A New Extended Group Mutual Exclusion Algorithm with Low Message Complexity in Distributed Systems

S. Dehghan and A.M. Rahmani

Abstract—The group mutual exclusion (GME) problem is an interesting generalization of the mutual exclusion problem. In the group mutual exclusion, multiple processes can enter a critical section simultaneously if they belong to the same group. In the extended group mutual exclusion, each process is a member of multiple groups at the same time. As a result, after the process by selecting a group enter critical section, other processes can select the same group with its belonging group and can enter critical section at the moment, so that it avoids their unnecessary blocking. This paper presents a quorum-based distributed algorithm for the extended group mutual exclusion problem. The message complexity of our algorithm is $O(\frac{1}{4}|Q|)$ in the best case and $O(\frac{5}{4}|Q|)$ in the worst case, where $|Q|$ is a quorum size.

Keywords—Group Mutual Exclusion (GME), Extended GME, Distributed systems.

I. INTRODUCTION

Distributed system is a set of processes (computers) connected by communication links. Mutual exclusion is a fundamental problem in distributed systems. When some resources (for example, a file, a communication channel, a printer) are shared among processes, two processes are not allowed to enter a critical section (CS) and use it at the same time. Recently, group mutual exclusion [1] has been proposed. There are multiple groups of processes. The processes in the same group can enter CS at the same time. An interesting application of the group mutual exclusion is presented in [2]. Consider large data sets stored in a secondary memory. A set of processes accesses to the data sets through a server. The server can be a CD jukebox. Using a regular mutual exclusive protocol, the server needs to repeatedly load and unload the data sets (e.g., the CDs) from the secondary memory to process the requests.

An efficient GME protocol would allow multiple processes to read the currently loaded data set (a CD) concurrently, while forcing processes requesting a different data set (another CD) to wait. For example, each process wants to read some data on the CDs. If CD $A$ is loaded, multiple processes which want to read data on CD $A$ can access it at the same time. These processes are in the same group. By contrast, the processes which want to read data on CD $B$ cannot do so when $A$ is loaded. These processes form a different group. In [3], the following extended definition has been introduced. Some processes might be members of multiple groups at the same time. In the CD jukebox example, the same data might be copied on CD $B$ and CD $C$. In this situation, the user can read the data if either CD $B$ or CD $C$ is currently loaded. A different example of GME is a server that can cache locally one page of the web. Clients of this server interested in the currently cached page can read it simultaneously; clients interested in a different page must wait [4].

For group mutual exclusion, shared memory system algorithms [4-7], token-based algorithms [8-11] and quorum-based algorithms [3, 12-14] have been proposed. The extended group mutual exclusion problems are discussed in [3] and [14] – both of the solutions are quorum-based protocols. Though [3] discusses this extended group mutual exclusion problem, it just notes that when a process belongs to multiple groups, it arbitrarily selects one group. The above algorithm is not sufficient, since when the process $p$ enters CS, another process $p'$, which can enter CS at the time might be blocked. Manabe et al. [14] called above situation as an unnecessary blocking.

In [14] a quorum-based extended group mutual exclusion algorithm is proposed that not have unnecessary blocking and the message complexity of this algorithm is $O(\frac{1}{6}|Q|)$ in the best case and $O(\frac{5}{4}|Q|)$ in the worst case where $|Q|$ is a quorum size. In the present study, we proposed a new algorithm for extended group mutual exclusion that avoids unnecessary blocking of processes. The message complexity of our algorithm is $O(\frac{1}{4}|Q|)$ in the best case and $O(\frac{5}{4}|Q|)$ in the worst case where $|Q|$ is a quorum size. Also the message complexity of our algorithm is less than [14].

The organization of this paper is as follows: In section two the system model and definition are presented. In section three the proposed algorithm introduced. In section four correctness of the algorithm is explained. In the fifth part message complexity of proposed algorithm is presented. In the sixth part, performance of proposed algorithm is compared with...
former algorithms. And finally in the seventh part the conclusion is given.

II. SYSTEM MODEL AND DEFINITION

B. The System Model

We assume an asynchronous distributed system which consists of a set of n processes $V = \{p_1, p_2, ..., p_n\}$. Each process has a unique identifier selected from a set of integers $\{1, 2, ..., n\}$. The processes communicate with each other by passing messages through first-in, first-out (FIFO), asynchronous, and reliable (no message loss occurs) channels. We assume that the system is error-free. In our algorithm, we have two classes of processes, director processes and applicant processes. A director process manages permissions and handles requests from applicant processes. An applicant process requests a set of resources and asks for permissions from director processes. In actual systems, one process can perform both roles simultaneously. Also, to avoid starvation and deadlock, each director process has a priority queue that if director process doesn’t grant permission to applicant process to enter CS, then a request is inserted to queue. Priority in this queue is assigned by a timestamp, which is a pair logical clock value on request and process identifier [15-16].

B. Coteries

We describe formal definition of coterie [17] below.

**Definition (Coterie [17]):**

Let $U = \{u_1, u_2, ..., u_n\}$ be a set. A set $C$ of subsets of $U$ is a coterie under $U$ if and only if the following three conditions are satisfied.

1. **Non-emptiness:** For each $Q \in C$, $Q$ is not empty and $\emptyset \subsetneq U$.
2. **Intersection property:** For any $Q, Q' \in C$, $Q \cap Q'$ is not empty.
3. **Minimality:** For any $Q, Q' \in C$, $Q$ is not a subset of $Q'$.

An element of $C$ is called a quorum.

Coteries are widely used for distributed mutual exclusion. Coteries are a set of quorums, and a quorum is a subset of processes $V$ such that any two quorums have non-empty intersection. A process wishing to enter its critical section sends a request to each process in a quorum. A process enters its critical section after it obtains permission from every process in a quorum, and releases the permission when it exits its critical section. Because the two quorums have nonempty intersection, mutual exclusion is guaranteed if each process never grants permission to more than one process at a time.

II. DESCRIPTION OF THE PROPOSED ALGORITHM

This section provides a brief description of our algorithm shown in figures 1 and 2. Let $C = \{Q_1, Q_2, ..., Q_k\}$ be a set of quorums, $A = \{p_1, p_2, ..., p_n\}$ be a set of applicant processes and $G_{\omega} = \{G_1, G_2, ..., G_k\}$ be a set of groups. When process $p_i$ wants to enter CS, it selects its group set $G_i$ from $G_{\omega} = \{G_1, G_2, ..., G_k\}$, and it selects a quorum $Q \in C$ and sends its request to every $q \in Q$.

When director processes receive this request, they grant permission to $p_i$ provided that no permission was granted before and CS is empty. When $p_i$ receives permission from each $q \in Q$, selects desired group from its own group set $G_i$ and enters CS. Other applicant processes that are in the same group with $p_i$ can enter CS concurrently. As a result, extended GME is guaranteed and their unnecessary blocking is avoided.

But this definition: "allowing another process to enter CS at the same time", leads to starvation [14]. Thus, to avoid starvation, the following method is used. When CS is empty, the first process which enters CS through selecting the desired group is as a master. Then, other applicant processes belonging to the same group can enter CS as slave simultaneously while master is in CS. As soon as master exits from CS, director processes do not grant permission to other processes. Then, after all slaves exit from CS, other processes enter and therefore starvation is avoided. In this case director processes select one request based on timestamp of suspended requests in queue and send permission to enter CS as master. After the master enters, director processes send permission to other suspended requests in queue that are in the same group with master to enter CS as slave.

A. Solutions of Proposed Algorithm for Improvement of Former Algorithm

This section presents solutions of proposed algorithm to decrease communication messages and the used memory by queues of director processes. In [14] each $q \in Q$ receives a request, first inserts it to queue. However only the suspended request can enter to queue and decrease used memory by queues. Also, when the pivot process wants to exit CS, it must waits for other process to exit from CS which cause improper wait of pivot process.

To eliminate improper wait of pivot process, the director processes must distinguish the exit of master and slave processes from CS. In fact guarantee of mutual exclusion is the duty of director processes.

For this purpose, four variables are used. Master is a string variable that saves master process name, Master_in is a boolean variable that shows the presence or absence of master in the CS, user is a set variable which shows the set of processes which serve as slaves in the CS and count is an integer variable which shows the number of processes that allowed them be as slaves but have not enter CS yet.

<table>
<thead>
<tr>
<th>Program</th>
<th>Applicant Process (p; process);</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-3.</td>
<td></td>
</tr>
</tbody>
</table>
The pseudo code for the applicant process algorithm is shown in Figure 1. When process \( p \) (group set is \( G \)) wants to enter CS: it sends a request to every process in \( Q \) with highest priority based on its timestamp and grants permission to it in order to enter CS as a master.

### B. The Applicant Processes Algorithm

The summary of the pseudo code for applicant processes shown in Figure 2 is as follows:

- **A-1.** When process \( p \), whose group set is \( G \), wants to enter CS, it selects a quorum \( Q \in C \) and sends a request to every process in \( Q \).
- **A-2.** If any of the \( q \in Q \) granted permission to enter as a slave, then it enters CS as a slave.
- **A-3.** If all \( q \in Q \) granted permission as master, it would select the desired group \( g \) from whose group set \( G \), and enter CS as master.
- **A-4.** If no permission is granted, it should wait. Also when \( p \) exiting CS, sends release message to every \( q \in Q \).

### C. The Director Processes Algorithm

The summary of the pseudo code for director processes shown in Figure 2 is as follows:

- **B-1.** When \( q \) receives a request from \( p \): 
  1. If permission is granted to other process to enter CS as a master, then it inserts its request to queue and doesn’t grant permission to \( p \).
  2. If no permission is sent to other process that enter CS as a master before, and if CS is empty, then it sends permission to \( p \) which enters CS as a master.
  3. Otherwise if master is in the CS, if \( p \) is in the same group with master, it allows \( p \) to enter as a slave and increments scount. But if master isn’t in the CS, it inserts its request to queue and doesn’t grant permission to \( p \).
- **B-2.** When \( q \) receives message that \( p \) has entered CS as a slave, then it adds \( p \) to user set and decrements scount.
- **B-3.** When \( q \) receives message that \( p \) has entered CS as a master, then \( q \) all process \( p \) in queue if \( M \in G \), it sends permission to enter CS as a slave and remove those from queue.
- **B-4.** When \( q \) receives release message from \( p \), if there is no other process in the CS and don’t send permission to other process that enter CS as a slave before (that is scount = 0), then it removes one request from queue with highest priority based on its timestamp and grant permission to it to enter CS as a master.
B.1. At arrival of "Request_Master(G_i, ts)" from $p_j$: 

```java
if (Dstatus = = wait_master) {
    insert ($p_j$, $G_i$, ts) to Queue;
    Send "No_Enter" to $p_j$;
}
if (Dstatus = = Idle) {
    if ($CS$ = = 'empty') {
        send "Enter_Master" to $p_j$;
        Dstatus = wait_master;
    }
    if (Dstatus = = Idle) ∨ (Dstatus = = wait_slave) {
        if ($Mg$ $∈$ $G_i$) ∧ (Master_in = = true) {
            send "Enter_Slave($Mg$)" to $p_j$;
            Dstatus = wait_slave;
        }
        else {
            insert ($p_j$, $G_i$, ts) to Queue;
            send "No_Enter" to $p_j$;
        }
    }
}
```

B.2. At arrival of "Slave($Mg$)" from $p_j$: 

```java
if (Dstatus = = wait_slave) {
    user = user $∪$ ($p_j$);
    scount $−$ ;
    if (scount = = 0) Dstatus = Idle;
}
```

B.3. At arrival of "Master($g$)" from $p_j$: 

```java
if (Dstatus = = wait_master) {
    Master = $p_j$;
    Master_in = = true;
    $Mg$ = $g$;
    $CS$ = 'full';
    ∀ all processes $p_i$ in Queue {
        if ($Mg$ $∈$ $G_i$) {
            Send "Enter_Slave($Mg$)" & remove it from Queue;
            Dstatus = wait_slave;
            scount $++$;
        }
        else
            Dstatus = Idle;
    }
}
```

B.4. At arrival of "Release" from $p_j$: 

```java
if (Dstatus = = waitSlave) {
    if ($CS$ = = $φ$ $∧$ (scount = = 0)) {
        $CS$ = 'empty';
        if (Queue is not empty) {
            Remove item ($p_j$, $G_i$, ts) from Queue
            with highest Priority
            & send "Enter_Master" to $p_j$;
            Dstatus = wait_master;
        }
    }
    else Dstatus = Idle;
}
else {
    user = user $\backslash$ ($p_j$);
    if ($user$ = = $φ$) $∧$ (Master_in = = false) $∧$
    (scount = = 0) {
        $CS$ = 'empty';
        if (Queue is not empty) {
            Remove item ($p_j$, $G_i$, ts) from Queue
            with highest Priority
            & send "Enter_Master" to $p_j$;
            Dstatus = wait_master;
        }
    }
    else Dstatus = Idle;
}
```

Fig. 2. The pseudo code of the director process algorithm

IV. PROOF OF CORRECTNESS

A. Safety

The mutual exclusion requirement in GME problem says that, no two processes requesting for a different group, must be in their CS simultaneously. Suppose $p_1$ and $p_2$ are applicant processes that request different groups. Let $q_1$ and $q_1$ be quorums that $p_1$ and $p_2$ select, respectively. Because any two quorums have non-empty intersection, we have $Q \cap Q' \neq 0$ then if $p_j$ be a director process in the intersection, since $p_j$ never sends permission for more than one group at a time, therefore $p_1$ and $p_2$ can not be granted by $p_j$ simultaneously.
B. Freedom from Starvation

Assume that starvation occurs. Let \( p_1 \) be the process in group \( g \) starves. We assume that \( \text{ts}(p_1) \) is the smallest among starved processes. It is clear that the request of \( p_1 \) is eventually granted if no process is its critical section and no other process is making a request.

Therefore we consider a case that some processes in group \( g \neq g \) repeat entering and exiting their CS.

According to proposed algorithm, director processes don't grant permission to other processes to enter CS as a slave after exiting master process from CS. Then, all slave processes exit from CS finally. Also the timestamp value for each request is increasing.

As a result, when all processes exit from CS, the director processes remove one request from queue with highest priority based on its timestamp and grant permission to it to enter CS as a master. Thus because \( \text{ts}(p_1) \) is the smallest among other request, is eventually granted permission and enters its critical section.

V. MESSAGE COMPLEXITY

Message complexity of the proposed algorithm is \( O(4|Q|) \) in the best case and \( O(5|Q|) \) in the worst case, where \(|Q|\) is the size of the smallest quorum in a coterie.

In the best case, only one request exists to enter CS. Then four following types of messages are exchanged between an applicant process and each director process in a quorum:

a. Applicant process \( p_i \) sends "Request Master(\( G, \text{ts} \))" to every \( q \in Q \).

b. \( p_i \) received "Enter_Master" from every \( q \in Q \).

c. \( p_i \) selects desired group \( g \) from whose group set \( G_i \) and sends "Master(\( g \))" to every \( q \in Q \), and enter CS as master.

d. When \( p_i \) exiting CS, sends "Release" to every \( q \in Q \).

Thus, message complexity is \( O(4|Q|) \) in the best case.

The worst case happens in two conditions. The first condition is the time when applicant process is not the same group as CS inside processes, the next condition is the time when applicant process is in the same group as processes CS inside but master process exits from CS. In these two conditions, one additional message in comparison with best case is communicated between applicant processes and quorum members. Then five following types of messages are exchanged between an applicant process and each director process in a quorum:

a. Applicant process \( p_i \) send "Request Master(\( G, \text{ts} \))" to every \( q \in Q \).

b. \( p_i \) received "No_Enter" from some of the \( q \in Q \), then wait for permission.

c. After a time \( p_i \) received one of this messages:

   - "Enter_Master" from every \( q \in Q \), then selects desired group \( g \) from whose group set \( G_i \) and sends "Master(\( g \))" to every \( q \in Q \), and enter CS as master.

   - "Enter_Slave(\( Mg \))" from some of the \( q \in Q \), then selects \( Mg \) as desired group and sends "Slave(\( Mg \))" to every \( q \in Q \), and enter CS as slave.

d. When \( p_i \) exiting CS, sends "Release" to every \( q \in Q \).

Thus, the message complexity of proposed algorithm is \( O(5|Q|) \) in the worst case.

VI. PERFORMANCE COMPARISON

In this section, the complexity and extended GME ability of proposed algorithm compared with three former algorithms: Joung [3], Toyomura et al. [13] and Manabe et al. [14]. Table I shows this comparison. The result shows that proposed algorithm has lower complexity in both best case and worst case with respect to other algorithms and has extended GME ability which after the process by selecting a group, enters critical section, other processes can select same group with its belonging group and can enter critical section at the moment, so avoid their unnecessary blocking.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Best case</th>
<th>Worst case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joung [3]</td>
<td>( \sqrt{\frac{2n}{m(m-1)}} )</td>
<td>( n )</td>
</tr>
<tr>
<td>Toyomura et al. [13]</td>
<td>( O(3</td>
<td>Q</td>
</tr>
<tr>
<td>Manabe et al. [14]</td>
<td>( O(6</td>
<td>Q</td>
</tr>
<tr>
<td>New Proposed Algorithm</td>
<td>( O(4</td>
<td>Q</td>
</tr>
</tbody>
</table>
VII. CONCLUSION

In the present paper, we proposed a new quorum based algorithm for the extended group mutual exclusion problem in distributed systems. The message complexity of our algorithm is $O(4|Q|)$ in the best case and $O(5|Q|)$ in the worst case. Also the message complexity of our algorithm is less than former algorithms.

REFERENCES