

B. Osmanoglu, *NASA Goddard Space Flight Center*, F. J. Meyer, *University of Alaska Fairbanks*, H. Caumont, *Terradue Srl*

The Future of SAR Processing is in the Cloud

Overview

Synthetic Aperture Radar (SAR) is an Earth observation tool growing in popularity because of an increase in free-and-open data and by the emergence of new sensing systems with global and regular observation strategies. Today SAR analyses can take advantage of growing image time series containing hundreds of regularly sampled and cloud-free images, creating new possibilities and new challenges for the radar remote sensing community. With each image being more than 1GB in size, it is becoming more and more difficult for end-users to manage the onslaught of SAR data on their personal computers and local storage devices. Hence, cloud computing platforms are starting to play a more prominent role as a convenient means for data processing and access. In this column we consider the underlying reasons for this transition to the cloud.

The volume of SAR data has been increasing, mainly due to the larger coverage and higher resolution of SAR products. There are also more SAR satellites collecting data – both from international civilian SAR agencies, and the commercial sector. Cloud computing costs are trending lower, while the costs of local server infrastructure have remained the same.

There are several SAR data processing platforms in development such as the European Space Agency's (ESA) Thematic Exploitation Platforms (TEP), Jet Propulsion Laboratory's (JPL) Advanced Rapid Imaging and Analysis (ARIA), and Alaska Satellite Facility (ASF) OpenSARLab and Hybrid Pluggable Processing Pipeline (HyP3) environments.

Trends

Looking at today's state-of-the-art and trends extrapolated from available data, one can begin to outline some expectations for the near future, while noting that predicting the future of any technology is difficult.

Volume of SAR Data

The volume of spaceborne SAR data has been growing as data downlink capabilities and sensor designs improve, allowing more regular global coverage at higher spatial resolutions. In Figure 1 we estimate the stripmap image size given instrument parameters, such as range and azimuth ground sampling, image swath and length, and single look complex (SLC) byte size. It must be noted that the Sentinel-1 and NISAR routinely use beam-steering methods for wider area coverage, and they are not stripmap in the strict sense of the word. The beam steering methods allow imaging at multiple look angles increasing swath width, without sacrificing resolution. It is also important to note that both polarizations of Sentinel-1 SLC are packaged together, resulting in doubling of the file size. Similarly, dual polarization NISAR data will be placed in a single HDF5 file, also resulting in the doubling of the file size. [Figure 1] shows that an exponential trendline fits the distributed file size rather well.

There might be minor differences between the data size reported here as we are not considering the metadata, or varying pixel size in radar geometry. The calculation is purely based on the coverage, pixel size, data format, and number of bands in the product ignoring data compression:

$$PackageSize = \frac{\Sigma az \times \Sigma rg}{\Delta az \times \Delta rg} \times Q \times B$$

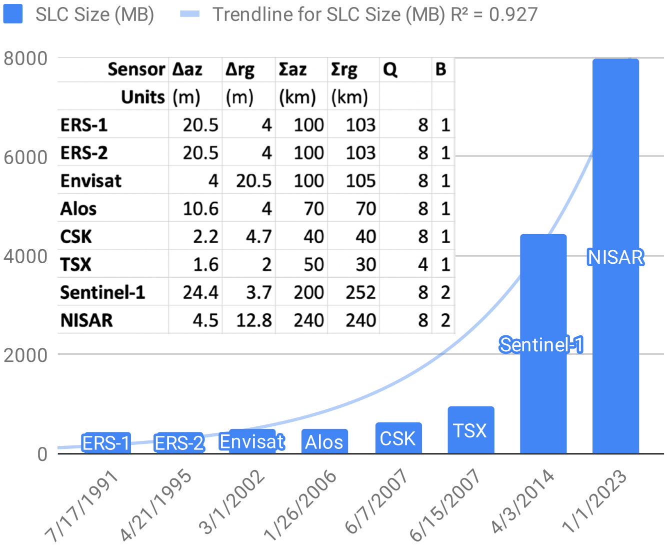
where Σaz and Σrg are the nominal size of imagery in meters for azimuth and range directions, Δaz and Δrg are the nominal sizes of azimuth and range pixels (i.e. sample distance, not resolution), Q is the number of bytes needed to represent the data type for each complex value (e.g. 8 bytes for Complex Float32), and B represent the number of imaging bands in the data package (e.g. 1 for single polarization data and 2 for dual polarization data).

Number of SAR Missions

Until 2000 there were one or two active SAR satellites in space, today there are at least ten SAR sensors collecting data. Including the commercial satellites, there will be many more in the near future. Furthermore, constellations and shorter repeat intervals will result in increased imaging opportunities. Figure 2 shows the time lines for various SAR programs from 1991 to 2025. ALOS-4 is expected to be launched in the next few months. Beyond that, ESA's BIOMASS, Sentinel-1C and

Photogrammetric Engineering & Remote Sensing
Vol. 87, No. 1, January 2021, pp. 5–8.
0099-1112/20/5–8

doi: 10.14358/PERS.87.1.5



Sensor	Δaz	Δrg	Σaz	Σrg	Q	B
Units	(m)	(m)	(km)	(km)		
ERS-1	20.5	4	100	103	8	1
ERS-2	20.5	4	100	103	8	1
Envisat	4	20.5	100	105	8	1
Alos	10.6	4	70	70	8	1
CSK	2.2	4.7	40	40	8	1
TSX	1.6	2	50	30	4	1
Sentinel-1	24.4	3.7	200	252	8	2
NISAR	4.5	12.8	240	240	8	2

Figure 1. The calculated file size vs. launch date of the missions show a clear trend towards larger files. In the table *Δ* is resolution, *Σ* is swath size, *az* is azimuth, *rg* is range *Q* is the number of bytes used to represent each pixel and B denotes number of channels in the package. TerraSAR-X uses Float16 and all others use Float32.

-1D are expected to launch in 2022, and NISAR is expected to launch in 2023. There are also future missions in planning stages such as ROSE-L (ESA), TanDEM-L (DLR).

Many commercial SAR companies are also planning constellations with many satellites. For example, as of November 2020 ICEYE has four satellites in orbit and is planning to have potentially 18 satellites. Capella Space has one satellite in orbit and is planning to have a constellation of 36 satellites. Timelines for deployment of complete constellations depend on funding. Availability of commercial SAR platforms will lead to more frequent observations anywhere on earth, which can especially be important for urgent response activities.

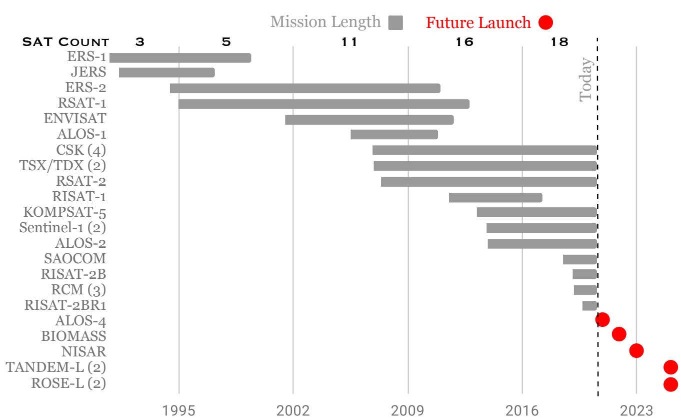


Figure 2. The timeframe of civilian space agency SAR missions show the increase in number of operational satellites. Number of satellites in missions are shown in parenthesis if more than 1. Number of operational satellites for each period are shown at the top.

Cloud Computing Costs

In addition to increasing file size, cloud computing costs have been declining over time and seem to follow an exponential relationship (Figure 3). Looking at the on-demand cost for

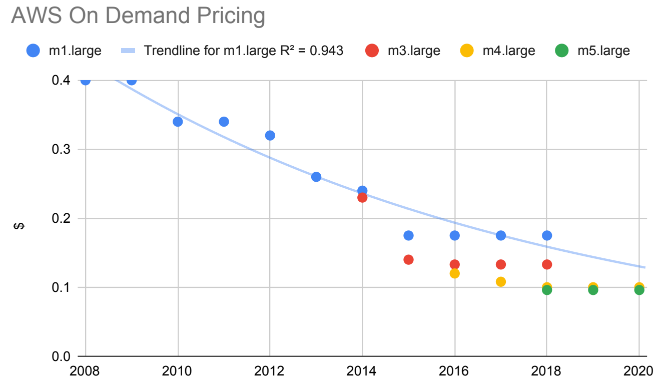


Figure 3. On demand pricing for Amazon Web Services platform indicates a decreasing trend in processing costs.

general purpose instances on Amazon, we see that the hourly cost on an m1.large instance reduced ~50% in a decade (1). This is likely due to competition and the lowering cost of hardware over time (e.g., technological advances in computational power). Both data upload and download from the cloud incur costs, the former usually being far cheaper. Data download costs can be significant, but strategic placement of data processing algorithms in the same cloud location as the data allows for the end-users to download only the final result, without incurring the cost of egress (downloading) of the SAR data.

Cloud Computing Software

Several SAR processing software elements (ISCE, SNAP, GAMMA, etc.) (2–4) are being implemented in cloud platforms, such as JPL's ARIA, ESA's TEP, ASF's HyP3 and OpenSARLab. JPL's ARIA is a cloud-based processing system with a web interface designed for rapid mapping of solid earth deformation. For event response, the results can be accessed anonymously at the ARIA-Share site. ARIA is also capable of creating standing orders for an area of interest and can also trigger processing of interferograms using the USGS shake map (5). Standard data products are accessible through the ARIA-products website.

ESA's Geohazards TEP (Thematic Exploitation Platform) is one of the several TEPs supported by ESA to allow for the sharing of data and algorithms among users (6). Geohazards-TEP or GEP, houses a collection of SAR data processing workflows. These workflows span a wide spectrum of applications such as simple interferometry, deformation time series, false color band combinations along with optical processing algorithms. There is also a social aspect of the platform as results and algorithms can be shared with other users. The TEP allows users to upload data and provides developer sandbox environments as virtual machines, including software frameworks such as JupyterLab and SDKs for standardized data discovery and access, for the integration of additional workflows "As-a-Service".

ASF's HyP3 has an easy-to-use web interface for several SAR data processing workflows (7). Similar to ARIA and TEP, it can generate single interferograms or time-series along with radiometric terrain corrected intensity time-series for change detection applications. ASF also provides OpenSARLab, where developers can author Python Notebooks to do their own analysis and go beyond what HyP3 currently offers. Such notebooks can be shared with other users through public GitHub repositories (8,9).

As noted above there are some similarities between ARIA, GEP and HyP3. They all provide a web interface to the cloud infrastructure that handles the SAR processing, and they all can create standing orders. Also, to some extent they use Python language for data analysis, even though underlying algorithms might also use other languages. Both GEP and HyP3 provide users with a mechanism to bring their own algorithm to the cloud. JPL is also working on a mechanism to allow users to bring their algorithms to the cloud, where massive NISAR data will be located (10).

Going Further

There are still ways to improve the user experience with cloud processing. Python is emerging as the de facto programming language with GEP and OpenSARLab already supporting python notebooks. GitHub is also a common denominator as all three platforms use GitHub for development. With a common data handling library, developers would be able to easily port algorithms between different cloud services. This common library would also allow for combining processing flows from different platforms.

One of the potential challenges faced using multiple cloud-based platforms jointly is the data egress cost. Both JPL and ASF are leveraging Amazon Web Services while GEP is using a Hybrid Cloud strategy. GEP uses Terradue Cloud Platform for the processing services' integration resources, and well-established Cloud Providers (AWS, Google Cloud, and in Europe the Copernicus DIAS, Hetzner, etc.) opportunistically for the processing services' production resources. For user convenience it is important to have a cloud-to-cloud data transfer mechanism, so that users do not have to download data or interim results to a personal computer. It must be noted that egress costs are not always paid by the end-user. Downloading data from a space agency results in the space agency incurring costs. Furthermore, moving data across buckets in the same lake (large data holding) does not incur egress costs. Therefore, developing processing algorithms that can work without downloading the data, or subsetting algorithms that can extract only the necessary part of the data files can further reduce costs.

There are also many other datasets beyond SAR that are being staged on the cloud that could help with analysis. Many earth science datasets (e.g., optical imagery, weather models) are already in the cloud, and their combination with SAR data is a very active research area for many applications. Finally, providing adequate training to users of cloud processing is necessary to get the most out of the cloud processing platforms. In the end, a tool is only as good as the hands that wield it.

Conclusion

Given the increasing amount of data SAR data analysis routinely uses these days, it is likely that the future of SAR processing will be in a cloud computing environment. ARIA, Hyp3, and GEP are three cloud-based processing environments to achieve this. They all try to provide the users a convenient interface for SAR processing for different workflows. Working towards a common data handling library to allow developers to generate common workflows and even allowing users to combine workflows from multiple platforms could further accelerate user adoption.

References

Supernor B. Why the cost of cloud computing is dropping dramatically [Internet]. [cited 2020 Oct 7]. Available from: https://appdevelopermagazine.com/why-the-cost-of-cloud-computing-is-dropping-dramatically/.

Rosen PA, Gurrola E, Sacco GF, Zebker H. The InSAR scientific computing environment. In: EUSAR 2012; 9th European Conference on Synthetic Aperture Radar. ieeexplore.ieee.org; 2012. p. 730–3.

Zuhlke M, Fomferra N, Brockmann C, Peters M, Veci L, Malik J, et al. SNAP (Sentinel Application Platform) and the ESA Sentinel 3 Toolbox. Vol. 734. ui.adsabs.harvard.edu; 2015. 21 p.

Werner C, Wegmüller U, Strozzi T, Wiesmann A. Gamma SAR and interferometric processing software. In: Proceedings of the ERS-Envisat Symposium, Gothenburg, Sweden. Citeseer; 2000. p. 1620.

Owen SE, Hua H, Pan L, Bekaert DP, Manipon G. Geodetic Imaging with Large SAR Data Sets using NASA High End Computing. In Proceedings of AGU Fall Meeting; 2019. p. G13C – 0573.

Fitrzyk M, Engdahl M, Fernandez D. ESA Copernicus Sentinel-1 Exploitation Activities. In: IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium. ieeexplore.ieee.org; 2019. p. 5389–92.

Meyer FJ, Ajadi OA, Schultz L, Bell J, Arnoult KM, Gens R, et al. An Automatic Flood Monitoring Service from Sentinel-1 SAR: Products, Delivery Pipelines, and Performance Assessment. In: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. ieeexplore.ieee.org; 2018. p. 6576–9.

Meyer FJ, Lewandowski A, Lundell E, Johnston A, Hogenson K, Agram P, et al. asf-jupyter-notebooks [Internet]. Alaska Satellite Facility; 2020 [cited 2020 Nov 4]. Available from: https://github.com/asfadmin/asf-jupyter-notebooks.

Burns R, Aivazis M, Gunter G, Fattahi H, Hawkins B, Shiroma G, et al. isce3 [Internet]. Github; 2020 [cited 2020 Nov 4]. Available from: https://github.com/isce-framework/isce3.

Moses JF, Baynes K, Behnke J. NASA Technical Reports Server (NTRS) [Internet]. ntrs.nasa.gov; 2018 [cited 2020 Nov 10]. Available from: https://ntrs.nasa.gov/search.jsp?R=20180008557.

Authors

Batuhan Osmanoglu is a Research Physical Scientist at NASA-GSFC, holds a Ph.D. in synthetic aperture radar (SAR) interferometry. He has been a part of Istanbul Technical University, University of Miami, University of Alaska Fairbanks, and he has worked on applications for observing surface deformation, measuring target velocities, and radar design. He serves as one of the NISAR Deputy Program Application Leads, candidate architecture deputy-lead for Surface Deformation and Change mission study, and one of the Disasters Program coordinators at GSFC.

Franz Meyer is a Professor for Radar Remote Sensing at the University of Alaska Fairbanks (UAF). He is also the Chief Scientist of the Alaska Satellite Facility, NASA's prime Distributed Active Archive Center (DAAC) for radar remote sensing data. He currently holds several leadership positions in internationally recognized professional organizations, including IEEE (Institute of Electrical and Electronics Engineers), ISPRS (International Society for Photogrammetry and Remote Sensing), and CEOS (Committee on Earth Observation Satellites).

Hervé Caumont is Terradue's Operations Officer in charge of supporting science and commercial users on Terradue Cloud Platform. As a senior architect, he's in charge of developing and maintaining Terradue's business relationships across international projects and institutions, managing corporate programs for business development, product line innovation, and solutions marketing. Hervé has long-standing expertise in the development and transfer to operations of scientific, environmental and military solutions delivering geospatial informations. He holds a Masters of Sciences in Electronics Engineering, and a Masters Degree applied in Remote Sensing and Image Processing from the University of Toulouse.