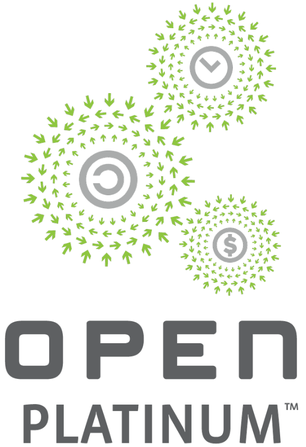
Open. Together.

OCP SUMMIT

SONiC Development for Large Scale Operations

Guohui Wang

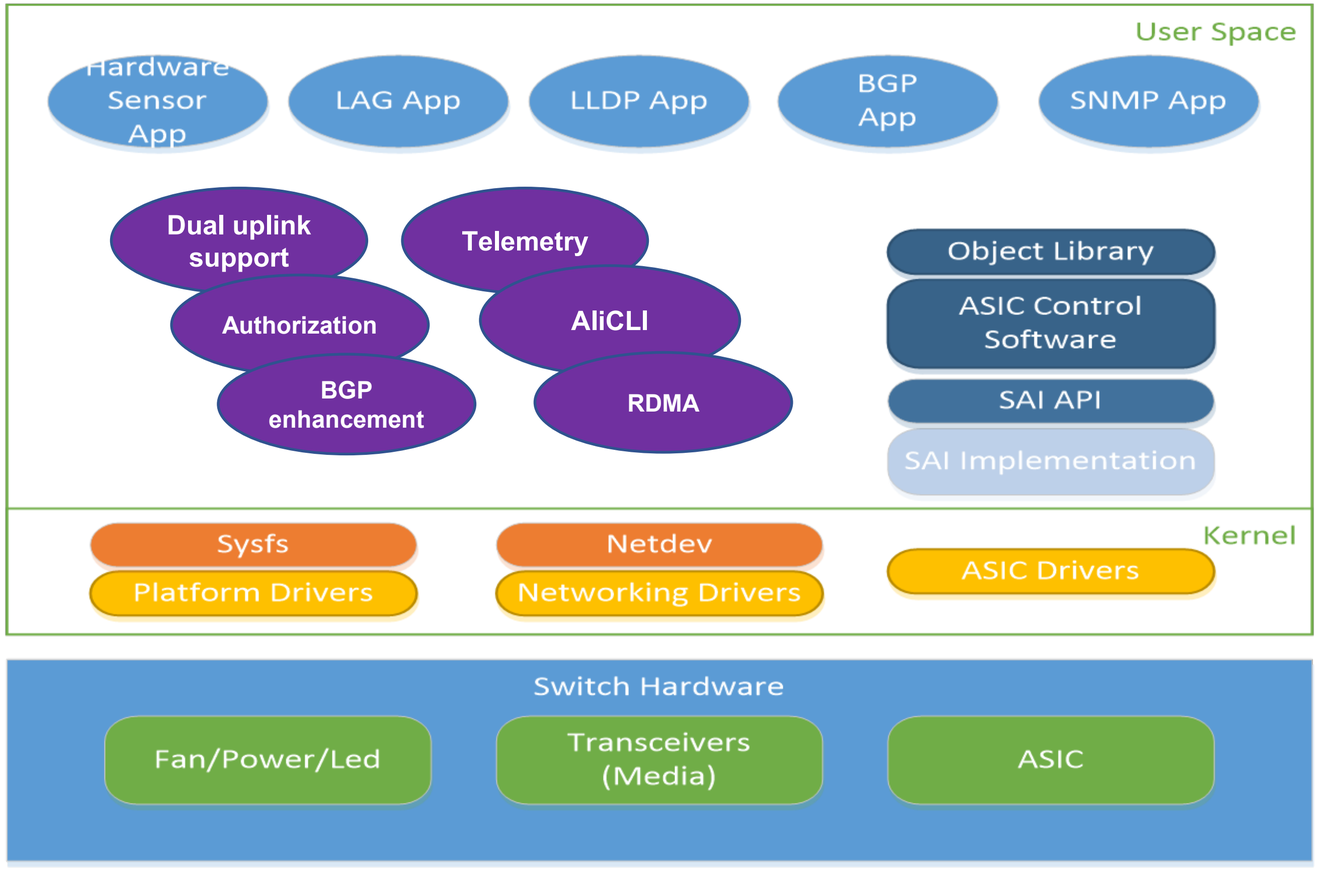
Engineering Director,
Alibaba Group



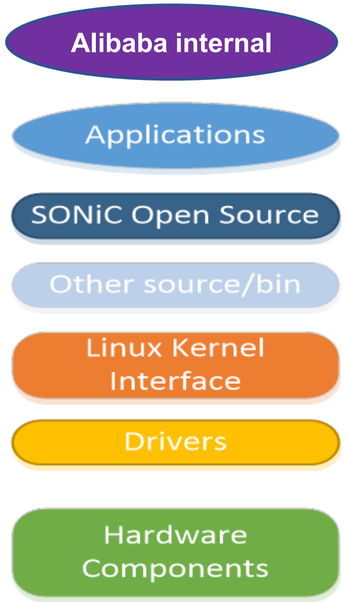
Alibaba and the SONiC community

- Alibaba joined SONiC community in 2017;
- Active participation and contributions
  - TACACS
  - VLAN trunking
  - sonic-telemetry
  - SWSS warm reboot
  - SysDB and routing performance optimization

SONiC Adoption in Production



- Picture from SONiC website



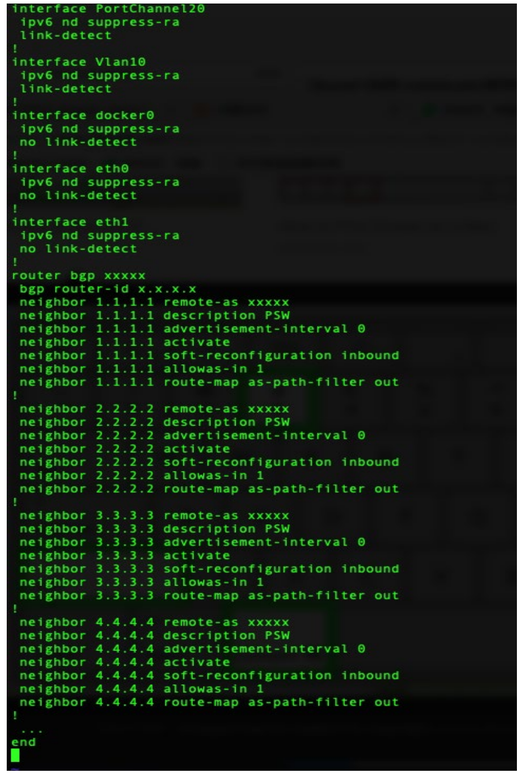
Running SONiC Fleets in Production

- Configuration
- Monitoring
- Maintenance and failure recovery
- Software upgrade and iterations

Do things in better ways with the open platform!

Configuration Management with structured APIs

- CLI-based configuration
  - Interaction latency in seconds
  - Hard to parse, program and verify



<u>BGP neighbor AS number update:</u>

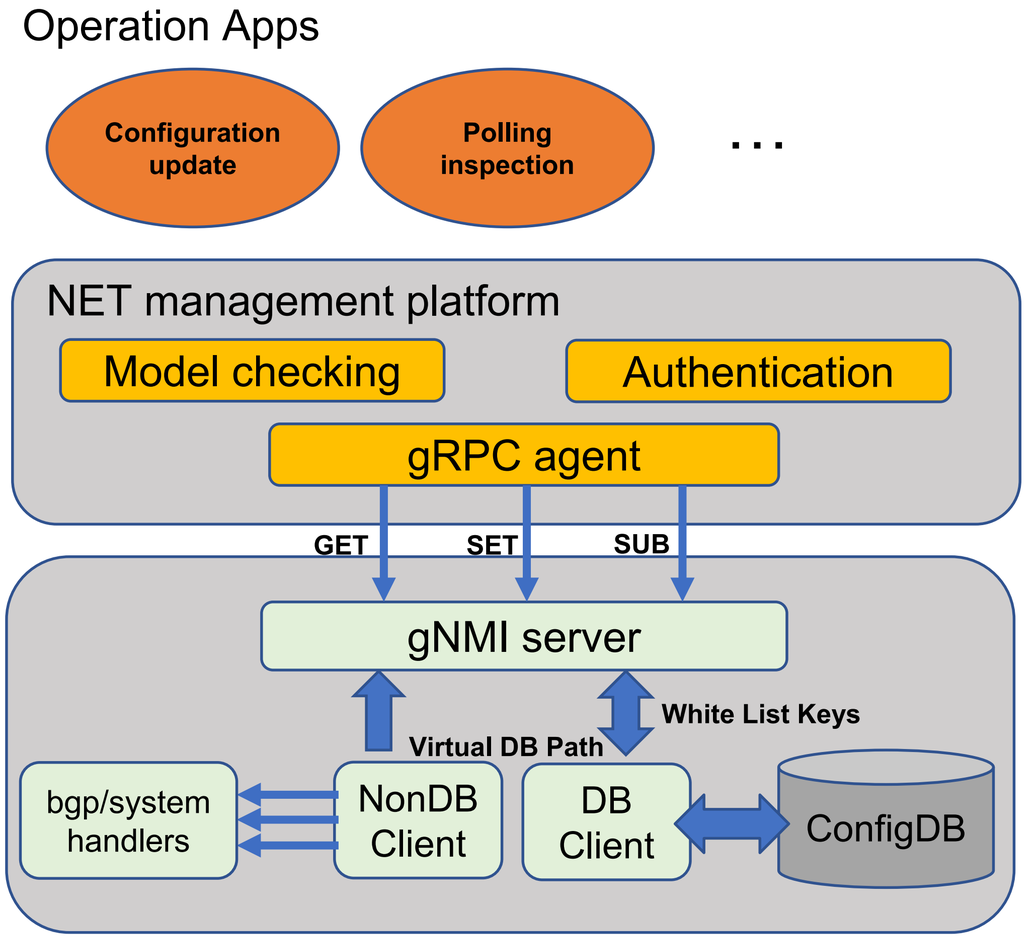
ASW-xxxx# config
ASW-xxxx (config) # bgp xxxxx
ASW-xxxx (config-router)# neighbor 1.1.1.1 remote-as yyyyy
ASW-xxxx (config-router)# do show running-config
…
Parse running config to verify
…
ASW-xxxx (config-router)# exit
ASW-xxxx(config)# exit

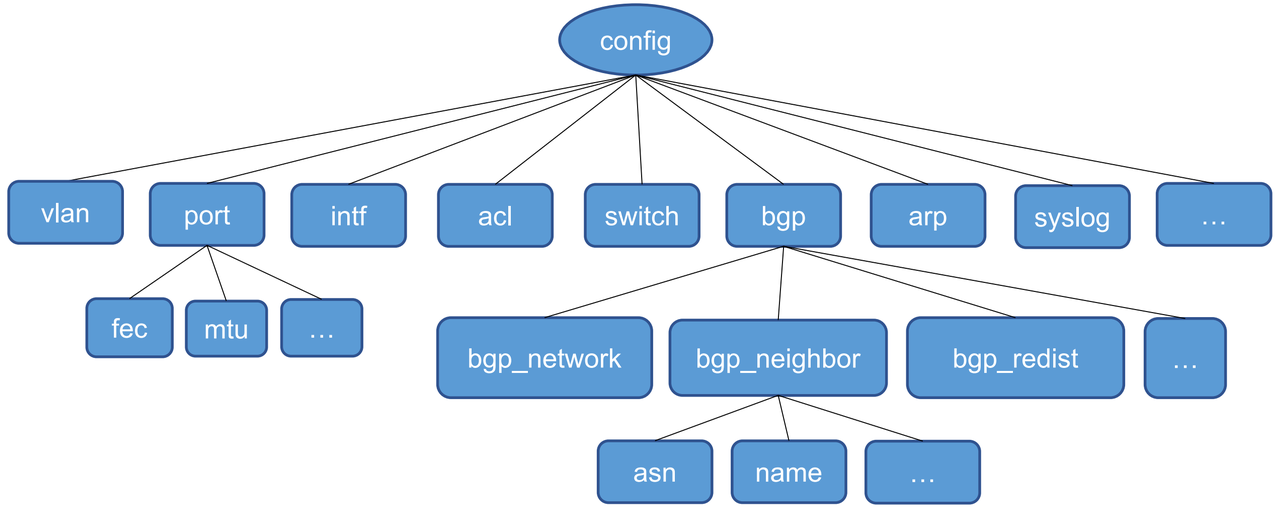
- Configuration with structure gRPC APIs
  - API latency in ms
  - Easy to model, program and verify

<u>BGP neighbor AS number update:</u>

SET BGP_NEIGHBOR/1.1.1.1/asn yyyyy;
GET BGP_NEIGHBOR/1.1.1.1/asn;

Configuration Management with structured APIs



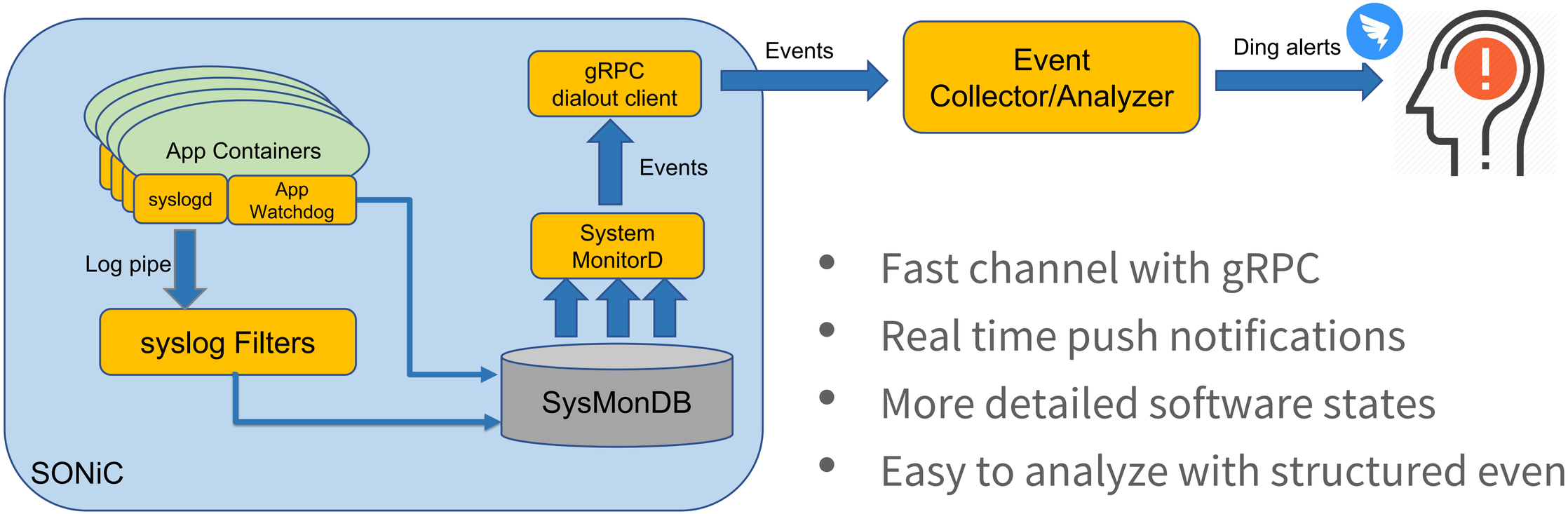


- gNMI-based data schema
- Device abstraction layer
  - virtual DB paths to hide system details;
  - white-list keys for access control;

Traditional Device Monitoring

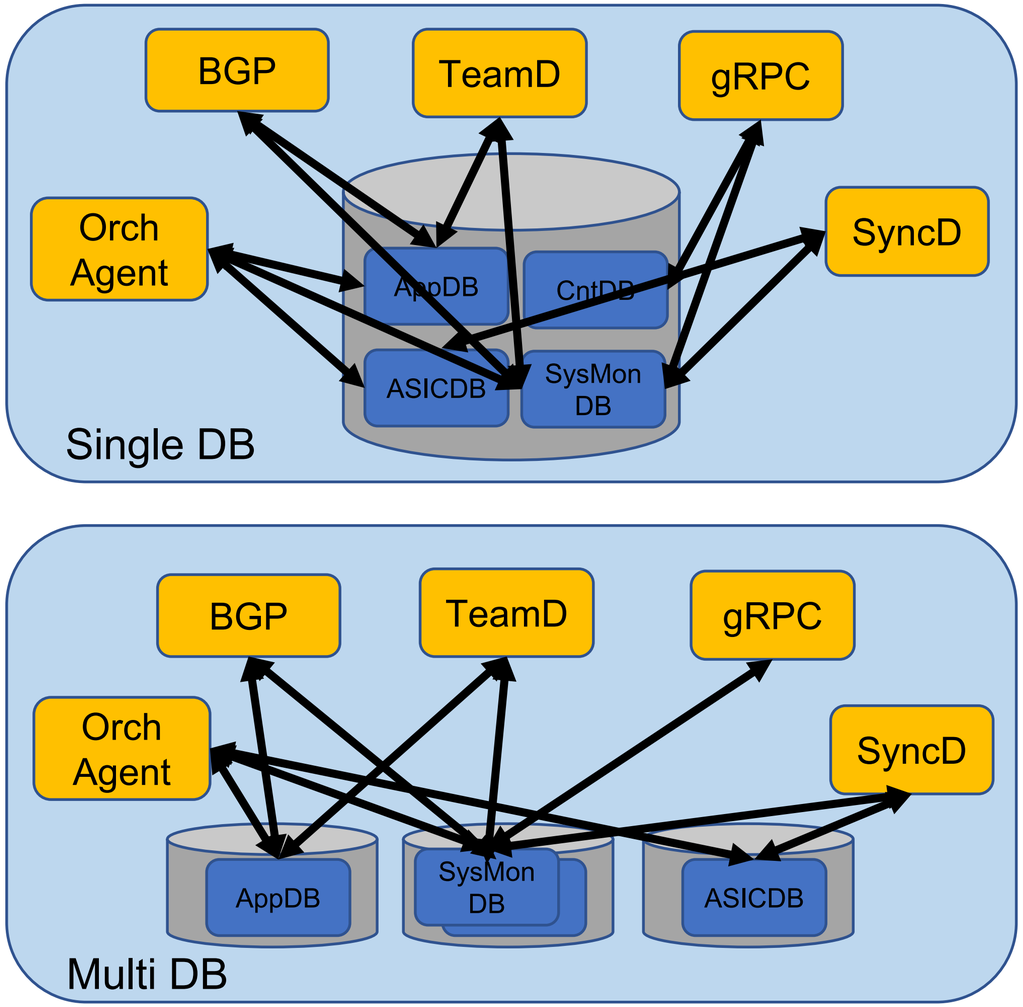
- SNMP polling inspection every 5 minutes
  - Long monitoring delay, inflexible structure, legacy code;
- Syslog monitor
  - Noisy data, hard to analyze;
- Black-box monitoring:
  - No internal software states;

Event-Based Device Monitoring



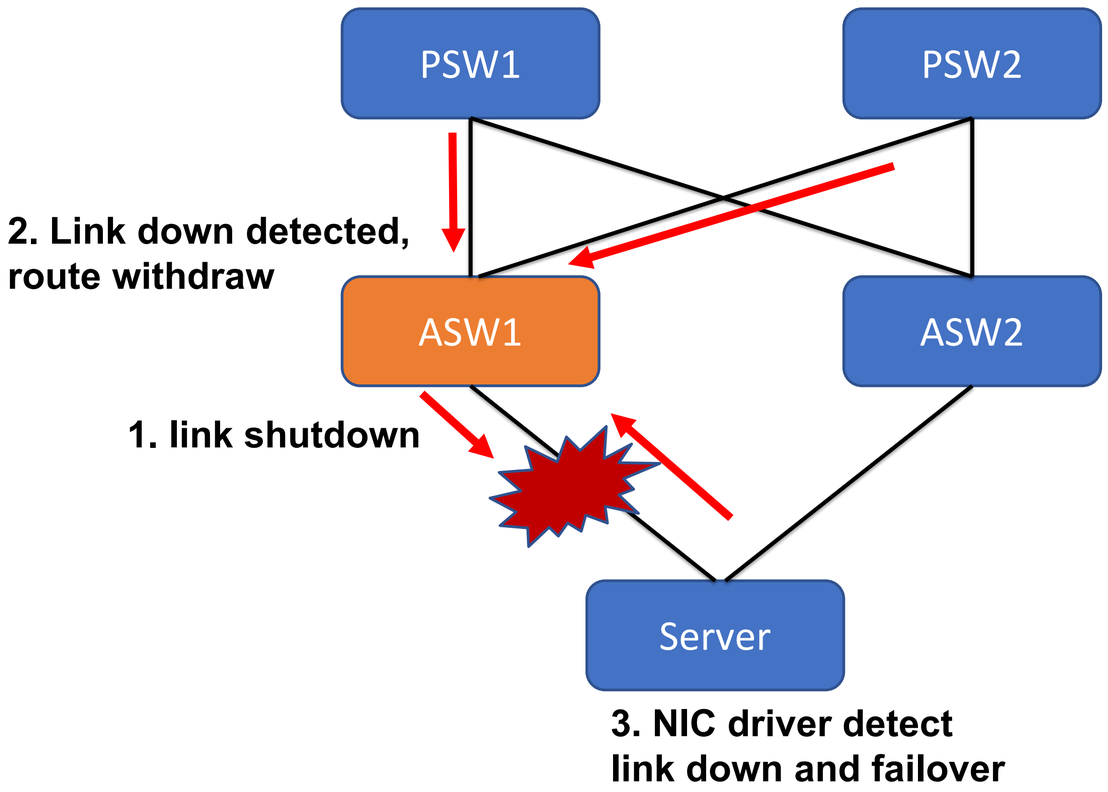
- Fast channel with gRPC
- Real time push notifications
- More detailed software states
- Easy to analyze with structured events

SONiC multi-DB optimization



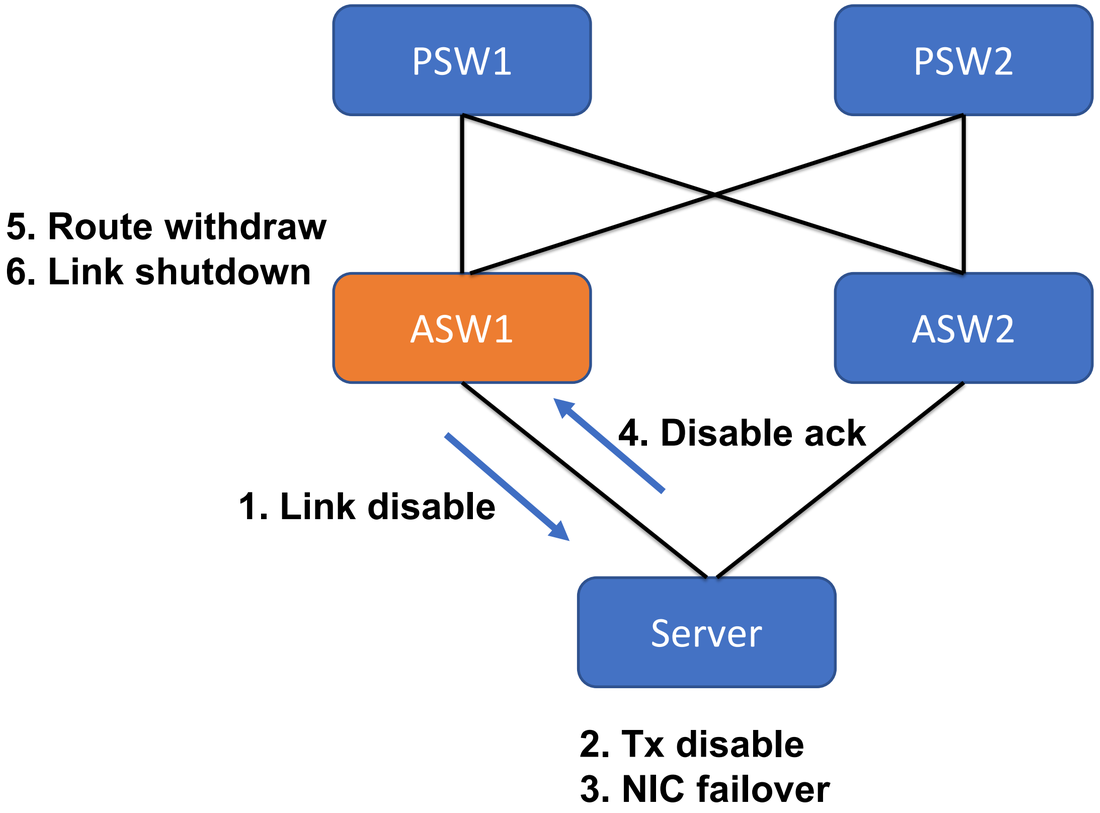
- Single-instance RedisDB is a system-wide bottleneck
- SONiC re-structured to support multi-DB instances
  - Multi-instance DB configuration at build time;
  - Dynamic binding of DB clients and instances;
  - Separate DB instances for routing/monitoring/management, over 50% improvement on route installation perf;

Maintenance Example: Device Isolation

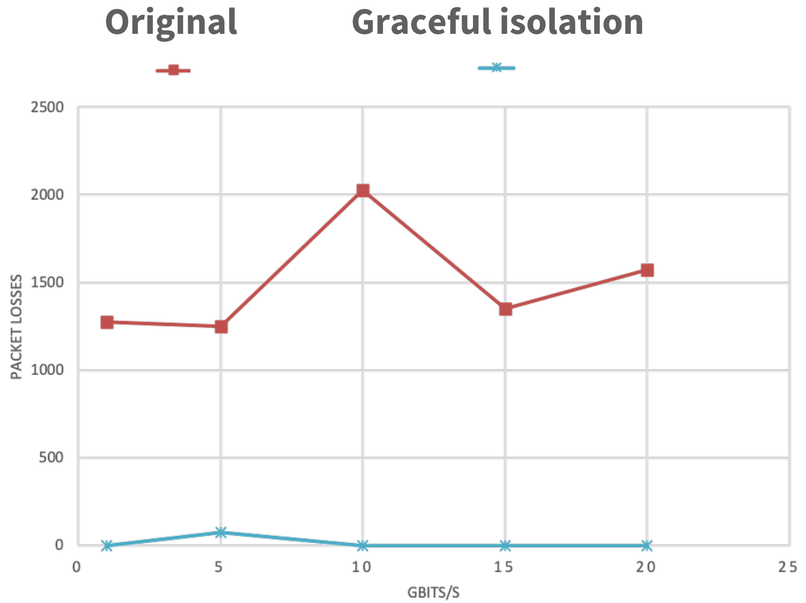


- Disruptive device isolation:
  - No switch and server coordination, purely rely on the propagation of link down event;
  - Large amount of packet losses between event 1, 2 and 3 on both outgoing and incoming traffic;

Graceful Device Isolation



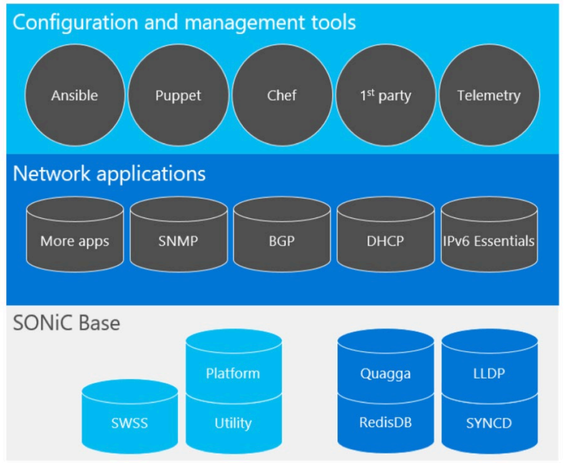
- Coordinated events between switch and server through customized protocol;
- graceful traffic failover and device maintenance;



Packet losses of a TCP flow during device isolation

Software Upgrade

- From hot-fix/cold-fix to modular software upgrade
  - Debian package and docker based upgrade;
  - Tooling for modular version control;
- Non-disruptive software upgrade with docker and system warm reboot

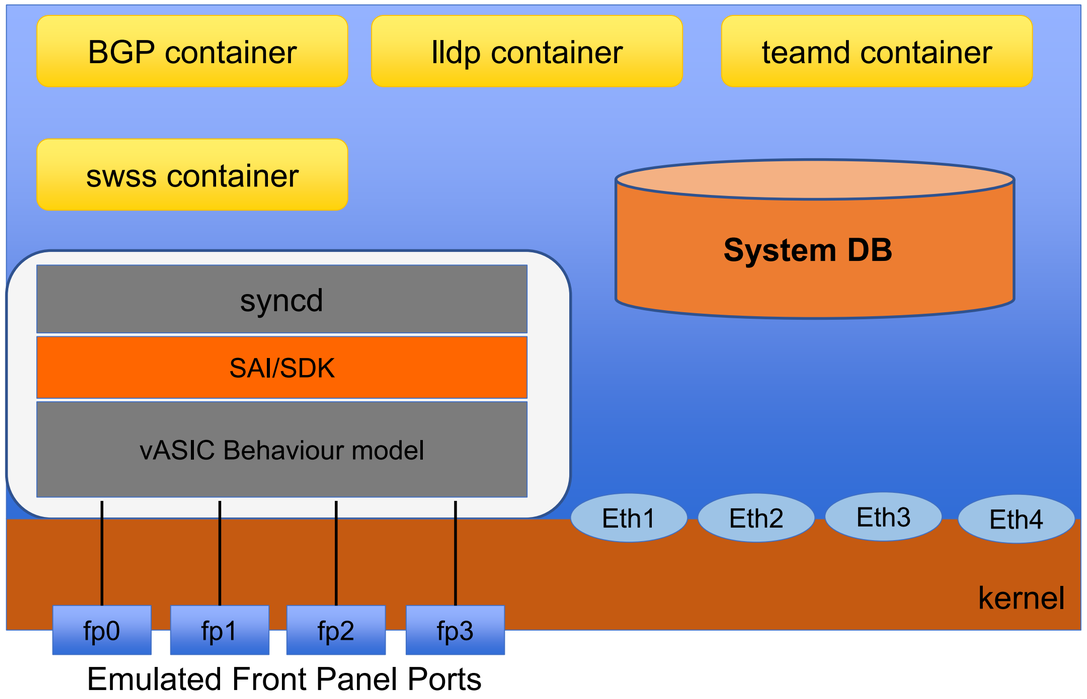



```
root@ASR-C2-4-C11-A.NA61:/home/admin# show version
SONiC Software Version: SONiC.SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 (*)
Distribution: Debian 8.11
Kernel: 3.16.0-4-amd64
Build commit: 80959af
Build date: Mon Aug 13 02:58:10 UTC 2018
Build number: 3
Built by: judong@sonic-build

Docker images:
REPOSITORY                  TAG                                 IMAGE ID            SIZE
docker-orchagent-bfn        SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 1c5075c388c       386.6 MB
docker-orchagent-bfn        latest                              1c5075c388c       386.6 MB
docker-syncd-bfn            SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 da411c5ab538      2.127 GB
docker-syncd-bfn            latest                              da411c5ab538      2.127 GB
docker-dhcp-relay           SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 c2b479d07783      378.1 MB
docker-dhcp-relay           latest                              c2b479d07783      378.1 MB
docker-database             SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 bfbcc2907024      377.7 MB
docker-database             latest                              bfbcc2907024      377.7 MB
docker-teamd                SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 9131b7b35f20      382.8 MB
docker-teamd                latest                              9131b7b35f20      382.8 MB
docker-snmp-sv2             SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 9ca93654beb5      443.9 MB
docker-sonic-telemetry      SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 1d928f467a59      485.7 MB
docker-sonic-telemetry      latest                              1d928f467a59      485.7 MB
docker-router-advertiser    SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 b54101e05021      375.3 MB
docker-router-advertiser    latest                              b54101e05021      375.3 MB
docker-platform-monitor     SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 58e9af8d7342      396.9 MB
docker-platform-monitor     latest                              58e9af8d7342      396.9 MB
docker-lldp-sv2             SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 dba62db9c4fb      382.6 MB
docker-lldp-sv2             latest                              dba62db9c4fb      382.6 MB
docker-fpm-quagga           SONiC-OS-AliNOS-rel-v1.3.1-22-846fae6 6ddba7e5a050      389.4 MB
docker-fpm-quagga           latest                              6ddba7e5a050      389.4 MB
docker-snmp-sv2             AliNOS-rel-v1.3.1-15-1880d2d        75e96e39ecc9      443.9 MB
docker-snmp-sv2             latest                              75e96e39ecc9      443.9 MB

Debian package version:
Name                                Version                 Dirty   Description
-----------------------------------------------------------------------------------------------------------------------
libnss-tacplus                      1.0.4-1+3-80959af               NSS module for TACACS+ authentication without local passwd entry
libpam-tacplus                      1.4.1-1+3-80959af               PAM module for using TACACS+ as an authentication service
libtac2                             1.4.1-1+3-80959af               TACACS+ protocol library
python-sonic-utilities              1.1-375-80959af                 Command-line utilities for SONiC
rsyslog                             8.4.2-1+deb8u2+2-80959af        reliable system and kernel logging daemon
sonic-cli                           1.0.0                   *       sonic-cli built using CMake
sonic-platform-ingrasys-s9280-64x   1.1.0+2-80959af                 This package contains S9280-64X platform driver utility for SONiC project.
```

Faster Iterations with AliNOS Emulator



- Fully virtualized SONiC device running in a VM
  - QEMU-kvm + ONIE x86_64-kvm;
  - SONiC + VM specific platform modules;
  - SDK/SAI + vASIC simulation model;
- Application Scenarios
  - Development and integration test;
  - Operation rehearsal;
  - Software verification and troubleshooting;

Real SONiC + **Real Config** + **vASIC model + x86** + **Virtual topo** + **Generated traffic**

Lessons learned

- Dangers are in the grey zone
  - Tricky issues come from platform/firmware/hardware-software interoperability problems;
  - Monitoring and real time fault detection is the key;
- Building systems with operations in mind
  - From CLI/SNMP to RPC-based system ；
  - Customized protocols for graceful operations;
  - Non-disruptive upgrade with docker and system warm reboot;
- Automated testing and operation rehearsal bring faster iterations;

Call to Action

- Building SONiC with strong operation supports
  - Operation tooling: failure handling, trouble-shooting;
  - Software-driven management interfaces;
  - System and network visibility;
  - Version control and software iterations;


Open. Together.
OCP Global Summit | March 14–15, 2019
OCP
SUMMIT