

Опция Compression продукта Sybase® Adaptive Server® Enterprise

Экономия благодаря оптимизации хранения данных

СПЕЦИФИКАЦИЯ ПРОДУКТА

ВОЗМОЖНОСТИ СЖАТИЯ ASE

- Сжимаются LOB размером до 2 Гбайт
- Оба применяемых алгоритма - FastLZ и ZLib - основаны на использовании словарей и обеспечивают сжатие без потерь
- Алгоритм и степень сжатия можно задавать отдельно для каждого столбца LOB
 - 9 степеней сжатия (2 для FastLZ и 9 для ZLib)
- Поддерживаются типы данных LOB
 - Сжатие типов Text/Image/UniText и объектов Java вне записей
- Сжатие возможно на уровне базы данных, таблиц и столбцов

Мы живем в эпоху экспоненциального роста объемов данных — процессы онлайн-трейдинга, интернет-продаж и использования мобильной информационной техники порождают триллионы транзакций, событий и процессов. Увеличению объемов данных способствуют ужесточающиеся нормативные требования и увеличение сроков хранения, а также потребность в ретроспективных исследованиях и анализе реального времени, являющихся основой для формирования информационных срезов, необходимых для управления бизнесом. Как справиться с растущими объемами данных и одновременно уложиться в бюджет?

В версии ASE 15.7 реализованы передовые возможности сжатия как структурированных, так и неструктурированных данных внутри базы — это позволяет, с одной стороны, сократить объем используемой дисковой памяти, с другой — уменьшить число операций ввода-вывода (благодаря тому, что при сжатии данных эффективный объем кэша увеличивается). Таким образом, ускоряется загрузка данных и повышается быстродействие. И самое главное, сжатие выполняется таким образом, что в прикладные программы, работающие с базами, не требуется вносить никаких изменений — так что внедрение новой версии не отнимает много времени.

СЖАТИЕ ДАННЫХ В ASE

Широкие возможности сжатия в ASE 15.7 позволяют достигать высоких степеней компрессии. Испытания свидетельствуют, что коэффициент сжатия может составлять 40-80%. Данные можно сжимать на уровне баз, таблиц и разделов; кроме того, они хранятся в упакованном виде на диске и в памяти. Компрессия может применяться по отношению к отдельным записям, когда сжимаются пустые поля/нули в столбцах постоянной длины. На уровне страниц и блоков применяется сжатие с помощью страничного словаря и страничного индекса. Повторяющиеся элементы данных, а также наборы элементов заменяются ссылками — это дает значительную экономию при хранении дублирующихся данных.



ПРЕИМУЩЕСТВА СЖАТИЯ В ASE

- Высокий коэффициент сжатия (40-80%) уменьшает объем занимаемой дисковой памяти
- Скорость загрузки данных увеличивается вплоть до 4 крат
- Гибкая и простая технология — сжимаются только те данные, которые необходимо упаковать. При этом не требуется вносить изменения в прикладные программы

Сжатие записей: сжимаются пустые поля/нули в столбцах постоянной длины

До

ABC	1	XX	“AC”
ABC	1	XX	“AC”

После

Страничное сжатие

- **Сжатие с помощью страничного словаря.** Повторяющиеся последовательности данных заменяются символическими ссылками, которая записывается в страничный словарь.
- **Сжатие с помощью страничного индекса.** Наборы столбцов с дублирующимися значениями сохраняются однократно, для ссылки на них в записи помещаются маркеры.

ASE имеет гибкие интерфейсы, позволяющие удобно выбирать подмножества данных, подлежащие сжатию. При создании сжатой таблицы или раздела система сжимает все данные, помещаемые в таблицу или раздел в ходе операций вставки или изменения. Если эффективно сжать вставляемые данные не удастся, то сохраняется исходная запись. Таблицы могут состоять одновременно из сжатых и несжатых данных. Если создается сжатая таблица, в нее загружаются данные, а затем режим сжатия для этой таблицы отключается, то ранее помещенные в нее данные остаются сжатыми, а вновь помещаемые не сжимаются. При этом нет необходимости распаковывать все данные для запуска запросов. Распаковываются лишь записи и столбцы, указанные в запросе, что повышает быстродействие. К сжатым данным можно применять операции вставки, изменения и удаления. Кроме того, можно отдельно задавать сжатие для активных или архивных данных. Такая гибкость позволяет легко настраивать систему на работу с существующими приложениями.

СЖАТИЕ БОЛЬШИХ ОБЪЕКТОВ (LOB)

ASE поддерживает также сжатие больших объектов (LOB) внутри базы. Поскольку эти объекты могут быть очень большого размера, их сжатие может дать весьма значительную экономию памяти. Применяются алгоритмы сжатия без потерь FastLZ и ZLib. Первый из них занимает меньше процессорных ресурсов и выполняется быстрее, второй обеспечивает более высокую степень сжатия.

Степень сжатия как обычных типов данных, так и LOB можно регулировать, например в зависимости от частоты обращения к данным. Например, к данным, которые требуются часто, целесообразно применять невысокую степень сжатия для экономии процессорных ресурсов.

ЗАКЛЮЧЕНИЕ

В 15 версии ASE реализован ряд улучшений в части хранения данных и быстродействия. Возможности сжатия структурированных данных и больших объектов позволяют компактнее хранить большие массивы данных и в то же время обеспечивают ускорение операций ввода-вывода — таким образом, при обработке даже самых больших массивов достигается высокое быстродействие. Описанные улучшения позволяют сократить затраты на хранение активно используемых данных, а также дополнительно сэкономить на вводе-выводе при кэшировании сжатых данных. Несмотря на тенденцию к снижению удельной стоимости физических систем хранения, с учетом опережающих ростов объемов данных совокупные издержки на управление ими растут. Предлагаемые ASE возможности сжатия позволяют удовлетворять возрастающие потребности благодаря эффективным хранению и загрузке больших массивов данных. ASE Compression лицензируется как опция.

Дополнительные сведения можно получить по адресу <http://www.sybase.ru/products/ase>.

Получить информацию о наших консультативных и образовательных услугах по поддержке внедрения технологий Sybase можно на странице www.sybase.ru