Referee Report

Paper Number:
Reviewer's Name: Jesse Bannon
Name of Paper: Characterizing Parallel Scientific Applications on Commodity Clusters: An Emperical Study of a Tapered Fat-Tree
Author(s): Edgar Leon, Ian Karlin, Abhinav Bhatele, Steven Langer, Chris Chambreau, Louis Howell, Trent D'Hooge, Matthew Leininger

Section I. Overview

A. Reader Interest

1. Which category describes this manuscript?

(x) Practice / Application / Case Study / Experience Report
( ) Research / Technology
( ) Survey / Tutorial / How-To

2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under IIIA.

( ) Very Relevant
( ) Relevant
(x) Interesting - but not very relevant
( ) Irrelevant

B. Content

1. Please explain how this manuscript advances this field of research and / or contributes something new to the literature. Please explain your answer under IIIA. Public Comments.

2. Is the manuscript technically sound? Please explain your answer under IIIA. Public Comments.

( ) Yes
(x) Appears to be - but didn't check completely
( ) Partially
( ) No

C. Presentation

1. Are the title, abstract, and keywords appropriate? Please explain your answer under IIIA. Public Comments.

(x) Yes
( ) No

2. Does the manuscript contain sufficient and appropriate references? Please explain your answer under IIIA.

(x) References are sufficient and appropriate
( ) Important references are missing; more references are needed
( ) Number of references are excessive

3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under IIIA. Public Comments.

(x) Yes
( ) Could be improved
( ) No

4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain your answer under IIIA. Public Comments.

(x) Satisfactory
( ) Could be improved
( ) Poor

5. Please rate and comment on the readability of this manuscript. Please explain your answer under IIIA.

(x) Easy to read
( ) Readable - but requires some effort to understand
( ) Difficult to read and understand
( ) Unreadable

Section II. Summary and Recommendation

A. Evaluation

Please rate the manuscript. Please explain your answer under IIIA. Public Comments.

( ) Award Quality
( ) Excellent
(x) Good
( ) Fair
( ) Poor

B. Recommendation

Please make your recommendation. Please explain your answer under IIIA. Public Comments.

( ) Accept with no changes
(x) Author should prepare a minor revision
( ) Author should prepare a major revision for a second review
( ) Reject

Section III. Detailed Comments

A. Public Comments (these will be made available to the author)

Explanation for the Recommendation

The paper makes a compelling argument for cluster computing environments to not invest in fully provisioned fat-trees for 7% savings. The organization of the argument provides insights to what typical job configurations look like in a large shared cluster and the communication required for different types of problems. These sections and their respective graphics provide a great introduction to the reader in understanding the impacts of reducing network bandwidth. Section III goes slightly out of scope with a page's worth of application descriptions. The physics aspect of each description could be significantly reduced, and instead focus on performance characteristics.

Analysis of the tapered fat-tree versus full fat-tree provided sufficient information for reproducibility and a surplus of runtime metrics related to communication. Characterizing the applications into categories based on message size made it easy for the reader to know what to expect in the performance results. The presentation of these results were simple to understand, and surprisingly had very little performance impact when removing half of the fat-tree.

Section VI provided a nice recap, but did not need its own section. The first part summarizes the findings in Section V; this should be the last subsection of Section V. Section VI A addresses the fact that GPUs and FPGAs will not have the same results as found in Section V. The information presented here is important, but should reside in the conclusion and future works section.

Summary of the Paper and Assessment

The paper presents a study of why commodity clusters do not require a fully provisioned fat-tree, and can settle with a tapered fat-tree for about 7% savings with little to no performance impact. It first describes the types of jobs ran at the Lawrence Livermore National Laboratory (LLNL), their node configurations, and communication requirements.

From observing the network link utilization, the fully provisioned fat-tree averaged less than 50% and peaked at just a little over 50%. With the tapered fat-tree, average utilization was still less than 50% and peaked just under 90%. No job(s) stressed the network enough to fully utilize the fat-tree switch, both full and tapered.

The paper later describes seven common applications ran on the cluster. This section went out of scope when describing the physics and maths involved. That aspect should have had a three sentence limit, and use this space to describe more about the application in terms of compute and communication intensity in detail.

Benchmarks describing the aggregate bandwidth based on message size shown in Fig. 6 is notably important. It shows how small message sizes are not affected in the tapered fat-tree but large message sizes are. This informs the reader that high message-sized applications are prone to performance loss on the proposed tapered fat-tree. Immediately following, the applications are categorized by their message size. 'Large' applications are expected to see performance loss based on Fig. 6.

In their performance analysis of the applications, the only slowdown was on a single application of 1-2% at large node counts when using the tapered fat-tree. This provided the evidence for their argument: the tapered fat-tree can save on cost with little-to-no performance impact. The authors subsequently note that these findings do not apply to more modern clusters containing GPUs. Their findings are still relevant considering there are many legacy CPU-based programs, and how commodity CPU-based clusters are not going away any time soon.