

Assignment No :A4

1 Title:

Cluster time.

2 Problem Definition

Write a program on an unloaded cluster for several different numbers of nodes and record the time taken in each case. Draw a graph of execution time against the number of nodes.

3 Learning Objectives

1. To understand the concept of clusters.
2. To understand the latency issue in cluster computing.
3. To be able to perform runtime analysis of a cluster computing system.

4 Learning Outcomes

1. Ability to form unloaded clusters and perform required computations on them.
2. Understanding of unloaded cluster computing.

5 Related Mathematics

Let S be the solution perspective of the given problem.

The set S is defined as:

$$S = \{ s, e, X, Y, F, DD, NDD, S_c, F_c | \emptyset_s \}$$

Where,

s= Start state, Such that $Y = \{\emptyset\}$

e= End state

X= Input Set.

$X = \{ \text{seq}(x) \mid x \in \text{Natural numbers} \}$

Y=Output set.

$Y = \{ \text{UnloadedClusterComputation} \}$

F= Set of functions used.

$F = \{ \text{master}(), \text{slave}(), \text{compute}(), \text{timeAnalyse}() \}$

master()= function for master purposes.

slave()= function for slave purposes.

compute()= function for computing the coverage of the input sequence.

timeAnalyse()= function for performing the time analysis of the cluster computation.

DD=Deterministic data.

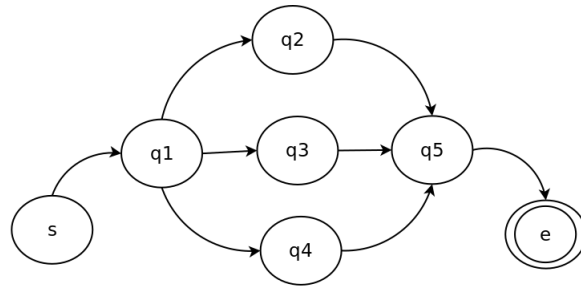
DD=

1. sequence follows proper constraints.
2. coverage is defined and exists.
3. sequence terminates.

NDD= Non-deterministic data.

NDD= \cup - DD

6 State Transition diagram



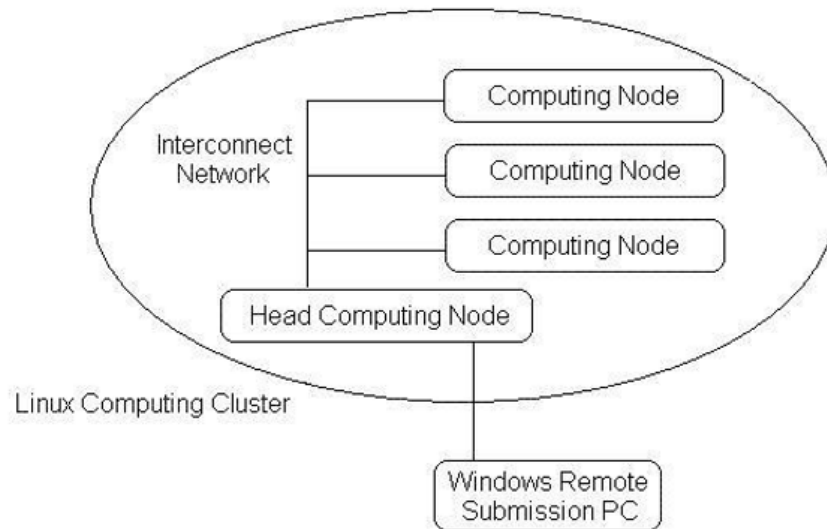
s = start state
q1 = master thread
q2 = slave thread
q3 = slave thread
q4 = slave thread
q5 = Perform shell and time analysis
e = end state

7 Concepts related theory

7.1 Clustering:

A computer cluster consists of a set of loosely or tightly connected computers that work together so that, in many respects, they can be viewed as a single system. Unlike grid computers, computer clusters have each node set to perform the same task, controlled and scheduled by software.

The components of a cluster are usually connected to each other through fast local area networks ("LAN"), with each node (computer used as a server) running its own instance of an operating system. In most circumstances, all of the nodes use the same hardware and the same operating system, although in some setups (i.e. using Open Source Cluster Application Resources (OSCAR)), different operating systems can be used on each computer, and/or different hardware.



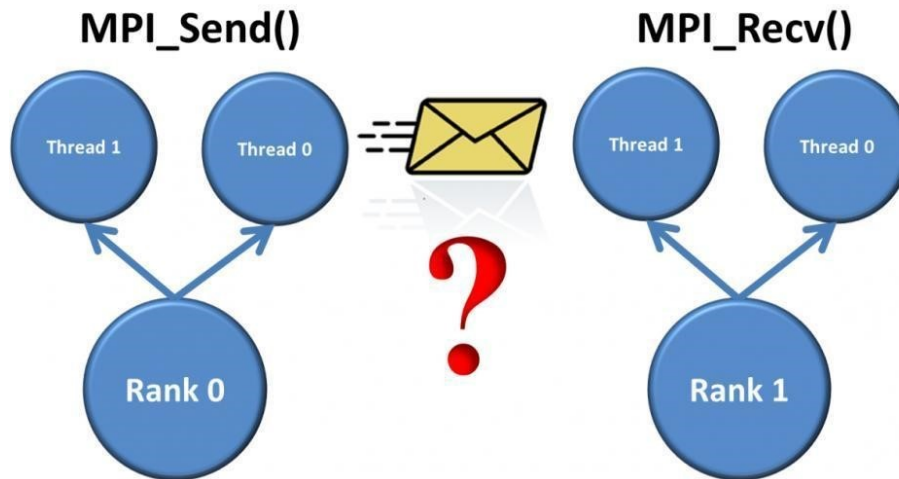
They are usually deployed to improve performance and availability over that of a single computer, while typically being much more cost-effective than single computers of comparable speed or availability.

Computer clusters emerged as a result of convergence of a number of computing trends including the availability of low-cost microprocessors, high speed networks, and software for high-performance distributed computing. They have a wide range of applicability and deployment, ranging from small business clusters with a handful of nodes to some of the fastest supercomputers in the world such as IBM's Sequoia. The applications that can be done however, are nonetheless limited, since the software needs to be purpose-built per task. It is hence not possible to use computer clusters for casual computing tasks.

A computer cluster may be a simple two-node system which just connects two personal computers, or may be a very fast supercomputer. A basic approach to building a cluster is that of a Beowulf cluster which may be built with a few personal computers to produce a cost-effective alternative to traditional high performance computing. An early project that showed the viability of the concept was the 133-node Stone Supercomputer. The developers used Linux, the Parallel Virtual Machine toolkit and the Message Passing Interface library to achieve high performance at a relatively low cost.

7.2 MPI:

MPI is a language-independent communications protocol used for programming parallel computers. Both point-to-point and collective communication are supported. MPI "is a message-passing application programmer interface, together with protocol and semantic specifications for how its features must behave in any implementation." MPI's goals are high performance, scalability, and portability. MPI remains the dominant model used in highperformance computing today.



MPI is not sanctioned by any major standards body; nevertheless, it has become a de facto standard for communication among processes that model a parallel program running on a distributed memory system. Actual distributed memory supercomputers such as computer clusters often run such programs. The principal MPI-1 model has no shared memory concept, and MPI-2 has only a limited distributed shared memory concept. Nonetheless, MPI programs are regularly run on shared memory computers. Designing programs around the MPI model (contrary to explicit shared memory models) has advantages over NUMA architectures since MPI encourages memory locality.

Although MPI belongs in layers 5 and higher of the OSI Reference Model, implementations may cover most layers, with sockets and Transmission Control Protocol (TCP) used in the transport layer.

Most MPI implementations consist of a specific set of routines (i.e., an API) directly callable from C, C++, Fortran and any language able to interface with such libraries, including C#, Java or Python. The advantages of MPI over older message passing libraries are portability (because MPI has been implemented for almost every distributed memory architecture) and speed (because each implementation is in principle optimized for the hardware on which it runs).

MPI uses Language Independent Specifications (LIS) for calls and language bindings. The first MPI standard specified ANSI C and Fortran-77 bindings together with the LIS. The draft was presented at Supercomputing 1994 (November 1994) and finalized soon thereafter. About 128 functions constitute the MPI-1.3 standard which was released as the final end of the MPI-1 series in 2008.

8 Program Listing

```
#include <stdio.h>
#include <stdlib.h>
#include <mpi.h>
#define v 1 /* verbose flag, output if 1, no output if 0 */
#define tag 100 /* tag for sending a number */
int main ( int argc, char *argv[] )
{
    int p, myid, i, f, *x;
    double start, end;
    MPI_Status status;
    MPI_Init(&argc, &argv);
    MPI_Comm_size(MPLCOMM_WORLD, &p);
    MPI_Comm_rank(MPLCOMM_WORLD, &myid);
    if(myid == 0) /* the manager allocates and initializes x */
    {
        x = (int*)calloc(p, sizeof(int));
        x[0] = 50;
        for (i=1; i<p; i++) x[i] = 2*x[i-1];
        if(v>0)
        {
            printf("The data to square : ");
            for (i=0; i<p; i++)
                printf(" %d", x[i]);
            printf("\n");
        }
    }
    if(myid == 0) /* the manager copies x[0] to f */
    {
        /* and sends the i-th element to the i-th processor */
        f = x[0];
        for(i=1; i<p; i++)
            MPI_Send(&x[i], 1, MPI_INT, i, tag, MPLCOMM_WORLD);
    }
}
```

```

}
else /* every worker receives its f from root */
{
MPI_Recv(&f,1,MPI_INT,0,tag,MPLCOMM_WORLD,&status);
if(v>0)
printf("Node %d will square %d\n",myid,f);
}
start = MPI_Wtime();
f *= f; /* every node does the squaring */
if(myid == 0) /* the manager receives f in x[i] from processor i */
for(i=1; i<p; i++)
MPI_Recv(&x[i],1,MPI_INT,i,tag,MPLCOMM_WORLD,&status);
else /* every worker sends f to the manager */
MPI_Send(&f,1,MPI_INT,0,tag,MPLCOMM_WORLD);
if(myid == 0) /* the manager prints results */
{
x[0] = f;
printf("The squared numbers : ");
for(i=0; i<p; i++)
printf(" %d",x[i]);
printf("\n");
end = MPI_Wtime();
printf("Runtime = %f\n", end-start);
}

MPI_Finalize();
return 0;
}

```

9 Output

```

pict@localhost:~/Documents/final/a4/code$ mpirun -np 3 ./abc
The data to square : 50 100 200
Node 2 will square 200
Node 1 will square 100
The squared numbers : 2500 10000 40000
Runtime = 0.000119
pict@localhost:~/Documents/final/a4/code$ mpirun -np 1 ./abc
The data to square : 50
The squared numbers : 2500
Runtime = 0.000014
pict@localhost:~/Documents/final/a4/code$ mpirun -np 5 ./abc
The data to square : 50 100 200 400 800
Node 2 will square 200
Node 3 will square 400
Node 4 will square 800
Node 1 will square 100
The squared numbers : 2500 10000 40000 160000 640000
Runtime = 0.000256
pict@localhost:~/Documents/final/a4/code$ mpirun -np 6 ./abc

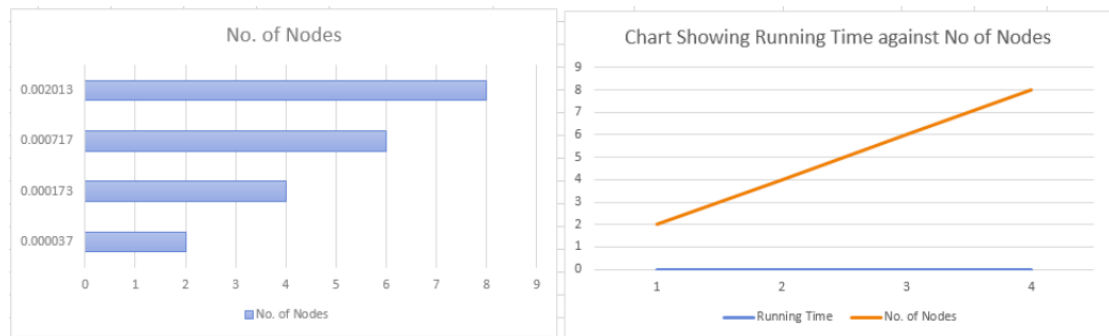
```

```

The data to square : 50 100 200 400 800 1600
Node 2 will square 200
Node 3 will square 400
Node 5 will square 1600
Node 1 will square 100
Node 4 will square 800
The squared numbers : 2500 10000 40000 160000 640000 2560000
Runtime = 0.000348
pict@localhost:~/Documents/final/a4/code$ ^C

```

9.1 Graph



10 TESTING :

10.1 Positive Testing

Sr. No.	Test Condition	Steps to be executed	Expected Result	Actual Result
1.	Enter the number of process	Press Enter	Time taken by processes to execute task	Same as Expected

10.2 Negative Testing

Sr. No.	Test Condition	Steps to be executed	Expected Result	Actual Result
1.	Entered input data is in character or symbol	Press enter	Error message	Display result

11 CONCLUSION :

Thus we studied concept of clustering, which consists of a set of loosely or tightly connected computers that work together and created a program on an unloaded cluster for several different numbers of nodes and record the time taken in each case.