

Chapter 9

Distributed Storage

File location and addressing

What is a file?



12DistributedStorage.pptx	14/05/2015 16:57	Microsoft PowerP...	4,184 KB
mClocks.pptx	13/05/2015 08:23	Microsoft PowerP...	105 KB

Normally we collapse.
Concepts: name;
contents; gui.

What about the backup of this file?

How do we distinguish?

File is an area on disk which contains data. The name of a file is a pointer to that area.

Directory is **not** a region of a disk which contains files.

Now we are going to
have to carefully
distinguish

A directory is a file which contains the names of files and pointers to the place on disk where the information about those files resides.

A file name is actually a reference to a pointer.
It contains a path

/home/username/teaching

and a name

Lecture2.ppt

3 inodes

root /

Starting at a fixed position on the disk are the inodes

The inodes are the place on disk where the metadata about a file is stored. Name, size, modification date, owner, ...

inode 2 contains the information about the root directory /

The file pointed to by inode 2 contains the names and inodes of the files (and directories) in the root directory.

So accessing in order to do

ls /user/home/kyberd

- Accessing inode 2
- Reading the file pointed to by inode 2 and finding the inode of the user directory.
- Reading the file pointed to by this and finding the inode of the home directory.
- Reading the file pointed to by this and finding the inode of the kyberd directory
- Reading the file pointed to by this and listing the file names in the file

4 *inodes*

The full pathname of a file is in fact an access to a data structure

This data structure which provides a pointer to the physical location on disk.

It is an index - one might even say a database.

Photo album software

Store the photo on disk

Create an album “Tracker Solenoid”

Place the photo in the album

Create an album “Publicity 2014”

Place the photo in the album

The organisation software places the photos on disk but adds an abstraction layer to allow you to organise the photos.

When you delete a photo from an album does it disappear from disk?

If you directly delete the photo does it disappear from the album?

A photo album is a database, with a simple DBMS



How many copies do we have?

6 <i>Distributed files</i> Add some pictures	How do we name them? They can be anywhere in the distributed system. There will often be duplicate copies of the files. Ideally we support a heterogeneous directory structure A solution was developed by the World-wide LHC Computing Grid WLCG Local File System Catalog LFC: provides a virtual interface to storage gives meaningful names hides the file system differences manipulated by user It understands duplicate files and will normally return the closest one. (Highest bandwidth – least contention) – when a job requests data SRM: protocol which governs access to the storage disk and tape hide complexities different underlying structures allows user to reserve physical space give a lifetime to a file reserve space in advance physical copy – global namespace	Rather than talk in generalities talk about a real system which has enormous amounts of data spread across the world Makes no reference to the file location Distributed storage
--	---	---

7 Manipulation

Ways of accessing files.

LCG: commands
operate on the data.
some of them have side effects in the LFC

lcg-cr: uploads a file to the storage and
registers it in the LFC
when using the lcg applications you refer to
the files on the storage using the SRM protocol

LFC:
manipulates the database
lfc-mkdir: creates a directory in the lfc (the
database) – but does affect an SE

you user interface machine connects to the
correct LFC server using the LFC_HOST logical name

Conventionally every VO has /grid/<vo> as a top
level directory

Can also define an LFC_HOME logical name

8 Manipulation

lgc-cr copy and register

```
lgc-cr --vo CMS
```

```
  -d srm://srm.grid.xx.yy:8443/<directory_Path><name>
```

```
  -l lfn:/<path>/<file>
```

```
  "file://<path><localFile>"
```

srm instruction says where on the data storage machine to place it

lfn: instruction says how to enter the name in the LFC data base

file: is the name on your local machine (UI)

The command creates a GUID – which is a hexadecimal identifier for this file.

unique identification
number

Replication

- shorten data transfer paths

- provide backup against data loss

- protect against system failure

The name of the file on a particular storage elements is referred to as the SURL – (SRM uniform resource locator)

Distributed storage

Automatic

This allows data in files to be automatically replicated as required.

Site which needs access to some data. The system can decide not to stream the data to that site, but make a copy available at a new storage location.

Information on that location associated with meta-data for that data set and future accesses will have the option of using that replica.

"The name of the song is called '*Haddocks' Eyes*'."

"Oh, that's the name of the song, is it?" Alice said, trying to feel interested.

"No, you don't understand," the Knight said, looking a little vexed. "That is what the name is *called*. The name really is '*The Aged Aged Man*'."

"Then I ought to have said 'That's what the *song* is called?' " Alice corrected herself.

"No, you oughtn't: that's quite another thing! The *song* is called '*Ways And Means*': but that's only what it's *called*, you know!"

"Well, what is the song, then?" said Alice, who was by this time completely bewildered.

"I was coming to that," the Knight said. "The song really is '*A-sitting On A Gate*': and the tune's my own invention."

Alice in through the looking glass.
Lewis Carroll



The song is

The song is called

The name of the song

The name of the song is called

Transparency

Like Alice we tend to mix up a thing and the name of a thing.

When we are planning distributed computing we need to realise.

The object is different from the name

We would like a way of referring to the contents of the file – which should make no reference to its physical location – replicas have the same name.
“Distributed Storage Lecture”

The management software might like a compact numeric way of referring to the same concept.

We need a way to talk about a particular replica

In the grid since a single file may be copied to number of temporary places simultaneously we create a separate identify for this transfer event