



COCONet Regional Data Center Request for Proposals



As part of the National Science Foundation- funded COCONet project, UNAVCO will provide startup funding for up to four host institutions to operate Regional Data Centers. As many as three institutions will be selected to host a **Regional Mirror Data Center**. One institution will be selected to host the more advanced **Regional Data Center**. Each institution selected will be eligible for up to two years of funding. Each institution will receive all required hardware and software from UNAVCO. Onsite installation and training support from UNAVCO will be provided during the first year of funding; ongoing operations support via phone and e-mail will be provided during the second year. Proposing institutions will be expected to provide Information Technology infrastructure (power, cooling, internet, networking) as an in-kind contribution to the Data Center.

Proposal Deadline for 2014 Funding: September 6, 2013

Funding Levels:

Regional Mirror: Up to \$10K/yr for 2 years

Regional Data Center: Up to \$20K/yr for 2 years

Start Date for Funding: 2014 funding starts October 15, 2013; 2015 funding starts October 15, 2014

Format: Proposals should be no more than 5 pages and should include the following:

- (1) Benefits to the COCONet community of hosting a data center at the proposing institution
- (2) Benefits to the proposing institution of hosting a data center
- (3) Description of the capabilities of the host institution including existing computational, systems administration, and software assets and resources available, and any current or former IT-dependent projects at the host institution.
- (4) Ability of the host institution to sustain data center hosting beyond the first two years of funding
- (5) Budget
- (6) Budget Justification

Curriculum Vitae for the Principal Investigator and Senior technical personnel (2 page limit in NSF format preferred) must be included but do not count within the 5-page limit.

Selection: The COCONet Steering Committee will evaluate proposals beginning September 16, 2013.

Notification: Notification of selected proposals will be September 20, 2013



Additional Information for Proposers

Specifications for COCONet Regional GPS/GNSS Data Center Hosting

Overview

To serve the needs of the COCONet community of scientists, educators, students at all levels, and the public, one or more Regional Data Centers (RDCs) are envisioned that would be hosted at an institution in the Caribbean region. As detailed in this document, a COCONet RDC can be configured as a Regional Mirror Data Center (RM) or a stand-alone Regional Data Center (RDC) depending on the desires and capabilities of the host institution to provide information technology (hardware and software support). A RM functions as a mirror for COCONet data and metadata holdings from the COCONet primary archive at UNAVCO; metadata and data would also be served by the RM. A RDC can include RM mirroring functions, but also has the capability for local data and metadata management including downloading stations and archiving GNSS data for the host institution and serving COCONet streaming data. At least one RDC is envisioned in the Caribbean region, along with one or more RMs.

UNAVCO will architect the systems, software, web, and database capabilities for each level of RDC. The hardware and software will be delivered as a turn-key system. Most of the components already exist within systems used operationally at UNAVCO.

Initial onsite installation support will be provided by UNAVCO, along with training of data center staff in operation and maintenance of the hardware and software systems. Updated software releases for functional aspects of the data center operation will be provided through time to address bugs and provide functionality enhancements as needed. Support for ongoing operation of the RM or RDC will also be provided by UNAVCO, principally through e-mail.



The RM and RDC require differing commitments of resources by the host institution, ranging from hardware hosting and very minimal local system administration and web support, to hardware hosting plus modest to significant system administration, software engineer, web administrator, and geodesy data engineer support for the RDC.

This document describes the requirements and guidelines for institutions from the Caribbean region participating in COCONet to respond to the Request for Proposals for Regional Data Center to be released in early-August, 2013.

Motivation

As part of the COCONet proposal, one or more Regional Data Centers for COCONet data are to be implemented in the Caribbean region with the support of UNAVCO. The purpose of hosting RDCs within COCONet is to build capacity so that locally-generated data and metadata are locally-managed, archived, curated and distributed from within the Caribbean region. The implementation and operation of a RDC will ensure that the capacity-building that is a signature element of the COCONet project will extend to the management of the data life cycle, without which the full benefits of the COCONet project for the region cannot be realized.

UNAVCO Support

UNAVCO has two decades of experience with data and metadata management and for data curation and distribution and has developed relevant software systems that can be scaled and employed by the COCONet Regional Data Centers. While any RM or RDC host institution will need to have some information technology capabilities, they will leverage database, software, web, and data management technologies developed by UNAVCO, leapfrogging years of development effort. The availability of technical expertise at the host institution will determine readiness for RM or RDC hosting. UNAVCO support will include delivering a turnkey system consisting of hardware and necessary software components, onsite installation, configuration, and startup, onsite training, and post implementation support.

Regional Mirror and Regional Data Center Specifications

The functions encompassed by the RM and RDC are indicated in Table 1. Both implementations leverage the Geodesy Seamless Archive Centers (GSAC) software technology developed at UNAVCO with support from USCD and CDDIS. The RM allows for local holding and distribution of data and metadata and relies on mirroring of data and metadata from a primary GSAC-enabled archive such as UNAVCO. The RDC includes mirroring some data and metadata from a primary GSAC-enabled archive and adds (1) local capability for downloading and data management and archiving from additional stations managed by the host institution; (2) management of metadata for locally managed stations; and (3) local stream management and broadcast capability.

Implementation of at least one RDC will build capacity so that locally-generated data and metadata are locally-managed, archived, curated and distributed from within the Caribbean region. UNAVCO will deliver to the RM/RDCs, systems that have been pre-configured at the level that the host institution is ready to support; this is the “Archive in a Box” concept involving turnkey components. Most of the components already exist within systems used operationally at UNAVCO. These components will be integrated and installed on the hardware to be delivered to the host institution as part of the implementation of the RM/RDC at the host institution.

| Table 1. Functions, systems and services of the Regional Mirror and Regional Data Centers | | |
|--|---|---|
| Level | Functions | Required systems and services |
| Regional Mirror | -host and serve metadata mirrored from UNAVCO -host and serve data mirrored from UNAVCO -GSAC Repository and web site | Virtual machine hardware, database, storage, software systems and web server, plus UNAVCO software components. |
| Regional Data Center | Regional Mirror functions plus: -local data download from stations -local streaming raw data -local metadata management -local data curation/preservation | Regional Mirror systems plus: Stream management system (e.g. VRS), Ntrip caster Cloud-based backup service Additional UNAVCO software components |

Tables 2 and 3 show the hardware and software components that support the functions implemented at the RM and RDC, respectively. In general terms, the hardware required is one server with virtual machine capability and storage for database and data holdings. Data holdings will be most conveniently managed with a separate RAID system. For the RDC a minimum of one additional virtual machine server system will be required. All components are delivered to the RDC at the installation phase. The increased functionality for the RDC will entail additional local expertise for ongoing operation of the RDC system as shown in Table 4.

| Table 2. Regional Mirror | | |
|---------------------------------|--|--------------------------------------|
| Hardware - All functions | System(s) | Licence(s) |
| | Virtual machine capable hardware (e.g. <i>TBD</i>) | VMWare or similar, CentOS (freeware) |
| | RAID Storage (e.g. <i>TBD</i>) | |
| Function | Software System | Technologies required |
| Host metadata | Site Metadata Manager database* | MySQL |
| Serve metadata | Geodesy Seamless Archive Centers web site and clients* | Java, Ant, Apache, tomcat |
| Host and serve data | | File system, FTP |
| Manage metadata mirroring | GSAC Management Clients* | Python |
| Manage data mirroring | GSAC Management Clients* | Python |
| * Provided by UNAVCO | | |

| Table 3. Regional Data Center | | |
|---|---|------------------------------|
| Hardware - All functions | System(s) | Licence(s) |
| | Same as RM | Same as RM |
| Function | Software System | Technologies required |
| All RM Functions | All RM Functions | All RM Functions |
| Data downloading – RDC managed sites | Download Clients* | Python |
| Metadata management – RDC managed sites | Site Metadata Manager Web Interface* | PHP, Javascript |
| Downloaded data Quality Checking, translation | TEQC* | |
| Archiving site metadata and data | Archiving management scripts* Archive database | Python MySQL |
| Cloud backup of site data and metadata | Backup management scripts* | Python |
| Streaming data management Stream broadcast | VRS? Ntrip | Windows... |
| * Provided by UNAVCO ** Commercial product | | |

Installation and Operations and Maintenance Costs

Some personnel costs to facilitate installation and all costs for ongoing operations and maintenance of the RM/RDC will be borne by the host institution. The additional complexity of components for the RDC will entail additional local expertise for installation, training time, and ongoing operation of the RDC system. Table 4 indicates personnel costs (expressed as expertise level and corresponding weeks of time required). The RM requires some local system administrator and web administrator time. The RDC requires software engineer, data engineer, and geodetic engineer time. For the RM or the RDC, a single person can provide all of the support required, if they have the appropriate background and expertise.

For the operations and maintenance phase, the assumption is that the RDC will manage 20 locally downloaded stations and 20 streamed stations.

| Table 4. Personnel Support required from RM/RDC | | |
|--|--|--|
| Level | Installation (includes training) | Operation and maintenance |
| RM | Systems administrator - 1 week Web Administrator – 1 week | System Administrator - 2 weeks/yr Web Administrator – 2 weeks/yr |
| RDC | Systems administrator - 2 weeks Web Administrator – 2 week Software Engineer – 2 week Data Engineer - 4 weeks | System Administrator - 2 weeks/yr Web Administrator – 2 weeks/yr Software Engineer – 2 weeks/yr Data Engineer – 13 weeks/yr |

Table 5 shows additional infrastructure capabilities required at each RDC level.

| Table 5. Infrastructure required at RM/RDC | |
|--|---|
| Level | |
| RM | Internet Networking and firewall Power Cooling |
| RDC | Same as for RM, but add Uninterruptible Power Supply and Conditioning |

Summary

As part of the COCONet project, support will be provided by UNAVCO to local host Regional Data Center institutions for installation and configuration of turnkey hardware systems to be provided by the COCONet Project funds, with software components pre-installed. Up to [two] Regional Mirror host institutions and one Regional Data Center host institution can be supported with hardware, installation, configuration and training of local personnel. During the first year, consultation support will be provided at the level required to ensure successful operation. Host institutions will provide infrastructure and personnel with required background and expertise.