

Research reproducibility and workflows with the Integrated Forecasting System

(featuring Belle the cat and friends)

Andrew Bennett

ECMWF

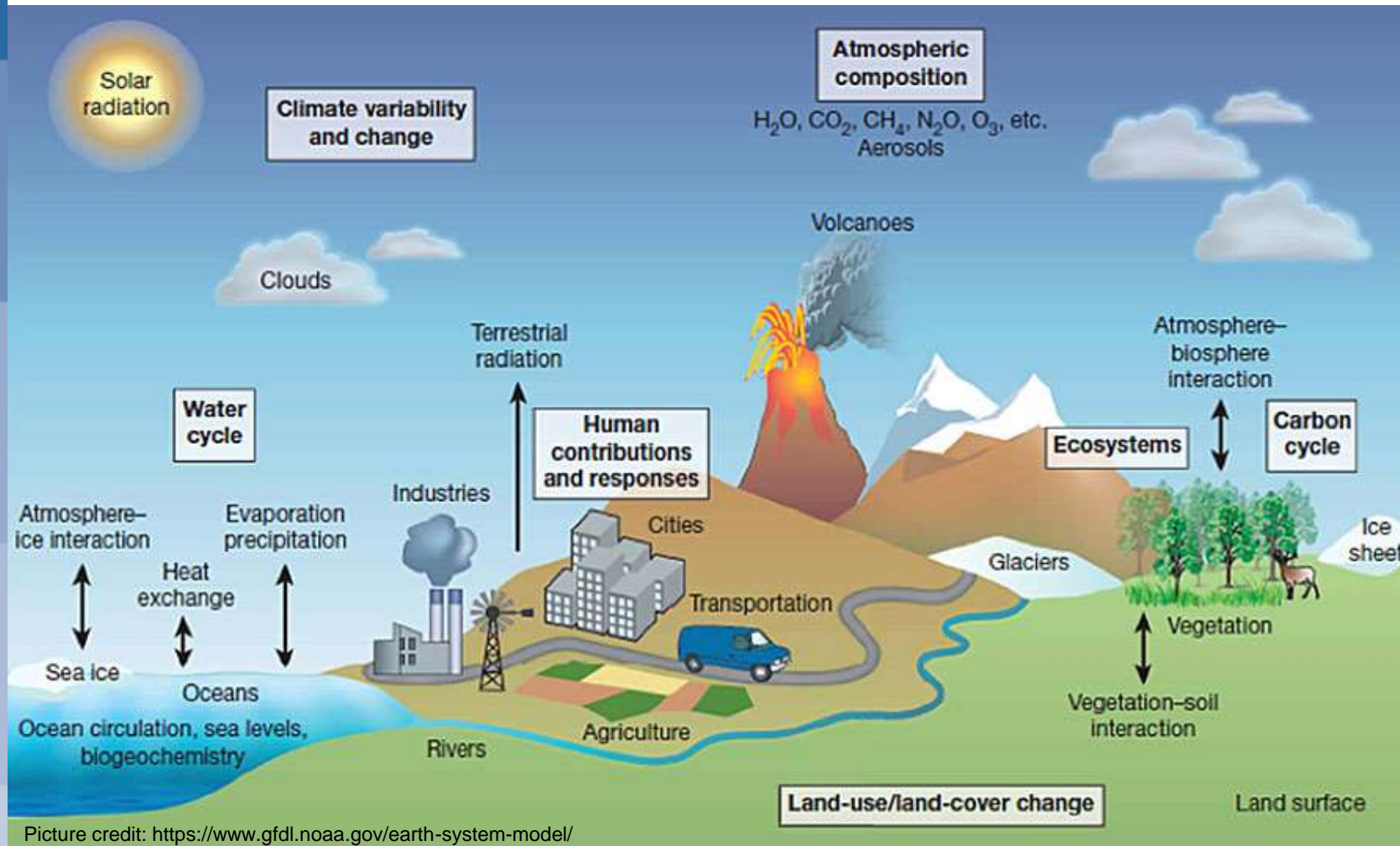
andrew.bennett@ecmwf.int



Belle



Integrated Forecasting System



- Non-linear feedbacks between components
- 6 million lines of code
- Future changes will increase complexity



diab:PrepIFS [1.9.1-837]

Window File Options Help

U diab

47r1

h72p

an

Data assimilation mode

Dates and times

Description

Forecast setup, 6hourly c

Git repos/tags, ECCODE

Initial data

Model uncertainty repres

NEMO Observation contr

NEMO perturbations con

NEMO physical settings t

NEMOVAR assimilation

Ocean general, resolutio

Ocean-atmosphere coup

Resolution, general setu

Submit options

Surface fields

Accounting, system prop

Add your own Variables,

Archiving

Augmented Control Vari

Bias correction

Blacklist files

Compiler options, debug

Copy history

Coupled wave model se

Dr Hook

Early delivery

Ensemble Data Assimila

Ensemble Kalman Filter

Forecast error handling

Gems/CTM coupling

MACC

Numerical Precision

ODB archiving

ODB configuration

Observations

Observations monitoring

Ozone, CO2

Plots, verification

Post-processing controls

Prognostic Hessian SV

47r1/h72p/an/libraries

Prebuilt IFS-bundle module

☐ IFS_BUNDLE_MODULE - Prebuilt IFS-bundle to use
 Build within experiment

GIT

☐ GIT_VERSION - Git environment module version
 2.12.0

☐ IFS_GIT_CENTRAL_URL - URL for ECMWF Bitbucket Git
 ssh://git@git.ecmwf.int:7999

Use same git tag for all IFS repositories

☐ IFS_ALL_GIT_TAG - Automatically apply this git tag for all the unset IFS_*_GIT_TAG below
 None

ifs-scripts

☒ IFS_SCRIPTS_GIT_REPO - URL for IFS scripts git repository
 Branch Owners Bitbucket Repository

☒ IFS_SCRIPTS_GIT_TAG - Git tag (branch name) for scripts
 dag_CY46R1_for_47R1_v6

ifs-suites

☒ IFS_SUITES_GIT_REPO - URL for IFS scripts git repository
 Branch Owners Bitbucket Repository

☒ IFS_SUITES_GIT_TAG - Git tag (branch name) for scripts
 diab_CY46R1_for_47R1_v6

ifs-source

☒ IFS_SOURCE_GIT_REPO - URL for IFS source git repository
 Main Central Repository

☒ IFS_SOURCE_GIT_TAG - Git tag (branch name) for source
 nat_CY46R1_for_47R1_v6

ifs-bundle

☒ IFS_BUNDLE_GIT_REPO - URL for ifs-bundle git repository
 Main Central Repository

☒ IFS_BUNDLE_GIT_TAG - Git tag (branch name) for ifs-bundle
 nat_CY46R1_for_47R1_v6

atlas

☐ ATLAS_GIT_REPO - URL for atlas git repository
 Main Central Repository

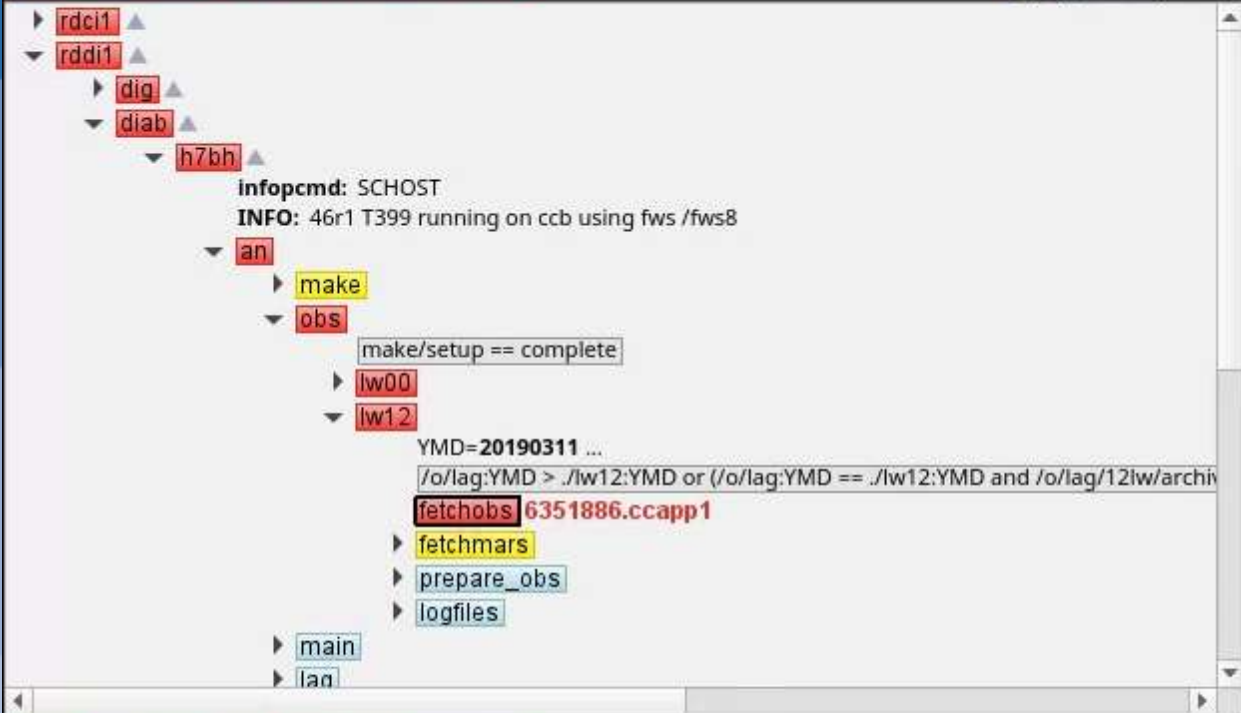
☐ ATLAS_GIT_TAG - Git tag (branch name) for atlas

Loading user folder for user diab, please wait

User folder for user diab loaded

Loading experiment type diab/47r1/h72p/an, please wait

Successfully loaded experiment type an



File: /vol/ifs_sms/rd/diab/h7bh/an/obs/lw12/fetchobs.1 Size: 585 KB
Source: served by ccb-log@9314 (took 0.1 s) at 2019-10-12 17:20:02

```

17:06:13 ERROR[12458] /usr/local/apps/ecflow/4.12.0/bin/ecflow_client --abort 6351886.ccapp1
17:06:13 ERROR[12436] [[ -n /fws8/sb/log/rd/diab/h7bh/an/obs/lw12/fetchobs.1 ]]
17:06:13 ERROR[12437] [[ -n /nfs/moms/ccappn098/var/spool/PBS/spool/6351886.ccapp1.0U ]]
17:06:13 ERROR[12437] [[ -f /var/spool/PBS/spool/6351886.ccapp1.0U ]]
17:06:13 ERROR[12437] cp /var/spool/PBS/spool/6351886.ccapp1.0U /fws8/sb/log/rd/diab/h7bh/an/obs/lw12/fetchobs.1
17:06:13 ERROR[12438] [[ -L /fws8/sb/log/rd/diab/h7bh/an/obs/lw12/fetchobs.1.running ]]
17:06:13 ERROR[12438] rm -f /fws8/sb/log/rd/diab/h7bh/an/obs/lw12/fetchobs.1.running
17:06:13 ERROR[12442] [[ -n /fws8/TMP/JTMP/81/rdx.6351886.ccapp1.20190909T170305/_child ]]
17:06:13 ERROR[12443] touch /fws8/TMP/JTMP/81/rdx.6351886.ccapp1.20190909T170305/_child
17:06:13 ERROR[12446] trap 0
17:06:13 ERROR[12447] date
  
```

Directory listing updated at 2019-10-12 17:20:02

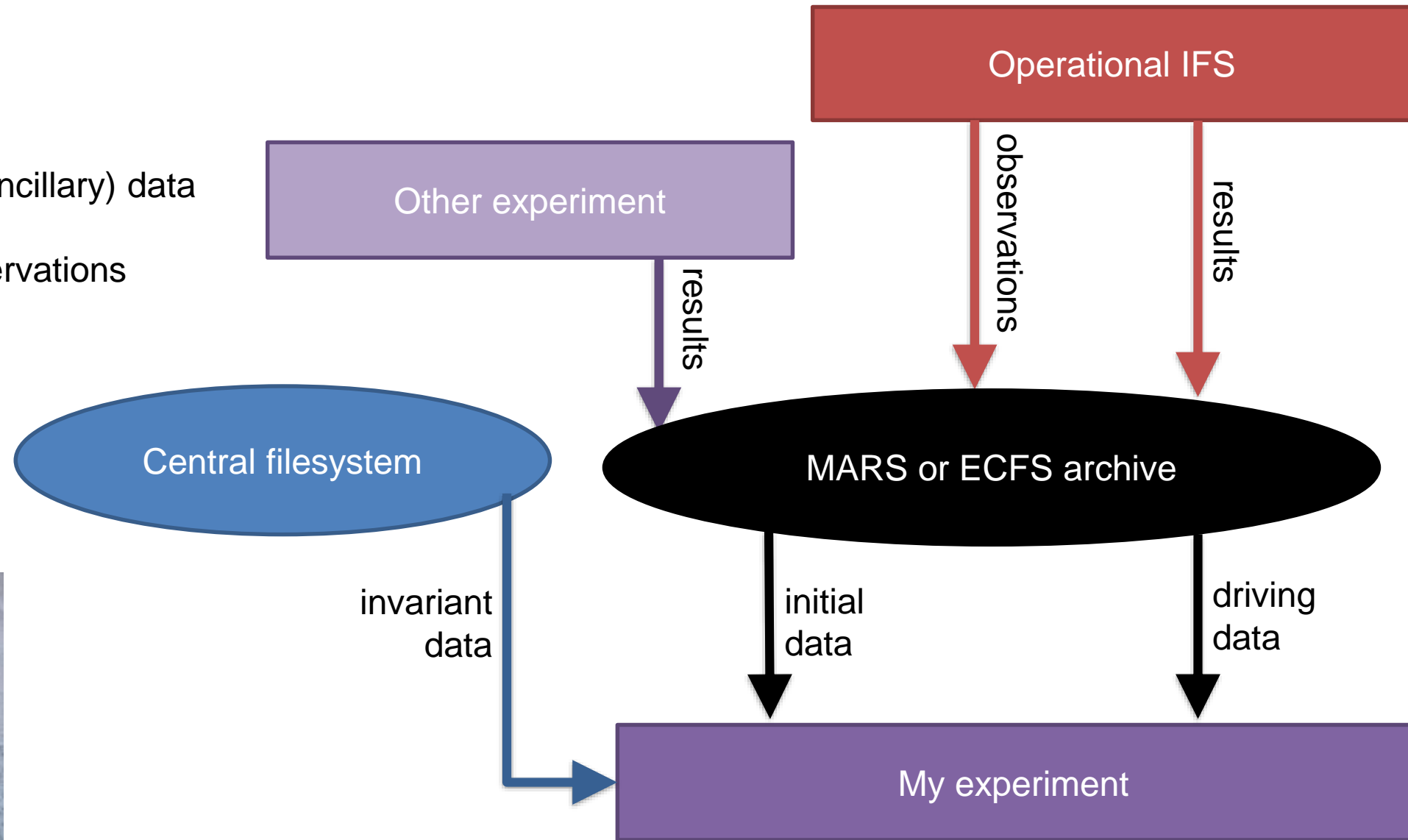
Name	Path	Size	Modified (ago)	Modified	Source
fetcho...	/fws8/...	585 KB	32 days ago	2019-09-09 18:06:14	served by ccb-log@9314
fetcho...	/fws8/...	372 KB	32 days ago	2019-09-09 18:03:05	served by ccb-log@9314

Filter: SELECT node

Node	Status	Type	Trigger	Label	Event	Meter	Status changed
/stc/h7wt/an/main	queued	family	uptraj_2 == complete				2019-Oct-12 11:37:00
/stc/h7wt/an/main	queued	task	ifstraj == complete				2019-Oct-12 11:37:00
/stc/h7wt/an/main	queued	task	ifstraj == complete				2019-Oct-12 11:37:00
/stc/h7wt/an/main	submitted	task					2019-Oct-12 16:18:00
/stc/h7wt/an/main	submitted	family	uptraj_1 == complete				2019-Oct-12 16:18:00
/stc/h7wt/an/main	complete	task	ifstraj == complete				2019-Oct-12 16:18:00
/stc/h7wt/an/main	complete	task	ifstraj == complete				2019-Oct-12 15:49:00
/stc/h7wt/an/main	complete	task					2019-Oct-12 15:48:00
/stc/h7wt/an/main	complete	family	uptraj_0 == complete				2019-Oct-12 16:18:00

Handling data

- Outputs: results
- Inputs:
 - invariant (static/ancillary) data
 - Initial data
 - Driving data/observations



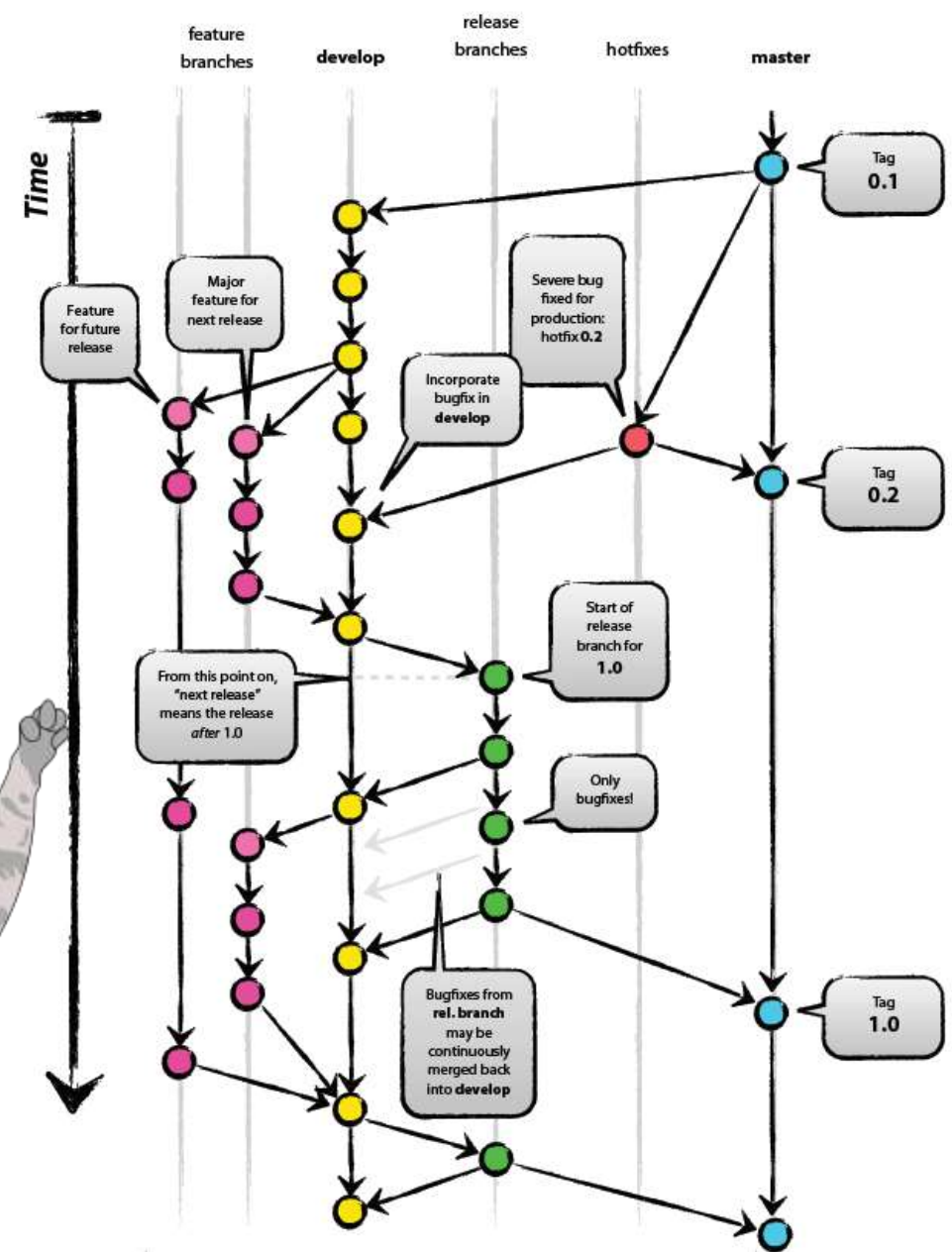
Code dependencies

- IFS code
 - ifs-source: Fortran source code
 - ifs-scripts: Korn shell scripts
 - ifs-defaults: templates for configuration of experiments
 - ifs-suites: definition of experiment pipeline / ecFlow suite



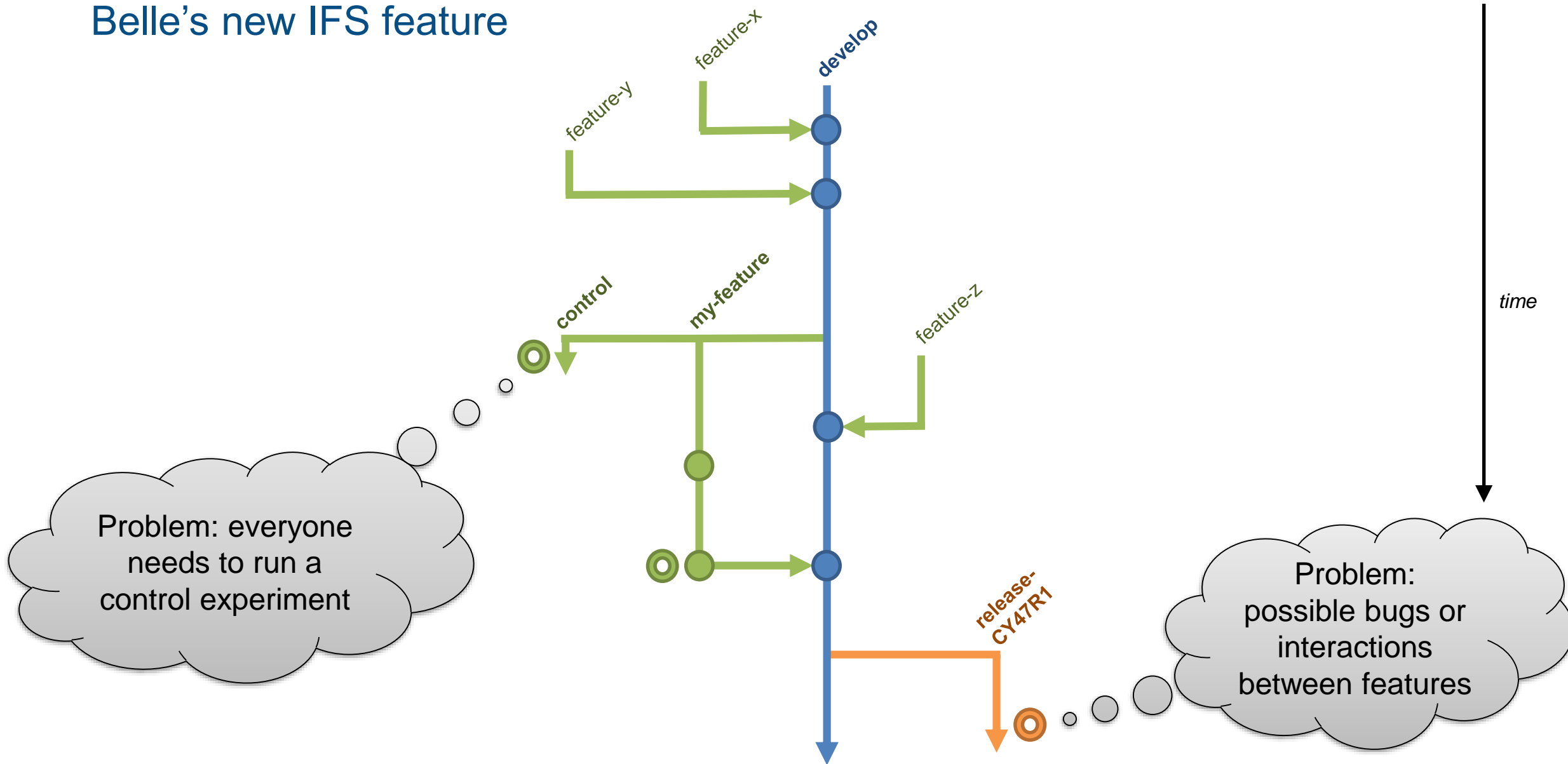
- Supporting libraries: ifs-bundle
 - Build machinery: ecbuild
 - In-house developed: eccodes, MIR, fckit, eckit, MARS client, ...

Branching models

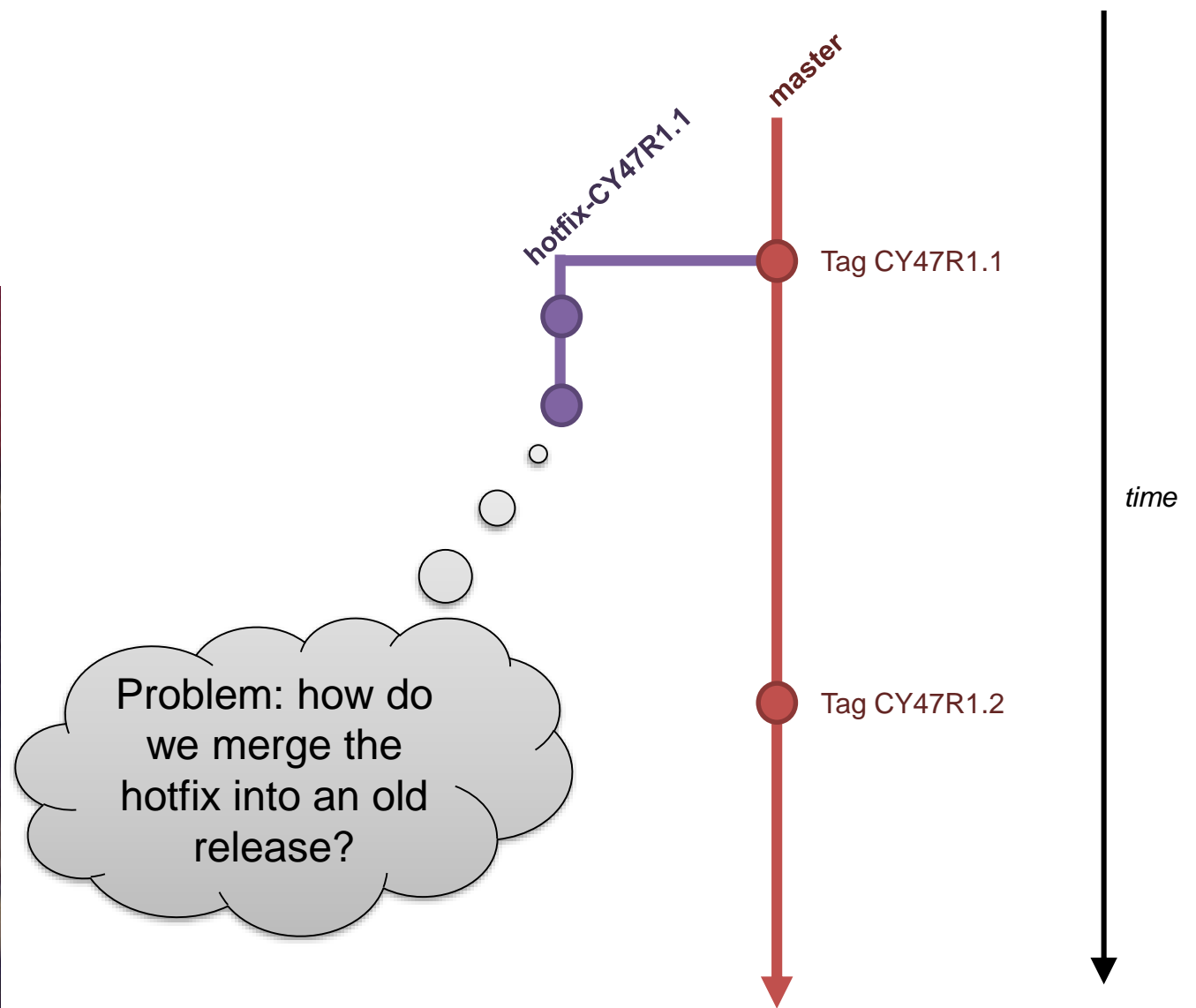
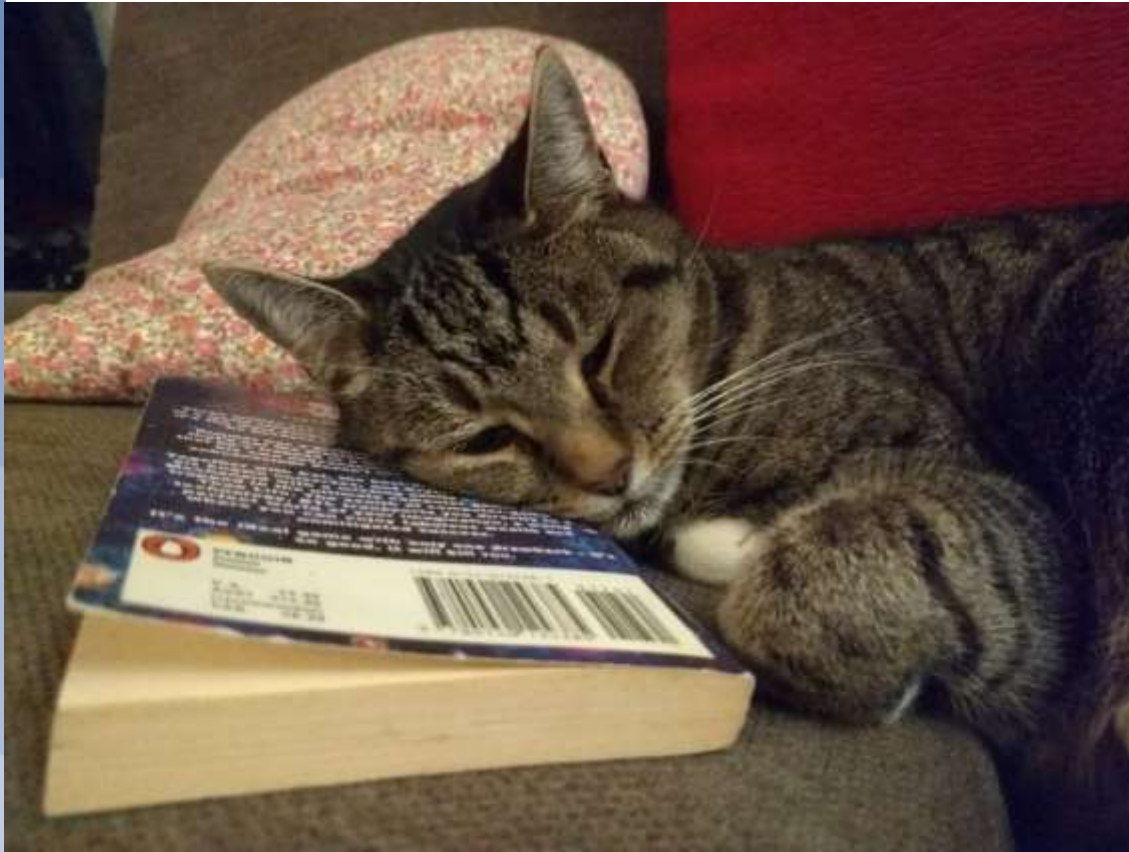


Author: Vincent Driessen
Original blog post: <http://nvie.com/posts/a-successful-git-branching-model>
License: Creative Commons BY-SA

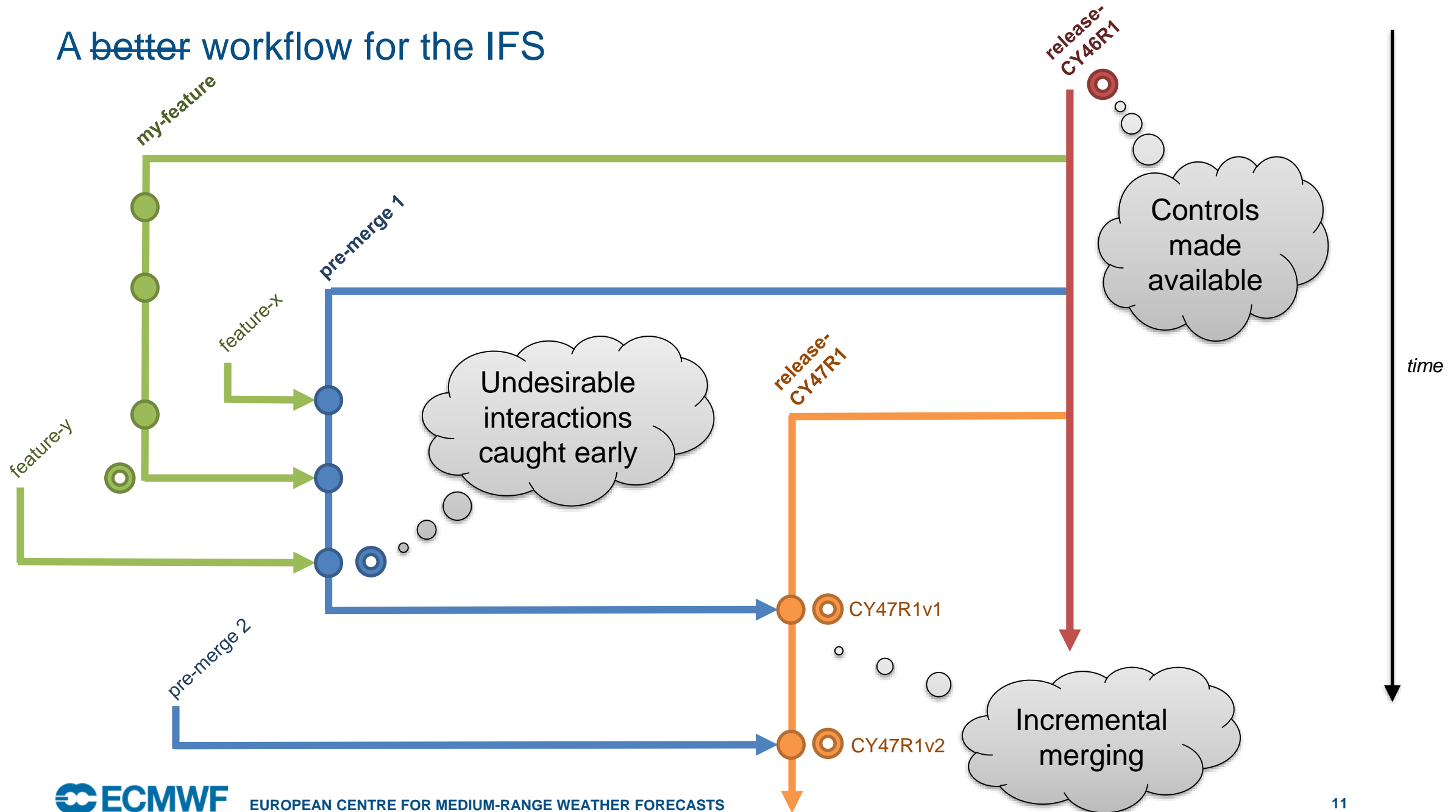
Belle's new IFS feature



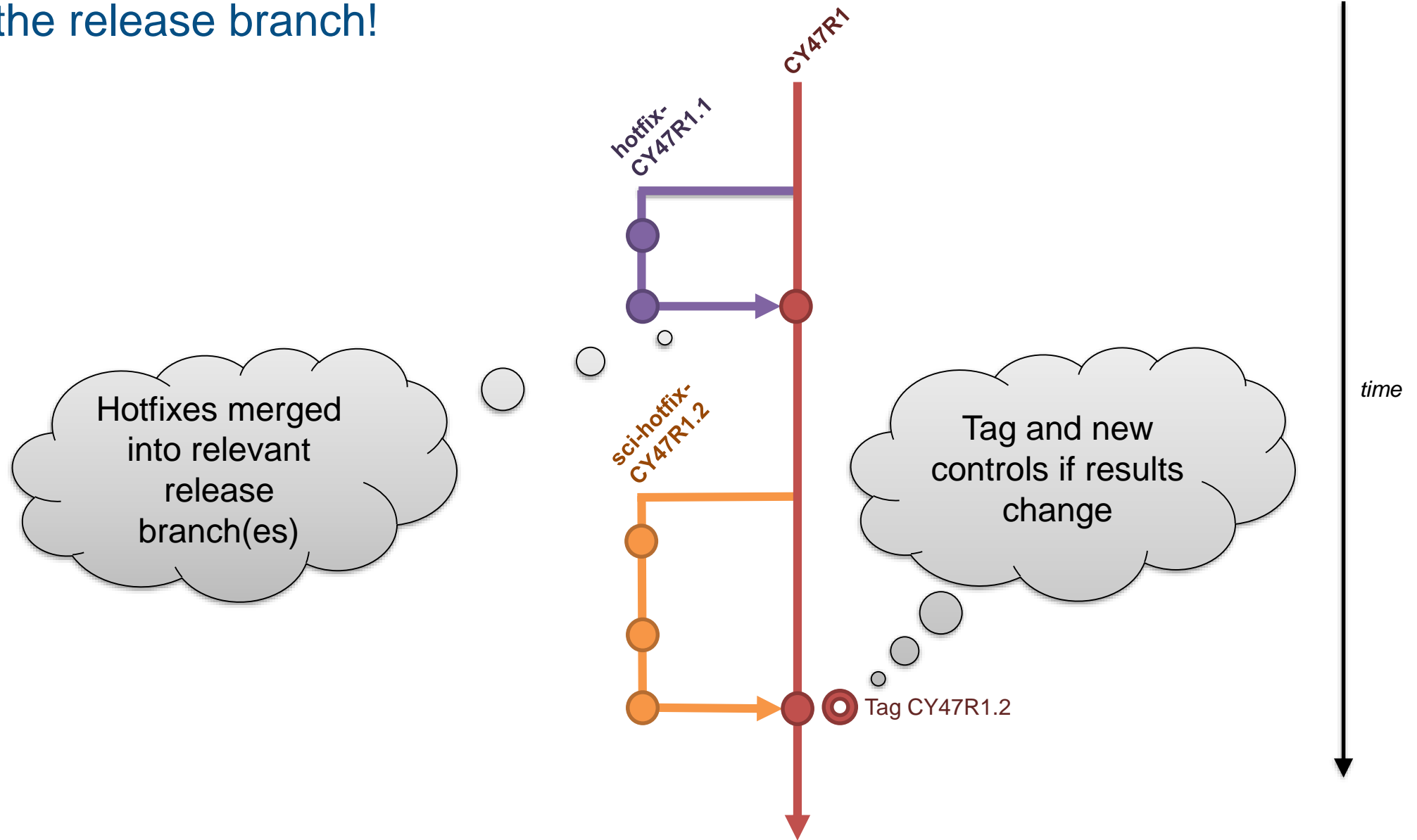
Belle fixes a bug



A better workflow for the IFS



Long live the release branch!

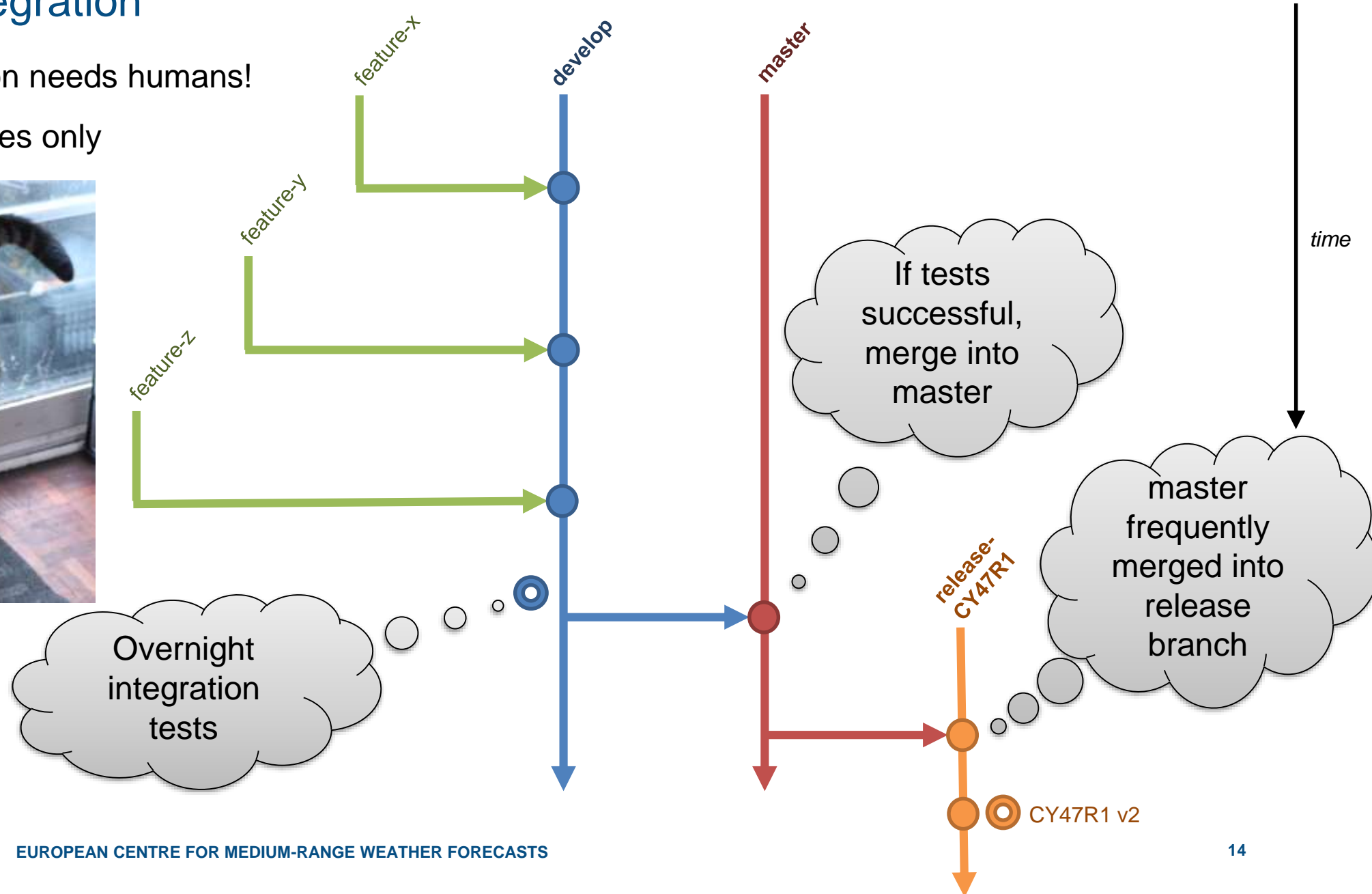


It's complicated!



Continuous integration

- Scientific evaluation needs humans!
- Bit-identical changes only



Workstation build and test system

```
Test project /tmp/tmpdir/diab/git/ifs-builds/CY46R1/build
  Start 898: ifs_t21_test_fc
1/18 Test #898: ifs_t21_test_fc ..... Passed    44.06 sec
  Start 899: ifs_t21_test_fc_tegen
2/18 Test #899: ifs_t21_test_fc_tegen ..... Passed     4.77 sec
  Start 900: ifs_t21_test_adj
3/18 Test #900: ifs_t21_test_adj ..... Passed     9.20 sec
  Start 901: ifs_t21_test_tl_taylor
4/18 Test #901: ifs_t21_test_tl_taylor ..... Passed     4.75 sec
  Start 902: ifs_t21_test_np2
5/18 Test #902: ifs_t21_test_np2 ..... Passed     6.12 sec
  Start 903: ifs_t21_test_nt2
6/18 Test #903: ifs_t21_test_nt2 ..... Passed     5.91 sec
  Start 904: ifs_t21_test_sv
7/18 Test #904: ifs_t21_test_sv ..... Passed    16.50 sec
  Start 905: ifs_t21_test_sppt
8/18 Test #905: ifs_t21_test_sppt ..... Passed     4.79 sec
  Start 906: ifs_t21_test_hybrid
9/18 Test #906: ifs_t21_test_hybrid ..... Passed     3.98 sec
  Start 907: ifs_t21_test_updclie
10/18 Test #907: ifs_t21_test_updclie ..... Passed    12.77 sec
  Start 908: ifs_t21_test_updclie2
11/18 Test #908: ifs_t21_test_updclie2 ..... Passed    12.87 sec
  Start 909: ifs_t21_test_compo_fc
12/18 Test #909: ifs_t21_test_compo_fc ..... Passed    16.80 sec
  Start 910: ifs_t21_test_compo_fc_climrad
```

Pre-built module

diab:PrepIFS [1.9.1.837]

Window File Options Help

47r1/h72p/an/libraries

Prebuilt IFS-bundle module

☐ IFS_BUNDLE_MODULE - Prebuilt IFS-bundle to use Build within experiment

GIT

☐ GIT_VERSION - Git environment module version 2.12.0

☐ IFS_GIT_CENTRAL_URL - URL for ECMWF Bitbucket Git ssh://git@git.ecmwf.int:7999

Use same git tag for all IFS repositories

☐ IFS_ALL_GIT_TAG - Automatically apply this git tag for all the unset IFS_*_GIT_TAG below None

ifs-scripts

☒ IFS_SCRIPTS_GIT_REPO - URL for IFS scripts git repository Branch Owners Bitbucket Repository

☒ IFS_SCRIPTS_GIT_TAG - Git tag (branch name) for scripts dag_CY46R1_for_47R1_v6

ifs-suites

☒ IFS_SUITES_GIT_REPO - URL for IFS scripts git repository Branch Owners Bitbucket Repository

☒ IFS_SUITES_GIT_TAG - Git tag (branch name) for scripts diab_CY46R1_for_47R1_v6

ifs-source

☒ IFS_SOURCE_GIT_REPO - URL for IFS source git repository Main Central Repository

☒ IFS_SOURCE_GIT_TAG - Git tag (branch name) for source nat_CY46R1_for_47R1_v6

ifs-bundle

diab

47r1

h72p

an

- Data assimilation mode
- Dates and times
- Description
- Forecast setup, 6hourly c
- Git repos/tags, ECCODE
- Initial data
- Model uncertainty repres
- NEMO Observation contr
- NEMO perturbations con
- NEMO physical settings t
- NEMOVAR assimilation
- Ocean general, resolutio
- Ocean-atmosphere coup
- Resolution, general setu
- Submit options
- Surface fields
- Accounting, system prop
- Add your own Variables,
- Archiving
- Augmented Control Varia
- Bias correction
- Blacklist files
- Compiler options, debug
- Copy history
- Coupled wave model se
- Dr Hook
- Early delivery
- Ensemble Data Assimila
- Ensemble Kalman Filter
- Forecast error handling
- Gems/CTM coupling

Possible future improvements

- Version control of invariant data files
- Run experiments inside a container



Summary

IFS tools and workflow adapted for:

- Ensuring reproducibility
- Long-running experiments for statistical significance
- Enabling scientific evaluation

But we're taking ideas from non-scientific workflow!

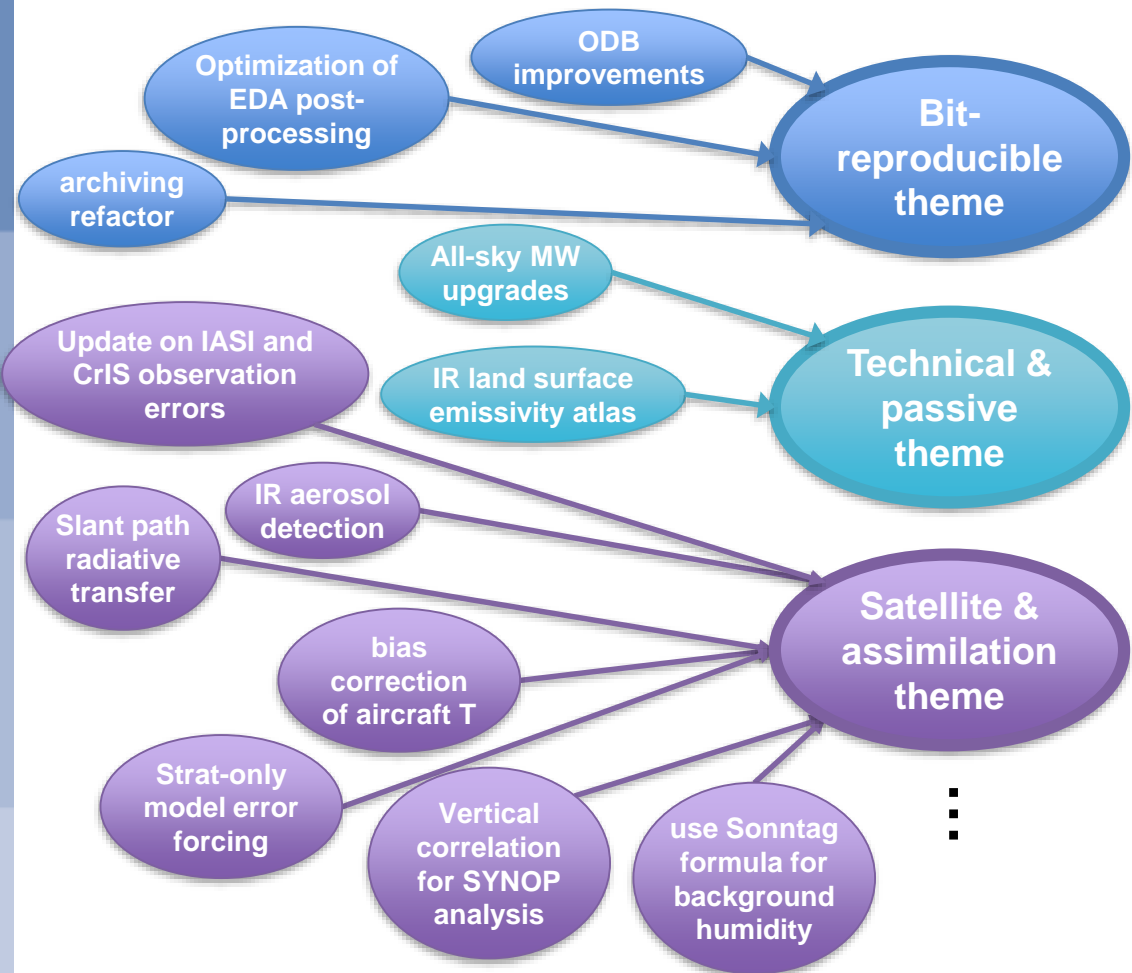


Thanks for listening!

Extra slides

Themed merging

Feature branches Pre-merge branches



Themed merging (2)

