H026

Improving Water Resources Modeling Efficiency using Live Data Communication

Christopher Tubbs[1], Nels J. Frazier[2], Trey Flowers[1]

NWS/Office of Water Prediction/National Water Center[1], Lynker Technologies, LLC[2]

OWP | OFFICE OF WATER PREDICTION

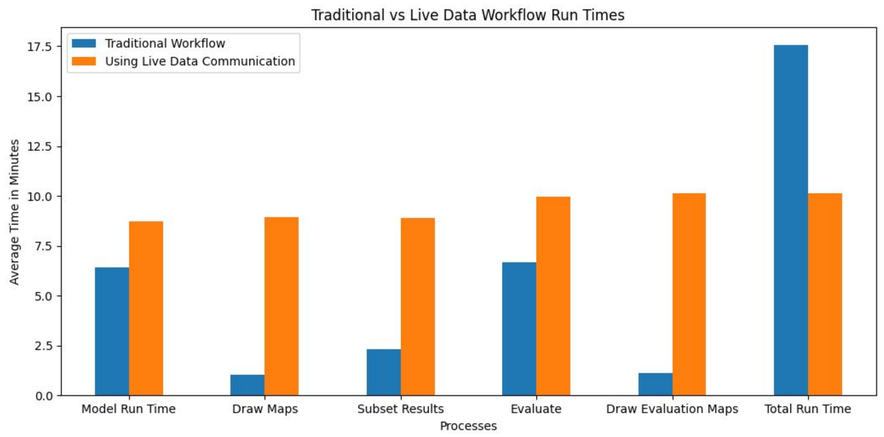
WHAT IS LIVE DATA COMMUNICATION?

- Allows a program to publish data to a channel
- Allows any number of other programs to subscribe to that channel
- Subscribers may act on any data sent through a channel in real time
- Follows the “Publisher-Subscriber” design pattern
- Takes advantage of the Message Passing Technique
- For Example,
  - A model publishes each value, time, and location to a channel upon generation and moves on to the next without waiting
  - A subscriber “hears” the message on the channel and uses the passed message to plot a point on a map immediately
  - Another subscriber “hears” that exact same message and inserts it into a database

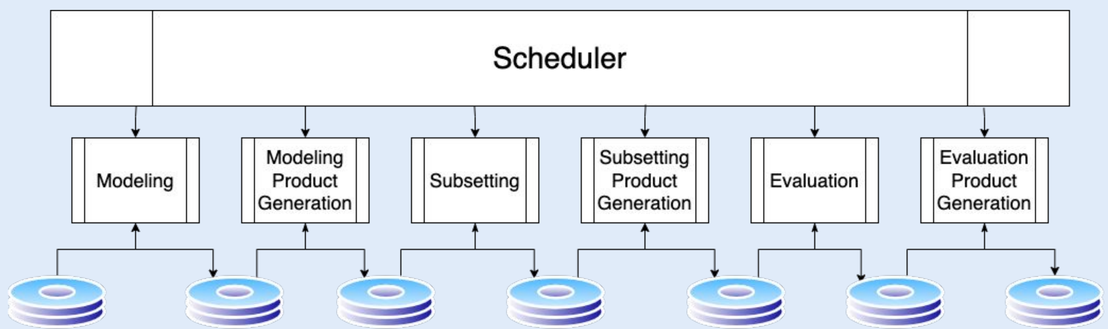
HOW IS THIS NEW OR DIFFERENT?

- Current workflows require one process to complete before another may begin
- Current workflows require dissemination on often large datasets to stage processing
- Current workflows require lengthy coordination between stakeholders
- Current Workflows allow for forecasts to become out of date before data may be transformed and distributed
- Currently,
  - A model must generate and save upwards of tens of millions of values
  - Those values must be disseminated to a common location
  - Other processes must find a way to find and organize the disseminated values
  - Each process must then load those tens of millions of values before any products may be generated

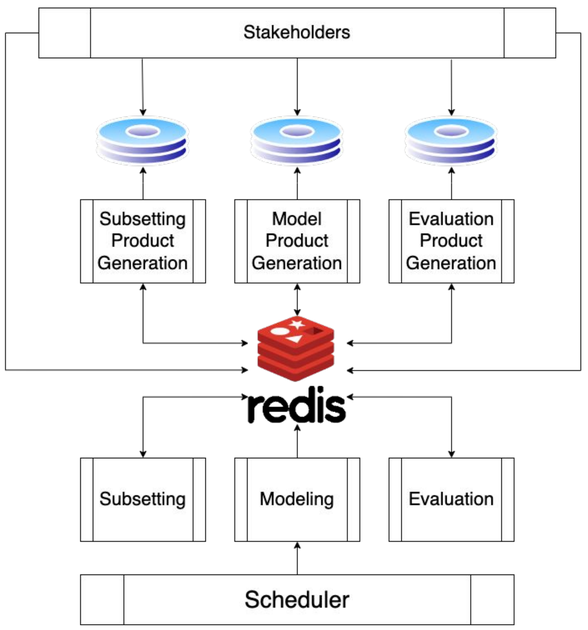
The Overall Processing Time for the Execution of a Water Model and the Generation of its Products is Reduced Dramatically Through the Use of Live Data Communication



Human Useable Products May Be Produced **42% Faster** When Generated with a Workflow that Utilizes Live Data Communication. The total run time of the traditional workflow is bound by the sum of all member processes while the total run time of the live data workflow is bound to the time of the last process to finish.



Classic workflow, with each step launched on schedule



Live Data Workflow, with only the modeling step launched on schedule

TAKEAWAYS

- Splitting processing into distinct units and distributing data immediately allows for new products and processes to be integrated rapidly and safely
- Any stakeholder with access to the network may immediately begin working with any and all data passing through it without interrupting any other users or processes
- Human friendly content may be distributed near instantaneously, bit by bit, instead of having to wait for long periods of time
- Fits comfortably within cloud and distributed environments

Source Code

CONTACT

Website: *https://water.noaa.gov*
Email: *nws.nwc@noaa.gov*
Repository: *https://github.com/NOAA-OWP/SyncVsAsyncModelPipeline*

View my poster and other AGU materials