IP SHOWCASE™ THEATER

CURATED BY VSF VIDEO SERVICES FORUM

Red & Blue, or Purple
Your network, your way

Gerard Phillips, Systems Engineer, Arista Networks

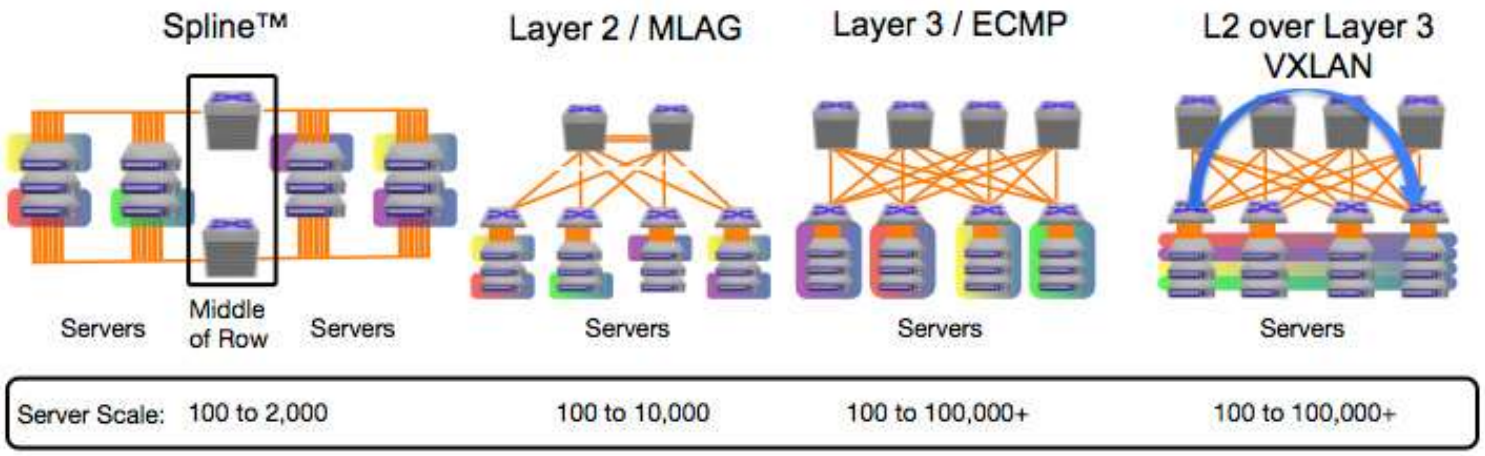
gp@arista.com, +44 7949 106098

ARISTA

IP SHOWCASE™ THEATER

What we'll cover

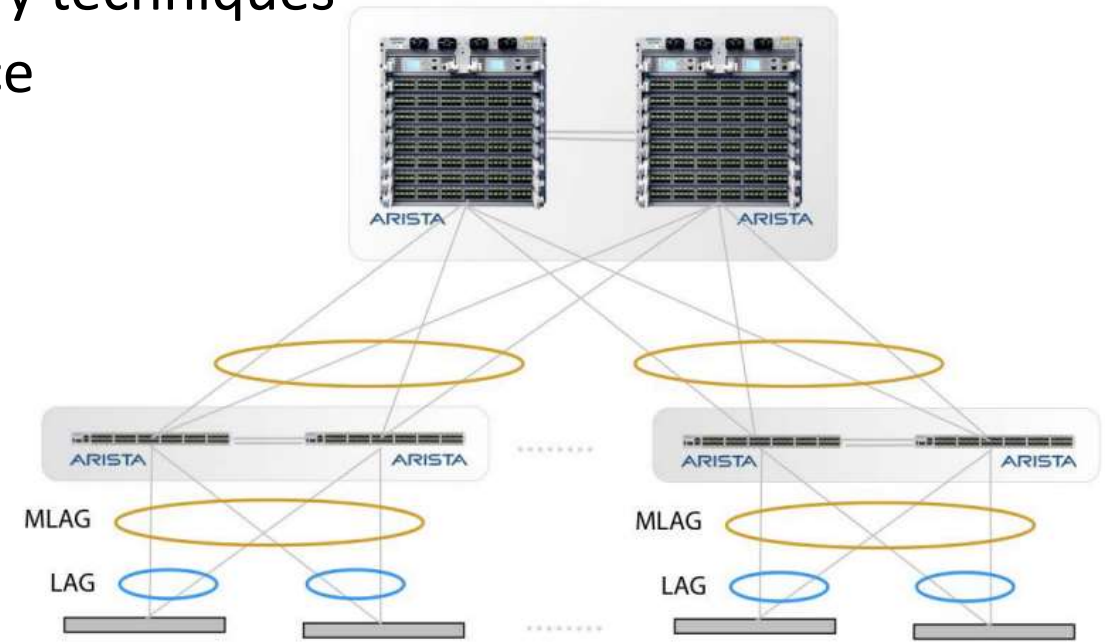
- Architectural Overview – L2 vs L3
- Designing for Resilience
- Architecture Options
  - Monolithic
  - Spine and Leaf - Hybrid
  - Spine and Leaf – Purple
- Conclusions



ARISTA

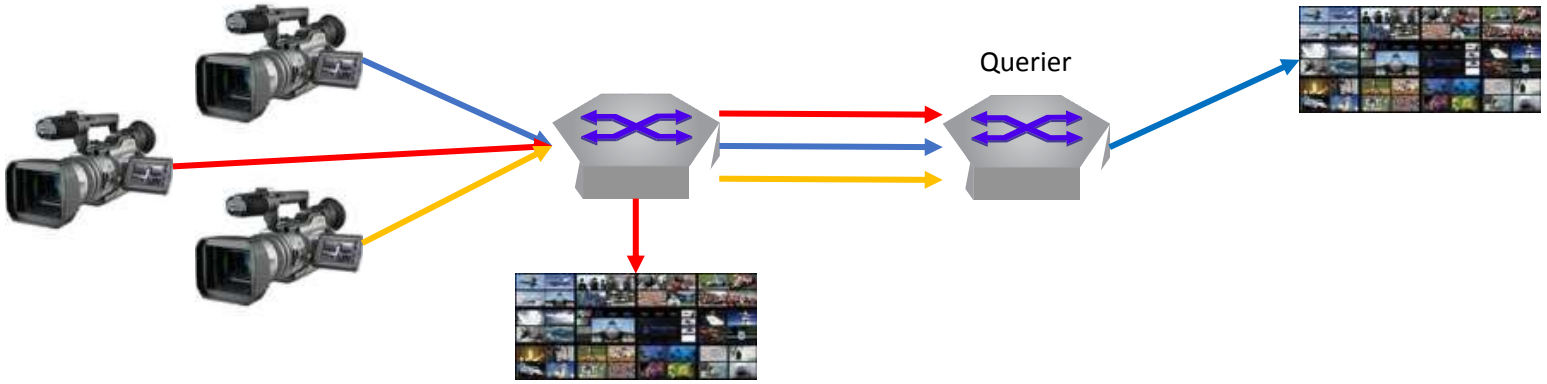
Architectural Overview – L2

- L2 networks typically deployed for audio installations
  - Low bit rates
  - Undersubscribed networks
  - Control systems used L2 scoped discovery techniques
  - MLAG provides scale, and spine resilience



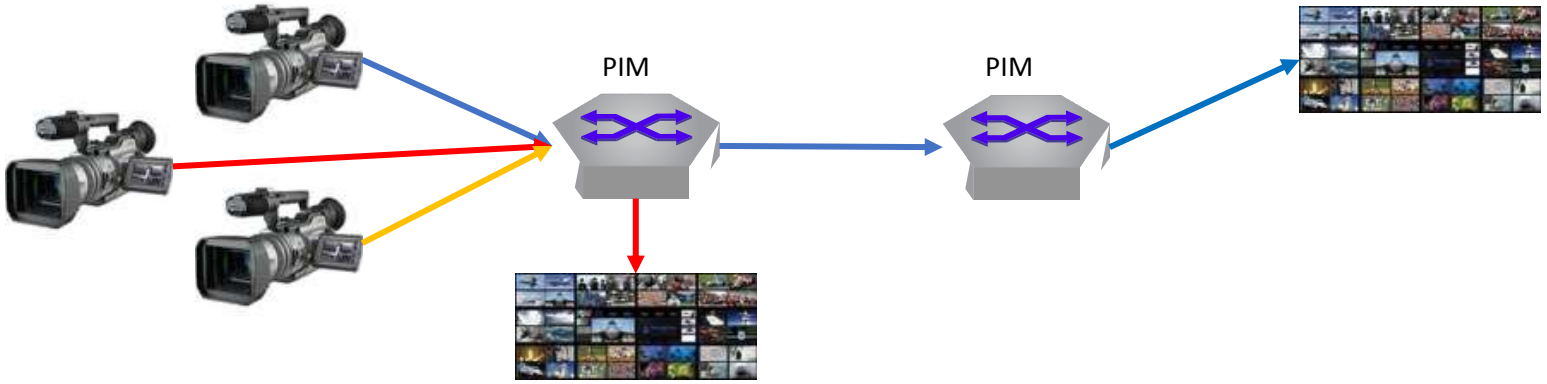
Architectural Overview – L2

- But L2 does not work for Live Production, high bit rate multicast
  - MLAG complex to configure for ASM Multicast (*,G) in a -7 environment
  - Flows originated in remote switches are flooded towards the querier
  - This potentially requires very large pipes!
  - The failure domain is very large
- You are also limited to 2 spines – potentially limiting scale



Architectural Overview – L3 is the answer

- This is the Datacenter architecture for scale and flexibility
- PIM allows multicast to be routed
- Failure domains are now able to be much smaller
- Flooding towards the querier is no longer required
  - Broadcast Controllers can be in charge of what transits any inter switch links



Unicast routing for L3

- Flexibility is one of the great benefits of the move to IP
- To facilitate this flexibility, we need a solid unicast routing capability
- This will under-pin any IGMP/PIM based multicast routing
- BUT, can provide security, control, resilience and flexibility
- Static routing can be used, but does scale…
  - Manually provisioning routes is error prone and slow
- BGP is the DC choice, scalable, fast convergence flexible, future proof
  - But other dynamic routing protocols are available – OSPF, ISIS etc.

IP SHOWCASE™ THEATER

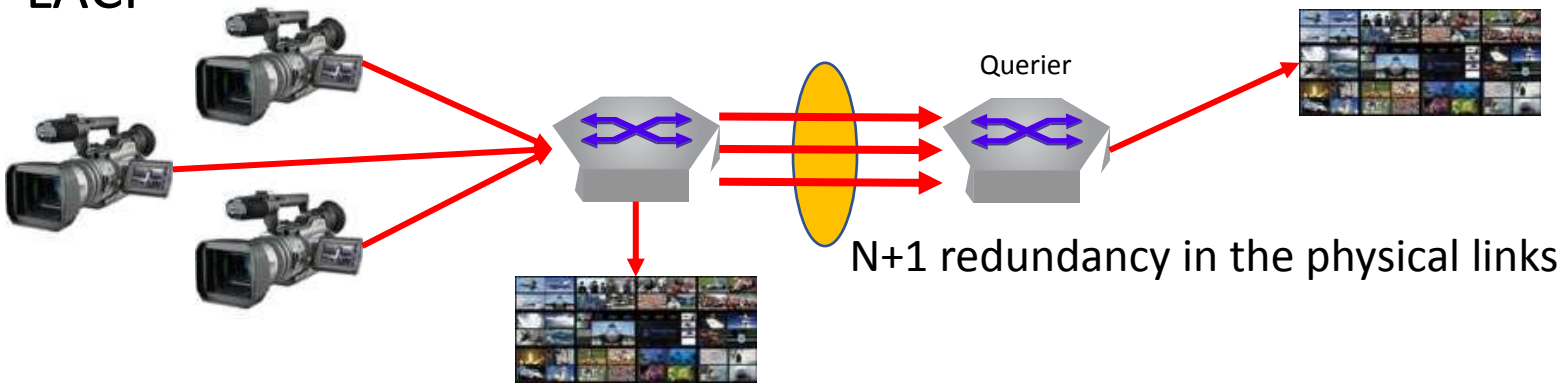
Designing for resilience

- Determined by how many failures your system should tolerate
- -7 Hitless merge provides the **capability** to provide:
  - RTP identical flows, on physically diverse NIC's
  - Physically diverse transport – optics, fibre, DAC, AOC etc
  - Physically diverse IP fabric
- You can survive the first failure, assuming you have a robust monitoring system that can provide quick, accurate, actionable info
- You also have a path to planned maintenance, upgrade, addition of new services etc

ARISTA

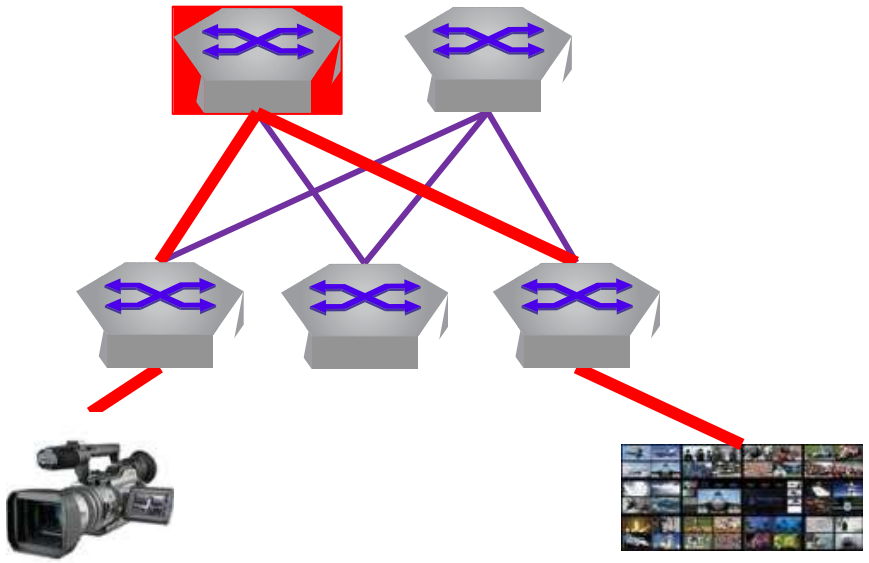
Designing for resilience – the 2nd failure

- How do we survive the 2nd failure?
- Choose quality components – switches, NOS, optics, fibre etc.
- Design in redundant PSU's, Fans, Supervisors, Fabric Modules
- Design in redundant Links between switches – N+1 or more
- Ensure routing protocols, or SDN can, and will use these "spares"
  - ECMP + LACP



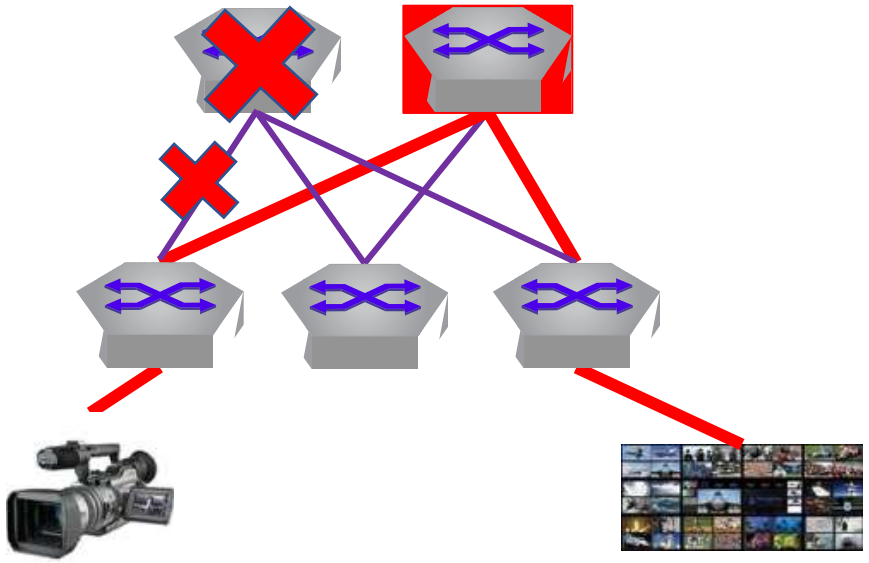
Designing for resilience – smaller failure domains

- Apply this physically as well as logically
- Monolithic switches allow line-cards, fabric modules & supervisors to be replaced in service



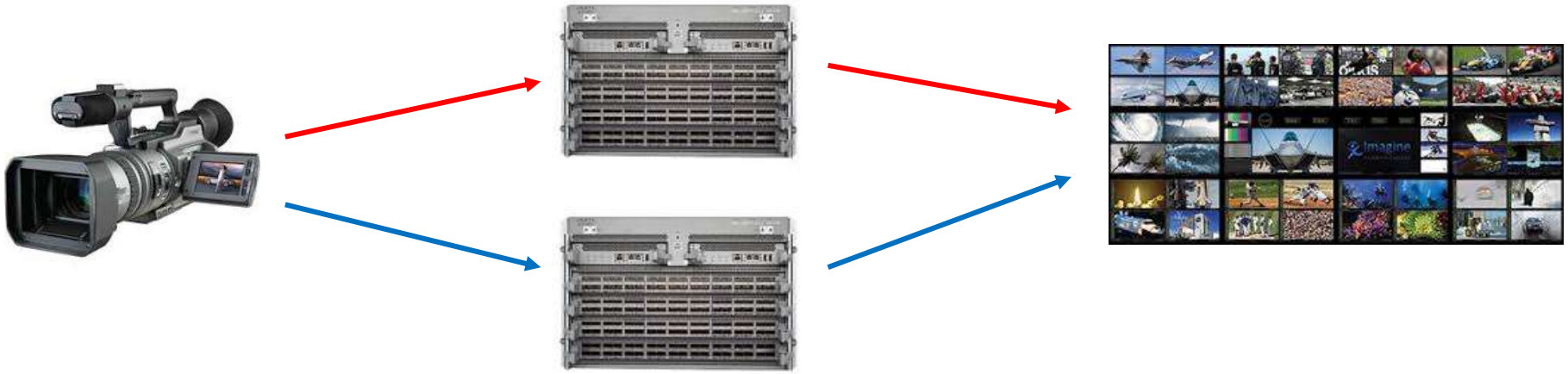
Designing for resilience – smaller failure domains

- Leaf and spine architectures allow you to manage smaller chunks:
  - Route around failed components
  - Route around devices under maintenance
  - Influence multicast routing tables
  - SDN



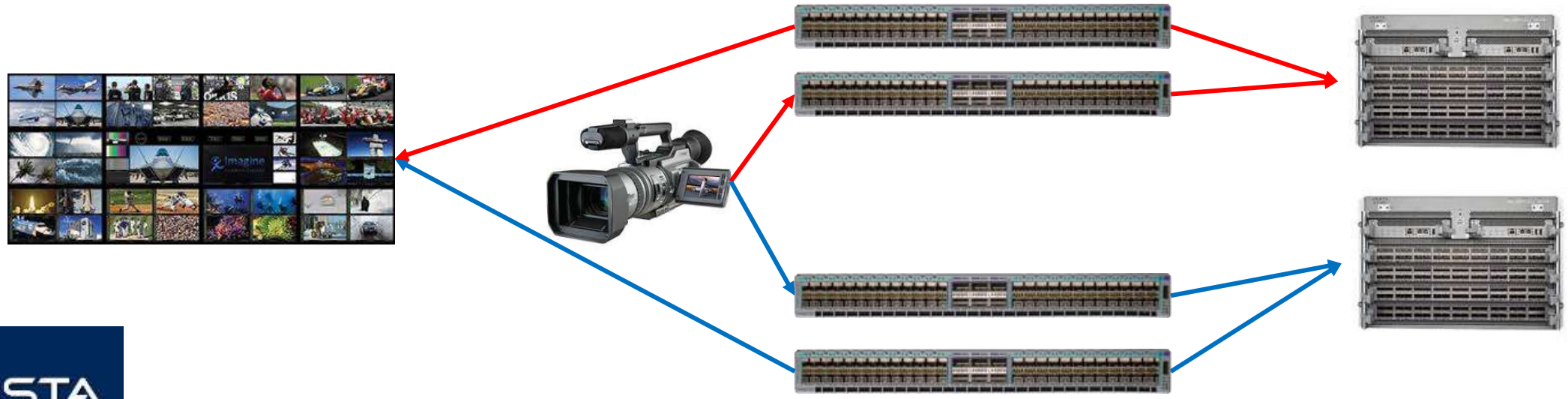
Monolithic

- Simplicity. Hitless merge -7 resilience
- SDN / flow orchestration is not necessary, IGMP can be used very successfully.
- Monolithic chassis' solutions can scale up to 16K[2] @ 3Gbe or 2304 hosts @ 25Gbe
- Redundancy is provided by 2 (essentially) air-gapped switches, redundant fans and PSU's, and optionally redundant switch supervisors



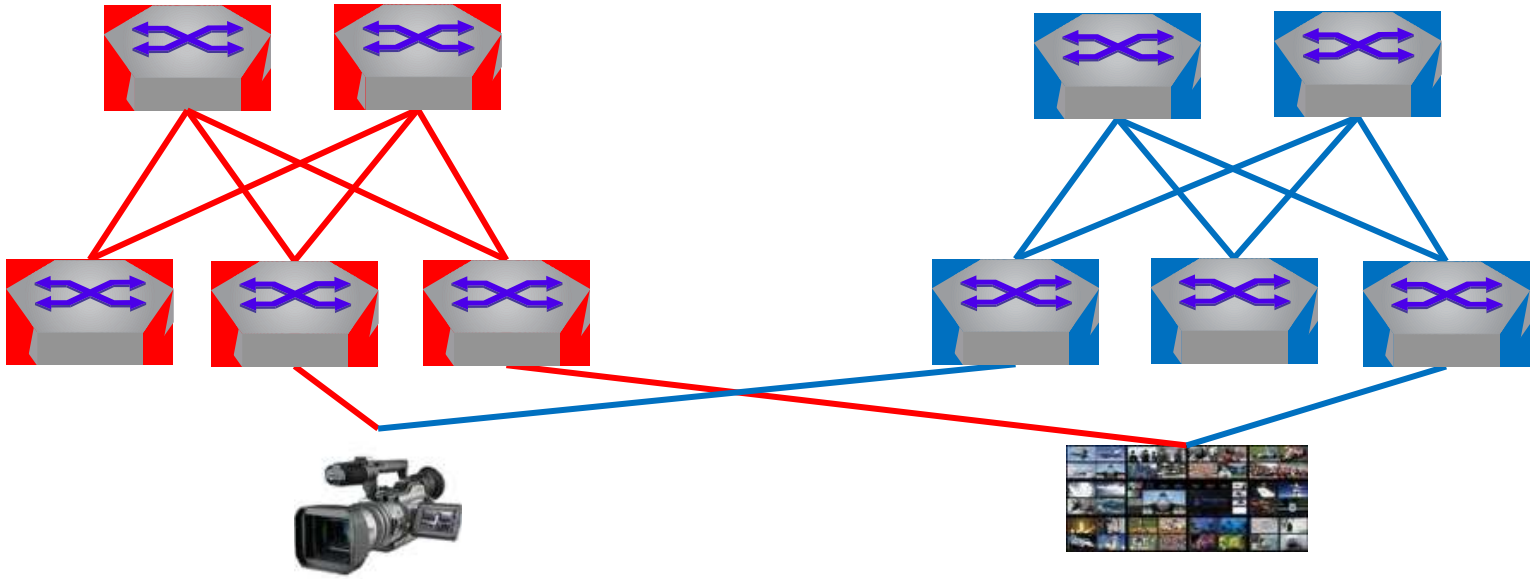
Monolithic - Expansion

- While this architecture is simple, it does have a scale limit
- Future expansion can build on a monolithic base, by using the monolithic switches as spine devices, adding SDN/orchestration, and hanging leaves from the “spine”
- This path opens up higher levels of future expansion, but provides a simple start point



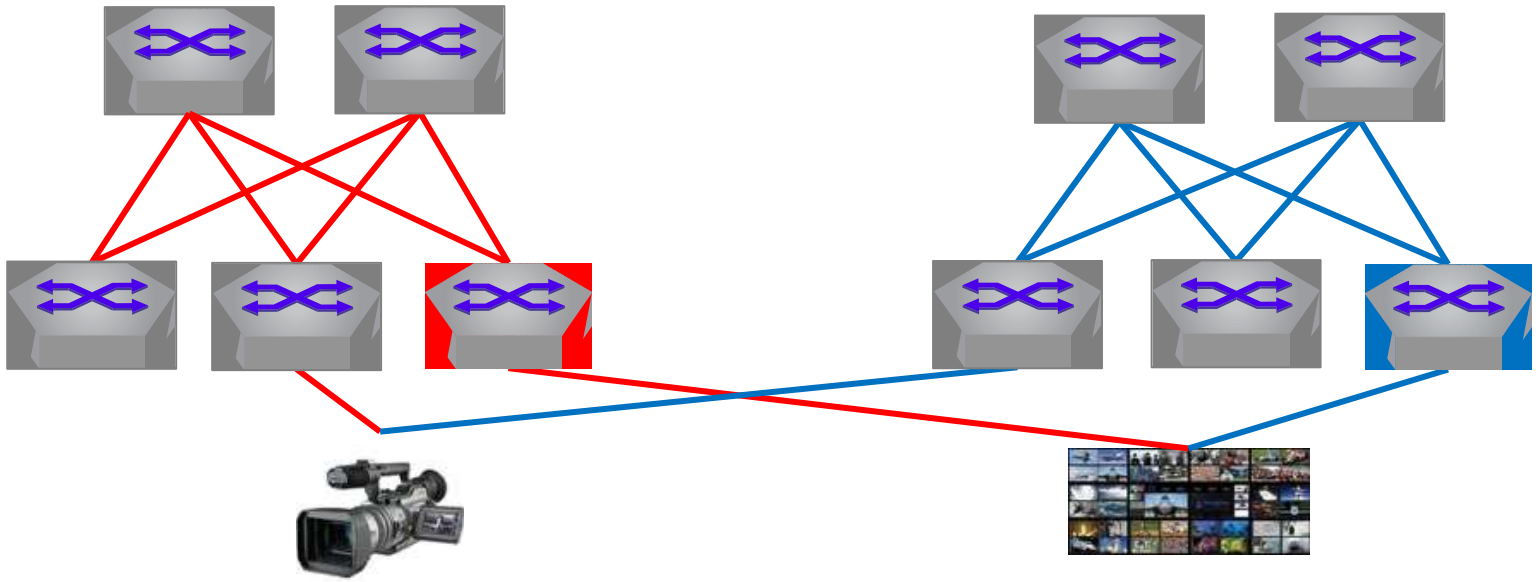
Spine and Leaf – Air-gapped Red and Blue

- L3 topology for cloud scale - supports future expansion
- Air-gapped provides flow security (-7)
- BGP routing for fast and reliable unicast convergence
- PTP Boundary Clocks in L&S provides scale and accuracy
- A Flow Orchestrator or SDN system is needed
- Simple -7 resilience still available
- Simple leaf pair could be a starting point!



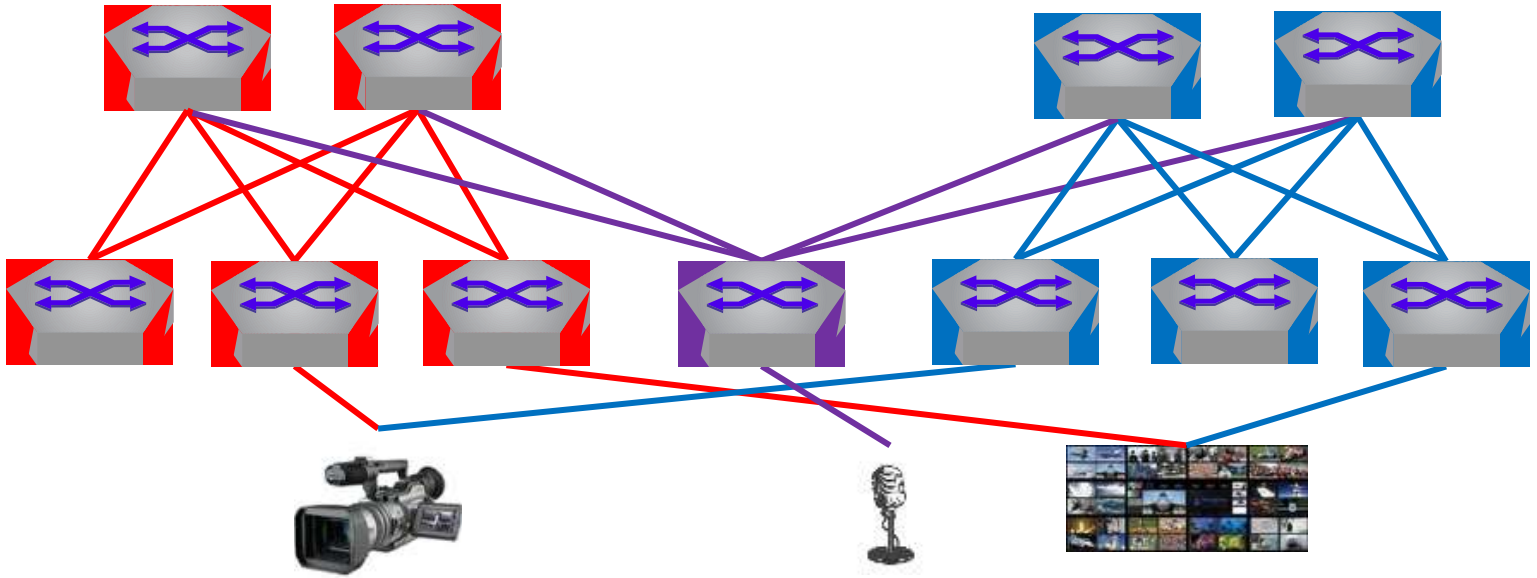
Spine and Leaf – Air-gapped Red and Blue

- L3 topology for cloud scale - supports future expansion
- Air-gapped provides flow security (-7)
- BGP routing for fast and reliable unicast convergence
- PTP Boundary Clocks in L&S provides scale and accuracy
- A Flow Orchestrator or SDN system is needed
- Simple -7 resilience still available
- **Simple leaf pair could be a starting point!**



Spine and Leaf – Air-gapped Red and Blue (Hybrid)

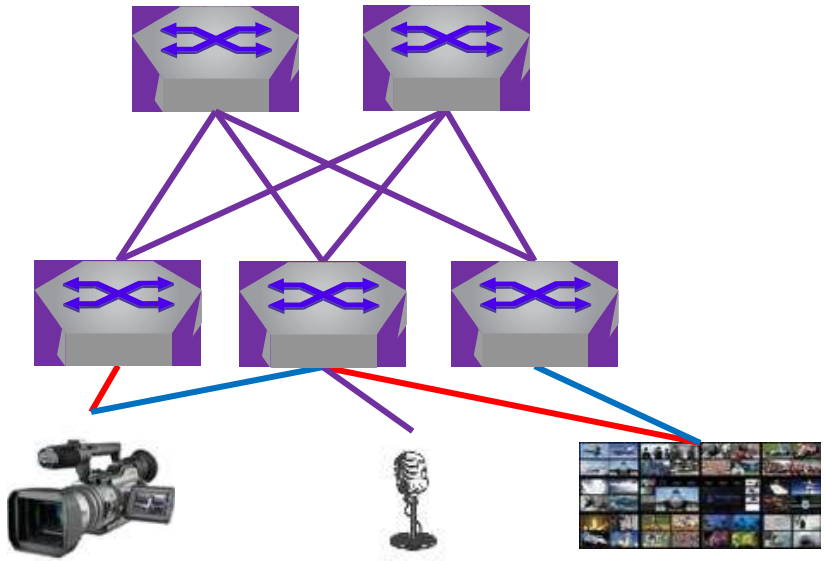
- Purple switches support single homed devices
- Add as many ”purple” switches as you need
- This architecture requires an SDN controller, BUT the dedicated Red/Blue spines make it a simpler device



Spine and Leaf – Purple

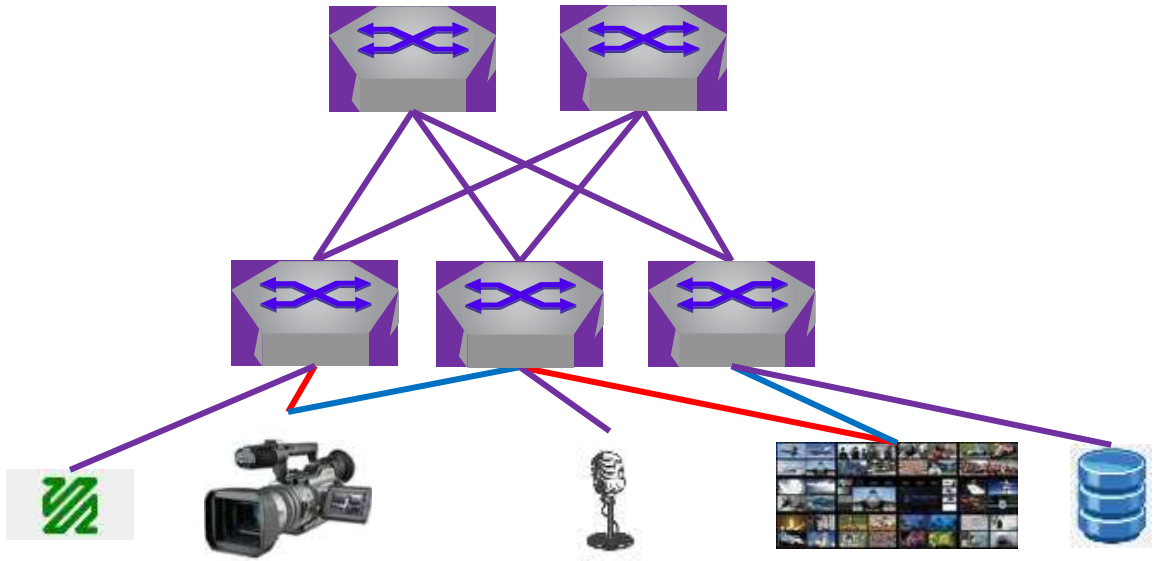
- L3 topology for cloud scale - supports future expansion
- Flow security (-7) provided **logically**, not physically
- BGP routing for fast and reliable unicast convergence
- BC PTP in both L&S provides scale and accuracy

- Any switch can support single homed devices
- A Flow Orchestrator or SDN system is needed
- Orchestrator is more complex than Red+Blue case
- Simple leaf pair could be a starting point!



Spine and Leaf – Purple

- L3 topology for cloud scale - supports future expansion
- Good starting point for a converged network later



IP SHOWCASE™ THEATER

Conclusions

- Choose your architecture for your needs
- Choose SDN or IGMP/PIM to solve your workflow challenges
- Choose Cloud Scale IP infrastructure
  - Provides many layers of resilience:
  - Focus on Quality = Reliable SW/HW = low TCO + high uptimes
  - Don’t let monitoring be an afterthought!
  - L3 provides this reliability and resilience at scale
    - .... and limits the failure domain size
  - Build in reliability, with redundancy
    - -7 Hitless merge
    - Redundant links (N+)
    - Resilient IP protocols – BGP, ECMP

ARISTA

IP SHOWCASE THEATER™

CURATED BY VSF VIDEO SERVICES FORUM

Thank You

Gerard Phillips, Systems Engineer, Arista Networks

gp@arista.com, +44 7949 106098

ARISTA