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

Lou Ricci, IBM
Howard L. Johnson, Brocade

8 August 2011 (3:00pm – 4:00pm)
Session 9931
Room Europe 7

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?
    - How do cascades Directors work?
- 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?
- Error Sensitivity
  - Is FICON more sensitive to errors than FCP?
    - How sensitive are FICON frames to loss or corruption?
    - What recovery actions are taken?
    - What are the differences with FCP?

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

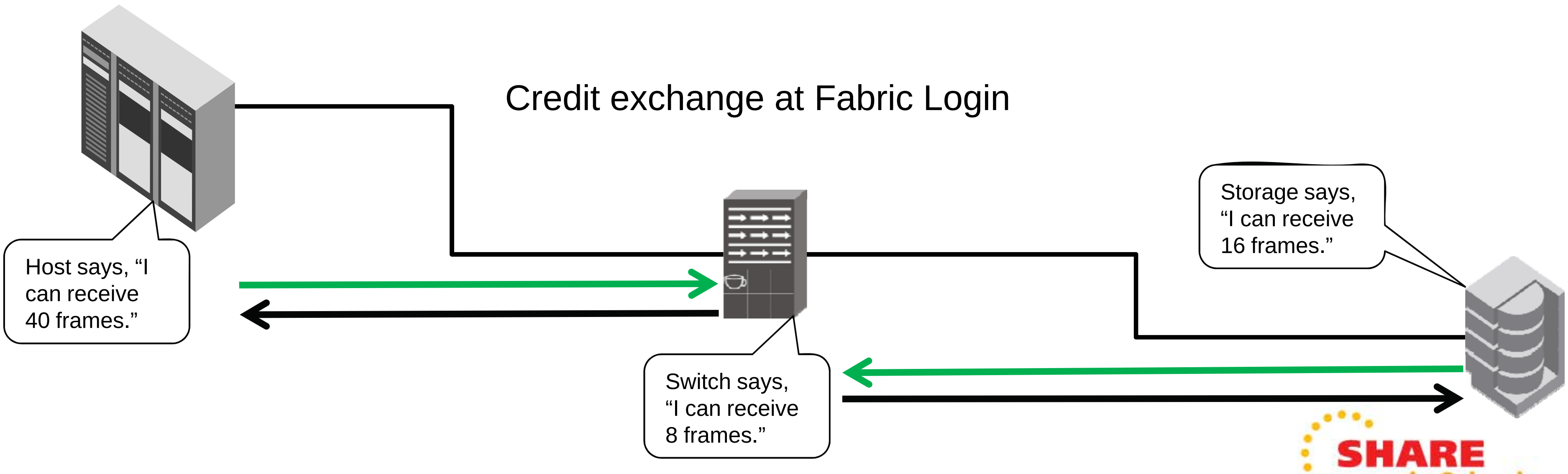
What is Buffer-to-Buffer Credit?

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

Flow Control

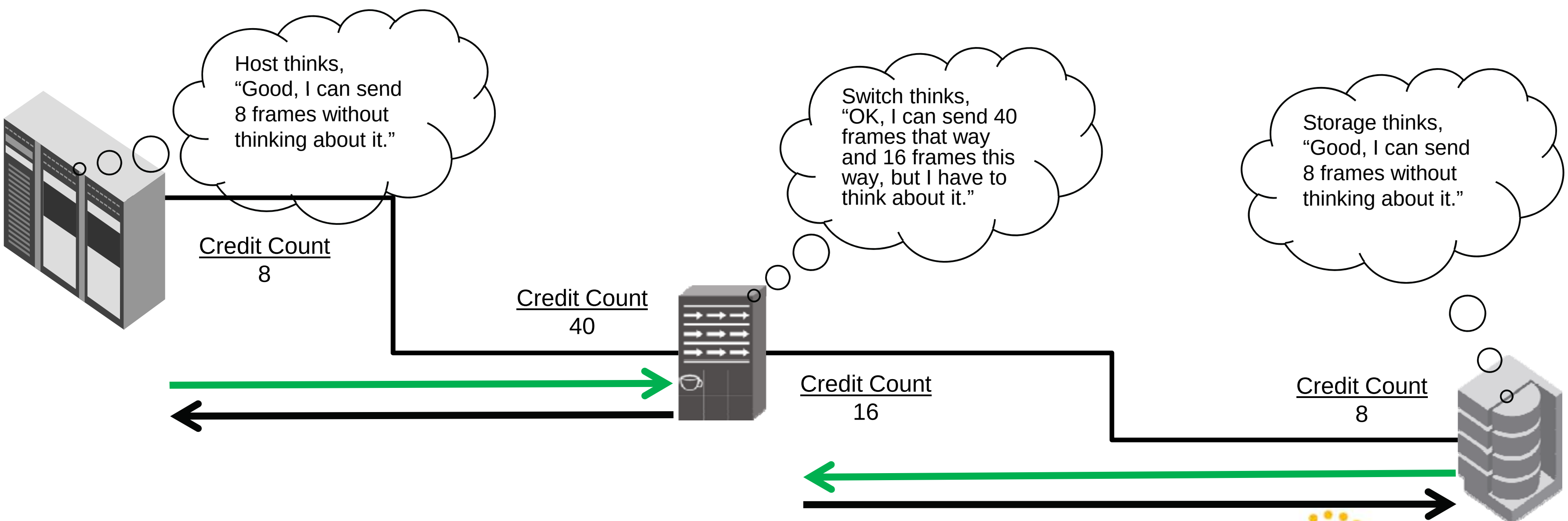
- Related to the devices' ability to receive and process frames
- Manages when frames are coming faster than they can be processed
- Dropped frames occur when frames are arriving too fast to be processed

- 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

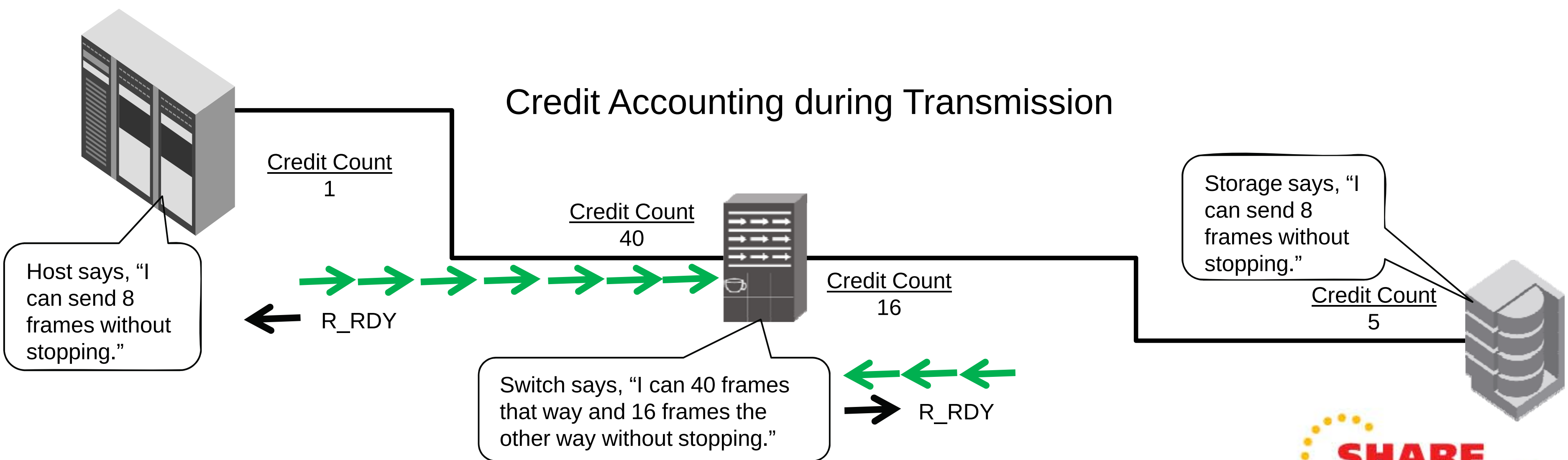


Credit accounting after Fabric Login

SHARE in Orlando 2011

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



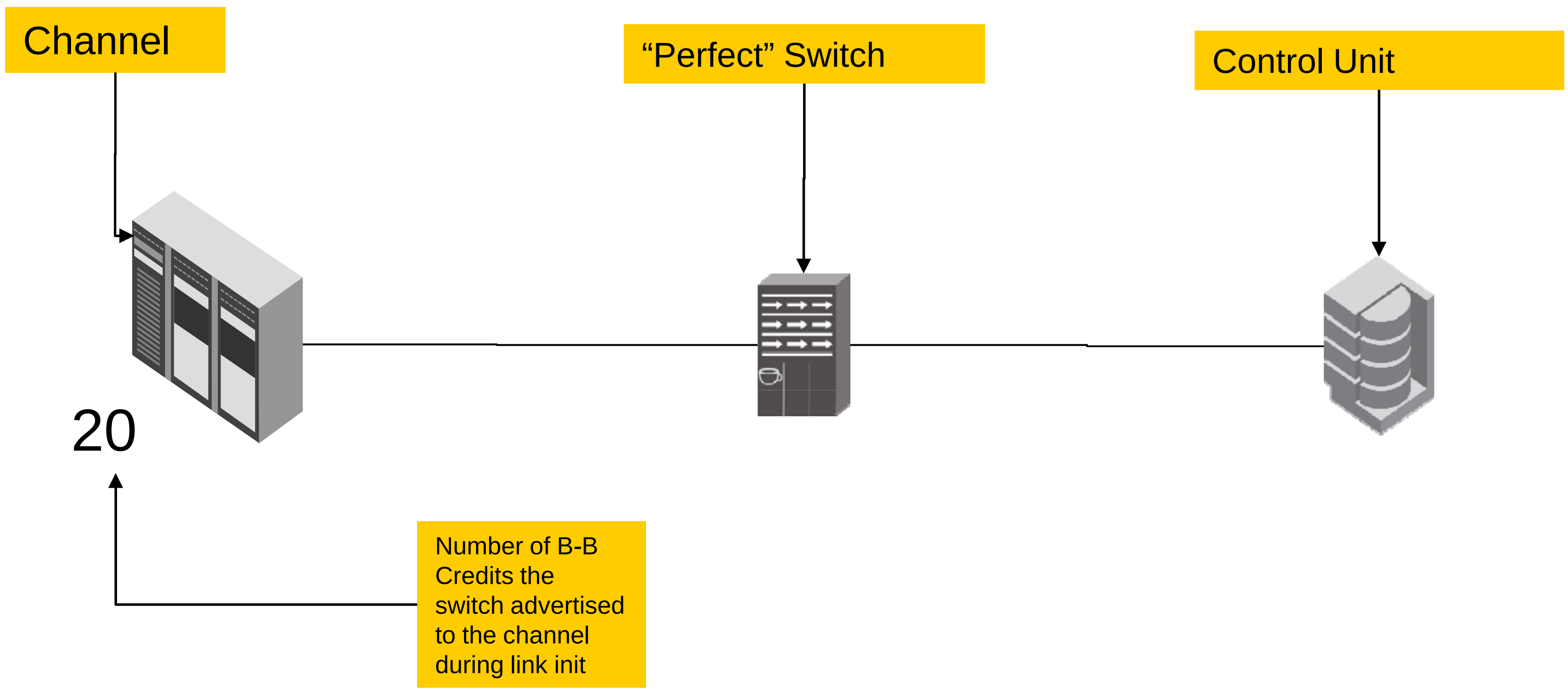
Urban Legend:
Buffer Credits at Zero are a Problem

- 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

Example: A full pipe

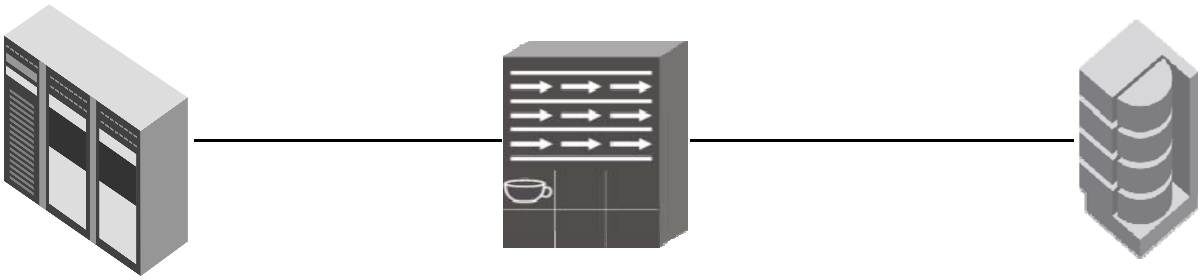
BUFFER CREDITS

Initial Conditions

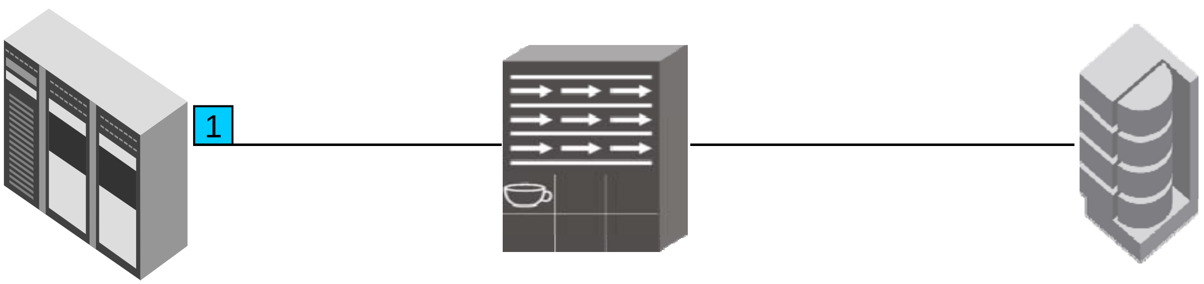


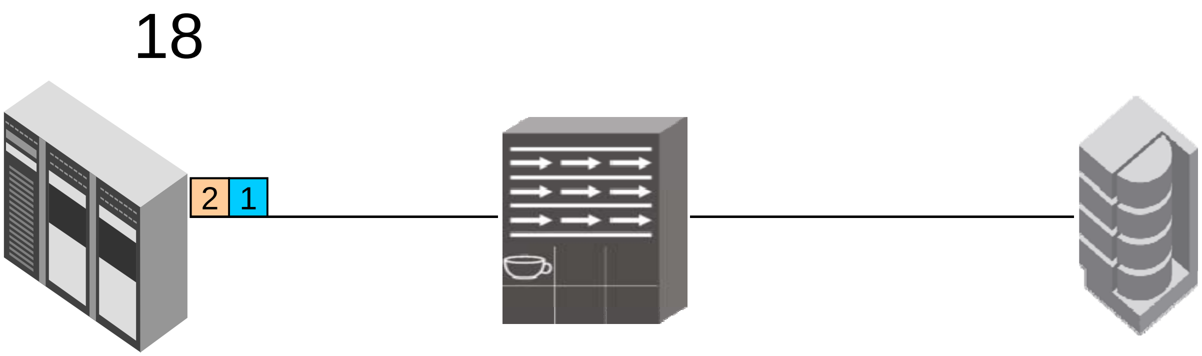
NOTE: In these animations, both the frames and R_RDY's are numbered. This is for illustrative purposes only. In reality, neither the frames nor the R_RDY's are numbered. The arrival of an R_RDY only informs the receiver that A frame has been forwarded, now WHICH frame has been forwarded.

20

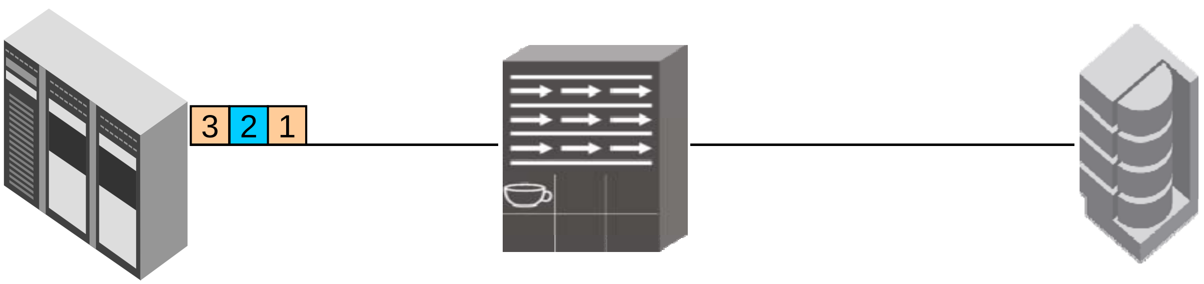


19

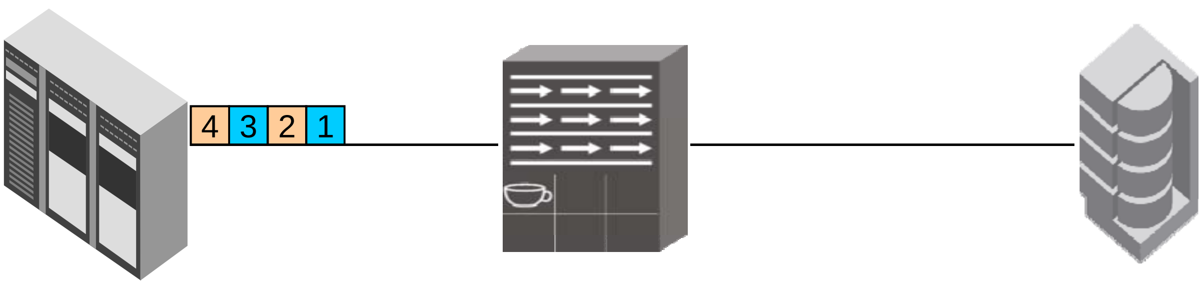



18
2
1

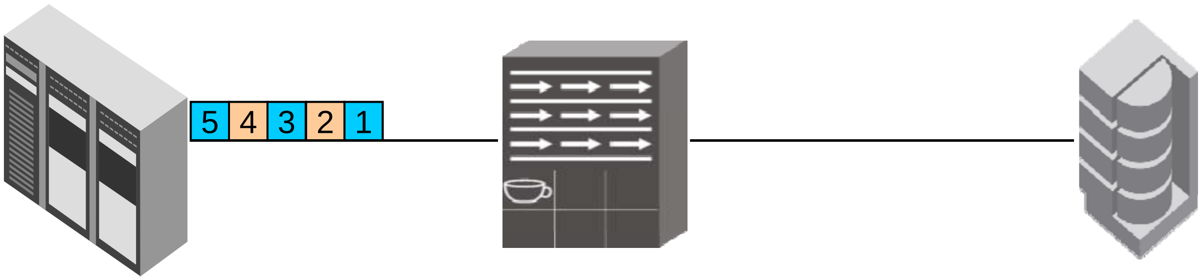
17



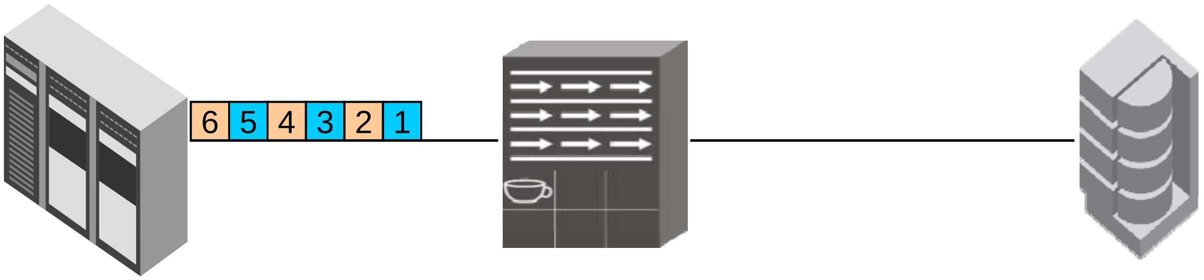
16



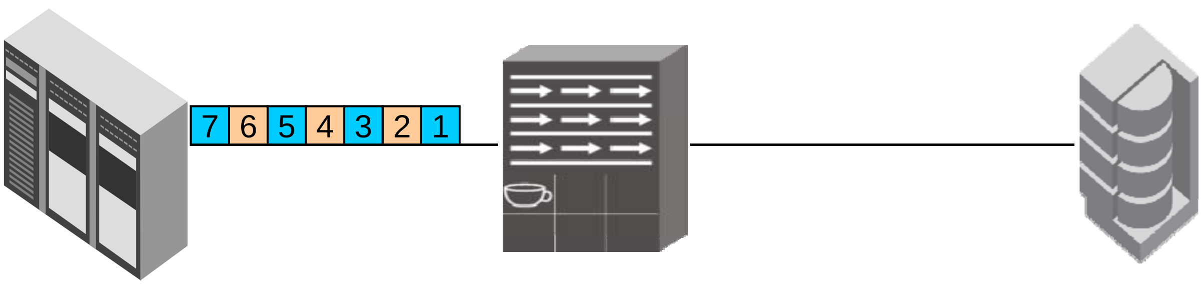
15

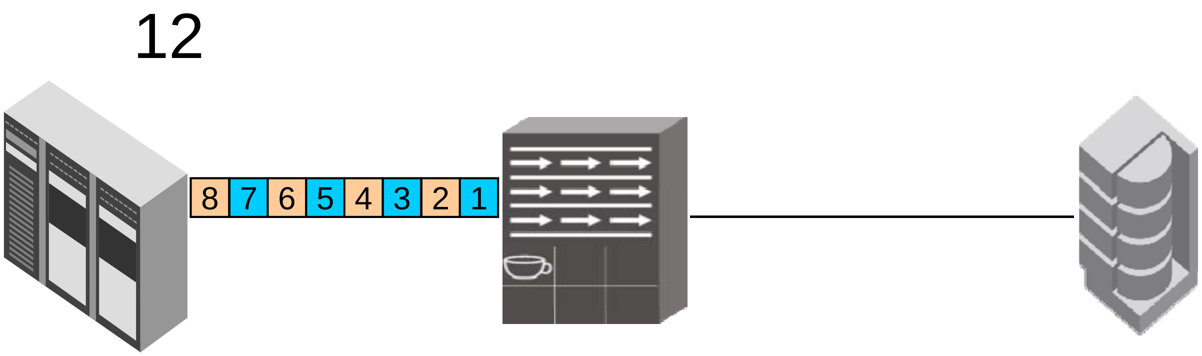


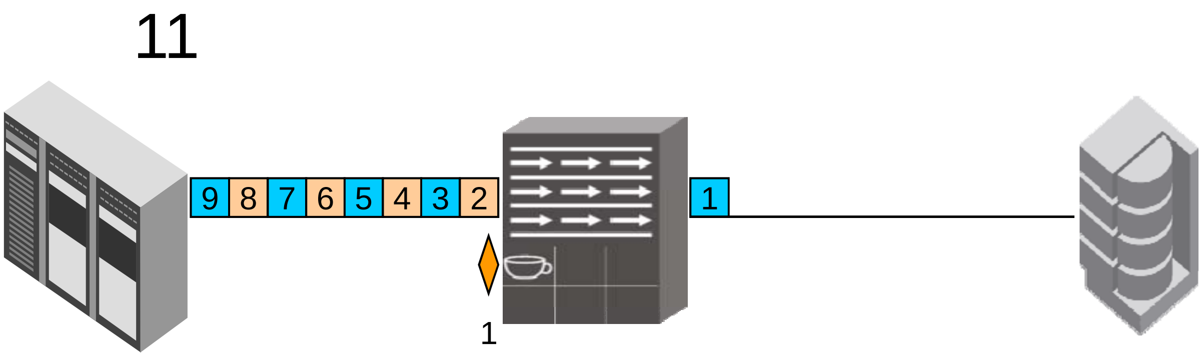
14



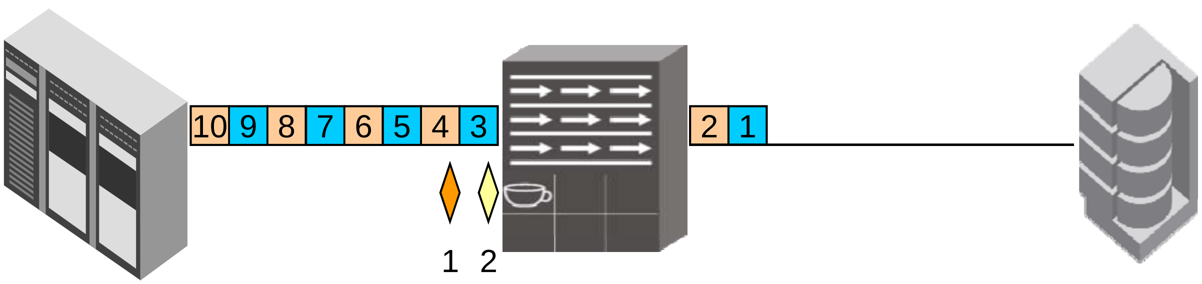
13



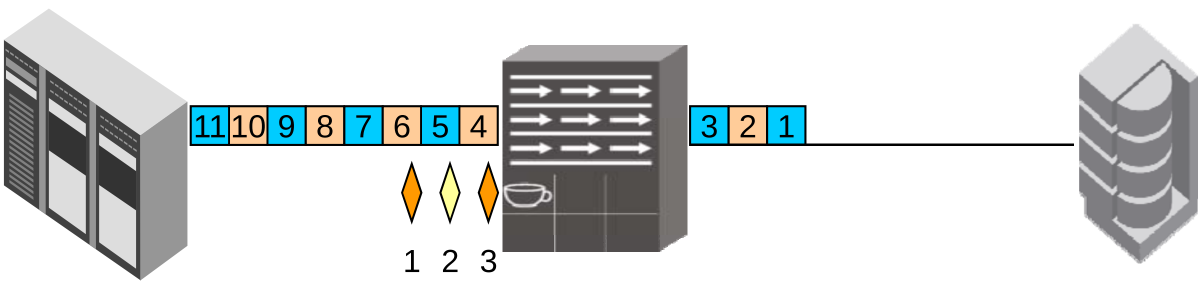

12
8
7
6
5
4
3
2
1


11
9
8
7
6
5
4
3
2
1
1

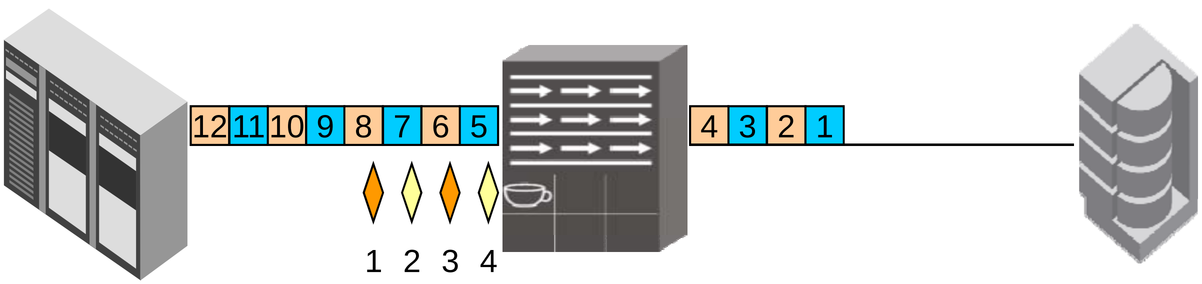
10

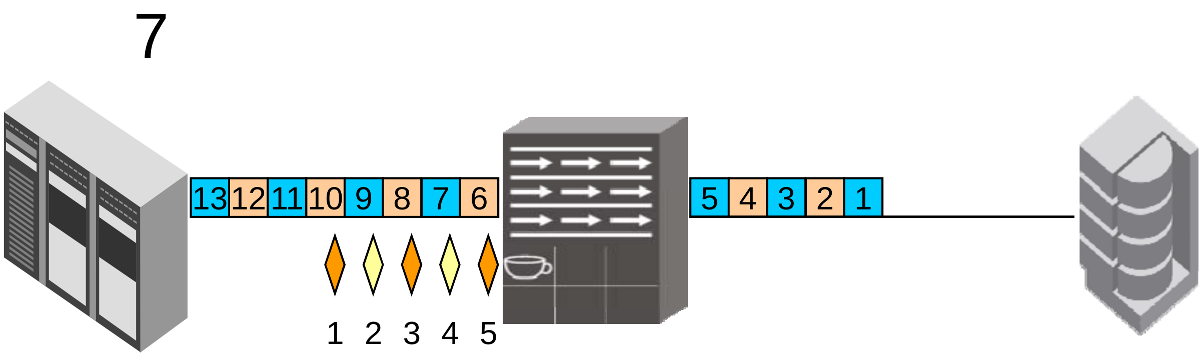


9

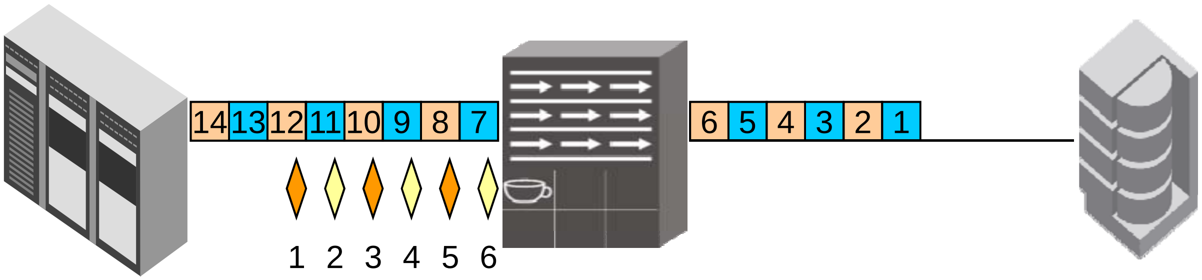


8

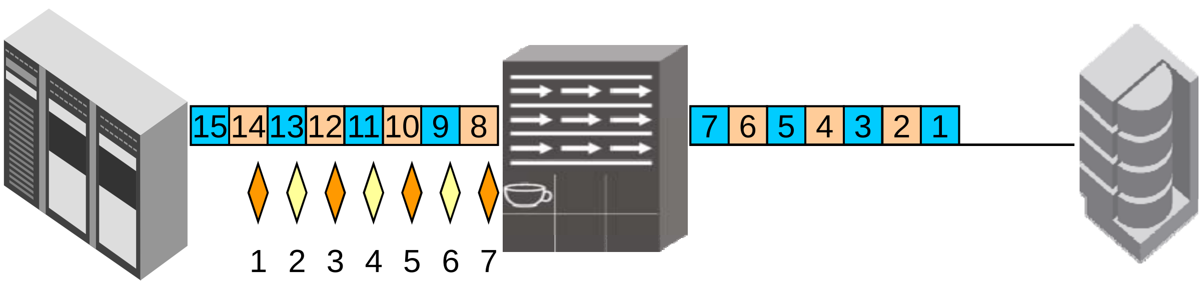



7
13 12 11 10 9 8 7 6
5 4 3 2 1
1 2 3 4 5

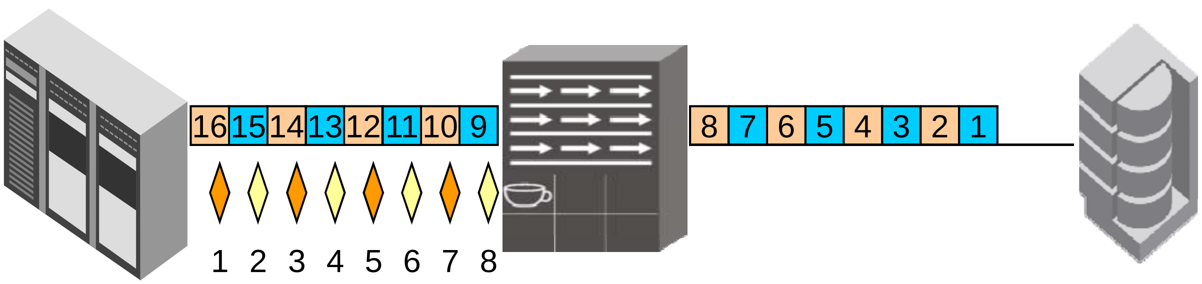
6



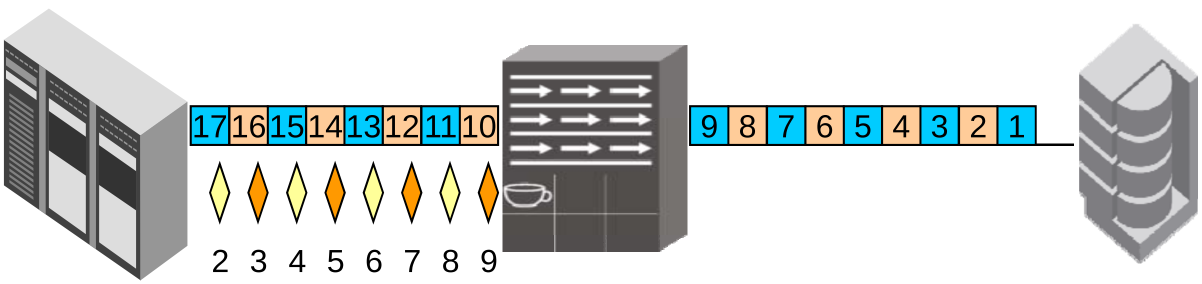
5



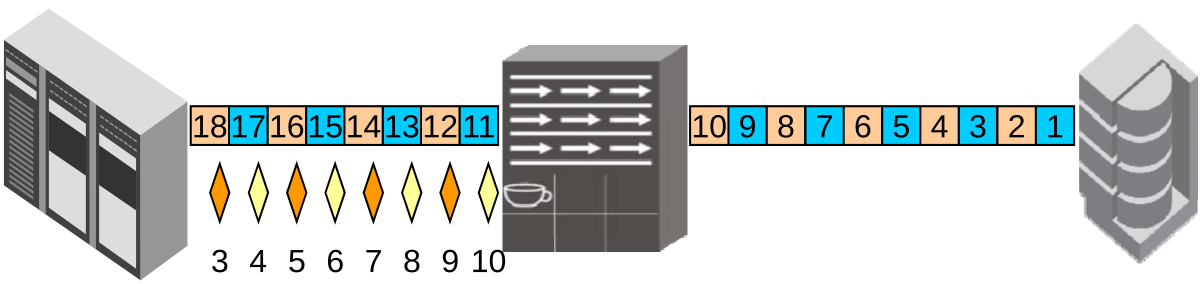
4



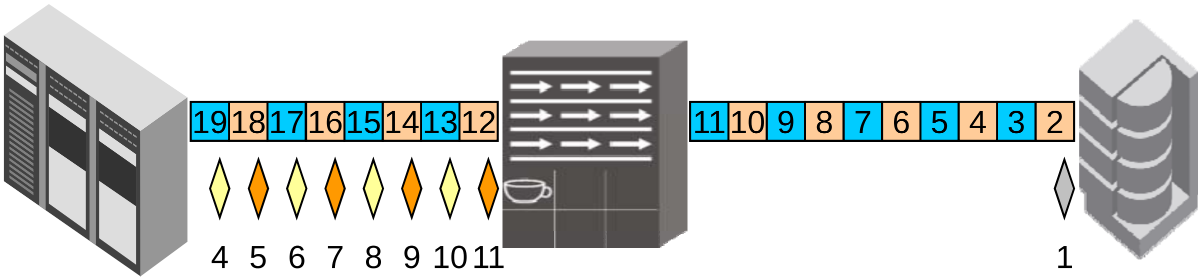
4



4



4

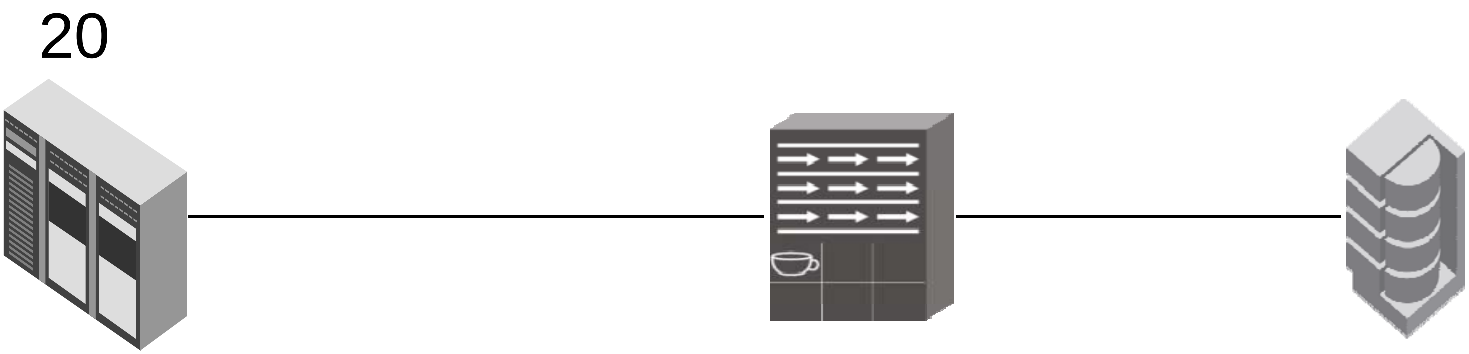


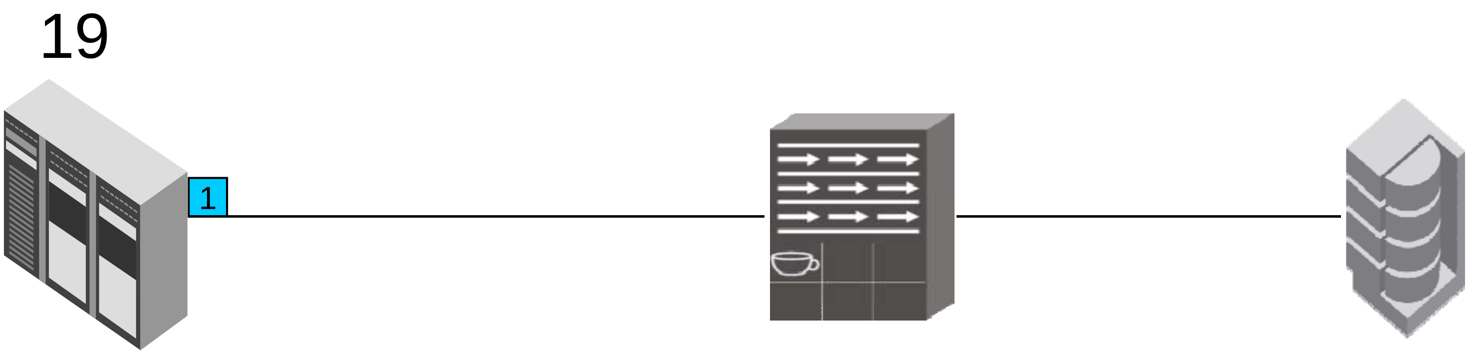
THIS PAGE INTENTIONALLY LEFT BLANK

Example: A not so full pipe

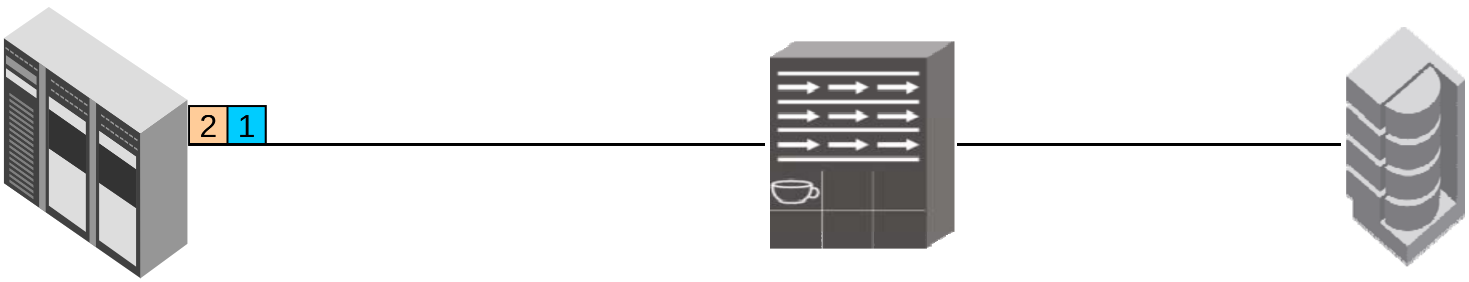
BUFFER CREDITS

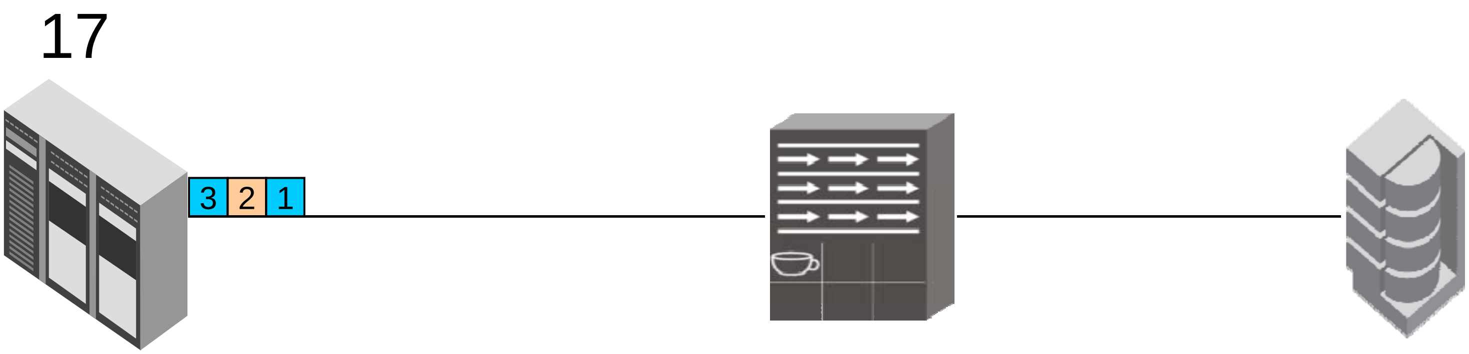
Suppose the switch is too far away from the channel for the B-B credit it advertised to the channel



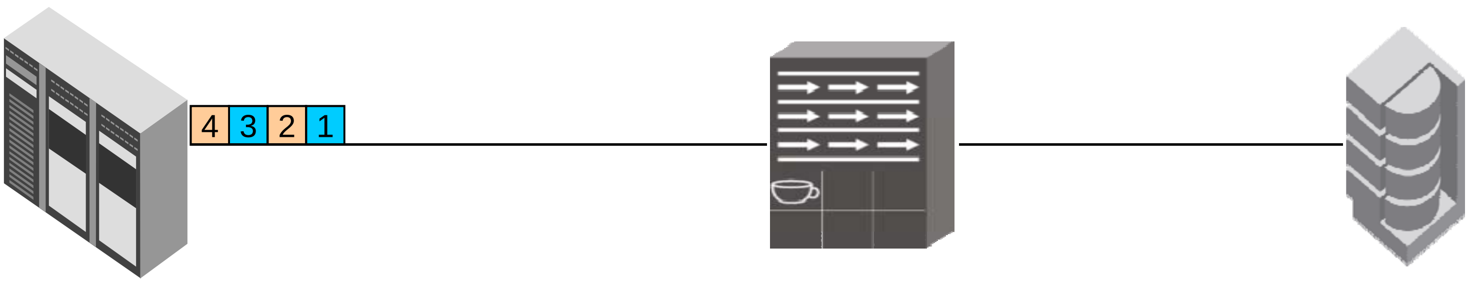

19
1

18

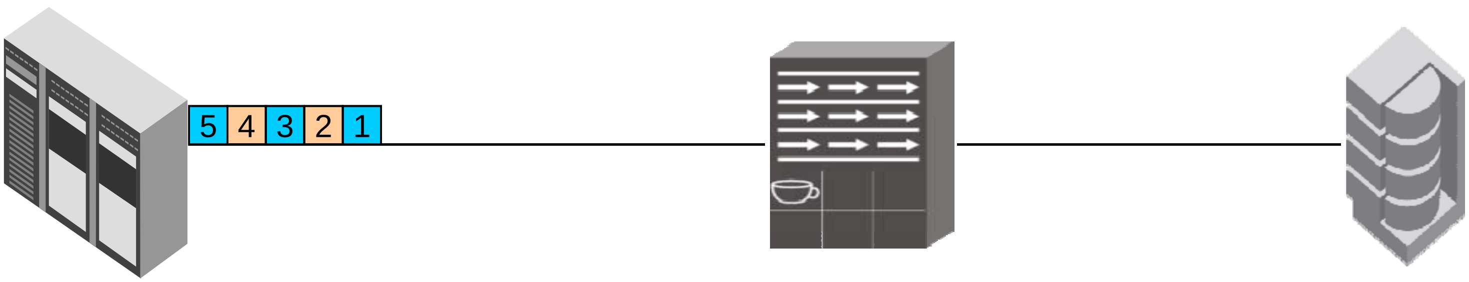


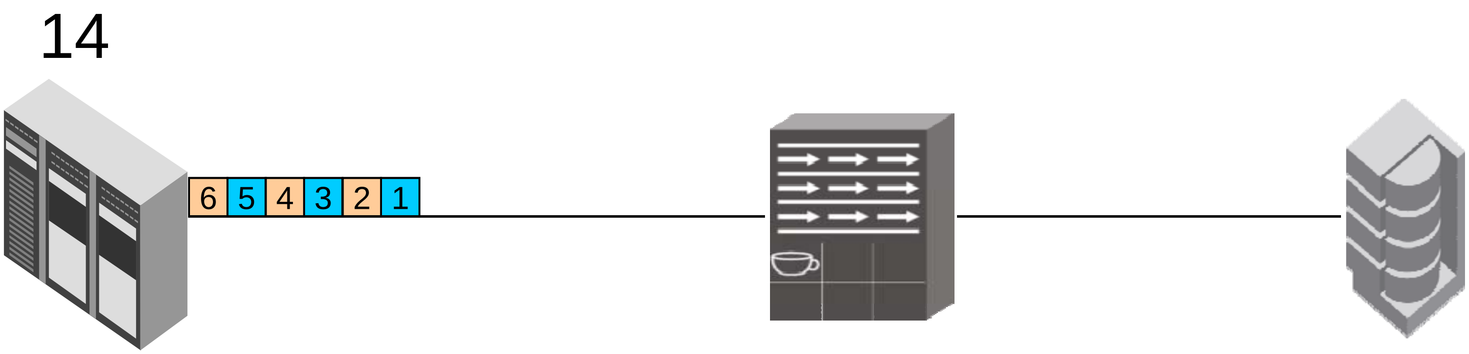

17
3
2
1

16

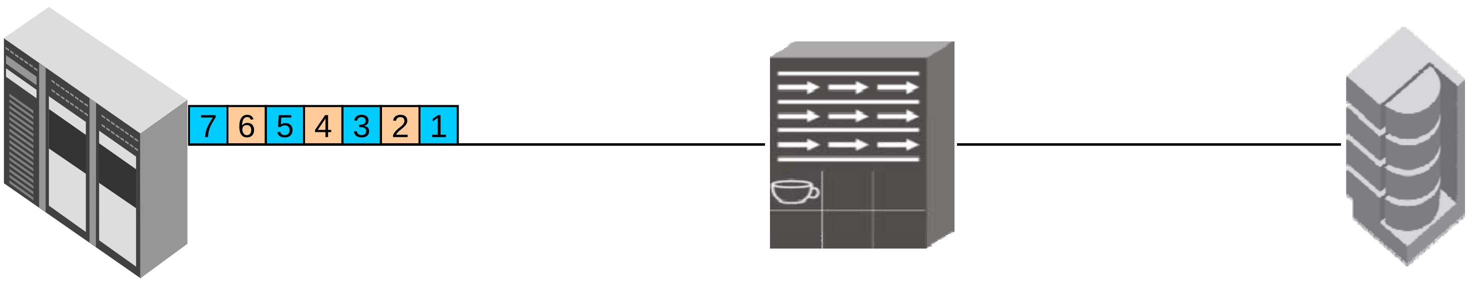


15

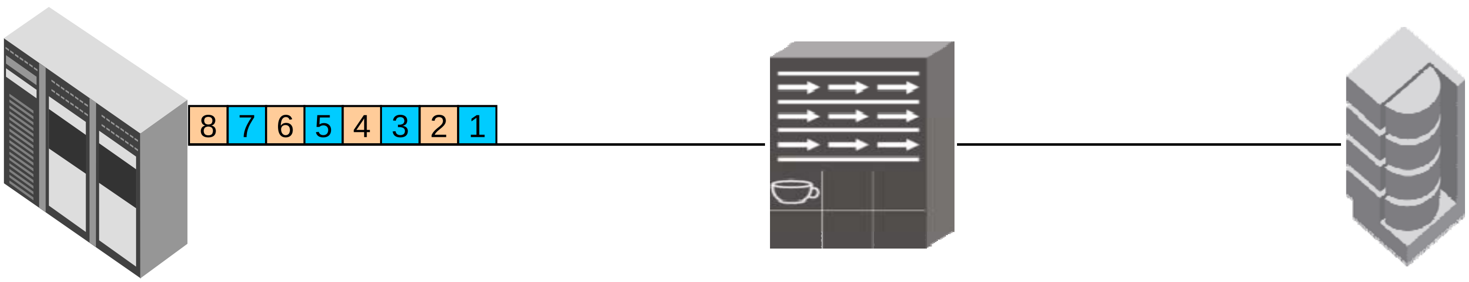


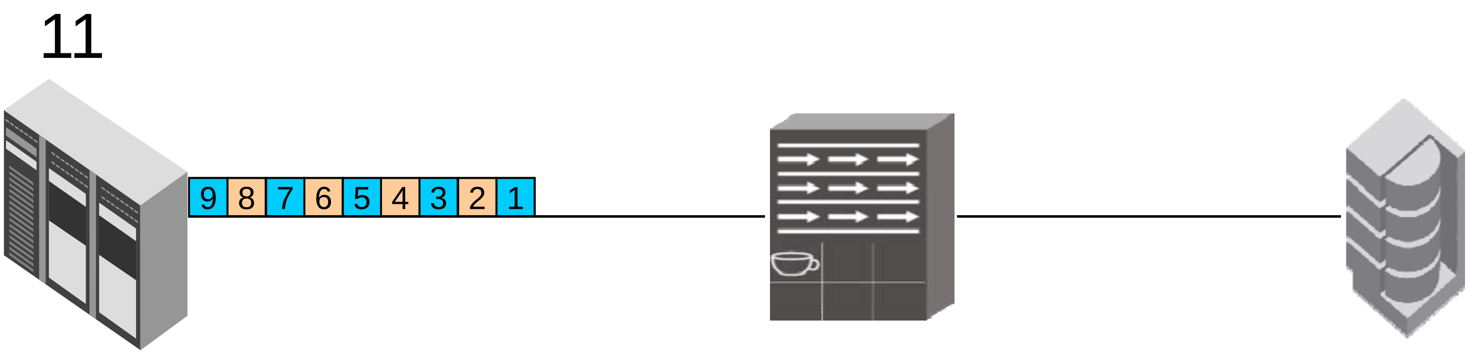

14
6 5 4 3 2 1

13

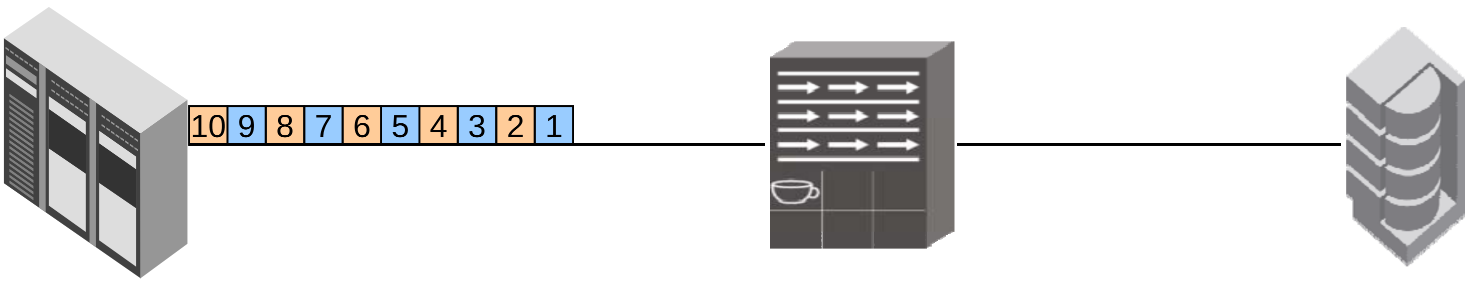


12

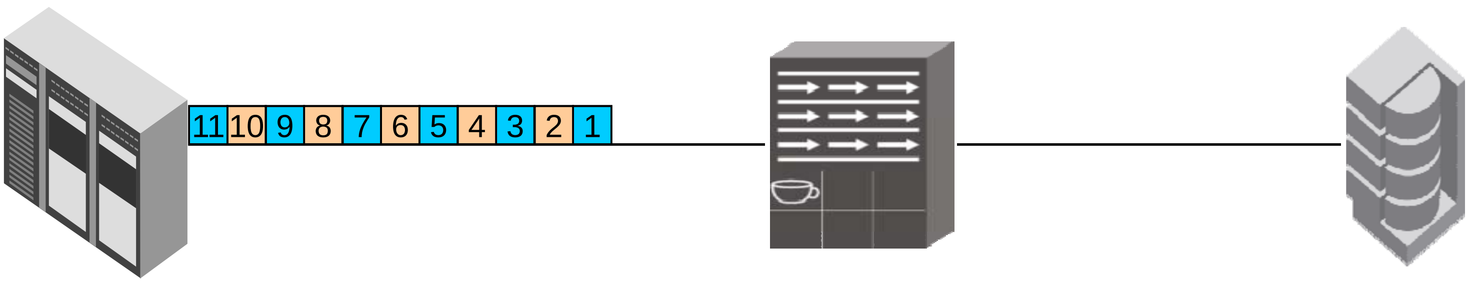



11
9
8
7
6
5
4
3
2
1

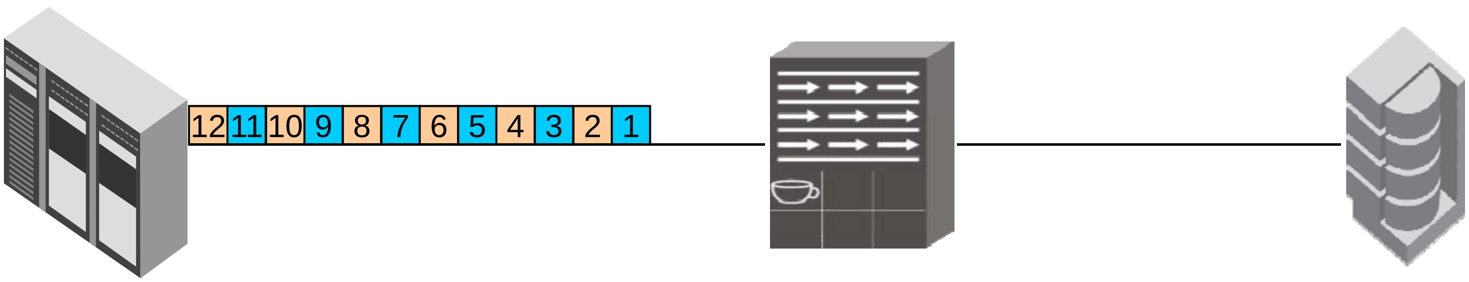
10



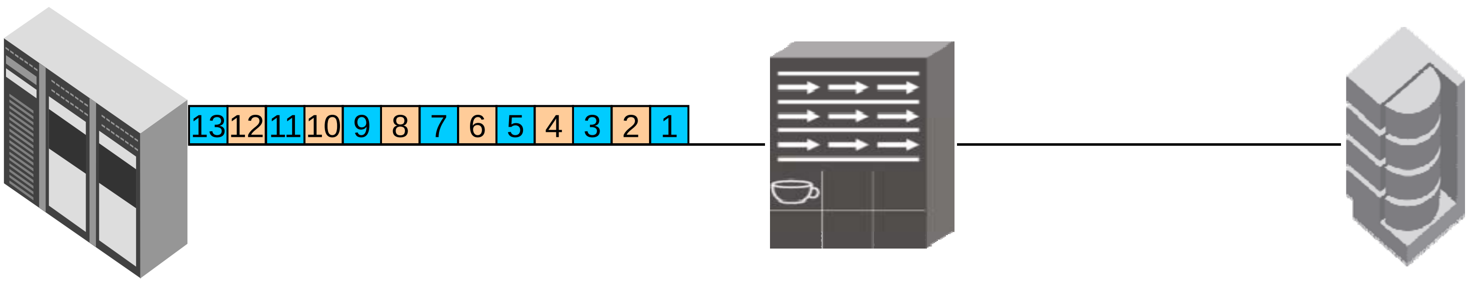
9



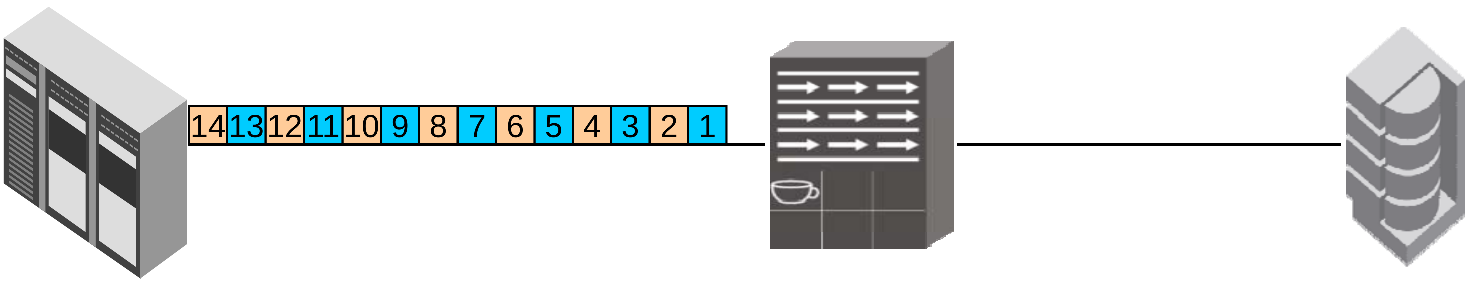
8



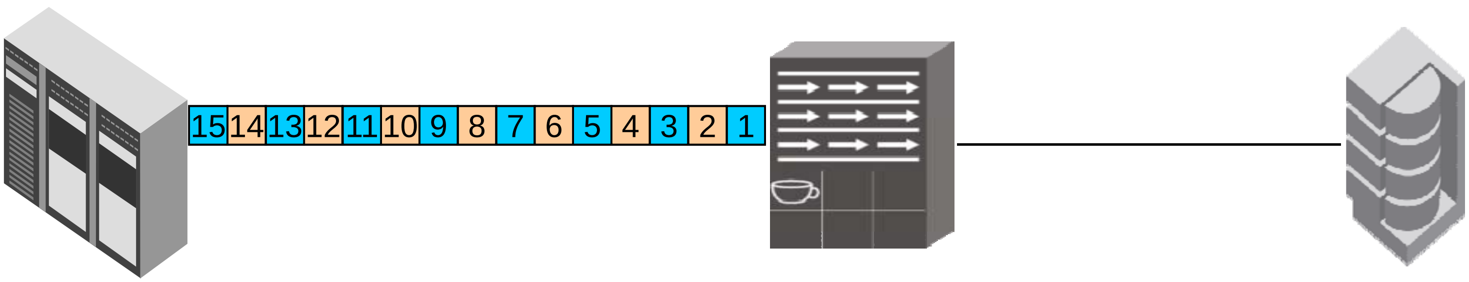
7

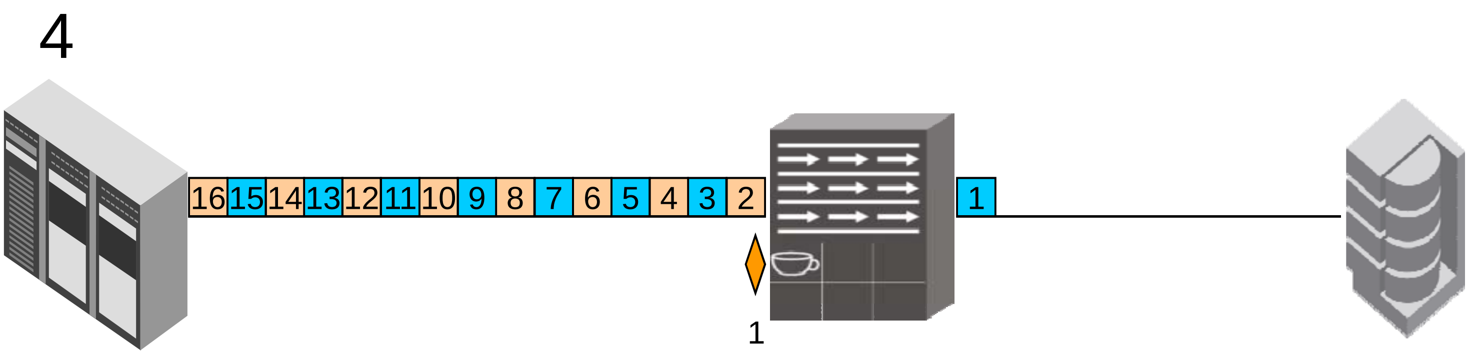


6



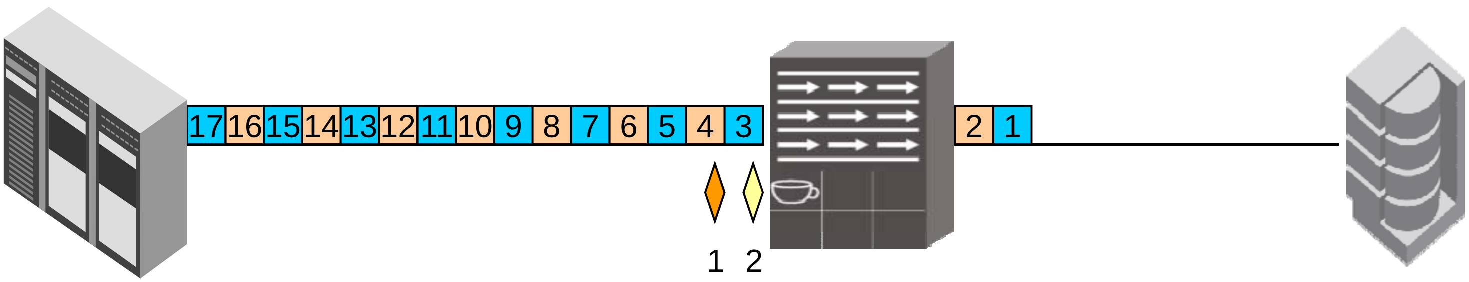
5

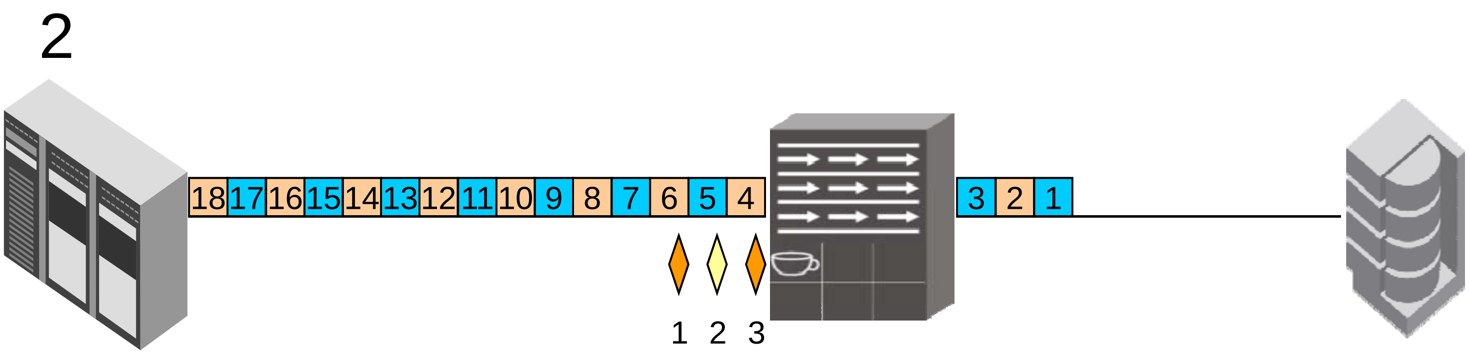




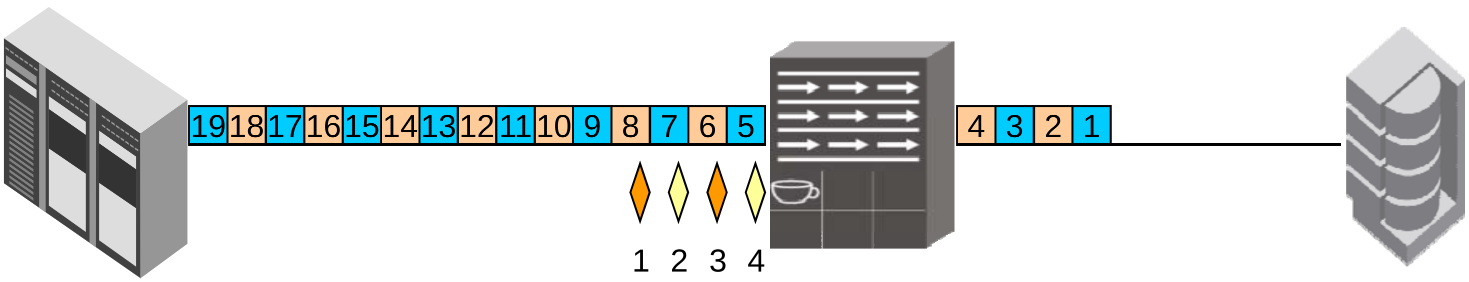
4
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
1

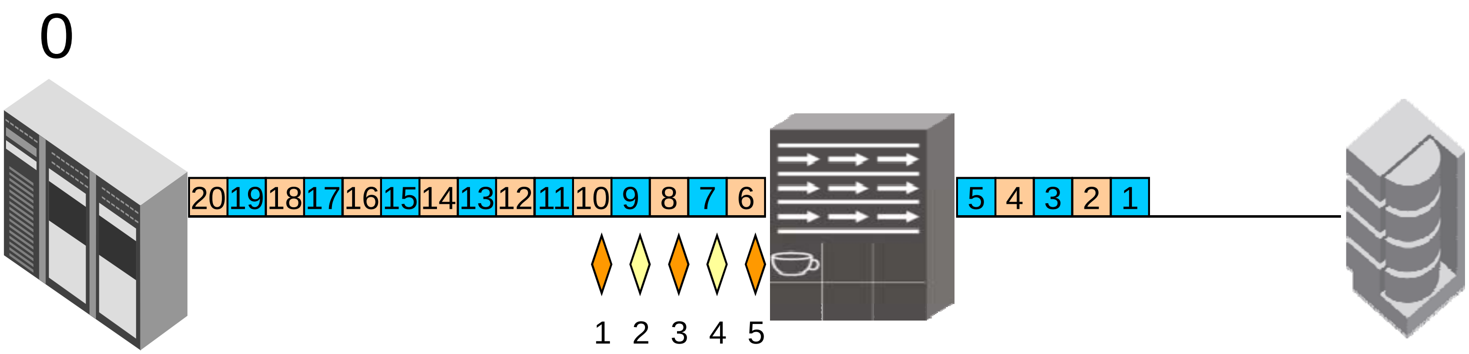
3



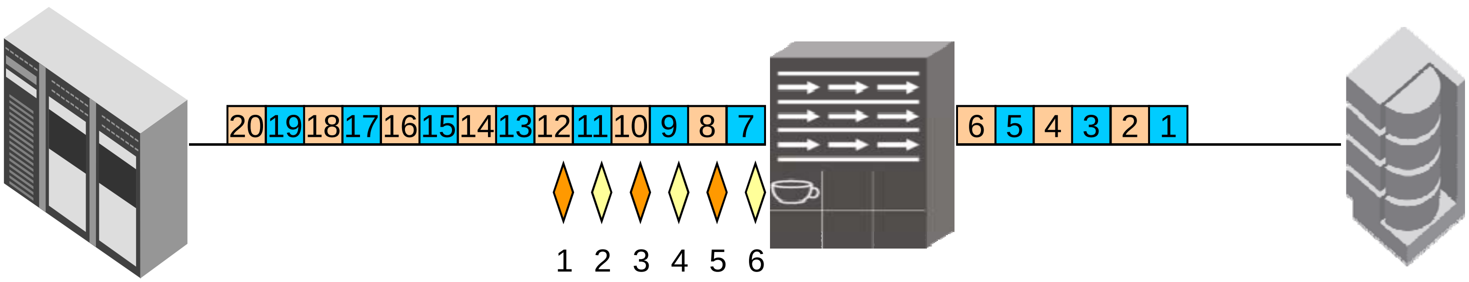

2
18 17 16 15 14 13 12 11 10 9 8 7 6 5 4
1 2 3
3 2 1

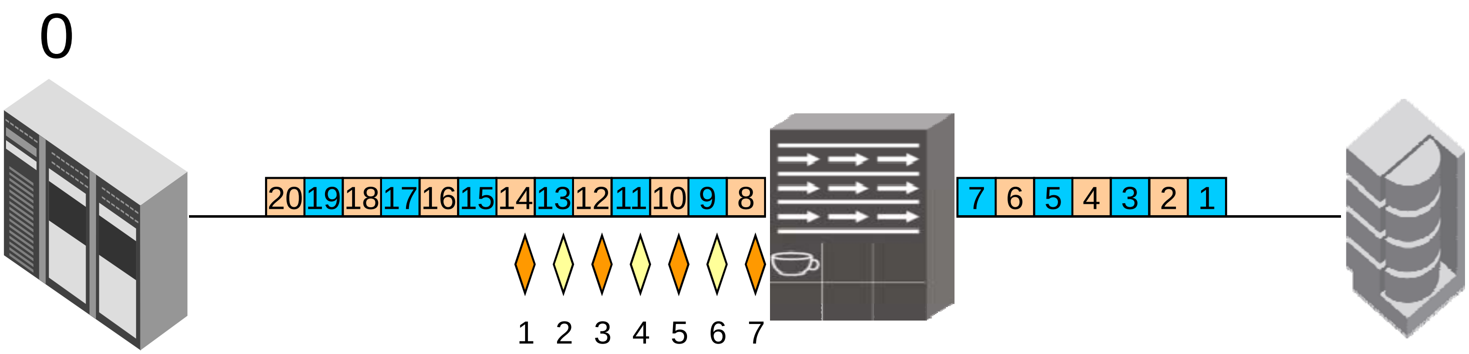
1

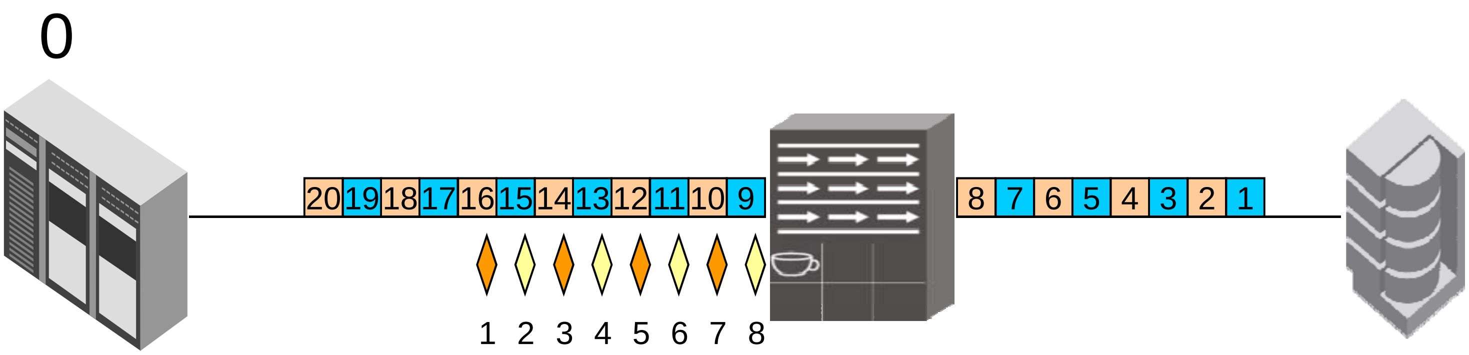


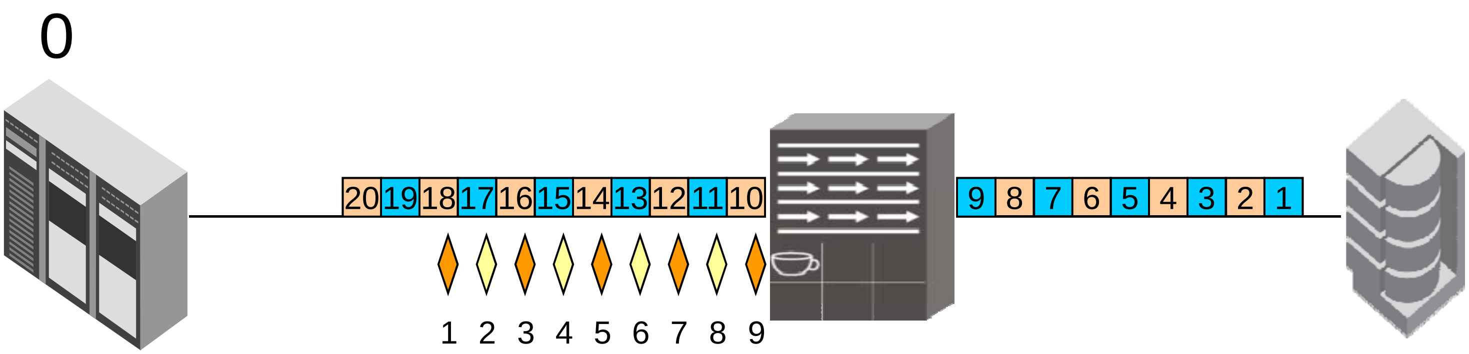

0
20 19 18 17 16 15 14 13 12 11 10 9 8 7 6
5 4 3 2 1
1 2 3 4 5

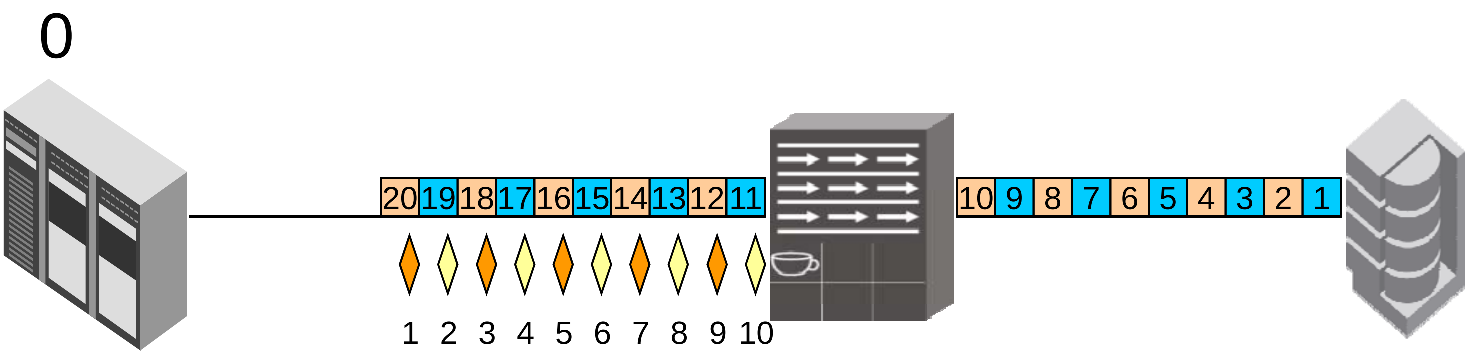
0

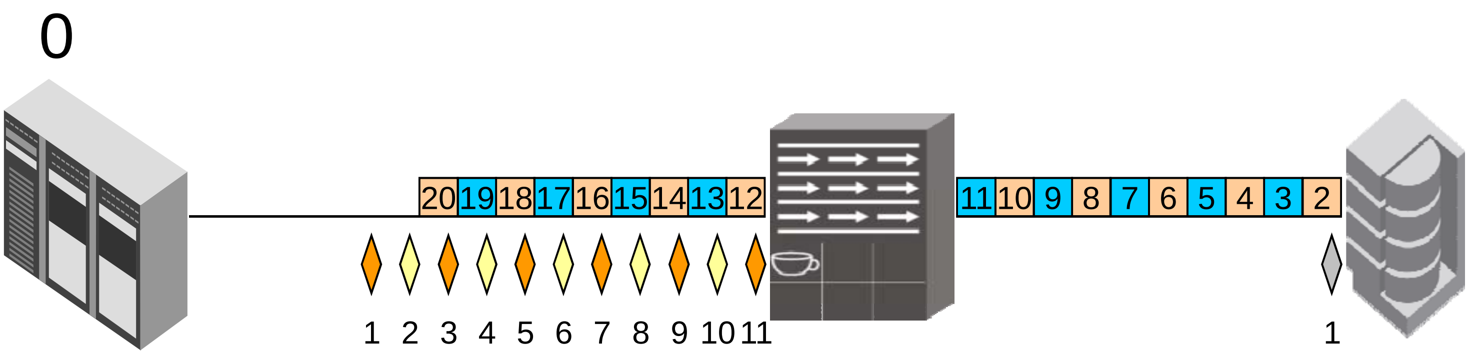



0
20 19 18 17 16 15 14 13 12 11 10 9 8
7 6 5 4 3 2 1
1 2 3 4 5 6 7

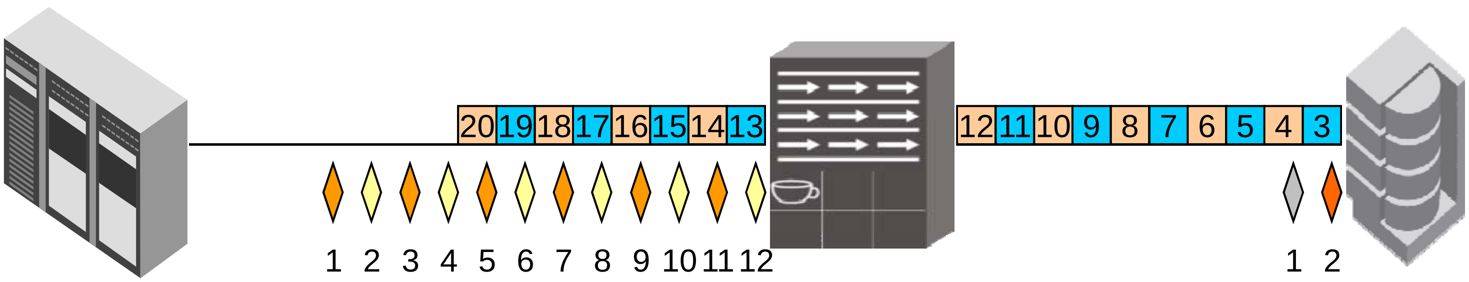

0
20 19 18 17 16 15 14 13 12 11 10 9
8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8


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

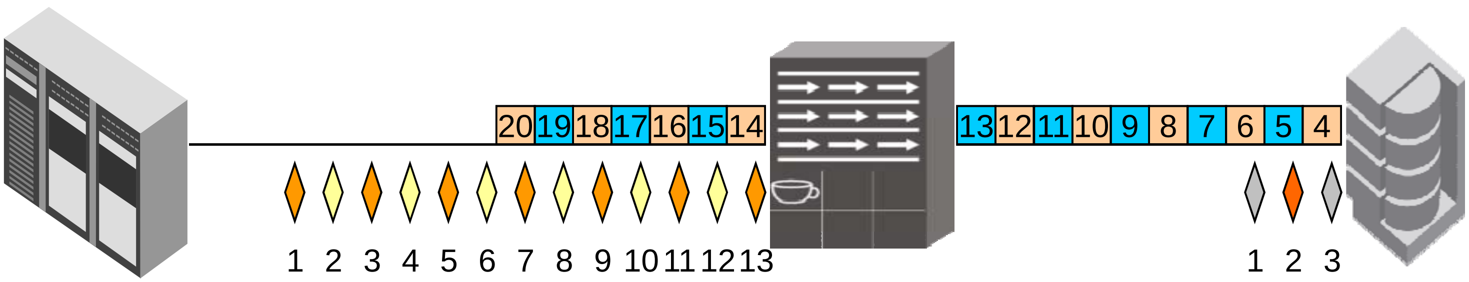

0
20 19 18 17 16 15 14 13 12 11
10 9 8 7 6 5 4 3 2 1
1 2 3 4 5 6 7 8 9 10


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

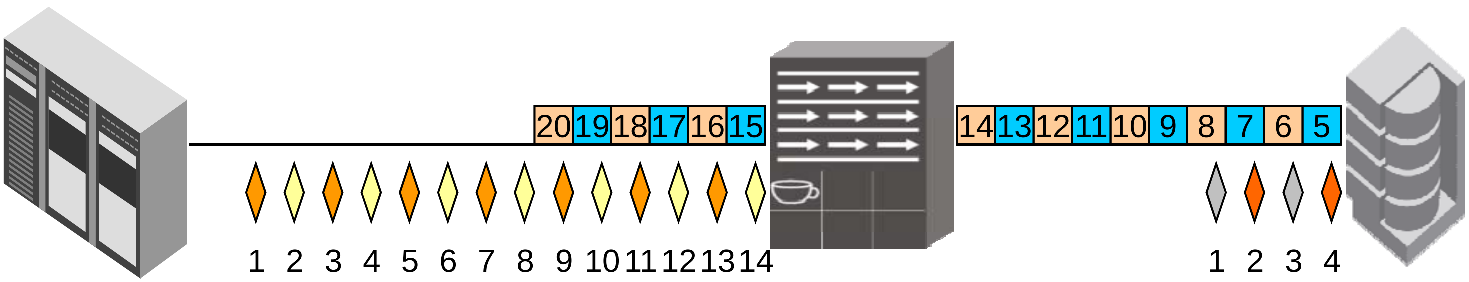
0



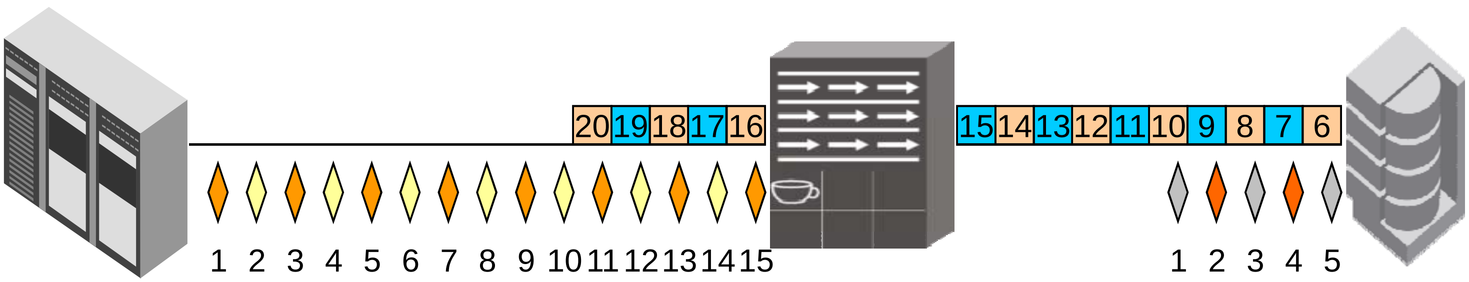
0



0

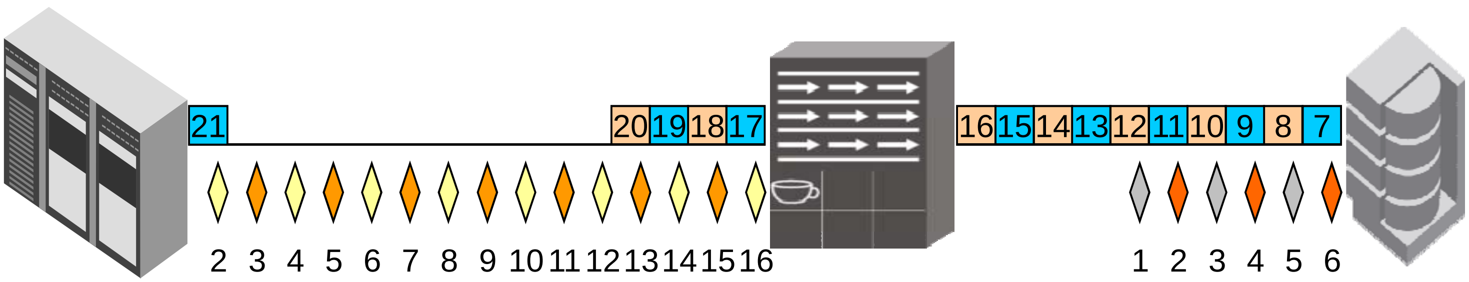

20 19 18 17 16 15
1 2 3 4 5 6 7 8 9 10 11 12 13 14
14 13 12 11 10 9 8 7 6 5
1 2 3 4

0

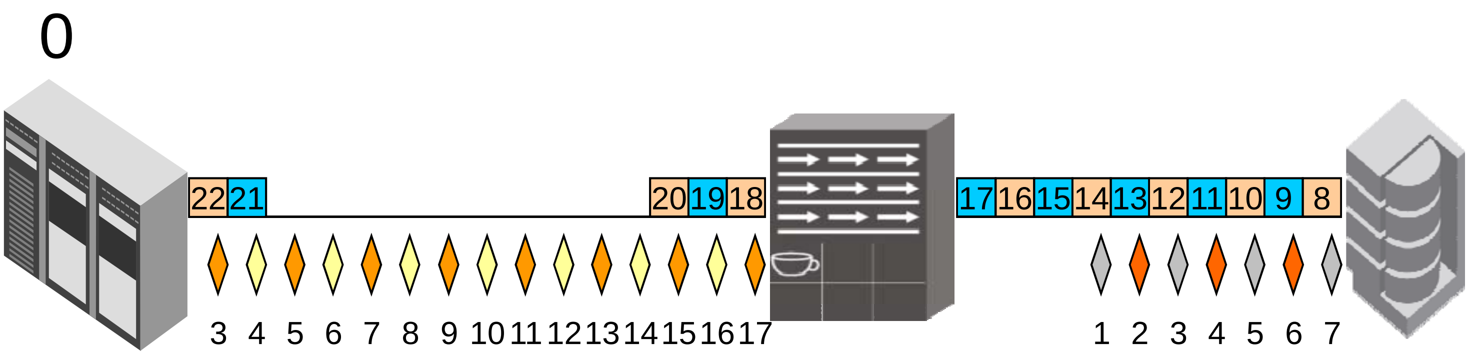


20 19 18 17 16
15 14 13 12 11 10 9 8 7 6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
1 2 3 4 5

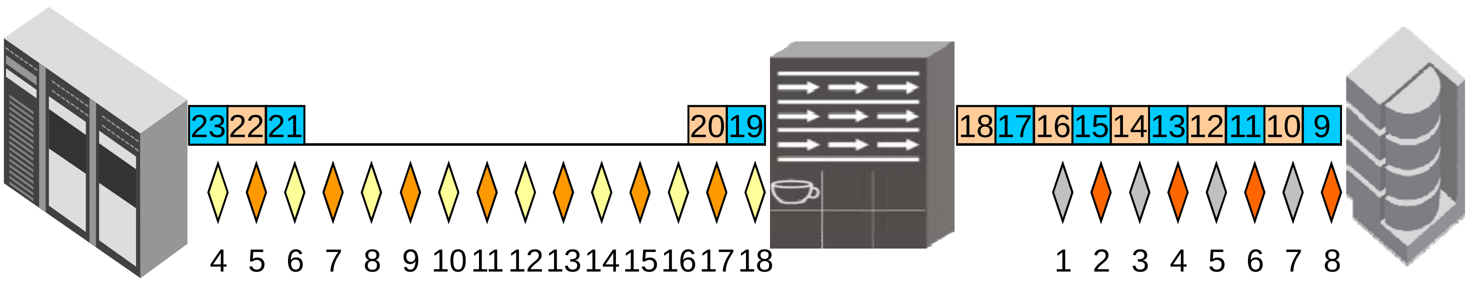
0

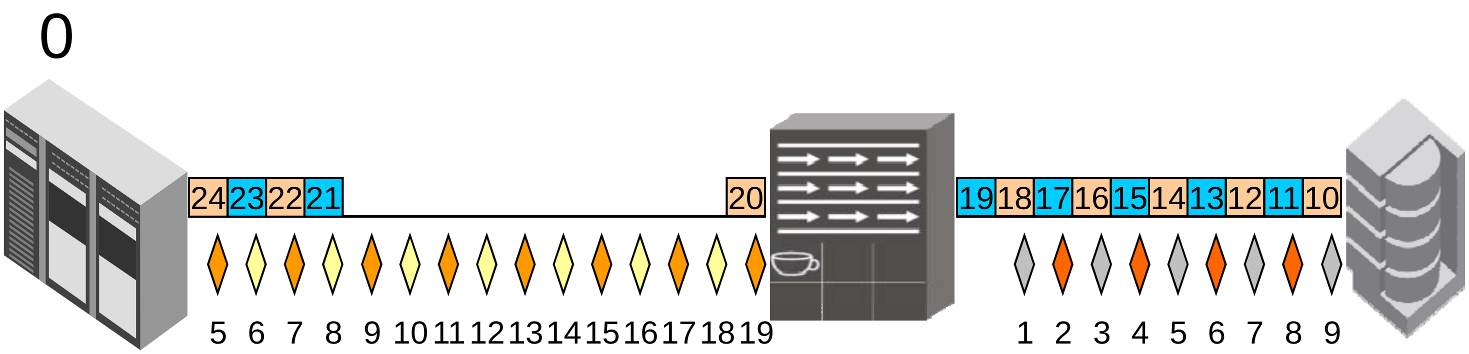


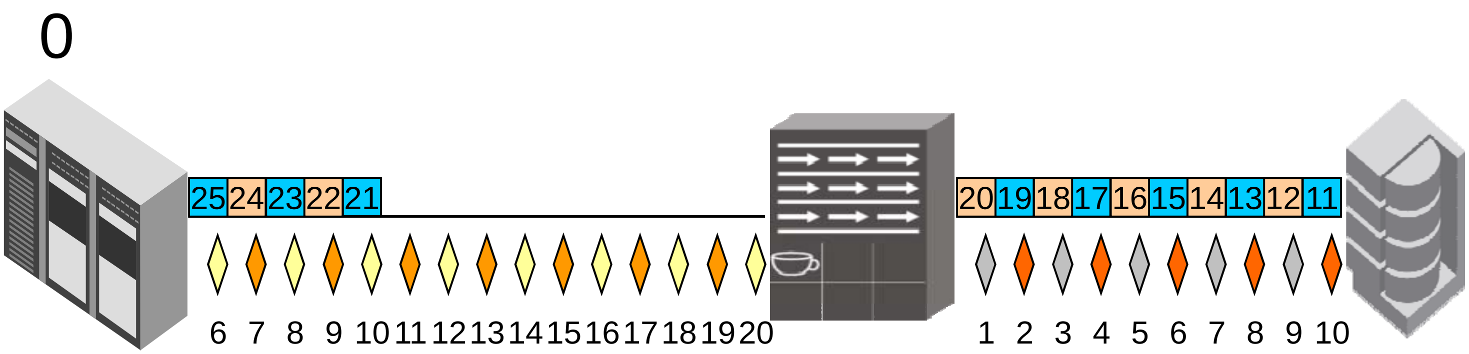
21
20
19
18
17
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
16
15
14
13
12
11
10
9
8
7
1 2 3 4 5 6


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

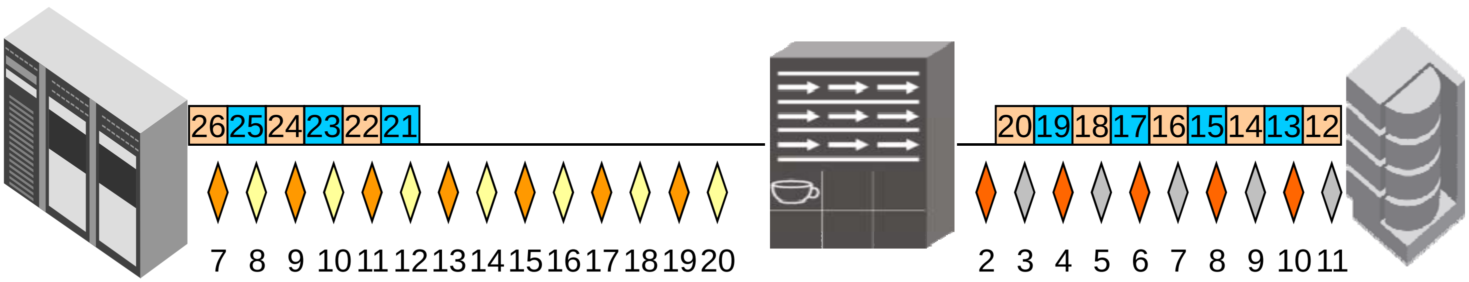
0


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

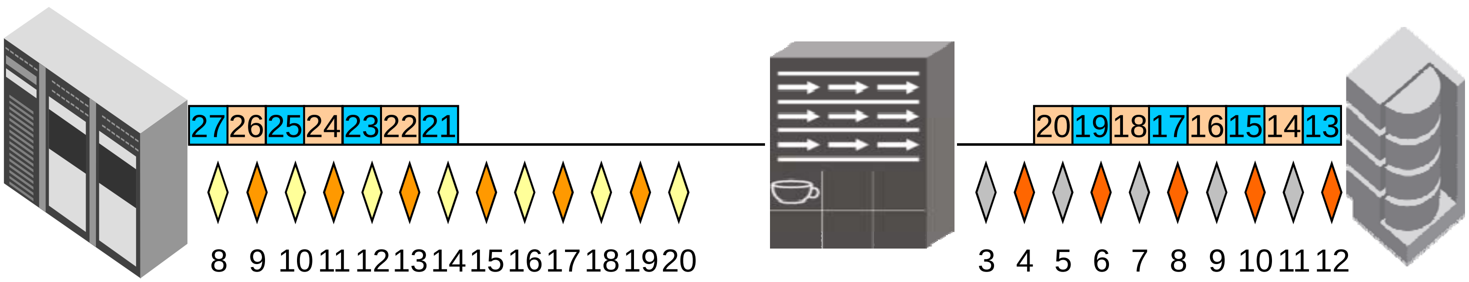

0
25 24 23 22 21
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
20 19 18 17 16 15 14 13 12 11
1 2 3 4 5 6 7 8 9 10

0

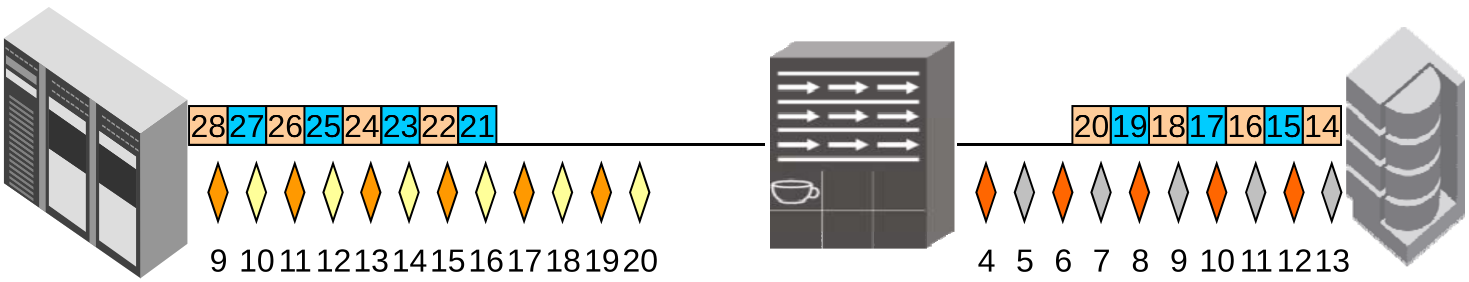


26 25 24 23 22 21
7 8 9 10 11 12 13 14 15 16 17 18 19 20
20 19 18 17 16 15 14 13 12
2 3 4 5 6 7 8 9 10 11

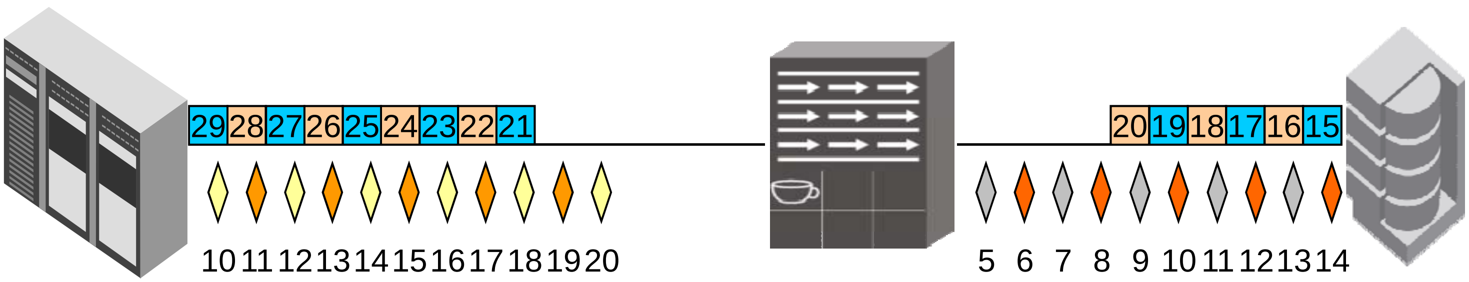
0


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

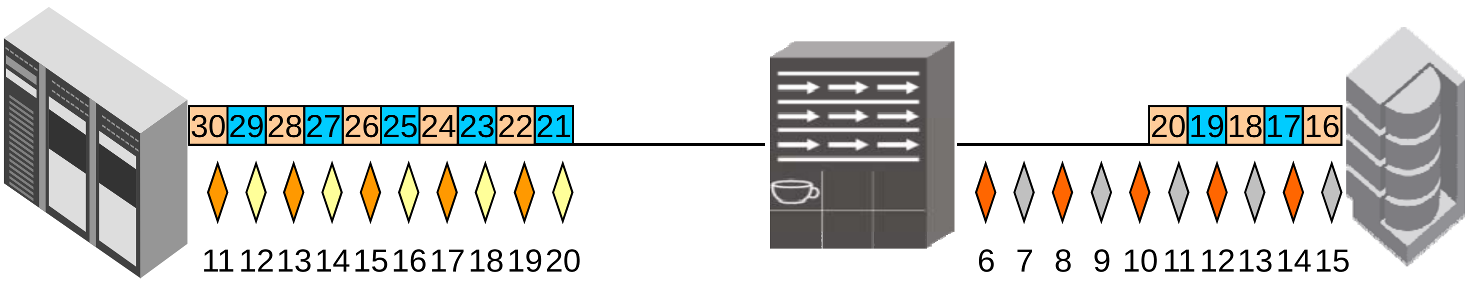
0


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

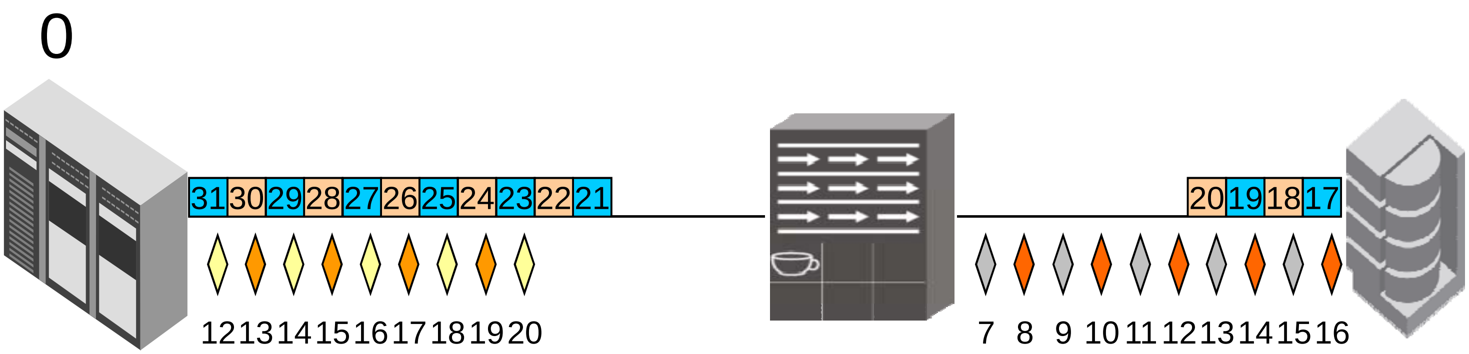
0


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

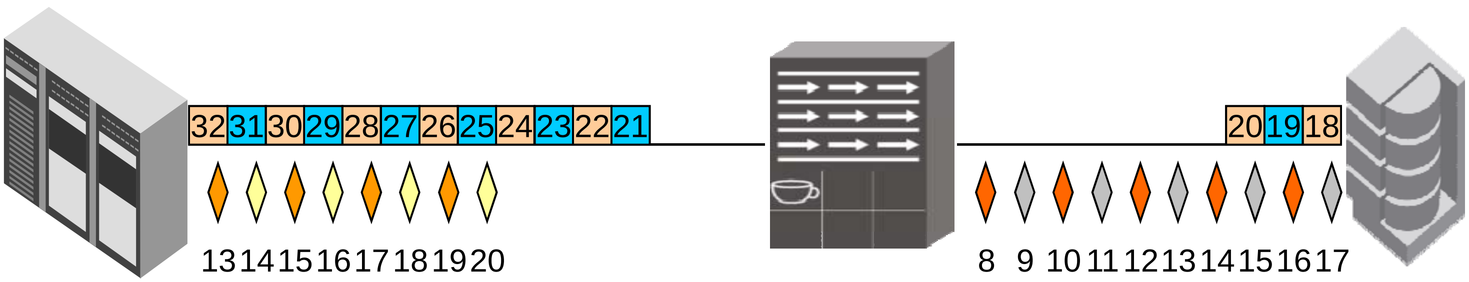
0



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

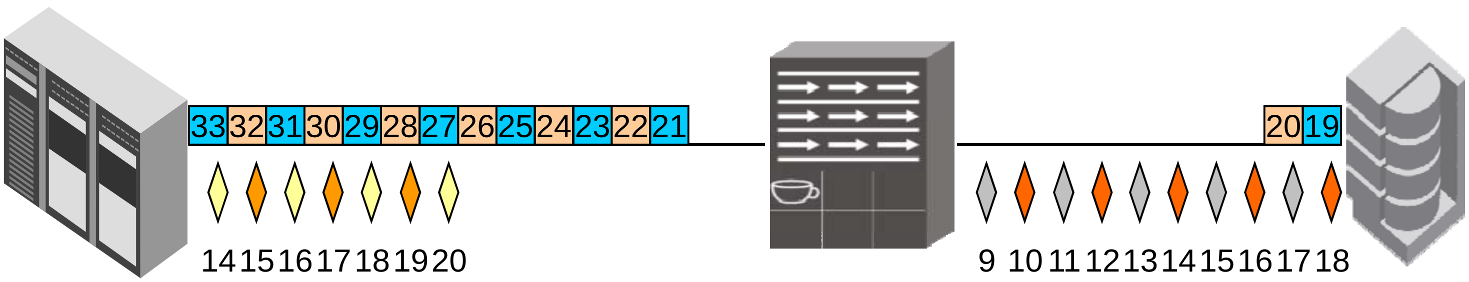

0
31 30 29 28 27 26 25 24 23 22 21
12 13 14 15 16 17 18 19 20
20 19 18 17
7 8 9 10 11 12 13 14 15 16

0

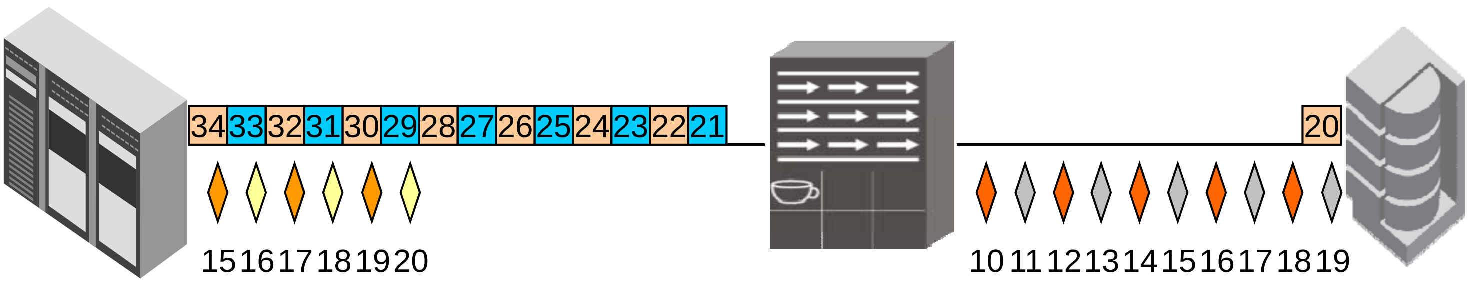


32 31 30 29 28 27 26 25 24 23 22 21
13 14 15 16 17 18 19 20
20 19 18
8 9 10 11 12 13 14 15 16 17

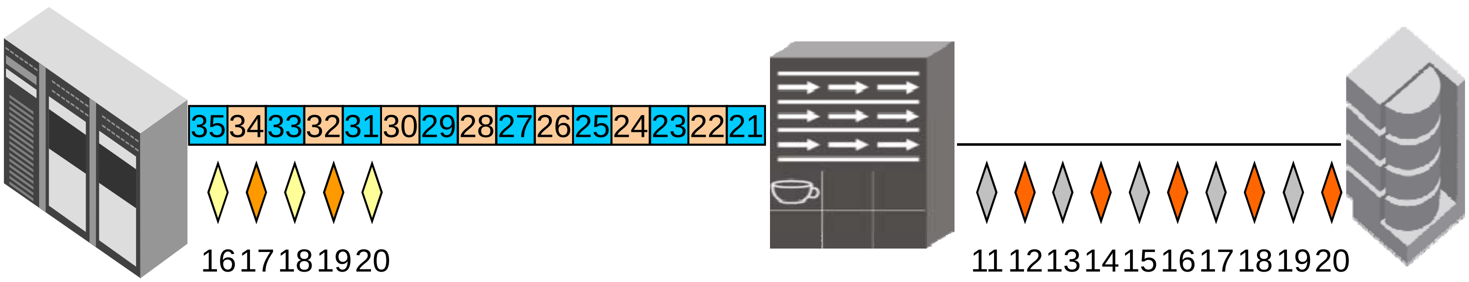
0


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

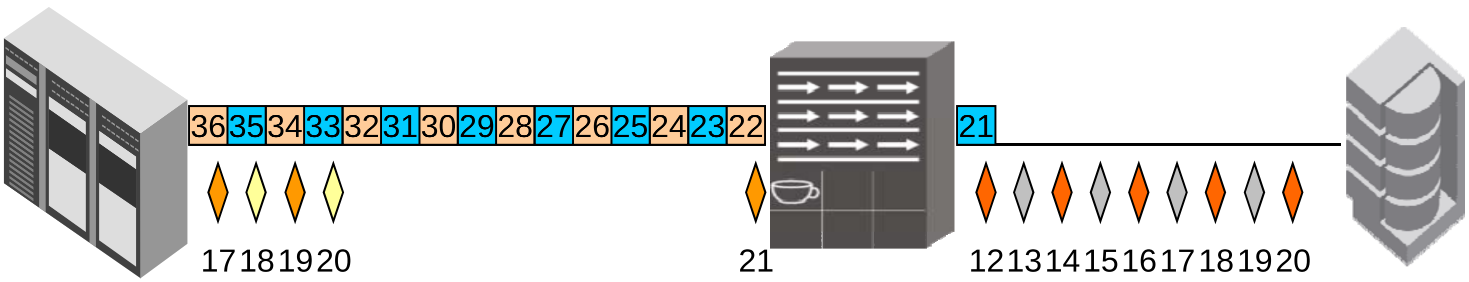

34 33 32 31 30 29 28 27 26 25 24 23 22 21
15 16 17 18 19 20
20
10 11 12 13 14 15 16 17 18 19

0



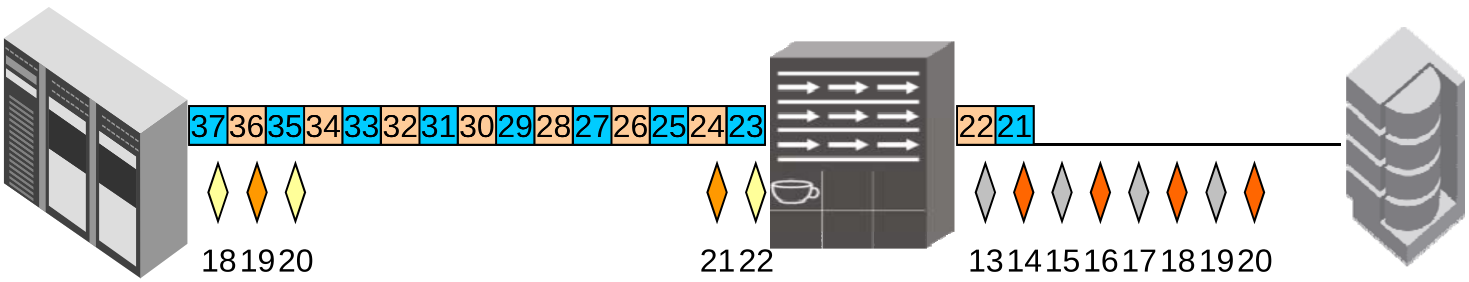
35 34 33 32 31 30 29 28 27 26 25 24 23 22 21
16 17 18 19 20
11 12 13 14 15 16 17 18 19 20

0

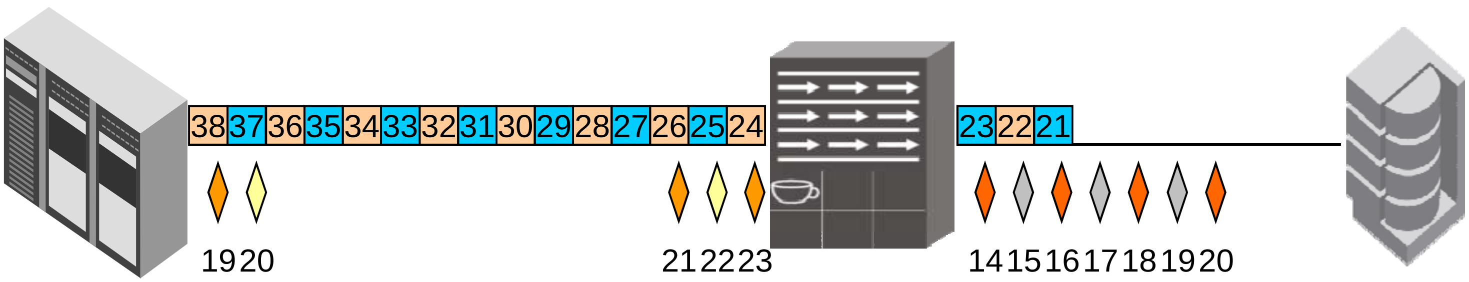


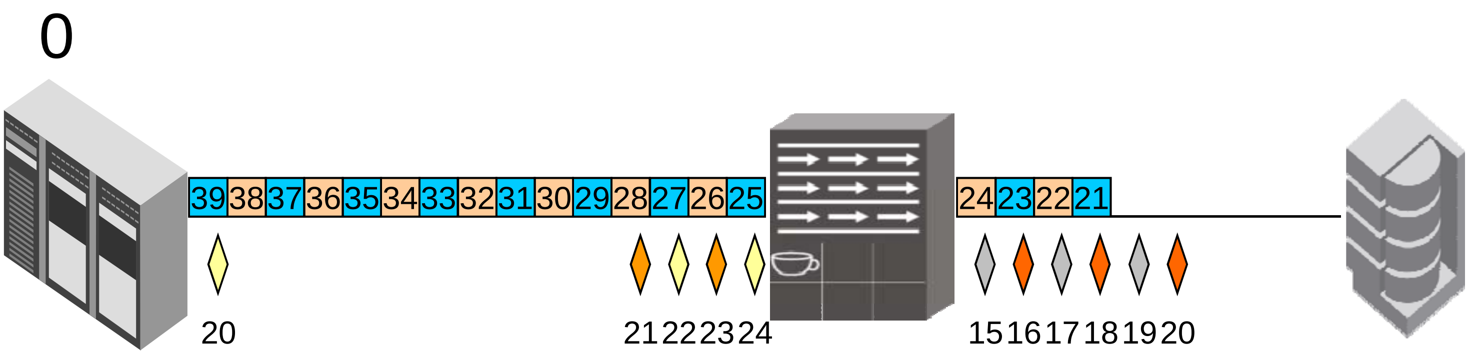
36 35 34 33 32 31 30 29 28 27 26 25 24 23 22
21
17 18 19 20
21
12 13 14 15 16 17 18 19 20

0

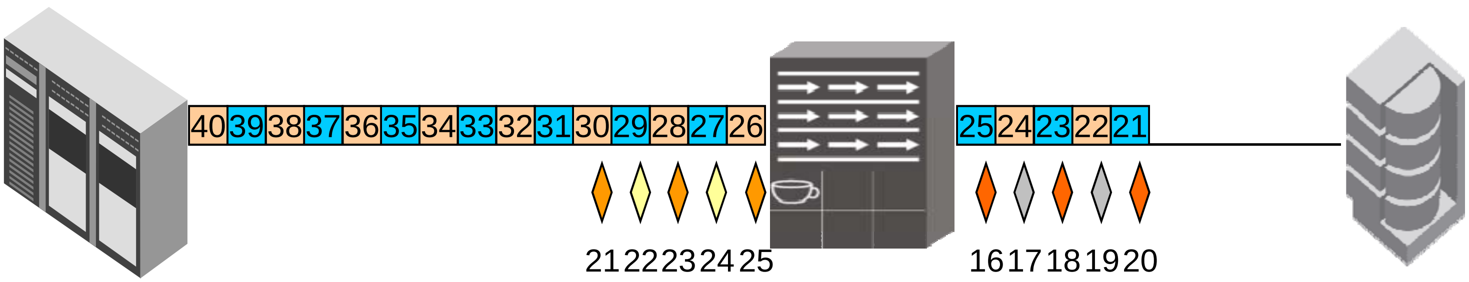

37 36 35 34 33 32 31 30 29 28 27 26 25 24 23
22 21
18 19 20
21 22
13 14 15 16 17 18 19 20

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

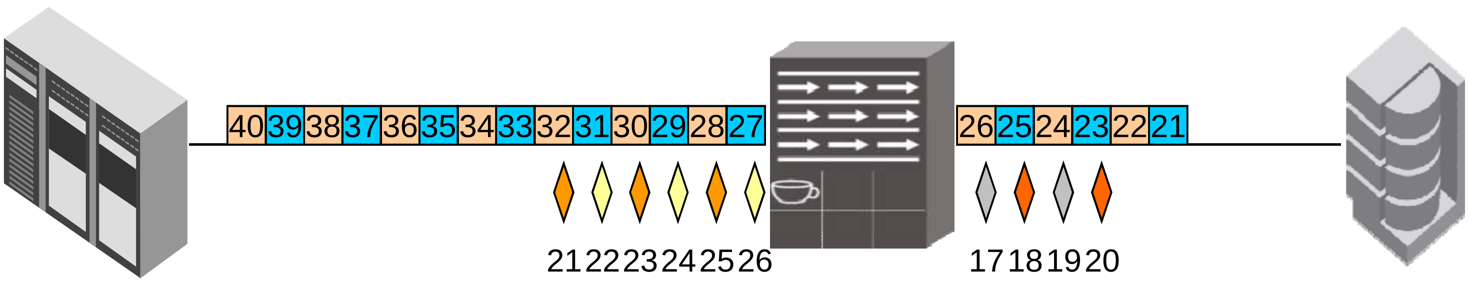

0
39 38 37 36 35 34 33 32 31 30 29 28 27 26 25
24 23 22 21
20
21 22 23 24
15 16 17 18 19 20

0

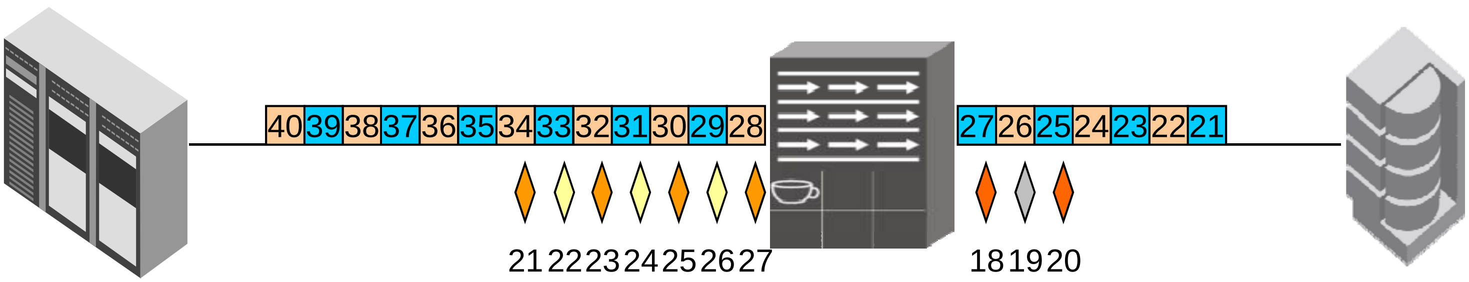


40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
21
22
23
24
25
25
24
23
22
21
16
17
18
19
20

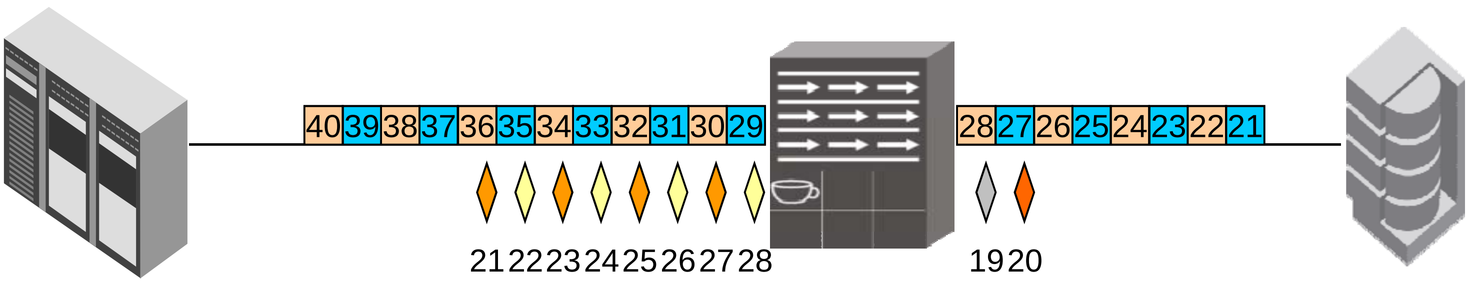
0


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

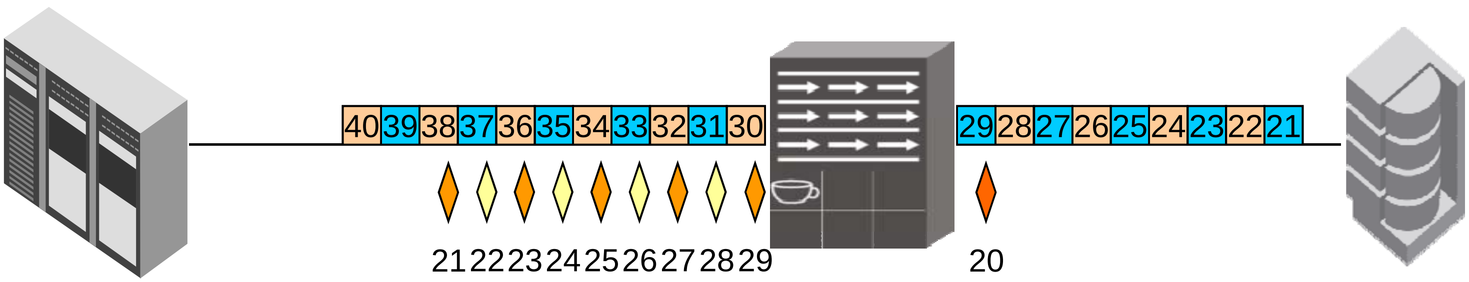
0


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

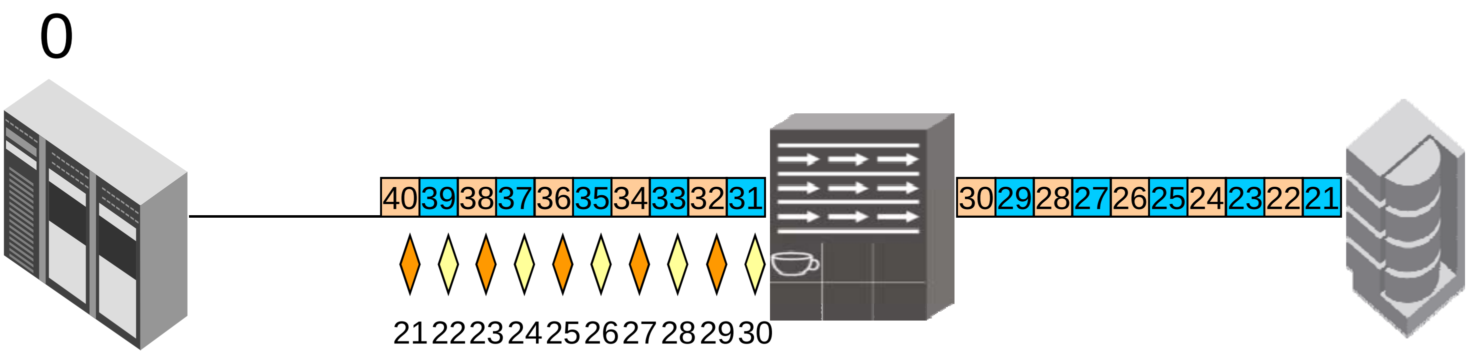
0


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

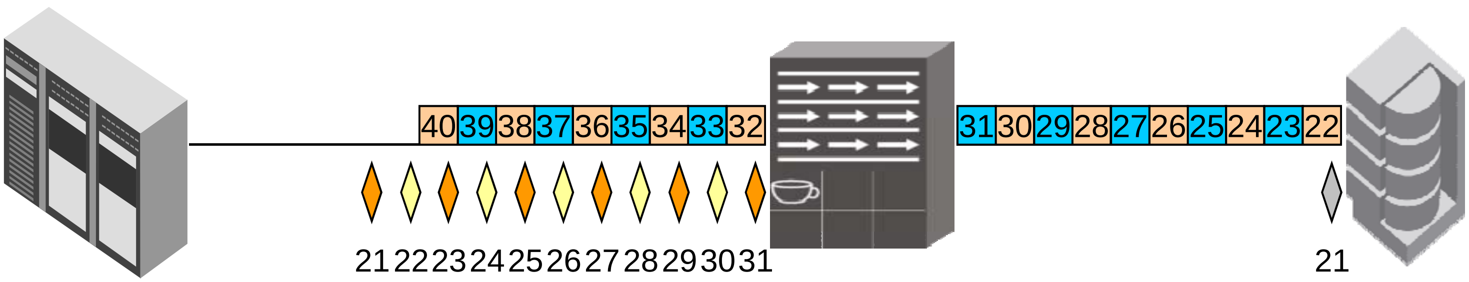
0

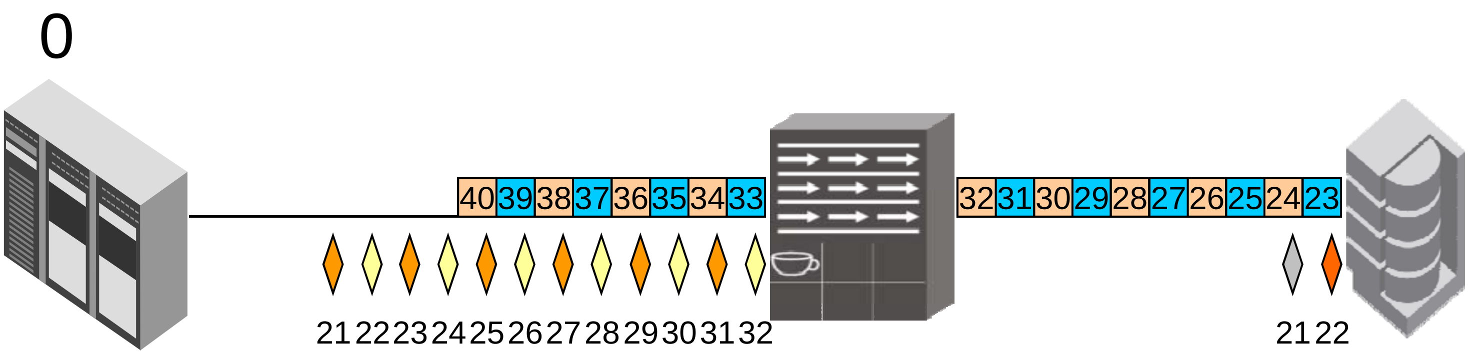


40 39 38 37 36 35 34 33 32 31 30
29 28 27 26 25 24 23 22 21
21 22 23 24 25 26 27 28 29
20


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

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

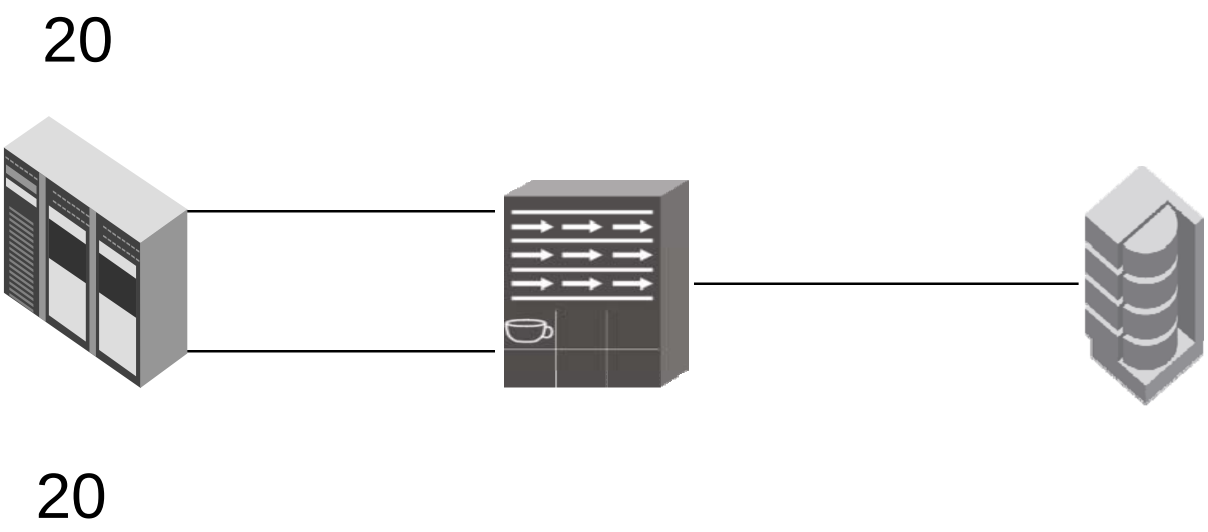
THIS PAGE INTENTIONALLY LEFT BLANK

Example: Multiple Senders and One Receiver

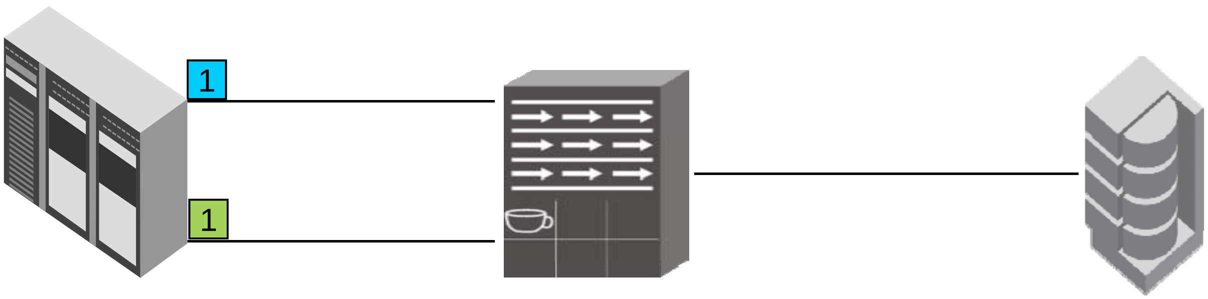
BUFFER CREDITS

Suppose there are multiple senders to one receiver

Each sender attempts to send at 100% link speed

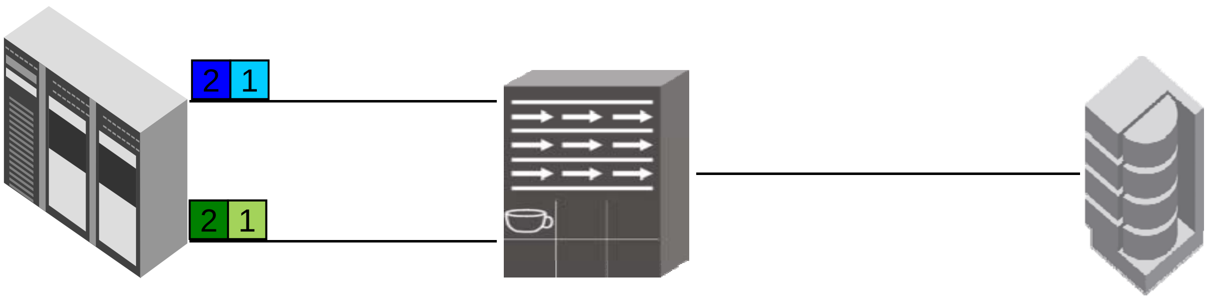


19



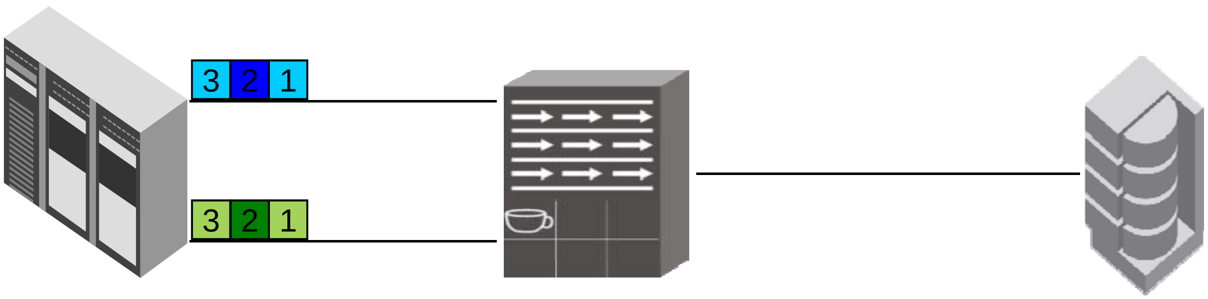
19

18



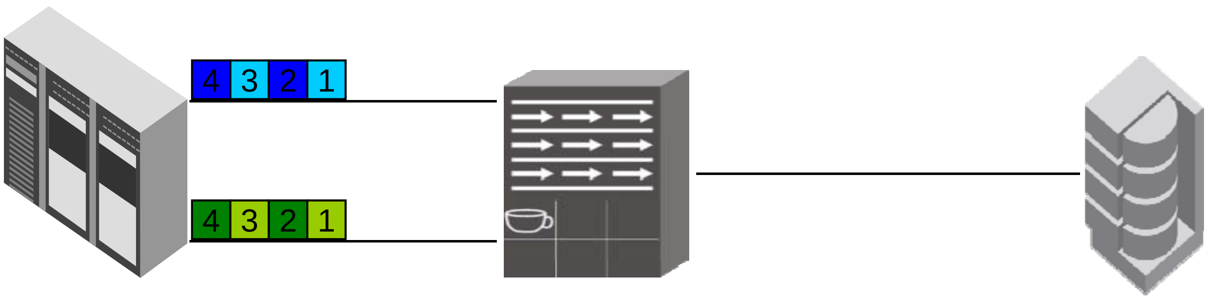
18

17



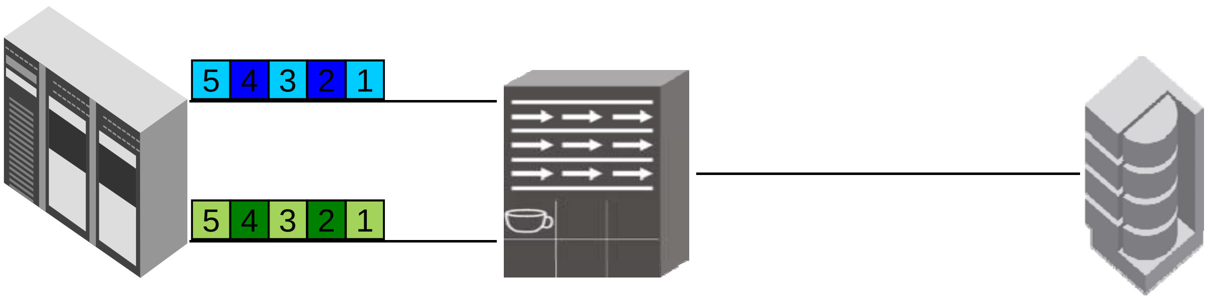
17

16



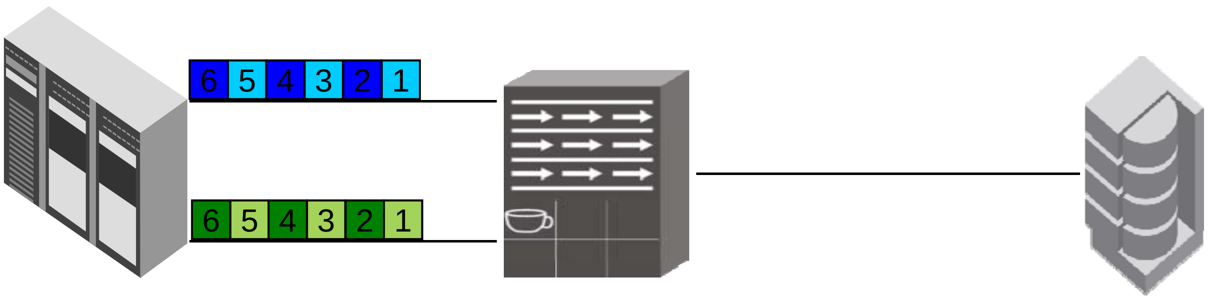
16

15



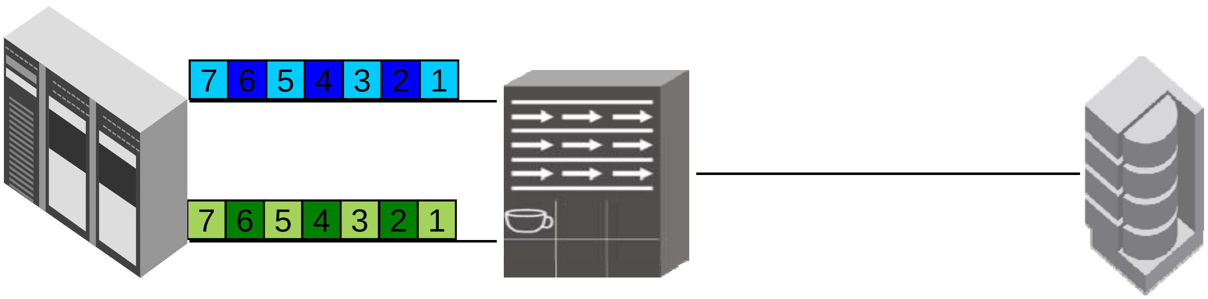
15

14



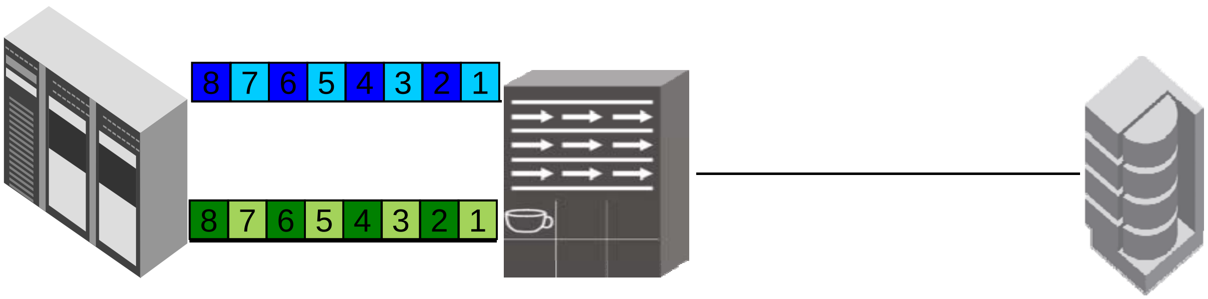
14

13

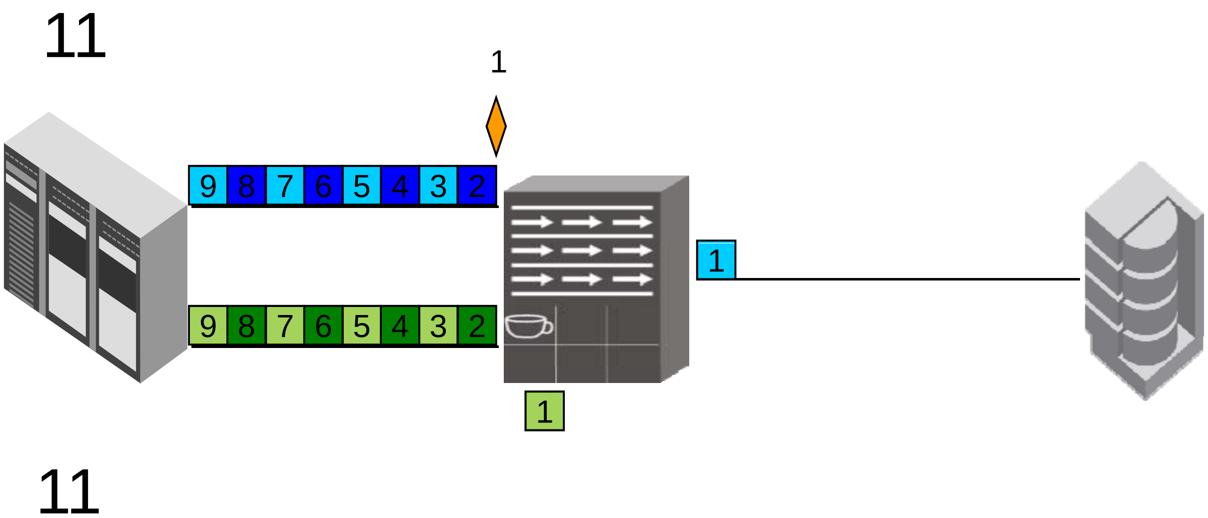


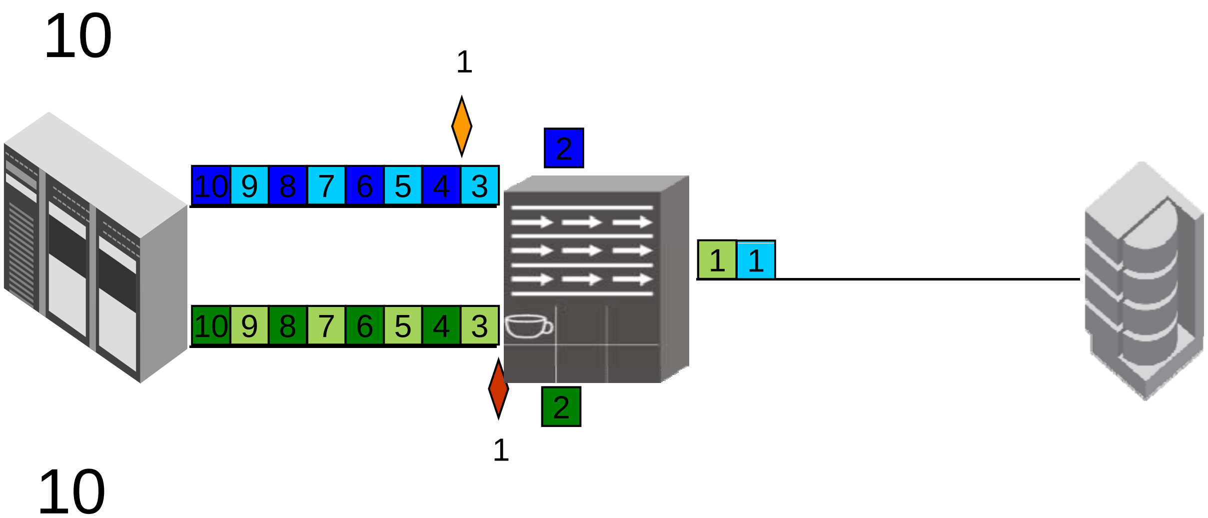
13

12

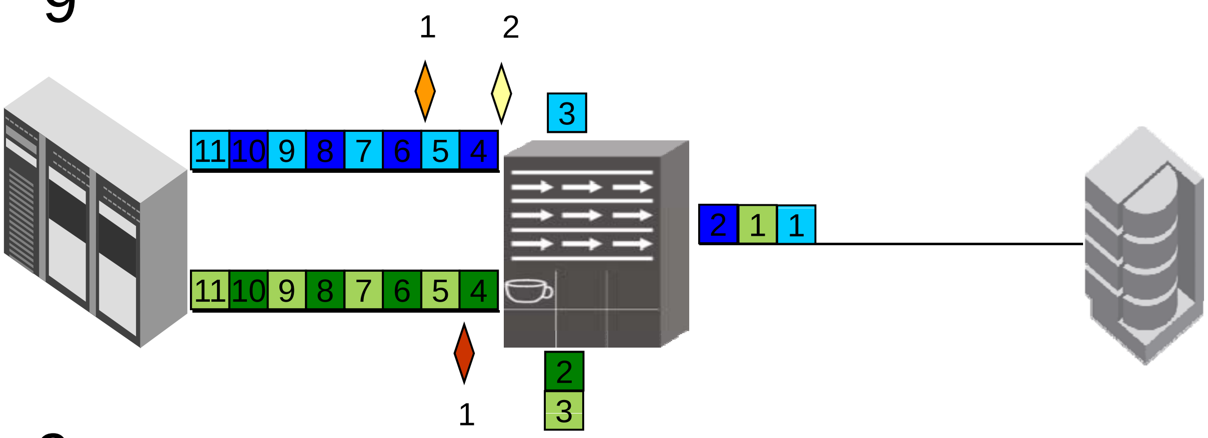


12


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

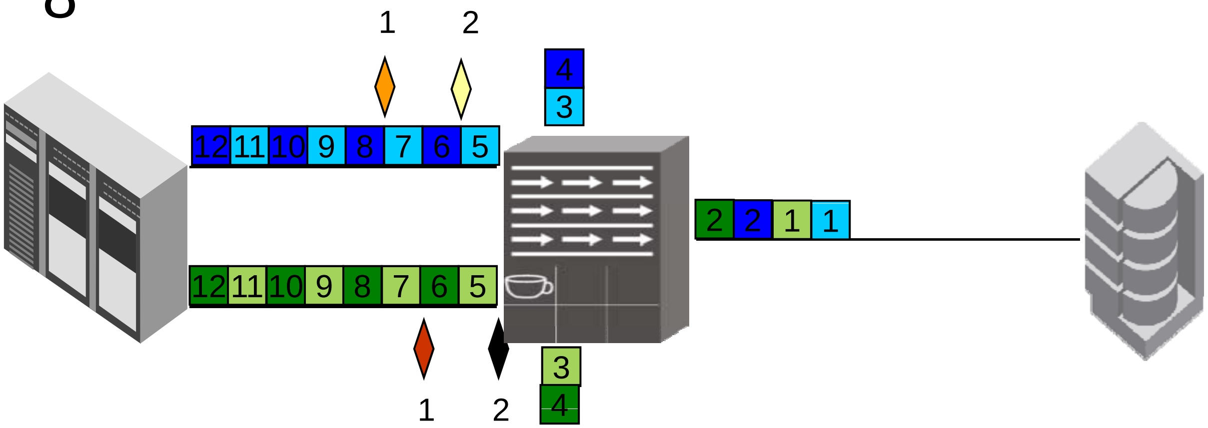

10
1
10 9 8 7 6 5 4 3
2
1 1
10 9 8 7 6 5 4 3
2
1
10

9

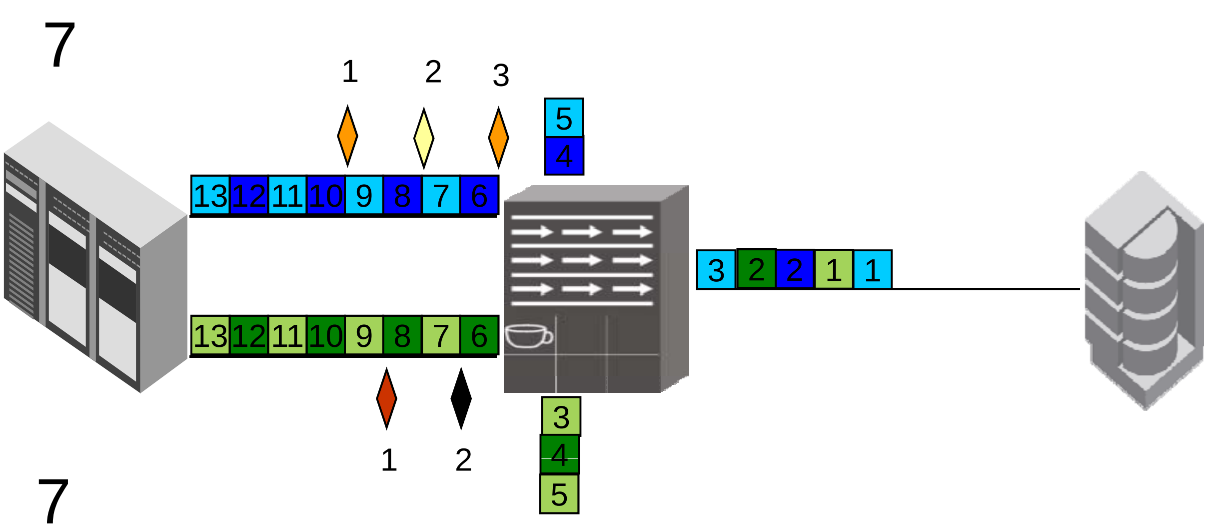


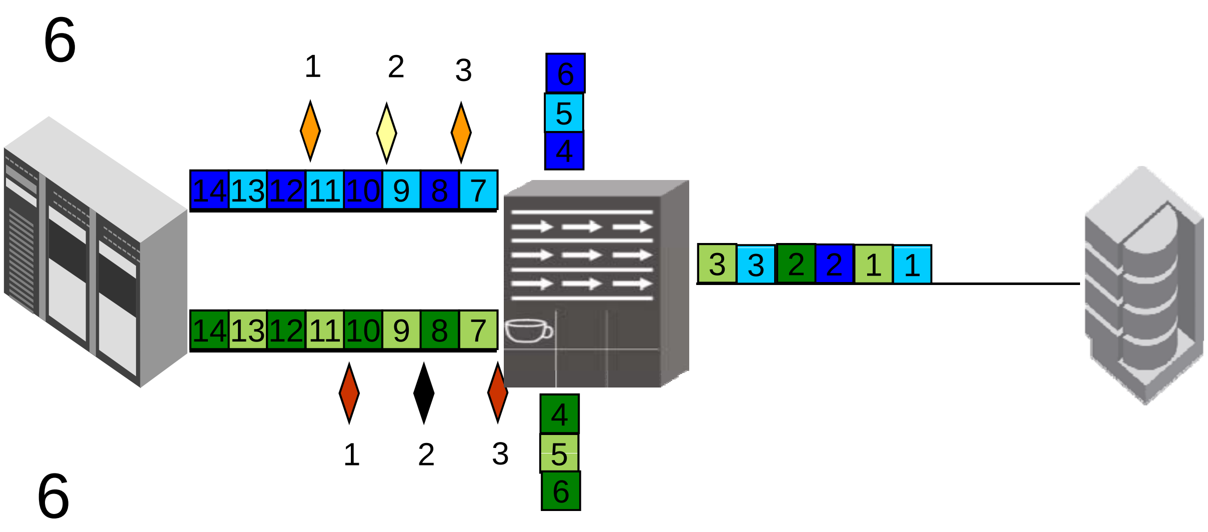
9

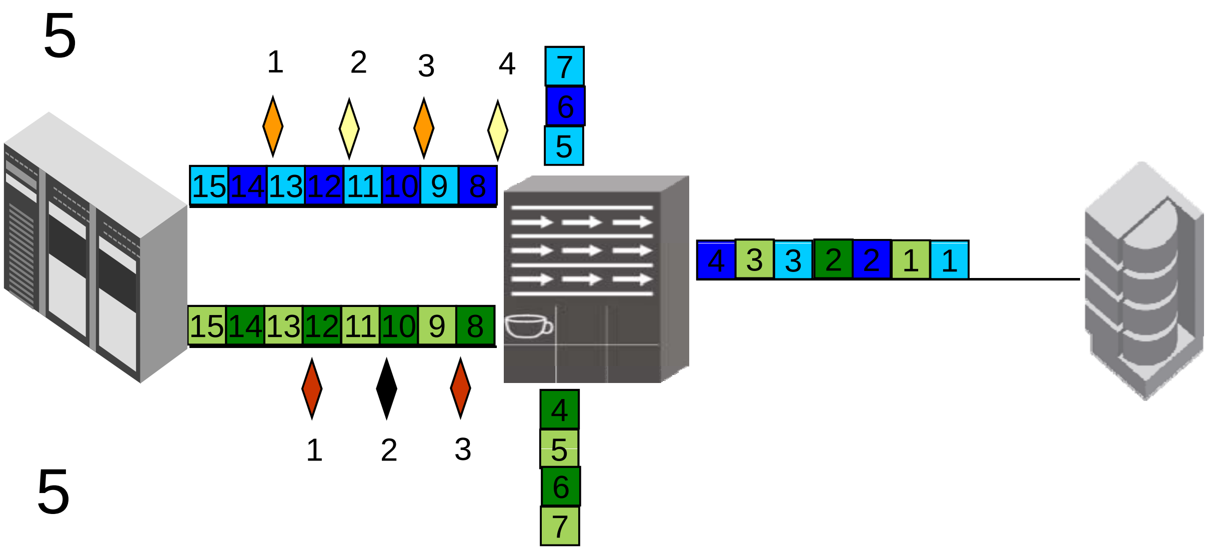
8

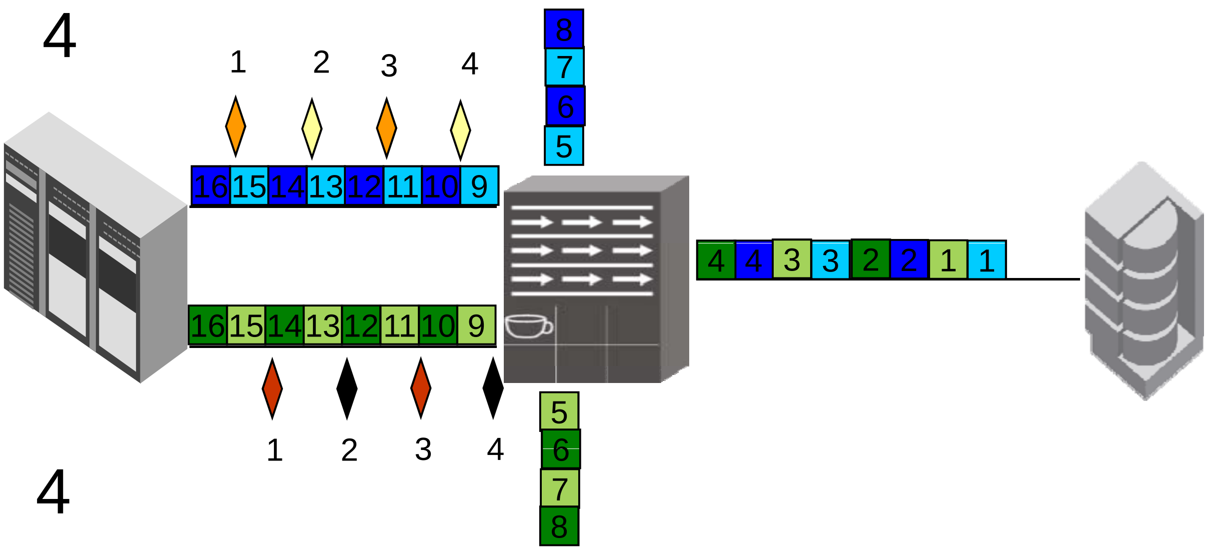


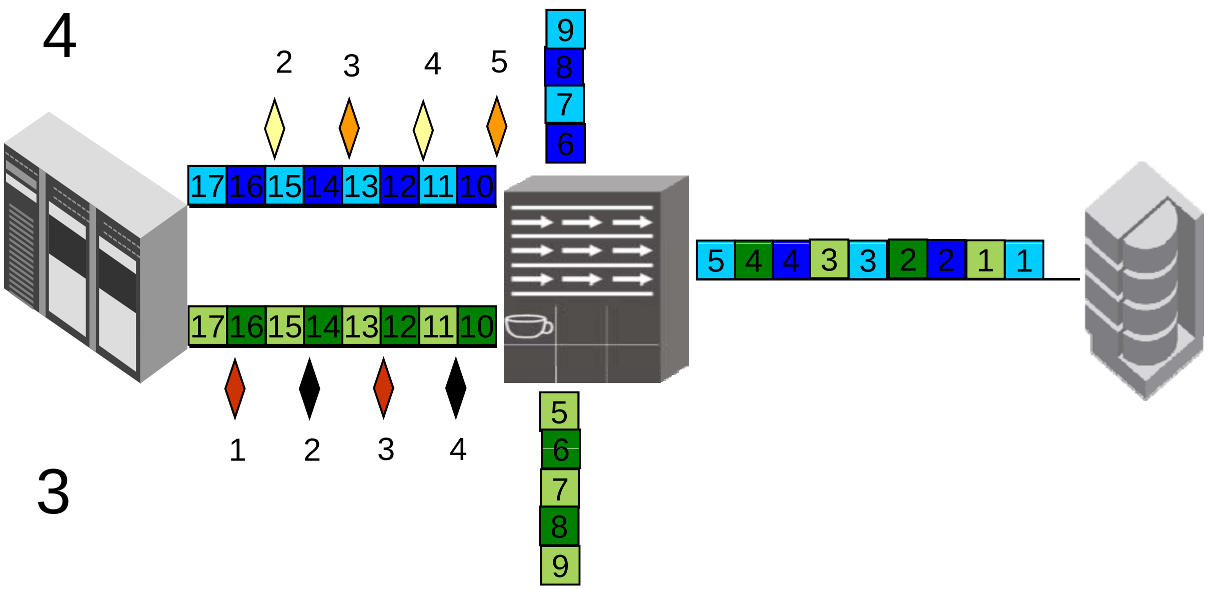
8

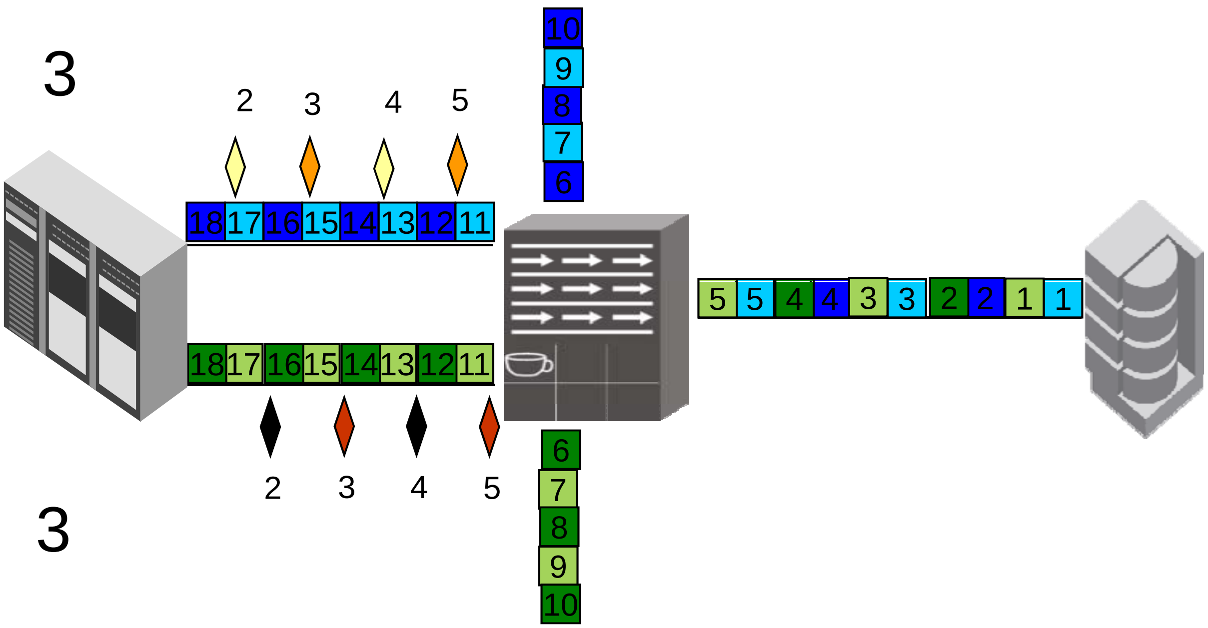

7
1
2
3
5
4
13
12
11
10
9
8
7
6
13
12
11
10
9
8
7
6
3
2
2
1
1
1
2
3
4
5
7

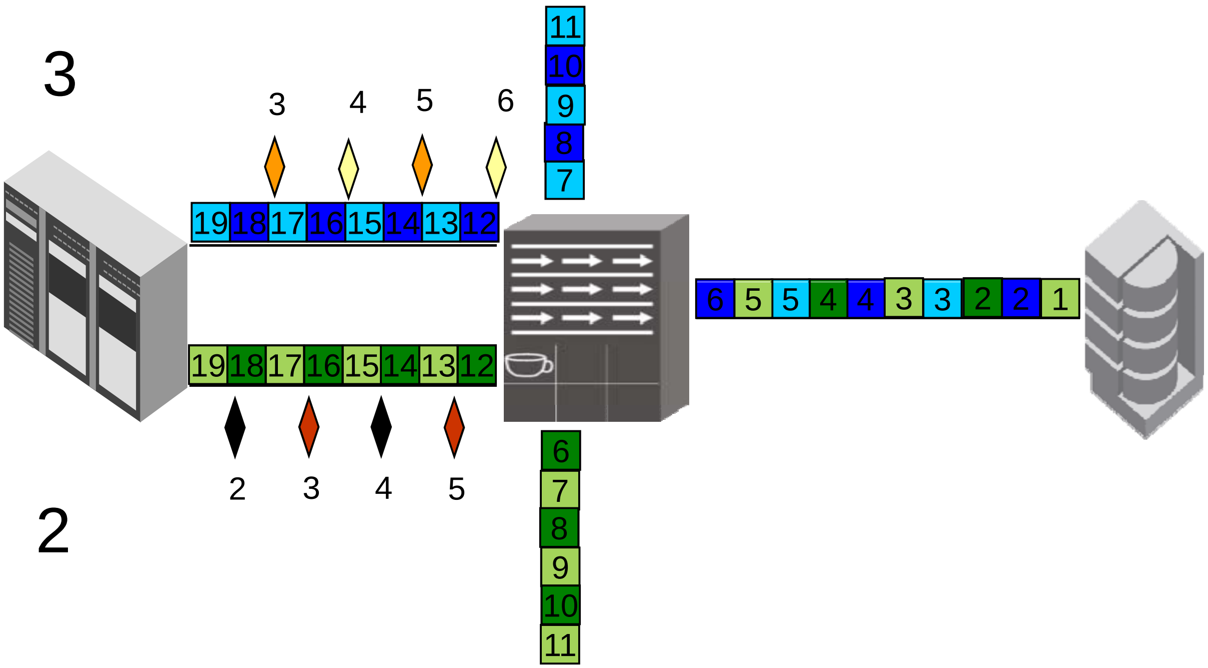

6
1
2
3
6
5
4
14 13 12 11 10 9 8 7
3 3 2 2 1 1
14 13 12 11 10 9 8 7
4
5
6
1
2
3
6

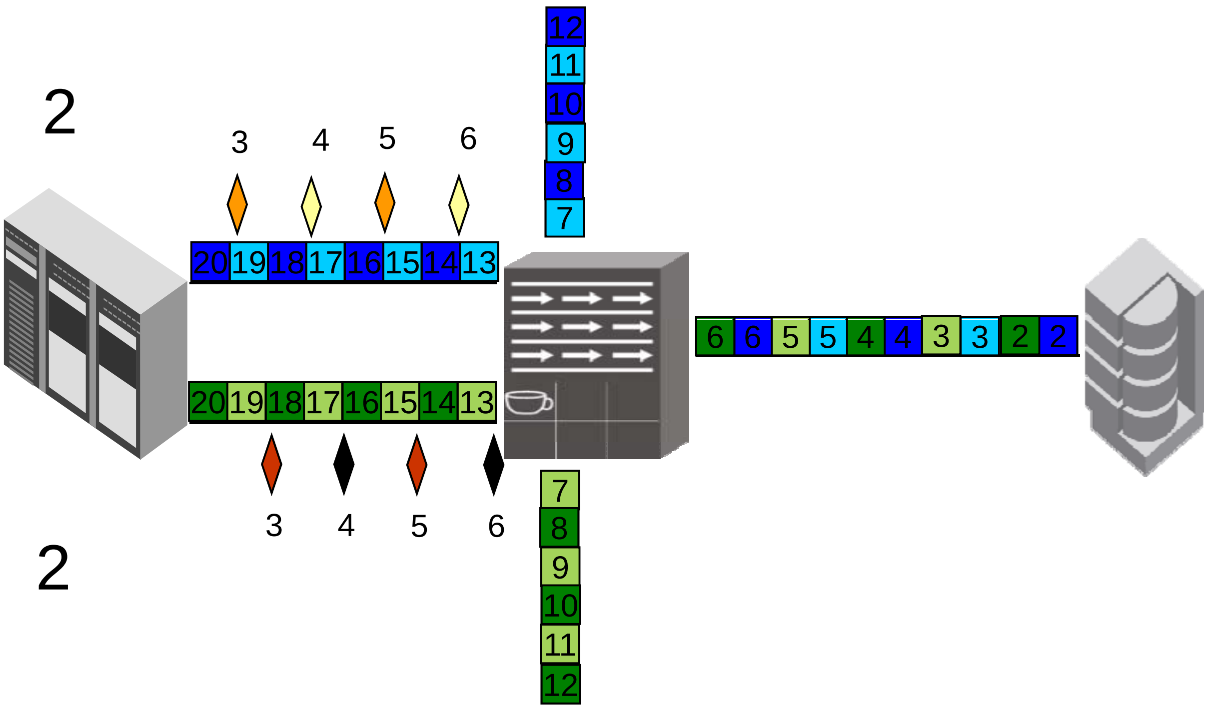

5
1
2
3
4
7
6
5
15 14 13 12 11 10 9 8
4 3 3 2 2 1 1
15 14 13 12 11 10 9 8
4
5
6
7
1
2
3
5

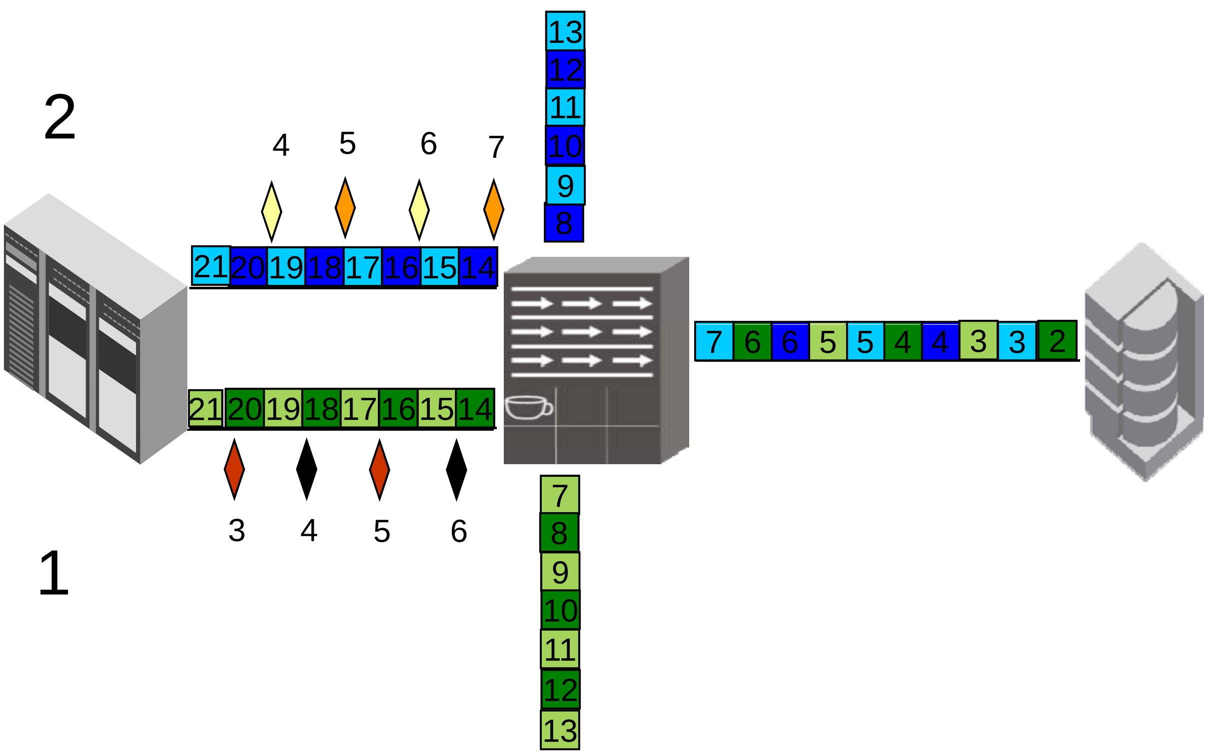

4
1
2
3
4
8
7
6
5
16 15 14 13 12 11 10 9
4 4 3 3 2 2 1 1
16 15 14 13 12 11 10 9
5
6
7
8
1
2
3
4
4


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

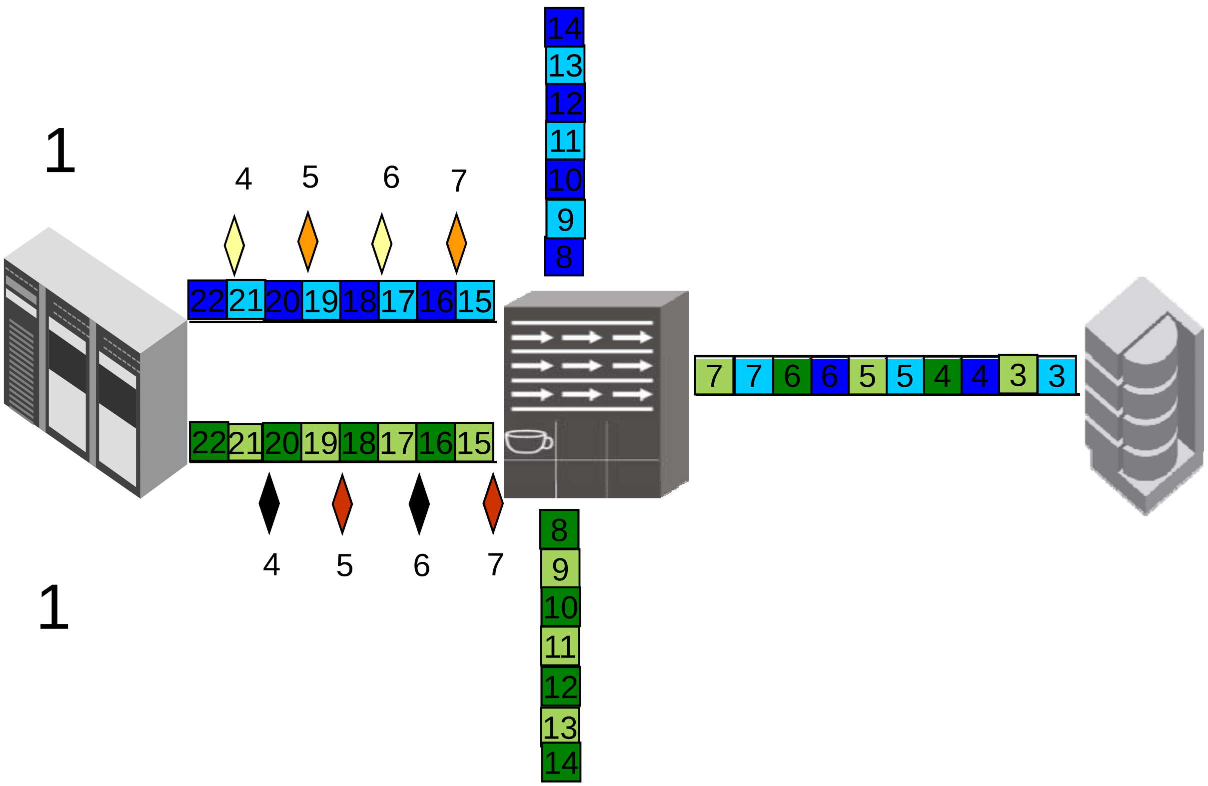

3
2
3
4
5
10
9
8
7
6
18 17 16 15 14 13 12 11
5 5 4 4 3 3 2 2 1 1
18 17 16 15 14 13 12 11
6
7
8
9
10
2
3
4
5
3


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

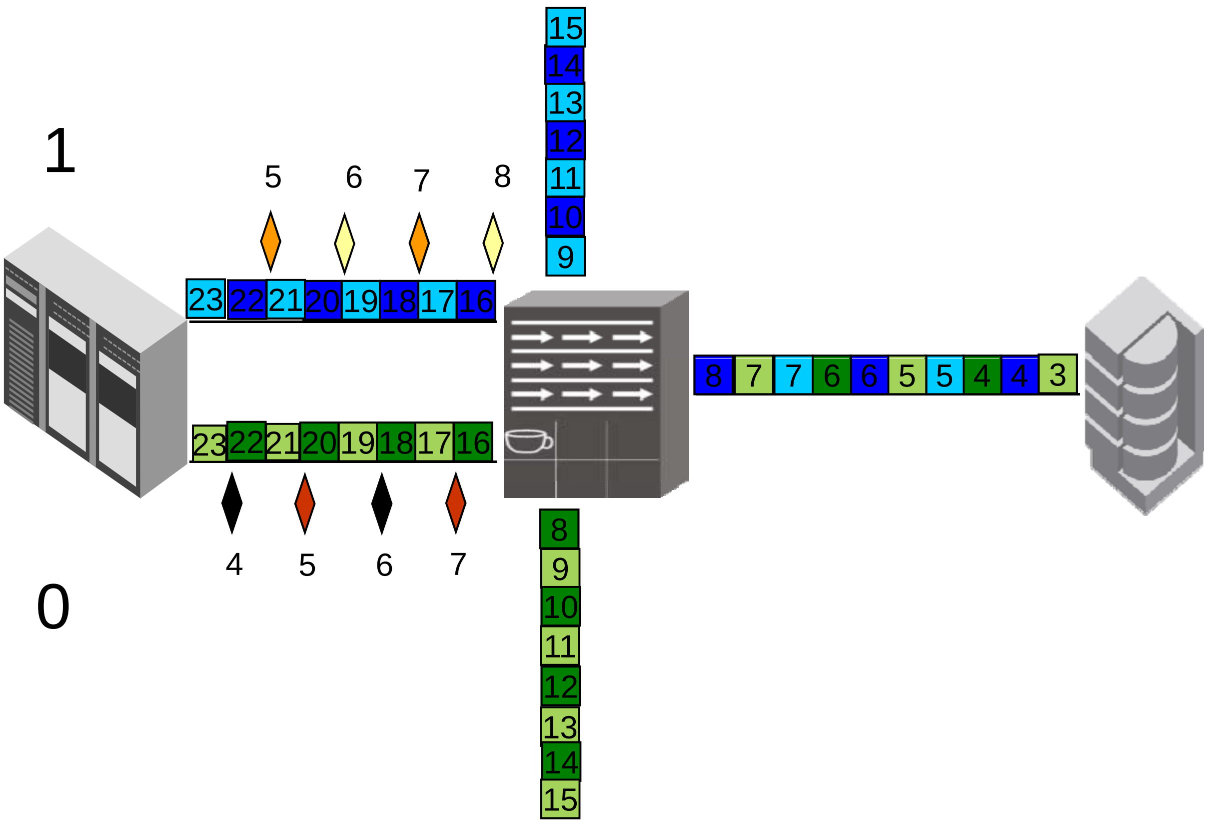

2
3
4
5
6
20 19 18 17 16 15 14 13
12 11 10 9 8 7
6 6 5 5 4 4 3 3 2 2
20 19 18 17 16 15 14 13
7 8 9 10 11 12
3
4
5
6
2

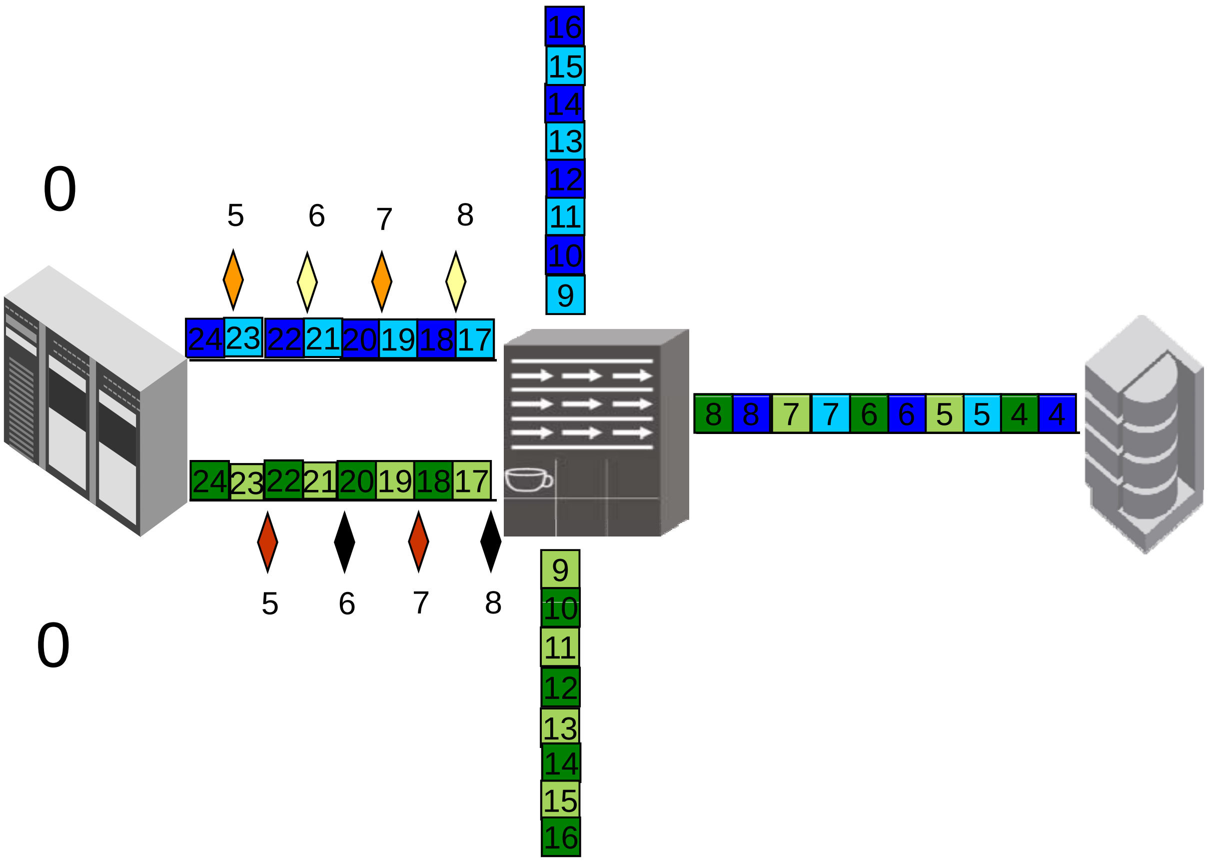


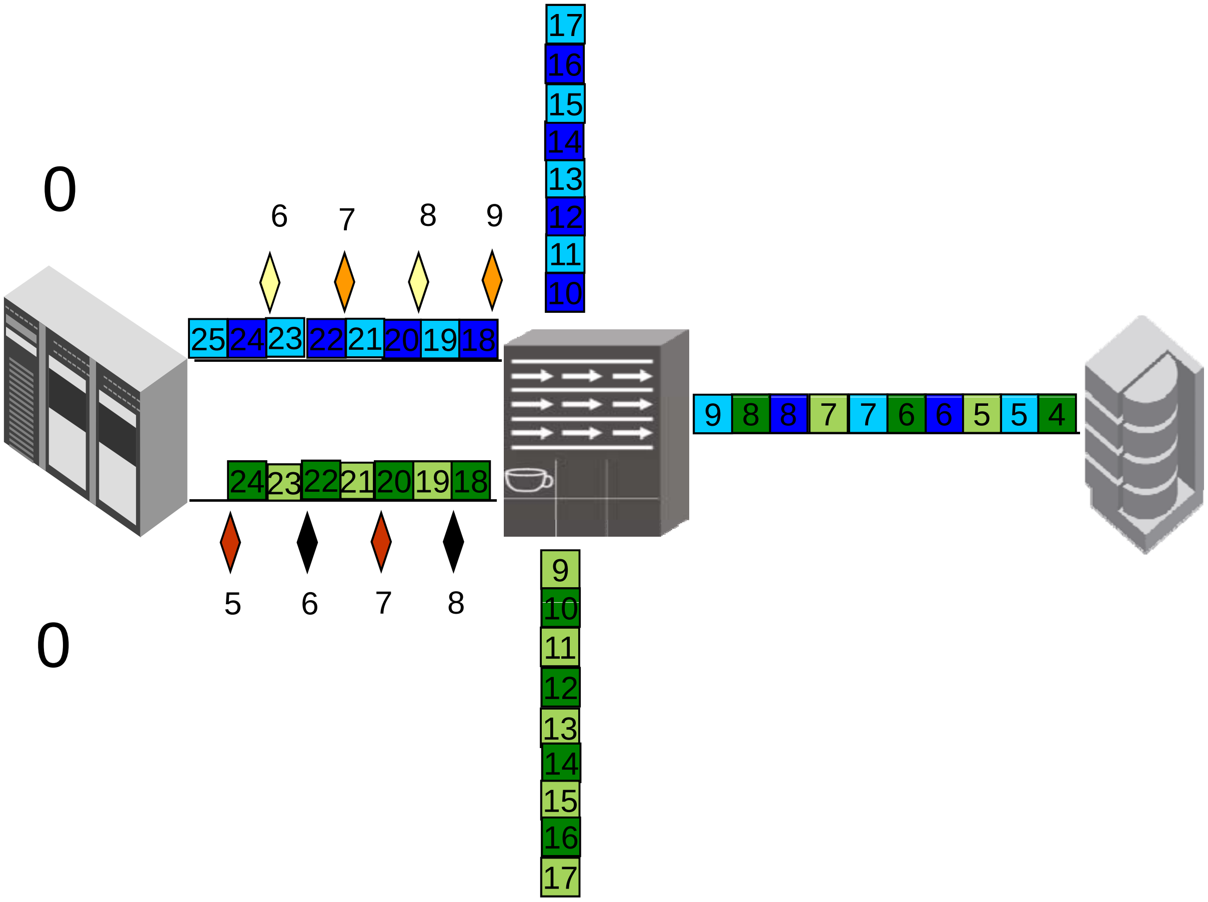
2
4 5 6 7
13 12 11 10 9 8
21 20 19 18 17 16 15 14
7 6 6 5 5 4 4 3 3 2
21 20 19 18 17 16 15 14
3 4 5 6
7 8 9 10 11 12 13
1

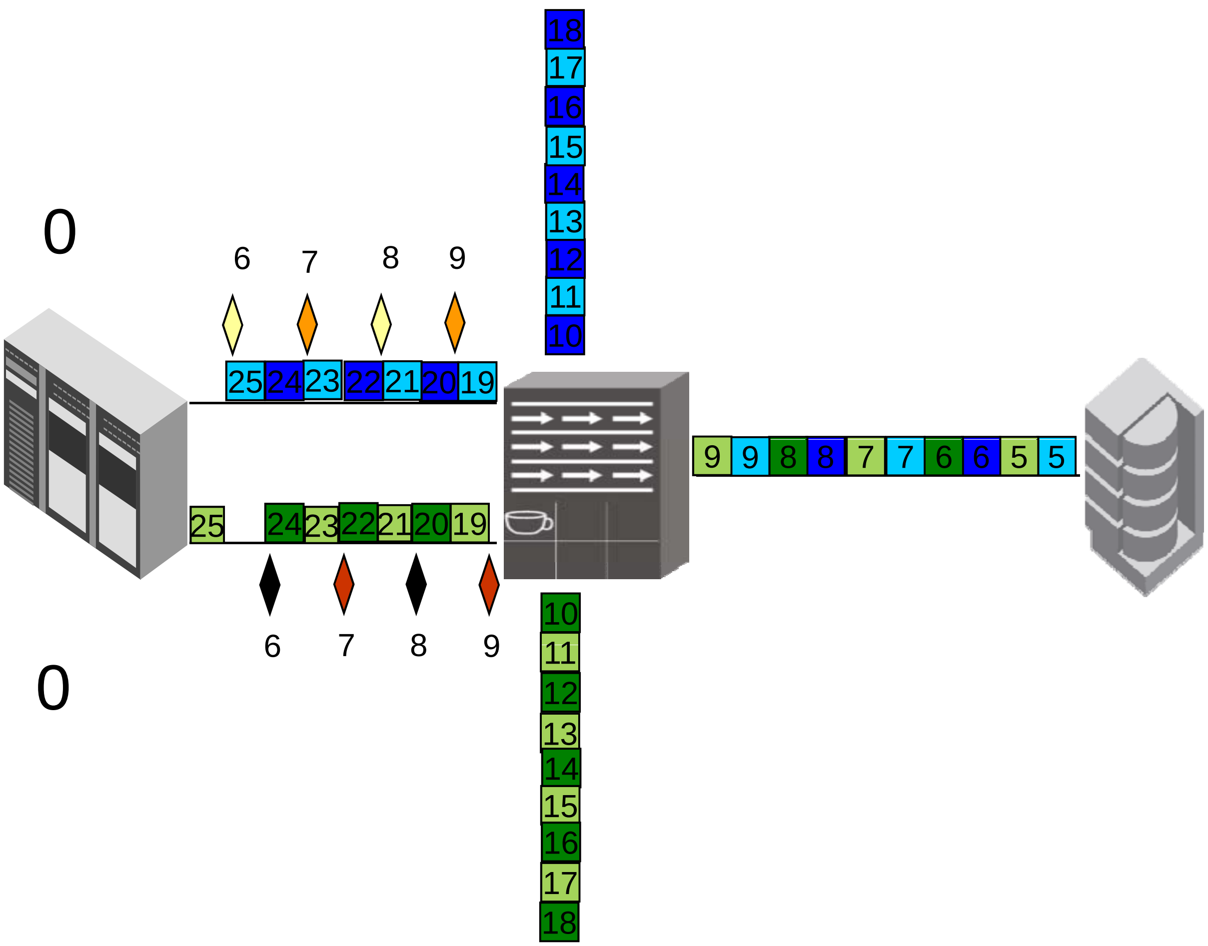


1
4 5 6 7
14
13
12
11
10
9
8
22 21 20 19 18 17 16 15
7 7 6 6 5 5 4 4 3 3
22 21 20 19 18 17 16 15
8
9
10
11
12
13
14
4 5 6 7
1

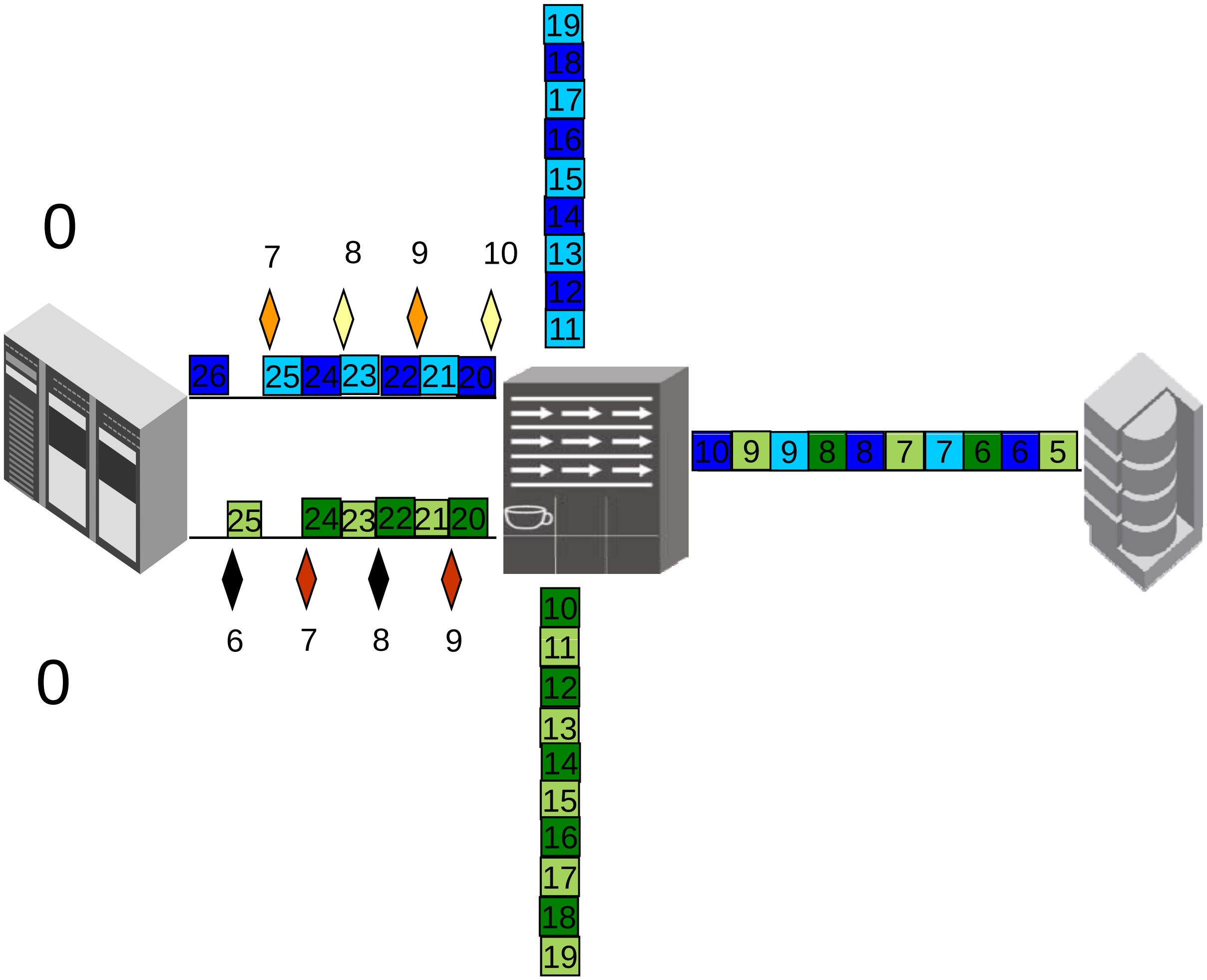

1
5
6
7
8
15
14
13
12
11
10
9
23 22 21 20 19 18 17 16
8 7 7 6 6 5 5 4 4 3
23 22 21 20 19 18 17 16
4
5
6
7
8
9
10
11
12
13
14
15
0

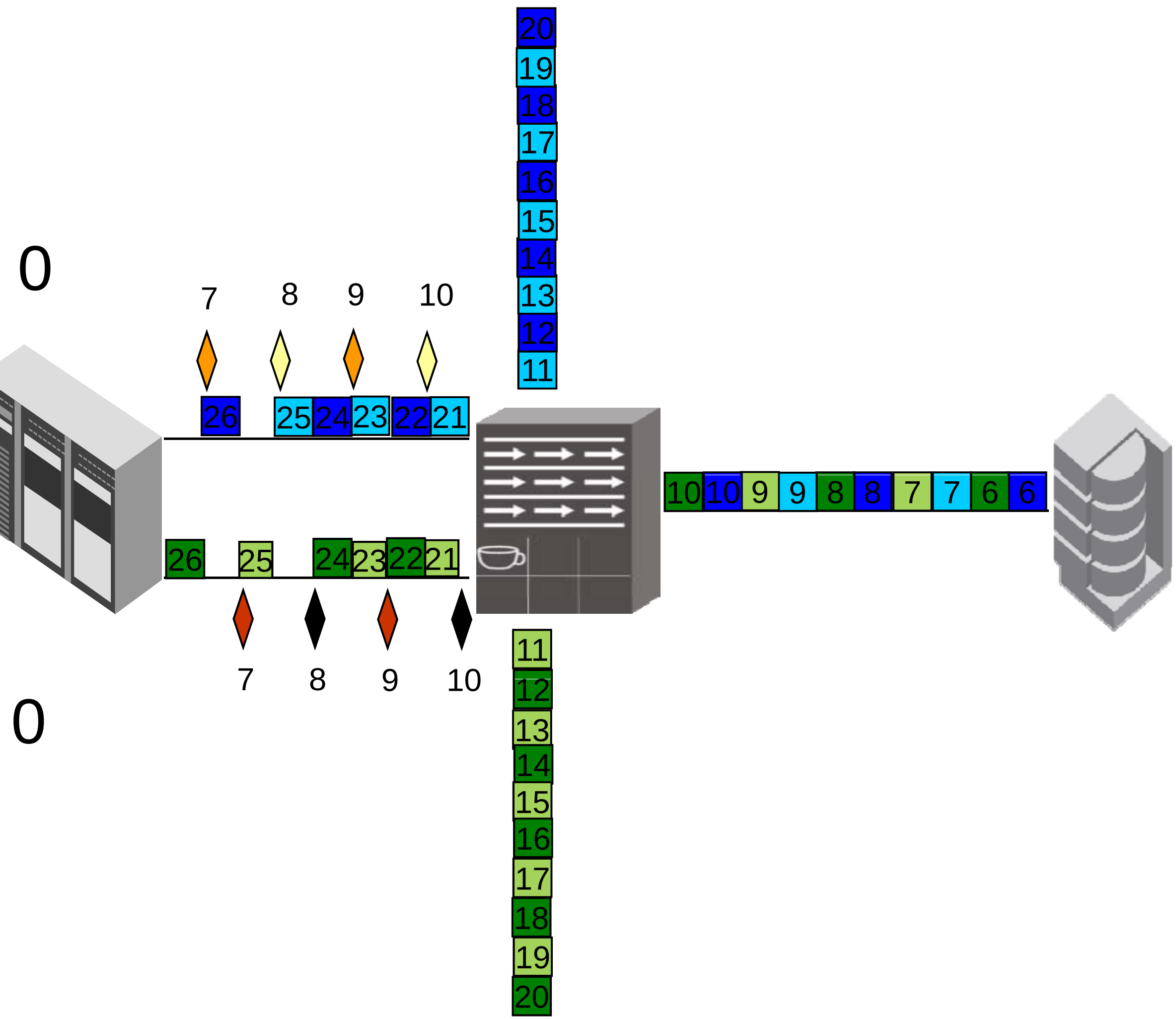

0
5
6
7
8
16
15
14
13
12
11
10
9
24 23 22 21 20 19 18 17
8 8 7 7 6 6 5 5 4 4
24 23 22 21 20 19 18 17
5
6
7
8
0
9
10
11
12
13
14
15
16


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

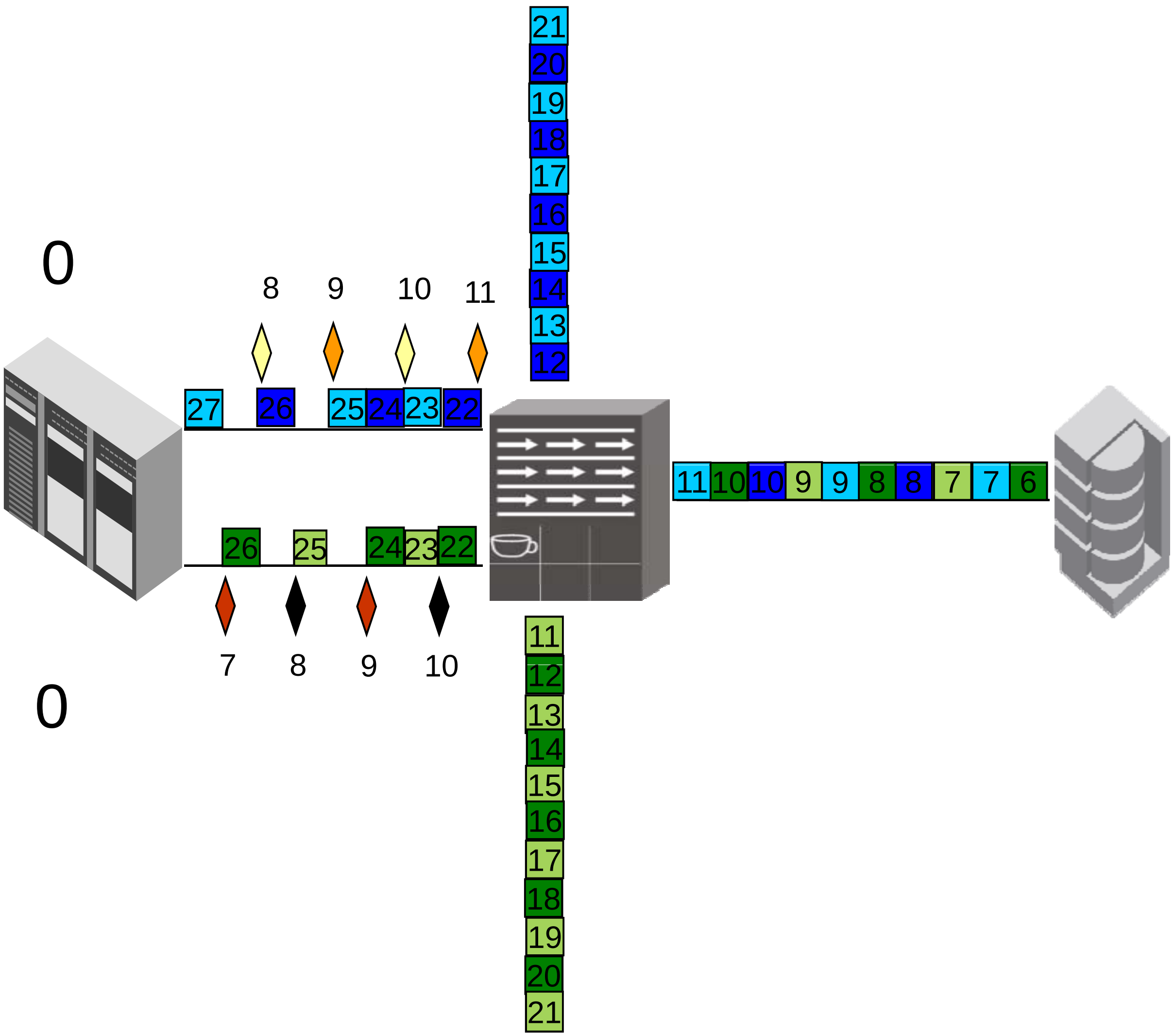


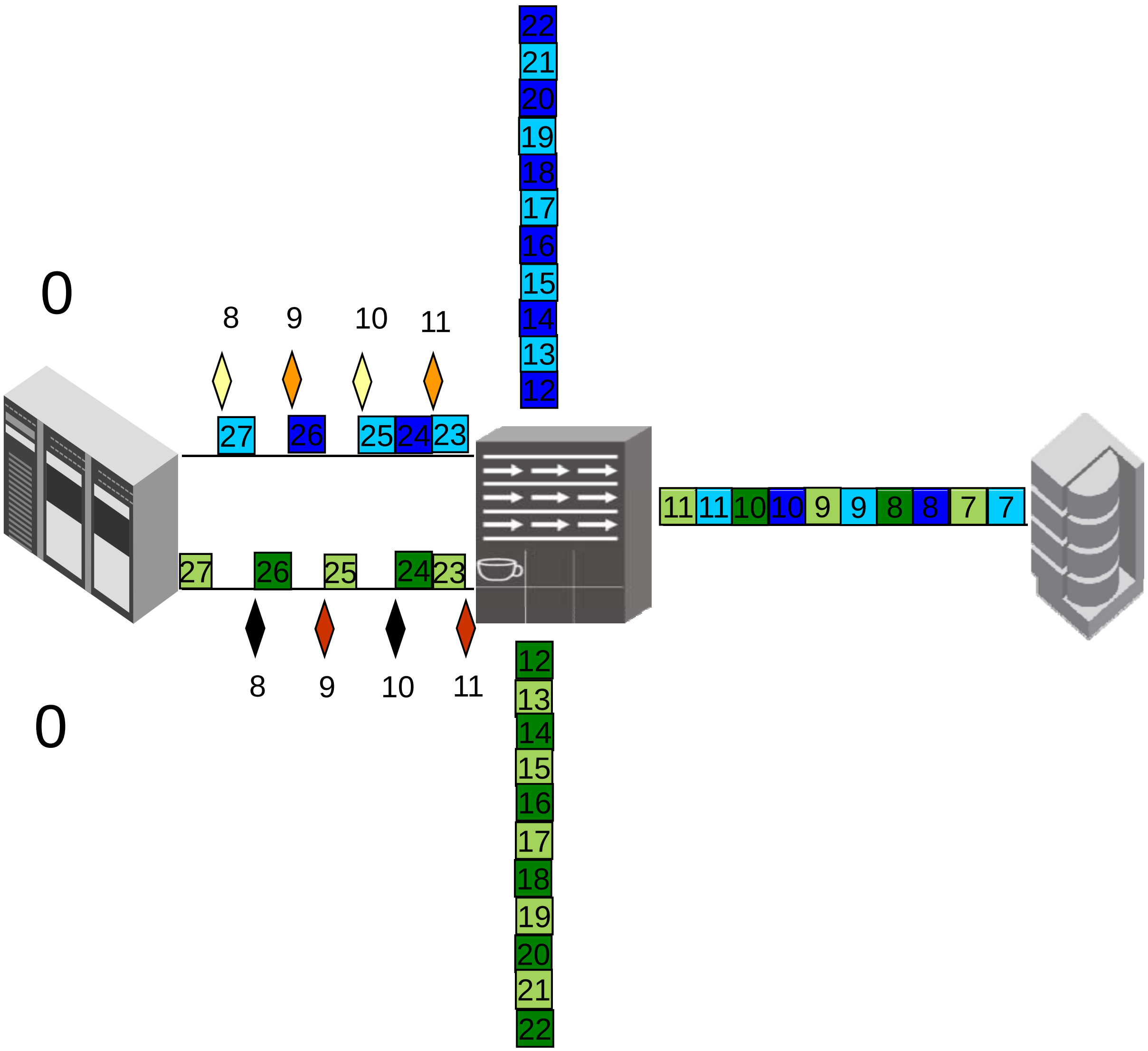
0
6
7
8
9
25 24 23 22 21 20 19
18 17 16 15 14 13 12 11 10
9 9 8 8 7 7 6 6 5 5
25 24 23 22 21 20 19
10 11 12 13 14 15 16 17 18
6
7
8
9
0

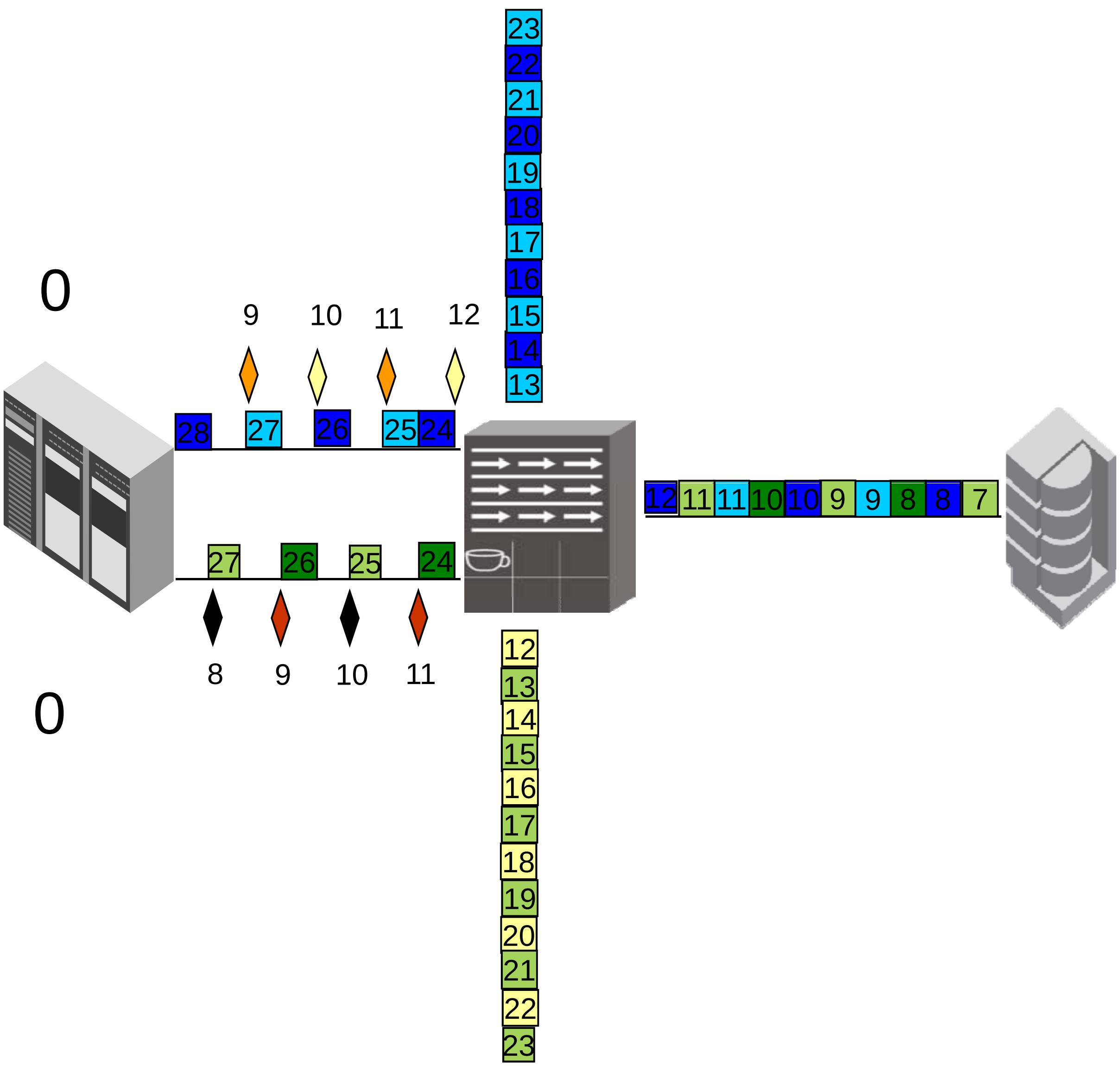

0
7 8 9 10
26 25 24 23 22 21 20
25 24 23 22 21 20
6 7 8 9
0
19 18 17 16 15 14 13 12 11
10 9 9 8 8 7 7 6 6 5
10 11 12 13 14 15 16 17 18 19



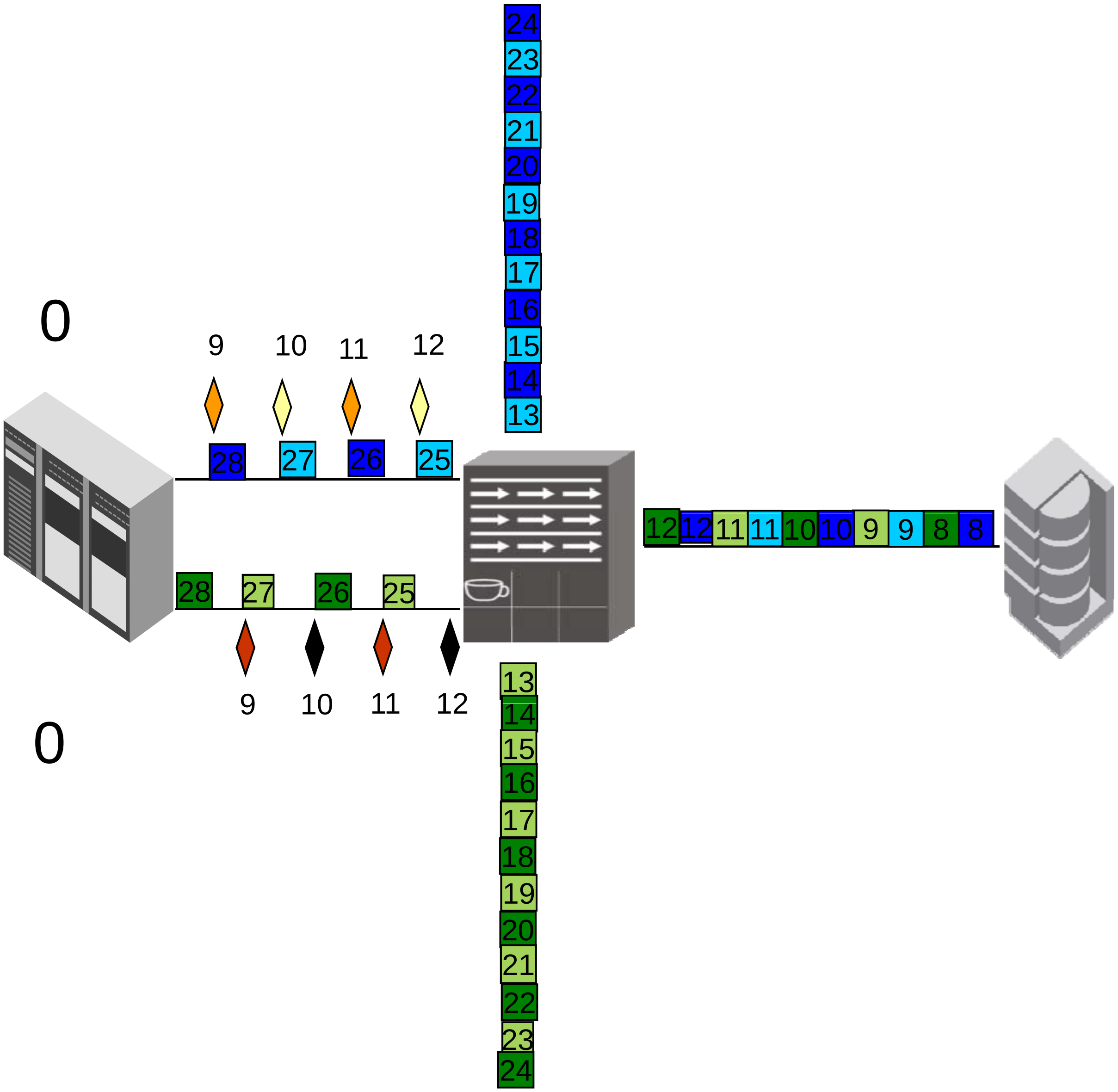
0
7
8
9
10
20
19
18
17
16
15
14
13
12
11
26
25
24
23
22
21
10
10
9
9
8
8
7
7
6
6
26
25
24
23
22
21
7
8
9
10
11
12
13
14
15
16
17
18
19
20
0


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

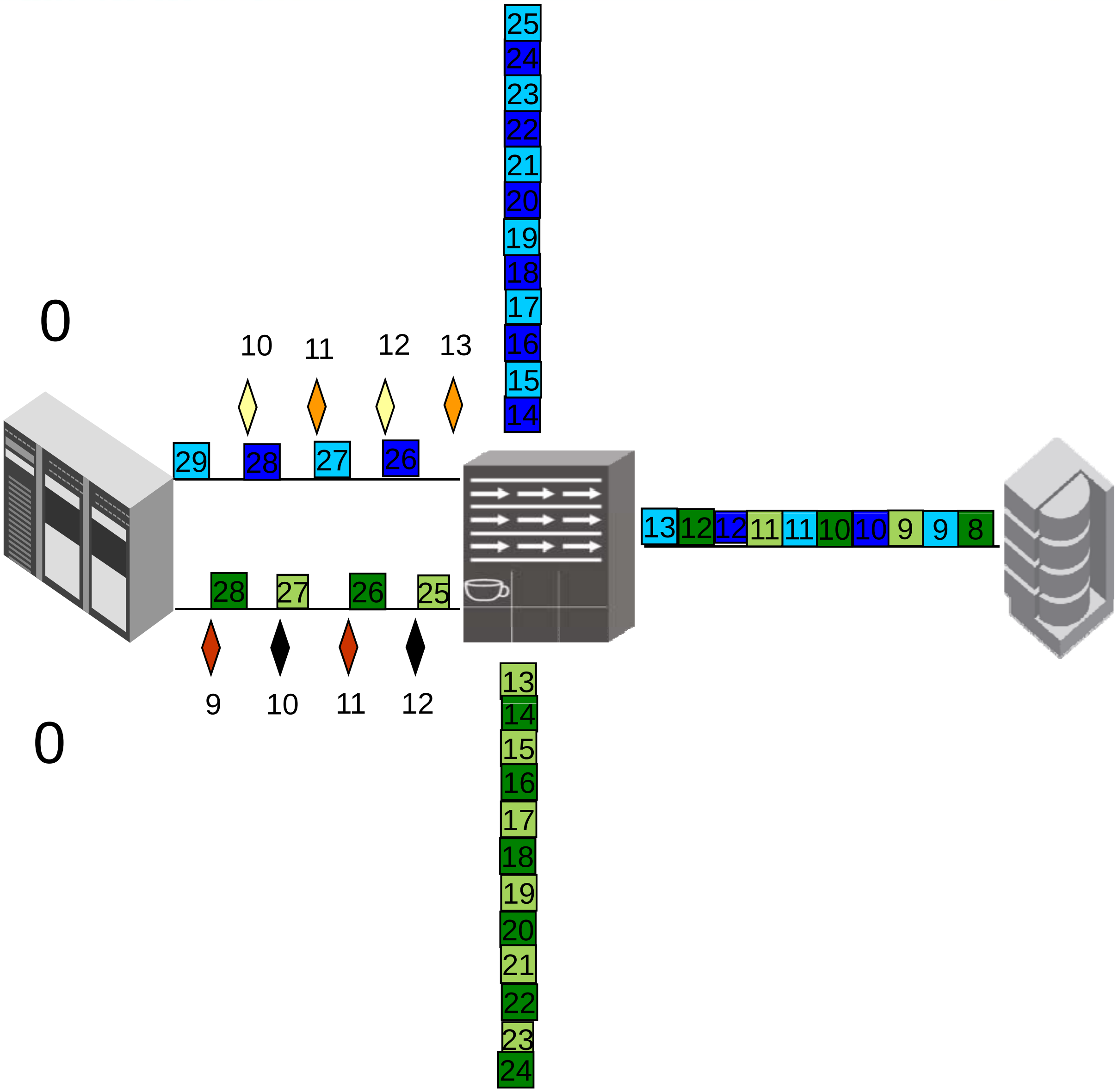

0
8 9 10 11
27 26 25 24 23
27 26 25 24 23
8 9 10 11
0
22 21 20 19 18 17 16 15 14 13 12
12 13 14 15 16 17 18 19 20 21 22
11 11 10 10 9 9 8 8 7 7



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



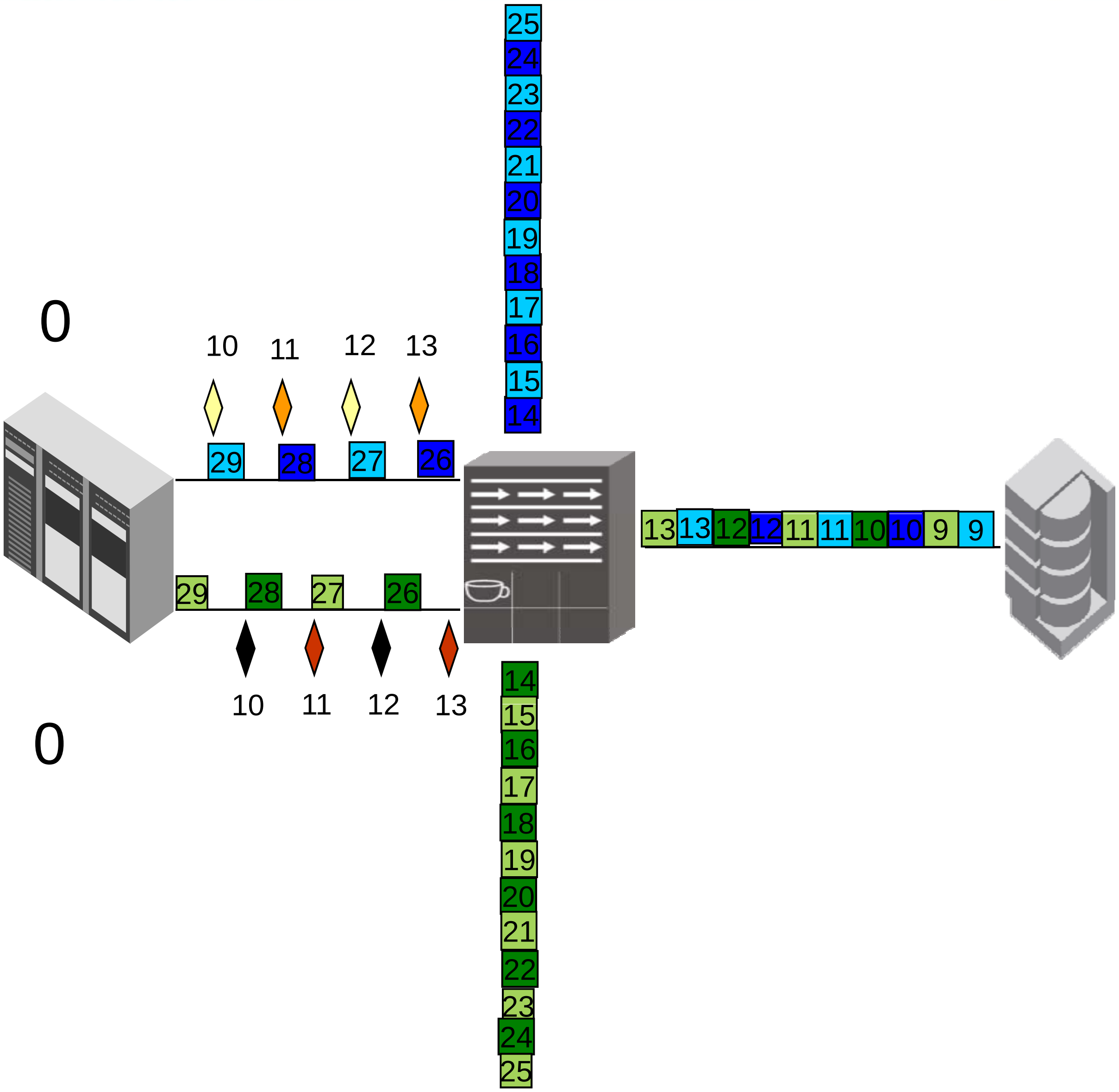
0
9
10
11
12
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12 12 11 11 10 10 9 9 8 8
28
27
26
25
9
10
11
12
0
13
14
15
16
17
18
19
20
21
22
23
24


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

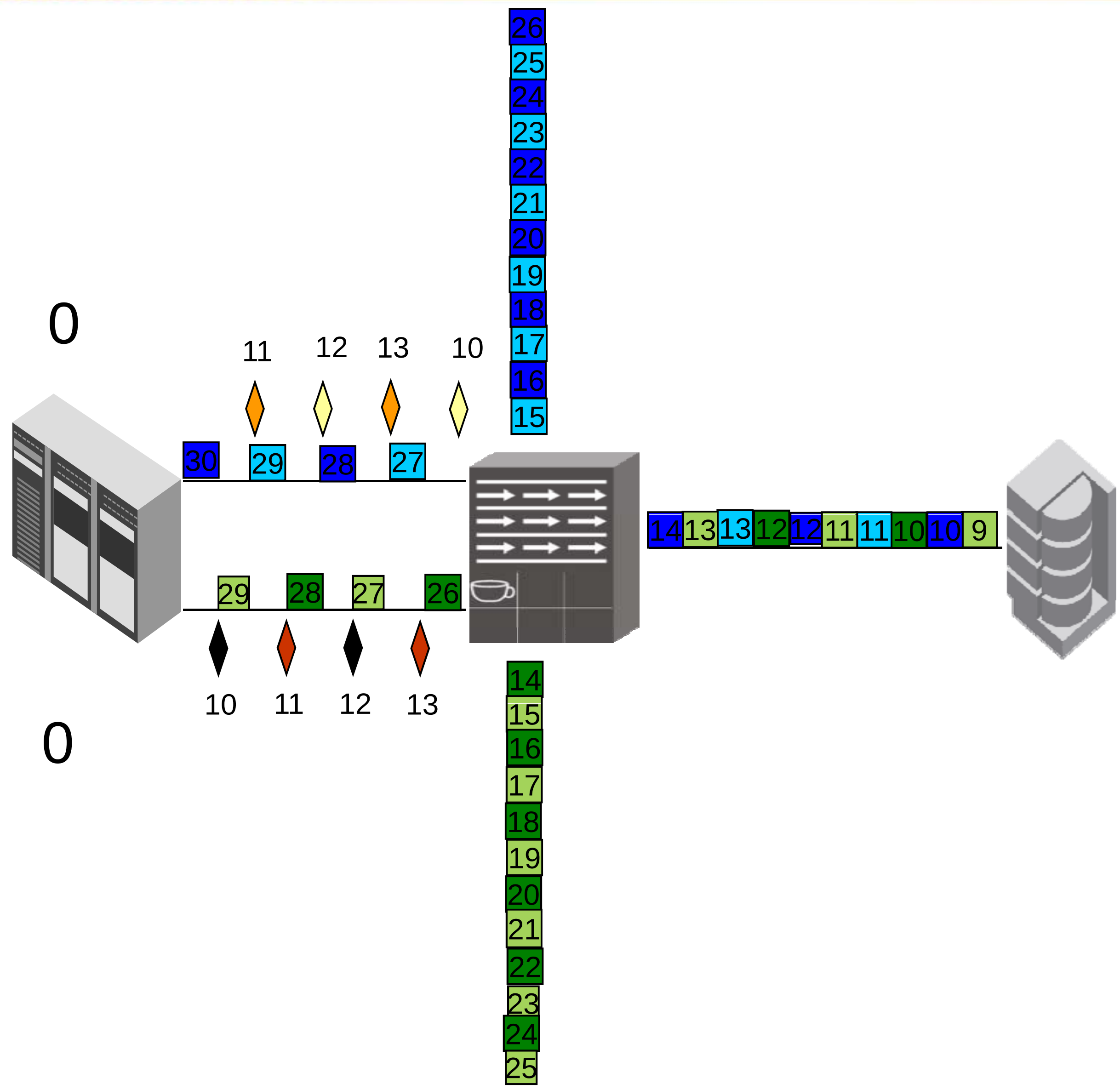
0

0

○



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



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

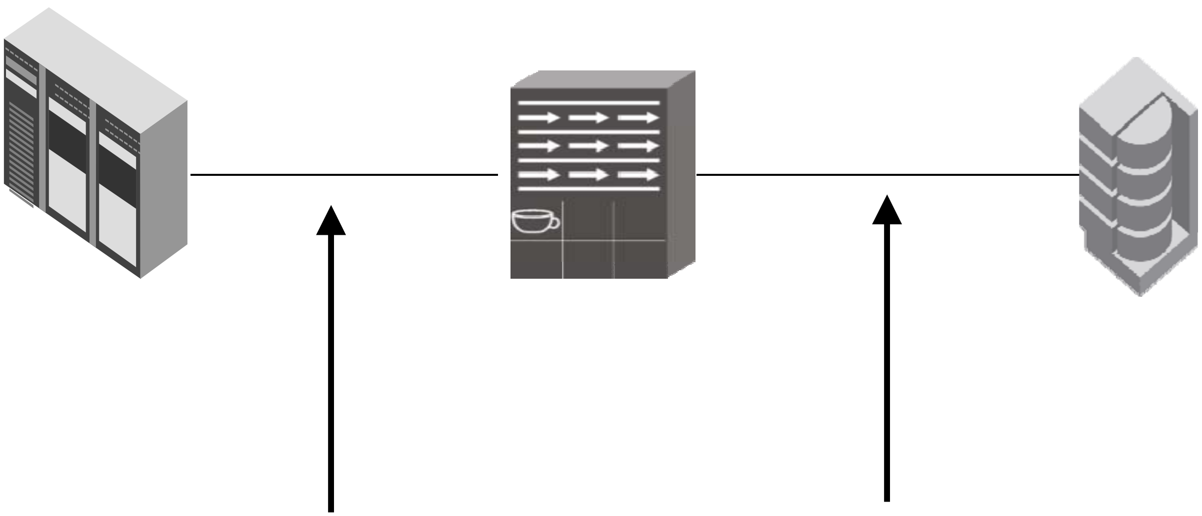
THIS PAGE INTENTIONALLY LEFT BLANK

Example: Different sized pipes

BUFFER CREDITS

Fat Pipe / Skinny Pipe

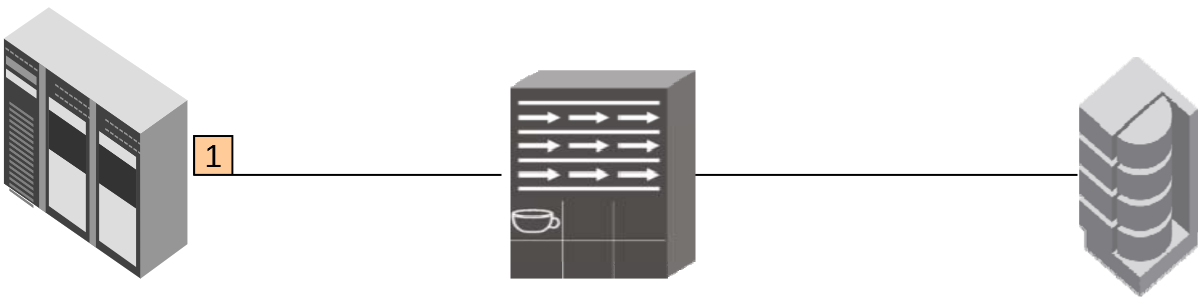
20



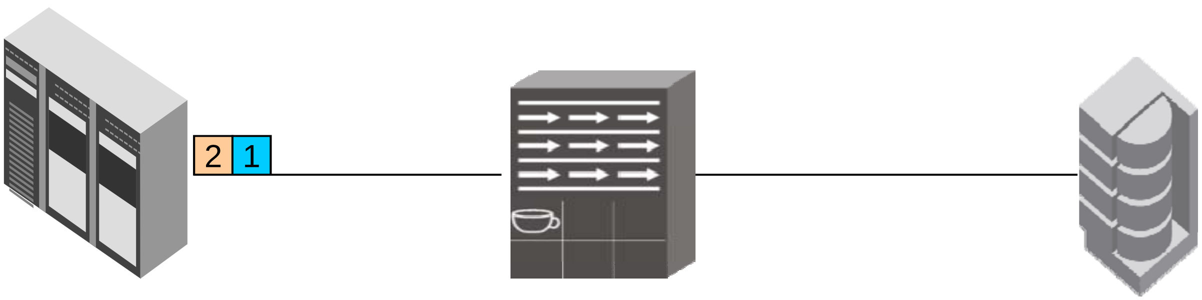
8 Gig Link

4 Gig Link

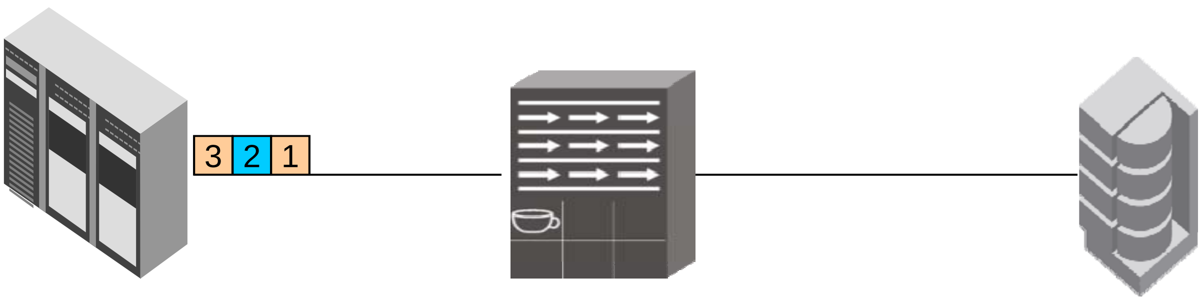
19



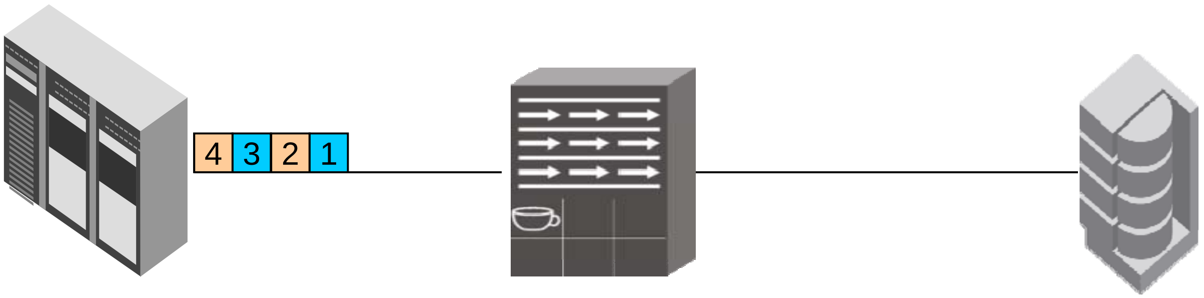
18



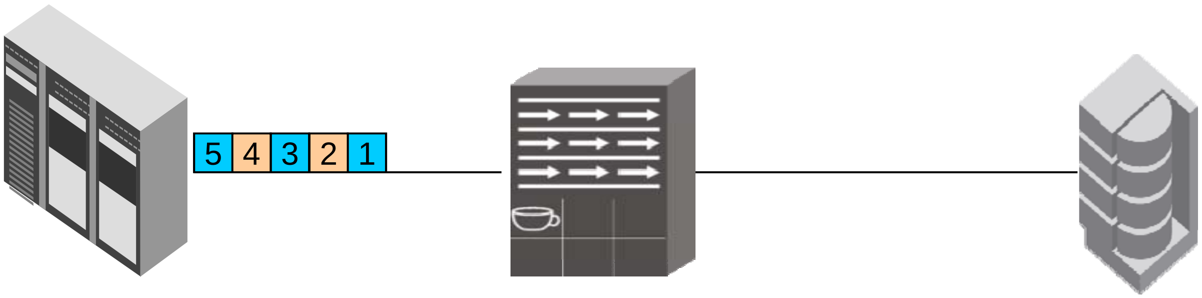
17



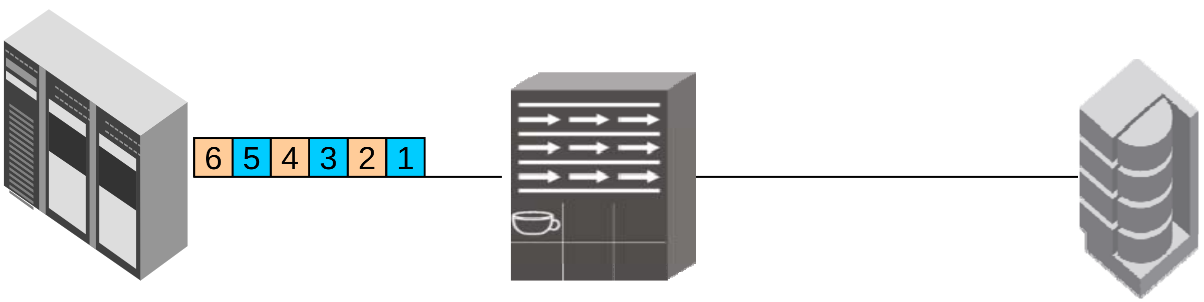
16



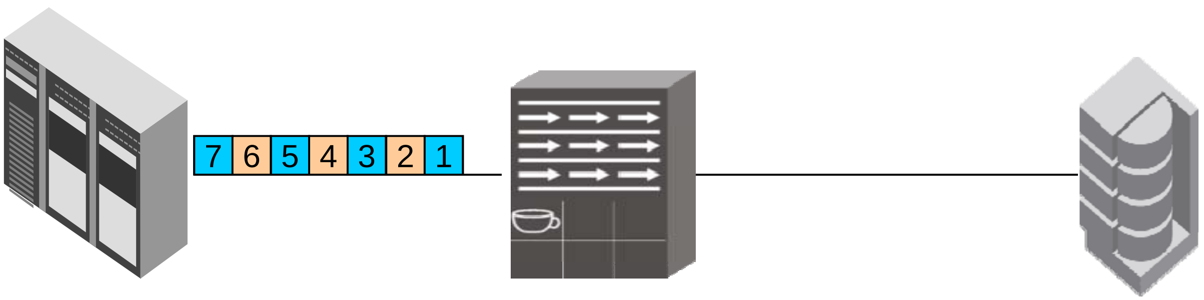
15



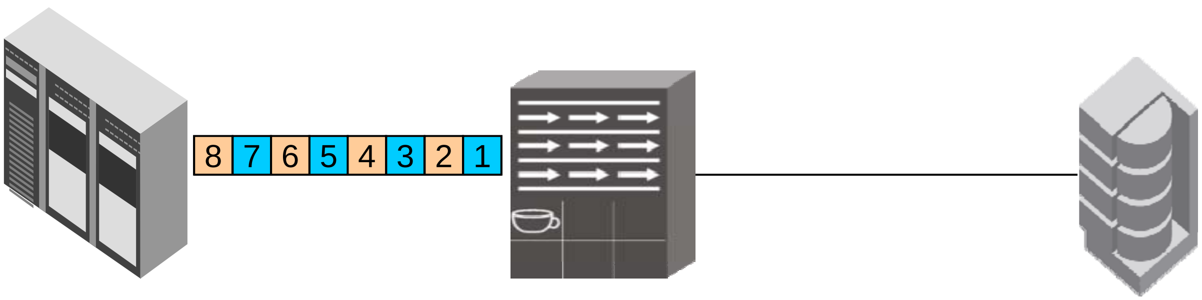
14



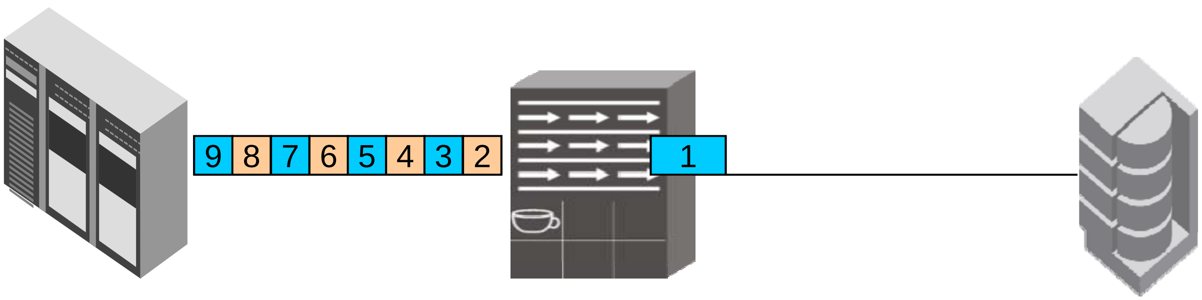
13



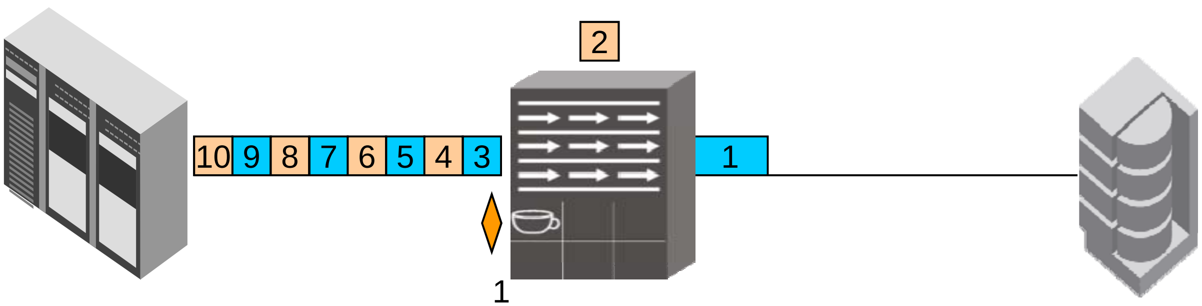
12



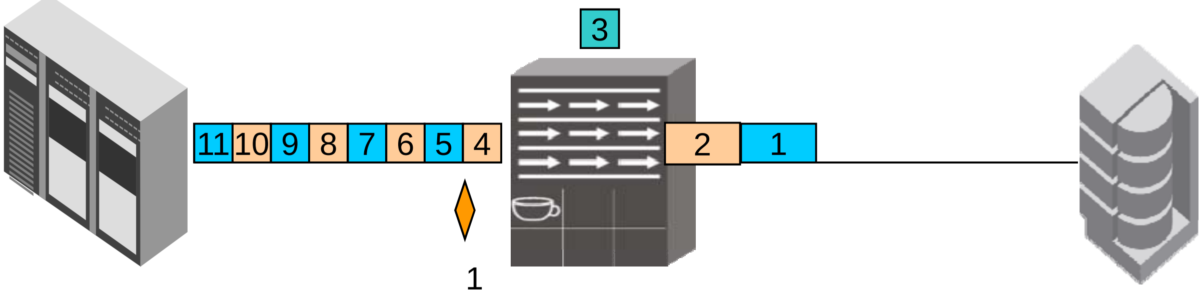
11



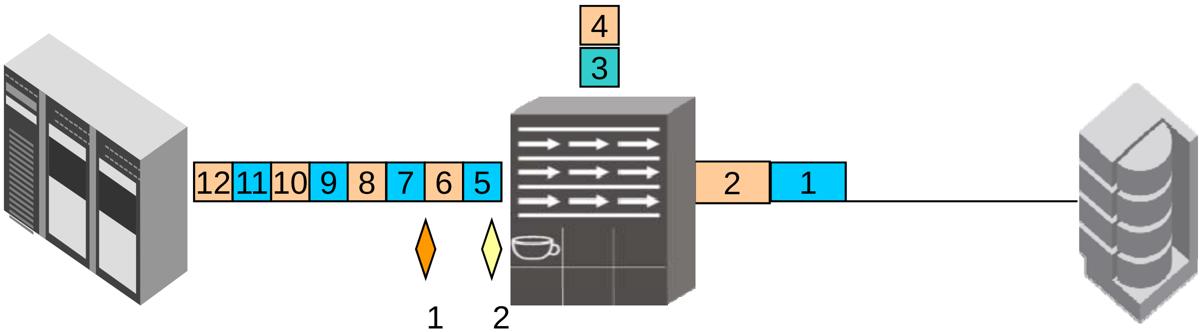
10



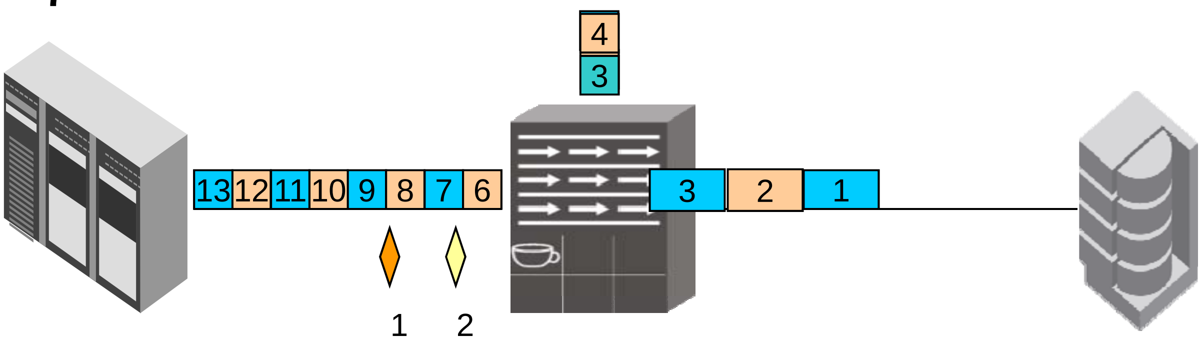
9



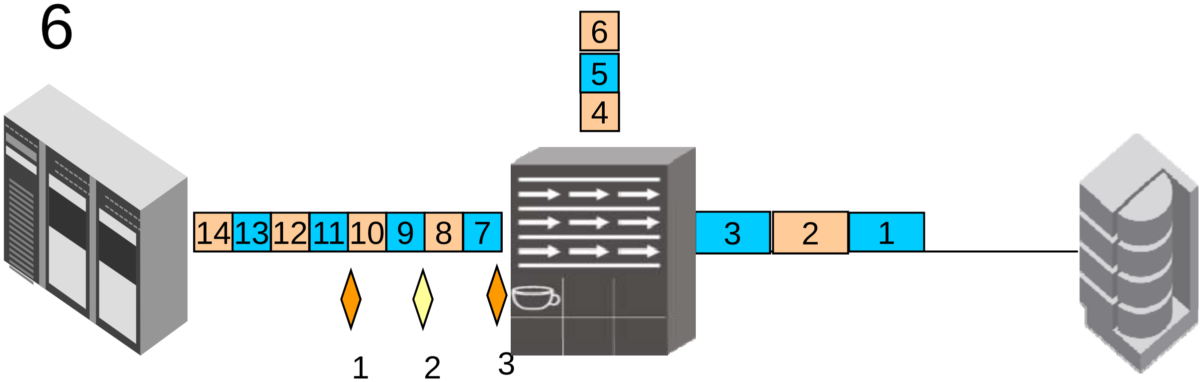
8



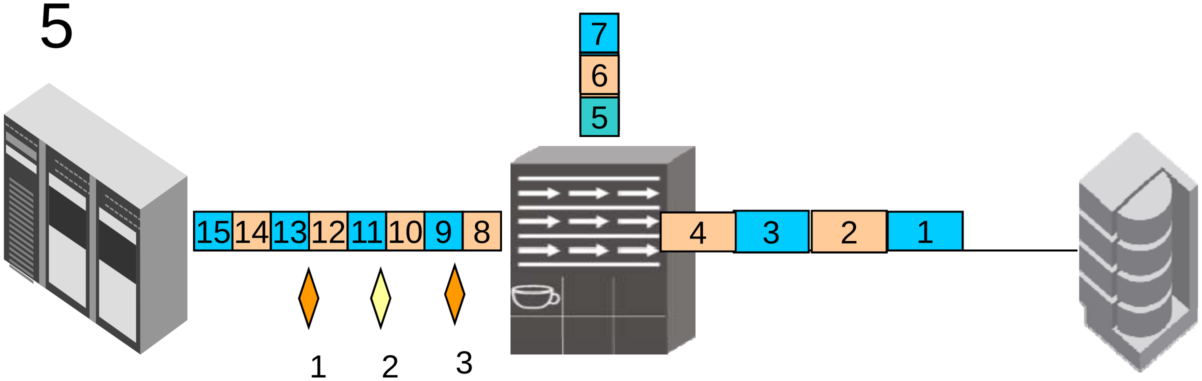
7



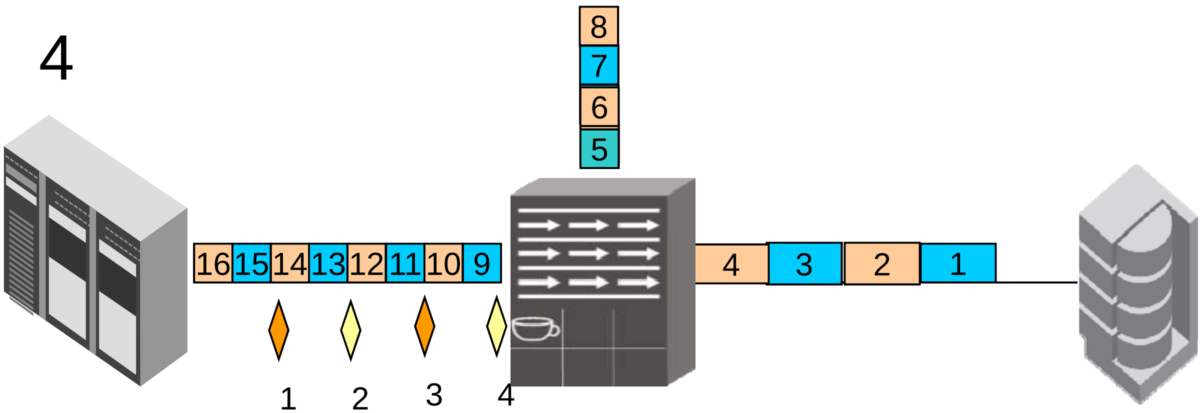
6



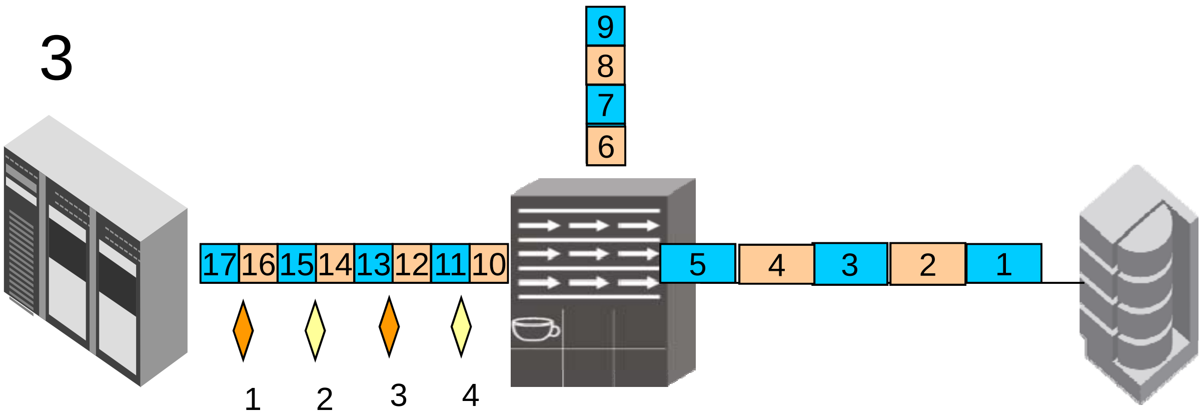
5



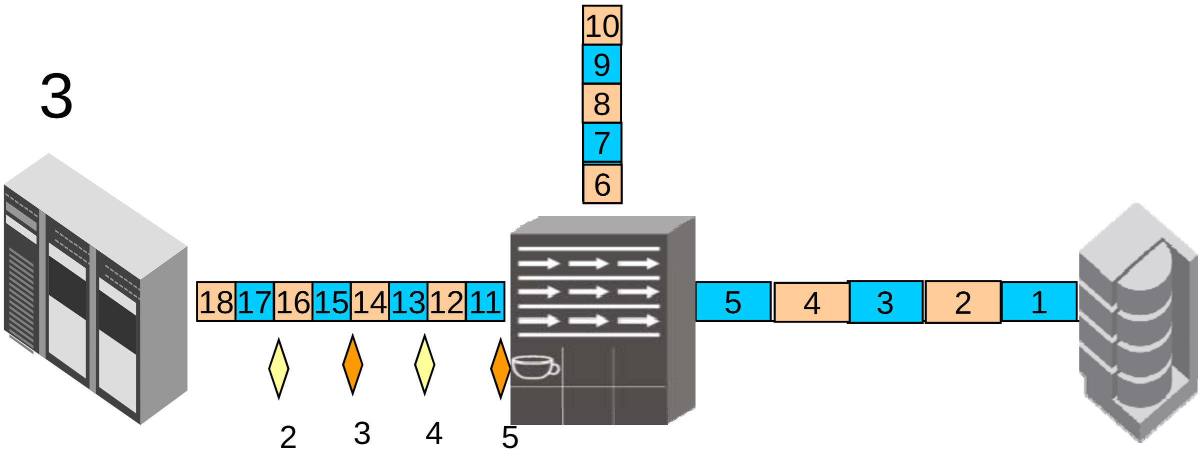
4



3



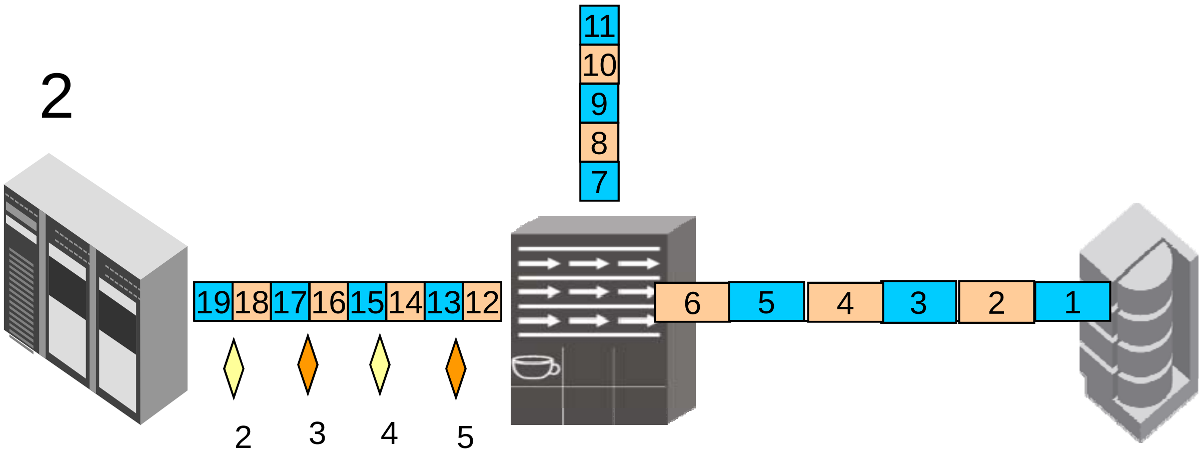
3



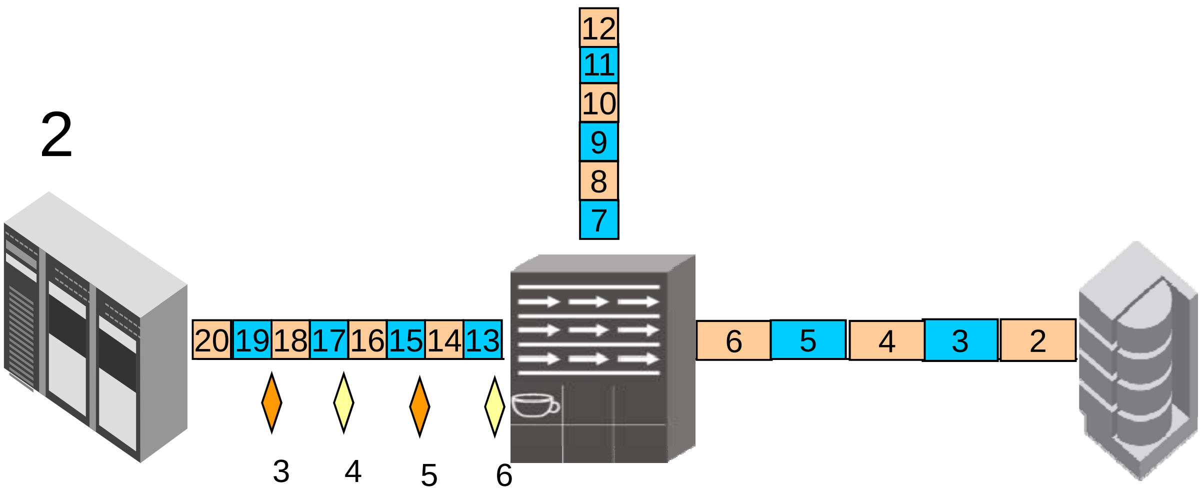
SHARE Technology · Connections · Results

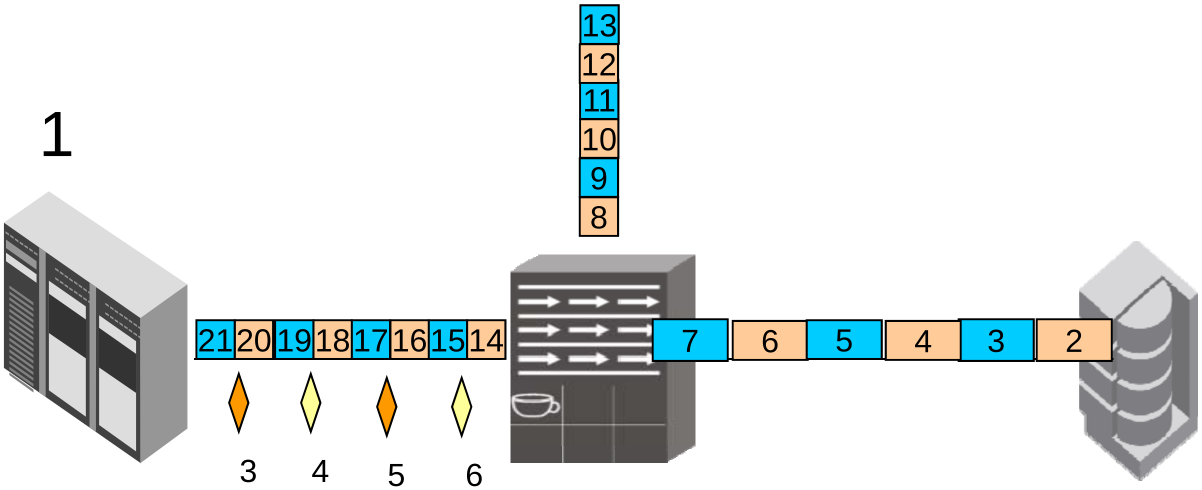
SHARE in Orlando 2011

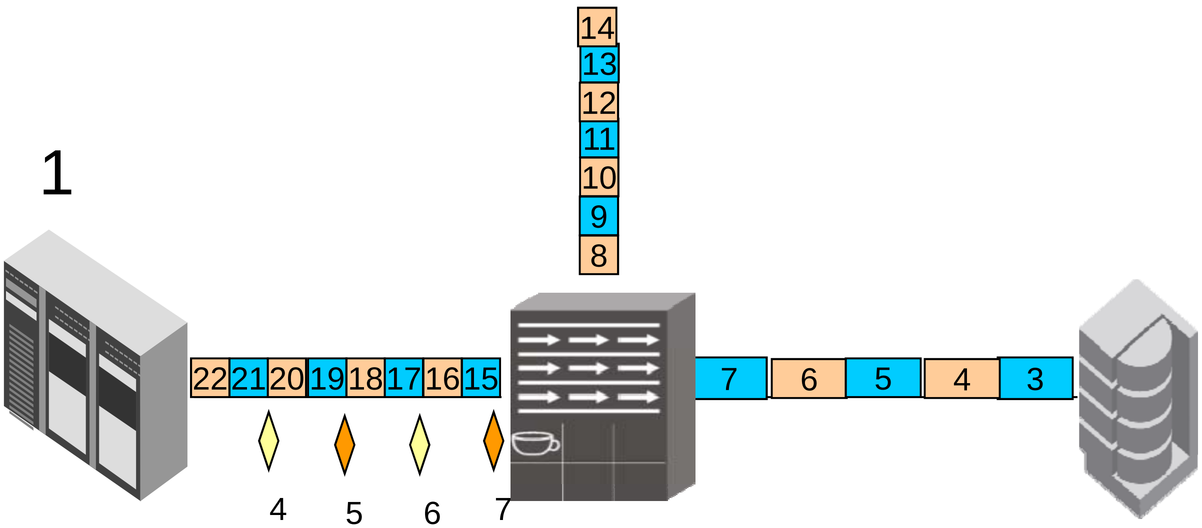
2

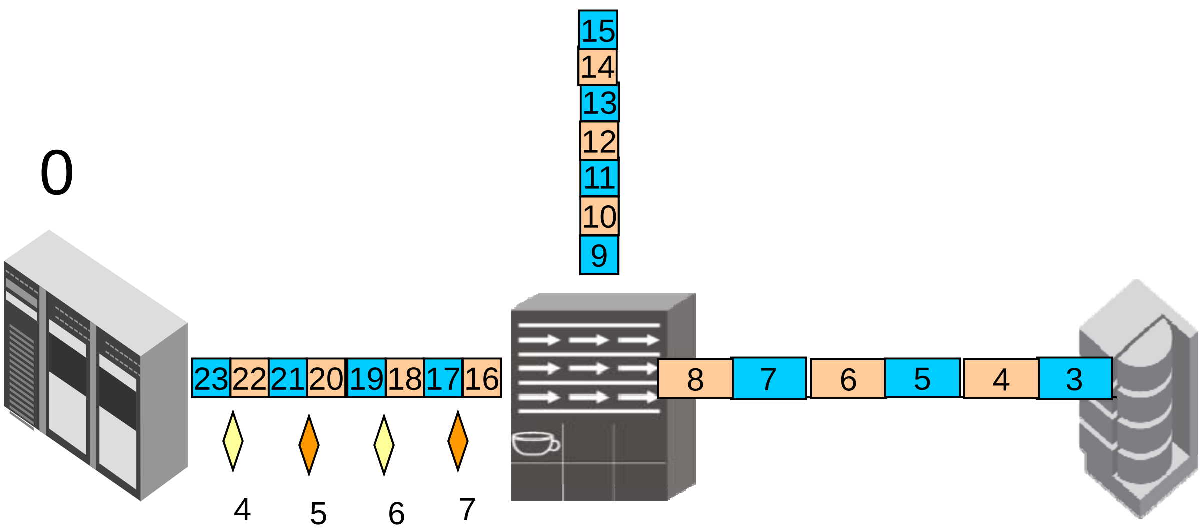


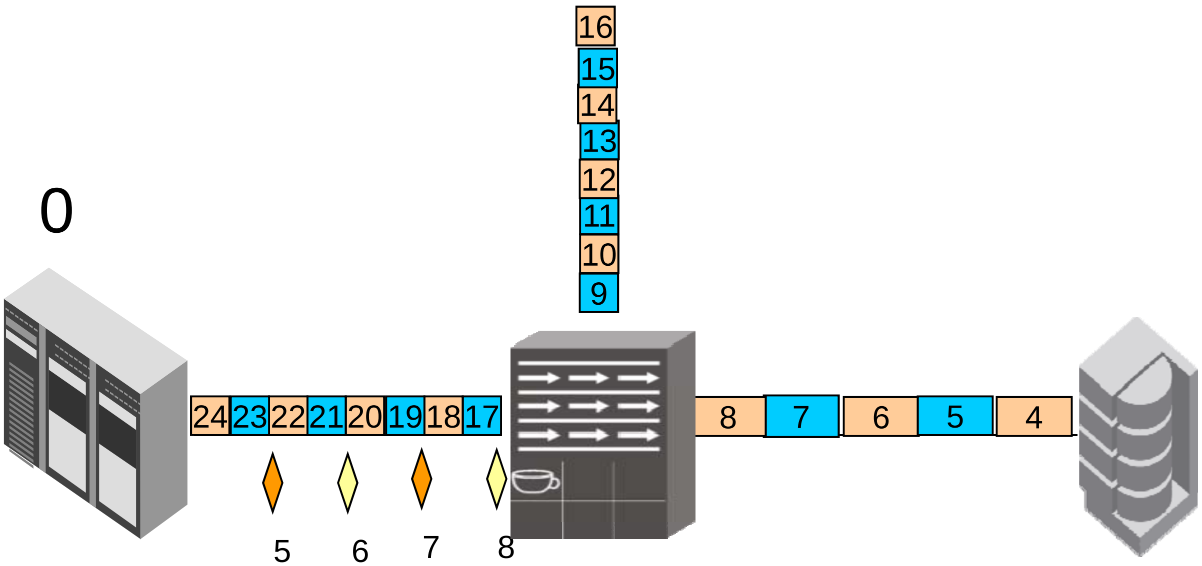
2

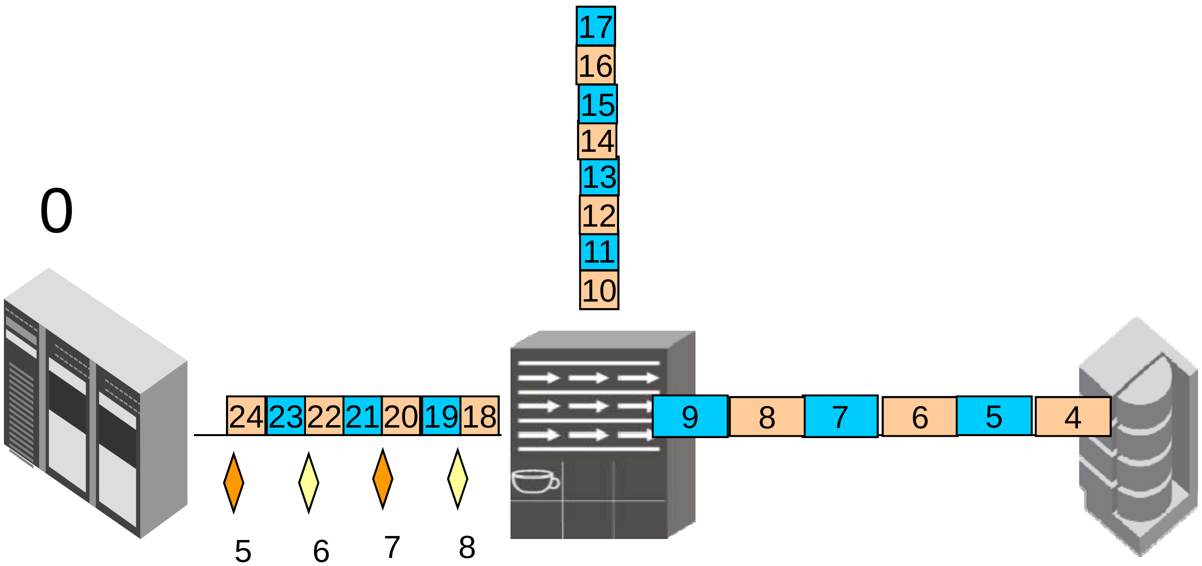


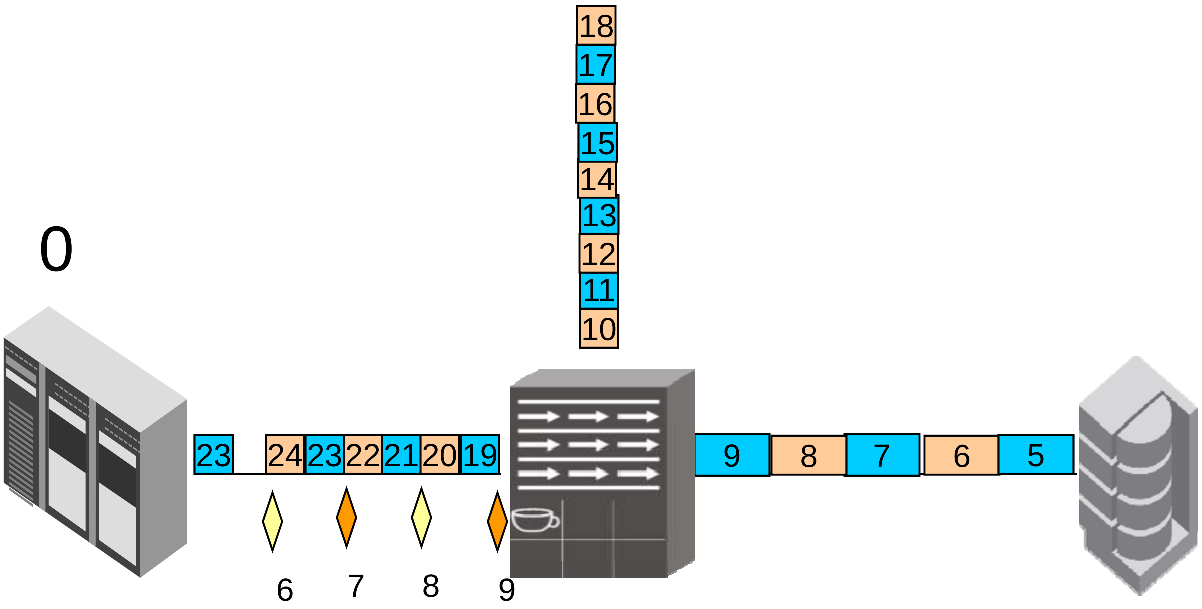

1
13
12
11
10
9
8
21 20 19 18 17 16 15 14
7 6 5 4 3 2
3 4 5 6


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

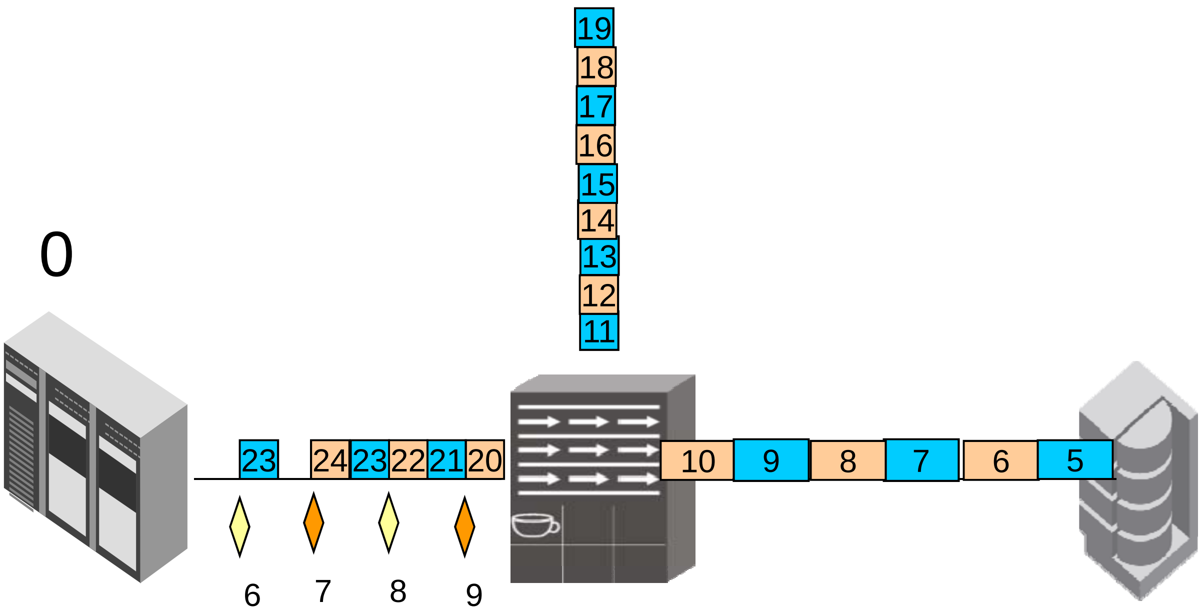

0
15
14
13
12
11
10
9
23 22 21 20 19 18 17 16
8 7 6 5 4 3
4
5
6
7


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

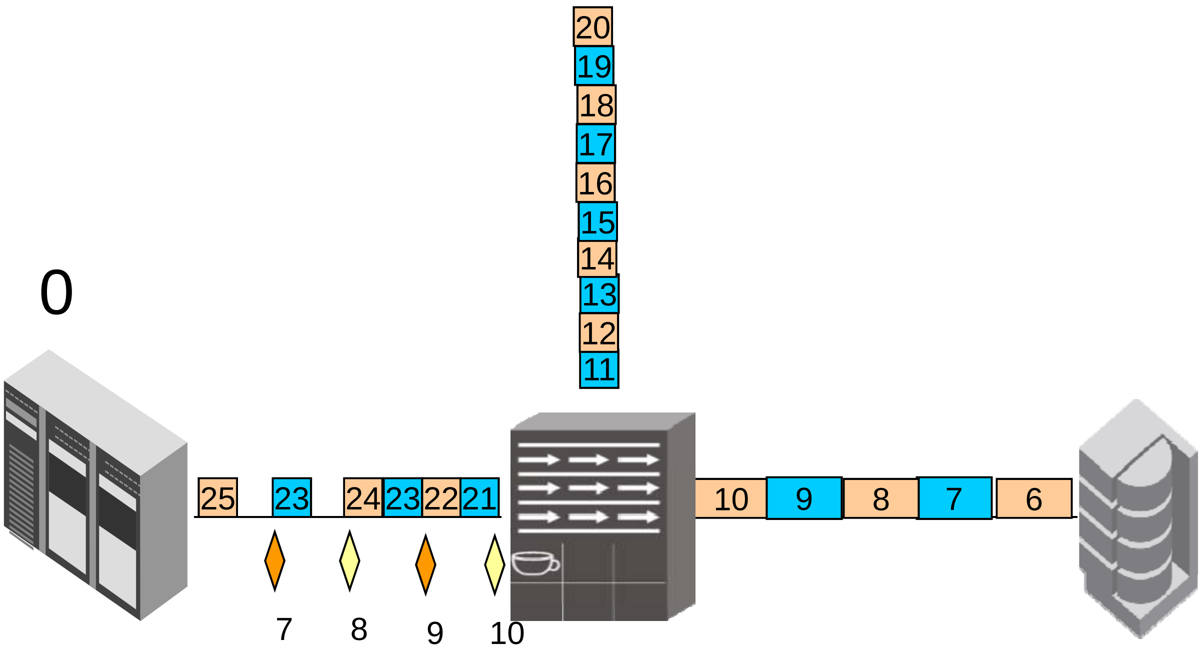

0
17
16
15
14
13
12
11
10
24 23 22 21 20 19 18
9 8 7 6 5 4
5 6 7 8



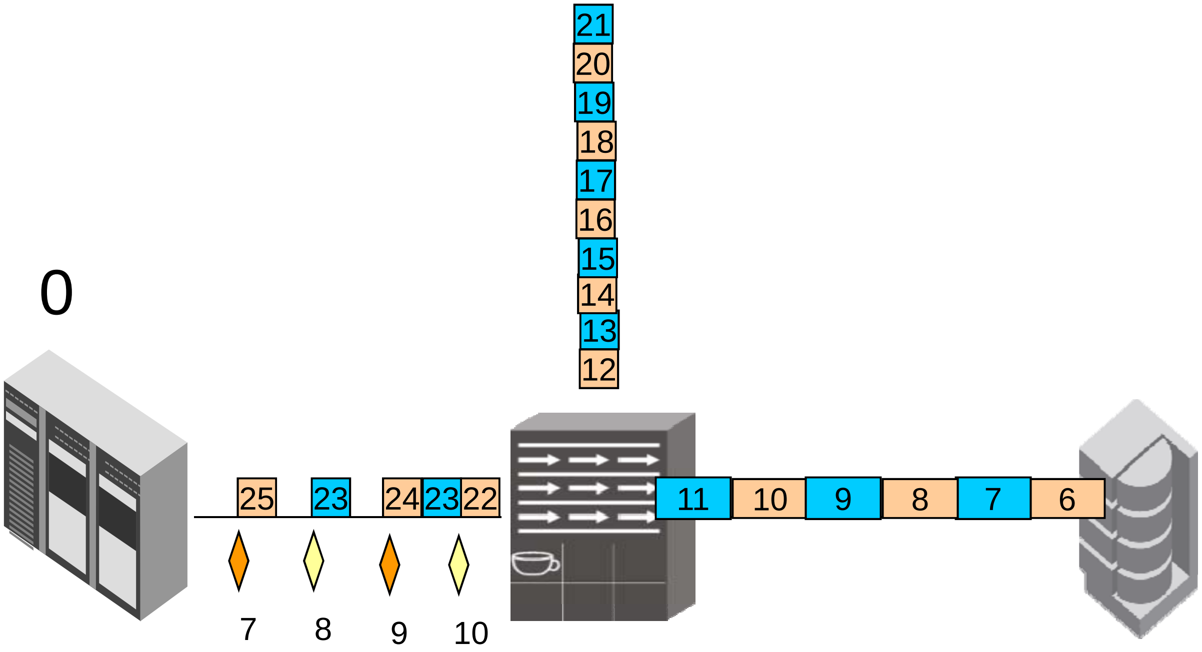
0
18
17
16
15
14
13
12
11
10
23
24
23
22
21
20
19
9
8
7
6
5
6
7
8
9



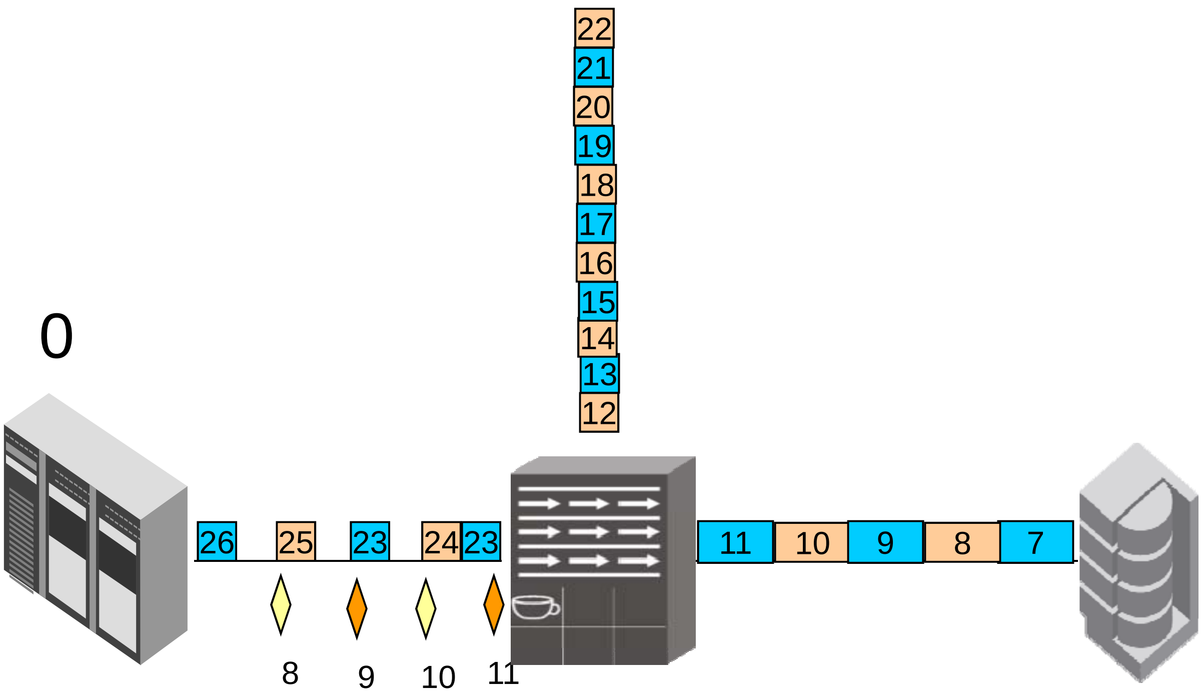
0
19
18
17
16
15
14
13
12
11
23
24
23
22
21
20
10
9
8
7
6
5
6
7
8
9

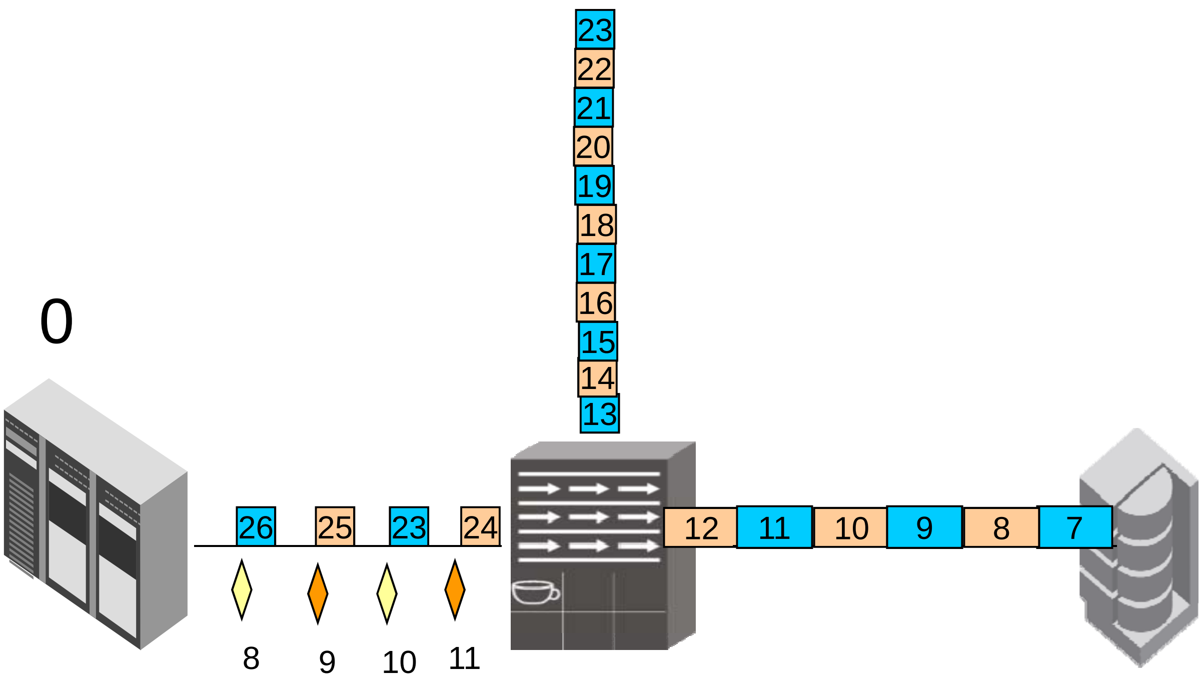

0
20
19
18
17
16
15
14
13
12
11
25
23
24
23
22
21
10
9
8
7
6
7
8
9
10

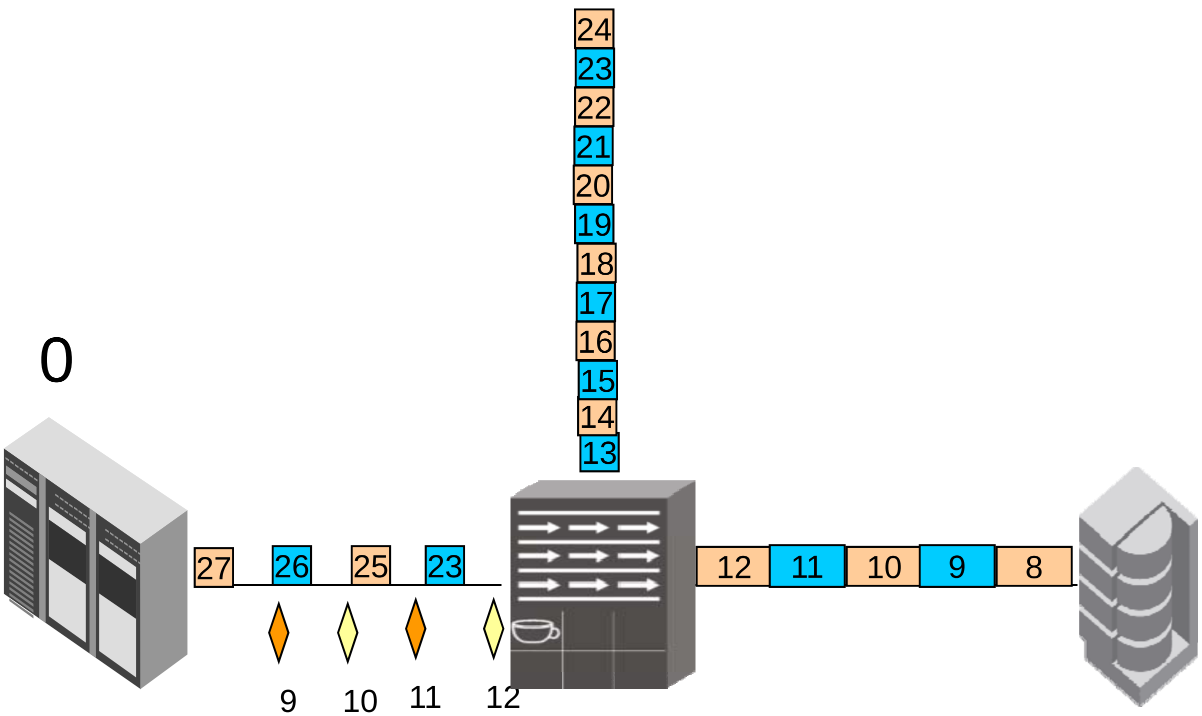
0



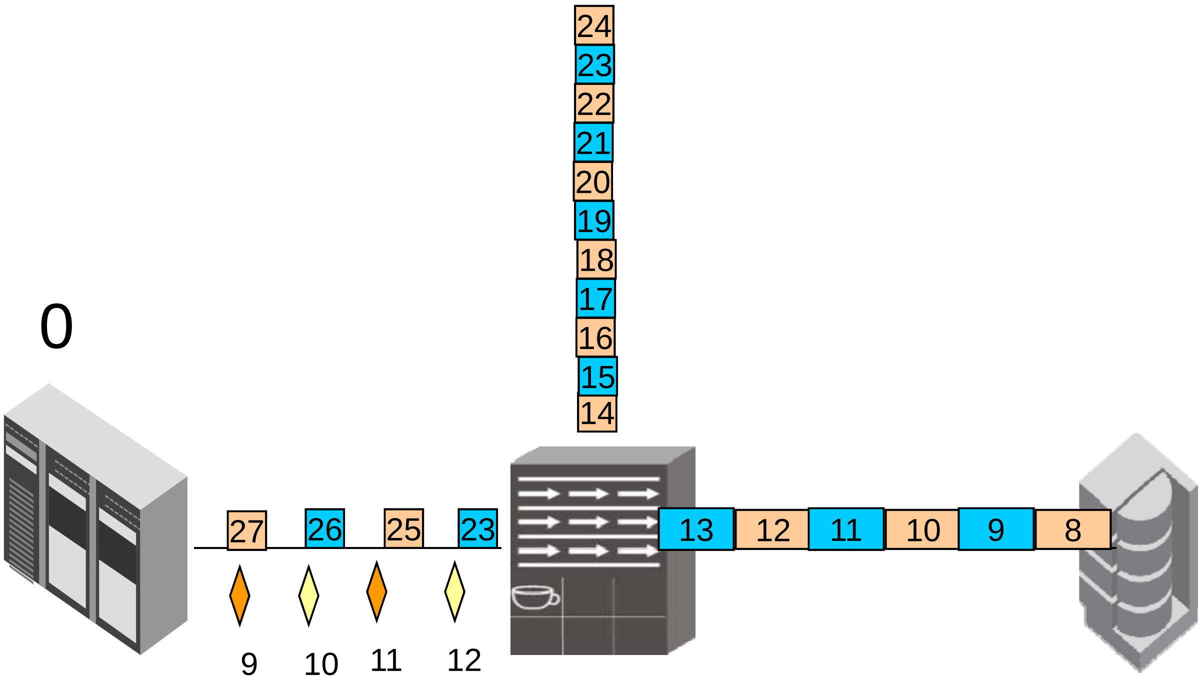
0




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

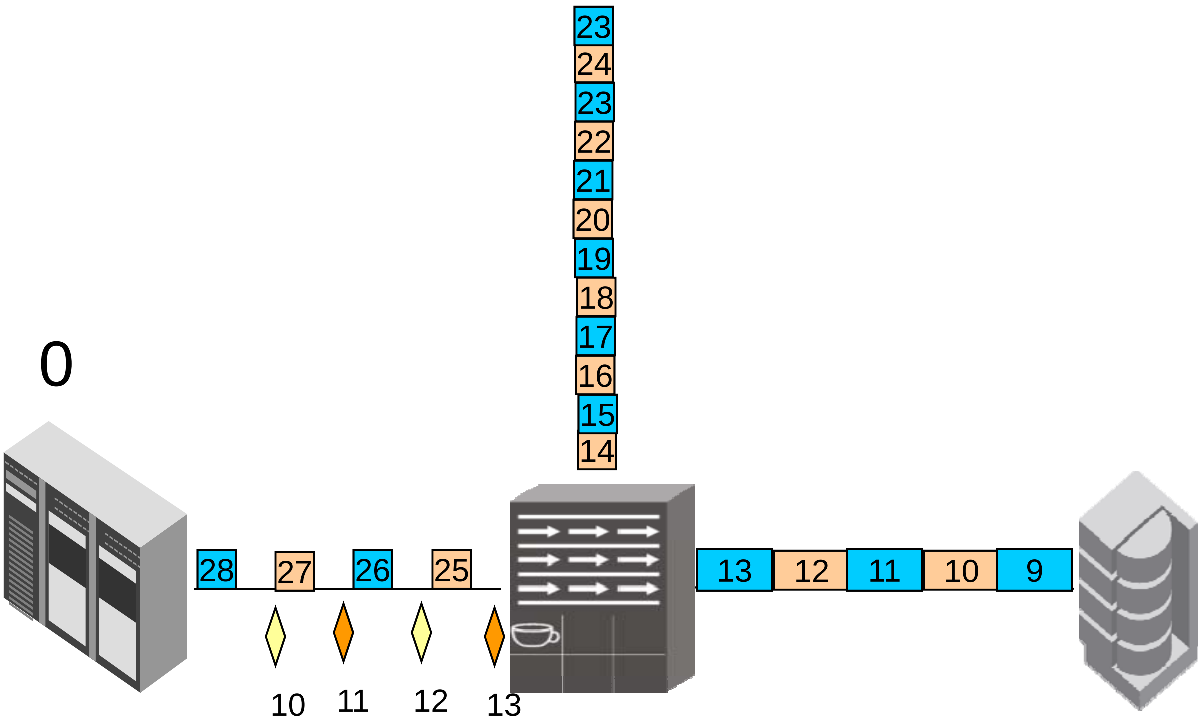


24
23
22
21
20
19
18
17
16
15
14
13
0
27
26
25
23
12
11
10
9
8
9
10
11
12

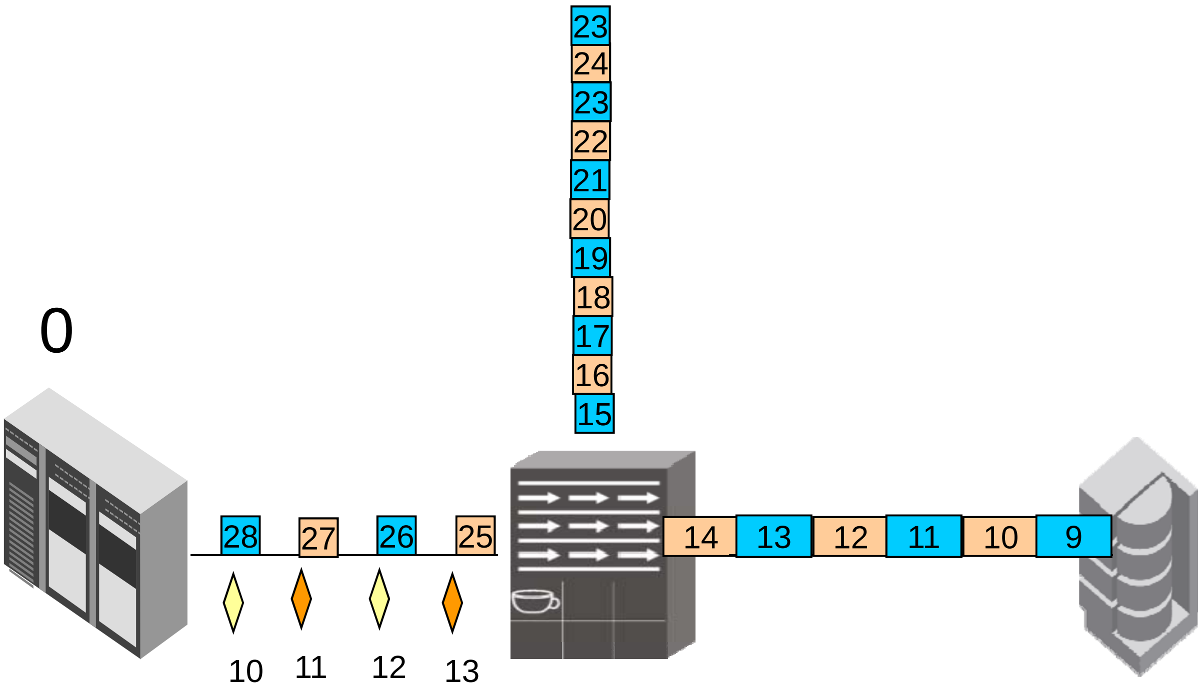


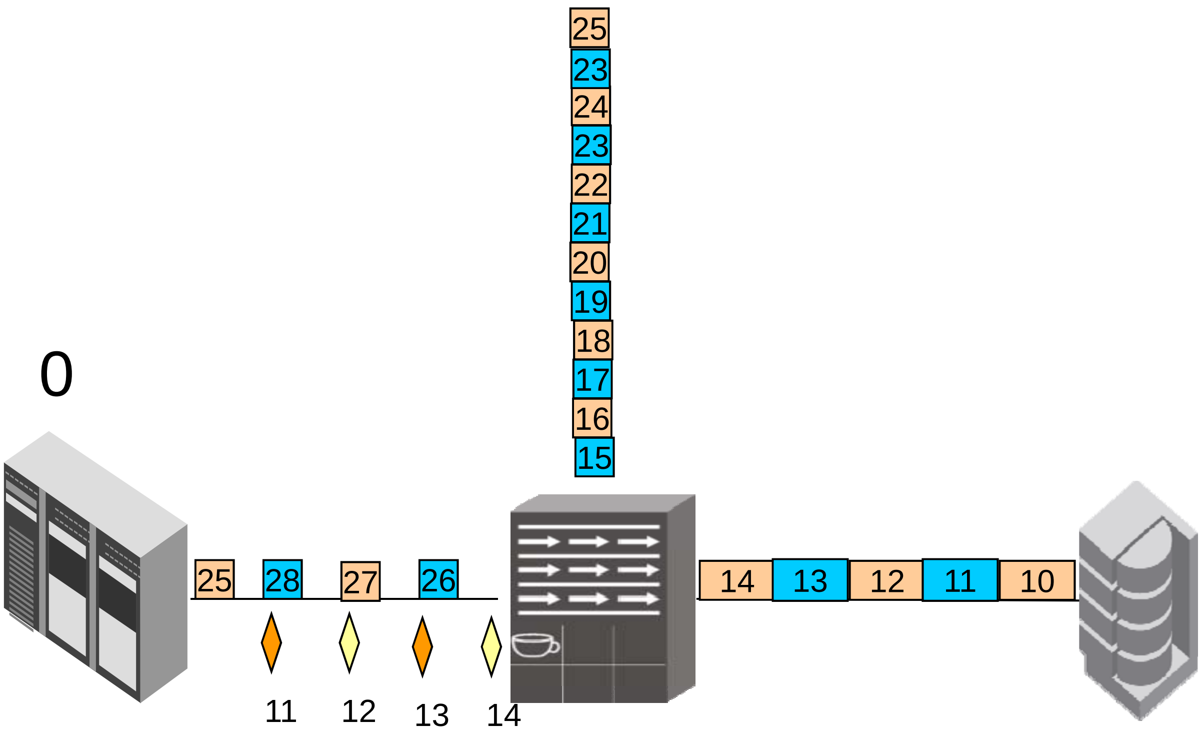
0
24
23
22
21
20
19
18
17
16
15
14
27
26
25
23
13
12
11
10
9
8
9
10
11
12

0



0




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

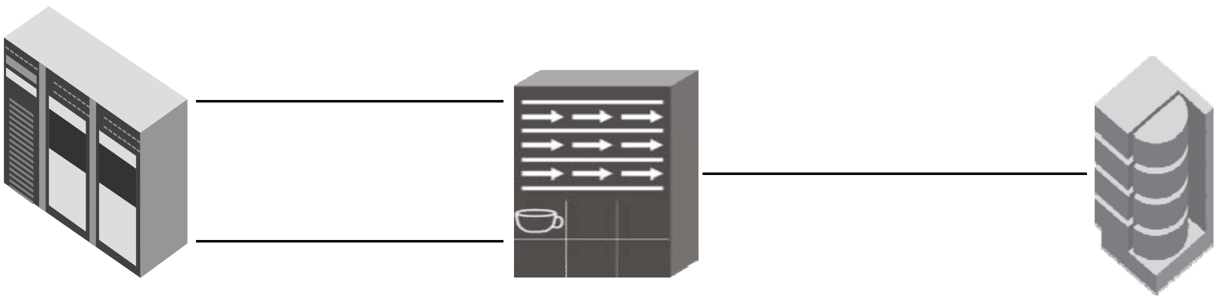
THIS PAGE INTENTIONALLY LEFT BLANK

Example: Real Life?

BUFFER CREDITS

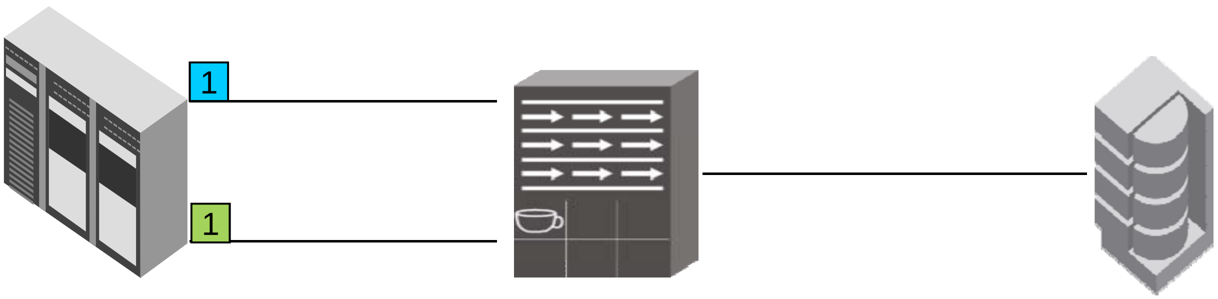
More real life example: Two senders sending at 30% - 50% link rate to one receiver

20



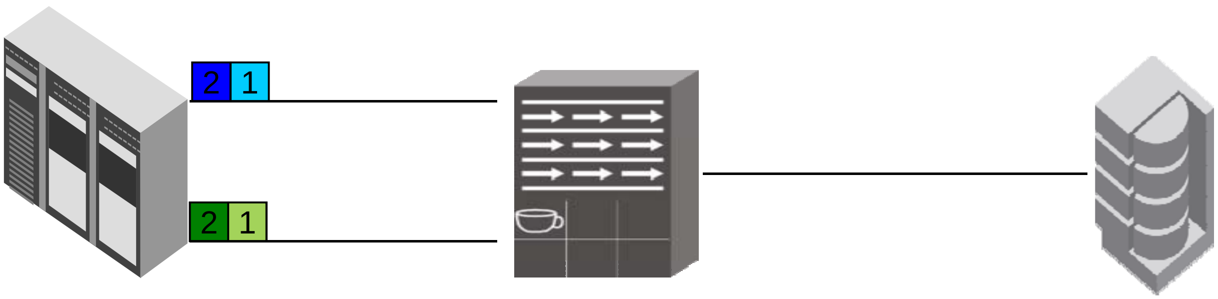
20

19



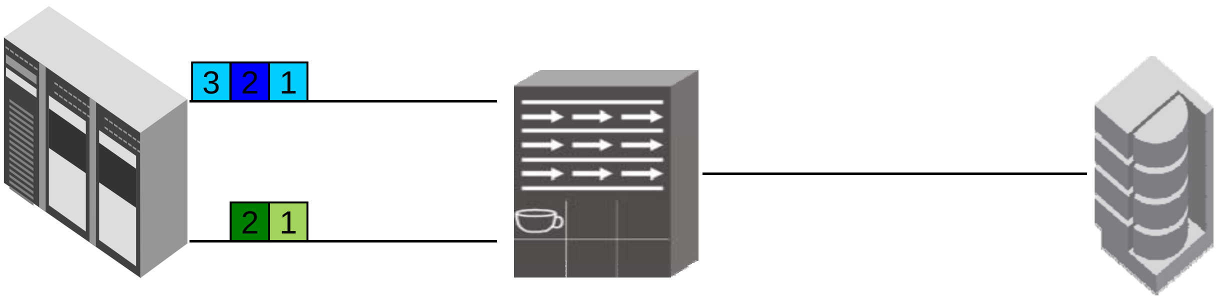
19

18



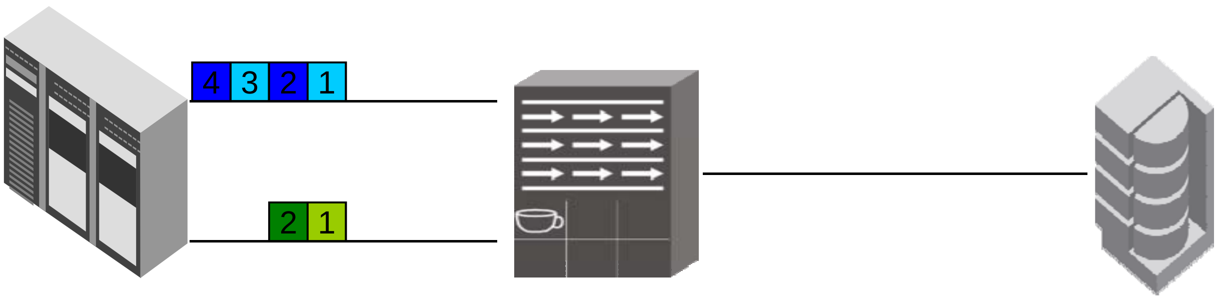
18

17



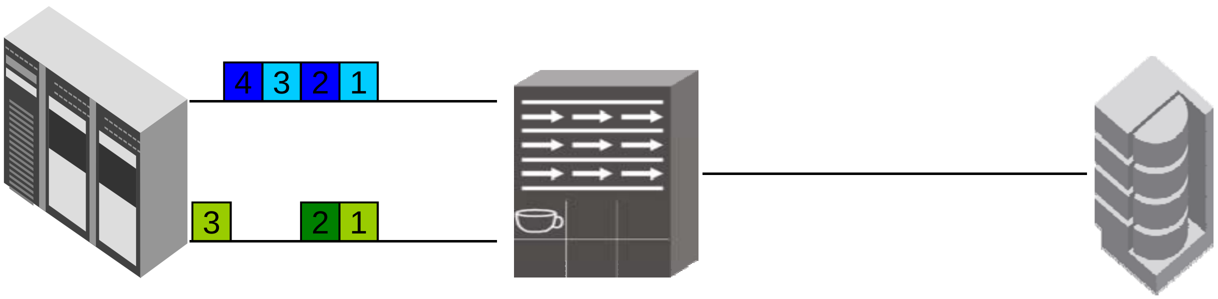
18

16



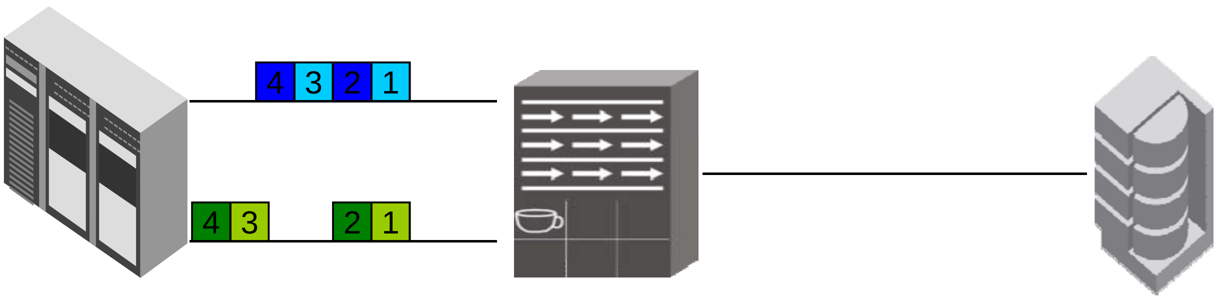
18

16



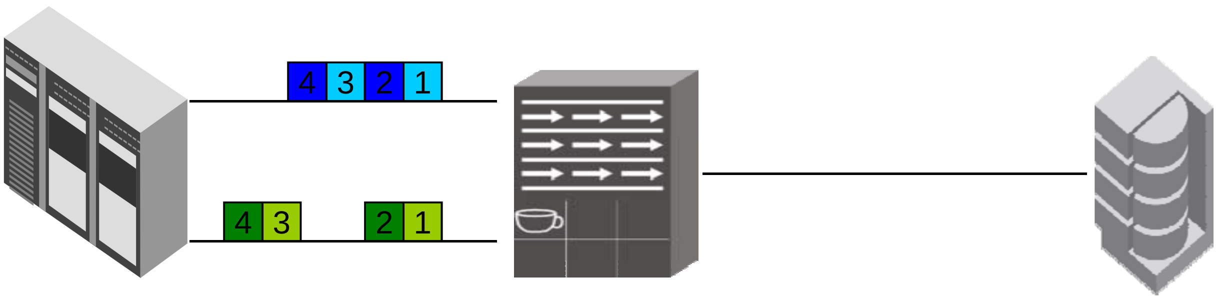
17

16



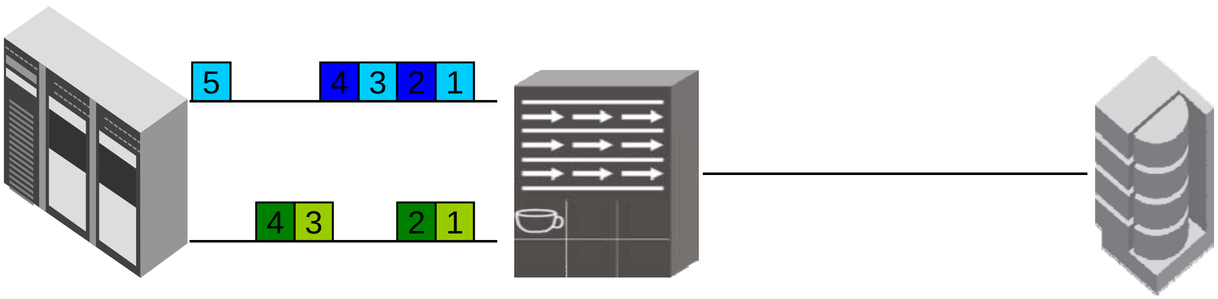
16

16



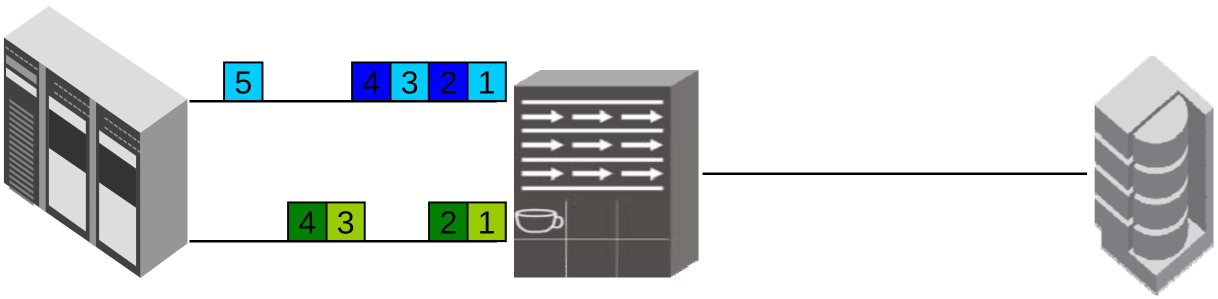
16

15

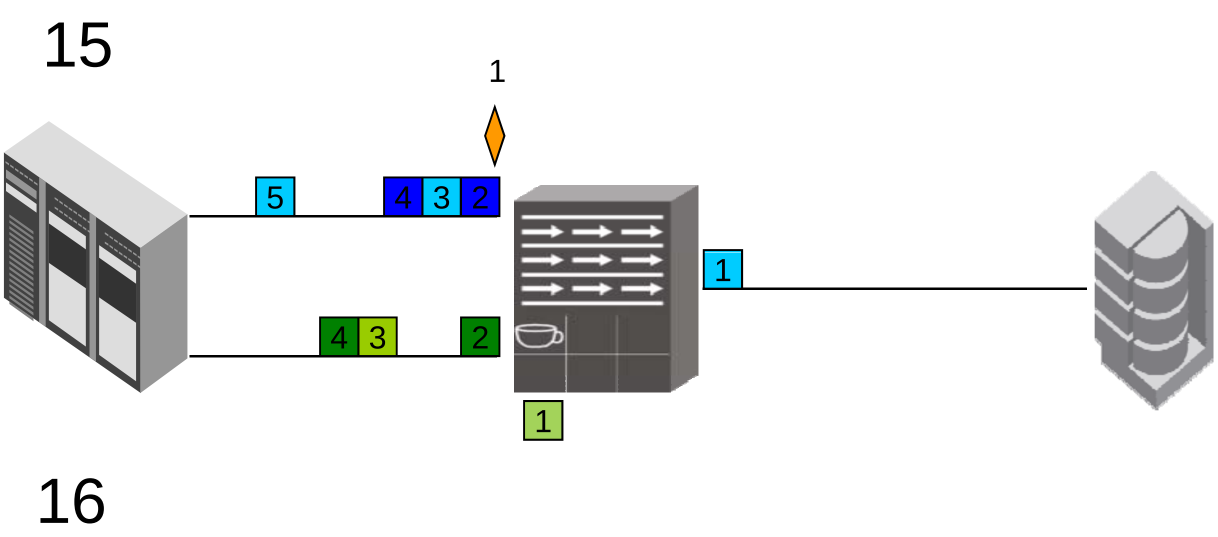


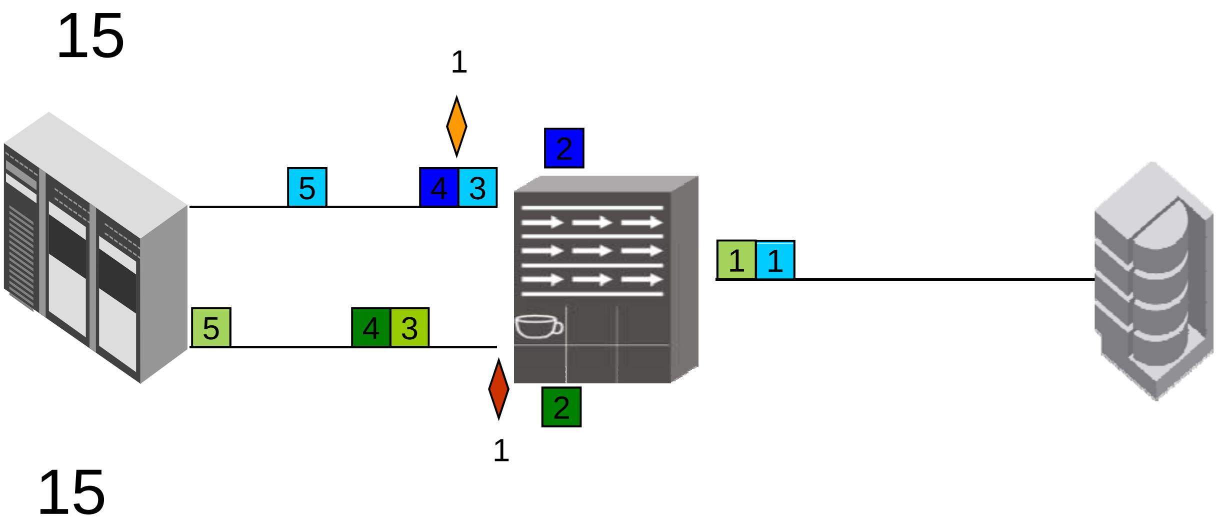
16

15

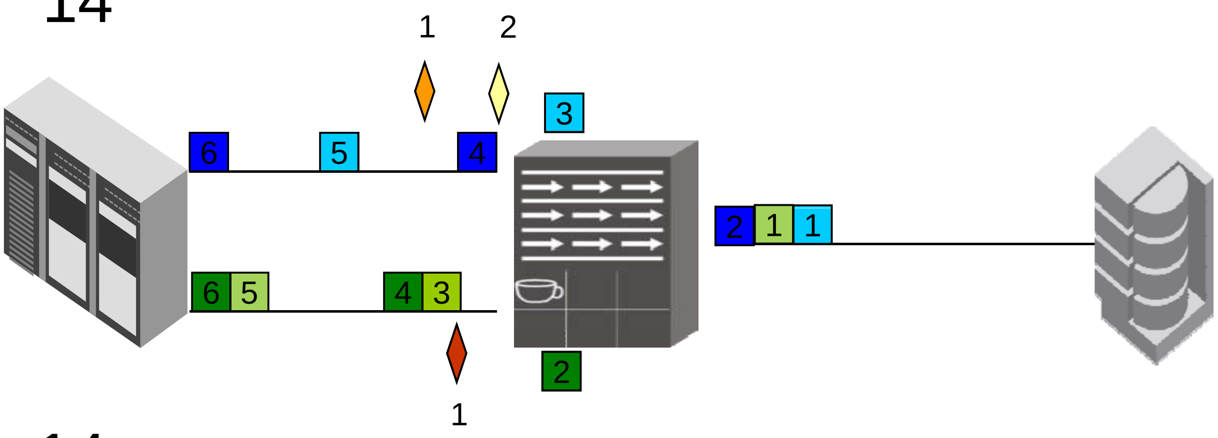


16

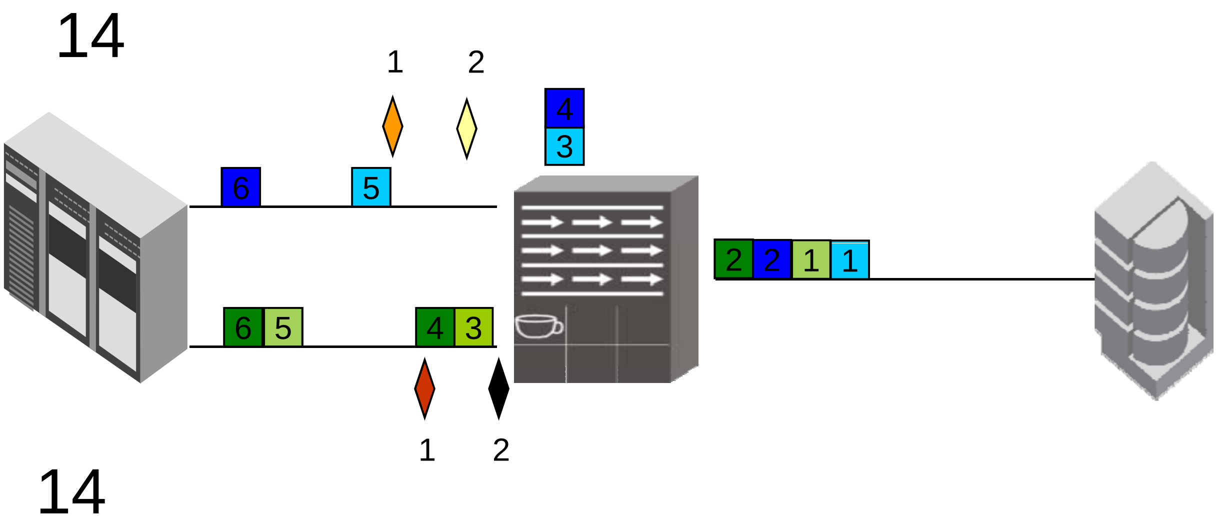

15
1
5
4
3
2
1
4
3
2
1
16

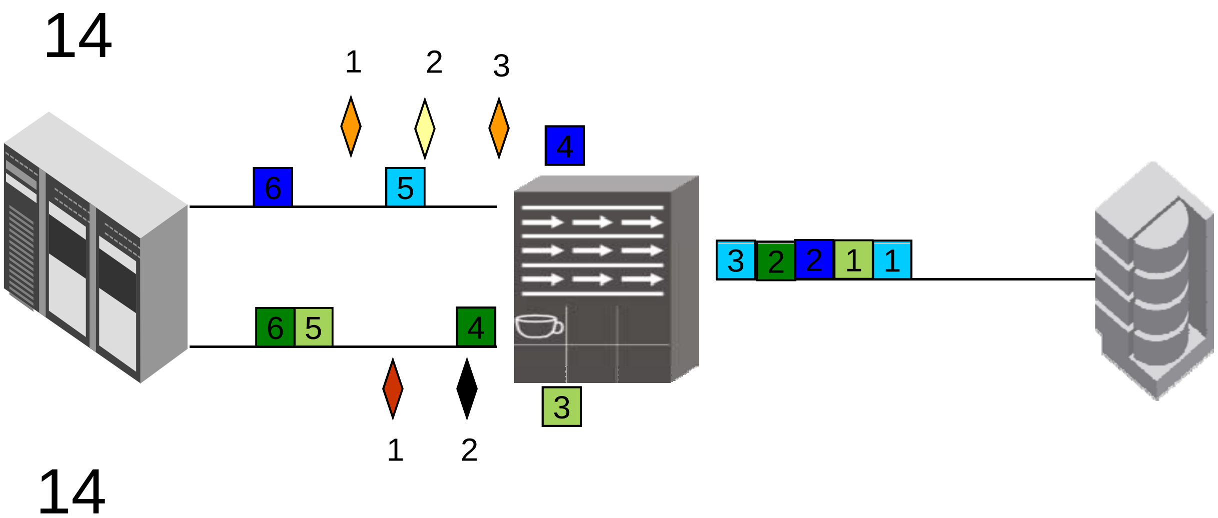

15
1
2
5
4
3
1
1
5
4
3
2
1
15

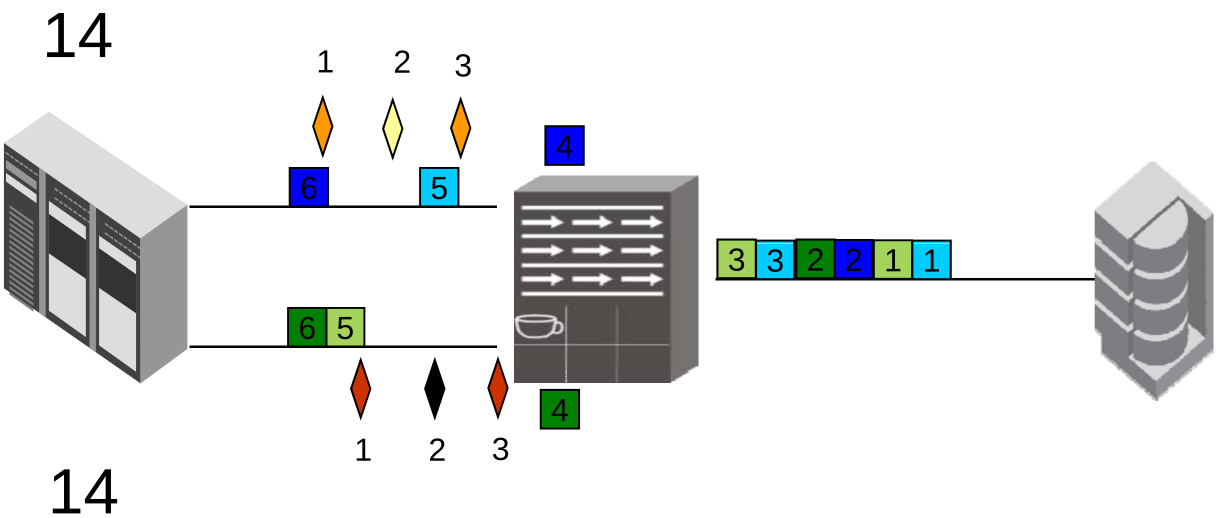
14

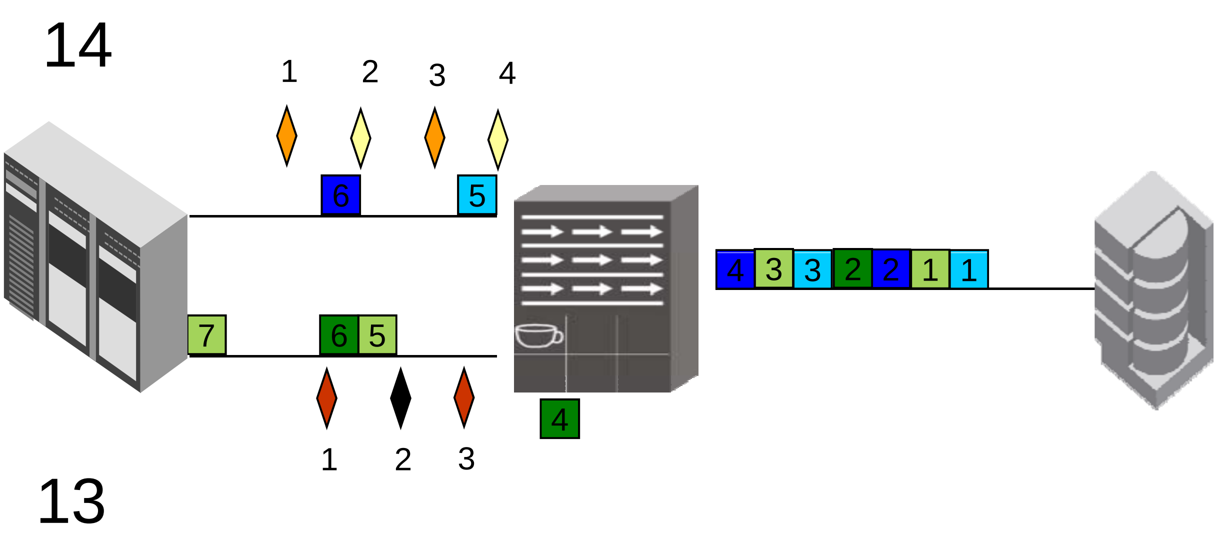


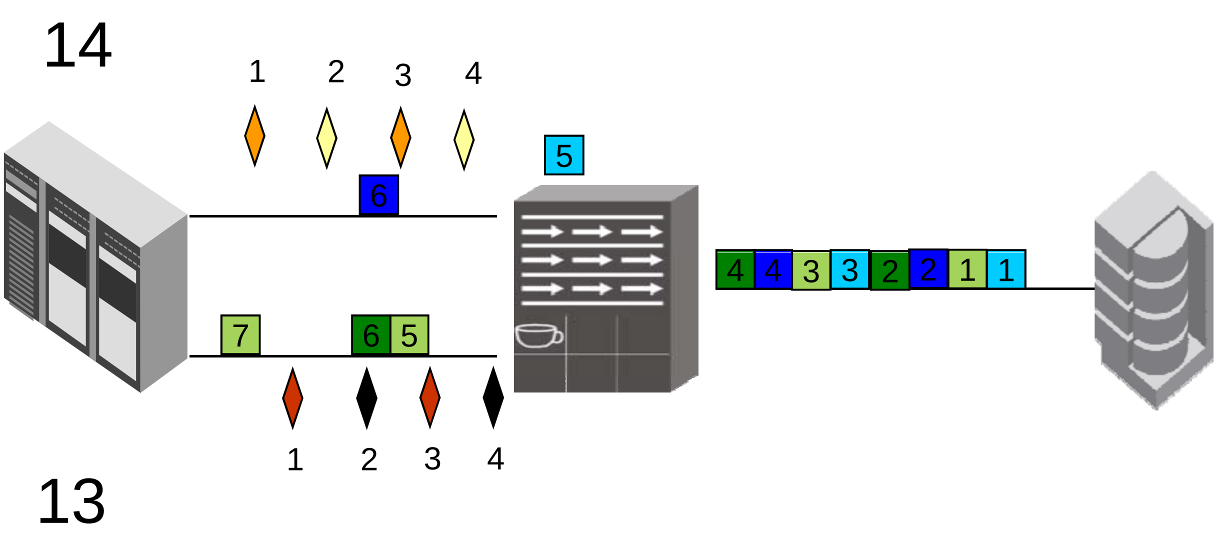
14


14
1
2
4
3
6
5
2
2
1
1
6
5
4
3
1
2
14

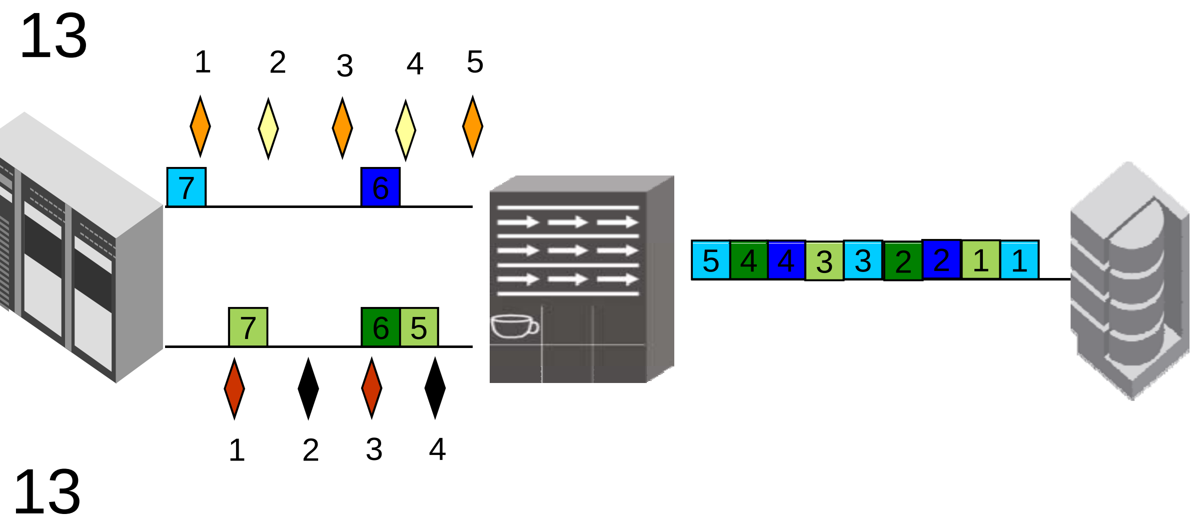

14
1
2
3
4
6
5
3
2
2
1
1
6
5
4
3
1
2
14

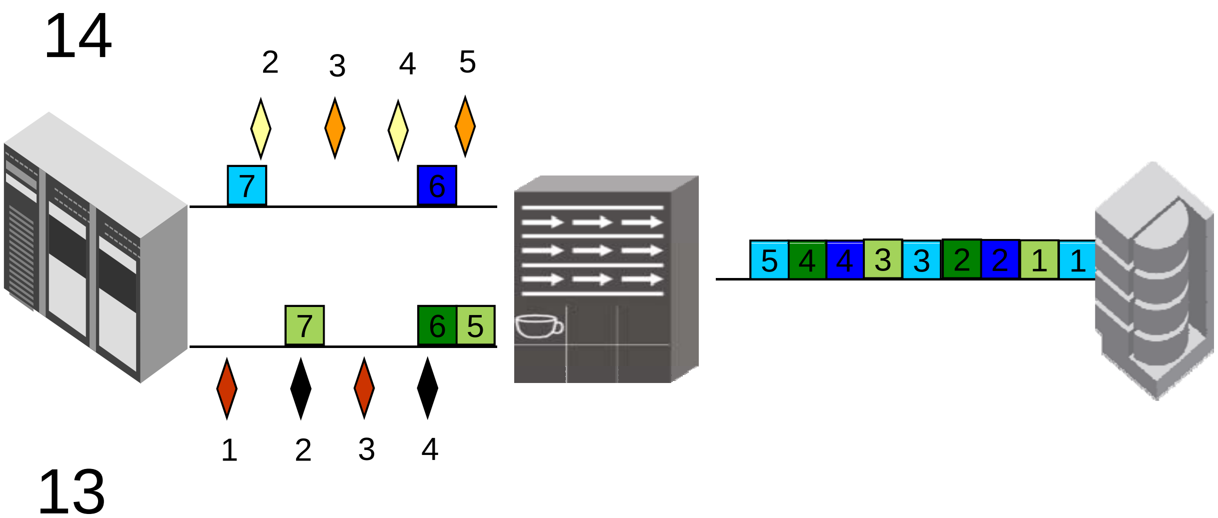

14
1
2
3
4
6
5
3
3
2
2
1
1
6
5
1
2
3
4
14

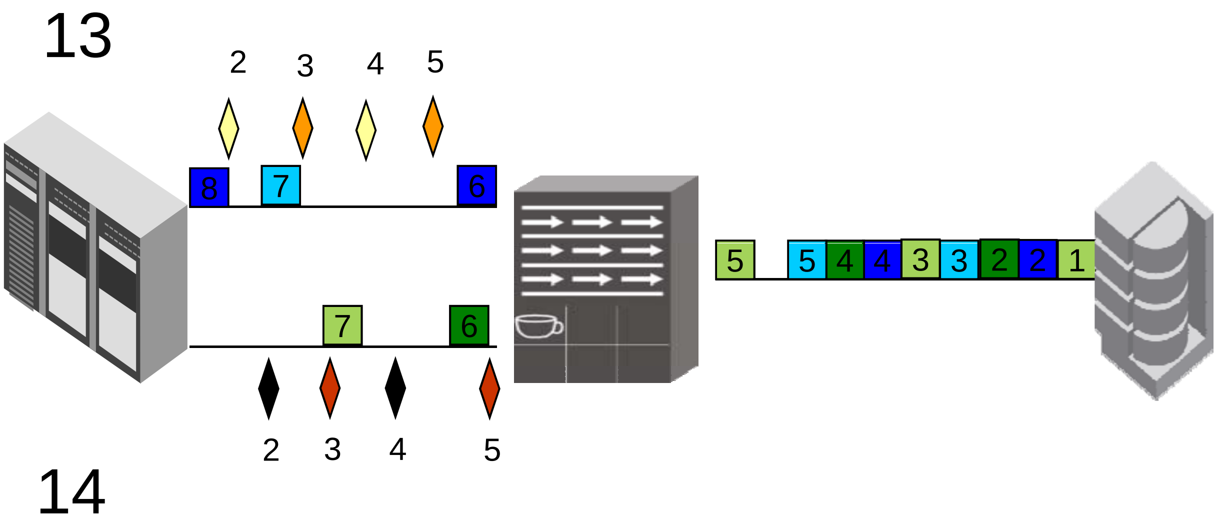

14
1
2
3
4
6
5
7
6
5
4
3
3
2
2
1
1
4
1
2
3
13



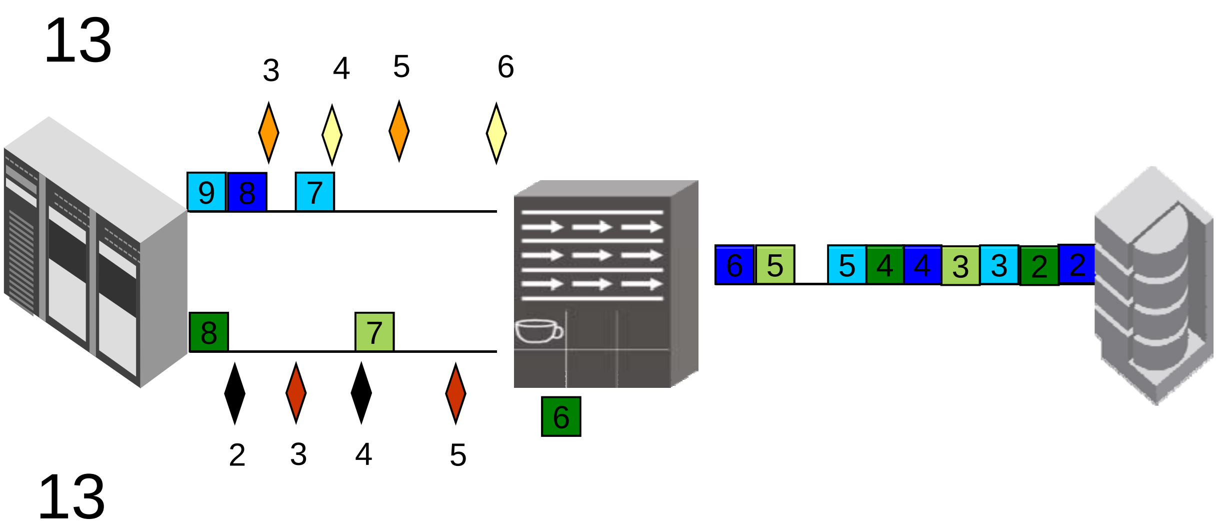
14
1
2
3
4
5
6
7
6
5
4
4
3
3
2
2
1
1
1
2
3
4
13

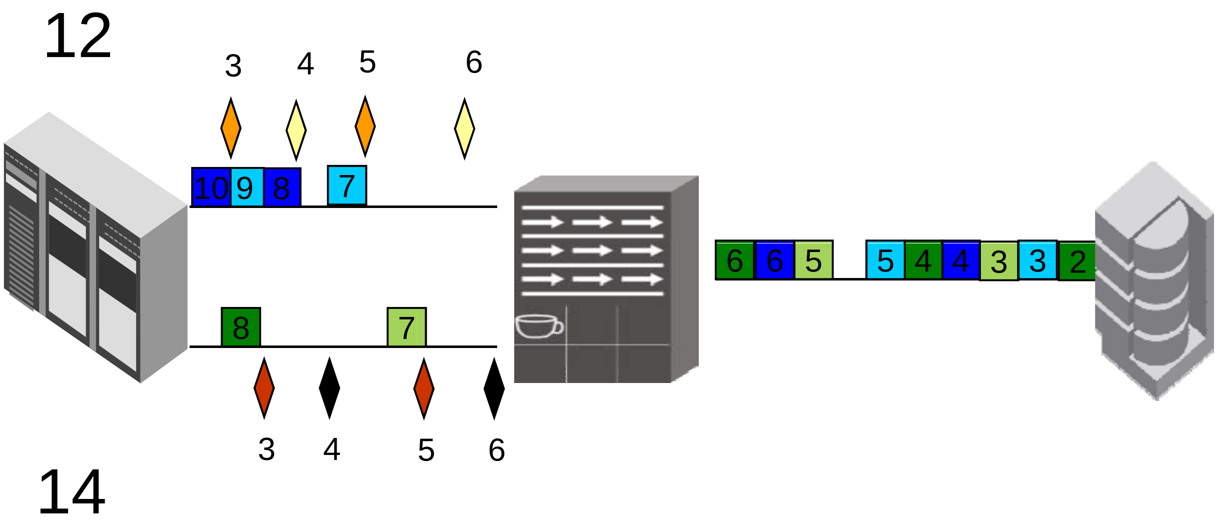

13
1
2
3
4
5
7
6
5
4
4
3
3
2
2
1
1
7
6
5
1
2
3
4
13

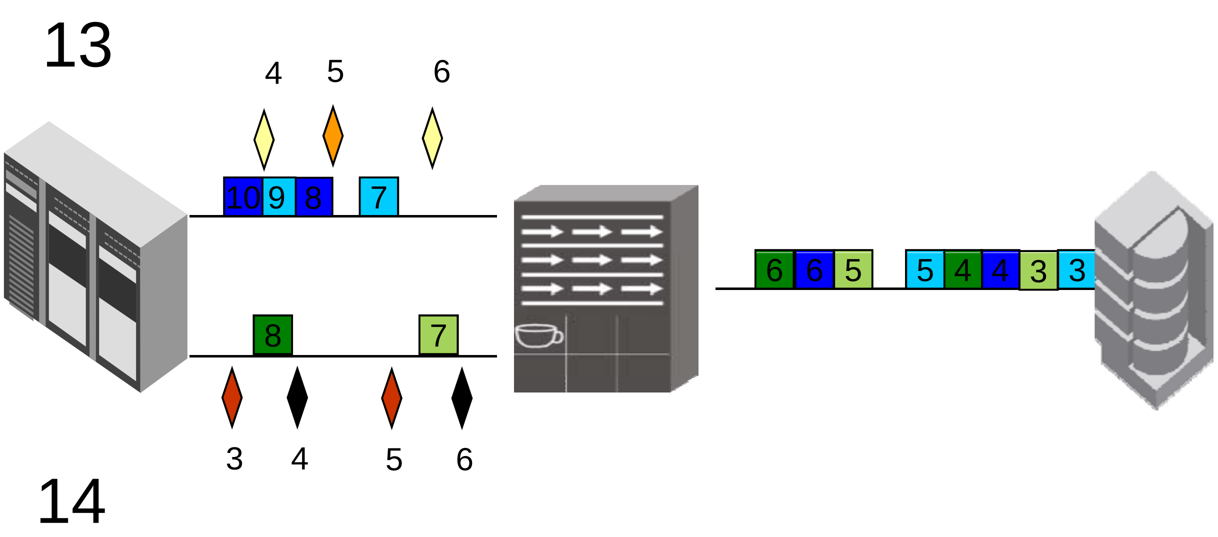

14
2
3
4
5
7
6
7
6
5
5
4
4
3
3
2
2
1
1
1
2
3
4
13

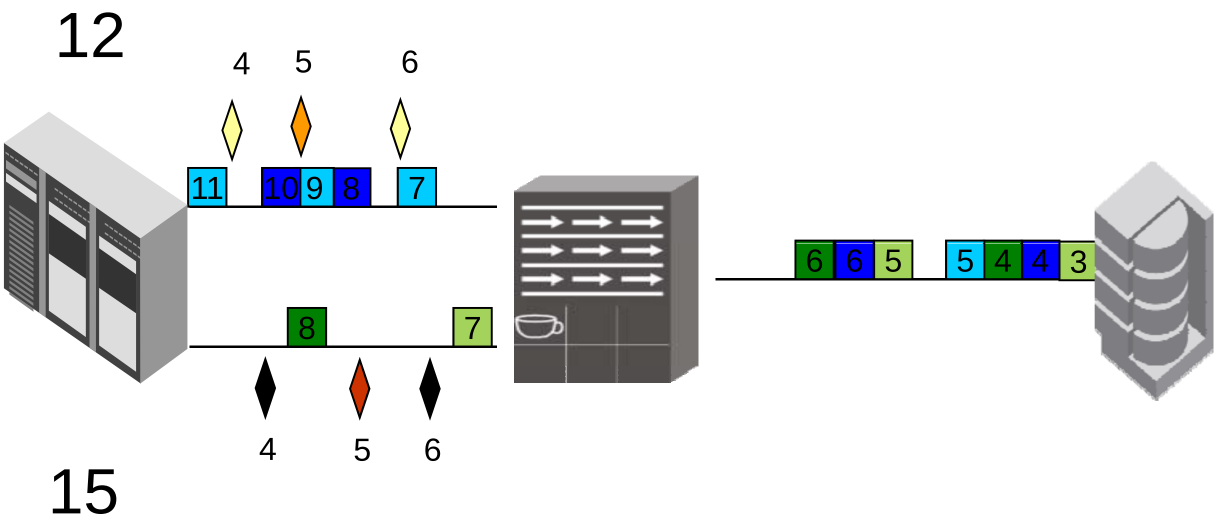


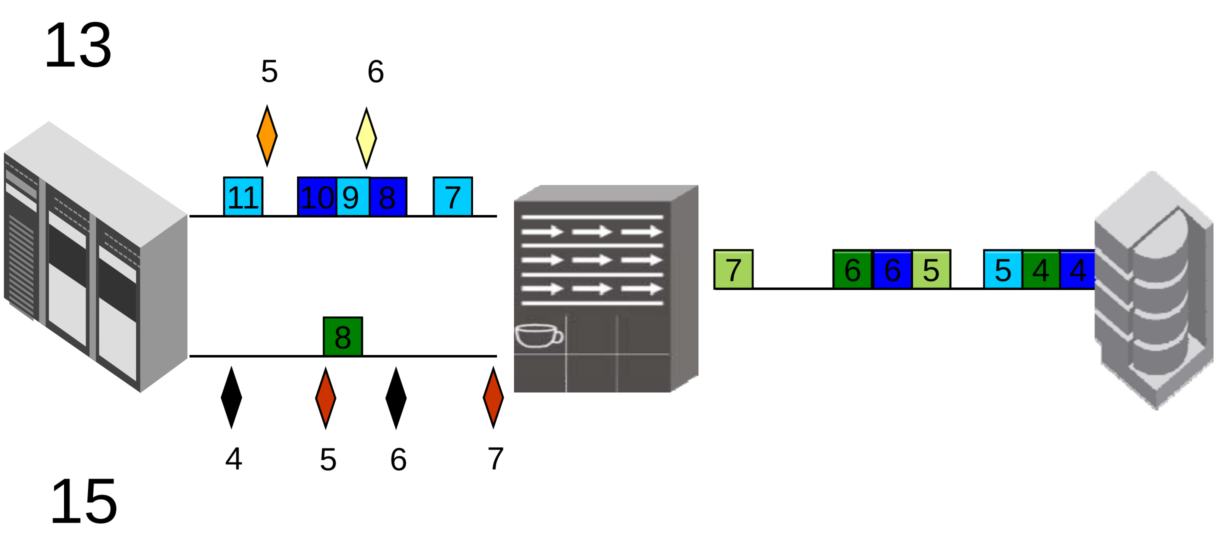
13
2
3
4
5
8
7
6
7
6
5
5
4
4
3
3
2
2
1
2
3
4
5
14


13
3
4
5
6
9
8
7
8
7
6
5
5
4
4
3
3
2
2
6
2
3
4
5
13

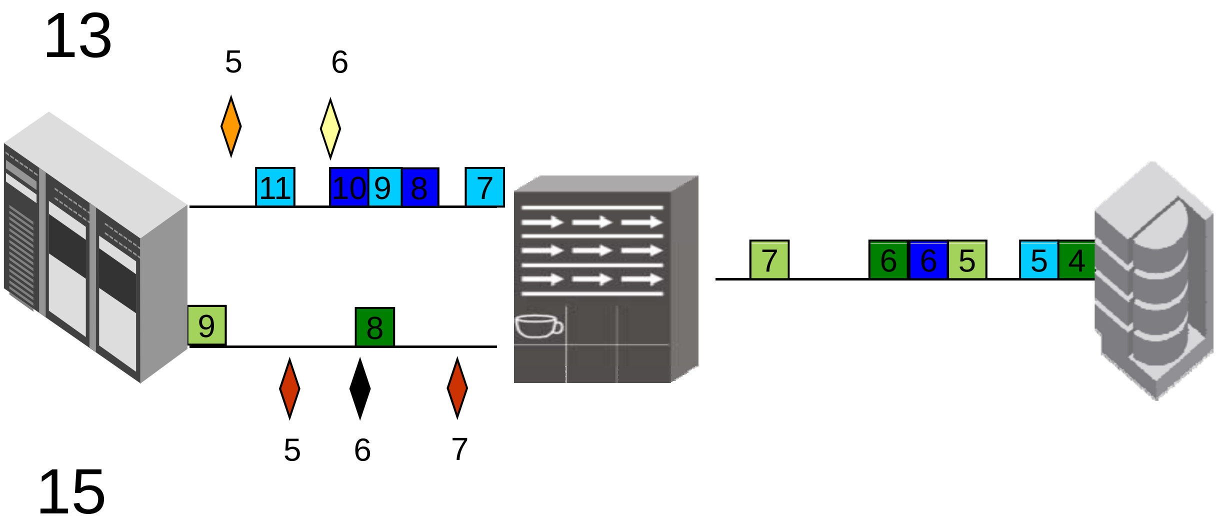

12
3
4
5
6
10
9
8
7
6
6
5
5
4
4
3
3
2
8
7
3
4
5
6
14

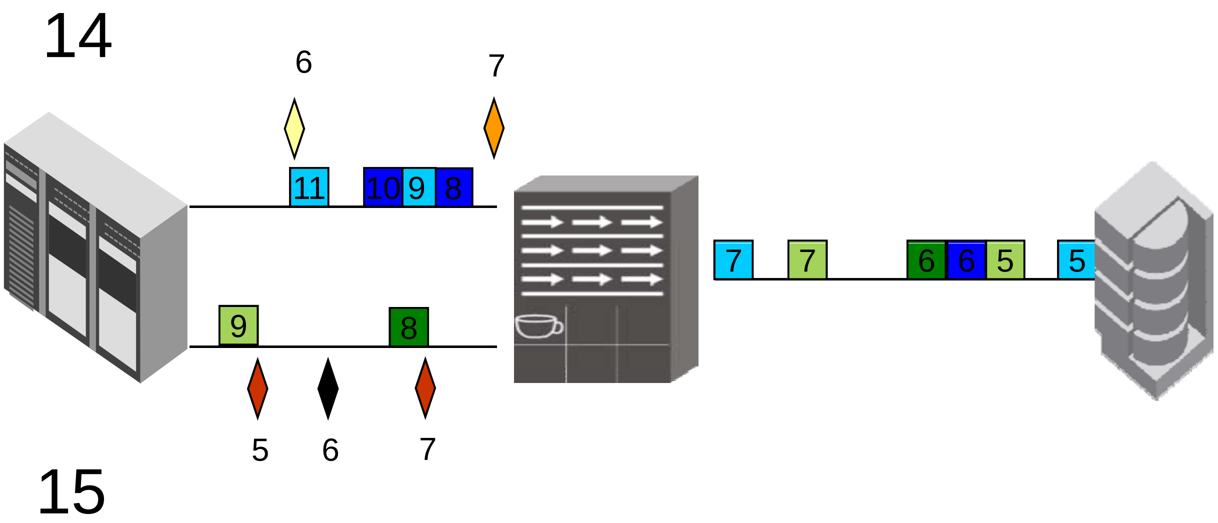

13
4
5
6
10
9
8
7
6
6
5
5
4
4
3
3
8
7
3
4
5
6
14


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

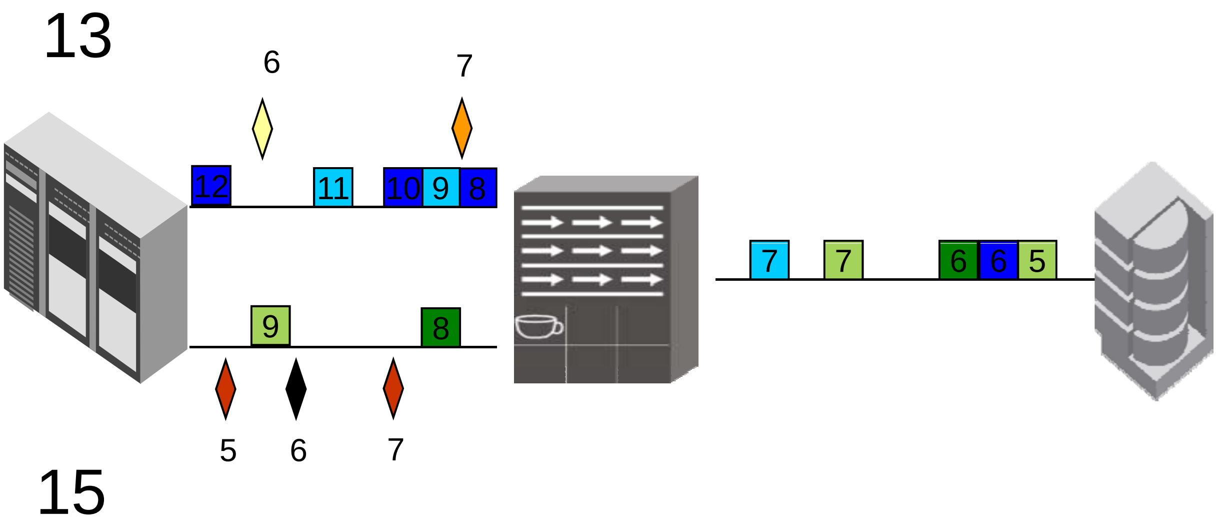


13
5
6
11
10
9
8
7
7
6
6
5
5
4
4
8
4
5
6
7
15

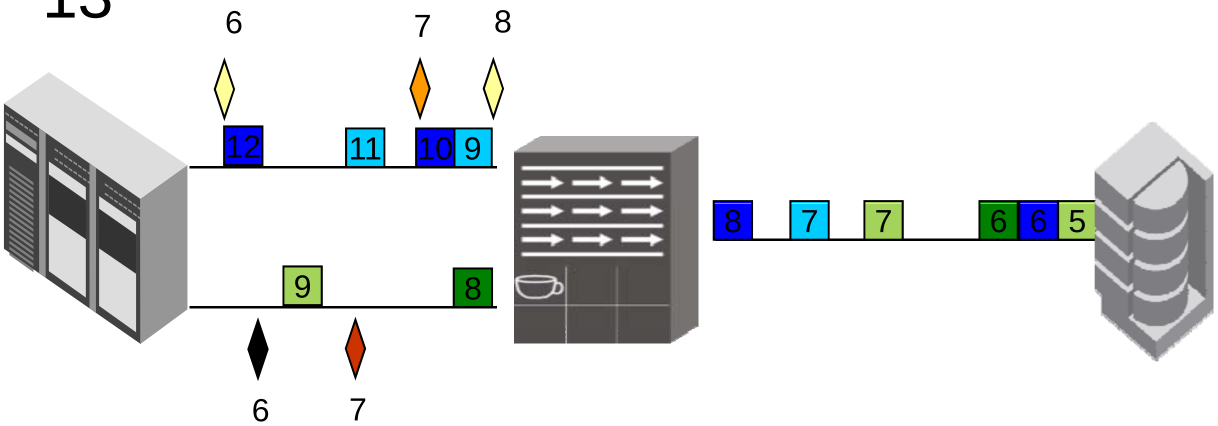

13
5
6
11
10
9
8
7
9
8
5
6
7
15
7
6
6
5
5
4



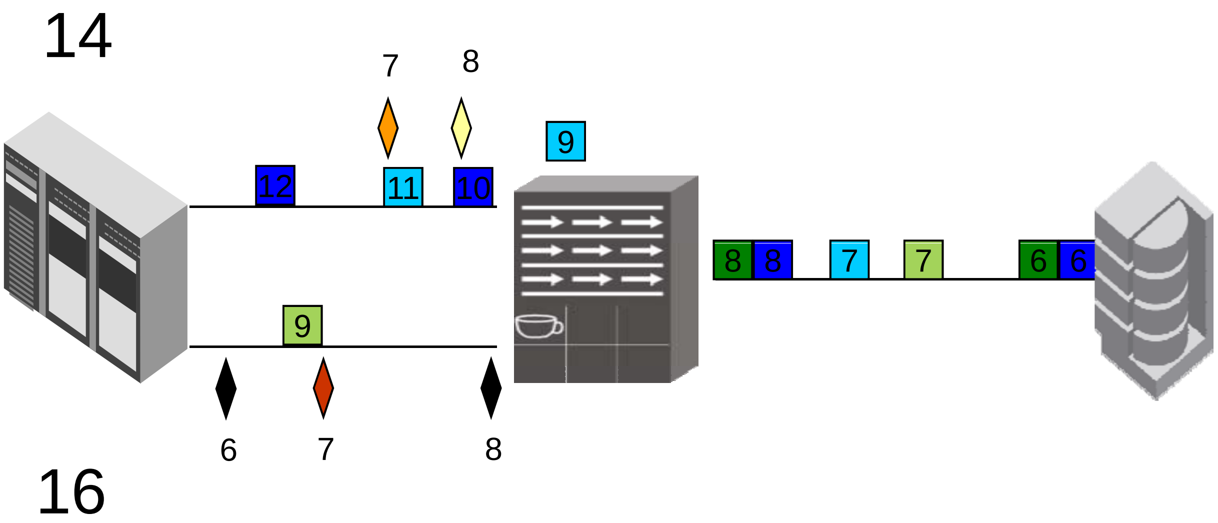
14
6
7
11
10
9
8
7
7
6
6
5
5
9
8
5
6
7
15


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

13



16

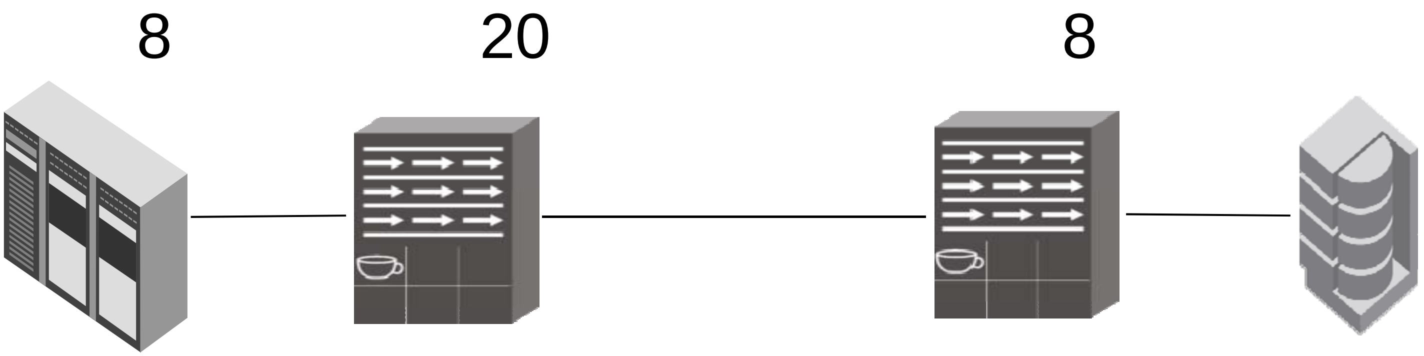


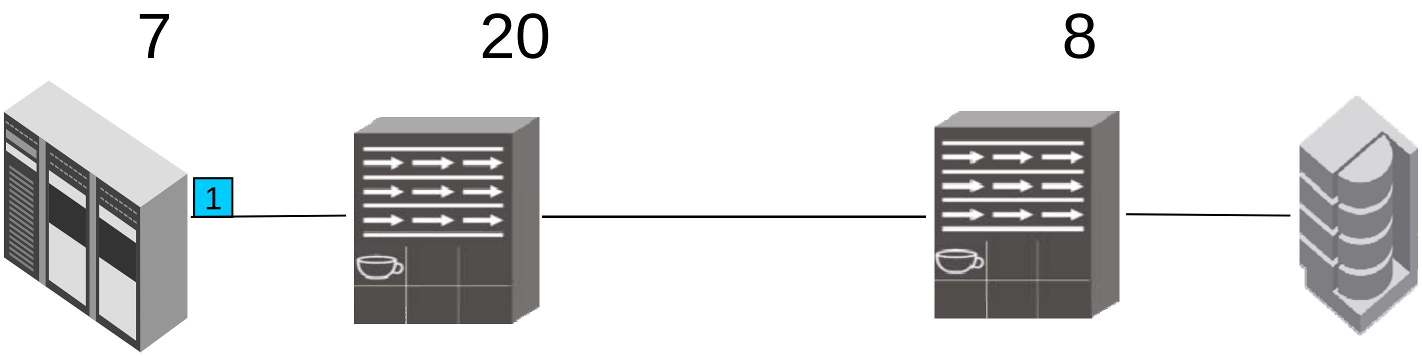
14
7
8
9
12
11
10
8
8
7
7
6
6
9
6
7
8
16

THIS PAGE INTENTIONALLY LEFT BLANK

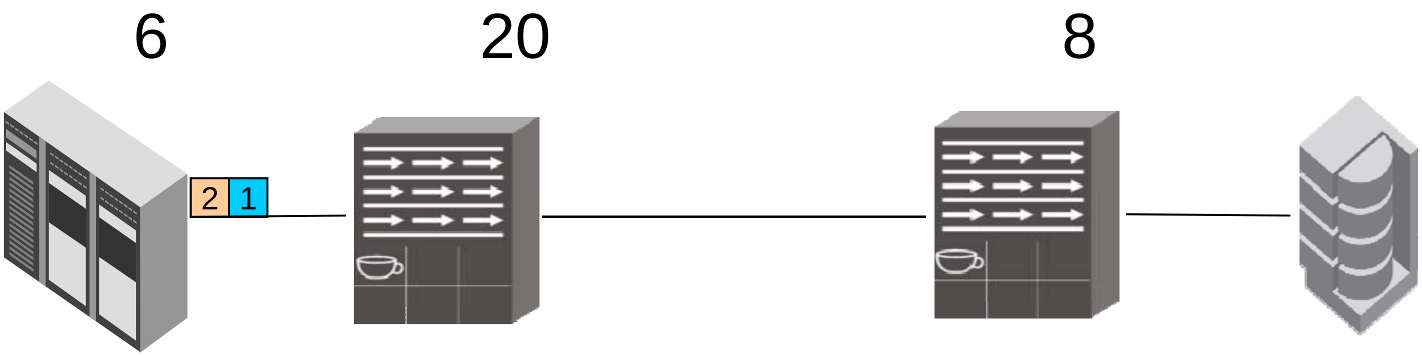
Example: Cascaded Directors

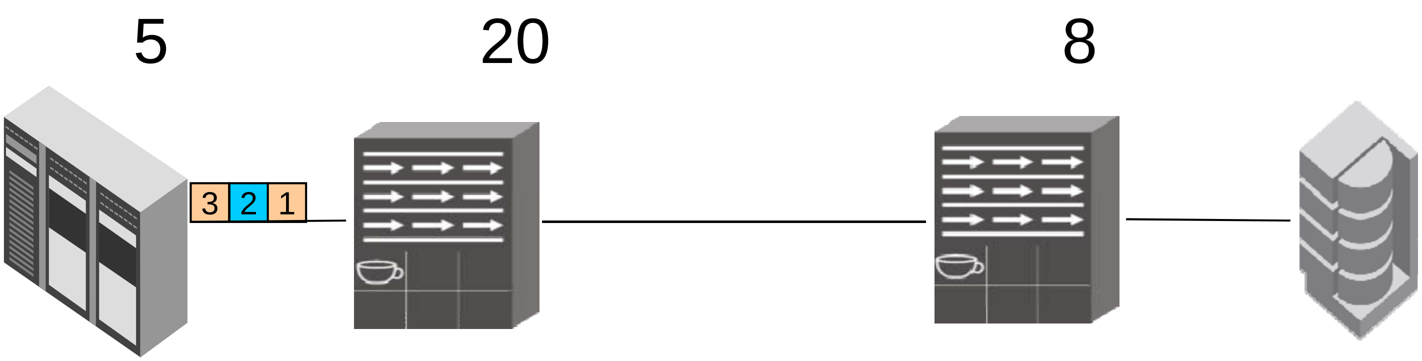
BUFFER CREDITS


8
20
8

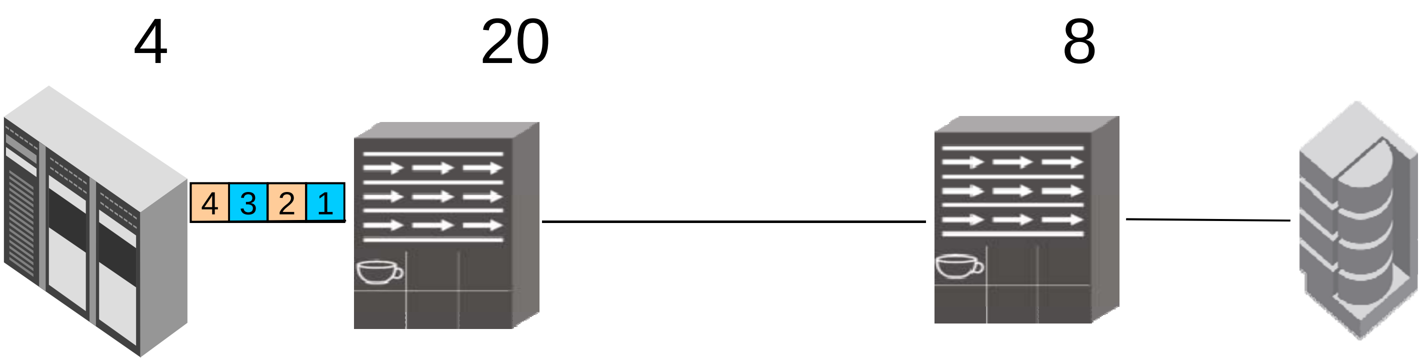


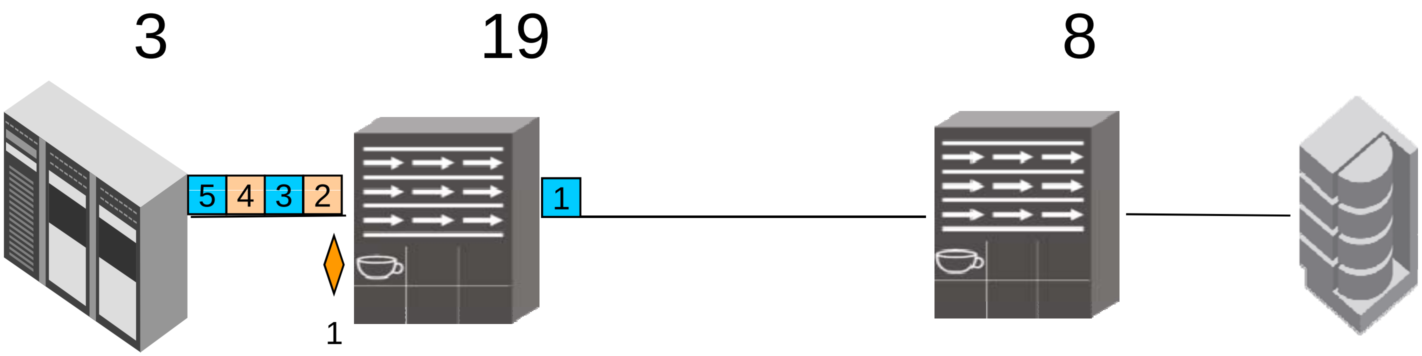
7
20
8
1

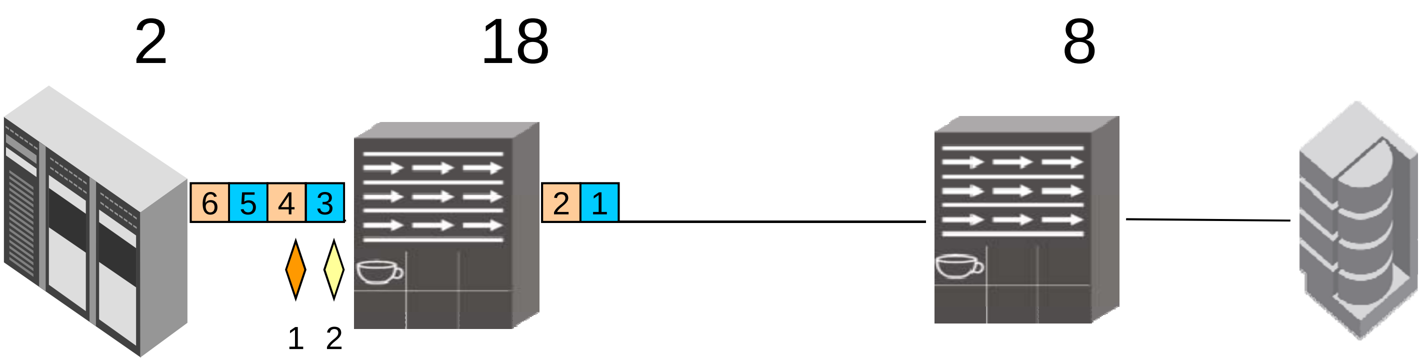

6
20
8
2
1

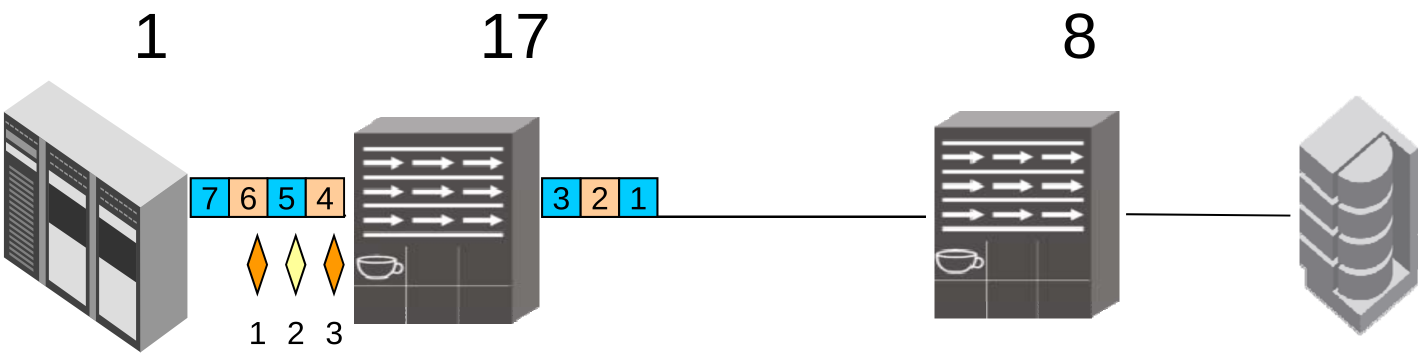


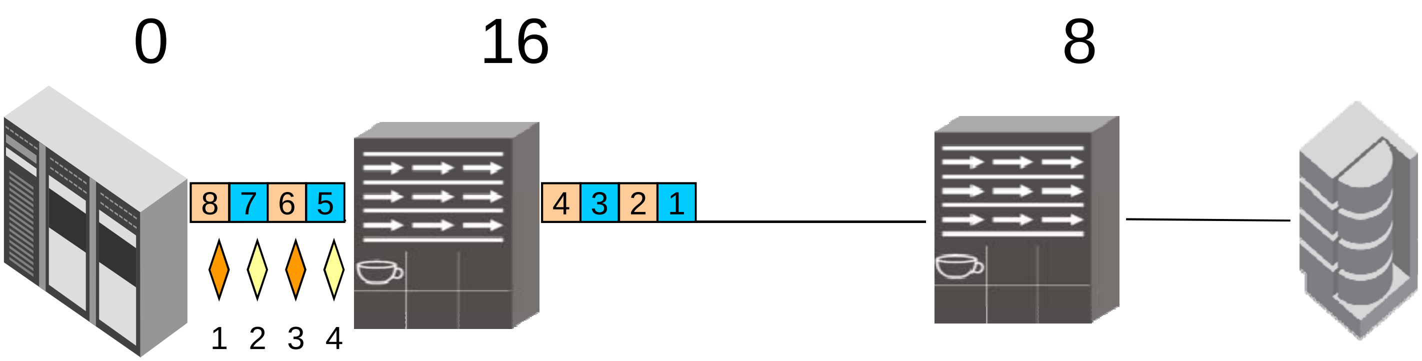
5
20
8
3
2
1

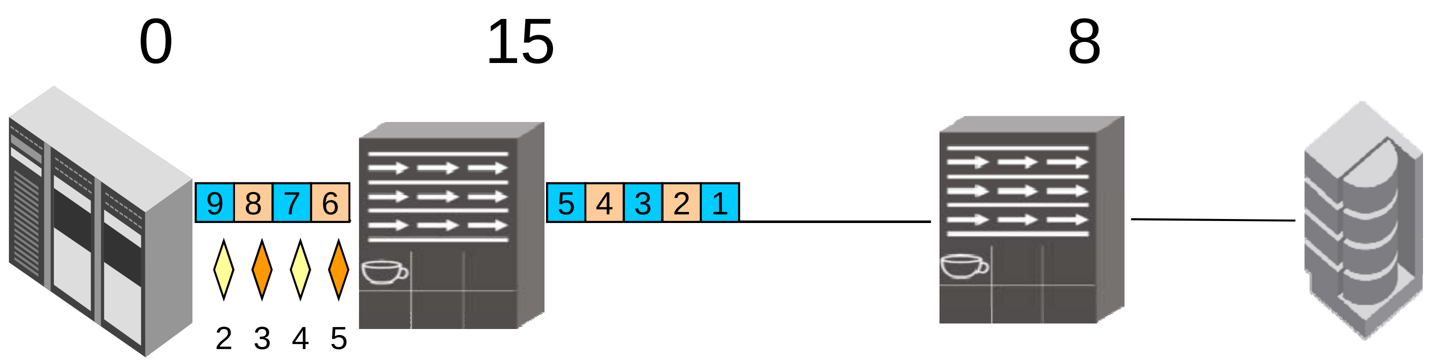

4
20
8
4
3
2
1

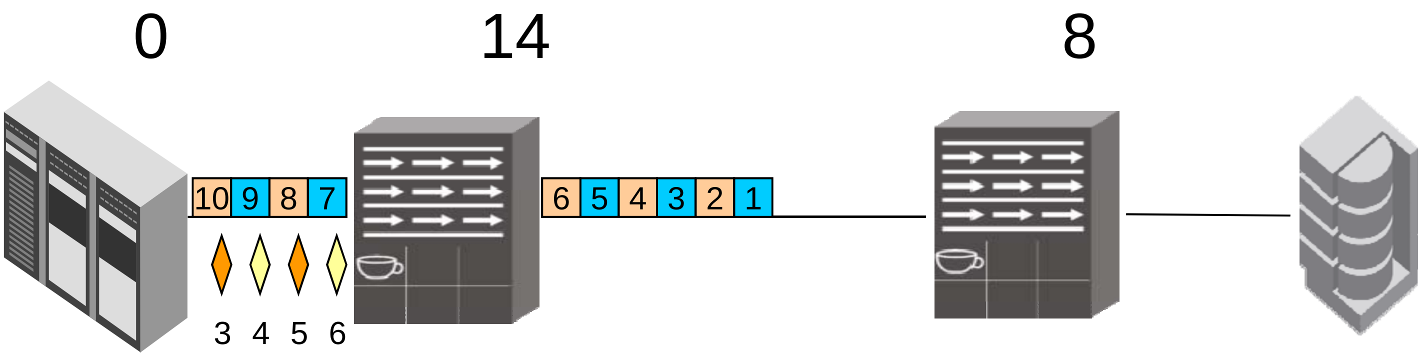

3
19
8
5
4
3
2
1
1


2
18
8
6
5
4
3
2
1
1
2

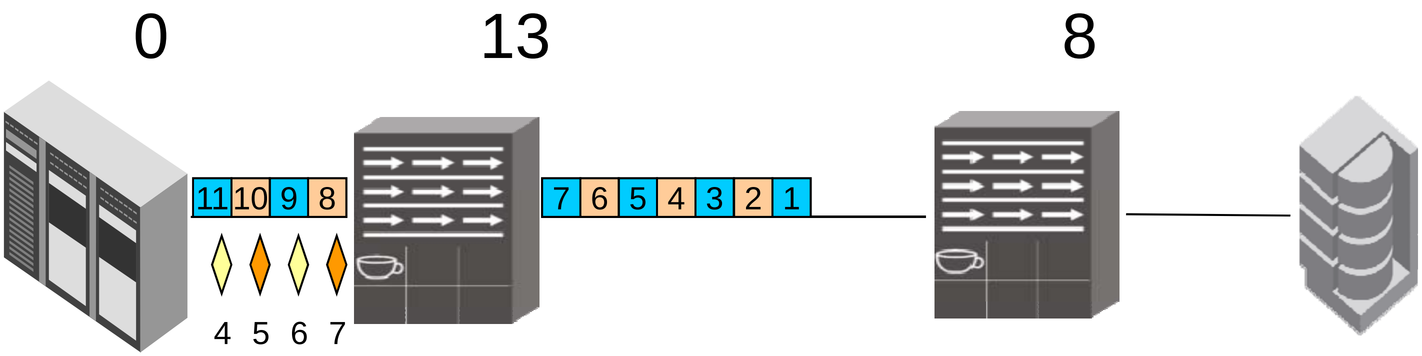

1
17
8
7 6 5 4
3 2 1
1 2 3

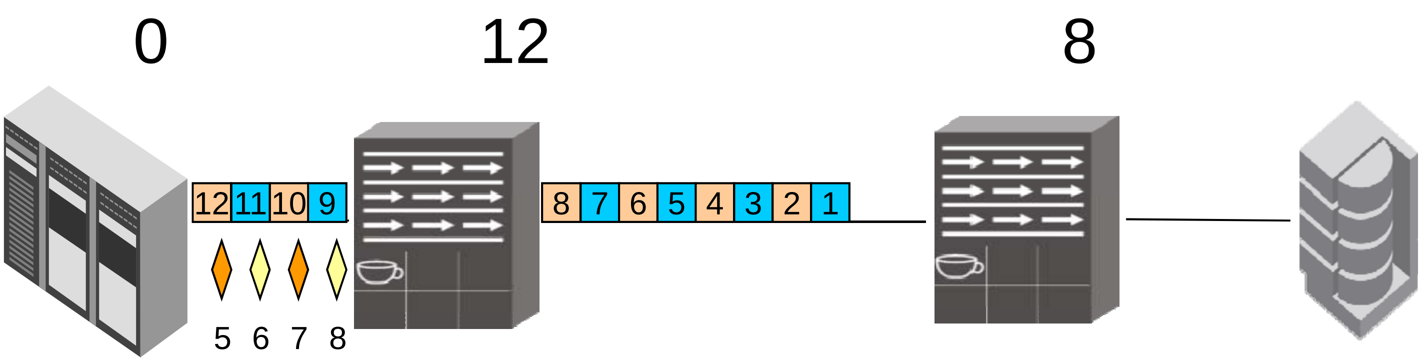

0
16
8
8 7 6 5
1 2 3 4
4 3 2 1

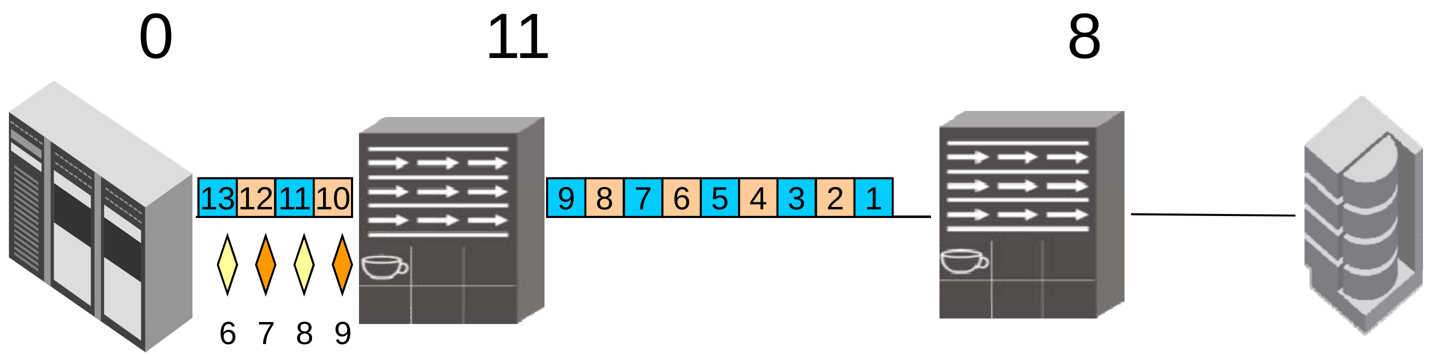

0
15
8
9 8 7 6
5 4 3 2 1
2 3 4 5

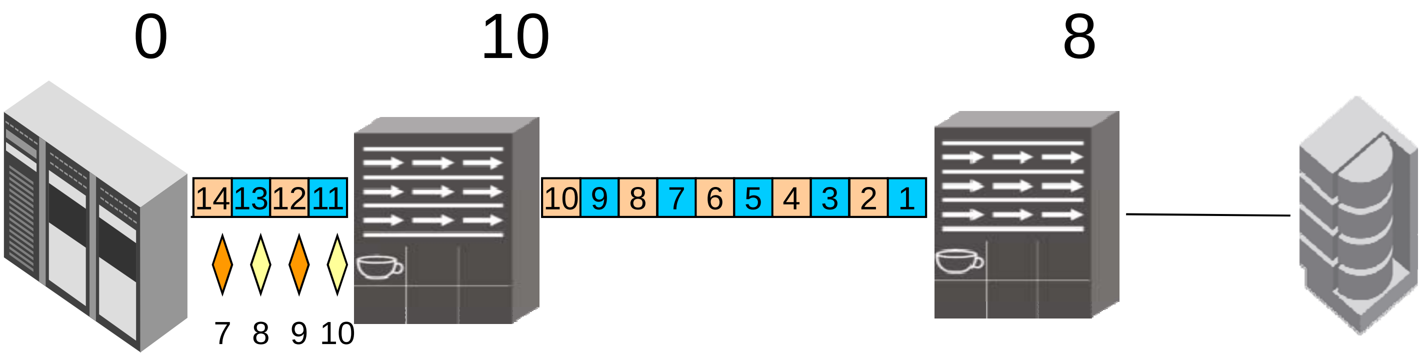


0
14
8
10 9 8 7
6 5 4 3 2 1
3 4 5 6

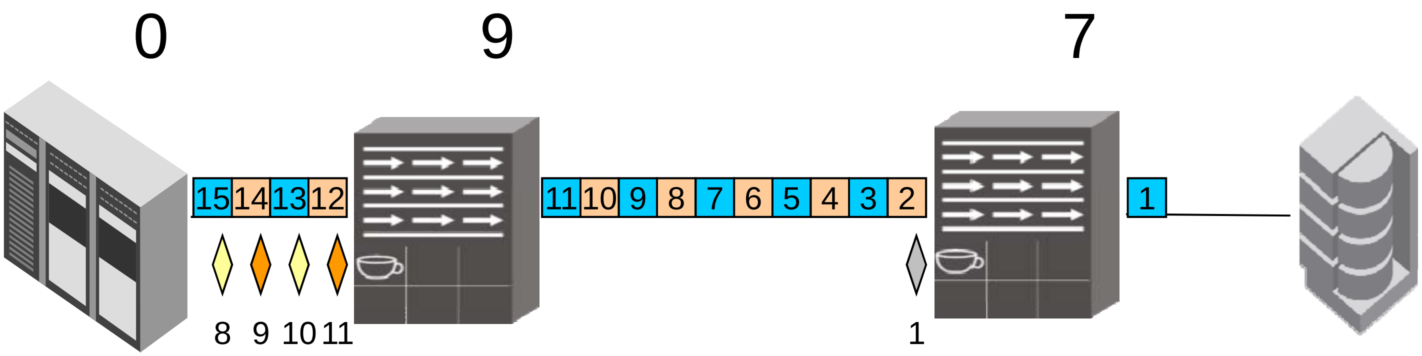

0
13
8
11 10 9 8
7 6 5 4 3 2 1
4 5 6 7


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

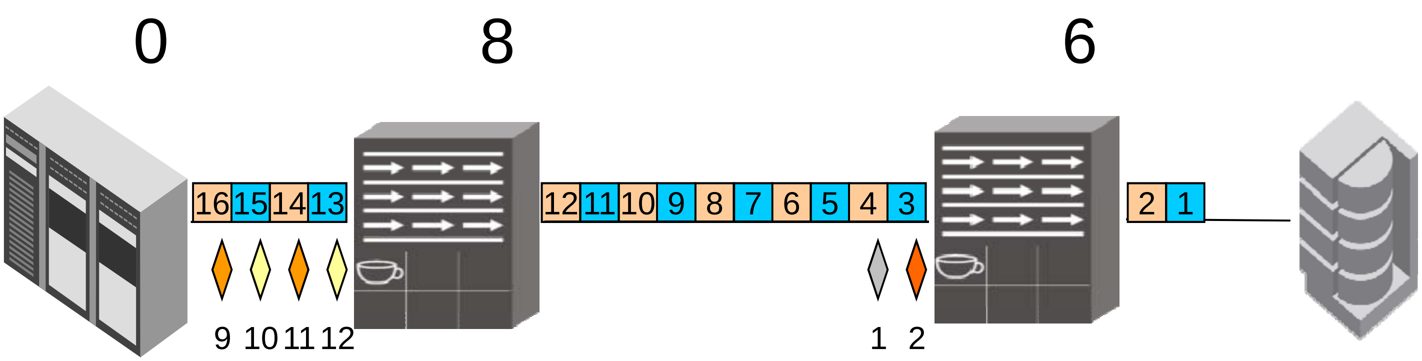

0
11
8
13 12 11 10
9 8 7 6 5 4 3 2 1
6 7 8 9



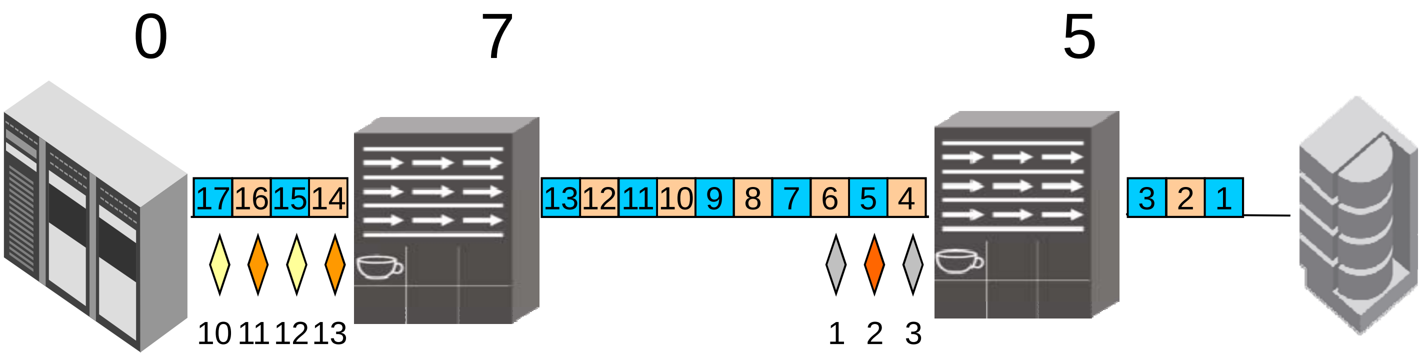
0
10
8
14 13 12 11
7 8 9 10
10 9 8 7 6 5 4 3 2 1

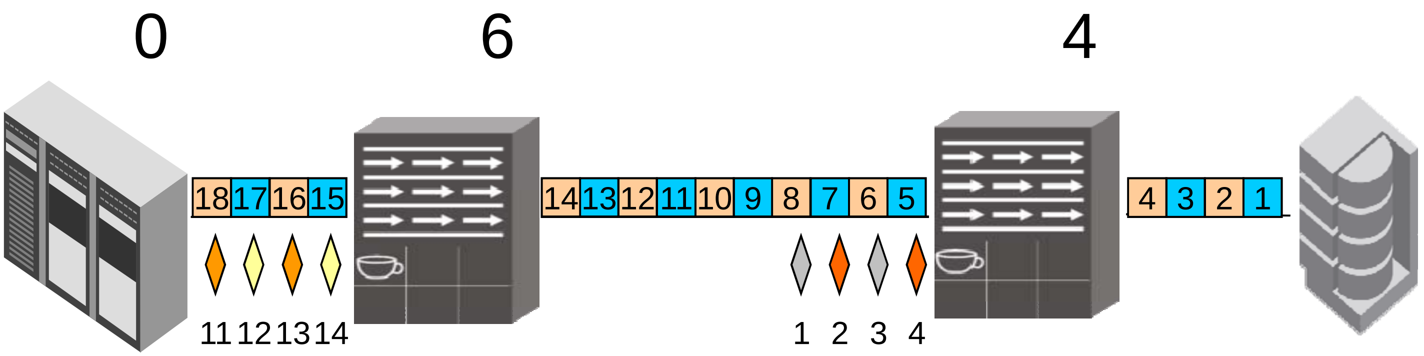


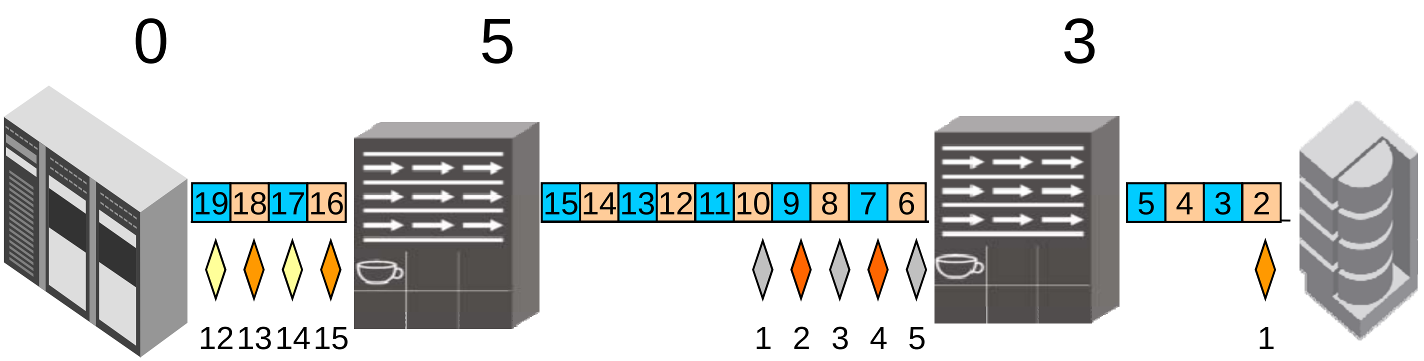
0
9
7
15 14 13 12
11 10 9 8 7 6 5 4 3 2
1
8 9 10 11
1

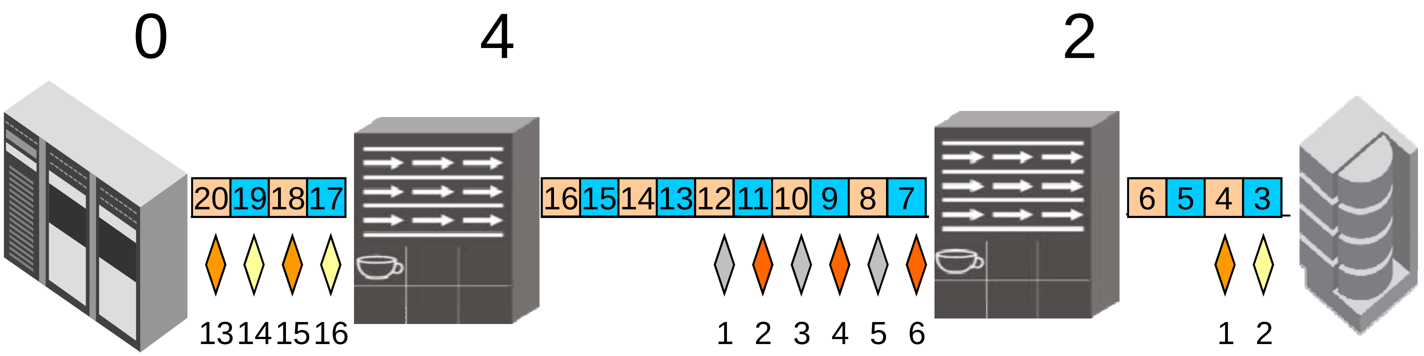


0
8
6
16 15 14 13
12 11 10 9 8 7 6 5 4 3
2 1
9 10 11 12
1 2

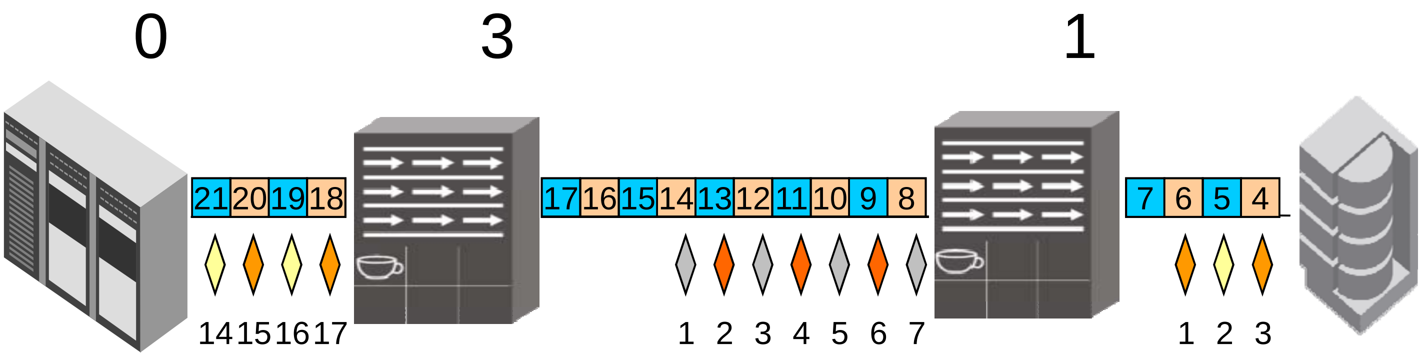

0
7
5
17 16 15 14
13 12 11 10 9 8 7 6 5 4
3 2 1
10 11 12 13
1 2 3

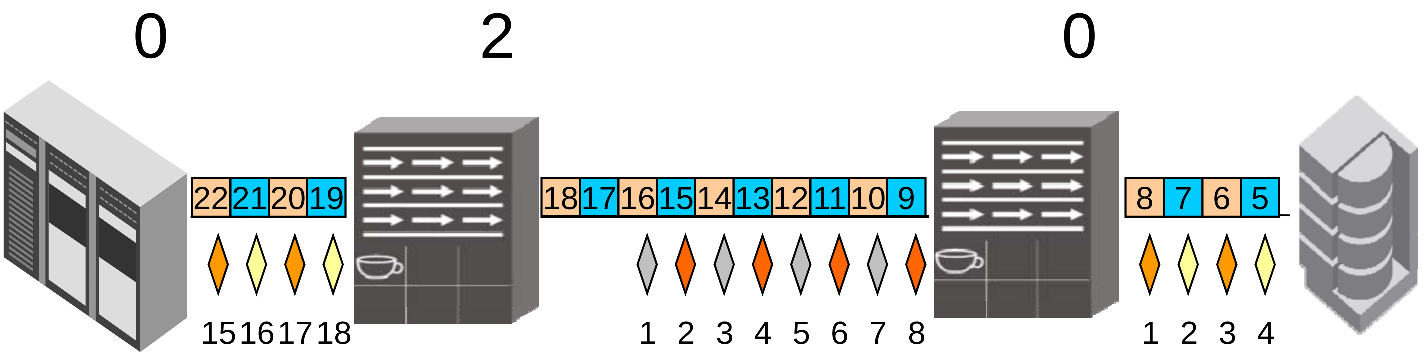

0
6
4
18 17 16 15
14 13 12 11 10 9 8 7 6 5
4 3 2 1
11 12 13 14
1 2 3 4

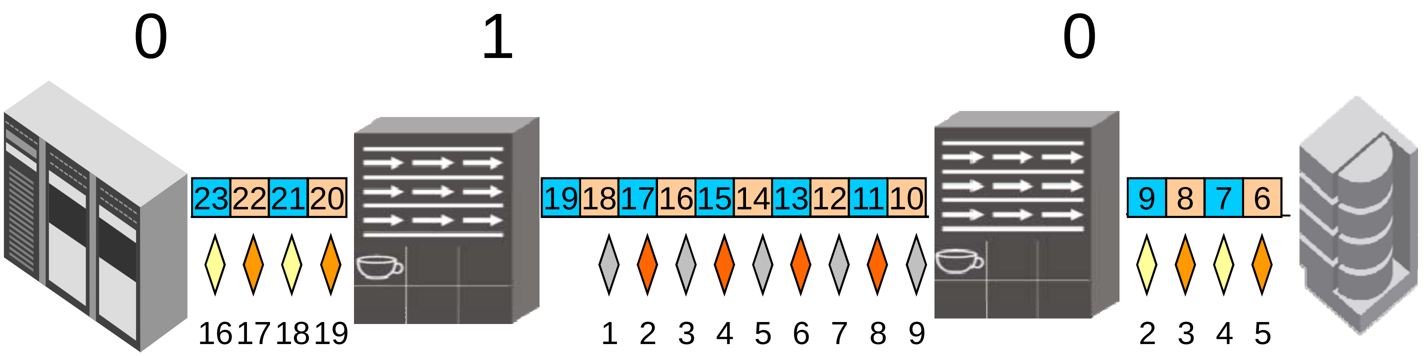

0
5
3
19 18 17 16
12 13 14 15
15 14 13 12 11 10 9 8 7 6
1 2 3 4 5
5 4 3 2
1

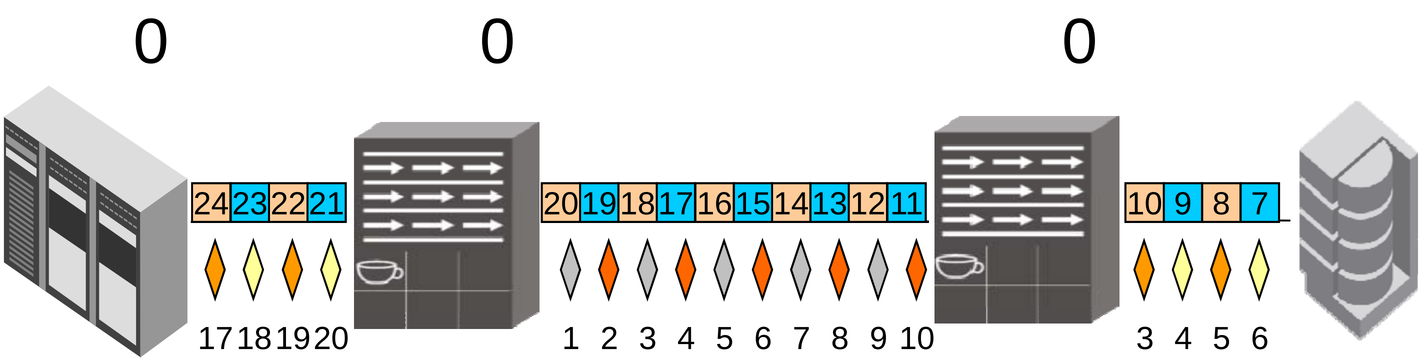


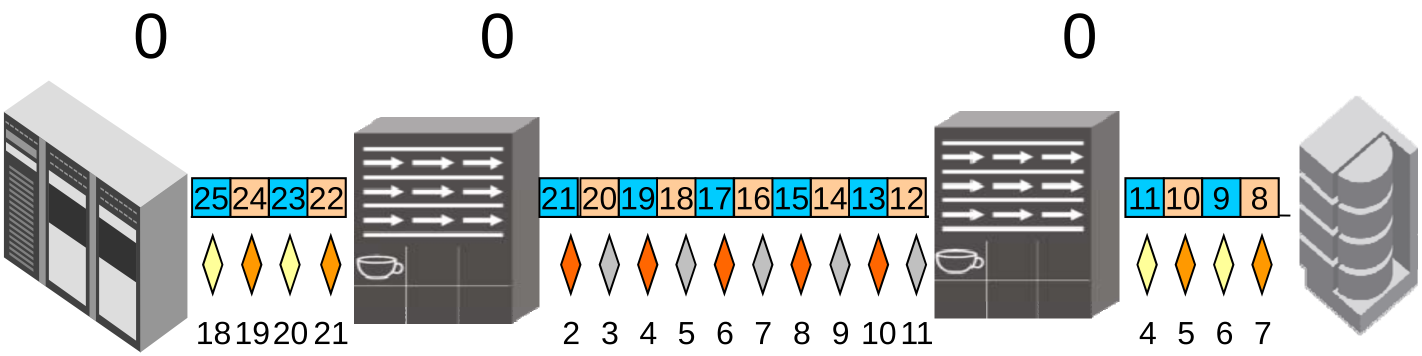
0
4
2
20
19
18
17
13
14
15
16
16
15
14
13
12
11
10
9
8
7
1
2
3
4
5
6
6
5
4
3
1
2


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


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


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


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



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

THIS PAGE INTENTIONALLY LEFT BLANK

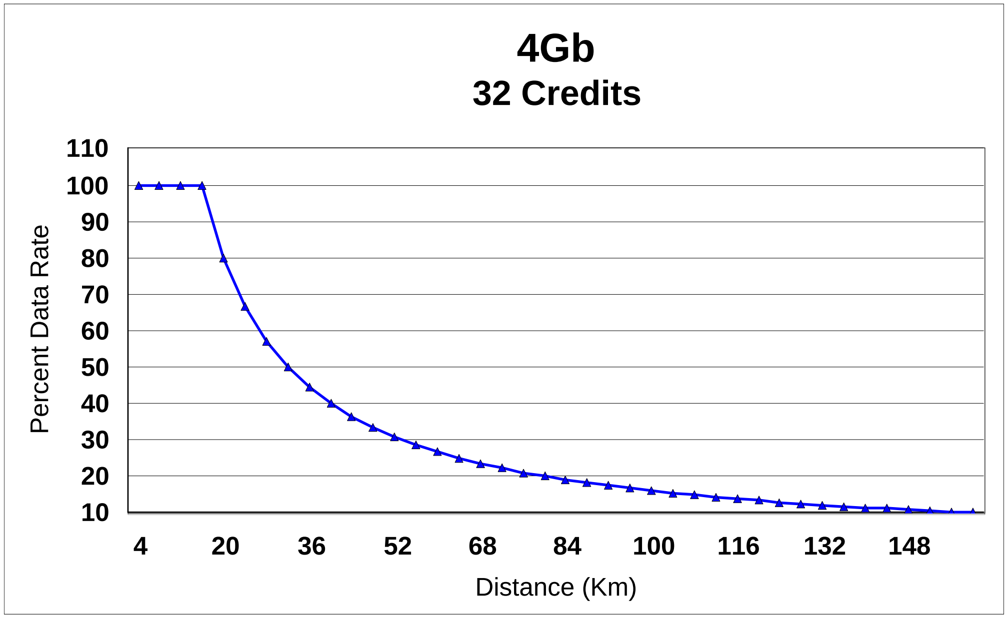
Other Neat Stuff

BUFFER CREDITS

How much credit do I need?

- Good “Rule of thumb”

Number of credits needed = $1 + \frac{\text{Link speed in Gb/s} * \text{Distance in km}}{\text{Frame Size in KB}}$



Example: 20 km at 1 Gb/s

$1 + \frac{1 * 20}{2} = 11$

Example: 10 km at 4 Gb/s

$1 + \frac{4 * 10}{2} = 21$

How “long” is a frame?

- Traveling at the speed of light, a frame can be very long
  - At 1G, the length of a frame is about 4-kilometers.
  - At 2G, the length of a frame is about 2-kilometers.
  - At 4G, the length of a frame is about 1-kilometer.
  - At 8G, the length of a frame is about 500-meters.
  - At 16G, the length of a frame is about 200-meters.

How “fast” is a frame?

- Speed of light in fibre
  - 200,000 km/second
  - 5-microseconds/km
- Transmission Rates in fibre
  - At 1G, a frame is sent in about 20-microseconds.
  - At 2G, a frame is sent in about 10-microseconds.
  - At 4G, a frame is sent in about 5-microseconds.
  - At 8G, a frame is sent in about 2.5-microseconds.
  - At 16G, a frame is sent in about 1-microsecond.

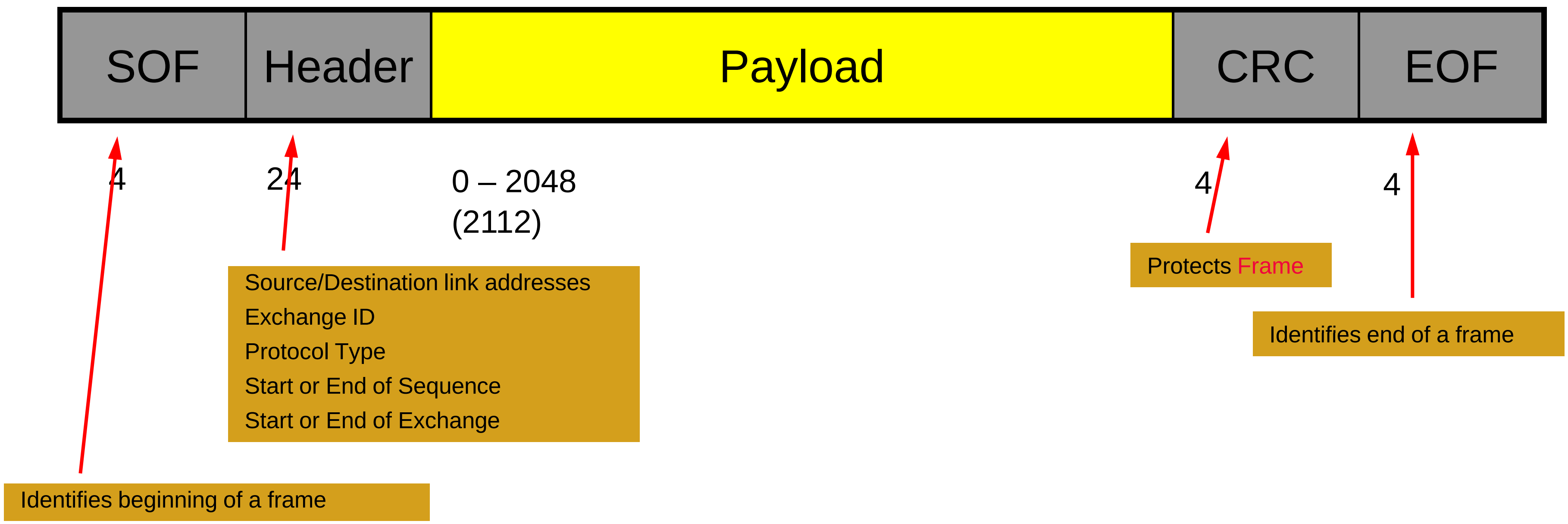
EXCHANGES

The parts of a transmission

- **Frame**
  - Building block of an Fibre Channel connection
  - Contains the information to be transmitted
  - The address of the source and destination
  - Control information
- **Sequence**
  - A set of one or more related Frames
  - Transmitted unidirectionally from one port to an other
- **Exchange**
  - An Exchange is one or more nonconcurrent sequences
  - A single operation
  - May be unidirectional or bidirectional

Fibre Channel Frame

The basic building block is the FRAME



Ficon IU Examples

1 Frame IU to transfer a Read CCW

SOF	Header	Ficon Header	CRC	EOF

3 Frame IU to transfer 4K of data

SOF	Header	Ficon Header	CCW Data (2016 bytes)	CRC	EOF

SOF	Header	CCW Data (2048 bytes)	CRC	EOF

SOF	Header	CCW Data (32 bytes)	Ficon CRC	CRC	EOF

Urban Legend:
FICON uses fewer Exchanges Than FCP

- Fibre Channel Architecture defines an Exchange as
  - "A mechanism for identifying and managing an operation between two ports"
- All IUs (a.k.a. Sequences) that make up a single I/O operation are part of an Exchange
- 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"

Sequences and IUs

- Each Upper Layer Protocol (ULP) defines the contents and format of it's own Information Units (IUs)
  - Commands
  - Data
  - Status
  - Control
  - Etc
- Ficon IUs can be up to 8K (8192) in size
  - 8160 (8K-32) bytes of data
    - 32 bytes contain Ficon Header information
  - 4 frames are needed for the largest IU
- The collection of frame(s) that make up a IU are called a Sequence
  - A Sequence may be as small as a single Frame

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

ERROR SENSITIVITY

Urban Legend: FICON is more sensitive than FCP

- Is Ficon More Sensitive to Errors than FCP?
  - Reasons for Link Errors are the same
- Is a Ficon frame more likely to get lost, damaged or corrupted than FCP?
  - The probability is the same
  - Frames are frames
- When a Ficon frame gets lost, damaged or corrupted, is the recovery action different from FCP?
  - Both protocols retry when errors occur
    - 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

Word \ Bits	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

Thank you

SUMMARY

Summary

- Buffer Credits
  - Distance
  - Flow Control
- Exchanges
  - Unidirectional
  - Bidirectional
- Error Sensitivity
  - Recovery
  - Reporting

SHARE, Orlando, August 2011

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

Session 9931

THANK YOU!

REFERENCES

Speaker Biography

- Lou Ricci
  - IBM
    - 32-years
    - 24-years in channel development
    - An inventor of FICON
  - FICON Firmware Team Leader
- Contact Information
  - lricci@us.ibm.com

Speaker Biography

- Patty Driever
  - IBM
  - System z I/O and Networking Technologist
- Contact Information
  - pgd@us.ibm.com

Speaker Biography

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

BONUS SLIDES

Speed of Light

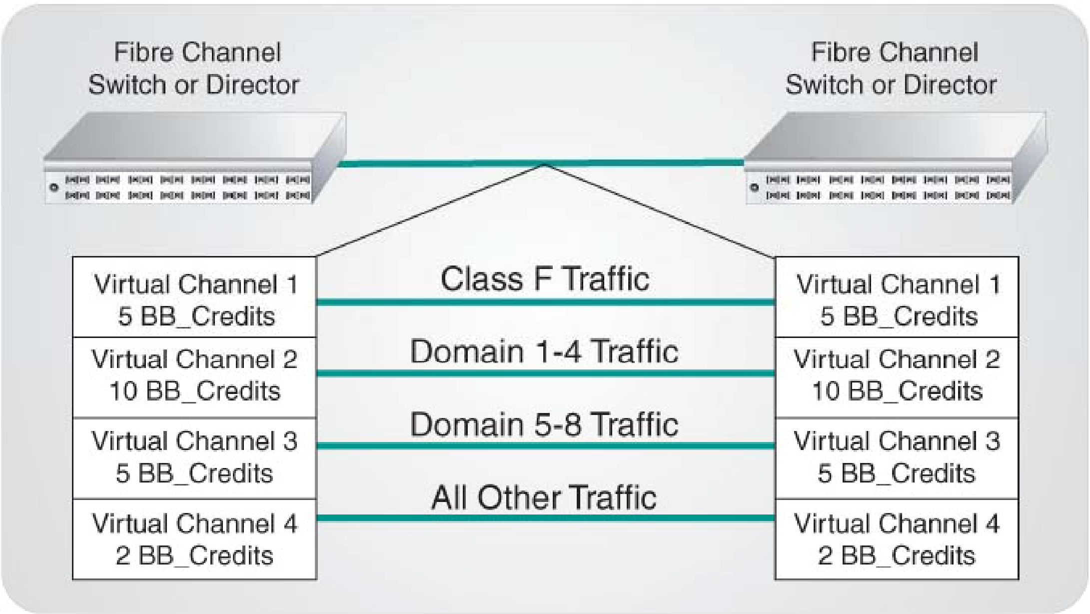
- 299 792 458 meters per second
  - In a vacuum
  - National Institute of Standards and Technology
    - http://www.nist.gov/pml/wmd/metric/length.cfm
- 300 000 000 meters per second
  - Common approximation
  - Wikipedia
    - http://en.wikipedia.org/wiki/Speed_of_light
- 1.44 – 1.46
  - Refractive index of 1300 nanometer fiber
  - Encyclopedia of Laser Physics and Technology
    - http://www.rp-photonics.com/refractive_index.html
- 1.5
  - Common approximation
  - Wikipedia
    - http://en.wikipedia.org/wiki/Optical_fiber#Materials
- **200 000 000 meters per second**
  - Calculated speed of light in single-mode fiber
  - V = C/N
  - 200 000 000 = 300 000 000 / 1.5

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



THIS SLIDE INTENTIONALLY LEFT BLANK