CVMFS IN EUCLID
OVERVIEW

P. CASENOVE

EC SGS SYSTEM MANAGER
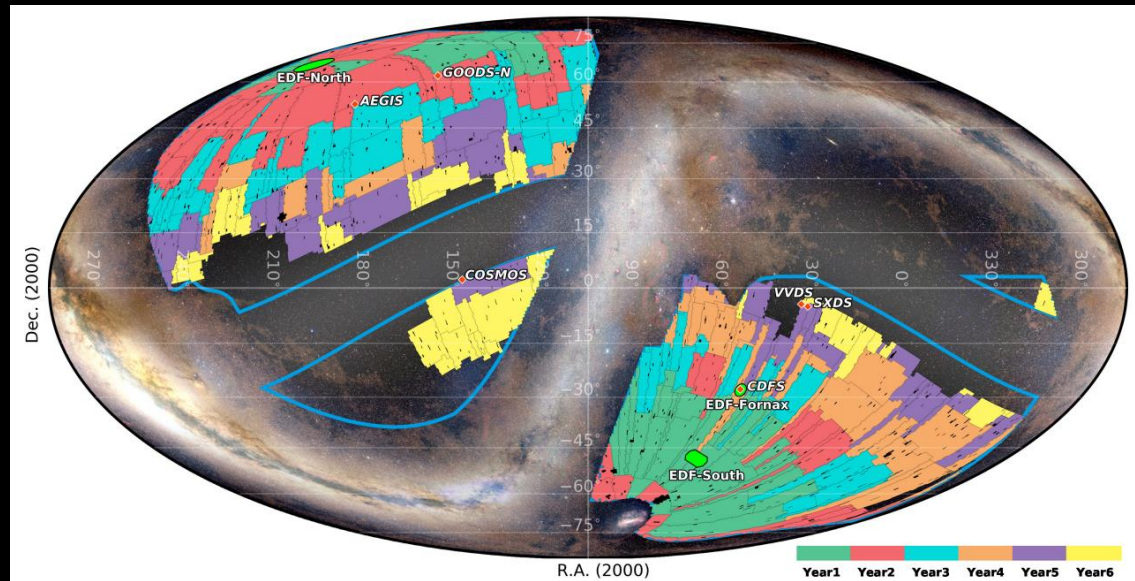
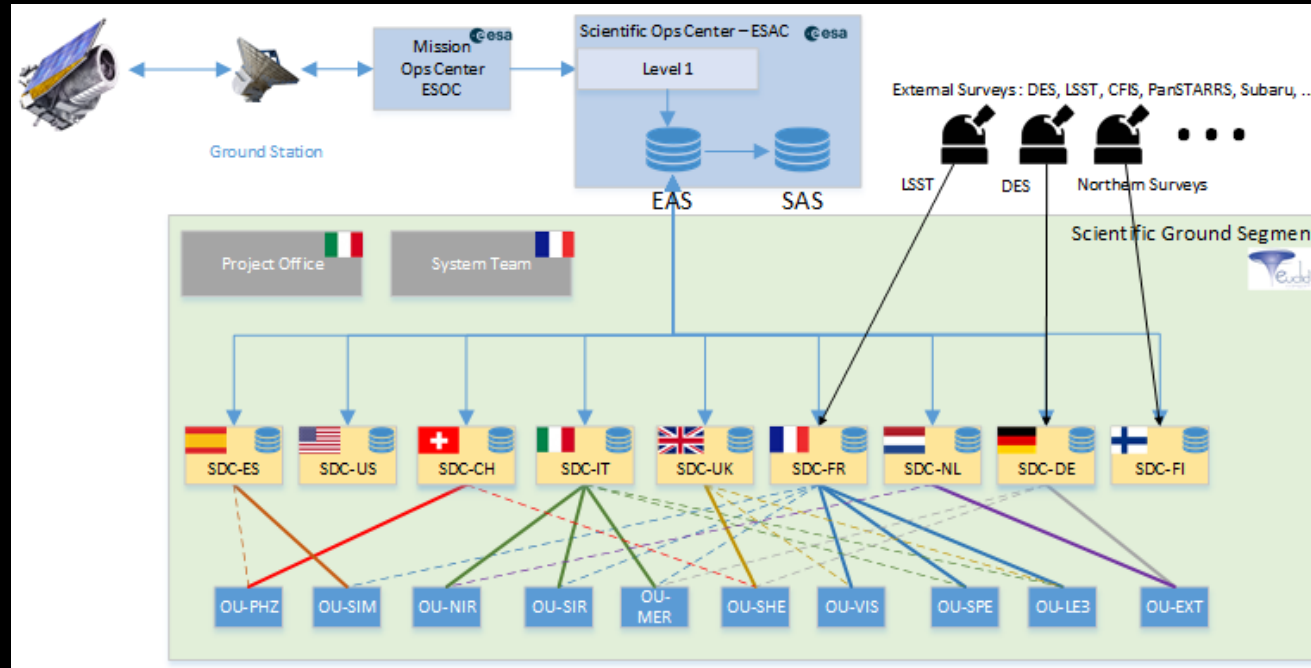17/09/2024

EUCLID MISSION

❖ M2 Mission of ESA Cosmic Vision program, designed to study the dark matter and energy through

❖ Space telescope launched on July 1st 2023 for a nominal mission of 6 years to survey 1/3 of the sky

❖ 850 Gbits of data downloaded from the spacecraft each day, to be processed by the ground segment

GROUND SEGMENT ARCHITECTURE

- ❖ Euclid data acquired from the Mission Operation Center (Darmstadt) and the Science Operation Center (Madrid)

- ❖ Processing done inside the Euclid Consortium Scientific Ground Segment, in 9 different data centers in the world

- ❖ Each data center receive a part of the sky and fully process this data, avoiding as much as possible inter data center transfers

  → Same pipeline (approx 1.5 million lines of code) is executed in all SDC

  → No incoming SSH connexion allowed in the data centers (obsviouly…)

CVMFS PROD ARCHITECTURE

- ❖ Final architecture set in 2020

- ❖ Added 1 S1 in US (bandwidth)

- ❖ Failover across the 3 S1 in any case

- ❖ GeoIP to connect to the nearest

- ❖ S0 migration in 2022 from COS7 to Debian 11
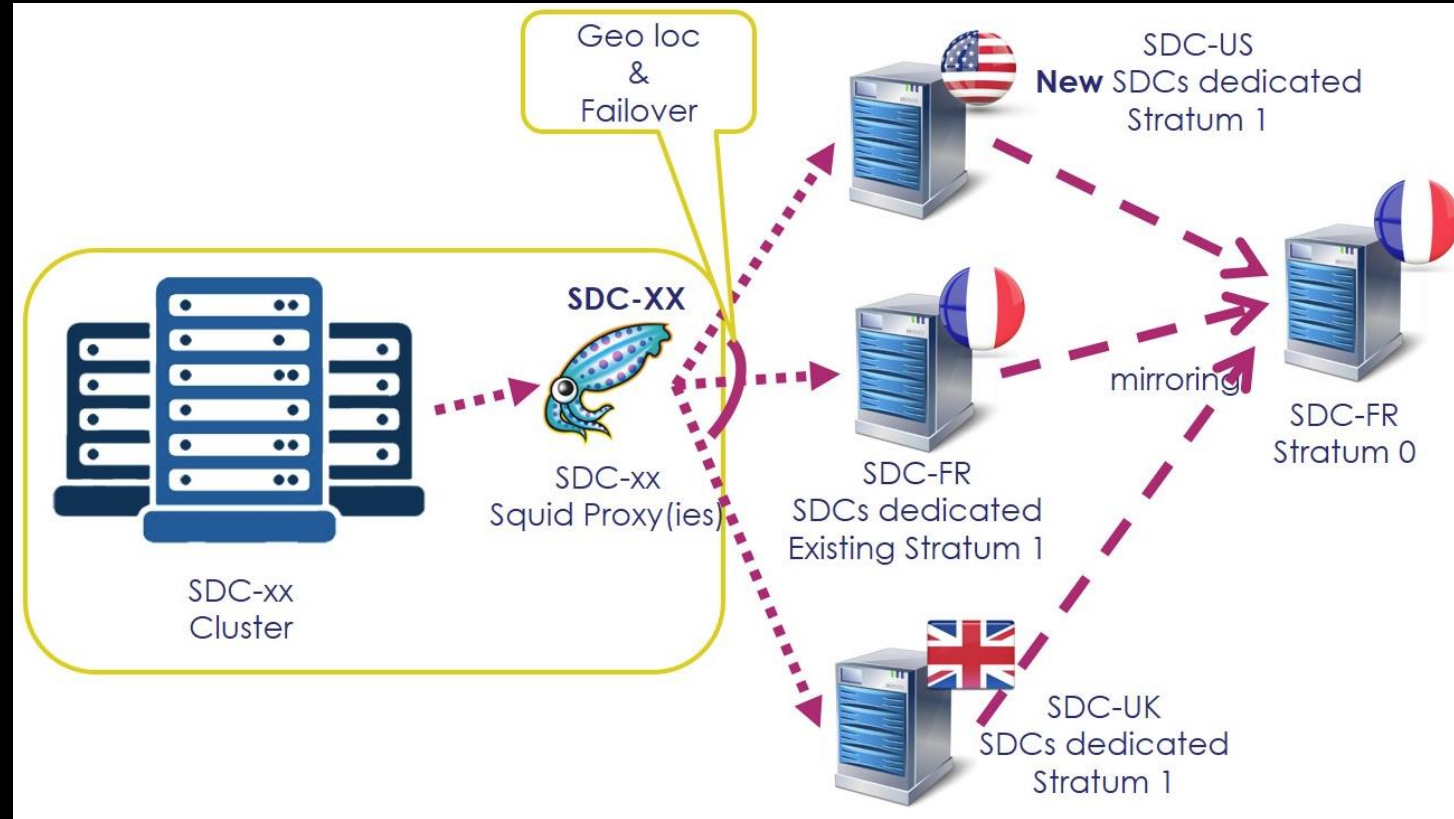
DEVOPS ARCHITECTURE

❖ Euclid code is in a self hosted gitlab instance

❖ CI/CD based on Jenkins

❖ Up to 2022:

   ❖ Code packaged as rpm

   ❖ Custom code running on S0 to install the RPM, taking a lock on the repo

   ❖ Highly unstable (RPM DB on cvmfs is not a good idea…)

❖ 2022: a new era

   ❖ Deployed CVMFS gateway

   ❖ Deployment using a simple "make install" from the publisher

   ❖ Publisher is just a specific Jenkins node

   ❖ Transaction only on the exact directory

   ❖ Highly stable after some initial debugging phases, and much faster

RÉPUBLIQUE
FRANÇAISE
Liberté
Égalité
Fraternité

cnes

CVMFS FOR THE DEVELOPERS

❖ Ground segment developers from over 100 institutes in EU and the US

❖ Working from many different places… including bad internet connections

❖ Up to 2020 : use of a public squid proxy

  ❖ Many complains due to 3128 port being blocked in many wifi

❖ Switched to a dedicated S1 server

❖ Failover to the other S1 in case of issue
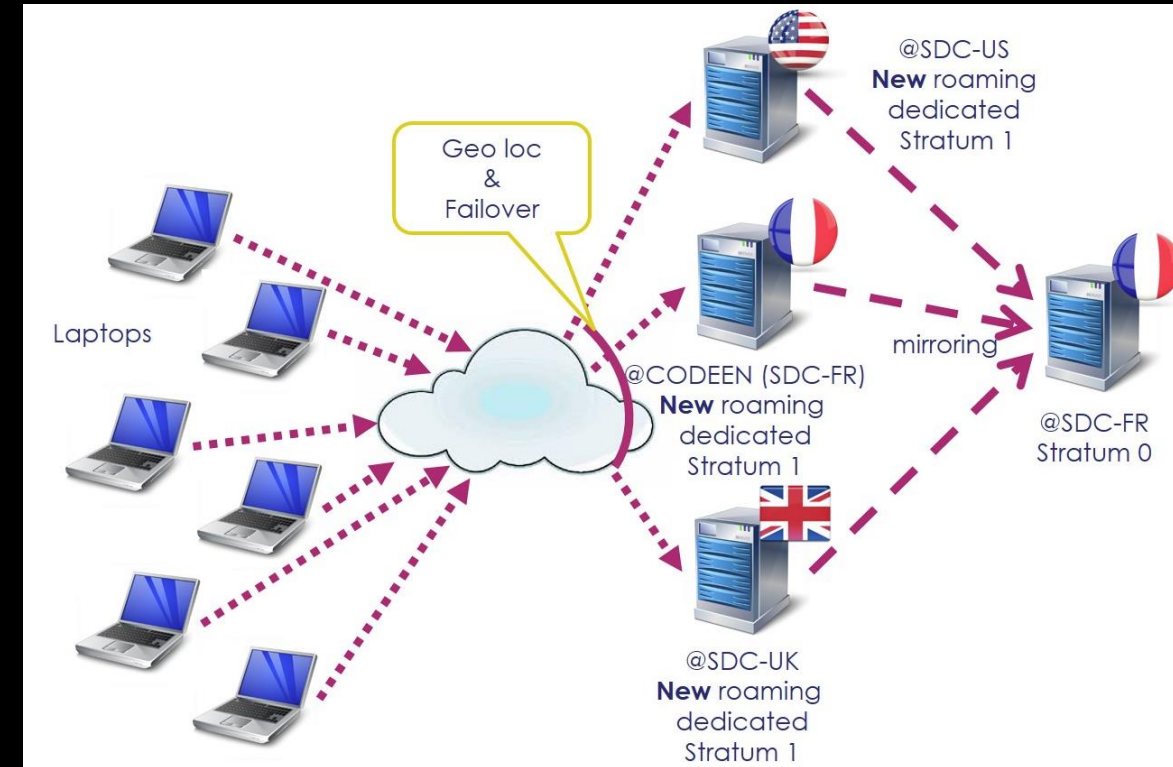
CVMFS IN EUCLID

❖ Bunch of updates to the initial (not so satisfactory) architecture

❖ S1 for the developers and publisher really stabilized the infrastructure

❖ Activated garbage collection for the dev repository having rolling updates

❖ Few to no containers: no investigation of DUCC and dedicated repo to deploy those images

❖ Only regular updates performed on the CVMFS servers since 2 years now

❖ Newcomers regularly impressed to see the code in 9 SDC across the world in 15 min after being built

 All of this with only outgoing HTTP traffic