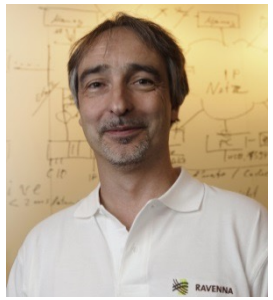
Understanding Latency in AoIP Systems

"Your Flight Will Be Delayed by 20 ms"

MoIP Pavilion @ AES New York 2022

Andreas Hildebrand, ALC NetworX



Andreas Hildebrand, RAVENNA Technology Evangelist

- more than 25 years in the professional audio / broadcasting industry
- graduate diploma in computer science
- R&D, project & product management experience
- member of AES67 TG and ST2110 DG

ALC NetworX GmbH, Munich / Germany

- established 2008
- R&D center
- developing & promoting RAVENNA
- Partnerships with > 40 manufacturers

RAVENNA

- IP media networking technology
- designed to meet requirements of professional audio / broadcasting applications
- open technology approach, license-free
- fully AES67- and SMPTE ST2110-compliant

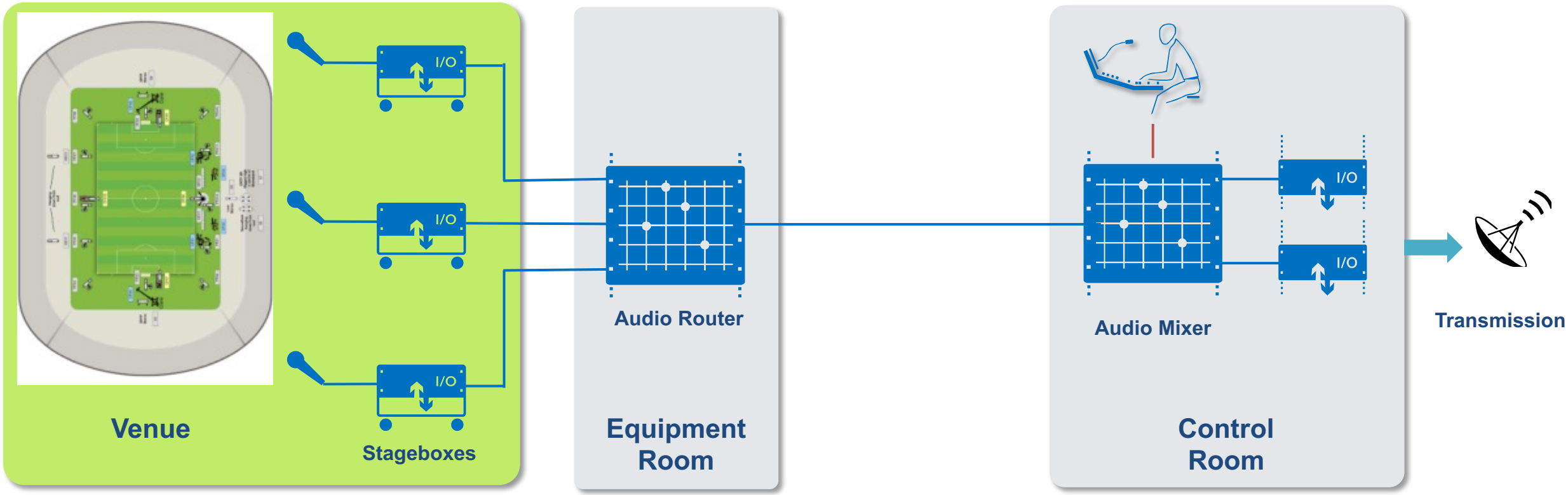
"... is a time **delay between** the **cause** and the **effect** of some physical change in the system being observed." [Wiki]

"... is a **consequence** of the **limited velocity** at which any physical interaction can propagate [...] which is always less than the **speed of light**." [Wiki]

RAVENNA
AES67 & ST 2110 built-in

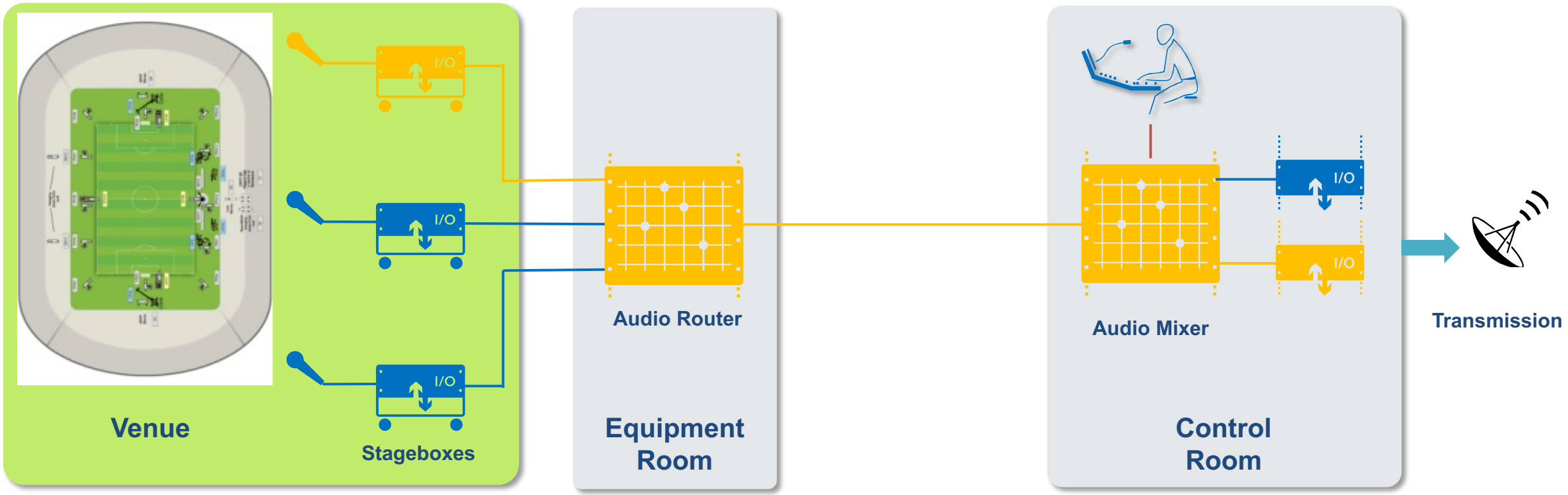
(Traditional) Digital Signal Distribution

(Traditional) Digital Signal Distribution



RAVENNA
AES67 & ST 2110 built-in

(Traditional) Digital Signal Distribution



fixed connection between microphone and mixer output

fixed / deterministic latency

= circuit switched routing

Approximate light signal travel times

Distance	Time
1 m	3.3 ns
1 km	3.3 µs
4000 km (NYC – LA)	13.3 ms
to geostationary orbit	119 ms
around Earth's equator	134 ms
Moon to Earth	1.3 s
Proxima Centauri (nearest star)	4.2 years
across the Milky Way	100,000 years
Andromeda Galaxy (nearest galaxy)	2.5 million years

RAVENNA
AES67 & ST 2110 built-in

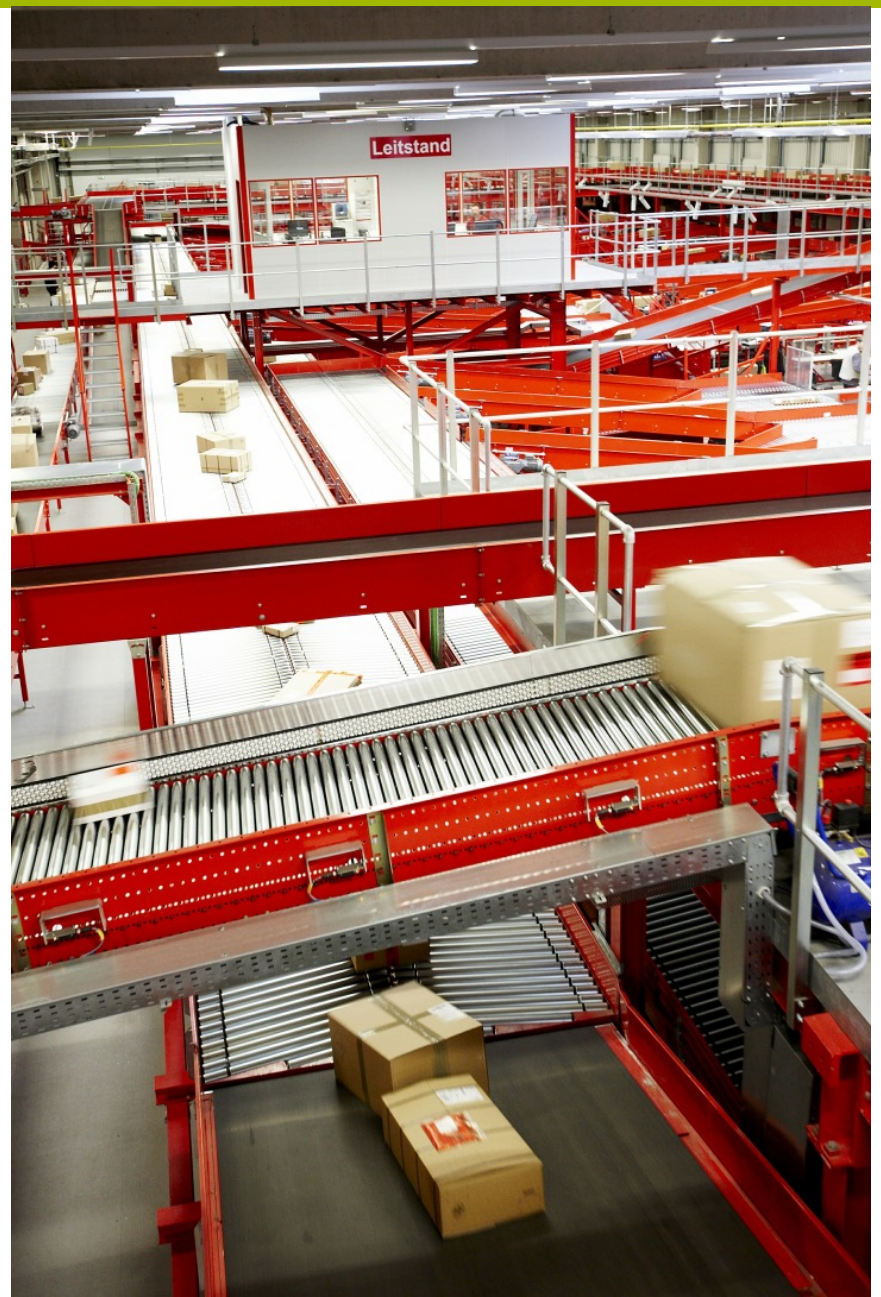
Networked Media Distribution

=

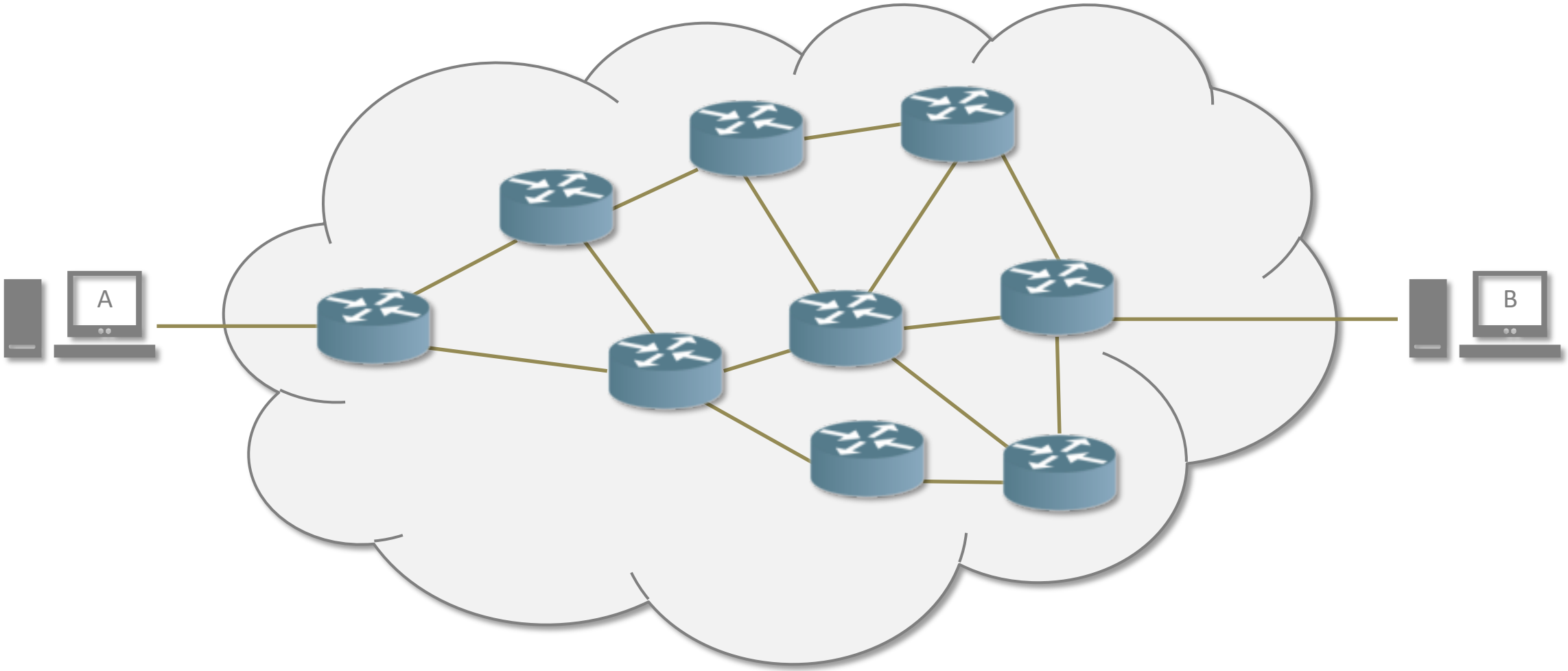
Packet Switching

Latency in AoIP Systems

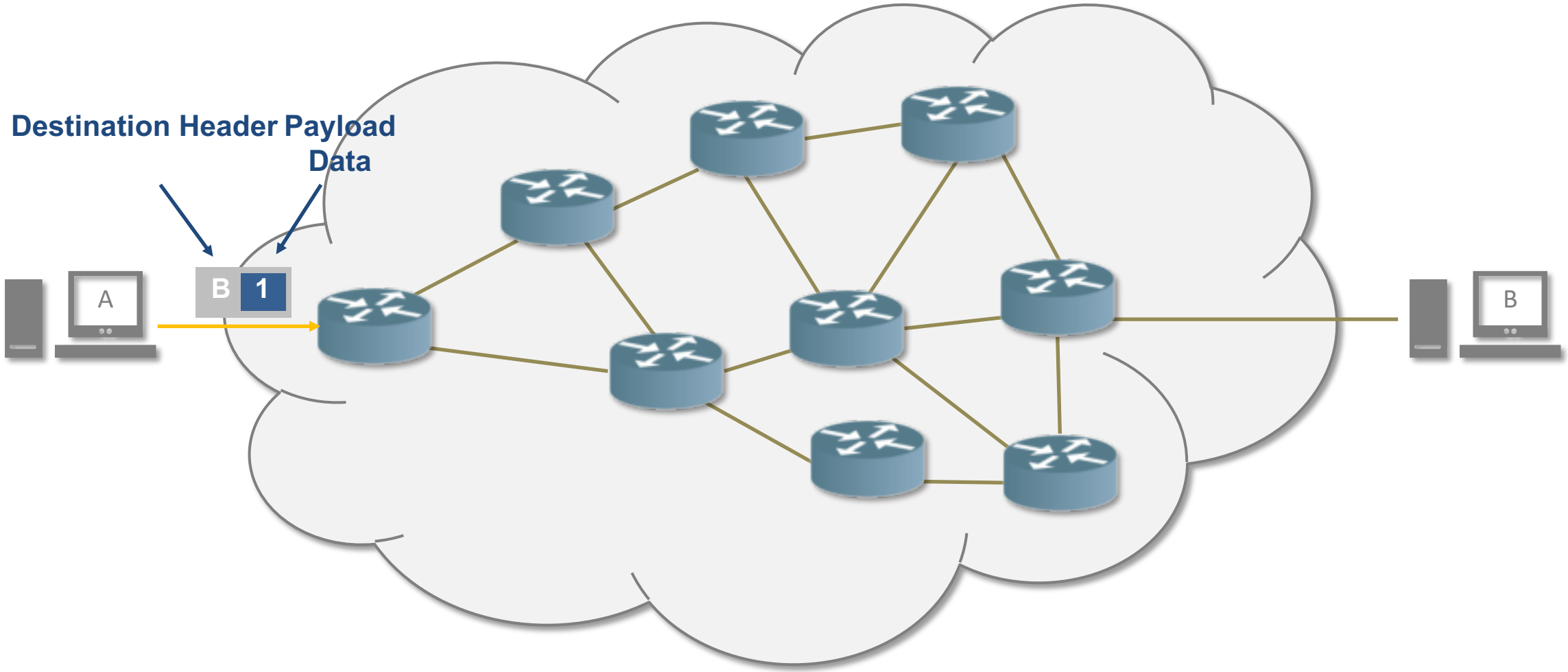
RAVENNA
AES67 & ST 2110 built-in



IP Packet Switching

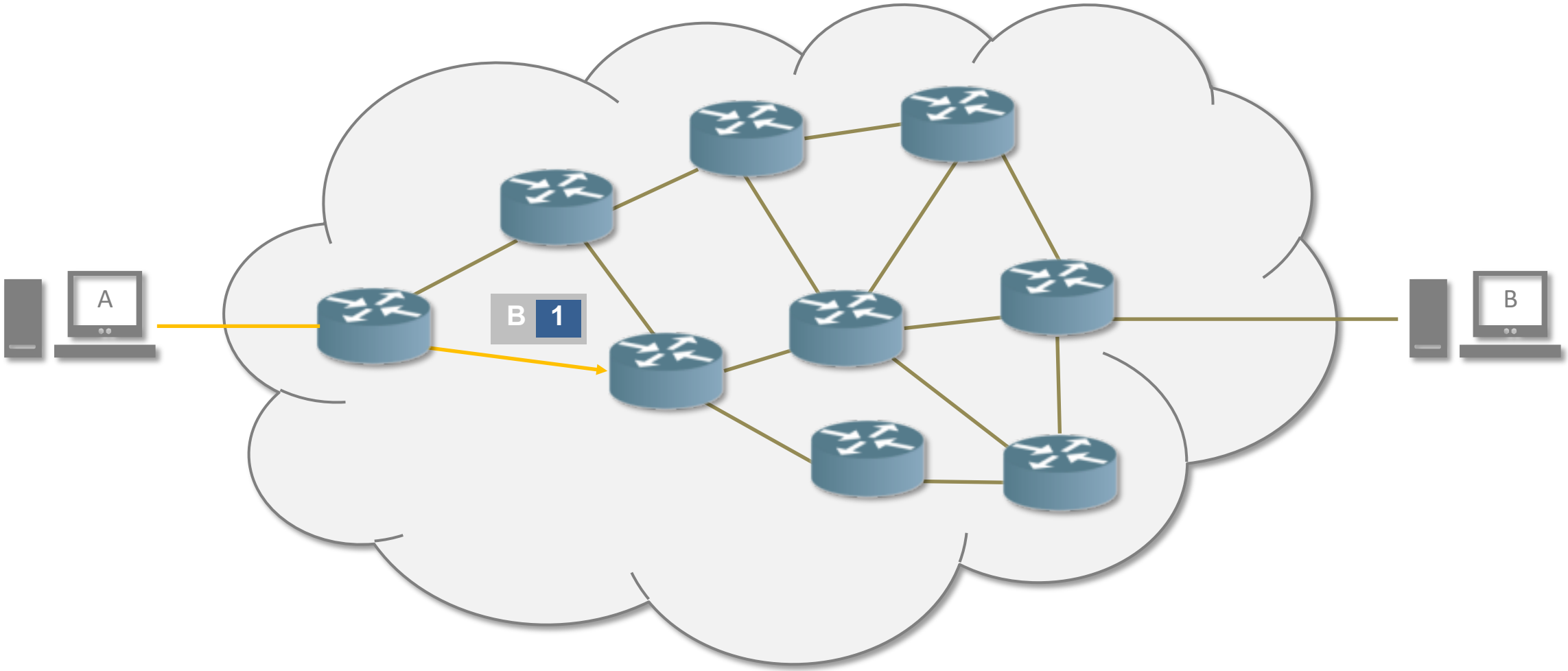


IP Packet Switching



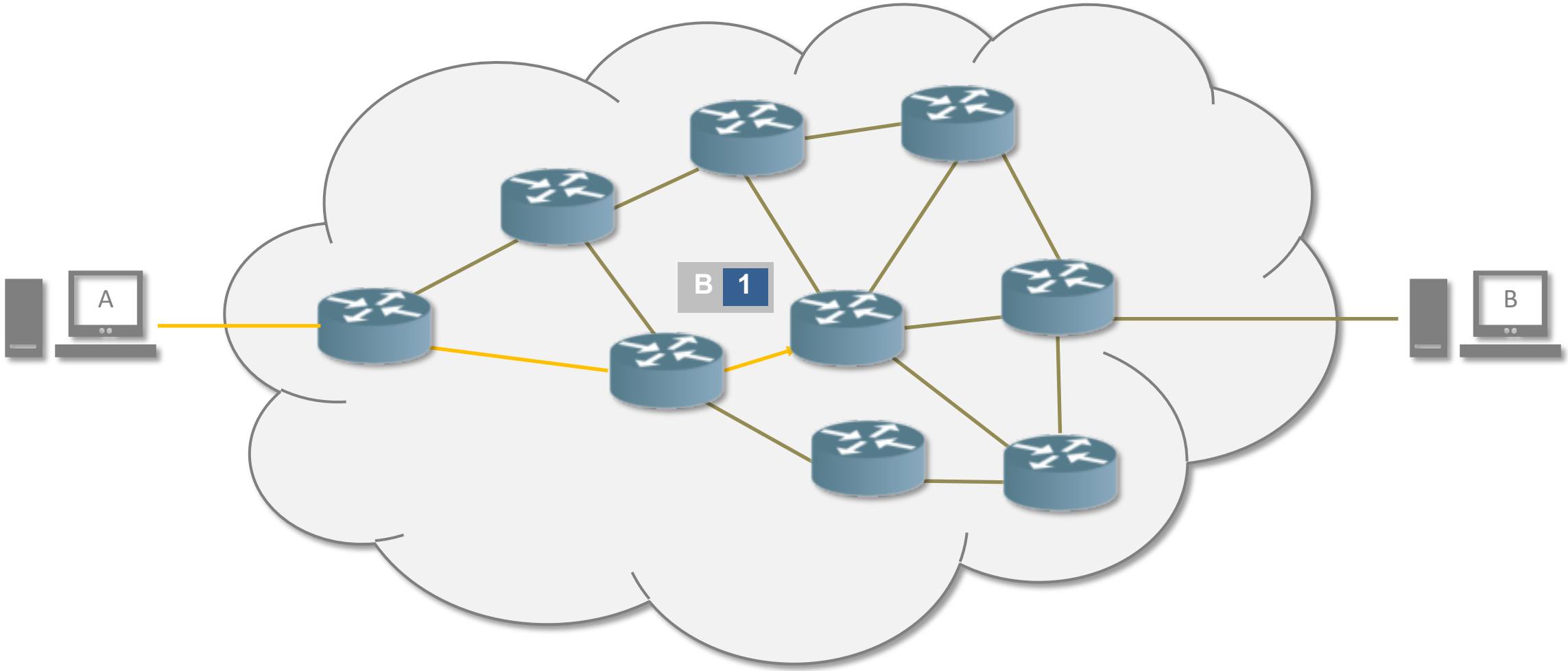
A sends data packet to B

IP Packet Switching



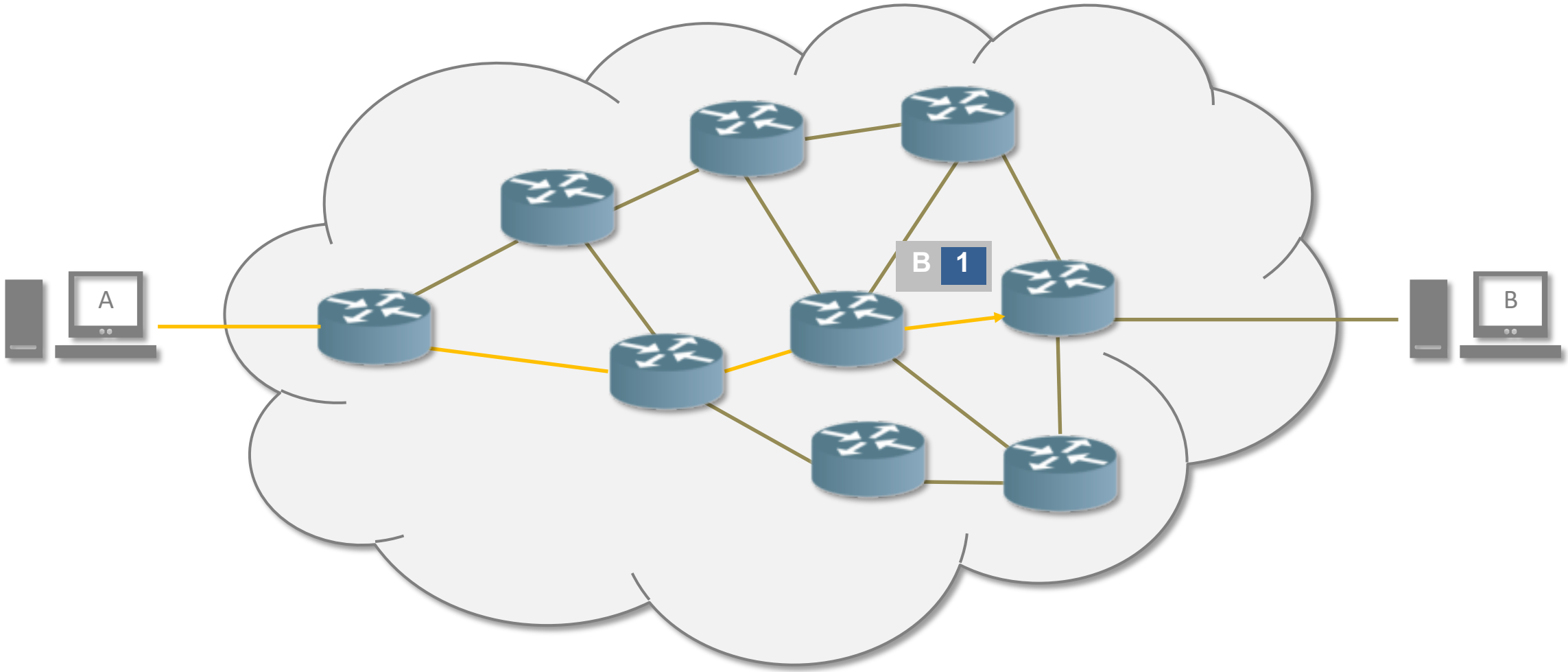
A sends data packet to B

IP Packet Switching



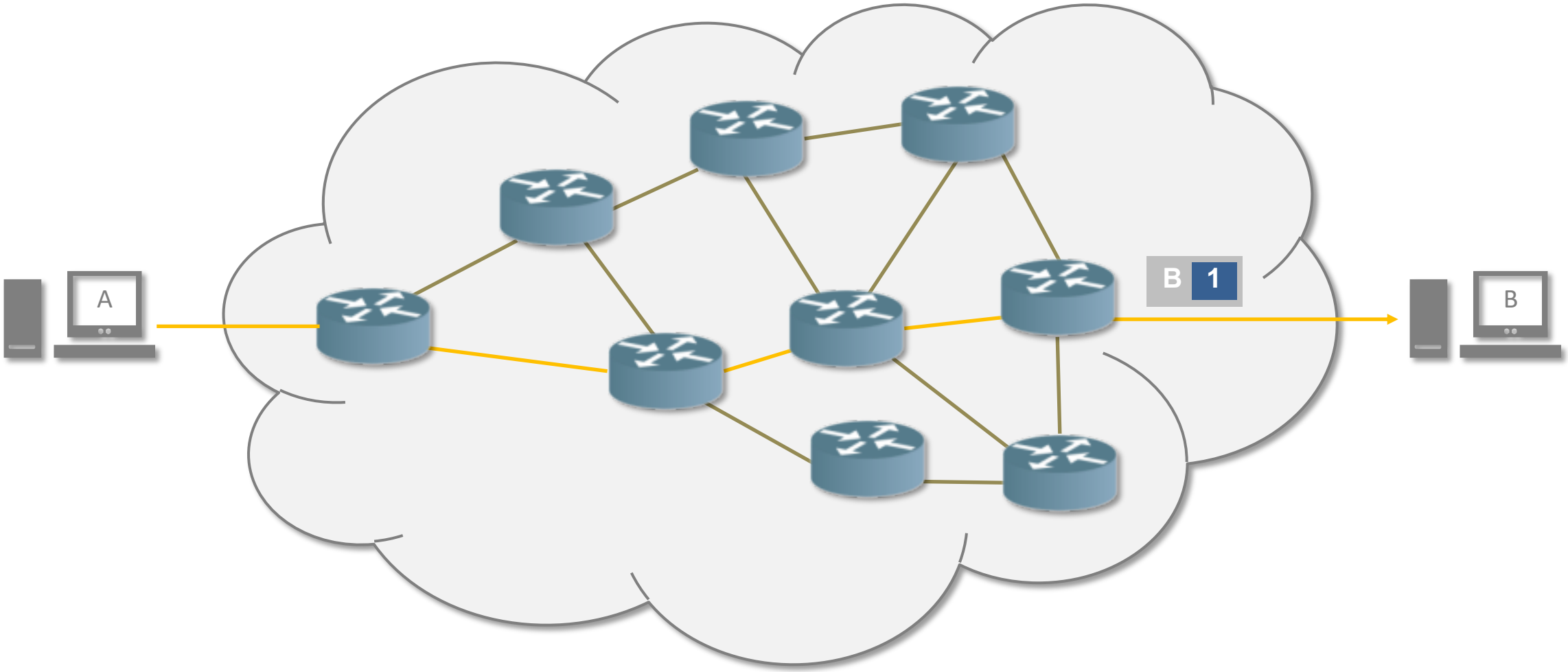
A sends data packet to B

IP Packet Switching



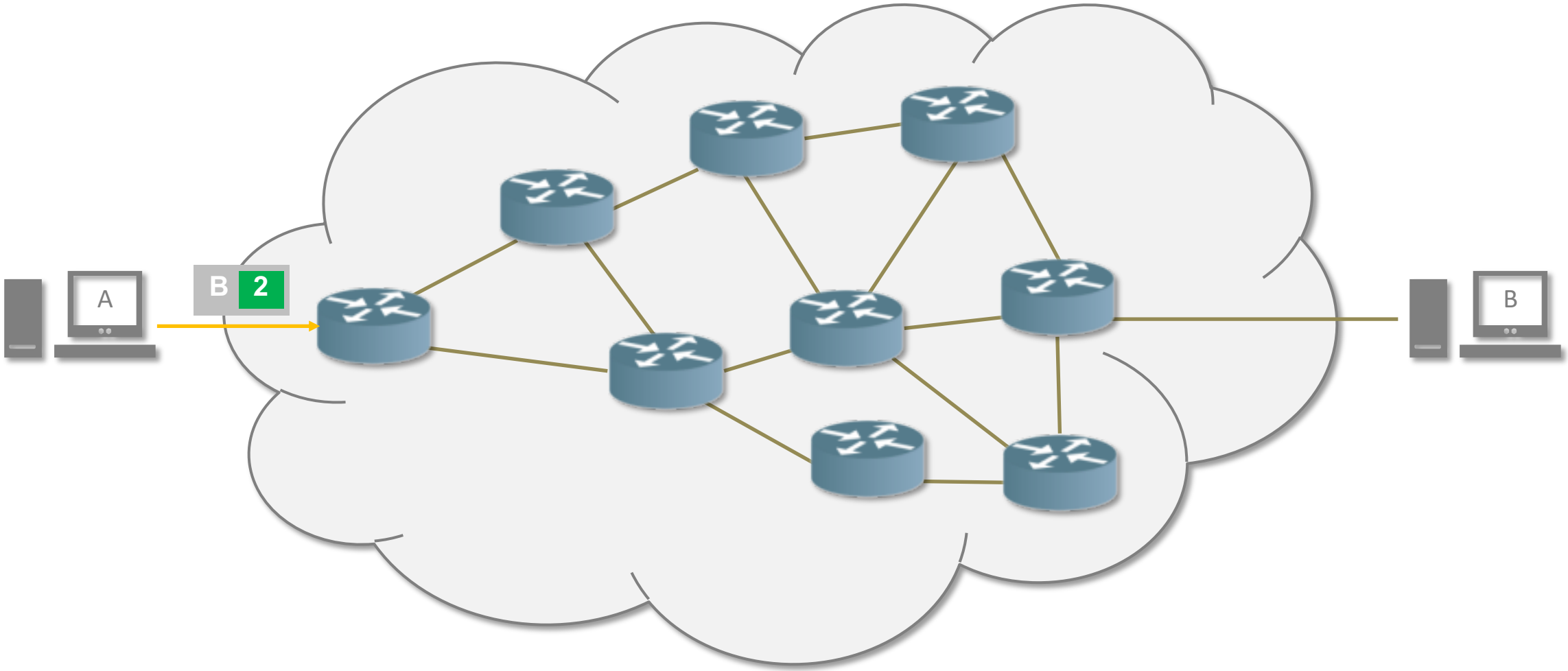
A sends data packet to B

IP Packet Switching



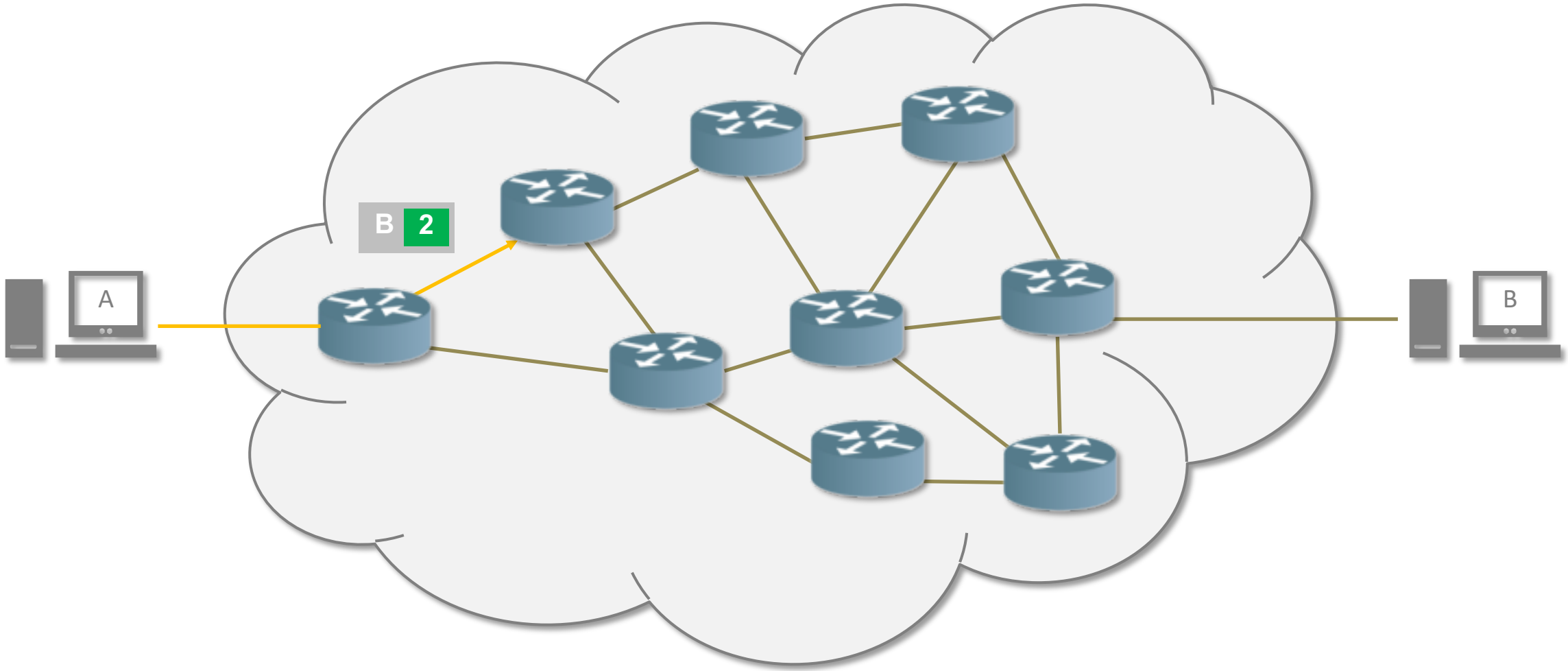
A sends data packet to B

IP Packet Switching



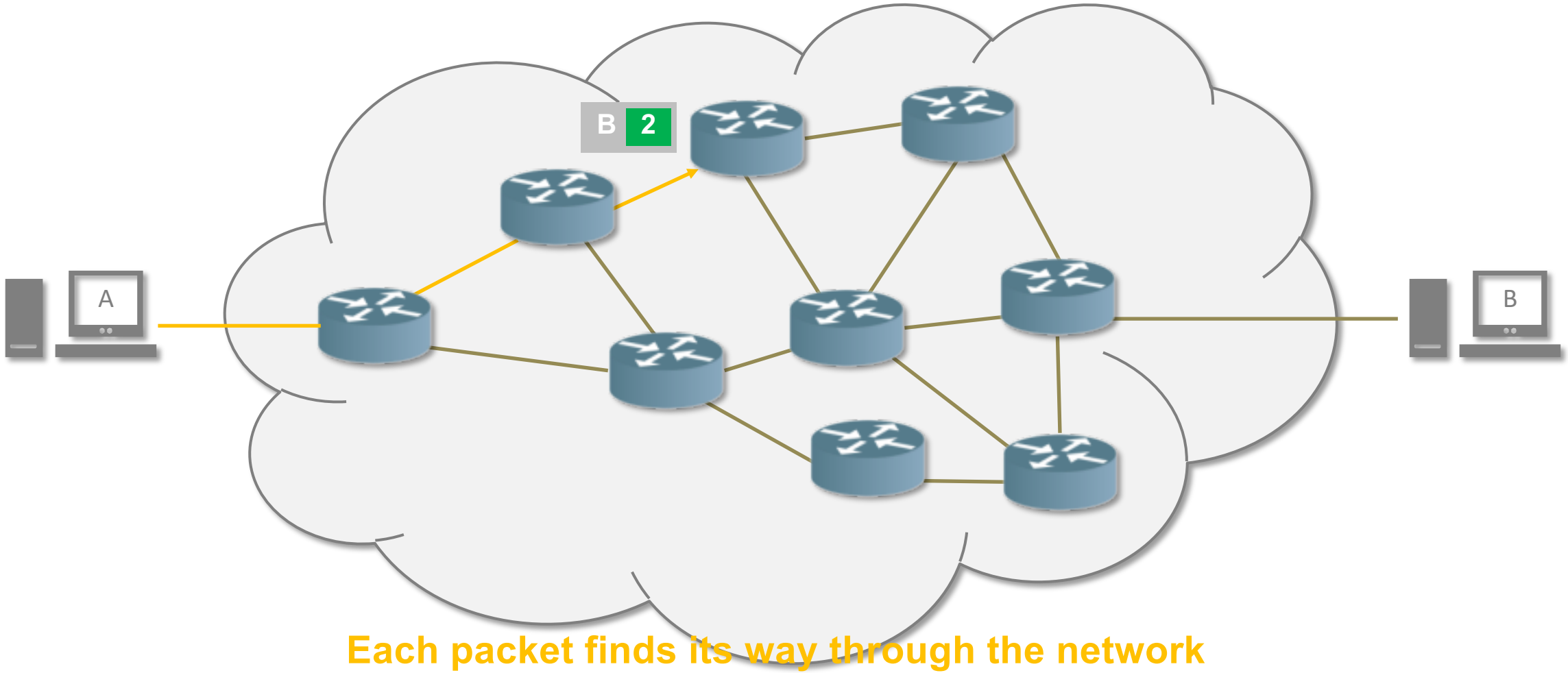
A sends another data packet to B

IP Packet Switching



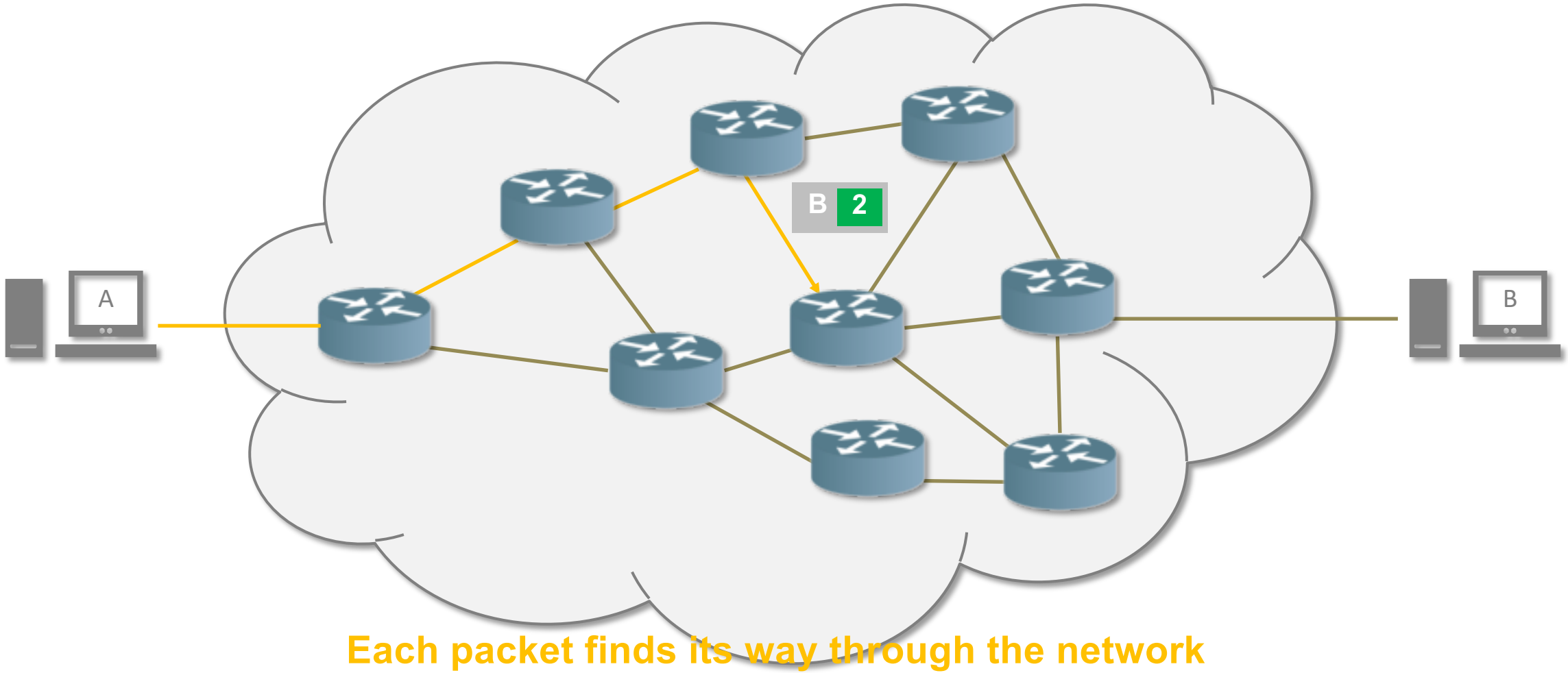
A sends another data packet to B

IP Packet Switching



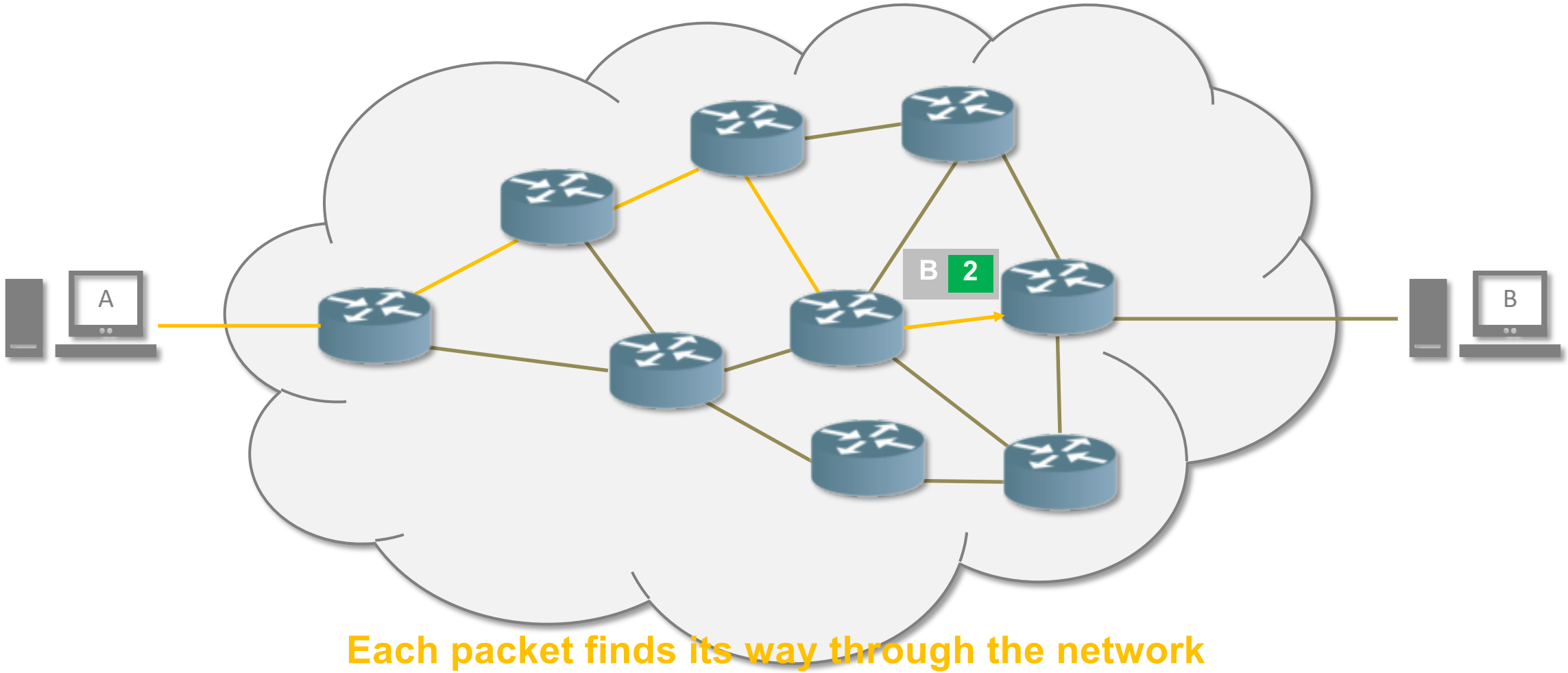
Each packet finds its way through the network by using its destination adress

IP Packet Switching



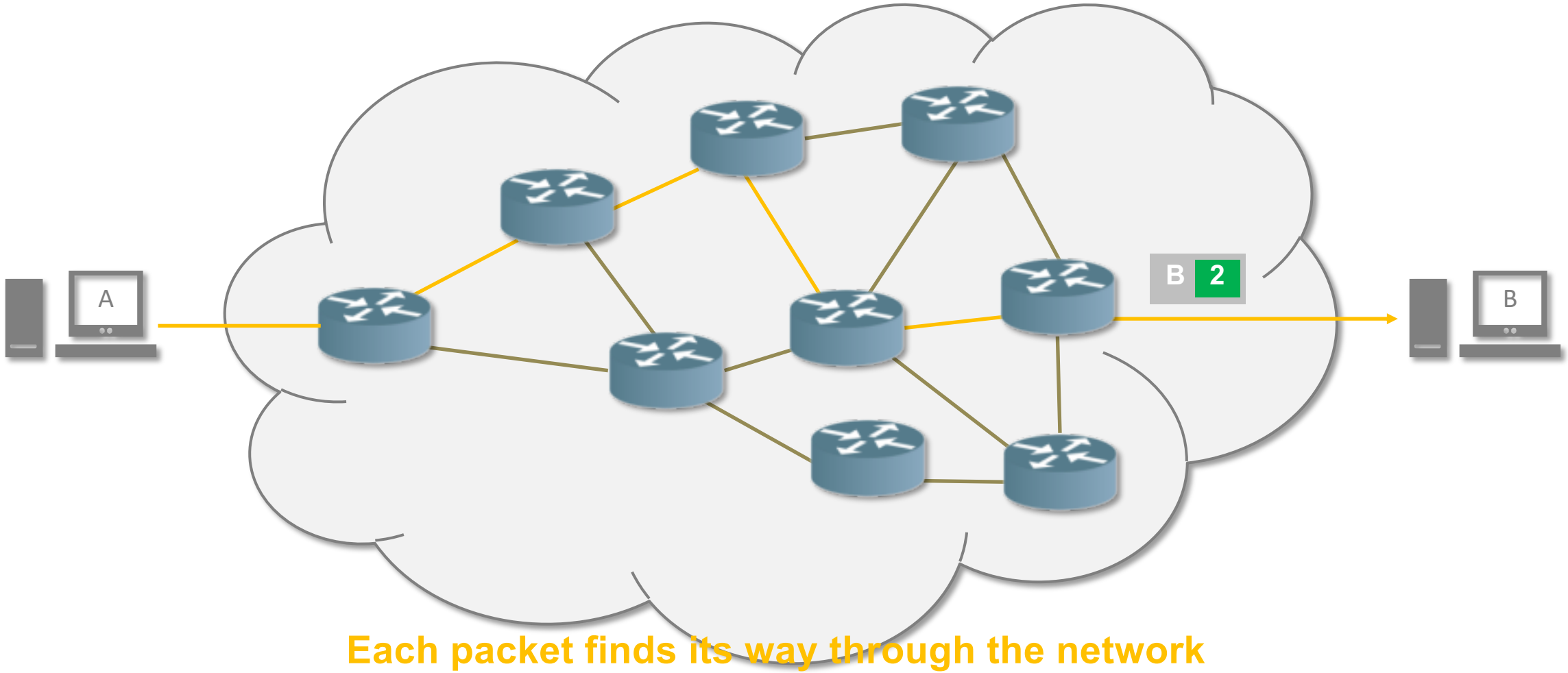
Each packet finds its way through the network by using its destination adress

IP Packet Switching



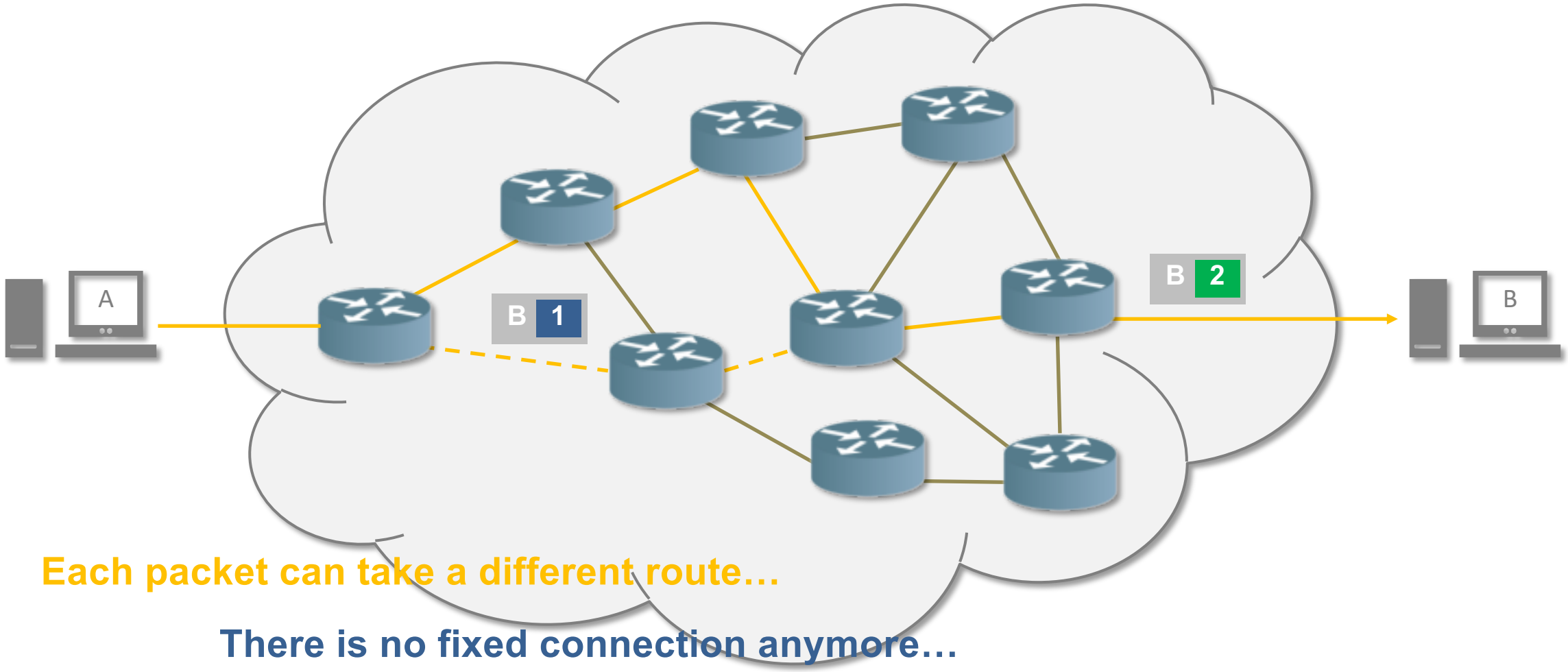
Each packet finds its way through the network by using its destination adress

IP Packet Switching



Each packet finds its way through the network by using its destination adress

IP Packet Switching



Each packet can take a different route…

There is no fixed connection anymore…

And thus, latency may vary…

Latency – determining factors

- Underlying network technology

Network technology	Network speed	Frame transmission time (MTU)
FE	100 Mbit/s	258,58 µs
GbE	**1 Gbit/s**	**25,86 µs (~ 1 sample time @ 48 kHz)**
10G	10 Gbit/s	2,6 µs
40G	40 Gbit/s	10,4 µs
100G	100 Gbit/s	260 ns

- Type of switching / routing equipment
  - Hardware / switching fabric (switches, most routers)
  - Software (some routers / firewalls etc.)
  - Mixed (routers w/ firewalls and / or advanced processing / filtering functions)
  - Store-and-forward vs. cut-through

Latency – determining factors

- Network topology
  - Number of hops
    - 1 hop @ GbE / switched: ~ 1 sample time @ 48 kHz
    - 4 hops @ FE / switched: ~ 1 ms
  - Unfortunately: delay not static due to dynamic traffic situation

Latency – determining factors

- Network topology
  - Number of hops
  - Distances (speed of light)

Distance	Time
1 m	3.3 ns
1 km	3.3 µs
4000 km (NYC – LA)	13.3 ms
to geostationary orbit	119 ms
around Earth's equator	134 ms
Moon to Earth	1.3 s

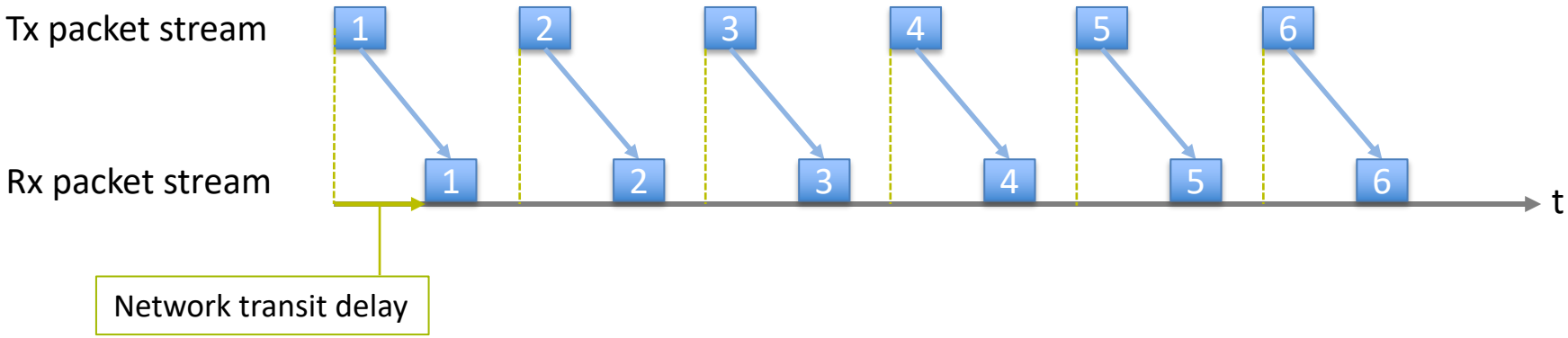
Latency – determining factors

- Network topology
  - number of hops
  - Distances (speed of light)

Distance	Time
1 m	3.3 ns
1 km	3.3 µs
6 km	**20 µs (1 sample @ 48 kHz)**
300 km	**1 ms (mandatory AES67 packet time)**
4000 km (NYC – LA)	13.3 ms
to geostationary orbit	119 ms
around Earth's equator	134 ms
Moon to Earth	1.3 s
3x round-trip Earth - Neptun	~ 1 d 51 m 19 s (RTP counter rollover @ 48 kHz)

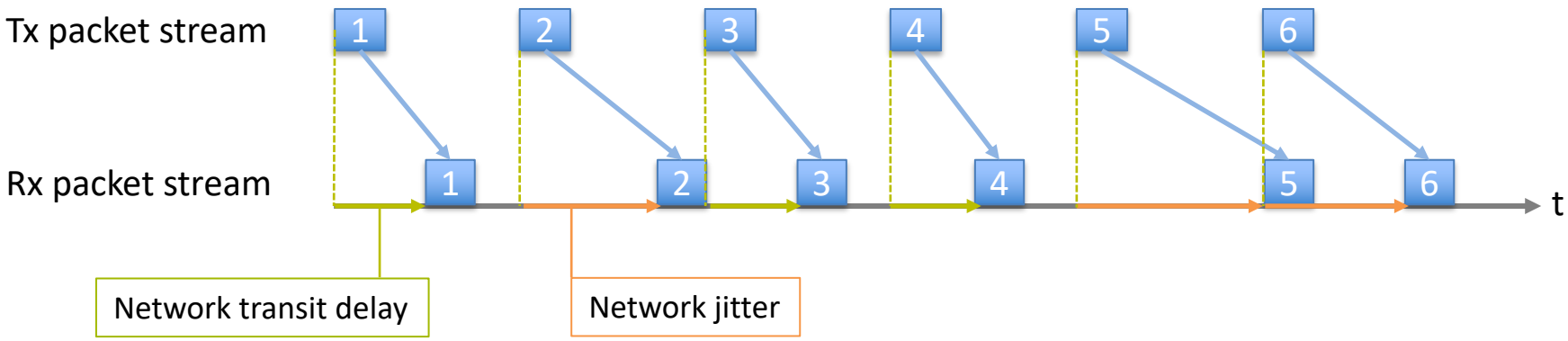
Latency – determining factors

- Network jitter (“PDV – packet delay variation”)



Latency – determining factors

- Network jitter (“PDV – packet delay variation”)

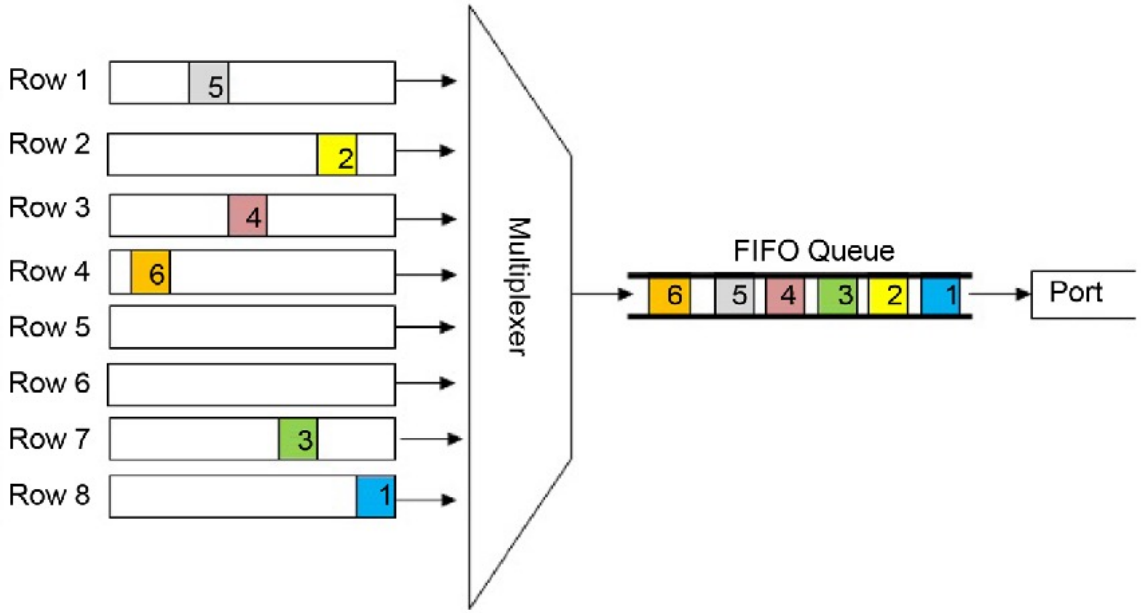


Latency – determining factors

- Network jitter (“PDV – packet delay variation”)
  - switch performance (hardware / software)
  - hops / routing (routing may change → no. of hops may change, network speed may change between hops)
  - bandwidth utilization

Latency – determining factors

- Network jitter (“PDV – packet delay variation”)
  - switch performance (hardware / software)
  - hops / routing (routing may change → no. of hops may change, network speed may change in between)
  - bandwidth utilization – without QoS:



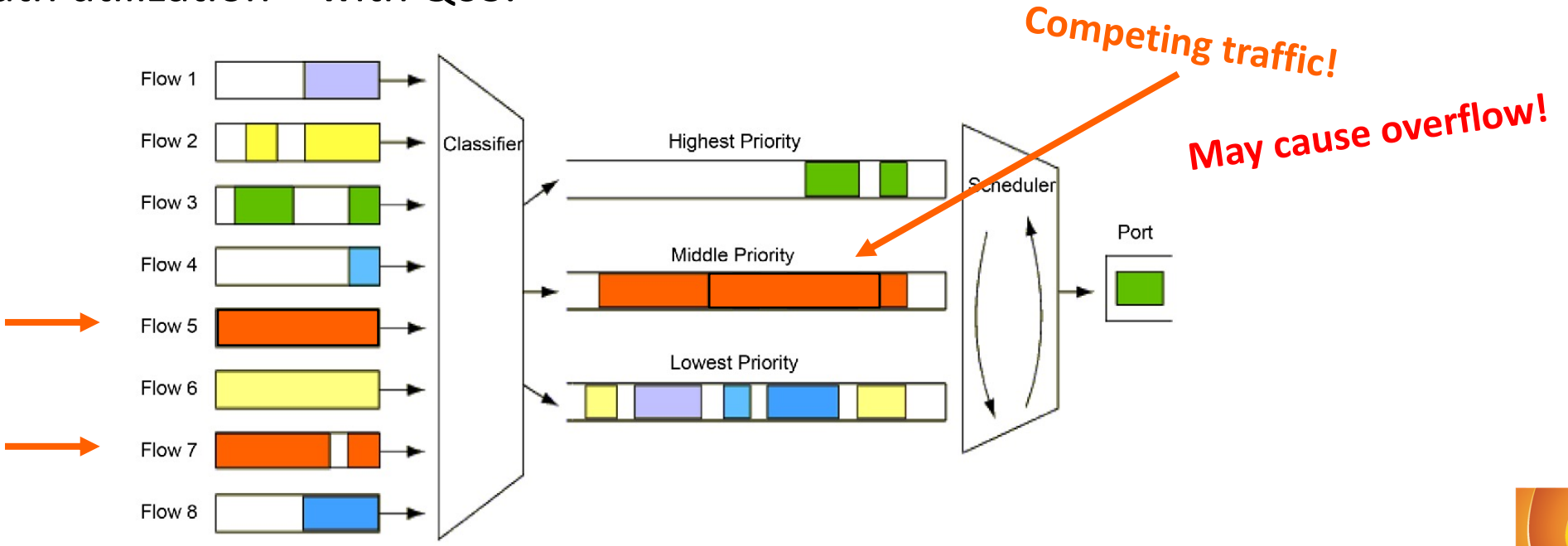
Latency – determining factors

- Network jitter (“PDV – packet delay variation”)
  - switch performance (hardware / software)
  - hops / routing (routing may change → no. of hops may change, network speed may change in between)
  - bandwidth utilization – with QoS:

QoS – Differentiated Services (DiffServ)

Latency – determining factors

- Network jitter (“PDV – packet delay variation”)
  - switch performance (hardware / software)
  - hops / routing (routing may change → no. of hops may change, network speed may change in between)
  - bandwidth utilization – with QoS:



RAVENNA AES67 & ST 2110 built-in

Latency – worst case multi-hop calculations

	Bytes		100	1000	Mbit/s			
MTU RTP	1460		Net Data Rate Ethernet					
MTU IP	1472		98,70	987,03	Mbit/s			
MTU Ethernet	1500		Transfer Time					
Ethernet Frame Size	1518	→	123,04	12,30	µs			
Ethernet Frame Size w/ VLAN	1522	→	123,36	12,34	µs			
Preamble:	8							
Interframe Gap:	12							
Formula for max. Bridge Latency	MaxLatency = t_BridgeDelay + t_MaxFrame + n * t_StreamFrame					n= number of concurrent RTP streams fro		
t_BridgeDelay			10,24	1,024	µs			
max. BridgeLatency			**258,58**	**25,86**	µs	n= 1		
Max. Ethernet Frame Size	1522							

	1. Hop			2. Hop			3. Hop	
	Bytes	Delay (µs)		Bytes	Delay (µs)		Bytes	Delay
1. stream		FE	GbE		FE	GbE		FE
channels	8			*8*			*8*	
bytes / sample	3			*3*			*3*	
frames/packet	48			*48*			*48*	
sample rate	48000			*48000*			*48000*	
payload	1152			*1152*			*1152*	
frame size	1214			*1214*			*1214*	
frame + pre + IFG	1234			*1234*			*1234*	
→ **Worst Case Latency**		**1713,12**	**171,31**		**3426,24**	**342,62**		**5139,36**
Competing streams								
number	15			15			15	
channels	8			8			8	
bytes / sample	3			3			3	
frames/packet	48			48			48	
sample rate	48000			48000			48000	
payload	1152			1152			1152	

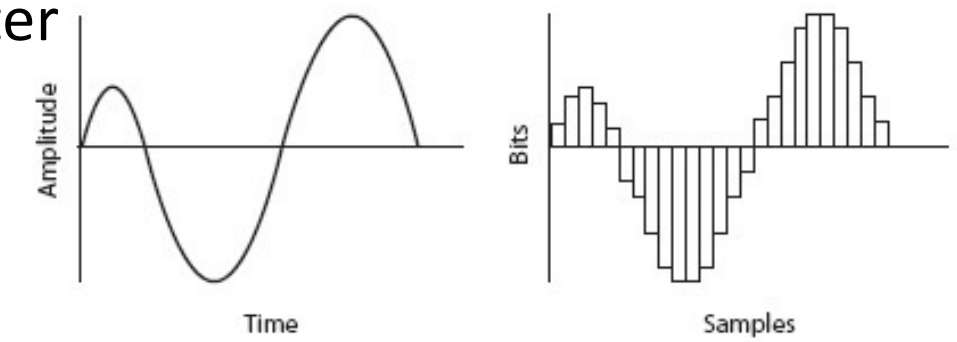
AES NEW YORK AUDIO ENGINEERING CONVENTION

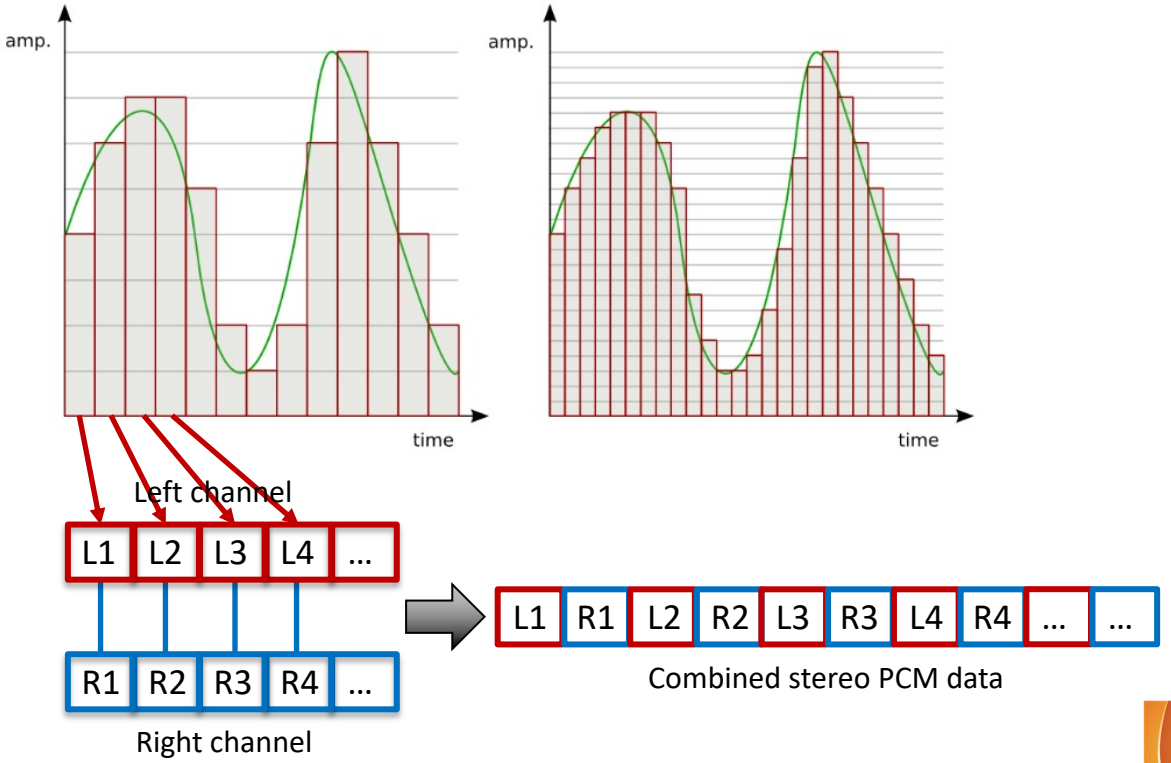
Latency – determining factors

- Stream / packet configuration & sampling rate

AES67 – (mandatory) encoding parameter

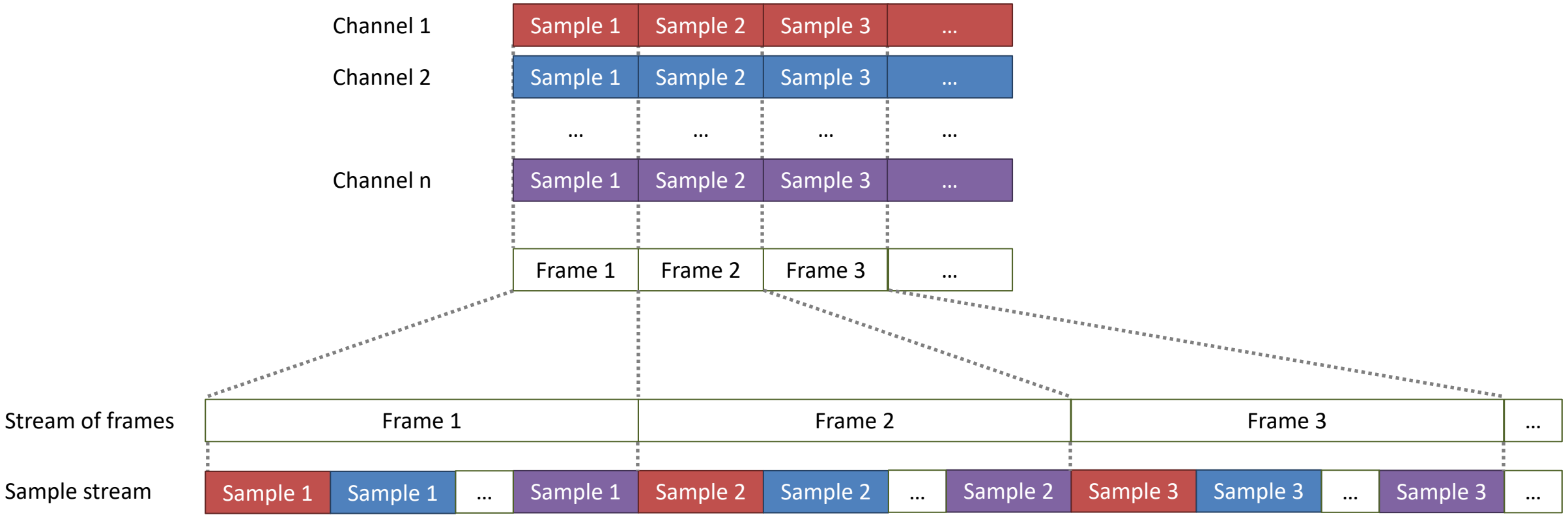
- **Sampling rate**: 48 kHz
  (optional: 44.1 / 88.2 / 96 kHz)
- PCM **bit width**: 16 and 24 bits
- **# of channels** per stream: 1..8





AES67 – (mandatory) encoding parameter

- **Packet time**: # of frames (samples x channels) per packet at given sample rate

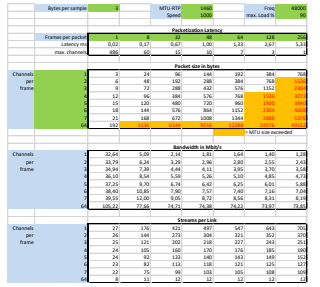


RAVENNA
AES67 & ST 2110 built-in

Latency – determining factors

- Stream / packet configuration & sampling rate

	Bytes per sample	3		MTU-RTP	1460		Freq	48000
	Frames per packet	1	8	12	32	48	128	256
	Latency ms	0,02	0,17	0,25	0,67	1,00	2,67	5,33
Channels per frame	1	3	24	36	96	144	384	768
	2	6	48	72	192	288	768	1536
	3	9	72	108	288	432	1152	2304
	4	12	96	144	384	576	1536	3072
	5	15	120	180	480	720	1920	3840
	6	18	144	216	576	864	2304	4608
	7	21	168	252	672	1008	2688	5376
	10	30	240	360	960	1440	3840	7680

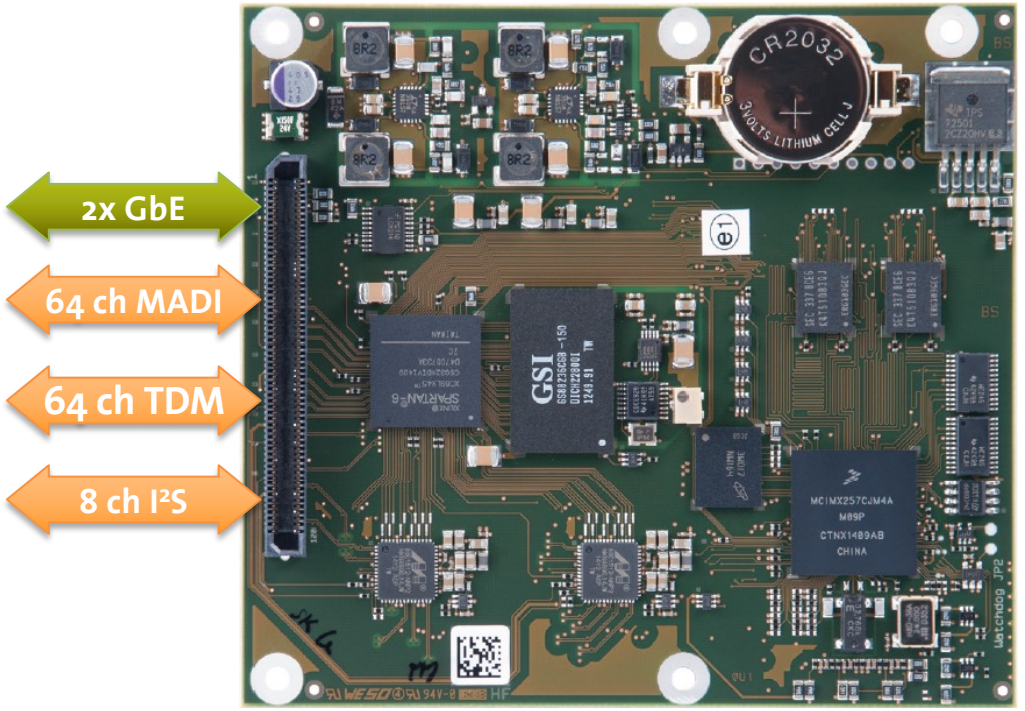


Latency – determining factors

- Sender / receiver implementation
  - Hardware
  - Embedded
  - Software (VSC)

COMi.MX – RAVENNA / AES67 SoM

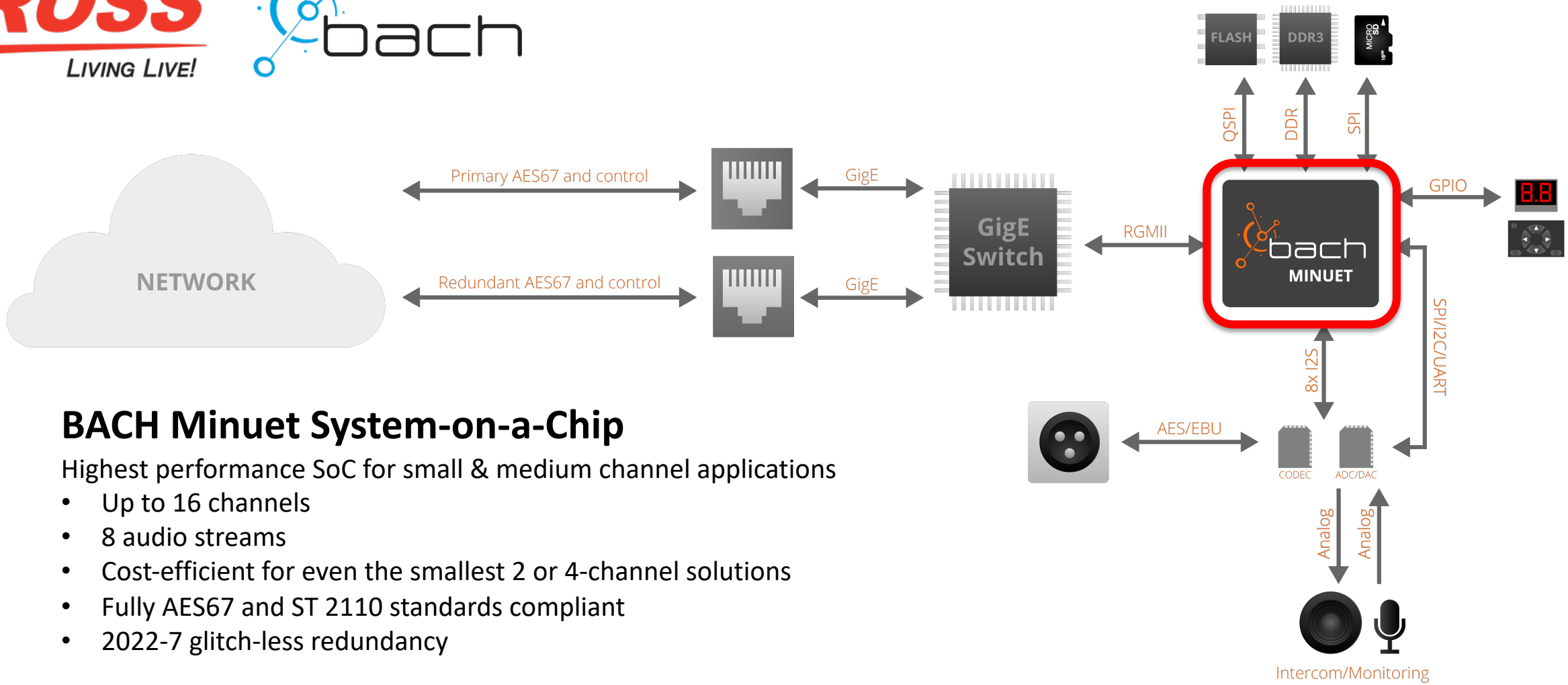
- Fully self-contained RAVENNA implementation
- **Audio interfaces: I²S (8 ch) / TDM, MADI (64 ch)**
- **Up to 192 kHz sampling rate**
- **Lowest latency support: down to 1 sample/packet!**
- **2 GbE NICs w/ ST2022-7 redundancy or load balancing**
- **2x 64 channels in & out**
- Full AES/EBU bit-transparent operation supported
- Jitter / delay buffer up to 40 ms per channel
- 4-tier 256 x 256 audio matrix
- Full AES67 & ST2110-30/-31 support



RAVENNA AES67 & ST 2110 built-in

ROSS LIVING LIVE!

bach



BACH Minuet System-on-a-Chip

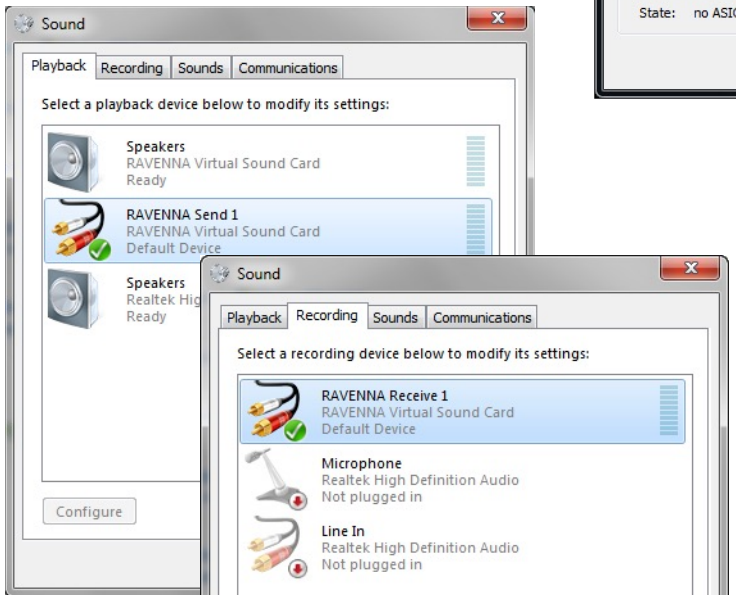
Highest performance SoC for small & medium channel applications

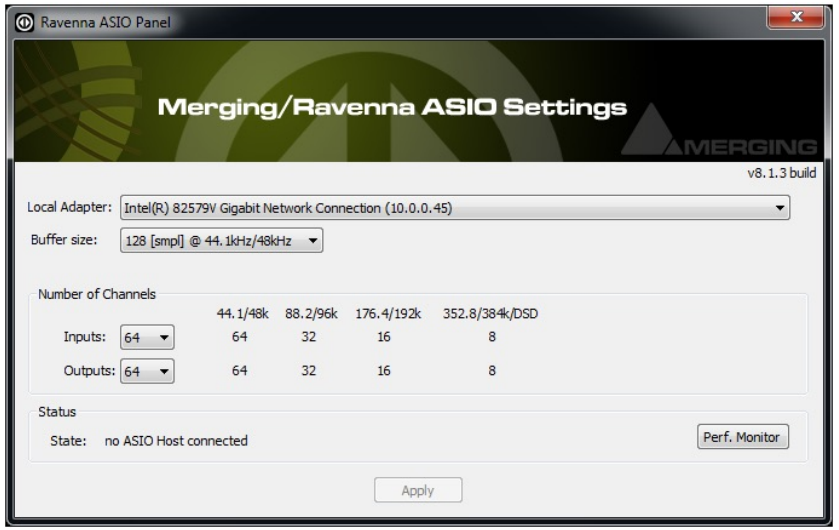
- Up to 16 channels
- 8 audio streams
- Cost-efficient for even the smallest 2 or 4-channel solutions
- Fully AES67 and ST 2110 standards compliant
- 2022-7 glitch-less redundancy

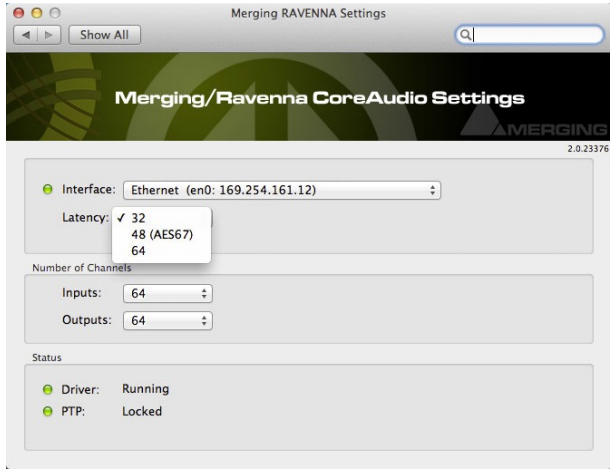
AES NEW YORK AUDIO ENGINEERING CONVENTION

VSC – Virtual Sound Cards

- Windows / MacOS / Linux
- Up to 64 channels playback / record
- Typ. processing latency: ~ 10 ms







Latency – determining factors

- Stream alignment

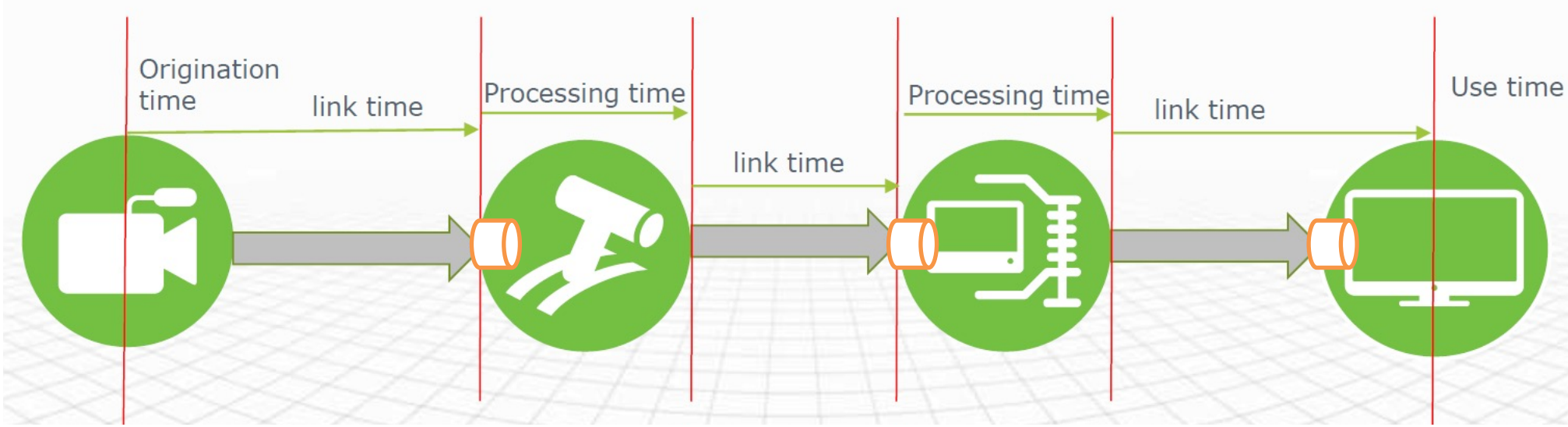
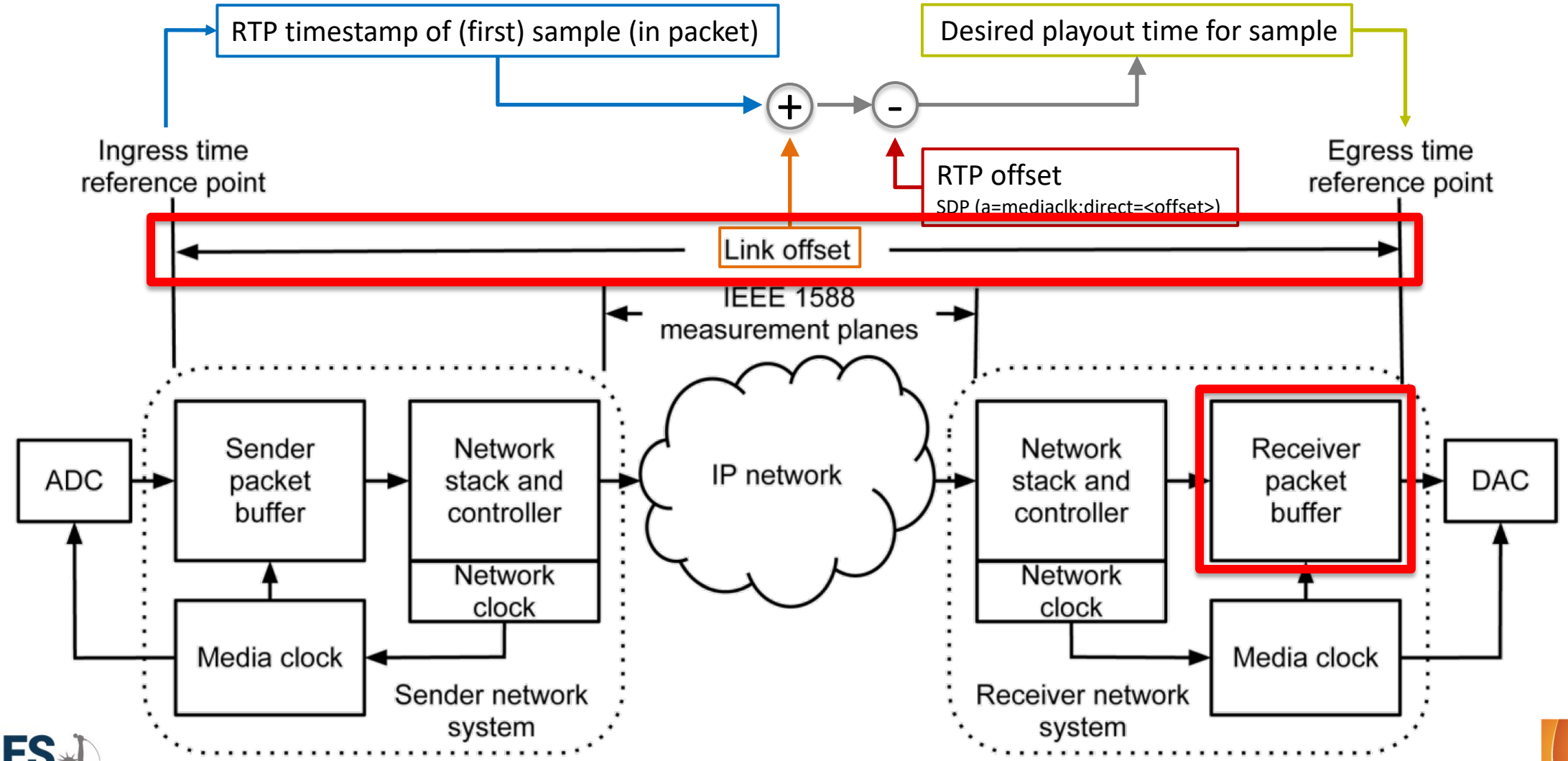


Image courtesy of Andy Rayner (Nevion)

Latency – determining factors

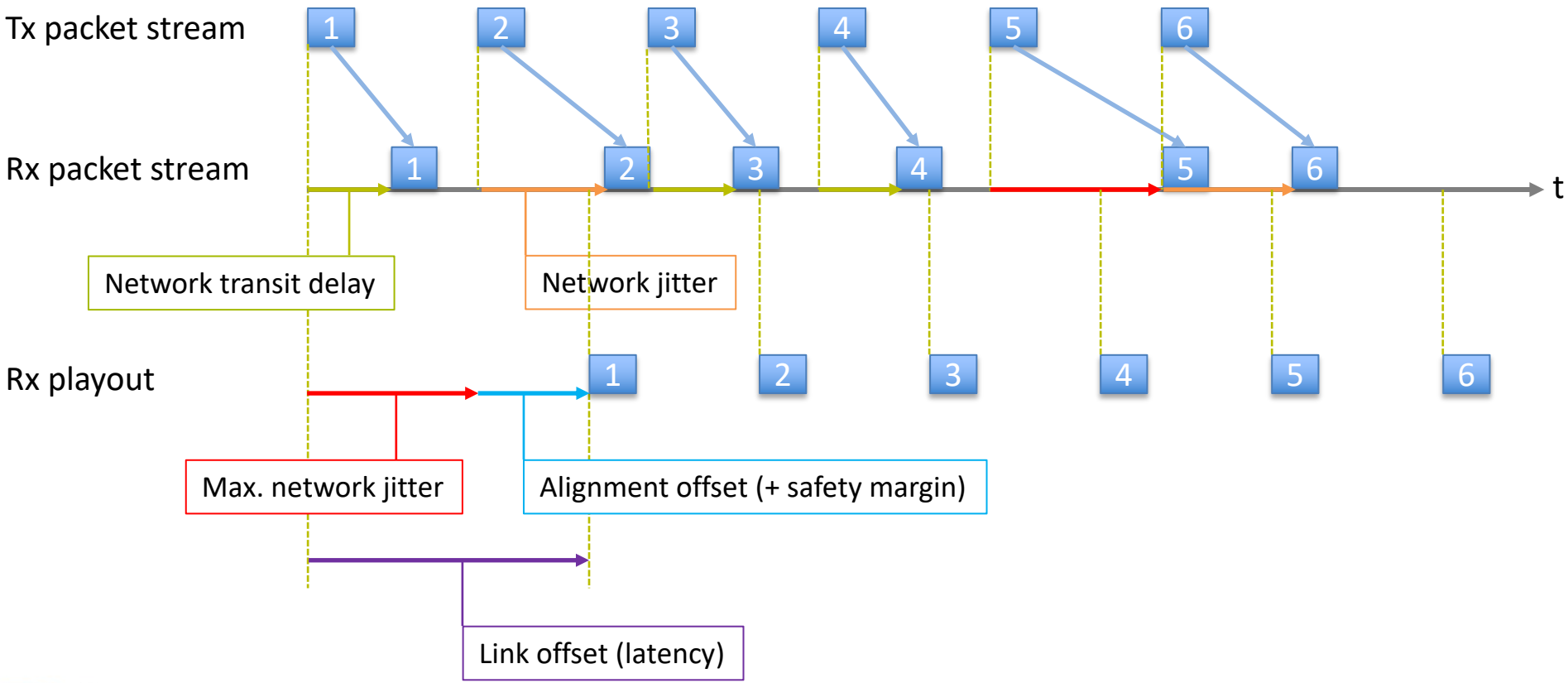
→ Link offset & receiver buffer size

AES67 synchronization - link offset (latency)



Latency – determining factors

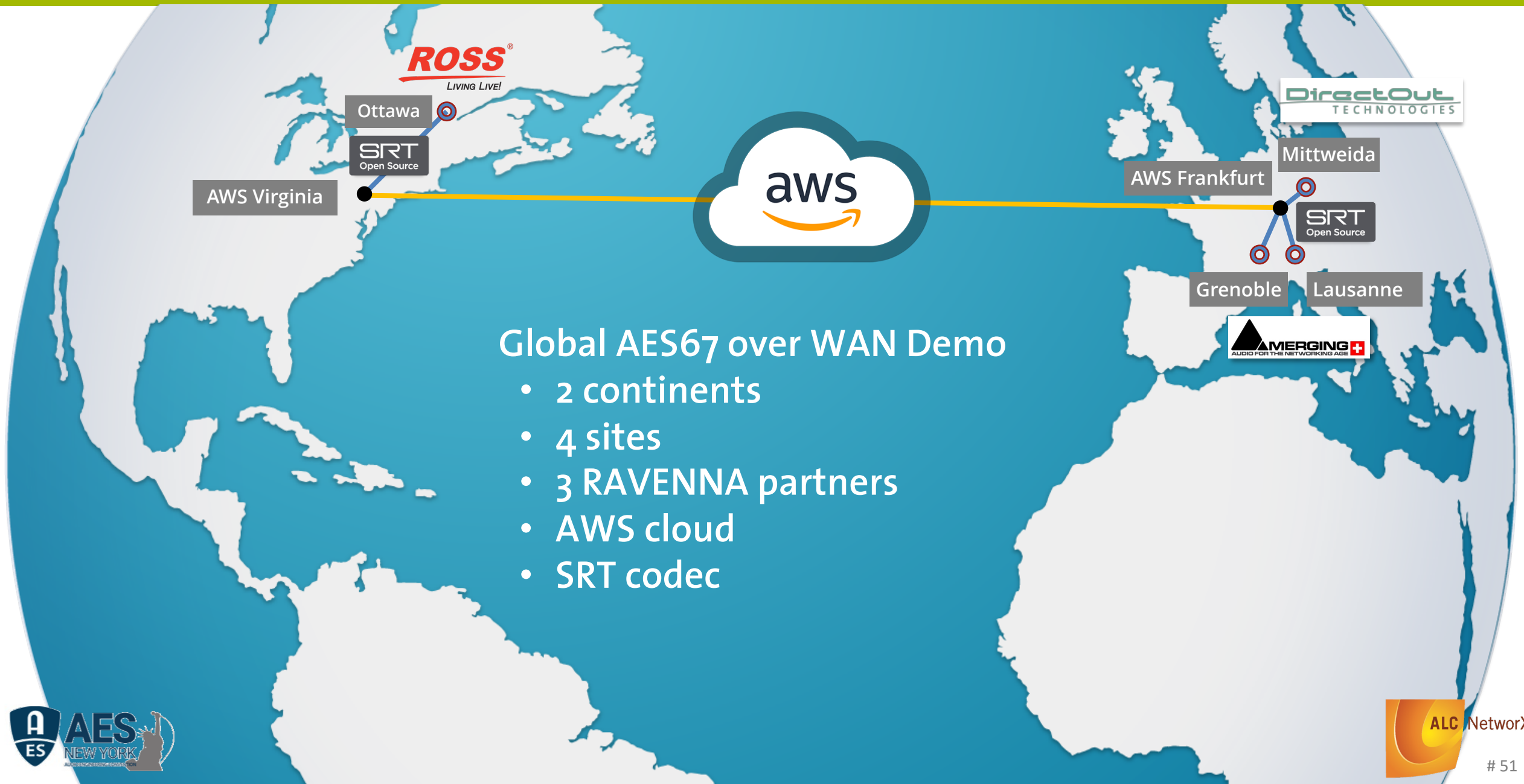
→ Link offset & receiver buffer size



Latency – determining factors

Global AES67 over WAN Demo

Latency in AoIP Systems



Global AES67 over WAN Demo

- 2 continents
- 4 sites
- 3 RAVENNA partners
- AWS cloud
- SRT codec

AES67 over WAN - lessions learned

- Synchronization works via GPS-synchronized local GMs (no need to transport PTP)
- There is packet loss but this can managed via SRT
- Latency ranged from 200 to 600 ms
- Receivers need to have deep receive buffers or mechanisms to compensate for the network delay and PDV

Latency – determining factors (summary)

- Underlying network technology
- Network topology
- Network jitter (PDV)
- Stream / packet configuration (packet time)
- Sender / receiver implementation
- Stream alignment

→ Link offset

→ Receiver buffer size

Receiver buffer size ≠ Link offset (latency)

Takeaways:

⇨ **sub-milliseconds** latency is achievable

⇨ req'd receiver buffer size depends on application

- AES67 requires 3 ms, but recommends 20 ms

Questions?

More information…

RAVENNA / AES67 / SMPTE ST 2110 Resources:

RAVENNA AES67 & ST 2110 built-in — www.ravenna-network.com/resources

AIMS Alliance for IP Media Solutions — www.aimsalliance.org (resources)

ravenna@alcnetworx.de

You've made it!

Contact information:

Andreas Hildebrand
ALC NetworX GmbH

ravenna@alcnetworx.de

CELEBRATING 12 YEARS
RAVENNA
AES67 & ST 2110 built-in

www.ravenna-network.com

Supplemental Slides

RAVENNA
AES67 & ST 2110 built-in

Approximate light signal travel times

Distance	Time
1 m	3.3 ns
1 km	3.3 µs
6 km	**20 µs (1 sample @ 48 kHz)**
300 km	**1 ms (mandatory AES67 packet time)**
4000 km (NYC – LA)	13.3 ms
to geostationary orbit	119 ms
around Earth's equator	134 ms
Moon to Earth	1.3 s
3x round-trip Earth - Neptun	~ 1 d 51 m 19 s (RTP counter rollover @ 48 kHz)
Proxima Centauri (nearest star)	4.2 years
across the Milky Way	100000 years
Andromeda Galaxy (nearest galaxy)	2.5 million years

AES NEW YORK AUDIO ENGINEERING CONVENTION

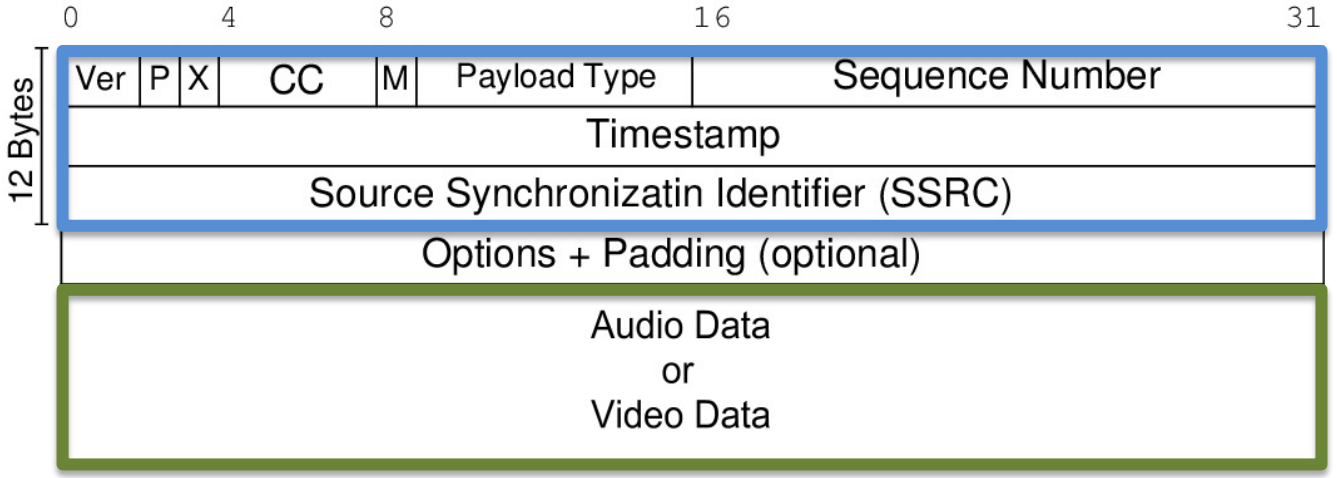

Google
Network process to application
7. Application
Data representation and encryption
6. Presentation
Interhost communication
5. Session
4. Transport
1. Physical
DNS , WWW/HTTP
P2P,EMAIL/POP,SMTP
Telnet, FTP
recognizing data
HTML, DOC, JPEG, MP3, AVI
Sockets, Session establishment
in TCP, SIP, RTP
RPC - Named pipes,
TCP, UDP, SCTP, SSL, TLS
IP, IPsec, ICMP, IGMP, OSPF
Ethernet, 802.11, MAC/LLC,VLAN, ATM, HDP
Fibre Channel Frame Relay, HDLC,
PPP, Q.921, Token Ring, ARP
RS-232, RJ45, V.34, 100BASE-TX, SDH, DSL, 802.11
7. Application
6. Presentation
5. Session
4. Transport
3. Network
2. Data link
1. Physical

RTP Packets (Layer 5)

- A format-agnostic transport protocol for real-time media data (RFC 3550)
- Consists of an RTP header and a payload area

RTP header

RTP payload

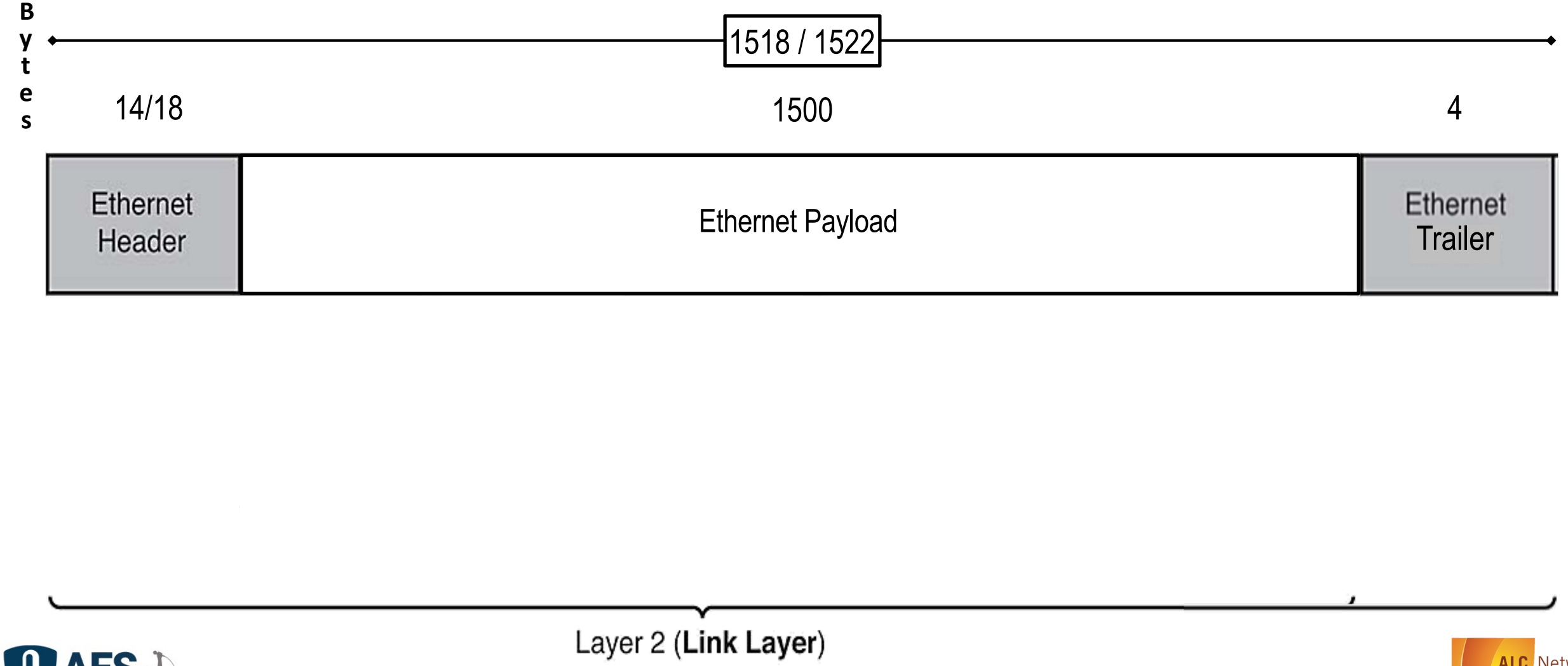


RTP Packets (Layer 5)

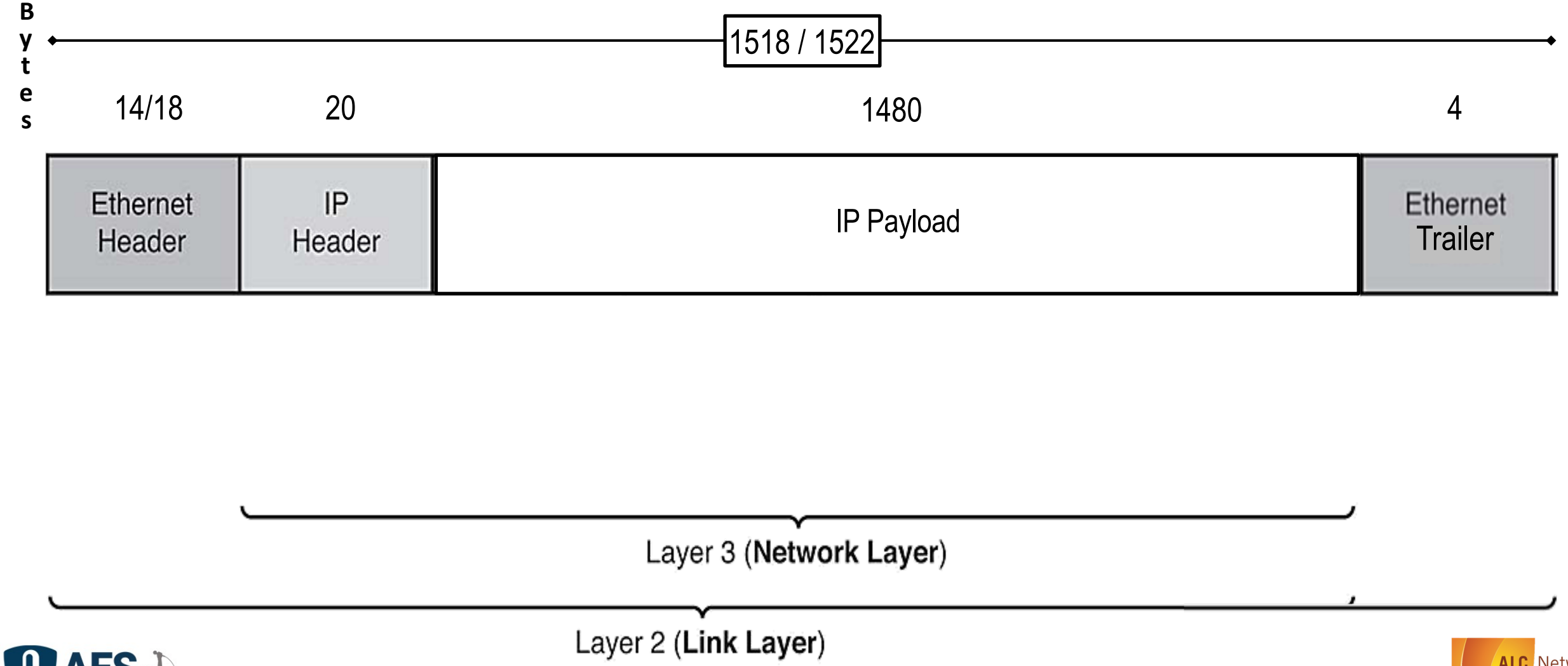
- A format-agnostic transport protocol for real-time media data (RFC 3550)
- Consists of an RTP header and a payload area
- RTP header (overhead) = **12 Bytes**
- IP + UDP + RTP overhead = 20 + 8 + 12 = **40 Bytes**
- MTU (maximum transmission unit, largest size of a packet that can be transmitted without being split): 1500 Bytes in an IP/Ethernet LAN

⇨ **0 to 1460 bytes** available for RTP **payload** data per packet

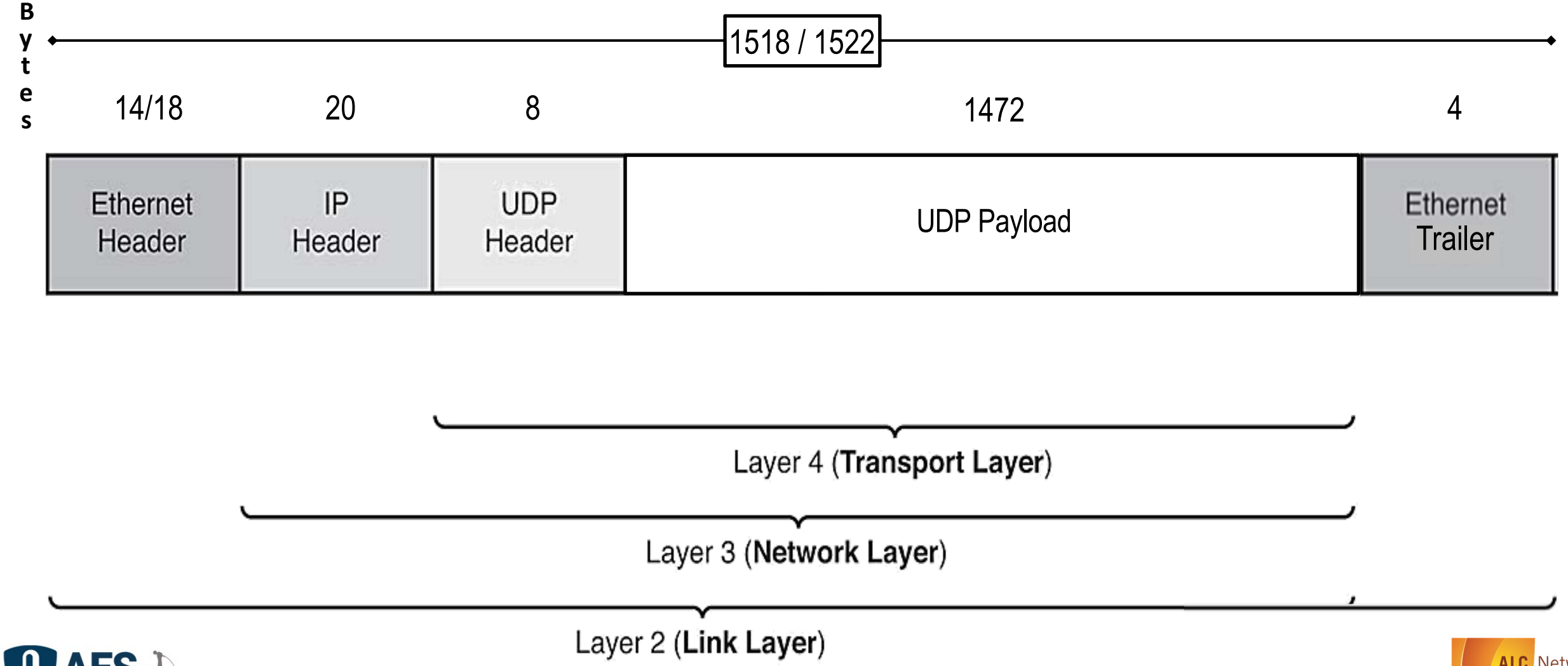
Layered Packet Encapsulation



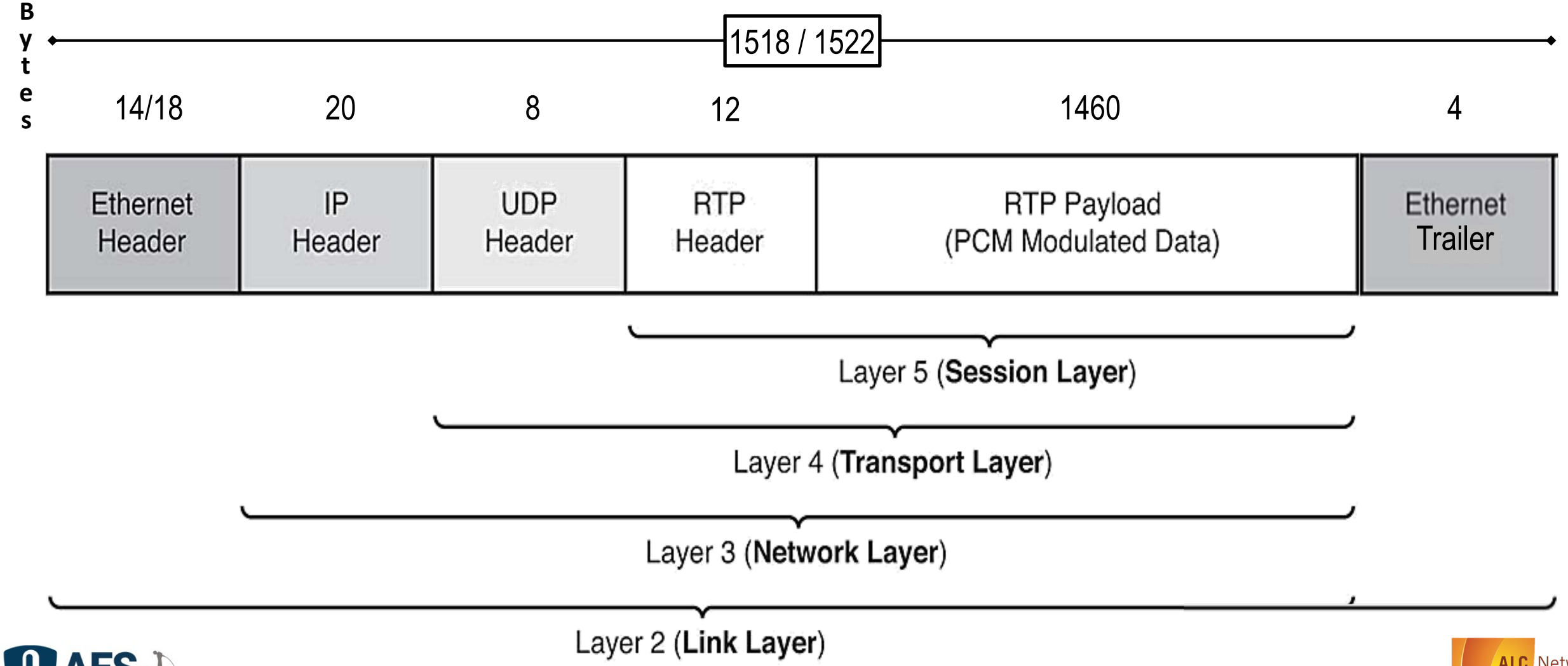
Layered Packet Encapsulation



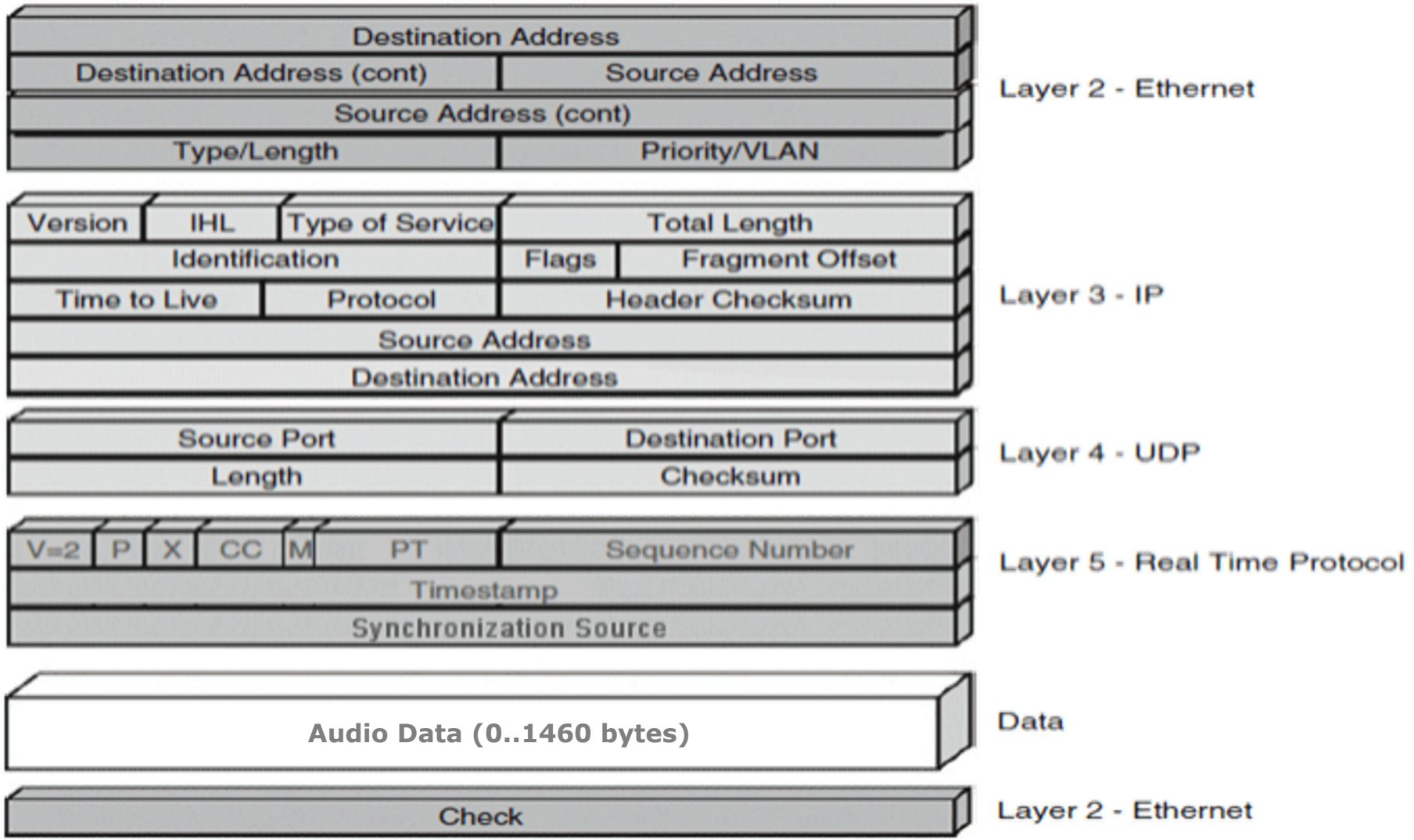
Layered Packet Encapsulation



Layered Packet Encapsulation



RTP - Layered Packet Encapsulation



RTP Packets - Payload

- Consist of RTP header, optional payload headers and the payload itself
- RTP overhead = **12 Bytes**
- IP + UDP + RTP overhead = 20 + 8 + 12 = **40 Bytes**
- MTU (maximum transmission unit, largest size of a packet that can be transmitted without being split) 1500 Bytes in an IP/Ethernet LAN: in principle **0 to 1460 bytes** available for RTP **payload** data per packet
- Examples:

  #1: 16 bit PCM, 2 channels, 96 samples (2 ms @ 48kHz): 384 bytes
  #2: 24 bit PCM, 8 channels, 48 samples (500 µs @ 96kHz): 1152 bytes
  #3: AES3 24 Bit PCM, 64 channels, 5 samples (104 µs @ 48 kHz): 1280 bytes

RAVENNA
AES67 & ST 2110 built-in

AES67 technology components:

- Network: IPv4 (IPv6), unicast / multicast & IGMPv2
- Transport: RTP/AVP (RFC 3550 & 3551) / UDP / IP
- Quality of Service: DiffServ w/ 3 suggested traffic classes (DSCP)

Quality of Service

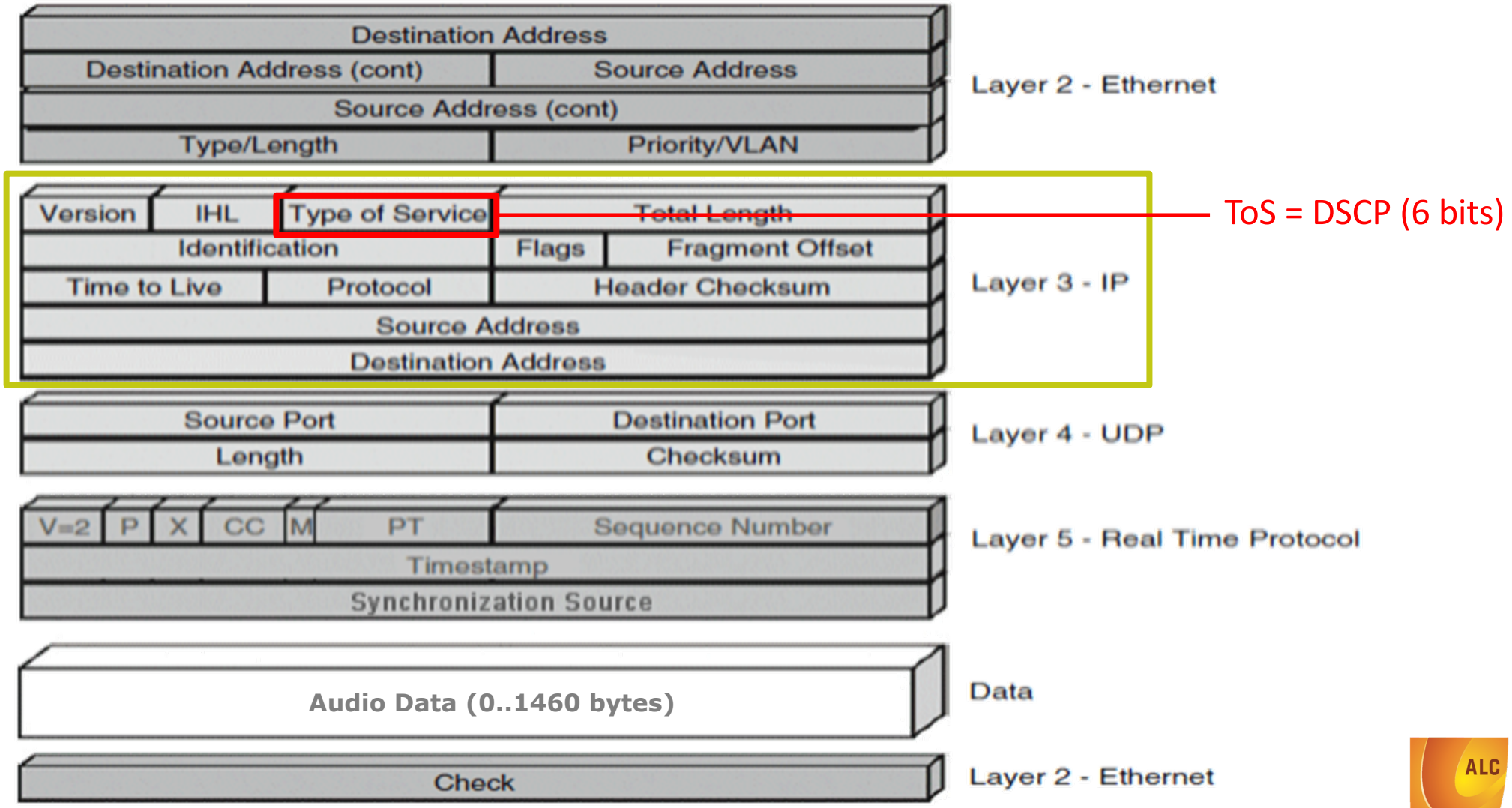
Coffee?

Lunch?

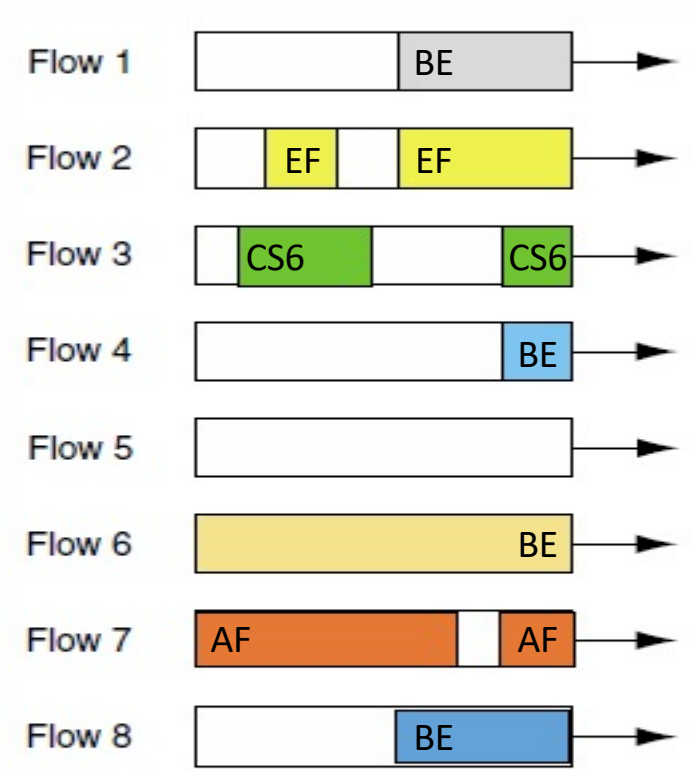
QoS – Differentiated Services (DiffServ)

- Defined in RFC 2474
- Defines up to 64 traffic classes (i.e. EF, AFx, CSx, BE etc.)
- Packets are tagged with DSCP value (0 – 63)

QoS – Differentiated Services (DiffServ)



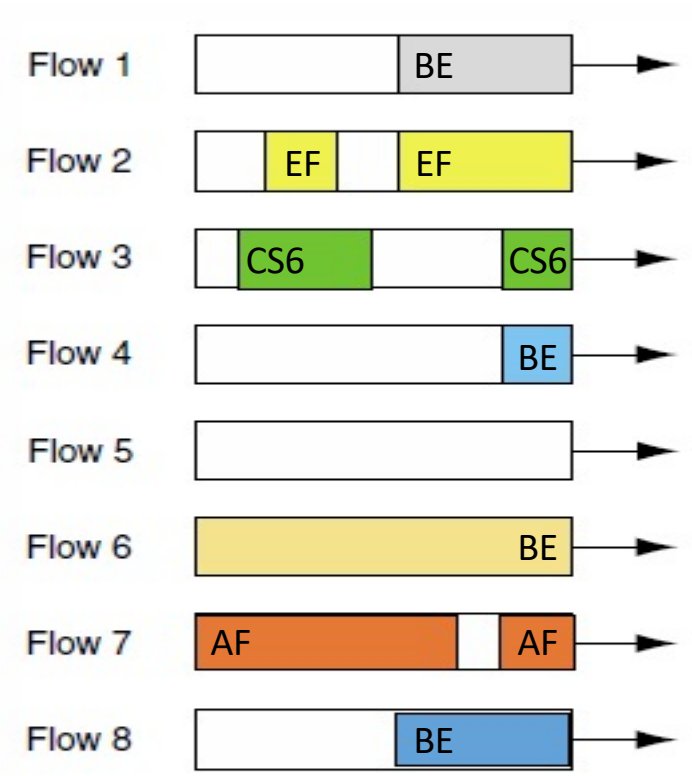
QoS – Differentiated Services (DiffServ)



QoS – Differentiated Services (DiffServ)

- Defined in RFC 2474
- Defines up to 64 traffic classes (i.e. EF, AFx, CSx, BE etc.)
- Packets are tagged with DSCP value (0 – 63)
- Switches store packets in different priority queues (requires proper configuration)

QoS – Differentiated Services (DiffServ)

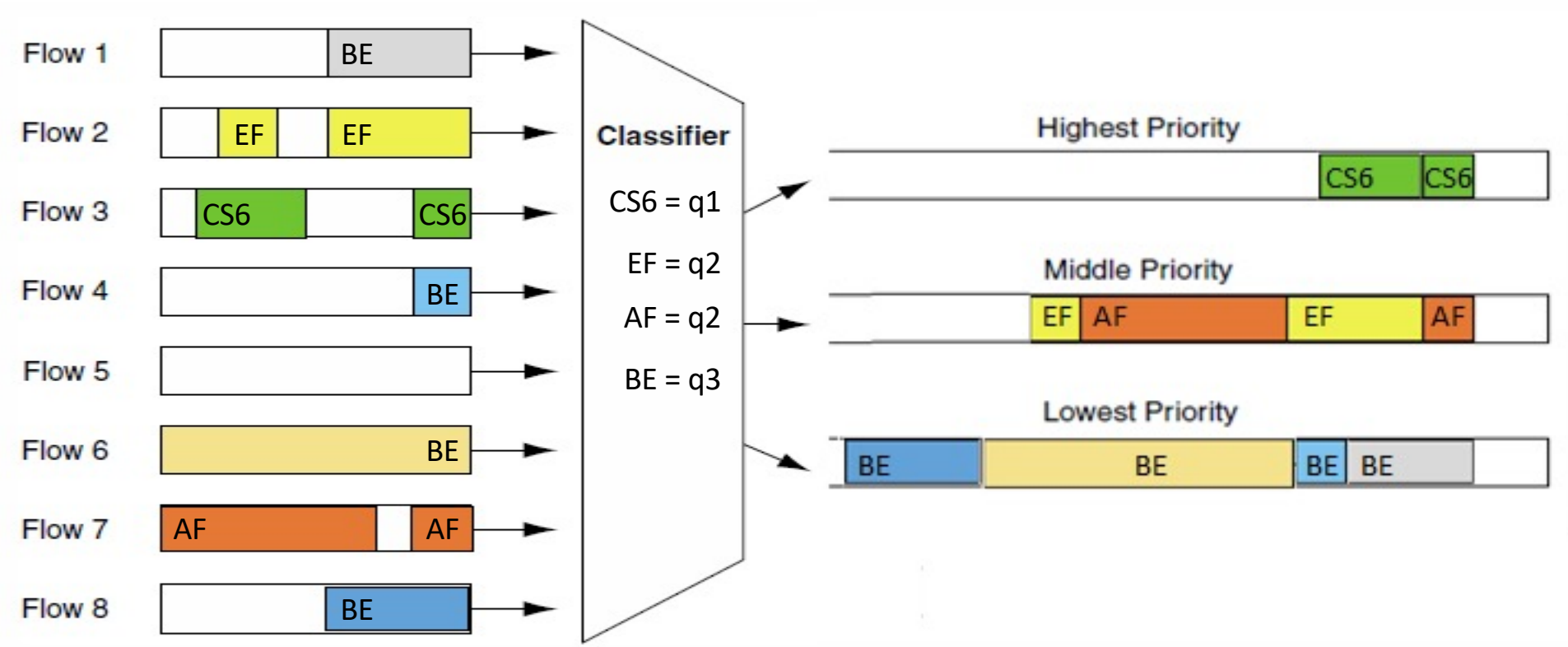


Queue 1 (High Priority)

Queue 2 (Medium Priority)

Queue 3 (Low Priority)

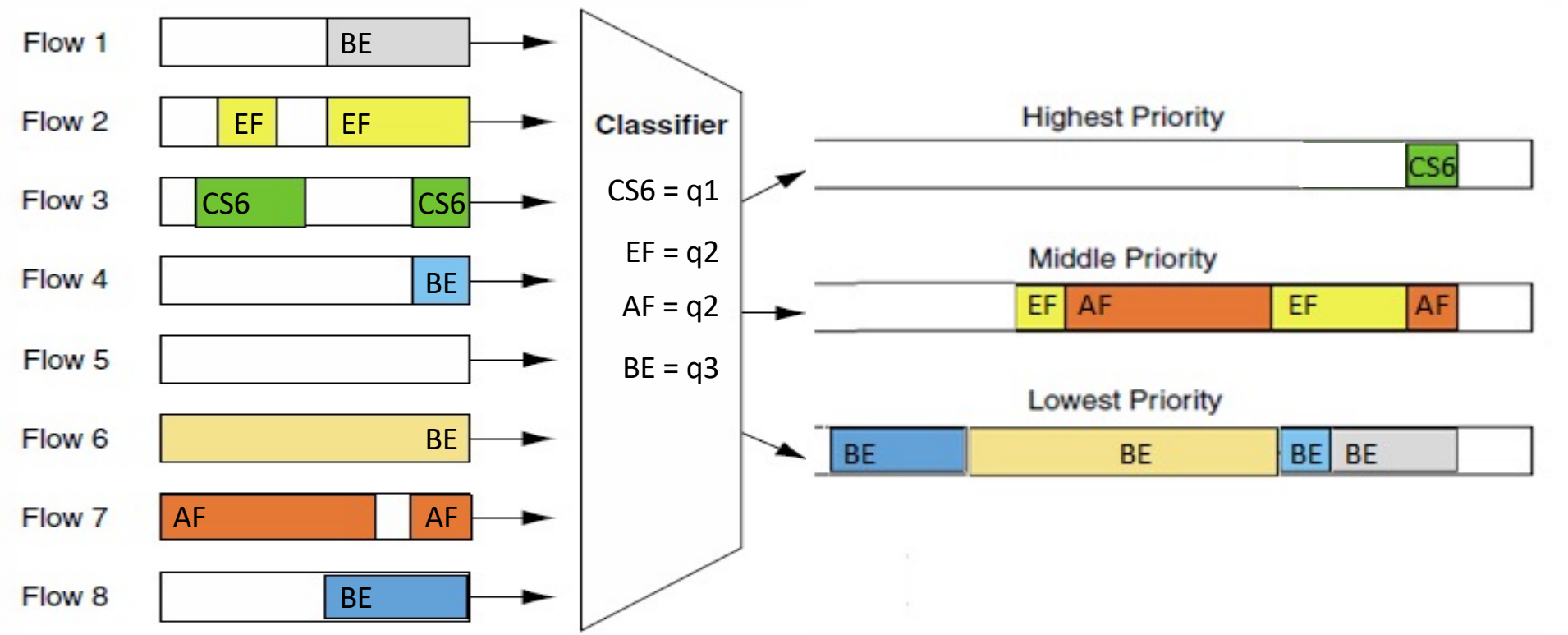
QoS – Differentiated Services (DiffServ)



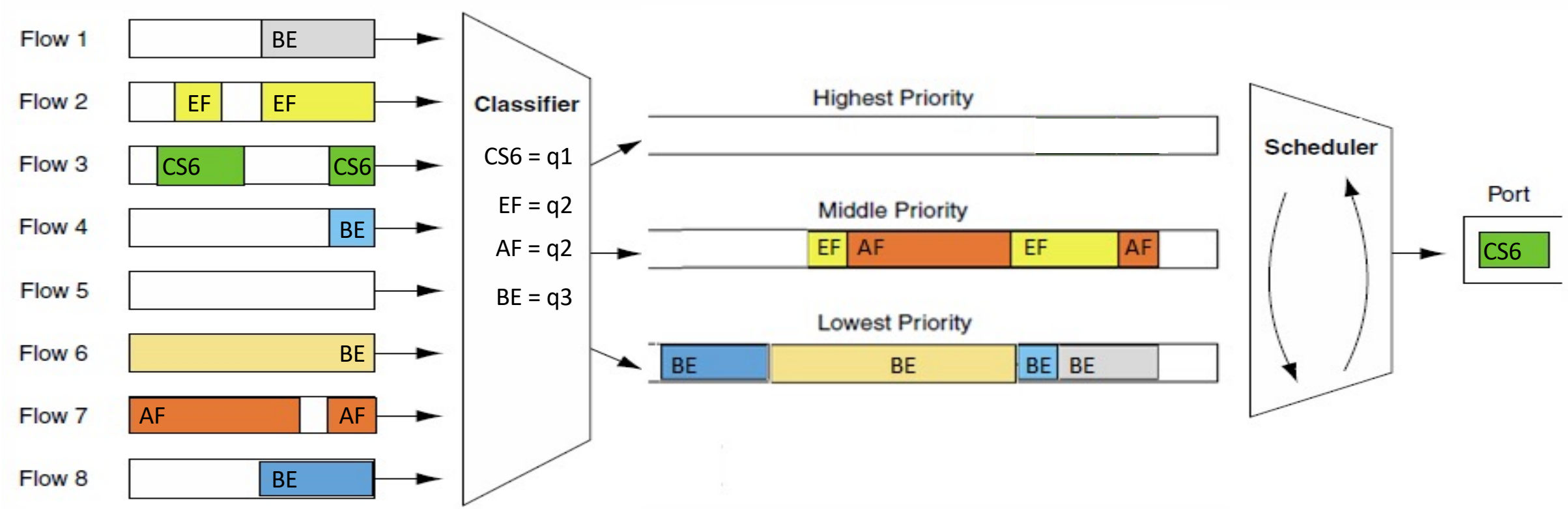
QoS – Differentiated Services (DiffServ)

- Defined in RFC 2474
- Defines up to 64 traffic classes (i.e. EF, AFx, CSx, BE etc.)
- Packets are tagged with DSCP value (0 – 63)
- Switches store packets in different priority queues (requires proper configuration)
- Egress scheduler forwards packets from higher prioritized queues first (strict priority / weighted round robin / guaranteed minimum bandwidth …)

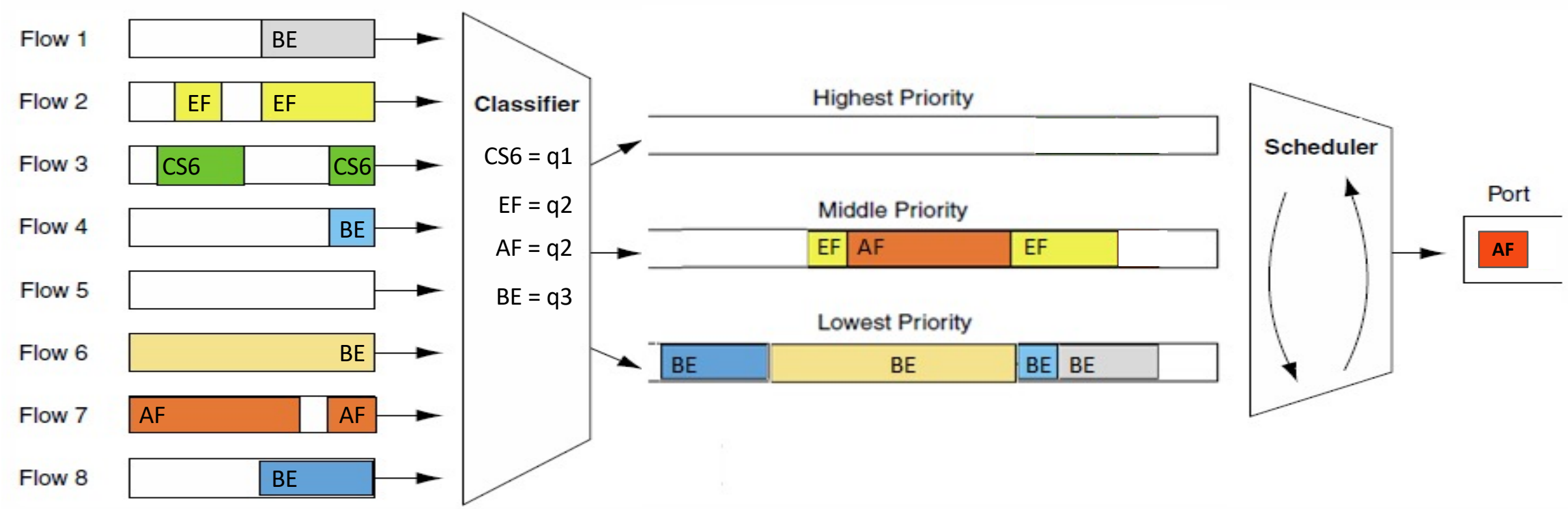
QoS – Differentiated Services (DiffServ)



QoS – Differentiated Services (DiffServ)



QoS – Differentiated Services (DiffServ)

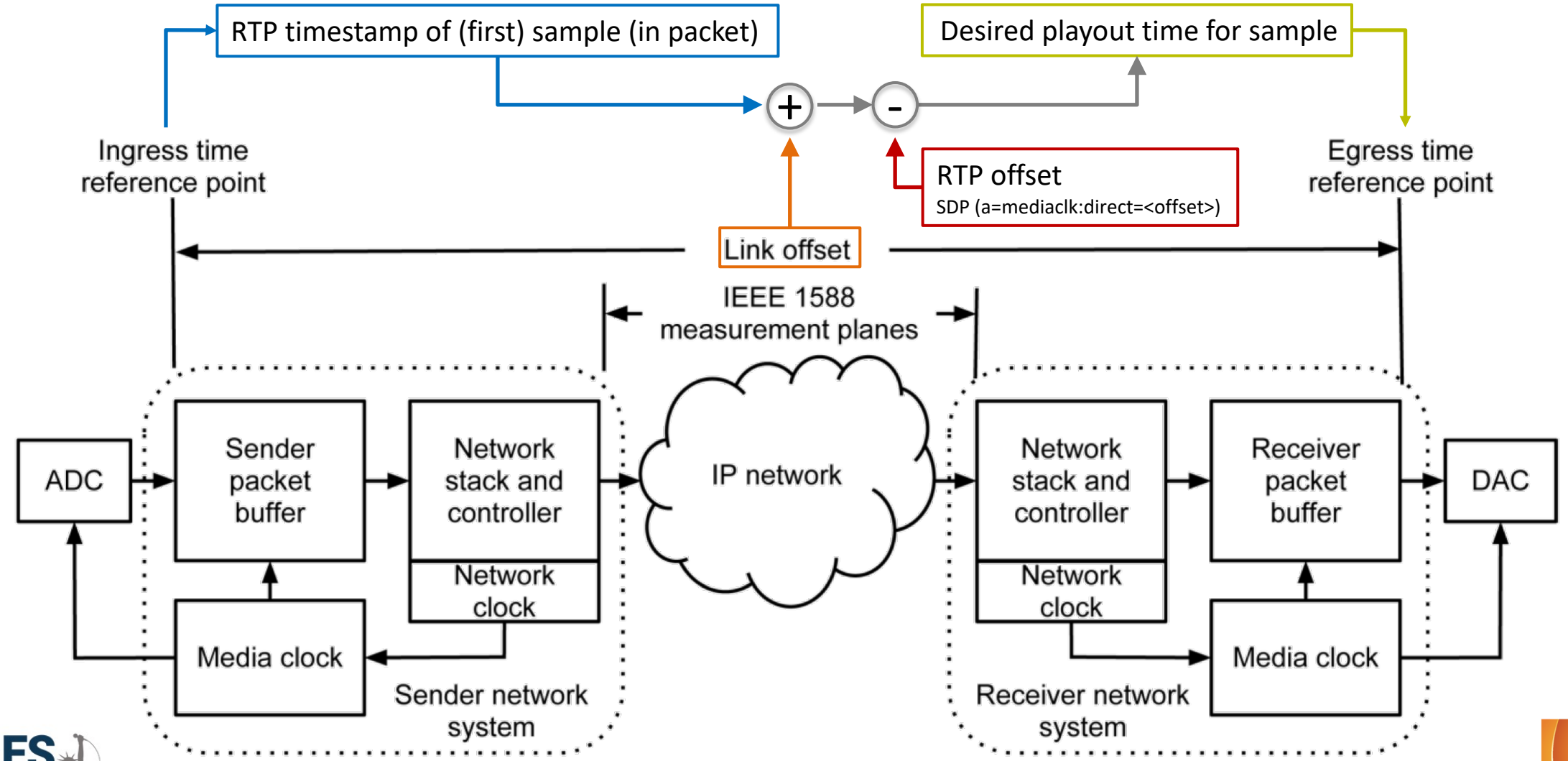


QoS – Differentiated Services (DiffServ)

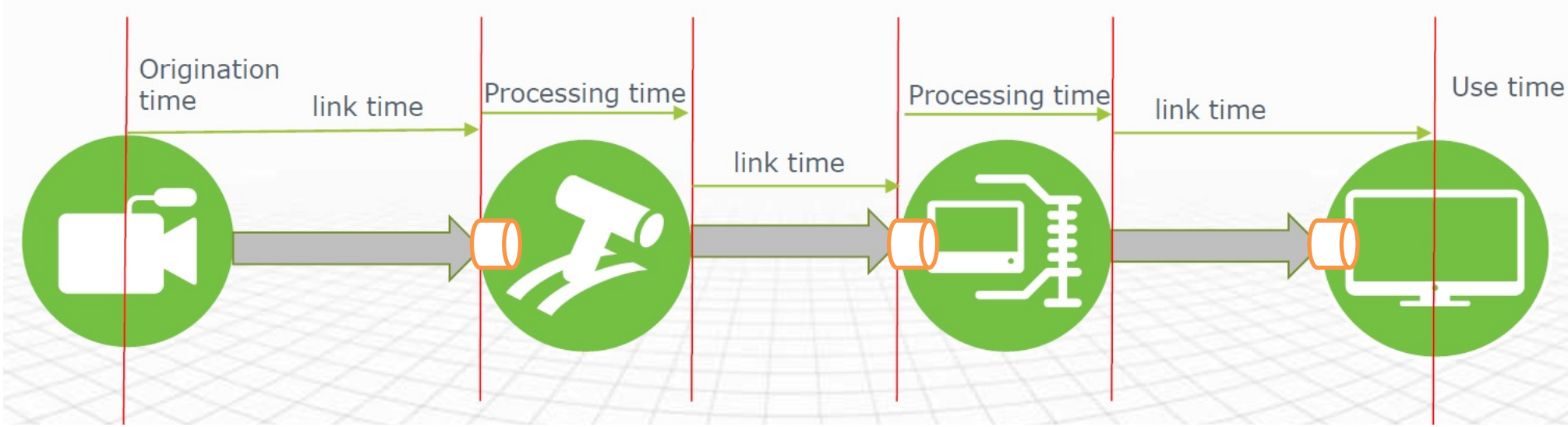
QoS – Differentiated Services (DiffServ)

- Defined in RFC 2474
- Defines up to 64 traffic classes (i.e. EF, AFx, CSx, BE etc.)
- Packets are tagged with DSCP value (0 – 63)
- Switches store packets in different priority queues (requires proper configuration)
- Egress scheduler forwards packets from higher prioritized queues first (strict priority, weighted round robin, guaranteed minimum bandwidth)
- Needs to be supported along full path from the transmitting to the receiving end
- No admission control → congestion / packet dropping possible when bandwidth is exceeded

AES67 synchronization - link offset (latency)



Production Workflow Timing



Buffer

Image courtesy of Andy Rayner (Nevion)