Fabric Expansion Study Group Draft minutes 03-691v0
October 7, 2003 1:00 PM – 6:00 PM
Oklahoma City, OK

Introductions

John Scheible opened the meeting at 10:20 AM. The group began with introductions. John thanked the host for the facilities

The agenda was approved as modified (03-475v1).

Approval of minutes

The minutes of the August meeting (03-605v0) were approved as written.

Approval of Agenda

The agenda was approved with one addition.

Review of Action items

- Switch vendors need to determine if the format proposed in 03-504v0 causes problems because F_CTL is moved later in the frame. If it does, then what is required to be at the head of the frame? (Closed 03-487v1 and 03-672v0)

5.1 Trace Route and Ping 03-508v0 Scott Kipp

Scott Kipp presented on how to manage an extended large fabric through the use of Trace Route and Ping operations. The desire is for the first version of this to contain this within a single fabric, not over an extended fabric. It uses a new CT request and a new Distributed Service. The question was raised whether there could be multiple ingress or egress ports. It was suggested that this indicated what the current path was, not all possible paths. This is consistent with Trace Route implementations in IP. It was pointed out that this is currently missing a timestamp. It was pointed out that this was limited by frame size, and consideration was not given to requirements for multi-frame sequences. There were questions raised about the use of distributed service vs. an SW_ILS. The switch would initiate an asynchronous CT command to the requestor to return the results of the Trace Route. The presentation does not yet cover error reporting/recovery yet. With an IP network there will not be any information about the IP network path. For B_Port operation, there are some issues about what is reported. Getting information about the tunneling that is occurring could be useful. One approach would be to flag B_Port connections to allow management to find information about the tunneling. PING is a subset of the TR operation. This could use the ECHO ILS operation. Scott will bring in additional textual descriptions in future meetings.

5.2 Classless FC-Ids 03-526v0 Claudio DeSanti

This proposal addresses the connection of separate fabrics to allow communication between devices on the separate fabrics. If the addressing space overlaps, then a NAT solution is required. Doing this without extending the header limits the solution to two gateways and the number of edge devices that can be connected. There was disagreement about whether the "disadvantages" were really disadvantages. Doing this with extended headers, requires a more complex backbone structure. Unique addresses can be achieved by extending the address space, which can be achieved by extending the header. Claudio's contention is that this requires an L3 layer, and the definition of the frame format is only a small part of the definition required. The point of Claudio's presentation was to explore the case where unique addresses are achieved without extending the address space. This hinges on whether a SAN with 239 switches is large enough. If it is then we need a Hierarchical Control Protocol.

Claudio's proposal was to improve the assignment of the address space using a variable length address assignment to switches. Switches may share a Fabric address, allowing address aggregation, and subdividing addresses based upon individual switch sizes. There was a request that the addressing would not be administratively assigned on each switch. Claudio quickly went through his last several slides, and will return with more details for further discussion.

5.3 Autonomous Regions & address trans 03-538v0 Alex Goral

The goal of this presentation was to flesh out what needs to be done to make Autonomous Regions functional in FC-SW-4. This approach is backwards compatible with legacy switches. It can use the existing LSRs without the requirements of summary LSRs. The functionality is absorbed by the Boarder Switch. Issues were raised as to whether there could be redundant AR0s (AR0 and AR0')? What Alex was proposing was that their redundancy was a "Hot Standby" implementation. The topology is useable for local/remote topologies, either through a WNA/LAN or separate departments. Without address translation this does not address the need for more address space. There was concern that without a mesh capability in the backbone then this approach was not feasible.

5.6 Big Fabrics, Virtual Fabrics and Gelatin 03-660v0 Joe Pelissier

Joe started by addressing the possible problems and how existing proposals address each of the problems. The two architectures that he addressed were ability to route between independent fabrics (Fabric Routing); and ability of a switch to host multiple independent fabrics (Virtual Fabrics). He then listed what required functionality was provided by each of the architectures. He followed this by listing the characteristics of each architecture. There was discussion but basic agreement with these lists. Joe then presented the relationship between Virtual fabrics and Fabric routing and how a single switch could support both of these. He included a diagram of a series of switches which each had varying degrees of implementing Virtual Fabrics and Fabric Routing and how

frames would move through the fabric. The "Gelatin" portion of the presentation was what needs to be standardized as compared to the individual implementations (e.g., VSAN, nDP, etc), which are built upon the standardized features. This was clarified that the standardization was of the tags that are defined even though the different implementations may use the tags differently.

5.4 Unified header for Fabric Extension 03-487v1 Silvano Gai

Silvano suggested that this proposal would solve most of the problems on the table if we go with an extended header. There is a new frame format required. He raised an issue that an encapsulation header required 3 words since CS_CTL required the proper bit(s) be set in the F_CTL field. Another issue was that an Inter-Fabric Header required that OX_ID be passed in the header. He went through a process of combining the Tagging header, Encapsulation Header, Inter-Fabric Header, and the original FC Header to get to a new header that contained the required fields in the locations that each function required them, and minimized by not the number of words required. This header is 5 words long, and includes all of the header information that is needed, but requires that both Encapsulation and inter-Fabric information are passed even if one of these is not required. The question was raised if only one of the two additional functions (Virtual fabric/Fabric Routing), what does it do with the fields that are not required for that functionality. Clarification was made that there is first routing to a Gateway, and then a Gateway may have multiple fabrics that it is routing to. Silvano's moldel uses only legacy switches and gateways. This did not address the model of a fabric that does not have a gateway. There was significant discussion of how the sets of requirements meshed, and there was finally agreement that we needed to start with a requirements document.

5.7 Inter- and Intra-Fabric Header proposal 03-672v0 Bob Snively

Bob again began with a set of requirements. There was disagreement about his first requirement, which was the requirement for Intra-Fabric routing for non-overlapping independent fabrics. His goal was to provide a header that meets the needs of each of the requirements and is still able to pass through a legacy fabric. Bob then showed how his header would be used in Intra Fabric routing and also in Inter Fabric routing. His proposal required that administratively there is not an overlap of VLAN tags if there is communication between the VLANs. If you are doing Fabric Routing and Virtual Fabric, you will require two of these headers. If you are only doing one of these, you will only have one header. The header is the same whether you are doing Fabric Routing or Virtual Fabric. There was a question raised about whether 2 different R_CTLs could be used; however, Bob's preference was to indicate the difference between Inter/Intra was based on a bit in the address. When routing a IIFR frame through a legacy switch/fabric, an Encapsulation header is added which is identical the current FC header. Proprietary headers can be added in front of this encapsulation header.

5.5 ESS Extention for Data Field Size 03-643v0 Scott Kipp

Scott presented a need due to encapsulated headers to communicate between switches what the Buffer-toBuffer Receive data field size is that should be reported in a FLOGI ACC response. The issue was raised as to how a legacy switch would deal with this issue since it would not communicate this. It was pointed out that a firmware upgrade on a Legacy switch would allow the use of this feature. There was discussion about the Buffer-toBuffer Receive data field size in the ELP, and it was noted that this was Switch-to-Switch, and this needed to be determined on a fabric basis. A higher value than what the standard allows can be reported. There was agreement on the direction of this proposal.

Action Items

Old Action Items

None

New Action Items

- Bob Nixon to send out a call on the reflector for requirements/problems that this group is attempting to address.
- Bob Nixon to gather a list of the requirements/problems that this group is trying to address.

Next meetings

The next meeting will be in December at Paradise Island, Bahamas during the next T11 plenary week, and 5 hours are requested.

Meeting adjourned at 5:45 PM.

Project status:

The project is still in the Study Group stage.

Attendance

Name	Company
Roy Elsbernd	AGILENT
Claudio DeSanti	ANDIAMO SYSTEMS, INC.
Tony DiCenzo	BROCADE COMMUNICATIONS
William R. Martin	BROCADE COMMUNICATIONS
Edward McClanahan	BROCADE COMMUNICATIONS
Steven L. Wilson	BROCADE COMMUNICATIONS
Chris Janz	CIENA
Silvano Gai	CISCO SYSTEMS, INC.
Stephen Cortese	CNT
Harry V. Paul	CNT CORPORATION
Bret Ketchum	COMPUTER NETWORK TECHNOLOGY

Name	Company
John Tyndall	CROSSROADS
David Black	EMC
Ron Stern	EMC
Bob Nixon	EMULEX
Ralph Weber	ENDL TEXAS
Mike Fitzpatrick	FUJITSU COMPUTER PRODUCTS
Vinod Bhat	HP
John Scheible	IBM AUSTIN
George Penokie	IBM TIVOLI SYSTEMS
Robert Sheffield	Intel
Neil Wanamaker	JNI
Alex Goral	LIGHTSAND COMMUNICATIONS
John Lohmeyer	LSI LOGIC
Louis Odenwald	LSI LOGIC
Peter Rivera	LSI LOGIC
Charles Monia	MCDATA
Michael O'Donnell	MCDATA
Joe Pelissier	McData
Gary Warden	NGS
Ed McGlaughlin	QLOGIC
Craig Carlson	QLOGIC CORP.
Tim Sheehan	SAN DIAL
James Coomes	SEAGATE TECHNOLOGY
Robert Kembel	SOLUTION TECHNOLOGY
Matt Gaffney	STORAGETEK
Steve Sletten	SUN MICROSYSTEMS
Horst Truestedt	TRUE FOCUS, INC
Ken Hirata	VIXEL