A Real-time QoS Measurer

Lisheng Huang
LSH@uestc.edu.cn

UESTC

Topics of Discussion

- Introduction
- On network QoS measuring
- Current methods & approaches
- Our method
- Consideration of accuracy
- Upcoming work

Introduction

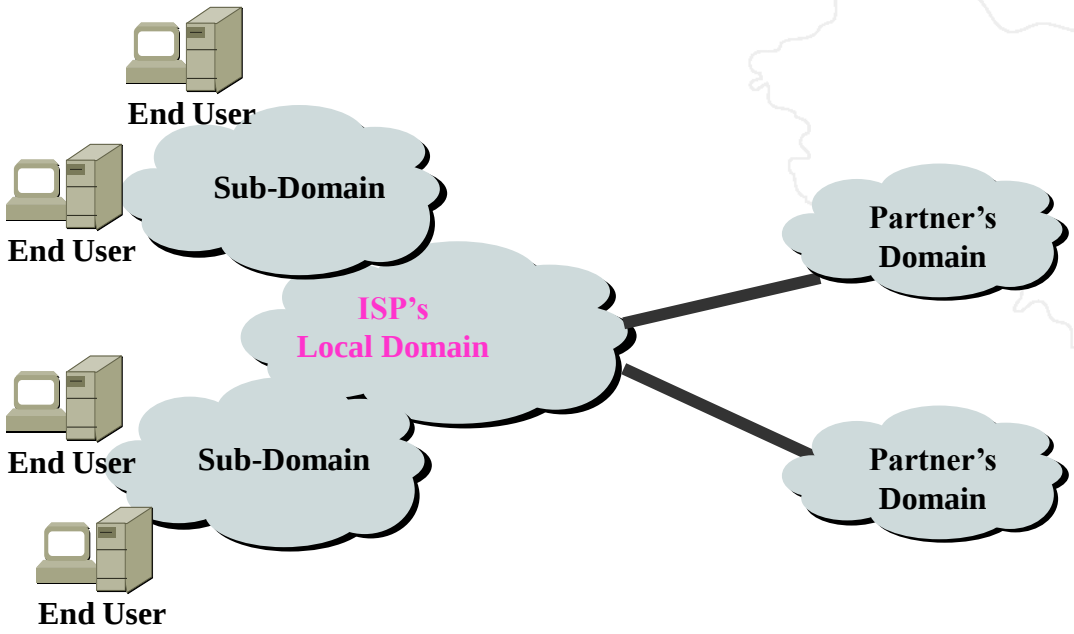
- A real-time, passive mode measurement tool based on high speed traffic monitoring & tracing mechanism.
- Make it possible for ISPs to obtain QoS metrics for most network paths within a short term at one measuring point.

On network QoS measuring

- Metrics of QoS
  - Transfer Delay, Loss Ratio, Path Bandwidth, Available Bandwidth, etc.
- Levels of QoS
  - Path-level: user-oriented
  - Domain-level: ISP-oriented
- Requirements from ISPs
  - SLA ensureance
  - Fault diagnose
  - Network / Application healthy
  - Measure->decision->traffic engineer

On network QoS measuring

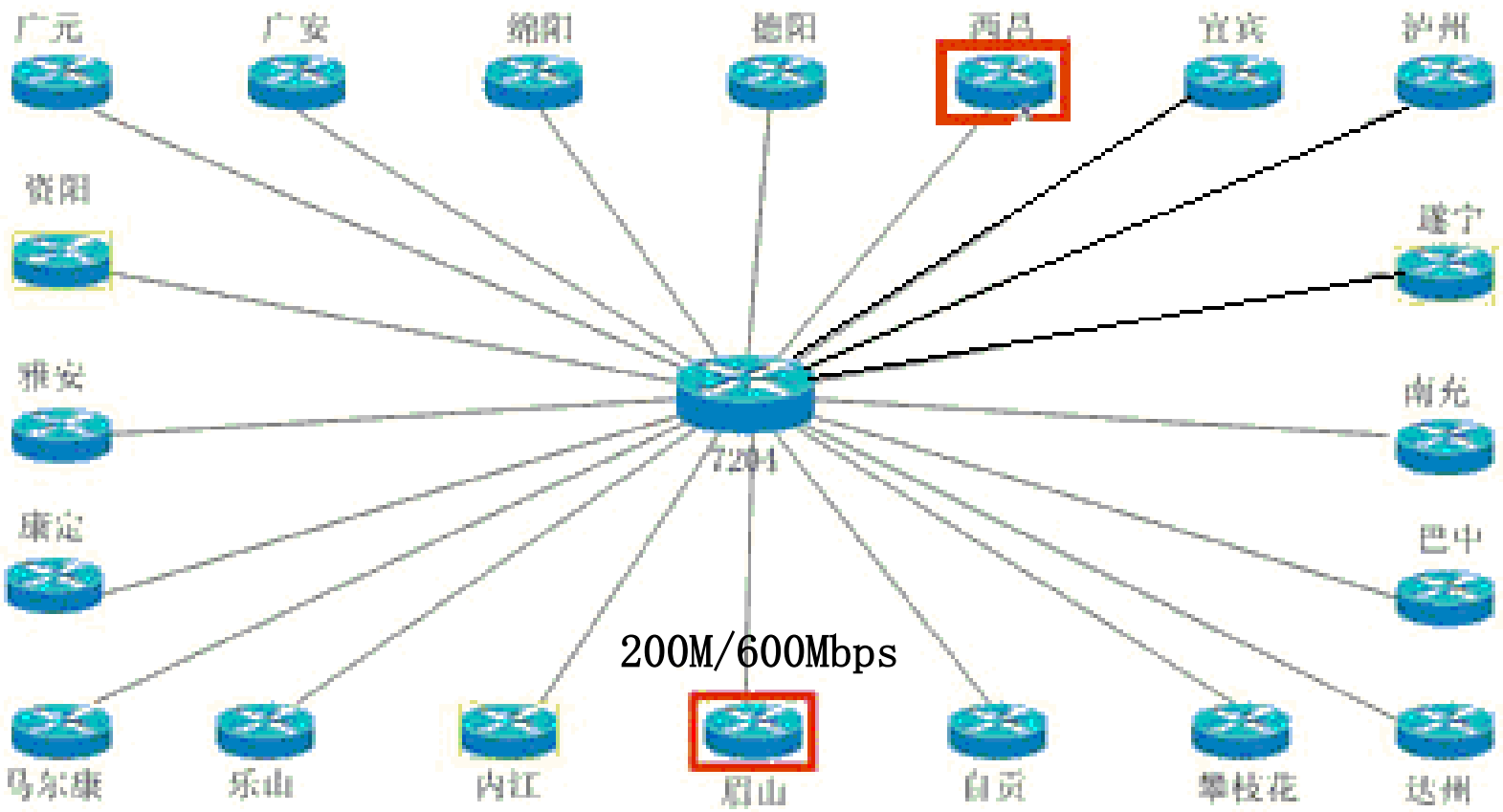
What do ISPs need?



- SLA verify
- Fault diagnose
- Network / Application healthy discover

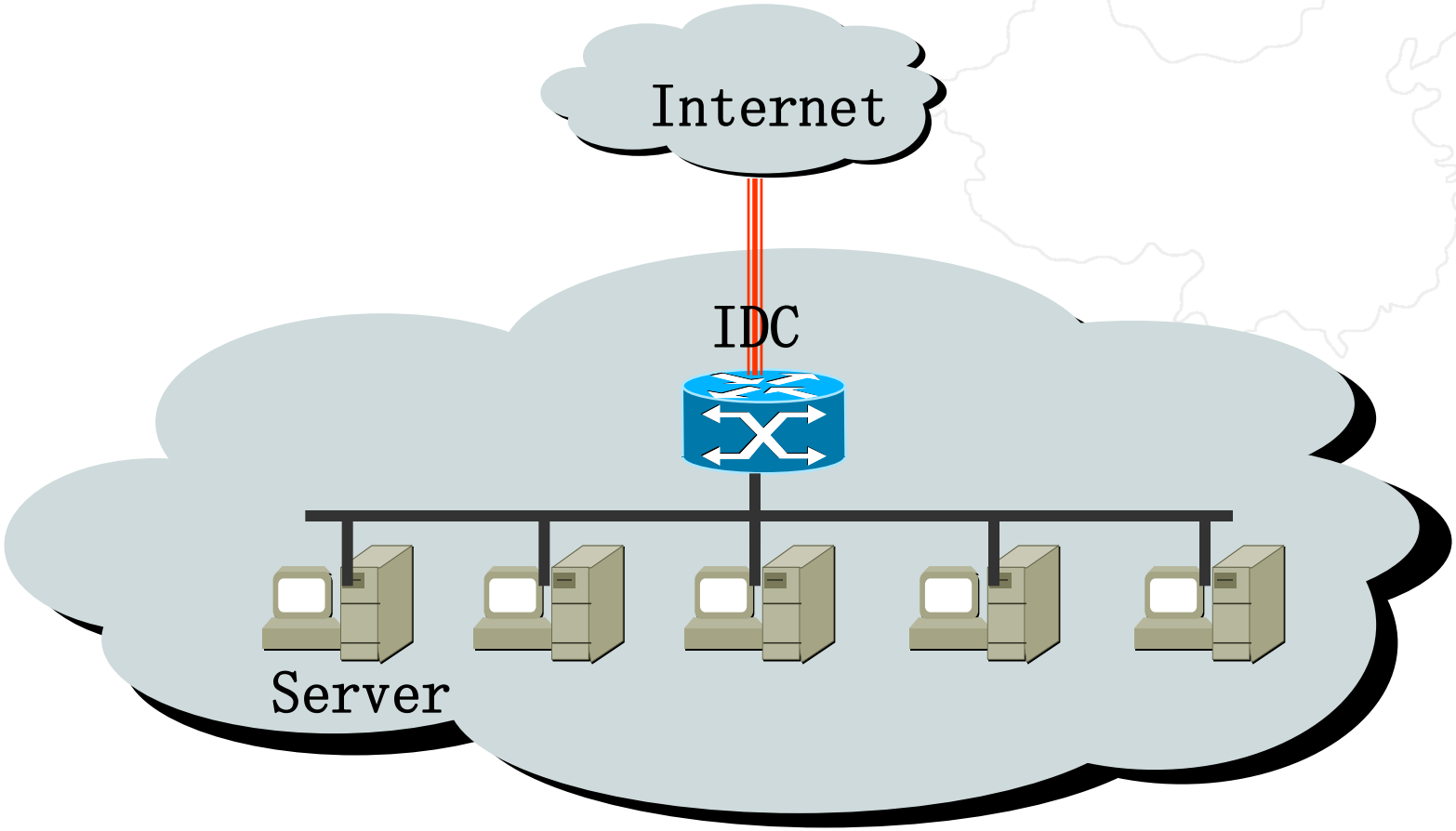
On network QoS measuring

What do ISPs need?



On network QoS measuring

What do ISPs need?



Current methods & approaches

- Active mode method
  - Probe packets
- Passive method
  - MRTG / NETFLOW
  - Traffic Monitoring

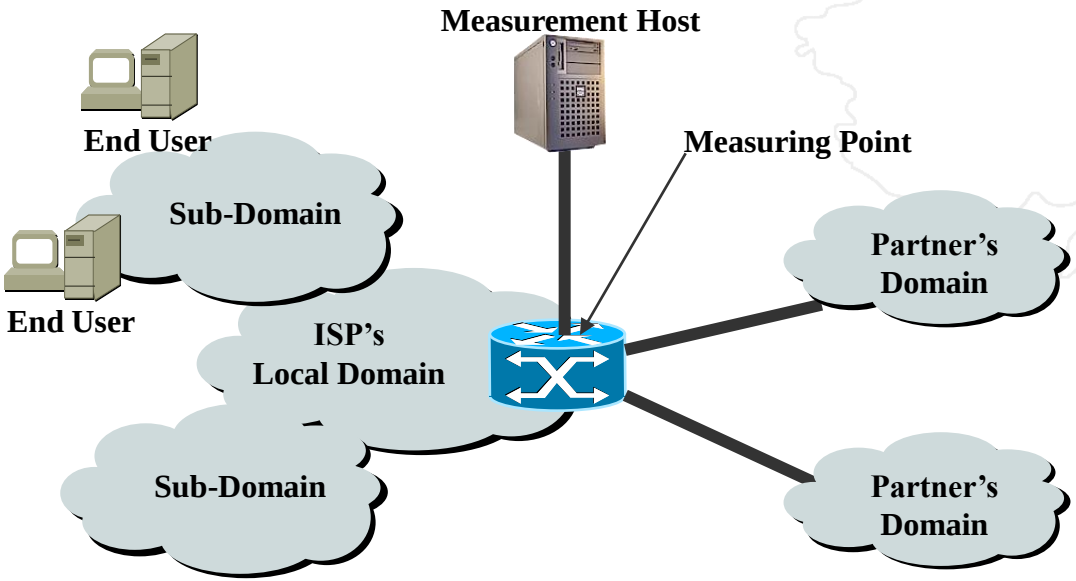
Current methods & approaches

Which method is most suitable for ISPs?

- Router-based methods
  - Do not provide any metric on end-to-end QoS
- Active mode methods
  - There are too many links / paths need to measure one-by-one
  - Extra traffic can hardly be accepted
- Traffic-monitor methods
  - Fast, automatic, easy, global, and accurate

Our method - measurement architecture

- Monitor network traffic at key point (measuring point)



Our method – TCP tracing

- Measure 3 main metrics (LR, TD, PB) by tracing all TCP conversations
  - widely used by network applications
  - reflects user's real perception
  - strictly "context-sensitive" , suitable for passive mode monitoring
  - contains multifarious protocol activities: as probe packets

Our method - LR

- Measuring Loss Ratio

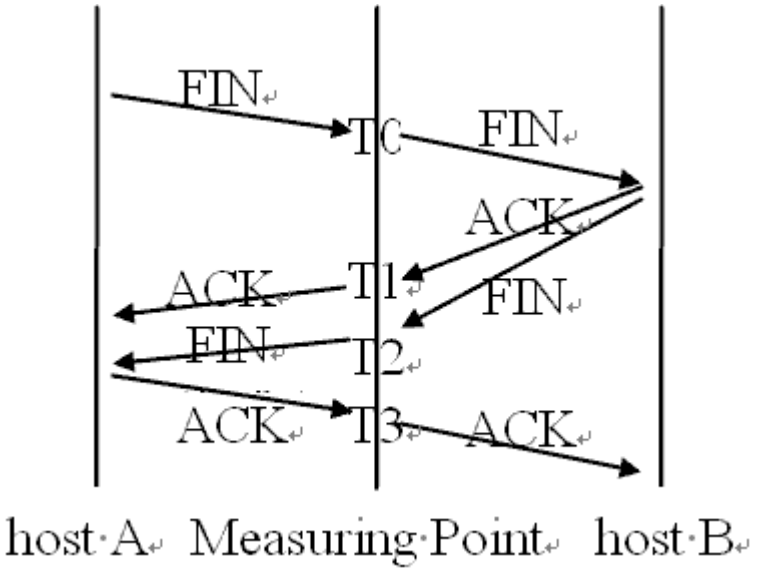
SEQ number of every TCP packet from a host can be predicted according to the previous TCP packet's SEQ number and TCP data length.

$$SEQ_{predict} = SEQ_{this} + DATALEN_{this}$$

If the actual SEQ number of a TCP packet (called SEQacture) is not the same as SEQpredict, there must be at least one packet was lost.

Our method - TD

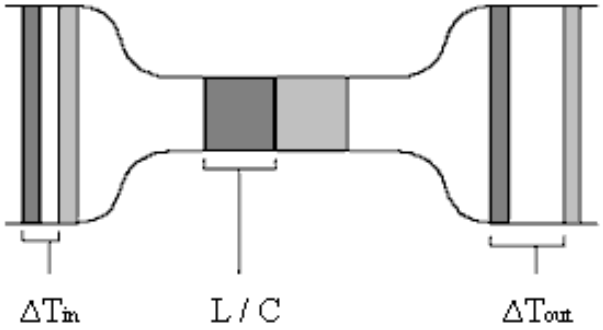
- Measuring Transfer Delay

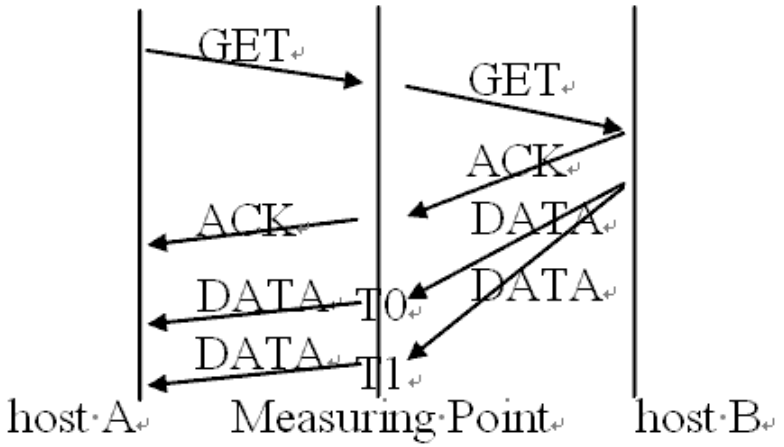


$$HPTD_{meas} = T1 - T0$$

Our method - PB

- Measuring Path Bandwidth(1)

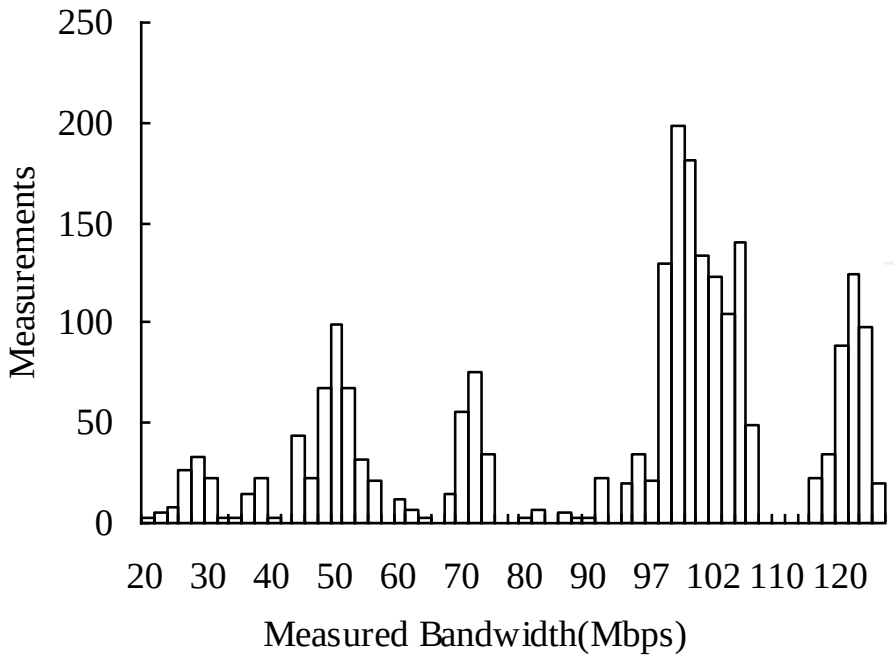




$$HPB = L / \Delta T$$

Our method - PB

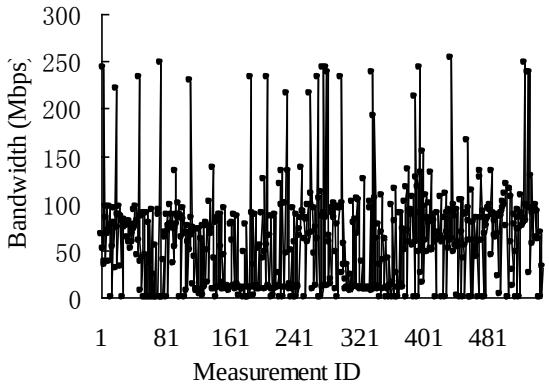
- multi-peak of measured path bandwidth



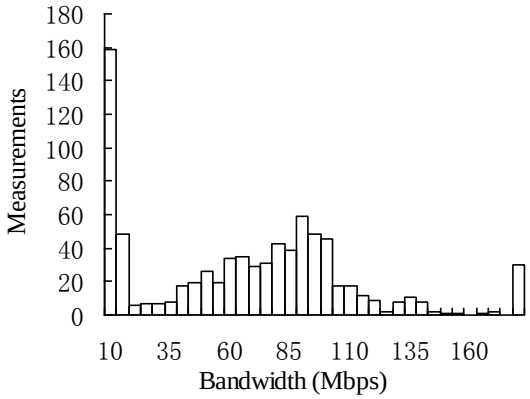
Filter noise with kernel density estimation

Our method - PB

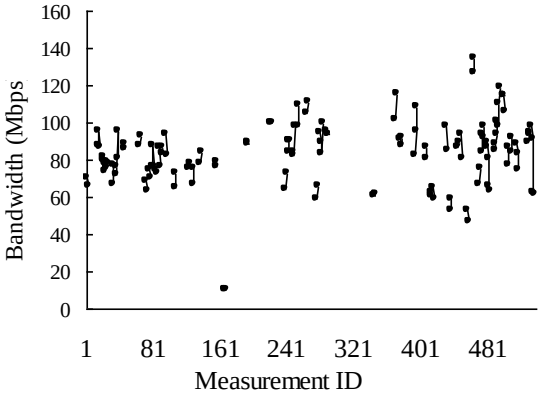
- "*sequence mode filtering*" (SMF)



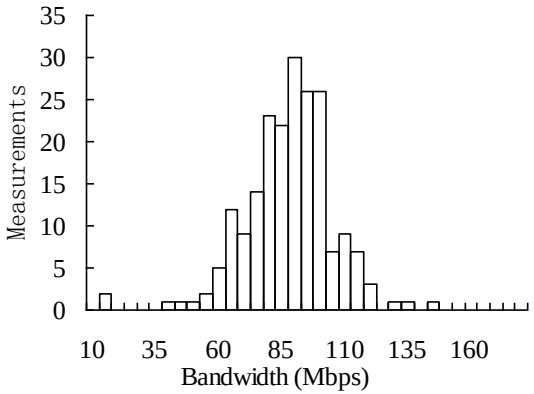
A. Serials of samples (Not filtered)



B. Not filtered Bandwidth



C. Serials of samples (Filtered with SMF)



D. Bandwidth Filtered with SMF

SMF filter works when measuring the bandwidth of a 100Mbps link

Our method - AVB

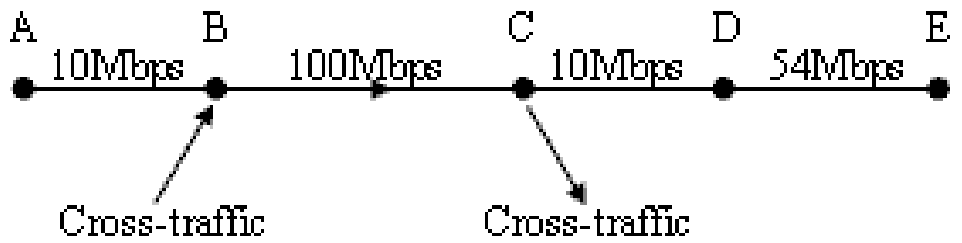
- Domain Path Idle Rate (DPIR, 0≤DPIR≤1)
  - DPIR is the idle rate of the path bandwidth from a network domain to the measuring point.
  - “long-time queuing” denotes busy status

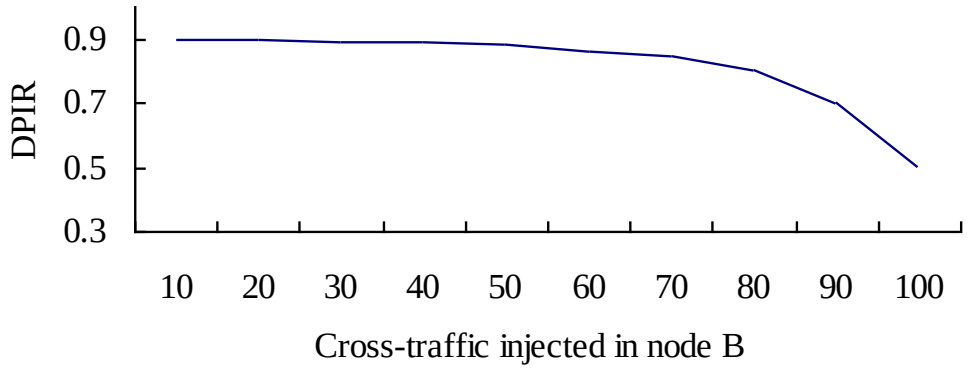
$$DPIR = \frac{N}{M}$$

While M is the amount of all PB measurement samples, and N is the quantity of valid samples (with SMF filter)

Our method - AVB

- Measuring DPIR

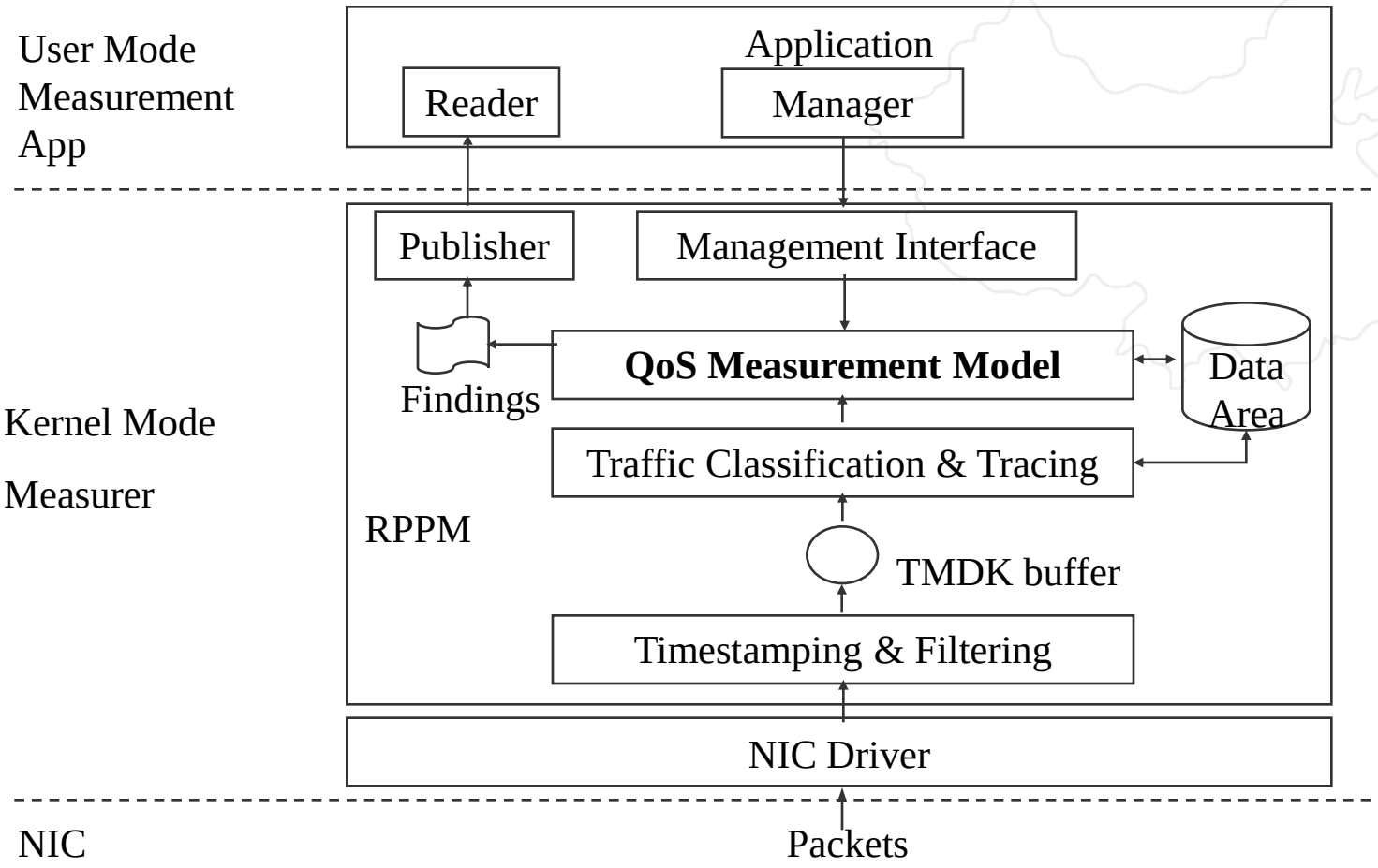




DPIR changes with the cross-traffic payload

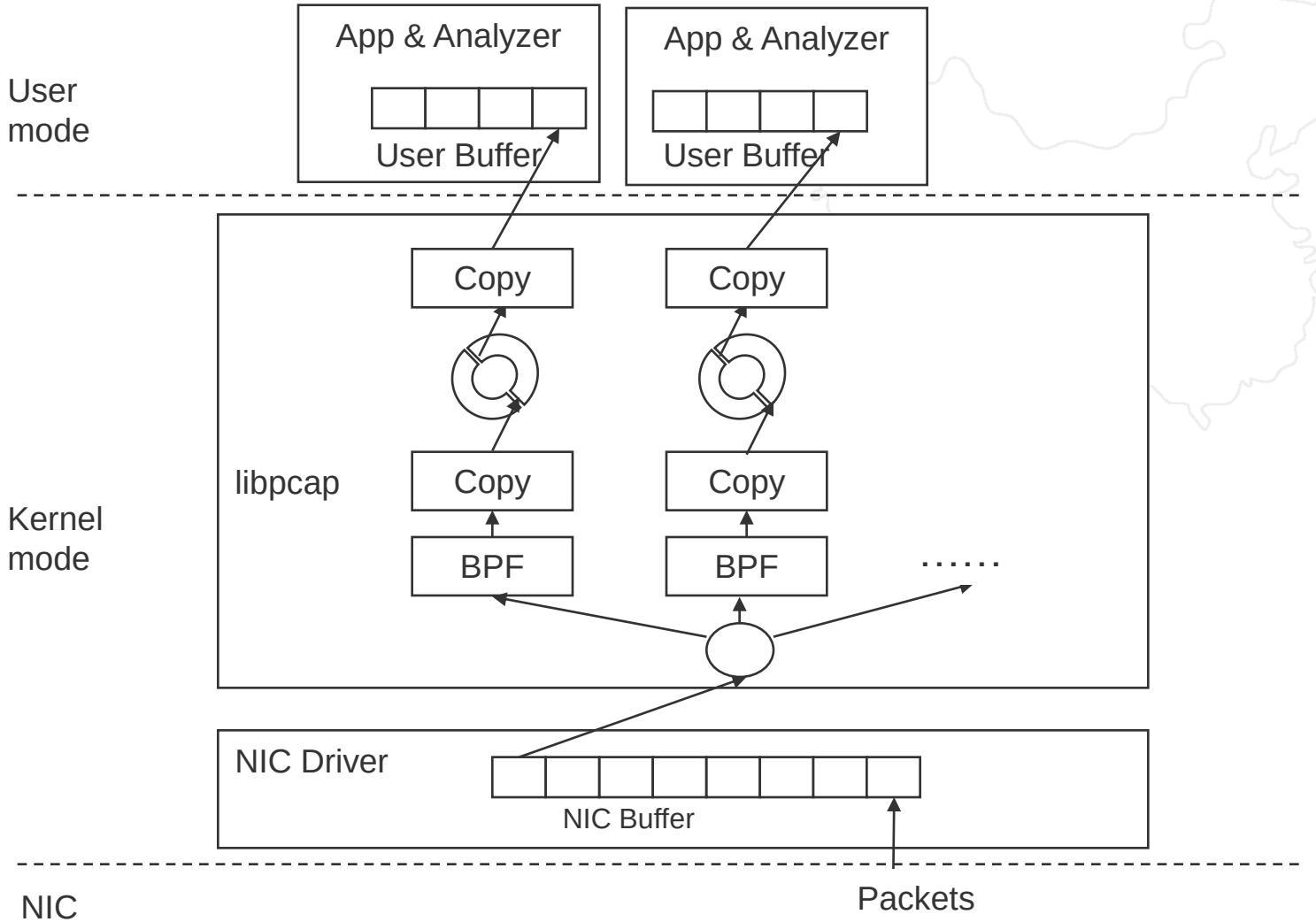
Our method – Software architecture

- RPPM works in kernel mode



Our method – Software architecture

- Comparing with libpcap packet capture

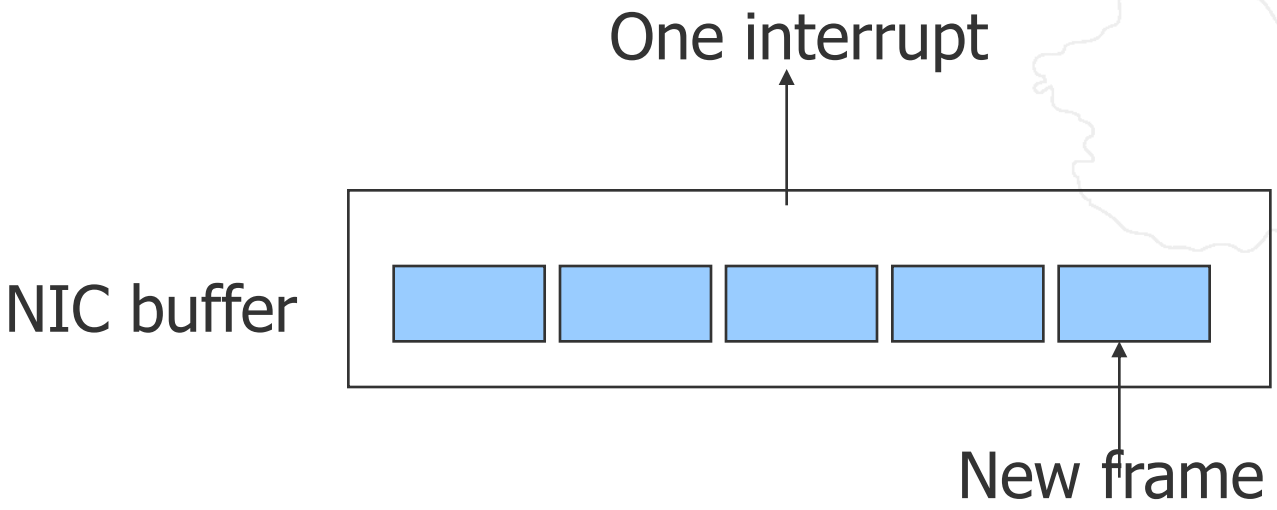


Consideration of accuracy

- accuracy is so important that it determines the availability of a measurement method

Consideration of accuracy

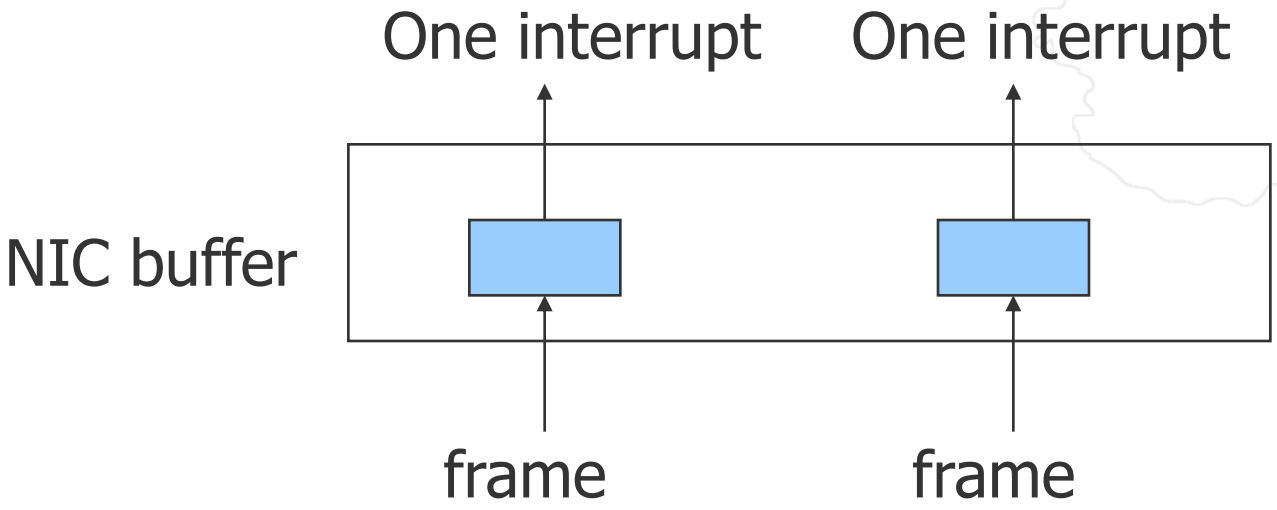
- NIC interrupt coalescence



NIC may reach an average value of 49.61 frames served per interrupt, these frames have the same timestamp !

Consideration of accuracy

- Disabling NIC interrupt coalescence



Our test results show: under the traffic load of 120-200K fps, a Gigabit NIC generates one interrupt for every packet (once per 4-8μs), and indicates only one packet in each interruption.

Consideration of accuracy

- Verification

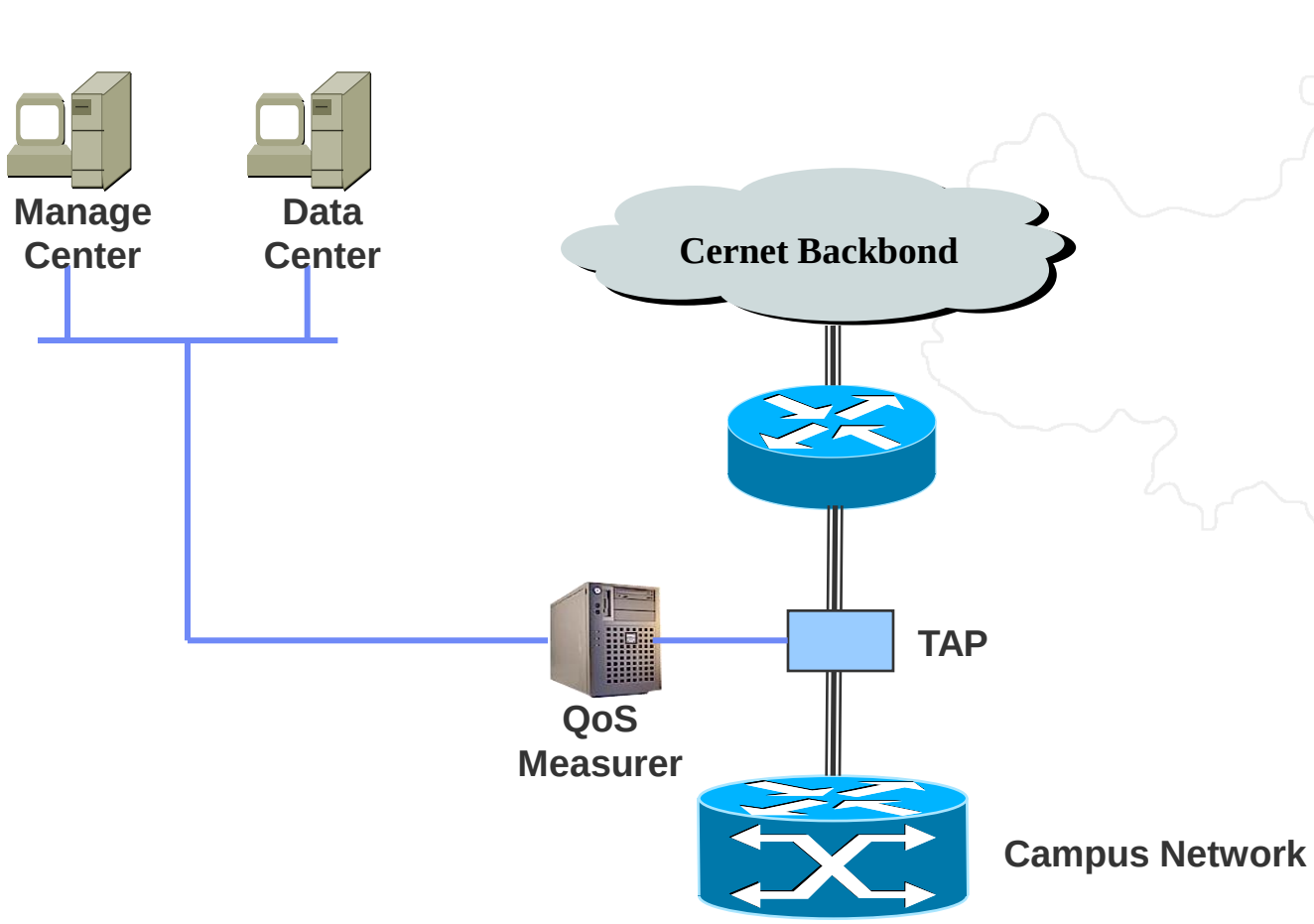
Path Bandwidth Measurement result

Host	PB to Measuring Point	Measured PB
202.115.25.X	10Mbps	9.9Mbps
202.112.14.X	100Mbps	102.1Mbps
222.197.188.X	1000Mbps	1280Mbps

Consideration of accuracy

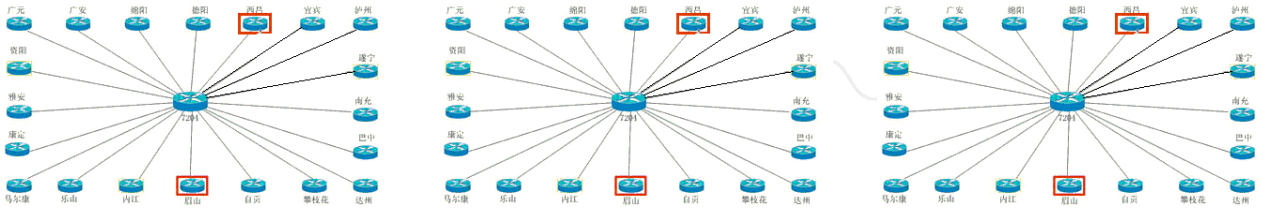
- More tests

RPPM is deployed into CERNET southwest backbone network (traffic flow between 850Mbps and 900Mbps, average alive TCP conversations between 50,000 and 200,000)


Manage Center
Data Center
Cernet Backbond
TAP
QoS Measurer
Campus Network

Upcoming work

Net View



Data Aggregation

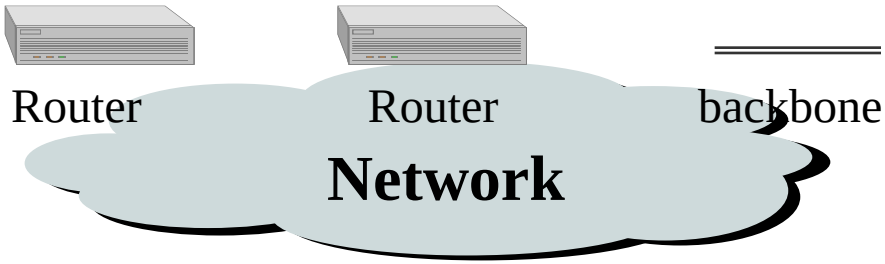
Flow Model / Performance / Security / User activity

Traffic Sensor

Sensor

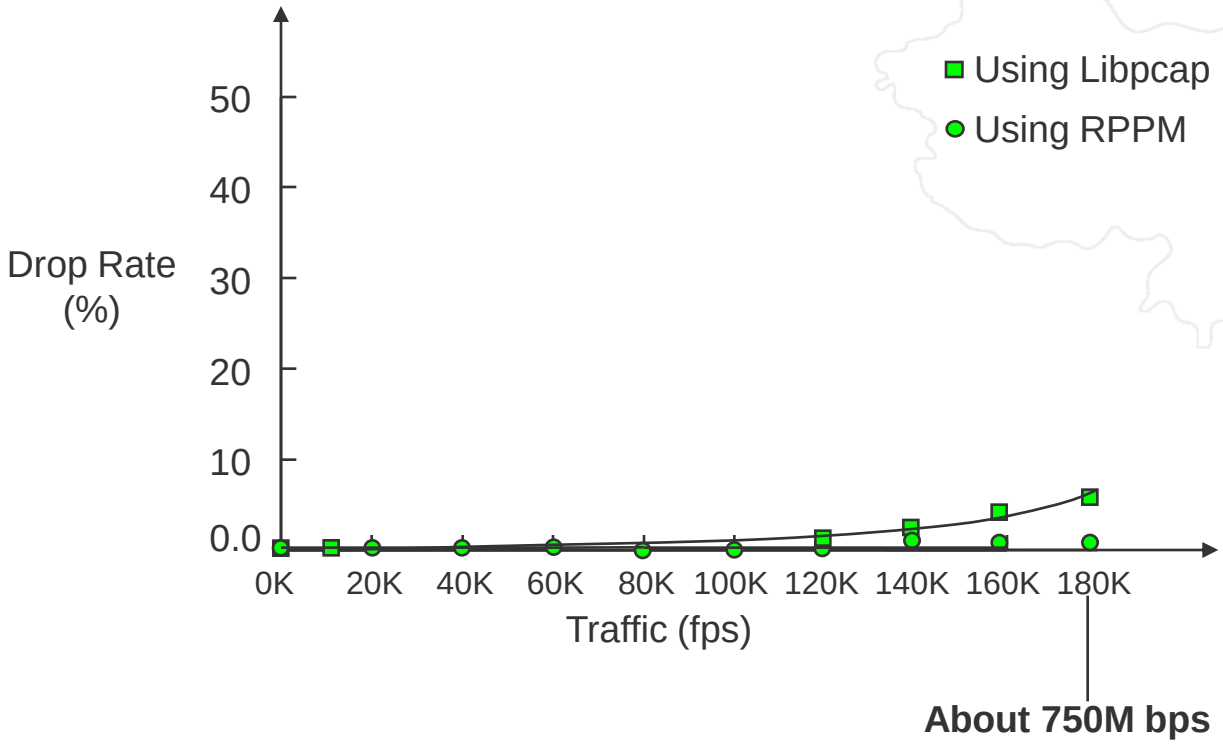
Sensor

Entity



Upcoming work

- **Traffic Sensor——Performance**



- The goal is > 7Gbps real-time measurement

Our other ralated work

- Method: Passive Network Measurement
- Performance: 100M -> 1G -> 2.5G -> 10G
- Goal: Flow Classification, Audit, IDS, Performance, Application, User…
- Standard: RTFM/IPPM

End of Talk. Thanks