EVERFLOW V2 – IPv4/IPv6 Portion 2.0

Shuotian Cheng <shuche@microsoft.com>

Motivation

Enable both IPv4 and IPv6 source and destination IP match and mirror functionality in SONiC.

Challenge

1. Hardware may or may not have the functionality of supporting IPv6;
2. Hardware may or may not have the functionality of supporting IPv4 and IPv6 in a single table;

Scenarios:

1. Legacy configurations in configuration database:
   The ACL table – "EVERFLOW" has type "MIRROR" and the behavior is maintained as only IPv4 supported;
2. acl-loader pipeline:
   acl-loader will put the rules into corresponding IPv4 and IPv6 tables depending on the table of the rules
3. CLI pipeline:
   CLI is to be extended to specify the mirror session is IPv4 or IPv6. Right now, no CLI is supported to create ACL tables or ACL rules.

Error handling:

> Both platforms' configuration database will have MIRROR and MIRRORV6 tables. No configuration conflicts will be generated since the configurations are aligned across different platforms.
> If one platform does not support IPv6, this issue will not be addressed in this part of the feature enhancement.

Design

Principal: Unify configuration on the platform level to ensure different platform could share same piece of configuration – MIRROR table stores IPv4 ACLs and MIRROV6 table stores IPv6 tables. In orchagent, only one table per type could be created and if the platform supports/requires IPv4 and IPv6 table combined, only one table will be created.

In orchagent:

1. Extend acl_table_type_t to have ACL_TABLE_MIRRORV6;
2. Query the hardware platform and get the capability of creating different types of ACL table for mirroring; store the capabilities into the state database;
3. If the platform is "Broadcom" or any other platform that supports IPv4 and IPv6 in one table, the default mirror table will support both IPv4 and IPv6. Only ONE mirror table will be created that covers all attributes. If the platform is "Mellanox" or any other platform that requires two

separate IPv4 and IPv6 tables, two mirror tables will be created with one having the type "ACL_TABLE_MIRROR" and one having the type "ACL_TABLE_MIRRORV6".

	Broadcom	**Mellanox**	
	MIRROR	**MIRROR**	**MIRRORV6**
SAI_ACL_TABLE_ATTR_FIELD_SRC_IP **SAI_ACL_TABLE_ATTR_FIELD_DST_IP**	**X**	**X**	**-**
SAI_ACL_TABLE_ATTR_FIELD_SRC_IPV6 **SAI_ACL_TABLE_ATTR_FIELD_DST_IPV6**	**X**	**-**	**X**
SAI_ACL_TABLE_ATTR_ACL_BIND_POINT_TYPE_LIST: SAI_ACL_BIND_POINT_TYPE_PORT SAI_ACL_BIND_POINT_TYPE_LAG	X	X	X
SAI_ACL_TABLE_ATTR_FIELD_ETHER_TYPE	**X**	**X**	**NOT SUPPORTED**
SAI_ACL_TABLE_ATTR_FIELD_ACL_IP_TYPE	**NOT USED**	**NOT USED**	**IPV6ANY**
SAI_ACL_TABLE_ATTR_FIELD_IP_PROTOCOL	X	X	X
SAI_ACL_TABLE_ATTR_FIELD_L4_SRC_PORT SAI_ACL_TABLE_ATTR_FIELD_L4_DST_PORT SAI_ACL_TABLE_ATTR_FIELD_ACL_RANGE_TYPE: SAI_ACL_RANGE_TYPE_L4_SRC_PORT_RANGE SAI_ACL_RANGE_TYPE_L4_DST_PORT_RANGE	X	X	X
SAI_ACL_TABLE_ATTR_FIELD_TCP_FLAGS	X	X	X
SAI_ACL_TABLE_ATTR_ACL_STAGE SAI_ACL_STAGE_INGRESS SAI_ACL_STAGE_EGRESS	X	X	X
SAI_ACL_TABLE_ATTR_FIELD_DSCP	X	X	X

<u>In state database:</u>

SWITCH capability table is added to indicate the switch capabilities:

```
"SWITCH_CAPABILITY_TABLE|switch": {
        "mirror": "TURE" | "FALSE",
        "mirrorv6": "TURE"|"FALSE",
}
```

<u>In configuration database:</u>

Two different ACL table configurations will be appearing in the database. In order for orchagent to distinguish two IPv4 and IPv6 tables are actually the same table for Broadcom, the name of two tables MUST BE EXACT THE SAME except that the IPv6 table's name has suffix "V6".

Old	New

"ACL_TABLE|EVERFLOW": {
"type": "MIRROR",
…
}
	"ACL_TABLE|EVERFLOW": {
	"type": "MIRROR",
	…
	},
	"ACL_TABLE|EVERFLOWV6": {
	"type": "MIRRORV6",
	…
	}

Other than the ACL_TABLE change, ACL_RULE also needs modifications. Below is the table for all supported fields for each of the tables.

	MIRROR	MIRRORV6
IN_PORTS	X	X
SRC_IP	**X**	
DST_IP	**X**	
SRC_IPV6		**X**
DST_IPV6		**X**
ETHER_TYPE (mandatory)	**X** **0x0800 for IPv4** **cannot be 0x86DD** **other values**	**X** **0x86DD for IPv6 only**
IP_PROTOCOL	X	X
TCP_FLAGS	X	X
DSCP	X	X
L4_SRC_PORT	X	X
L4_DST_PORT	X	X
L4_SRC_PORT_RANGE	X	X
L4_DST_PORT_RANGE	X	X

Flow Diagram

Example:

```
"ACL_TABLE": {
        "EVERFLOW": {
                "type": "MIRROR",
                "policy_desc": "EVERFLOW",
                "ports": [
                        "PortChannel0001",
                        "PortChannel0002",
                        …
                ]
        },
        “EVERFLOWV6”: {
                “type”: “MIRRORV6”,
                “policy_desc”: “EVERFLOWV6”,
                “ports”: [
                        “PortChannel0001”,
                        “PortChannel0002”,
                        …
                ]
        }
},
“ACL_RULE”: {
        “EVERFLOW|RULE_1”: {
        },
        “EVERFLOWV6|RULE_1”: {
        }
},
…
```

The above configuration will be read by orchagent will it is inserted into the database. Orchagent will determine whether to separately creates two different ACL tables or combine them together.

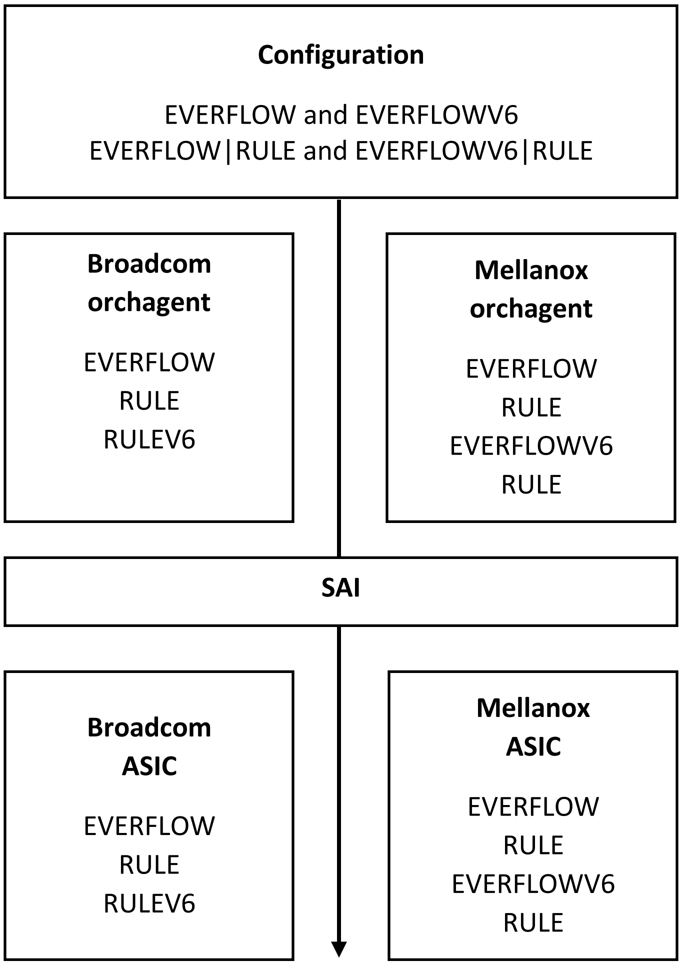
For Broadcom, whatever table will support both IPv4 and IPv6 at the same time. If the type is MIRROR, the table will be created in orchagent with exact name as it is. If the type is MIRRORV6, orchagent will check the name of the table having suffix “V6” and check if the corresponding table with MIRROR type is already created – if yes, orchagent will skip the current table; otherwise it will create a corresponding table with MIRROR type that supports both IPv4 and IPv6. For all rules that are associated with IPv6 table (table name having suffix “V6”), the rules are to be inserted in the general table.

Potential issue:

1. If we have only MIRROR table, not MIRRORV6 table in the configuration database, can we still push rules associated with IPv6 table name?
2. If we have only MIRRORV6 table, not MIRROR table in the configuration database, can we still push IPv4 rules associated with general table name?

For Mellanox, different tables will be created separately. If only one table exists in configuration database, Mellanox platform will not support IPv4 and IPv6 functionalities at the same time. Thus, the above potential issues for Broadcom will not appear here.

Once orchagent gets all the configurations from the database, it will create either one or two tables logically in the software and then call corresponding SAI APIs to create either one or two physical tables in the ASIC.


Configuration
EVERFLOW and EVERFLOWV6
EVERFLOW|RULE and EVERFLOWV6|RULE
Broadcom orchagent
EVERFLOW
RULE
RULEV6
Mellanox orchagent
EVERFLOW
RULE
EVERFLOWV6
RULE
SAI
Broadcom ASIC
EVERFLOW
RULE
RULEV6
Mellanox ASIC
EVERFLOW
RULE
EVERFLOWV6
RULE