Cooperative Institute for Research in the Atmosphere (CIRA) Requirements Review

August 11th, 2016

Disclaimer

Cooperative Institute for Research in the Atmosphere Requirements Review

Final Report

RMCMOA Workshop
Fort Collins, CO
August 11th, 2016

ESnet is funded by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research. Vince Datoria is the ESnet Program Manager.

ESnet is operated by Lawrence Berkeley National Laboratory, which is operated by the University of California for the U.S. Department of Energy under contract DE-AC02-05CH11231.

This is LBNL report XXX

Contents

Disclaimer **2**

Cooperative Institute for Research in the Atmosphere Requirements Review **3**

Final Report **3**

Contents **4**

Participants and Contributors **5**
Report Editors *5*

Executive Summary **6**

Findings **9**

Action Items **12**

ESnet SC Requirements Review Background and Structure **13**

Colorado State University Overview **15**

Cooperative Institute for Research in the Atmosphere (CIRA) Case Study **16**
Background *16*
Network and Data Architecture *16*
Collaborators *16*
Instruments and Facilities *17*
Process of Science *18*
Remote Science Activities *18*
Software Infrastructure *19*
Cloud Services *19*
Outstanding Issues *19*

Appendix – Network Diagram **21**

Participants and Contributors

Scott Baily, Colorado State University, Central IT
Joe Breen, University of Utah, Networking
Patrick Burns, Colorado State University, Vice President for Information Technology
Scot Colburn, NCAR, Networking
Dave Hoffman, Colorado State University, Central IT
Thomas Hauser, University of Colorado Boulder, Director of Research Computing
John Knaff, NOAA, CIRA
Marla Meehl, NCAR, Networking
Kevin Micke, Colorado State University, CIRA
Kate Petersen, ESnet, Networking
Greg Redder, Colorado State University, Central IT
Jason Zurawski, ESnet, Networking

Report Editors

Kate Petersen, ESnet: kate@es.net
Jason Zurawski, ESnet: zurawski@es.net

Executive Summary

The Energy Sciences Network (ESnet) is the primary provider of network connectivity for the U.S. Department of Energy (DOE) Office of Science (SC), the single largest supporter of basic research in the physical sciences in the United States. In support of the Office of Science programs, ESnet regularly updates and refreshes its understanding of the networking requirements of the instruments, facilities, scientists, and science programs that it serves. This focus has helped ESnet to be a highly successful enabler of scientific discovery for over 25 years.

In August 2016 ESnet and Colorado State University organized a review to characterize the networking requirements of the Cooperative Institute for Research in the Atmosphere (CIRA) located on the campus of Colorado State University (CSU).

CIRA is a cooperative institute that is also a research department within CSU's College of Engineering, in partnership with the Department of Atmospheric Science (ATS). Its vision is to conduct interdisciplinary research in the atmospheric sciences by entraining skills beyond the meteorological disciplines, exploiting advances in engineering and computer science, facilitating transitional activity between pure and applied research, leveraging both national and international resources and partnerships, and assisting the National Oceanic and Atmospheric Administration (NOAA), Colorado State University, the State of Colorado, and the Nation through the application of our research to areas of societal benefit.

Several key findings highlighting the results from the review are noted below:

1. While some fields such as high energy physics and astrophysics have data sets that allow for significant data reduction (e.g., filtering events in high energy physics, or eliminating the dark portions of a telescope image), climate data sets are not easy to reduce in this way. Because of this, climate data analysts need access to significant volumes of data. In general, data reduction in climate science means eliminating variables from a data set which are not needed for a particular analysis; however, that does not mean those variables are not important or valuable to other climate research and scientific discovery.

2. Some workflows involve the creation of high-value derived data products as a result of the analysis of climate and weather data from NOAA repositories. Currently, these workflows require the download of the input data as files from the NOAA repository to a local filesystem, followed by the execution of the analysis code on a local computational resources connected to the local filesystem. It would be more efficient if the analysis code could read the input data directly from the network into memory, without relying on copying the raw data to local storage first. The Large Hadron Collider (LHC) experiments

are in the process of transitioning to this remote I/O model—climatic and weather research could take advantage of this model as well.

3. The products of CIRA (models of high-interest data, such as the path of a typhoon) have a limited shelf life, but are very popular to a subset of highly distributed and remote users during a limited period of time. It is common for the CIRA products to become so popular, that a DDoS-like attack pattern may occur causing problems for their resources as well as the CSU campus network. A strategy to manage the demands on CIRA and CSU resources during busy times must be devised.

4. Storage is a problem for CIRA due to the large volumes of data and is increased as funding and time allow. The use of NAS devices results in a fragmented collection of data storage that puts heightened emphasis on each user knowing where a data set may be stored. As the data set sizes increase, it is recommended that CIRA and CSU IT partner on a scalable storage solution that offers a large and unified way to archive research data.

5. The workflow to create products within CIRA is heavily researcher/operator driven, and relies on a complex environment of code, storage, and processing hardware. Modernizing this supply chain through the use of an intelligent workflow manager, HPC/HTC techniques, and unified storage could result in increased productivity and a more stable environment that scales with data and demand.

6. The Front Range GigaPoP (FRGP) regional network provides ample capacity to CSU, and has the ability to add more as the need arises. Currently the campus is connected by 5 x 10Gbps of connectivity, with 10Gbps devoted to a special purpose Science DMZ regional network.

7. Upstream data repositories managed by NOAA have several bottlenecks, including the link between satellite and ground station, and the use of a Trusted Internet Connections (TIC) to filter all outgoing traffic. These bottlenecks limit the overall speed at which CIRA can ingest data, and maximum wall-clock time at which it can produce research products.

8. LAN bandwidth within CIRA and CSU is being augmented to facilitate some of the data forecast needs. In particular, a strategy to reduce local congestion is being devised wherein high-profile devices (e.g. the CIRA web server) can be migrated to the CSU DMZ network. This migration will mitigate LAN congestion, and better serve the distributed and remote set of users that consult CIRA products.

9. A joint NSF proposal (possibly to the Data Infrastructure Building Blocks, DIBBS, solicitation) between CSU IT and CIRA is suggested as a way to fund some of the software, workflow, and storage upgrades. With CIRA as an anchor user, CSU IT can better support other forms of research.

This report expands on these points, and covers other collaborations at Colorado State University. The report contains a Findings section and documents the entire CIRA case study discussed during the Requirements Review.

Findings

Below are the findings of the ESnet and Colorado State University Requirements review focusing on CIRA. This review was held August 11th 2016 in Fort Collins CO. These points summarize important information gathered during the review.

- Data consumed by CIRA will increase in the coming years due to radical increases in the capabilities of the satellites (GOES-R by 2017, and GOES-S and GOES-T in the future) and climate models. The center, and CSU IT, should discuss ways to handle this from both a storage and network perspective.
- Climate/Weather data do not 'trigger' (e.g. throw away) pieces of a data set that are known to be of lesser value; this results in larger raw data sets to process and store. It has been established that these data sets retain their research value over time, and are kept in local and distributed storage. This increased data load points to needing a solution for short, medium, and long term storage at CSU and CIRA that will scale over the coming years.
- Migrating the CIRA web server to a location in the CSU DMZ would result in better local traffic management. This benefit will be most realized during periods of extreme demand (e.g. when a CIRA product is valuable to a particular remote community).
- Moving the CIRA web server from a single machine to a more durable setup (e.g. multiple real or virtual hosts, use of a high capacity web server like Apache) will result in better performance and ability to handle demand during busy times. This hardware/software change should occur at or around the time of the migration to the CSU DMZ.
- Storage for research data will become a serious issue in the coming years as the resolution and quantity of weather/climate information increases. CSU IT and CIRA should devise a plan for more durable/plentiful storage to be used by the center and/or campus.
- The current CIRA workflow is not very sophisticated, but highly functional. Modifying the tools used (e.g. workflow manager software, the use of HPC/HTC resources) has the potential to automate the science and increase productivity significantly.
- Use of workflow manager software (e.g. Pegasus, etc.) could facilitate an immediate productivity increase for CIRA, and allow for greater portability of the process of creating models.
- It is suggested that CSU IT and CIRA pursue a future NSF grant (e.g. in the DIBBS program) to address the need for storage, software, and processing assistance in this science workflow.

- CIRA products become very popular during specific and predictable time periods. E.g. if CIRA is modeling a tropical cyclone, it is predicted that there will be an increased demand from people within the region about to experience the tropical cyclone. This results in repeated data flows from CIRA to potentially remote locations. CIRA can work out a strategy with CSU IT to alert them when a 'busy time' is coming, to better allow CSU IT to prepare network, storage, and processing resources to support the demand.
- Upstream data repositories that CIRA uses that are provided by NOAA are a significant bottleneck. The first item (that cannot be solved) is the time required to download data from the satellite to the repository. The second item (that may be addressed) is the use of legacy data transfer tools (FTP) and the construction of a TIC bottleneck. NOAA and CIRA may be able to work together to find alternative solutions – in particular the use of GridFTP/Globus should be considered.
- The Foothills campus infrastructure is constrained by an antiquated power delivery system. CIRA/ATS typically experience 5-6 power outages a year on average. These power outages typically last anywhere from 30-120 minutes and severely disrupts the data processing, hardware and IT personnel time and effort. Thousands of dollars are wasted in hardware and IT personnel time recovering from these types of outages. CSU IT is in a unique position to be able to facilitate the hosting of key resources (e.g. the CIRA web server) to prevent service interruption.
- The GOES program will produce satellites that stand to increase the order of magnitude of data on a regular basis, with the next major increase coming in late 2016. This implies a greater need for storage and processing capabilities.
- CIRA IT has explored cloud technologies and determined it is not a feasible solution for most of the data processing algorithms being developed at CIRA. Cloud solutions could provide a benefit for imagery being processed for the RAMMB web server due to the high volume of traffic the website receives. Kevin Micke has been exploring cloud technologies through the Open Commons Consortium (OCC)/ Open Science Data Cloud (OSDC) platform, but the immense cost of data ingress and egress make any of the private sector solutions cost prohibitive. CloudLab, a product of the University of Utah, may be a solution to consider in future years.
- CIRA has eight full time dedicated IT personnel (one Infrastructure Manager/hardware specialist, four Linux system administrators, two Windows system administrators and one web developer). Some of these personnel also perform research software engineering along with several other research personnel at CIRA.
- CSU IT staff could learn a lot more about the process of science by spending a day observing CIRA operations.

- After a recent upgrade from CSU IT, WAN bandwidth is sufficient for current CIRA needs. MAN/LAN bandwidth has been a problem in the past. Taking care of the congested web server at the foothills campus will be a good first step to alleviate congestion.
- CSU IT has an incredible opportunity to take the lessons of this requirements review to work with their research community. The information devised in this review has shown there are several easy wins available, and ways to build trust and confidence with researchers to advance the mission of campus science. Even minor improvements will result in a higher level of competitiveness, and the ability to sell the campus network as more than capable of supporting science, when pursuing funding opportunities.

Action Items

ESnet and Colorado State University recorded a set of action items from the August 11th 2016 Requirements Review, continuing the ongoing support and collaboration.

- CSU IT will work with CIRA to offer a solution for the congested web server. This may include:
  - Adoption of new hardware (either real or virtual)
  - Use of a new web server environment (Apache)
  - Exploring load balancing
  - Migration of resource to edge of campus in the CSU DMZ.
  - Establishing two paths – internal data path (e.g. ability of CIRA staff to upload new products) and external data path (e.g. for outside users to download products via HTTP)
- CSU IT will discuss storage solutions for CIRA and beyond. NAS devices purchased on an individual basis will not scale. A campus-wide storage solution (paid for via institutional overhead during the grant process) will allow all research groups to scale, and know that the storage is being managed in a professional manner.
- CSU IT and CIRA should explore a joint proposal to the NSF DIBBS program. The goal would be to show a science use case, and propose a new way to handle the workflow (new processing, storage, and software development).
- CSU IT will discuss ways to take the lessons from this review, and apply them to other research groups on campus. It is suggested that a person in the engineering team be given a time allocation to perform the outreach and interview process on a pilot basis. If successful, the position could be funded by the NSF directly (CI Engineer) or as a back-charge via overhead on NSF grants the school is awarded in other departments.

ESnet SC Requirements Review Background and Structure

Funded by the Office of Advanced Scientific Computing Research (ASCR) Facilities Division, ESnet's mission is to operate and maintain a network dedicated to accelerating science discovery. ESnet's mission covers three areas:

1. Working with the DOE SC-funded science community to identify the networking implications of instruments and supercomputers and the evolving process of how science is done.
2. Developing an approach to building a network environment to enable the distributed aspects of SC science and to continuously reassess and update the approach as new requirements become clear.
3. Continuing to anticipate future network capabilities to meet new science requirements with an active program of R&D and advanced development.

Addressing point (1), the requirements of the SC science programs are determined by:

a) A review of major stakeholders' plans and processes, including the data characteristics of scientific instruments and facilities, in order to investigate what data will be generated by instruments and supercomputers coming online over the next 5–10 years. In addition, the future process of science must be examined: How and where will the new data be analyzed and used? How will the process of doing science change over the next 5–10 years?
b) Observing current and historical network traffic patterns to determine how trends in network patterns predict future network needs.

The primary mechanism to accomplish (a) is through the SC Network Requirements Reviews, which are organized by ASCR in collaboration with the SC Program Offices. SC conducts two requirements reviews per year, in a cycle that assesses requirements for each of the six program offices every three years. The review reports are published at http://www.es.net/requirements/.

The other role of requirements reviews is to help ensure that ESnet and ASCR have a common understanding of the issues that face ESnet and the solutions that it undertakes.

In August 2016, ESnet organized a review in collaboration with Colorado State University to characterize the networking requirements for the facilities and science programs of Cooperative Institute for Research in the Atmosphere (CIRA).

The CIRA representatives were asked to communicate and document their requirements in a case-study format that included a network-centric narrative describing the science, instruments, and facilities currently used or anticipated for future programs; the network services needed; and how the network is used.

Participants considered three timescales on the topics enumerated below: the near-term (immediately and up to two years in the future); the medium-term (two to five years in the future); and the long-term (greater than five years in the future).

More specifically, the structure of a case study was as follows:

- ***Background***—an overview description of the site, facility, or collaboration described in the case study.
- ***Collaborators***—a list or description of key collaborators for the science or facility described in the case study (the list need not be exhaustive).
- ***Network and Data Architecture***—description of the network and/or data architecture for the science or facility. This is meant to understand how data moves in and out of the facility or laboratory focusing on local infrastructure configuration, bandwidth speed(s), hardware, etc.
- ***Instruments and Facilities***—a description of the network, compute, instruments, and storage resources used for the science collaboration/program/project, or a description of the resources made available to the facility users, or resources that users deploy at the facility.
- ***Process of Science***—a description of the way the instruments and facilities are used for knowledge discovery. Examples might include workflows, data analysis, data reduction, integration of experimental data with simulation data, etc.
- ***Remote Science Activities***—a description of any remote instruments or collaborations, and how this work does or may have an impact on your network traffic.
- ***Software Infrastructure***—a discussion focused on the software used in daily activities of the scientific process including tools that are used to locally or remotely to manage data resources, facilitate the transfer of data sets from or to remote collaborators, or process the raw results into final and intermediate formats.
- ***Cloud Services***—discussion around how cloud services may be used for data analysis, data storage, computing, or other purposes. The case studies included an open-ended section asking for any unresolved issues, comments or concerns to catch all remaining requirements that may be addressed by ESnet.

Colorado State University Overview

Colorado State University is a public research university located in Fort Collins, Colorado. The university is the state's land grant university, and the flagship university of the Colorado State University System.

The current enrollment is approximately 33,000 students, including resident and non-residents. The university consists of approximately 1,800 faculty in eight colleges and 53 academic departments. Bachelor's degrees are offered in 72 fields of study, with master's degrees in 55 fields. Colorado State confers doctoral degrees in 45 fields of study, in addition to a professional degree in veterinary medicine.

In fiscal year 2014, CSU spent $308.0 million on research and development, ranking 75th[*] in the nation overall and 42nd[†] when excluding medical school spending.

The Cooperative Institute for Research in the Atmosphere (CIRA) and the Department of Atmospheric Science (ATS) bring in almost $35 million annually to the CSU research budget—almost 10% of total CSU research budget.

---

[*] https://ncsesdata.nsf.gov/profiles/site?method=rankingBySource&ds=herd

[†] https://ncsesdata.nsf.gov/herd/2014/html/HERD2014_DST_32.html

Cooperative Institute for Research in the Atmosphere (CIRA) Case Study

Background

The Cooperative Institute for Research in the Atmosphere (CIRA) is an institute that is also a research department within CSU's College of Engineering, in partnership with the Department of Atmospheric Science. The center's vision is to conduct interdisciplinary research in the atmospheric sciences by entraining skills beyond the meteorological disciplines, exploiting advances in engineering and computer science, facilitating transitional activity between pure and applied research, leveraging both national and international resources and partnerships, and assisting NOAA, Colorado State University, the State of Colorado, and the Nation through the application of our research to areas of societal benefit. The CIRA facility is located at the CSU foothills campus.

Network and Data Architecture

To accommodate the research demand for access to many next-generation large volume data sets in use at CIRA, private networks are used to isolate traffic to processing clusters. Included in each cluster are 10Gbps switches and dual network adapters on all servers and associated NAS storage devices. This isolates the heavy volume data traffic from the overall LAN.

Operational control of the entire CIRA network infrastructure is being transitioned to Central IT (Academic Computing and Networking Services, or ACNS, referred elsewhere in this document as "CSU IT") as part of an IT centralization initiative. One architectural change resulting from this effort will be the addition of a redundant 10Gbps circuit to the main campus and 10Gbps connectivity to all CIRA computer rooms. Desktop workstations will still run on 1Gbps network links, so the compute infrastructure of the near future will transition from powerful Xeon workstations at researchers' desks to a client/server environment using the high-speed private networks.

Collaborators

Two federal groups are collocated in CIRA: the Regional and Mesoscale Meteorology Branch (RAMMB) and a National Parks Service (NPS) group that studies air quality research and monitoring. These groups are augmented with CIRA researchers doing joint research.

The Regional and Mesoscale Meteorology Branch (RAMMB) of NOAA/NESDIS conducts research on the use of satellite data to improve analysis, forecasts and warnings for regional and mesoscale meteorological events. It also collaborates with CSU's Atmospheric Science department, the National Weather Service, and the US Department of Defense. Virtual organizations include NOAA National Environmental Satellite, Data, and Information Service (NESDIS) – including our sister cooperative institutes, NOAA National Centers for Environmental Prediction

(NCEP) and the DoD. Remote data available from these sources complement the ground station activities and allow for the development and real-time testing of applications and applied research, archives of weather events, data assimilation, and development of training material that is key to CIRA and RAMMB's missions.
VO key endpoints include Suitland, Maryland, Silver Spring Maryland, College Park Maryland, Miami, Florida, Honolulu, Hawaii, Madison, Wisconsin, and Gulf Port, Mississippi.

National Park Service Air Quality Group and CIRA associated does a variety of research on air quality, including:

- WRAP Technical Support System - Providing technical support to the Western Regional Air Partnership (WRAP) visibility improvement program.
- IMPROVE - Contributing to the monitoring of air quality in protected environments through the Interagency Monitoring of PROtected Visual Environments (IMPROVE).
- VIEWS - Supporting the Environmental Protection Agency through the Visibility Information Exchange Web System (VIEWS) for regional haze research.
- AirToxics - Online access to data about toxic substances in the air.

This case study will be from the RAMMB prospective only.

Instruments and Facilities

***Present*:**
CIRA has ground stations that access the Geostationary Operational Environmental Satellites (GOES) with several antennas pointing toward the existing satellites (GOES-13, GOES-14 and GOES-15). It also is the home to the CloudSat Data Processing Center. Currently all GOES data is saved, though this will become more challenging as the volumes increase significantly in late 2016. There is a dedicated internet (T1) from CIRA to CSU at the present time. Several (order 30) NAS devices serve as storage, and Blu-ray discs provide permanent backup. Most researchers have their own workstations (order 50 on site/20 RAMMB), there are multiple high-powered servers (order 20) generally used by multiple researchers for data processing, and several CIRA clusters available for high end computing.

***Next 2-5 years*:**
The next generation of GOES satellites will increase data storage and network needs by 10-20 times present. Additional polar orbiting data will also be needed for ongoing research projects. The number of people and machines is expected to grow 20%, storage by 2000%. This is still being explored and some data will not be saved.

***Beyond 5-years*:**
Global satellite and model data is expected to grow another 200%

Process of Science

Present: Satellite data are the key to the scientific research done at CIRA, but there are several ancillary data required to do accomplish CIRA's interdisciplinary research. In addition to the satellite data received via the ground station, low earth orbiting satellite data are also collected from remote sources. Examples include the global constellation of geostationary data (3-4 satellites), NPOES Preparatory Project (NPP), CloudSat, Global Precipitation Mission (GPM;NASA) and others tied to specific funding opportunities. To complement those data, environmental data is also collected including things like weather data, global model analyses and forecasts, operational hurricane models, real-time hurricane information and hurricane aircraft data. Some modeling and simulation activities are done at CIRA to support applied research and data assimilation activities.

Of these data the NPP, GPM, and the global constellation of geostationary satellites consume most of the network usage (order 200-500 Gb per day).

Satellite applications, data archive, real-time dissemination of data and applications are our primary output and interface with users (public, governmental and private). Examples include real-time GOES displays and loop, real-time hurricane information, experimental proving grounds/testbed activities. Much of the sharing is done through RAMMB's web server at http://rammb.cira.colostate.edu/ , but there are a number of products that are pushed to operational centers for their real-time assessment.

Next 2-5 years:
These activities will continue as the mission remains, but the volume of data (model and satellite) is expected to increase 2000% as new geostationary and low-earth satellites and model resolutions increase as does the overall use of said data.

There is also a larger emphasis toward data assimilation and modeling that will likely take place on at high performance centers.

Beyond 5 years:
More of the same, our infrastructure and tasks will increase, though the data volumes will likely level off until the next generation of satellites/data/models becomes available.

Remote Science Activities

Testbed and proving ground activities will continue with the National Weather Service, but as the next generation of satellites becomes available, we will be required to acquire data and applications that are produced in Washington DC. In addition NPP and its follow on satellites (JPSS-1 &-2) will continue to be collected from remote locations. The methods currently being explored are mostly internet driven solutions where CIRA ingests subsets of the data from NOAA HPC systems.

Collaborations with the National Weather Service, NASA (i.e. GPM), and the DoD will continue. This will require acquiring data from systems and facilities that are still being designed to distribute these volumes of data.

Working with data on US government facilities is becoming harder as security is a moving target and the political environment is requiring more back ground verification to access those systems. Cyber security and data restrictions are limiting CIRA's ability to download data from some US government sites.

Software Infrastructure

***Present*:**

Much of the satellite data currently comes in three formats HDF5, NetCDF, and McIDAS. However the ancillary data needed to design applications comes in a variety of text and World Meteorological Organization formats (GRIB2, BUFR, HDOBS, etc.). The data are collected using a combination of McIDAS, Python, and shell scripts using various protocols (ftp, http, https, ftps, sftp, etc.)

Software consists of McIDAS, Python, FORTRAN, C, C++, shell scripts, Perl, etc. Display capabilities come with McIDAS and NWS operational platforms (AWIPS-II).

Processing will consist of legacy software (FORTRAN, Python, etc.) formatted to our customer's needs (e.g., for AWIPS-2, GRIB, NetCDF, etc.)

***Next 2-5 years*:**

McIDAS is likely going to be phased out in this time and other methods to display and work with the data will likely be adopted although there is no clear path at this time.

Data will be accessed from remote centers using data management systems that are still being developed. It is still to be determined the protocols that will be used. Data formats for different types of products are still being defined.

***Beyond 5 years*:**

More of the same, unless there is an effort to upgrade/adapt. There is hope that design, formats, and infrastructure will be decided in the 2-5 year period.

Cloud Services

This is an open question as physical storage will likely be too costly, and the alternatives are not collecting/saving as much data on permanent storage or moving to cloud systems. CIRA's primary issue with cloud systems is the immense cost of data ingress and egress for data volumes of these sizes.

Outstanding Issues

***Facts*:**

- Data volumes are going to increase 2000%

***Challenges*:**

- Networks may become saturated (LAN, and MAN mostly as dealing with data becomes more complex). WAN growth has kept pace thus far.
- NOAA's systems are not yet well defined for getting the raw data to CSU, and are expected to change due to more complex security protocols and restricted access to data.
- Security at the NOAA level may makes data sharing near impossible.
- The climate and weather data formats are still, and always seem to be, in flux.
- Computing power and storage capabilities are starting to level off. Costs are increasing, resulting in only small upgrades in capabilities.
- Uncertain funding environment - CIRA is largely supported by federal funding (NOAA, NASA grants)

Appendix – Network Diagram