Application-Awareness in SDN

Zafar Qazi*, Jeongkeun "JK" Lee, Tao Jin!, Gowtham Bellala, Manfred Arndt#, Guevara Noubir+
*Stony Brook University, HP Labs, !Qualcomm Research, #HP Networking, +Northeastern university

Rapid deployment of new SDN service requires application-awareness

- Trend: mobile and cloud bring in myriads of new applications
- SDN APIs of today are capable of L2/3/4-based virtualization, but **lack L7-awareness**

Existing solutions

QoS marking by application: untrusted

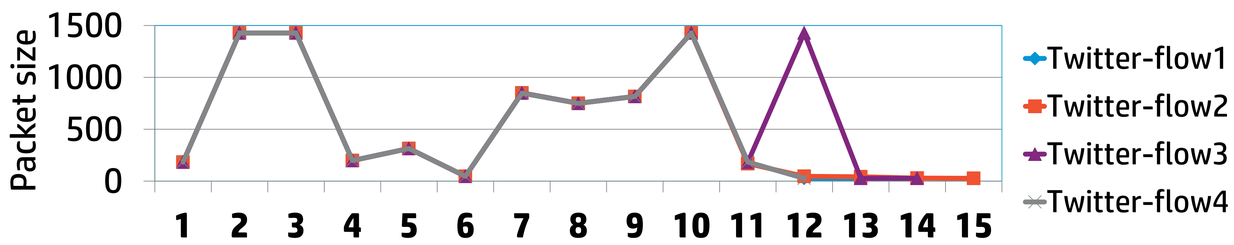
Port-based classification: too coarse grained

Deep Packet Inspection (DPI)

- Exploit signatures in packet payload, widely used
- Require intensive human effort to maintain signature DB
- Computationally expensive, not scalable
- Low accuracy with encryption

Machine Learning (ML) based approach

- Exploit flow features: e.g., sizes of first "N" pkts



- Lightweight, scalable, handles encrypted traffic
- **Challenge: fine-grained ground truth** for ML training

Our Solution: *Atlas*

Crowd-sourcing ground truth collection

- Quickly detect new apps and rapid app updates

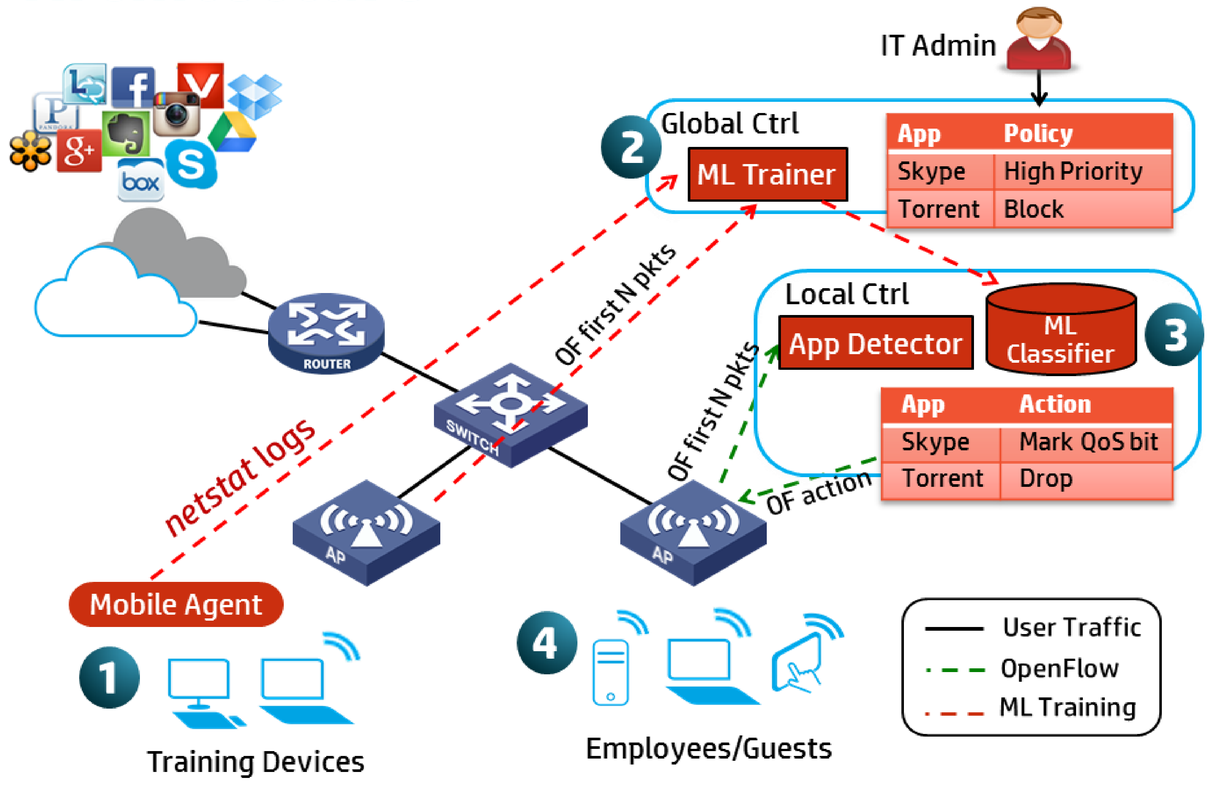
Fine-grained app-detection, not just traffic type

- E.g., Differentiate Google Talk vs. Kik Messenger

Scalable & efficient

- Extend OpenFlow for flow feature collection
- Pre-program application policy, reduce control overhead
- Emerge app-awareness into network edges

Architecture



1. Mobile Agents on a few devices send *netstat* logs
   - Application name & N-tuple
2. ML trainer maps application name to flow features
   - Extended OpenFlow stats: first "N" packet sizes
3. Train and distribute ML classifier
4. Application on any device properly identified by AP
   - Used to enforce per-app policy

Prototype

- >10 **Android** phones/tablets on HP Labs WLAN
- 4 weeks of data for ML training & evaluation
- **Avg 96%** detection accuracy over >30 popular apps
- *c5.0* classifier, 1.4M flows/sec on 3.3GHz CPU