ABSTRACT
This paper introduces a new Replica control protocol named “Diagonal Replication with Intersection of Quorums in 2D Mesh Protocol (IQ2DM)” for grid environment and compares its performance with the previous protocols. The performance in this paper is data availability for read and writes operation, which are compared to the Read-One Write-All (ROWA), Voting (VT), Grid Configuration (GC), Three Dimensional Grid Structure (TDGS), and Diagonal Replication on Grid (DRG). This paper discusses the protocol of replicating data for grid environment, putting the protocol in a logical 2D mesh structure by employing the quorums and voting techniques. The data file is copied in a selected replica from the diagonal sites in each Quorum. The selection of a replica depends on the diagonal location of the structured 2D mesh network where the middle replica is selected because it is the shortest path to get a copy of the data from most of the direction in the quorum. The algorithm in this paper also calculates the optimized number of nodes to be grouped in each quorum and how many quorums are needed for the number of nodes, \( N \) in a network. IQ2DM protocol also ensures that the data for read and write operations are consistent, by ensuring the quorum must not have a nonempty intersection quorum. To evaluate the IQ2DM protocol, we developed a simulation model in Java.

**Keywords:** Data replication, Grid, Data management, Availability, 2D Mesh protocol

I. INTRODUCTION
A grid is a distributed network computing system, a virtual computer formed by a networked set of heterogeneous machines that agree to share their local resources with each other. A grid is a very large scale, generalized distributed network computing system that can scale to internet size environment with machines distributed across multiple organizations and administrative domains [1][2]. The availability of a data in a large network and replicating data at a minimum communication cost are also some of the issues [3][4][5][6]. EU data grid and HEP in CERN used “Reptor” as the prototype to manage replica in grid[7]. We investigate on the replica selection for optimizing and improving data accessing by using replica control protocol in distributed database to the grid environment. It is concerned to efficient utilization of a pool of heterogeneous systems with optimal workload management utilizing an enterprise's entire computational resources (servers, networks, storage, and information) acting together to create one or more large pools of computing resources. There is no limitation of users or originations in grid computing.

Quorums improved the performance of fault tolerant and availability of data [8][9]. Quorums reduce the number of copies for reading or writing data. To implement quorum, a protocol must satisfy two constraints which are total of quorum for read, \( qr \) and write quorum, \( qw \) must be larger than the total number of votes, \( v \) assigned to the copies of the data object and the quorum for write, \( qw \) is larger than \( v/2 \) [10]. To address the availability, IQ2DM replicates data on the middle node of a quorum of read or write in the logical structured of 2D mesh topology network. Java is used to run this replication protocol.

The paper is organized as follows: In Section 2, EDR2M protocol and its algorithm are introduced. In Section 3, we present the previous replica control
protocols. This section includes the formulation for read/write availability for the previous protocols in distributed database and grid computing. Section 4 describes the modification in the structure of EDR2M protocol by reducing the number of nodes and the number of quorums based on the algorithm defined in section 1 and the proposed protocol is called as IQ2DM and Section 5 discusses the conclusion of IQ2DM protocol.

Figure 1. The main component in Replica Management System

II. ENHANCED DIAGONAL REPLICA TION IN 2D MESH (EDR2M) PROTOCOL

In EDR2M protocol [22] [23], all nodes are logically organized into two dimensional Mesh structure. The data are replicated to only one node of the diagonal site which is the middle node of the diagonal site in each quorum. All replicas are operational meaning that the copies at all replicas are always available. This protocol uses quorum to arrange nodes in cluster. Quorum is grouping the nodes or databases into small cluster to manage the replica for read or write operations. Figure 5 illustrates how the quorums for network size of 25 nodes are grouped by nodes of 3 x 3 in each quorum. Nodes which are formed in a quorum intersect with other quorums. This is to ensure that each quorum can communicate or read other data from other nodes which is in another quorum. The number of nodes grouped in quorum, R must be odd so that only one middle node from the diagonal sites can be selected such as the black circle in Figure 3, which reduces the communication cost. Example, s (3, 2) in Figure 3 is selected to have the copy of data. It is formed by identifying number of quorums needed in the network. Each quorum must intersect with each other. The algorithm [23] of the model is as in Fig. 2 it illustrates how the algorithm is designed and implemented.

Main
Input number of row or column, N
n = next odd number after \( \sqrt{N} \)
If n is an odd integer then
Find the number of quorum, Q
\[ q = \left\lfloor \frac{\sqrt{n} - n/10} \right\rfloor \]
Q = q^2
Find the number of nodes X in each quorum,
\[ X = \frac{n}{q} \]
Get the next odd integer after X
Select the middle replica, R
Copy the data file at R
Else add one virtual column and row, Col new + Row New
N = Col new * Row New then
Return N to Main

Figure 2. Algorithm of the data replication in EDR2M protocol

III. RELATED WORK

There are few protocols to replicate a data in distributed database and grid computing as discussed in the following subsections, and we compare our enhanced diagonal structure of 2D mesh protocol with the other protocols corresponding to its availability.

1. Read-One Write-All (ROWA) Protocol

ROWA is a simple and straightforward protocol [11]. It requires all copies of all logical data items that are updated by a transaction be accessible for the
transaction to terminate. Failure of one site may block a transaction and in ROWA, a read operation needs only one copy, while a write operation needs to access the \( n \) number of copies. Therefore, the availability for a read operation can be represented as one out of \( n \) and for a write operation as \( n \) out of \( n \) [12]. Thus, the formulation for read availability for ROWA, \( A(\text{ROWA}, R) \) is as given in Eq. (1)

\[
A(\text{ROWA}, R) = \sum_{i=1}^{n} \binom{n}{i} p^i (1-p)^{n-i}
\]  

(1)

And the formulation for write availability \( A(\text{ROWA}, W) \) is as given in Eq. (2)

\[
A(\text{ROWA}, W) = \sum_{i=1}^{n} \binom{n}{i} p^i (1-p)^{n-i}
\]  

(2)

Where \( p \) is the probability that a copy is accessible and \( p \) is from 0.1 to 0.9. ROWA reduces the availability of the database in case of failure since the transaction may not complete unless it reflects the effects of the write operation on all copies. Therefore, there have been a number of algorithms that have attempted to maintain mutual consistency without employing the ROWA protocol [13].

2. **Voting (VT) Protocol**

In VT approach, every copy of replicated data object is assigned to a certain number of votes and a transaction has to collect a read quorum of \( r \) votes to read a data object, and a write quorum of \( w \) votes to write the data object. Quorum must satisfy two constraints which are \( r + w \) must be larger than the total number of votes, \( v \) assigned to the copies of the data object and \( w > \sqrt{v/2} \), where the total of write quorum of \( w \) votes must be larger than half of the total number of votes, \( v \).

To avoid the read availability becomes expensive, the read quorum \( k \) is selected where \( k \) is smaller than the majority quorum.

[10] The formulation for read availability in VT, AVT, \( R \) is as given in Eq. (3), where \( n \) is the total number of nodes that has the votes or sometime it is called replica.

\[
A(\text{VT}, R) = \sum_{i=1}^{n} \binom{n}{i} p^i (1-p)^{n-i}
\]  

(3)

And the corresponding formulation for write availability in VT, AVT, \( W \) is as in Eq. (4), where \( k \) is the number of votes for read or write quorum.

\[
A(\text{VT}, W) = \sum_{i=1}^{n} \binom{n}{i} p^i (1-p)^{n-i}
\]  

(4)

This protocol is popular and easy to implement but writing an object is fairly expensive[10] where write quorum \( w \) of copies must be larger than the majority votes \( v \), \( w > \sqrt{v/2} \).

3. **Grid Configuration (GC) Protocol**

GC protocol [14] has \( n \) copies of data objects that are logically organized in the form of \( n \times n \) grid as shown in Figure. In Figure 5, the number of nodes is 25 in \( 5 \times 5 \) grid network where \( n \) are five. The figure shows three grey circles, which represent nodes that are down or not active. The nodes which are downed can be placed logically anywhere in the grid structure. Read operations on the data item are executed by acquiring a read quorum that consists of data copy from each column in the grid, while write operations are executed by acquiring a write quorum that consists of all copies in one column and a copy from each remaining column.

The formulation of read availability in the GC protocol [15], AGC, \( R \) is as given in Eq. (5)

\[
A(\text{GC}, R) = \sum_{i=1}^{n} \binom{n^{1/2}}{i} p^i (1-p)^{n^{1/2}-i} \sqrt{n}
\]  

(5)

Where \( p \) is the probability that a copy is accessible and \( p \) is from 0.1 to 0.9. While, the formulation of write availability in the GC, \( A_{GC,W} \) is as given in Eq.(6)

\[
A_{GC,W} = [1-(1-p)^{n}]^{p} -[1-(1-p)^{n}] -p^{n}
\]  

(6)

This protocol requires a bigger number of read and write quorum for the read operations to be executed. Read quorum must be at every column and for write operations to be executed, write quorum must exist at one of the entire column and exist at least once at other columns. Thus, this decreases the data availability. It is also vulnerable to the failure of entire column or row in the grid.

4. **Three Dimension Grid Structure (TDGS) Protocol**

TDGS protocol replicated its data in logical box shape structure with four planes [10] [16]. Figure 3, illustrates eight copies of data object. The read operations in TDGS are executed by acquiring a read quorum that consists of any hypotenuse copies.
For the example shown in Figure 4, hypotenuse copies are \{A, H\}, \{B, G\}, \{C, F\}, and \{D, E\}. Read operations are executable by these pairs of hypotenuse copies. Write operations are executable from any planes that consist of hypotenuse copy. Planes in Figure 3, are \{H, A, B, C, D\}, \{C, F, E, G, H\}, and etc. Example, to execute read operation, copies from \{A, H\} must be accessible and to execute write operation, copies from \{H, A, B, C, D\} must also be accessible. In TDGS protocol, read quorum can be constructed from four hypotenuse copies. Formulation of read availability \[A_{TDGS,R}\] is as given in Eq.(7) where \(p\) is the probability that a copy is accessible and \(p\) is from 0.1 to 0.9.

\[A_{TDGS,R} = (1- p^2)^4\] (7)

Whereas, formulation of write availability, \(A_{TDGS,W}\) is as given in Eq.(8)

\[A_{TDGS,W} = 1-(1-\beta)\] (8)

Where \(\beta = \varphi + \varphi - p^2 (\varphi \times \varphi)\)

\[\varphi = p^4 (1+p-p^2)\quad \text{and } \varphi = p^4 (2-p^2)\]

Read and write quorum must intersect otherwise the hypotenuse of read quorum is not accessible and the write quorum is also not accessible to update the latest data. This affects the consistency of the data.

5. Diagonal Replication on Grid (DRG) Protocol

Diagonal Replication on Grid (DRG) is a protocol [9] which is logically organized in a two dimensional grid structure. For example 4, if a DRG consists of twenty-five sites, the network is logically formed into 5 x5 grids as shown in Figure 4. Each site has a master data file. A site is either operational or failed and the state (operational or failed) of each site is statistically independent to the others. When a site is operational, the copy at the site is available; otherwise it is unavailable. Sites can be down or not active. In Figure 6, sites 23, 24, and 25 are not active or fail. Sets of diagonal sites in the grid are selected such as set \{1, 7, 13, 19, and 25\}. After the diagonal sites are identified, the primary copy of the data is placed on the replica which is distributed diagonally [9]. For example in figure 6, diagonal set \(D^2(s)\) is \{s(2), s(8), s(14), s(20), s(21)\} and \(D^3(s)\) is \{s(3), s(9), s(15), s(16), s(22)\}.
needed in $N$ size of network where $n$ is the number of row or columns. The total number of quorums, $qr$ for the whole network, $N$ is $q2$ where $q$ is the number of quorum in a row or columns as shown in Eq. (9). This is to ensure that each quorum can communicate or read other data from other nodes which is in another quorum. We create a structure of IQ2DM, in fig.6 n is 5, for 3 x 3 nodes of grid network size and $p$ is the probability of data available which is between 0 to 1, qr and qw are the number of quorums for read and write operations respectively. Thus, the formulation for read availability for IQ2DM, A (IQ2DM), is as given in Eq. (10).

$$q = \sqrt{n-n/10}$$

$$qr= q2^{0.5}$$  \hspace{1cm} (9)

To calculate the data availability [22], all copies are assumed to have the same probability $p$ of availability. Refer Eq. (10).

$$A \text{ (IQ2DM)} = \sum_{i=qr}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$  \hspace{1cm} (10)

V. CONCLUSION

In this paper, a new protocol, called Diagonal Replication with intersection of quorums in 2D Mesh (IQ2DM) protocol for grid environment has been proposed to manage the data replication in a large network size such as in distributed system and especially in grid environment. IQ2DM protocol, selects one replica in a diagonal site of a quorum in a 2D mesh logical structure. The number of nodes in each quorum is odd so it is easy to select only one node from the diagonal site in each quorum. The IQ2DM protocol provides a convenient approach for write operations. This is due to the minimum number of quorum size required. DR2M has overcome the read and write availability issues of TDGS which is the latest replica control protocol. IQ2DM protocol is the best protocol for replication control protocol as it provides the highest write availability among the protocols discussed in the previous sections and also reduces the communication cost.

VI. REFERENCES


22. Rohaya Latip, Mohamed Othman, Azizol Abdullah, Department of Technology Communication and Network, Faculty of Computer Science and Information Technology, University Putra Malaysia, “Quorum-based Data Replication in Grid Environment”, International Journal of Computational Intelligence Systems, Vol.2, No. 4 (December, 2009), 386-397.

23. Rohaya Latip1, Tazquia Mizan College of Computer Science Najran University Najran, KSA, Nur Faezah Ghazali, Rabiah Abd Kadir Faculty of Computer Science and Information Technology, University Putra Malaysia Serdang, Malaysia, Feras Ahmad Hanandeh, Information Technology, Hashemite University, Jordon. (2013). Replica Control Protocol: Triple Quorum Replication (TQR) in Data Grid, 5th International Conference on Computer Science and Information Technology (CSIT).