BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF ENERGY
Office of Science

Intelligent, Scalable and Real-time Network Monitoring for Scientific Workflows

Anna Giannakou, Wenji Wu, Chin Guok, Jeronimo Bezerra
Lawrence Berkeley National Lab, ESnet, FIU

Contact: agiannakou@lbl.gov

ESnet

Overview

- Motivation
  - Science workflows and data requirements
    - Example 1: Emerging lightsource workflows
    - Example 2: CMS workflows
  - Large Science Data transfer challenges/anomalies
  - Existing solutions
  - Why existing solutions are insufficient
- Background (data sources)
  - High-touch
  - Q factor
- Our Solution
  - Using in-network telemetry data for real-time network monitoring
    - High-performance data transfer
    - Network security
    - Network planning
  - High-Touch: precision network telemetry services
  - Early results
- Future work
- Summary

Big data science workflows and data requirements

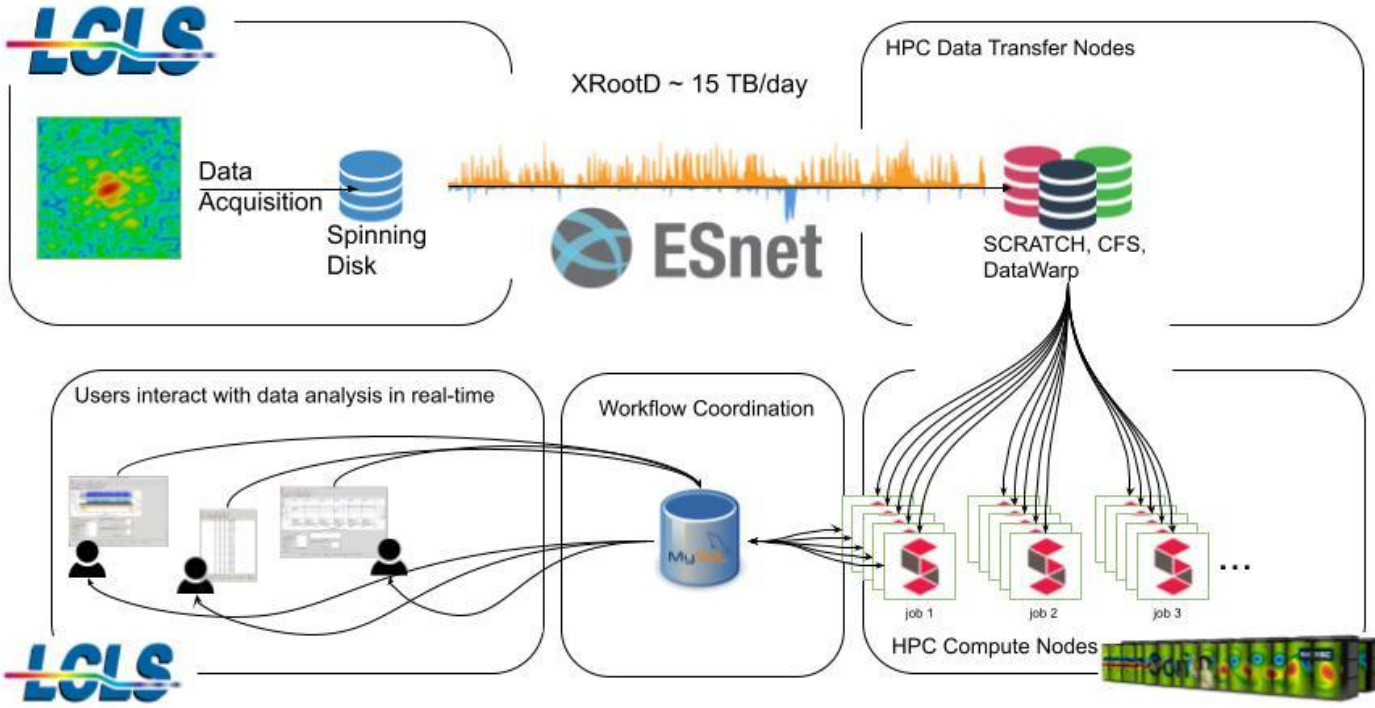
The LHC Accelerator Complex

DOE BES Structural Biology Resources

Today…

- Science instruments generate vast amounts of data
- Data must be collected, stored and analyzed in a distributed manner
- Emerging class of workflows requires fast result turnaround

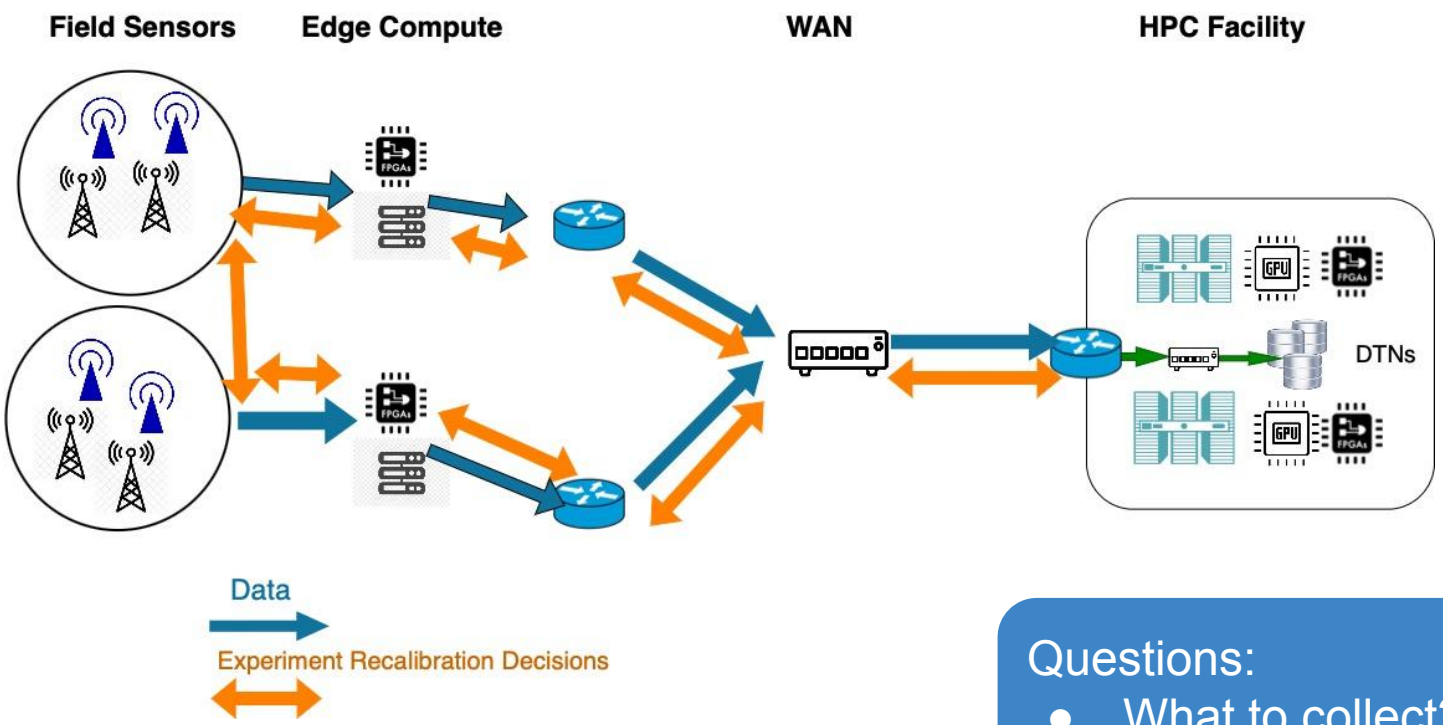
Emerging workflow: fast result turnaround for experiment calibration



Lightsource workflow class:

- 15TB/day data
- Near-real time result turnaround required
- Experiment reconfiguration based on results

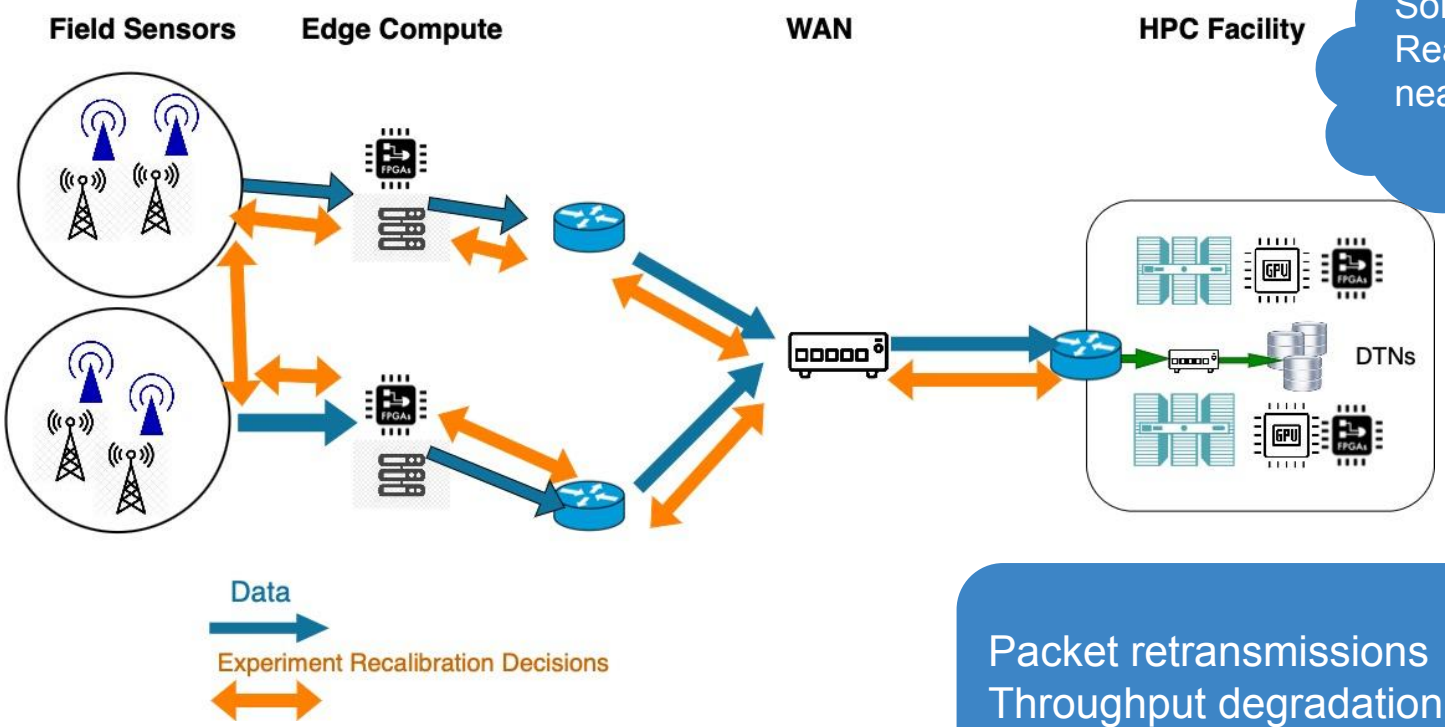
Monitoring data transfers is a challenging task



Questions:

- What to collect?
- When and where?

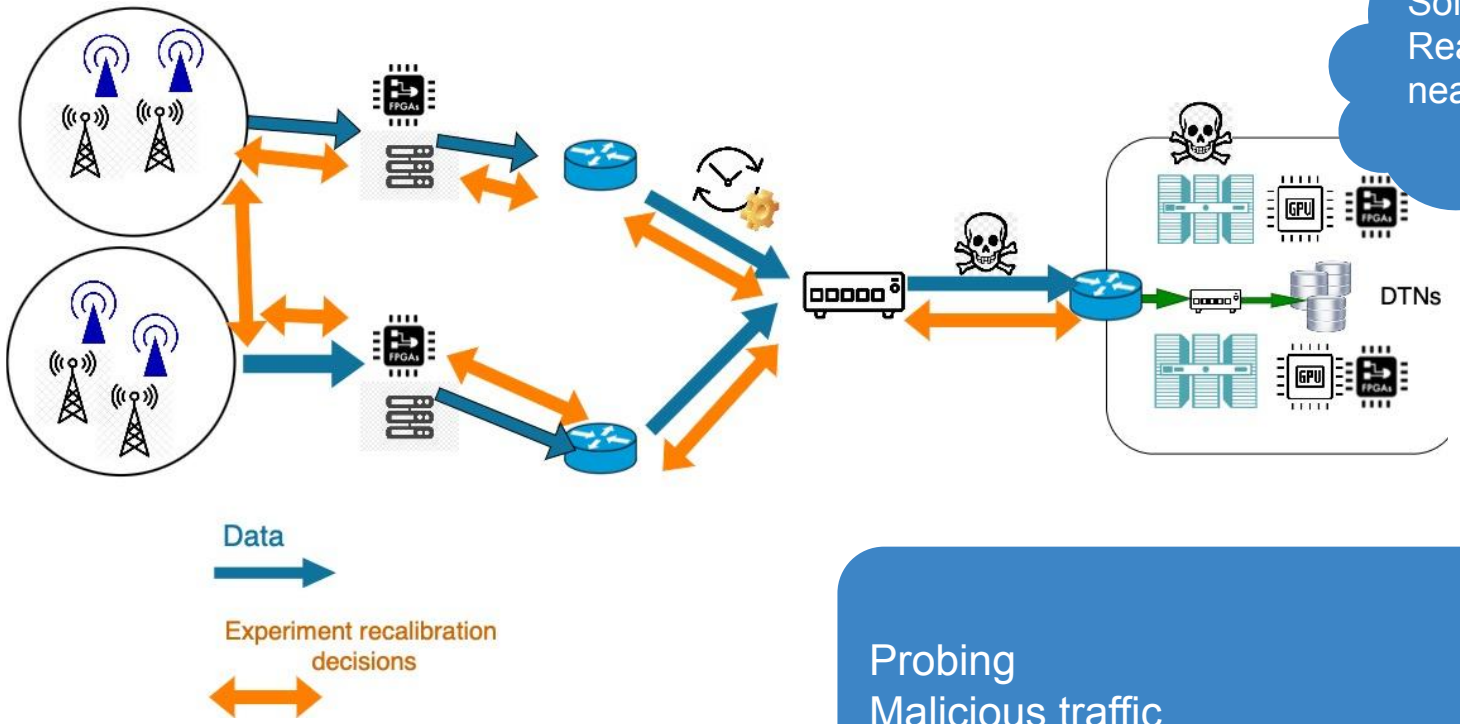
Data transfers experience performance degradation challenges



Solutions in Real-time or near real-time

Packet retransmissions
Throughput degradation
Long completion times

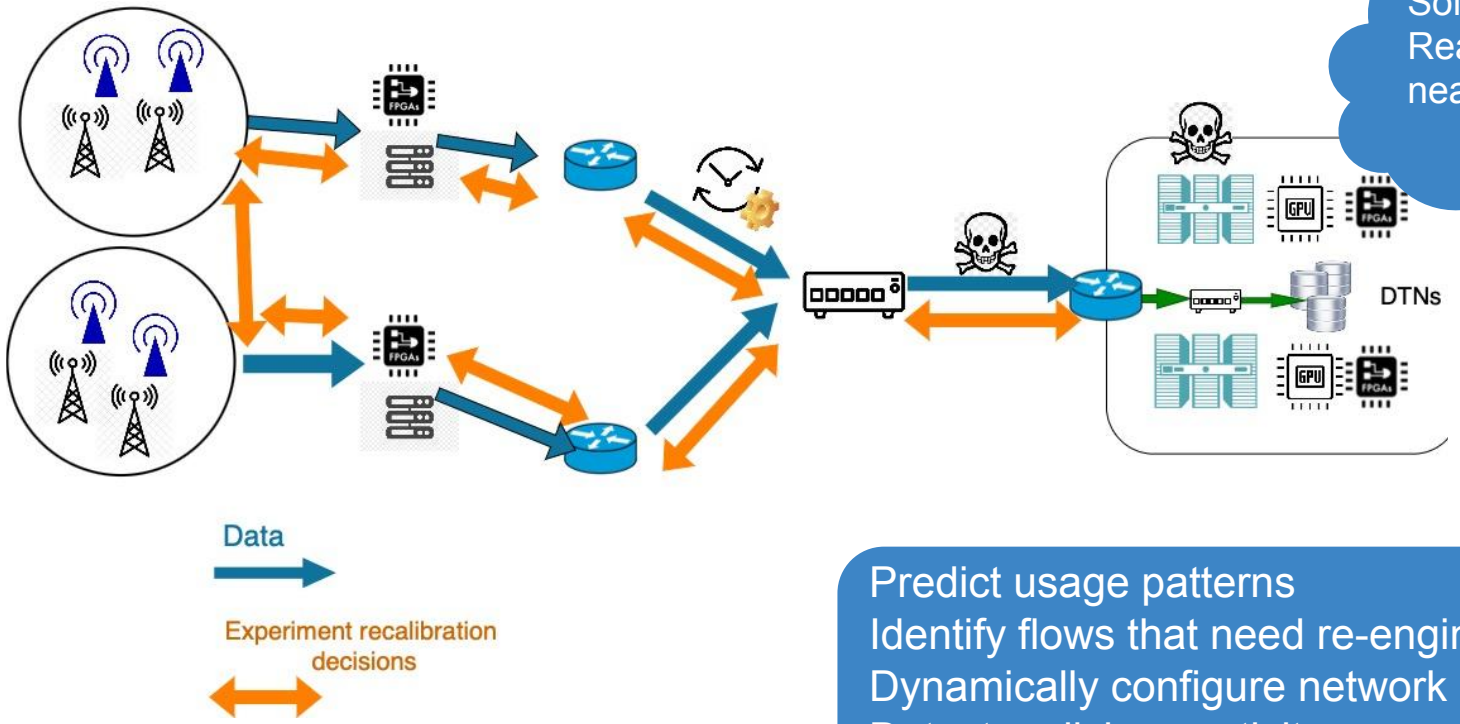
Data transfers experience security events at different levels



Solutions in Real-time or near real-time

Probing
Malicious traffic
Data exfiltration attempts

Anomaly detection and performance prediction is critical for seamless data movement

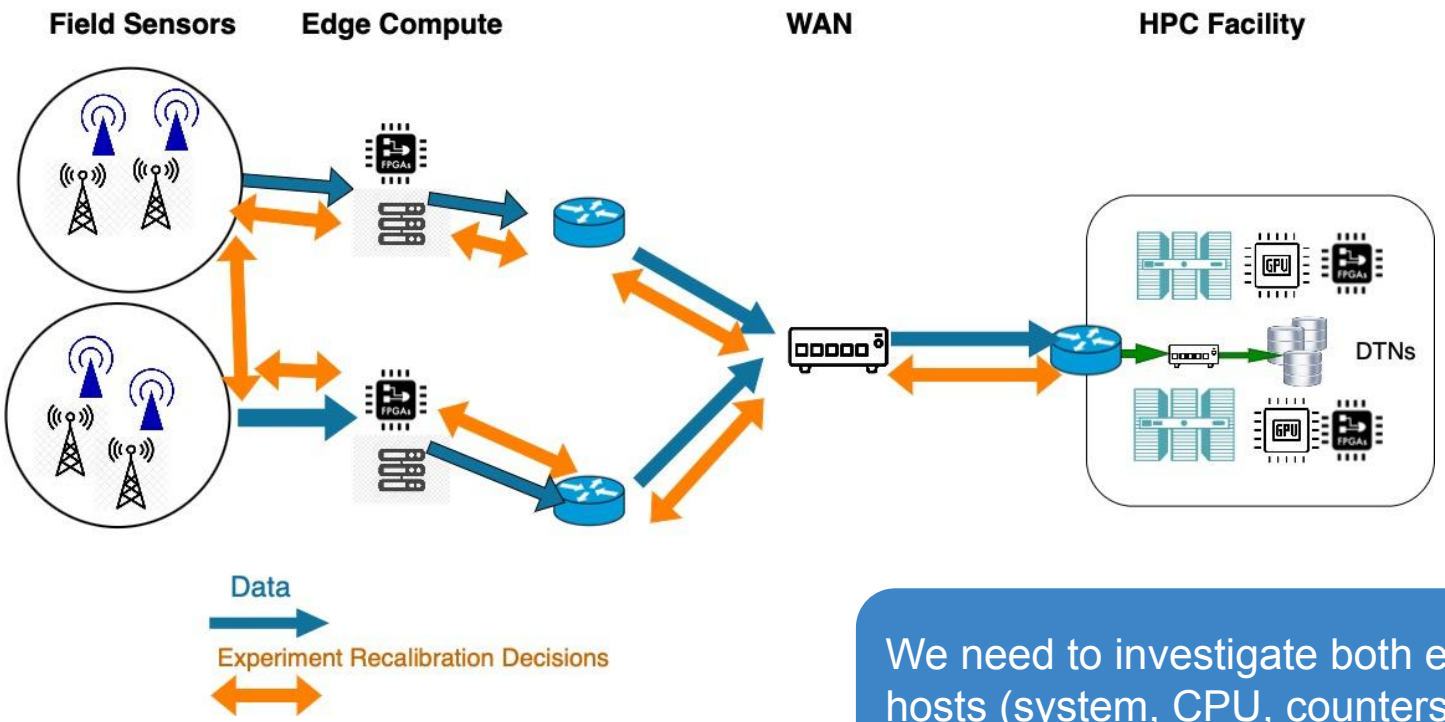


Predict usage patterns
Identify flows that need re-engineering
Dynamically configure network paths
Detect malicious activity

Existing anomaly detection solutions perform poorly on scientific network traffic

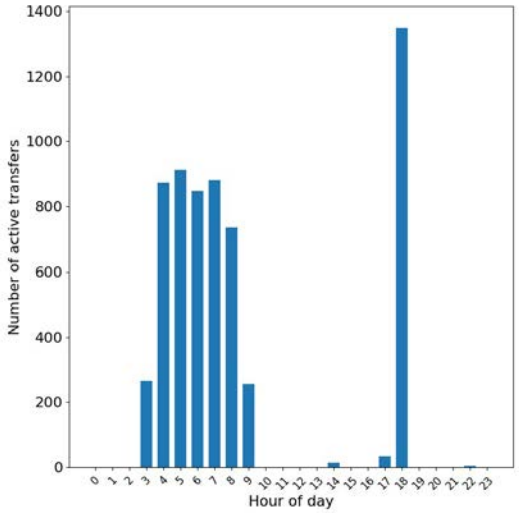
- **High variability of data transfers**
- **Multi-dimensional feature dependencies**
- Limited visibility
  - Sampled data (SNMP)
    - Per interface
    - Aggregate
  - End-host data (tstat)
    - Statistics summary
  - Flow-based (sFlow, Netflow)
    - approximate
- Offline solutions
  - Post mortem analysis

High variability: data transfer performance depends both on network conditions and end-host system performance

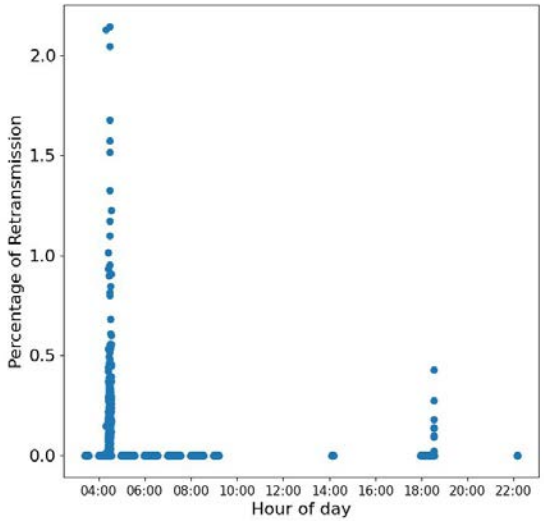


We need to investigate both end hosts (system, CPU, counters) and network state (packets, flows)

Data transfer variability: flow retransmissions increase during busy time windows



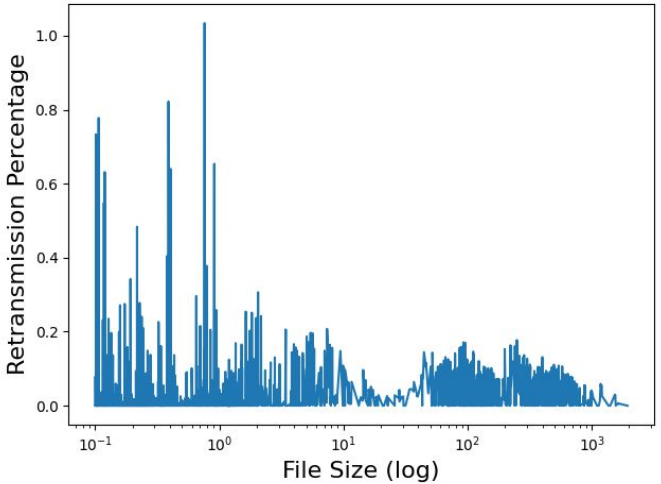
Number of active incoming transfers over 24h period on a DTN



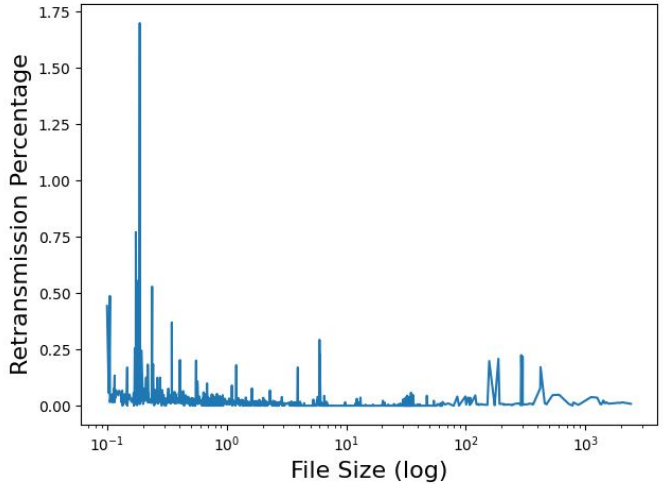
Percentage of retransmissions for incoming transfers over 24h period on a DTN

Retransmissions spike and then even out suggests of NIC load balancing (still under investigation)

Data transfer variability: small flows experience larger retransmission percentages



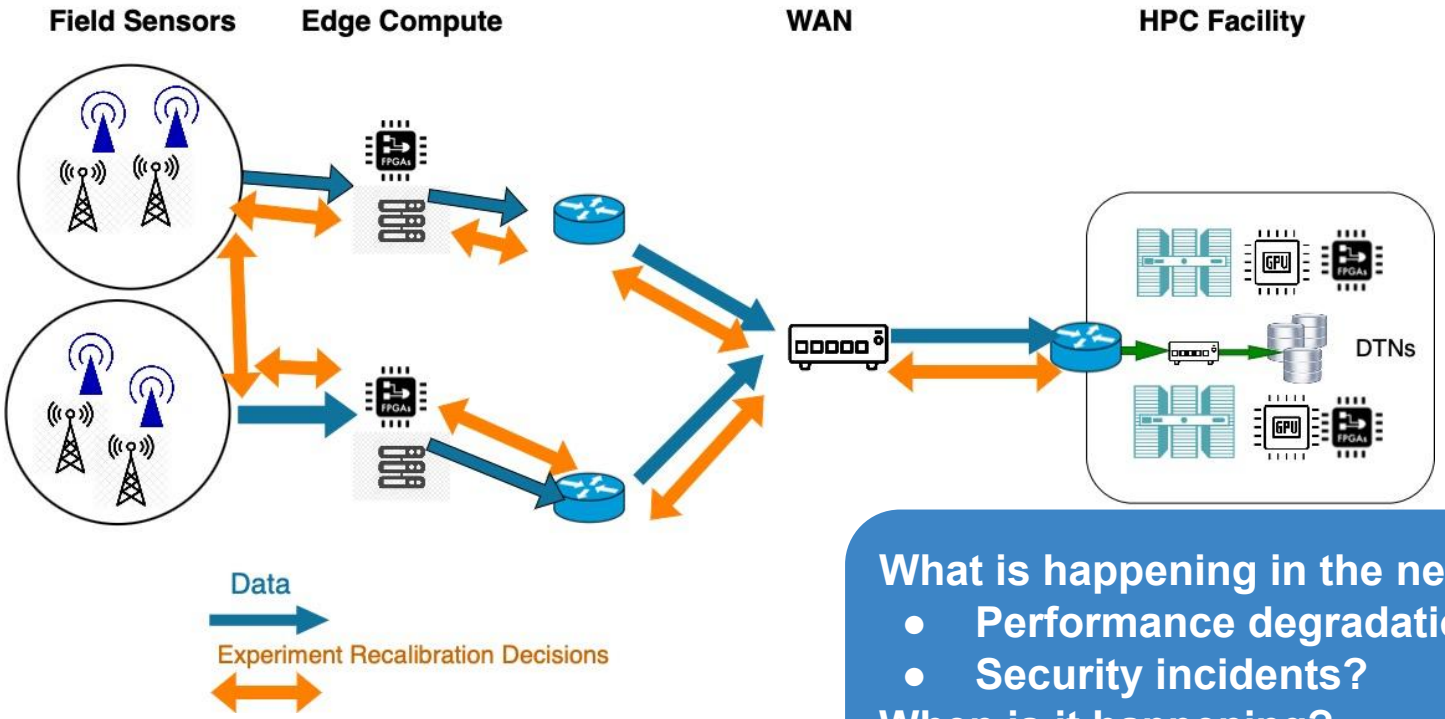
Incoming flow sizes on a DTN



Outgoing flow sizes on a DTN

Larger flows are more stable (retransmission <0.2%)
Do we care about small flows?
Yes! For DTNs that users can choose approx. 90% of transfers are below 100GB

Data transfer performance depends both on network conditions and end-host system performance



What is happening in the network?
- Performance degradation events?
- Security incidents?

When is it happening?

Why?

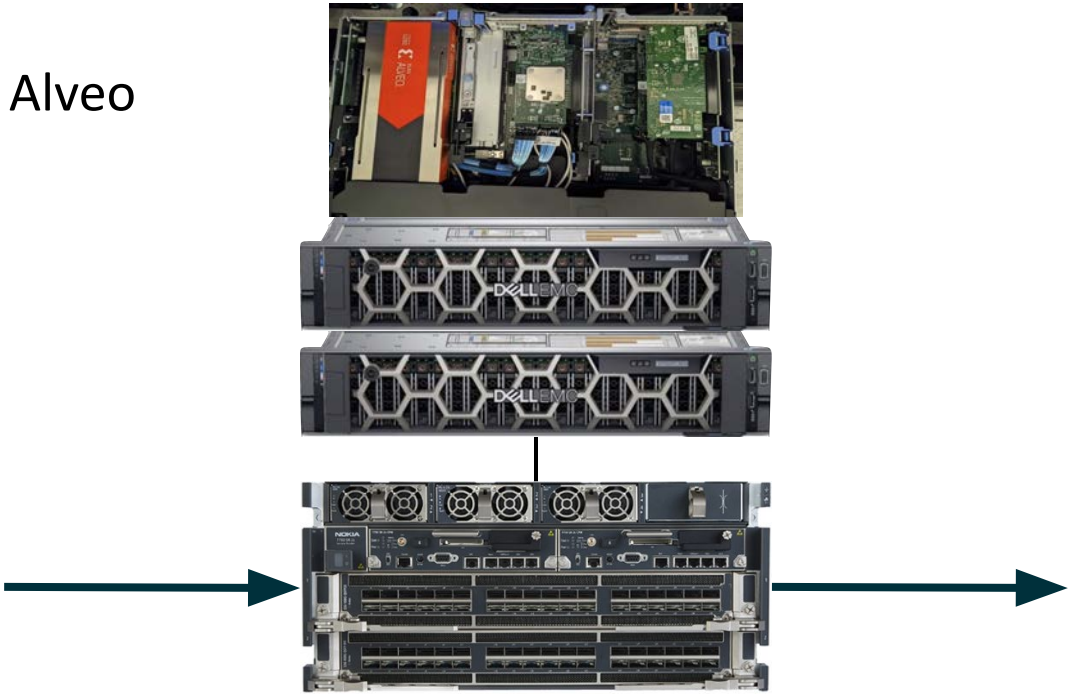
Data sources: ESnet6 High-Touch and Q-Factor

ESnet6 High-Touch: architecture and design

FPGA (Xilinx Alveo U280)

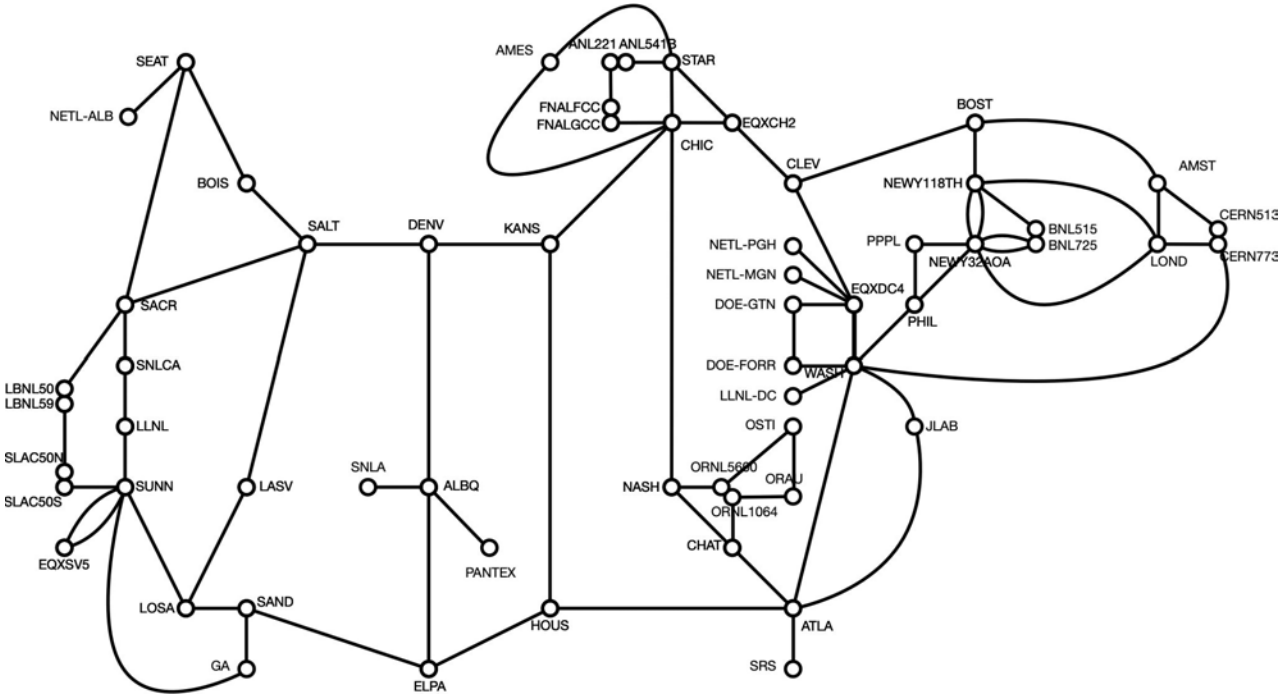
HT Servers

Router



ESnet6 High-Touch: platform and deployment plan

42 deployment locations, each location will have 2 High-Touch servers



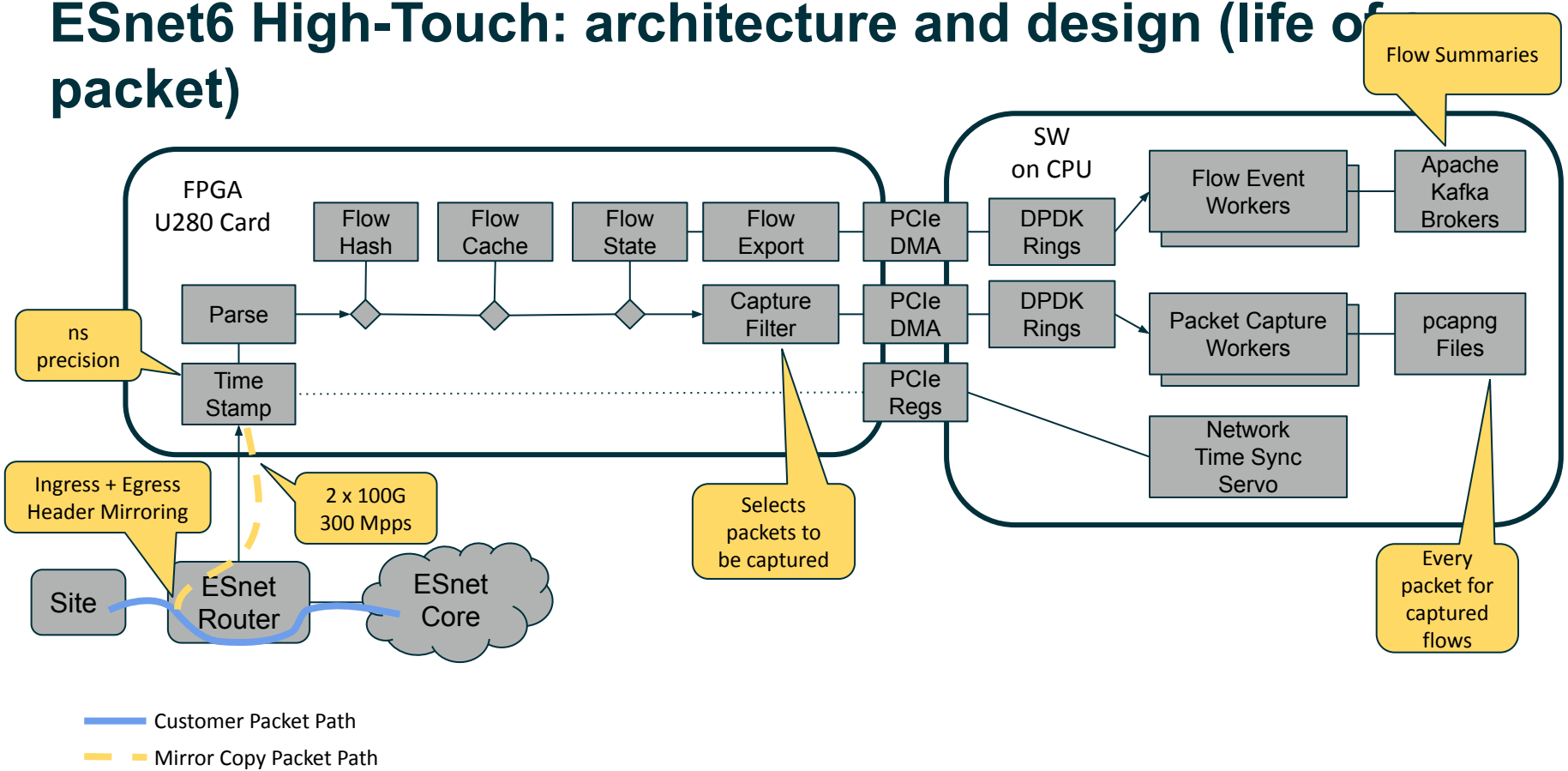
High-Touch Server (x2)

Xilinx Alevo U280 FPGA:

- 2x100G port
- 1.2M logic cells
- 32GB DDR4
- 8GB HBM2 memory (3.2 Tbps I/O)

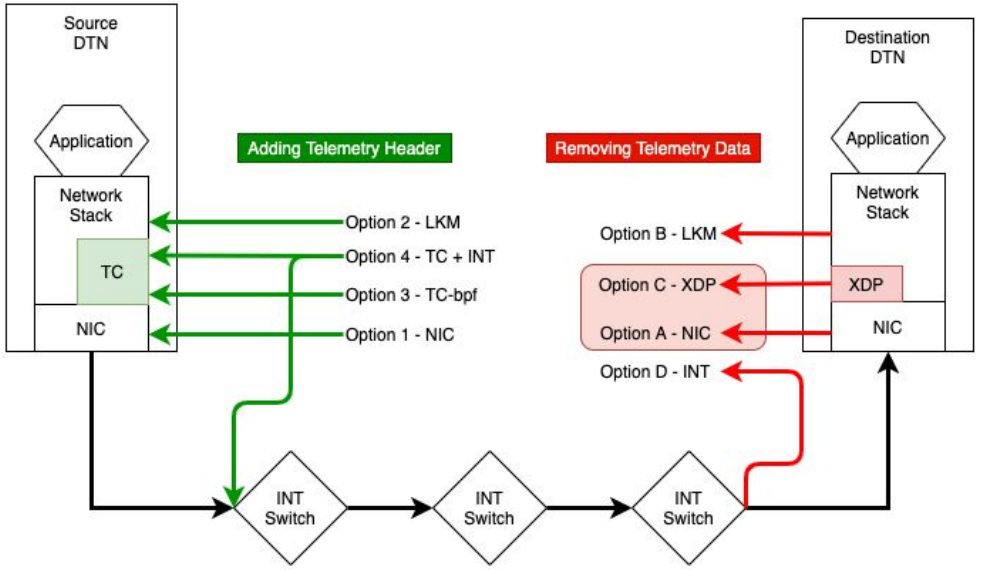
Compute node

ESnet6 High-Touch: architecture and design (life of packet)

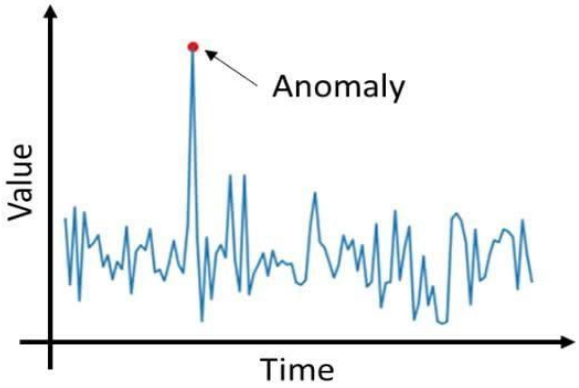


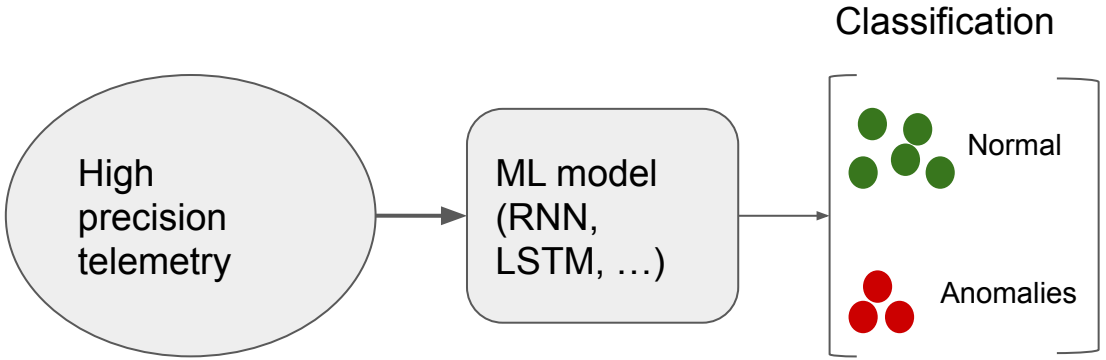
Q-Factor for host-based INT and tunning

- Q-Factor will extend the hosts' capabilities by:
  - Adding support for INT to hosts
  - Supporting a Telemetry Agent to tune the host
- INT at hosts will operate via two approaches:
  - If a programmable NIC is available, INT will be done at the NIC
  - Otherwise, eBPF/XDP/TC will be used before the Linux TCP/IP stack for performance
- Hosts' applications will NOT need to be changed:
  - Q-Factor will operate before the Linux TCP/IP stack
  - Completely hidden from the upper layers



Solution: detecting network anomalies using high precision telemetry data and machine learning



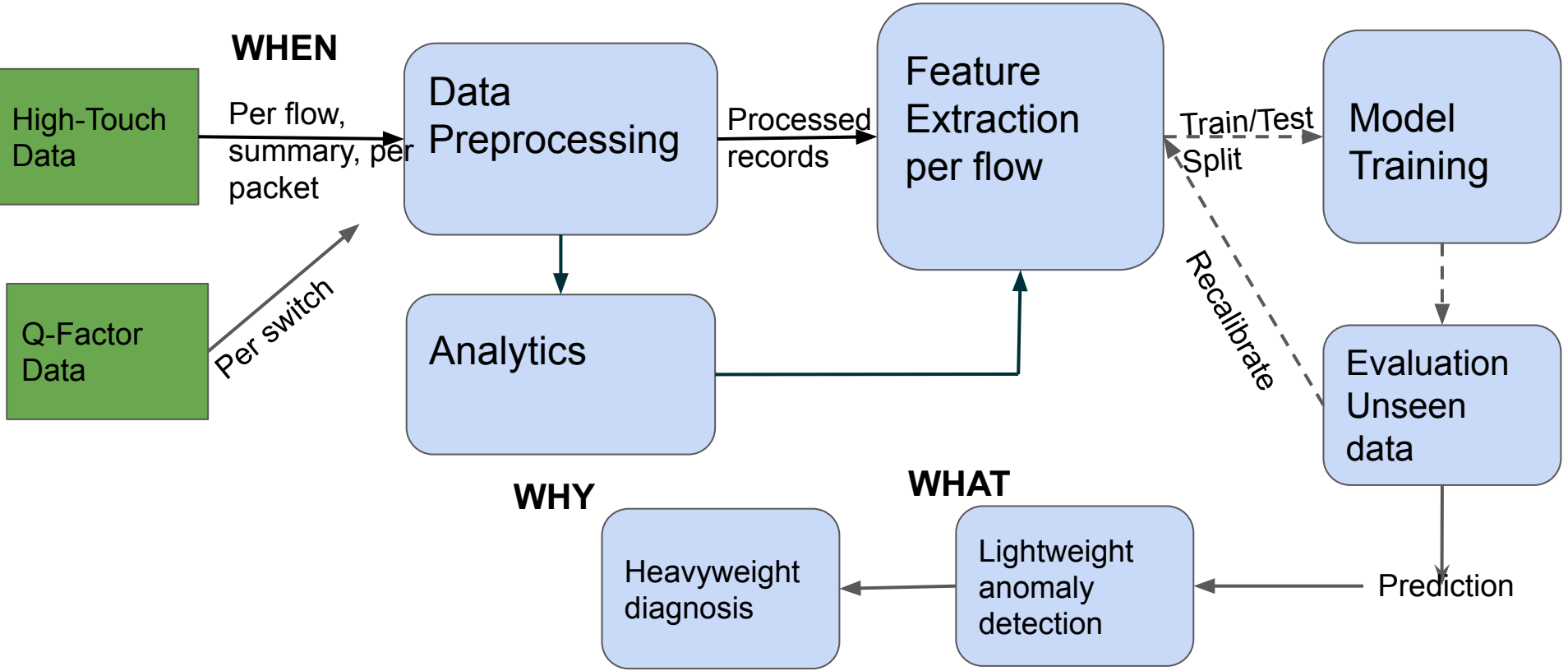


Why use high precision telemetry:
Unprecedented visibility

- Flow level
- Packet level

Near real time network microscope enables fast traffic reengineering

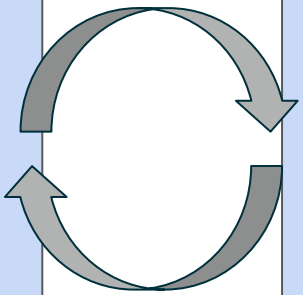
Detecting network anomalies using high precision telemetry data and machine learning



Data preprocessing and analytics using high-touch data

Data Preprocessing:

- Flow grouping based on 5 tuple (sIP, dIP,sp,dp,proto)
- Incoming/outgoing traffic filtering
- Additional field computation:
  - Duration
  - Rate, etc
- Time window binning

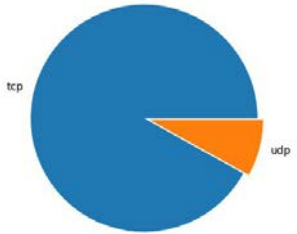


Analytics:

- Discover subnets of interest:
  - Volume
  - Duration
  - Rate
  - Number of connections
- Discover time windows of interest:
  - Busy
- Specific flow filters:
  - Single packet
- Subnet/org correlation

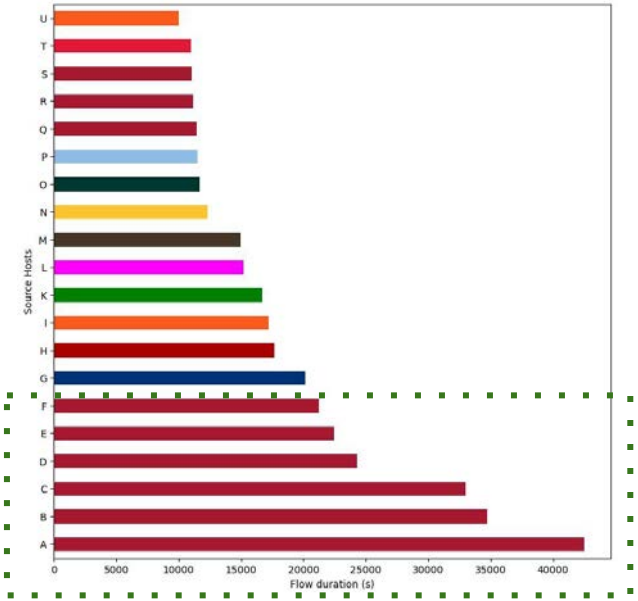
Data analytics example: longest and heaviest flows originate from hosts in a particular subnet

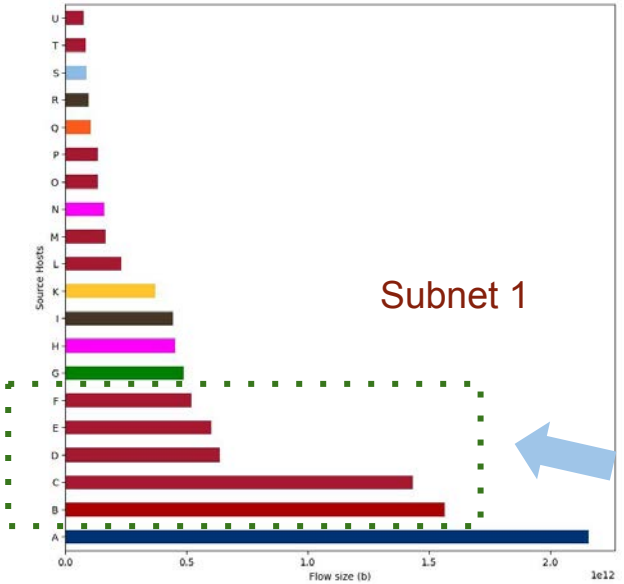
4-day dataset from high-touch instance monitoring NERSC border router



TCP dominant (95% of overall traffic)

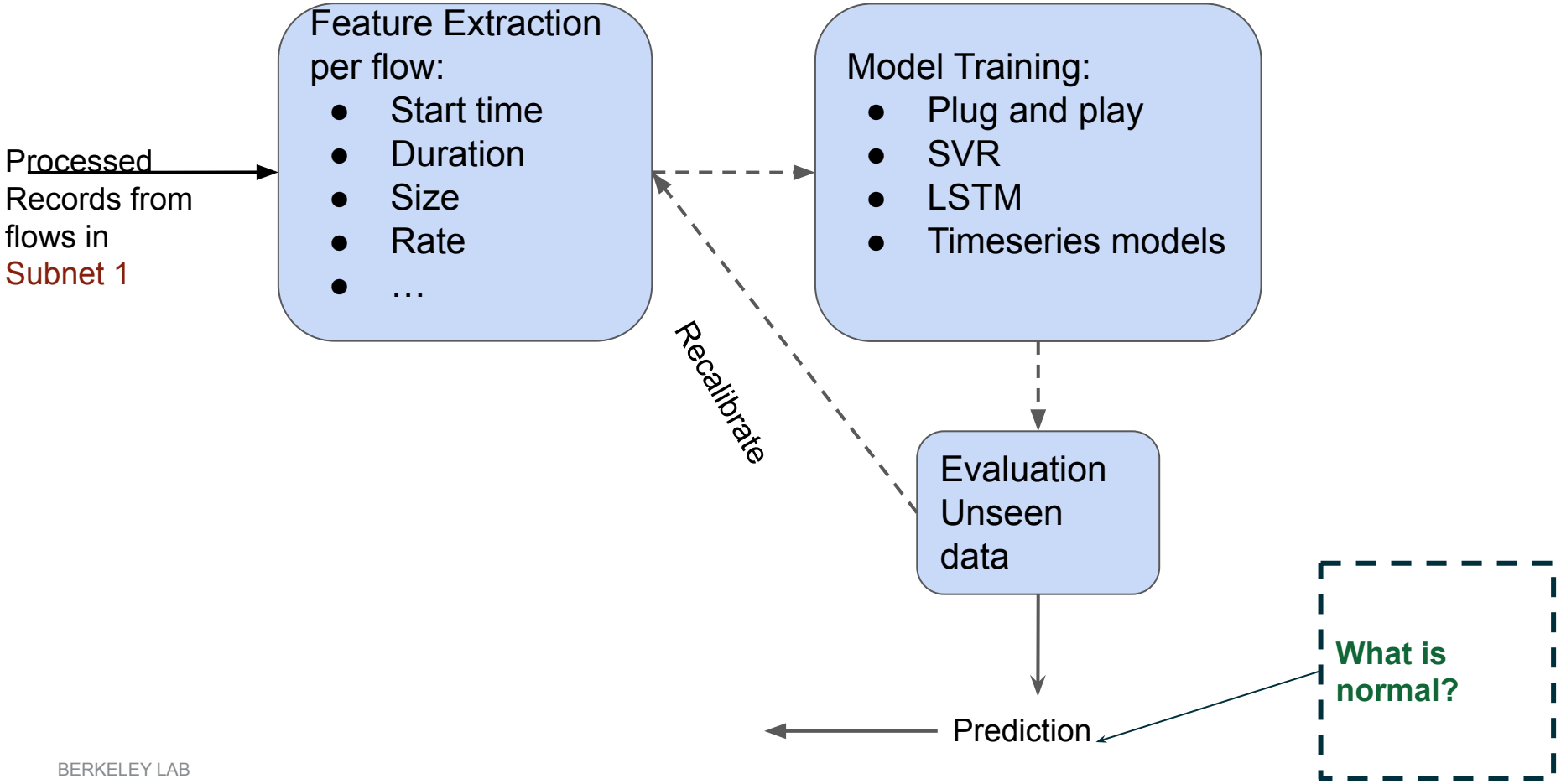
Subnet 1



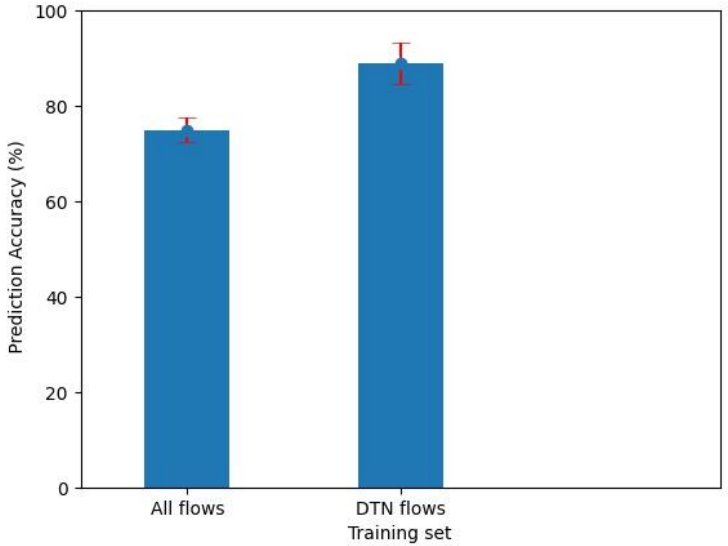


Subnet 1

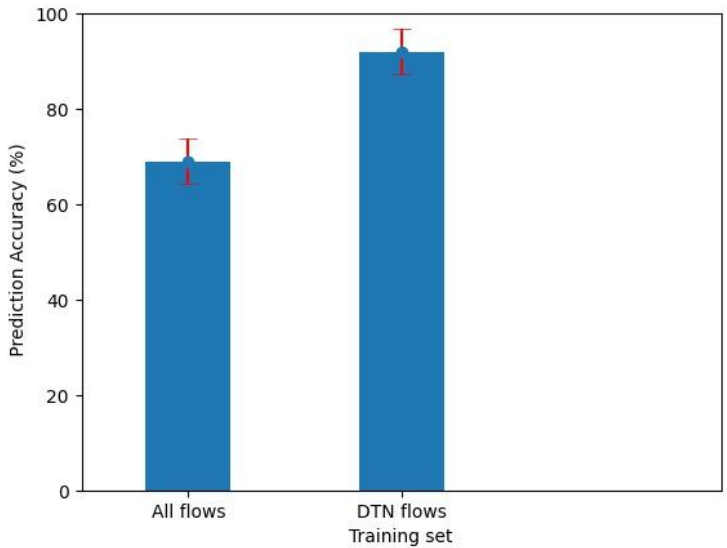
What do we want to predict and why?



Early results: flow size and duration prediction accuracy



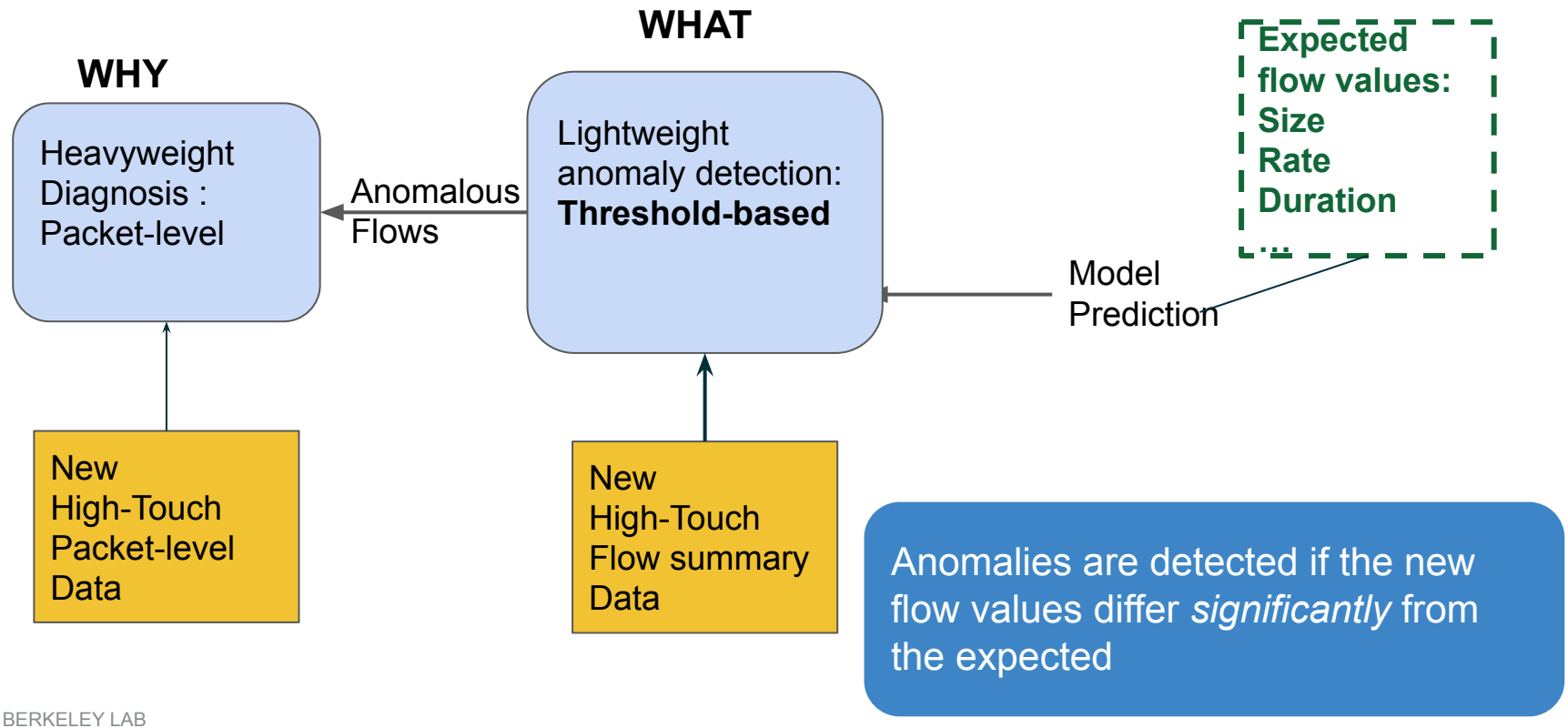
Flow **duration** prediction in 4 day window of NERSC traffic



Flow **size** prediction in 4 day window of NERSC traffic

Training subnet specific SVR models for predicting flow duration and flow size yields improved accuracy ()

Detecting network anomalies based on deviation from “normal”



Future Work: ml improvements

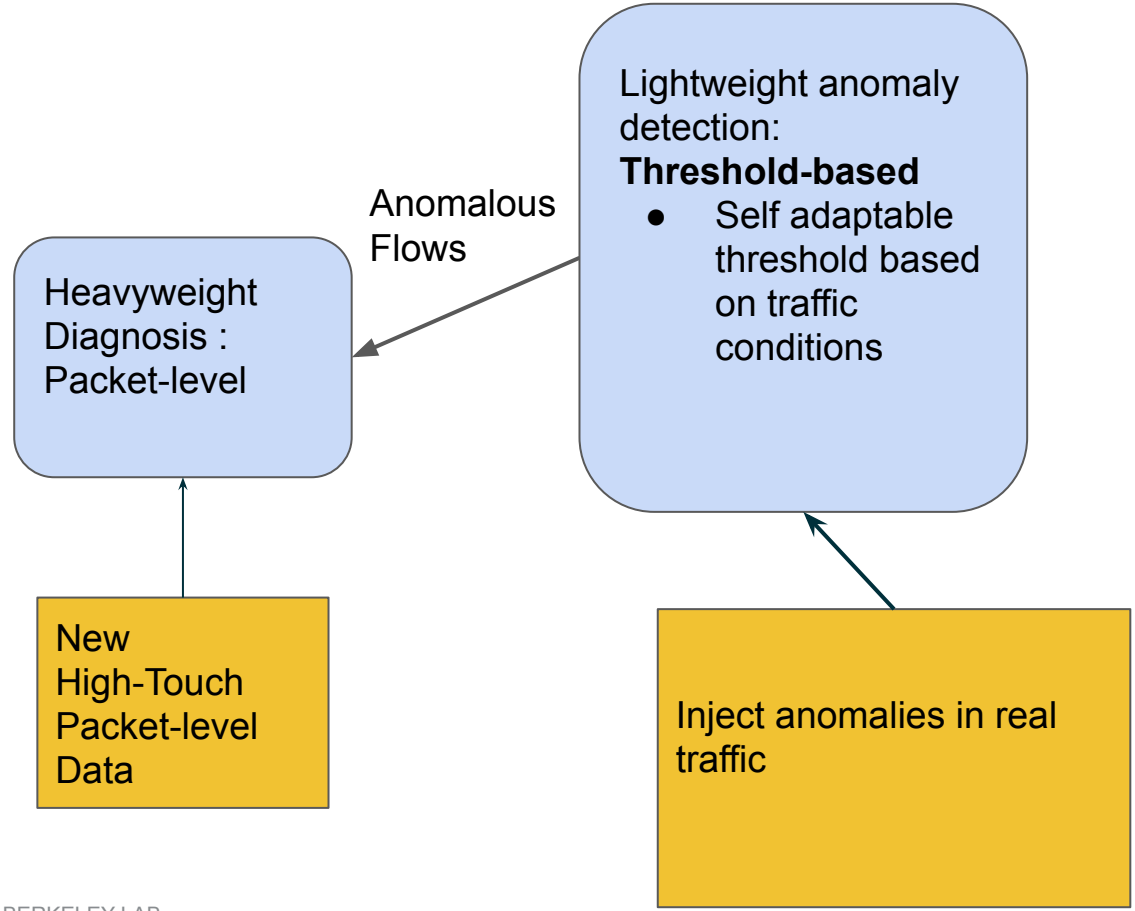
Model Training:

- Experiment with RNN models
- Predict additional features
  - Rate
  - Loss
- Other datasets
  - Org to org traffic
  - application-specific

Model Tunning:

- Evaluate retraining strategies
  - Fixed time
  - Accuracy threshold
- Online/offline

Future work: detection



Future Work- Detection

Model Training:

- Experiment with RNN models
- Predict additional features
  - Rate
  - Loss
- Other datasets
  - Org to org traffic
  - application-specific

Model Tunning:

- Evaluate retraining strategies
  - Fixed time
  - Accuracy threshold
- Online/offline

Conclusion

- Fast and reliable movement of data across facilities is instrumental in modern workflows and scientific discovery
- Data transfer performance is a multi-factorial issue
- Data transfers experience anomalies that remain undetected/unmitigated due to lack of visibility
- We can use INT data to train AI/ML models that successfully detect anomalies in near real-time and enable traffic reengineering

Acknowledgement

Special thanks to Bruce Mah, Yatish Kumar, Stacey Sheldon, Lavanya Ramakrishnan, Nick Wright, Taylor Groves

Questions?