

# Fixity Checking a Large Climate Data Archive

#### Sam Pepler

**Fixity** - the state of being

unchanging or permanent

Centre for Environmental Data Analysis, STFC Rutherford Appleton Laboratory, Didcot, United Kingdom

# Motivation

Climate and Earth Observation sciences rely on consistent large data sets. The Centre for Environmental Data Analysis (CEDA) has an archive of over 3PB of data, the bulk of which is the output from climate models and numerical weather prediction systems, and earth observation datasets.

It is important to make sure they these data are fixity checked because:

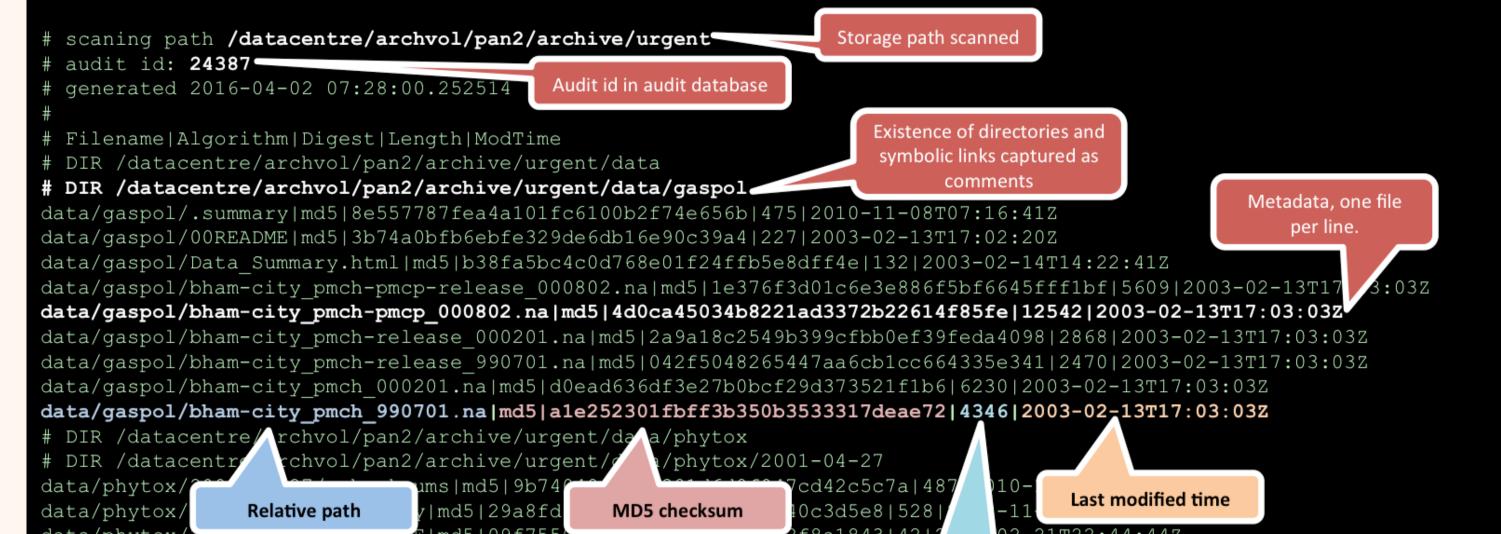
- Corruption can disrupt of bulk processing or services that rely on data consistency.
- Checks gives a quantifiable measure of trust in repository systems.
- Auditing provides traceability for data results, for example proving a file was the same 3 years ago.

### File sets

The CEDA archive is divided into file sets for the purpose of file management operations, such as backup, media migration and file fixity checking. File sets are nonoverlapping; every file is in one and only one file set. To ensure practical management each file set is limited to no more than 40TB in size and no more than around 3 million files. File sets going over these limits are split into small units.

#### **Checkm files**

Checkm files are used as a simple file manifest format, see example below. This provides a record of archive state and can be compared to the previous audit.



## Audits

An audit is a complete read operation of all the files in a file set. We aim to audit all file sets every 3 months. The audit provides a record of what is in the archive, a check for file corruption and flags files with other problems, for example bad permission attributes.

#### The audit cycle

Every hour an audit is submitted to the queue. For each audit:

- 1. The file set with the least up to date audit is found.
- 2. The file set is divided into file lists so that no list has more than 100000 files or is bigger than 100GB.
- 3. Each file list is used as input for a job constructing a portion of the checkm file.
- 4. A compilation process stiches together the final checkm file and submits it to the audit database.
- 5. The results are compared to the last audit, noting numbers of new, modified, deleted and corrupt files.
- A daily digest is emailed to the archive manager drawing attention to problems.

lata/phytox/2<del>001-04-27/00KLADM</del>E|md5|09f7550<del>2Cauresroaerreoao</del>2f8a1843|4<u>2|</u> data/phytox/2001-04-27/car\_pm10-instillation\_010430.na|md5|4b44140c8 data/phytox/2001-04-27/car pm10-instillation\_010430.pdf|md5|799b70736

<u>02-21T22:4</u>4:44Z Size

Audit report for /badc/cmip5/data/cmip5/output1/MOHC/HadGEM2-A/amip

e63|8176|2003-02-22T22:25:03Z 44|141|2003-02-21T22:57:23Z

Welcome, <mark>Sam</mark>. Change password / Log ou

data/phytox/2001-04-27/car pm10-toxicity 010430.na|md5|685523916de702aebc07df197fab651a|6468|2003-02-22T22:11:38Z data/phytox/2001-04-27/car\_rat-weight\_010427.na/md5/098e5e4f6e4eef006794d2fb7e4e5a63/1816/2003-02-22T22:10:54Z data/phytox/00README|md5|c19ece966db83465cf1441886315aa11|318|2003-02-21T22:51:47Z

**CEDA database administration** 

CEDAinfoDB | UserDB | CEDA Team Trac | Security Database

#### File set and audit database

These screen shots from the CEDA database show how information from audits is recorded

and analysed.							Audited fileset				
							Storage /datacentre/archvol/pan41/archive/spot-247 path				
Section CEDA database administration							Fileset /badc/cmip5/data/cmip5/output1/MOHC/HadGEM2-A/amip				
CEDAinfoDB   UserDB   CEDA Team Trac							Fileset Allocated partition /disks/fir1 (24 Tb free) (2011-05-05 10:52:29.990030) Reduced allocation after consultation with Met Office. Ag notes Stephens (2012-01-19) Migrated /disks/fir1 (19 Tb free) -> /datacentre/archvol/pan41 (0 Tb free) (2012-09-24 15:09:18.698475) [do_audit=False] Audit 20202 flagged corruption. This was resolved and the audit flag cleared on 20160413-0820. Replaced corrupt file from backup.				
Home > Cedainfoapp > Audits											
Select audit to change											
Search							Partition /datacentre/archvol/pan41 (7 Tb free)				
Action: $\frown$ Go 0 of 100 selected							Storage- True (last done:Fileset: spot-2472-06 Processing completed: 201606201004 )				
	Pk	Fileset	Auditstate	Starttime	Total files	Corrupted file	5				
	25673	/badc/cmip5/data/cmip5/output1/NOAA-GFDL/GFDL- CM2p1/decadal2010	analysed	16-May-2016 12:25		0	Audit comp	Audit comparision			
	25672	/badc/cmip5/data/cmip5/output1/NOAA-GFDL/GFDL-	analysed	16-May-2016 11:24	494	0		This Audit	Last Audit		
	25671	CM2p1/decadal1990 /badc/cmip5/data/cmip5/output1/NOAA-GFDL/GFDL-	analysed	16-May-2016 10:24	494	0	Audit number	20202	17825		
	23071	CM2p1/decadal2009	anaryseu	10-May-2010 10.24	494	0	State	analysed	analysed		
	25670	/badc/cmip5/data/cmip5/output1/NCAR/CCSM4/amip4xCO2	analysed	16-May-2016 09:24	60	0	Total files	7084	7084		
C	25669	/badc/cmip5/data/cmip5/output1/MOHC/HadCM3/decadal1997	analysed	16-May-2016 08:24	1540	0	Total volume	3877396398772	3877396398772		
	25668	/badc/cmip5/data/cmip5/output1/MOHC/HadCM3/decadal1994	analysed	14-May-2016 14:24	1540	0	New files	0	0		
	25667	/badc/cmip5/data/cmip5/output1/MOHC/HadCM3/decadal1993	analysed	14-May-2016 13:24	1540	0	Deleted files	0	0		
	25666	/badc/cmip5/data/cmip5/output1/MOHC/HadCM3/decadal1992	analysed	14-May-2016 12:24	1540	0	Modified files	0	0		
	25665	/badc/cmip5/data/cmip5/output1/CCCma/CanAM4/amipFuture	analysed	14-May-2016 11:24	158	0	Unchanged files	-	7084		
	25664	/badc/cmip5/data/cmip5/output1/MOHC/HadCM3/decadal1991	analysed	14-May-2016 10:24	1540	0	-	03-Oct-2015 15:23			
C	25663	/badc/cmip5/data/cmip5/output1/BCC/bcc-csm1-1/rcp26	analysed	14-May-2016 09:24	860	0	Start time				
	25662	/badc/cmip5/data/cmip5/output1/MOHC/HadCM3/decadal1989	analysed	14-May-2016 08:24	1540	0	End time	03-Oct-2015 15:49	-		
	25661	/badc/cmip5/data/cmip5/output1/MOHC/HadCM3/decadal1988	analysed	14-May-2016 07:24	1540	0	Corrupt files	1	0		
	25660	/badc/cmip5/data/cmip5/output1/NOAA-NCEP/CFSv2- 2011/decadal2005	analysed	14-May-2016 06:24	443	0	<b>corrupt files</b> 6hr/atmos/6hrPlev/r5i3p1/files/ua_20110630/ua_6hrPlev_HadGEM2-A_amip_r5i3p1_2007120106- 2008120100.nc md5 c4a73f8dfeb8d7b0c5ecbd7571f787dd 477786188 2011-06-30T14:10:05Z				
	25659	/badc/cmip5/data/cmip5/output1/BCC/bcc-csm1- 1/decadal2005	analysed	14-May-2016 05:24	652	0					
	25658	/badc/cmip5/data/cmip5/output1/CCCma/CanAM4/amip	analysed	14-May-2016 04:24	1518	0					

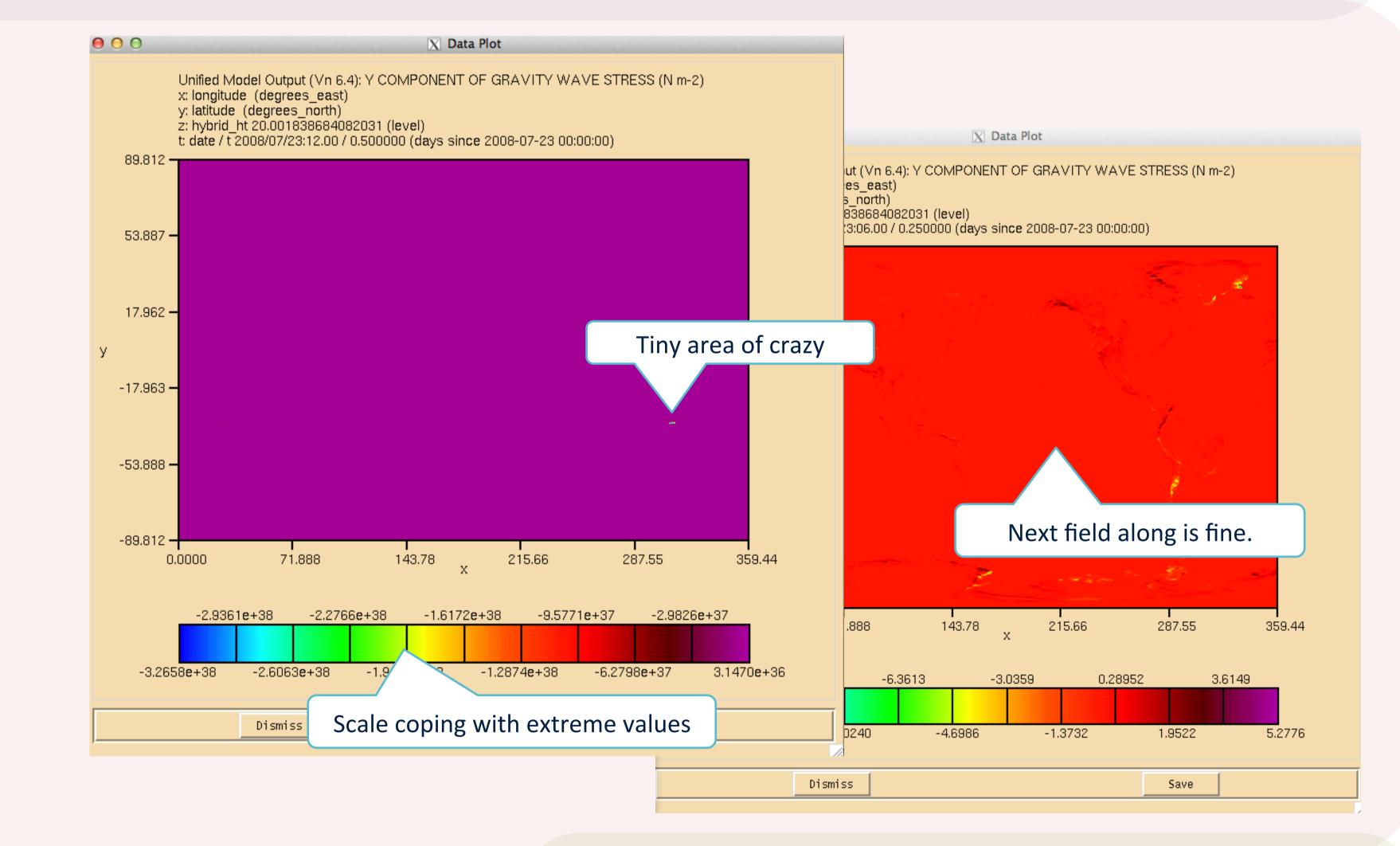
#### Results

In 2014, 13 corrupt files were detected, and in 2015, 10 corrupt files were detected among the 180 million files in the archive. It is hard to pinpoint the causes of the corruption. In the past year several corruption events have shown a 2k block of bad data suggesting a systematic problem. It is still relatively rare, and easily recoverable from backups.

The images shown here are an example of how the corruption can manifest in a single area of a single field in file.

#### Conclusions

- It is important not to simply rely on the internal integrity checking mechanisms in hardware systems. These systems are fallible like everything else.
- Corruption is rare on the current JASMIN storage system.
- Future systems need to be monitored closely if we are to have confidence in their fixity.





#### **Further Information**

#### Hardware Used

- JASMIN offers petascale storage, batch compute and cloud computing for big data challenges in environmental science, hosted at STFC Rutherford Appleton Laboratory in the UK.
- Lotus is the batch compute service provided JASMIN, with around 3000 cores available.
- The CEDA archive uses the Panasas petascale parallel storage as the primary storage for is archive.
- The CEDA info database is hosted on a virtual machine hosted on JASMIN
- JASMIN data analysis facility: <u>http://jasmin.ac.uk/</u>
- CEDA: http://www.ceda.ac.uk/
- Sam.pepler@stfc.ac.uk
- Checkm spec: J. Kunza, S Abrams, D Loy, Chechm: a checksum-based manifest format (v0.7), 2010, http://www.cdlib.org/services/uc3/ docs/checkmspec.html





NATURAL ENVIRONMENT RESEARCH COUNCIL



#### **Centre for Environmental Data Analysis**

SCIENCE AND TECHNOLOGY FACILITIES COUNCIL NATURAL ENVIRONMENT RESEARCH COUNCIL