USENIX ATC '25

TGW: Operating an Efficient and Resilient Cloud Gateway at Scale

Yifan Yang, Lin He, Jiasheng Zhou, Xiaoyi Shi, Yichi Xu, Shicheng Wang, Jinlong E, Ying Liu, Junwei Zhang, Zhuang Yuan, Hengyang Xu



清華大學
Tsinghua University

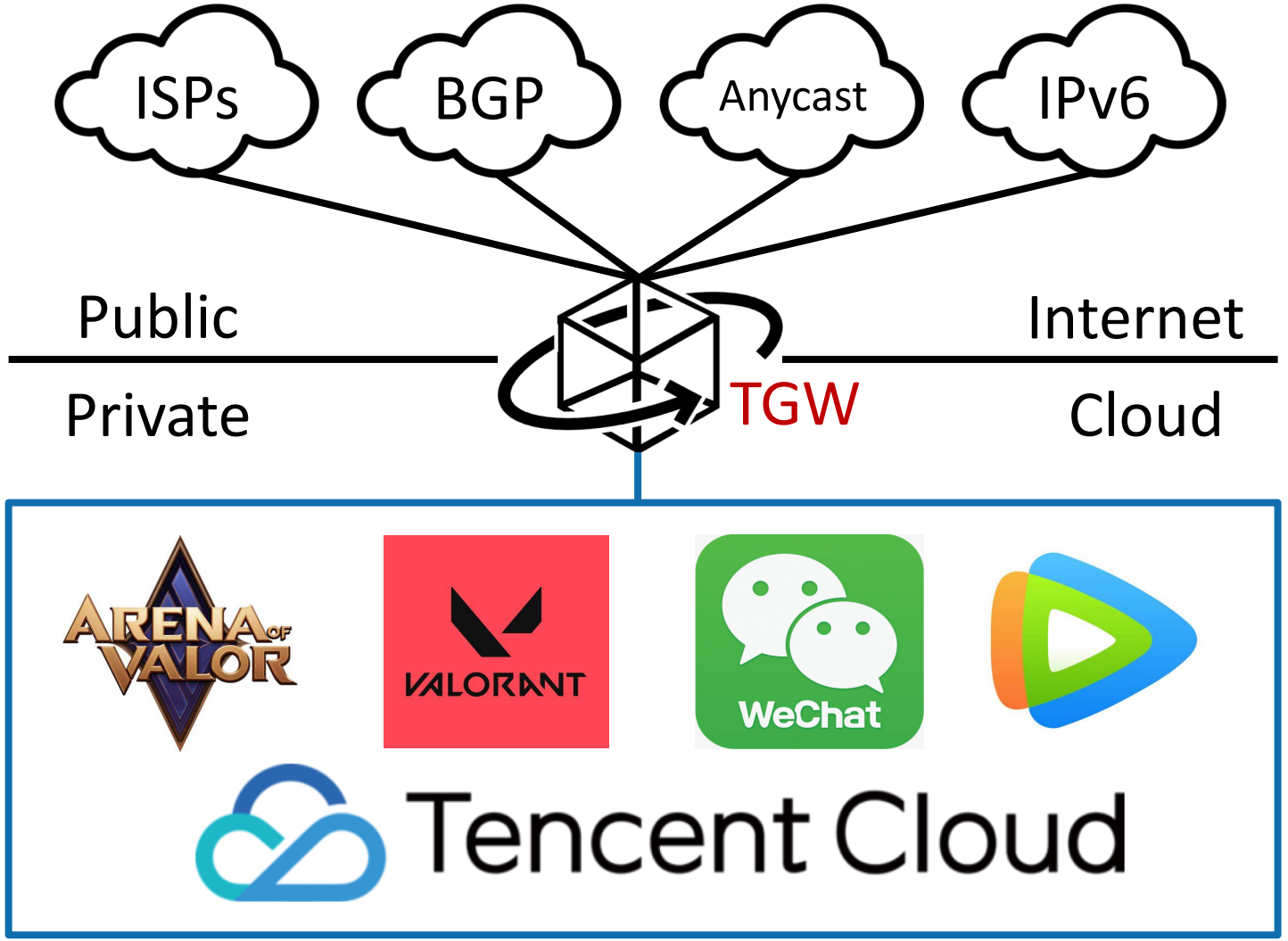


中國人民大學
RENMIN UNIVERSITY OF CHINA

Tencent

What is TGW?

- Tencent GateWay (TGW) is the "bridgehead" of Tencent Cloud
- TGW carries both the north-south and east-west traffic



TGW's Evolution

- Tencent Cloud's service number and gateway traffic have grown at an average rate of over 100% and 60% per year, respectively

Start

2011-2012
Public backend servers
Lacking IP addresses

Scale: service

2013-2015
Steady development
Service growth
Simple migration mechanism
Low failure resilience

New Arch: DPDK

2016-2017
DPDK
40G NIC

Scale: deployment

2018-2019
Large-scale deployment
Efficiency improvement
IPv6 and Anycast support
100G NIC

New Arch: Regional EIP

2020-now
Regional EIP
Dynamic migration
100% availability
200G NIC

- DPDK-based TGW has been operating for 8 years
- Today's TGW is deployed in more than 30 regions
- TGW scales to hundreds of clusters and thousands of machines per region

Why DPDK?

Traditional gateway: fixed-function hardware switches

- ✓ Cheap and high performance
- ✗ Very inflexible

Software: XDP (eXpress Data Path)

- ✓ Simple for development and operation
- ✗ Relatively poor performance

- Deployed at the edge host of our cloud CDN
- Achieve load balancing and anti-DDoS functions at 10 Mpps per host, utilizing just 5% CPU

Software: VPP (Vector Packet Processor)

- ✓ Rich routing features and good performance
- ✗ Introduce more management cost

Why DPDK?

Hardware: DPUs, FPGAs, and smartNICs

- ✓ Flexible programmability and CPU offloading
- ✗ Low cost performance (Gbps/$) and high power consumption

Hardware: programmable switches

- ✓ Flexible programmability
- ✓ Provide the same throughput with minimum cost and power consumption

- Deployed at our testbed
- Each switch manages 50K IP addresses and 200K rules
- Each switch provides up to 2 Tbps throughput
  - ✗ Highly resource-constrained
  - ✗ Unclear outlook

Motivation

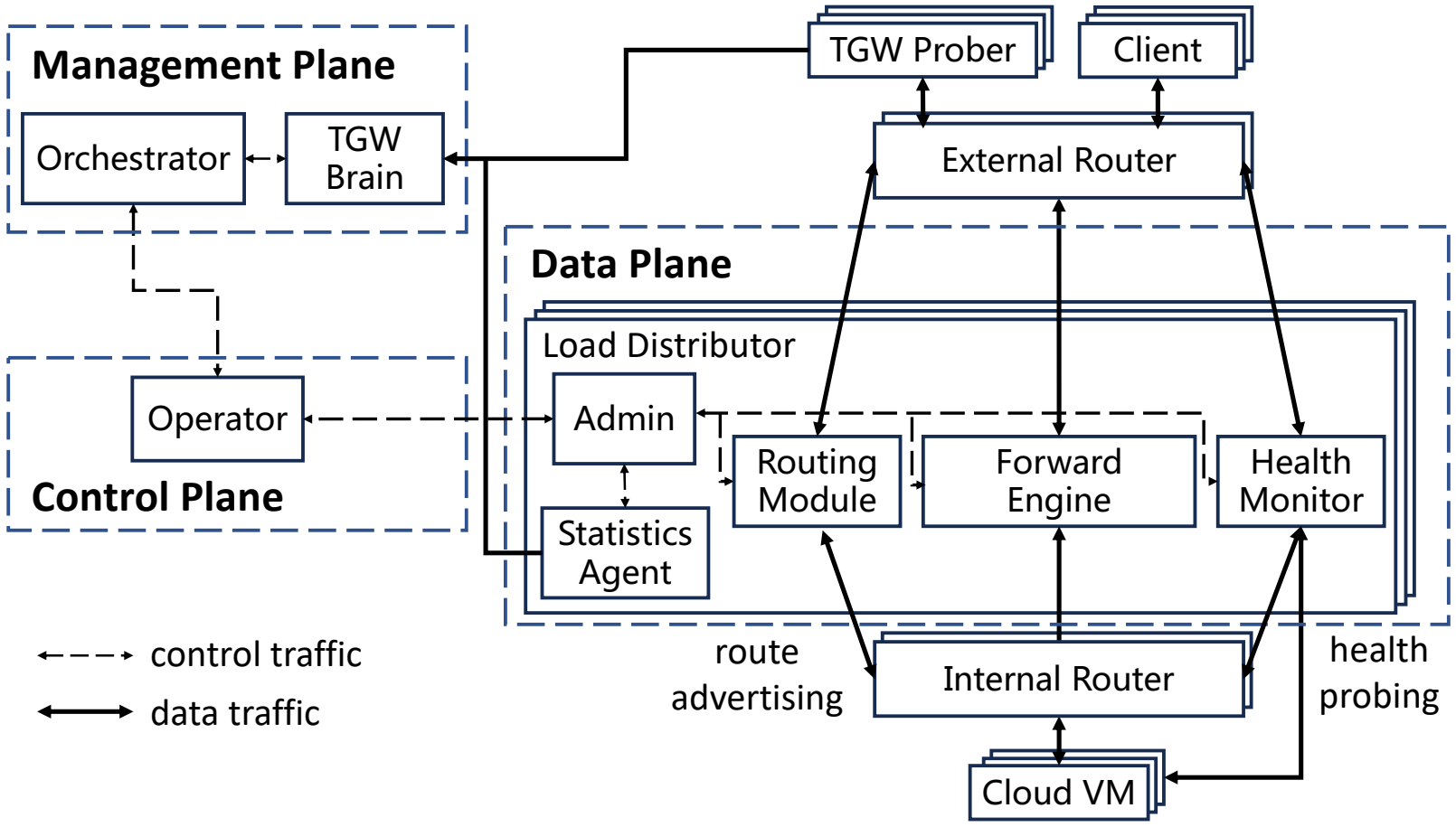
Key design principle of TGW: low latency, low jitter, and low packet loss

- TGW is built for Tencent Cloud services
- Tencent Cloud's "killer" services are online gaming and live streaming
- Online gaming and live streaming customers strongly require a low-latency, low-jitter, and low-packet-loss network

Motivations for building TGW

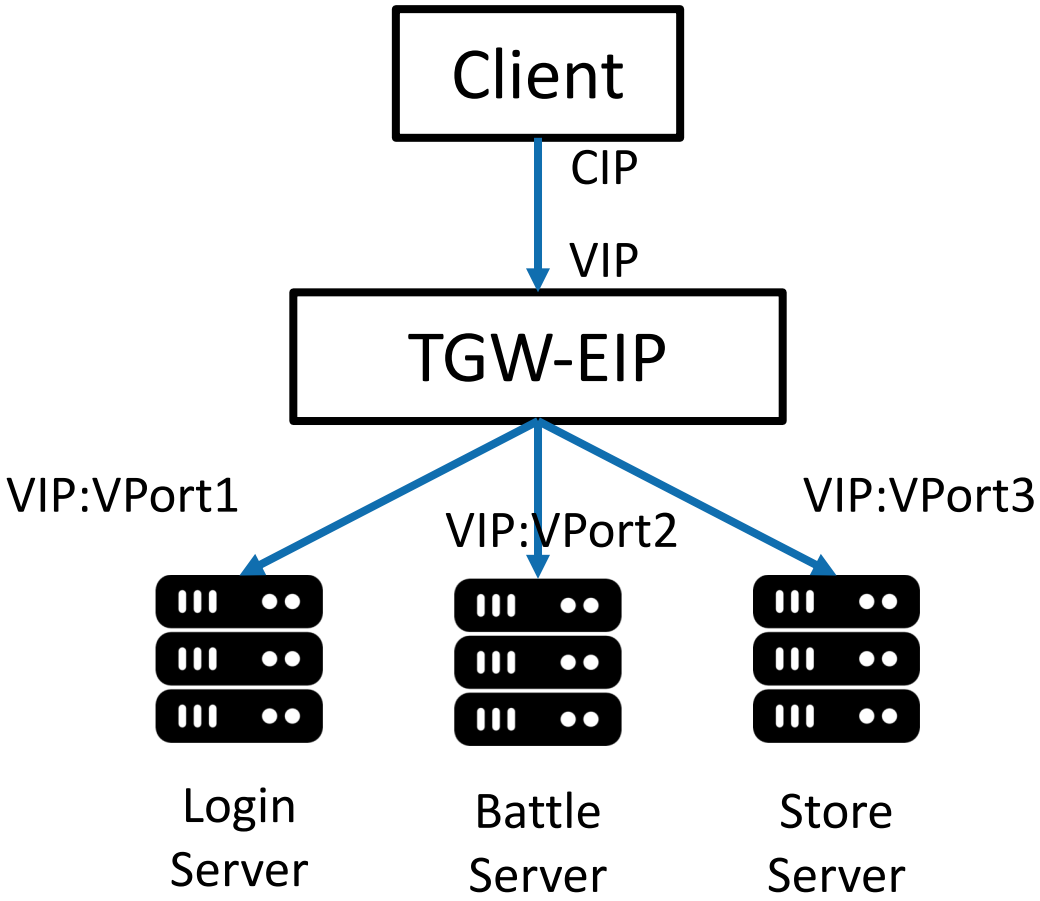
- Efficient packet processing
- Scalable state management
- Fast failure recovery

Architecture

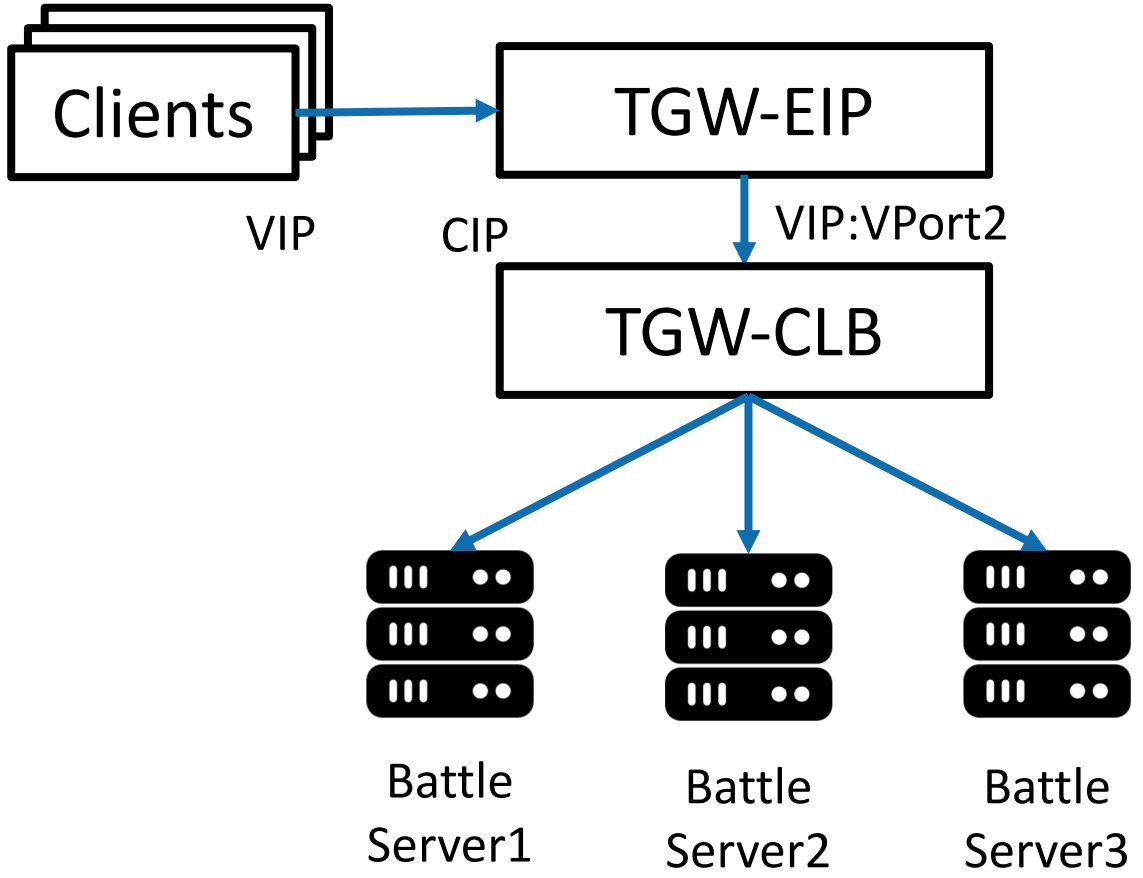


- Routers
- Load distributors (LD)
  - Routing module
  - Forward engine
  - Health Monitor
  - Statistics Agent
  - Admin
- TGW probers and TGW brains
- Operator
- Orchestrator

Workflow



Elastic IP Access

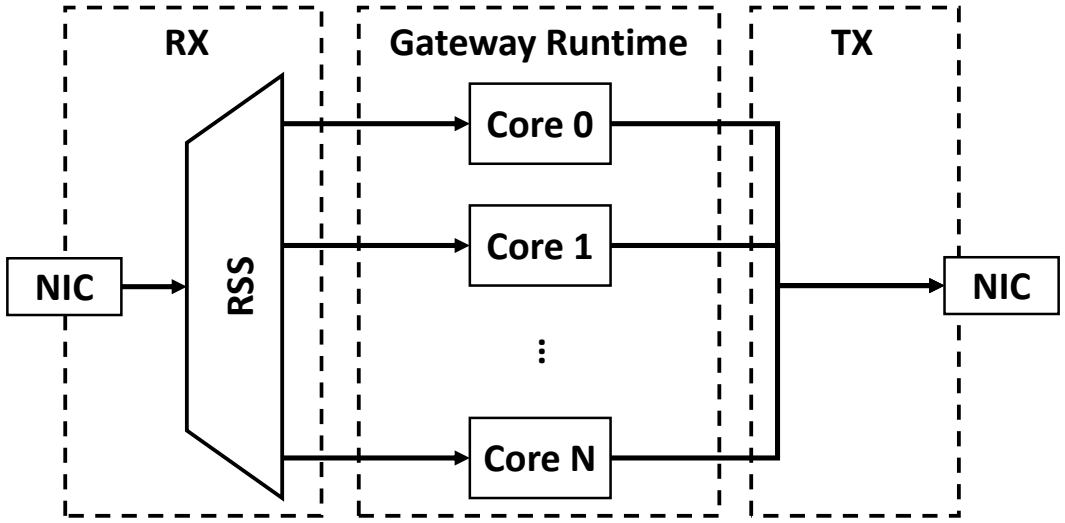


Cloud-scale Load Balancing

Design Details

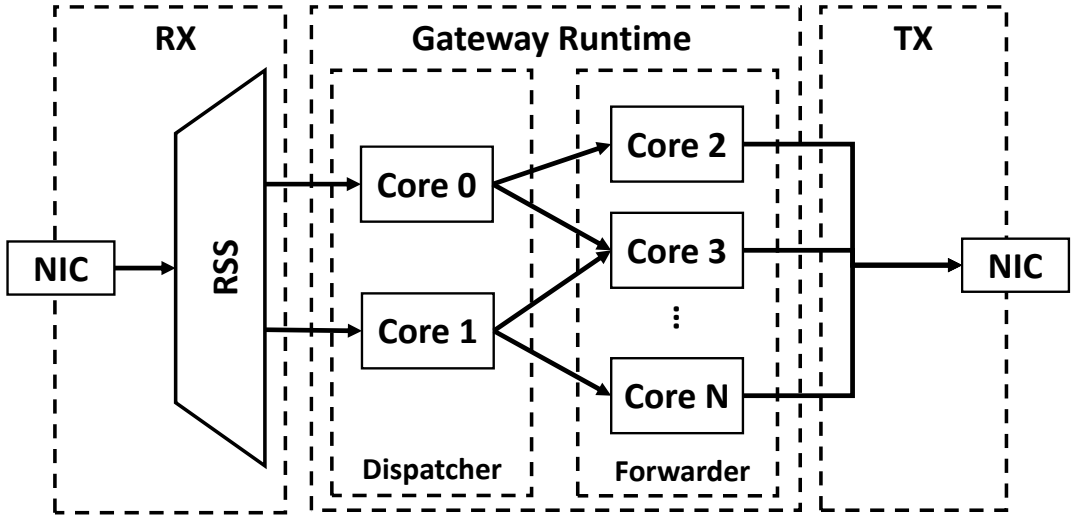
- Packet Forwarding
- State Migration
- Failure Resilience
- Failure Detection and Localization

Packet Forwarding



TGW-EIP: RTC model

- Stateless
- Minimize latency
- easy to scale



TGW-CLB: pipeline model

- Stateful
- Service-specific identifier dispatching
  - Prevent delay jitter in gaming services
  - QUIC-aware load balancing

Live Migration

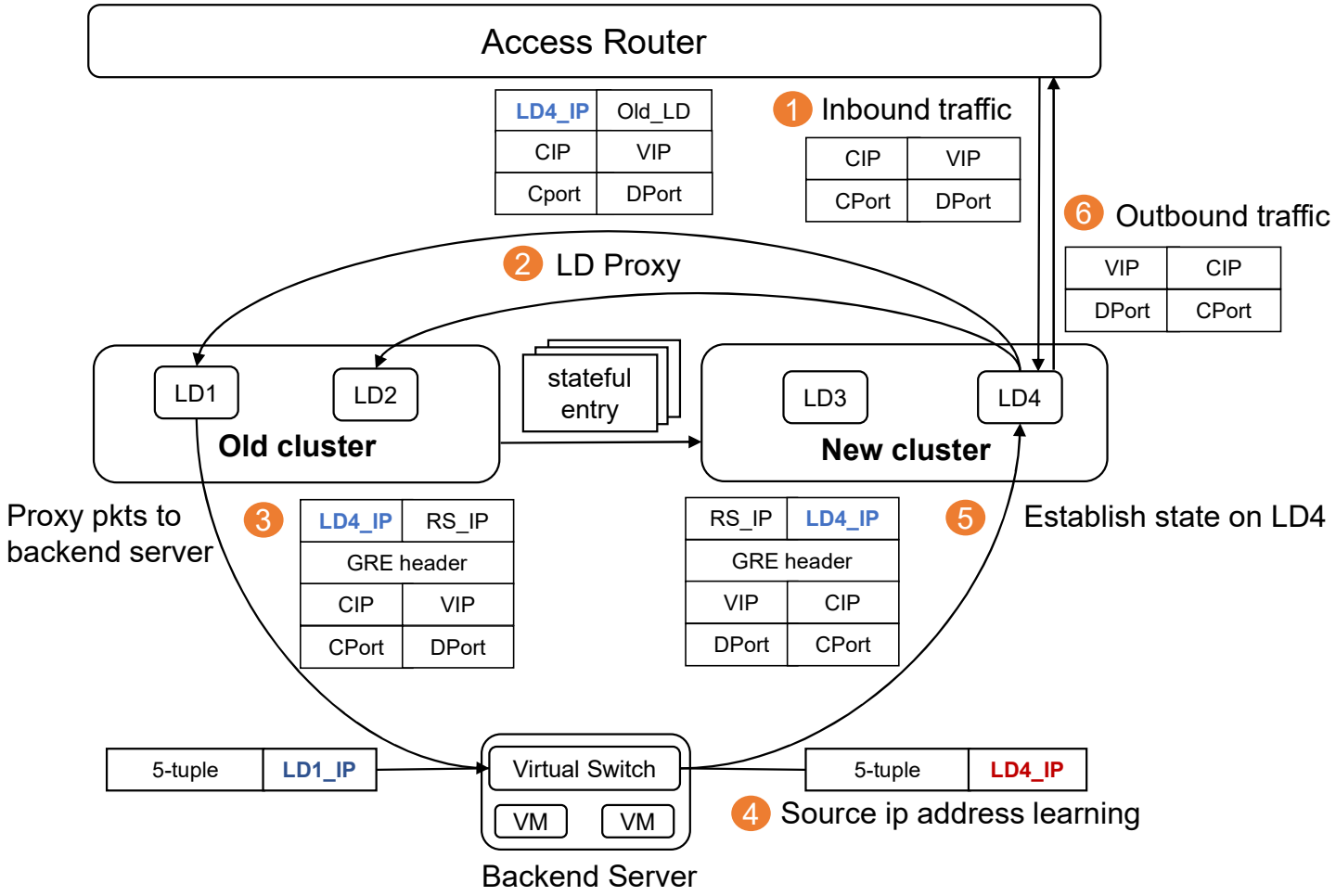
Tricky "80/20" rule: 80% of the traffic is carried by 20% of the forwarding clusters

Solution: Dynamically migrate the traffic from overloaded clusters to more capable clusters

Challenge: How to keep the service uninterrupted during migration?

Solution: Design a lossless, unperceived, and fast-converging live migration mechanism

LD proxy

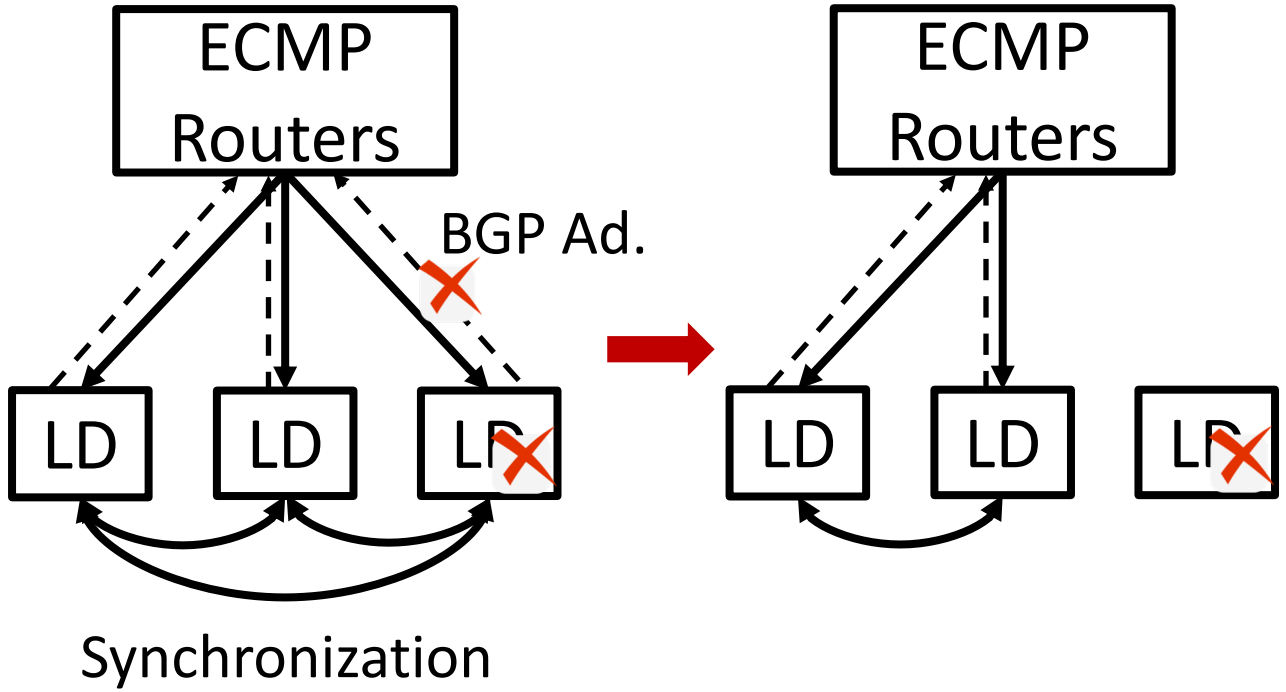


Problem in consistency:
Not migrated states or new states are only on the old cluster, but subsequent packets are forwarded to the new cluster

Solution: LD proxy and Source IP address learning

Intra-AZ

We employ different **fault tolerance models** at different levels

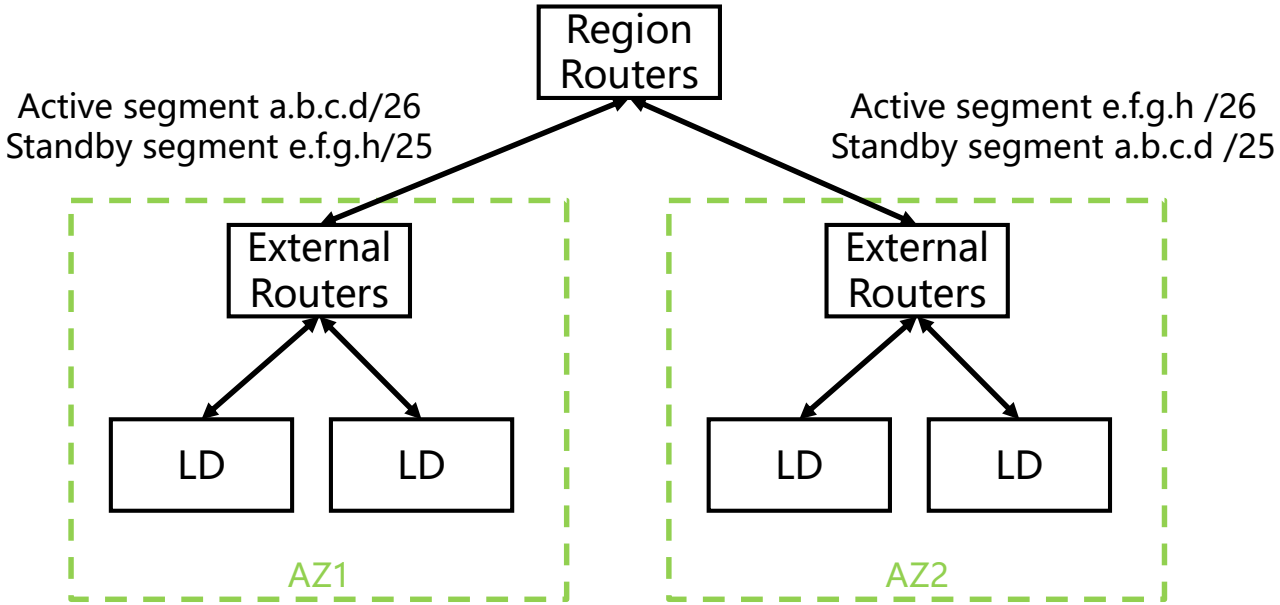


Intra-AZ Active-Active model

- Lower synchronization costs
- Millisecond-level failure detection and second-level failover delay

Inter-AZ

We employ different **fault tolerance models** at different levels

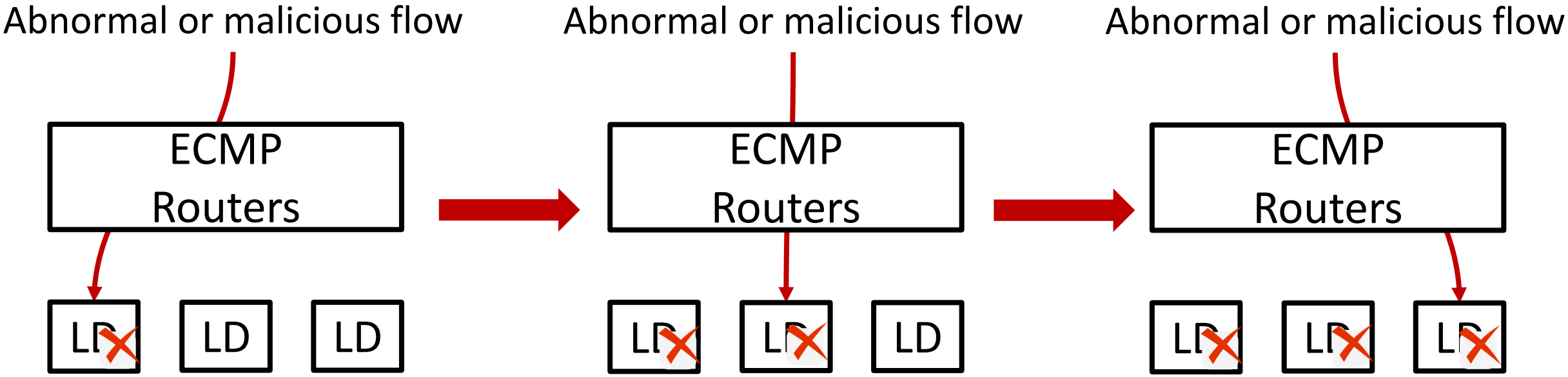


Inter-AZ Active-Standby model

- Prevent jitter caused by equivalent AZs
- Minimize synchronization cost

Decentralized Migration

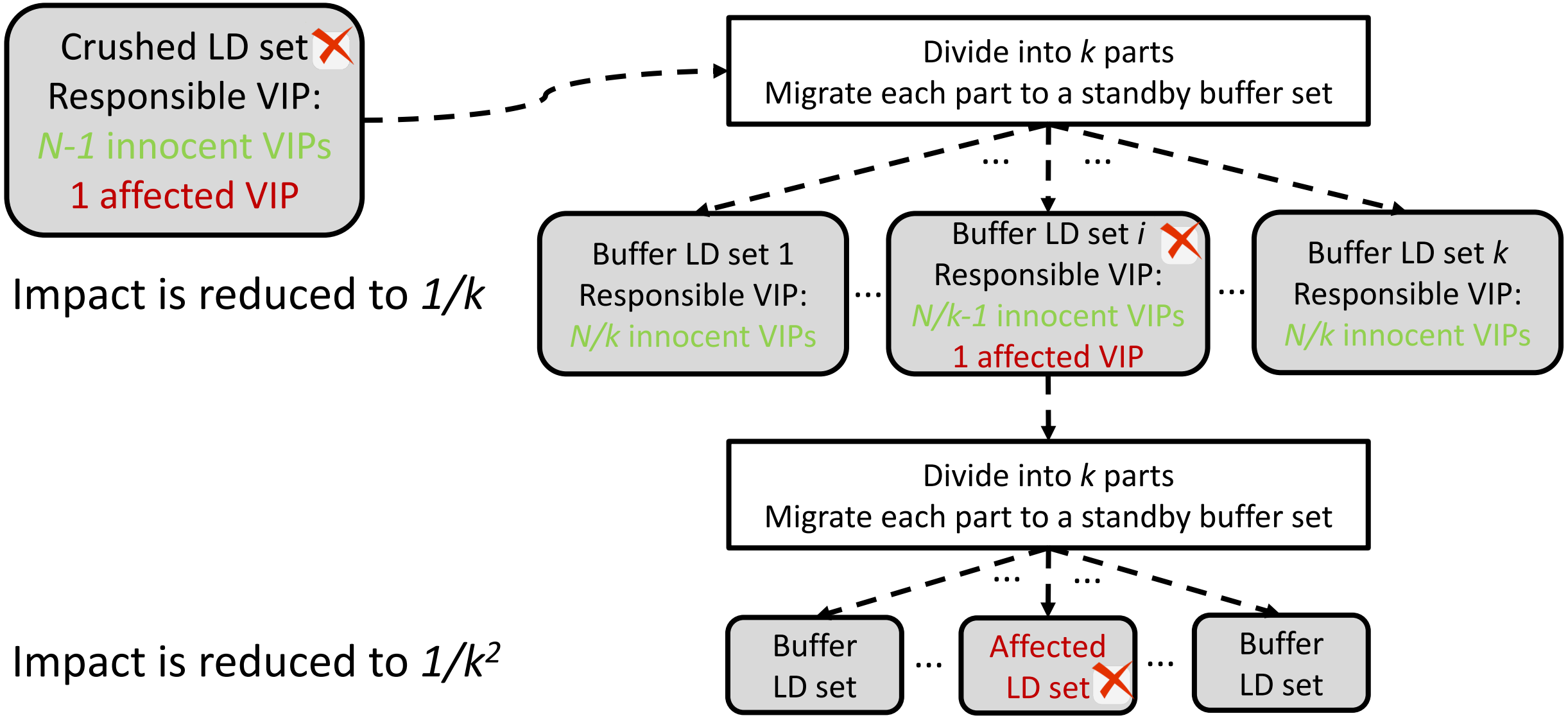
Problem: Flow-triggered failures cannot be resolved by our backup-based fault tolerance model



Observation: Such failures are usually caused by traffic destined for a single VIP

Solution: Decentralized migration

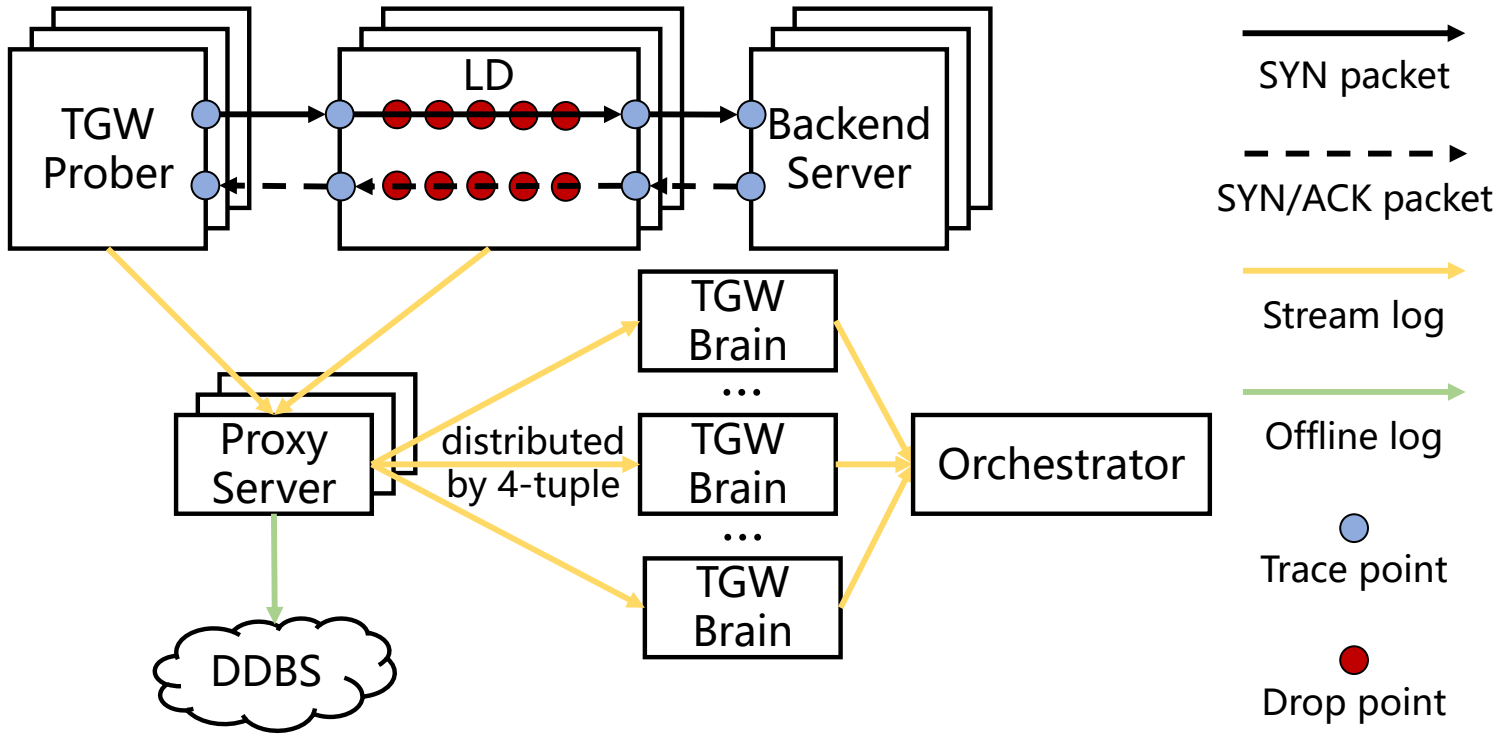
Decentralized Migration



Impact is reduced to $1/k$

Impact is reduced to $1/k^2$

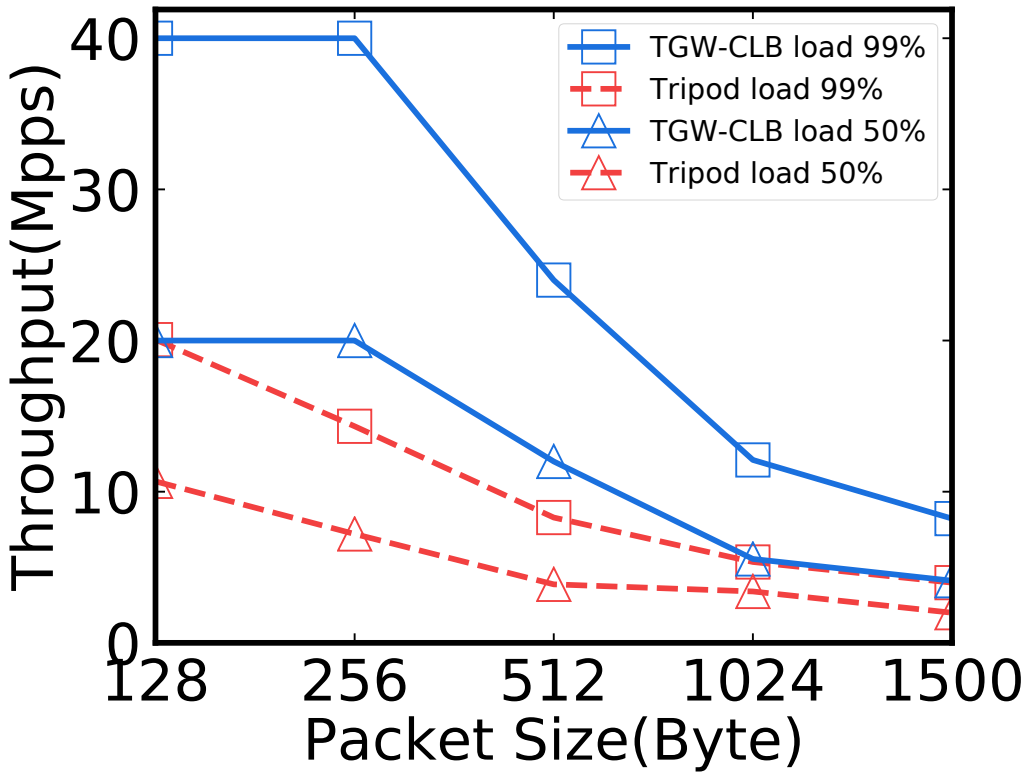
Failure Detection and Localization



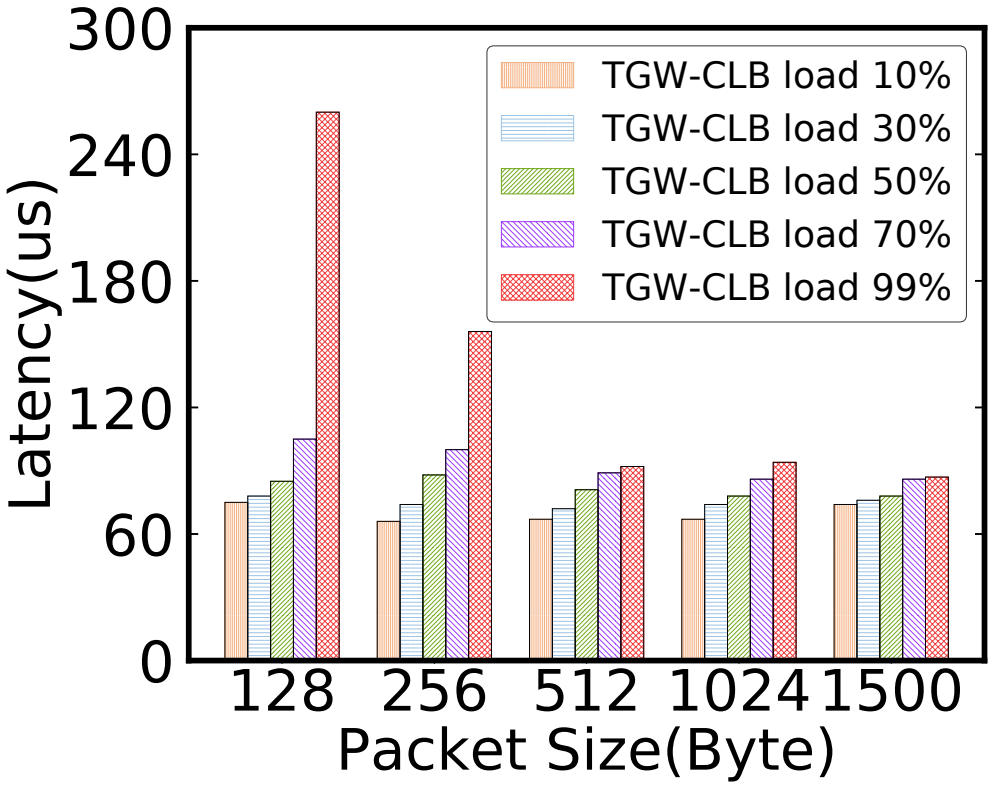
- TGW probers: Send probing packets to the backend servers and get the returned result
- Proxy servers: Aggregate probing logs and LD logs from TGW probers and LDs, and distribute them to TGW brains by 4-tuple
- TGW brains: Analyze logs in parallel and generate warnings and suggestions
- Orchestrator: Apply failure recovery plans
- Offline logs: Are stored in DDBS for future query

Evaluation: Testbed

Single-LD node performance



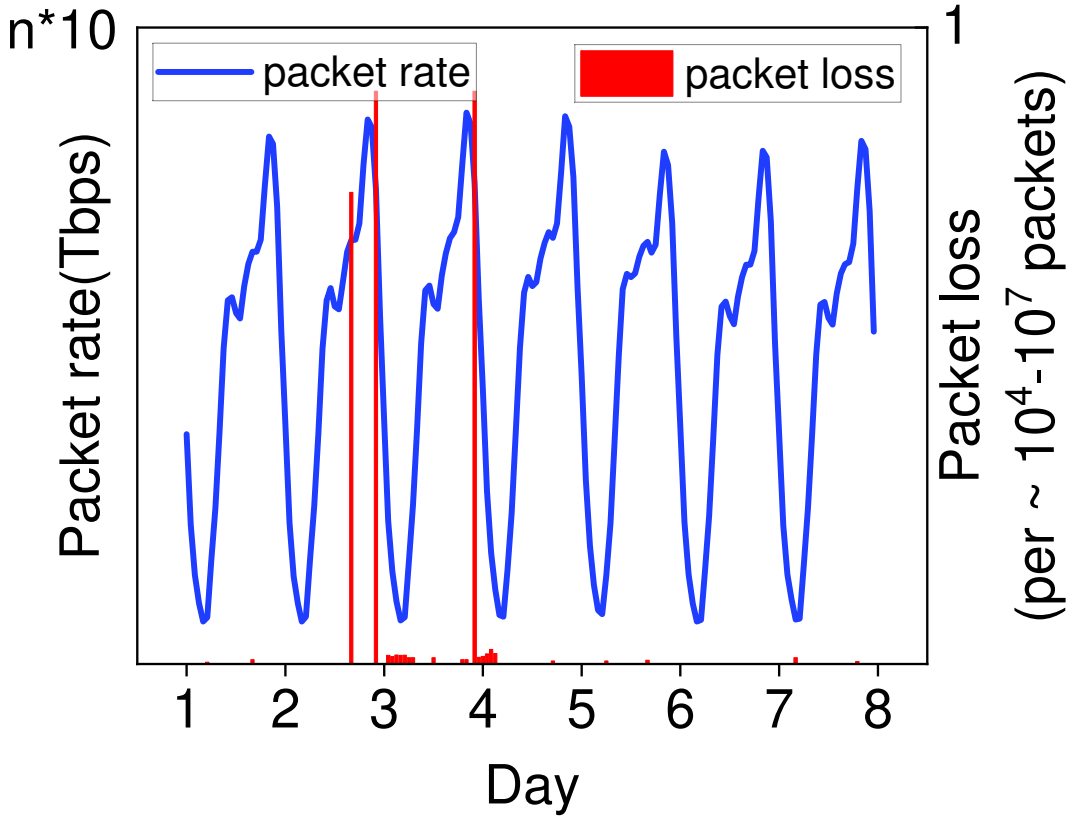
Single-LD throughput under different packet sizes and CPU load



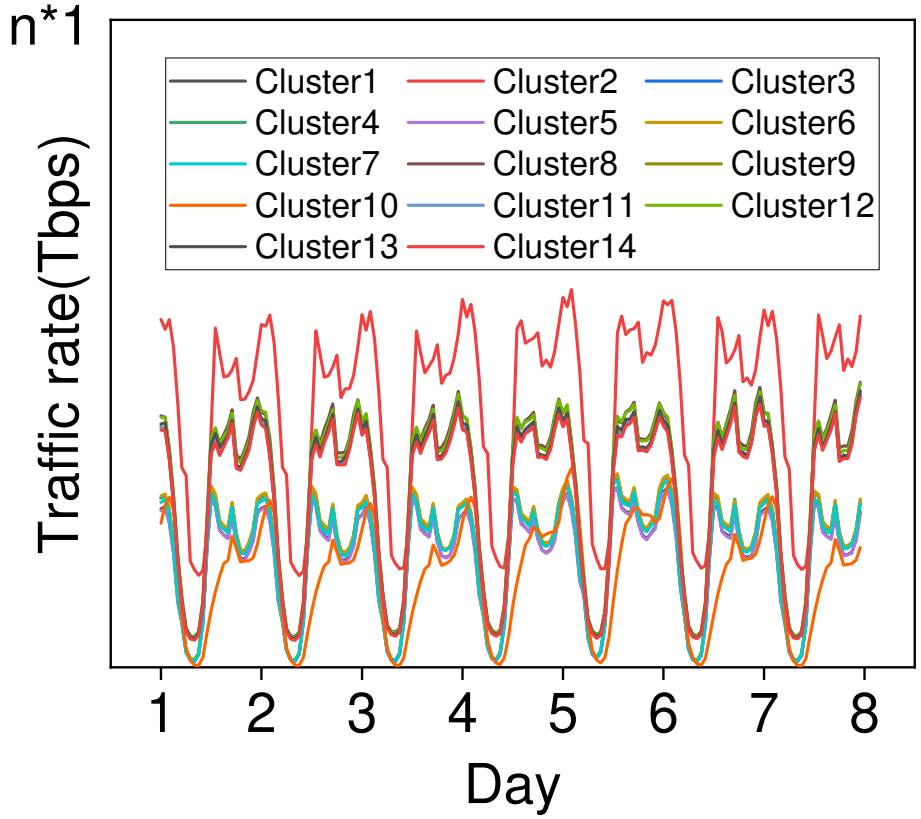
Single-LD latency under different packet sizes and CPU load

Evaluation: Real World

Real-world performance of a TGW region



Throughput and packet loss rate of a TGW region during a one-week shopping festival



Load distribution between clusters of a TGW region during a one-week shopping festival

Experiences and Lessons

Blast radius isolation

- Restrict the fault domain by a loosely coupled architecture
- Ad-hoc failure reaction (e.g., decentralized migration)
- Regional isolation

Redundancy: 50%

- AZ level: Two AZs are a redundant pair
- Rack level: Machines in a cluster are distributed across more than 2 racks
- Machine level: Half of the machines can carry all the traffic of a cluster
- Link level: Tolerating half of the NICs' or fibers' failure

Experiences and Lessons

Scalability and portability

- Distributed and modular design
- TGW is scaled to more than 30 regions and thousands of machines per region
- Mature technology and maintenance system
- TGW has been deployed for dozens of customer companies, including famous banks, TV stations, and car groups

The service type determines the properties of the infrastructure

- TGW is built for Tencent Cloud services
- Tencent Cloud's "killer" services are online gaming and live streaming
- Online gaming and live streaming customers strongly require a low latency, low jitter, and low packet loss network

More experiences and lessons: See our paper for details

Conclusion

TGW is a mature cloud gateway system for **Tencent Cloud**

- Mature TGW has been operated for more than 8 years and is scaled to more than 30 regions and thousands of machines per region
- TGW is composed of a highly optimized DPDK-based forwarding plane, a lossless live migration mechanism, a multi-level failure resilience model, and a telemetry-based failure detection and localization system
- TGW has been proven efficient, scalable, and highly resilient through our testbed experiments and real-world data.

This presentation is supported by the ***Deng Feng Fund*** of Tsinghua University

Thanks!

Contact: *yangyifan0202@gmail.com*