

Оценка среднего времени задержки в сетях IP с помощью тензорной методологии.

Пономарев Д.Ю.

18 сентября 2008 г.

Аннотация

В данной работе предлагается применить для оценки среднего времени задержки в сетях IP математический аппарат преобразования систем координат, рассматривая данные сети, как сети массового обслуживания, которые можно представить, как геометрические объекты, проекции которых в различных системах координат различны, но физические свойства самих объектов при этом не меняются. В качестве основного уравнения выбрано выражение, связывающее загрузку в системе и интенсивность поступления вызовов в систему через время обслуживания.

Abstract

There is present mathematical method for mean time delay analysis for IP networks. The mathematical model of these networks is queueing nets. Queueing networks may be define as geometrical objects which have various projections in different systems of coordinates but physical properties is not change. The basic equation is expression for system load definition by the use of stream intensity and time service.

На сегодняшний день независимо от технологий нижних уровней модели взаимодействия открытых систем (ВОС) основными протоколами, обеспечивающим процессы обработки информационных потоков в современных сетях связи (в т.ч. и с поддержкой механизмов поддержки качества обслуживания), является стек протоколов TCP/IP (Transmission Control Protocol/Internet Protocol). Массовый характер использования данного стека в качестве транспорта данных обусловлен как широким спектром и распространенностью оборудования с его поддержкой, так и тем, что его программная реализация предусмотрена в большинстве современных операционных систем. Однако, следует отметить, что данный этап развития современного общества предъявляет более высокие требования к объемам и качеству информации передаваемой по телекоммуникационным сетям. При построении любой сети связи, помимо выбора технологии передачи информации, определяется способы и возможности организации качественного обслуживания передаваемой информации. Для решения данной задачи в протоколе передачи данных по сетям связи предусмотрены механизмы обеспечения качества обслуживания (QoS – Quality of Service) [1]. Основной задачей данного сервиса является

разделение входящего потока информации по категориям (классам) обслуживания и, в зависимости от присвоенной категории, обслуживание данного информационного потока по определенным правилам. Одним из основных показателей оценки QoS является среднее время задержки [2].

Однако, использование только средств протокола IP недостаточно для обработки информационных потоков с заданными показателями качества обслуживания (QoS), так как в протокол IP изначально не была заложена необходимость обеспечения механизмов поддержки QoS в отличие от ATM. В дальнейшем, механизмы интегрального (IntServ) и дифференциального (DiffServ) обслуживания позволили решать задачи поддержки QoS в сетях IP [2, 3].

В сетях с IP протоколом могут использоваться следующие принципы поддержки QoS: на базе интегрированных служб (IntServ – Integrated Services Architecture) и на базе дифференцированных служб (DiffServ – Differentiated Services Framework). Механизм IntServ ориентирован больше на периферийное сетевое оборудование, а более масштабируемая архитектура DiffServ используется в магистральной части глобальных сетей. Модель IntServ применяется совместно с протоколом резервирования ресурсов (RSVP) в качестве рабочего сигнального протокола, а, следовательно, требует решения вопросов управления приоритетами запросов о резервировании ресурсов и проблемы резервирования в случае отсутствия поддержки протокола сигнализации некоторыми устройствами сети. На всем пути передачи информационного потока с определенной категорией QoS необходимо поддерживать как определение «сквозных» параметров сигнализации так и гибкость алгоритмов управления пропускной способностью, что определяется временным характером резервируемого ресурса и необходимостью периодических обновлений с помощью RSVP-сообщений. Модель IntServ для поддержки QoS определяет три класса обслуживания: гарантированное обслуживание, максимально доступное качество и контролируемая загрузка. В DiffServ по сравнению с IntServ акцент в большей степени сделан не на сигнализации, а на способах обработки потоков в каждом участке сети на основе заданного класса обслуживания данного информационного потока. Потоки согласно predetermined правилам классифицируются таким образом, что множество прикладных потоков объединяется в ограниченный набор классифицированных потоков.

Кроме рассмотренных технологий можно определить достаточно большой класс технологий и сетей, построенных на их базе, где существует задача нахождения вероятностно-временных характеристик систем с целью поддержания необходимого качества и надежности обслуживания. Кроме того, следует отметить, что сложность решения такой состоит не только в самих техноло-

гиях, но и в структурах сетей, использующих данные технологии, а также в необходимости принятия быстрых решений для задач управления ресурсами сети, как в локальном, так и глобальном масштабах.

Благодаря реализации механизмов QoS в сетях IP реализуется возможность обработки различных информационных потоков, источниками которых являются сетевые приложения, со своими требованиями к качеству обслуживания. Для обеспечения требований качества обслуживания каждого сетевого приложения необходимо знать какую пропускную способность и время задержки должна будет гарантировать сеть. Однако, проектирование рассматриваемых сетей, а в дальнейшем и их эксплуатация, может быть затруднена трудоемкостью расчетов параметров оборудования, при которых необходимое качество обслуживания будет поддерживаться на необходимом уровне, что усложняется как сложной топологией таких сетей, так и использованием различных дополнительных протоколов обработки на различных уровнях модели ВОС. Возможным решением этой проблемы может являться применение тензорного метода. Применение тензорного метода к расчету сетей работающих на базе стека протоколов TCP/IP позволит определить загрузку и задержку, создаваемую сетевыми приложениями в транзитных и оконечных узлах сети связи.

В данной работе предлагается применить для оценки вероятностно - временных характеристик (ВВХ) IP сетей математический аппарат преобразования систем координат, рассматривая данные сети, как сети массового обслуживания (СМО), которые в свою очередь рассматриваются, как геометрические объекты, проекции которых в различных системах координат различны, но физические свойства самих объектов при этом не меняются. Анализ сетей, основанный на данном подходе, получил название тензорного . Основателем тензорной методологии анализа систем является известный американский ученый и инженер Г. Крон. Дальнейшее развитие идеи тензорного анализа для информационных систем получили в работах Петрова А.Е, Арменского А.Е., Кузнецова О.Л., Петрова М.Н. и др.

Для инженерного анализа ВВХ сетей массового обслуживания в данной работе предлагается использовать в качестве инвариантного уравнения известное выражение для определения коэффициента использования устройств (ρ), дающее связь между интенсивностью поступления вызовов (λ) и средним временем обслуживания: $\rho = \lambda t_{\text{обсл}}$. В данной работе предлагается использовать метод анализа вероятностно-временных характеристик систем, основанный на следующих предположениях. Во-первых, что поток вызовов с одной и той же интенсивностью (λ) поступления вызовет при неизменной интенсивности обслуживания одну и ту же загрузку (ρ) устройств при изменении

структуры и можно считать, что будет выполняться соотношение (инвариант): $\rho\lambda = \rho'\lambda'$, где переменные со штрихом для одной структуры сети, без штриха для другой. Во-вторых, объединение систем в единую сеть не вызывает никаких изменений процесса обслуживания информационного потока. В-третьих, изменение структуры сети не предполагает качественное изменение основных соотношений между физическими величинами, описывающими простейший элемент, а определяет только их численное изменение. Данные предположения в литературе получили название постулатов обобщения.

Тогда, исходя из инварианта $\rho\lambda = \rho'\lambda'$, применив соотношение между интенсивностями исходной и примитивной сетей: $\bar{\lambda}' = \bar{C}\bar{\lambda}$ и используя известное соотношение $\rho = \lambda t_{\text{обсл}}$, где $t_{\text{обсл}}$ - среднее время обслуживания одного требования, можно записать: $\bar{\rho}'\bar{C}\bar{\lambda} = \bar{\rho}\bar{\lambda}$. Далее, находим соотношение между загрузками примитивной и исходной сетей, как $\bar{\rho} = \bar{C}^T\bar{\rho}'$ или $\bar{\rho}' = (\bar{C}^T)^{-1}\bar{\rho}$, и получаем: $(\bar{C}^T)^{-1}\bar{\rho} = \bar{t}'\bar{C}\bar{\lambda}$. Окончательно уравнение будет иметь вид: $(\bar{C}^T\bar{t}'\bar{C})\bar{\lambda} = \bar{C}^T\bar{\rho}'$, в котором исходная сеть описана в символах примитивной, а, следовательно, известной сети. Далее, решая полученное уравнение относительно λ , находим коэффициенты использования устройств в исходной сети ρ . Данный подход позволяет при минимальных затратах оценить загрузку сетей, обеспечив тем самым определение остальных характеристик сети: распределение вероятностей состояний по отдельным системам: $p_n = f(\rho)$, а также значения средней очереди \bar{N} и среднего времени задержки \bar{T} .

Кроме описанного выше контурного метода, можно использовать несколько другой подход, связанный с модификацией инвариантного уравнения $\rho = \lambda t_{\text{обсл}}$, а именно, в виде: $\lambda = \rho\mu$, где μ – интенсивность обслуживания. Определяя примитивную сеть, состоящую из такого же количества систем и описываемых инвариантным уравнением: $\bar{\lambda}' = \bar{\rho}'\bar{\mu}'$, находим матрицу перехода (\bar{A}) от одной проекции к другой: $\bar{\rho}' = \bar{A}\bar{\rho}$, и определяем составляющие матричного уравнения, как: $\bar{A}^T\bar{\mu}'\bar{A}\bar{\rho} = \bar{A}^T\bar{\lambda}'$.

Рассмотрим простейший пример применения второго метода (условно назовем его узловым) к расчету сети связи (рис. 1). К сети, состоящей из двух коммутаторов и двух маршрутизаторов, подключается абонентский концентратор (АК), от которого в пакетную сеть поступают информационные потоки от абонентских устройств.

На представленной ниже модели сети АК представлен в виде трех систем массового обслуживания, первая из которых соответствует входящему направлению, а две другие – исходящим направлениям к коммутаторам. Интенсивность потоков вызовов со стороны входа АК λ_a распределяется соответственно на интенсивности λ_{a1} и λ_{a2} , которые зависят от числа абонентов

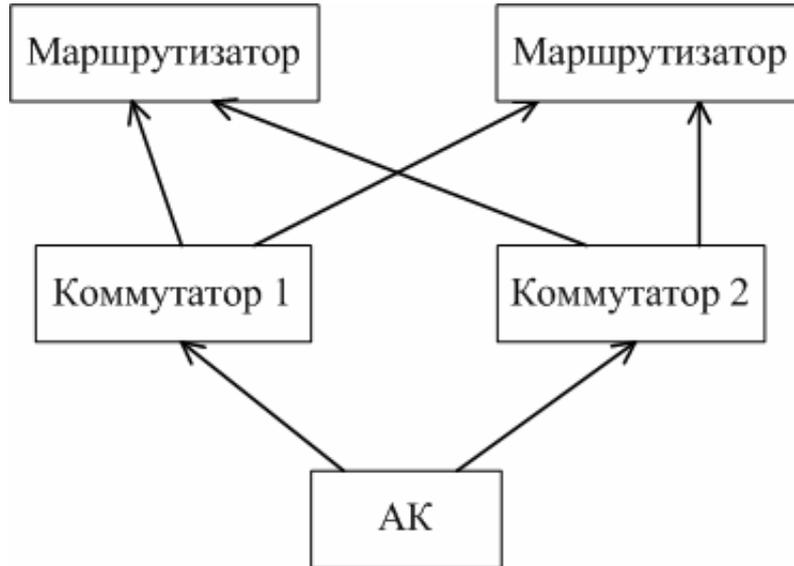


Рис. 1. Пример сети

подключенных к каждому коммутатору.

Интенсивность потоков вызовов поступающих на вход коммутаторов равны λ_{sw1} и λ_{sw2} . Так как в коммутаторе происходит разделение информационных потоков, выходу каждого коммутатора соответствует две системы массового обслуживания с интенсивностями λ_{s1} , λ_{s2} и потоков вызовов на маршрутизаторы λ_{sw12} , λ_{sw21} . На вход каждого коммутатора поступает информация от АК λ_{a1} или λ_{a2} . Интенсивность поступления вызовов на маршрутизаторы λ_{R1} и λ_{R2} . Таким образом, должно выполняться следующее условие:

$$\begin{cases} \lambda_a = \lambda_{a1} + \lambda_{a2} \\ \lambda_{sw1} = \lambda_{sw12} + \lambda_{s1} \\ \lambda_{sw2} = \lambda_{sw21} + \lambda_{s2} \\ \lambda_{R1} = \lambda_{s1} + \lambda_{sw21} \\ \lambda_{R2} = \lambda_{s2} + \lambda_{sw12} \end{cases}$$

Примитивная узловая сеть из n ветвей состоит из n не замкнутых обособленных ветвей. Определение компонент геометрических объектов примитивной сети:

$$\bar{\lambda}' = (\lambda_1 \lambda_2 \lambda_3 \lambda_4 \lambda_5 \lambda_6 \lambda_7 \lambda_8 \lambda_9 \lambda_{10} \lambda_{11})^T,$$

$$\bar{\rho}' = (\rho_1 \rho_2 \rho_3 \rho_4 \rho_5 \rho_6 \rho_7 \rho_8 \rho_9 \rho_{10} \rho_{11})^T,$$

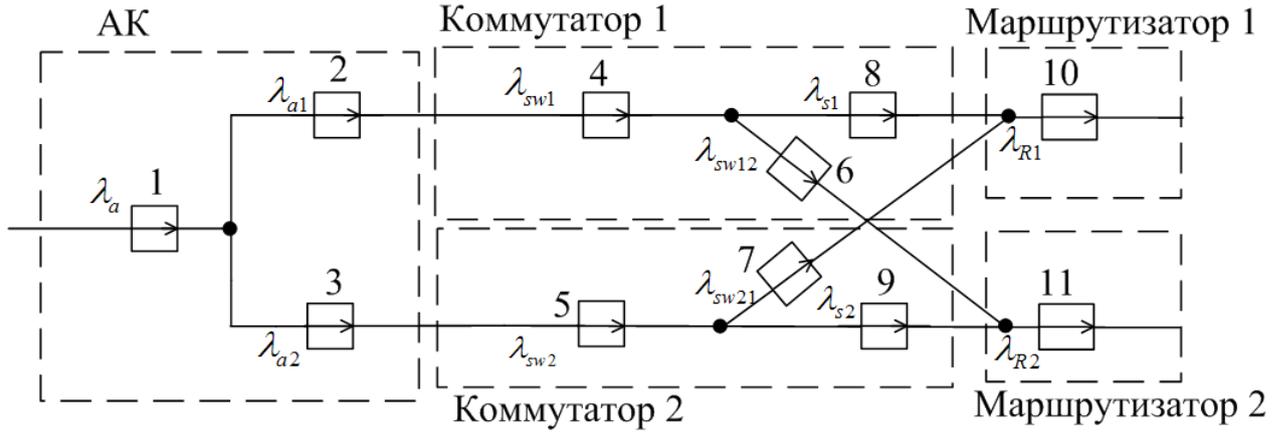


Рис. 2. Модель исследуемой сети

$$\bar{\mu}' = \begin{bmatrix} \mu_1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & \mu_2 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \mu_3 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \mu_4 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \mu_5 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu_6 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu_{11} \end{bmatrix}.$$

Находя соотношение между нагрузками ветвей примитивной сети и нагрузками в исходной сети, находим матрицу перехода :

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix}$$

Интенсивности потоков вызовов систем массового обслуживания для исходной сети находятся, как:

$$\bar{\lambda} = \begin{bmatrix} \lambda_1 - \lambda_2 - \lambda_3 \\ \lambda_2 - \lambda_4 \\ \lambda_3 - \lambda_5 \\ \lambda_4 - \lambda_6 - \lambda_8 \\ \lambda_5 - \lambda_7 - \lambda_9 \\ \lambda_6 + \lambda_9 - \lambda_{11} \\ \lambda_7 + \lambda_8 - \lambda_{10} \\ \lambda_{10} \\ \lambda_{11} \end{bmatrix}$$

Производя замену уравнения исходной сети $\bar{\lambda} = \bar{\mu}\bar{\rho}$ на уравнение $\bar{A}^T \bar{\mu}' \bar{A} \bar{\rho} = \bar{A}^T \bar{\lambda}'$, получаем систему линейных уравнений, позволяющую определить загрузки исходной сети, а далее могут быть найдены значения загрузок в ветвях при заданных значениях вероятностей поступления вызовов в отдельные системы массового обслуживания. При известных параметрах устройств с использованием рассчитанных тензорным методом значениях загрузки можно найти параметры QoS для каждого маршрута. Время задержки для маршрута будет определяться, как: $T = \sum_{i=1}^{\alpha} T_i$, где α – число систем обработки для маршрута; T_i – среднее время задержки в отдельной системе. Рассмотрим следующий случай для сети, представленной на рисунке 1. Вызов абонента, подключенный к первому коммутатору, должен быть обслужен на втором маршрутизаторе. Тогда информация будет проходить последовательно через системы массового обслуживания 1-2-4-6-11, и общее время задержки при передаче информации из конца в конец $T_{\text{инф}}$ равно сумме задержек на каждом участке маршрута (в данном случае, в качестве СМО рассматриваются системы М/М/1):

$$T_{\text{инф}} = T_1 + T_2 + T_4 + T_6 + T_{11} = \frac{1}{1-\rho_1} + \frac{1}{1-\rho_2} + \frac{1}{1-\rho_4} + \frac{1}{1-\rho_6} + \frac{1}{1-\rho_{11}}$$

При проектировании или эксплуатации телекоммуникационной сети, работающей на базе стека протоколов ТСР/IP согласно приведенным выше теоретическим выкладкам можно учитывать, как распределять потоки в сети, чтобы каждый поток получил требуемое качество обслуживания. На сегодняшний день в руководящих документах и рекомендациях МСЭ для сетей передачи данных по протоколу IP определены рекомендуемые нормы задержек для различных классов обслуживания в службах передачи данных [4].

В заключении, можно отметить следующее:

протокол IP является наиболее распространенным и позволяет объединить практически все существующие на сегодняшний момент услуги сетей с обеспечением заданного уровня QoS;

тензорный метод позволяет достаточно просто формализовать проектные

процедуры для сетей такого типа;

сложные топологии, необходимость учета особенностей передачи в различных технологиях, динамическое управление маршрутами потоков: всё это приводит к невозможности определения показателей качества классическими методами;

тензорный метод позволяет оценивать требуемые показатели качества при приемлемых вычислительных затратах.

Список литературы

1. Griffin, D.; Spencer, J.; Griem, J.; Boucadair, M.; Morand, P.; Howarth, M.; Ning Wang; Pavlou, G.; Asgari, A.; Georgatsos, P. Interdomain routing through QoS-class planes [Quality-of-Service-Based Routing Algorithms for Heterogeneous Networks] // IEEE Communications Magazine. – 2007. – Vol. 45(2). – pp. 88 – 95.
2. Jongtae Song; Mi Young Chang; Soon Seok Lee; Jinoo Joung. Overview of ITU-T NGN QoS Control // IEEE Communications Magazine. – 2007. – Vol. 45(9). – pp. 116 – 123.
3. J. Metzler. Accelerating a diverse mix of traffic across the wan // Webtorials: brief. – 2008. – August. – P.7.
4. Яновский Г. Г. Качество обслуживания в IP сетях // Вестник связи – 2008. – №1. – С. 65-74.

Благодарности

Работа поддержана грантом Президента Российской Федерации для государственной поддержки молодых российских ученых МК-2070.2008.9

Предметный указатель

ВВХ, 3
время задержки для маршрута, 7
загрузка, 3
инвариант, 4
матрица перехода, 6
рекомендуемые нормы задержек, 7
среднее время задержки, 2
тензорный анализ, 3
уравнение для контурного метода, 4
уравнение для узлового метода, 4
DiffServ, 2
IntServ, 2
QoS, 1
RSVP, 2