



IDEA GROUP PUBLISHING

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA

Tel: 717/533-8845; Fax 717/533-8661; URL-<http://www.idea-group.com>

ITB10828

Chapter III

Using Peer-to-Peer Systems for Data Management

Dinesh C. Verma

IBM T.J. Watson Research Center, USA

Abstract

This chapter describes a peer-to-peer approach for managing data backup and recovery in an enterprise environment. Data management systems in enterprises constitute a significant portion of the total cost of management of data in enterprise systems. Maintaining data with a high degree of availability and reliability is typically done by having a centralized backup system that maintains backup copies of data. The maintenance of a large dedicated backup server for data management requires a highly scalable network and storage infrastructure, leading to a major expense center within the enterprise. With the current trends in workstation disk storage, an alternative peer-to-peer paradigm for data management can offer an approach that provides equivalent performance at a fraction of the cost of the centralized backup system. The author hopes that the contents of the chapter would lead to the development of more solutions that harness the power of peer-to-peer networks.

This chapter appears in the book *Peer-to-Peer Computing: The Evolution of a Disruptive Technology* by Ramesh Subramanian and Brian D. Goodman. Copyright © 2005, Idea Group Inc. Copying or distributing in print or electronic forms without written permission of Idea Group Inc. is prohibited.

Introduction

Peer-to-peer (P2P) technology has become mainstream media story due to its widespread use in exchanging popular files, and has almost become synonymous with music swapping over the Internet. The popularity of peer-to-peer systems on the Internet underscores the amount of scalability and ease of administration that is offered by that technology. These features enable peer to peer technology to be exploited for many other applications, both over the Internet as well as within enterprise networks. In this chapter, we look at one such application of peer-to-peer systems which can result in substantial cost savings for many enterprises.

One of the key support functions expected of an IT department in most medium- and large-scale companies is that of backup and recovery of files stored on the various computers within the enterprise. Data stored in a computer may be lost due to various causes, such as user mistakes, power outages, disk crashes, accidental erasures, viruses, and so forth. Data management systems within an enterprise facilitate the recovery of lost data. The standard modus operandi of data management systems is to make automated periodic copies of data, restoring files from the most recent backup copy when required.

Most enterprises operate a backup server to maintain backup copies of data across the different machines in an enterprise. Such backup servers must have the capacity to store all of the data present in a large number of client computers while providing a very high level of availability. Due to the volume of data in an enterprise, backup servers tend to be high-end, expensive machines with high-capacity data storage devices. Many vendors provide software products that provide data backup and restore functions.

The operation and maintenance of backup servers is an expensive operation for most enterprises. For every dollar spent on new storage hardware, enterprises are estimated to spend three dollars to manage data (Clark, 2002). In many enterprises, the growing cost of IT operations is a significant concern. Furthermore, industrial surveys indicate that up to 60% of storage capacity in an enterprise typically remains unused.

The excess storage capacity available in the devices can be exploited to provide data backup services to other devices in the enterprise. This can enable machines in an enterprise to provide collaborative backup service to each other, eliminating the need for an expensive backup server. The caveat, of course, would be to ensure that the collaborative backup remains as reliable as the traditional backup server.

This chapter describes a peer-to-peer system for data backup and recovery services, and analyzes its reliability and availability. Our calculations show that

11 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/using-peer-peer-systems-data/28042

Related Content

The Cyber Talent Gap and Cybersecurity Professionalizing

Calvin Nobles (2018). *International Journal of Hyperconnectivity and the Internet of Things* (pp. 42-51).

www.irma-international.org/article/the-cyber-talent-gap-and-cybersecurity-professionalizing/210627

Centrality Metrics, Measures, and Real-World Network Graphs: Node and Edge Centrality Metrics, Correlation Measures, and Real-World Network Graphs

(2018). *Centrality Metrics for Complex Network Analysis: Emerging Research and Opportunities* (pp. 1-33).

www.irma-international.org/chapter/centrality-metrics-measures-and-real-world-network-graphs/204776

Key Generation for Wireless Sensor Networks Using Symmetric Balanced Incomplete Block Design

K. G. Srinivasa, V. Archana, V. Poornima and C. Reshma (2012). *Technological Advancements and Applications in Mobile Ad-Hoc Networks: Research Trends* (pp. 202-218).

www.irma-international.org/chapter/key-generation-wireless-sensor-networks/64703

Protecting Data Confidentiality in the Cloud of Things

Bashar Alohal and Vassilios G. Vassilakis (2017). *International Journal of Hyperconnectivity and the Internet of Things* (pp. 29-46).

www.irma-international.org/article/protecting-data-confidentiality-in-the-cloud-of-things/179896

An Ontology-Based P2P Network for Semantic Search

Tao Gu, Daqing Zhang and Hung Keng Pung (2010). *Networking and Telecommunications: Concepts, Methodologies, Tools, and Applications* (pp. 1714-1727).

www.irma-international.org/chapter/ontology-based-p2p-network-semantic/49835