CYCLIC SPRING PROTOCOL - A QUORUM-BASED APPROACH FOR REPLICA CONTROL

Cheng-Hong Cho and Shi-Hong Lin

Department of Electrical Engineering
National Taiwan University of Science and Technology
Email:chc@olemap.com

ABSTRACT

Data replication is an important technique in distributed system. Most replication techniques are quorum-based approaches. These approaches employ logical structures or mathematical methods to solve the consistent problem of data replication. They have some improvements in getting smaller quorum size, higher availability, and better loading balance. However, most of them have some of the following disadvantages: (1) The load distribution is unbalance. (2) These methods do not applied to any arbitrary number of sites. This paper presents a new approach called cyclic spring protocol, an approach based on a circular numbering system. This protocol is symmetric that every site in the system bears the same responsibility. In addition, it is applicable to any arbitrary number of sites.

1. INTRODUCTION

A distributed system is a system that consists of a number of interconnected computers. These computers share their resources with one another so that they can work together to accomplish users' requirements. To cooperate more efficiently, data in a distributed system are often spread over different computers (sites). This approach is; however, less reliable because a successful data access must rely on a good network. To improve the reliability, data are replicated and scattered around so that they are available at various sites. Even if some copies of the needed data are temporary out of reach, alternative copies are still available at other sites.

The major problem in data replication is how to guarantee that people can always get the most up-to-date information in such a distributed system. A simple approach is to replicate all data at all the sites. This allows its users to access the same data at any available site. However, it is very expensive in maintaining the data since every update request must be performed at all the sites. Alternative approaches reduce the number of data copies that need to be updated while increasing the number of accessed data copies to ensure a most recent copy.

The replica control to maintain the consistency is that any two write operations or any pair of read and write operations must access at least one common data copy in executing time. Thus it can guarantee that a read operation can always retrieve the most current data and a write operation can be admitted unless it conflicts with others. The goal of replication is to provide good efficiency, high availability, load balance, and fault tolerance in distributed systems.

Most replication techniques are quorum-based approaches. The quorum-based replica control defines a set of quorums, where each quorum consists of a set of nodes. The term node is used to refer to a physical copy of replicated data in a distributed system. In this system, each copy of the node is tagged with a version number. This protocol defines the read quorums and write quorums for read and write operations respectively. If the operation is read, the copy with highest version number in the quorum is the most up-to-date one. If the operation is write, all the copies in the quorum are updated. In order to ensure the data consistency, every write quorum must intersect any read quorum or any write quorum.

Some quorum-based protocols impose logical structures on their systems such as tree protocol [1, 2], grid protocol [3, 4], triangular lattice protocol [5], and triangular grid protocol [6], etc. Some quorum-based protocols don't employ the logical structure on their system. They may use mathematical method to solve the problem such as the read-write difference pair protocol [7]. All these methods attempt to reduce the quorum size, to improve the availability, or to balance the load among the sites. However, it is very hard to be successful at all aspects because every approach has its priority.

In this paper, we propose an efficient approach to get the access quorums (uniform read and write quorums). This approach is based on a circular numbering system. A distributed system with N nodes is organized as a logical circle, denoted by N-ring. Every node in the system is assigned a distinct number between 1 to N and is arranged by its number sequentially. Some sequences of particular patterns are employed as read and write quorums. This approach is not only highly available but also symmetric. However, to achieve the high availability, we need to slightly increase the quorum size.

This paper is organized as follows. In Section 2, we describe some quorum-based replica control protocols that have been proposed. In Section 3, we propose cyclic spring protocol. We present its definitions, algorithms and properties. In Section 4, we analyze our protocol. In Section 5, we conclude the result.

2. RELATED WORKS

Consider a distributed system consisting of N distinct nodes (sites) that are linked by a communication network. Every node has its own computing capability. These nodes communicate with one another in the network by exchanging messages.

This section briefly introduces some replica control approaches: read-one-write-all protocol, weighted voting protocol [8, 9], grid protocol, triangular grid protocol, circle grid protocol [10], and read-write difference pair protocol.

2.1 Read-one-write-all Protocol

The simplest protocol for managing replicated data is read-one-write-all protocol (ROWA). In this protocol, read operations can be executed on any copy of the replicated data, but write operations is required to write all the copies of the replicated data. It provides the best reading technique but the worst writing circumstances for its users. Sites that carry copies of the replicated data may fail at any time for some reasons. These failed sites will prohibit database users from updating data since write operations are required to perform over all the copies of the replicated data.

2.2 Weighted Voting Protocol

In weighted voting protocol, each copy of a replicated data object is assigned a certain number of votes. A read operation has to collect a read quorum of r votes to read a data item and a write operation has to collect a write quorum of w votes to write a data item. There are two restrictions in this protocol:

(1) 2w > total number of votes.

(2) r + w > total number of votes.

The major achievement of this protocol is that it is fault tolerant. The system can keep working even if there are some failure sites. A copy that was failed will not be accessed in the successive transaction and will not have the largest version number. Hence, after it has recovered, it will not be read until it has been written at least once.

2.3 Grid Protocol

Grid protocol arranges all the nodes of an N-node system into an m×n(=N) grid structure. In this protocol, a read operation has to collect a read quorum formed by a column-cover (at least a node from each column) or a complete column of the grid. A write quorum is composed of a read quorum and a complete column of the grid.

For example, a 3×4 grid with 12 nodes is shown in Figure 2.1. In this example, {1, 2, 3, 4}, {5, 10, 3, 8}, {9, 10, 7, 4} are read quorums, and {1, 2, 3, 4, 5, 9}, {3, 4, 5, 8, 10, 12}, {3, 4, 7, 9, 10, 11} are write quorums.

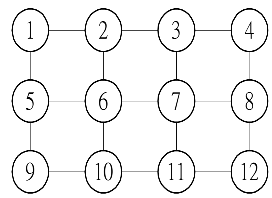


Figure 2.1. a 3×4 grid structure with 12 nodes

2.4 Triangular Grid Protocol

The geometric topology shown in Figure 2.2 is denoted as a triangular grid. This protocol arranges the N-node system into a triangular grid of height h where h = $\left\lceil \frac{\sqrt{8N+1}-1}{2} \right\rceil$. The term height refers to the total layers of a triangular grid. A boundary-cover-tree (BCT) of a triangular grid is a tree which inside the triangular grid and contains at least one node from each boundary line of the triangular grid. A minimal BCT is a BCT, which contains just h nodes. In this protocol, all read quorums are also write quorums, and they are therefore called access quorums. The quorum size of this protocol is h, the same as the height of triangular grid structure.

For example, a triangular grid with 15 nodes is shown in Figure 2.2. The set {1, 2, 4, 7, 11}, {2, 3, 5, 8, 12}, {7, 8, 9, 10, 15} are some access quorums in this protocol.

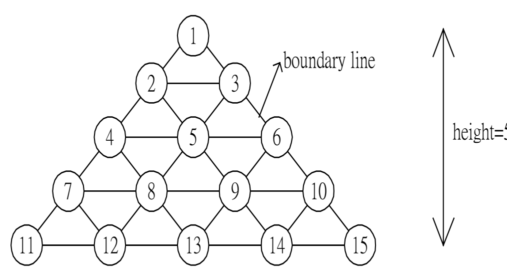


Figure2.2. a 15-node triangular grid

The advantages of triangular grid protocol are that it provides high availability, low quorum size, and uniform access quorums.

2.5 Circle Grid Protocol

In circle grid protocol, nodes are numbered from 0 to N-1. Let m be $\lceil \sqrt{N} \rceil$. It organizes the N nodes into a logical grid in the way that a node with number x is placed at the row numbered $\lfloor x/m \rfloor$ and the column numbered (x mod m). Let r be the number of rows in the logical structure,

$$r = \begin{cases} m-1, & \text{if } (m-1)^2+1 \leqq N \leqq m(m-1) \\ m, & \text{if } m(m-1)+1 \leqq N \leqq m^2 \end{cases}$$

Read quorum $R_{1,i}$, $R_{2,i}$, and write quorum $W_{i,a}$ are defined as follows:

$R_{1,i} = \{x|x=(i+j) \bmod N, 0 \leqq j<m \}$, for $0 \leqq i <N$.

$R_{2,i} = \{x|x=(i+jm) \bmod N, 0 \leqq j<r \}$, for $0 \leqq i<m$.

$W_{i,a} = \{x|x=(i+j) \bmod N, 0 \leqq j<m \} \cup \{x|x=(i+jm+a) \bmod N, j>0, (jm+a)<N \}$, for $0 \leqq i<N, 0 \leqq a<m$.

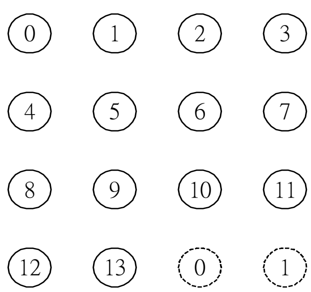


Figure 2.3. a 14-node circle grid

Figure 2.3 shows an example of 14 nodes. Some read quorums of this example are $R_{1,0} = \{0, 1, 2, 3\}$, $R_{1,3} = \{3, 4, 5, 6\}$, $R_{1,12} = \{12, 13, 0, 1\}$, $R_{2,1} = \{1, 5, 9, 13\}$, $R_{2,2} = \{2, 6, 10, 0\}$, $R_{2,3} = \{3, 7, 11, 1\}$, etc. Some write quorums of this example are $W_{0,0} = \{0, 1, 2, 3, 4, 8, 12\}$, $W_{0,1} = \{0, 1, 2, 3, 5, 9, 13\}$, $W_{6,3} = \{6, 7, 8, 9, 13, 3\}$, $W_{9,0} = \{9, 10, 11, 12, 13, 3, 7\}$, $W_{11,1} = \{11, 12, 13, 0, 2, 6, 10\}$, $W_{13,2} = \{13, 0, 1, 2, 5, 9\}$,

2.6 Read-write Difference Pair Protocol

The read-write difference pair protocol is based on the idea of cyclic block design and cyclic difference set in combinatorial theory. Let N be the number of nodes in the system. Some definitions and a theorem are given below.

[Definition 2.1] A cyclic group set G(A) under U, is cyclically generated from A, is a subset of U, as follow: $G(A) = \{Q_i \mid Q_i = \{ q \mid q = (a + i) \bmod N, \forall a \in A \}, i \in U\}$.

[Definition 2.2] The ordered pair (R, W) is called a read-write coterie under U if and only if the following two properties hold: (1) (R, W) is a bicoterie under U; (2) W is a coterie under U.

[Definition 2.3] A pair (C, D), where $C = \{c_0, c_1,\ldots, c_{n-1}\}$ and $D = \{d_0, d_1,\ldots, d_{p-1}\}$, is said to be a relaxed (N, n, p)-difference pair if for every $d \in U$, there exists at least one ordered pairs (c_i, d_j), where $c_i \in C$ and $d_j \in D$, such that $c_i - d_j \equiv d \pmod N$.

[Definition 2.4] The ordered pair (R, W) is called a read-write difference pair (rw-difference pair) under U if and only if both of the following properties hold: (1) (R, W) is a relaxed difference pair under U; (2) W is a relaxed difference set under U.

[Theorem 2.1] The ordered pair $(G(C_0), G(D_0))$ is a read-write coterie under U if and only if the ordered pair (C_0, D_0) is a read-write difference pair.

Let N = 15, $C_0 = \{0, 4, 8, 12\}$, and $D_0 = \{0, 1, 2, 3, 7\}$, then $G(C_0) = \{\{0, 4, 8, 12\}, \{1, 5, 9, 13\}, \{2, 6, 10, 14\}, \{3, 7, 11, 0\}, \{4, 8, 12, 1\}, \{5, 9, 13, 2\}, \{6, 10, 14, 3\}, \{7, 11, 0, 4\}, \{8, 12, 1, 5\}, \{9, 13, 2, 6\}, \{10, 14, 3, 7\}, \{11, 0, 4, 8\}, \{12, 1, 5, 9\}, \{13, 2, 6, 10\}, \{14, 3, 7,11\}\}$

$G(D_0) = \{\{0, 1, 2, 3, 7\}, \{1, 2, 3, 4, 8\}, \{2, 3, 4, 5, 9\}, \{3, 4, 5, 6, 10\}, \{4, 5, 6, 7, 11\}, \{5, 6, 7, 8, 12\}, \{6, 7, 8, 9, 13\}, \{7, 8, 9, 10, 14\}, \{8, 9,10, 11, 0\}, \{9, 10, 11, 12, 1\}, \{10, 11, 12, 13, 2\}, \{11, 12, 13, 14, 3\}, \{12, 13, 14, 0, 4\}, \{13, 14, 0, 1, 5\}, \{14, 0, 1, 2, 6\}\}$

The elements of $G(C_0)$ are read quorums and the elements of the $G(D_0)$ are write quorums. This protocol is symmetric and it is applicable to arbitrary number of nodes.

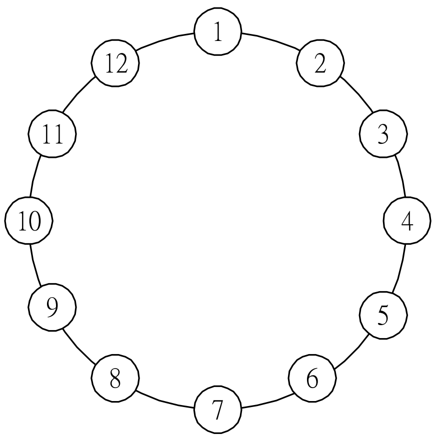
3. CYCLIC SPRING PROTOCOL

In this section, we propose a new protocol to get read and write quorum. This approach is originated from the circle grid protocol. In our protocol, the leading sub-sequence of an m-comet sequence is variable. It means that m, the number of consecutive nodes, is a variable. Therefore, the size of access quorum is not necessary the same. It is flexibility, like a spring.

In this section, we give the definitions, present an algorithm to get access quorums, and prove its correctness.

3.1 Definitions

This approach is based on a circular numbering system. A distributed system with N nodes is organized as a logical circle, denoted by *N*-ring. Every node in the system is assigned a distinct number from 1 to *N* and is arranged by its number sequentially. The term node is used to refer to the physical copy of replicated data in a distributed system. Some sequences of particular patterns are employed as read



and/or write quorums. The *N*-ring system with 12 nodes is shown in Figure 3.1.

Figure 3.1. a 12-ring system.

[Definition 3.1] The distance of an ordered pair of nodes v_i, v_j in an *N* -ring is denoted by dis(v_i, v_j), where

$$\mathrm{dis}(v_i, v_j) = \begin{cases} v_j - v_i\ , & \text{if } v_i \leqq v_j \\ v_j + N - v_i\ , & \text{if } v_I > v_j \end{cases}$$

In Figure 3.1, the distance of an ordered pair of nodes 1, 6 is denoted by dis(1, 6) = 5. The distance of another ordered pair of nodes 9, 4 is denoted by dis(9, 4) = 7.

[Definition 3.2] An ordered set of nodes S = {v_1, v_2,…, v_n} in an *N*-ring is said to be a circularly ascending sequence if for any two adjacent nodes of S, either $v_i < v_{i+1}$ or v_i and v_{i+1} are the largest and the smallest ones in S, respectively.

The sequence {8, 12, 1, 4} in Figure 3.1 is a circularly ascending sequence, but the sequence {1, 4, 8, 11, 2} is not.

[Definition 3.3] A circularly ascending sequence S = {v_1, v_2,…, v_n} in an *N*-ring is said to be an *m*-jump sequence (m $\geqq$1) if for any pair of adjacent nodes v_i, v_{i+1} of S, dis(v_i, v_{i+1})$\leqq$m.

The sequence {1, 4, 6, 7} in Figure 3.1 is a 3-jump sequence.

[Definition 3.4] An *m*-jump sequence S = {v_1,v_2,…,v_n} is said to be an *m*-jump circle if dis(v_n, v_1)$\leqq$m .

The sequence {1, 4, 7, 10} in Figure 3.1 is a 3-jump circle.

[Definition 3.5] A circularly ascending sequence C = {v_1, v_2,…, v_m, v_{m+1},…, v_n} in an *N*-ring is said to be an *m*-train sequence if {v_1,v_2,…, v_m} is a 1-jump sequence and {v_m, v_{m+1},…, v_n}is an *m*-jump sequence. {v_1,v_2,…,v_m } is called the leading sub-sequence, and {v_m, v_{m+1},…, v_n} is called the trailing sub-sequence of C.

The sequence {1, 2, 3, 4, 8} in Figure 3.1 is a 4-train sequence.

[Definition 3.6] The size of a *m*-train (any arbitrary m) sequence C = {v_1,v_2,…,v_n} is denoted by size(C), where size(C)= 1+dis(v_1, v_n)

The size of 4-train sequence C = {1, 2, 3, 4, 8} in Figure 3.1 is size(C) = 1 + dis(1, 8) = 8.

[Definition 3.7] An interval [v_i, v_j] in an *N*-ring is a set of nodes defined as the following:

$$[v_i, v_j] = \begin{cases} \{u \mid v_i \leqq u \leqq v_j\}, & \text{for } v_i \leqq v_j \\ \{u \mid u \leqq v_j \text{ or } u \geqq v_i\}, & \text{for } v_i > v_j \end{cases}$$

In Figure 3.1, interval [1, 5] = {1, 2, 3, 4, 5}, and interval [10, 4] = {10, 11, 12, 1, 2, 3, 4}.

[Definition 3.8] A pair of *m*-train sequences C_1 = {u_1, u_2,…, u_n}, C_2 = {v_1, v_2,…, v_p} are said to be scope overlapping if at least one node of C_1 is in the interval [v_1, v_p].

In Figure 3.1, the pair of 4-train sequence C_1 = {1, 2, 3, 4, 6, 8} and C_2 = {7, 8, 9, 10, 2} are scope overlapping .

[Definition 3.9] A write quorum W is the union of an arbitrary number of *m*-train sequences C_1, C_2 ,…, and C_n and satisfies that $\sum_i size(C_i) \geq \left\lceil \frac{N+1}{2} \right\rceil$. A read quorum R is an *m*-jump circle or a write quorum.

[Definition 3.10] An m-comet sequence C = {v_1, v_2, …, v_n} is said to be an m-comet circle if dis(v_n, v_1)$\leqq$m.

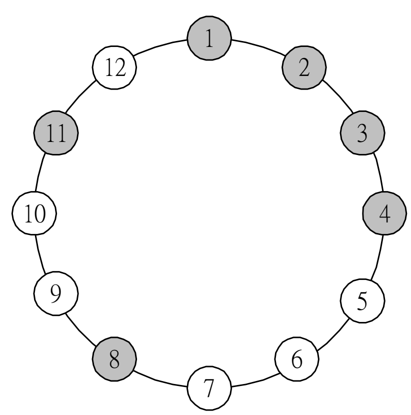


Figure 3.2. a 12-ring system.

For example, in Figure 3.2, the sequence {1, 2, 3, 4, 8, 11} is a 4-comet circle.

[Definition 3.11] An access quorum is defined as an m-comet circle and it is employed as read or write quorum.

3.2 Algorithms

In this section, we present an algorithm to get access quorums that are employed as read or write quorums. This protocol is shown in algorithm 3.1. Get_access_quorum (N): find an m-comet circle to be access quorum where m is not a pre-defined number.

[Algorithm 3.1] Find an access quorum.

Get_access_quorum (*N*){

In an N-rig, find the longest 1-jump sequence with p consecutive nodes where $p \leqq (N+1)/2$;

Find the pair of adjacent nodes v_i, v_j *with the largest distance* $dis(v_i, v_j)$ *and let* $max\_dis = dis(v_i, v_j)$;

If ($p \geqq$ max_dis) {

If (max_dis>$\lceil \sqrt{N} \rceil$) {

Let m = max_dis and find an m-comet circle C;

Return(C); /* C is the access quorum */

}

else if ($p \geqq \lceil \sqrt{N} \rceil$) {

Let m = $\lceil \sqrt{N} \rceil$ *and find an m-comet circle C*;

Return(C); /* C is the access quorum */

}

else {

Let m = p and find an m-comet circle C;

Return(C); /* C is the access quorum */

}

}

else *there is no proper access quorum*;

}

[Example 3.1] Consider the 24-ring system shown in Figure 3.3. Let m = 5. The failure nodes {9, 10, 11, 18, 21, 23, 24} are marked black. We know that the max_dis = 4 and p = 8. That is, the maximal number of consecutive failure nodes is 3 and the maximal number of consecutive active nodes is 8. Since $p \geqq$ max_dis and $p \geqq \lceil \sqrt{N} \rceil = 5$. Therefore, let m = 5, we can obtain the access quorum C = {1, 2, 3, 4, 5, 8, 13, 17, 22}.

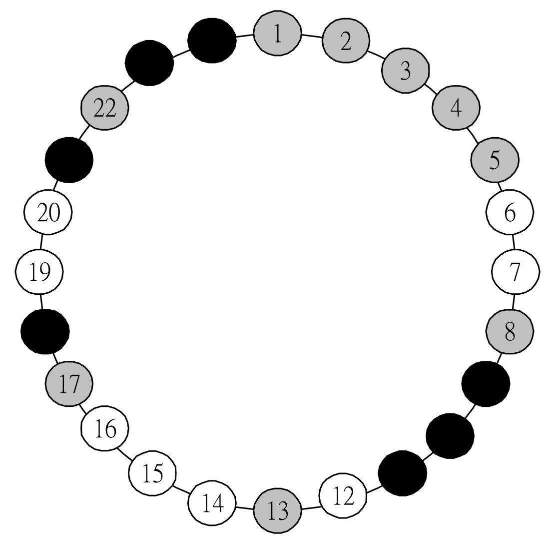


Figure 3.3. a24-ring system with failure nodes {9, 10, 11, 18, 21, 23, 24}.

[Example 3.2] In Figure 3.4, the failure nodes {10, 11, 12, 13, 14, 15, 24} are marked black. We know that the max_dis = 7 and p = 9. That is, the maximal number of consecutive failure nodes is 6 and the maximal number of consecutive active nodes is 9. Since $p \geqq$ max_dis and $p > \lceil \sqrt{N} \rceil = 5$. Therefore, let m = 5, we can obtain the access quorum C = {1, 2, 3, 4, 5, 8, 13, 17, 22}.

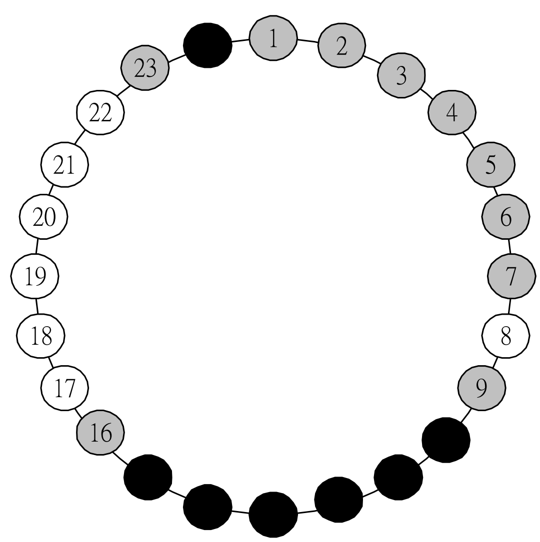


Figure 3.4. a 24-ring system with failure nodes {10, 11, 12, 13, 14, 15, 24}.

3.3 Proof of Correctness

In the following, we prove the intersection between two access quorums. It is used to prove the data consistency.

[Theorem 3.1]

In circular spring protocol, any two access quorums intersect.

Proof: Let $C_1=\{u_1, u_2, \ldots, u_p\}$, $C_2=\{v_1, v_2, \ldots, v_q\}$ be any two access quorums.The C_1 is an m-comet circle and C_2 is an n-comet circle.

(1) If $m \geqq n$

Since $dis(v_i, v_j) \leqq n$, where v_i, v_j be any two adjacent nodes in C_2. Therefore, v_i is at most n-1 nodes apart from v_j in C_2.We know that $u_1, u_2, \ldots, u_m$ in C_1 are m consecutive nodes. We can conclude that $v_j = u_k$, $1 \leqq k \leqq m$ or $v_i = u_k$, $1 \leqq k \leqq m$. i.e. C_1 intersects C_2.

(2) If m<n

This can be easily proved the other way around.

∴In circular spring protocol, any pair of access quorums intersect.

4. PERFORMANCE ANALYSIS

In this section, we first show that the proposed protocol is symmetric. We then analyze the access quorum sizes and the availability of this protocol. We also compare this protocol with some well-known quorum-based protocols.

4.1 Symmetric Property

A replica control protocol is said to be symmetric or fully distributed if every node in the system bears the same responsibility. Quorum-based protocols impose various logical structures in order to reduce their sizes of quorums. However, most of the protocols have unbalanced load sharing.

The proposed approach is based on the *N*-ring system. We arrange nodes into a logical circle and number them from 1 to *N* sequentially, where N is the number of nodes of the given system. Since circle is a symmetric structure, all nodes in a circle are identical except their identifiers, the assigned numbers. In other words, numbers assigned to these nodes are for identifying purpose only. Therefore, every node in an *N*-ring is included in the same number of quorums, and this approach is symmetric.

Our protocol is applied to arbitrary number of nodes. The nodes are arranged into an N-ring system, no matter what number the N is.

4.2 Quorum Size

In this section, we analyze the size of the access quorums. The size of quorums in these protocols is depended on the value of m. The m is the number of leading sub-sequence in an m-comet sequence or in an m-comet circle on the N-ring system. In Section 3, the m is variable and it is depended on the longest distance between two adjacent nodes.

In circular spring protocol, in the best case, the minimal access quorum size is equal to $\lceil \frac{N+1}{m} \rceil + m - 1$ for a given m. However, the m is variable, depended on the longest distance between two adjacent nodes. In the worst case, the maximal size of access quorum is equal to $\lceil \frac{N+1}{2} \rceil$. That is, the lower bound for the size of the access quorum generated with a given m is equal to $\lceil \frac{N+1}{m} \rceil + m - 1$. The upper bound for the size of the access quorum is equal to $\lceil \frac{N+1}{2} \rceil$.

In the worst case, the maximal sizes of read quorum or write quorum may be larger than most quorum-based techniques. However, in that situation, most protocols may not find out any one quorum. That is, the system does not work properly.

4.3 Availability

In this section, we analyze the availability of our presented protocol in Section 3. Assume that each node is available with identical probability p and the state of each node is independent and will not change when an operation is in progress. The availability of a replica control protocol is defined to be the probability that at least one quorum can be constructed.

In Section 3, we proposed circular spring protocol. Here, we analyze its availability and compare with circle grid protocol. In an N nodes system, Circle grid protocol defined two type of read quorum. In the circle structure, the read quorum is that each node is less than m nodes apart from another one. The write quorum is that consists of m consecutive nodes and followed nodes such that each node is less than m nodes apart from another one. The m is equal to $\lceil \sqrt{N} \rceil$ and is fixed.

The write and read quorum in circular spring protocol is defined as access quorum. The access quorum is an m-jump circle and m is variable. The read quorum and write quorum in circle grid is included in m-comet circle. Therefore, the availability of our protocol is better than circle grid protocol.

Finally, we analyze the availability of circular spring protocol. We calculate the availability by simulating the conditions where all nodes are with the same probability of their existence. Let the probability that a node is available for service be p. Assume that the possibility is uniform among the nodes of this system. The availabilities of circular spring protocol calculated from this simulation are shown in Table 4.1 for the cases that the probabilities of each node to be available are $p = 0.95$ and $p = 0.9$. According to the result, this approach shows satisfactory performance. Moreover, a system in this protocol becomes more reliable as we add more nodes into it.

Table 4.1. The availability of a system

Number of nodes	the probability of a available node is	
	$p = 0.95$	$p = 0.9$
4	0.99046	0.96398
5	0.99182	0.97072
6	0.99828	0.96990
7	0.99834	0.99004
8	0.99944	0.99394
9	0.99978	0.99610
10	0.99984	0.99732
11	0.99990	0.99814
12	0.99994	0.99878
13	0.99998	0.99916
14	1.00000	0.99930
15	1.00000	0.99958
16	0.99998	0.99960
17	0.99996	0.99974
18	1.00000	0.99968
19	1.00000	0.99976
20	1.00000	0.99988
21	1.00000	0.99996
22	1.00000	0.99990
23	1.00000	0.99994
24	1.00000	0.99996
25	1.00000	0.99998
26	1.00000	0.99998
27	1.00000	0.99998
28	1.00000	1.00000

5. CONCLUSIONS

In this paper, we proposed a new quorum-based approach, cyclic spring protocol, for replica control. In cyclic spring protocol, multiple copies of replicated data are organized into a logical circle structure, N-ring. This protocol has some advantages: (1) It is applicable to systems with any arbitrary numbers of nodes. (2) Every node in the system bears the same responsibility to read and write operations. In other words, it is symmetric and load balanced. (3) The availability of cyclic spring protocol is better than that of most quorum-based protocols.

6. REFERENCES

[1] D. Agrawal and A. EI Abbadi, "The Tree Quorum Protocol: An efficient approach for managing replicated data", in Proc. of Very Large Database Conf., pp. 243~254, 1990.

[2] D. Agrawal and A. EI Abbadi, "The Generalized Tree Quorum Protocol: An efficient approach for managing replicated data", *ACM Trans. Database syst.17(4)*, pp. 689~717, 1992.

[3] S.Y. Cheung, M.H. Ammar, and M. Ahamad, "The Grid protocol: a high performance scheme for maintaining replicated data," in Proceedings of the IEEE 6th Int. Conf. on Data Engineering, pp. 438~445, 1990.

[4] A. Kumar, M. Robinovich, and R. Sinha, "A Performance Study of a New Grid protocol and General Grid Structures for Replicated Data," Tech. Report 93-03-02, Department of Computer Science and Engineering, University of Washington, March 1993.

[5] C. Wu and G. G. Belford, "The Triangular Lattice Protocol: A Highly Fault Tolerant and Highly Efficient Protocol for Replicated Data", in IEEE 11th Proc. Symp. On Reliable Distributed System, pp. 66~73, 1992.

[6] Cheng-Hong Cho and Jer-Tsang Wang, "Triangular Grid protocol: An Efficient Scheme for Replica Control with Uniform Access Quorums," *Theoretical Computer Science, 196*, pp. 259~288, Aug. 1998.

[7] Ching-Ming Lin, Ge-Ming Chiu, and Cheng-Hong Cho, "An Efficient Quorum-Based Scheme for Managing Replicated Data in Distributed Systems", in Int'l Conf. on Parallel Processing, pp.328~335, 1999.

[8] R.H. Thomas, "A Majority Consensus Approach to Concurrency Control for Multiple Copy Database". *ACM Transactions on Database System*, 4(2): 180~209, Jun. 1979.

[9] D.K. Gifford, "Weighted voting for replicated data," in Proc. of the 7th ACM Symposium on Operating System Principles, pp. 150~159, 1979.

[10] Y. T. Wu, Y. J. Chang, S. H. Yuan, and H. K. Chang, "A New Quorum-Based Replica Control Protocol," in Proceedings. Pacific Rim International Symposium on Fault-tolerant System, pp. 116~121, Dec. 1997.