Measurement Artifacts in NetFlow Data

Rick Hofstede, Idilio Drago,
Anna Sperotto, Ramin Sadre, Aiko Pras

UNIVERSITY OF TWENTE.

Introduction

- Flow monitoring technologies are a scalable alternative to packet-based solutions and have been widely deployed
- Analysis always depends on the quality of the input data
- Can we trust our NetFlow data?
- How widespread are artifacts in NetFlow data?

Introduction

Trammell *et al.*: Peeling Away Timing Error in NetFlow Data
In: Proceedings of the 12th International Conference on Passive and Active Network Measurement (PAM 2011)

Kögel *et al.*: One-way Delay Measurement based on Flow Data: Quantification and Compensation of Errors by Exporter Profiling
In: Proceedings of the 25th International Conference on Information Networking (ICOIN 2011)

Cunha *et al.*: Uncovering Artifacts of Flow Measurement Tools
In: Proceedings of the 10th International Conference on Passive and Active Network Measurement (PAM 2009)

Case Study

- Our operational experience has been gained over the years:
  Cisco Catalyst 6500 (SUP720-3B)

- We discuss five artifacts (non-comprehensive)

Case Study

1. Imprecise flow record expiration

Case Study

1. Imprecise flow record expiration
2. TCP flows without flag information

Case Study

1. Imprecise flow record expiration
2. TCP flows without flag information
3. Invalid byte counters

Case Study

1. Imprecise flow record expiration
2. TCP flows without flag information
3. Invalid byte counters
4. Non-TCP flow records with TCP ACK flag set

Case Study

1. Imprecise flow record expiration
2. TCP flows without flag information
3. Invalid byte counters
4. Non-TCP flow records with TCP ACK flag set
5. Gaps

Experiment Setup

- Are the identified artifacts also present in flow data from other flow exporters?
- Can the artifacts be identified in flow data without having access to exporter statistics?

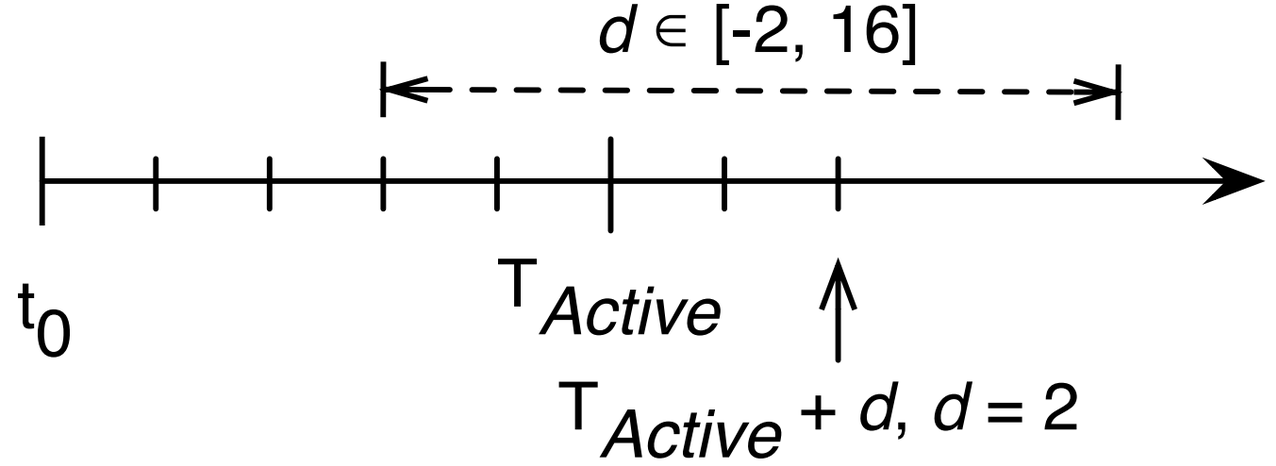
Experiment Setup

No.	Model	Modules	Software version
1.	Cisco Catalyst 6500	WS-SUP720-3B (PFC3B, MSFC3)	IOS 12.2(33)SXI5
2.	Cisco Catalyst 6500	WS-SUP720-3B (PFC3B, MSFC3)	IOS 12.2(33)SXI2a
3.	Cisco Catalyst 6500	VS-SUP2T-10G-XL (PFC4XL, MSFC5) + WS-X6904-40G	IOS 15.0(1)SY1
4.	Cisco Catalyst 7600	RSP720-3C-GE (PFC3C, MSFC4)	IOS 15.2(1)S
5.	Juniper T1600	MultiServices PIC 500	JUNOS 10.4R8.5
6.	INVEA-TECH FlowMon	-	3.01.02

Three vendors, wide range of Cisco models...

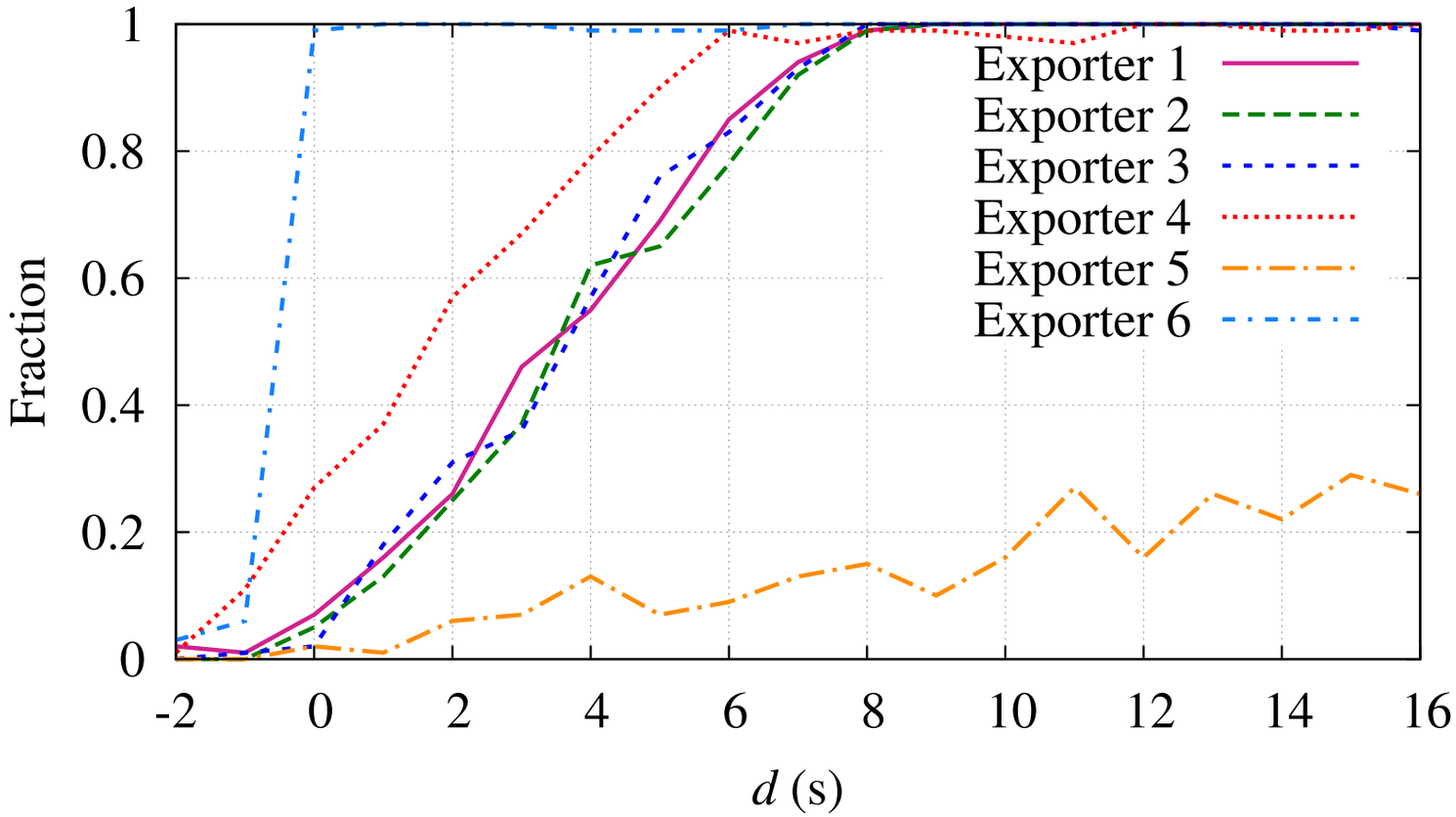
Artifact Analysis

Imprecise flow record expiration



Artifact Analysis

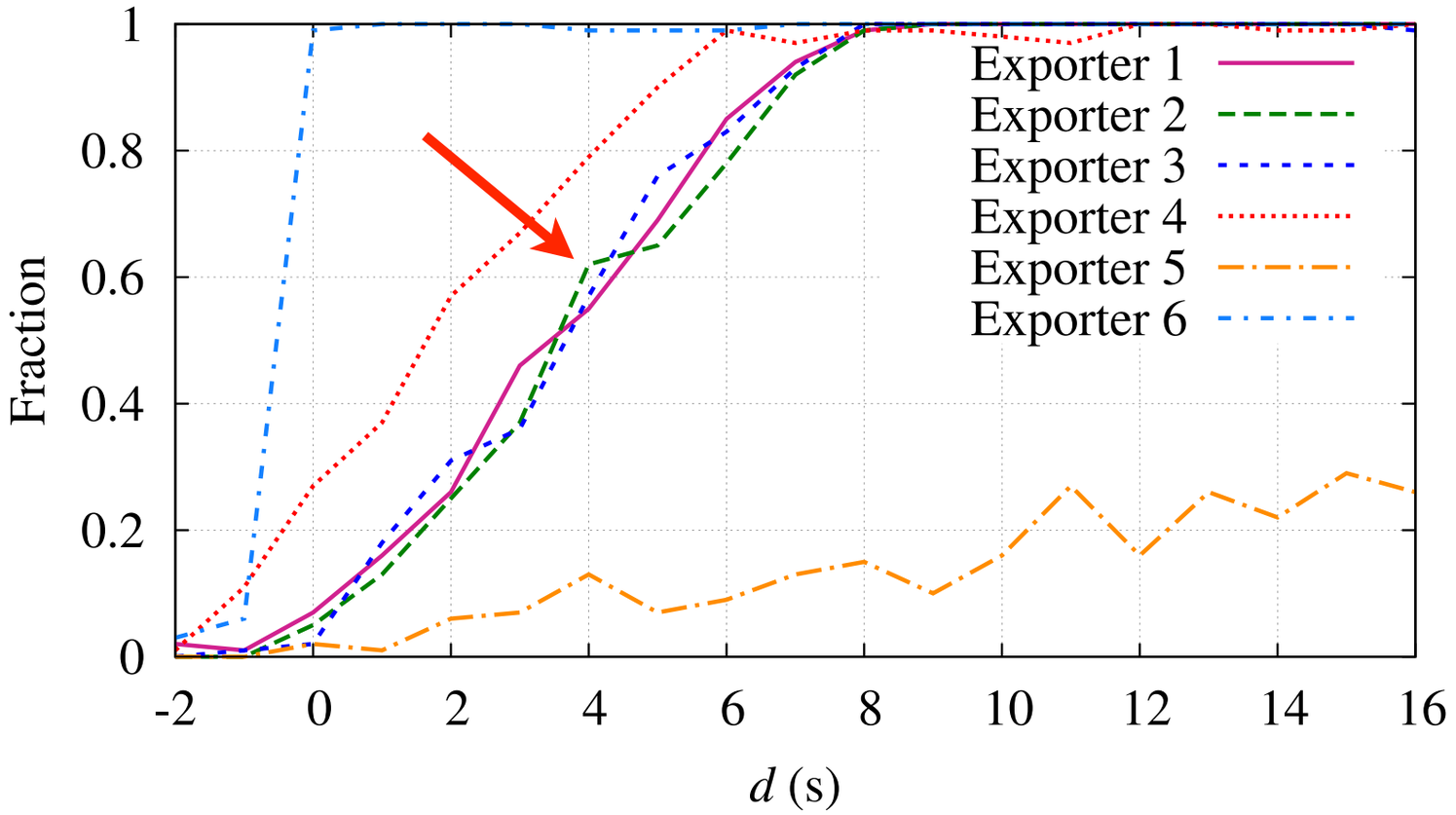
Imprecise flow record expiration



- Exporter 1-4 don't expire flow records according to Cisco documentation
- Exporter 5 shows incorrect flow record starting times and does not stabilize

Artifact Analysis

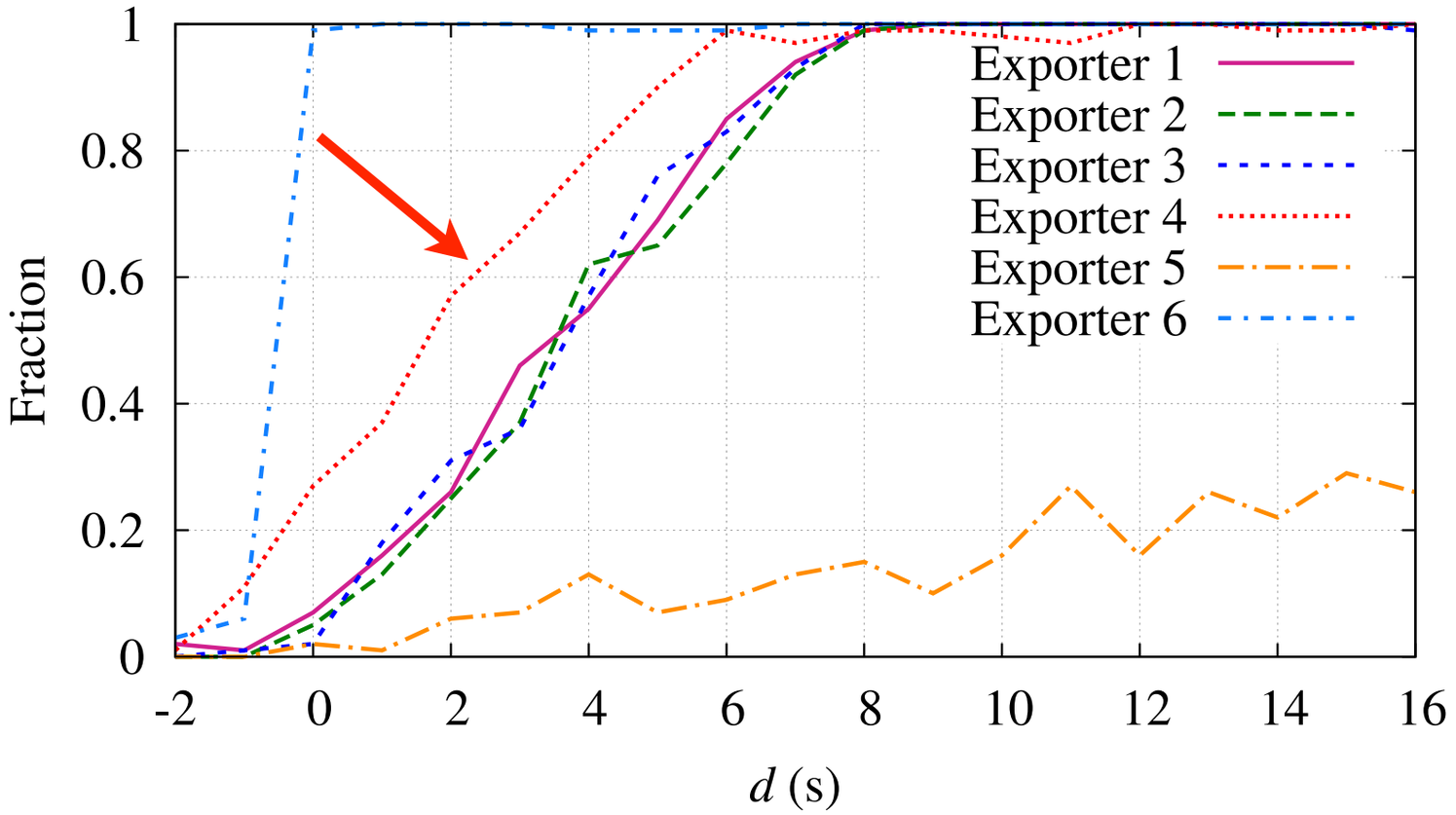
Imprecise flow record expiration



- Exporter 1-4 don't expire flow records according to Cisco documentation
- Exporter 5 shows incorrect flow record starting times and does not stabilize

Artifact Analysis

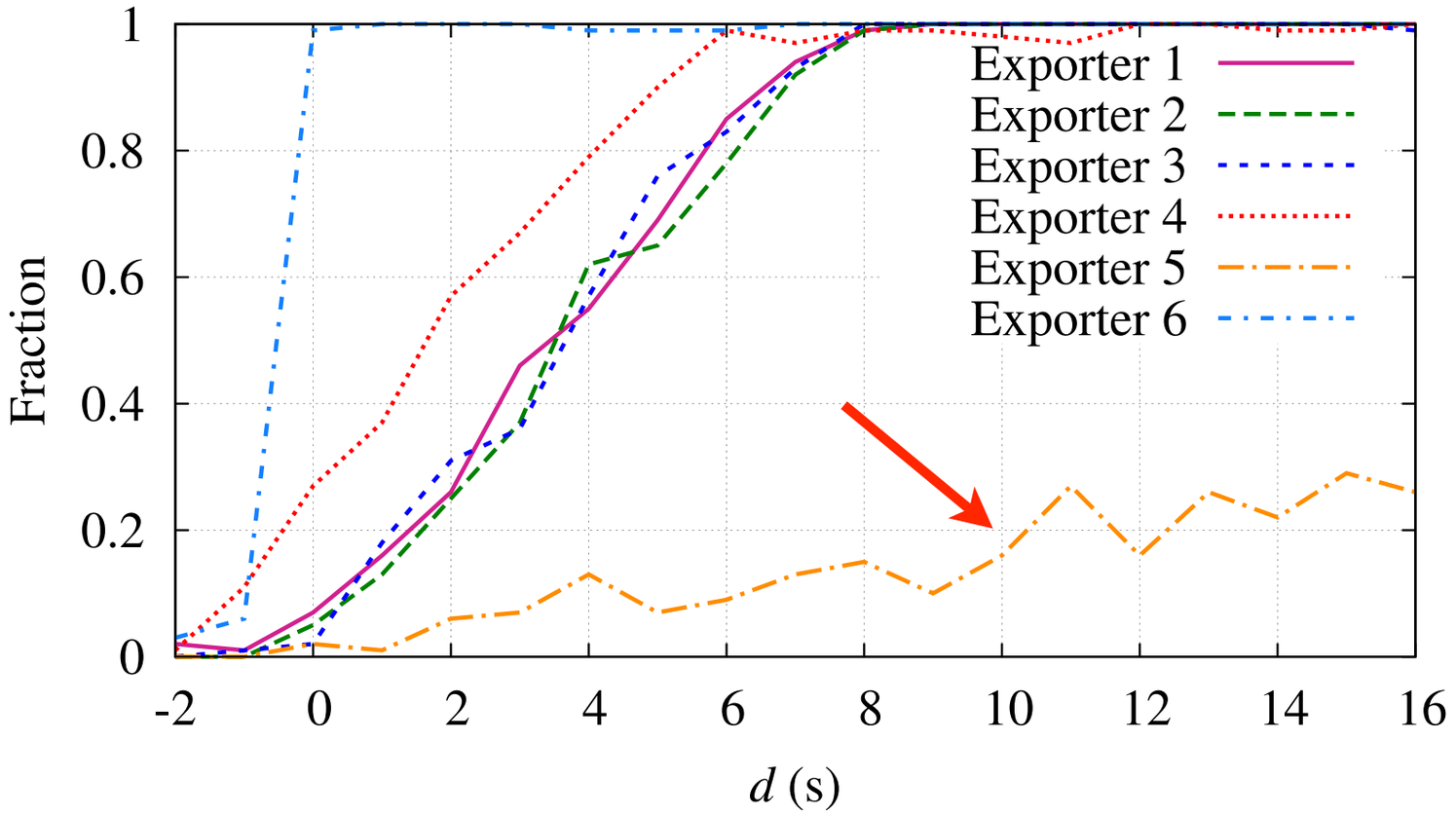
Imprecise flow record expiration



- Exporter 1-4 don't expire flow records according to Cisco documentation
- Exporter 5 shows incorrect flow record starting times and does not stabilize

Artifact Analysis

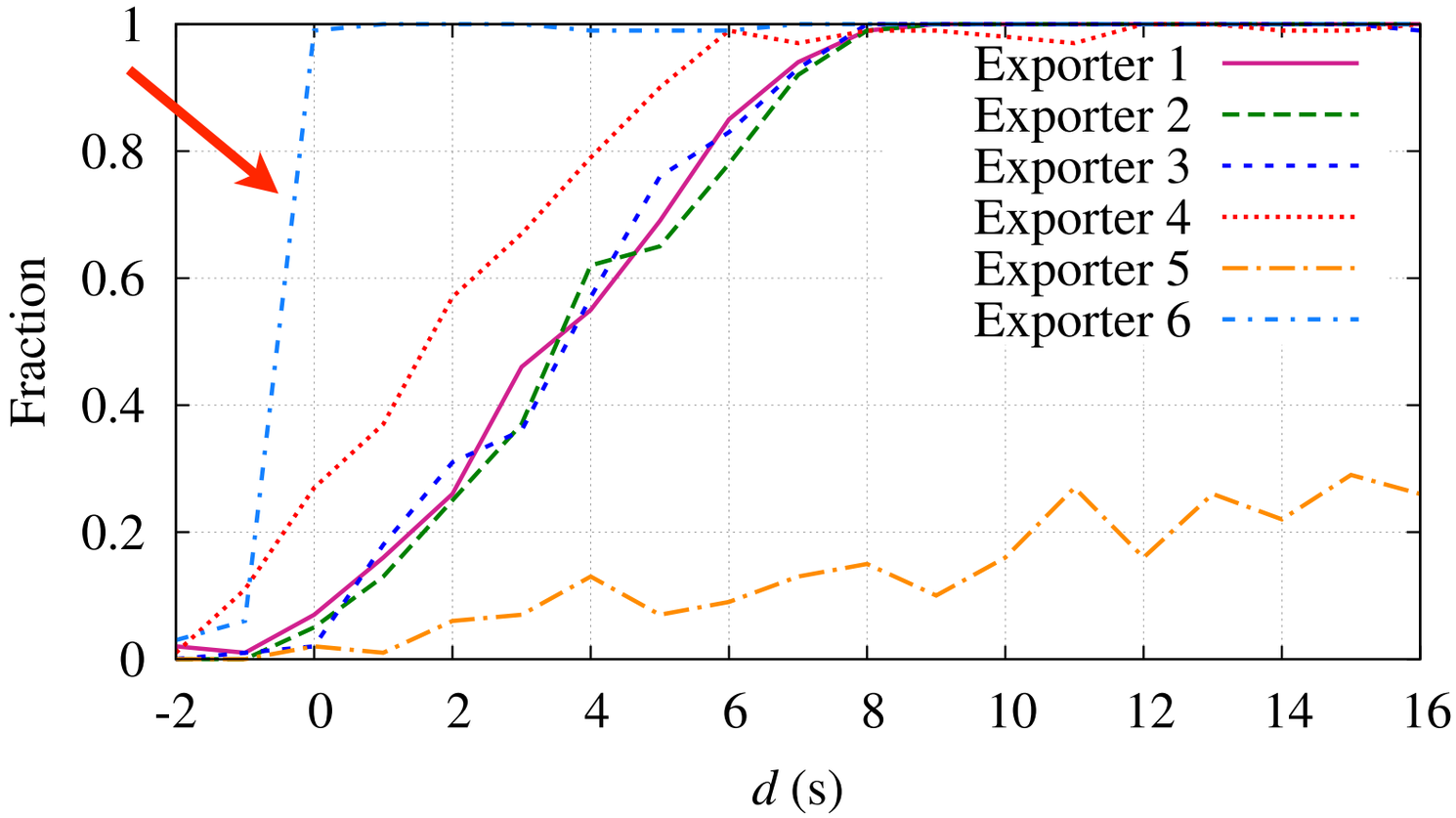
Imprecise flow record expiration



- Exporter 1-4 don't expire flow records according to Cisco documentation
- Exporter 5 shows incorrect flow record starting times and does not stabilize

Artifact Analysis

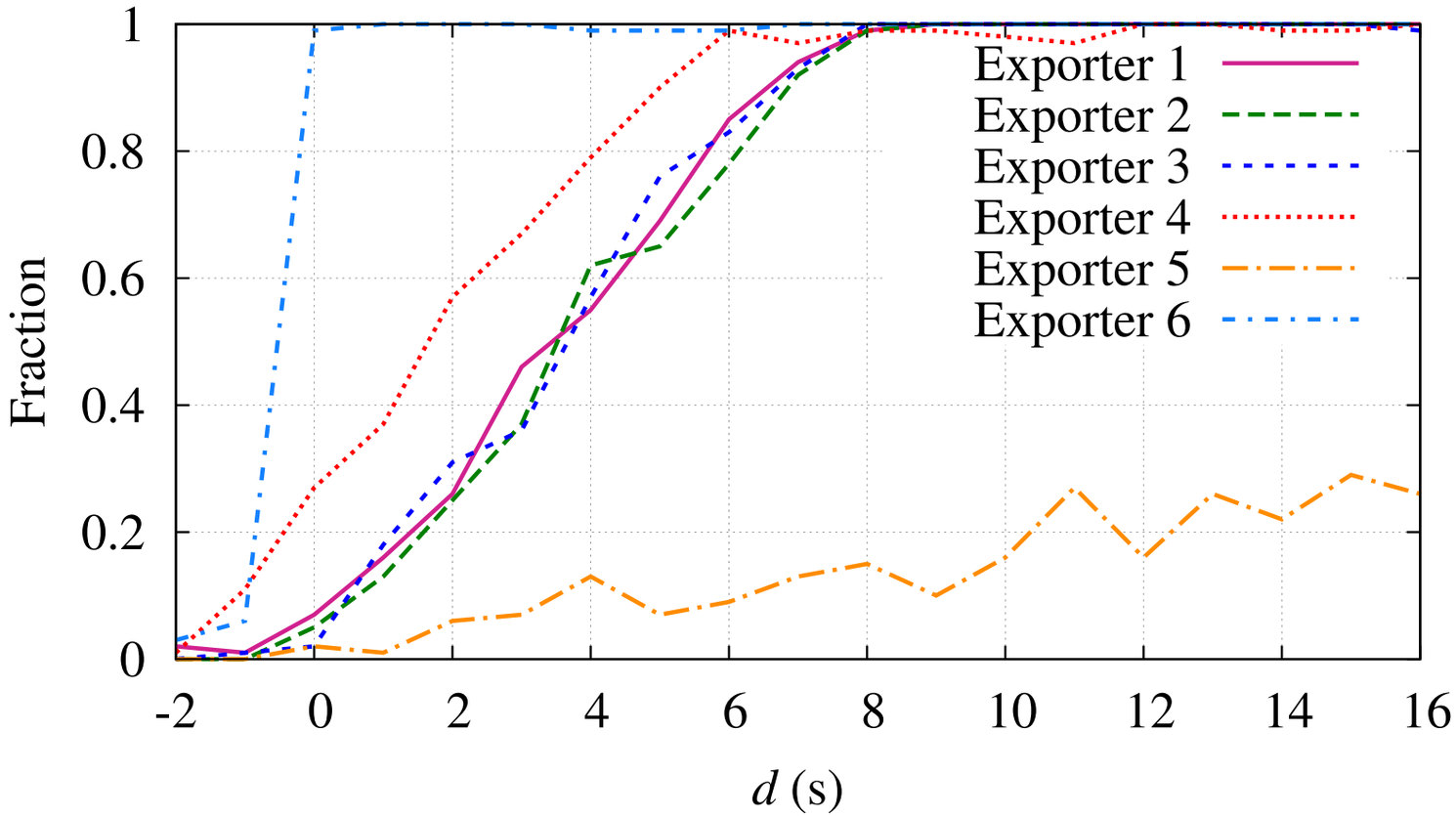
Imprecise flow record expiration



- Exporter 1-4 don't expire flow records according to Cisco documentation
- Exporter 5 shows incorrect flow record starting times and does not stabilize

Artifact Analysis

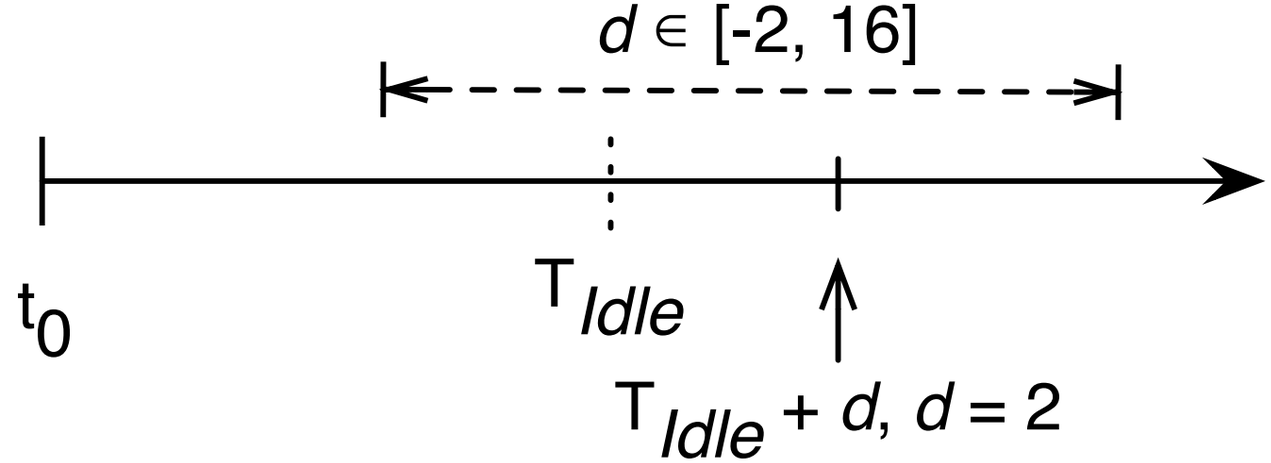
Imprecise flow record expiration



- Exporter 1-4 don't expire flow records according to Cisco documentation
- Exporter 5 shows incorrect flow record starting times and does not stabilize

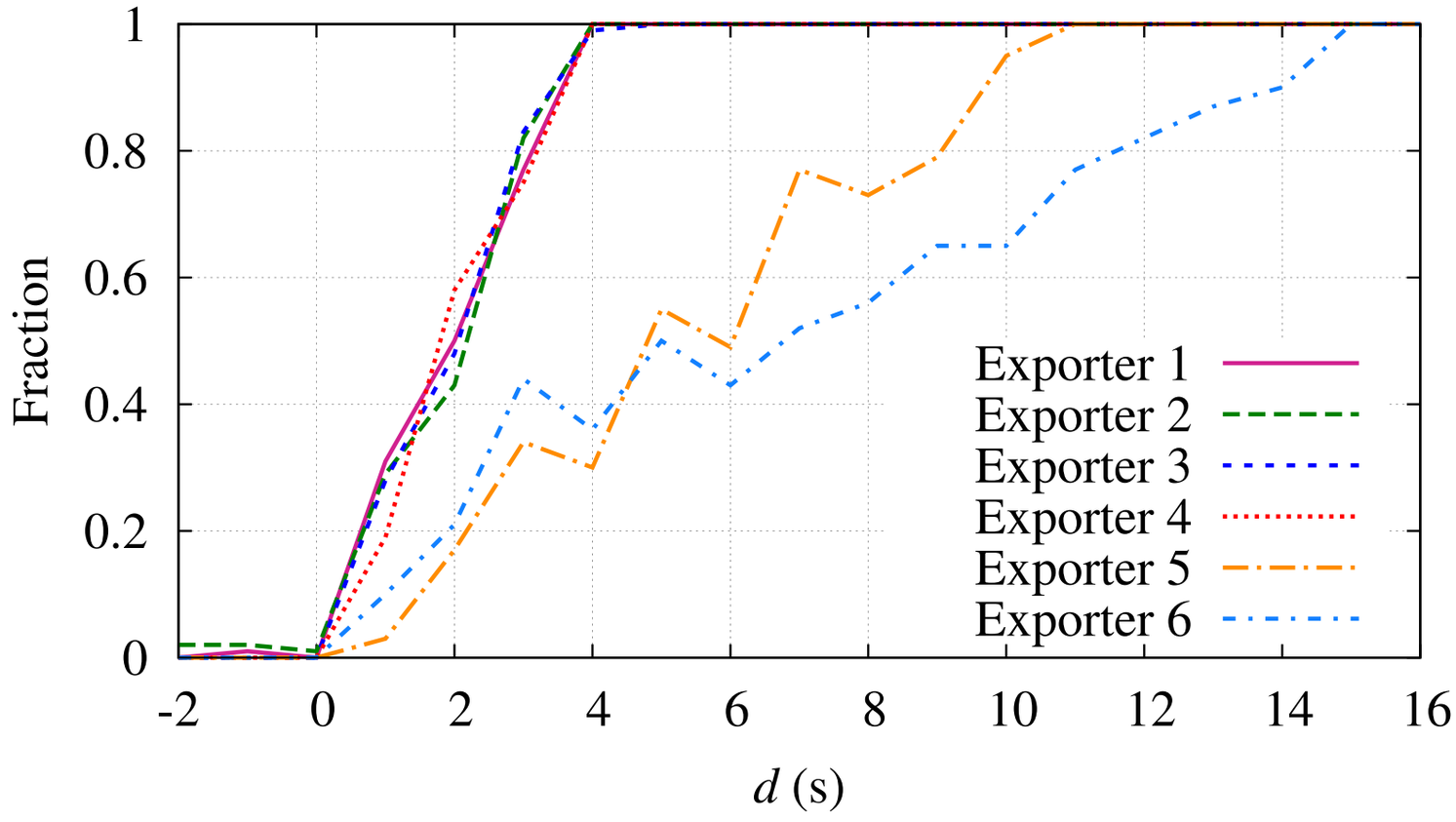
Artifact Analysis

Imprecise flow record expiration



Artifact Analysis

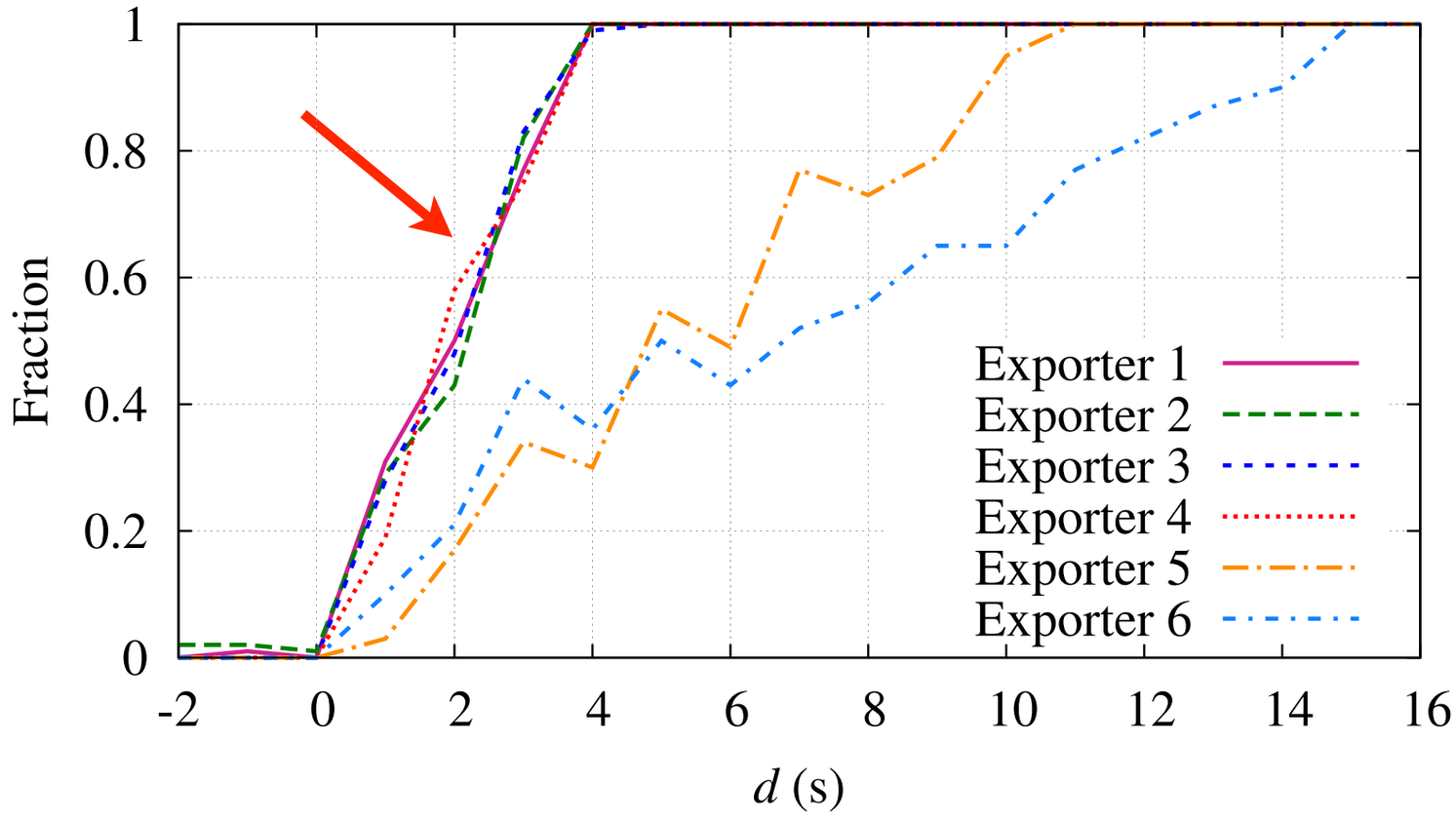
Imprecise flow record expiration



- Exporter 1-4 behave according to Cisco documentation
- Exporter 5 and 6 expire flow record up to 11 and 15 seconds after timeout, respectively -- depends on absolute timeout value

Artifact Analysis

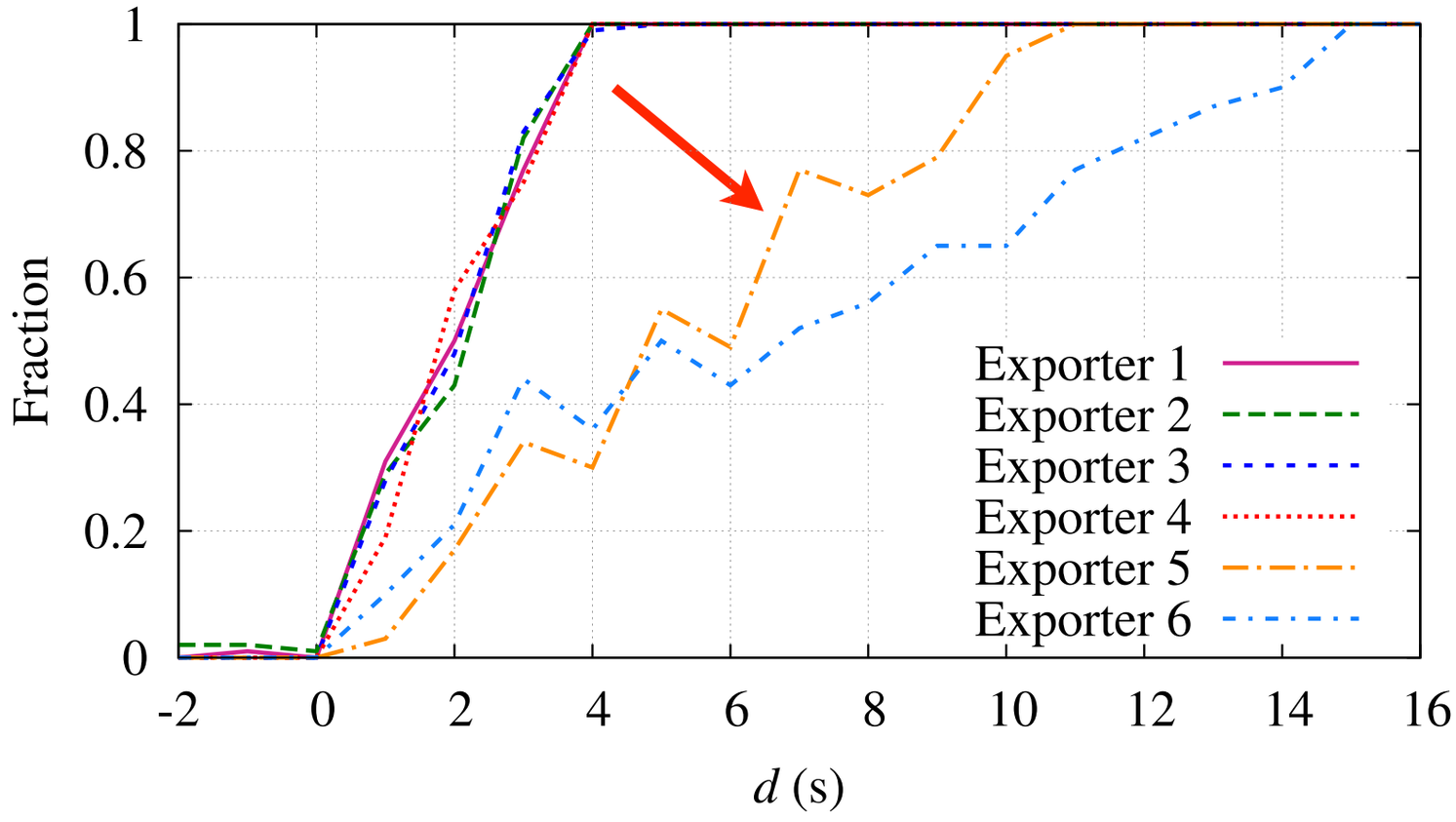
Imprecise flow record expiration



- Exporter 1-4 behave according to Cisco documentation
- Exporter 5 and 6 expire flow record up to 11 and 15 seconds after timeout, respectively -- depends on absolute timeout value

Artifact Analysis

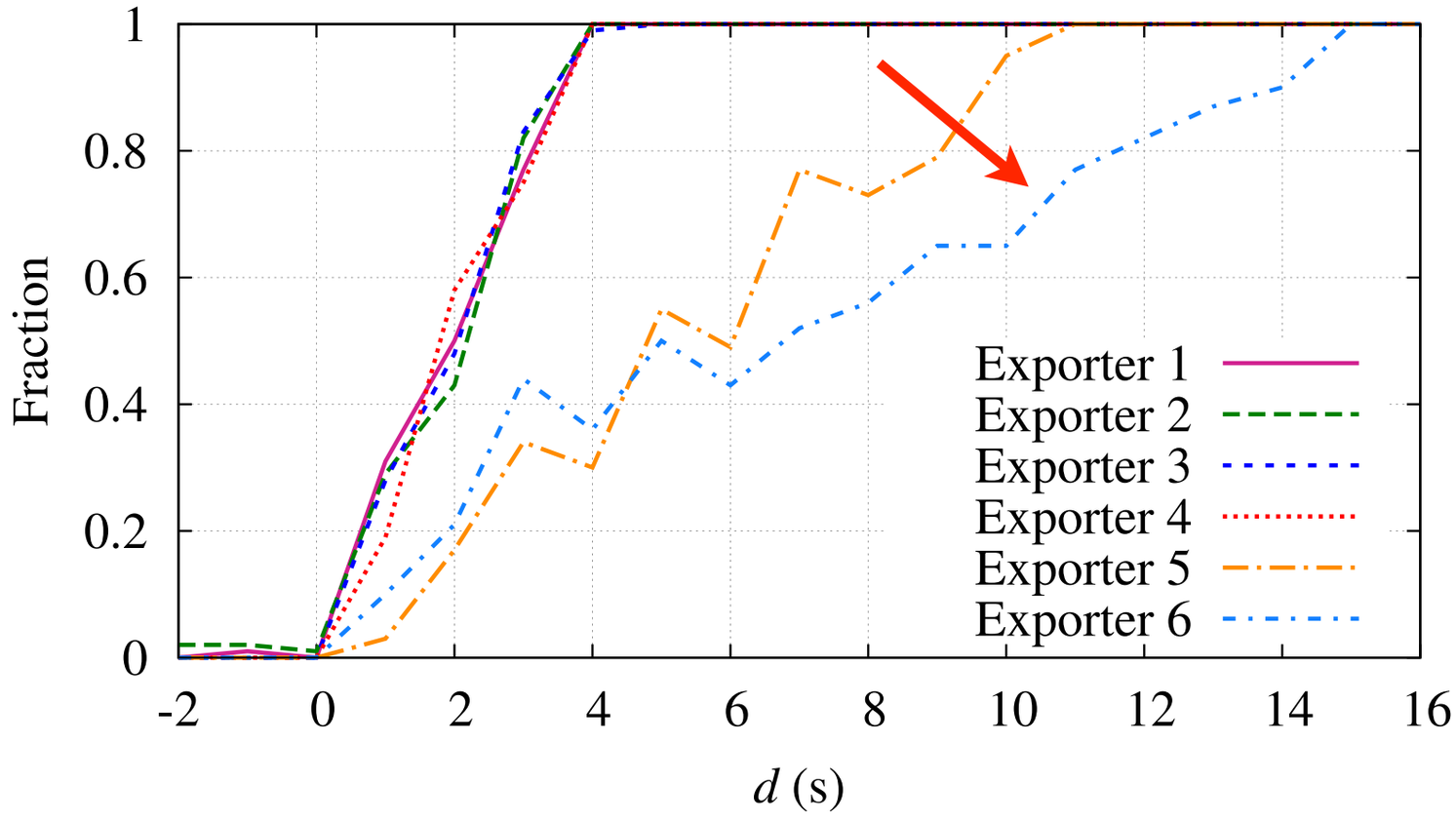
Imprecise flow record expiration



- Exporter 1-4 behave according to Cisco documentation
- Exporter 5 and 6 expire flow record up to 11 and 15 seconds after timeout, respectively -- depends on absolute timeout value

Artifact Analysis

Imprecise flow record expiration



- Exporter 1-4 behave according to Cisco documentation
- Exporter 5 and 6 expire flow record up to 11 and 15 seconds after timeout, respectively -- depends on absolute timeout value

Artifact Analysis

TCP flows without flag information

No.	Model	Modules	Software version
1.	Cisco Catalyst 6500	WS-SUP720-3B (PFC3B, MSFC3)	IOS 12.2(33)SXI5
2.	Cisco Catalyst 6500	WS-SUP720-3B (PFC3B, MSFC3)	IOS 12.2(33)SXI2a
3.	Cisco Catalyst 6500	VS-SUP2T-10G-XL (PFC4XL, MSFC5) + WS-X6904-40G	IOS 15.0(1)SY1
4.	Cisco Catalyst 7600	RSP720-3C-GE (PFC3C, MSFC4)	IOS 15.2(1)S
5.	Juniper T1600	MultiServices PIC 500	JUNOS 10.4R8.5
6.	INVEA-TECH FlowMon	-	3.01.02

- Exporter 1, 2 and 4 don't export (but respect!) TCP flags for hardware-switched flows
- 99.6% of TCP flow records from Exporter 1, 2 and 4 lack TCP flag information

Artifact Analysis

Invalid byte counters

No.	Model	Modules	Software version
1.	Cisco Catalyst 6500	WS-SUP720-3B (PFC3B, MSFC3)	IOS 12.2(33)SXI5
2.	Cisco Catalyst 6500	WS-SUP720-3B (PFC3B, MSFC3)	IOS 12.2(33)SXI2a
3.	Cisco Catalyst 6500	VS-SUP2T-10G-XL (PFC4XL, MSFC5) + WS-X6904-40G	IOS 15.0(1)SY1
4.	Cisco Catalyst 7600	RSP720-3C-GE (PFC3C, MSFC4)	IOS 15.2(1)S
5.	Juniper T1600	MultiServices PIC 500	JUNOS 10.4R8.5
6.	INVEA-TECH FlowMon	-	3.01.02

- Exporter 1-4 export records with invalid byte counters
- 20% of all frames* have less than 46 bytes of payload, which would be reported incorrectly

* Based on traces of 1 day from the UT campus network and CAIDA 'equinix-sanjose'

Artifact Analysis

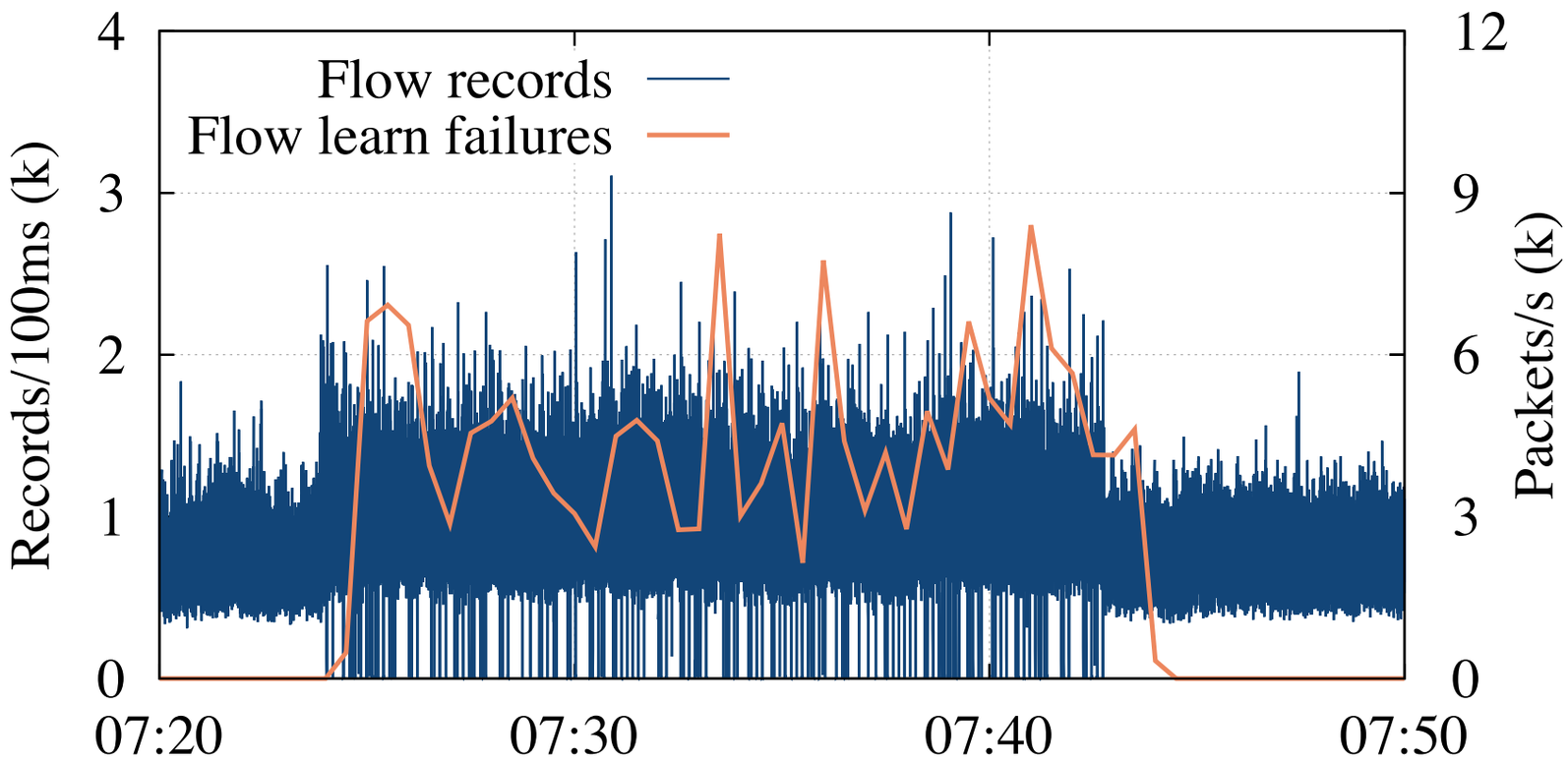
Non-TCP flow records with TCP ACK set

No.	Model	Modules	Software version
1.	Cisco Catalyst 6500	WS-SUP720-3B (PFC3B, MSFC3)	IOS 12.2(33)SXI5
2.	Cisco Catalyst 6500	WS-SUP720-3B (PFC3B, MSFC3)	IOS 12.2(33)SXI2a
3.	Cisco Catalyst 6500	VS-SUP2T-10G-XL (PFC4XL, MSFC5) + WS-X6904-40G	IOS 15.0(1)SY1
4.	Cisco Catalyst 7600	RSP720-3C-GE (PFC3C, MSFC4)	IOS 15.2(1)S
5.	Juniper T1600	MultiServices PIC 500	JUNOS 10.4R8.5
6.	INVEA-TECH FlowMon	-	3.01.02

- Exporter 1, 2 and 4 export flow data containing this artifact
- 1% of flow records is non-TCP with TCP ACK flag information set

Artifact Analysis

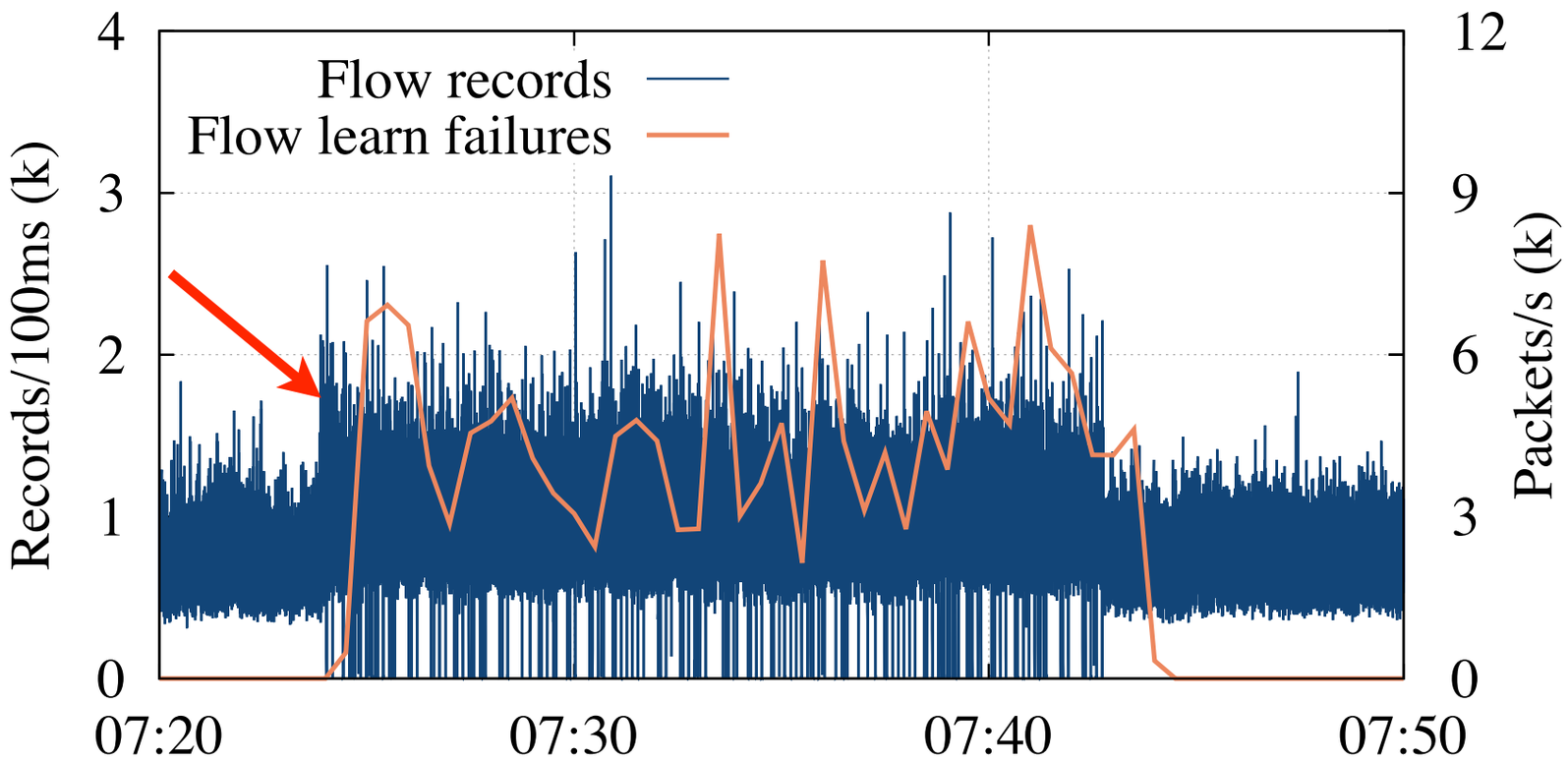
Gaps



- Increase in traffic volume causes ‘flow learn failures’
- Overloaded exporters may introduce more artifacts!

Artifact Analysis

Gaps



- Increase in traffic volume causes ‘flow learn failures’
- Overloaded exporters may introduce more artifacts!

Conclusions

- We have identified several artifacts in NetFlow data from a range of exporters from various vendors
- Some artifacts can be repaired easily, while others adversely impact the data quality
- We believe that flow data applications cannot be designed and developed to be generic and applicable to any flow data
- Future work:
  - Impact of packet sampling on flow data artifacts
  - Data cleanup tool for detecting and repairing artifacts in flow data

Questions?