ECFS

ECMWF's **F**ile **S**torage system

Bojan Kasic
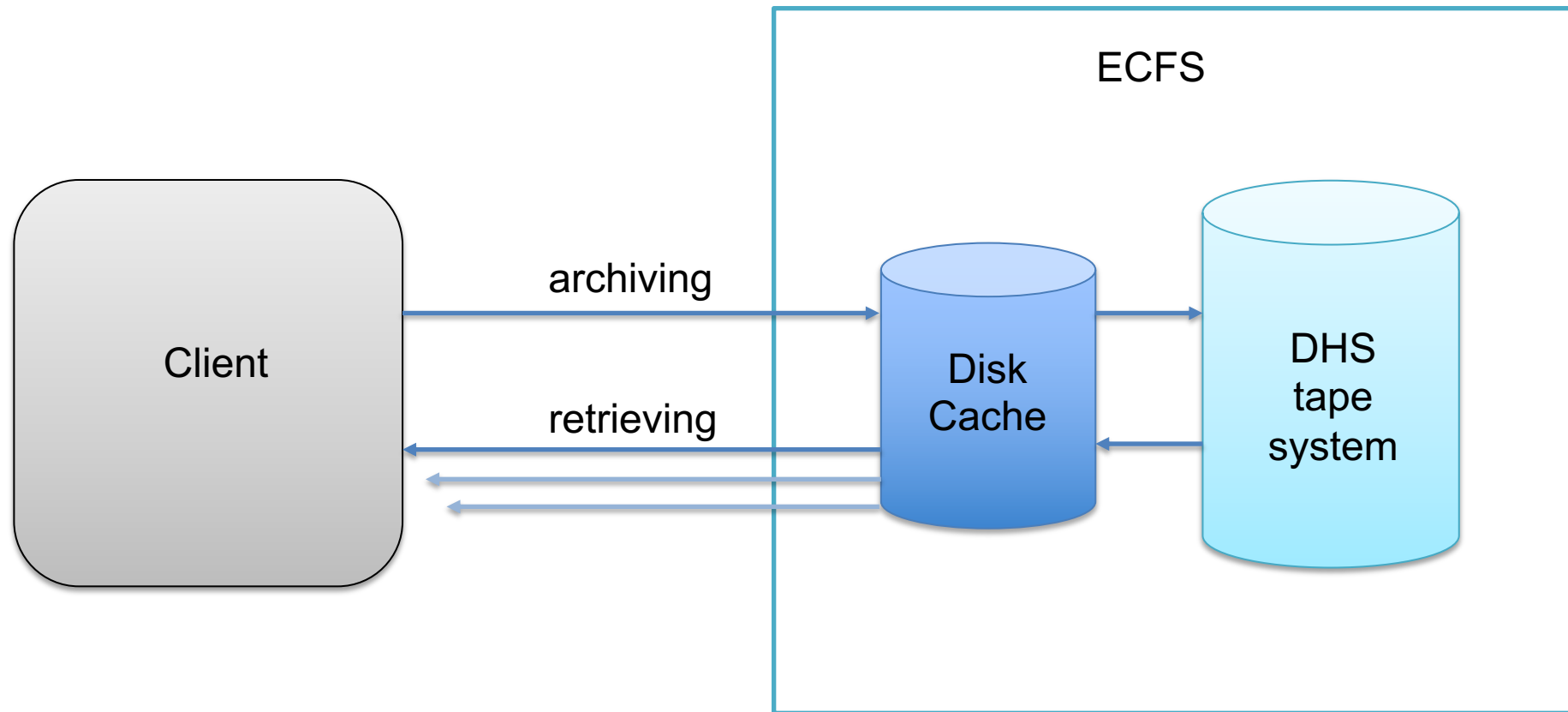
bojan.kasic@ecmwf.int

ECMWF

Overview

- 181 PB of data (~25% of total archived data at ECMWF)

- 545 M files (96% of total archived files at ECMWF)

- 324 MB/file

- 80 PB is Member State data in primary, **less than 2 TB of backup data**

- Total MS data in MARS is 7 PB

- Total MS DHS allocation for 2023, primary data is 92 PB and **37 PB for backup**

- ECFS can store any data format

ECMWF   EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

ECFS Architecture

- Client-server application

- ECFS has two domains where the files can be stored:

- Member State users have access to both systems

- Co-operating state users have access to ectmp only

ec:

permanent

ECFS

ectmp:

temporarly domain
Files stored for 90 days

The ECFS architecture

ECMWF   EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

ECFS client / commands

ECFS command	Description
els	List ECFS files
erm	Remove ECFS files
emkdir	Create empty ECFS directories
ecd	Change the current ECFS working directory
epwd	Display the current ECFS working directory for the relevant domain (ec: or ectmp:)
echmod	Change the permissions of a ECFS file or directory
echgrp	Change the group ownership of an ECFS file or directory
ecp	Copy files across ECFS domains, including UNIX
emv	Move files across ECFS domains, including UNIX
ecat	Copy files between a domain and STDIN/STDOUT
etest	Check file types and compare file attributes
etouch	Change file timestamps
emove	Rename files or directories within an ECFS domain
ecfs_status	Get status on ECFS usage
ecfsdir	Archive or retrieve a complete UNIX directory as one ECFS file

ECFS commands and options

- https://confluence.ecmwf.int/display/UDOC/ECFS+user+documentation

- man <ECFS command>

- Backup support option: Backup copy is not created by default!

```
ecp –b {filename} ec:{path}/

els –l
brw-r-----   1 username    group      5646 Nov  1 18:44 {filename}
```

- **"-b" option** also supported with **emv** and **ecfsdir** commands

- **erm** does soft-delete only. It is possible to recover deleted data for a limited time (1 month). You can contact https://confluence.ecmwf.int/site/support if you accidently delete any ECFS files and want them to be restored.

ECFS commands and options

- Wildcard is supported for rightmost ECFS file element of a path, but not for the directories in the path

```
>ecd ec:dir1/dir2
>ecp file* $SCRATCH/
```

```
>ecp ec:dir1*/dir2/filename $SCRATCH
```

- "--parents" option with ecp:

```
>ecp --parents subdir1/subdir2/file  ec:dir/
The target pathname is "ec:dir/subdir1/subdir2/file"
```

- ECaccess supports direct file transfers from ECFS and your local computer

ECFS commands

- echmod

> To make your file readable to everyone
> >echmod 644 ectmp:myfile

- eumask

> >eumask 022

- echgrp

> >echgrp mysecgrp ectmp:folder/

- ecfsdir

> >ecfsdir $SCRATCH/dir/ ectmp:model/dir_backup

Optimising ECFS read access

- Downloading files in the order they are written to the taps with "--order=tape"

```
>ecd ec:dir1/dir2
>els -l --order=tape
file1          volser:J12345
file2          volser:J12345
file3          volser:J11111
dir3/file4   volser:J12345
dir4/file5   volser:J12345

>cat>sourcelist<<eof
ec:file1
ec:file2
ec:file3
ec:dir3/file4
ec:dir4/file5
eof

>ecp --order=tape –F sourcelist $SCRATCH
```

ECMWF   EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Tips

- https://confluence.ecmwf.int/display/UDOC/ECFS+user+documentation

- Archive large files. The file size limit os 128GiB. Always tar your small files to a larger file before archiving in ECFS

- Do not create double archive of files already available in MARS or ECFS

- Do not transfer the same files frequently

- If you need to store files for short period of time, archive them in **ectmp**

- Think about who needs to read your files (permissions, eumask)

- Avoid using ECFS commands in parallel jobs on HPCF

- Use "-b" option for files which cannot be recreated such as your programs

- Do not use "-b" for model outputs which can be re-created

Questions?

Practicals:

- https://confluence.ecmwf.int/display/UDOC/MS+computing+training+course+practicals