Research and implementation of IHEP Network performance analysis platform

Shan Zeng

on behalf of IHEP network group

Funded by NSFC (No. 12175258)

zengshan@ihep.ac.cn

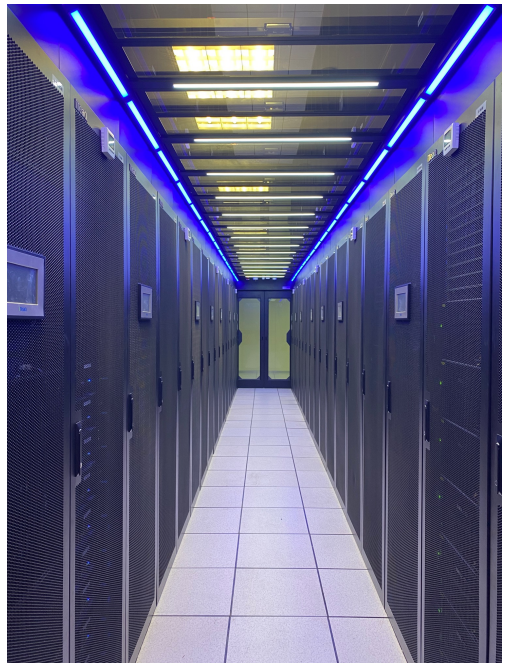
3-25-2022

Outline

- Background
- IHEP network overview
- Platform architecture
- Analysis method and process
- Use cases
- Future plan
- Summary

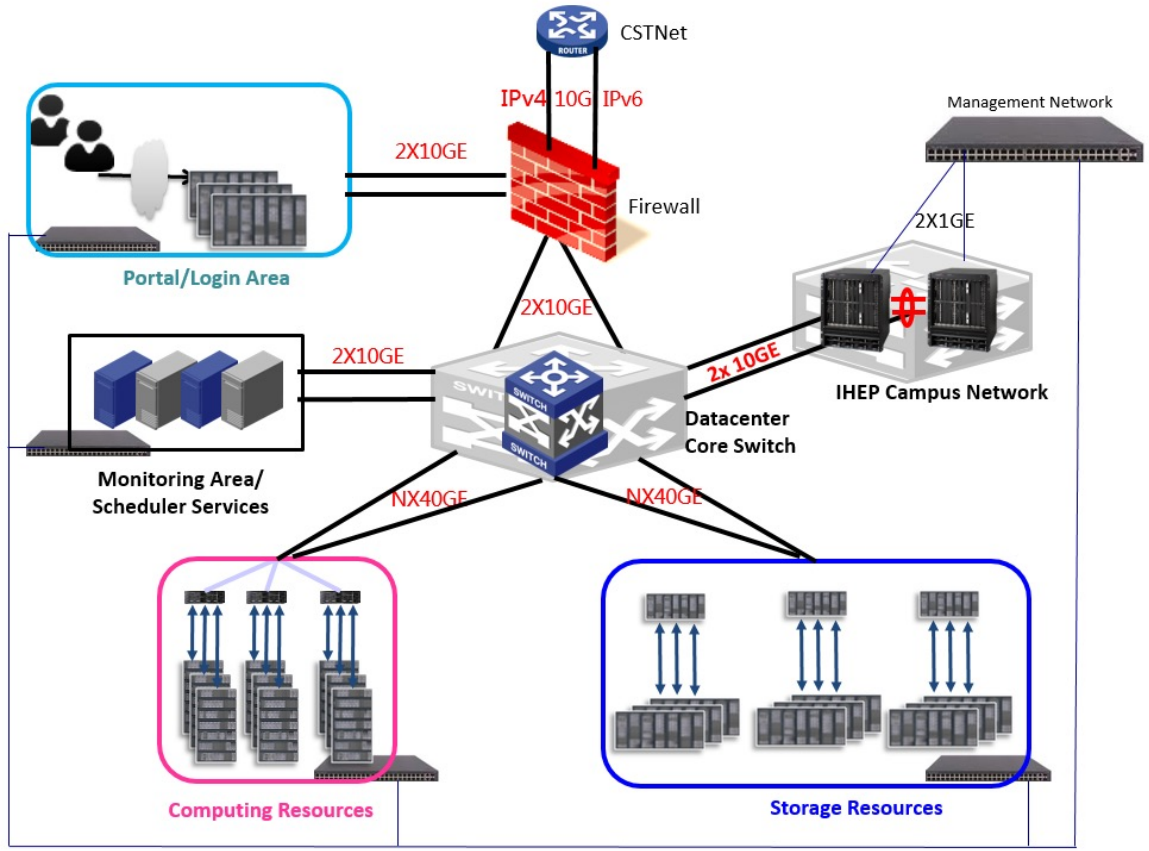
Background

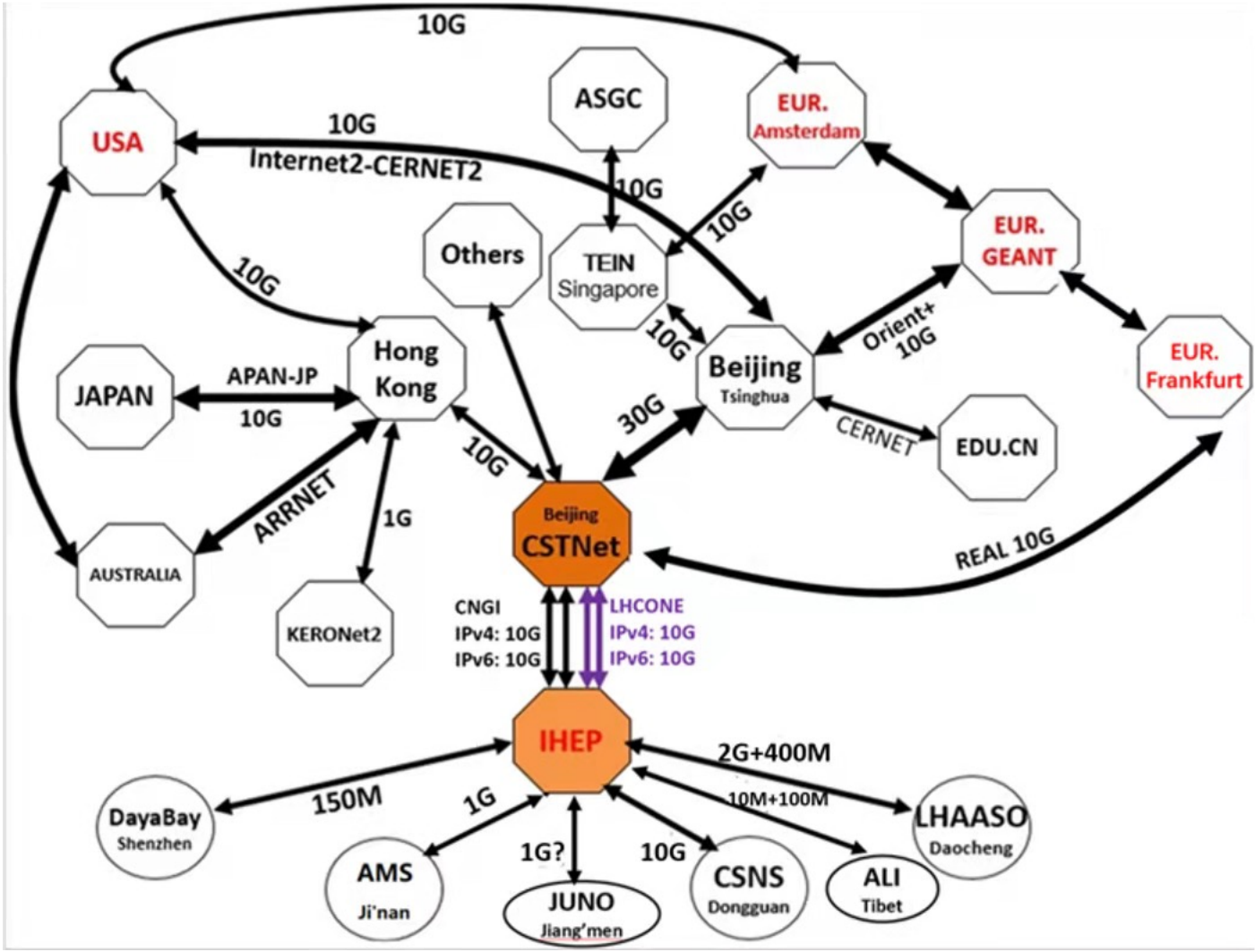
- **More and more large scientific facilities are being built or running in IHEP**
- **Network becomes more and more important in:**
  - Data transferring from onsite to offline data center
  - Offline data analysis: computing and storage
- **More features are needed in high performance network**
  - Design philosophy
    - High bandwidth and low latency
    - Stable, scalable and Flexible
  - Performance visualization ability
    - Network Performance analysis bind with application performance
    - Network capabilities by real-time analysis



IHEP network overview

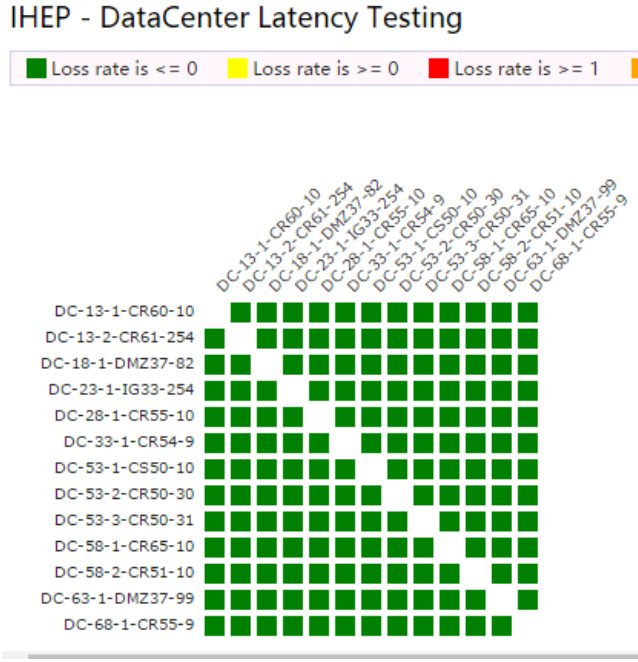
- **With 40Gb/s WAN bandwidth**
- **Data center network**
  - Backbone bandwidth: 800Gb/s
  - Support IP/IB network

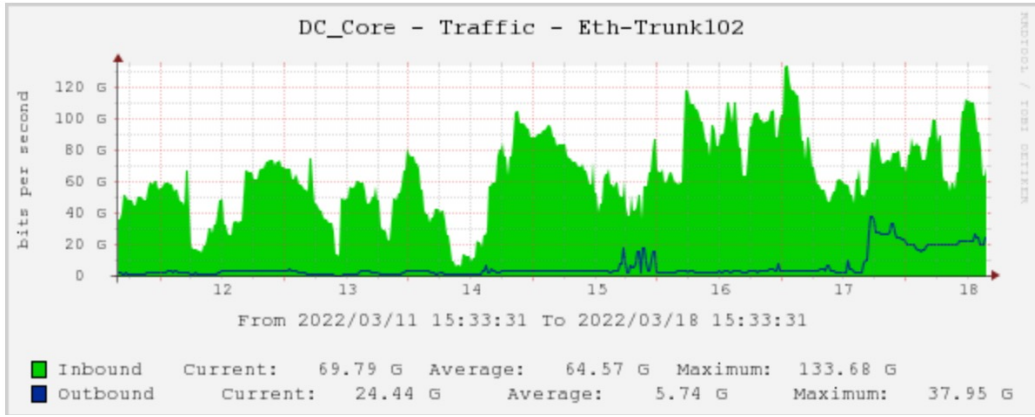




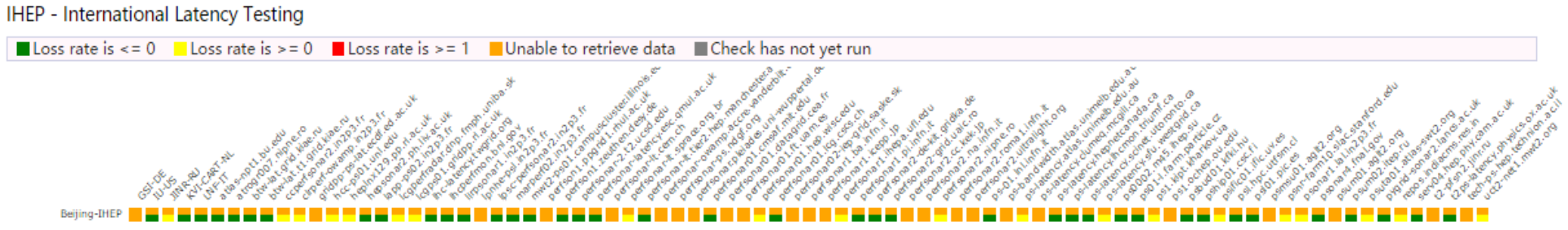
Network monitoring tools and methods

- **Network device status monitoring**
  - Port up/down
  - Port traffic
  - Zabbix/cacti/nagios
- **Connectivity monitoring**
  - Loss
  - Latency
  - PerfSonar



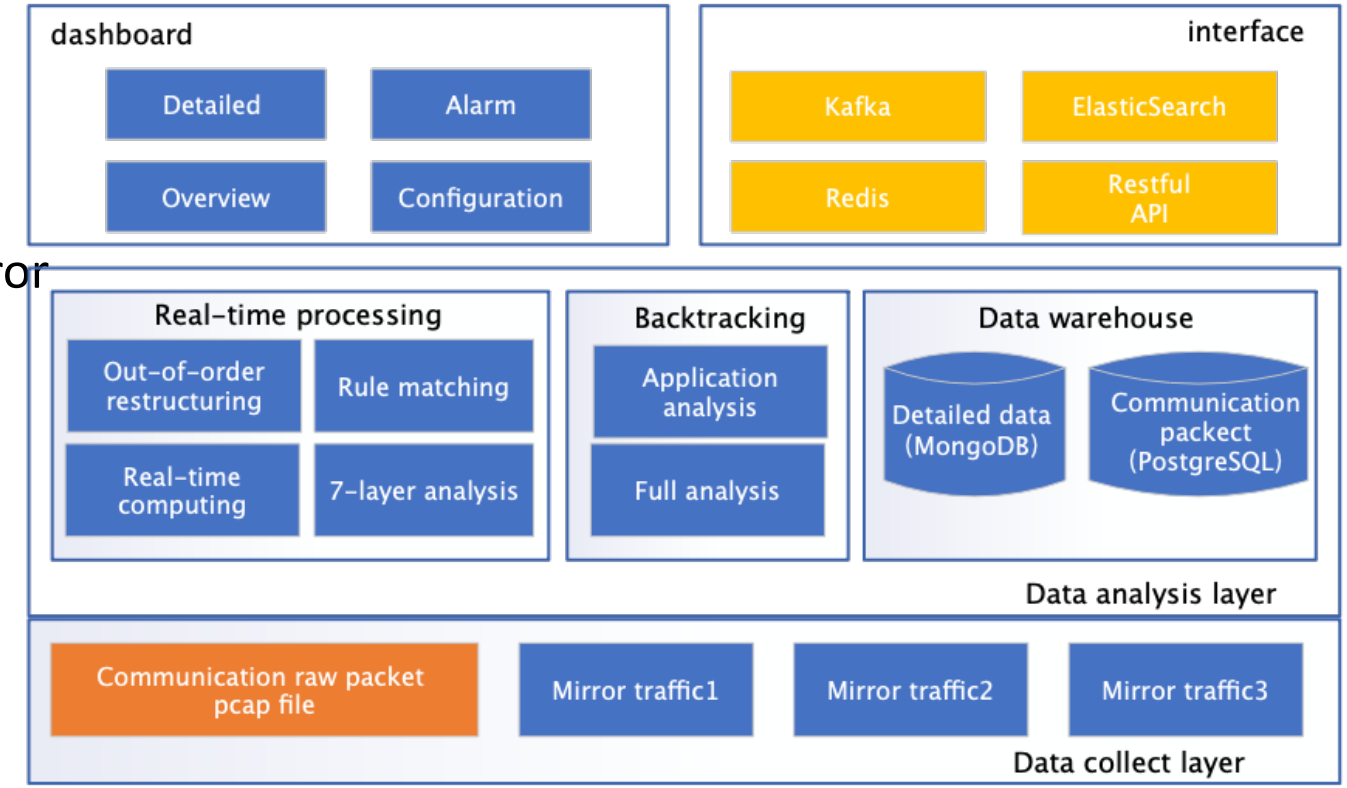


- **Lack the network performance analysis related to application performance**



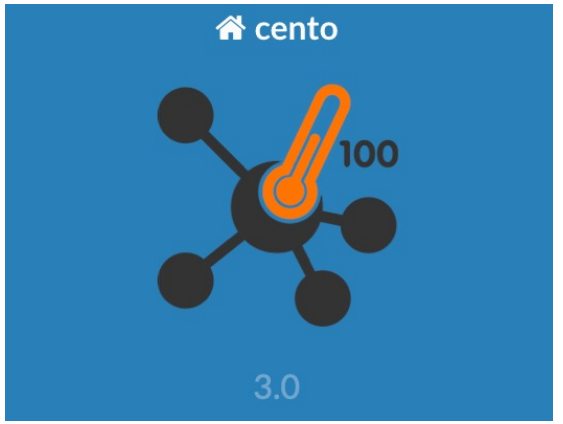
Platform architecture

- **What we get?**
  - Network traffic
  - Communication raw packet
- **Where are they from?**
  - Network traffic: router/core switch traffic mirror
  - Communication raw packet: tcpdump
- **How we did?**
  - Real-time processing
    - Rule matching
    - 7-layer analysis
  - Application related traffic and packet analysis
    - Full traffic
  - Data warehouse
    - MongoDB: Detailed data
    - PostgreSQL: Communication packet



nProbe cento

- nProbe is both a netFlow v5/v9/IPFIX Probe and a collector that can be used to play with NetFlow flows for IPv4/v6
- nProbe cento is a 1/10/40/100 Gbit NetFlow/IPFIX Probe, traffic classifier, and packet shunter
  - A high-speed NetFlow probe able to keep up with 1/10/40/100 Gbit.
  - Flow export to JSON, Text, Kafka, Syslog, ntopng
  - Native PF_RING and PF_RING ZC support for high-speed packet processing
  - Flow-based Load Balancing to IDS/IPS
  - Layer-7 application visibility using nDPI (Deep Packet Inspection) or micro-nDPI (a lightweight DPI library supporting the most important protocols such as HTTP/HTTPS/DNS) for improved performance
  - Flow-based Load Balancing to IDS/IPS

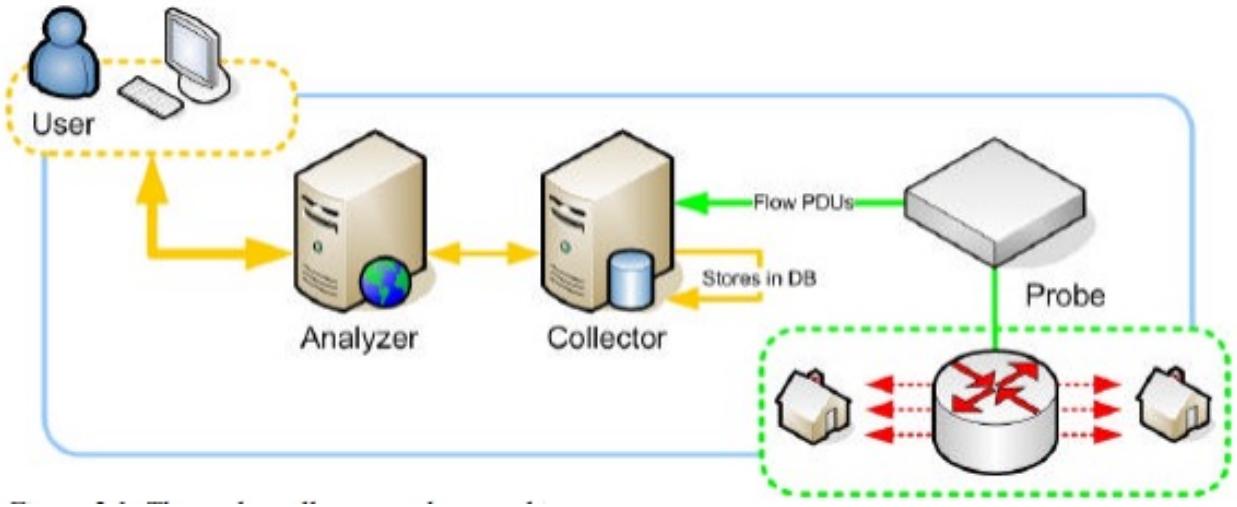


Probe and Flow Exporter

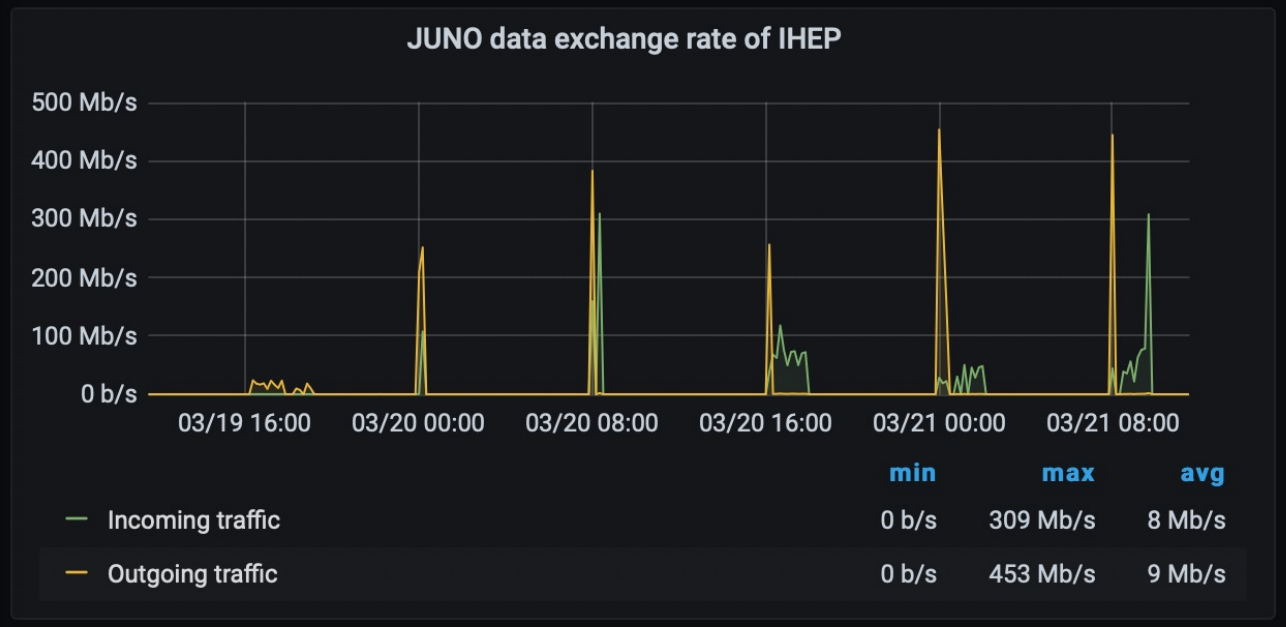
```
cento -i zc:eth1 --v9 192.168.1.200:2055
```

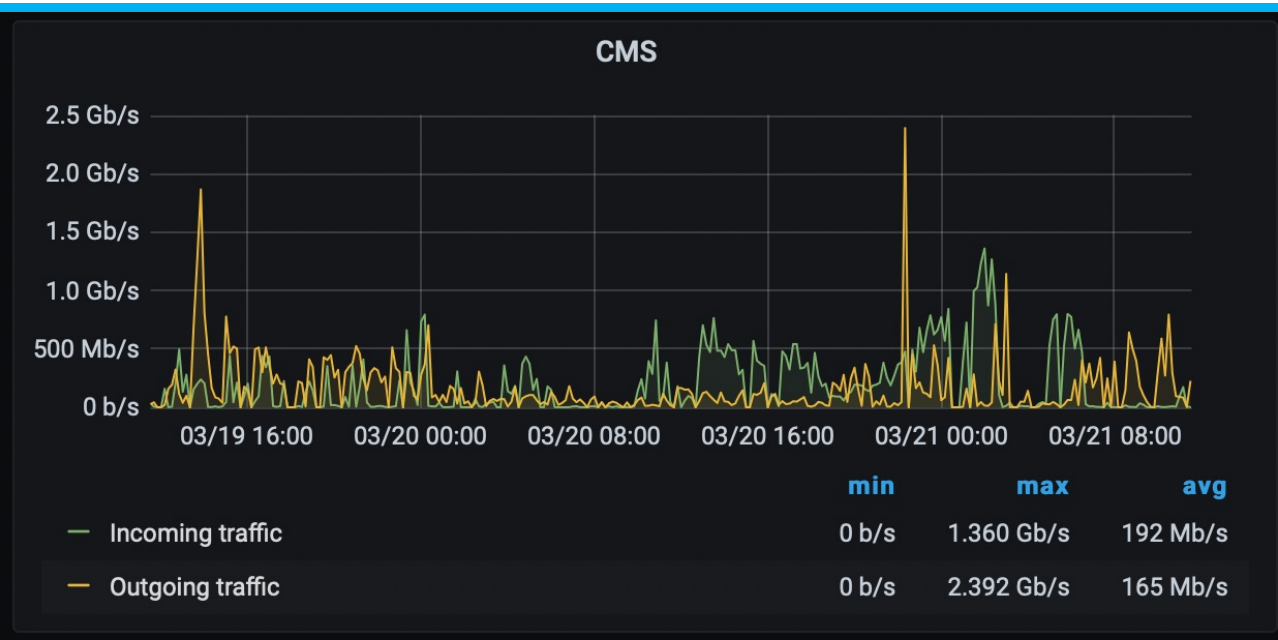
Netflow analysis process

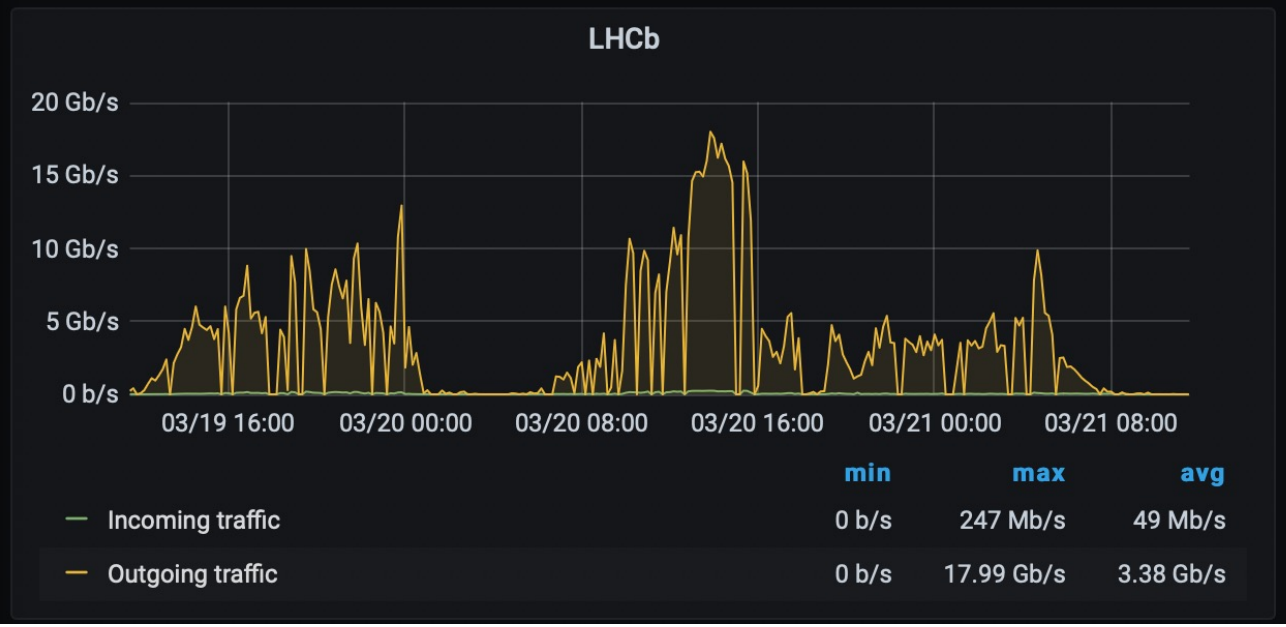
- Analyze incoming and outgoing traffic of high-energy physics experiments in IHEP
- Including: CMS/ATLAS/LHCb/JUNO
- Right now focus on the IPv4 traffic
  - Mirror the router traffic to analysis node
  - Capture and store the traffic in the analysis node by nfcapd
    - nfcapd_file/IHEP/2022/03/18/nfcapd.202203180827
  - Dump the nfcapd file by nfdump
  - Classify experiments by IP peers and ports
  - Calculate traffic count by cluster analysis
  - Export the result to elasticsearch
  - Shown in Grafana dashboard

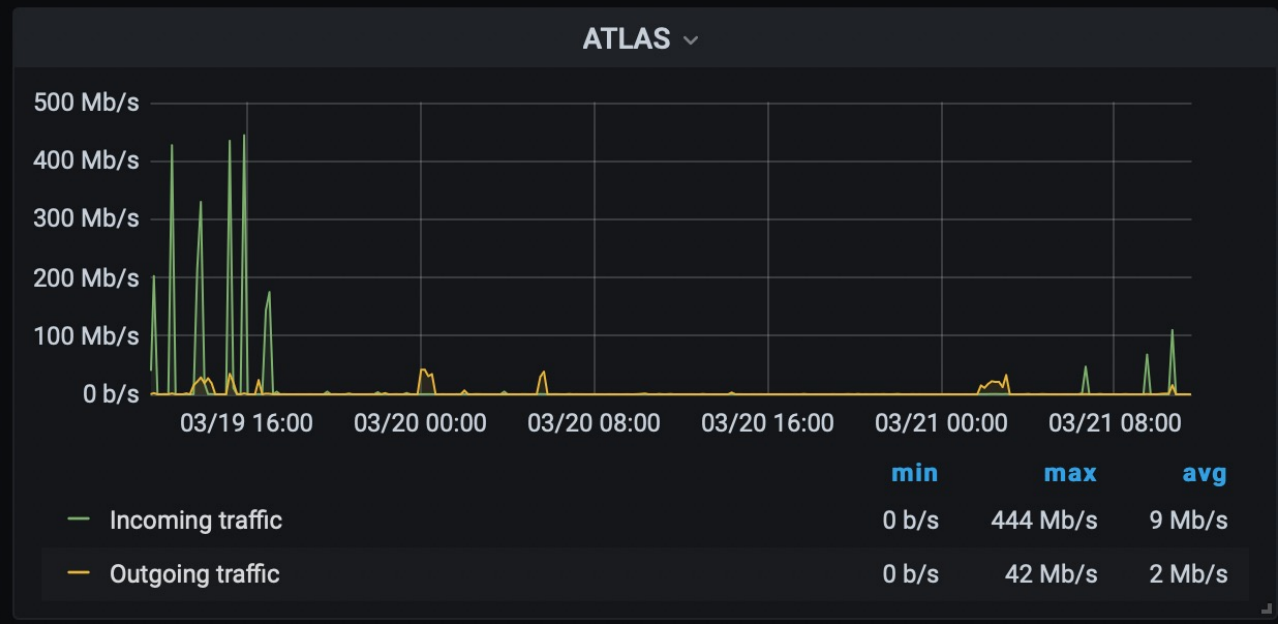


Use case


JUNO data exchange rate of IHEP
500 Mb/s
400 Mb/s
300 Mb/s
200 Mb/s
100 Mb/s
0 b/s
03/19 16:00
03/20 00:00
03/20 08:00
03/20 16:00
03/21 00:00
03/21 08:00
min max avg
Incoming traffic 0 b/s 309 Mb/s 8 Mb/s
Outgoing traffic 0 b/s 453 Mb/s 9 Mb/s

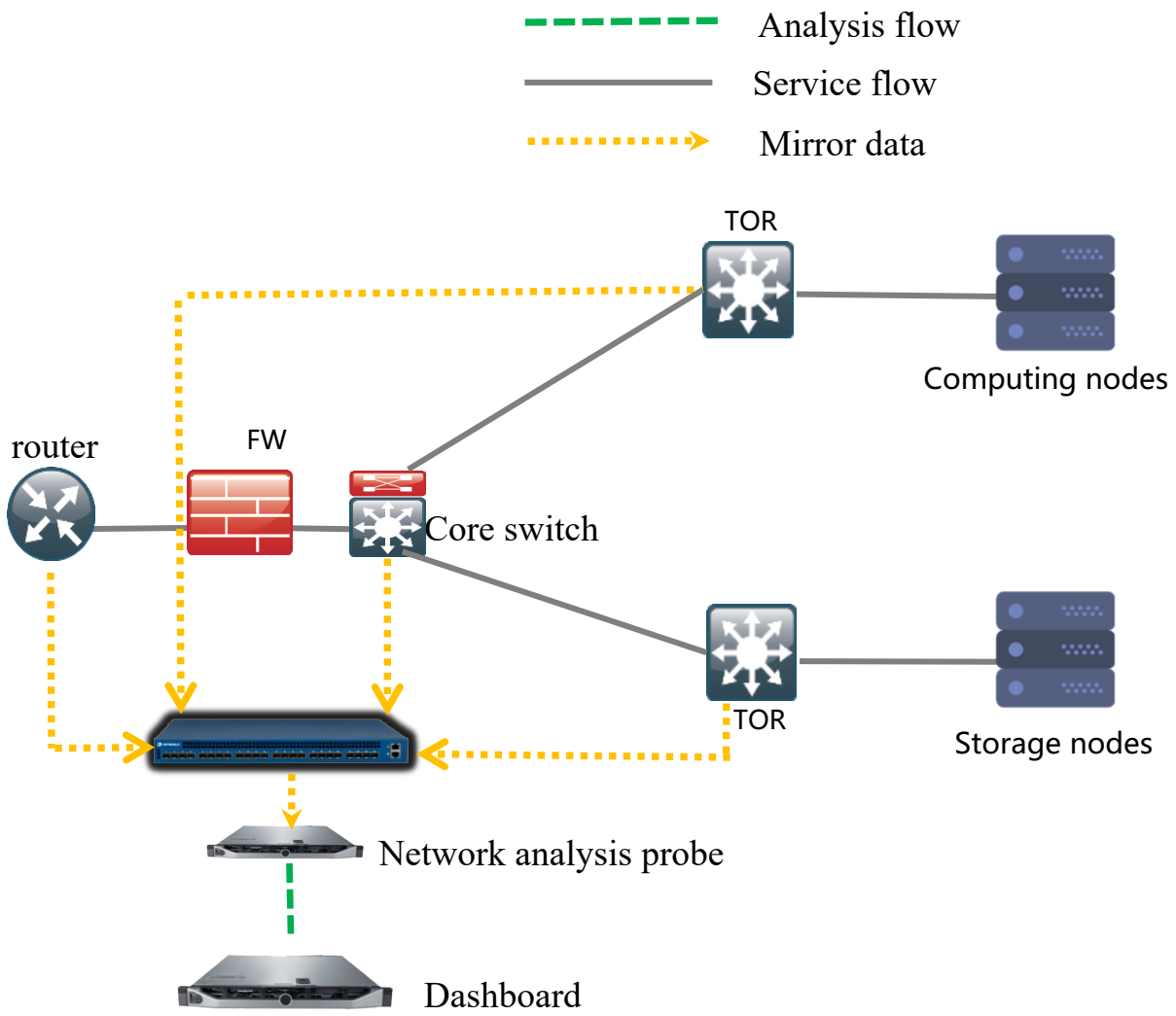

CMS
2.5 Gb/s
2.0 Gb/s
1.5 Gb/s
1.0 Gb/s
500 Mb/s
0 b/s
03/19 16:00
03/20 00:00
03/20 08:00
03/20 16:00
03/21 00:00
03/21 08:00
min max avg
Incoming traffic 0 b/s 1.360 Gb/s 192 Mb/s
Outgoing traffic 0 b/s 2.392 Gb/s 165 Mb/s


LHCb
20 Gb/s
15 Gb/s
10 Gb/s
5 Gb/s
0 b/s
03/19 16:00
03/20 00:00
03/20 08:00
03/20 16:00
03/21 00:00
03/21 08:00
min max avg
Incoming traffic 0 b/s 247 Mb/s 49 Mb/s
Outgoing traffic 0 b/s 17.99 Gb/s 3.38 Gb/s


ATLAS
500 Mb/s
400 Mb/s
300 Mb/s
200 Mb/s
100 Mb/s
0 b/s
03/19 16:00
03/20 00:00
03/20 08:00
03/20 16:00
03/21 00:00
03/21 08:00
min max avg
Incoming traffic 0 b/s 444 Mb/s 9 Mb/s
Outgoing traffic 0 b/s 42 Mb/s 2 Mb/s

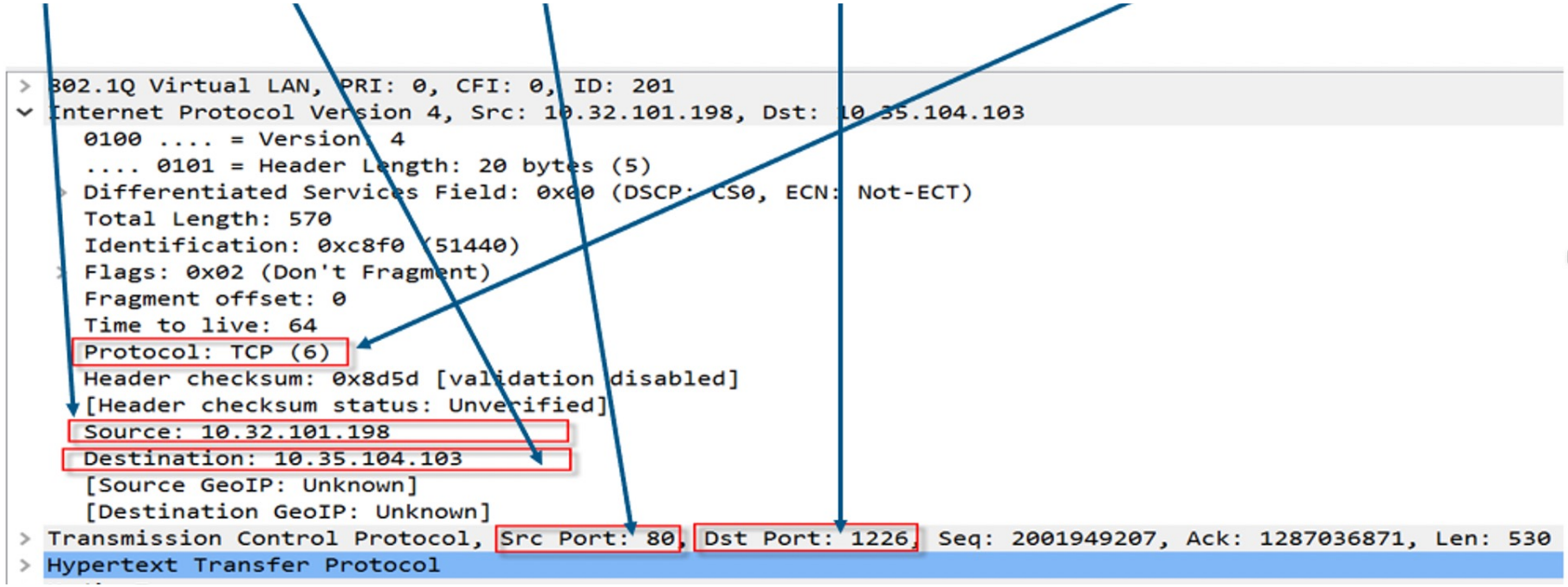
Data center application performance analysis from network view

- **Bypass deployment**
  - Mirror key network traffic
  - Analyze the traffic by a probe
- **Performance metrics**
  - Connection established analysis
    - Amount of retransmission
    - TCP retransmission rate
  - Response time
    - Server response time
    - Client response time
    - Service response time
  - Latency
    - Server latency
    - Client latency
  - Load analysis
    - Amount of concurrency



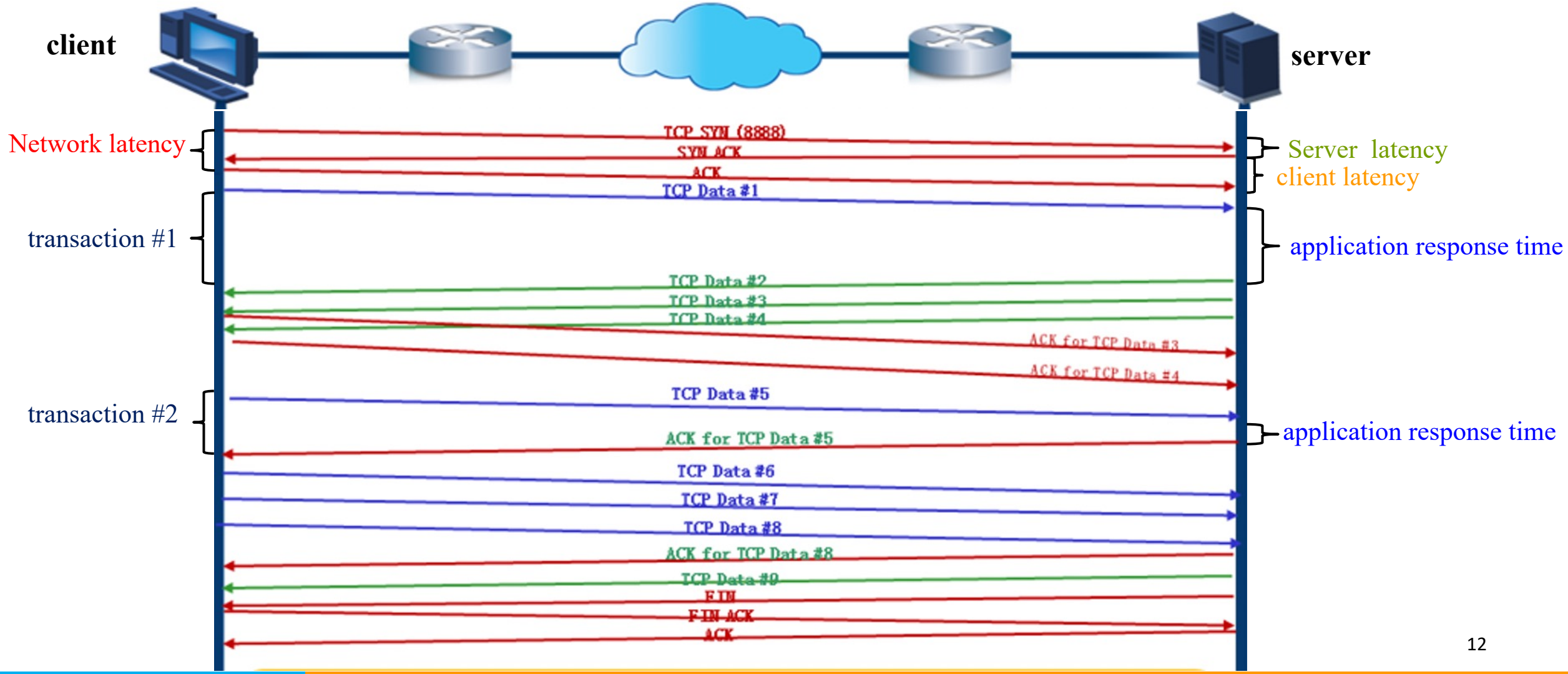
Key technologies of analysis probe (I)

- **High performance real-time processing**
  - Based on TCP/IP session
  - Application and service matching through five-tuple information
    - Source IP, Dest IP, Source Port, Dest Port, Protocol identity



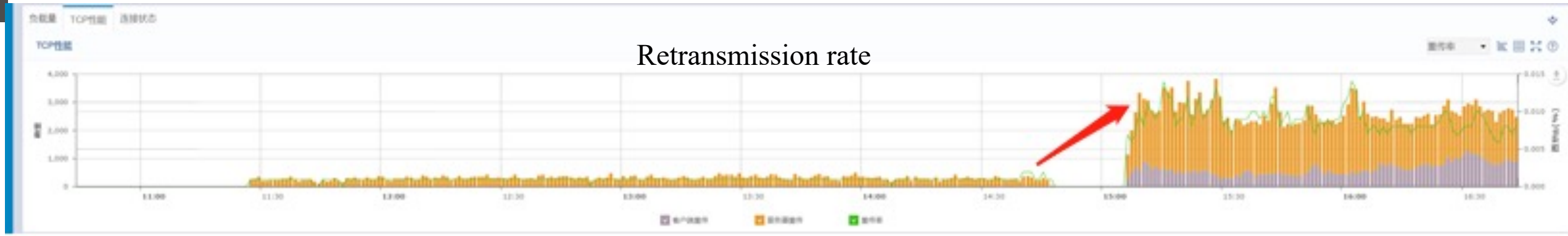
Key technologies of analysis probe (II)

- **Analysis the response time for both network and application/service by**
  - Deep analysis of three-way handshake and transmission process for TCP/IP



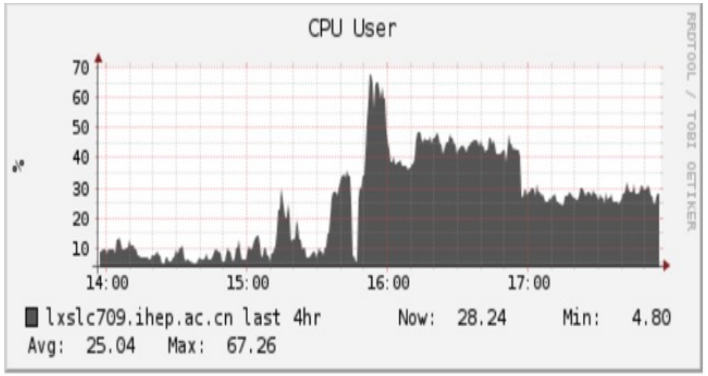
Use case

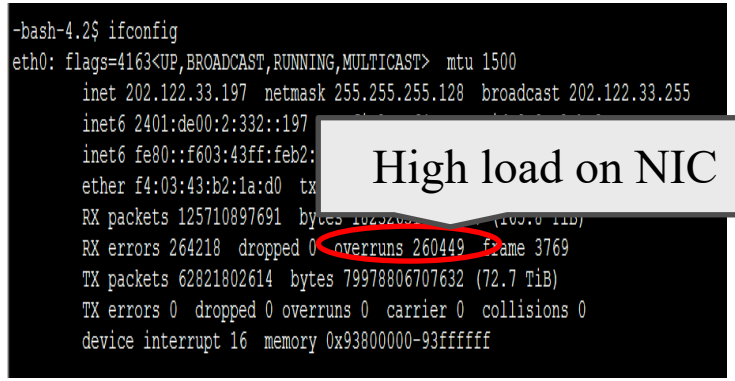
Retransmission rate



Eos server

Login node





High load on NIC

Future plan

- **IPv6 network traffic analysis based on experiments**
- **Deep and detailed data center network performance will be analyzed**
  - More application/service performance will be considered and combined
  - Backtracking will be added
- **Cross data center network performance analysis**
  - More metrics will be considered and designed
- **Dashboard should be optimized**

Conclusion

- **The purpose is to find the bottleneck of application performance**
- **IHEP started to research and implement the platform since the middle of last year**
  - Architecture design is finished
  - Some functions have been developed, such as WAN traffic analysis and DC application performance monitoring
  - More functions can be added since we capture enough traffic and packets
- **Future work**
  - IPv6 network traffic analysis based on experiments
  - Deep and detailed data center network performance will be analyzed
  - More metrics will be considered and designed
- **Any suggestions and cooperation are welcomed**

Thanks for your attention