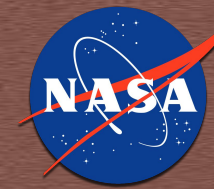


8th annual EARTH SYSTEM GRID FEDERATION



SYNDA: *The show must go on...*



Atef Ben Nasser
Guillaume Levavasseur
Sébastien Denvil



Institut
**Pierre
Simon
Laplace**

December 5th, 2018

Plan:


- Overview
- SYNDA today
- SYNDA & the community
- Roadmap



SYNDA in a nutshell



<https://github.com/Prodiguer/synda>

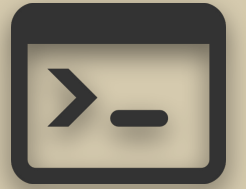
- 
- ESGF mirroring
• Command-line
• Parameter discovery
• Data discovery
• Similar to
 "apt-get"
 but asynchronous

- Data management
• Use nearest replica
• GridFTP
• Post-processing
• Downloading supervision

SYNDA current features

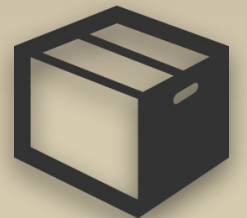
- DOCKER package,
- **RPM package,**
- `install.sh` script.

INSTALLATION PROCEDURE



SYSTEM PACKAGE (v3.9)

- **Daemon mode running in background,**
- Per-user configuration file (e.g., `~/.syndarc` file),
- Online help.



SEARCH

- Support most ESGF project (CMIP5, CORDEX, etc.),
- **Faceted data selection (experiment, variable, etc.),**
- **Incremental search** (download only what's new),
- **Nearest replica selection.**



DOWNLOAD

- **Parallel downloads (using HTTP or GridFTP protocol),**
- Checksums control,
- Transfer priority and statistics,
- **Transparent x509 certificate renewal.**

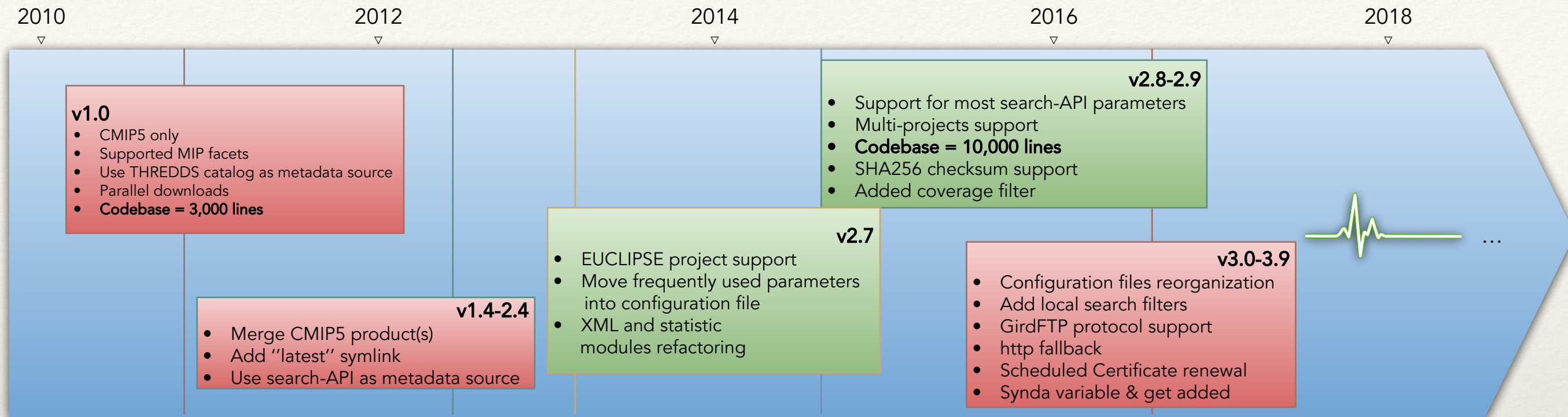


DATA MANAGEMENT

- Default storage following DRS tree format,
- Template based data management,
- Auto-remove old dataset versions.



SYNDA history



SYNDA position in ESGF environment

- ESGF portal (data search)
- pyclient (data search)
- wget (data download)
- ESGPrep (versioning and Data Reference Syntax management)

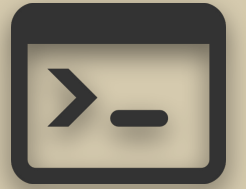
SYNDA in 2018 ?

- Latest release Oct. 25th 2017 (sdt 3.9/sdp 1.4)
- Issue count on github: 49
- Pending pull requests: 1
- Number of forks: 6

SYNDA known issues

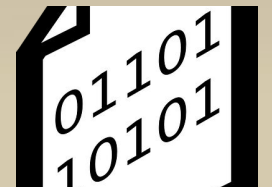
INSTALLATION PROCEDURE

- No pip installable package.
- or Conda environment.



CODE BASE

- Poor *initial* code-coverage results.
- Absence of a test-suite.
- Deprecated python 2.6.6
- Code and DB architecture too CMIP5 specific



USER XP

- CMIP5 specific legacy code (variable management within dataset id)
- Difficult to configure to suit local needs



DOWNLOAD

- Performance optimization
- Gridftp misuse



SYNDA IDENTITY

- Jack of all trades, master of none
- (trying to achieve many tasks, maybe too many?)
- Community contributions must be within set boundaries
- Some meta-dev work required



- Bugs and releases :
 - Bug fixes will start rolling as soon as possible (start of 2019)
 - Bugs will be treated based on the order of priority (you could help by properly labeling the issue you created by severity on GH).
 - Priority will be given to refactoring, some fixes will be therefore delayed.
- Pull Requests :
 - Just like the bug fixes they will individually be treated as soon as humanly possible
 - They'll also depend on refactoring and priority
 - Should be within the yet to be defined perimeter of synda
 - SYNDA Test-Suite will play a major part in streamlining PR ingestion
- SYNDA Test-Suite :
 - Vital in the transition to python 3.x (tbd and possibly keep 2.7 support if needed and not costly)
 - Will contribute to defining SYNDA's perimeter but will also be defined by it
- Documentation :
 - Update and maintain

- Defining SYNDA's perimeter :
 - What's expected and what's not
- Refactoring :
 - Getting SYNDA CMIP6 compliant at first,
 - Making sure SYNDA more project-dynamic on the long run.
 - Improving the code metrics score (e.g. code coverage, current metrics are biased due to test-suite absence).
- Technologies :
 - Current solution is based on 2.6.6 -> unacceptable.
 - Highest priority to move to 3.x
 - Could consider backwards compatibility with 2.7 if needed
 - sqlite db (?)
 - pyclient (?)
 - Globus client (?)
- Distribution :
 - RPM packaging to PIP and/or Conda package
- Test-Suite :
 - 2-step plan:
 - Basic test suite for obvious reasons
 - A comprehensive test suite later on.
 - Continuous integration of new developments & PRs on the long run (via Jenkins or other)

Thank you for your attention

- Core Development :
 - Atef BN.¹
- Refactoring :
 - Atef BN.¹
- PR & Bug fixes :
 - Atef BN.¹
- Support :
 - Guillaume L.², Sébastien D.¹, Atef BN.¹
- Test-Suite:
 - Atef BN.¹
- Documentation:
 - Guillaume L.², Atef BN.¹

¹: IPSL/CNRS

²: IPSL/Sorbonne Universités

② <https://github.com/Prodiguer/synda>

Contact info:

Github:

Atef BN. : <https://github.com/AtefBN>

Guillaume L.: <https://github.com/glevava>

Sébastien D.: <https://github.com/SebastienDenvil>

Email:

Atef BN.: atef.bennasser@ipsl.fr

Guillaume L.: glipsl@ipsl.fr

Sébastien D.: sebastien.denvil@ipsl.fr

Join us on Slack !