

Вариант 8. Федоренко Екатерина. 22505

Задача 1. Сделайте общее описание известной вам информационной или вычислительной системы. Для выбранной системы составьте описание нагрузки, предложите прогноз ее эволюции. Сформулируйте для этой системы задачу планирования мощности.

Компания имеет облачное решение «Сайты УК», которое предлагает управляющим компаниям приобрести сайты «Под ключ» - основной функционал уже разработан, и включает в себя создание закрытых личных кабинетов пользователей – жильцов и общую витрину для всех желающих с основной информацией о компании.

В личных кабинетах жильцов существует возможность ввести показания счетчиков, оплатить квитанции и многое другое. Облачное решение подразумевает хранение на единственном сервере, общую базу данных и общую память для всех сайтов. Общее число сайтов в настоящий момент равно 45. В каждом из них в среднем 100 личных кабинетов. То есть сейчас 4500 пользователей.

Одновременно сервер может обрабатывать 12000 сессий, допустимое время отклика равняется 1 секунде.

Руководством компании принято решение о привлечении клиентов со всей страны. Предполагается, что число пользователей увеличится на 60%.

Необходимо определить, справится ли облачное решение с такой нагрузкой, и если нет, то придумать варианты решения проблемы.

У провайдера облачного решения есть следующие пути:

1. Добавить мощности к текущему серверу
2. Модернизировать архитектуру системы так, чтобы была возможность использовать дополнительный сервер

Задача 2. Для выбранной вами в задаче 1 системы постройте ее подробное общее описание и предложите модель нагрузки. Составьте административный план развития системы и на его основе уточните составленный вами прогноз эволюции модели нагрузки.

Система состоит из сервера, обслуживающего сессии клиентов, и входной очереди сессий. Активным устройством является сервер, его основная характеристика – число одновременно поддерживаемых пользовательских сессий, 12000.



До привлечения клиентов:

- Загрузка составляет: $4500 / 12000 = 37,5\%$
- Пропускная способность: 4500 tps
- Средняя длина очереди: $4500 / (12000 - 4500) = 0,6 \text{ t}$
- Время отклика: $0,6 / 12000 = 0,00005 \text{ s}$

После привлечения новых клиентов число пользователей станет равным $4500 * 1,6 = 7200$

- Загрузка составит: $5850 / 12000 = 60 \%$
- $60\% < 100\% \Rightarrow$ мощности существующего сервера достаточно, дополнительные меры принимать не нужно.

Административный план развития системы будет осуществляться в направлении модернизации информационной системы вследствие ежегодного увеличения числа организаций – управляющих компаний примерно на 100 единиц (в информационную систему будут добавляться новые регионы РФ).

Очевидно, если сервер не модернизировался с начальной конфигурации, то его производительности в пиковую нагрузку хватать не будет, следовательно, производительность необходимо увеличивать. Есть несколько вариантов усовершенствования:

1. Модернизация облачного сервера (увеличение производительности минимум в 2 раза, проведение данной операции раз в несколько лет);
2. Приобретение нового (с бóльшей производительностью) или дополнительного сервера;

Задача 3. Для выбранной вами в задаче 1 системы определите основные параметры модели производительности; сформулируйте обоснованную спецификацию измерений, необходимых для их идентификации; укажите наиболее подходящий для этого монитор.

Основным измеряемым параметром для сервиса будет количество активных сессий в единицу времени. Необходимо использовать программный монитор, с определенными интервалами времени регистрирующий число активных сессий. Монитор будет работать в режиме выборки, при этом интервал между сессиями мониторинга будет минимизирован, чтобы не потерять данные о редких событиях.

Задача 4. Рассмотрим однопроцессорную модель, описанную выше:

Исходная производительность системы:

Интенсивность потока транзакций: $\lambda = 30 \text{ tps}$

Интенсивность обработки заявок: $\mu = 50 \text{ tps}$

Загрузка системы: $U = \frac{\lambda}{\mu} = 60\%$

Пропускная способность: $T = \lambda = 30 \text{ tps}$

Средняя длина очереди: $\overline{n_q} = \frac{\lambda}{\mu - \lambda} = 1.5 \text{ t}$

$$\text{Время отклика: } R = \frac{1}{\mu - \lambda} = 0.05 \text{ s}$$

а) Как изменятся метрики производительности системы, если ее нагрузка увеличится в $N/10$ раз, а мощность центрального процессора в $N/10 + 1$ раз, где

$N = 1 \dots 25$ — номер варианта?

$$\Delta\lambda = \frac{N}{10} = \frac{8}{10} = 0.8$$

$$\Delta\mu = \frac{N}{10} + 1 = 1.8$$

Новые значения интенсивности:

$$\lambda = 1.8 * 30 = 54 \text{ tps}$$

$$\mu = 2.8 * 50 = 140 \text{ tps}$$

Новые значения метрик:

$$U = \frac{\lambda}{\mu} = \frac{54}{140} \approx 39\%$$

$$T = \lambda = 54 \text{ tps}$$

$$\overline{n_q} = \frac{\lambda}{\mu - \lambda} = \frac{54}{140 - 54} \approx 0.63 \text{ t}$$

$$R = \frac{1}{\mu - \lambda} \approx 0.0116 \text{ s}$$

Вывод: интенсивность поступления заявок выросла в 1.8 раза, интенсивность их обработки – в 2.8 раза; загрузка снизилась на 21%, пропускная способность выросла в 1.8 раза, средняя длина очереди сократилась на 0.87 транзакцию, время отклика снизилось практически в 4,3 раза.

б) Пусть p % пользователей системы начали использовать новое приложение и больше не обращаются к системе. Администрация решила установить менее мощный центральный процессор. Какой должна быть минимальная мощность нового процессора для того, чтобы среднее время отклика системы не превысило $1/20$ с ($p = 3N$)?

$$p = 3N = 3 * 8 (\%) = 24\% = 0.24$$

Новое значение интенсивности поступления запросов:

$$\lambda = (1 - 0.24) * 30 \approx 22.8 \text{ tps}$$

Время отклика должно быть не более 0.05 с:

$$R = \frac{1}{\mu - \lambda} \leq 0.05 \text{ s}$$

$$0.05(\mu - \lambda) \geq 1 \leftrightarrow \mu \geq 42,8 \text{ tps}$$

Новый процессор должен обеспечивать обработку не менее 42,8 транзакций в минуту.

в) Предположим, что центральный процессор системы был усилен на q %. Какое минимальное увеличение нагрузки приведет к тому, что среднее время отклика превысит $1/20$ с ($q = 4N$)?

$$q = 4N = 4 * 8 (\%) = 32\% = 0.32$$

Новое значение интенсивности обработки запросов:

$$\mu = (1 + 0.32) * 50 \approx 66 \text{ tps}$$

Время отклика должно превысить 0.05 с:

$$R = \frac{1}{\mu - \lambda} > 0.05 \text{ s}$$

$$0.05(\mu - \lambda) < 1 \leftrightarrow \lambda > 46 \text{ tps}; \min \lambda = 47 \text{ tps}$$

Чтобы время обработки транзакции усиленным процессором стало больше 0.05 с, необходимо поступление не менее 47 транзакций в минуту.

Задача 5. Рассмотрим модель центрального сервера:

Исходные данные:

Средние интенсивности обслуживания транзакций:

- на ЦП $\mu_c = 6 \text{ tpm}$;
- на медленном диске $\mu_s = 2 \text{ tpm}$;
- на быстром диске $\mu_f = 3 \text{ tpm}$;

Маршрутизация транзакций м/ду дисками производится с вероятностями p и $1 - p$, $p = 1/2$.

Уровень мультипрограммирования равен 2.

$$U_f = \frac{\mu_c p}{\mu_f} = \frac{6 * 0.5}{3} = 1$$

$$U_s = \frac{\mu_c (1 - p)}{\mu_s} = \frac{6 * 0.5}{2} = 1.5$$

$$P_{200} = (1 + U_f + U_s + U_f^2 + U_s^2 + U_f U_s)^{-1} \approx 0.12$$

$$P_{110} = U_f P_{200} = 0.12$$

$$P_{020} = U_f^2 P_{200} = 0.12$$

$$P_{011} = U_s U_f P_{200} = 0.18$$

$$P_{101} = U_s P_{200} = 0.18$$

$$P_{002} = U_s^2 P_{200} = 0.27$$

Быстрый диск:

$$U^f = P_{110} + P_{020} + P_{011} = 0.42 = 42\%$$

$$T^f = \mu_f U^f = 1.26 \text{ tpm}$$

$$N_q^f = 1 * P_{110} + 2 * P_{020} + 1 * P_{011} = 0.54 \text{ t}$$

$$R^f = \frac{N_q^f}{T^f} \approx 0.43 \text{ min}$$

Медленный диск:

$$U^s = P_{101} + P_{011} + P_{002} = 0.63 = 63\%$$

$$T^s = \mu_s U^s = 1.26 \text{ tpm}$$

$$N_q^s = 1 * P_{101} + 1 * P_{011} + 2 * P_{002} = 0.9 \text{ t}$$

$$R^s = \frac{N_q^s}{T^s} \approx 0.71 \text{ min}$$

Центральный процессор:

$$U^c = P_{200} + P_{110} + P_{101} = 0.42 = 42\%$$

$$T^c = \mu_c U^c = 2.52 \text{ tpm}$$

$$N_q^c = 2 * P_{200} + 1 * P_{110} + 1 * P_{101} = 0.54 \text{ t}$$

$$R^c = \frac{N_q^c}{T^c} \approx 0.21 \text{ min}$$

	U	T	N_s	R
Быстрый диск	42%	1.26 tpm	0.54 t	0.43 min
Медленный диск	63%	1.26 tpm	0.9 t	0.71 min
Процессор	42%	2.52 tpm	0.54 t	0.21 min

в) На сколько изменится производительность системы, если производительность процессора возрастет на p %, а q % файлов будут перемещены с медленного диска на быстрый ($p = 4N$, $q = 2N$)?

$$p_{CPU} = 4N = 4 * 8(\%) = 32\% = 0.32$$

$$q = 2N = 2 * 8(\%) = 16\% = 0.16$$

Новая производительность процессора:

$$\mu_c = 1.32 * 6 = 7.92 \text{ tpm}$$

Вероятность обращения к «быстрому» диску после перемещения файлов:

$$p = 0.5 + q = 0.66$$

$$U_f = \frac{\mu_c p}{\mu_f} = \frac{7.92 * 0.66}{3} \approx 1.74$$

$$U_s = \frac{\mu_c (1 - p)}{\mu_s} = \frac{7.92 * 0.34}{2} \approx 1.35$$

$$P_{200} = (1 + U_f + U_s + U_f^2 + U_s^2 + U_f U_s)^{-1} \approx 0,0886$$

$$P_{110} = U_f P_{200} = 0.154$$

$$P_{020} = U_f^2 P_{200} = 0.27$$

$$P_{011} = U_s U_f P_{200} = 0.208$$

$$P_{101} = U_s P_{200} = 0.12$$

$$P_{002} = U_s^2 P_{200} = 0.16$$

Быстрый диск:

$$U^f = P_{110} + P_{020} + P_{011} = 0.63 = 63\%$$

$$T^f = \mu_f U^f = 1.89 \text{ tpm}$$

$$N_q^f = 1 * P_{110} + 2 * P_{020} + 1 * P_{011} = 0.902 \text{ t}$$

$$R^f = \frac{N_q^f}{T^f} \approx 0.48 \text{ min}$$

Медленный диск:

$$U^s = P_{101} + P_{011} + P_{002} = 0.53 = 53\%$$

$$T^s = \mu_s U^s = 1.06 \text{ tpm}$$

$$N_q^s = 1 * P_{101} + 1 * P_{011} + 2 * P_{002} = 0.65 \text{ t}$$

$$R^s = \frac{N_q^s}{T^s} \approx 0.61 \text{ min}$$

ТТХ процессора:

$$U^c = P_{200} + P_{110} + P_{101} = 0.36 = 36\%$$

$$T^c = \mu_c U^c = 2.85 \text{ tpm}$$

$$N_q^c = 2 * P_{200} + 1 * P_{110} + 1 * P_{101} = 0.6 \text{ t}$$

$$R^c = \frac{N_q^c}{T^c} \approx 0.21 \text{ min}$$

	U	T	N_s	R
Быстрый диск	63%	1.89 tpm	0.902 t	0.48 min
Медленный диск	53%	1.06 tpm	0.65 t	0.61 min
Процессор	36%	2.85 tpm	0.6 t	0.21 min

Табличка с сравнением:

	U	T	N_s	R
Быстрый диск	↑ на 21%	↓ на 0,6 tpm	↑ на 0,362t	↑ на 0.05 min
Медленный диск	↓ на 10%	↓ на 0,2 tpm	↓ на 0.25 t	↓ на 0.1 min
Процессор	↓ на 6%	↑ на 0,33 tpm	↑ на 0,06 t	Не изм

Благодаря ускорению процессора и перемещению файлов удалось добиться более равномерной загрузки обоих дисков.