SHARKFEST '14

WIRESHARK DEVELOPER AND USER CONFERENCE

JUNE 16-20 2014 · DOMINICAN UNIVERSITY

I1: Best Practices for Packet Collection, Aggregation & Distribution in the Enterprise

J. Scott Haugdahl

Architect, Blue Cross Blue Shield MN; scott_haugdahl@bluecrossmn.com

Formerly Asst. VP, Architect, and NAA founder at US Bank

(NOT a best practice!)

The Blue Cross Blue Shield Experience

- Who is Blue Cross Blue Shield of Minnesota?
  - The first “Blue” health plan in the nation & the largest insurer in Minnesota
  - 2.6 million members across all 50 states, 3,500 employees, nonprofit
  - Administrative cost less than 10 cents on the dollar, among lowest in the country
- What is Enterprise Systems Management (ESM)?
  - A group that collectively manages enterprise wide performance metrics and reporting from system level to multi-tier application performance
  - Built up in recent years with strong support from senior leadership
  - Owns event monitoring and reporting, capacity planning, **Shared Visibility Fabric (SVF)**, packet level analysis tools, Application Performance Monitoring (APM)
- Why the SVF?
  - ***Created a third generation and scalable packet collection, aggregation, and distribution system with flexible mapping rules, fabric services, end-to-end virtual connectivity, and scalability for new data center growth***

The US Bank Experience

- Who is US Bank (symbol: USB)?
  - 5th largest U.S. commercial bank, 3,100 branches, 67,000 employees, $364B assets
  - Recognized for its strong financial performance and prudent risk management
- What is Network Application Analysis (NAA)?
  - Founded in 2008 as part of US Bank's Network Planning and Engineering to adapt new thinking around methods, tools, process, and collaboration in order to focus on resolving potential or chronic application performance problems
  - Solutions oriented, not only the lower network layers 1 – 3 (i.e. infrastructure)
  - Gained credibility during pre-migration analysis to a new data center and created a unique opportunity to architect a large **Shared Data Access Network (SDAN)**
- Why the SDAN?
  - ***The only solution able to collect and aggregate multiple streams simultaneously from several tiers in real-time to feed Application Performance Monitoring (APM), fraud detection, IDS, and sniffer tools***

Packet Collection, Aggregation, & Distribution

… Is comprised of

1. Packet sourcing via taps and SPANs
   - Physical in-line media taps (fiber and copper)
   - Switch mirror ports (SPAN), blade chassis mirror ports (blade mirroring), firewall mirror ports, etc.
2. Switching fabric of intelligent packet aggregation matrix switches
3. Configuration and control software

… Is NOT

Sniffers, IDS appliances, Application Performance Management (APM) nodes, fraud monitoring tools, etc.

… These are the consumers of the sourced packets

Also referred to as a Data Access Network, Visibility Fabric (Switching Array), Network Packet Broker, Monitoring Switch Fabric, etc.

Best Practices

- Include the word *shared* when branding your monitoring fabric
  - Using Shared Data Access Network, Shared Visibility Fabric, etc. sends a strong message across the enterprise & leads to collaboration
- Separate out the consumers when budgeting and forecasting
  - Adding costs of large capacity packet capture appliances can skew your capital expenditure (CAPEX) and long term maintenance (OPEX) which may be perceived as a monitoring fabric cost
- Create an NAA group or broad spectrum ESM team
  - Need to bridge the gap between network and application teams for problem determination and resolution
  - Determining the right methods to capture and process packet flows is critical and tool dependent
  - Required unless you own all of the capacity and performance tools
  - Tough politically as managers tend to protect the status quo
  - Requires recognition and strong support from senior leadership

Shared Monitoring Fabric
Selling It – The Big Three

1 **Stream Sharing**

Stream sources (ingress or **network ports**) can service many consumers (egress or **tool ports**) critical to protecting your customers and improving the end-user experience

2 **Multi-tier Stream Aggregation**

Several streams from multiple tiers can be **aggregated** to one or more outputs, in order to monitor complex applications and save on tool ports – the so-called “one to many” and “many to one”

3 **Filtering**

Streams can be filtered by MAC, IP or other criteria, allowing **focused** analysis or specific web front ends with a significant drop in resource requirements on the tool or appliance

The Dark Ages

"Technicians had to physically unplug and move tools from one tap or SPAN port to another. That necessitated change orders and scheduling during off hours, slowing the group's agility and flexibility to monitor effectively."

- Royal Bank of Canada

***NOT** a best practice!*

Some Early Challenges

- Sharing Cisco SPANs (or R/ER* RSPANs) was a big problem
  - Contention, prioritizing, and managing across multiple teams and analysts
  - Limited span ports, typically 2 max per switch
  - SPAN technology has not kept up with switching technology, such as port channels
- 10 Gbps was about to explode
  - How to handle all that data?
- Cisco NAM's were EOL and too limited
- Security was using dedicated taps and mirror ports
- Blade servers were quickly emerging – lack of visibility
- Applications were growing in number and complexity
  - Myriad of infrastructure and application tiers, as well as asymmetrical routing often requires multiple simultaneous data stream capture points
- First generation physical layer matrix switches were not suitable for emerging technology and requirements
  - Basic "A-B" switch
  - How do you scale to 10 Gbps with only layer 1 control and visibility?

* Not a best practice.

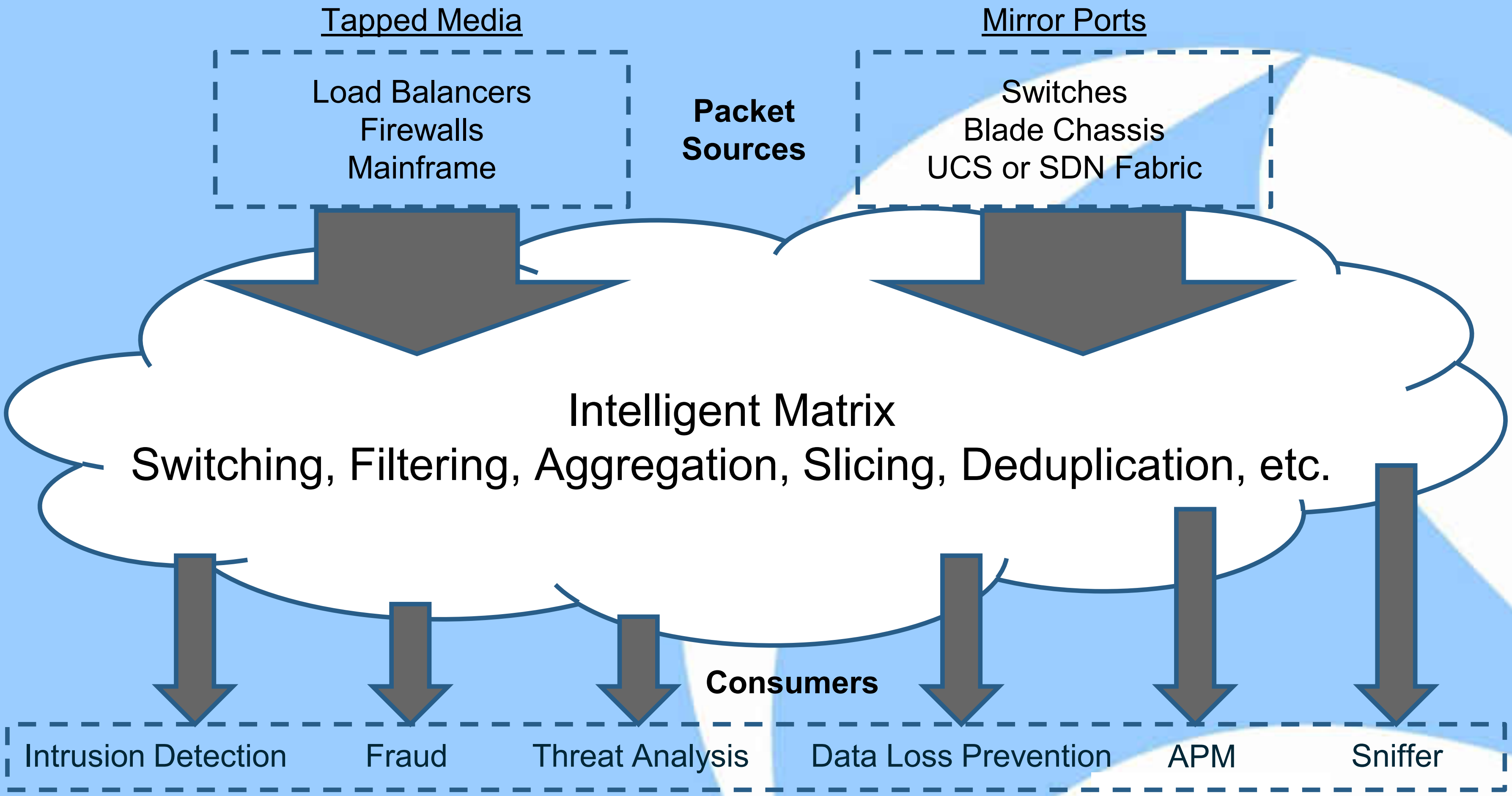
Sharing SPANs got ugly!

***NOT** a best practice!*

Best Practices

- Set goals up front such as…
  - Ease demand for SPAN ports
  - Allow for ad-hoc as well as permanent monitoring
  - Share packet streams with analysis, security, compliance tools, and ?
  - Architect from the ground up for a new data center or migration?
- Define specific requirements for a POC such as…
  - Can handle a variety of physical media types (copper and fiber) and data rates (1, 10, 40, 100 Gbps?)
  - Able to aggregate streams to higher speed or load balanced output ports
  - Filter streams by MAC, IP, IP + Ports, VLAN, pattern match, etc.
  - Able to stack or cluster the matrix switches to work and as one unit
  - Provide a central point of management and access control
  - Grow to handle up to 1,000 ports, roughly 10:1 network:tool port ratio
  - Derive a lab test plan for cross-blade & cross-chassis packet blasting

Fast Forward: The Modern Shared Monitoring Fabric (SMF)



Tapped Media

Load Balancers
Firewalls
Mainframe

Packet Sources

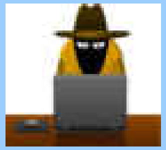
Mirror Ports

Switches
Blade Chassis
UCS or SDN Fabric

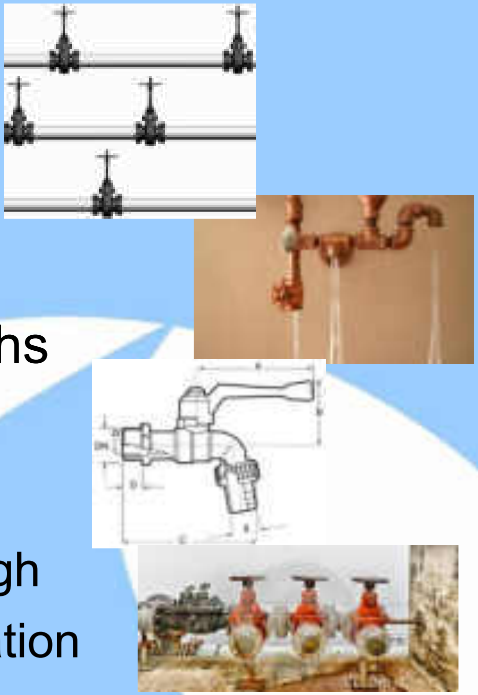
Intelligent Matrix
Switching, Filtering, Aggregation, Slicing, Deduplication, etc.

Consumers

Intrusion Detection | Fraud | Threat Analysis | Data Loss Prevention | APM | Sniffer



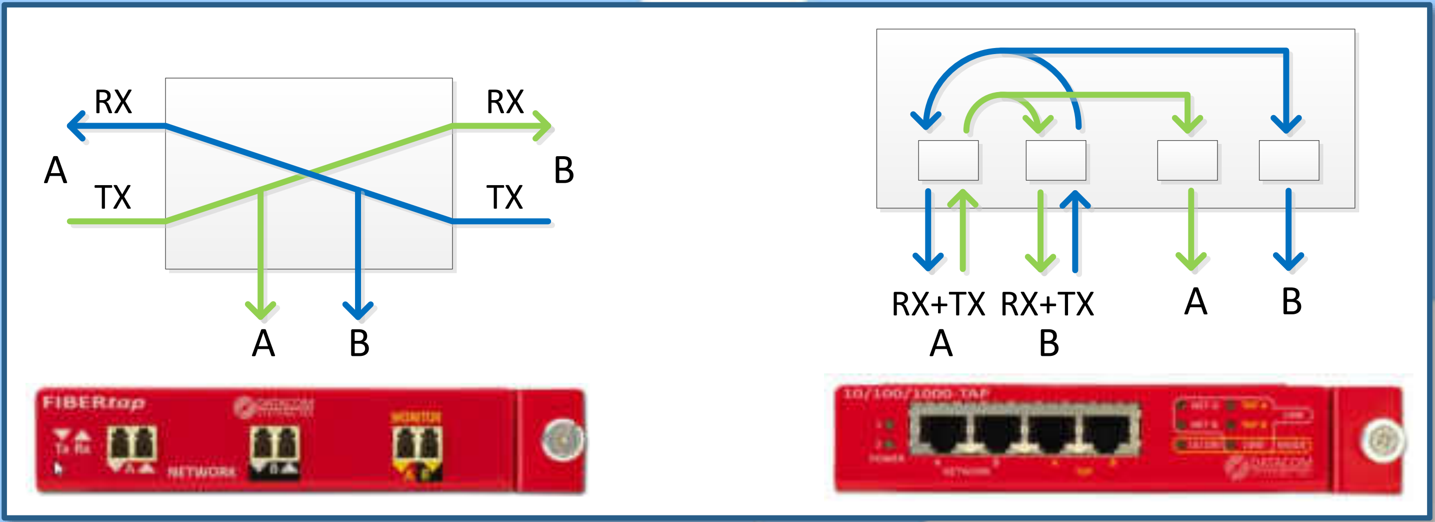
A 60 Second Tap Primer



- Fiber taps are “passive” and split the light into two paths
  - Most are a 50:50 split ratio & thus distances are halved
- Copper taps are regenerative and thus require power
  - Internal mechanical relays provide power fail pass through
  - Many options like link-state propagation, packet aggregation (with buffering), statistics, manageability

Fiber

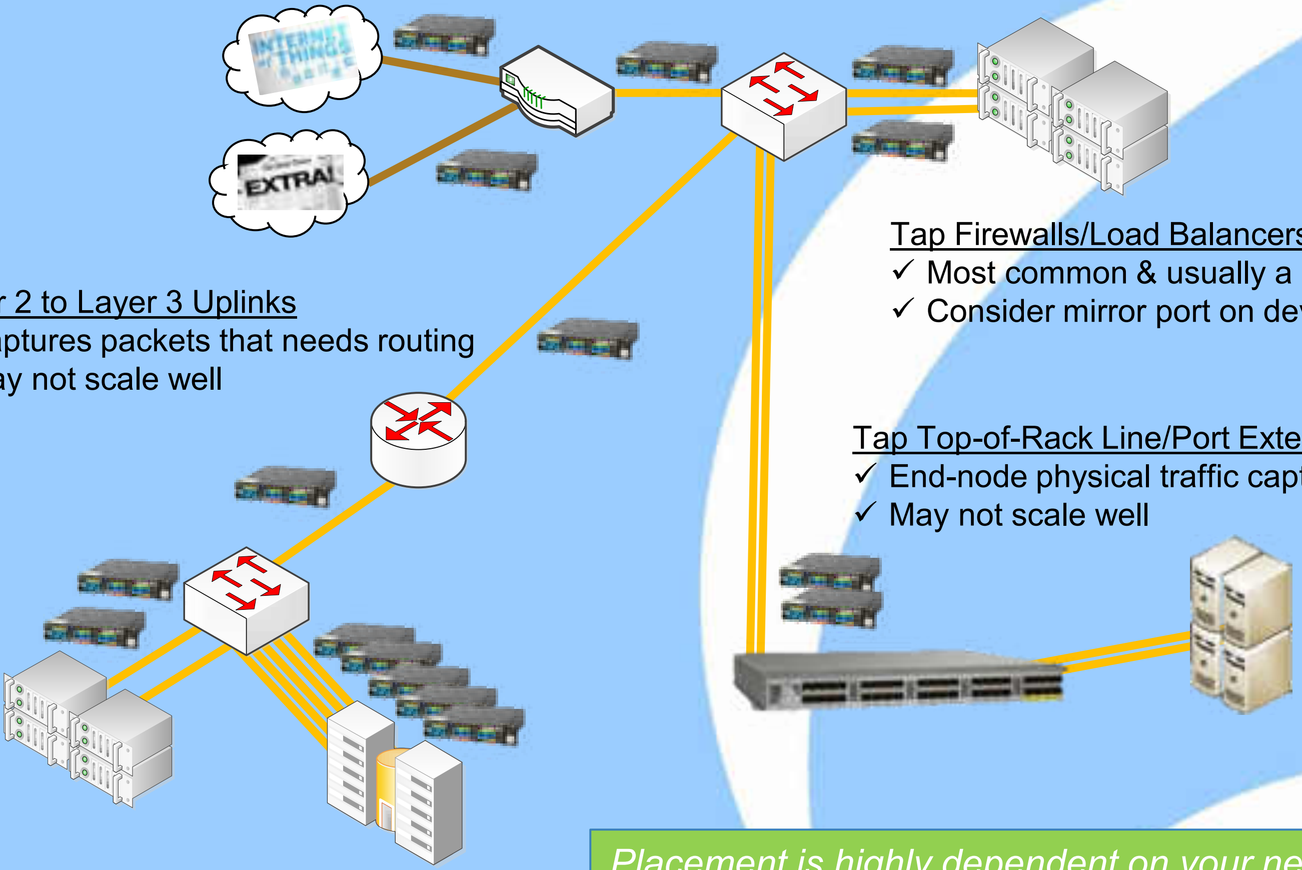
Copper



Best Practices

- SPANs (and mirror ports) usefulness is diminishing, so avoid if possible
  - Easy to over subscribe, especially with port channel or full duplex aggregation
  - Eliminate the old practice of using aggregation taps and use fiber where possible
  - *Be mindful that each one tap takes up two monitoring ports when operating in non-aggregation mode*
- Tap related network points into high density traffic aggregators (**TA**) and send aggregated flows to core cluster (**CC**) for tool consumption
  - Perimeter Firewall Taps →**TA** → Firewall Aggregate Uplink → **CC** → IDS
  - Top of Rack (TOR) Cisco 2232/2248's Taps → **TA** → Server Farm Aggregate Uplink → **CC** → Fraud Detection
  - Mainframe OSA Taps → **TA** → Mainframe Aggregate Uplink → **CC** → Auditing Tool
- For branch WAN connections, Consider preserving separate send/receive full duplex tap ports all the way through to your tools
  - Preserving full duplex tapped connections from taps to tools helps to preserve incoming vs. outgoing traffic

Tap Placement Considerations



Tap Inside and Outside WAN Routers

- ✓ Outside may be subject to carrier approval

Tap Firewalls/Load Balancers/IPS/WAF/Web Proxies

- ✓ Most common & usually a best practice
- ✓ Consider mirror port on device (not switch) if supported

Layer 2 to Layer 3 Uplinks

- ✓ Captures packets that needs routing
- ✓ May not scale well

Tap Top-of-Rack Line/Port Extender Uplinks

- ✓ End-node physical traffic capture
- ✓ May not scale well

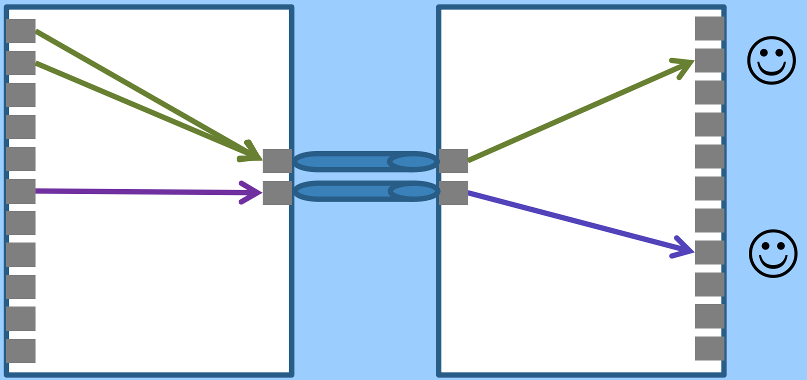
Placement is highly dependent on your needs and architecture.

High Density Server and MF Interfaces

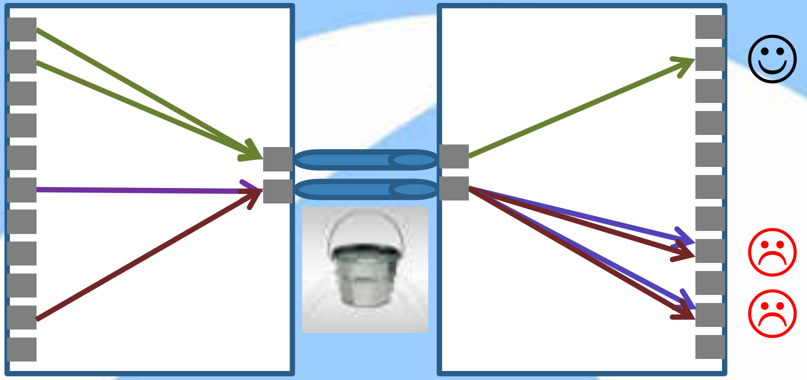
- ✓ May not scale well for distributed
- ✓ *May still need a virtual server tap solution*

Shared Monitoring Fabric Switches Cascading vs. “One Pane of Glass”

Cascading

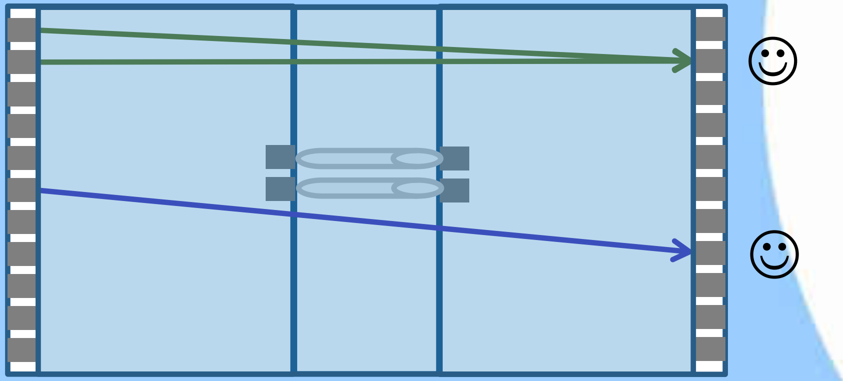


Connections are managed switch-by-switch.

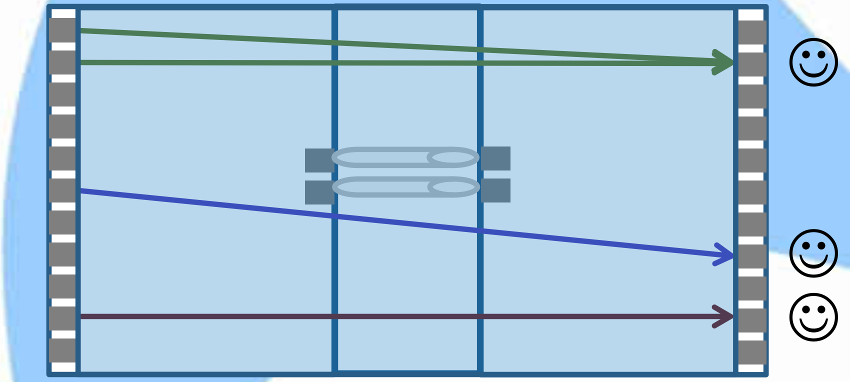


A cascaded fabric can get messy very quickly, especially attempts to filter and sort traffic.
Adding more uplinks does not scale well.

Stacking/Clustering– Best Practice!



The system becomes “one pane of glass” such that we only need to specify the end ports.



Far easier to manage but still need to be conscious of interconnect bandwidth; imagine a stack of 30 units!

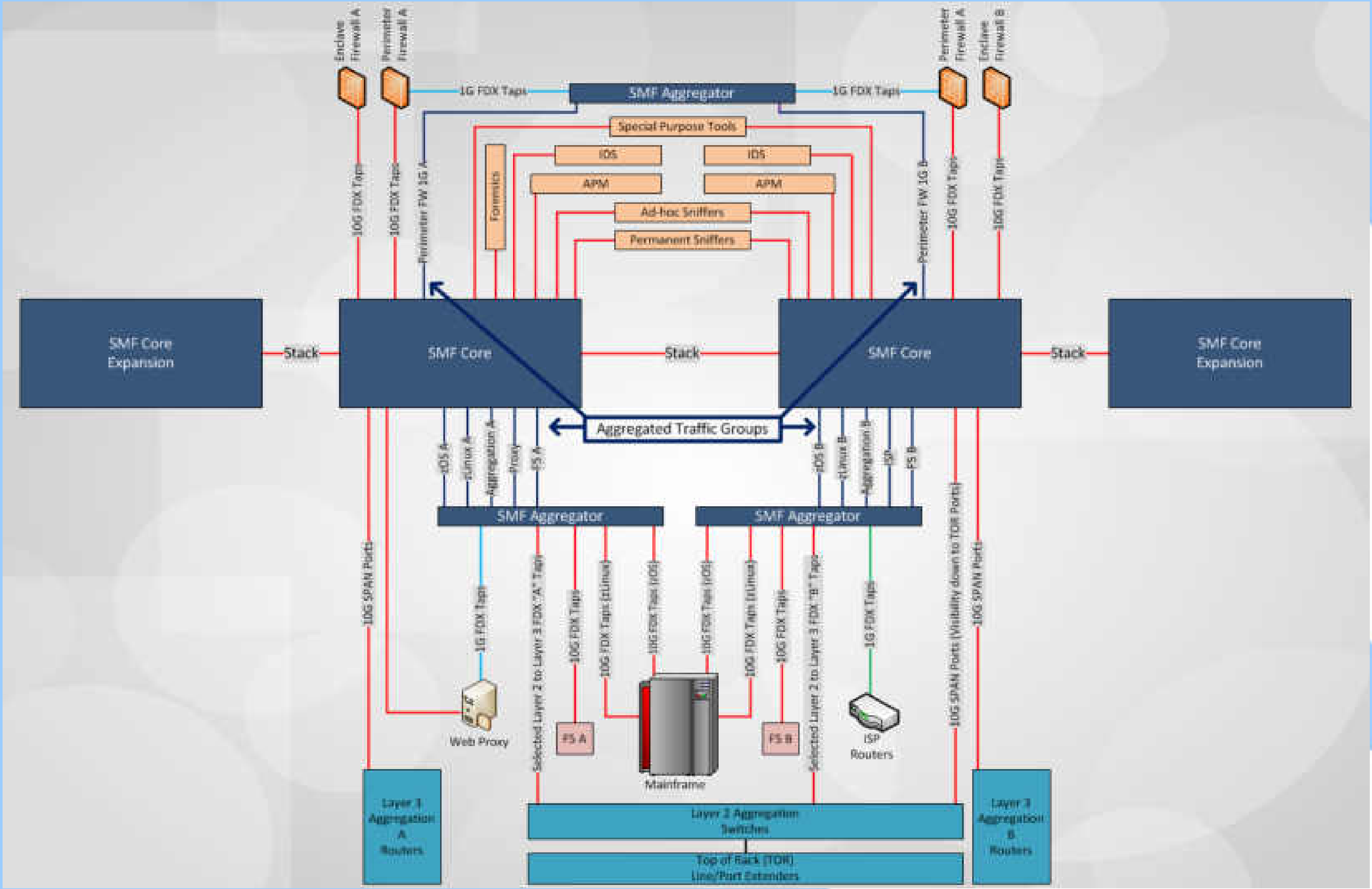
Best Practices

- Use less expensive highly concentrated traffic aggregators to consolidate “like” packet sources (such as perimeter firewall interfaces, mainframe OSAs, etc.), then channelize the groups to uplinks and use full fabric stacking or clustering at the core
  - Aggregators with built-in taps are good IF you can double back to the source vs. more optimal near mid-point tapping
- Use rules and filtering to greatly reduce load on the appliance
  - Security and APM appliances do not need to waste cycles filtering irrelevant data
  - Reducing unnecessary intake can also increase post processing performance
- Copy APM flows to permanent sniffers for tool validation & post mortem analysis
  - Also feed security tool flows to your sniffer to validate setup and operation

And last but not least…

Connect high performance, high volume sniffers for data capture and selectively mine packets to Wireshark for analysis!

Sample Shared Monitoring Fabric (SMF)



Bonus: More Best Practices!

- Use a naming convention that indicates the names of devices and their exact ports from/to which packets are entering/exiting
  - Ensure that such information is self-evident without consulting an external resource
- Design to eliminate packet loss in the monitoring fabric
  - Over subscription often due to excessive aggregation or uplink capacity
  - *If packets go missing there, you've got an unreliable and potentially misleading view of your network not to mention credibility*. Corollary: monitor for and alert on packet drops and breached utilization thresholds
- **Every** vendor's filtering implementation has gotchas. Understand them!
- Obvious: Always validate that the filtering/dispatching changes you make are working correctly!

How do we get packets from VM-to-VM or blade-to-blade?

Poor

- On the server packet capture agent/analyzer*
  - Limited capture throughput and retention
  - Does not enter the monitoring fabric; save it for workstations
- Blade chassis mirror port
  - Basically the same issues as SPAN; does not address VM's
- Virtual taps or virtual switch port mirroring
  - Consumes real resources to get packets to the monitoring fabric
  - May require large coverage across platforms and VMs
- Integrate packet-based APM with agent-based APM
  - Can work extremely well if well understood & validated
- Architect your infrastructure to force packets out of the box
  - High capacity bandwidth in the data center is cheap
  - Requires strong network + server + APM collaboration

Best

*Or Wireshark in an analyzer VM if using VMWare VDS

NOT Best Practices!

Thank…

You!

↓

Appendix: Switching Fabric Definition

Switching fabric is the combination of hardware and software that moves data coming in to a network node out by the correct port (door) to the next node in the network.

Switching fabric includes the switching units (individual boxes) in a node, the integrated circuits that they contain, and the programming that allows switching paths to be controlled. The switching fabric is independent of the bus technology and infrastructure used to move data between nodes and also separate from the router. The term is sometimes used to mean collectively all switching hardware and software in a network.

The term uses a fabric metaphor to suggest the possible complexity and web-like structure of switching paths and ports within a node. The switching fabric typically includes data buffers and the use of shared memory.

Source: techtarget.com