一种基于阀值的服务器主动复制策略的实现

邹淑芳,李　民
(云南广播电视大学 职业技术学院,云南 昆明 650223)

摘要: 采用阀值来控制复制副本数量的设想,对基于阀值的服务器主动复制策略进行了分析,并用 Java语言设计实现了一种动态复制方法,其中考虑了网络和节点的负载动态变化对复制产生的影响. 该方法可以较好地控制复制副本数量,降低副本一致性维护的代价.

关键词: 动态复制;阀值;分布式系统;服务器
中图分类号: TP302　　**文献标识码**: A　　**文章编号**: 1007 - 855X(2008)04 - 0038 - 04

Design of Initiative Server Replication with Threshold

ZOU Shu-fang, LIM in
(Vocational College, Yunnan Radio and TV University, Kunming 650223, China)

Abstract: In the distributed systems, replication is a key technique to its usability, fault - tolerance as well as efficiency. Threshold value is adopted to control number of duplicate. Initiative server replication with threshold is then analyzed. A dynamic replication method is realized by Java, which takes load variety of net and node into consideration. This method can control numbers of duplicate and reduce maintenance cost of coherence.

Key words: dynamic replication; threshold; distributed systems; server

0 引 言

在分布式系统中,复制是将一个文件在多个节点上放置几份副本. 复制是提高可用性和容错的关键技术[1]. 由于受移动计算和与此相关的断接操作的影响,高可用性正越来越引起人们广泛的兴趣,而容错在安全系统中通常作为一项必备的服务.

进行复制主要出于两个目的:可靠性和性能. 首先,数据复制可以提高系统的可靠性. 如果一个文件系统已经实现数据复制,那么当一个副本被破坏后,文件系统只需要转换到另一个对象副本就可以继续运行下去.

进行数据复制的另一个目的是提高性能. 当分布式系统需要在服务器数量和地理区域上进行扩展时,复制对于提高性能是相当重要的. 当需要访问由单一服务器管理的对象进程数量不断增加时,系统就需要进行服务器数量上的扩充.

可扩展性[2]问题也常常以性能问题的形式出现. 将对象的拷贝放置在使用的进程的附近,可以降低访问时间,提高性能,从而解决可扩展性问题.

目前,国内外对动态复制策略有很多研究,提出了多种策略,如:无条件复制,最老的文件删除;最佳客服端策略;瀑布复制策略;混和瀑布复制;快速传播复制;无复制或缓存;层叠复制无缓存;缓存 +层叠复制;快速扩展,基于发布 /订阅的分片复制技术等.

尽管提出了很多复制策略,但可归结为两种:服务器启动的复制策略和客户启动的复制策略. 然而,上述复制策略都存在一些不足,例如有的复制策略会导致副本数量过多,且当副本更新时要保持副本一致性需要花费更多的代价.

收稿日期: 2008 - 02 - 25.
第一作者简介:邹淑芳 (1962 -),女,副教授. 主要研究方向:分布式计算,图像处理.
E - mail: zoushufang@yntvu.edu.cn

针对以上不足,采用了阀值来控制副本的数量,并且在复制中考虑了网络和节点的负载动态变化对复制产生的影响.该方法可以较好地控制复制副本数量,降低副本一致性维护的代价.

1 复制方法与动态策略

复制方法分为静态复制和动态复制两种.静态复制策略[3]是在分布式系统初始状态时,将副本放置在系统的各个结点上(或者放置在各个服务器上).这种方法的缺点是不能适应用户需求的变化,表现在:最初的副本可能稍候就不再受用户欢迎,而最受用户欢迎的文件却根本没有复制到自身的节点上,这样导致这个文件的站点过载,不但浪费了带宽,而且还可能成为网络的瓶颈.因此现在基本不采用静态复制方法.

动态对象复制[4]就是复制的创建、删除和管理都可以动态且自动地进行,也就是说随着用户行为改变而相应改变其复制策略,并且能够根据存储模式的改变自动地创建和删除文件副本,相比静态对象复制,动态复制显然更能满足分布式系统的需要,而动态复制的核心就是动态复制策略.

动态复制策略需要回答 3个基本问题[5]:何时开始创建副本?哪个文件应该被复制?复制放置到哪个节点上?对上面 3个问题的不同回答产生不同的复制策略.复制的产生依据下面的事实:复制能否减少文件存取的花费.对于动态复制策略的评价一般是通过平均反映时间和整体带宽占有量来进行的.

动态复制策略有无条件复制、最佳节点策略[6]、瀑布复制策略[7]、快速传播、串联复制策略[7]等,其中串联复制策略与最优客户端策略相似,但不同的是该策略不直接复制到请求该文件次数最多的节点 B,而是先复制到本地节点 A到节点 B最短路径上的第一个节点 C,然后节点 C可以向下一个节点 D复制,经过如此串联,最终复制到节点 B.如图 1所示.

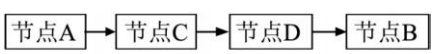

图1　串联复制策略图

Fig.1　Replication in series strategy

最常使用的是最佳节点复制策略和串联复制策略,但这些策略都没有完整的考虑网络带宽的动态变化,因此我们提出并设计了一种基于阀值的服务器主动复制方法,综合考虑网络负载的动态变化对复制的一些影响,并采用了复制的选取、复制的获得、复制的释放等策略,较好地在动态变化的网络环境中实现了复制和副本数量的控制.

2 基于阀值的服务器主动复制

3.1 复制策略

1)复制副本策略:每台服务器跟踪每个文件的访问计数器以及提出这些访问请求的位置(如 IP地址).假设对于一个给定的客户 P,对服务器 Q上的文件 F的访问请求被 Q计入到访问计数器 Count$_Q$(*P*, *F*),如果访问计数值大于阀值 max(*P*, *F*),则把文件 *F*复制到 *P*.

2)删除副本策略:当对服务器 S上的指定文件 F的请求数量下降到删除阀值 min(*S*, *F*)之下时,服务器 S可以删除该文件 F.用一个特殊的记号标记原始副本,防止每个文件至少存在一个拷贝(原始文件).

3)移动策略:如果服务器 P对服务器 Q上的文件 F的请求次数超过了对 Q上文件 F的所有请求次数(Add (F))的一半以上,考虑把文件 F从 Q上移动到 P上,让 P来接管 Q,但这个时候必须考虑 P的负载情况以及存储空间问题,如果负载小,有存储空间,Q将文件 F拷贝到 P上,然后删除 F.如果 P的负载过重,考虑将文件 F移动到离 P较近的服务器.

3.2 复制设计

基于服务器的主动复制各部分功能关系如图 2所示.

功能说明:

Replication Manager(RM):对复制的选择,创建复制,删除复制文件,跟踪复制过程.

Group Manager(GM):存储每个结点的信息,如网络带宽,CPU性能,存储空间等.

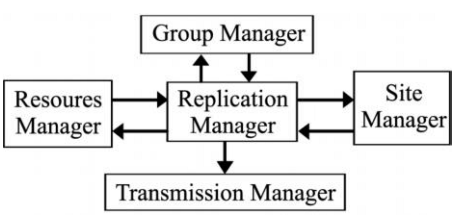

图2　复制功能模块图

Fig.2　Function module of replication

Site Manager(SM):每个结点(计算机)运行一个 SM,它负责跟 RM 交互.

Resources Manager(ReM):文件信息管理.

Transmission Manager(TM):文件传输模块.

3.2.1 复制与删除功能流程图

复制与删除功能流程见图 3、图 4.

3.2.2 Replice Management

1)复制选择. 当节点向主服务器提出对对象 A 的访问请求时,RM 对 ReM 中对象 A 的访问次数进行加 1操作,并跟复制阀值进行比较. 如果大于复制阀值,则 RM 调用 TM 将文件复制到该节点上.

当有多个节点对多个文件访问的阀值次数同时达到了复制要求,这时候把这些节点放置在一个队列中,通过对 GM 的访问获得对应节点信息(如网络带宽等),通过对节点的响应时间从小到大重新进行排队,根据队列节点顺序进行复制操作. 该队列建立在预期的响应时间之上,要选择适当的响应时间.

响应时间:t=服务时间 +等待时间 +交互时间.

等待时间 =服务时间 ×复制队列的长度.

2)复制获得. 当服务过载或网络阻塞时,终端用户将忍受增加的响应时间. 当响应时间大于客户容忍的时间时,SM 通过向 RM 动态追加复制要求,此时 RM 重新对该节点提供一个复制,这样可以提高复制的性能. 是否重新复制的决策通过一个算法进行判别并实现.

ReplicaAcquisition(R_{curr}, R_{prev}, $T_{threshold}$, P, Q, M)

{ /*

假设当前时间是 t;

R_{curr}:当前复制的平均响应时间;

R_{prev}:前一段时间的平均响应时间; $T_{threshold}$:

平均响应时间的最大界限;

……

}

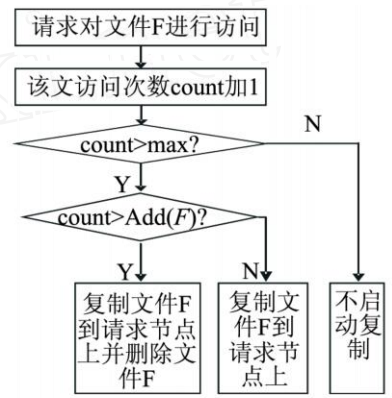

图3 复制副本流程图

Fig.3 Flow chart of duplicate replication

3)复制释放. RM 中的 Cache 有很多复制要求,会造成 Cache 的阻塞,导致一些复制无法进行而使系统资源被浪费. 此时 GM 将释放一些复制. 哪些复制要求将被释放,可通过一个算法来判别.

算法:

ReplicaRelease($U_{threshold}$, P, Q) {

…

}

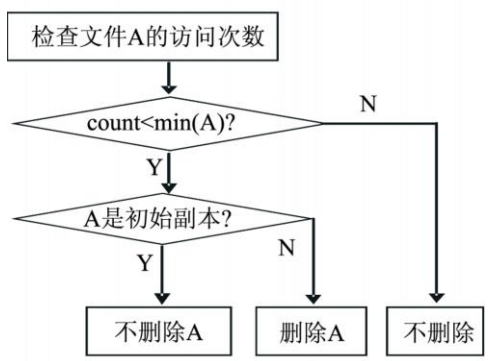

图4 删除副本流程图

Fig.4 Flow chart of delete duplicate

3.2.3 Group Manager

通过网络的路由表和资源查找过程可以得到该表,关键问题是必须动态获取网络情况,如带宽、节点 CPU 处理能力以及磁盘剩余空间等.

3.2.4 Resources Manager

该表是对文件情况的一个存储,采用 Java 对象进行存储. 抽象结构为一个表结构表. 如表 1所示.

表 1　对象信息表

Tab. 1　Information chart for objects

序号	逻辑文件名	文件大小	创建时间	逻辑文件类型	最近访问的时间	存放的位置	访问的位置 IP	访问的次数	版本号码	是否为原始版本
1	Linux操作手册	89KB	2007/05/09	文本文档	2007/05/02	192.168.54.52	192.168.54.40	1	1	1

3 仿真实验

为了验证上述居于阀值的服务器主动复制的性能,我们设计了一个仿真实验.

实验环境:组建了一个小型局域网,在该局域网中有 4台机器,我们把其中的一台机器定为服务器,其余 3台作为主机,在 3台主机上安装了 SM程序.而在服务器中选取的复制文件大小为 7M.在实验中 3台主机通过 SM对服务器中的几个文件发起了多次访问请求,并让请求次数达到复制阀值,或者移动阀值,通过对 GM的调用获取这些节点的相关信息从而建立复制队列,从复制队列中的这些请求节点依次进行复制.我们假设对本地文件访问耗时为 0.实验结果如表 2所示.

表 2　实验结果

Tab. 2　Result of experiment

复制节点数	对服务器访问耗时/s	复制阀值 max	传输耗时/s	复制后本地平均访问耗时/s
1	0.91×4	4	10.4	10.4
1	0.89×2	2	10.4	10.4
3	1.23×4	4	27.04	27.04
3	1.21×2	2	26.9	26.9

由于在小型局域网里,没有发生网络阻塞现象.而且文件的访问较快,因此复制的阀值过小不太好.如果在大量节点广域分布式环境下其阀值可能有所变化,该阀值的确定是和网络的带宽以及服务器的处理能力等有关联.通过实验,我们可以看到基于阀值的服务器主动复制系统实现了预期的复制效果.

4 结 论

对现今分布式系统中的一些复制技术进行了分析,提出并设计了一种基于服务器的主动复制方法,实验表明基于服务器的主动复制可以较好地控制副本复制数量,降低副本一致性维护的代价,但阀值的确定必须依赖于网络的具体带宽、服务器的响应时间等情况.

参考文献:

[1] GEORGE COULOURIS等.分布式系统概念与设计 [M].金蓓弘,译.北京:机械工业出版社,中信出版社, 2004.

[2] 陈琨,陈福明.基于网格计算的文件与对象复制服务 [J].微型电脑应用, 2003, 19(9): 36 - 37.

[3] ELLEN W, ZEGURA et al. Application - Layer Anycasting: A Server Selection Architecure and Use in a Replicated Web Service[J]. IEEE/ACM Transactions on Networking, 2000, 8(4): 10 - 16.

[4] B. Noble et al. Fluid Replication[M]. Proceedings of the Network Symposium, 1999.

[5] 何炎祥,范清风等.网格计算中动态复制策略的设计 [J].计算机工程, 2004, 30(3): 1 - 2.

[6] 谢储晖.空间信息网格中的复制技术 [J].闽江学院学报, 2004, 25(5): 2.

[7] 吴豪,曾国荪,张季平.数据网格关键技术分析 [J].计算机工程与应用, 2003, (35): 1 - 2.