

The Joint Effort for Data assimilation Integration (JEDI)

T. Auligné, Y. Tremolet

Joint Center for Satellite Data Assimilation (JCSDA)

The Joint Effort for Data assimilation Integration (JEDI) aims at developing a community next generation data assimilation framework.

As Earth observing systems are constantly evolving and new instruments are launched, continuous scientific developments are necessary for exploiting the full potential of all available data. Data volumes are also increasing with time, requiring more and more efficient data assimilation. Given the cost and limited lifetime of new observing systems, it is important that this process happens quickly. One goal of the new framework is to enable efficient research work and accelerated transition from research to operations.

As weather and environmental forecasting progress, they are also evolving towards a more comprehensive representation of the Earth system. The framework will have to cater to all components of the Earth system, enabling the evolution towards a fully coupled Earth system.

Data assimilation algorithms themselves are also evolving and progressing to better exploit available data. The framework will include existing operational DA algorithms and facilitate exploration of new DA science across domains and applications.

The supercomputers where data assimilation systems are run are becoming more and more complex with always more and possibly heterogeneous processing elements. Using them efficiently is a growing concern in the community, with most centers exploring ways to improve scalability of their forecasting models and data assimilation systems. The framework will have to take scalability into account and facilitate adaptation to new architectures.

Requirements for operational and research uses are different. Generally, the system will have to be flexible for development and research, and easy for users to learn and test. At the same time, it is imperative that it satisfies operational requirements for robustness, efficiency, coding standards, and maintainability.

The long term objective of the Joint Effort for Data assimilation Integration (JEDI) is to provide a unified data assimilation framework for research and operational use, for different components of the Earth system, and for different applications, with the objective of reducing or avoiding redundant work within the community and increasing efficiency of research and of the transition from development teams to operations. It is important to note that a unified system does not mean a single configuration to be imposed on all partners as each partner can use or develop different applications within the framework.

The requirements above are diverse and many aspects are very complex in themselves or together. However, over the last decade or two, software development technology has advanced significantly, making routine the use of complex software in everyday life. Most recently, communication technology has also progressed to a level where working on the same project across the country has become common practice in the software industry. Together, these technologies make the goals set forward feasible.

The key concept in modern software development for complex systems is the separation of concerns. In a well-designed architecture, teams can develop different aspects in parallel without interfering with other teams work and without breaking the components they are not working on. Scientists can be more efficient focusing on their area of expertise without having to understand all aspects of the system. This is similar to the concept of modularity. However, modern techniques (such as Object Oriented programming) extend this concept and, just as importantly, enforce it automatically and uniformly throughout a code.

In order to facilitate collaborative work, modern software development tools have been and will continue to be used. These tools include version control, bug and feature development tracking, automated regression testing and provide utilities for exchanging this information. This is essential for working across agencies and across time zones. They will be used both for initial development and long term evolution and maintenance.

A summary of the infrastructure and design choices will be presented.