A NETWORK LAB EXPERIMENT OF MULTI-CLONING OF OS BY USING CLONEZILLA

Bishnu Prasad Gautam and Dambar Raj Paudel

● **Abstract**

OS

Clonezilla   OS

● **Key words**

Diskless cloning
Cloud Architecture
Clone Metering
Diskless Client Virtual Cluster

1. Introduction

Disaster reducing multi-cloning is an important technology for large and corporate organization. Sooner or later we are going to need a more robust methodology for backup of a server and a system to restore. Whether it's an upgrade that goes bad, corruption of key operating system (OS) files from hardware failure, or simply rolling out a number of identical servers, using smart backup strategies can drastically reduce the downtime required by any of these tasks. Further, mass installation of operating systems with different kinds of software packages save time, electrical energy and human resources that solve efficient system installation in the organization and energy saving problem. In this paper, we discuss mass installation of disaster free partition cloning useful for academic, corporate or governmental organization. We also successfully experimented these techniques for server rollout, and for upgrading of operating system with supporting packages.

Multiple cloning of system enables the hardware resource optimization and rapid deployment of service in the network. Further it plays an important role in saving energy helping environmental problems for human beings. We discuss Diskless Remote Boot in Linux (DRBL) which provides a diskless or systemless environment for client machines. It works on Debian, Ubuntu, Red Hat, Fedora, CentOS and SuSE. DRBL uses distributed hardware resources and makes it possible for clients to fully access local hardware.

2. Problem: Lengthy process of System Installation

Imagine that you need to install the systems in intranet having 100s of clients and you are alone to complete the task. Installation of OS in those machines within limited time is very time consuming task. This would be the typical problem for any system administrator who needs to complete the entire system installation in his networks within given time frame.

3. Solution: Disk Imaging Technique

Disk image is also called as disk clone, an exact copy of a computer hard drive. The copy includes all the partition information, boot sectors, the file allocation table, operating system installation and application software. Disk images are used to transfer hard drives contents during a hardware upgrade, to restore a hard drives contents during disaster recovery or when a hard drive is erased, and to transfer the contents of a hard drive from one computer to another.Typically, special disk imaging software is required to copy the hard drive data [1]. It saves the entire data from the disk, including the file structure and all files and folders from the disk, in a single file. Because disk images are exact copies, or "clones," of original disks, they can be used to duplicate disks or serve as full backups in case a system restore must be done.

Most disk image files store data in a raw, binary format. This means they do not have a file system, which tells the computer how to access the files and folders in the disk image. Therefore, in order for the data making readable, disk image must be mounted first by either the operating system or a disk utility program.

4. Current Issues

System installation process from single system to mass installation of operating system required utility

software. There are few cloning software such as Symantec Ghost [6], Drive Cloner [7] to name a few, however, the speed of cloning still takes substantial amount of time. As OS data volume is increasing, it directly affect in the rate of imaging time wise and resource utilization. Moreover, most of the available tools clone the entire sectors of the hard disk containing the empty sectors also.

5. Experimented tool: Diskless Remote Boot in Linux (DRBL)

Diskless Remote Boot in Linux (DRBL) is an open source solution to manage the deployment of the GNU/Linux operating system across many clients [3]. DRBL supports lots of popular GNU/Linux distributions, and it is developed based on diskless and systemless environment for client machines. DRBL uses PXE/Etherboot, DHCP, TFTP, NFS and NIS to provide services to client machines so that it is not necessary to install GNU/Linux on the client hard drives individually [3]. Preparation is required to set up a server machineas a DRBL server which store and send template image to clients during operation, and follow the DRBL installation wizard to configure and push the environment for client machines step by step.
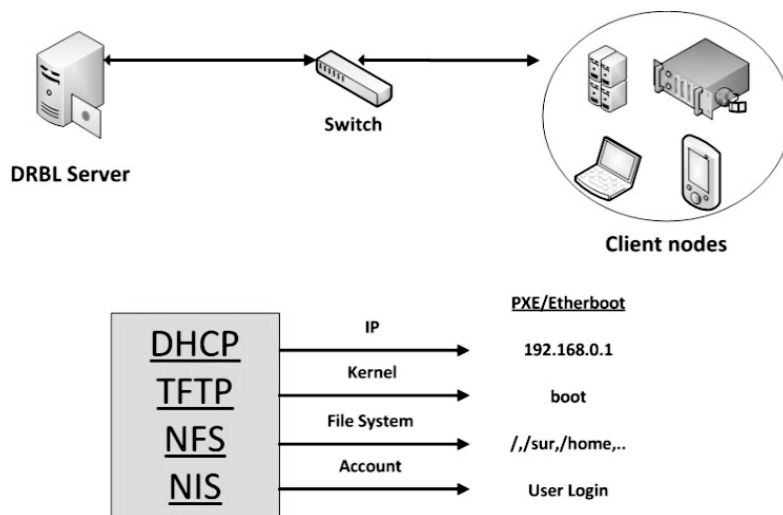

Figure: DRBL Process [3]

Thus, cross-platform and user-friendly are the key factors that make the DRBL become a superior imaging tool. DRBL can efficiently deploy diskless or diskfull cluster environment, and manage client. It configures these services (TFTP, NIS, DHCP, and NFS) to build a cluster environment. According to this implementation, an administrator just needs two steps to deploy cluster environment. In the first step he/she installs DRBL packages and generates kernel and initrd for client; secondly setup environment parameters, such as IP address, and set the numbers of clients that are going to be installed. It also provides cluster management and cluster system transformation (diskfull or diskless system). It focuses on effective configuration and deployment for Cluster environment; the package based deployment tools are the best choice. As DRBL is package-based deployment tools, it does not need complex pre configuration. DRBL also provides Cluster management tools to manage client nodes.
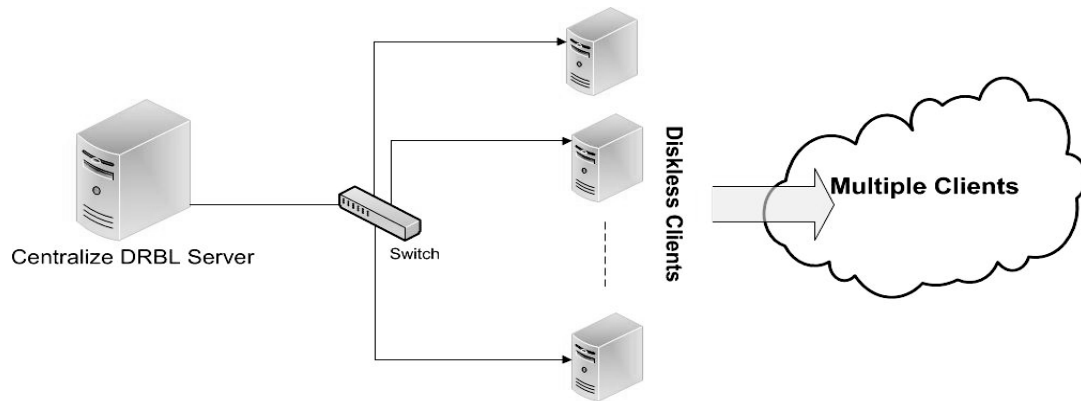
Figure 1: DRBL Cluster System Architecture

6. Experimented Lab Environment

We have used University LAB consisting 64 workstations for mass deployment of multi-Operating System. Windows7 and Ubuntu 10.04 were cloned by making one of the computers as DRBL server with software supporting libraries and packages. It's very flexible to transform between two different modes cluster environment (diskfull and diskless) through DRBL. The cluster has 1 server, 64 client's machines. Client machine does not need to install any things in prior, however it needs PXE enabled connection to the DRBL server and manage the deployment operating system.

For transporting image simultaneously to all computers; multicast is a good choice. The multicast function allows sending the data only to those receivers that requested it. Ethernet cards of machines which don't participate in the transmission automatically block out the packets at the hardware level. Moreover, network switches are able to selectively transmit the packets only to those network ports to which receivers are connected. Both features thus allow a much more efficient operation than broadcast.Therefore, if switch does not block multicast packets, it's better to use multicast function.

Clonezilla, based on DRBL, Partclone and udpcast, allows to do bare metal backup and recovery. Two types of Clonezilla are available, Clonezilla live and Clonezilla SE (server edition). Clonezilla live is suitable for single machine backup and restore. While Clonezilla SE is for massive deployment, it can clone many computers simultaneously. Clonezilla saves and restores only used blocks in the hard disk. This increases the clone efficiency. At the University's LAB, Clonezilla SE was used to clone 64 computers simultaneously. It took only about an hour to clone a 20 GB system image to all 64 computers via multicasting.

7. Mode of Clonezilla

Cloning a Hard disk or single system backup which can be used during crash recovery is developed as Clonezilla Live. It works on hard drive or PXE boot. Similarly, Clonezilla Server (DRBL) is used for mass cloning. In Clonezilla mode the template system image is copied into DRBL Server and secondly it is cloned to multiple clients. Both of these systems can be used according to requirements. When a server or a client crashes then the written image in disk can be written into the crashed system. So Clonezilla is disaster recovery backup

software designed to clone or backup entire drives. It can backup hard disks regardless of the data or operating system contained in them. In the figure below two types of cloning method are presented. The first Clonezilla live copies entire image into storage disk while the Clonezilla server mode is working in two steps. First it makes images of template system into DRBL server secondly it export image to multiple clients at the same time.
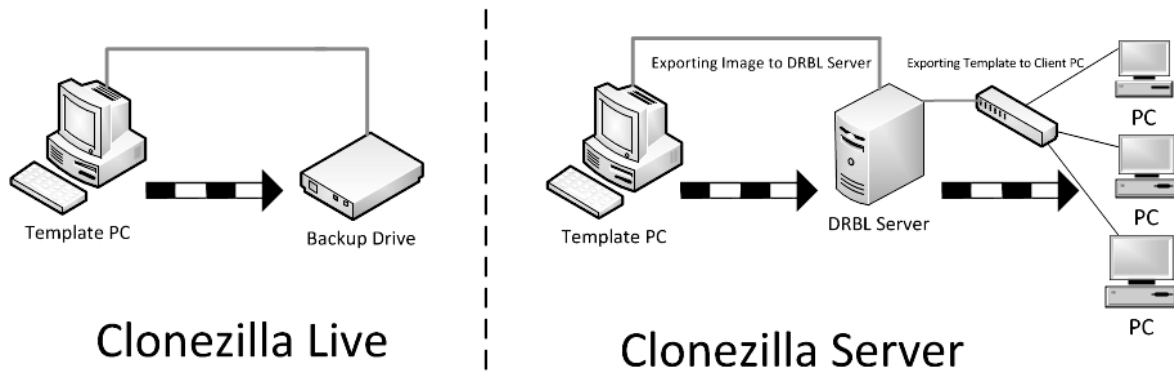

Figure 2: Clonezilla live and Clonezilla Server Mode

8. DRBL Installation

First step, it has to install DRBL software in server. If system OS is Debian/Ubuntu package system, it just installed l package from DRBL website. Then it needs to execute DRBL configuration command "drplsrv -i" to choose your Kernel version for nodes and automatic installs the packages that DRBL required, such as DHCP, NFS, NIS and TFTP [3]. Then, using DRBL deployment command "drblpush -i" to push Kerrighed environment to all nodes. Kerrighed is a Single System Image operating system for clusters. DRBL offers interactive dialog to help users to build DRBL environment and it automatically configures and starts all the services required to
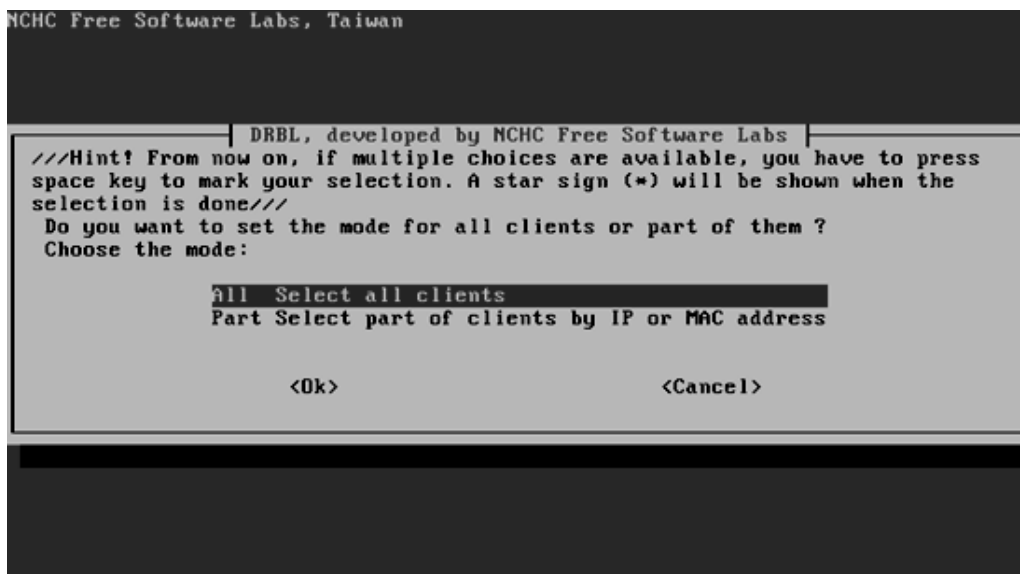

Figure 3: DRBL Installation screen many of one

make the Cluster work [2]. It automatically detects the network interfaces that have private IP addresses assigned to them and asks users how many clients are required to set up. DRBL provides two methods for nodes IP address: (1) fix IP address (binding MAC address): this feature is useful to setting up system for security; (2) dynamic IP address (range of IP address) in the open environment where anyone can add a new machine dynamically.

9. Configuring and Managing environment for DRBL-Kerrighed

When above steps accomplished, then the DRBL with Kerrighed is complete. For specific purposes, some system configuration and environment tuning may be necessary. The command "dcs" of DRBL pops out DRBL management graphic user interface to manage nodes. In this case, it is suitable to choose DRBL to deploy nodes massively [2]. DRBL can fast and efficiently deploy nodes, the deployment procedure just need two commands (drplsrv .i, drblpush .i). DRBL can automatically setup required services (DHCP, NFS, NIS and TFTP). DRBL offers central management interface to effective managed and configured nodes. In addition, DRBL also offers a lot of management commands.

10. Related Work

Previously many attempts [1] have been made in disk cloning and mass installation. To help understand the requirement of large and enterprise organization, even single system backup for minimizing disaster different commercial vendors kept their production in the market. Simultaneously open source community was also in process to bring their latest technology featuring easy use and open community support [5]. Vendors like Symantec with its commercial product Ghost was a milestone for the recent requirement for disk cloning and data backup though there were limitations. Clonezilla of NCHC Lab was another open source milestone project that was discovered to solve the Ghost limitation as well as for open source community.

11. Limitations

In recent days most of the users may think to install OS with required utilities without any hardware or software boundaries. The documents on Clonezilla shows that the destination partition must be equal or larger than the source one is one of the major limitation as user may not have similar kind of storage device. Secondly, differential/incremental backup is not implemented yet. The next challenge is that online imaging/cloning which is not implemented which can be taken as major limitation. The partition to be imaged or cloned has to be un-mounted. Software RAID/fake RAID is not supported by default. Due to the image format limitation, the image cannot be explored or mounted [2]. It cannot be recovered single file from the image in case a single file is broken. However, it still has workaround to make it. Recovery Clonezilla live with multiple CDs or DVDs is not implemented yet. Now all the files have to be in one CD or DVD if it is required to choose to create the recovery iso file.

12. Result and Discussion

DRBL belongs to the package based automatic installers because it has packed several necessary packages for

users to install easily by running the DRBL scripts. Lots of GNU/ Linux distributions support DRBL, so that it is convenient for users to build up a Cluster from those supported platforms. Users can have additional choices on different platforms, and the diskless environment makes Cluster deployment and management easier. DRBL provides lots of practical commands with brief instructions for users to modify those configuration files and necessary changes from client nodes, so that users can control all these client nodes and let them act as users need at the same time through DRBL related commands [3]. That's the reason of choosing DRBL for deployment and management of Cluster environment, because it performs excellent in both node configuration and Cluster management.

13. Conclusion

Clonezilla is an efficient, networkable, piece of software that can clone a single machine to multiple machines simultaneously. DRBL can be a convenient Clustering tool for users to deploy Single System Image Clusters easily and quickly. The ability to quickly and reliably image the machines in lab is a precious achievement. An open source Clonezilla can be used and recommended for installation of systems in academic institutions, in order to install an efficient, networkable, piece of software in a single machine or several machines simultaneously.

● **References**

[1] Steven S., Ceasar S., Jazz W.,Thomas T., Kuo-Lien H., Jean-Francois N., "The massive deployment tools: DRBL and Clonezilla", *2009 WPCG (Workshop on PC-Grid)*, HPC Asia 2009.

[2] Jiun-Yu W., Yao-Tsung W., Steven S., Hui C. W., "Using DRBL to Deploy MPICH2 and CUDA on Green Computing", *INFOCOMP 2011 : The First International Conference on Advanced Communications and Computation*.

[3] Wen-ChiehKuo, Che-Yuan Tu, Yao-Tsung Wang, "Deploy Kerrighed SSI massively using DRBL", *National Center for High-Performance Computing*, Taiwan

[4] G. Vallee, Stephen L. Scott, C.Morin, J.-Y. Berthou, and H. Prisker. "SSI-OSCAR: a Cluster Distribution for High Performance Computing Using a Single System Image. "*Proceedings of the 19th International Symposium on High Performance Computing Systems and Applications (HPCS'05) 1550- 5243/05*, IEEE 2005. [7] J. Parpaillon.

[5] HaiyangQian and Deep Medhi, "Server Operational Cost Optimization for Cloud Computing Service Providers over a Time Horizon", University of Missouri-Kansas City, Kansas City, 2011, MO, USA

[6] Symantec Ghost, http://www.symantec.com/

[7] Drive Cloner RX, http://www.horizondatasys.com/363895.ihtml

●英文タイトル

A NETWORK LAB EXPERIMENT OF MULTI-CLONING OF OS BY USING CLONEZILLA

●要約

In order to reduce unwanted resource consumption during mass OS installation, optimized cloning techniques with minimized environmental impact and power consumption are becoming increasingly important research

issues for academic institutions, organizations, and technological architect leaders.

In this paper we are going to describe the lab experiment of disk cloning technique that we experimented in our network lab for mass installation of OS with the help of Clonezilla.