

Chapter 2

End-to-End Dataflow Parallelism for Transfer Throughput Optimization

Esma Yildirim

Louisiana State University, USA

Tevfik Kosar

University at Buffalo (SUNY), USA

ABSTRACT

The emerging petascale increase in the data produced by large-scale scientific applications necessitates innovative solutions for efficient transfer of data through the advanced infrastructure provided by today's high-speed networks and complex computer-architectures (e.g. supercomputers, parallel storage systems). Although the current optical networking technology reached transport speeds of 100Gbps, the applications still suffer from the inadequate transport protocols and end-system bottlenecks such as processor speed, disk I/O speed and network interface card limits that cause underutilization of the existing network infrastructure and let the application achieve only a small portion of the theoretical performance. Fortunately, with the parallelism provided by usage of multiple CPUs/nodes and multiple disks present in today's systems, these bottlenecks could be eliminated. However it is necessary to understand the characteristics of the end-systems and the transport protocol used. In this book chapter, we analyze methodologies that will improve the data transfer speed of applications and provide maximal speeds that could be obtained from the available end-system resources and high-speed networks through usage of end-to-end dataflow parallelism.

DOI: 10.4018/978-1-61350-110-8.ch002

INTRODUCTION

The data transfer throughput is a major factor that affects the performance of applications from many scientific areas (e.g. high-energy physics, bioinformatics, numerical relativity and computational fluid dynamics). The advancements in optical networking technology have gone beyond the achievable throughput values the applications get, however the same speed up is not seen in the application performance due to many reasons such as the protocol inadequacy, poorly tuned protocol parameters and underutilized capacity of the end-systems. The current protocols that are highly common (e.g. TCP) were originally not designed for high-bandwidth networks. Due to its additive increase multiplicative decrease policy, TCP takes a long time to fill the pipe of long-fat network pipes. Many protocols have been designed for high-bandwidth networks in the transport layer (Kola & Vernon, 2007; Jin et al, 2005; Floyd, 2003) to overcome this problem however they fail to replace TCP.

Other than transport layer protocols, some application-level solutions are proposed as well. Two of the very common methods are tuning buffer size and using parallel streams. While some buffer tuning methods need modification to the kernel (Cohen & Cohen, 2002; Semke, Madavi & Mathis, 1998; Torvalds et al, 2010; Weigle & Feng, 2001), the others are done at the application level (Jain, Prasad & Davrolis, 2003; Prasad, Jain & Davrolis, 2004; Hasegawa et al 2001; Morajko, 2004). Although the buffer size parameter is properly tuned, it does not show a better performance than using parallel streams because parallel streams recover from packet loss quickly rather than a single stream with tuned buffer. They achieve high throughput by mimicking the behavior of individual streams and get an unfair share of the available bandwidth (Sivakumar, Bailey & Grossman, 2000; Lee et al, 2001; Balakrishnan et al, 1998; Hacker, Noble & Atley, 2005; Egger, Heideman & Touch, 2000; Karrer, Park &

Kim, 2006; Lu, Qiao & Dinda, 2005). However excessive usage of parallel streams reaches the network to a congestion point and it is hard to predict this point. The studies that try to find the optimal number of streams are so few and mostly based on approximate theoretical models (Hacker, Noble & Atley, 2002; Lu et al, 2005; Altman et al, 2006; Kola & Vernon, 2007}. They all have specific constraints and assumptions. Also the correctness of the proposed models is mostly proved with simulation results.

The foretold solutions to improve the throughput only remove the disadvantages of the protocols used. However, at some point the end-system resources become the source of bottleneck such as CPU, disk and NIC itself. Additional parallelism is needed through striping but the optimal level of striping is an open research area. The existing tools such as the GridFTP striped server (Allcock et al, 2005) and Dmover (Nathan et al, 2010) provide a means to utilize striping through multiple CPUs and nodes of an end-system architecture but they give the preference to the user to define the parallelism level. A dynamic and autonomic system that will decide this level depends on many factors.

In this book chapter we discuss many factors that affect the end-to-end application throughput such as the buffer size, parallel streams, CPU speed, disk speed and access methods in systems that use high-speed networks. The major purpose of this chapter is to provide insight to the characteristics of the end-systems that cause the bottleneck for throughput and to discuss future directions. We also present a method to optimize the parallel stream number and we have seen that this model gives very accurate results regardless of the type of the network.

BACKGROUND

TCP is the most widely used transport protocol but it may also be a major performance bottleneck in high-bandwidth networks because of its design

15 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/end-end-dataflow-parallelism-transfer/59676

Related Content

JIRL: A C++ Toolkit for JPEG Compressed Domain Image Retrieval

David Edmundson and Gerald Schaefer (2013). *International Journal of Multimedia Data Engineering and Management* (pp. 1-12).

www.irma-international.org/article/jirl/84022

Media Literacy Organizations

Iram Mukhtar Mahajan, Mudasir Rather, Huma Shafiq and Uzma Qadri (2018). *Digital Multimedia: Concepts, Methodologies, Tools, and Applications* (pp. 1394-1406).

www.irma-international.org/chapter/media-literacy-organizations/189533

An Adaptation Architecture Dedicated to Personalized Management of Multimedia Documents

Farida Bettou and Mahmoud Boufaïda (2017). *International Journal of Multimedia Data Engineering and Management* (pp. 21-41).

www.irma-international.org/article/an-adaptation-architecture-dedicated-to-personalized-management-of-multimedia-documents/176639

Video Summarization Based on Multimodal Features

Yu Zhang, Ju Liu, Xiaoxi Liu and Xuesong Gao (2020). *International Journal of Multimedia Data Engineering and Management* (pp. 60-76).

www.irma-international.org/article/video-summarization-based-on-multimodal-features/267767

Digital Watermarking Based on Neural Network Technology for Grayscale Images

Jeanne Chen, Tung-Shou Chen, Keh-Jian Ma and Pin-Hsin Wang (2005). *Encyclopedia of Multimedia Technology and Networking* (pp. 204-212).

www.irma-international.org/chapter/digital-watermarking-based-neural-network/17247