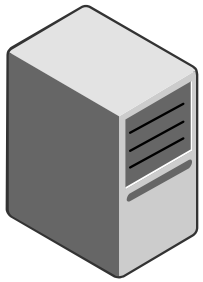
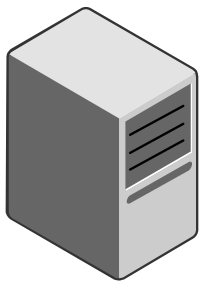
Persistent Reservation Issue #346

Roger Cummings
Symantec

Basic Situation

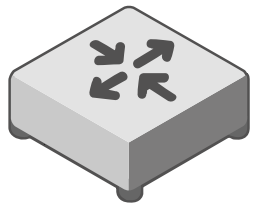


Init A

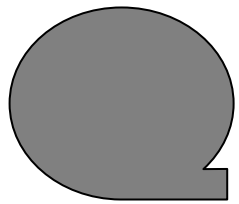


Init B

Two Initiators

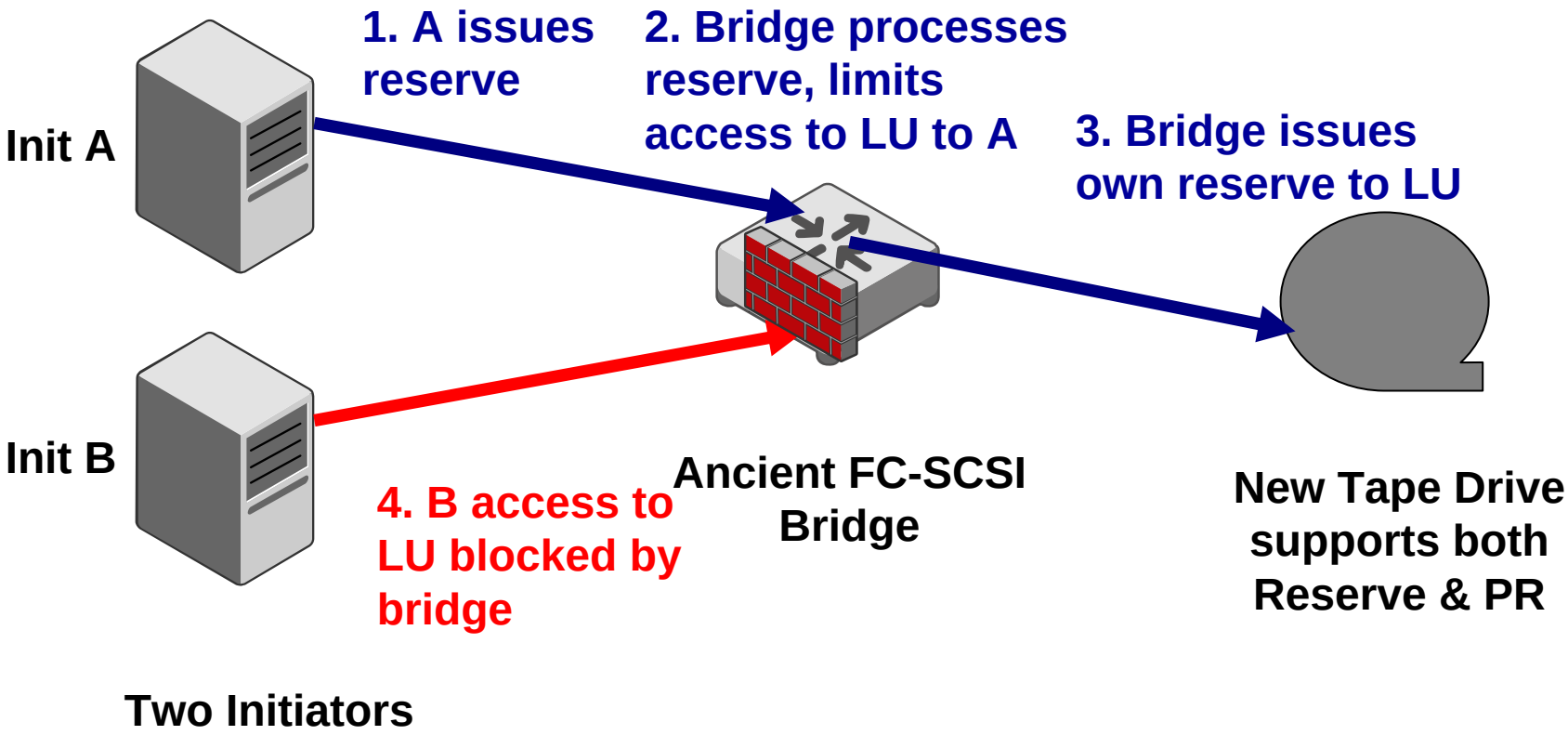


Ancient FC-SCSI Bridge

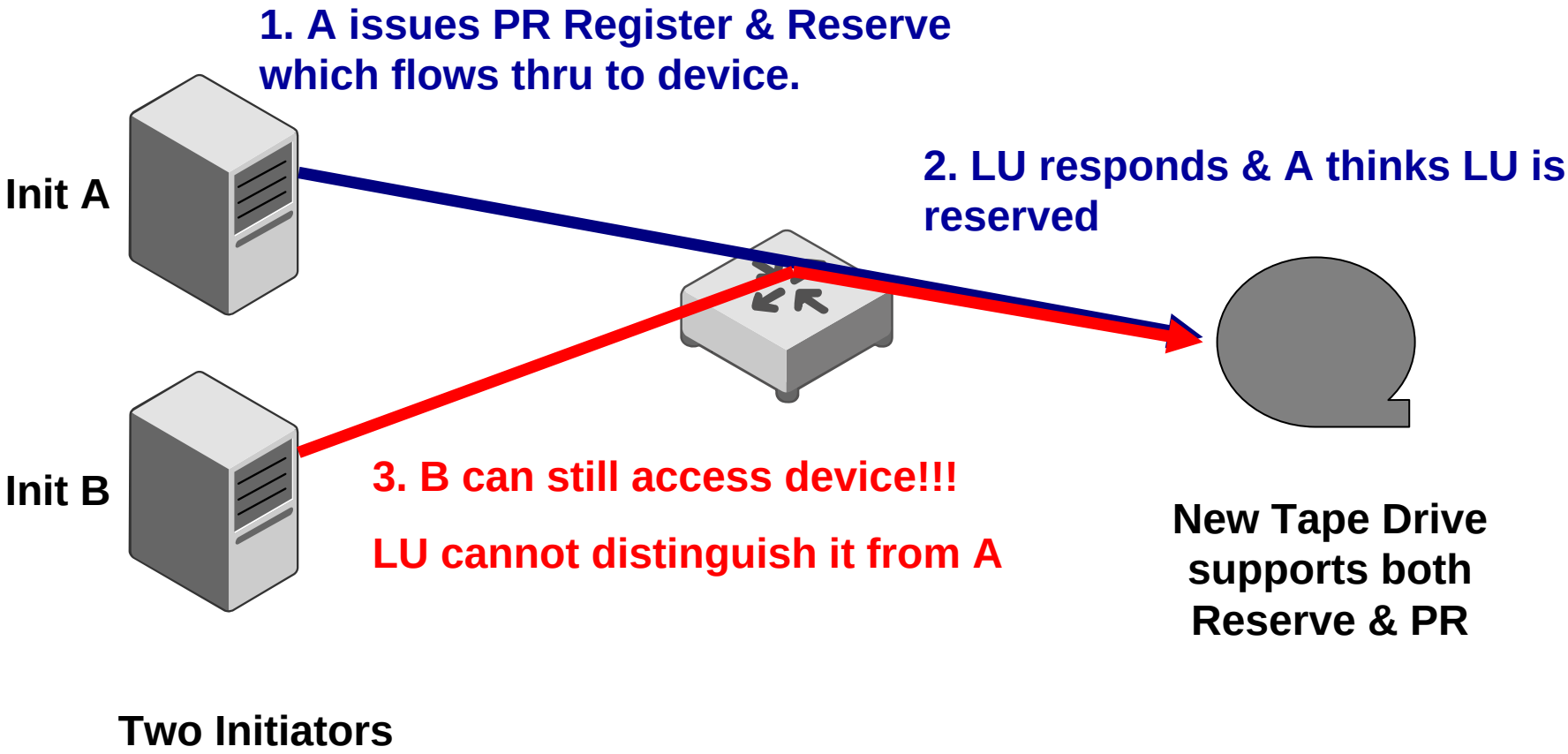


New Tape Drive supports both Reserve & PR

Reserve works well



PR fails horribly



And what's even worse

- We believe there's no way to detect that the bridge is there, much less what it's characteristics are!
  - We'd very much like to be proven wrong on this point!
- Need to get the message out ASAP that people deploying PR face a data corruption exposure in systems with old bridge products
  - In short term need all bridges that still have development support to be modified to either discard or reject PR commands if they don't support PR
  - And here's an idea for a longer-term solution……

Traceroute example

```
Tracing route to t11.org [66.155.124.38]
over a maximum of 30 hops:

  1    <1 ms    <1 ms    <1 ms  198.18.0.1
  2     1 ms     1 ms     1 ms  209.226.87.1
  3    27 ms     3 ms     3 ms  10.37.37.201
  4     3 ms    23 ms     3 ms  core1-vancouver-pos1-0.in.bellnexxia.net [206.10.8.101.13]
  5    71 ms    88 ms    97 ms  64.230.229.37
  6    71 ms    75 ms    73 ms  core2-chicago23-pos0-0.in.bellnexxia.net [206.10.8.103.114]
  7    74 ms    71 ms    76 ms  bx1-chicago23-pos11-0.in.bellnexxia.net [206.108.103.125]
  8    71 ms    72 ms    89 ms  p13-0.core01.ord01.atlas.cogentco.com [154.54.11.29]
  9   160 ms   140 ms   127 ms  p15-0.core02.ord01.atlas.cogentco.com [66.28.4.62]
 10   423 ms   381 ms   181 ms  p6-0.core02.jfk02.atlas.cogentco.com [66.28.4.85]
 11   171 ms   201 ms   200 ms  p13-0.core01.phl01.atlas.cogentco.com [66.28.4.2]
 12   244 ms   243 ms    99 ms  p4-0.core01.dca01.atlas.cogentco.com [66.28.4.17]
 13    92 ms   114 ms   108 ms  p15-0.core02.dca01.atlas.cogentco.com [66.28.4.22]
 14   104 ms   224 ms   104 ms  p14-0.core01.atl01.atlas.cogentco.com [66.28.4.161]
```

- Based on the same ICMP protocol as Ping
- Utilizes a “Time To Live” field in the IP Header that’s decremented by each router
- Provides a simple elegant scheme that allows an application to completely map out an infrastructure

Add Field to INQUIRY

Bit / Byte	7	6	5	4	3	2	1	0
0	OPERATION CODE (12H)							
1	reserved			**LEVEL COUNT (LC)**			Obsolete	EVPD

- All current bridges can access byte 0 & 1 of CDB (to detect Reserve)
- Define new Level Count field (LC) containing unsigned integer in byte 1
- Rules as follows
  - If LC=0, existing behavior
  - If LC>0, decrement LC by 1
    - If result=0, respond to Inquiry command
      - Else, relay the command to the next level in the hierarchy as per SAM-3
- Would allow an application to map a hierarchy, and induce a “transparent” bridge to identify itself
  - Even supporting only EVPD=0 would be enough initially to locate the bridges

Feedback requested

- Is the situation with ancient bridges a candidate for an “info letter” sent to all T10 member companies?
- Does pursuing the “traceroute” Inquiry make any sense?