Appendix – Partitioning

Table of contents

Partitioning enforces isolation of the fabric. The default partition is created on all managed devices. Devices that are running an SM, all switches, routers, and gateways are added to the default partition with full membership. By default, all the HCA ports are also added to the default partition with FULL membership.

Partitioning is provisioned to the Subnet Manager via the partitions.conf configuration file, which cannot be removed or manually modified.

> **Note**
>
> For those who use NVIDIA gateway systems, for proper system functionality, disable the automatic partitioning by changing the attribute gateway_port_partitioning = none in the /opt/ufm/files/conf/gv.cfg configuration. Restart UFM for the change to take effect.

If required, you can add an extension to the *partitions.conf* file that is generated by UFM. You can edit the file, */opt/ufm/files/conf/partitions.conf.user_ext,* and the content of this extension file will be added to the *partitions.conf* file. Files synchronization is done by UFM on every logical model change. However, it can also be triggered manually by running the */opt/ufm/scripts/sync_partitions_conf.sh* script. The script validates and merges the */opt/ufm/files/conf/partitions.conf.user_ext* file into the */opt/ufm/files/conf/opensm/partitions.conf* file and starts the heavy sweep on the Subnet Manager.

> **Note**
>
> The maximum length of the line in the partitions.conf file is 4096 characters. However, to enable long PKeys, it is possible to split the pkey membership to multiple lines:
>
> IOPartition=0x4, ipoib, sl=0, defmember=full : <port-guid1> , <port-guid2> ;

```
IOPartition=0x4, ipoib, sl=0, defmember=full : <port-guid3> , <port-
guid4> ;
```

The *partitions.conf.user_ext* uses the same format as the *partitions.conf* file. See SM Partitions.conf File Format for the format of the *partitions.conf* file.

For example, to add server ports to PKey 4:

```
IOPartition=0x4, ipoib, sl=0, defmember=full : 0x8f10001072a41;
```

SM Partitions.conf File Format

This appendix presents the content and format of the SM *partitions.conf* file.

```
OpenSM Partition configuration
==============================

The default partition will be created by OpenSM unconditionally even
when partition configuration file does not exist or cannot be accessed.

The default partition has P_Key value 0x7fff. OpenSM's port will always
have full membership in default partition. All other end ports will have
full membership if the partition configuration file is not found or cannot
be accessed, or limited membership if the file exists and can be accessed
but there is no rule for the Default partition.

Effectively, this amounts to the same as if one of the following rules
below appear in the partition configuration file:
In the case of no rule for the Default partition:
Default=0x7fff : ALL=limited, SELF=full ;
In the case of no partition configuration file or file cannot be accessed:
Default=0x7fff : ALL=full ;


File Format
```

```
===========

Comments:

Line content followed after \'#\' character is comment and ignored by
parser.

General file format:

<Partition Definition>:[<newline>]<Partition Properties>;

Partition Definition:
[PartitionName][=PKey][,ipoib_bc_flags][,defmember=full|limited]

PartitionName - string, will be used with logging. When omitted
empty string will be used.
PKey - P_Key value for this partition. Only low 15 bits will
be used. When omitted will be autogenerated.
ipoib_bc_flags - used to indicate/specify IPoIB capability of this partition.

defmember=full|limited - specifies default membership for port guid
list. Default is limited.

ipoib_bc_flags:
ipoib_flag|[mgroup_flag]*

ipoib_flag - indicates that this partition may be used for IPoIB, as
a result the IPoIB broadcast group will be created with
the flags given, if any.

Partition Properties:
[<Port list>|<MCast Group>]* | <Port list>

Port list:
<Port Specifier>[,<Port Specifier>]
```

Port Specifier:
<PortGUID>[=[full|limited]]

PortGUID - GUID of partition member EndPort. Hexadecimal
numbers should start from 0x, decimal numbers
are accepted too.
full or limited - indicates full or limited membership for this
port. When omitted (or unrecognized) limited
membership is assumed.

MCast Group:
mgid=gid[,mgroup_flag]*<newline>

- gid specified is verified to be a Multicast address
IP groups are verified to match the rate and mtu of the
broadcast group. The P_Key bits of the mgid for IP
groups are verified to either match the P_Key specified
in by "Partition Definition" or if they are 0x0000 the
P_Key will be copied into those bits.

mgroup_flag:
rate=<val> - specifies rate for this MC group
(default is 3 (10GBps))
mtu=<val> - specifies MTU for this MC group
(default is 4 (2048))
sl=<val> - specifies SL for this MC group
(default is 0)
scope=<val> - specifies scope for this MC group
(default is 2 (link local)). Multiple scope settings
are permitted for a partition.
NOTE: This overwrites the scope nibble of the specified
mgid. Furthermore specifying multiple scope
settings will result in multiple MC groups
being created.
qkey=<val> - specifies the Q_Key for this MC group
(default: 0x0b1b for IP groups, 0 for other groups)

WARNING: changing this for the broadcast group may
break IPoIB on client nodes!!!
tclass=<val> - specifies tclass for this MC group
(default is 0)
FlowLabel=<val> - specifies FlowLabel for this MC group
(default is 0)

newline: '\n'

Note that values for rate, mtu, and scope, for both partitions and multicast groups, should be specified as defined in the IBTA specification (for example, mtu=4 for 2048).

There are several useful keywords for PortGUID definition:

- 'ALL' means all end ports in this subnet.
- 'ALL_CAS' means all Channel Adapter end ports in this subnet.
- 'ALL_SWITCHES' means all Switch end ports in this subnet.
- 'ALL_ROUTERS' means all Router end ports in this subnet.
- 'SELF' means subnet manager's port.

Empty list means no ports in this partition.

Notes:

White space is permitted between delimiters ('=', ',', ':', ';').

PartitionName does not need to be unique, PKey does need to be unique.
If PKey is repeated then those partition configurations will be merged
and first PartitionName will be used (see also next note).

It is possible to split partition configuration in more than one

definition, but then PKey should be explicitly specified (otherwise
different PKey values will be generated for those definitions).

Examples:

Default=0x7fff : ALL, SELF=full ;
Default=0x7fff : ALL, ALL_SWITCHES=full, SELF=full ;

NewPartition , ipoib : 0x123456=full, 0x3456789034=limited, 0x2134af2306 ;

YetAnotherOne = 0x300 : SELF=full ;
YetAnotherOne = 0x300 : ALL=limited ;

ShareIO = 0x80 , defmember=full : 0x123451, 0x123452;
# 0x123453, 0x123454 will be limited
ShareIO = 0x80 : 0x123453, 0x123454, 0x123455=full;
# 0x123456, 0x123457 will be limited
ShareIO = 0x80 : defmember=limited : 0x123456, 0x123457, 0x123458=full;
ShareIO = 0x80 , defmember=full : 0x123459, 0x12345a;
ShareIO = 0x80 , defmember=full : 0x12345b, 0x12345c=limited, 0x12345d;

# multicast groups added to default
Default=0x7fff,ipoib:
mgid=ff12:401b::0707,sl=1 # random IPv4 group
mgid=ff12:601b::16 # MLDv2-capable routers
mgid=ff12:401b::16 # IGMP
mgid=ff12:601b::2 # All routers
mgid=ff12::1,sl=1,Q_Key=0xDEADBEEF,rate=3,mtu=2 # random group
ALL=full;

Note:

The following rule is equivalent to how OpenSM used to run prior to the partition manager:

Default=0x7fff,ipoib:ALL=full;