

Draft Minutes

T11.3 FC-BB-5 ad hoc work group meeting

9 November 2007 - 9 AM to 5 PM PST

Las Vegas NV

The FC-BB-5 ad hoc work group of the INCITS Fibre Channel Protocols T11.3 Task Group held a meeting at Las Vegas NV on 9 November 2007, hosted by Brocade, Cisco, and Hitachi GST. Attendance was 34 people from 20 companies and is tabulated at the end of this document.

Minutes were taken by Bob Nixon (bob.nixon@emulex.com). Please report any corrections by email to the T11.3 reflector at T11_3@mail.T11.org.

1 Opening remarks

1.1 Introductions

Richard DiPasquale opened the meeting Friday, 9 November 2007 at 9:05 AM PST. He thanked our host organization, Brocade, Cisco, and Hitachi GST, and led a round of introductions.

2 Meeting Policy

2.1 Attendance and Membership

Richard DiPasquale stated that all persons present are considered members of this meeting and may vote on questions, limited to one vote per company present. He explained that attendance is recorded electronically at www.t11.org/att, and explained the procedure. Attendance at this meeting does not count toward attendance at the plenaries of T11 and its task groups (i.e., being here will not get you out or keep you out of membership jeopardy).

Richard DiPasquale advised that although T11 does not limit participation in the activities of its work groups to representatives of T11 member organizations, T11 encourages organizations that are regularly represented at T11 work group activities to become members of T11.

2.2 Patents

Richard DiPasquale indicated that among the rules and policies under which this working group operates are the INCITS intellectual property policies. He requested persons wishing to make statements or ask questions relevant to this policy not to pursue discussion at this work group meeting, but instead to do so with an officer of T11.3 or T11.

The INCITS intellectual property policies are specified in subclause 8.4 of RD-1, "Policies and Guidelines", available at

<http://www.incits.org/rd1/rd1.htm#p8-4>

A summary of the INCITS intellectual property policies as applicable to the work of this meeting is given in document T11/05-385v1, which is available on www.t11.org.

2.3 Antitrust

Richard DiPasquale indicated that among the rules and policies under which this working group operates are the INCITS Antitrust Guidelines. Any member of the meeting is responsible for objecting if he believes discussion in the meeting violates those guidelines. As examples, there should never be discussion of the following topics at any INCITS or INCITS subgroup meeting:

- Any company's prices or pricing policies;
- Specific R&D, sales and marketing plans;
- Any company's confidential product, product development or production strategies;
- Whether certain suppliers or customers will be served;
- Prices paid to input sources; or
- Complaints about individual firms or other actions that might tend to hinder a competitor in any market.

If such discussion is not immediately terminated, it is the chairperson's responsibility to terminate the meeting. The INCITS Antitrust Guidelines are available at

<http://www.incits.org/inatrust.htm>

3 Approval of Agenda

An agenda for the FC-BB-5 ad hoc work group meeting 9 November 2007 has been distributed via email. The order of presentation was adjusted slightly.

Ralph Weber (ENDL) moved and Landon Noll (NeoScale) seconded to accept the emailed agenda as modified as the agenda for this meeting. Approved unanimously.

4 Scheduled Business

4.1 Discovery

T11/07-572v1

DeSanti, et al

This document presented joint work and agreement among several member companies on a method for discovery of multiple FCFs (FCoE routers) in an Ethernet from a single FCoE NIC. It therefore provides a way of resolving one of the requirements posed for this project: multipathing/failover opportunities through a single NIC.

It is a refinement of work presented at the prior meeting by Cisco. Among the improvements included

- a) more completely elaborated process for both FCF-FCF discovery and ENode-FCF discovery;
- b) discovery of only jumbo frame capable paths;
- c) avoidance of jumbo frame multicasts (depends on symmetric jumbo frame capability);
- d) protection against interconnect of instances of a Fabric that have identical identity, but are intended to be independent;
- e) several specific discovery parameters supporting possibly useful optional behaviors;
- f) ability to detect a failed VN_Port among one or more functional VN_Port (i.e., a failed virtual link); and
- g) stabilizes the routing tables in learning switches.

Open issues included

- a) timeouts for determining completion of discovery processes;
- b) size of the padding to force advertisements to be maximum necessary FC frame size, realizing that the maximum necessary FC frame size may be either greater or less than 2112 bytes;

- c) need/method for an end-of-list indication, apart from the list length;
- d) detailed behaviors of the discovery methods with respect to the requirements for and effects of the recommended descriptors, especially with respect to advertisement priority;
- e) need for completely deterministic discovery behavior, especially with respect to advertisement priority;
- f) significance of a zero FC-MAP descriptor value;
- g) resolution of the use of, or need for, the FC-MAP descriptor; and
- h) behavior for change of FCF priority (it was agreed there are probably sufficient tools, but no specified behavior).

Because of the importance of this presentation in establishing a critical area of general technical accord, it was by consensus granted extra presentation time.

Claudio DeSanti (Cisco) moved and Bob Snively (Brocade) seconded to accept 07-572v1 as the basis for FCF discovery in FC-BB-5, reflecting DDD and the agreements at the FC-BB-5 ad hoc work group meeting 9 November 2007. The motion passed 15 Favoring, 0 Opposing, 2 Abstaining.

Claudio DeSanti (Cisco) moved and Bob Snively (Brocade) seconded to accept that VF_Ports and VE_Ports are addressed by Universal MAC addresses belonging to the FCF.

In discussion, the restriction to "Universal" was questioned. It was described as attempting to indicate a "factory-assigned" or "burned-in" MAC address.

A friendly amendment was made to change the motion to

to accept that VF_Ports and VE_Ports do not use mapped MAC addresses.

The question was called. There was no objection to

The motion passed 11 Favoring, 1 Opposing, 4 Abstaining.

Bob Snively (Brocade) moved and Bob Nixon (Emulex) seconded to accept that VF_Port and VE_Port MAC address uniqueness is the basis for FCoE.

After discussion, the motion was withdrawn.

4.2 Avoiding Data Corruption

T11/07-546v1

Gai (Nuova)

This presentation described a way to use switch access control lists (ACLs) in combination with the FC-ID mapping method of MAC address assignment to defend against data corruption resulting from several classes of (accidental and/or deliberate) misconfigurations.

An issue was raised with the protection provided against misconfigurations during the discovery process. Responses included that the window for such misconfigurations was exceptionally narrow, and that discovery errors due to these would not lead to data corruption, since the discovery process includes further protection.

An issue was raised with the wisdom of reliance on software-assigned MAC addresses. Responses included that these were now a fact of life on Ethernet.

The need for this protection was questioned, since misconfiguration is a basic Ethernet issue, not a specifically Fibre Channel problem. Responses included that the objective was to provide the same level of protection as a native FC Fabric.

The need for this protection was questioned, since issues require unlikely combinations of events. Responses included the ability of deliberate attackers to make the combinations likely, and the exceedingly high value of the data being protected.

The presenter emphasized that any ACL-reliant scheme shall not interfere with any possible non-FCoE ACLs for the same port. Among the implications are that an FCoE ACL can NOT include unconditional rules, unless they are at the end of the ACL list.

The presenter emphasized that a MAC learning misconfiguration requires no more than one frame every minute or so to cause denial of service, since it would cause a few frames to be lost, which would in turn disrupt typical SCSI operations for more than a minute.

The presenter emphasized that learning misconfigurations could not be prevented by any Ethertype-limiting ACL. MAC learning can be compromised by any frame type. This leads to the possibility of a MAC learning misconfiguration that requires no more than one frame every minute or so, since it would cause a few frames to be lost, which would in turn disrupt typical SCSI operations for more than a minute.

The presenter recommended that Ethernet switches used in FCoE environments support the ability to dynamically adapt their ACLs by inspection of FCoE FLOGI/FDISC LS_ACC.

Positive factors included:

- a) ability to identify the type of source and destination of FCoE frames only by examination of the SA and DA, and therefore to prevent certain routing loops by appropriate ACLs;
- b) ability to extend protection provided by discovery to the subsequent FLOGIs/FDISCs; and
- c) ability to protect against rogue-host MAC learning attacks.

Open issues raised included:

- a) lack of standard ACL behaviors for switches, especially dynamically modified ACLs;
- b) the need to enhance Ethernet switches, especially in support of dynamic ACL management; and
- c) the ability to recognize a mapped address in the presence of (mis)configurations that include more than one FC-MAP.

4.3 FCoE Addressing

T11/07-547v0

Gai (Nuova)

Because the intense interest in prior presentations caused a substantial slip in the agenda schedule, Silvano relinquished this time in order to allow other speakers to proceed.

4.4 Ethernet Direct Attached Fabrics

T11/07-608v0

Carlson (IBM)

The presentation proposed that direct communication between FCoE endpoints (i.e., without passing through an FCF) should be accepted as an additional goal of this project. It relied on the accepted discovery and login processes and suggested extensions to current processes that provide the necessary routing information and zoning behavior. The preferred alternative for zoning was proposed to be HBA-enforced. A longer-term desire was said to be switchless FCoE Fabrics.

Positive factors included:

- a) shorter data paths in large Ethernet fabrics with few FCFs.

Open issues raised included:

- a) the impacts of allowing mixed direct and FCF-routed traffic between the same two end points; and
- b) distrust of HBA-based zoning.

4.5 Preserve FCoE Ethernet ACLs

T11/07-569v0

Tsao (Intel)

This presentation observed that migration of virtual machines in native Fibre Channel Fabrics often required reversion to “open zoning” because the migration would cause traffic that was blocked by established zoning rules. It proposed that this project should accept the goal of avoiding this defect in FCoE. It followed to propose that FCFs attempt to make the associations of WWPN-FC_ID-MAC persist across changes in physical connectivity. One proposed prerequisite was abandoning the hierarchical structure of FC_IDs so that a group of FCFs implementing a single Fabric could share a large pool of FC_IDs.

It was claimed that this may be more easily done in FCoE than in native FC because the VF_Port assignment of a migrating VN_Port could be retained. The migration of ACLs was not discussed.

4.6 Server Provided MAC addresses

T11/07-591v1

Hufferd (Brocade)

This presentation gave a detailed introduction to the constraints presumed by the server-assigned MAC address approach, and that these constraints are presumed to be required of any FCoE approach:

- a) The FCoE adapter should be seen at the Host system as a normal FC adapter;
- b) No changes to the way servers define & use MAC addresses
- c) FCoE should not change the way Servers (and customers) manage MAC addresses today

The presentation presumed and did not elaborate on the discovery model presented and approved earlier in this meeting as T11/07-572v1.

The presentation then summarized several single OS and multiple-OS system architectures, and in each, how the existing “burned-in” MAC addresses would be used for FCoE within the constraints presumed. His cases included architectures with software implementations of FCoE.

The presentation made special note of a problem of quiescing SCSI activity when a software implementation of FCoE in a virtual OS is used: only the hypervisor knows the need to quiesce, but only the guest OS knows what SCSI IO is in progress.

The presentation further observed that the architecture(s) of SRIOV in support of virtual machine systems is not yet well-known and so it is unclear what challenges it raises. The hope is that if the issue is resolved for NICs and HBAs, it will be similarly solved for FCoE, IF the existing FC HBA interfaces and drivers can be carried over (i.e., constraint a above).

4.7 Pathing in the FCoE adapter

T11/07-655v0

Hufferd (Brocade)

This presentation introduced the issues raised by the need for large numbers of Fibre Channel logins from a multiple virtual OS system. The problem long ago scaled past economical resolution by multiple physical NICs. It is now resolved in native FC by assignment of multiple N_Port_IDs (NPIV). NPIV implies frame routing within a Fibre Channel N_Port (HW or SW doesn't matter). It then stepped through extending that model to FCoE. It concluded

that server-assigned MAC addresses posed no serious obstacles to that extension, given several presumptions about the exact nature of the server assignment methodology. It further concluded that it had no disadvantages, and probably some advantages (e.g., HW does not need to support a large number of MAC addresses), for FCoE HBA design versus FC_ID mapped MAC addresses.

An issue was raised that SRIOV leads to the requirement for many MAC addresses for any Ethernet attachment, whether FCoE is involved or not. This might reduce one of the posed advantages of the server assigned address method (i.e., fewer MAC addresses), especially for software FCoE. The responses included that SRIOV designs are insufficiently well-known to project whether and how it may impact either approach.

4.8 Ethernet Switch ACLs for FCoE

T11/07-656v0

Hufferd (Brocade)

This presentation described a way to use switch access control lists (ACLs) in combination with server assigned MAC addresses to defend against data corruption resulting from several classes of (accidental and/or deliberate) misconfigurations. It posed serious concern with reliance on dynamic ACL management. It concluded that ACL usage would be equally effective for either MAC address assignment method, but the server-assigned method required less ACL complexity and did not require dynamic ACL management.

The presentation added that dynamic MAC address assignment may lead to difficulties versus 802.1x compliance.

4.9 FCoE Connections (Updated)

T11/07-574v1

DeSanti (Cisco)

This presentation was deferred due to lack of time.

5 Unscheduled Business

No unscheduled business was presented.

6 Next meeting

Time has been scheduled at the next plenary meeting, in Lake Mary FL in December 2007. No time will be requested for an interim meeting (e.g., with T10) in January.

7 Review of Action Items

No new actions were assigned at this meeting.

8 Adjournment

It was moved by Claudio DeSanti (Cisco) and seconded by Roger Hathorn (IBM) to adjourn. Approved unanimously.

The meeting was adjourned at 4:54 PM PST on 9 November 2007.

9 Attendance

Organization	Representative
54TH STREET SYSTEMS	John Hayes
ALCATEL-LUCENT	Richard DiPasquale
BROADCOM	Pat Thaler
BROADCOM	Uri Elzur
BROCADE	Anoop Ghanwani
BROCADE	David Peterson
BROCADE	Glenn Wenig
BROCADE	John Hufferd
BROCADE	Robert Snively
BROCADE	Scott Kipp
BROCADE	Steven L. Wilson
BROCADE	Suresh Vobbilisetty
CISCO SYSTEMS	Claudio DeSanti
EMC	David Black
EMC	Stuart Miniman
EMULEX	Bob Nixon
EMULEX	Tuan Nguyen
ENDL TEXAS	Ralph Weber
FINISAR	David Freeman
HEWLETT-PACKARD	Jerry Aguren
HEWLETT-PACKARD	Robert Elliott
HEWLETT PACKARD	Don Fraser
HEWLETT PACKARD	Krishna Babu Puttagunta
HITACHI GLOBAL STORAGE TECHNOLOGIES	Dan Colegrove
IBM	Roger Hathorn
IBM	Scott Carlson
INTEL	Gary Tsao
MELLANOX TECHNOLOGIES	Diego Crupnicoff
NEOSCALE SYSTEMS	Landon Noll
NUOVA SYSTEMS	Silvano Gai
PMC SIERRA	David Sheehy
QLOGIC	Craig W. Carlson
SUN MICOSYSTEMS	Ariel Hendel
VMWARE	Lawrence Lamers

