Pain Points of Various Data Center Network Designs
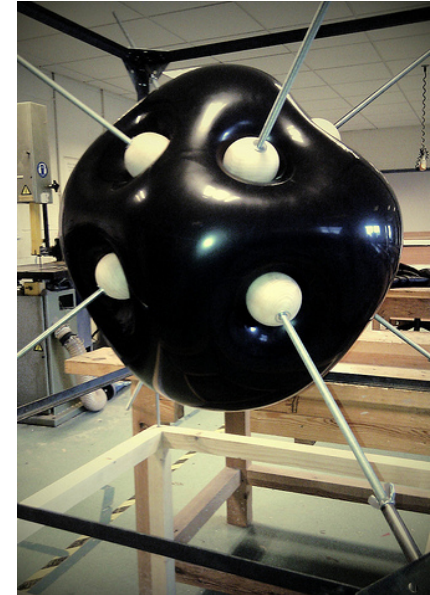
ARMD Update

Linda Dunbar: ldunbar@huawei.com

Benson Schliesser: bschlies@cisco.com

Background

- IETF Working Group
  - ARMD = "Address Resolution for Massive numbers of hosts in the Data center"
  - ARMD track in June NANOG
- Goal of this talk: solicit operators to challenge (publicly or privately) the generic DC network designs and their associated pain points

Scenario #1: L3 to Access (ToR)

- A single rack is it's own L2 domain, has its own IP subnet:

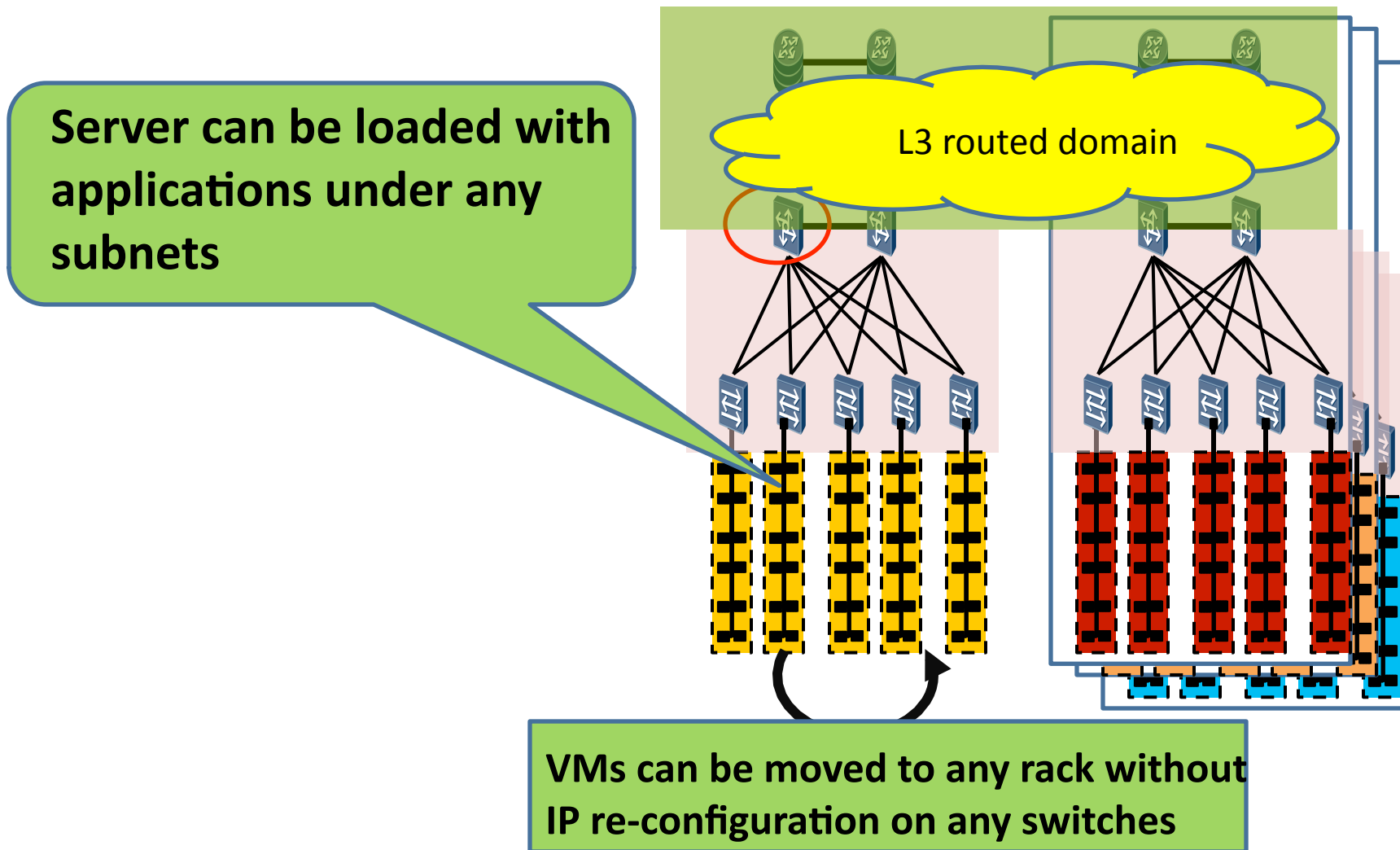- Benefits: ARP/ND scale very well. No problem.

L3 Routed Access

When server is loaded with new applications, it has to inherit the same IP subnet
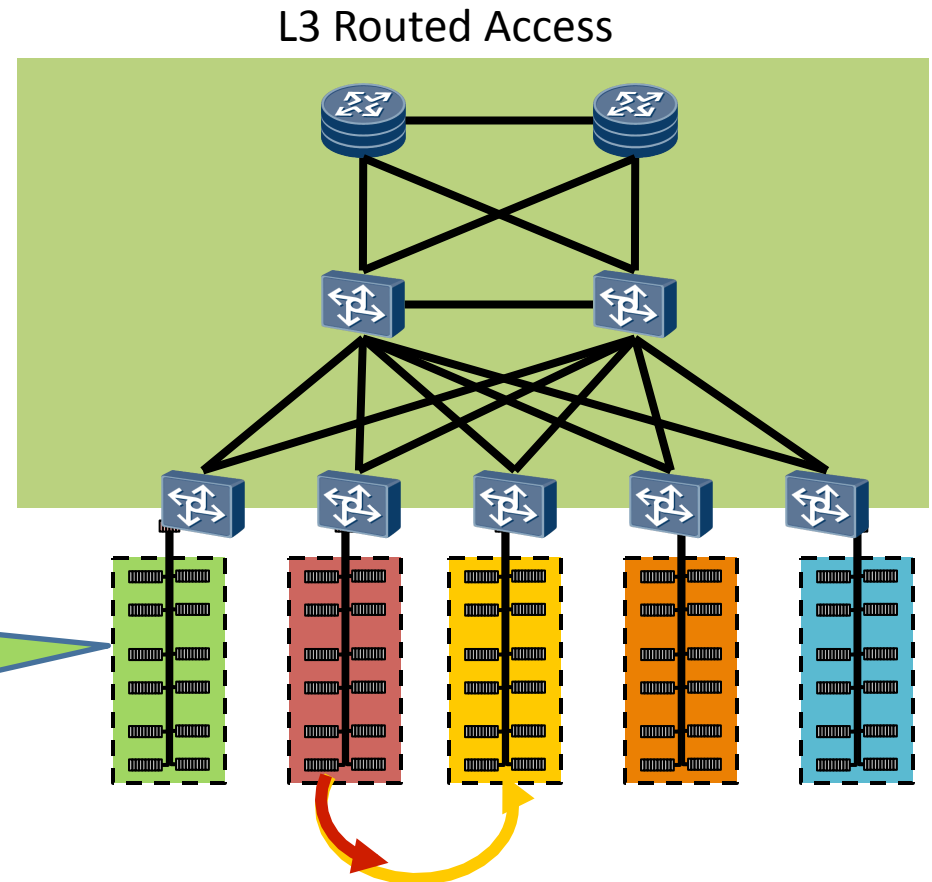
IP addresses have to be reconfigured when VMs move to a different rack

Scenario #2A: L3 to Aggregation

Traditional L2/L3 Tiered

L3 routed domain

Server can be loaded with applications under any subnets

VMs can be moved to any rack without IP re-configuration on any switches

Scenario #2B: L3 to Gateway Only

Server can be loaded with applications under any subnets

Minimal IP re-config needed

What triggers the need?
•Reduce or increase the number of racks when demand changes.
•Allow servers to be re-loaded with different applications under different subnets without any physical moving or IP re-configuration.

L3 routed domain

L2 Domain with many VLANs enabled
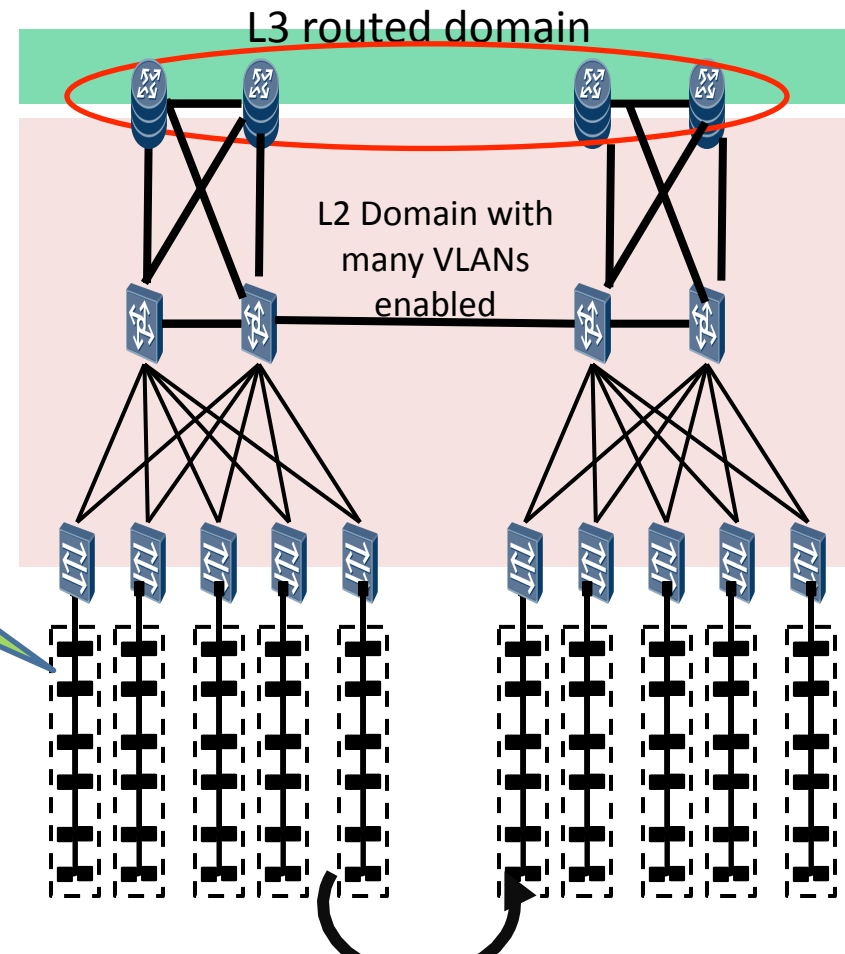
VMs can be moved to any rack without IP re-configuration on any switches

Pain Point #A for Scenario #2:
When external peers initiate communication with hosts inside data center

-Router needs to hold data frames,

-Trigger ARP/ND to validate if the target exists in the L2 domain

-When response is received from the target, send the data frames to the target
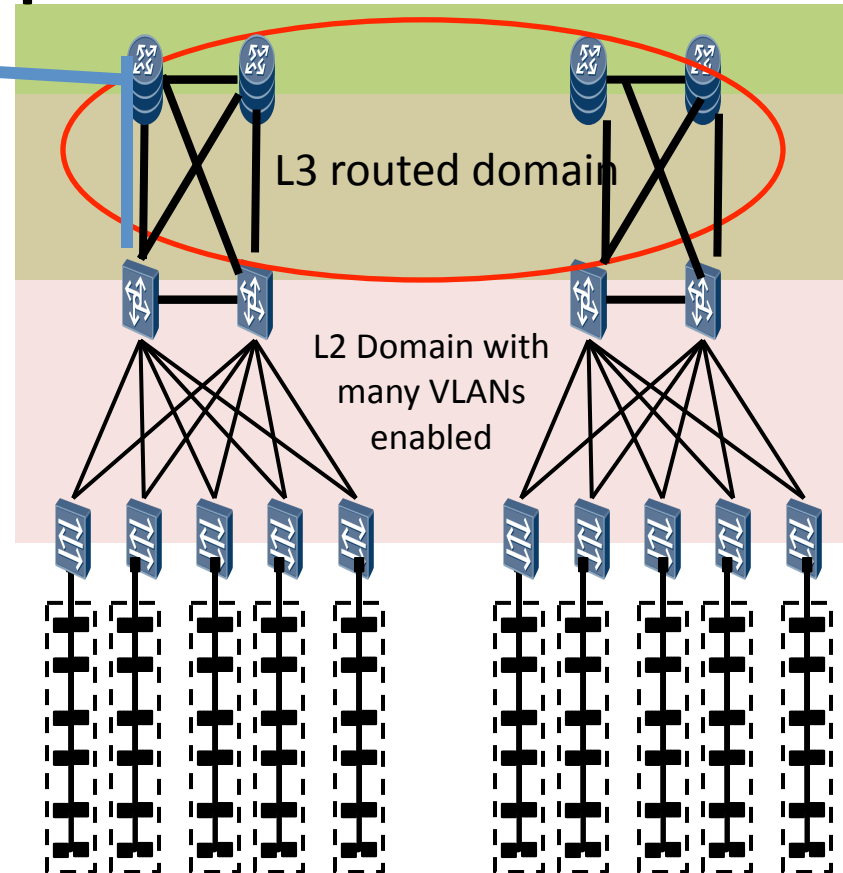
→ CPU & buffer intensive.

L3 routed domain

L2 Domain with many VLANs enabled
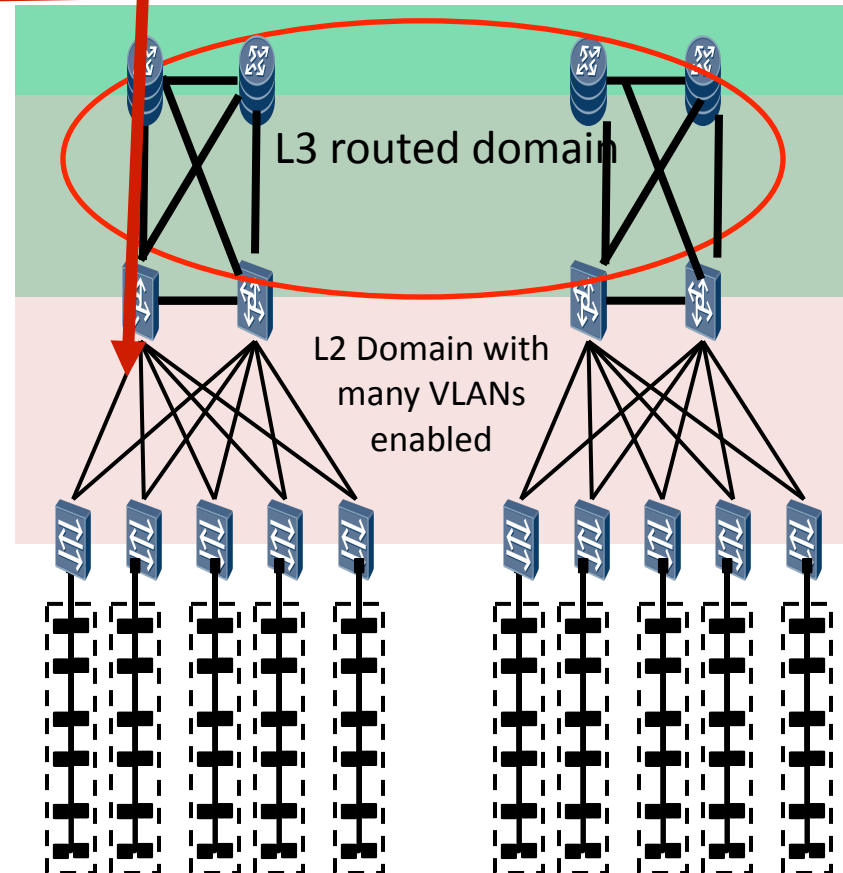
Pain Point #B for Scenario #2:
When internal hosts need to communicate with external peers

Hosts send ARP/ND to default gateways frequently

IPv4 solution: frequent gratuitous ARP by gateway.

IPv6: none.

L3 routed domain

L2 Domain with many VLANs enabled

Pain Point #C for Scenario #2:
Hosts in two different subnets communicate with each other within data center

Gateway router is impacted twice:
- one for applications in subnet-A initiating ARP/ND request to the gateway (#1 above), and

- the second for the gateway to initiate ARP/ND requests for the targets in subnet-B (#2 above).
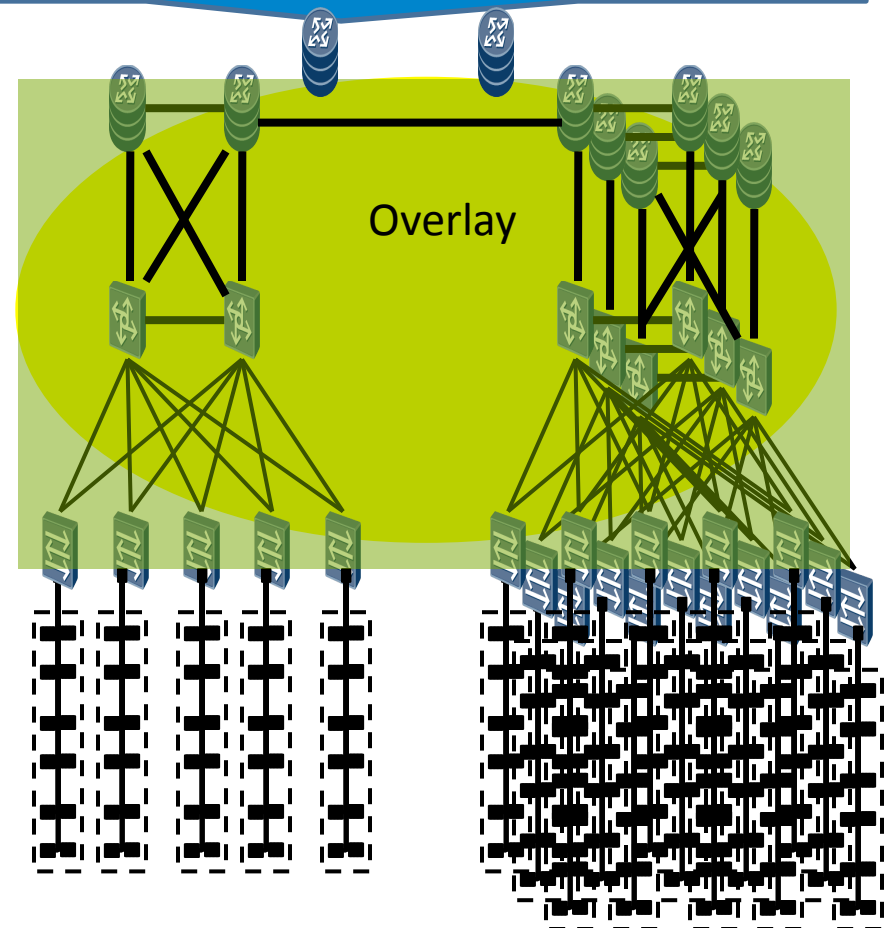

L3 routed domain

L2 Domain with many VLANs enabled
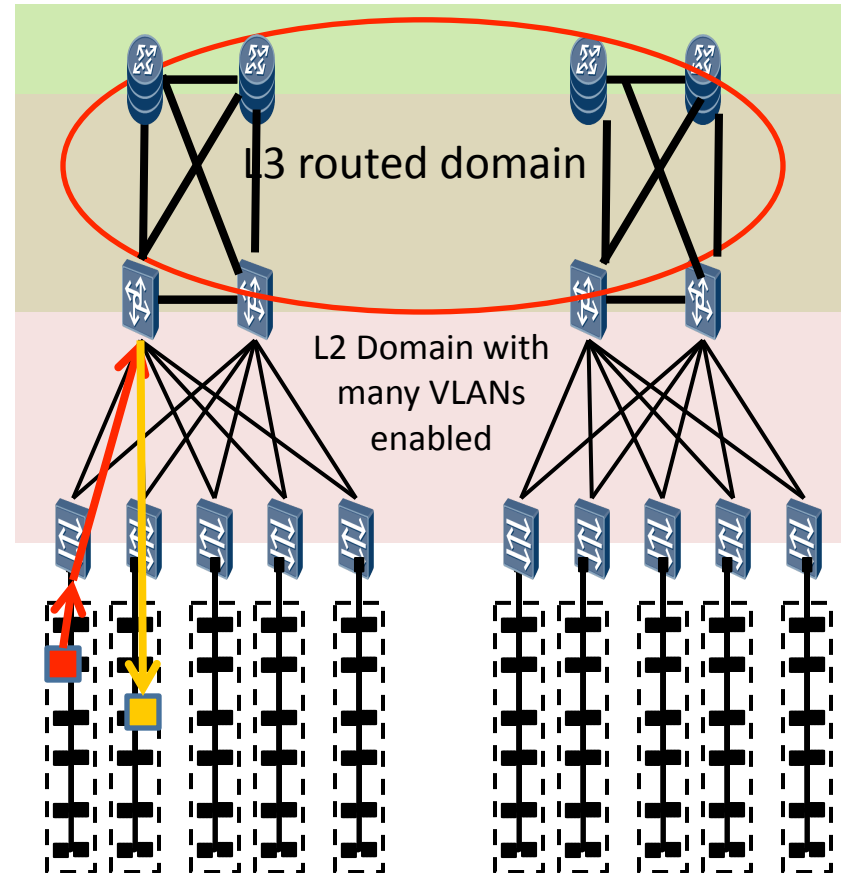
Overlay Network
Hosts have different addresses than network addresses

When external peers communicate with internal hosts:
   Gateway routers have to resolve target address, plus Network Edge node

- ToR or Hypervisor perform network address encapsulation
  - GRE encapsulation
  - TRILL
  - MAC-in-MAC
  - http://tools.ietf.org/html/draft-mahalingam-dutt-dcops-vxlan-00
  - http://tools.ietf.org/html/draft-sridharan-virtualization-nvgre-00
  - http://tools.ietf.org/html/draft-wkumari-dcops-l3-vmmobility-00

Overlay

Wanting your feedback!

http://tools.ietf.org/wg/armd

Thank you!