

# Новшества в Windows Server 2008: стек TCP/IP

Андрей Бирюков

**В новой версии ОС Windows Server 2008 стек TCP/IP был полностью переработан. Рассмотрим, какие новшества появились в нем.**

**W**indows Server 2008 содержит множество различных нововведений. Такие давно используемые и хорошо всем знакомые вещи, как реализация стека TCP/IP в новой операционной системе, также претерпела весьма существенные изменения.

## Основные изменения и структура

Приведу список основных изменений в новом сетевом стеке. Прежде всего это одновременная поддержка как IPv4, так и IPv6. Шестую версию IP сейчас поддерживают практически все выпускаемые операционные системы, так что поддержка в Windows 2008 была вполне ожидаема. Новая технология поддержки IPv4 и IPv6 называется Dual-IP.

Другим нововведением является поддержка так называемой Strong Host Model (усиленной модели хоста). Что это такое, лучше объяснить на при-

мере. Когда на какой-либо узел приходит обычный unicast-пакет, сетевой интерфейс принимает его, если IP-адрес назначения совпадает с одним из адресов, который присвоен сетевым интерфейсам на данной машине. Такая обработка пакетов называется слабой моделью хоста (Weak Host Model). В случае усиленной модели будет обработан пакет только с IP-адресом, соответствующим адресу интерфейса, на который он пришел. Использование слабой модели может сделать не только конкретный узел, но и всю локальную сеть менее устойчивой к несанкционированному проникновению, так как в случае, если на внешний интерфейс узла, выполняющего роль маршрутизатора, послать пакет с адресом из внутренней, то при неаккуратной конфигурации такой пакет вполне может достичь точки назначения. С использованием усиленной модели даже в случае неаккуратной настройки правил маршрутизации такого не случит-

ся. В текущих реализациях TCP/IP на Windows Server 2003 и Windows XP используется слабая модель. Более подробно с моделями узлов можно ознакомиться в статье [1].

Next Generation TCP/IP стек снижает нагрузку на драйвер Network Driver Interface Specification (NDIS) и сетевые адаптеры при обработке TCP и других видов трафика. Данный функционал весьма полезен при работе в сетях с высокой пропускной способностью. Также в новой версии стека появилась возможность обработки трафика на многопроцессорных машинах с распределением нагрузки. Для высокопроизводительных сетевых приложений, прежде всего различных видеотрансляций, многопроцессорная обработка также будет полезной.

Еще одним важным нововведением в стеке Next Generation является возможность автоматической настройки определенных компонентов стека (например, масштабируемость TCP

receive window и других) без ручного внесения изменений в конфигурацию. Тут следует отметить, что в UNIX-системах подобные функции существуют уже давно и существенно упрощают жизнь системным администраторам. Так что реализация данного функционала в Next Generation Stack является несомненным шагом вперед.

Схематически новую архитектуру можно изобразить с помощью **рисунка**. Поясню работу стека последовательно, снизу вверх. За отправку и получение пакетов на интерфейс отвечает драйвер NDIS. Затем данные передаются драйверу tcpip.sys.

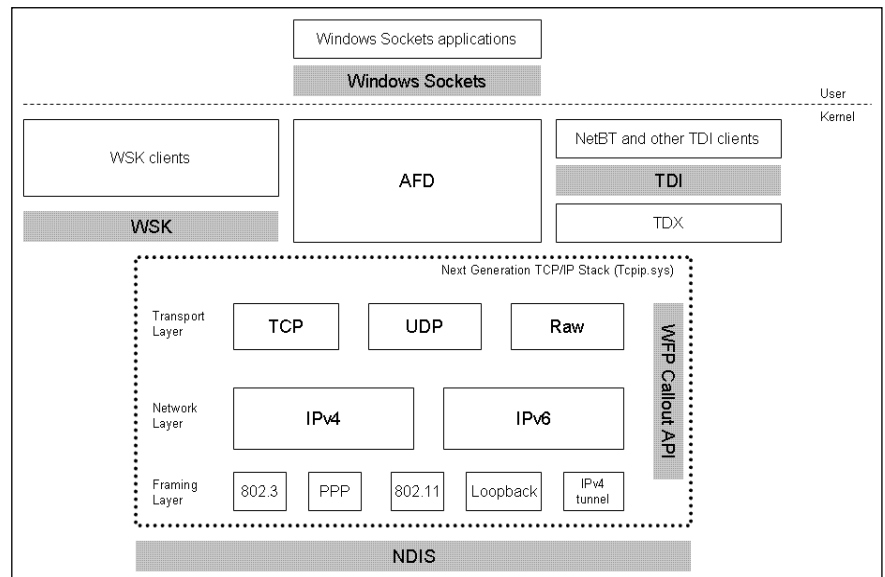
Архитектура драйвера новой версии стека TCP/IP, реализованного в файле tcpip.sys, состоит из следующих уровней:

- **Транспортный уровень** – содержит реализации протоколов TCP и UDP, а также механизм для отсылки базовых IP-пакетов, которым не требуется наличие TCP или UDP-заголовков.
- **Сетевой уровень** – содержит реализации протоколов IPv4 и IPv6 в виде уровня Dual IP layer (о данной технологии я уже писал выше).
- **Фреймовый уровень** – содержит модули для фрейминга пакетов IPv4 и IPv6. Существуют модули для интерфейсов IEEE 802.3 (Ethernet), IEEE 802.11 и Point-to-Point Protocol (PPP). Помимо этого существуют модули и для логических интерфейсов, таких как Loopback Interface, а также интерфейсов туннелирования на основе IPv4, которые часто используются в технологиях поддержки переключения с IPv4 на IPv6.

После такой обработки полученные данные передаются Winsock Kernel и другим обработчикам уровня ядра. И, наконец, в верхней части архитектуры находятся приложения Windows Socket, которые извлекают непосредственные данные и передают их приложениям.

## Новые API

Также, рассказывая о новых функциях сетевого стека, нельзя не упомянуть о новых возможностях для программистов. Прежде всего это появление новых API для безопасности



Архитектура Next Generation TCP/IP

и фильтрации пакетов. Новый набор компонентов для программирования получил название Windows Filtering Platform (WFP). WFP обеспечивает возможность фильтрации на всех уровнях стека протоколов TCP/IP. Также следует отметить, что WFP более тесно интегрирован в реализацию TCP/IP для Windows Server 2008. Это позволяет сторонним разработчикам быстрее и легче разрабатывать сетевые драйверы, службы и приложения для работы с TCP/IP трафиком. Более подробно с Windows Filtering Platform можно ознакомиться в статье [2].

В продолжении темы API хотелось бы упомянуть три основных программных интерфейса, используемых приложениями, сервисами и другими системными компонентами для доступа к новому стеку TCP/IP – это WSK (Winsock Kernel), используемый клиентскими приложениями, Windows Sockets, используемый приложениями и сервисами (программный интерфейс Windows Sockets использует драйвер AFD – Ancillary Function Driver для выполнения функций на уровне сокетов через TCP/IP) и TDI (Transport Driver Interface), используемый NetBIOS over TCP/IP (NetBT) и другими устаревшими клиентами. Для трансляции запросов между TDI и новым стеком TCP/IP используется интерфейс TDX. По замыслу разработчиков WSK должен со временем заменить TDI.

Новый стек TCP/IP поддерживает программный интерфейс Callout API, который является унифицированным

способом для встраивания в стек и модификации данных на уровне пакетов. Программный интерфейс является частью WFP, обеспечивающей доступ к обработке пакетов на сетевом и транспортном уровнях. Новый стек также поддерживает отсылку и получение фреймов (frames), используя NDIS.

Завершая детальное описание нововведений, хочу привести краткий список функций, еще не упомянутых, однако заявленных разработчиком. Это улучшенная поддержка производительности и коррекция ошибок, поддержка аппаратных конфигураций и автонастройки, богатые функции расширяемости на уровне нового программного интерфейса, улучшенная поддержка рабочих станций мобильных пользователей, часто переключающихся из одной сети в другую, взаимодействие с Windows CE, Xbox и Windows Embedded, устойчивость к известным DoS и другим сетевым атакам.

## Новые решения старых проблем

Теперь сравним архитектуру Next Generation TCP/IP Stack с предыдущими реализациями TCP/IP. Прежде всего вспомним некоторые особенности работы стека в предыдущей версии операционной системы и их недостатки. В пакетах обновлений (service packs) для Windows Server 2003 были включены дополнительные возможности: в TCP появилось уже упо-



минавшееся сегодня масштабирование окон (window scaling), определение нерабочих шлюзов (dead gateway detection) и другие тонкие настройки, которые были выполнены лучше и более надежно, по сравнению со стеком в Windows 2000. Справедливости ради следует также отметить, что реализация стека TCP/IP была практически полностью скопирована с ОС FreeBSD.

Однако в стеке не было поддержки IPv6. Поддержка стека была реализована простым добавлением второго драйвера транспортного уровня Tcpip6.sys для существующего драйвера Tcpip.sys для IPv4. Результатом этого явился двойной стек, в котором транспорт для IPv4 и IPv6 поддерживался двумя различными драйверами, и для включения IPv6 администраторы должны были установить компонент протокола IPv6 на своих машинах с Windows XP и Windows Server 2003. Очевидно, что такой подход имеет несколько недостатков. Например, функциональность разделенного фильтра пакета для каждого протокола IP приводит к тому, что сложно настраивать брандмауэр на машинах, на которых работают оба протокола. Также это усложняет жизнь разработчикам сетевых приложений, которые необходимы для поддержки как IPv4, так и IPv6. А это означает дублирование кода и менее эффективный канал для выполнения стека.

Еще одним недостатком прежней реализации стека является механизм определения нерабочих шлюзов (dead gateway detection), который был введен как способ поддержания работоспособности, когда основной шлюз (gateway) или маршрутизатор (router) вышел из строя. В больших компаниях часто есть резервный маршрутизатор, который может работать в случае выхода из строя основного маршрутизатора, а определение нерабочих шлюзов позволяет использовать такие резервные шлюзы. Проблема реализации данного механизма в Windows Server 2003 заключалась в том, что не было возможности отката, другими словами, когда основной шлюз восстанавливал свою работоспособность, система продолжала отправлять пакеты на резервный шлюз до следующей перезагрузки. Так как

зачастую резервный маршрутизатор обладает меньшей производительностью, а многие серверы не перезагружаются месяцами, отсутствие возможности возвращения на основной шлюз несколько неприятно. Также данный механизм в Windows Server 2003 может определять аварии только на локальном шлюзе по умолчанию, аварии на удаленных шлюзах определить нельзя.

Также в стеке TCP/IP в Windows Server 2003 иногда происходит медленная обработка пакетов в каналах с высокой пропускной способностью. Это вызвано тем, что стек настраивает TCP с помощью медленного старта (slow start) и устранения перегрузок (congestion avoidance) – два стандартных алгоритма TCP/IP, которые были спроектированы еще до того, как гигабайтные сети стали нормой. Например, при использовании DFS (Distributed File System) вы можете иметь сервер вместо концентратора, который может копировать большое количество данных, проходящих по высокоскоростной WAN через множество маршрутизаторов. Результатом является то, что WAN-соединение перегружается, и копирование занимает больше времени, чем должно, иногда намного больше.

Завершая перечисление недостатков старого стека, хотелось бы упомянуть о проблемах с безопасностью. Дело в том, что одна таблица маршрутов (routing table) используется для разделения трафика для каждой из сетей, к которым подключен сервер, в том числе и VPN-соединения, что в свою очередь означает, что если у пользователя есть права локального администратора на машине, то он имеет возможность для изменения таблицы маршрутов таким образом, чтобы создать угрозу безопасности локальной сети, разрешив доступ из Интернета.

Теперь рассмотрим, каким образом приведенные выше недостатки были устранены в новой версии стека.


Проблема взаимодействия IPv4 и IPv6 решена с помощью уже упоминавшейся выше технологии Dual IP layer, которая существенно упрощает взаимодействие этих двух протоколов.

В Windows Server 2008 полностью решена проблема гарантированной

изоляции трафика VPN-соединением сессий входа, таблиц маршрута и сетевых интерфейсов в логическую конструкцию под названием подсистема выбора маршрута (routing compartment). Такое деление означает, что, например, соединение от ноутбука к Интернету и корпоративной сети с помощью VPN полностью изолированы одно от другого, что существенно увеличивает защищенность корпоративной сети.

Для решения проблемы отката в механизме определения нерабочих шлюзов (dead gateway detection) используется похожий набор методов, которые IPv6 использует для определения недоступных соседей в сетях IPv4, что в основном включает мониторинг сессий TCP и обмен пакетами ARP. Когда система обнаруживает такую ситуацию, стек переключается на следующий шлюз по умолчанию из списка, но стек продолжает мониторинг предыдущего шлюза, и если тот возобновляет работу, то система переключается обратно на этот основной шлюз. Это позволяет обеспечить оптимальную связь с удаленными офисами, гарантируя, что резервные шлюзы будут использоваться, только пока не работает основной шлюз.

В новом стеке проблема низкой производительности в сетях с высокой пропускной способностью решена с помощью нового алгоритма вместо алгоритмов, использовавшихся на других платформах. Этот новый подход называется Compound TCP (CTCP), и он может значительно повысить производительность при использовании в различных сетях.

Надеюсь, что информация о новшествах в стеке протокола TCP/IP будет вам полезна при работе с Windows Server 2008. 

1. <http://technet.microsoft.com/en-us/magazine/cc137807.aspx> – описание Weak And Strong Host Models.
2. <http://www.microsoft.com/whdc/device/network/WFP.msp> – описание Windows Filtering Platform.
3. [http://technet.microsoft.com/ru-ru/library/bb878108\(en-us\).aspx](http://technet.microsoft.com/ru-ru/library/bb878108(en-us).aspx) – статья на TechNet, посвященная Next Generation TCP/IP.
4. <http://www.netdocs.ru/articles/TCP-IP-Networking-Windows-Vista.html> – сеть TCP/IP в Windows Vista.