Chapter 12
Cyberinfrastructure, Cloud Computing, Science Gateways, Visualization, and Cyberinfrastructure Ease of Use

Craig A. Stewart
Indiana University, USA

Richard Knepper
Indiana University, USA

Matthew R. Link
Indiana University, USA

Marlon Pierce
Indiana University, USA

Eric Wernert
Indiana University, USA

Nancy Wilkins-Diehr
San Diego Supercomputer Center, USA

ABSTRACT

Computers accelerate our ability to achieve scientific breakthroughs. As technology evolves and new research needs come to light, the role for cyberinfrastructure as "knowledge" infrastructure continues to expand. In essence, cyberinfrastructure can be thought of as the integration of supercomputers, data resources, visualization, and people that extends the impact and utility of information technology. This chapter discusses cyberinfrastructure, the related topics of science gateways and campus bridging, and identifies future challenges and opportunities in cyberinfrastructure.

INTRODUCTION

Computers accelerate our ability to achieve scientific breakthroughs. As technology evolves and new research needs come to light, the role for cyberinfrastructure as "knowledge" infrastructure continues to expand. In essence, cyberinfrastructure can be thought of as the integration of supercomputers, data resources, visualization, and people that extends the impact and utility of information technology. This article defines and discusses cyberinfrastructure and the related topics of science gateways and campus bridging, identifies future challenges in cyberinfrastructure, and discusses challenges and opportunities related to the evolution of cyberinfrastructure and cloud computing.

BACKGROUND

Today's US national cyberinfrastructure ecosystem grew out from the National Science Foundation-funded supercomputer centers program of the 1980s (National Science Foundation, 2006). Four centers provided supercomputers and support for their use by the US research community. Researchers generally accessed one supercomputer at a time, sometimes logging into a front-end interface. At this time, the focus of the research computing community was centered on supercomputers – traditionally defined as computers that are among the fastest in existence. Over time there have been several different supercomputer architectures, but the key points were that supercomputers were monolithic systems that were among the fastest in the world. At present we can think of supercomputers as being a subset of the more general term high performance computer (HPC) – where HPC means that many computer processors work together, in concert, to solve large computational challenges and where the computer processors communicate via very fast, networks internal to the HPC system. HPC focuses on computing problems where a high degree of communication is needed among the processors working together on a particular problem. HPC is a more general term than supercomputers because there are many HPC systems that are modest in total processing capacity relative to the fastest supercomputers in the world (cf. Top500.Org, 2016).

In the early days of supercomputing, using multiple supercomputers in concert was not possible. In the late 1980s, the National Research and Education Network initiative created several testbeds for distributed computing, including the CASA testbed which linked geographically distributed supercomputers to solve large-scale scientific challenges (U.S. Congress Office of Technology Assessment, 1993). A turning point in distributed high performance computing was the I-WAY project – a short-term demonstration of innovative science enabled by linking multiple supercomputers with high performance networks (Korab & Brown, 1995). It demonstrated the possibilities to advance science and engineering by linking supercomputers using high-speed networks.

In the late 1990s, the NASA Information Power Grid provided a production grid of multiple supercomputers connected by a high-speed network (Johnston, Gannon, & Nitzberg, 1999). Around this time began also the concept of high throughput computing (HTC) with a software system called Condor (Litzkow, Livny, & Mutka, 1988). HTC takes the approach of breaking a problem up into small pieces of work and distributing them to multiple CPUs over network connections that may be relatively slow. HTC best suits problems where relatively little communication is needed among the processors working together on a particular problem or simulation. Because HTC applications can operate relatively efficiently on processors with little communication among the processors, HTC applications have always

Related Content

Fast HEVC Inter-Prediction Algorithm Based on Matching Block Features
Meifeng Liu, Guoyun Zhong, Yueshun He, Kai Zhong, Hongmao Chenand Mingliang Gao (2018).
International Journal of Mobile Computing and Multimedia Communications (pp. 40-61).
www.irma-international.org/article/fast-hevc-inter-prediction-algorithm-based-on-matching-block-features/198385

In-Memory Analytics
Jorge Manjarrez-Sanchez (2019). *Advanced Methodologies and Technologies in Network Architecture, Mobile Computing, and Data Analytics (pp. 381-389).*
www.irma-international.org/chapter/in-memory-analytics/214629

Intelligent Data Mining-Based Method for Efficient English Teaching and Cultural Analysis
Qing Aiand Hongyu Guo (2022). *International Journal of Mobile Computing and Multimedia Communications (pp. 1-14).*
www.irma-international.org/article/intelligent-data-mining-based-method/293745

Indian Healthcare Service Management Through Data Mining: Datamining for Healthcare Services
Manaswini Pradhan (2018). *Next-Generation Mobile and Pervasive Healthcare Solutions (pp. 219-233).*
www.irma-international.org/chapter/indian-healthcare-service-management-through-data-mining/187525

Ontology-Based Knowledge Model for Multi-View KDD Process
EL Moukhtar Zemmouri, Hicham Behja, Abdelaziz Marzakand Brigitte Trousse (2012). *International Journal of Mobile Computing and Multimedia Communications (pp. 21-33).*
www.irma-international.org/article/ontology-based-knowledge-model-multi/69531