Servers in Action: Towards Distributed Traffic Measurement in Data Centers

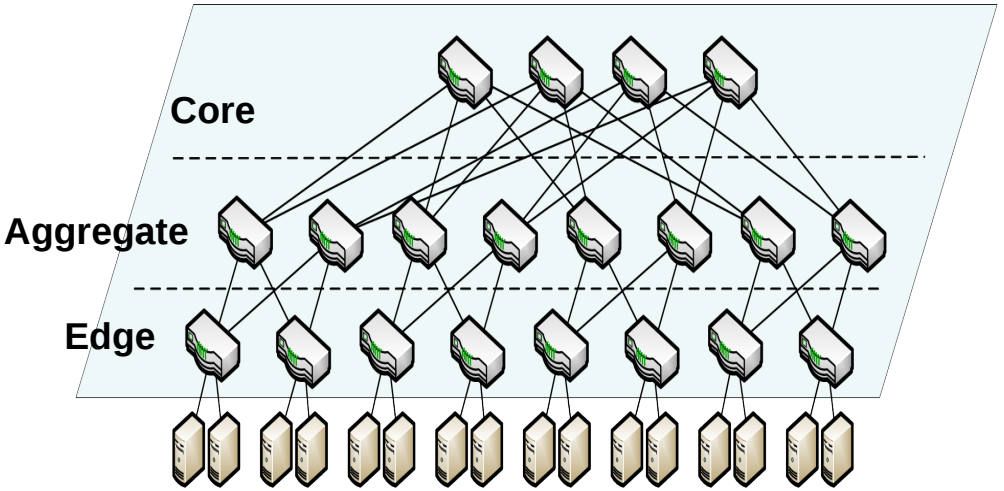
Praveen Tammana & Myungjin Lee
School of Informatics
University of Edinburgh

MSN'14 - Cosener's House, Abingdon

Network Measurement in Data Centers

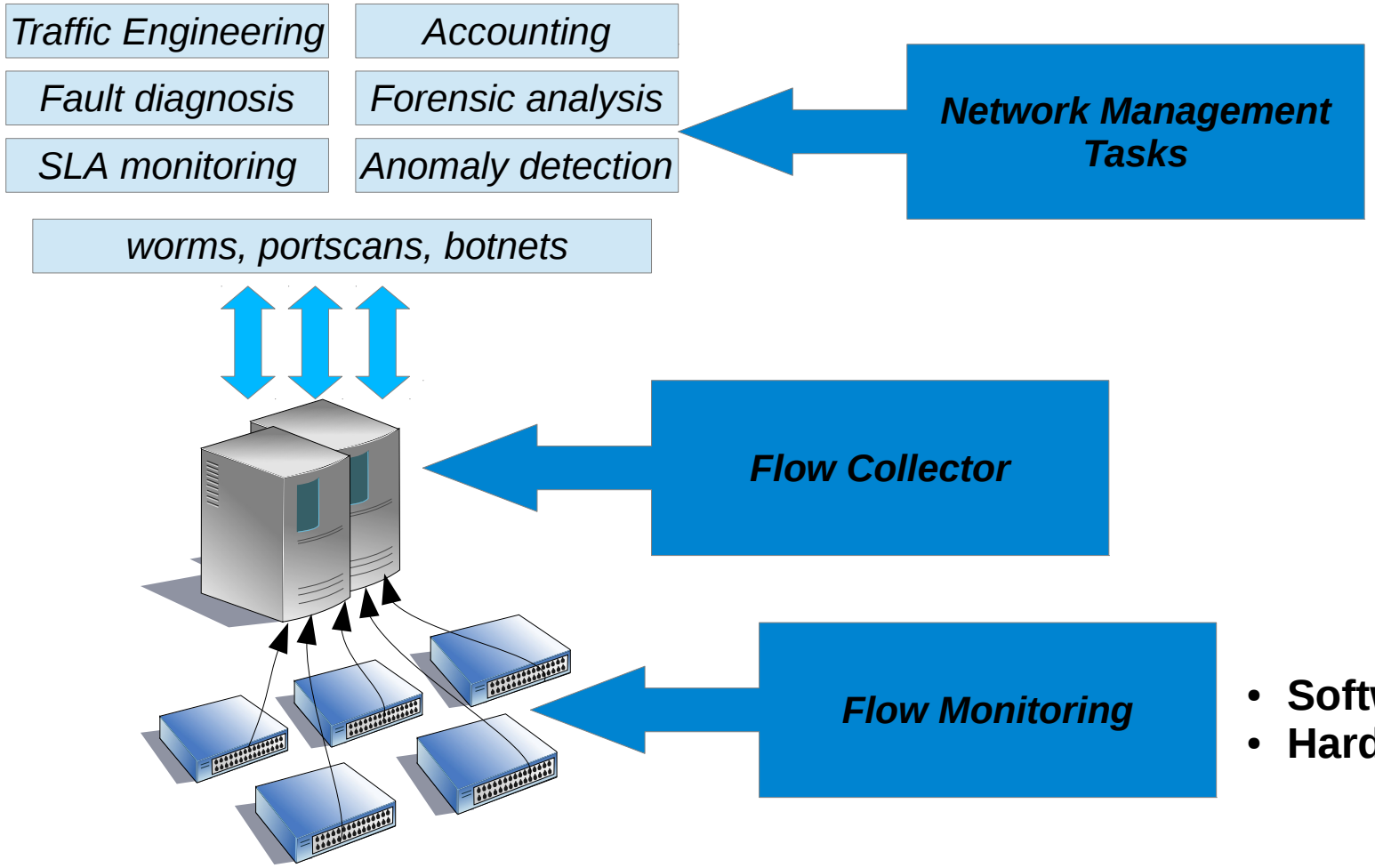
- Data center applications
  - High Bandwidth
  - Latency sensitive
- Network Management
  - Control
    - Routing, Access control
  - Measurement
    - Traffic engineering, Security applications
    - Basic requirements
      - Accuracy, Scalability
    - Data center requirements
      - Programmable, Responsive, Evolvable

Data Center Measurement Requirements

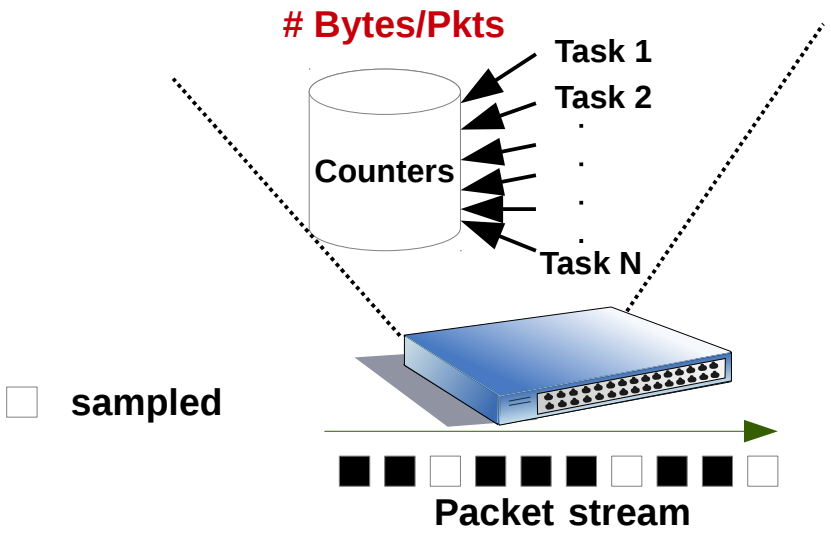


- **Responsiveness** : Quick control loop decisions (Heavy flow scheduling)
- **Programmability**: Adaptable to dynamic workloads
- **Evolvability:** Software based measurement module (bloom filter, trie, hashtable)

Network Measurement Framework



Software Based - Flow Monitoring

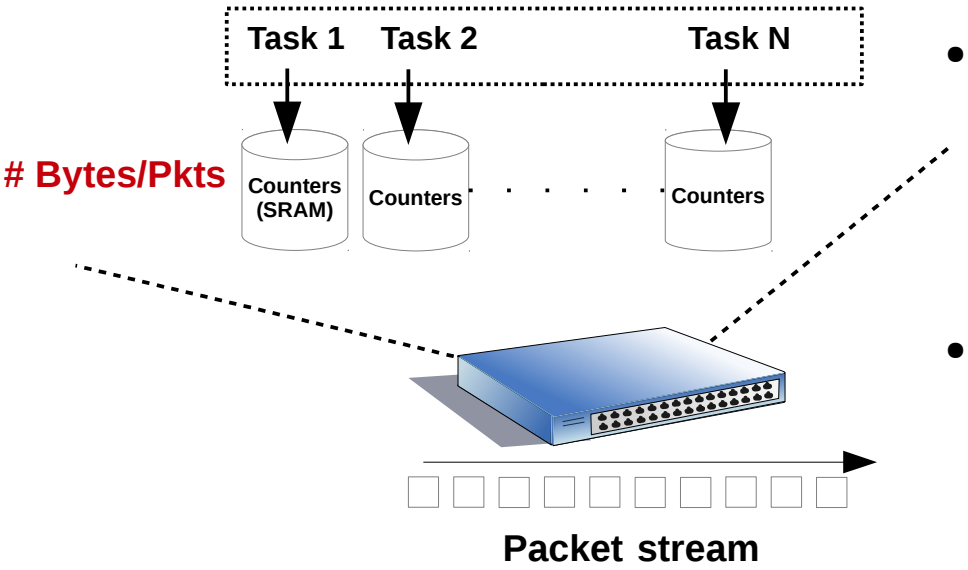


- **Sampling**
  - NetFlow, sFlow
  - High traffic rates compliance with limited switch resources (SRAM, CPU)
- **Problem**
  - **Not Accurate (Basic Requirement)**
    - ➔ Flow coverage and accuracy are compromised.
    - ➔ Not suitable for management tasks that requires fine grained flow details.

Management Tasks

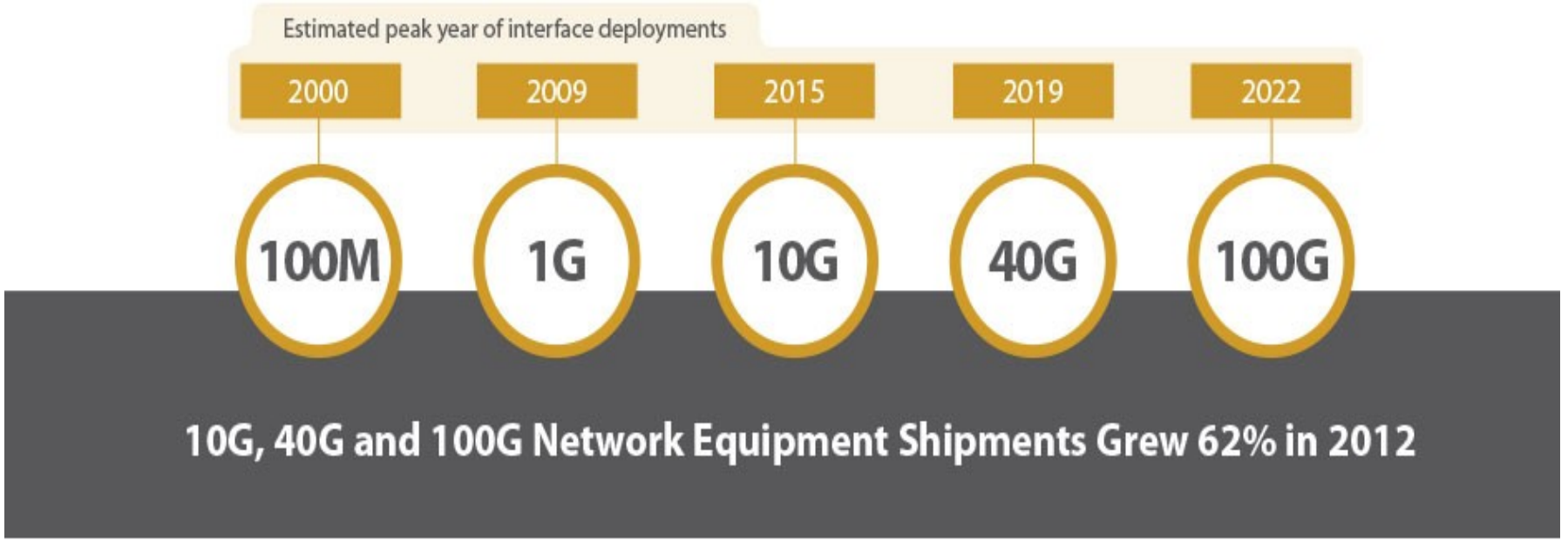
Traffic Engineering	*Accounting*
Fault diagnosis	Forensic analysis
SLA monitoring	Anomaly detection
worms, portscans, botnets	

Hardware Based - Flow Monitoring



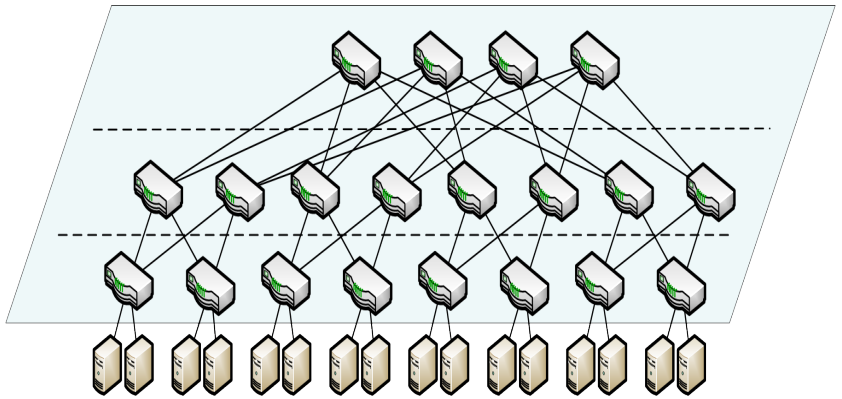
- **Task Oriented**
  - Task 1 : Anomaly Detection
  - Task 2 : Traffic Engineering
- **Problem**
  - **Not Evolvable (DC Requirement)**
    - Higher speed links (40/100 Gbps)
    - SLA monitoring in data centers

Data Center Network Is Evolving

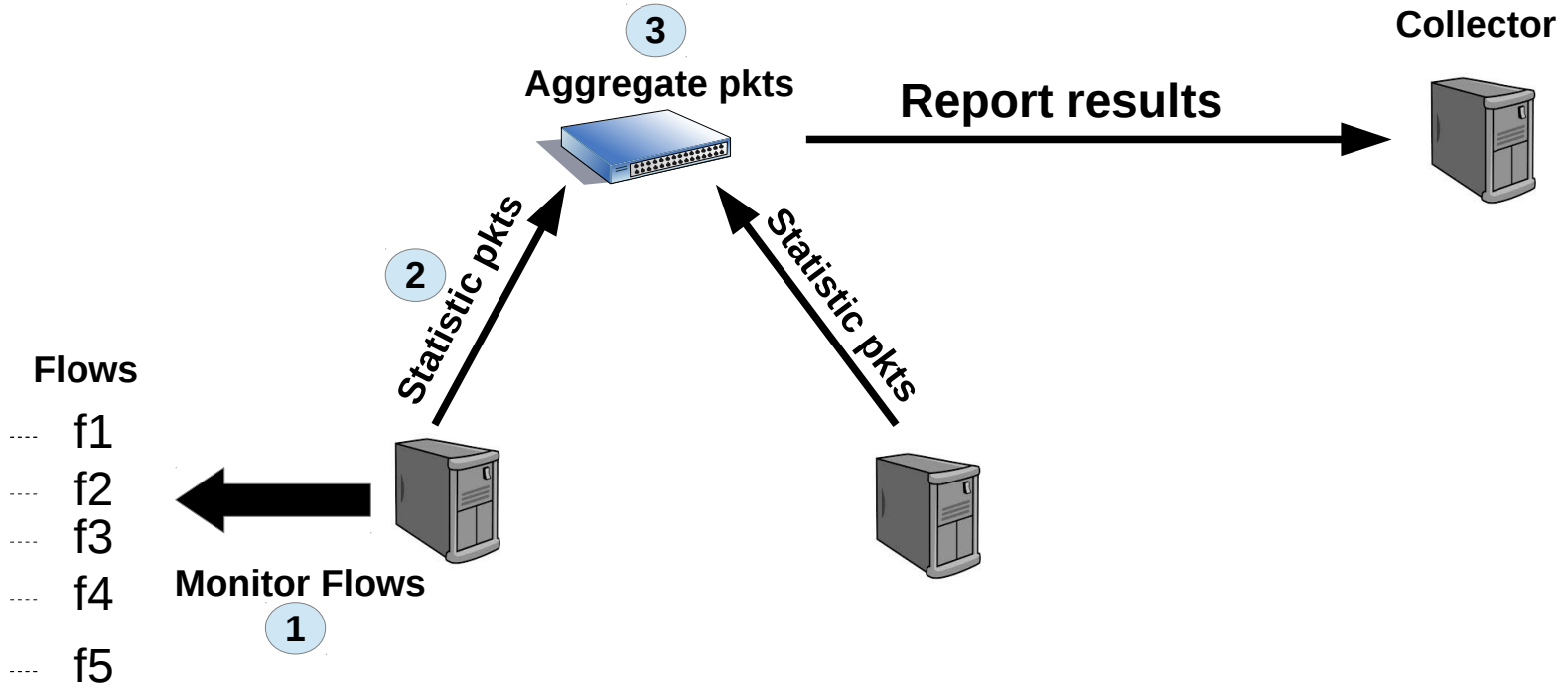


(Net Optics 2013)

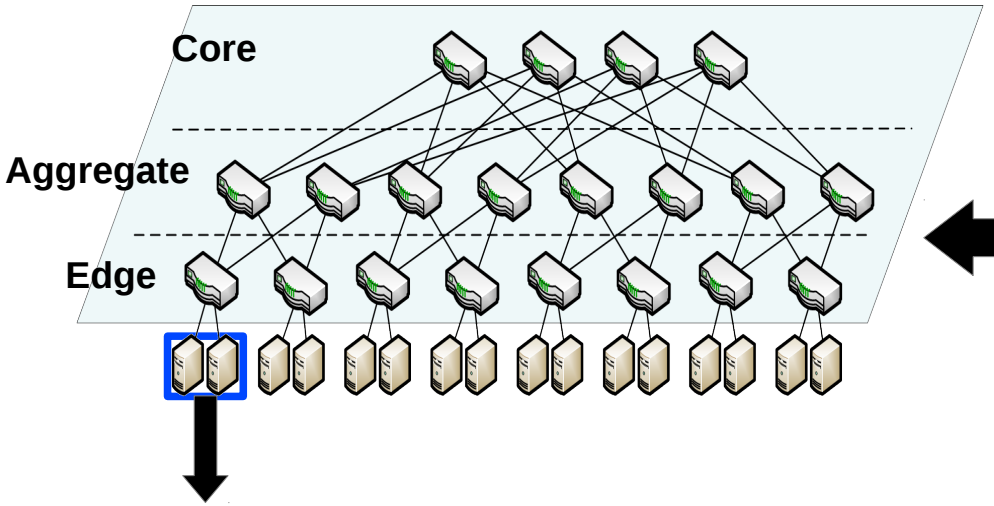
Distributed Traffic Measurement



- **Our approach :**
  - Distribute flow monitoring overhead between switches and servers



Distributed Traffic Measurement



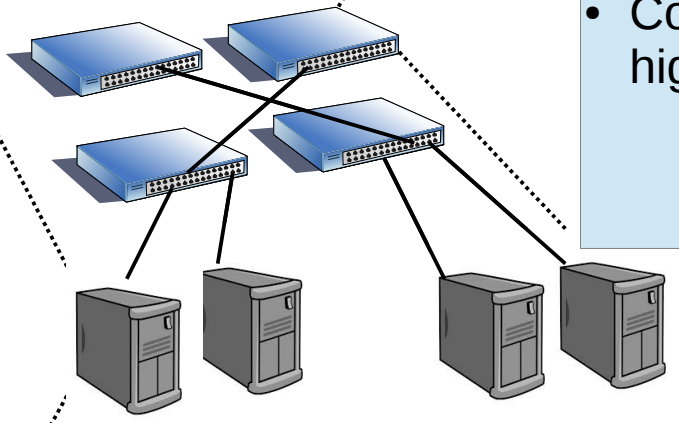
Administrators have complete control of switches and servers

- High computational resources (multiple cores, large memory)
- Hosts observe relevant traffic of running services
- Monitors less traffic than switch

Proposed Framework

Aggregation module

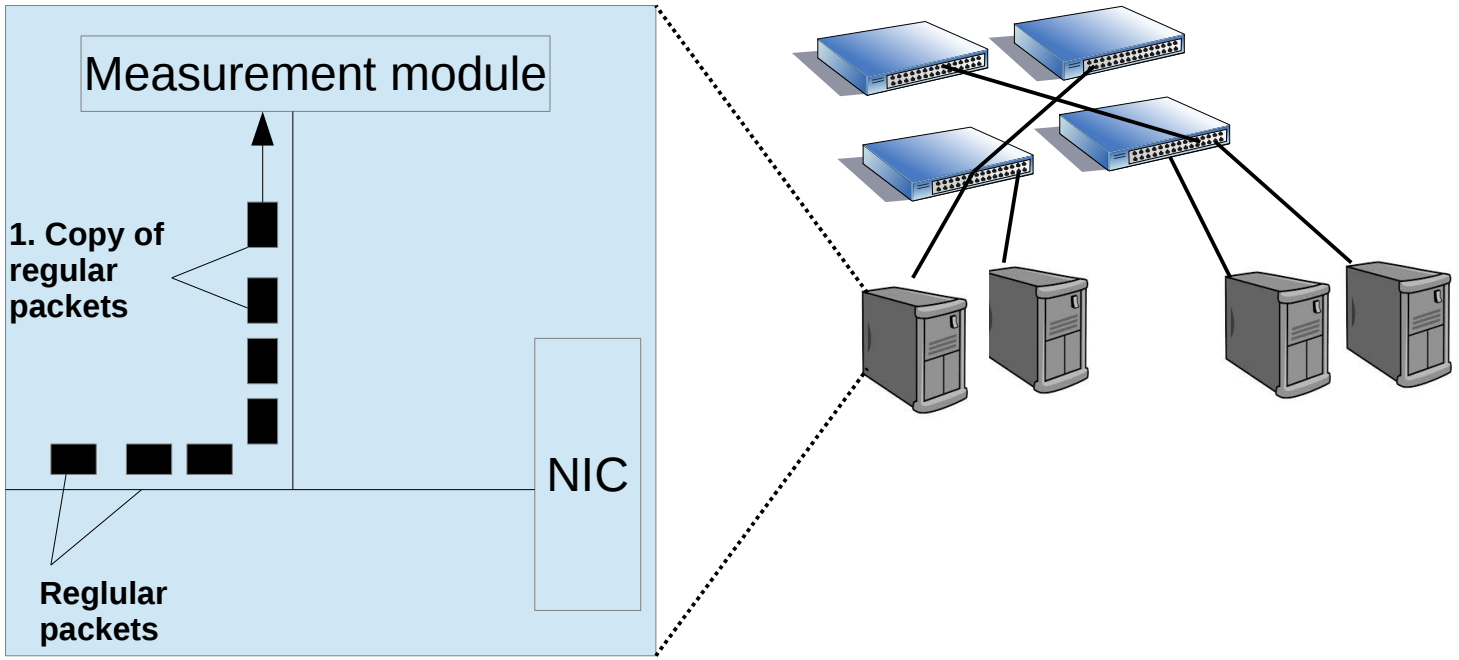
- Consumers
- Aggregates statistic packets (s-pkts)
- Counters can be stored in high density DRAM



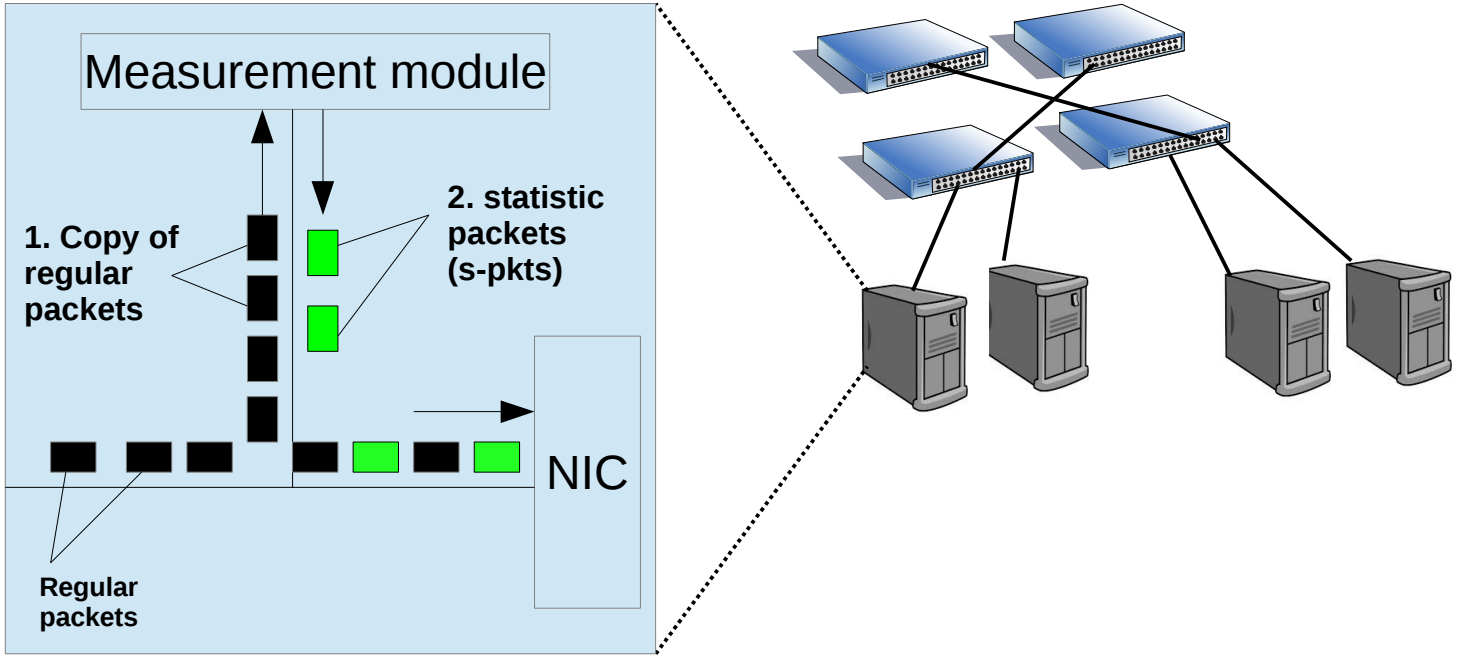
Measurement module

- Producers
- Monitor traffic and generates statistic packets (s-pkts) (e.g., per flow record)
- Feeds s-pkts to ToR switch

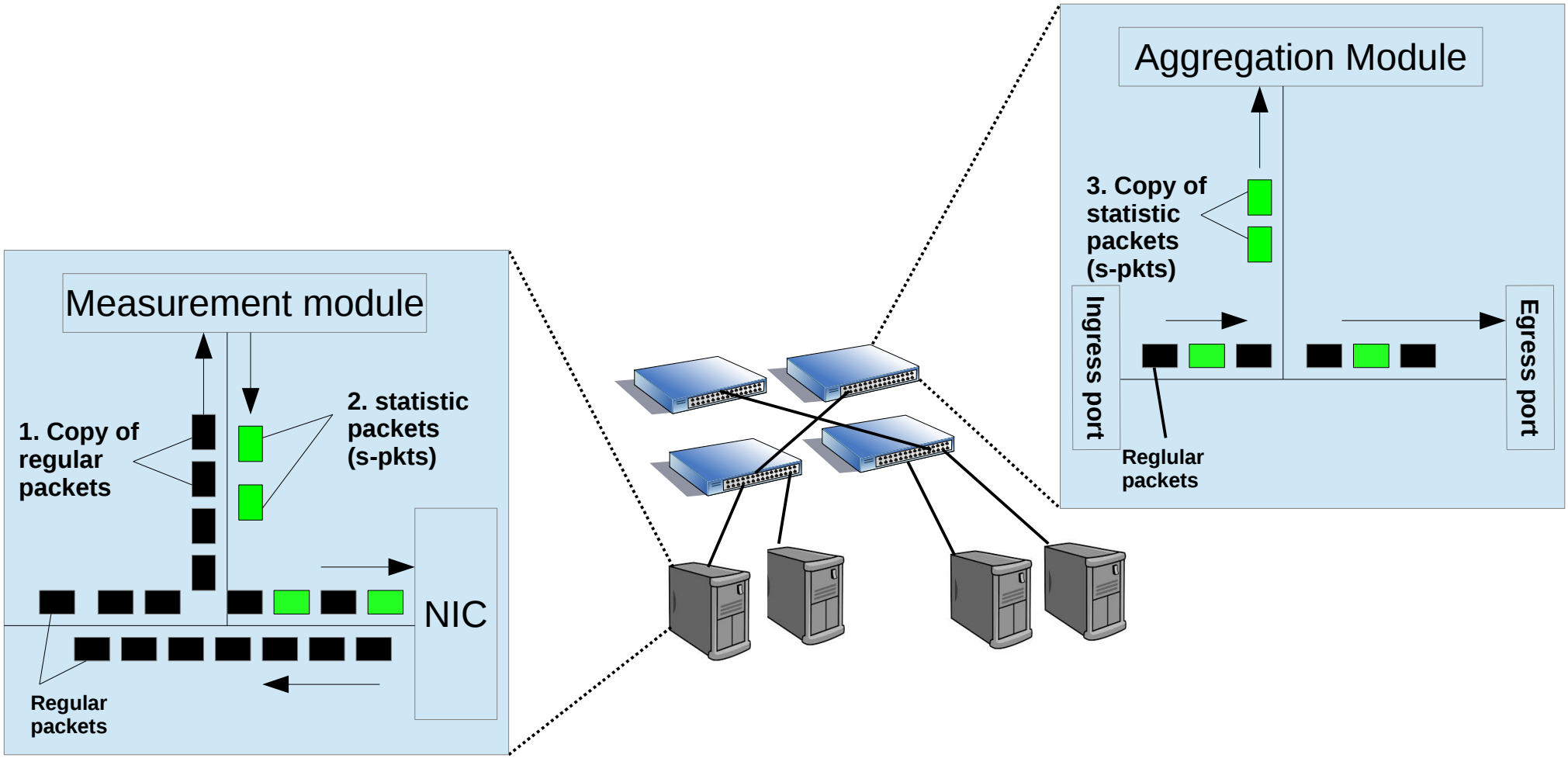
Proposed Framework – Packet Processing



Proposed Framework – Packet Processing



Proposed Framework – Packet Processing



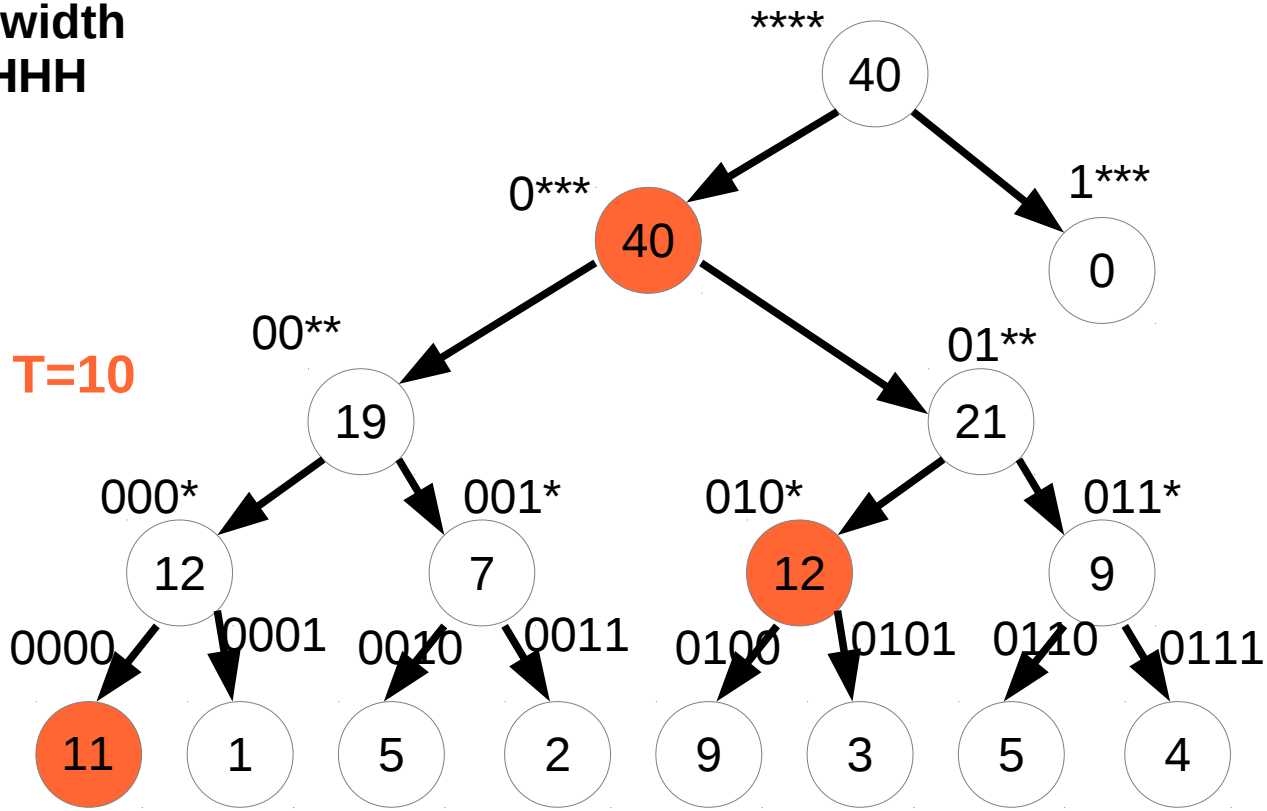
On going work : Statistic Packet Forwarding

- How to forward statistic packet ?
  - Packet path encoding and IP source route option
  - Use switch forwarding table

Usecase – Hierarchical Heavy Hitter (HHH)

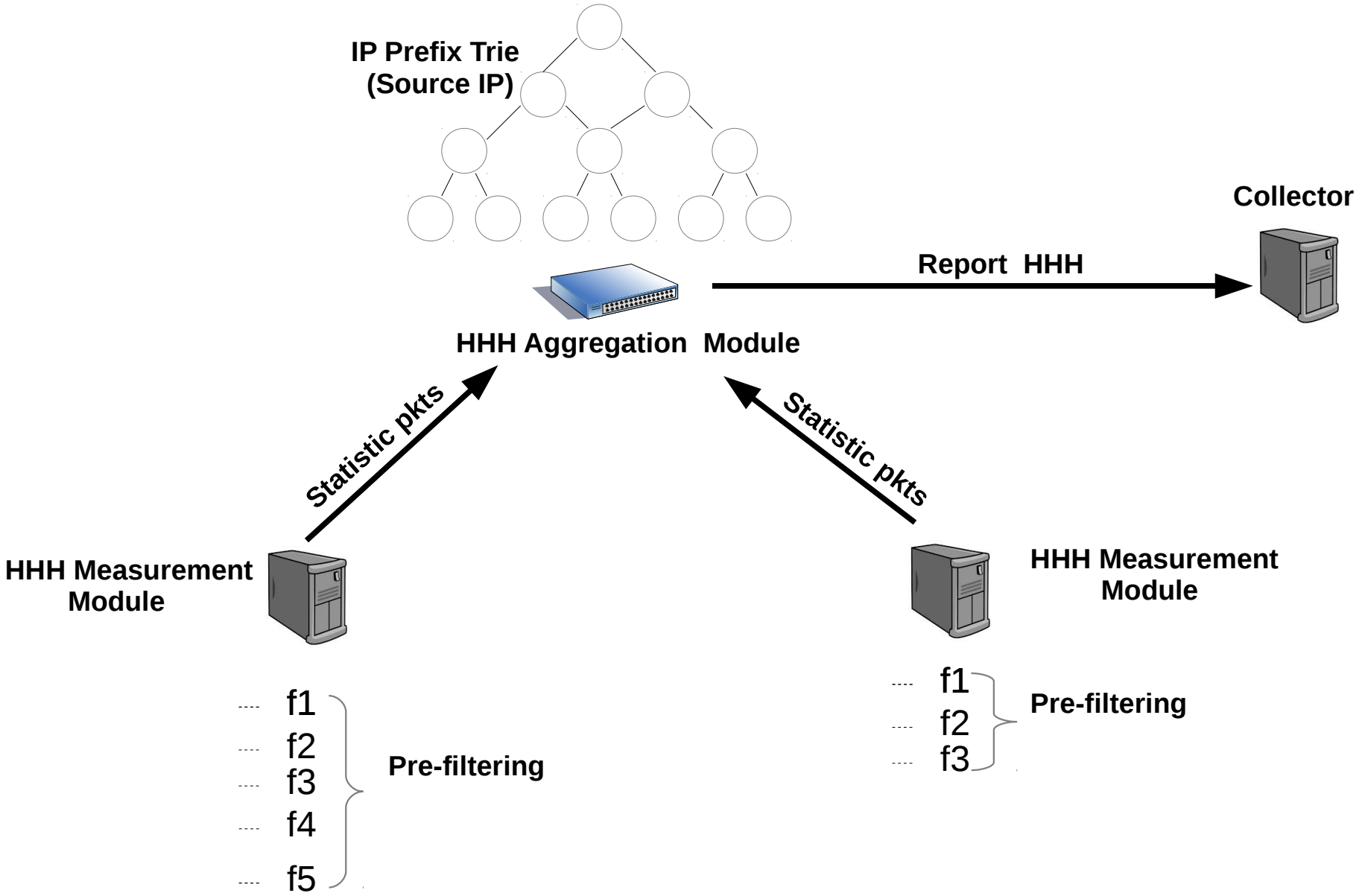
HHH: Longest IP prefix occupies more than fraction T of link bandwidth after excluding any descendant HHH

Threshold : T=10



Traffic volume for each IP Prefix

HHH Detection

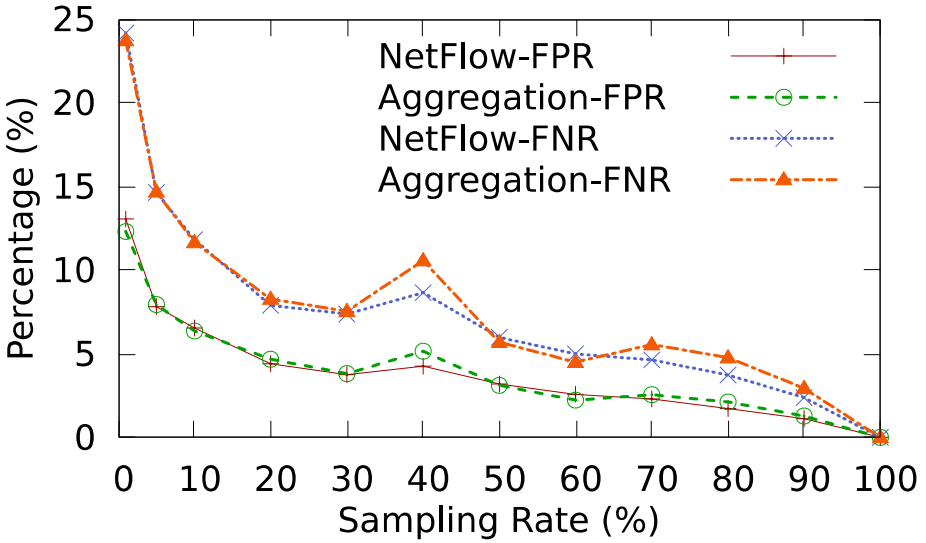


Evaluation

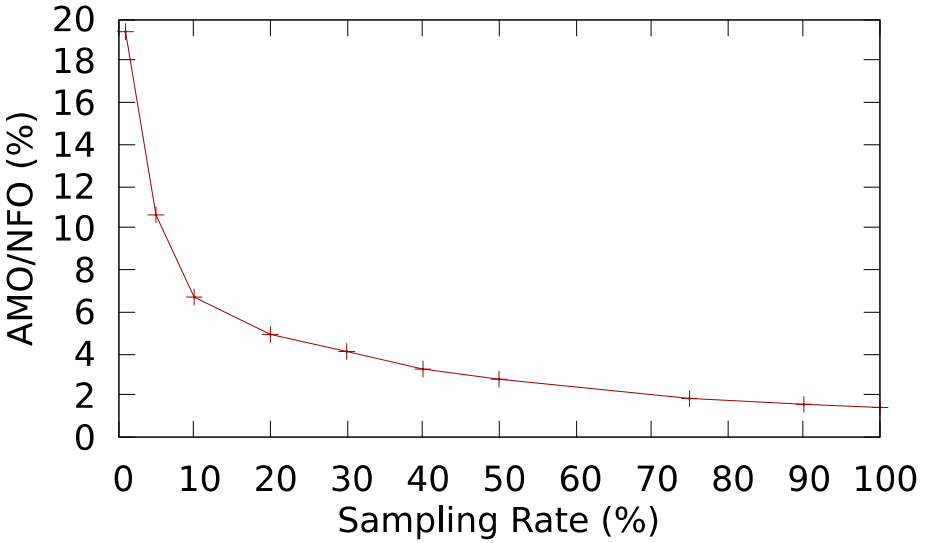
- **Simulation setup**
  - Measurement module : **Customized YAF**
  - Aggregation module : **IP Prefix Trie**
  - Packet trace – T. Benson : University data center
- **Aggregation module performance**
  - HHH Accuracy
  - Computation overhead on Servers and switches
  - Compared with NetFlow

Preliminary Results

FPR : False Positive Rate
FNR : False Negative Rate



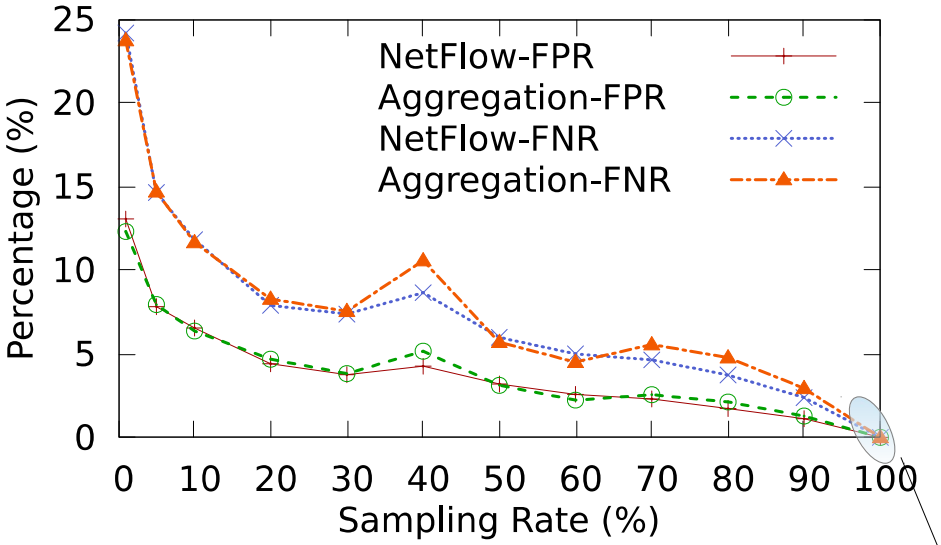
HHH Accuracy Varying Sampling Rates



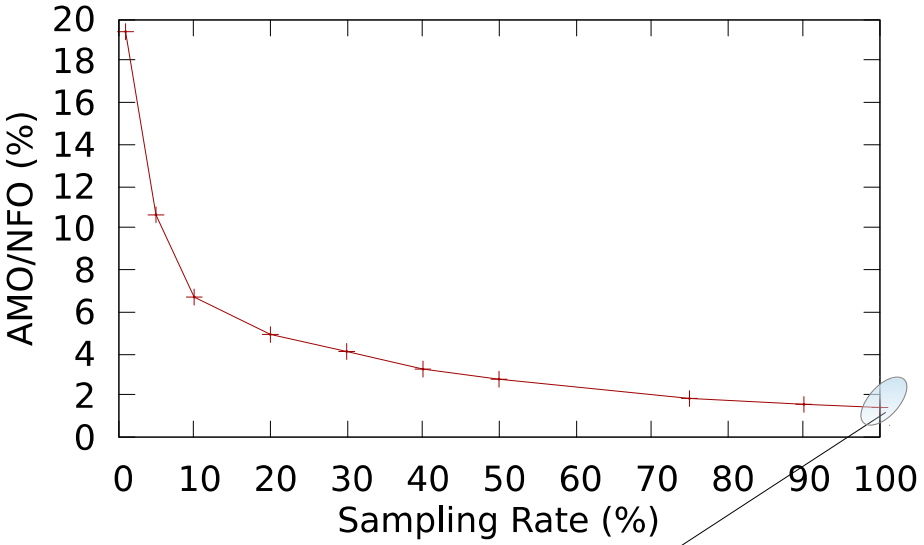
Aggregation module overhead (*AMO*) over NetFlow overhead (*NFO*)

Preliminary Results

FPR : False Positive Rate
FNR : False Negative Rate



HHH Accuracy Varying Sampling Rates



Aggregation module overhead (*AMO*) over NetFlow overhead (*NFO*)

Correctness : 100% Sampling rate – 100% Accuracy
Overhead: AMO is just < 2% of NFO

Conclusions and Future work

- **Conclusions**
  - Our framework offloads overhead on switch
  - Evolves along with data center traffic volume
  - Provides more flexibility to data centre operators
- **Future Work**
  - Prototyping proposed framework
  - Exploring performance across different measurement tasks
  - Endhost based network trouble shooting (e.g., packet loss, delay)
  - Impact of packet loss on accuracy
  - Distributing measurement task overhead across network

Thank You

Questions

Praveen Tammana
praveen.tammana@ed.ac.uk
University of Edinburgh

Challenges

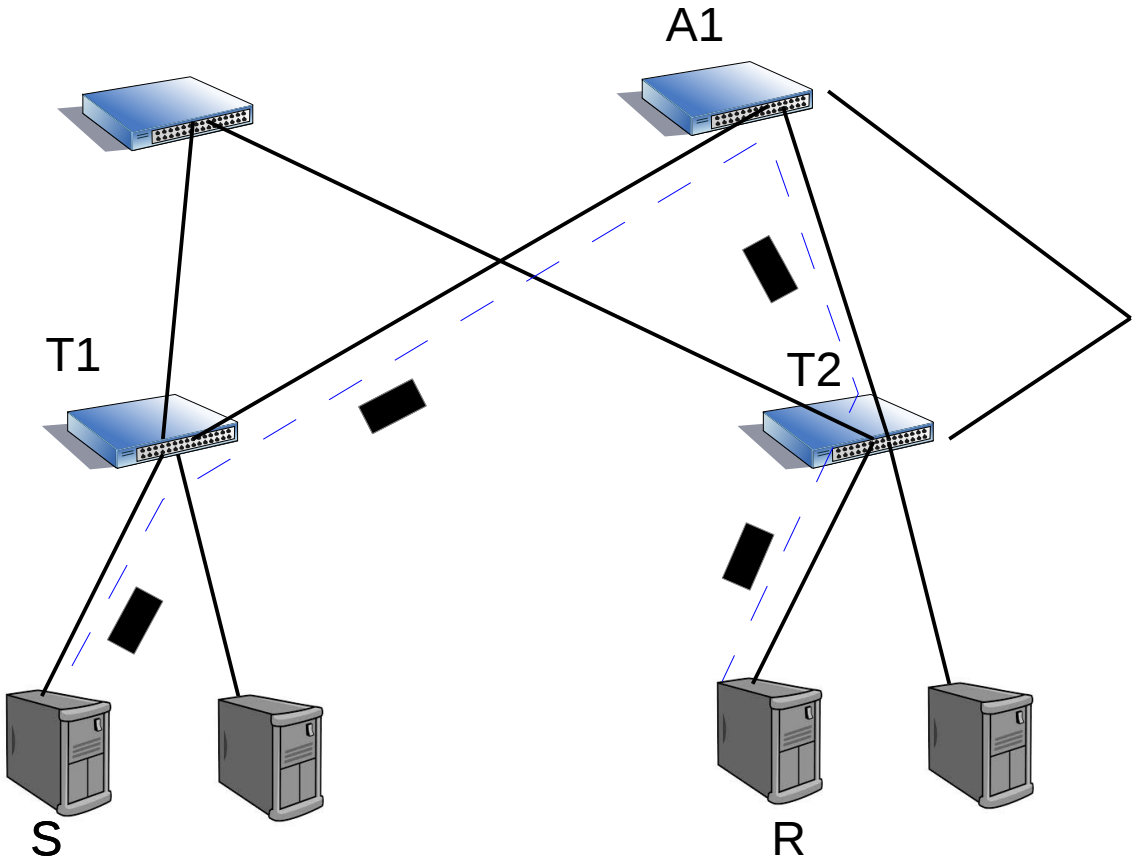
- Handling multiple paths between End Hosts
- Consistency with forwarding rules update

Measurement Tasks

- Hierarchical Heavy Hitter (HHH)
- Heavy Hitter
- Superspreader
- Flow Size Distribution
- DDoS

Proposed Framework : s-pkt forwarding

Flow Path : S → T1 → A1 → T2 → R

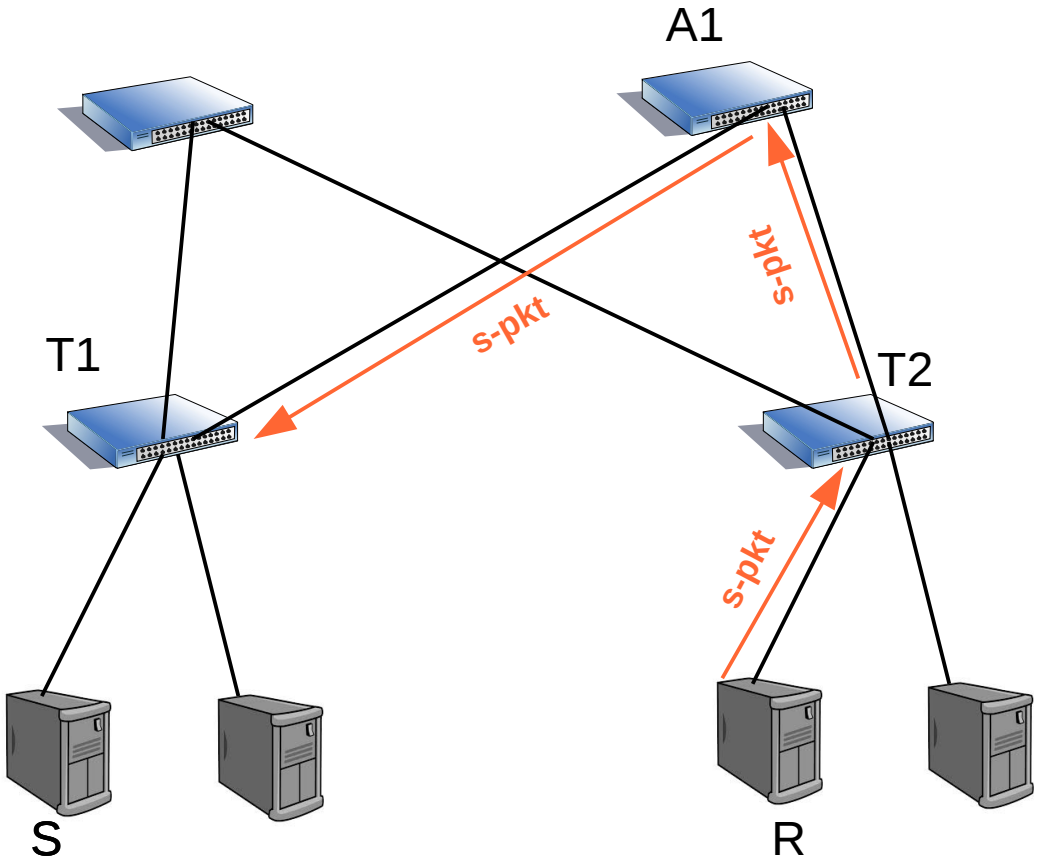


Encodes path information into packet

Proposed Framework : s-pkt Forwarding

Flow Path : S → T1 → A1 → T2 → R

s-pkt : R → T2 → A1 → T1



1. Generate s-pkt

2. Enables IP source routing option