Poster presentation at the Library & Information Services (LISA) VIII Conference "Astronomy Librarianship in the era of Big Data and Open Science", 6 - 9 June 2017, European Doctoral College, Strasbourg, France

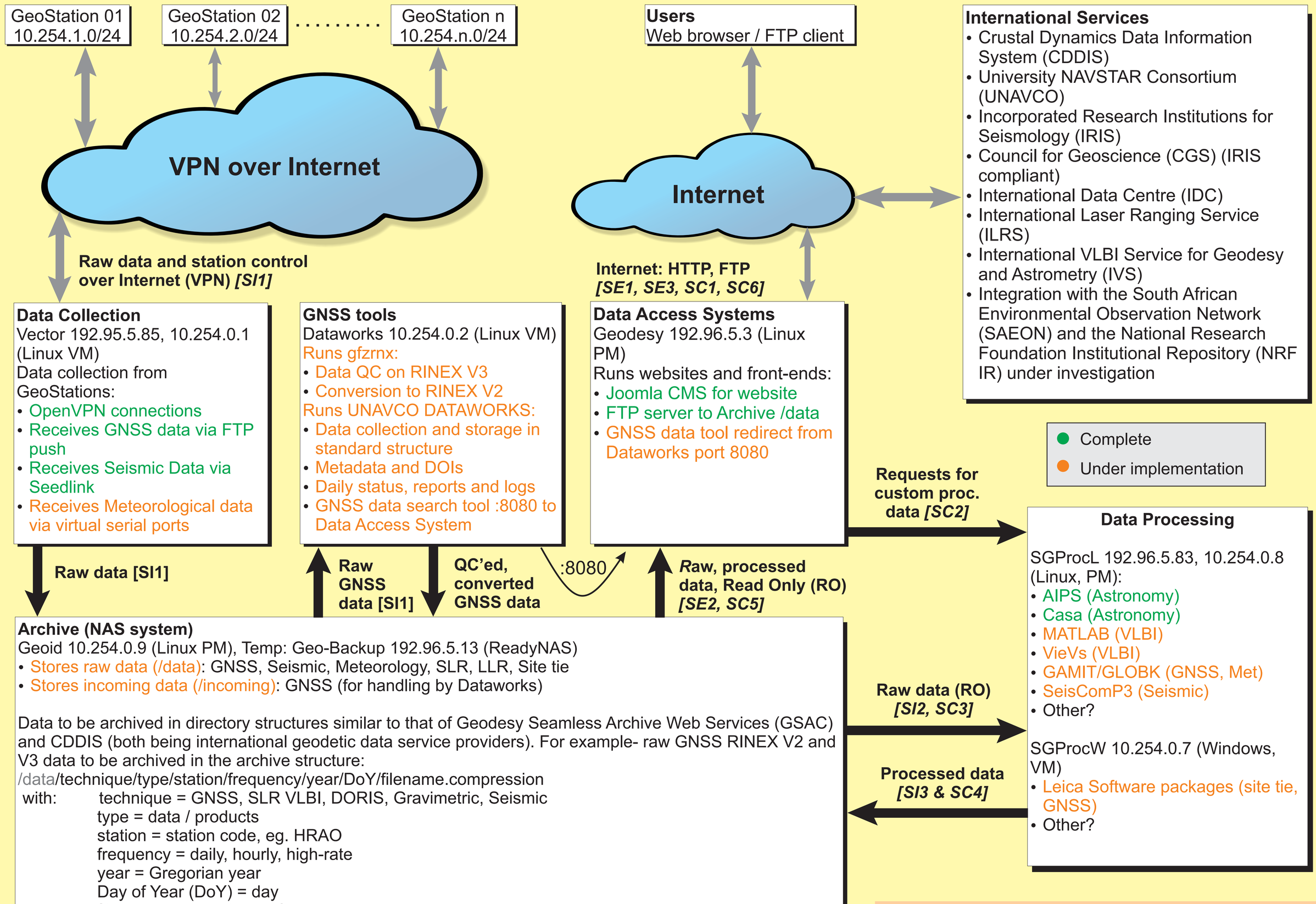
Progress with the new research data management system at HartRAO

Glenda Coetzer[1, 2], Roelf Botha[1] and Lorette Jacobs[1, 2]

[1]Hartebeesthoek Radio Astronomy Observatory, South Africa
[2]University of South Africa

Corresponding Author: glenda@hartrao.ac.za

The Hartebeesthoek Radio Astronomy Observatory (HartRAO) participates in global astronomic and geodetic research activities. Astronomy activities focus largely on the use of the 26-m radio telescope to conduct astronomical observations- single-dish observations and astronomical Very Long Baseline Interferometry (VLBI), whilst geodetic activities focus on using the HartRAO 15-m and 26-m radio telescopes for geodetic VLBI, Global Navigation Satellite Systems (GNSS) positional reference stations, weather data, seismic systems, satellite laser ranging (SLR) and Doppler Orbitography and Radiopositioning (DORIS) systems. Added to the existing instruments and techniques are gravimetric instruments, a Lunar Laser Ranger (LLR) and, in the near future, a VLBI Global Observing System (VGOS) telescope. HartRAO is also participating in the African VLBI Network (AVN), a network of radio telescopes in Africa. Data and data products produced by HartRAO's expanded range of on-site and off-site instruments must be archived and stored at HartRAO and made accessible to the scientific community. The data management and storage systems currently being used have certain drawbacks, such as being distributed and outdated as well as having a limited capacity to manage additional large data volumes, types and user requirements. This necessitated the design and implementation of a new, next-generation Geodetic Research Data Management System (GRDMS), which will comply with internationally accepted standards. Main objectives of the system are to organise, structure and store geodesy and geodynamics related data and data products in a central data bank, maintain information about the archival of the data and disseminate data, data products and information in a timely manner to the global research community. Components of our data management system will be similar to and incorporate the same software as that of the Crustal Dynamics Data Information System (CDDIS) and University NAVSTAR Consortium (UNAVCO). Data structures and file-naming conventions of the CDDIS and UNAVCO will be used for all geodetic data. Each dataset will receive persistent interoperable identifiers, Digital Object Identifiers (DOIs). A web-based graphic user interface (GUI) for the dissemination of data and data products will be provided to users. We present progress to date on various sub-systems as well as a top-level conceptual model of the GRDMS.



Internal Steps [SI] data flow cycle:
- [SI1]: The data collection (Vector) and storage of raw data from the various stations on the Archive
- [SI2]: Raw data is streamed to the Data Processing Unit for processing towards data products
- [SI3]: Processed data are sent and stored in specified formats and structures in the Archive

External Steps [SE] data flow cycle:
- [SE1]: The scientific community can interact with and request data via a website (HTTP) and/or FTP
- [SE2]: The Data Access System obtains the requested data from the Archive
- [SE3]: Requested data are packaged into a single compressed file and made available

Custom Steps [SC] data flow cycles (for simple requests via the online interface):
- [SC1]: Requests are submitted on the website via a special interface
- [SC2]: Requests are translated to a script and sent to the Data Processing Unit
- [SC3]: The Data Processing Unit obtains the required raw data and processes it
- [SC4]: Processed data is sent to the Archive for storage
- [SC5]: The Data Access System retrieves processed data
- [SC6]: Results are sent to the requesting user

Acknowledgements

The authors would like to acknowledge funding awarded by the National Equipment Programme (NEP) of the National Research Foundation (NRF) for the development of the co-located academic network.

References

Coetzer, G.L. et al. 2015, Library and Information Services in Astronomy (LISA) VII. ASP Conference Series, Vol. 492, p22

Noll, C. 2017, IVS 2015-2016 Biennial Report, https://cddis.nasa.gov/docs/2016/cddis_IVSreport_2015_2016.pdf

Noll, C. & Michael, P. 2016, IGS Workshop "GNSS Futures", February 08-12, 2016, Sydney NSW, Australia.