Versioning and Replication for CMIP6

Authors: Stephan Kindermann, Sebastien Denvil

Executive Summary

(Requirements summary for ESGF / CDNOT / modeling groups / data users)

ESGF infrastructure:

Requirement: Improvement of versioning support and control in the ESGF publisher
Priority: high (should be in production state when first CMIP6 publications start)
- CMIP6 requires a standard version string. Proposal: "YYYYMMDDxx" (year,month,day, consecutive number for multiple version publications at the same day, details are worked out and implemented in the publisher by the ESGF publication working team, concrete examples will be provided in the proposed publication best practices document)
- Enforcement of consistent version changes (e.g. only incrementation of versions possible)
- Map-file generation is required and enforced in the first publication step. Map-file includes versioning information
- Concise "CMIP6 Data publication and data versioning best practices" document which give clear instructions to CMIP6 publishing sites. This could be part of an overall CMIP6 data management document.
- Un-publication of individual versions should be possible
- Re-publication of older (then current) versions should be possible

Requirement: Agreement on versioning granularity in CMIP6
Priority: high
- Proposal: "atomic dataset level" = set of files describing time series of a variable (same granularity as used in the CORDEX project, finer granularity than in CMIP5)

Requirement: Decision on CMIP6 policy for access and storage of older versions of data
Priority: medium
- Is there a general CMIP6 guideline to store and provide access to older versions ?
- Probably different policies at ESGF publishing sites. Make these policies transparent to end users.

Requirement: Versioning consistency checking tools
Priority: medium
- Define concise checklist for data node managers with respect to versioning (consistency of file level, thredds/solr level)
- (could be operationally supported by a generic file/thredds/solr consistency checking tool, to be used by data centers (and CDNOT team))

Requirement: Define replication responsibilities of ESGF sites
Priority: medium

- Define "core replication sites" and their responsibilities (update frequency, spreading replication load across centers to ensure ESGF published data has replicas published at at least one site as early as possible..)
- Establish and organize replication team and replication team communication

CDNOT team:

Requirement: Enforce and overview consistent versioning at sites
Priority: medium
- Detect deviations from the policies outlined in the proposed "CMIP6 Data publication and data versioning best practices" document
- Contact data center with instructions to resolve inconsistencies
- Decide on actions to take if inconsistencies persist based on agreed upon measures.

Requirement: Establish a replication team (part of CDNOT?) and enforce and overview consistent replication at sites
Priority: high
- Close collaboration with ICNWG team

Data Users:

Requirement: easy access to version related information
Priority: medium
- Information about available data versions (via ESGF portal and search API)
- Information about no longer available data versions (via portal and search API)

Requirement: notification about version changes
Priority: medium
- Definition of version change information notification mechanisms for end users.
- Notification mechanisms of different granularity (and implementation complexity) are possible:
  - User specific mail notification
  - Generic mail notifications
  - Information feeds users have actively to subscribe to
- See also related requirement in errata system WIP paper

Requirement: easy exploitation of replica access possibility for end users
Priority: medium
- Generation of data access scripts with replica link alternatives included
- Possibility to tailor generation of data access scripts to specific regions to exploit replica sites from these regions (users from EUROPE etc.)
- Selection in GUI based download of preferred replica sites

- The replica notion could possibly be made transparent for data users. A replica will be presented to them if the "original" dataset is not available.

Modeling Center:

Requirement: provide information about version changes

Priority: high

- Communicate with ESGF publisher to provide information about data involved and about reason of new versions and the importance for end users -- details are outlined in the "CMIP6_errata_system" WIP paper (https://drive.google.com/open?id=1LNhw23tQotVomnkWQqfCPrBWspZi0bpdv2a1IegvHKI)

Motivation and Scope

Versioning for CMIP6

Keeping older data versions and maintaining versioning history information is of key importance for providing sustainable, trusted ESGF end user services for the distributed CMIP6 archive. Data retrieved from the CMIP6 data archive for data analysis by end users is often replaced by newer versions before the analysis results are published. On the other hand inconsistent data versioning related data management activities at data centers in the ESGF data federation, quickly leads to an unrecoverable inconsistent overall data space provided to end users (as well as long term archival sites). Thus use cases motivating consistent CMIP6 versioning originate both from the end user side as well as from the ESGF infrastructure side:

- Larger data centers need to keep older versions of important and often used datasets
- End users need transparent version information of accessed datasets as one important constituent to be able to generate trustable and reproducible results.
- A consistent CMIP6 versioning practice across the distributed ESGF sites is necessary to enable consistent overall versioning related services (e.g. search, end user notification, replication etc.)

A collection of versioning related requirements as well as solution approaches is collected and maintained in the (living) document at: https://docs.google.com/document/d/1tOaFQEXFyjqAOOIvcdiaX3XrXuxIv_nlE5_FCme1id4/edit)

Replication for CMIP6

Because of the decentralized architecture of ESGF replication is required to provide stable and efficient access to ESGF hosted CMIP6 data products. This paper summarizes the key requirements as well as necessary steps to move towards a stable replication procedure for CMIP6. Replication for CMIP6 is motivated by various use cases, e.g.:

- **Move data to larger ESGF sites providing stable, efficient and long term data storage and distribution services.** Often data is first published at smaller ESGF sites administered by local research groups. Replication to larger well established data centers with clear data

management processes, service level agreements and skilled data managers improves data accessibility.

- **Improve stability of data access in case of individual data server problems.** Data servers face regular downtimes as well as infrequent emergency situations. Data replicas provide redundancy and the possibility to access data also in case of server problems.
- **Large scale data analysis.** End-users and research groups often need to access large portions of CMIP6 data. This normally involves the replication of ESGF files from various sites to a local CMIP6 file cache. Such caches often provide replicas to local research groups whereas they quickly become out of sync with the data available on ESGF. This can be avoided by automated replication tasks strongly depending on common versioning policies across ESGF sites.
- **Better data access performance and load distribution.** End users have the choice to download replicas of files which are nearest to their home institute. On the other hand download burden is distributed to multiple servers holding replicas, exploiting available network bandwidths better.
- **Reducing manual workload related to replication.** The number of ESGF nodes, number of data files and modelling groups is increasing with CMIP6, yet the available resources at each site remain limited. A promising strategy for the federation to deal with this is to unify replication by using the same tools and workflows and further automate replication tasks.

Goals

The goal is to define and establish stable and coordinated CMIP6 versioning and replication policies for ESGF sites.

Versioning:

Versioning is closely related to the low level (file level) data management activities at sites. It is also closely related to the ESGF publication procedure making new versions known to the ESGF infrastructure (thredds metadata and solr indexes). Inconsistent file level and ESGF level versioning practices at ESGF sites for CMIP5 resulted in a number of severe overall data space consistency problems which are difficult or even impossible to resolve. To enable long term sustainable version information a strict adherence to consistent ESGF versioning practices at ESGF sites is required (supervised by the CDNOT). Additionally an extension of the ESGF infrastructure is proposed, which stores and maintains version links as part of the PID metadata associated to CMIP6 data with PID based tracking ids.

Replication:

There is a need for agreements on what, when and where to replicate, including the definition of "core" replication sites and the consistency requirements for replica sets. On the other hand the technical infrastructure has to be established to better support automatic replication procedures; this includes the core replication software used as well as its operational integration into the ESGF CMIP6 data management procedures and close cooperation with the international network working team (ICNWG). The experiences gathered in CMIP5 replication at DKRZ, BADC, IPSL, PCMDI and NCI indicate the strong need for agreements and coordinated developments to be able to meet CMIP6 requirements.

Current status

Versioning:

Consistent and correct versioning currently strongly depends on established "good practices" at the individual data centers: Versioning related modifications at file level have to be accompanied by corresponding ESGF publication/un-publication activities, making the modifications known to ESGF. Thus, for example, currently there are "ESGF unknown" file modifications in ESGF (e.g. reflected in inconsistent checksums of offered files with respect to published ESGF metadata) and inconsistent versioning at ESGF level for some datasets as the version number is a freely configurable option in the ESGF publication process.

Replication:

ESGF currently only supports an "offline, on demand" replication procedure, where dedicated replication sites pull replica-sets from ESGF sites, then reorganize them to fit into their internal ESGF data organization structure and finally publish them as "replicas" into the ESGF data federation. No automatic replica synchronization or notification mechanisms are supported. Thereby in general original data can be unpublished or modified without effects on replica sites etc. Different software is in use at the different sites to support and manage parallel download streams, handle error situations, and check consistency requirements (e.g. checksums). A tool, currently in broader use, is named "synda" and was developed at IPSL (https://github.com/Prodiguer/synda) . This tool is mostly used in an uncoordinated way to establish and maintain a "local" ESGF data set cache at institutes (a cache which is not published into ESGF and thus not visible as "replica site" in the ESGF data federation. Also replication software based on globus online is currently worked on.

Replica repositories which are not published to ESGF are also referred to as "dark repositories". We expect such repositories to persist into CMIP6.

Effort

Versioning

Managing and maintaining version information as part of the ESGF publication workflow (e.g. mapfile generation etc.) has to be improved for better consistency of the overall CMIP6 data storage; this is in work. Additionally versioning related documentation has to be maintained and made accessible; this aspect is covered in the proposed ESGF errata WIP document (https://docs.google.com/document/d/1uP_8HD9fPeH20wDWfJPIcqkI54kCj1hy_djmZ4ss994/edit). On the other hand consistent versioning strongly depends on "good practices" at the ESGF data publication sites. These practices have to be enforced and overviewed by the CMIP Data Node Operations Team (CDNOT).

The proposed persistent identifier infrastructure for CMIP6 data (see document https://docs.google.com/document/d/13VjI377yNRnBE9fHkAqRY5o-QlUQvlFaoTkPdwZaNio/edit) can be exploited to maintain versioning related metadata of datasets (at different granularity), thus enabling a sustainable backend to provide stable versioning related end user services (e.g. version history navigation of CMIP6 files). In order to achieve this, PID associated versioning metadata has to be maintained as integral part of the ESGF data publication/un-publication procedure. Automatic integrity checks can be implemented later on, in order to detect version information mismatches between file systems, ESGF metadata, and PID metadata.

Replication

Core sites have to join replication related efforts. In the first phase these include the old core replication sites like DKRZ, BADC, IPSL and PCMDI. Developments are needed to establish and support a data replication software infrastructure (e.g based on "synda" or/and globus online). The management of replication activities (what, when etc.) has to be coordinated in a team of data managers at core sites, closely cooperating with the ICNWG network team.
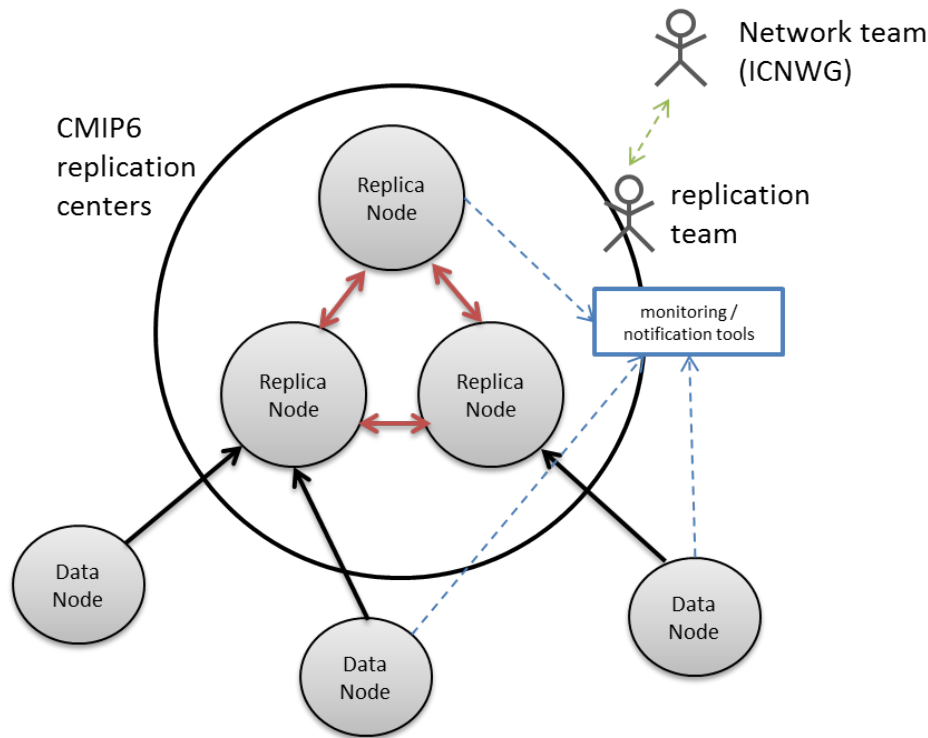

Figure 1: CMIP6 replication from data nodes to replica centers and between replica centers coordinated by a CMIP6 replication team

Developments and contributions to the ESGF software stack are required to establish the "hooks" necessary to support fully automatic replication procedures in the future. Therefore stable non-automatic replication between core sites is the first priority. The mentioned "hooks" mainly are related to file publication and versioning events which should trigger future automatic rule based replication procedures.

In figure 2 an overview of the core replication steps and components are illustrated:

- Data to be replicated is published to an ESGF index node
- Based on a specification of the data sets to be replicated (specific facet selections as well as e.g. version and date information) the replica set is determined (using the search API of the ESGF index node)
- The replication tool manages the network transport to the target data-transfer node.
- The replication site integrates the transferred data into their ESGF data pool (e.g. involving site specific data management activities)
- The replication site publishes the replica set to the ESGF index node

In general the data transfer involves dedicated "data-transfer nodes" which are separated from the target replica node for many operational reasons (security, performance, data center policies,..). Independent of the use of dedicated data-transfer nodes end-to-end data replication in general involves local data management activities at the replica nodes to integrate the transferred replica sets into their ESGF exposed data pool (e.g. consistency checks, directory structure adaptations etc.).
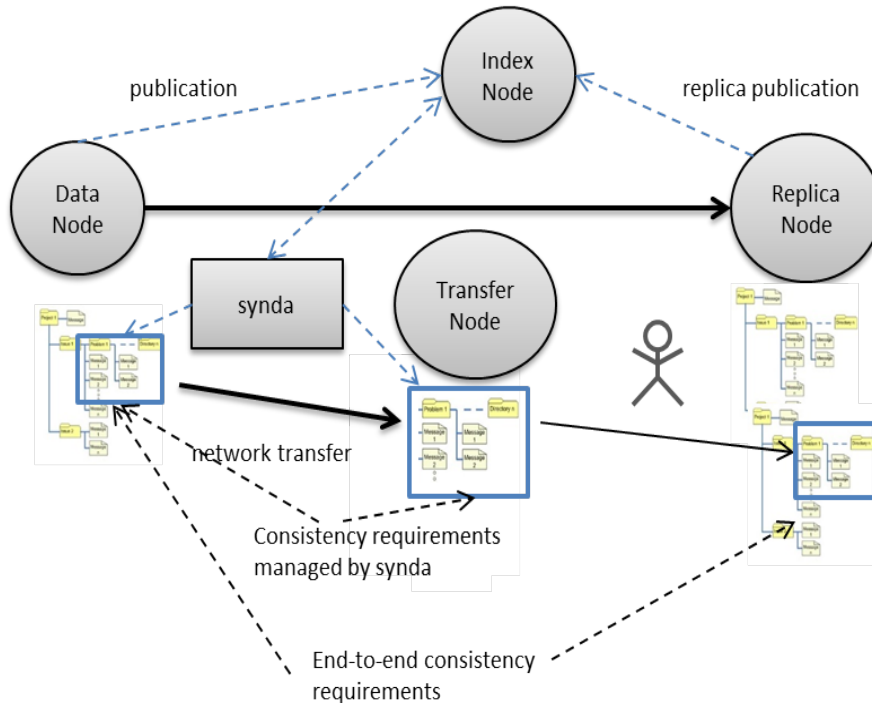

Figure 2: CMIP6 replication procedure involving data-transfer nodes

The proposed PID infrastructure will provide a stable backend for future PID based replication procedures, if replication metadata is maintained as part of the PID records. The only core requirement for this is the integration of the PID API (see the PID WIP white paper) into the publication process at replication centers in order to add replica links to the PID metadata of published data sets.

Related documents

https://docs.google.com/document/d/1tOaFQEXFyjqAOOlvcdiaX3XrXuxIv_nlE5_FCme1id4/edit

Contact

Stephan Kindermann (kindermann@dkrz.de)

Sebastien Denvil (sebastien.denvil@ipsl.jussieu.fr)

Contributions from: Katharina Berger, Guillaume Levavasseur, Ag Stephens.