Virtual Large-Scale Disk System for PC-Room

Erianto Chai[1], Minoru Uehara[1], Hideki Mori[1], and Nobuyoshi Sato[2]

[1] Dept. of Information and Computer Sciences, Toyo University, 2100 Kujirai, Kawagoe, Saitama, 350-8585, Japan
[2] Faculty of Software and Information Science, Iwate Prefectural University, 152-52, Sugo, Takizawa, Iwate, 020-0193 Japan
eriaji@gmail.com, uehara@toyonet.toyo.ac.jp, mori@toyonet.toyo.ac.jp, nobu-s@iwate-pu.ac.jp

Abstract. There are many PCs in a PC room. For example, there are 500 PCs in our University. Each PC has a HDD, which is typically not full. If the disk utilization is 50% and each PC has a 240GB HDD, there is 60TB (500x120GB) free disk space. The total size of the unused capacity of these HDDs is nearly equal to the capacity of a file server. Institutions, however, tend to buy expensive appliance file servers. In this paper, we propose an efficient large-scale storage system that combines client free disk space. We have developed a java-based toolkit to construct a virtual large-scale storage system, which we call VLSD (Virtual Large-scale Disk). This toolkit is implemented in Java and consists of RAID (Redundant Arrays of Inexpensive/Independent Disks) and NBDs (Network Block Device). Using VLSD, we show how to construct a large disk that consists of multiple free spaces distributed over networks. VLSD supports typical RAID and other utility classes. These can be combined freely with one another.

Keywords: RAID, Network Block Device, Storage System.

1 Introduction

The requirements of higher education are growing. Companies expect graduates to maintain Linux servers and to administer networks, something which only privileged users are allowed to do. However, it is difficult for a beginner to become a privileged user. Consequently, we have developed a plan to use a virtual machine, such as VMware, Microsoft Virtual PC 2004, and so on, to allow students to learn how to maintain Linux servers in a safe environment. In this way, virtual machines are very useful in advanced education. These virtual machines, however, require much disk space. VMware recommends 8GB as the HDD capacity for Linux and 16GB for Windows. In our experience, Windows actually requires 24GB HDD capacity. This means that 40-120TB disk capacity is required for 5000 students.

An appliance file server with a 60TB capacity, was considered for this purpose. A commercial company quoted 250 million yen (about 2 million dollars) for such a file server. However, we already have 500 PCs. If we add 120GB HDDs to all our PCs at

T. Enokido, L. Barolli, and M. Takizawa (Eds.): NBiS 2007, LNCS 4658, pp. 476–485, 2007.

a cost of 150 dollars per HDD, a 60TB distributed file server can be realized at a cost of just 75 thousand dollars. The ratio of cost saving is 1:33, making this more important than any other factor, such as performance, dependability or maintainability. We therefore propose a distributed storage system and develop a toolkit to aid in the construction of such a system. We call this toolkit VLSD (Virtual Large-scale Disk).

We assume that a VLSD is used in the following educational environment. There are 512 client PCs, each running both Windows and Linux. Linux runs on VMware in Windows as a guest OS. Each client has at least 170GB free disk space. Therefore, the total available space is 70TB. We construct the distributed storage system using this free space. We can use the file system via a Linux file server, which supports NFS for Linux clients and SMB/CIFS for Windows. The Linux file server can be either real or virtual.

VLSD is a toolkit for developing virtual large scale disks, and can run on any platform that supports Java, since VLSD is 100% pure Java. Using VLSD, both Linux and Windows can be either a client or a server. A VLSD disk server creates a disk file on its own file system, and then publishes it as a remote disk for VLSD clients to share. VLSD is independent of the file system, and can combine Linux disk files with Windows disk files. In addition, VLSD can construct a virtual 64 bit file system on a 32 bit system.

VLSD can concatenate multiple virtual disks, thereby increasing the storage capacity. VLSD supports typical RAID classes, such as RAID0, 1, 4, 5, and 6, and can construct multiple hierarchies of RAID combining any RAID class with any other RAID class.

The remainder of this paper is organized as follows. In Section 2, we introduce several related works and give details of the problem. In Section 3, we introduce our VLSD toolkit. We then evaluate the VLSD and discuss the results. Finally, we offer our conclusions.

2 Related Works

2.1 64 Bit File Systems

We need more than a 64 bit file system to construct a large-scale storage system. Here, a 64 bit file system means that the seek pointer of a file is represented as a 64 bit integer. Initially we wish to construct a 70TB storage system, but in the future, we intend constructing 4EB storage systems. A 64 bit file system theoretically allows us to construct a 16EB storage system. However, this is not really feasible because Java's long is a 64 bit signed integer. In practice, the maximum disk capacity is less than 8GB if Java is used.

SGI's XFS is a popular 64 bit file system running on Linux. UFS2 is also popular in FreeBSD. Sun's ZFS is provided for Solaris. In a sense, ZFS is the greatest file system because it uses 128 bits. In this paper, we use XFS because the file server runs on Linux.