

结构化P2P网络一致性维护策略

齐小刚¹, 杨伟^{1†}, 刘立芳², 高蓉¹

(1. 西安电子科技大学 数学与统计学院, 西安 710071; 2. 西安电子科技大学 计算机学院, 西安 710071)

摘要: 一致性维护策略是解决结构化peer-to-peer(P2P)网络中数据可靠性和一致性的关键技术,基于该技术,数据可以在不同时间、不同地域得到有效的管理.然而,不同节点的空间异构性、能力异构性和功能异构性对数据一致性维护提出了挑战.对此,首先介绍结构化P2P网络一致性维护的基本概念和特点,并阐述一致性维护策略设计面临的挑战以及需考虑的性能指标;然后提出一致性维护策略的分类标准和系统模型,根据该分类标准,详细剖析当前一致性维护策略相关的研究工作;最后,讨论结构化P2P网络一致性维护研究的开放性问题,并指出未来研究的方向和重点.

关键词: 对等网络; 结构化P2P网络; 一致性维护; 分布式系统

中图分类号: TP393

文献标志码: A

Consistency maintenance strategy of structured P2P network

QI Xiao-gang¹, YANG Wei^{1†}, LIU Li-fang², GAO Rong¹

(1. School of Mathematics and Statistics, Xidian University, Xi'an 710071, China; 2. School of Computer Science and Technology, Xidian University, Xi'an 710071, China)

Abstract: The consistency maintenance strategy is the key technology to solve the reliability and consistency of data in the structured peer-to-peer(P2P) network. Based on this technology, the data can be managed at different time and in different regions effectively. However, the spatial heterogeneity, capacity heterogeneity and functional heterogeneity of different nodes pose challenges to data consistency maintenance. Therefore, this paper introduces the basic concept and characteristics of consistency maintenance in the structured P2P network and dissects the challenges and concerned performance metrics in design of the consistency maintenance strategy. According to the proposed classification criteria and system models, the related research work on consistency maintenance strategy are analyzed in depth. Finally, the problems in consistency maintenance strategy design are discussed to point out future research trends and focuses.

Keywords: peer-to-peer; structured P2P; consistency maintenance; distributed system

0 引言

随着计算机网络技术和应用的迅猛发展,P2P网络由于其可扩展性和分布式等特性得到了广泛的研究和重视,但许多P2P应用只支持静态文件的读写,而对于许多动态改变的文件并不支持^[1].随着计算机技术的不断发展,对于新的P2P应用,人们要求这些应用能够支持可变文件的读写功能^[2-4];在P2P网络中,复制和缓存这两种方法常被用来提高数据的可扩展性和可用性^[5].然而,随着这两个技术不断被发掘和关注,另一个新的问题随之浮出水面,即数据一致

性维护.现有的许多研究过多地关注于P2P文件复制技术^[2-3,6-12],而对于数据一致性维护关注甚少.结构化P2P即对数据的存储和查询采用一种固定而非随机的方法.大多采用基于分布式哈希表(DHTs)^[13]的方法来存储和查询数据.在基于DHTs的系统中,会在逻辑上存在一个命名空间,数据和节点将被分配到这个空间上^[14].结构化P2P系统往往十分庞大,而且系统中的节点性能是不均匀的,使得一致性维护方法的设计变得十分困难^[15];并且结构化P2P系统允许节点自由加入和退出,这给结构化P2P系统带来了

收稿日期: 2017-03-08; 修回日期: 2017-08-25.

基金项目: 国家自然科学基金项目(61572435,61472305); 陕西省自然科学基金项目(2015JZ002,2015JM6311); 宁波市自然科学基金项目(2016A610035,2017A610119); 复杂电子系统仿真重点实验室基础研究基金项目(DXZT-JC-ZZ-2015-015).

责任编辑: 张文安.

作者简介: 齐小刚(1973-),男,教授,博士,从事系统建模与故障诊断及其应用等研究; 杨伟(1993-),男,硕士生,从事分布式数据存储和管理的研究.

†通讯作者. E-mail: yangw@stu.xidian.edu.cn

很高的动荡性. 现有的结构化P2P系统利用分布式散列表(DHTs)^[13]分配资源给不同的节点, 由于分布式散列表的设计, 使得P2P系统在拓扑层面上达到了负载均衡. 然而, 现实情况却复杂得多, 由于内容流行程度的不同, 某些内容被访问的次数非常多, 而某些内容被访问的次数又非常少, 这样, 使得被访问次数多的节点经常会出现负载过重等问题^[16], 在这种情况下, 提供复制节点便成为减少节点负载的一个有效方法. 但随着文件内容的改变, 相应复制节点的内容也需改变, 这给一致性维护带来许多困难. 因此, 一致性维护研究对于提高数据可用性以及提高结构化P2P网络适应性意义重大.

本文首先介绍一致性维护的基本概念, 提出新的一致性维护策略分类标准; 然后指出在结构化P2P网络中设计一致性维护策略面临的挑战, 详细介绍现有的一致性维护策略并深入剖析解决挑战的主要方法; 最后, 提出结构化P2P网络一致性维护策略设计的相关开放性问题的, 指出未来研究的方向和重点.

本文的主要贡献如下:

1) 提出一致性维护策略的分类标准和系统模型, 根据该分类标准, 详细剖析当前一致性维护策略相关的研究工作, 并将当前的一致性维护策略进行分类对比.

2) 阐述一致性维护策略设计面临的挑战以及应考虑的性能指标, 并分析各性能指标对一致性维护策略制定的影响程度.

3) 基于上述两个方面详细剖析当前一致性维护策略相关的研究工作并进行对比, 指出其在研究内容上的优点和存在的挑战, 以及改进的方向.

4) 讨论结构化P2P网络一致性维护研究的开放性问题, 为未来的研究指出可能的方向和重点.

1 一致性维护的基本概念

定义1(一致性维护) P2P系统中两个或多个文件副本节点更新同一文件副本, 并使其达到版本同步性的过程称为一致性维护.

定义1是本文通过总结P2P系统一致性维护的目的、过程和达到的结果而提出的. 为了简单地介绍和了解一致性维护的特点, 并对比P2P网络与客户机和服务器结构(C/S)网络的区别, 本文用一个简单的例子做类比阐述. 图1是一个C/S网络, 其中: 节点0表示文件主节点, 其他节点表示副本节点. 当节点0的文件发生更新时, 其他节点也要作出相应的更新, 这时, 每个节点的更新就必须依靠服务器来分发.

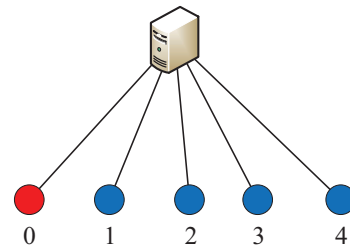


图1 C/S网络

图2表示一个结构化P2P系统, 其中: 节点7是文件 k 的数据主节点, 节点1、2、4、6是文件 k 的副本节点. 当文件 k 发生更新为 k' 时, 这个新的文件 k' 必须尽可能快和全面地替代节点1、2、4、6中的文件 k . 由于网络中没有服务器, 更新的传播主要依靠节点之间进行数据传输.

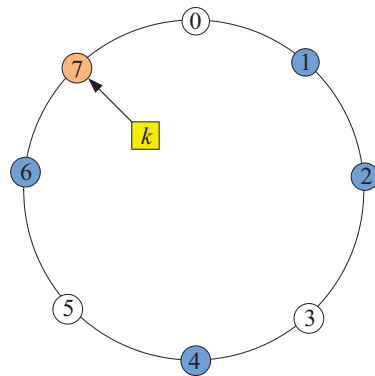


图2 结构化P2P系统

文件复制技术在P2P系统中是提高文件可扩展性和数据可用性的一个有效手段, 数据复制技术在分布式系统中的研究也得到了充分发掘. 在结构化P2P系统中, 由于文件的流行程度不同, 某些文件的访问频率可能非常高, 在这种情况下, 存储该文件的节点很有可能过载. 为了解决这个问题, 可以在其他节点建立文件的副本, 从而减轻节点的负载. 然而, 由于P2P环境的动荡性和复杂性, 往往无法确定系统中副本的数量以及存储位置, 多个节点可能并发更新同一数据的不同副本, 恶意节点可能违反协议, 从而导致更新失败.

目前, 结构化P2P系统中解决一致性维护问题主要有两类思想: 1) 预设中心节点以控制所有副本节点的一致性维护^[17]; 2) 无中心节点的一致性维护. 预设中心节点的一致性维护即预设中心索引服务器存储所有节点的信息, 更新时只需控制中心索引服务器即可^[18]; 无中心节点的一致性维护即每个副本节点存储少量其他副本节点的信息, 这些副本节点会被构建成一个逻辑拓扑, 然后按照某种既定的方式传播更新.

2 策略分类标准及实现系统模型

2.1 分类标准

以第1节一致性维护策略分类思想为基础,并结合文献[16]中一致性维护策略分类思想的优点,将一致性维护策略分为中心点控制策略^[9-21]和分布式更新策略^[1,16,22]两大类(如图3所示).该分类基于一致性维护策略实现的系统依赖条件,即是否依赖预设中心节点控制所有副本节点信息.从系统依赖条件出发对一致性维护策略进行分类的好处是,可以凸显一致性维护策略中的循环依赖问题,即节点拥有的信息依赖于节点间已有的信息传递,该问题常常被研究者忽视.

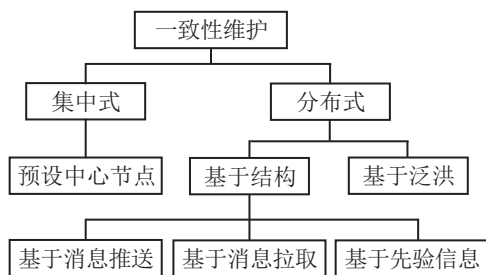


图3 结构化P2P网络一致性维护策略分类标准

中心点控制策略依赖于预设的中心节点控制所有副本节点的信息,因而又称为集中式一致性维护策略.该策略与许多集中式网络一样,是一种静态、主动、快捷的一致性维护策略.由于预设的中心节点的存在,并且中心节点已知所有副本节点的位置,中心点控制策略极大地简化了一致性维护过程中寻找副本节点的过程.目前,有少数小规模P2P系统采用这种策略.但是随着P2P系统的扩展越来越迅速,中心节点需要记录的副本节点信息也成倍增加,中心节点的处理速度和存储空间大小成为限制中心点控制策略一致性维护效率的瓶颈^[15].更为严重的是,中心节点可能会因为受到恶意攻击损坏或者遭受拒绝服务(DoS)攻击而瘫痪,给整个一致性维护系统带来单点失效的风险^[23].

分布式更新策略不依赖于预设的中心节点控制所有副本节点的位置信息,是一种分布式可扩展的一致性维护策略,非常适合P2P网络的大规模性和高度动荡性,并且也适用于大多数的结构化P2P系统,如CAN^[17],Chord^[24],Pastry^[25]和Tapestry^[26].该策略根据消息传播方式可细分为:基于泛洪(Flooding-based)的一致性维护策略^[27]和基于结构(Structured-based)的一致性维护策略^[1].基于泛洪的一致性维护策略采用泛洪的方式传播更新,这种传播方式的好处是在节点动荡的条件下有效地完成更新传播过

程,具有较好的节点失效容忍度;但同时这种方式的缺点也很明显,泛洪的传播方式会产生大量的数据冗余,从而造成花费过多和网络风暴等问题^[28].为了减小数据冗余,其他一致性维护方式采用基于结构的一致性维护策略,这类方法通过把副本节点组织成一个特殊的结构,使得存在于这个结构中的副本节点只会接收一次更新,从而减少了数据冗余.但是,节点的损坏或者离开会破坏结构,造成更新传播失败^[29].基于结构的一致性维护策略又可细分为基于推送的一致性维护策略(Push-based)、基于询问的一致性维护策略(Pull-based)和基于先验信息的一致性维护策略(Prior information-based).基于推送的一致性维护策略通过消息推送的方式传播更新,副本节点只能被动等待更新;基于询问的一致性维护策略通过副本节点询问的方式传播更新,即需要更新的节点接收更新,在一定程度上减小了开销和网络资源的占用率;基于先验信息的一致性维护策略通过收集某些信息(如地理位置信息、兴趣等)构建某种结构来传播更新,这样,在构建结构和分发消息方面可以更多地考虑节点自身因素.

2.2 系统模型

综合分析结构化P2P网络一致性维护策略实现的前提条件不难发现,网络功能和网络环境的不同会对一致性维护策略的设计产生一定的影响.本文提出设计结构化P2P网络一致性维护策略需要考虑3类影响因素:网络支持能力、设备能力和网络环境影响.通常,这3类因素是设计一致性维护策略的假设前提条件.网络支持能力是网络自身属性特征对一致性维护策略所能提供的帮助,决定了一致性维护策略设计的基本思想和构架;设备能力即网络中节点能力的反映,决定了一致性维护策略的性能和实现难度;网络环境反映了网络的整体环境,如网络的动荡性,决定了一致性维护策略对网络环境影响的鲁棒性.具体内容如下.

1) 网络支持能力.

i) 路由信息更新速度: P2P网络的高度动荡性决定了路由信息的准确程度取决于路由信息的更新速度,路由信息更新速度快的P2P网络有利于一致性维护过程的准确性.

ii) 错误检查方案: 保存数据的节点会不可避免地出现节点的失效,因此,需要一套有效的机制来发现系统中节点的失效,这就是错误检测.好的错误检查机制可以提高一致性维护的性能.

iii) 网络规模: 结构化P2P网络中文件的副本节

点数量反映了一致性维护策略需要考虑的规模和可扩展性. 一致性维护策略可扩展性越好, 越有利于策略的适用范围和可移植性, 但也存在开销过大和管理困难的问题.

iv) 角色异构: 节点在一致性维护过程中扮演的角色不同(如主节点/副本节点, 发送节点/接收节点, 父亲节点/子孙节点等), 其执行的操作也不同(存储节点/转发节点等).

2) 副本节点能力.

i) 节点在线时间: 节点存在于P2P网络中的时间反映了节点对于网络结构稳定性的贡献, 在线时间越长, 一致性维护方法所需的费用越低.

ii) 并行处理: 并行处理即节点可以同时处理多个任务, 在一定程度上反映了节点处理数据的能力, 网络中节点拥有并行处理可以使得一致性维护策略的设计有更大的伸缩空间. 目前大多数节点都拥有并行处理能力.

iii) 节点上/下行带宽: 反映节点接收数据速度和节点传递数据的速度. 节点的上/下行带宽可以反映网络拥塞的程度, 这是许多一致性维护方法需要考虑的因素.

3) 网络环境影响.

i) 网络恶意行为: 考虑到网络攻击对一致性维护产生的影响, 应提升一致性维护策略对网络攻击的鲁

棒性. 对于网络中常见的攻击, 如DoS攻击, 可以使节点过载而瘫痪^[23]. 一致性维护策略需要拥有一定的抗毁能力, 在面对DoS攻击时可以在不影响一致性维护过程的情况下快速调整.

ii) 网络动荡程度: 节点加入和离开对一致性维护策略产生的影响. 结构化P2P网络允许节点自由加入和离开, 因此, 节点加入和离开对一致性维护策略产生的影响必须考虑在内.

iii) 结构匹配度: 定义为节点间逻辑跳数和物理跳数之比. 该比值越高, 结构匹配程度越高, 即用逻辑结构反映物理网络的可行性越强. 在一致性维护策略的构建过程中, 需要考虑节点间逻辑距离与物理距离之间的关系, 以提高一致性维护策略的性能.

根据一致性维护策略的实现对上述影响因素的依赖强度(如表1所示. 其中: \checkmark 表示依赖, $-$ 表示无法得知), 本文将一致性维护策略划分为4种不同的系统模型(见表1): 有辅助的中心系统模型(Assisted)^[30]、信息系统模型(Informational)^[31]、结构系统模型(Structured)^[1, 29]、无依赖系统模型(Free-for-all)^[32]. 它们从一致性维护策略实现所依赖假设条件的角度反映不同一致性维护策略的设计特点和复杂程度. 由于这4种系统模型并不能完美区分所有的一致性维护策略, 本文根据“最符合”的原则将现有的一致性维护策略研究工作归类到不同的系统模型中.

表1 一致性维护策略的实现系统模型

假设条件		系统模型			
		Assisted	Informational	Structured	Free-for-all
网络支持能力	路由信息更新速度	高	高	高	低
	错误检查方案	\checkmark	\checkmark	\checkmark	$-$
	网络规模	有限	分布式	分布式	分布式
	角色异构	\checkmark	\checkmark	\checkmark	\checkmark
副本节点能力	节点在线时间	$-$	\checkmark	\checkmark	$-$
	并行处理	\checkmark	\checkmark	\checkmark	\checkmark
	节点上/下行带宽	\checkmark	\checkmark	\checkmark	\checkmark
网络环境影响	网络恶意行为	\checkmark	\checkmark	\checkmark	$-$
	网络动荡程度	\checkmark	\checkmark	\checkmark	$-$
	结构匹配度	$-$	\checkmark	\checkmark	$-$

有辅助的中心系统模型可以看作分布式P2P网络中一个路由数据汇集节点, 所有节点的寻路和资源发现都由这个节点来处理, 可以看作半分布式P2P网络场景. 信息系统模型表现为对节点某些信息的依赖, 如节点的兴趣、地理位置^[29]等信息, 然后根据这些信息设计一致性维护策略. 结构系统模型应用于当前多个一致性维护策略, 在一致性维护过程中, 会把副本节点按照某种给定的结构组成相应的逻辑网络, 不同的结构有着不同的特点. 例如, 树形结构具

有随着树高增加节点呈指数增长的特点, 利用这个特点, 可以在不增加节点负载的基础上加快数据的传播速度, 但由于树形结构连通度为1, 且容易受到单点性能退化的影响, 目前许多一致性维护方法都针对此类问题作出了改进. 无依赖系统模型对假设条件的依赖性最小, 通常这类系统模型依靠泛洪来传播更新, 这类系统有着维护简单、可靠性高等特点, 但考虑到泛洪的缺点, 信息传播过程中可能会对网络造成极大的负担, 甚至会引起网络风暴等问题.

2.3 小结

综上所述, 本文将现有的一致性维护策略研究工作从6个方面进行分类, 选取方式按照相似性原则, 即如果两个一致性维护策略在这几个方面没有

很明显的区分, 则只选取其中之一(见表2). 具体如下: 网络场景(Scen), 纯分布式和半分布式; 系统模型(Sys); 是否依赖先验信息(Inf); 是否需要维护结构(Mant); 更新获得方式(Acq), 推送、拉取; 是否缓存更新(Cache).

表2 结构化P2P网络一致性维护设计策略分类表

Schemes	Scen	Sys	Inf	Mant	Acq	Cache
BCoM ^[1]	Semi-distributed	Structured	No	Yes	Push	Yes
CUP ^[33]	Distributed	Structured	No	No	Push	Yes
Buchegger 等 ^[34]	Distributed	Free-for-all	No	No	Hybrid	No
GeWave ^[22]	Semi-distributed	Informational	Yes	Yes	Pull	No
SCOPE ^[16]	Semi-distributed	Structured	No	Yes	Push	No
Swarm Intelligence ^[29]	Semi-distributed	Informational	Yes	No	Push	No
Gnutella ^[35]	Semi-distributed	Assisted	No	No	Push	No
Locality-aware ^[5]	Semi-distributed	Informational	Yes	Yes	Push	No
Lan 等 ^[36]	Distributed	Free-for-all	No	No	Hybrid	No
Shen 等 ^[22]	Semi-distributed	Informational	Yes	Yes	Push	No
Online I ^[37]	Sistributed	Free-for-all	No	No	Pull	No
Nakashima 等 ^[38]	Semi-distributed	Structured	No	Yes	Push	No
Nobre 等 ^[39]	Distributed	Assisted	No	No	Push	No

3 研究挑战及性能指标

设计结构化P2P网络一致性维护策略主要面临3个方面的挑战: 一致性维护效率、结构维护、一致性强度. 由于分布式更新策略是目前的研究趋势, 本文主要针对分布式更新策略提出相应性能指标以应对这些挑战.

3.1 一致性维护效率

定义2(一致性维护效率) 一致性维护效率是指一致性维护过程中更新传播至所有副本节点所需要的时间^[40].

由于P2P网络的高动荡性, 更新传播速度越快, 一致性维护效率越高, 其更新丢包的发生也就越少.

在中心点控制策略的一致性维护中, 一致性维护效率取决于中心点的处理能力和中心点与副本节点之间的带宽. 所以在中心点控制策略中, 一致性维护效率一般很难优化.

在分布式更新策略中, 有两种不同的更新传播方式: 一种是基于泛洪的无方向更新传播; 一种是基于结构的有方向更新传播. 在基于泛洪的无方向更新传播中, 由于其消息传播方式的特殊性, 一致性维护效率一般都比较低, 但这种高效率同时伴随着高的开销和其他问题, 如网络风暴和带宽占用等问题. 目前多采用基于结构的有方向更新传播策略, 这种方式通常把副本节点组织成一种结构(比如树), 在更新的传播中, 更新沿着规定的方向传播(在树中从上到下), 这种方法的好处是可以保证一致性维护效率的同时减少开销. 基于结构的有方向更新传播通常要求副

本节点只能收到一次更新, 这样可以减少网络中的数据冗余. 但这种方法也存在一些问题, 如在结构的设计上, 需要保证节点加入的公平性, 即所有副本节点都有加入结构的权利. 结构的构建方式要求尽量简单, 不能有很高的复杂性, 因为P2P网络的高度动荡性, 结构通常需要在被破坏后尽快恢复, 此时, 一个轻量级的结构恢复算法是必要的. 对结构的性能也需要有相应的衡量指标, 如树形结构的树高, 由于树形结构是单连通的, 破坏其中任意一个节点都会引起结构的损坏; 而且在树形结构中, 由于更新是自上向下传播的, 树的高度也决定着更新的传播速度. 在分布式更新策略中, 结构的构建方式和维护结构的开销决定了一致性维护效率, 因而在保证一致性维护效率的同时很难保证其他因素(如开销、带宽占用等), 这使得基于结构的有方向更新传播的一致性维护策略成为目前结构化P2P网络的研究热点. 为了达到高的一致性维护效率并达到其他要求, 一致性维护策略需要考虑如下性能指标:

1) 冗余消息数(Redundancy): 在一致性维护过程中, 网络中所有的更新消息数量即为一致性维护策略的冗余消息数. 冗余消息数越高, 说明一致性维护策略的开销越大.

2) 更新传播时延(Delay): 从更新发起到所有副本节点收到更新的时间间隔即为更新传播时延. 时延越高, 说明一致性维护的效率越低.

3) 丢包率(Discard rate): 更新传播过程中的丢包数与总更新数的比值即为丢包率. 丢包率越高, 意味

着节点收到更新的概率越低,相应的一致性维护效率也会变低.

4) 带宽占用率:一致性维护过程中的带宽占用量与网络总带宽的比值即为带宽占用率.带宽占用率越高,表明网络中冗余消息数过高,或者是更新消息经过的路径太长.低的带宽占用率则可以减轻网络的负载.

5) 平均路径长度:平均路径长度即更新传播路径长度与副本节点数的比值.平均路径长度越短,说明一致性维护效率越高,相应的带宽占用率越低,开销越小.

3.2 结构维护

在副本节点损坏或者离开后如何恢复一致性维护策略的结构,就是结构维护.这个问题是基于结构的一致性维护策略需要重点考虑的问题.通常,执行结构建立算法和结构维护算法是实现结构维护的两种主要途径.

结构建立算法即当结构被破坏时,一致性维护策略根据结构开始建立的算法重新组织节点.这种方式适用于结构简单、副本规模较小的一致性维护策略;当副本节点数较多时,这种方式由于需要重新建立结构而不再适用.当网络动荡性较大时,反复重新构建结构在时间复杂度和空间复杂度上都比较高.

结构维护算法即当结构被破坏时,一致性维护策略根据被破坏节点的位置,在不对结构做大规模修改的情况下改动小部分节点从而维护整个结构.这种方法由于其改动小而被大多数研究者选择.在结构维护算法中,算法的时间复杂度影响结构的修复时间,空间复杂度影响节点负载.

在中心控制策略中,中心节点会持有所有副本节点的地理信息,整个一致性维护结构可以抽象为一个星形网络:中心节点位于网络中央,其他节点连接中心节点.在这个结构中,某个节点损坏并不会影响一致性维护策略的性能,所以在中心控制策略中,中心节点可以通过发送心跳消息来检测节点是否损坏.

在分布式更新策略中,通常会把副本节点组织成一个结构,所以当某个副本节点损坏时,结构通常会被破坏,在这种情况下,会设计一个结构恢复算法,通过改动损坏节点周围小部分节点来恢复结构.好的结构维护算法可以提高一致性维护策略鲁棒性.在BCoM^[1]中,让子孙节点缓存祖先节点的信息可以更快速地恢复结构.采用结构维护算法进行结构恢复时,需要考虑如下性能指标:

1) 结构特征(Structural features):一致性维护策略

用来组织副本节点的结构特点,如树高、节点度大小等.在树形结构中,较低的树高可以减小更新传播时延,降低开销,增强抗毁性.

2) 算法时间复杂度(Time complexity):时间复杂度用来衡量结构恢复时间.由于P2P网络的高度动荡性,较低的算法时间复杂度意味着较快的结构恢复速度.

3) 算法空间复杂度(Space complexity):空间复杂度用来衡量算法运行过程中所占存储大小.较低的算法空间复杂度给一致性维护策略带来的运算负载低.

4) 算法开销(Overhead):即结构恢复过程的开销.由于P2P网络的高度动荡性,结构维护算法要求开销小,速度快.

5) 失效节点数(Numbe of failed nodes):失效节点数用来衡量结构维护算法的适用性.在一致性维护过程中,单点失效和多点失效并存,所以结构维护算法必须适用于多点失效环境.

3.3 一致性强度

综合分析数据一致性在时间同步上的差别.本文对一致性强度定义如下.

定义3(一致性强度) 复制节点持有的平均更新版本与最新版本的距离的倒数称为一致性强度.

在中心控制策略中,由于中心节点的存在,更新可以发送给每个需要的副本节点,而且中心控制策略的一致性强度比较高,但中心控制策略依赖于中心节点的性能、状态等,这便给中心节点带来很大的负载,尤其是在副本节点数较多的情况下.

在分布式更新策略中,基于泛洪的无向传播因受TTL的限制并不能传播给所有需要更新的副本节点,并且这样会造成开销过大的问题,所以这种方式不再适用于目前的结构化P2P网络环境.基于结构的有向传播通过建立某种结构来传播更新,通常这种方法会设定一个节点加入算法,使得有需要的副本节点可以加入结构接收更新,结构的设定和更新传播方式的确决定了一致性强度.为了增加一致性强度,多采用树形结构来传播更新,因为随着树高的增长节点数呈指数增长.也有的一致性维护策略考虑地理位置,因为地理位置相近的节点路径长度短,更新传播速度也快.一致性维护策略在提高一致性强度方面,需要考虑如下性能指标:

1) 版本偏差(Version deviation):版本偏差即各个副本节点持有的文件版本和最新文件版本的差异大小.版本偏差间接反映了一致性强度的大小,在一致

性维护过程中,应尽可能使得最新版本的更新传播至副本节点,减小版本偏差.

2) 更新覆盖范围:更新覆盖范围定义为接收到更新的节点与全部副本节点的比值.更新覆盖范围可以反映一致性维护策略的丢包率和加入结构算法的执行效果.

4 一致性维护策略

本节将重点介绍分布式更新策略,并深入剖析相关研究工作的基本原理以及面临的具体挑战.根据第2节提出的一致性维护策略分类标准,介绍几种典型的一致性维护策略:基于预设中心的一致性维护策略^[41],基于泛洪的一致性维护策略^[28],基于消息推送的一致性维护策略SCOPE^[16]和BCoM^[11],基于消息拉取的一致性维护策略GeWave^[22],基于先验信息的一致性维护策略^[5,22,29].考虑到目前数据量增长速度快,数据类型多样性增加,基于预设中心节点和基于泛洪的一致性维护策略分别由于其单点性能瓶颈问题和花费过高问题已不适合当前的网络环境,所以本节对这两种策略只作简单介绍,主要介绍其他三类一致性维护策略.

4.1 基于预设中心节点和泛洪的一致性维护策略

基于预设中心节点的一致性维护策略是P2P系统发展早期提出的一种一致性维护策略.早期的P2P系统规模小,存储数据量也较少,容易管理,因此,预设中心节点的一致性维护效率能显著地提高一致性强度.文献[41]采用预设中心节点的一致性维护策略,通过虚拟服务器提供一致性维护服务,每个动态文件的一致性都需要通过虚拟服务器进行一致性维护.这种方法的缺点十分明显,即中心服务器能力有限,并不能适用大规模网络.

基于泛洪的无向传播策略是人们对一致性维护过程中研究出来的一种策略,由于P2P网络是分布式的,节点可获得的其他节点的信息较少,在更新传播方面只能依靠泛洪的方式来传播.文献[28]采用基于泛洪的方式传播更新,节点在接收到更新后再把更新发送给自己的所有邻居节点,通过这样的方式把更新传播至副本节点.这种传播方式具有丢包少、传播速度快的优点.但缺点也非常明显:泛洪的方式由于其数据冗余量大可能引起网络风暴问题;为了限制消息的生命周期,通过附加TTL的方式来限制消息的传播,但TTL过大时数据冗余量过多,网络资源占用较多,而TTL过小又会使更新传播不到副本节点^[28].随着网络资源的紧缺和网络规模的增大,基于泛洪的一致性维护方式不再适用于当前P2P网络.

4.2 基于消息推送的一致性维护策略

在一致性维护过程中,当文件内容发生改变时,采用将更新主动推送给其他节点从而达到一致性维护的方式称为基于消息推送的一致性维护策略.基于推送的一致性维护策略相比其他一致性维护策略的优点是一致性强度高,一旦文件发生改变,更新信息便会推送给每个副本节点,使得更新可以及时地发送给每一个节点.

4.2.1 SCOPE和BCoM

SCOPE^[16]和BCoM^[11]拥有数据冗余小、一致性维护效率高的特点,他们通过把副本节点组织成树形结构来传播更新.在传播更新时,采用消息推送的方式分发更新,即只要文件发生改变,这次更新便会从树根推送至树中每个节点,这样可以达到较强的一致性强度.随着树高增加,每一层节点数呈指数增长,因此,更新会以很快的速度传至副本节点,提高了一致性维护效率.另外,由于树形结构的单连通性,每个节点只会接收到一份更新,减小了数据冗余.采用消息推送的方式分发更新存在的问题是:由于不是所有的节点都需要更新,而消息推送的方式无法根据每个节点的需求选择分发更新,消息推送更新会造成不必要的资源浪费.

这两种方法的不同点在于,SCOPE选择划分命名空间^[42],然后从每次划分的部分中选择代表节点作为树中的节点,这种划分一直持续到所有的副本节点都位于不同的分支中,这样划分命名空间构成的树称为副本分裂树(RPT)(如图4、图5所示).

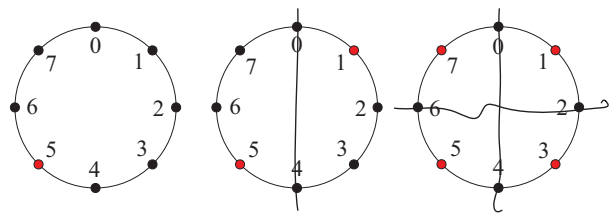


图4 SCOPE空间划分示意

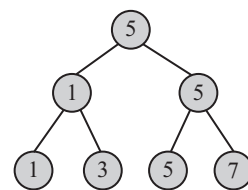


图5 SCOPE副本分裂树构建示意

BCoM的基本思想是根据副本节点的能力构建树形结构,能力越强的节点位置越接近根节点,这样构建成的树称为d度传播树(dDT)(见图6).为了维护dDT结构并优化其性能,BCoM又提出了两种方法:一种是快速恢复机制以提升整个结构对节点和链路

损坏的鲁棒性;一种是节点移动策略,使得性能好的节点移动到树的上层以提高一致性维护效率。

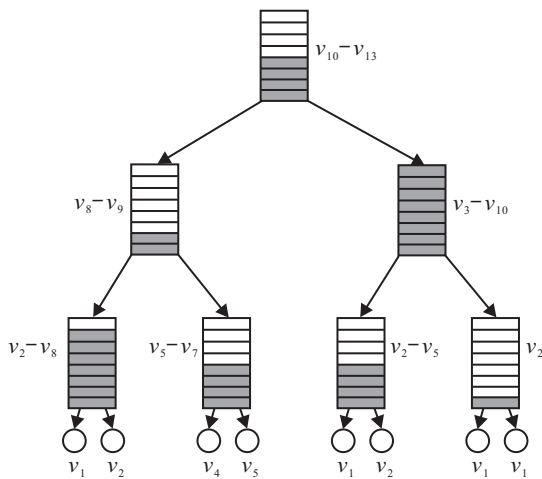


图6 BCoM传播树(2DT)

BCoM策略在一致性维护中提出了在节点中开辟缓存区缓存历史更新,这样做的好处是可以保证更新传播到每个有需要的节点,并且可以减少丢包,增强服务质量(QoS),使得一致性维护更有弹性。

在上述两种方法中,均采用树形结构来分发更新,但这两种方法都有各自的缺点:在SCOPE中,由于采用划分命名空间的方法,某些节点可能会多次存在于树中,当这类节点收到网络攻击或者自身故障时,更新便不能送达下层节点,从而引发丢包问题,降低QoS;在BCoM中,为了解决这个问题,采用了滑动窗口策略来缓存各个历史版本的更新,但同时也存在单点拥塞^[43]的问题,当某个节点由于性能低下而产生拥塞时,更新便会拥塞至根节点,当面对之后的更新时,由于根节点缓存区已满从而只能丢弃更新,这样会使得所有节点不能接收到更新从而引发更严重的问题。

4.2.2 方法分析和主要挑战

1) 方法分析:基于推送的一致性维护策略通过主动推送更新的方式达到维护一致性的目的。即文件一经改变,更新会以较快的速度主动推送至每个副本节点,使得每个副本节点都能在第一时间接收到更新。采用这种方法的一致性维护策略通常具有较高的一致性强度和较快的一致性维护效率。

2) 主要挑战:基于推送的一致性维护策略可以保持较高的一致性强度,但是这种方法会增加开销。在网络动荡较高的环境中,节点可能并不需要实时更新,在这种情况下,应根据节点的需求传播更新,以减轻网络负担并减小开销。当然也可以采取定时更新的方式减轻网络的负担,但更新间隔需要根据多方面

的因素确定。如实时数据传输、网络社交等对更新的渴求度较高,而网络工具的下载和使用等并不需要最新版本的文件,所以对更新的渴求度较低。总之,基于推送的一致性维护策略面临的最大挑战是:如何在满足节点对更新需求的情况下减少开销。如果采取文件改变就推送更新的方法,则会产生不必要的网络负载和开销;如果更新间隔设置过大,则会使得节点长时间得不到需要的更新,降低一致性强度。所以,如何根据网络环境和节点需要设置一个合适的更新间隔是一个有待研究的问题。

4.3 基于消息拉取的一致性维护策略

考虑到不同副本节点对于更新的需求情况不同,某些节点可能需要保持较高的一致性强度,但某些节点对于更新的渴求度并不高。在这种情况下,为了满足不同节点对于更新的需求同时保持较低的开销,副本节点应根据自身对于更新的需求主动询问并取得更新,即副本节点认为需要进行一次文件更新则主动向其他节点询问并取得更新,这种副本节点主动获得更新的方式称为基于消息拉取的一致性维护策略。该一致性维护策略相比其他一致性维护策略的优势是实现了供求平衡,即当节点需要更新时才会传播更新。在高动荡性的P2P网络中,节点可能并不需要实时更新,在这种情况下,比起基于消息推送的一致性维护策略,基于消息拉取的一致性维护策略可以达到更好的对网络扰动的容忍度和更低的开销。

4.3.1 GeWave

GeWave采用树形结构传播更新,但更新的传播方式是基于消息拉取的一致性维护策略,即由树中节点根据自身需要自主地发送更新请求消息,从而减小不必要的开销。GeWave应用于高动荡性和网络资源少的结构化P2P网络时,在树的构建中,节点通过两种属性决定自身所处的位置:一是更新时间(TTR),即节点对更新的渴求度,当节点比较频繁地询问更新时,其TTR越短,即更新渴求度越高,所以把TTR高的节点放置在下层,TTR低的节点放置在上层,这样可以在不产生不必要数据冗余的情况下满足节点的更新需求;二是地理位置,地理位置相近的节点之间数据传送速度较快,所以GeWave选择地理位置相近的节点在同一支,其好处是这一支的节点地理位置都比较接近,所以更新的传播速度较快。在经过这两个特点组织成树形结构以后,该树形结构具有以下特点:同一支的节点地理位置接近,数据传输效率高,自上向下节点的TTR依次增高,即对更新的渴求度降低,如图7所示。

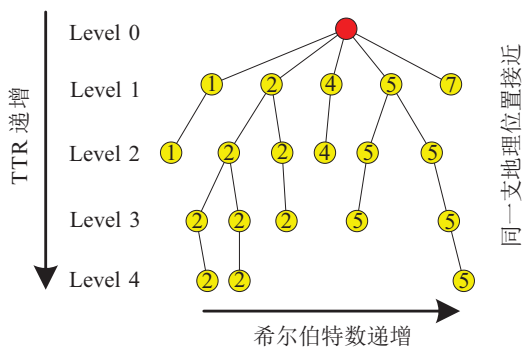


图7 GeWave分发树构建示意

GeWave运用自适应的消息拉取方式,当某个节点感觉自身的文件已过时,便向父亲节点询问更新信息^[22]. 记文件的最大更新率为 $1/\Delta_{max}$,意味着节点每 Δ 时间更新一次文件,即最快更新时间. 也就是说,节点可以通过每 Δ_{max} 时间更新一次文件以保证文件不过时. 记 $1/\Delta_{min}$ 为文件的最小更新率,文件拥有者可以通过历史记录获得 Δ_{max} 和 Δ_{min} ,当他复制文件给其他节点时,也会把这两个数值一并分发给其他节点. TTR的值根据最大更新率和最小更新率确定,采取线性增加乘法减少算法(LIMD)^[44]. 当某个节点需要更新时,节点会发送询问消息给父亲节点,父亲节点再转发这个消息给祖父节点,这样层层递进,直到询问消息到达根节点,根节点就会分发更新.

4.3.2 方法分析和主要挑战

1) 方法分析:基于消息拉取的一致性维护策略通过副本节点主动获取更新的方式完成一致性维护. 考虑到副本节点对于更新的需求不同,这种方式能够在满足不同节点对于更新需求的情况下减少网络资源占用和开销. 在数据冗余方面,该一致性维护策略的数据冗余通常大大低于基于消息推送的一致性维护策略.

2) 主要挑战:基于询问的一致性维护策略通过询问-分发的方式传播更新,所以这个方式面临的最大挑战是一致性强度较弱,不适用于更新频率较大的结构化P2P系统. 当更新频率较大时,节点需要首先发送询问消息然后等待更新到来,而文件主节点需要等待其他节点的询问消息才会分发更新,这种方式会浪费很多时间,从而更新的速度并不能满足节点所需的更新频率;并且当更新频率较高时,节点发送询问信息也会频繁,这样会给网络带来很大负载.

4.4 基于先验信息的一致性维护策略

为了解决一致性维护过程中的最优路径选择问题并减少一致性维护开销,人们利用结构化P2P网络中的许多先验信息来确定一致性维护策略^[31]. 通过先验信息确定副本节点的组成结构、聚类方式和更

新的分发方式的一致性维护策略即为基于先验信息的一致性维护策略. 在这些先验信息中,较为常用的有地理位置信息和用户兴趣信息. 由于覆盖网络的路径和现实的物理路径差别很大,可以根据节点的地理位置信息选择较短的数据传播路径从而达到较高的一致性效率和一致性强度,根据节点地理位置构建一致性维护结构可以减小开销,提高结构稳定性. 用户兴趣信息可以代表节点的活动范围和节点存储数据的类别,可以作为节点活动的预测基准. 文献[5, 22]提出的是基于地理位置信息的一致性维护策略,文献[29]则是基于兴趣的一致性维护策略.

4.4.1 基于地理信息的一致性维护策略

由于覆盖网络的链路并不是真实的物理链路,覆盖网络中的节点距离并不能刻画节点之间的物理距离. 基于覆盖网络的这种特点,在基于地理信息的一致性维护策略中,节点的地理位置信息通常被设定为已知的,所以,在之后的结构建立中,可以充分考虑节点的地理位置信息,使得数据传输发生在物理距离接近的节点之间. 文献[22]根据节点的地理信息把位置相近的节点布设在树的同一支上,使得在整个树中,地理位置接近的节点都处在同一支上,这样不论是在更新传播还是结构维护上都能以较短的通信距离完成(如图8所示).

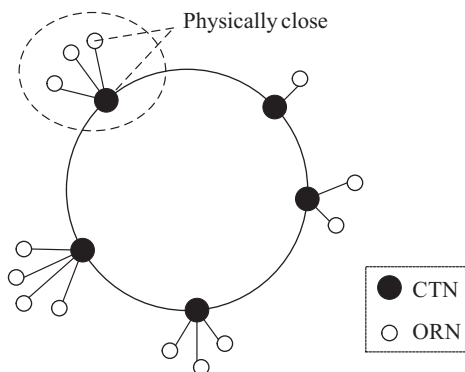


图8 分层结构

在地理位置数学化方面,采用地理位置标记法,给每个节点基于 m 维笛卡尔坐标空间生成一个 m 维的坐标,根据地理相近的节点拥有相近的地理坐标向量来确定节点位置和近似程度^[45]. 由于 m 向量计算复杂,算法采用空间填充曲线^[44]来简化地理位置信息. 利用希尔伯特曲线^[46]把 m 维笛卡尔空间映射到一维空间并且不损失近似度,这个一维空间中的数称为希尔伯特数,所以,两个地理位置接近的节点拥有近似的希尔伯特数.

文献[5]采用节点分类的方式组织副本节点,根据节点的地理位置把节点分组,并将每个组中的节

点分为两类CRN和ORN.其中:CRN具有更高的可靠性和更强的性能,ORN是与CRN距离相近的节点.然后组建分层网络,上层为各个CRN组成的网络,下层表现为其余ORN与CRN相连的结构(见图8).传播更新可以只考虑代表节点,当代表节点收到更新时会把更新分发给组内其他节点,从而减小开销和结构规模.

4.4.2 基于兴趣聚类的一致性维护策略

基于兴趣聚类的一致性维护策略通过节点的兴趣把节点分为不同的组,由于同一组中的节点有相同的兴趣,在之后的一致性维护和信息查询中可以根据兴趣选择不同的节点组进行一致性维护或者信息查询,这样可以提高效率并减少开销^[31].文献[29]采用这种方式分组节点,在对节点分组时,根据某个特定的关键字把节点分类,并将兴趣相同的节点分为同一集群,然后在这一组中根据节点所处地理位置把节点分为若干组,在每个组中选出一个代表节点(见图9).

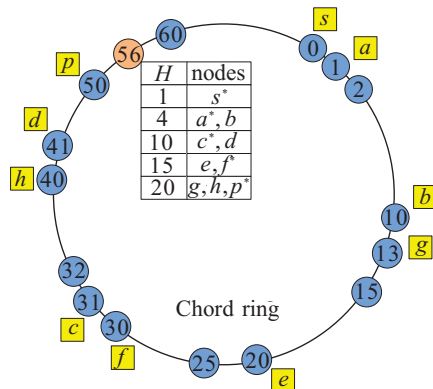


图9 基于地理位置和兴趣聚类信息示意

图9中所有节点拥有同样的兴趣,根据地理位置可以把它们分为5组,其中节点 s 是数据主节点,然后在每个组中选出一个代表节点分发更新给组中其他节点.在一致性维护中,把每个组中的代表节点组织成一个树,称为地理可知的平衡 d 度树(LBDT).

4.4.3 方法分析和主要挑战

1) 方法分析:基于先验信息的一致性维护策略充分利用P2P组网过程中的先验信息来设计一致性维护策略.在节点地理位置信息的帮助下,可以将节点按照地理最短距离的方式组织起来,从而在更新分发过程中减小网络资源占用并增强一致性维护效率.在节点存储内容和节点兴趣信息的帮助下,可以将相同兴趣的节点聚类成簇,从而减小信息查询时间并增加一致性维护效率.由于网络中先验信息种类很多,可以根据具体需求利用不同的先验信息达到一致性维护的目的.

2) 主要挑战:基于先验信息的一致性维护策略可以在很大程度上增强一致性维护效率并减小开销.这种方法的主要挑战是先验信息收集.在结构化P2P网络中,节点可以自由加入和离开,节点所存储的信息也经常变更,并且节点的某些信息会关系到节点的安全和隐私(如地理位置),因此节点可能会拒绝分享这类信息.另外,节点的某些信息的收集需要经历一个较长的阶段(如兴趣)才可以稳定下来,而这类信息又经常会发生变动,使得对节点活动的预测显得不是特别准确.因此,基于先验信息的一致性维护策略需要一个较好的信息收集和过滤算法来达到以上目的,并且如何缩短信息收集时间也是研究者需要考虑的问题.

4.5 适用场景和综合性能分析

通过对以上几种一致性维护策略的介绍和分析可以看到,不同的一致性维护策略在适用场景方面、实现的难易程度、优化目标等方面各有不同.本节将对上述一致性维护策略在适用场景和综合性能角度进行分析.

4.5.1 适用场景分析

不同的网络场景对一致性维护策略的要求不同,并且不同用户对于一致性维护策略的需求也不同.本文针对不同的网络场景和需求,对上述4种一致性维护策略进行分析,明确各个一致性维护策略在不同网络场景下的表现,对一致性维护策略的选择提出建议和帮助.

结合前文对于各种一致性维护策略的分析,这里对一致性维护策略在以下几个场景和需求下进行分析:一致性强度(Intensity)、对网络动荡的适应性(Adaptability to network turbulence)、信息查询和交换速度(Speed of information inquiry and exchange)、网络资源占用情况(Network resource occupancy)、一致性效率(Efficiency).

表3给出了这5个方面的对比分析,下面对几个场景和优化目标进行说明.

网络动荡适应性:由于P2P网络允许节点自由加入和退出,当节点加入和退出频繁时,路由信息更新可能会出现不匹配情况,在这种情况下分发更新会导致数据丢失.如何克服由于节点加入和退出而导致的一致性维护过程失败的能力称为一致性维护策略对于网络动荡的适应性.

信息查询和获取速度:网络中副本节点会根据自身需要查询并获取相应的更新,这个查询和获取过程的耗时随着不同的一致性维护策略而不同.

表3 一致性维护策略适用场景分析

分类	一致性强度	网络动荡适应性	信息查询和获取速度	网络资源占用	一致性维护效率
基于泛洪的一致性维护策略	低	高	低	高	低
基于消息推送的一致性维护策略	高	高	适中	低	高
基于消息拉取的一致性维护策略	适中	低	适中	适中	适中
基于先验信息的一致性维护策略	高	适中	高	低	高

网络资源占用情况:在更新传播过程中对于网络资源额占用(如带宽)大小。

4.5.2 综合性能分析

一致性维护策略旨在达到一致性维护的安全性、可靠性、快速性和低开销的目的,然而,当前的一致性

维护算法只能达到以上一个或几个目的,各种一致性维护策略都有着各自的优点和缺点. 本文从实现费用(Design cost)、计算复杂度(Complex)、实现难度(Difficulty)、数据冗余量(Redundancy)这4个方面对上述几类一致性维护策略进行分析,如表4所示。

表4 一致性维护策略综合性能分析

策略分类	实现费用	计算复杂度	实现难度	数据冗余量
基于泛洪的一致性维护策略	低	低	低	高
基于消息推送的一致性维护策略	适中	适中	适中	低
基于消息拉取的一致性维护策略	适中	适中	适中	低
基于先验信息的一致性维护策略	高	高	高	低

5 开放性问题

结构化P2P网络一致性维护策略的研究还存在以下开放性问题有待研究解决。

5.1 服务质量(QoS)

结构化P2P网络具有高度动荡性,使得保障QoS成为P2P网络组网策略的难点,而当前文件的动态变化更进一步提高了QoS保障的难度. 如P2P在线游戏^[47]、P2P在线社交^[48]等均需要一致性维护策略具有较好的QoS. 目前,一致性维护策略研究的趋势就是安全性、可靠性、快速性和低开销. 现有的许多工作仅优化了上述4个条件中的部分条件,并没有做到全面优化. 如BCoM^[11]利用树形结构传播更新达到了低时延,利用节点建立的缓存区提高了可靠性,但其开销较大;Swarm^[29]利用兴趣聚类的思想使得相同兴趣的节点在同一组,可以达到开销低、时延低的目的,但由于其传播更新时采用树形结构,其可靠性较差. 因此需要设计能够满足上述几个要求的一致性维护策略,从而提高QoS.

5.2 抗毁性设计

在一致性维护过程中,为了减少开销和提高速度,往往采用树形结构传播更新,但树形存在结构稳定性差、容易遭到破坏、破坏后修复难度大等缺点. 所以在更新传播过程中,一旦网络遭受恶意攻击或者节点离开,一致性维护效率会由于丢包和结构破

坏而变低. 为了提高一致性维护策略的抗毁性,结合当前计算机和通信技术的发展方向,可以预见在未来,存储设备的价格会逐步下跌,而网络带宽资源等会变得越来越紧缺^[49-50]. 所以在一致性维护策略的设计上,可以利用节点的存储设备来增强一致性维护策略的抗毁性,提高一致性维护强度,减小对网络资源的占用. 目前,对于利用节点存储空间来增加一致性维护的抗毁性已逐渐得到研究者重视. 如BCoM^[11]利用节点的存储空间构建一个更新滑动窗口来存储更新,在某些节点损坏的情况下仍然可以传播更新而不至于丢失更新信息,从而减少了丢包率,增强了一致性维护策略的抗毁性.

5.3 智能一致性维护策略

在一致性维护过程中,由于更新大小和重要性不同,副本节点对更新的需求程度也不同,现有的一致性维护策略都无法适应所有的网络场景,并且不能随着更新的大小和重要性动态调整. 例如:某次更新可能只更新很小的部分,而这次更新除了个别节点需求但其他大部分节点都不需要,此时,如何使更新只发送给需要的节点是一个值得研究的问题;另外,副本节点对每次更新的需求不同,所以每次传播更新时需要本次更新的节点数目都是一个动态变化的量,如何使得一致性维护策略能适应和解决这种动态性是目前存在的问题. 因此,需要设计一个具有学习能力的一致性维护策略,实现按需分发,平衡各节点负载.

表5 一致性维护策略各研究方向分析

分类	策略	优化目标	复杂度	优点	缺点
性能优化	负载均衡	使得系统的存储、带宽、一致性强度等要求基本一致	较低	资源必须得到基本保障	主要从保障性能方面考虑,其他方面验证以兼顾
	高可靠性	使各类资源的可靠性达到指定的具体要求	较低	业务可靠性得到保障	依据可靠性要求不同,可能增加相当数量的资源
	最大化满足用户需求	无具体目标函数表	低	按用户级别(优先级、安全级等)分类	只能定性分配,很难量化
节能优先	节能算法	最大程度降低一致性维护的资源消耗	高	能够降低一致性维护的资源消耗,节约成本	难以满足其他相关的性能(负载均衡、可靠性、服务质量)等需求
成本优先	最大化利用率	资源利用率最大,使得所有网络资源和节点资源得到最大程度的利用	较高	资源利用率较高(相应的成本较低)	难以满足相关的性能(负载均衡、可靠性、服务质量)等需求

5.4 更新重要性分析

由于一致性维护策略只负责传播更新,对更新不做分析,从而使得各种信息都会一视同仁地分发给副本节点.然而,更新信息并不是面向所有副本节点,可能仅仅针对某些节点,所以,对更新重要性的分析显得至关重要.由于更新重要性不同,重要的更新需要以更安全方式被传播至需求节点,为了达到安全性和可靠性,可以采用多点传输或者冗余的方式,这样必定会增加开销.另一方面,为了减少开销,重要性较弱的更新传播时可能会采取安全性较弱的方式.如何权衡更新的重要性与更新开销之间的关系需要人们做进一步的研究.因此,需要设计一种可以分析更新重要性的一致性维护策略,对更新信息进行优先级分类,使得重要更新可以更可靠地传播至需要的节点.

5.5 恶意节点检测

在一致性维护过程中,更新的传输大多采用存储-转发的方式,这就不可避免地产生了恶意节点篡改数据或者丢弃数据的问题^[51-52].因此,一套恶意节点检测方案必须应用于一致性维护策略中,以评估整个数据存储系统的可靠性和安全性.当前其他网络中大多采用基于信任值的可靠性评估^[53],基于投票的可靠性评估^[54],基于蒙特卡洛故障树的可靠性评估^[55],基于网络性能的可靠性评估^[56]等方法.恶意节点检测的加入可以提高一致性维护策略的可靠性,使得更新能够完整地分发到相应的节点上.

考虑上述开放性问题,并综合当前一致性维护策略的研究方向,本文将一致性维护策略的研究方向分为三大类:性能优先、节能优先和成本优先.结合各自的特点,从策略、优化目标、复杂度、优点、缺点等

几个方面进行分析(见表5).该分析结果对今后的一致性维护策略的研究和发展提供了方向和优化目标.

6 结语

随着对结构化P2P网络研究的深入,如何适应信息技术的发展,成为P2P网络设计和优化需要解决的迫切问题.作为结构化P2P网络最基础且必不可少的部分,一致性维护研究意义重大.本文对结构化P2P网络一致性维护策略相关研究工作从一致性维护研究的挑战、分类标准及系统模型、实现策略以及开放性问题等方面进行了分析和总结,重点对分布式一致性维护策略这个热点研究方向进行了详细剖析,并指出策略设计中需要注意的问题,希望能为研究者设计更加实用的一致性维护策略直至最终推动结构化P2P网络组网及实际部署给出有意义的启示.

参考文献(References)

- [1] Hu Y, Bhuyan L N, Feng M. Maintaining data consistency in structured P2P systems[J]. *IEEE Trans on Parallel and Distributed Systems*, 2012, 23(11): 2125-2137.
- [2] Khan M A, Debnath H, Borcea C. Balanced content replication in peer-to-peer online social networks[C]. *IEEE Int Conf on Big Data and Cloud Computing*. Atlanta: IEEE, 2016: 274-283.
- [3] Bharambe A, Douceur J R, Lorch J R, et al. Donnybrook: Enabling large-scale, high-speed, peer-to-peer games[J]. *ACM SIGCOMM Computer Communication Review*, 2008, 38(4): 389-400.
- [4] Meng X, Zhang C. An ant colony model based replica consistency maintenance strategy in unstructured P2P networks[J]. *Computer Networks*, 2014, 62: 1-11.
- [5] Li Z, Xie G, Li Z. Locality-aware consistency maintenance for heterogeneous P2P systems[C]. *Parallel and Distributed Processing Symposium*. California:

- IEEE, 2007: 1-10.
- [6] Hassanzadeh-Nazarabadi Y, Küpçü A, Özkasap Ö. Awake: Decentralized and availability aware replication for P2P cloud storage[C]. Int Conf on Smart Cloud. New York: IEEE, 2016: 289-294.
- [7] Trifa Z, Khemakhem M. A novel replication technique to attenuate churn effects[J]. Peer-to-Peer Networking and Applications, 2016, 9(2): 344-355.
- [8] Monni S L. Adaptive media replication in unstructured P2P file sharing systems based on geographical properties and query distributions[C]. Int Conf on Automated Solutions for Cross Media Content and Multi-channel Distribution. Lucca: IEEE, 2008: 171-179.
- [9] Navarro-Esteva A, Xhafa F, Caballe S. A P2P replication-aware approach for content distribution in E-learning systems[C]. Int Conf on Complex Systems. Krakow: IEEE, 2012: 917-922.
- [10] Kageyama J, Kobayashi M, Shibusawa S, et al. A file replication method based on demand forecasting in P2P networks[C]. The 2nd Int Conf on Applications of Digital Information and Web Technologies. London: IEEE, 2009: 268-274.
- [11] Guo L, Yang S, Zhang R, et al. RBMA: Replication based on multilevel-agent for P2P systems[C]. Int Symposium on Computer Network and Multimedia Technology. Wuhan: IEEE, 2009: 1-5.
- [12] Xhafa F, Potlog A D, Spaho E, et al. Evaluation of intra-group optimistic data replication in P2P groupware systems[J]. Concurrency and Computation: Practice and Experience, 2015, 27(4): 870-881.
- [13] Zhang H, Wen Y, Xie H, et al. Distributed Hash table[M]. New York: Springer, 2013.
- [14] Balakrishnan H, Kaashoek M F, Karger D, et al. Looking up data in P2P systems[J]. Communications of the ACM, 2003, 46(2): 43-48.
- [15] Rahmani M, Benchaïba M. A comparative study of replication schemes for structured P2P networks[C]. Proc of the 9th Int Conf on Internet and Web Applications and Services. Lisbon, 2014: 147-158.
- [16] Chen X, Ren S, Wang H. SCOPE: Scalable consistency maintenance in structured P2P systems[C]. The 24th Annual Joint Conf of the IEEE Computer and Communications Societies. Miami: IEEE, 2005, 3: 1502-1513.
- [17] Li Z, Weis T. CAN tree routing for content-addressable network[J]. Sensors & Transducers, 2014, 162(1): 124.
- [18] Shen H. IRM: Integrated file replication and consistency maintenance in P2P systems[J]. IEEE Trans on Parallel and Distributed Systems, 2010, 21(1): 100-113.
- [19] Xiong J, Hu Y, Li G, et al. Metadata distribution and consistency techniques for large-scale cluster file systems[J]. IEEE Trans on Parallel and Distributed Systems, 2011, 22(5): 803-816.
- [20] Tang X, Zhou S. Update scheduling for improving consistency in distributed virtual environments[J]. IEEE Trans on Parallel and Distributed Systems, 2010, 21(6): 765-777.
- [21] Li Z, Xie G, Li Z. Efficient and scalable consistency maintenance for heterogeneous peer-to-peer systems[J]. IEEE Trans on Parallel and Distributed Systems, 2008, 19(12): 1695-1708.
- [22] Shen H, Liu G. A geographically aware poll-based distributed file consistency maintenance method for P2P systems[J]. IEEE Trans on Parallel and Distributed Systems, 2013, 24(11): 2148-2159.
- [23] Suto K, Nishiyama H, Kato N, et al. THUP: A P2P network robust to churn and DoS attack based on bimodal degree distribution[J]. IEEE J on Selected Areas in Communications, 2013, 31(9): 247-256.
- [24] Elbreiki W, Hassan S, Habbal A, et al. A Comparative study of chord and pastry for the name resolution system implementation in information centric networks[C]. The 4th Int Conf on Internet Applications, Protocols and Services. Kuala Lumpur: IEEE, 2015: 359-367.
- [25] Eichert F A, Monhof M, Graffi K. The impact of routing attacks on pastry-based P2P online social networks[C]. European Conf on Parallel Processing. Cham: Springer, 2014: 347-358.
- [26] Zhao B Y, Huang L, Stribling J, et al. Tapestry: A resilient global-scale overlay for service deployment[J]. IEEE J on Selected Areas in Communications, 2004, 22(1): 41-53.
- [27] Seddiki M, Benchaïba M. SWS: A smart walk mechanism for resources search in unstructured mobile P2P networks[C]. The 1st Int Conf on New Technologies of Information and Communication. Cairo: IEEE, 2015: 1-6.
- [28] Lee W, Kim T Y, Kang S, et al. Revised P2P data sharing scheme over distributed cloud networks[C]. Information Science and Applications. Berlin: Springer, 2015: 165-171.
- [29] Shen H, Liu G, Chandler H. Swarm intelligence based file replication and consistency maintenance in structured P2P file sharing systems[J]. IEEE Trans on Computers, 2015, 64(10): 2953-2967.
- [30] Ng C H, Sia K C, King I. Peer clustering and firework query model in the peer-to-peer network[C]. The 12th Int World Wide Web Conf. Budapest, 2003: 237-249.
- [31] Shen H, Liu G, Ward L. A proximity-aware interest-clustered P2P file sharing system[J]. IEEE Trans on Parallel & Distributed Systems, 2015, 26(6): 1509-1523.
- [32] Seddiki M, Benchaïba M. 2P-Lookup: Popularity and proximity based P2P Lookup mechanism over MANETs[J]. J of Network and Computer Applications, 2016, 71: 181-193.
- [33] Roussopoulos M, Baker M. CUP: Controlled update propagation in Peer-to-Peer networks[J]. Computer

- Science, 2002, 21(6): 1-6.
- [34] Buchegger S, Schiöberg D, Vu L H, et al. PeerSoN: P2P social networking—Early experiences and insights[C]. Proc of the 2nd ACM Workshop on Social Network Systems. Nuremberg: ACM, 2009: 46-52.
- [35] Althobaiti K, Alotaibi S J, Alqahtani H. Random walk with jumps: A new query search method based on analysing Gnutella protocol[C]. The 2015 World Congress on Internet Security. Cambridge: IEEE, 2015: 125-130.
- [36] Lan J, Liu X, Shenoy P, et al. Consistency maintenance in peer-to-peer file sharing networks[C]. The 3rd IEEE Workshop on Internet Applications. California: IEEE, 2003: 90-94.
- [37] Thangasamy P, Jenni S, Aiswarya. An efficient approach of integrated file replication and consistency maintenance in Peer-to-Peer systems[J]. Int J of Computational Engineering Research, 2014, 11(4): 10-14.
- [38] Nakashima T, Fujita S. Scalable tree-based consistency maintenance in heterogeneous P2P file sharing systems[C]. The 44th Int Conf on Parallel Processing Workshops. Bristol: IEEE, 2015: 250-256.
- [39] Nobre J C, Granville L Z. Consistency maintenance of state of management data in P2P-based autonomic network management[J]. Peer-to-Peer Networking and Applications, 2016, 11(3): 213-226.
- [40] Kalaivani E, Kumar J S, Coimbatore C. An efficient caching scheme and consistency maintenance in hybrid P2P system[J]. Int J of Computer Networks and Wireless Communication, 2012, 2(1): 71-80.
- [41] Wang Z, Kumar M, Das S K, et al. File consistency maintenance through virtual servers in P2P systems[C]. The 11th IEEE Symposium on Computers and Communications. Sardinia: IEEE, 2006: 435-441.
- [42] Woungang I, Tseng F H, Lin Y H, et al. Mr-chord: Improved chord lookup performance in structured mobile p2p networks[J]. IEEE Systems J, 2015, 9(3): 743-751.
- [43] Ramasubramanian V, Sireer E G. Beehive: Exploiting power law query distributions for $O(1)$ lookup performance in peer to peer overlays[C]. Symposium on Networked Systems Design and Implementation. San Francisco, 2004: 864-875.
- [44] Khayyat Z, Awara K, Alonazi A, et al. Mizan: A system for dynamic load balancing in large-scale graph processing[C]. Proc of the 8th ACM European Conf on Computer Systems. Amsterdam: ACM, 2013: 169-182.
- [45] Li-Haixia, Lu-Chuiwei, Sun-Sheng. Intelligent network storage on the big data[C]. Int Conf on Multisensor Fusion and Information Integration for Intelligent Systems. Baden-Baden: IEEE, 2014: 1-6.
- [46] Su T, Wang W, Lv Z, et al. Rapid delaunay triangulation for randomly distributed point cloud data using adaptive Hilbert curve[J]. Computers & Graphics, 2016, 54: 65-74.
- [47] Gibson M, Vasconcelos W W. Effects of knowledge base quality on peer-to-peer information propagation[J]. Procedia Computer Science, 2015, 52: 350-357.
- [48] Buchegger S, Schioberg D, Vu L H, et al. Implementing a p2p social network-early experiences and insights from peerson[C]. The 2nd ACM Workshop on Social Network Systems. New York, 2014: 16-24.
- [49] Li X, Pan M, Song Y, et al. Economic-robust session based spectrum trading in multi-hop cognitive radio networks[C]. 2015 IEEE Global Communications Conf. San Diego: IEEE, 2015: 1-6.
- [50] Kim Y, Lee J, Jeong J, et al. Multi-flow rate control in delayed Wi-Fi offloading systems[C]. 2016 Int Conf on Information Networking. Chiang Mai: IEEE, 2016: 274-279.
- [51] Tsuda T, Komai Y, Sasaki Y, et al. Top- k query processing and malicious node identification against data replacement attack in MANETs[C]. The 15th Int Conf on Mobile Data Management. Shenzhen: IEEE, 2014, 1: 279-288.
- [52] Li S, Su W, Li H. The research of security of P2P network file sharing system[C]. The 29th Int Conf on Advanced Information Networking and Applications Workshops. Gwangju: IEEE, 2015: 250-254.
- [53] Bao F, Chen R, Chang M J, et al. Hierarchical trust management for wireless sensor networks and its applications to trust-based routing and intrusion detection[J]. IEEE Trans on Network and Service Management, 2012, 9(2): 169-183.
- [54] Wang L, Li Z, Ren S, et al. Optimal voting strategy against rational attackers[C]. The 6th Int Conf on Risk and Security of Internet and Systems. New South Wales: IEEE, 2011: 1-8.
- [55] Zhang P, Chan K W. Reliability evaluation of phasor measurement unit using Monte Carlo dynamic fault tree method[J]. IEEE Trans on Smart Grid, 2012, 3(3): 1235-1243.
- [56] Zhang H, Huang N, Liu H. Network performance reliability evaluation based on network reduction[C]. 2014 Annual Reliability and Maintainability Symposium. Washington: IEEE, 2014: 1-6.

(责任编辑: 李君玲)