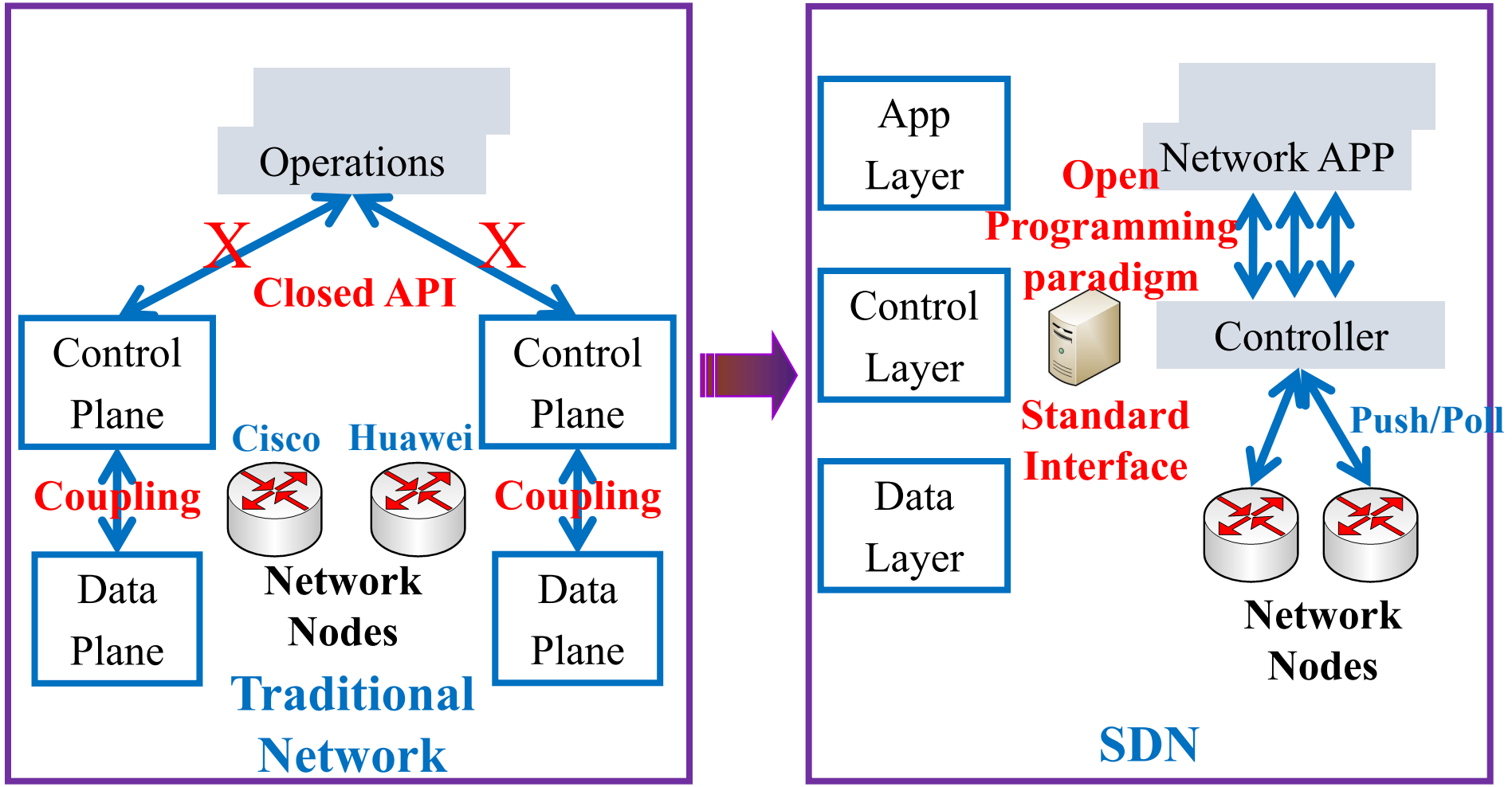
RAID 2018

Control Plane Reflection Attacks in SDNs: New Attacks and Countermeasures

Menghao Zhang[1] Guanyu Li[1] Lei Xu[2] Jun Bi[1]
Guofei Gu[2] Jiasong Bai[1]
[1]Tsinghua University [2]Texas A&M University

清華大學
Tsinghua University

Software Defined Networking



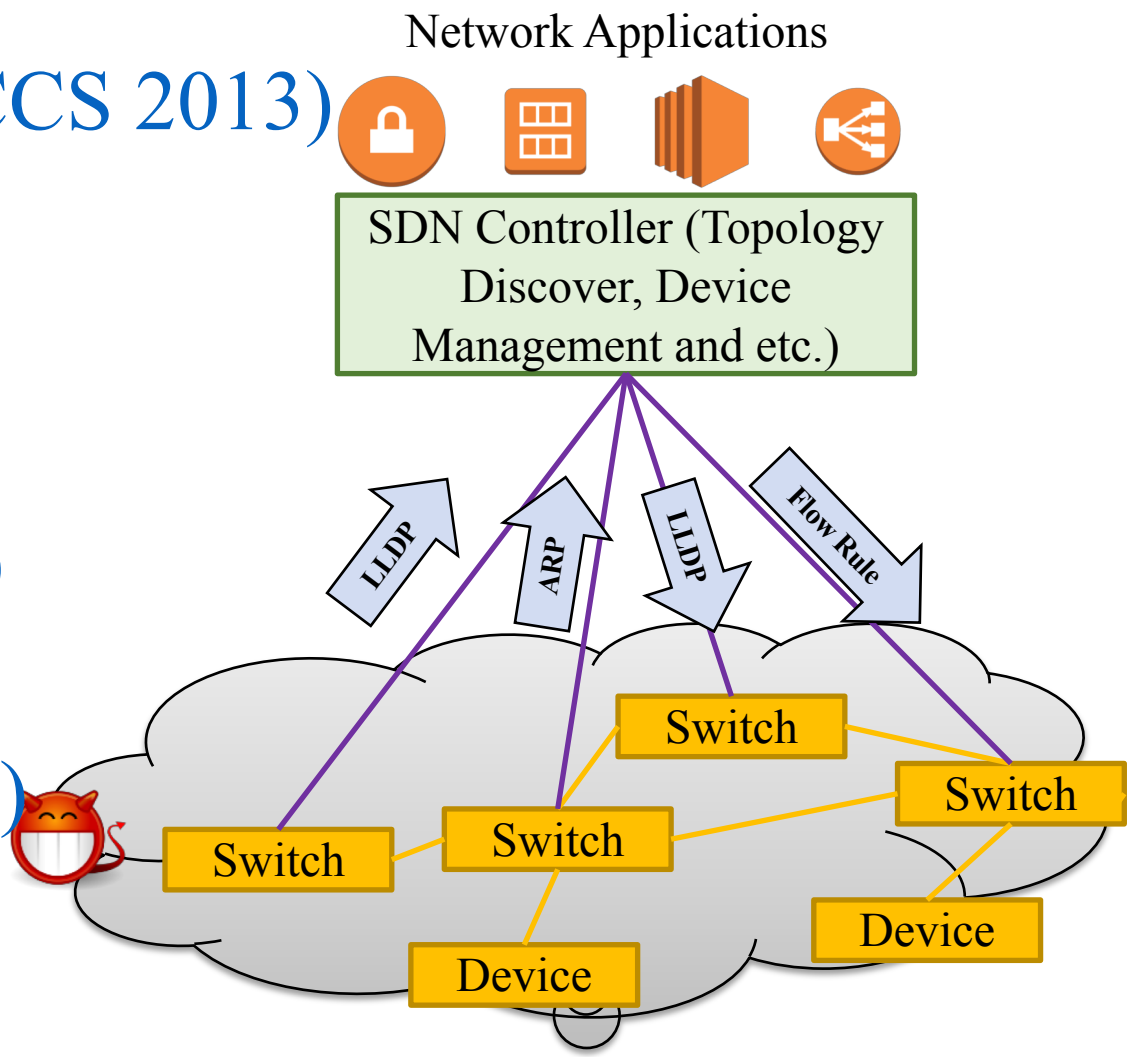
Decoupling of control plane and data plane

Event-driven network applications

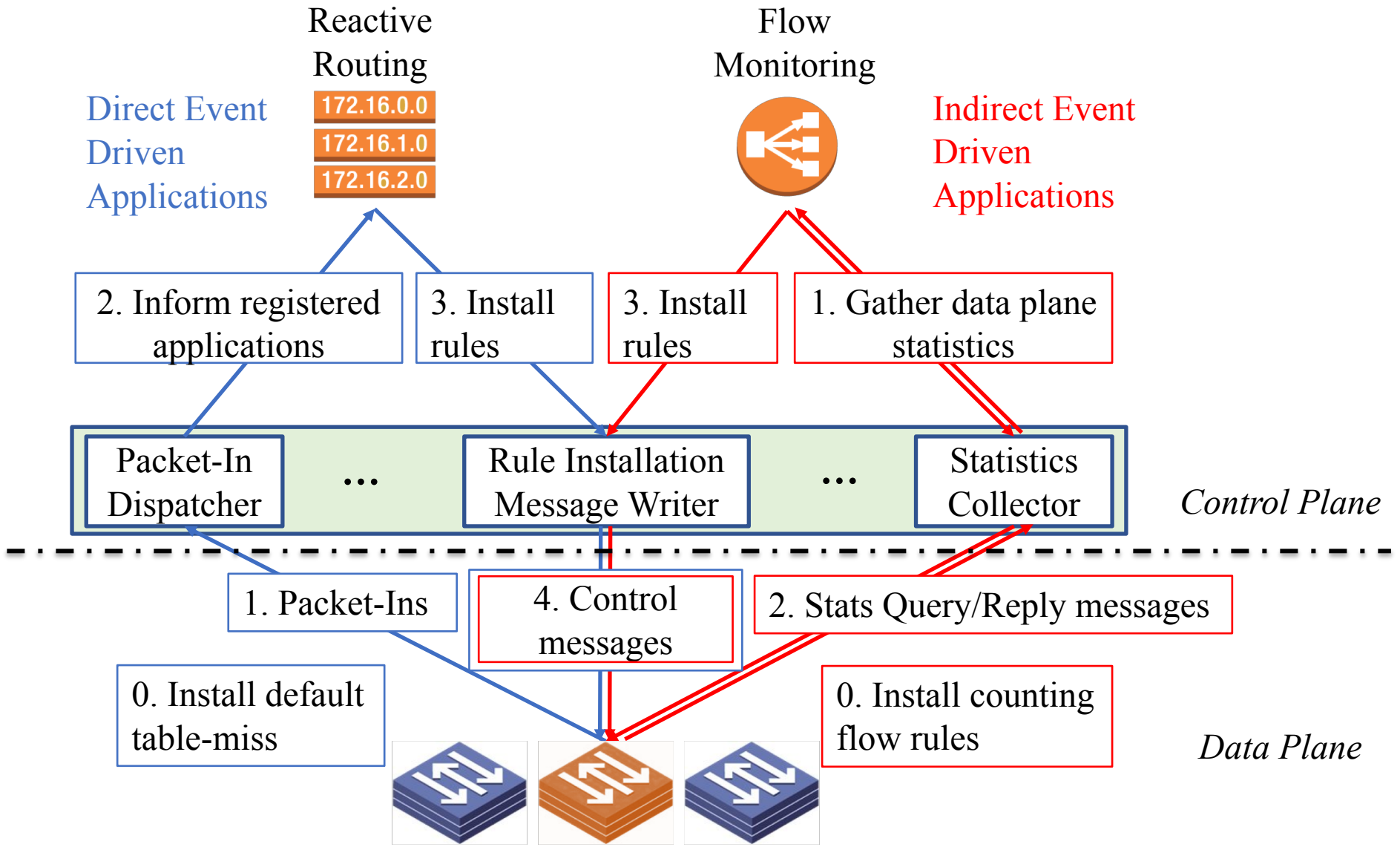
Quick and timely event response

Numerous SDN Attacks

- Data-to-control Plane Saturation (CCS 2013)
  - Controller processing capability
- Topology Poisoning (NDSS 2015)
  - Topology management
- Persona Hijacking (USENIX 2017)
  - Identity binding
- State Manipulation (USENIX 2017)
  - Harmful race conditions



Processing Logic of Data Plane Event



Motivation and Observation

- The processing logic of these two kinds of data plane events could be exploited by the attacker to reflect the control messages towards the switches.
  - Wide usage of direct/indirect data plane events

Table 1: Examples of indirect event driven applications in aca

Category	Application name	Description
Optimization	Hedera [18], MicroTE [19]	SDN controller polls the data plan flow statistics to detect large flows source estimation and flow scheduli
Monitoring	OpenSketch [20], Scream [21], Mozart [22]	SDN controller configures data plar measurement, and then gets the dat tics through a query and report pr
Security	FloodDefender [23], DDoS Detection [24], CloudWatcher [25]	SDN controller monitors the pack each/multiple flow(s) with process extracts traffic features for classifica

Table 2: Examples of indirect/direct event driven applications in industry

	Indirect event driven	**Direct event driven**
OpenDaylight	Statistics Application Load Balancing LACP(Link Aggregation Control Protocol)	Forwarding Firewall L2Switch
ONOS	Port Statistics Performance Metrics Collection	Learning Switch Reactive Forwarding Proxy ARP/NDP
Floodlight	Link Bandwidth Utilization FlowDiff	Firewall Forwarding
Ryu	Traffic Monitor QoS	Switching Hub Firewall Router
Nox	Load Balancing DDoS Detection	FlowACL
Pox	DoS Mitigation	FlowQos

Motivation and Observation

- The processing logic of these two kinds of data plane event could be exploited by the attacker to reflect the control messages towards the switches.
  - Wide usage of direct/indirect data plane events
  - Limited control message processing capability of SDN-enabled hardware switches

Limited flow entries capacity, e.g., 8192

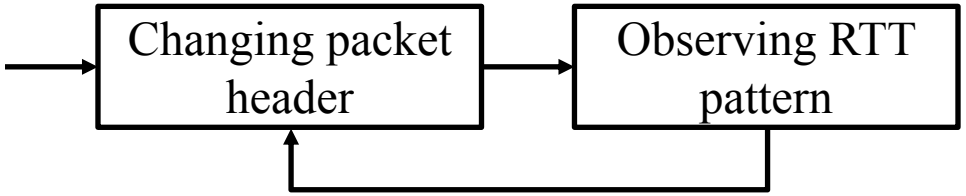
Limited flow rule update capability, e.g., <= 200

Control Plane Reflection Attacks

- Threat Model
  - Attacker controls a collection of compromised hosts or VMs (e.g., by malware infection) in SDN-based network
- Target
  - Limited control message processing capability of SDN-enabled hardware switches, to further degrade the network performance
    - Vector 1: Table-miss Striking Attack
    - Vector 2: Counter Manipulation Attack

Vector 1: Table-miss Striking Attack

- Enhanced attack vector from data-to-control plane saturation attack
  - Random packet generation method to trigger table-miss
  - Utilizing probing and triggering phases to determine the forwarding grain, then could strike table-misses in a more accurate and cost-efficient manner
    - Probing phase
    - Triggering phase
      - craft attack packet stream based on probed grains



Vector 2: Counter Manipulation Attack

- Probing phase: Using three types of packet streams to accurately infer the concrete usage of the indirect data plane events
  - Timing probing packets
  - Test packets
  - Data plane stream
- Triggering phase: Permuting the packet interval and packet size of each flow, to deliberately manipulate the counter value to trigger a large number of Flow-Mods every period

Vector 2: Counter Manipulation Attack(cont.)

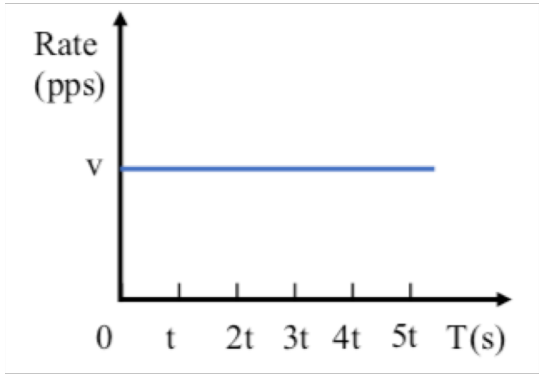
- Timing probing packets
  - Goal: Measuring the workload of software agent of a switch
  - Properties
    - Hitting the table-miss flow rule in the switch, and triggering the operations of the corresponding applications (e.g., Flow-Mod or Packet-Out).
    - Evoking a response from the SDN-based network to compute the RTT for each timing probing.
    - Being sent in an extremely low rate (10 pps is enough), and putting as low loads as possible to the switch software agent.
  - Alternatives: ARP request/reply, ICMP request/reply, TCP SYN or UDP (details in paper)

Vector 2: Counter Manipulation Attack(cont.)

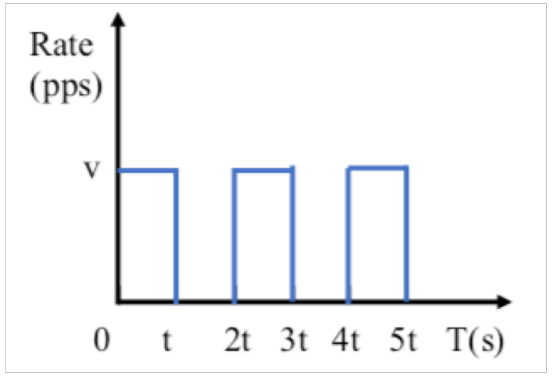
- Test packets
  - Goal: Strengthening the effects of timing probing packets by adding extra loads to the software agent of the switch
  - Property
    - Incurring appropriate loads on switch CPU
  - Alternative
    - Packets with a random destination IP address and broadcast destination MAC address

Vector 2: Counter Manipulation Attack(cont.)

- Data plane stream
  - Goal: obtaining more advanced information such as the specific conditions to trigger the control of indirect event-driven applications
  - Properties
    - Going directly through the data plane, without explicit interaction with the controller
  - Templates



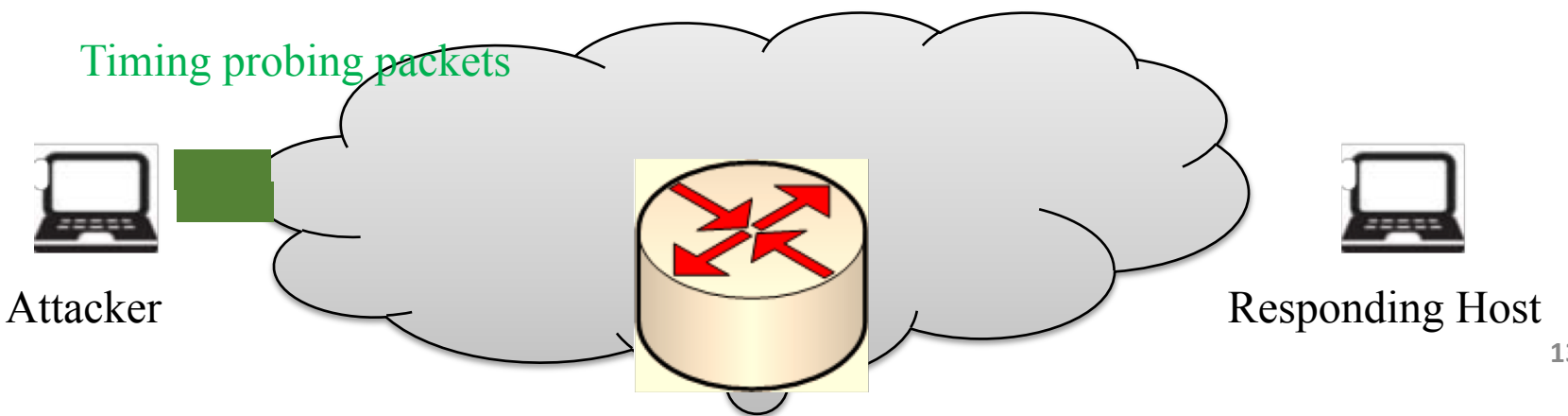
Template 1: Data plane stream with steady rate



Template 2: Data plane stream with 0-1 rate

Vector 2: Counter Manipulation Attack(cont.)

- Insight and observation
  - different kinds of downlink messages have diverse expenses for the control channel
    - Flow-Mod >= Statistics Query >> Packet-Out.
  - Timing-based side channel effect



Attack Evaluations

- 5 Test Applications
  - Reactive Routing
    - five-tuples grained forwarding policy
  - Flow Monitoring
    - Heavy hitter, Microburst, PIAS, and DDoS Detection
- Controller: Floodlight
- Switch: Pica8 P-3290
  - Widely used in academia/industry, supports many advanced OpenFlow data plane features, and applicable for other switch brands

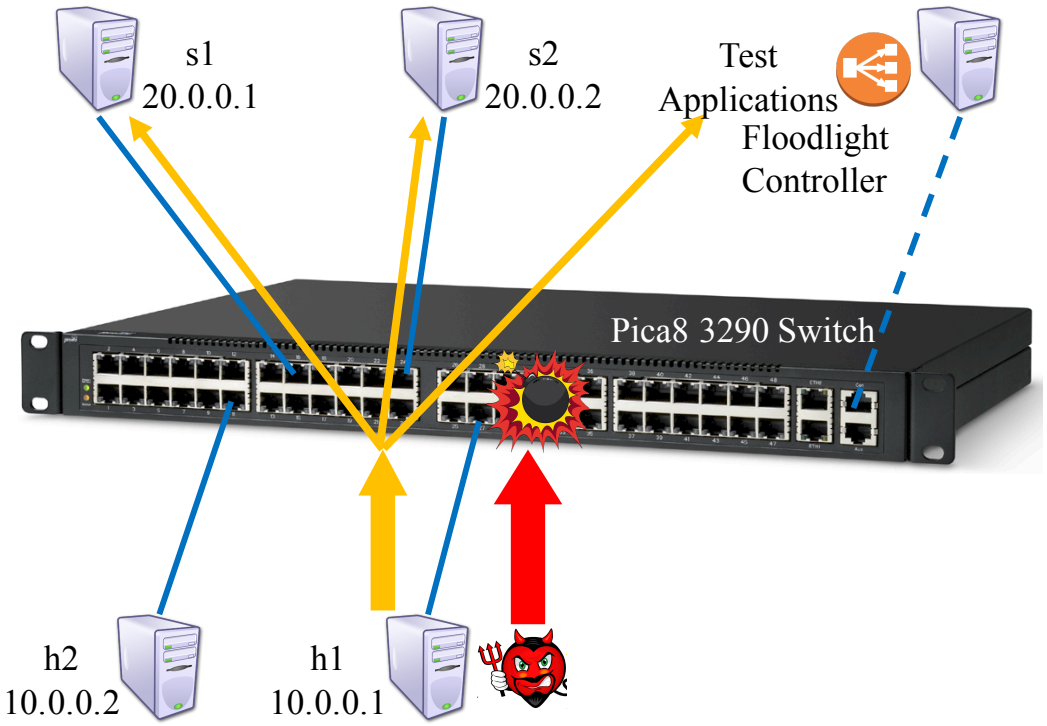
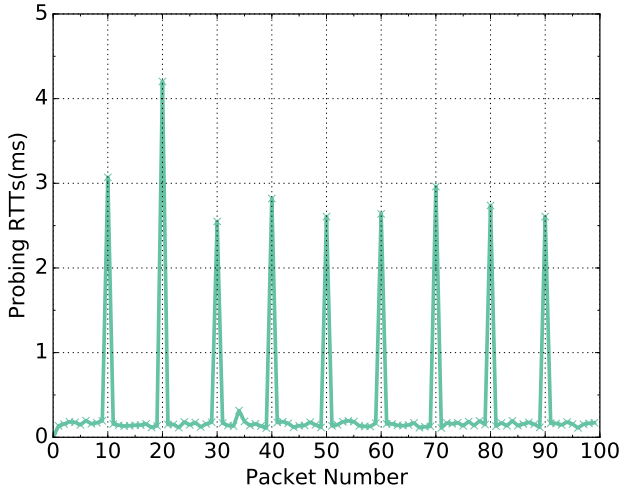
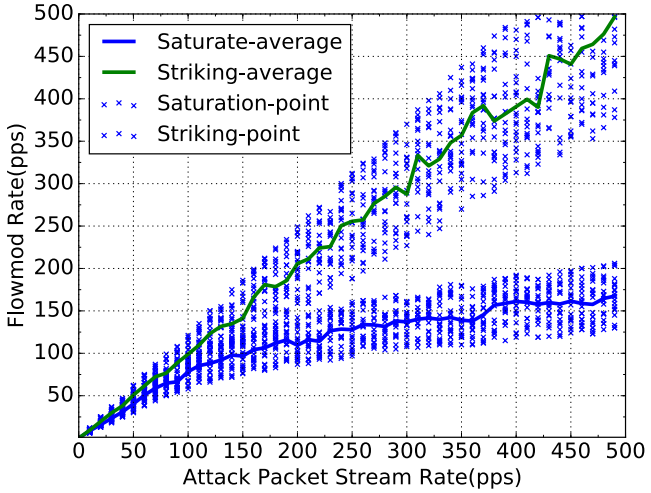


Table-miss Striking Attack

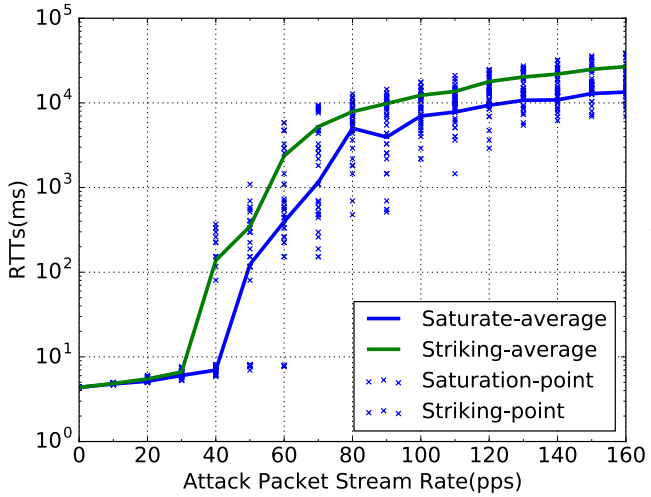


RTTs for Reactive Routing when UDP port changes

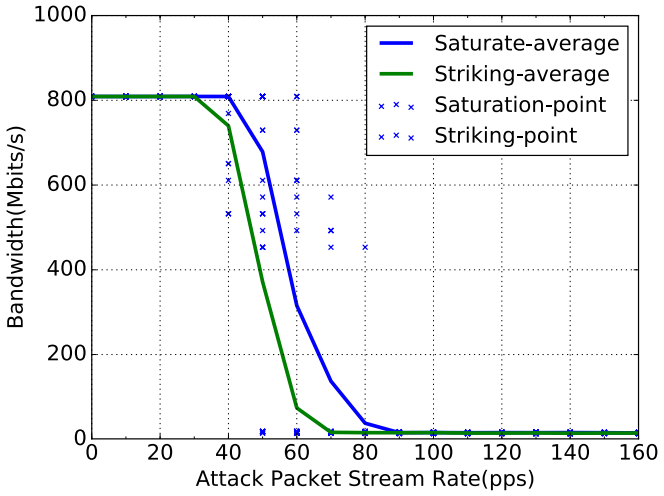


Attack efficiency for Reactive Routing

Cost-efficient and Powerful



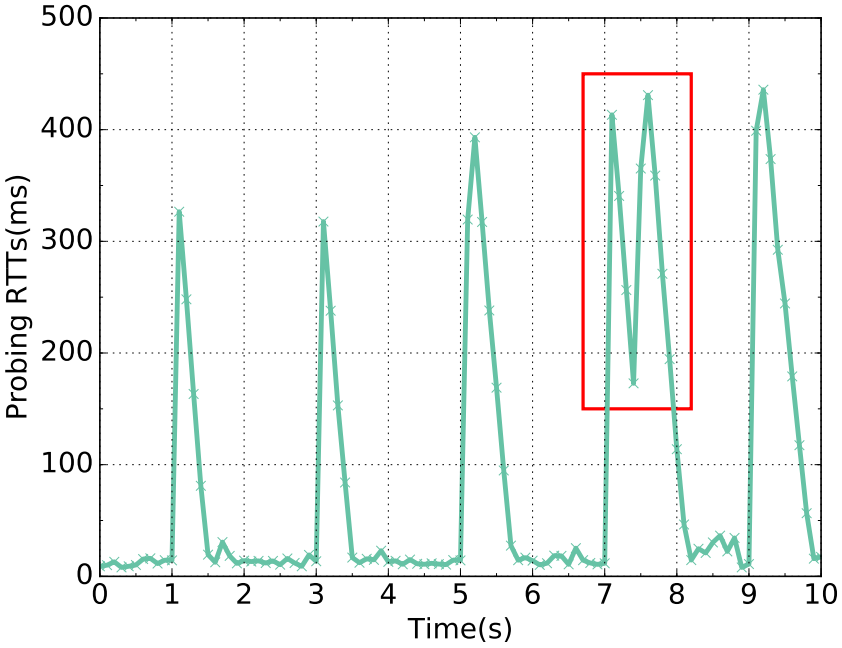
RTTs for normal users under the saturation attack and the striking attack



Bandwidth for normal users under the saturation attack and the striking attack

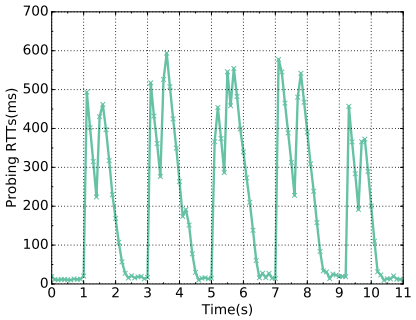
Counter Manipulation Attack

- Parameter settings
  - test packets: 300 pps
  - timing probing packets: 10 pps
- Inference
  - Flow Monitoring-based applications poll the switch for statistics every 2 seconds.
  - The double peaks in red rectangle (double-peak phenomenon) denote two expensive downlink messages are issued successively.

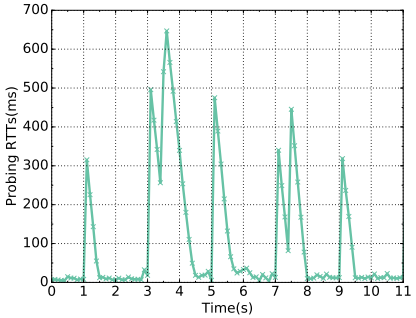


Counter Manipulation Attack(cont.)

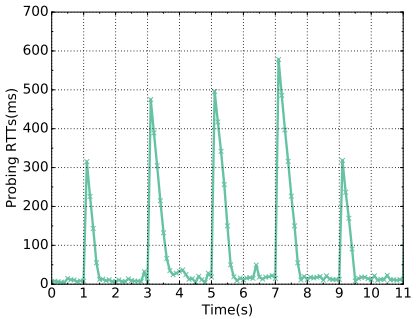
- More inference behind double-peak phenomenon



Timing probe RTTs patterns 1

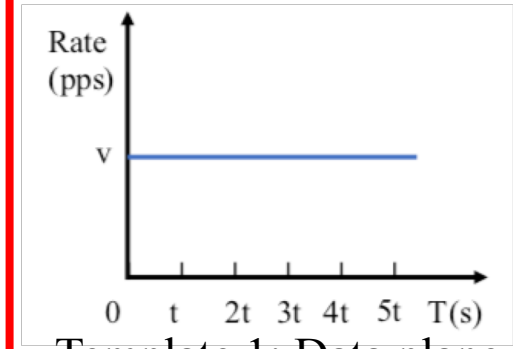


Timing probe RTTs patterns 2

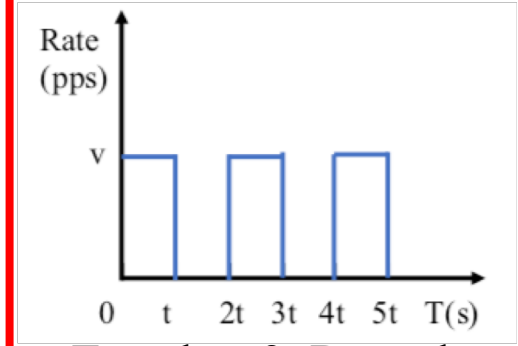


Timing probe RTTs patterns 3

Question: Could we reveal more confidential information with the joint trials and analysis of data plane stream and double-peak phenomenon?



Template 1: Data plane stream with steady rate



Template 2: Data plane stream with 0-1 rate

Counter Manipulation Attack(cont.)

- Application control logics revisiting
  - 4 fundamental dimensions
  - (Packet Number, Packet Byte) x (Volume-based, Rate-based)
- Case study
  - If the attacker obtains RTT pattern 1 with the input of data plane stream template 1, where v is a big value, and obtains RTT pattern 2 with template 2, where v is also a big value, she/he could determine that packet number volume-based statistic calculation method is adopted.
  - Key insight: packet number volume-based statistic calculation approach is sensitive to stream with a high pps.

Counter Manipulation Attack(cont.)

- Generalizations

Microburst

	Volume-based	Rate-based
Packet Number	Template1(v↑, p)→ patterns 1 Template2(v↑, p)→ patterns 2	Template1(v↑, p)→ patterns 3 Template2(v↑, p)→ patterns 1
Packet Byte	Template1(v, p↑)→ patterns 1 Template2(v, p↑)→ patterns 2	Template1(v, p↑)→ patterns 3 Template2(v, p↑)→ patterns 1

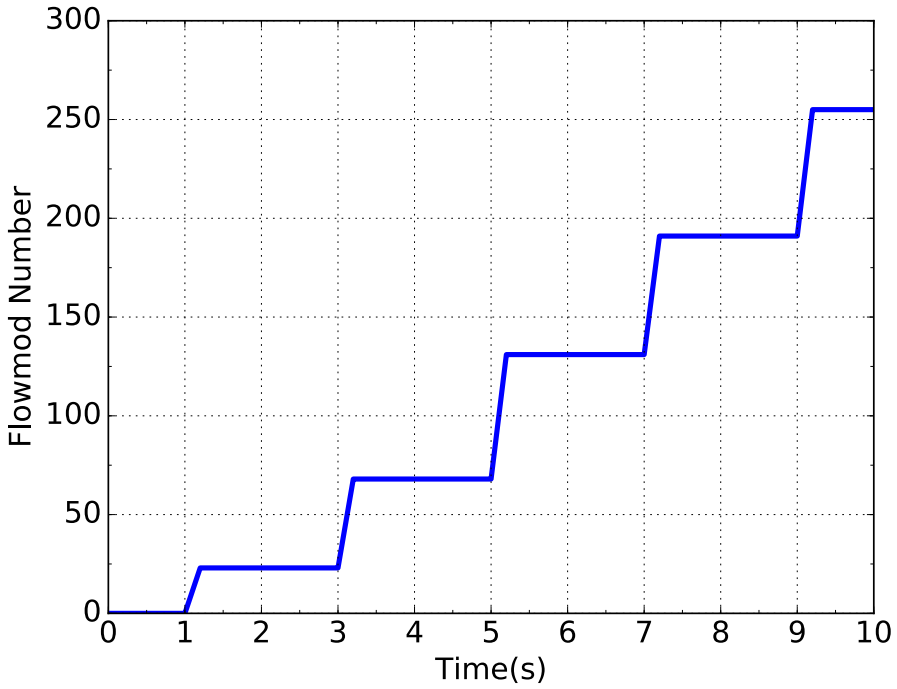
Heavy Hitter
PIAS
DDoS Detection

DDoS Detection

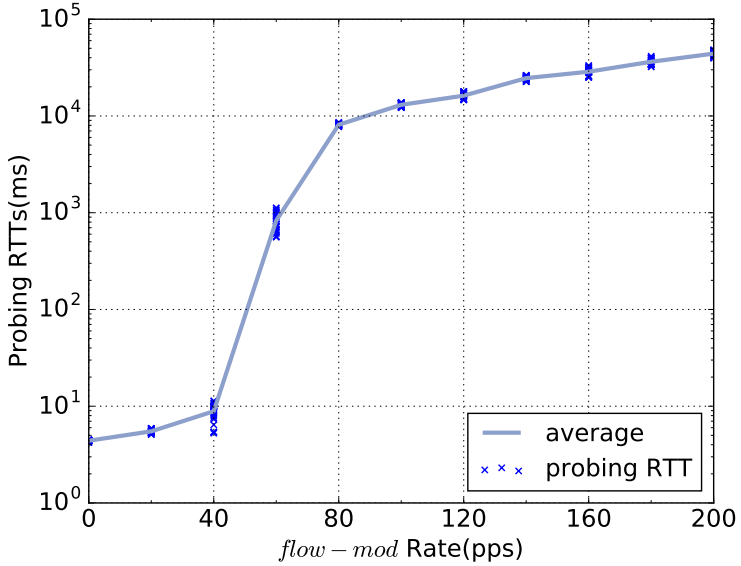
- Threshold inference
  - Variations of v and p in a binary search manner to infer the critical value of volume or rate.

Counter Manipulation Attack(cont.)

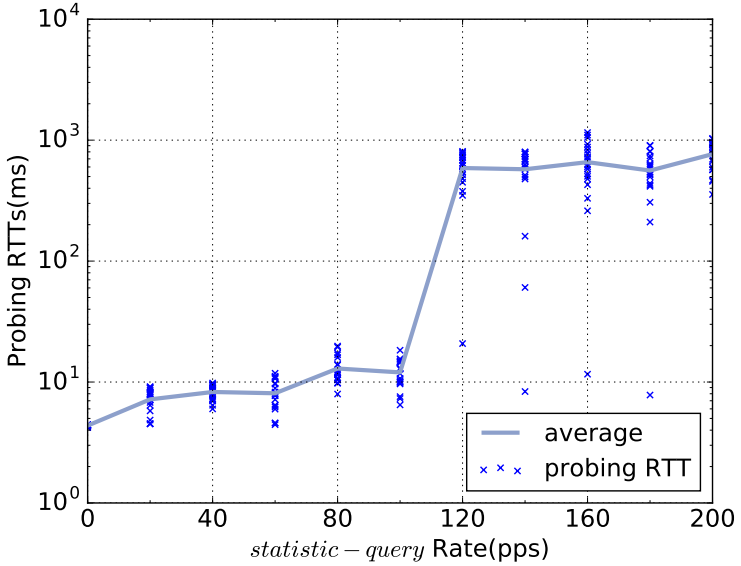
- A case study for triggering phase
  - PIAS with 3 level priority
  - Attack procedure
    - Initiating 10 new flows per second, and carefully setting the sent bytes of each flow in each period (2s) bigger than the critical value we probed.
  - Attack effect
    - A number of Flow-Mod messages are issued to the switch when statistic query/reply occurs.



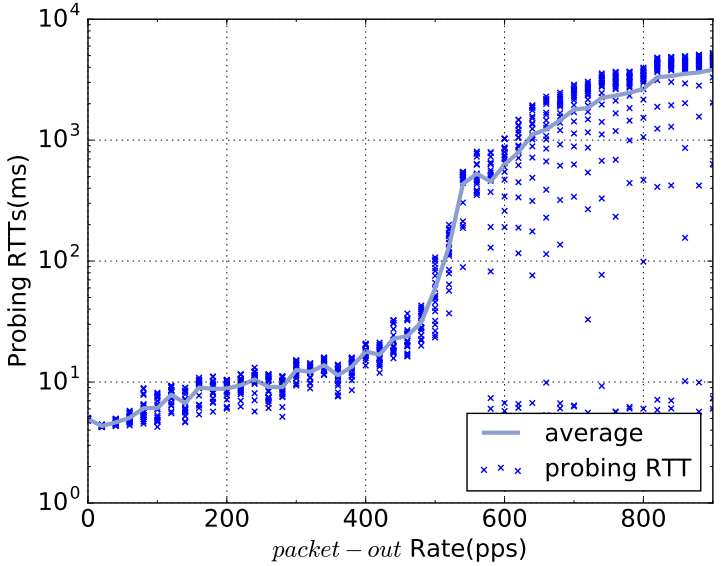
Attack Fundamentals and Analysis



Timing probe RTTs as Flow-Mod rate varies



Timing probe RTTs as statistic query rate varies



Timing probe RTTs as Packet-Out rate varies

- Different downlink messages have diverse expenses for the downlink channel, and all of the three scenarios encounter a significant **nonlinear jump**.

Attack Fundamentals and Analysis(cont.)

- The background traffic can boost the accuracy of probing phase sometimes.
  - A moderate rate of background traffic would amplify the probing effect.
    - The effect of background traffic is somewhat like the role played by test packets, and it would put some baseline loads to the switch protocol agent, which would make the probing more accurate.
  - An excessively high rate of background traffic would make the RTT patterns random and irregular.
- The background traffic also boost the attack effect of trigger phase.

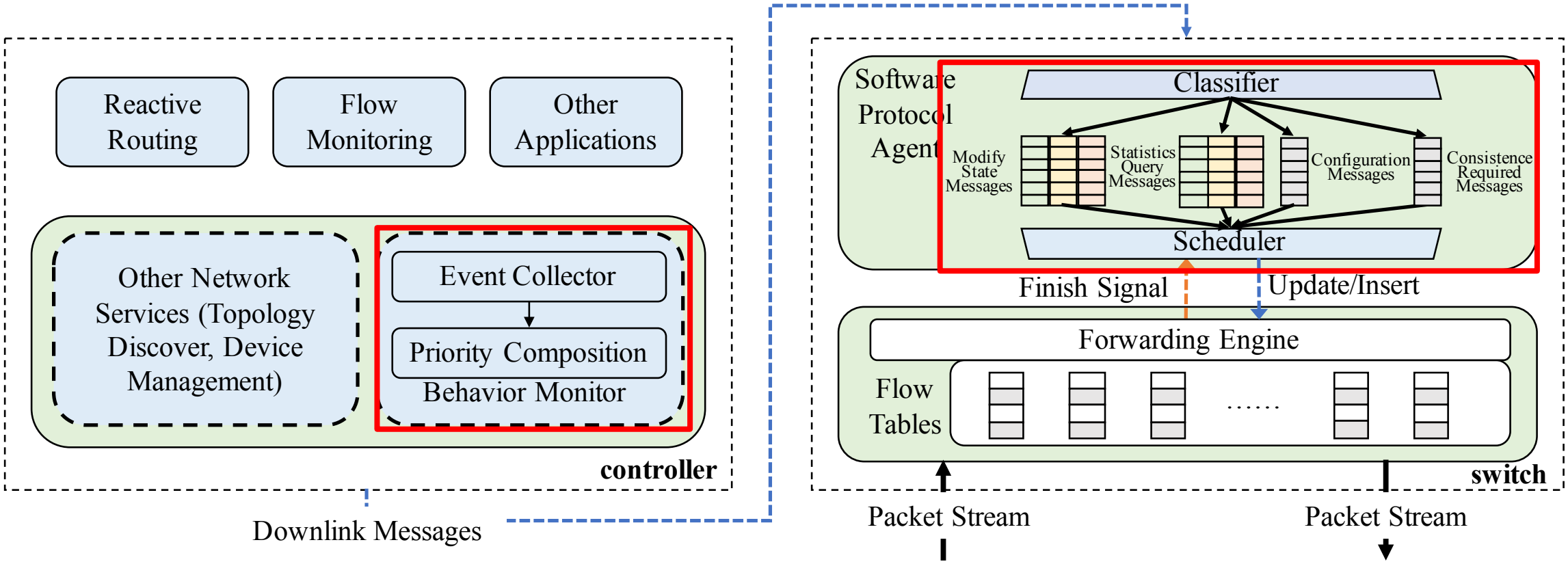
Defense Alternatives

- Fundamental reason
  - The performance of existing commodity SDN-enabled hardware switches could not suffice the need of the SDN applications.
- Alternatives
  - limiting the use of dynamic features for network applications
  - limiting the downlink message transmission rate directly in the controller
  - Adding some latency to random downlink messages

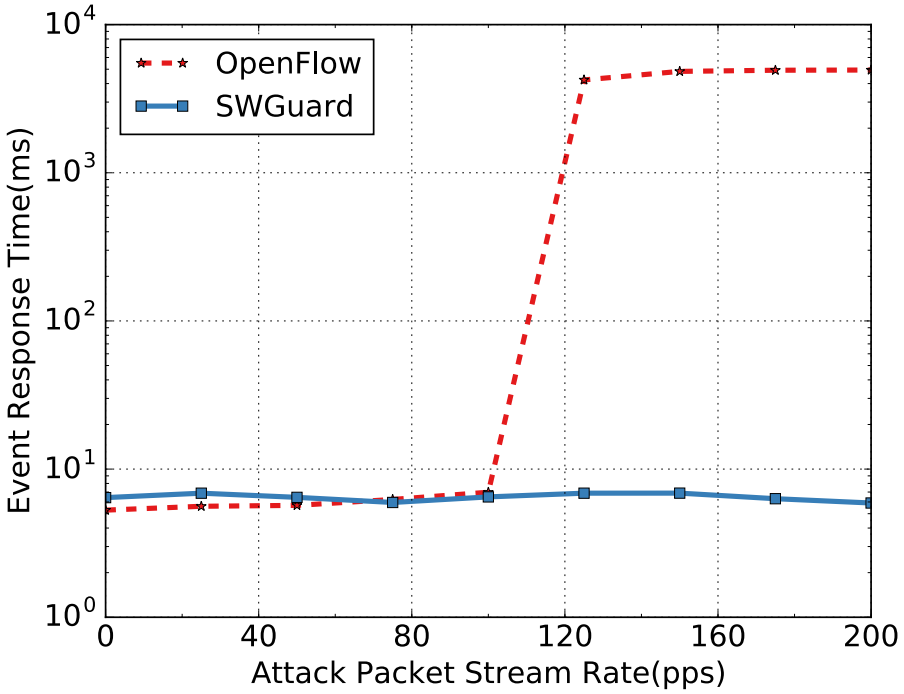
Our Approach: SWGuard

- Key idea
  - Discriminate good from evil, and prioritize downlink messages with discrimination results
- Key design points
  - Multi-queue scheduling strategy
  - Novel monitoring granularity: Host-application pair(HAP)

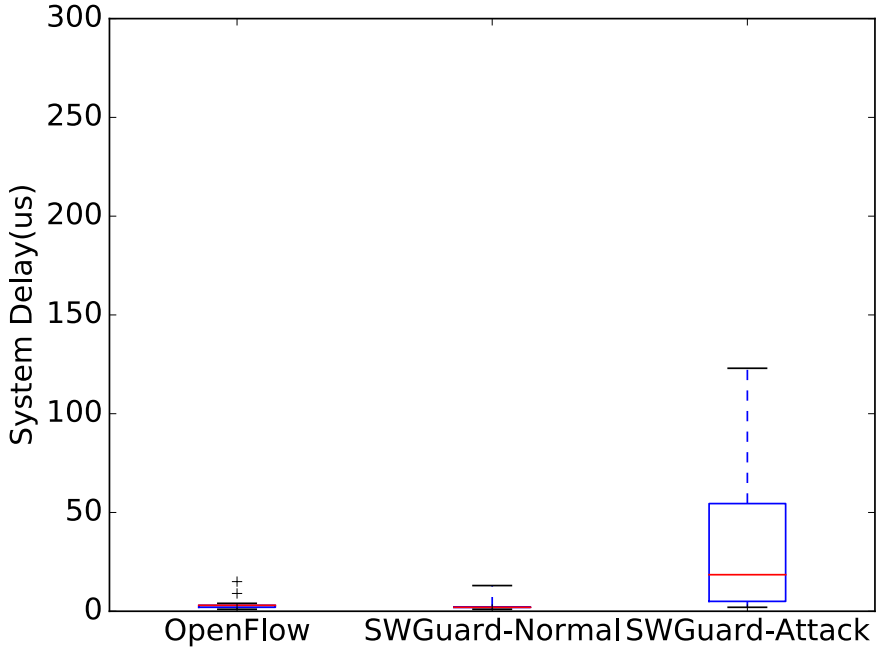
SWGuard Framework



Defense Evaluation



Defense Effect



Defense Overhead

SWGuard provide effective protection for control messages with only minor overheads.

Conclusion

- The processing logic of direct/indirect data plane events can be exploited combined with the limitation of SDN-enabled hardware switches.
- Two control plane reflection attacks can hurt the usability of SDN in an efficient, stealthy and powerful way.
- New security extensions to SDN, SWGuard can mitigate the threat effectively.

Thanks!
Q&A

zhangmh16@mails.tsinghua.edu.cn

Workload test on SDN components

- We measure the resource usage of the hardware switch and the controller, and find that the CPU usage of the switch could reach above 90% at the point of the nonlinear jump, while the memory usage of the switch, the CPU and memory usage of the control server is relatively low (at most 30%).