

UFS Infrastructure: Repositories Sub-Group

Presentation to UFS Steering Committee

08 June, 2018

Group Membership

Rusty Benson – NOAA/GFDL (co-chair)

Arun Chawla – NCEP/EMC (co-chair)

Cristiana Stan – GMU (co-chair)

Cecelia DeLuca – NOAA/GSD & CIRES/ESMF

Bill Sacks – NCAR

Gerhard Theurich – NRL/SAIC/ESMF

Seth Underwood – NOAA/GFDL

Mariana Vertenstein – NCAR

Jun Wang – NCEP/EMC

Glossary

Repository – a specific location referenced via URL or other identifier that acts as an archive with tracking capability managed through some revision control system (branches, forks, etc.). For software development a repository may contain source code, an integrated regression test suite/system, documentation, etc.

Authoritative Repository – a repository defined by the presence of a governance group and well-defined process to manage development and periodic releases of reference versions. It exists as the definitive source for a given software development project and must contain regression suite and documentation.

Component Repository – a repository that contains at a minimum source code for a (model) component

UFS Application – model that falls under the definition put forth by the UFS Steering Committee. Examples include: *FV3GFS*; *FV3GFS+CHEM*; *FV3GFS_coupled* (*FV3GFS+CICE5+MOM6+WWIII*); *Regional FV3*; etc.

Guiding Principles

Develop a community-friendly, repository management strategy for the UFS application suite

Flexible enough to meet mission deliverables of agencies while encouraging community contributions

Needs to work with established community development efforts for various model components (land, ice, ocean, etc.)

Utilize best practices demonstrated by successful open-development model components used within the UFS application suite

Identifies a clear pathway for R2O and O2R for UFS applications and their corresponding components

Development for a constituent component in a UFS application must occur in the individual authoritative repository to prevent duplication

UFS Repository Approach

Umbrella strategy to define a UFS application through external references

UFS Umbrella Repository:

- created for each UFS application

- is an *authoritative repository*

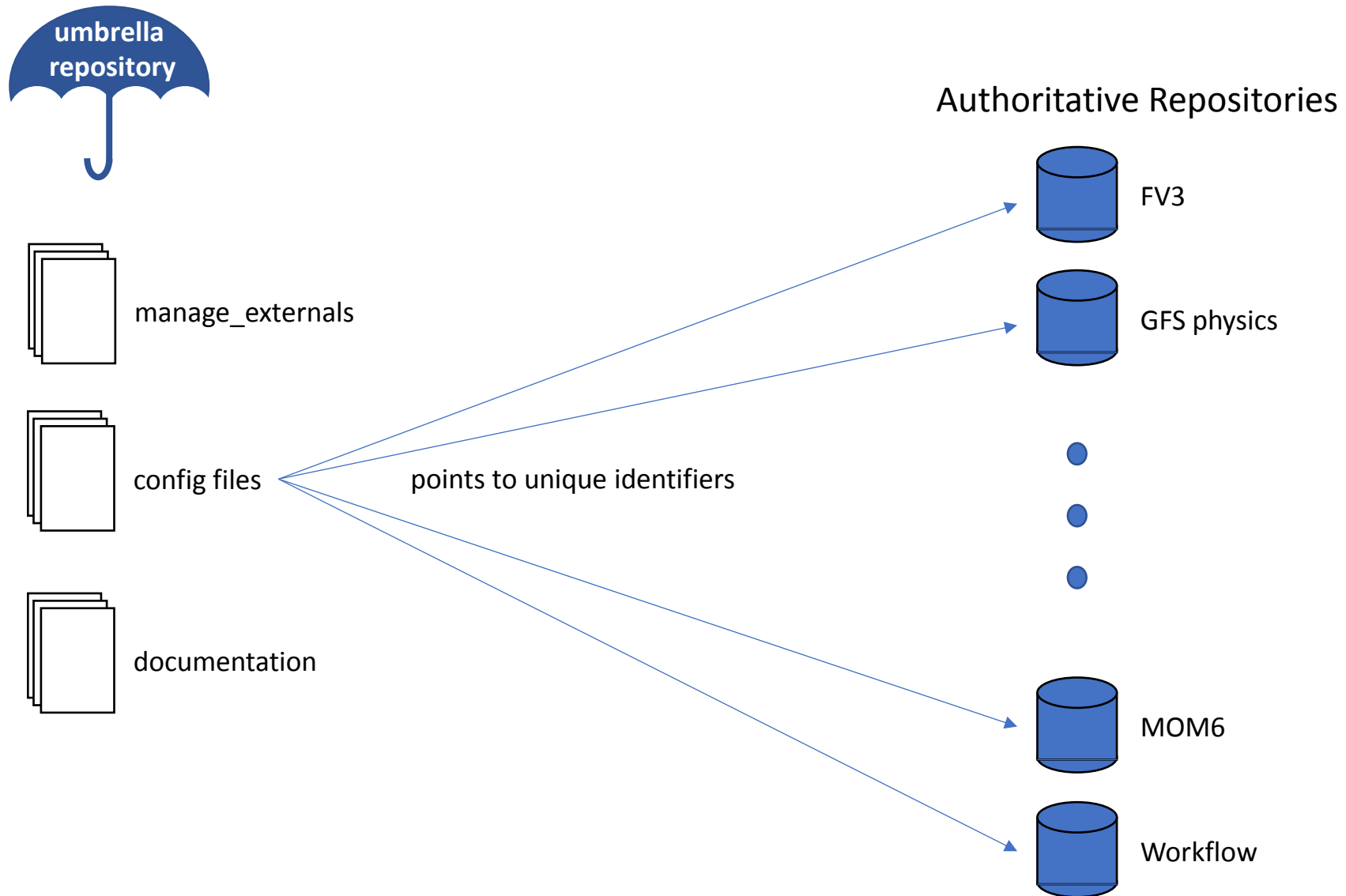
- contains no application source code or data

- populated with configuration files

- must have proper documentation

- relies on connecting tool *manage_externals*

UFS Umbrella Repository



NEMSFv3gfs Use Case

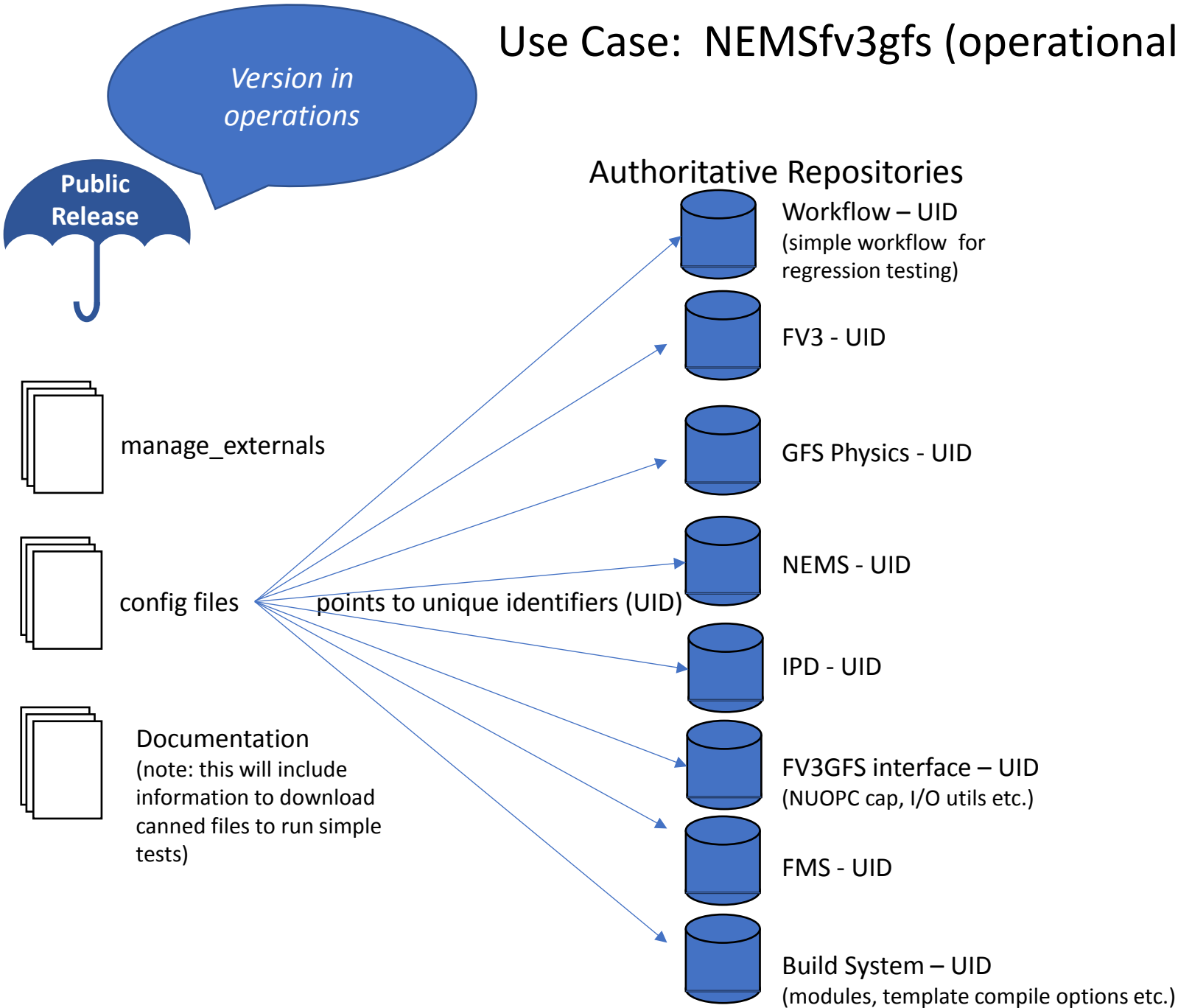
Start from Public Release of operational version

Create a development trunk of the umbrella repository

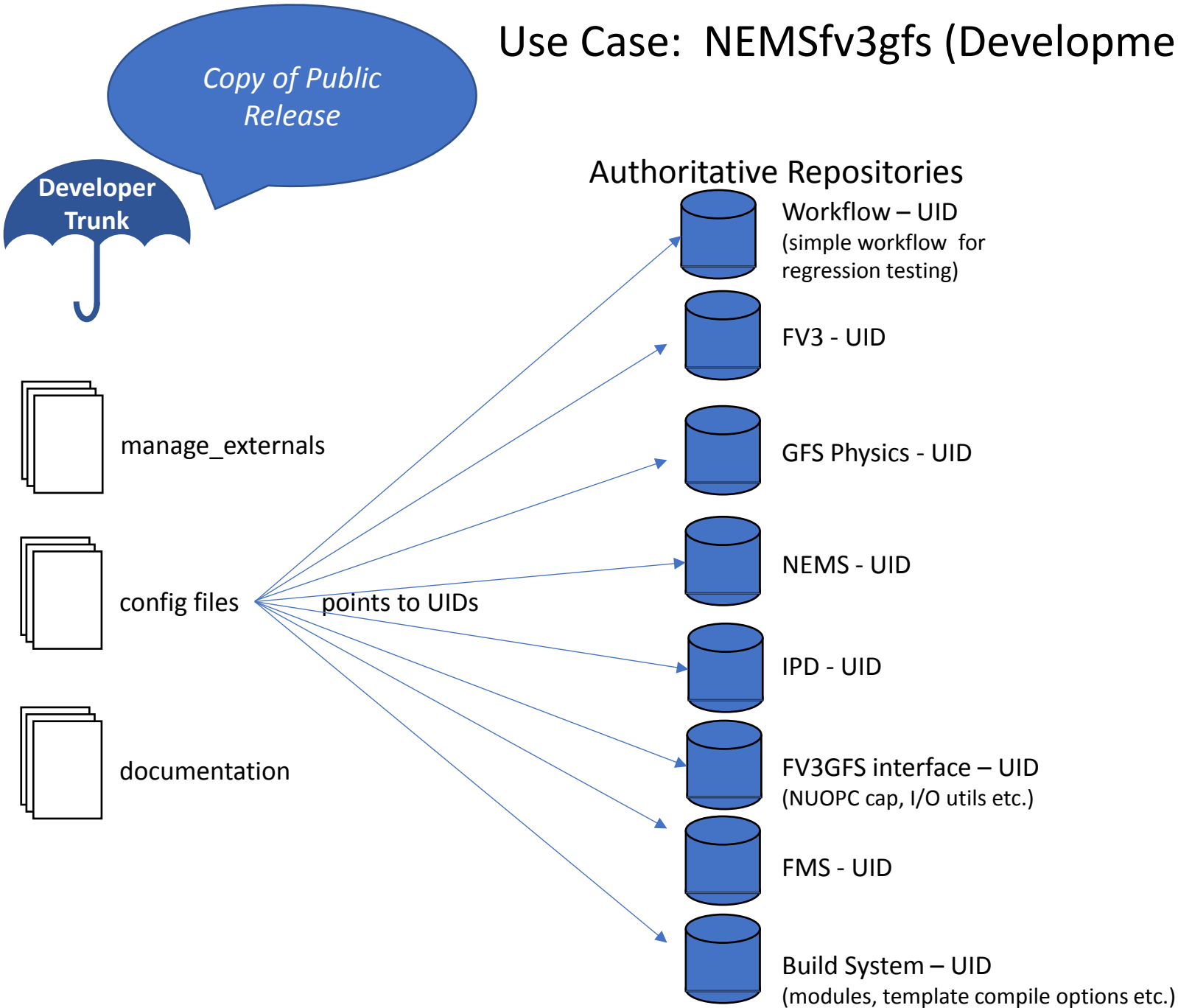
Make changes within a model component following the rules and governance of the specific component authoritative repository(-ies)

Update the config files within the UFS umbrella repository development trunk

Use Case: NEMSFv3gfs (operational release)



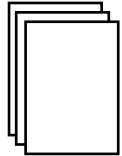
Use Case: NEMSFV3GFS (Development)



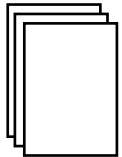
Use Case: NEMSFv3gfs (Development, new physics)

Development follows rules of component(s)

Developer Trunk

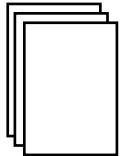


manage_externals



config files

points to UIDs



documentation

Authoritative Repositories



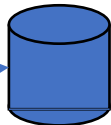
Workflow – UID
(simple workflow for regression testing)



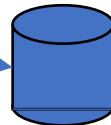
FV3 - UID



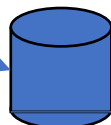
NEMS - UID



IPD - UID



FV3GFS interface – UID
(NUOPC cap, I/O utils etc.)



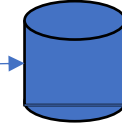
FMS - UID



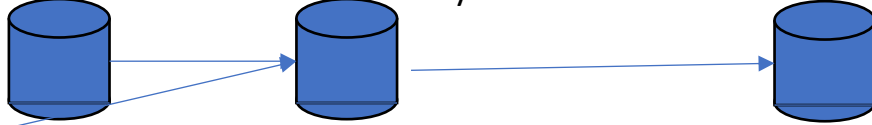
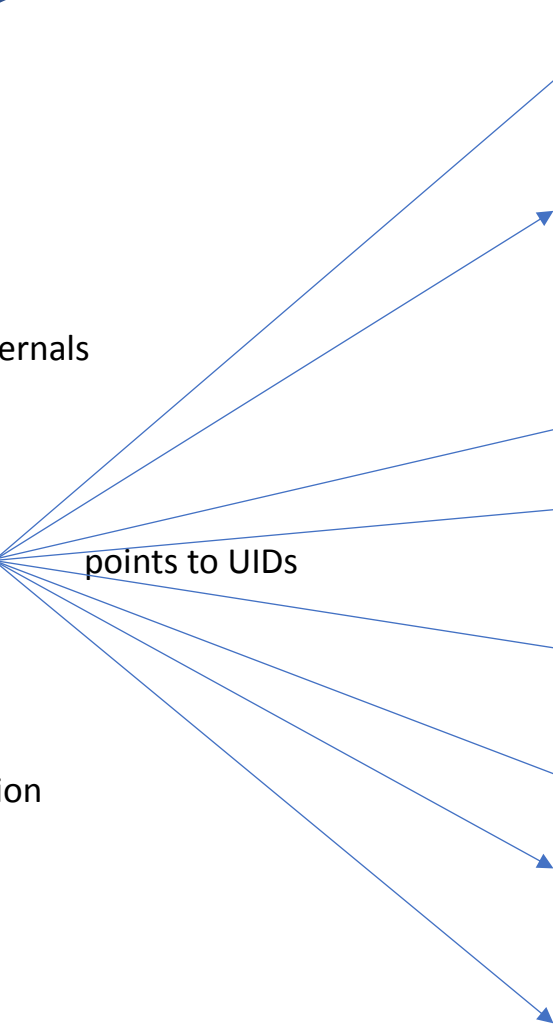
Build System – UID
(modules, template compile options etc.)



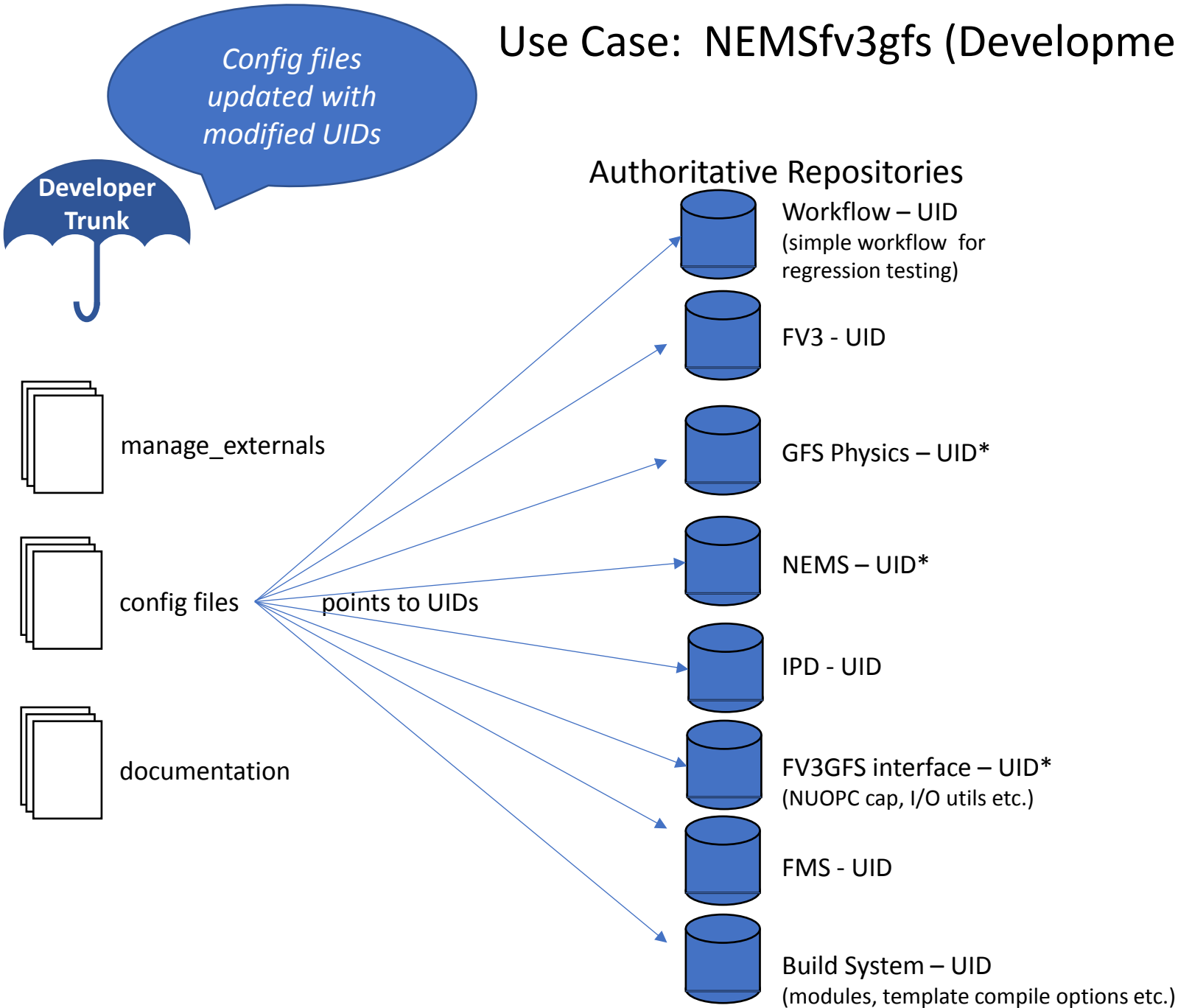
Physics branch – UID*



Pull request back to repo trunk

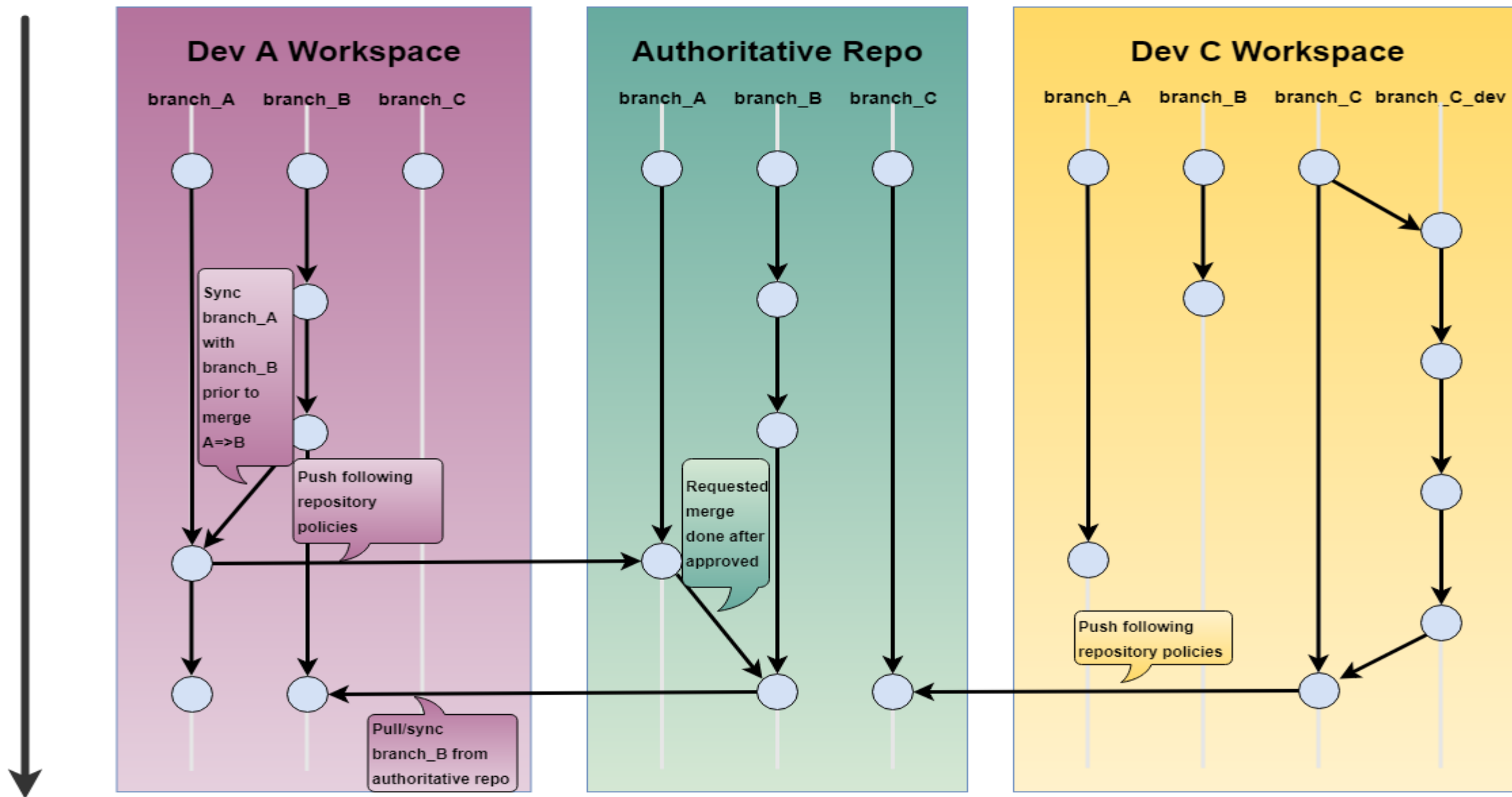


Use Case: NEMSFV3GFS (Development, trunk updated)



Authoritative *Umbrella* Repository.

Time



FAQ

Q. Will the UFS umbrella repository allow anyone to mimic what is in operations?

A. While every attempt will be made to provide operational versions to the community, there may be cases where this is not feasible. Examples, include operational datasets and/or source code with restricted distribution rights.

Q. Is a UFS umbrella repository limited to containing only links to source code?

A. No. All of the elements necessary for execution of a given UFS application, will be provided in some fashion - including supporting application libraries and access to initial conditions and other necessary datasets via a publicly-accessible data portal.

FAQ

Q. What are the resources needed for maintaining umbrella repositories?

A. This depends on the applications that make up the UFS and their complexity. As a rule of thumb, the more complex the development repository and the number of active developers, the larger the number of code managers needed. For example, the WAVEWATCH III repository has a single code manager, but the audience is getting large enough the governance committee is thinking of adding another code manager. For MOM6, the code management duties are split between 2 people. Back of the envelope calculations for NEMSfv3gfs and a coupled FV3-MOM6-CICE5-WWIII system indicates a 6 - 10 FTE effort. This does not include the integration team needed to test the end-to-end modeling systems. This is only sustainable if all the agencies involved make a commitment to share the support functionality

FAQ

Q. Is there expected to be an authoritative repository for each model component? Where are model component issues expected to be submitted?

A. Yes. Model component issues are expected to be resolved in the respective repository where the issue is being tracked. In the event a model component does not exist in an open repository but there is an expectation of active development by the community, an open repository should be created to hold periodic releases with approved community development manually merged into a given release. For utilities (application libraries, etc.) that are not involved in active development we will rely on public releases.

Action Items

Get the agreement of Steering Committee on the umbrella approach

Identify the central repository service: GitHub? GitLab? Atlassian? other?

Codify the governance practices for umbrella repositories

Identify the UFS application suite constituents and assign responsible leads/actors

Explore umbrella repository prototypes where possible for existing UFS applications

Develop a list of requirements for the authoritative component repositories

Open Questions

Who will mobilize the community to become active participants in adoption/creation of a community workflow (necessary for the graduate student test) for reference within the umbrella repository?

What are the plans for a WG or SG to tackle creation of a data portal?

What entity(-ies) will be tasked with training the community-at-large on the umbrella repository rules and procedures?

Who will define the actual process by which development will be considered for operations? Will it be via an annual report put out by the UFS-SC specifying research areas of interest for improvements to the individual UFS applications?