Tuesday, July 29, 2025 1:26 AM

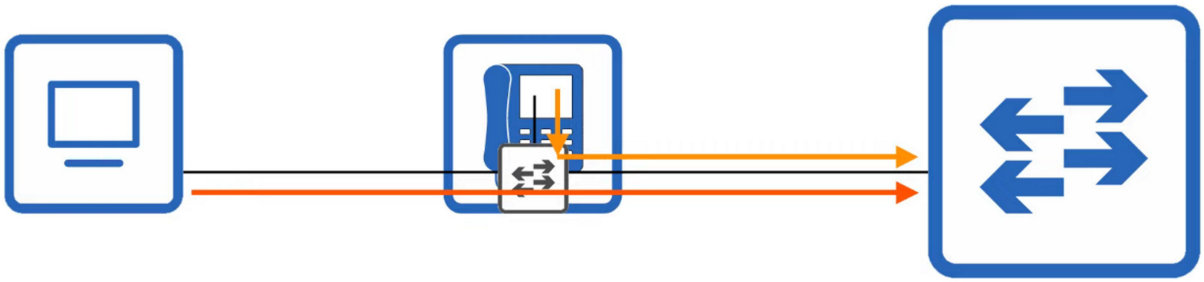
QoS (Voice VLANs) : PART 1

IP PHONES / VOICE LANS

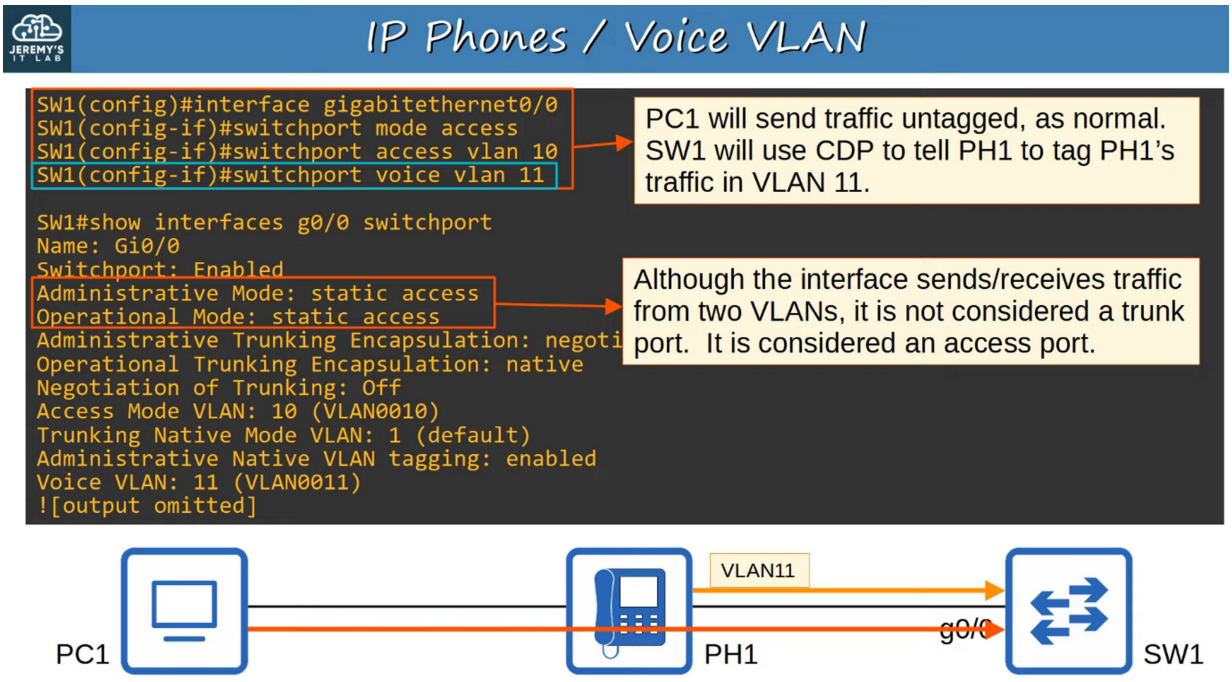
- Traditional phones operate over the *public switched telephone network* (PSTN)
  - Sometimes, this is called POTS (Plain Old Telephone System)
- IP PHONES use VoIP (Voice Over IP) technologies to enable phone calls over an IP NETWORK, such as the INTERNET
- IP PHONES are connected to a SWITCH, just like any other end HOST

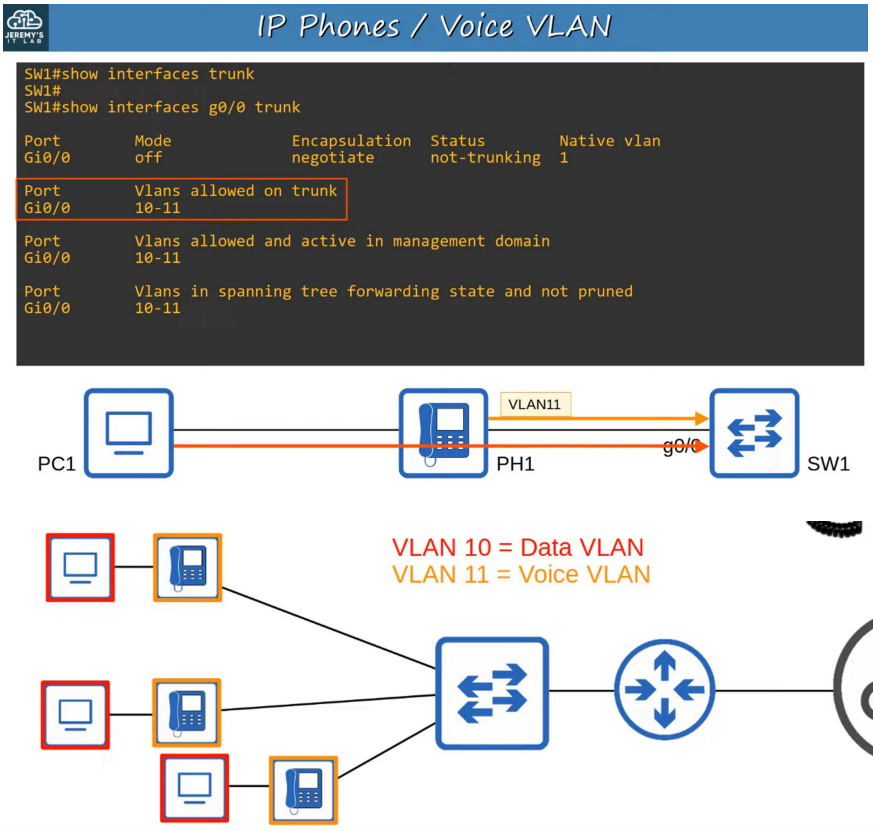
IP PHONES

- Have an internal 3-PORT SWITCH
  - 1 PORT is the “UPLINK” to the EXTERNAL SWITCH
  - 1 PORT is the “DOWNLINK” to the PC
  - 1 PORT connects internally to the PHONE itself



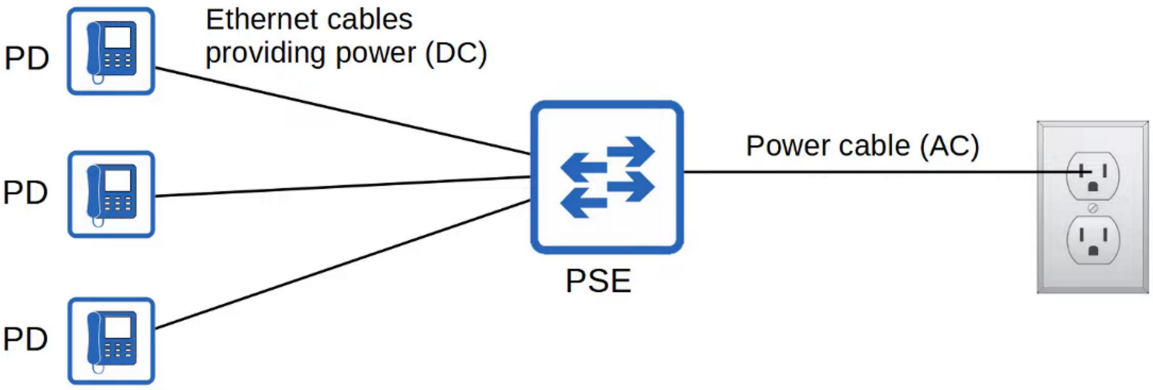
- This allows the PC and the IP PHONE to share a single SWITCH PORT. Traffic from the PC passes through the IP PHONE to the SWITCH
- It is RECOMMENDED to separate “VOICE” traffic (from IP PHONE) and “DATA TRAFFIC” (from the PC) by placing them into SEPARATE VLANS (!)
  - This can be accomplished using a VOICE VLAN
  - Traffic from the PC will be UNTAGGED - but traffic from the PHONE will be tagged with a VLAN ID





POWER OVER ETHERNET (PoE)

- PoE allows Power Sourcing Equipment (PSE) to provide POWER to Powered Devices (PD) over an ETHERNET cable
- Typically, the PSE is a SWITCH and the PDs are IP PHONES, IP CAMERAS, WIRELESS ACCESS POINTS, etc.
- The PSE receives AC POWER from the outlet, converts it to DC POWER, and supplies that DC POWER to the PDs



- TOO much electrical current can damage electrical DEVICES
- PoE has a process to determine if a CONNECTED DEVICE needs power and how much it needs.
    - When a DEVICE is connected to a PoE-Enabled PORT, the PSE (SWITCH) sends LOW POWER SIGNALS, monitors the response, and determines how much power the PD needs

- If the DEVICE needs POWER, the PSE supplies the POWER to allow the PD to boot
- The PSE continues to monitor the PD and SUPPLY the required amount of POWER (but not too much!)

- *POWER POLICING* can be configured to prevent a PD from taking TOO much POWER
  - 'power inline police' configures power policing with the default settings: disable the PORT and send a SYSLOG message if a PD draws too much power
    - Equivalent to 'power inline police action err-disable'
    - The INTERFACE will be put in an ‘error-disabled’ state and can be re-enabled with 'shutdown' followed by 'no shutdown'

```
SW1# conf t
Enter configuration commands, one per line. End with CNTL/Z.
SW1(config)# int g0/0
SW1(config-if)# power inline police
SW1(config-if)# end
SW1# show power inline police g0/0
Available:800(w) Used:32(w) Remaining:768(w)
Interface Admin  Oper       Admin      Oper       Cutoff Oper
          State  State      Police     Police     Power  Power
--------- ------ ---------- ---------- ---------- ------ -----
Gi2/1     auto   on         errdisable ok          17.2  16.7
```

  - 'power inline police action log' does NOT shut down the INTERFACE if the PD draws too much power. It WILL restart the INTERFACE and send a SYSLOG message

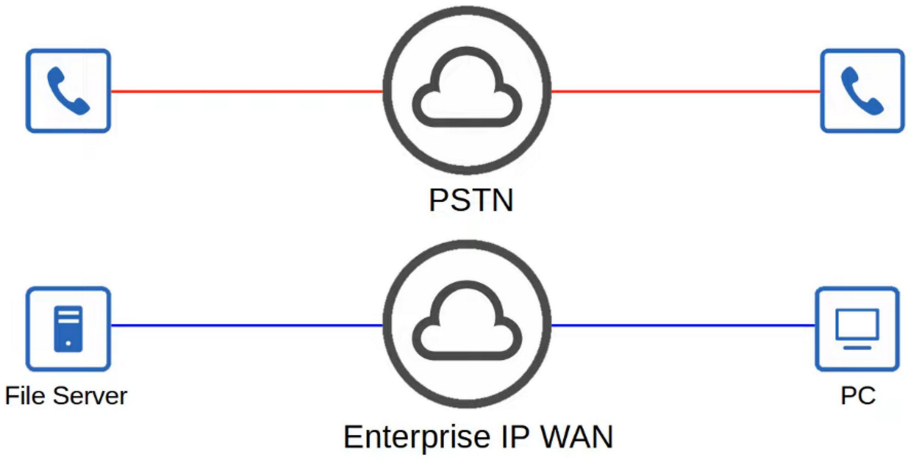
```
SW1# conf t
Enter configuration commands, one per line. End with CNTL/Z.
SW1(config)# int g0/0
SW1(config-if)# power inline police action log
SW1(config-if)# end
SW1# show power inline police g0/0
Available:800(w) Used:32(w) Remaining:768(w)
Interface Admin  Oper       Admin      Oper       Cutoff Oper
          State  State      Police     Police     Power  Power
--------- ------ ---------- ---------- ---------- ------ -----
Gi0/0     auto   on         log        ok          17.2  16.7
```

Power over Ethernet (PoE)

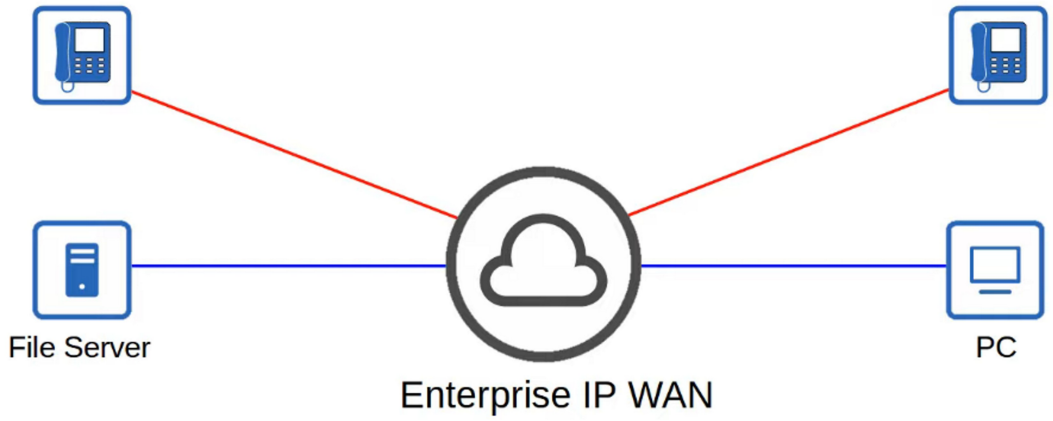
Name	Standard #	Watts	Powered Wire Pairs
Cisco Inline Power (ILP)	Made by Cisco, not standard	7	2
PoE (Type 1)	802.3af	15	2
PoE+ (Type 2)	802.3at	30	2
UPoE (Type 3)	802.3bt	60	4
UPoE+ (Type 4)	802.3bt	100	4

INTRO TO QUALITY OF SERVICE (QoS)

- VOICE traffic and DATA traffic used to use entirely separate NETWORKS
  - VOICE TRAFFIC used the PSTN(Public switch telephone network)
  - DATA TRAFFIC used the IP NETWORK (Enterprise WAN, Internet, etc)
- QoS wasn't necessary as the different kinds of TRAFFIC didn't compete for BANDWIDTH



- Modern NETWORKS are typically *converged networks* in which IP PHONES, VIDEO TRAFFIC, REGULAR TRAFFIC, etc. all share the same IP NETWORK
- This enables COST SAVINGS as well as more ADVANCED FEATURES for VOICE and VIDEO TRAFFIC (Example : Collaboration Software like Cisco WebEx, MS Teams, etc)
- HOWEVER, the different kinds of TRAFFIC now have to compete for BANDWIDTH
- QoS is a set of TOOLS used by NETWORK DEVICES to apply different TREATMENT to different PACKETS



QUALITY OF SERVICE (QoS)

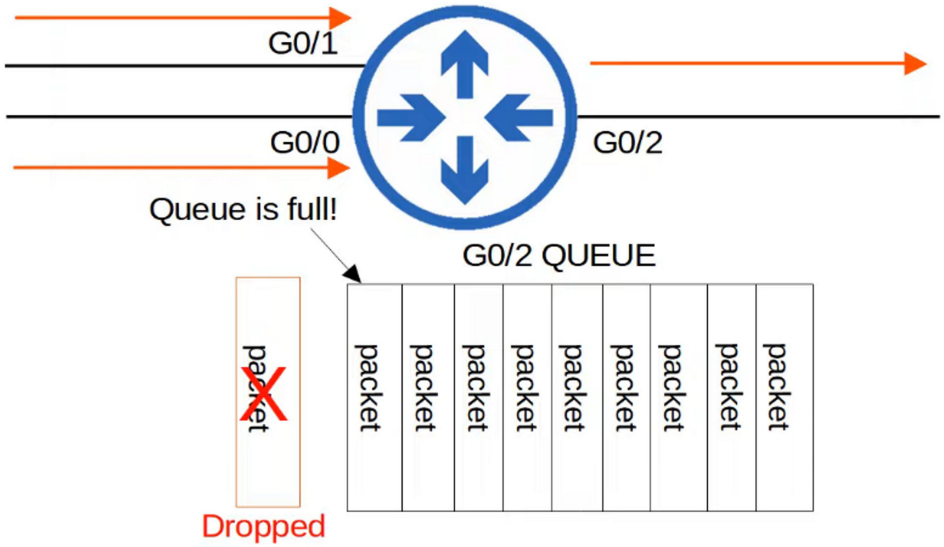
- QoS is used to manage the following characteristics of NETWORK TRAFFIC
  - BANDWIDTH
    - Overall CAPACITY of the LINK (measured in *bits per second*)
    - QoS TOOLS allow you to RESERVE a certain amount of a link's BANDWIDTH for specific kinds of traffic
  - DELAY
    - One-Way Delay = Time it takes traffic to go from SOURCE to DESTINATION
    - Two-Way Delay = Time it takes traffic to go from SOURCE to DESTINATION and return

- JITTER
  - The variation in ONE-WAY DELAY between PACKETS SENT by the same APPLICATION
  - IP PHONES have a 'jitter buffer' to provide a FIXED DELAY to audio PACKETS
- LOSS
  - The % of PACKETS sent that DO NOT reach their DESTINATION
  - Can be caused by FAULTY CABLES
  - Can also be caused when a DEVICE'S PACKET QUEUES get full and the DEVICE starts discarding PACKETS

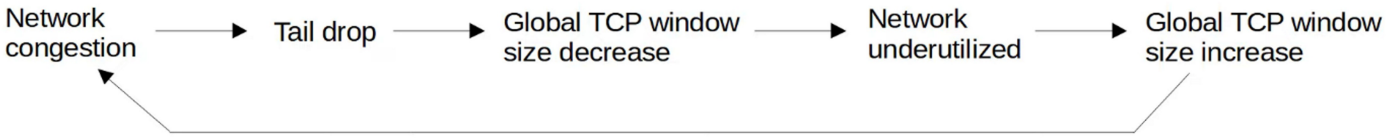
- The FOLLOWING STANDARDS are recommended for ACCEPTABLE INTERACTIVE AUDIO quality:
  - ONE-WAY DELAY : 150 milliseconds or less
  - JITTER : 30 milliseconds or less
  - LOSS : 1% or less
- If these STANDARDS are not met, there could be a noticeable reduction in the QUALITY of the phone call

QoS QUEUING

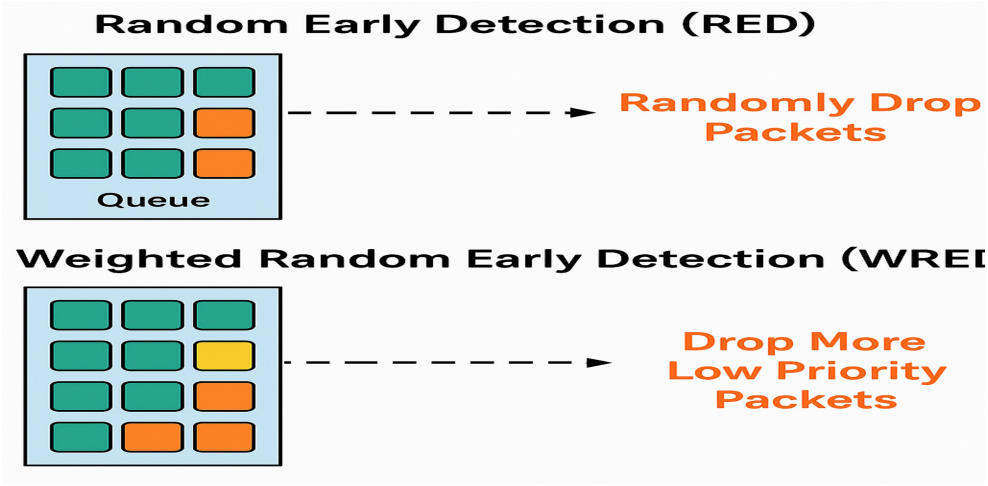
- If a NETWORK DEVICE receives messages FASTER than it can FORWARD them out of the appropriate INTERFACE, the MESSAGES are placed in the QUEUE
- **By default, the QUEUED MESSAGES will be FORWARDED in a FIRST IN FIRST OUT** (FIFO) manner
  - Message will be SENT in the ORDER they are RECEIVED
- If the QUEUE is FULL, new PACKETS will be DROPPED
- **The is called *tail drop***



- TAIL DROP is harmful because it can lead to **TCP GLOBAL SYNCHRONIZATION**
- Review of the **TCP sliding window**:
  - → Hosts using TCP use the 'sliding window' increase/decrease the rate at which they send traffic as needed.
  - → When a packet is dropped it will be re-transmitted.
  - → When a drop occurs, the sender will reduce the rate it sends traffic.
  - → It will then gradually increase the rate again.
- When the QUEUE fills UP and TAIL DROP occurs, ALL TCP HOSTS sending traffic will SLOW DOWN the rate at which they SEND TRAFFIC
- They will ALL then INCREASE the RATE at which they send TRAFFIC, which rapidly leads to MORE CONGESTION, dropped PACKETS, and the process REPEATS…



- A SOLUTION to prevent TAIL DROP and TCP GLOBAL SYNCHRONIZATION is RANDOM EARLY DETECTION (RED)
- When the amount of TRAFFIC in the QUEUE reaches a certain THRESHOLD, the DEVICE will start RANDOMLY dropping PACKETS from select TCP FLOWS
- Those TCP FLOWS that dropped PACKETS will reduce the RATE at which TRAFFIC is sent, but you will avoid TCP GLOBAL SYNCHRONIZATION, in which ALL TCP FLOWS reduce and then increase the rate of transmission at the same time, in waves.
- In STANDARD RED, all kinds of TRAFFIC are treated the SAME
- WEIGHTED RANDOM EARLY DETECTION (WRED) - an improved version of RED, allows you control which PACKETS are dropped depending on the TRAFFIC CLASS


Random Early Detection (RED)
Queue
Randomly Drop Packets
Weighted Random Early Detection (WRE
Drop More Low Priority Packets