UDC 621.39

M.M. Klymash, I.V. Demydov, O.A. Lavriv, Yu.D. Dobush

[1]*Lviv Polytechnic National University, Lviv*

ANALYSIS OF SERVICE WORKFLOWS DISTRIBUTION AND SERVICE DELIVERY PLATFORM PARAMETERS

Developing the newest incarnations of E-business over information technology (IT) industry we stress challenges with traditional networking approach that was basic one for modern communications in very last years. To fully enjoy the benefits of best business applications built on the basis of adopted powerful Service Delivery Platforms (SDP) it is necessary to ensure that QoS with service workflows' distribution and applied interaction is sufficient; whereas whole the system is survivable and effective. This paper is focused on key technical issues of service workflow distribution with qualitative ensuring approach to obtain SDP parameters needed to get service reliability under specified SLAs.

Key words: *Service Delivery Platform, Quality of Service, workflows distribution, service-oriented resource planning.*

Introduction

E-business has emerged as a new paradigm for the use and delivery of network IT, and is evolving the wide development and support of services in very deep integration with most of human activities in dynamic realization to fit consumers' necessities around the world. In the network context, a wide range of mentioned IT resources and capabilities were developed and realized as modern grid and cloud technologies implementation. Interconnected servers, networking, storages are operating with middleware, some kinds of data, making secure applications interaction, supporting much of known business processes. To provide these processes with excellent concurrent qualities we need ensure, that there are available ways for provisioning, economical positioning, and flexible scaling over distributed SDP. Developing these newest incarnations of IT industry we stress a challenges with traditional workflow management approach that was basic for modern communications in very last years. To fully enjoy the benefits of best business applications built on the basis of adopted powerful SDP it is necessary to ensure that QoS with service workflows' distribution and applied interaction is sufficient; whereas the system is survivable and effective. This paper is focused on key technical issues of service workflow management with qualitative ensuring approach to obtain SDP parameters needed to get service reliability under specified SLAs.

In the chapter 1.1 we describe main relations between parameters of SDP and QoS indexes. In the chapter 1.2 we analyze two modes of multiservice flows forming and their impact on QoS. In the chapter 1.3 we present a simulation approach to formalize the infrastructure of distributed SDP as plurality of service nodes with service interfaces and service workflows' routers. Chapter 2 describes the solution of the service workflows prediction and distribution tasks within method of SDP parameters calculation.

QoS and SDP workflow management approach

Either of technologies implemented as SDP faced with management of service workflows. There are several parameters of SDP that are necessary to provide QoS by SLA on the specified level. One of them is the workflow delay that mainly consists of routing delay and operational delay on service interfaces. Workflow management into distributed service platform was studied in numerous papers such as [1, 2, 11]; a service interfaces operation analysis was carried out in [3-6] both with main architectural conceptions for SDP and ESB technologies. Next problem lies in service and infrastructure reliability. Packed switched network guarantees that with correct parameters selection of the infrastructure to be realized over them, some level of reliability could be achieved *a priory*. This could be clearly verified even by early US DARPA investigations [7], but the main task is to guarantee the service reliability by correct operation of the service nodes. The service node resources are similar to routing nodes' buffers and technologically limited by quantity.

The hierarchical structure of SDP is caused by the demands of high scalability and diversification of access, aggregation, and content distribution challenges. The direct result of the use of this kind of structure causes the difference of traffic properties at each level of the network model. Consequently, the parameters of the nodes might be different to provide appropriate QoS parameters. So, to choose the node parameters is a very important task of service quality.

It is clear, that packed switched networks are the basis for any service infrastructure for next years. One of the main problems of service workflow management in networks with packet switching is the absence of the possibility to guarantee transmission delay of the informational service workflows, whereas in such infrastructures the number of users, that are simultaneously users of underlying grid service composition, increasing dramatically, generating more and more service workload.

For the multiservice traffic it is important to define the maximum service quality requirements for the particular application in the common flow. These parameters have to be accepted as basic in the process of the network's resource planning.

To control situations when service workflows' delays and, respectively, delays caused by the workload on the service interface shall exceed permitted (in a case of insufficient resources), under the condition of the increasing quantity of the users we could define two main ways:

a) to use more faster service infrastructure and more productive service interfaces, that is not economically confirmed as a rule;

b) to use several alternative workflow configurations for service workflows, which generally with sufficient number of service interfaces that could prove acceptable end-to-end service delay and service reliability due to SLA restrictions.

Such service workflow management policies and service delivery platform configurations could be obtained by some routing, or in other words, workflow management algorithm using. This situation needs elaborating approach with study of the network infrastructure SDP workload and its spatial structure-topology is more survivable for freely scalable network in our opinion. By spatial service nodes location and their configuration there should be formed underloaded areas with the least service workflows' runtime delay [1, 10, 11]. Thus we need to review service flows distribution more detail.

Resources Utilization by the Multiservice Flows

Let's us consider the problem of resources utilization for each service quality parameter in more detail. The generalized model of traffic distribution is shown in Fig. 1. This scheme demonstrates forming a multiservice packet flow and its processing by a servicing network node.

Every application of the end-user generates the ordered sequence of the packets with equal processing priorities. The aggregation of these arranged sequences in the single multiservice flow occurs at the next stage. The processing priorities in this flow are different. And the total sequence has stochastic applications allocation.

The servicing node gives proper resources to guarantee the quality of service. Thus, the model also can reach the task of defining the number of applications in the multiservice flow or, if the number of applications is fixed, it answers the question about the total amount of multiservice flows which can be processed by the node. We propose to differentiate the two modes of the multiservice flow forming. The first one is similar to the virtual channel switching. It is represented in Fig. 2. The total flow is the effect of consequent resource utilization by each application flow. In this kind of servicing system there are no priorities for different packets and they are lost only in the crisis situation (some crashes).

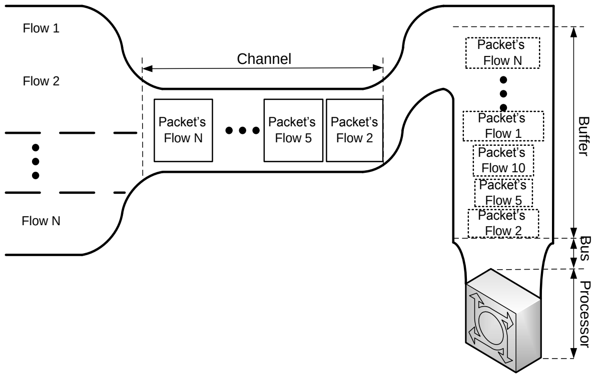

Fig. 1. Multiservice traffic distribution model

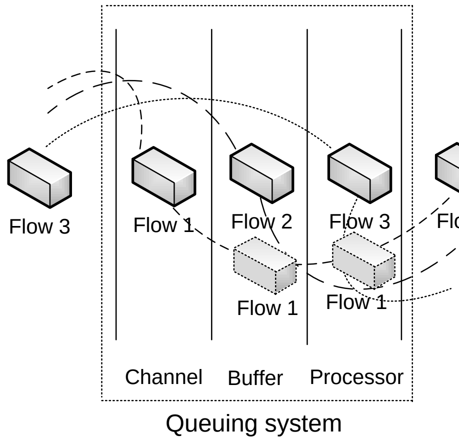

Fig. 2. Multiservice flow formation within a virtual channel mode

The second one represents the situation of simultaneous packet generation by the applications and their combination into a single flow. In that case different packets of a particular service may use different node resources at the same time (see Fig. 3). This process is absolutely stochastic. Thus, the service quality parameters cannot be guaranteed.

We must say that the service quality parameters for particular traffic may be provided by the selection of the appropriate structural and functional parameters of servicing node.

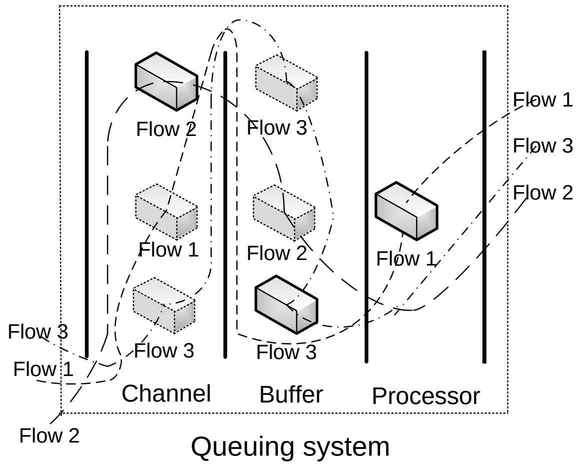

Fig. 3. Parallel packet generation and their combination into a single flow

Simulation approach for SDP infrastructure

The SDP is generally based on the principles of highways and hierarchies - in order to minimize the service workflow communication (the number of recource restricted components). Therefore, service infrastructure should be designed to maximize efficiency in the expensive SDP portal segment that matches the high level of expensive communication systems, which use to be maximally effective (see Fig. 4).

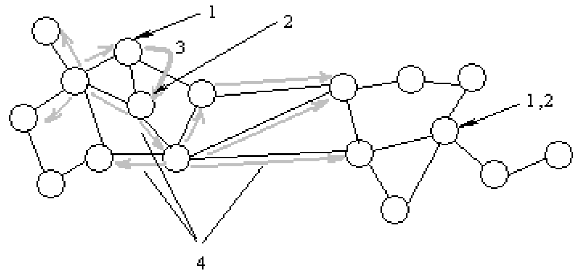

Fig. 4. An example of distributed SDP implementation as network grid with service workflows concentration on the basis of portal allocation resulted after heuristic analysis in the service node 2 by delay and topology metrics

The middleware interaction workflow (3) via service interface 1 and peripheral service workflows (4) requests are shown on the Fig. 4. The service infrastructure could be formalized as adjacent matrix of the net interconnections in generally between service interfaces and service workflow transit nodes [1, 11].

The specific subjective of this work is determine SDP parameters with providing QoS by minimal delay of the service workflows distribution given with prediction of SDP resources workload and its real indexes.

SDP parameters analysis and calculation

A brief analysis of service workflows distribution

Given adjacent matrix $\|c\|$ describing service interconnections topology into SDP infrastructure we could review some net service system with queuing (for example packet-switched reliable communication bus) with N nodes. Let us describe service workflows between service interfaces nodes by the workflow indexes matrix λ_{ij}. In general case we predefined that SDP service infrastructure was designed matching of services' consumers SLA requirements.

Therefore, a priory is clear that for service interface nodes and transit workflow nodes there available resources are characterized by:

- The adequacy of buffer size and productive resource amount for processing of service workflows requested by users;
- Their restricted amount, is the feature of real SDP infrastructure.

After that, we could note, that for each workflow in the packed-switched SDP communication bus the following equation comes true describing service workflow's delay in interconnection between service nodes *i* and *j*:

$$t_{ij} \approx k\Lambda_0\lambda_{ij}, \quad (1)$$

where $k = 1$ is the proportionality index.

Λ_0 – the index, depend on the service logics, and respectively algorithm, into some buffer resource of the service node or server, and its amount. In this case we could define:

$$\Lambda_0 = \left[t_{proc} + t_{wait}\left(\frac{\overline{d_q}}{d_{max}}\right)\right] t_{proc}^{-1} = 1 + \frac{t_{wait}}{t_{proc}}\frac{\overline{d_q}}{d_{max}}, \quad (2)$$

where t_{proc} – is the time of the one request processing into service workflow with not loaded service node (resources are unused and buffer is empty, general requirement for service processing or service workflows transiting trough the service bus); t_{wait} – is the time of delay until service node resource could be ready to process service request (service workflow in queue waiting, frequently is specified in the most of SLAs, mainly for real-time oriented tasks, exceeding this time could be equal to DoS failure); $\overline{d_q}$ – is the average workload of the service node (the average service requests' queue length, this workload could be estimated or required by SLA and fixed for calculations); d_{max} – is the maximal possible workload of the service node (maximal service requests' queue length, as a rule d_{max} is fixed as SDP basic equipment parameter).

We assume assumption that working efficiently the SDP is not overloaded. Therefore we described service nodes of SDP infrastructure by the queuing model in Kendall's notation [8] *M/M/*1*/N* at $N \to \infty$ (buffer going to be empty or service resources unused).

To calculate average service workflow delay for packet switched infrastructure of SDP we used Kleinrock formula withdrawal from Little's law [9, 10, 11]:

$$T = \frac{1}{\gamma}\sum_{k=1}^{N}\sum_{l=1}^{N}\lambda_{kl}\cdot t_{kl}, \quad (3)$$

where γ – the full service workload of distributed SDP grid, that could be calculated by following formula:

$$\gamma = \sum_{i=1}^{N}\sum_{j=1}^{N}\gamma_{ij}, \quad (4)$$

where γ_{ij} – is the sum of service workflows between nodes *i* and *j*.

Taking into account the balanced symmetric character of the service requests and reply via interfaces interconnection both with transit distribution of service workflows we could modify (3) as following:

$$T = \frac{1}{\gamma}\sum_{k=1}^{N}\sum_{l=1}^{N}\lambda_{lk}\cdot t_{kl}\,. \qquad (5)$$

With regard to (1) the value of service workflows' delays was written as:

$$T = \frac{\Lambda_0}{\gamma}\sum_{k=1}^{N}\sum_{l=1}^{N}\lambda_{lk}\cdot\lambda_{kl}\,. \qquad (6)$$

Solving service workflows distribution task

Optimization of the service infrastructure configuration on the criterion of minimum service workflow delay could be made if considering the system workload dependent proportionally on SDP sub topological properties (calculating by adjacent matrix transforming for service workload prediction), that resulted as matrix elements set $\|c\|$, raised to power *N-1*. Commonly the service workflow interconnection could be written by following functional relation [10]:

$$T_{min} = \frac{\Lambda_0}{\gamma}\sum_{k=1}^{N}\sum_{l=1}^{N}\lambda_{lk}\cdot\lambda_{kl} - \frac{\beta_{-1}}{\gamma}\sum_{i=1}^{N}\sum_{j=1}^{N}c_{ij}^{N-1}\cdot c_{ji}^{N-1}\,. \qquad (7)$$

The criterion of minimum service workflow delivery delay:

$$T \to T_{min}\,.$$

There β_{-1} – is the index of the workload proportionality in formula (7), generalized could be presented as:

$$\beta_{-1} = \frac{c^{N-1}}{\max\left\{c^{N-1}\right\}} = N^{2-N}\,.$$

For homogeneous unified SDP architecture we assume that $\forall\beta_{-1} = \text{const}$. Let us define:

$$S(\lambda_{ij},\lambda_{ji}) = \sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_{ij}\cdot\lambda_{ji}\,,$$

$$D(c_{ij}^{N},c_{ji}^{N}) = \sum_{i=1}^{N}\sum_{j=1}^{N}c_{ij}^{N}\cdot c_{ji}^{N}\,.$$

Service workflow delivery time minimization $T \to T_{min}$, and in the limiting case $T \to 0$ at the optimal workflow management with loaded distributed service infrastructure could be described as minimization condition (7). Given the notation introduced, this condition will become a following view:

$$\lim_{\|\lambda\|\to\|c\|} T = 0\,,$$

and respectively:

$$\lim_{\|\lambda\|\to\|c\|}\left[\frac{\Lambda_0}{\gamma}\sum_{k=1}^{N}\sum_{l=1}^{N}\lambda_{lk}\cdot\lambda_{kl} - \frac{\beta_{-1}}{\gamma}\sum_{i=1}^{N}\sum_{j=1}^{N}c_{ij}^{N-1}\cdot c_{ji}^{N-1}\right] =$$

$$= \lim_{\|\lambda\|\to\|c\|}\left[\frac{\Lambda_0}{\gamma}S(\lambda_{lk},\lambda_{kl}) - \frac{\beta_{-1}}{\gamma}D(c_{ij}^{N-1},c_{ji}^{N-1})\right] = 0.$$

Implicating indexes of the adjacent matrix $\|c\|$ elements relatively to the service workflows interconnection indexes $\|\lambda\|$ in the limiting case $\|\lambda\| \to \|c\|$, $S(\lambda_{lk},\lambda_{kl}) \to D(c_{ij}^{N-1},c_{ji}^{N-1})$ we obtain the following polynomial functional equation:

$$\Lambda_0 S(\lambda_{ij},\lambda_{ji}) - \beta_{-1}\left[S(\lambda_{ij},\lambda_{ji})\right]^{\frac{1}{N-1}} = 0\,. \qquad (8)$$

This equation could be interpreted as optimality condition of the SDP service workflow management. By substituting to (8) the functional $S(\lambda_{ij},\lambda_{ji})$ we obtain a solution for optimal service workflows distribution task into SDP infrastructure:

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_{ij}\cdot\lambda_{ji} = \left(\frac{\Lambda_0}{\beta_{-1}}\right)^{\frac{2-N}{N-1}}\,. \qquad (9)$$

Given amounts λ of service workflows and some technical requirements, specified in SLA we could calculate all necessary SDP parameters in some scaled service infrastructure by equation with formula (2) substituted to (9). Thus, for correct and effective service workflow management we represent following criterion:

$$\sum_{k=1}^{N}\sum_{l=1}^{N}\lambda_{kl}\cdot\lambda_{lk} \Leftrightarrow \sum_{k=1}^{N}\sum_{l=1}^{N}\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\Delta\gamma_{ij}\cdot x_{kl}^{(i,j)}\right)^2 - \left(\frac{\Lambda_0}{\beta_{-1}}\right)^{\frac{2-N}{N-1}} \to \min \qquad (10)$$

where $x_{kl}^{(i,j)}$ – is the part of service workflow γ_{ij} between service nodes i and j that distributing from node k to l, this amount could be calculated on the each service node into SDP. The guard conditions of the service workflow $x_{kl}^{(i,j)}$ should be applied for (10):

$$\sum_{k=1}^{N}x_{kl}^{(i,j)} - \sum_{k=1}^{N}x_{lk}^{(i,j)} = \begin{cases}-1, l = i,\\ 0, l \ne i, j,\\ 1, l = j.\end{cases}$$

Thus, at $N \to \infty$ the service flows management realizing with minimal resources utilization $\overline{d_q}$, i.e. minimum Λ_0 or otherwise, at $\Lambda_0 \approx 1$ (a time unit of service processing).

Conclusions

Based on achieved results (10), we conclude that global-scalable increasing and complication of the SDP distributed structure, requires the fulfillment of the condition for optimal service flows management: the SDP should provide sufficient increase of the service processing resources for respective interaction indexes $\Delta\lambda_{lk}$ between user and service interfaces while some custom service application is in progress.

An effective and correct SDP configuring (both physical interfaces interconnection structure and workflow management) could be realized by calculation its necessary parameters to maximize service infrastructure productivity and provide service workflows distribution processes with minimal delays, that significantly increases QoS indexes of the scalable distributed service delivery system, avoids DoS failures when quantity of active users and their demands dramatically grows up. This work presents corresponding criterions, task solving in strict correspondence with formulated hypothesis by mathematical transformations with obtaining practically implementable formulas and approaches.

References

1. *Demydov I. Enterprise Distributed Service Network Architecture and Topology Optimization / I. Demydov, M. Klymash, N. Kryvinska, C. Strauss // J. Space-Based and Situated Computing. – 2012. – Vol. 2, № 1. – P. 23-30.*
2. *. Klymash M. Traffic routing in telecommunication nets and its diakoptics representation / M. Klymash, B. Strykhalyuk, M. Kaidan, I. Demydov // Computation Problems of Electrical Engineering.– 2011. - №1(1). – P. 15 - 1(http://jcpee.lp.edu.ua/images/authors/004_Klumash.pdf)*
3. *Kryvinska N. Analytical Approach for the Modeling of Real-time Services over IP Network Elsevier Transactions of IMACS / N. Kryvinska // Mathematics and Computers in Simulation, ISSN: 0378-4754. – 2008. – Vol.79. – P. 980 – 990.*
4. *Kryvinska N. Intelligent Network Analysis by Closed Queueing Models / N. Kryvinska // Telecommunication Systems, ISSN: 1018-4864 (Print) 1572-9451 (Online). – September 2004. – Vol. 27(1). – P. 85-98.*
5. *Papazoglou Michael P. Service-Oriented Computing: a Research Roadmap / Michael P. Papazoglou, P. Traverso, F. Leymann. // International Journal of Cooperative Information Systems. – 2008. – Vol. 17(2). – P. 223–255*
6. *Debanjan Saha , Sambit Sahu , Anees Shaikh, A service platform for on-line games, Proceedings of the 2nd workshop on Network and system support for games, p.180-184, May 22-23, 2003, Redwood City, California*
7. *Leonard Kleinrock: An early history of the internet [History of Communications]. IEEE Communications Magazine 48(8): 26-36 (2010).*
8. *Kendall, David G. (September 1953). "Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain". Annals of Mathematical Statistics 24 (3): 338–354. doi:10.1214/aoms/1177728975.*
9. *Berezko M.P., Vishnevskiy V.M., Levner Ye.V., Fedotov Ye.V. Mathematical models of the routing study in data transmission networks // Informational processes. – 2001. – Vol. 1, № 2. – P. 103–125. (http://ict.informika.ru/ft/004415/4.pdf)*
10. *Tymchenko O.V., Demydov I.V. Routing optimization of informational workflows in the net structures by the minimal packet delay criterion // J. on Institute for Modeling in Energy Engineering named after G. Pukhov, NASc of Ukraine. – Kiev.:2006. – Vol.37. –P.194–201.*
11. *Demydov I.V., Kryvinska N., Klymash M.M. An Approach to the Flexible Information/Service Workflow Managing in Distributed Networked Architectures // The International Workshop on Design, Optimization and Management of Heterogeneous Networked Systems (DOM-HetNetS'09), in conjunction with the 38th International Conference on Parallel Processing (ICPP-2009), September 22nd -25th, 2009. - Vienna, Austria, 2009. - P. 236-242.*

Поступила в редколлегию 23.06.2012

Рецензент: д-р техн. наук, проф. В.О. Пелішок, Національний університет «Львівська політехніка».

АНАЛІЗ РОЗПОДІЛУ СЕРВІСНИХ ПОТОКІВ ТА ПАРАМЕТРІВ ПЛАТФОРМ НАДАННЯ ПОСЛУГ

М.М. Климаш, І.В. Демидов, О.А. Лаврів, Ю.Д. Добуш

Розробка новітніх застосувань Е-бізнесу через інформаційні технології (ІТ) підкреслює проблеми з традиційним підходом до побудови мереж, характерним для комунікації останніх років. Щоб повною мірою скористатися перевагами бізнес-застосувань, побудованих на основі потужних платформ надання послуг (SDP), необхідно переконатися, що забезпечено QoS в процесі розподілу сервісних потоків і живучість та ефективність системи. Стаття спрямована на вирішення ключових технічних питань розподілу сервісних потоків із забезпеченням якості при розрахунку параметрів SDP, необхідних для забезпечення експлуатаційної надійності та дотримання угод про рівень обслуговування.

Ключові слова: *платформа надання послуг, якість обслуговування, розподіл потоків, сервісно-орієнтоване планування ресурсів.*

АНАЛИЗ РАСПРЕДЕЛЕНИЯ СЕРВИСНЫХ ПОТОКОВ И ПАРАМЕТРОВ ПЛАТФОРМ ПРЕДОСТАВЛЕНИЯ УСЛУГ

М.М. Клымаш, И.В. Демидов, О.А. Лаврив, Ю.Д. Добуш

Разработка новых приложений Е-бизнеса через ин-формационные технологии (ИТ) подчеркивает проблемы с традиционных подходом к построению сетей, характерным для коммуникации последних лет. Чтобы в полной мере пользоваться преимуществами бизнес-приложений, построенных на основе мощных платформ предоставления услуг (SDP), необходимо убедиться, что обеспечены QoS в процессе распределения сервисных потоков, живучесть, а также эффективность системы. Статья направлена на решение ключевых технических вопросов распределения сервисных потоков с обеспечением качества при расчете параметров SDP, необходимых для обеспечения эксплуатационной надежности и соблюдения соглашений об уровне обслуживания.

Ключевые слова: *платформа предоставления услуг, качество обслуживания, распределение потоков, сервис-ориентированное планирование ресурсов.*