Economic Model for Consistency Management of Replicas in Data Grids with OptorSim Simulator

Ghalem Belalem

Dept. of Computer Science, Faculty of Sciences,
University of Oran – Es Senia, Oran, Algeria
Ghalem1dz@gmail.com

Abstract. Data Grids are currently solutions suggested meeting the needs for the scale large systems. They provide whole of resources varied, geographically distributed of which the goal is to ensure a fast access and effective to the data, to improve the availability, and to tolerate the breakdowns. In such systems, these advantages are not possible that by the use of the replication. Nevertheless, several problems can appear with the use of replication techniques, most significant is the maintaining consistency of modified data. The strategies of replication and scheduling of the jobs were tested by simulation. Several simulators of grids were born. One of the important simulators for Data Grids is the OptorSim tool. In this work, we present an extension of OptorSim by consistency management service of replicas in data Grid. This service is based on hybrid approach, combining between pessimistic, optimistic approaches and economic model, articulated on a hierarchical model.

Keywords: Data Grid, Replication, Consistency, OptorSim, Optimistic approach, Pessimistic approach, Economics models.

1 Introduction

Replication in the Grids [3] is one of the popular research topics and is a widely accepted way to improve data availability and fault tolerance in the Grids. Broadly speaking, we can say that the technique of replication raises several problems [4, 5, 7], such as: *(i)* Degree of replicas; *(ii)* Choosing the replicas; *(iii)* Placement of replicas; *(iv)* Service of consistency management of replicas. Two approaches are generally used for the consistency management: optimistic approach and pessimistic approach [6]. In this present paper, we propose service based on a hybrid approach, combining between the two approaches of coherence (pessimist and optimist) and the economic model of the market. The approach proposed rests on a hierarchical model with two layers, it aims principally at maintaining the consistency of large scale systems. The remaining part of this article is organized as follows: in Section 2, we present the fundamental principles of the consistency management approaches. Section 3, describes the hybrid approach and its model proposed in terms of structure, functionalities, features and principal algorithm of our approach based on economic model for consistency management of the

P. Primet et al. (Eds.): GridNets 2008, LNICST 2, pp. 121 – 129, 2009.

replicas. The various preliminary experiments discussed in section 4. Finally, we finish this article by a conclusion and some future work perspectives.

2 Approaches of Consistency Management

The replica of data is made up of multiple copies, on the separate computers. It is a significant technology which improves the availability and the execution. But the various replica of the same object must be coherent, i.e. seem only one copy. There are many models of consistency. All does not offer the same performances and do not impose the same constraints to the application programmers. The consistency management of replicas can be done either in a synchronous way with use of algorithms known as pessimistic, or in an asynchronous way with use of algorithms known as optimistic [6]. The synchronous model makes it possible to make an immediate propagation of the updates towards the replicas with blocking of the access, whereas the asynchronous model differs the operation from propagation of the updates, which introduces a certain level of divergence between replicas.

1. *Pessimistic approach:* This forbids any access to a replica provided that it is up to data [1, 6], what gives an illusion to the user to have a single substantial copy. The main advantage of this approach is that all replicas converge at the same time, what allows guaranteeing a strong consistency, so avoiding any problem linked to the stage of the reconciliation. This type of approach is adapted well to systems of small and average scales and it becomes very complex to implement for systems with wide scale. So, we can raise three major inconveniences of this type of approach: *(i)* It is very badly adapted to vague or unstable environments, such as the mobile systems or the grids at strong rate of change; *(ii)* It cannot support the cost of update when the degree of replication is very high; *(iii)* It cannot support the cost of update when the degree of replication is very high.
2. *Optimistic approach:* This approach authorizes the access to any replica and all the time. In this way, it is then possible to access a replica which is not inevitably coherent [1, 6]. So, this approach tolerates a certain difference between replicas. On the other hand, it requires a phase of detection of difference between replicas then a phase of correction of this difference by converging the replicas on a coherent state. Although it does not guarantee coherence strong as in the pessimistic case, it possesses nevertheless certain number of advantages which we can summarize as follows [1]: *(i)* They improve the availability: because the access to data is never blocked; *(ii)* They are flexible as regards the network management which does not need to be completely connected so that the replicas are completely accessible, like the mobile environments; *(iii)* They can support large number replicas by the fact that they require not enough synchronization between replicas; *(iv)* Its algorithms are well adapted to large scale systems.

3 Service of Consistency Management Based on Economic Models

To improve the quality of services (QoS) in consistency management of replicas in data grids, we deem valuable for extending work presented in [1] by an economic model of the market when resolving conflicts between replicas. In the real world market, there exist various economic models for setting the price of services based on supply-and-demand and their value to users. These real world economy models such as commodity market model, market bidding, auction model, bargaining model etc can also be applied to allocate resources and tasks in distributed computing [2]. In our work, we regard a grid as a distributed collection of elements of calculation (CE's) and elements of storage (SE's). These elements are dependent whole by a network to form a Site. This hierarchical model is composed of two levels (see Figure 1), level 0 is charged to control and to manage local consistency within each site, level 1 consists in ensuring the global consistency of the grid.

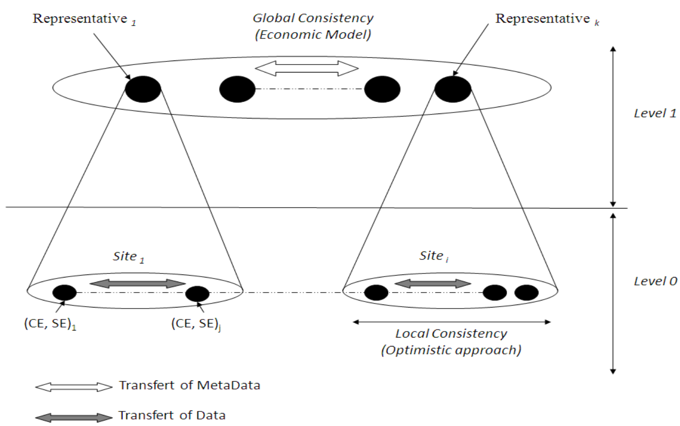

Fig. 1. Proposed Model

The main stages of global process of consistency management of replicas proposed is defined by the following Figure 2.

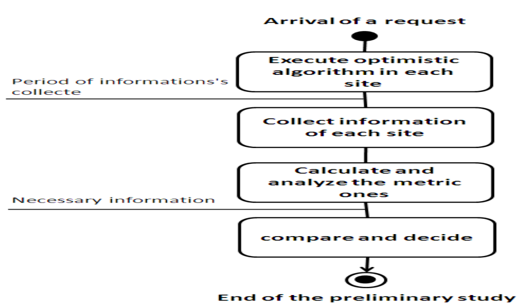

Fig. 2. Principal stages of consistency management

- Stage 1: Collect of information (collection): example of information, we have the number of conflicts, divergences per site as well as current time.
- Stage 2: Calculate metrics (analyzes): consist in analyzing and comparing the information collected with the tolerated thresholds (threshold of conflicts and threshold of divergence,…).
- Stage 3: Decision: decide the adequate algorithm of consistency management of replicas.

To study collaboration between the various entities of service of consistency management, we used the diagram of activity as we can see it in this Figure 3.

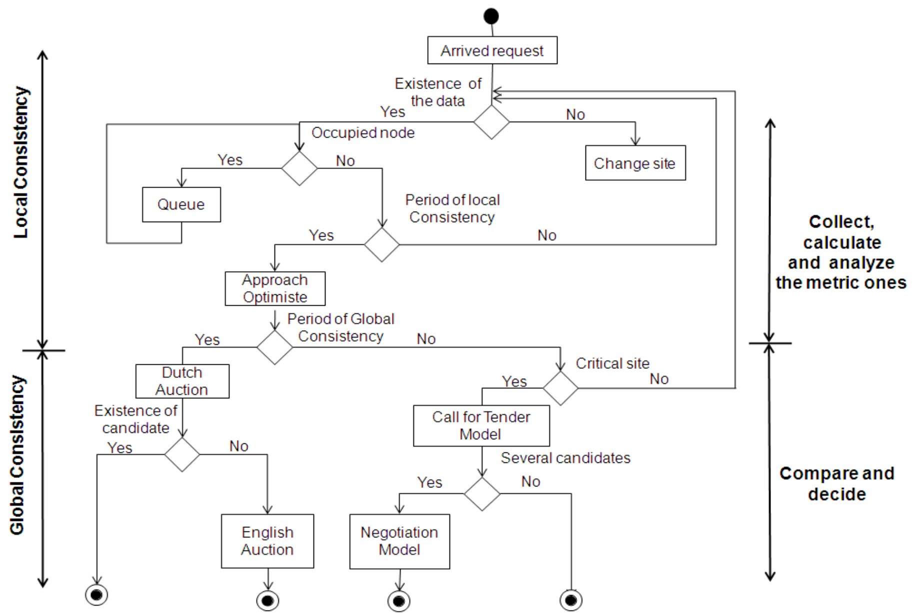

Fig. 3. Diagram of activity of service consistency management

3.1 Local Consistency

Local consistency also called coherence intra-site, in this level we find the sites which make the grid. Each site contains CE's and SE's. The retorted data are stored on the SE's and are accessible by CE's via the operations from reading or writing. Each replica attached to additional information is called metadata (Timestamp, of the indices, the versions,…). The control and the management of coherence are ensured by representatives elected by a mechanism of election, the goal of these representatives is to treat the requests of readings and writings coming from the external environment, and to also control the degree of divergence within the site, by calculating measurements appearing in Table 1: Rate of the number of conflicts per site (τ_i), distance within a site (D_{local_i}), dispersion of version (σ_i).

Table 1. Measurements calculated in local consistency

Measure	Definition
n_i	Number of replicas of same object inside $Site_i$
D_{local_i}	$Version_{max}$ - $Version_{min}$ of replicas inside $Site_i$
τ_i	Rate of number of conflicts inside $Site_i$
σ_i	Standard deviation of $Site_i$ = $\sqrt{1/n_i \sum_{t=1}^{n_i} (V_{it} - \overline{V}_i)^2}$
	Where $\overline{V}_i$ is average version inside $Site_i$

We detect critical situations of one site to meet of one of the following cases:

- τ_i > Rate of Conflicts number tolerated;
- D_{local_i} > Distance tolerated;
- $\sigma_i > \sigma_m$ tolerance rate for dispersion of versions.

3.2 Global Consistency

This level is also called coherence of inter-sites, it is responsible for global consistency of data grid, each site cooperates with the other sites via the representatives who communicate between them. The goal of representatives is to control the degree of divergence and of conflict within the site, by calculating measurements appearing in Table 2.

Table 2. Measurements calculated in global consistency

Measure	Definition
$D_{global}(i,j)$	If $n_i < n_j$ then $\sqrt{1/n_i \sum_{t=1}^{n_i} \lvert V_{it} - V_{jt} \rvert^2}$ else $\sqrt{1/n_j \sum_{t=1}^{n_j} \lvert V_{it} - V_{jt} \rvert^2}$
τ_{ij}	$(\tau_i + \tau_j)/(n_i + n_j)$: Rate of conflicts between $Site_i$ and $Site_j$
Cov(i,j)	If $n_i < n_j$ then $\frac{1}{n_i}\sum_{t=1}^{n_i} (V_{it} - \overline{V}_i)(V_{jt} - \overline{V}_j)$ else $\frac{1}{n_j}\sum_{t=1}^{n_j} (V_{it} - \overline{V}_i)(V_{jt} - \overline{V}_j)$
ρ_{ij}	$Cov(i,j)/ \sigma_i \sigma_j$: Coefficient of correlation between $Site_i$ and $Site_j$

Two situations can be treated, the first situation corresponds to the competing negotiation, started after the presence of a critical situation of a site which is translated by the existence of:

- τ_{ij} > Threshold of a tolerated number of conflicts;
- $D_{global}(i,j)$ > Threshold of tolerated divergence;
- $|\rho_{ij}| < \varepsilon$ Where $\varepsilon \ll 1$.

This negotiation proceeds between two groups (group of stable representatives who are not in critical situation and a group of representatives in crisis), that is to say by a call for tender, where a representative in crisis receives offers of the other stable representatives by providing described measurements above. This representative seeks the representative who has measurements close with his; algorithm below (algorithm 1) described the mechanism to this negotiation:

Algorithm 1. CALL FOR TENDER

Calculate τ_i , D_{local_i}, σ_i /* measurements of *i*th representative */
candidates $\leftarrow$ false, Nbr_candidates $\leftarrow$ 0
for all elements of the group of sites in crisis **do**
Representative$_a$ $\leftarrow$ representative of site in crisis; j $\leftarrow$ 1
repeat
  Representative$_b$ $\leftarrow$ representative of the stable site
    if ($|\tau_a - \tau_b| < \varepsilon$) $\vee$ ($|D_{local\_a} - D_{local\_b}| < \varepsilon$) $\vee$ ($|\sigma_a > \sigma_b| < \varepsilon$) **then**
      Nbr_candidates $\leftarrow$ Nbr_candidates+1
    end if
    j $\leftarrow$ j +1
until j $\geq$ Number of elements of the group of stable sites
if (Nbr_candidates > 1) **then**
  candidates $\leftarrow$ true; Algorithm Negotiation /* Negotiation of candidates */
else
  Propagation of the updates of the stable site to the sites in crisis
end if
end for

If there are two or more representatives candidates (*candidates = true*), being able to put the site in crisis in stable state one passes to a negotiation between these candidates (see Algorithm 1), the best of them is that which is more stable, and will proceed to update the site in crisis. The second situation corresponds to the co-operative negotiation (see Algorithm 2) between the representatives of each site, started by the exhaustion of the period defined for total coherence.

Algorithm 2. NEGOTIATION

τ_i , D_{local_i}, σ_i /* measurements of *i*th representative for stable site */
winner $\leftarrow$ first of all sites stable /* Candidate supposed winner */
τ_i , D_{local_i}, σ_i /*measurements of ith representative*/ ; no_candidate $\leftarrow$ false
for all representatives – { candidate$_{winner}$ } **do**
  if ($\tau_i < \tau_{winner}$) $\wedge$ ($D_i < D_{winner}$) $\wedge$ ($\sigma_i < \sigma_{winner}$) **then**
    winner $\leftarrow$ reprensentative$_i$
    end if
  Winner propagates its updates with the sites in crisis
end if

For that, two algorithms inspired of the economic model can be used to put the whole of the replicas in a coherent state.

The Dutch auction (Algorithm 3) ensures the quality of service of which the aim is to obtain the most recent replica, it consists in seeking the representative having the largest version and to compare its measurements with those of the other representatives for goal to have a representative having less conflicts and divergences, and to compare the version of this representative with the average of all vectors of version. If its version is lower or equal "*no_candidate = true*" (see Algorithm 3), then we proceed to the Algorithm of English Auction.

Algorithm 3. DUTCH AUCTION

```
representative_max ← representative having the most recent version
version_reserves  /* the average of all vectors of versions */
τ_i , D_local_i, σ_i  /* measurements of i^th representative */
no_candidate ← false
for all representatives – {representative_max} do
    if (τ_max > τ_i) ∨ (D_local_max > D_local_i) ∨ (σ_max > σ_i) then
        if version_i ≤ version_reserves then
            no_candidate ← true
        else
            representative_max ← representative_i ; no_candidate ← false
        end if
    end if
end for
if (no_candidates = false ) then
representative_max propagates the updates with the whole of representatives
else
    Algorithm English Auction
end if
```

The English auction (Algorithm 4) has the same principle that the Dutch auction except that it is ascending starting with the representative having the oldest replica and increases according to measurements calculated within each site.

Algorithm 4. ENGLISH AUCTION

```
representative_min ← representative having the oldest version
version_reserves  /* the average of all vectors of versions */
τ_i , D_local_i, σ_i  /* measurements of i^th representative */
for all representatives – {representative_min} do
    if (τ_min > τ_i) ∨ (D_local_min > D_local_i) ∨ (σ_min > σ_i) then
            representative_min ← representative_i
    end if
end for
representative_min propagates the updates with the whole of the representatives
```

4 Experimental Study

We have implemented a simulation tool for our approach and their component, where entire architecture was discussed in the previous sections. The tool is based on the OptorSim Grid simulator [1], extended with our hybrid approach, in order to validate and to evaluate our approach of consistency management of replicas compared to pessimistic and optimistic approaches, we carried out series of experiments whose results and interpretations are covered in this section. In order to analyze the results relating to the experimentation of our approach, we used three metrics to know the response time, the number of divergences and number conflicts among replicas. These experiments were carried out with the following parameters of simulation: 10 sites, 100 nodes, retorted data 10 times in each site according to the strategy multi-Masters, the bandwidth being fixed at 100Mb/ms for star topology, and we varied the number of requests (requests of writings) between 100 to 500 requests per step of 100.

Fig. 4. Average response time / #Requests

Fig. 5. Average response time / #Requests

We notice that in graph of the Figure 4, some is it a number of requests carried out, response time of the two approaches pessimists: ROWA and Quorum, are obviously more than optimistic approach, hybrid [1] like our approach. Us let us notice in the Figure 5, that the optimistic approach which does not block the requests, it who causes to reduce the execution time contrary to the two other approaches which are slightly confused, this increase is explained owing to the fact that local and global consistency, were started to propagate the updates in the two approaches. From this Figure and although two curves are not very different, the curve of our approach remains in below of that of the hybrid approach. This experiment allows us to evaluate and compare the QoS, by taking account of the divergences and the conflicts a number of conflicts measured during the execution of the writing request to the counterparts, between our suggested approach and optimistic one. The results presented in Figures 6, 7, 8, and 9 show in significant way that our approach solves quickly divergences and conflicts by using economic models.

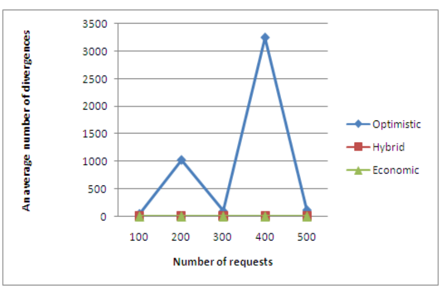

Fig. 6. Number of divergences / #Requests

Fig. 7. Number of divergences / #Requests

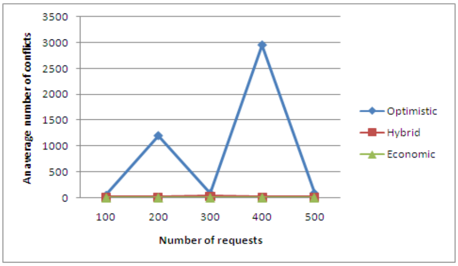

Fig. 8. Number of conflicts / #Requests

Fig. 9. Number of conflicts / #Requests

5 Conclusion and Future Works

In this work, we tried to seek balance between the performance and the quality of service. This led us to propose a service of consistency management based on a hybrid approach, combining between consistency approaches (pessimistic and optimistic) and the economic model of the market, which allows at the same time to ensure local consistency for each site and global consistency of the entire grid. The results of these experiments are rather encouraging and showed that the objective of compromise between quality of service and performance was achieved. There are a number of directions which we think are interesting, We can mention:

- We propose to integrate our approach in the form of Web service in the Globus environment by using technology WSDL [3];
- Currently, we place a replica randomly. It is useful to explore the possibility of making static or dynamic placement to improve QoS.

References

1. Belalem, G., Slimani, Y.: Hybrid Approach for Consistency Management in OptorSim Simulator. International Journal of Multimedia and Ubiquitous Engineering (IJMUE) 2(2), 103–118 (2007)
2. Buyya, R., Murshed, M., Gridsim, M.: A Toolkit for the Modeling and Simulation of Distributed Resource Management and Scheduling for Grid Computing. The Journal of Concurrency and Computation: Practice and Experience 4(13-15), 1175–1220 (2002)
3. Foster, I., Kesselman, C.: The Grid 2: Blueprint for a New Computing Infrastructure. Elsevier Series in Grid Computing. Morgan Kaufmann Publishers, San Francisco (2004)
4. Gray, J., Helland, P., O'Neil, P., Shasha, D.: The Dangers of Replication and a Solution. In: Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data, Montreal, Quebec, Canada, June 4-6, 1996, pp. 173–182 (1996)
5. Ranganathan, K., Foster, I.: Identifying Dynamic Replication Strategies for a High-Performance Data Grid. In: Lee, C.A. (ed.) GRID 2001. LNCS, vol. 2242, pp. 75–86. Springer, Heidelberg (2001)
6. Saito, Y., Shapiro, M.: Optimistic Replication. ACM Computing Surveys 37(1), 42–81 (2005)
7. Xu, J., Li, B., Li, D.: Placement Problems for Transparent Data Replication Proxy Services. IEEE Journal on Selected areas in Communications 20(7), 1383–1398 (2002)