

Research on Reality of Distributed Storage System Based on Cloud Model

ZHANG Min-Qing¹, DONG Bin¹, ZHANG Wei^{1,2}, YANG Xiao-Yuan^{1,2}

1. Key Laboratory of Network & Information Security of APF, Engineering College of APF, Xi'an 710086, China

2. Key Laboratory of Network & Information Security of the Ministry of Education, Xidian University, Xi'an 710071, China

1.wj_dongbin@126.com

Abstract: Survival of Distributed Storage System is becoming the latest important concerns along with extensive application of Storage System. Nowadays, There are so many kinds of Storage System, such as DAS, NAS, SAN and so on, all of these systems must work in a well state, but how to describe this state accurately. People proposes the concept of Reality in order to depict persistent service ability of a Storage system. Reality of distributed storage system can be affected by many factors, such as invalidation frequency of node, Storage strategy and so on. Nowadays, there isn't a mature algorithm to measure the reality of distributed storage system. This paper proposes a quick measure algorithm and a graded cloud model. The results of experiments show that this graded cloud model proposed is valid.

Keywords: Distributed storage system; Cloud model; Reality

基于云模型的分布式存储可靠性模型研究

张敏情¹, 董斌¹, 张薇^{1,2}, 杨晓元^{1,2}

1. 武警工程学院电子技术系网络与信息安全武警部队重点实验室, 西安, 中国, 710086

2. 西安电子科技大学网络信息安全教育部重点实验室, 西安, 中国, 710071

【摘要】: 随着分布式存储系统的广泛应用, 系统的可生存性成为当今研究的热点问题。如今 DAS, NAS, SAN 等存储结构被大量使用, 如何使系统处于良好工作并且精确描述系统状态, 人们提出了系统可靠性的概念。分布式存储系统的可靠性受节点失效概率、存储策略等诸多因素影响, 因此目前对分布式存储系统的可靠性的度量的算法还不成熟。根据隶属云理论, 提出了不同等级的可靠云模型以及计算云可靠性的快速算法, 给出了分布式存储系统可靠性的计算公式。模拟实验表明该度量模型是有效的。

【关键词】: 分布式存储系统; 云模型; 可靠性;

1 引言

分布式存储系统是现代存储技术与计算机网络技术相结合的产物, 它以网络技术为基础, 将服务器系统的数据处理与数据存储相分离, 实现对数据的海量存储。文献[1]提出门限Quorum系统的交汇性可以较好地满足分布式系统中对共享数据一致性的要求, 因而常常被用于构建对可靠性要求较高的分布式数据服务。然而对于分布式存储系统可靠性的度量目前还不成熟。

文献[2, 3]给出了两种基于信任评估模型, 通过建立模型对主体的信任度进行定量分析, 从而将模

糊问题转化为具体问题。90年代初期李德毅院士在传统模糊数学和概率统计的基础上提出了定性定量互换模型即云模型, 它主要反映宇宙中事物或人类知识中概念的2种不确定性: 模糊性和随机性, 它把模糊性和随机性有机集合在一起, 研究自然语言中的最基本的语言值所蕴含的不确定性的普遍规律, 使得人们可以把自己对事物的认识转化为定量的数值来描述, 使得所描述的信息更加适合统计和分析^[4, 5, 6]。

本文主要研究云模型在分析分布式存储系统可靠性, 通过建立模型, 将人们对Quorum系统各项性能的分析转化为定量的描述, 再将其置入数学模型

基金项目: 国家自然科学基金资助项目(60573032)。

中分析,最后得出系统可靠性的结论。这种度量分布式系统可靠性的方法经过模拟实验测试结果是有有效的。

2 预备知识

2.1 云的基本概念

设U是一个用精确数据值表示的定量论域,C是U上的定性概念,若定量值 $x \in U$,且T是定性概念C的一次随机实现,x对T的隶属度(确定度)为是具有稳定倾向的随机数,则x在论域U上的分布称为云,每个x称为一个云滴($x, \mu_T(x)$)。云是用语言值表示的某个概念与其定量表示之间的不确定性转换模型,用以反映自然语言中概念的不确定性,反映了随机性和模糊性的关联性,构成定量与定性间的映射。

2.2 云的特征

云的数字特征用期望Ex,熵En和超熵He来体现,它们反映了定性概念的定量特征。

期望Ex(Expectation):在论域空间中最能够代表这个定性概念的点,是这个概念量化的最典型样点。

熵En(Entropy):代表一个定性概念的可度量粒度,通常熵越大概念越宏观。熵还反映了定性概念的不确定性,表示在论域空间可以被定性概念接受的取值范围大小即模糊度。

超熵He(Hyper entropy):熵的不确定性的度量,它反映代表定性概念值的样本出现的随机性,揭示了模糊性和随机性的关联。

由中心极限定理可知在客观世界中有许多随机变量,它们由大量的相互独立的随机因素的综合影响所形成的。而其中的每一个单位因素在总的影响中所起的作用都是微小的。这种随机变量可以近似看成服从正态分布^[7]。本文所构造的模型为正态云模型。

2.3 正态云分布定理

根据正态分布特性,若论域中所有元素构成的云滴组成概率为1的整体云,则云滴落入 $[Ex-3En, Ex+3En]$ 之外区域可视为小概率事件,概率可以忽略。

2.4 云的算法

云有两个基本算法为正向云发生器和逆向云发生器。正向云发生器是将主体对事物主观描述转化为定量描述,又称为云滴,输入为期望值Ex、熵En、超熵He、云滴数N,输出是N个云滴在数域空间的定量位置及每个云滴代表该概念的确定度。

逆向云发生器是将定量描述转化为主观认识,

输入为其N个云滴在数域空间的精确位置和每个云滴代表该概念的确定度,输出是这N个云滴表示的定性概念的期望值Ex,熵En和超熵He。

云发生器实现方法:

1)生成一个期望值为En,标准差为He的正态随机数 En' ;

2)生成一个期望值为Ex,标准差为 En' 绝对值的正态随机数x;

3)令x为定性概念的一次具体量化值,称为云滴;

4)计算 $y = e^{-\frac{(x-Ex)^2}{2(En)^2}}$;

5)令y为x属于该定性概念的确定度;

重复1)~5)步,直到生成N个所需云滴。

逆云发生器实现方法:

根据N个云滴计算:样本均值 $B = \frac{1}{N} \sum_{i=1}^N |x_i - \bar{x}|$;

计算样本方差: $S^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2$, $E\hat{x} = \bar{x}$,

$En = (\pi/2)^{1/2} \times B$, $He = (S^2 - En^2)^{1/2}$ 。

3 可靠性度量模型

3.1 基本思想

在分布式存储网络中,可靠性的度量是根据用户对存储系统在存取效率、安全性、容错能力、平均无故障时间(MTBF)、平均故障修复时间(MTTR)等方面对系统进行综合评价,根据这些评价数据建立模型。由于Quorum系统是由若干个Quorum构成,每个Quorum中还存在多个存取节点,所以在度量整个Quorum系统的可靠性时,要先计算每个Quorum的可靠性,而后再综合整体的可靠性。

3.2 参考云模型

定义1: 设 $\Omega = \{x\}$ 为研究论域,存在算法 τ ,使得 $\tau(\Omega) = (\Omega_1, \dots, \Omega_j)$ and $\forall i, j \in Z \Omega_i \cap \Omega_j = \emptyset$,则称 Ω_i 为一个参考论域,在参考论域上建立的云模型称为参考云。

定义2: 可靠度空间 $Grd(Q_i) = [0, 1]$, 其中若 $Grd(Q_i) = 0$, 则称系统 Q_i 完全不可靠, 若 $Grd(Q_i) = 1$, 则称系统 Q_i 完全可靠。

定义3: 可靠性等级 反映分布式存储系统可靠性的主要指标,是系统服务能力的综合体现。本文将系统可靠性分为五个等级。

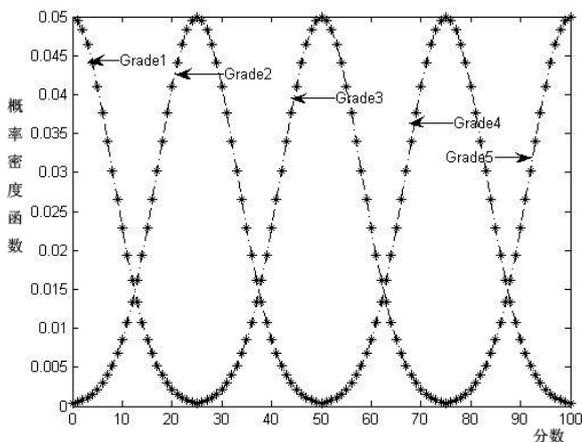
Grade1: 系统完全可靠。

Grade2:系统非常可靠,能够完成几乎全部存储任务。

Grade3:系统比较可靠,可以完成存储任务。

Grade4:系统需要进行可靠性维护,但仍能提供服务。

Grade5:系统不可靠,需要立即维护。



图一: 参考云图

3.3 可靠性度量算法

在分布式存储可靠性度量模型中,用户对系统的不同属性进行量化打分,再乘以相应属性权值(Wi),根据分数进行建模归类,本文采用文献[8]中逆向云算法。

输入: n个用户对Quorum第j属性的评价值 $\Omega_i = (X1j, X2j, \dots, Xnj)$ 。

输出: 产生Quorum第j属性的可靠云模型 $R_j(Ex_j, En_j, He_j)$ 。

根据输出产生的可靠云模型,给出以下三个可靠性计算方法:

算法1: 若 $R_j(Ex_j, En_j, He_j)$ 为可靠模型云, $R_c(Exc, Enc, Hec)$ 为基云, 令

$$\eta_{jc} = 1 - |(Ex_j - Exc) + 3(Enc - En_j)| / 6Enc$$

。可见若建立模型中有大量的云滴落入基云中,则两云越接近, η_{jc}

越接近与1。但是 η_{jc} 为1只是说明云模型 R_j 与基云 R_c

最接近,并不意味着两者相同。算法1的定义只给出了一种快速确定可靠度的算法,不适合用来精确衡量

两者的相似度。若某一模型与相邻两基云 η_{jc} 相

等,则可选择较高可靠性等级。

算法2: 单个Quorum可靠性算法

$$Grd(Q_i) = \sum_{j=1}^m W_j R(Q_{ij}), \quad \sum_{j=1}^m W_j = 1$$

其中 $R(Q_{ij})$ 为 Q_i 第j个属性的可靠性, w_j 为第j个属性的权值, m为此

Quorum中节点个数。

算法3: 分布式系统可靠性计算公式

$$Grd(S) = \sum_{i=1}^n P_i Grd(Q_i), \quad \sum_{i=1}^n P_i = 1$$

其中 P_i 为 Q_i 的使用概率, n为分布式系统中包含Quorum个数。

4 实验与讨论

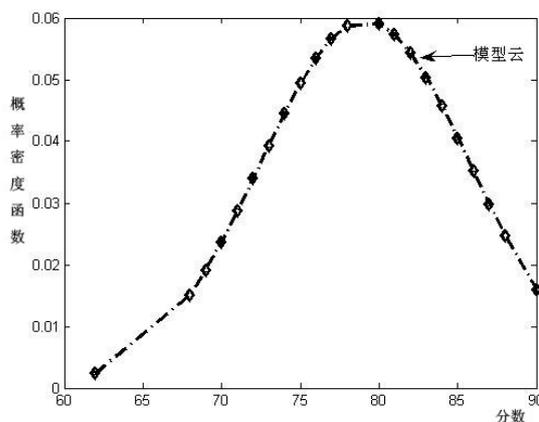
4.1 基本思想

实验假设一个分布式存储系统有100组Quorum组成,每次客户请求对若干个Quorum进行读写操作,读写操作后客户给所用Quorum可靠性评价,根据所有客户的可靠性评价进行量化,形成可靠云模型,再与基云相比较,得到此Quorum的可靠性。再根据本分布式存储系统的存取策略,利用算法3计算出整个系统的可靠性。

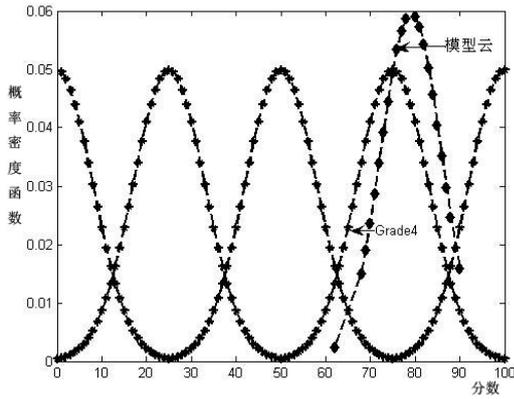
4.2 实验内容

4.2.1 存储系统可靠性度量

本文选择100组Quorum构成存储系统,每组Quorum有100个客户进行可靠性评价,若用户认为此Quorum可靠则选择[50-100]随机数,若用户认为不可靠则选择[0-50]随机数。可靠性分为五个等级,分别为 $RC1(0,8,0.5), RC2(25,8,0.5), RC3(50,8,0.5), RC4(75,8,0.5), RC5(100,8,0.5)$ 。本实验采用Matlab7.0作为实验工具,共10轮每组100个可靠性度量数据,对单个Quorum进行统计分析。图一为一组模型云图,图二为此模型云与基云的对比图。

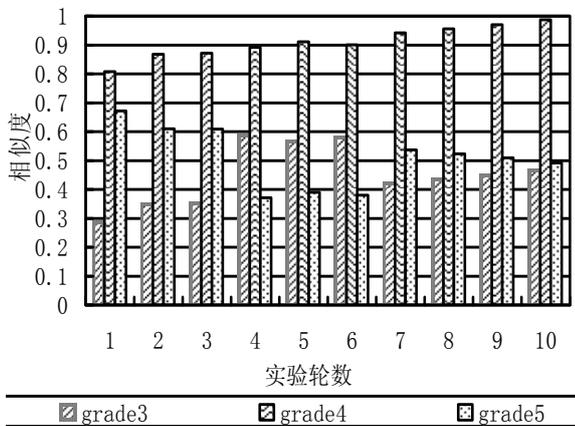


图二 模型云图



图三 模型云与基云对比图

从图三中可知模型云与Grade4基云最接近，因此本次试验可靠性结果可以认为是Grade4。图四给出了十次上述试验得出结果。



图四 十次试验结果比较

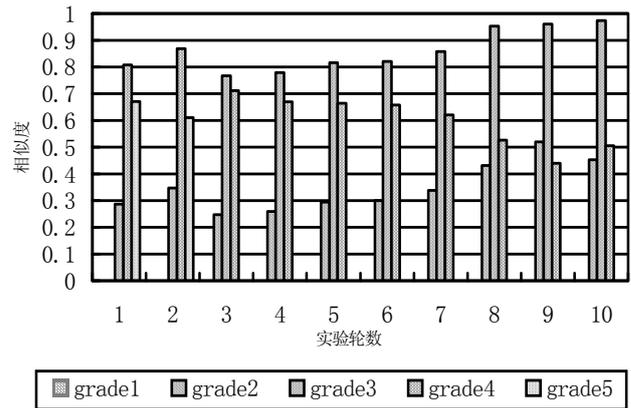
由图四可以看出，随着实验轮数的增加可靠性结果趋于稳定，由此可见此可靠模型是有效的。整个存储系统可靠性可由4.1节算法2得出，其中 $n=100$ 。

4.2.2 突发故障模型

实验仍在4.2.1假设前提下进行，若某一时刻一Quorum突发的Byzantine故障数大于可容忍故障节点数，系统将无法提供服务。模拟实验结果如图五所示，可以看出此Quorum在第三次实验中发生Byzantine故障，可靠性等级由grade4降低至grade3。

5 结束语

本文根据云模型理论结合分布式存储实际，给出了一种分布式存储系统可靠性度量模型，结合模拟实验结果表明，该模型在正常状态和突发Byzantine故障时对存储系统可靠性的度量是有效的。



图五 Byzantine 故障情况下可靠性变化图

参考文献:

- [1] Wei ZHANG, Jianfeng MA, Threshold Byzantine Quorum System and Distributed Storage[J], ACTA ELECTRONICA SINIC,2008(36):314-319.
张 薇, 马建峰等, 门限 Byzantine Quorum 系统及其在分布式存储中的应用[J]电子学报 2008 (36): 314-319.
- [2] Haisheng HUANG,Ruchuan WANG, Subjective trust evaluation model based on membership cloud theory[J], Journal on Communications,2008 (29):13-19.
黄海生, 王汝传, 基于隶属云理论的主观信任评估模型研究[J]通信学报 2008 29(4):13-19
- [3] Xiangyi MENG,Guangwei ZHANG, et al. Research on Subjective Trust Management Model Based on Cloud Model [J], Journal of System Simulation, 2007 (14): 3310-3317.
孟祥怡, 张光卫等, 基于云模型的主观信任管理模型研究[J]系统仿真学报 2007 19(14): 3310-3317
- [4] Deyi LI, Yi DU. Artificial Intelligence with Uncertainty [M]. Beijing: National Defence Industry Press, 2005.
李德毅, 杜鹤. 不确定性人工智能[M], 北京: 国防工业出版社, 2005.
- [5] Deyi LI, DW Cheung, Xuemei Shi, et al. Uncertainty reasoning base on cloud models in controllers[J]. Computers and Mathematics with Applications, 1998, 35(3): 99-123.
- [6] Deyi LI, Jiawei HAN, Xuemei Shi, et al. Knowledge representation and discovery based on linguistic atoms[J]. Knowledge-based Systems, 1998, 10(7):431-440.
- [7] Zhou Sheng,Shiqian Xie,Chengyi Pan, Probability theory and mathematical statistics(the second edition)[M].Beijing:Higher Education Press,1989.
盛骤, 谢式千, 潘承毅 概率论与数理统计(第二版)[M], 北京:高等教育出版社,1989.
- [8] Huijun Lv, Ye Wang, Deyi LI, The application of backward cloud in qualitative evaluation[J], Chinese Journal of Computers.吕辉军, 王晔, 李德毅. 逆向云在定性评价中的应用[J].计算机学报, 2003, 26(8): 1009-1014.