Arista Networks

Designing Media Networks

Gerard Phillips – SE Media & Entertainment
gp@arista.com

ARISTA

Designing Media Networks

- The move to IP
- Synchronisation
- A brief history of standards
- Media Network Architectures
- SDN / Flow orchestration
- Network Visibility / Programmability
- Wrapping up

The move to IP

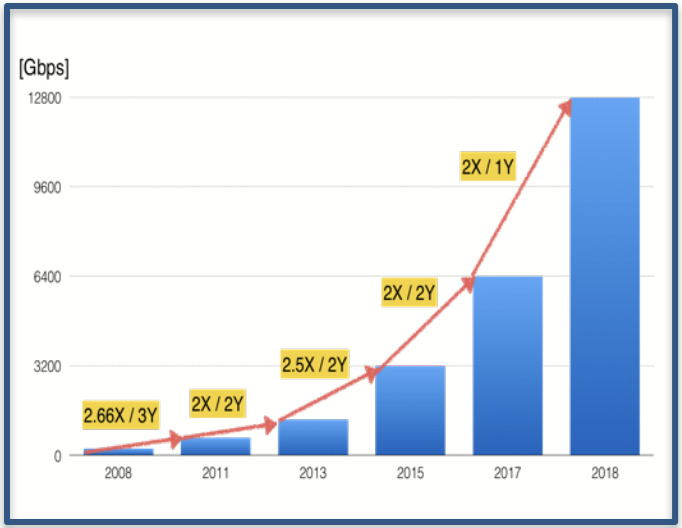
ARISTA

Why make the move to IP?

- SDI, tried and tested, understood
- Single duplex, single composite flow / BNC

- IP, full duplex, multiple flows / fiber
- Leverage COTS:-
  - New merchant silicon every 12 months
  - Faster than Moores law right now!
  - Reliability / product quality – drive by cloud / HFT

- Piggy back off the Cloud Titans – 400Gbe shipping now…



ARISTA

Why make the move to IP?

- Now, its cheaper @ scale > 576^2 (Lawo estimate)
- Enables the scale you'd need at UHD
- Massive space and power savings for OB applications

- Flexibility – IP is just the plumbing
- Converged / multi-tenancy / anybody anywhere

- IP end-to-end – glass to glass - live production, playout, contribution, distribution, repurposing, editing, OTT….

- …. enables new / innovative workflows, production techniques, etc

27RU = 576*576

7RU = 4K*4K @ 3G

Why make the move to IP?

- Becomes cheaper as more native IP end points / processing elements become available – SNP / Neuron / MV's / Playout servers….
- IP…. It is tried and tested & understood!
- Reliability – 24/7/365 operation is an expectation
- No performance limits – Line-rate, non-blocking, fast accurate switching, PTP
- Anything else?
  - Reduce operating costs
  - Easily spin up new services / clients
  - An on ramp to virtualised workflows
  - Flex into the cloud

Flexible connectivity

SFP/SFP+/SFP28
1/10/25Gbe

QSFP/QSFP28
40/100Gbe

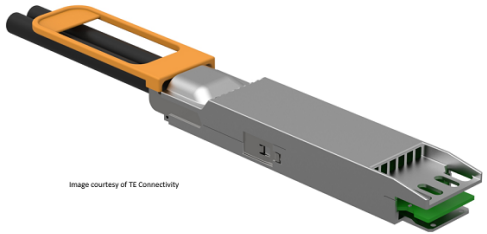
QSFP/QSFP28 -> 4xLC

75Ohm, Single Duplex
3Gbps, maybe 12Gbps?

Direct Attached Copper
10/25/40/100Gbe

Active Optical Cable
10/25/40/100Gbe

OSFP or QSFP-DD
400Gbps



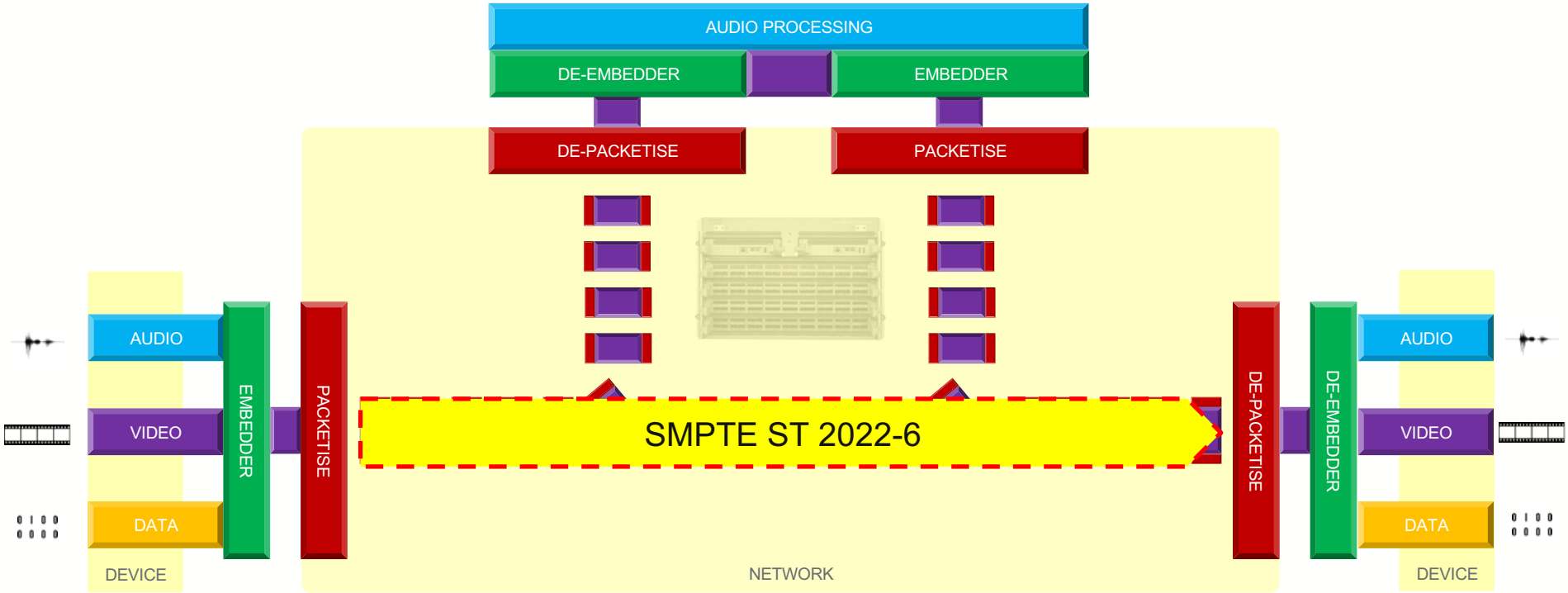
A brief history of standards…

A brief history…

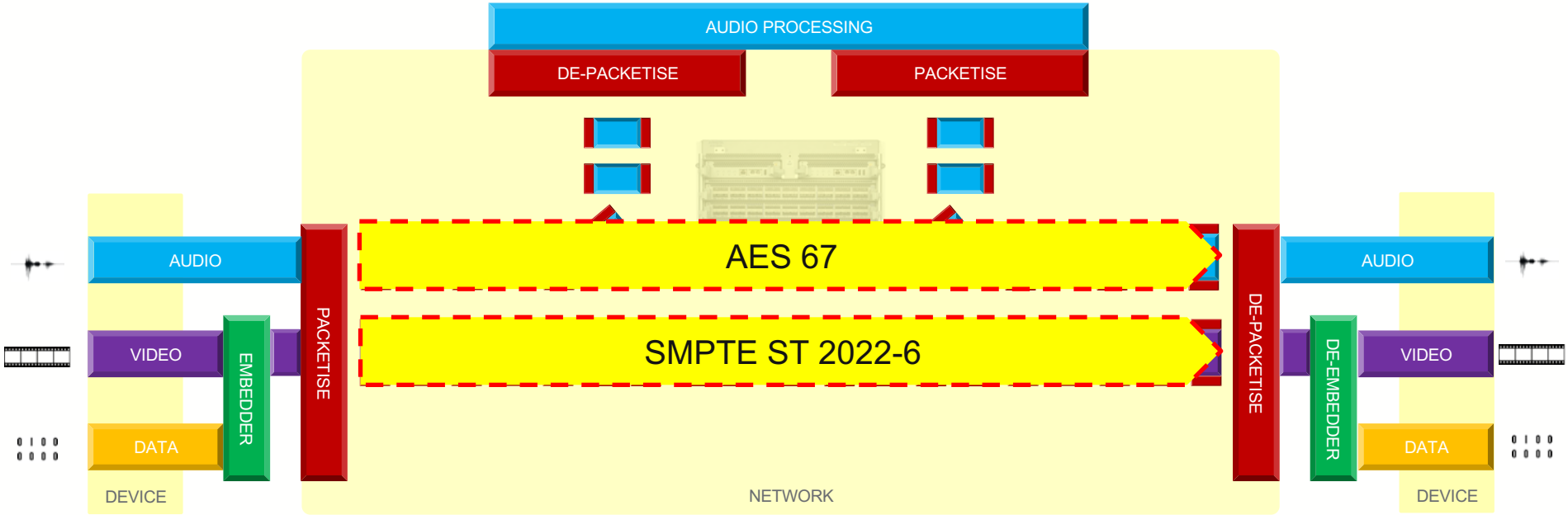
- Audio has been IP for ages – Dante, Ravenna, Livewire
  - AES67 describes the bits that are common…
  - ST2110-30 is …..almost AES67

- The VSF's work lead to the ST2022 standards (CBR/VBR/HBR + FEC + "-7")
  - Very often used for contribution / hand-off
  - 2022-6 encapsulates the whole SDI stream
  - Guarantees lip-sync, and metadata delivery
  - Can be used (amongst other things) to tunnel TICO over SDI over IP!
  - No UHD variant yet – so you're stuck with quad HD / 2SI for uncompressed
- SMPTE's ST2059 provides a PTP profile for high bandwidth media
- The VSF creates TR-03 & TR-04
- SMPTE adopts and re-writes into ST2110
- **All underwritten by the AIMS roadmap and within the JT-NM framework**

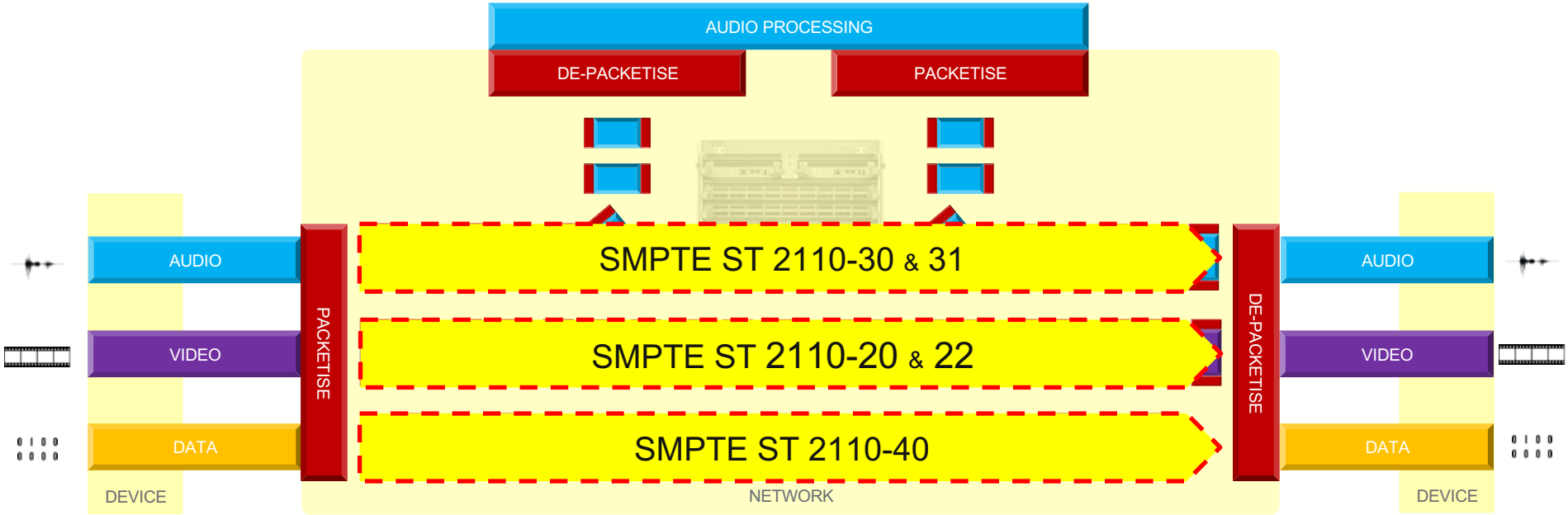
SMPTE ST 2022-6
Multiplexed flows



SMPTE ST 2022-6 & AES 67
Multiplexed flows



SMPTE ST 2110
Essence flows



ST2110 == Flexibility

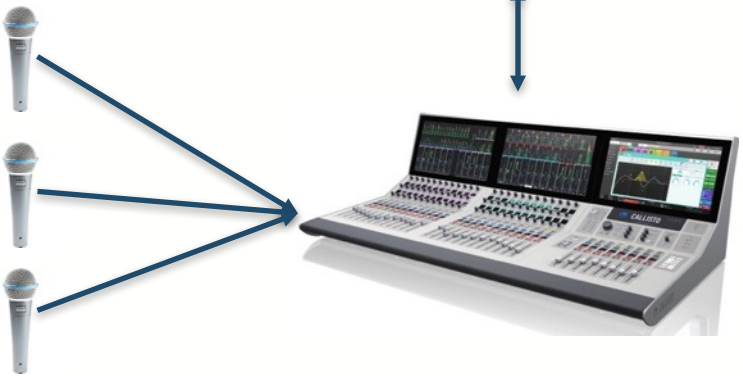
- Separate flows for individual essences
- No blanking sent
- Decoupled from underlying SDI formats
- Implicit synchronisation (ST2059)
- 4:1:1 -> 4:4:4
- 8->16 bits colour depth
- Format flexibility to 32k[2]

- Extensible
  - Compressed video / audio
  - Native IP metadata (script / GPS co-ords)

- ST2110-10, System Timing
- ST2110-20, Uncompressed video
- ST2110-21, Traffic shaping (-20)
- ST2110-22, Compressed
- ST2110-30, PCM Audio
- ST2110-31, AES3 Transparent Transport
- ST2110-40, ANC Data

SDI Live production

TX Feed



Low latency
Synchronised
Separate essences where required

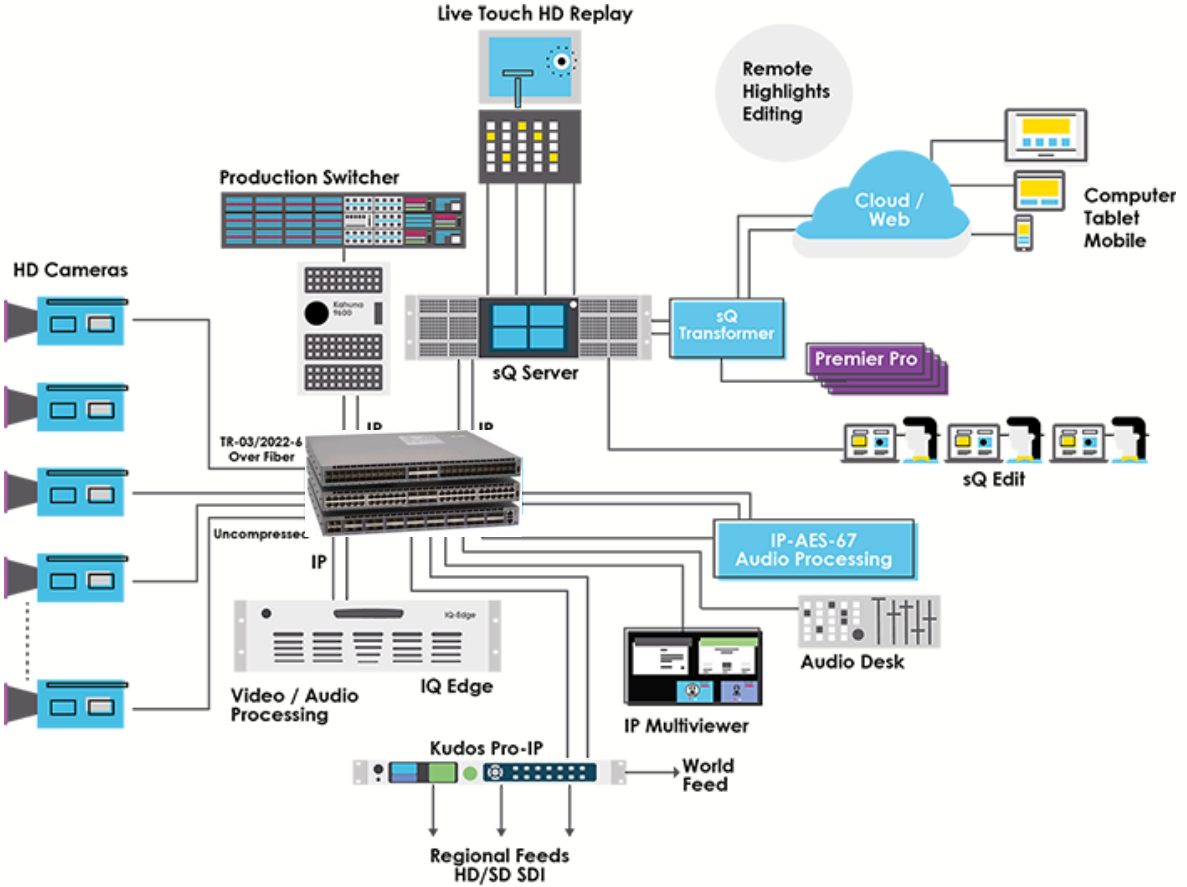
- SDI
- MADI
- AES67
- AES3
- ....

IP Live production

- Looks much the same as now!
- Efficient essence workflows

- Hybrid systems possible
  - SDI sea with IP islands
  - IP sea with SDI islands

- Naturally enables S/W nodes
  - On-ramp to virtualised studio

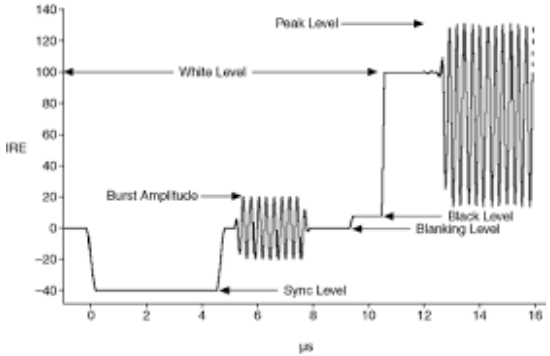


ARISTA

Synchronisation

Before PTP

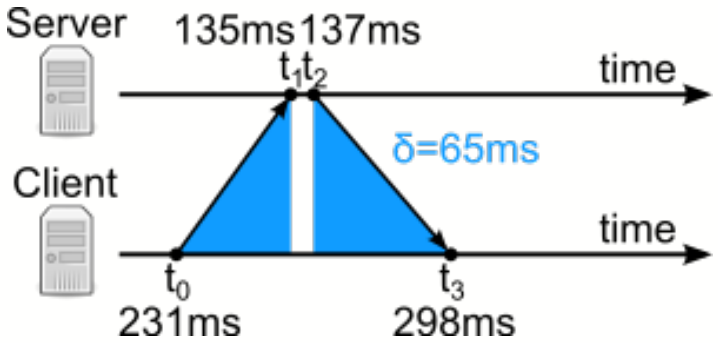
- Broadcast systems require synchronization (a shared time-base)
  - Minimizes latency
  - Prerequisite for accurate on air cuts / effects
  - Essential for multi-channel audio formats
  - Raises quality (frame drops, audio glitches)
  - Enables lower cost kit (no frame stores)



- BB / Tri-syncs / Word clock / DARS provide this synchronisation

And in the ST2110 age? We still need to….

- Provide frequency and phase synchronization between all components in a system
  - Allows all clocks to be locked together
- Provides media element identification
  - Video frames
  - **Audio samples**
  - Metadata packets



- Take complex essence like SDI/Audio/Captions and tag these elements as we split them. So that we can then re-combine later
- Tag disparate sampled elements like camera / microphone and combine

PTP to the rescue

- IEEE 1588 / PTPv2
- **P**recision **T**ime **P**rotocol

- Widely used
  - Industrial Automation
  - Financial trading
  - Power generation / distribution
  - Mobile backhaul synchronization
  - Basis of White Rabbit – used at CERN
  - AVB / Dante / Ravenna

- Typically locked to GPS
- High accuracy possible (10's ns)

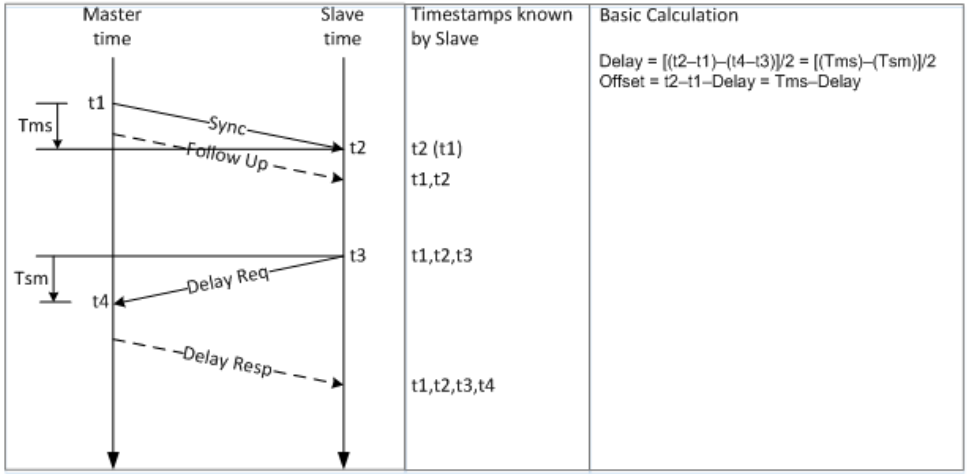
- **SMPTE ST2059** defines the epoch as midnight, 1st Jan 1970
- **SMPTE ST2059** defines how video frames and audio samples map to this
- **SMPTE ST2059** provides a "profile" that aims to allow +/-500ns to be achieved

IEEE1588 / PTPv2 for Broadcast

- SMPTE ST2059
  - Specifies an epoch
  - Specifies target accuracy
  - Specifies how video and audio are related to the epoch
- ST2110 (20 / 30 / 40 etc), AES67, ST2022-6(TR-04)
  - RTP derived from PTP
  - Enables elements to be tied back together
  - Provides frequency, phase and wall clock time
  - AES67 and SMPTE2059 have different setting ranges in their “profiles”
- AES67 / ST2110 overlap (AES-R16-2016)
  - Fortunately, there is common ground – you can use one profile to rule then both!

How does PTP work?

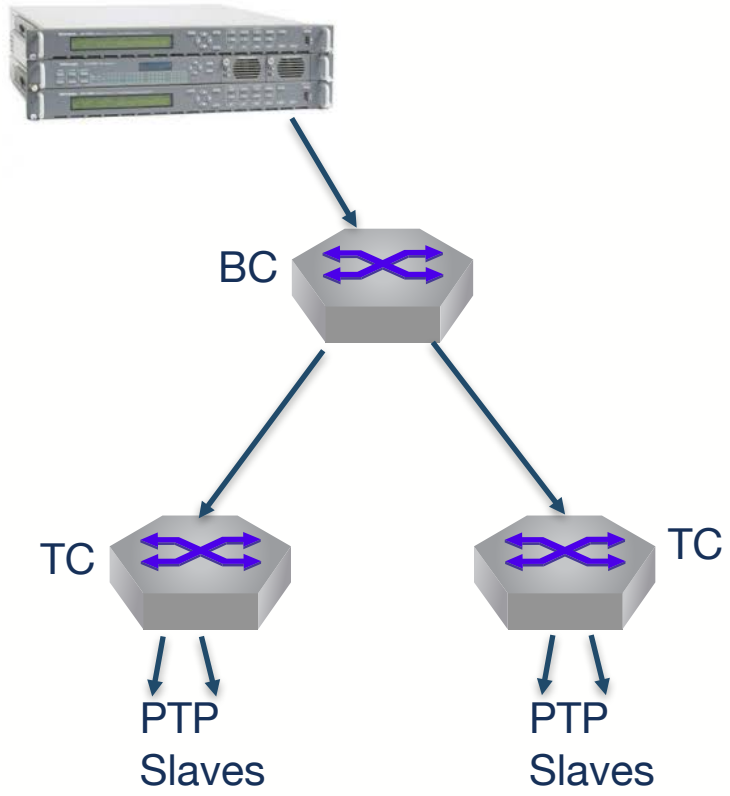
- Announce messages sent by the master
  - Received by all slaves (and potential masters)
  - Typically 1 per second
  - Used in the BMCA process
- Syncs sent periodically by the master
  - Received by all slaves
  - Typically 8 per second
  - Not dissimilar to NTP!
- Delay requests sent by slaves
  - Typically 8 per second
- Delay Response back from master



- Hardware timestamping is essential
- Reliable delivery + simple maths + complex filtering = accurate slave time!

Maximising PTP performance across a network

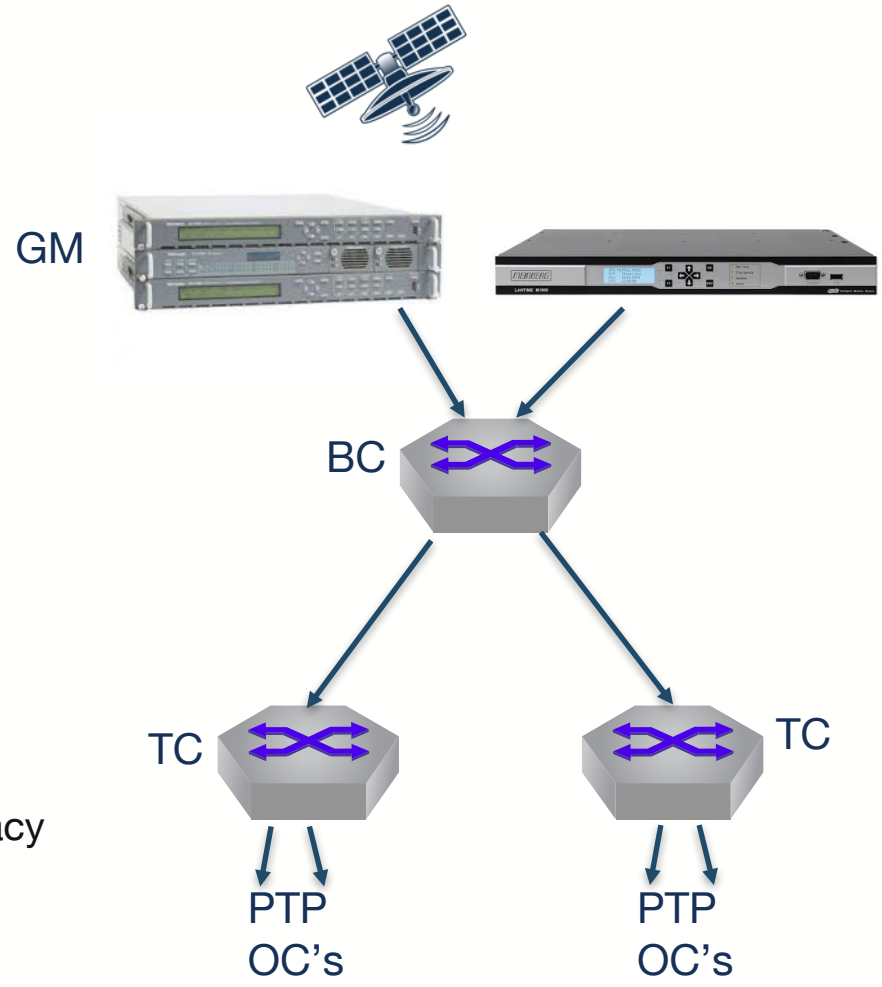
- Use GPS to lock GM's if practical
- In a hybrid system PTP and BB coexist
- Rich network PTP functionality provides;
  - Boundary and Transparent clock capability
  - SMPTE ST 2059 + AES67 profile support
  - Scalability – 1000's of endpoints
  - Architectural simplification



ARISTA

PTP clock types

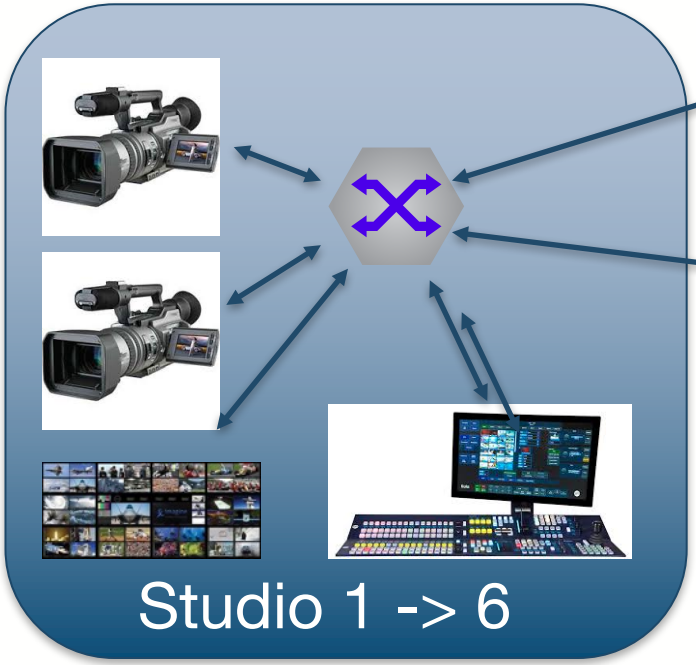
- Ordinary Clock
  - Grand Master – Typically GPS locked
  - Slave Only
  - Slave or master
- Boundary Clock
  - Eliminates switch delay (== jitter)
  - Switch acts as both Slave and Master
  - Run host ports at the rate you need
- Transparent Clock
  - Eliminates switch delay (== jitter)
  - Messages forwarded through switch
  - Slaves use correction field to improve accuracy



PTP and the LAN

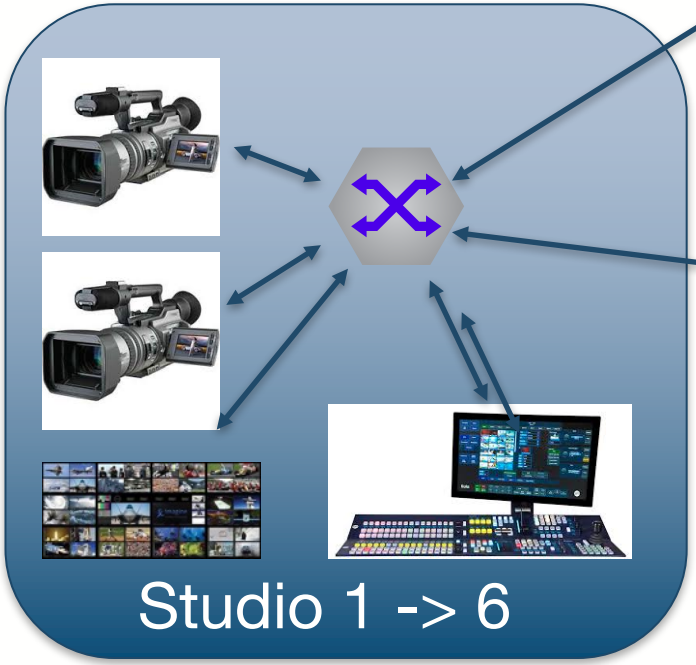
Remote facility

LAN



- LAN connectivity provides low jitter
- PTP services *can be* shared
- Consider resilience levels
- Consider resource flexibility

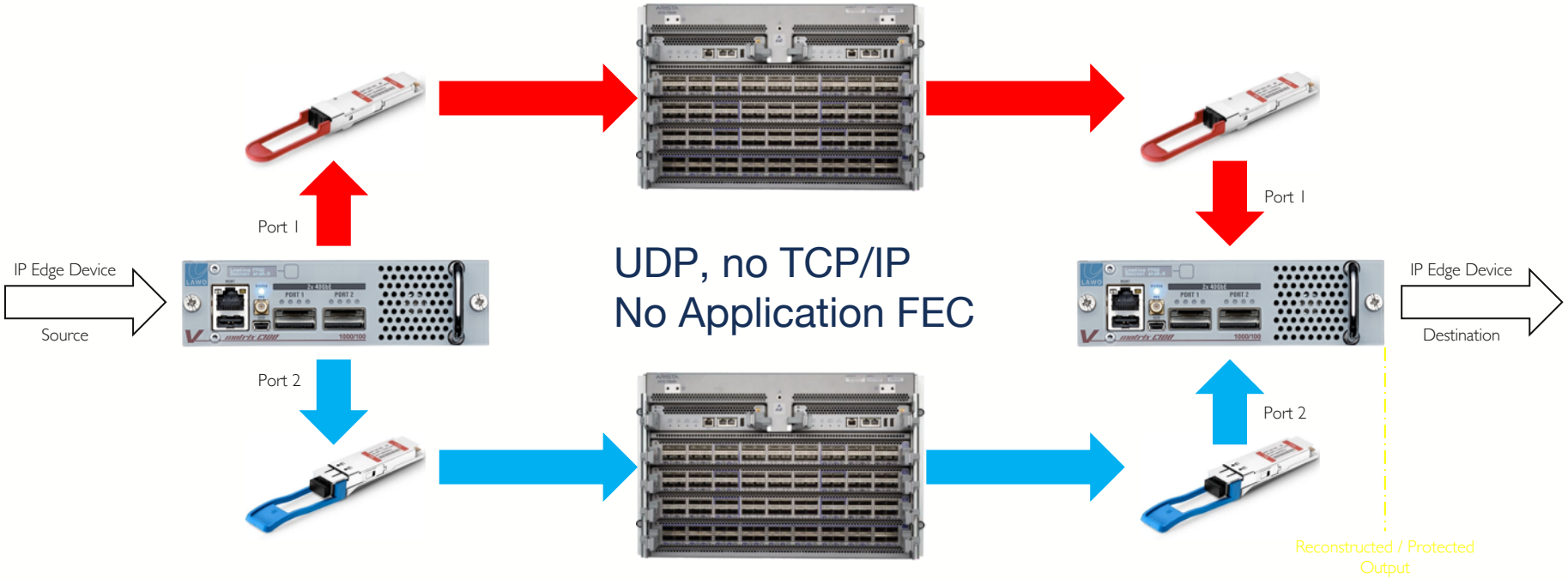
PTP and the WAN



- PTP will not tolerate jitter
- Would not recommend sharing PTP
- Design as if remote facility is NOT synchronous
- Like we typically do now!

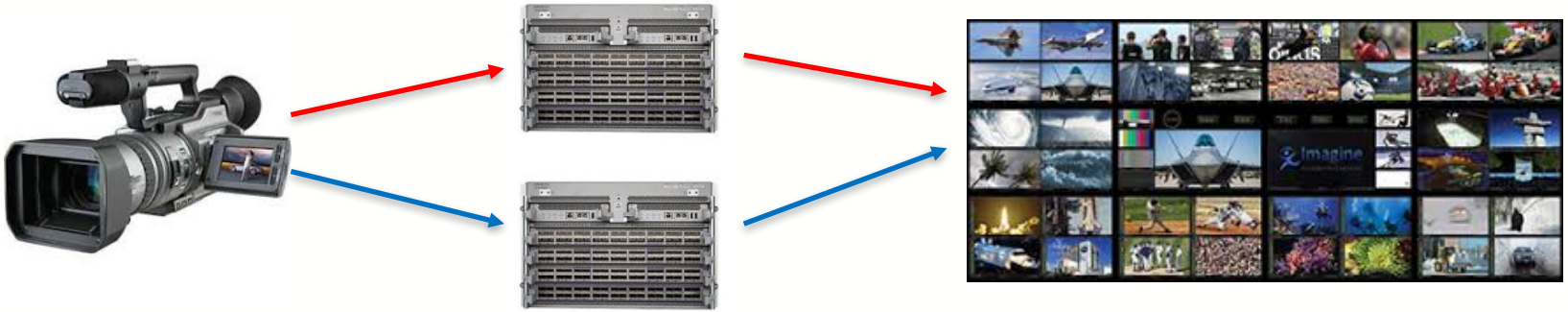
Media Network Architectures

Seamless protection switching (-7)



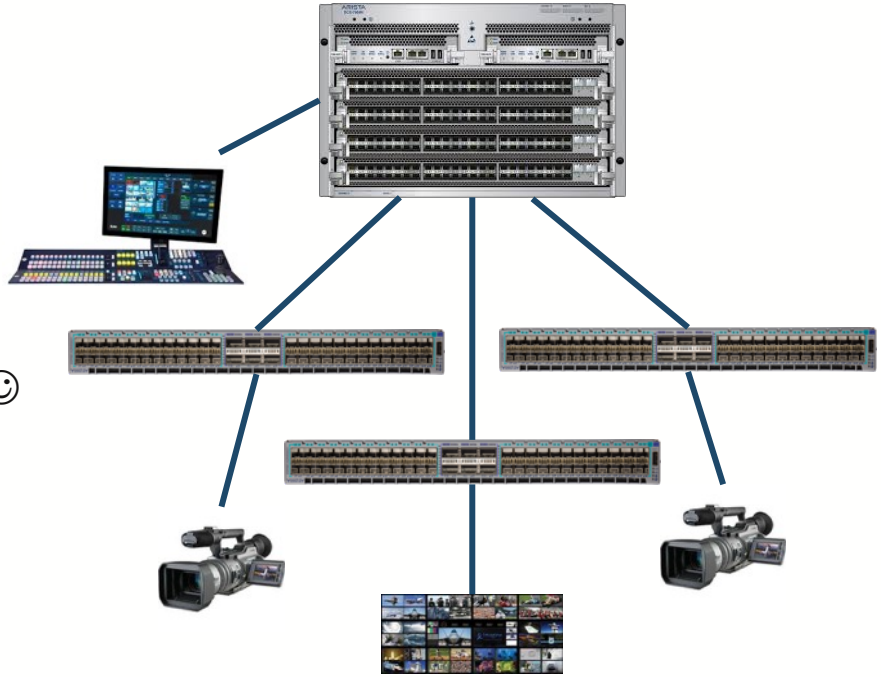
Media Network Architectures, Single big switch

- Choose the right switch, and this is line-rate non-blocking and multicast non-blocking
- Very simple to envisage and provision, IGMP works fine
- Very scalable 32x100Gbe -> 576x100Gbe -> X4 @ 400Gbe!
- **896* 896 @ 3Gbe -> 16K square @ 3Gbe**
- > 2k hosts @ 25Gbe



Hub and spoke

- >16k hosts @ 25Gbe, 1:10 provisioning
- 400Gbe around the corner….
- Scale at the rate you want to
- Just like traditional broadcast Tie-lines ☺
- Non-blocking no longer makes sense
  - Does that fit your workflow?
- Flow orchestration (SDN) should be considered

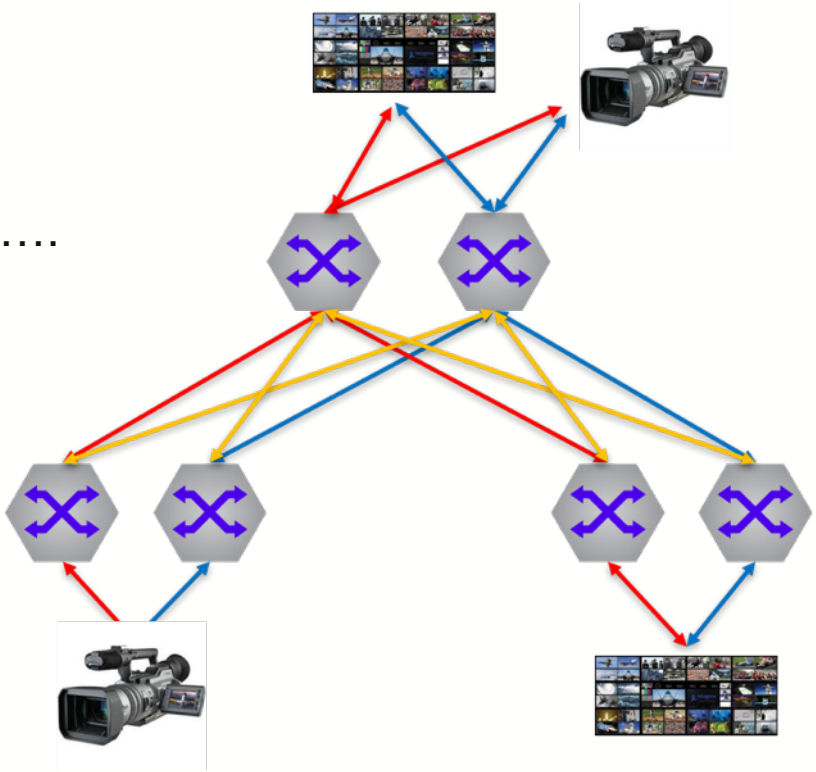


Leaf and Spine

- A great architecture for future thinking converged network
- ...think Virtual Network Functionality
- ST2110, Storage, Compute, Transcode….

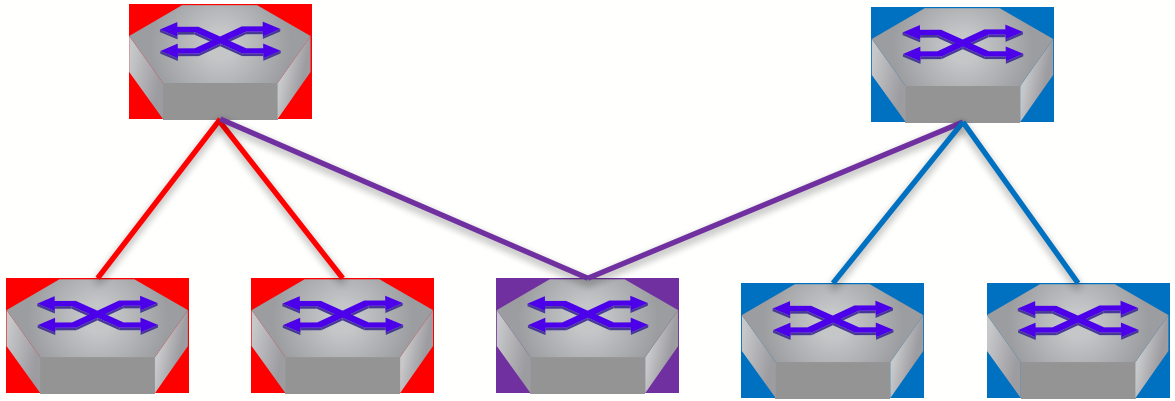
But….

- Now, flow orchestration is essential
  - Protects ST2110 / ST2022 / AES67
  - Enforces workflow security
  - Contributes to Multi-tenancy operation



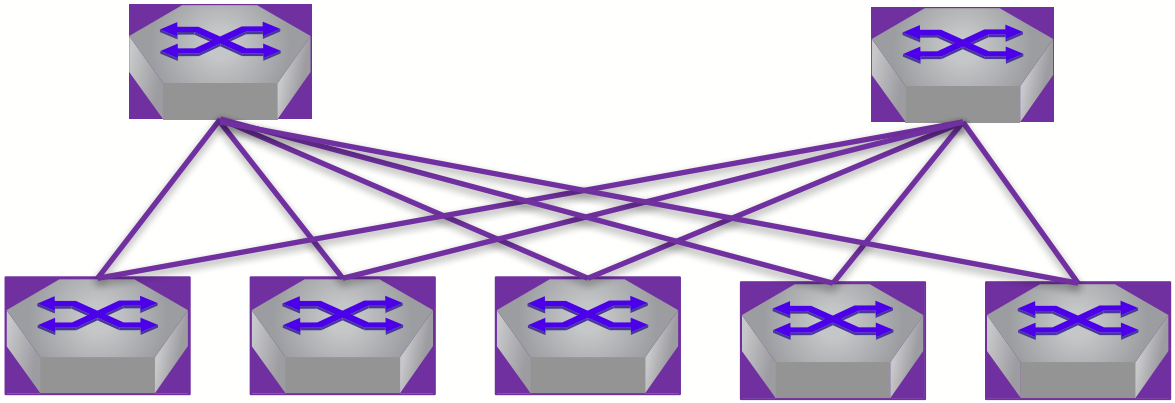
Purple / hybrid approach

- Red / Blue spines – Simplifies SDN controller
- Red / Blue or Purple leafs
- -7 Diversity guaranteed by Red / Blue spine
- -7 enables relatively simple upgrades, maintenance
- -7 handles failure gracefully
- Single homed devices accommodated



Pure Purple approach

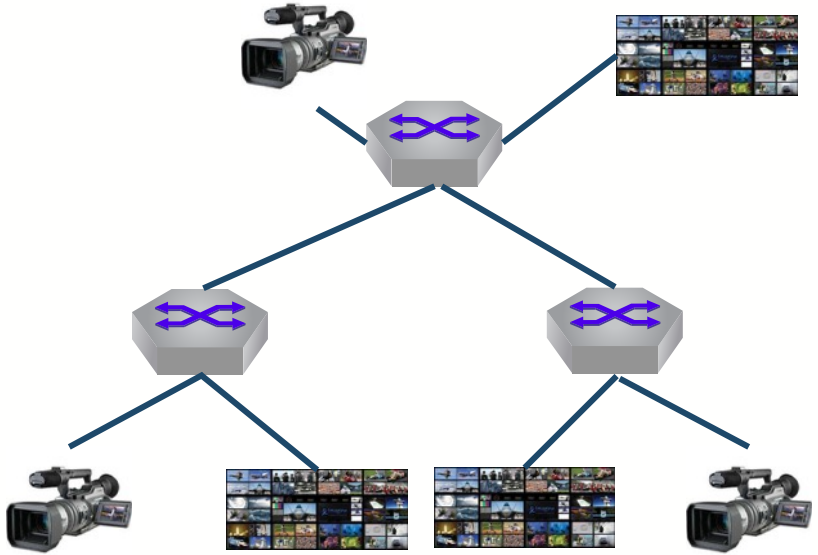
- Spines and leafs are purple
- -7 Diversity provided logically by SDN controller
- -7 flows can be maintained in case of spine failure (but not complete physical diversity)
- Trade-off's can be made in failure case – how to use remaining spine B/W
- SDN controller needs to take proactive action for upgrades, maintenance, failure
- Single homed devices accommodated



Expansion…

…. if you want to start with a single switch (pair -7) and grow

- Start with a leaf
- Either add a sp(l)ine…
- Or promote the leaf pair to “spline”



- Start with IGMP or SDN
- Your decision – your network

SDN / Flow orchestration

ARISTA

Who needs SDN? Who doesn't?

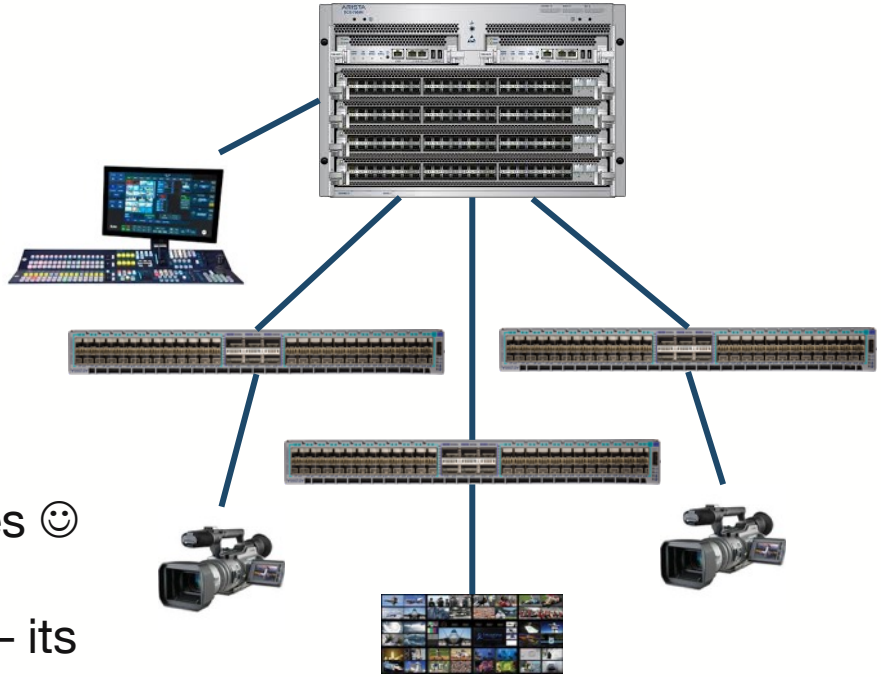
- **Large, distributed, multi switch oversubscribed media networks**
  - More source bandwidth than any switch to switch interface can transport
  - Cannot use existing protocols and configurations to define network forwarding
  - Implementing a BC that can talk to multi-vendor API's (network and endpoint)

    » **This is KEY for 2110 deployments!!!**

- **Non-blocking network designs**
  - Enough bandwidth for required flows to be distributed over available links with no risk of oversubscription (may be single link or statically defined multi-link)
  - Examples: audio breakouts, lower bitrate file/compressed, defined I/O paths
- **Single switch, non-blocking deployments**
  - Bandwidth control not required
  - IGMP for subscribers works fine in single
  - Chassis scale to 500+ 100G ports or 2000+ 25G ports

SDN / Flow orchestration

- Flow orchestration (SDN) should be considered:
  - For multi-switch topology
  - For multi-link connectivity
- Non-blocking no longer makes sense
- Just like traditional broadcast Tie-lines ☺
- Don’t forget security and scheduling – its all resource management!



NOS Choice: The Key for SDN and more….

- SDN is founded on fast, secure, flexible API's
- State based NOS'…
  - Simplify extensibility
  - Lead naturally to state based streaming telemetry
  - Facilitate many API options – the right tool…
- Open standard API's and tool bring…
  - Re-use, efficiency, higher up-time
  - Reduce vendor lock-in
  - Drive down Opex

OPENCONFIG

State Driven Replication and Streaming

ARISTA

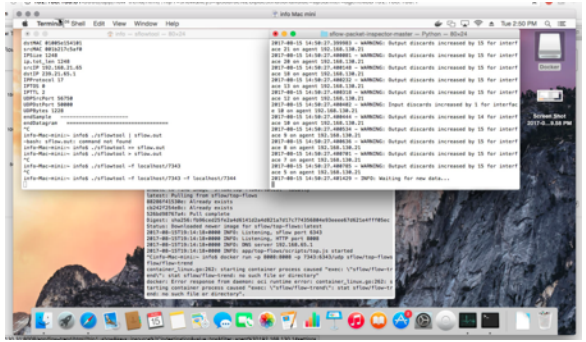
Network Visibility / Programmability

ARISTA

Monitored Data -> Information -> Knowledge -> Control

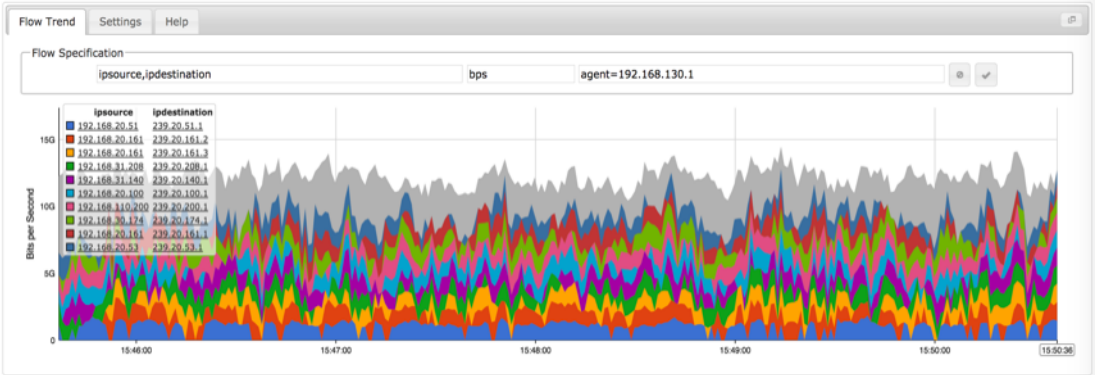
- The OSI model is 7 layer – SDI is 3?
- Good telemetry bridges this gap and some…

- Rich (open) API's enable 3rd parties to build Application specific tools, to leverage network health and status



- Rich (open) API's provide choice and flexibility

- Flexibility enables Business Agility

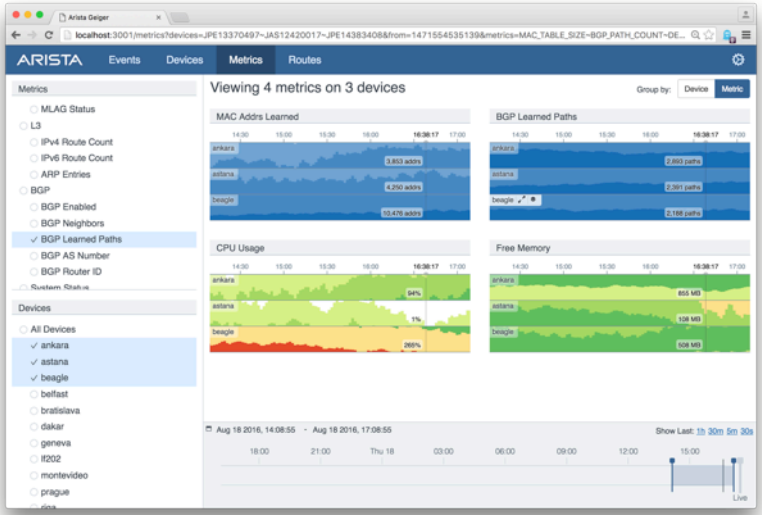


Programmability example – Lawo smartDASH

- Arista's eAPI exposes a rich network data set…
- … transformed through the smartDASH Broadcast lens

CloudVision Telemetry App

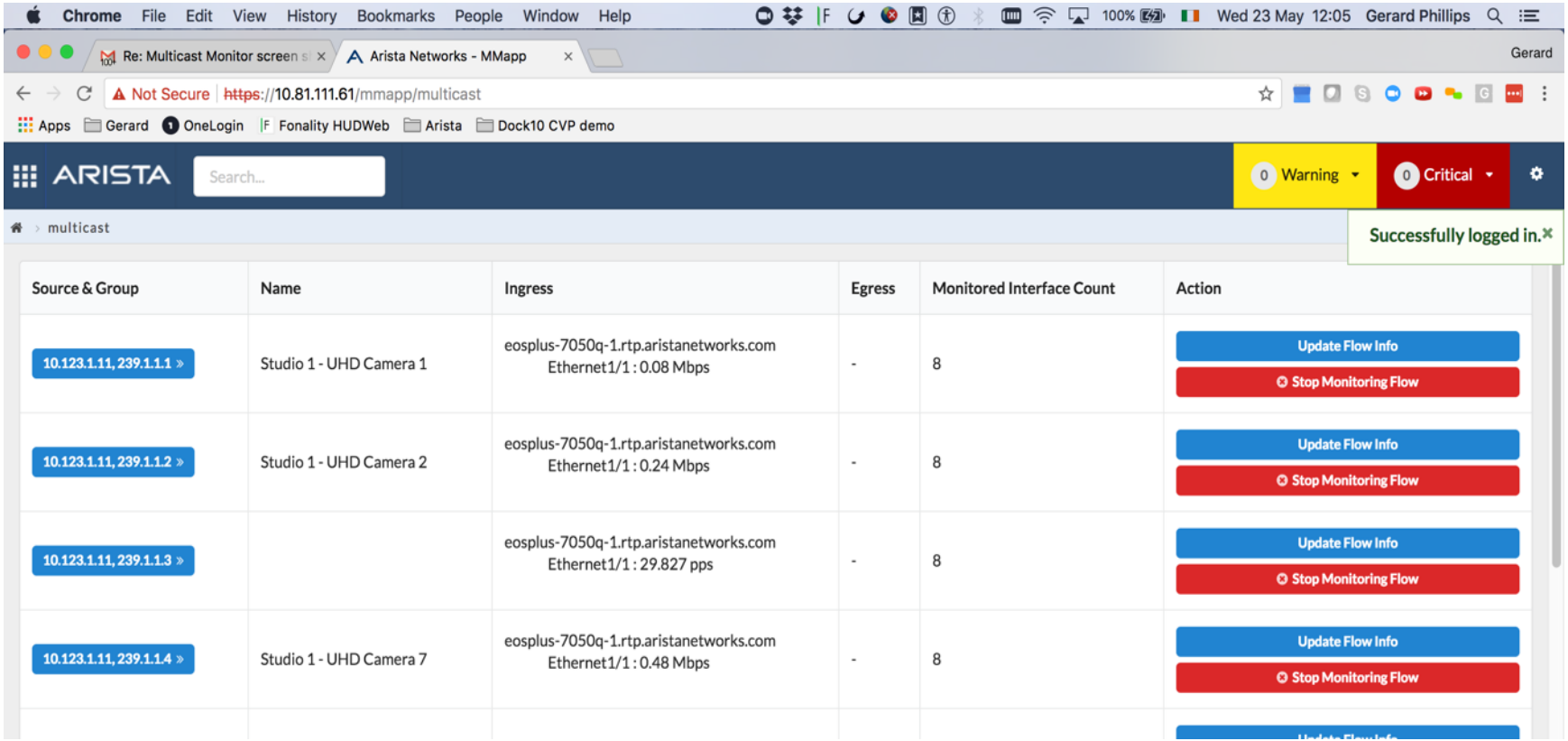
Workstream Analytics Viewer



Complete, real-time NetDB state streaming

- CloudVision Telemetry Apps provide front-end for visibility network state
- ***Workstream Analytics Viewer***
  - 1st CloudVision Telemetry App
  - Correlation of network-wide data
  - Views: Event, Device, Metric, and more
  - Timeline view for better historic troubleshooting
- More apps to follow:
  - Other CV-based apps
  - APIs for customer & partner apps

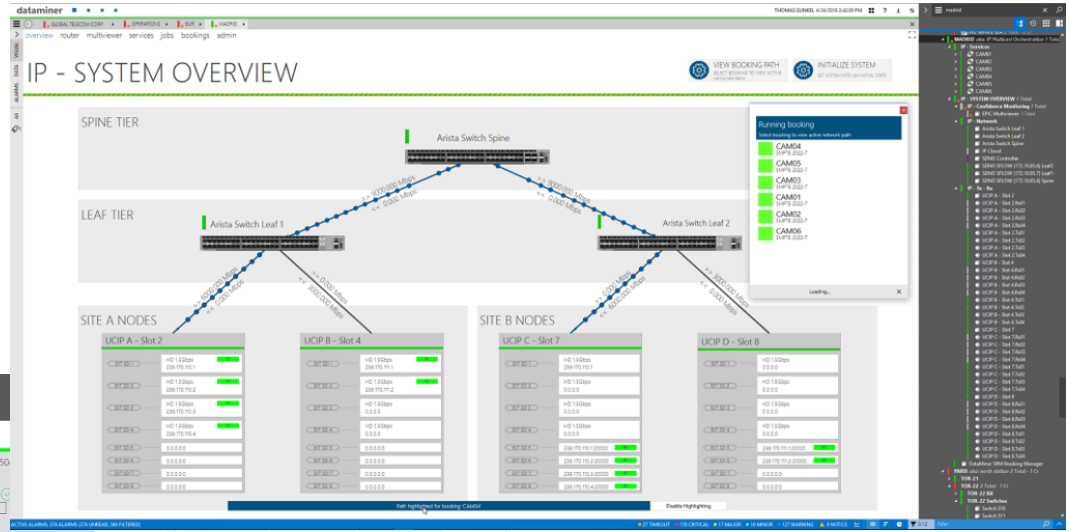
Multicast monitor – Real-time flow monitoring

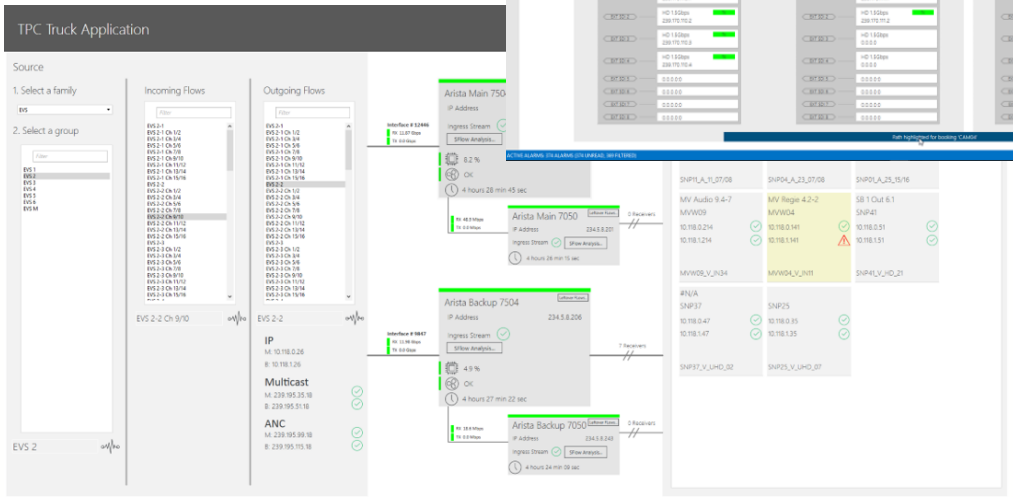


ARISTA

Programmability example – Skyline/Dataminer

- Arista's eAPI exposes a rich network data set…
- … transformed to Broadcast Knowledge





Wrapping it up

Live Production needs that to be addressed in IP

- Multicast
  - One to many
  - Non-blocking, Deterministic
  - Traditional Multicast - IGMPv3 / IGMPv2 / PIM-SSM
    - Hardened in the HFT market
    - IGMPv3 on-air performance proven with many customers – TimelineTV / NEP Wimbedon / rtl/BCE
    - Proven at huge scale with NEP Australia, and others
- SDN / Orchestration
- Scale – 10's of thousands of multicast groups for ST2110
    - Camera grading – fast, reliable
    - Workflow changes - reliable @ scale
    - Simultaneous routing changes – quick, reliable
- High availability / Reliability
- Visibility + Programmability

ARISTA

Moving to IP….

- Be clear about your objectives – what does IP bring to your installation?
- Its not for every project, but hybrid SDI / IP works well
- Don’t think of this as SDI over IP – its live production / playout over IP
- Xfunctional teams – broadcast guys and IT guys

So……

- Leverage subject experts – Broadcast vendor / SI / network vendor
- Your Broadcast Controller – its what will unlock value in your IP infrastructure.
- Don’t leave the monitoring of IP as an after-thought

But don’t be put off – its real!

sky

euro media group
UNITED BROADCAST FACILITIES

ARISTA

Thank you!

Q&A

ARISTA