Copernicus Climate Change Service

"Data Node software evolution for 2020"

Guillaume Levavasseur
(C3S_34a - WP3)

Copernicus

European Commission

Climate Change

ESGF installation

- Traditionally done by a mix of bash scripts and manual admin actions.
- ESGF nodes deployed for CDS have been managed this way.
- Recently, IPSL successfully upgraded its data node using the alternative approach based on the Ansible framework (developed by LLNL)
  - ➡ The load-balancing architecture designed for CDS supports the Ansible installation without any particular configuration.
- From the IS-ENES ESGF Workshop conclusions:
  - ➡ ESGF "bash installer" is now deprecated
  - ➡ Continue development of the ESGF Ansible installer to support Docker Images.

The Data Node software for CDS will now rely on ESGF Ansible installer.

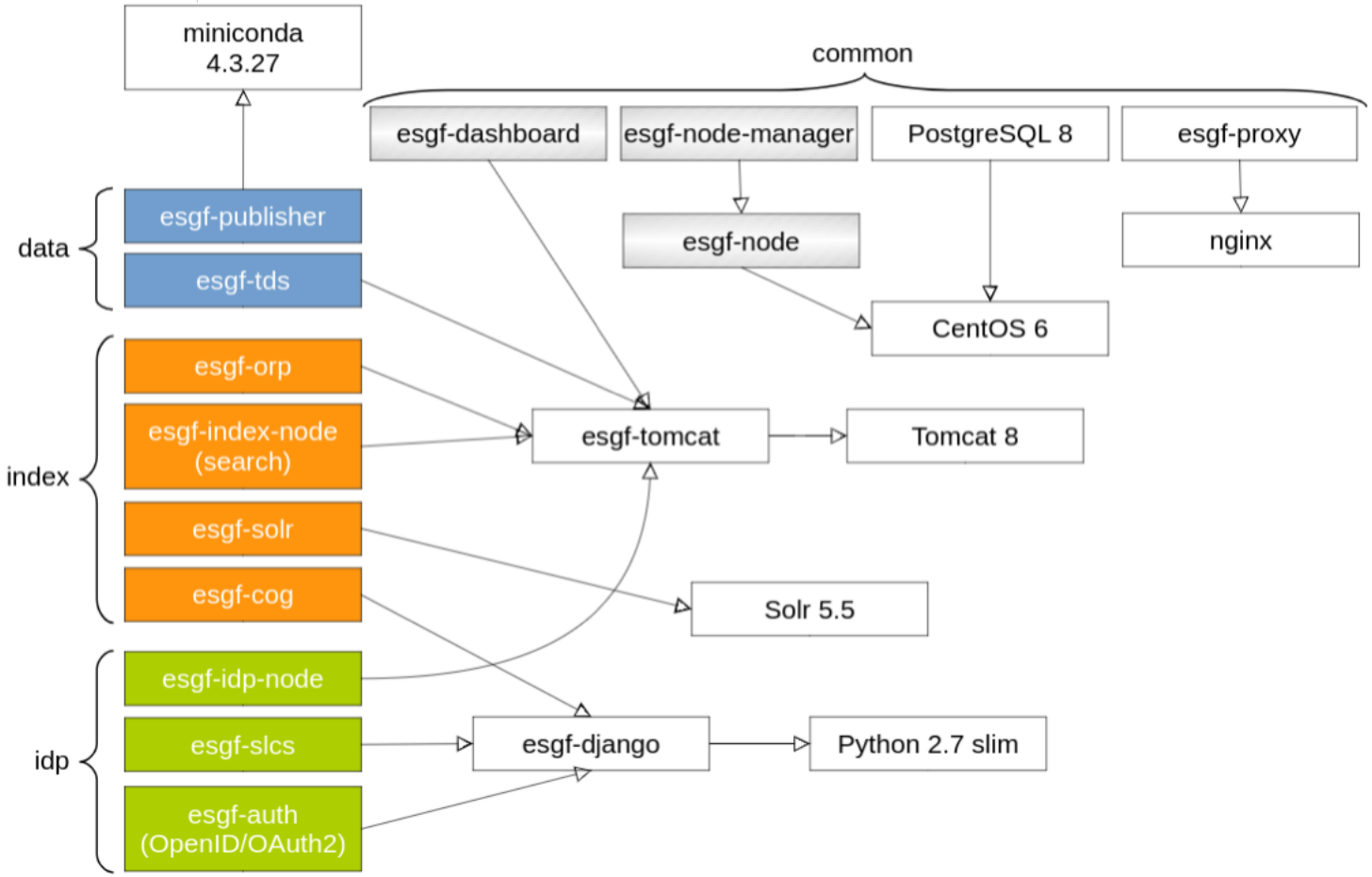
Copernicus

European Commission

Climate Change

Issue requiting future developments: containerization efforts

- A first attempt was made in 2017 to containerize ESGF that has been completely revisited in the beginning of 2018 by Matt Pritchard (CEDA) and Sébastien Gardoll (IPSL), to ease future containerization strategy.



Climate Change

Issue requiting future developments: containerization efforts

- A first attempt was made in 2017 to containerize ESGF that has been completely revisited in the beginning of 2018 by Matt Pritchard (CEDA) and Sébastien Gardoll (IPSL), to ease future containerization strategy.
- This solution has been tested by CEDA, and IPSL as an end-to-end ESGF Docker deployment with successful publication.
- From the IS-ENES ESGF Workshop conclusions:
  - ➡ ESGF will move from the current "RPM-based" deployment to Docker Images in 2020.
  - ➡ Those Docker Images could be used through Kubernetes or through Ansible installer to orchestrate containers across multiple nodes.

In the CDS context, it depends on level of sophistication the three sites are willing to attempt. As Tier 1 nodes, relying on Kubernetes is recommended for failover capacity.

Copernicus European Commission

Climate Change

Issue requiting future developments: containerization efforts

- Blocker to completion on ESGF containerization:
  - Realign to a baseline the parallel Docker deployment strategies.
  - Hard-coded configurations in some ESGF components.
  - Manage/update custom code in attribute service.
  - Review the data node "fabric" to provide Docker containers for the community. The IPSL has a build platform based on Jenkins and two dedicated VMs for Jenkins jobs also able to build Docker Images.
  - Sébastien Gardoll left the IPSL on mid-October ; Pierre Logerais at IPSL replaces Sébastien Gardoll and is training on about containerization purposes. For the time being most of containerization efforts and skills rely on Matt Pritchard (CEDA).

Climate Change

Issues for future maintenance and service continuation

- Python 2 will be retired at the end of 2019 but it is still used the ESGF stack especially in the publication tools.
- Security updates of CentOS6 will stop in November 2020. The IPSL wrote a procedure to upgrade from ESGF 4.0.3 to 4.0.4 including OS migration from CentOS 6 to 7 :

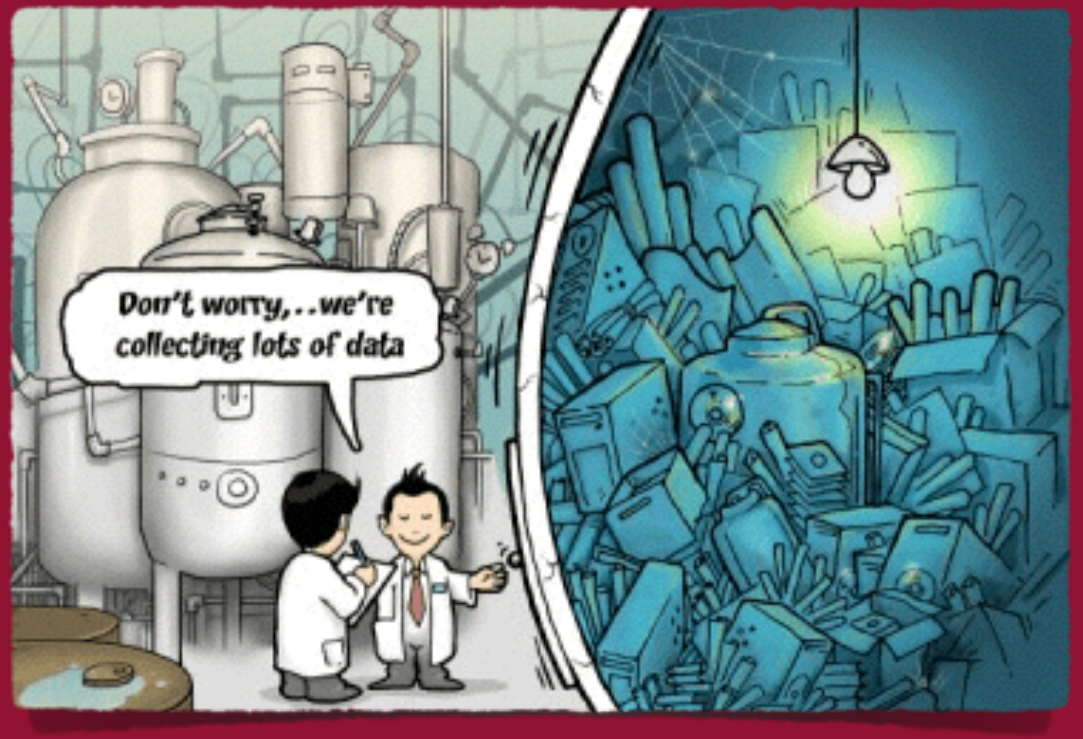
  http://vesgdev-docker2.ipsl.upmc.fr/esgf/migration_centos_7.html
- From the IS-ENES ESGF Workshop conclusions:
  - ➡ On-going efforts to migrate to Python 3 as soon as possible. This requires that ESGF community defines the earlier Python 3 version to be supported by the stack (3.6 is recommended).
  - ➡ Install ESGF stack on CentOS7 is highly recommended (which is supported by the Ansible installer).

The ESGF nodes deployed in the CDS context will benefit ESGF evolutions by regular upgrades.

Copernicus

European Commission

Copernicus Climate Change Service



Thank you for your attention.

Copernicus

European Commission