# SONiC VOQ System

March 2020

BROADCOM

VOQ Switch System - Overview

- VOQ Switch system consists of several Line Cards populated with Switch devices, and Fabric Cards populated with Fabric devices
    - Note: a single Switch device, with no Fabric is also a valid VOQ Switch System

- Each Line Card is running a single SONiC image with separate BGP, ASIC-DB, SAI, etc. per Switch device

- Fabric devices are generally treated as "transparent" full mesh connectivity Fabric SONiC does not participate in BGP
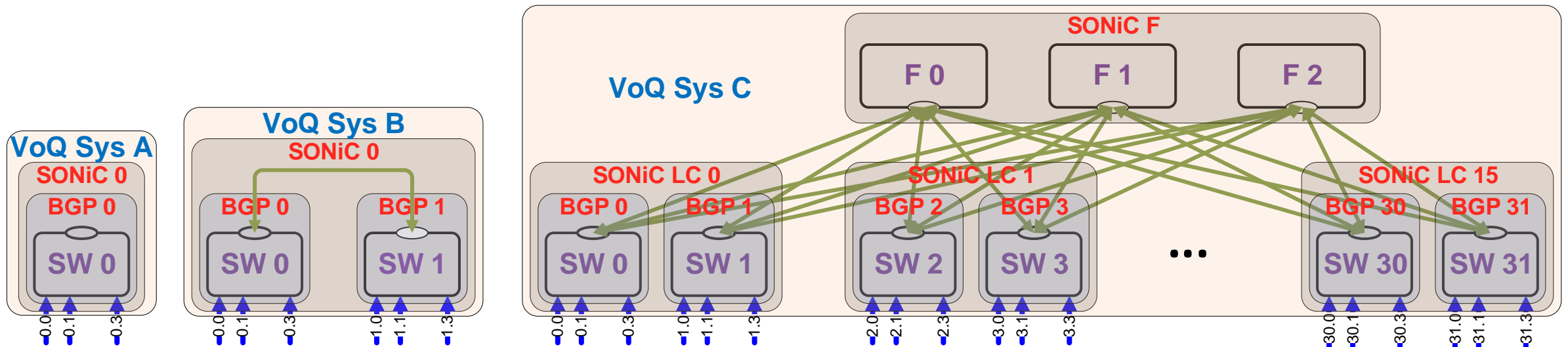
VOQ System - Unfolded view, packet walk

- Unfolded view better explains the VOQ Switch System data plane (traffic flow from left to right)

- We define all Network facing ports and CPUs (or other internal units) ports as System-Ports System-Port IDs are globally assigned

- Packets are received (from Network) and processed on an Ingress Device, where packet's destination System-Port and quality of service parameters (TC) are resolved

- The Ingress device includes an array of Virtual Output Queues (VOQ), each VOQ store traffic destined to a specific System-Port and Traffic Class (TC)

- The destination System-Port and TC are resolved to a VOQ

- System-Ports (on Egress devices) schedule traffic pointing to it from all relevant VOQs on Ingress devices (and cores)

- Packets sent from Ingress to Egress devices uses the "Fabric" as a fully connected self routed cross connect media

- "Fabric" is the collection of all Fabric ports, links, and devices
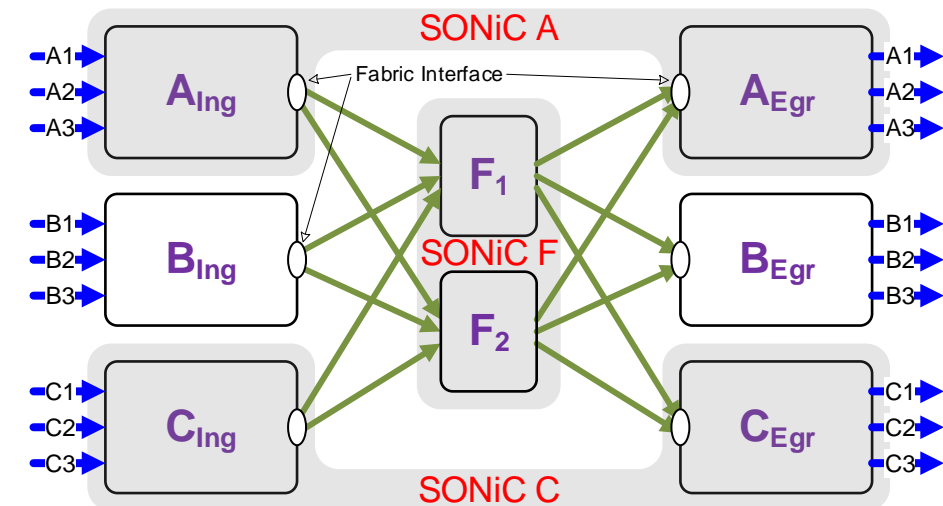
VOQ Scheduling

- Each switch have in its ingress a bundle of VOQs (one per TC) per each System-Port

- Traffic from all ingress VOQs pass the fabric to remote Egress-Queues, per System-Port and TC

- Credit Scheduler dispense credits to its VOQs, according to a configured QoS policy
  E.g. Strict-Priority, Weighted-Fair-Queue

- The egress ports arbitrates between its Egress-Queues.
  The arbitration policy is the same QoS policy as used by the Credit Scheduler

- Congestion control, Drop/Color statistic are applicable only to the ingress VOQs
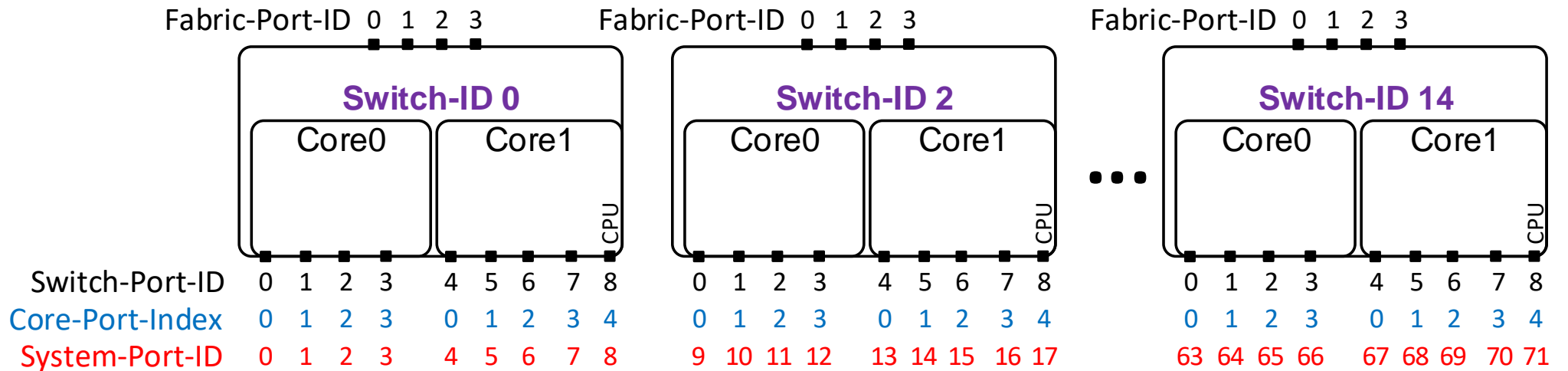
BROADCOM®
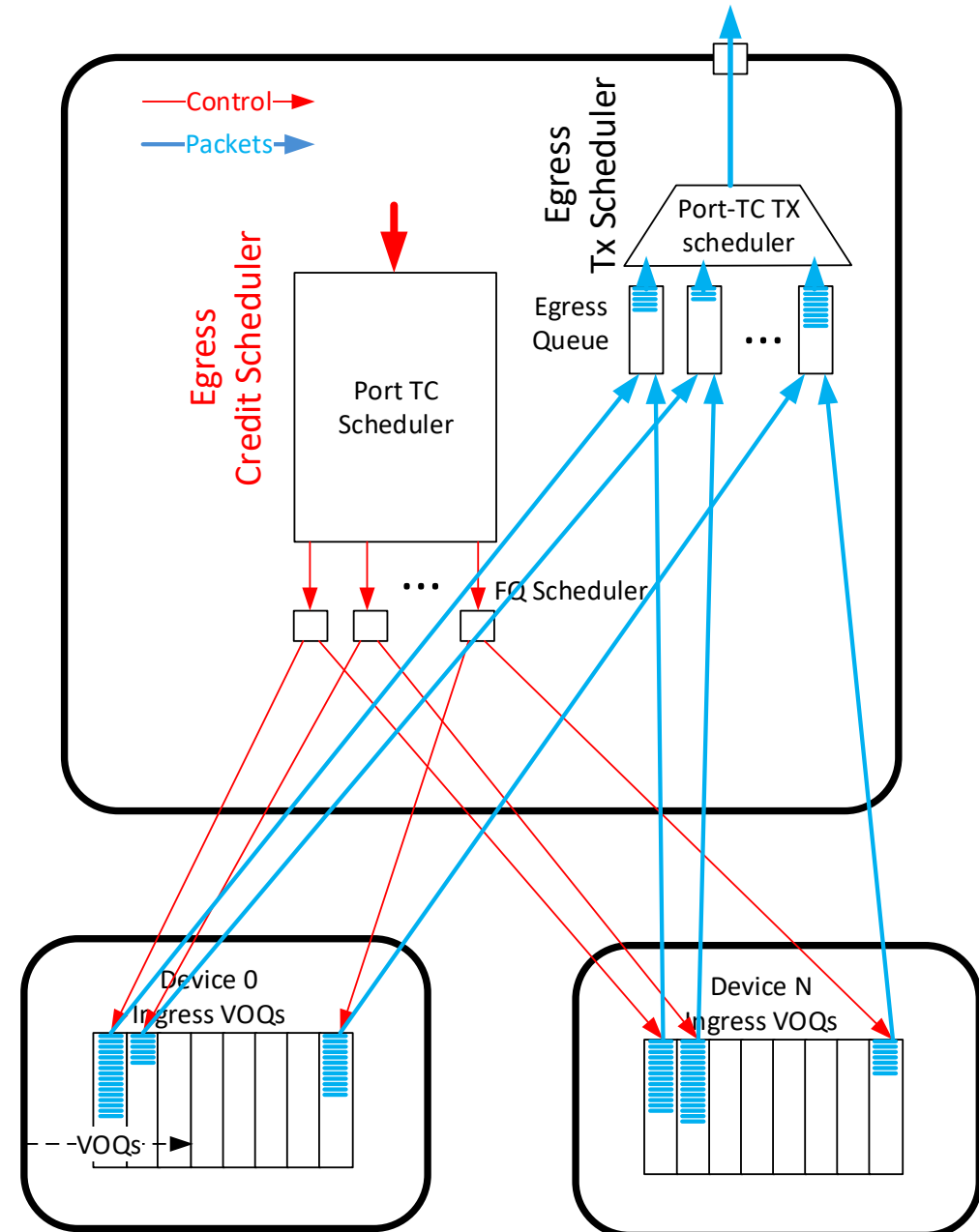
VOQ System - SAI extension proposal (1/2)

- SAI System-Port - A new SAI object globally assigned to all Network facing ports and CPUs (or other internal units) ports on a VOQ system

- A single Switch device may be composed of multiple Cores (A.K.A pipelines), the Core number, and total number of Cores in the system are required for some system creation processes

- Fabric devices and Switches on the line cards are connected through Fabric port
  Fabric-Port is a new subtype of the Port object
  Fabric-Ports are mainly used in system diagnostic procedures

- A VOQ line card Switch has 3 separate lists of ports:
  – Local Ports (A.K.A Ports)
  – System-Ports
  – Fabric Ports

- Note: all switches have the same System-Ports list consisting of all Network ports in the entire chassis, including its own ports, and all CPU ports.

BROADCOM®

VOQ System - SAI extension proposal (2/2)

- In a VOQ Switch, the SAI_QUEUE_TYPE_UNICAST object refers to the Egress Queues. While the ingress VOQs use a new sub-type SAI_QUEUE_TYPE_UNICAST_VOQ.

- A System-Port definition comprises the identification of its location in the chassis: Device-ID, Core-Index, Local-Core-Port-ID, and a list of VOQs, corresponding to the Traffic Classes.

- Each VOQ is associated with a WRED-Profile. Typically Drop and Color statistics are applied to the VOQs (rather than the Egress Queues)

- Fabric port is associated with one to three TX- Fabric-Queues, whose IDs are the object handles for statistics collection.

BROADCOM®

VOQ System - SAI extension, example

- The drawing below depicts an 8 line cards switches VOQ system

- The 8th Switch, with Switch-ID=14 will have the following SAI Ports' lists:
    - Local Ports (0, 1, 2, 3, 4, 5, 6, 7, 8)
    - Fabric Ports (0, 1, 2, 3)
    - System-Ports (0, 1, 2, 3, 4, …, 68, 69, 70, 71)

Thank You