Network Technical Summit

Cloud Security AI

Inline inferencing for networking functions with Traffic Analytics Development Kits (TADK)

Qiu, Kun

intel®

Agenda

- Background
  - Cloud security
    - Why AI-based security?
- Problem
- Solution
- Optimization

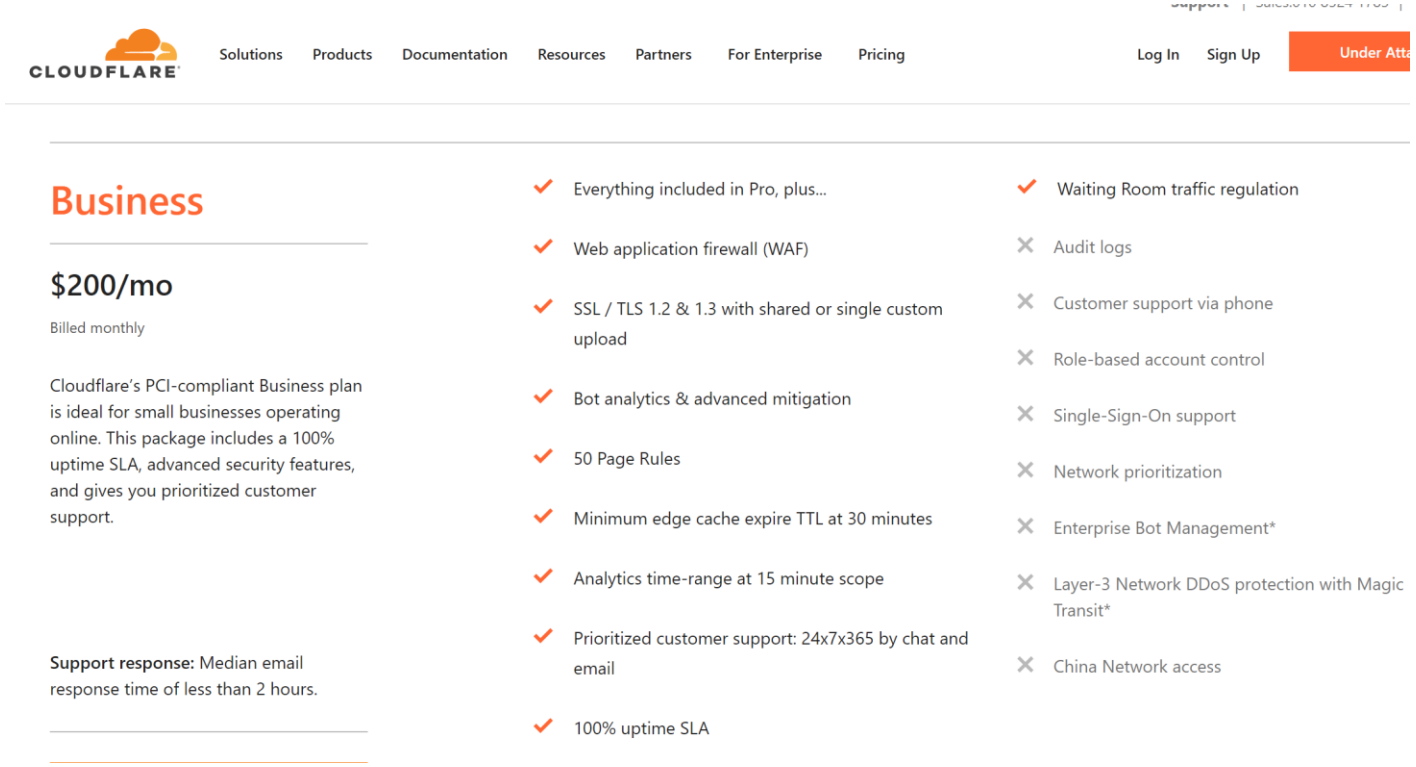
Cloud Security

Magic Quadrant for Network Firewalls

- Capabilities of network firewalls include:
  - Application awareness and control
  - Intrusion detection and prevention
  - Advanced malware detection
  - Logging and reporting
  - Reference:
    - https://www.gartner.com/doc/reprints?id=1-27UP5WXW&ct=211102&st=sb

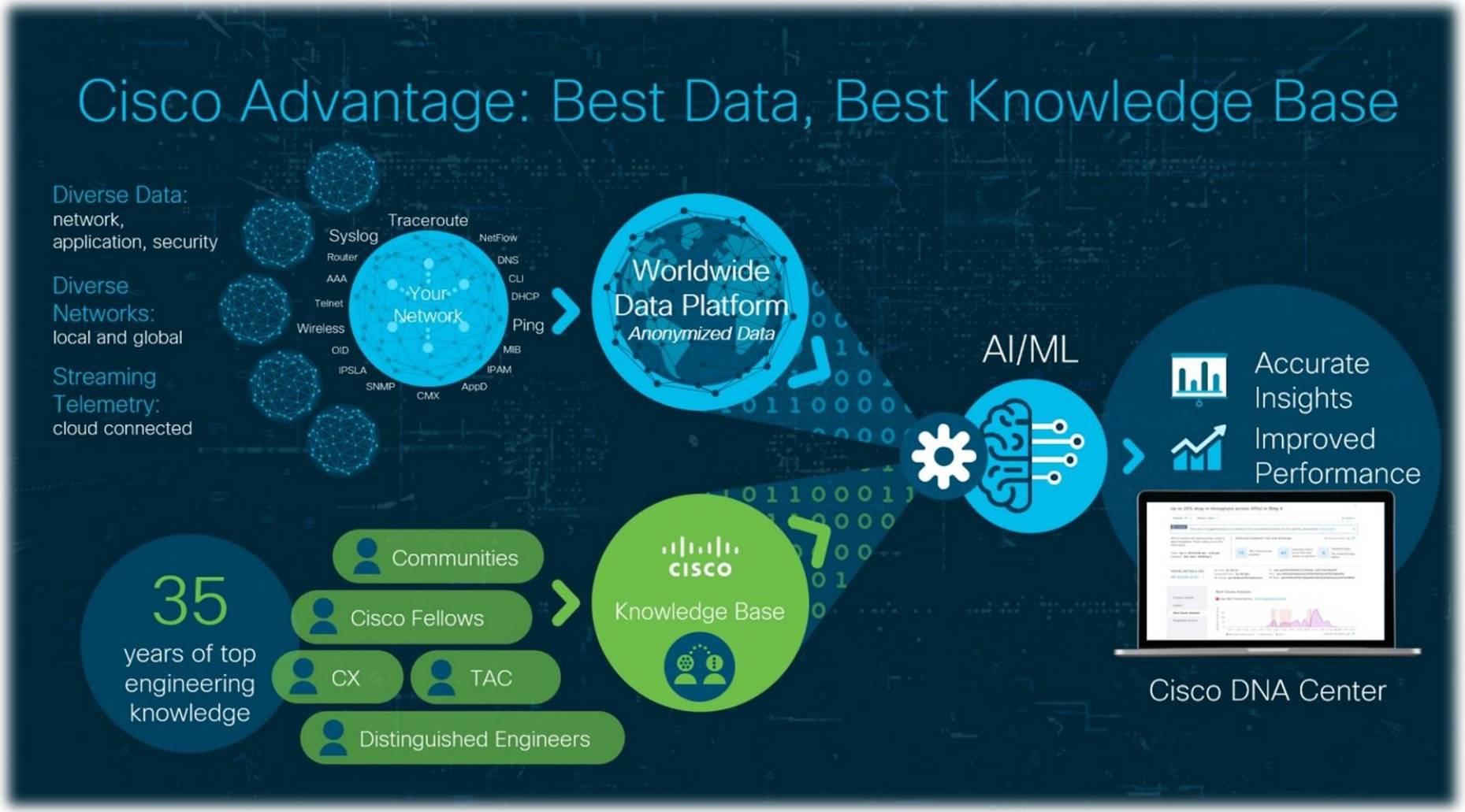
Why AI-Based Security?

Example: Web Application Firewall



- Rules:
  - Open source?
    - Core rule sets (basic)
  - Regular expression?
    - The Witcher's Language
    - \b[A-Z0-9._%+-]+@[A-Z0-9.-]+\.[A-Z]{2,}\b
  - How about encrypted traffic?

Why AI-Based Security?

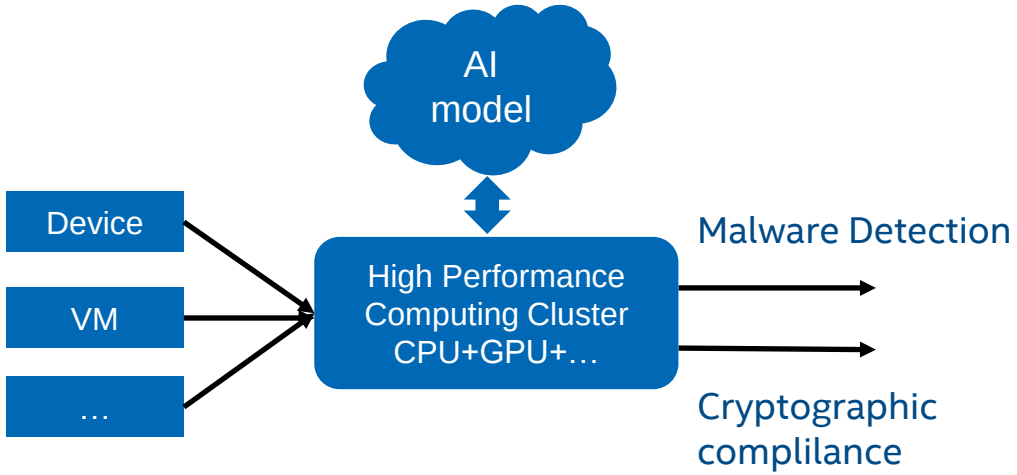


Agenda

- Background
- Problem
  - Why network AI deployment so hard?
    - Performance issue/Data
    - Developing tools
- Solution
- Optimization

Why Network AI Deployment So Hard

Performance



- Real time
- How about cloud?
  - Cost?

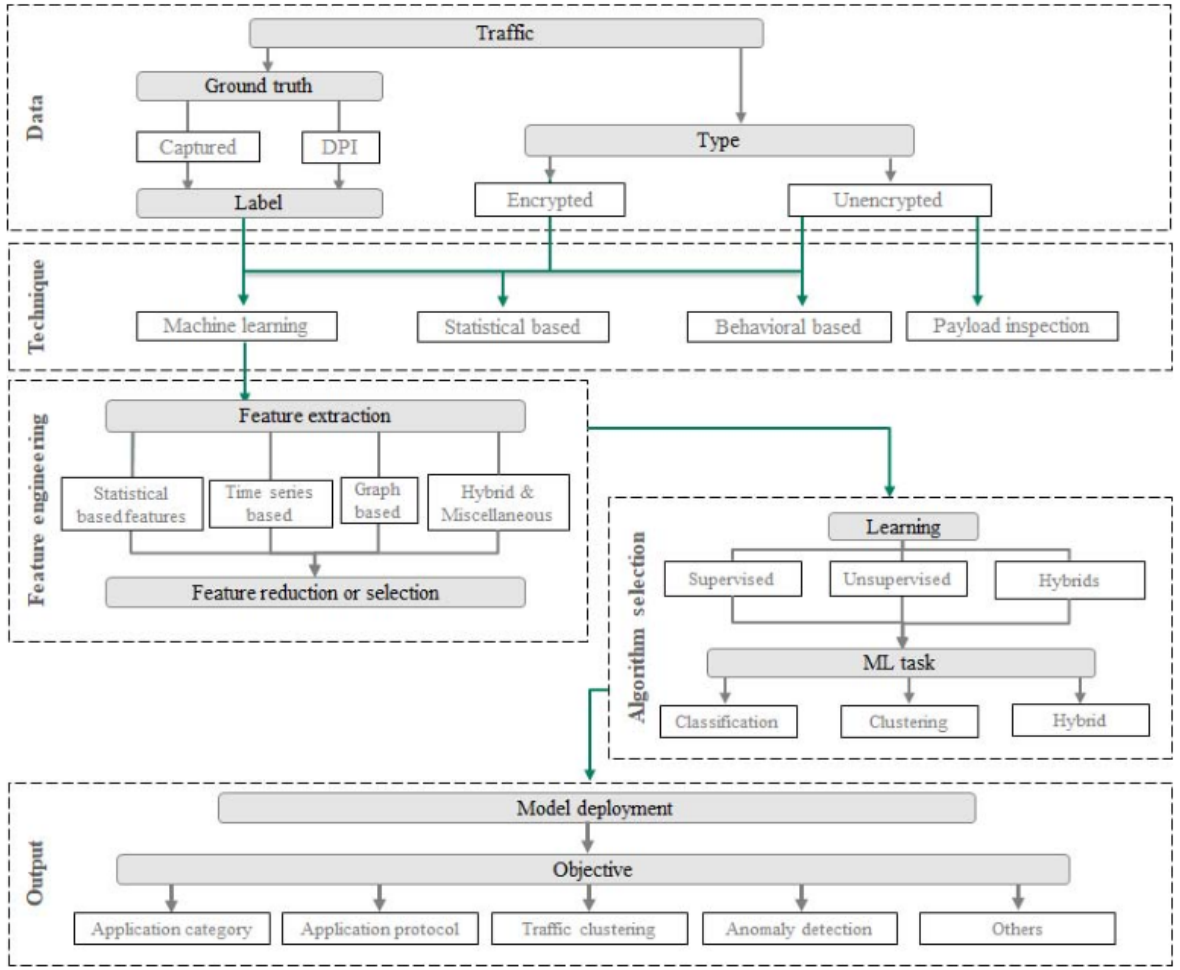
Data

- Different from NLP/Voice/Vision..
- How to collect?
  - Full packets?
    - Performance/Storage
  - Features?
- How to label?

Why Network AI Deployment So Hard

Development Tools

- Traffic data collection
- Feature extraction
- Feature reduction
  - Feature selection
- Algorithm selection
  - Model construction
- Validation



Reference: Towards the deployment of Machine Learning solutions in network traffic classification: A systematic survey

Agenda

- Background
- Problem
- Solution
  - Development Tools: Traffic Analytics Development Kits (TADK)
    - “White Box” Software
    - Example: network traffic analytics with AI
  - Performance
  - Data
- Optimization

Solution

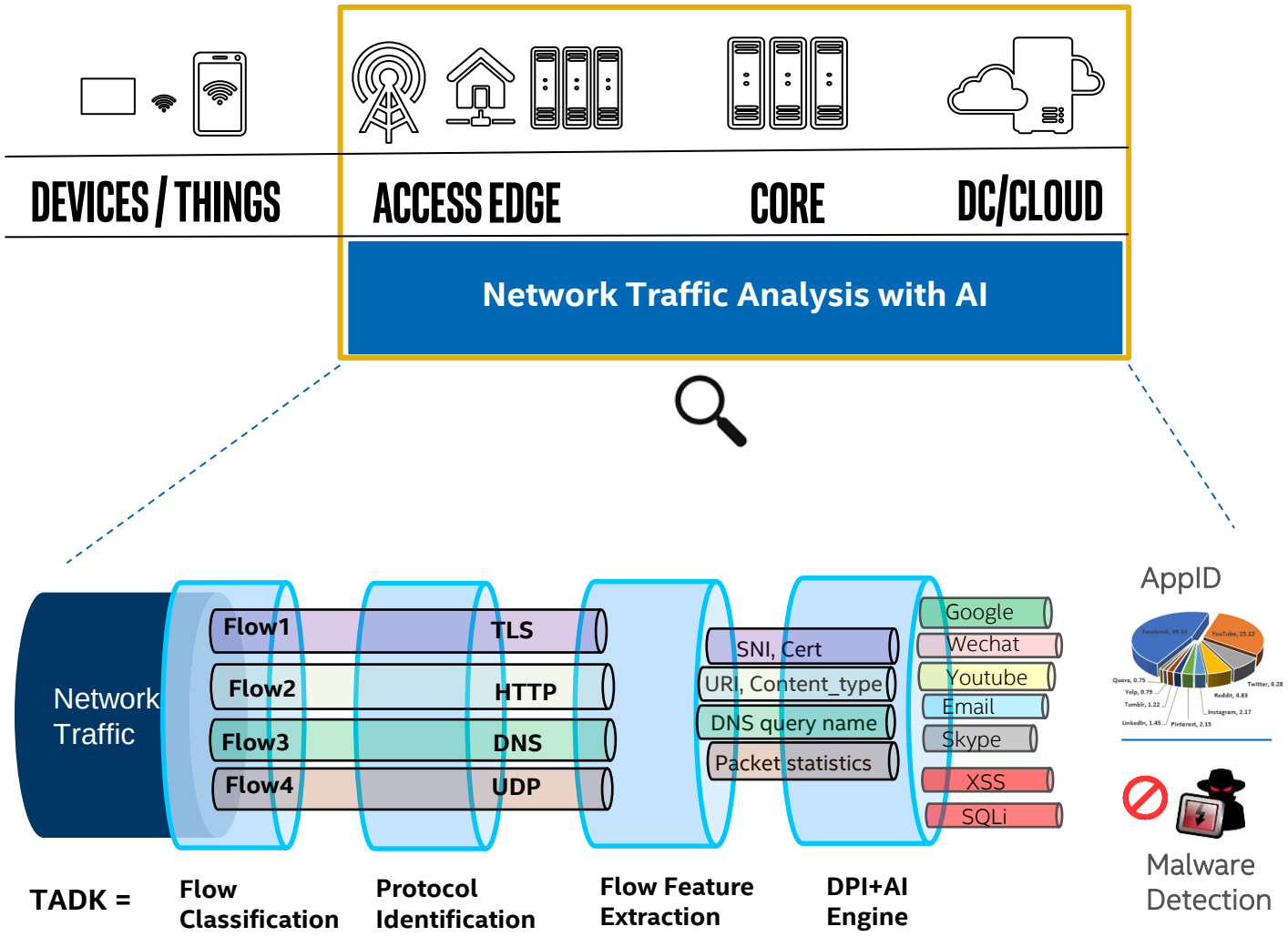
DPDK

- White box equipment/software
  - Packet-level
  - High-performance router
  - Load-balance
  - Switch
  - …

TADK

- “White Box” Software
  - Traffic-level
  - Traffic classification
    - Traffic probe/Sensor
    - Encrypted traffic analytics
    - Quality of service
    - ….
  - Next generation firewall
    - Web application firewall
    - …

Traffic Analytics Development Kit (TADK)

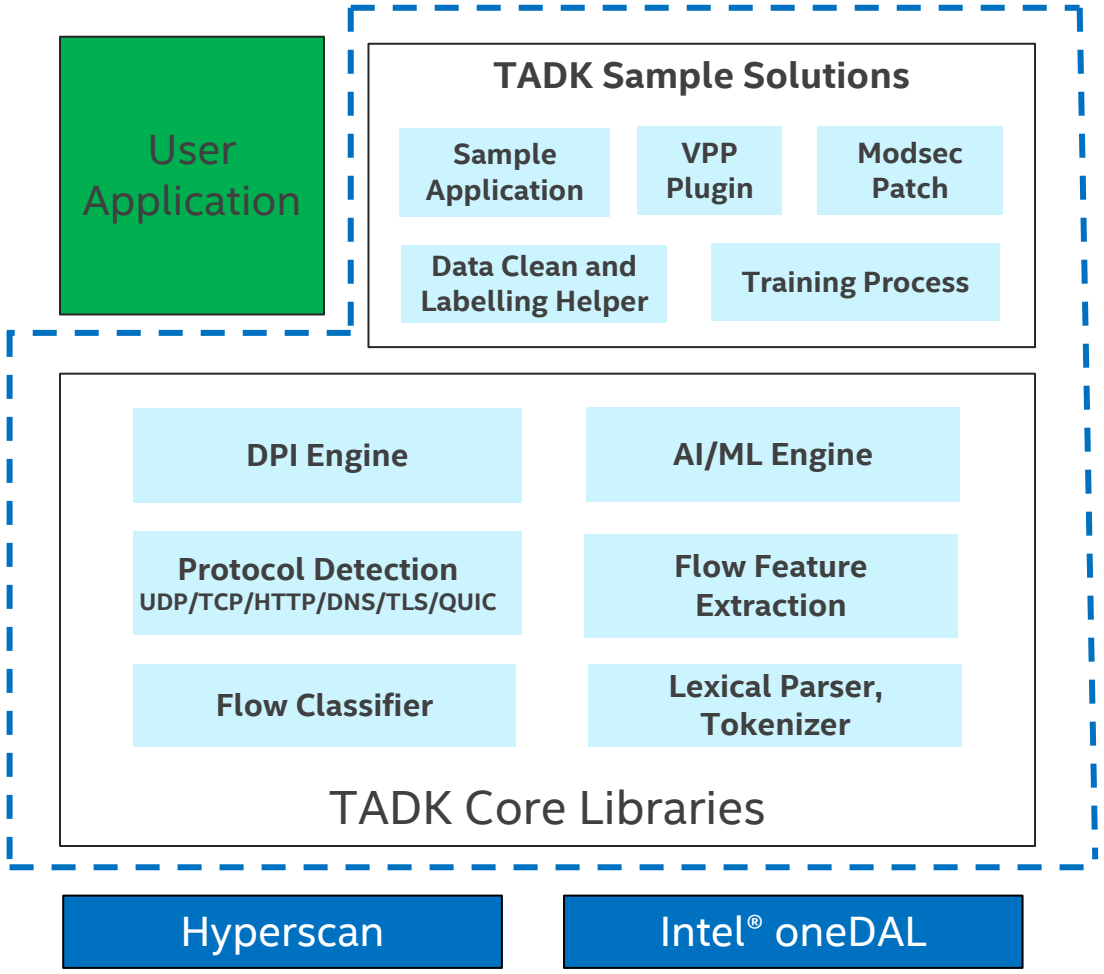


- Network Traffic Analysis is critical in various workloads such as:
  - Content inspection in SD-WAN/CPE
  - Next-gen Firewall
  - Encrypted Traffic Analysis
  - Web Application Firewall
  - Network Visualization
- TADK accelerating traffic analysis pipeline by:
  - Rule-based Deep Packet Inspection
  - AI and Machine learning

Traffic Analysis Development Kit (TADK): Highly optimized libraries on Intel's platform

Traffic Analytics Development Kit (TADK)

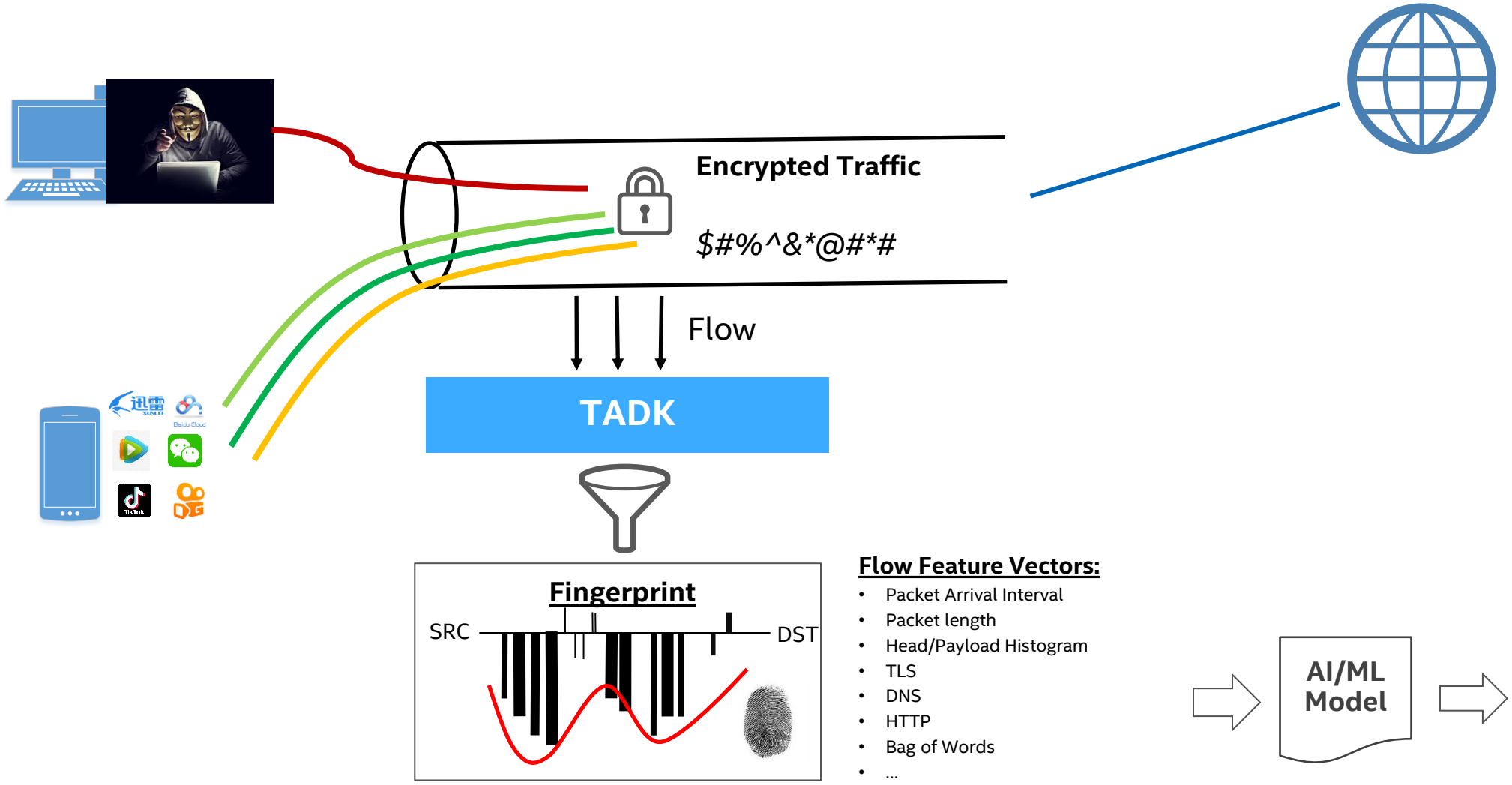
- Provide fundamental building-blocks for Traffic analysis workloads
- Highly optimized libraries on IA platform with modular design
- Support DPI and AI engines



Agenda

- Background
- Problem
- Solution
  - Development Tools: Traffic Analytics Development Kits (TADK)
  - Performance
    - Feature extraction
    - Model prediction
  - Data
- Optimization

Example: Encrypted Traffic Analysis



Flow Feature Vectors:

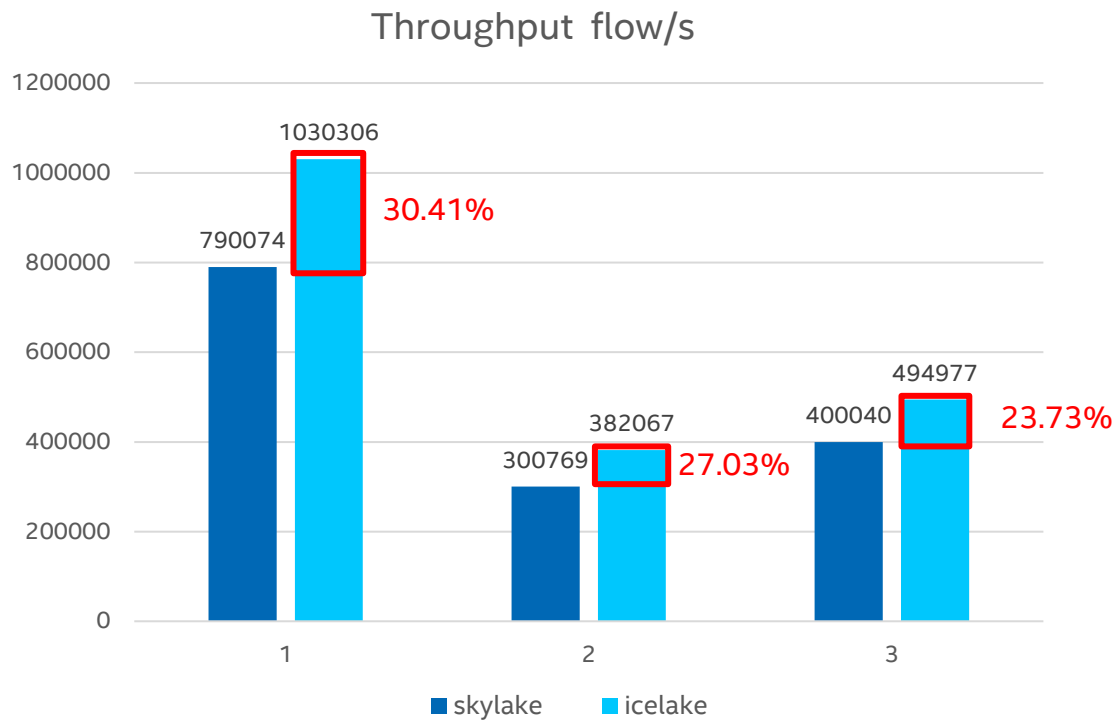
- Packet Arrival Interval
- Packet length
- Head/Payload Histogram
- TLS
- DNS
- HTTP
- Bag of Words
- …

*Analysis Encrypted Traffic **without decryption***

Throughput (Feature Extraction)

- Platform:
- Intel(R) Xeon(R) Gold 6148 CPU @ 2.40GHz
- Intel(R) Xeon(R) Platinum 8358 CPU @ 2.60GHz

- PCAP:
- 1 - dns.pcap
- 2 - http.pcap
- 3 - tls.pcap



pcap	flow count	packet count	packet per flow (average)	byte count
dns.pcap	25,112	52,589	2	4,895,727
http.pcap	1792	14,320	8	12,806,007
tls.pcap	6341	82,005	13	48,271,549

Example Traffics

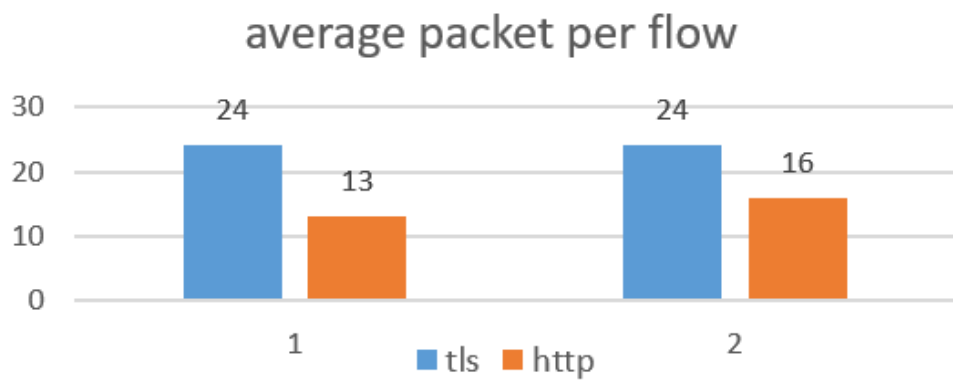
flow table + feature extraction + prediction

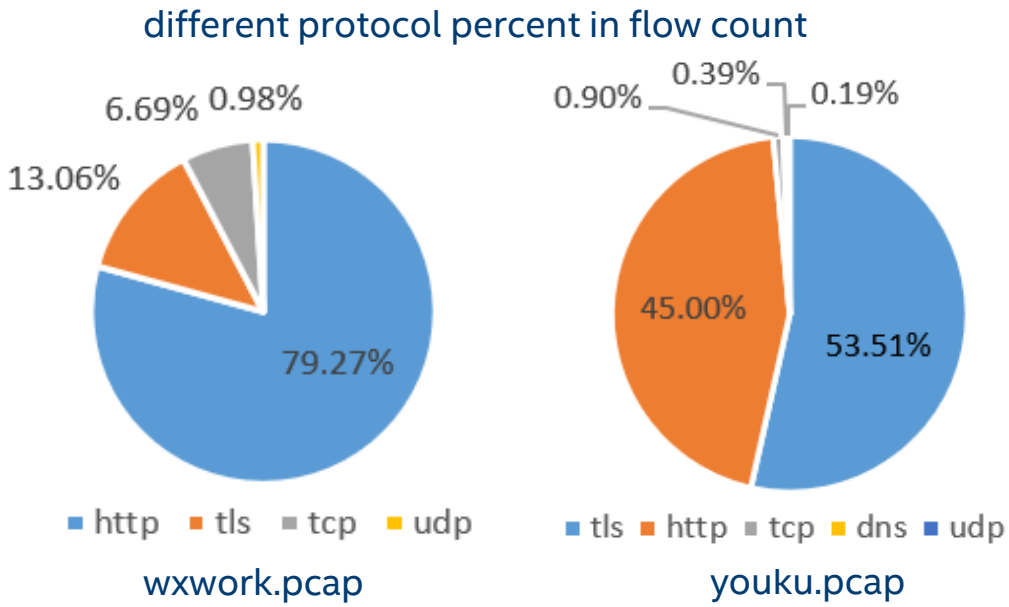
PCAP:

1 - wxwork.pcap

2 - youku.pcap

pcap	flow count	packet count	packet per flow	byte count
wxwork.pcap	1524	21,995	15	5,082,409
youku.pcap	1551	31,071	20	15,474,984



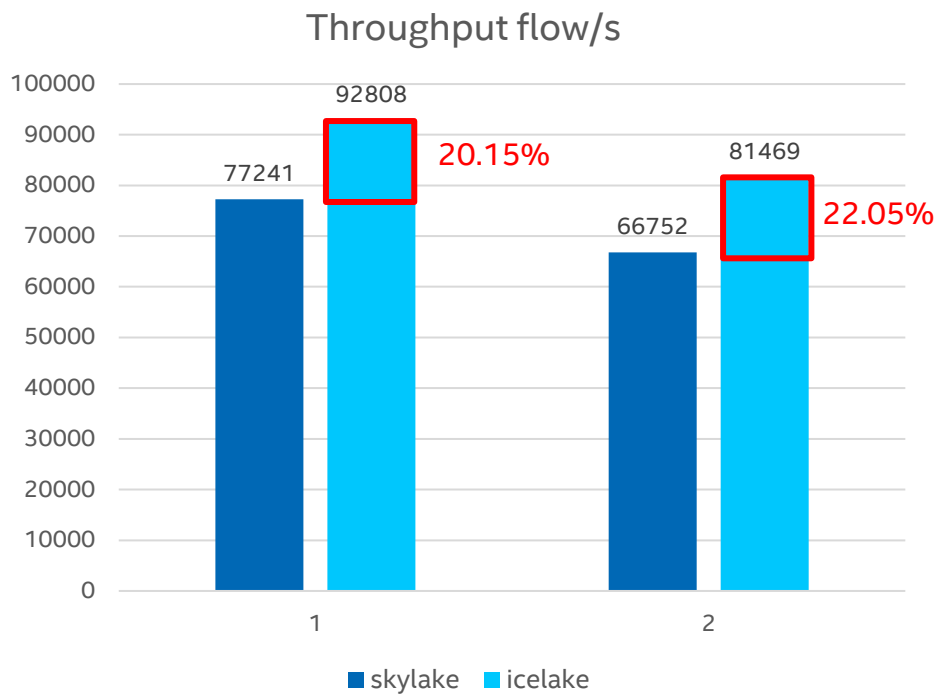


Throughput (Overall)

flow table + feature extraction + prediction

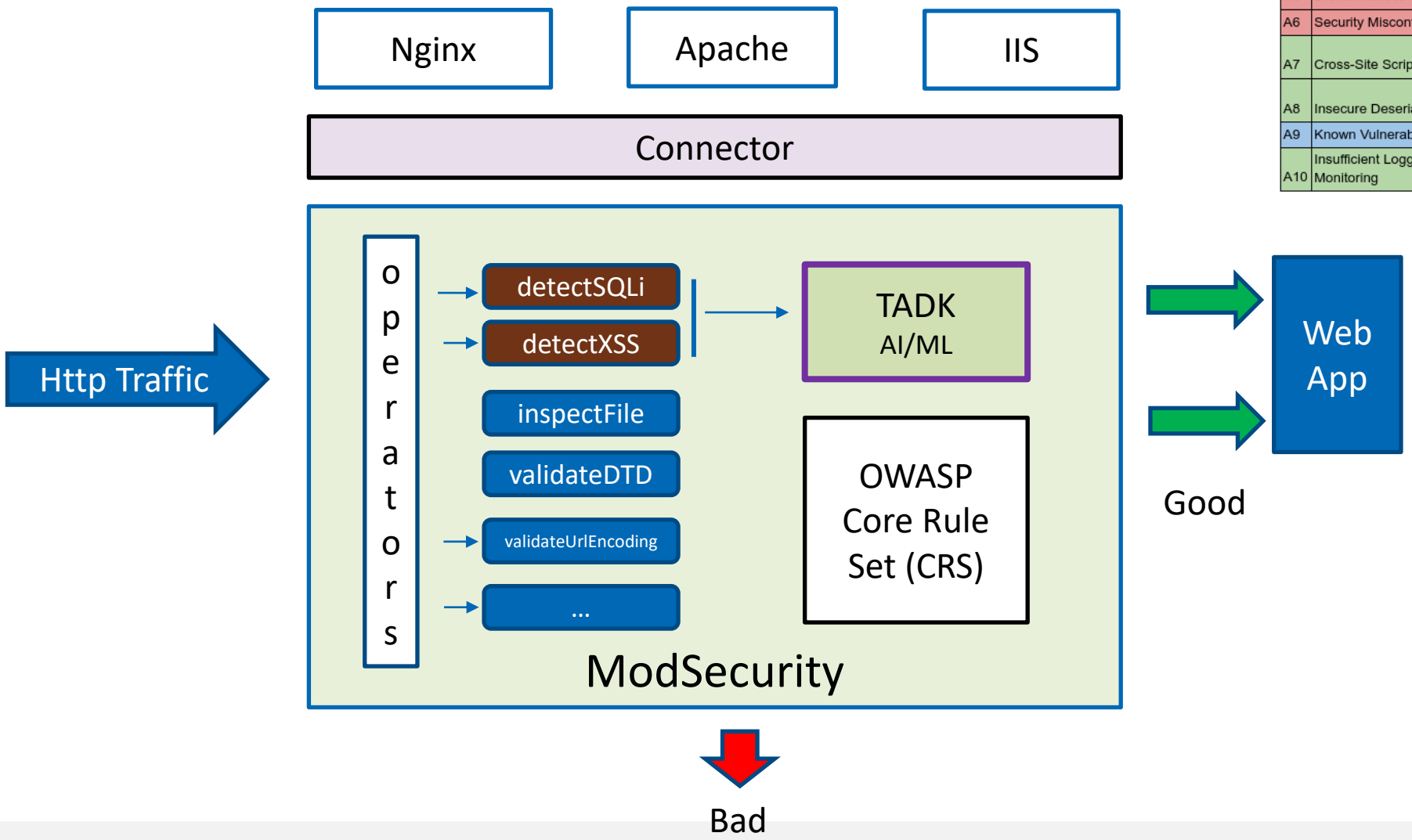
- Platform:
- Intel(R) Xeon(R) Gold 6148 CPU @ 2.40GHz
- Intel(R) Xeon(R) Platinum 8358 CPU @ 2.60GHz

- PCAP:
- 1 - wxwork.pcap
- 2 - youku.pcap

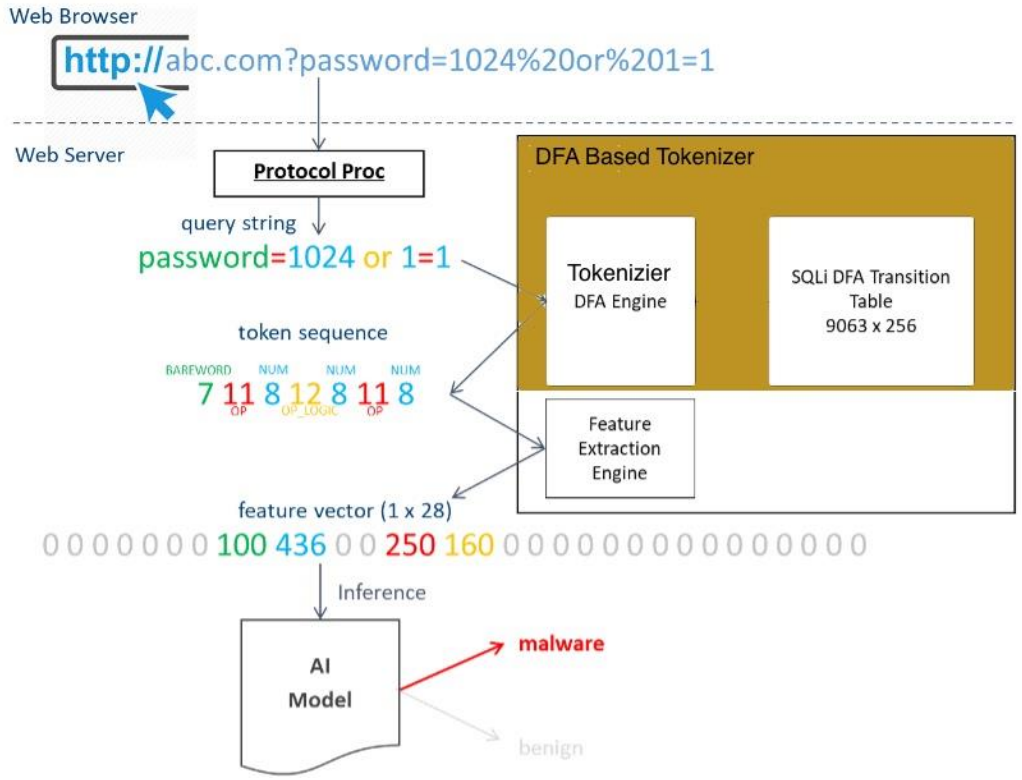


Usage: WAF

OWASP Top 10 2017		change	OWASP Top 10 2021 proposal	
A1	Injections	**as is**	A1	Injections
A2	Broken Authentication	**as is**	A2	Broken Authentication
A3	Sensitive Data Exposure	**down 1**	A3	Cross-Site Scripting (XSS)
A4	XML eXternal Entities (XXE)	**down 1 + A8**	A4	Sensitive Data Exposure
A5	Broken Access Control	**down 1**	A5	Insecure Deserialization **(merged with XXE)**
A6	Security Misconfiguration	**down 4**	A6	Broken Access Control
A7	Cross-Site Scripting (XSS)	**up 4**	A7	Insufficient Logging & Monitoring
A8	Insecure Deserialization	**up 3 + A4**	A8	**NEW:** Server Side Request Forgery (SSRF)
A9	Known Vulnerabilities	**as is**	A9	Known Vulnerabilities
A10	Insufficient Logging & Monitoring	**up 3**	A10	Security Misconfiguration

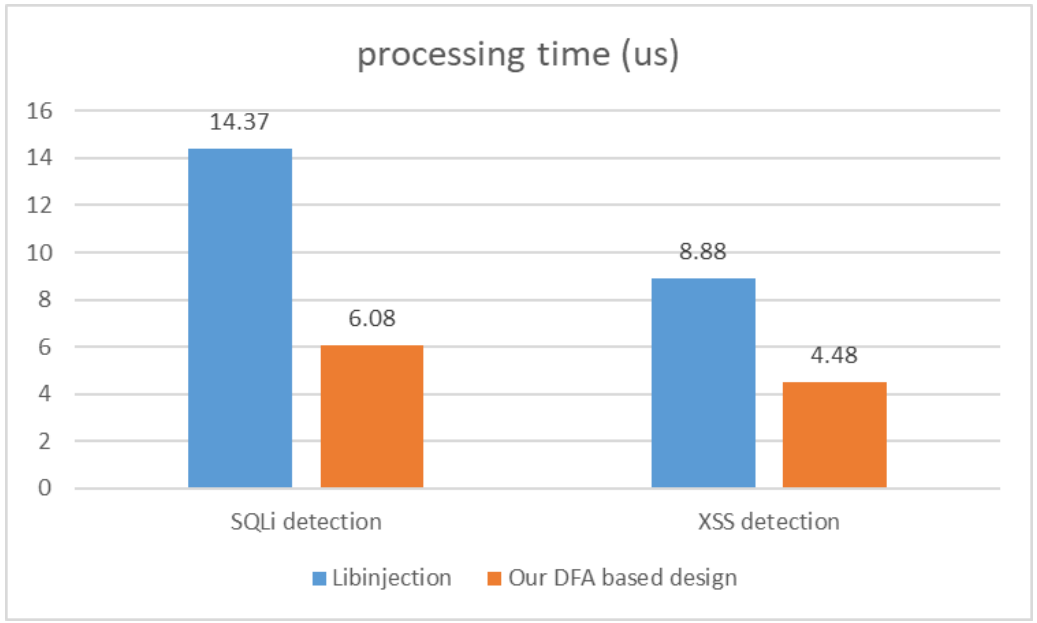


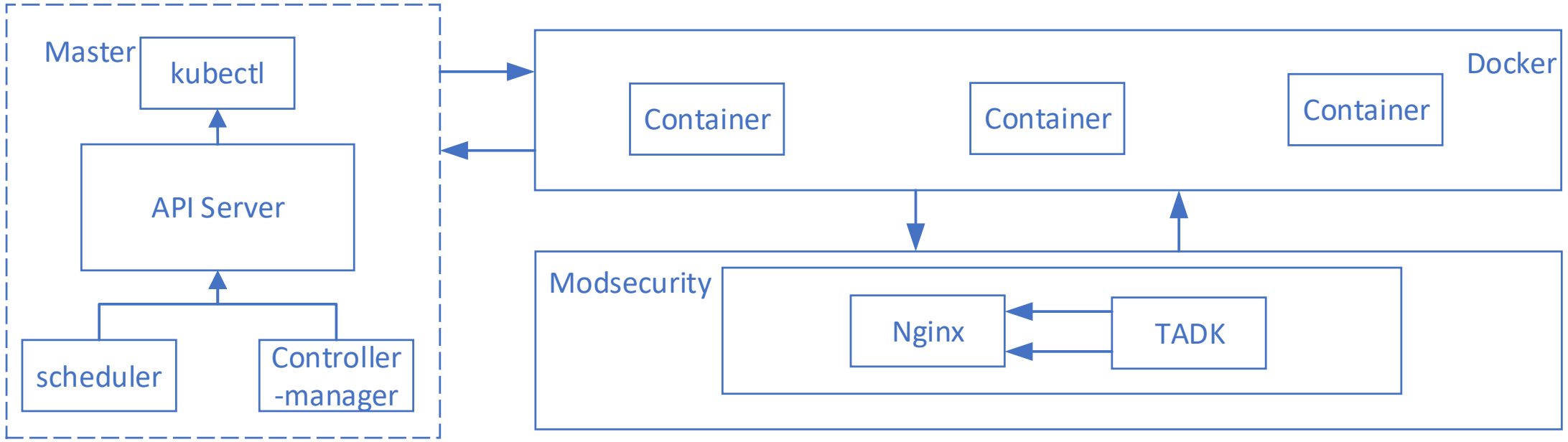
HTTP Feature Extraction



Intel(R) Xeon(R) Platinum 8358 CPU

feature extraction + prediction



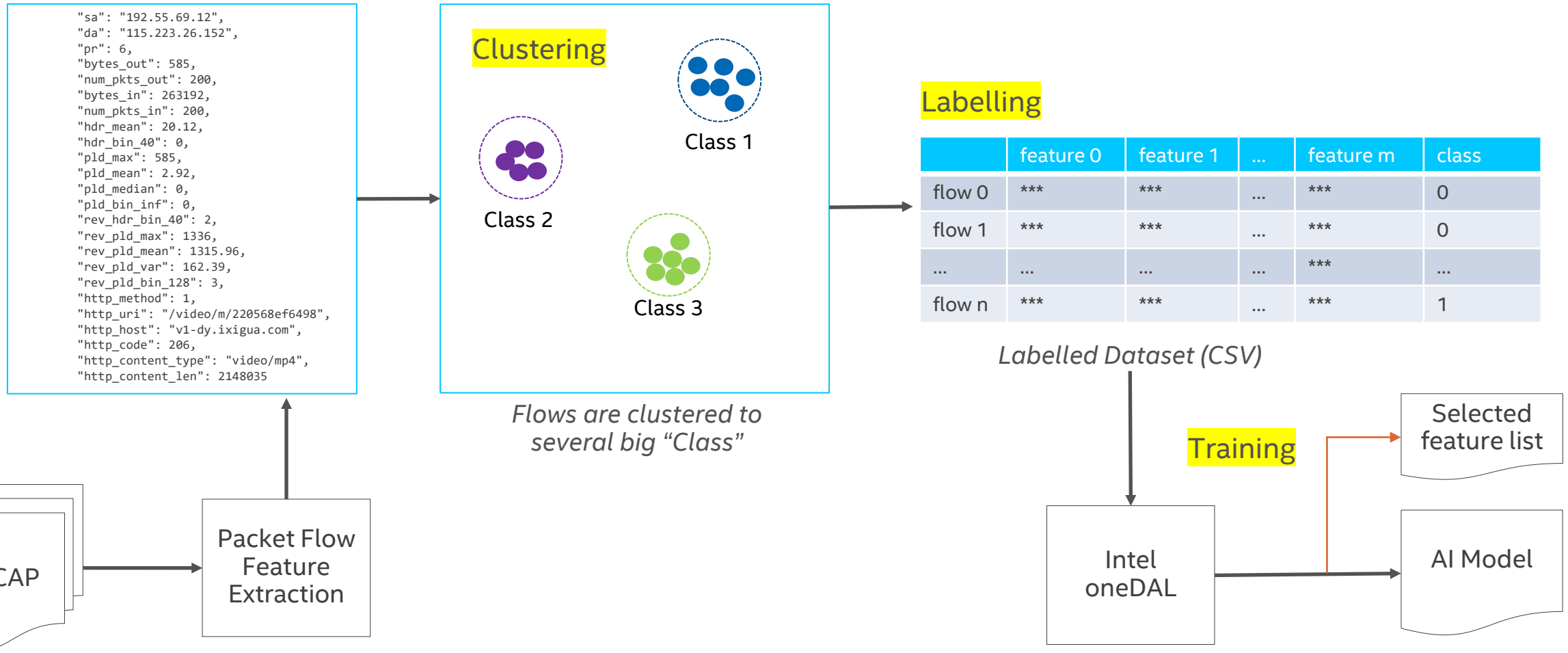



```
new@new-S2600STQ:~/container-experience-kits$ sudo kubectl get all -n modsec-tadk -o wide ;
NAME                                          READY   STATUS    RESTARTS   AGE   IP             NODE          NOMINATED NODE   READINESS GATES
pod/tadk-intel-tadkchart-7594bd49b9-7747m     1/1     Running   0          24s   10.244.0.134   new-s2600stq   <none>           <none>
pod/tadk-intel-tadkchart-7594bd49b9-9v72g     1/1     Running   0          24s   10.244.0.128   new-s2600stq   <none>           <none>
pod/tadk-intel-tadkchart-7594bd49b9-bwlr6     1/1     Running   0          24s   10.244.0.131   new-s2600stq   <none>           <none>
pod/tadk-intel-tadkchart-7594bd49b9-c68ss     1/1     Running   0          24s   10.244.0.130   new-s2600stq   <none>           <none>
pod/tadk-intel-tadkchart-7594bd49b9-f2ls5     1/1     Running   0          24s   10.244.0.127   new-s2600stq   <none>           <none>
pod/tadk-intel-tadkchart-7594bd49b9-hvm2c     1/1     Running   0          24s   10.244.0.126   new-s2600stq   <none>           <none>
pod/tadk-intel-tadkchart-7594bd49b9-n6hnn     1/1     Running   0          24s   10.244.0.129   new-s2600stq   <none>           <none>
pod/tadk-intel-tadkchart-7594bd49b9-p7dfp     1/1     Running   0          17h   10.244.0.125   new-s2600stq   <none>           <none>
pod/tadk-intel-tadkchart-7594bd49b9-xkj2m     1/1     Running   0          24s   10.244.0.132   new-s2600stq   <none>           <none>
pod/tadk-intel-tadkchart-7594bd49b9-xqgsg     1/1     Running   0          24s   10.244.0.133   new-s2600stq   <none>           <none>

NAME                            TYPE       CLUSTER-IP     EXTERNAL-IP   PORT(S)          AGE   SELECTOR
service/tadk-intel-tadkchart    NodePort   10.101.24.56   <none>        8005:31412/TCP   17h   app.kubernetes.io/instance=tadk-intel,app.kubernetes.io/nam
e=tadkchart

NAME                                    READY   UP-TO-DATE   AVAILABLE   AGE   CONTAINERS   IMAGES                                  SELECTOR
deployment.apps/tadk-intel-tadkchart    10/10   10           10          17h   tadkchart    localhost:30500/tadk-image:v21.08.2     app.kubernetes.io/instan
ce=tadk-intel,app.kubernetes.io/name=tadkchart

NAME                                               DESIRED   CURRENT   READY   AGE   CONTAINERS   IMAGES                                  SELECTOR
replicaset.apps/tadk-intel-tadkchart-7594bd49b9    10        10        10      17h   tadkchart    localhost:30500/tadk-image:v21.08.2     app.kubernetes.io/
instance=tadk-intel,app.kubernetes.io/name=tadkchart,pod-template-hash=7594bd49b9
```

Agenda

- Background
- Problem
- Solution
  - Development Tools: Traffic Analytics Development Kits (TADK)
  - Performance
  - Data
    - Labeling helper
    - Feature selection
- Optimization
- Example usage

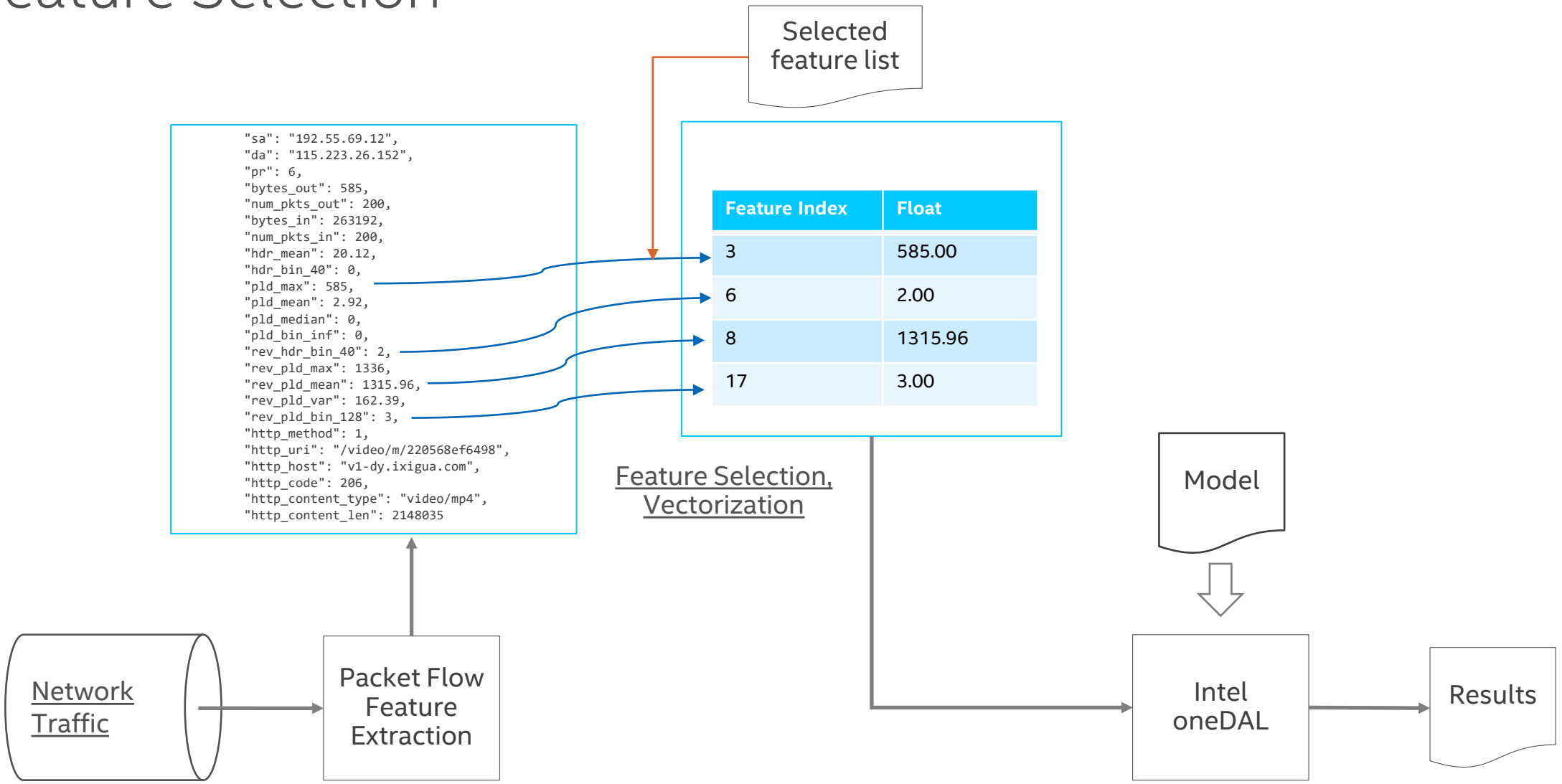
Labeling Helper



Pre-processing

Training, Feature selection, Serialization

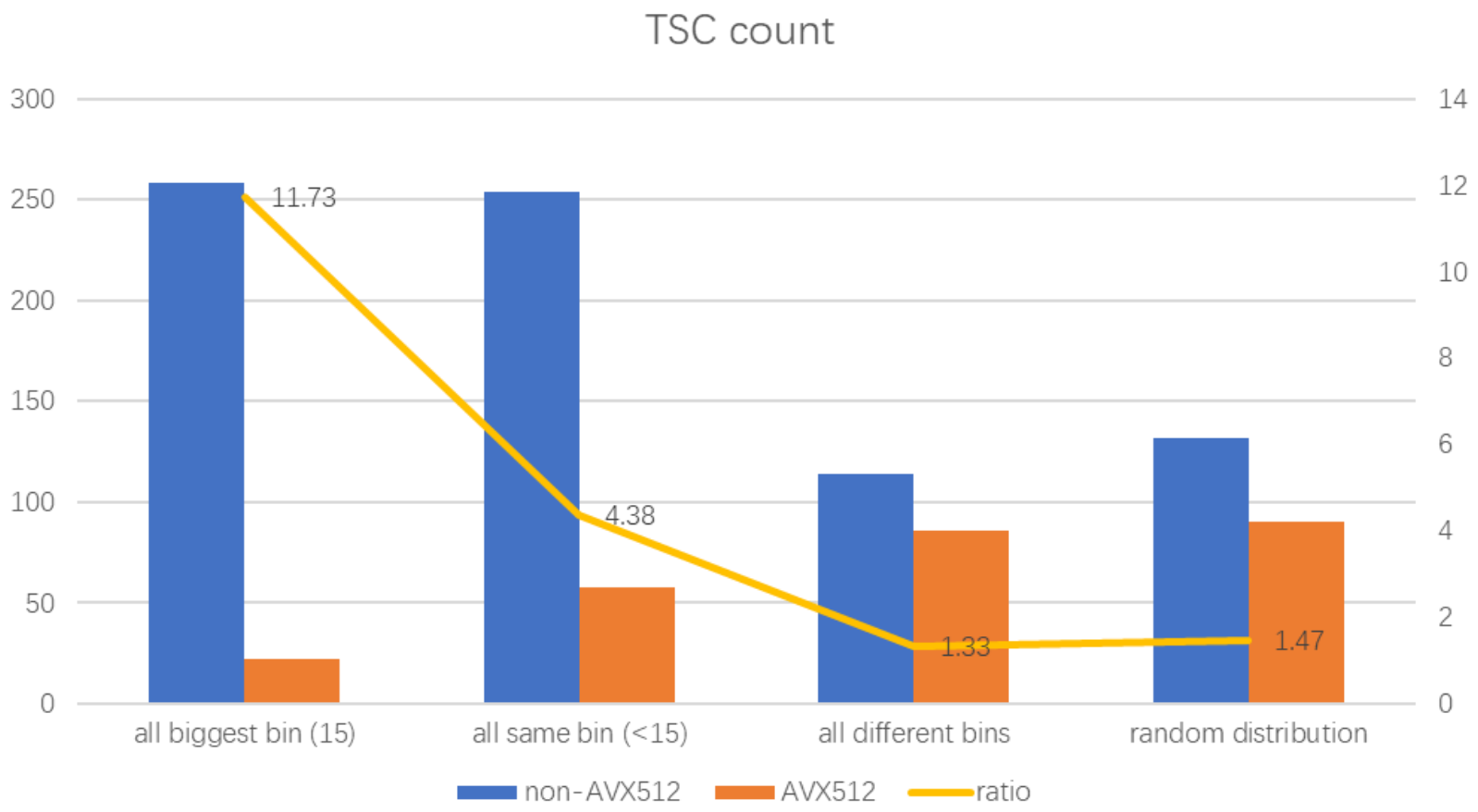
Feature Selection



Agenda

- Background
- Problem
- Solution
- Optimization
  - SIMD: AVX512
- Example usage

Histogram calculation performance w/ AVX512
48 packets, 100000 loops



Conclusion

- TADK
  - A toolbox for building network AI applications
    - Traffic classification
    - Web application firewall
    - Intrusion detection
    - ……
  - High-performance
    - Fully SIMD optimized
  - Flexible
    - Customize
- Refernce samples provided
  - Probe
  - Nginx/ModSecurity plugin
- Easy deployment

Q&A


intel