CDNOT
A developers point of view

Katharina Berger

DKRZ

December 7, 2017

ESGF Installation

- ▶ Very dependent on institutional setup
  - ▶ Different firewall, server configuration and hardware
  - ▶ Docker: Solution for server-based issues?
  - ▶ Develop pre-installation tests to identify firewall constraints
  - ▶ Be more flexible, e.g. option to download repos beforehand
- ▶ Identify problems during installation
  - ▶ Improve error messages
  - ▶ Develop fully automated post-installation tests
- ▶ Problems for certificate setup and renewal
  - ▶ Add 'new' (root-)certificates to ESGF truststore
  - ▶ Need notification service for expired certificates
- ▶ Establish a mentor program

ESGF Publication

- ▶ Initial setup of configuration files and permissions challenging
  - ▶ Make project configuration files optional for local projects?
  - ▶ Keep the vocabulary / config files centralized?
- ▶ Inflexible
  - ▶ Publish other data formats than netCDF, e.g. GRIB?
  - ▶ Publish additional documents (e.g. PDF) and link to data?
- ▶ Misleading error messages
- ▶ Use CDNOT mailing list to collect federation-wide requirements

User issues

- ▶ Issues with wget scripts
    - ▶ Improve error messages
    - ▶ '-H' does not show error messages by default
    - ▶ Missing group membership not obvious
    - ▶ Retire the old-style wget script? Possible replacement?
- ▶ GridFTP/Globus download not widely supported
    - ▶ Malfunctioned GridFTP urls published

General

- ▶ Most components are not well documented
- ▶ Need more automated tests
  - ▶ Pre-Installation 'dry-run' tests
  - ▶ Post-Installation tests
  - ▶ Post-Publication tests (including download and permissions)
- ▶ Operations
  - ▶ Setup automated health-check and certificate tests
  - ▶ Notification service
- ▶ Need more communication between ESGF and externally maintained components, e.g. THREDDS, Globus

Open questions

- ▶ Health-checks
  - ▶ Run locally on each node or centralized?
  - ▶ How to notify the admins?
  - ▶ Take off malfunctioned nodes from the federation (temporary)?
- ▶ Solutions for the (firewall-based) installation issues?
- ▶ Simplify configuration of new projects for publication?
- ▶ Replacement for wget script?
- ▶ Do we need more communication between working groups?
- ▶ How to announce new data and index nodes in the ESGF?