



OSTIS-2015

(Open Semantic Technologies for Intelligent Systems)

УДК 004.432.4

РЕШЕНИЕ ЗАДАЧИ ПРОГНОЗИРОВАНИЯ ВОЗРАСТАНИЯ НАГРУЗКИ В СИСТЕМЕ ДИСТАНЦИОННОГО ОБУЧЕНИЯ «ISPRING LEARN»

Сокольников А.М., Сидоркина И.Г.

*Поволжский государственный технологический университет,
г. Йошкар-Ола, Республика Марий Эл, Россия*

sokolnikov.alexey@gmail.com

igs592000@mail.ru

В статье исследуются способы прогнозирования высокой нагрузки для web-сервисов. Обосновывается необходимость их использования в системе дистанционного обучения iSpring Learn [iSpring Learn, 2014]. Раскрываются проблемы, которые возможно решить с помощью методов прогнозирования возрастания нагрузки и преимуществ их использования в связке с балансировщиком нагрузки.

Ключевые слова: Высоконагруженные системы; Система дистанционного обучения iSpring Learn; Балансирование нагрузки, Прогнозирование возрастания нагрузки.

Введение

Для разработчиков программного обеспечения сегодня все острее встает проблема обработки и хранения больших объемов данных. Количество просмотров страниц в социальной сети Facebook превышает 200 миллиардов в месяц [Блинков, 2010]. Для ускорения работы сервис использует несколько серверов, а обращения клиентов между ними распределяет особый класс программ, называемый «Балансировщик нагрузки».

Балансировка нагрузки — метод распределения заданий между несколькими сетевыми устройствами с целью оптимизации использования ресурсов, сокращения времени обслуживания запросов, горизонтального масштабирования кластера, а также обеспечения отказоустойчивости.

Большинство крупных сервисов используют данный класс программного обеспечения, но проблема заключается в том, что сервис не будет готов к резкому возрастанию нагрузки. Такая ситуация приведет к задержкам в выполнении запросов и возрастанию нагрузки на отдельно взятом сервере. Как один из вариантов решения этой проблемы предложено использование средства превентивного анализа высоких нагрузок.

1. Средство превентивного анализа нагрузки

Методы прогнозирования возрастания нагрузки

анализируют и обрабатывают запросы к серверу и сохраняют их метаданные. Кроме этого, они опрашивают состояние загруженности серверов. На основе этих данных производится поиск закономерностей выявления высоких нагрузок и вынесение решения о необходимости разворачивания дополнительных экземпляров сервиса.

Преимущество данного подхода заключается в том, что при использовании методов прогнозирования возрастания нагрузки в связке с балансировщиком, количество пиков запросов на один сервер значительно сократится, поскольку об авральной ситуации будет известно заранее до ее фактического возникновения.

Такие методы назовем средством превентивного анализа нагрузки, то есть средой, опережающей действия противной стороны [Комлев, 2006]. Превентивный анализ осуществляет обработку данных в запросах и пытается выявить закономерности в скачках нагрузок. Выделим четыре основных типа зависимостей, влияющих на нагрузку: календарные, временные, геолокационные и зависимости от источников трафика.

1.1. Метод выявления календарных закономерностей

Одной из закономерностей скачков нагрузки может служить зависимость сервиса от конкретных дат. Наиболее очевидный пример сервисов, нагрузка на которые зависит от календарных дат -

Автоматическая Телефонная Станция. В большинство будничных дней она справляется с возлагаемыми на нее нагрузками, однако в дни праздников у клиентов возникают проблемы с соединением.

Подобные проблемы можно наблюдать у таких сервисов, как социальные сети. Однако не стоит отрицать, что велика вероятность того, что проблема заключается не только на стороне сервиса, но и на стороне провайдера, предоставляющего доступ к сети в интернет.

Рассмотрим влияние календарных зависимостей на систему дистанционного обучения (СДО) iSpring Learn. Данную СДО используют более чем 2 000 ВУЗов, школ и компаний по всему миру. Минимальная численность организации – 50 пользователей. Самая крупная организация – более 23 000 пользователей. Годовой график нагрузки представлен на рисунке 1:

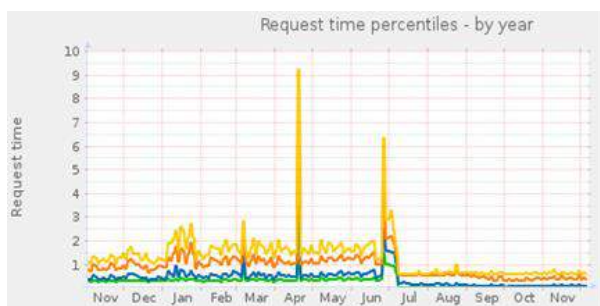


Рисунок 1 - Распределение нагрузки iSpring Learn за 2014 год

Помимо социальных сетей, зависимость от даты имеют периодические новостные издания. Примером может служить сайт журнала или блог, записи в котором публикуются с некой периодичностью. К примеру, каждую пятницу. Система превентивного анализа выявит эту закономерность и примет решение о необходимости дополнительных экземпляров сервиса.

Для выявления календарных закономерностей предложено использовать накопленные сервисом данные о нагрузке в разрезе дня. Для получения более точных прогнозов рекомендовано использовать синхронизацию с календарем событий. Примерами событий могут являться дни выпуска новой версии продукта, праздники, либо дни старта рекламной компании продукта.

1.2. Метод выявления временных закономерностей

Помимо календарных закономерностей необходимо брать в учет зависимости от времени суток. Время работы клиента с web-сервисом зависит от специфики сервиса. Таким образом, нагрузка распределяется неравномерно все 24 часов. В случае системы дистанционного обучения iSpring Learn график зависимости среднего времени выполнения скрипта от времени суток представлен на рисунке 2:

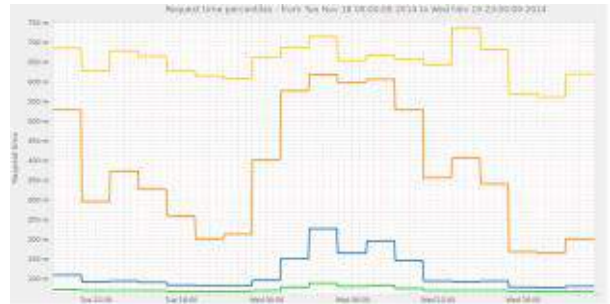


Рисунок 2 - Зависимость времени работы скрипта iSpring Learn от времени суток

На рисунке 2 отчетливо наблюдается скачок нагрузки в середине графика. Это связано с тем, что такие сервисы, как системы дистанционного обучения преимущественно используются клиентами в рабочее время. Время на сервере отличается от времени у клиента, и по графику заметно, что нагрузка нарастает с 00:00 и убывает к 12:00. Такой эффект возникает из-за разницы часовых поясов клиентов. В 3:00 проявляется ярко выраженный скачок – в этот момент наступает пересечение «рабочего времени» у клиентов с разными часовыми поясами.

Ссылаясь на описанные выше факты, можно утверждать, что нагрузка на сервере может отличаться не только в зависимости от дней, но и в зависимости от часов. По этой причине предварительный анализ трафика будет полезен и благодаря использованию системы прогнозирования появится возможность избежать случаев, когда нагрузка на сервере резко возрастает. Система выявит закономерности, проявляющиеся за день или неделю, и даст необходимые указания балансировщику нагрузки.

1.3. Метод выявления закономерностей источников трафика

При отслеживании календарных зависимостей сервиса, необходимо учитывать тот факт, что нагрузка может возрастать не только с периодичностью, связанной с событиями на анализируемом сервисе. Существуют варианты, когда периодичность загрузки одного сервиса зависит от периодичности другого. Примером такого взаимодействия может служить ссылка на анализируемый сайт, размещенная на ресурсе со значительным числом пользователей. Сервис Google Analytics [GoogleAnalytics, 2014] позволяет отслеживать адрес, с которого пришел клиент. Предоставляется несколько вариантов отчета: название поисковой системы, адрес сайта-источника, название одной из новостей и посещения тек, кто ввел адрес непосредственно в адресной строке.

Возросший трафик от поисковой системы будет говорить о том, что данный запрос сейчас является популярным. В большинстве случаев, такая популярность длится не более недели. Инструментальное средство превентивного анализа при увеличении доли трафика с поисковых систем

обеспечивает принятие решения о подготовке к высоким нагрузкам. Для анализа количества запросов компанией Google был разработан инструмент Google Trends [GoogleTrends, 2014], он позволяет оценивать количество использования ключевых слов не только в зависимости от времени, но и в зависимости от географического положения пользователей.

В случае резкого возрастания трафика с постороннего ресурса сети интернет, существует большая вероятность, что на нем была размещена ссылка на ваш ресурс. Если большого количества переходов с него до этого момента не наблюдалось, средство будет считать этот ресурс новостным. Новостные ресурсы имеют ленту новостей, которая в среднем обновляется за 1-2 дня. В таком случае средство примет решение об увеличении количества экземпляров сервиса на данный период. Если приток клиентов с этого ресурса уже наблюдался, то за период повышенной нагрузки будет принят средний период для этого ресурса.

Необходимо учитывать тот факт, что зависимость от источников трафика тесно связана с календарной зависимостью.

Таким образом, для реализации метода выявления закономерностей источника трафика, необходимо отслеживать адрес ресурса, с которого началось резкое увеличение количества пользователей и в случае, если этот источник – поисковая система – выяснить ключевое слово и определить популярность этого запроса. Основываясь на этих данных, появится возможность прогнозировать период времени высокой нагрузки сервиса.

1.4. Метод выявления геолокационных закономерностей

С целью повышения значимости данных прогноза, необходимо отслеживать географическое положение пользователя. Автором предложено определение региона с помощью IP-адреса запроса. Обладая данными географии запросов пользователей, в рекомендации, отсылаемой средством прогнозирования балансировщику нагрузки, появится возможность уточнения региона, в котором желательно подготовиться к скачку нагрузки.

Современные высоконагруженные сервисы используют сети дата-центров, расположенные по всему миру. К примеру, компания Facebook пользуется услугами дата-центров Силиконовой Долины в США и Шведскими дата-центрами в Европе [GigomResearch, 2013]. Компания iSpring использует два дата-центра. Первый находится в городе Даллас, США. Второй – в Москве. Такой подход позволяет избавить сервер, используемые российскими клиентами от нагрузок, связанными с запросами пользователей США. Благодаря такой организации работы инструментального средства существует возможность в случае чрезмерного

возрастания нагрузки в одном регионе увеличить для него производительные мощности серверов.

Таким образом, средство помимо календарной зависимости должно учитывать еще и геолокационную зависимость и при комбинации этих факторов принять решение об увеличении количества серверов в дата-центре определенного региона, к примеру США или Европы.

1.5. Прогнозирование нагрузок на основе выявленных закономерностей

На рисунке 3 показано схематическое представление web-сервиса, использующего методы прогнозирования возрастания нагрузки совместно с балансировщиком:

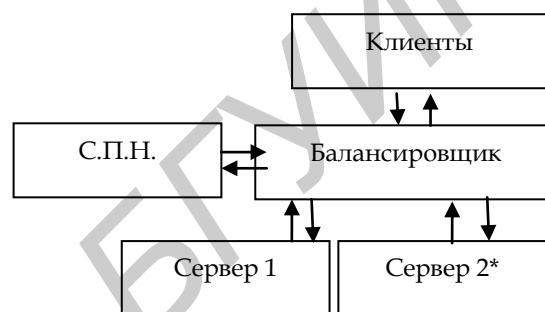


Рисунок 3 - Схематическое представление сервиса, использующего предложенную средство прогнозирования нагрузки.

Балансировщик нагрузки имеет статистику занятости серверов. Предположим, что в данный момент нагрузка на сервис не велика и с ней успешно справляется Сервер 1. Сервер 2* находится в режиме ожидания. Запросы на него не идут.

Введем обозначение С.П.Н. – средство, реализующее методы прогнозирования нагрузки. Если прогноз будет положительным и с потенциальной нагрузкой, рассчитанной анализатором сервис не справится, средство прогноза отправит запрос балансировщику на пробуждение Сервера 2*. Он выйдет из режима ожидания и начнет свою работу до того, как произойдет пик нагрузки.

При первом запуске средства прогнозирования ему задается максимальный приемлемый уровень загрузки сервера. Все запросы клиентов направляются на балансировщик. Он передает средству прогнозирования данные запросов.

Средство прогнозирования получает следующие данные: дата и время запроса, геолокация пользователя, количество и уровень загруженности серверов. Если доступны данные об источнике перехода, сохраняет и их. Сервис постоянно обучается, сохраняя новые данные. Критически важно отслеживать уровень загруженности серверов.

Для принятия решения системой используется метод экстраполяции для функции $L = f(t)$, где L –

уровень нагрузки сервера, а t – время. Если полученное в результате экстраполяции значение L больше, чем максимально заданное для сервера, будет принято решение об увеличении количества серверов.

При положительном решении для уточнения данных используются дополнительные параметры (источник перехода, географическое положение пользователя). В случае определения источника перехода будет определен его тип и в зависимости от него будет указана предположительная длительность увеличения нагрузки. Если скачок нагрузки происходит из конкретного региона, то средство прогнозирования нагрузки учтет этот факт.

Таким образом, в случае положительного прогноза, балансирующий получит данные в формате $\{T, G\}$. Где T – интервал времени, на который стоит увеличить количество серверов, G – регион, в котором необходимо это сделать.

С появлением средства, прогнозирующего появление высоких нагрузок, будет реализована возможность решения проблемы большого количества запросов заранее, до фактического столкновения с ней. Это положительно скажется на скорости и надежности работы высоконагруженного web-сервиса, вне зависимости от его предназначения.

Заключение

В настоящее время анализируется статистика нагрузки системы дистанционного обучения iSpring Learn для поиска закономерностей и выявления предпосылок высокой нагрузки.

Но уже сейчас можно говорить об острой практической необходимости в средстве прогнозирования, имеющем способность принимать решения, основываясь на данных от запросов клиентов СДО и данных о загруженности, полученных от серверов.

Библиографический список

- [iSpring Learn, 2014] Система дистанционного обучения iSpring Learn, 2014, <http://www.ispringsolutions.com/ispring-learn>
- [Блинков, 2010] Архитектура Facebook / Блинков И.; 2010 – <http://www.insight-it.ru/masshtabiruemost/arkhitektura-facebook/>
- [Комлев, 2006] Словарь иностранных слов / Комлев Н.Г.; 2006
- [GoogleAnalytics, 2014] Источники трафика/ Google.; 2014 - <https://support.google.com/analytics/answer/1033173?hl=ru>
- [GoogleTrends, 2014] Сервис анализа частоты запросов в поисковой системе Google/ Google.; 2014 - <https://www.google.ru/trends/>
- [GigamResearch, 2013] Facebook's first European data-center goes live in Sweden/David Meyer.; 2014 - <https://gigaom.com/2013/06/12/facebooks-first-european-data-center-goes-live-in-sweden/>

SOLUTION TO THE PROBLEM OF PREDICTING LOAD INCREASES IN LEARNING MANAGEMENT SYSTEM "ISPRING LEARN"

Sokolnikov A.M., Sidorkina I.G.

*Volga State University of Technology, Mari El,
Russian Federation*

sokolnikov.alexey@gmail.com

igs592000@mail.ru

The article examines ways to predict high load web-services. Justify of their use in learning management system iSpring Learn. Reveals the problems that can be solved using the methods of forecasting load increases and benefits of their usage in conjunction with a load balancer.

Introduction

To solve the problems with high load web-services most developers used load balancers. The problem with this approach is that the service is ready to be a sharp increase in load. This situation will lead to delays in query execution and increase the load on a particular server. As a solution to this problem is proposed use of preventive means of analysis of high loads..

Tool for preventive load analysis

Increase in load forecasting methods analyze and process requests to the server and store their metadata. In addition, they poll current state of the server load. Based on these data, it searches for patterns of occurrence of high loads and a decision on the need to deploy additional instances of the service.

The advantage of this approach is that the use of forecasting methods of increasing the load balancer in conjunction with, the number of peaks queries on a single server is significantly reduced because of an emergency situation will be known in advance of its actual occurrence. Such methods will be called preventive load analysis.

Proactive analysis processes the data in a query and tries to identify patterns in the surge loads. There are four main types of dependencies that affect the load: calendar, time, location-based and depending on the sources of traffic.

Conclusion

With the tool, predicts the emergence of high loads, will be realized the possibility of solving the problem of a large number of requests in advance of the actual collision with it. This has a positive impact on the speed and reliability of a heavily web-service, regardless of its destination.