Module 5: Data Management
EVA meeting, Bergen 8 May 2014

Ingo Bethke

EVA

uni Research

Data Storage

Resources in EVA

- EVA depends on the national data storage facilities (NorStore)
- NorStore allocations are per calendar year
- requested storage in Tb (without backup):

	2014	2015	2016	2017
DISK	159	180	334	769
TAPE	334	+278	+798	+1614

- uncertain that NorStore can meet EVA's storage demands

Data Storage

Resources in EVA

- EVA depends on the national data storage facilities (NorStore)
- NorStore allocations are per calendar year
- requested storage in Tb (without backup):

	2014	2015	2016	2017
DISK	159	180	334	769
TAPE	334	+278	+798	+1614

- uncertain that NorStore can meet EVA's storage demands

Current availability with "dusage -p ns2345k"

```
=========================================================================
Project      Account    Resource       Type    Usage     Limit
NS2345K      PROJECT    norstore_osl   Disk    147TB     200TB
NS2345K      PROJECT    tapestore      Tape    130TB     334TB
NS2345K      PROJECT    norstore_osl   Files   3264780   None
NS2345K      PROJECT    tapestore      Files   8170      33400
-------------------------------------------------------------------------
```

→ plan to move Earthclim legacy data to new National Data Archive

Data Storage

How do we internally manage the resources?

- allocation per project requires discipline from users
- keep your WP leader updated on your storage needs
- communicate storage issues as follows:
  researchers $\rightarrow$ WP leaders $\rightarrow$ project leader $\rightarrow$ NorStore
- reduce **DISK** storage where possible:
  1) consciously select output parameters, 2) use compression,
  3) migrate data to tape, 4) delete obselete data
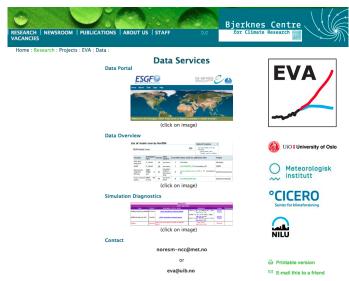
Data Storage

How do we internally manage the resources?

- allocation per project requires discipline from users
- keep your WP leader updated on your storage needs
- communicate storage issues as follows:
  researchers $\rightarrow$ WP leaders $\rightarrow$ project leader $\rightarrow$ NorStore
- reduce **DISK** storage where possible:
  1) consciously select output parameters, 2) use compression,
  3) migrate data to tape, 4) delete obselete data

Tape and compression guides

http://wiki.met.no/noresm/norstoretape

http://wiki.met.no/noresm/noresm2nc4mpi

Data Services

EVA hompage: `http://www.bjerknes.uib.no/pages.asp?kat=192&lang=2`

Data Services

Data Portal `http://noresg.norstore.no`

Data Services

Data Overview (Alf G.) `http://wiki.met.no/noresm/listofruns`

List of model runs by NorESM

Performed runs   Edit

Table of Contents ▲
· List of model runs by NorESM
   · Performed runs
   · Codes used in tables

Purpose	Simulation type	Contact	Disk (space)	Ensemble	Name (click for additional info)	Project
SO2 time evolution	C, 20thC	OS	norstore	Y	1850SOx	PEGASOS
CMIP	C, 20thC	OS	norstore	Y	N20TRAERCN_X (3 simulations X)	CMIP5
Aerosols & radiative forcing	AMIP, PD&PI (7 yrs)	AK	norstore (289 G) AeroCom disk	N	aerocomA2noresm_r128_X (7 simulations X)	EarthClim/AeroCom
Arctic aerosol optics	AMIP, PD (3 yrs)	AK	norstore (9 G)	N	NF2006to2008r164	EarthClim/ACCESS

Diagnostics (Detelina I.) `http://noresg.norstore.no/public`

Project EVA					
Fully Coupled Runs					
Name	Model	Description (click for details)	Diagnostics	Contact	Publications
N1850_f19_tn11_01_default	NorESM	AMOC sensitivity to isopycnal mixing	N1850_f19_tn11_01_default-Obs, 1-200y: Atm Ocn Ice Lnd	Detelina Ivanova	
N1850_f19_tn11_01_E12	NorESM	AMOC sensitivity to isopycnal mixing	N1850_f19_tn11_01_E12-Obs, 1-200y: Atm Ocn Ice Lnd N1850_f19_tn11_01_E12-default, 1-200y: Atm Ocn Ice Lnd	Detelina Ivanova	
MyExp	MyModel	MyExp Short Description (ReadME with details linked)	Type, Period: Atm Ocn Ice Lnd	MyEmail	

Data Sharing

Norstore project area

- access limited to members of Earthclim/Eva

Data Sharing

Norstore project area

- access limited to members of Earthclim/Eva

ESG data portal – standard publication

- default option for model intercomparison projects (e.g., CMIP)
- strict requirements on post-processing and publishing format
- $\rightarrow$ data duplication, requires time and human resources
- contact: alfg@met.no (Oslo), ingo.bethke@uni.no (Bergen)

Data Sharing

Norstore project area

- access limited to members of Earthclim/Eva

ESG data portal – standard publication

- default option for model intercomparison projects (e.g., CMIP)
- strict requirements on post-processing and publishing format
- $\rightarrow$ data duplication, requires time and human resources
- contact: alfg@met.no (Oslo), ingo.bethke@uni.no (Bergen)

ESG data portal – simplified publication

- data published "as is" with help of single command ($\sim$5 min)
- no need for post-processing $\rightarrow$ no data duplication
- no access control, i.e., all data freely downloadable
- "unlisted" option $=$ data downloadable but not exposed on portal

ESG data portal – simplified publication

Step 1: choose data folder or file that you want to share and find an acronym for the dataset

Example for data folder: `/projects/NS2345K/noresm/thredds/CORE2/ARCTIC`

Example for dataset id: `CORE2-ARCTIC`

ESG data portal – simplified publication

Step 1: choose data folder or file that you want to share and find an acronym for the dataset

Example for data folder: `/projects/NS2345K/noresm/thredds/CORE2/ARCTIC`

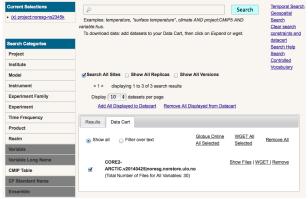Example for dataset id: `CORE2-ARCTIC`

Step 2: run publisher script

Syntax: `esgpublish <path to data> <dataset id> [unlisted]`

Example: `esgpublish /projects/NS2345K/.../CORE2/ARCTIC CORE2-ARCTIC`

ESG data portal – simplified publication

Access option 1: Use wget-script from `http://noresg.norstore.no/`
`thredds/fileServer/esg_dataroot/ns2345k/CORE2-AMOC/wget_CORE2-AMOC.sh`

ESG data portal – simplified publication

Access option 1: Use wget-script from `http://noresg.norstore.no/`
`thredds/fileServer/esg_dataroot/ns2345k/CORE2-AMOC/wget_CORE2-AMOC.sh`

Access option 2: Visit portal page `http://noresg.norstore.no`
and search for dataset with project-id=noresg-ns2345k and
dataset-id=CORE2-ARCTIC

ESG data portal – simplified publication

Access option 3: Browse dataset catalogue on thredds server

`http://noresg.norstore.no/thredds/esgcet/3/CORE2-AMOC.v20140505.html`

Dataset	Size
Catalog http://noresg.norstore.no/thredds/esgcet/3/CORE2-AMOC.v20140505.html	
📁 CORE2-AMOC	
NOIIA_T62_tn11_sr10m60d_01.cice.March_241-300.hi.nc	38.71 Mbytes
NOIIA_T62_tn11_sr10m60d_01.mld0030_march_241-300.nc	39.26 Mbytes
NOIIA_T62_tn11_sr10m60d_01.spg_heat_content_241-300.nc	1.104 Kbytes
NOIIA_T62_tn11_sr10m60d_01.atlantic_zonalmean_timemean_281-300.nc	142.1 Kbytes
NOIIA_T62_tn11_sr10m60d_01.atlantic_meridional_heat_flux_annual_241-300.nc	42.08 Kbytes
NOIIA_T62_tn11_sr10m60d_01.bering_strait_transport_annual_241-300.nc	1.044 Kbytes
NOIIA_T62_tn11_sr10m60d_01.micom.hy.241-300.Atlantic.msflx.nc	90.60 Kbytes
NOIIA_T62_tn11_sr10m60d_01.sst_annual_241-300.nc	39.26 Kbytes
NOIIA_T62_tn11_sr10m60d_01.aice_september_241-300.nc	39.26 Kbytes
NOIIA_T62_tn11_sr10m60d_01.gyreindexes_241-300.nc	2.004 Kbytes
NOIIA_T62_tn11_sr10m60d_01.micom.hy.1-300.Atlantic.mmflxd.nc	27.92 Kbytes
NOIIA_T62_tn11_sr10m60d_01.micom.hy.297-300.Atlantic.mmflxd.nc	4.932 Kbytes
NOIIA_T62_tn11_sr10m60d_01.micom.hy.1-300.Atlantic.mmflxl.nc	21.15 Kbytes
NOIIA_T62_tn11_sr10m60d_01.barotropic_streamfunction_annual_241-300.nc	45.34 Kbytes
wget_CORE2-AMOC.sh	2.769 Kbytes
NOIIA_T62_tn11_sr10m60d_01.mld0125_march_241-300.nc	39.26 Kbytes
NOIIA_T62_tn11_sr10m60d_01.depthmean_0-700_timemean_281-300.nc	8.849 Mbytes
NOIIA_T62_tn11_sr10m60d_01.atlantic_meridional_salt_flux_annual_241-300.nc	42.09 Kbytes
NOIIA_T62_tn11_sr10m60d_01.micom.hy.297-300.Atlantic.mmflxl.nc	4.704 Kbytes
NOIIA_T62_tn11_sr10m60d_01.micom.hy.241-300.Atlantic.mhflx.nc	90.59 Kbytes

Initial TDS Installation · My Group
THREDDS Data Server [Version 4.2.10 - 20120417.2151] Documentation

guide at `http://wiki.met.no/noresm/norstoreesg`