# Monitorización de red: Captura de tráfico

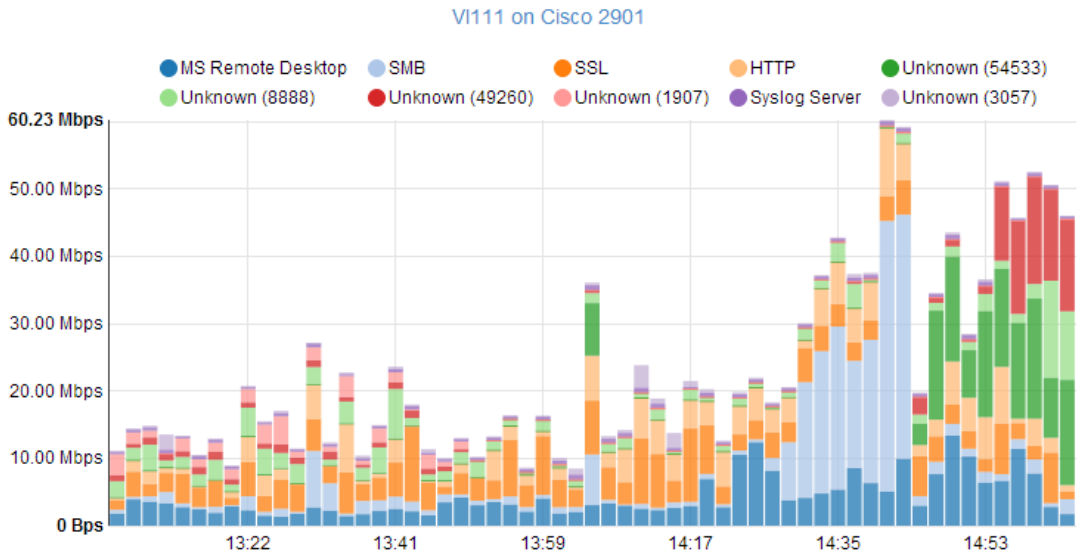
Area de Ingeniería Telemática
http://www.tlm.unavarra.es

Grado en Ingeniería en Tecnologías de Telecomunicación, 4º

# Medición de flujos

# ¿Flujos?

- Normalmente definidos a nivel de transporte
- IPsrc : puertoSrc : IPdst : puertoDst : protocolo
- “Protocolo” para distinguir UDP de TCP
- Si es ICMP, evidentemente no hay puertos
- Varias alternativas
  - NetFlow
  - IPFIX
  - J-Flow
  - sFlow

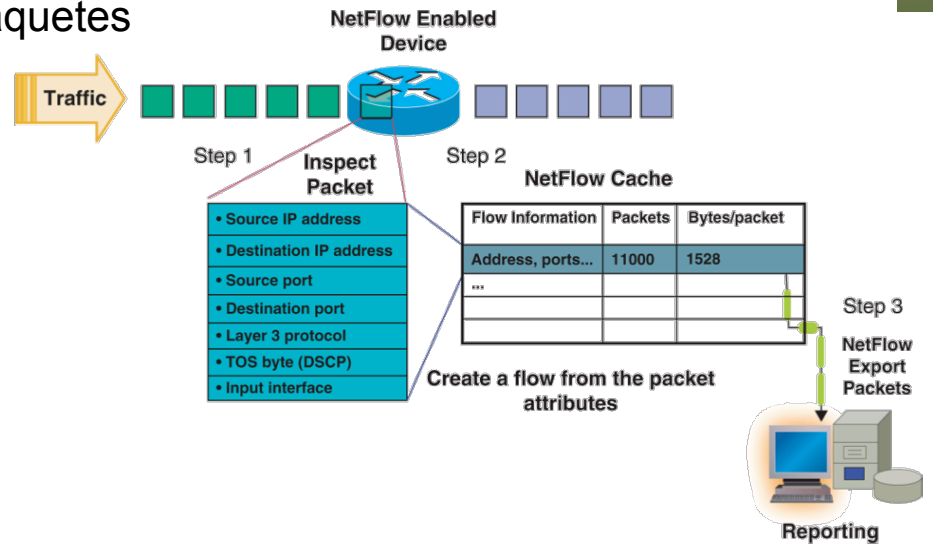


# NetFlow

- Cisco (patentado)
- Origen:
  - IOS mantiene una cache de flujos activos
  - Cache para acelerar la toma de decisiones de reenvío (CEF = *Cisco Express Forwarding*)
  - Con contadores sobre cada uno (bytes, paquetes, etc) en *flow records*
- Para paquetes IPv4 e IPv6 (y MPLS), unicast y multicast
- Se puede emplear para
  - Monitorización y planificación de red
  - Accounting/billing
  - Traffic matrix
  - Detectar ataques
- Ligeramente diferente en routers y switches (switches gama alta)
- Tiene efecto en el uso de CPU del equipo

# NetFlow: flujos

- RFC 3954 “Cisco Systems NetFlow Services Export Version 9”:
  - *“An IP Flow, also called a Flow, is defined as a set of IP packets passing an Observation Point in the network during a certain time interval. All packets that belong to a particular Flow have a set of common properties derived from the data contained in the packet and from the packet treatment at the Observation Point.”*
- Flujos unidireccionales
- Una serie de valores (*keys*) del paquete determinan el flujo:
  - Direcciones IP origen y destino
  - Protocolo sobre IP
  - Puertos de transporte origen y destino
  - Interfaz por el que llegan los paquetes
  - DSCP



# Netflow: valores

- Número de paquetes y bytes
- Timestamp de primer y último paquete
- Interfaz de entrada y salida
- Next-hop
- ASN origen y destino
- Puede añadir Layer 2
  - Dirección MAC origen y VLAN ID de tramas recibidas
  - Dirección MAC destino y VLAN ID de tramas transmitidas
- Para seguridad puede añadir
  - Máximo y mínimo TTL
  - Máxima y mínima longitud de paquete
  - IPID
  - Código y tipo ICMP
  - Flags TCP acumulados
- Versión 9 provee una definición más flexible del registro para poder añadir campos (*templates*)

# NetFlow: agregación

- Versiones: 1 (*legacy*), 5, 7 (solo Catalyst), 8 (agregación), 9 (flexible y extensible), 10 (IPFIX)
- Agregación:
  - Por AS origen y destino
  - Por dirección o prefijo IP origen y destino
  - Por protocolo y puerto
  - etc.

**1. Create and update flows in NetFlow cache**

| SrcIf | SrcIPadd | DstIf | DstIPadd | Protocol | TOS | Flgs | Pkts | Src Port | Src Msk | Src As | Dst Port | Dst Msk | Dst As | NextHop | Bytes /Pkt | Active | Idle |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Fa1/0 | 173.100.21.2 | Fa0/0 | 10.0.227.12 | 11 | 80 | 10 | 11000 | 00A2 | /24 | 5 | 00A2 | /24 | 15 | 10.0.23.2 | 1528 | 1745 | 4 |
| Fa1/0 | 173.100.3.2 | Fa0/0 | 10.0.227.12 | 6 | 40 | 0 | 2491 | 15 | /26 | 196 | 15 | /24 | 15 | 10.0.23.2 | 740 | 14.5 | 1 |
| Fa1/0 | 173.100.20.2 | Fa0/0 | 10.0.227.12 | 11 | 80 | 10 | 10000 | 00A1 | /24 | 180 | 00A1 | /24 | 15 | 10.0.23.2 | 1428 | 1145.5 | 3 |
| Fa1/0 | 173.100.6.2 | Fa0/0 | 10.0.227.12 | 6 | 40 | 0 | 2210 | 19 | /30 | 180 | 19 | /24 | 15 | 10.0.23.2 | 1040 | 24.5 | 14 |

**2. Expiration**

- Inactive Timer Expired (15 Sec Is Default)
- Active Timer Expired (30 Min Is Default)
- NetFlow Cache Is Full (Oldest Flows Are Expired)
- RST or FIN TCP Flag

| SrcIf | SrcIPadd | DstIf | DstIPadd | Protocol | TOS | Flgs | Pkts | Src Port | Src Msk | Src As | Dst Port | Dst Msk | Dst As | NextHop | Bytes /Pkt | Active | Idle |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Fa1/0 | 173.100.21.2 | Fa0/0 | 10.0.227.12 | 11 | 0 | 10 | 11000 | 00A2 | /24 | 5 | 00A2 | /24 | 15 | 10.0.23.2 | 1528 | 1800 | 4 |

**3. Aggregation?** (No / Yes)

**4. Export version**

Non-aggregated flows—export version 5 or 9

**5. Transport protocol**

Export packet: Header | Payload (Flows)

**E.g. Protocol-Port Aggregation Scheme Becomes**

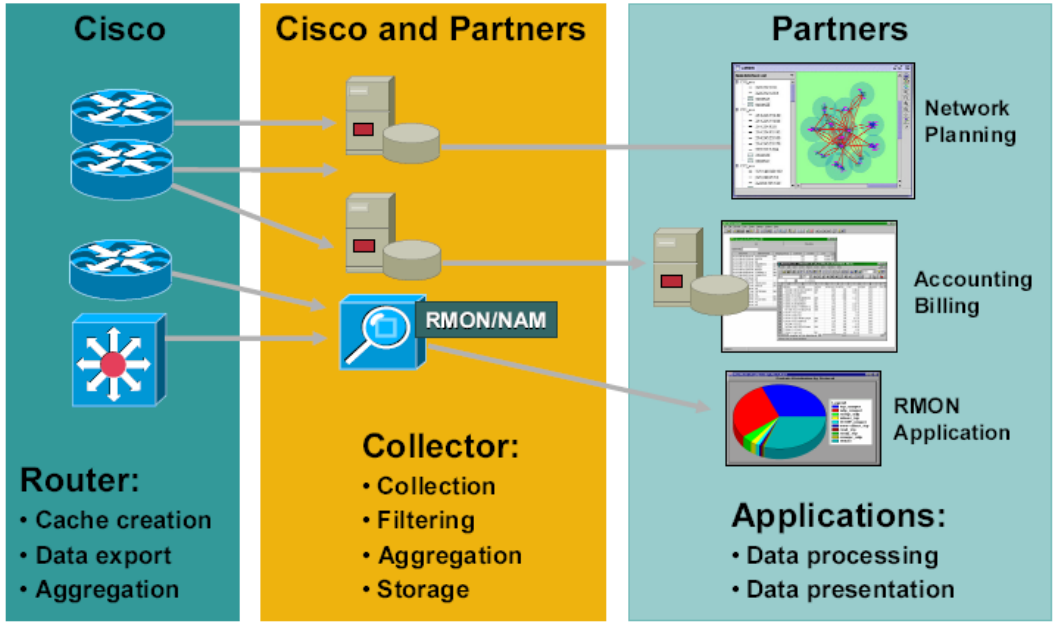
| Protocol | Pkts | Srcport | Dstport | Bytes/Pkt |
|---|---|---|---|---|
| 11 | 11000 | 00A2 | 00A2 | 1528 |

Aggregated Flows—Export Version 8 or 9

B. Claise y R. Wolter, "Network Management: Accounting and Performance Strategies", Cisco Press

# NetFlow: exportación

- Los flujos se eliminan de la cache por inactividad
- Con actividad, llegado un tiempo máximo también se eliminan
- Flujos TCP se eliminan ante banderas de FIN o RST
- Si se llena la cache eliminan flujos aunque no hayan caducado
- Estos *flow records* se exportan a un *collector* por UDP o SCTP (RFC 2960)
- Puede ser un ordenador o un módulo en un router/switch
- Exporta múltiples registros en un paquete
- Una aplicación de gestión puede analizar esos registros o exportarse a MIB RMON

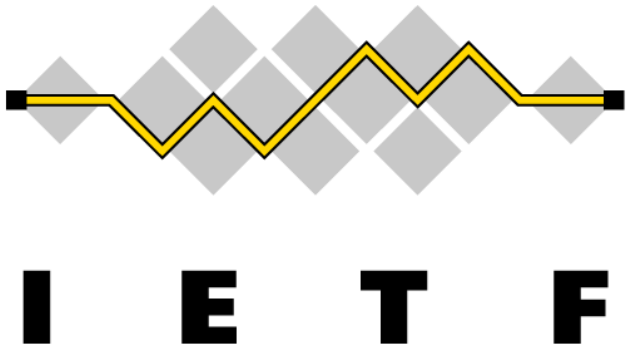


# Otras alternativas para flujos

**IPFIX**

- Cisco publicó RFC 3954 “Cisco Systems NetFlow Services Export Version 9”
- Informativa, simplemente para ser el punto de arranque del desarrollo de IPFIX
- De hecho en la cabecera IPFIX dice que es la versión 10
- RFC 5101 “Specification of the IP Flow Information Export (IPFIX) Protocolo for the Exchange of IP Traffic Flow Information”
- RFC 5102, 5470, 5610, 5655,...
- También se exportan los flujos en modo *push* al *collector*
- Un *template* describe el formato del registro y se envía también al *collector*

**(...)**



# Otras alternativas para flujos

**IPFIX**

**J-Flow**

- Juniper
- Paquetes muestreados

**sFlow**

- InMon Corporation
- Mide flujos hasta L7 con gran cantiad de interfaces mediante packet sampling

**IPDR (TM Forum), LFAP (Riverstone), CRANE (RFC 3423)...**

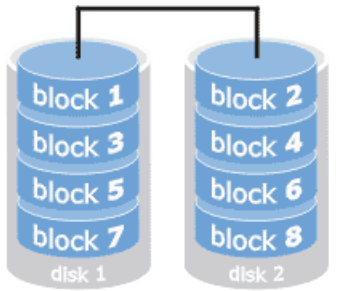
# Captura de tráfico

# ¿Los flujos son suficiente?

- Los contadores en la MIB son una medida muy agregada que no distingue origen y destino
- Los flujos (NetFlow o conversaciones RMON) distinguen a los extremos pero siguen siendo contadores
- Hay cierta temporalidad pues aun estando activo el flujo se envía el registro cada cierto tiempo
- Pero sigue siendo información muy agregada
- Por ejemplo una serie temporal de tráfico tiene como mínimo la escala de exportación
- ¿Qué hacer si queremos la máxima información?

# Recoger todo el tráfico

- Ventaja: Permite máximo detalle en el análisis
- Inconveniente: La cantidad de datos e información puede ser descomunal
- ¿Análisis online u offline?
  - 1Gbps = 125MB/s = 450GB/h
  - Digamos utilización del 60% durante 4h + utilización del 30% durante otras 4h y resto 0% = 450x0.6x4 + 450x0.3x4 = 1.6 TB/día
  - Online
    - No requiere gran cantidad de espacio en disco
    - Requiere (según el análisis) considerable capacidad de CPU y RAM (llegando a necesitar múltiples cores o incluso miles en GPGPUs)
  - Offline
    - Requiere almacenar en disco
    - Según el tipo de disco duro hablamos de unas velocidades sostenidas de 100-200MB/s, hasta 300MB/s (enterprise), 500-800MB/s SSD
    - ¿Un interfaz a 10Gbps con picos? Puede requerir varios discos en paralelo

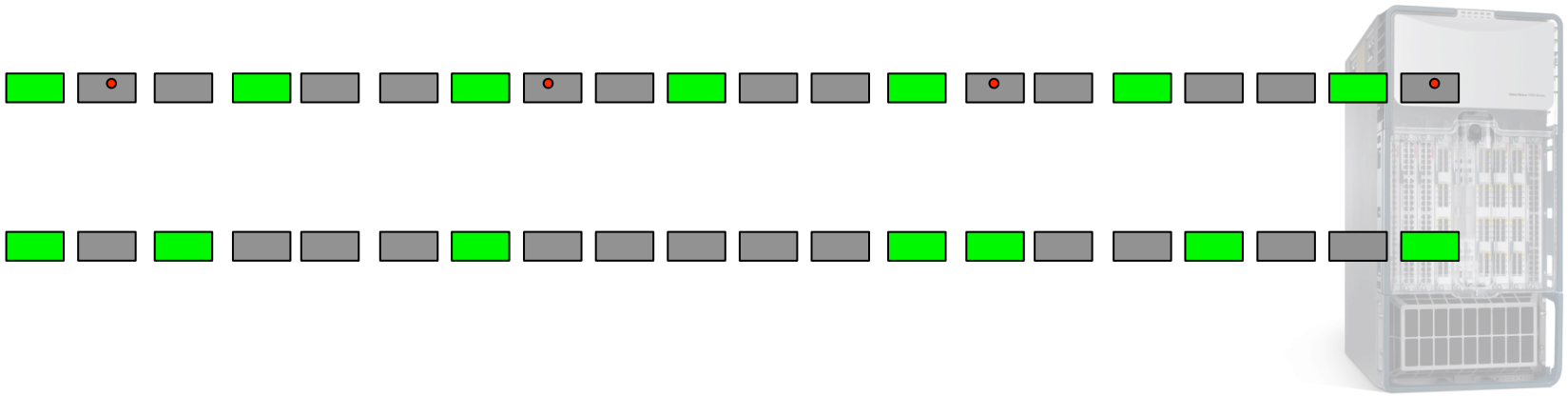


# Completo vs muestreado

- Para medición de flujos o registros de paquetes
- Podemos analizar/recoger todos los paquetes o solo una fracción
- Completo
  - No nos dejamos nada, no hay error de medida
  - Podemos seguir el estado de sesiones
  - Podemos inspeccionar contenido de aplicación
  - Pero gran cantidad de información recogida (número de paquetes o flujos)
  - Puede llegar a tener requisitos serios de CPU y throughput a disco, throughput de registros en red (NetFlow), etc
- Muestreado (...)

# Completo vs muestreado

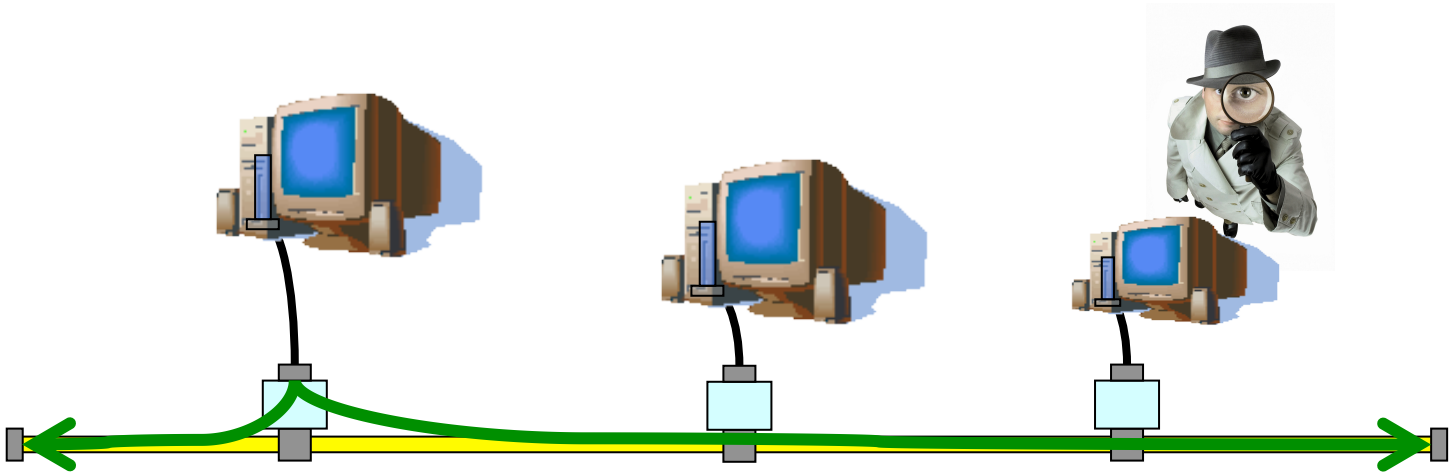
- Muestreado
  - En conmutadores/routers de gama alta puede que la medición tenga que ser “muestreada”
  - Se analiza/recoge 1 de cada N paquetes o cada N ms
  - Escala mejor para grandes tasas de tráfico
  - Pero podemos dejarnos algo importante (por ejemplo para facturar)
  - Muestreo determinista o probabilístico (1 de cada N al azar o para cada uno una probabilidad)
  - Determinista puede sesgar los resultados ante patrones periódicos
  - IETF PSAMP (RFC 5474 “A Framework for Packet Selection and Reporting”)



# Modo promiscuo

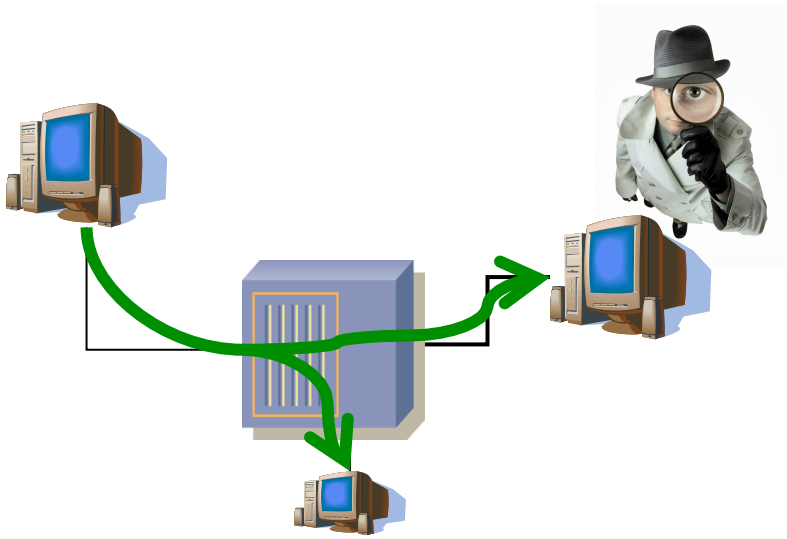
# ¿Cómo tener acceso al tráfico?

- Sencillo en la Ethernet original pues todos los hosts veían todas las tramas de la LAN
- Solo requería una NIC capaz de trabajar en modo *promiscuo*



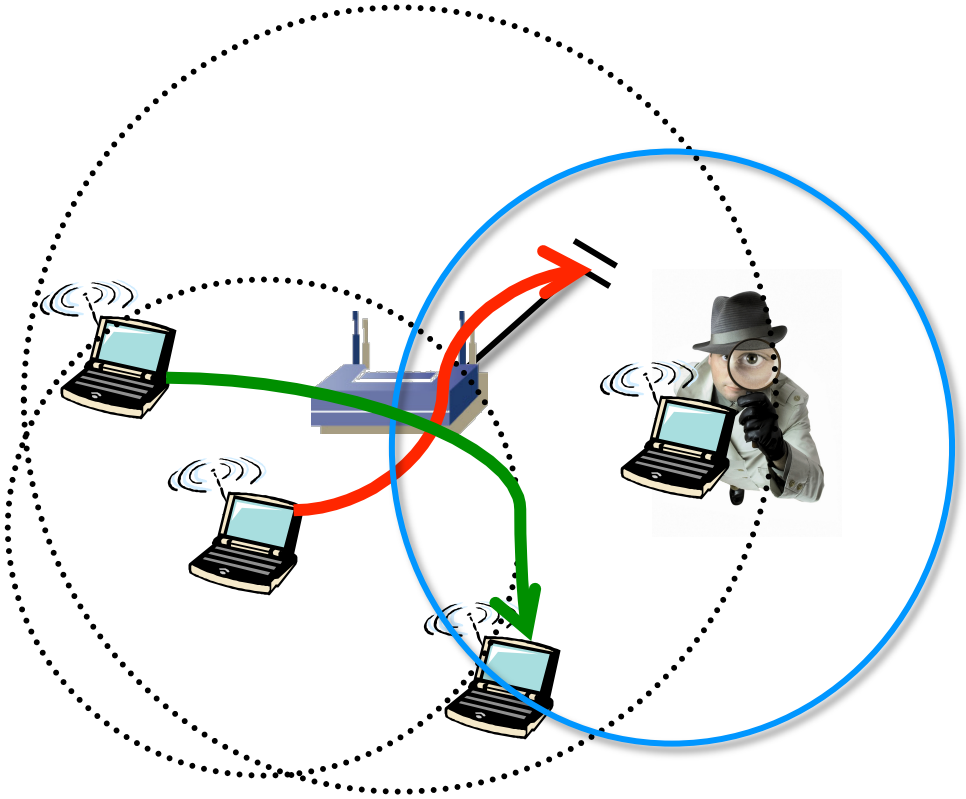
# ¿Cómo tener acceso al tráfico?

- En Ethernet sobre par de cobre con Hubs, similar



# ¿Cómo tener acceso al tráfico?

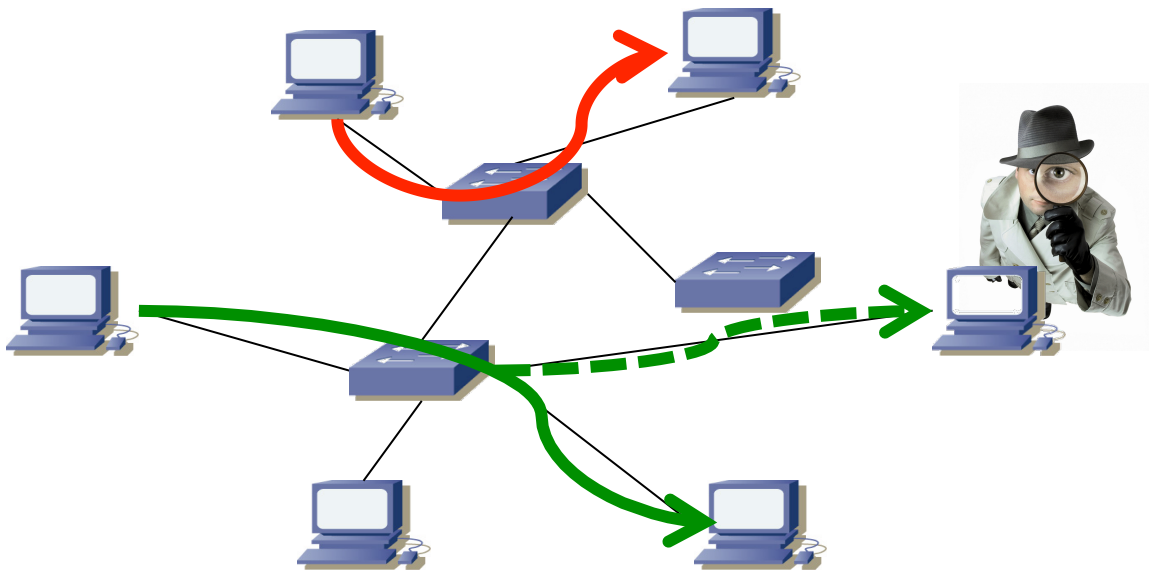
- En WLANs de nuevo un interfaz ve lo que envían otros
- Pero solo si está dentro del alcance
- ¿ En escenario BSS (con AP) ?
- Debería ver todo el tráfico reenviado por el AP a la WLAN
- Pero podría no ver el que vaya al DS (sí los ACK)



# *Mirrors* y *taps*

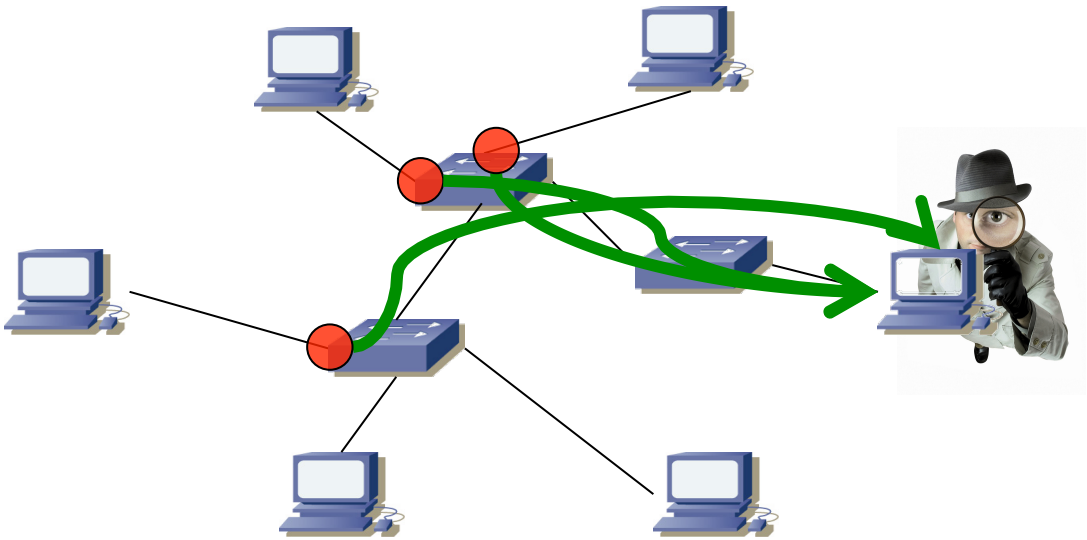
# ¿Cómo tener acceso al tráfico?

- Al llegar la conmutación Ethernet desaparece el dominio de colisión
- Un host solo puede ver las tramas que se dirijan a él o para las que se haga inundación
- Soluciones:
  - Mirror (...)



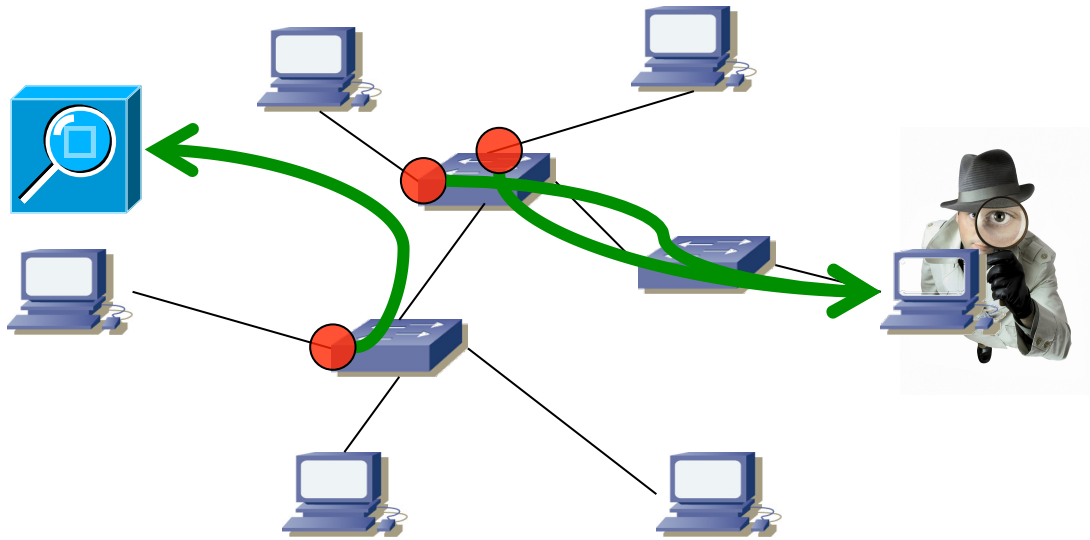
# Mirror/SPAN

- Cisco habla de “SPAN” (Switched Port ANalyzer)
- Port-based SPAN (PSPAN):
  - Se especifican uno o varios puertos origen y uno destino
  - Para cada puerto origen también si monitorizar rx, tx o ambos
- VLAN-based SPAN (VSPAN):
  - Origen todos los puertos que pertenecen a una VLAN
- Local SPAN: puertos monitorizados y destino en el mismo switch
- Remote SPAN (RSPAN)
  - Algunos puertos no están en el mismo switch que el puerto destino
  - Requiere una VLAN para transportar el tráfico monitorizado entre switches



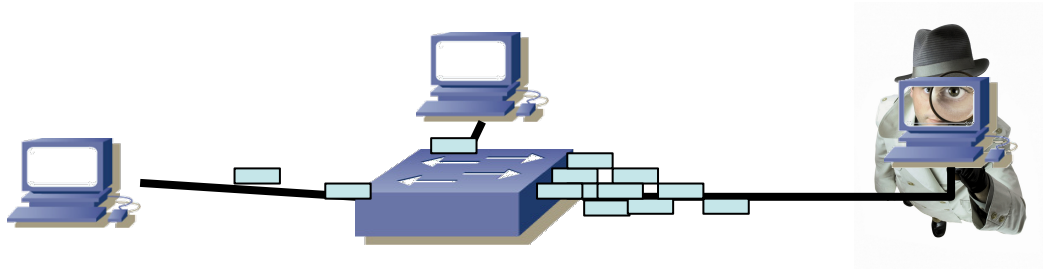
# Mirror/SPAN

- En un entorno conmutado una sonda RMON tiene dos alternativas:
  - Ser un módulo de un conmutador
  - Recibir el tráfico por un puerto que haga *mirroring*



# Mirror/SPAN: Limitaciones

- El puerto que recibe el tráfico puede congestionarse
  - Porque recibe de varios puertos o de uno pero los dos sentidos
  - Se podría dirigir el tráfico a una agregación de puertos
  - En algunos equipos se pueden crear reglas para indicar los paquetes a copiar al puerto de mirror y así restringirlo a un subconjunto
- Suele haber limitaciones en el número de sesiones de SPAN
- No reenvía paquetes estropeados pues no pasan del interfaz entrante
- Pueden aparecer duplicados



# Mirror: Ejemplo

- Juniper EX2200 Ethernet Switch

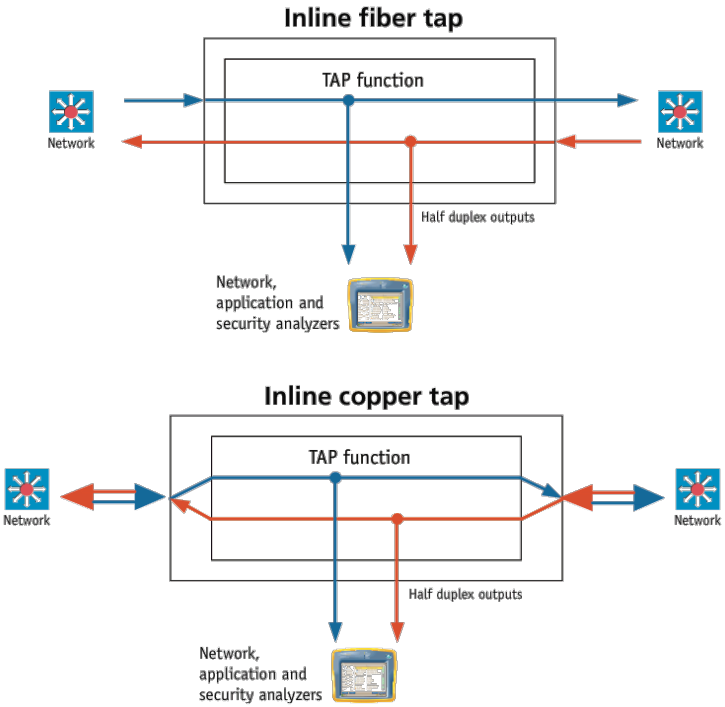
## Troubleshooting

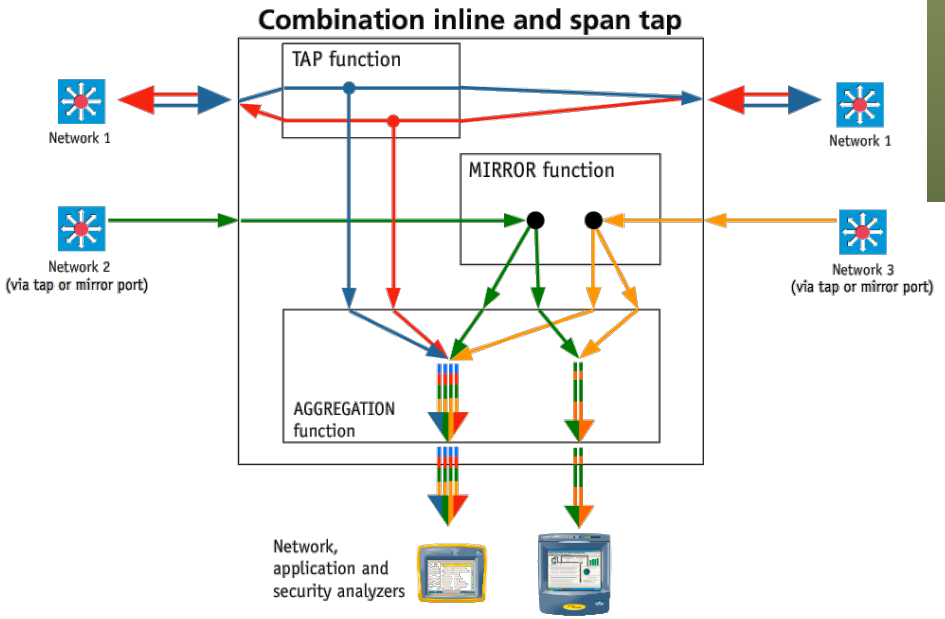
- Debugging: CLI via console, telnet, or SSH
- Diagnostics: Show and debug command statistics
- Traffic mirroring (port)
- Traffic mirroring (VLAN)
- ACL-based mirroring
- Mirroring destination ports per system: 1
- LAG port monitoring
- Multiple destination ports monitored to 1 mirror (N:1)
- Maximum number of mirroring sessions: 1
- Mirroring to remote destination (over L2): 1 destination VLAN
- IP tools: Extended ping and trace
- Juniper Networks commit and rollback

LAG = Link Aggregation Group

# ¿Cómo tener acceso al tráfico?

- Soluciones:
  - **Mirror**
  - **Network Tap**
    - Solo para el tráfico que circule por un enlace
    - Ante fallo de corriente mantienen el enlace de datos
    - En fibra, *splitters*




http://es.flukenetworks.com/enterprise-network/network-monitoring/Tap-Solutions

# Network Tap

NetOptics®

## Gig Zero Delay Tap


The Gig Zero Delay Tap is the industry's only 10/100/1000BaseT Tap with true Zero Delay operation. Using innovative new technology, this Tap guarantees absolutely zero packet loss on the network link even during power outages. With the Gig Zero Delay Tap, power glitches and failures no longer mean dropped packets and lengthy renegotiation sequences. Your network operates more smoothly and your critical business applications remain responsive with the Gig Zero Delay Tap in your monitoring infrastructure.

Get total traffic visibility for 10/100/1000 monitoring and security devices by deploying Net Optics Gig Zero Delay Taps on critical network links as permanent monitoring access ports. The Gig Zero Delay Tap sends copies of traffic moving in each direction on the link to a separate NIC on the monitoring device for comprehensive full-duplex monitoring. The Tap has no IP address, so monitoring devices are isolated from the network, dramatically reducing their exposure to attacks. The monitoring device connected to the Tap sees all full-duplex traffic as if it were in-line, including Layer 1 and Layer 2 errors.