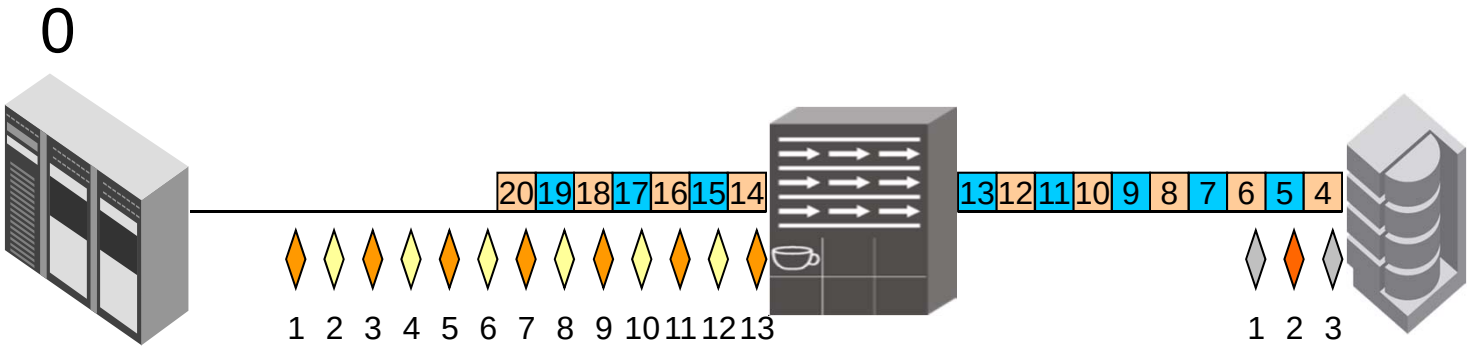

0
20 19 18 17 16 15 14 13 12 11 10 9
8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8

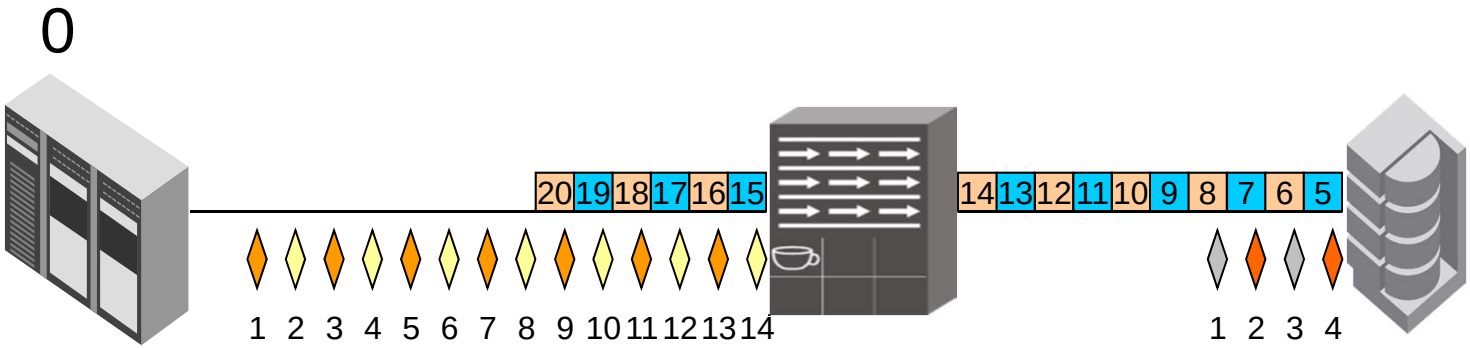

0
20 19 18 17 16 15 14 13 12 11 10
9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9

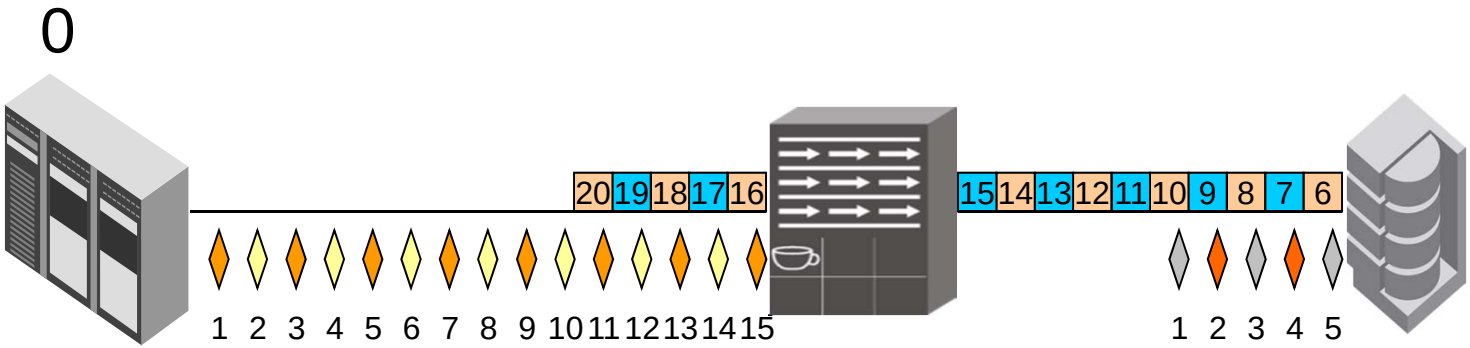

0
20 19 18 17 16 15 14 13 12 11
10 9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10

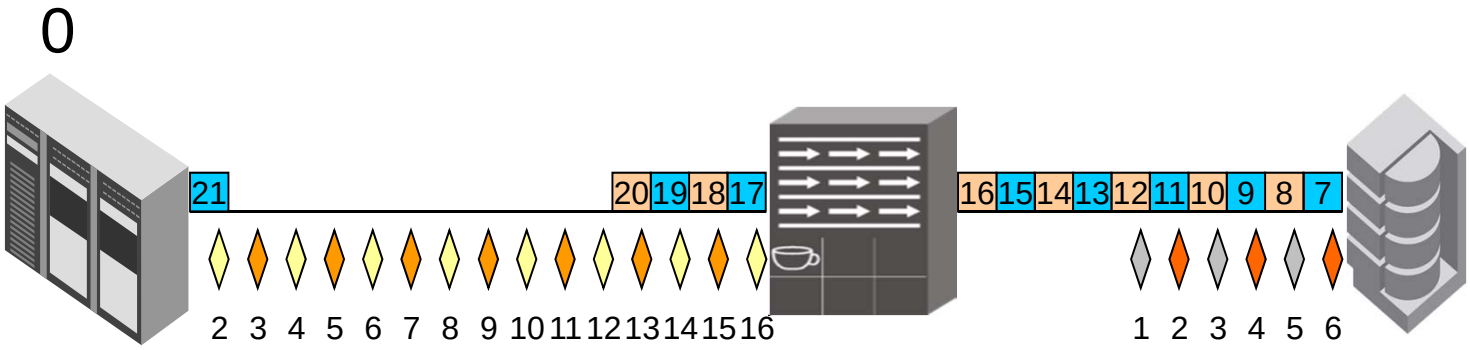

0
20 19 18 17 16 15 14 13 12
11 10 9 8 7 6 5 4 3 2
1 2 3 4 5 6 7 8 9 10 11
1

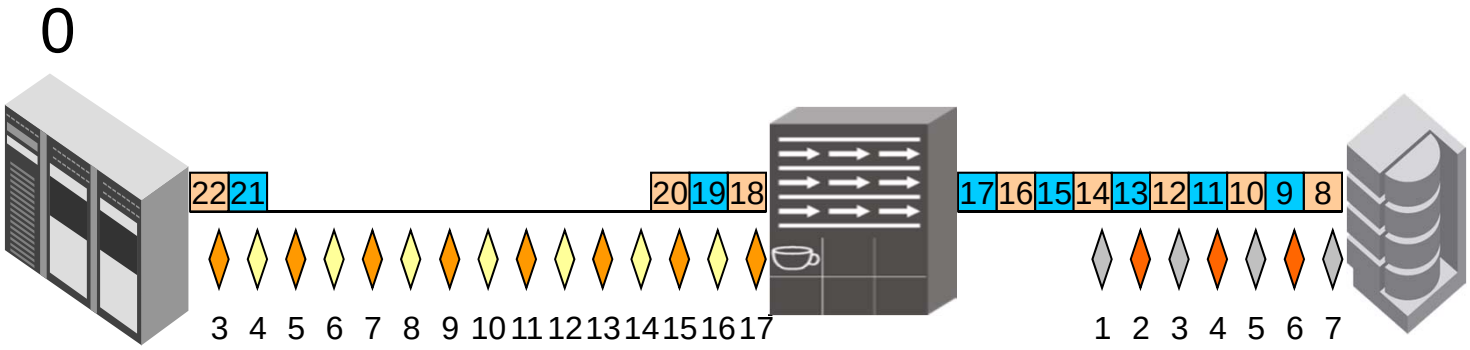

0
20 19 18 17 16 15 14 13
12 11 10 9 8 7 6 5 4 3
1 2 3 4 5 6 7 8 9 10 11 12
1 2

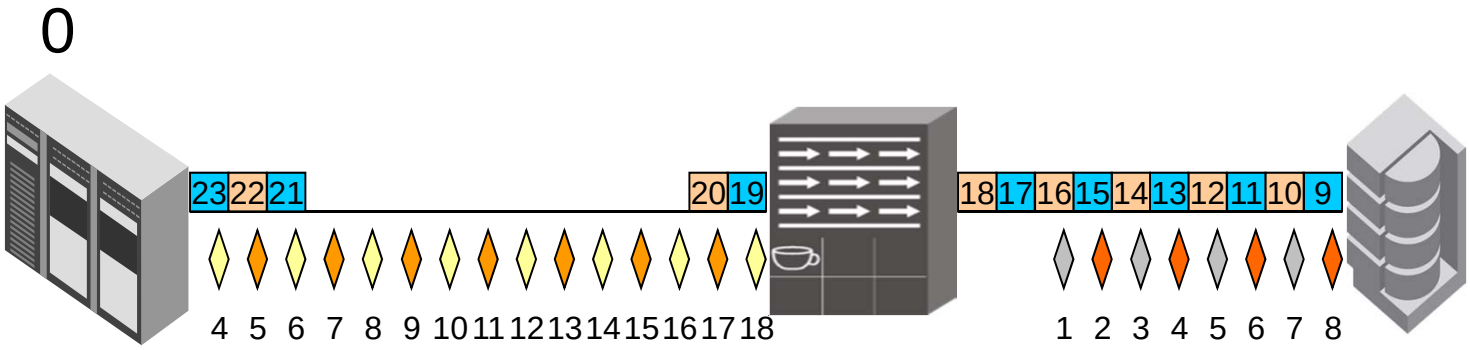

0
20 19 18 17 16 15 14
13 12 11 10 9 8 7 6 5 4
1 2 3 4 5 6 7 8 9 10 11 12 13
1 2 3

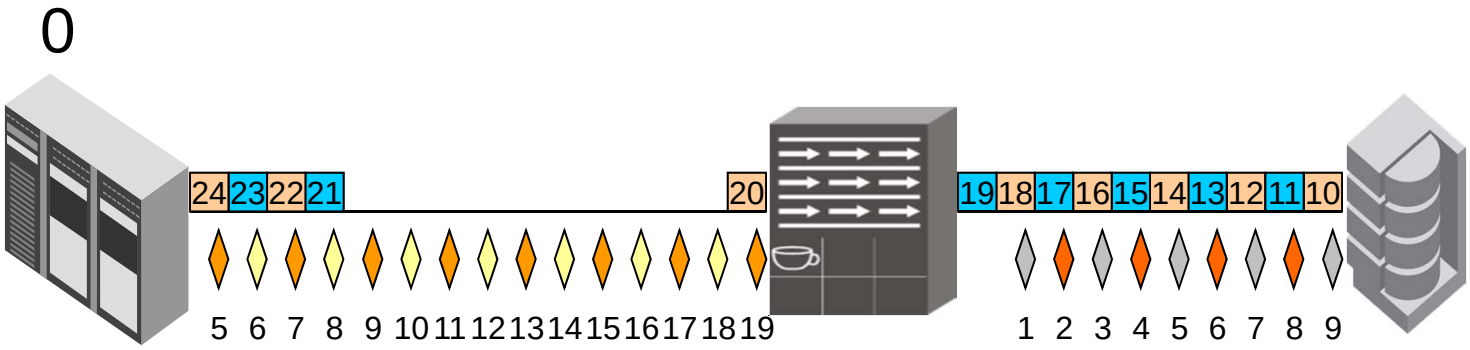

0
20 19 18 17 16 15
14 13 12 11 10 9 8 7 6 5
1 2 3 4 5 6 7 8 9 10 11 12 13 14
1 2 3 4

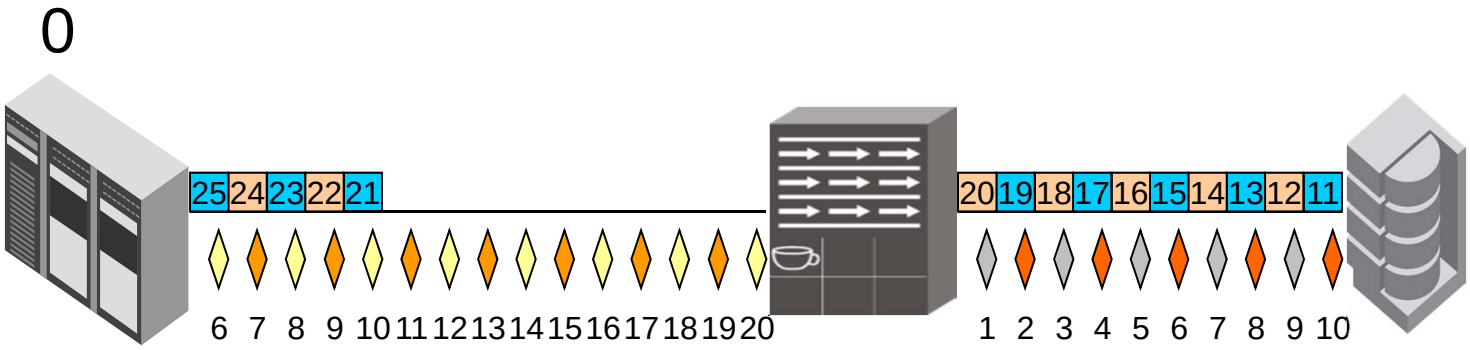

0
20 19 18 17 16
15 14 13 12 11 10 9 8 7 6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
1 2 3 4 5

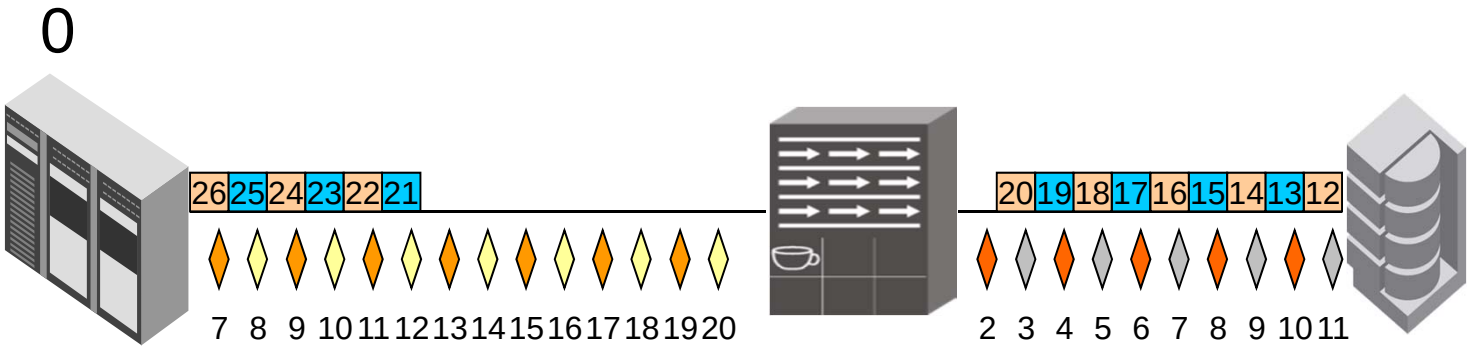

0
21
20 19 18 17
16 15 14 13 12 11 10 9 8 7
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
1 2 3 4 5 6

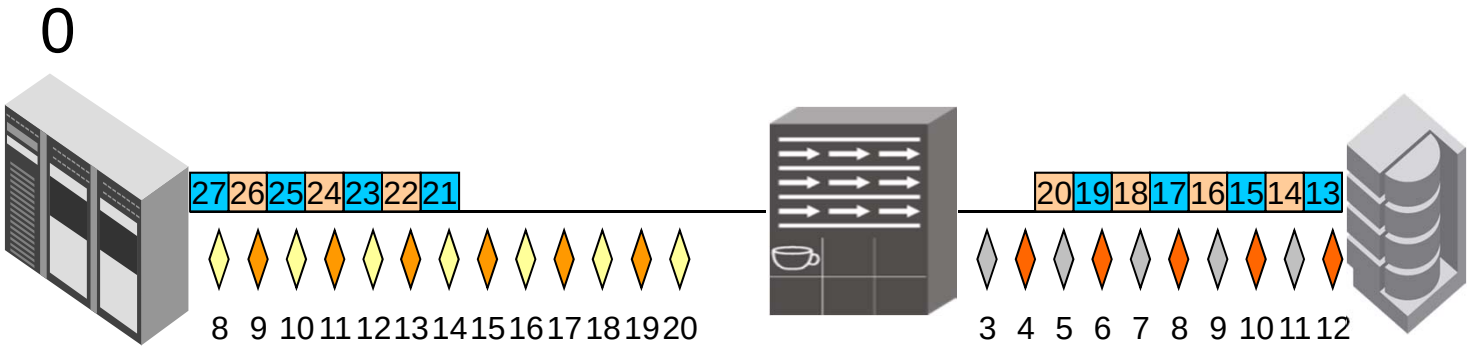

0
22 21
20 19 18
17 16 15 14 13 12 11 10 9 8
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
1 2 3 4 5 6 7

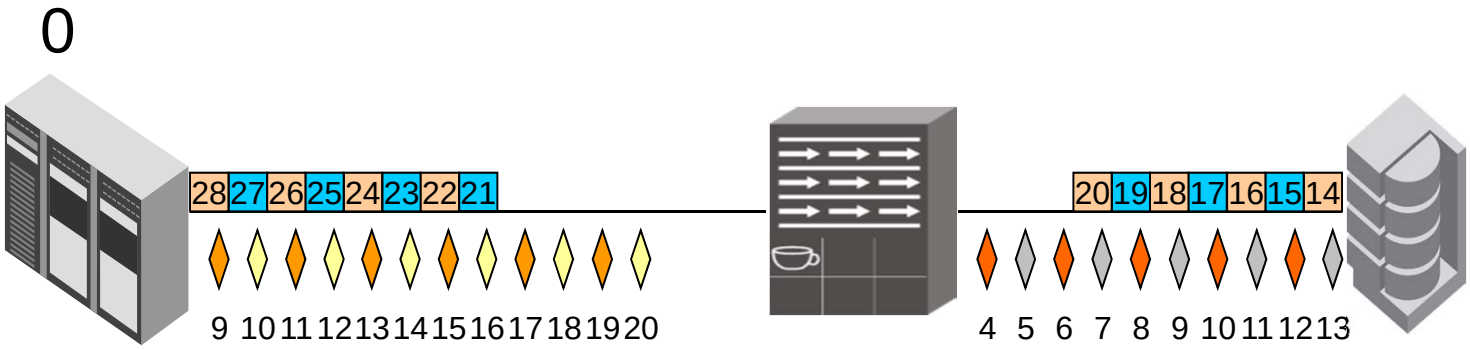

0
23 22 21
20 19
18 17 16 15 14 13 12 11 10 9
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
1 2 3 4 5 6 7 8

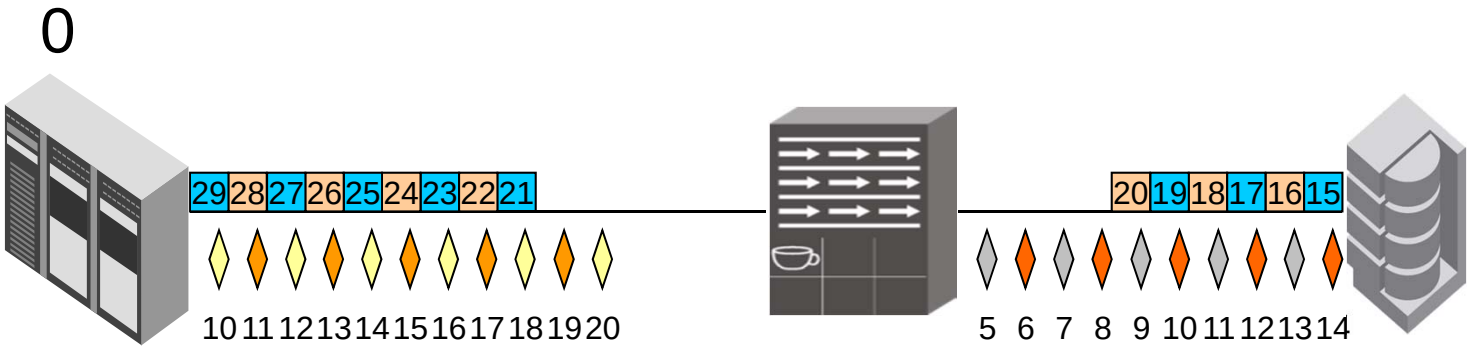

0
24 23 22 21
20
19 18 17 16 15 14 13 12 11 10
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
1 2 3 4 5 6 7 8 9

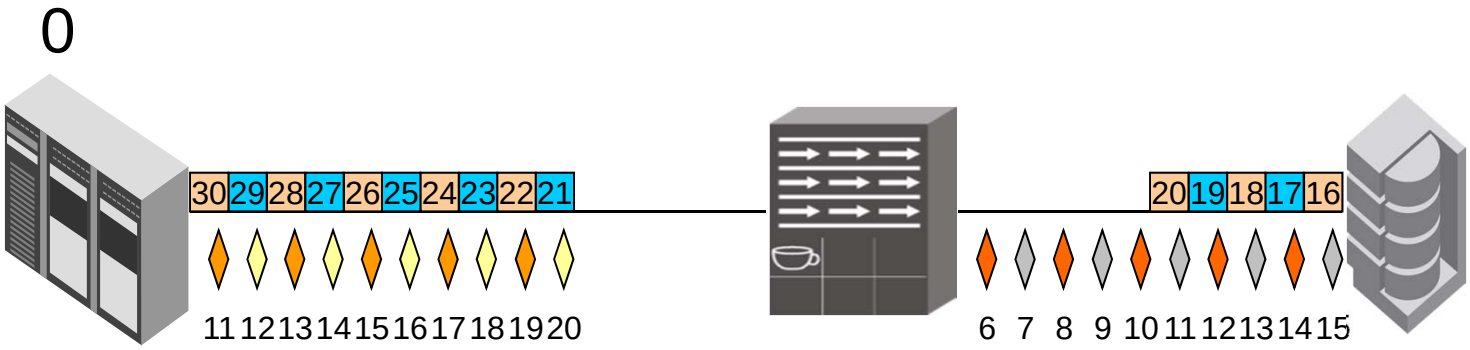

0
25 24 23 22 21
20 19 18 17 16 15 14 13 12 11
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
1 2 3 4 5 6 7 8 9 10

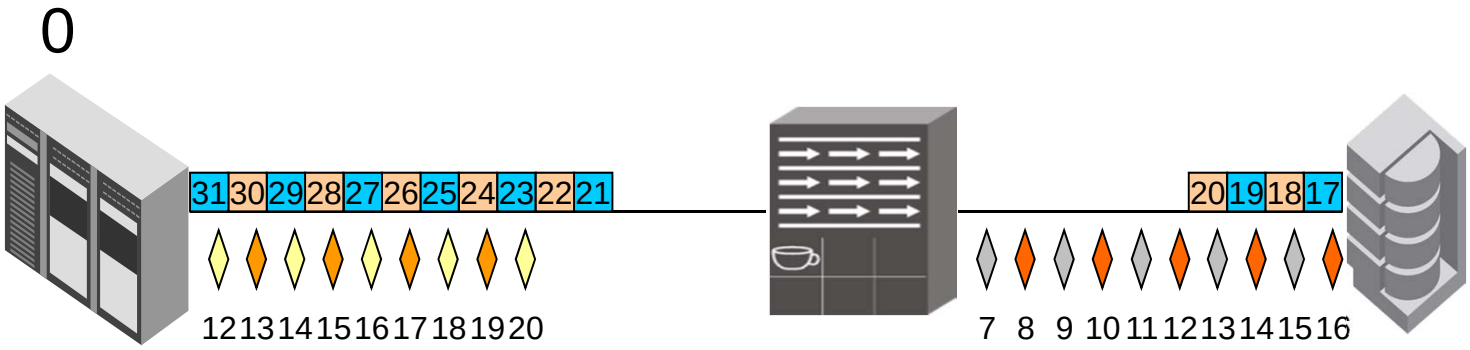

0
26 25 24 23 22 21
7 8 9 10 11 12 13 14 15 16 17 18 19 20
20 19 18 17 16 15 14 13 12
2 3 4 5 6 7 8 9 10 11

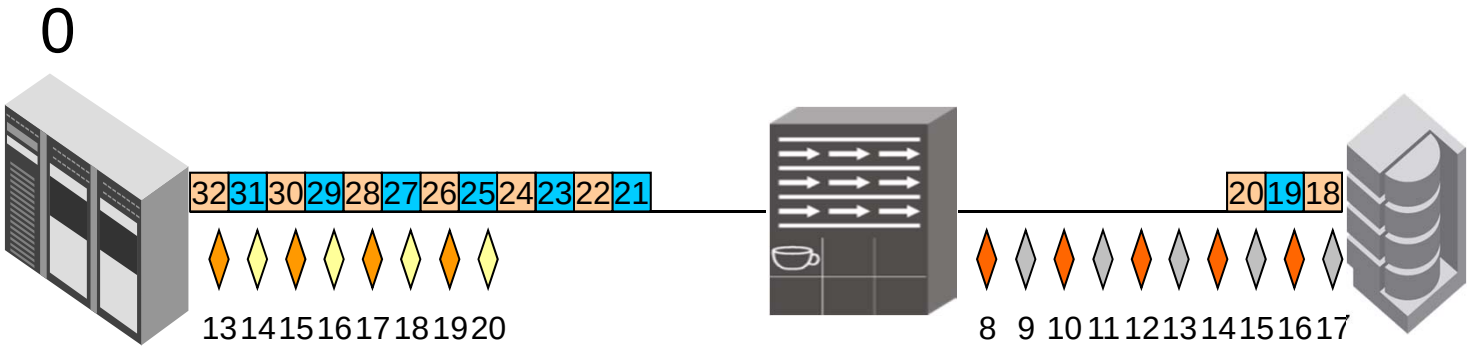

0
27 26 25 24 23 22 21
8 9 10 11 12 13 14 15 16 17 18 19 20
20 19 18 17 16 15 14 13
3 4 5 6 7 8 9 10 11 12

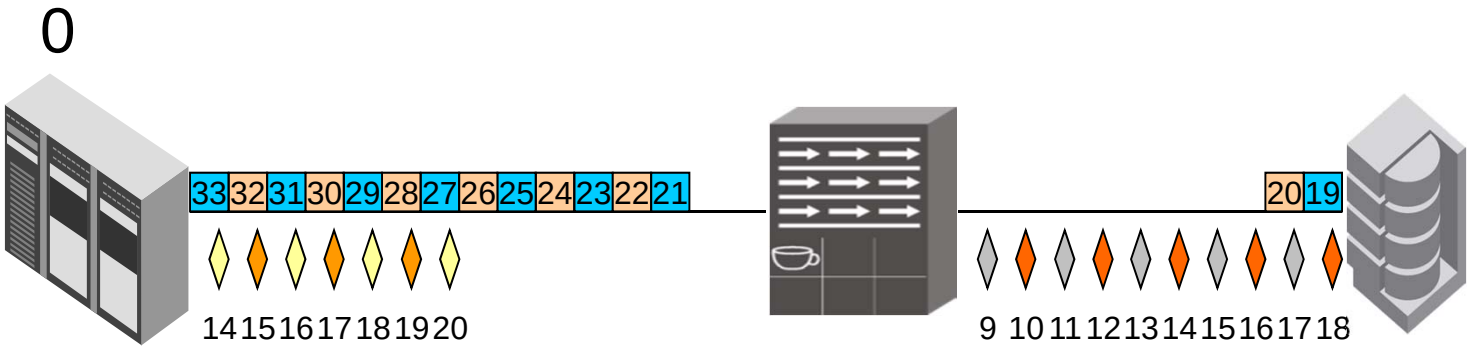

0
28 27 26 25 24 23 22 21
9 10 11 12 13 14 15 16 17 18 19 20
20 19 18 17 16 15 14
4 5 6 7 8 9 10 11 12 13

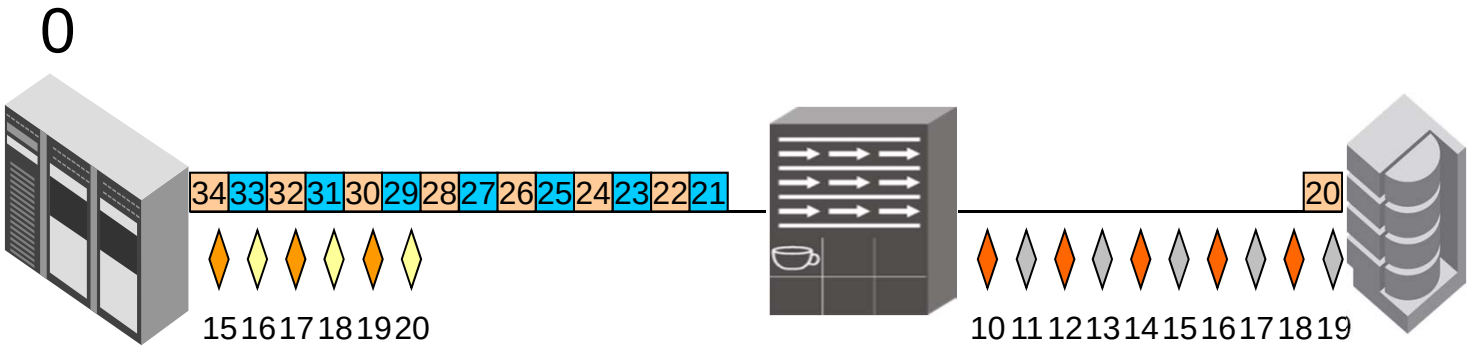

0
29 28 27 26 25 24 23 22 21
10 11 12 13 14 15 16 17 18 19 20
20 19 18 17 16 15
5 6 7 8 9 10 11 12 13 14

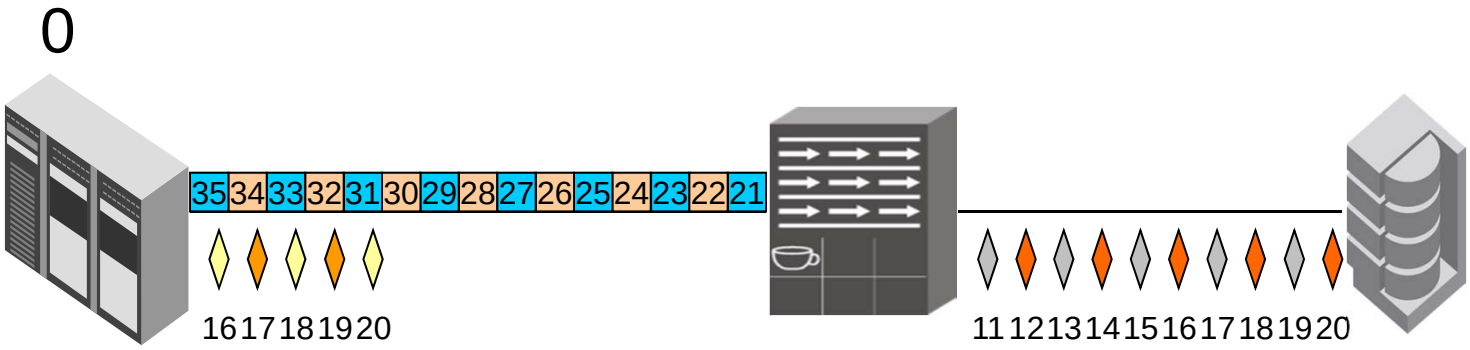

0
30 29 28 27 26 25 24 23 22 21
11 12 13 14 15 16 17 18 19 20
20 19 18 17 16
6 7 8 9 10 11 12 13 14 15

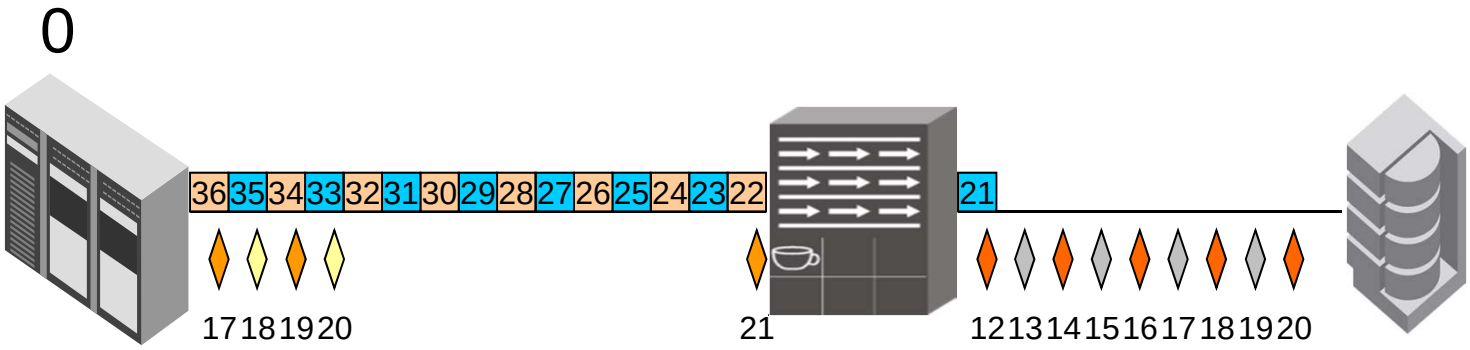

0
31 30 29 28 27 26 25 24 23 22 21
12 13 14 15 16 17 18 19 20
20 19 18 17
7 8 9 10 11 12 13 14 15 16

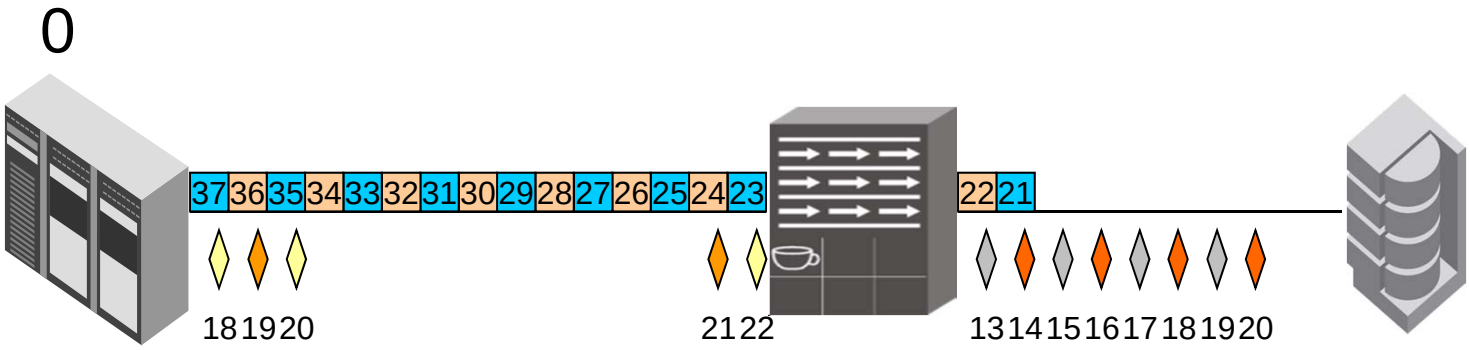

0
32 31 30 29 28 27 26 25 24 23 22 21
13 14 15 16 17 18 19 20
20 19 18
8 9 10 11 12 13 14 15 16 17

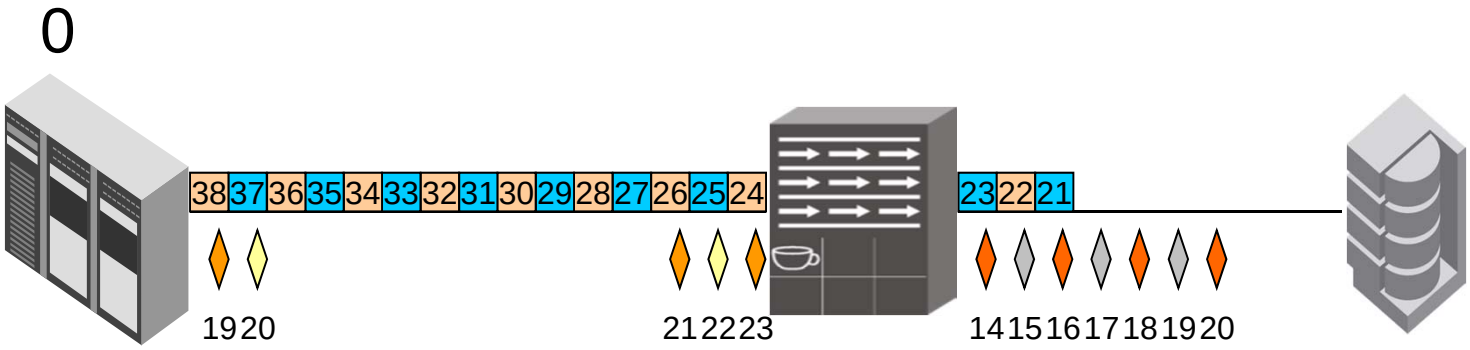

0
33 32 31 30 29 28 27 26 25 24 23 22 21
14 15 16 17 18 19 20
20 19
9 10 11 12 13 14 15 16 17 18

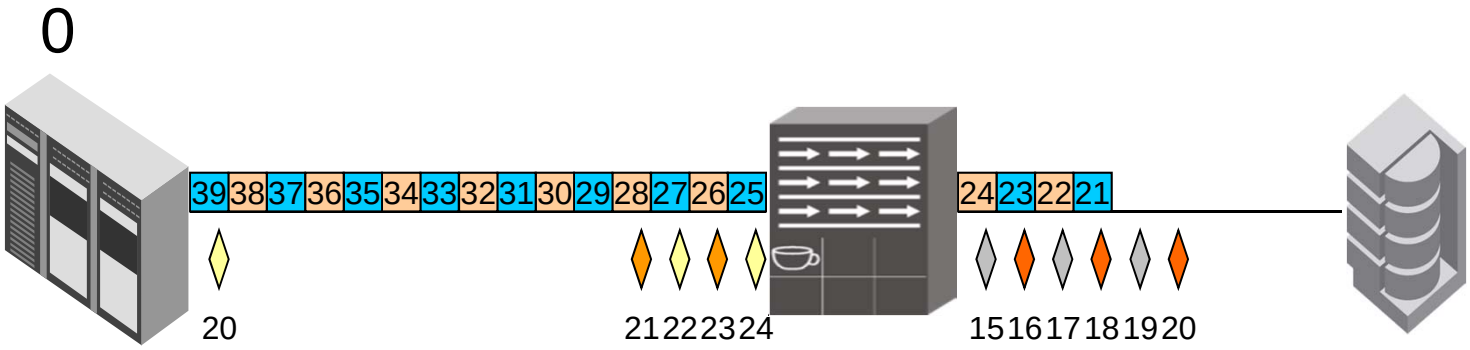

0
34 33 32 31 30 29 28 27 26 25 24 23 22 21
15 16 17 18 19 20
20
10 11 12 13 14 15 16 17 18 19

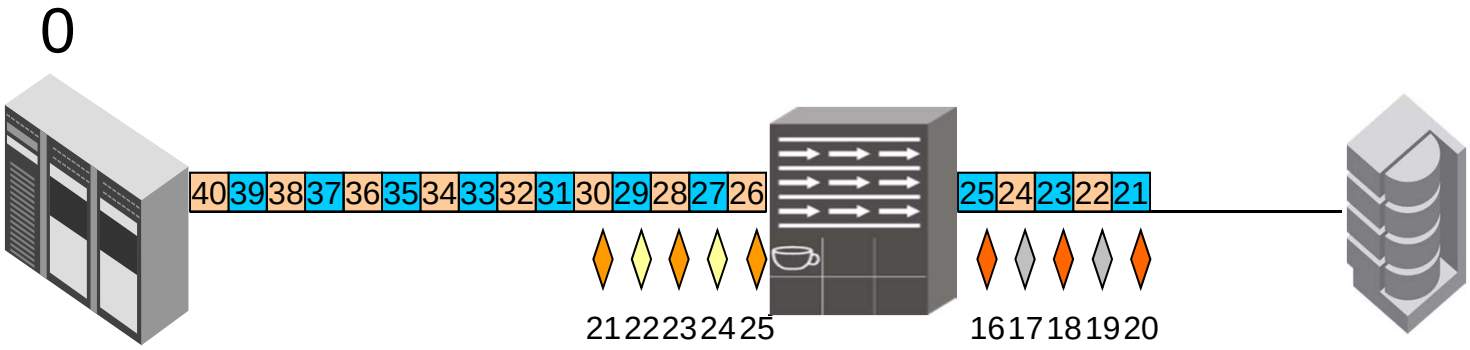

0
35 34 33 32 31 30 29 28 27 26 25 24 23 22 21
16 17 18 19 20
11 12 13 14 15 16 17 18 19 20

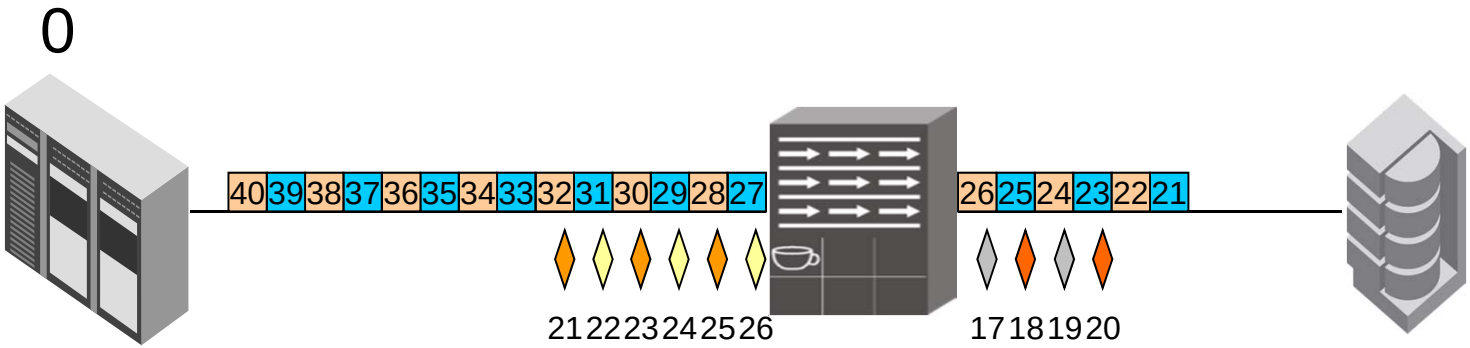

0
36 35 34 33 32 31 30 29 28 27 26 25 24 23 22
21
17 18 19 20
21
12 13 14 15 16 17 18 19 20

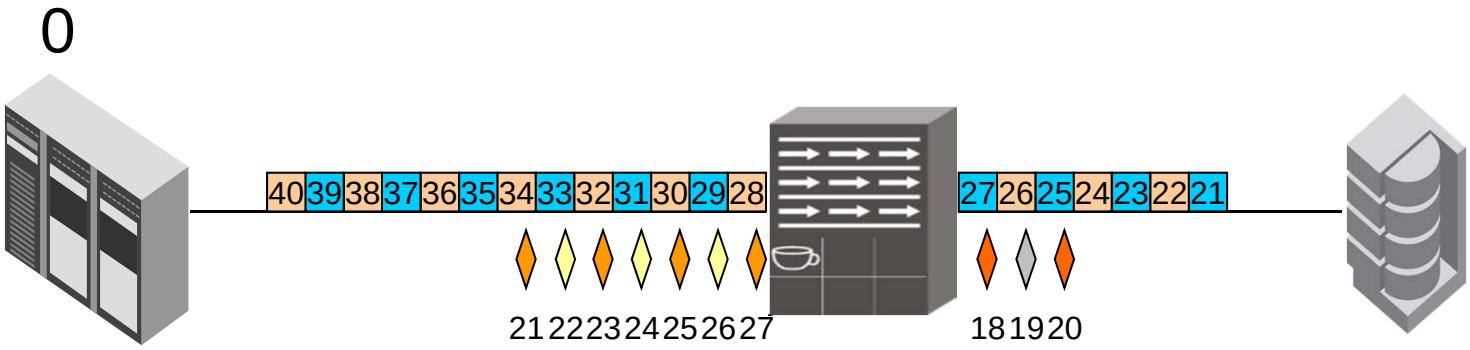

0
37 36 35 34 33 32 31 30 29 28 27 26 25 24 23
22 21
18 19 20
21 22
13 14 15 16 17 18 19 20

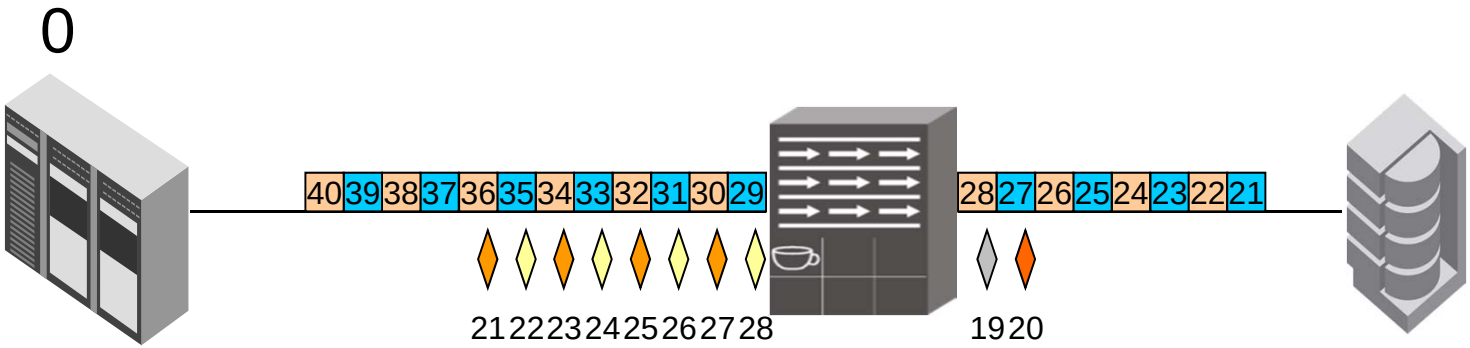

0
38 37 36 35 34 33 32 31 30 29 28 27 26 25 24
23 22 21
19 20
21 22 23
14 15 16 17 18 19 20

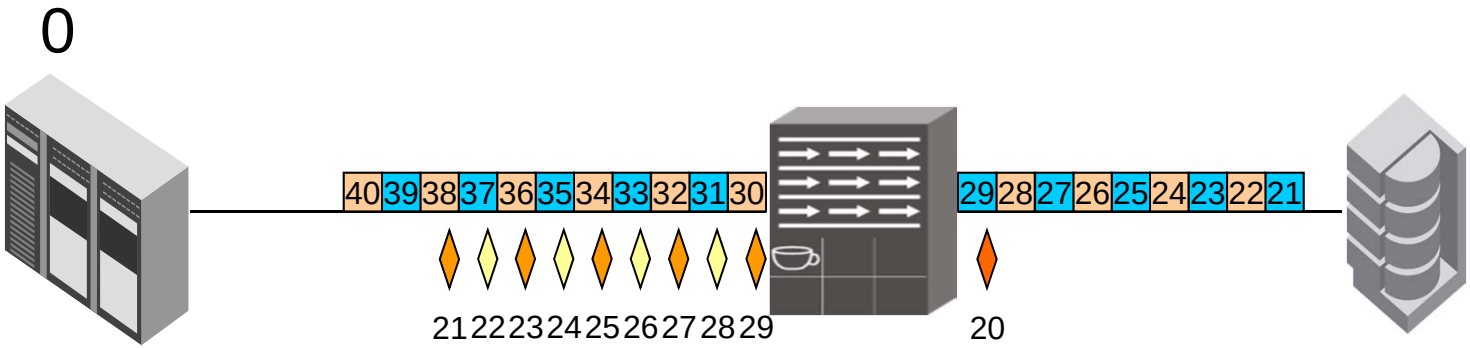

0
39 38 37 36 35 34 33 32 31 30 29 28 27 26 25
24 23 22 21
20
21 22 23 24
15 16 17 18 19 20

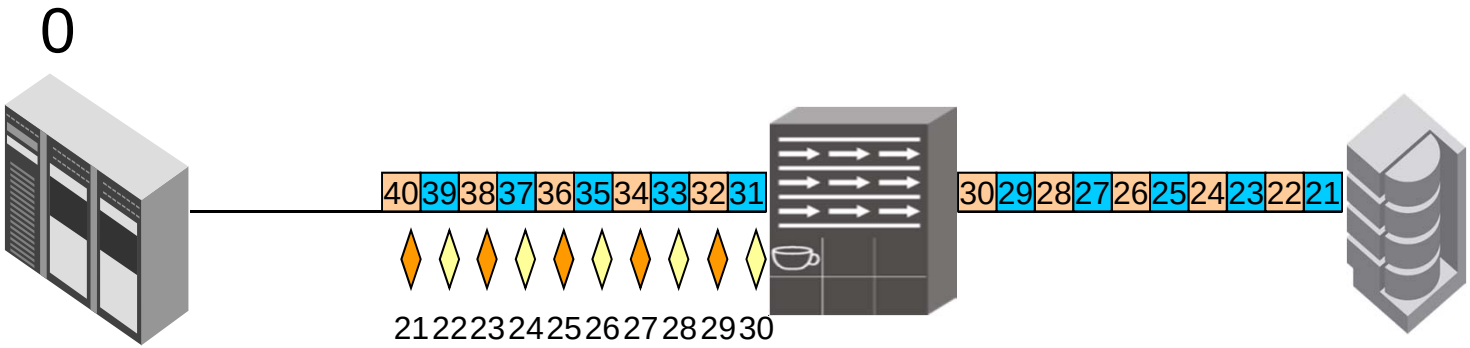

0
40 39 38 37 36 35 34 33 32 31 30 29 28 27 26
25 24 23 22 21
21 22 22 23 24 25
16 17 18 19 20

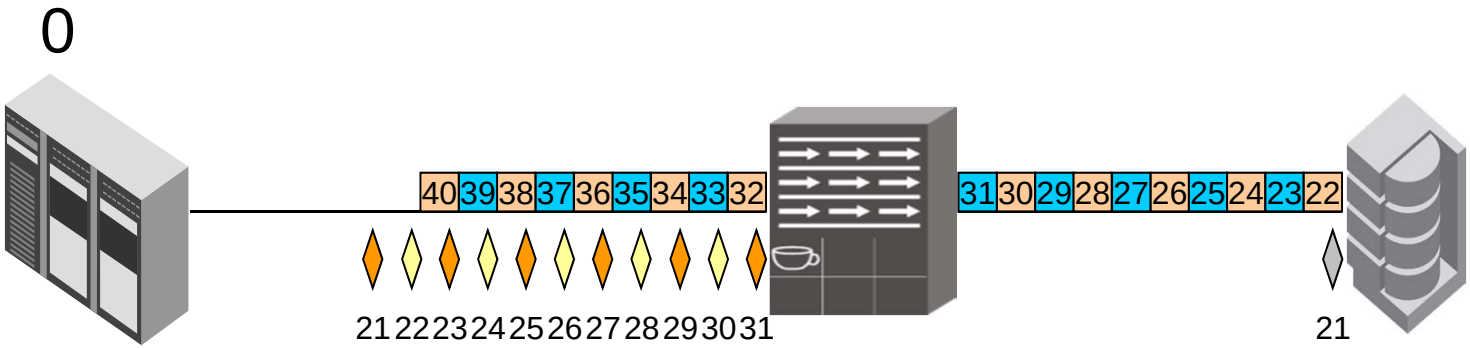

0
40 39 38 37 36 35 34 33 32 31 30 29 28 27
26 25 24 23 22 21
21 22 23 24 25 26
17 18 19 20

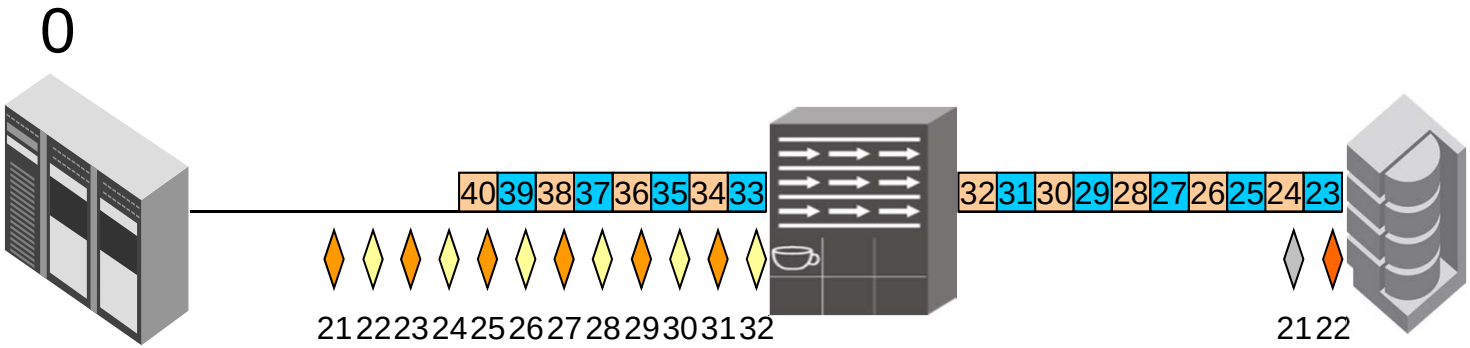

0
40 39 38 37 36 35 34 33 32 31 30 29 28
27 26 25 24 23 22 21
21 22 23 24 25 26 27
18 19 20


0
40 39 38 37 36 35 34 33 32 31 30 29
28 27 26 25 24 23 22 21
21 22 23 24 25 26 27 28
19 20


0
40 39 38 37 36 35 34 33 32 31 30
29 28 27 26 25 24 23 22 21
21 22 23 24 25 26 27 28 29
20


0
40 39 38 37 36 35 34 33 32 31
30 29 28 27 26 25 24 23 22 21
21 22 23 24 25 26 27 28 29 30


0
40 39 38 37 36 35 34 33 32
31 30 29 28 27 26 25 24 23 22
21 22 23 24 25 26 27 28 29 30 31
21


0
40 39 38 37 36 35 34 33
32 31 30 29 28 27 26 25 24 23
21 22 23 24 25 26 27 28 29 30 31 32
21 22

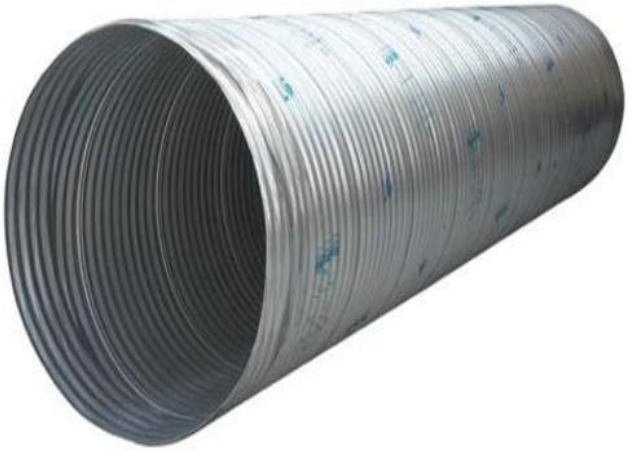
THIS PAGE INTENTIONALLY LEFT BLANK

Example: Different sized pipes

BUFFER CREDITS

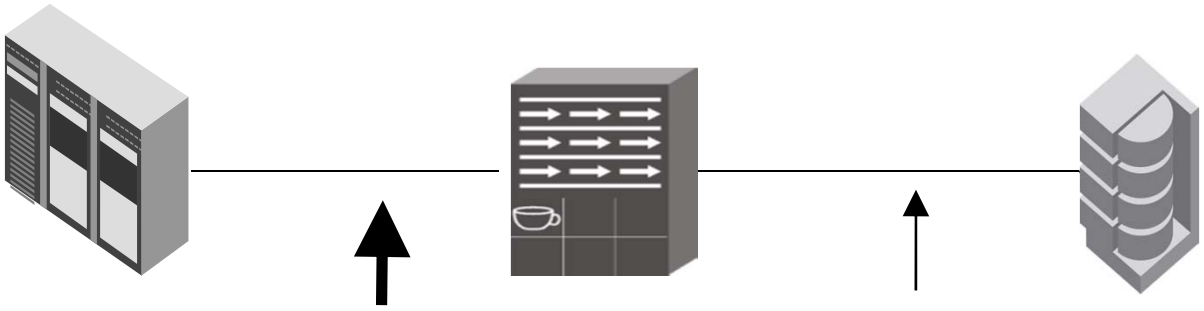
Fat Pipe

Skinny Pipe



Fat Pipe / Skinny Pipe

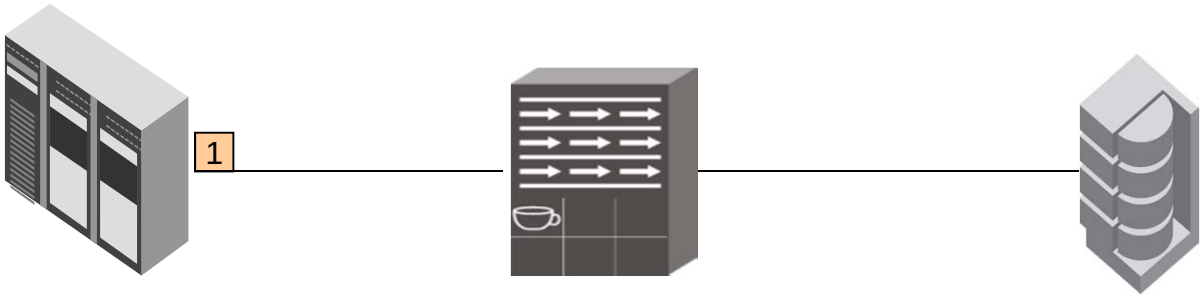
20



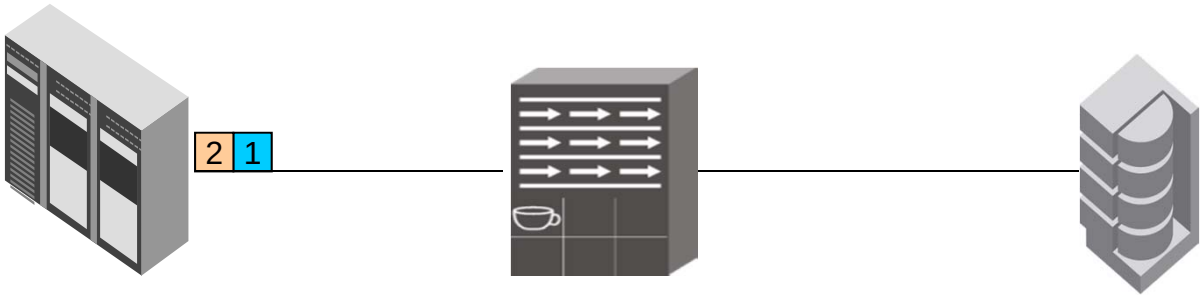
Fat Pipe

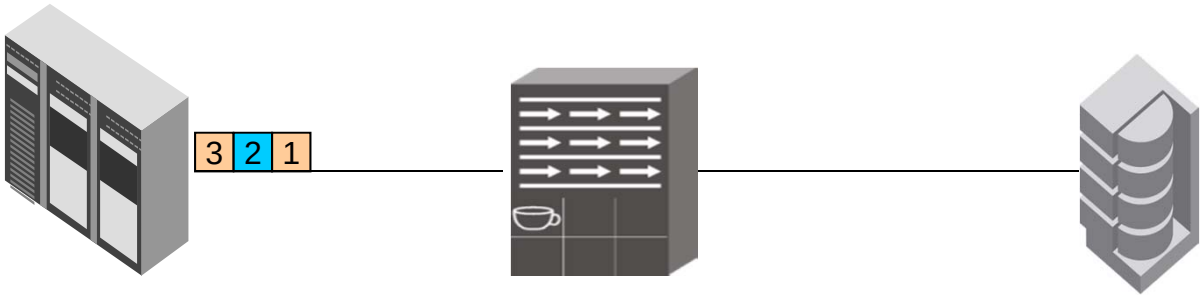
Skinny Pipe

19

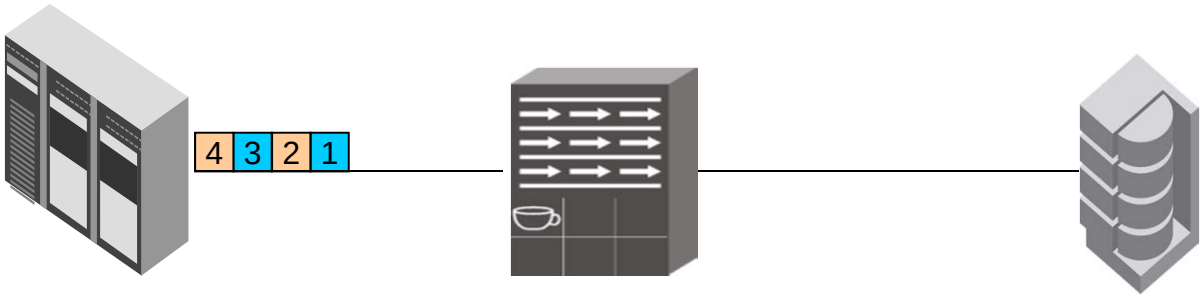


18

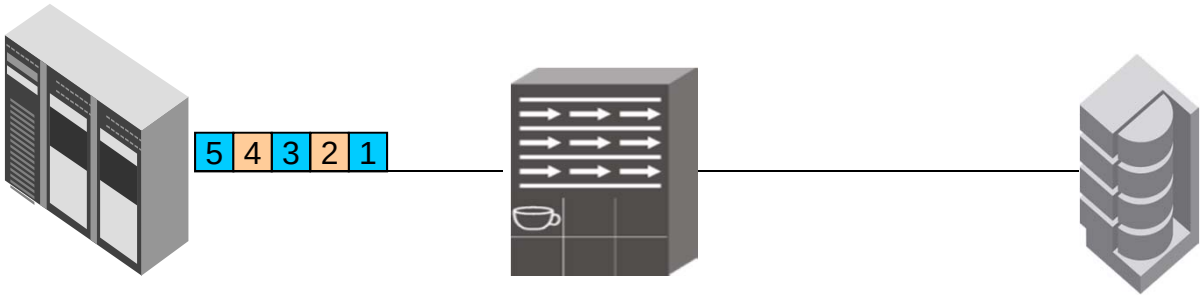


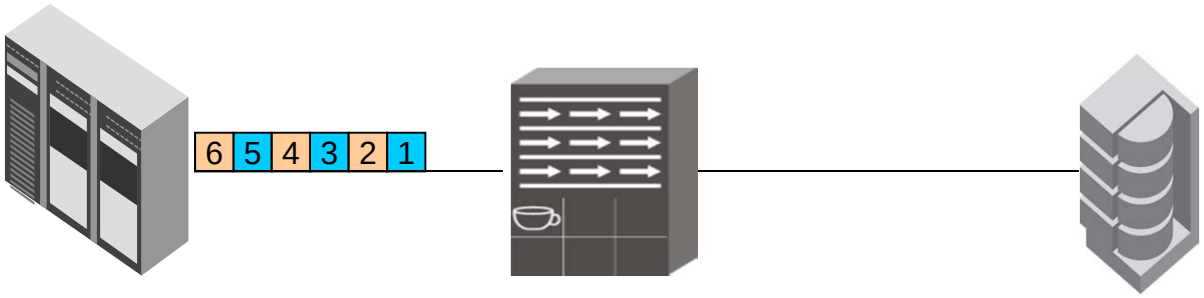

3
2
1

16

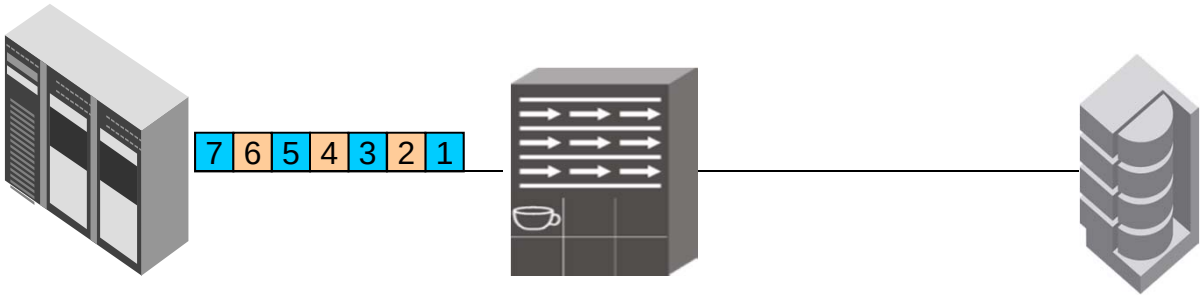


15

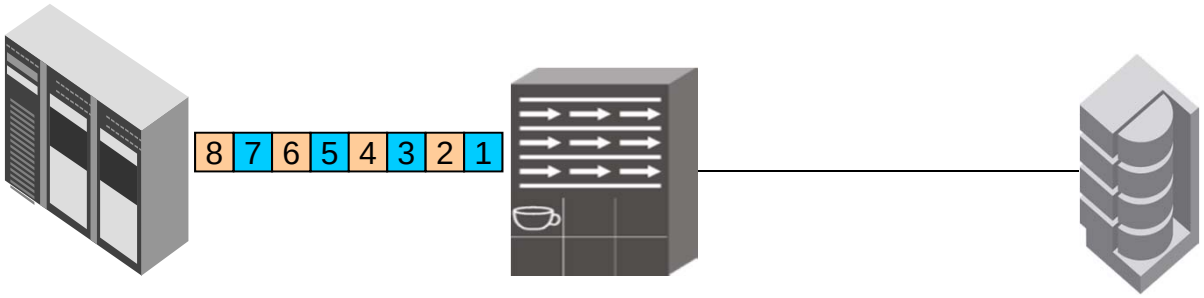



6
5
4
3
2
1

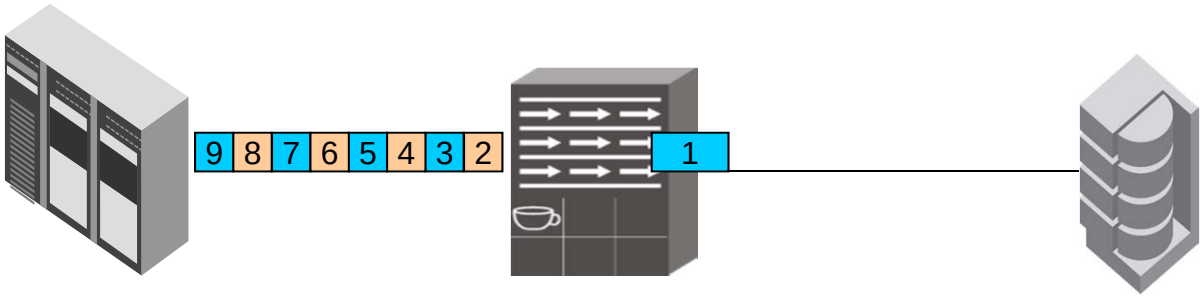
13



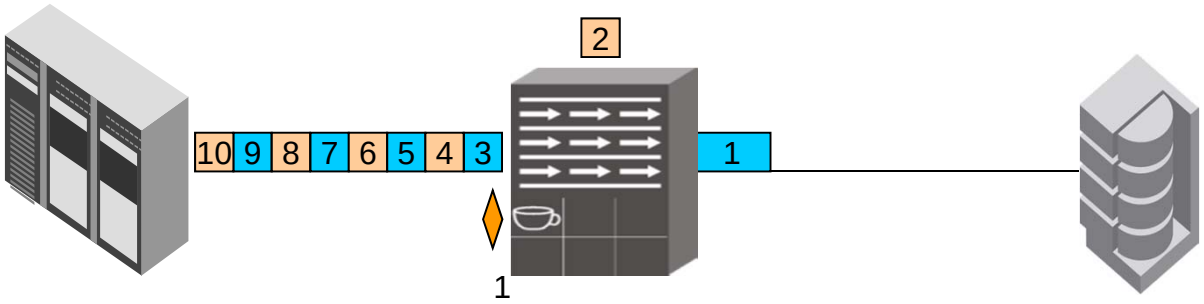
12



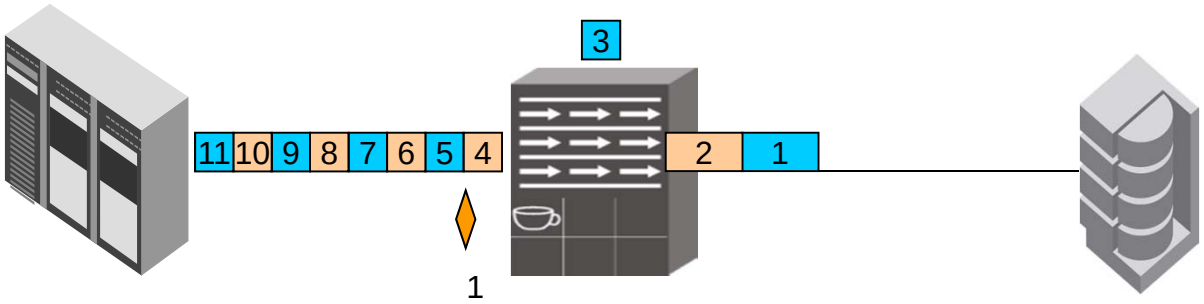
11

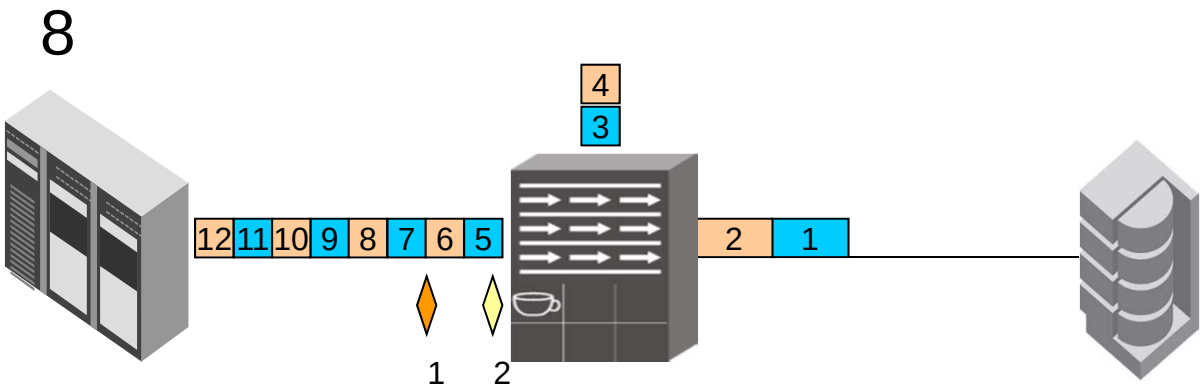


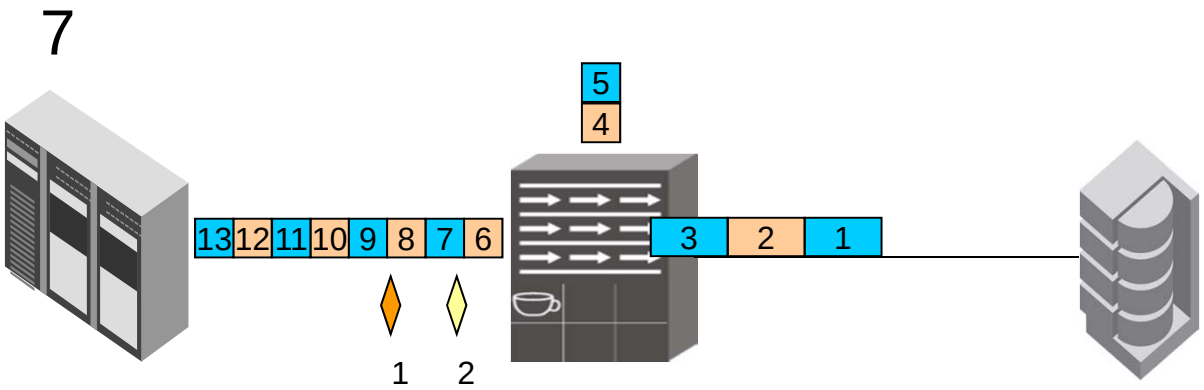
10

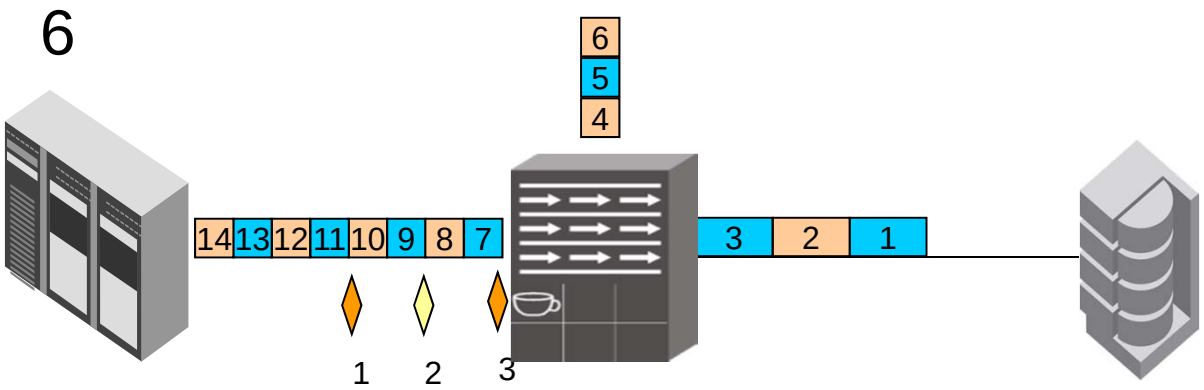


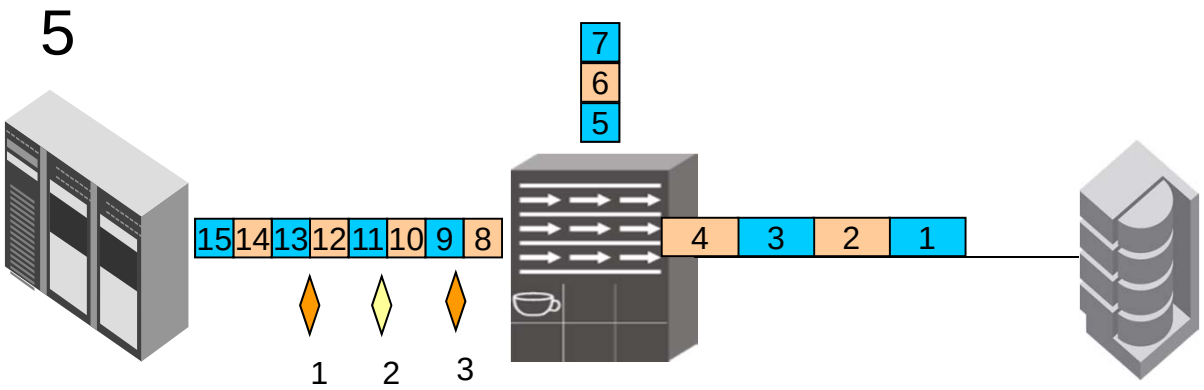
9

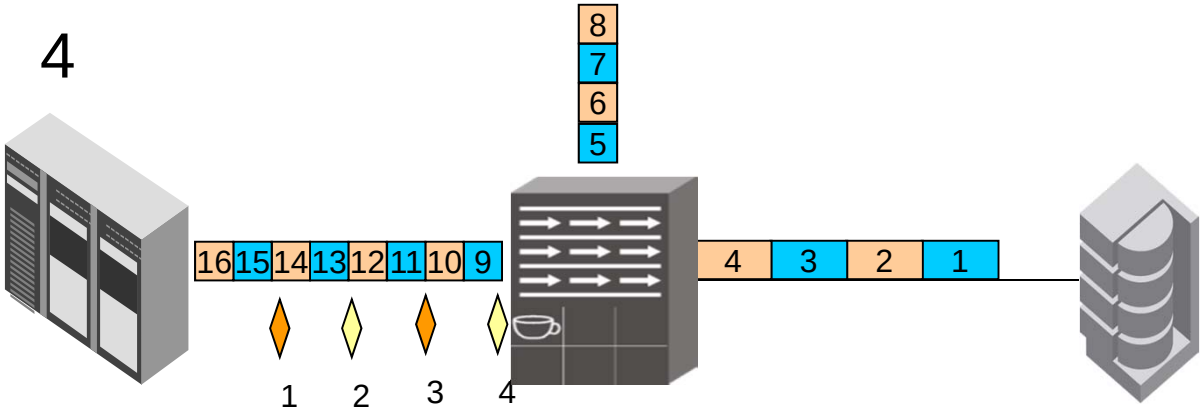


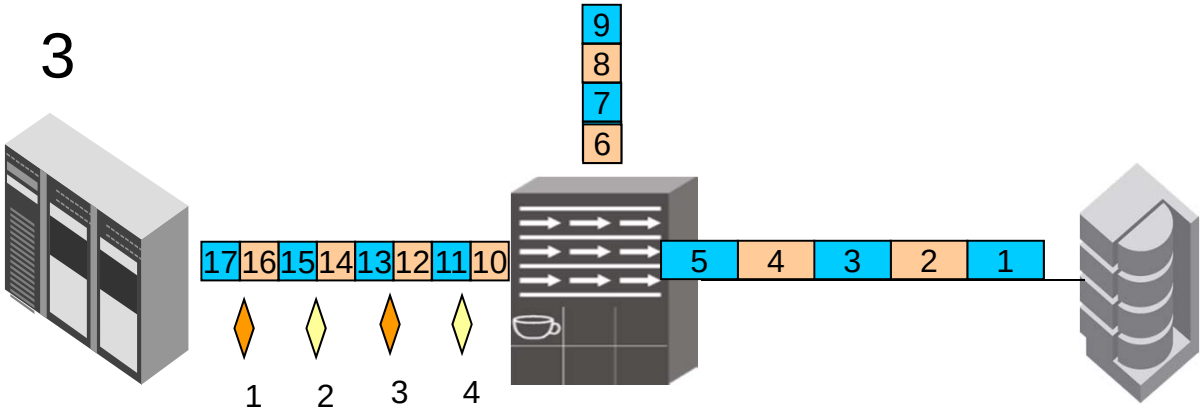

8
4
3
12
11
10
9
8
7
6
5
2
1
1
2

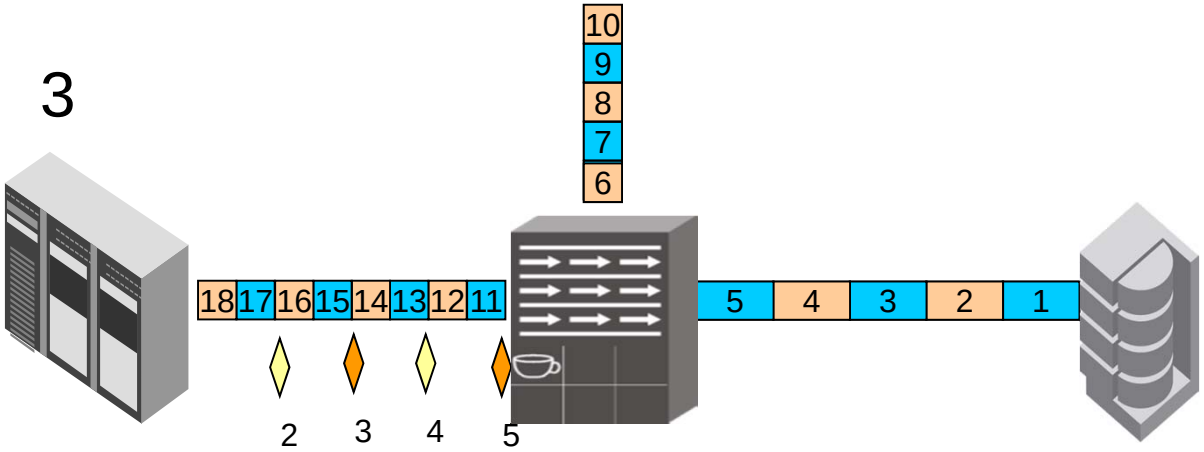

7
5
4
13
12
11
10
9
8
7
6
3
2
1
1
2

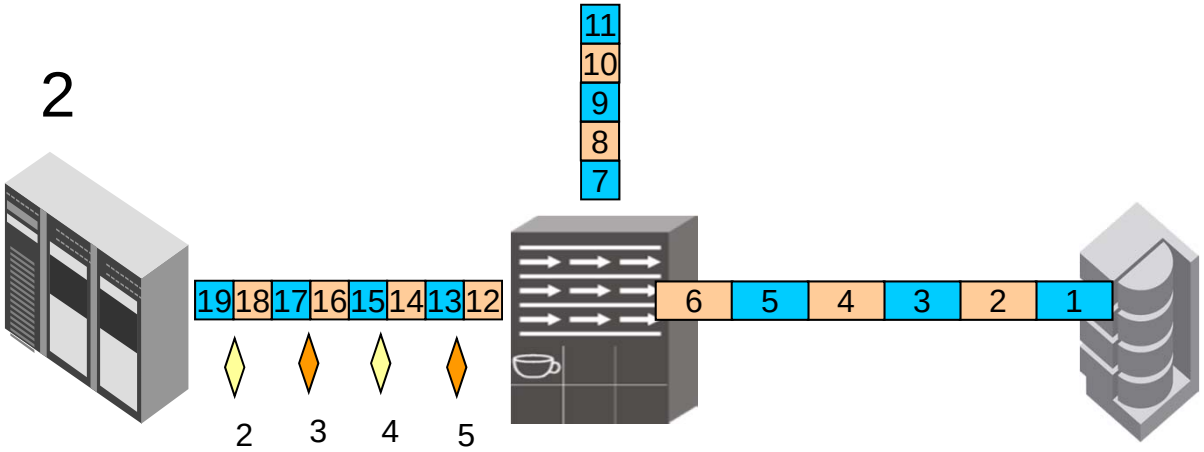

6
6
5
4
14
13
12
11
10
9
8
7
3
2
1
1
2
3

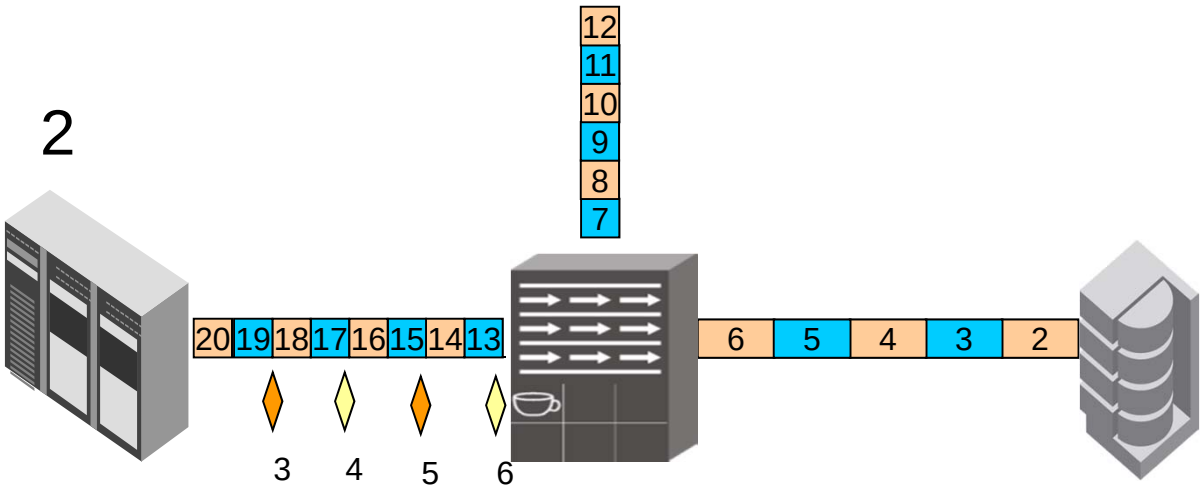

5
7
6
5
15
14
13
12
11
10
9
8
4
3
2
1
1
2
3

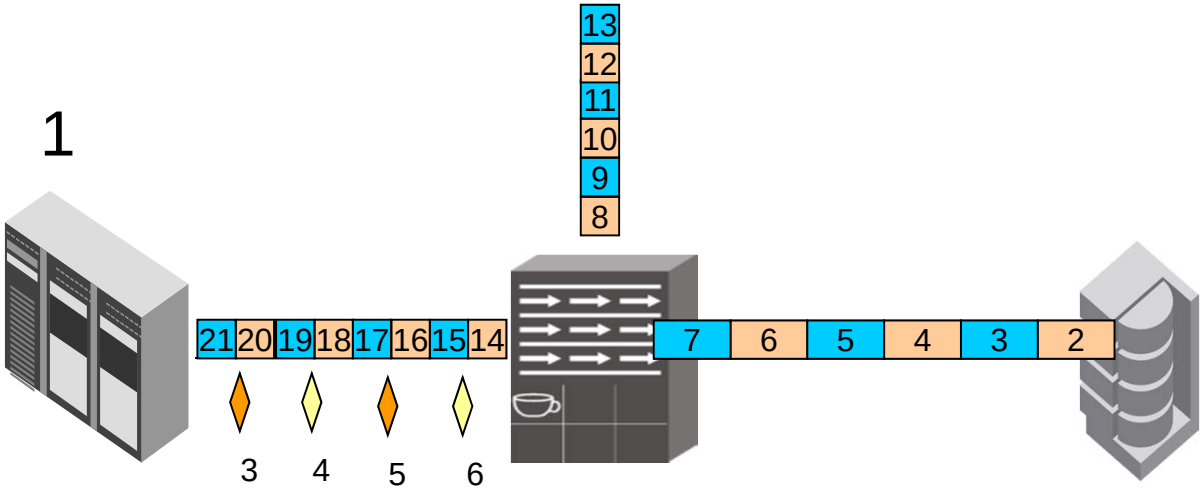

4
8
7
6
5
16 15 14 13 12 11 10 9
4 3 2 1
1
2
3
4

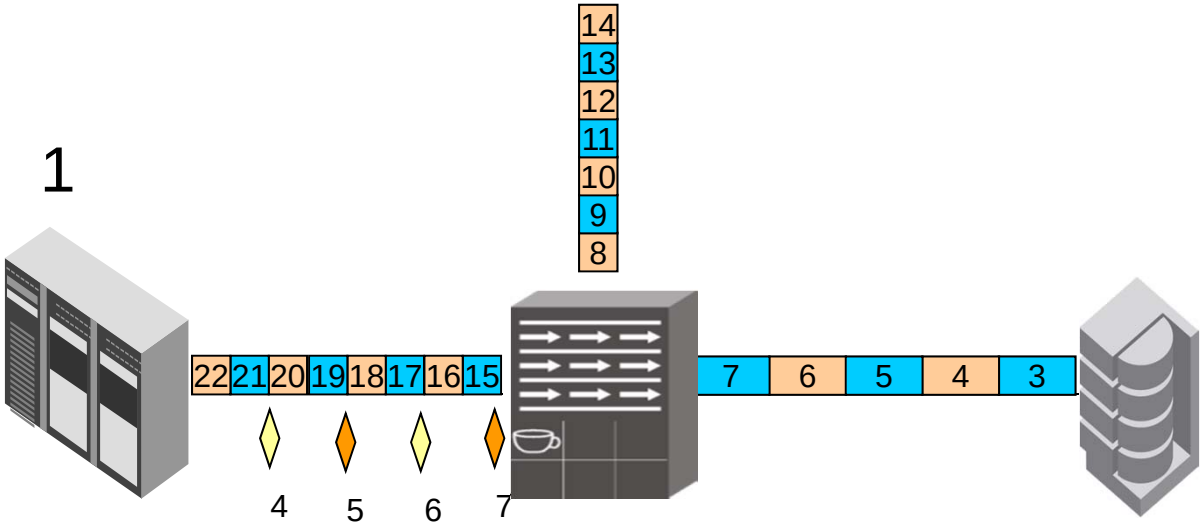

3
9
8
7
6
17
16
15
14
13
12
11
10
5
4
3
2
1
1
2
3
4

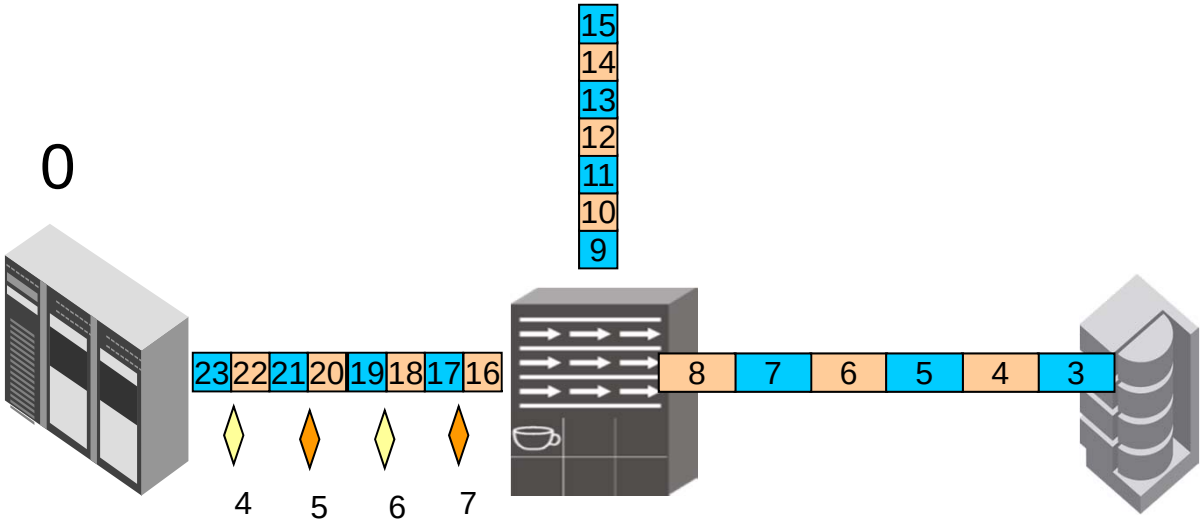

3
10
9
8
7
6
18 17 16 15 14 13 12 11
5 4 3 2 1
2 3 4 5

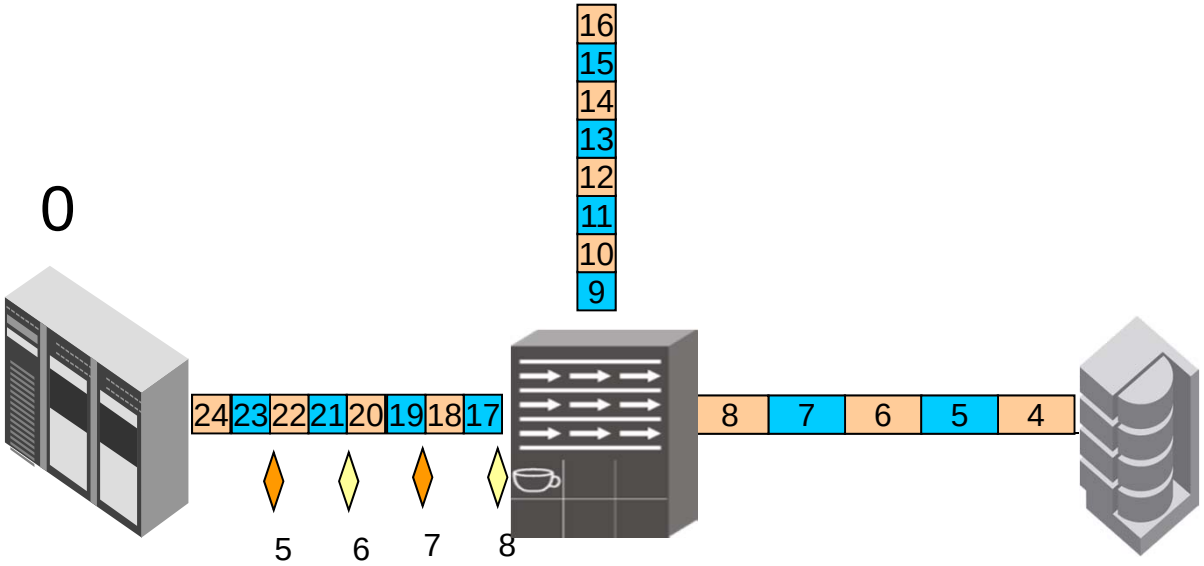

2
11
10
9
8
7
19 18 17 16 15 14 13 12
6 5 4 3 2 1
2 3 4 5

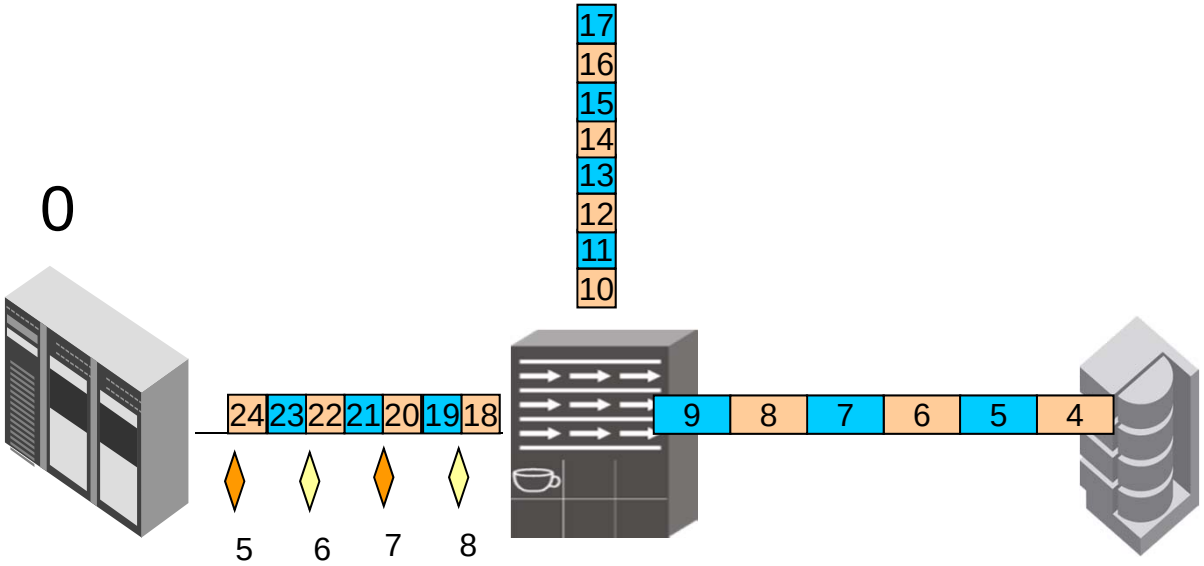

2
12
11
10
9
8
7
20 19 18 17 16 15 14 13
6 5 4 3 2
3 4 5 6

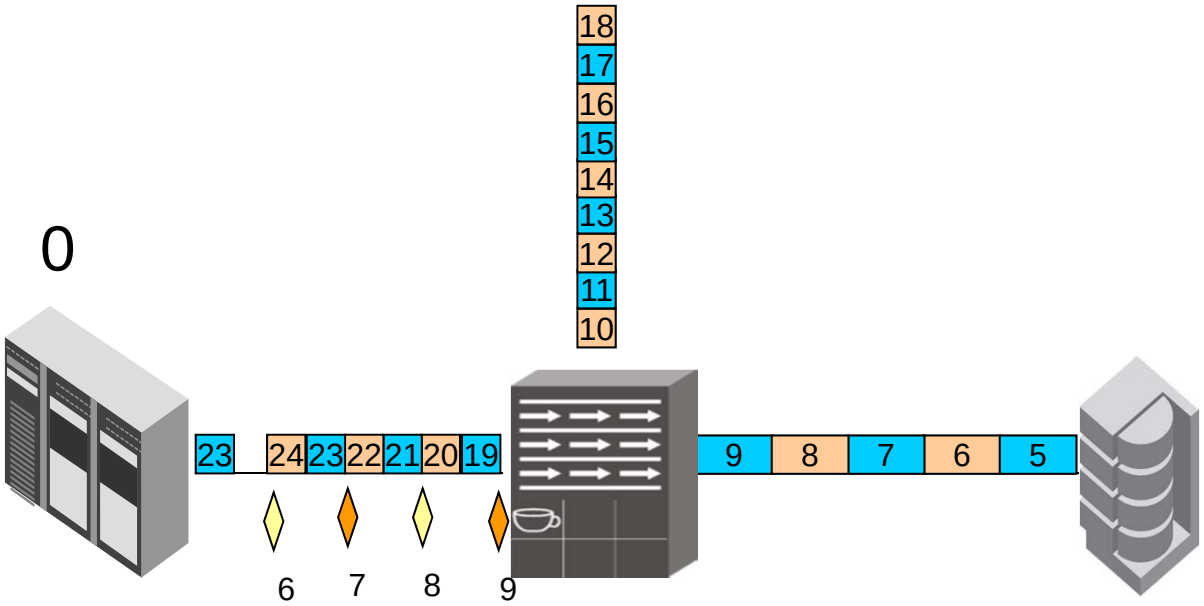

1
13
12
11
10
9
8
21
20
19
18
17
16
15
14
7
6
5
4
3
2
3
4
5
6

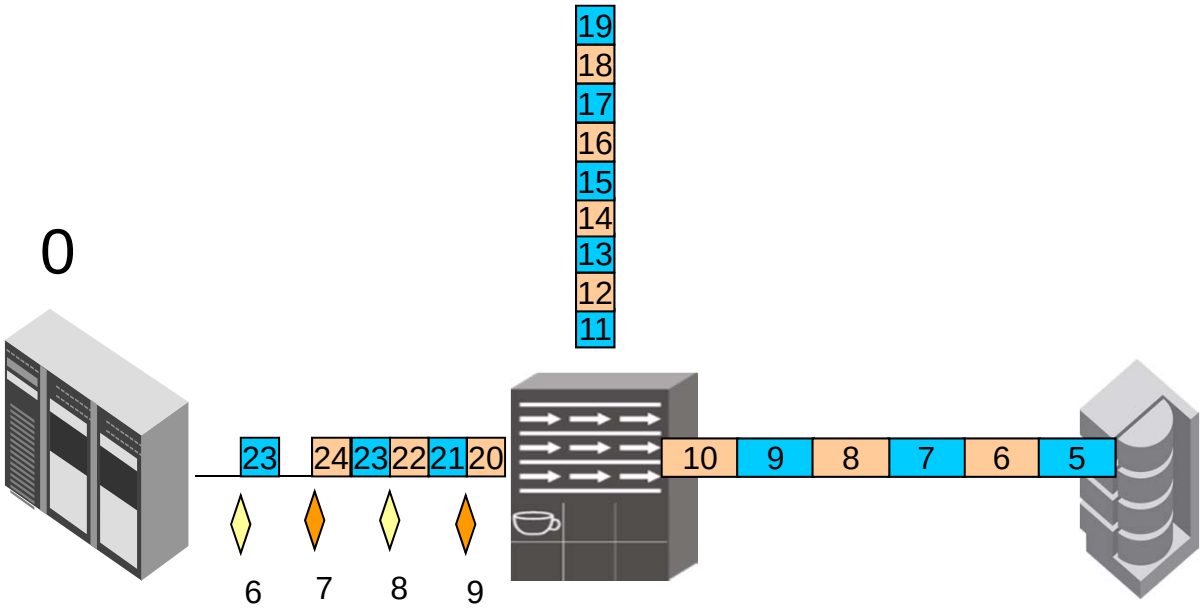

1
14
13
12
11
10
9
8
22
21
20
19
18
17
16
15
7
6
5
4
3
4
5
6
7

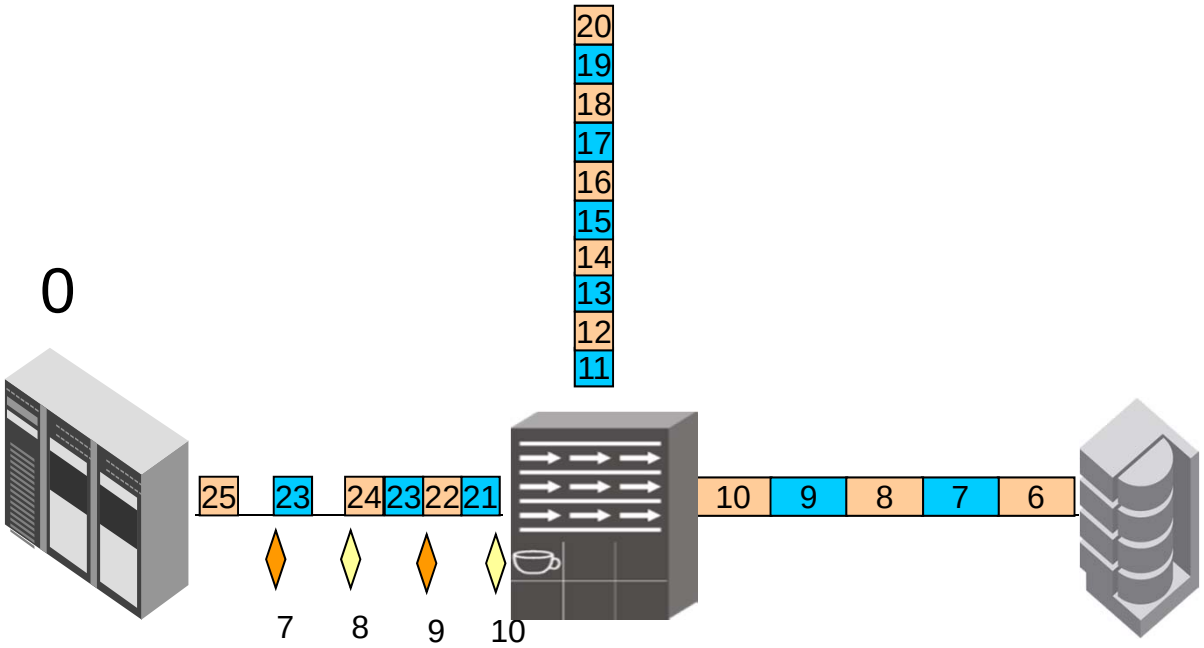

0
15
14
13
12
11
10
9
23
22
21
20
19
18
17
16
8
7
6
5
4
3
4
5
6
7

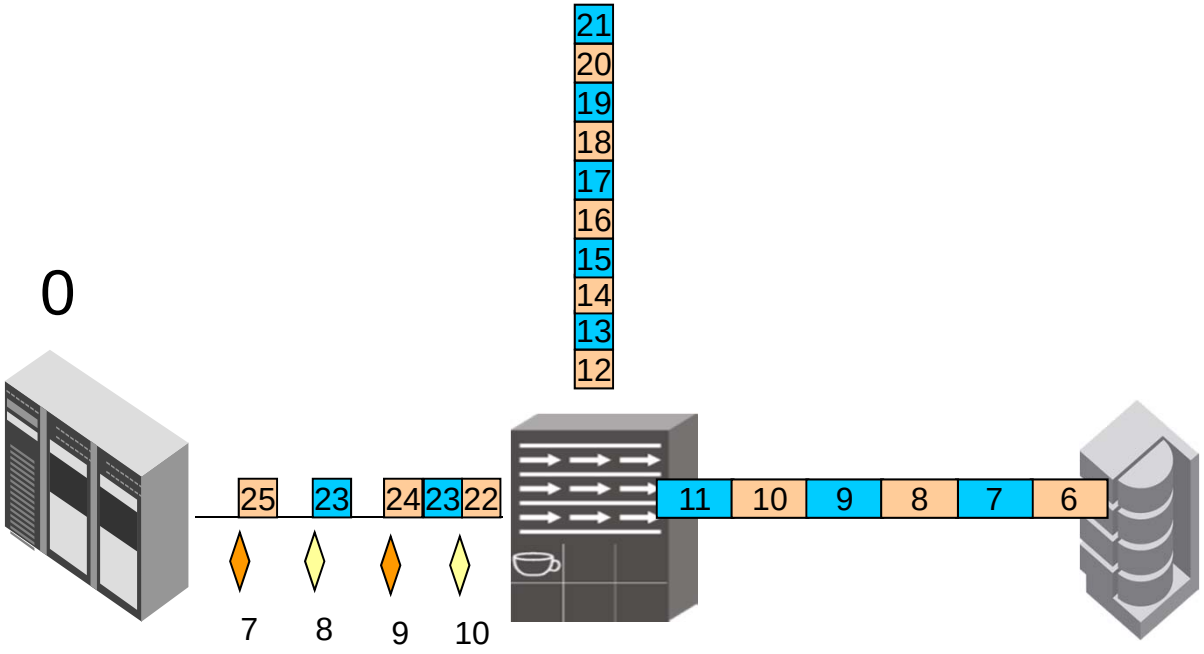

0
16
15
14
13
12
11
10
9
24
23
22
21
20
19
18
17
8
7
6
5
4
5
6
7
8

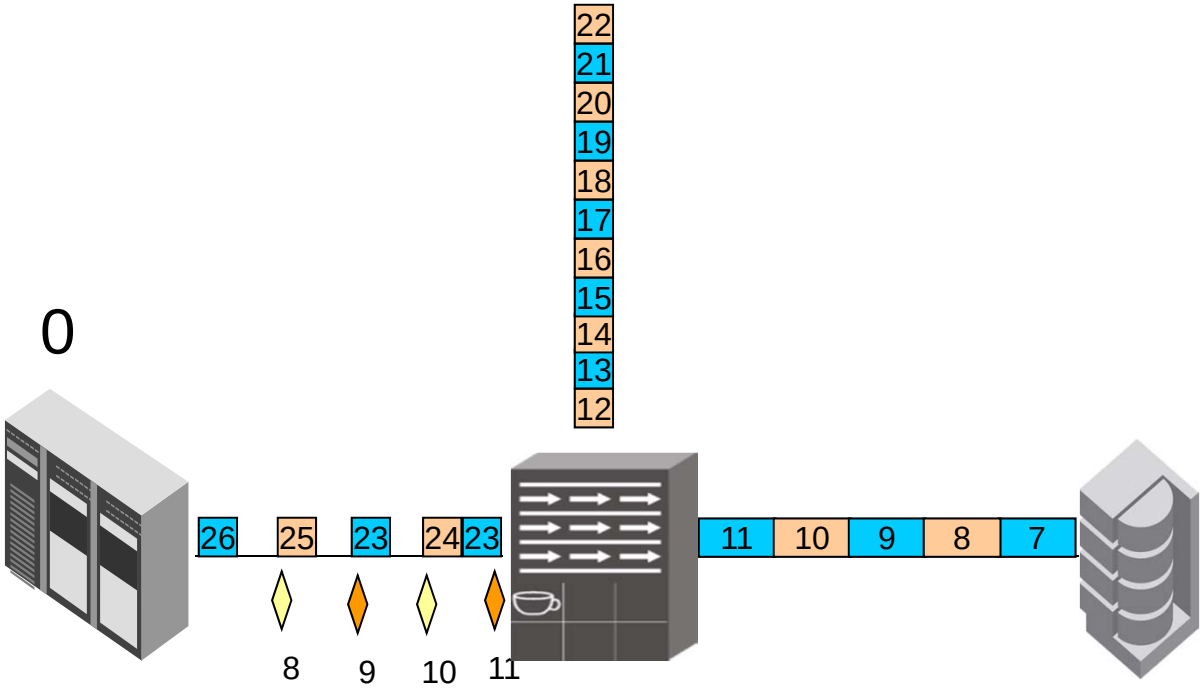

0
17
16
15
14
13
12
11
10
24
23
22
21
20
19
18
9
8
7
6
5
4
5
6
7
8

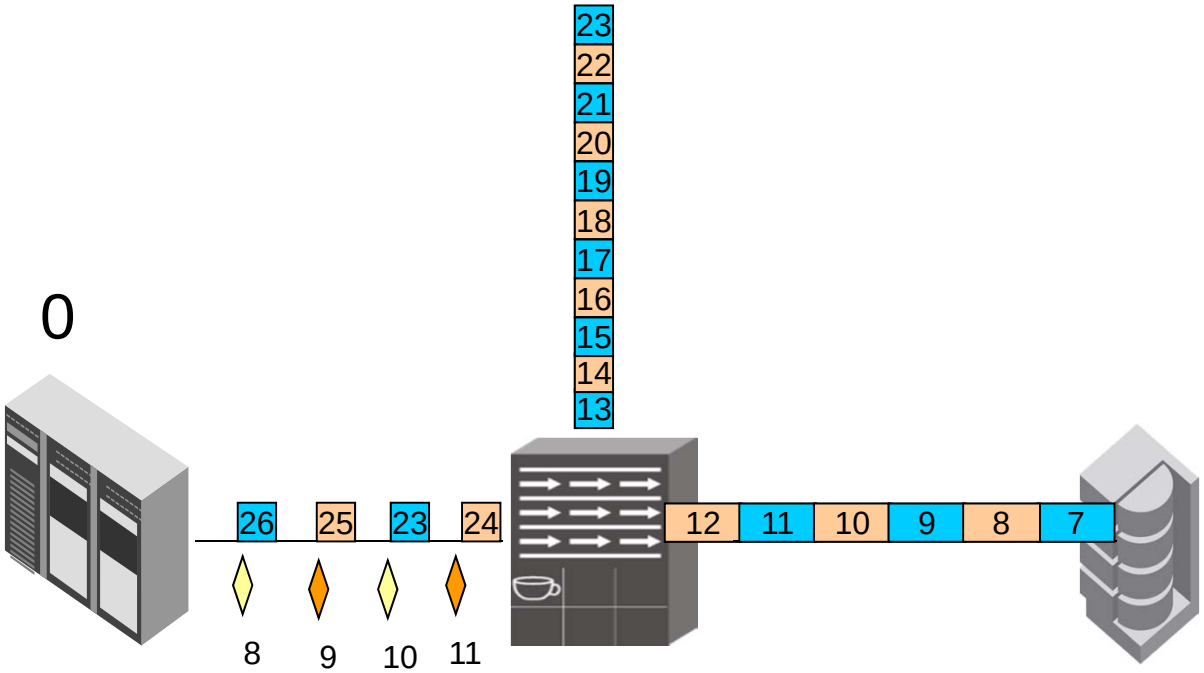

0
18
17
16
15
14
13
12
11
10
23
24
23
22
21
20
19
9
8
7
6
5
6
7
8
9

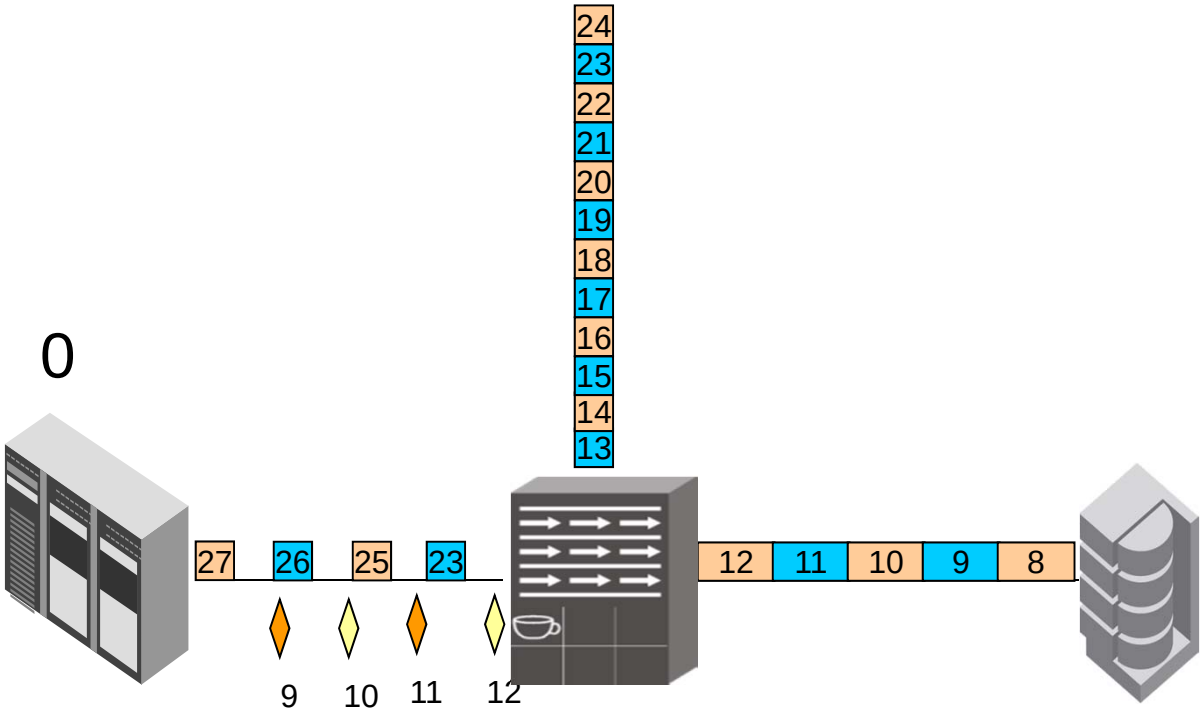

0
19
18
17
16
15
14
13
12
11
23
24
23
22
21
20
10
9
8
7
6
5
6
7
8
9

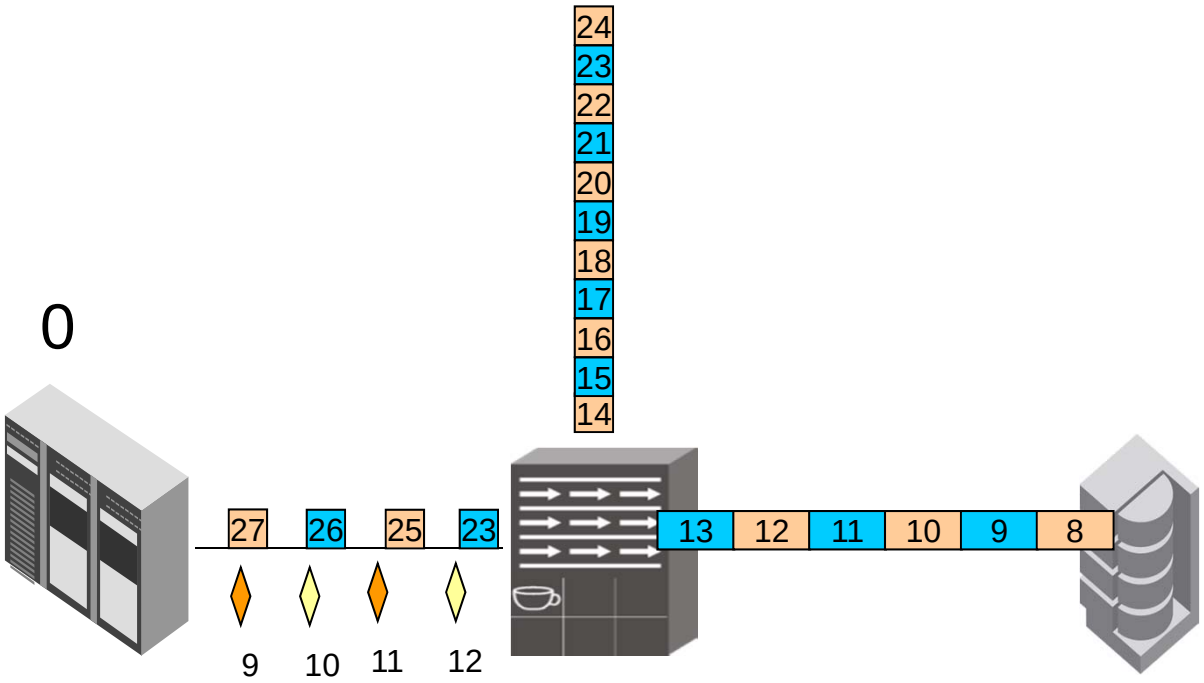

20
19
18
17
16
15
14
13
12
11
0
25
23
24
23
22
21
10
9
8
7
6
7
8
9
10

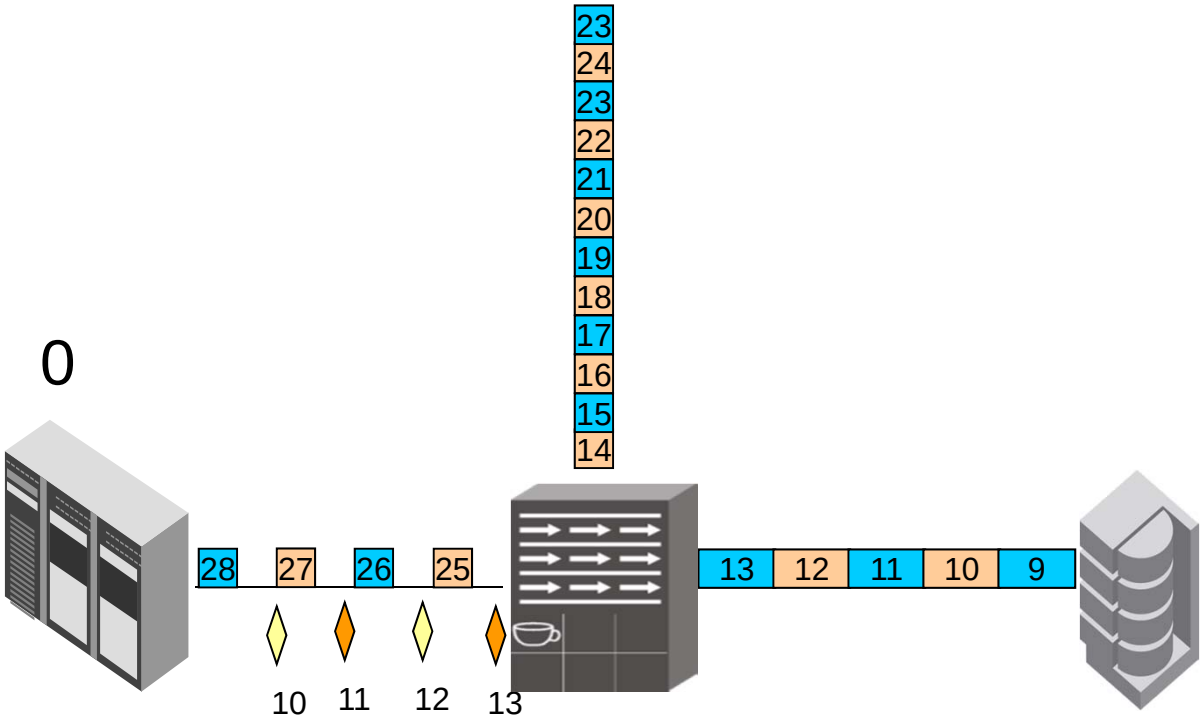

21
20
19
18
17
16
15
14
13
12
0
25
23
24
23
22
11
10
9
8
7
6
7
8
9
10

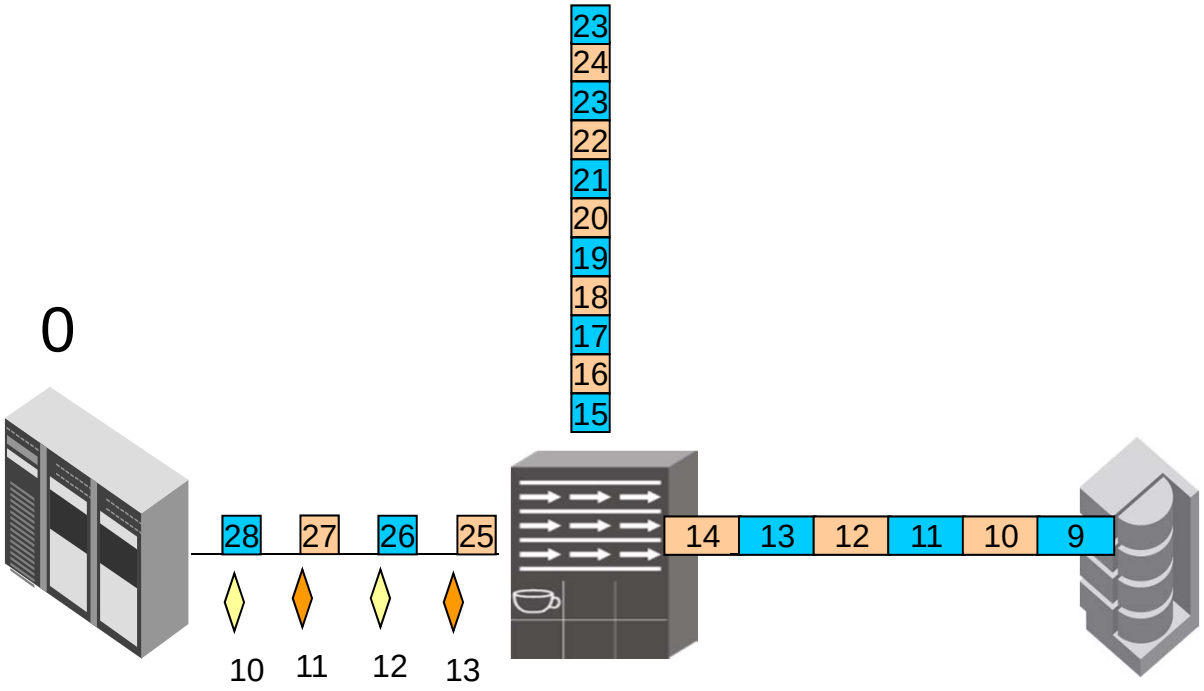

22
21
20
19
18
17
16
15
14
13
12
0
26
25
23
24
23
11
10
9
8
7
8
9
10
11

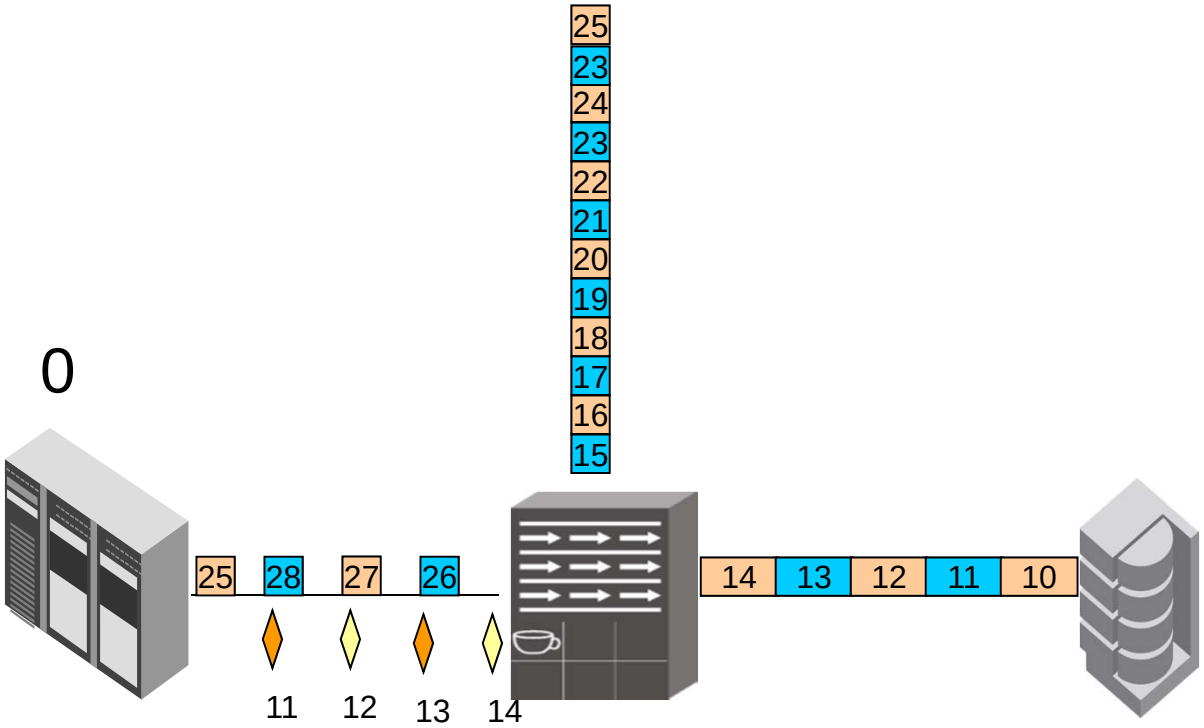

23
22
21
20
19
18
17
16
15
14
13
0
26
25
23
24
12
11
10
9
8
7
8
9
10
11


24
23
22
21
20
19
18
17
16
15
14
13
0
27
26
25
23
12
11
10
9
8
9
10
11
12


24
23
22
21
20
19
18
17
16
15
14
0
27
26
25
23
13
12
11
10
9
8
9
10
11
12


23
24
23
22
21
20
19
18
17
16
15
14
0
28
27
26
25
13
12
11
10
9
10
11
12
13


23
24
23
22
21
20
19
18
17
16
15
0
28
27
26
25
14
13
12
11
10
9
10
11
12
13


25
23
24
23
22
21
20
19
18
17
16
15
0
25
28
27
26
14
13
12
11
10
11
12
13
14

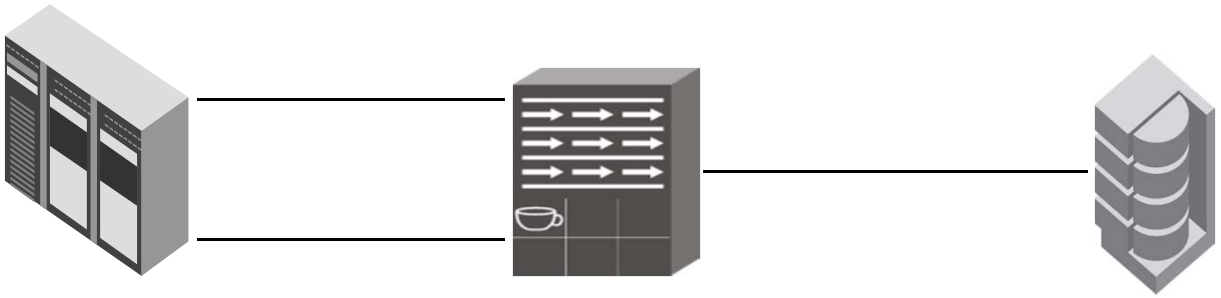
THIS PAGE INTENTIONALLY LEFT BLANK

Example: Real Life?

BUFFER CREDITS

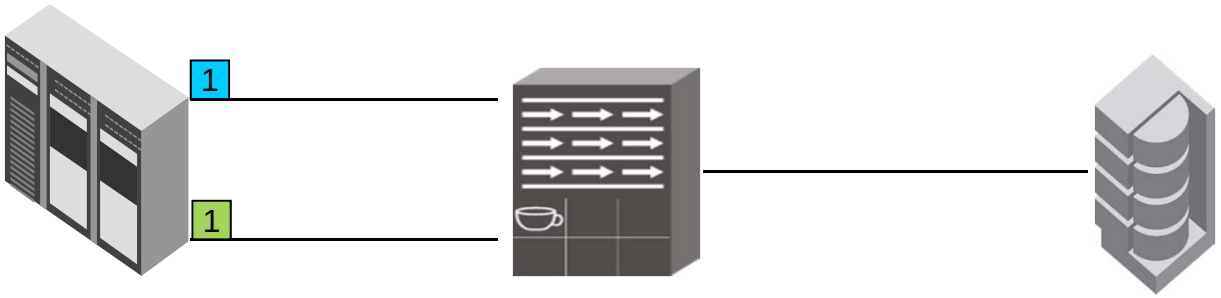
More real life example: Two senders sending at 30% - 50% link rate to one receiver

20



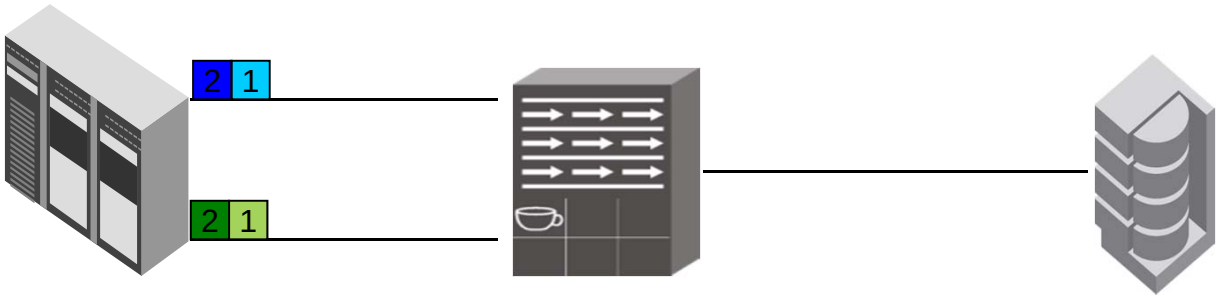
20

19



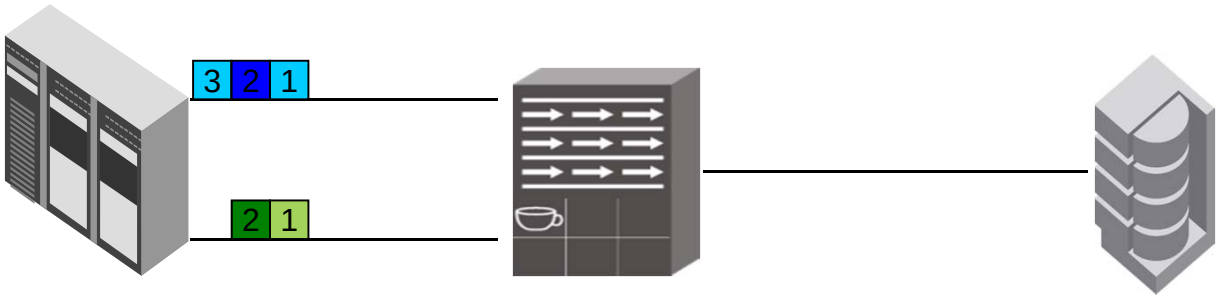
19

18



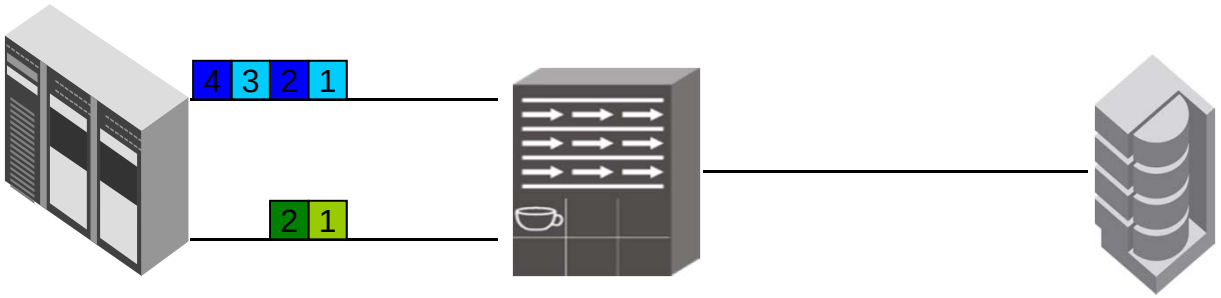
18

17



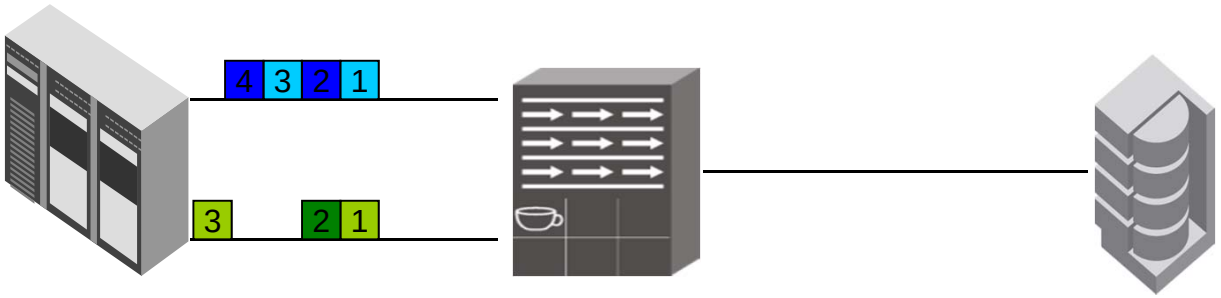
18

16



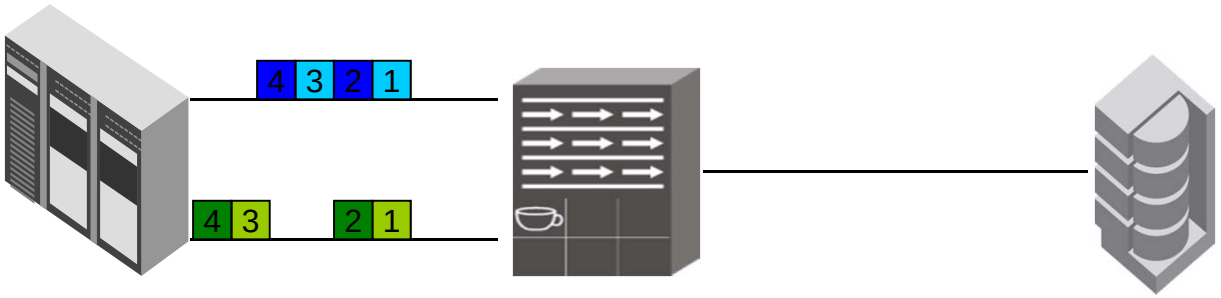
18

16



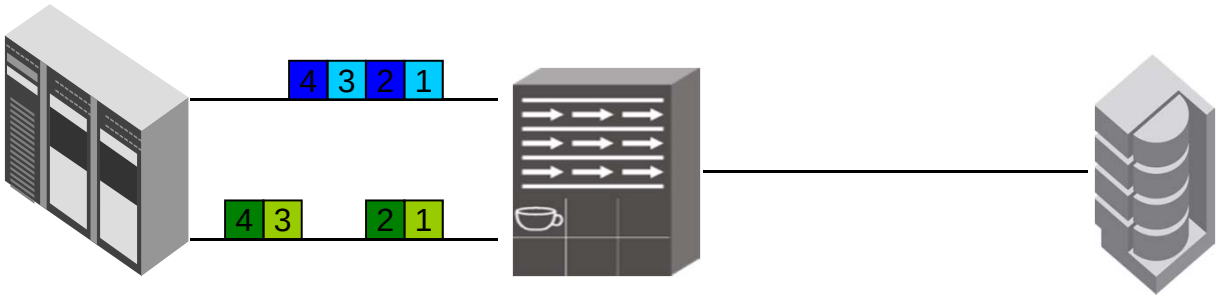
17

16



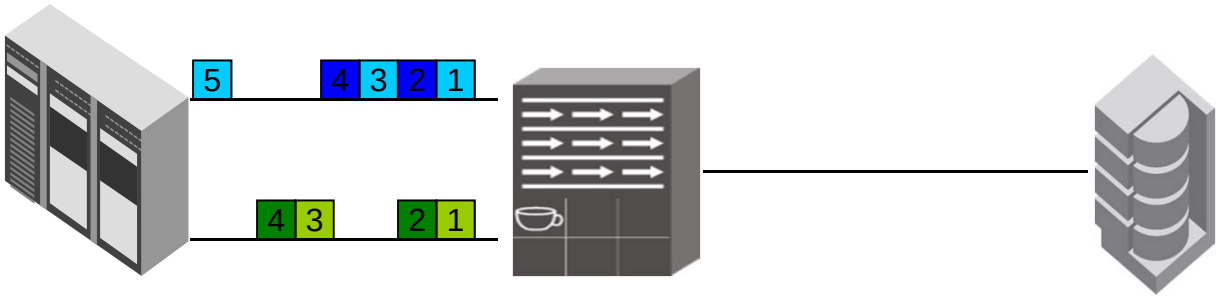
16

16



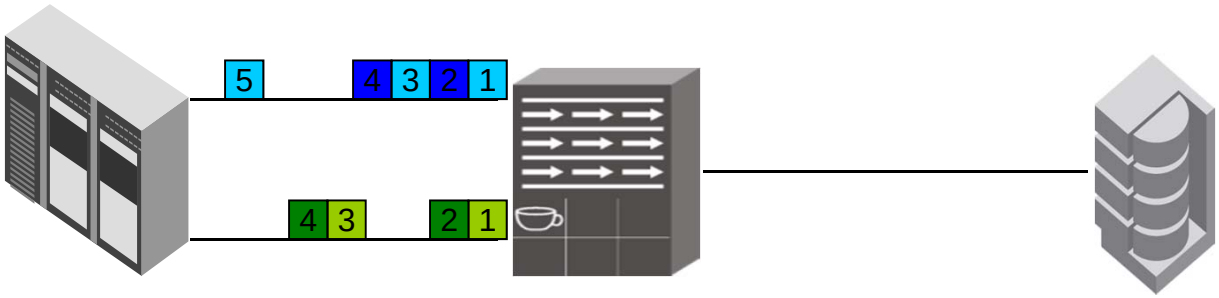
16

15



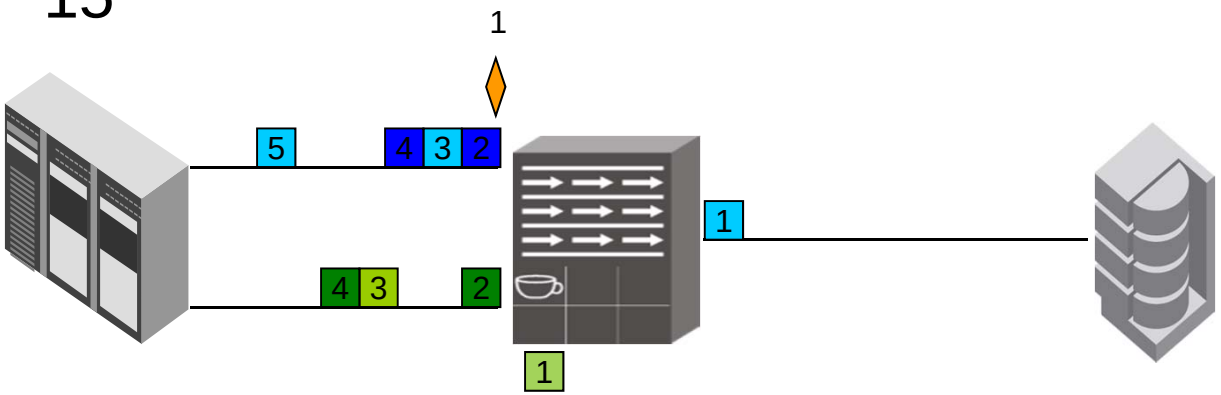
16

15

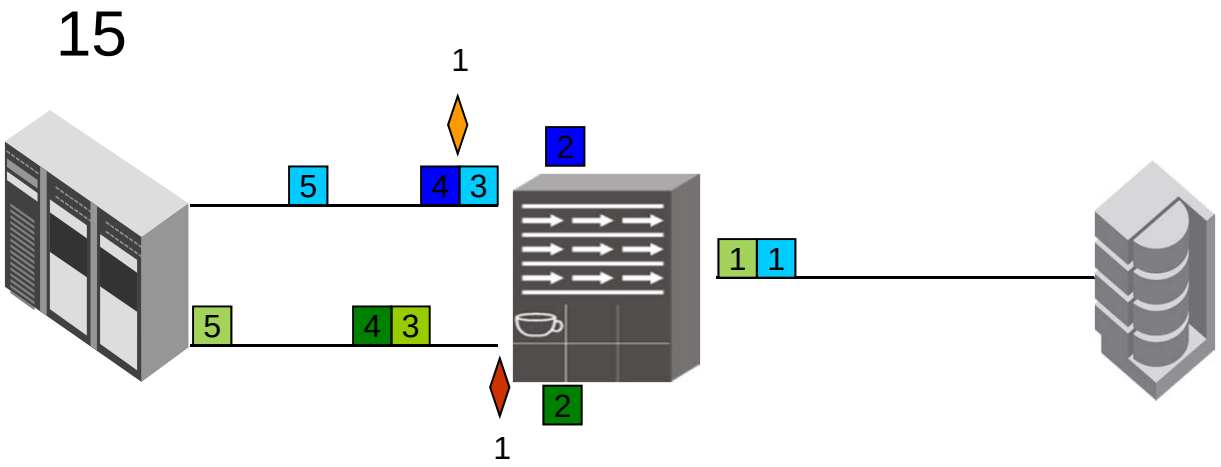


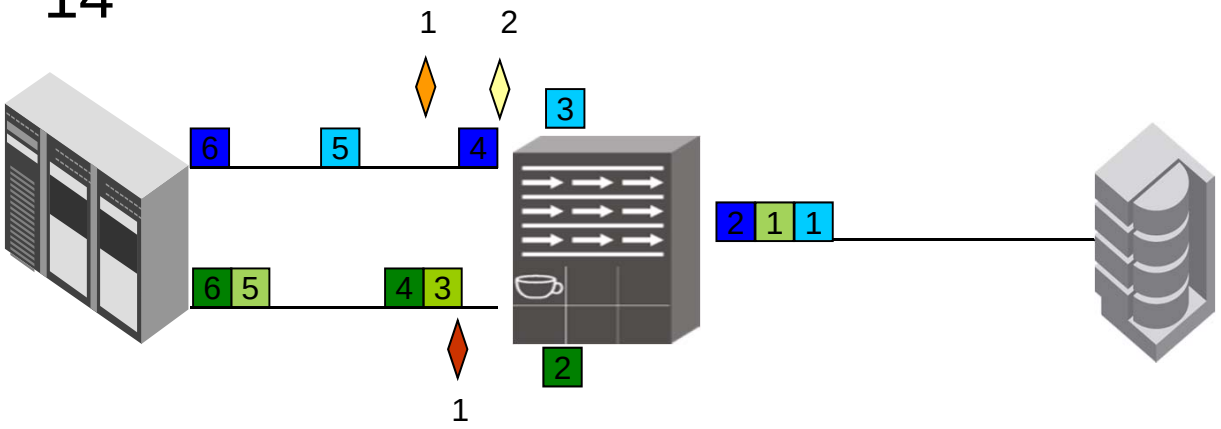
16

15



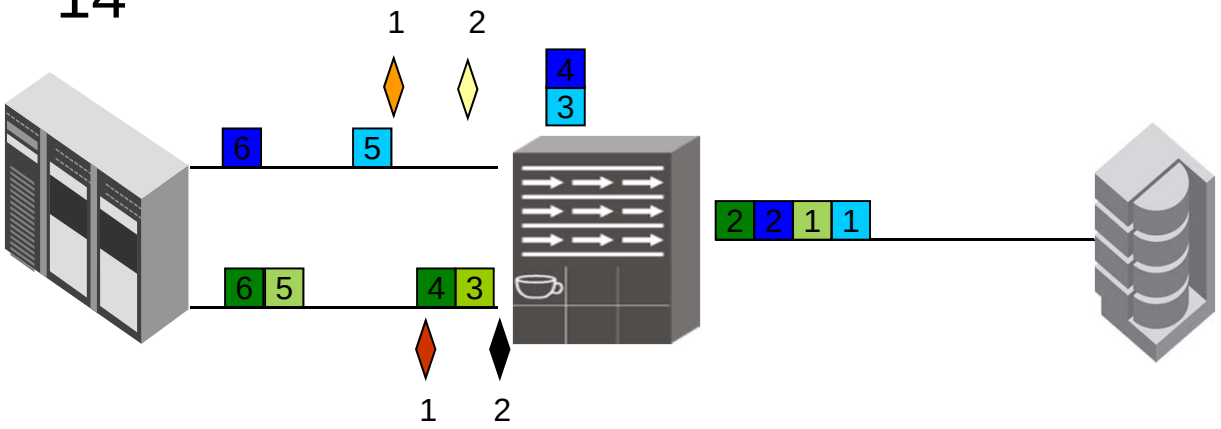
16


15
1
2
5
4
3
5
4
3
1
1
2
1
15



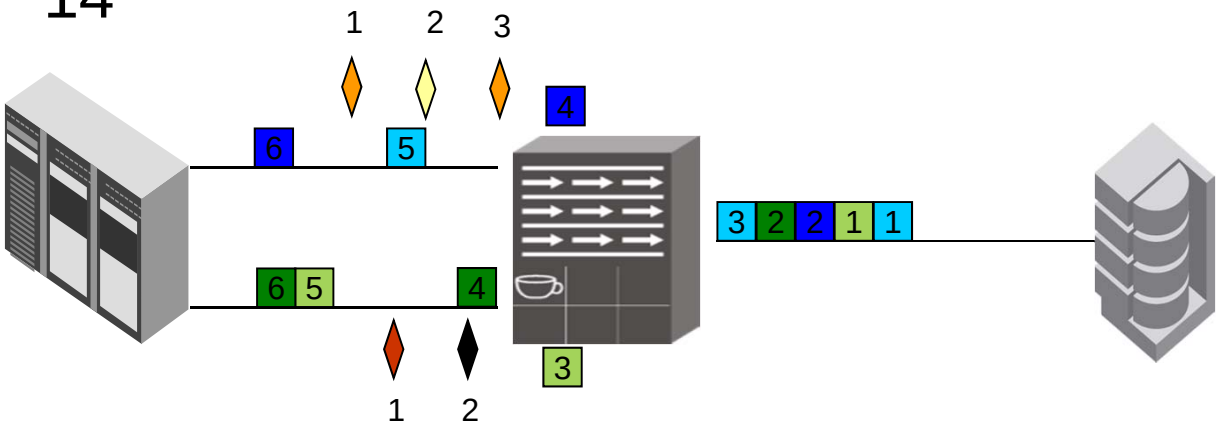
14

14



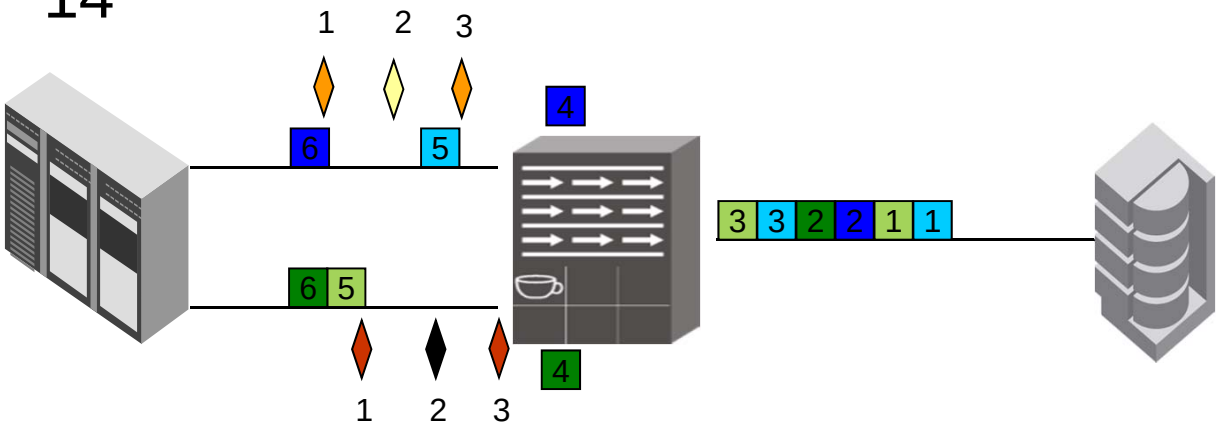
14

14



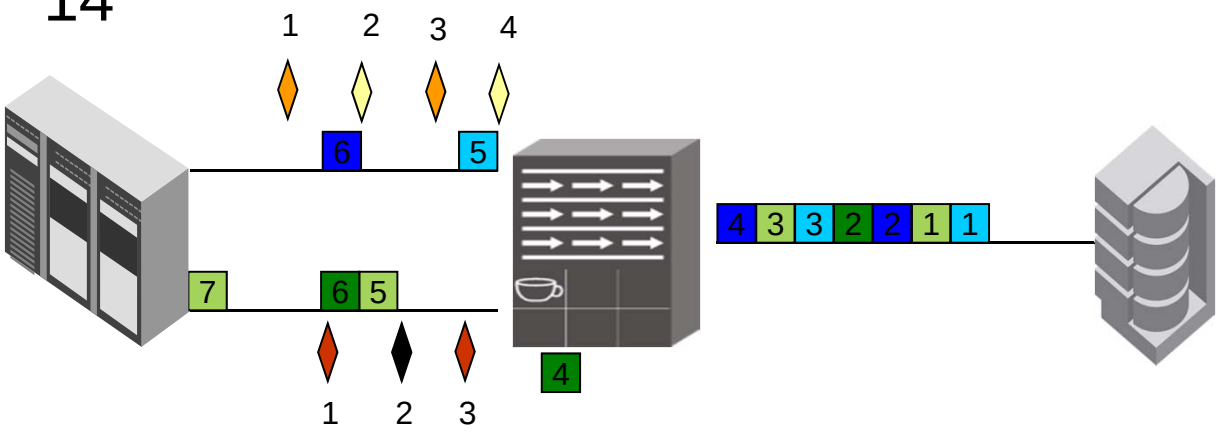
14

14

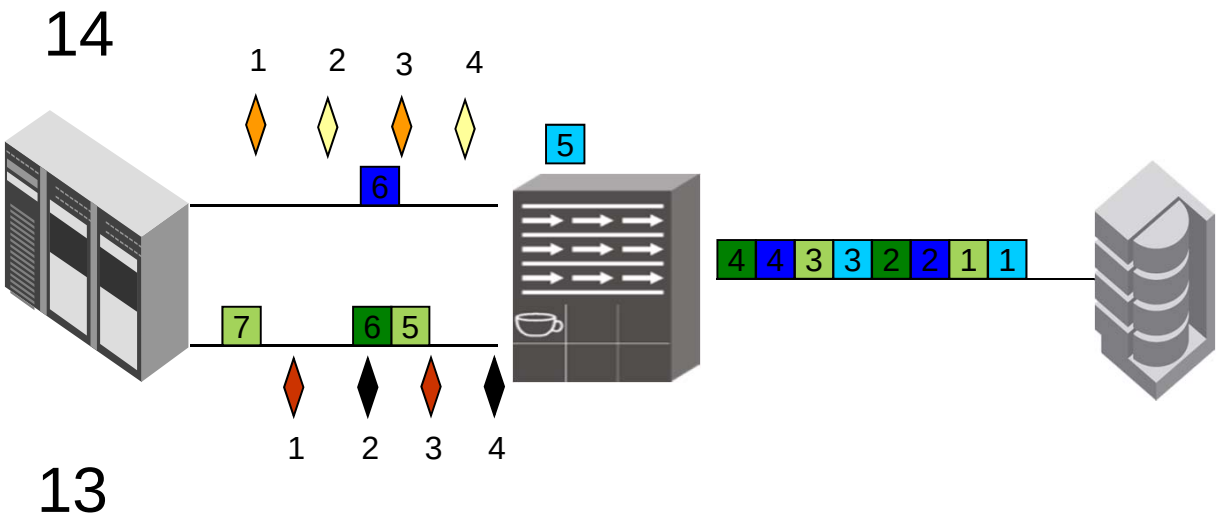


14

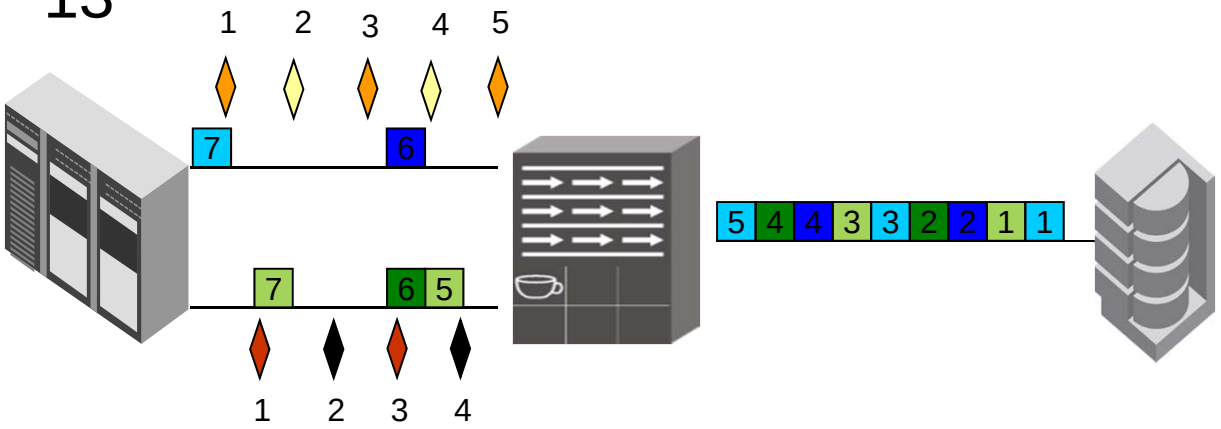
14



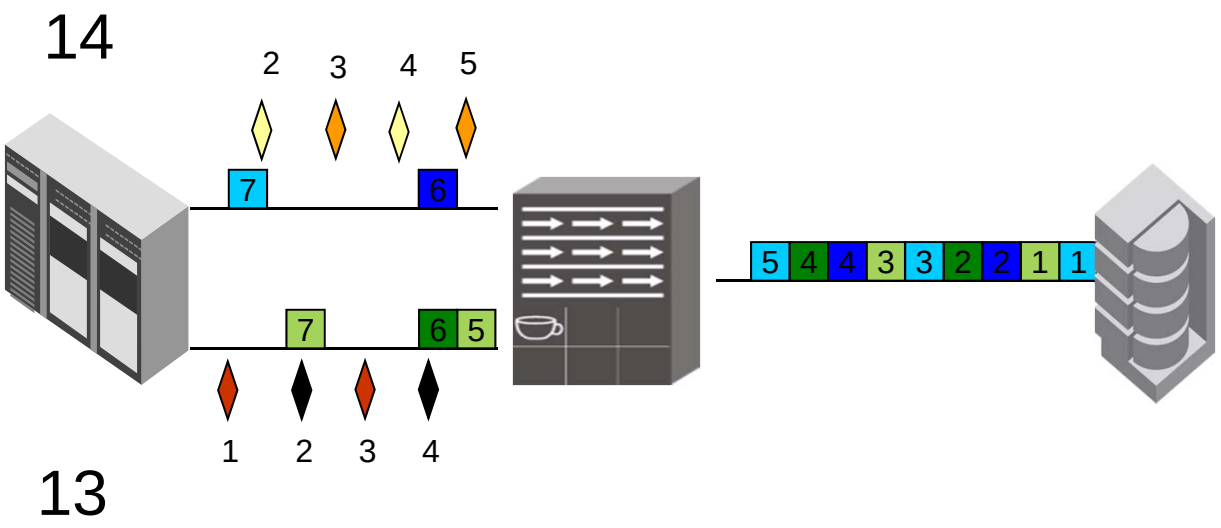
13


14
1
2
3
4
5
6
7
6
5
4
4
3
3
2
2
1
1
1
2
3
4
13

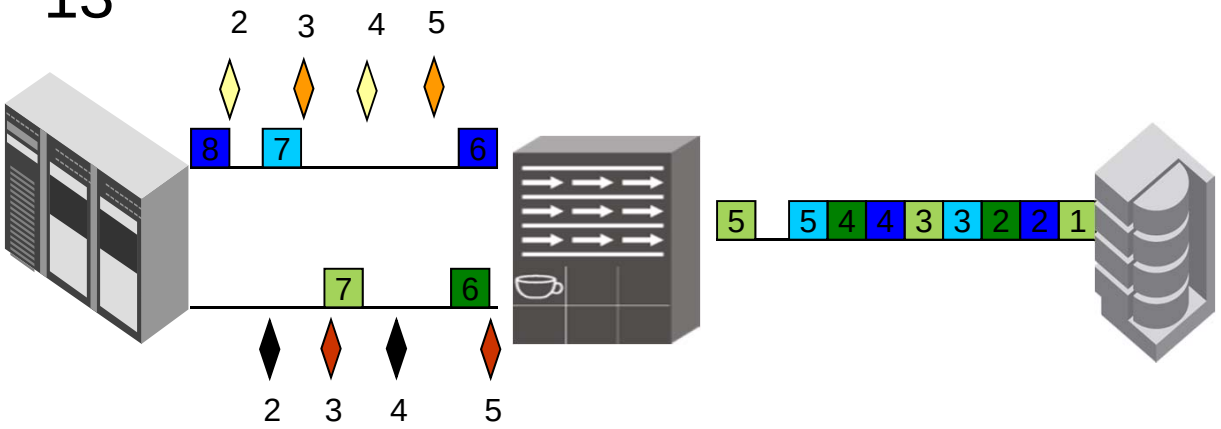
13



13

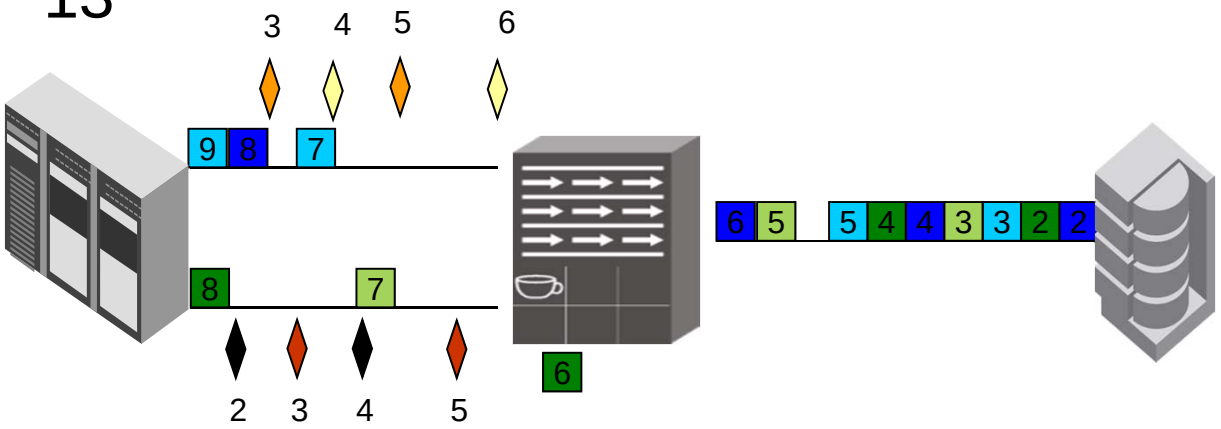

14
2
3
4
5
7
6
7
6
5
1
2
3
4
13
5
4
4
3
3
2
2
1
1

13

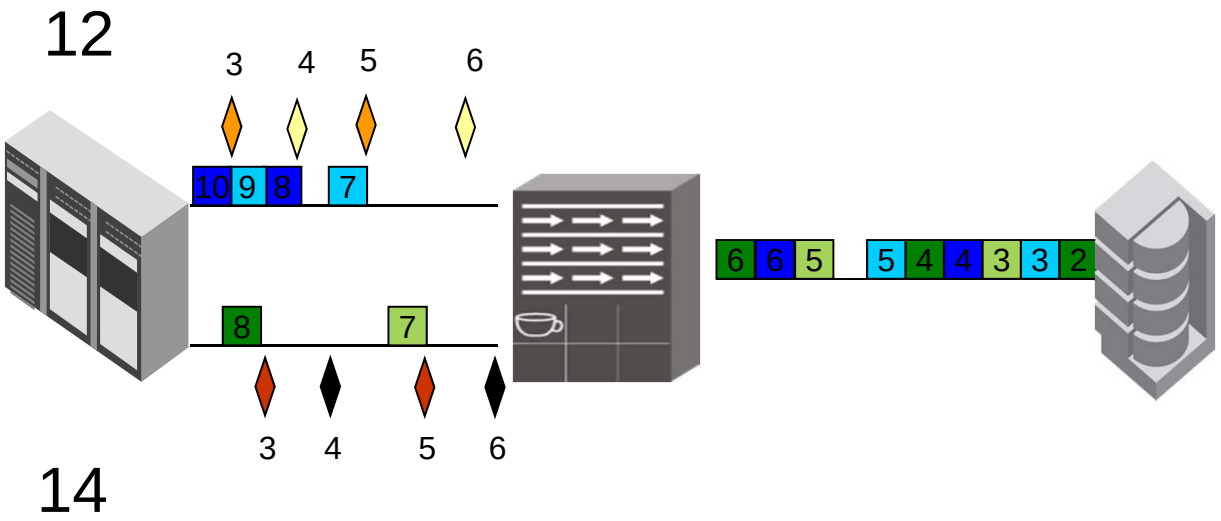


14

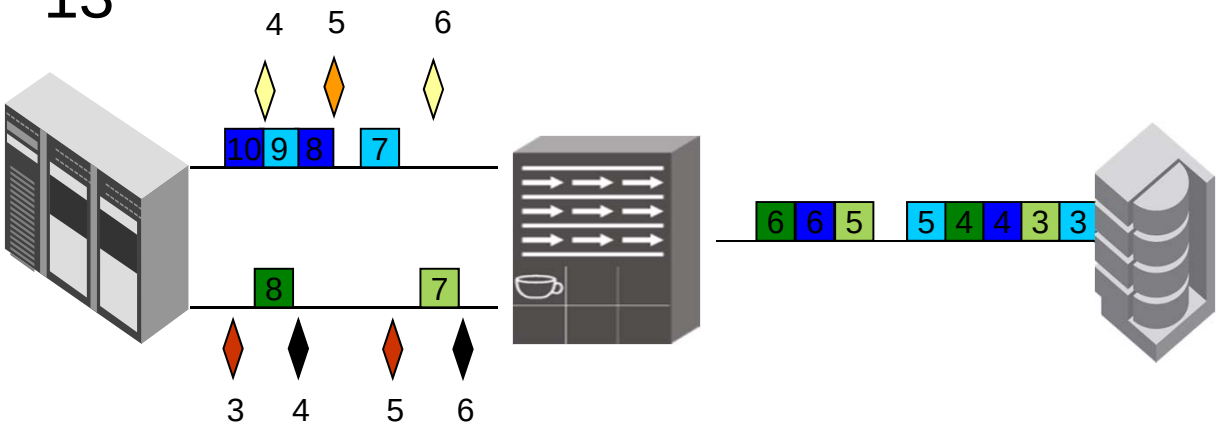
13



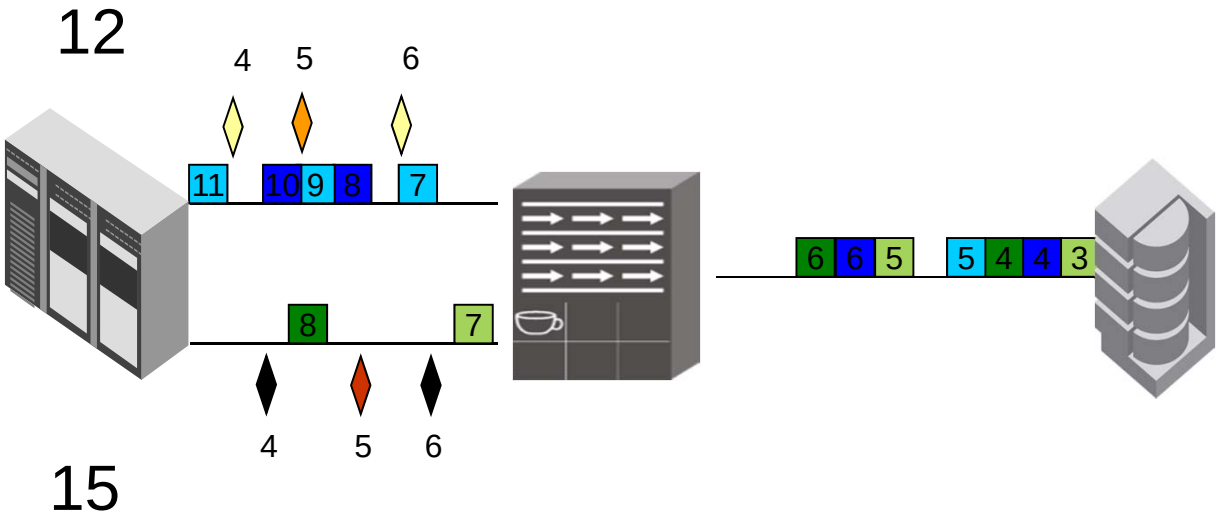
13

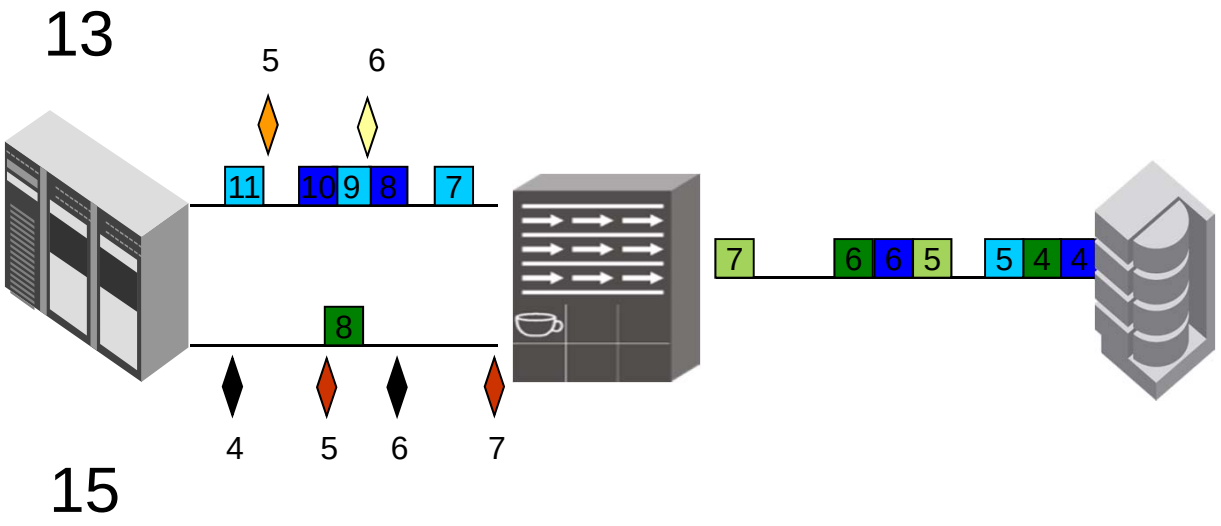

12
3
4
5
6
10
9
8
7
8
7
3
4
5
6
14
6
6
5
5
4
4
3
3
2

13

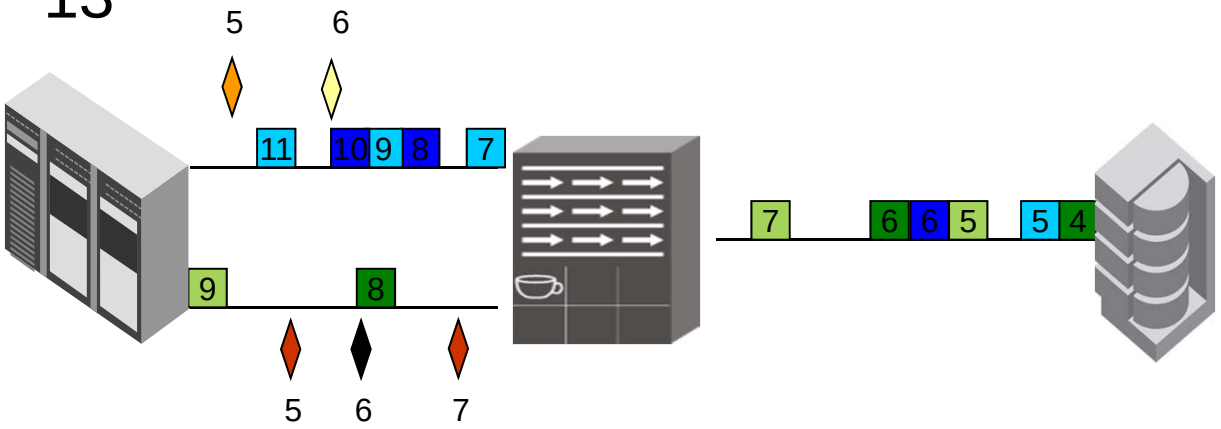


14


12
4
5
6
11
10
9
8
7
8
7
4
5
6
15
6
6
5
5
4
4
3

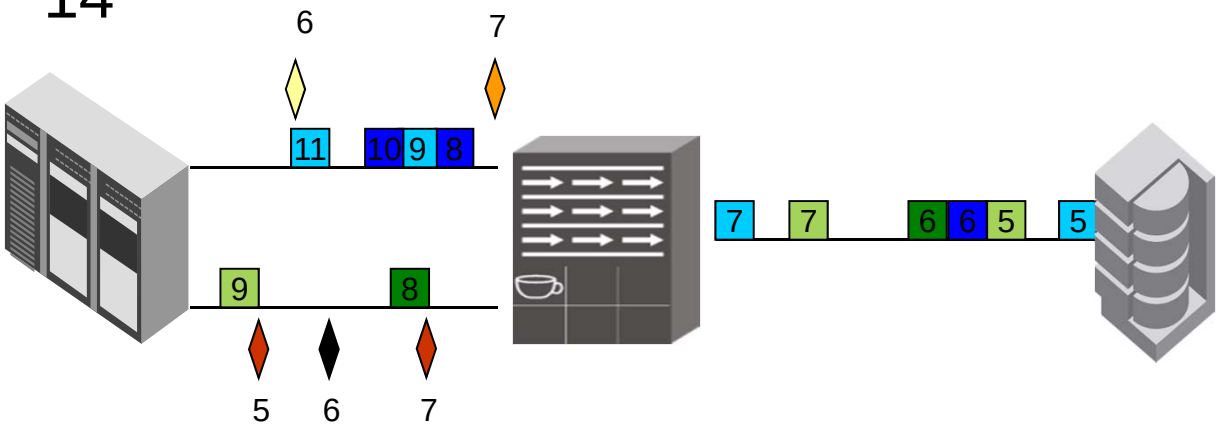

13
5
6
11
10
9
8
7
8
4
5
6
7
15
7
6
6
5
5
4
4

13

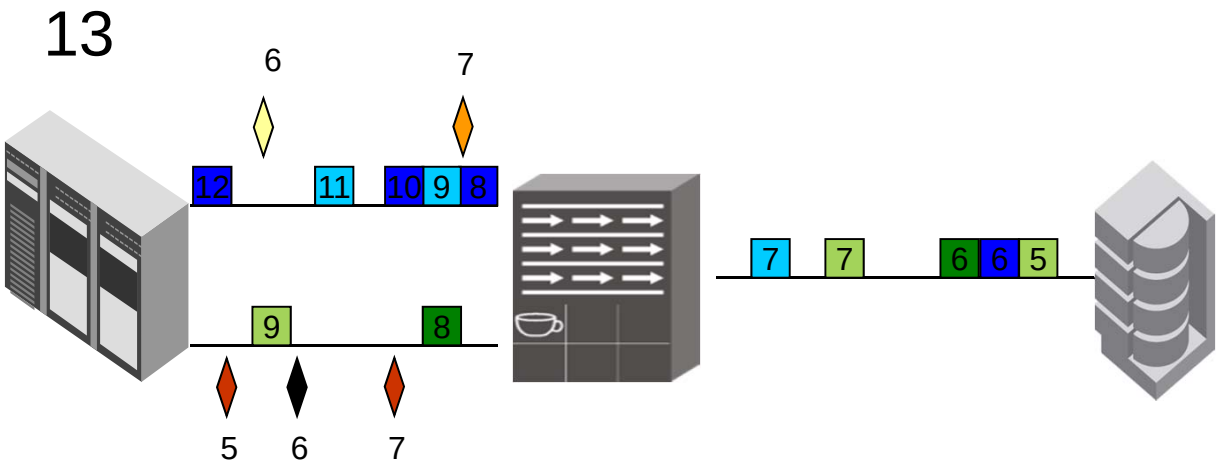


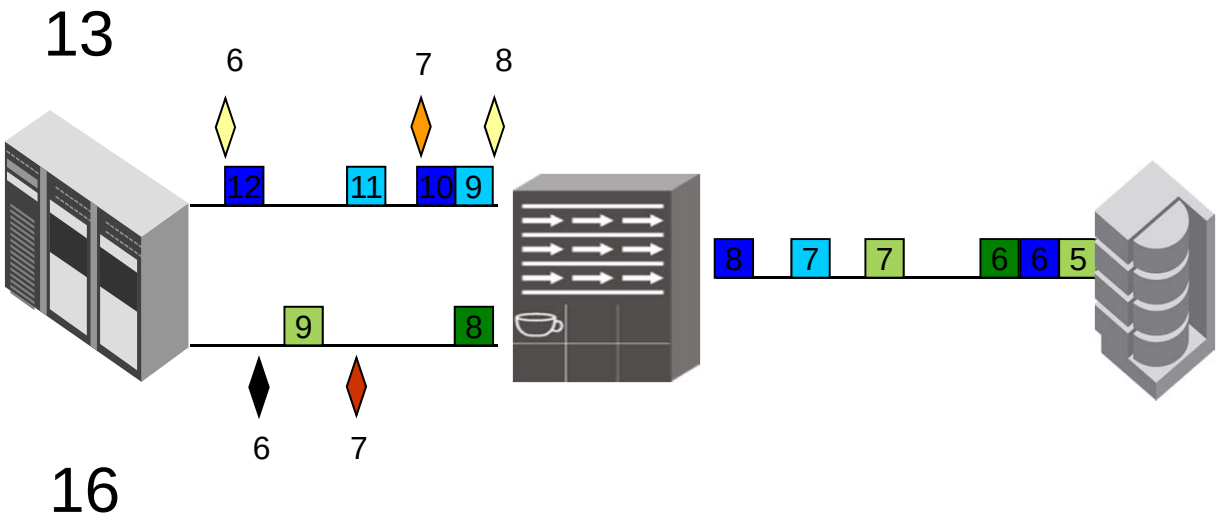
15

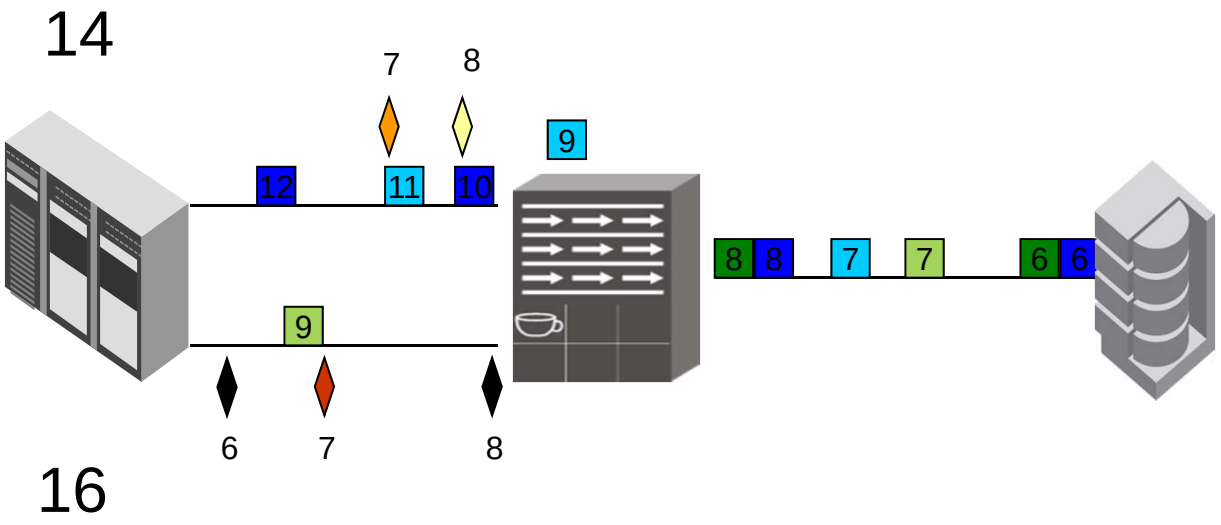
14



15

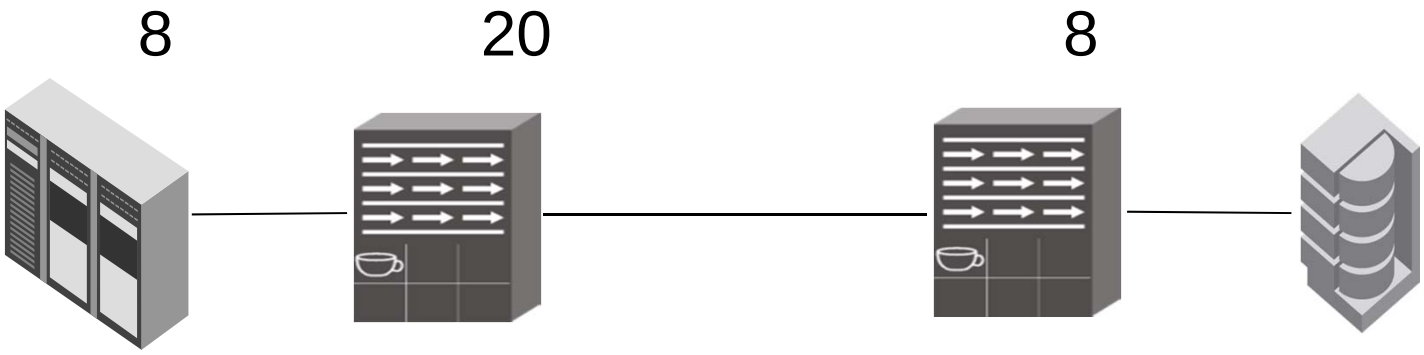

13
6
7
12
11
10
9
8
9
8
5
6
7
7
7
6
6
5
15

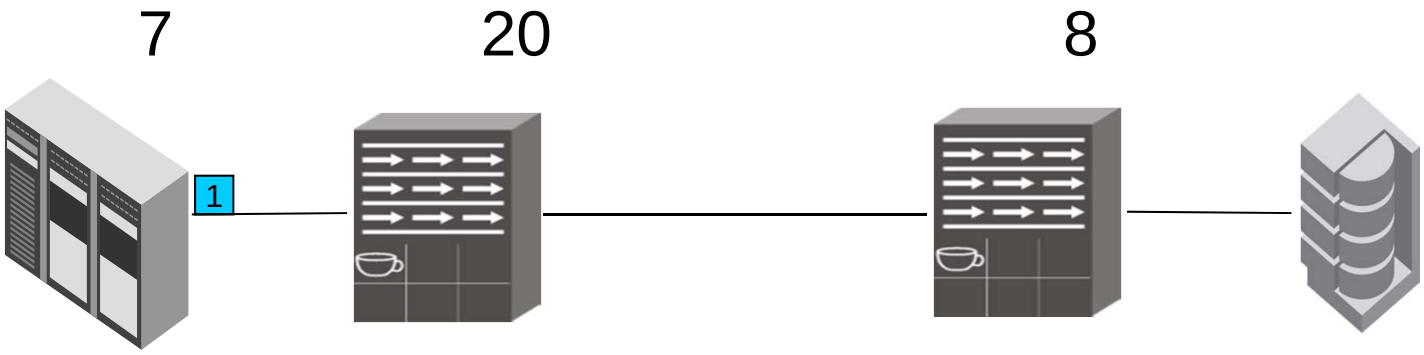

13
6
7
8
12
11
10
9
9
8
6
7
16
8
7
7
6
6
5

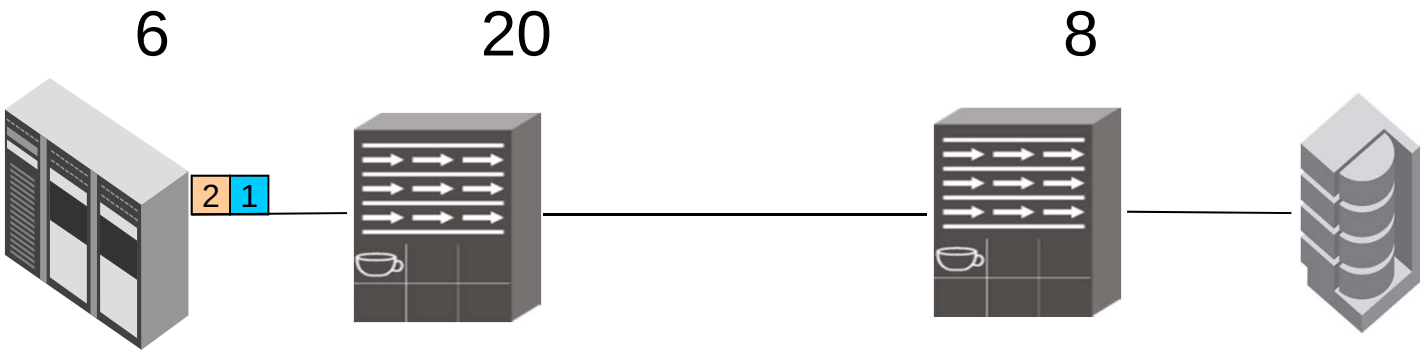

14
7
8
9
12
11
10
9
8
8
7
7
6
6
6
7
8
16

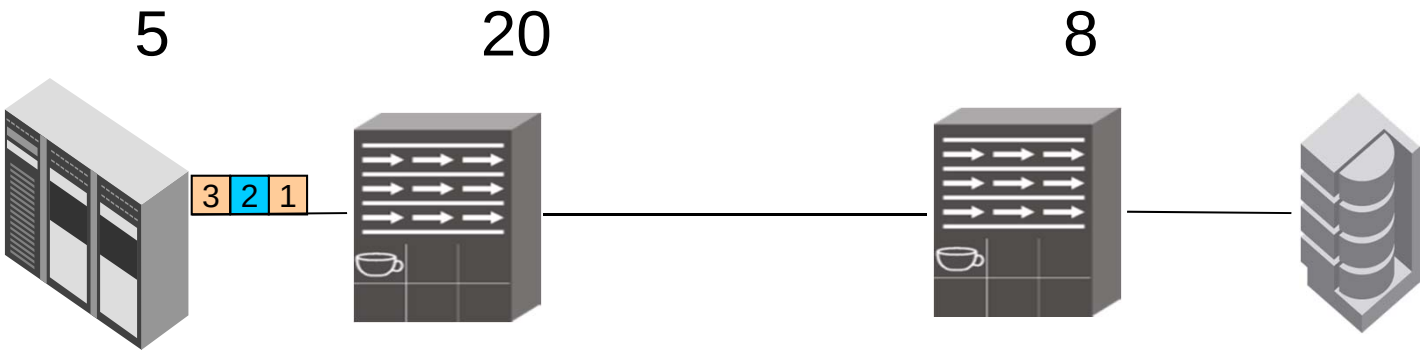
Example: Cascaded Directors

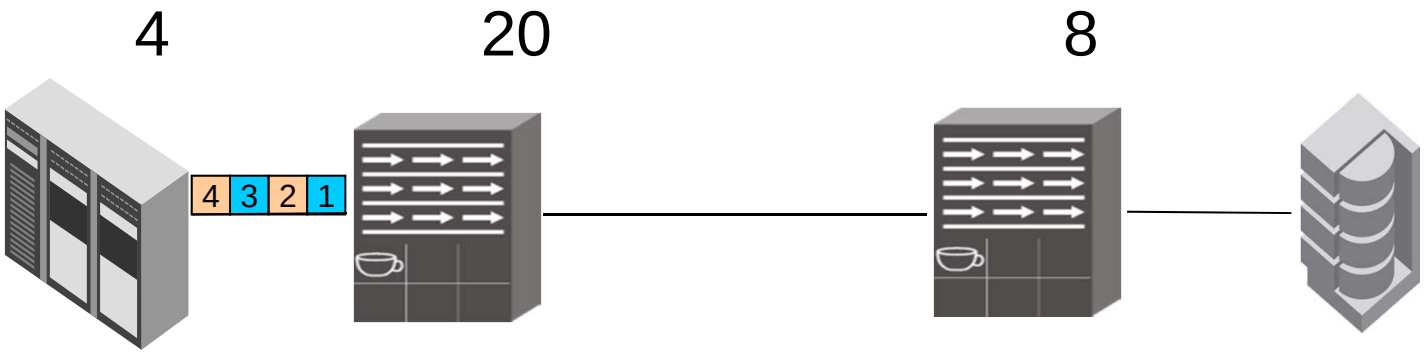
BUFFER CREDITS

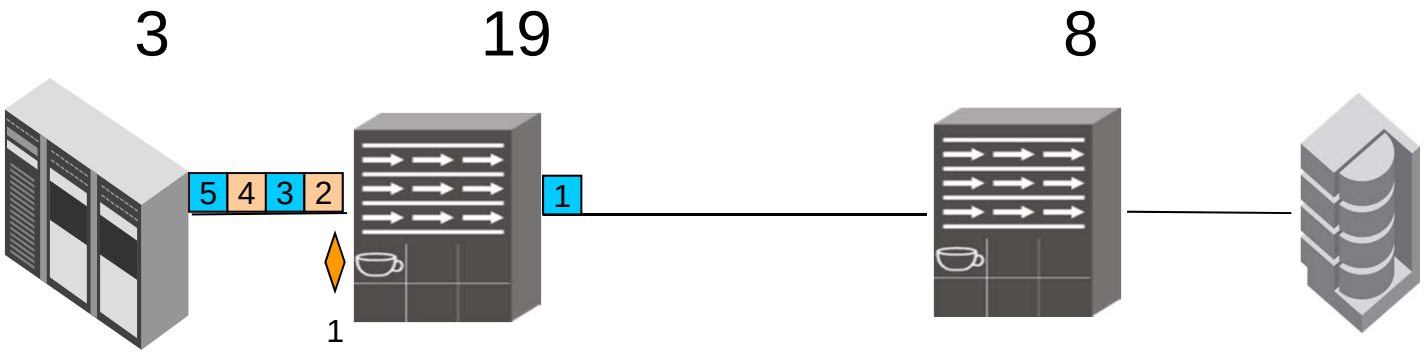

8
20
8

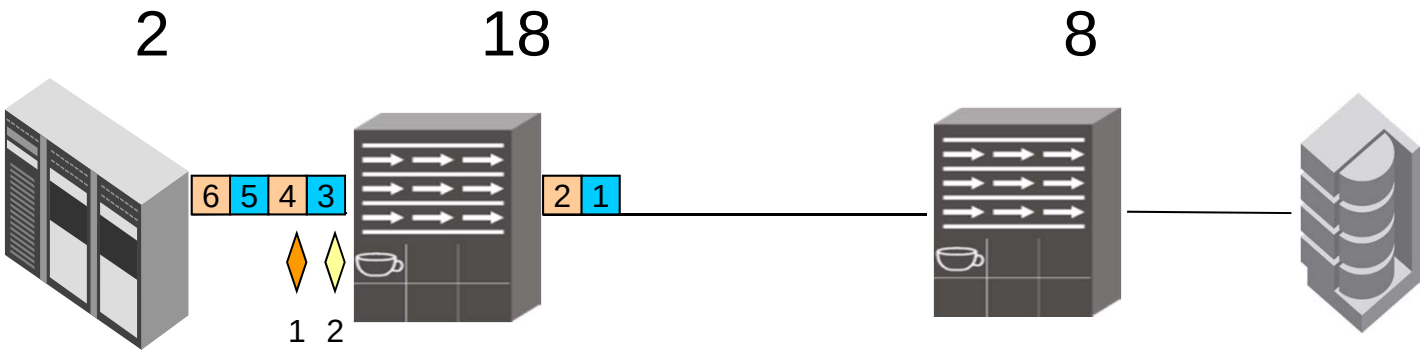

7
20
8
1

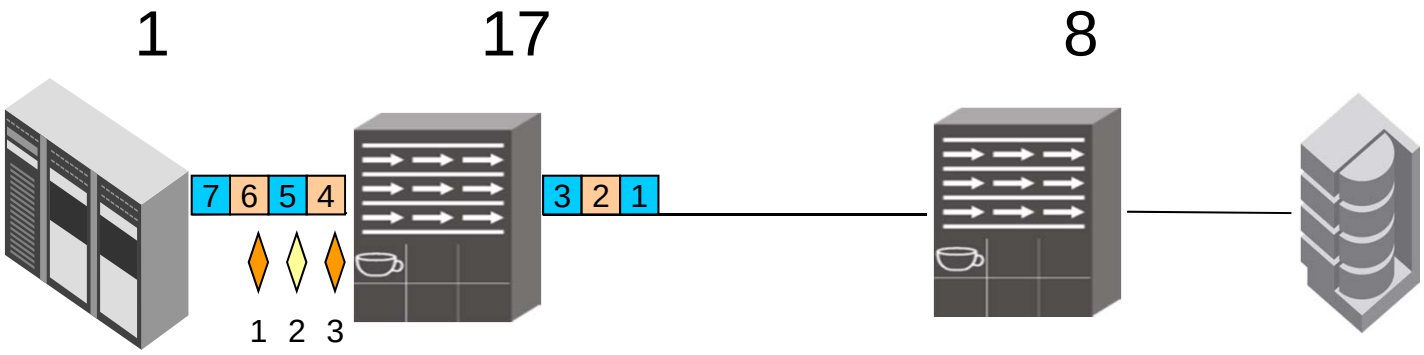

6
20
8
2
1

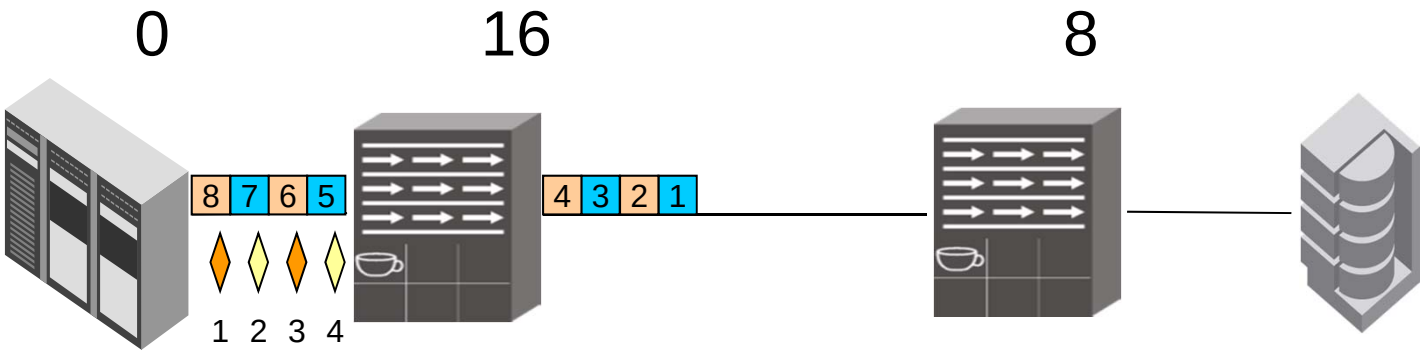

5
20
8
3
2
1

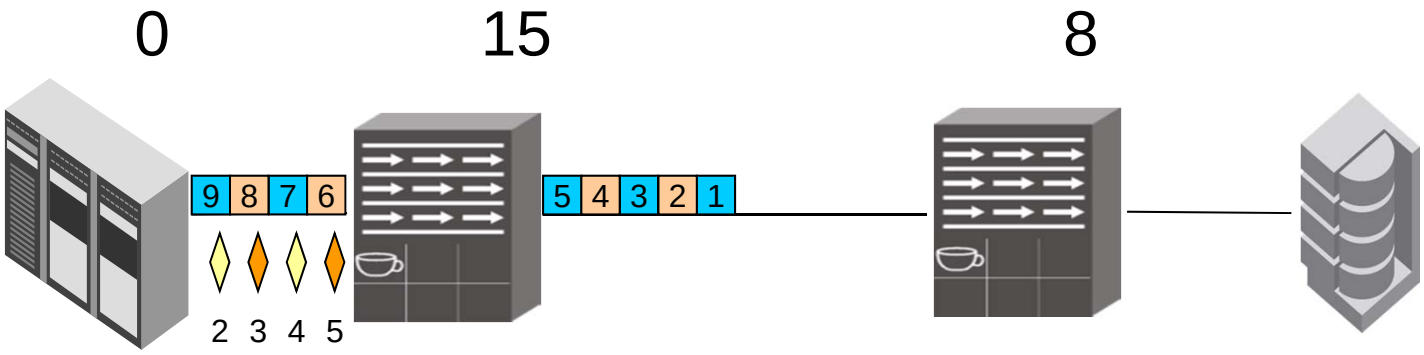

4
20
8
4
3
2
1

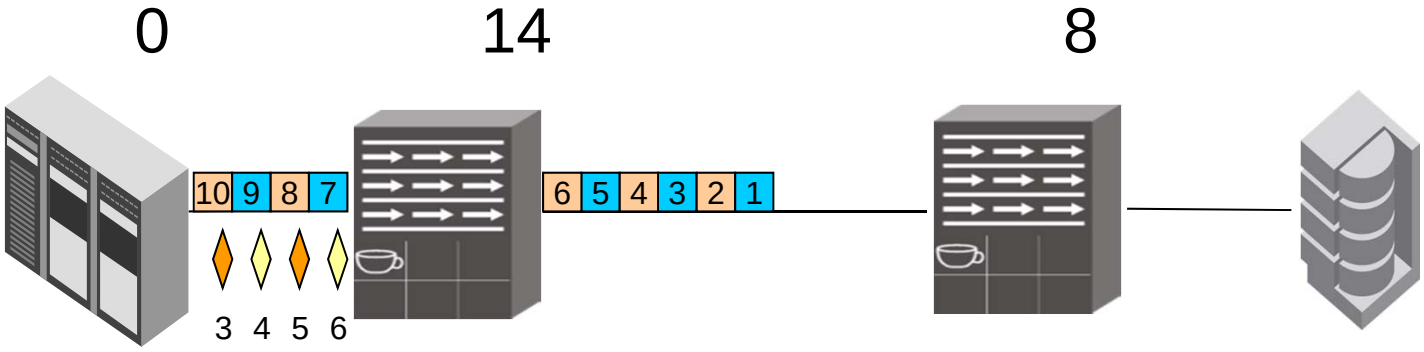

3
19
8
5
4
3
2
1
1

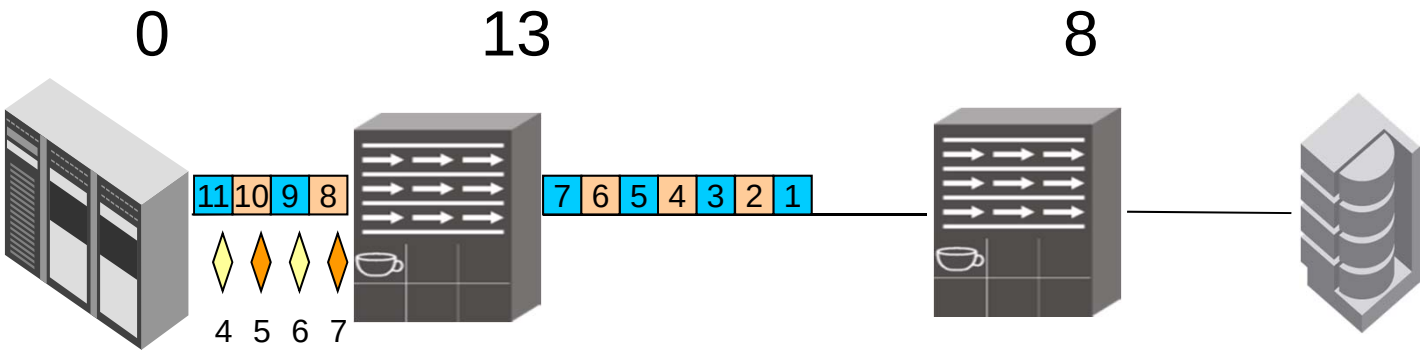

2
18
8
6 5 4 3
2 1
1 2

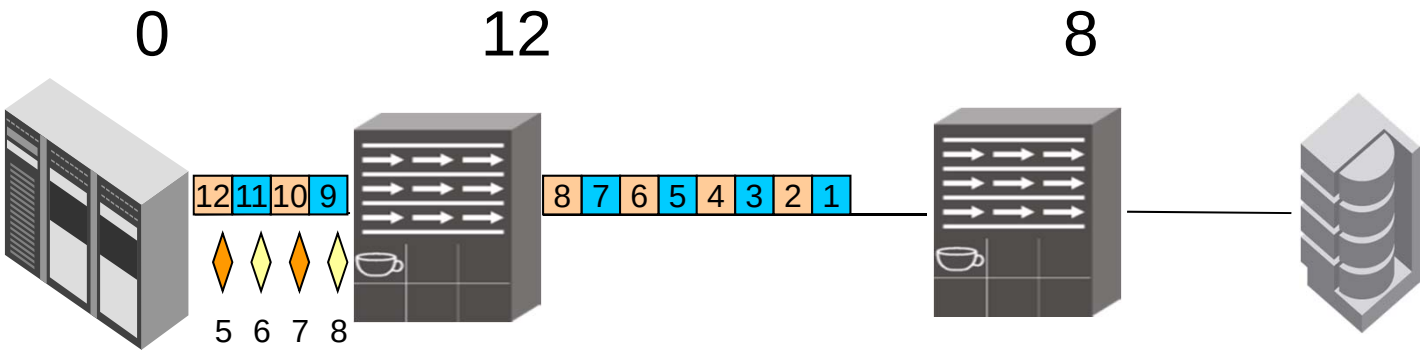

1
17
8
7 6 5 4
3 2 1
1 2 3

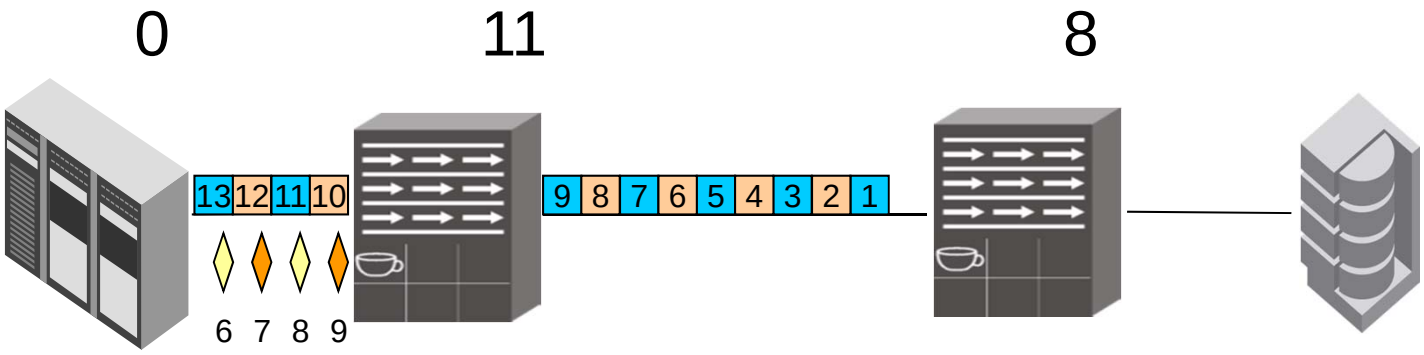

0
16
8
8 7 6 5
4 3 2 1
1 2 3 4

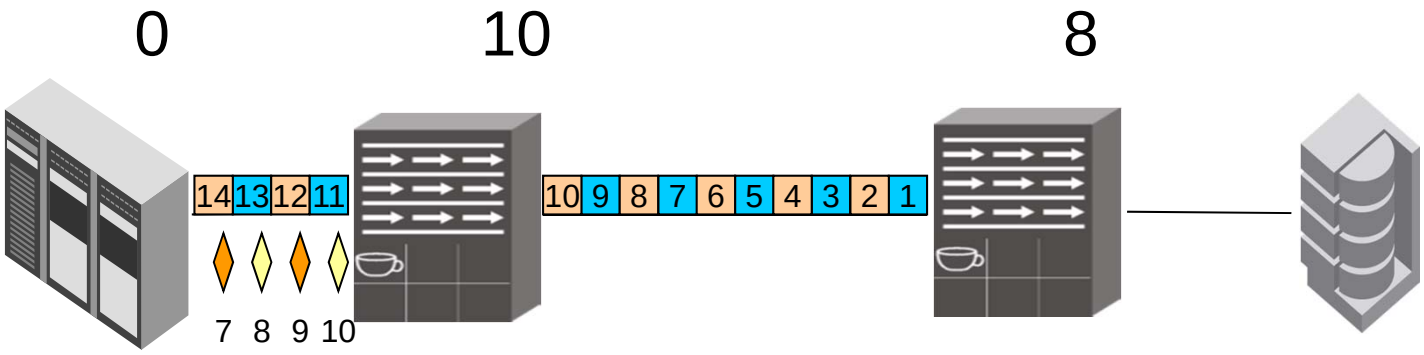

0
15
8
9 8 7 6
5 4 3 2 1
2 3 4 5

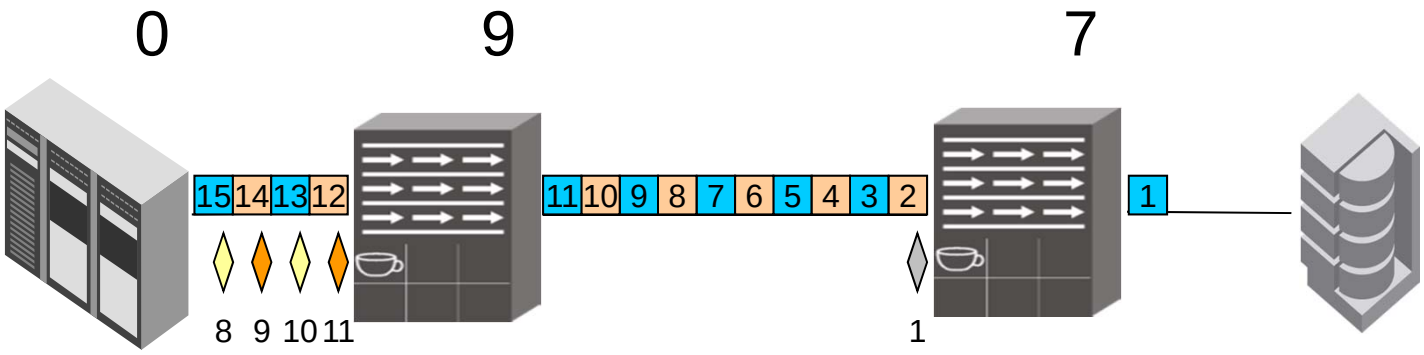

0
14
8
10 9 8 7
6 5 4 3 2 1
3 4 5 6

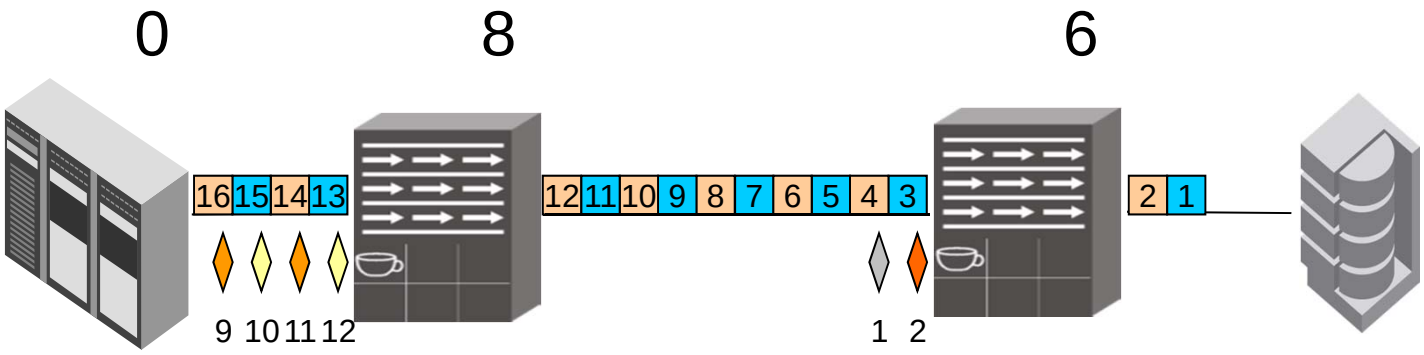

0
13
8
11 10 9 8
7 6 5 4 3 2 1
4 5 6 7

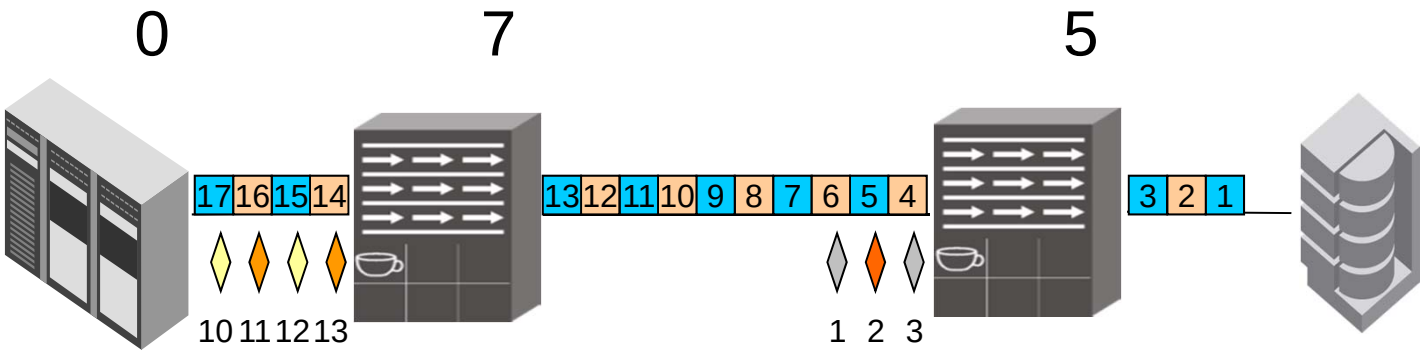

0
12
8
12 11 10 9
8 7 6 5 4 3 2 1
5 6 7 8

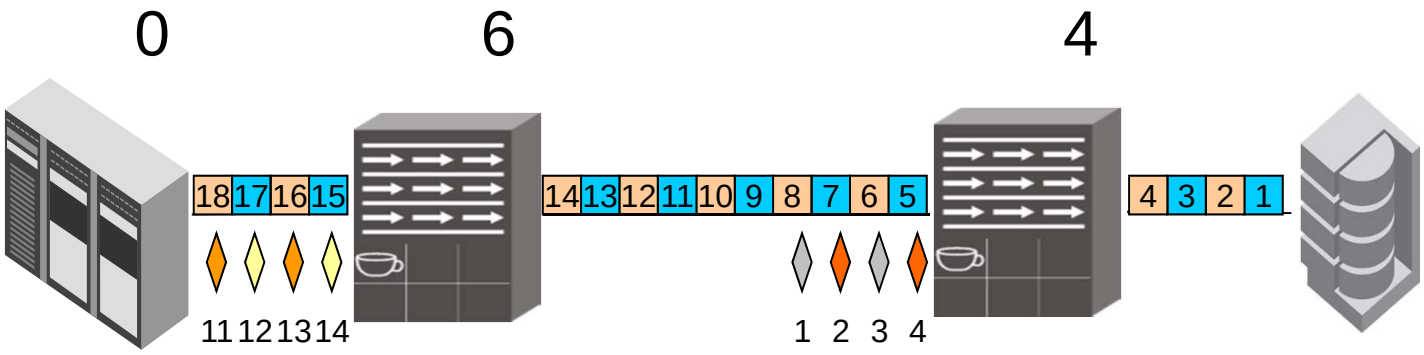

0
11
8
13 12 11 10
9 8 7 6 5 4 3 2 1
6 7 8 9

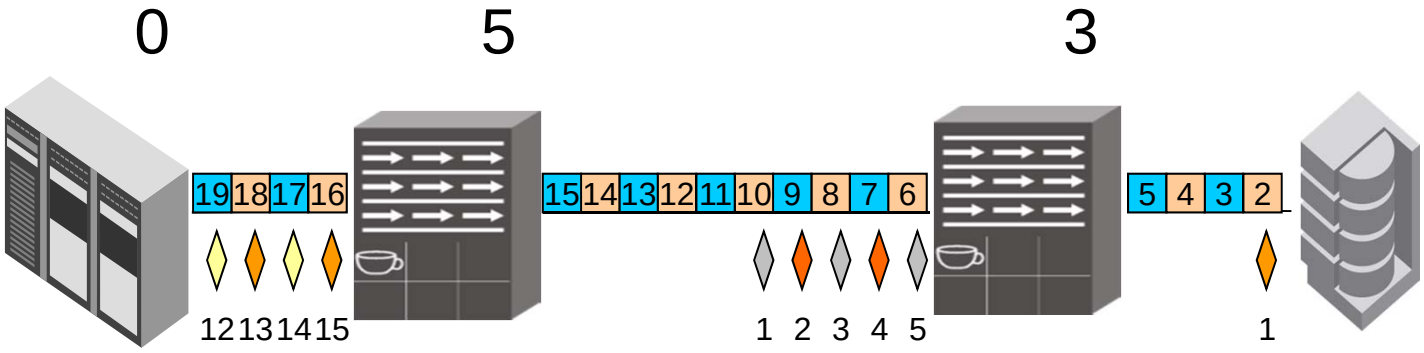

0
10
8
14 13 12 11
10 9 8 7 6 5 4 3 2 1
7 8 9 10

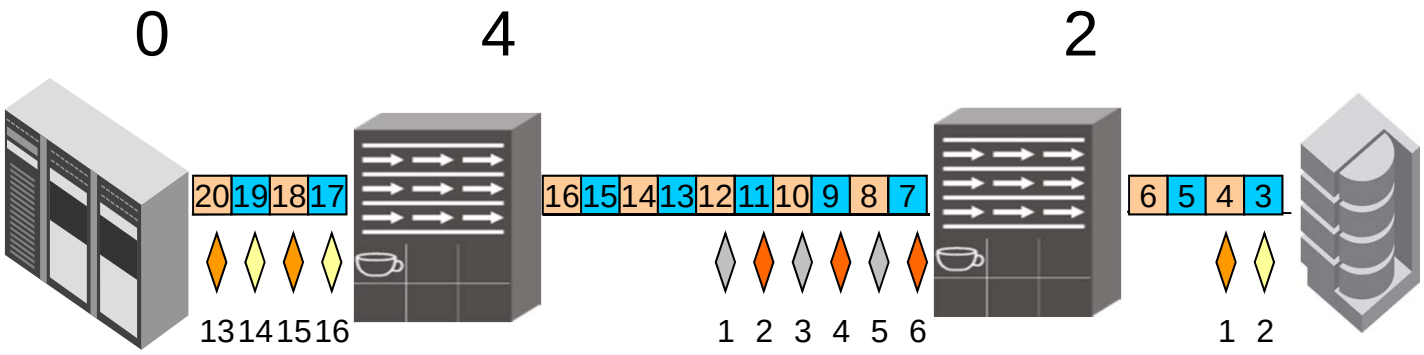

0
9
7
15 14 13 12
11 10 9 8 7 6 5 4 3 2
1
8 9 10 11
1

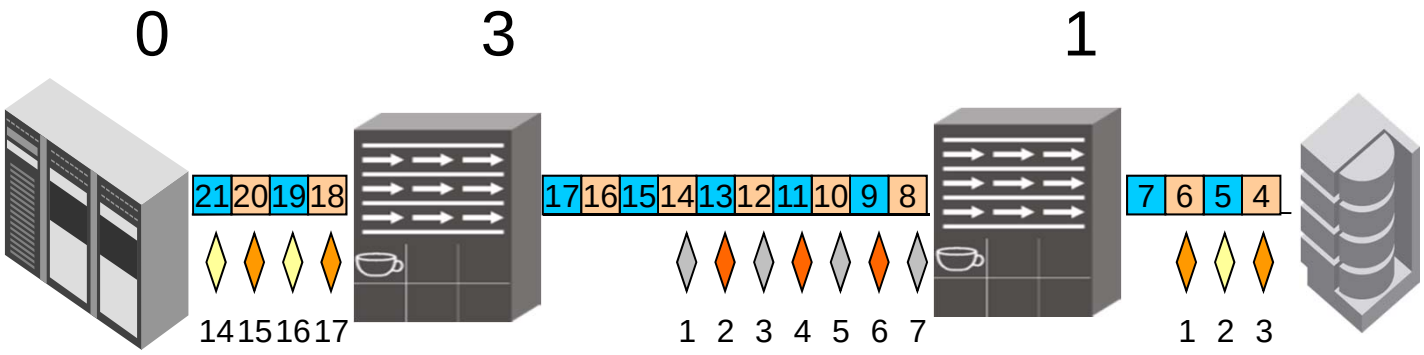

0
8
6
16 15 14 13
12 11 10 9 8 7 6 5 4 3
2 1
9 10 11 12
1 2

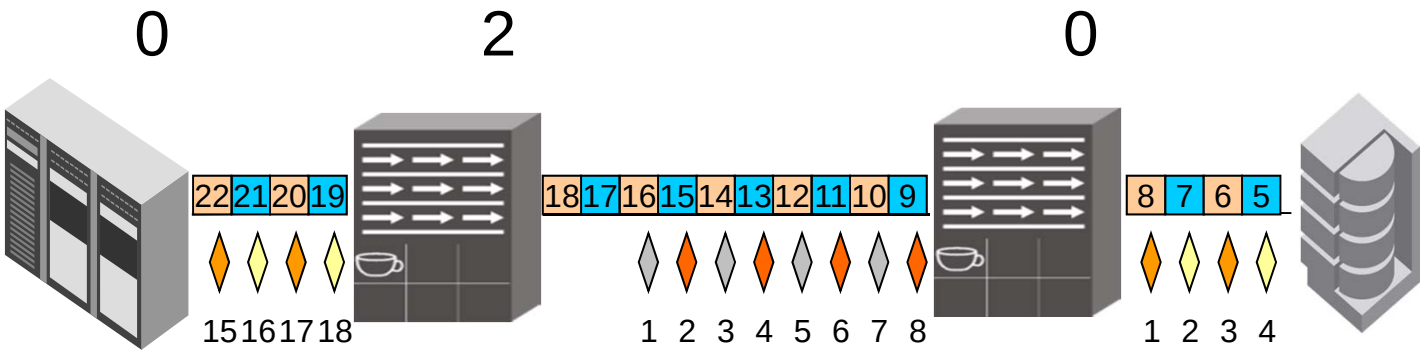

0
7
5
17 16 15 14
13 12 11 10 9 8 7 6 5 4
3 2 1
10 11 12 13
1 2 3

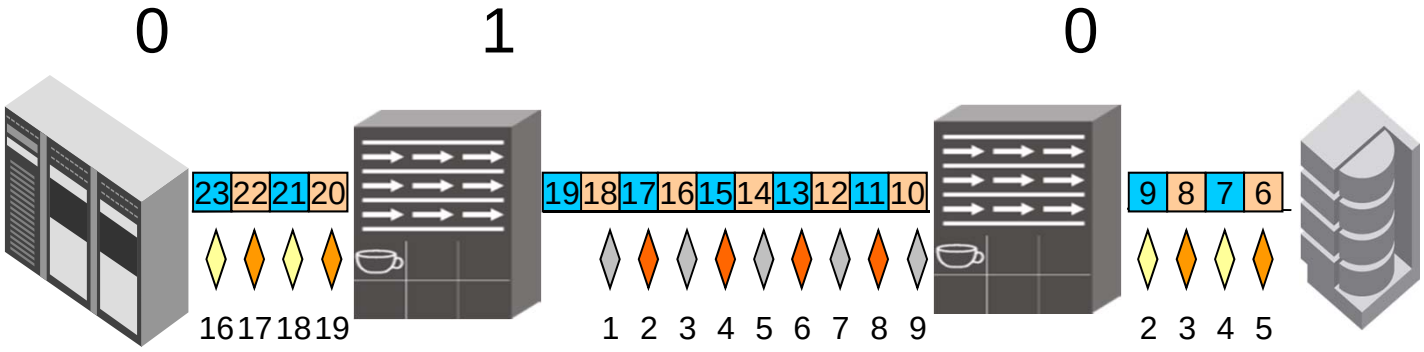

0
6
4
18 17 16 15
14 13 12 11 10 9 8 7 6 5
4 3 2 1
11 12 13 14
1 2 3 4

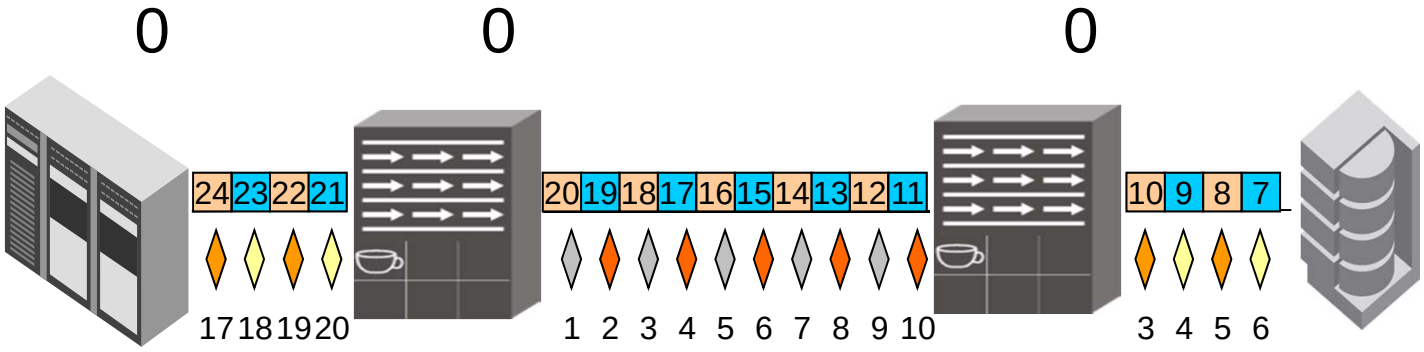

0
5
3
19 18 17 16
15 14 13 12 11 10 9 8 7 6
5 4 3 2
12 13 14 15
1 2 3 4 5
1

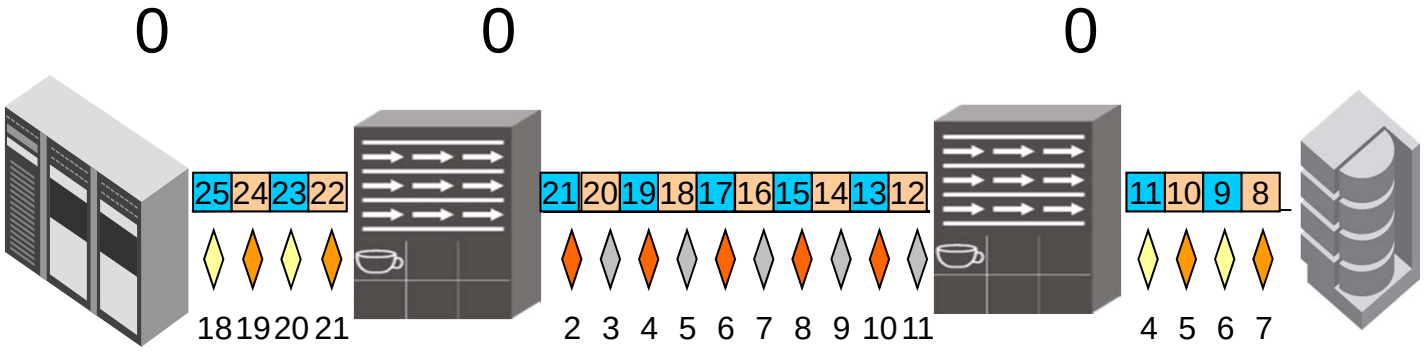

0
4
2
20 19 18 17
16 15 14 13 12 11 10 9 8 7
6 5 4 3
13 14 15 16
1 2 3 4 5 6
1 2


0
3
1
21 20 19 18
17 16 15 14 13 12 11 10 9 8
7 6 5 4
14 15 16 17
1 2 3 4 5 6 7
1 2 3


0
2
0
22 21 20 19
18 17 16 15 14 13 12 11 10 9
8 7 6 5
15 16 17 18
1 2 3 4 5 6 7 8
1 2 3 4


0
1
0
23 22 21 20
19 18 17 16 15 14 13 12 11 10
9 8 7 6
16 17 18 19
1 2 3 4 5 6 7 8 9
2 3 4 5

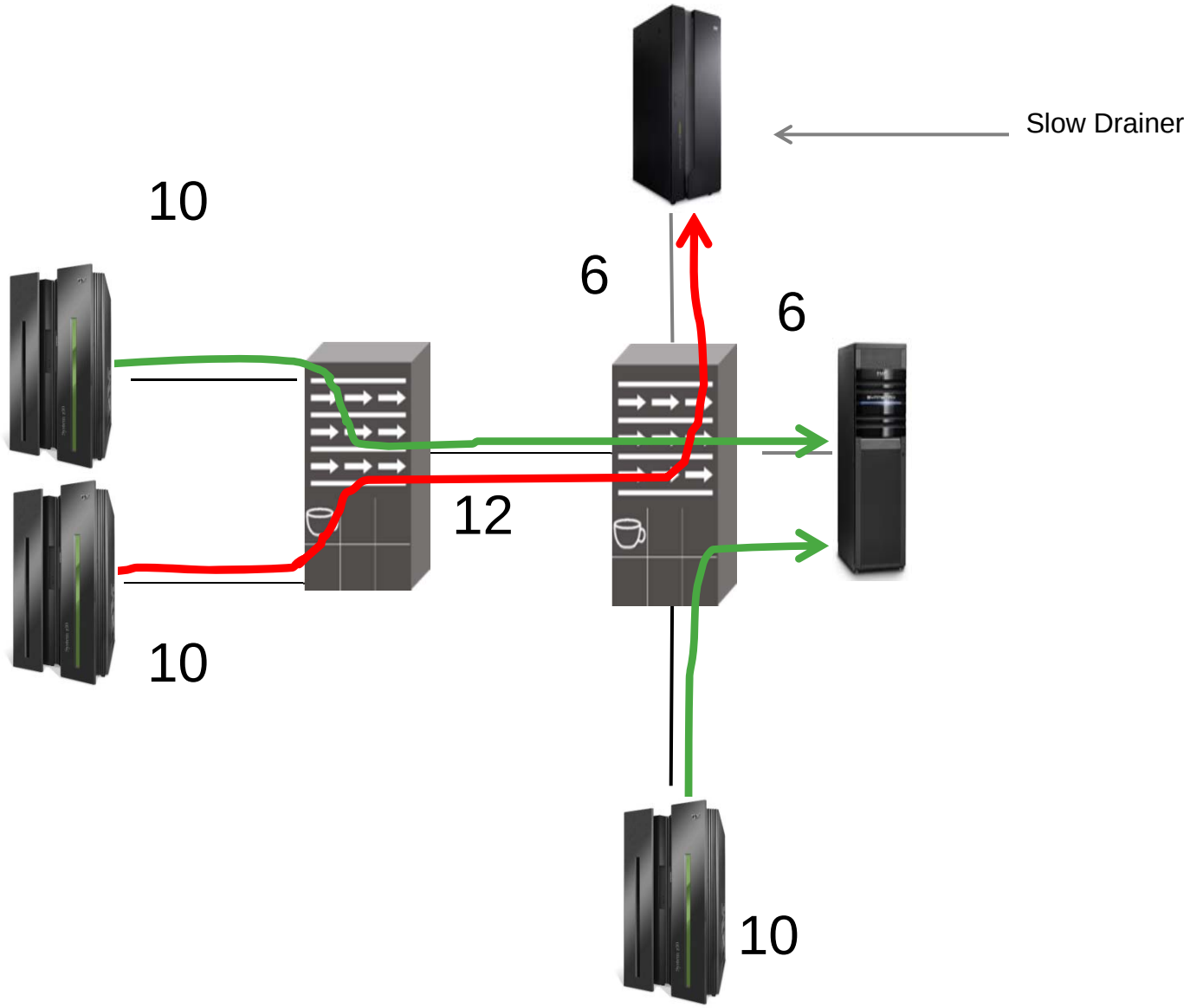

0
0
0
24 23 22 21
20 19 18 17 16 15 14 13 12 11
10 9 8 7
17 18 19 20
1 2 3 4 5 6 7 8 9 10
3 4 5 6

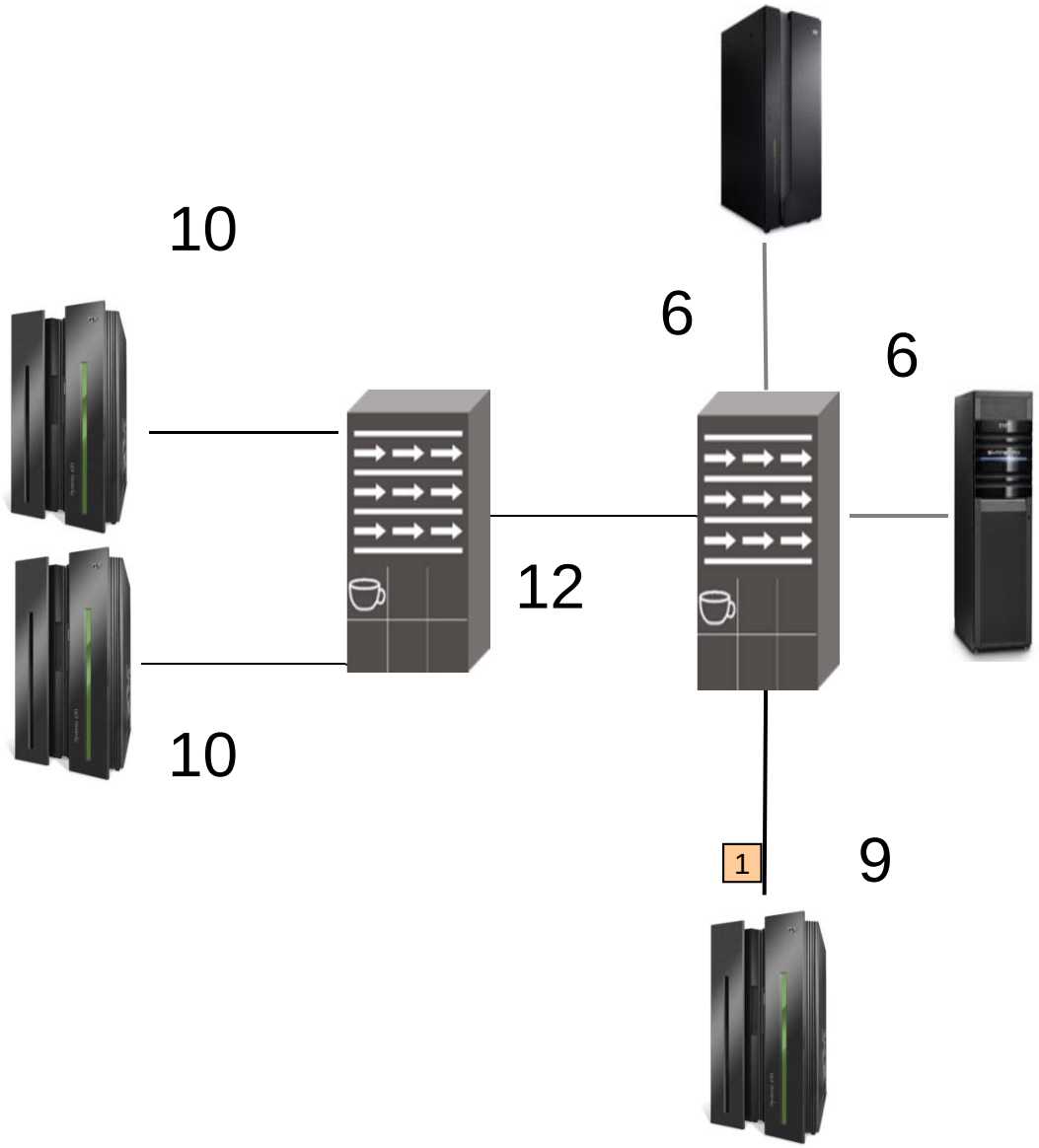

0
0
0
25 24 23 22
21 20 19 18 17 16 15 14 13 12
11 10 9 8
18 19 20 21
2 3 4 5 6 7 8 9 10 11
4 5 6 7

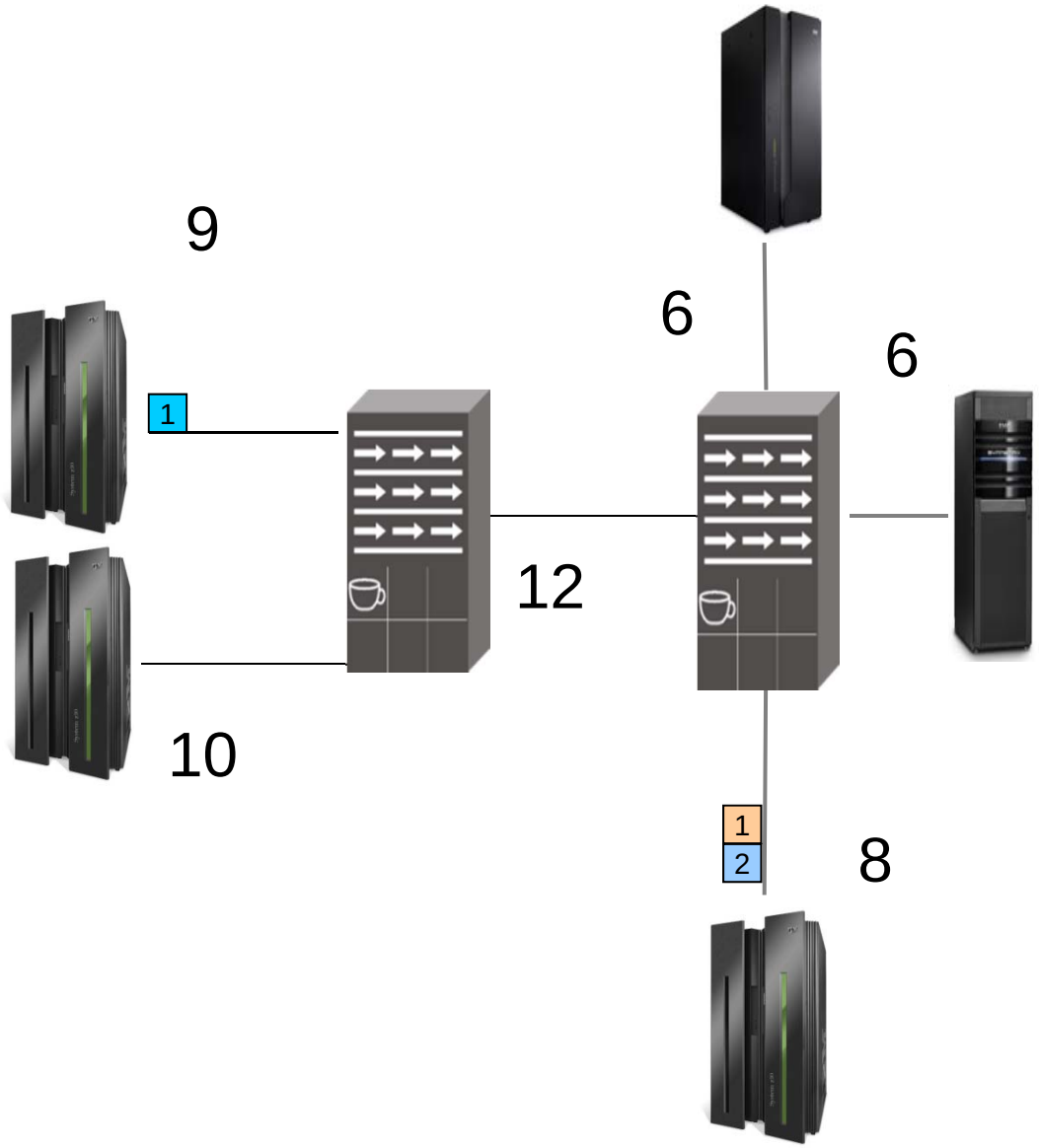
THIS PAGE INTENTIONALLY LEFT BLANK

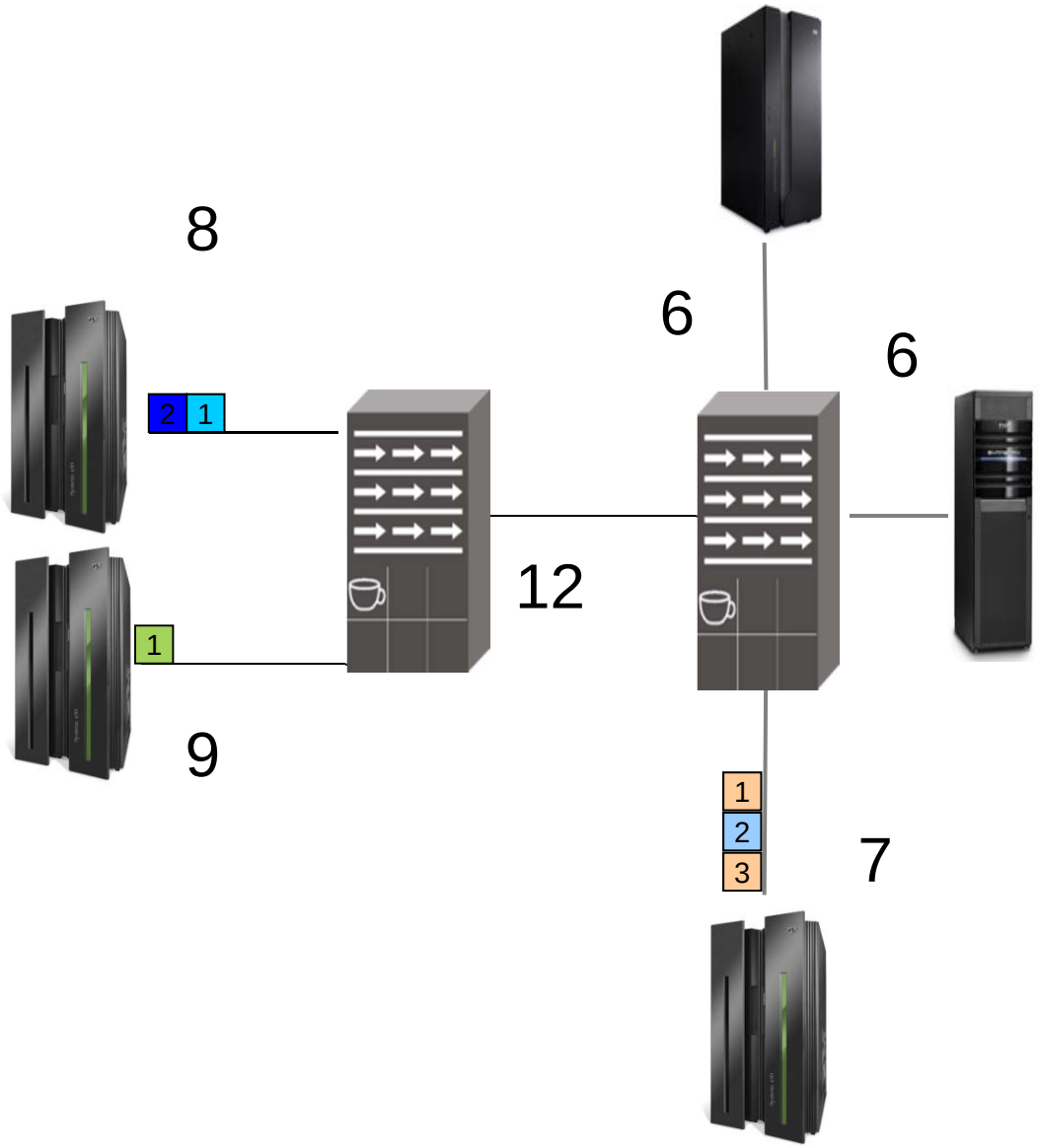
Slow Drainers

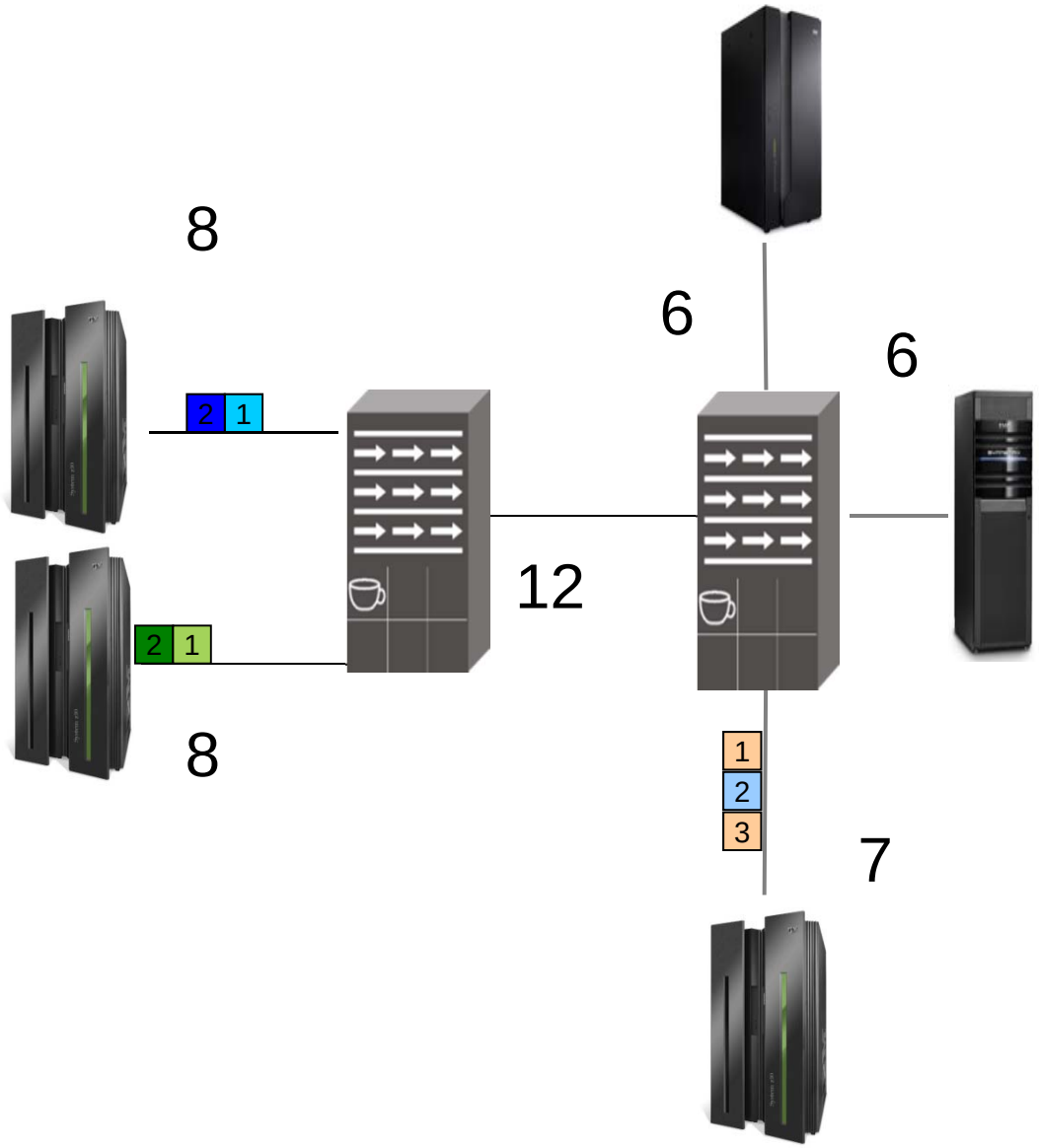
- A device that delays the return of BB credit significantly slower than the desired load
- Moderate delays can cause performance impacts even to other uninvolved devices
- Severe delays can cause errors (IFCCs), even to other uninvolved devices

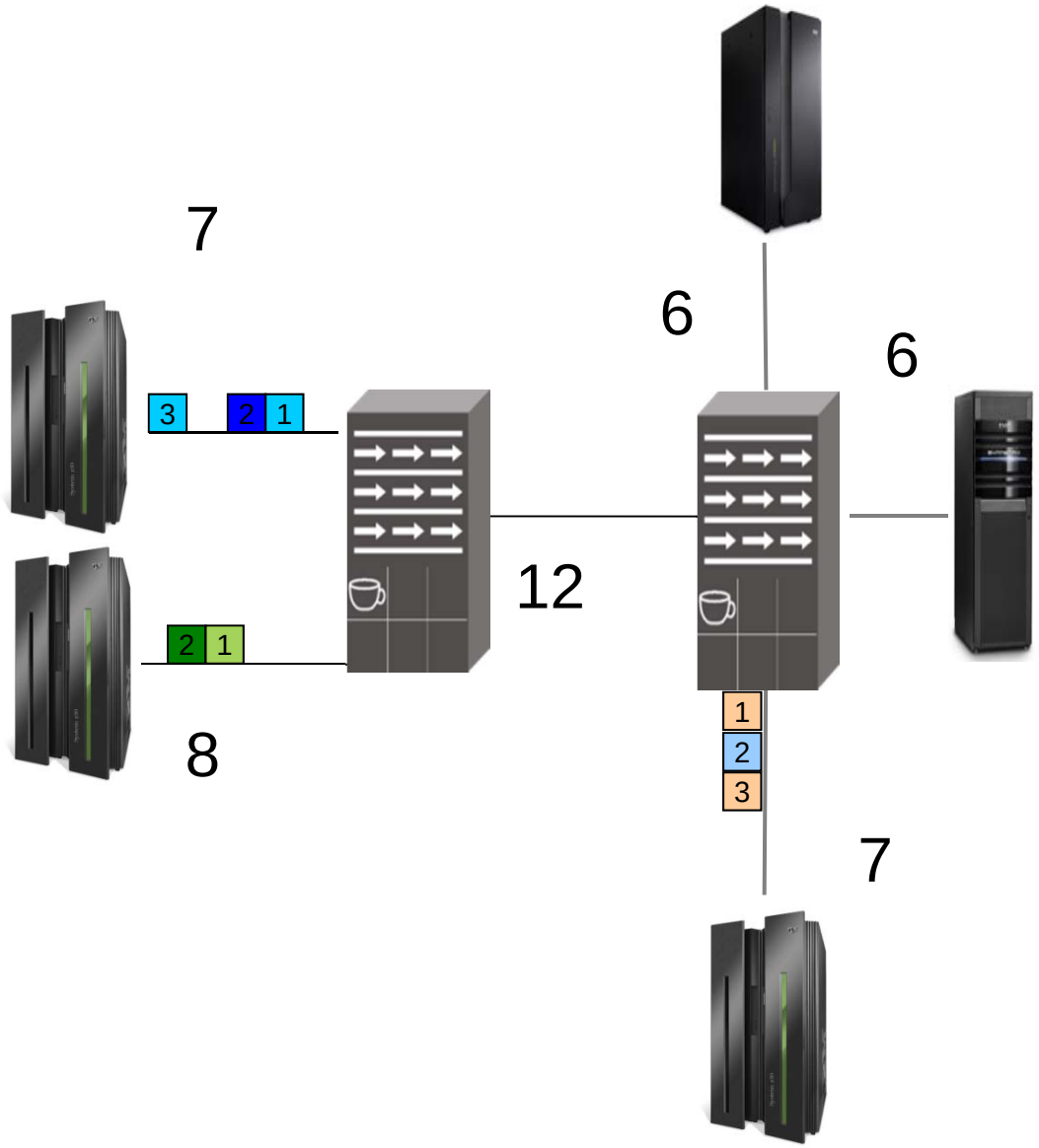

Slow Drainer
10
6
6
12
10
10

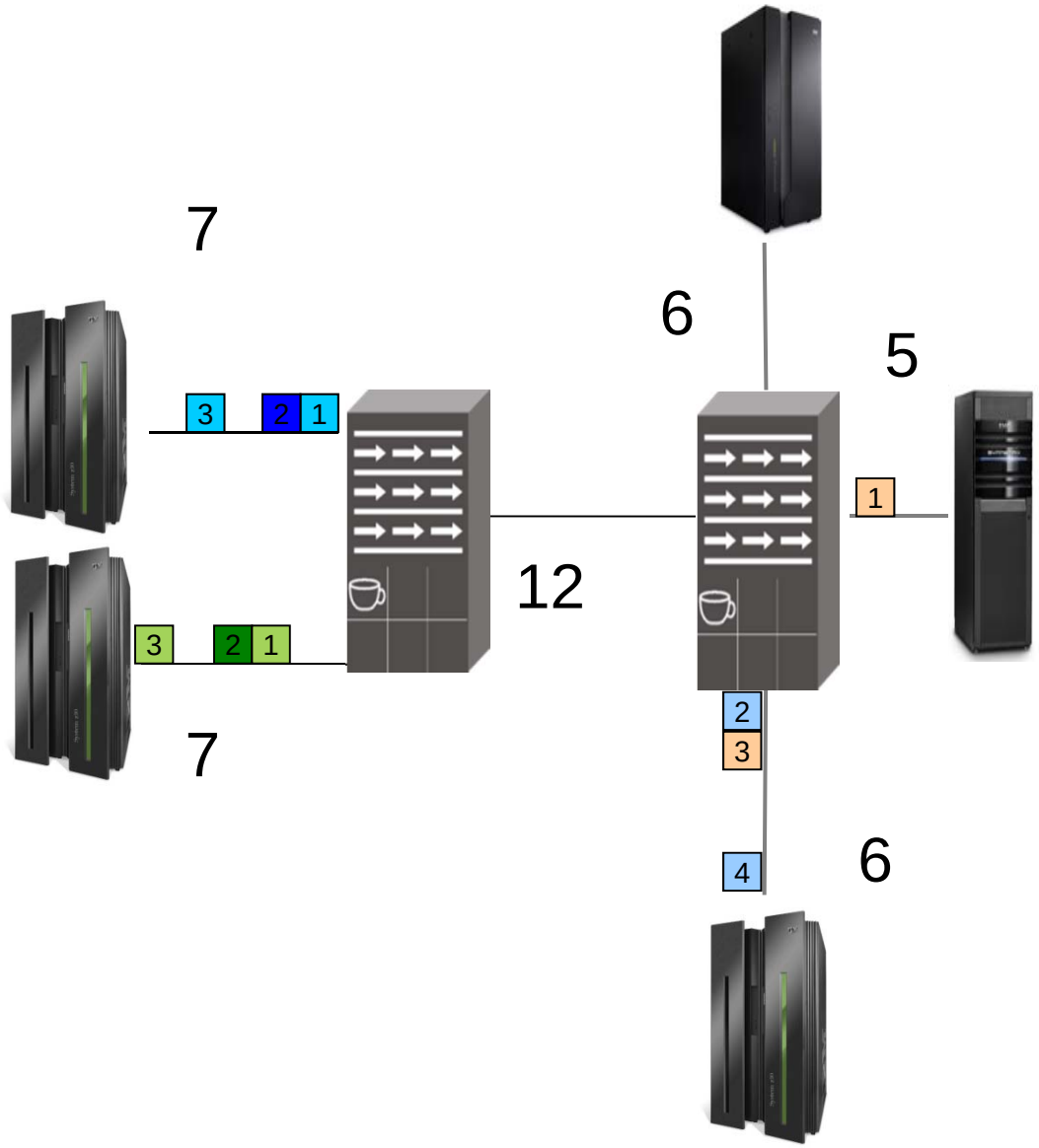

10
6
6
12
10
1
9

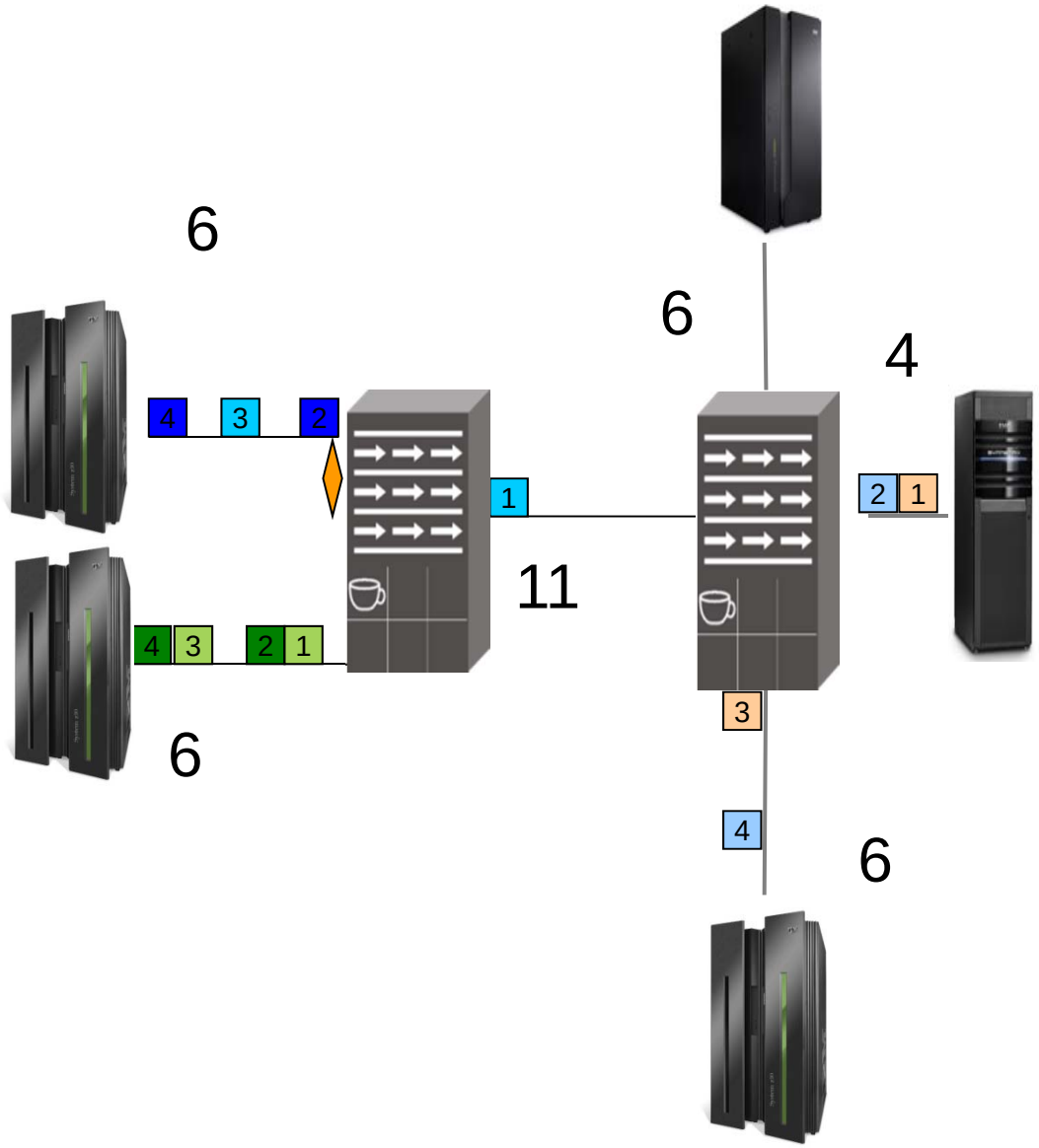

9
6
6
1
12
10
1
2
8

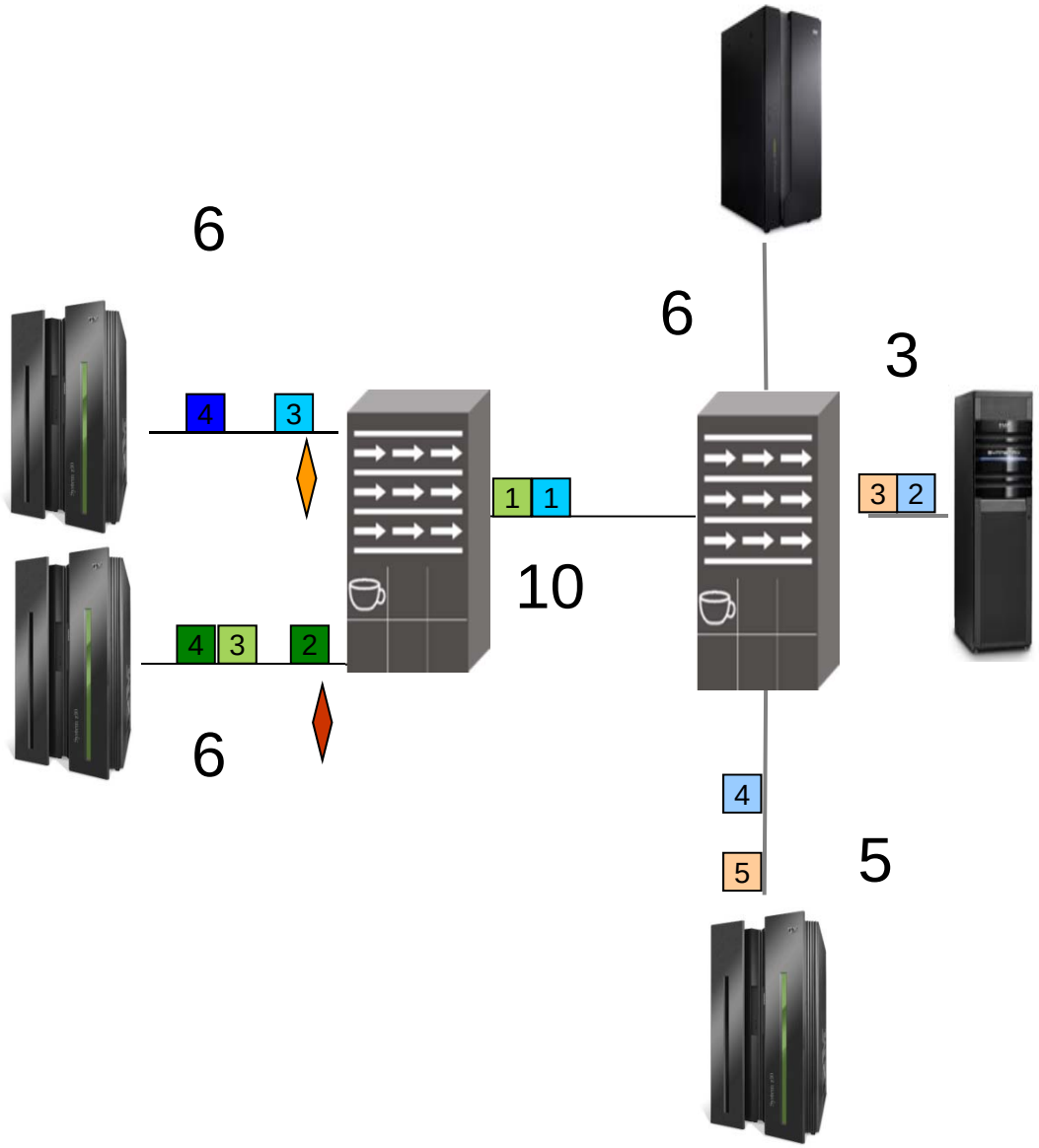

8
6
6
2
1
12
1
9
1
2
3
7

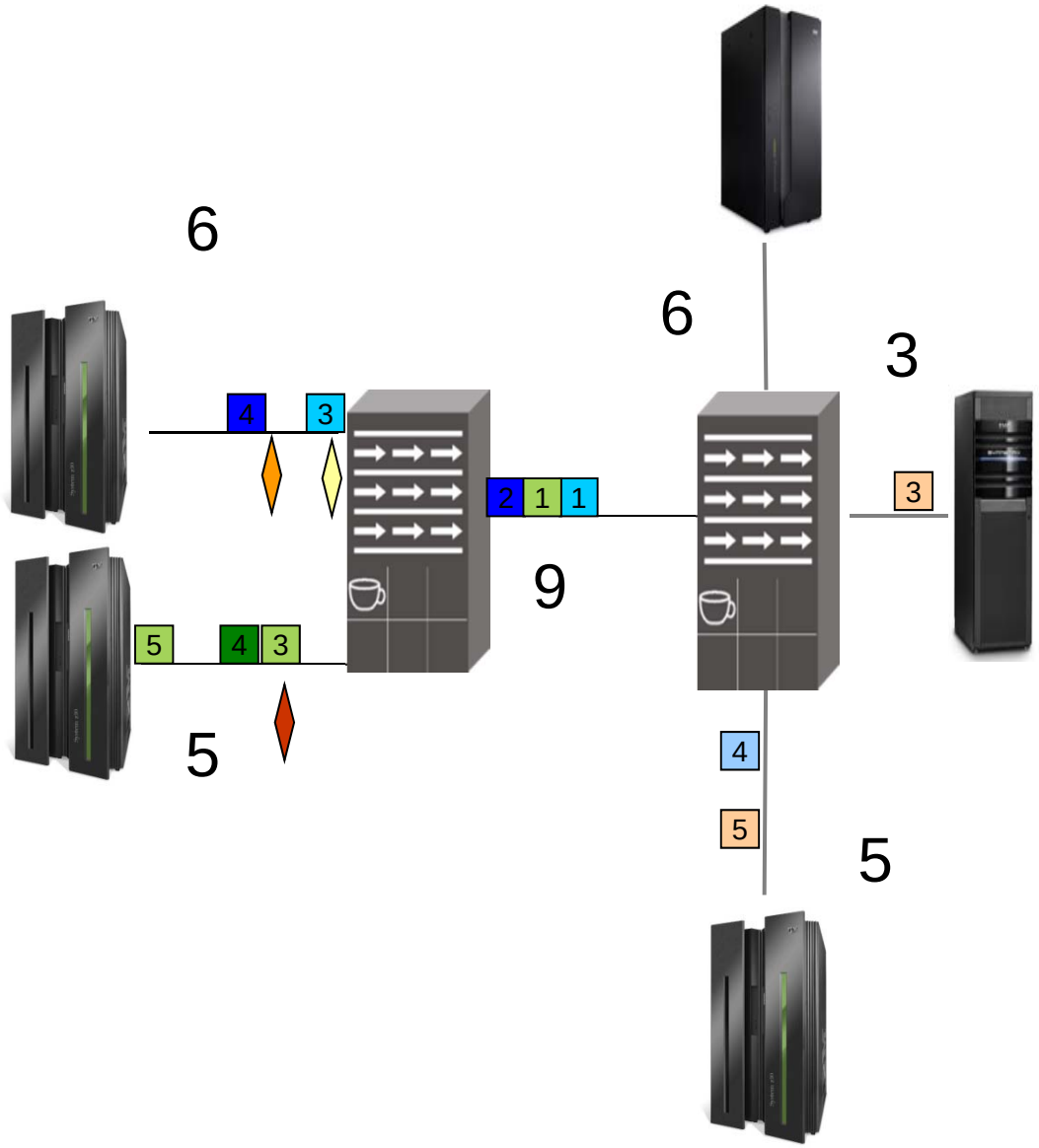

8
6
6
2
1
12
2
1
8
1
2
3
7

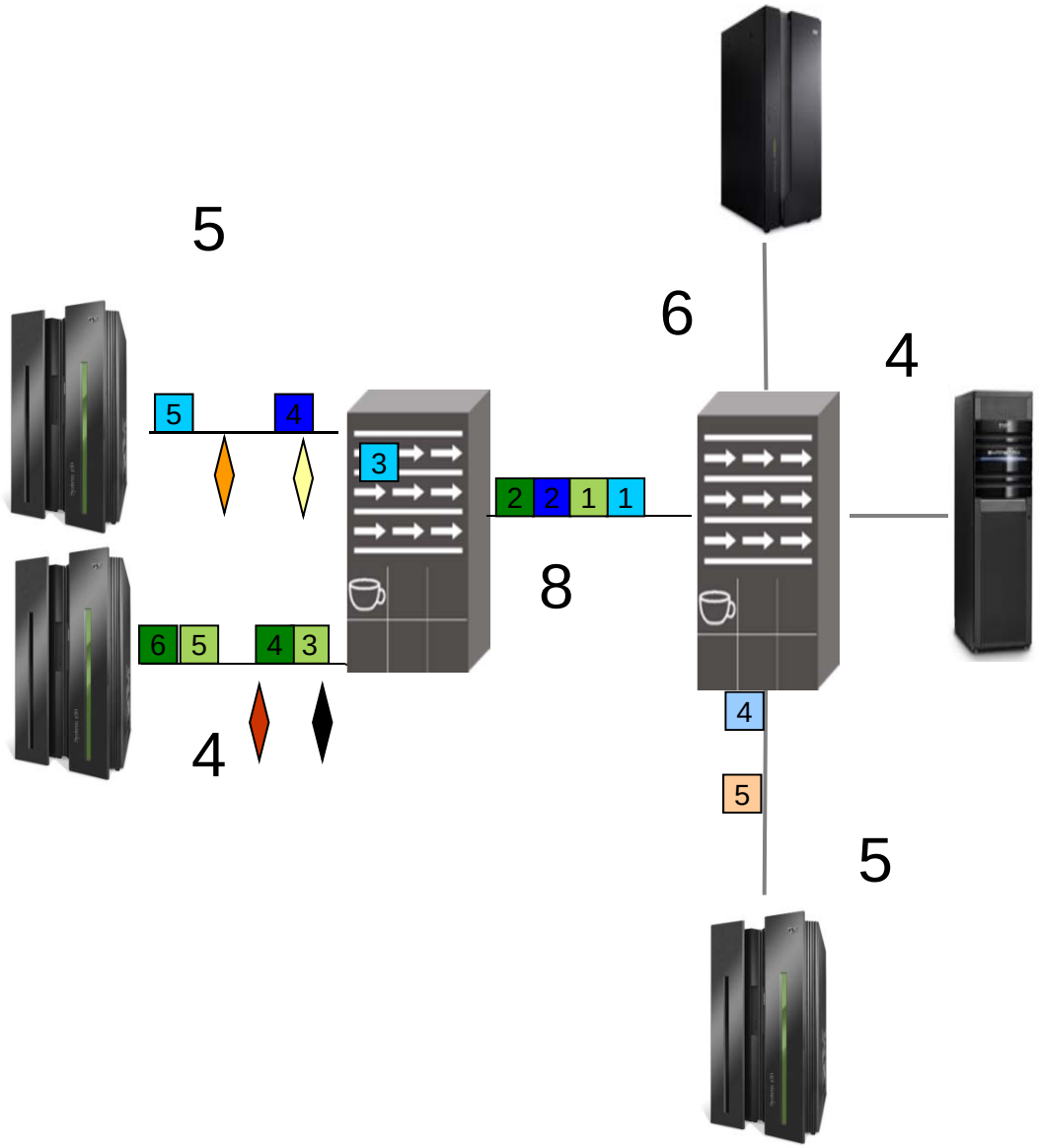

7
6
6
3
2
1
12
2
1
8
1
2
3
7

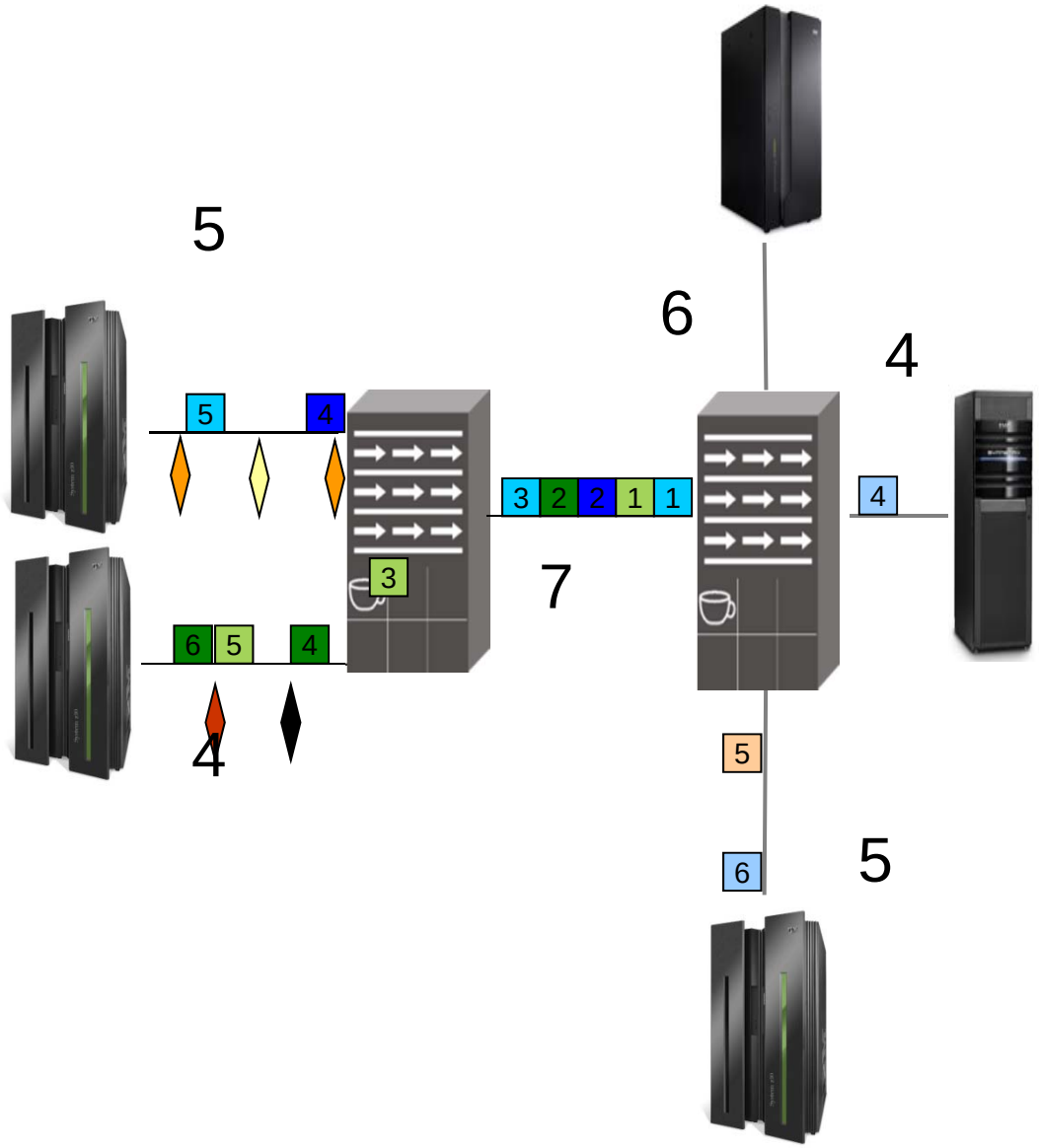

7
6
5
3
2
1
12
1
3
2
1
2
3
7
4
6

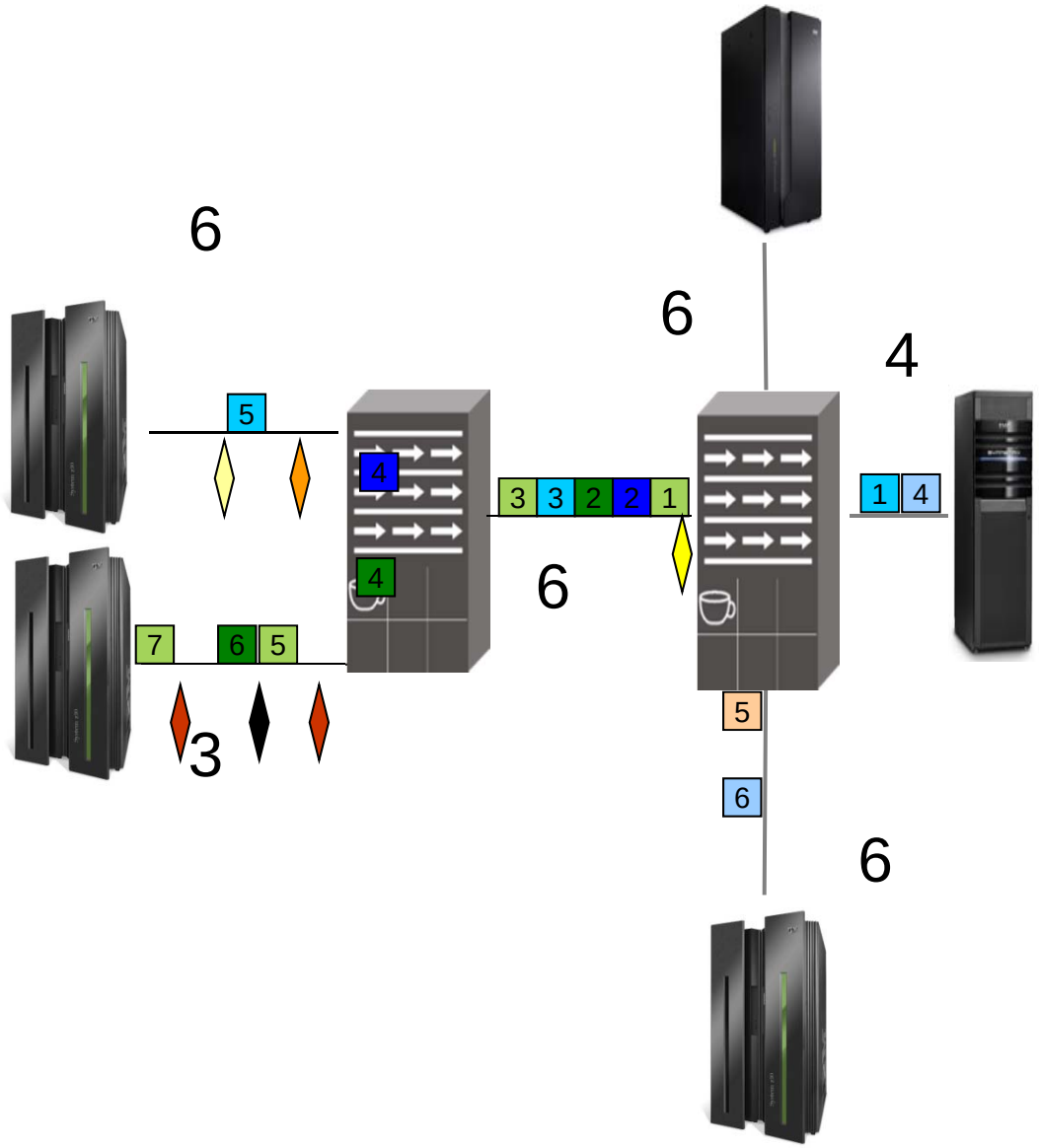

6
6
4
4 3 2
1
11
2 1
4 3 2 1
6
3
4
6

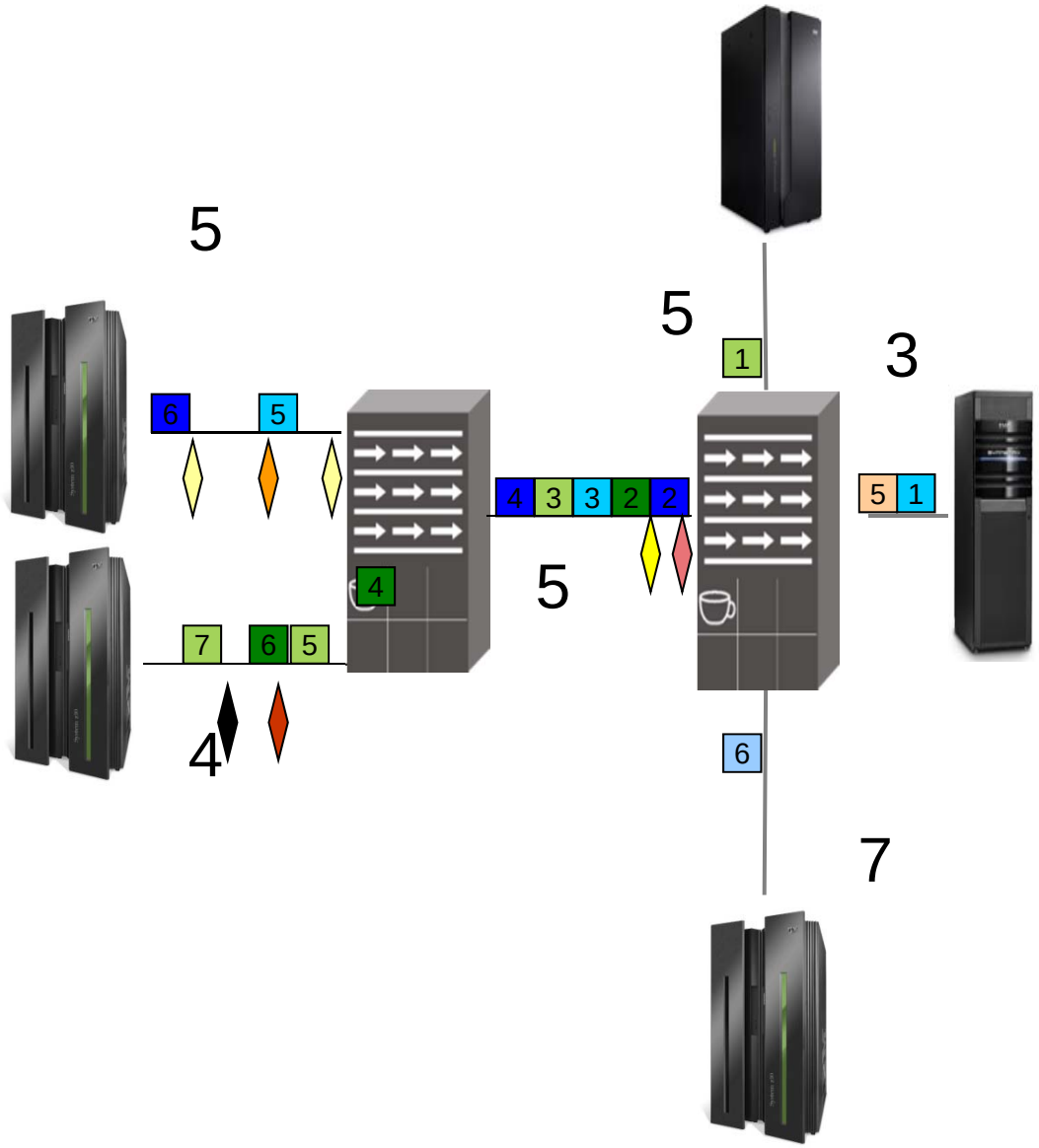

6
6
3
4
3
1
1
10
3
2
4
3
2
6
4
5
5

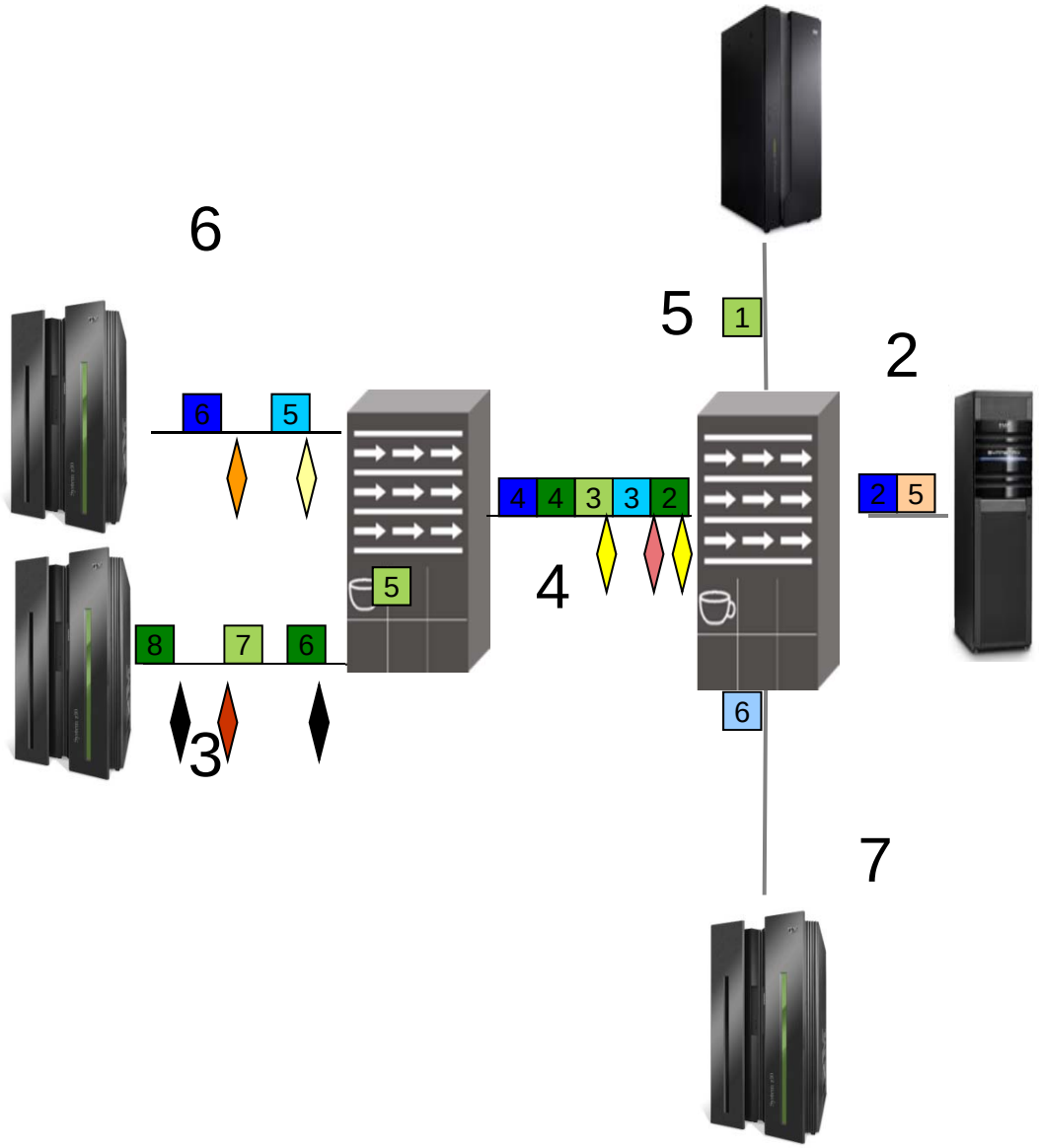

6
6
3
4
3
2
1
1
3
9
5
4
3
5
4
5
5

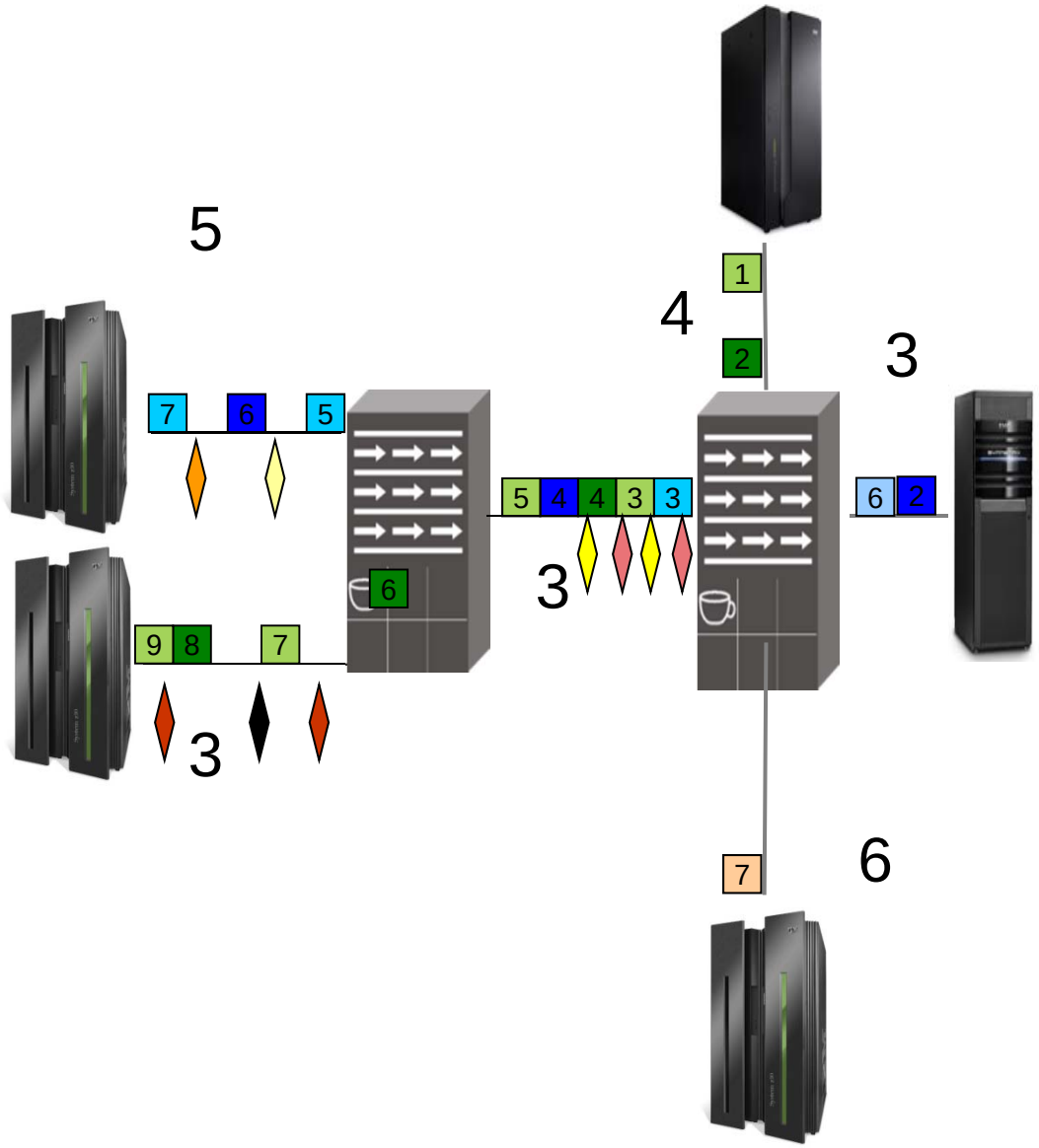

5
6
4
5
4
3
2
2
1
1
8
6
5
4
3
4
4
5
5

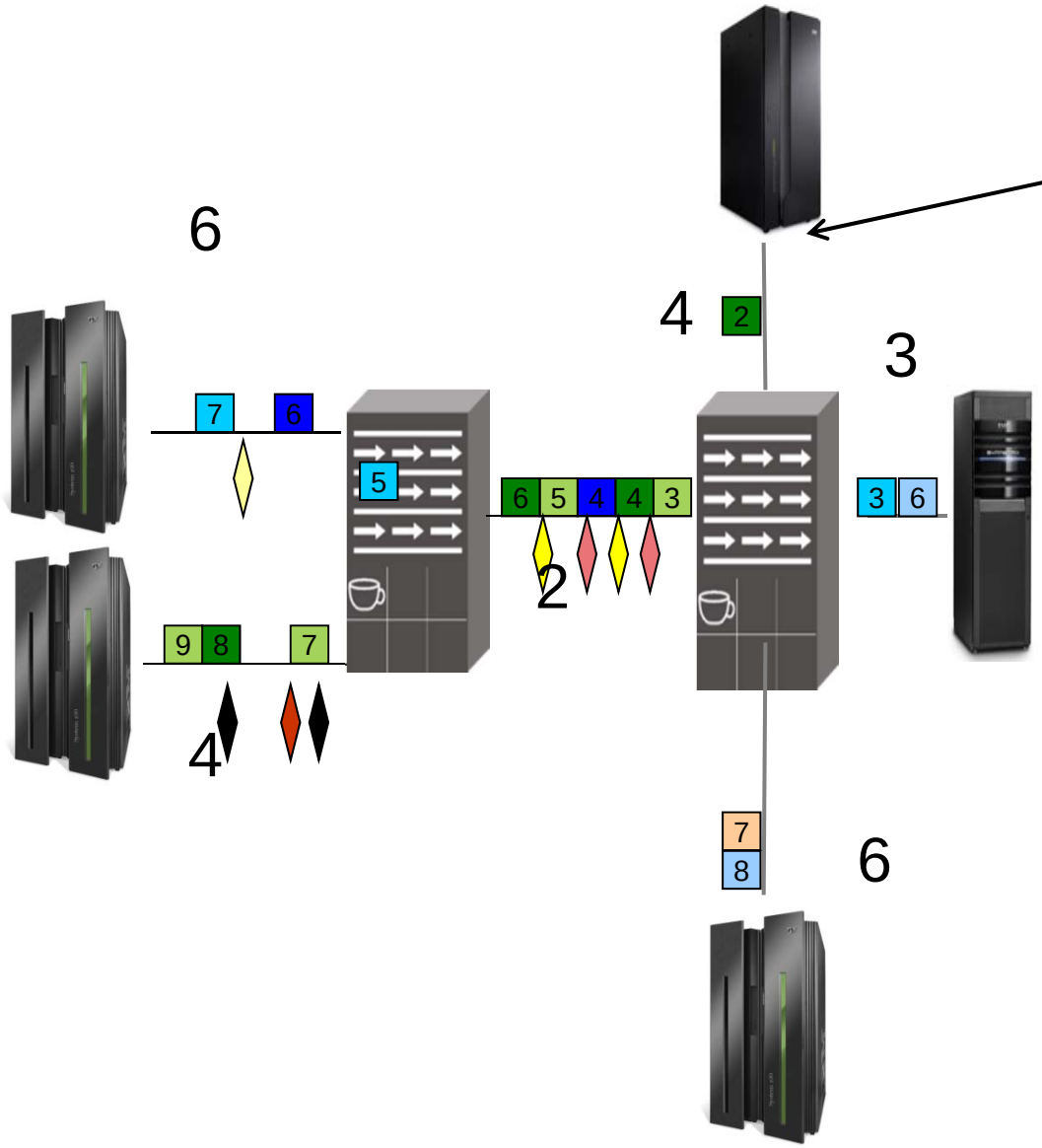

5
6
4
5
4
3 2 2 1 1
4
3
7
6 5 4
4
5
6
5

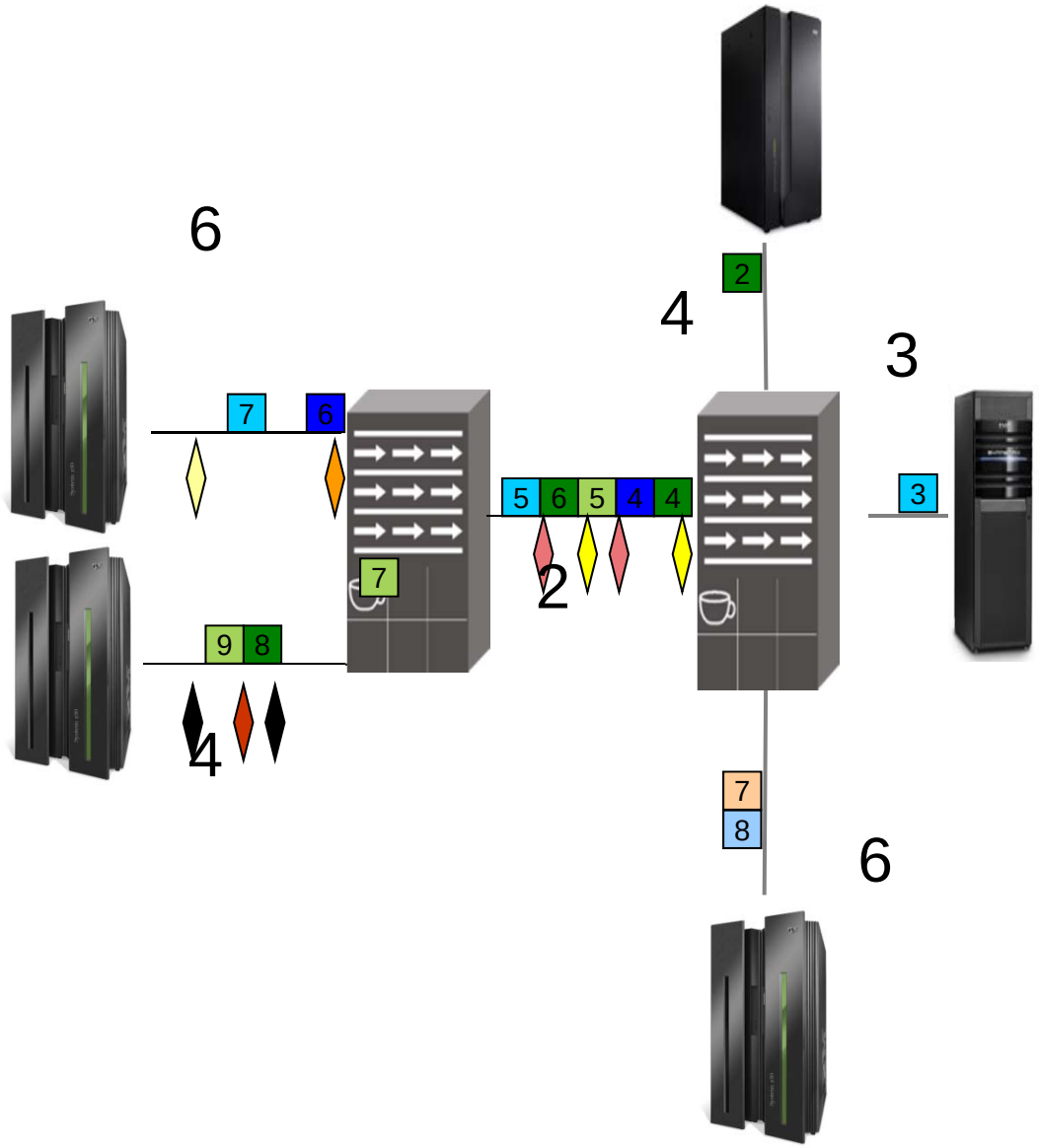

6
6
4
5
4
3 3 2 2 1
1 4
4
6
7 6 5
3
5
6
6

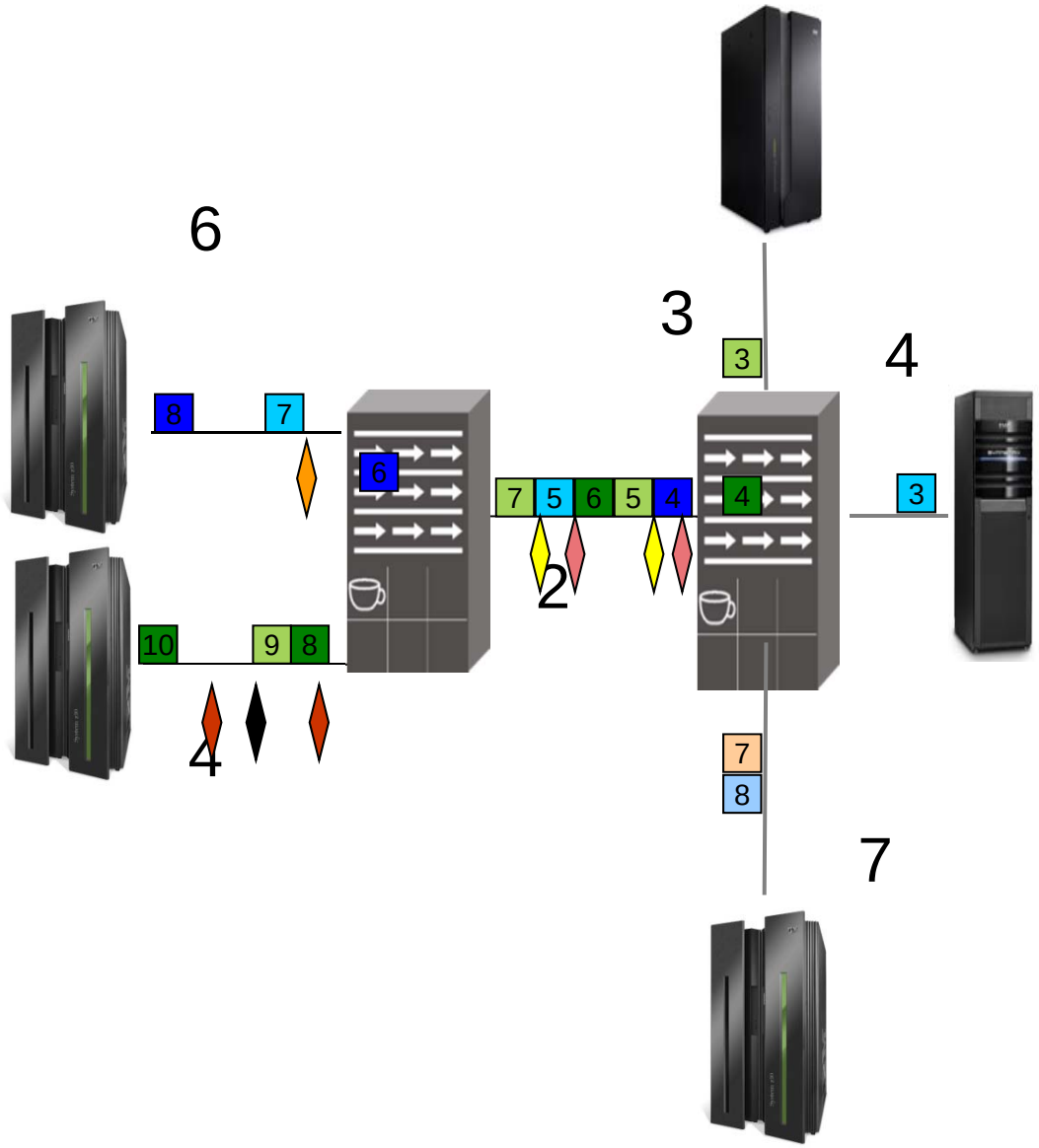

5
5
3
1
6
5
4 3 3 2 2
5 1
4
5
7 6 5
4
6
7

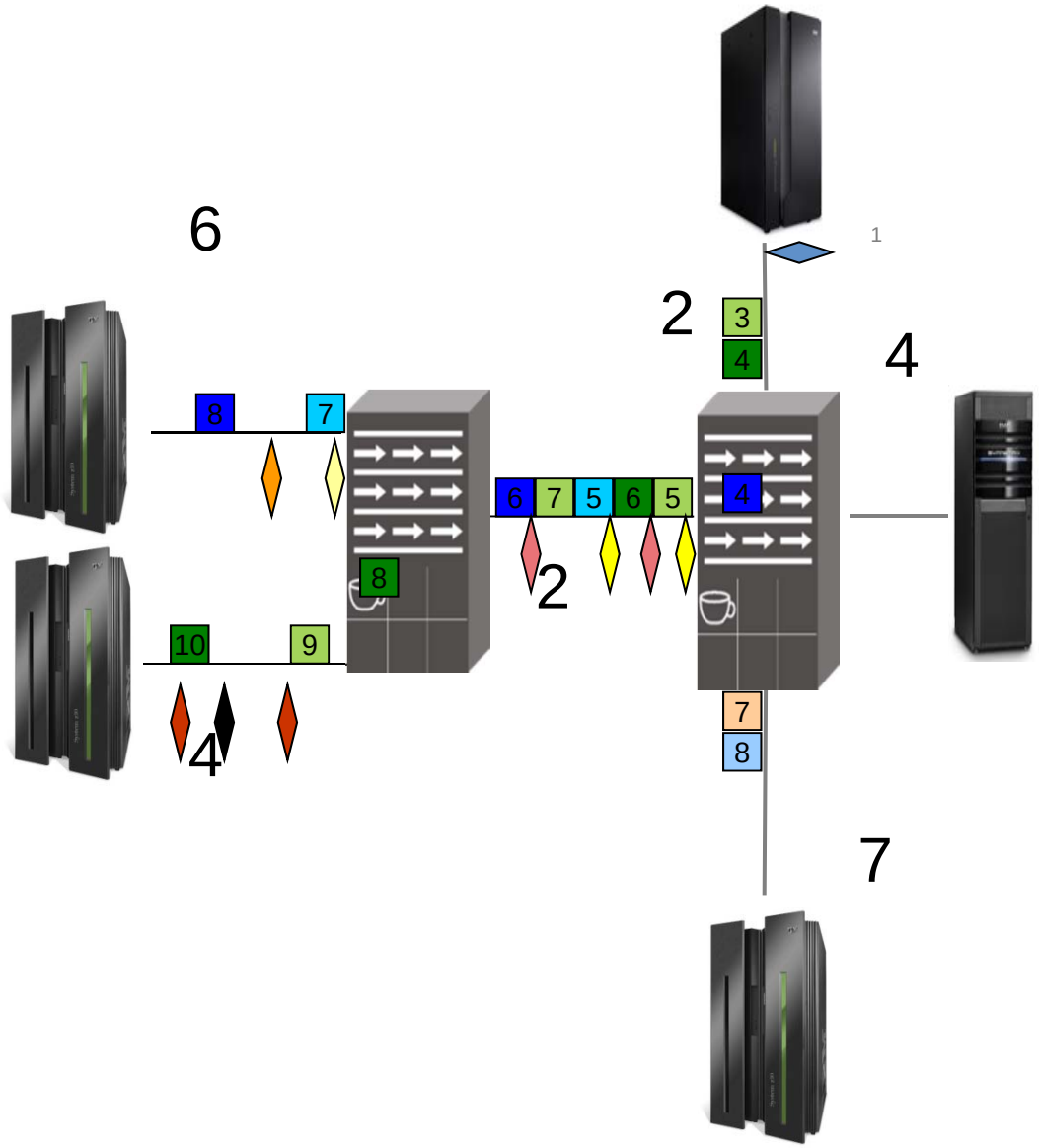

6
5
1
2
6
5
4 4 3 3 2
2 5
5
4
8
7
6
3
6
7

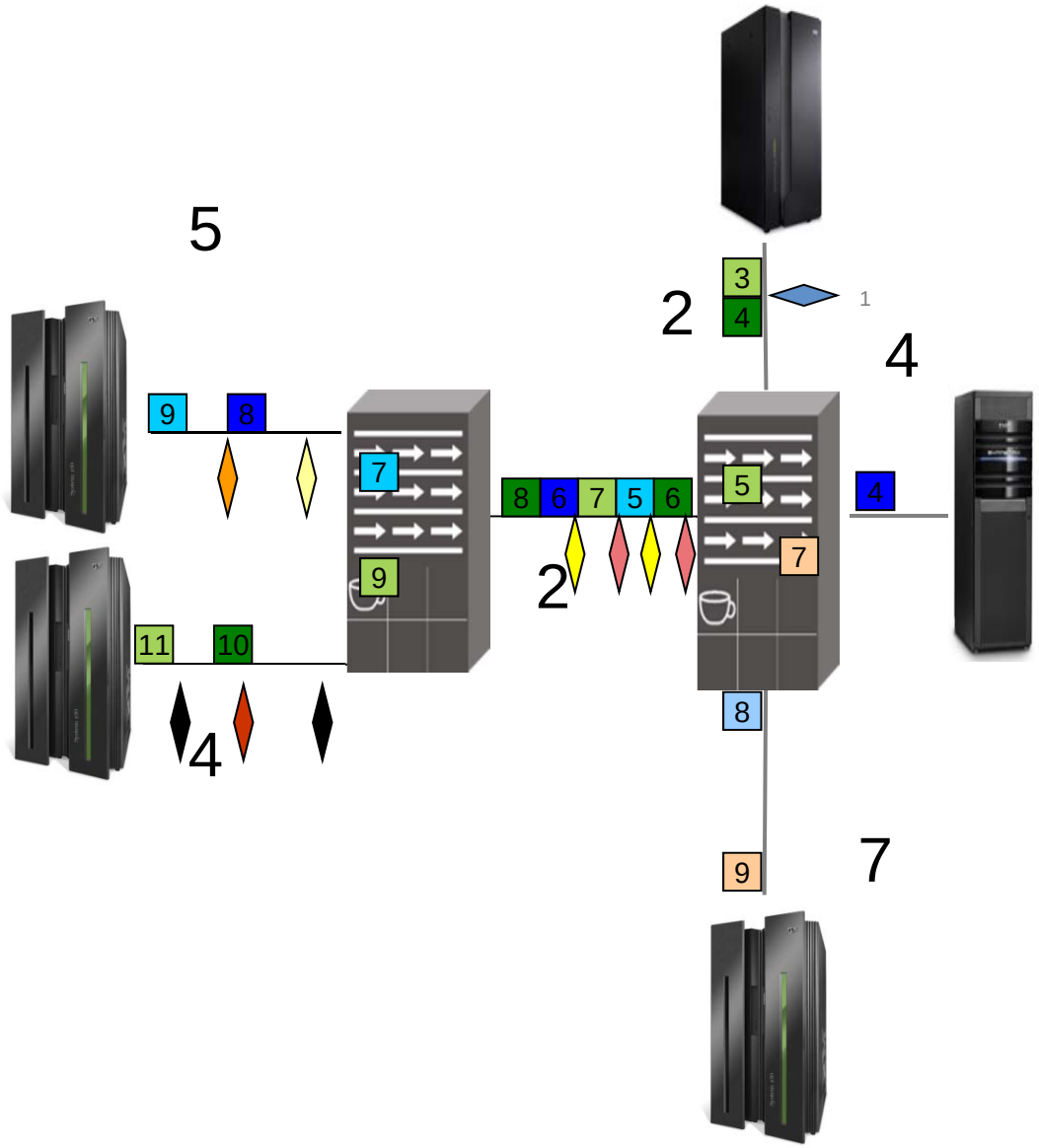

5
4
3
6
7
6
5
5
4
4
3
3
6
2
6
9
8
7
3
1
2
7

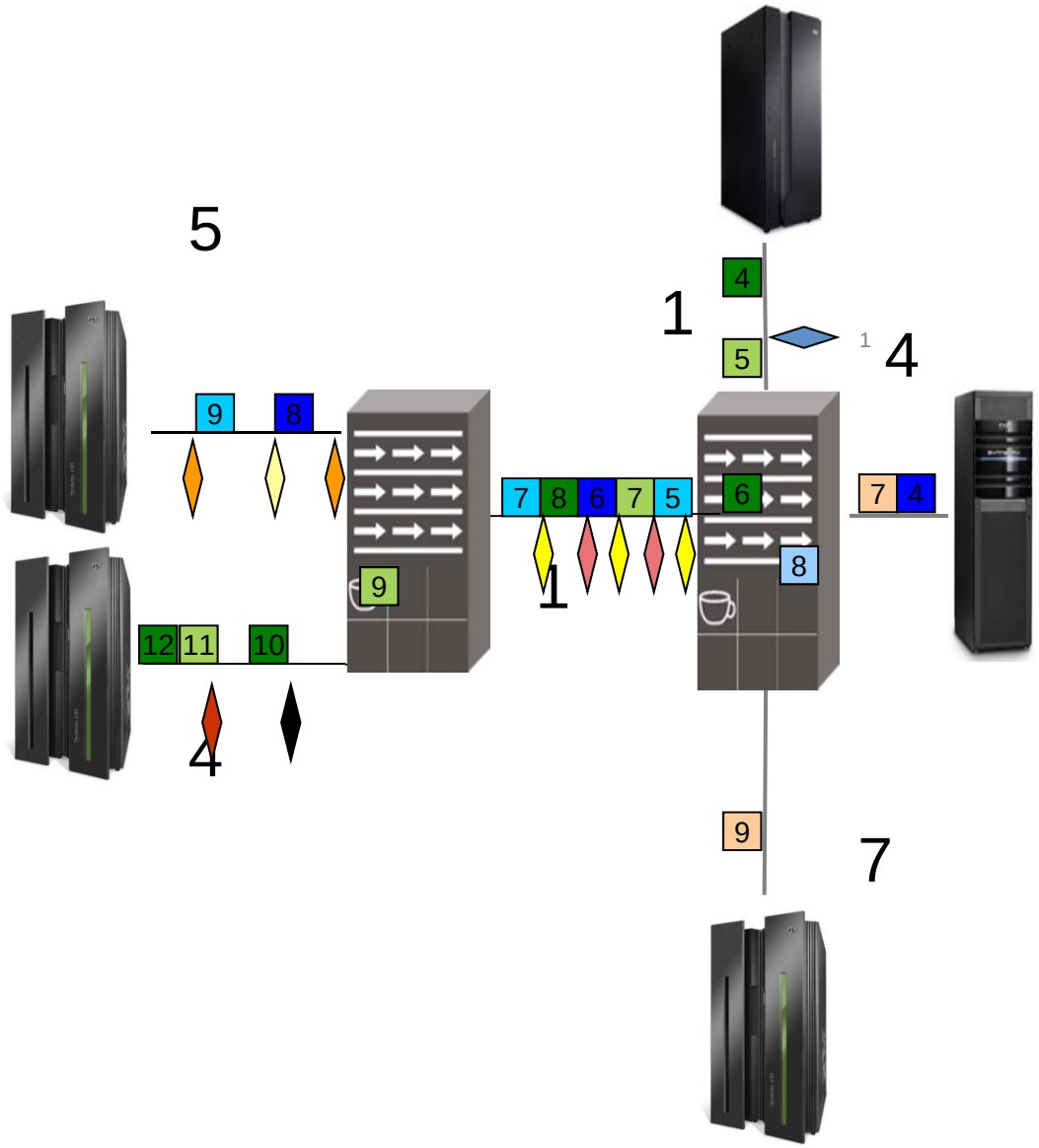

Notice NO R_RDY
6
4
2
3
7
6
5
6
5
4
4
3
3
6
2
9
8
7
4
7
8
6

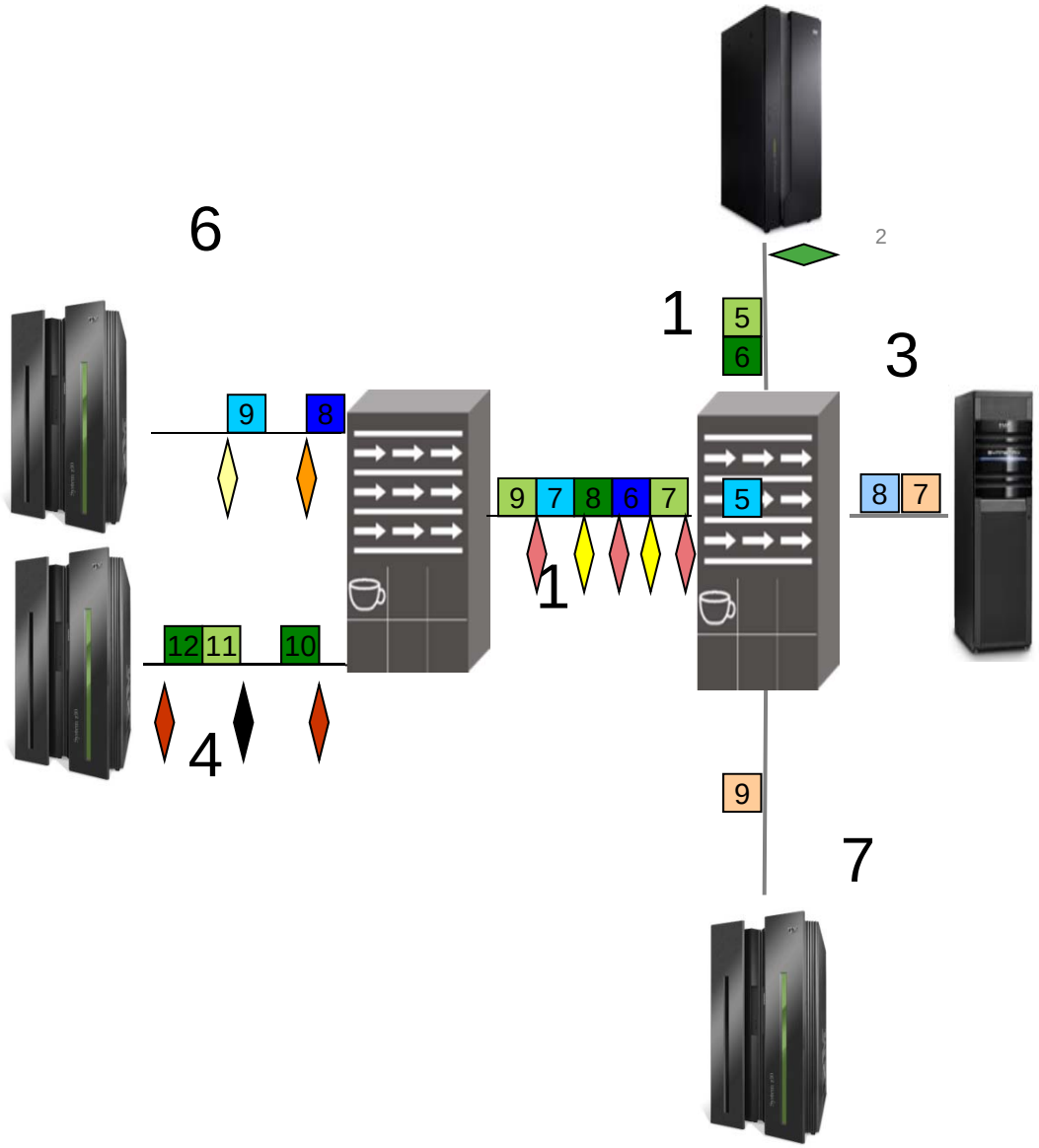

6
4
3
2
7
6
5
6
5
4
4
3
7
2
9
8
4
7
8
6

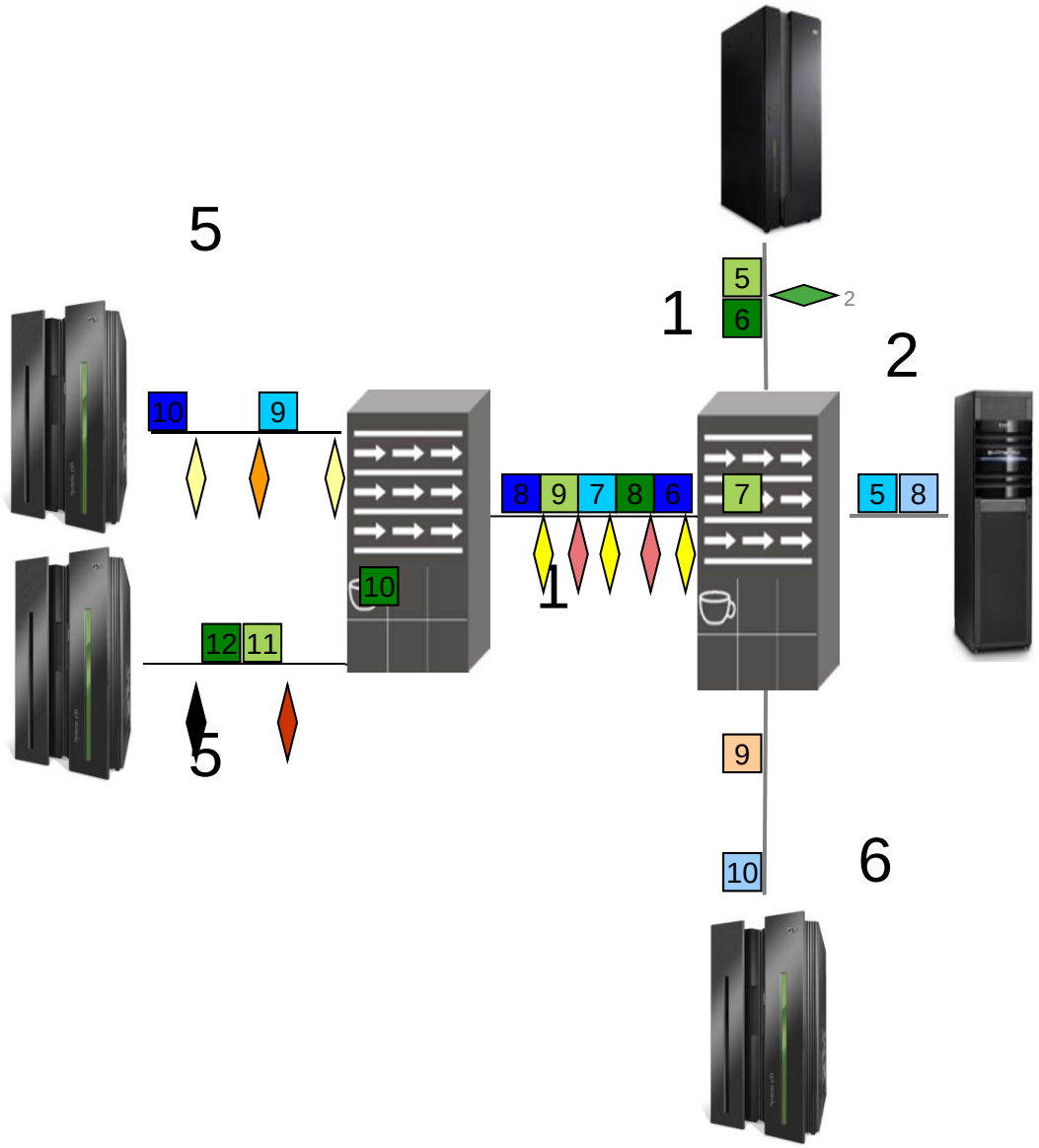

6
3
4
8
7
6
7
5
6
5
4
4
3
2
10
9
8
4
7
8
7

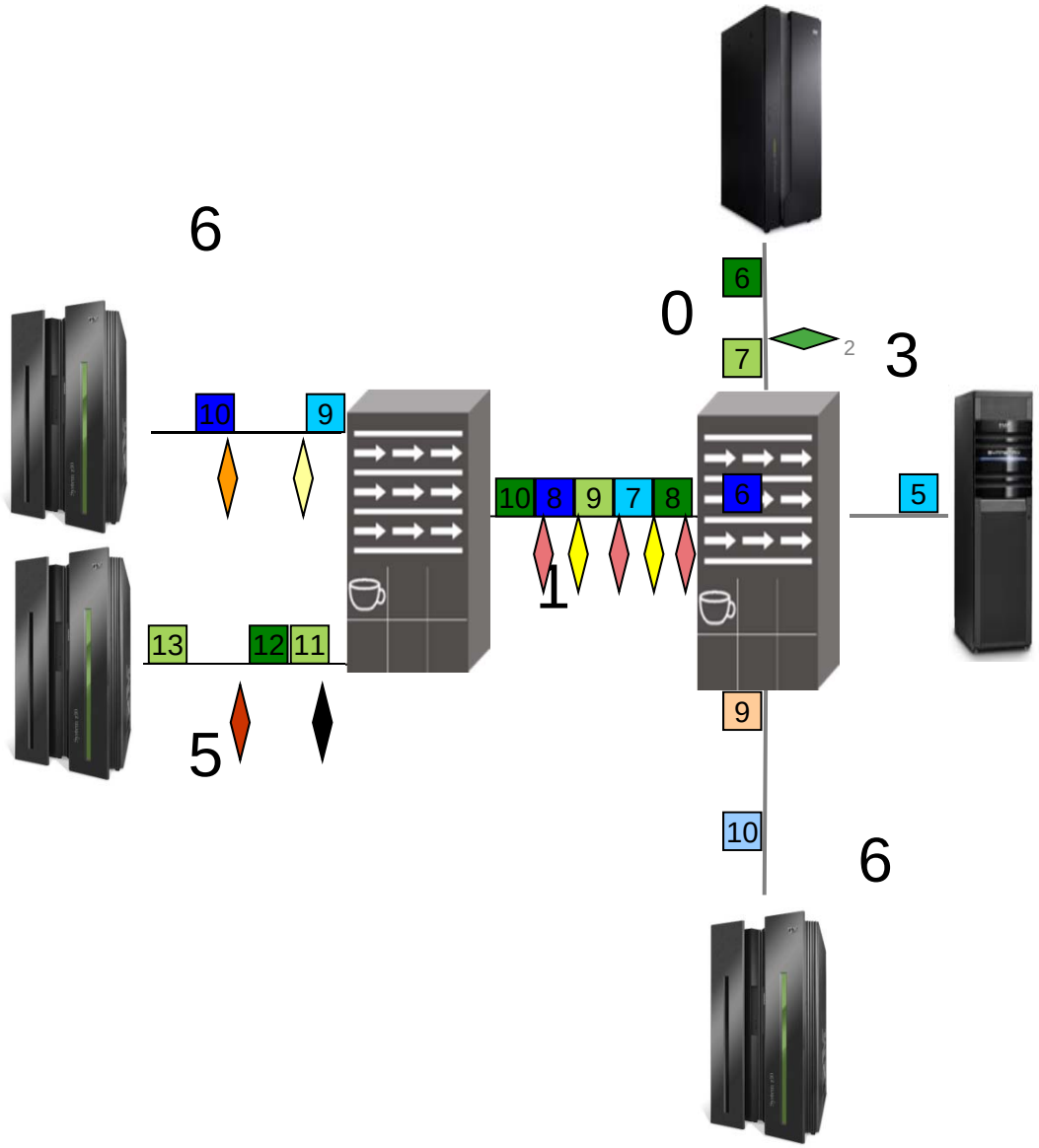

6
1
2
3
4
4
8
7
6
7
5
6
5
4
8
2
10
9
4
7
8
7

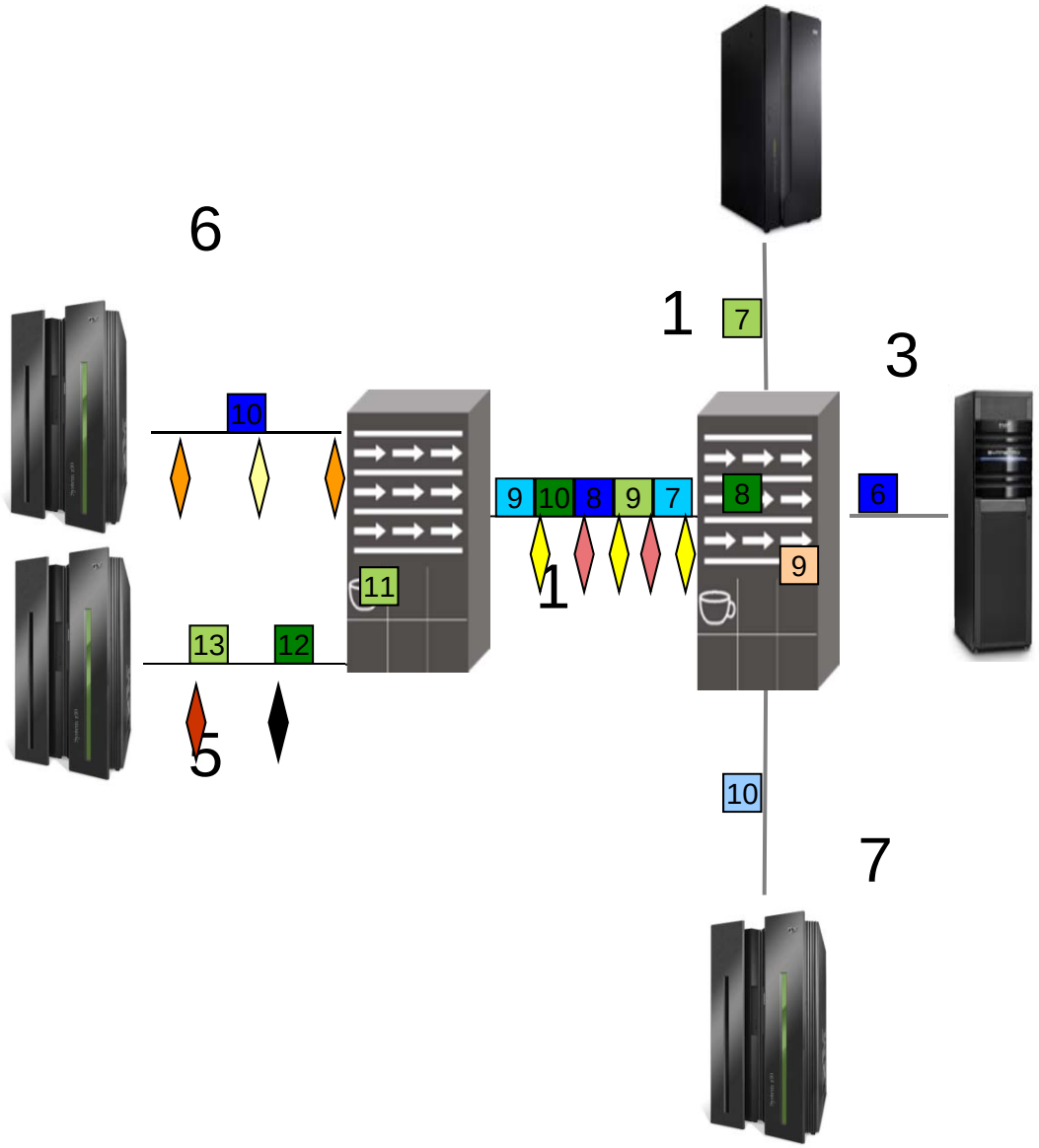

5
2
3
4
1
4
9
8
7
8
6
7
5
6
5
4
7
9
2
11
10
4
8
9
7

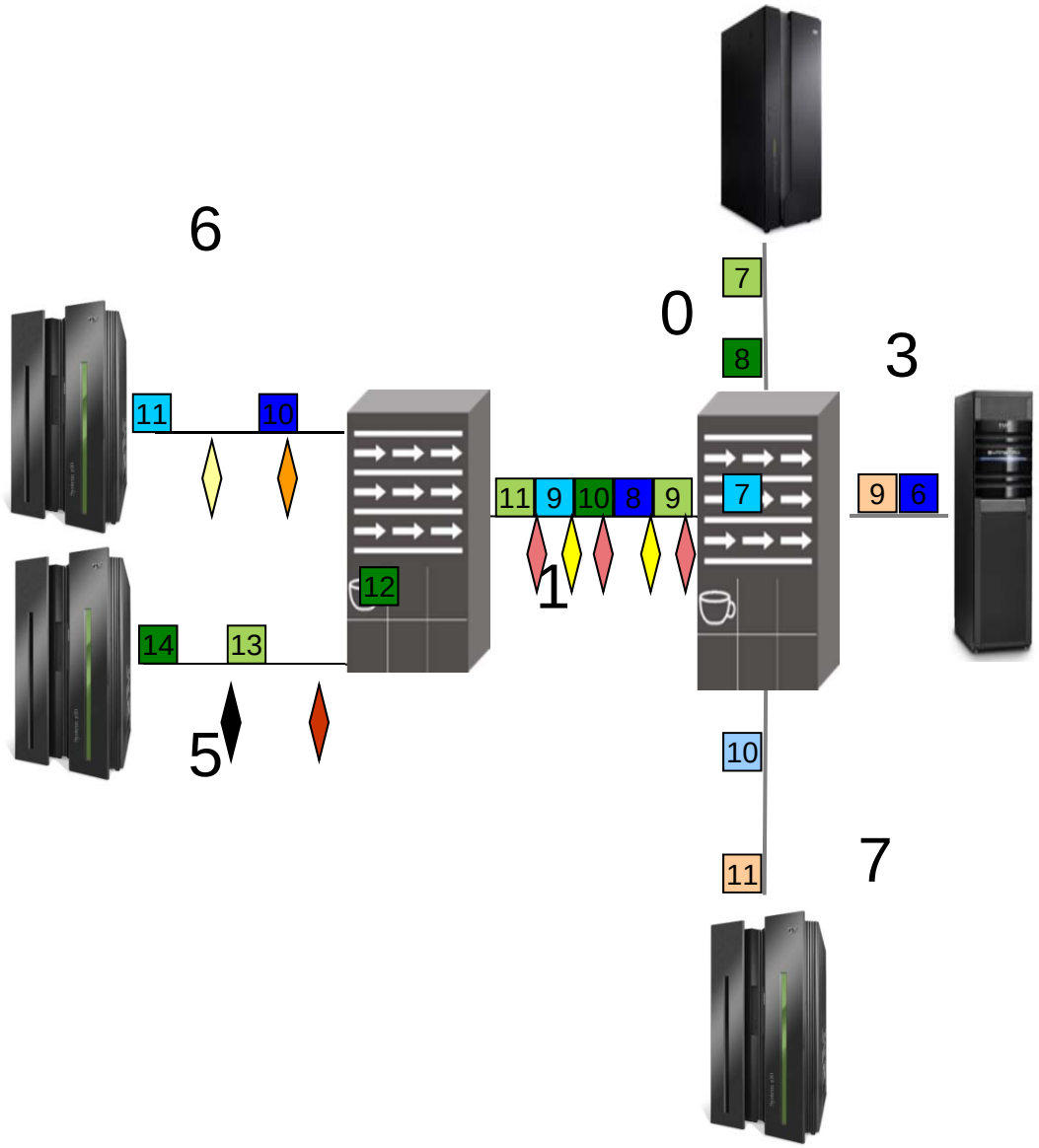

5
1
4
1
4
9
8
7
8
6
7
5
6
7
4
8
9
1
12
11
10
4
9
7

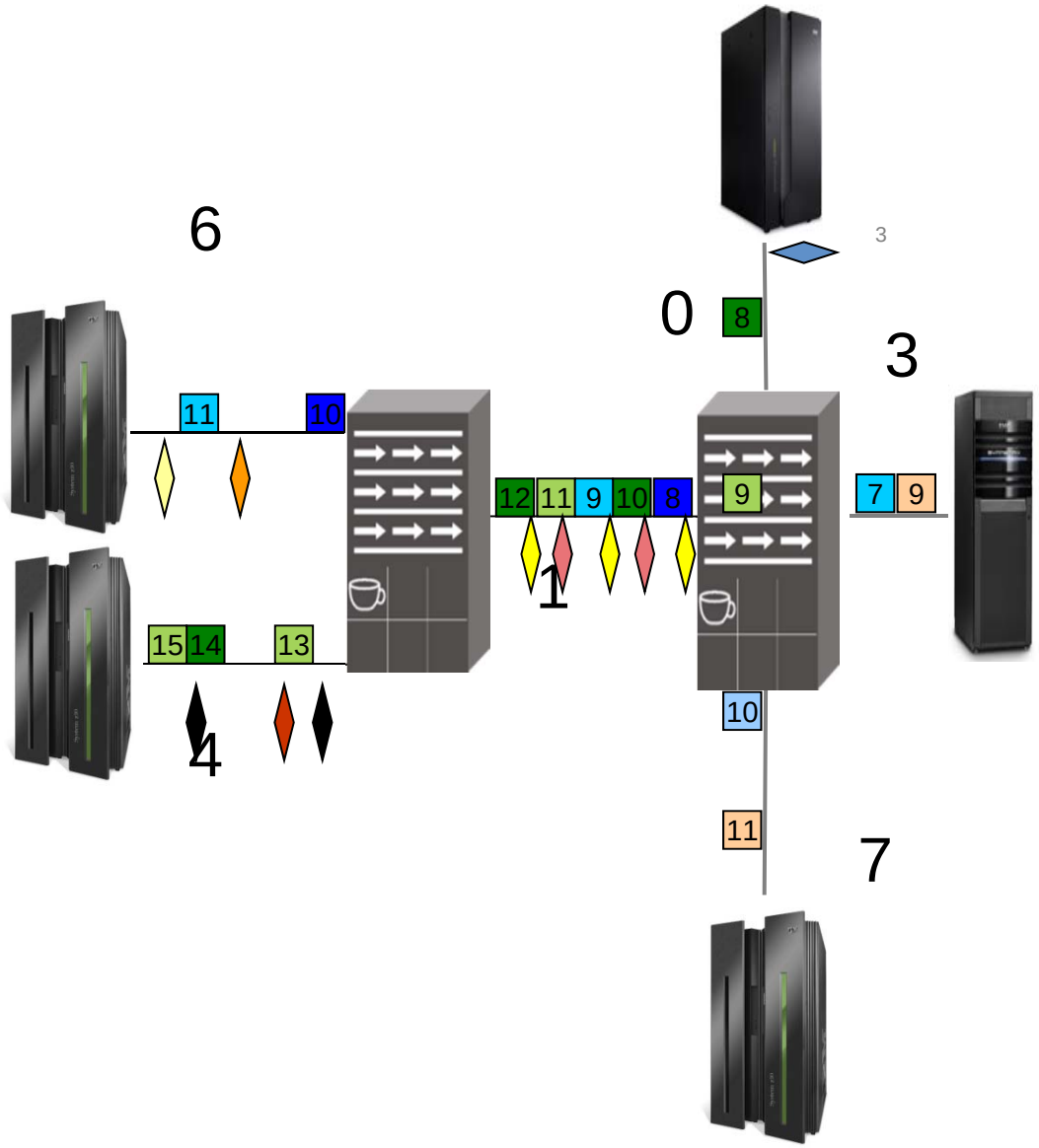

6
2
1
5
6
3
9
8
9 7 8 6 7
5
8 7
1
12 11
10
4
9
7

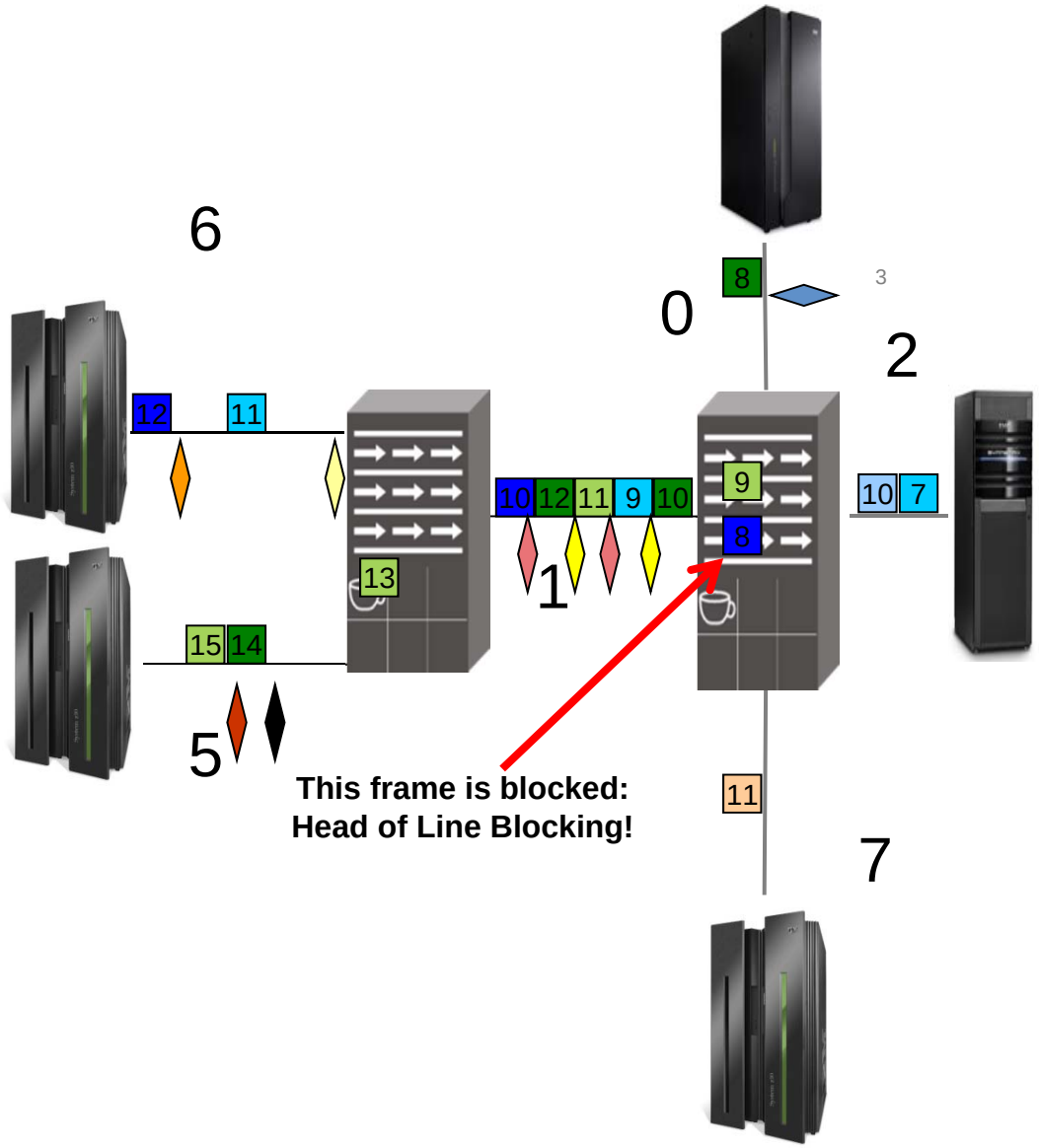

5
1
2
6
10
9
8 9 7 8 6
7
5 8
10
12 11
1
5
9
10

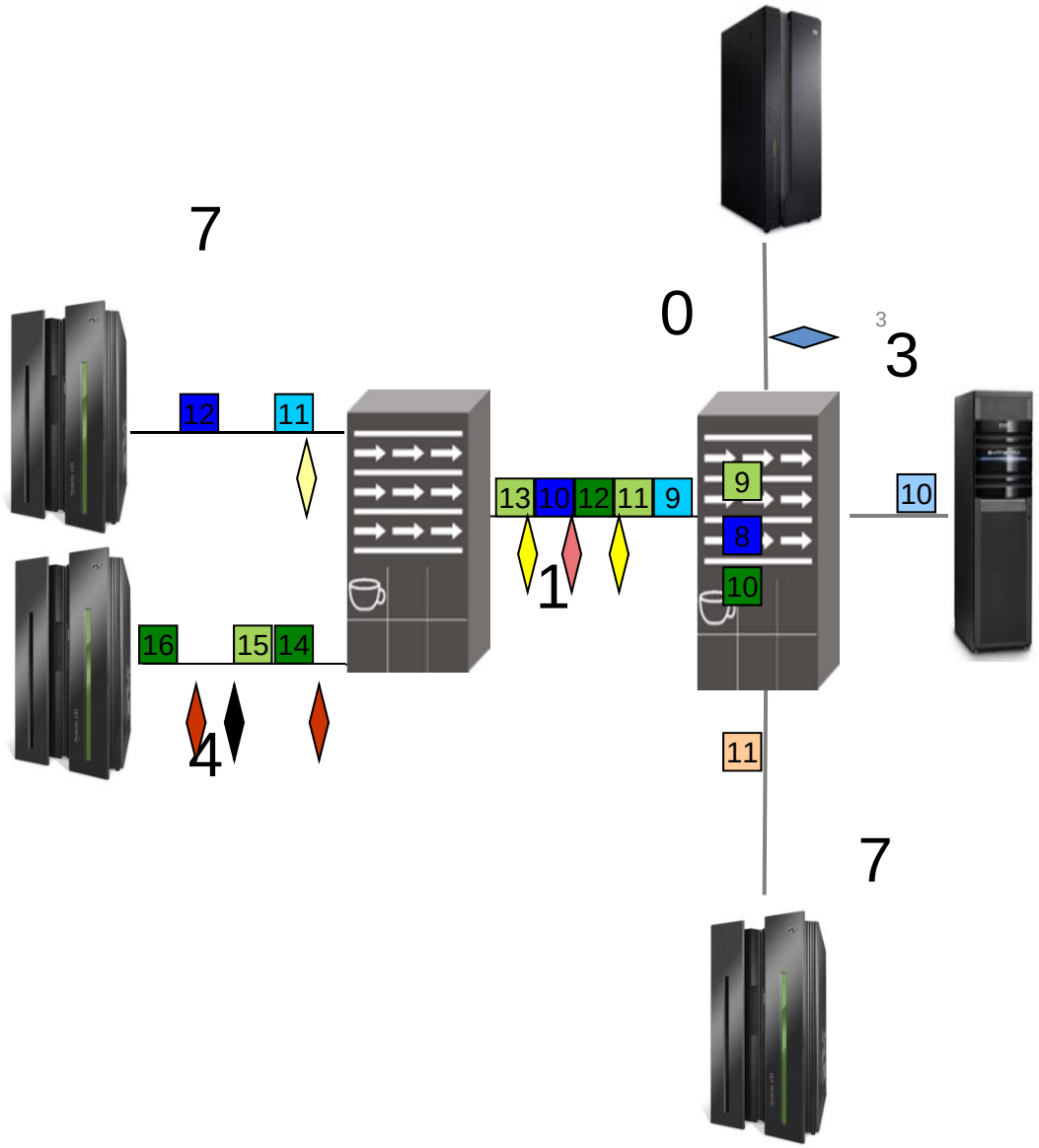

6
10
9
13
12
11
5
10
8
9
7
8
1
6
0
6
7
2
3
5
9
10
6

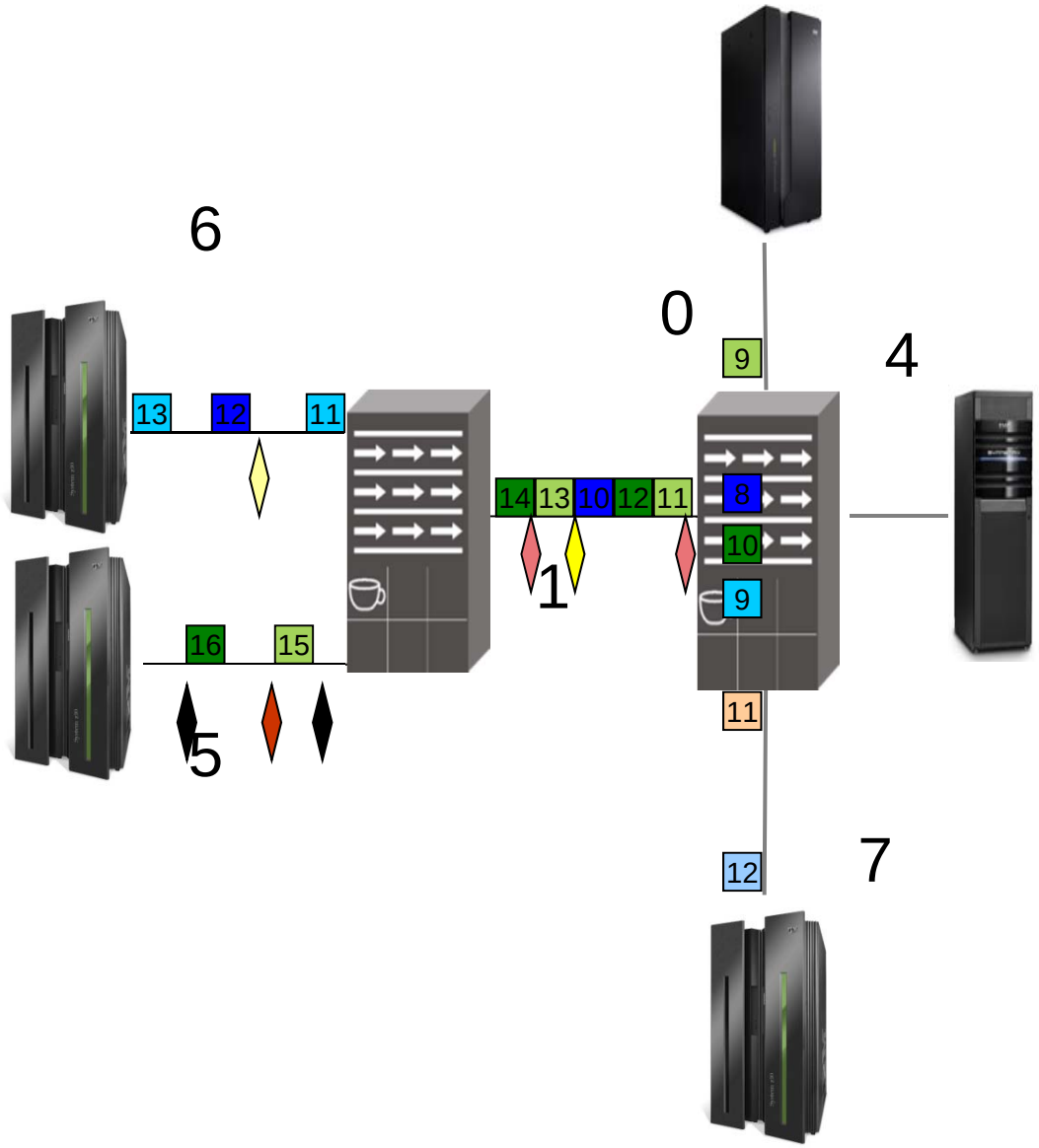

6
1
7
3
10
9
10
8
9
7
8
6
11
1
9
13
12
5
10
7

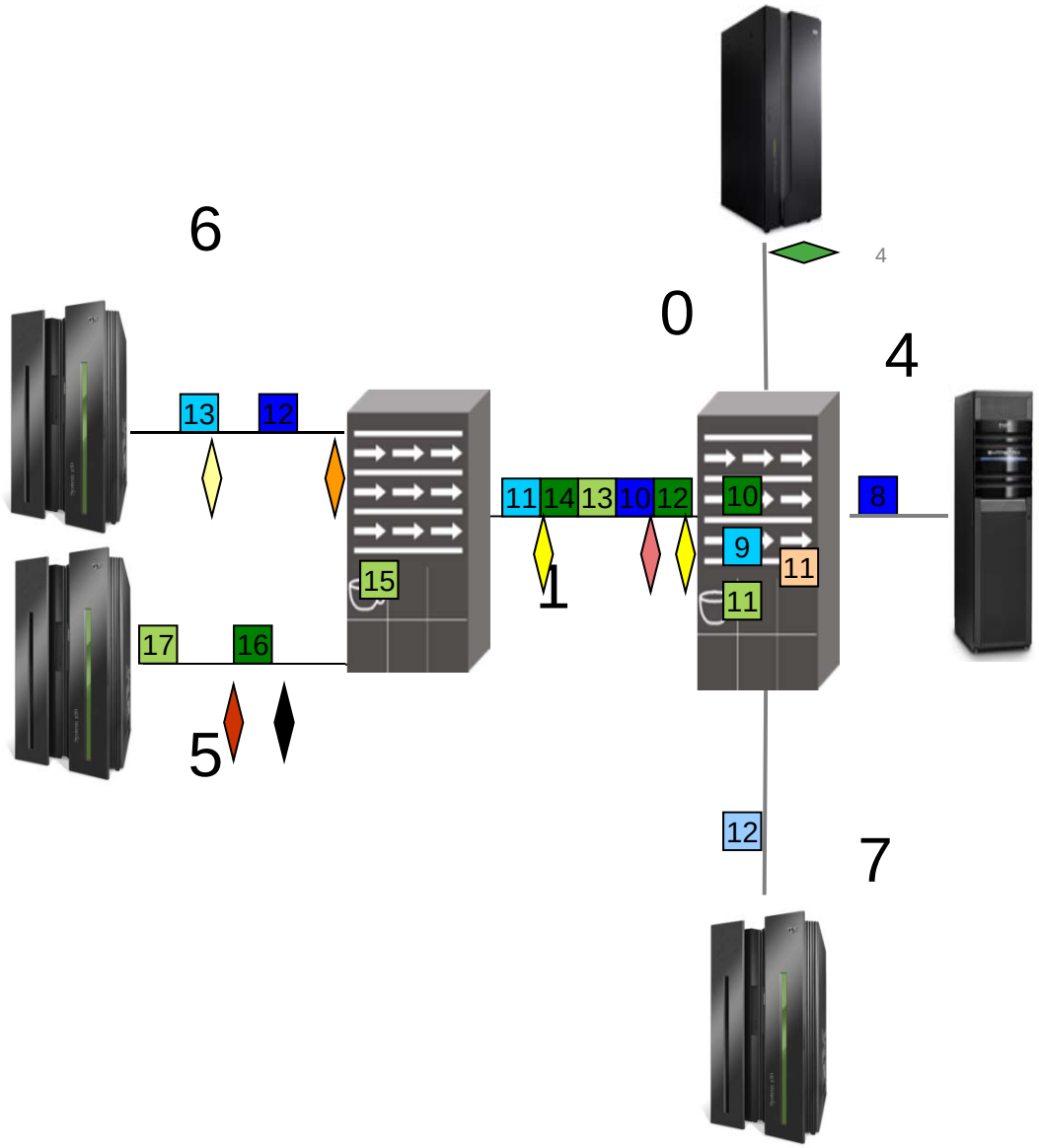

6
0
3
1
5
7
11
10
12
14
13
11
9
10
8
9
7
9
6
7
8
10
11

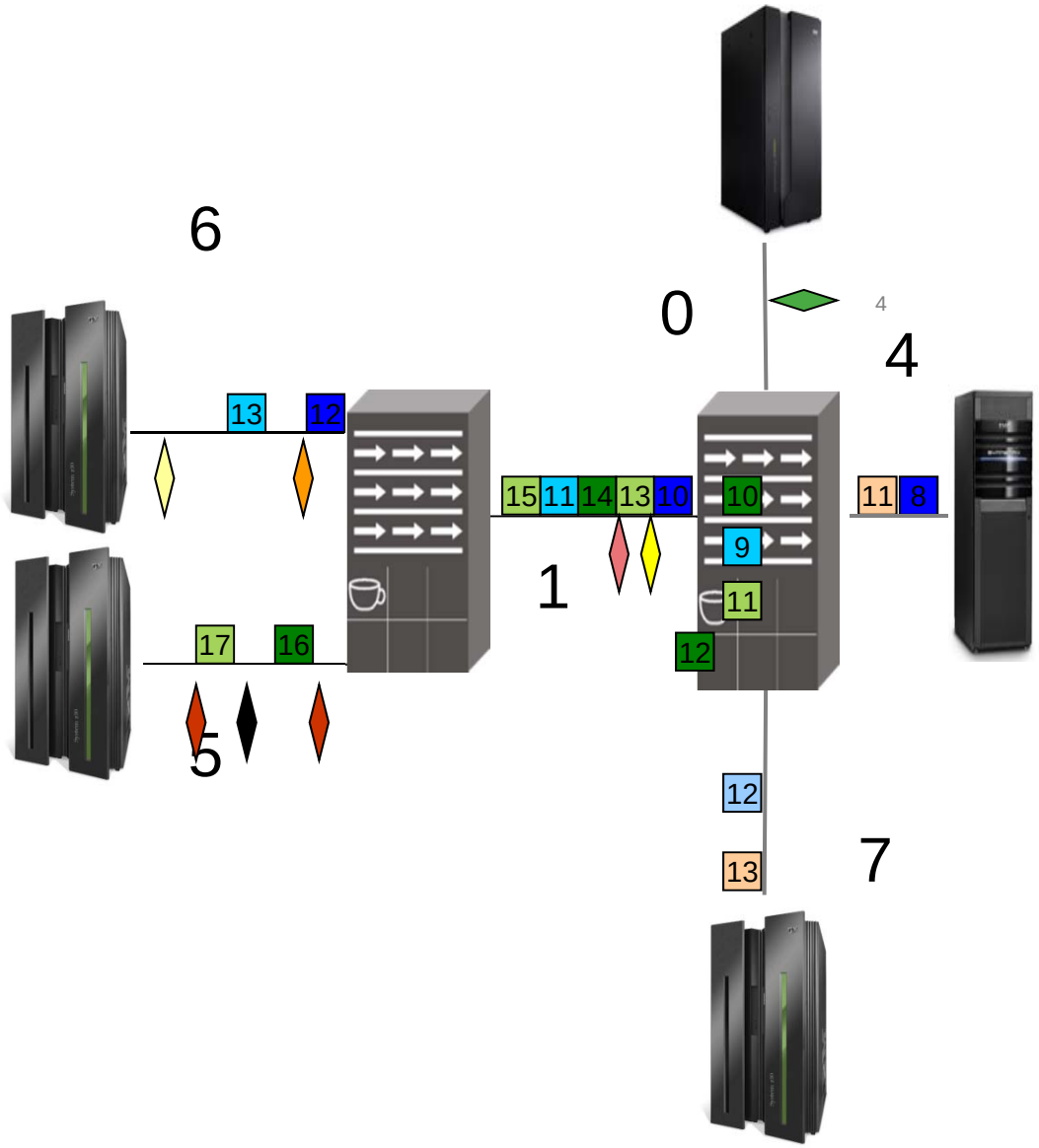

6
3
0
8
3
11
10
12
11
9
10
8
9
7
9
1
15
14
13
4
10
11
7

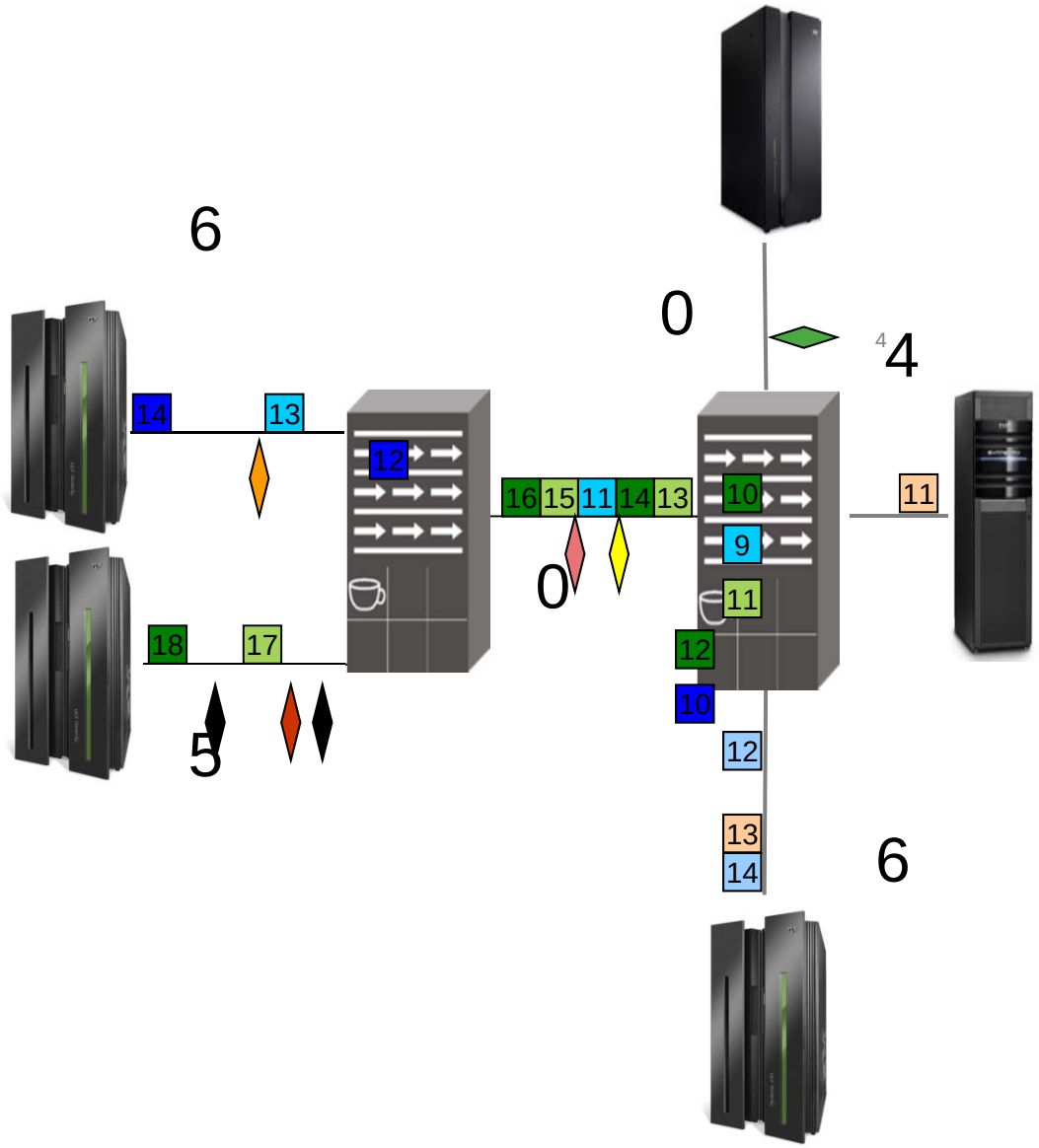

6
0
3
2
1
5
7
12
11
13
15
14
10
12
11
9
10
8
9
10
7
11
This frame is blocked:
Head of Line Blocking!

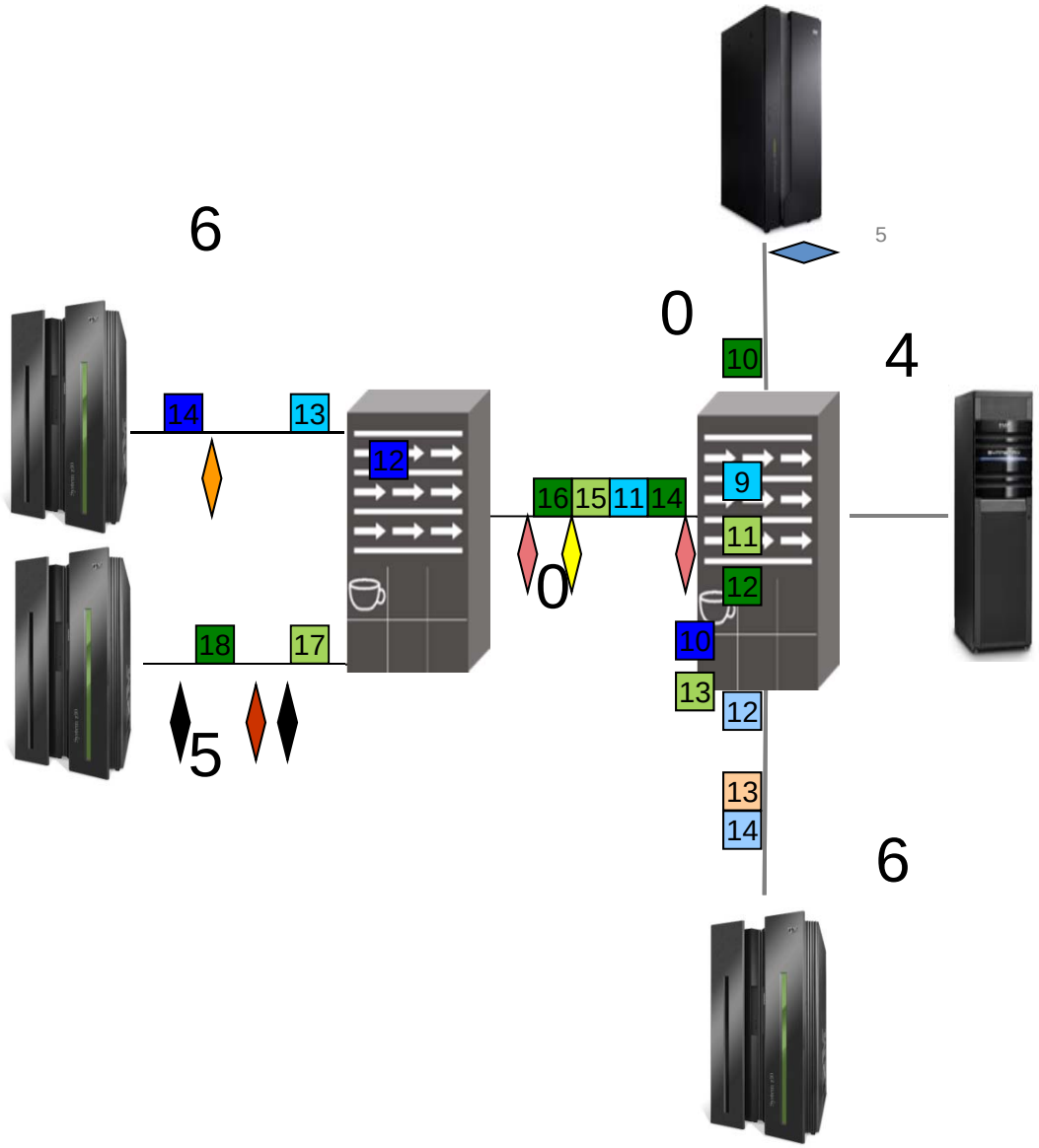

7
0
3
3
12
11
13
10
12
11
9
9
8
10
10
16
15
14
1
4
11
7

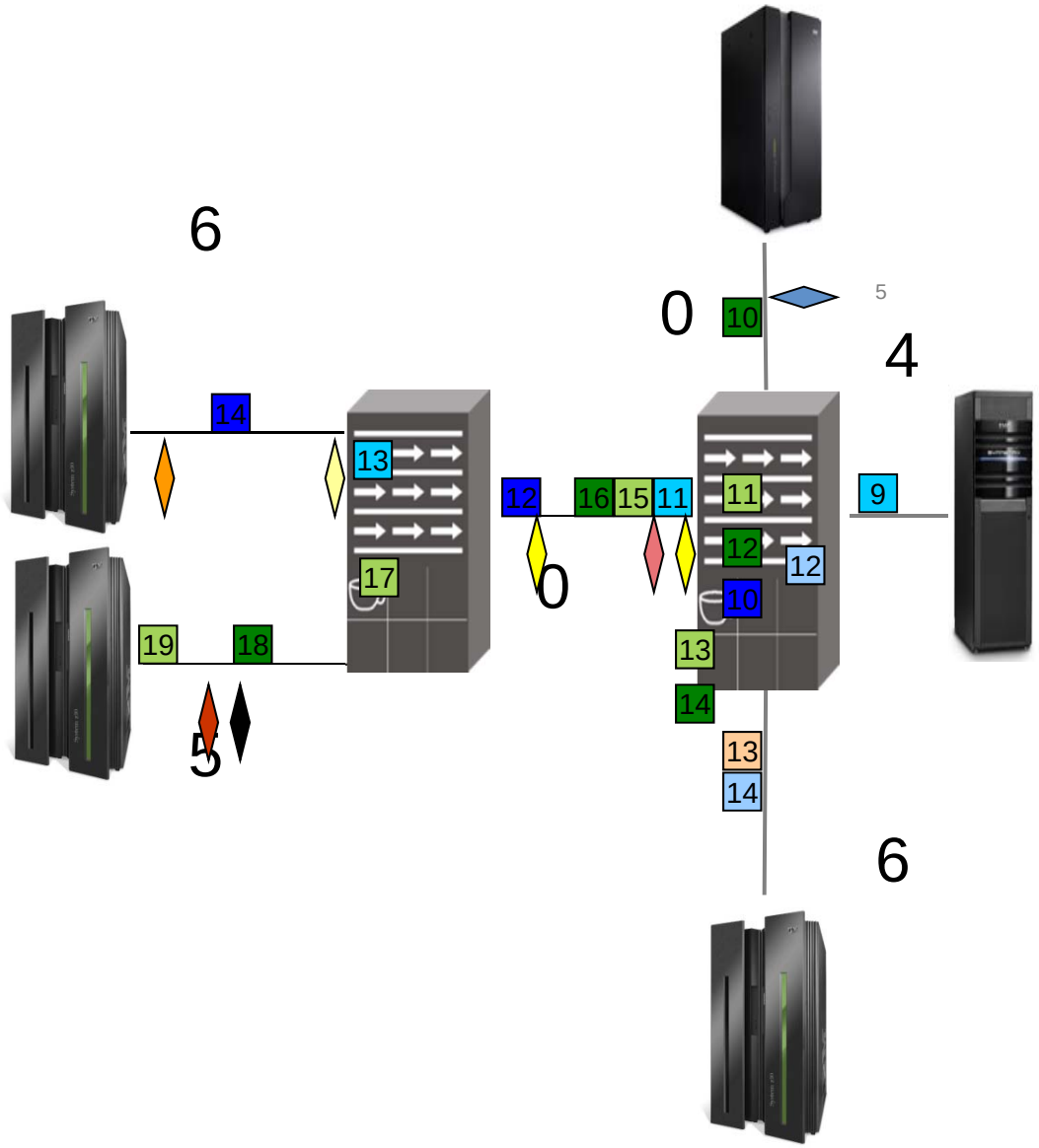

6
0
4
9
13
12
11
14
13
10
12
11
8
10
9
1
16
15
11
5
12
7

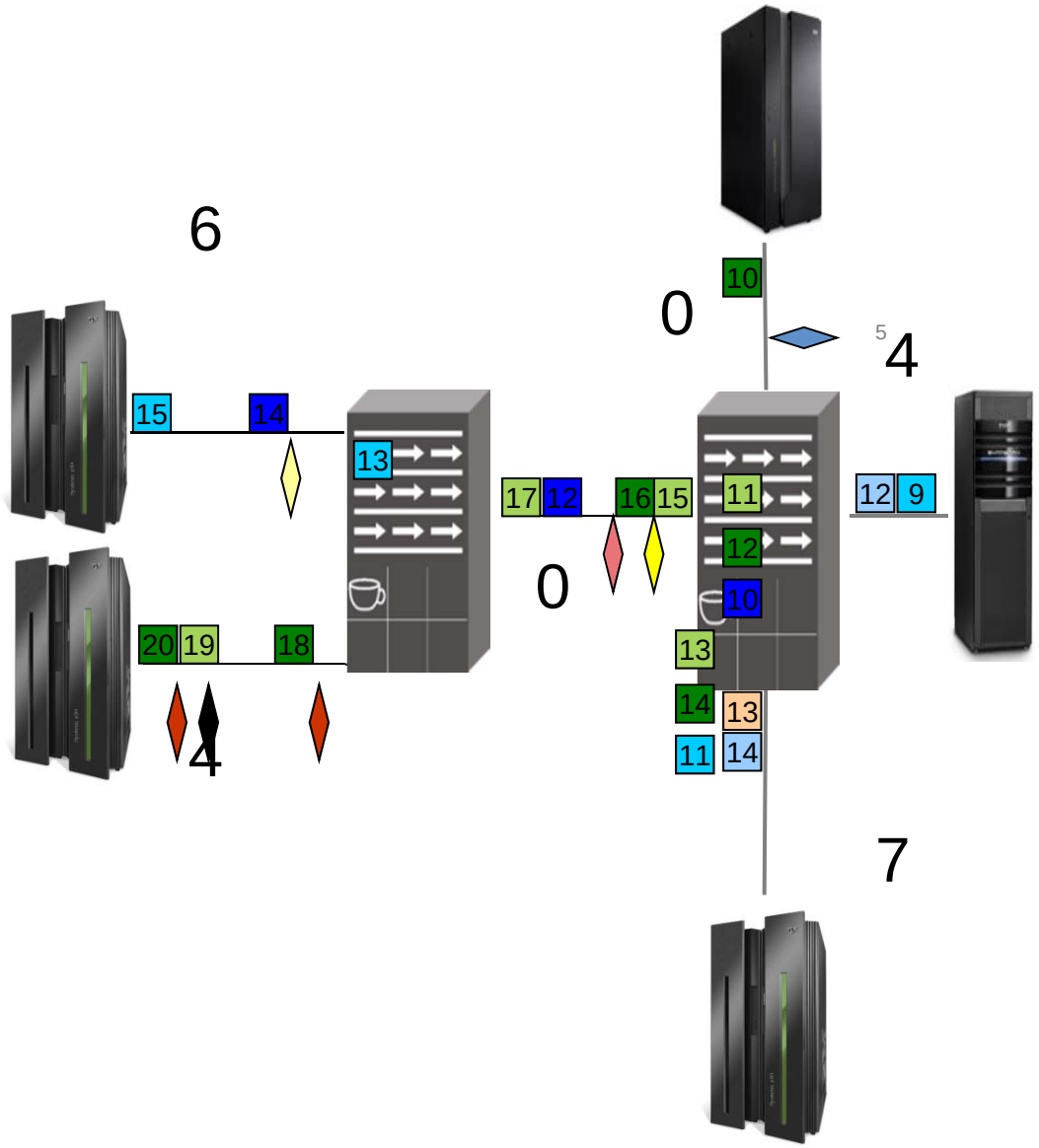

6
0
4
4
13
12
11
14
13
10
12
10
8
9
11
15
1
11
17
16
5
12
7

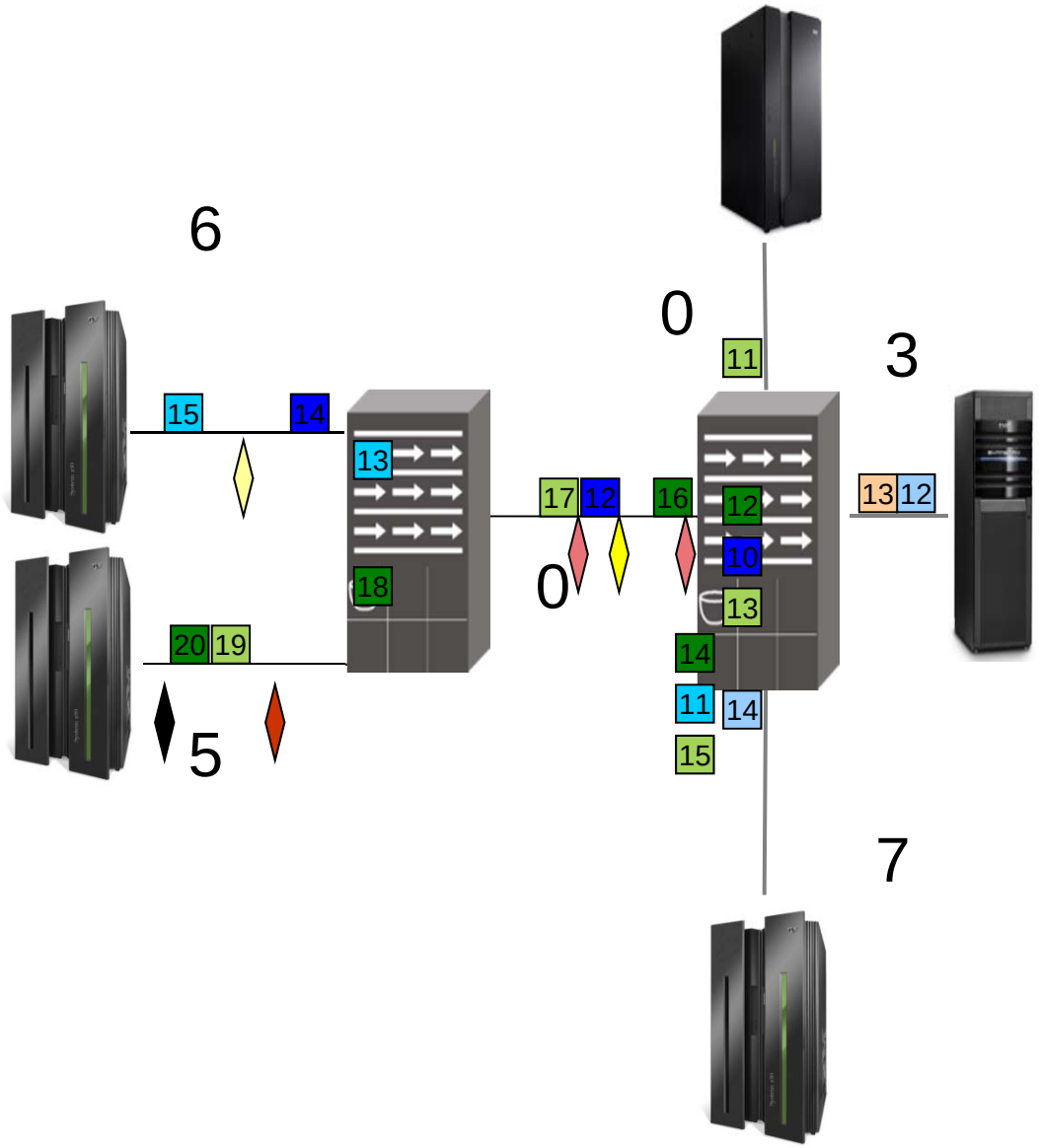

6
0
4
13
12
15
11
14
13
10
10
11
8
9
11
1
12
17
16
5
12
13
7

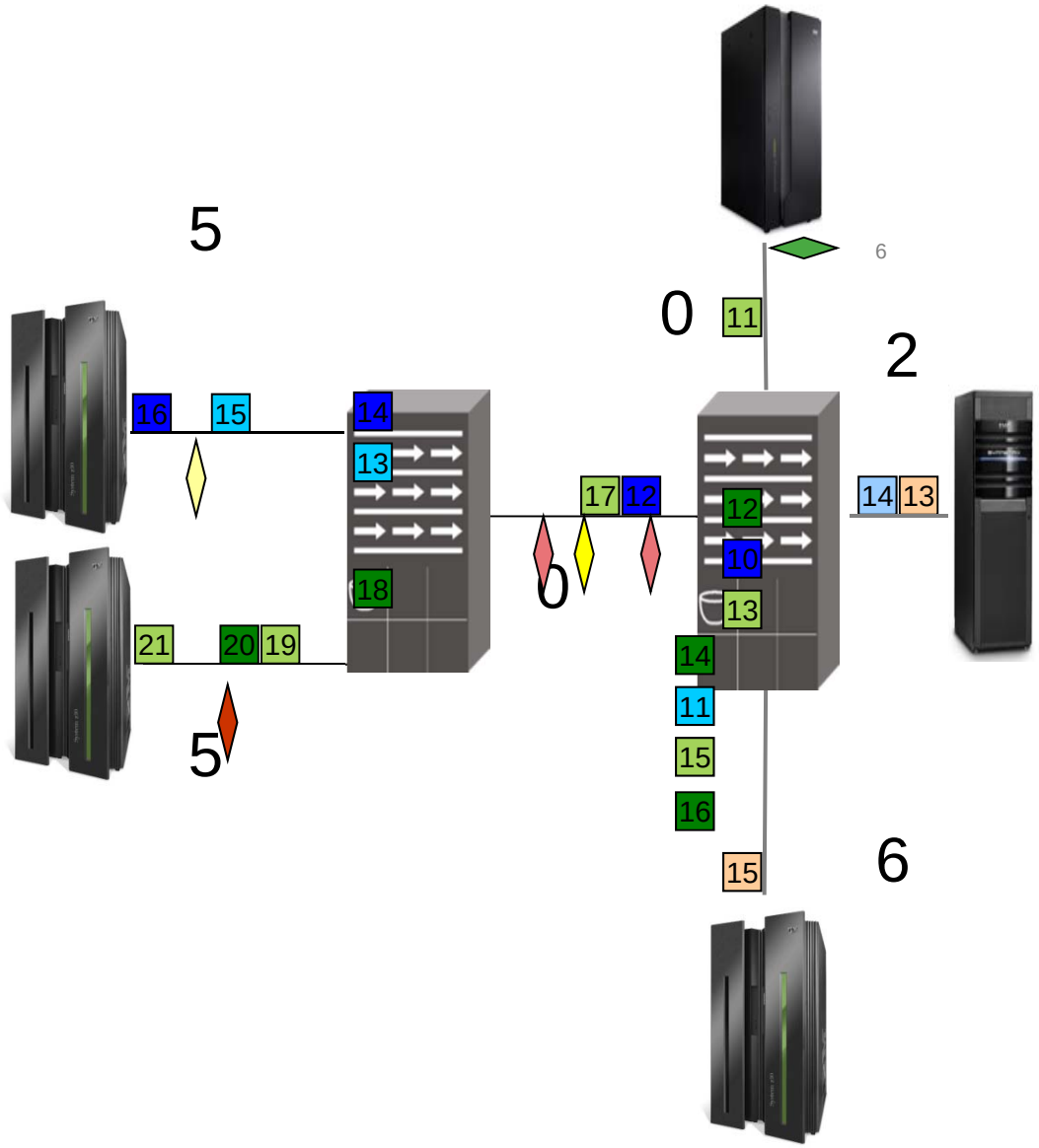

6
0
4
14
13
12
16
15
11
14
13
10
11
9
11
0
18
17
12
10
5
12
13
14
6

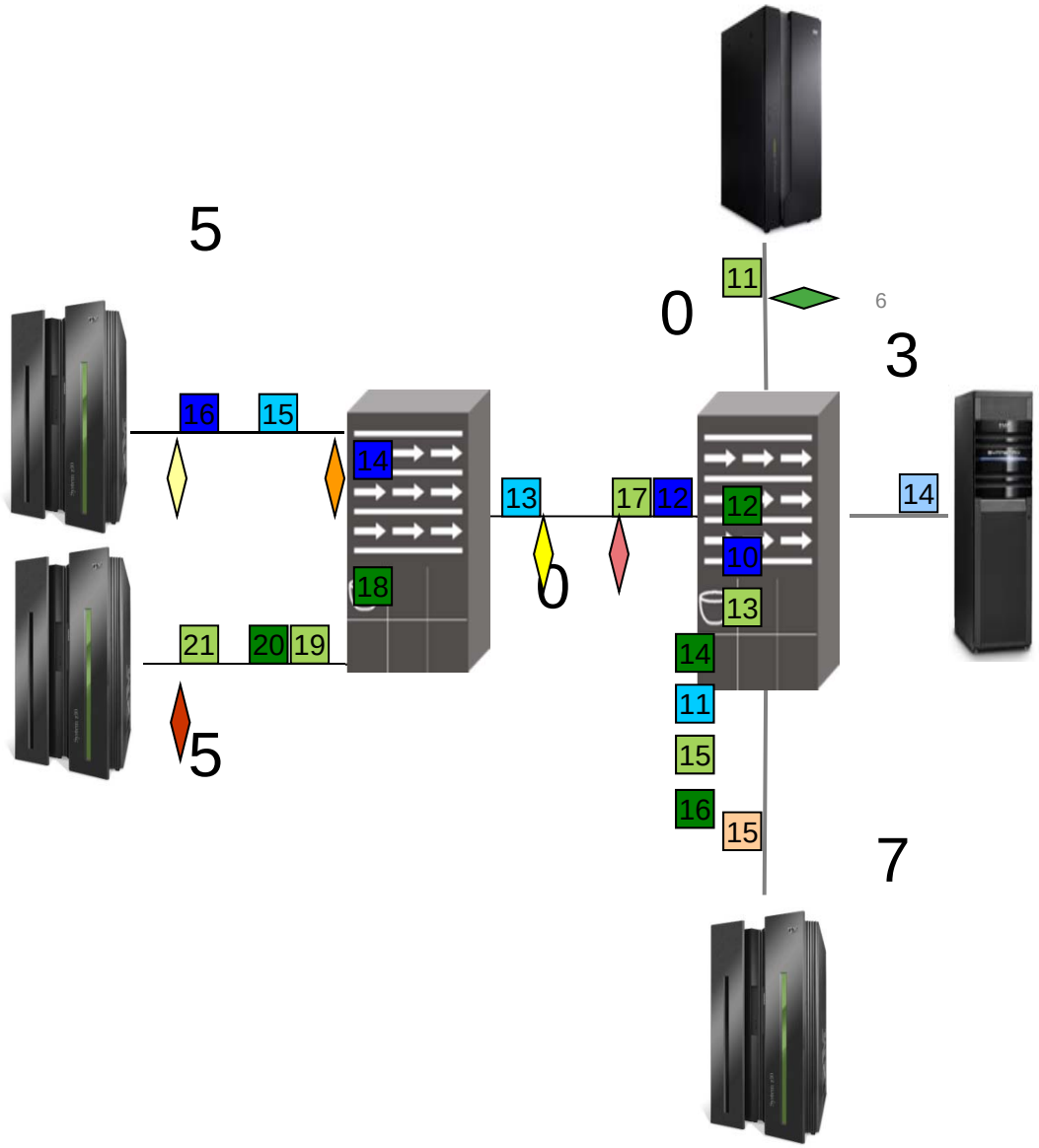

6
5
0
4
10
14
13
12
16
15
11
14
9
11
12
0
10
13
12
18
17
5
13
14
6

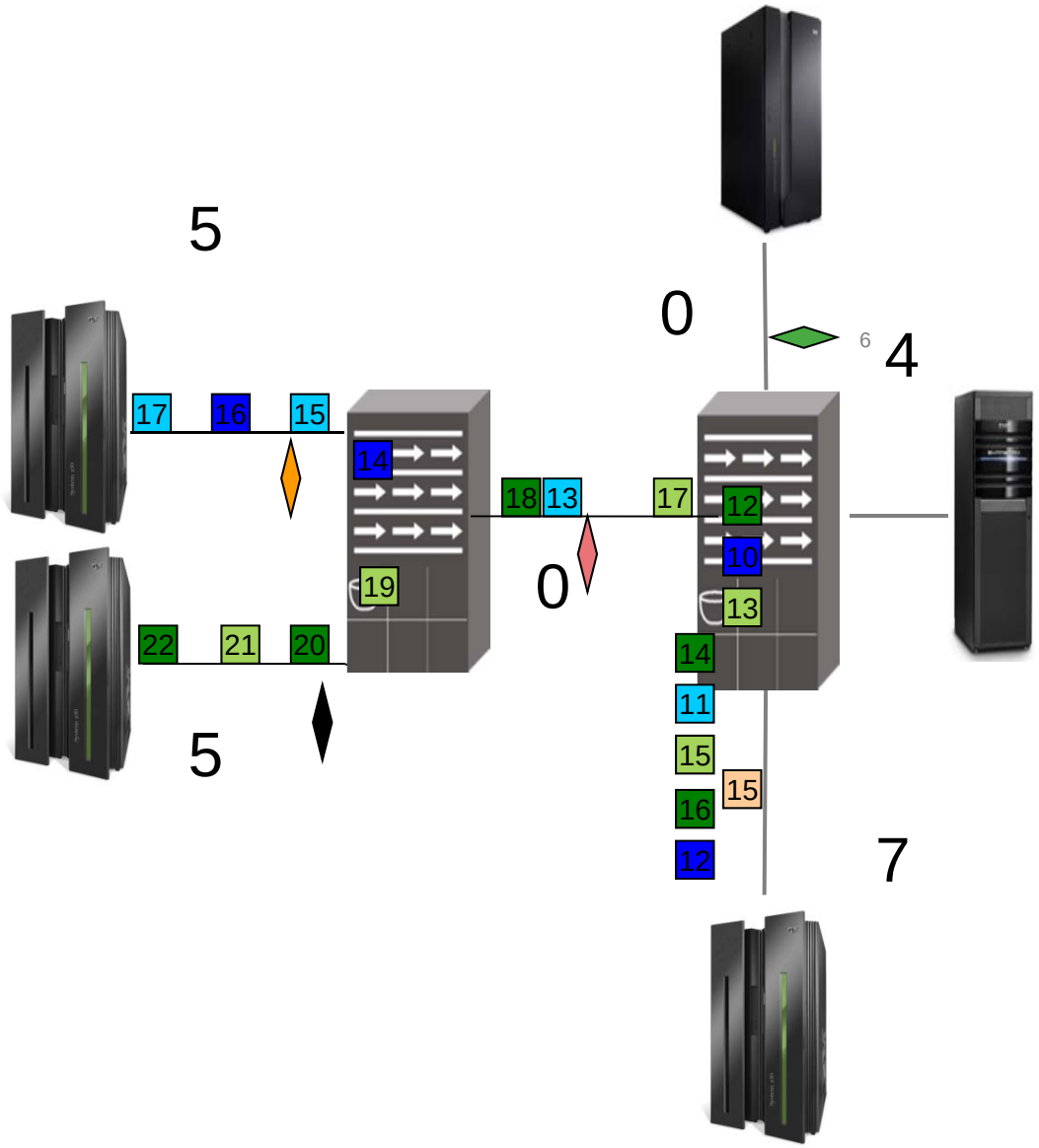

6
0
10
5
4
14
13
12
16
15
11
11
9
12
12
17
0
10
19
18
13
14
5
13
14
6

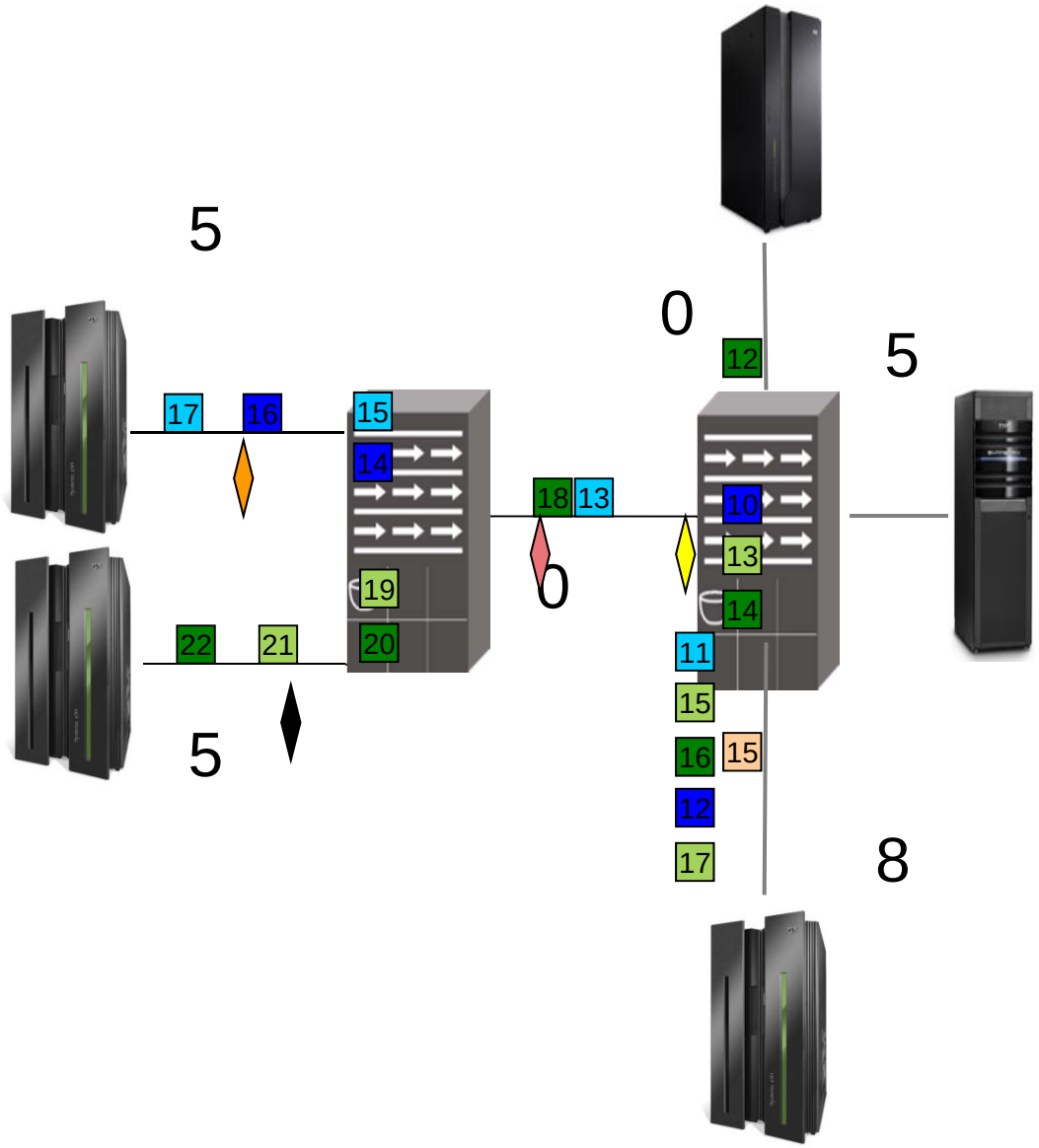

6
0
5
4
15
14
13
17
12
16
15
11
12
9
12
10
0
20
19
18
13
14
13
11
14
4
7

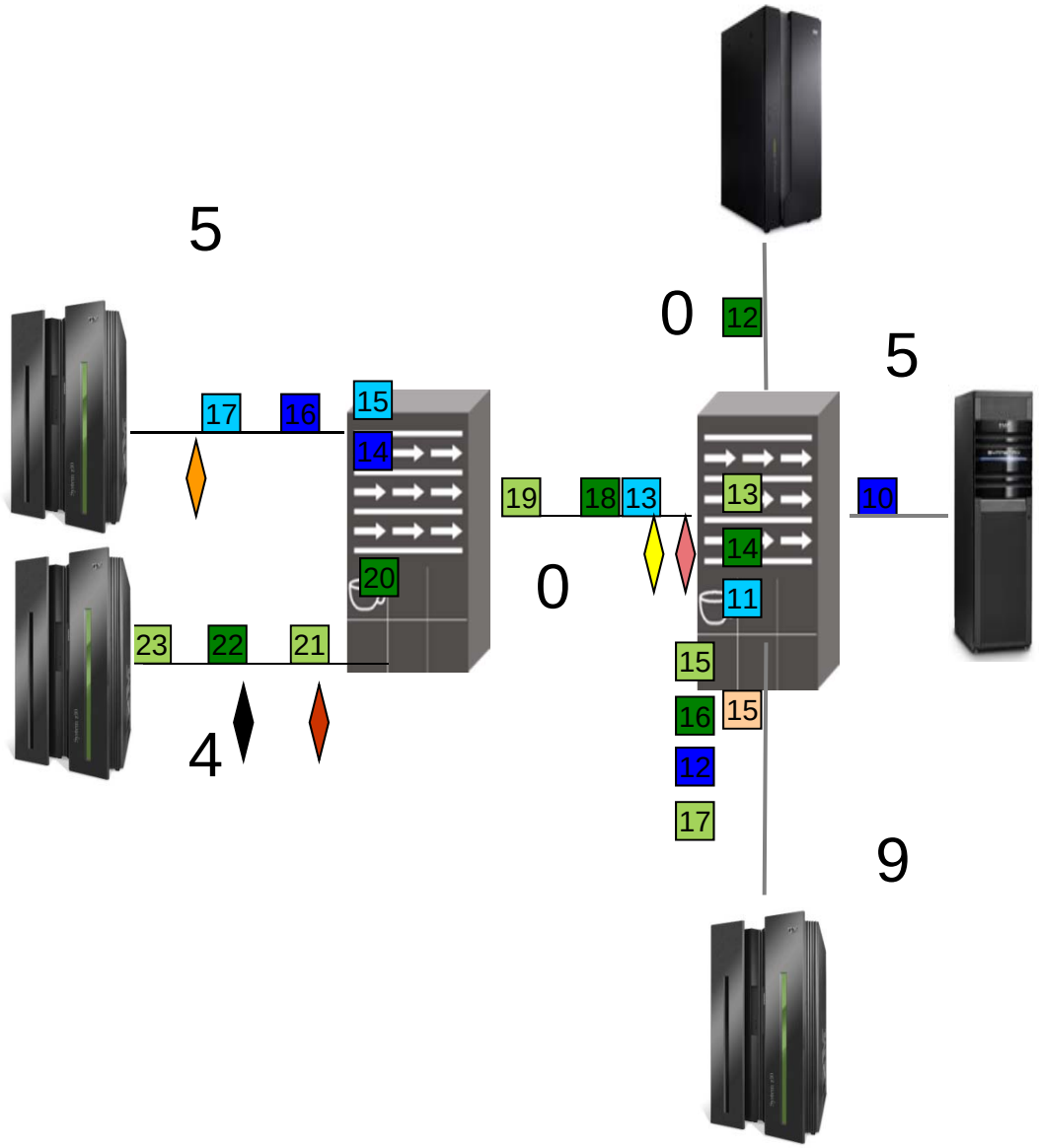

6
0
3
11
15
14
13
17
12
16
12
13 12
10
18
0
20
19
13
14
11
14
15
5
7

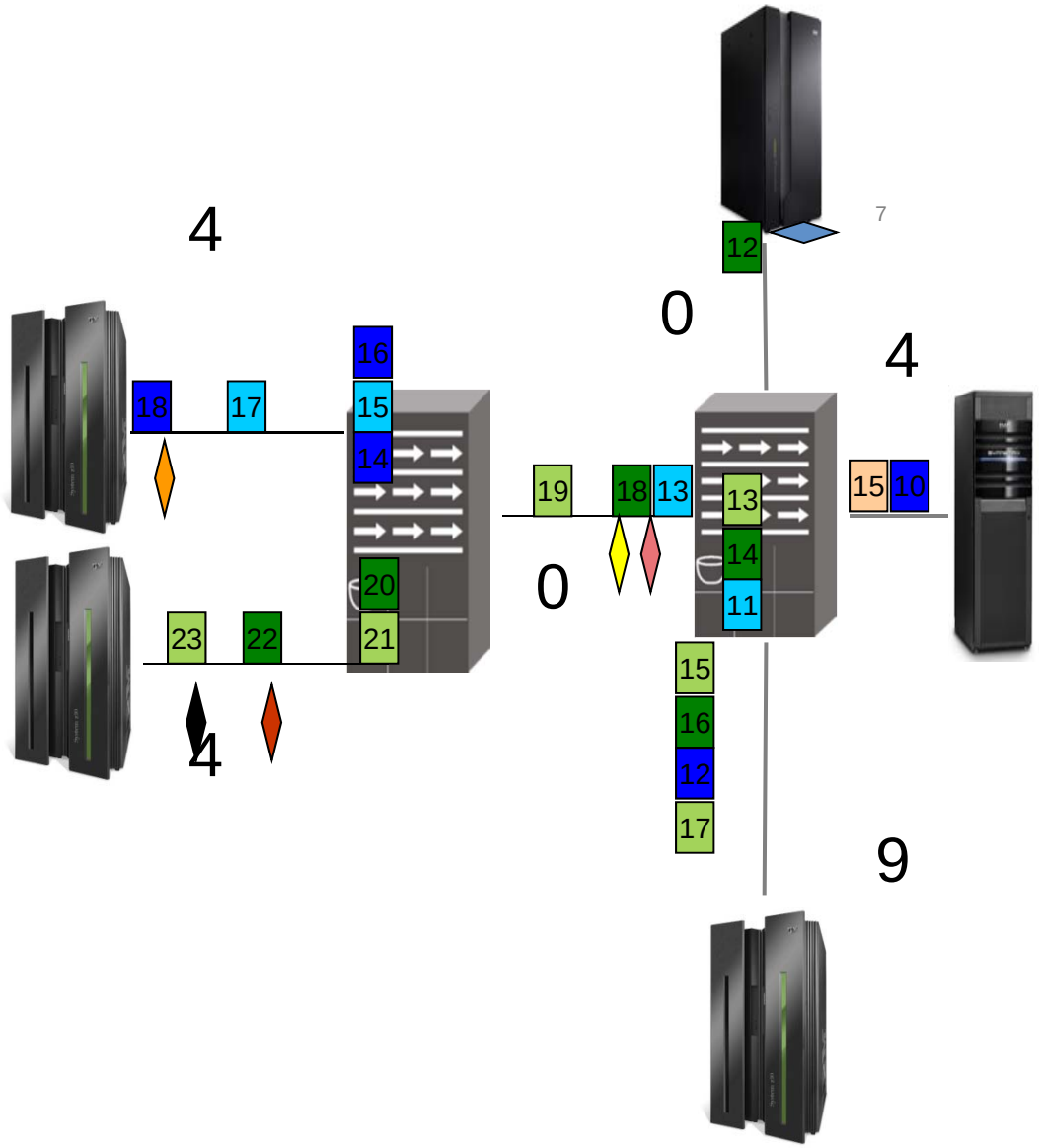

5
6
0
11
2
16
15
14
13
17
12
14
13
12
10
13
18
0
21
20
19
14
11
15
16
5
15
6

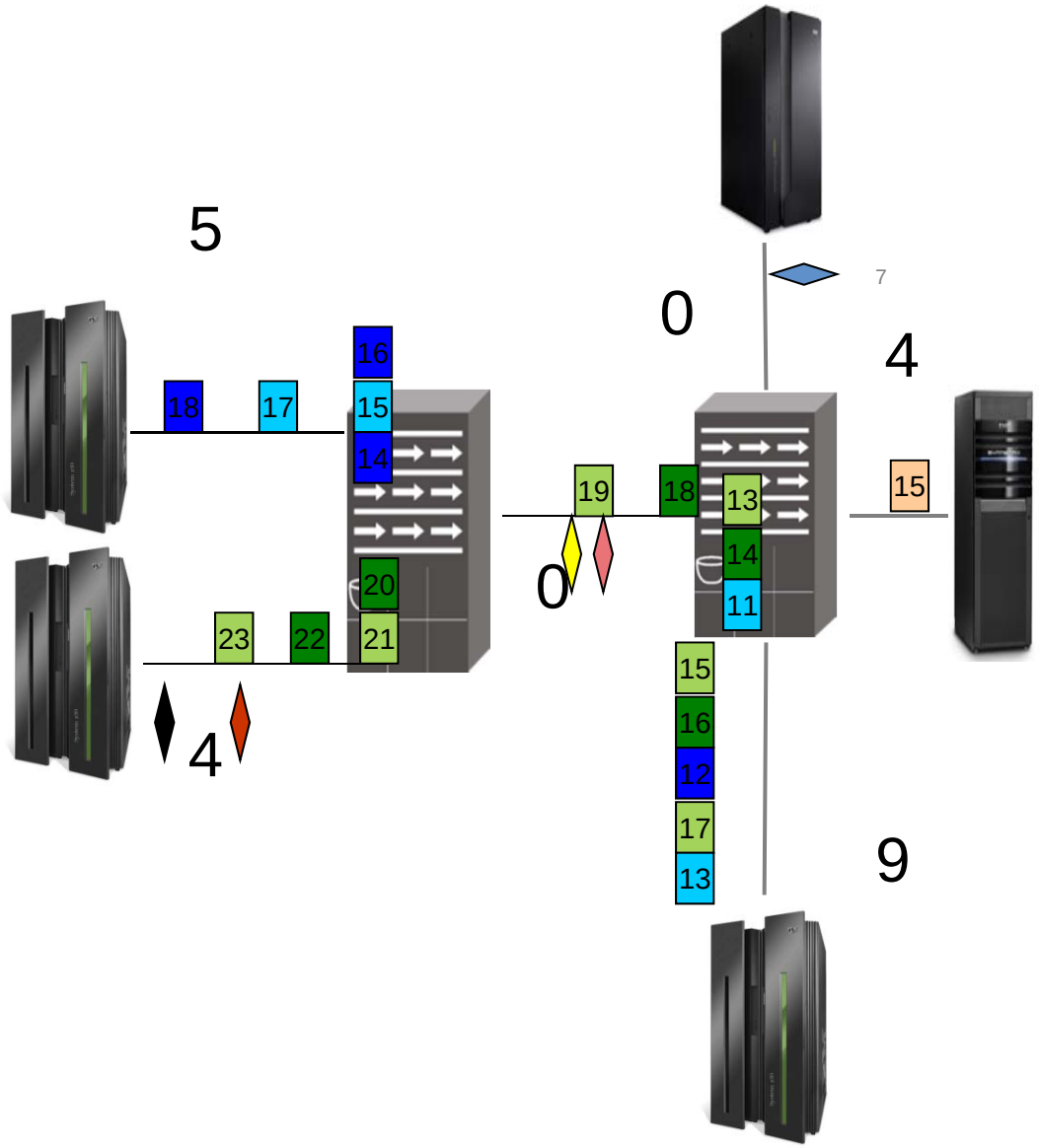

5
0
6
3
16
15
14
13
17
12
12
14
10
18
0
13
21
20
19
14
11
5
15
16
15
7

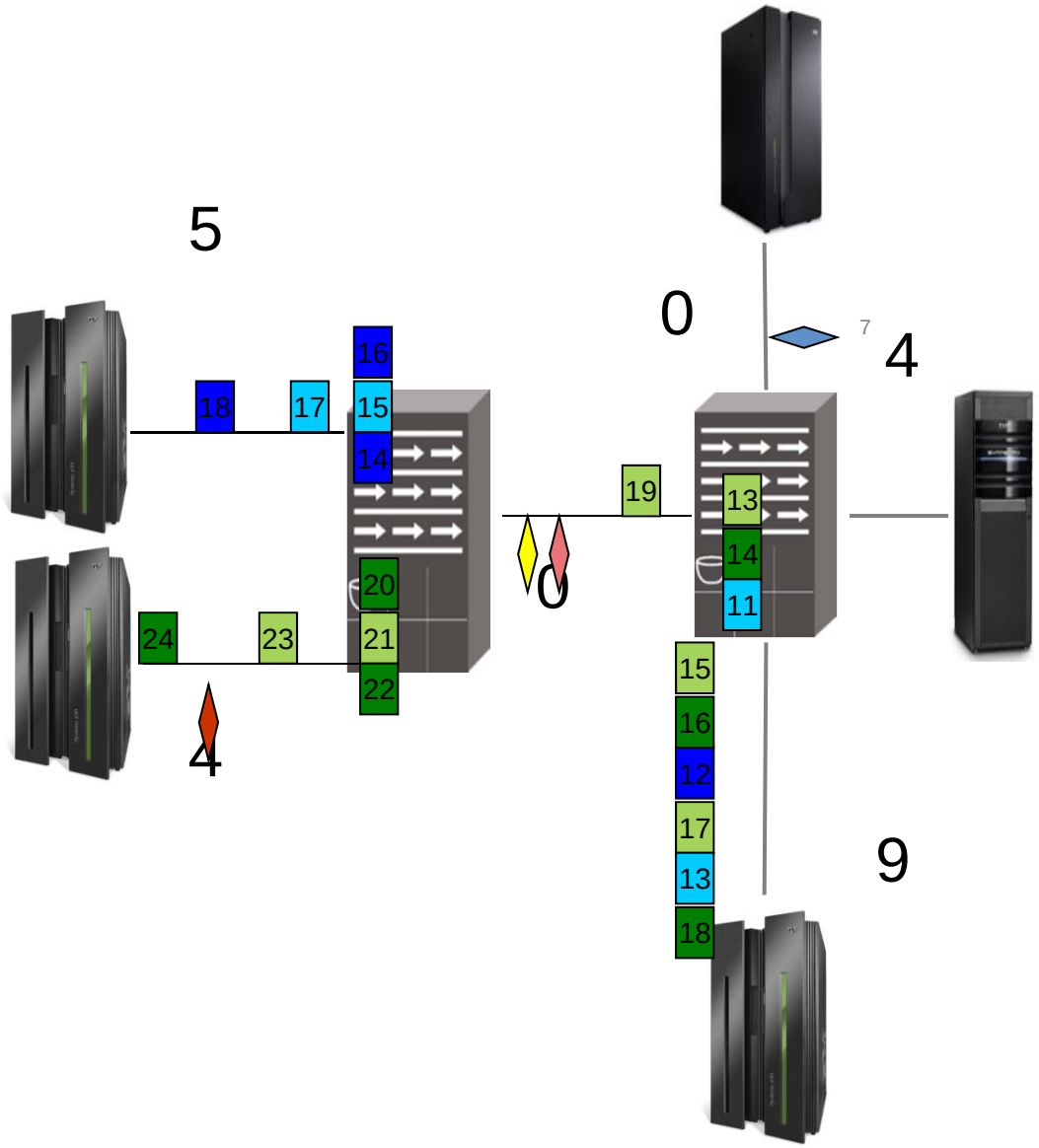

5
0
6
4
17
16
15
14
18
13
17
12
10
13
19
0
22
21
20
14
11
15
15
16
12
5
7


5
0
5
17
16
15
14
18
13
10
13
19
0
14
22
21
20
11
15
16
15
12
17
5
8


5
17
16
15
14
20
23
22
21
4
19
18
13
0
0
12
13
14
11
15
16
15
12
17
5
10
9


4
18
17
16
15
14
20
21
23
22
4
0
19
18
13
13
14
11
15
10
4
12
0
15
16
12
17
9


5
18
17
16
15
14
20
23
22
21
4
0
19
18
13
14
11
0
7
4
15
15
16
12
17
13
9


5
0
7
4
16
18
17
15
14
19
13
14
11
20
24
23
21
22
0
4
15
16
12
17
13
18
9

What Can Be Done About A Slow Drainer?

- Difficult to identify the culprit
  - Sometimes issues are intermittent
  - There may be more than one gremlin

- Tools are available to help!
  - “Bottleneck Detection”

See Also:
https://elabnavigator.emc.com/vault/pdf/FC_SAN.pdf
https://www.ibm.com/developerworks/mydeveloperworks/blogs/sanblog/entry/how_to_deal_with_slow_drain_devices20?lang=en
http://www.brocade.com/downloads/documents/html_product_manuals/NA_SAN_1130/wwhelp/wwhimpl/common/html/wwhelp.htm#href=Ch_Performance.31.7.html&single=true

What Can Be Done About A Slow Drainer?

- Creates an alert when a port(s) have zero credit for a user selectable amount of time
- Use detection mechanisms to monitor ISLs ports and F_Ports

SAN Performance Gotcha's

- Oversubscription
  - Trying to send more than can be received
- Slow draining
  - Unable to receive what is being sent in a timely manner

Culprits

- Oversubscription
  - Big pipe to small pipe
- Signal degradation
  - Failing SFP
  - Poor connections
  - Insufficient cable hygiene
- Unexpected slow down
  - Big pipe to small pipe (16G to 2G)
  - Narrow bridge (reduced ISL bandwidth)
- Odd stuff
  - Long links
  - Zoning conflicts
  - Invalid Attachment / Decommissioned Channel/CU Port
  - Queue depth

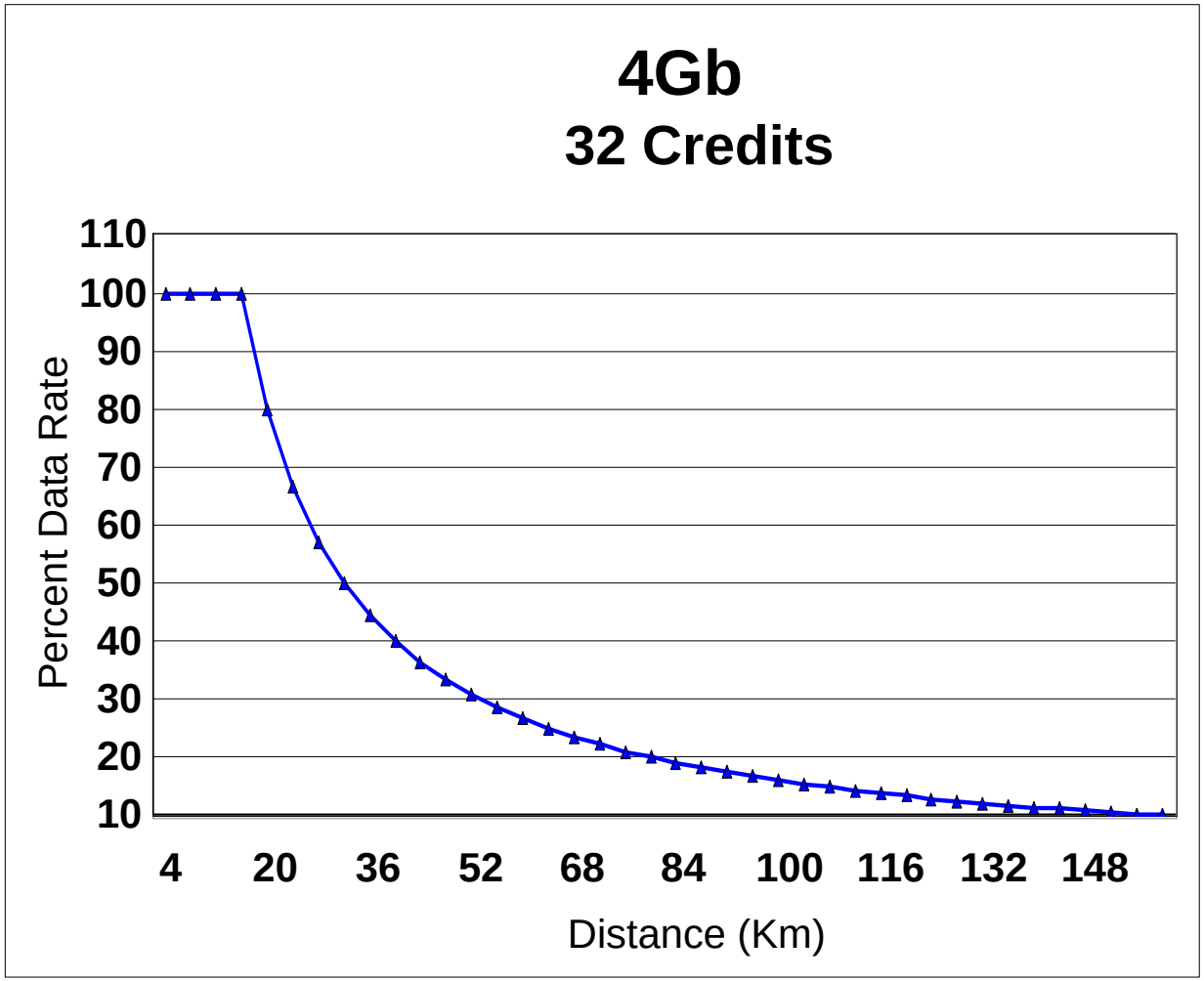
Stuff to Look For

- Physical Errors
  - Detected (Obvious stuff)
    - Logins & Logouts
    - Link Failures
  - Undetected (Not Obvious)
    - CRC Errors
      - *Encoding errors (inside of a frame)*
    - Code Violation Errors
      - *Loss of synchronization*
      - *Loss of signal*
    - Loss of Sync
      - *Encoding errors (outside of a frame)*
- Logical Errors
  - Credit Starvation
    - Time at zero credit
    - Link Resets
    - Class 3 Discards
  - Response Times
    - High CMR
    - Exchange Completion Times
    - SCSI Reservation Conflicts

Other Neat Stuff

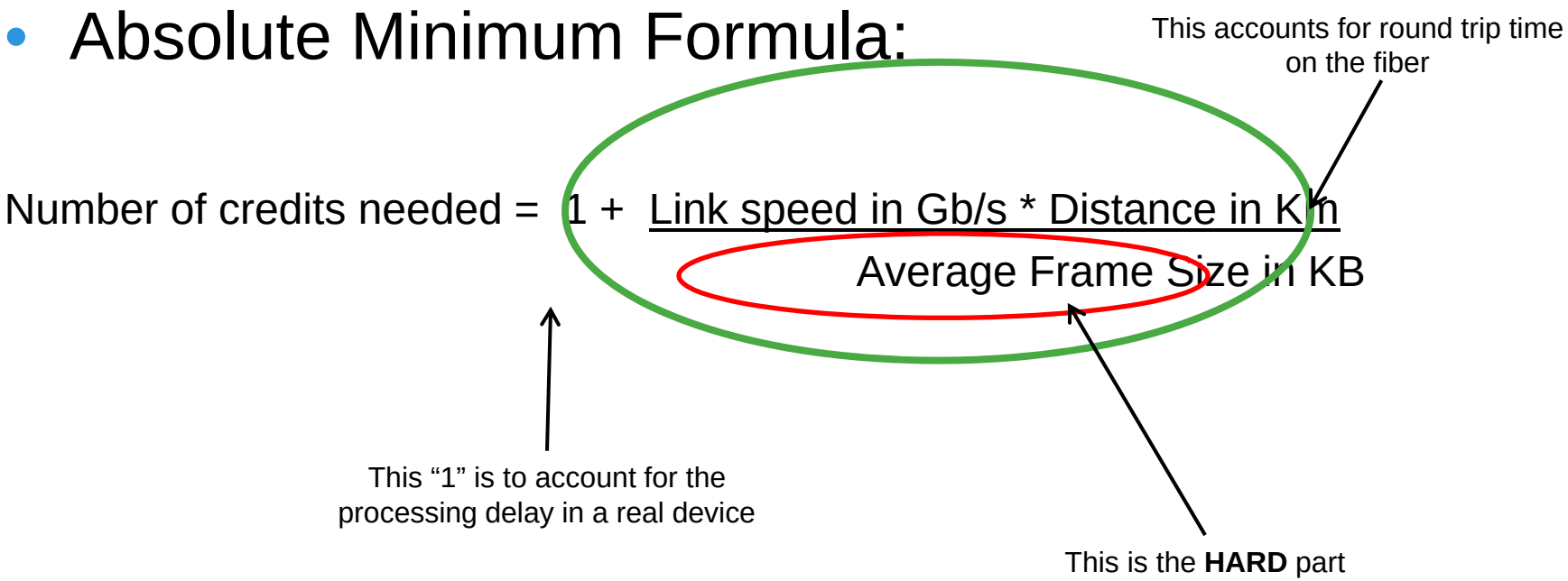
BUFFER CREDITS

How Does Credit Affect Throughput?

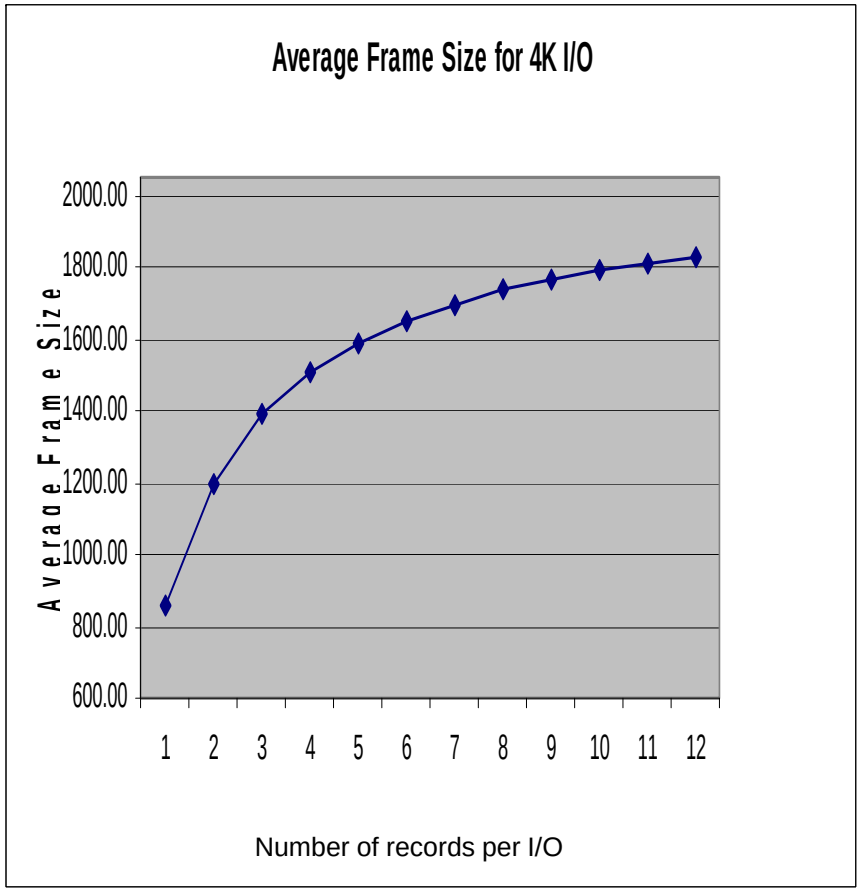


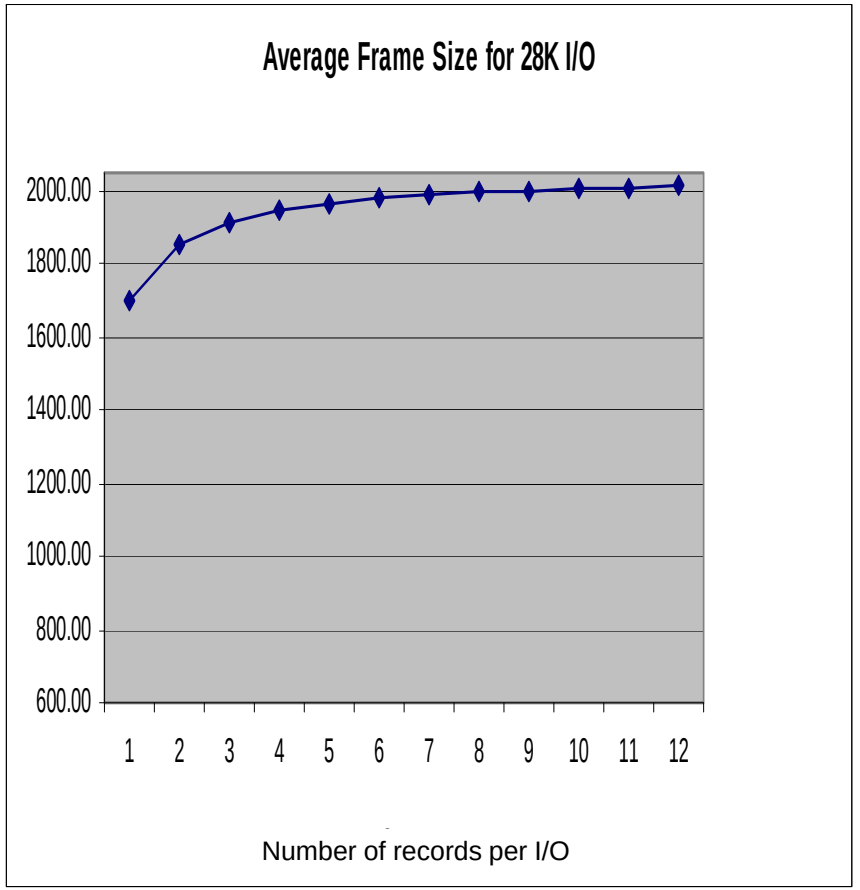
How Much Credit Do I Need For A Given Distance?

- Absolute Minimum Formula:



Average Ficon Frame Size vs Block Size





How Should I Configure My BB Credit?

- Local links under 1Km in length:
  - Even at 8G and small frame size, < 16 credits required

1 + ((8*1) / .8) = 11

- Provision ISLs with extra headroom
  - Use formula, and round UP

1 + ((8*10) / 1.5) = 54

EXCHANGES

Urban Legend:
FICON uses fewer Exchanges Than FCP

- In Ficon, each concurrent I/O operation uses two Exchanges
  - One unidirectional Exchange for IUs from the Channel to the CU
  - A different unidirectional Exchange for IUs from the CU to the Channel
- The PAIR is commonly know as a “Ficon Exchange”

How many Exchanges do I need?

- Little's Law states:
  - *The number of "things" in a system can be determined by multiplying the average arrival rate of those "things" by the average time each "thing" stays in the system.*
  - Applied to Ficon:
    - The average number of Exchanges active at any given time = Average I/O rate * Average response time
    - Example: 5000 Ficon I/Os / Second on a given channel with .4ms service time[1] needs 2 Active Exchanges (pairs) at any given time

[1] The amount of time the I/O is active in the channel

Urban Legend:
Ficon is More Sensitive to Errors than FCP

- Is a Ficon frame more likely to get lost, damaged or corrupted than FCP?
  - No, the probability is the same
- When a Ficon frame gets lost, damaged or corrupted, is the recovery action different from FCP?
  - Not really. They are both retried, FCP by the Device Driver, Ficon by IOS/ERP

So What are the Differences?

- z Operating Systems tend to provide more detailed messages
- Ficon does provide additional debug data and actions
  - RNID
  - Link Error Status Blocks
  - Extensive State Change Processing

Table 89 - Link Error Status Block format for RLS command

Bits / Word	31 .. 00
0	Link Failure Count
1	Loss-of-Synchronization Count
2	Loss-of-Signal Count
3	Primitive Sequence Protocol Error
4	Invalid Transmission Word
5	Invalid CRC Count

Source: FC-FS-3 INCITS/T11 Draft Standard v0.92

See www.t11.org

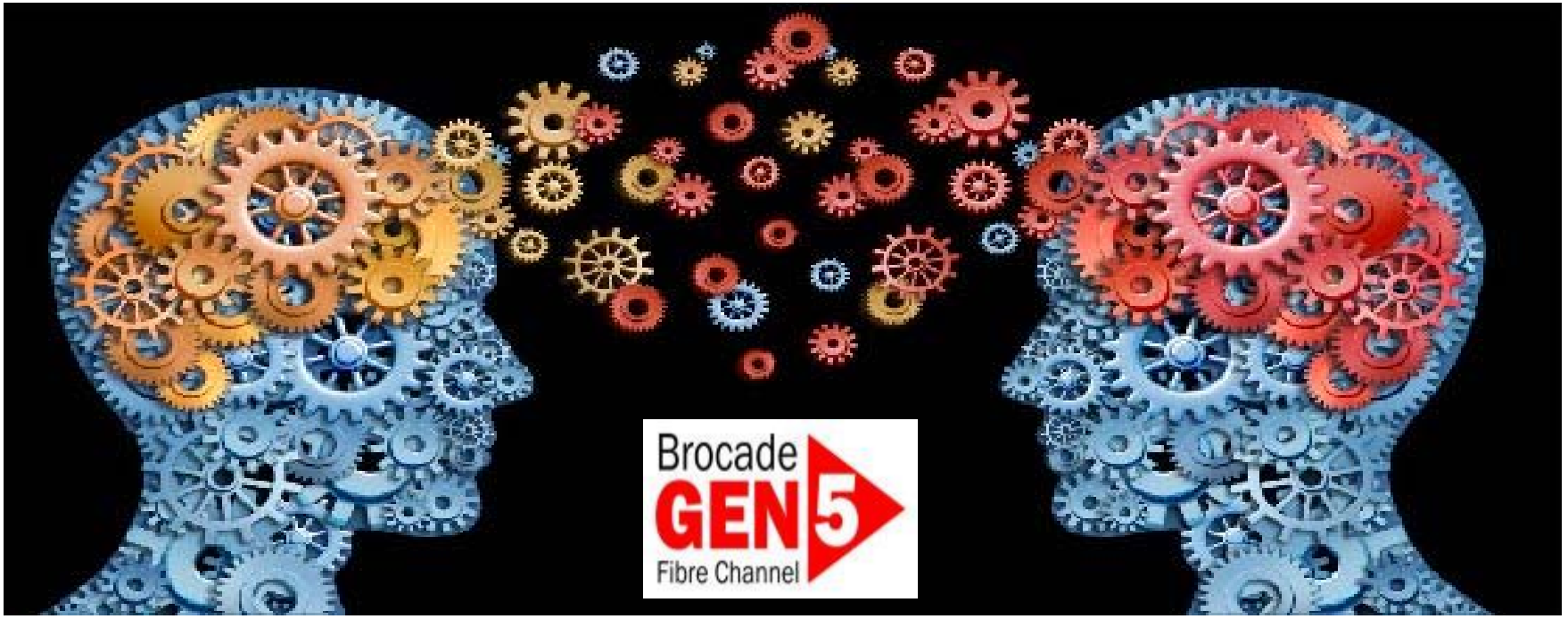
Summary

- Buffer Credits
  - Distance
  - Flow Control
- Exchanges
  - Unidirectional
  - Bidirectional

.....Newly Scheduled Presentation.....

Brocade SAN and FICON Update

Please consider attending to discover the innovation of Brocade's Gen 5 Fibre Channel Architecture



Wednesday August 14, 2013 – 11:00pm to 12:00pm -- Session 14482

SHARE, Boston, August 2013

Buffer-to-Buffer Credits, Exchanges, and Urban Legends

Session 14281

THANK YOU!

QR Code

REFERENCES

Speaker Biography

- Lou Ricci
  - EMC
    - 1 Year
  - IBM
    - 34-years
    - 24-years in channel development
    - An inventor of FICON
    - FICON Firmware Team Leader
- Contact Information
  - Louis.ricci@emc.com

Speaker Biography

- Howard L. Johnson
  - BROCADE
  - Technology Architect, FICON
    - 29 years technical development and management
- Contact Information
  - howard.johnson@brocade.com

BONUS SLIDES

End to End Credit

- Device to Device Flow Control
  - Between source and destination
    - Not the links
  - Similar to buffer-to-buffer flow control
    - At N_Port Login
    - Report available receive buffers (EE_Credit)
    - Transmitter counts buffers transmitted (EE_Credit_CNT)
    - Receiver acknowledges frame (ACK)
      - *ACK 1 (a single data frame in a sequence) – most common*
      - *ACK n (several (N) consecutive data frames in a sequence)*
      - *ACK 0 (all data frames in a sequence) – not used*

Virtual Channels

- Technology to allocate BB_Credits to particular data flows
  - Class F traffic has one data flow
  - Assigned with Zoning by using special Zone names