Modeling Considerations in BPO Multiplexing Environment

A.Q. Ansari
Department of Electrical Engineering
Jamia Millia Islamia, New Delhi, India
aqansari@ieee.org

Tapasya Patki
Department of Computer Science
University of Arizona, Tucson, AZ
tpatki@email.arizona.edu

Abstract

Outsourcing, during its establishment phase, focused on the socio-economic and the managerial concerns of the companies, countries and people involved. As a step towards the next generation outsourcing, we need to ascertain a technology framework to outsource in a secure and optimized manner with reduced reliance on litigations approach. This paper discusses issues involved in simulation and modeling studies for Business Processing Outsourcing (BPO). Multiplexing of BPO infrastructure using Customer Interface Array (CIA), is a distributed environment across various service centers, to minimize delays, maximize efficiency and improve customer satisfaction. The paper discusses a Fuzzy Logic based methodology for customer load balancing across the various service centers. This algorithm is for deployment at the level of the Inter-Dialoguing Processor (IDP), which is used as a resource for facilitating communication between the CIA and the various service centers. IDP is useful to establish context-awareness in BPO environment.

Index-Terms: BPO, Outsourcing, Parallel and Distributed processing, Fuzzy Logic, Performance Evaluation, Middleware Resource Management

1. Introduction

As we progress towards the next generation of outsourcing, the burden to ensure Quality Assurance in Business Process Outsourcing (BPO) environment is also distributed and the requirement for effective certification process needs to be revisited since the local static controls are either removed (or reduced) in the outsourcing projects. Therefore, mere reliance on QA is not felt adequate and a more comprehensive and integrated approach is needed for BPO outsourcing environments [14]. A recent outsourcing case study reports that the innovation and the quality concerns are important in outsourcing domain and application of Fuzzy Set theory is effective method for industries to evaluate the performance of outsourcing activities [3, 11]. Outsourcing companies become more dependent on outsourcers and the quality and delivery performance of the company depends totally on its outsourcers. The limitations of existing QA methodologies are arising out of the following factors:

(i) Inadequacy of distributed support for continuous quality assurance [16]
(ii) Absence of remote data collection facilities.

The need to assess a particular city-based BPO infrastructure where the outsourcing project is being executed has been established from the security point of view and to reduce legal conflicts also. Three primary parameters that can be deployed to achieve this uniformity are Customer Satisfaction Rate (CSR), Cyber Crime Index (CCI) and Civil Infrastructure Status (CIS). These parameters take into account various social, economical, cultural and geographical aspects of the areas in which the BPO service centers are located [14]. This is in addition to the negotiated and accepted Service Level Agreement (SLA). The protection of core competencies during outsourcing and the trade-off of knowledge transfer/sharing is an important consideration. In order to assess these parameters, we have to introduce an object oriented based software agent as a client software module and to that extent a provision needs to be incorporated in outsourcing contractual arrangements [12]. This calls for additional computational load at the project execution infrastructure. The probing daemon is not a intrusion detection module [15]. Some of the above concerns are addressed by resorting to Multiplexing of BPO infrastructure, to ensure smooth and simultaneous operation of an array of BPO service centers in an optimized fashion, such that the loading of these service centers is uniform and optimized. Performance modeling issues of BPO multiplexing infrastructure have not been addressed in the
literature in the past. Usually, these factors introduce (i) uncertainty in the periodicity of data availability, since the state information messages for individual nodes are exchanged at discrete intervals (and hence introduce variable latencies) and (ii) imprecision in the data quality. In order to overcome these issues, fuzzy logic based modeling approach is suggested which tolerates data imprecision.

We need to study the BPO multiplexing infrastructure from the parallel and distributed processing perspective for addressing issues like (i) Quality of Service (QoS), (ii) technological support to the existing legislative procedures and (iii) to enhance performance of professional outsourcing as a corporate level activity. Further, in order to provide insight to assess the applicability of IEEE standard 1516 for modeling and simulation high-level architecture, we examine with a broader outlook, the scope of parallel and distributed processing for BPO multiplexing infrastructure in Section 2. The rest of the paper is organized as follows. Section 3 introduces the fuzzy logic approach to load balancing in BPO infrastructure. Section 4 brings out the Inter-Dialoguing processor (IDP) related issues. Section 5 gives conclusions and suggestions for future work.

2. Parallel and Distributed Processing Considerations in BPO Environment

A typical small scale BPO infrastructure may be operational in a city in the form of a LAN or in multiple cities using intranet. In the past, BPO infrastructures were examined more as standalone infrastructural local units for their performance. However, with the next generation corporate outsourcing practices, the need to apply simulation and modeling studies for performance analysis using parallel and distributed computing techniques is on rise for multiplexing of BPO infrastructures. Such BPO multiplexing need not be restricted to a single country; but spread over multiple countries also. Since outsourcing usually involves diverse Quality-of-Service (QoS) demands involving databases, web services, multimedia and network centric applications, the computational load requirements at outsourcing infrastructure, is different than that of a scientific setup (CPU bound) or large commercial data centers (I/O bound) [2, 7, 17]. Distribution is inherent in the BPO environment with associated inevitable interactions.

From modeling perspective, we consider the BPO functions from three different viewpoints (i) Structural (ii) Behavioral and (iii) Service. While a structural view is usually adopted for construction (evolution), a behavioral view is used for modeling (analysis) and a service view is recommended for implementation (performance) [16,17]. A typical outsourcing infrastructure is organized as a cluster-computational setup, using interconnected PC's and workstations in a network environment. BPO establishment is usually a mix of CPU bound, I/O bound, network bound and memory bound applications, and their characterization from the performance viewpoint, in the presence of imprecise data, is facilitated using fuzzy logic. Further, such a computational cluster has to provide an image of a single integrated computing resource, though internally it is spread over several nodes (in a single or across multiple countries). In this context, the uncertainty in the global state is fuzzy in nature, since the actual global state in a BPO multiplexed environment cannot be measured. A distributed application in outsourcing infrastructure consists of many software components that get executed on one or multiple nodes. A proper scheduling needs to be undertaken to break an application into one or more components, to ensure uniform loading of the nodes. Load balancing involves assigning tasks to processors to complete the work in least amount of time. Simulation and modeling techniques have to assess the effectiveness of the load balancing algorithms from dynamic behavior and for scalable enhanced performance.

Scalability indicates the successful functioning of an algorithm that is independent of physical topology as well as system node size [9]. Software methodologies supporting Parallel Virtual Machine (PVM), Message Passing Interface (MPI) and distributed Shared memory (DSM), are being widely used in parallel computing domain to treat heterogeneous network of computers as a parallel machine. [5,6]. Message passing is a part of the design of MPI program and the programmer can fine tune the performance of the program by reducing the unnecessary message passing, since message passing is a significant cost for applications running on distributed systems. Unlike MPI, in the DSM systems, consistency maintenance is related with the consistency of the whole shared memory. Several unnecessary messages are required to be handled in DSM, and the programmer is helpless when designing DSM program. Since the capacities of various machines in BPO establishments are not
identical (due to heterogeneity), the usable capacities vary in the time domain according to the load from multiple users. Parallel programming in distributed computing environment using Distributed Object Migration Environment (Dome) uses automatic partitioning and distribution of objects over a heterogeneous network and also caters for load balancing and fault tolerance against network failures [4]. Dome achieves load balancing through automatic mapping of objects at run time as against conventional message passing programs in imbalanced system. Dome and ECO address issues of load balancing using processor execution rate and network topology (partitioning of networks into subnets) respectively [13]. Multithreaded, Parallel and Distributed (MPD) programming is a concurrent programming language to support parallel and distributed computing and is widely adopted in academic and University environment [1]. MPD supports two performance evaluation packages (context switching primitives and language primitives). Although, heterogeneous support is not available in MPD for supporting BPO multiplexing, the pilot study to support modeling of homogeneous clusters could be attempted.

Performance aspects of service oriented computing facilities have been examined using knowledge enhanced Bayesian network models to study response time [23]. The load balancing approaches in operating systems are unsuitable for BPO systems as little information is available about the system’s capacities about inter-task relationships [17]. The probing daemon execution (described in the earlier section) requirement arising out of quality assessment and service verification concerns, to meet data collection needs for security and related purposes like CCI, CSR and CIS, introduces a computational requirement as data acquisition overload. Inter-Dialoguing Processor (IDP), which is used as a resource for facilitating communication between the Customer Interface Array (CIA) and the various service centers, is useful to establish context-awareness in BPO environment [14, 18, 20]. In the past, a fuzzy decision tree is used to form a search mechanism for vague knowledge in Design for Outsourcing (DFO) using information obtained from outsourced maintenance, with index for classifying vague knowledge. [3]. Similarly, Fuzzy Goal Programming (FGP) has been used for outsourcer evaluation and management system for a textile company. FGP model selects the most appropriate outsourcers suitable to be strategic partners with the company and simultaneously allocates the quantities to be ordered to them [2].

3. Application of Fuzzy Logic in Customer Load Balancing for BPO Infrastructure

Load balancing decision in BPO environment needs to be examined from two separate angles viz. global perspective (advocating master or central control) and local view point. While, both these options could be supported (as summarized in Table 1), it is suggested that the BPO infrastructure operators should be given the choice of selection, either as a static practice or as an organizational culture where no fixed bias is shown for any decision methodology. In fact, it is advisable to treat it as a matter of negotiation between the agency awarding the outsourcing contract and the organization executing the work.

<table>
<thead>
<tr>
<th>Particulars</th>
<th>Local</th>
<th>Global</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Exchange</td>
<td>Master oriented approach where</td>
</tr>
<tr>
<td></td>
<td>Control data locally with</td>
<td>computational time of all tasks</td>
</tr>
<tr>
<td></td>
<td>neighbors</td>
<td>are assessed for ideal data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>distribution</td>
</tr>
<tr>
<td>Performance</td>
<td>After each load balancing phase, there is no optimal global mapping</td>
<td>Small control information exchange</td>
</tr>
<tr>
<td>Re-mapping</td>
<td>Re-mapping requires less data movement</td>
<td>Re-mapping involves large data movement</td>
</tr>
<tr>
<td>Scalable to a</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>larger population of</td>
<td></td>
</tr>
<tr>
<td></td>
<td>processors</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Load Balancing Decision in BPO Environment

BPO load balancing needs to be explored using several parameters like processor utilization,
processor queue length, memory utilization, and network traffic flow. This approach is necessary due to the multiplexing of BPO infrastructure on the geographical locations to provide and serve larger user community.

3.1 Simple Fuzzy Control Technique

Out of the four parameters indicated for BPO load balancing, we discuss the network traffic flow through a simple example of control of flow of packets on a communication link. One of the important considerations in Quality of Service (QoS) algorithm is to support consistent amount of bandwidth of the link. We should adjust the rate at which the packets are released depending upon the link utilization. The following fuzzy rule can solve the problem [8].

Fuzzy Rule: If the application utilization of the link is HIGH (LOW),
then REDUCE (INCREASE) the rate of flow of packets through the gate for the application.

In this example, we have two Fuzzy linguistic terms viz. HIGH and LOW for the only fuzzy variable Link Utilization [22]. The membership function for the Link utilization (Network Traffic) is shown in Figure 1.

Mathematically, we represent the two membership functions as follows [21].

\[
\begin{align*}
\mu_{LOW} &= 1, & \text{if rate} < 80; \\
&= (100 - \text{rate})/20, & \text{if rate} > 80 \text{ and } < 100 \\
&= 0, & \text{if rate} > 100 \\
\mu_{HIGH} &= 0, & \text{if rate} < 100; \\
&= (\text{rate} - 100)/20, & \text{if rate} > 100 \text{ and } < 120 \\
&= 1, & \text{if rate} > 120 
\end{align*}
\]

Since the base rate is to be increased (i.e. summation) if the utilization is low and is required to be reduced (i.e. subtraction), when the utilization is high, we update the new rate from the existing rate using the following formula

\[
\text{rate} = \text{rate} + (\mu_{LOW} \cdot \text{pdelta}) - (\mu_{HIGH} \cdot \text{pdelta}), \text{where pdelta is a tunable parameter (pdelta = 10, assumed)}
\]

In order to exercise the fuzzy control, we assume that at a particular time, the rate of packets is 110. Then using the above rate computation, we get the updated rate as 105. When the existing rate is 105, then the updated rate is 102.5. Likewise, if the rate is 80, then the updated rate is 90.

3.2 Fuzzy Load Balancing

In principle, the load balancing attempts to organize the total cluster system load by transferring (or commencing) processes on the idle or lightly loaded nodes, in preference to heavily loaded nodes. In order to overcome the limitations of the static load balancing approach, the dynamic load balancing uses the current system state information to improve cluster performance in the heterogeneous applications prevailing in BPO multiplexing infrastructure. For a BPO infrastructure, Load Information Vector is a function of CPU utilization, CPU queue length, network traffic, and memory utilization. Heuristic load balancing is an accepted methodology [10]. The computation of load of a particular node requires processing of fuzzy rules to infer the load using fuzzy sets representations for each of the above four parameters in the Load Information Vector. Thus, the modeling methodology is a powerful mechanism to allow individual nodes to incorporate flexible decisions in the BPO multiplexing architecture. The fuzzy sets could be on the lines of the representation given in Section 3.1 for packet flow for network traffic. A load-sharing algorithm partitions a system.
into domains consisting of sets of nodes. A Load State of a node is represented as a fuzzy term using fuzzy linguistic variables IDLE, LOW, NORMAL and HIGH. A similarity relation is used at node computation, to decide which other node to include in its domain depending on the node’s Load State. We use a mix of triangular and trapezoidal representation.

A typical fuzzy rule set is of the following form:

Rule 1: If the application CPU utilization of the link is VERY LOW, then the Node Load is IDLE.
Rule 2: If the CPU utilization is LOW AND Network Traffic is LOW, then the Node Load is LOW.
Rule 3: If the CPU utilization is MEDIUM AND Network Traffic is MEDIUM, then the Node Load is NORMAL.
Rule 4: If the Memory utilization is VERY HIGH OR (CPU utilization is HIGH AND Network Traffic is HIGH), then the Node Load is HIGH.

While there is no unique method for describing the fuzzy membership functions of individual parameters, the triangular (and/or mix of triangular and trapezoidal) representations are preferred over S and Pi type of curves, especially from computational simplicity angle [8]. We use standard hedges like very and more-or-less as Concentration (squaring) and Dilatation (square rooting) operations [8, 21, 22].

4. Role of Inter-Dialoguing Processor

Multiplexing of BPO infrastructure on the basis of geographical locations uses a three-schema architecture at the grass root level, comprising of a physical (internal) schema, a conceptual (logical) schema and an application (external) schema. Resource allocation should be carried out to cater to the increasing traffic through the network at the peak hours due to availability of efficient resources, resulting into the optimal utilization of the BPO infrastructure. We describe the concept of an Inter-Dialoguing Processor (IDP) that acts as a transfer-establishing resource link between the service center and the Customer Interface Array (CIA), dynamically. The main task of the IDP is to merge and route the traffic of the N service centers via a single transfer path to the CIA, which is responsible for interactions with the actual customer. The CCI value is computed for each service center. This value is used for handling the security issues. At the IDP level, the CIS value is calculated. The customer interface array uses the CIS value for routing information to various customers. It is also responsible for finding the CSR with the help of user feedback. The control-system-managing module of Inter-Dialoguing Processor (IDP) of BPO Multiplexer acts as a load balancer and changes the FMAP by increasing the membership value of lesser-used BPO infrastructure machines and reducing those of heavily loaded machines. FMAP is a resource relation that allows customers to determine which service center stores or caches which files. This representation and replication (FMAP is globally replicated and present in different machines with different entries) of FMAP enables the customers to be serviced with minimum number of network hops using BPO multiplexing, by taking into account parameters like link speed, bandwidth, loading status etc when assigning membership values in the FMAP. As and when a new machine joins the multiplexing infrastructure, the FMAP of the closest link is read, and the system-updating module of IDP modifies its own map accordingly. Updating of FMAPs takes place synchronously. This helps in preventing network congestion.

5. Conclusions and Future Work

This paper has focused on simulation and modeling issues in the BPO infrastructure from parallel and distributed computing perspective. The application of Fuzzy logic for load balancing has been discussed. The role and scope of Inter-Dialoguing-Processor is brought out as a unique methodology. In our analysis, we have deployed Type-1 fuzzy set representation where, the degree of membership (i.e. belongingness), is indicated by a number in the range from 0 to 1 on (Y- axis). It is crisp. Type-2 fuzzy logic represents uncertainty using a function, which is itself a type-1 fuzzy number. Since the value at each point in a type-2 fuzzy set is given as a function, type-2 fuzzy sets are three-dimensional. A number of researchers have recently shown that in certain applications that type-2 representation can outperform type-1. Type-2 fuzzy sets are recommended for situations where

(i) Data generating system is known to be time varying and the mathematical description of the time variability is unknown

(ii) Linguistic terms are used that have a non-measurable domain.

It is interesting to explore the parallel and distributed computing applied to BPO environment using Type-2
fuzzy set representation. CCI, CSR and CIS have the above intrinsic characteristic features and need to be handled using type-2 fuzzy sets.

**Acknowledgements**

The guidance, suggestions and the feedback on the initial manuscript received from Dr. Amar Gupta, Eller College of Management, University of Arizona has helped in improving the depth and coverage of this paper. Encouragement received from Dr Greg Andrews, Department of Computer Science, University of Arizona, in pursuing Professional Outsourcing as elective subject is highly appreciated.

**References**