

# Когда WAN быстрее LAN, или Способы консолидации WAN-офисов

Данная публикация — обзор представленных на российском рынке решений для оптимизации трафика WAN-сетей (Wide Area Networks), дающих существенные преимущества по централизации и упрощению управления распределенными офисами как для крупных, так и для небольших компаний.

## Введение

Решения, связанные с задачами оптимизации управления распределенными IT-инфраструктурами, стали активно развиваться в последние 1,5 года. Это вызвано 2 причинами. Во-первых, вследствие того, что у многих крупных компаний IT-инфраструктура значительно усложнилась и превратилась в географически распределенную глобальную структуру. Этот процесс в силу объективных причин затрагивает все большее число компаний. Такая IT-инфраструктура постепенно становится плохо управляемой и плохо синхронизируемой с центральным офисом, требующей при этом значительных затрат на ее поддержание. Заметим, что уже сложившееся мнение о решениях по оптимизации трафика WAN как о продуктах, в основном ориентированных на средний и малый бизнес, неправильно, т.к. их внедрение на нескольких десятках и сотнях удаленных офисах крупной компании превращает их в крупный проект, в ряде случаев существенно меняющий бизнес-процессы.

Во-вторых, требования современного бизнеса заставляют даже мелкие компании иметь более одного офиса, но при наличии необходимости централизации своей деятельности они не в состоянии делать значительные инвестиции в IT-инфраструктуру. Внедрение в подобных компаниях решений по оптимизации WAN-тра-

фика позволяет решить и эту проблему. При этом, чем выше уровень централизации управления, тем меньше затрат идет на установку и поддержание функционирования этой инфраструктуры (например, за счет отказа поддержания в удаленных офисах функционала по резервному копированию/восстановлению данных, почтовых серверов и др.). Если развивать эту аналогию далее, то по сути, удаленные офисы превращаются в тонких клиентов вместо рабочих станций в клиент-серверной архитектуре.

IT-инфраструктура в среде WAN, в отличие от LAN, требует повышенного внимания с точки зрения оптимизации трафика потому, что имеет гораздо большие ограничения по пропускной способности. Если LAN обычно имеет полосу от 100 Mbps до

1 Gbps и больше, то WAN — от 64 kbps до T1 (1,5 Mbps) и много реже — выше. Также WAN имеет значительно большую величину задержки, чем LAN — от 100 до 1000 раз (из-за гораздо больших физических расстояний), определяемую физической скоростью распространения электромагнитных волн.

Как может скорость света влиять на файловый обмен? При наличии удаленного офиса в Санкт-Петербурге и необходимости доступа от него к файловой системе в Москве, задержка на путешествие света туда и обратно (1300 км) — порядка 10 мс. При этом код практически всех систем (в целях ускорения разработки, снижения стоимости, увеличения совместимости, обеспечения мобильности и т.д.) написан с использованием высокоуровневых библиотек, которые выполняют любую простейшую операцию (чтение байта из файла) максимально затратным способом с проверкой всех возможных ситуаций и генерацией до сотни обращений к сетевому протоколу (CIFS в сетях Microsoft). В результате получаем секунды на одну операцию, и как итог — минуты на открытие файла редактором Word. А есть еще большие задержки, например Москва—Владивосток—Москва: 90 мс. Реальное снижение пропускной способности канала связи в зависимости от задержки представлено на рис. 1<sup>\*)</sup>.

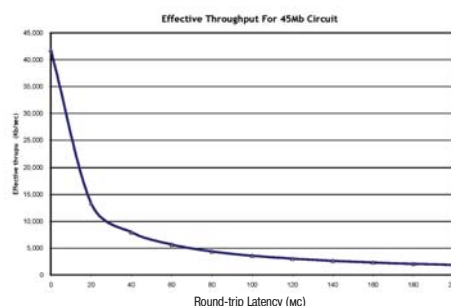


Рис. 1. Изменение пропускной способности канала T3 (45 Мбит/с — максимум) в зависимости от задержки. При задержке 100 мс реальная полоса снижается в 10 раз.

\*) "It is not about bandwidth white paper", HP (5983-2465EN), 13.05.2005.

Вследствие всего вышесказанного, на рынке стали появляться как чисто программные решения, так и аппаратно-программные, оптимизирующие работу существующих решений в WAN-средах. Последние из них выделяются в отдельную группу решений, имеющих аббревиатуру WAFS (Wide Area Files Services – глобально распределенные файловые сервисы или WAN-акселераторы).

Как показали недавно проведенные исследования компанией Gartner, рынок решений<sup>\*\*\*</sup>, связанных с оптимизацией трафика WAN-сетей – один из самых быстроразвивающихся с ежегодным ростом 40% (об одном из таких решений читайте в SN № 1/22, 2005 – прим. ред.). Теоретически и практически решения для централизации распределенной IT-инфраструктуры давно существуют. Они связаны с увеличением полосы пропускания каналов связи, расширением функциональности удаленных серверов и т.п., но все это – прежде всего финансовые затраты. Рассматриваемые классы как программных, так и WAFS-решений, отличаются от них прежде всего оптимизацией подхода, т.е. при минимуме вложений достигаются максимальное расширение функциональности и, соответственно, более широкая доступность для потребителей. Ценовой диапазон таких решений (на один удаленный офис) лежит в пределах \$1–15 тыс.

Рассмотрим три программно-аппаратных решения, анонсированных в течение 2005 г. по настоящий момент, и одно дополнение к программному продукту NetBackup, т.н. PureDisk.

## Программно-аппаратные решения оптимизации WAN-трафика

Из данной группы решений на российском рынке представлены 3 вендора с соответствующими предложениями (в последовательности выхода на рынок):

- Cisco – на базе т.н. файлового ускорителя – File Engine, предлагаемого в нескольких вариантах исполнения: FE-511, FE-611 и FE-7326 (SN № 1/22, 2005). Решение WAFS от Cisco может поставляться как в комплекте с хранилищем NAS, так и без него (предполагается интеграция с имеющимся). Одними из первых это решение поддержала компания EMC. В соответствии с соглашением, достигнутым между Cisco и EMC, Cisco получила возможность “оемить” EMC NS500 – систему на базе упрощенной программы сетевого хранения и управления данными, которая объединила платформы Celerra и EMC CLARiON®, что обеспечило высокое качество, простоту в использовании и доступность;
- HP – на базе HP StorageWorks Enterprise File Services WAN Accelerator. Первая версия была анонсирована в мае прошлого года, в феврале с.г. уже объявлен релиз 2.1;
- Brocade – Tapestry WAFS (одно из трех решений в рамках технологического направления Tapestry) – анонсировано в сентябре 2005 г. и в на-

стоящее время, помимо партнеров Brocade, продвигается такими вендорами, как Fujitsu Siemens Computers (решение FibreCAT® N20-WAFS) и Nortel Networks.

Кроме упомянутых, есть еще несколько разработок, которые в России практически не представлены и не продвигаются.

Общая цель решений по оптимизации WAN-трафика – максимальное упрощение управлением удаленными офисами как в отношении доступности данных к/от удаленным/ых офисов, так и снижения затрат на персонал и оборудование. Например, за счет внедрения WAFS-решений можно полностью отказаться от развертывания в удаленных офисах файловых серверов, почтовых серверов, ленточных систем для резервного копирования/восстановления (превращая их в полностью бессерверные центры), используя для этого ресурсы центрального офиса, не ухудшая при этом показателей производительности приложений и повышая уровень функциональности. Помимо этого, управление удаленными центрами/офисами полностью централизуется и они превращаются, с точки зрения администрирования и восприятия, из удаленных в локальные центры.

С архитектурной точки зрения, все решения подобны, но каждое имеет ряд особенностей, определяющих его позиционирование. В частности, это поддерживаемые типы оптимизации WAN-трафика и протоколы, которые оптимизируются; поддерживаемая среда и др.

### HP StorageWorks Enterprise File Services (EFS) WAN Accelerator 2.1

Решение EFS WAN Accelerator строится на базе 3-х типов HP ProLiant-серверов и использует различные технологии: автообнаружение, компрессию, оптимизацию задержки, алгоритмы предсказания загрузки WAN. Семейство серверов решения EFS WAN Accelerators включает EFS WAN Accelerator Manager (DL320-M25), который может использоваться для централизованного управления и контроля пула до 500 удаленных серверов на базе DL320-2010. Помимо этого, в центральном офисе для работы с удаленными офисами требуется установка более мощного сервера – DL380-5010 (рис. 2).

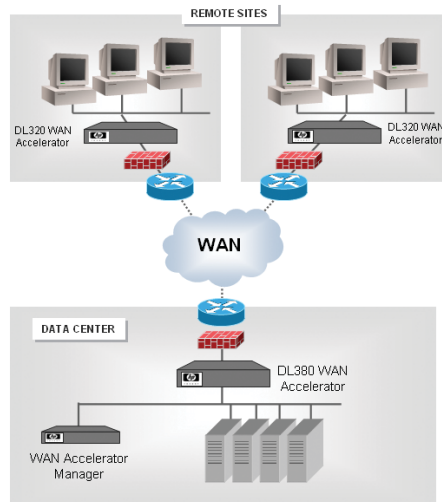


Рис. 2. Общая схема построения решения HP StorageWorks EFS WAN Accelerator.

Основная направленность EFS WAN Accelerators – снижение трафика WAN для всех приложений, использующих TCP (включая CIFS, NFS, HTTP, FTP, MAPI и MS-SQL), а также консолидация трех типов WAN-оптимизации (устранение избыточности данных, оптимизация задержки и предсказание профиля транзакций) в одном продукте.

Данное решение полностью прозрачно для серверов и клиентов, не требует никакой отдельной настройки на них и обеспечивает, прежде всего, решение следующих основных задач в WAN-среде: повышение производительности приложений и консолидацию серверов и ресурсов хранения.

Это достигается за счет:

- удаления всего избыточного трафика WAN – файлы, почтовые сообщения, резервные копии, web-трафик, FTP и др. Технология имеет отдельное название – Scalable Data Referencing – SDR и работает на всех TCP-приложениях, включая Microsoft Office, Lotus Notes, CIFS, NFS, CAD, ERP, FTP, HTTP и др. При этом данные никогда не посылаются дважды за счет того, что создается “библиотека” трафика в виде небольших меток для данных, которые используются для повторения пересылки только отдельных частей данных. Это дает возможность снизить трафик в среднем до 95%, или 20-кратное увеличение полосы пропускания;
- оптимизации “поведения” TCP за счет устранения эффектов ожидания при передаче пакета (механизм реализуется на базе т.н. технологий Virtual Window Expansion (VWE) и Window Scaling);
- оптимизации “поведения” протокола приложений (например, устранение большинства запросов приложений, использующих Windows CIFS или MAPI-email). Дополнительно EFS WAN Accelerator может быть доступен через MS-SQL Acceleration Module, что позволяет в значительной степени его оптимизировать для Microsoft Project 2003 (запросы и ответы между клиентами и серверами). Также возможна оптимизация и для других СУБД.

В целом все эти усилия могут дать более чем 100-кратное увеличение производительности WAN. Возможные коэффициенты ускорений при использовании EFS WAN Accelerator для разных применений приведены в табл. 1.

Табл. 1. Коэффициенты ускорения для различных применений

| Типы приложений       | Ускорение                      |
|-----------------------|--------------------------------|
| File Sharing          | Windows (CIFS) 4 – 170x        |
|                       | Unix (NFS) 2 – 5x              |
| Mail                  | Exchange (MAPI) 3 – 50x        |
|                       | Notes 2 – 20x                  |
| Web                   | HTTP, Intranet, ERP 3 – 50x    |
|                       | FTP 3 – 120x                   |
| Backup & Replication  | SnapMirror 3 – 50x             |
|                       | Veritas 2 – 25x                |
|                       | MS Backup 2 – 25x              |
| Software Distribution | SMS 3 – 60x                    |
| Database              | MS-SQL 1 – 20x                 |
|                       | Oracle SQL 1 – 3x              |
| ERP                   | Fat Client, SAP, etc... 1 – 3x |

Большое значение при внедрении WAFS и EFS WAN Accelerator, в частности,

имеет масштабируемость решения. В табл. 2. даны возможные диапазоны IT-инфраструктуры в зависимости от типа сервера. При этом максимально возможное число WAN Accelerator appliances, которым может управлять HP EFS WAN Accelerator Manager, составляет 500.

Табл. 2. Показатели масштабируемости EFS WAN Accelerator в зависимости от типа используемого сервера

| Сервер/показатель                  | DL320-510 | DL320-2010 | DL380-5010 |
|------------------------------------|-----------|------------|------------|
| WAN capacity                       | 512 kbps  | 4 Mbps     | 45Mbps     |
| Raw disk store                     | 250 GB    | 500 GB     | 1.8 TB     |
| Optimized TCP connections          | 200       | 1300       | 4000-5000  |
| User capacity (consolidated)*      | до 25     | до 200     | до 1600    |
| User capacity (non consolidated)** | до 150    | до 1100    | до 6000    |
| Fan-out***                         | —         | до 20:1    | до 150:1   |
| RAID                               | No        | Yes        | Yes        |
| Hot-swappable disks                | No        | No         | Yes        |
| GigE fail-through                  | Yes       | Yes        | Yes        |
| Auto-discovery                     | Yes       | Yes        | Yes        |

### Решение Brocade Tapestry WAFS

Решение Brocade Tapestry WAFS было анонсировано в сентябре 2005 г. Основная его направленность – дополнение других решений из пакета Microsoft Office с целью усиления их возможностей в части повышения производительности при работе в среде WAN.

Tapestry WAFS разработано для использования преимуществ технологий, заложенных в платформу Microsoft Windows Storage Server 2003 R2, и, в частности, возможности по оптимизации передачи данных через WAN. Данная функциональность развивается Microsoft в рамках т.н. концепции Universal Distributed Storage Vision и на текущий момент доступна через решение System Center Data Protection Manager (анонсированное почти вместе с Tapestry WAFS в сентябре 2005 г.) на базе новой файловой системы – Distributed File System Replication (DFS-R), включенной в Windows Storage Server 2003 R2.

Tapestry WAFS работает на Windows Storage Server 2003 ОС (в значительной степени урезанная специализированная и более надежная версия ОС в сравнении со стандартной Windows) и применяет WAN-оптимизированный протокол, известный еще как “Storage Caching over IP (SC/IP)”, который использует широкий диапазон технологий для существенного улучшения работы приложений Microsoft Office в распределенной среде. Данное решение состоит всего из двух устройств, одно из которых устанавливается в центральном офисе (Tapestry WAFS Core Appliance), другое (Tapestry WAFS Edge Appliance) – в удаленном (рис. 3). В отличие от предыдущего решения, здесь отсутствует управляющий сервер.

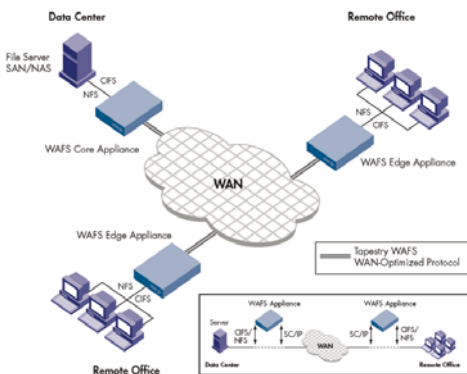


Рис. 3. Структурная схема решения Tapestry WAFS.

Табл. 3. Сервисы Tapestry WAFS значительно снижают сложность управления и стоимость удаленных офисов

| Stackable Services             | Бизнес-преимущества   |
|--------------------------------|---|
| <b>Exchange services</b>       | Оптимизация трафика Microsoft Exchange по глобальной сети (WAN), сокращение потребности в удаленных серверах Exchange   |
| <b>Management services</b>     | Уменьшение сложности за счет поддержки технологии Systems Management Server (SMS) распространения обновлений ПО, исправлений и приложений к удаленным офисам  |
| <b>Web caching services</b>    | Ускорение производительности приложения web-приложений при сохранении сетевой полосы пропускания  |
| <b>Network services</b>        | Обеспечение DNS/DHCP-сервисов, а также сервисов управления доменом без необходимости инсталляции локальных серверов   |
| <b>Print services</b>          | Доступность совместного использования ресурсов печати для всех сетевых компьютеров в филиалах   |
| <b>Wide area file services</b> | Максимальное использование удаленных файловых серверов, при уменьшении требований к сетевой полосе, повышении надежности, централизации управления, увеличения производительности, повышении целостности данных и уменьшения затрат на хранение |

Tapestry WAFS – одно из немногих решений, полностью сертифицированное Microsoft, дающее возможность консолидации различных сервисов управления удаленных офисов (табл. 3).

Минимизация WAN-трафика достигается в Tapestry WAFS за счет следующих особенностей:

- передачи только изменений файла. При совместном использовании файла и проведении изменений в нем в удаленном офисе в центральный офис передаются (с компрессией и в потоковом режиме) только его изменения;
- потоковой передачи файлов с компрессией. При передаче Microsoft документов в WAN-среде они обычно “нарезаются” на страницы. Это вызывает дополнительные задержки и лишние пересылки данных, когда требуется следующая страница. Чтобы избежать этого, в Tapestry WAFS используется многостраничная потоковая передача с компрессией;
- оптимизации временных файлов. Tapestry WAFS “знает” специфику временных файлов и минимизирует лишние их пересылки в WAN;
- агрегирования данных. Tapestry WAFS при передаче группирует данные от разных приложений в один пакет, что позволяет минимизировать задержки и лишние пересылки служебной информации.

Табл. 4. Компания из списка Fortune 500. Передача Microsoft-файлов (в сек) между Германией и США (bandwidth: 413 kbps; latency: 140 ms)

| File Type and Size | WAN-Rd без TWAFS | WAN-Rd с TWAFS “cold” | WAN-Rd с TWAFS “warm” | WAN-Wr без TWAFS | WAN-Wr с TWAFS |
|--------------------|------------------|-----------------------|-----------------------|------------------|----------------|
| .doc, 645 KB       | 45               | 11                    | 3                     | 115              | 6              |
| .xls, 3488 KB      | 160              | 13                    | 3                     | 246              | 12             |
| .doc, 2878 KB      | 78               | 35                    | 3                     | 468              | 8              |
| .doc, 1671 KB      | 58               | 21                    | 3                     | 318              | 6              |
| .doc, 19 KB        | 36               | 3                     | 3                     | 46               | 6              |
| .xls, 14 KB        | 41               | 4                     | 3                     | 39               | 13             |
| Total              | 6 мин. 58 сек    | 1 мин. 27 сек         | 18 сек                | 20 мин 32 сек    | 51 сек         |

Табл. 5. Компания розничной торговли. Передача данных (в сек) между USA и Hong Kong (bandwidth: 232 kbps; latency: ~280 ms)

| File Type and Size    | WAN-Rd без TWAFS | WAN-Rd с TWAFS “cold” | WAN-Rd с TWAFS “warm” | WAN-Wr без TWAFS | WAN-Wr с TWAFS |
|-----------------------|------------------|-----------------------|-----------------------|------------------|----------------|
| .doc, 24 KB (drawing) | 54               | 15                    | 14                    | 55               | 10             |
| .doc, 160 KB (tables) | 122              | 18                    | 10                    | 69               | 10             |
| .doc, 79 KB           | 60               | 18                    | 11                    | 75               | 11             |
| .xls, 319 KB          | 40               | 15                    | 9                     | 43               | 11             |
| .xls, 1.25 MB         | >180             | 25                    | 13                    | 82               | 9              |
| .ppt, 3 MB            | >150             | 100                   | 13                    | 166              | 9              |
| Total                 | 9 мин. 56 сек    | 3 мин. 11 сек         | 1 мин 10 сек          | 8 мин 10 сек     | 1 мин 10 сек   |

Все эти особенности позволяют значительно повысить производительность Microsoft-приложений в WAN-среде, что иллюстрируется табл. 4 и 5.

Масштабируемость – до 500 пользователей на WAFS Edge Appliance.

### Решение Cisco WAFS (File Engine)

Во многом решение Cisco WAFS похоже на первое из рассмотренных нами решений – HP StorageWorks EFS WAN Accelerator, но в отличие от него, Cisco WAFS строится на базе одного устройства – файлового ускорителя, которое в зависимости от загруженного ПО могут исполнять следующие роли (рис. 4):

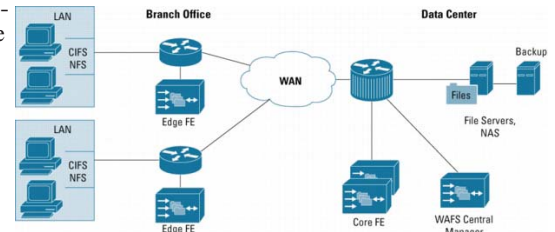


Рис. 4. Структурная схема развертывания решения Cisco WAFS.

- удаленного файлового ускорителя (Edge File Engine) – устанавливается в удаленных офисах, заменяет локальные файловые и принт-серверы; рекомендуется для обслуживания до 150 пользователей на текущей аппаратной платформе;
- центрального файлового ускорителя (Core File Engine) – устанавливается в центре обработки информации в непосредственной близости от файловых серверов; рекомендуется для обслуживания до 4000 пользователей на текущей аппаратной платформе;
- системы управления WAFS (WAFS Central Manager) – нужна только для настройки системы;
- системы репликации WAFS (WAFS File Replicator) – нужна только в начальный период эксплуатации для перемещения содержимого локального файлового сервера на централизованный.

Аппаратная платформа для Cisco WAFS представлена тремя устройствами: FE-511, FE-611 и FE-7326, ранжированными по производительности. ПО может быть настроено на работу в качестве всех трех компонентов сети (удаленный или централизованный файловый ускоритель, репликатор и менеджер) в один и тот же момент времени.

При необходимости наращивания мощности в удаленном офисе могут быть установлены несколько устройств с балансировкой нагрузки между ними (автоматической или ручной направленным групп клиентов на конкретные кэширующие устройства). Центральные устройства допускают кластеризацию с автоматической балансировкой нагрузки и практически линейным ростом производительности по причине отсутствия необходимости в запоминании состояний сессии.

Данное решение не требует инсталляции на клиентские компьютеры и никакое дополнительное ПО. Это решение

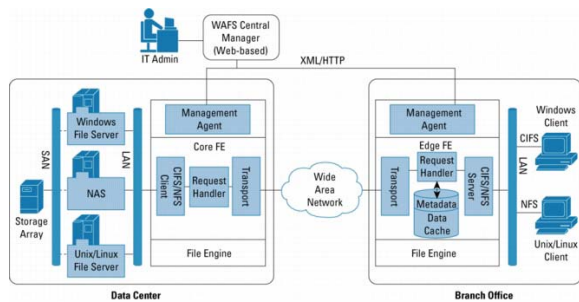


Рис. 5. Архитектура решения Cisco WAFS.

прозрачно для конечных пользователей и легко интегрируется в любую среду.

С архитектурной точки зрения, обычное решение Cisco WAFS состоит из множества Edge File Engines, которые распределены в удаленных сайтах, и одном или более Core File Engines, обеспечивающих доступ к центральным файловым серверам. С логической точки зрения, Windows- и UNIX-пользователи и приложения работают с центральным файловым сервером, используя стандартные CIFS и NFS протоколы, соответственно. Они делают это тем же самым способом, как и при обращении к любым другим CIFS или NFS файловым серверам или NAS (рис. 5).

В файловом ускорителе применяется новый подход к кэшированию — на уровне протокола (CIFS, NFS). Кэшируется не только сама информация, но и ее среда (контекст). На рынке много продуктов (в том числе и условно бесплатных) по довольно качественному кэшированию файлов, но это первый продукт по кэшированию протокола.

Кэширование только файлов (безо всех служебных сообщений протокола файлового обмена) позволяет снизить нагрузку на сеть, но не решает главную проблему — сниженные задержки. Учитывая рассмотренную практику программирования, кэшированием протокола удастся снизить задержку на 93% (документ Word на глобальных каналах начинает открываться не за минуты, а за секунды).

Кэширование на уровне протокола предполагает понимание принципов работы сети (в данном случае — TCP/IP) и тесное взаимодействие с сетью. Например, сеть TCP/IP допускает потерю пакетов. Если использовать для передачи одну сессию TCP, то на каналах большой производительности придется применять глубокую буферизацию (TCP — протокол с подтверждением), и в случае сброса единственного пакета придется переподать большой объем информации. WAFS открывает множество сессий (пропорционально мощности каналов) и дробит крупные пакеты на части.

В дополнение к новым протокольным принципам кэширования не забыты и хорошо известные файловые: упреждающее чтение, асинхронная запись через буфер, передача через сеть только измененных блоков, компрессия информации (средний коэффициент 2:1, но при определенных видах информации — вплоть до 20:1).

Основные функциональные особенности Cisco WAFS приведены в табл. 6.

Табл. 6. Основные функциональные особенности WAFS

| Функциональность   | Описание  |
|--|---|
| Полная замена файлового сервера в удаленном офисе                  | Обеспечиваются файловый и принтерный сервисы, включая распространение программ и патчей (таких, как Microsoft Systems Management Server)  |
| Кэширование на уровне протоколов                                   | Обеспечивает доступ на запись и чтение с производительностью и задержкой, как в LAN, при хранении информации в центре обработки информации, удаленном по WAN <ul style="list-style-type: none"> <li>Локальное кэширование протоколов CIFS и NFS — изменения файлов и каталогов немедленно кэшируются локально и затем передаются на центральный файловый сервер</li> <li>Локальная обработка метаданных — метаданные (такие, как атрибуты файлов и директорий) обслуживаются локально</li> <li>Частичное кэширование файлов — при записи передается только измененная часть файла, если при чтении указывается не весь файл, то и чтение производится только части файла, что серьезно экономит полосу пропускания каналов связи</li> </ul>   |
| Снижение задержки за счет оптимизации протоколов                   | <ul style="list-style-type: none"> <li>Кэширование Write-Back — центральный файловый ускоритель буферизирует операции записи от многих удаленных файловых ускорителей без риска нарушить целостность данных</li> <li>Read-ahead — если приложение читает файл последовательно, то файловый ускоритель начинает буферизировать данные, не дожидаясь очередного запроса от пользователя</li> <li>Кэширование отрицательных ответов — информация об отсутствующих файлах и каталогах запоминается</li> <li>Оптимизация Microsoft RPC (remote procedure call) — кэширование запросов и ответов позволяет значительно снизить задержки</li> <li>Алгоритм предсказания и уменьшения управляющих сообщений — значительно снижает задержку</li> </ul> |
| Оптимизация используемой полосы на каналах связи                   | Минимизация количества информации, передаваемой через WAN, что особенно важно для каналов связи низкой производительности <ul style="list-style-type: none"> <li>Конвейерная обработка</li> <li>Оптимизация работы TCP</li> <li>Увеличенные сегменты при потоковом чтении</li> <li>Ограничение на используемую полосу — чтобы оставить ресурсы и для других приложений</li> <li>Компрессия пакетов</li> <li>Использование IP QoS — интерактивные запросы помечаются более высоким приоритетом</li> </ul>  |
| Целостность данных   | Сброс данных на диск из КЭШа после операции закрытия файла для дополнительной защиты против неизвестных проблем (сбои по питанию, программные ошибки...)  |
| Многоуровневая связность   | WAFS полностью соответствует семантике CIFS по обеспечению актуальности данных и метаданных. Существует возможность одновременной работы из нескольких удаленных офисов с одним файлом  |
| Заблаговременное кэширование в соответствии с заданными политиками | Администратор системы может указать, какая информация должна быть кэширована всегда на конкретных удаленных файловых ускорителях  |
| Высокая доступность  | Несколько уровней <ul style="list-style-type: none"> <li>Центральные файловые ускорители — могут быть объединены в кластеры с резервированием N+1 и балансировкой нагрузки</li> <li>Возможность запоминания (Backup) и восстановления (Restore) конфигурации файловых ускорителей для их быстрой замены</li> <li>Удаленные файловые ускорители — кластеризация может быть обеспечена за счет поддержки DFS (Microsoft Distributed File System)</li> </ul>   |

## Программные решения оптимизации WAN-трафика — NetBackup PureDisk

Как показывают исследования, традиционные технологии управления и защиты данных намного отстают от современных методов обработки данных. В большинстве организаций концептуально технологии резервного копирования не меняются в течение многих десятилетий (!). Между тем, ежегодный 60% рост корпоративных данных и глобализация бизнеса, приводящая к распределенным IT-WAN-инфраструктурам, требуют новых подходов и решений вследствие того, что в WAN-средах даже поддерживать уже существующий уровень доступности и надежности данных старыми методами становится невозможным.

ПО NetBackup PureDisk компании Symantec ориентировано на реализацию одной задачи — оптимизации процессов резервного копирования в среде WAN. NetBackup PureDisk — продукт, который относится к сектору, развиваемому на т.н. технологии маршрутизации контента (content routing), и который поддерживает различные типы доступа к данным (SAN, DAS, NAS) на любой платформе от любого вендора.

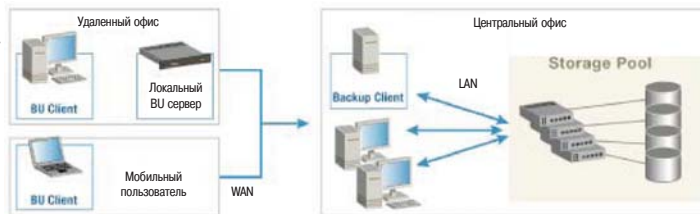
PureDisk использует патентованную технологию сигнатур (Single Instance Store — SIS), которая формирует уникальную 128-битную последовательность на основе контента файла и позволяет отличать уникальные файлы от множественных его копий. Устранение излишнего хранения копий и их пересылка дает возможность экономить значительные ресурсы и резко уменьшать сетевой трафик.

В обычной системе приблизительно 15% файлов/байт изменяются и, соответственно, резервируются каждый день. При использовании технологии, реализованной в PureDisk, эту величину удастся уменьшить до 1–2%. Например, при общем объеме данных 100 Гбайт и еженедельном выполнении 6 инкрементальных резервных копий и одной полной общей копией для backup'a, — более чем 100 Гбайт. При использовании PureDisk на это потребуется менее чем 10 Гбайт.

PureDisk также делит большие файлы на меньшие сегменты. Это дает очевидное преимущество для резервной копии mail.pst файлов, содержание которых изменяется незначительно каждый день. Вырезая большие файлы в маленькие порции данных, система будет резервировать только маленькие части, которые содержат новые данные, в то время как основная часть файла остается неизменной и, таким образом, не включается в резервную копию.

Функционирование PureDisk строится на концепции распределенных агентов. Агент PureDisk постоянно находится на каждой защищенной системе (настольный компьютер, ноутбук, сервер и др.) и вычисляет «отпечатки» файлов. При обнаружении новых уникальных файлов агент пересылает их на сервер резервного копирования. Агенты управляются системным администратором и выполняют процедуры резервного копирования на каждом PC и каждом сервере — автоматически в соответствии с корпоративными рекомендациями/правилами.

Все агенты независимы, соответственно, множественные резервные задачи могут быть выполнены одновременно. Агенты на каждом компьютере работают в фоновом режиме, но их приоритет можно менять.



Типичная среда, защищаемая PureDisk, имеет резервный сервер на каждом сайте для локальных резервных копий и центральный сервер — для реплицирования локальных серверов. Агенты установлены на всех системах на удаленных сайтах (включая портативные компьютеры). Через агенты удаленные сайты также автоматически “следуют” за корпоративными резервными рекомендациями/политиками. Агенты на всех системах посылают новые, уникальные файлы локальному резервному серверу, который проверяет уникальность файлов на всех местных системах сайта и только тогда посылает (данные перед пересылкой шифруются) действительно уникальные файлы на основной резервный сервер (рис. 6). Это минимизирует требования к полосе пропускания WAN и обеспечивает долгосрочную масштабируемость системы. Всегда существуют две полных резервных копии удаленного сайта — одна на локальном сервере и вторая — на центральном сервере. Таким образом на удаленном сайте можно полностью отказаться от хранения данных на ленте, а дополнительную копию данных (для обеспечения катастрофоустойчивости) хранить только централизованно.

Решение, основанное на PureDisk, гарантирует, что все данные от всех удаленных сайтов имеют резервную копию, работает полностью автоматически, сокращая таким образом возможные человеческие ошибки и в соответствии с заданными политиками. Через прямой интерфейс каждый пользователь имеет возможность восстановления любых

**Рис. 6.** Концепция распределенных агентов обеспечивает уровневую проверку на уникальность и пересылку файлов: от локальных агентов к локальному серверу и только затем — к центральному серверу резервного копирования.

файлов даже на своем домашнем PC с центральной резервной системы.

### Заключение

*Представленные в обзоре решения достаточно сильно различаются по архитектуре построения и, зачастую, по позиционированию на рынке, тем не менее, решают одну и ту же задачу: ускорение работы пользователей и уменьшение загрузки канала передачи данных.*

*Решения компаний HP и Cisco выглядят наиболее функциональными и законченными, что, впрочем, неудивительно — соответствующие продукты появились на рынке раньше других. Вследствие своей универсальности, они хорошо подходят клиентам, если в удаленных офисах есть ряд разнородных приложений, часть из которых не используют платформу от Microsoft, а канал загружен достаточно сильно. Необходимо только отметить, что решение от Cisco изначально было сфокусировано на файловых сервисах, в то время как решение от HP является более комплексным.*

*Следует учесть, что для принятия решения по реализации большого проекта следует провести тестирование подобных систем от нескольких поставщиков. При этом следует отметить, что за счет большого объема внутренних дисков решение от компании HP позволяет более глубоко осуществлять кэширование, а у решение от компании Cisco хорошая масшта-*

*бируемость архитектуры, удобная для увеличения производительности при обслуживании удаленных офисов.*

*Решение от компании Brocade появилось относительно недавно и сфокусировано на тех компаниях, у которых в удаленных офисах*

*используется только ПО Microsoft, что встречается довольно часто. Такая фокусировка имеет смысл и потенциально означает лучшее и оптимальное решение поставленной задачи с точки зрения цены. Заметим еще раз, что каждый заказчик имеет свой набор приложений и требований к ним, и выбор техники следует производить только после тестирования. При небольших проектах это не всегда возможно. Поэтому в таких случаях заказчику придется определять перечень приложений и требований к задержкам/загрузке канала, а также оценивать решения на основе ориентировочных показателей, полученных от поставщиков, для такого рода приложений.*

*Решение от компании Symantec стоит несколько особняком — его можно использовать одновременно с другими решениями по оптимизации трафика, т.к. оно интегрировано с системой резервного копирования и решает достаточно узкую задачу. Оно очень хорошо подходит тем компаниям, которые хотят оптимизировать инфраструктуру резервного копирования на местах и обеспечить надежное хранение и быстрый доступ к резервируемой информации. Над таким решением также следует задуматься, если пропускной способности канала не хватает для ночного резервного копирования на центральную площадку, особенно, если на предприятии уже используется ПО NetBackup.*

**Василий Кострюков, vkostr@lanit.ru**  
Руководитель направления систем хранения данных, ЛАНИТ