

СПОСОБ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ GRID СИСТЕМ НА БАЗЕ СЕТЕЙ MPLS

Предложен способ повышения эффективности функционирования GRID систем. Проведен анализ изменений пропускной способности каналов передачи данных на базе объединенных деревьев доставки информации туннелей LSP сетей MPLS. Показано, что распределение нагрузки по маршрутизаторам DR, а также балансировка пропускной способности при многоабонентской доставке информации существенно зависит не только от физической, но и от логической организации распределенной системы

A way to improve the effectiveness of the GRID systems is proposed. The analysis of changes in channel bandwidth capacity based on the combined trees information delivery of LSP tunnels MPLS networks is adopted. It has been shown that load distribution of the DR as well as balancing bandwidth capacity of multicast information delivery depends not only on the physical but also on the logical organization of a distributed system

Введение

В связи с расширением области применения GRID систем, увеличения доли мультимедийного и высокоскоростного трафика в компьютерных сетях и системах, актуальной становится задача организации быстрой и эффективной доставки информации.

GRID – это аппаратно-программная инфраструктура на основе компьютерной сети, которая обеспечивает эффективный доступ к высокопроизводительным компьютерным ресурсам [1]. В настоящее время для построения GRID систем все чаще используется многопротокольная коммутация на основе меток (MPLS – Multiprotocol Label Switching). Технология MPLS сочетает в себе возможности управления трафиком, присущие технологиям канального уровня [2] (уровень 2 модели OSI), масштабируемость и гибкость протоколов, характерные для сетевого уровня [3] (уровень 3 модели OSI). В соответствии с протоколом MPLS маршрутизаторы и коммутаторы присваивают на каждой точке входа в таблицу маршрутизации особую метку и сообщают эту метку соседним устройствам.

Сеть MPLS делится на две функционально различные области – ядро и граничную область. Ядро образуют устройства, минимальным требованием к которым является поддержка MPLS и участие в процессе маршрутизации трафика для того протокола, который коммутируется с помощью MPLS. Маршрутизаторы ядра занимаются только коммутацией. Все функции классификации пакетов по различным классам FEC (Forwarding Equivalence Class), а также реализацию дополнительных

сервисов берут на себя граничные маршрутизаторы LER (Label Edge Router) и маршрутизаторы переключений LSR (Label Switching Router). В результате интенсивные вычисления приходится на граничную область, а коммутация выполняется на уровне ядра. В связи с этим актуальной становится задача повышения эффективности вычислений в граничной области MPLS.

Механизмы MPLS могут быть реализованы двумя способами: на основе модели перекрытий (overlay) или одноранговой (peer) модели. В модели перекрытий, называемой также UNI, маршрутизатор является клиентом оптического домена и взаимодействует только с непосредственно примыкающим к нему оптическим узлом. В одноранговой модели уровень IP/MPLS обладает теми же полномочиями, что и уровень оптической передачи. Основная задача MPLS в обоих случаях заключается в расширении сферы действия технологии коммутации по меткам и ее переносе от маршрутизаторов на оптический уровень, где решения о дальнейшей пересылке пакетов принимаются на основании временных интервалов, длин волн и физических портов «неявных меток», а не содержимого пакета. Технология MPLS устанавливает равноправные отношения между оптическими доменами за счет поддержки новых классов LSR.

В настоящее время для построения GRID систем все чаще используется многопротокольная коммутация на основе меток (MPLS – Multiprotocol Label Switching). Технология MPLS сочетает в себе возможности управления трафиком, присущие технологиям каналь-

ного уровня (уровень 2 модели OSI), масштабируемость и гибкость протоколов, характерные для сетевого уровня (уровень 3 модели OSI). В соответствии с протоколом MPLS маршрутизаторы и коммутаторы присваивают на каждой точке входа в таблицу маршрутизации особую метку и сообщают эту метку соседним устройствам.

Сеть MPLS делится на две функционально различные области – ядро и граничную область. Ядро образуют устройства, минимальным требованием к которым является поддержка MPLS и участие в процессе маршрутизации трафика для того протокола, который коммутируется с помощью MPLS. Маршрутизаторы ядра занимаются только коммутацией. Все функции классификации пакетов по различным классам FEC (Forwarding Equivalence Class), а также реализацию дополнительных сервисов берут на себя граничные маршрутизаторы LER (Label Edge Router) и маршрутизаторы ядра LSR (Label Switching Router). В результате интенсивные вычисления приходятся на граничную область, а высокопроизводительная коммутация выполняется в ядре, что позволяет оптимизировать конфигурацию устройств MPLS в зависимости от их местоположения в сети.

Главная особенность MPLS – отделение процесса коммутации пакета от анализа IP-адресов в его заголовке. Очевидным следствием описанного подхода является тот факт, что очередной сегмент LSP может не совпадать с очередным сегментом маршрута, который был бы выбран при использовании традиционных алгоритмов маршрутизации. Использование явно задаваемого маршрута в сети MPLS свободно от недостатков стандартной IP-маршрутизации от источника, поскольку вся информация о маршруте содержится в метке и пакету не требуется нести адреса промежуточных узлов, что улучшает управление распределением нагрузки в сети. По значению метки пакета определяется его принадлежность к определенному классу FEC на каждом из участков коммутируемого маршрута. Метка должна быть уникальной лишь в пределах соединения между каждой парой логически соседних LSR. Поэтому одно и то же ее значение может использоваться LSR для связи с различными соседними маршрутизаторами, если только имеется возможность определить, от какого из них пришел пакет с данной мет-

кой. В рамках архитектуры MPLS вместе с пакетом разрешено передавать не одну метку, а целый стек. Результат коммутации задает лишь верхняя метка стека, нижние же передаются прозрачно до операции изъятия верхней. Такой подход позволяет создавать иерархию потоков в сети MPLS и организовывать туннельные передачи. Коммутируемый путь LSP одного уровня состоит из последовательного набора участков, коммутация на которых происходит с помощью метки данного уровня.

Для динамической маршрутизации в GRID системах наиболее эффективно использовать многоабонентскую передачу информации, на базе протокола «точка – многоточие» с использованием меток. Существуют два протокола рассматриваемые IETF (Internet Engineering Task Force) для построения LSP в сетях MPLS [4]: протокол резервирования ресурсов RSVP-TE (Resource Reservation Protocol – Traffic Engineering) и протокол распределения меток LDP (Label Distribution Protocol). Оба протокола могут быть расширены для передачи информации в LSP режиме «точка – многоточие», однако, RSVP-TE строит деревья «точка – многоточие» от корня к конечным вершинам, а LDP наоборот. В случае одновременной передачи одного IP-пакета по нескольким адресам, распределение меток LDP формируется после групповой передачи. В свою очередь, протокол RSVP-TE специально разработан для масштабируемого группового обслуживания с использованием сетей VPN и мульти-VPN на основе группового дерева. Поэтому, использование протокола RSVP-TE для совмещения резервирования ресурсов и организации LSP для различных потоков данных является более эффективным для поставленной задачи.

Применение RSVP-TE позволяет оптимизировать GRID систему под конкретные задачи путем создания туннелей LSP на основе VPN. Сообщения этих протоколов передаются от одного узла сети к другому в соответствии с данными об IP-адресах маршрута.

Необходимо отметить, что каждый граничный маршрутизатор должен располагаться как можно ближе к корню дерева (Root), обычно, такой граничный маршрутизатор является абонентом многоабонентской доставки или широковещательного потока который был послан для распределения. В этом случае, все граничные маршрутизаторы (PE) посылают

единому Root свои пакеты через LSP туннель ядра для получения внешней метки, осуществляя проверку предварительно назначенной внутренней метки соответствующему маршрутизатору распределения (DR) для данной VPN группы, и заменяют метку VPN группы

на метку всего дерева. В таблице маршрутизатора DR существует вторая метка (вых. метка) для передачи информации вниз по дереву распределения другим VPN, как показано на рис. 1.

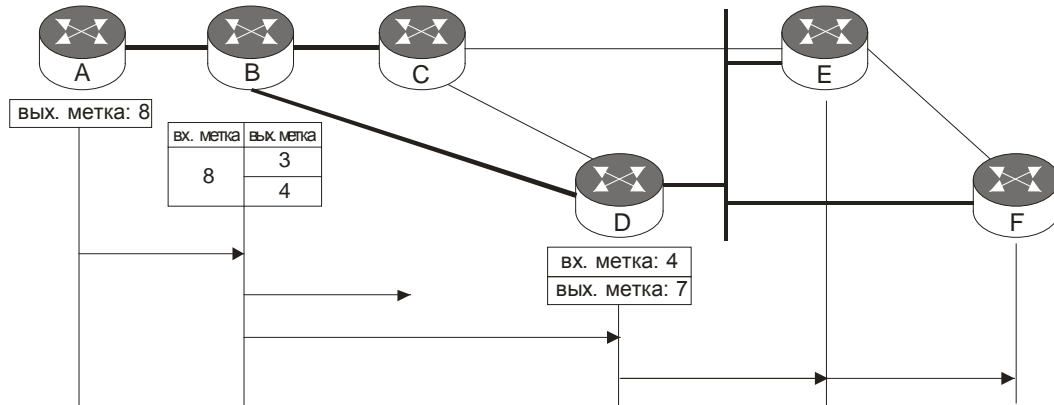


Рис. 1. Процесс передачи информации вниз по дереву распределения другим VPN

Выбор между единичным деревом и объединенным деревом распределения зависит от доступных ресурсов. Объединенное дерево обеспечивает надежность и минимальную задержку, за счет своей сложности и состояния передачи, поэтому защищенное одиночное групповое дерево в нашем случае является наиболее приемлемым и перестраиваемым. Перестройка дерева осуществляется при помощи сигнального протокола «точка-многоточие» LSP поддерживаемого RSVP-TE.

В связи с тем, что возрастает интерес к использованию GRID систем существенным требованием является повышение пропускной способности канала без повышения стоимости ресурсов системы. Для оценки состояния пропускной способности будем рассматривать простую модель, полученную в результате объединения маршрутизаторов PE с абонентами в сети VPN по не которым признакам. Предлагается не изменять физическую топологию системы, а применять VPN сети для поддержки необходимой структуры. Информация, предназначенная всем узлам GRID системы, которые принадлежат множеству сетей VPN объединенных в дерево распределения, передается при помощи протокола «точка-многоточие». Для этого необходимо, чтобы все маршрутизаторы PE передавали информацию корню дерева (Root). Если маршрутизатор не владеет информацией о составе группы, то передача информации осуществля-

ется вниз по дереву граничным маршрутизаторам.

Для этого необходимо отметить, что VPN сети A и B объединены между собой маршрутизаторами DR, которые в свою очередь связаны между собой корневыми маршрутизаторами CR. Для передачи информации абонент AMb VPN сети A передает информацию в маршрутизатор CR Root через маршрутизаторы PE и DR.

Корневой маршрутизатор передает информацию маршрутизаторам DR, которые входят в состав VPN B, и далее осуществляется передача информации вниз по дереву через маршрутизаторы PE до абонентов.

Общую пропускную способность (BW_{LSP}) туннеля LSP сети VPN можно вычислить аналитически относительно соединения «точка-точка». Если свести многоадресную и широковещательную передачу всех VPN к единице, то пропускную способность можно определить по формуле:

$$BW_{LSP} = K * \left(N * M * 2 + M \sum_{i=1}^{\log_2(N)} 2^{i-1} * (2i-1) \right), \quad (1)$$

где: K – количество корневых маршрутизаторов CR;

N – количество маршрутизаторов DR,

M – количество абонентов.

В (1) определена стоимость репликации одного пакета в одной из вершин графа и пе-

ресылка его по туннелю LSP VPN соединения «точка-точка» другим вершинам. Первое слагаемое, $N * M * 2$ определяет стоимость связи от маршрутизаторов DR до маршрутизаторов PE. Второе слагаемое

$$M \sum_{i=1}^{\log_2(N)} 2^{i-1} * (2i - 1)$$

определяет стоимость от корня дерева до маршрутизаторов DR.

С другой стороны, если дерево передает информацию по соединению «точка-многоточие» для каждого из туннелей LSP сетей VPN, общая пропускная способность (BW_{mLSP}) может быть представлена в виде:

$$BW_{mLSP} = K * \left(1 + \log_2(N) + \sum_{i=1}^{\log_2(N)} 2^i + M * N \right), \quad (2)$$

где: K – количество корневых маршрутизаторов CR;

N – количество маршрутизаторов DR;

M – количество абонентов.

Слагаемое $1 + \log_2(N)$ определяет стоимость отправки пакета от маршрутизатора PE к корню дерева «точка-многоточие». Сумма $\sum_{i=1}^{\log_2(N)} 2^i$ определяет стоимости передачи информации от корня дерева всем маршрутизаторам PE, которые входят в состав туннеля LSP сети VPN.

После объединения сетей VPN в дерево распределения, изменяется только последнее слагаемое в (2), которое принимает вид $Z * N$, где Z ($Z \geq M$) обозначает среднее количество вершин, которые зависят от маршрутизатора DR в «точка-многоточие» туннеля LSP, после равномерного распределения K деревьев на W групповых деревьев ($W \leq K$).

Увеличение пропускной способности (BW_{abLSP}) возникает при условии, когда объединенное дерево VPN А не содержит в себе маршрутизаторов PE из группы VPN В. При этом передача информации через такой маршрутизатор PE считается бесполезной. В этой ситуации, маршрутизаторы DR поддерживают количество передающих связей равное $W * Z$. Необходимо отметить, что $M \leq Z \leq L$.

$$BW_{abLSP} = K * \left(1 + \log_2(N) + \sum_{i=1}^{\log_2(N)} 2^i + Z * W \right), \quad (3)$$

где: K – количество корневых маршрутизаторов CR;

N – количество маршрутизаторов DR;

Z – среднее количество вершин, которые зависят от маршрутизатора DR;

W – количество групповых деревьев.

В результате значение пропускной способности ($BW_{(ab)}$), которое не эффективно используется во время передачи от абонента АМВ VPN сети А всем абонентам VPN сети В можно определить как разность (2) от (3), представленное в виде:

$$BW_{(ab)} = K * (M * N - Z * W) \quad (4)$$

Необходимо отметить, что распределение сетей VPN случайным образом является не эффективным для построения больших распределенных систем. Невзирая на то, что пропускная способность зависит от вида топологии сети, необходимо проводить анализ распределения VPN сетей для повышения скорости процесса передачи данных. Таким образом, можно сделать вывод, что анализ общей пропускной способности туннелей LSP сетей VPN во многом зависит от правильной организации распределенной системы.

Моделирование

Для оценки преимущества объединенных деревьев, проведем моделирование распределенной системы при условии, что количество VPN равно 1000. Возьмем по 20 граничных маршрутизаторов для каждого из P маршрутизаторов DR. Расположим сети VPN по случайному закону распределения таким образом, чтобы они были равномерно разделены между всем маршрутизаторами DR. Для большей эффективности будем рассматривать дополнительные параметр - плотность распределения LSP туннелей VPN по маршрутизаторами DR (VPN density). Предположим, что все VPN сети имеют одинаковую плотность. Определим, что количество объединенных деревьев W изменяется в пределах от 1 до 1000. На каждой итерации, значение Z рассчитывается путем группировки случайным образом разных VPN сетей в объединенное дерево, которое в свою очередь относится к другой группе VPN сети в маршрутизаторе DR.

Результаты моделирование представлены на рис. 2. Рисунок 2.а показывает отношения использования пропускной способности от количества объединенных деревьев системы. Отметим, что параллельные линии отображают состояние пропускной способности при

использовании метода передачи информации «точка-точка», а кривые – «точка-многоточие» с использованием групповых деревьев.

На рисунке 2.б показано количество исходящих связей маршрутизаторов DR при многоабонентской доставке информации.

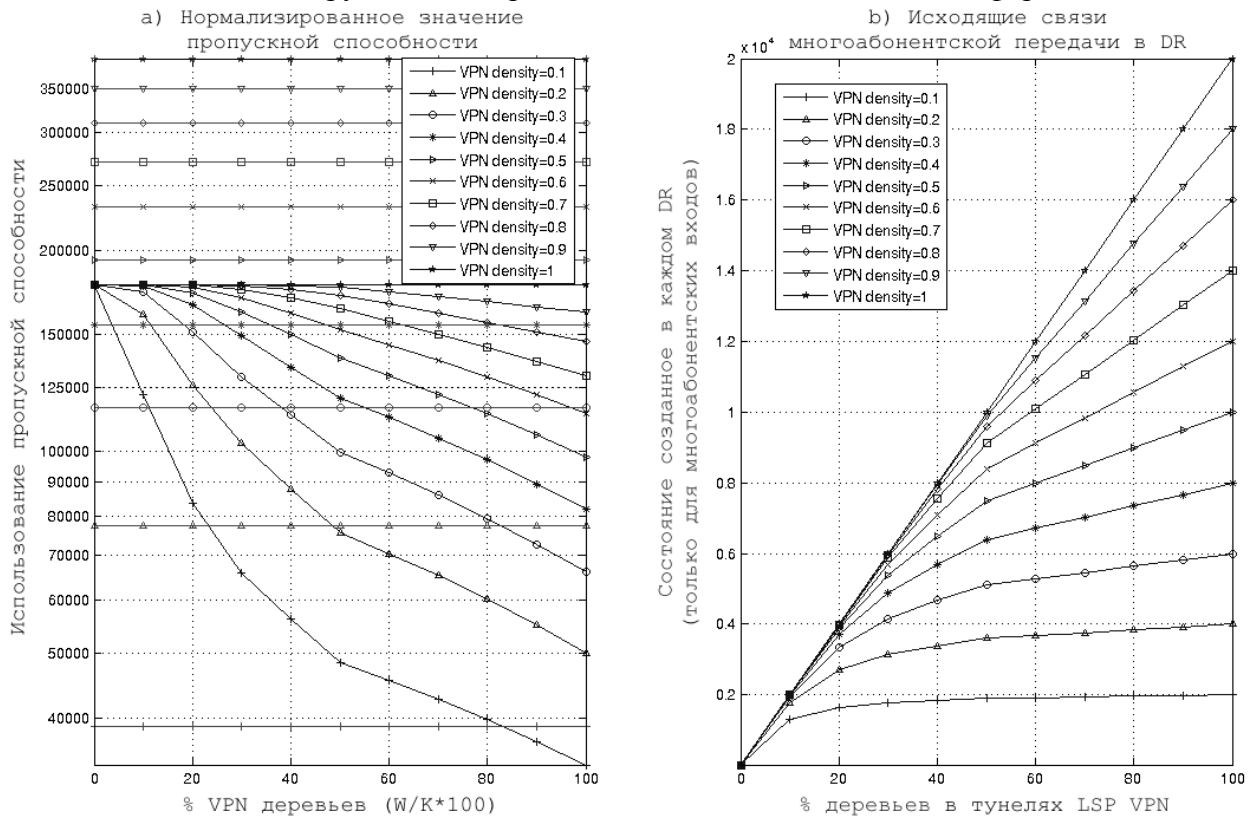


Рис. 2. Оценка пропускной способности и количества исходящих связей в GRID системе с использованием объединенных деревьев на базе туннелей LSP сетей MPLS.

Выводы

Предложенный метод многоабонентской доставки информации позволяет повысить эффективность функционирования GRID систем, за счет более эффективного распределения пропускной способности каналов передачи данных на базе объединенных деревьев доставки информации туннелей LSP

сетей MPLS. Проведя анализ результатов моделирования можно сказать, что распределение нагрузки по маршрутизаторам DR, а также балансировка пропускной способности при многоабонентской доставке информации существенно зависит не только от физической, но и от логической организации распределенной системы.

Список литературы

1. Ian Foster. What is the Grid? A Three Point Checklist, Argonne National Laboratory & University of Chicago, 2002, p.4
2. Столлингс В., Современные компьютерные сети. 2-е изд. – СПб.: Питер, 2003. - 783с.
3. Шиллер Й., Мобильные коммуникации.: Пер. с.англ. - М.: «Вильямс», 2002.-384 с.
4. A.Chaak, "Quality of service and Traffic Engineering in Consolidated Core and Metro Networks", University of Toronto, September 2004