vehere

IMPACT ON METADATA SIZING WITH TRAFFIC AGGREGATION

Content

1. Introduction

This document explores the concept of traffic aggregation, delving into its scenarios and impacts. We will examine how traffic aggregation techniques can affect system performance and the sizing of overall solution components. Furthermore, it will address key challenges associated with traffic aggregation and provide best practices for successful implementation.

We have used a 10G link as a reference in this document; however, the concept of aggregation remains the same whether it is a 10G, 100G, or 400G link.

2. 10G Full-Duplex Link Throughput in Transmission Network

Theoretically, a 10G full duplex link has a throughput of 10 Gbps in each direction, resulting in a total capacity of 20 Gbps.

However, in real-world scenarios, the actual throughput of 10Gbps duplex link will always have less than 20Gbps due to the following reasons:

- In a typical client-server interaction, the client's communication with the server is minimal, focusing primarily on issuing queries or requests. The server, in contrast, often responds with comprehensive data or results, reflecting the more substantial payload that the server needs to convey back to the client.
- In network design and operation, a critical principle is to avoid configuring any network link to utilize more than 80% of its total capacity due to the following fundamental reasons:

    - **Performance and Latency:** When a network link approaches its maximum capacity, the risk of performance degradation increases. High utilization levels can lead to congestion, causing increased latency and jitter, which negatively impact the quality of service. By keeping utilization below 80%, network operators create a buffer that helps maintain smooth and responsive network performance.

    - **Avoiding Congestion:** Network links operating near full capacity are more susceptible to congestion during peak usage times or when additional traffic is introduced. This can result in packet loss, retransmissions, and reduced throughput. Keeping utilization under 80% helps mitigate these risks and ensures that the network can handle unexpected spikes in traffic without significant issues.

    - **Error Handling and Resilience:** Lower utilization levels provide room for handling errors and network anomalies. It allows the network to manage retransmissions and error recovery processes more efficiently, contributing to overall network reliability and resilience.

    - **Quality of Service (QoS):** For networks providing critical services, such as real-time applications or mission-critical operations, adhering to this principle ensures that QoS requirements are consistently met, delivering a better experience for end-users.

3. Metadata and Traffic Patterns in Standard 10G Duplex Links

Consider we have two 10G links (Link-1 and Link-2 as per the given diagram, Fig 1), each link is carrying traffic of 15Gbps in both directions (A to B traffic is 6 Gbps and B to A is 9 Gbps).

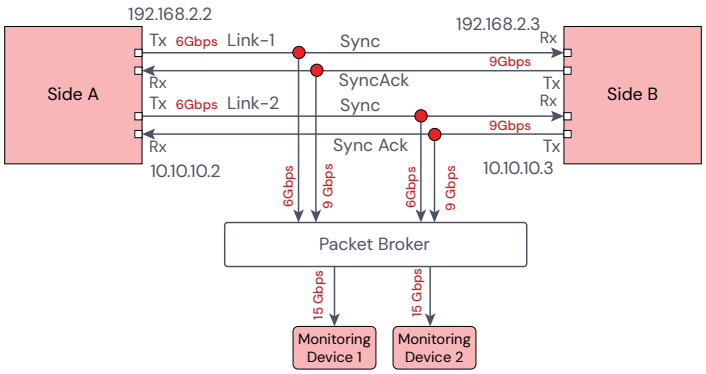

Fig 1. Aggregation of Multiple links

With the above-mentioned scenario, each TCP session establishes a connection between a client and server, characterized by unique attributes such as session identifiers, source and destination IP addresses, and port numbers. Metadata is created based on these unique TCP sessions, capturing relevant details about each connection as metadata of that TCP session.

Hence, despite the high throughput of 15 Gbps on each 10G duplex link, the total number of metadata sessions can still be considered optimal as it reflects the number of unique TCP sessions.

Traffic Pattern
In a 10G full-duplex link, different protocols and applications distribute traffic according to their specific needs and characteristics. TCP and UDP are the primary protocols, each with distinct usage patterns and effects on bandwidth utilization. Applications such as web browsing, file transfers, streaming, and VoIP all leverage the high-speed capabilities of the link in different ways, necessitating careful management to maintain optimal performance and avoid congestion.

4. Simplex Vs Full-Duplex

In simplex mode, data transmission occurs only in one direction. There is no provision for the receiver to send data back to the sender over the same channel. Simplex links will be used for monitoring purpose as monitoring device only receives the data, but does not respond back.

In full-duplex communication, data can be sent and received simultaneously over the same communication channel. Both ends of the communication link can transmit and receive data at the same time. All the network traffic in real communication will be used in full-duplex mode as the data travels in both directions simultaneously.

With the above understanding, the throughput of the monitoring system will always be considered in simpler terms, as the monitoring system only receives data and does not send any data back. Additionally, if the user wishes to aggregate multiple links, they should explicitly consider the aggregation of low-utilized links to avoid any impact on performance and capacity.

5. Different Types of Aggregations and its Impact

5.1. Aggregating one side of multiple 10G duplex links

When aggregating simplex links of each 10G duplex link using a packet broker, it is common to see the number of meta sessions double, even though the overall bandwidth remains within the 10G full-duplex limit, the aggregation process involves handling multiple logical connections, hence the apparent doubling of meta sessions.

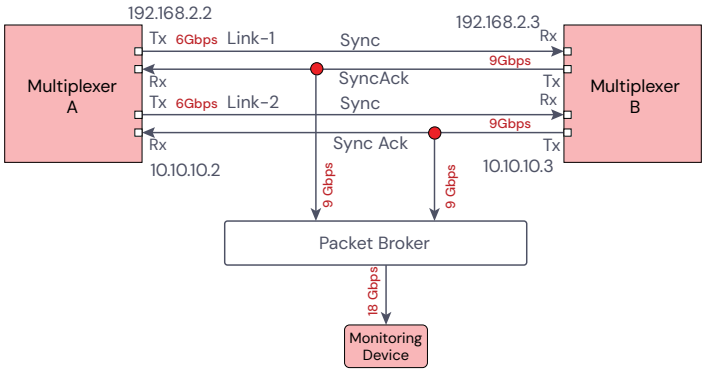

Fig 2. Aggregation of two simplex links

As the number of unique sessions/ metadata increases, it will have the following impact on the monitoring system-

- **Reduced performance:** The overall performance of monitoring device server gets impacted as it required a greater number of resources to handle increased metadata at monitoring device level and logstash level.
- **Impact on metadata storage:** As the unique TCP sessions increased, respective metadata index size also increased. Hence, its storge gets doubled although the overall input bandwidth/throughput remains equivalent to 10G link.
- **Impact on logstash/messaging layer:** Since the metadata is doubling compared to the actual 10G link, this will result in a higher number of metadata files. Consequently, there will be an impact on logstash performance due to the increased number of metadata files it needs to handle.

5.2. Aggregation of many links to one links with overflow

Aggregating multiple links into one to increase the throughput of the input traffic for monitoring, as shown in the figure 3 below, can sometimes result in traffic exceeding the capacity of the output link consequently, some traffic may be dropped at the packet broker level.

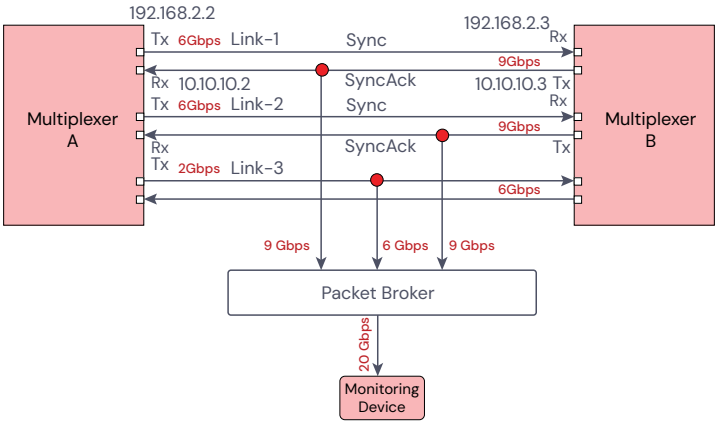

Fig 3. Aggregation multiple links to one output link

In the above-mentioned scenarios, it will have the following impact on the solution-

- **Impact on Logstash/Messaging Layer:** Since the metadata is doubling compared to the actual 10G link, this will result in a higher number of metadata files. Consequently, there will be an impact on logstash performance due to the increased number of metadata files it needs to handle.

- **Impact on Metadata Storage:** As the unique TCP sessions increased, the respective metadata index size increased. Hence, its storage got doubled although the overall input bandwidth/throughput remained equivalent to 10G link.

- **Impact on IP Decoding/Reconstruction Capability :** As there would be packet drops of the session, it may lead to incomplete reconstruction or may have errors in reconstruction of cleartext data.

5.3. Aggregating multiple links based unencrypted application or protocols

Each network link carries a mix of different application protocols. All links connect to a packet broker, which aggregates only clear-text protocols or routes all VoIP calls through a single 10G link by filtering out other unwanted protocols.

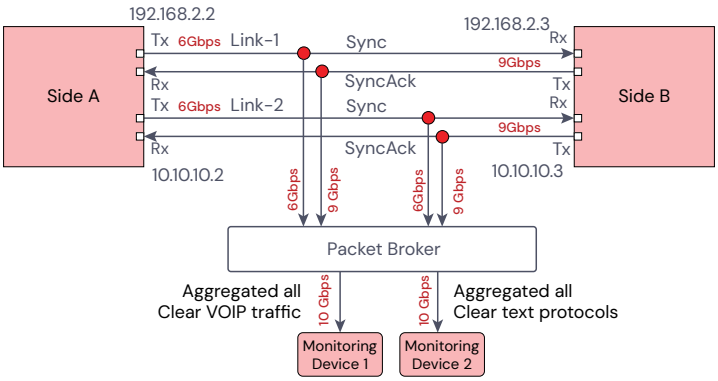

Fig 4. Aggregation specific traffic based on application/protocol

In the above-mentioned aggregation scenario, if all the VoIP traffic or clear-text traffic is directed to a specific monitoring device, it will have the following impacts on the solution:

- **Impact on IP decoding/reconstruction capability:** As large number of packets/VoIP sessions have to be processed, the decoder might struggle to keep up, which can result in delayed or dropped packets. Additionally, an increased load on the reconstruction engine can lead to decoding errors, corrupting packet data and rendering it useless for analysis.

- **Impact on raw storage of the monitoring device:** As the raw storage is being stored at monitoring device, if specific monitoring device receives only clear text or VoIP traffic, then raw storage of the monitoring device may get utilized faster and can lead to shortage in raw storage.

- **Impact on metadata size:** As we are also extracting all the textual content and indexing as part of metadata for making it searchable, having a complete clear text data can lead to increase in metadata which results into shortage of storage retention period.

5.4. Aggregation of multiple low utilized links

If each duplex link is underutilized compared to its actual capacity, we can connect all those under-utilized links to the packet broker. This setup would aggregate their throughput to 10G without filtering any unwanted traffic or modifying its traffic distribution.

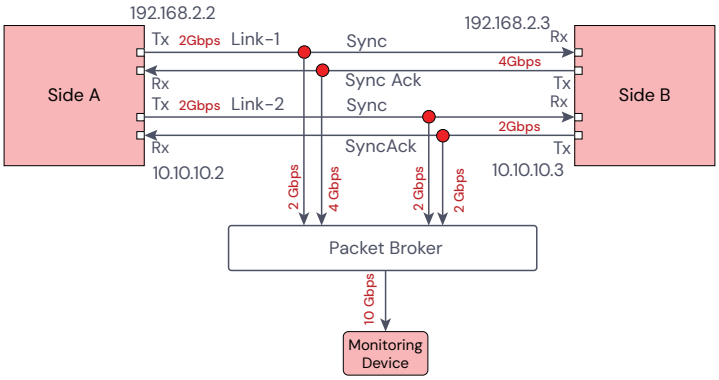

Fig 5. Aggregation of multiple low utilized links

In this scenario, there would not be any impact on the solution performance and its sizing as the overall traffic and its uniqueness of metadata still being maintained as per actual network 10G link and its throughput.

6. Conclusion

Based on the analysis presented in sections 5.1, 5.2 and 5.3, aggregating traffic from multiple links is not recommended due to the significant performance challenges it introduces. Aggregation using multiple 10G/100G/400G links should only be considered when actual link utilization is significantly lower than the available capacity.

The summary of aggregation types and its impact is given in the following table.

S.No	Aggregation Type	Impact
1	Aggregating one side of multiple 10G duplex links	▪ Reduced monitoring device and logstash performance. ▪ Impact on metadata storage
2	Aggregation of many links to one/ few links with overflow	▪ Reduced monitoring device and logstash performance. ▪ Impact on metadata storage ▪ Impact on IP decoding/reconstruction capability
3	Aggregating multiple links based application or protocols	▪ Impact on IP decoding/reconstruction capability ▪ Impact on raw storage of the monitoring device ▪ Impact on metadata storage
4	Aggregation of multiple low utilized links	No impact on the system performance