Upgrade of KEK Campus network

Soh Y. Suzuki, Tadashi Murakami, Fukuko Yuasa,
Ryouichi Baba, Toshiaki Kaneko,
Teiji Nakamura, Kiyoharu Hashimoto, Mitsuo Nishiguchi,
Hirofumi Maeda, Atsushi Manabe
Takanori Hara, Tomoaki Nakamura

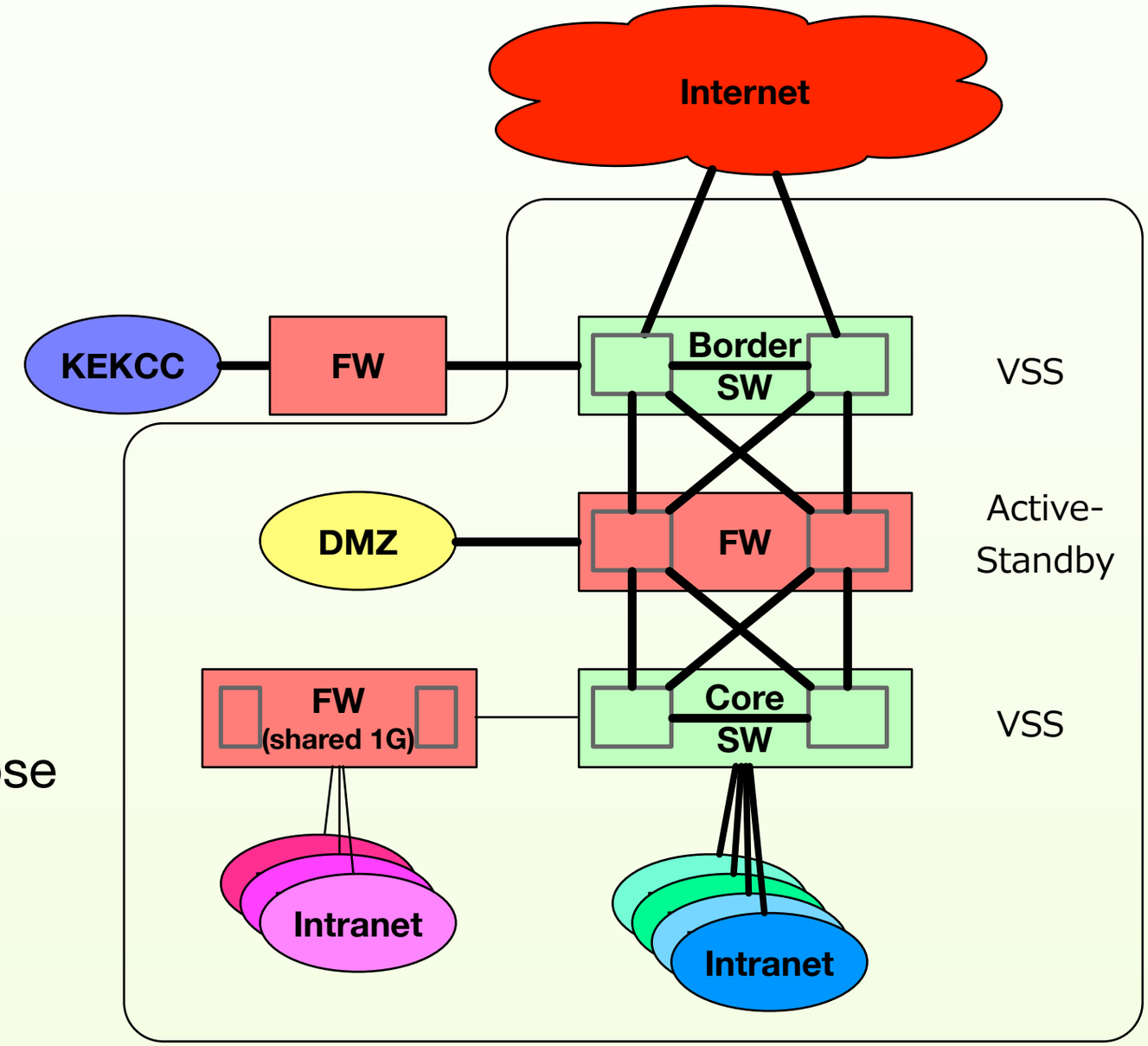
Computing Research Center, KEK

Contents

- Overview
- Migration from Catalyst6506 to ARISTA7280SR
- Migration from Catalyst6509 to Nexus9506
- Migration of ACL from Nexus9506 to Firewall
- Troubles on distribution switch

Previous campus network

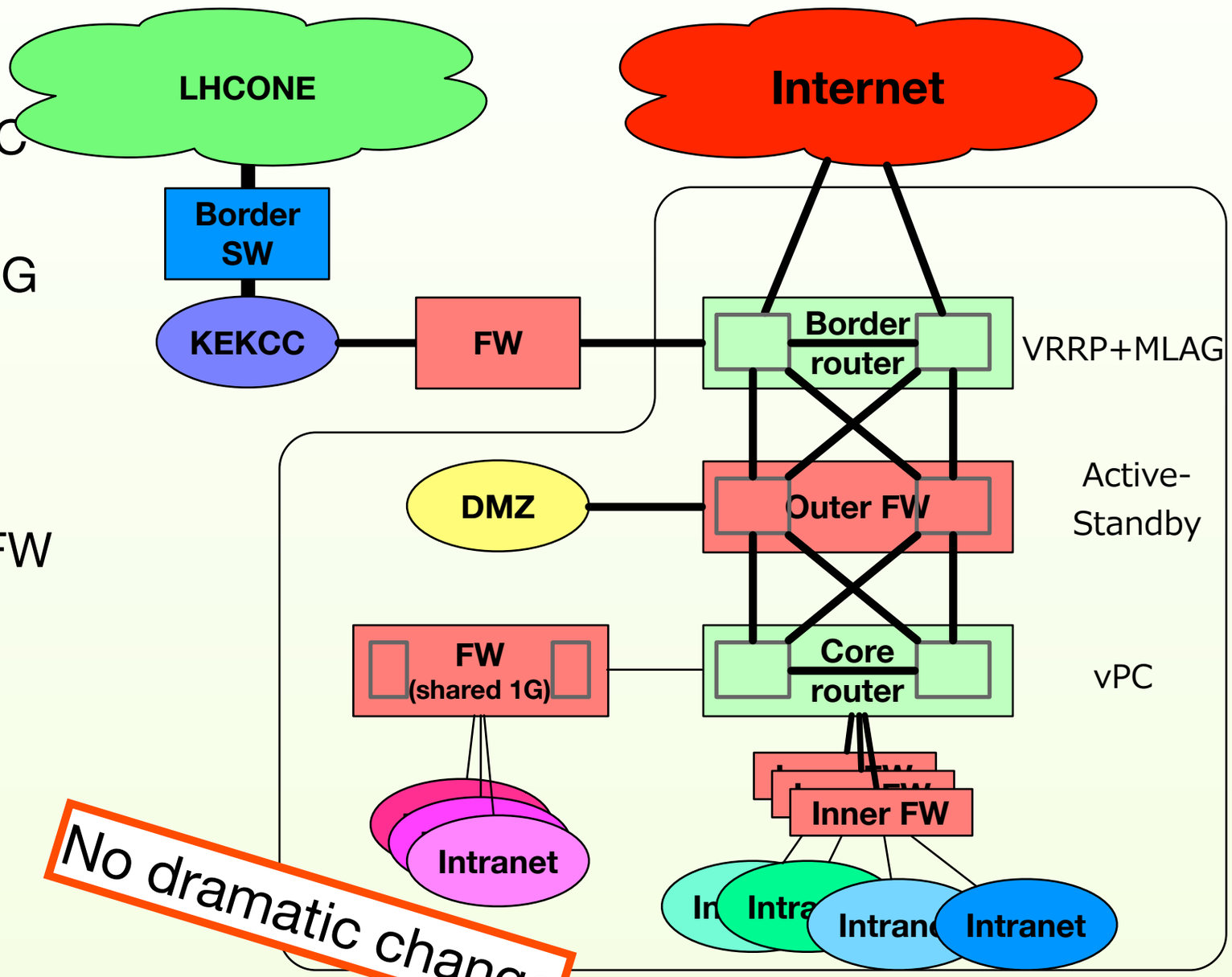
- connection to Internet
  - 2x10G(1st, 2nd) + 1x40G(3rd)
    - redundant, not load-balancing
    - Campus network shares single 10G with computing facility (KEKCC)
  - 10G IDS
  - LHCONE has another 40G
- Core part
  - redundant
  - 10G on FW is shared by all purpose and limits BW
- Distribution
  - 1G: > 90%
  - reliability > bandwidth



Campus network

Present campus network

- Connection to Internet
  - 1x40G(1st) + 2x10G
    - still shared with KEKCC
  - 40G IDS
  - LHCONE has another 40G
- Core part
  - redundant by vPC
  - Sufficient 10G ports on FW
  - Still 1G for distribution
- Distribution
  - 1G: > 90%
  - reliability > bandwidth
- Inner Firewall
  - to control session among intranets.

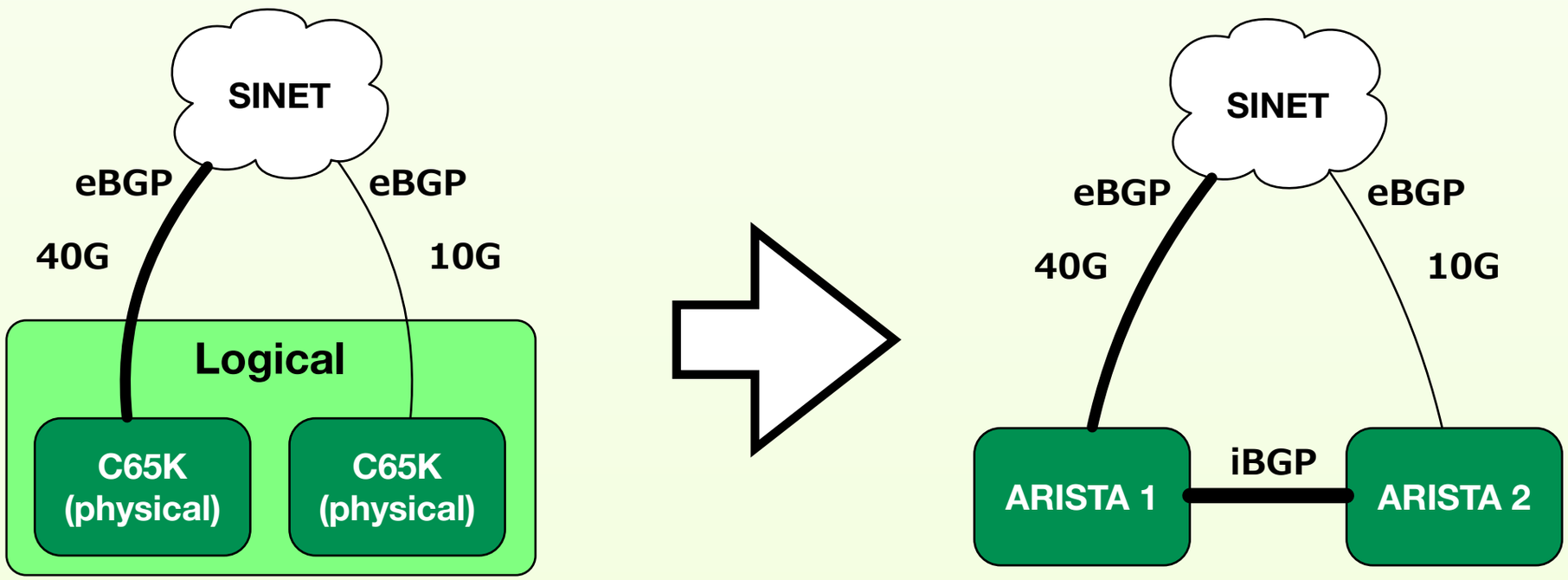


Campus network

Border switches: Catalyst 6506 → ARISTA 7280

- **MLAG instead of VSS by Catalyst 6500**
- **We have to split configuration to two switches.**
  - Hand-synchronization
  - Splitting eBGP peers into two devices needs iBGP
  - Relatively complex network

- We have multiple peers only to SINET, all other peers are moved to the primary.

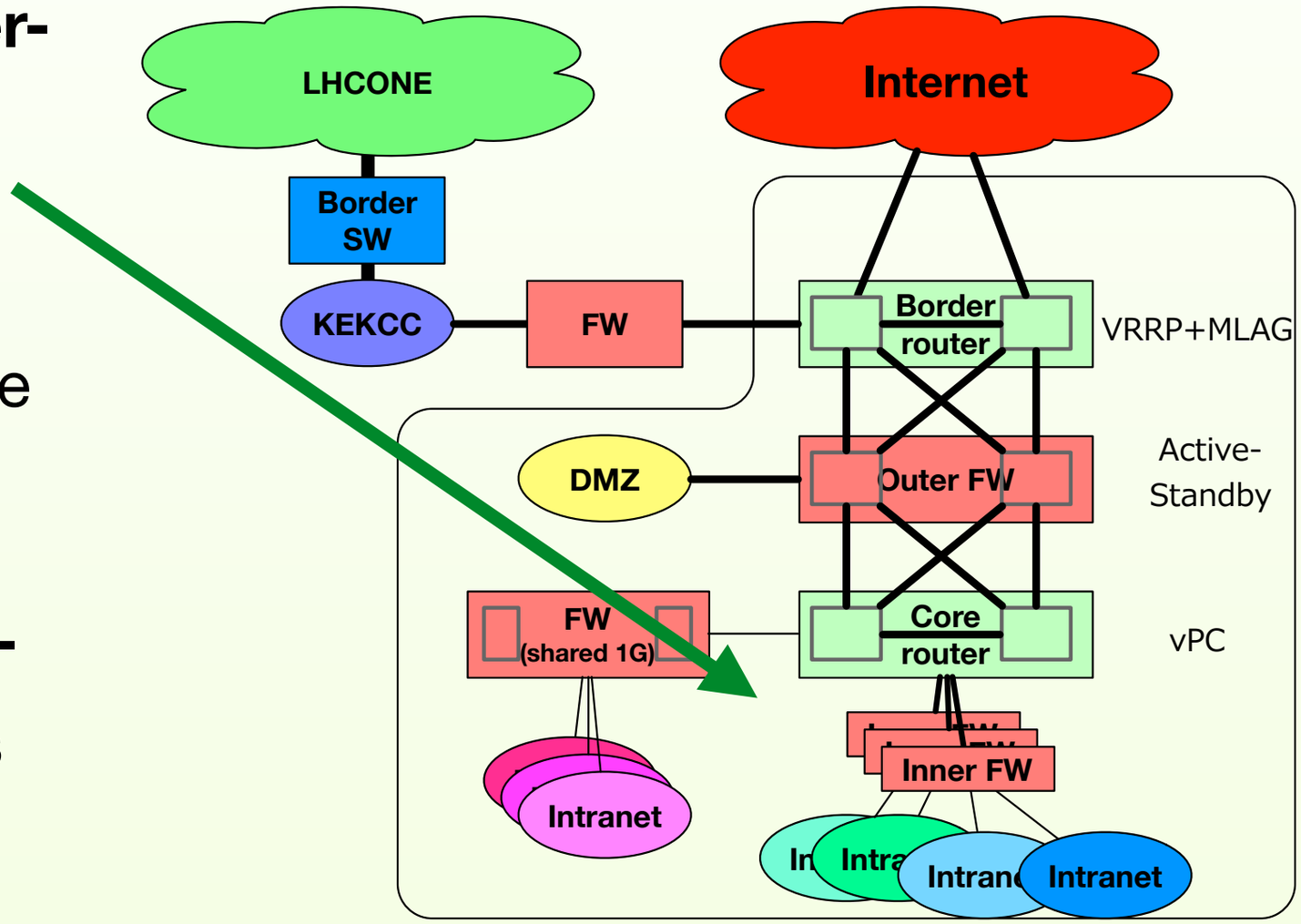


Core switches: Catalyst 6509 → Nexus 9508

- **Nexus 9500 provides vPC mechanism to synchronize configuration, but it is not VSS.**
  - We use HSRP for the redundancy
  - Big pitfall: vPC needs two living SUP at boot.
    - Don't reboot one when the other is dead
- **We choose ACL mode instead of ACI mode**
  - to migrate awful lines of ACL from Catalyst
  - In a test before installation, we confirmed all ACLs can be migrated if we remove obsolete ACLs.
  - But no newer ACLs can be added.

Migration of ACL from router to FW

- **We newly introduce inner-FW to control session among intranets.**
  - formerly ACL on the core router was used for that
- **We had to review all ACL lines by finding requests from old emails.**



Troubles

Small pitfalls in new border SW

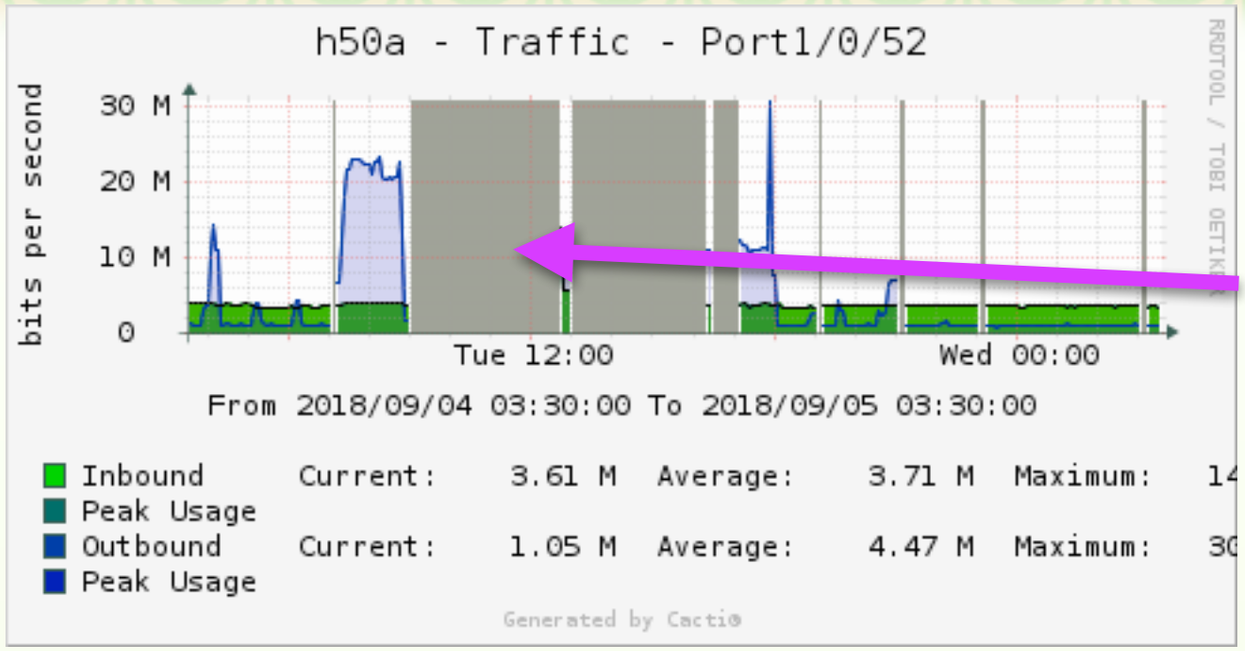
- Setting ACL containing TTL filter
  - Accepted although it replied that "it isn't supported in data-plane"
  - It actually discards all packets reaching the line
  - **We must not put ACL for Catalyst by copy-and-paste.**

- Default limit of BGP routes per peer is smaller than Catalyst
  - Only 12k routes
  - Once it exceeds, the peer becomes down and doesn't recover automatically.
  - **We extended it and need monitor.**

Pitfall in distribution switch

- New distribution switch uses totally rewritten OS.
- Load average of management CPU is **always** very high **only after** connected our campus network
  - Often ignores initial ARP to the management IP
    - takes 20 seconds to respond retries of ARP
  - slow response to SNMP request, timeout on SSH by scripts
- Unable to monitor all switches (~250) within 5 minutes
  - Parallel monitoring didn't solve it because of disk throttling.

Actually, it is not always!



cacti gave up get data
as it didn't respond in time

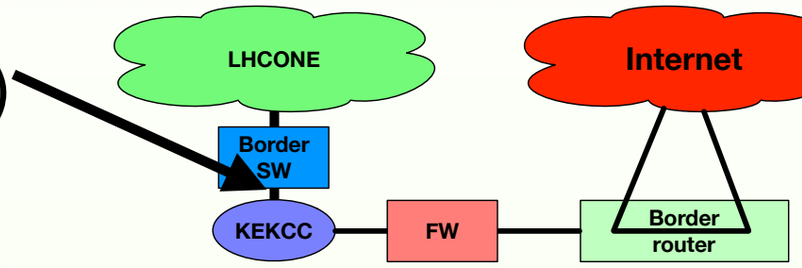
- Mostly the problem appears in working time.
  - During lunch time, it disappears.
  - When the working time is over (~17:15), immediately it disappears.
- The problem is caused by incorrect processing of unnecessary broadcast packets. (forwarded to management CPU)
- Half year later, OS update mostly solved the problem.

Now solved

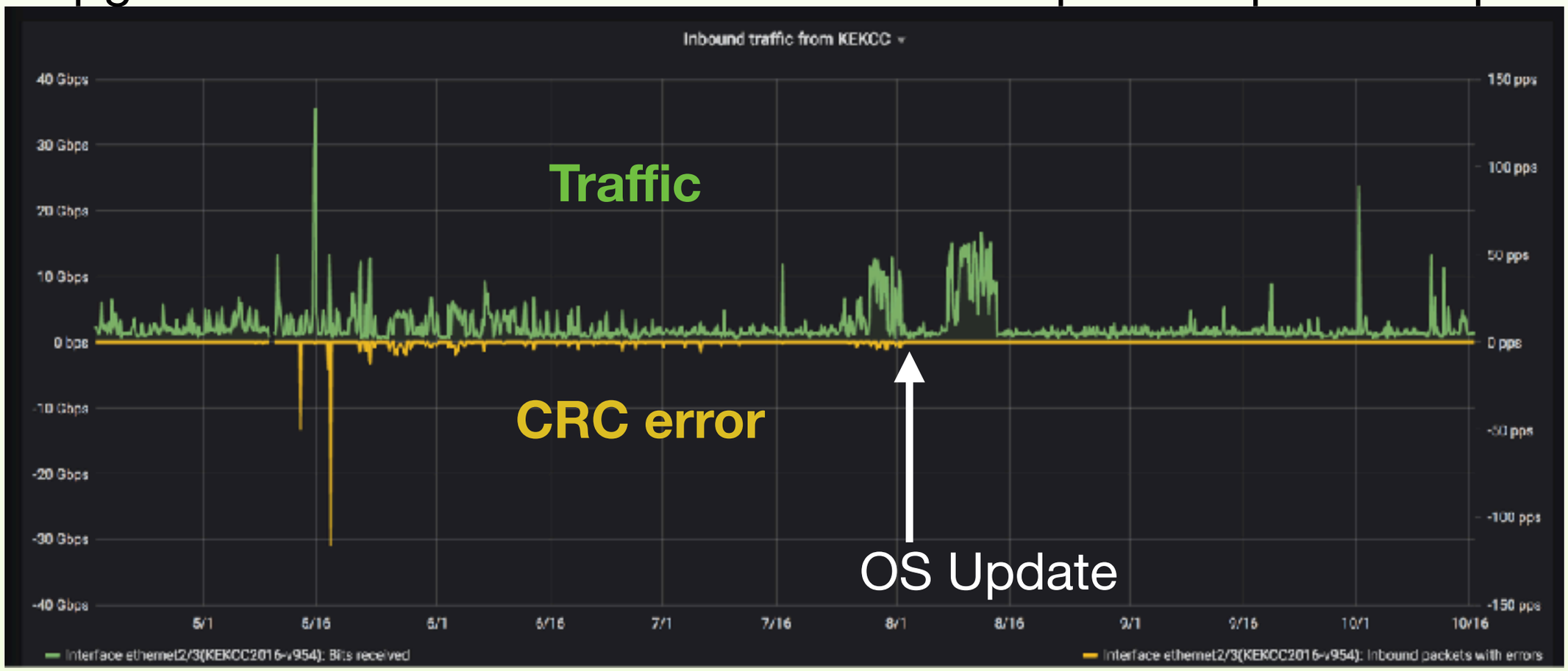
Traffic / metric history

- **Cacti → nearly giving up**
  - Tools based on RRDTools require high random I/O performance when parallel data collection is enabled.
  - In our case, it should write 134GB in every 5 minutes (> 400MB/s)
    - to hold 5 years history of samples in 5 minutes interval
- **Zabbix (3.4 → 4.0 → 4.2)**
  - Initially the performance was good, but housekeeping made high disk IO after few months later.
  - Using 4.2 and TimeScaleDB solves it... but I worry about the situation in 1 year later.
- **Visualization by Grafana**
  - more impressive dashboard
  - requires relatively higher random I/O
  - Caching by SSD and dm-cache significantly mitigate it, but SSD burn-out stops recording data at all. (Last weekend I met it)

CRC error on link for LHCONE



- Just on **our** 40G-LR4 link (MLXe4 and Nexus7K)
- Observed at MLXe side
- After starting Zabbix, we found continuous CRC error
- correlation to traffic is very low
- Replacing both optics didn't solve at all, but disappeared after OS upgrade and reboot of MLXe. No need to replace expensive optics?



Summary

- We have successfully upgraded campus network with small pitfalls
  - IDS and the primary link are upgraded to 40G.
  - Traffic from campus network to internet doesn't interfere that from KEKCC anymore.
- Storing traffic data with 5 minutes interval more than few years is difficult rather than we expected.