

Системы хранения данных (по материалам зарубежной печати)

Кузнецов И.М., ОАО СКБ СП, г. Саратов

Тенденции увеличения объема информации, получаемой при разведке и обработке данных в полевой и промысловой геофизике, существенно влияют на темпы и направление технического прогресса в области сбора, систематизации и хранения исходных, промежуточных и итоговых материалов на всех этапах современных технологий поиска и разведки полезных ископаемых.

Процесс сбора данных переходит от 2D к 3D технологии. Описание и мониторинг коллектора требуют использования через определенные временные промежутки 3D (4D) технологий сбора данных, что увеличивает количество пакетов данных в несколько раз. Подход к обработке изменился от последовательной к высокой интерактивности и интеграции. В настоящее время широко распространенной практикой является применение промежуточных стадий обработки для обеспечения дополнительного контроля качества и сохранности материалов. По-трассовой обработке переходит к MPP (Параллельная Обработка Массива). После увлечения широким использованием рабочих станций необходимость в больших ЭВМ существенно увеличилась и снова приобрела вес во всей схеме обработки. Использование центрального процессора (CPU) более чем удвоилось за последние три года и этот темп роста, похоже, сохранится. Одновременно, за тот же период объем 3D наблюдений увеличился в 10 раз. При существующем подъеме в промышленности этот рост, по-видимому, продолжится. Регистрирующие системы от нескольких сотен каналов перешли к тысячам каналов и каждый канал имеет повышенную разрешенность в битах и скорости квантования (рис.1).

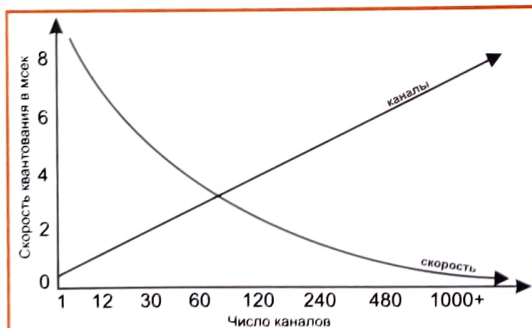


Рис. 1 Тенденции в регистрации сейсмоданных

Более того, параллельно и другие используемые средства и приборы, такие как: гравиметрические, магнитные, геоэлектрические, сейсмологические, дистанционного контроля и скважинной геофизики, получили те же преимущества от внедрения цифровой техники в традиционно аналоговые области. Емкость памяти RAM увеличилась с нескольких сотен мегабайт до нескольких терабайт, что также позволяет использовать значительно большие рабочие разрезы, файлы и модели в процессе обработки и интерпретации. Другими важными моментами

являются создание копий, перезапись и архивация. Обычные системы и резервирование данных требуют постоянного увеличения объема хранения данных. С учетом всех этих тенденций и дополнительных факторов мы сталкиваемся с быстрым ростом объема данных и необходимостью в новых подходах к проблеме управления этим ростом.

Как максимально эффективно решить проблему архивов, резервного копирования и оперативного доступа к столь значительным объемам информации? Прежде всего следует разобраться в существующих технологических решениях и способах организации систем хранения данных.

Хранение данных в корпоративной информационной системе.

Распределенная корпоративная система хранения данных современной геофизической компании, как правило, имеет: корпоративный электронный архив, включающий: дисковый массив и библиотеки электронных носителей; архив данных стеллажного хранения на бумажных и электронных носителях; систему ремастеринга данных.

1. Корпоративный электронный архив

1.1. Дисковый массив

В информационных системах первичным накопителем информации служит жесткий диск. Объемы баз данных, как правило, всегда превышали емкости одиночных дисков, поэтому информация сначала записывалась на один диск, а затем, когда свободное пространство на нем заканчивалось, - на другой. При таком подходе к размещению информации можно создавать большие тома, но ускорения работы дисковой системы не происходит. В 1987 г. была предложена идея объединения нескольких дисков в единый массив - технология RAID. В основе технологии RAID лежат три основных метода записи и защиты информации: распределение последовательности сегментов данных по дискам с определенной циклической очередностью (поочередное размещение); зеркальное отображение дисков; вычисление контрольных сумм. Поочередное размещение сегментов на разных дисках позволяет создавать большие тома и ускорять выполнение операций ввода-вывода; "зеркалирование" накопителей и вычисление контрольных сумм обеспечивают избыточность информации, позволяющую в случае сбоя



Рис. 2 Архитектура DAS

ЭЛЕКТРОННЫЕ АРХИВЫ В ГЕОФИЗИКЕ
ТЕМАТИЧЕСКИЙ ОБЗОР

Табл. 1 Преимущества и недостатки NAS

Преимущества	Недостатки
• Для расширения емкости системы не требуется переводить серверы в режим off-line • Емкость системы расширяется очень просто • Доступ к файлам по существующим сетям протокола TCP/IP • Клиенты в среде различных ОС могут совместно использовать файлы • Затраты ниже, чем при использовании обычных серверов	• Единственная точка подключения для разделов устройства хранения может привести к возникновению узких мест • Каждое устройство управляется как отдельный том • Нет разделения данных между устройствами • Архитектура пригодна только для передачи файлов; нет поддержки данных на уровне блоков

восстановить утраченные данные.

Организация хранилища в виде дискового массива при всей его технической мощи требует серьезных затрат и подходит далеко не для любой цели.

В первые годы нового века почти три четверти общего числа продаж систем хранения данных пришлось на системы прямого подключения к серверу типа DAS (Direct-Attached Storage), не всегда соответствующие требованиям в области надежности и доступности данных. Решения DAS (рис. 2) были разработаны более сорока лет назад для расширения емкости хранения существующих серверов. Дисковая система DAS, ассоциированная с сервером, доступна только через этот сервер, что приводит к созданию отдельных "островов" данных и оказывается серьезным недостатком с точки зрения сегодняшней ИТ-инфраструктуры. Кроме того, DAS не позволяет нескольким серверам совместно использовать файлы, а для добавления дисковых накопителей к сети серверы должны переводиться в автономный режим. По мере роста требований к системам хранения приходится просто добавлять диски к серверам и серверы к сетям, что существенно усложняет инфраструктуру системы хранения данных.

Для преодоления проблем, характерных для DAS, в 90-х годах были разработаны архитектуры сетевых систем хранения данных NAS (устройства хранения данных, подключаемые к сети) и SAN (сети хранения данных).

Основная задача архитектуры NAS (Network Attached Storage) - упростить совместное использование файлов. На базовом уровне системы NAS (рис.3а) представляют собой оборудование, подключаемое непосредственно к локальной сети. В отличие от архитектуры DAS, в системах NAS не нужно переводить серверы в автономный

режим для увеличения общей емкости, диски можно добавлять в структуру NAS, просто подключив устройство в сеть. Преимущества и недостатки NAS, приведены в табл. 1.

Системы SAN (Storage Area Networks) решают не только многие проблемы DAS и NAS, но, что более важно, и проблемы перегрузки каналов связи и задержки в локальных IP-сетях. Архитектура SAN (рис.3б), переносящая всю систему хранения в выделенную подсеть, ориентирована на интенсивную работу в режиме резервного копирования и обмена данными. Преимущества и недостатки SAN приведены в табл. 2.

Как показывает практика, значительная часть (до 80%) хранимой в предприятии информации используется достаточно редко. Ясно, что, освобождая RAID-системы от подобной информации, перемещая ее на менее дорогие сменные носители, мы получаем сразу целый ряд преимуществ: уменьшается стоимость подсистемы хранения, повышается производительность дисковых массивов, снижается объем и время резервного копирования, а значит, и потребность в оборудовании.

1.2. Библиотеки электронных носителей

• Роботизированные библиотеки

Одним из решений, позволяющих повысить эффективность применения RAID-систем и обеспечить прозрачный доступ к архивным данным, являются роботизированные библиотеки. Роботизированная библиотека представляет собой устройство, в корпусе которого размещаются магазины для хранения сменных носителей (нескольких сотен дисков или лент), приводы, транспортный механизм для смены носителей в приводах. Для этих архивных накопителей характерны такие достоинства, как высокая емкость, стабильность и долговечность используемых носителей при их невысокой стоимости, возможность автоматизированного управления, возможность создания физически защищенных от удаления записей, простота внедрения, легкость масштабирования, невысокие затраты при модернизации. Эти накопители эффективны для работы с большими объемами (десятки терабайт) редко используемой информации, но по скорости доступа они уступают RAID-системам, и поэтому их называют накопителями near-line ("почти на линии"). RAID-системы, в свою очередь, обеспечивают оперативный доступ к информации, на них хранятся наиболее активно используемые данные.

В качестве вторичных накопителей используются архивные накопители на магнитных лентах, оптических дисках (CD и DVD) и магнитооптических дисках. В табл. 3 приведены базовые характеристики носителей. Базовые



Табл. 2 Преимущества и недостатки SAN

Преимущества	Недостатки
• Высокая пропускная способность • Простота и отсутствие ограничений на масштабирование • Снижение вероятности перегрузок в локальной сети • Поддержка блочной передачи данных • Поддержка совместного использования данных различными серверами	• Ограничение на расстояние препятствует использованию зеркалирования или разбиения данных с различных геофизических точек • Сети на основе протокола Fibre Channel дороги и сложны в поддержке • Совместное использование файлов в общем случае не возможно • Закрытые архитектуры собственной разработки

характеристики накопителей для различных носителей приведены в табл. 4 (данные взяты по разным производителям).

В библиотеках геофизических компаний, где хранятся чрезвычайно большие файлы и необходима высокая потоковая скорость последовательной передачи данных, а потребности в одновременном доступе к данным невысокие, эффективными устройствами near-line служат ленточные библиотеки.

На рынке представлены различные компании, разрабатывающие и изготавливающие роботизированные библиотеки магнитных лент IBM (IBM 3583, 3584 Ultrium Tape Library), Overland Data (LibraryPro, LibraryXpress Neo Series), Fujitsu Siemens Computers (CentricStor), HP (HP StorageWorks MSL6000, ESL) и т.д. Практически все производители библиотечных систем поставляют устройства всевозможных конфигураций: с различным числом и набором накопителей, магазинов и гнезд, с возможностью установки гнезда обмена или без него, допускающие или нет "горячую" замену накопителя. Различны также и интерфейсные соединения.

• Ленточные накопители

Несмотря на удешевление дискового пространства, ленточные накопители до настоящего времени все равно значительно выигрывают в цене в расчете на единицу хранимой емкости, и этот разрыв в ближайшем будущем сохранится. Поэтому отрасль ленточных накопителей по-прежнему развивается: увеличивается емкость лент, скорость записи и другие основные параметры этих систем. В соответствии с используемой технологией записи выпускаемые в настоящее время накопители делятся на три основные категории:

1. накопители с линейной технологией записи;

2. накопители с продольной серпантинной технологией записи;

3. накопители со спиральной технологией записи.

Консорциум, объединяющий корпорации IBM, Hewlett-Packard и Seagate Technology, разработал открытый стандарт записи на ленту для продольной серпантинной технологии записи - Linear Tape Open (LTO). Существует два варианта этой технологии - Ultrium (однокатушечный картридж) и Accelis (двухкатушечный картридж). В своих разработках IBM использует стандарт Ultrium, обеспечивающий высокую скорость записи и надежность хранения.

Первое поколение устройств (имеются планы развития технологии вплоть до четвертого поколения) обладает скоростью записи до 108 Гбайт/час на один лентопротяжный механизм (драйв), предусмотрена возможность чтения во время записи (Read While Write, RWV). Тестовое чтение данных сразу после записи (read after write verification) гарантирует проверку данных в масштабе реального времени и позволяет динамически перезаписывать сбойные блоки. На ленте шириной менее 1,3 см, благодаря применению технологии многоканальной серпантинной записи и серводорожек, используемых для точного позиционирования головки, удастся записать 384 дорожки данных. Микросхема памяти емкостью 4 Кбайт на каждом картридже хранит информацию о состоянии носителя. Информация с этой микросхемы используется лентопротяжным устройством при записи на картридж. Если в дальнейшем лента будет использоваться в другой библиотеке, эта библиотека получит всю необходимую информацию о ленте. Даже если 12,5% ленты окажется испорченной (хотя представить такое

Табл. 3 Базовые характеристики носителей

Тип носителя	Скорость чтения	Скорость записи	Время доступа	Возможность записи	Возможность перезаписи	Удельн. стоим. хранения 1Гбайта, долл.	Срок хранения, лет
CD	Выше средней	Нет	Среднее	Нет	Нет	0,8-1,5	Более 100 лет
CD-R	Средняя	Средняя	Среднее	Да	Нет	0,4-1,5	Более 100 лет
CD-RW	Средняя	Средняя	Среднее	Да	Да	1,0-2,5	100 лет
DVD-RAM	Средняя	Средняя	Среднее	Да	Да	1,9-6,0	Более 100 лет
Магнитн. лента	Высокая	Высокая	Большое	Да	Да	0,7-1,7	30 лет
Магнито-оптический диск	Высокая	Высокая	Малое	Да	Да	1,6-2,5	50-100 лет
Жесткий диск	Очень высокая	Очень высокая	Очень малое	Да	Да	3,5-30	Не регламент.

ЭЛЕКТРОННЫЕ АРХИВЫ В ГЕОФИЗИКЕ ТЕМАТИЧЕСКИЙ ОБЗОР

Табл. 4 Базовые характеристики накопителей

Накопитель	Потоковая скорость чтения, Мбайт/с	Потоковая скорость записи, Кбайт/с	Потоковая скорость перезаписи, Кбайт/с	Время доступа, мс	Емкость носителя	Время наработки на отказ (MTBF), тыс. ч.
CD-ROM	0,9-2,4	-	-	100-170	650, 740 Мбайт	100
CD-R	0,9-2,4	0,9-1,8	-	125-180	650, 740, 800 Мбайт	100
CD-RW	0,9-2,4	0,9-1,8	0,9-1,2	125-180	650, 740, 800 Мбайт	100
DVD-RAM	1,35-2,8	-	-	80-180	1,46; 2,6; 4,7 Гбайт	100
Tape	1,1-6,0	1,1-6,0	1,1-6,0	1000 для самых быстрых	512 Мбайт, 2,5; 4; 8; 24; 40; 80 Гбайт	250 (читающая головка 50)
MO	2,1-4,6	1,1-2,3	1,6-2,3	15-35	128; 230; 640 Мбайт, 1,3; 2,6; 4,2; 9,1; Гбайт	250
HDD	До 39,4	До 39,4	До 39,4	4,8-16	До 250 Гбайт	До 1200

повреждение достаточно сложно, учитывая, что картридж рассчитан на 228 лет), стример продолжит работать. Емкость картриджа составляет 100 Гбайт без сжатия или 200 Гбайт со сжатием, скорость записи достигает 15 Мбайт/с, скорость поиска - 6 м/с. Такие параметры выглядят очень привлекательно на фоне других устройств хранения.

В Табл. 5 приведен ряд параметров накопителей, использующих различные технологии записи.

• HSM - иерархические хранилища

Сам по себе объем всех устройств хранения - это лишь часть проблемы управления корпоративными данными. Результатом реализации систем хранения на предприятиях часто оказывается появление отдельных "островов" архивных данных на разных носителях в разных устройствах, работающих под управлением разных операционных систем. Для максимально эффективного использования всех ресурсов необходимы программные решения, которые строго связывали бы жесткий диск с носителями других классов. Наиболее эффективную, прозрачную для пользователей и приложений одновременную работу со всеми типами накопителей корпоративного хранилища обеспечивает ПО, работающее в соответствии с концепцией сетевого хранения HSM (Hierarchical Storage -Management, управление иерархическими системами хранения). При использовании программных продуктов HSM подключение библиотеки к информационной системе выглядит как расширение уже существующего тома RAID-массива - не требуется никакой перенастройки запросов в системе. Таким образом, вся совокупность накопителей превращается в единое виртуальное хранилище информации.

На сегодня основные разработчики ПО HSM - компании OTG Software, BakBone Software, FalconStor и Sanbolic. Ключевое преимущество HSM-технологии - возможность автоматической оптимизации размещения данных по различным типам носителей информации. Благодаря этому удается минимизировать стоимость владения данными: использование дорогих систем хранения становится более эффективным при сохранении доступности и увеличении надежности информации. Накопители в HSM организованы в иерархию и разделены по функциональ-

ному признаку: первичные накопители - высокопроизводительные RAID-массивы; вторичные - более емкие и надежные роботизированные библиотеки. Процесс оптимизации напрямую связан с традиционными процедурами архивирования: резервированием (перемещением редко используемых данных на недорогие носители архивных накопителей) и восстановлением (возвращением затребованных данных на RAID-массив). Для пользователей и приложений эти операции протекают прозрачно, в фоновом режиме; миграция данных осуществляется автоматически, на основе правил, заданных при настройке ПО. Как уже говорилось, в качестве вторичных используются архивные накопители на магнитных лентах, оптических дисках и магнитооптических дисках.

2. Архив стеллажного хранения на бумажных и электронных носителях

В настоящее время в архивах стеллажного хранения информации геофизических предприятий хранится огромное количество геофизических материалов на различных носителях, накопленных за долгий период времени. Данные материалы включают как первичные материалы, полученные в ходе проведения геофизических работ (сейсмозаписи, параметрическая информация, карты, схемы, SPS-файлы, рапорты операций и т.д.), так и материалы, являющиеся результатом обработки первичных материалов и интерпретации.

Хранение данных в архиве стеллажного хранения является существенным этапом процесса управления данными в течение их жизненного цикла (Life Cycle Data Management) в ходе которого решаются следующие задачи:

1. приемка информации на различных носителях от поставщиков;
2. регистрация данных в каталоге архива;
3. ремастеринг данных;
4. контроль состояния носителя и перенос (при необходимости) данных на новый носитель.

Одной из важнейших задач, решаемых в процессе хранения данных в архиве, является проведение ремастеринга данных. В процессе ремастеринга архивные данные подготавливаются к использованию, включая контроль качества записанных на носители данных, оцифровка (при необходимости) материалов, перевод данных в цифровые обменные форматы.

Таблица 5 Накопители на магнитной ленте

Технологии/накопители	Производитель	Макс. скорость переноса несжатых данных
Линейная технология		
7 дор., 1/2", кассеты		6.25-7.25 KB
21 дор., 1/2", кассеты		128 KB
9 дор., 1/2", кассеты	IBM, 3M, Mtmorex, Graham	20-780 KB
18 дор., 1/2", 3480 картридж	IBM, Storage Tec, 3M	1/5 MB
18/36дор., 1/2", 3490 картридж	IBM, Storage Tec, 3M	2 MB
36дор., 1/2", 3490E картридж*	IBM, Storage Tec, 3M	3 MB*
Продольная серпантинная технология		
128-дор. DLT (1/2") 4000 картридж	DEC	1.25 MB
128-дор. NTP 3590 (1/2") картридж	IBM	6-9 MB
128-дор. NCTP 3600 (1/2") картридж	Philips LMS	10 MB
208-дор. DLT 8000 (1/2") картридж	Quantum	6 MB
384-дор. LTO Ultrium (1/2") картридж	IBM, HP, Seagate	15 MB
448-дор. SDLT-220 (1/2") картридж	Quantum	11 MB
Спирально сканирующая технология		
DAT (4mm) картридж	Gigatape, HP	185-500 KB
CAM (8mm) картридж	Exabyte	240-500 KB
VHS (13mm) картридж	Mass Data Systems **	250 KB
VHS (13mm) картридж	Metrum/Mountain Gate	2 MB
DDS-3/DDS-4 (4mm) картридж	Seagate	S-4/8 MB
SD3 (1/2") картридж	Seagate Tek, 3M	11.25 MB
DTF-1/DTF-2 (1/2") картридж	Sony	12-20 MB
AIT-2/AIT-3 (8mm) картридж	Sony	6-12 MB
D1-s,m,i (19mm) картридж	Sony	14.4 MB
D2-s,m,i (19mm) картридж	Ampex, Fuji, Sony	>15 MB
DDD2-s,m,i (19mm)(дв.пл-ть.)картридж	Ampex	>20 MB
QDD2-s,m,i (19mm)(дв.пл-ть.)картридж	Ampex	>20MB

* 3490E пишет сжатые данные

**Устройство может использоваться в полевых системах

(например, SEG-Y для сейсмических данных, LAS для данных каротажа, ASCII для картографических материалов и т.д.), проверка координат наблюдения и т.д.

Ни один носитель записи не может существовать вечно. Таким образом, перенос данных со старого носителя на новый неизбежен. Вопрос в том, когда этот момент наступает? Современные ленточные накопители в процессе записи/чтения вырабатывают информацию, позволяющую быстро отличить неисправность носителя от неисправности накопителя без использования роботов, других накопителей или системного времени. Набор пороговых значений укажет, что файл должен быть скопирован с данной ленты на новый носитель данных.

Контроль качества носителя можно также осуществлять, ориентируясь на следующие статистические данные:

1. число обращений к файлу;
2. число установок в накопитель картриджа или кассеты;
3. возраст (время нахождения) файла/объема в системе;
4. число проведенных коррекций ошибок, отражающее легкость/трудность чтения.

Эти данные могут контролироваться для идентификации файла/объема и предупреждения системных операторов, чтобы они запланировали процедуру копирования в удобное для системы время. Однако рациональнее привлечь к выполнению этой задачи отдельный сервер/робот архива, чтобы эта работа была автоматизирована, и процедуры тестирования и копирования выполнялись без вмешательства в обычный процесс работы. То

же самое относится и к процедуре выполнения ремастеринга.

Обеспечение сохранности и круглосуточного доступа к данным без узких мест очень сложная задача, но ее можно решить при наличии желания, знаний и достаточных денежных средств. Компании, успешно использующие свою информацию, полностью отдают себе отчет в преимуществах, проистекающих из способности управлять данными, вот почему они остаются впереди своих конкурентов в сегодняшней быстро меняющейся рыночной среде.

Литература

1. Savage, P. Recommended practices for storage and archiving of exploration data. The Leading Edge, 13(2), 1994
2. Garry Nikola Kolafa, Modern Storage architectures, mass storage systems, their components and their application in geophysics. First Break, V.21, 2003
3. Еремин С. Хранилища данных, или где и как хранить терабайт? PC Magazine № 3, 2002
4. Борзенко А. Технология RAID. BYTE, № 8, 2002
5. Борзенко Ф. Унифицированное сетевое хранение данных. BYTE, № 8, 2002
6. Воробьев М. Решения IBM TotalStorage. BYTE, №10, 2002
7. Васильев С. Хранение данных в КИС. Byte, № 1, 2002
8. Гермашев Б. Выбор корпоративной системы хранения информации. Byte, №10, 2002

ЭЛЕКТРОННЫЕ АРХИВЫ В ГЕОФИЗИКЕ ТЕМАТИЧЕСКИЙ ОБЗОР