OCP Tech Talk 2022

Rita Hui, Principal Engineering Manager, Microsoft
Shrikrishna Khare, Software Engineer, Meta
Switch Abstraction Interface

- CRUD operation over extensible Entity/Attribute/Value data model
- Reference data-plane behavior model supports various devices
- Significant feature/partner growth since announcement in 2014
- [https://github.com/opencomputeproject/SAI](https://github.com/opencomputeproject/SAI)
SAI Function Calls & Pipelines

- Register SAI APIs to achieve specific functions and control logging separately
- Construct the right order to execute SAI function calls
- Refer to SAI Pipelines to create the right packet action flow
SAI Metadata

SAI metadata is set of validation checks, enhancements and auto generated data from SAI headers, that are used at compilation time or at run time.

Few highlights:

• Information about each attribute and enum
• Serialization methods
• API and notification structures
• Object dependency graph
• Style check
• Attribute and object validation

Predefined: saimetadatatype.h

Auto generated: saimetadata.c and saimetadata.h

To generate the metadata:

$ cd SAI/meta
$ make
## SAI Release Highlights since Last OCP Update in Mar 2020

### SAI 1.7

**Nov 2020**
- Policy Based Hash
- Masked Hash with Optional Ordering
- MP2P and MP2MP tunnel termination table entry
- User Defined ACL
- Support max accumulative headroom size
- PFC pause duration in microseconds
- Gearbox port failover
- Multiple interface types on a port
- Mirroring Sampled Packets into Sflow Tunnel
- Query Statistics capability

### SAI 1.8

**Mar 2021**
- Packet header based VRF classification
- Query available packet DMA pool size
- Add IPv6 flow label hash attribute
- Override VRF
- Switch scoped tunnel attributes
- MPLS Enhancements
- Support for ACL extensions in metadata
- More Metadata Checks
# SAI Release Highlights since Last OCP Update in Mar 2020

<table>
<thead>
<tr>
<th>SAI 1.9</th>
<th>SAI 1.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sept 2021</td>
<td>Mar 2022</td>
</tr>
<tr>
<td>• IPSec</td>
<td>• Autogenerated framework for PTF SAI Tests</td>
</tr>
<tr>
<td>• SRv6 programming model updates</td>
<td>• 900+ New SAI PTF test cases covering 28+ features</td>
</tr>
<tr>
<td>• NVGRE tunnel</td>
<td>• Mirror Sessions Counter</td>
</tr>
<tr>
<td>• VXLAN user defined UDP source ports</td>
<td>• SAI NAT aging notification</td>
</tr>
<tr>
<td>• Class-based Forwarding</td>
<td>• Tunnel Scoped TC MAP and Remarking</td>
</tr>
<tr>
<td>• Trap and drop counter for MPLS lookup miss</td>
<td>• Attribute for PFC Deadlock</td>
</tr>
<tr>
<td>• Assign values to enum fields</td>
<td>• 8-lanes interface types</td>
</tr>
<tr>
<td>• Enable/disable forwarding pause frames</td>
<td>• Get Bulk Stats API</td>
</tr>
<tr>
<td>• Additional FEC modes for PAM4 ports</td>
<td>• Bulk Set for NH Group Object</td>
</tr>
<tr>
<td></td>
<td>• Bulk API for setting Port Attributes</td>
</tr>
<tr>
<td>SAI Roadmap</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>SyncE for Phy</td>
<td></td>
</tr>
<tr>
<td>VLAN Stacking and Translations</td>
<td></td>
</tr>
<tr>
<td>MPLS Scaling</td>
<td></td>
</tr>
<tr>
<td>DPU</td>
<td></td>
</tr>
<tr>
<td>SAI Generic Extension</td>
<td></td>
</tr>
<tr>
<td>SAI Warmboot PTF Test and More PTF Improvements</td>
<td></td>
</tr>
</tbody>
</table>
PINS - SAI Generic Extensions

- P4 used to model the SAI pipeline
- PINS Capabilities
  - SDN managed SAI tables
  - SAI Generic Extension tables (New)
- SAI Generic Extensions
  - Allows realization of new use cases
  - Via addition of new match-action table
  - Underlying SAI pipeline as baseline
- SDN protocol: P4Runtime
  - Enables runtime-control of data plane objects

Slide courtesy of Reshma Sudarshan, Ravi Vantipalli, Shitanshu Shah from Intel
New SAI object type - **SAI_OBJECT_TYPE_GENERIC**

A Key-based object can be described using

1> Object name – Is the name of match action table and can be either a string or integer.

2> Object Attribute – This node is a list of fields describing a set of attributes that carry the essential data required to program the HW block.

**SAI_GENERIC_PROGRAMMABLE_ATTR_ENTRY** is in the form of `sai_json_t` and it defines a generic data type using JSON as the data format with a predefined grammar allowed to be used within.

```plaintext
Programmable_object_attr.id = SAI_GENERIC_PROGRAMMABLE_ATTR_OBJECT_NAME
Programmable_object_attr.value = "Table Name"
Programmable_object_attr.id = SAI_GENERIC_PROGRAMMABLE_ATTR_ENTRY
Programmable_object_attr.value = {
  "attributes": [
    {
      <attribute_name>: {
        "sai_metadata": {
          "sai_attr_value_type": "<SAI_ATTR_VALUE_TYPE_T>",
          "brief": "Brief Attribute Description",
          "sai_attr_flags": "<SAI_ATTR_FLAGS_T>",
          "allowed_object_types": [ "<LIST_OF_ALLOWED_OBJECT_TYPES>" ],
          "default_value": "<DEFAULT_ATTR_VALUE>"
        },
      "value": <VALUE of the attribute>
    }
  ]
}
```
Meta SAI Update
FBOSS

- Facebook Open Switching System (FBOSS)
- Meta’s software stack for managing network switches in Meta’s datacenters
Meta’s SAI Spec Contributions

- [srikrishnagopu] enable/disable decrement ttl support for port, rif and nexthop (#1050)
- [srikrishnagopu] Add queue index to hostif packet attribute (#1049)
- [parvezshaikh] MPLS Host-If traps for packets with expiring TTL and Router Alert Label (#1062)
- [parvezshaikh] Add new hash fields for MPLS labels (#1058)
- [parvezshaikh] Add MPLS QoS Maps (#1060)
- [parvezshaikh] MPLS ACL filters (#1064)
- [parvezshaikh] Provide TTL and QoS treatment during MPLS encap and decap (#1079)
- [parvezshaikh] Additional Port Interface Types (#1098)
- [parvezshaikh] Add CR2/SR2 interface types (#1115)

8 Feature Enhancements
8 MPLS Enhancements
3 Warmboot Enhancements
Meta’s SAI Spec Contributions (contd).

- [srikrishnagopu] mirroring sampled packets into sflow tunnel (#1082)
- [srikrishnagopu] renaming decrement ttl enum to disable decrement ttl enum (#1114)
- [jasmeetbagga] Add packet allocate, free function to allow for 0 copy packet TX (#1137)
- [jasmeetbagga] Add attribute to query available packet DMA pool size (#1198)
- [Midhun Somasundaran] Adding counter id attribute for mpls insegment (#1228)
- [Midhun Somasundaran] Add drop counter for mpls lookup miss (#1227)
- [Midhun Somasundaran] Add a trap type for mpls lookup miss (#1252)
- [shri-khare] Assign values to enum fields (#1259)
- [Midhun Somasundaran] Add label attribute for counter (#1407)
- [shri-khare] Add label attribute for ACL counter (#1430)

8 Feature Enhancements
8 MPLS Enhancements
3 Warmboot Enhancements
Meta’s SAI Spec Contributions: 2022

- Bulk Set for NH Group Object (#1356): BRCM PR in collaboration with Meta
  - Bulk get/set for Next Hop Group
- Add label attribute for counter (#1407)
  - Used to uniquely identify Counter during Warmboot
- Add label attribute for ACL counter (#1430)
  - Used to uniquely identify ACL Counter during Warmboot
- ECMP Member Capability and Configuration (#1461): BRCM PR in collaboration with Meta
  - ASIC agnostic uniform programming interface for ECMP capability query and configuration
- Worked with the SAI community for Minor releases to include urgent PRs.
Meta’s SAI Spec Contributions: 2022 (contd.)

- Near future SAI Spec contributions Meta is exploring:
  - SAI support for PHY Diagnostics: enhancing SAI spec with ~20 new SAI attrs
  - SAI support for MPLS tunneling enhancements for sFlow packets
  - SAI support for Average buffer utilization per queue
Call to Action

• Invite contributions in all areas for proposal
  • New SAI Features (Objects, Attributes, APIs…)
  • SAI PTF Tests

• Increase SAI API Coverage in ASIC Vendor SAI Adaptors

Repo: [opencomputeproject/SAl: Switch Abstraction Interface (github.com)]
Mailing list: [OCP-SAI@OCP-All.groups.io | Home]
Wiki: [Networking/SAI - OpenCompute]