

Унифицированная модель междоменной маршрутизации A Unified Approach to Inter-Domain Routing

Статус документа

Этот документ содержит информацию для сообщества Internet. Документ не задает каких-либо стандартов Internet. Допускается свободное распространение документа.

Тезисы

Этот документ относится к категории информационных RFC и описывает один из вариантов междоменной маршрутизации в междоменной среде будущего. Основное внимание уделено масштабированию сетей до очень больших размеров и функциональности, которая при обеспечении масштабирования поддерживает маршрутизацию в среде разнородных услуг, требований и критериев выбора маршрута.

Примечание: Работа D. Estrin и S. Hotz финансировалась Национальным научным фондом (National Science Foundation) по контракту NCR-9011279 с GTE Laboratories. Работа Y. Rekhter финансировалась Агентством перспективных оборонных проектов DARPA¹ по контракту DABT63-91-C-0019. Выводы и заключения, содержащиеся в этом документе могут отличаться от точки зрения DARPA и NSF.

1.0 Постановка задачи

Глобальную среду междоменного взаимодействия² можно представить в виде модели из множества хостов, соединенных между собой через каналы связи и коммутационные узлы³. Управление хостами, каналами и коммутаторами, которые совместно образуют ресурсы глобальной сети², не является однородным и распределено между множеством областей администрирования. Ресурсы, находящиеся под единым управлением, образуют домен. Для поддержки автономности и гетерогенности доменов маршрутизацию делят на две компоненты – внутридоменную⁴ (внутреннюю) и междоменную⁵ (внешнюю). Внутренняя маршрутизация обеспечивает поддержку обмена данными между хостами, которые связаны каналами и коммутаторами, относящимися к одному домену. Междоменная маршрутизация обеспечивает поддержку обмена данными между хостами в тех случаях, когда данные передаются через каналы и коммутаторы множества доменов. Устройства, пересылающие пакеты через границу домена называются граничными маршрутизаторами (BR⁶). Устройства, отвечающие за обмен данными междоменной маршрутизации, называют маршрутными серверами (RS⁷). Функции RS и BR могут быть реализованы в одном устройстве⁸.

По мере развития глобальной сети в части размера и усложнения требований к маршрутизации, обеспечение междоменной маршрутизации с учетом этих критериев становится все более важной задачей. Число и сложность требований к маршрутизации растут по нескольким причинам:

- (a) ограничения на выбор транзитных сетей, вносимые отправителями, получателями и транзитными сетями;
- (b) различные типы запрашиваемого и предлагаемого сервиса;
- (c) наличие множества операторов с различными схемами оплаты.

Наличие многочисленных комбинаций перечисленных критериев с разными весами значительно усложняет механизмы, обеспечиваемые традиционными архитектурами поэтапной маршрутизации⁹ [ISIS10589, OSPF, Hedrick88, EGP].

Текущий подход к междоменной маршрутизации в сообществе можно разделить по двум направлениям, одно из которых лучше всего представлено архитектурой BGP¹⁰/IDRP¹¹ [BGP91, Honig90, IDRP91], а другое — архитектурой IDPR¹² [IDPR90, Clark90]. В этом документе предложено рассматривать эти две архитектуры, как дополняющие, а не взаимоисключающие.

¹ Defense Advanced Research Projects Agency

² В оригинале «global internet». *Прим. перев.*

³ В оригинале «switching facilities». Это понятие существенно шире, нежели термин «коммутатор» в современном понимании. Тем не менее, для краткости будет использоваться термин «коммутатор». *Прим. перев.*

⁴ intra-domain (interior).

⁵ inter-domain (exterior).

⁶ border router.

⁷ route server.

⁸ В настоящее время это является нормальной практикой и функции граничного маршрутизатора реализуются в одном устройстве с функциями маршрутного сервера. *Прим. перев.*

⁹ hop-by-hop routing.

¹⁰ Border Gateway Protocol — протокол граничного шлюза.

¹¹ Inter-Domain Routing Protocol — протокол междоменной маршрутизации.

¹² Inter-Domain Policy Routing — междоменная маршрутизация на базе правил.

Предполагается, что в течение следующих 5 - 10 лет будут появляться новые типы сервиса, для обеспечения которых будут развернуты специальные устройства. Хотя число и разнообразие маршрутов, обеспечиваемых архитектурами поэтапной маршрутизации с поддержкой TOS¹ (например, множественные маршруты с метками), достаточно для большей части трафика, важно отметить востребованность механизмов поддержки эффективной маршрутизации специализированных типов трафика через особые маршруты. Примерами особых маршрутов являются:

- (1) маршруты, проходящие через один или множество транзитных доменов, которые нежелательны с точки зрения домена отправителя;
- (2) маршруты через транзитные домены, которые поддерживают недостаточно широко используемый сервис.

Все прочие маршруты будем относить к базовому типу.

Желание эффективно поддерживать специальные маршруты ведет к исследованию вопроса о динамической организации маршрутов [Breslau-Estrin91, Clark90, IDPR90]. В работе [Breslau-Estrin91] рассмотрен выбор алгоритма междоменной маршрутизации на базе правил и особо рассмотрены вопросы, связанные со специфическими для отправителя маршрутами и другими «специальными» маршрутами. По результатам рассмотрения сделан вывод о том, что для поддержки специальных маршрутов лучше всего подходят алгоритмы с заданием маршрута отправителем и расширенные алгоритмы маршрутизации по состоянию каналов²; далее эта модель будет называться «маршрутизацией по запросу отправителя»³. Однако архитектура с маршрутизацией по запросу отправителя, если ее сделать единственной архитектурой междоменной маршрутизации, приводит к возникновению проблем масштабирования, поскольку эта архитектура не включает иерархической кластеризации и агрегирования маршрутной информации. Например, если маршрут из промежуточного транзитного домена X к домену получателя Y используется 1000 доменов-отправителей, IDPR требует организации и поддержки состояний для каждого из 1000 маршрутов в транзитных граничных шлюзах между доменами X и Y. В противоположность этому, другая модель междоменной маршрутизации, основанная на поэтапной маршрутизации и распределенной системе расчета маршрутов (см. ниже), обеспечивает широкую поддержку агрегирования, а также абстрагирование информации о доступности, топологии и пересылке пакетов. Такую модель используют протоколы BGP и IDRP [BGP91, IDRP91]. Хотя архитектура BGP/IDRP способна обеспечить маршрутизацию для очень большого числа сетей, основанных на использовании дейтаграмм, эта архитектура не обеспечивает поддержки специальных требований к маршрутам и гибкости маршрутизации IDPR.

1.1 Обзор унифицированной архитектуры

Мы хотим обеспечить поддержку специальных маршрутов, а при отсутствии необходимости в таких маршрутах — возможность использовать агрегирование. Следовательно, наша масштабируемая архитектура междоменной маршрутизации состоит из двух компонент — маршрутизации по запросу отправителя (SDR³) и узловой маршрутизации (NR⁴). Компонента NR обеспечивает расчет и организацию маршрутов, совместно используемых большим числом отправителей. Эти обычные маршруты широко используются и распространяются, следовательно для них имеет важное значение возможность агрегирования. Компонента SDR рассчитывает и организует специальные маршруты, используемые небольшим количеством отправителей, когда нет смысла использовать NR⁵. Потенциально большое число различных специальных маршрутов вкуче с их редким использованием делают для таких маршрутов использование механизма NR неоправданно дорогим.

Хорошей аналогией может служить производство продукции общего назначения. Когда имеется предсказуемая картина спроса, компания может выпускать и продавать серийную продукцию. В рассматриваемой архитектуре NR обеспечивает поддержку «серийных» маршрутов. Если запросы невозможно предсказать, производитель принимает заказы и выпускает продукцию по мере необходимости. Кроме того, если продукция настолько специфична, что ее потребляет один или несколько заказчиков, потребители могут повторно запрашивать продукцию, даже если потребность в ней предсказуема, поскольку потребители не создают достаточно большого рынка для серийного выпуска продукции. SDR обеспечивает для таких специальных случаев маршрутизацию по запросам.

Объединяя NR и SDR мы обеспечим поддержку междоменной маршрутизации в глобальных сетях достаточно большого размера с одновременной эффективной поддержкой маршрутов со специализированными требованиями.

При разработке архитектуры предполагалось существование различных требований к маршрутизации и необходимость поддержки специальных маршрутов. С другой стороны, архитектура не включает допущений о запрашиваемых типах маршрутов и их распространении. Маршрутизация будет адаптироваться естественным путем в течение времени с учетом изменения картины трафика и появления новых типов сервиса путем распределения расчета и организации отдельных типов маршрутов между двумя компонентами гибридной архитектуры⁶.

Хотя функции пересылки пакетов в компонентах NR и SDR слабо связаны между собой, механизм обмена информацией о связности для SDR опирается на сервис, обеспечиваемый компонентой NR.

¹ type of service — тип обслуживания.

² extended link-state algorithm.

³ source-demand routing — маршрутизация по запросу отправителя. Такая модель используется в архитектуре междоменной маршрутизации на базе правил (IDPR).

⁴ node routing — маршрутизация в узлах (сети).

⁵ Для маршрутов, которые используются лишь спорадически (например, запросы на организацию таких маршрутов не являются постоянными или возникновение запросов трудно предсказуемо), зачастую целесообразно применять SDR.

⁶ Прежде, чем продолжить рассмотрение архитектуры, следует отметить, что поддержка сильно специализированных маршрутов для всех пар «отправитель — получатель» в глобальной сети или даже приближение к этому является неосуществимой задачей для любой архитектуры маршрутизации, которую можно предположить. Иными словами, мы не предполагаем, что какая-либо предсказуемая архитектура маршрутизации сможет поддерживать естественный рост требований пользователей и сетевых служб. Однако проблемы здесь не возникает. Возможности архитектуры могут превосходить потребности пользователей. Более того, некоторые требования, которые мы считаем невыполнимыми с точки зрения междоменной маршрутизации, могут быть полностью выполнены за пределами маршрутизации. Тем не менее, высказанные здесь предостережения избавляют от нереальных ожиданий.

1.2 Структура документа

В разделе 2 описаны требования и приоритеты для разработки компонент NR и SDR. В разделах 3 и 4 описано устройство NR и SDR, соответственно, в свете приведенных здесь требований. Раздел 5 описывает поддержку протокола для унифицированной архитектуры и кратко рассматривает процесс перехода к ее использованию. Последний раздел содержит краткое резюме.

2.0 Архитектурные требования и приоритеты

Для обоснования устройства масштабируемой архитектуры междоменной маршрутизации, нужно четко сформулировать приоритеты и критерии оценки. В этом разделе определены требования к сложности, уровню абстракции, правилам и типам сервиса.

2.1 Сложность

Сложность междоменной маршрутизации должна оцениваться на основе следующих метрик производительности:

- (1) издержки на хранение (storage overhead);
- (2) издержки на расчеты (computational overhead);
- (3) издержки на сообщения (message overhead).

Такая оценка имеет существенное значение для определения масштабируемости любой архитектуры.

2.1.1 Издержки на хранение

Издержки на хранение для объекта, принимающего участие в междоменной маршрутизации, связаны с необходимостью хранения базы маршрутной информации RIB¹ и базы информации о пересылке FIB². База RIB содержит маршрутную информацию которой объекты обмениваются с использованием протокола междоменной маршрутизации; данные RIB являются источником информации для расчета маршрутов. База FIB содержит информацию, которую объекты используют для пересылки междоменного трафика; данные FIB являются результатом расчета маршрутов. Для того, чтобы уровень издержек на хранение был приемлемым, базы FIB и RIB должны по мере увеличения числа доменов в глобальной сети расти существенно медленней линейной функции (например, близко к логарифмической). Для удовлетворения этих требований в части RIB архитектура должна обеспечивать механизмы агрегирования и абстрагирования маршрутной информации и информации о пересылке или нахождения подмножества этой информации по запросу. Для удовлетворения требований в части FIB архитектура должна обеспечивать механизмы агрегирования информации о пересылке (для рассчитанных маршрутов NR) или динамической организации/отзыва этой информации (для рассчитанных маршрутов SDR).

Кроме того, будучи действительно важной метрикой оценки, издержки на хранение оказывают прямое влияние на издержки, связанные с расчетами и расходом полосы. Пока сложность расчетов не является фиксированной (независимой от общего числа доменов), издержки на хранение оказывают прямое влияние на присущую архитектуре сложность расчетов, поскольку маршрутная информация является входными данными для расчета маршрутов. Более того, пока архитектура не поддерживает нарастающих обновлений, при распространении маршрутной информации издержки на хранение оказывают прямое влияние на размер занимаемой полосы, поскольку передача маршрутной информации занимает большую часть используемой полосы.

2.1.2 Издержки на расчеты

Работа компоненты NR основана преимущественно на предварительном расчете маршрутов. Если междоменная маршрутизация используется повсеместно, предварительно рассчитанные маршруты включают все достижимые точки назначения. Следовательно, сложность таких расчетов следует минимизировать. В частности, весьма желательна реализация в архитектуре той или иной формы частичного расчета, при котором расчет результатов изменения топологии будет требовать меньших ресурсов, нежели полный расчет. Даже при необходимости проведения полного расчета сложность его должна расти в зависимости от числа доменов медленней, чем линейная функция.

Компонента SDR использует расчет по запросу и кэширование. Следовательно, сложность расчетов не имеет столь важного значения. Другой причиной снижения требований к сложности расчетов для SDR является то, что в SDR предположительно рассчитывается значительно меньшее число маршрутов, нежели в NR (хотя расчет маршрутов SDR может происходить чаще).

Ни при каких обстоятельствах не допускается экспоненциальный рост сложности расчетов в зависимости от числа доменов (ни для NR, ни для SDR).

2.1.3 Дополнительный расход полосы

Полоса, расходуемая на распространение маршрутной информации, должна быть ограничена. Однако рост производительности сетевых соединений и возможность использования компрессии, делают это требование менее важным по сравнению с требованиями к издержкам на хранение и расчеты. Дополнительное снижение расхода полосы может быть обеспечено за счет передачи нарастающих обновлений маршрутной информации взамен передачи маршрутных данных целиком.

Хотя издержки на хранение и расход полосы могут быть взаимосвязаны, при использовании нарастающих обновлений дополнительный расход полосы в установившемся состоянии (нет изменений в топологии) становится пренебрежимо малым и не зависит от издержек на хранение. Иными словами, использование нарастающих обновлений снижает дополнительный расход полосы, связанный с динамикой сети. Следовательно, повышение уровня стабильности физических соединений в купе с методами подавления влияния топологических нестабильностей, позволяют существенно понизить значимость дополнительного расхода полосы.

¹ Routing Information Base

² Forwarding Information Base

2.2 Агрегирование

Агрегирование и абстрагирование маршрутной информации и информации о пересылке обеспечивает очень мощный механизм снижения издержек на хранение, вычисления и расход полосы. Возможность агрегировать и, впоследствии, абстрагировать маршрутные данные и информацию о пересылке имеет важное значение для масштабируемости архитектуры¹. Это, в частности, верно по отношению к компоненте NR, поскольку эта компонента должна обеспечивать возможность предоставления маршрутов ко всем или почти всем достижимым адресатам.

В то же время, поскольку сохранение независимости и автономии каждого домена является одним из ключевых требований междоменной маршрутизации, архитектура должна стремиться к максимальной гибкости схемы агрегирования (например, вносить как можно меньше предварительных условий и требовать от участвующих доменов минимальной координации в глобальном масштабе).

База маршрутной информации RIB содержит три типа данных:

- (1) топология (например, соединения между доменами или группами доменов);
- (2) доступность на сетевом уровне;
- (3) ограничения на транзит.

Агрегирование маршрутных данных должно приводить к снижению размера всех трех компонент. Агрегирование информации о пересылке будет следовать за агрегированием информации о доступности.

Кластеризация (путем формирования конфедераций доменов) позволяет решить следующие задачи агрегирования:

- (1) сокрытие части реальной физической топологии и связанное с этим абстрагирование топологических данных;
- (2) объединение множества достижимых объектов в один объект для снижения издержек на хранение;
- (3) выражение ограничений на транзит в терминах кластеров, а не отдельных доменов.

Как отмечено в работе [Breslau-Estrin91], архитектура должна разрешать формирование и изменение конфедераций без излишних операций по настройке и координированию. В частности, создание конфедерации не должно требовать какой-либо глобальной координации (например, требуемой в ECMA [ECMA89]). Кроме того, агрегирование не должно требовать явного указания взаимного расположения каждого домена относительно других доменов (например, от доменов или конфедераций не должно требоваться согласия на частичное упорядочение — кто за кем и т. п.).

Архитектура должна позволять различным доменам использовать разные методы агрегирования и абстрагирования. Например, исследователь из IBM может организовать маршрут в USC как объект доменного уровня для получения преимуществ в результате специальной обработки TOS для получения доступа к USC или даже транзита через этот домен. В то же время некий сотрудник Digital Equipment Corporation может видеть информацию на уровне California Educational Institutions Confederation и знать лишь, что USC является членом этой конфедерации. В дополнение к сказанному USC может видеть часть внутренней структуры IBM Confederation (на уровне доменов), тогда как UCLA может маршрутизировать конфедерацию доменов IBM только как единое целое.

Поддержка конфедераций должна быть гибкой. В частности, архитектура должна позволять перекрытие конфедераций без их полной вложенности (например, домен или группа доменов могут быть частью одной или множества конфедераций). Например, USC может быть частью California Educational Institutions Confederation и частью US R&D Institutions Confederation, хотя эти конфедерации не являются подмножеством по отношению одна к другой. Другой пример, T.J. Watson Research Center может быть частью NYSERNET Confederation и частью IBM-R&D-US Confederation. Хотя в обоих примерах зона перекрытия представляет собой один домен, возможны случаи, когда зона перекрытия состоит из множества доменов. В качестве примера рассмотрим множество доменов, формирующих IBM Confederation, и другое множество, образующее DEC Confederation. В конфедерации IBM есть домен IBM-Research, а в конфедерации DEC — DEC-Research. Оба эти домена могут участвовать в некой совместной работе с организацией между ними прямого соединения. В таком случае архитектура должна позволять ограничить использование этого канала так, чтобы другие домены IBM Confederation не могли использовать канал для связи с другими доменами DEC Confederation. Подобные примеры возникают и в тех случаях, когда транснациональная корпорация формирует конфедерацию, а отдельные филиалы в каждой стране могут входить в соответствующие конфедерации своей страны. Корпорации в таких случаях может потребоваться предотвращение использования своего домена в качестве транзитного внутри страны (в соответствии с локальной или международной политикой). Все приведенные выше примеры показывают ситуации с перекрытием конфедераций при необходимости контроля трафика, который проходит через перекрывающиеся ресурсы.

Когда агрегирование приспособливается к любой междоменной архитектуре, зона доступности такого агрегирования оказывает прямое влияние на связанную с агрегированием масштабируемость. В то же время, использование такой возможности зависит от способа выделения адресов. В частности, масштабирование, связанное с информацией о пересылке существенно зависит от допущения о наличии общего соответствия между иерархией агентств по распределению адресов и способами организации доменов маршрутизации и их конфедераций (см. параграф 2.6).

2.3 Правила маршрутизации

Правила маршрутизации, которые должна поддерживать архитектура, можно в общем случае, отнести к правилам транзита и правилам выбора маршрутов [Breslau-Estrin 91, Estrin89].

Вносимые правилами транзита ограничения могут быть обусловлены разными причинами. Архитектура должна поддерживать ограничения по адресам отправителя/получателя, типу обслуживания (TOS), идентификации класса пользователя (UCI²), оплате и пути [Estrin89, Little89]. Архитектура должна обеспечивать возможность задания правил транзита для всех маршрутов (NR и SDR). Даже для широко используемых маршрутов NR с незначительными ограничениями по отправителям и получателям, маршруты NR могут иметь некую транзитную классификацию

¹ Хотя мы не говорим об отсутствии иных путей обеспечить масштабирование, нам неизвестны какие-либо иные механизмы, обеспечивающие агрегирование/абстрагирование, за исключением кластеризации. Следовательно, в остальной части документа предполагается, что для агрегирования/абстрагирования используется кластеризация.

² user class identification.

(например, ограничения по TOS, стоимости, классу пользователей). Если широко используемый маршрут к адресатам имеет классификацию политики, последняя должна анонсироваться NR с явным указанием транзитных ограничений.

Правила выбора маршрутов позволяют каждому домену выбрать один маршрут к адресату из множества доступных маршрутов к нему. Для обеспечения максимальной автономии и независимости доменов архитектура должна поддерживать гетерогенные правила выбора маршрутов, когда каждый домен может использовать свои критерии выбора. Этот аргумент уже приводился в более ранних документах [IDPR90, Clark90, Breslau-Estrin91] применительно к выбору маршрута, задаваемого отправителем. Кроме того, каждый транзитный домен должен иметь возможность задания своих критериев выбора маршрутов, предоставляемых ему соседями. На практике это является прямым следствием приведенного выше требования — если мы позволяем выражать правила выбора для маршрутов NR, мы не можем предполагать что все домены будут использовать такие же правила. Для поддержки разнородных правил выбора маршрутов одним из ключевых моментов является разграничение междоменной и внутridoменной маршрутизации [ESMA89, Estrin89]. Однако поддержка всех возможных правил выбора маршрутов не относится к числу задач архитектуры. Список неподдерживаемых правил выбора маршрутов приведен ниже в параграфе 2.3.2.

2.3.1 Соккрытие маршрутной информации

Архитектура не должна требовать от всех доменов глобальной сети показа их связности и ограничений на транзит. Домены должны иметь возможность выполнения своих правил транзита при минимальном анонсировании своих соединений и правил транзита. В архитектуре не следует требовать, чтобы все домены в междоменной среде (internet) выявляли связи между собой и транзитные ограничения; анонсирование следует выполнять на основе критерия «необходимо знать». Более того, анонсирование правил транзита доменам, которые не могут воспользоваться этими правилами, будет увеличивать объем информации, которую требуется хранить, обрабатывать и распространять. Для маршрутизаторов поддержка полной информации о междоменной топологии и правилах транзита может быть не только неудобной, но и просто запрещенной.

2.3.2 Неподдерживаемые правила

В этом документе и более ранних работах авторы отмечают, что архитектура глобальной междоменной маршрутизации должна поддерживать широкий диапазон правил. В этом параграфе перечислены несколько типов правил, которые явно не рекомендуются поддерживать. Причины отказа от такой поддержки в общем случае сугубо практические — предполагается, что такие правила будут значительно повышать издержки или могут вызывать нестабильности маршрутизации.

1. **Выбор маршрута под влиянием внешних условий.** Процесс выбора маршрута можно моделировать функцией, которая присваивает маршруту неотрицательное значение в зависимости от уровня предпочтительности этого маршрута. Выбор маршрута осуществляется путем применения этой функции ко всем возможным маршрутам и выделения маршрута в наибольшем значении. Для обеспечения стабильности, функция ранжирования не должна принимать на входе состояние и атрибуты других маршрутов (к тому же или к другим адресатам).
2. **Влияние внешних условий на правила транзита.** Для обеспечения стабильности на правила транзита, принятые в домене, не должна автоматически оказывать влияние информация, являющаяся внешней по отношению к домену. В частности, правила не должны автоматически меняться в ответ на получение информации о других доменах или маршрутах, выбранных локальным доменом или другими доменами. Правила транзита также не могут меняться автоматически в ответ на получение информации о параметрах производительности локального домена или других доменов.
3. **Влияние правил на внешние состояния (например, на нагрузку).** Задачей архитектуры не является поддержка зависимой от загрузки маршрутизации для обычных маршрутов. Однако возможно использование некоторыми типами сервиса информации о загрузке для выбора маршрутов SDR.
4. **Очень большое количество одновременных маршрутов SDR.** Задачей архитектуры не является поддержка очень большого числа одновременных маршрутов SDR через какой-либо отдельный маршрутизатор. В частности, издержки на хранение FIB, связанные с маршрутами SDR, должны быть одного порядка с издержками на хранение маршрутов NR, а также во всех случаях эти издержки должны быть ограничены требованиями к сложности, рассмотренными выше¹.

2.4 Поддержка TOS-маршрутизации

В этом документе упоминается поддержка TOS²-маршрутизации. Ведется много исследований и разработок в части сетевой архитектуры и протоколов для обработки широкополосного multimedia-трафика. Часть такого трафика чувствительна к задержкам, другая часть требует широкой полосы. Неразумно предполагать, что одна коммуникационная система, развернутая в масштабах глобальной сети (включая все городские, региональные и магистральные сети) обеспечит однородную поддержку всех типов трафика. Для поддержки различающихся требований в гетерогенной среде будут использоваться различные механизмы управления ресурсами (например, различные варианты резервирования ресурсов) в разных частях глобальной сети [ST2-90, Zhang91].

В этом контексте поиск маршрутов, которые могут обеспечивать запрошенные параметры TOS, является задачей протокола маршрутизации. Явно не относится к числу задач протоколов маршрутизации общего назначения поиск маршрутов, которые гарантируют наличие доступных ресурсов в нужное время. Иными словами, неразумно предполагать, что оперативная информация о доступности ресурса может быть достоянием всех удаленных точек глобальной сети. Вместо этого после идентификации TOS-маршрута приложение, которому требуется определенный уровень обслуживания, будет пытаться использовать этот маршрут (например, с помощью явного запроса на организацию маршрута, если такой запрос нужен для нижележащих протоколов). В разделе 4 описаны дополнительные службы, которые могут обеспечиваться для поддержки более адаптивного выбора маршрута с определенным TOS³.

¹ Как было отмечено выше, теоретически издержки могут расти по закону $O(N^2)$, где N показывает число доменов в сети. Однако не очевидно, что это будет служить фактором, ограничивающим широкое использование SDR, при условии, что для обычных маршрутов доступна компонента NR.

² Type of service — тип обслуживания (сервиса).

2.5 Общее в компонентах маршрутизации

Хотя различие в требованиях к компонентам NR и SDR допускается, такое решение (при прочих равных) является наименее предпочтительным. В теории использование похожих алгоритмов и механизмов для компонент NR и SDR дает определенные преимущества. В этом случае обеспечивается возможность совместного использования программного кода и баз данных, а архитектура становится более управляемой и полнофункциональной. С другой стороны, использование в архитектуре маршрутизации гетерогенных компонент с некоторой независимостью одной от другой также может давать определенные преимущества (например, повышение устойчивости к ошибкам). В разделе 5 рассматривается несколько областей, в которых могут использоваться предложения по поводу повышения уровня унификации.

2.6 Взаимодействие с адресацией

Предлагаемая архитектура должна быть применима для различных схем адресации. В документе не принимается каких-либо допущений по поводу конкретной структуры адресов за исключением того, что эта структура должна обеспечивать возможность агрегирования без использования строгой иерархической маршрутизации.

За исключением этого требования, большинство предложений по расширению пространства адресов IP, например, могут использоваться вместе с предложенной архитектурой. Однако архитектура сама по себе не обеспечивает (не предлагает) какого-либо решения проблемы адресации.

3.0 Устройство компоненты NR

В этом разделе рассматривается компонента NR с учетом рассмотренных выше архитектурных требований и приоритетов. При обсуждении NR предполагается поэтапная (hop-by-hop) маршрутизация. Задаваемая отправителем маршрутизация имеет иные ограничения и используется для компоненты SDR.

3.1 Обзор NR

Компонента NR разработана и оптимизирована для сред, в которых значительная часть пакетов будет передаваться по маршрутам, совместно используемым многими отправителями, а картина трафика предсказуема. Эффективность NR-компоненты повышается при увеличении числа доменов-отправителей, совместно использующих маршрут от некоей промежуточной точки к адресату. Более того, NR лучше всего подходит для маршрутизации междоменного трафика, который имеет достаточно установившийся характер, что позволяет поддерживать маршрут постоянно, или характер трафика предсказуем и маршруты можно организовывать по необходимости (скорей на основе прогнозов, нежели по реальному трафику). Такие маршруты сами включаются в процедуры распределенного расчета и активизации маршрутов.

Маршруты, организованные в маршрутизаторах, и информация, которая была использована для расчета этих маршрутов, отражают рабочее состояние средств маршрутизации¹ (а так же вносимые правилами ограничения) на момент расчета маршрутов. Расчет маршрутов вызывается изменениями в рабочем состоянии маршрутизации или ограничениях правил. Компонента NR поддерживает расчет маршрутов, который динамически адаптируется к изменениям как в топологии, так и в правилах. Компонента NR позволяет зависимый от времени выбор маршрутов или их удаление. Однако временная зависимость должна быть предсказуемой (например, анонсирование того или иного маршрута только после окончания рабочего времени) и маршруты должны использоваться достаточно широко для включения в NR.

Предлагаемая архитектура основана на допущении о том, что большая часть междоменного обмена данными происходит с использованием маршрутов, рассчитанных и поддерживаемых компонентой NR.

Более того, обмен маршрутной информацией, требуемой для компоненты SDR, зависит от средств, обеспечиваемых компонентой NR, поэтому правила NR должны разрешать передачу информации о доступности SDR. Следовательно, архитектура требует, чтобы все домены в глобальной сети поддерживали компоненту NR. Поскольку масштабируемость (в плане размеров глобальной сети) является одним из фундаментальных требований для компоненты NR, нужно обеспечить множество механизмов в различной степени изоэдренности для агрегирования и абстрагирования информации.

Потенциальное снижение объемов маршрутной информации и информации о пересылке находится в сильной зависимости от способа распределения адресов и структурирования доменов и конфедераций. "Если отсутствует корреляция между иерархией распределения адресов и организацией доменов маршрутизации, ... для каждого домена маршрутизации OSIE потребуются включить запись в таблицу маршрутизации на каждой граничной IS² всех остальных доменов маршрутизации" [Ogan89]. Бесспорно, что масштабирование компоненты NR почти полностью предсказуемо на основе допущения о наличии общего соответствия между иерархией агентств по распределению адресов и способами организации доменов маршрутизации, при котором становится возможным эффективное и частое агрегирование и абстрагирование данных маршрутизации и пересылки. Следовательно, по природе междоменной маршрутизации в целом и NR-компоненты, в частности, масштабируемость архитектуры сильно зависит от гибкости схемы для агрегирования и абстрагирования, а также вносимых такой схемой предварительных условий. Более того, благодаря гибкости архитектуры, эффективность работы (масштабируемость) глобальной сети в целом или отдельных ее частей будет зависеть от соотношения между гибкостью и агрегированием.

Хотя компонента NR оптимизируется в соответствии с общими требованиями к маршрутизации для экстремально большого числа пользователей, это не означает, что маршруты, создаваемые компонентой NR, не могут использоваться для различных TOS. Напротив, по мере расширения использования TOS (например, во множестве доменов или предсказуемо по времени) маршруты TOS будут гарантированно рассчитываться компонентой NR.

В заключение отметим, что компонента NR лучше всего подходит для обеспечения маршрутов, которые используются множеством доменов. Т. е., эффективность компоненты NR растет с увеличением числа доменов-отправителей, совместно использующих маршрут от некоей промежуточной точки к адресату. Более того, NR лучше всего подходит для

³ Адаптивный выбор маршрута предполагает адаптивность только в процессе выбора. После того, как маршрут выбран, он не подвергается каким-либо изменениям за исключением отзыва устаревших маршрутов.

¹ В оригинале «routing facilities». *Прим. перев.*

² Intermediate System — промежуточная система. *Прим. перев.*

маршрутизации междоменного трафика, который находится в достаточно установившемся состоянии или предсказуем для организации маршрутов по мере необходимости (на основе прогноза, а не реального трафика).

3.2 Выбор алгоритма маршрутизации для NR

С учетом того, что компонента NR, основанная на поэтапной маршрутизации, требуется для обеспечения масштабируемой и эффективной междоменной маршрутизации, остальная часть этого раздела посвящена фундаментальным вопросам выбора алгоритма маршрутизации для NR.

Обычно при обсуждении алгоритмов маршрутизации рассматриваются алгоритмы на основе состояния канала (link state - LS) и векторов удаления (distance vector - DV). Однако простые протоколы DV (например, Routing Information Protocol [Hedrick88]) не обеспечивают должного масштабирования в связи с возникновением проблем сходимости¹. Для решения этой проблемы были разработаны усовершенствованные протоколы DV [Jaffee82, Zaumen91, Shin87], использующие механизмы синхронизации или дополнительную информацию о пути. Для случая междоменной маршрутизации добавление информации о пути имеет важное значение для политики поддержки. Следовательно, алгоритмы, рассматриваемые для NR, будут работать на основе состояния каналов и один из таких алгоритмов мы назвали PV². Хотя характеристики алгоритмов LS в общем случае понятны (см., например, [Zaumen 91]), мы должны отвлечься от обсуждения для краткого описания концепций алгоритма PV³.

3.2.1 Обзор протокола Path Vector

Протоколы маршрутизации BGP⁴ [BGP91] и IDRP⁵ [IDRP91] являются примерами использования алгоритма PV. BGP представляет собой протокол маршрутизации между автономными системами в сетях TCP/IP. IDRP представляет собой протокол междоменной маршрутизации OSI, который планируется стандартизировать в ISO⁶. Поскольку в терминах функциональности BGP представляет собой подмножество IDRP, в оставшейся части документа будет обсуждаться только протокол IDRP.

Алгоритм маршрутизации, реализованный в PV, имеет некоторое сходство с традиционным алгоритмом маршрутизации Беллмана-Форда (Bellman-Ford) в том смысле, что каждый граничный маршрутизатор анонсирует соседним BR⁷ доступных для него адресатов. Однако алгоритм маршрутизации PV дополняет анонсы доступных адресатов информацией о свойствах пути к этим адресатам.

Эта дополнительная информация представляется в форме атрибутов пути. Для подчеркивания тесной связи пути к адресату и свойств этого пути PV определяет маршрут как пару «путь — атрибуты пути». Это объясняет название протокола «path-vector» - BR получает от соседних BR вектор, содержащий пути к множеству адресатов⁸. Путь, выражаемый в терминах доменов (или конфедераций), через которые он проходит, передается как специальный атрибут, содержащий последовательность доменов маршрутизации, через которые прошла информация о доступности адресатов. Подавление маршрутных петель реализуется с помощью специального атрибута пути в отличие от алгоритмов LS⁹ и DV¹⁰, использующих монотонно возрастающую глобальную метрику для выбора маршрута [Shin87].

Поскольку PV не требует от всех доменов маршрутизации поддержки однотипных критериев (правил) выбора маршрутов, используемые конкретным доменом правила выбора могут быть неизвестны другим доменам. Для обеспечения максимального уровня автономности и независимости доменов маршрутизации, каждый домен, участвующий в PV, может иметь собственное представление об оптимальности маршрута. Это представление основывается исключительно на локальных правилах выбора маршрутов и информации, содержащейся в атрибутах пути для маршрутов. PV стандартизует только результаты процедуры выбора маршрута, допуская использование в самом процессе выбора нестандартизованных правил¹¹.

3.3 Сложность

На основе приведенного выше описания алгоритма PV мы можем сравнить его с алгоритмами LS в терминах определенных выше параметров сложности.

3.3.1 Издержки на хранение

Без учета агрегирования маршрутной информации и издержек, связанных с ограничениями на транзит можно показать, что в достаточно типовых условиях средние издержки на хранение RIB для схемы PV при одном значении TOS принимают значение в диапазоне от $O(N)$ до $O(N \log(N))$, где N задает общее число доменов маршрутизации [Rekhter91]. LS RIB без агрегирования маршрутной информации, без учета ограничений на транзит при общем для всех доменов наборе правил выбора маршрутов для одного значения TOS потребуются топологическое отображение домена в целом, размер которого составит $O(N)$.

Поддержка гетерогенных правил транзита и выбора маршрута с поэтапной пересылкой и LS требует, чтобы каждый домен знал правила выбора маршрутов и транзита для всех прочих доменов. Это может существенно увеличивать объем маршрутной информации, которая должна храниться в каждом домене. Если число анонсируемых правил растет по мере роста числа доменов, требования по хранению информации могут стать невыполнимыми. Напротив, поддержка гетерогенных правил выбора маршрута не оказывает влияния на сложность хранения информации для схемы PV (нужно хранить только локальную информацию о правилах). Наличие ограничений на транзит в PV приводит

¹ В оригинале «convergence problem». *Прим. перев.*

² Path Vector - вектор путей.

³ Мы предполагаем что для расчета путей по топологической базе при использовании алгоритмов LS будет применяться та или иная форма алгоритма SPF [Dijkstra59, OSPF].

⁴ Border Gateway Protocol — протокол граничного шлюза.

⁵ Inter Domain Routing Protocol — протокол междоменной маршрутизации.

⁶ В настоящее время протокол IDRP стандартизован, как ISO/IEC 10747:1994. *Прим. перев.*

⁷ Border Router -граничный маршрутизатор. *Прим. перев.*

⁸ Термин «path-vector protocol» имеет случайное сходство с термином «distance-vector protocol». В последнем случае BR получает от каждого из своих соседей вектор, содержащий «расстояния» по множества адресатов.

⁹ Link State — состояние канала. *Прим. перев.*

¹⁰ Distance Vector — вектор удаленности. *Прим. перев.*

¹¹ Это краткое пояснение адресовано Ross Callon (Digital Equipment Corporation).

к ограничению распространения маршрутной информации, что дополнительно снижает требования по хранению. В противоположность этому в LS такое снижение требований невозможно, поскольку каждый домен должен знать правила транзита для всех остальных доменов. Наконец, некоторые из транзитных ограничений (например, ограничения, связанные с путем), в PV могут выражаться более компактно (см. обсуждение агрегирования ниже).

Способность дополнительного снижения издержек на хранение обеспечивается алгоритмом PV за счет того, что расчет маршрутов выполняется до их распространения и в соседние домены передается только информация о маршрутах, выбранных доменом. В противоположность этому выбор маршрута с использованием LS оторван от распространения маршрутной информации и не оказывает влияния на это распространение.

Хотя в теории агрегирование маршрутной информации может снижать издержки на хранение как для LS, так и для PV, только агрегирование топологической информации дает одинаковый эффект в обоих алгоритмах. Агрегирование транзитных ограничений при использовании LS может приводить к потере информации, но при использовании PV этого не происходит.

Агрегирование гетерогенных правил выбора маршрутов весьма проблематично для LS. В PV распространения правил выбора маршрутов не происходит, поэтому агрегирование правил выбора не создает проблем¹.

Поддержка разных значений TOS имеет одинаковое влияние на LS и PV.

Потенциально размер LS FIB можно уменьшить при расчете маршрутов на каждом узле по запросу. Однако реальный эффект будет сильно зависеть от картины трафика, размера памяти и стратегии кэширования. Если трафик не является в значительной степени локальным, может наблюдаться некоторое снижение производительности за счет избыточного расхода ресурсов на вычисление воздействий при использовании PV может быть меньше, чем при традиционной поэтапной маршрутизации на основе LS. При использовании PV необходимость пересчета маршрутов возникает только в тех доменах, чьи маршруты были изменены, тогда как при традиционной поэтапной маршрутизации с использованием LS требуется пересчет во всех доменах. Хотя в LS также можно реализовать частичный пересчет (например, при изменении топологии пересчитываются только измененные маршруты), "изучение показывает, что изменение незначительного числа каналов (возможно 2) ожидаемая сложность расчетов нарастающих обновлений превышает сложность полного пересчета" [ANSI87-150R]. Однако такие проверки существенно проще, нежели полное выполнение процессов SPF.

3.3.2 Сложность расчета маршрутов

Даже при использовании одинаковых правил выбора маршрутов во всех доменах сложность расчета маршрутов в PV меньше, нежели в LS. Расчет в PV состоит из оценки недавно полученного маршрута и сравнения его с имеющимся², тогда как расчет в LS требует использования алгоритма SPF³ (например, Dijkstra).

При использовании PV изменения топологии вызывают необходимость пересчета только для тех маршрутов, на которые воздействуют эти изменения, что более эффективно, нежели полный пересчет. Однако за счет того, что в каждый вектор удаленности включается полная информация о пути, эффект изменения топологии может простираться существенно дальше, нежели для традиционных алгоритмов на основе вектора удаленности. С другой стороны, число доменов, на которые изменения воздействуют при использовании PV может быть меньше, чем при традиционной поэтапной маршрутизации на основе LS. При использовании PV необходимость пересчета маршрутов возникает только в тех доменах, чьи маршруты были изменены, тогда как при традиционной поэтапной маршрутизации с использованием LS требуется пересчет во всех доменах. Хотя в LS также можно реализовать частичный пересчет (например, при изменении топологии пересчитываются только измененные маршруты), "изучение показывает, что изменение незначительного числа каналов (возможно 2) ожидаемая сложность расчетов нарастающих обновлений превышает сложность полного пересчета" [ANSI87-150R]. Однако такие проверки существенно проще, нежели полное выполнение процессов SPF.

Поддержка гетерогенных правил выбора маршрутов оказывает значительное влияние на сложность расчетов. Основная проблема поддержки гетерогенных маршрутов с LS заключается в сложности расчетов при выборе маршрута. В частности, нет уверенности в том, что какой-либо алгоритм сможет обеспечить менее, чем экспоненциальный рост времени расчета маршрутов ко всем адресатам при использовании поэтапной маршрутизации LS с гетерогенными правилами выбора маршрутов. В противоположность этому PV позволяет каждому домену выбирать маршруты автономно, на основе локальных правил. Следовательно, поддержка различных правил выбора маршрутов не оказывает влияния на сложность расчета маршрутов в PV. Аналогично, поддержка зависимых от пути правил транзита в LS ведет к экспоненциальному росту сложности, тогда как в PV такая поддержка не оказывает влияния на сложность расчета маршрутов.

Поскольку NR опирается прежде всего на предварительный расчет маршрутов, агрегирование оказывает существенное влияние на долгосрочную масштабируемость архитектуры. В известной степени агрегирование упрощается при использовании PV за счет снижения сложности вычислений. Хотя подобные аргументы можно использовать и по отношению к LS, возможности агрегирования в LS более проблематичны, поскольку LS опирается на согласованные топологические планы каждого узла. Не очевидно также повышение сложности расчетов, связанное с транзитными ограничениями и неоднородными правилами выбора маршрутов в LS.

3.3.3 Дополнительный расход полосы

Маршрутные обновления PV распространяются в открытом виде⁴. Анонсируется полный маршрут для каждого поддерживаемого TOS. В LS значение TOS лишь несколько увеличивает размер маршрутного анонса, что значительно меньше числа анонсируемых маршрутов. Хотя значения TOS в LS могут кодироваться более эффективно, нежели в PV, информация о состоянии каналов лавинно рассылается во все домены, тогда как в PV обновления маршрутов распространяются только в те домены, которые эти маршруты используют. Следовательно, трудно в общем виде сравнить дополнительный расход полосы для разных схем, когда все прочие условия совпадают.

Более того, эта разница может совсем не иметь значения, поскольку интерес представляет скорей число сообщений, чем их размер в битах (благодаря компрессии и расширению полосы каналов, а также повышению их стабильности).

¹ Хотя принятые в домене правила выбора маршрутов не распространяются в явном виде, они оказывают влияние на маршруты, доступные для других доменов. Маршрут, который может быть предпочтителен для отдельного домена и не запрещен правилами транзита, может оказаться недоступным в результате применения правил выбора некоего промежуточного домена. Возможность расчета и организации маршрутов, которые могут быть утраченными при поэтапной маршрутизации (как LS, так и PV), является важной особенностью SDR.

² При получении обновлений требуется выполнить некие дополнительные проверки для предотвращения петель.

³ Shortest Path First — сначала кратчайший путь. *Прим. перев.*

⁴ В оригинале «fully-expanded information». *Прим. перев.*

3.4 Агрегирование

Формирование конфедераций доменов в целях согласованного расчета поэтапных маршрутов LS требует согласованности правил между доменами конфедерации. Кроме того для конфедераций LS требуется, чтобы объект некоего уровня был членом единственного объекта более высокого уровня. В общем случае внутренние данные конфедерации не могут быть доступны за пределами этой конфедерации, поскольку в противном случае могут возникать петли в результате использования несогласованных планов сетей в разных доменах. Следовательно, использование конфедераций с поэтапной маршрутизацией LS ограничено, поскольку каждый домен (или конфедерация) может быть частью единственной конфедерации более высокого уровня и может экспортировать только правила, согласованные с этой конфедерацией (см. примеры в параграфе 2.2). Очевидно, что эти ограничения достаточно сильно влияют на масштабируемость архитектуры.

Для сравнения PV может принимать разные определения конфедераций, поскольку предотвращение петель обеспечивается за счет передачи полной информации о пути. Согласованные планы сетей не требуются на каждом сервере маршрутов, поскольку расчет маршрутов предшествует их распространению. Следовательно, конфедерации PV могут быть вложенными, разорванными¹ и перекрывающимися. Домен (или конфедерация) может экспортировать разные правила или TOS как часть разных конфедераций, что обеспечивает максимальную гибкость, являющуюся критически важной с точки зрения масштабирования и расширяемости архитектуры.

Таким образом, агрегирование важно для достижения приемлемых пределов сложности, а гибкость важна для обеспечения приемлемого агрегирования в пределах глобальной, децентрализованной сети. Основные преимущества алгоритма PV обусловлены его гибкостью.

3.5 Правила

Необходимость выражения правил транзитных ограничений для любого маршрута (как NR, так и SDR), сама по себе, может поддерживаться в LS и PV. Однако связанная с этим сложность поддержки для LS существенно выше, чем для PV. Это обусловлено необходимостью лавинной рассылки всех транзитных правил в LS, тогда как в PV распространение правил транзита контролируется ограниченным распространением маршрутной информации. Ограниченное распространение всегда связано с меньшими издержками, нежели лавинная рассылка.

Все транзитные ограничения LS могут поддерживаться в PV, но не наоборот. Иными словами, существуют некоторые транзитные ограничения (например, связанные с конкретным путем), которые легко реализуются в PV, а их поддержка в LS связана со значительными издержками. Например, не очевидно как NR при использовании LS сможет поддерживать правила в стиле ECMA, основанные на иерархических отношениях между доменами, хотя такие правила достаточно просто поддерживаются в PV.

Как показано выше, поддержка гетерогенных правил выбора маршрутов непрактична для поэтапной маршрутизации LS по причине значительных издержек на расчеты и хранение маршрутов. Напротив, PV поддерживает гетерогенный выбор маршрутов без дополнительных издержек.

3.6 Соккрытие информации

PV имеет явные преимущества в части сокрытия указанной части информации. Поэтапная маршрутизация LS основана на возможности поддерживать во всех доменах в точности совпадающую информацию — это вступает в явное противоречие с возможностью сокрытия заданной части информации. Т. е., требование сокрытия части информации не может быть удовлетворено для поэтапной маршрутизации LS.

3.7 Общее в компонентах NR и SDR

В работе [Breslau-Estrin91] рассмотрено использование алгоритма LS для маршрутизации SDR. Следовательно, имеются некоторые предпочтения для использования LS с маршрутизацией NR. Однако в силу некоторых причин маршрутная информация NR и SDR не совпадает в точности даже при использовании одного алгоритма. Более того, имеется несколько предложений по унификации обработки (распространение и хранение) маршрутной информации и информации о пересылке даже при использовании разных алгоритмов.

При рассмотрении различий между NR и SDR мы должны обратить внимание на несколько аспектов:

1. **Маршрутная информация и протокол распространения:** LS для SDR в достаточной степени отличается от использования LS при маршрутизации NR. Например, в SDR алгоритму LS не требуется агрегировать домены; напротив в SDR алгоритму LS требуется детальная информация для генерации специальных маршрутов. В дополнение к этому существенные для NR требования согласованности не имеют значения для компоненты SDR. Следовательно информация LS для компоненты SDR может быть получена по запросу, тогда как компонента NR должна использовать лавинную рассылку топологической информации.
2. **Алгоритм расчета маршрутов:** Не очевидно в каких случаях компоненты SDR и NR могут использовать общие алгоритмы расчета маршрутов в связи с трудностями поддержки гетерогенных правил выбора маршрутов в NR.
3. **Информация о пересылке:** Использование различных алгоритмов расчета маршрутов не препятствует однотипной пересылке пакетов в обеих компонентах. Даже при использовании LS с компонентой NR, требования будут такими же, т. е., агент пересылки может определить, когда следует использовать предварительно рассчитанные маршруты NR, а когда - организовать маршрут SDR для пересылки конкретных пакетов данных.

В заключение отметим, что использование похожих алгоритмов и механизмов в компонентах SDR и NR будет давать преимущества. Однако эти преимущества не будут доминировать над другими факторами, рассмотренными выше.

3.8 Резюме

С учетом издержек, связанных с глобальной маршрутизацией, агрегирование маршрутной информации имеет важное значение; в то же время, мы отмечали, что такое агрегирование должно быть гибким. В связи с трудностями поддержки

¹ Состоящими из несмежных частей. *Прим. перев.*

поэтапной маршрутизации LS при наличии (а) гибкого агрегирования, (b) гетерогенных правил выбора маршрутов и (c) неполноты или несогласованности маршрутной информации, мы не видим альтернативы использованию алгоритма PV для компоненты NR в предлагаемой архитектуре.

С учетом отмеченного выше наша компонента NR использует архитектуру PV, в которой расчет и организация маршрутов осуществляется в распределенной манере маршрутизаторами, принимающими участие в работе NR-компоненты¹.

Распределенный алгоритм сочетает в себе некоторые свойства алгоритмов на основе состояния канала и алгоритмов на основе вектора удаленности — в дополнение к информации о следующем интервале (next hop) компонента NR поддерживает для каждого маршрута атрибуты пути (например, список доменов или доменных конфедераций, через которые проходит маршрутная информация). Атрибуты пути, передаваемые по маршруту выражают широкий спектр правил маршрутизации и делают явным весь маршрут к получателю. При использовании агрегирования это является надмножеством доменов, которые формируют путь к получателю.

4.0 Маршрутизация по запросу отправителя (SDR)

Междоменные маршрутизаторы, участвующие в работе компонента SDR, пересылают пакеты в соответствии с маршрутной информацией, рассчитанной и организованной доменами, из которых исходит трафик (доменами source routing).

Важно, чтобы требуемая от исходного домена организация маршрута была не вопросом выбора, а необходимостью. Если конкретный маршрут используется незначительным числом доменов (в пределах одним доменом), более приемлемым вариантом является расчет и организация специального маршрута вместо того, чтобы нагружать промежуточные узлы задачей поиска и организации маршрута в соответствии со специальными требованиями. В дополнение к этому, если запрашиваемый маршрут не является предсказуемым и, следовательно, может быть определен только отправителем, источник пакетов сам должен организовать такой маршрут.

В общем случае информация, используемая доменами source routing для расчета маршрута по запросу отправителя, отражает административное (но не рабочее) состояние элементов маршрутизации (т. е., заданной топологии и правил)². Следовательно, для домена source routing возможен расчет маршрута, который не является работающим на момент организации. Компонента SDR пытается уведомить исходный домен об отказе при попытке организации маршрута. Подобно этому, компонента SDR обеспечивает механизмы для уведомления домена source routing об отказах с ранее организованными активными маршрутами. Иными словами, компонента SDR выполняет маршрутизацию, которая адаптируется к изменениям топологии; однако такая адаптивность достигается, как следствие механизмов организации маршрутов и управления ими. Это отличается от поведения компонента NR, где изменения состояния распространяются и принимаются узлами по мере возможности. Следовательно, для ускорения адаптации к изменениям в рабочем состоянии элементов маршрутизации, компонента SDR позволяет исходному домену переключаться на маршрут, рассчитанный компонентой NR, если обнаружен отказ на запрошенном отправителем маршруте (на фазе организации маршрута или после его установки) и правила разрешают использовать маршрут NR.

Для решения задач масштабирования компонента NR будет группировать модемы в конфедерации [IDRP91]. Напротив, SDR будет разрешать использование маршрутов уровня AD³ отдельными доменами, не разрешая использовать такой маршрут всей конфедерации, к которой домен относится. Подобно этому, один транзитный домен может поддерживать правило или специальное значение TOS, которые не поддерживаются другими доменами в их конфедерации(ях). Иными словами, архитектура использует SDR для поддержки неиерархических, неагрегируемых правил тогда и там, когда и где они требуются. Следовательно, компонента SDR сама по себе не обеспечивает свойств масштабирования, присущих NR. В компенсацию SDR не требует полного, глобального плана доменов если отдельные фрагменты сети никогда не пересекаются пакетами и не передаются туда. В результате утраты структуры маршрутизации SDR не гарантирует, что принимающие участие в маршрутизации домены всегда будут иметь информацию, требуемую для расчета пути к адресату. Кроме того, если домен имеет всю требуемую информацию, ее объем может оказаться слишком велик для своевременного сохранения и/или расчета маршрута. Однако, несмотря на отсутствие гарантий, целью архитектуры является обеспечение эффективных методов, с помощью которых отправитель может получить информацию, требуемую для расчета нужных маршрутов⁴.

Существенным для SDR является допущение о том, что маршруты, организованные по запросу, будут использоваться бережливо. Архитектура предполагает, что в любой заданный момент времени набор из всех установленных по запросам маршрутов в глобальной сети, образует малую часть маршрутов по запросу, которые могут быть потенциально организованы всеми доменами. Архитектура предполагает, что число маршрутов, установленных в BR⁵ компонентой SDR, должно иметь значение порядка $\log N$ (где N — общее число доменов маршрутизации в Internet), чтобы параметры масштабирования компонента SDR были сравнимы с параметрами масштабирования NR. Компонента NR обеспечивает масштабируемость за счет иерархии.

Отметим, что приведенное выше требование не означает, что лишь малое число доменов может принимать участие в SDR или что маршруты, организованные компонентой SDR должны иметь малое время жизни. Все, что следует из этого требования — это то, что произведение числа маршрутов, задаваемых доменами, которые участвуют в SDR, на среднее время жизни маршрута SDR является ограниченным. Например, архитектура позволяет существовать небольшому числу маршрутов SDR со сравнительно долгим временем жизни или множеству короткоживущих

¹ Пересылка пакетов по маршрутам, создаваемым компонентой NR, может осуществляться с использованием маршрутизации, задаваемой отправителем (source routing) или поэтапной (hop-by-hop) маршрутизации. Второй вариант является более предпочтительным, поскольку он обеспечивает снижение числа состояний в маршрутизаторах (если используется организация маршрута) и позволяет избежать явного включения маршрута в пакеты сетевого уровня. Однако архитектура не запрещает маршрутизации, задаваемой отправителем (или организации маршрута) по маршрутам рассчитанным, но не организованным компонентой NR.

² Если SDR использует информацию NR, рабочее состояние можно рассматривать как выбор некоего маршрута.

³ Административный домен — Administrative domain. *Прим. перев.*

⁴ Основной целью правила или TOS-маршрутизации является расчет маршрута, который удовлетворяет набору заданных требований, и эти требования имеют преимущество по отношению к оптимизации маршрута. Иными словами, даже если домен, участвующий в SDR или NR, имеет достаточно информации для расчета маршрута, с учетом заданного набора требований архитектура не гарантирует оптимальности рассчитанного маршрута.

⁵ Граничный маршрутизатор — Border router. *Прим. перев.*

маршрутов. Однако архитектура явно запрещает установку множества маршрутов SDR с относительно большим временем жизни. Число маршрутов SDR является функцией от числа доменов, использующих маршруты SDR, и числа маршрутов, используемых в расчете на один домен-отправитель.

В заключение отметим, что SDR хорошо подходит для трафика, который (1) используется недостаточно широко (или слабо предсказуем, или не имеет устоявшегося характера) для использования компоненты NR и (2) имеет продолжительность, существенно превышающую время на организацию маршрута.

Архитектура не требует от всех доменов Internet принимать участие в SDR. Следовательно, вопросы масштабирования (в части размера Internet) становятся менее критичными (хотя остаются достаточно важными) для компоненты SDR. Вместо этого основной фокус SDR смещается в направлении возможности расчет маршрутов, удовлетворяющих заданным требованиям, в предположении, что общее число доменов, которым одновременно требуются специальные маршруты невелико по сравнению с общим размером глобальной сети.

4.1 Сравнение алгоритмов Path Vector и Link State для SDR

При задаваемой отправителем маршрутизации для расчета маршрутов можно использовать как методы, основанные на состоянии каналов, так и методы, основанные на векторах удаленности. Целесообразно использовать метод вектора удаленности или состояния канала при расчете пути с задаваемой отправителем маршрутизацией (source routing). Например, протокол типа BGP, в котором отправитель использует полную информацию AD о пути, полученную в маршрутных обновлениях, для создания source route. Такой протокол позволяет преодолеть некоторые недостатки, связанные с поэтапной работой алгоритмов на основе векторов удаленности. Однако мы не будем более обсуждать такие протоколы, поскольку можно добиться больших результатов при использовании задаваемой отправителем маршрутизации без применения алгоритмов на основе состояния канала. Возможности задаваемой отправителем маршрутизации в контексте внутримоментной (AD) маршрутизации на базе правил дают отправителю контроль над всем маршрутом. Эта задача не может быть решена полностью, когда промежуточные узлы контролируют анонсирование маршрутов соседям и, следовательно, отправителю.

Иными словами, промежуточные узлы должны быть способны предотвратить использование маршрута путем задания правил транзита, но если маршрут не заблокирован (т. е., является корректным для всех AD на маршруте), маршрут следует сделать доступным для отправителя независимо от предпочтений промежуточных доменов для их собственных потоков трафика.

Следовательно, компонента SDR использует архитектуру типа IDPR, в которой обновления информации о состоянии каналов распространяются с явным указанием правил в каждом обновлении вместе с анонсированием связности узлов.

4.2 Распространение маршрутной информации

За счет использования поэтапной маршрутизации компоненты NR, основанной на алгоритме PV в дополнение к задаваемой отправителем маршрутизации компоненты SDR мы уменьшаем сложности, связанные с агрегированием информации о пересылке SDR — большой процент междоменного трафика, передаваемого в каждый момент любым граничным маршрутизатором, будет пересылаться с использованием агрегированной информации о пересылке компоненты NR. Однако использование NR не решает другой важной проблемы масштабирования, связанной с SDR, — распределения, хранения и расчета с использованием полного топологического плана на уровне домена. В этом параграфе описаны перспективные предложения по расширению возможностей масштабирования при распределении, хранении и расчете маршрутной информации SDR.

Отметим, что для решения этой проблемы мы не предлагаем такой же путь, который используется для NR. А priori здесь не применима, поскольку разные домены могут требовать различных методов абстрагирования маршрутной информации. Например, если агрегируется маршрутная информация доменов, в которых не используются одинаковые правила и характеристики TOS (т. е., службы), тогда за пределами агрегата будут анонсироваться только те службы, которые анонсируются всеми доменами агрегата. Для размещения специальных маршрутов SDR использует агрегаты только тогда, когда домены-компоненты (и агрегат в целом) анонсируют требуемые параметры TOS и правила. Когда требуемые параметры TOS или правила не обеспечиваются агрегатом, используется полная информация о доменах-компонентах для создания маршрута через те домены, которые поддерживают требуемые характеристики. Следовательно, требуется иной, более гибкий, способ снижения объема хранимой и распространяемой информации. Таким образом решается два вопроса — распространение заданной конфигурацией топологической информации и правил, а также распространение динамической информации о состоянии.

4.2.1 Конфигурационные параметры

Информация о существовании междоменных каналов и поддерживаемых доменами правилах изменяется достаточно медленно. Такого типа информация относится к конфигурационным параметрам. В текущей спецификации IDPR полная, глобальная конфигурационная информация хранится сервером маршрутов каждого домена. Серверы маршрутов (RS¹) представляют собой объекты, которые рассчитывают маршруты, задаваемые отправителем. При старте RS может загружать базу данных о соединениях с соседнего RS, по мере изменения доменов, междоменных соединений или правил осуществляется лавинная рассылка данных об изменениях маршрутным серверам в каждом домене.

Мы еще не задали точный механизм распределения конфигурационных параметров связности для SDR. Однако, в отличие от текущей спецификации IDPR, компонента SDR не будет использовать глобальную лавинную рассылку всех конфигурационных параметров. Исследуется несколько методов организации и распространения такой информации.

Конфигурационные параметры могут регулярно распространяться за пределами основного канала (например, на компакт-дисках). Можно также передавать эту информацию в несколько заранее согласованных мест, например, на серверы FTP. В промежутках между такими «серьезными» обновлениями агрегированные наборы обновлений могут глобально рассылаться в лавинном режиме. Более того, может использоваться ограниченная лавинная рассылка (например, на заданное число интервалов) в зависимости от важности внесенных изменений — информация об изменении правил для основной магистрали может рассылаться глобально, информация о новом межрегиональном

¹ Route server — сервер маршрутов. Прим. перев.

канале может рассылаться только в связанных с этим каналом регионах, а информация о новом канале в домене, не являющемся транзитным, станет доступной только после очередного регулярного «серьезного» обновления.

Изменения, информация о которых не распространяется сразу же, нет необходимости отыскивать. Однако сервер маршрутов может получить соответствующую информацию путем прямого запроса от удаленных серверов или из сообщения об ошибке, возникшей в результате передачи трафика по некорректному маршруту. Полной глобальной лавинной рассылки можно избежать за счет использования комбинации таких механизмов.

Даже если в начальной реализации используется простая глобальная лавинная рассылка, мы должны изучить проблему структурирования информации о связности, чтобы эти данные можно было отыскивать и распределять более селективно с сохранением для отправителей возможности нахождения желаемых маршрутов. Например, можно рассмотреть RS, запрашивающие один у другого отфильтрованную информацию. Как в таких случаях RS следует определять фильтры, чтобы получить информацию, достаточную для нахождения специального маршрута при эффективном ограничении объема передаваемой информации? Другой вопрос связан с эффективным предвидением и описанием информации, которая может потребоваться для расчета маршрута.

Проблема состоит еще и в том, что сети не имеют согласованной географической и топологической организации, следовательно исходный домен не знает, какая информация ему нужна (что следует запрашивать), пока он не найдет нужный путь. Даже при наличии централизованной базы данных нужны методы структурирования конфигурационных параметров так, чтобы пути с наибольшей вероятностью использования обнаруживались первыми, что позволило бы снизить расход времени на расчет маршрутов.

Одним из перспективных предложений является организация информации с использованием фрагментов маршрутов¹ (частичных путей). Хотя число маршрутных фрагментов растет быстрее, чем число доменов (по крайней мере, как $O(N^2)$), мы можем выбрать те фрагменты, которые будут полезны при расчете маршрутов. В частности, для каждого тупикового домена фрагменты будут содержать пути к нескольким хорошо известным магистральям². К преимуществам этого способа относятся агрегирование доменной информации в манере, полезной для расчета маршрутов, задаваемых отправителем, и индексирование по адресатам, которое обеспечивает ссылки на информацию и нахождение данных, требуемых для расчета конкретного маршрута. В настоящий момент пока не очевидно, как фрагменты маршрутов будут влиять на способность SDR обнаруживать неиерархические маршруты.

4.2.2 Динамические данные о состоянии

Кроме метода глобального или частичного распространения конфигурационных параметров второй проблемой является решение вопроса о том, кто и как будет распространять динамическую информацию о состоянии (т. е., активизации или разрыве междоменных соединений).

В текущей версии IDPR динамическая информация о состоянии рассылается глобально в лавинном режиме в дополнение к конфигурационной информации. Мы предлагаем распространять информацию о состоянии в жесткой зависимости от местоположения. Во-первых, динамическая информация будет анонсироваться в малой окрестности (незначительное число интервалов). Этот простой механизм с малыми издержками использует механизм топологической локализации. В дополнение к лавинной рассылке обновлений состояния в ближайшей окрестности мы также хотим обеспечить более точную маршрутную информацию для более протяженной окрестности (дальше нескольких интервалов). RPU³ представляет собой механизм передачи динамической информации о состоянии узлам, находящимися за пределами окрестности радиусом k интервалов (hop), используемого для обновлений, но этот механизм никогда не будет обеспечивать лучшего сервиса (меньшего числа установок некорректных маршрутов) чем получение доступа к динамической информации о состоянии [Estrinetal91].

RPU использует существующие активные маршруты (представленные состоянием организованных маршрутов или содержимым кэша недавно использованных маршрутов source route, переданных через искомый узел) в качестве рекомендации для распространения уведомлений о событиях. Вместо отчетов лишь о состоянии маршрутов, которые будут использоваться, RPU сообщает о состоянии других междоменных соединений домена. Если заданная отправителем маршрутизация показывает местоположение маршрута, для отправителя более очевидно использование других маршрутов через искомый узел; в любом случае избыточная информация о других узлах имеет минимальный объем.

В этом случае отправители будут получать информацию о состоянии из тех областей сети, через которые они поддерживают активные маршруты, даже если эти области находятся за пределами k интервалов. При использовании такой схемы значение k может быть достаточно малым для обеспечения эффективности, а RPU можно использовать для снижения числа инцидентов, связанных с некорректными маршрутами, созданными благодаря неточным данным о состоянии. Это будет полезно, если при длинных путях будет показываться местоположение маршрутов относительно областей, которые ближе к точке назначения (и, следовательно, находятся за пределами k -окрестности лавинной рассылки информации). В таких ситуациях лавинная рассылка информации источникам длинных маршрутов будет неэффективной, поскольку значение k будет равно длине маршрута и почти во всех случаях процент узлов, использующих информацию значительно снижается по мере роста значения k .

4.3 Управление SDR

SDR можно построить поверх сетевого уровня, поддерживаемого компонентой NR, или параллельно этому уровню. Пересылка SDR будет поддерживаться двумя методами - loose source-routing⁴ и route setup⁵.

Первый метод будет позволять источнику пакета задать последовательность доменов, через которые пакет должен проходить по пути к адресату. Пересылка такого пакета внутри домена или даже между доменами одной конфедерации будет выполняться средствами внутридоменной маршрутизации. Это позволяет избавиться от избыточного трафика, связанного с поддержкой информации о состояниях соединений и поддержкой транзакций.

¹ Понятие маршрутных фрагментов изначально предложили Dave Clark и Noel Chiappa.

² Фрагменты маршрута могут быть рассчитаны сервером маршрутов на стороне получателя и сделаны доступными через информационные запросы или путем глобальной информационной рассылки. В дополнение к этому маршруты, рассчитанные компонентой NR, также могут использоваться в качестве фрагментов маршрутов SDR.

³ Reverse path update — обновление обратного пути.

⁴ Нечеткая маршрутизация со стороны отправителя.

⁵ Организация маршрута.

Второй метод основан на механизмах, предложенных для IDPR и описанных в [IDPR90]. Этот метод хорошо подходит для сеансов, продолжительность которых значительно превышает время кругового обхода. Протокол организации маршрута определяет форматы пакетов и обрабатывает пакеты организации маршрута (т. е., пакеты setup). Когда источник генерирует пакет setup, первый граничный маршрутизатор вдоль заданного source route проверяет запрос на организацию маршрута и, если этот запрос приемлем, устанавливает маршрутную информацию, которая включает идентификатор пути, предыдущий и следующий маршрутизаторы и другую связанную с учетом информацию, которую требует конкретный домен. Пакет setup передается следующему BR в source route доменного уровня, где описанная процедура повторяется¹. Когда установочный пакет достигает адресата, в обратном направлении отправляется подтверждение², которое поэтапно проходит через маршрутизаторы и каждый BR на пути активизирует свою маршрутную информацию. Последующие пакеты данных, проходящие по тому же пути к адресату, будут включать в себя идентификатор пути. Этот идентификатор используется для нахождения соответствующей информации о следующем интервале для каждого пакета.

Граничные маршрутизаторы, которые поддерживают обе компоненты - NR и SDR, должны быть способны определить, какой механизм следует использовать. При наличии соответствующей информации в PDU сетевого уровня такой BR должен быть способен принять однозначное решение о пересылке данного PDU с использованием компоненты NR или SDR. Механизм определения компоненты зависит от того, реализован сетевой уровень компоненты SDR поверх сетевого уровня NR или параллельно ему. После того, как решение принято, пакеты, которые пересылаются с использованием маршрутов, организованных компонентой SDR, пересылаются в выходной порт, связанный со значением Path ID в заголовке пакета. Пакеты, которые будут пересылаться с использованием маршрутов NR, пересылаются в выходной порт, связанный с соответствующим адресатом и значением ToS (если оно присутствует) в заголовках пакетов.

Далее мы опишем основные различия между процедурой организации маршрутов IDPR и процедурой, предложенной здесь для развития гибридной архитектуры.

Если в процессе организации маршрута BR обнаруживает, что остальная часть (после BR) указанного маршрута от BR до адресата совпадает с маршрутом NR от BR до адресата, данный BR может отключить маршрут SDR в данной точке и отобразить его на маршрут NR. Чтобы это произошло, спецификация маршрута SDR должна полностью соответствовать маршруту NR. В дополнение к этому весь маршрут пересылки (т. е., оставшиеся до адресата интервалы) должен совпадать.

Более того, если маршрут NR изменяется на протяжении активного маршрута SDR и новый маршрут NR не совпадает с маршрутом SDR, тогда для оставшейся части пути к адресату должен быть организован маршрут SDR. Следовательно, при отображении маршрута SDR на маршрут NR исходный пакет setup должен быть сохранен. Пакеты на пути от отправителя к получателю могут, следовательно, проходить сегменты маршрутов SDR и NR, однако пакет не будет проходить других сегментов SDR после его прохождения через сегмент NR. Тем не менее, в течение переходного периода пакеты могут идти по некорректным маршрутам и, следовательно, эта функция должна быть опциональной и управляемой.

Отправитель может также запросить уведомление, если через некоторое время после неудачной организации маршрута неработоспособный канал или узел возвращается в работоспособное состояние. Если BR на маршруте обнаруживает, что на следующем интервале BR недоступен, обнаруживший проблему BR может добавить отправителя в список уведомлений и когда следующий BR снова станет доступным, отправителю может быть передано уведомление. Этот механизм обеспечивает возможность получения информации о прекращении неполадок. Например, домен может принять решение о маршрутизации через резервный путь при отказе на основном маршруте — упомянутый выше механизм будет своевременно информировать отправителя о восстановлении предпочтительного маршрута.

Третья опция обеспечивает адаптацию после организации маршрута. Если BR в процессе пересылки пакетов по активному маршруту SDR обнаруживает отказ на маршруте SDR, он может перенаправить трафик к адресату в существующий маршрут NR. Такая адаптация допустима лишь в том случае, когда используемый маршрут NR не противоречит правилам (например, на маршруте могут быть неприемлемые условия обслуживания — ToS). Адаптация происходит лишь в тех случаях, когда отправитель выбрал соответствующую опцию при организации маршрута. Отправитель также задает для себя режим уведомления о случаях описанной выше смены маршрута.

При отказе на маршруте SDR обнаруживший отказ BR передает уведомления отправителям на активных маршрутах, на которые воздействует отказ. В дополнение к этому BR может включить в уведомление информацию о состоянии других BR того же домена. В частности, BR может включить наиболее свежее обновление для своего домена, показывающее состояние междоменных соединений и правил. Это может быть полезно для обхода сбойных узлов, т. е., показывает возможность использования альтернативных маршрутов, проходящих через нужный домен, но не затронутых отказом.

В заключение отметим, что при организации маршрута отправитель имеет несколько опций, представляемых в виде флагов в пакете setup:

1. Если доступен маршрут NR, удовлетворяющий всем локальным правилам и TOS, использовать этот маршрут.
2. Флаг согласия отправителя на использование принятого по умолчанию маршрута NR при отказе в процессе организации маршрута SDR.
3. Запрос уведомлений об отображении в маршрут NR.
4. Запрос дополнительной конфигурационной информации при отказе.
5. Запрос на включение в список рассылки уведомлений о доступности ресурсов.
6. Разрешение на перемаршрутизацию пакетов в NR при отказе после организации маршрута (в течение времени, пока не будут нарушаться правила).
7. Запрос уведомления при перемаршрутизации пакетов данных.
8. Запрос дополнительной конфигурационной информации при отказе, когда маршрут активен.

¹ Для снижения задержки начальный пакет (setup packet) может быть передан до завершения проверок.

² accept message — сообщение о восприятии.

9. Запрос на добавление в список рассылки уведомлений при отказе на активном маршруте.

5.0 Унифицированная архитектура

В дополнение к намеченным здесь направлениям развития и реализации архитектуры будущие исследования должны включать работы по унификации компонент предложенной архитектуры. Ниже перечислен ряд предложений общего характера:

1. **База маршрутной информации:** Возможно удастся использовать одну RIB для NR и SDR. Маршруты NR могут представляться как прямые графы, помеченные флагами (на узлах или каналах), соответствующими базовым ограничениям на транзит. SDR требует, чтобы к этому графу были добавлены каналы с отличными от базовых правилами, которые были найдены и организованы для специальных маршрутов. Кроме того, могут добавляться специальные флаги правил к каналам, уже организованным компонентой NR.
2. **Распространение информации: Векторы** пути NR могут включать адреса репозитория обновлений SDR для каждого AD (или конфедерации), чтобы помочь компоненте SDR в получении требуемой информации по запросу. Для доменов с минимальными правилами, где пространство для хранения информации меньше, чем пространство, требуемое для хранения адресов репозитория (например, если правила для домена выражены простым шаблоном), векторы пути NR могут включать соответствующий флаг.
3. **Пересылка пакетов:** Следует рассмотреть вопрос замены текущего заголовка сетевого уровня в стиле IDPR (который содержит глобальный идентификатор пути, используемый при пересылке пакетов данных следующему шлюзу правил на маршруте IDPR) стандартным заголовком (например, IP или CLNP), дополненным некоторыми опциональными полями. Это позволит унифицировать разбор заголовков пакетов и функции пересылки для SDR и NR, а также (возможно) снизить издержки на инкапсуляцию.
4. **Информация о доступности:** В настоящее время IDPR распространяет информацию о доступности сетей в обновлениях, тогда как IDPR распространяет только информацию о доступности доменов. IDPR использует службу доменных имен для отображения номеров сетей на номера доменов, требуемые для принятия решения о маршрутизации. Следует рассмотреть вопрос о получении информации о доступности сетей и доменах в одинаковой манере.

5.1 Применимость для различных протоколов сетевого уровня

Предложенная здесь архитектура разработана для аккомодации таких протоколов сетевого уровня, как IP [Postel81], CLNP [ISO-473-88] и ST-II [ST2-90]. Кроме того, эта архитектура предназначается для поддержки будущих механизмов сетевого уровня, например, предложений Clark и Jacobson или интегрированного сервиса IP, предложенного Braden и Casneg. Однако мы не можем дать никаких гарантий относительно самих механизмов. В любом случае не все упомянутые протоколы будут способны использовать возможности, обеспечиваемые архитектурой. Например, пока нет соответствия между растущим числом предлагаемых сетевых служб и способностью конкретного протокола сетевого уровня однозначно выражать запросы таких служб, способность архитектуры поддерживать маршрутизацию при наличии множества разнотипных служб остается по большей части академической. Т. е., не все компоненты архитектуры будут одинаково важны для разных протоколов сетевого уровня. С другой стороны, эта архитектура разработана для работы в глобальной сетевой среде будущего. Ведущиеся в настоящее время исследования позволят реализовать и оценить сетевые механизмы для разных типов сервиса, которые могут использоваться в сетях будущего, предлагающих такой сервис.

Одной из фундаментальных проблем предложенной архитектуры является решение вопроса о поддержке одного или множества протоколов. Архитектура не содержит каких-либо предположений о протоколах сетевого уровня для поддержки протокола междоменной маршрутизации в каждом домене или едином протоколе междоменной маршрутизации, который будет работать с множеством протоколов сетевого уровня¹. Т. е., предлагаемая архитектура может быть реализована на основе одного протокола междоменной маршрутизации, поддерживающего множество протоколов сетевого уровня или на основе множества протоколов междоменной маршрутизации (с одинаковой архитектурой), предназначенных для конкретных протоколов сетевого уровня².

5.2 Переходный период

Предложенная архитектура не предназначена для полномасштабного развертывания в короткие сроки. Мы предполагаем, что эта архитектура обозначает цель, на достижение которой следует направлять исследования и инвестиции в ближайшие 5 лет.

В то же время, архитектура не требует полномасштабного развертывания в существующей сети Internet. Компонента NR может развертываться постепенно. Например, NR для протокола IP можно развернуть путем замены существующей маршрутизации EGP-2 маршрутизацией на основе BGP. После установки компоненты NR она может быть дополнена элементами, обеспечиваемыми компонентой SDR.

Наиболее критичные компоненты архитектуры, требуемые для поддержки SDR, включают организацию маршрутов и пересылку пакетов в маршрутизаторах, которые поддерживают SDR. Участие в маршрутизации в качестве транзитного домена потребует распространения локальных конфигурационных параметров (LCI³), а от некоторых маршрутизаторов домена будет требоваться также реализация протоколов организации маршрутов и их поддержки. Участие в качестве

¹ Подобная проблема уже существует для протоколов внутridoменной маршрутизации и порождает достаточно серьезную полемику в сообществе Internet. По нашему мнению, проблема выбора между использованием одного или множества протоколов для междоменной маршрутизации является более сложной, нежели для внутridoменной.

² Если будет принято решение об использовании одного протокола, очевидно, что IDPR будет использоваться в качестве базового протокола для компоненты NR. Поскольку в настоящее время IDPR развивается в направлении CLNP, требуется дополнительная работа по добавлению поддержки протоколов IP и ST-II. При выборе поддержки множества протоколов очевидно, что протокол BGP будет использоваться в качестве базового для компоненты NR в сетях IP, а IDPR будет использоваться в качестве базового для компоненты NR в среде CLNP. Потребуется дальнейшая работа по поддержке компоненты NR для протокола ST-II. Может также потребоваться дополнительная работа по спецификации новых функций, которые могут быть добавлены в BGP.

³ local configuration information — локальная конфигурационная информация.

исходного домена потребует от домена доступа к серверу маршрутов RS для расчета маршрутов, а также в исходном домене потребуются наличие маршрутизатора, реализующего протоколы организации и поддержки маршрутов. В дополнение к этому элемент управления сетью должен будет описывать локальные конфигурационные параметры и передавать их в центральные репозитории. Потребуется также механизм сбора и распространения конфигурационной информации даже при ее централизованном хранении.

6.0 Заключение и перспективы продолжения работы

Предложенная здесь архитектура объединяет протоколы поэтапной маршрутизации на основе векторов пути (hop-by-hop path-vector) и задаваемой отправителем маршрутизации по состоянию каналов (source-routed link-state), используя каждый из этих протоколов в тех случаях, когда протокол подходит лучше всего. Компонента NR использует алгоритм PV и многоуровневые гибкие конфедерации доменов для поддержки эффективной маршрутизации базовых пакетов с использованием обычных маршрутов. Компонента SDR использует алгоритм LS и рассчитывает маршруты на основе базы данных о конфигурационных и динамических параметрах для передачи специфического трафика через специальные маршруты. В прошлом эти два подхода к маршрутизации рассматривались как взаимоисключающие, хотя они могут хорошо дополнять друг друга. В настоящее время оба подхода к маршрутизации уже используются в параллель. Два механизма маршрутизации при их совместном использовании обеспечивают гибкую и эффективную поддержку маршрутизации с использованием TOS и правил в масштабах глобальной сети Internet.

Настало время рассмотреть вопросы, связанные объединением и интеграцией двух компонент. Мы должны вернуться назад и пристально рассмотреть обе архитектуры и их связанные с ними протоколы, избегая повторений и обеспечивая бесшовную интеграцию.

7.0 Благодарности

Авторы выражают благодарность Hans-Werner Braun (San Diego Supercomputer Center), Lee Breslau (USC), Scott Brim (Cornell University), Tony Li (Cisco Systems), Doug Montgomery (NIST), Tassos Nakassis (NIST), Martha Steenstrup (BBN) и Daniel Zappala (USC) за их комментарии к предварительным вариантам документа.

8.0 Литература

- [ANSI 87-150R] "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol", ANSI X3S3.3/87-150R.
- [BGP 91] Lougheed, K., and Y. Rekhter, "A Border Gateway Protocol 3 (BGP-3)", RFC 1267, Cisco Systems, T.J. Watson Research Center, IBM Corp., October 1991.
- [Breslau-Estrin 91] Breslau, L., and D. Estrin, "Design and Evaluation of Inter-Domain Policy Routing Protocols", To appear in Journal of Internetworking Research and Experience, 1991. (Earlier version appeared in ACM Sigcomm 1990.)
- [Clark 90] Clark, D., "Policy Routing in Internetworks", Journal of Internetworking Research and Experience, Vol. 1, pp. 35-52, 1990.
- [Dijkstra 59] Dijkstra, E., "A Note on Two Problems in Connection with Graphs", Numer. Math., Vol. 1, 1959, pp. 269-271.
- [ECMA89] "Inter-Domain Intermediate Systems Routing", Draft Technical Report ECMA TR/ISR, ECMA/TC32-TG 10/89/56, May 1989.
- [EGP] Rosen, E., "Exterior Gateway Protocol (EGP)", RFC 827, BBN, October 1982.
- [Estrin 89] Estrin, D., "Policy Requirements for Inter Administrative Domain Routing", RFC 1125, USC Computer Science Department, November 1989.
- [Estrin-etal91] Estrin, D., Breslau, L., and L. Zhang, "Protocol Mechanisms for Adaptive Routing in Global Multimedia Internets", University of Southern California, Computer Science Department Technical Report, CS-SYS-91-04, November 1991.
- [Hedrick 88] Hedrick, C., "Routing Information Protocol", RFC 1058, Rutgers University, June 1988.
- [Honig 90] Honig, J., Katz, D., Mathis, M., Rekhter, Y., and J. Yu, "Application of the Border Gateway Protocol in the Internet", RFC 1164¹, Cornell Univ. Theory Center, Merit/NSFNET, Pittsburgh Supercomputing Center, T.J. Watson Research Center, IBM Corp., June 1990.
- [IDPR90] Steenstrup, M., "Inter-Domain Policy Routing Protocol Specification and Usage: Version 1", Work in Progress², February 1991.
- [IDRP91] "Intermediate System to Intermediate System Inter-domain Routing Exchange Protocol", ISO/IEC/JTC1/SC6 CD10747.
- [ISIS10589] "Information Processing Systems - Telecommunications and Information Exchange between Systems - Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol for use in Conjunction with the protocol for providing the Connectionless-mode Network Service (ISO 8473)", ISO/IEC 10589.
- [ISO-473 88] "Protocol for providing the connectionless-mode network service", ISO 8473, 1988.
- [Jaffee 82] Jaffee, J., and F. Moss, "A Responsive Distributed Routing Algorithm for Computer Networks", IEEE Transactions on Communications, July 1982.
- [Little 89] Little, M., "Goals and Functional Requirements for Inter-Autonomous System Routing", RFC 1126, SAIC, October 1989.

¹Этот документ признан устаревшим и заменен RFC 1268. *Прим. перев.*

²Работа опубликована в RFC 1479. *Прим. перев.*

- [Oran 89] Oran, D., "Expert's Paper: The Relationship between Addressing and Routing", ISO/JTC1/SC6/WG2, 1989.
- [OSPF] Moy, J., "The Open Shortest Path First (OSPF) Specification", RFC 1131, Proteon, October 1989.
- [Postel 81] Postel, J., "Internet Protocol", RFC 791¹, DARPA, September 1981.
- [Rekhter 91] Rekhter, Y., "IDRP protocol analysis: storage complexity", IBM Research Report RC17298(#76515), October 1991.
- [Shin87] Shin, K., and M. Chen, "Performance Analysis of Distributed Routing Strategies Free of Ping-Pong-Type Looping", IEEE Transactions on Computers, February 1987.
- [ST2-90] Topolcic, C., "Experimental Internet Stream Protocol, version 2 (ST II)", RFC 1190², CIP Working Group, October 1990.
- [Zaumen 91] Zaumen, W., and J. Garcia-Luna-Aceves, "Dynamics of Link State and Loop-free Distance-Vector Routing Algorithms", ACM Sigcomm '91, Zurich, Switzerland, September 1991.
- [Zhang 91] Zhang, L., "Virtual Clock: A New Traffic Control Algorithm for Packet Switching Networks".

Вопросы безопасности

Вопросы безопасности в этом документе не обсуждаются.

Адреса авторов

Deborah Estrin

University of Southern California
Computer Science Department, MC 0782
Los Angeles, California 90089-0782
Phone: (310) 740-4524
EMail: estrin@usc.edu

Yakov Rekhter

IBM T.J. Watson Research Center
P.O. Box 218
Yorktown Heights, New York 10598
Phone: (914) 945-3896
EMail: yakov@ibm.com

Steven Hotz

University of Southern California
Computer Science Department, MC 0782
Los Angeles, California 90089-0782
Phone: (310) 822-1511
EMail: hotz@usc.edu

Перевод на русский язык

Николай Малых

nmalykh@protocols.ru

¹Перевод этого документа на русский язык имеется на сайте www.protocols.ru. Прим. перев.

²Этот документ признан устаревшим и заменен RFC 1819. Прим. перев.