

УДК 621.397

МОНИТОРИНГ СОСТОЯНИЯ ЭКЗЕМПЛЯРОВ СЕРВИСОВ И МЕТОД ОБЕСПЕЧЕНИЯ ТРЕБУЕМОГО УРОВНЯ НАДЕЖНОСТИ



[Е.А. КОПЫТОВА](#),
[САИД ХАЛАВА ФАУАЗ](#)
Харьковский национальный
университет радиоэлектроники

Запропоновано метод забезпечення надійності надання послуг в мультисервісній мережі, оснований на сервіс-орієнтованій архітектурі. Даний метод дозволяє підвищити ефективність роботи системи мережного управління шляхом реалізації вибору екземпляра сервісу на основі інформації про стан екземплярів, отриманій системою моніторингу.

The method of ensuring reliability of service in multi-service networks based on service-oriented architecture was proposed. This method allows to increase the efficiency of network management through the implementation of choice based on the service instance information on copies of obtained by the monitoring system.

Предложен метод обеспечения надежности предоставления услуг в мультисервисной сети, основанной на сервис-ориентированной архитектуре. Данный метод позволяет повысить эффективность работы системы сетевого управления путем реализации выбора экземпляра сервиса на основании информации о состоянии экземпляров, полученной системой мониторинга.

Введение

В настоящее время в Украине развитие телекоммуникаций происходит в направлении расширения спектра предоставляемых услуг, что влечет за собой переход к мультисервисным сетям. Одной из важнейших задач развития мультисервисных сетей является обеспечение требуемого уровня качества обслуживания (Quality of Service, QoS). Согласно стандартам организации OASIS (Organization for the Advancement of Structured Information Standards) [1] надежность является одним из показателей QoS. Следовательно, для достижения требуемого уровня QoS необходимо обладать средствами управления надежностью предоставляемых сервисов. В настоящее время провайдеры обладают достаточно ограниченным набором средств в данной области, а достижение требуемого уровня надежности, как правило, обеспечивается простым дублированием ресурсов. Для конечных пользователей это приводит к увеличению стоимости услуг.

Практическая реализация управления услугами и бизнесом современных мультисервисных сетей на сегодняшний день выполняется в рамках нескольких технологических подходов: CORBA, COM, DCOM, SOA [2, 3, 4]. Каждый из указанных

подходов обладает своими особенностями предоставления услуг (сервисов). При этом вопросы обеспечения надежности предоставления услуг в данных технологиях практически не затрагиваются.

Среди существующих подходов, применяемых для организации взаимодействия элементов сети, наиболее распространенным является подход, основанный на сервис-ориентированной архитектуре (Service-Oriented Architecture, SOA) [2]. В данном подходе при публикации сервиса в репозитории провайдером заполняется лишь декларированный уровень надежности, который можно существенно изменить в процессе работы. Однако это негативно сказывается на обеспечении заданного качества обслуживания, что является существенным недостатком SOA.

При проектировании системы сетевого управления мультисервисных сетей на основе SOA гарантируется, что заявленные требования к надежности будут выполнены [5]. Это относится к стабильному, рабочему состоянию системы, но учитывая возможность сбоев, возникает необходимость в разработке метода, позволяющего управлять надежностью предоставляемых услуг в мультисервисных сетях.

I. Постановка проблемы и анализ публикаций

При решении задачи обеспечения требуемого уровня надежности предоставления услуг современной сетью можно выделить несколько ключевых вопросов, требующих решения.

Первый заключается в том, что существующие системы сетевого управления основаны на использовании протокола SMNP [6] и позволяют выполнять мониторинг только параметров элементов сетевой инфраструктуры. При этом такой мониторинг не позволяет оценить взаимные влияния параметров и режимов работы отдельных элементов, что существенно затрудняет диагностику и поиск причин ухудшения работы сети.

Второй вопрос заключается в том, что системы сетевого управления не позволяют оценивать параметры QoS при предоставлении сервисов, что существенно затрудняет решение задачи управления конфигурацией сервисов.

Учитывая описанные трудности, существующие системы сетевого управления обеспечивают предоставление услуг с требуемым качеством путем резервирования ресурсов, что влияет на стоимость услуг для конечного пользователя.

Следовательно, для предоставления услуг с требуемым качеством необходимо, чтобы система сетевого управления позволяла:

- проводить мониторинг всех элементов сети, что позволит оценить взаимные влияния параметров QoS и режимов работы всех элементов мультисервисной сети;
- упростить задачи управления конфигурацией сервисов на основании результатов мониторинга всех элементов сети;
- обеспечить требуемый уровень надежности без резервирования дополнительных ресурсов.

Существующие системы сетевого управления мультисервисными сетями не могут обеспечить выполнение данных требований. Следовательно, возникает необходимость в разработке метода обеспечения требуемого уровня надежности, который позволит усовершенствовать существующие системы сетевого управления мультисервисными сетями с учетом выдвинутых требований.

В настоящей работе предлагается метод обеспечения требуемого качества обслуживания, который включает в себя алгоритм диагностики экземпляров сервисов и обработки информации о состоянии узлов мультисервисной сети с последующим выбором экземпляра сервиса, на который перенаправляется запрос пользователя. Предлагаемый метод основан на системе сетевого управления мультисервисными сетями на базе SOA [7].

II. Метод обеспечения требуемого уровня надежности предоставления услуг в мультисервисных сетях

При разработке метода обеспечения требуемого уровня надежности предоставления услуг в мультисервисных сетях предполагается наличие специального сервера мониторинга [8], который располагается на проху-сервере. Предлагаемый метод включает следующие алгоритмы: алгоритм мониторинга экземпляров сервисов, алгоритм работы проху-сервера и алгоритм работы системы мониторинга при выборе экземпляра сервиса.

Алгоритм мониторинга экземпляров сервисов.

В рамках предлагаемого метода сервер мониторинга осуществляет мониторинг экземпляров сервиса согласно предложенному алгоритму, на основании которого сервер мониторинга посылает тестовые запросы на экземпляры сервисов, предоставляющие запрашиваемые услуги, и сравнивает полученный ответ с эталонным. Если ответ идентичен, то система мониторинга пошлет следующий запрос на экземпляр сервиса через установленное время ($ServiceCheckInterval, SChI$). Также система обнуляет значение SCh_f ($ServiceCheckFail$). Если ответ от сервера отличается от эталонного, то система увеличивает значение счетчика SCh_f и сравнивает это значение с пороговым ($ServiceCheckFailMaximum, SCh_{fmax}$). Пороговое значение SCh_{fmax} зависит от типа сервиса, мониторинг которого выполняется. Это необходимо для того, чтобы избежать «ложной тревоги» о неработоспособности сервиса. Если $SCh_f = SCh_{fmax}$, то система генерирует предупреждение о том, что заданный экземпляр сервиса неработоспособен. Новая попытка проверки сервиса выполняется после истечения таймера $SChI$. На рис. 1 представлен предложенный алгоритм мониторинга экземпляров сервисов.

При реализации данного алгоритма мониторинга важным является выбор интервала мониторинга, так как некорректный выбор может привести либо к избытку служебного трафика в сети и как следствие — к чрезмерным задержкам, либо к несвоевременному обнаружению сбоев и повышению вероятности потерь.

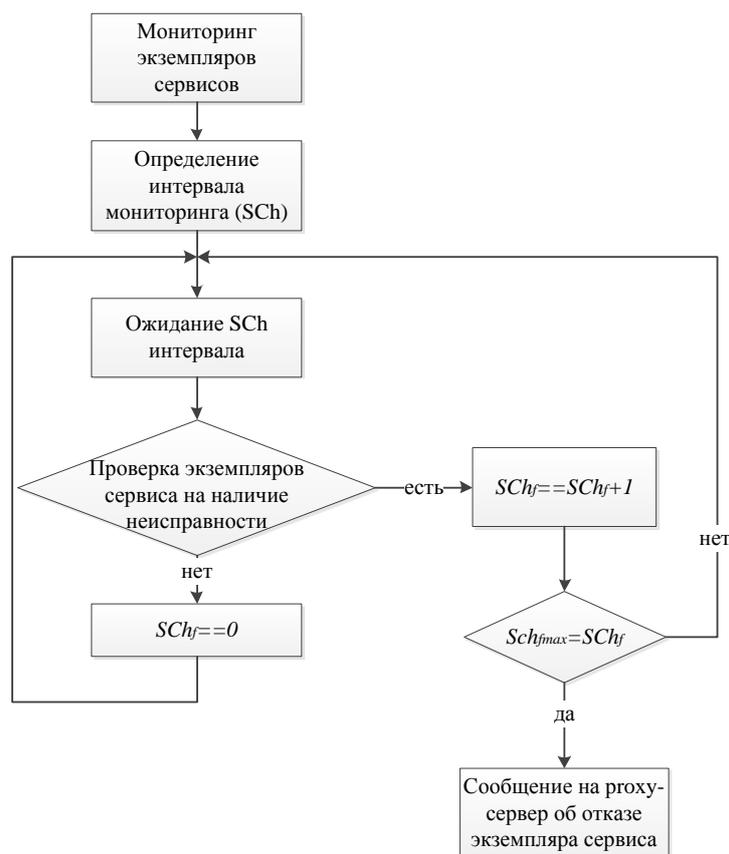


Рис. 1. Алгоритм мониторинга экземпляров сервисов

Определение частоты мониторинга системы

В результате анализа логики функционирования системы сетевого управления и проведя декомпозицию состояний системы, получен вероятностно-временной граф (ВВГ) (рис.2) [9]. Основой для построения ВВГ является граф состояний системы сетевого управления мультисервисными сетями на базе WBEM (Web-Based Oriented Architecture), отражающий состояния системы в работоспособном состоянии, от ожидания запроса до формирования ответа пользователю. В системе сетевого управления на базе WBEM, согласно [7], таких состояний семь.

В данном графе используется сквозная нумерация состояний. Состояния, описывающие процесс распределения потоков при функционировании сети без возникновения неисправностей, представлены состояниями 1 - 7. Состояния явных отказов - состояний 8 - 12, определены числом типов неисправностей, число неработоспособных состояний количественно равны числу состояний явных отказов – состояния скрытых отказов (13 - 17), и состояние мониторинга системы – состояние 18.

Переходы между работоспособными состояниями размечены как $\mu_1 - \mu_7$. Интенсивности переходов в состояние диагностики системы размечены как λ_p .

Значения интенсивностей обработки запросов системой в работоспособном состоянии $\rho_1 - \rho_7$ при отсутствии задержек в обслуживании равны между собой.

Вероятность того, что сеть останется в рабочем состоянии, обозначена как P_p , P_m – вероятность перехода в режим мониторинга сети, P_c – вероятность возникновения сбоев в работе сети, P_b – вероятность восстановления работоспособности элементов сети, $\lambda_i D$ – интенсивность переходов в состояние отказов ($i=1,2,..,5$).

Путем эквивалентных преобразований этот ВВГ (рис. 2) приводится к виду, представленному на рис. 3, где $s1$ – состояние ожидания системы, $s2$ – состояние обработки запроса, $s3$ – состояние явных отказов, $s4$ – состояние неявных отказов, $s5$ – состояние мониторинга системы.

Функции переходов для ВВГ (рис. 3) имеют следующий вид:

$$f_1(z) = (P_{c2}z^{1/\lambda 1D} + P_{c2}(z^{1/\lambda 2D} + z^{1/\lambda 5D}))P_p z^{1/\mu 2} P_{c3}(z^{1/\lambda 2D} + z^{1/\lambda 3D} + z^{1/\lambda 4D}) \times \\ \times P_p z^{1/\mu 3} P_{c4}(z^{1/\lambda 1D} + z^{1/\lambda 5D})P_p z^{1/\mu 4} P_{c5}(z^{1/\lambda 1D} + z^{1/\lambda 5D})P_p z^{1/\mu 5} P_{c6} \times \\ \times (z^{1/\lambda 2D} + z^{1/\lambda 3D} + z^{1/\lambda 4D})P_p z^{1/\mu 6} P_{c7}(z^{1/\lambda 1D} + z^{1/\lambda 5D});$$

$$f_2(z) = (P_{c2}z^{1/\lambda 1D} + P_{c2}(z^{1/\lambda 2D} + z^{1/\lambda 5D}))P_p z^{1/\mu 2} P_{c3}(z^{1/\lambda 2D} + z^{1/\lambda 3D} + z^{1/\lambda 4D}) \times \\ \times P_p z^{1/\mu 3} P_{c4}(z^{1/\lambda 1D} + z^{1/\lambda 5D})P_p z^{1/\mu 4} P_{c5}(z^{1/\lambda 1D} + z^{1/\lambda 5D})P_p z^{1/\mu 5} P_{c6} \times \\ \times (z^{1/\lambda 2D} + z^{1/\lambda 3D} + z^{1/\lambda 4D})P_p z^{1/\mu 6} P_{c7}(z^{1/\lambda 1D} + z^{1/\lambda 5D});$$

$$f_3(z) = P_{c8}z^{1/\rho 1} + P_{c9}z^{1/\rho 2} + P_{c10}z^{1/\rho 3} + P_{c11}z^{1/\rho 4} + P_{c12}z^{1/\rho 5};$$

$$f_4(z) = P_{c13}z^{1/\lambda 1D} + P_{c14}z^{1/\lambda 2D} + P_{c15}z^{1/\lambda 3D} + P_{c16}z^{1/\lambda 4D} + P_{c17}z^{1/\lambda 5D};$$

$$f_5(z) = (P_{c13} + P_{c14} + P_{c15} + P_{c16} + P_{c17})z^{1/\lambda d}.$$

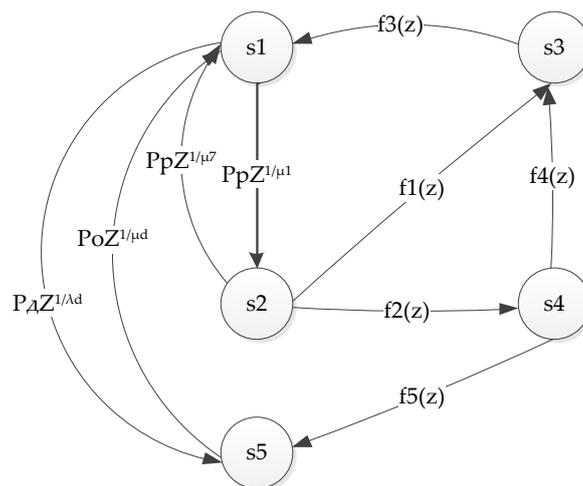


Рис. 3. Промежуточные преобразования вероятностно-временного графа функционирования системы сетевого управления на основе WBEM

Путем эквивалентных преобразований этот ВВГ (рис. 3) будет приведен к виду, представленному на рис. 4, где s_i – состояние отказа.

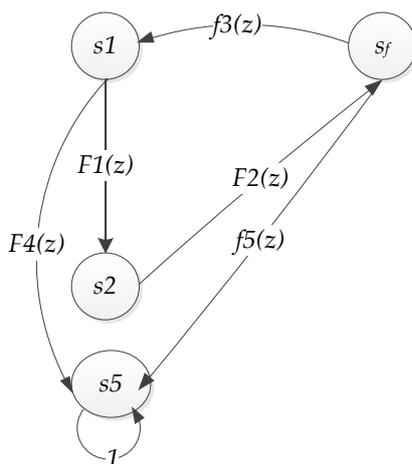


Рис.4. Окончательный вид преобразованного ВВГ

Производящая функция графа, представленного на рис. 4, имеет вид:

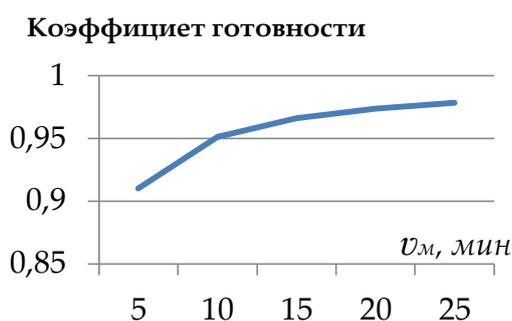
$$F(z) = F1(z) + F2(z) + f3(z) + F4(z) + f5(z).$$

Основываясь на вероятностно-временном графе (рис.4), время нахождения системы в состоянии мониторинга равно:

$$T_m = \frac{dF_m(z)}{dz} \Big|_{z=1},$$

где $F(z) = F4(z) + f5(z)$.

На основании аналитического моделирования по предложенному ВВГ проведен расчет времени нахождения сети в состоянии мониторинга, что позволяет определить частоту мониторинга системы, а также ее влияние на вероятности отказов и коэффициент готовности. Результаты моделирования представлены на рис.5.



а)



б)

Рис. 5. Зависимость коэффициента готовности (а) и вероятности отказов системы (б) от интервала мониторинга (v_m)

Анализ полученных данных показал, что с увеличением интервала проведения мониторинга системы увеличивается коэффициент готовности, но при этом возрастает

вероятность отказов. Следовательно, при выборе интервала мониторинга системы на предмет надежности предоставляемых услуг ключевыми являются требования рассматриваемого типа трафика к вероятности потерь и коэффициенту готовности.

Алгоритмы выбора экземпляра сервиса

В рамках предложенного метода для обеспечения заданного уровня надежности предлагаются алгоритмы выбора экземпляра сервиса для распределения запросов в системе. Согласно предложенной модели управления [3] запрос от frontend-сервера передается на проху-сервер, который выполняет задачу распределения запросов по экземплярам сервиса. Алгоритм работы проху-сервера приведен на рис.6.



Рис. 6. Алгоритм работы проху-сервера при выборе экземпляра сервиса

Задачей системы мониторинга является определение наименее загруженного экземпляра сервиса и передача данных проху-серверу для выбора экземпляра сервиса. Для решения данной задачи предлагается алгоритм работы системы мониторинга, представленный на рис. 7.

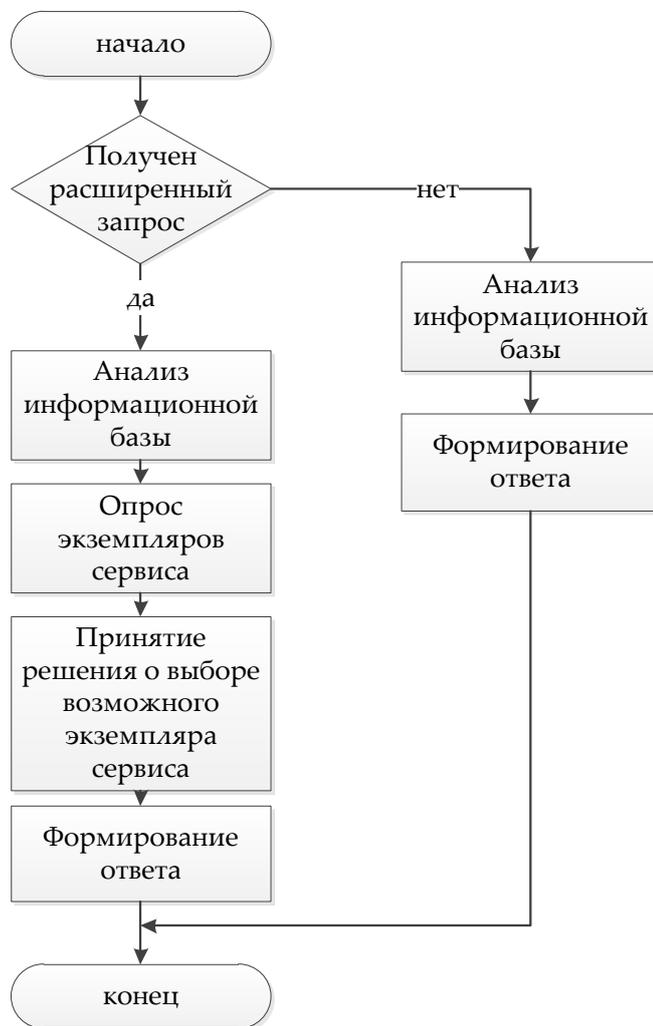


Рис. 7. Алгоритм работы системы мониторинга при выборе экземпляра сервиса

Система может получить запросы двух типов: обычный и расширенный. При получении обычного запроса производится поиск необходимых данных в информационной базе. В случае удачного поиска система передает требуемые данные проху-серверу, в обратном случае — отрицательный ответ. При получении расширенного запроса система мониторинга производит поиск в информационной базе своего узла и посылает запрос работающим экземплярам сервисов для определения степени их загруженности в настоящий момент времени. После обработки полученной информации о состоянии экземпляров сервисов, система мониторинга принимает решение и отправляет соответствующий ответ проху-серверу.

Основной задачей предложенных алгоритмов является обеспечение максимальной эффективности работы системы при выборе экземпляра сервиса. В качестве

такого показателя эффективности может выступать некая функция f , которая связывает между собой время поиска данных и загруженность системы связи.

Оценить загруженность сети можно произвести с использованием коэффициента относительного объема служебного трафика:

$$k_{от} = \frac{V_c}{V_{общ}},$$

где V_c — служебный трафик в системе поиска текущих данных,
 $V_{общ}$ — общий трафик системы.

Временные параметры можно оценить следующим коэффициентом:

$$k_{вр} = \frac{t_n}{t_{обр}},$$

где t_n — время выбора экземпляра сервиса,
 $t_{обр}$ — время обработки соответствующего запроса.

Следовательно, общий вид функции эффективности определяется особенностями системы и относительной важностью учитываемых параметров. В простейшем случае функция может быть линейной:

$$f = \alpha k_{зг} + \beta k_{вр},$$

где α , β — коэффициенты, учитывающие относительную важность каждого параметра. Для более точного учета влияния указанных параметров на эффективность работы системы управления возможно использование функций более сложного вида.

Выводы

В статье предложен метод обеспечения надежности предоставления услуг в мультисервисной сети, основанной на веб-ориентированной архитектуре. В отличие от известных предложенный метод основан на мониторинге и анализе состояния мультисервисной сети на уровне управления сервисами. Состояние экземпляра сервиса проверяется системой мониторинга, которая реализуется на проху-сервере. Использование данного подхода позволит поддерживать требуемый уровень надежности предоставления сервиса без повышения требований к ресурсам сети.

Одним из ключевых вопросов при реализации систем управления сервисами является выбор частоты формирования тестовых запросов. Предложенный метод позволяет определить частоту мониторинга мультисервисной сети. Разработанный метод также позволяет оценить влияние объема передаваемого служебного трафика на коэффициент готовности и вероятность отказов сервисов в мультисервисной сети. Согласно проведенному анализу с увеличением интервала проведения мониторинга сети увеличивается коэффициент готовности, однако это негативно сказывается на вероятности отказов. Следовательно, при выборе интервала мониторинга системы

ключевыми являются требования рассматриваемого типа сервиса к вероятности потерь и коэффициенту готовности.

Список литературы:

1. OASIS CIQ V3.0 Party Committee Specification 02. — OASIS, 2008. — 70 p.
2. Бибештейн Н., Боуз С., Джонс К. Компас в мире сервис-ориентированной архитектуры (SOA): ценность для бизнеса, планирование и план развития предприятия. — М.: КУДИЦ-ПРЕСС, 2007. — 256 с.
3. Гребешков А.Ю. Стандарты и технологии управления сетями связи. — СПб.: ЭкоТрендз, 2003. — 288 с.
4. Oberg R.G. Understanding and Programming COM. — М: «Вильямс», 2000. — 480 p.
5. Kharchenko V., Odarushenko V., Odarushenko E. Basic Multifragment Macro models For Reliability Assessment Of Fault-Tolerant Computer Systems // Radio-electronic Computer System. — 2006. — №5(17). — P.105 – 109.
6. Stallings W. SNMP, SNMP v2, SNMP v3, and RMON 1 and 2. Third edition. — Addison-Wesley, 1998. — 619 p.
7. Копытова Е.А., Гладий Л.В. Архитектура управления сервисами в WEB-центрированной системе // Материалы IX международной научно-технической конференции «Перспективные технологии в средствах передачи информации ПТСПИ'2011». Том.1. — Владимир: ВлГУ, 2011. — С. 57 – 60.
8. Chan P., Lyu M., Malek M. Making Services Fault Tolerant // Service Availability: Third International Service Availability Symposium, ISAS 2006, Helsinki, Finland, May 15-16, 2006 : Revised Selected Papers. Lecture Notes in Computer Science. Vol. 4328. — Berlin Heidelberg: Springer, 2006. — P. 43 – 61.
9. Лосев Ю.И., Бердников А.Г., Гойхман Э.Ш., Сизов Б.Д. Адаптивная компенсация помех в каналах связи / Под.ред. Ю.И. Лосева. — М.: Радио и связь, 1988. — 208 с.