

IRIS Collaboration Meeting

Data Distribution using CVMFS and Pelican

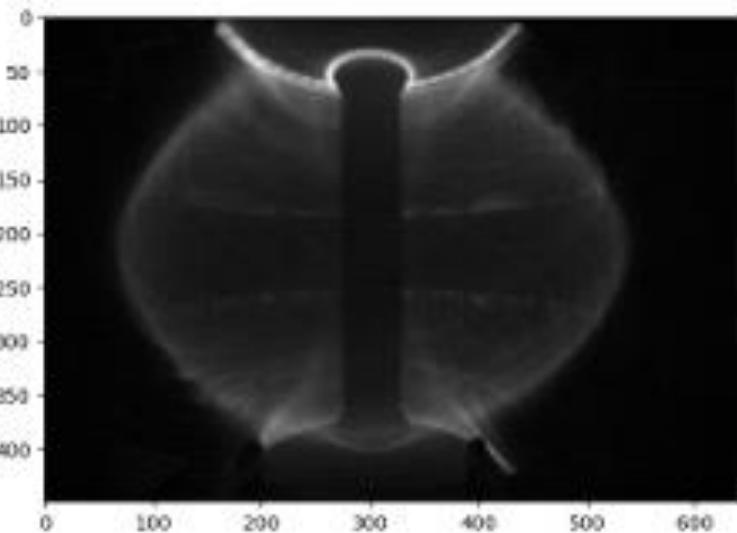
Author: Jonathan Hollocombe, Adam Parker, Stephen Dixon

Date: 14/01/2026

MAST(-U) and DIII-D

MAST (Mega Ampere Spherical Tokamak)

- Spherical tokamak design commissioned by EURATOM/UKAEA
- Built at Culham Centre for Fusion Energy, Oxfordshire, UK
- Experiments ran from 1999 through to 2013
- Produced **~30,000 shots** over its history
- Succeeded by MAST Upgrade (MAST-U) in 2020



DIII-D

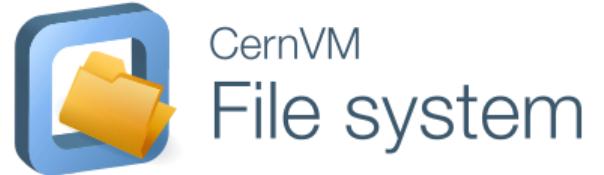
- Major advanced tokamak research device operated in the USA
- Operated by **General Atomics for the U.S. Department of Energy (San Diego)**
- Commissioned as an upgrade of the earlier **Doublet III** design, with first operation in **1986** and continuing active research campaigns now (mid-1980s - present)
- Has achieved **over 200,000 plasma shots** over its operating history, marking a major milestone in experimental campaigns



CVMFS and Pelican

CVMFS (CernVM File System) enables efficient global data distribution via a tiered mirror architecture, minimising long-distance data transfers. It presents remote data as a familiar POSIX file system, transferring (HTTP) files only when accessed (efficiency) and caching (performance) them locally for high performance.

Historically, CVMFS has been used primarily for **software distribution**, particularly in large-scale scientific computing environments.

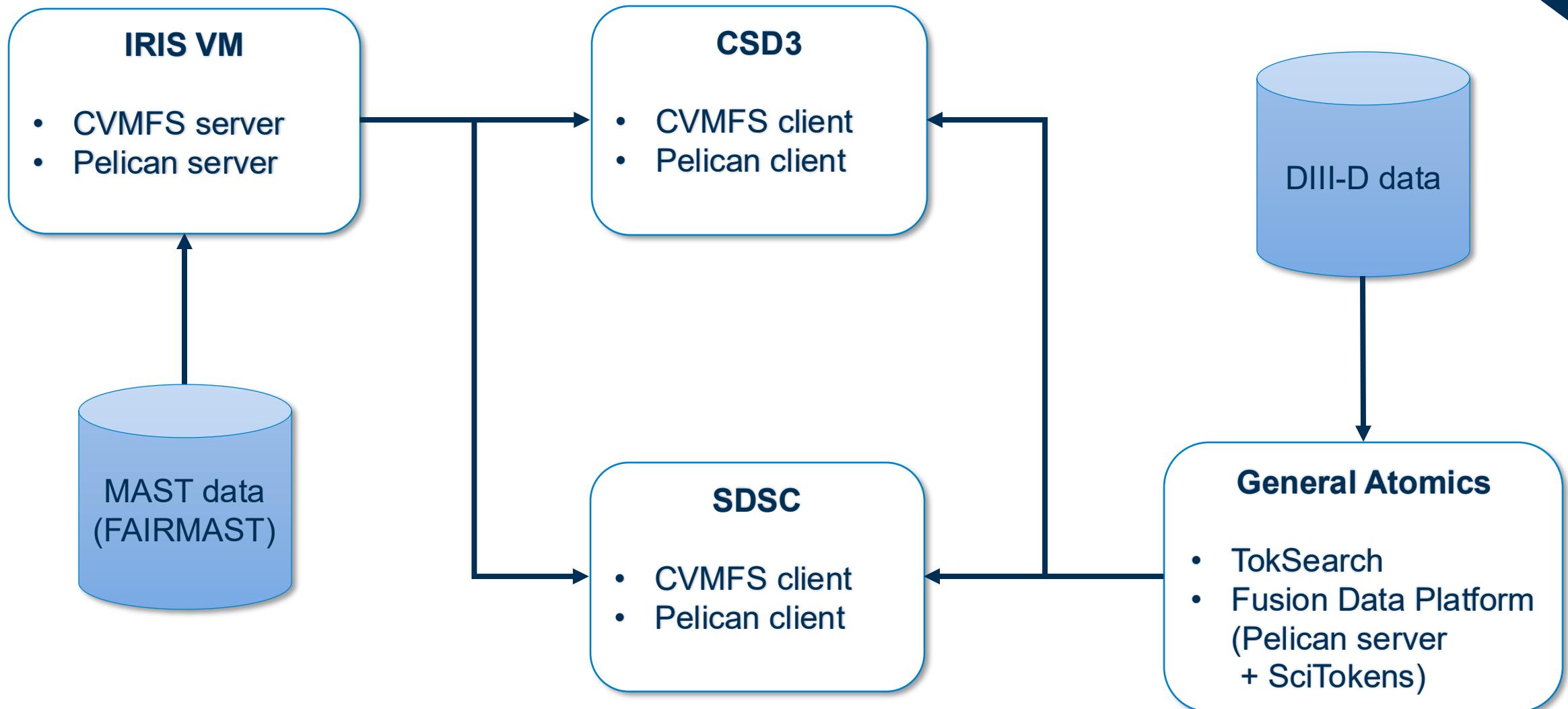


Pelican is the technology behind the Open Science Data Federation (OSDF), enabling globally distributed data networks.

- Provides a common interface to diverse storage backends (e.g. across different fusion experiments).
- Supports flexible authentication and authorisation (e.g. via SciTokens) for managing open (MAST), mixed, or embargoed (MAST-U) data.



Current Infrastructure



Initial Comparisons

- **CVMFS**
 - Presents as POSIX file system
 - Aggressive caching
 - Option to chunk data transfer
- **Pelican**
 - Ease of use, ease of configuration
 - Authentication/federated access
 - More supported storage backends
- Comparing '*apples*' to '*oranges*' but the next steps are to *compare* these with different access patterns and maybe other technologies

[1] CVMFS: <https://cvmfs.readthedocs.io/en/stable/>

[2] Pelican: <https://pelicanplatform.org/>

