INTRODUCTION TO ETHERNET & TCP/IP NETWORKING

GEOFF WATERS
DIGITAL NETWORKING SYSTEMS ENGINEERING

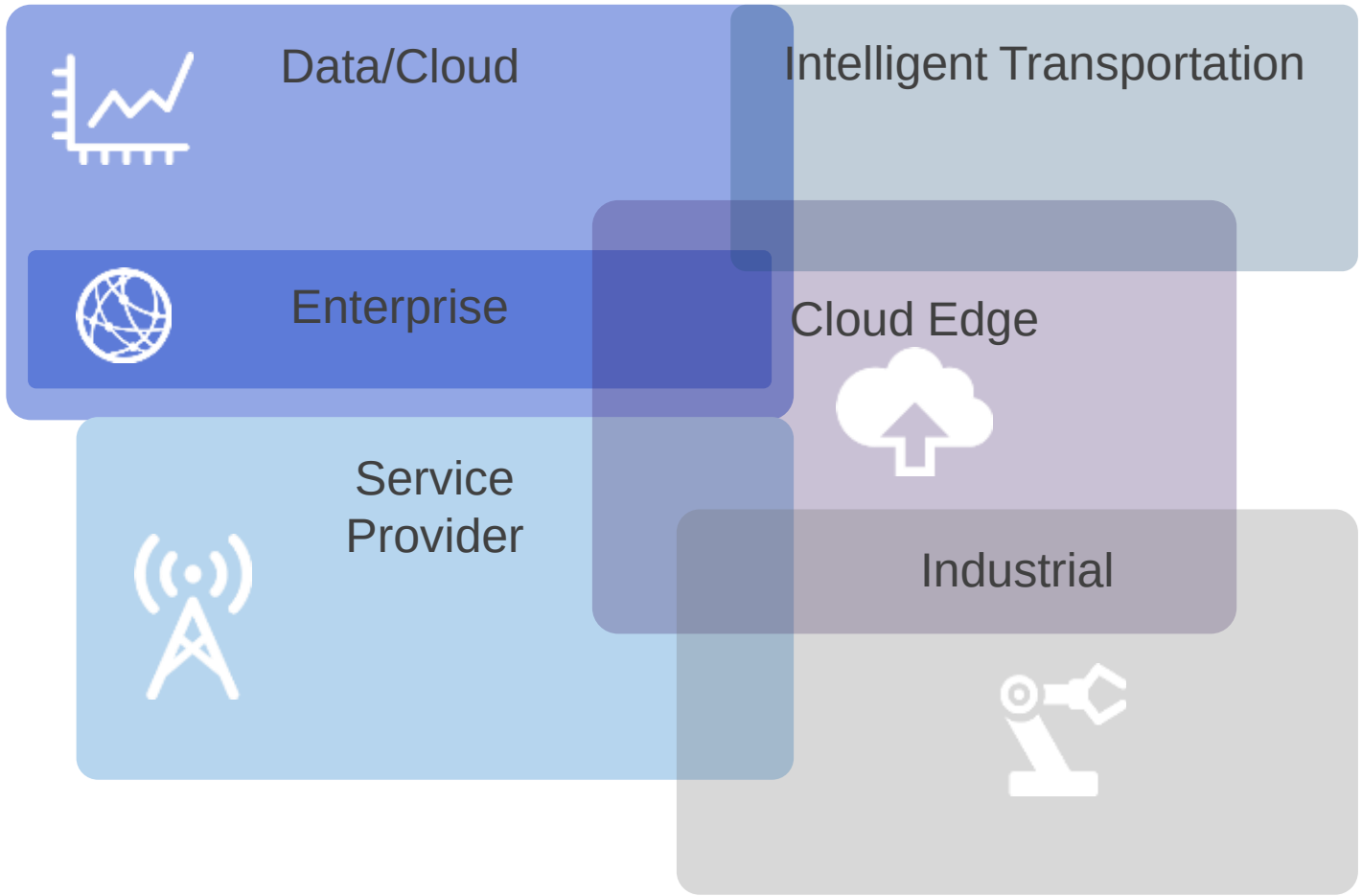
AMF-AUT-T2816 | AUGUST 2017

SECURE CONNECTIONS FOR A SMARTER WORLD

Abstract

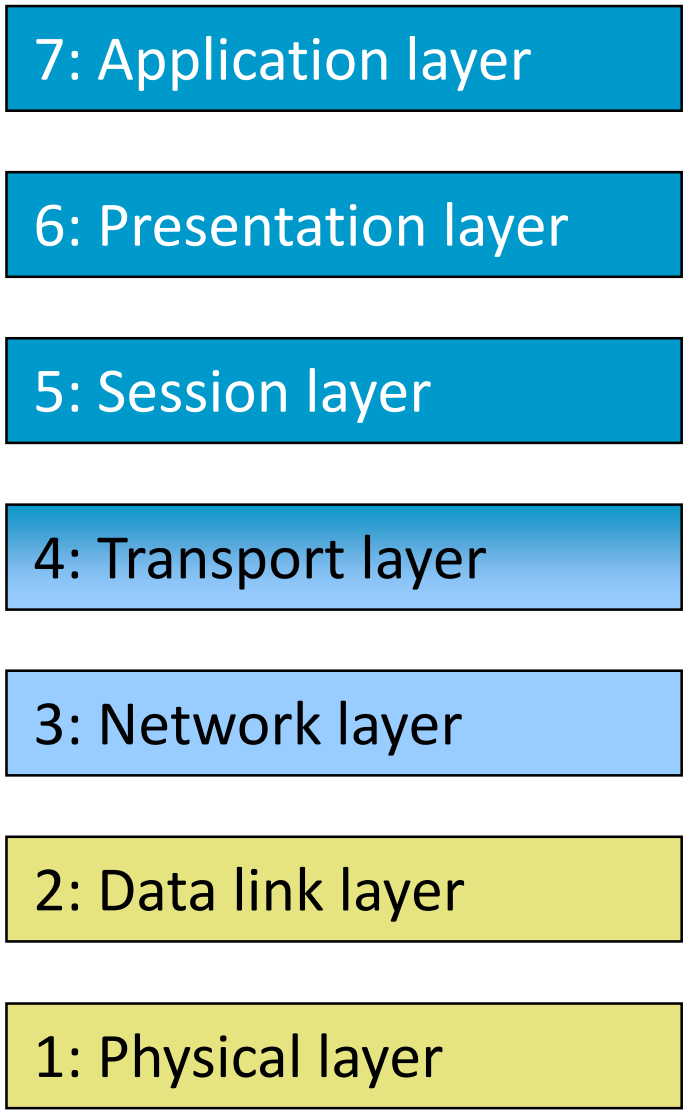
Planning to do any switching or routing? This course provides an introduction to the principles of Ethernet and TCP/IP networking for non-IT applications. Topics to be reviewed include frame & packet formats, the purpose of various header & trailer fields, security extensions, and reliability mechanisms. New developments in ‘soft’ switching will also be covered.

NXP DN (Digital Networking) Business



Digital Networking high performance networking and computing processors offer server class performance for real time control and high touch data services in wireless & wireline infrastructure

The OSI Model of Networking



First introduced in the late 1970s…

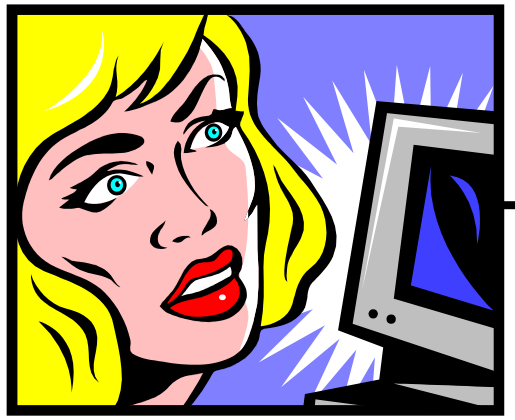
The OSI Model of Networking

7: Application layer	Ex. Apache Web Server
6: Presentation layer	
5: Session layer	
4: Transport layer	TCP, UDP, SCTP
3: Network layer	IP (Internet Protocol), IPSec
2: Data link layer	Ethernet MAC
1: Physical layer	Ethernet PHY

Services Provided By Layers (But Not All By All!)

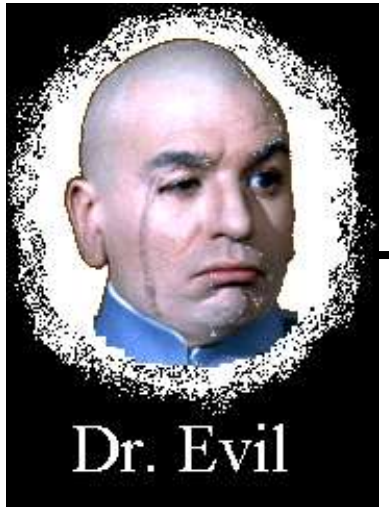
- Analog to Digital Conversion
- Error Detection and Correction
- Auto Discovery and Negotiation
- Source and Destination Addressing
- Access/Admission Control
- Reliable Delivery including Retransmission
- Flow Control, Congestion Management, and Prioritization
- Security

Security Basics – Firewall/Access Control List



Alice uses her identity to access Bob's Bank.

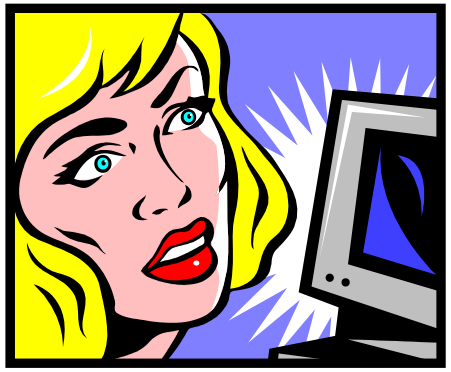
Dr. Evil uses his identity to access Bob's Bank.



Bob the banker can either block access to specific Evil parties (Blacklist) or approve access to specific Good parties (Whitelist).

The access list can be static (rarely is ever updated) or dynamic (continually updated).

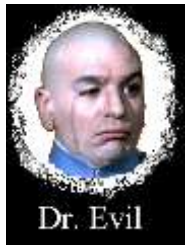
Security Basics – Data Integrity



Transfer $1000 from acct# 1234567
to acct# 7654321 (MAC=3127)

Transfer $100,000 from acct# 1234567
to acct# 10101010 (MAC=3127)

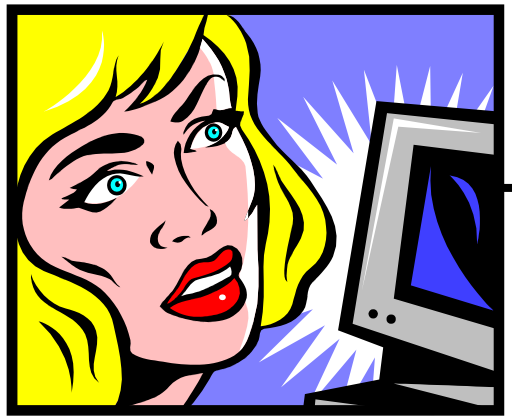
Alice sends her request along with a Message Authentication Code derived from both the message and a secret key



Attacker can't generate the proper MAC for the altered message because he doesn't know the secret key.

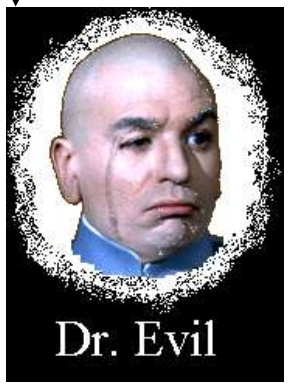
Bob the banker recalculates the Message Authentication Code over the message and compares to the received MAC. The data manipulation is detected, and the fraudulent transaction is not processed.

Security Basics – Anti-Replay



Request #1: Transfer $1000 from acct# 1234567 to acct# 10101010 (MAC=7576)

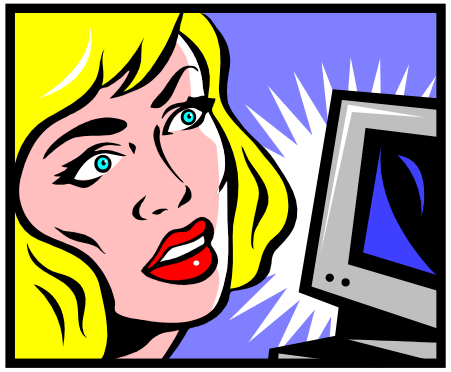
Request #1: Transfer $1000 from acct# 1234567 to acct# 10101010 (MAC=7576)



Alice sends a single legitimate transfer request to her bank, including a Sequence Number, along with a Message Authentication Code derived from both the message and a secret key

Bob the banker keeps a record of the sequence numbers sent by Alice. If he sees duplicate sequence numbers, he discards the replayed transactions. If Dr. Evil tries to replay a transaction with a different sequence number, the Message Authentication Code will fail because he doesn't have the key required to create the MAC.

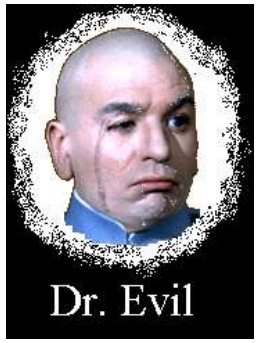
Security Basics - Encryption



Transmission inside encrypted tunnel, combining message with secret key

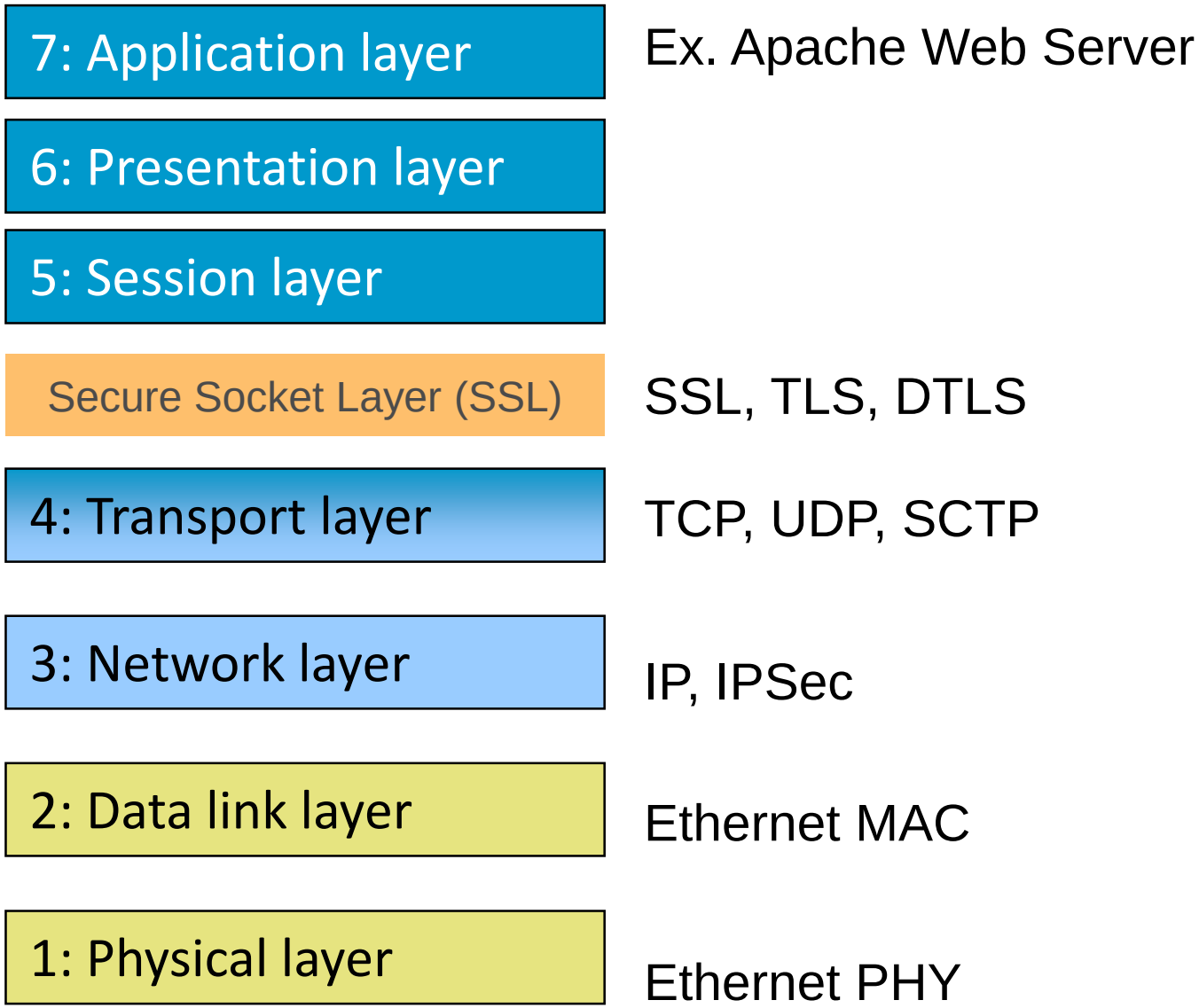
Transfer $1000 from acct# 1234567 to acct# 7654321

Qw;o9yf4 0hnm4mc, niofen *#$ hn23n()* 1n

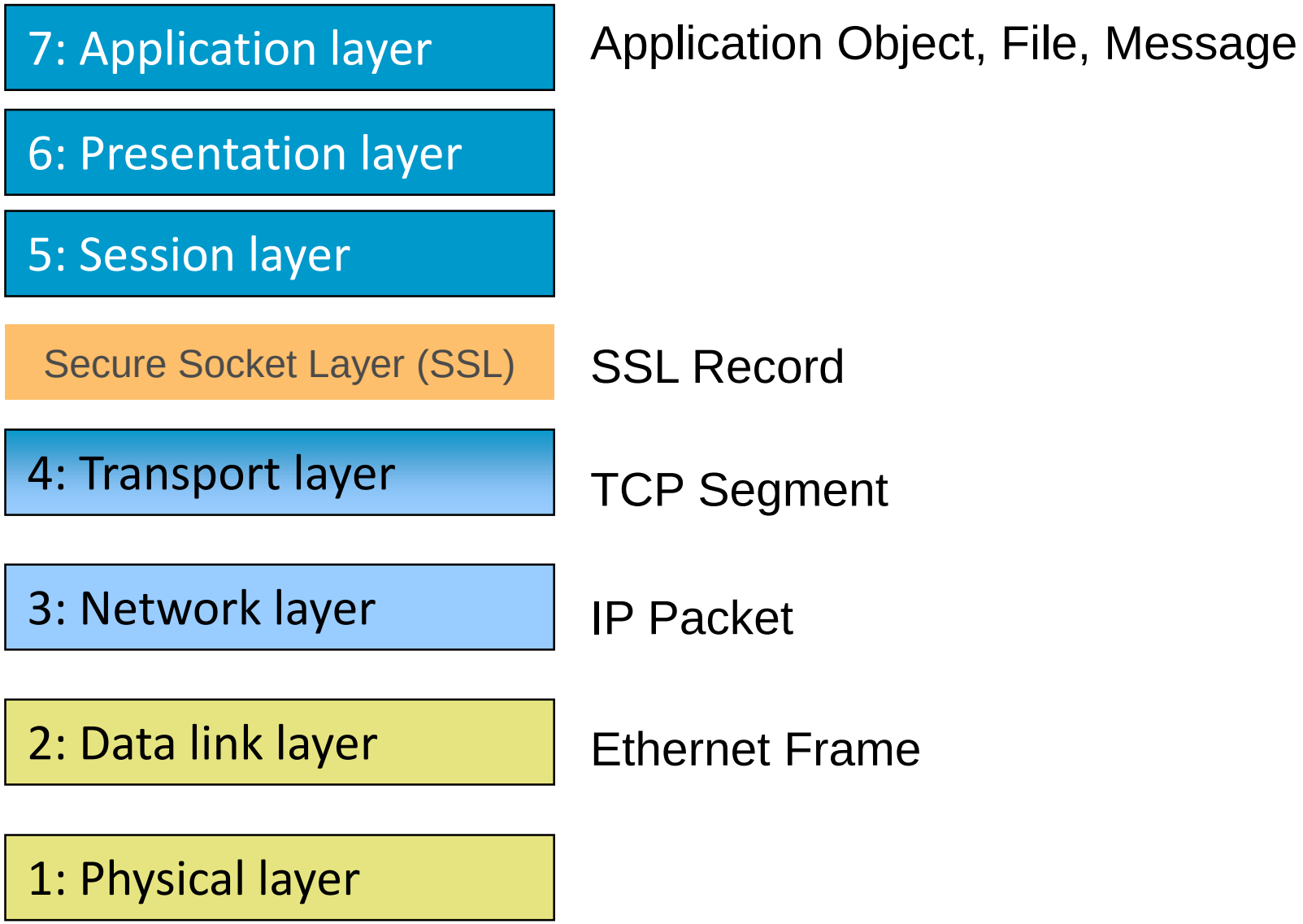


Attacker doesn’t know secret key, can’t eavesdrop

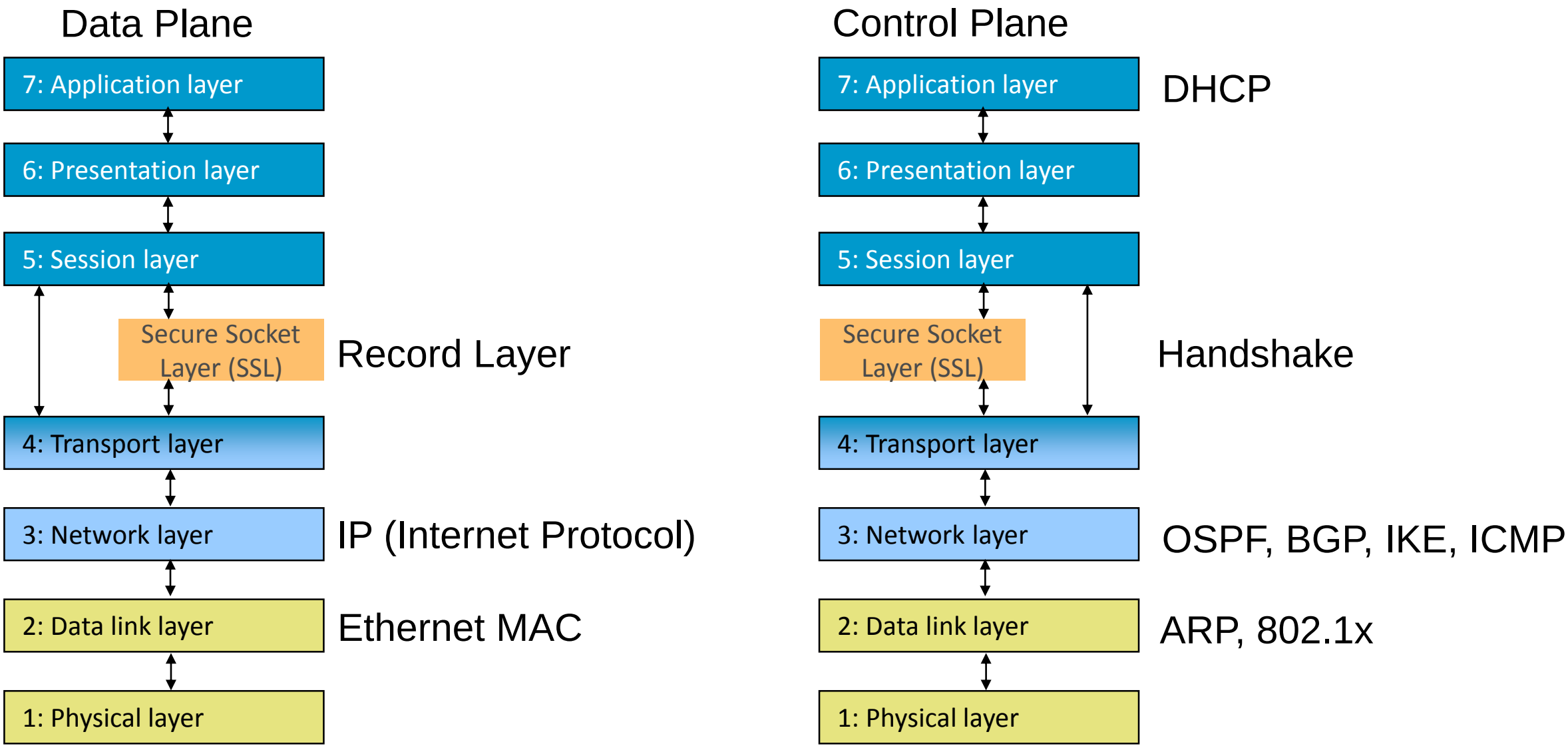
According to IETF



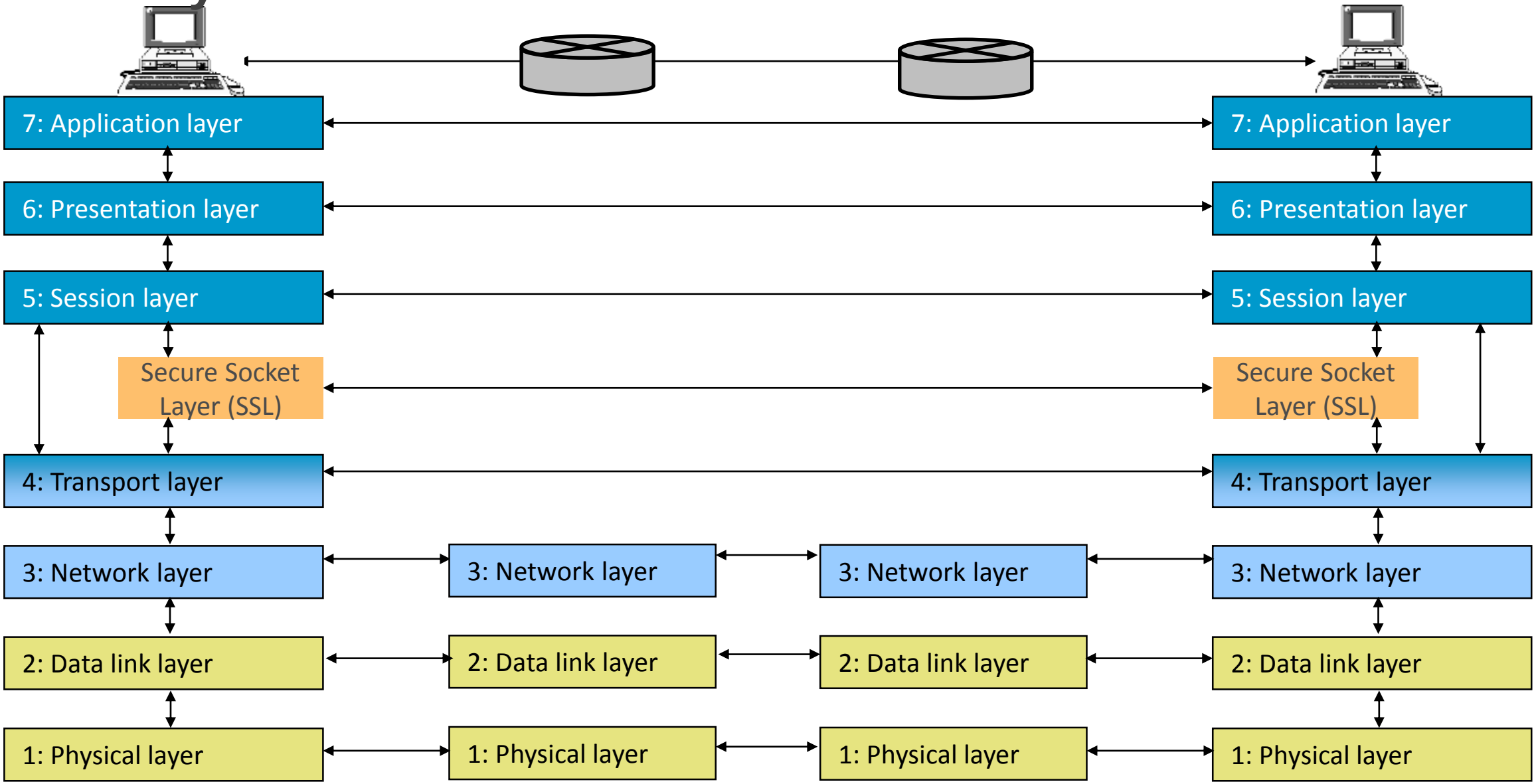
They aren’t all ‘packets’…



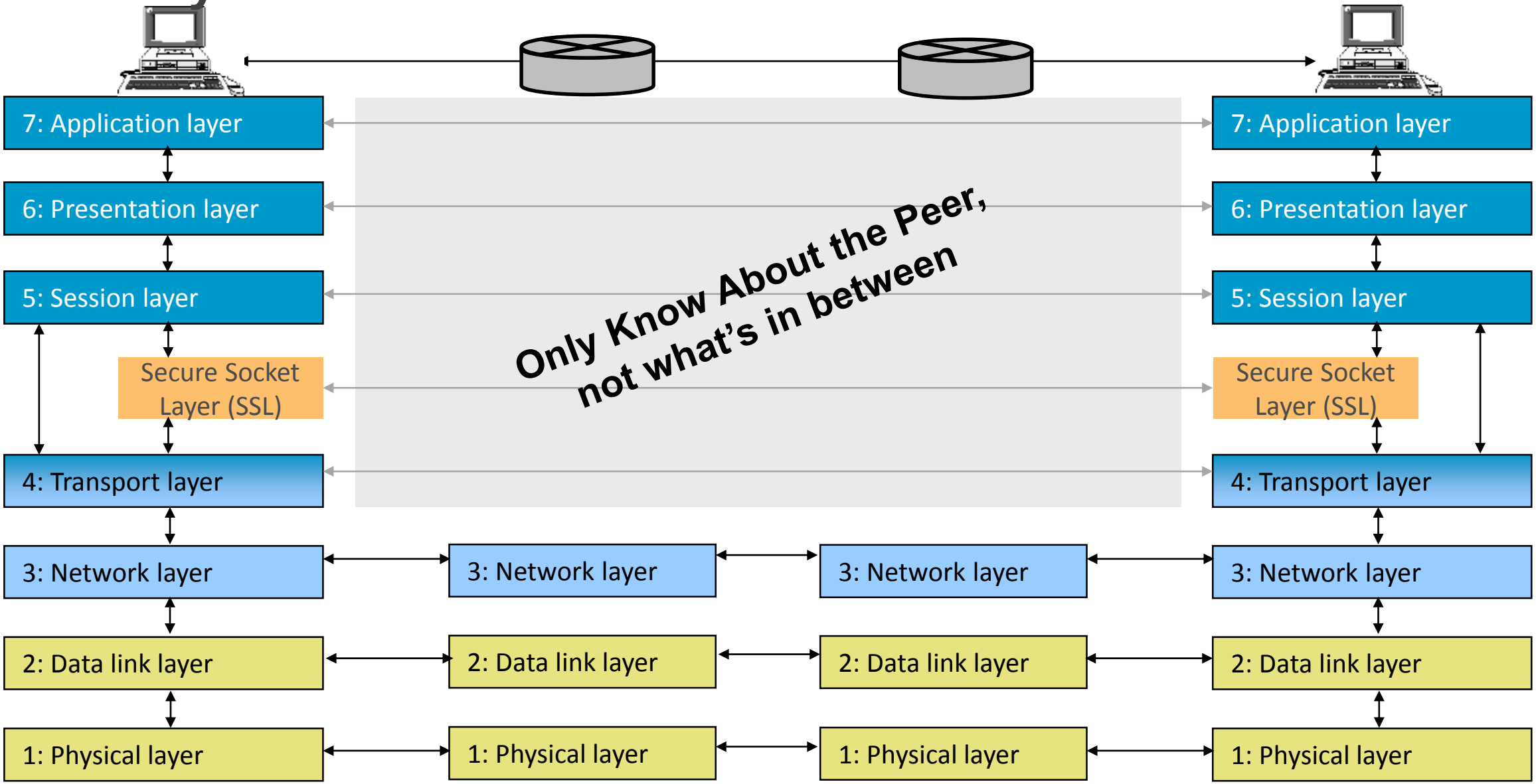
Control Yourselves



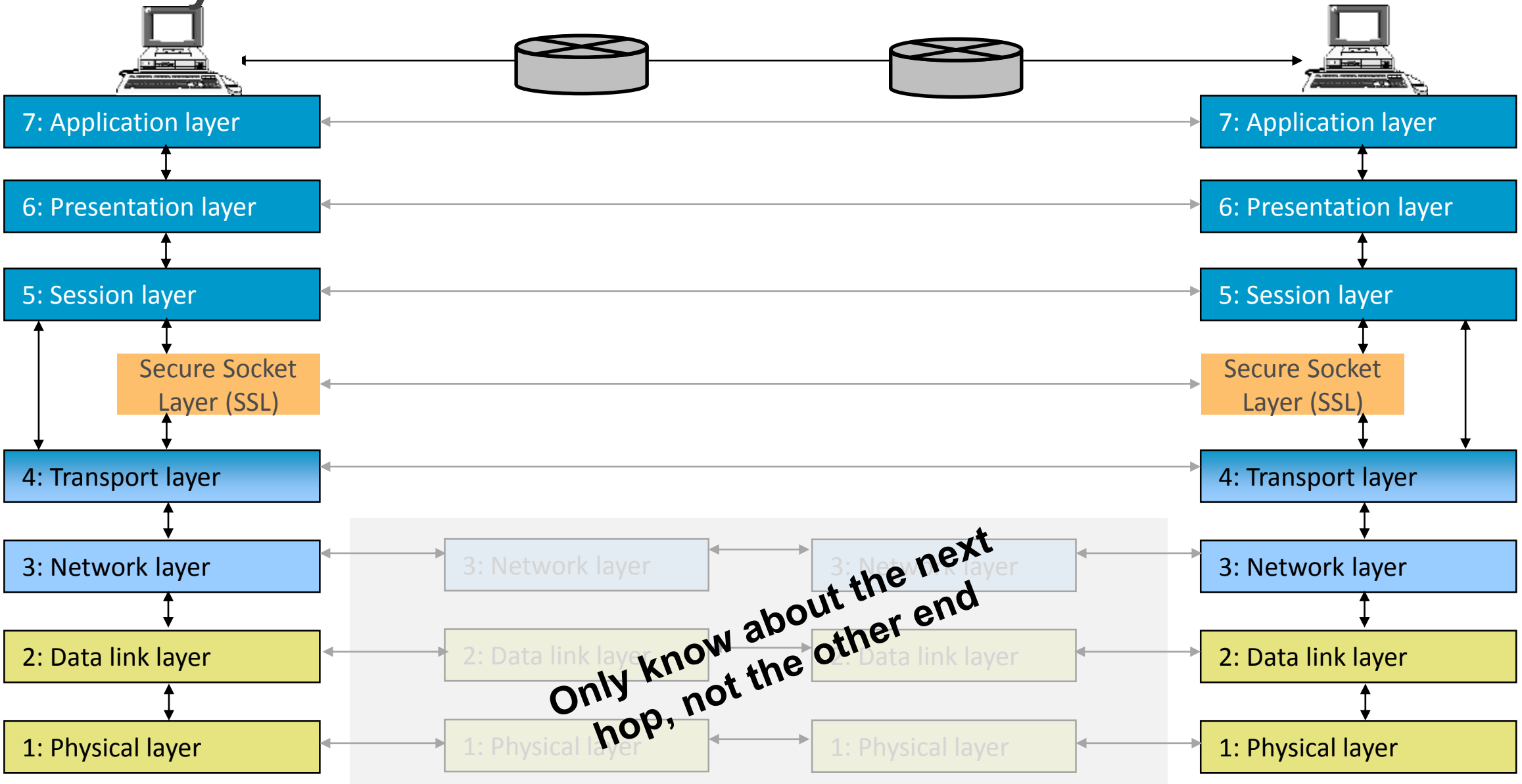
The Layers in Action



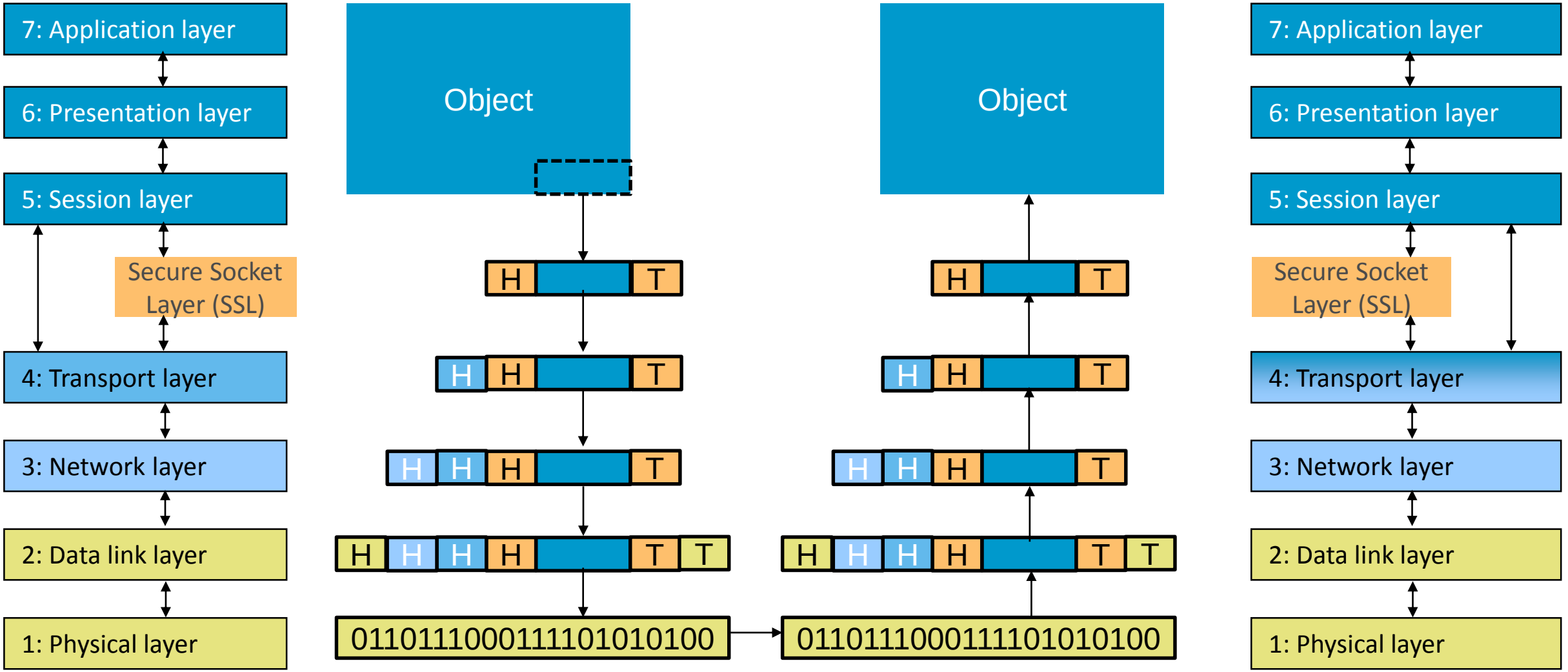
The Layers in Action



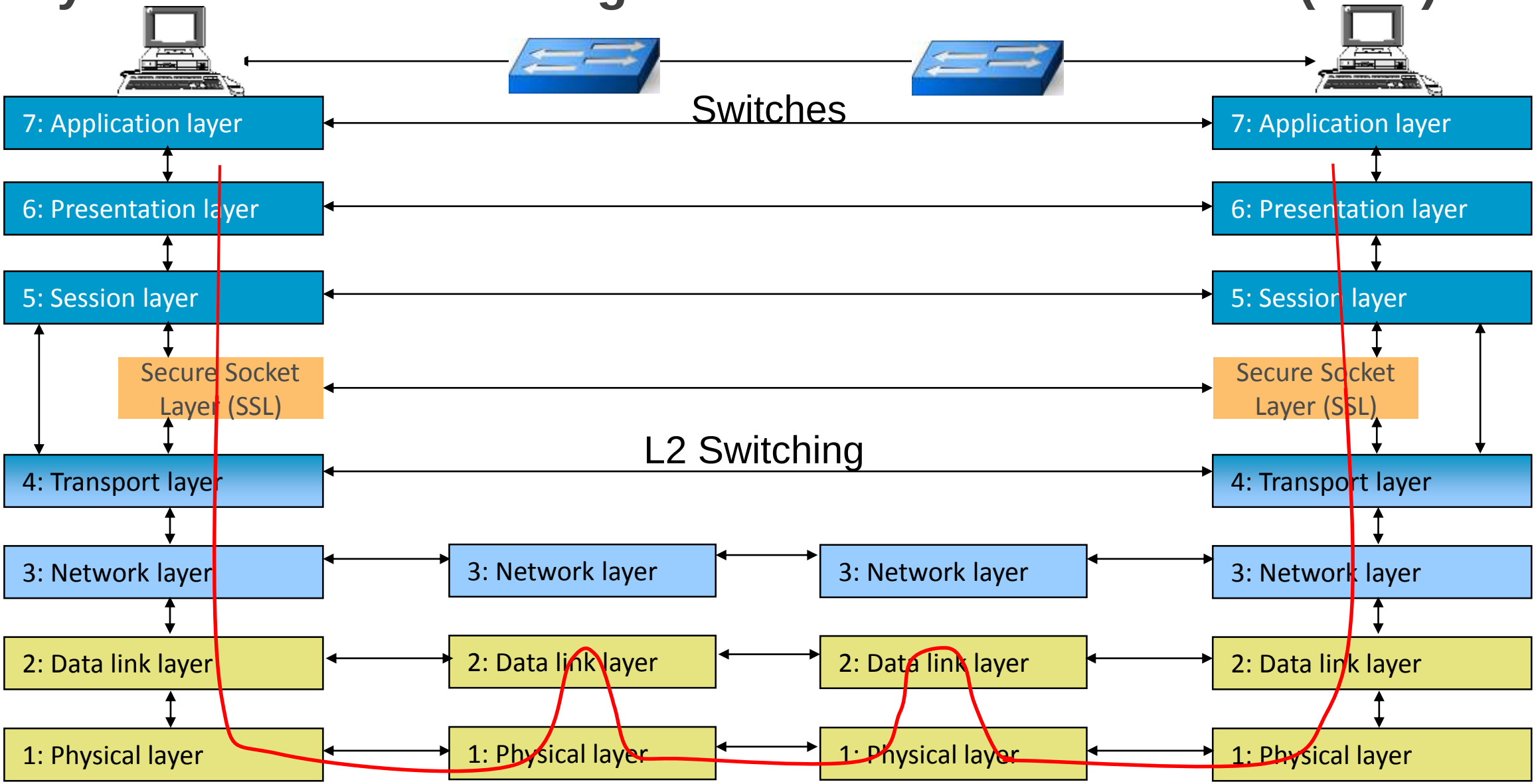
The Layers in Action



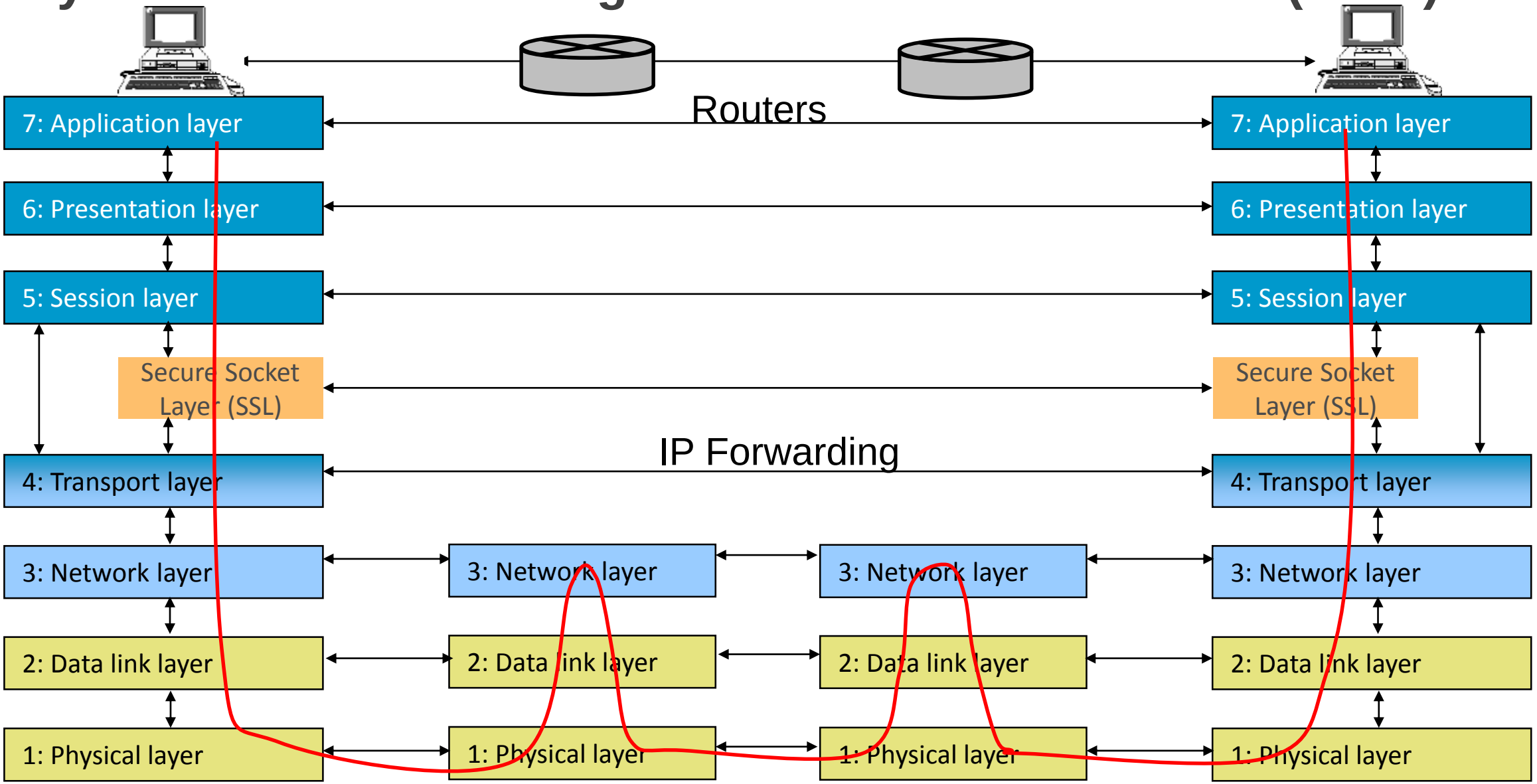
Payloads, Headers, & Trailers



Systems communicating over a Local Area Network (LAN)



Systems communicating over a Wide Area Network (WAN)



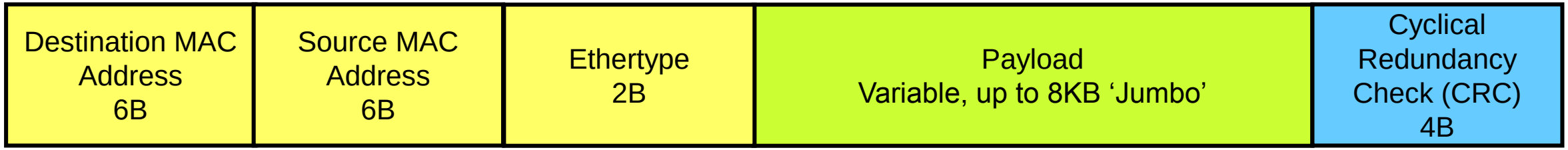
IP Forwarding vs Routing

- Sending packets to the next router isn't routing, it's forwarding. Or more precisely, IP Forwarding.

- Routing is the act of determining the network path from IP Source Address to IP Destination Address. Routing is performed by L3 control protocols such as OSPF (Open Shortest Path First) and BGP (Border Gateway Protocol).

Layers 1 & 2
Ethernet

Basic Ethernet



Next Hop Address

Current Hop Address

Ethernet 'Flavor' and Priority (Class of Service)

Polynomial based checksum for error detection

Analog to Digital Conversion	Layer 1 Ethernet PHY
Auto Discovery and Negotiation	Layer 1 Ethernet PHY
Source and Destination Addressing	Layer 2 Ethernet Header, changes each hop. Broadcast/Multicase supported.
Error Detection and Correction	Layer 2 Ethernet CRC; detection only
Access/Admission Control	Virtual LAN, 802.1x Network Admission Control; Control Plane
Reliable Delivery including Retransmission	Not supported
Flow Control, Congestion Mgt, and Prioritization	Pause Frames, Data Center Bridging, Time Sensitive Networking
Security	MACSEC

Ethernet Physical Layer

- PHYs can be integrated with the MAC, or implemented as a separate chip.
- The PHY determines the speed and the range of the Ethernet link.
- IEEE standardizes the interface between a MAC and its external PHY, and the PHY's interface to its peer PHY.

PHYs may support;

- Transmission distances of inches to miles.
- Low to high cost transmission media.
- Low (100Mbps) to high (100Gbps) data rates.

PHY interfaces;

- Parallel (MII, GMII, RGMII), Serial (SGMII, XFI).
- Single to many (QSGMII).

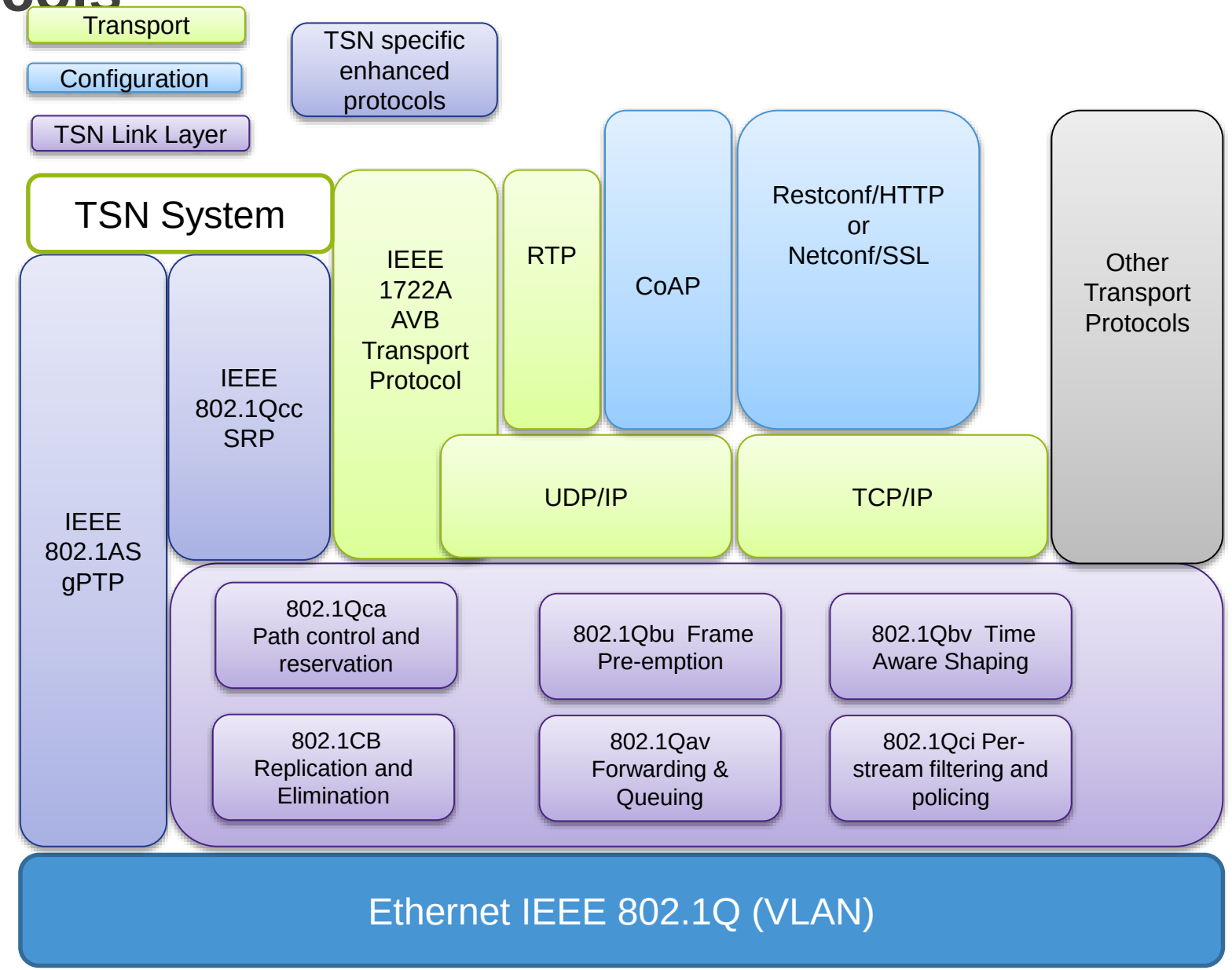
BroadReach™ refers to both the MAC to PHY and PHY to PHY interfaces.

Advanced Ethernet Options

Physical LANs are generally shared infrastructure. Virtual LAN – Adds a 'Tag' to the header identifying the 'virtual' LAN the frame is allowed to be switched within. Combined with 802.1x Network Admission Control, and you can assign guests on your network to 1 VLAN, while internal users are on another. Or separate the engineering network from HR. Or Infotainment from Vehicle Control.

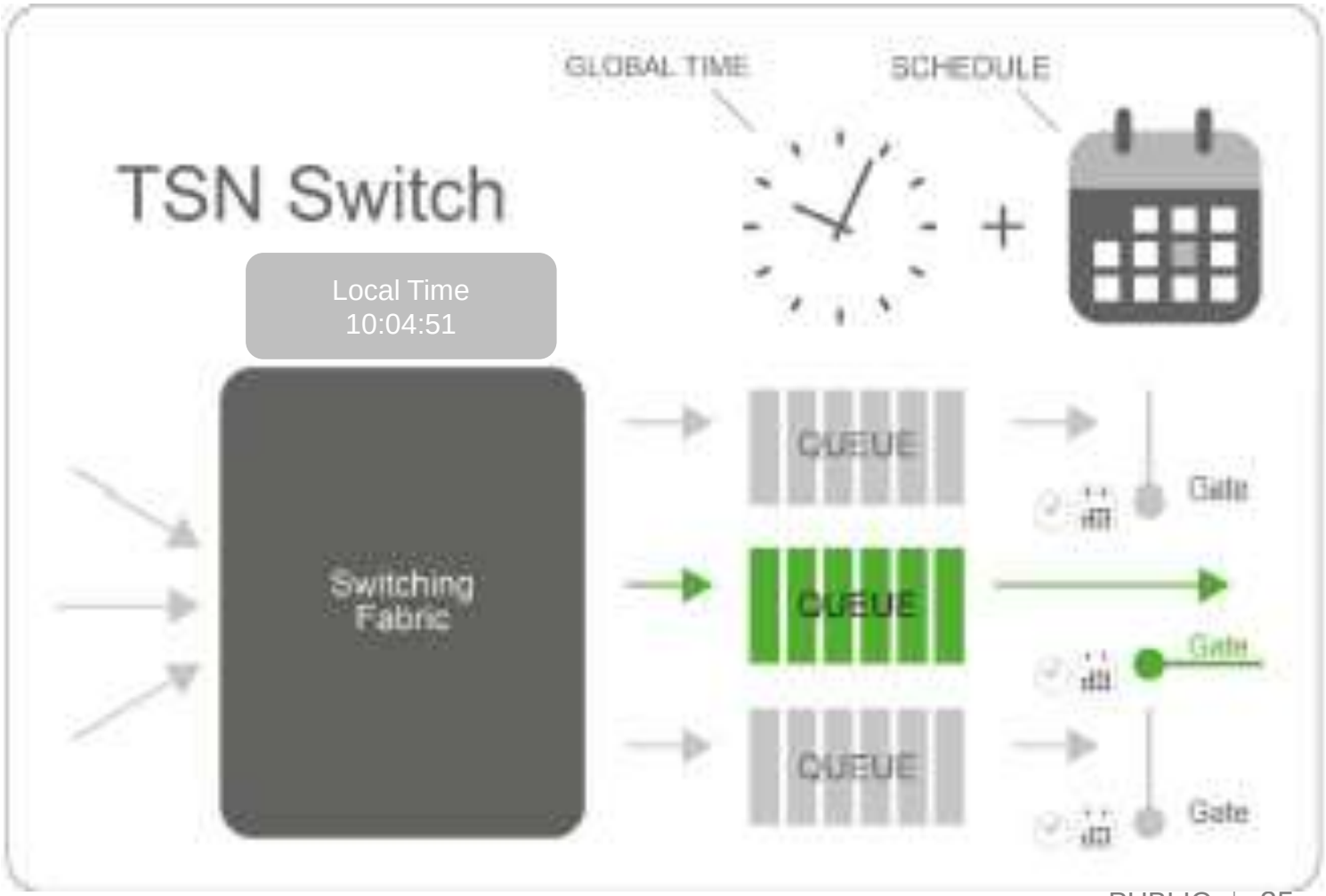
Virtual LANs rely on tag based access control. The frames are still subject to snooping, injection, or manipulation. MACSEC provides for encryption of the frame payload, anti-replay, and strong cryptographic integrity check value (ICV).

TSN Protocols



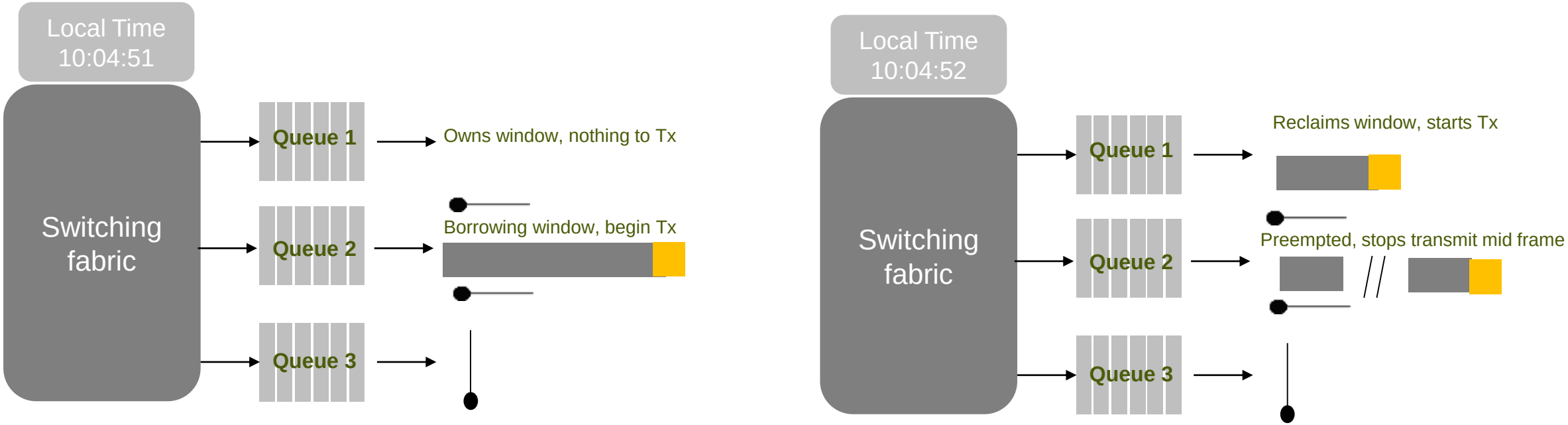
802.1Qbv: TAS – Time Aware Shaping

- Time Triggered transmissions from each of several queues.
- Queue selection based on priority field of VLAN tag



802.1Qbu: Frame Pre-emption

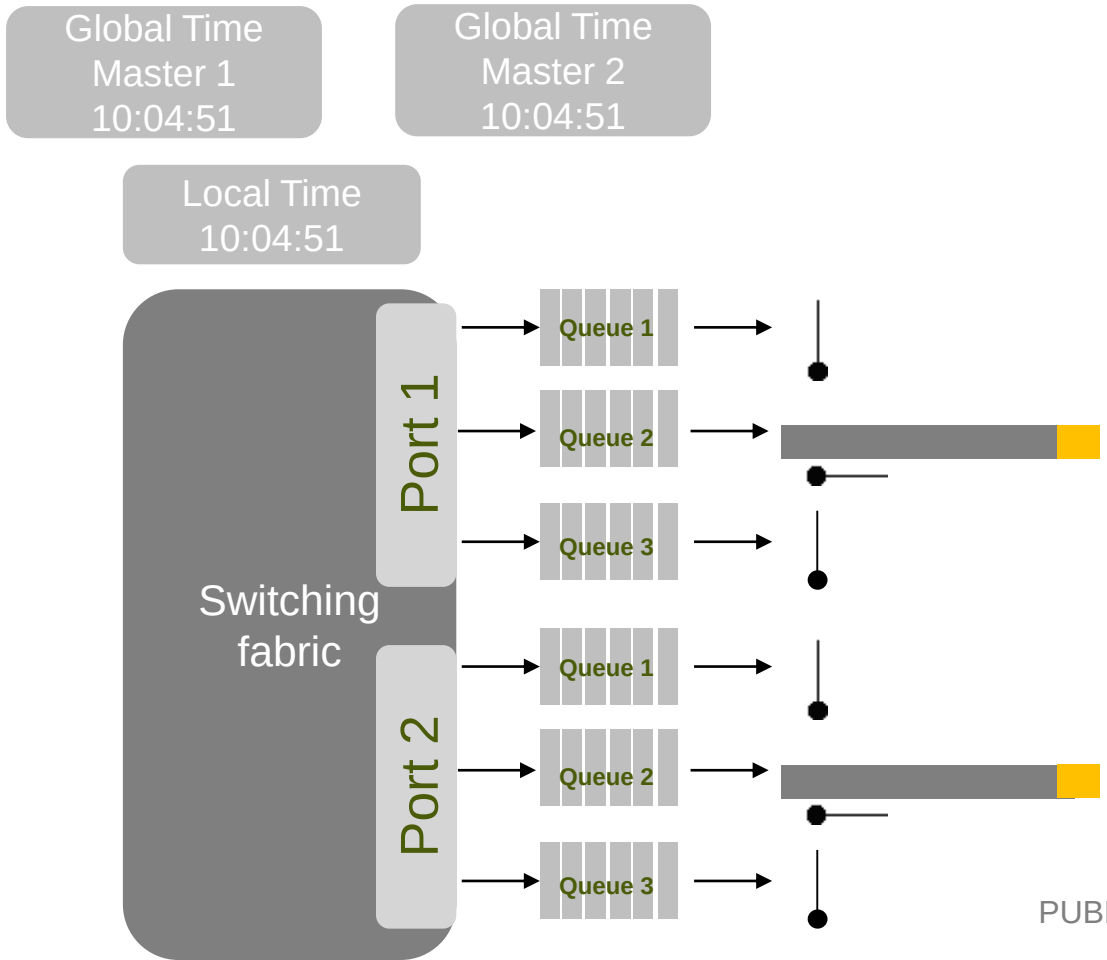
- Pure Time Aware Shaping only opens 1 queue gate at a time. With frame pre-emption, if the queue whose turn it is to transmit (ie, Queue 1) doesn't have any data, another queue (ie, Queue 2) can start transmitting. However if data suddenly appears in Queue 1, Queue 1 can pre-empt Queue 2. Queue 2's frame transmission is terminated instantly, regardless of frame boundary, and Q1 begins transmission.



Qbu deals with how queues determine when they can borrow someone else's window, how the owner takes the window back, and how to stop and later complete the frame that was pre-empted.

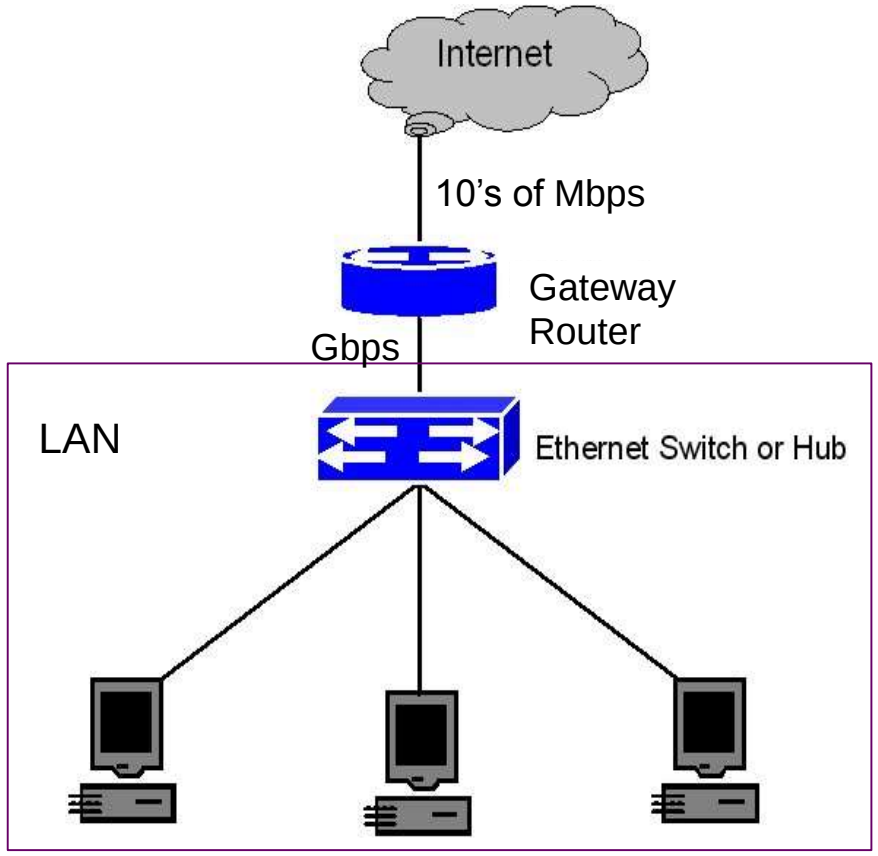
802.1CB: Redundancy (frame replication and elimination)

- From a management perspective, defines how to configure redundant transmissions of frames (presumably from different ports on a switch) to ensure that the target receives at least one copy, even if one transmission path is disrupted. Similar to HSP and PRP.
- Works with 802.1CA: Path Control & Reservation.
- On the datapath side, defines how to handle it when both paths are functioning, and you get 2 copies of the frame.



Layer 3
Internet Protocol

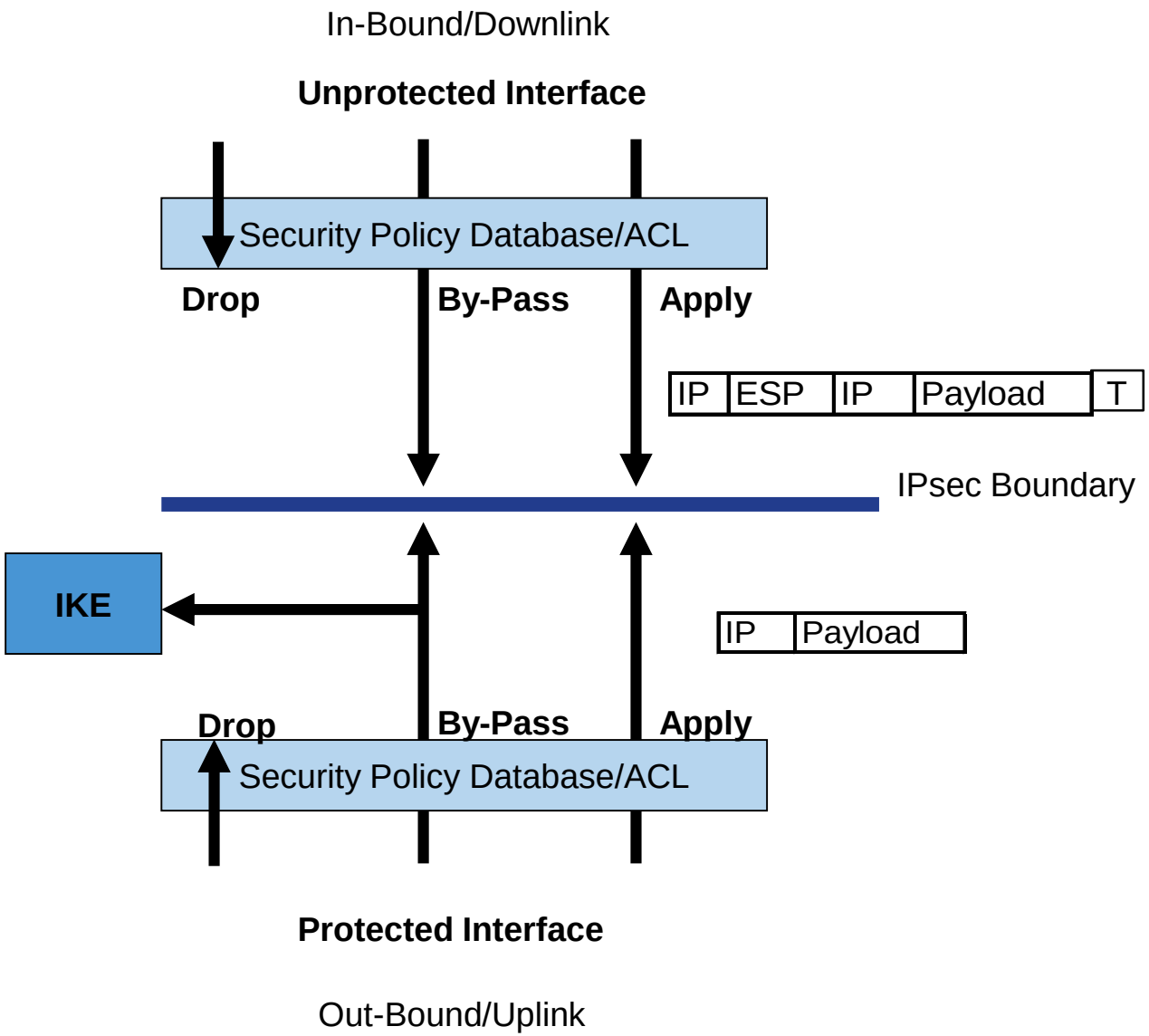
Forwarding isn't the issue



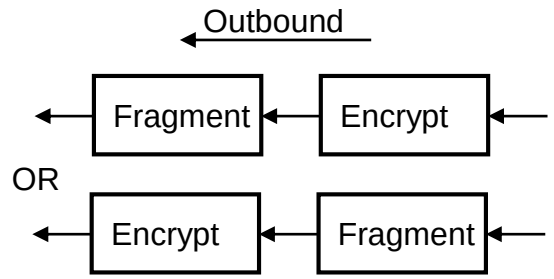
- The network layer is where your data leaves the physical security of your LAN and gets routed through the internet. Consequently, a lot of the interesting processing at the network layer is security related.
- WAN interfaces are often 1-2 orders of magnitude lower bandwidth than the WAN, which means congestion and prioritization are major concerns.

Analog to Digital Conversion	Not supported
Auto Discovery and Negotiation	Routing control protocols
Source and Destination Addressing	IP Header, only a TTL field changes per hop
Error Detection and Correction	IP Header covered by simple checksum, payload not covered
Access/Admission Control	Firewall rules, IKE
Reliable Delivery including Retransmission	Not supported
Flow Control, Congestion Mgt, and Prioritization	Extensive policing, shaping options
Security	IPsec

IPSec & Firewall



- IPsec and access control list/firewall were originally treated separately, but today most implementations treat IPsec as a special firewall rule.
- IPsec is a family of specifications, covering multiple modes of operation.
- Encapsulating Security Payload (ESP) Tunnel Mode, which makes the original IP packet into the payload of the IPsec packet, is the most commonly used.
- ESP Tunnel grows the packet by ~80B, which can cause the IP packet to exceed the typical Ethernet payload allowance. This leads to fragmentation. It is better to fragment before encapsulation, so that each packet has a security header.



- IPsec RFCs were weak in this area, and different implementations can take different approaches to fragmentation.

IPsec Events and Statistics

IPsec is also weak (compared to IEEE specifications) in specifying how various events should be dealt with. This is partly because different applications have different security concerns, and the appropriate action for one system might be wrong for another.

IPsec events:

- Key Lifetime Expiration (hard limit)
- HMAC failures
- Detection of replayed packet
- Detection of packet outside anti-replay window
- Padding error

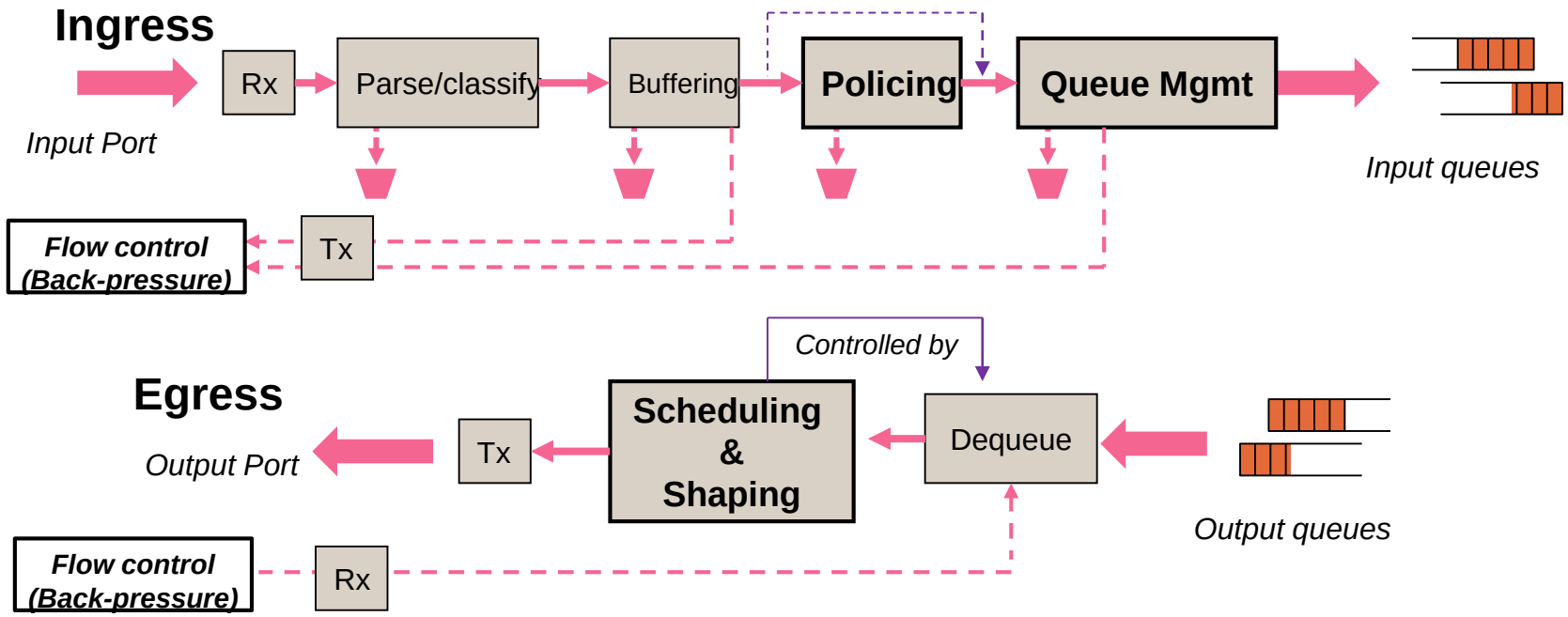
The general reaction to these events is to drop the packet, and in some implementations, drop the session. This can expose the system to denial of service attacks (attacker injects a replayed packet to cause a legitimate session to be dropped). More sophisticated implementations may have event thresholds for killing the session.

Traffic Management

Traffic Management - Intro

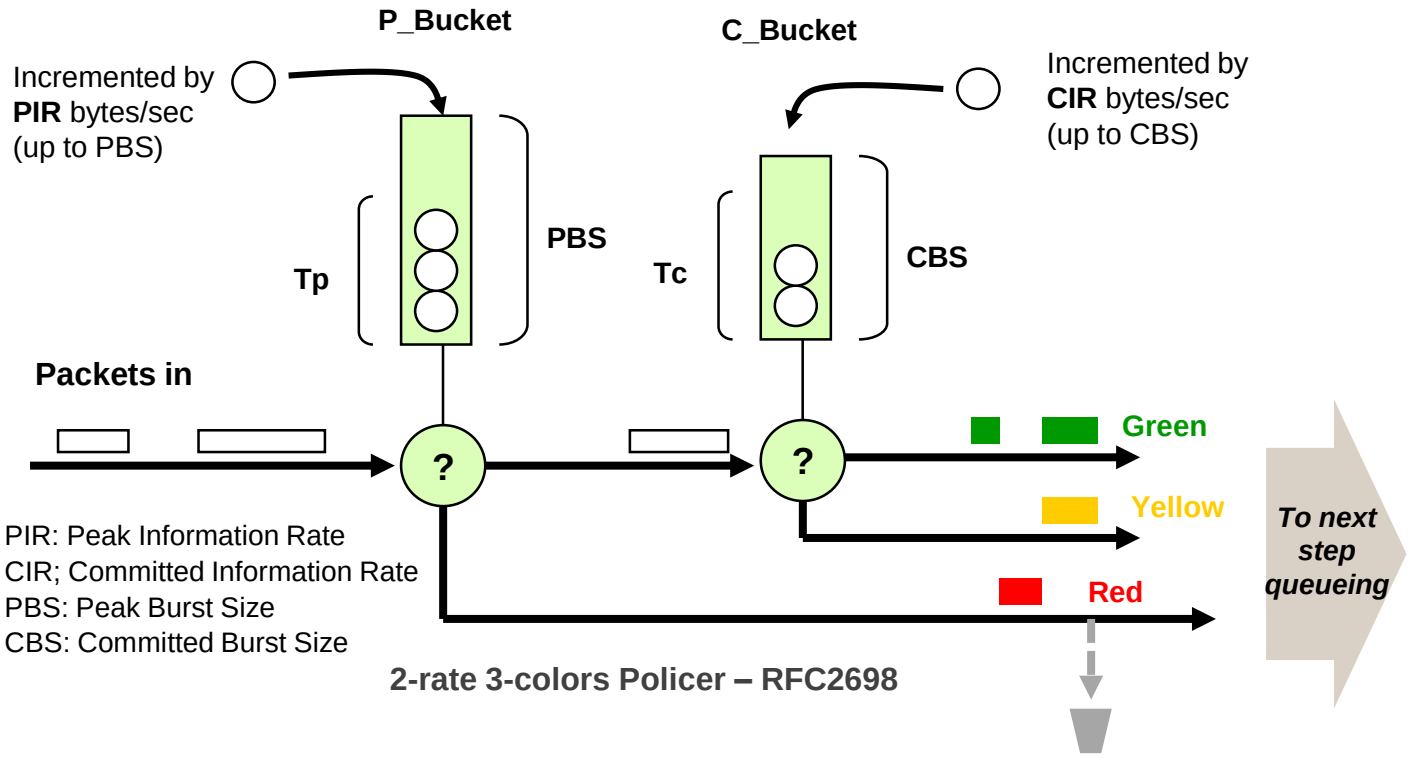
TM purposes

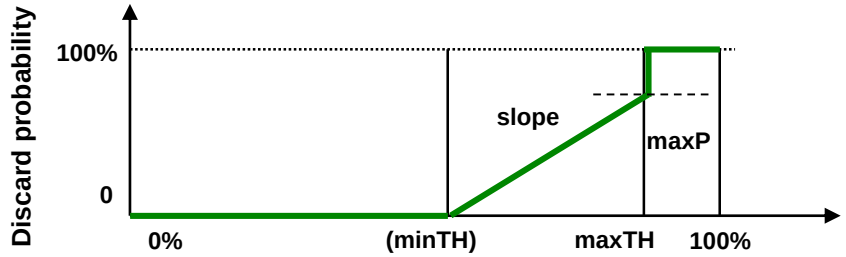
- Manage traffic in accordance with network usage policies/rules
- Manage congestion situations
  - **Congestion management** (Discard packet on Tail Drop threshold, per queue or group of queues)
  - **Congestion avoidance** (RED/WRED Random Early Discard)
- Provide QoS
- Not tied to a network layer, can be applied at any queue



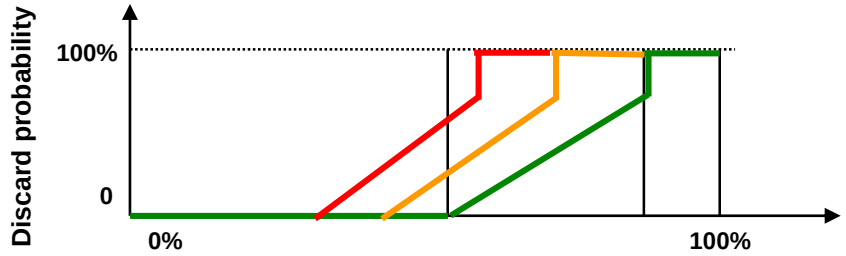
Congestion Avoidance – RED principle

RED (Random Early Discard) primarily for managing congestion in a way appropriate to TCP end-to-end flow control





(simple) RED



W(eighted)RED

Apply different threshold + slope, depending on how each packet is pre-marked /colored through classification or policing

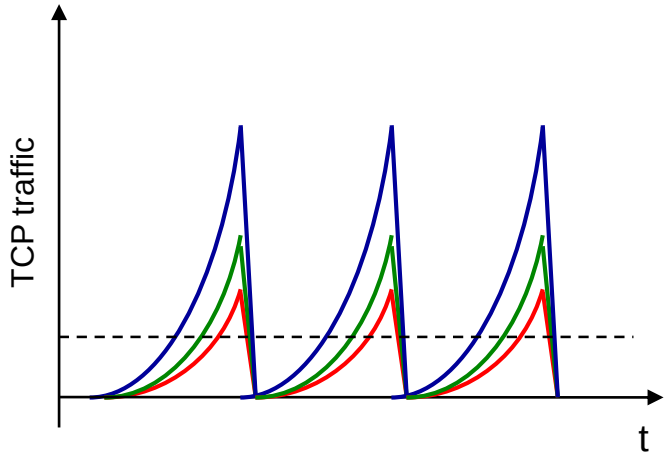
Principle:
- Dual token bucket meters a packet stream according to two rates, a Peak Rate and a Committed Rate and their associated Burst size.
- Mark as Red if exceeds PIR, otherwise mark as Yellow or Green depending on whether it exceeds or not CIR.

Congestion Avoidance – RED Use Case

RED (Random Early Discard) primarily for managing congestion in a way appropriate to TCP end-to-end flow control

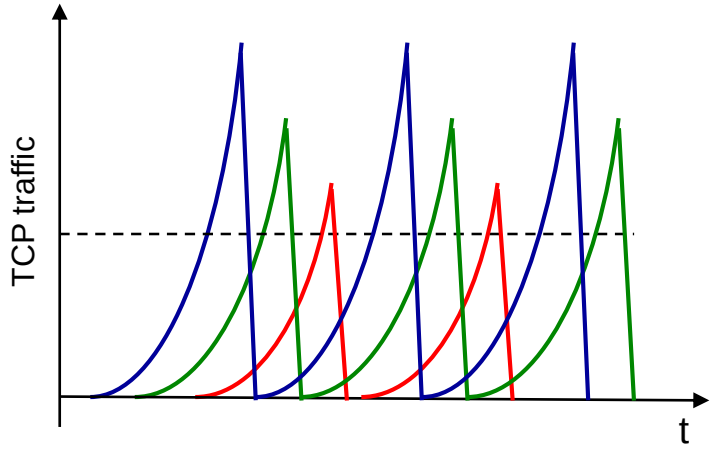
TCP Global Synchronization ‘’Issue”

Without Congestion Avoidance



multiple flows are throttling back followed by a sustained period of lowered link utilization, thus causing queue oscillation

With Congestion Avoidance (RED)



packets are partially dropped before the queues fill up, this gives flows such as TCP connections the opportunity to slow down the sending rate before the queues get full.

Layer 4
Transmission Control Protocol (TCP)
&
User Datagram Protocol (UDP)

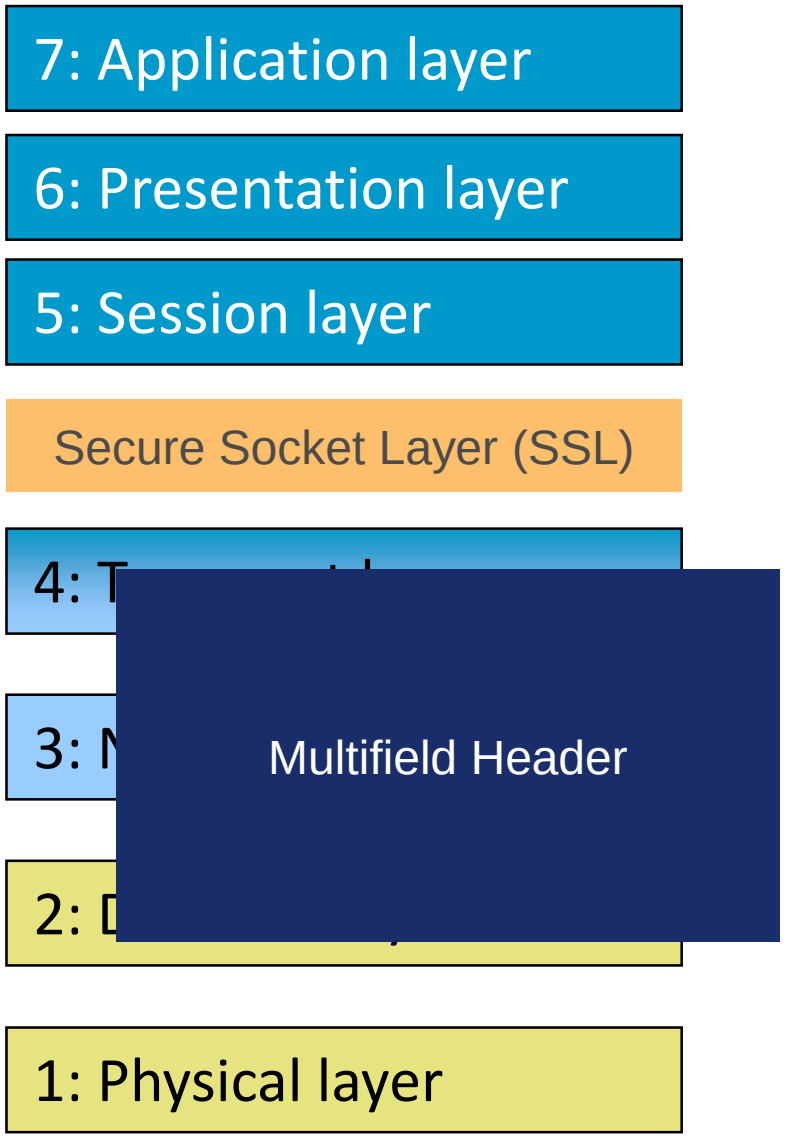
TCP & UDP

- The Transport Layer is in the end-point where network traffic is terminated, not forwarded. TCP and UDP are the dominant transport layer protocols, with SCTP becoming more important.
- TCP is used for reliable data delivery. The receiver confirms delivery of each TCP segment. If the sender doesn't get confirmation (TCP ACK), it will pause, then resend the data.
- UDP is used when 100% of the data doesn't have to arrive. Generally used for voice and video.

Analog to Digital Conversion	Not supported
Auto Discovery and Negotiation	Not supported
Source and Destination Addressing	Source & destination ports
Error Detection and Correction	TCP header covered by simple checksum, payload not covered
Access/Admission Control	Not supported
Reliable Delivery including Retransmission	TCP detects and retransmits lost segments, UDP doesn't
Flow Control, Congestion Mgt, and Prioritization	TCP includes flow control (back-off algorithm), UDP doesn't
Security	TCP can use TLS, UDP can use DTLS

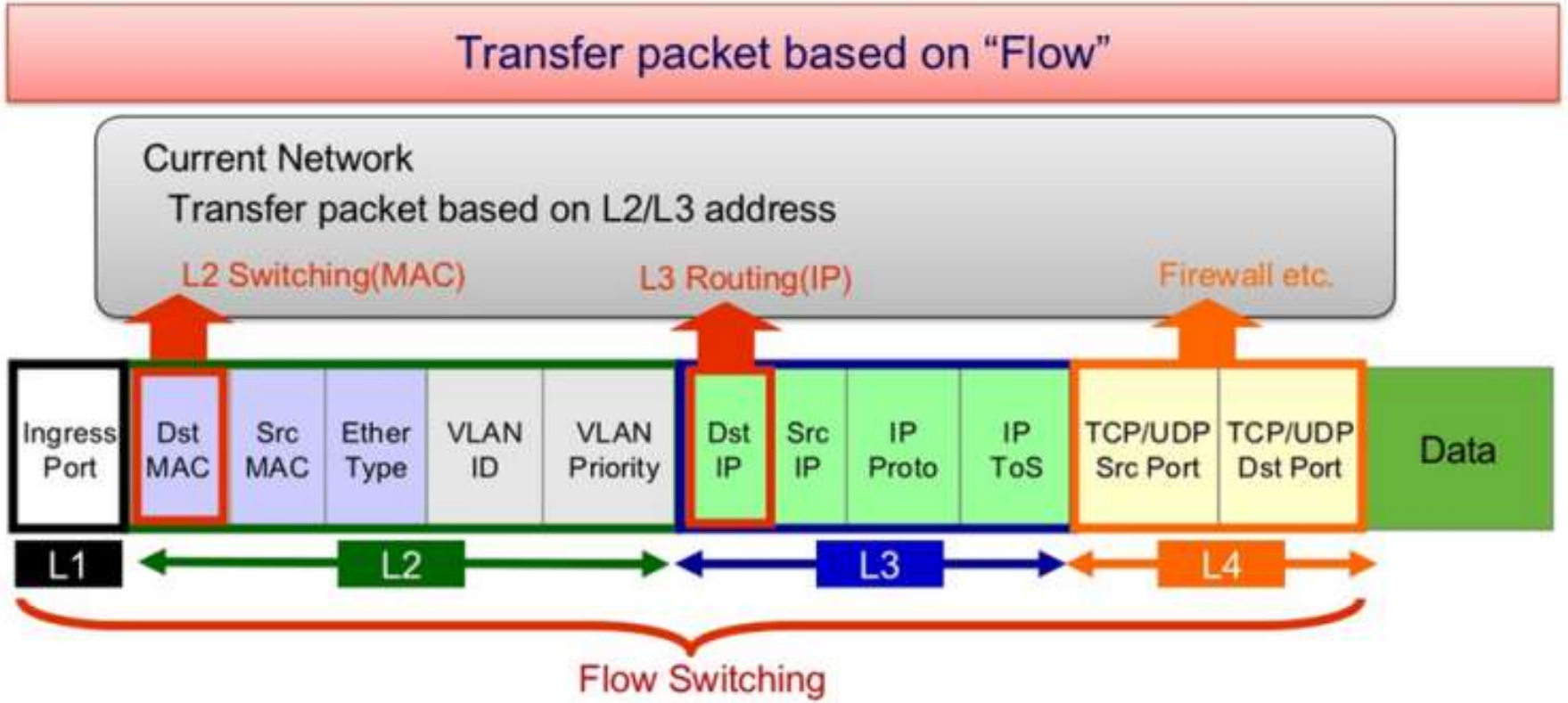
Software Defined Networking

Bye Bye OSI, Hello SDN!



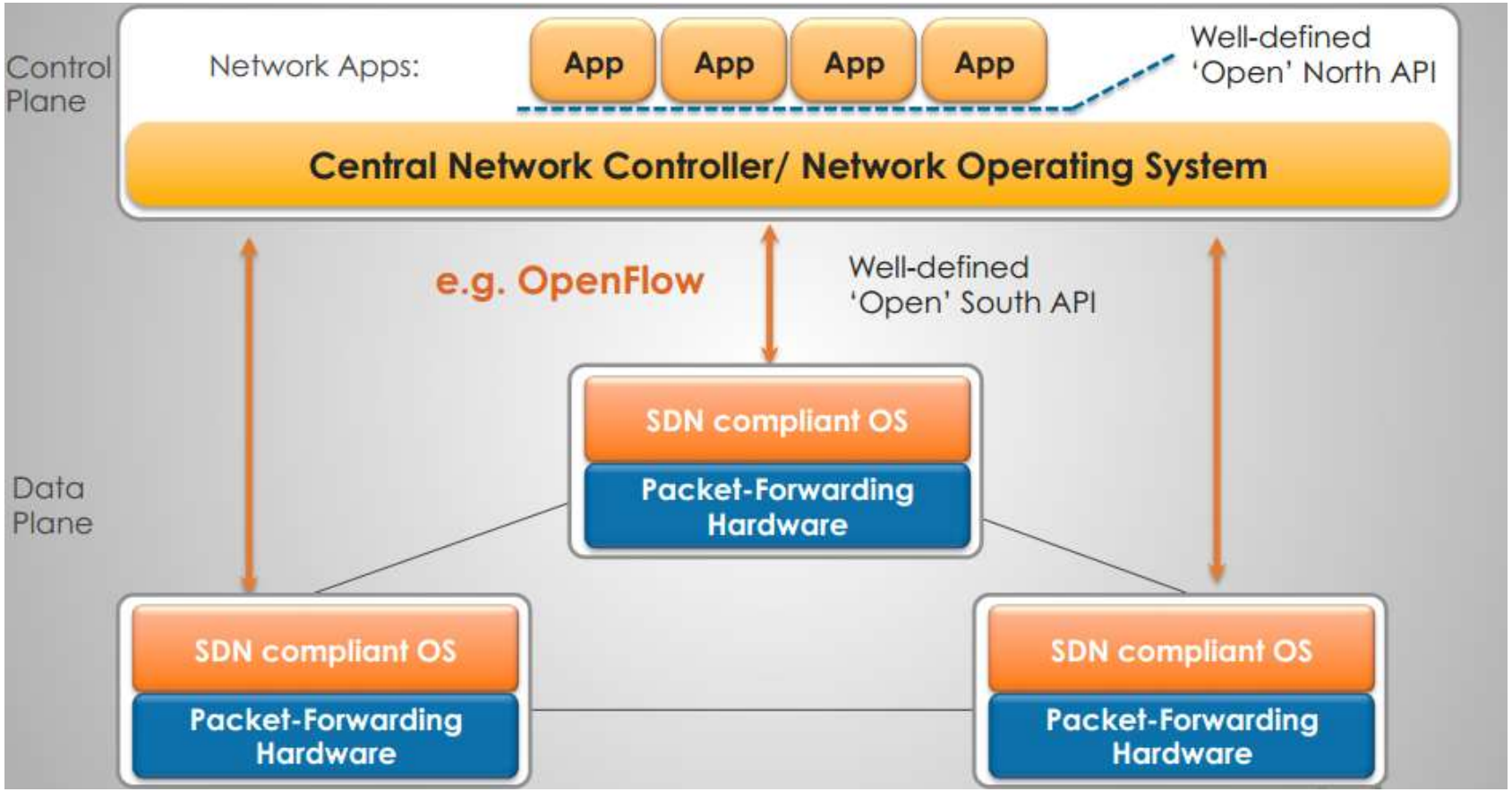
- SDN treats layer 2-4 headers as a pool of address and attributes against which a single forwarding table can be applied.
- By interface, the SDN data plane used configured rules to extract the fields to be used as a ‘key’ for a look-up.
- Each look-up has an associated action;
  - Drop
  - Add/remove header fields
  - Forward

SDN Flow Switching



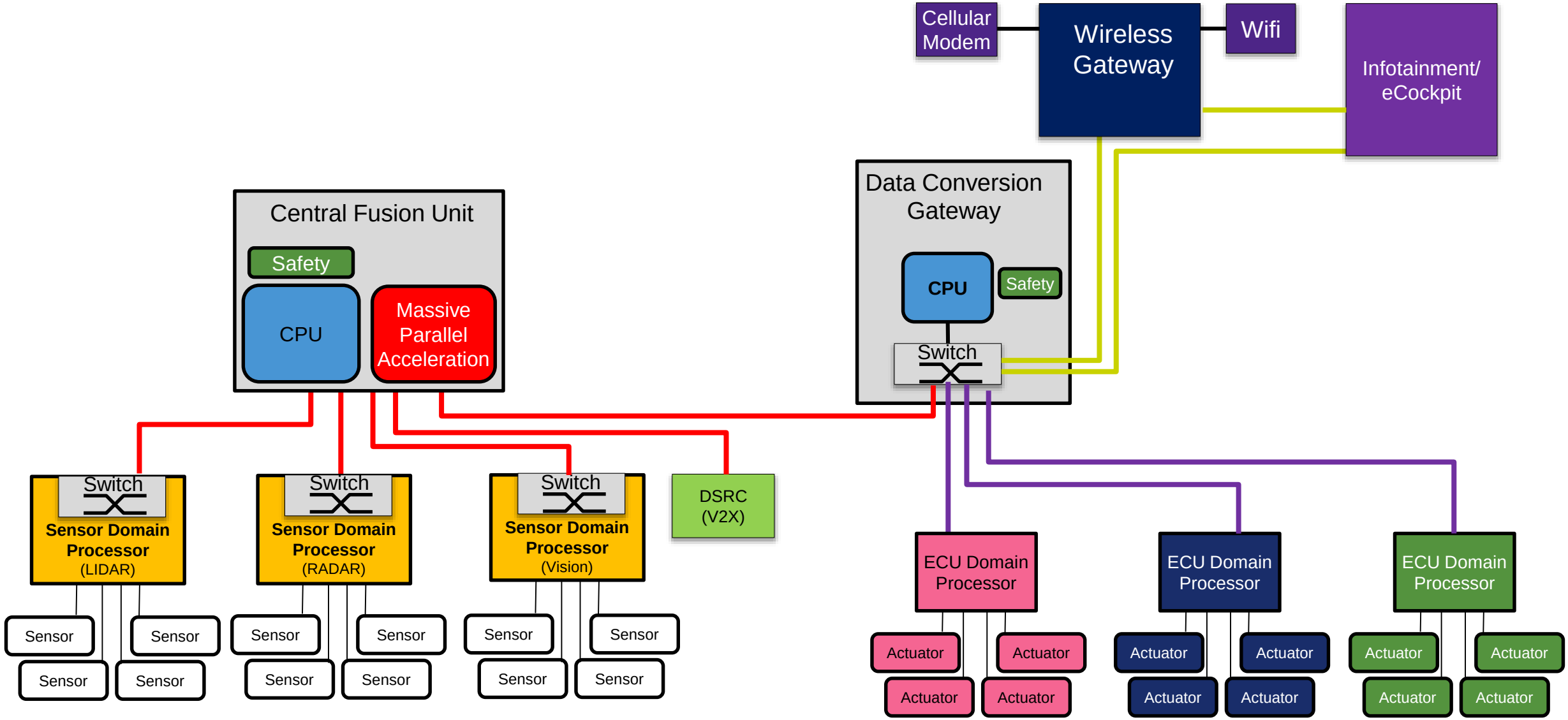
Flow is distinguished by rule of combination through L1(port), L2(MAC), L3(IP), L4(port).
Transferring method that use flow is called flow switching.

SDN Controller



- SDN standardizes the interface between the control plane which configures the look-up/action tables and the dataplane that uses the tables.

Conceptual Vehicle Architecture



SECURE CONNECTIONS
FOR A SMARTER WORLD