



TEXAS TECH UNIVERSITY

Edward E. Whitacre Jr.
College of Engineering

Computer Science

Data-Intensive Scalable Computing Laboratory (DISCL)

Technical Report

Department of Computer Science

Texas Tech University

Rethinking High Performance Computing System Architecture for Big Data Applications

Yong Chen, Chao Chen, Yanlong Yin, Xian-He Sun, William D. Gropp, Rajeev Thakur

yong.chen@ttu.edu, chao.chen@ttu.edu, yyin2@iit.edu, sun@iit.edu, wgropp@illinois.edu,
thakur@mcs.anl.gov

10/08/2012

Technical Report № TTU/DISCL-2012-02

<http://discl.cs.ttu.edu>

LIMITED DISTRIBUTION NOTICE: This report has been submitted for publication outside of TTU-DISCL and will probably be copyrighted if accepted for publication. It has been issued as a Technical Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of TTU-DISCL prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g. payment of royalties).

Rethinking High Performance Computing System Architecture for Big Data Applications

Yong Chen, Chao Chen, Yanlong Yin, Xian-He Sun, Rajeev Thakur, William D. Gropp

Department of Computer Science, Texas Tech University, Lubbock, TX

yong.chen@ttu.edu, chao.chen@ttu.edu, yyin2@iit.edu, sun@iit.edu, wgropp@illinois.edu, thakur@mcs.anl.gov

Abstract

Big data applications present a critical question to the high performance computing (HPC) community - how to efficiently support these applications with HPC systems that are traditionally designed for big compute applications? The conventional HPC systems are computing-centric and designed for computation-intensive applications. Big data problems have drastic different characteristics compared to big compute applications. These problems, however, will still largely rely on HPC systems to be solved. In this research, we try to answer this question with a rethinking of HPC systems architecture. We study and analyze the potential of a decoupled system architecture for big data applications. The new concept is to decouple conventional compute nodes and dynamically provision as data processing nodes that focus on data access capability for data-intensive applications. We present theoretical studies and analyses for such a decoupled HPC system architecture. The current results have shown its promising potential for big data applications. Its data-centric architecture can have an impact in designing and developing future HPC systems for data-intensive scientific discoveries and innovations.

Keywords: High performance computing, data-intensive computing, big data, storage