

基于 IP 网络的 Linux 集群系统备份、恢复模型

梁宝龙^{1,2}, 王 旭^{1,2}, 叶 钊^{1,2}, 郑虹晖^{1,2}, 程洪涛^{1,2}

(1. 海南省气象信息中心, 海南 海口 570203;

2. 海南省南海气象防灾减灾重点实验室, 海南 海口 570203)

摘 要: 为了提高 Linux 系统在集群业务中运行的可靠性, 对 Linux 集群系统的备份与恢复进行了研究。设计了一种基于 IP 网络的 Linux 系统的备份与恢复模型。该模型分服务器端和客户端, 通过服务器端将客户端的系统备份, 并在实际的环境中验证该模型的有效性。实验结果表明, 该模型可用于实时 Linux 集群系统的备份与恢复, 有效地提高了 Linux 集群系统恢复正常运行的可靠性, 有助于 Linux 集群实时业务系统的尽快恢复。在实时性业务要求较高的 Linux 集群业务运行中, 有很好的指导意义。

关键词: Linux 系统; 集群; 备份; 恢复; 模型

中图分类号: TP31

文献标识码: A

文章编号: 1673-629X(2015)08-0053-05

doi:10.3969/j.issn.1673-629X.2015.08.011

A Backup and Recovery Model Based on IP Network for Linux Cluster System

LIANG Bao-long^{1,2}, WANG Xu^{1,2}, YE Fan^{1,2}, ZHENG Hong-hui^{1,2}, CHENG Hong-tao^{1,2}

(1. Hainan Meteorological Information Center, Haikou 570203, China;

2. Key Laboratory of South China Sea Meteorological Disaster Prevention and Mitigation of Hainan Province, Haikou 570203, China)

Abstract: In order to improve the reliability of Linux system running in the business of cluster, the backup and recovery of Linux cluster system is researched, and a backup and recovery model based on IP network for Linux cluster systems is proposed. This model divides server and client between backup and recovery on the Linux cluster systems, and the validity of this model in the real environment is verified. According to the experimental result, this model can be used to real-time backup and recovery to improve the reliability of running for Linux cluster system effectively. This model has valuable reference to the higher requirements of real-time Linux clusters business.

Key words: Linux system; cluster; backup; recovery; model

0 引 言

在日益发展的互联网时代, 高性能 Linux 集群系统在 7 * 24 小时运行的实时业务中需要稳定的数据收集及传输, 对服务器的稳定运行要求极高。在对硬件的要求方面加强之后, 对系统的稳定运行也有很高的要求。如何进一步保证系统正常运行, 目前较为广泛应用的集群技术能有效提高 Linux 系统的冗余性。但是极个别系统的宕机、人为的误操作导致配置文件丢失等等, 在一定程度上会对业务的正常运行造成影响。在实时业务系统出现异常后, 如何以最快的速度对系

统进行恢复, 逐渐成为业务维护人员首要考虑的问题。

如何保障业务系统的正常稳定运行, 备份与恢复技术近年来成为研究热点。对系统备份与恢复的研究, 前人做了很多有指导性的工作^[1-5]。王德军等^[6]提出的裸机备份与恢复技术研究、任欣等^[7]提出的远程文件备份与恢复系统的设计与实现, 分别从裸机备份恢复技术和远程备份的方法对系统的备份与恢复进行研究。郭京等^[8]提出的基于 Linux 的异步网络备份系统的设计与实现、刘仕一等^[9]提出的异地备份系统数据一致性检测方法、邱龙金等^[10]提出的安全灾难备

收稿日期: 2014-07-16

修回日期: 2014-11-05

网络出版时间: 2015-06-23

基金项目: 国家科技支撑计划课题(2013BAK05B03); 中国气象局气象关键技术集成与应用项目(CMAGJ2012Z04); 海南省气象局科技创新项目(HN2013MS22)

作者简介: 梁宝龙(1986-), 男, 硕士, CCF 会员, 研究方向为计算机系统结构。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20150623.1005.002.html>

份系统,分别从异地备份、数据一致性、安全的容灾备份等方面对系统的备份和恢复提出了很好的建议。敖莉等^[11]提出的重复数据删除技术、浦海挺等^[12]对 Linux 下系统备份/恢复软件的实现为系统备份的设计与恢复研究做了很多有益的贡献。

针对气象行业实时业务传输系统的情况,文中提出一种基于 IP 网络的 Linux 集群系统的备份、还原模型,并在海南气象通信系统的真实 Linux 集群系统环境中进行备份与恢复测试。测试结果表明,该模型对 Linux 集群系统中单机系统的备份、恢复具有很好的指导意义,对 Linux 集群系统的高可靠性运行具有很好的促进作用。

1 基于 Linux 的网络集群通信结构

1.1 网络体系结构

当前的气象 Linux 集群服务器通信系统的结构如图 1 所示。备份服务器 1 台,型号为 IBM x3850 X5,作

为备份服务端并部署服务器端软件,简称 backup-server。客户端服务器 6 台,作为系统备份还原的客户端,分别部署客户端软件。其中 4 台 HP DL380G5 服务器,组成新一代国内气象通信系统海南省级系统(简称新一代通信系统),承担海南省实时报文业务数据的收集和上行传输业务,分别简称 behk-srv1、behk-srv2、behk-srv3、behk-srv4;2 台 IBM X3650 服务器,组成中国气象局卫星广播系统(CMACast)海南分系统(简称 CMACast 系统),负责实时接收中国气象局卫星广播播发的各类气象数据,并存储和分发到研究所、气象台等各业务部门使用。CMACast 客户端分别简称 cmacast-recv1、cmacast-recv2。6 台客户端服务器通过 SAN(Storage Area Network)^[13]存储网络和 IP 网络组成一个集群系统,实现业务数据的快速交换。备份服务器通过 IP 网络和其他 6 台集群客户端服务器百兆网络全双工互通。

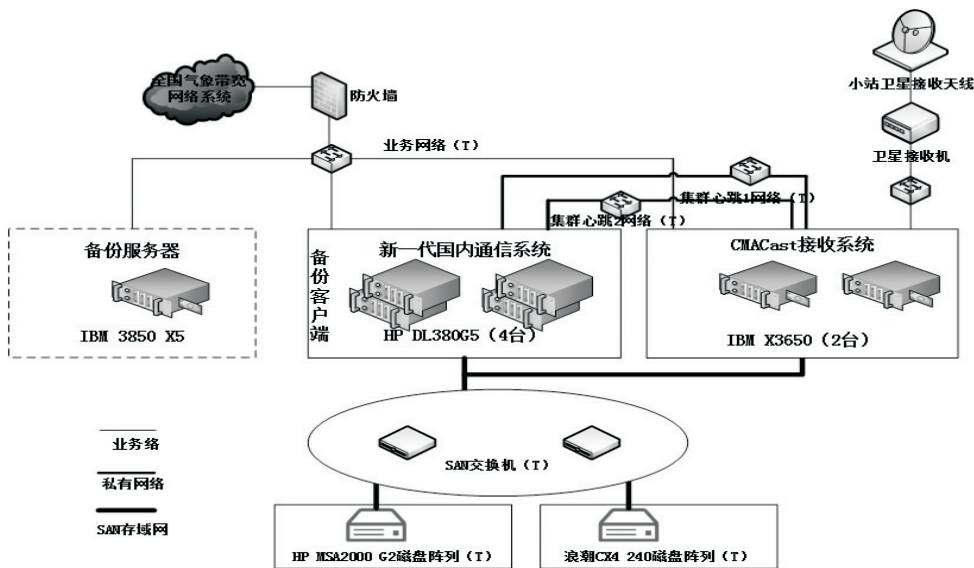


图 1 通信系统结构

1.2 集群体系结构

在新一代通信系统和 CMACast 系统组成的集群系统中,集群各服务器有两组通信链路,分别是心跳网络和业务数据传输网络。心跳链路检测该服务器是否在集群中,并确保该集群服务器在集群系统通信,维护集群中心跳网络的连通性;业务数据传输链路,在承担业务数据传输的同时,也作为集群的通信链路。只要集群中这两组通信链路中的一组链路正常,都可承担其集群的正常工作。平时也可通过查看该集群的网络视图,可随时关注集群各节点服务器的通信链路状态是否正常,绿色表示链路正常,红色表示链路中断。

集群的业务应用视图如图 2 所示。集群的应用包括集群管理服务(cvm)、负载均衡服务(LVS_VS)、mysql(数据库服务)、数据收发服务(mms_app_ha)、文

件传输服务(ftp-service)等等,各个应用的状态分别为在线/离线(online/offline)。

整个集群的各个应用分布在 6 个服务器节点中运行,各自承担相应的业务,共同协作完成海南气象资料的上行传输和 CMACast 下行资料的接收处理。新一代系统集群应用承担全省气象资料的收集和上行传输,尤其是整点国家基准站报文、精细化预报报文、雷达产品等资料的收集和上传业务,传输考核要求比较严格,要求 10 分钟以内必须要传到国家局服务器。节点服务器一旦出现软件、系统等问题时,短期内会影响业务的正常运行,影响全省气象资料传输考核质量。因此,集群系统的正常可靠运行尤为重要。在硬件、软件方面加强之后,系统的备份、快速还原也逐渐成为一种重要的保障手段。

➡ Status (All Groups status on systems) : behk_cluster

Systems ▶						
Groups ▼	behk-srv1	behk-srv2	behk-srv3	behk-srv4	cmacast...	cmacast.
LVS_RS	Online	Online	Online	Online		
behk_grp	Online	Online	Online	Online	Online	Online
cvm	Online	Online	Online	Online	Online	Online

图 2 集群的业务应用视图

2 总体设计

2.1 技术方案

新一代通信系统和 CMACast 系统承担 7 * 24 小时业务不间断运行,软件和配置文件一旦确定之后,变更不大。因此,首要考虑的指标是系统出现崩溃之后如何快速恢复。基于快速恢复考虑,文中拟采用 BMR^[14-15](Bare Metal Restore,裸机恢复)技术实现对客户端节点服务器的系统整体备份和恢复。裸机恢复优点是在灾难性故障发生之后,擦除对计算机进行重新格式化的恢复过程,比较符合当前气象业务的快速恢复的需求。同时通过 BMR 技术在每次系统配置做

变更之后,均可在不影响系统正常运行情况下,对当前的系统做备份。备份镜像文件通过 BMR 实例备份到服务器端存储,恢复时光盘引导申请恢复通过 BMR 技术对系统进行整体恢复。

2.2 模型结构

模型的体系结构由备份服务器端和客户端两部分组成,集群系统备份模型结构如图 3 所示。客户端首次收集系统配置信息时,为了不影响当前的实时传输业务,需要将该服务器的业务切换到其他集群服务器中,并将该服务器退出当前的通信集群,收集信息结束后可手工启动应用并开启集群服务恢复当前服务器的业务传输。

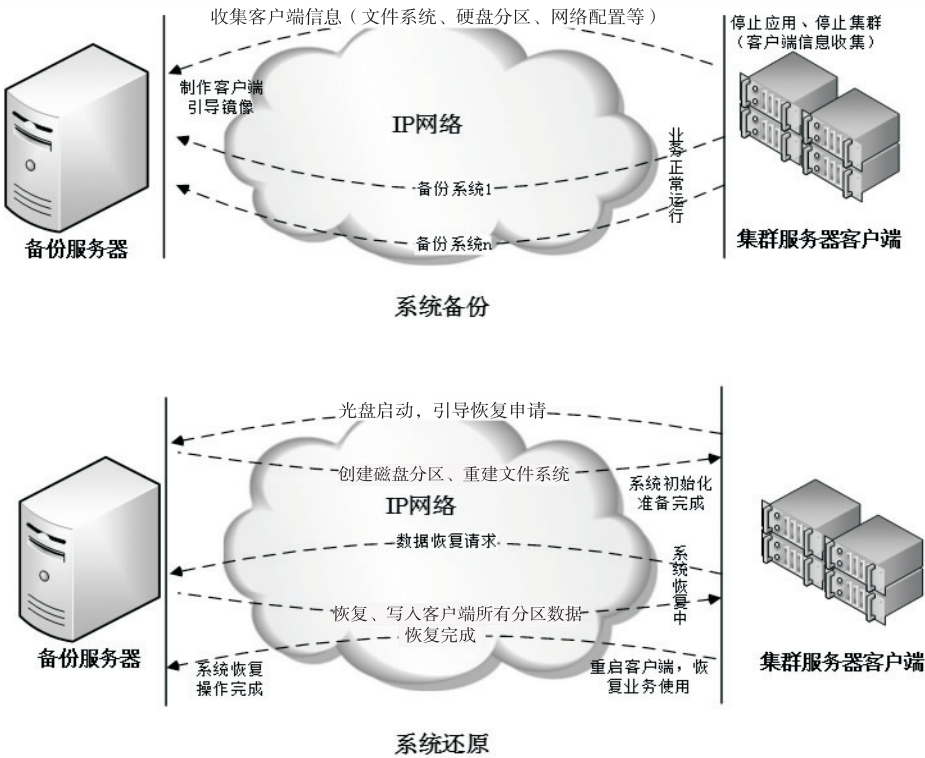


图 3 集群系统备份、还原

备份操作,通过备份实例将客户端的配置信息及数据通过 IP 网络备份到服务器端。通过服务器端逐一收集群集客户端的系统配置信息,包括将客户端的

操作系统、硬件驱动、文件系统、应用配置信息及本地全部数据备份,在服务器端为每个客户端建立一个备份实例,并制作客户端引导镜像。

还原操作,当对客户端通过光盘引导发起还原操作申请时,将进行如下操作:

服务器端通过客户端实例为客户端创建磁盘分区、重建文件系统,完成系统初始化准备。通过 IP 网络将服务器端对应备份的实例数据按规则还原写入到客户端服务器的对应分区中,数据还原写入完毕后重启服务器,客户端系统恢复并加入集群运行。

3 具体实现

3.1 备份收集准备

备份与恢复模型的实现分为服务器端和客户端两部分。服务器端通过收集客户端配置信息、分区等将客户端备份实例添加到服务器端。新一代通信系统和 CMACast 系统客户端服务器上分别安装客户端软件,服务器端 backup-server 安装服务器端软件,客户端和服务端同在一个可互通的 IP 网络环境里。备份、恢复操作均由服务器端发起启动指令。

3.2 收集客户端配置信息

收集客户端服务器的配置到服务器端的根目录,需要停止单机的集群应用、收集客户端的硬件信息、配置信息、开启集群应用共四个步骤,例如收集 behk-srv1 配置的操作步骤如下:

Step1:停止集群服务,脚本命令操作如下:

```
hastoplocal
/etc/init.d/vxodm stop
/etc/init.d/vxfen stop
```

Step2:收集客户端服务器硬件信息。

命令 fdisklu>hardinfo.txt #查看并将磁盘信息输出到 hardinfo.txt 文件中,硬盘的信息在前 15 行之中

命令备份 fdisk,cp/sbin/fdisk/sbin/fdisk.bak #收集客户端信息之后恢复 fdisk 执行命令:cp/sbin/fdisk.bak /sbin/fdisk

命令 echo“fdisk.baklu | head 15”>/sbin/fdisk

Step3:收集客户端信息,执行命令/usr/opensv/net-back/bin/bmrsavecfg infoonly 后,在/usr/opensv/net-back/baremetal/client/data 目录下生成 bundle.bat,将 bundle.bat 重新命名为 behk-srv1_bundle.dat 并拷贝到备份服务器端 backup-server 的根目录上,命令 scp behk-srv1_bundle.bat backup-server:/。

Step4:收集完成之后,开启客户端集群服务:

```
/etc/init.d/vxfen start
/etc/init.d/vxodm start
hastart
```

至此,集群中的 behk-srv1 客户端服务器信息收集完毕。重复以上步骤,分别完成收集其他客户端服务器信息到 backup-server 备份服务器的根目录上。

3.3 备份服务器加入客户端实例

Step1:启动 BMR 并通过 BMR 来建立连接客户端备份,脚本如下:

```
/usr/opensv/netbackup/bin/rc.bmr start
#启动 BMR 服务
/usr/opensv/netbackup/bin/bmrsetupmaster
#创建 BMR 数据库
/usr/opensv/netbackup/bin/bmrsetupboot-register
#启动 boot server
```

Step2:将客户端 behk-srv1 加入到服务器端,把节点 behk-srv1 的客户端信息加入服务端,脚本如下:

```
/usr/opensv/netbackup/bin/bmrs -o import -r config -path/behk-srv1_bundle.dat -client behk-srv1
```

Step3:在服务器端配置备份策略,添加设备备份类型,备份的分区,自动备份时间,修改备份记录属性等。

至此,在服务器 backup-server 建立了一个客户端 behk-srv1 的备份策略 BMR。

重复以上操作,逐一为其他客户端建立相应的备份策略,并将策略加入至服务器端。

4 实验与分析

实验环境:客户端 6 台服务器组成集群,操作系统均为 SUSE Linux Enterprise Server 10 SP2,内核为 Linux SuSE 2.6.16,其中 4 台组成新一代国内气象通信系统海南省级系统,负责实时业务数据上行传输业务;其余 2 台组成气象业务下行的 CMACast 海南分系统,负责实时对 CMACast 卫星数据下行接收处理;服务器端操作系统为 SUSE Linux Enterprise Server 10 SP4,内核为 Linux SuSE 2.6.16。备份客户端服务器 CPU、内存等基本参数如表 1 所示。

表 1 备份还原服务器 CPU、内存参数

服务器名称	CPU 参数	内存/GB
behk-srv1 ~ behk-srv4	8 核 Intel(R) Xeon(R) E5405 @ 2.00 GHz	32
cmacast-recv1 ~ cmacast-recv2	4 核 Intel(R) Xeon(R) E5440 @ 2.83 GHz	8
backup-server	48 核 Intel(R) Xeon(R) E7530 @ 1.87 GHz	32

备份操作:在服务器 backup-server 端逐个对客户端进行备份操作,备份期间不影响客户端的正常运行。查看备份日志统计备份所需的时间并作记录。表 2 分别统计了各节点第一次备份的镜像文件、配置文件的存储空间。各节点应用进程在备份时所承担的业务不一致等原因,镜像文件大小不一致,配置文件存储空间

差别不大。

表 2 客户端节点备份镜像、配置文件

主机名称	C1_F1 文件/GB	C1_TIR 文件/MB
behk-srv1	7.8	19
behk-srv2	7.3	19
behk-srv3	7.2	19
behk-srv4	7.3	19
cmacast-srv1	6.2	16
cmacast-srv2	6.4	17

还原操作:停客户端集群,停客户端应用,服务器端发起还原操作,插入客户端光盘重启服务器并按照提示输入服务器端的名称 backup-server、IP 地址、掩码、网关等服务器信息,之后无人值守系统还原直至完成系统还原操作。还原结束确认之后,自动重启系统,系统启动后加入到集群中正常运行。查看还原日志并逐个统计系统还原时间。

为测试文中提出的备份、还原模型的有效性,对 6 台客户端 Linux 集群服务器的系统备份、恢复进行测试。集群系统的备份与恢复的测试时间如图 4 所示。各系统的备份所需的平均时间小于 7 min,并且备份期间不影响正在运行 Linux 集群系统的实时业务。系统恢复时间小于 10 min,可以满足气象行业传输的时效要求。

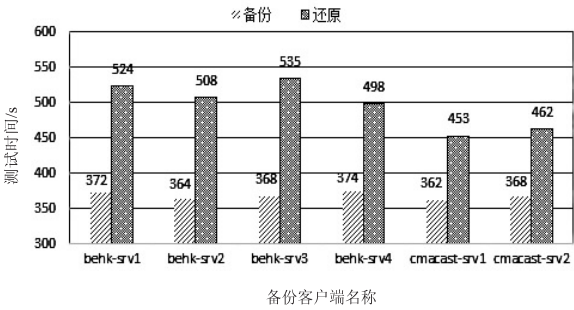


图 4 集群系统备份、还原时间

5 结束语

文中设计并实现了一种基于 IP 网络的 Linux 集群系统备份、恢复模型,并在海南气象通信系统的真实环境中做了备份、恢复测试,结果表明了提出的模型有效性。该模型有效提高了海南省气象通信集群系统异常故障恢复速度,为新一代国内气象通信系统海南省级系统的日常维护工作提供了一种快速有效的系统备份、恢复机制。通过擦除对计算机进行重新格式化的

快速恢复过程,对 Linux 集群系统的备份、还原有一定借鉴和推广作用。

参考文献:

[1] Wang Y,Li Z,Lin W. A fast disaster recovery mechanism for volume replication systems[C]//Proc of international conf on high performance computing and communications. [s. l.]:[s. n.],2007:732-743.

[2] Lawler C M,Harper M A,Szygenda S A,et al. Components of disaster-tolerant computing:analysis of disaster recovery,IT application downtime and executive visibility[J]. International Journal of Business Information Systems,2008,3(3):317-331.

[3] Shao B B M. Optimal redundancy allocation for information technology disaster recovery in the network economy [J]. IEEE Transactions on Dependable and Secure Computing, 2005,2(3):262-267.

[4] 韩德志,谢长生,李怀阳. 存储备份技术探析[J]. 计算机应用研究,2004,21(6):1-4.

[5] 瞿志伟,谢康林,陈 琰,等. 磁盘备份系统的快速数据恢复方法[J]. 计算机工程,2005,31(9):198-200.

[6] 王德军,王丽娜,郝永芳. 裸机备份与恢复技术研究[J]. 计算机应用研究,2008,25(12):3735-3738.

[7] 任 欣,李 涛,胡晓勤. 远程文件备份与恢复系统的设计与实现[J]. 计算机工程,2009,35(10):112-114.

[8] 郭 京,李 涛,赵 奎,等. 基于 Linux 的异步网络备份系统的设计与实现[J]. 计算机工程,2006,32(8):280-282.

[9] 刘仕一,李 涛,刘哲智,等. 异地备份系统数据一致性检测方法[J]. 计算机工程与设计,2010,31(17):3766-3768.

[10] 邱龙金,刘晓洁,赵 奎. 安全的灾难备份系统[J]. 计算机工程与设计,2011,32(10):3258-3261.

[11] 敖 莉,舒继武,李明强. 重复数据删除技术[J]. 软件学报,2010,21(5):916-929.

[12] 浦海挺,李 涛,伍良富,等. Recovery Genius:Linux 下系统备份/恢复软件的实现[J]. 四川大学学报:工程科学版,2004,36(2):110-112.

[13] 熊 琦,王丽娜,王德军,等. 基于磁盘和 SAN 的网络数据备份模型[J]. 计算机工程,2007,33(4):233-235.

[14] VERITAS NetBackupbare metal Restore™ 6. 0. [EB/OL]. (2014-07-29) [2005-09-29]. http://www.symantec.com/business/su pport/resources/sites/BUSINESS/content/live/TECHNICAL_ SOLUTION/46000/TECH46415/en_US/281956.pdf.

[15] Symantec NetBackupbare metal restore 管理指南[EB/OL]. (2012-06-29) [2014-07-29]. <http://www.doc88.com/p-3127135888068.html>.

基于IP网络的Linux集群系统备份、恢复模型

作者：

[梁宝龙](#)，[王旭](#)，[叶钊](#)，[郑虹晖](#)，[程洪涛](#)，[LIANG Bao-long](#)，[WANG Xu](#)，[YE Fan](#)，[ZHENG Hong-hui](#)，[CHENG Hong-tao](#)

作者单位：

[海南省气象信息中心](#)，[海南 海口 570203](#)；[海南省南海气象防灾减灾重点实验室](#)，[海南 海口 570203](#)

刊名：

[计算机技术与发展](#)

英文刊名：

[Computer Technology and Development](#)

年，卷(期)：

2015(8)

引用本文格式：[梁宝龙](#).[王旭](#).[叶钊](#).[郑虹晖](#).[程洪涛](#).[LIANG Bao-long](#).[WANG Xu](#).[YE Fan](#).[ZHENG Hong-hui](#).[CHENG Hong-tao](#) [基于IP网络的Linux集群系统备份、恢复模型](#)[期刊论文]-[计算机技术与发展](#) 2015(8)