

# IPv6 over IB Encapsulation Unicast and Multicast

**draft-hkchu-ipoib-ipv6oib-00.txt**

**Jerry Chu**

**jerry.chu@eng.sun.com**

**51<sup>th</sup> IETF, London**

---

# Agenda

- What delimits an IPoIB link?
- Support IP multicast on a link
- Issues with IB 1.0 spec. on multicast
- Link layer address for IPoIB
- Q & A

# What delimits an IPoIB link?

- IP link (RFC2460): a communication facility or medium over which nodes can communicate at the LINK layer
- IB subnet?
- IB partition?
  - across IB routers? (Is P\_Key preserved?)
- IB network provides full connectivity ⇒ any set of nodes (subject to partitioning rule) can form a link

# How to support IP multicast (incl. broadcast) on a link?

- Map IP multicast to link multicast
- Standardize a fixed mapping to enable interoperability
- Link multicast must have a scope matching the link boundary (*link-local* scope)
- IP multicast routing works independently of link type

# IB multicast

- Layer-3 ⇨ MGID
  - IPv6 multicast address
  - Multicast routing not defined yet.
- Layer-2 ⇨ MLID
  - used by switches to program fwding tables
  - MC GroupRecord defines a link-local multicast group (within an IB subnet)
  - contains MGID, Q\_Key, MLID, MTU, TClass, P\_Key, FlowLabel, HopLimit

# How to map MGID to MLID?

- Determined by a SM-managed per-subnet MCGroupRecord table
- Many-to-many associations create inconsistencies and ambiguities
  - MGID: multiple MCGroupRecords w/ different P\_Key or Q\_Key or MTU or TClass..., which record to use?
  - can MLID be shared? Must define and perform multicast validity check at receiving QPs.
  - First must define what a link IB multicast group really is!

# IP to IB multicast mapping

- IPv6 multicast address  $\Rightarrow$  MGID
  - but what about scope bits?
- An IP link is delimited by
  - an IB partition?
  - P\_Key + Q\_Key?
  - a site-local scope with all nodes in it?
- Within an IB subnet
  - use P\_Key (and Q\_Key?) as link-id
  - lookup MCGroupRecord with matching MGID and link-id
  - rest (MTU, FlowLabel...) are link attributes

# IPoIB link layer addresses

- Only include attributes required to
  - uniquely identify an IP endpoint
  - to interoperate with other nodes⇒ GUID (or GID) and QPN!
- What else is really needed?
- Leave the rest to SM



# Summary

- Must define the boundary of an IPoIB link first
- IB multicast needs clarification
- Keep link layer address simple
- The rest is easy
- Q & A