

民航气象数据库系统典型磁盘故障分析

刘希之

(中国民用航空中南地区空中交通管理局湖北分局, 湖北 武汉 430000)

摘要:影响民航气象数据库系统运行性能的一个重要因素是数据存储性能。阐述了独立冗余磁盘阵列(redundant arrays of independent disks, RAID)的工作原理及其在民航气象数据库系统数据存储设备中的应用,并结合典型磁盘故障案例,详细阐述故障处置过程,最后对磁盘出现故障的原因进行分析,并对日常工作提出检修意见。

关键词:存储性能;RAID;故障分析

中图分类号:TP311.13

文献标识码:A

文章编号:1004-7344(2023)17-0148-03

0 引言

民航气象数据库系统通过收集存储全国各地机场的气象情报和气象资料,为民航客机飞行提供及时、准确、规范的气象服务,是集通信传输、数据处理、信息服务为一体的数据库应用系统。数据存储设备性能的好坏、访问速度快慢和可靠性的高低直接影响民航气象数据库的运行效率和运行质量。近几年,随着民航气象数据库所发挥的作用越来越大,逐渐成为航空气象服务中重要的组成部分,其所能提供的信息资料在不断完善,具体如民航报、卫星云图资料、9210资料等多达数十种的信息。民航气象数据库最早在1994年12月建立,最初是由7个民航地区气象中心共同组成,并在1999年1月的时候投入使用。

本文主要研究独立冗余磁盘阵列(redundant arrays of independent disks, RAID)在民航气象数据库系统中的应用,详细分析了民航气象数据库系统中磁盘故障产生原因,并给出典型磁盘故障处置过程,最后结合故障实例和过往研究结论提出关于数据库系统磁盘维护的建议,为民航气象机务人员在数据库系统维护中提供参考。

1 RAID技术在民航气象数据库系统的应用

1.1 RAID5技术概述

磁盘及磁盘阵列是民航气象数据库系统的底层存储设备。磁盘主要是利用电磁特性在磁片上记录信息,将磁头定位到指定位置进行记录的读写操作,在需要读写信息的时候通过磁臂和磁片的旋转进行。而RAID^[1]则是通过冗余容错技术,将多个独立的物理磁盘按照不同的方式组合成一个存储性能优于单一磁盘的磁盘组,同时也能提高数据的可靠性和可用性。

不同的数据和冗余信息在磁盘阵列上的组织方

式,在数据存储性能和存取速度等方面都存在着差异。RAID5是一种采用块交叉和奇偶校验信息旋转分布的磁盘阵列。它是将组成RAID5的每一块磁盘上分别存储数据和每一块磁盘所产生的奇偶校验信息,并将奇偶校验信息和对应的数据分别存储在不同的磁盘上,其中完整的数据存储在任意N-1的磁盘上,即有一个空间与一块磁盘的容量相当,用于奇偶校验信息的存储。以4个硬盘组成的RAID5为例,其存储方式如图1所示。

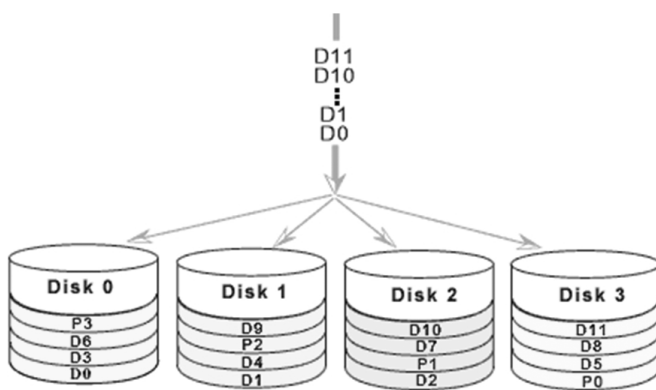


图1 RAID5存储方式

图1中,P0为D0,D1和D2的奇偶校验信息,P1为D0,D1,D3的奇偶校验信息,P2为D0,D2和D3的奇偶校验信息,P3为D1,D2和D3的奇偶校验信息。从图中可以看出,RAID5并没有对存储的数据进行备份,而是在组成RAID5的每一块磁盘上都存储了数据及对应的奇偶校验信息。由此可见,当RAID5的一个磁盘出现损坏时,数据的完整性不会受到影响,从而使数据的安全性得到保障。当损坏的盘面被替换掉后,RAID也会自动在这个盘面上利用剩余的奇偶校验信息进行数据重建,从而保持RAID5的高可靠性。

1.2 RAID5 技术在民航气象数据库系统中的应用

1.2.1 民航气象数据库系统构成

民航气象数据库系统是一套建立于 IBM 双机服务器上、Oracle 数据库管理系统进行管理、一个 24h 不间断运行的规模庞大的实时联机气象信息系统,由北京、广州、上海、沈阳、成都、西安、乌鲁木齐 7 个地区气象中心以及 37 个空管中心站气象数据库系统组成。气象信息系统三级网络如图 2 所示。

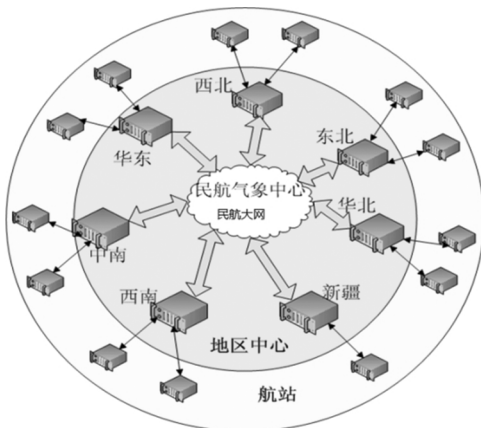


图 2 气象信息系统三级网络

民航气象数据库系统的硬件构成中,为了确保交换资料的一致性,软件的配备需要完全相一致,但是软件的功能配置可以根据实际的要求进行相应的替换,以第三级数据库系统的配置为例。三级数据库主要由一台 IBM system p5 作为主服务器,一台 IBM system p5 作为应用端口的服务器,接着采用 CISCO2960 交换机用于核心交换机,CISCO2851 路由器和本地区的气象中心数据库相连接。在软件结构上,主要由通信分系统、资料处理子系统与数据库管理子系统等构成。

(1)通信分系统。通信分系统在气象数据库系统中主要负责用于接收各类的气象资料,以及对资料进行检查、处理与发送。同时,能够负责编写接收到的各类气象资料、转发、告警和查询等多个功能。为了确保能够获取全面的民航气象数据,通信分系统需要 24h 持续运作,实时地对接收过来的资料和信息进行处理,并将处理好的资料按照统一的格式上交到上层系统等待下一步的处理。

(2)资料处理子系统。资料处理子系统主要根据系统设定好的程序,将通信分系统中发过来的各类资料,按照一定的形式进行分析、要素分解并完成入库。在资料处理子系统中,共含有 12 个进程,各个处理的进程都由一个总的监控进程实施监控,确保资料的处理持续运作。

(3)数据库管理子系统。该系统中主要采用的是客

户机服务形式,能够全天候监管资料处理与数据库等系统。其主要的功能包括:登录系统、处理资料、管理数据库、恢复数据库与系统配置、窗口等。其中,数据库管理子系统中的告警功能能够对系统运作中存在的服务器端运行异常、监控进程运行异常等进行告警,从而帮助工作人员及时发现问题。

(4)预报业务应用平台与气象服务应用平台。预报业务应用平台建立主要的作用是为了帮助预报人员提供综合信息咨询等。其底层属于是数据库系统,通过预报业务应用平台,既可以进行数据的检索,获取各类资料,也可以将平台中所生成的产品通过通信分系统进行存档。

气象服务应用平台的目的是为了用户更为便捷地进行信息的浏览、检索,更好地推动气象服务的标准化。

1.2.2 气象数据库系统的维护与管理

(1)通讯机的维护。在维护通讯机的时候,可以通过运用命令 `dsqmq` 检查 MQ 管理器是否已运行起来了,如果还没有运行起来的话,可以通过输入专门的命令启动队列管理器。

在检查 FTP 服务是否运行正常的时候,可以通过输入 `ftp 172.22.130.2` 查看当前 FTP 服务是否在正常运作。

(2)服务器的维护。为了确保资料能正常接收,查看服务器 `data` 目录下子目录中的产品接收是否处于正常的状态。

(3)控制平台与网络的维护。控制平台的维护措施中,主要通过运行 `CTMS.EXE` 软件检查进程状态、通道状态是否有告警信息。

网络维护中,主要通过通过对路由器、防火墙面板指示灯等进行检查,确认是否正常。

1.2.3 在民航气象数据库系统中 RAID5 技术的应用

以湖北空管分局气象台为例。气象数据库系统配置 1 台通信前置服务器,安装接收、检查、转换各类气象资料的 Redhat Linux 系统,并将数据按预先设定的规则上传至中南地区中心,实现分局站与地区中心之间的数据交换;配置 3 台 IBM P520 数据库服务器,用于实现气象资料的入库和管理。其中,数据库服务器磁盘阵列配置采用 RAID5 技术,配置 4 块 146.5G 的 IBM 硬盘。由上述可知,RAID5 技术相当于只使用一块磁盘的存储容量用于存放校验数据,其他存储容量用于存储数据,且允许单磁盘故障,出现故障时,磁盘阵列控制器可通过冗余差错技术从其他硬盘中恢复/重建生成故障盘的丢失数据。

因此,在高可靠性、高安全性的民用航空气象数据库系统中,广泛应用了兼顾存储性能、数据安全和存储成本的 RAID5 技术。

2 典型磁盘故障

根据磁盘厂商公布的磁盘数据表显示,磁盘的年故障率在 0.58%~0.88%^[2]。然而美国卡内基梅隆大学并行数据实验室的研究结果却指出在真实应用环境中,磁盘的年更换率往往都超过了 1%,多数情况下为 2%~4%,在某些系统中甚至高达 13%,远远高于磁盘厂商所公布的磁盘年故障率,而且磁盘的年更换率随着磁盘使用年限的增长而持续增长^[3]。

除了磁盘的完全故障外,潜在扇区错误也是一种常见的磁盘错误,它是指磁盘内部发生了扇区错误但是磁盘没有及时侦测到,直到用户下次访问时才发现。也就是说当存储系统中一个磁盘发生故障时,邻近的磁盘在未来较短时间内发生故障的概率将大大提高。对于 24h 不间断运行的民航气象数据库系统而言,磁盘时刻处于读写状态之中,磁盘故障概率也大大提高。下面将结合典型故障案例,分析磁盘故障过程,给出气象数据库系统磁盘维护建议。

(1)现象:通信前置服务器至某台数据库服务器的 MQ 通道异常。

(2)处置过程:①检查通信前置服务器网络及硬件状态,正常。②检查通信前置服务器至数据库服务器的网络,正常。③检查数据库服务器网络状态,正常。④检查数据库服务器硬件状态。使用 AIX 系统的 SMIT 工具 smitty pxdam->list-PCI,进入磁盘状态界面,发现磁盘 pdisk0 状态显示为 failed,磁盘阵列状态显示为 degraded。⑤热插拔更换磁盘。使用 SMIT 工具 smitty pxdam->Diagnostics and Recovery Options->确定 pdisk0 的卡槽,使用热插拔操作指引更换磁盘后进行数据重建。⑥数据重建过程中多次跳转至 1%重新开始。检查数据同步报错信息,发现均为磁盘 pdisk3 导致,但磁盘 pdisk3 状态显示为 actived,怀疑为磁盘 pdisk3 存在潜在扇区故障导致数据重建失败。⑦该台数据库服务器磁盘阵列采用 RAID5 技术,只允许单盘故障,因而不支持同时更换两块磁盘。在进行多次数据重建后,数据同步成功,随即使用热插拔操作更换存在潜在扇区故障的磁盘 pdisk3,并进行数据重建。⑧数据重建成功后,重启通信前置服务器至数据库服务器的 MQ 通道,成功。

此次数据库系统故障由数据库服务器磁盘故障引起,且在磁盘阵列数据恢复过程中,由于邻近磁盘 pdisk3 存在潜在扇区故障,从而导致数据恢复时间拉长。美国卡内基梅隆大学的一项研究曾指出,在磁盘阵

列数据恢复过程中,数据恢复时间长,第二个磁盘发生故障的概率越大。数据显示数据恢复时间为 1h,第二个磁盘发生故障的概率为 0.05%;数据恢复时间 3h,第二个磁盘发生故障的概率为 1.0%;数据恢复时间 6h,第二个磁盘发生故障的概率达到 1.4%^[4]。

(3)故障总结及建议。

由此可见,潜在扇区故障的磁盘会大大影响数据库系统的可靠性和可用性。为避免磁盘同时故障导致的 RAID5 阵列崩溃,根据此次磁盘故障过程,结合过往研究结论,给出以下 4 个建议:①磁盘型号尽量选择不同品牌相同容量或者相同品牌不同批次,可以在一定程度上避免 2 块甚至 2 块以上磁盘同时故障。因为相同品牌同一批次的磁盘,几乎是同一流水线上的产品,老化程度几乎一致,在磁盘使用环境、读写频次几乎一致的情况下,磁盘出现同时故障的概率较高。②为避免在数据重建过程中,由于潜在扇区故障磁盘的存在,导致数据重建失败,进而导致原始数据丢失,在进行数据重建之前应做好数据备份。③不要随意移动磁盘。磁头和磁盘间的间隙仅有 0.015~0.025 μm ,如果磁盘受到强烈颠簸,则会导致磁盘产生物理坏道。④定期扫描磁盘,检测出存在潜在扇区故障的磁盘^[5]。

3 结语

本文详细介绍了 RAID5 工作原理及其在民航气象数据库系统的应用,从一个典型故障案例入手,详细阐述故障处置过程,分析磁盘故障原因,并结合此次故障原因和过往研究结论,提出了民航气象数据库系统磁盘维护工作的建议,为民航气象机务人员维护工作提供参考。

参考文献

- [1] 乔旭坤,李顺,李君,等.基于机器学习的硬盘故障预测研究[J].计算机技术与发展,2022,32(6):215-220.
- [2] 吴文波.基于大数据技术的服务器磁盘故障诊断方法[J].信息记录材料,2022,23(6):170-172.
- [3] 王辉登.基于无线传感网络的网络磁盘设备故障监测方法研究[J].信息与电脑(理论版),2022,34(1):38-40.
- [4] 朱洪斌,李悦.国产人工智能平台的磁盘故障预测应用[J].电子世界,2021(18):19-20.
- [5] 李新鹏,高欣,何杨,等.不平衡数据集下基于自适应加权 Bagging-GBDT 算法的磁盘故障预测模型[J].微电子学与计算机,2020,37(3):14-19.

作者简介:刘希之(1995—),女,汉族,湖北黄石人,助理工程师,硕士研究生,主要从事民航气象机务工作。