

数据网格中的副本定位^①

刘高嵩 张小波 龙 军 (中南大学信息科学与工程学院 湖南 长沙 410075)

摘要: 在数据网格中,数据副本的定位是副本管理的一个关键问题,是数据网格环境中数据访问和管理的基础。本文分析了当前对数据副本定位所采用的集中式和分布式的定位方法中存在的不足,提出了一种混合式的副本定位策略,将数据副本的定位信息分散地存储于不同的域定位节点中,通过逻辑上集中的逻辑文件定位索引实现了对数据副本的准确定位。

关键词: 数据网格;副本定位;副本目录;数据副本;逻辑文件;物理文件

Replica Location in Data Grid

LIU Gao-Song, ZHANG Xiao-Bo, LONG Jun

(Institute of Information Science and Engineering, Central South University, Changsha 410075, China)

Abstract: Replica location is the key problem of replica management in a data grid. It is also the basis of data access and management. This paper analyzes the pitfall of centralized and distributed replica locations that are used in replica location. By proposing a hybrid strategy of data replica location, scattering replica location information in different domain location nodes, logical file name location index, which is logically centralized, can be used to locate replica information.

Keywords: data grid; replica location; replica catalog; data replica; logical file; physical file

1 引言

数据网格主要面向广域的异构环境,以命名透明性、定位透明性、协议透明性为设计目标,以数据管理为中心,在底层将异构的、分布的数据资源汇聚起来,为上层应用提供通用、可靠、可维护的服务^[1]。由于这些数据资源的分布性,用户访问这些资源往往比较困难,创建数据副本可以减少数据访问延迟、降低网络带宽的消耗,从而提高可用性^[2]。但数据副本的建立也带来了一些新的问题,例如:当某一个副本发生更新时,为了使其余的数据副本也进行同样的更新操作,如何确定其余数据副本的位置,就成为一个不得不面对的问题,这就是数据副本的定位问题。

目前,副本定位的策略主要有集中式副本定位和分布式副本定位。集中式副本定位策略实现简单,但随着应用规模的增大,集中式副本定位逐渐成为系统性能和可用性的瓶颈,限制了系统的可扩展性和可靠

性,比较有代表性的网格项目有 Globus 和 SRB^[3]。Ripeanu 和 Foster 提出了一种分布的副本定位策略,利用 Bloom Filter 技术在本地的各个节点上压缩存储系统全部副本定位信息,这种方法有着良好的查询性能,但带来很大的存储开销^[4]。由于集中式副本定位和分布式副本定位在数据网格系统中的应用都存在明显的不足,因此迫切需要一种新的副本定位模型,以便对数据网格系统的副本进行快速有效的定位。

本文通过分析以上两种副本定位策略,以分布式副本定位策略为基础,提出了一种混合式的副本定位模型,这种定位模型采用集中式副本定位的方式,在数据网格中建立逻辑文件的全局唯一索引,但并不会将全局唯一索引集中地存储到一个索引服务器中,而是通过映射技术分散的存储到网格中的特定节点上;对于副本定位信息的存储参考分布式副本定位的方式,但不会在各个节点中存储系统全部的副本定位信

^① 基金项目:国家自然科学基金(60873081, U0835003, M0921005);高等学校博士学科点专项科研基金(20090162110072)

收稿时间:2010-01-18;收到修改稿时间:2010-02-12

息,而是在特定的节点中,只是存储当前网格域中所有逻辑文件对应的数据副本的定位信息。混合式副本定位模型既具有集中式副本定位的简易性,又具备分布式副本定位的可扩展性和可靠性,而且拥有较高的系统性能。

2 副本定位模型

2.1 相关概念

数据副本 (Replica): 对同一个文件的多个拷贝,这些和源文件完全相同的文件称之为副本,这里的文件可以是数据库,也可以是一般的普通文件等。

网格域 (Grid Domain, GD): 由网格系统中物理上或功能上相近的节点所构成。

逻辑文件名 (Logical File Name, LFN): 是文件在数据网格的名称空间中全局唯一标识符,它代表一个文件的所有副本,所有的逻辑文件都必须在网格中注册后才能使用。为了保证文件的逻辑名称的唯一性,在产生文件的逻辑名称时,以其物理名、IP地址、存储路径为参数对其采用 MD5 算法产生一个 128 位的数字标识,作为该文件的 LFN,即:

$LFN = MD5(\text{"物理名"} + \text{"IP 地址"} + \text{"存储路径"})$ 。

物理文件名(Physical File Name, PFN): 用于唯一标识存储于某个位置的单个物理文件。

副本目录(Replica Catalog, RC): 是一个数据结构或数据存储,用于存储逻辑文件名和物理文件名、提供逻辑文件名到物理文件名的映射信息。

2.2 副本定位模型

在混合式副本定位模型中,包含了三类节点(如图 1 所示):存储节点(Storage Node, SN)、域定位节点(Domain Location Node, DLN)以及全局定位节点(Global Location Node, GLN)。

2.2.1 存储节点 SN

SN 存储了实际的物理数据。另外存储节点还保存了以下的信息:

(1) 为了对每个网格节点上的副本信息进行管理,对每个节点设计了本地副本目录 LRC(Local Replica Catalog),负责管理本地节点所存储的副本文件的物理文件名 PFN 与逻辑文件名 LFN 之间的映射 (PFN, LFN, Version),其中 Version 为当前 LFN 的版本号,为后续的保持数据副本的一致性提供支持。

LRC 支持本地副本查询能力,即网格节点通过查询本地 LRC,可以得知本地节点上是否存有某个逻辑文件 LFN 的副本。由于每个 LRC 负责一个网格节点,因此将 LRC 存储到对应的网格节点上。

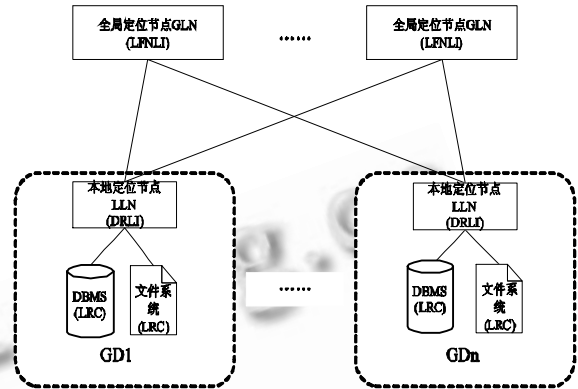


图 1 副本定位模型的架构

(2) 当前存储节点中已注册的逻辑文件的描述信息(LFN, PFN, 注册日期, 逻辑文件的存储位置, Version, 副本的个数);其中副本的个数是指当前的网格系统中存在多少个 LFN 的数据副本。

2.2.2 域定位节点 DLN

域定位节点 DLN 主要保存以下的信息:

(1) 已注册逻辑文件到对应各个副本的定位信息 (LFN, Version, PFN, 数据副本存储的位置, 副本节点的 IP 地址, 副本节点的存储系统类别, 访问参数),即域副本定位索引 DRLI(Domain Replica Location Index),是网格系统副本目录的重要组成部分,为数据副本提供定位信息。例如当前域内的一个逻辑文件有两个数据副本,则 DRLI 会存储两个映射信息,如表 1 所示。

表 1 DRLI 上的副本定位信息

LFN1	PFN1	属性(Version, IP, 存储系统类型, 目录位置等)
LFN1	PFN2	属性(Version, IP, 存储系统类型, 目录位置等)
⋮	⋮	⋯

与文献[4]提出的在本地各个节点上压缩存储系统全部副本定位信息相比,DRLI 中只存储以当前域内的逻辑文件为数据源的数据副本信息,极大地减少了存储上的开销。DRLI 也提供了本地查询能力,当一个

副本文件发生改变后, 如果该副本文件恰好和它的源文件在同一个域内, 那么通过查询 **DRLI** 就可以定位该副本文件的所有其它副本文件的位置。

(2) 当前域内已注册的网格节点信息(网格节点的名称, 节点 IP, 节点的物理位置, 节点的所有者, 操作系统类型, 存储系统类别), 为网格节点的动态加入和退出网格系统提供支持。

(3) 当前域内已注册逻辑文件与其对应网格节点的映射(LFN, Version, PFN, 网格节点名称, 注册日期, 网格节点地址)。

在每个域内部选择一个网格节点作为该域的域定位节点, 一般选择那些可靠性高、存储能力强以及网络带宽较高的网格节点作为域定位节点。

2.2.3 全局定位节点 GLN

全局定位节点 **GLN** 主要保存以下的信息:

(1) 各个域节点的注册信息(域节点的名称, IP 地址, 域节点的物理位置, 拥有者, 注册日期), 为域节点的动态加入和退出网格系统提供支持。

(2) **LRC** 只负责对本地节点上的副本信息进行管理, **DRLI** 只负责对当前域内的逻辑文件所对应的副本文件进行管理。在实际应用中, 可能需要定位其它域中的某个逻辑文件对应的全部副本, 为了能在全局范围内实现对逻辑文件 **LFN** 的统一查询和管理, 通过建立逻辑上集中的逻辑文件定位索引 **LFNLI(Logical File Name Location Index)** 提供全局范围的查询支持。**LFNLI** 负责管理逻辑文件 **LFN** 与其所在的节点域之间的映射。例如逻辑文件 **LFN1, LFN2** 到其所在域的映射信息(**LFN1, DLN1**), (**LFN2, DLNn**), 如表 2 所示。通过访问 **LFNLI** 可以定位逻辑文件所在的域, 然后访问该域中的 **DRLI** 就可以得到逻辑文件 **LFN1** 在整个网格系统中的所有数据副本。

表 2 **LFNLI** 上的映射信息

LFN1	Version	DLN1 (IP 地址, 域节点名称, 注册日期等)
LFN2	Version	DLNn (IP 地址, 域节点名称, 注册日期等)
⋮	⋮	⋯

在混合式副本定位模型中, 本地副本目录 **LRC** 以及域副本定位索引 **DRLI** 分布的存储在各个存储节点

SN 和域定位节点 **DLN** 中, 副本数据的定位信息由 **LRC** 和 **DRLI** 共同提供, 域定位节点 **DLN** 和存储节点 **SN** 可以自由的维护本节点中的副本定位信息, 使得对副本文件定位信息的存储和管理, 不会影响网络系统的整体性能。**SN, DLN, GLN** 是逻辑上的节点, 在实际中它们可能是同一个网格节点。由于逻辑文件定位索引 **LFNLI** 在逻辑上仍然是一个整体, 所以对逻辑文件 **LFN** 的查询和管理具有集中式副本定位的特点。

2.3 副本目录

存储节点中的本地副本目录 **LRC**、域定位节点中的域副本定位索引 **DRLI** 以及全局定位节点中的逻辑文件定位索引 **LFNLI**, 共同构成了网格系统中副本目录的逻辑结构。如图 2 所示:

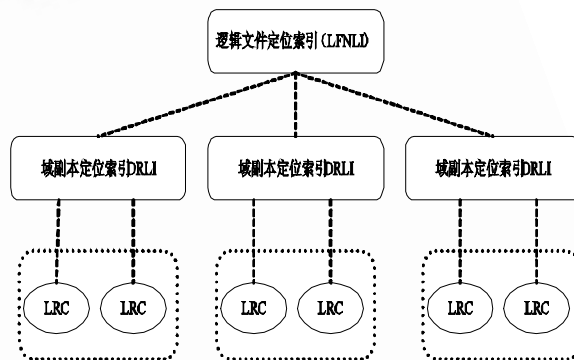


图 2 副本目录逻辑结构

为了避免逻辑上集中的逻辑文件定位索引 **LFNLI** 成为网格系统的性能瓶颈, 有必要将 **LFNLI** 均衡的映射在特定的网格节点中, 使得逻辑文件定位索引 **LFNLI** 在逻辑上采用集中式的管理方式, 而在存储时则采用分布式的存储方式。

2.3.1 **LFNLI** 在网格节点中的存储

为了将 **LFNLI** 均衡的存储在特定的网格节点中, 规定: 只有那些域中网格节点数目大于 **N** 的域定位节点才能作为 **LFNLI** 的存储节点, 即全局定位节点 **GLN**。

分布式存储 **LFNLI** 的算法描述如下:

(1) 对全局定位节点进行二进制编码。编码方法如下: 假设有 **M** 个全局定位节点 **GLN**, 则需要构造有 $(2^M - 1)$ 个节点的二叉树, 按照左 0 右 1 的原则, 对二叉树的叶子节点进行编码, 得到全局定位节点的二进制编码 (**Global Location Node Binary Code, GLNBC**), 用数组 **GLNBC[M]** 存储。举例如下:

假设有5个全局定位节点,则需要构造有9个节点的二叉树(如图3所示),则GLN1的二进制编码为:000;GLN2的二进制编码为:001;依次类推可以得到其余节点的二进制编码。

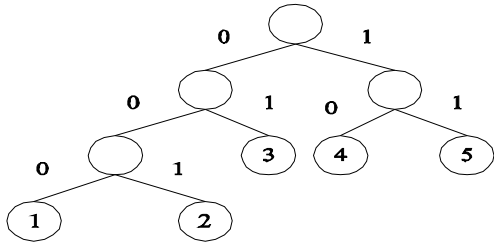


图3 GLN的编码过程

(2) 为LFNLI选择全局定位节点。选择逻辑文件LFN的前几位二进制编码等于GLN二进制编码的节点,存储相应的LFNLI。从LFN到GLN的映射过程可以描述为:LFN->GLNBC->GLN。

这样就达到了将逻辑上集中的逻辑文件定位索引LFNLI,均衡存储在网格节点中的目的。

3 副本定位过程

基于图2所示的副本目录,提出了层次式的副本定位过程,用户的副本定位请求在副本目录中由底层向高层传播,然后再由高层向底层定位副本文件,具体包括如下步骤(如图4):

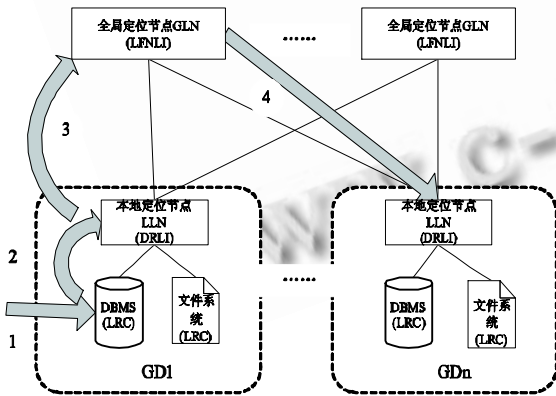


图4 副本定位过程

假设网格系统的任意节点n上的副本文件发生了更改,节点n所属的域为GD1。

(1) 获取发生改变的副本文件的PFN,在本地LRC中查找该PFN对应的LFN,然后执行操作2;

(2) 根据步骤1中找到的LFN,向域定位节点DLN1发出定位LFN副本的请求;若域副本定位索引DRLI中存在对应的LFN项,说明该副本文件与它的源文件在同一个域的范围,该LFN项中就包含了该LFN的全部副本文件的网络位置信息,副本定位过程结束;若域副本定位索引中不存在对应的LFN项,表示该副本文件的源文件并不存在于当前的域中,则执行操作3;

(3) 域定位节点根据得到的LFN,利用分布式存储LFNLI的算法,可以找到存储该逻辑文件名的LFNLI的全局定位节点GLN,然后向该全局定位节点GLN发出定位LFN副本的请求;

(4) 全局定位节点GLN根据得到的LFN,在逻辑文件定位索引LFNLI中查找该LFN;若找到对应的LFN项,就可以根据全局定位节点GLN中保存的各个域定位节点的地址,将定位LFN副本的请求转发到相应的域定位节点,这里假设该域定位节点为DLNn,然后执行操作5;若没有找到对应的LFN项,就向请求定位LFN副本的域定位节点发送错误报告。

(5) 域定位节点DLNn在收到全局定位节点GLN的定位请求后,在当前域的域副本定位索引DRLI中查找该LFN,若存在对应的LFN项,说明已经找到了该LFN对应的所有文件副本,将这些副本的位置信息返回给节点n,副本定位过程结束;若不存在对应的LFN项,则向全局定位节点发出错误信息,副本定位过程结束。

上述定位过程的算法描述如下:

FindReplicas()

```

{
  If(LRC 中存在该 PFN 对应的 LFN)
  {
    获取该 PFN 的 LFN;
    向该节点所在的 DLN 发出该 LFN 的副本定位请求;
    If(DRLI 中存在对应的 LFN 项)
    {
      Return 该 LFN 对应的全部副本位置信息。
    }else
    {
      利用分布式存储 LFNLI 算法,可以找到存储该 LFN
      的 LFNLI 的全局定位节点 GLN;
      向 GLN 发出该 LFN 的副本定位请求;
    }
  }
}
    
```

```

    If(逻辑文件定位索引 LFNLI 中存在该 LFN 项)
    {
        根据 LFNLI 中的 LFN 项定位该 LFN 所在
        的 DLNn。
        向该 DLNn 发出 LFN 的副本定位请求;
        If(DRLIn 中存在对应的该 LFN 项)
        {
            Return LFN 对应的全部副本位置信息。
        }else
        {
            Return 错误信息;
        }
    }else
    {
        Return 错误信息;
    }
}
Exit;
}

```

4 性能分析

(1) 负载均衡性。假设网格系统中全局定位节点 GLN 的个数为 $N(2m \leq N < 2m+1)$, 其中 m 表示完全二叉树的深度), 逻辑文件 LFN 的个数为 W , 则每个全局定位节点 GLN 上的逻辑文件的记录个数小于 $2W/N$, 每个全局定位节点中维护的最大逻辑文件的个数不超过最小逻辑文件个数的 1 倍。与集中式副本定位相比, 在本定位模型中, 副本定位信息以及逻辑文件定位索引 LFNLI 都分散地存储在各个特定网格节点中, 有效的降低了网格系统对某个特定节点的强依赖性, 实现了负载均衡。

(2) 可靠性。每个 LFN 的逻辑文件定位索引对应两个全局定位节点 GLN1 和 GLN2, 假设 GLN1 失效的概率为 $M1$, GLN 失效的概率为 $M2$, 则 GLN1 和 GLN2 同时失效的概率为 $M1 * M2 (0 < M1 * M2 < M1, M2 < 1)$, 两个全局定位节点同时失效的概率远小于任何一个节点失效的概率。由于副本定位信息是通过特定的方式分散的存储在网格节点中, 某一个网格节点的突然失效, 不会对整个网格系统的正常运行有太大的影响, 增强了系统的容错性。

(3) 可扩展性。考虑到数据网格中的动态性, 混

合式的副本定位以分布式副本定位方法为基础, 将逻辑文件定位索引均衡的映射在不同的网格节点中, 数据副本的定位信息分散在各个域定位节点中, 因此对逻辑文件 LFN 以及副本文件数目的可扩展性支持较好。与集中式副本定位相比, 混合式副本定位只是记录了逻辑文件 LFN 到其所在的域定位节点 DLN 的映射, 并没有存储逻辑文件到其所有副本文件的映射, 因此, 对逻辑文件 LFN 的管理不会成为整个网格系统的性能瓶颈。另外混合式的副本定位对存储节点 SN 和域定位节点 DLN 的数目也没有限制, 系统可以根据实际的需要进行灵活的配置和扩展。

(4) 降低网格节点的存储负担。与 Ripeanu 和 Foster 提出在各个网格节点上压缩存储系统全部副本定位信息的副本定位策略相比, 在本定位模型中, 逻辑文件定位索引 LFNLI 均衡分布在特定的网格节点中, 各个域定位节点只存储当前域中逻辑文件对应的副本定位信息, 极大的减少了副本定位信息的存储开销。

(5) 可用性。混合式的副本定位充分考虑到因特网的网络特点, 在进行副本定位过程中, 由近到远展开, 充分利用了局部性优势, 不仅很好的避免了因特网的缺点, 而且减少了相应时间, 提高了工作效率。

5 总结

本文围绕怎样在复杂网络环境中准确定位文件副本问题, 首先研究了当前数据网格中文件副本定位主要采用的两种方法, 分析了它们的不足, 进而提出了混合式副本定位模型, 并在此基础上给出了副本目录的逻辑结构。该副本定位模型具有层次式的副本定位特点, 分析表明, 该副本定位模型应用与基于因特网的数据网格中, 具有减少系统响应时间, 实现简单, 提高系统性能等优点。

参考文献

- 1 高改梅, 白尚旺, 党伟超. 一种分布式数据网格副本定位机制的研究. 太原科技大学学报, 2007, 28(4): 262.
- 2 宋宝燕, 毛艳英, 李旭峰, 奚丽娜, 申德荣, 于戈. 一种基于负载均衡的分层副本定位方法. 计算机工程, 2008, 34(18): 99.
- 3 王驰. 数据网格副本策略研究[硕士学位论文]. 上海: 上海交通大学, 2007.
- 4 田宋阳. 数据网格中副本定位及选择服务[硕士学位论文]. 重庆: 重庆大学, 2006.