

УДК 681.3.06

А.Л.Хотелева (5 курс, каф. РТиТК), В.С.Синепол, к.т.н., проф.

## ИССЛЕДОВАНИЕ ЭФФЕКТИВНОСТИ ПОЛИТИК КЭШИРОВАНИЯ WEB-ОБЪЕКТОВ

**ABSTRACT:** Caching in the World Wide Web is a widely used technique to reduce bandwidth demand and improve response times for popular objects. There are several available cache implementations. The replacing policy used by a proxy server effects on its efficiency. The results published in the related works don't give certain information about cache efficiency nowadays. This paper discusses the results of implementation of two replacement policies on base of proxy server Squid in the real network.

Как объем, так и удельный вес Web-трафика в компьютерных сетях постоянно растет. Сегодня он составляет 70 – 90% общего объема входящего трафика для большинства корпоративных сетей. Кэширование Web-объектов является одним из возможных путей уменьшения загрузки сетевых каналов этим трафиком и повышения доступности информации.

Эффективность технологии кэширования, наряду с производительностью процессора и системы ввода-вывода прокси-сервера, существенно зависит от алгоритма, в соответствии с которым производится помещение web-объектов в память и их замена (удаление). Попытки учесть разнообразные характеристики web-объектов в условиях отсутствия общепризнанного критерия эффективности прокси-сервера привели к созданию достаточно большого числа как однопараметрических, так и многопараметрических алгоритмов замещения объектов в кэше. Приводимые в литературе [1-4] данные о влиянии правила замещения объектов в памяти прокси-серверов на показатели его эффективности носят противоречивый характер. Кроме этого, условия, при которых применение этой технологии может быть эффективным, исследованы также недостаточно полно.

Подробный анализ 9 политик замещения и изучение их влияния на эффективность кэширования web-объектов позволяет сделать следующие выводы.

1. Эффективность кэширования, в смысле параметров hit rate (HR) и byte hit rate (BHR), существенно зависит от объема кэш-памяти и применяемой политики замещения объектов.

2. При размере памяти менее 256 ГБ выбор оптимальной политики замещения может повысить эффективность прокси-сервера на 5 – 10 %.

3. При размере кэш-памяти более 256 ГБ политика замещения практически не влияет на эффективность кэширования.

4. Абсолютные значения HR и BHR для одних и тех же политик замещения и одинаковых объемов кэша значительно (до 30%) изменяются при применении их на прокси-серверах в различных сетях, что свидетельствует о существенном влиянии на эти показатели структуры запросов пользователей.

5. Многопараметрические правила кэширования обычно превосходят однопараметрические по обоим показателям эффективности. В числе однопараметрических правил замещения наиболее предпочтительным представляется правило LRU (Least-Recently-Used), которое, несколько уступая правилам GDSF (Greedy-Dual-Size-Frequency) и

LFUDA (Least Frequently Used with dynamic aging) по показателю HR, не менее эффективно по показателю BHR и существенно проще в реализации.

Экспериментальное сравнение эффективности web-кэширования при использовании различных политик замещения было проведено в условиях компьютерной сети 14 корпуса общежития СПбГПУ, насчитывающей примерно 200 хостов. Условия и результаты эксперимента приведены в табл. 1.

Таблица 1.

Применяемая политика замещения	LRU	GDSF
Время проведения эксперимента	13.05.04 -25.05.04	07.06.04 – 19.06.04
Количество дней	13	13
Количество запросов	1 278 079	1 541 411
Общий объем запрошенных документов, байт	5 919 581 809	6 013 236 694
Объем кэша, МБ	1000	1000
Количество документов, доставленных из кэша	372 732	598 624
Общий объем документов, доставленных из кэша, байт	535 913 203	548 099 882
Hit Rate, %	29,16	38,84
Byte Hit Rate, %	9,05	9,11

В качестве прокси-сервера был использован сервер Squid 2.5, позволяющий выбрать одну из трех политик замещения: LRU, GDSF, или LFUDA.

Полученные результаты подтверждают преимущество политики GDSF над LRU по показателю HR (39% и 29%, соответственно). Однако значения показателя BHR для обеих рассмотренных политик практически не отличаются и оказались относительно низкими (9%).

Экспериментальные значения HR и, особенно, BHR оказались заметно ниже показателей, получаемых на основании моделирования [4]. Анализ этого обстоятельства указывает на необходимость более полного учета параметров web-трафика при моделировании процедур кэширования. На основании анализа опубликованных в [1-3,5] данных можно утверждать, что для адекватного моделирования http прокси-сервера в набор существенных характеристик web-трафика и содержимого кэша прокси-сервера следует включать следующие характеристики:

- распределение размеров уникальных запрошенных web-объектов;
- распределения размеров всех запрошенных web-объектов;
- распределение количества  $k$ -кратных запросов объектов  $k=1,2,3,\dots$ (популярность);
- вероятность повторных запросов в зависимости от размера объектов, числа запросов каждого объекта и от длительности их хранения в КЭШе;
- зависимость частоты запросов объектов от их размера.

Для некоторых из указанных зависимостей уже существуют хорошо апробированные модели, но большая их часть пока изучена недостаточно. Построение аналитических и имитационных моделей современного web-трафика является одной из задач дальнейшего исследования.

#### ЛИТЕРАТУРА:

1. Barford Paul, Crovella Mark, “An Architecture for a WWW Workload Generator”, Computer Science Department, Boston University, 1997.
2. Rizzo Luigi, Vicisano Lorenzo, “Replacement policies for a proxy cache”, UCL-CS Research Note RN/98/13.
3. Barford Paul, Bestavros Azer, Bradley Adam, Crovella Mark, “Changes in Web Client Access Patterns”, Computer Science Department, Boston University, 1998.

4. Arlitt Martin , Dilley John, “Improving Proxy Cache Performance-Analyzing Three Cache Replacement Policies”, Internet System and Applications Laboratory, HP Laboratories, HPL-1999-142, 1999.
5. Barford Paul, Crovella Mark, “Generating Representative Web Workloads for Network and Server Evaluation”, Computer Science Department, Boston University, 1998