BGP Session Culling

https://tools.ietf.org/html/draft-iops-grow-bgp-session-culling

Will Hargrave, LONAP
Mat Griswold, 20C
Nick Hilliard, INEX
Job Snijders, NTT

Abstract

This document outlines an approach to mitigate negative impact on networks resulting from maintenance activities. It includes guidance for both IP networks and Internet Exchange Points (IXPs). The approach is to ensure BGP-4 sessions affected by the maintenance are forcefully torn down before the actual maintenance activities commence.

Why are we doing this

- Some people are culling, some people aren’t, we think they should

- Gain input from network operators everywhere

- Have a referenceable document to point people at

“Voluntary” or “Involuntary”

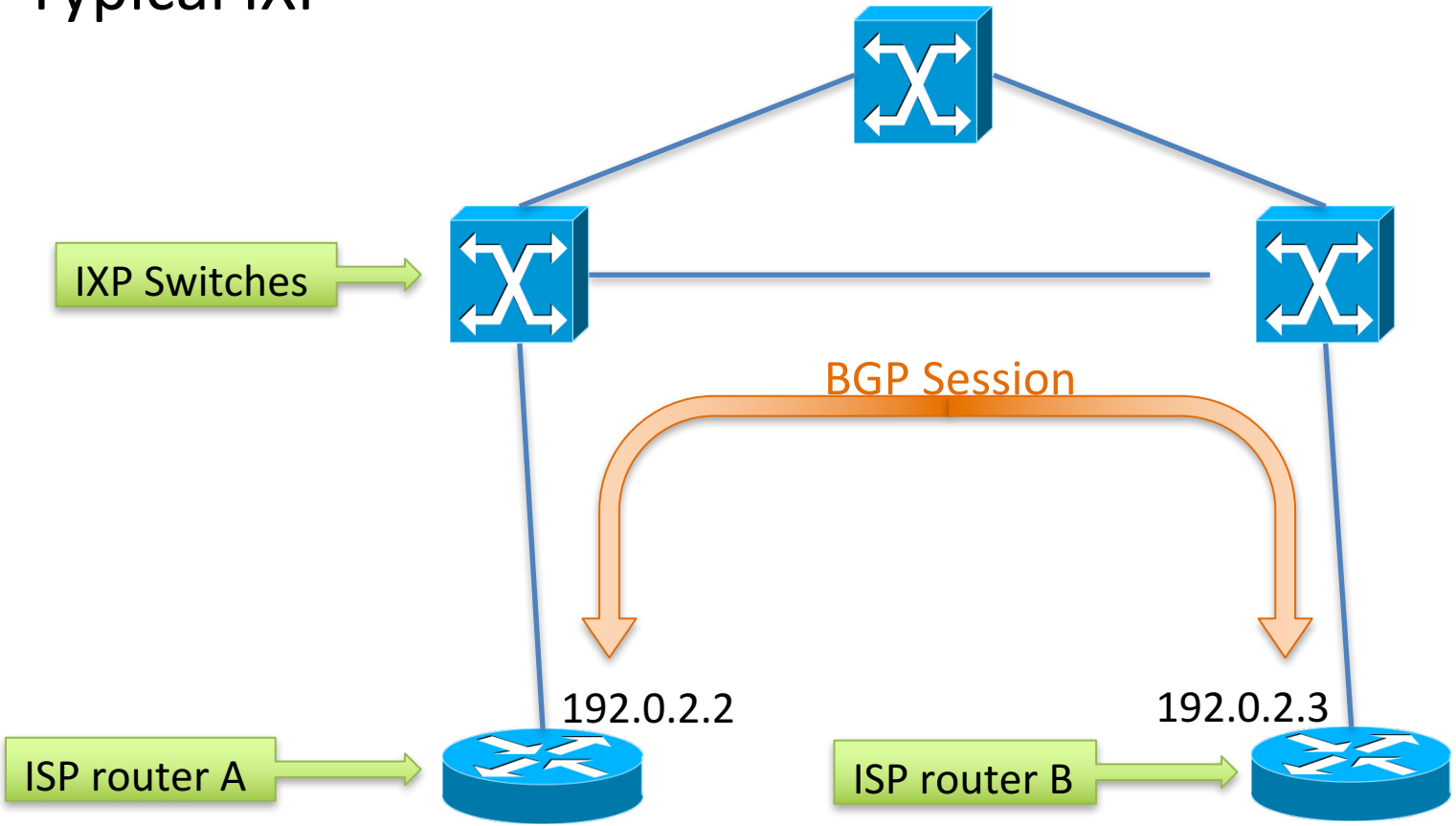
Voluntary: The BGP operator tears down potentially affected sessions, usually with an administrative shutdown.

Involuntary: The Caretaker of the lower level network disrupts BGP control-plane traffic, generally with an L4 ACL.

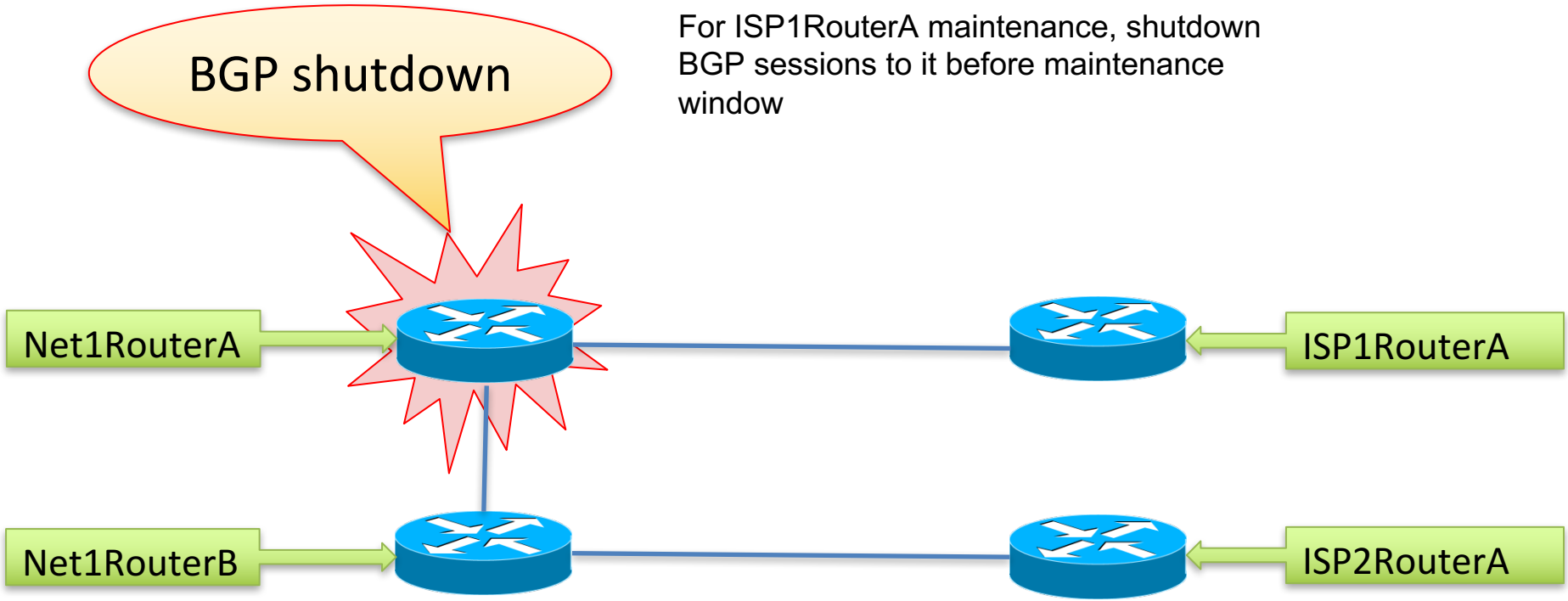
Suitable for cases where multilateral BGP is facilitated through a switched layer 2 fabric, notably IXPs.

Both result in a smooth drainage of traffic prior to losing data-plane and affecting end-user traffic.

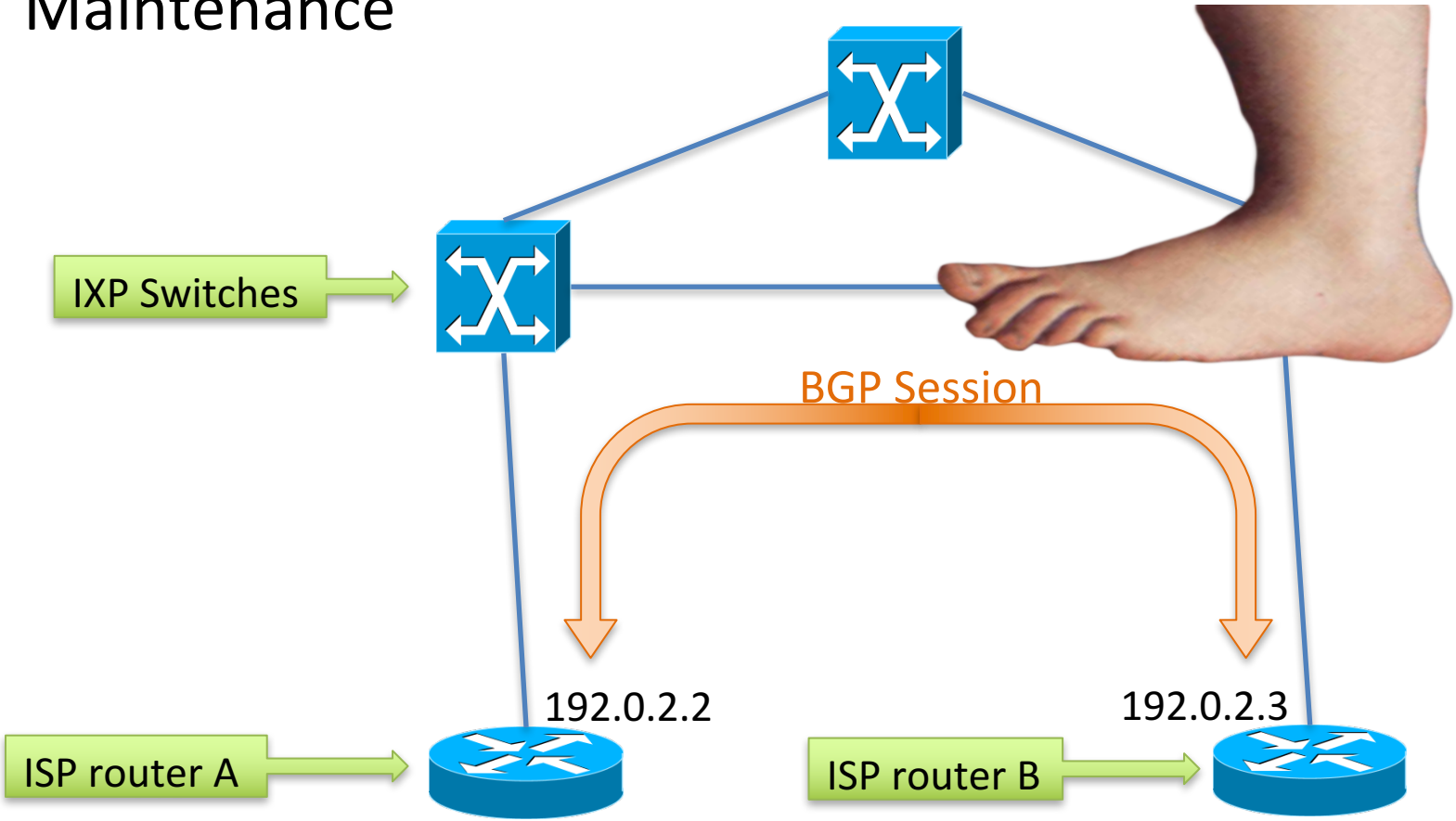
Typical IXP


IXP Switches
BGP Session
192.0.2.2
192.0.2.3
ISP router A
ISP router B

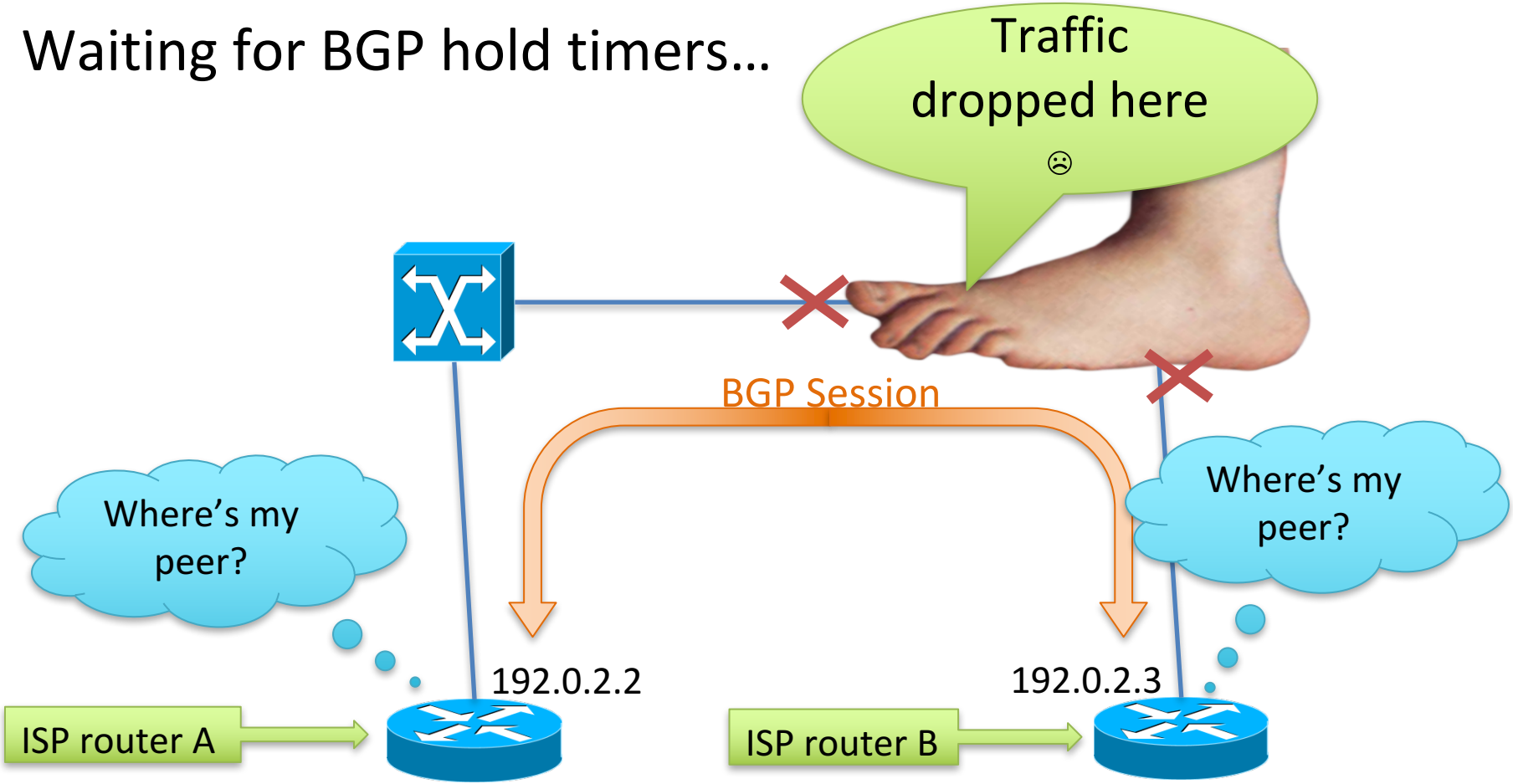
Voluntary Example



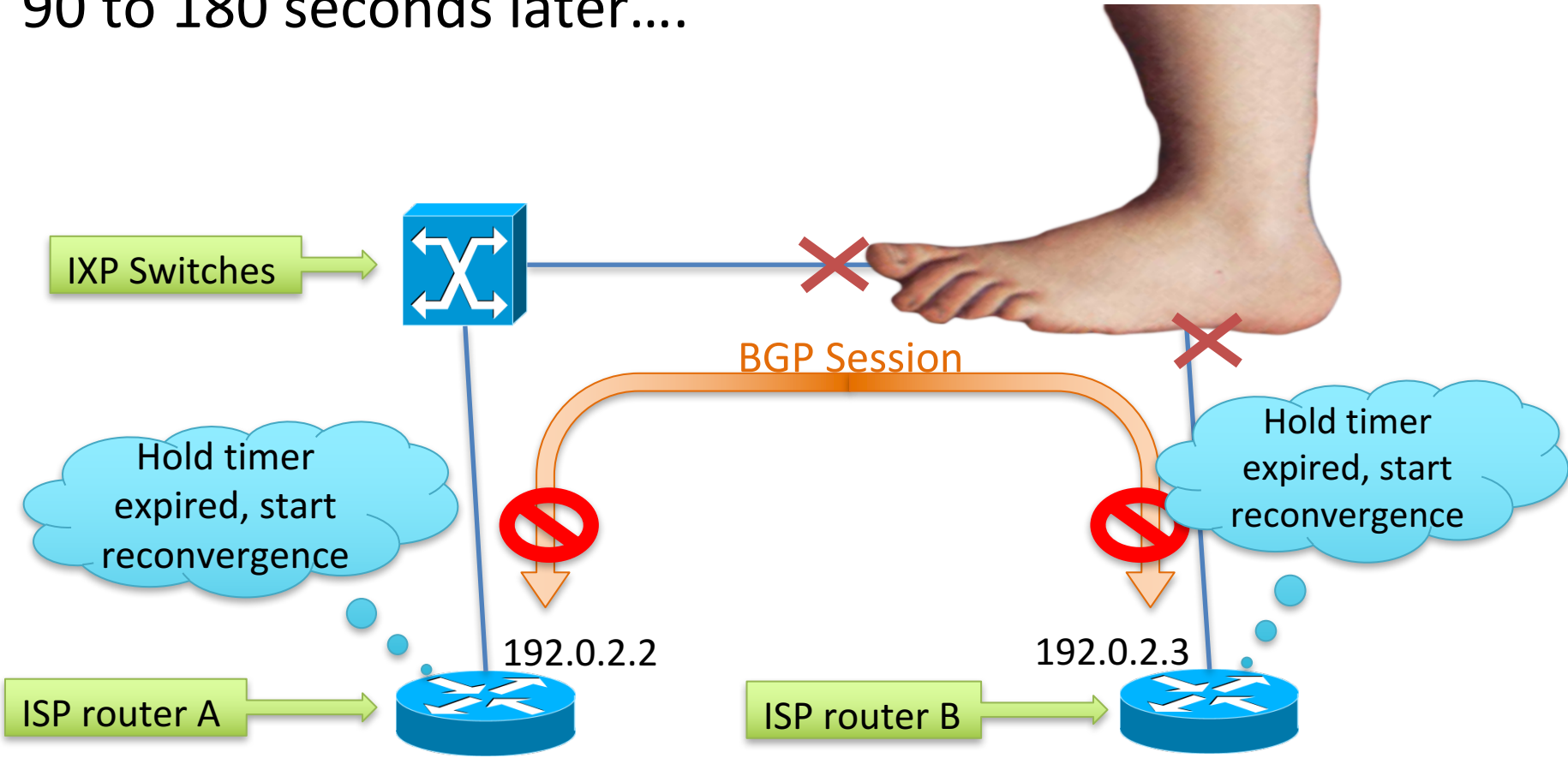
Maintenance



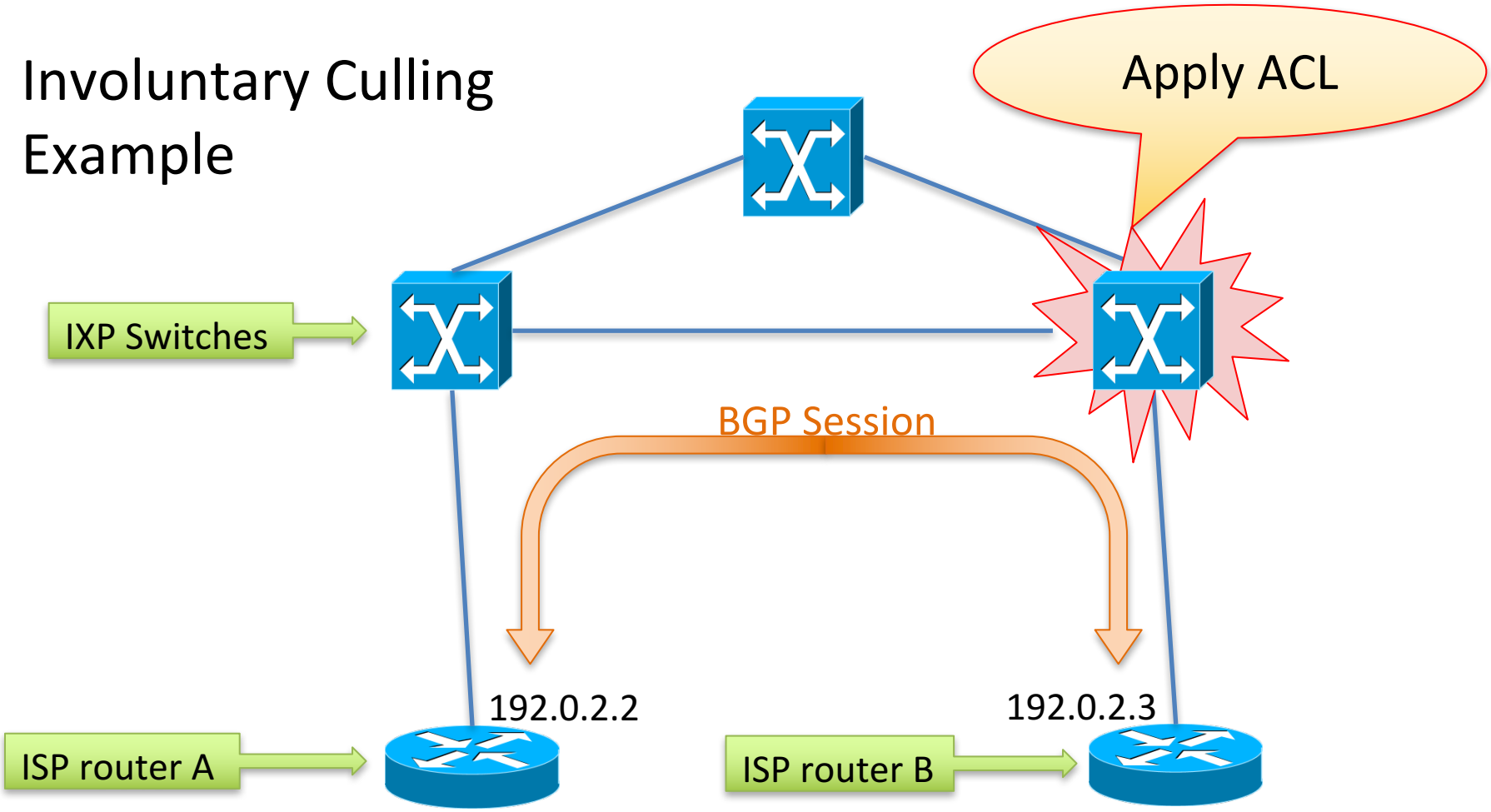
Waiting for BGP hold timers…



90 to 180 seconds later….



Involuntary Culling Example



L4 Packet Filter

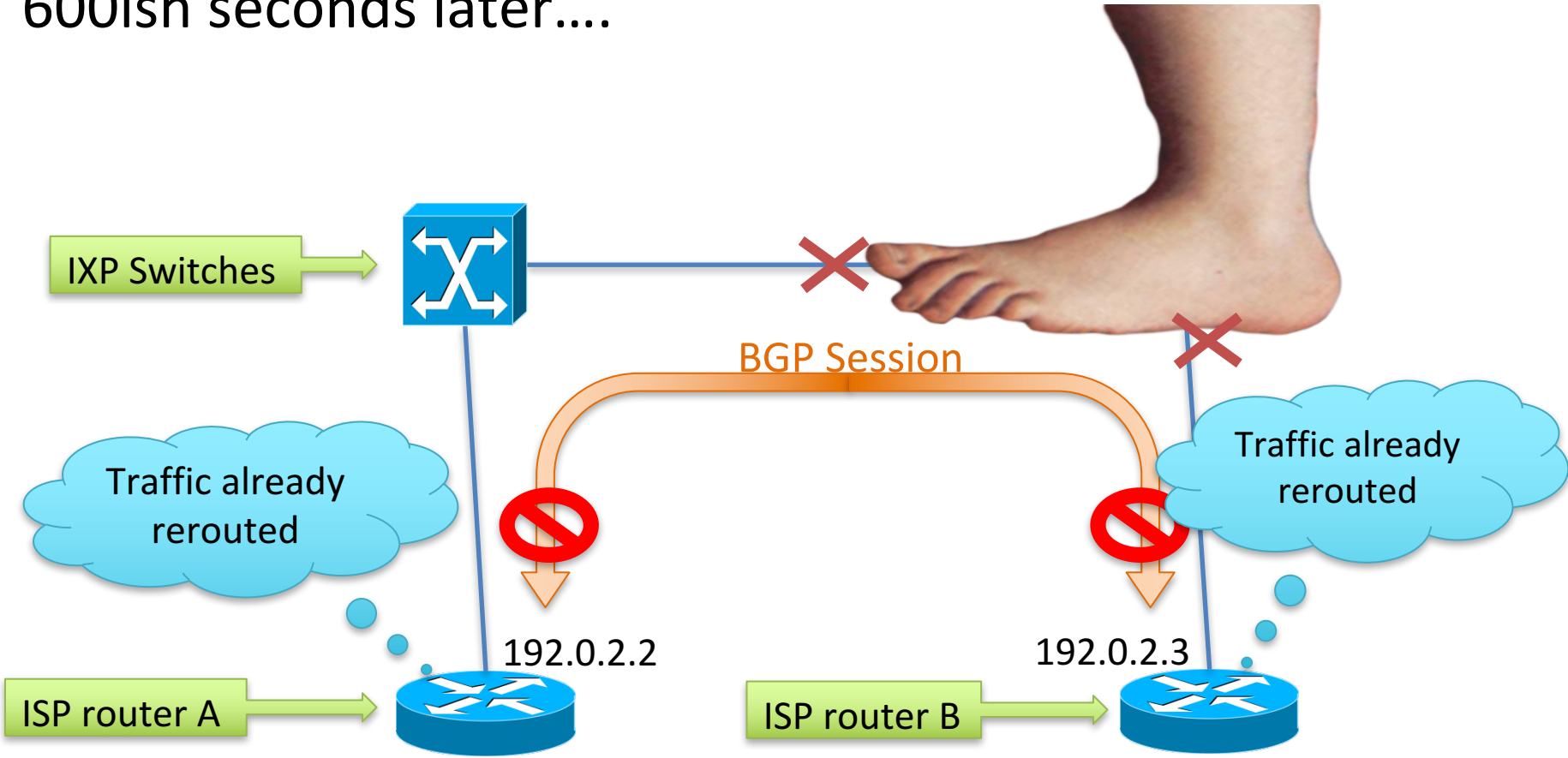
```
> show configuration firewall family ethernet-switching filter cull
  term cull-v4 {
      from {
          ip-version {
              ipv4 {
                  port bgp;
                  ip-source-address {
                      192.0.2.0/24;
                  }
                  ip-destination-address {
                      192.0.2.0/24;
                  }
                  ip-protocol tcp;
              }
          }
      }
      then discard;
  }
term cull-v6 {
```

Filter in both directions

IXP Subnet

Don't forget IPv6!

600ish seconds later….



Use and Monitoring

After applying culling, one needs to wait for BGP hold timer expiration to trigger Reconvergence and then monitor traffic across the affected ports to determine when it’s safe for maintenance.

Time can vary a lot depending on size of routing tables, hardware, et al, so the caretaker needs to use best judgement at when there is enough traffic drained.

Questions?

References:

- https://tools.ietf.org/id/draft-iops-grow-bgp-session-culling-00.txt

- https://ripe67.ripe.net/presentations/374-WH-IXPMaintReduce.pdf