ВУЗ: МГТУ им. Баумана

Группа: ИУ5-61Б

Студент: Белоусов Евгений

Near-line and offline storage (often called Hierarchical Storage Management) is the modern way of dealing with current storage needs. Hard disks are becoming cheaper, but data storage s requirements are higher, so it's better to plan for HSM than assume disks can continually be added to systems.

HSM is essentially the automatic movement of data between media, the media type used depending on when it was last accessed. Many software and hardware vendors have HSM solutions, and all are based on the same basic techniques.

The most common HSM setup is where there's online storage (the hard disk), near-line storage (some sort of fast media from where a file can be quickly retrieved), and offline storage (slower media that might take some time for files to be recovered, but it is cheaper for a long-term storage). This arrangement is the major thrust of today's systems. Most of the time these systems will comprise optical media for near-line and tape media for offline storage.

Data is automatically moved from the online disk to the near-line optical media if it hasn't been accessed for a definable period of time. This is typically three months (depending on your business). This near-line system is likely to be erasable optical disks in some form of jukebox.

The system has to operate on the basis that a user won't know that a file has been moved into near-line storage. Therefore some marker is left in the directory structure on the disk so that the user can still see the file. If the user then tries to open it, the file will automatically be copied from near-line to online storage, and opened for the user. All the user notices is a slight time delay while the file is opened.

Moving data from near-line to offline storage can be done using a similar mechanism, but more often the marker left in the directory for the user to see will just contain a reference. This gives the user the facility to request the file back from the systems administrator, and could have information like "This file has been archived to offline media' and a reference to the tape number that the file is on. This is then sent to the systems administrator and the file can be recovered from tape in the usual way.

Some modern systems have the ability to keep multiple tapes in a tape changer or jukebox system, so retrieval from offline to online storage can be automatic. However, it is more likely that when a file goes into offline storage it s will never be recovered, as it has probably been untouched for several months (again depending on the business). Therefore the requirement to recover from offline to online is reasonably infrequent.

The choice of storage media type is a crucial aspect of HSM. The cheapest is undoubtedly tape (be it digital, analogue or digital linear), so this tends to be used for offline storage. However, tape has no guarantee of data integrity beyond

as one or two years, whereas optical systems, such as CDs, WORMs and MO disks, have much better data integrity over a longer period of time. Depending on the precise application, archiving systems are usually based on the media type

w that has the best integrity. The major suppliers within the HSM market are totally open about the media that can be used with their software. Current HSM systems support most hardware devices, so you can mix and match media to suit requirements. Given the fact that media choice depends on the length of time you want your data to remain intact, and also the speed at which you want to recover it, the choice for many system managers is as follows.

Tape is used for backup systems where large amounts of data need to be backed up on a regular basis. Tape is cheap, integrity is good over the short to medium term, and retrieval from a backup can be made acceptable with good

ss tape storage practices.

Near-line storage should be based on erasable optical disks. This is because access is random, so the access speed to find and retrieve a particular file needs to be fast, and data integrity so is also good.

Archiving systems should probably be CD- or WORM-based, as again access speeds are good, media costs are reasonably cheap and, importantly, the integrity of the media over the medium to long term is good.

One important thing to remember with archiving systems is the stored data's format. The data might be held perfectly for 10 or 15 years, but when you need to get it back, it's essential that you maintain appropriate hardware and software to enable you to read it.

Ближнее и автономное хранение данных (часто называемое иерархическим управлением хранилищами) - это современный способ удовлетворения текущих потребностей в хранении данных. Жесткие диски становятся дешевле, но требования к хранилищу данных выше, поэтому лучше планировать HSM, чем предполагать, что диски могут постоянно добавляться в системы.

HSM-это, по сути, автоматическое перемещение данных между носителями, тип носителя, используемый в зависимости от того, когда он был получен в последний раз. Многие производители программного и аппаратного обеспечения имеют решения HSM, и все они основаны на одних и тех же базовых методах.

Наиболее распространенная установка HSM - это когда есть оперативное хранилище (жесткий диск), ближнее хранилище (своего рода быстрый носитель, из которого можно быстро извлечь файл) и автономное хранилище (более медленный носитель, который может занять некоторое время для восстановления файлов, но это дешевле для долгосрочного хранения). Это устройство является главной движущей силой современных систем. Большую часть времени эти системы будут включать оптические носители для ближней линии и ленточные носители для автономного хранения.

Данные автоматически перемещаются с сетевого диска на Ближний оптический носитель, если они не были доступны в течение определенного периода времени. Это обычно составляет три месяца (в зависимости от вашего бизнеса). Эта почти линейная система, скорее всего, представляет собой стираемые оптические диски в какой-то форме музыкального автомата.

Система должна работать на основе того, что пользователь не будет знать, что файл был перемещен в ближайшее хранилище. Поэтому некоторый маркер остается в структуре каталогов на диске, так что пользователь все еще может видеть файл. Если пользователь затем попытается открыть его, файл будет автоматически скопирован из ближней линии в онлайн-хранилище и открыт для пользователя. Все, что замечает пользователь, - это небольшая временная задержка при открытии файла.

Перемещение данных из почти линейного хранилища в автономное можно выполнить с помощью аналогичного механизма, но чаще всего маркер, оставленный в Каталоге для просмотра пользователем, будет содержать только ссылку. Это дает пользователю возможность запросить файл обратно у системного администратора и может содержать информацию типа "этот файл был заархивирован на автономный носитель" и ссылку на номер ленты, на которой находится файл. Затем он отправляется системному администратору, и файл может быть восстановлен с ленты обычным способом.

Некоторые современные системы имеют возможность хранить несколько лент в лентопротяжном устройстве или музыкальном автомате, поэтому извлечение из офлайн-хранилища в онлайн-хранилище может быть автоматическим. Однако более вероятно, что когда файл попадает в автономное хранилище, он никогда не будет восстановлен, так как он, вероятно, был нетронутым в течение нескольких месяцев (опять же в зависимости от бизнеса). Поэтому требование о восстановлении из оффлайна в онлайн-режим достаточно редко встречается.

Выбор типа носителя информации является важнейшим аспектом HSM. Самым дешевым, несомненно, является лента (будь то цифровая, аналоговая или цифровая линейная), поэтому она, как правило, используется для автономного хранения. Однако лента не гарантирует целостности данных более одного или двух лет, в то время как оптические системы, такие как компакт-диски, черви и МО-диски, имеют гораздо лучшую целостность данных в течение более длительного периода времени. В зависимости от конкретного применения системы архивирования обычно основываются на типе носителя, который обладает наилучшей целостностью. Основные поставщики на рынке HSM полностью открыты в отношении средств массовой информации, которые могут быть использованы с их программным обеспечением. Современные системы HSM поддерживают большинство аппаратных устройств, поэтому вы можете смешивать и сопоставлять носители в соответствии с требованиями. Учитывая тот факт, что выбор носителя зависит от того, сколько времени вы хотите, чтобы ваши данные оставались нетронутыми, а также от скорости, с которой вы хотите их восстановить, выбор для многих системных менеджеров выглядит следующим образом.

Лента используется для систем резервного копирования, где необходимо регулярно создавать резервные копии больших объемов данных. Лента дешева, целостность хороша в краткосрочной и среднесрочной перспективе,а извлечение из резервной копии можно сделать приемлемым при хорошей практике хранения на ленте.

Ближняя память должна быть основана на стираемых оптических дисках. Это происходит потому, что доступ является случайным, поэтому скорость доступа для поиска и извлечения конкретного файла должна быть быстрой, а целостность данных также хороша.

Системы архивирования, вероятно, должны быть основаны на компакт - дисках или червях, поскольку опять же скорость доступа хороша, затраты на носители достаточно дешевы и, что немаловажно, целостность носителей в среднесрочной и долгосрочной перспективе хороша.

Одна важная вещь, которую следует помнить в системах архивирования, - это формат хранимых данных. Эти данные могут храниться идеально в течение 10 или 15 лет, но когда вам нужно получить их обратно, очень важно, чтобы вы поддерживали соответствующее оборудование и программное обеспечение, чтобы вы могли их прочитать.

Summary

1. This text is about storages of informations.
2. HSM is the modern way of deal with storage need.
3. It is essentially the automatic movement data between media.
4. The most common HSM setup is where there is online storage, near-line storage and offline storage.
5. It is recommend to use tape for backup systems, optical disks for near-line storages and CD or WORM for archiving systems.
6. Else important thing is format for kept data.
7. You should read it when you will need the data.
8. To sum up, I would say, that data storing is very important.

A.

1. data storage requirements
2. table:
   1. cheap, good integrity/ tape/ offline
   2. fast access, good integrity/ optical/ near line
   3. very fast/ CD, WORM/ online
3. It moved to nearline storage and to offline storages
4. Via a especial marker
5. it copied to the users`s system
6. reference to the type, where the file is on
7. systems administrator
8. digital, analog and digital linear
9. answers
   1. 2
   2. 2
   3. 1
   4. 2
10. recover speed and time you want to keep your data
11. hardware and software to enable you to read the data

B.

a false

b true

c true

d false

e false

f true

g true

h false

i false

j true