

文章编号：1005-8451（2017）03-0033-04

并行文件系统在车站系统备份中的应用

王 健，王启明

（北京铁路局 信息技术所，北京 100860）

摘要：分析车站系统数据库和应用程序备份需求，深入研究GPFS并行文件系统的基本框架和功能特点，将新投产主机加入到既有的GPFS集群，实现对集群内共享存储的实时并发访问，解决了车站系统生产数据第三方集中备份以及应用应急恢复问题，达到了整合备份资源、规避单点隐患、节约备份成本的目的，进一步确保了铁路的信息安全生产。

关键词：GPFS；高可用性；存储共享；数据备份

中图分类号：U291.6 : TP39 **文献标识码：**A

Application of general parallel file system in station system backup

WANG Jian, WANG Qiming

(Department of Information Technology, Beijing Railway Administration, Beijing 100860, China)

Abstract: This article analyzed the demand of backup for the database of station system and application program, studied on the basic framework and features of the general parallel file system(GPFS), added the new host to the existing GPFS cluster, implemented the real-time concurrency access to shared storage in a cluster. It was solved the problem of third party backup for the production data of the station system and application crash recovery, achieved the purpose of integrating backup resources, avoiding single point, hidden danger, saving backup cost, further ensured the safety production of railway information.

Keywords: general parallel file system(GPFS); high availability; storage sharing; data backup

随着铁路信息化的迅速发展，各种应用系统应运而生，其硬件架构多数采用双机加单存储模式，为了防止存储单点故障、同时满足各系统数据库和应用程序的备份需求，选用一套第三方集中备份环境，利用已经部署了GPFS软件的一套小型机（双机+双存储）作为服务节点，搭建高可用并行文件系统集群，逐步将新投产的车站系统小型机作为客户端节点加入其中，各节点能对集群中的两台存储进行并发双写，实现了车站系统存储空间集中共享及冗余备份。

1 GPFS架构

GPFS文件系统由3层架构组成：磁盘，网络共享磁盘(NSD)和GPFS文件设备，如图1所示。

2 方案实现目标

(1) 将车站系统主机节点加入到既有的GPFS集群，所有节点通过GPFS协议对集群内的共享存储进

收稿日期：2016-11-01

作者简介：王 健，高级工程师；王启明，高级工程师。

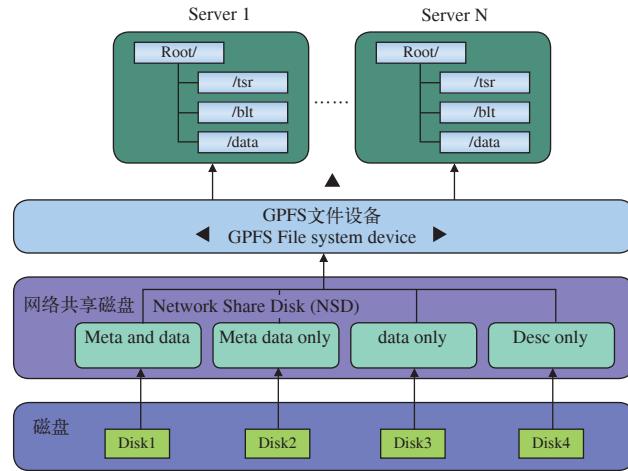


图1 GPFS 基本结构

行数据块级的访问，并且集群内的两台共享存储可镜像双写，达到备份冗余；

(2) 对共享存储进行合理空间分配，做到车站系统备份存储的集中共享，实现资源整合，提高设备利用率；

(3) 降低车站系统存储总体备份成本，减轻机房的配电压力，节省机房空间，减少系统管理人员工作量；

(4) 为车站系统的数据备份和恢复引入新的技术路线, 提高系统性能和管理水平。

3 方案实现步骤

3.1 GPFS集群规划

图2为GPFS集群方案逻辑结构图。地区数据中心处理系统是既有的GPFS集群,包含两台IBMP740小型机和两台EMC VNX5300存储,其中:Server1、Server2作为该集群的服务节点(即NSD Server或IO Server),通过光纤直接连接到SAN存储网络;车站系统1和车站系统2是两套车站系统。在Server1、Server2和这4个Client节点上进行相应的GPFS配置,使4台车站系统小型机作为NSD客户端通过以太网连接加入GPFS集群,从而能够在本地挂载GPFS文件系统,存储各自的应用程序和数据库RMAN备份,各节点主机配置如表1所示。

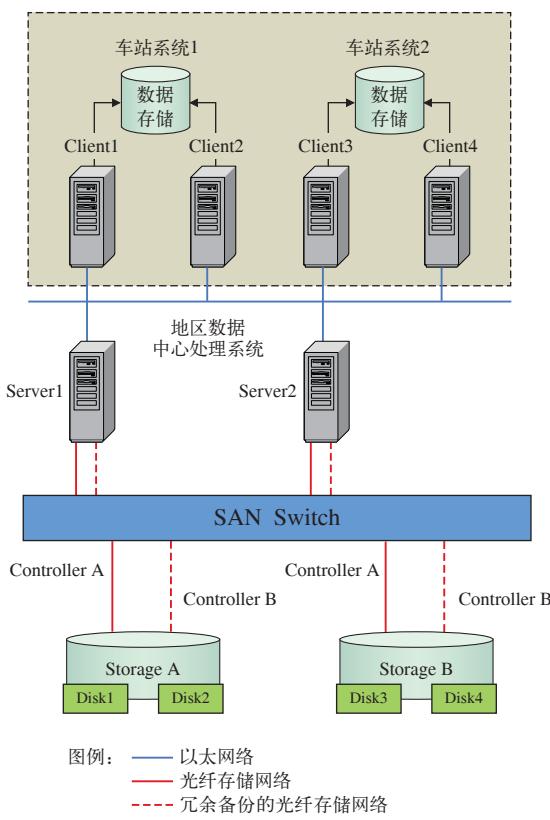


图2 GPFS集群规划示意图

3.2 GPFS集群扩充

3.2.1 前期准备

(1) 按照两套车站系统的实际应用及数据库使用空间,在GPFS集群两台共享存储上各划分4个

表1 机器配置表

序号	主机名	用途	IP地址	系统环境
1	Server1	主NSD服务器	192.16.1.1	P7408 CPU/48 G MEM/AIX 7102
2	Server2	备NSD服务器	192.16.1.2	P7408 CPU/48 G MEM/AIX 7102
3	Client1	车站系统服务器1	192.16.1.3	P74012 CPU/64 G MEM/AIX 7100
4	Client2	车站系统服务器2	192.16.1.4	P74012 CPU/64 G MEM/AIX 7100
5	Client3	车站系统服务器3	192.16.1.5	P7408 CPU/16 G MEM/AIX 6104
6	Client4	车站系统服务器4	192.16.1.6	P7408 CPU/16G MEM/AIX 6104

100 G 逻辑盘映射给 Server1、Server2 两个节点。

(2) 在车站系统 Client 节点上配置好 GPFS 集群内部使用 IP 地址。

(3) 修改所有机器的 /etc/hosts, 使得集群中涉及到的所有主机之间互相能解析主机名, 并配置好各节点的 RSH, 确保彼此之间通信正常。

(4) 编辑 /etc/profile 文件, 添加 GPFS 软件路径。

(5) 在所有的 Client 节点安装 GPFS 软件。

3.2.2 在GPFS主服务节点Server1上创建NSD磁盘

(1) 编写 NSD 磁盘配置文件, 然后通过 mmcrnsd 命令创建, 以车站系统 1 为例, 为其创建 4 块 NSD 磁盘。

(2) 创建 NSD 磁盘。

```
#mmcrnsd -F /etc/gpfs/cz1nsdfile -v no
```

(3) 通过 mmlsnsd 命令可以查看 NSD 磁盘
#mmlsnsd -aL。

(4) 同理, 为车站系统 2 创建 4 块 NSD 磁盘。

3.2.3 创建GPFS文件系统

(1) 编写 GPFS 文件系统配置文件, 以车站系统 1 为例, 从两台存储各取一块 100 G NSD 盘创建 /cz1smisbf1 文件系统, 实现存储双写, 同理, 创建 /cz1smisbf2 文件系统。

(2) 为车站系统 1 创建两个文件系统, 实际大小均为 100 GB :

```
#mmcrfsgpfs_cz1smisbf1 -F /etc/gpfs/cz1smisbf1.
```

```
cfg -B512K -Q yes -m 2 -r 2 -T /cz1smisbf1
```

```
#mmcrfsgpfs_cz1smisbf2 -F /etc/gpfs/cz1smisbf2.
```

```
cfg -B512K -Q yes -m 2 -r 2 -T /cz1smisbf2
```

3.2.4 在既有集群中增加Client节点

(1) 增加节点 Client1

```
#mmaddnode -N Client1
```

(2) 检查增加结果

```
mmchlicense client --accept -N Client1
```

(3) 同理，增加 Client2、Client3 和 Client4 共 3 个节点，并进行检查。

(4) 启动 4 个 Client 节点

```
#mmstartup -N Client1Client2 Client3 Client4
```

(5) 在节点上挂载新建的文件系统

```
#mmmount /cz1smisbf1/cz1smisbf2/cz2smisbf1/  
cz2smisbf2
```

(6) 检查节点状态

```
#mmgetstate-a
```

(7) 在 Client 节点上配置开机后自动启动 GPFS

```
#mmchconfigautoload=yes
```

(8) 查看 GPFS 集群状态信息

```
#mmlscluster
```

3.3 数据备份和恢复方法

3.3.1 数据备份方法

以车站系统 1 为例，将现车应用程序打包拷贝到 /cz1smisbf1 文件系统 smisbf 目录中，将 Oracle 数据库 RMAN 全库备份及归档文件路径指向 /cz1smisbf1 的 orabf 目录进行数据库备份，这样，应用和数据库的备份数据可直接存储 GPFS 文件系统上。

3.3.2 数据恢复方法

如果其自身的磁盘阵列损坏，系统不能识别共享卷组，生产数据全部失效，利用该 /cz1smisbf1 文件系统中的应用和数据库备份数据就地进行恢复：

(1) 卸载 /cz1smisbf2 文件系统。

```
#mmumount /cz1smisbf2-a
```

(2) 车站系统 1 数据库程序安装在本地 ROOTVG，数据库结构文件存储在 /oradata 文件系统下，通过在 GPFS 集群 Server 端变换 GPFS 文件系统挂载点的，将 /cz1smisbf2 挂载到 /oradata 目录下。

```
#mmchfs/dev/gpfs_cz1smisbf2 -T/oradata
```

```
#mmmount /oradata-a
```

(3) 将数据库启动到 MOUNT 状态，利用 RMAN 恢复管理器进行数据库恢复启动。

```
#rman target /
```

```
rman>startup mount
```

```
rman>restroe database;
```

```
rman>recover database;
```

```
rman>alter database open resetlogs;
```

(4) 利用 TAR 命令将 /cz1smisbf1 文件系统 smisbf 目录下备份的应用程序包进行解压，然后正常启动现车程序，应用恢复。

(5) 此时如果车站系统 1 本地存储恢复使用，先做一次数据库全库备份，然后停掉数据库，卸载 /oradata 文件系统，将原先的逻辑卷重新挂载，通过最新的全库备份在该文件系统上进行数据库恢复，将 /cz1smisbf1 文件系统中的应用程序恢复到原应用文件系统，即可完全恢复现场生产。

4 方案实现效果

通过实施此方案，将新投产小型机安装 GPFS 软件，作为 Client 节点加入到已有的 GPFS 集群环境，无需增加 HBA 光纤卡连接第三方 SAN 网络，不改变现有存储网络结构，而是通过以太网络连接 GPFS 集群，通过挂载新创建的 GPFS 文件系统，作为车站系统数据和程序备份环境使用。同时，模拟存储、文件系统或数据库故障，改变 GPFS 文件系统挂载点通过数据库 RMAN 恢复或文件拷贝方式进行应用恢复，经过试验证明，有了第三方集中备份环境，可节约备份和运维成本，GPFS 软件能实现对两台共享存储的并行双写，实现备份冗余。并且 GPFS 集群最多支持上千个主机节点，支持存储和节点的在线扩展，使用灵活、扩展节点实现简单。

5 结束语

通过采用 GPFS 软件搭建高可用并行文件系统集群，实现了车站系统存储空间集中共享及备份，既整合了资源，又提高了车站系统的安全保障。从 GPFS V3.4 开始增加 Windows 集群支持，使其适用范围进一步扩展，后续我们还会做进一步研究，为铁路的安全运输生产提供新的技术保障。

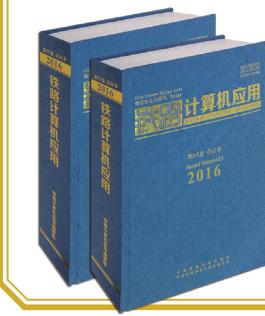
参考文献：

- [1] IBM Redbooks. Linux Clustering With Csm and Gpfs[Z]. IBM, 2004.
- [2] 文 平 .AIX UNIX 系统管理、维护与高可用集群建设 [M]. 北京：机械工业出版社，2012.
- [3] (美) 弗里曼 .Oracle Database 11g RMAN 备份与恢复 [M].

- 北京：清华大学出版社，2011.
- [4] 孙晓南. 网络存储与数据备份 [M]. 北京：清华大学出版社，2010.
- [5] 刘 峰. 现代铁路信息技术导论 [M]. 北京：中国铁道出版社，2010.
- [6] 魏新宇. GPFS 3_4 的配置变更与性能评估 [DB/OL]. [2012-10-22]. http://www.ibm.com/developerworks/cn/aix/library/1210_weixy_gpfs43/.
- [7] 谷 珊，帅 炜，陈志阳. 基于 Linux 集群环境上 GPFS 的问题诊断 [DB/OL]. [2010-05-14]. <http://www.ibm.com/developerworks/cn/linux/l-cn-gpfs-ts/>.
- [8] Boas Betzler,Nitin Koshti. 在多个 IBM Cloud 实例上创建 GPFS 集群 [DB/OL]. [2011-07-18]. <http://www.ibm.com/developerworks/cn/cloud/library/cl-GPFSclustersoncloud/>.

责任编辑 陈 蓉

《铁路计算机应用》 2016年合订本（限量版）出版发行



合订本为大16开精装本，全彩印刷，每册定价 **160** 元。

**限量发行 100 套
从速订阅**

订购热线：010-51849236
<http://www.tljsjyy.com>

微信公众订阅号：tljsjyy

企业专用铁路可视化运输生产调度指挥系统 (iTMIS) 免费应用

iTMIS 不仅是一套企业铁路运输生产管控系统软件，更是针对拥有专用铁路的大型企业“整合资源、流程优化，建设企业专用铁路运输管控一体化信息平台”的全面解决方案，它是企业专用铁路物流系统的 MES (生产执行管理系统) 。它可以使您在更短时间内，更好地完成企业对货物及铁路运输车辆的实时管控，降低铁路车辆停时费用，加快企业过轨自备车的周转，提高运输生产作业的效率，降低企业的运输成本，并为企业各部 门管理及决策提供系统的信息服务与决策支持。该系统获得 2010 年度国家科技部创新基金的无偿资助。

推出基于互联网+的合作方式：

iTMIS 系统免费使用，有偿运维！

1. 无需投入，保证第一年至少降低 10% 以上的铁路延时费。 2. 无建设周期，投入即使用，专业运维、贴心服务。
3. 优化人员构置，可替代或精简部分岗位人员工作，降低企业人力成本。 4. 低运维成本，高效益回报。



成都劳杰斯信息技术有限公司

研发中心：西南交通大学科技园现代工业中心 A 座 211 号

营销中心：成都市二环路北一段 10 号万科加州湾 V 派 1712 室

邮箱：logistics_it@126.com

邮编：610031

咨询电话：18980847041 18980847040 服务热线：400-000-5129