SPAN port vs network TAP

for usage in APM/IM environments

Contents

Change History

Revision	Name	Email	Reason for Change
0.2	Joerg Mertin	Joerg.mertin@ca.com	Updated, added change due to comments/reviews
0.1	Joerg Mertin	Joerg.mertin@ca.com	Initial release

Document Purpose/Summary

This main purpose of this document is to raise our field staff's awareness of the differences when using a SPAN port instead of a network TAP and help them take the right decision when doing a POC or a final installation in a production environment.

Intended Audience

Technical field staff having to deal with a CA Software installation, notably in the APM/ADA environment.

Introduction

In the Monitoring worlds, one thinks that a SPAN port of a switch provides a 100% accurate copy of the traffic. This is only true in properly configured environments, and please make the distinction between accurate and mirror (100% exact copy) of the traffic.

To understand the difference in a network TAP vs. a SPAN port, a picture says more than a thousand words:

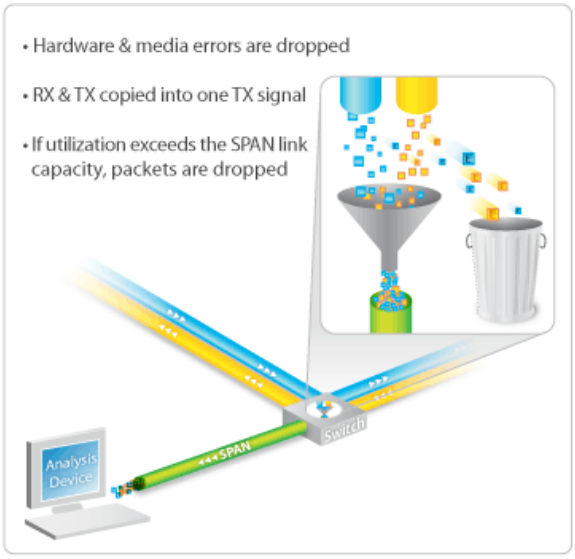

Fig. 1The SPAN Port will modify the forwarded data

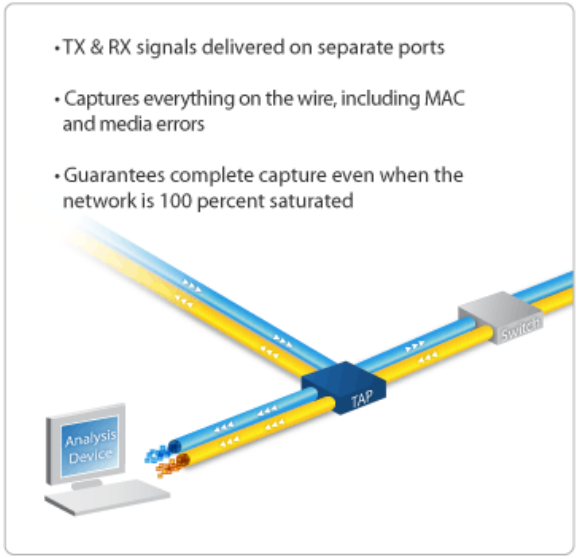

Fig. 2 The Network TAP will forward a real copy of the data.

There are however way more implications according to the switch configuration, number of ports to be SPAN'd to the output SPAN port, and if remote SPAN traffic flows will be used and integrated into the one output span port.
SPAN Ports are usually found in network Switches. One network switch has 2 to many (up to 100 and/or more) ports to service.
In very short terms, a Network TAP will provide a real copy of the network traffic, while the SPAN port will just forward us traffic that has been identified as “forward worthy”.

Network TAP

The network TAP is a technology that will, when installed correctly provide a 100% copy of the network traffic going over a physical link. To understand this, see Fig. 3.

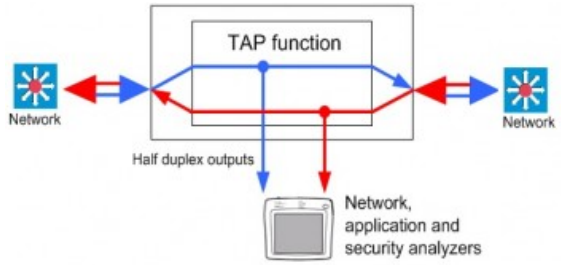

Fig. 3 – Network TAP function

It will really be a copy of the real traffic.
Due to the nature of the capture, the max throughput cannot be more than the throughput on the link the TAP has been tapped into !
This is the cleanest way to obtain a copy of network traffic. It will be 100% identical to the one passing on the tapped link .

Aggregation Network TAP's

Aggregation network TAP's are a mix between a Network TAP and a SPAN Port. However, these devices have been designed for exactly this purpose (which is not the case on standard SPAN Ports on the common Switch).

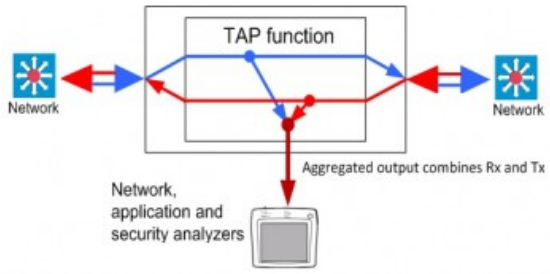

Fig. 4 – Network Aggregation TAP function

This technology will be subject, as the SPAN Port, to over-subscription of the Aggregated output port. The main difference to SPAN Ports is that this device has been designed for this, and can buffer the peaks for a certain amount of time. This can however affect performance metrics and cause extra latency to be reported as being due to either network or server (depending upon direction of data flow). In short, this can affect the accuracy of the monitored metrics.
If, for whatever reason the total aggregated traffic (Full Duplex of by example 1GB, in every direction) is higher than 1GB for a long time, the Aggregation TAP will drop data.

For high volume networks, a regular network TAP will be the better choice.

Best practice on setting up a Mirror port

The really best setup would be Passive network TAP's configured in-line in front of the devices to be monitored, as shown in figure 5 (Note: This does apply to TAP's in general, and this example is suitable for a NIDS sniffer).

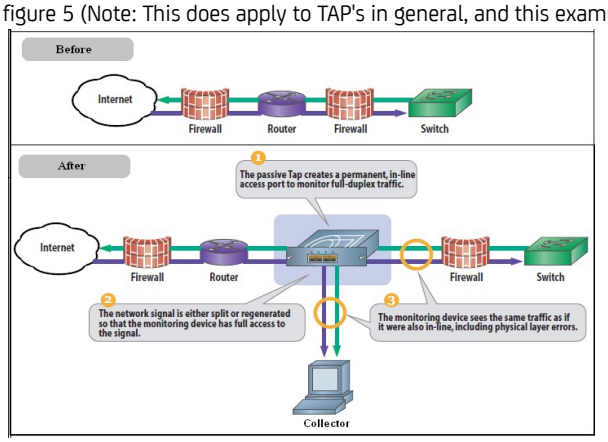

Fig 5. Passive TAP installed in-line.

This avoids any kind of errors due to wrong configuration, will never have an issue with oversubscribed network output SPAN ports.

In case – the collector device has only one port, but you need more than one passive TAP to provide monitoring data, place a specialized Network Aggregation device in front of the collector. These devices are specialized in doing exactly this, and utilize their entire CPU and memory resources to the task !

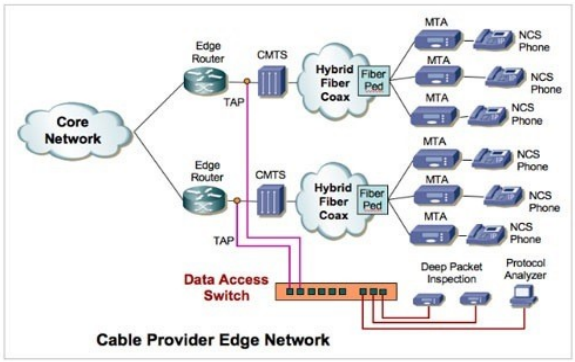

Fig 6. Passive TAP in-line hooked up to an Aggregation device

Good examples of such devices would be the Gigamon or Anue devices. NetOptics has also some very good devices specialized for that task !

SPAN Ports

The Switch SPAN port provides the Monitoring device connected to it a look alike copy of the traffic data it is “told” to forward. This also means that the quality of the forwarded traffic will depend on many factors induced by the technique used to SPAN traffic, but also by the purpose SPAN ports have been created in the first place.
SPAN ports initial usage was troubleshooting. This means that whatever traffic was broken, will not be added to the SPAN output port. However – the nature of the SPAN Port to be one usual physical port on a switch, or sometimes a dedicated Port with a higher throughput – will limit the amount of traffic this one can forward to the analyser.
Note also – that a Switch can usually forward traffic from all it's connected ports (except the SPAN output port), and has to do the operation as shown in fig. 1 – for each and every input port. But the Switch was designed for something else – like forward packets from one port to another one – without aggregating the traffic.

In contrary to a network TAP, a Switch is installed in any location one needs a connection point to a network. This makes the Switch SPAN Port very easy to apply/use. However – the implications can be devastating to the quality of the forwarded traffic to the SPAN output port, especially when Daisy-Chaining SPAN Port outputs (RSPAN).

The following factors will have to be taken into account when configuring a SPAN port.

1. Spanning will change the timing of the frame interaction (You will have a look alike copy, but not a real copy). No switch or Router can handle/replicating all data at high throughputs and handle its primary job of switching and/or routing in real time. In times of 10GB Full Duplex links, this has become more obvious.
2. Spanning is not the main priority on a Switch. If replicating a frame becomes an issue, the hardware will temporarily drop the SPAN process.
   By example Cisco warns that SPAN data is treated with a lower priority than regular port-to-port data. If transporting remote SPAN (RSPAN) traffic through a Inter Switch Link (ISL) which shares the ISL Bandwidth with regular network traffic, the network traffic takes priority. You will have to equip the Cisco switches with dedicated Monitoring cards (with own CPU and Memory) and use a dedicated ISL physical connection/cable to provide a decent copy of network traffic through SPAN technology that is not too modified.
3. If the output SPAN port becomes oversubscribed, frames are dropped (Spanning 2 full duplex 1GB links results in potentially forwarding 4GB of traffic through one 1GB SPAN port. We potentially loose 75% of the overall traffic).
   This is a physical limitation due to the maximum traffic the output port can forward !
4. Correct configuration of the SPAN port is not without its dangers. A misconfiguration can take out the entire switch operation. If a net-worker performs a change on a switch and the switch reloads its configuration, the time it takes the Switch to become operational again, no SPAN data will be forwarded (Depends on models though).
5. Depending on the SPAN port configuration, the SPAN output will show the data as seen while entering the Switch, while inside the Switch or while exiting the Switch. On the application layer this may not be of importance as the application layer is analysed and the remaining data is mostly ignored, however troubleshooting issues and taking into account the SYN/ACK messages and traffic directions, will change according to the configuration.
   Essentially, the admin who sets up the SPAN port has to decide if he wants to copy traffic in to the SPAN port, out of the SPAN port, or in and out of the SPAN port. If the decision is made to copy in and out of the SPAN port, duplicate packets will appear when intra-switch traffic is carried.
6. SPAN port drops packets that are corrupt or do not meet a minimum size. So all frames won't be passed to the end device. This means especially for the NetQOS devices providing results on the TCP/IP packets, that these “discarded” informations cannot be analysed, hence are hidden from our Monitoring (these are usually used in congestion analysis).
   The CEM/TIM from APM has less issues with that, but one needs to be aware of this for troubleshooting purposes.

7. RSPAN is handled in the way that the source switch is taking frames from the source port, encapsulates these frames in newly created RSPAN frames. The switch will then send these RSPAN frames in a Inter Switch Link (ISL) through the production network to a destination switch which then removes the encapsulating headers and presents the frames to a capture device on a destination Switch SPAN output port. From a simple frame timing and performance Monitoring perspective, what appears at the last destination will not look like what actually occurred at the source port or on the network being monitored. This is an enhanced version of point 6 !
8. If the RSPAN frame is now "re-aligned" with data from the local switch into a common SPAN port, we are mixing "alien" RSPAN Frames where the timing information is already very inaccurate, with the SPAN Frames from the current switch with all limitations as stated before. This daisy chaining is inducing an even stronger jitter (timing misalignment, missing information etc.) than a normal SPAN would be capable of. RSPAN is the same as a SPAN Port, but adds more delay, more loss, more filtering and more data modification (Timings). This is just unreliable for any type of serious monitoring or analysis technique.
9. RSPAN configuration very often induce double, triple traffic loops, in a way that "click" interfaces make it easy to configure something without thinking. On average - every RSPAN implementation will show at least one VLan, or RSPAN port to show up twice on the SPAN output port.

To make long explanations short. Using Switch SPAN port technology to analyse data can be done if the technician implementing it knows the network, what maximum traffic it can ask the Switch to actually SPAN.
However, if these factors are unknown – a Network TAP is definitely the technology to use.
A SPAN port will perform well on low-utilized networks, or if analysis is not affected by dropped packets !

Best practice in setting up a SPAN Port

- Clearly define which traffic needs to be SPAN'd. Place the switch that will do the SPAN Session as near to the device (that needs to be Monitored) as possible.
- Make sure the output SPAN port will not be over-subscribed (which will lead to packet loss).
- Make sure the Switch providing the output SPAN port session has not a high CPU load already.
- Definitely avoid using RSPAN.
- Avoid using SPAN sessions that have been dedicated for IDS, Opnet type systems, as this tends to provide duplicate traffic (usually configured to SPAN in and out traffic as seen from the switch).
- In case you need to forward traffic from one Switch to another – use VLAN techniques to forward some specific traffic from one Switch to another switch (TCP/IP is used here, and the VLAN is handled as real traffic, hence won't be discarded in case of traffic overload), and configure the output SPAN session on the last Switch only !
- Make sure all session information is forwarded in any setup.

Implications for CA equipments (CEM from APM, ADA/MTP Collector).

The main issue of SPAN/RSPAN forwarded traffic to the TIM's becomes nowadays the packet misalignment.

A SPAN configuration on one Switch can already delay traffic (all headers need to be modified), realign traffic (put Tx and Rx into the Tx of the SPAN Output port), drop packets due to congestions (using 20 input ports to feed one SPAN output port of 1GBps, and the maximum traffic induced by these 20 full duplex GB (=40Gps) ports is higher than 1Gbps !).

Packet misalignment is happening anywhere in the path from the client to the user, influenced by the overall traffic volume existing in all routing/switching devices in that path. The higher the traffic volume a switch has to handle, the higher the burden placed to a Switch CPU, the more QoS (Quality of Service) traffic is flowing on the network, the more packets will be misaligned. The regular traffic on the internet is growing steadily – and the packet misalignment increases !

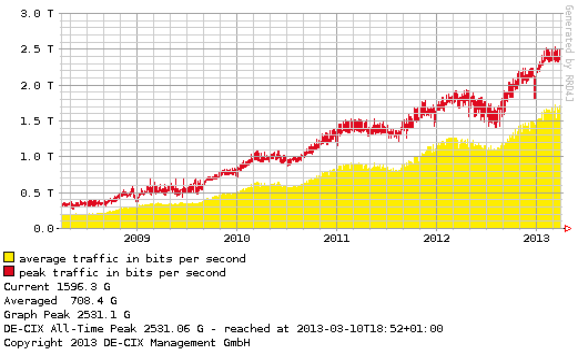

Fig. 7 DE-CIX Peering point, 5 year period.

Figure 7 shows the increase of traffic in the last years (also called the iPhone/Siri effect). All this increased traffic lead to more packet misalignment.

On the packet level analysis, this will not be an issue – even though one can only analyse what the Switch really forwards – and in this special case, one has to ask himself if the Switch is really forwarding all the data it sees.
However when analysing data on the application layer (like HTTP protocol, extract the transported HTML code interpreted by browsers), one needs to have the entire data flow from A to Z for one particular HTML page*. This also means the TIM needs to see the 3Way Handshake of a communication initialization, and the end tag (FIN) to know the connection is closed.
As it is quite common that traffic is not reaching its destination in the right order, the TIM will allocate buffers to keep the misaligned packets in cache until all data is available, realign it, and extract the required information for analysis out of the packets payload.
For one communication flow (client Web Browser to one Server and back), the amount of required memory is quite small, as every OS handles this on different levels. The TIM however handles that realignment for many 100 or 1000 connections at the same time. The required memory to buffer the misaligned packets starts to become an issue and due to Cache size and time constrains need to free buffer space for the next incoming misaligned data. Usually these will be seen as “Out of Order Queued Bytes” in the Tim logs.

** There is an option that tells TIM to try to synchronize with data in the middle of a TCP connection. It is not recommended because TIM may match a request with the wrong response if HTTP pipelining is in use. Also, this option only works if the request and response start at the beginning of a packet. But it may be useful for some B2B applications that keep the connection between the client and the server open for a long time. It doesn't work for HTTPS because TIM needs to see the SSL handshake.*
To enable it, define a setting called MonitorAlreadyOpenedConnections and set it to 1.

In addition to the memory used, managing out-of-order data seems to be slow, causing TIM to spend more time handling out-of-order data, which makes it more likely that the packet buffer will fill up and the OS will drop packets. This becomes a vicious cycle because then there are missing packets that TIM will never see, causing it to spend even more time managing the out-of-order queues.

One other place the packet misalignment has proven to make issues is SSL data decryption. Clean traffic (aligned, and only the traffic the TIM requires) prevents bad surprises. One missing packet can invalidate an entire data-stream required to look into one html-page !

Information/data Sources

- Various CEM Engagements, old experience from work at international ISP
- Wikipedia
- Network – discussion forums around networking excluding Hardware Manufacturers publicity ! (the real stuff)
  SPAN Port or TAP ? CSO Beware (by Tim O' Neil)
  RSPAN ... Friend or Foe ? (by Tim O' Neil)
- Network Instruments TAP vs. SPAN
- SPAN Port or TAP by Gigamon
- Multi TAP Network Packet Capturing from networksecurity Wiki

Credits & Acknowledgements

Thanks to Hallett German, Gary Jones, Rob Webb and Steven Tepper for reviewing the document and making suggestions on things to add/change.

Joerg Mertin | Sr. Engineering Services Architect | CA Technologies
Solutions SWAT | joerg.mertin@ca.com