An Analytic Quantitative Estimation of the Performance of the Mediator for Interaction Towards Automation of Internet Protocol and Carrier-Grade Management Ecosystems[†]

*Raufu Oluwatoyin Raheem**

The paper deals with the analytic estimation of the performance of a newly invented device for the interaction of the Internet Protocol network management system and the carrier-grade management ecosystem. The Erlang formula was used to estimate the performance of the new device in terms of black box and white box. The results showed that the new device would be of tremendous benefit to the network industry in terms of the management of the interaction and service creation in the industry, with the capacity to handle many requests within the shortest time with less cost, despite the effect of rollback situation that may be encountered.

Keywords: Analytic quantitative estimation, Performance, Grade of service, Rollback effect, Block effect, Erlang (B) extension

---

Introduction

The global Internet Protocol (IP) traffic was predicted to increase 4.3-fold in volume from 2009 to 2014 and reach 63.9 exabytes per month in 2014 (CISCO, 2012). With this prediction and the current rate of higher demand for Internet access, the IP network has been facing challenges of finding solutions that would manage the available network capacity. This management is in flexible manner that adds no cost to its present operating and capital expenditure. Also, there is a wide accord among Telecommunication Vendors and Operators (TVO) that the future generation network would be combined of Internet architectures embedded into high-bandwidth technologies and carrier-grade systems for control and management of the network (Juna, 2010).

Presently, the segmentation of IP and carrier-grade technologies has not only produced the carrier's organizational separation, but a kind of fragmentation of the technical competence through separate Network Management Systems (NMSs) (Yannuzzi *et al.,* 2012). The interaction between two NMSs is static and most of the operations were done manually. When there is a request for the creation of a service,

---

[†] This research was carried out by the author when he was a Ph.D. Research student at Technology Management Economics and Policy Program (TEMEP), College of Engineering, Seoul National University (SNU), Seoul, South Korea.

* Senior Engineer, Ministry of Communications, Abuja, Nigeria. E-mail: omotoso73@yahoo.com

© 2018 IUP. All Rights Reserved.

المنـــارة للاستشارات

www.manaraa.com

for example service provisioning, link restoration and other related services are initiated by the IP network management. The time taken to serve the request by the transport network management is too long: it takes hours and even days. This resulted in too much delay and business value loss in the Internet ecosystem.

Having identified this problem, the ONE-consortium via European Union (EU) Commission project (European Project (FP7) ONE {INFSO-ICT-258300}) invented a mediator called 'ONE-Adapter'. ONE-Adapter acts as a mediator for automatic interaction, coordinating and facilitating the communication between the two network management systems. However, mediator was implemented with Web Service Business Process Execution Language (WS-BPEL) as one of the platforms in order to wire a number of services during its mediating process (Chamania *et al.,* 2012). This paper perceived that an important factor for the performance of the mediator invented in such arena is whether the activities in the mediator are arranged to run with as much concurrency as possible. In practice, the developer may not arrange the activities correctly or efficiently. This problematic activity arrangement may lead to WS-BPEL program fault.

The performance evaluation carried out by ONE-Consortium is not compressive enough to convince the network management system to foresee the potentiality of newly invented mediator. ONE-Consortium evaluated the mediator based on shipping performance approach, instead of stress performance evaluation. The stress performance analysis would reveal the Quality of Service (QoS), and predict and give better understanding of the performance properties to determine the maximum capacity, to identify the performance of critical components and to remove the performance of bottleneck (Koziolek, 2010) of mediator that warrants managerial decision. However, the object of this paper is analytic quantitative estimation of the quality of the mediator for the interaction towards the automation of the Internet and the carrier-grade management systems in terms of stress performance evaluation. By doing so, the prospective value of mediator would become visible to the operators in the network management industry. Because in today's fourth industrial revolution era, network managers become impressed with a new technology only when it has economic viability and better performance compared to their present counterpart. In this paper, the following question was addressed: What is the performance of the new device (mediator) in terms of service level? For this question, the paper reviewed analytic method in determining the performance of a device and then the performance of the mediator was estimated based on the service level using analytic approach.

The remainder of this paper is organized as follows. Section 2 reviews related literature of traditional and mediator approaches, and the performance method. Section 3 presents the method in evaluation of the mediator with the use of analytic approach, comparing the traditional approach with the mediator with use of examples, and performs experiment using Mitan phoneCalc to test the quality of mediator in terms of Grade of Service (GoS). Section 4 presents the results and discussion. Finally, the paper ends with conclusion.

المنارة للاستشارات
www.manaraa.com

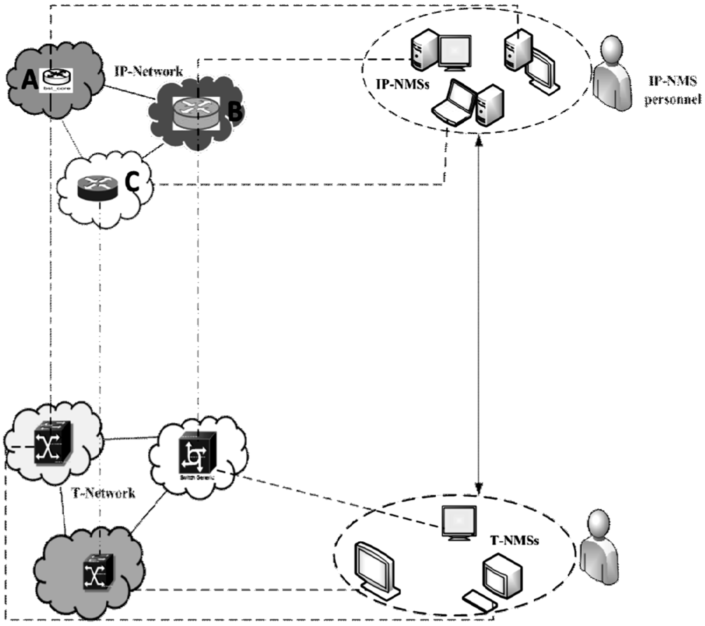
In terms of the contribution of this paper, it is the first approach to evaluate or estimate a newly invented device using the Erlang (B) extension in business process management, incorporating rollback effect in performance evaluation metric so that its viability becomes visible to network industry.

2. Literature Review

2.1 The Traditional Approach

The core operator's network is divided into two layers: Transport and IP, as shown in Figure 1. There are different network departments involved in configuration operations. The steps involved in creating a new IP link based on traditional approach are as follows (European project {FP7} ONE (INFSO-ICT-258300): The IP planning department elaborates a technical document defining the network creation request in which some of the necessary deployment interventions include the creation of new links. The IP provisioning department evaluates the network resources availability and, if it is necessary to buy new resource, it draws up a plan for purchasing of new equipment. This takes up to a week. After that, the network management department demands the request process information in order to execute the request. This demand may include information such as: Routers A, B, and C, as shown in Figure 1. Also, it takes one hour to a week.

Figure 1: Traditional Approach for the Interaction Between and Coordination of IP-NMSs and T-NMs



www.manaraa.com

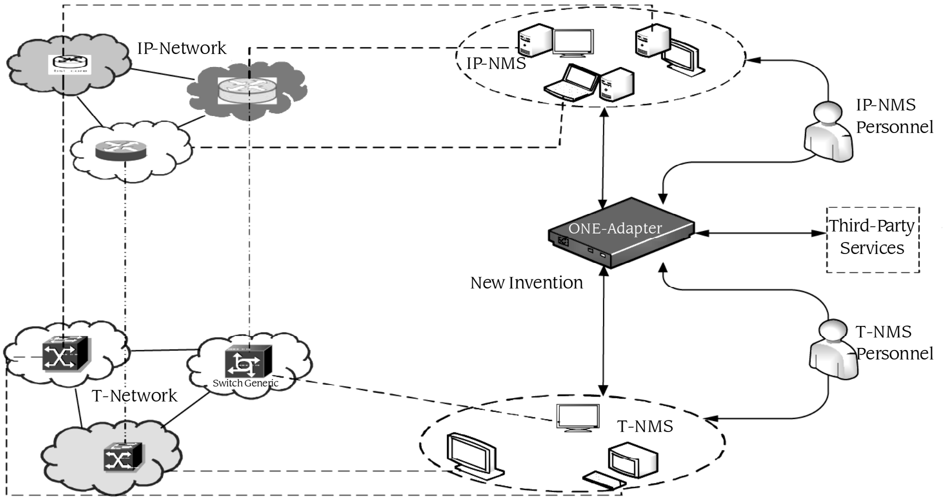
Request Processing at IP Level: If the acquisition of new IP equipment is necessary, the purchasing department sends the requests to the corresponding manufacturers to supply that equipment. It takes up one to three months. Also, IP engineering department installs the new equipment (time taken from 7-15 working days for the installation).

Request Processing at Transport Level: Transport planning and provisioning department checks the existence of the required equipment, and in case it required additional new equipment, it plans for the purchase of new equipment. If it already exists, it just goes to circuit configuration; otherwise, the purchasing department asks the corresponding manufacturers for the supply of all necessary equipment (duration: from one to three months). Transport engineering department installs and/or configures the equipment (duration: from one to two weeks).Transport engineering department carries out the interconnection of the managing equipment (duration: up to one week). Lastly the IP circuit is configured, validated and activated (this takes almost a week).

2.2 The Mediator Approach

Having identified the delay and longer service time in the traditional method of interaction and communication between the IP and transport network management systems, the ONE consortium invented a device called 'ONE-Adapter' (European project {FP7}. The ONE-Adapter approach is being developed to facilitate the interaction between the IP and transport network management systems. The significant innovation of the ONE-Consortium is the design and implementation of a communication adapter between the IP and transport network management systems via automatic management functions. The framework of the ONE is based on a set of semantics used for automated configurations. ONE-Adapter mediates between the IP Network Management System (IP-NMS) and the Transport Network Management System (T-NMS), as shown in Figure 2.

Figure 2: Mediator Approach for the Interaction Between and Coordination of IP-NMS and T-NMS



www.manaraa.com

Assuming that ONE-Adapter offers categories of services including link provisioning, IP service provisioning, IP-offloading and post-failure recovery in mediating between the IP-NMS and T-NMS, IP-NMS invokes ONE-Adapter with its input (input parameters) request via invocation.

The aforementioned parameters are related to the following requested services: link provisioning, IP service provisioning, IP-offloading and post-failure recovery. Upon receiving this request, ONE-Adapter processes the request internally with its element modules and then invokes T-NMS. T-NMS processes the request and invokes back ONE-Adapter with its response. ONE-Adapter then invokes back its response to IP-NMS. All the interfaces of the partner links are defined in Web Service Descriptive Language (WSDL). These allow ONE-Adapter to automatically interact or communicate with the other parties' elements, such as IP-NMS and T-NMS, via web service invocation. The ONE-Adapter element module consists of core and auxiliary modules. The core modules of ONE-Adapter comprises Management Controller (MC), Ontology Mapper (OM) and Workflow Processor (WP), while the auxiliary modules include: ONE-admin, T-NMS module, IP-NMS module, measurement module, topology module and Trigger Module (TM).

2.3 Performance

A system should be designed for optimum performance. The first paper released by Erlang in 1909 was regarded as the beginning of performance in network engineering, and since then, the work on performance has been extended by many scholars (Ghanbari *et al.,* 1997). Most of the works on performance have common important features: (a) the estimation of performance is based partly on statistical data and/or assumptions; and (b) some aspects of the estimated performance are stated in terms of probabilities. According to Ghanbari *et al.* (1997), the best approach to determine the performance is combining the analytical method with simulation. "Performance analysis of real-world systems often relies on analytical performance models such as queuing networks to capture limits on the maximum concurrency levels and pooling at software and hardware servers" (Casale *et al.,* 2015).

In the same manner, most of the researches on performance evaluation were based on the turnaround time (Marinescu, 2002) or performance turning (Mazanek and Hanus, 2011). Tan *et al.* (2007) itemized the process performance evaluation criteria into time (time index: process duration utility), service (service index: customer satisfaction), quality (quality index: cost structure utility), speed (speed index: product heap utility) and efficiency (efficiency index: resource usage utility).

Over several years, most models were proposed to analyze the performance of Distributed Software System Architecture (DSSA) with a view to avoiding the pitfalls of poor QoS (Olabiyisi *et al.,* 2010). Furthermore, Olabiyisi *et al.* (2010) suggested that the performance of software architecture evaluation needs to focus on both objective and subjective variables of the client and end users as input parameters.

www.manaraa.com

In addition, Koziolek (2010) conducted a comprehensive survey of performance approach that was component-based and identified that performance in the literature was always dealt with performance prediction and performance measurement. Koziolek (2010) distinguished the two performance approaches by defining that the performance prediction normally deals with the anticipated performance of a component-based design to avert a problem in the system implementation that may lead to substantial costs for redesigning the component architecture. While performance measurement is dealing with analysis, identification and observation of the performance implemented through component-based systems at running time to know their properties. It was noted that most authors who worked on component based performance were influenced by the following factors: component implementation, required services, usage file and deployment platform.

In summary, performance management systems are usually implemented as balanced and dynamic solutions for offering support to the decision-making process (Vukšiæ *et al.,* 2013). Therefore, the selection of a performance metric is important because the metrics are often in rivalry.

2.4 Erlang Approach

The Erlang approach has been successfully applied in the areas of telephone call centers (Mehrotra, 1997; Mehrotra and Fama, 2003; Franzese *et al.,* 2009; and Chromy *et al.,* 2011); ships at ports (Kalashnikov, 1994); hospitals (de Bruin *et al.,* 2007; Restrepo *et al.,* 2009; and de Bruin *et al.,* 2010) and the Internet (Bonald and Roberts, 2012). The importance of the Erlang formula has been analyzed by Bonald and Roberts (2012). Bonald and Roberts (2012) applied the formula to the Internet traffic to create an awareness of its usefulness for the Internet traffic, apart from the telephone network. Table 1 shows the summary of areas of successful application of Erlang approach with their purpose.

Table 1: Areas of Successful Application of Erlang Approach

Application Areas	Purpose	Reference
Telephone Call Center	Talking on the telephone to customers or prospects for principal businesses. Comparison of analytical and simulation model for call center in order to identify the advantage of using simulation.	Mehrotra (1997), Mehrotra and Fama (2003), Franzese *et al.* (2009) and Chromy *et al.* (2011)
Ships at Port	Ships arriving at the port to be moored (anchored).	Kalashnikov (1994)
Hospital	Investigation of the bottlenecks in the emergency care chain of cardiac inpatient flow: The primary goal is to determine the optimal bed allocation.	de Bruin *et al.* (2007)

المنارة للاستشارات
www.manaraa.com

Table 1 (Cont.)

Application Areas	Purpose	Reference
	Developed a decision support system based on the Erlang loss model, which can be used to evaluate the current size of nursing units.	de Bruin *et al*. (2010)
	Proposed model for pre-screening tools to identify promising ambulance allocations.	Restrepo *et al*. (2009)
Internet	Demonstration that the Internet has a formula linking demand, capacity and performance that in many ways is the analog the Erlang loss formula of telephony.	Bonald and Roberts (2012)

"Besides communication and computer networks, call centers and risk management are still hot application domains for queueing theory" (Boxma and Walraevens, 2017) with Erlang approach.

In general, Erlang distribution is used for incoming request that follows Poisson distribution, while service rate follows exponential distribution. Remarkably, Erlang is used effectively for handling any kind of time distribution (Franzese *et al.,* 2009). Based on its aforementioned successes, Erlang approach was used in this work to predict GoS of the newly invented device. The formulae for Erlang B, Erlang C and Erlang B extension are presented in Table 2.

Table 2: Type of Erlang Queue Model

Type of Model	Model Formula	Reference
Erlang B	$Erlang\,(B) = \frac{\frac{\rho^N}{N!}}{\Sigma_{i=0}^{N} \frac{\rho^i}{i!}}$	Bonald and Roberts (2012); and Garg and Goyal (2012)
Erlang C	$Erlang\,(C) = \frac{\frac{\rho^N}{N!(N-\rho)}}{\Sigma_{i=0}^{N-1} \frac{\rho^i}{i!} + \frac{\rho^N N}{N!(N-\rho)}}$	Chromy *et al.* (2011); and Bonald and Roberts (2012)
Erlang B Extension	$Erlang\,(B)\,Extention = \frac{\frac{\left(\frac{K_T}{k}\right)^N}{N!}}{\Sigma_{i=0}^{N} \frac{\left(\frac{K_T}{k}\right)^i}{i!}}$	Garg and Goyal (2012)

المنارة للاستشارات
www.manaraa.com

Table 2 shows the types of Erlang models with their model formula: Erlang C formula is based on the assumption that the client request or invocation is not attended to immediately but it needs to wait in queue. It is originally defined as: N-number of server or number of channel and Pc-probability that incoming request would not be served immediately, or it needs to wait in queue before it is served. While Erlang (B) extension is similar to Erlang B, it needs takes account of the fact that a percentage of the requests is immediately represented to the server.

3. Methodology

3.1 Analytic Model of Mediator as White Box

Mediator system consists of group process with formulation of sequence flow within the element of its modules and the tasks in process are subjected to the intertask sequencing rules to ensure effective interaction during the mediation.

Mediator system consists of group process (Z) with formulation of sequence flow within the element of its modules, and the tasks in process are subjected to the intertask sequencing rules. The group business process includes: event notification, workflow ID assigns activity, path computation activities, generating new IP links activity, creating connection tunnel and configure routing rules at the tunnel endpoints.

Event notification activity (Z1) is, in this case, to trigger the workflow; an event is generated by the IP-NMS or other Graphic Users Interface (GUI) and received by the TM. In order to receive triggers from the IP-NMS, the ONE-Adapter needs to implement an event listener for the IP-NMS which would be integrated into the TM design. Upon receiving the event, the TM performs authentication; some basic preprocessing is required to check for other preconditions needed for other correlated events in the event log.

After this, a trigger would be composed by the TM based on the information in the preconditions and the information carried by the event notification.

In workflow ID assigns activity (Z_2), the trigger is sent from the TM to the MC. The MC contacts the Otology Mapper (OM) with the trigger to verify the validity of the trigger and perform classification. This additional step ensures that the incoming trigger is valid, and in case of errors, it can point to an inconsistent trigger composition in the TM. In the next step, the MC assigns a workflow to the trigger based on the trigger classification and the workflow selection policies. The policies would assign priority to the workflow for queuing before processing in the Workflow Processing (WFP). The MC is cross-checked with the Authorization Authentication and Accounting (AAA) module for authorization of the user issuing the event (embedded in the trigger) to initiate the specified workflows. After authorization, the workflow is put in the waiting queue, from where it is taken by the WFP for initializing the workflow execution.

In path computation activities (Z_3), processing in the WFP generates a request for computation of the VPN service based on the information available in the trigger. The

للاستشارات المنارة
www.manaraa.com

WFP contacts the Programmable Logic Module (PLM), which in turn communicates with the Path Computation Element (PCE) with a request for path computation. During this communication, the WFP communicates with the OM which transforms the request so that it is understood by the session management interfaces in the PLM. The PCE sends the computed path to the PLM which then sends it to the OM, which transforms the response into the standard ontology and sends it to the WFP.

In generating new IP links activity (Z_4), in this case, there is a need to generate one or more new IP links as computed by the PCE response. The WFP communicates with the topology module (over the OM) to request free interface at the router endpoints and the corresponding client interfaces in the transport switches. Upon receiving this information, the WFP communicates with the T-NMS control module to request a circuit between the client interfaces (again over the OM). If the circuit setup is successful, the WFP communicates with the IP-NMS control module to configure the IP endpoint and includes the IP interface configuration information obtained above. It is important to note that the IP interface configuration requires that the IP interface be assigned IP addresses from a private address pool. To this end, two possibilities can be assumed: (i) the operator provides a pool of private IP addresses inside the event notification which is then forwarded in by the trigger; and (ii) this information is already available via an existing module such as the topology module (or possibly a separate configuration module).

Creating connection tunnel (Z_5), after forming the necessary circuits, the WFP updates the ERO received from the PCE via the PLM to incorporate the interface addresses assigned to the new links. The WFP then sends a request to the IP-NMS control module to create an MPLS tunnel based on the Explicit Route Object (ERO) provided. The IP-NMS control module sends confirmation that the tunnel has been established.

Configure routing rules at the tunnel endpoints (Z_6): after receiving the tunnel information in the response, the WFP uses the information available in the trigger and the tunnel information to create requests for the configuration of routing rules to forward VPN data into the tunnel. The WFP then requests the IP-NMS control module to configure routing rules at the tunnel endpoints. The IP-NMS control module responds with a confirmation after the rules have been set.

IP-offloading as a case: Based on the used case defined by Yannuzz *et al.* (2011), IP-offloading is an innovative multilayer operation targeted to reduce the IP traffic across unnecessary intermediate routers. They stated further that the traffic of specific services or a set of traffic flows become higher than the desired throughput; the ONE-Adapter would drive network resource requests in order to bypass the intermediate routers affected by the increase in traffic load. Thereby, the process involved in performing IP-offloading by ONE-Adapter consists of five group processes of workflows that include: Event notification, path computation, generating new IP links, creating connection tunnel and configure routing rules at the tunnel endpoints, as shown in Figure 3.

المنارة للاستشارات
www.manaraa.com

Figure 3 shows that each group process consists of subprocess workflow of element modules of ONE-Adapter that orchestrated via WS-BPEL platform by service composition of wiring processes.

Figure 3: Example of IP-Offloading Workflow Process with ONE-Adapter

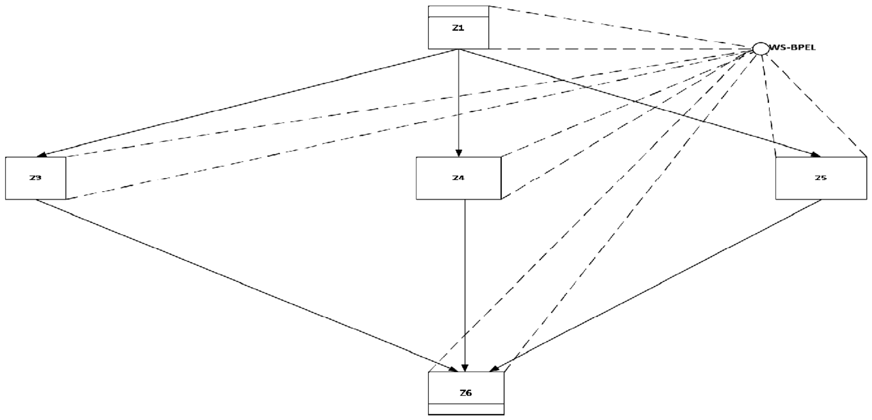
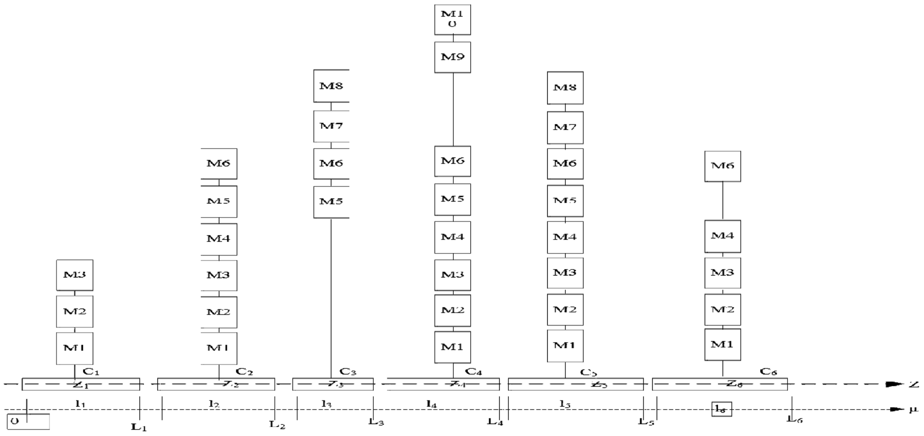


Figure 4 shows that M1 to M10 represent element modules of ONE-Adapter: M_1 is IP-NMS module, M_2 is IP-NMS module, M_3 is AAA, M_4 is MC, M_5 is OM, M_6 is workflow processor, M_7 is path computation element, M_8 is path configuration, M_9 is T-NMS control module and M_{10} is topology module.

Figure 4: Mediator as White Box as Group of Process Set (Multistage Process)



Giving a set of customized rules and user profiles, the customization process generates an individualized sequence of activity based on centralized scheduling system control by MC that acts as headquarter or nucleus in mediator system. It selects only the

www.manaraa.com

relevant component services (tasks) and obeys other constrains such as sequence ordering rule that is based on predefined business rule and type of services requested for and user profiles; customization process automatically selects tasks and arranges them. These properties distinguish mediator from other server that had been represented in queue or traffic modeling. Therefore, this paper considers mediator as white box and later as black box. The white box is perceived as a multistage process, as shown in Figure 4.

3.1.1 Mediator as White Box

Let Z represent the group process as indicated in Equation (1).

$$Z = \{Z_1, Z_2, \ldots Z_n\} \qquad \ldots(1)$$

Let M represent the corresponding module that executes the process as a set of modules available to it, as expressed in Equation (2).

$$M = \{m_1, m_2, \ldots m_j\} \qquad \ldots(2)$$

Let the corresponding task schedules set to be *H,* as indicated in Equation (3).

$$H = \{H_1, H_2, \ldots H_k, \ldots H_q\} \qquad \ldots(3)$$

Furthermore, the cost is very vital in the business process. This paper minimizes overhead cost of activities of mediator modules and its group process. Such overhead cost includes compensation handler cost, compensation scope cost. Let $C_{i,k}^{j}$ be the cost of allocating m_i to process Z_j and using the module for completion of task of process Z_j under schedule of H_k. In case H_k schedule need not process Z_j to use module m_i, thus the cost of $C_{i,k}^{j}$ is equal to zero. The H_k schedule gives optimal solution, if its cost is minimal, as expressed in Equation (4).

$$C = \min C(H_k) = \min \Sigma_{j=1}^{n} \Sigma_{i}^{q} C_{i,k}^{j} \qquad \ldots(4)$$

Also, in order to maximize the mediator's modules utilization (i.e., maximization of mediator utilization as system), there is a need to complete all process group's sets as fast as possible with error free, each module availability and with lower cost.

All the basic activities' costs are equal to zero except <invoke> activity. The invocation activities' costs are computed more on the lines of incremental and numbers of invocation made to other partner web service by mediator; the cost-in Equation (4) is taken as fixed cost for any scheduling process. Assume that invoking a single web service costs $0.1 and there is a discount rate of 10% for invoking more than one partner at a time. The Total Processing Cost (TPC) by mediator is computed as the sum of cost of basic processing cost by mediator modules (Fixed Cost-FC) and Variable Cost (VC) of invoking other partner web service by mediator and is expressed in Equation (5).

المنارة للاستشارات
www.manaraa.com

$$TPC_{\text{ONE-Adapter}} = FC_{\text{basic process by mediator modules}} + VC_{\text{invoke}} \quad ...(5)$$

where total cost of invoked is given in Equation (6).

$$VC_{\text{invoke}} = \Sigma_{i=1}^{q} Cost_{\text{invoke}} \quad ...(6)$$

where $i = 1, 2, \ldots, q$, q is the number of partners invoked at a given interval or time.

To estimate this, the study estimated average schedule completion time of group process Z; $Z_j \, \varepsilon \, Z$, as summarized in Table 3 with its corresponding processing time represented by T_{m1} to T_{m10} according to mediator's module elements (M).

3.1.2 *Average Schedule Completion Time*

The deadline for each process in Z is given in Table 3, while the execution time is given in Equations (7) and (8).

Table 3: Summary of Notation

Z	**Meaning**	**Modules**	**Process Time (T)**
Z_1	Event Notification	M_1, M_2, M_3	$L_1 = T_{M1} + T_{M2} + T_{M3}$
Z_2	Workflow Assign ID	$M_1, M_2, M_3, M_4, M_5, M_6$	$L_2 = T_{M1} + T_{M2} + T_{M3} + T_{M4} + T_{M5} + T_{M6}$
Z_3	Path Computation	M_5, M_6, M_7, M_8	$L_3 = T_{M5} + T_{M6} + T_{M7} + T_{M8}$
Z_4	Generating New IP Links	$M_1, M_2, M_3, M_4, M_5, M_6, M_9, M_{10}$	$L_4 = T_{M1} + T_{M2} + T_{M3} + T_{M4} + T_{M5} + T_{M6} + T_{M9} + T_{M10}$
Z_5	Creating Connection Tunnel	$M_1, M_2, M_3, M_4, M_5, M_6, M_7, M_8$	$L_5 = T_{M1} + T_{M2} + T_{M3} + T_{M4} + T_{M5} + T_{M6} + T_{M7} + T_{M8}$
Z_6	Configure Routing Rules at the Tunnel End-Points	M_1, M_2, M_3, M_4, M_6	$L_6 = T_{M1} + T_{M2} + T_{M3} + T_{M4} + T_{M6}$

$$L_i = \{L_1, L_2, \ldots L_n\} \quad ...(7)$$

$$l_i = \{l_1, l_2, \ldots, l_n\} \quad ...(8)$$

Therefore, slack time for process Z_j is expressed in Equation (9).

$$S_j = L_j - l_j \quad ...(9)$$

3.1.3 *Wrap Time*

The begin time of process Z_{j+1} is L_{j+1}^{1} and deadline of process Z_j is L_j, therefore, the wrap time between process Z_{j+1} and process Z_j is expressed in Equation (10).

المنارة للاستشارات
www.manaraa.com

$$Wrap\,Time = L^{1}_{j+1} - L_j \qquad \text{...(10)}$$

3.1.4 Stability Estimation

This is an estimation to detect the schedule of mediator if it is stable. Supposing that all processes Z_j are element of Z with the periodically occurred at period τ_j in mediator with N modules, necessary schedule is test mediator stability, the utilization of each N modules is as a maximum of one and is expressed in Equation (11).

$$\mu = \Sigma_i^N \mu_i = \Sigma_i^N \frac{l_i}{\tau_i} \le N \qquad \text{...(11)}$$

Also, the rate at which traffic is presented to the process Z_j in the element of Z periodically occurred as T_j in mediator with N modules, therefore, mediator stability is given as in Equation (12).

$$\lambda = \Sigma_i^N \lambda_i \qquad \text{...(12)}$$

Therefore, overall utilization of mediator with a given set of Z process of N modules is expressed in Equation (13).

$$\rho = \frac{\Sigma_i^N \lambda_i}{\Sigma_i^N \mu_i} = \frac{\lambda}{\mu} \qquad \text{...(13)}$$

3.1.5 Rollback Effect

All processes that perform successfully up to $(N - i)$ modules in Z without completion of process Z_j of schedule H_k would be rollback (undo); let the utilization of each $(N - i)$ module that will rollback is $d\mu$, as given in Equation (14).

$$d\mu = \Sigma_{j-1}^{N-i} \mu_j = \Sigma_j^{N-i} \frac{l_j}{\tau_j} \le (N - i) \qquad \text{...(14)}$$

Thus, new utilization of each N module that successfully performs operation before failure occurred is expressed in Equation (15).

$$\mu + d\mu \qquad \text{...(15)}$$

With the same rate at which traffic is presented to modules M in Z group set of process, the new overall utilization when rollback occurred is expressed in Equation (16).

$$\rho^* = \frac{\Sigma_i^N \lambda_i}{\Sigma_i^N \mu_i + d\mu} = \frac{\lambda}{\mu + d\mu} \qquad \text{...(16)}$$

المنارة للاستشارات
www.manaraa.com

3.1.6 Blocking Effect

When there is heavy traffic, any module M of process Z may be busy running and job scheduler sends the incoming task to group process Z_j to enqueue for certain short period of time (t) and later inqueue and execution. Also, such incoming task would be blocked from entry Module M_i of group process of Z_j in scheduler of H_k and such a task may be abandonment or cancel and retry.

Assuming that modules M of process Z are occupied due to certain events such as rollback, and based on group set of process of mediator modules that perform execution of service request information of multistage process, therefore overall blocking probability $P(b)$ as product of all modules that are busy in H_k schedule in the process Z_j is expressed in Equation (17).

$$P(b) = \Pi_{i=1}^{N} b_i \qquad \text{...(17)}$$

3.1.7 Grade of Service

The probability that modules M of mediator in a group set of process Z are occupied and equal to probability that incoming traffic is blocked. Consequently, GoS is expressed as in Equation (18) in terms of blocking probability.

$$GoS = 1 - \Pi_{i=1}^{N} b_i \qquad \text{...(18)}$$

3.2 Analytic Model of Mediator as Black Box

Mediator is a kind of automatic single server when considered as black box system. The service request is coming from IP-NMS or other stakeholder; when the ONE-Adapter is not occupied, the request is accepted and served incoming request immediately. In other manner, when the ONE-Adapter is occupied, the request is blocked. In case of blocking, IP-NMS has two options: either abandon the request (by emit request or cancel) or by making reentry the request again as represented in Figure 5. Figure 5 shows the analytical model of ONE-Adapter with Erlang (B) extension. For the purpose of this paper, it is assumed that the algorithm followed is: $M/M/1$ (for normal request without blocking) and $M/M/1 + du$ (for request with blocking or rollback) queue model. The request arrives following the Erlang distribution arrival rate and the service rate, with a rate of λ, while the service time of each request follows exponential distribution, with a mean of $1/\mu$. When ONE-Adapter is busy, any other request that comes in would be blocked automatically.

The requests are immediately represented to the server, ONE-Adapter if the requester encounters blocking during a busy time. ONE-Adapter accepts a request when it is not busy, as shown in Figure 6. Let the probability that ONE-Adapter is idle be P_0 and the probability that ONE-Adapter is occupied (busy) be P_1.

For the equilibrium state to establish λP_0 must be equal to μP_1 as stated in Equation (19) and utilization (ρ) of ONE-Adapter is expressed in Equation (20).

المنارة للاستشارات
www.manaraa.com

Figure 5: Analytical Model of ONE-Adapter with Erlang (B) Extension

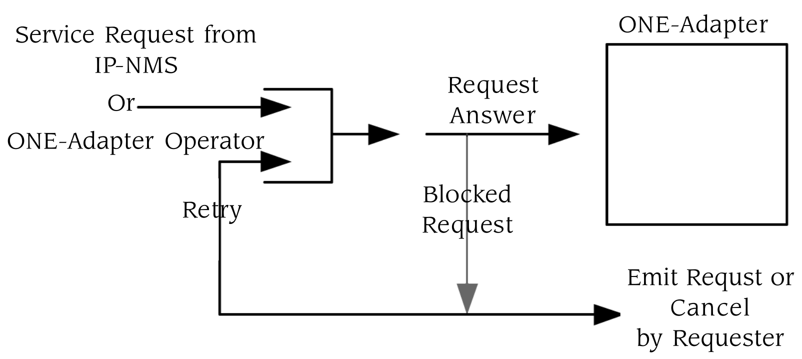
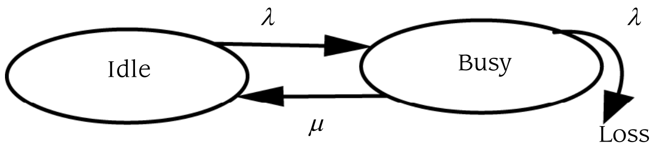


Figure 6: State Diagram of ONE-Adapter



$$\mu P_1 = \lambda P_0 \qquad \text{...(19)}$$

$$\frac{P_1}{P_0} = \frac{\lambda}{\mu} + \rho \qquad \text{...(20)}$$

From the general formula of Erlang (B) (Mazanek and Hanus, 2011), the probability of loss can be known and expressed as in Equation (21).

$$Erlang\,(B) = \frac{\frac{\rho^n}{n!}}{\Sigma_{i=0}^{n} \frac{\rho^i}{i!}} \qquad \text{where } i = 0, 1, \ldots, n \qquad \text{...(21)}$$

Therefore, the blocking probability, $E(B)$, of a request for service by ONE-Adapter can be expressed in terms of the probability state, as in Equation (22), where n is number of requests that came in at a particular interval or number of requests that ONE-Adapter can handle at a time.

المنارة للاستشارات
www.manaraa.com

$$E(B) = \frac{\dfrac{\left(\dfrac{P_1}{P_0}\right)^N}{N!}}{\Sigma_{i=0}^{n} \dfrac{\left(\dfrac{P_1}{P_0}\right)^i}{i!}} \qquad \text{...(22)}$$

The probability of the GoS of a request accepted by ONE-Adapter is estimated as in Equation (23).

$$GoS = 1 - Blocking\ Probability$$

$$GoS = 1 - \frac{\dfrac{\left(\dfrac{P_1}{P_0}\right)^N}{n!}}{\Sigma_{i=0}^{N} \dfrac{\left(\dfrac{P_1}{P_0}\right)^i}{i!}} \qquad \text{...(23)}$$

Equation (23) can be expressed in terms of percentage also. Thus, Equations (18) and (23) are compared, which results in Equation (24).

$$\Pi_{i=1}^{N} b_i - \frac{\dfrac{\left(\dfrac{P_1}{P_0}\right)^N}{n!}}{\Sigma_{i=0}^{N} \dfrac{\left(\dfrac{P_1}{P_0}\right)^i}{i!}} \qquad \text{...(24)}$$

For long process due to effect of rollback, Erlang C formula is employed from Table 3, as many requests would be queued up and only changes time for wrap time.

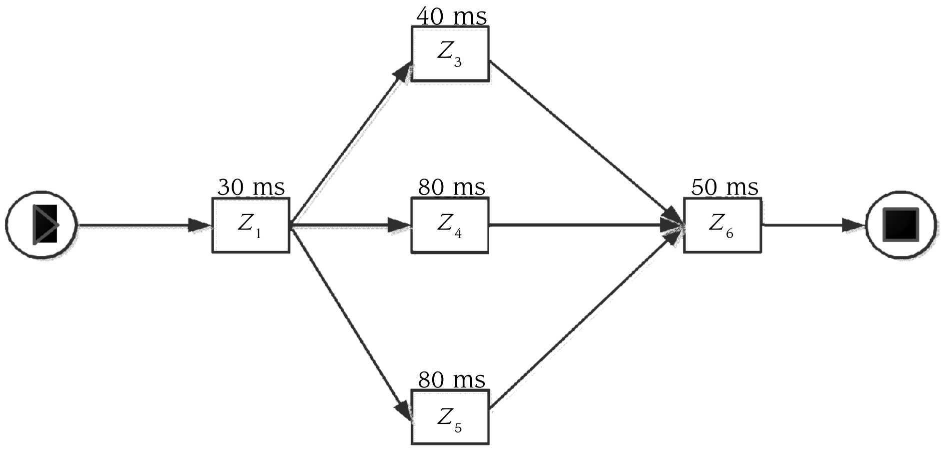
3.3 Mediator Approach: IP-Offloading Service as an Example

IP-offloading consists of five group processes: Event notification (Z_1), path computation (Z_3), generating new IP link (Z_4), creating connection tunnel (Z_5), and configures routing rules at the tunnel endpoints (Z_6). Each group process contains subprocess in mediator modules, as shown in Figure 7.

Figure 7 shows that group processes Z_3, Z_4 and Z_5 are to be performed in parallel, while Z_1 and Z_6 are group processes that start and end process respectively. This is

المنارة للاستشارات
www.manaraa.com

Figure 7: Example of IP-Offloading Workflow Process Activity on Node by Mediator



shown in Table 4, the total process time is computed for a single IP-offloading service requested by IP-NMS from mediator services.

Table 4: Scheduling of Process Activities by Mediator Approach

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}	Sum
Z_1	10	10	10	0	0	0	0	0	0	0	30
Z_3	0	0	0	0	10	10	10	10	0	0	40
Z_4	10	10	10	10	10	10	0	0	10	10	80
Z_5	10	10	10	10	10	10	10	10	0	0	80
Z_6	10	10	10	10	0	10	0	0	0	0	50

Note: IP-offloading ready: Total estimated time = 160 m/s.

In order to compute the time taken, it is assumed that each module of mediator takes equal time to process any task invoked to it. For the purpose of the paper, 10 ms are used for each module. Mediator component response time for use case: IP-offloading; assuming that each module takes 10 ms for processing any job that push/pass to it. Total time to process IP-offloading (Figure 4) and values *M*'s and *Z*'s (Table 4) are estimated as 160 m/s.

3.4 Traditional Appraoch: IP-Offloading Service as an Example

The IP-offloading process in traditional approach, as shown in Figure 8, is normally carried out between the requester, IP-NMS and TP-NMS.

المنارة للاستشارات
www.manaraa.com

From Figure 8, it is noted that the requester forwards its IP-offloading request parameter to IP-department via manual process and it took an hour before IP-department received the notification for such request. The IP-department checked IP network resource availability for path computation; this took three hours. After that, IP-department asked transport department for transport layer generating new IP link; this took two working days (48 h). The transport department checked transport resource and creating connection tunnel and it normally takes 5 h. Afterward, the transport department informed IP-department to configure routing takes at the tunnel endpoints, for the configuration of routing rules takes almost 48 h before message is conveyed to IP-offloading requester that the link is ready for offloading.

Figure 8: Example of IP-Offloading Workflow Process Activity by Traditional Approach

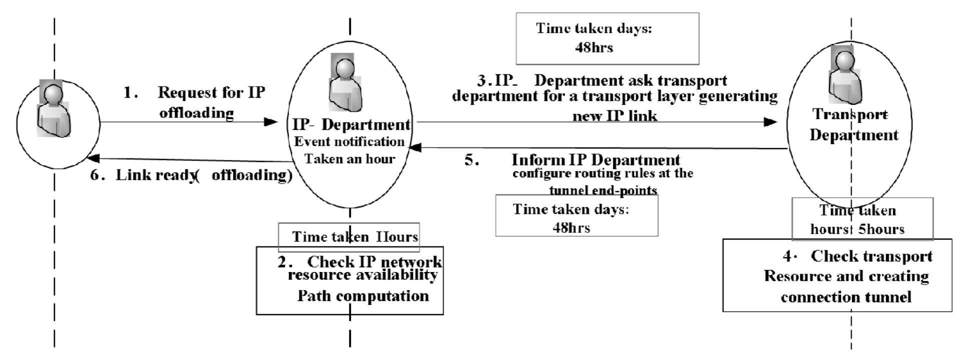


Table 5 shows that the total time for capture event/processing time of IP-service offloading via traditional approach is 105 h. Based on the two approaches (Tables 4 and 5), it can be deduced that in traditional approach there is a great loss of value in terms of time delay in capturing event from one transaction of an activity to other. This process is referred to as value lost through latency in event utility by Palmer (2011) and negatively exponential reduced to zero in solving requester problem. A lot of values are lost in traditional approach from business relevance of event occurred, event captured, event analyzed and action taken due to manual processing approach that increases total reaction time in handling of event, in contrast to quick handling and capture of event by

Table 5: Scheduling of Process Activities by Traditional Approach

S. No.	Events	Time Taken, T_n
1.	Event Notification	$T_1 = 1$ h
2.	Check Resource Availability/Path Configuration	$T_2 = 3$ h
3.	Generating IP Link	$T_3 = 48$ h
4.	Creating Connection Tunnel	$T_4 = 5$ h
5.	Configuration of Routing Rules	$T_4 = 48$ h
	IP-Offloading Ready: ($T_1 + T_2 + T_3 + T_4 + T_5$)	**= 105 h**

المنارة للاستشارات
www.manaraa.com

mediator via automated process. This reduced time in event captured, event analyzed and action taken eventually reduced total reaction time for IP-offloading to 160 m/s, compared to traditional approach's 105 h, before IP-offloading service can be served.

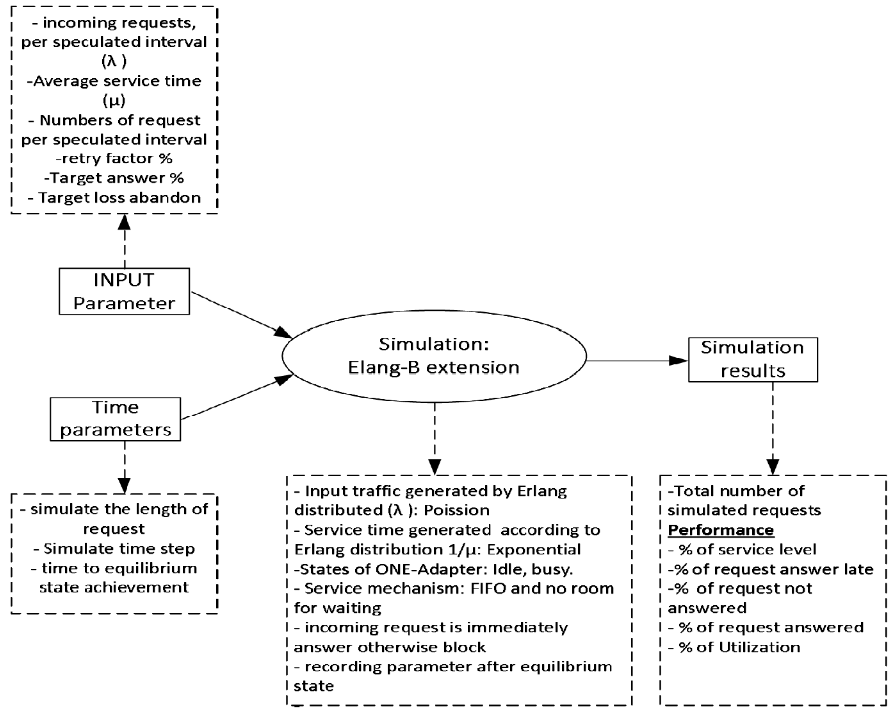
Next section shows the experiment using MITAN PhoneCalc Queueing Models (Tanner, 2013), to test the quality of mediator further by more traffic load. This allowed evaluating mediator performance based on the analytic model in section three.

3.5 Experiment Using Mitan PhoneCalc to Test the Quality of Mediator: GoS

In order to visualize the traffic distribution, blocking probability, GoS, as expressed in Equations (16) and (21), and considering other parameters that cannot be incorporated in Equation (21) such as reentry factor, interval factor, the paper uses MITAN phoneCalc T-calc V7.0 simulation as a virtual software packages. The tool is capable of handling simulation part of Erlang model (Erlang B, Erlang C and Erlang B extension) via T-cal and T-scope.

The following are the input parameters, as shown in Figure 9—workload inputted into T-calc program package: average service time is 160 ms (estimated from Table 1); reentry factor of 4.7; and target answer percentage is 74% and target loss abandon/loss rate of 5%, is used based on the recommendation of Restrepo *et al.* (2009).

Figure 9: Process of Simulation for ONE-Adapter Using Erlang (B) Extension Model



المنارة للاستشارات
www.manaraa.com

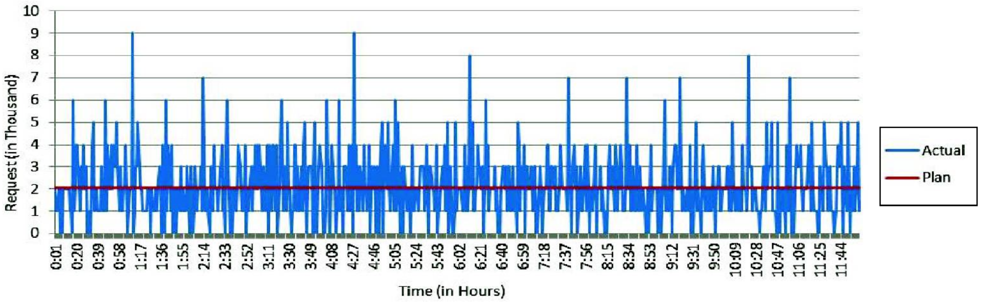
Figure 9 shows that the input number of requests loaded are one thousand requests within interval of 5 min, 30 min and 60 min, respectively. The requests later increase in 1,000 and up to 30,000 requests are recorded accordingly for each of speculated interval in each Erlang model type (Erlang B, Erlang B extension and Erlang C).

From the experiment, it is observed that Erlang B and Erlang C are not suitable for queue modeling of mediator due to rollback process of BPEL that warrant blocking and reenter of incoming request for service from IP-NMS to mediator. Erlang B extension is suitable for queue modeling of mediator because it deals with blocking and reenters of incoming request. Therefore, Erlang B and Erlang C are served as control experiment.

4. Results and Discussion

Figure 10 represents the outcome of T-scope simulation which showed how the arrival distribution would occur with maximum of 10,000 requests plan to be handed within 12 h. The plan was two thousand requests would come into ONE-Adapter within a few minutes interval, as indicated in red in Figure 10, while actual requests are indicated in blue color within the speculated time. Vertical axis represents number of requests (in thousand), while horizontal axis represents time (period of 12 h). Number of requests per minute within 12 h, where 1,000 to 2,000 requests have highest number of requests that come-in into mediator within the planned interval.

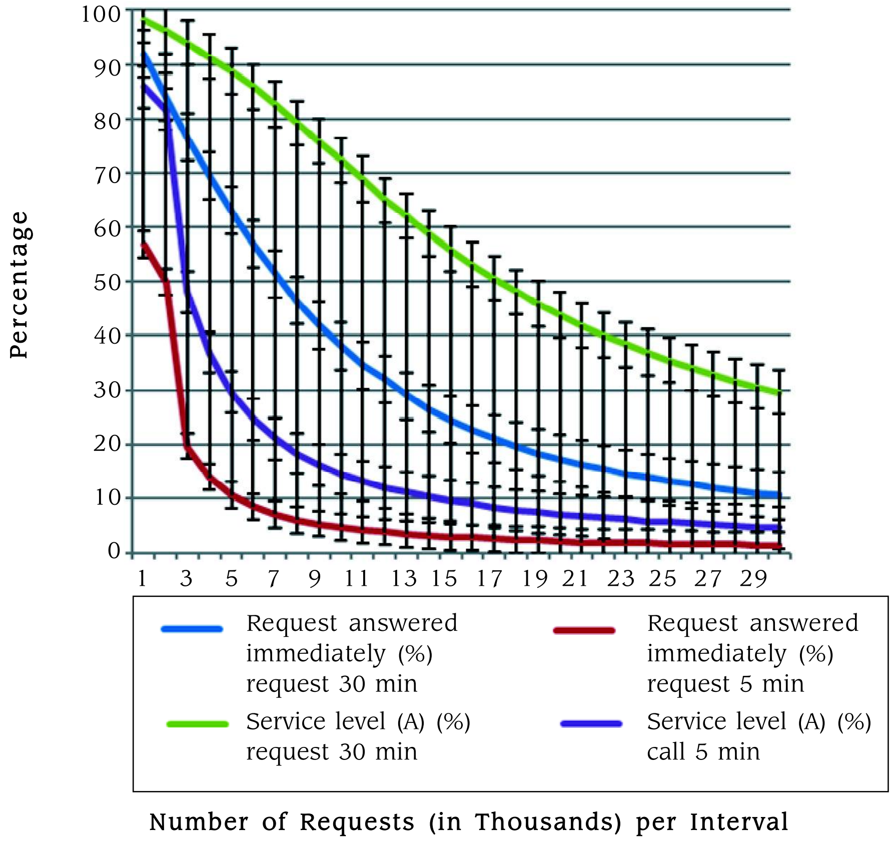
Figure 10: Arrival Distribution of Request from T-Scope Simulation



Having identified the T-scope results in Figure 10, the report was obtained from outcome of loaded parameters as input into T-calc and these parameters were increased along with the time interval of 5 min and later to 30 min, for each of the one thousand requests up to 30 thousand requests. This T-calc helped in visualizing the stress performance in terms of the service level or GoS and in terms of the requests that would probably not be answered: as the number of requests became higher, the service level decreased. The same can be said regarding the number of requests that were answered immediately, but the number was lower than the GoS. Also, the GoS percentage was higher in the 30 min interval than in the 5 min interval. This is shown in Figure 11.

المنارة للاستشارات
www.manaraa.com

Figure 11: Percentage of Service Level and the Requests Answered Immediately Within 5 min Versus 30 min Interval



The percentage of requests that were answered in 30 min interval late is lower than the percentage of requests that were answered 5 min interval late (Figure 12).

Figure 13 shows the percentage of requests that were answered and those that were not answered within the speculated interval. The sum of the percentages of requests answered and not answered was equal to 100% in each time interval used.

In the same manner, the percentage of utilization was higher in the 5 min interval than in the 30 min interval, as indicated in Figure 14.

Figure 15 shows the cost distribution of fixed cost of mediator modules, variable cost due to support mediator with WS and total processing cost.

The results imply that there is a need to balance the service level and the number of requests to be handled within a particular time interval with the utilization. With these results, this paper has helped in proper planning of the actual requests to be handled within a given time interval to lower the minimum blocking probability and to reduce the number of unanswered requests.

المنارة للاستشارات
www.manaraa.com

Figure 12: Percentage of Requests Answered Late: 5 min Versus 30 min

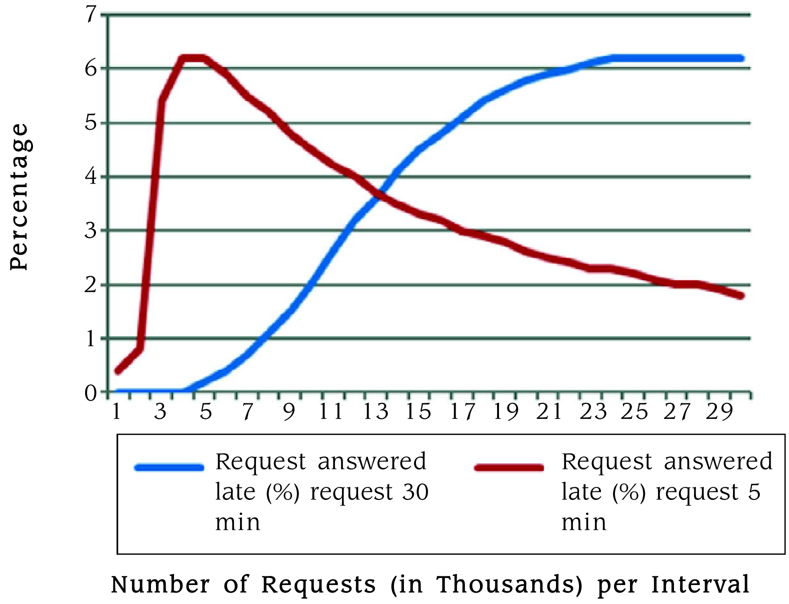
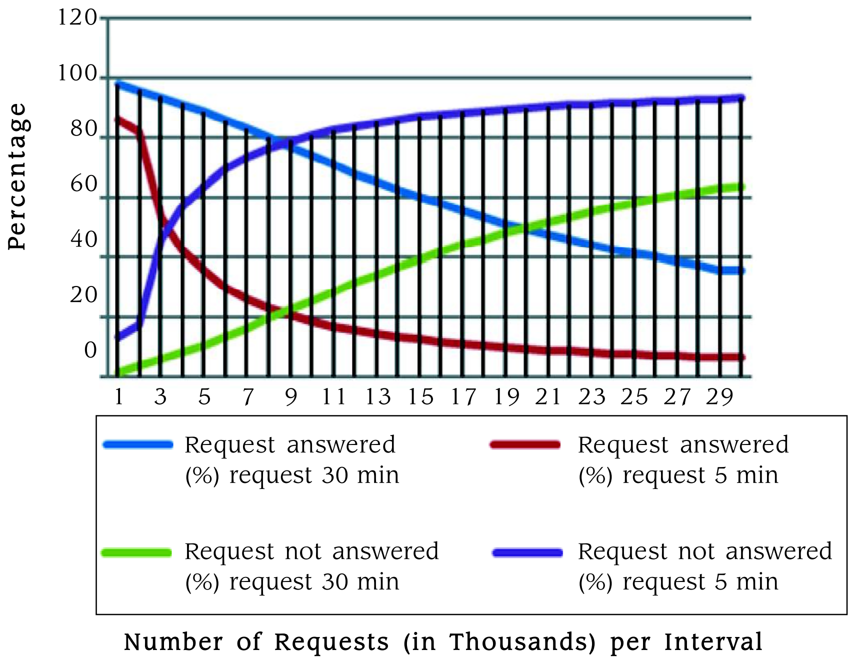


Figure 13: Percentage of Requests Answered and not Answered Within 5 min Interval Versus 30 min Interval



المنارة للاستشارات
www.manaraa.com

Figure 14: Utilization (%) Request 5 min Interval Versus 30 min Interval

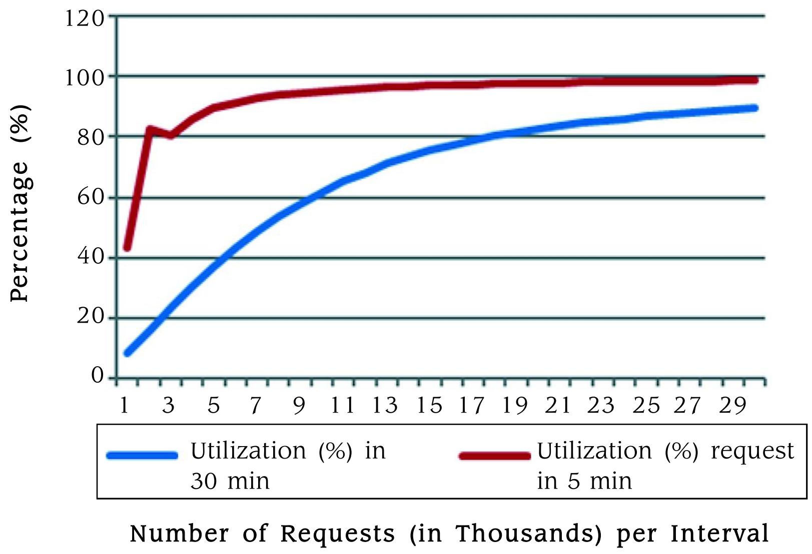
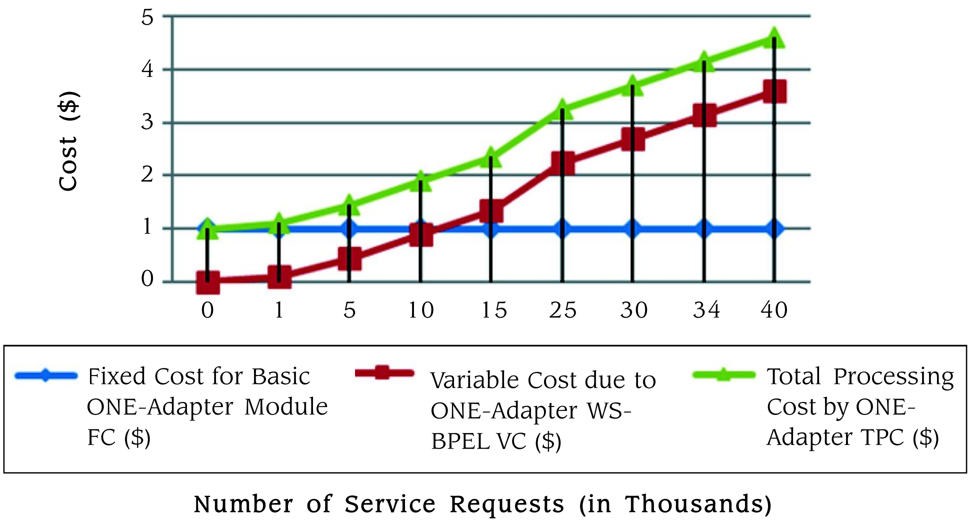


Figure 15: Total Cost Estimation for Support Mediator with WS-BPEL



Conclusion

In this paper, Erlang (B) extension is used to represent the behavior of mediator in real time using a single scenario with performance evaluation and cost estimation. The work would help the operators make managerial decision as to how to balance between

www.manaraa.com

performance metric parameters when mediator is deployed. Also, it would assist in evaluating how many requests to be handled by mediator per minute, hour, or week. Lastly, the work allows the network industry to see the potentiality of ONE-Adapter and how it can help the industry deal with its present situation.

The limitation of this work is that personnel cost for those who would be incharge of monitoring the work of ONE-Adapter and other overhead costs that may be encountered when putting it into operation were not considered. It is recommended that such costs be included in the performance evaluation in further studies. 🖥

References

1. Bonald T and Roberts J W (2012), "Internet and the Erlang Formula", *ACM SIGCOMM Computer Communication Review*, Vol. 42, No. 1, pp. 23-30.
2. Boxma O and Walraevens J (2017), "Computational Methods and Applications in Queueing Theory", *Annals of Operations Research*, pp. 1-2.
3. Casale G, Pérez J F and Wang W (2015), "QD-AMVA: Evaluating Systems with Queue-Dependent Service Requirements", *Performance Evaluation*, Vol. 91, pp. 80-98.
4. Chamania M, Demirbilek E and Jukan A (2012), "Using BPEL Workflow Processing for Cross-Layer Orchestrations in IP-Over-Optical Networks: A Proof of Concept", *Network Operations and Management Symposium (NOMS), IEEE,* pp. 1191-1194.
5. Chromy E, Misuth T and Kavacky M (2011), "Erlang C Formula and Its Use in the Call Centers", *Advances in Electrical and Electronic Engineering*, Vol. 9, No. 1, pp. 7-13.
6. CISCO (2012), "CISCO Visual Network Index: Forecast and Methodology, 2011-2016", Unpublished, White Paper, available at www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns705/ns827/white_paper_c11-481360.pdf. Accessed on December 30, 2012.
7. de Bruin A M, Van Rossum A C and Visser M C (2007), "Modeling the Emergency Cardiac in-Patient Flow: An Application of Queuing Theory", *Health Care Management Science*, Vol. 10, No. 2, pp. 125-137.
8. de Bruin A M, Bekker R and Van Zanten L (2010), "Dimensioning Hospital Wards Using the Erlang Loss Model", *Annals of Operations Research*, Vol. 178, No. 1, pp. 23-43.
9. European Project (FP7) ONE (INFSO-ICT-258300), *Deliverable D5.2: Report on Techno-Economic Validation Testbed Setup and Experimental Validation*, Unpublished. Accessed on August 30, 2013.
10. European Project (FP7) ONE (INFSO-ICT-258300), *Towards Automated Interactions Between the Internet and the Carrier-Grade Management Ecosystems*, available at http://www.ict-one.eu/. Accessed on November 23, 2011.

المنارة للاستشارات

www.manaraa.com

11. Franzese L A G, Fioroni M M and Botter R C (2009), "Comparison of Call Center Models", in *Winter Simulation Conference*, December, pp. 2963-2970, Winter Simulation Conference.

12. Garg P and Goyal N (2012), "Congestion Control Using Modified Erlangs Loss Formulae", *International Journal of Engineering Research and Technology,* Vol. 1, No. 6, pp. 1-3. ESRSA Publications.

13. Ghanbari M, Hughes C J, Sinclair M C *et al.* (1997), "Principles of Performance Engineering for Telecommunication and Information Systems", *IEE Telecommunications Series*, available at http://www.ict-one.eu/wiki/doku.php?id=oneproject:wp5: deliverables:d5.2& rev=1376381424

14. Juna A (2010), "ONE Towards Automated Interactions Between the Internet and the Carrier-Grade Management Ecosystems", *Future Network, European Commission Information Society and Media*, Unpublished. Accessed on March, 2011.

15. Kalashnikov V V (1994), "Mathematical Methods in Queuing Theory", *Mathematics and Its Application,* Vol. 271, pp. 5-15.

16. Koziolek H (2010), "Performance Evaluation of Component-Based Software Systems: A Survey", *Performance Evaluation*, Vol. 67, No. 8, pp. 634-658.

17. Marinescu D C (2002), *Internet-Based Workflow Management, Toward a Semantic Web*, p. 3, John Wiley & Sons Inc. Publication, New York.

18. Mazanek S and Hanus M (2011), "Constructing a Bidirectional Transformation Between BPMN and BPEL with a Functional Logic Programming Language", *Journal of Visual Languages & Computing*, Vol. 22, No. 1, pp. 66-89.

19. Mehrotra V (1997), "Ringing Up Big Business", *OR/MS Today*, Vol. 24, No. 4, pp. 18-25.

20. Mehrotra V and Fama J (2003), "Call Center Simulations: Call Center Simulation Modeling: Methods, Challenges, and Opportunities", *Proceedings of the 35th Conference on Winter Simulation: Driving Innovation,* December, pp. 135-143, Winter Simulation Conference.

21. Olabiyisi S O, Omidiora E O, Uzoka F M E *et al*. (2010), "A Survey of Performance Evaluation Models for Distributed Software System Architecture", March, *World Congress on Engineering,* July 4-6, 2012, Vol. 2186, pp. 35-43, International Association of Engineers, London, UK.

22. Palmer N (2011), "The Role of Trust and Reputation in Social BPM", *Social BPM: Work, Planning, and Collaboration Under the Impact of Social Technology*, pp. 35-43.

23. Restrepo M, Henderson S G and Topaloglu H (2009), "Erlang Loss Models for the Static Deployment of Ambulances", *Health Care Management Science*, Vol. 12, No. 1, pp. 67-79.

المنارة للاستشارات

www.manaraa.com

24. Tan W, Shen W and Zhao J (2007), "A Methodology for Dynamic Enterprise Process Performance Evaluation", *Computers in Industry*, Vol. 58, No. 5, pp. 474-485.

25. Tanner M (2013), "PhoneCalc Queueing Models", MITAN Ltd., available at http://www.mitan.co.uk/mainhome.htm. Accessed on July 15, 2013.

26. Vukšiæ V B, Bach M P and Popoviè A (2013), "Supporting Performance Management with Business Process Management and Business Intelligence: A Case Analysis of Integration and Orchestration", *International Journal of Information Management*, Vol. 33, No. 4, pp. 613-619.

27. Yannuzzi M, Masip-Bruin X, De Dios O G *et al.* (2011), "Bridging the Interoperability Gap Between the Internet and Optical Network Management Systems", pp. 177-180, in Networks and Optical Communications (NOC), 2011 16th European Conference, IEEE.

28. Yannuzzi M, Jukan A, Masip-Bruin X *et al.* (2012), "The Internet and Transport Network Management Ecosystems: A Roadmap Toward Convergence", 16th International Conference on Optical Network Design and Modeling (ONDM), April, pp. 1-6, IEEE.

Reference # 56J-2018-01-01-01

المنارة للاستشارات
www.manaraa.com

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.

www.manaraa.com