

МОДЕЛИРОВАНИЕ ПОДСИСТЕМЫ ТЕРМИНАЛЬНОЙ СЕССИИ УСЛУГИ ТИПА «ВИРТУАЛЬНЫЙ РАБОЧИЙ СТОЛ»

© 2016 г. А.А. СУЛЕЙМАНОВ

Московский технический университет связи и информатики

Облачные услуги давно стали одним из основных направлений развития современных инфокоммуникаций. Такие услуги находят свое применение в бизнесе, государственных учреждениях, промышленности, мультимедиа и т.д. Находя различные сферы применения и предоставляя различные варианты реализаций, облачные услуги, однако, имеют общую задачу предоставления сервиса пользователю с обеспечением приемлемого качества.

В данной работе рассмотрена облачная услуга типа «виртуальный рабочий стол» (англ. Desktop as a Service, DaaS), основная идея которой заключается в предоставлении пользователю полноценного рабочего места «из облака» на терминальное пользовательское устройство. Реализации услуги могут быть различными, одной из наиболее популярных стала схема, при которой каждому пользователю предоставляется доступ к его индивидуальной виртуальной машине, запущенной на сервере провайдера. Таким образом, рассматриваемая услуга подразумевает замену пользовательских персональных компьютеров на терминальные устройства – тонкие клиенты (малошумящие, компактные рабочие станции, предназначенные для подключения к серверу услуги). В качестве терминального устройства могут выступать также смартфоны, планшетные или стационарные компьютеры при условии наличия на них программного DaaS агента.

Доступ к виртуальной машине предоставляется по сети передачи данных и осуществляется за счет протокола доставки виртуального рабочего стола. Каждая фирма-производитель программного обеспечения для облачных систем, разработала и развивает собственный протокол. Более подробно такие протоколы рассмотрены в [1]. Они работают поверх TCP, используют алгоритмы компрессии/декомпрессии изображений при передаче изображений рабочего стола, имеют специализированные механизмы коррекции ошибок.

Основными клиентами услуги являются офисные сотрудники, операторы контакт-центров, госслужащие, сотрудники технической поддержки и др. Обычно такие пользователи работают с приложениями, которые требуют выделения некоторых заранее предсказуемых ресурсов сервера, пользовательского терминала и сети. Однако имеются и такие пользователи услуги, которые в силу специфики своей работы вынуждены активно использовать мультимедиа контент: сервисы видео-телефонии, видеоролики высокого качества, онлайн-игры. Следовательно, провайдеру необходимо предоставлять таким пользователям большие ресурсы (сервер, сеть, пользовательское устройство).

В общем виде терминальная сессия между сервером и клиентом может быть охарактеризована, как непрерывная передача снимков удаленного рабочего стола, развернутого в «облаке», на пользовательское устройство и передача команд, кодирующих события движения мыши и нажатия клавиш клавиатуры в обратном направлении. Внутри системы «сервер – клиентские терминалы» можно логически выделить подсистему инициализации терминальной сессии и подсистему терминальной сессии. Аналитическая модель первой была показана в [2]. В подсистему терминальной сессии

входят следующие основные компоненты: агент пользовательского устройства, облачная платформа, служба каталогов, агент виртуальной машины, файловый репозиторий. Упрощенная схема подсистемы терминальной сессии показана на рис. 1.

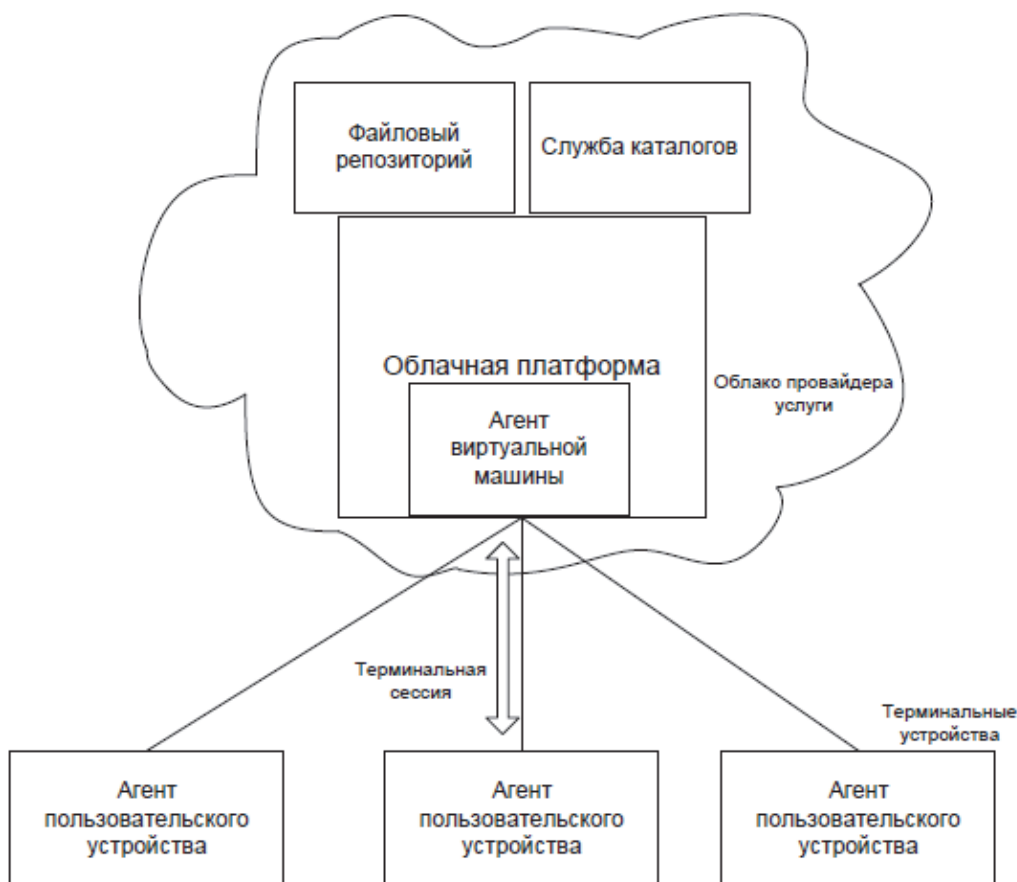


Рис. 1. Упрощенная схема подсистемы терминальной сессии.

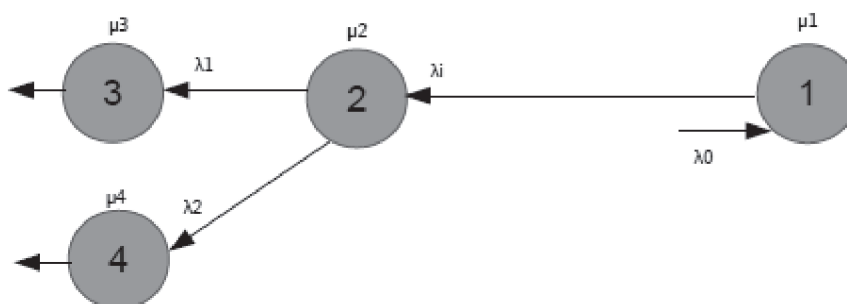


Рис. 2. Граф CeMO.

Для построения математической модели будем использовать сеть массового обслуживания (CeMO), известную, как сеть Джексона [3]. Согласно аппарату сетей Джексона в узлах сети расположим системы массового обслуживания (СМО) вида M/M/1 с дисциплиной FCFS. Модель терминальной сессии, представленная графом на рис. 2, содержит следующие компоненты:

- 1 – вершина, описывающая агент ВМ;
- 2 – вершина, описывающая агент пользовательского устройства;
- 3, 4 – вершины, описывающие ядра процессора.

Введем следующие обозначения:

λ_0 – интенсивность входящего в узел 1 потока;

λ_i – интенсивность входящего в узел 2 потока;
 λ_1 – интенсивность входящего в узел 3 потока;
 λ_2 – интенсивность входящего в узел 4 потока;
 μ_1 – интенсивность обслуживания узла 1;
 μ_2 – интенсивность обслуживания узла 2;
 μ_3 – интенсивность обслуживания узла 3;
 μ_4 – интенсивность обслуживания узла 4;
 p_{ij} – переходные вероятности.

Приведем далее аналитические соотношения для временных характеристик сети. Будем полагать, что режим работы сети стационарный, следовательно, интенсивности поступающего в узел i потока и выходящего из него потока должны совпадать ($\lambda_0 = \lambda_i$). Система уравнений равновесия согласно [4] будет иметь вид:

$$\lambda_j = \sum_{i=0}^4 \lambda_i \cdot p_{ij}, j = \overline{0,4}$$

Среднее число заявок в узле i обозначим N_i . Его можно определить следующим образом согласно [5]: $N_i = \frac{\rho_i}{1 - \rho_i}$. Среднее время пребывания заявки в узле T_i можно

найти из формулы Литтла: $\lambda_i \cdot T_i = N_i$. Откуда $T_i = \frac{N_i}{\lambda_i} = \frac{1}{\mu_i - \lambda_i}$, где $\rho_i = \frac{\lambda_i}{\mu_i}$. Среднее

время ожидания в узле i : $W_i = T_i - \frac{1}{\mu_i} = \frac{\rho_i}{1 - \rho_i} \cdot \frac{1}{\mu_i}$. Суммарное среднее число заявок во

всей сети можно найти из формулы: $N = \sum_{i=1}^4 N_i$.

Под временем отклика будем понимать среднее время, проходящее с момента вхождения заявки в систему до момента ее выхода обслуженной. Именно это время является одним из наиболее важных параметров услуги, поскольку напрямую влияет на воспринимаемое пользователем качество услуги. Суммарное среднее время пребывания заявки во всей сети (время отклика) T можно выразить из: $\lambda_0 \cdot T = N$, тогда

$$T = \frac{N}{\lambda_0} = \frac{1}{\lambda_0} \cdot \sum_{i=1}^4 \frac{\lambda_i}{\mu_i - \lambda_i} \quad (1)$$

Рассмотрим численный пример, используя значения, близкие к реальным. Пусть $\lambda_0 = 3.3 \text{ с}^{-1}$, $W_1 = 0.1 \text{ с}$, $T_1 = 0.3 \text{ с}$. Пусть по условию задачи узел 2 распределяет поступившие на него заявки на узлы 3 и 4 поочередно. Тогда переходные вероятности p_{23} и p_{24} будут равны 0.5.

Из системы уравнения равновесия найдем $\lambda_1 = \lambda_2 = \lambda_i \cdot p_{23} = 3.3 \cdot 0.5 = 1.65 \text{ с}^{-1}$.

Найдем $\mu_1 = \frac{1}{T_1 - W_1} = 5 \text{ с}^{-1}$. Далее рассчитаем среднее время отклика по фор-

муле (1), используя различные значения μ_2 , μ_3 и μ_4 . График зависимости среднего времени отклика от времени обслуживания в узле 2 (агент пользовательского устройства) показан на рис. 3. Этот узел выбран исходя из того, что на практике изменения, направленные на ускорение работы услуги на пользовательской стороне, как правило, вносят именно на этом узле, поскольку процессоры имеют конкретные спецификации и их параметры редко могут быть изменены. Разработкой и дальнейшим улучшением агента пользовательского устройства занимаются производители программного обеспечения облачных услуг.

Таким образом, в первом приближении показана математическая модель подсистемы терминальной сессии услуги «виртуальный рабочий стол». Параметры подсистемы описаны в терминах теории массового обслуживания, получены выражения

для среднего времени пребывания заявки в узле, времени ожидания в узле, времени отклика.

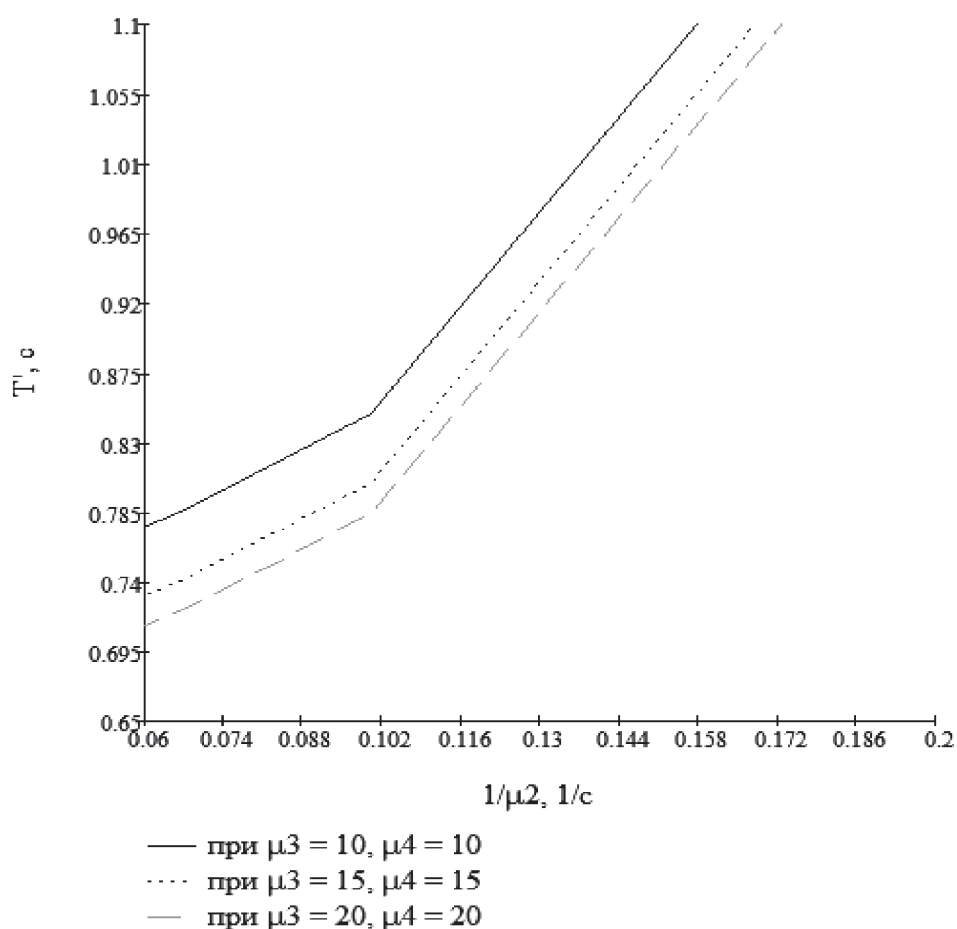


Рис. 3. График зависимости среднего времени отклика от времени обслуживания в агенте пользовательского устройства.

Дальнейший анализ характера зависимостей между найденными параметрами поможет сформулировать методику обеспечения приемлемого качества услуги, который может быть применен провайдером на этапе проектирования и эксплуатации инфраструктуры.

СПИСОК ЛИТЕРАТУРЫ

1. Сулейманов А.А. Воспринимаемое качество при использовании тонкого клиента на базе облачных платформ // «ИНФОКОМ-2014» / Труды Северо-Кавказского филиала Московского технического университета связи и информатики, 22-25 апреля 2014 г. Ростов-на-Дону.: ПЦ «Университет», 2014, ч. 1, с. 124 – 126.
2. Сулейманов А.А., Нетес В.А. Анализ времени подключения к облачной услуге "виртуальный рабочий стол" // Т-Comm: Телекоммуникации и транспорт. – 2016. – Т. 10. – № 7. – С. 41-46.
3. Jackson J.R. Networks of waiting lines / J.R. Jackson // Oper. Res. – 1957. – Vol. 5, № 4. – P. 518–521.
4. Бочаров П.П., Печинкин А.В. Теория массового обслуживания. М.: РУДН, 1995. – 529 с.
5. Virtamo J. Queueing Course. Complete lecture notes. 2005. Электронный ресурс, режим доступа: http://www.netlab.tkk.fi/opetus/s383143/kalvot/E_qnets.pdf