The Parallel Architecture Approach, Single Program Multiple Data (Spmd) Implementation on Clusters of Terminals Using Java RMI

Sudhir Kumar Meesala\(^1\), Dr. Pabitra Mohan Khilar\(^2\), Dr. A. K. Shrivistava\(^3\)

\(^1\)Ph.D. Scholar, Dept. Of Computer Science & Engineering, Dr. C.V. Raman University, Kota, Bilaspur(Cg), India

\(^2\)Assistant Professor, Department Of Computer Science And Engineering, Nit, Rourkela(Orrissa), India

\(^3\)Professor & Head, Department Of Physics, Dr. C. V. Raman University, Kota, Bilaspur(Cg), India

Abstract


Keywords: Single Program Multiple Data(SPMD), Remote Method Invocation(RMI), Parallel Processing, Distributed Technology, Molecular Biology, Weather Forecast, Climate Modeling, Circuit Design, Biometric, Re-Engineering, Recycling Engineering

I. Introduction

In Computer Calculations, SPMD (Single Program, Multiple Data) Is A Technique Employed To Achieve Parallelism; It Is A Subcategory Or One Of The Basic Type Of MIMD. Tasks Are Split Up And Run Simultaneously On Multiple Processors With Different Input In Order To Obtain Results Faster. SPMD Is The Most Common Style Of Parallel Programming.[1] It Is Also A Essential For Research Concepts Such As Active Messages And Distributed Shared Memory.

Reference Between SPMD & SIMD

In SPMD, Multiple Autonomous Processors Simultaneously Execute The Same Program At Independent Points, Rather Than In The Lockstep That SIMD Imposes On Different Data. With SPMD, Tasks Can Be Executed On General Purpose Cpus; SIMD Requires Vector Processors To Manipulate Data Streams. Note That The Two Are Not Mutually Exclusive.

Concept Of Distributed Memory

Starts Its Own Program And Communicates With Other Nodes By Sending And Receiving Messages, Calling Send/Receive Routines For That Purpose. Barrier Synchronization May Also Be Implemented By Messages. The Messages Can Be Sent By A Number Of Communication Mechanisms, Such As TCP/IP Over Ethernet, Or Specialized High-Speed Interconnects Such As Myrinet And Supercomputer Interconnect. Serial Sections Of The Program Are Implemented By Identical Computation On All Nodes Rather Than Computing The Result On One Node And Sending It To The Others. Nowadays, The Programmer Is Isolated From The Details Of The Message Passing By Standard Interfaces, Such As PVM And MPI. Distributed Memory Is The Programming Style Used On Parallel Supercomputers From Homegrown Beowulf Clusters To The Largest Clusters On The Teragrid.

Concept Of Shared Memory

Combination Of Levels Of Parallelism
Current Computers Allow Exploiting Of Many Parallel Modes At The Same Time For Maximum Combined Effect. A Distributed Memory Program Using MPI May Run On A Collection Of Nodes. Each Node May Be A Shared Memory Computer And Execute In Parallel On Multiple Cpus Using Openmp. Within Each CPU, SIMD Vector Instructions (Usually Generated Automatically By The Compiler) And Superscalar Instruction Execution (Usually Handled Transparently By The CPU Itself), Such As Pipelining And The Use Of Multiple Parallel Functional Units, Are Used For Maximum Single CPU Speed.

SPMD History

Parallel & Distributed Architecture
Distributed Computing Is Method Of Computer Processing In Which Different Parts Of A Program Run Simultaneously On Two Or More Computers That Are Communicating With Each Other Over A Network. Distributed Computing Is A Type Of Parallel Computing.[7] But The Latter Term Is Most Commonly Used To Refer To Processing In Which Different Parts Of A Program Run Simultaneously On Two Or More Processor That Are Part Of The Same Computer. While Both Types Of Processing Require That A Program Be Parallelized - Divided Into Sections That Can Run Simultaneously, Distributed Computing Also Requires That The Division Of The Program Take Into Account The Different Environments On Which The Different Sections Of The Program Will Be Running. For Example, Two Computers Are Likely To Have Different File Systems And Different Hardware Components.[3] Distributed Computing Is A Natural Result Of The Use Of Network To Allow Computers To Efficiently Communicate. But Distributed Computing Is Distinct From Networking. The Latter Refers To Two Or More Computers Interacting With Each Other, But Not, Typically, Sharing The Processing Of A Single Program. The World Wide Web Is An Example Of A Network, But Not An Example Of Distributed Computing.[14] There Are Numerous Technologies And Standards Used To Construct Distributed Computations, Including Some Which Are Specially Designed And Optimize For That Purpose, Such As Remote Procedure Calls (RPC), Remote Method Invocation (RMI) Or Net Remoting.[5]
Organizing The Interaction Between Each Computer Is Of Prime Importance. In Order To Be Able To Use The Widest Possible Range And Types Of Computers, The Protocol Or Communication Channel Should Not Contain Or Use Any Information That May Not Be Understood By Certain Machines. Special Care Must Also Be Taken That Messages Are Indeed Delivered Correctly And That Invalid Messages Are Rejected Which Would Otherwise Bring Down The System And Perhaps The Rest Of The Network.

Various Hardware And Software Architectures Are Used For Distributed Computing. At A Lower Level, It Is Necessary To Interconnect Multiple CPUs With Some Sort Of Network, Regardless Of Whether That Network Is Printed Onto A Circuit Board Or Made Up Of Loosely Coupled Devices And Cables. At A Higher Level, It Is Necessary To Interconnect Processes Running On Those CPUs With Some Sort Of Communication System. Distributed Programming Typically Falls Into One Of Several Basic Architecture Or Categories:

- **Client-Server**
- **3-Tier Architecture**
- **N-Tier Architecture**
- **Distributed Objects**
- **Loose Coupling Or Tight Coupling.**

1) **Client-Server**— Smart Client Code The Server For Data, Then Formats And Displays It To The User. Input At The Client Is Committed Back To The Server When It Represents A Permanent Change.
2) **3-Tier Architecture**— Three Tier Systems Move The Client Intelligence To A Middle Tier So That Stateless Clients Can Be Used. This Simplifies Application Deployment. Most Web Applications Are 3-Tier.
3) **N-Tier Architecture**— N-Tier Refers Typically To Web Application Which Further Forward Their Request To Other Enterprise Services. This Type Of Application Is The One Most Responsible For The Success Of Application Servers.

1. **Tightly Coupled (Clustered)**— Refers Typically To A Set Of Highly Integrated Machines That Run The Same Process In Parallel, Subdividing The Task In Part That Are Made Individually By Each One, And Then Put Back Together To Make The Final Result.
2. **Peer-To-Peer**— Architecture Where There Is No Special Machine Of Machines That Provide A Service Or Manage The Network Resources. Instead All Responsibilities Are Uniformly Divided Among All Machines, Known As Peers.


The Types Of Distributed Systems Are Based On Flynn's Taxonomy Of Systems:

- **Single Instruction Single Data (SISD)**
- **Single Instruction Multiple Data (SIMD)**
- **Multiple Instruction Single Data (MISD)**
- **Multiple Instruction Multiple Data (MIMD)**
- **Single Program Multiple Data (SPMD)**

We Are Implementing Client-Server Architecture And Single Program Multiple Date (SPMD) Taxonomy.

**Remote Method Invocation (RMI)**


**A. The Stub/Skeleton Layer**

This Layer Intercepts Method Calls Made By The Client To The Interface Reference And Redirects These Calls To A Remote Object. Stubs Are Specific To The Client Side, Whereas Skeletons Are Found On The Server Side. To Achieve Location Transparency, RMI Introduces Two Special Kinds Of Objects Known As Stubs And Skeletons That Serve As An Interface Between An Application And Rest Of The RMI System. This Layer’s Purpose Is To Transfer Data To The Remote Reference Layer Via Marshalling And Unmarshalling. Marshalling Refers To The Process Of Converting The Data Or Object Being Transferred Into A Byte Stream And
Unmarshalling is the reverse – converting the stream into an object or data. This conversion is achieved via object serialization.

The stub/skeleton layer of the RMI lies just below the actual application and is based on the proxy design pattern. In the RMI use of the proxy pattern, the stub class plays the role of the proxy for the remote service implementation. The skeleton is a helper class that is generated by RMI to help the object communicate with the stub; it reads the parameters for the method call from the link, makes the call to the remote service implementation object, accepts the return value and then writes the return value back to the stub. In short, the proxy pattern forces method calls to occur through a proxy that acts as a surrogate, delegating all calls to the actual object in a manner transparent to the original caller.

**Stub**

The stub is a client-side object that represents (or acts as a proxy for) the remote object. The stub has the same interface, or list of methods, as the remote object. However, when the client calls a stub method, the stub forwards the request via the RMI infrastructure to the remote object (via the skeleton), which actually executes it.

**Sequence of Events Performed by the Stub:**

1. Initiates a connection with the remote VM containing the remote object.
2. Marshals (writes and transmits) the parameters to the remote.
3. VM waits for the result of the method invocation.
4. Unmarshals (reads) the return value or exception returned.
5. Return the value to the caller.

In the remote VM, each remote object may have a corresponding skeleton [16].

**Skeleton**

On the server side, the skeleton object takes care of all the details of “remoteness” so that the actual remote object does not need to worry about them. In other words, we can pretty much code a remote object the same way as if it were local; the skeleton insulates the remote object from the RMI infrastructure.

**Sequence of Events Performed by the Skeleton**

- Unmarshals (reads) the parameters for the remote method (remember that these were marshaled by the stub on the client side).
- Invokes the method on the actual remote object implementation.
- Marshals (writes and transmits) the result (return value or exception) to the caller (which is then unmarshalled by the stub).

The diagram shows the RMI architecture (Fig 3 & 4).

**B. The Remote Reference Layer:** The remote reference layer defines and supports the invocation semantics of the RMI connection. This layer maintains the session during the method call.
C. The Transport Layer

Single Program Multiple Data (SPMD) Algorithm
We Are Implementing Remote Method Invocation From JAVA Language As Platform To Apply Parallel Processing Concept Single Program Stream Multiple Data Stream (SIMD) In Distributed Network; Here We Are Using Client/Server Architecture. Server Is The Class Where The Distribution Process Occurs. We Are Having A Set Of Randomly Generated Numbers. Here As We Have Single Client We Retrieve Three Numbers From Client And Give Them To Server For Factorial Calculation And Summation. Client Has Job Of Distribution Of Numbers. There Can Be Many Servers And They Can Have Different Methods, Which Can Be Applied Concurrently, Result Will Be Returned To Client For Further Operations. As Many Server Are Present In This Application, So, We Have To Implement Thread To Bring Access Of Server To One Server At One Time. This Will Not Cause Corruption Of Data And Thus The Work Produce Satisfactorily Results.

RMI Is A Simple Method Used For Developing And Deploying Distributed Object Application In A Java Environment. Creating Distributed Object Application Using RMI Is A Simple As Writing A Stand-Alone Java Application. RMI Enables A Programmer To Create Distributed Java Application, In Which The Methods Of Remote Java Object Can Be Called From Other Java Virtual Machines Running Either On The Same Host Or On Different Hosts Scattered Across A Network.

A Call To Remote Object Using RMI Is Identical To A Call Made To A Local Object With The Following Exceptions:
1. An Object Passed As A Parameter To A Remote Method Or Returned From The Method Must Be Serialization Or Be Another Remote Object.
2. An Object Passed As A Parameter To A Remote Method Or Returned From The Method Called Is Passed By Value And Not By Reference.


In Any Distributed Application, For The Client Side Of The Application To Make The Call To Remote Object, That Client Object Would First Be Able To Locate The Remote Object RMI Provide The Registry Services N Or The Name Services To Make This Possible.

We Register Any Remote Object That It Is Exporting With A Name Server Called A Registry. We Can Maintain A Registry Server That Is Running On A Well-Known Pre Defined Port Number. An Application Can Register With The Registry If It Is On Same Physical Machine.

Steps For Creating RMI Applications:
- Define An Interface Of The Remote Classes.
- Implement The Interface In Server-Side Application.
- Bind Objects To Registry Service.
- Create Stubs And Skeleton Classes.
- Create And Compile Client Program To Access The Remote Objects.
- Install Files On Client And Server Machines.
- Start The RMI Registry
Steps Involved In Running The RMI Application:
In case The Server Application And Client Application Is Run In The Same Machine:
- Run The RMI Registry At Specified Port, If Not Specified, It Runs At The Default Port 1099.
- Run The Server Application In Another DOS Window.
- Run The Client Application From The Same Machine.
In case The Server Application And Client Application Is Run On The Separate Machine:
- Run The RMI Registry At Specified Port, If Not Specified, It Runs At The Default Port 1099.
- Run The Server Application In Another DOS Window.
- Run The Client Application From A Separate Machine.
Following These Steps RMI Application Can Be Implemented.

Algorithm for Developing and Running the RMI Application for Distributed System.
Step 1: Enter And Compile The Source Code
Step 2: Generate Stubs And Skeletons
Step 3: Install Files On The Client And Server Machines.
Copy Addclient.Class, Addserverimpl_Stub, And Addserverintf. Class To A Directory On The Client Machine.
Step 4: Start The RMI Registry On The Server Machine.Start Rmiregistry
Step 5: Start The Server
Java Addserver
Step 6: Start The Client
For Calculating Serially (Run At Each And Individual Machine). The Addclient Software Requires Four Arguments: The Name Or IP Address Of The Servermachine And The Three Numbers That Are To Be Summed Together Of First Two Number And Factorial Of Third Number. You May Invoke It From The Command Line By Using One Of The Two Formats Shown Here. (Ex: Java Addclient 172.16.16.14 458 475 5 Or Java Addclient Server1 485 475 5) For Calculating Parallel (Run All At Same Time) The Addclient Software Requires Arguments: Three Numbers That Are To Be Summed Together Of First Two Number And Factorial Of Third Number. In This Process We Never Use The IP Address Because We Already Use All IP Address In Addclient.Java. You May Invoke It From The Command Line By Using One Of The Two Formats Shown Here Ex: Java Addclient 458 475 5).

II. Results and Conclusion
We Successes In Implementing Remote Method Invocation From JAVA Language As A Platform To Apply Single Program Stream Multiple Data Stream(SPMD) On Clusters Of Terminal’s (COT’s). Here We Are Using Client/Server Architecture. Server Is The Class Where The Distribution Process Occurs. We Are Having A Set Of Randomly Generated Numbers. Here As We Have Single Client We Retrieve Nth Tasks And Give Them To Nth Server For Various Complex Calculations. Client Has Job Of Distribution Of Numbers. There Can Be Many Nodes As A Servers And They Can Have Different Methods, Which Can Be Applied Concurrently, Result Will Be Returned To Client For Further Operations.
To Estimate The Performance Of The Distributed System The Time For The Computation Of The Task Solved By Different Servers Has To Be Measured In The Sequential And Parallel Case. The Result Was Examined Only Within The Area 1 To 20 Terminal’s (See Table1)

Table I: Serial And Parallel Time In Seconds On Number Of Terminal

<table>
<thead>
<tr>
<th>No Of Terminals</th>
<th>Serial Time( Sec.)</th>
<th>Parallel Time( Sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3.856</td>
<td>2.382</td>
</tr>
<tr>
<td>4</td>
<td>5.785</td>
<td>2.437</td>
</tr>
<tr>
<td>6</td>
<td>7.754</td>
<td>2.492</td>
</tr>
<tr>
<td>8</td>
<td>9.723</td>
<td>2.579</td>
</tr>
<tr>
<td>10</td>
<td>11.842</td>
<td>2.592</td>
</tr>
<tr>
<td>12</td>
<td>13.142</td>
<td>2.719</td>
</tr>
<tr>
<td>14</td>
<td>15.752</td>
<td>2.827</td>
</tr>
<tr>
<td>16</td>
<td>17.625</td>
<td>2.843</td>
</tr>
<tr>
<td>18</td>
<td>19.679</td>
<td>2.906</td>
</tr>
</tbody>
</table>
Following Charts Are Representing The Graphically Performance Of Serial And Parallel Distribution And Compare Between Both Process (See Fig 3,4,5 And 6)

III. Conclusion
An Advantage Of Using Parallel Processing Instead Of Serial Processing Is Low Cost, High Efficiency Resulting From Use Of Multiprocessing Technique. Using Parallel Processing With Distributed Network Provides Additional Advantage Of Flexibility And Speed Up In Complex Calculations. Using RMI We Can Further Enhance The Application By Performing File Transfer Remotely. We Can Also Use The Output Given By The Server To Client For Further Calculations

![Fig 5: Comparison Chart Of Parallel And Serial Time Using MIMD](image)

![Fig 6: Comparison Chart Of Parallel And Serial Time Using MIMD](image)

IV. Acknowledgement
My Express Thanks And Gratitude To All The Departments’ Personals And Sponsors Who Give Me A Opportunity To Present And Express My Paper On This Level. I Wish To Place On My Record My Deep Sense Of Gratitude To All Reference Papers Authors For Them Valuable Help Through Their Papers, Books, Websites Etc.

Author’s Profile
Sudhir Kumar Meesala Has Received His Master Of Technology Degree In Computer Technology From National Institute Of Technology, Raipur(CG) The Year 2007. At Present He Is Pursuing Ph.D. With The Specialization Of Computer Science And Engineering College. His Area Of Interest Parallel Processing, Distributed Technology, Compiler Design, Image Processing, Operating System, Network Programming And Structured Computer Engineering Etc.

Dr. Pabitra Mohan Khilar Has Received His Ph.D. In Computer Science And Engineering From IIT Kharagpur(WB) India In The Year 2009. At Present He Is Working As An Associate Professor At Department Of Computer Science And Engineering, National Institute Of Technology, Rourkela(Orissa). His Areas Of Interests Are Parallel And Distributed Computing, Cloud And Grid Computing, Distributed Wireless Network, Distributed Embedded Network Etc.
Dr. A. K. Shrivastava has received his Ph.D. in Physics (Microwave Propagation) from B. R. Ambedkar Bihar University, Muzafferpur (Bihar) in the Year 2003. At present he is working as an Professor and Head at Dr. C. V. Raman University, Kota, Bilaspur(CG) in the Department of Physics. His areas of interests are power electronics, electrical drives, power systems, renewable energy sources and custom power devices.

References


