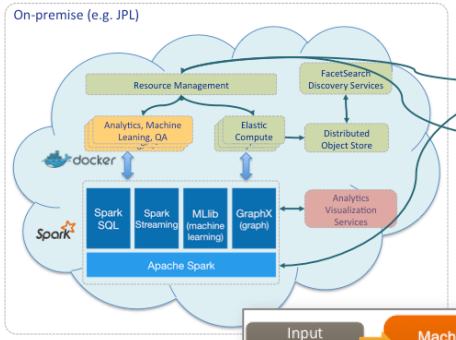
Agile Big Data Analytics of High-Volume Geodetic Data Products for Improving Science and Hazard Response

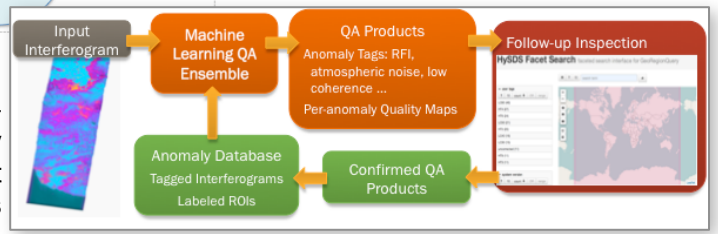
PI: Hook Hua, JPL

Objective

- Develop an advanced hybrid-cloud computing science data system for performing massive-scale analytics of geodetic data products. This includes:
  - Improving the quality of automated data product generation of high-volume and low-latency NASA solid Earth science data products to support hazards monitoring
  - Enabling end-user analysis to be performed on increasing collections of InSAR and GPS data in order to improve the understanding, quality, and features of the data



Scalable Big Data Analytics with hybrid-cloud computing



Machine learning-based quality assurance to detect anomalies

Accomplishments

- Developed a cloud-based system for agile and large-scale SAR analysis that applies machine learning and data analytics to assess and deliver high-quality data products.
- Enabled production of earthquake, flood, and volcano monitoring response products, including automated processing of proxy maps and urgent response inteferograms from Sentinel-1A observations (and others), while lowering delivery latencies from 20+ hours to ~2 hours.
- Infused capability into missions supporting OCO-2 L2 full physics bulk processing, NISAR science data processing (85-430 TB/day expected), and SWOT science data processing (minimum 25 TB/day expected).
- Actively used in supporting analysis for Hurricanes Harvey and Irma through the Applied Science Disaster response program including autonomous response to the M8.1 Mexico earthquake, all in September, 2017.
- Received 2017 AGU Charles D. Falkenberg award in recognition of the impact of this, and related, work.

Co-Is/Partners: Piyush Agram, Sang-Ho Yun, Tom Farr, Eric Fielding, Zhen Liu, Paul Lundgren, Angelyn Moore, Susan Owen, Giangi Sacco, Paul Rosen, JPL; Mark Simons, Caltech; Pietro Milillo, University of Basilicata

$TRL_{in} = 3$ $TRL_{out} = 5$