

# ERA-CLIM2: Service developments (WP5)



General Assembly  
Patrick Laloyaux on behalf of Manuel Fuentes  
16 January 2017



- Development of public data services
- NetCDF support in MARS
- Data services usage

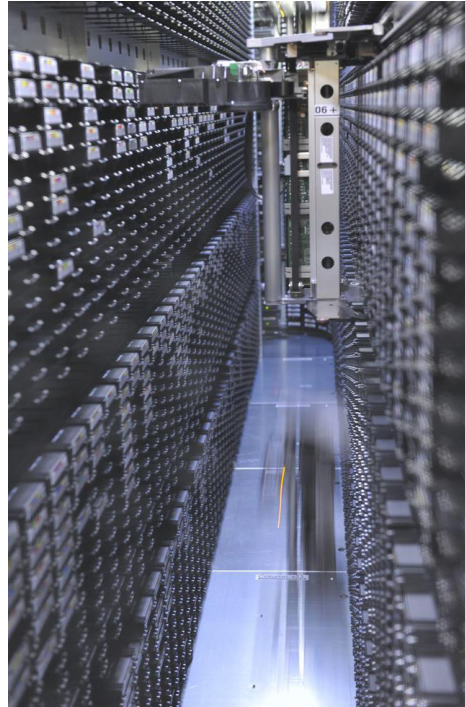
# Development of public data services (D5.2)

## CERA-20C atmospheric data in GRIB format archived on MARS (Meteorological Archive and Retrieval System)

- Data is accessed via a meteorological meta-language interface

retrieve,

```
class=ep,  
date=2004-01-01,  
expver=2379,  
levtype=sfc,  
number=0/1/2/3/4/5/6/7/8/9,  
param=34.128,  
stream=enda,  
time=00/03/06/09/12/15/18/21,  
type=an,  
target="data.grib"
```



## CERA-20C ocean data in NetCDF format archived on ECFS (ECMWF File System)

- large file system for files that are not suitable for storing in MARS
- UNIX-like commands enable users to copy data

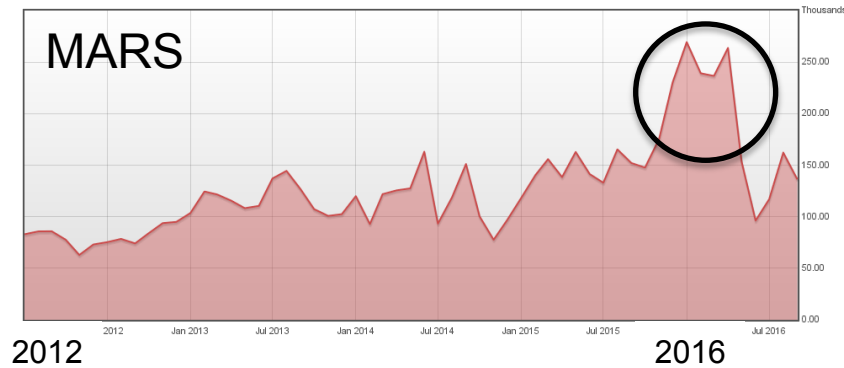
```
ecp ec:/ERAS/cera20c/${EXPID}/an/output/${YYYY}/* $SCRATCH
```

# Data archiving of CERA-20C reanalysis

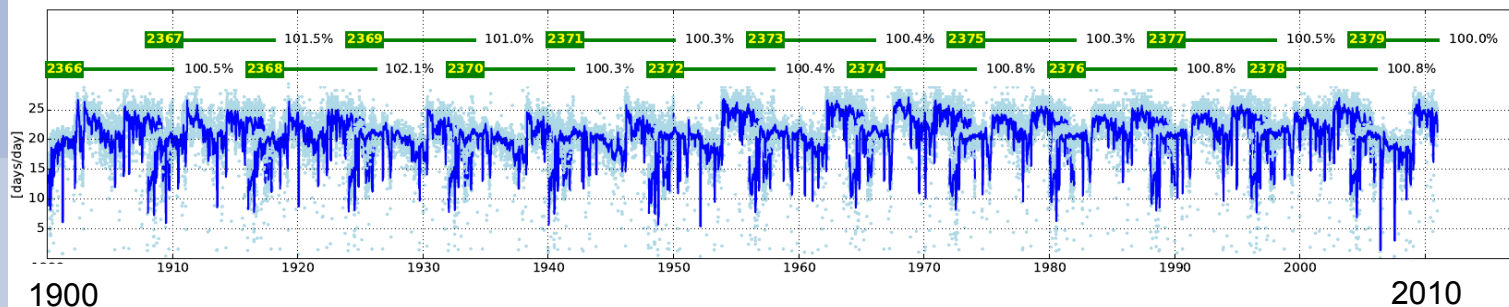
## Key facts and numbers:

- production from Nov 2015 to Jun 2016 using 5% of HPC system
- 500,000 4D-Var problems to solve, one every 30 seconds
- 1600 Tb of climate data (massive throughput for the file systems)

## Number of fields/files archived in MARS and ECFS



Production speed: lot of variability, many manual actions with on-the-fly optimisation for the different production streams



# Development of public data services

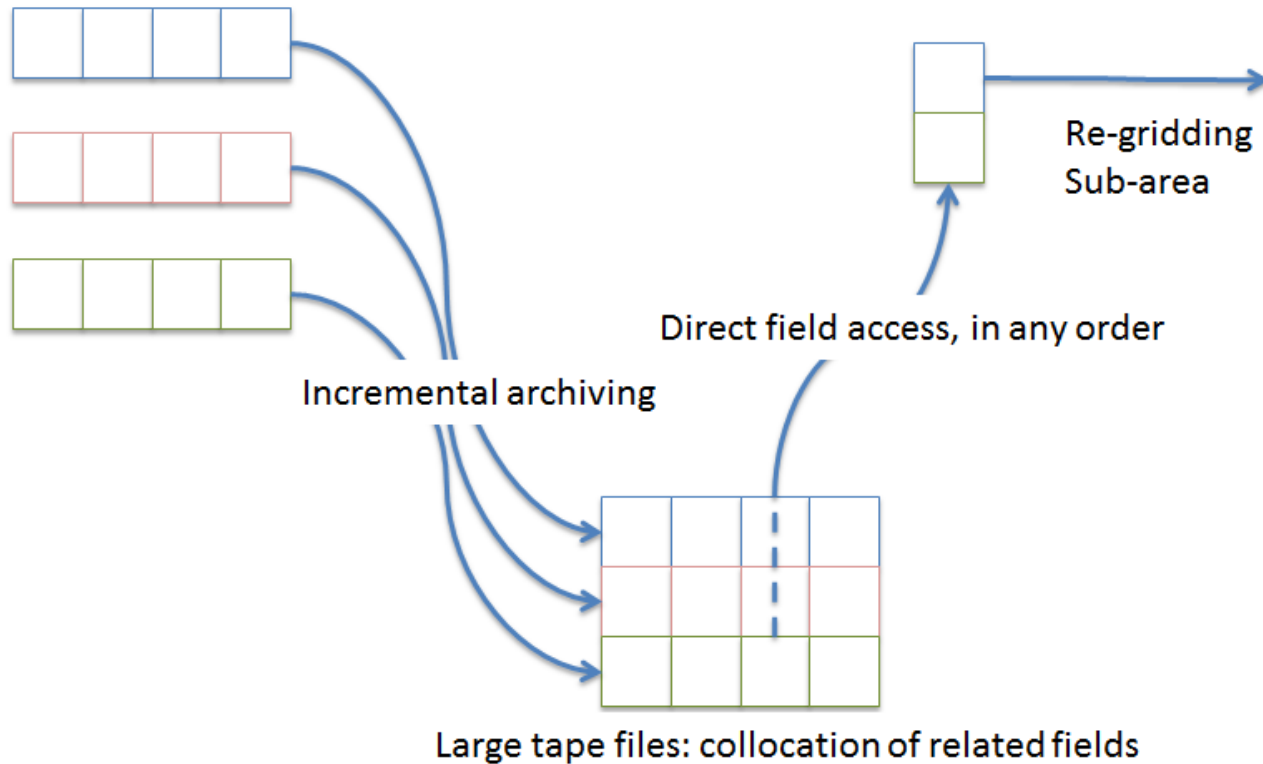
## Consolidation of CERA-20C into user version

Production is complete and quality of data has been checked, the various CERA-20C streams will be consolidated into a single version:

- Retrieve all relevant data (GRIB/NetCDF Fields, ODB feedback)
- Re-badge headers with the user version (0001)
- Archive back into MARS

This process is ongoing and expected to take several months

# Services offered by MARS



- MARS archives atmospheric field of CERA-20C (incremental archiving)
- CERA-20C consolidation is ongoing to offer the required data service: GRIBs are reorganised into larger files, to minimise the total number of files and collocate related fields
- On retrievals, MARS find the required fields, reads them from tape, and re-assembles them according to the user's request

# CERA-20C goes public for 1901–2010

## Browse reanalysis datasets

[Research homepage](#)

[Data Assimilation](#)

[Modelling and prediction](#)

[Climate reanalysis](#)

### Reanalysis datasets

[ERA-Interim](#)

[ERA-Interim/Land](#)

[ERA-20C](#)

[ERA-20CM](#)

[CERA-20C](#)

[Coupled Earth-system reanalysis](#)

[Reanalysis for climate monitoring](#)

[Ocean reanalysis](#)

[Projects](#)

[Publications](#)

[Special projects](#)

Dataset	Archive	Time period	Atmosphere	Atmospheric composition	Ocean waves	Ocean sub-surface	Land surface	Sea Ice	Observation Feedback Archive
<a href="#">ERA-Interim</a>	<a href="#">Download</a>	1979-present	✓		✓		✓		Expected soon...
<a href="#">ERA-Interim/Land</a>	<a href="#">Download</a>	1979-2010					✓		
<a href="#">CERA-20C</a>	Expected soon...	1901-2010	✓		✓	✓	✓	✓	✓
<a href="#">ERA-20CM</a>	<a href="#">Download</a>	1900-2010	✓		✓		✓		
<a href="#">ERA-20C</a>	<a href="#">Download</a>	1900-2010	✓		✓		✓		✓
<a href="#">ERA-20CL</a>	Expected soon...	1900-2010					✓		
<a href="#">ERA-40</a>	<a href="#">Download</a>	1957-2002	✓		✓		✓		
<a href="#">ERA-15</a>	<a href="#">Download</a>	1979-1993	✓				✓		
<a href="#">ORAS4</a>	<a href="#">Download</a> ↗	1958-2015				✓			
<a href="#">ORAP5</a>	<a href="#">Download</a> ↗	1979-2013				✓		✓	
<a href="#">ORAS5</a>	Expected soon...					✓		✓	

Public Data Server has been extended to include CERA-20C  
Public interface to data stored in MARS

# CERA-20C goes public for 1901–2010

## CERA-20C

[+ Expand all](#) [- Collapse all](#)

CERA-20C is the ECMWF 10-member ensemble of coupled climate reanalyses of the 20th century, from 1901-2010. It is based on the [CERA](#) assimilation system, which assimilates only surface pressure and marine wind observations as well as ocean temperature and salinity profiles. It is an outcome of the [ERA-CLIM2](#) project.

[▶ Product description](#)

[▶ Spatial and temporal resolution](#)

### ▼ Forecast steps

All forecasts are integrated daily, from 18 UTC, for +step hours. The significance of the forecast step depends on whether the forecast parameter is instantaneous or accumulated (from the beginning of the forecast):

Step	3	6	9	12	15	18	21	24	27
<b>Valid time, for instantaneous forecast parameters</b>	21UTC	00UTC next day	03UTC next day	06UTC next day	09UTC next day	12UTC next day	15UTC next day	18UTC next day	21UTC next day
<b>Accumulation period, for accumulated forecast parameters</b>	18UTC to 21UTC	18UTC to 00UTC	18UTC to 03UTC	18UTC to 06UTC	18UTC to 09UTC	18UTC to 12UTC	18UTC to 15UTC	18UTC to 18UTC	18UTC to 21UTC

[Research homepage](#)

[Data Assimilation](#)

[Modelling and prediction](#)

[Climate reanalysis](#)

[Reanalysis datasets](#)

[ERA-Interim](#)

[ERA-Interim/Land](#)

[ERA-20C](#)

[ERA-20CM](#)

**CERA-20C**

[Coupled Earth-system reanalysis](#)

[Reanalysis for climate monitoring](#)

[Ocean reanalysis](#)

[Projects](#)

[Publications](#)

[Special projects](#)



# CERA-20C goes public for 1901–2010



## Navigation

- [Home](#)
- [Public Datasets](#)
- [Job list](#)

## See also...

- [Access Public Datasets](#)
- [General FAQ](#)
- [WebAPI FAQ](#)
- [Accessing forecasts](#)
- [GRIB decoder](#)

## Public Datasets

Access to these datasets is provided free of charge. Terms and conditions may apply, please check with each individual dataset.

### Global Reanalyses

- ▶ [ERA5 test \(Jan 2016 - Feb 2016\) \*\*NEW\*\*](#)
- ▶ [CERA-20C \(Jan 1901 - Dec 2010\) \*\*NEW\*\*](#)
- ▶ [ERA-20C \(Jan 1900 - Dec 2010\)](#)
- ▶ [ERA-Interim \(Jan 1979 - present\)](#)
- ▶ [ERA-Interim/LAND \(Jan 1979 - Dec 2010\)](#)
- ▶ [ERA-20CM \(Jan 1900 - Dec 2010\) Final](#)
- ▶ [ERA-20CM \(Jan 1900 - Dec 2010\) Experimental](#)
- ▶ [ERA-40 \(Sep 1957 - Aug 2002\)](#)
- ▶ [ERA-15 \(Jan 1979 - Dec 1993\)](#)

### Observation Feedback

- ▶ [ERA-20C \(Jan 1900 - Dec 2010\)](#)
- ▶ [ISPD v2.2](#)
- ▶ [ICOADS v2.5.1 with interpolated NOAA 20CR feedback](#)

### Multi-model

- ▶ [S2S](#)
- ▶ [TIGGE](#)
- ▶ [TIGGE LAM](#)

### Atmospheric composition

- ▶ [MACC Reanalysis](#)
- ▶ [CAM5 Near-real-time](#)
- ▶ [CAM5 Global Fire Assimilation System](#)
- ▶ [MACC GHG flux inversions](#)
- ▶ [CAM5 GHG flux inversions](#)

### Miscellaneous

- ▶ [DEMETER Project](#)
- ▶ [ENSEMBLES project](#)
- ▶ [YOTC](#)



# CERA-20C goes public for 1901–2010

## Select time

00:00:00  03:00:00  06:00:00  09:00:00  12:00:00  15:00:00  18:00:00  21:00:00

[Select All](#) or [Clear](#)

## Select number

0  1  2  3  4  5  6  7  8  9

[Select All](#) or [Clear](#)

## Select parameter

- |   |   |
|---|---|
| <input type="checkbox"/> 2 metre dewpoint temperature                     | <input type="checkbox"/> 2 metre temperature                              |
| <input type="checkbox"/> 10 metre U wind component                        | <input type="checkbox"/> 10 metre V wind component                        |
| <input type="checkbox"/> 10 metre wind speed                              | <input type="checkbox"/> 100 metre U wind component                       |
| <input type="checkbox"/> 100 metre V wind component                       | <input type="checkbox"/> Albedo   |
| <input type="checkbox"/> Boundary layer height                            | <input type="checkbox"/> Charnock   |
| <input type="checkbox"/> Convective available potential energy            | <input type="checkbox"/> Forecast albedo                                  |
| <input type="checkbox"/> Forecast logarithm of surface roughness for heat | <input type="checkbox"/> Forecast surface roughness                       |
| <input type="checkbox"/> High cloud cover                                 | <input type="checkbox"/> Ice temperature layer 1                          |
| <input type="checkbox"/> Ice temperature layer 2                          | <input type="checkbox"/> Ice temperature layer 3                          |
| <input type="checkbox"/> Ice temperature layer 4                          | <input type="checkbox"/> Instantaneous eastward turbulent surface stress  |
| <input type="checkbox"/> Instantaneous moisture flux                      | <input type="checkbox"/> Instantaneous northward turbulent surface stress |
| <input type="checkbox"/> Instantaneous surface sensible heat flux         | <input type="checkbox"/> Lake bottom temperature                          |
| <input type="checkbox"/> Lake cover                                       | <input checked="" type="checkbox"/> Lake depth                            |
| <input type="checkbox"/> Lake ice depth                                   | <input type="checkbox"/> Lake ice temperature                             |
| <input type="checkbox"/> Lake mix-layer depth                             | <input type="checkbox"/> Lake mix-layer temperature                       |
| <input type="checkbox"/> Lake shape factor                                | <input type="checkbox"/> Lake total layer temperature                     |
| <input type="checkbox"/> Leaf area index, high vegetation                 | <input type="checkbox"/> Leaf area index, low vegetation                  |
| <input type="checkbox"/> Low cloud cover                                  | <input type="checkbox"/> Mean sea level pressure                          |
| <input type="checkbox"/> Medium cloud cover                               | <input type="checkbox"/> Near IR albedo for diffuse radiation             |

# CERA-20C goes public for 1901–2010

Instantaneous surface sensible heat flux

Lake cover

Lake ice depth

Lake mix-layer depth

Lake shape factor

Leaf area index, high vegetation

Low cloud cover

Medium cloud cover

Near IR albedo for direct radiation

Neutral wind at 10 m v-component

Sea-ice cover

Skin temperature

Snow density

Soil temperature level 1

Soil temperature level 3

Surface pressure

Total cloud cover

Total column liquid water

Total column rain water

Total column water

UV visible albedo for diffuse radiation

Volumetric soil water layer 1

Volumetric soil water layer 3

Lake bottom temperature

Lake depth

Lake ice temperature

Lake mix-layer temperature

Lake total layer temperature

Leaf area index, low vegetation

Mean sea level pressure

Near IR albedo for diffuse radiation

Neutral wind at 10 m u-component

Sea surface temperature

Skin reservoir content

Snow albedo

Snow depth

Soil temperature level 2

Soil temperature level 4

Temperature of snow layer

Total column ice water

Total column ozone

Total column snow water

Total column water vapour

UV visible albedo for direct radiation

Volumetric soil water layer 2

Volumetric soil water layer 4

[Select All](#) or [Clear](#)

Monthly means already available

Analysis and forecast available very soon (2-3 weeks)

# Example with a Python interface (Web API)

- Install a simple library
- Download data via scripts

```
from ecmwfapi import ECMWFDataServer  
server = ECMWFDataServer()  
server.retrieve({  
    'dataset' : 'cera20c',  
    'stream' : 'edmo',  
    'levtype' : 'sfc',  
    'param' : '165.128',  
    'date' : '20091201',  
    'type' : 'an',  
    'number' : 0,  
    'target' : "data.grib"  
})
```

- Development of public data services
- **NetCDF support in MARS**
- Data services usage

# NetCDF support in MARS (D5.1)

Original NEMO output files contain many variables (2D, 3D), feedback files, restart files, ocean observations, all annotated with NEMO's own convention

Technical developments in MARS to support archiving and retrieval of data in NetCDF

## Solution considered:

NetCDF files are split into individual NetCDF files, 2D or 3D

- Resulting NetCDF files must follow an agreed convention based on CF
- Resulting NetCDF files are annotated with MARS specific information. These attributes are used by MARS to index the NetCDF files, and treat them as simple binary records

On retrieval, those records will be assembled in a single NetCDF file to be delivered to the user

# Prototype of NetCDF Ocean output in MARS



[Home](#) [Chart dashboard](#) [Contact](#)

[Manuel Fuentes](#) | [Sign out](#)

[About](#) [Forecasts](#) [Computing](#) [Research](#) [Learning](#)

## Navigation

[Public Datasets](#)

[Job list](#)

[< Return to selection](#)

## Additional filtering

### Current request

*Stream:* Atmospheric model

*Parameter:* Ocean salinity

*Year:* 1900

*Number:* 1

*Month:* February

*Version:* 2366

*Type of level:* Depth

*Time:* 00:00:00, 03:00:00, 06:00:00, 09:00:00, 12:00:00, 15:00:00, 18:00:00, 21:00:00

*Date:* 19000203

*Type:* Analysis

*Class:* ERA-CLIM2 coupled reanalysis of the 20th-century (CERA-20C)

### The request will be done using the following attributes:

**Area:** Default (as archived) ([change](#))

**Grid:** 1x1 ([change](#))

[▲ Top of page](#)

copyright © ECMWF



# Support for NetCDF at ECMWF

Other projects require NetCDF support

- Sub-seasonal to Seasonal project (S2S) requires archiving in MARS of Ocean output from 11 production Centres (near real-time + reforecast)
- Other forecasting systems producing ocean output (HRES, ENS, ....)
- All data being served from ECMWF Data Portals within a common framework

Define common metadata that will allow tools to seamlessly handle all the above

- Define a MARS/NetCDF convention

## Challenges:

Define what variables to archive, focusing on user service

- not all output is interesting to users

Find CF standard names (not only for NEMO, but for any ocean field)

- About 60% variables have a CF standard name

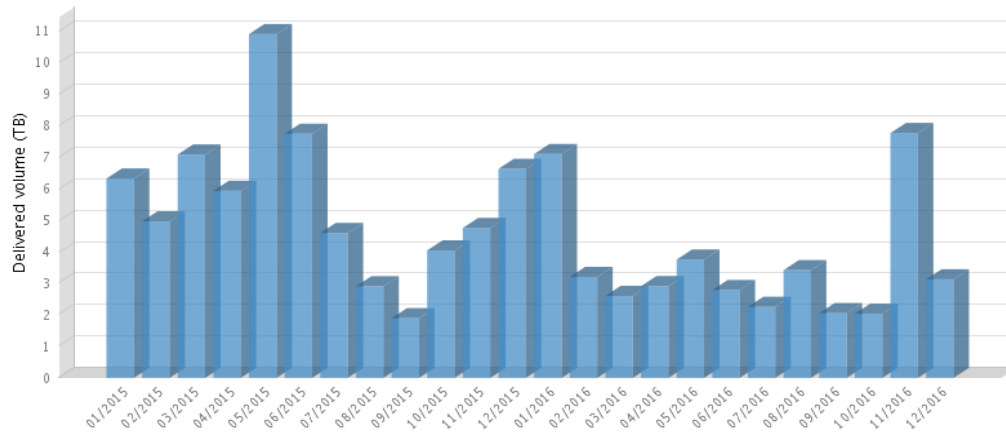
Find correct metadata to enable assembling records on retrieval (monthly means)

- Development of public data services
- NetCDF support in MARS
- **Data services usage**

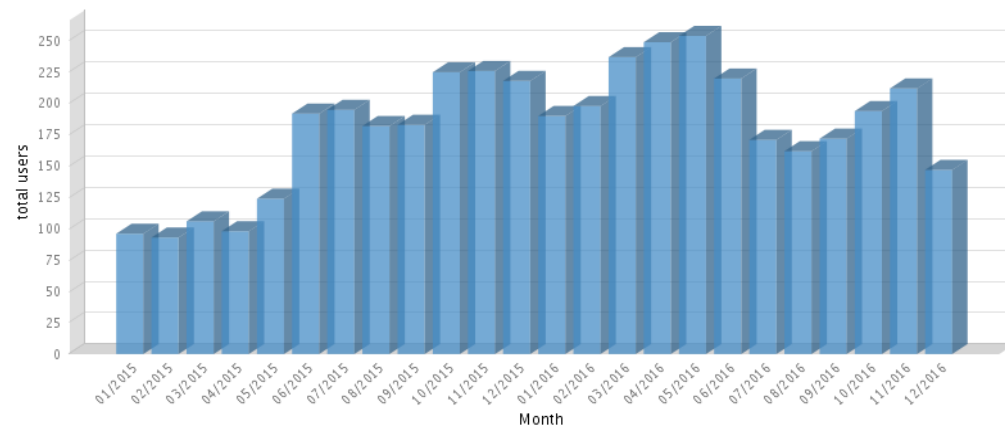
# Data service usage (D5.3)

- Establish user requirements for reanalysis data and measure service usage and performance
- Set up the users support
- Experience gained from ERA-20C released on 1 October 2014

Delivered Volume in TB



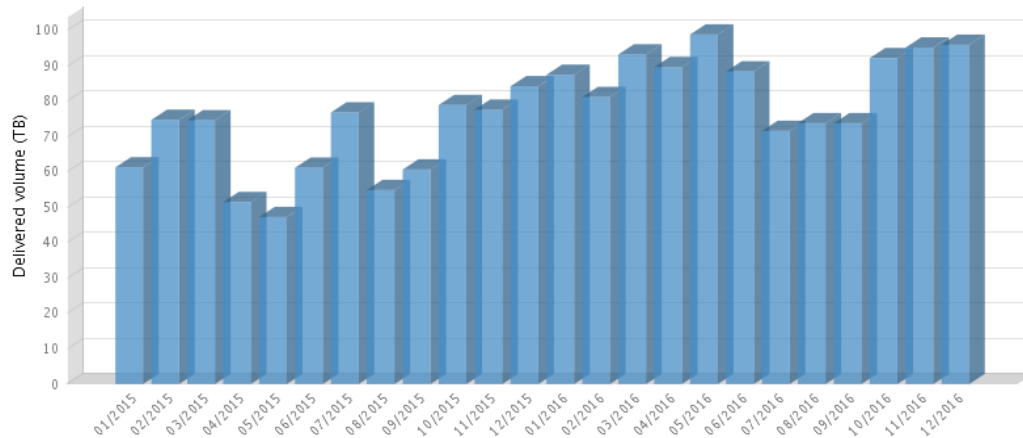
total users



# Data service usage

- Establish user requirements for reanalysis data and measure service usage and performance
- Set up the users support
- Experience gained from ERA-Interim

Delivered Volume in TB



Total users

