

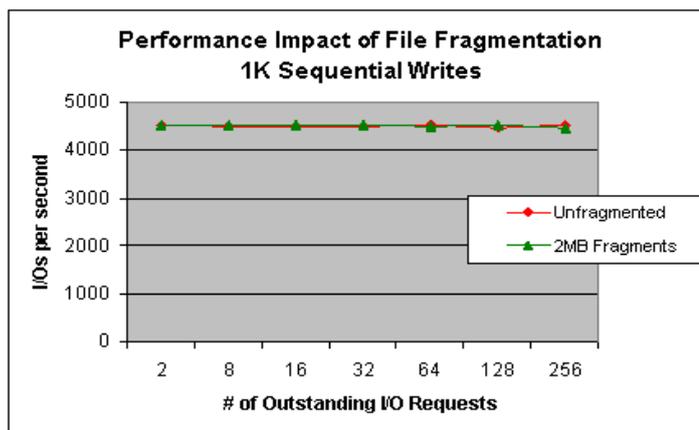
Влияет ли фрагментация на скорость доступа к данным.

Доступ к данным может быть случайным (random) и последовательным (sequential). Рассмотрим последовательный доступ. Отследим вопрос фрагментации, и выясним какое влияние она имеет на производительность.

В интернет ресурсах можно найти описание довольно интересного эксперимента. Для эксперимента было создано два файла размером 10 GB . Один из них поделили на 5 тысяч фрагментов размером по 2 MB, второй на 60 тысяч фрагментов размером по 128 KB. С помощью этих файлов проводили тестирование чтения и записи в последовательном порядке. Для хранения использовали диски сервера (на графиках обозначенные DAS), а также хранилище EMC Symmetrix DMX-2 (на графиках обозначенные SAN).

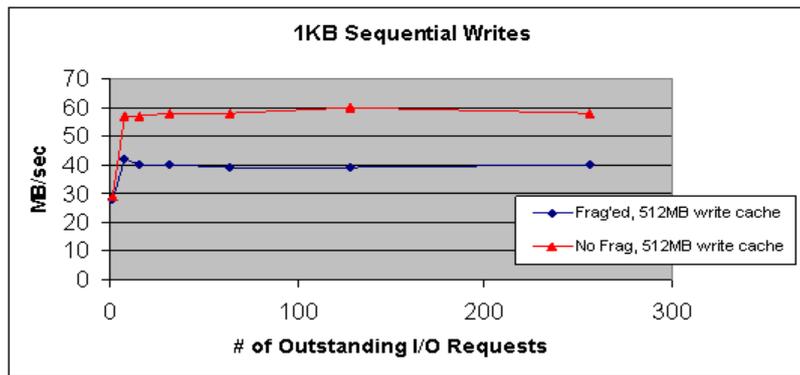
Со слов автора описание эксперимента выглядит так:

The server was a standard DL580 with 4 single-core sockets (2.0GHz Xeon) with 4GB of RAM. The disk array was a DMX2 with 10K rpm 146GB FC drives in a RAID10 config. The test LUN was 96GB in size (with 8GB disk slices from 24 spindles). Two Emulex LP8000 HBAs were load balanced with PowerPath.



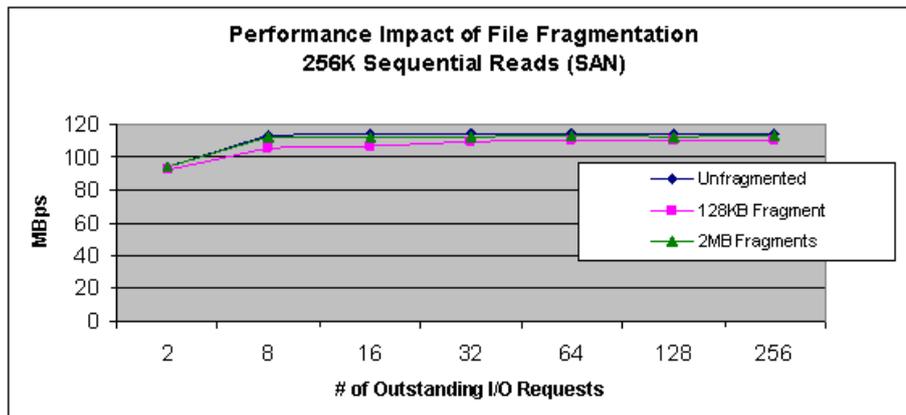
Результат первой части эксперимента получился довольно неожиданный. При использовании SAN – стораджа с высокой производительностью. Результат тестирования почти не отличался. Дефрагментированный и фрагментированный файл в результате тестирования последовательной записи фрагментами по 1К. Результативное кэширование и быстрое действие DMX сгладили эффект снижения скорости из-за множества фрагментов.

Следующим этапом эксперимента было создание точно такого же раздела на дисках серверов (DAS) и тестирование.



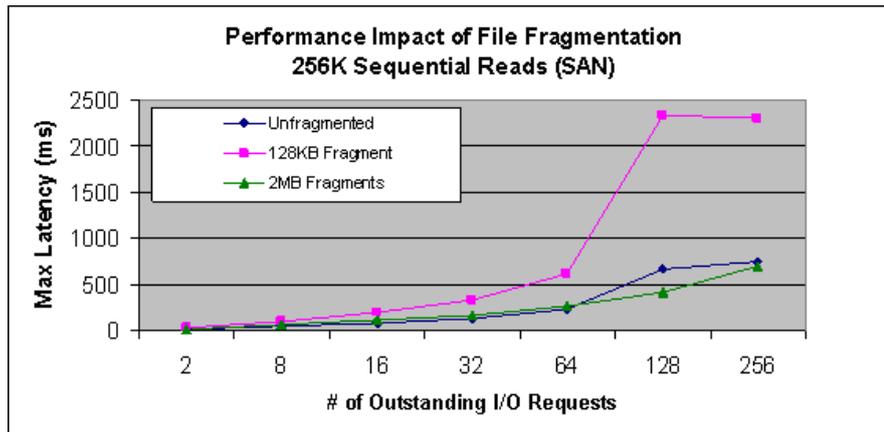
В этом случае на графике мы четко видим, что фрагментированность файла привела к снижению скорости действий на 30%.

Следующим шагом было уменьшение фрагментов до 128 КВ. Файл разделенный на 60 тысяч фрагментов протестировали используя последовательное чтение.



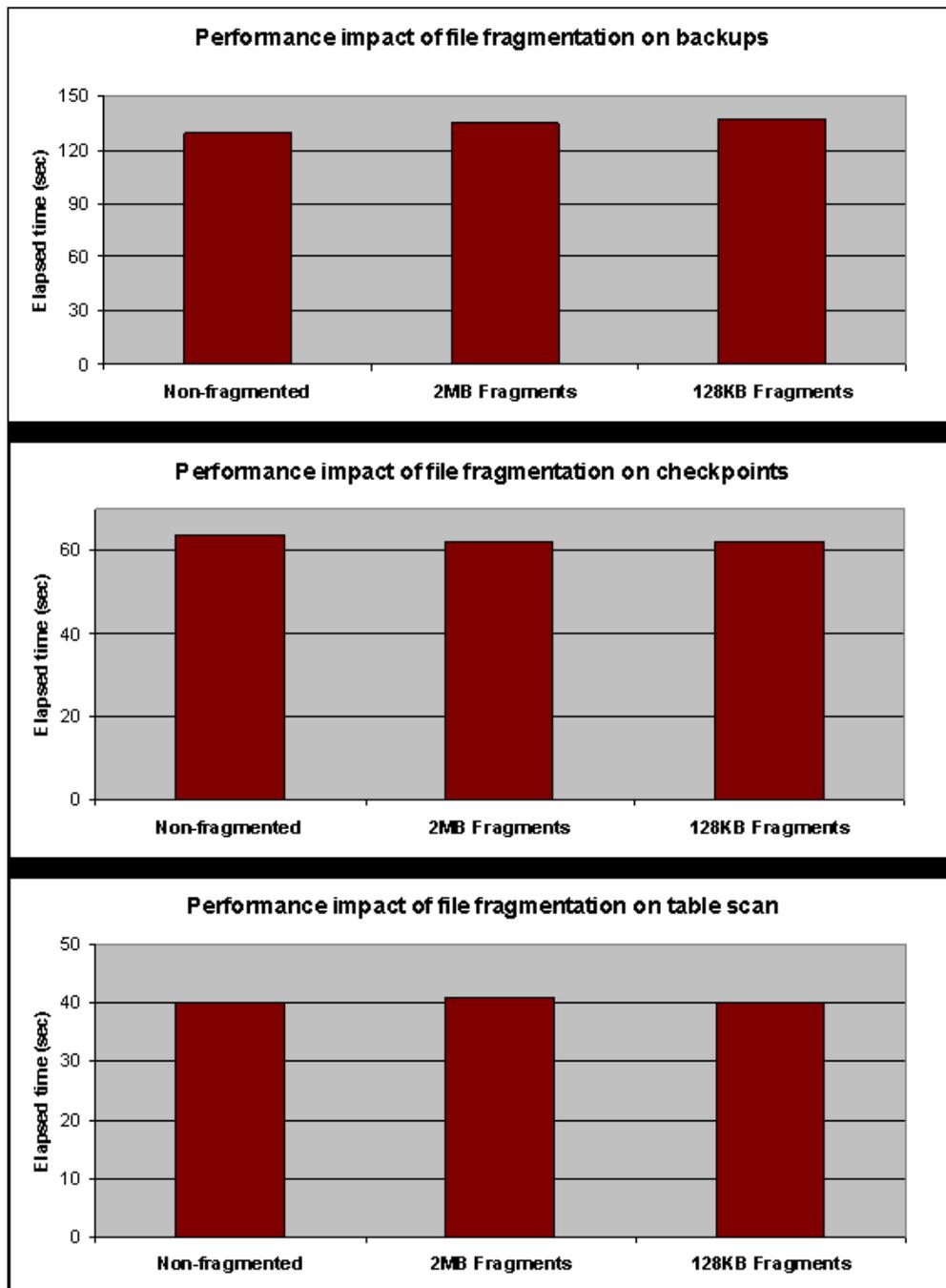
В результате выяснилось, что эффекта от фрагментации почти нет на высокопроизводительном массиве SAN. Зато на DAS фрагментирование файла принесло свои заметные плоды.

С помощью этого эксперимента был найден лишь один параметр.



В случае деления файла на 60 тысяч фрагментов, беспорядочно разбросанных на диске, увеличилось время Latency (задержка чтения). Самое интересное, что при разделение файла на 5 тысяч фрагментов такого результата не было.

Последним шагом в эксперименте было измерение времени, затраченного на выполнение ряда операций.



Цель этого эксперимента является не отрицание воздействия фрагментации на производительность записи и чтения информации при последовательном доступе. Степень влияния фрагментации зависит от многих факторов: режим доступа, размер файлов, блоков чтения и т.д.

Вывод можно сделать такой. Чаще всего влияние фрагментации на скорость работы системы слишком переоценивается. Из выше описанного мы видим, что используя последовательный доступ, этот режим более всего подвержен влиянию фрагментации, влияние ее мало значительно.