Switch Fabric Benchmarking Framework

November 18, 2002

Revision 1.0

Editor(s):

Chris Bergen, ZettaCom., chris@zettacom.com

Ali Poursepanj, Agere Systems, alip@agere.com

For additional information contact:
The Network Processing Forum, 39355 California Street,
Suite 307, Fremont, CA 94538
+1 510 608-5990 phone ✦ info@npforum.org

Table of Contents

1 Revision History

Revision	Date	Reason for Changes
1.1	11/18/2002	Reformatted the document using the official NPF template
1.1	11/21/2002	Fixed the Figure 1 diagram. The content has not changed

2 About this Document

This document presents the framework of activity for the NP Forum fabric benchmarking task group. This work is part of the Network Processing Forum benchmark workgroup activity.

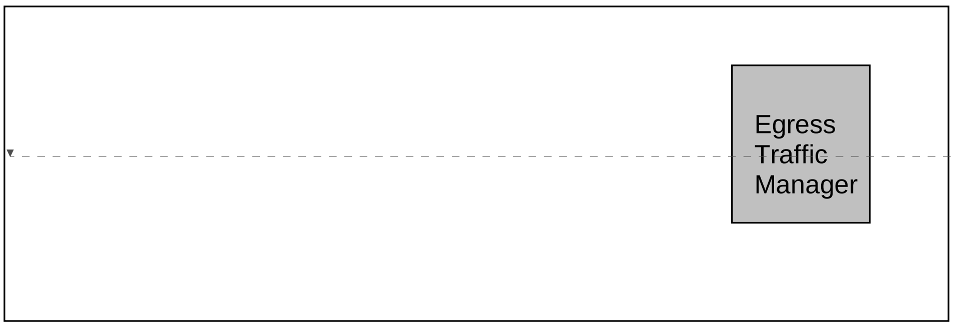
3 Preface

The fabric benchmarking activities of the NPF started in April of 2001, when a number of NPF members requested a meeting and discussed the switch fabric benchmarks. As a result, a task group was formed to develop a methodology and set of metrics for the benchmarking of switch fabrics. This document contains the framework for a general approach to fabric benchmarking, within which a benchmarking specification will be developed.

4 Introduction

The switch fabric is the central building block in network switch systems and routers. Recently, the Common Switch Interface forum (CSIX) was formed in order to provide a common interface for the industry to follow. This forum shows the importance of the switch fabric and the interest of suppliers of these components.

Logically, a typical fabric has three main sections. The first section is the ingress data path interface (CSIX). It connects ingress line cards to the switching section. The second section is the actual switching function that connects any ingress ports to any egress ports. The third section consists of the egress data path (CSIX) interface on the egress line card, which concentrates traffic from the switching function and forwards the traffic to an egress line card's network processing subsystem on the same line card. The fabric is treated as a black box during the development of the benchmarking specification (Figure 1).



Deleted: <sp><sp>**<sp>**<sp>¶
¶

The performance of a switch fabric depends on several factors, including the fabric internal architecture, fabric configuration, application and traffic characteristics, and the data flow mechanism between the traffic manager and the fabric. The user of a switch fabric, the system designer, is interested in understanding the performance of the fabric under various conditions. Currently, there does not exist a standard means by which switch fabric vendors specify the performance of their switch fabrics. Consequently, it is is not feasible for a fabric customer to directly compare performance data provided by a given fabric vendor with a fabric from another vendor. (Switch fabric features, such as the number of ports and number of class of service levels supported by the fabric, are relatively unambiguous.)

5 Guidelines for the Development of a Fabric Benchmarking Specification:

1- The Fabric Benchmarking Task Group (FBTG) is only providing a standard method and metrics for measuring performance. The FBTG is not creating a “Consumer's Guide” to switch fabrics, nor is it classifying or categorizing fabrics (for example, into such categories as "enterprise" or "core" router fabrics, by application or by architecture).

2- The FBTG is not performing the actual tests. It is assumed that the audience of the Fabric Benchmarking Specification (e.g., a given switch fabric customer) will request the actual performance data from the fabric vendor, either directly or via a third party, and the fabric vendor (or third party) will perform the measurements and provide the information.

3- The FBTG is defining a set of test methods for determining the performance of a switch fabric under various scenarios. For example, if the fabric vendor states that their fabric TM has a given performance under bursty traffic, the FBTG provides the exact traffic arrival pattern (or a set of possible traffic arrival patterns) that should be used to represent bursty traffic. As another example, if the fabric vendor states that their fabric has the same performance regardless of the number of classes of service being used, then the FBTG provides the method by which the traffic should be distributed among different classes in order to determine the performance under the condition of multiple classes of service.

4- The FBTG is confining its scope of work to switch fabrics that use a CSIX interface in the first phase of work. It is possible to extend the benchmarking methods to other interfaces, such as the Streaming Interface, once such a standard is complete and adopted as a fabric interface, at which time the FBTG will review this document to determine its relevance and make the appropriate updates The FBTG is specifically not attempting to specify tests for a non-standard interface, such as SPI4, phase 2, because it is not possible for to do so.

5- The FBTG will need to make some decisions regarding the test environment, such as assuming that the NPU has a pragmatic value for response to backpressure (flow control) messages (for example, a small value or zero), because in most cases, the test signals will be generated from test equipment and simulation environments. Even though it might not represent a real-world scenario, it still provides valid information for comparison purposes. Also, if a given fabric customer wants the switch fabric vendor to provide performance criteria based on a different assumption, such as changing the external device’s response latency to 10 cycles, for example, that fabric customer is able to do so in the course of his business negotiations with the fabric vendor.

6- The FBTG focuses on and limits its work to that which is reasonable to accomplish. For example, since not all fabrics implement a similar queuing structure, it is not reasonable to make any assumptions about what types of thresholds are available to be configured, and as such it is necessary to leave that to the fabric vendor to decide. Likewise, we treat the fabric as a black box, meaning that we do not specify anything in the tests that require knowledge of or visibility into the internal implementation of the fabric or the real-time monitoring of internal states. The I/O of the black box is the standard fabric interface (i.e., CSIX).

7- The FBTG will provide performance metrics for each individual traffic scenario rather than an average over several scenarios. The FBTG does not determine which measurement parameters are important. That is left to the fabric customer to decide, based on their application. (For example, we do not decide if throughput is more important than latency, or if jitter should be include or not. Instead, we define *how* to measure throughput, how to measure latency, and how to measure jitter.) The fabric customer can decide if they care to have such information included in the performance information. In other words, we do not spend time on "application-dependent" properties. Likewise, issues such as bandwidth loss due to segmentation into fixed-size cells (i.e., "cell tax") are beyond the scope of our task. (If that is an issue for the fabric vendor, the fabric vendor can ask the cell size used and do an assessment of the impact of the given cell size.)

8- The FBTG does not need to discuss or itemize technical implementation details that may impact the performance of the fabric, such as the possibility of Head-Of-Line Blocking and the conditions under which it may occur, because that will be reflected in the tests. (We may, however, want to include a traffic pattern that would result in the worst possible HOLB, for example, and allow the fabric customer to choose that scenario for evaluation.) The FBTG limits its efforts to scenarios that are relevant to provide a benchmark comparison of switch fabrics and does not attempt to provide a complete characterization of the fabrics.

9- The FBTG does not need to describe how to implement the tests, since it is up to the fabric vendor or third party to figure this out. For example, there is no need for a "standard" connector, since the fabric vendor may even provide performance results based on simulation, or from traffic generated from an FPGA in a reference system using actual silicon. (Nonetheless, there might be some secondary benefit in eventually specifying one.)

APPENDIX A LIST OF COMPANIES BELONGING TO NPF DURING APPROVAL PROCESS

Actel
Agere Systems
Alcatel
Altera
AMCC
Analog Devices
Avici
Azanda Network Devices
Bay Microsystems
Broadcom
Chiaro Networks
Chip Engines
Cisco
Clearspeed
Conexant
Corrent Corporation
C-Port Corporation
Cypress Semiconductor
Dune Networks
Ericsson
Erlang Technology
ETRI
EZ Chip
Fast-Chip
Firebit Ltd.
Flextronics Semiconductor
Fujitsu Microelectric America
Fulcrum Microsystems
Future Communications Software
Greenfield Networks
GSI Technology
HCL Technologies
Hi/fn
Huawei Technology Corporation
IBM
IDT
Infineon Technologies
Intel
Integrated Silicon Solution, Inc.
Internet Machines
IP Flex
IP Infusion
Kawasaki LSI
LSI Logic
LVL7
Marvell International
Memcall
Micron Technologies
Modelware, Inc.
Mosaid
NEC
NetLogic
NetOctave
Next Hop Technologies
Nokia
OKI Electronic
Paion Co., Ltd.
PetaSwitch Solutions
PMC-Sierra
Propulsion Networks
Protean Devices
QQ Technology
RadiSys
Samsung Electronics
Sandburst Corporation
Sibercore Technologies
Silicon & Software Systems
Silicon Access
Silicon Logic Engineering
SiPackets
SiSilk
Solidum
Sony Electronics
Spirent Communications
STMicroelectronics
Sun Microsystems
Tau Networks
Teja Technologies
TeraBlaze
TeraChip
TeraCross
Teradiant Networks
TranSwitch
Tundra Semiconductor
U4EA Group
Vitesse
Xelerated
Xilinx
Zagros Networks
Zettacom
ZTE