



**Отчет по результатам
Нагрузочного тестирования**

СОДЕРЖАНИЕ

1	Список сокращений и терминов	3
2	Назначение документа.....	3
3	Основные положения.....	3
3.1	Цели тестирования	3
3.2	Отступления от методики тестирования.....	4
3.3	Ограничения тестирования.....	4
3.4	Выводы	4
4	Рекомендации.....	5
5	Результаты тестирования	6
5.1	Подтверждение выводов.....	6
5.2	Тестирование на промышленной базе данных	12
5.3	Тестирование на увеличенной базе данных.....	17

1 Список сокращений и терминов

Сокращение/термин	Определение
НТ	Нагрузочное тестирование.
СНТ	Средства нагрузочного тестирования.
КТС	Комплекс технических средств.
ПО	Программное обеспечение.
ОС	Операционная система.
БД	СУБД – система управления базами данных.
АС	Автоматизированная система.
Внешние (или сопряженные) системы	Автоматизированные системы, обрабатывающие соответствующие запросы, получаемые из шины.
LoadRunner	Средство нагрузочного тестирования компании «Hewlett-Packard» (HP).
, (в числах)	Разделитель групп разрядов числа.
. (в числах)	Разделитель целой и дробной части числа.

2 Назначение документа

Основная цель данного документа – предоставить обработанные и систематизированные результаты нагрузочного тестирования АС, описать отклонения при проведении тестирования от методики и ограничения тестирования.

3 Основные положения

3.1 Цели тестирования

Основной целью проведения нагрузочного тестирования АС является оценка возможностей системы. Для определения производительности АС планируется:

- Определение максимальной производительности (количество запросов/час) АС, на существующей конфигурации (в соответствии с настроенными правилами).

- Проверка надежности АС во время тестирования в течение длительного времени.
- Определить отказоустойчивость АС при отказе одной из площадок серверов приложения (останов 2-х из 4-х серверов приложения)
- Проверка производительности АС при увеличении объема БД (рассчитывается на ближайший год). Определение зависимости между производительностью АС и размером БД.
- Выявление потенциально «узких» мест АС.

3.2 Отступления от методики тестирования

Принципиальных отступлений от Методики тестирования не было. По аналогии с работой промышленной среды, запросы к АС отправлялись без задержек с максимально возможной интенсивностью (ограничением интенсивности запросов была только производительность самой АС). Для второго длительного теста на 24 часа был изменен состав серверов приложений, а именно, увеличен до 5 (в Методике в составе КТС для НТ было определено 3 сервера приложений). Данный длительный тест и оказался успешным.

3.3 Ограничения тестирования

Не было зафиксировано.

3.4 Выводы

- Максимальная производительность АС на существующей конфигурации равна **10 соединений / 3688 обработанных запросов/час** на объеме текущей промышленной базы данных. Критерием для определения максимальной производительности в данном случае является появление значительного количества ошибок (количество неуспешных транзакций превышает 10 %), кроме того заметна тенденция к исчерпанию аппаратных ресурсов сервера.
- Длительный тест надежности (на 24 часа) с использованием планового КТС (5 сервера приложений по Методике) является неуспешным по тем же

критериям: значительное количество ошибок, приводящих к отказу обслуживания, кроме того также исчерпаны аппаратные ресурсы сервера. При добавлении дополнительных 2 серверов приложений (суммарное количество серверов приложений стало равно 7) тест завершается успешно с результатом **6 соединений/ 62695 обработанных запросов/ 24 часа**.

- Тестирование на отказоустойчивость показало нестабильность развития ситуации после отключения 4 серверов приложений (50% ресурса), а именно, оставшиеся 2 сервера приложения могут не справиться с нагрузкой. Критерии признания эксперимента неуспешным все те же, что и в предыдущих случаях: значительное количество ошибок, приводящих к отказу обслуживания, кроме того происходит значительное расходование выделенной памяти (исчерпаны аппаратные ресурсы сервера).
- Максимальная производительность АС с базой данных увеличенного объема равна **6 соединений/ 3048 обработанных запросов/час**. Критериями для определения максимальной производительности является появление значительного количества ошибок (количество неуспешных транзакций достигает примерно 6 %). Времена отклика слабо зависят от нагрузки при этом значение самой длительной транзакции (OnlineMatching) увеличились до 7,5 секунд в среднем (до 14,1 по критерию 90 перцентилей). Предполагается что в этом эксперименте «узким местом» становится сервер базы данных.
- Под нагрузкой происходит увеличение объема выделенной (Committed Memory) памяти, а также памяти исключительного использования (Private Bytes) для процессов на серверах приложений, что может указывать на «утечку» памяти.
- Если «утечка» памяти имеет место, то в этом случае она является блокирующим условием на пути определения максимальной производительности, так как полученные в относительно кратковременных экспериментах (определения/подтверждения) результаты могут ниже в длительной промышленной эксплуатации АС.

4 Рекомендации

- С помощью разработчиков разобраться в том, имеют ли действительно место «утечки памяти» на серверах приложений АС. Влияние этой проблемы может быть таковым, что его сложно компенсировать даже приростом аппаратных ресурсов.

5 Результаты тестирования

Раздел «Результаты тестирования» состоит из трех подразделов. В первом подразделе приводятся результаты/данные подтверждающие выводы в разделе 3. Во втором и третьем подразделах иллюстрируются все эксперименты, проведенные в данном тестировании вместе с анализом.

5.1 Подтверждение выводов

5.1.1 Определение максимальной производительности на текущей промышленной базе данных

Прирост ошибок на фоне увеличения нагрузки иллюстрируется рисунком 5.1.1

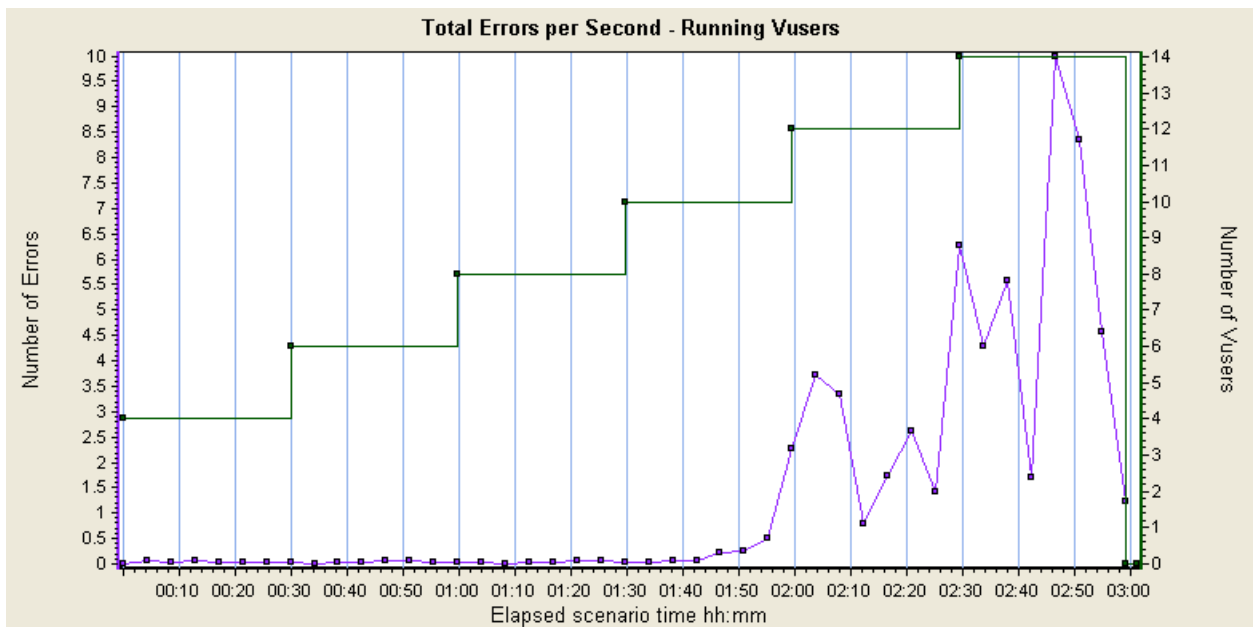


Рисунок 5.1.1 Нарастание нагрузки и интенсивность возникновения ошибок

Таблица 5.1.1

Название транзакций со статусом «Failed»	Кол -во	% ¹
OnlineMatching	2 075	15,4
onlineMatchingChange	146	12,6
updateMetadata	1 234	14,0

5.1.2 Проведение длительного теста надежности

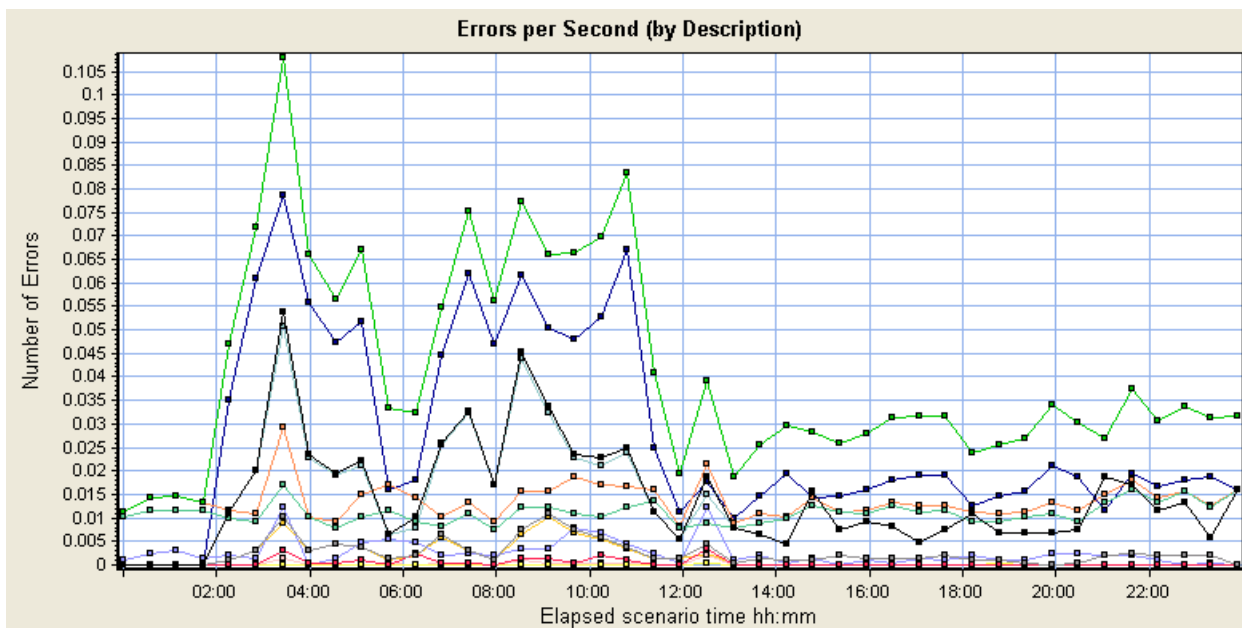


Рисунок 5.1.3 Интенсивность возникновения ошибок

¹ Рассчитывался по соотношению ошибочных транзакций к общему количеству транзакций, так что процент ошибок на «ступенях» где они появились еще больше.



Рисунок 5.1.4. Мониторинг ресурсов сервера приложений 1

Рисунки 5.1.3 и 5.1.4 подтверждают большое количество ошибок (первые 12 часов эксперимента) совместно с ростом показателей использования памяти (коричневая, синяя и зеленая линии). Предполагается, что именно «съедание» памяти процессами является причиной возникновения ошибок. Более подробно эксперименты по проведению длительного тестирования описаны в подразделе 5.2.4.

5.1.3 Проведение тестирования отказоустойчивости

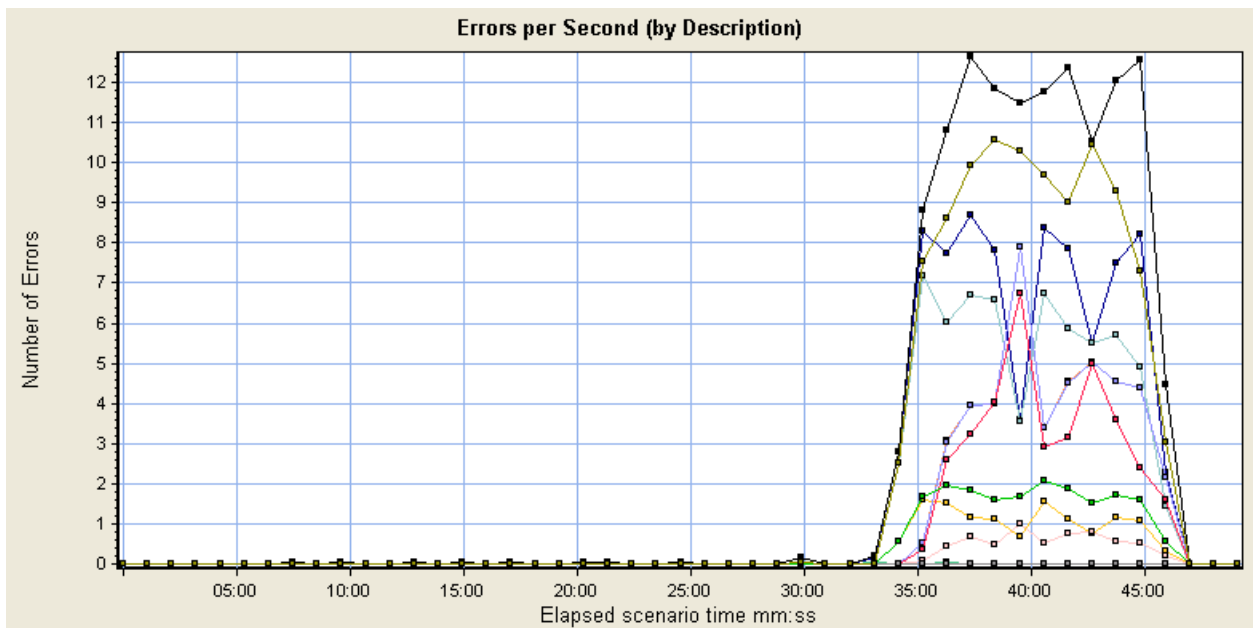


Рисунок 5.1.5 Интенсивность возникновения ошибок после отключения 2х серверов приложений.

Рисунок 5.1.5 иллюстрирует отказ двух оставшихся в работе серверов. Более подробно эксперименты отказоустойчивости описаны в подразделе 5.2.3.

5.1.4 Определение максимальной производительности на увеличенной базе данных

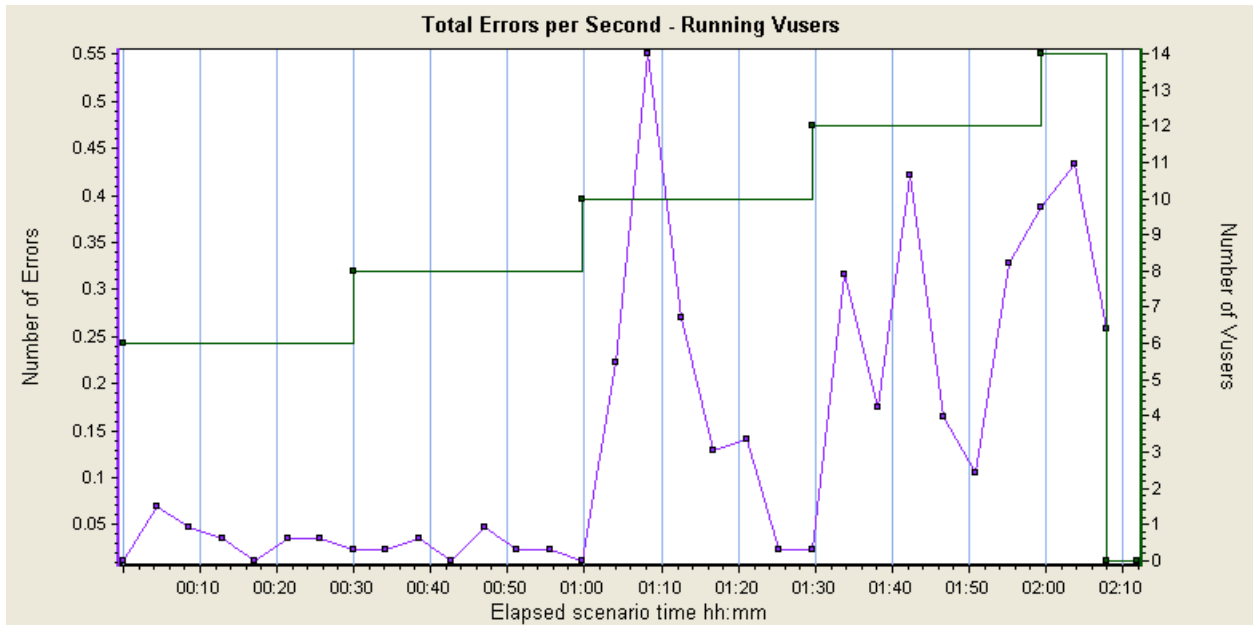


Рисунок 5.1.6

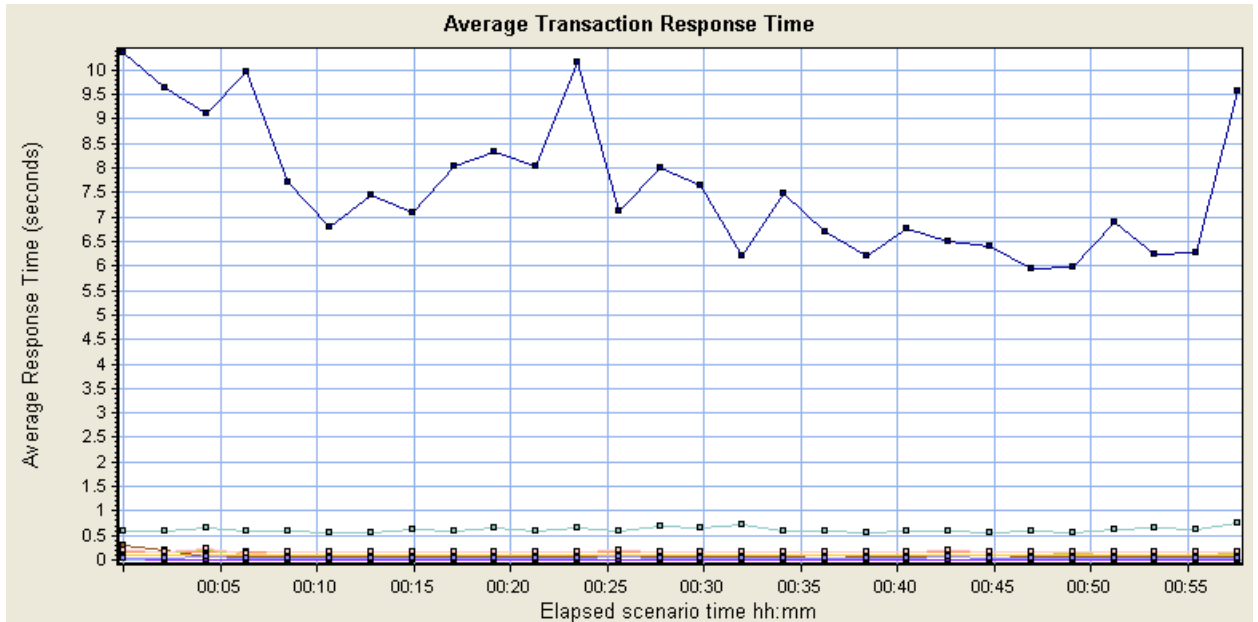


Рисунок 5.1.7.

Времена отклика (OnlineMatching) в тестировании на увеличенной базе данных выросли примерно в 2-2,5 раза, на что видимо, повлияло увеличение

времени обработки запросов непосредственно СУБД. Анализ результатов мониторинга сервера БД показывает, что время обслуживания дисковой подсистемы временами достигает значений 100 и более миллисекунд (Таблица 5.1.2). Увеличившийся под нагрузкой отклик дисковой подсистемы вполне может приводить к увеличению времен отклика по транзакции OnlineMatching.

Таблица 5.1.2

Дата время	Диск	Время сервиса		
02/16/2012 11:57:00	Disk#	22.07	23.3	1.66
02/16/2012 11:58:00	Disk##	7.13	7.21	6.66
02/16/2012 11:58:00	Disk*	102.58	106.04	3.33
02/16/2012 11:59:00	Disk##	8.17	8.29	3.33
02/16/2012 11:59:00	Disk*	98.85	101.66	55
02/16/2012 11:59:00	c0t0d0	0	0	0
02/16/2012 12:00:00	Disk##	7.15	7.53	3.52
02/16/2012 12:00:00	Disk*	75.28	83.77	24
02/16/2012 12:01:00	Disk##	20.44	20.62	5
02/16/2012 12:01:00	Disk*	70.99	71.78	58.33
02/16/2012 12:02:00	Disk##	7.22	7.31	6.66
02/16/2012 12:02:00	Disk*	30.5	33.33	3.33
02/16/2012 12:03:00	Disk##	6.96	6.97	3.33

5.1.5 Предположение о возможной утечке памяти

Во всех экспериментах (под постоянной и нарастающей нагрузкой) было отмечено значительное увеличение выделяемой памяти (\Committed Bytes), памяти выделенной в исключительное пользование процессам (\Private Bytes) и памяти образующей рабочие наборы процессов (\Working Set).

Таблица 5.1.3

Тип счетчика	Минимум	Максимум
\Committed Bytes	573157376	1657032704
\Private Bytes	238751744	1299607552
\Working Set	341467136	1411055616

В таблице 5.1.3 приводятся данные по длительному неуспешному эксперименту проверки надежности. Видно, что значения используемой памяти вырастают в 3-5 раз в течение эксперимента. Примерно такая же динамика потребления памяти для всех неуспешных экспериментов данного тестирования.

5.2 Тестирование на промышленной базе данных

5.2.1 Тест определения максимальной производительности (№6)

В соответствии с Методикой первым был проведен эксперимент по достижению максимального уровня производительности. Нагрузка увеличивалась путем добавления 2 новых соединений (виртуальных пользователей) каждые 30 минут. Начиная с нагрузки равной 10 соединениям (виртуальным пользователям) АС начинает возвращать ошибки, говорящие об отказах обслуживания запросов «Expected SOAP result, received SOAP fault», «Failed to retrieve output arguments/checkpoints - SOAP fault occurred».

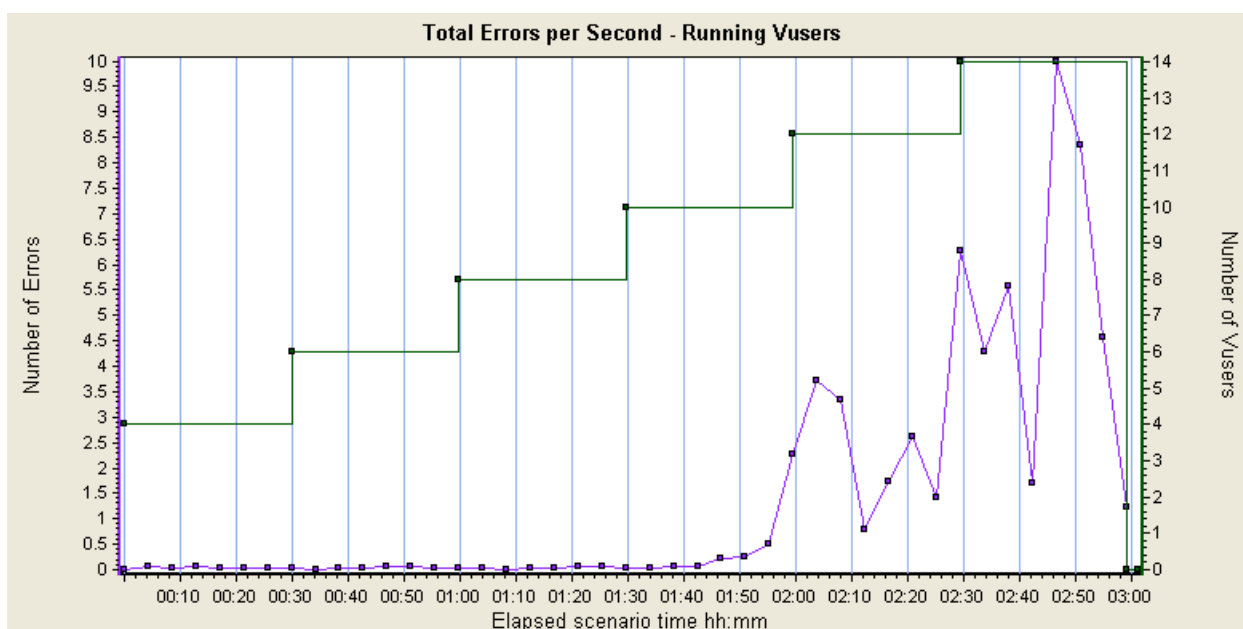


Рисунок 5.2.1 Увеличение нагрузки (зеленый график) и интенсивность возникновения ошибок (сиреневый график)

Рисунок 5.2.1 показывает появление значительного количества ошибок (в том числе и по отказу обслуживания) совместно с ростом нагрузки. При этом времена отклика не ухудшаются и примерно одинаковы на протяжении всего теста (Рисунок 5.2.).

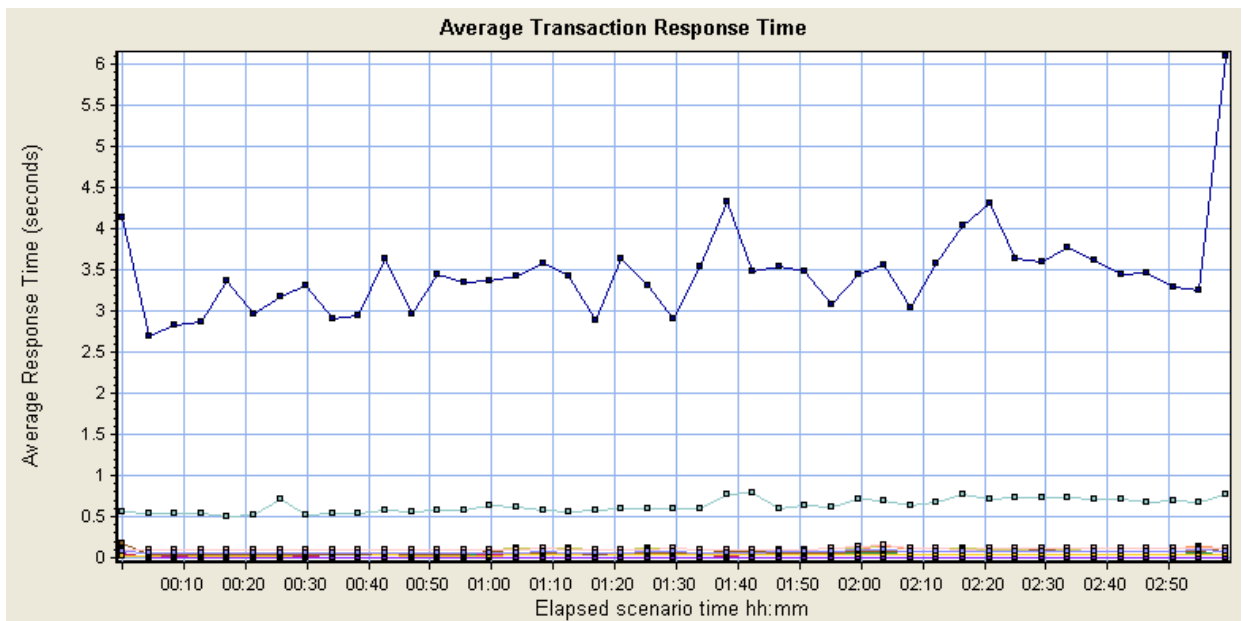


Рисунок 5.2.2 Времена отклика по операциям в течение эксперимента определения максимума производительности.

5.2.2 Тест подтверждения максимальной производительности (№7)

Анализируя выполненный тест, в качестве нагрузки подтверждения максимальной производительности выбирается нагрузка, состоящая из 8 соединений, отправляющих запросы к АС с максимально возможной интенсивностью. По результатам тестирования определяется максимально возможная производительность АС на данном оборудовании, выраженная в количестве запросов в час.

Таблица 5.2.1.

Название запроса	Кол-во/час
OnlineMatching	6020
onlineMatchingChange	480
updateMetadata	2343

Итого, за час успешно было выполнено 8843 запроса, что и является верхней границей производительности тестируемой версии АС. Количество запросов в каждой категории представлено в таблице 5.2.1.



Рисунок 5.2.3. Мониторинг сервера приложений 3

Мониторинг серверов приложений в течение 1 часа не выявляет серьезных проблем. Несколько настораживает довольно значительный рост выделяемой памяти на трех серверах приложений из четырех (график желтого цвета на рисунке 5.2.3, сервер 3 взят для примера).

Таблица 5.2.2

Название сервера	Min (выделенной памяти, bytes)	Max (выделенной памяти, bytes)
1	550801408	1355800576
2	553017344	727826432
3	681259008	1710649344
4	653373440	1771896832

Анализ таблицы 5.2.2 показывает, что на 3х серверах приложений объем выделенной (committed) памяти вырос в течение теста в 2 с лишним раза. Столбец «Min (выделенной памяти, bytes)» показывает значения памяти в начале теста, а столбец «Max (выделенной памяти, bytes)» показывает значения выделенной памяти в конце теста.

5.2.3 Тест отказоустойчивости

В процессе тестов выполнялось отключение (останов) 2х из 4х серверов приложений с целью понять реакцию АС на отказ. Было проведено 2 эксперимента с порядковыми номерами №8 и №9, в обоих экспериментах нагрузку создавали 5 соединений (виртуальных пользователей) отправляющие запросы на АС с максимальной интенсивностью. В эксперименте №8 оставшиеся два сервера не смогли «удержать» сконцентрированную на них нагрузку и самопроизвольно прекратили обслуживание запросов. Было принято решение повторить тест с теми же условиями. В эксперименте №9, система (АС) смогла обслуживать запросы, имея в своем составе только 2 сервера приложений, и, при возвращении в строй первых 2х, продолжила обработку запросов, таким образом, продемонстрировав отказоустойчивость.

Неуспешный эксперимент Тест (№8)

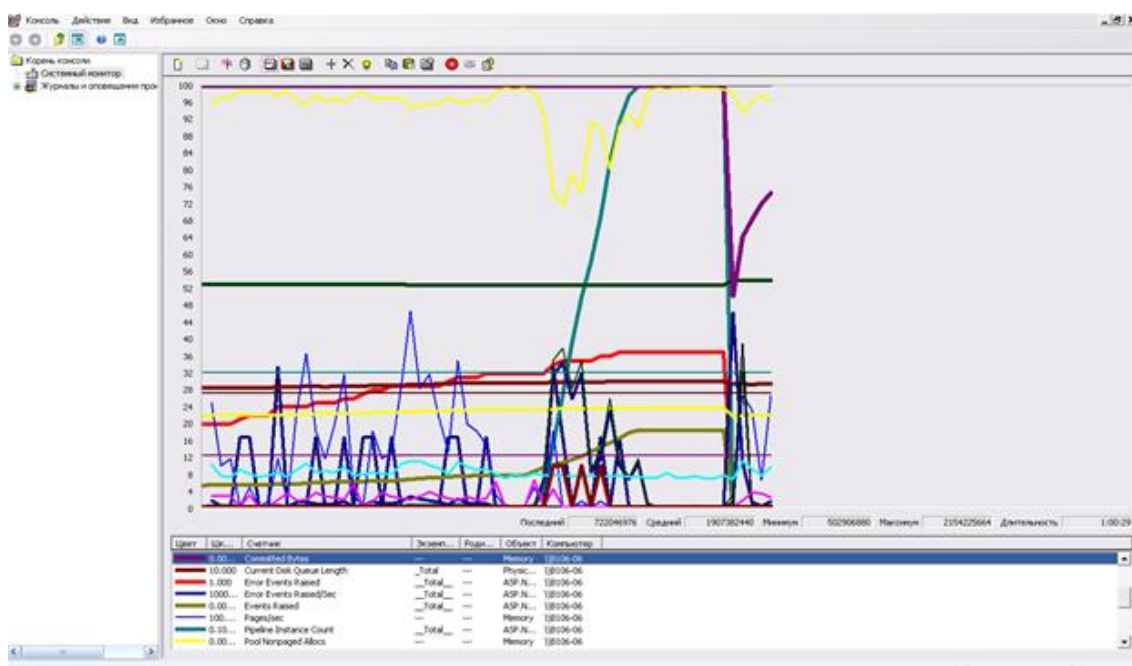


Рисунок 5.2.5 Эксперимент №8, мониторинг системных ресурсов одного из двух оставшихся включенными серверов (3).

Два оставшихся сервера не выдерживают нагрузки предназначенной для 4 серверов приложений.

5.2.4 Длительное тестирование (24 часа)

Длительные эксперименты проводились с нагрузкой 6 соединений (виртуальных пользователей) в течение длительного времени (24 часа). Эксперимент №10 выполнялся с 4мя серверами приложений и завершился неуспешно с большим количеством ошибок, связанных с отказом обслуживания. Эксперимент №11 перепроводился на увеличенном (до 6) количестве серверов приложений, при этом ошибок-отказов обслуживания со стороны АС обнаружено не было.

Таблица 5.2.3. Результаты успешного длительного теста №11 на 6 серверах приложений

Название запроса	Количество запросов за 24 часа (6 серверов приложений) Тест №11
OnlineMatching	114448
OnlineMatchingChange	12340
UpdateMetadata	64907
ИТОГО:	191695

Сравнительные результаты по обоим длительным экспериментам рассматриваются ниже.

Неуспешный длительный Тест (№10)

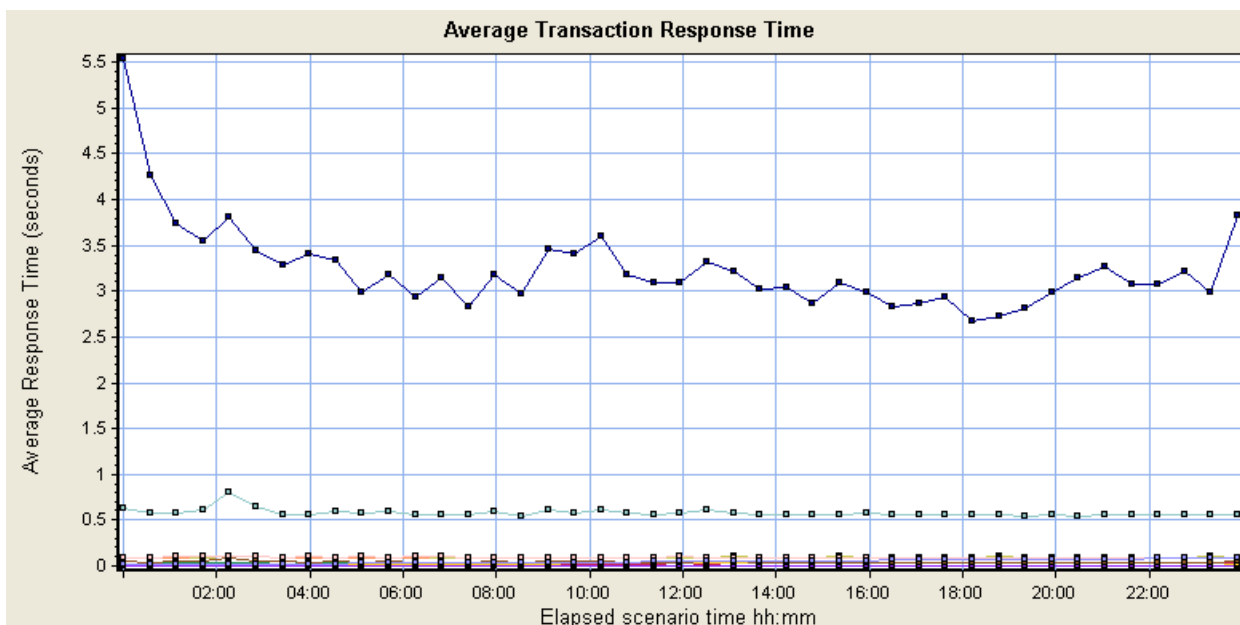


Рисунок 5.2.8. Тест №10, времена отклика

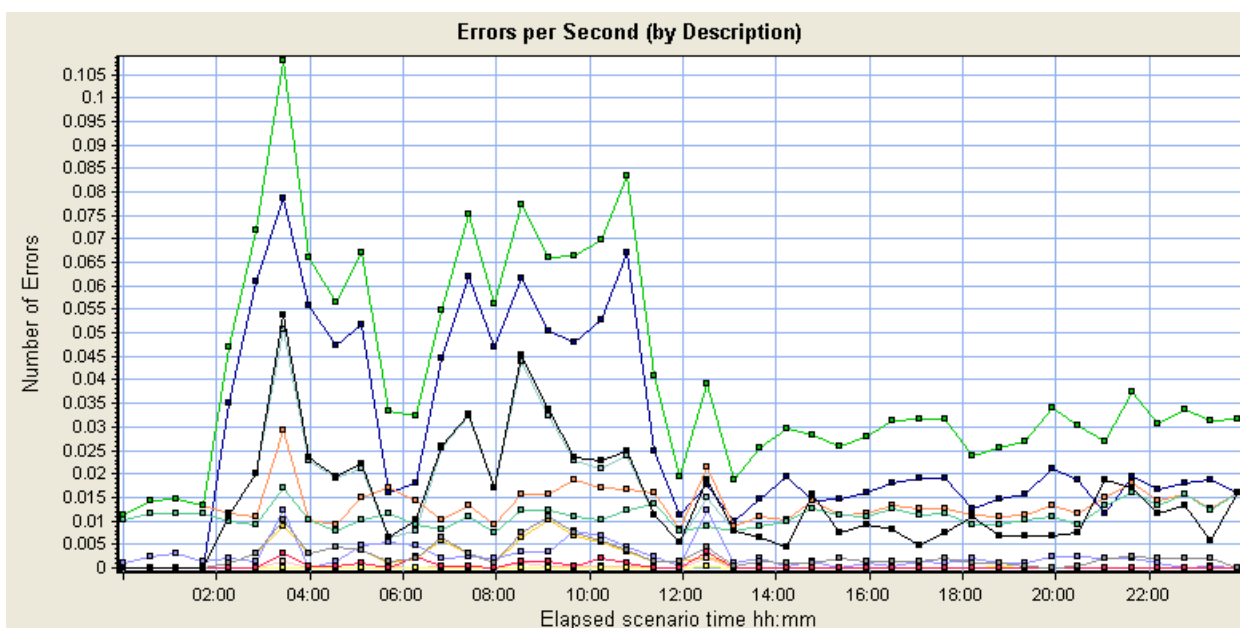


Рисунок 5.2.9. Тест №10, возникновение ошибок в течение теста

5.3 Тестирование на увеличенной базе данных

5.3.1 Тест определения максимальной производительности(№12)

Тест определения максимальной производительности полностью повторяет такой же тест, описываемый в п.5.1.1. Тем не менее, видно, что результаты данного теста несколько хуже. Это видно и по началу возникновения ошибок и по временам отклика на рисунках 5.3.1 и 5.3.2.

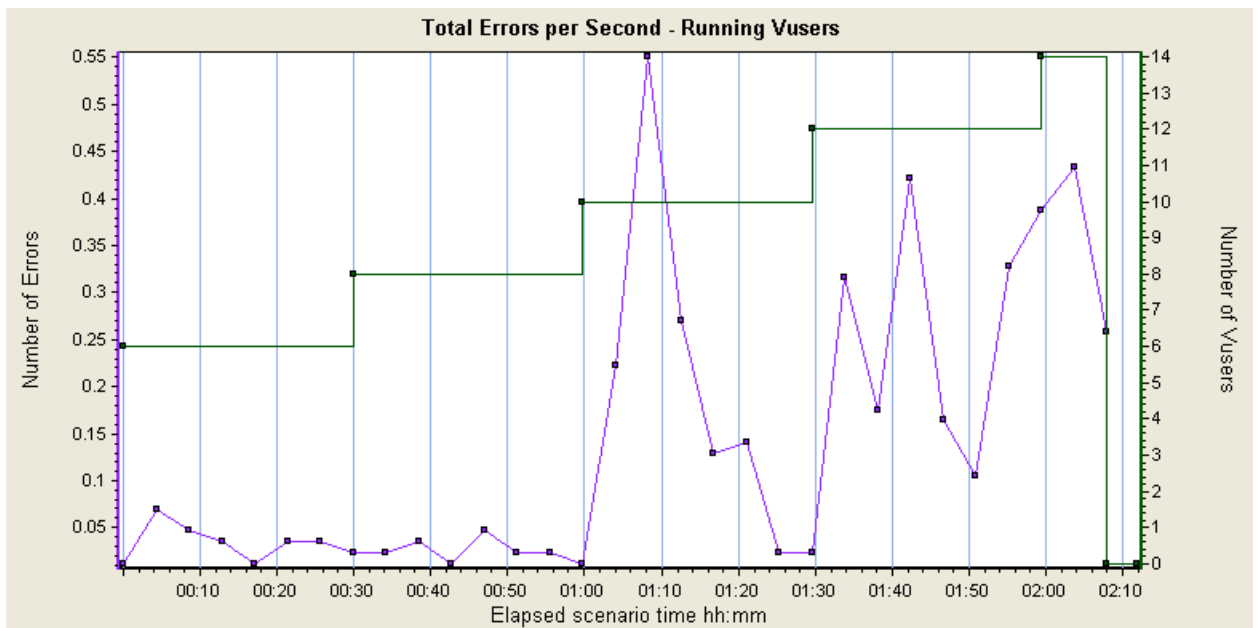


Рисунок 5.3.1 Увеличение нагрузки (зеленый график) и интенсивность возникновения ошибок (сиреневый график)

Видно, что высокая интенсивность появления ошибок имеет место уже на границе перехода с 8 соединений на 10. Это значит, что уже при тестировании на 8 соединениях мы находимся на пределе допустимой производительности.

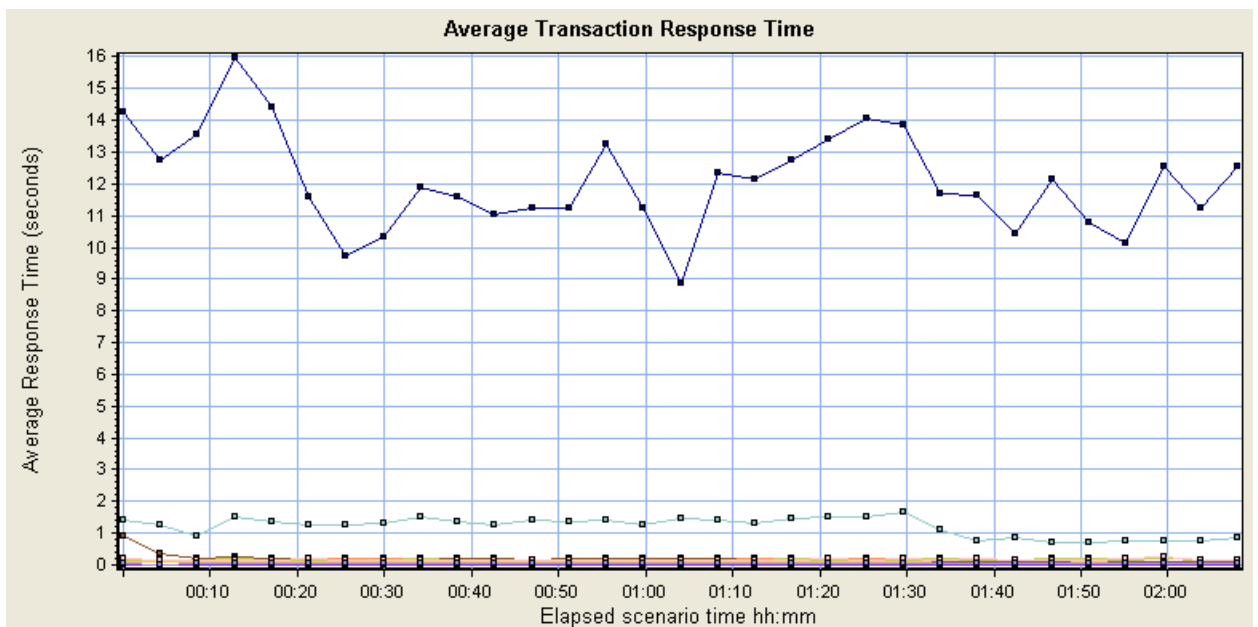


Рисунок 5.3.2 Времена отклика по операциям в течение эксперимента определения максимума производительности.

Времена отклика так же увеличились примерно в 3 раза на увеличенной базе данных.

5.3.2 Тест подтверждения максимальной производительности (№13)

Было принято решение проводить тест подтверждения на 6 соединениях, так как результаты в целом ухудшились. Рисунок 5.3.3 иллюстрирует времена отклика для теста подтверждения максимальной производительности.

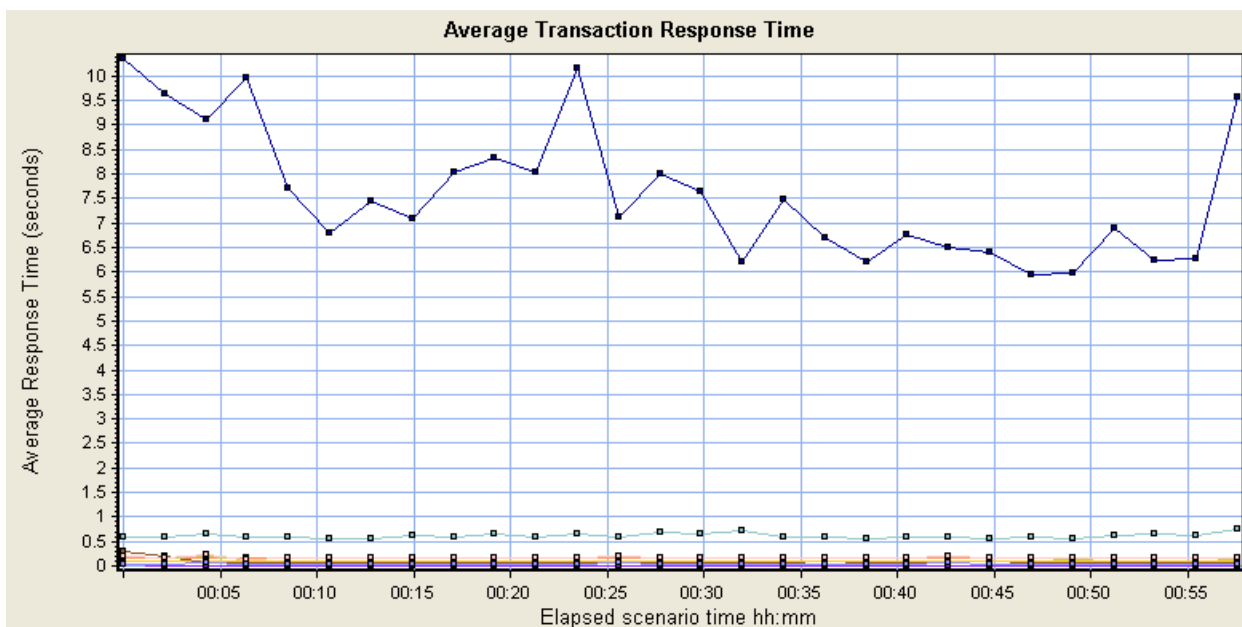


Рисунок 5.3.3. Времена отклика по операциям в течение эксперимента определения максимума производительности.

Таблица 5.3.1.

Название запроса	Кол-во/час
OnlineMatching	1814
onlineMatchingChange	236
updateMetadata	998
Итого:	3048

Таким образом, производительность, выраженная в обработанных запросах, на увеличенной базе данных снижается до 3048 операций в час. Анализируя времена обслуживания на дисковой подсистеме сервера БД можно предположить, что снижение производительности и увеличение времен отклика вызвано перегруженностью сервера БД. Если говорить точнее, то некоторой перегрузкой именно дисковой подсистемы, так как загрузка ЦПУ на сервере БД не превышала

50%. В качестве подтверждения может служить фрагмент теста в таблице 5.3.2. с временами обслуживания доходящими до 100 мс. Идеально эти времена не должны превышать 10 мс.