YAF: Yet Another Flowmeter

Chris Inacio <inacio@cert.org>
Brian Trammell <trammell@tik.ee.ethz.ch>

CERT Software Engineering Institute | Carnegie Mellon®

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Yet Another Flowmeter

- Flowmeter
  - What is flow
  - Why do you want flow
  - So why YAF

flow

- The simple version: a very brief summarization of a network connection
  - The key values
    - IP address source & destination
    - Protocol
    - Transport source & destination port

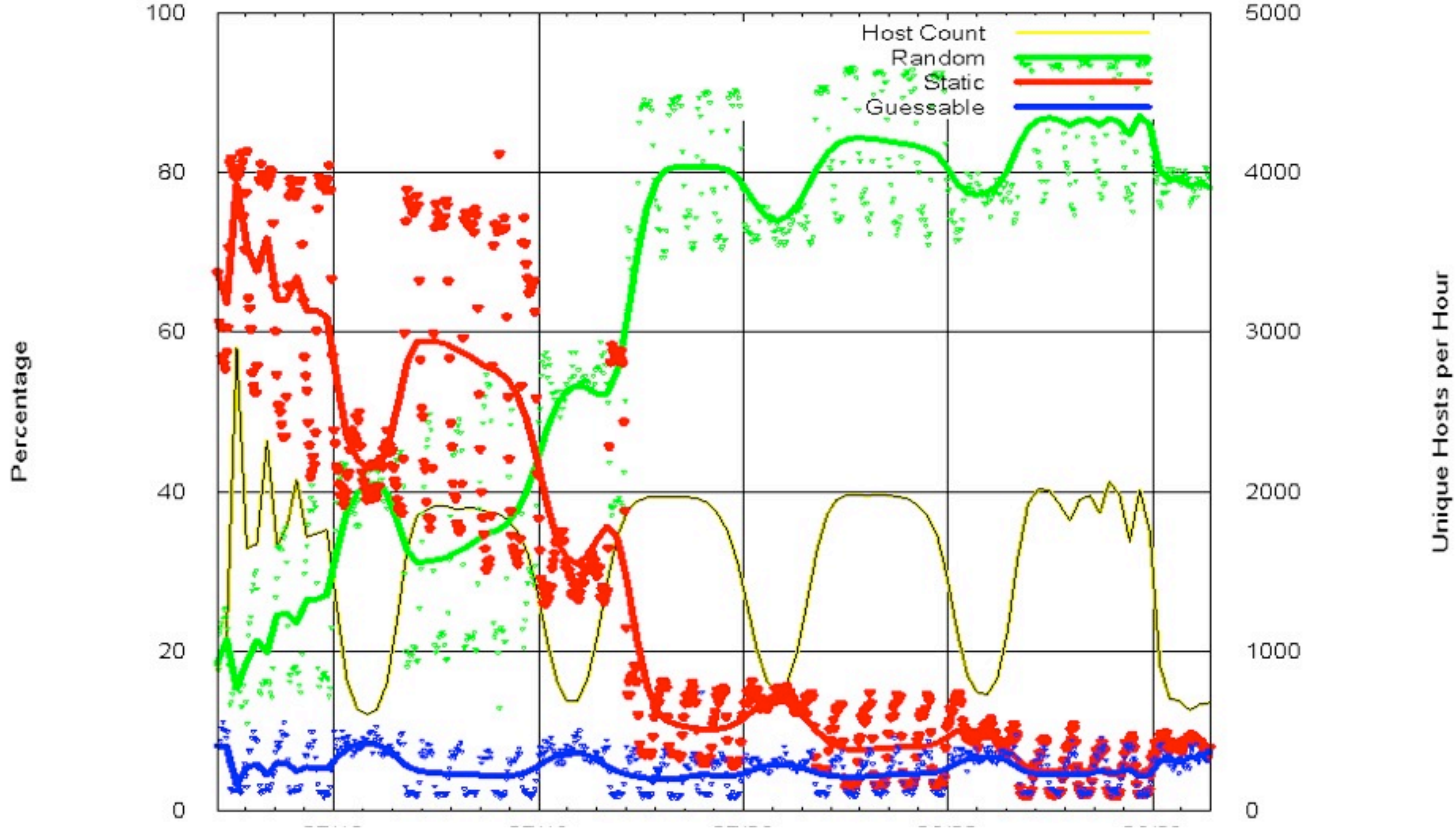
flow

- And the rest…
  - Time / Date etc.
- Lots of variations / possibilities here
  - Number of packets sent / received
  - Number of bytes sent / received

But I don’t do billing?
(or even if you do)

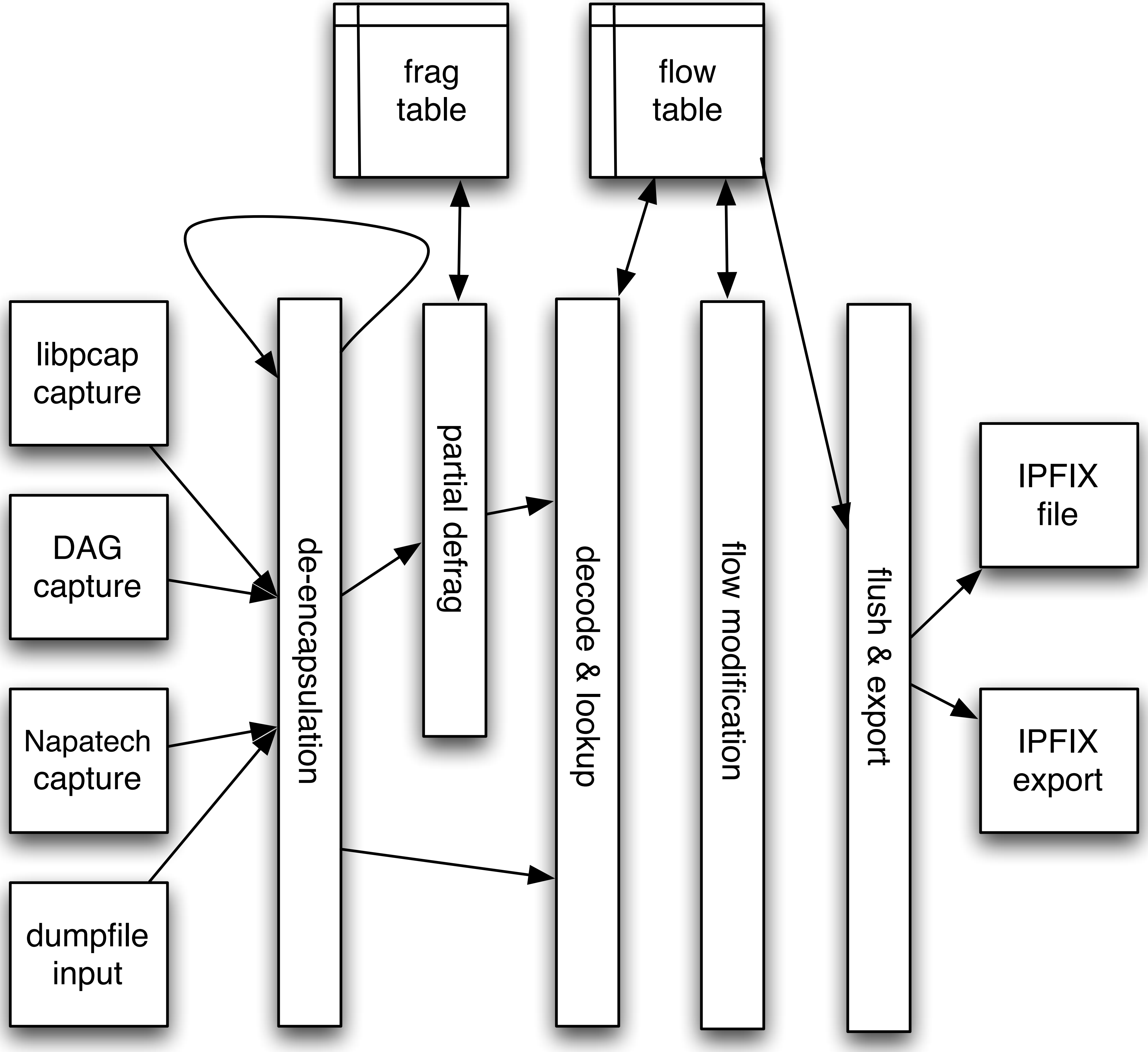
Kaminsky DNS protocol vulnerability

- Cache poisoning via DNS transaction ID guessing
- Not enough randomness, makes guessing easy

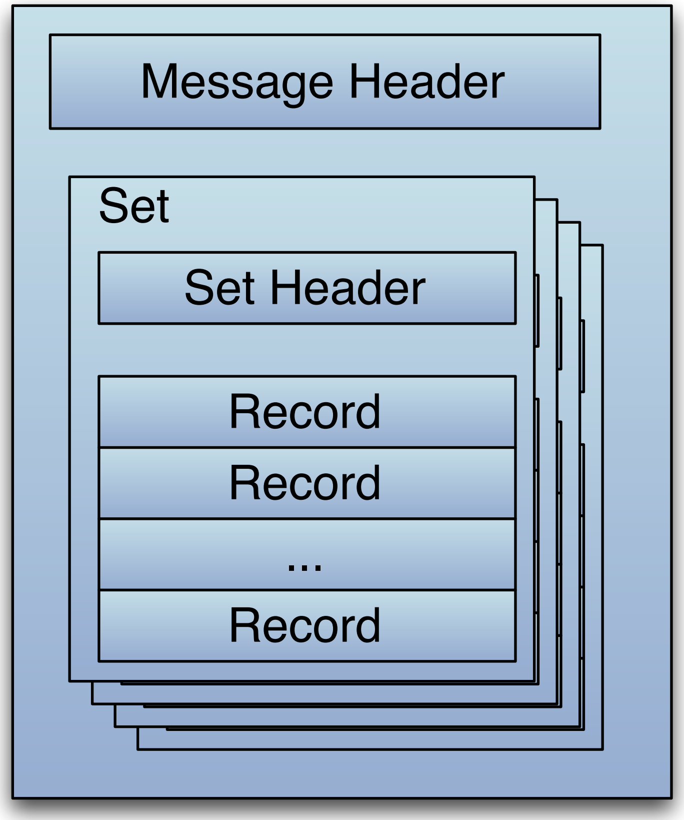

Host Count
Random
Static
Guessable
Percentage
Unique Hosts per Hour
100
80
60
40
20
0
5000
4000
3000
2000
1000
0

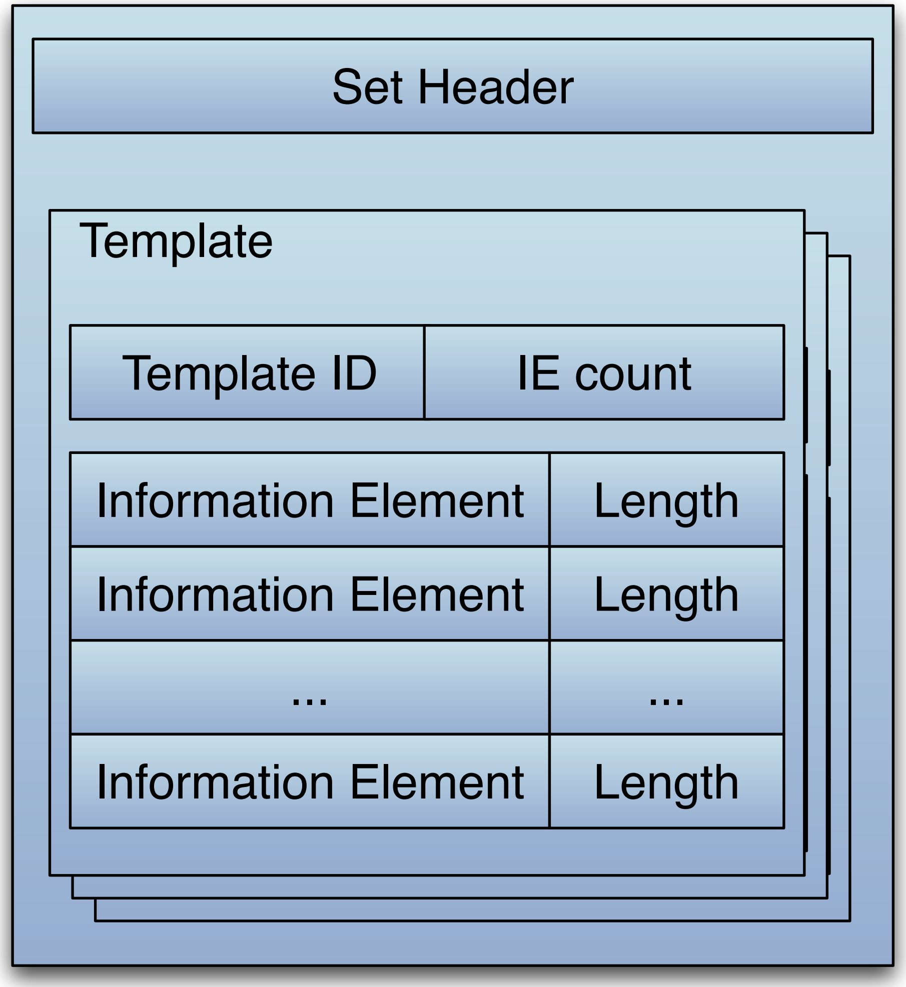
Objectives in YAF's construction

- Compliant to standard for flow, IPFIX
- Biflow based construction
- High performance (based on profiling)
- Flexible L2 decoding
- Open design for adding enhancements

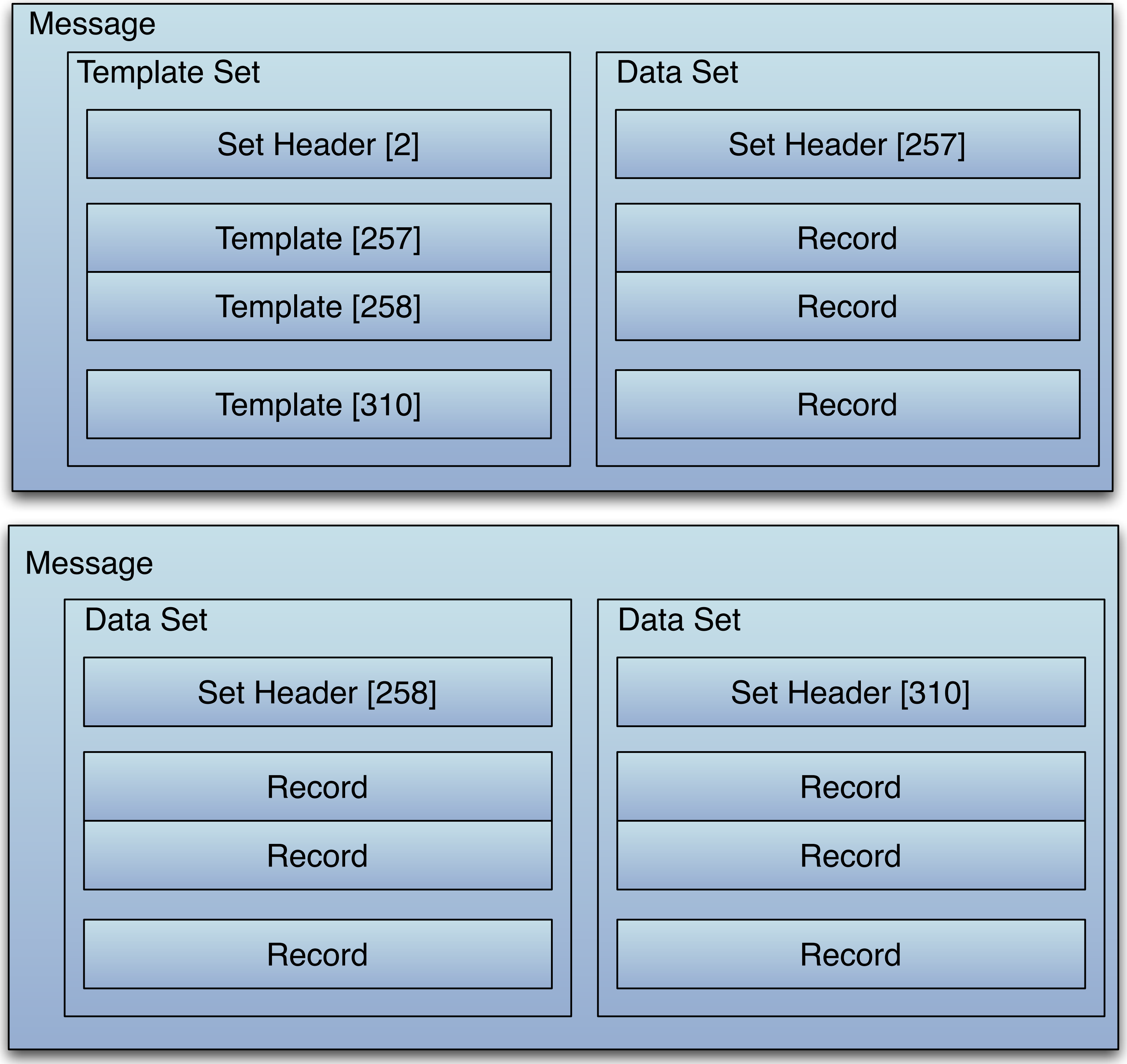

frag table
flow table
libpcap capture
DAG capture
Napatech capture
dumpfile input
de-encapsulation
partial defrag
decode & lookup
flow modification
flush & export
IPFIX file
IPFIX export

Condensed IPFIX Primer





Condensed IPFIX Primer



Message

Template Set

Set Header [2]

Template [257]

Template [258]

Template [310]

Data Set

Set Header [257]

Record

Record

Record

Message

Data Set

Set Header [258]

Record

Record

Record

Data Set

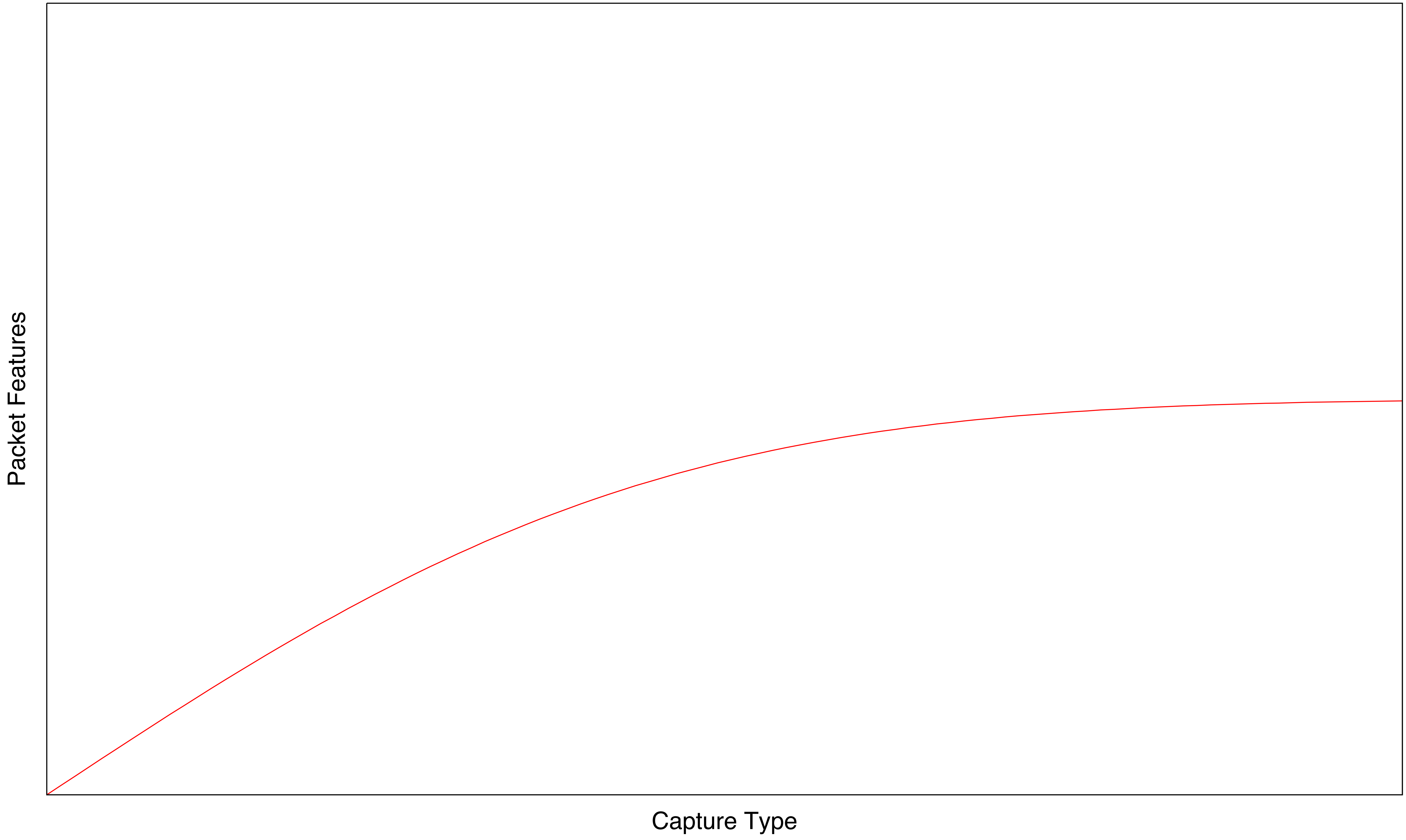
Set Header [310]

Record

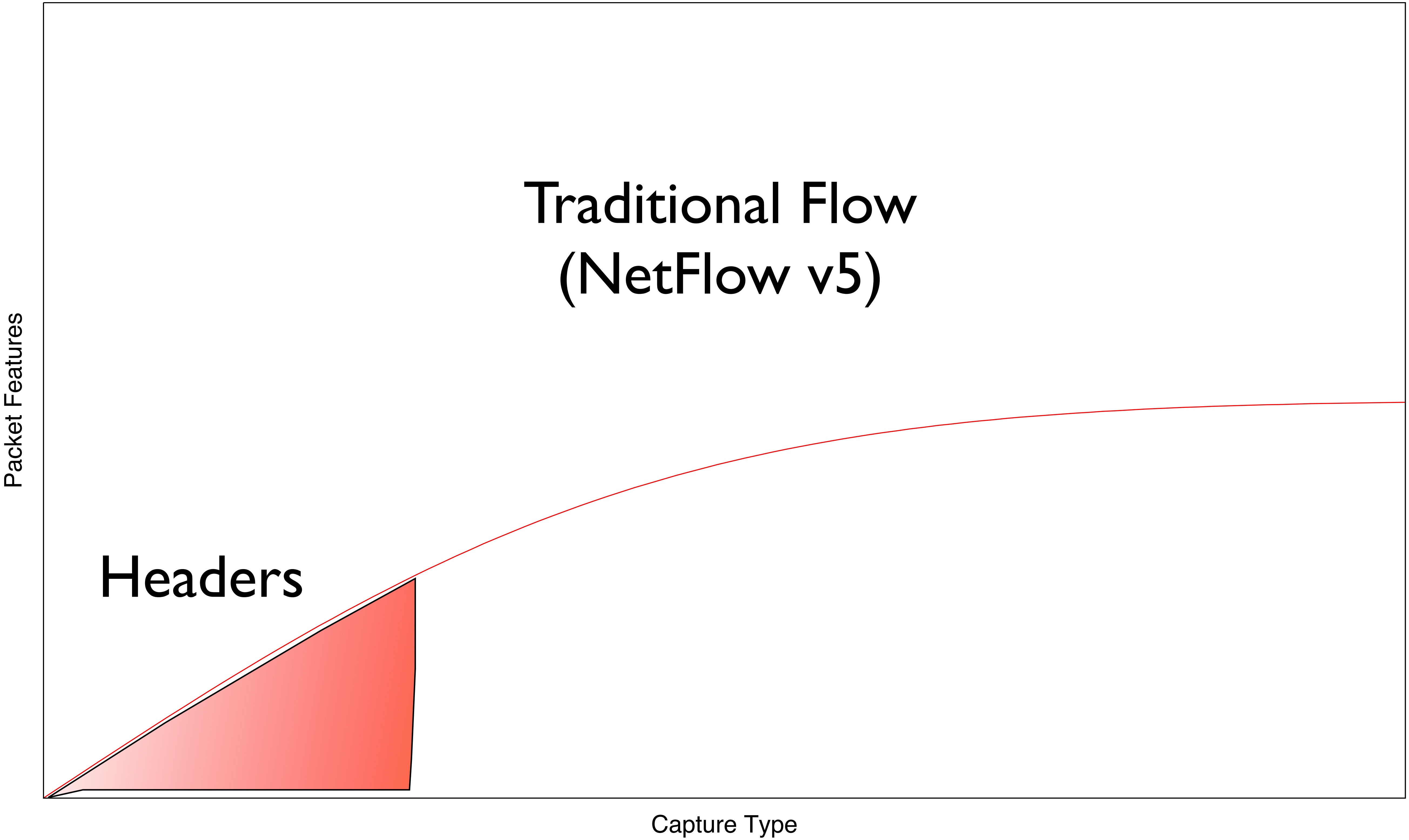
Record

Record

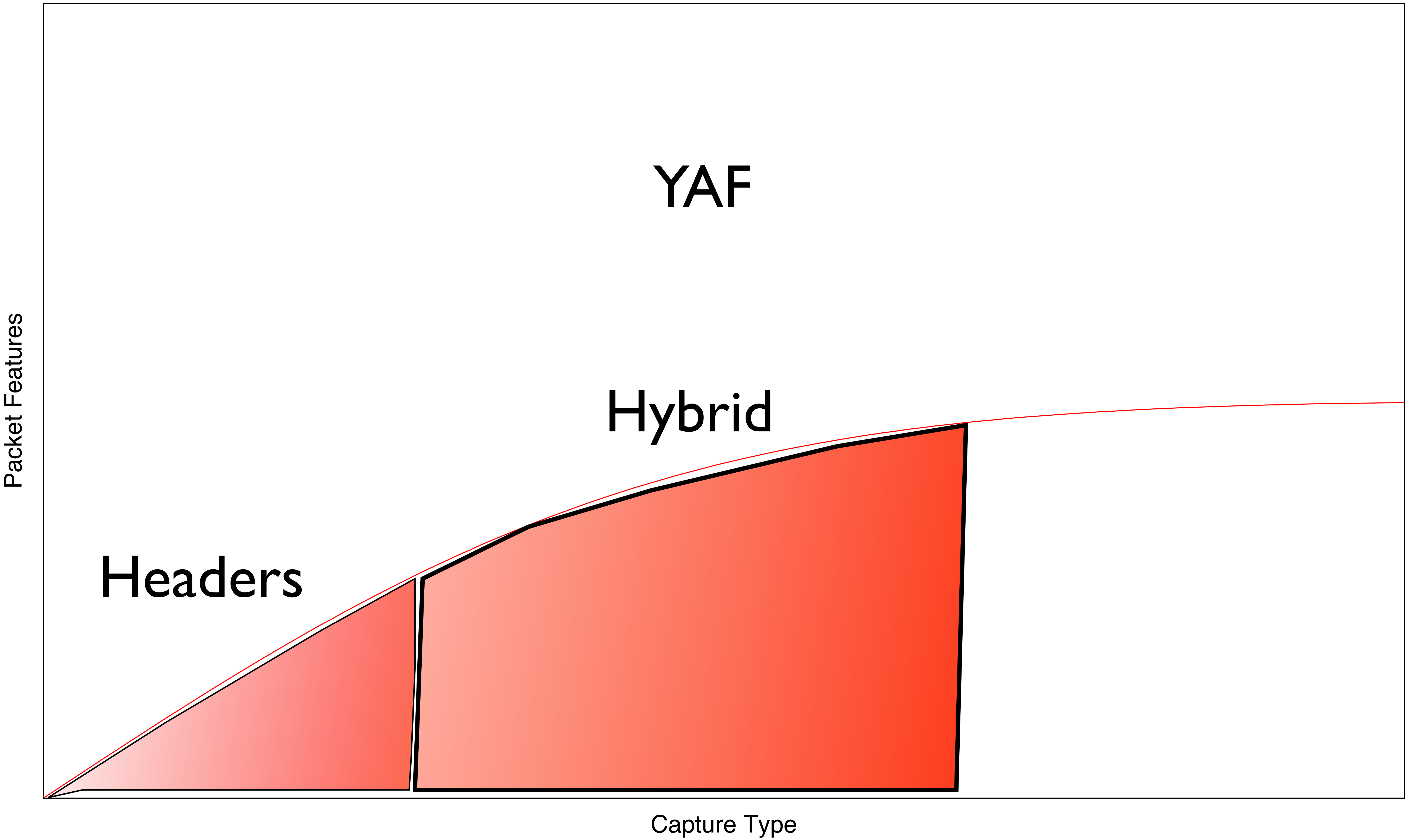
Network Capture Spectrum



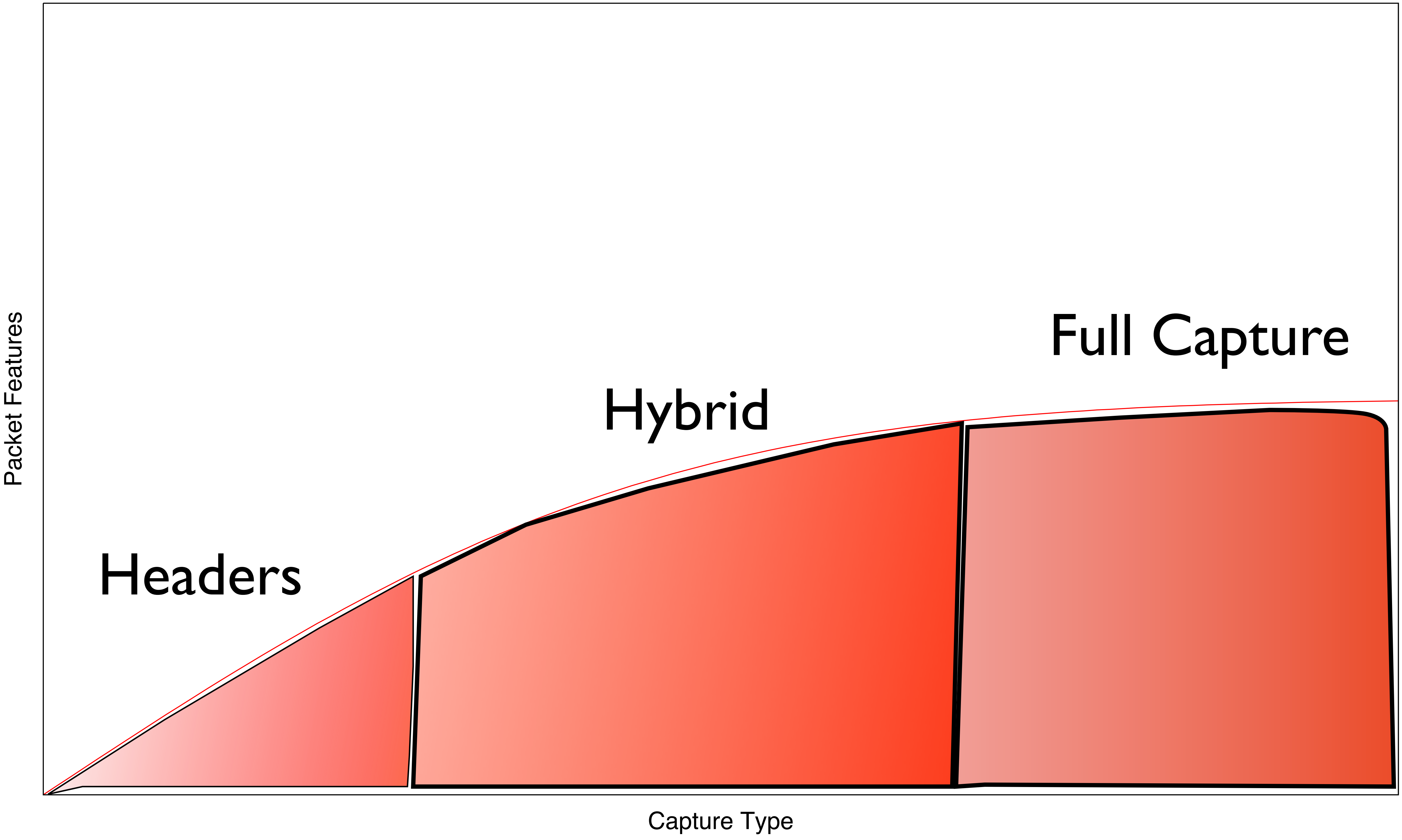
Network Capture Spectrum



Network Capture Spectrum



Network Capture Spectrum



Current YAF Capture (minimal privacy impact)

- Balancing Act Between Understanding Our Network and Privacy
  - Basic flow information:
    - Who talked to whom, how much, when
  - Application labeling:
    - Banner analysis for port independent protocol checking

Current YAF capture (minimal privacy impact)

- Application labeling (continued)
  - can recognize:
    - HTTP, SSH, SMTP, Gnutella, Yahoo Messenger, DNS, FTP, SSL/TLS, SLP, IMAP, IRC, RTSP, SIP, RSYNC, PPTP, NNTP, TFTP, Teredo, MySQL, POP3

Current YAF capture (minimal privacy impact)

- Entropy analysis
  - Good indication if traffic is encrypted or compressed

Current YAF Capture

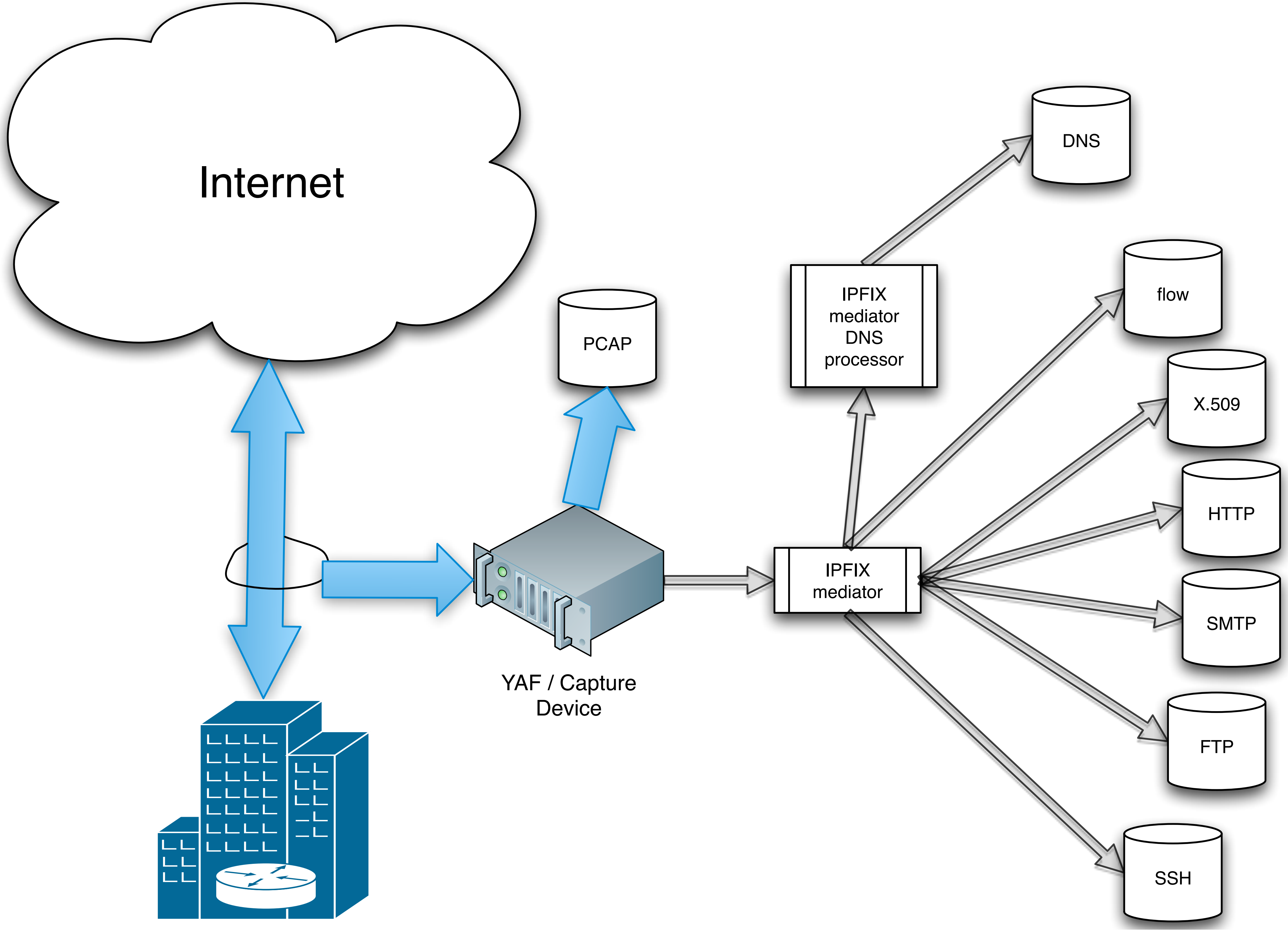
- DNS capture
  - Because it is the root of almost all valid network transactions
  - We can limit capture to just Authoritative and NXDomain responses
  - Or capture all DNS transaction information

Current YAF Capture

- Highly detailed capture for specific protocols:
  - HTTP
    - Server, User-Agent, GET, Connection
    - HTTP, Referer, Location, Host
    - Content-Length, Age, Content-Type
    - Accept, Accept-Language,(Result Code)

Current YAF Capture

- Other in depth protocols
  - FTP, IMAP, RTSP, SIP, SMTP, SSH
- Soon to be added
  - X.509 Certificates
  - Primarily from recognized SSL/TLS protocol negotiations


Internet
PCAP
YAF / Capture Device
IPFIX mediator DNS processor
IPFIX mediator
DNS
flow
X.509
HTTP
SMTP
FTP
SSH

Capturing Flow (and others) using IPFIX

- Using the IPFIX model, we can turn on many features in YAF, and filter with mediators
- We can enhance our handling of specific data types, still carry the information in IPFIX, and send to future places

Finishing the Full Deployment

- We have some of the backend tools to handle the various different data types from YAF now. (Storage and analysis)
- Working on the simple/dumb backend (probably MySQL based) to just capture data (may not scale well enough)
- IPFIX mediator toolkit materials are available

Objectives Met?

- YAF is deployed in LARGE scale environments now
- We have been able to quickly add both network encapsulation types and specific network traffic data decoders quickly
- IPFIX has proven to be both compact and flexible

Where do you fit in?

- It is available for you to use
- You can enhance and extend it - we are willing to take contributions
- Adding certain new detectors (especially for text based protocols) is *really* easy
- You tell me

Getting YAF

http://tools.netsa.cert.org

netsa-help@cert.org

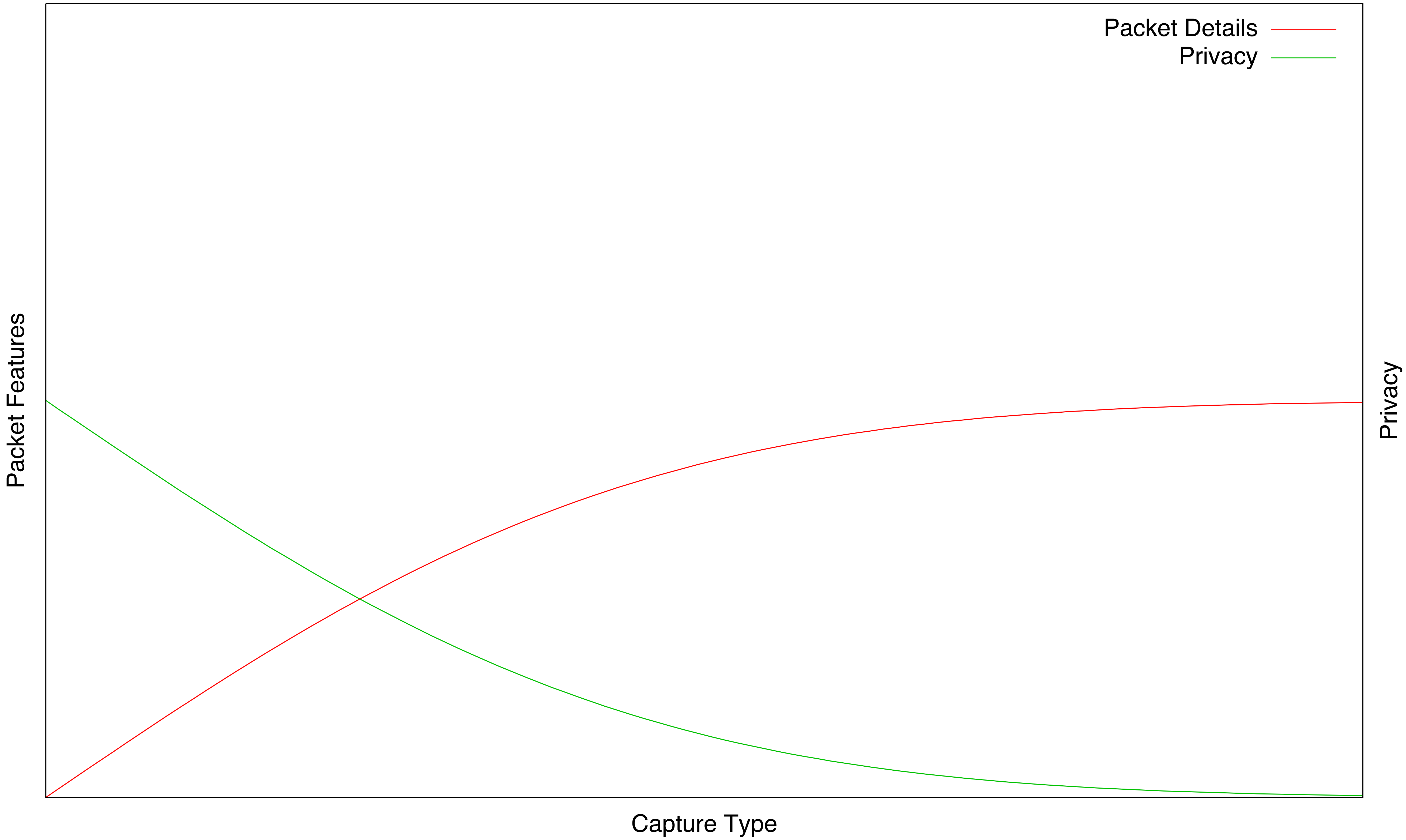
Questions?
Comments?

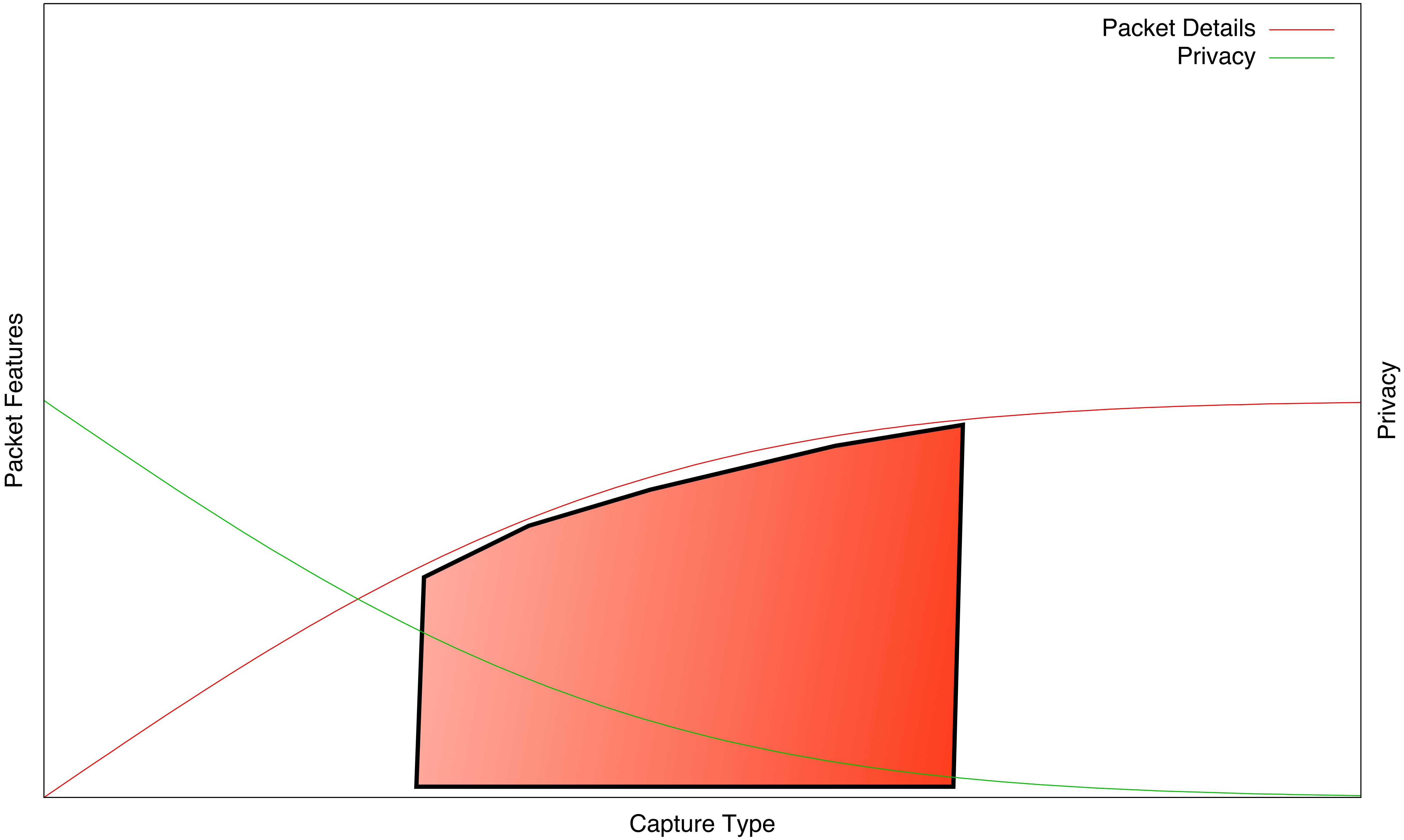
Gratuitous plug:

FloCon®2011

Salt Lake City Marriott Downtown
Salt Lake City, Utah
January 10-13, 2011

Backups


Packet Details
Privacy
Packet Features
Privacy
Capture Type


Packet Details
Privacy
Packet Features
Privacy
Capture Type