Design and Development of the Platform for Network Traffic Statistics and Analysis

Hao Hu, Luo Qi, Fazhi Qi

IHEP

22 Mar. 2018

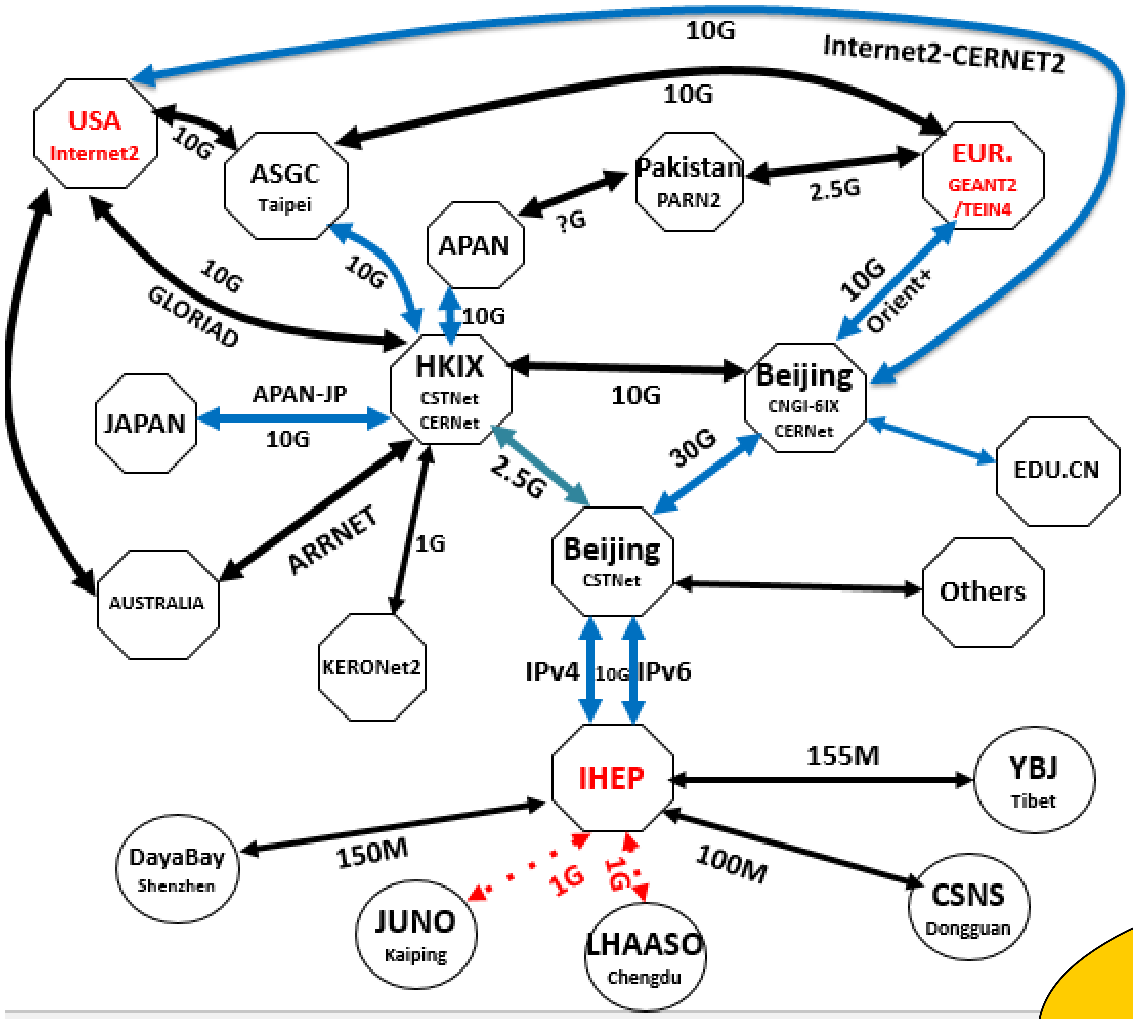
Outline

1. Motivation
2. Platform design
3. Function modules
4. Future Plan
5. Summary

Outline

1. Motivation
2. Platform design
3. Function modules
4. Future Plan
5. Summary

IHEP WAN Topology

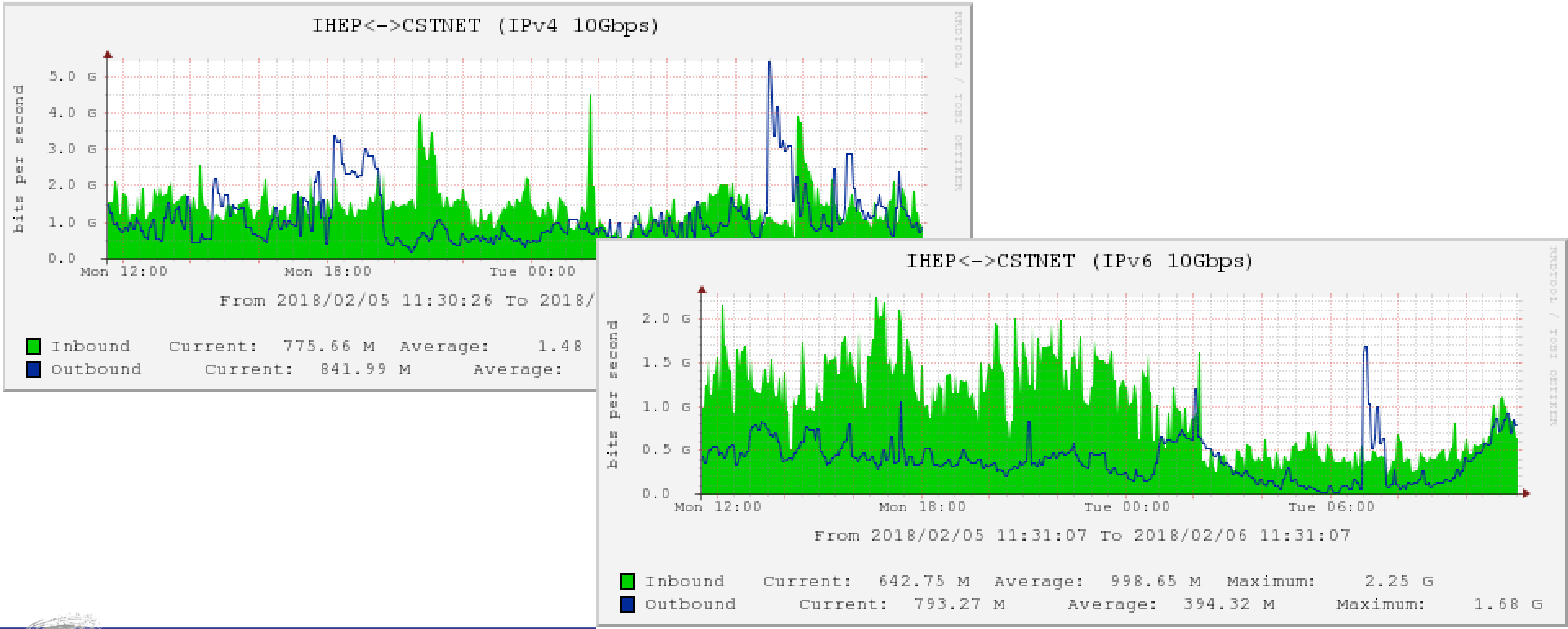


- **IHEP- USA**
  - IHEP-CSTNet-CERNet-USA
  - 10Gbps
- **IHEP- EUR**
  - IHEP-CSTNet-CERNet-London-EUR
  - 10Gbps
- **IHEP- Asia**
  - IHEP-CSTNet-HKIX-Asia
  - 2.5Gbps
- **IHEP- Domestic Univ**
  - IHEP-CSTNet-CERNet-Univ
  - 10Gbps

Bandwidth utilization of the links between IHEP and USA/EUR/ASIA?

Motivation--IHEP Traffic Status

- IPv4 average traffic: 2.5Gbps(in+out), Max. 7Gbps
- IPv6 average traffic: 1Gbps(in+out), Max. 1.5Gbps
- Data exchange: over 7PB/year



Motivation

- Know clearly for: Who, When, What, Where, Why in the network traffic

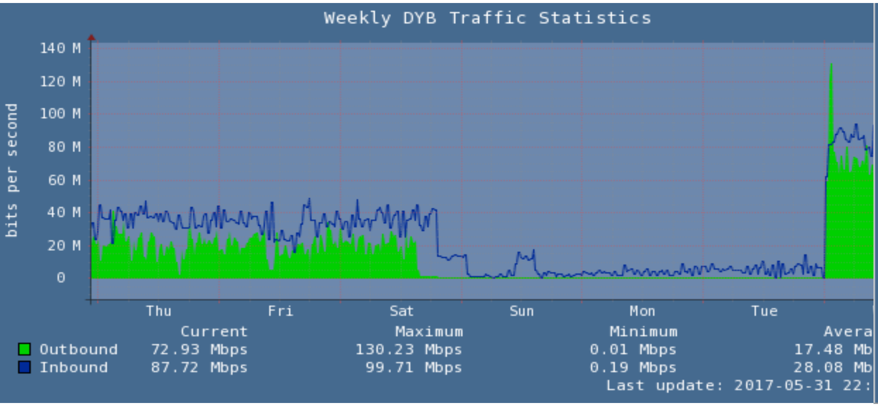
WHY
WHO
5W
WHAT
WHAT
WHERE

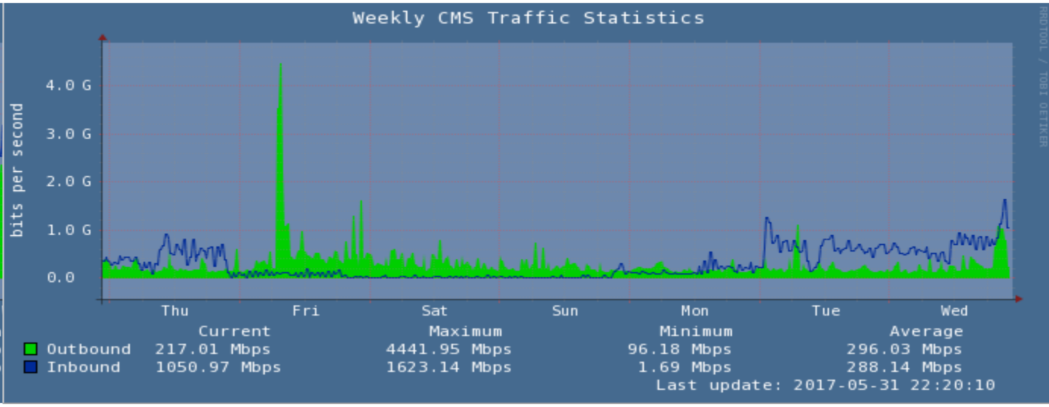
- Network optimization--Reliability/Performance/Efficiency
- Historical trends for strategic planning
- Network security analysis
- ....

Who: IP addresses / Users
When: on which time
What: protocols, ports, data traffic, applications, etc.
Where: flow direction, which countries/regions (max.volume)
Why: malicious attacks or normal data transfer

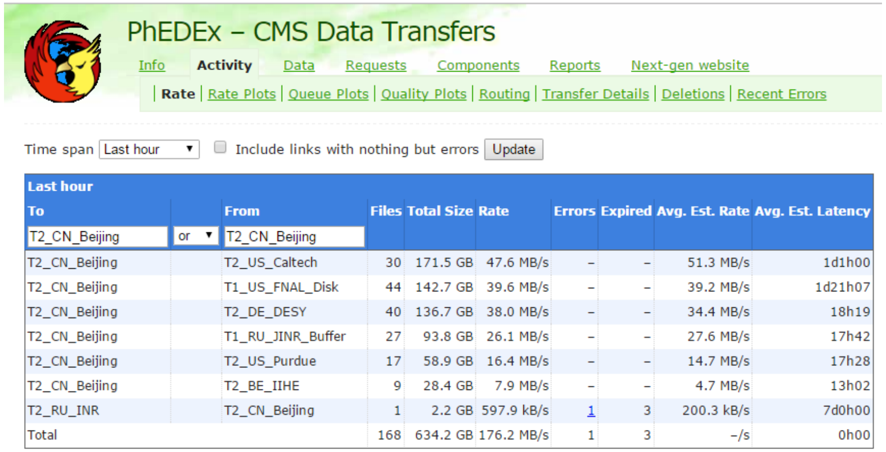
Motivation--IHEP traffic statistics status

- Traffic statistics based on IP address range： Daya Bay、CMS、ATLAS





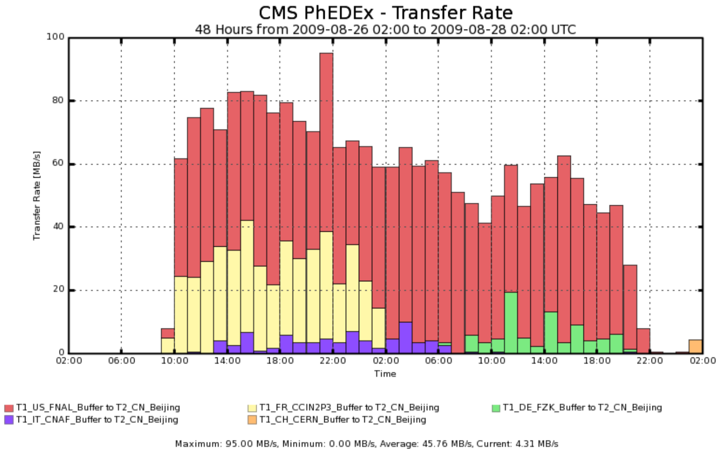
- Traffic statistics based on experimental data transfer system： PhEDEx



Time span Last hour ▾ □ Include links with nothing but errors

Last hour

To		From	Files	Total Size	Rate	Errors	Expired	Avg. Est. Rate	Avg. Est. Latency
T2_CN_Beijing	or ▾	T2_CN_Beijing							
T2_CN_Beijing		T2_US_Caltech	30	171.5 GB	47.6 MB/s	–	–	51.3 MB/s	1d1h00
T2_CN_Beijing		T1_US_FNAL_Disk	44	142.7 GB	39.6 MB/s	–	–	39.2 MB/s	1d21h07
T2_CN_Beijing		T2_DE_DESY	40	136.7 GB	38.0 MB/s	–	–	34.4 MB/s	18h19
T2_CN_Beijing		T1_RU_JINR_Buffer	27	93.8 GB	26.1 MB/s	–	–	27.6 MB/s	17h42
T2_CN_Beijing		T2_US_Purdue	17	58.9 GB	16.4 MB/s	–	–	14.7 MB/s	17h28
T2_CN_Beijing		T2_BE_IIHE	9	28.4 GB	7.9 MB/s	–	–	4.7 MB/s	13h02
T2_RU_INR		T2_CN_Beijing	1	2.2 GB	597.9 kB/s	1	3	200.3 kB/s	7d0h00
Total			168	634.2 GB	176.2 MB/s	1	3	–/s	0h00



Lack of overall fine-grained network traffic statistics and analysis

Lack of user behavior analysis and intrusion detection in cyber security

Outline

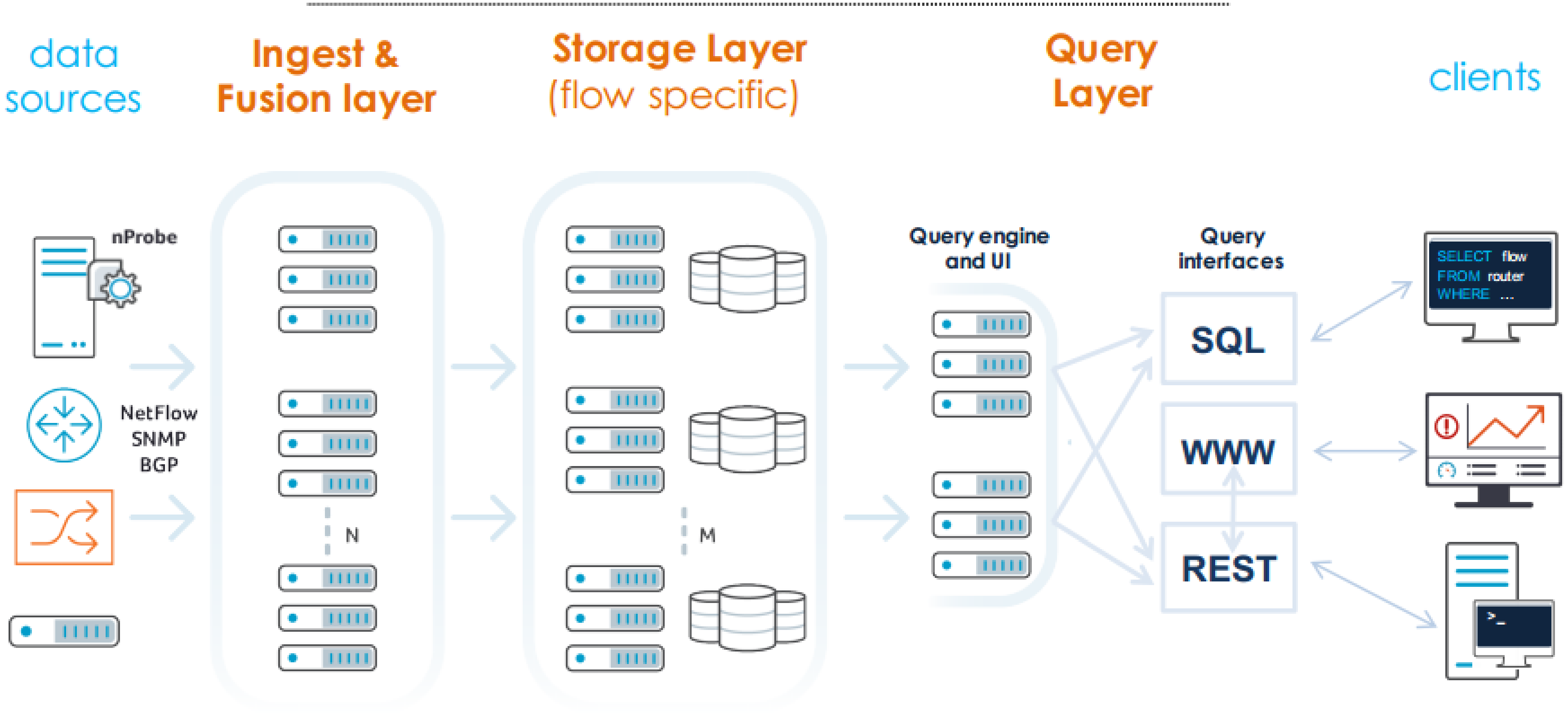
1. Motivation
2. Platform design
3. Function modules
4. Future Plan
5. Summary

General design principles

- Traffic of high-speed network can be captured without missing： 10Gbps
- Traffic flow records should include the following elements: 5-Tuple(src_ip, src_port, dst_ip, dst_port, protocal)
- Large amount of historical data can be stored and queried efficiently: at least 1 year raw data
- Data analysis module should be extensible (add or remove analysis plugins)
- Flexible and friendly user interface

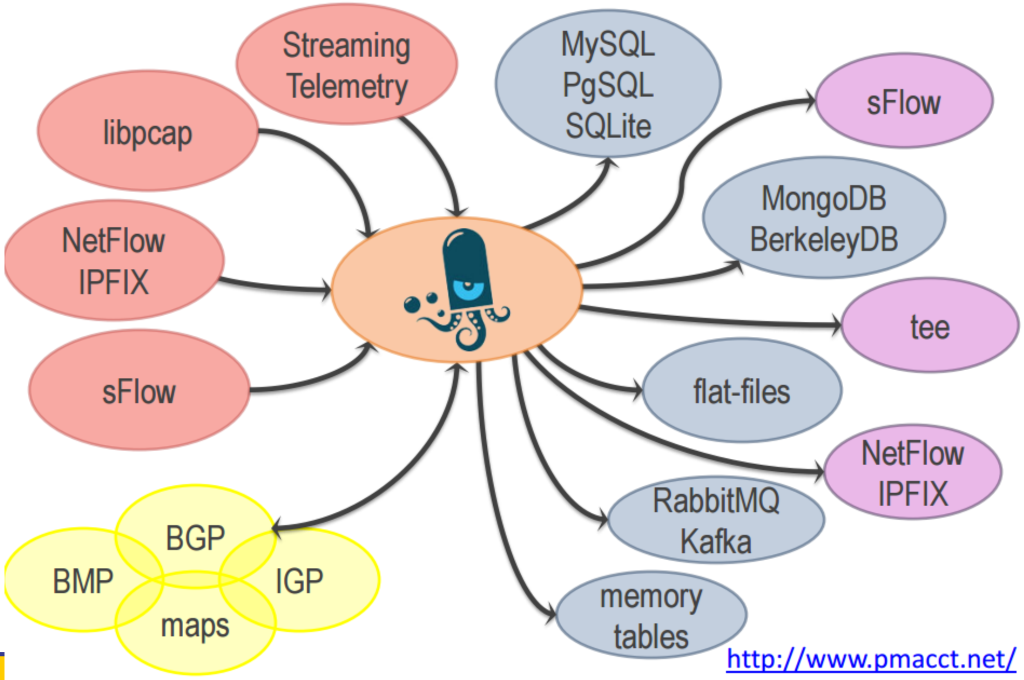
Architecture

- Data Sources + Data preprocessing (ingest & fusion) + Storage + Data Query + Graphic Display
- Principle: loose coupling between layers--extensible



Key tool— pmacct

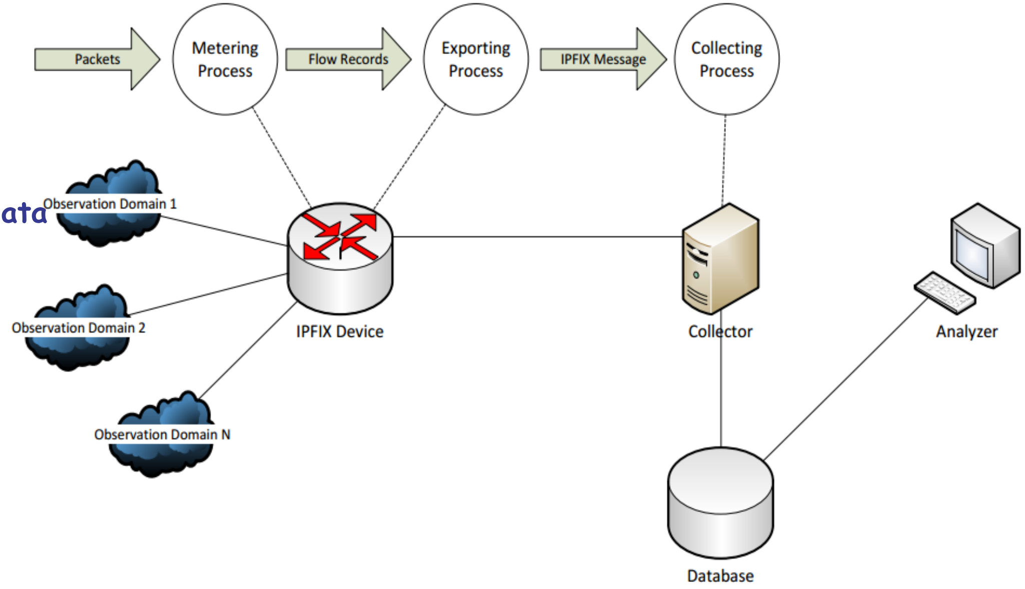
- Open source software
- A small set of multi-purpose passive network monitoring tools which can account, classify, aggregate, replicate and export forwarding-plane data, ie. IPv4 and IPv6 traffic;
- Collect data through: libpcap, Netlink/NFLOG, NetFlow v1/v5/v7/v8/v9, sFlow v2/v4/v5 and IPFIX
- Save data to backends including: Relational Databases, NoSQL databases, RabbitMQ, Kafka, memory tables, flat files

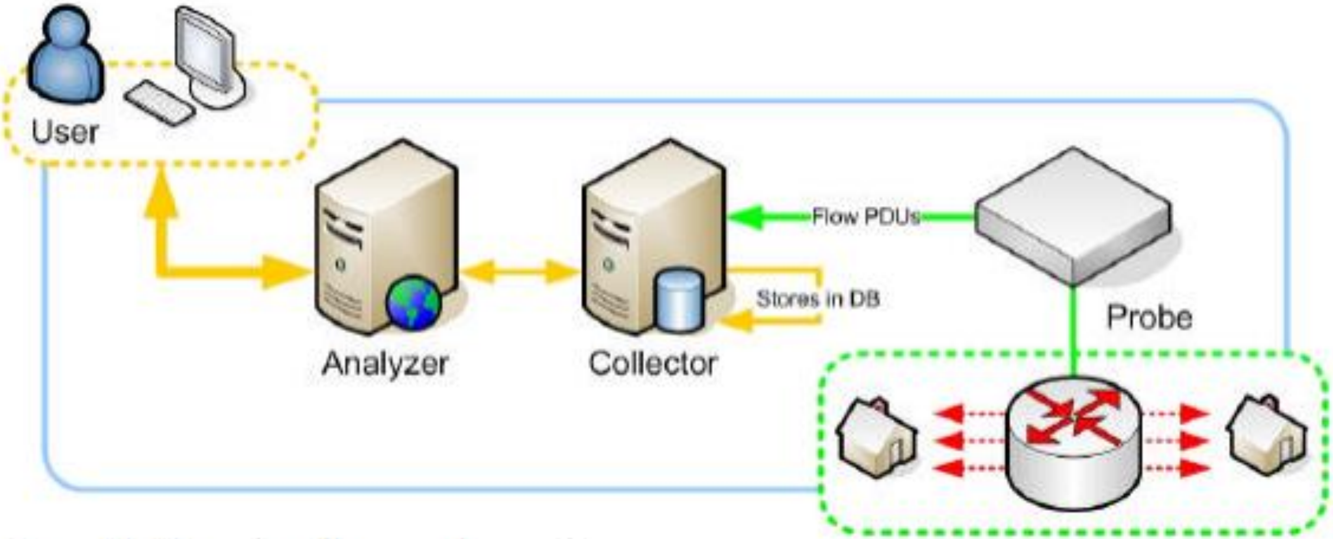


http://www.pmacct.net/

pmacct workflow

- **Workflow:**
  - » Receives sFlow/netflow/IPFIX data from devices
  - » Network flow aggregate
    - » Reduces data diversity
    - » Precompiles statistics -> reduces amount of data
    - » Looks like RRD (round robin database)
  - » Resolution of statistics
    - » sFlow 15 seconds
    - » IPFIX 5 minutes
  - » Output with tab-spaced/CSV/Avro/JSON format





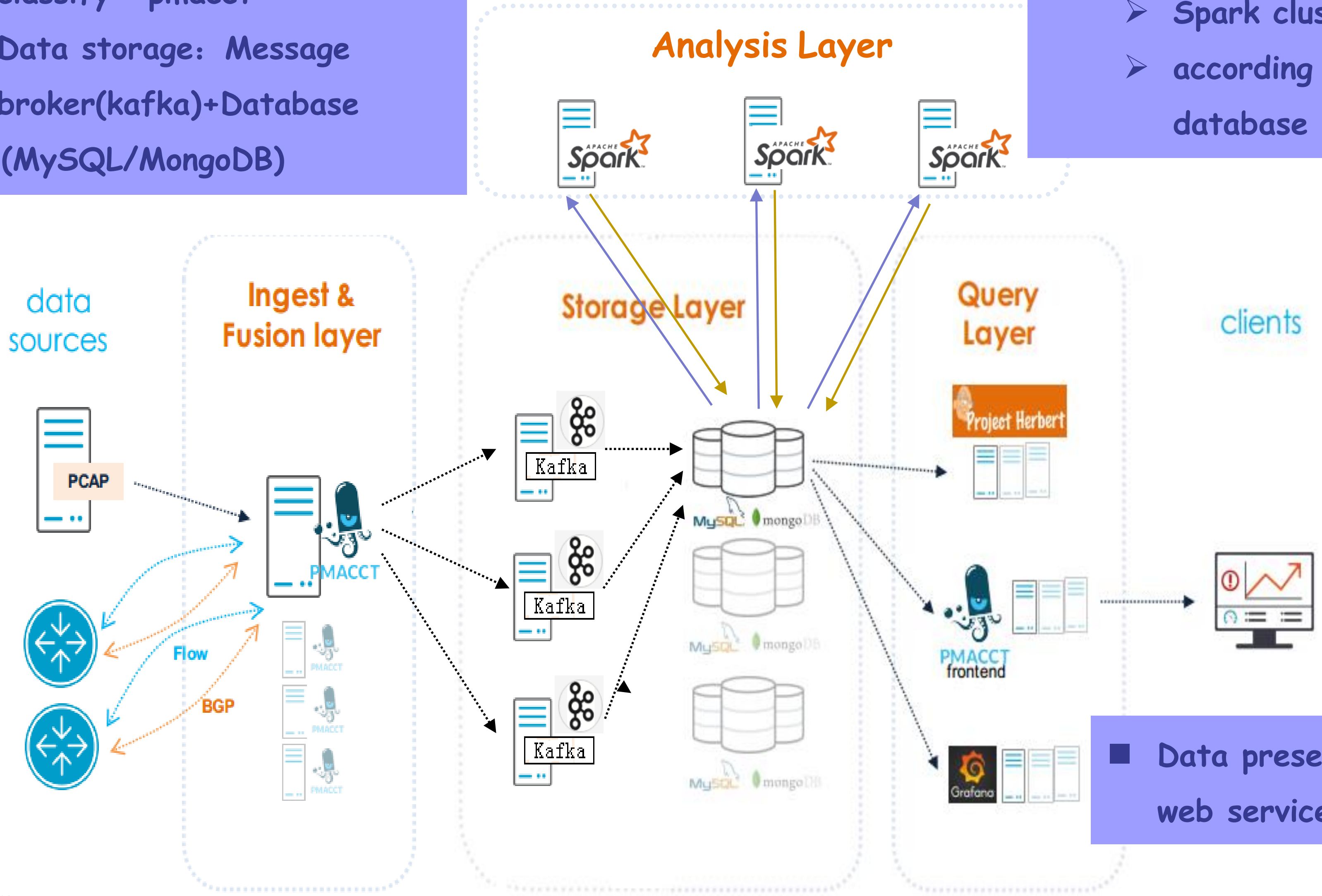
Outline

1. Motivation
2. Platform design
3. Function modules
4. Future Plan
5. Summary

Realization: Function modules

- Data acquisition: aggregate, classify--pmacct
- Data storage: Message broker(kafka)+Database (MySQL/MongoDB)

- Data analysis:
  - ➢ Spark cluster
  - ➢ according to GeoIP database



- Data presentation: web service + Echarts

Data Acquisition--pmacct

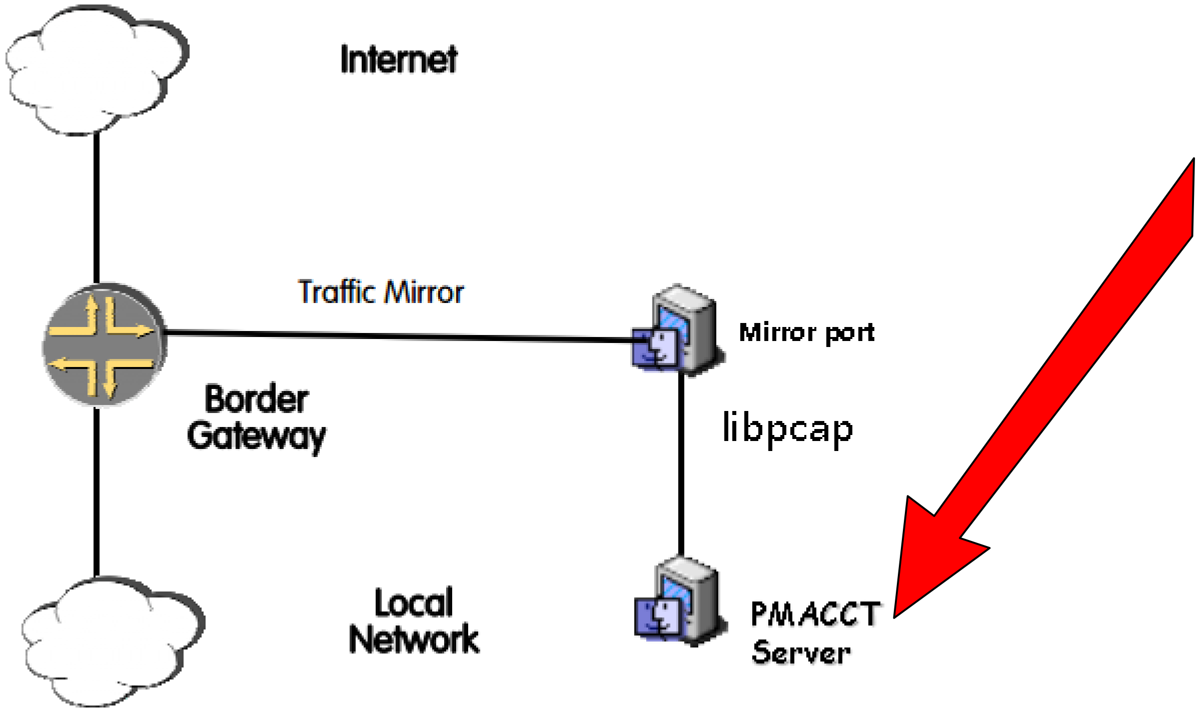
- Data source：

  collected from border router of IHEP

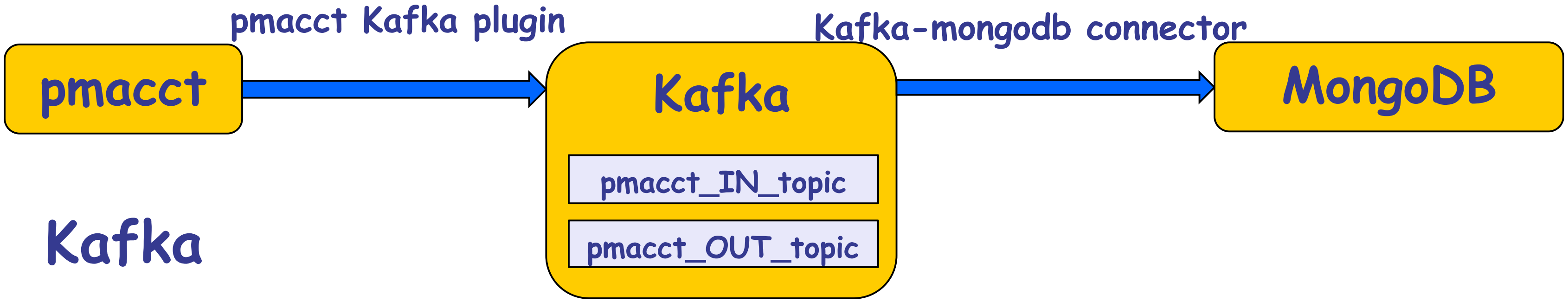
- Key points:
  - PMACCT server collects data through libpcap (according to the network device)
  - Mirror port: avoids the impact on the performance of network devices

Data preprocessing

- purpose：improving the efficiency of data storage and data reading
- PMACCT configuration：aggregation, filtering, classification
- aggregation：5mins
- rules：src_ip, dst_ip, src_port, dst_port, proto
- filtering：pcap_filter（collect IN/OUT source data）



Data storage: Kafka + MongoDB

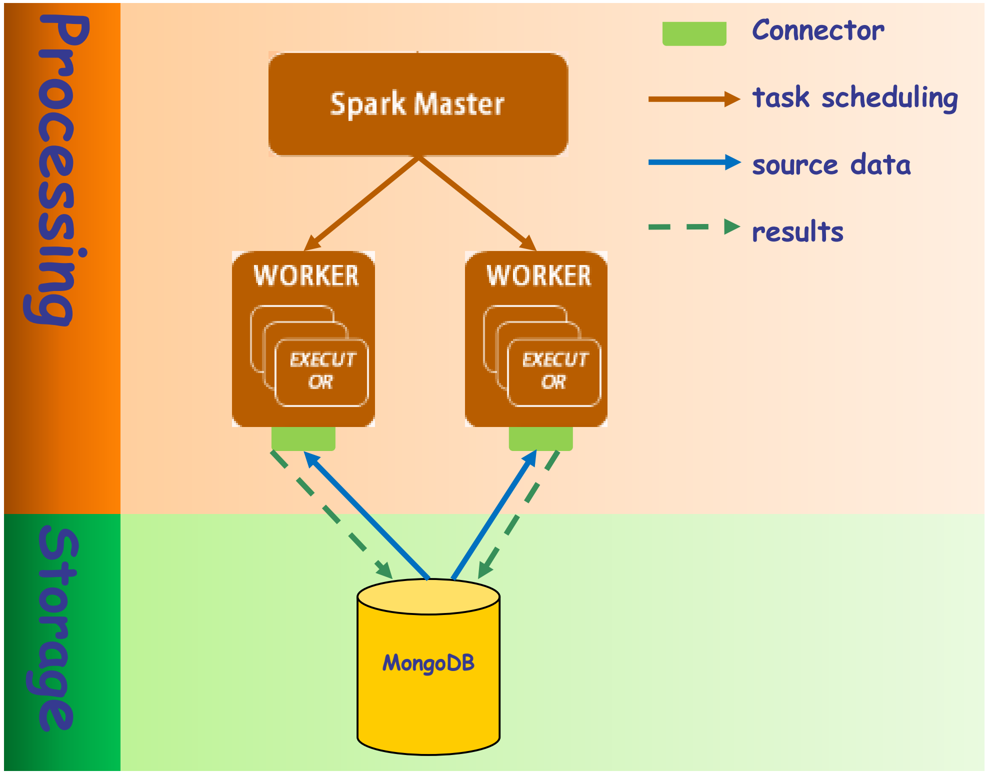


- **Kafka**
  - fast, scalable, durable and distributed
  - for big data: Millions of records within 5 minutes(IN+OUT)
  - The data within 1 week are preserved
- **MongoDB**
  - Distributed and document-oriented database(15G/month)
  - With high insert and query performance, the average insert speed is up to 7000 records/s
  - Source data are saved as:

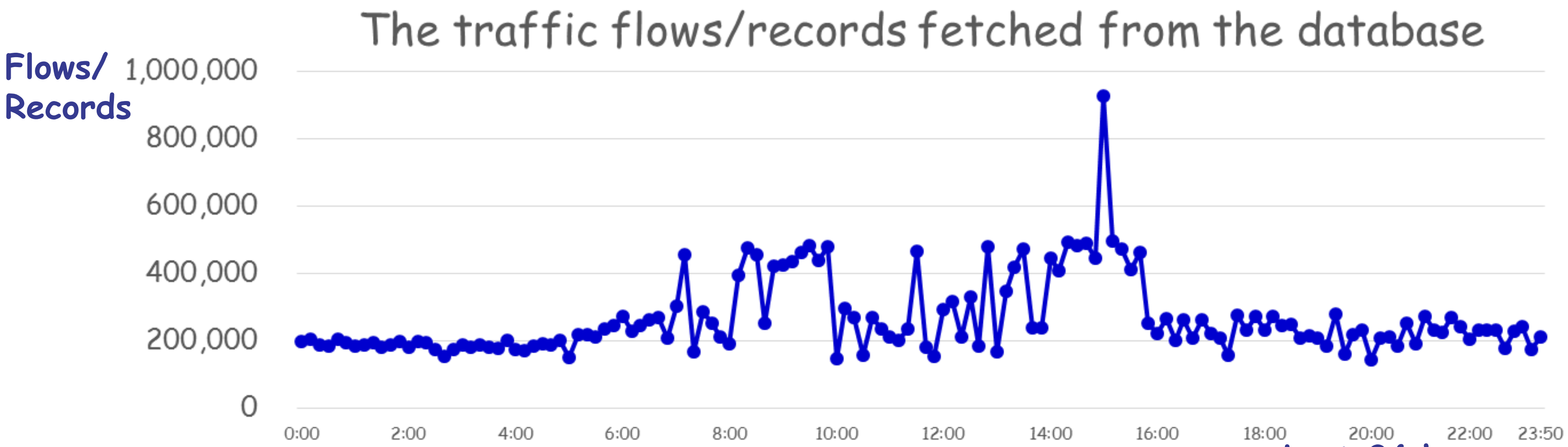
ip_src	ip_dst	src_port	dst_port	ip_proto	packets	bytes	stamp_ inserted	stamp_ updated

Data analysis

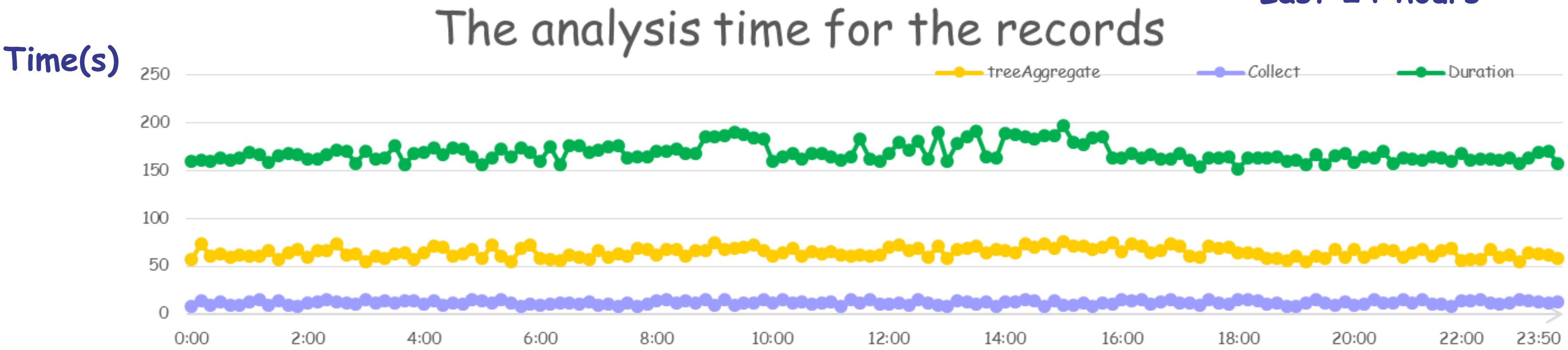
- Open-source & cluster-computing framework: Spark cluster.
- Task: timed cron jobs are set to calculate IN/OUT cumulative traffic between IHEP and domestic/international IP addresses within 10mins, 1hour, 1day separately.
- GeoLite2 database is used to classify the regional information of the traffic.



Data analysis—processing efficiency



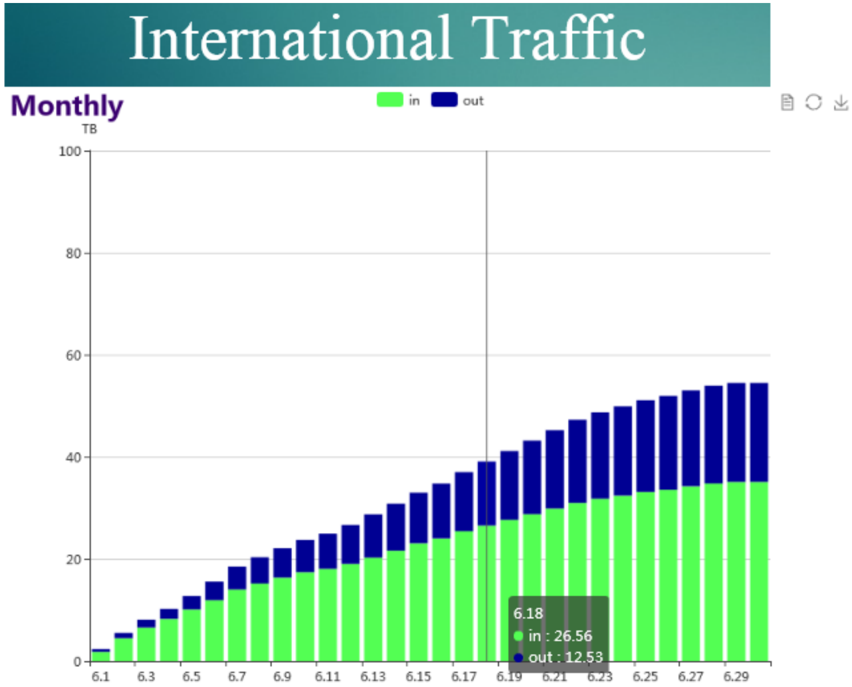
Last 24 hours



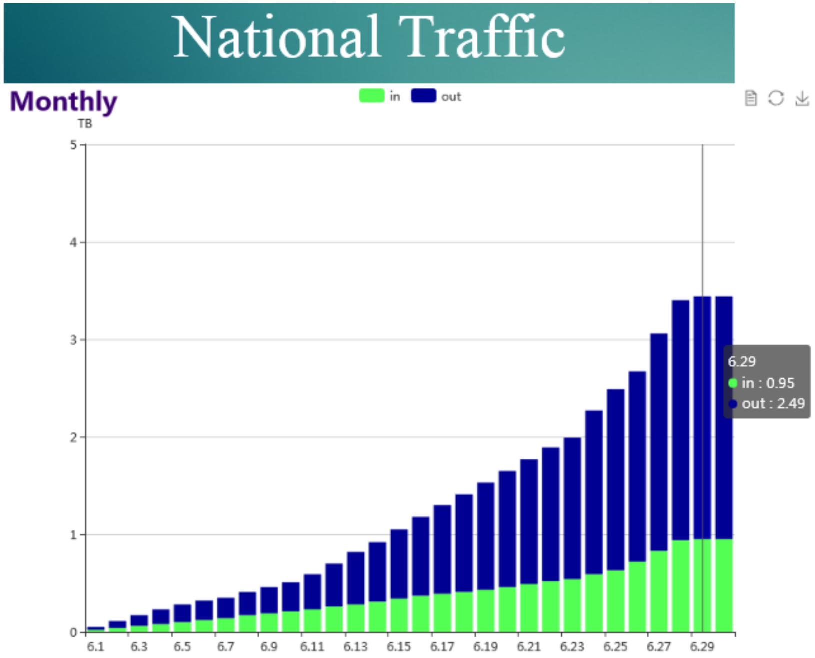
Last 24 hours

- The amount of data records every 10 mins (IN): 200,000 – 900,000
- Processing time is stable: less than 200 seconds
- Computing source can be increased for larger volumes of data

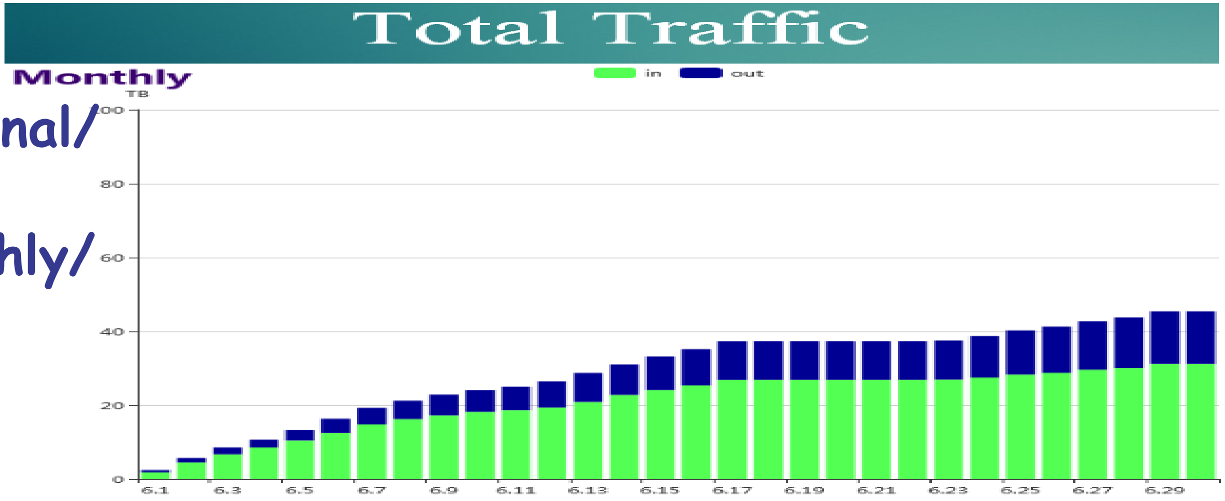
Data Presentation



International Traffic



National Traffic



Total Traffic

1. International/National/ Total traffic
2. Daily/Weekly/Monthly/ Yearly

Outline

1. Motivation
2. Platform design
3. Function modules
4. Future Plan
5. Summary

Future Plan

- Data Analysis
  - Rule-based network intrusion detection module will be added to identify the malicious action in network(DDoS attack, Scans, Worms)
  - User behavior analysis(P2P Apps, Botnets)
- Display
  - GeoIP plugin will be used to display regional traffic data on a map.

Summary

- A framework with network traffic data acquisition, storage, analysis and graphic interface has been finished.
- The network traffic statistics based on IP prefix and GeoLite2 database has been realized.
- Network security detection plugin and network traffic statistics display on map are under developing.

Thank you for your attention!