Software Defined Privacy-Preserving Measurement Instrument and Services

Yan Luo, Univ. of Massachusetts Lowell
Cody Bumgardner, Univ. of Kentucky
Gabriel Ghinita, Univ. of Massachusetts Boston
Michael McGarry, Univ. of Texas El Paso
Outline

- Overview of IRNC AMIS Project
- Design of IRNC AMIS Instrumentation Framework
  - Overview
  - AMIS measurement functions
  - Plugin management
  - Privacy preservation
  - Data analytics interface
- Q&A
Objectives

- 40+Gbps flow-granularity network measurement instrument
- Software defined measurement
- Preserving privacy of network flow info
- In-depth flow analytics

NSF Funded Project Team:

- Yan Luo, PI, University of Massachusetts Lowell
- Cody Bumgardner, Co-PI, University of Kentucky
- Gabriel Ghinita, Co-PI, Univ. of Massachusetts Boston
- Michael McGarry, Co-PI, University of Texas El Paso
Overview of IRNC AMIS Framework

**UTEP**
1. Configure and manage measurement tasks
2. Annotate instrument data with auxiliary data for analytics
3. Measurement data visualization and analytics

**UKY**
1. Config management
2. Dispatch mtask to AMIS nodes
3. Query processing on netflow records

**Privacy Protection**

**UMB**
1. Syntactic Privacy
2. Differential Privacy with BigData tools

**AMIS Measurement Functions**
1. Netflow generation
2. Packet tracing
3. Packet validation
4. Software defined measurement
5. Distributed measurement

**Web-based Network Data Visualization and Analytics**
AMIS Measurement Functions

Current functions

• Netflow
  ▪ Generate NetFlow v5 record and push to NetFlow collector
  ▪ Analyze traffic characteristics from NetFlow records

• Packet Tracing
  ▪ Trace the occurrence of flows/packets on links monitored by (distributed) AMIS instrument

• Packet Validation
  ▪ Validate the packets on the link (IP level validation) and report good/bad packets

Ongoing Work

• Software Defined Measurement
  ▪ P4 based protocol parser and matching table
  ▪ Measurement function composition

• Distributed and Collaborative Measurement
1. NetFlow: OVS  
2. Packet Validation: PCAP  
3. Packet Tracing: PCAP
IRNC AMIS: Distributed Control System

- Project components managed through a distributed control system comprised of agents and plugins.
- Plugins implemented to manage control of data processing components.
- Agents allow for anonymous operation with centralized control.
Device View

Key:
- Control and Config Data
- Measurement Data

Service Module View:
AMIS Device

Central Services
- Agent Controller
  - eQueue_0

Direct Client

Probe Agent

AMIS Device
- mQuery_0
- mQueue_0
- mQuery_1
- mQueue_1
- lQuery_1
- lQueue_0
- lQueue_1

Measurements
- Local Processing
System View

Key:
- Control and Config Data
- Measurement Data

Remote Control and Aggregation

Device 0
- IQueue_0
- Probe Agent

Device 1
- IQueue_0

Central Controller
- eQueue_0

Agent Controller
- HTTP
- Dashboard /API

Query
- eQueue_1

Direct Client
AMIS: Supported Privacy Modes

Privacy

Online Mode

Syntactic Privacy
Best-effort Protection
Fast Performance
Flexible Query Capabilities
Low Storage Requirements

Offline Mode

Differential Privacy
Strong Protection
Moderate Performance
Statistical Queries Only
High Storage Requirements
AMIS Privacy: Online Mode

Online Privacy Module

- Flow Window
- Fast Indexer
  - Geo Mapping
  - AS Mapping

k-anonymizer Engine
- Hilbert Fractals
- Linear Runtime

Sanitized flows

Sensitive Raw Flows

Rabbit MQ

Sensitive
Sanitized

Student
Network Engineer
Researcher
AMIS Privacy: Offline Mode

Offline Privacy Module

- Differential Privacy Engine
- Query Optimizer
- Hadoop Analytics

Sensitive
- Raw Flows
- Rabbit MQ

Sanitized
- Sanitized Query Results

Sensitive
- HBase Repository

IRNC AMIS, Internet2 TechX, 2016
Empower understanding of network exchange link utilization
- who is communicating with who? what types of data are they communicating?
- which institutions are utilizing the exchange links?

Empower network management (autonomous?)
- **FCAPS** model of network management
  - **F** – fault detection and correction
    - troubleshooting, monitoring, anomaly detection
  - **C** – configuration and operation
  - **A** – accounting and billing
  - **P** – performance assessment and optimization
    - performance monitoring, capacity planning
  - **S** – security assurance and protection
Most of our visualizations are driven by Netflow records.

We annotate Netflow records with AS, geographic, and application data, *Netflow+*.

We turn Netflow+ records into data viz objects to be rendered in the browser.
Thank You!

Q&A
Visualization of Anonymized Data

Hyper-Rectangle View

Analytics View

IRNC AMIS, Internet2 TechX, 2016