

地震资料处理环境中数据写策略研究^①

金 弟, 庄锡进, 王启迪, 曹晓初, 王宗仁

(中国石油杭州地质研究院 计算机应用研究所, 杭州 310023)

摘 要: 在地震资料处理环境中, 用户频繁地大规模地震数据写操作后, 针对文件系统池中存在地震数据写与分布不均衡的缺陷, 剖析地震资料处理流程, 探讨分析各种常规地震数据写策略, 提出了一种基于概率的地震数据写策略, 以满足文件系统池在不同初始状态下地震数据写均衡与优化分布. 实验结果验证了其优越性.

关键词: 地震资料处理; 文件系统池; 地震数据写策略; 分布优化

引用格式: 金弟, 庄锡进, 王启迪, 曹晓初, 王宗仁. 地震资料处理环境中数据写策略研究. 计算机系统应用, 2018, 27(2): 58-63. <http://www.c-s-a.org.cn/1003-3254/6181.html>

Research on Data Writing Strategy in Seismic Data Processing Environment

JIN Di, ZHUANG Xi-Jin, WANG Qi-Di, CAO Xiao-Chu, WANG Zong-Ren

(Department of Computer Application, Research Institute of Hangzhou Geology, Hangzhou 310023, China)

Abstract: In seismic data processing environment, the seismic data writing has imbalance without seismic data distribution optimization in the file system pool when large scale seismic data frequently perform writing and deleting operations. This study analyzes the seismic data processing classification and various conventional writing strategies of seismic data. It proposes a seismic data writing strategy based on the idea of probability which meets the optimization distribution of seismic data writing under different initial files in system pool. The experimental results show that the method is effective.

Key words: seismic data process; file system pool; seismic data writing strategy; distribution optimization

在油气勘探过程中, 利用人工地震波激发采集的地震资料是一种海量数据^[1]. 利用高性能计算集群系统与超大规模并行处理应用软件对地震资料数据进行叠前偏移成像处理过程中, 负责地震资料大数据读写等操作的存储子系统成为性能瓶颈的核心环节.

在存储子系统中, 当地震资料处理系统中的计算节点客户端大规模 I/O 并行请求地震数据写所形成的高聚合 I/O 带宽时, 存储主机集群技术^[2] 能避免存储主机节点端出现拥堵与单点故障; 动态存储多路径技术^[2] 能有效解决在存储主机与存储控制器之间的数据链路通道出现拥堵与单点故障; 并行文件系统减少在 Linux 操作系统环境下的文件系统层出现大量文件读写的

I/O 排队等待时间^[2]. 但是这些技术解决的是存储子系统体系结构中单个文件系统层及以下底层的 I/O 性能提升. 而对于文件系统池中各个文件系统的层面上, 存在各个文件系统的的海量数据写吞吐量、写频率、写负载以及写容量等差别较大; 各个文件系统的存储剩余容量、存储剩余容量百分比值等存在较大差异; 地震数据文件容量、数量等在各个文件系统上的存储不均衡. 因此无法实现文件系统池中各个文件系统之间的地震数据写负载均衡与数据分布优化. 本文首先阐述地震资料处理流程, 接着分析不同常规地震数据写策略的优缺点, 最后提出基于概率的地震数据写策略的优越性并进行实验验证.

① 收稿时间: 2017-04-25; 修改时间: 2017-05-09; 采用时间: 2017-05-19

1 地震资料处理流程

1.1 系统典型部署

地震资料处理是典型的数据密集型应用,其特点是对陆地、海洋采集的原始海量地震数据进行常规处理、偏移成像处理等形成准确的地下构造图像,以便发现油气藏.目前主流的地震资料处理环境中,普遍采用基于集群的高性能并行计算机系统^[3],实现其计算与存储分离、存储共享的方式,典型的系统结构如图1.由图1知,海量地震数据存储于文件系统池中,计算节点通过高速网络向I/O节点请求I/O读写,多个I/O节点并行对文件系统池读写地震数据.文件系统池由多个文件系统构成,通常采用高性能磁盘阵列、并行文件系统构建,各个节点像访问多个本地文件系统一样来读写地震数据.地震资料处理系统的软件体系结构采用分层模式,如图2所示.应用功能层部署在多个计算节点上,完成常规处理、并行偏移处理等.服务控制层是其核心服务器,执行整个软件系统的控制与管理等功能.数据服务层部署在I/O节点上,进行地震数据与非地震数据的读写与管理,直接与数据存储层的文件系统池与关系数据库进行数据交互.

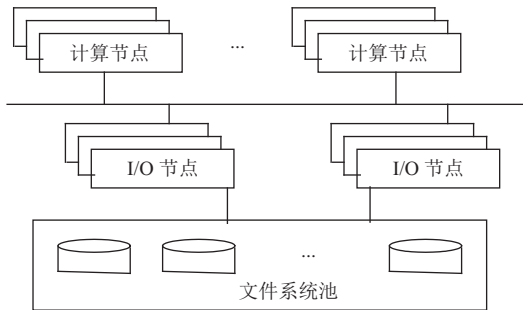


图1 地震资料处理系统典型结构

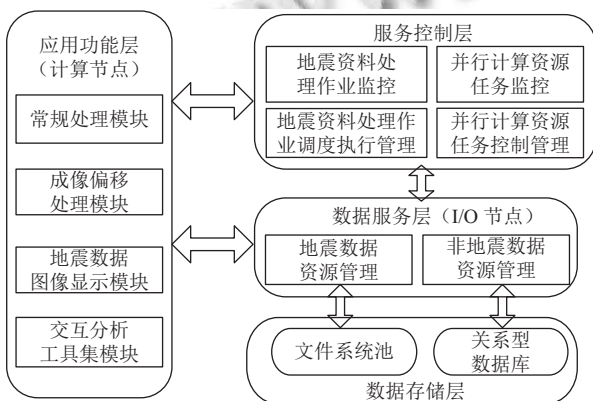


图2 地震资料处理系统的软件体系结构

作者所在单位从2010年开始采用并行文件系统StorNext^[4]部署文件系统池,目前的文件系统池由21个文件系统组成,规模达到PB级存储容量.采用StorNext部署的文件系统池能很好的实现单个文件系统内地震数据的均衡物理分布与多I/O节点并行高效读写.但在一个地震资料处理作业过程中,对请求写的地震数据必须预先静态的指定向哪个文件系统写,无法根据多个文件系统的的海量地震数据分布信息变化动态的进行修改调整.所以要满足多个文件系统之间的地震数据负载均衡写,数据服务层与数据存储层不能感知,需要依靠地震资料处理人员的主观判断与手工操作干预.

1.2 流程分析

地震资料处理的主要流程如图3所示.检波器产生的原始地震数据存储于文件系统池中;地震资料处理软件通过解编生成特定私有格式地震数据;读取解编后的地震数据,经过常规处理与偏移处理产生若干份中间成果地震数据;读取中间成果地震数据,进行交互分析,产生最终处理后的成果地震数据.

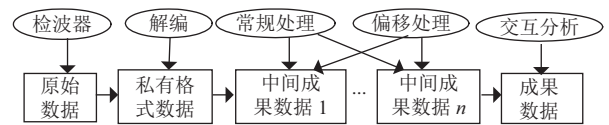


图3 地震资料数据处理流程

由图3知,地震资料处理流程主要包括原始地震数据解编、常规处理、偏移处理以及交互分析4个主要阶段.原始地震数据解编消耗CPU较少,对I/O需求较大.常规处理包括静校正、反褶积、噪音压制等方法,需要多次读入、计算、删除和写大量中间地震数据,对I/O和CPU需求都比较大.偏移处理包括时间偏移、深度偏移、逆时偏移等复杂方法,需要多次并行读入、计算、删除和写大量中间地震数据,对存储子系统的并行I/O与计算子系统的并行计算要求非常高.交互分析是读取常规处理和偏移处理后的相关地震数据,做分析修正.对I/O和CPU消耗不多.因此除交互分析外,其它三个阶段都有大规模的读、写及删除大量地震数据文件.这种频繁的大规模地震数据写、删除等操作导致文件系统池中各个文件系统之间的地震数据写不平衡与地震数据分布不平衡.

2 常规地震数据写策略与分析

针对文件系统池中各个文件系统之间的地震数据写与分布不平衡,结合地震资料处理的4个重要阶段

及其特点,常规地震数据写策略主要有三种。

2.1 最长时间间隔写策略

对于文件系统池中所有可用的文件系统,根据文件系统写操作的时间戳,最近一次被选中写操作以后的时间间隔值作为选择某个文件系统进行写的依据^[5],最长时间间隔值的文件系统优先被选中,确保各个文件系统之间地震数据写负载均衡、地震数据分布优化。

这种写策略的优点是实现方法简单。缺点是由于没有考虑文件系统的总容量与可用剩余容量等因素。在文件系统池中,当大容量文件系统与小容量文件系统混合、各个文件系统之间的剩余容量差值较大的情况下,大量的地震数据写操作会存在小容量文件系统优先写满,导致可用文件系统快速减少。同时对于成批大容量地震数据文件写操作时,小容量文件系统容易出现重负载,大容量文件系统容易出现轻负载。因此该策略在地震数据写操作容量规模不大的情况下效果比较好。数据解编是将野外采集的地震数据正确加载到地震资料处理软件系统中,进行观测系统定义,并对地震数据进行编辑和校正的过程;交互分析主要对地震剖面、地震速度谱等进行分析调整。这二个阶段的地震数据写规模比较小,并且是串行写操作,比较适合使用该地震数据写策略。而常规处理、偏移处理会产生大规模地震数据写现象,不宜使用该地震数据写策略。

2.2 剩余容量最大百分比写策略

针对最长时间间隔写策略的缺陷,在可用的文件系统池中,选取可用剩余容量最大百分比的文件系统进行地震数据写操作。单个文件系统剩余容量百分比=(单个文件系统总容量-单个文件系统已被使用容量)/单个文件系统总容量*100。与最长时间间隔写策略相比,该策略考虑了各个文件系统的使用状态,即可用剩余容量的信息,能确保文件系统池中,各个文件系统的可用剩余容量的百分比保持均衡。避免小容量文件系统优先写满,达到地震数据写平衡与优化分布。

该策略也存在缺点,当文件系统池中一个文件系统的可用剩余容量百分比明显高于文件系统池中其它所有的文件系统的时候,会产生单个文件系统写热点瓶颈问题。例如当前文件系统池中有3个文件系统d1、d2、d3,其可用剩余容量的百分比为50%、40%、40%。由于文件系统池扩容,新增一个新的文件系统d4,其可用剩余容量的百分比为99%。在这种情况下,以后所有的数据写操作都选择文件系统d4,其它

的文件系统d1、d2、d3均处于空闲状态,直到文件系统d4的可用剩余容量的百分比为50%以下。因此该策略比较适合在文件系统池中,各个文件系统的剩余容量最大百分比在初始状态时就比较接近的情况。在地震资料处理环境中,只要文件系统池的状态满足这个条件,地震资料处理的4个主要阶段都可以使用该地震数据写策略。但在实际地震资料处理流程中,由于常规处理、偏移处理这二个阶段存在反复多次大规模中间结果的地震数据输出,而且这二个阶段也会存在自身多次执行,要满足该数据写策略的条件比较困难。因此该地震数据写策略对于地震数据写规模较小的数据解编、交互分析可以使用,而对地震数据写规模比较大的常规处理、偏移处理不适合。

2.3 基于时间与容量因素的写策略

2.1节的写策略仅考虑时间因素,存在小容量文件系统优先写满、文件系统写操作负载不均衡的缺点,2.2节的写策略仅考虑容量因素,存在单个文件系统写热点瓶颈问题。基于时间与容量因素的写策略同时考虑各个文件系统的地震数据写操作最近一次被选中以后的时间间隔值这个时间因素和各个文件系统可用剩余容量值这个容量因素,结合加权可调参数计算出最终值作为选择文件系统进行地震数据写操作的判别依据。

在文件系统池中,假设文件系统的数量为 n ,四个可调参数为 a, b, c, d 。文件系统池容量集合 $S_n = \{s_1, s_2, \dots, s_n\}$;文件系统池可用剩余容量的百分比集合 $B_n = \{b_1, b_2, \dots, b_n\}$;文件系统池可用剩余容量集合 $V_n = \{v_1, v_2, \dots, v_n\}$,其中 $v_i = b_i \times s_i$;文件系统池最近一次被选中以后的时间间隔集合 $T_n = \{t_1, t_2, \dots, t_n\}$;基于时间与容量因素的文件系统池写策略最终值集合 $W_n = \{w_1, w_2, \dots, w_n\}$,其中 $w_i = \{a + b \times t_i\} \times \{c + d \times v_i\}$ 。参数 a, b 调整的是时间因素的权重,参数 c, d 调整的是容量因素的权重。通过参数值调整,其适应性比较强,例如增大参数 c 及 d 的值,同时减少参数 a 及 b 的值,此时该写策略趋向于基于最长时间间隔写策略^[6]。

该写策略同时考虑了时间因素与容量因素,且通过加权可调参数,避免了2.1节与2.2节策略的缺点,理论上能适合应用于地震资料处理的4个主要阶段。但该策略也存在缺陷,由于引入了4个可调参数,存在需要手工干预而且调整起来相对较复杂,在实现过程中,4个参数的优化组合存在一定的难度与需要有一定

的经验,可操作性差.同时该策略在参数的调整过程中,时间因素与容量因素的权值都不高的这些文件系统容易出现饥饿现象,即在那些文件系统上的写地震数据量很小或没有写地震数据.因此在地震资料处理流程中,交互分析阶段地震数据量写较小,该策略的4个参数的选择相对较易,可以适用.而数据解编、常规处理以及偏移处理这三个阶段的地震数据量写逐渐变大,尤其是偏移处理都是并行地震资料处理作业,涉及到多次并行地震数据写操作,该策略的参数调整组合优化比较困难,不太适合.

3 基于概率的地震数据写策略

该策略运用基于概率的思想、方法,通过计算概率值作为优先级别值来判别优先选择哪个文件系统.根据文件系统的写频繁程度、文件系统写操作的动态容量状态信息进行统筹优化考虑.采用最长时间间隔、绝对剩余容量及相对剩余容量三个关键变量因素分别计算概率值,然后根据地震资料处理环境中4个主要阶段的实际需求,通过权重因子参数进行调整这三个变量概率值的权重,来满足文件系统池的地震数据写平衡与优化分布.该策略实现方法过程如下.

文件系统池中,假设文件系统的数量为 n ;文件系统池可用文件系统集合 $D_n = \{d_1, d_2, \dots, d_n\}$;文件系统池最近一次被选中以后的时间间隔集合 $T_n = \{t_1, t_2, \dots, t_n\}$;文件系统池基于最近一次被选中以后的时间间隔概率集合 $PT_n = \{pt_1, pt_2, \dots, pt_n\}$;在文件系统池中,基于最长时间间隔的第 i 个文件系统概率计算公式为:

$$pt_i = t_i / \sum_{i=1}^n t_i \quad (1)$$

文件系统池容量集合 $S_n = \{s_1, s_2, \dots, s_n\}$;文件系统池可用剩余容量百分比集合 $B_n = \{b_1, b_2, \dots, b_n\}$;文件系统池绝对剩余容量概率集合 $PS_n = \{ps_1, ps_2, \dots, ps_n\}$;在文件系统池中,第 i 个文件系统绝对剩余容量概率计算公式为:

$$ps_i = (s_i \times b_i) / \sum_{i=1}^n (s_i \times b_i) \quad (2)$$

文件系统池可用剩余容量最小百分比 $B_{\min} = \text{MIN}\{b_1, b_2, \dots, b_n\}$;文件系统池可用剩余容量最大百分比 $B_{\max} = \text{MAX}\{b_1, b_2, \dots, b_n\}$;文件系统池可用剩余容量

的百分比和如下:

$$b_{\text{sum}} = \sum_{i=1}^n b_i$$

文件系统池可用相对剩余容量概率集合 $PB_n = \{pb_1, pb_2, \dots, pb_n\}$;在文件系统池中,第 i 个文件系统可用相对剩余容量概率计算公式为:

$$pb_i = (B_{\max} - b_i) / (n \times B_{\max} - B_{\text{sum}}) \quad (3)$$

定义概率权重因子 WF , $0 \leq WF \leq 1$;文件系统池内选择文件系统进行地震数据写的概率集合 $P_n = \{p_1, p_2, \dots, p_n\}$;根据式(1)、(2)、(3),在文件系统池中,第 i 个文件系统进行地震数据写的概率计算公式如下:

$$p_i = [(1-WF)/2] \times pt_i + [(1-WF)/2] \times ps_i + WF \times pb_i \quad (4)$$

公式(4)由三项构成,第一项用时间比例方法计算基于最长时间间隔的文件系统概率值 pt_i ,所占的权重是 $(1-WF)/2$;第二项用剩余容量比例方法计算文件系统绝对剩余容量概率值 ps_i ,所占的权重是 $(1-WF)/2$;第三项用最值差方法计算文件系统可用相对剩余容量概率值 pb_i ,所占的权重是 WF .第 i 个文件系统进行地震数据写的最终概率值由这三项各自的概率值加权再相加得到.对于 WF 取任何0至1之间的值,三项的权重和为1,各项所占的权重比值由 WF 控制与调整.

该策略考虑文件系统池的最长时间间隔、绝对剩余容量及相对剩余容量分别作为概率变量的分布值,其值为在文件系统池中选择哪个文件系统进行地震数据写的主要判别依据.在此基础上,运用调整概率权重因子 WF 这个参数的值大小,权衡这三个主控因素的概率变量值权重,最终达到文件系统池内各个文件系统之间地震数据写平衡与优化分布.而且该策略仅引进一个概率权重因子参数 WF ,参数优化调整相对简单.该策略实现算法流程如图4.

4 实验及结果分析

为了验证本文提出的基于概率写策略的效果,对该写策略进行实验测试分析.实验环境如下,16个IBM HS22刀片计算节点,4个IBM x3650 I/O节点,StorNext4.0并行文件系统构建的文件系统池,采用地震资料处理软件GeoVation2015^[6]的并行地震数据写软件子模块模拟用户频繁的大规模地震数据写操作.选择国内某盆地采集的实际地震数据作为实验数据,其信息如表1所示.为方便测试,采用4个文件系统fs1、fs2、fs3及fs4组成的文件系统池,4个文件系统

的初始状态信息如表2;抽取所测试的地震数据文件样本如表3,60个地震数据文件共计约1TB地震数据容量.

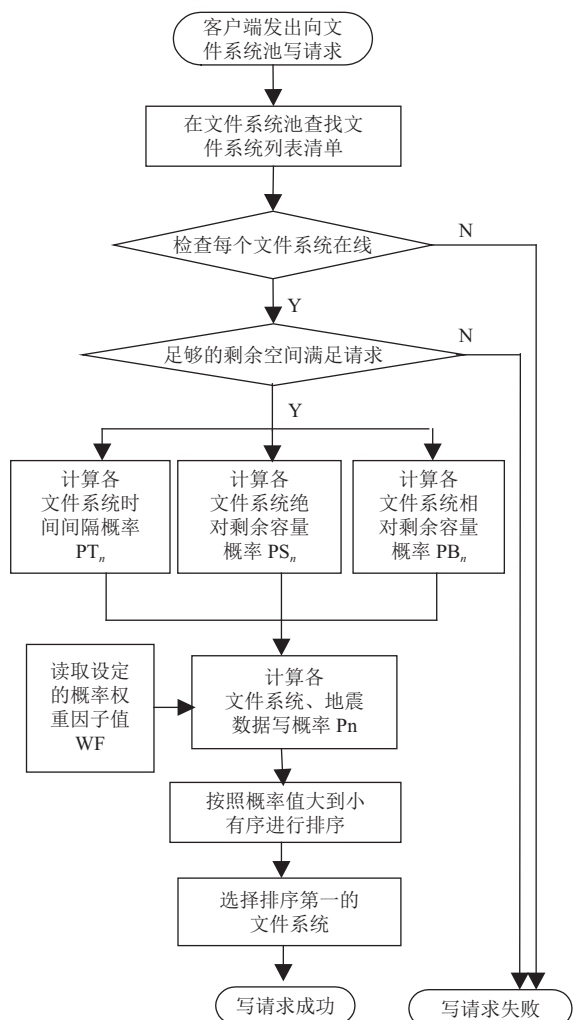


图4 基于概率地震数据写策略算法流程

表1 实验地震数据信息

参数名称	参数值	参数名称	参数值
覆盖面积	120 km ²	线束	40
总炮数	18 500 炮	接受线	55
覆盖次数	40	采样间隔	10 ms
面元	25*25 m	偏移深度	1200 m

表2 文件系统池初始条件状态

文件系统名称	总容量(GB)	剩余容量百分比(%)
fs1	600	80
fs2	600	80
fs3	600	40
fs4	200	90

表3 地震数据文件样本信息

类型	单个文件大小范围	数量
二维地震数据	20 MB~50 MB	20
二维地震数据	400 MB~600 MB	20
三维地震数据	30 GB~60 GB	20

4.1 文件系统写地震数据对比

图5给出了在不同概率权重因子值条件下,1TB地震数据写操作完成后,4个文件系统被写地震数据的容量分布.文件系统池初始状态fs4剩余容量180GB,fs1与fs2的剩余容量480GB,fs4与fs1(fs2)剩余容量差值比较大,2倍多.由图5可知,对于不同的概率权重因子值,向fs1写地震数据容量分布范围379.8GB至390.6GB;向fs2写地震数据容量分布范围372.6GB至386.4GB;向fs3写地震数据容量分布范围142.2GB至156.6GB;向fs4写地震数据容量分布范围81.4GB至128.8GB.

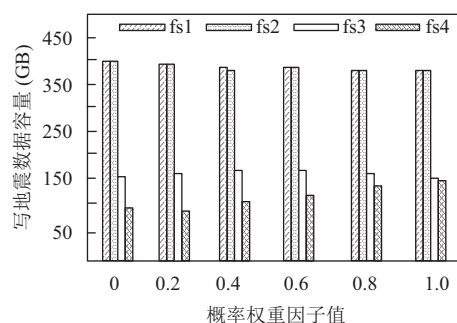


图5 写地震数据容量分布

根据不同文件系统的写地震数据容量分布范围,可知该策略避免2.1节所描述的在各个文件系统之间剩余容量差值较大的条件下,小容量文件系统优先写满情况.也没有出现小容量文件系统容易出现重负载,而大容量文件系统容易出现轻负载的现象.同时在不同权重因子值条件下,写地震数据容量最小的文件系统也有81.4GB数据写入,避免2.3节所描述某些文件系统可能出现写地震数据容量很小或没有写地震数据这种饥饿现象.

4.2 文件系统剩余容量百分比对比

文件系统池初始状态中,剩余容量百分比最小的fs3值为40%,剩余容量百分比最大的fs4值为90%,其差值较大.图6显示1TB地震数据写操作完成后,4个文件系统的剩余容量百分比分布.

由图6可知当概率权重因子值小于等于0.6时,

fs1、fs2、fs3及fs4的剩余容量百分比区间分别为14.9%至16.6%、15.6%至17.5%、13.9%至14.4%及37.9%至48.1%。当概率权重因子值大于0.6时,fs1、fs2、fs3及fs4的剩余容量百分比区间分别为18.2%至18.8%、17.9%至18.1%、16.3%至17.9%及18.4%至25.6%。可以看出当概率权重因子值小于等于0.6时,fs3和fs4的剩余容量百分比与其初始状态一样仍然保持其较大的差值;当概率权重因子值大于0.6时,fs3和fs4的剩余容量百分比比较接近,向fs4写地震数据比较频繁。因此为了解决2.2节所描述产生单个文件系统热点写瓶颈问题,建议调整概率权重因子值小于等于0.6。

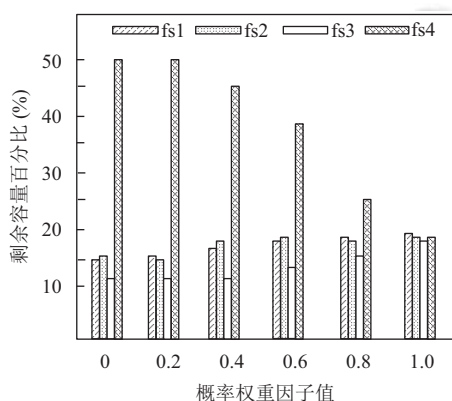


图6 文件系统池剩余容量百分比分布

4.3 文件系统剩余容量对比

文件系统池初始状态中,最大剩余容量为fs1与fs2,都是480 GB,最小剩余容量fs4为180 GB。图7显示1 TB地震数据写操作完成后,4个文件系统在不同概率权重因子值下的剩余容量分布。在不同概率权重因子值条件下,fs1与fs2的剩余容量都基本相同,写数据非常均衡。

由图7可知,当概率权重因子值小于等于0.4时,fs1的剩余容量区间89.4 GB至98.4 GB,fs4的剩余容量区间81.5 GB至99.8 GB。当概率权重因子值大于0.4时,fs1的剩余容量区间99.6 GB至112.8 GB,fs4的剩余容量区间36.8 GB至75.8 GB。可见当概率权重因子值大于0.4时,fs1与fs4的剩余容量值与初始状态一样仍然保持较大的差值;当概率权重因子值小于等于0.4时,fs1和fs4的剩余容量值比较接近。

因此在文件系统池的初始状态为各个文件系统的剩余容量相差很大的情况下,如果保持各个文件

系统的剩余容量均衡,建议调整概率权重因子值小于等于0.4;如果保持各个文件系统的写地震数据量均衡,建议调整概率权重因子值大于0.4。

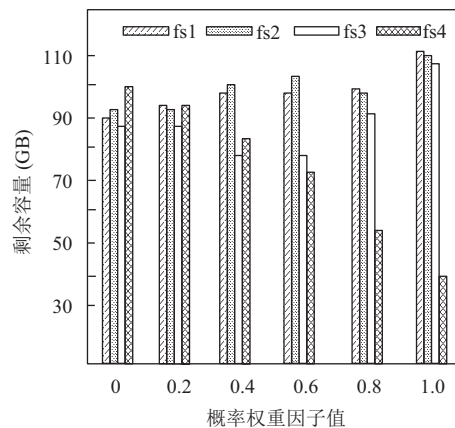


图7 文件系统池剩余容量分布

通过上述实验结果数据分析,根据实际需要,通过调整概率权重因子,基于概率写策略能很好的确保文件系统池中各个文件系统之间地震数据写平衡与优化分布,满足地震资料处理流程中的4个主要阶段的地震数据写需求。

5 结论

本文提出基于概率思想的地震数据写策略,设置不同的概率权重因子值,对地震数据写平衡与优化分布有较大的影响。因此结合不同的文件系统池状态、地震数据文件个数与容量以及地震数据写顺序,自动学习调整优化概率权重因子值是后续需要关注和研究的问题。

参考文献

- 1 赵满. 地震数据并行访问策略的研究[硕士学位论文]. 大庆: 东北石油大学, 2013.
- 2 金弟, 庄锡进, 王启迪, 等. 存储框架模型在地震资料大数据中的应用. 计算机系统应用, 2016, 25(2): 45-51.
- 3 张军华, 臧胜涛, 单联瑜, 等. 高性能计算的发展现状及趋势. 石油地球物理勘探, 2010, 45(6): 918-925.
- 4 Quantum. Stor NextFile System Users Guide. Seattle, USA: Quantum Press, 2010.
- 5 Schlumberger. Omega 2015. 1 Administration Guide. Houston, TX, USA: Schlumberger, 2015.
- 6 CGG. Geovation 2015 Administration Guide. Cedex FRANCE: CGG, 2015.