**СПОСОБЫ ОЦЕНКИ ТРЕБОВАНИЙ К АППАРАТНОМУ**

**ОБЕСПЕЧЕНИЮ ПРИ ПРОЕКТИРОВАНИИ СЕРВЕРА**

**ДЛЯ ПОТОКОВОГО ВИДЕОСЕРВИСА**

*Пукита М.Г.,*

*научный руководитель Шакирзянова А.А.,*

*ассистент кафедры АСОИУ*

Казанский национальный исследовательский

технический университет им. А.Н. Туполева – КАИ,

г. Казань, РТ

Оценка ожидаемых требований к аппаратному обеспечению для сервера потокового видеосервиса – важный процесс, который включает в себя анализ различных факторов для обеспечения оптимальной производительности и надежности. Поскольку потоковые сервисы продолжают набирать популярность, понимание базовых аппаратных потребностей становится все важнее для обеспечения высокого качества обслуживания пользователей. Проблемы, связанные с оценкой требований к оборудованию, включают изменчивость пользовательского спроса, ограниченность ресурсов и необходимость точного прогнозирования.

При прогнозировании требований к оборудованию для сервера потокового веб-сервиса необходимо учитывать несколько ключевых факторов для обеспечения оптимальной производительности и высокой надежности. Эти факторы включают в себя ресурсы сервера, пропускную способность, решения для хранения данных и масштабируемость, каждый из которых играет важную роль в общей инфраструктуре потокового сервиса [1].

*Ресурсы сервера*

Центральный процессор (ЦП) и оперативная память – основополагающие компоненты, которые существенно влияют на производительность сервера. Увеличение объема памяти может повысить производительность за счет уменьшения подкачки и обеспечения более быстрого доступа к данным, а такие конфигурации, как двухканальные или четырехканальные архитектуры, обеспечивают лучшую пропускную способность для приложений, интенсивно использующих память [1, 3].

*Емкость и тип хранилища данных*

Решения для хранения данных играют ключевую роль в определении способности сервера эффективно управлять данными. Организации часто используют внешние системы хранения данных, но внутренняя система хранения также должна соответствовать прогнозируемым потребностям. Выбор между твердотельными (SSD) и жесткими дисками (HDD) влияет как на скорость, так и на емкость, причем SSD обычно обеспечивают более высокую скорость чтения и записи. Гибридный подход, сочетающий оба типа хранилищ, позволяет оптимизировать производительность для часто используемых данных и сохранить экономическую эффективность для больших объемов. Кроме того, использование RAID-контроллеров может повысить защиту данных и производительность за счет дублирования хранящейся информации.

*Учет пропускной способности сети*

Пропускная способность сети – критический элемент, влияющий на количество одновременных зрителей, которое может поддерживать сервер при потоковой передаче видео. Недостаточная пропускная способность может привести к буферизации и ухудшению качества работы пользователей, особенно для контента высокой четкости [1, 2].

*Масштабируемость и будущий рост*

Масштабируемость – важнейший фактор при выборе серверного оборудования, поскольку она определяет способность системы адаптироваться к растущим требованиям с течением времени. Планирование серверной инфраструктуры с расчетом на будущее путем инвестирования в оборудование, совместимое с новейшими технологиями и стандартами, такими как PCIe 4.0 или PCIe 5.0, может способствовать более плавному переходу по мере развития потребностей [4].

Существуют различные методы непосредственной оценки системных требований. Рассмотрим некоторые из них.

*Анализ профилей нагрузки*

Определение типов потоков (например, простые стримы или потоки с транскодированием) играет большую роль в выборе последующих действий при проектировании. Потоки с транскодированием расходуют значительно больше ресурсов ЦП и ОЗУ, чем простые потоки. Важно помнить, что для обеспечения стабильного качества обслуживания нагрузка на сервер не должна превышать 80% [2].

*Расчет пропускной способности сети*

Оценка необходимой пропускной способности проводится на основе ожидаемого числа зрителей и допустимого качества видео. Например, для трансляции в качестве 480p потребуется около 1 Mbps на один поток. Соответственно, для 1000 аналогичных трансляций потребуется 1000 Mbps [2].

*Нагрузочное тестирование*

Нагрузочное тестирование с применением таких инструментов, как Prometheus и Grafana, позволяет отслеживать метрики производительности сервера с различным количеством подключенных зрителей. На основе результатов такого тестирования можно принять решение о необходимых параметрах [1, 2].

*Аппаратное ускорение*

Нагрузку на центральный процессор можно снизить, включив аппаратное ускорение шифрования (например, инструкции AES), тем самым повысив производительность обработки данных.

*Моделирование рабочих нагрузок*

Модели очередей и методы машинного обучения для оценки потребностей в ресурсах на основе исторических данных о нагрузках и производительности системы могут быть результативным решением для серверов, ожидающих большой нагрузки со стороны пользователей [1].

*Предиктивное моделирование*

Предиктивное моделирование может быть использовано для прогнозирования будущих потребностей в ресурсах на основе анализа текущих данных о производительности и загруженности системы путем предположения о кратном увеличении нагрузки [4].

Из всего вышеизложенного можно сделать вывод, что оценка требований к аппаратному обеспечению для сервера потокового веб-сервиса – это важный процесс в условиях растущего спроса на высококачественные услуги потоковых трансляций. Правильное понимание ключевых концепций – производительности, надежности, пропускной способности и гибридных моделей хранения данных – позволяет организациям более эффективно планировать использование своих ресурсов и адаптироваться к изменяющимся условиям рынка. Использование современных методов анализа и прогнозирования, включая предиктивное моделирование и методы исследования операций, способствует улучшению качества обслуживания пользователей и повышению конкурентоспособности на рынке потоковых услуг.

1. Ardagna, D., Casale, G., Ciavotta, M. *и др.* Quality-of-service in cloud computing: modeling techniques and their applications. *// Internet Serv. Appl.* **5**, 11 (2014). – [Электр.ресурс]. Режим доступа: https://jisajournal.springeropen.com/articles/10.1186/s13174-014-0011-3 (дата обращения: 30.10.2024).
2. Подбираем сервер для 1000 WebRTC стримов // Блог компании Flashphoner (Habr) (2021). – [Электр.ресурс]. Режим доступа: https://habr.com/ru/companies/flashphoner/articles/570266/ (дата обращения: 1.11.2024).
3. How should I choose a new server hardware configuration? // FS Knowledge Center (2024). – [Электр.ресурс]. Режим доступа: https://community.fs.com/article/how-should-i-choose-a-new-server-hardware-configuration.html (дата обращения: 30.10.2024).
4. Nikita S., Track 11 Server Performance Metrics and Optimize Server Health // CloudPanel (2023). – [Электр.ресурс]. Режим доступа: https://www.cloudpanel.io/blog/server-performance-metrics/ (дата обращения: 30.10.2024).