FRAMEWORK FOR DISTRIBUTED COMPUTING

PROJECT WORK DONE AT

Lucid Technologies Pvt. Ltd., Coimbatore

PROJECT REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE DEGREE OF

MASTER OF COMPUTER APPLICATIONS

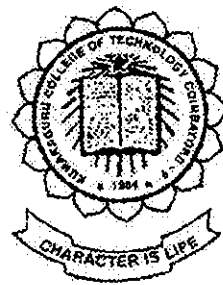OF BHARATHIAR UNIVERSITY, COIMBATORE.

SUBMITTED BY

M.Ranjith Prakash

9938MO628

GUIDED BY

Mr. S. Chandrasekaran M.C.A, M.B.A   **Mr. S.GaneshBabu M.C.A**

EXTERNAL GUIDE   INTERNAL GUIDE

Department of Computer Science and Engineering

KUMARAGURU COLLEGE OF TECHNOLOGY

Coimbatore – 641 006

May 2002

CERTIFICATES

Department of Computer Science and Engineering

KUMARAGURU COLLEGE OF TECHNOLOGY

(Affiliated to the Bharathiar University)
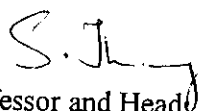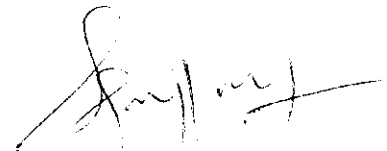
Coimbatore – 641 006

CERTIFICATE

This is to certify that the project work entitled

"FRAMEWORK FOR DISTRIBUTED COMPUTING"

Done by

M.Ranjith Prakash
9938MO628

Submitted in the partial fulfillment of the requirements for the award of the degree of

Master of Computer Applications of Bharathiar University

Professor and Head   Internal Guide
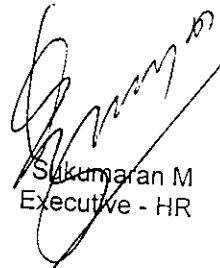
Submitted to the university Examination held on : 09 - 05 - 2002

Internal Examiner   External Examiner

Lucid Technologies Pvt. Ltd.

Bonafide Certificate

Certified that this thesis on "General Framework for distributed computing" is the bonafied work of **Mr. Ranjith Prakash M, Register Number: 9938M0628,** who carried out the project in our organization during the period December 2001 to April 2002. Certified further that to the best of my knowledge, the work reported there-in does not form part of any other thesis or work on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.
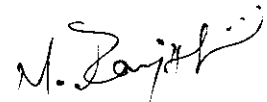
Sukumaran M
Executive - HR

DECLARATION

I here by declare that the project entitled "**FRAMEWORK FOR DISTRIBUTED COMPUTING**" Submitted to **Bharathiar University** as the project work of Master of Computer Applications Degree, is a record of original work done by me under the supervision and guidance of **Mr.S.Chandrasekaran, Senior Software QA Engineer, Lucid Technologies Pvt. Ltd, Coimbatore** and **Mr.S.GaneshBabu, Lecturer, Kumaraguru College of Technology, Coimbatore** and this project work has not found the basis for the award of any Degree/ Diploma / Associateship/ Fellowship or similar title to any candidate of any university.

Place: Coimbatore

Date: 30-04-2002

Signature of the Student

Countersigned by

INTERNAL GUIDE

ACKNOWLEDGEMENT

ACKNOWLEDGEMENT

My heartfelt thanks to **Dr. K.K. Padmanabhan B.Sc.,(Engg),M.Tech,Ph.D, Kumaraguru college of Technology, Coimbatore** for having granted permission to undertake this project work.

My heartfelt gratitude to Dr. **S.Thangasamy Ph.D., Professor and Head of Department, Kumaraguru College of Technology, Coimbatore** for guiding me to complete the project successfully.
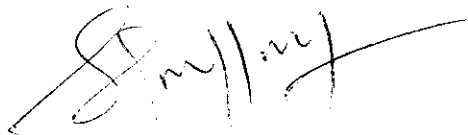
I am greatly indebted to Lucid Technologies Pvt. Ltd., for providing me the opportunity to do this project at their organization. I take an immense pleasure in thanking Managing Director of Lucid Technologies Pvt. Ltd.

I express my profound gratitude to my external guide **Mr. S.ChandraSekaran, Senior Software QA Engineer, Lucid Technologies, Coimbatore** for providing excellent guidance for the timely completion of project. His sound expertise and experience helped me gain a lot through the period of this project.

My deepest acknowledgement to my internal guide **Mr. S.GaneshBabu,** for his whole hearted assistance and support all round the project. His valuable guidance has an unmatched impact on project. His trust and confidence in my abilities gave me immense strength to complete the project successfully.

My sincere word of thanks to course coordinator **Mrs. V.Geetha**, and the project review team , for their excellent guidance for timely completion of project.

My acknowledgement would be incomplete without thanking my mother and friends for their support, love and understanding throughout the endeavor. And special thanks to my friend Tijo who helped me during the course of project.

M. Ranjith Prakash

SYNOPSIS

Distributed Computing is a term used to denote a large class of techniques that are used to provide simultaneous data processing tasks for the purpose of increasing the computational speed of a system.

My project is an attempt to creating a General Framework for the Distributed Applications for windows platform. This project aims at efficiently utilizing the network resources by distributing a complex computation, which exhibits parallelism.

Conventionally, to solve problems involving intensive computations researchers made use of powerful computers, which were exorbitantly expensive. Generally PC users do not make optimum use of the processor all the time. The processor is kept idle most of the time. My project aims at using the CPU time for parallel evaluation of the task.

The system has been developed on a LAN consisting of PC's running in windows operating system. The system software necessary has been developed and the system was put in test by assigning it a computationally intensive task.

This project is successfully completed and demonstrated with the difference in processing time when the application was run on stand alone and on the cluster. This project has given me the opportunity to learn the fundamental concepts of Distributed Computing and also the network programming using windows sockets.

TABLE OF CONTENTS

Page No.

INTRODUCTION

1.1 PROJECT OVERVIEW:

The System can be defined as "**a parallel or distributed system that consists of a collection of interconnected whole computers, that is utilized as a single, unified computing resource**". In general, the goal of a cluster is to make it possible to share a computing load over several systems without either the users or system administrators needing to know that more than one system is involved. Ideally, if more processing power is needed the user simply "plugs in a new component" (Computer), and presto, the performance of the system as a whole improves.

With extremely powerful processors available in the market there is an increasing tendency to upgrade our systems with the latest processors available. But honestly how many of us really use the full power of these processors. Ultimately we end up with powerful number crunching processors sitting in our systems waiting idly for some job to be given.

Well irrespective of whether we give the computer some job to do or not, it keeps computing at the same rate of 566 MHz (or whatever is its clock speed) the clock never slows down. The processor keeps going in a loop waiting for some job to be assigned to it; this only means that's we are wasting a lot of computing power.

So is there any way we can use this excess power of the processors. Thankfully people have done it called the parallel processing. It is the technique of putting together cheap low cost chips and motherboards meant for normal desktop PC's to construct a powerful parallel processing architecture. NASA did the most pioneering work in this type of architecture. The application developed by them is called seti@home and it was a splendid success.

1

This project is really exciting and interesting and my aim was to take a computationally intensive process distribute it over a cluster of systems and demonstrate the improvement in time efficiency and that is exactly what I achieved.

I took up the process of database application payroll, with a huge number of records, and I distributed among clusters and results were achieved very quicker than stand alone system application.

1.2 ORGANIZATION PROFILE:

Lucid Technologies Pvt. Ltd.
Coimbatore

Company specializes in **eServices** like Development of Web Applications, Integration of legacy systems, Software Quality Assurance and Product Support. Company empowers clients with cutting edge solutions and enables them to gain competitive advantage needed to overcome e-business challenges. The company differentiates itself by becoming a one-stop eServices shop by offering quality services to global customers.

eServices :

Company offers eServices for development, deployment and management of ebusiness applications. They work closely with our customers providing them with cutting edge solutions; services and support to enable them gain competitive advantage. Company leverages its expertise and technology to help the business gain technical advantage and improve at the speed of echange.

Company can help to transform the business into a successful ebusiness by using proven methodologies. They also offer reliable turnkey solutions to transform your business plans into ecommerce sites. The personalized support that they do on customer behalf can help increase the customer satisfaction.

Engineering Services:

With state-of-the-art CAD/CAE development facility at Chennai, India, they offer a diverse range of CAD/CAE services. Using powerful products like CATIA, IDEAS, Pro-E, Unigraphics, ANSYS, NASTRAN combined with the extensive engineering expertise, Lucid offers you best-in-class services and optimum value for money.

They offer the following services on the CAD/CAE platforms mentioned above

* Drafting and Documentation
* Digitizing
* Component design, assembly design
* Solid & surface modeling

Products:

Integrated Business Solution (IBS) Version 1.2

IBS is a fully integrated multi-user application targeted at small and medium manufacturing companies, designed and developed by Lucid Technologies. The software supports multiple companies, each with multiple locations with multiple stocking points at each location. The software supports user-defined security for access control (at many levels), which can be configured by the administrator.

3

Other Services:

Company offers the following additional services:

* Migration from Novel to NT

* Software development using C/S and n-tier Architecture based on MS Visual Studio Tools

* Design and implement email solution based on MS Exchange 5.x

* Leased Internet connectivity based on MS back office components

* Developing applications on MS Office platform using VBA

* Intranet and Extranet development

SYSTEM STUDY & ANALYSIS

2.1 EXISTING SYSTEM:

Conventionally, to solve problems involving intensive computations researchers made use of powerful computers, like Super computers, which were exorbitantly expensive. When we want to reduce the cost, we have to go for simple systems PC's but the time to execute the computations will increase dramatically.

Also here the operations are performed in the single system, and if there is any hardware or software failure occurs we lose the valuable data, so the availability is very less in this type of system. When the time progresses the database size increases and there will result in slowness of the system. These are the characteristics of the existing system.

2.2 PROPOSED SYSTEM:

Proposed System can be defined as "a parallel or distributed system that consists of a collection of interconnected whole computers, that is utilized as a single, unified computing resource ". The goal of a cluster is to make it possible to share a computing load over several systems without either the users or system administrators needing to know that more than one system is involved.

Marketing studies showed a huge demand for higher availability in small businesses as databases and electronic mail have become essential to their daily operation. These businesses cannot afford specialized computer operations staff, so ease of installation and management was defined as key product advantages.

2.3 REQUIREMENTS ON NEW SYSTEM:

Load Sharing:

The goal of a cluster is to make it possible to share a computing load over several systems without either the users or system administrators needing to know that more than one system is involved.

Availability:

The System is highly available. Even though a hardware or software failure occurs, we can execute the program as programs and the database are replicated over the clusters, and any cluster at any time can act as a sever, to invoke the operations.

Expandability:

If more processing power is needed the user simply "adds a cluster to the network", and, the performance of the system as a whole improves.

Shared Nothing Clusters:

The characteristic of the cluster used in this project is Shared Nothing Clusters, which means during the execution of the process, a cluster does not depend on other cluster, or wait for other cluster to complete the operation.

2.4 USER CHARACTERISTICS:

Any one who is interested to make the application distributed can use this System. All functions required to make a distributed application is implemented in a DLL (Dynamic Link Library). Any project can refer to this DLL file and can use these functions implemented in it.

This project is implemented on a payroll application. This application is used by the management for the purpose of calculating pay slip of the employees in the organization. This has the very good user interface, and efficient coding making the customer very satisfied. The company records are maintained safely and securely as authentication is checked for each and every user. Authentication is done by checking password for every user.

3.1 HARDWARE CONFIGURATION:

- Minimum of three PC's with Pentium processors running in Windows operating system.

- The Systems should be inter-connected in a network.

- All the systems should have a minimum of 32 MB RAM.

- Processor speed 200 MHz or above

- Hard disk space 2 GB or above

3.2 SOFTWARES AND TOOLS USED:

Microsoft Visual Basic 6.0:

Visual basic is the software used as the front end as it is highly user interactive. In visual basic, creating an application is very much easy and faster than that of the other software like VC++. Even though VC++ is much more powerful, I have selected visual basic due to the limitation of time and also visual basic is very well suited for my project.

Microsoft Access:

Access is the back end, which I have used for my project. This comes with the Microsoft office setup. Even though powerful software's like oracle and SQL server is available, I have selected Access because, In distributed computing we are going to use many computers which can be incrementally added to the network, we cant be sure that Oracle or SQL server is available in that computer. Due to the highly availability of Access I have chosen this as back end.

Help Author:

This is the tool used to create the help file for the user. It is very user interactive and an easy way to create the help file. Help file created through this tool looks alike that of that windows help format, so user is more convenient in using help file created using this Help Author Tool.

Adobe Photoshop 5.0:

This is best tool available in the market for creating and editing images. For the view point of the user the project should look pleasant .In this project I have used Photoshop to create the backgrounds and also the buttons.

4.1 INPUT DESIGN:

The input design is the important phase in the design of software because the design for handling input data specifies how data area accepted for computer processing. Generally the computer system has intensive interaction with the outside world, mostly this interaction takes place through visual terminals such as monitors. So while developing a software system we have to take care of efficient development of interactive and user-friendly interface.

The quality of input system, determines the quality of the system output. Input design features can ensure reliability of the system and produce results from accurate data. All the data validation is done during the input from the user, and thus we do not allow any wrong entries to the system.

4.2 OUPUT DESIGN:

Outputs from the system are required to communicate the results of processing to users. They are also used to provide a permanent copy for later consultation. The Output of this system is purely an INTERNAL output, as the destination is with in the organization. In this system, the outputs are given in the form of reports, which the user can view in the monitor itself, or he can take hard copy of it.

Type of the output is in report format. The contents like headings, data members, format changes from report to report. The frequency of taking reports depends on the management instruction. The location of data required to take the output is available in the local machine itself as the results are send back to the server machine after the process is completed in the cluster machine. The reports are attached at the end of this documentation.

10

4.3 <u>DATABASE DESIGN:</u>

The Database is defined as ' an integrated collection of data '. The overall objective in the development of database technology has been to treat data as an organizational resource and as a integrated whole. In a database, information from several files is coordinated, accesses and operated upon as though it is in a single file. Logically, the information is centralized; physically, the data may be located on different devices and in widely scattered geographical locations, connected through data communication facilities.

Database Design of Distribution Framework :

The database of framework contains five tables. The design of each table is given below.

1. Cluster Details Table:

This table contains the information about the various clients , their name, whether they are online and whether its Database is update.

Field Name	Data Type	Description
Cluster Name	Text	Name of Cluster got from the user during registration
Status	Boolean	States whether Cluster is online / offline
DbUpdate	Boolean	States whether Cluster Database is update or not

2. Fragment Table:

These tables contain the information that how many records are allocated to each cluster for the processing.

Field Name	Data Type	Description
Cluster Name	Text	Name of Cluster got from the user during registration
Start Record Number	Number	Gives the detail about the start record number given for that cluster
Limit	Number	The number of record the particular cluster to execute

3. Query Table:

All the operation performed in the application is stored in a form of query in this table for the purpose of updating the cluster database .

Field Name	Data Type	Description
Query Number	Number	Automatic number generated for each query
Query	Text	The actual query is stored here.

4. Query Update:

Here it contains the information, from which query it has to be send to the cluster for the purpose of updating the cluster databse while it joins in the network.

Field Name	Data Type	Description
Cluster Name	Text	Got during the registration
Start Query Number	Number	From this number the query in query table is transferred to cluster.

5. User Detail:

This is the table available in the Cluster side. The tables listed above are all present in the server side. This contains the data about the name of its computer given during the registration phase and also the number of records it has to execute when server asks to do it.

Field Name	Data Type	Description
Cluster Name	Text	Name given by the cluster user during registration
Start Record Number	Number	Gives the detail about the start record number given for that cluster
Limit	Number	The number of record the particular cluster to execute

13

Database design for Payroll System :

The database of payroll system contains seven tables. The structure of each table is explained below.

1. Employee Detail:

   This contains the complete information about the employees available in the company.

Field Name	Data Type	Description
MstEmpName	Text	Name of employee
MstEmpUcode	Text	Code given to employee
MstEmpDegCode	Text	Designation code assigned to employee
MstEmpAddress	Text	Address of employee
MstEmpEmail	Text	Email of employee
MstEmpPhone	Number	Phone number of employee
MstEmpMobile	Number	Mobile number of employee
MstEmpDoj	Date	Employee joining date

2. Designation Detail:

   This contains all the information about the various designations available in the organization.

Field Name	Data Type	Description
MstDesgUcode	Text	User Code entered by the Operator
MstDesgTitle	Text	Title of the Designation
MstDesgBasicPay	Number	Basic Pay of Deisgnation
MstDesgDa	Number	Dearness Allowances
MstDesgTa	Number	Travel Allowances
MstDesgHra	Number	House Rent Allowances
MstDesgCca	Number	City Allowances
MstDesgPf	Number	Provident Fund
MstDesgLic	Number	Insurance
MstDesgEsi	Number	ESI of Employee
MstDesgOtherAllow	Number	Other Allowances
MstDesgOtherDeduct	Number	Other Deductions
MstDesgCasualLeave	Number	Total Allowed casual days
MstDesgMedicalLeave	Number	Total Allowed medical days

3. Login Details:

This maintains the various users of this system, their passwords, their rights on each form and also the last access date.

Field Name	Data Type	Description
MstLoginName	Text	Login name of user
MstLoginPwd	Text	Password of the user
MstAccessDate	Date/Time	Date of last access
FrmRead	Number	Denotes rights given to forms

4. Leave Taken:

This table contains the information about the leave taken by the employees.
The structure is as follows

Field Name	Data Type	Description
TrnEmpucode	Text	Employee Code
TrnEmpDesgucode	Text	Designation Code
TrnLevFrmDt	Date/Time	Leave Taken From date
TrnLevFrmSess	Number	Leave Taken From Session
TrnLevToDt	Date/Time	Leave Taken To Date
TrnLevToSess	Number	Leave Taken To Session
TrnNoofDays	Number	Total Days Availed
TrnLeaveType	Number	Medical or Casaul Leave

5. Working Days:

This table gives the details about number of working in every month of the
year

Field Name	Data Type	Description
MstMonthId	Number	Month
MstNoofWorknDays	Number	Number of working days

6. MonthWise Leave Taken:

This table contains the data about every employee , the total leave taken by
him for the entire month.

16

Field Name	Data Type	Description
TmEmpUcode	Text	Employee Code
TmMonthId	Number	Month Id
TmLeaveTaken	Number	Number of leave taken days
TmOverflow	Number	Sets when leave exeeds the allowed days

7. Payslip Table:

The final result is stored after the distributed computation is over. The structure of table is given as,

Field Name	Data Type	Description
TmEmpUcode	Text	Employee Code
TmDesgUcode	Text	Designation Code
TmEmpBasicPay	Number	Basic Pay
MstDesgDa	Number	Dearness Allowances
MstDesgTa	Number	Travel Allowances
MstDesgHra	Number	House Rent Allowances
MstDesgCca	Number	City Allowances
MstDesgPf	Number	Provident Fund
MstDesgLic	Number	Insurance
MstDesgEsi	Number	ESI of Employee
MstDesgOtherAllow	Number	Other Allowances
MstDesgOtherDeduct	Number	Other Deductions
TmMonthId	Number	Month ID
TmYear	Number	Year
TmEmpGross	Number	Gross Pay
TmEmpNet	Number	Net Pay

17

Thus we have two databases, one for the payroll system and another for the distributed framework and totally we have eleven tables. The structures of all the tables are explained clearly above.

4.4 PROCESS DESIGN:

Process design gives the series of operations designed to manipulate data to produce output from a computer system. Here the process at lower levels of detail, which will define the detailed steps, to be taken to produce the specified computer output from the initial input of data. Here we use the design tools like flow charts, Data flow Diagram to explain the process.

Process design has to be done separately for the cluster process and the server process. Before that we have to look at the architecture of the system to be developed.

The Server Machine :

This is a central machine, which controls the overall execution of the process. This machine requires good processing configuration, as it has to perform many functions. This splits the problem in to small parallel executable process according to the number of site machines available, and assigns the independently executable process to the site machines with the data required. Then after the process completed at the site machine, it will sent the solution to the server machine. After this the Server will reform the solution from various clients to get the overall solution for the problem.

This Server system is the one, which is used by the END USER. He just submits the job to this simulator machine and then Simulator will take care of all the things. This provides the user a *single system image*, which is a feature of the project.

18

The several modules in the server machine are given below.

1. Monitor :

This module performs the following functions of

- Maintaining the status of execution

- Maintaining the database about the various cluster machines available and which are in execution etc.

2. Process Decomposer:

Here we have the procedure to split the program in to small parallel executable process and assign this process to various cluster together with the data required.

3. Solution Extractor:

After the cluster machines complete the execution, it will return the solution to the Server, and here the individual solutions are compiled to get the overall solution.

4. Database Controller:

This contains the data that are necessary to execute the process and also has the control mechanisms over the data with respect to the clusters.

The Cluster Machine :

These are the system, which has the capability to performing the job assigned to it with its own processing resources. These machines should have the resources to connect to the communication media. We have to develop the application for site machine to get the data from the Server machine, and to execute it properly and to send back the solution to the Server after the processing is completed.

Any computer in the world, no matter where it is can act as the site machine by registering in to the server and can start the processing by downloading the software required. This is a good feature, as even though we do not have much number of systems available locally to act as a cluster machine, we can use another persons system, or some other remote LAN, after it gets registered in the Server.

The several modules in the cluster machine are given below.

1. Manager :

This module performs the various functions of

- Maintaining the status of the process
- Maintaining the database required for the process
- Handling the exception conditions of network

2. Application:

This is the actual application that executes the process assigned to it by   the server.

ARCHITECTURE OF PROPOSED SYSTEM
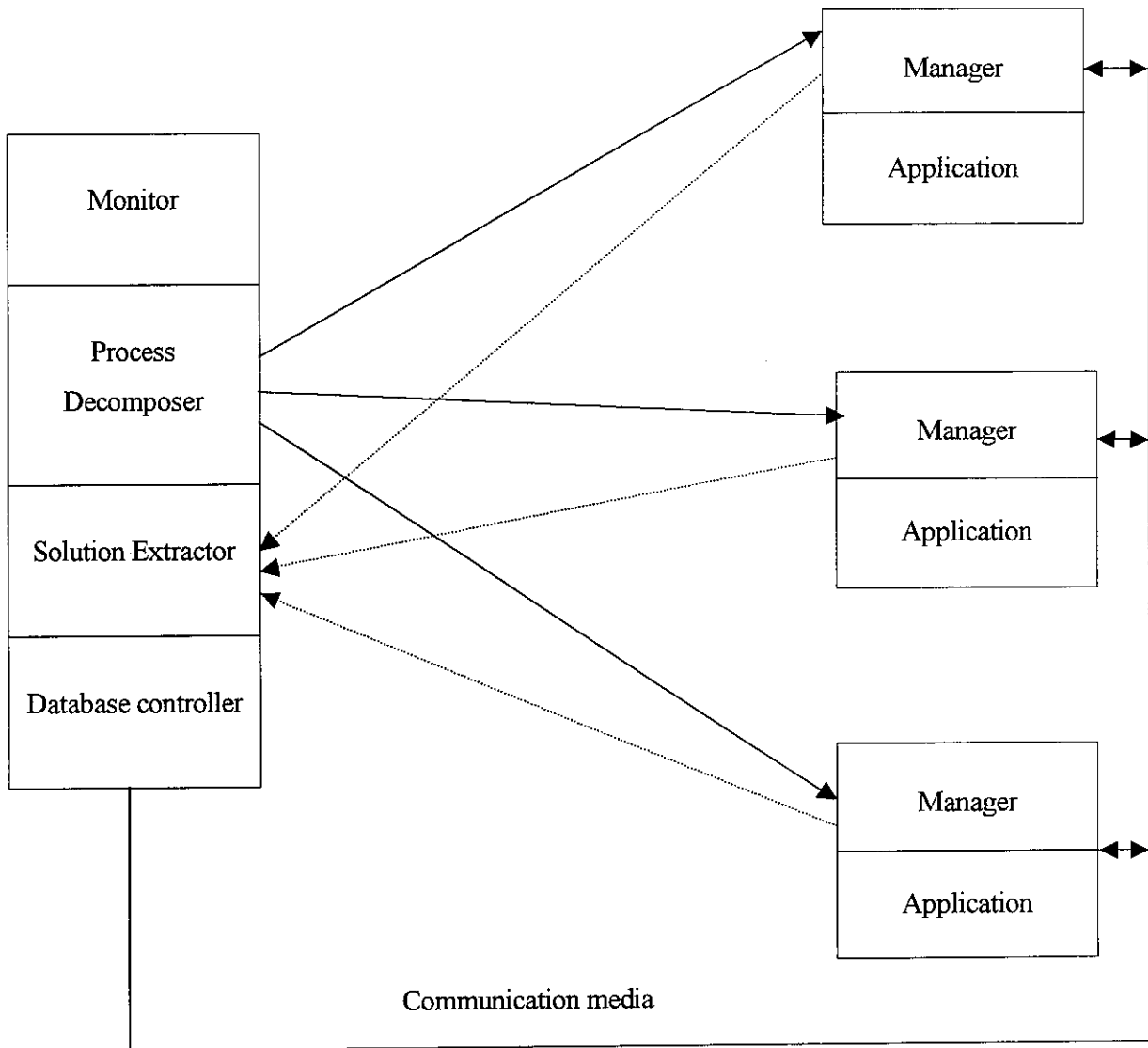
--- General Architecture of the proposed system ---

STAGES OF THE CLUSTER:

This diagram gives the link between the modules available in the cluster machine.
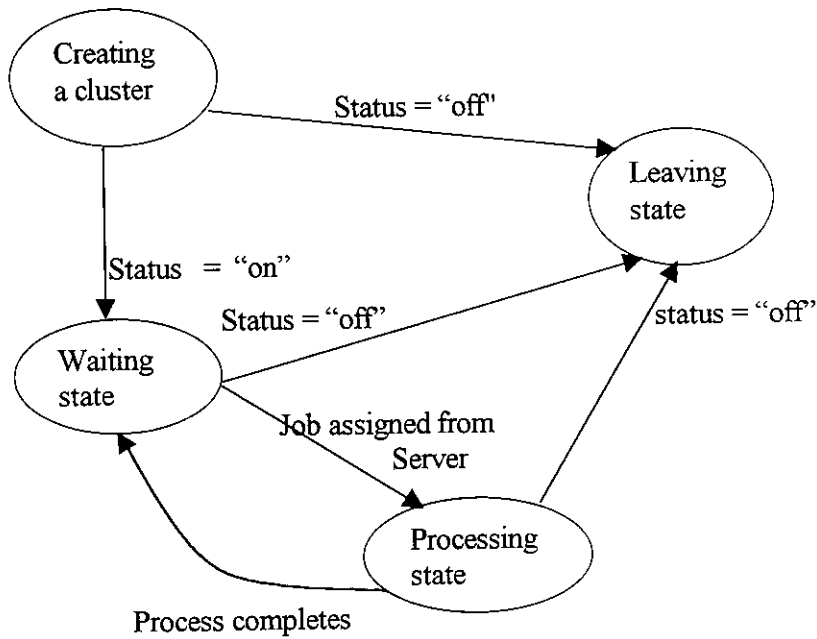
CREATING A CLUSTER :

When a machine has to be converted to a cluster, first we have to install the application. The process of registering the cluster to the central server is called Creating a cluster.

The process design of creating a cluster is given as flowchart below,

Flow chart for process of creating cluster :

```
                        ╭──────────────────────╮
                        │   Setup prg. Started  │
                        ╰──────────────────────╯
                                   │
                                   ▼
                                 ◇◇◇◇◇
            NO                 Is entry                 YES
         ┌─────────────────   already   ───────────────────┐
         │                    available                    │
         ▼                      ◇◇◇◇                        │
  ┌──────────────────┐                                      │
  │ Send the user name│                                     │
  │ from the cluster to the │                               │
  │ server            │                                     │
  └──────────────────┘                                      │
         │                                                  │
         ▼                                                  │
  ┌──────────────────┐                                      │
  │ Make an entry in the local │                            │
  │ table USER DETAIL with the │──────────────┐             │
  │ cluster name      │                        │            │
  └──────────────────┘                         └──────────▶ │
         │                                                  │
         ▼                                                  │
  ┌──────────────────┐                                      │
  │ Update table Cluster │                                  │
  │ Detail in the server. │                                 │
  │ With the cluster name. │                                │
  └──────────────────┘                                      │
                                                            ▼
                                          ╭──────────────────────╮
                                          │   End of creating     │
                                          │      cluster          │
                                          ╰──────────────────────╯
```

23
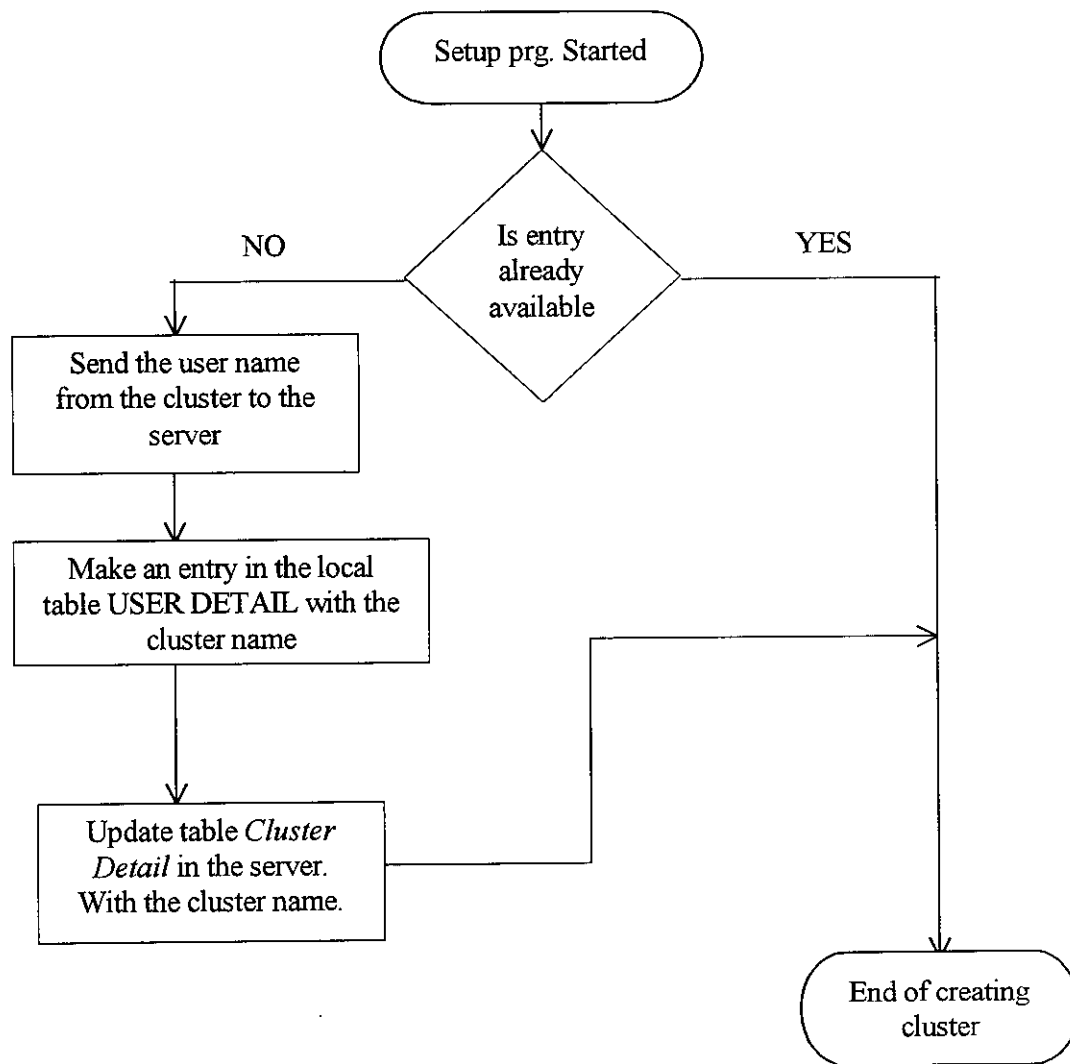
JOINING THE CLUSTER:

When the user gives the connect request to the server , the process to be done is given in the following flowchart.

```
                    ┌─────────────────────┐
                    │   Connect request   │
                    └──────────┬──────────┘
                               │
                               ▼
              ┌────────────────────────────────┐
              │  Check whether the user information │
              │  is available in the server database │
              └────────────────┬───────────────┘
                               │
                               ▼
                        ◇                      No    ┌──────────────────────┐
                    Is Avail ? ──────────────────────▶│ The cluster cannot join │
                        ◇                              │ the process            │
                               │                       └───────────┬──────────┘
                             Yes                                    │
                               ▼                                    ▼
         ┌──────────────────────────────────┐              ╭──────────╮
         │  The cluster machine Looks the    │              │   STOP   │
         │  table CLUSTER_DETAILS and        │              ╰──────────╯
         │  sets the field STATUS as "on" for │
         │  the particular CLUSTER_ID.       │
         └──────────────────┬───────────────┘
                            │
                            ▼
         ┌──────────────────────────────────┐
         │  Check whether the database_update │
         │  field is true or false.          │
         └──────────────────┬───────────────┘
                            │
                            ▼
                        ◇                  YES      ( B )
                    Db_upate ? ─────────────────────
                        ◇
                            │
                          No
                            ▼
                          ( A )
```

24

(A)

Execute the queries in the Queryfile_table
from the start query number available in
the table QUERY.

Change the field "db_update" of
CLUSTER_DETAILS to "true".

(B)

Wait for the server
machine to submit the job

LEAVING THE CLUSTER:

Disconnect
request

In the server locate the
CLUSTER_DETAIL table and then
change the field of STATUS to
false

Close the cluster application

STOP

DISTRIBUTING THE RECORDS:

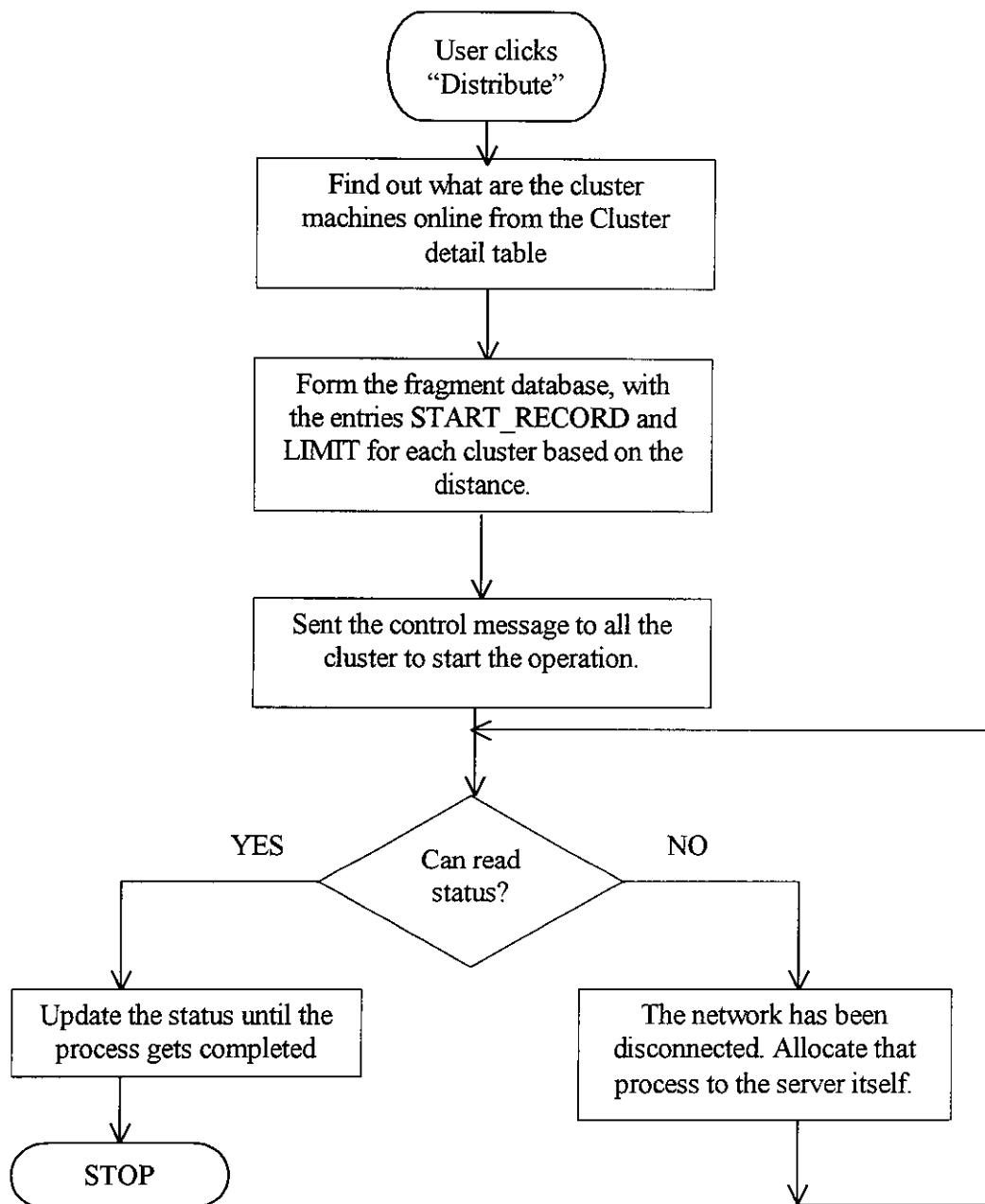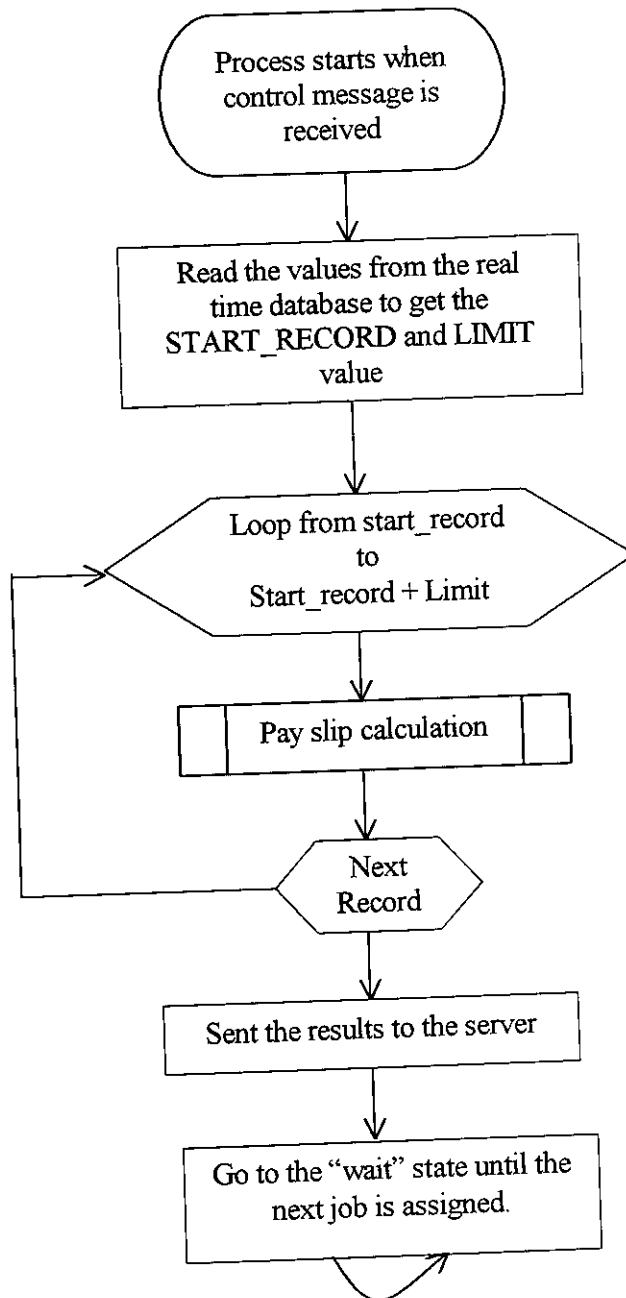This is process of distributing the records among the various clusters online by splitting the whole table and the split contents in the Fragment table, whose structure is discussed in the database design.

CLUSTER MACHINE PROCESS:

When the cluster machine receives the control message from the server, it starts the processing, completes it and then sends the result to server.

```
          ┌─────────────────────┐
         ( Process starts when   )
         ( control message is    )
         ( received              )
          └──────────┬──────────┘
                     │
                     ▼
          ┌─────────────────────┐
          │ Read the values from │
          │ the real time        │
          │ database to get the  │
          │ START_RECORD and     │
          │ LIMIT value          │
          └──────────┬──────────┘
                     │
                     ▼
          ╱─────────────────────╲
         ╱ Loop from start_record ╲
         ╲ to                     ╱
          ╲ Start_record + Limit ╱
           ╲───────────┬───────╱
                       │
                       ▼
          ┌─┬───────────────────┬─┐
          │ │ Pay slip calculation│ │
          └─┴───────────────────┴─┘
                       │
                       ▼
                ╱──────────╲
               ╱   Next     ╲
               ╲   Record   ╱
                ╲──────────╱
                       │
                       ▼
          ┌─────────────────────┐
          │ Sent the results to  │
          │ the server           │
          └──────────┬──────────┘
                     │
                     ▼
          ┌─────────────────────┐
          │ Go to the "wait"     │
          │ state until the      │
          │ next job is assigned. │
          └─────────────────────┘
```

27

SYSTEM IMPLEMENTATION & TESTING

5.1 SYSTEM IMPLEMENTATION:

Implementation is the stage of the project when the design is turned into a working system. At the stage the main workload, the greatest upheaval and the major impact on existing practices shift to the user department. If the implementation is not carefully planned and controlled, it can cause chaos. Thus it can be considered to be the most crucial stage in achieving a successful new system and in giving the users confidence that the new system will work and be effective.

Implementation involves careful planning, investing of the current system and its constraints on the implementation, design of the procedures to achieve the changeover, training the staff to use this system. The main task of the implementation is the implement planning and the training of the user.

Implementation Planning:

The implementation of this system involves people from different department. We have to plan carefully for accessing the respective staff members. This involves meeting the respective staff members only with the permission of their managers and without disturbing their usual routines. And our implementation process should not disturb or collapse the existing system. Once this planning is over, the major effort is to ensure that the project is working properly.

Implementing The System :

When the implementation planning is completed, our first step is to convert the existing files of the existing system, to the format required by our system. This step requires very careful attention; we need control procedures to implement this step. When the files are converted, we can now use the new system.

User Training :

To achieve the benefits of the new system, a good training must be given to the person who is going to use the system. The help file created for both the server and the cluster can be used for this training purpose. This makes the user to understand the complete system and so he can carry on his work effectively.

5.2 SYSTEM TESTING:

System testing is the stage of implementation, which is aimed at ensuring that the system works accurately and efficiently before live operation commences. Thus the system test in implementation should be a confirmation that all is correct and an opportunity to show the users that the system works.

Black Box Testing:

This testing focuses on the functional requirements of the software. It finds the errors of Incorrect or missing functions, interface errors, errors in database access, performance errors, and initialization errors.

White Box Testing:

Using the white box testing all the logical decisions on the true and false side of the product is tested. All the loops within the operational bunds are checked. Logical errors and incorrect assumptions are identified and rectified.

Unit Testing:

Unit testing is done at the development phase itself. When ever a unit of code is written, it is tested with the test cases during the development itself, so that it will be easy for us during the other testing done at the implement phase. The following table states the test done on the form EMPLOYEE DETAIL.

| Sno	Test case	Expected output	Actual result	Remark
1	View data (Employee Number)	To list the detail of the employee	Employee details	Correct
2	add data (Employee details)	To add a employee in to database	Data is added to the table	Correct
3	delete data (Employee details)	To delete a employee from table	Data is deleted from table	Correct

Path Testing:

This technique is used to confirm test effectiveness, based on control topology of the code. This technique demands knowledge of the internal structure of the code, as such; it is a form of white-box testing.

Performance Testing:

This technique is used to test the system performance, to check the response time to the user from the system. This involves the old test to be rerun, and is used to demonstrate that the newly corrected system still perform everything the old one did.

The following chart gives the response time for the user, with the single system and also says how quick is the response time when number of clusters adds in to the system.

Response time

0
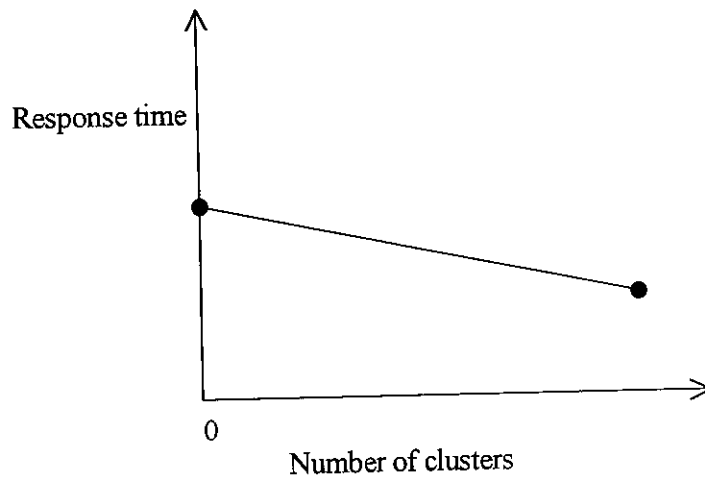
Number of clusters

When we analyze the graph we can know that, when there is no clusters involved in processing, the response from the system is very slow and, as number of clusters adds in to the system we can see the response time is coming down, which means we are getting the response from the system very quick.

The following table gives the approximate values collected during the testing process of the system. The testing is done with the following configuration.

- Pentium 200 MHz
- 32 Mb RAM
- 2 GB hard disk

Number of systems involved in the testing is three. One machine acts as a server and the other two machines as clusters. I have kept the number of records as the constant that is 1000 records.

Number of clusters online	Response time in sec
Zero	20
One	14
Two	8

Stress Testing:

By this testing we overload the system by giving high value at extreme conditions, and we see how the system reacts. Here we will know the limitations of our software. Here we have the following limitations

- Does not support more than 64 clusters

- Data transferred between the server and the cluster not to exceed two billion character (which is the size of variable length string)

The error found during the stress testing was,

- Data loss occurs when many clusters uploads the results to the server after completing the process.

This error was corrected by using Two-way handshaking method that is by sending the acknowledgement between the cluster and the server, whenever the uploading takes place from the cluster side.

5.3 REFINEMENT BASED ON FEEDBACK:

Based on the feedback given by the users, the user interactivity has been improved by assigning keyboard shortcuts to the operations available. Also the status of the client machine like online, offline and also the operations performed on the cluster machine is reported back to the server machine, so that user can know what is happening in the cluster machine.

6. <u>CONCLUSION</u>:

All the Objectives of the proposed system were achieved successfully. The distributed framework has been created. All the functions have been written in a DLL file. So any one who wishes to develop a distributed system with no sharing clusters can use this function available in the DLL. He just wants to know about the definition of the various functions available in the DLL, and he does not want to take care about the actual implementation of these functions. The project mainly focuses on the Database applications, which can be distributed.

The payroll application developed to explain this concept, has achieved the goals of,

- Distributed environment
- Highly user friendly and interactive
- Faster response
- Portable and flexible for future development

Thus this project has achieved all its objectives, which was listed in the proposed system. This DISTRIBUTED FRAMEWORK will be helpful for Lucid technologies and also for individuals who are interested to develop distributed computing applications.

7.0 SCOPE FOR FUTURE DEVELOPMENT:

This project is developed only for the No Sharing Clusters, and hence there is no communication between the clusters. There is only communication between the server and the clusters. In the future, this product can be enhanced to work with sharing clusters.

Although the project title is "Framework for Distributed computing", this product cannot be used for all distributed computing applications. This can be used only for the database applications to share the load of the server. In future, this product can be enhanced to work with other type of applications also.

This product cannot be used for large applications, since this product supports maximum of only 64 (maximum supported by Winsock API of windows) clusters. In future this limitation can be eliminated and can be used for the commercial purpose.

Even though this product can be used with communication media as Internet, it does not provide much security to the user. High level of security like encryption can be added to the system in the future.

BIBLIOGRAPHY

1. Book Name : **Winsock 2.0**

   Author   : Lewis Napper

   Publisher   : Comdex Computer Publishing

2. Book Name : **Windows NT Clustering**

   Author   : Mark. A. Sportack

   Publisher : SAMS

3. www.itpapers.com as on December 20 .

4. Msdn.microsoft.com as on From February 10 to March 30.

APPENDICES