

МОДЕЛЬ МАРШРУТИЗАЦИИ В ГИБРИДНЫХ СЕТЯХ ДОСТАВКИ КОНТЕНТА С ГАРАНТИРОВАННЫМ КАЧЕСТВОМ ОБСЛУЖИВАНИЯ

Евсеева О.Ю., Мохаммед Б. Кадер

Харьковский национальный университет
радиоэлектроники, Украина.

E-mail: evseeva.o.yu@gmail.com

Abstract

Under growth of the Internet and increasing volume of video and other large-scale multimedia contents, key challenge is related to delivering the content to users. According the problem Content Delivery Network (CDN) and Peer-to-Peer (P2P) have been proposed and deployed. CDN and P2P have complementary advantages: high scalability and low cost of P2P and high reliability and good stability within CDN. As a result different hybrid approaches were proposed. In the article we'll assume CDN as basic architecture and that one user can connect to several content servers in same time (in parallel). One of the important problems that are solved in the network is request routing which combines choosing of the content servers (for sourcing) and traditional routing from the chosen servers to the user. The paper proposes a mathematical model which allows to formulate the problem of request routing in hybrid content delivery network as optimization problem and to resolve it in concordance with the traditional task of multipath routing in transport telecommunications network. In order to realize rigid guarantees proposed mathematical model includes tensor QoS-condition.

Растущая популярность таких сервисов, как, например, YouTube и Hulu, связанных с передачей потокового видео и загрузкой видеоконтента, по прогнозам ведущих аналитиков рынка телекоммуникационных услуг приведет к преобладанию данного типа трафика в сети Internet. Например, к 2017 году ожидается, что доля видео в суммарном объеме Internet-трафика составит 65% [1]. Особенность видеотрафика заключается, во-первых, в больших объемах передаваемых данных, во-вторых, в высоких скоростях передачи. На сегодняшний день основными решениями, обеспечивающими хранение и доставку мультимедийного контента потребителям, являются сети доставки контента (Content Delivery Network, CDN) и пиринговые сети (Peer-to-Peer, P2P). Идеология построения сетей CDN основана на использовании для хранения контента множества зеркалирующих серверов, что обеспечивает гарантированный доступ к контенту и стабильную скорость его загрузки. В то же время функционирование CDN связано с решением ряда задач, таких как репликация контента и синхронизация серверов, авторизация и маршрутизация запросов пользователей на тот или иной сервер, что ведет к усложнению системы и, соответственно, повышает стоимость предоставляемых услуг. С другой стороны, пиринговые сети, основанные на совместном использовании ресурсов всех подключенных в данный момент к сети пользователей (клиентов), наряду с неоспоримыми преимуществами, такими как масштабируемость и низкая стоимость, обладают рядом существенных недостатков: низкая надежность, неконтролируемая пропускная способность системы, невозможность гарантировать скорость загрузки контента, сложность контроля как самого контента, так и пользователей. Как результат, в литературе описана и на практике уже реализуется идея гибридного построения сетей доставки контента, предполагающая совместное использование CDN и P2P в рамках единой инфраструктуры, где пользователь может получить контент как с сервера CDN, так и от другого пользователя (пира) [2].

Поскольку на данный момент существуют различные варианты комбинирования CDN и P2P, при выборе архитектуры построения сети доставки контента будет исходить из необходимости гарантированного качества обслуживания пользователей, которое оценивается, среди прочего, через время отклика сервиса, надежность сервиса и скорость передачи запрашиваемого контента. Реализация услуги с гарантиями по данным показателям предполагает предпочтительное использование серверов в качестве источника контента.

С целью повышения качества обслуживания, снижения нагрузки на отдельно взятые сервера (балансировка загруженности серверов) и сбалансированной загруженности телекоммуникационной сети, обеспечивающей непосредственную транспортировку трафика, дополним традиционную CDN возможностью подключения одного пользователя к нескольким контент-серверам одновременно. В таком случае задача маршрутизации запросов может быть сформулирована как задача поиска совокупности серверов, которые будут выступать в качестве источниками контента, а также множества маршрутов доставки данного контента от каждого сервера до потребителя (пользователя) с выполнением гарантий относительно качества предоставляемой услуги.

С целью решения задачи маршрутизации запросов в описанной постановке предлагается использовать следующую математическую модель [3]:

$$\sum_{j \in N} x_{ij}^{kl} - \sum_{j \in N} x_{ji}^{kl} - \sum_{g \in G_l} y_g^{kl} x_{gi}^{kl} = \begin{cases} 0, & \text{если } i \neq s_k, t_k, \\ -1, & \text{если } i = t_k, \end{cases} \quad (1)$$

$$\sum_{j \in N} x_{gj}^{kl} = \begin{cases} 0, & \text{если } g \neq s_k, \\ y_g^{kl}, & \text{если } g = s_k, \end{cases} \quad (2)$$

$$\begin{cases} y_g^{kl} = 0, & \text{если } g \neq s_k, \\ 0 < y_g^{kl} \leq 1, & \text{если } g = s_k, \end{cases} \quad \sum_{g \in G_l} y_g^{kl} = 1, \quad (3)$$

$$0 \leq x_{ij}^{kl} \leq 1, \quad (4)$$

$$\sum_{k \in K} \sum_{l \in L} r^{kl} x_{ij}^{kl} \leq c_{ij}, \quad \sum_{k \in K} y_g^{kl} \leq S_g^l, \quad (5)$$

где x_{ij}^{kl} – доля трафика, передаваемого по тракту передачи $(i, j) \in E$ от i -го маршрутизатора к j -му (индекс l отражает тип передаваемого контента, а индекс k связан с парой отправитель-получатель $\{s_k, t_k\}$, $s_k \in G_l$, $t_k \in N$); N – множество узлов (маршрутизаторов) транспортной телекоммуникационной сети; y_g^{kl} – переменная использования g -го сервера, $g \in G_l$, в качестве источника контента l -го типа в k -й паре отправитель-получатель, G_l – множество периферийных серверов, на которых доступен контент l -го типа; K – множество пар отправитель-получатель в обслуживаемой сети доставки контента; L – множество возможных типов контента, распространяемых данной сетью; E – множество трактов передачи транспортной телекоммуникационной сети; r^{kl} – интенсивность трафика, порождаемого при передаче контента l -го типа между k -й парой сервер-пользователь; S_g^l – максимальное число сессий l -го типа, которое способен обслужить g -й сервер (производительность сервера).

Выражения (1) – (5) представляют собой потоковую модель телекоммуникационной сети, адаптированную для случая динамического выбора источника трафика (сервера контента) и ориентированную на многопутевой способ доставки. Здесь (1) – (2) представляют собой закон сохранения потока для i -го маршрутизатора транспортной телекоммуникационной сети, выражения (3) обеспечивают целостность передаваемого от нескольких серверов контента, условие (4) ориентирует на многопутевой способ передачи трафика, выражения (5) связаны с ограниченными объемами сетевых и вычислительных ресурсов.

В рамках модели (1) – (5) процесс маршрутизации запросов в CDN связан с поиском переменных y_g^{kl} , а наличие неизвестных переменных x_{ij}^{kl} позволяет решить эту задачу совместно с традиционной задачей многопутевой маршрутизации в транспортной телекоммуникационной сети. При этом в качестве целевой функции может выступать стоимостная функция

$$W = Q_x \bar{x} + Q_y \bar{y} \rightarrow \min, \quad (6)$$

где \bar{x} – вектор, объединяющий в себе переменные x_{ij}^{kl} ; \bar{y} – вектор, объединяющий переменные y_g^{kl} ; Q_x , Q_y – векторы весовых коэффициентов, определяющие стоимость использования сетевых и вычислительных ресурсов соответственно.

Одновременное использование совокупности серверов в качестве источника контента способствует повышению эффективности использования сетевых и вычислительных ресурсов в целом, однако не

гарантирует требуемые значения скорости загрузки контента и задержки передачи. С целью реализации обслуживания гарантированного качества введем в модель (1) – (6) дополнительное условие [4]

$$\lambda^{(req)} \leq \left(G_{\pi\eta}^{(4,1)} - G_{\pi\eta}^{(4,2)} \left[G_{\pi\eta}^{(4,4)} \right]^{-1} G_{\pi\eta}^{(4,3)} \right) \tau_{(req)}, \quad (7)$$

где $\lambda^{(req)}$ и $\tau_{(req)}$ – требуемые значения скорости загрузки контента и задержки передачи соответственно; $G_{\pi\eta}^{(4,1)}$, $G_{\pi\eta}^{(4,2)}$, $G_{\pi\eta}^{(4,3)}$, $G_{\pi\eta}^{(4,4)}$ – подматрицы проекции метрического тензора системы, построенного по правилам, описанным в [5]. Заметим, что проекция метрического тензора является функцией интенсивности трафика, передаваемого по трактам передачи транспортной телекоммуникационной сети, что в рамках представленной модели отражено переменными x_{ij}^{kl} .

Таким образом, предлагаемая математическая модель позволяет сформулировать и найти оптимальное решение задачи маршрутизации запросов пользователей; достичь высокой согласованности при решении задач выбора контент-сервера и маршрутизации контента к запрашивавшему его пользователю; обеспечить гарантированное качество обслуживания по скоростным и временным показателям сетевого уровня; достичь высокой эффективности использования сетевых и вычислительных ресурсов за счет многопутевого способа маршрутизации и перераспределения запросов между контент-серверами.

Литература:

1. Cisco Visual Networking Index: Forecast and Methodology, 2012–2017. White paper. – Cisco, 2013. – 15 p. – Available: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf
2. ZhiHui Lu, Ye Wang, Yang Richard Yang An Analysis and Comparison of CDN-P2P-hybrid Content Delivery System and Model // Journal of Communications. – 2012. – Vol. 7, N. 3. – P. 232 – 245.
3. Евсеева О.Ю., Кадер М.Б. Математическая модель маршрутизации запросов в сетях доставки контента // Системи обробки інформації. – 2012. – № 9 (107). – С. 165 – 170.
4. Лемешко А.В., Евсеева О.Ю. Тензорная модель многопутевой маршрутизации с гарантиями качества обслуживания одновременно по множеству разнородных показателей [Электронный ресурс] // Проблеми телекомунікацій. – 2012. – № 4 (9). – С. 16 – 31. – Режим доступа: http://pt.journal.kh.ua/2012/4/1/124_lemeshko_tensor.pdf.