

## ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ КЭШИРОВАНИЯ В КОРПОРАТИВНЫХ СЕТЯХ

© А.В. Пономарев, В.Ю. Носков, 2012

ФГАОУ ВПО «Уральский федеральный университет  
имени первого Президента России Б.Н. Ельцина», г. Екатеринбург

### Предпосылки к кэшированию трафика

Как развивается Интернет, и в какую сторону, можно приблизительно знать, анализируя статистические данные. Статистика Интернет говорит об общем неуклонном росте и количества новых пользователей и количества новых ресурсов [1].

Таблица 1

#### Рост интернет-пользователей

Мировая статистика роста интернет пользователей на март 2011 г.					
Регион	Население (2011 г.)	Интернет пользователь и (2001 г.)	Интернет пользователи на текущий момент	Рост числа пользователей (% населения)	Прогресс 2000-2011 гг.
Африка	1,037,524,058	4,514,400	118,609,620	11,4 %	2,527,4 %
Азия	3,879,740,877	114,304,000	922,329,554	23,8 %	706,9 %
Европа	816,426,346	105,096,093	476,213,935	58,3 %	353,1 %
Средний Восток	216,258,843	3,284,800	68,553,666	31,7 %	1,987,0 %
Северная Америка	347,394,870	108,096,800	272,066,000	78,3 %	151,7 %
Латинская Америка	597,283,165	18,068,919	215,939,400	36,2 %	1,037,4 %
Океания/Австр алия	35,426,995	7,620,480	21,293,830	60,1 %	179,4 %
<b>Итого</b>	<b>6,930,055,154</b>	<b>360,985,492</b>	<b>2,095,006,005</b>	<b>30,2 %</b>	<b>480,4 %</b>

### Постоянный рост потребностей в производительности Интернет

Сегодня все чаще и чаще возникают повышенные требования к пропускной способности каналов между клиентами Сети и серверами. Это происходит по разным причинам:

- повышение производительности клиентских компьютеров;
- увеличение числа пользователей в Сети;
- появление приложений, работающих с мультимедийной информацией;
- увеличение числа сервисов, работающих в реальном масштабе времени.

Следовательно, имеется потребность в экономичном решении, предоставляющем нужную пропускную способность во всех перечисленных случаях. Ситуация усложняется еще и тем, что нужны различные технологические решения – для организации магистралей сети и подключения серверов одни, а для подключения настольных клиентов – другие. Стоит отметить, что за последнее время средняя скорость Интернета на планете возросла и составляет около 1,7 Мб/с. Южная Корея, согласно последним результатам тестирования скорости Интернета, подтвердила статус лидера, увеличив среднюю скорость более чем на 20 % по отношению к прошлому году (14,7 Мб/с). Такой значительный рост из всех стран первой десятки наблюдался только у Ирландии. Румыния, Швеция и Чехия снизили средние результаты за последний квартал, хотя ранее там наблюдался только рост. Главным же

являются тот факт, что средняя скорость Интернета растет благодаря постоянному усовершенствованию оборудования и прокладке новых магистралей.

### **Увеличенное время доступа к ресурсам**

Но переоборудование сетей операторов связи достаточно дорого и требует больших инвестиций. К сожалению, появление современных технологий у пользователей происходит быстрее, чем обновление программно-аппаратной платформы у интернет-провайдеров. Даже проектирование Сети с «запасом» на рост потребностей пользователей не обеспечивает ожидаемый эффект в силу того, что зачастую глобальные узлы не способны удовлетворить требования всех пользователей. Проблема появляется на этапе извлечения контента для пользователей. Современные сайты строятся на основе генерируемых страниц, что само по себе увеличивает нагрузку на сервер. А также все чаще использует весомые медиа-файлы, пересылка которых по Глобальной сети происходит для каждого обратившегося пользователя. Получается определенный замкнутый круг. Пользователи получают контент медленнее, чем ожидают, и увеличивают свою тарифную пропускную способность, вследствие чего глобальный узел должен использовать режим с повышенной вычислительной нагрузкой, что замедляет обработку запросов других пользователей. Те пользователи, запросы которых обрабатываются медленнее, обвиняют во всем оператора связи и фактически выбирают тариф с большей скоростью канала. Эта ситуация распространена даже в тех сетях, где провайдер может обеспечить номинальную скорость внутри своей сети, а об ответственности за скорость доступа к глобальным узлам говорить не приходится.

### **Кэширование трафика в корпоративных сетях**

#### ***Краткое описание технологии кэширования***

Кэширование – технология, которая вносит элемент интеллектуальности в концепцию хранения и пересылки данных, определяя, какие информационные ресурсы будут скопированы из ядра на периферию сети и как часто они будут обновляться. В основу технологии кэширования положен тот факт, что данные дешевле и эффективнее хранить, чем передавать. На самом деле, эта идея не нова: в компьютерах и других электронных вычислительных устройствах данные хранятся локально для уменьшения времени доступа к ним, тот же принцип реализуется и в браузерах Web, где последние просмотренные страницы Web находятся в кэш-памяти на жестком диске пользователя. Кэширующие серверы могут также применяться в корпоративной локальной сети, способствуя повышению производительности Intranet., причем установка кэширующих серверов предполагается даже в филиалах компаний, где сегодня они применяются очень редко. По словам представителя Meta Group Питера Ферстбрука, технология кэширования в корпоративной сети позволяет экономить денежные средства, так как появляется возможность отложить на некоторый срок наращивание мощности брандмауэров и устройств доступа к глобальной сети.

#### ***Методы оценки эффективности кэширования в корпоративных сетях***

Снижение нагрузки зависит от отношения интенсивности запросов страницы пользователями и интенсивности обновления фрагментов. В веб-системах в большинстве случаев интенсивность запросов страницы значительно превосходит интенсивность обновлений фрагментов. При решении проблемы ограниченного пространства кэша используются алгоритмы замещения объектов в кэше, оставляющие самые «ценные» данные. Ценность данных – это обычно функция от нескольких параметров, таких как давность последнего запроса, частота запросов, размер, стоимость получения объекта, частота обновлений объекта. Политика замещения принимает решения, основываясь на этих значениях.  $D$  – множество запрашиваемых объектов  $\{d_1, d_2, \dots, d_n\}$ .  $R_d$  – общее число обращений к объекту  $d$  за период времени,  $H_d$  – число обращений к объекту, который уже находится в кэше, за период времени,  $S_d$  – размер объекта и  $T_d$  – время, требуемое для генерации объекта (стоимость получения объекта). Критерии оценки эффективности работы кэша:

1. Соотношение обращений к объекту, обслуженных из кэша, к общему количеству запросов объекта
2. Соотношение размеров объектов обращения, к которым обслужены из кэша к общему размеру запрашиваемых объектов
3. Соотношение времени получения объектов, обращения к которым обслужены из кэша к общему времени получения запрашиваемых объектов

$$R_h = \frac{\sum_{d \in D} Hd}{\sum_{d \in D} Rd}$$

$$R_s = \frac{\sum_{d \in D} Sd * Hd}{\sum_{d \in D} Sd * Rd}$$

$$R_t = \frac{\sum_{d \in D} Td * Hd}{\sum_{d \in D} Td * Rd}$$

### Эффективность средств кэширования в зависимости от численности сотрудников и интенсивности использования Интернет

Как видно из формул

$$R_h = \frac{\sum_{d \in D} Hd}{\sum_{d \in D} Rd} \quad R_s = \frac{\sum_{d \in D} Sd * Hd}{\sum_{d \in D} Sd * Rd} \quad R_t = \frac{\sum_{d \in D} Td * Hd}{\sum_{d \in D} Td * Rd}$$

(где Rd – общее число обращений к объекту d за период времени, Hd – число обращений к объекту, который уже находится в кэше, за период времени, Sd – размер объекта и Td – время), с ростом числа запросов к файлу (т.е. с ростом числа персонала) эффективность кэша будет увеличиваться, т.к. увеличится величина Hd – число обращений к объекту, который уже находится в кэше. Также увеличится эффективность кэширующего аппарата, если многие пользователи будут работать с одним и тем же файлом большого размера. Пример ситуации: если в Сети появляется новая серия любимого сериала, скаченный одним пользователем материал будет транслироваться с локальных узлов на гораздо большей скорости. Если рассматривать более рабочий пример: с ростом числа сотрудников на предприятии использующим общую корпоративную информационную систему, увеличится число запросов на сервер, который генерирует рабочий интерфейс. Если данные, отвечающие за интерфейс, кэшировать, это приведет к существенной экономии времени на подготовку в рабочему процессу.

Приблизительная эффективность использования кэширующих устройств в зависимости от частоты обращения к системе представлена в табл. 2.

Таблица 2

Эффективность кэширования Http-трафика от числа обращений

Число обращений к системе в секунду	% кэшированного трафика
1	15
3	39
4	38
6	53
9	68
14	34
100	55

### Продукты, выполняющие кэширование трафика

#### Классификация средств кэширования

Эволюция функций, особенности применения и ценовой политики привели к естественному разделению рынка на две категории:

- программные средства на универсальных платформах;
- специализированные аппаратные комплексы.

Все они выполняют одни и те же задачи, но разными способами с различными начальными данными.

### ***Программные средства на универсальных платформах***

Полностью программные средства кэширования поставляются большинством производителей операционных систем. Примерами могут служить squid cache и HandyCache. Некоторые системные администраторы предпочитают иметь дело с ПО кэширования, работающим под управлением знакомых им операционных систем или с продуктами одного производителя. В дополнение к относительной дешевизне, простоте установки и конфигурирования, к достоинствам программных систем кэширования можно отнести также возможность их работы на существующих серверах UNIX, или Windows NT. Хотя эта особенность позволяет обойтись при установке системы кэширования без изучения незнакомой операционной системы, таким продуктам не хватает масштабируемости и гибкости – качеств, необходимых для того, чтобы кэш работал с требуемой производительностью в специфических конфигурациях сети. Тем не менее программный подход вполне применим в локальных сетях средних размеров.

Второй категорией продуктов являются предустановленные серверы кэширования. В комплекс обычно входят ОС UNIX или NetWare, специальное программное обеспечение и готовое стандартное оборудование. Такие комплексы (например, Dell PowerApp.cache) обладают более высокой степенью масштабируемости и управляемости, а также просты в установке. Однако они узко специализированы и не решают никаких других задач, кроме кэширования.

### **Специализированные аппаратно/программные решения**

Рассматриваемая категория продуктов предназначена для крупных сетей, так как здесь базовые функции кэширования сочетаются с высокой надежностью и отказоустойчивостью. Данный подход целесообразно использовать для долгосрочных проектов. Эта категория оборудования работает под управлением собственной ОС, специально разработанной для оптимизации кэширования. Модификация ядра ОС повышает эффективность обработки продолжительных потоков данных, файлов большого размера и трафика длительных сеансов связи. Среди прочих на рынке специализированных кэширующих устройств фигурируют компании CacheFlow, Oversi, Juniper, PeerApp. Специализация продуктов приводит к появлению более развитых и специфических решений: проху для Web, управления контентом, сетевого резервирования и других функций, относящихся к доставке информационных ресурсов. Пока что рынок плохо принимает специализированные аппаратные системы кэширования, особенно с учетом конкуренции со стороны недорогих многоцелевых серверов, традиционно доминирующих по объемам продаж. Все еще существует обширная категория сетевых администраторов, избегающих применения специализированных устройств кэширования, несмотря на предоставляемые ими возможности централизованного управления.

### **Список использованных источников**

1. Информационный ресурс «INTERNET USAGE STATISTICS». [Электронный ресурс]. Режим доступа <http://www.internetworldstats.com/stats.htm>.
2. Информационный ресурс «Кэширование ресурсов Internet». [Электронный ресурс]. Режим доступа <http://kunegin.narod.ru/ref2/bal/kesh.htm>.