CIRA

VOLUME 32, FALL 2009

A New Chapter Envisioned

Colorado State University

CONTENTS

CIRA

Cooperative Institute for Research in the Atmosphere

www.cira.colostate.edu

On the Cover: Architect, Art Hoy's vision for the upcoming CIRA building addition.

NOAA

CIRA's Re-competition Experience with NOAA

Mary McInnis-Efaw and Steven Miller

CIRA's founding as an institute can be traced to the cooperative agreement established with the National Oceanic and Atmospheric Administration (NOAA) in 1980. CIRA was among the very first Cooperative Institutes (CIs) to be launched in the nation with the purpose of engaging university faculty at the cutting edge of environmental research with the scientific staff at NOAA. The collaboration has been fruitful and productive over the last nearly 30 years. In fact, today CIRA is a thriving interdisciplinary research institute comprised of multi-agency cooperative agreements, expenditures north of $15M per year and nearly 150 employees located around the country. The success of CIRA's research endeavors is due in no small measure to the track record cultivated through our association with NOAA, which remains a vital core of our operations both in terms of dollars and breadth of collaborations in which our scientists engage.

Several years ago, the research arm of NOAA responded to a Congressional mandate to implement an open re-competition process for its CIs, whereby any institution of higher learning could propose to host a CI responding to NOAA's articulated needs. Given the limited resources within NOAA to administer the CI program, this re-competition necessarily applied to all incumbent members, including CIRA. While confident in the strength of our program, CIRA's leadership nevertheless began a process early on to strategize and position ourselves to emerge the clear choice in NOAA's call for a new "Cooperative Institute to Investigate Satellite Applications for Regional/Global-Scale Forecasts." It's important to note that prior to this re-competition process, the CIs had to submit to periodic, rigorous scientific reviews. The outcome of that review would then serve as the means by which the cooperative agreement would be renewed for an additional five year term. This open re-competition process was entirely new to all of the NOAA CIs and harbored considerable unknowns to the staff at CIRA and to the University.

CIRA, Fort Collins group.

With the end date of CIRA's cooperative agreement in sight, a new team in place under the leadership of Director Graeme Stephens and Deputy Director Steven Miller and with the assistance of the outgoing leadership team including Tom Vonder Haar and Ken Eis, CIRA began to assess NOAA's needs and direction on the 5 and 10 year horizon. This "needs review" combined with a thorough assessment of the University's capabilities and resources, and a recapping of CIRA's venerable track record all were synthesized and became a plan of attack. CIRA could anticipate the RFP (Request for Proposals) and begin writing to the strengths and capabilities of its team and ultimately communicate to NOAA how renewal would be the logical best outcome for both itself and for the University.

Included in this strategy were early negotiations with the University on significant cost-sharing measures that would demonstrate our commitment to the NOAA CI program. These

NOAA

CIRA, Boulder group.

measures included 1) funding for a new Atmospheric Science faculty line connected to CIRA research themes, 2) significant reductions in the on and off-site overhead rates assessed to NOAA projects conducted within the cooperative agreement, 3) and a plan to finance an approximately 4000 sq ft addition to the building.

The RFP for the re-competition of CIRA was due to be released in the summer of 2008, however the actual date it appeared was early October. All of the groundwork in the preceding months paid off when it became clear that the CIRA team had done an accurate job in anticipating the RFP content. Given the short turn-around time of 90 days for submission and the desire to submit the complete proposal before much of the University infrastructure shut down for the holidays, a busy eight weeks ensued to wrap up the final version. CIRA submitted its proposal a week before Christmas and officially learned the good news of its selection in May of 2009. Below is an excerpt from the official NOAA press release:

"NOAA announced today its renewed affiliation with the Cooperative Institute for Research in the Atmosphere (CIRA) at Colorado State University in Fort Collins, Colo. Chosen through a competitive process, the cooperative institute will continue to investigate satellite applications for improving regional and global-scale weather forecasts, water resource forecasts, and provide integrated weather information to meet future aviation and surface transportation needs.

"Our partnership with CIRA provides professors and students of CSU an exciting and challenging opportunity to collaborate with NOAA scientists on cutting-edge research," said Richard Spinrad, NOAA assistant administrator for oceanic and atmospheric research. "Improving the accuracy of weather forecast warnings and looking at short-term climate forecasts are important efforts toward NOAA's overall mission to monitor and enhance weather and water information, improve decision making, and promote environmental stewardship."

We look forward to continuing and expanding upon our successful program with NOAA over the course of the five-year award. A noncompetitive scientific review will take place toward the end of this term which will provide CIRA the opportunity to secure an additional five-year extension of the cooperative agreement upon the successful outcome. At the time of this writing, the CIRA team is working closely with a talented architect from Denver on the design for our new building addition. The concept drawings have energized the staff and have us all thinking about the promise of our future. With new leadership in place, new/clarified research directions to pursue, a growing facility, strong cooperation with our partner departments, and the ongoing support of the University, NOAA and our other major sponsors in the Department of Defense and National Park Service, CIRA is well poised to deliver on its new research vision.

GOES-R GLM

Is Lightning a Predictor of Hurricane Intensity Change?

Mark DeMaria, John A. Knaff (NOAA/NESDIS) and Robert T. DeMaria (CIRA)

The next generation Geostationary Operational Environmental Satellite (GOES) systems starting with GOES-R (scheduled for late 2015) will include a geostationary lightning mapper (GLM). The GLM will provide nearly continuous locations of total lightning (cloud to ground and intra-cloud) with an accuracy of about 10 km over most of the field of view of GOES-east and -west. This coverage will include the vast majority of the regions where tropical cyclones occur in the Atlantic and northern East Pacific. This lightning data will provide new information about the convective structure in tropical cyclones and their environments. Several studies with ground-based detection systems suggest that increased lightning activity near the storm center is a precursor to storm intensification (e.g., Squires and Businger 2008; Price et al. 2009). However, these and other studies were limited to cases studies of individual storms, small samples within the range of land-based lighting networks, or only those cases that intensified into strong hurricanes. In preparation for GOES-R, a very large sample of tropical cyclones is being examined that includes all phases of the storm life-cycle and storms that did not intensify very much.

Lightning production is a complicated process that is not completely understood. Generally speaking, lightning occurs in convective clouds. The electric charge is generated by the simultaneous presence of various types of cloud ice and super-cooled water droplets (drops colder than 0°C (the usual freezing point of water), but which have not yet frozen due to the small droplet size). The generation of charge also depends on the cloud updraft speed. Because tropical cyclones usually have weaker updrafts and fewer regions with the simultaneous occurrence of super-cooled water and ice than the typical thunderstorms that occur over land, they tend to have less lightning. However, because of the dependence on updrafts and cloud properties, lightning still has the poten-
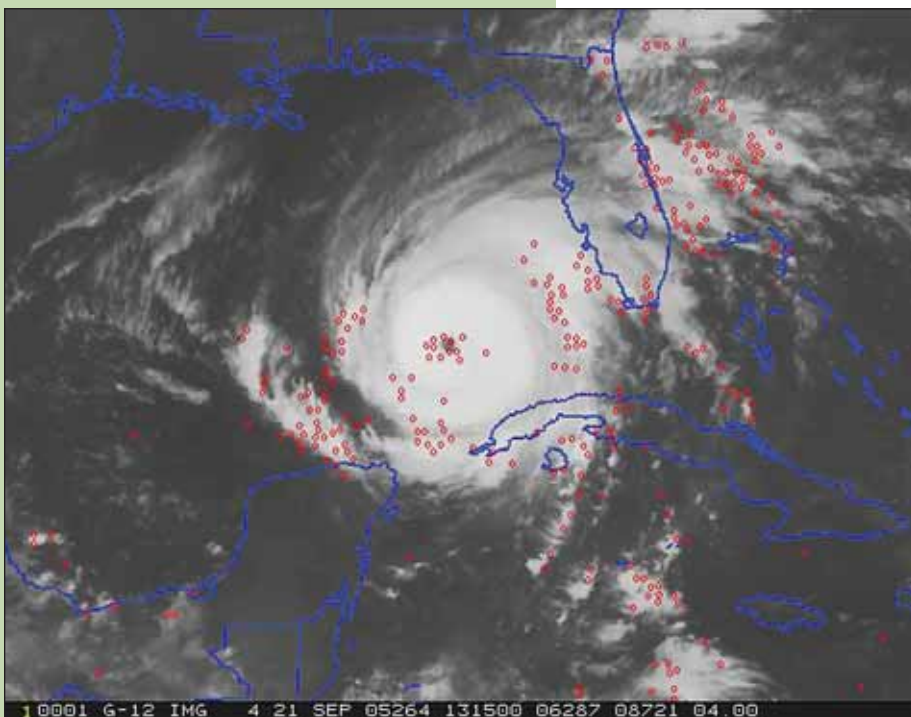
Figure 1. GOES Infrared image of Hurricane Rita at 1315 UTC on 21 September 2005. The locations of the lighting strikes over the 6 hr period centered at 1315 UTC detected by the World Wide Lightning Location Network (WWLLN) are indicated by the red circles.

tial to provide valuable information about tropical cyclone structure and evolution.

There are a number of currently available lightning observing systems, including space- and ground-based instruments. However, none of these systems has the capability to monitor the total lightning over the large expanses of the tropical oceans with the time resolution needed for tropical cyclone analysis. The Lightning Imaging Sensor (LIS; 1997-present) and Optical Transient Detector (OTD; 1995-2000) are on low-earth orbiting satellites and use instruments that detect optical pulses from lightning (Boccippio et al. 2002). Although these instruments detect total lightning, they only provide an occasional snapshot and so are not well-suited to study the time evolution of lightning in individual tropical cyclones. There are some specialized regional ground-based lightning mapping arrays (LMAs) such as the one near Washington, D.C. (http://
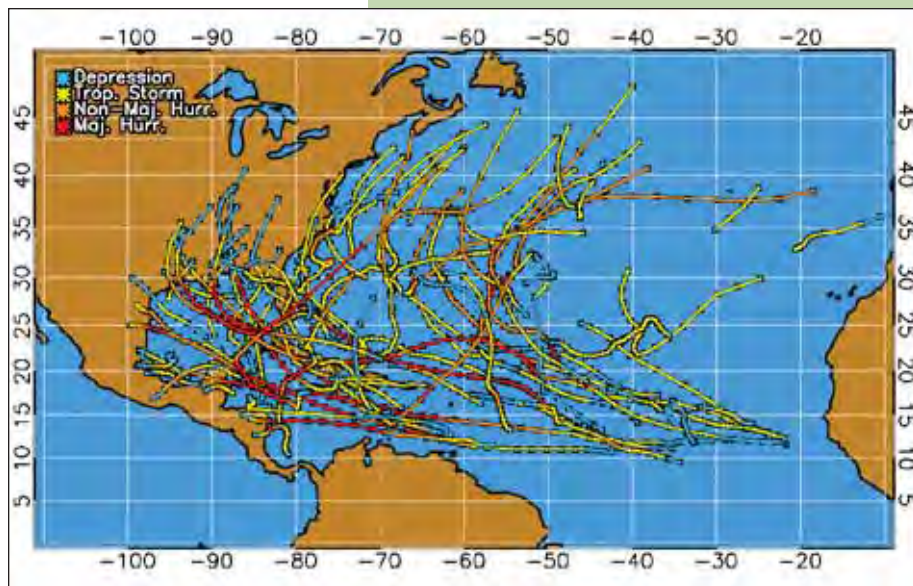
Figure 2. The tracks of all Atlantic tropical cyclones from 2004-2008. The storm intensity classification is indicated by the color of the track.

branch.nsstc.nasa.gov/PUBLIC/DCLMA) that can detect total lightning with high time and spatial resolution, but these do not provide much coverage over the ocean. Other ground based networks such as the National Lightning Detection Network (NLDN) and World Wide Lightning Location Network (WWLLN) utilize very low-frequency electromagnetic radiation to identify the timing and location of cloud to ground lightning with a fairly wide coverage. The NLDN and WWLLN data have been utilized in the majority of previous studies of tropical cyclone lightning. The WWLLN data provides coverage over the global tropics, so it will be used in this study. The range of the NLDN data was also recently expanded, but a long-term climatology is not yet available.

The properties of the WWLLN data are described by Rodger et al. (2006) and at http://webflash.ess.washington.edu. Basically, this data provides the timing and location of the cloud to ground lightning over the entire globe. Although the detection rate is fairly low, it does not have much spatial variability. The WWLLN data was recently reprocessed with a new algorithm that improved the reliability. In this study, the reprocessed data for 2004-2008 is used. To account

for the low detection rate and as a crude adjustment to account for the fact that only cloud to ground strikes are detected, a correction factor is applied to the lightning density (described below) for each year to make the total strikes over the Atlantic tropics and subtropics equal to that from the well-calibrated OTD/LIS lightning climatology over this same region. Figure 1 shows the WWLLN strike locations over a six hour period for Hurricane Rita centered at 12 UTC on 21 September 2005.

The WWLLN data were obtained for the full life cycles of all Atlantic tropical cyclones from 2004-2008. Figure 2 shows the tracks of all the storms included in the sample. Since the primary interest is the storm intensity change before storms make land-fall, only those cases where the storm remained over water for at least the next 24 hours are included. The lightning strike data for each storm are composited over 6 hr periods. The dataset contains 1057 6 hr periods from 66 tropical cyclones. As shown in Figure 2, storms ranging in intensity from tropical depressions (maximum sustained surface winds less than 34 kt) to major hurricanes (maximum sustained surface winds 100 kt or greater) are included.

Since the structure of tropical cyclones depends strongly on the distance from the storm center, the lightning strike data over each 6 hr period was analyzed in a storm-centered cylindrical coordinate system. The strikes in grid cells in 100 km radial intervals out to 1000 km and azimuthal intervals of 45° were determined for each case. Because the area of the grid cells increase with distance from the center, the lightning density (strikes per unit area and time) was calculated. The correction factor to account for the low detection rate and conversion from cloud to ground to total lightning described above was applied to the lightning density rather than the strike counts. To be consistent with the OTD/LIS climatology, lightning density was calculated in units of strikes/km^2-year.

Figure 3 shows an example of the time evolution lightning density in the inner 100 km and the maximum wind for Hurricane Rita from 2005

starting from the tropical depression stage to 24 hours before landfall. This figure shows that the lightning activity was very transient, with a large burst of activity starting late on September 19th. After the lightning burst, Rita rapidly intensified to a category 5 hurricane (maximum winds of at least 140 kt) in the central Gulf of Mexico. This result is consistent with the analysis by Squires and Businger (2008). A more general study of the lightning activity for all tropical cyclones around the globe that reached at least category 4 intensity (maximum sustained surface winds of 115 or greater) over a three year period was performed by Price et al (2008). They showed that a lightning burst was associated with the intensification of nearly all of these storms, although the time lag between the lightning and intensification was highly variable. In some cases the lightning peak occurred after the maximum wind increased. A similar behavior can be seen for Hurricane Rita in Figure 3. Although the largest lightning burst occurs before the intensification, there is another burst of lightning starting on September 22nd as the storm began to weaken. During this period, Rita began to interact with a region of higher vertical wind shear (where the winds surrounding the storm vary with height, which tends to limit storm intensification). These results suggest that while bursts of lightning might be a precursor to intensification, the relationship is more complicated than indicated by previous studies.
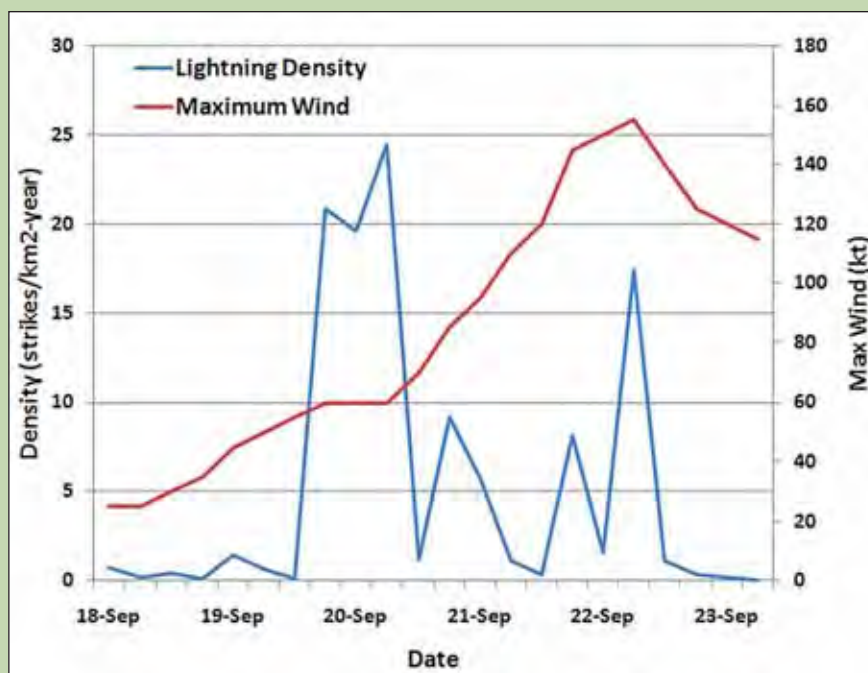
Figure 3. The time evolution of the 0-100 km lightning density and maximum one-minute sustained surface wind for Hurricane Rita (2005) from the time of its initial formation east of the Bahamas on 18 September to 24 hr before its landfall near the Texas-Louisiana border.

To further investigate whether or not there is a relationship between lightning and tropical cyclone intensification, a correlation analysis was performed. The lightning density over each 6 hr period was correlated with the change in maximum wind in the following 48 hours. This correlation was performed for each 100 km radial interval out to 600 km. Figure 4 shows the correlation coefficients as a function of radius. The correlation for the total sample (blue curve) was slightly negative for the inner radii and was not statistically significant at the 95% level. This result indicates that there is no significant relationship between inner core lightning and subsequent intensity changes. The correlation coefficient was greater for the larger radii, and the relationship was statistically significant for the radial interval centered at 250 km.

To better understand the statistical relationships shown in Figure 4, the environmental vertical wind shear for each case was also calculated. It was found that the lightning activity generally increased as the vertical shear increased. Because vertical shear tends to have a negative influence on tropical cyclone intensity change, this relation-

GOES-R GLM

Figure 4. The coefficient of correlation between lightning density composited over 6 hr intervals and the change in maximum surface winds in the following 48 hr for the total 2004-2008 sample and for the sample with low vertical shear. The correlations were performed for each 100 km radial interval out to 600 km. The x-axis is the center of each radial interval.
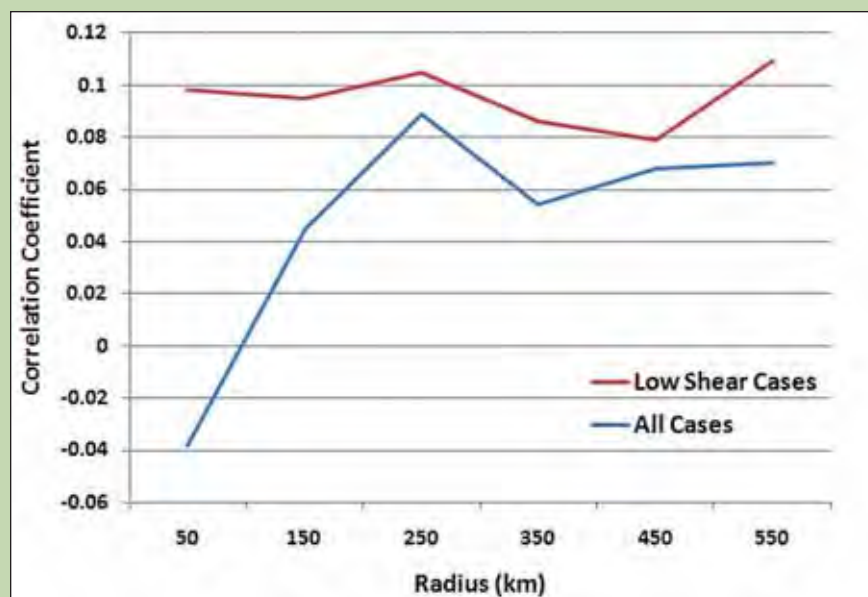
ship reduces the correlation between lightning and intensity change. To help separate these two effects, the total sample was stratified into low and high shear cases. The vertical shear in the storm environment was determined from global forecast model analyses and is magnitude of the shear vector between 200 and 850 hPa averaged over an area within 500 km of the storm center. Figure 4 shows the correlation between lightning density and the subsequent 48 hr maximum wind change for the cases where the vertical shear was low. The correlation coefficients are greater than those for the total sample, and the relationships were statistically significant for the first four radial intervals.

These results indicate that there is a significant relationship between lightning activity and tropical cyclone intensity change that may prove to be useful for improved intensity fore-casts. However, it is important to note that this relationship is more complicated than indicated by previous studies. Increasing vertical shear tends to increase lightning but reduce tropical cyclone intensity, which partially cancels the positive correlation between lightning and intensity changes, especially in the storm inner core. This research is continuing to further the understanding of lightning activity in tropical cyclones and explore potential forecast applications of ground- and GLM-based information to intensity prediction.

References

Boccippio, D.J., W.J. Koshak, and R.J. Blakeslee, 2002: Performance assessment of the Optical Transient Detector and Lightning Imaging Sensor. Part I: Predicted Diurnal Variability. *J. Atmos. Oceanic Tech.*, **19**, 1318-1332.

Price, C., M. Asfur, and Y. Yair, 2009: Maximum hurricane intensity preceded by increase in lightning frequency. *Nature Geoscience*, **2**, 329-332.

Rodger, C.J., S. Werner, J.B. Brundell, E.H. Lay, N.R. Thomson, R.H. Holtzworth, and R.L. Dowden, 2006: Detection efficiency of the VLF World-Wide Lightning Location Network (WWLLN): initial case study. *Ann. Geophys.*, **24**, 3197-3214.

Squires, K., and S. Businger, 2008: The morphology of eyewall lightning outbreaks in two category 5 hurricanes. *Mon. Wea. Rev.*, **136**, 1706-1726.

VIEWS

The Visibility Information Exchange Web System (VIEWS)

http://views.cira.colostate.edu
Shawn McClure

The Visibility Information Exchange Web System (VIEWS) is a website and database system designed to provide easy access to a wide variety of air quality data through an integrated suite of visualization and analysis tools. VIEWS began as a collaborative effort funded by EPA between the five Regional Planning Organizations (RPOs) and the National Park Service Visibility Program at CIRA to help states, tribes, federal land managers, researchers, and planners evaluate air quality and visibility in federally-protected ecosystems according to the stringent requirements of the EPA's Regional Haze Rule and the National Ambient Air Quality Standards. As a result of its early success and growing inventory of data and tools, VIEWS began to attract a wider audience in the years following its debut, and its original focus on regional haze has since been expanded to include climate change, health effects, emissions control strategies, and general environmental impacts.

VIEWS provides access to over five dozen different air quality datasets, including data from the EPA's Air Quality System (AQS), the Clean Air Status and Trends Network (CASTNet), the National Atmospheric Deposition Program (NADP), AIRNow, and the National Emissions Inventory. VIEWS is also the primary dissemination point for data from the Interagency Monitoring of Protected Visual Environments (IMPROVE), a national monitoring initiative designated by the EPA as the official source of data for use in calculating visibility impairment from fine particulate matter in federal Class I Areas in accordance with the Regional Haze Rule and the Clean Air Act. The VIEWS team is also starting to incorporate results from a variety of air quality and meteorological models, as well as data and images from several NASA satellites. A primary goal of VIEWS is to bring together ground-based observations, modeling results, emissions inventories, and satellite data into a single, integrated system so that they can be compared and analyzed side-by-side in order to give the user a more comprehensive understanding and a "one atmosphere" view of air quality.

VIEWS employs an advanced data acquisition and import system to consolidate air quality data from a variety of sources into a single, highly-optimized data warehouse. Ground-based measurements from dozens of monitoring networks, air quality modeling results, and detailed emissions inventories are imported and updated on a regular basis using a systematic data model and carefully standardized metadata. The names, codes, units, and flags used in the source datasets are carefully mapped to a unified protocol, and

Figure 1. The VIEWS home page provides a guided tour of features, information on recent data updates, local weather news, maps of visibility and weather webcam locations, and local air quality news.

VIEWS

Figure 2. The VIEWS site browser uses Google Maps to allow users to select and view metadata details about data collection locations.

native formats and organizations are transformed into a common, normalized database schema. This design enables users to explore, merge, and analyze datasets of widely-varying origin in a consistent manner using a single set of tools and web services. This degree of integration allows decision-makers to analyze diverse datasets side-by-side and focus on high-level planning strategies without having to contend with the details of data management and manipulation.

At present, VIEWS has over 1500 registered users from over 100 different countries and 500 different organizations, including scientists, students, researchers, planners, and stakeholders. In

2008, the VIEWS website averaged around 1200 unique visitors per day and had a yearly total of 2,343,012 individual page views from around 85,000 unique IP addresses, a volume of traffic that has been steadily increasing every year since the first launch of the system in late 2002. While the majority of visits come from users in the United States, an increasing number are coming from other countries, and the development team has received several inquiries about the possibility of adding international air quality data to the VIEWS database.

In 2006, VIEWS was selected by the Western Regional Air Partnership (WRAP), the largest RPO administered by the Western Governors' Association and the National Tribal Environmental Council, to serve as the infrastructure for the WRAP's Technical Support System (TSS). The TSS is an extended suite of analysis and planning tools designed to help planners develop long term emissions control strategies for achieving natural visibility conditions in Class I Areas by 2064. The TSS was designed upon the foundational database and software framework developed for VIEWS, and leverages this infrastructure to provide a unique decision support system for Western states and tribes that are seeking to develop emissions control strategies and track the effectiveness of those strategies over time. The TSS also consolidates the data resources of the WRAP's Emissions Data Management System (EDMS), Fire Emissions Tracking System (FETS), Regional Modeling Center (RMC), and causes of Haze Assessment project. The integrated systems approach that has been used to design the framework of the VIEWS/TSS allows developers to leverage and add value to these additional significant investments in data, expertise, and information technology.

The WRAP has been instrumental to the continued development of VIEWS, both from a funding and planning standpoint. As the largest of the RPOs, the WRAP represents a history of almost 20 years of collaboration on the technical aspects of western air quality issues, and has established valuable precedents for air quality planning and tracking that have often been facili-

tated by its ongoing investment in VIEWS/TSS. This investment, bolstered by the participation and in-kind support of states, tribes, and federal agencies, has been instrumental in establishing a broad basis for developing a regional understanding of air quality issues related to environmental health, ecological and aesthetic protection of our natural areas, and beginning to evaluate the air quality impacts of climate change and adaptation. The VIEWS team plans to continue working with the WRAP to evolve the VIEWS/TSS into an enterprise-level decision support system that can serve the air quality analysis and planning needs of air quality researchers worldwide.

The VIEWS/TSS users are typically asking questions like, "What pollutants are impacting a given area?" and, "Where are these pollutants coming from?" States are further mandated to answer the question, "What can be done to reduce these impacts?" because the Regional Haze Rule requires states and tribes to develop implementation plans for reducing emissions and demonstrating reasonable progress towards doing so. These plans must provide for an improvement during the 20% worst visibility days while also ensuring no degradation during the 20% best visibility days. To accomplish this, users must identify the pollutants impacting a given area, quantify their amounts, and determine the sources of anthropogenic emissions that contribute to this pollution, on both the "best" and the "worst" visibility days in the area. They must then determine available control measures for each source and evaluate these measures on the basis of costs, time, energy and environmental impacts, and the remaining life of the source. Planners then employ these analyses to make decisions about what controls to implement, to estimate projected improvements, and to track their progress in reaching these goals. The resulting decisions have obvious ecological impacts, but can also have important political and economic impacts in the sense that deciding which sources to control is a significant issue for such an aesthetic standard, and the process of controlling

emissions and tracking progress costs money and takes time.

This rather complex decision-making process requires an array of analytical tools and a multiplicity of data. Historically, planners had to assemble this data from a wide variety of segregated sources, reorganize and reformat it, and then employ an inconsistent and heterogeneous collection of tools to examine and analyze it. With VIEWS/TSS, planners can now use a single, standardized database and an integrated suite of tools to access all the resources they need and answer the questions, perform the analyses, and make the decisions required for the development and acceptance of their state and tribal implementation plans.

To further enhance the value of VIEWS for the purposes of air quality decision support, the team collaborated in 2007 with the Institute for the Environment at the University of North Carolina, Chapel Hill to submit a NASA ROSES proposal to incorporate satellite data into VIEWS/TSS. The proposal was awarded, and work began at CIRA in mid-2008 to seamlessly integrate a wide variety of NASA satellite data into VIEWS. It is hoped that the satellite data, through extensive temporal frequency and geographic coverage, will yield important insights into the temporal evolution and the three-dimensional distribution of atmospheric aerosols. In addition, air quality

Figure 3. An integrated suite of interactive visualization and analysis tools provide quick exploration of patterns and trends in the data.
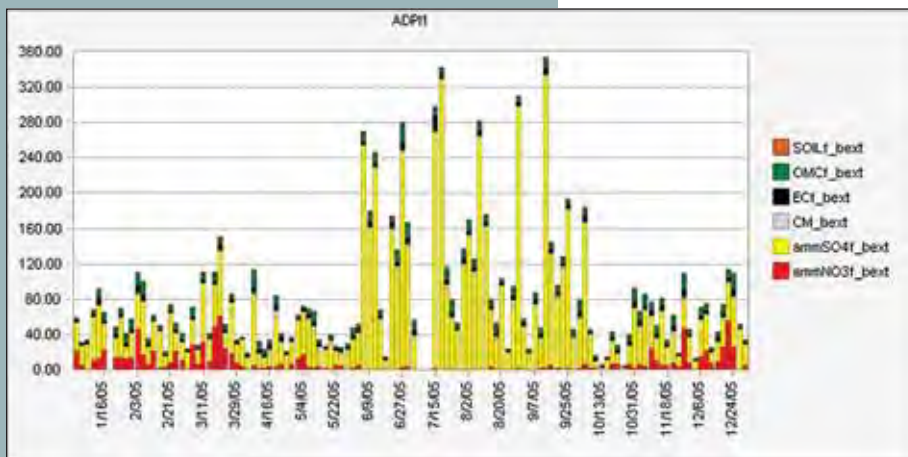
VIEWS
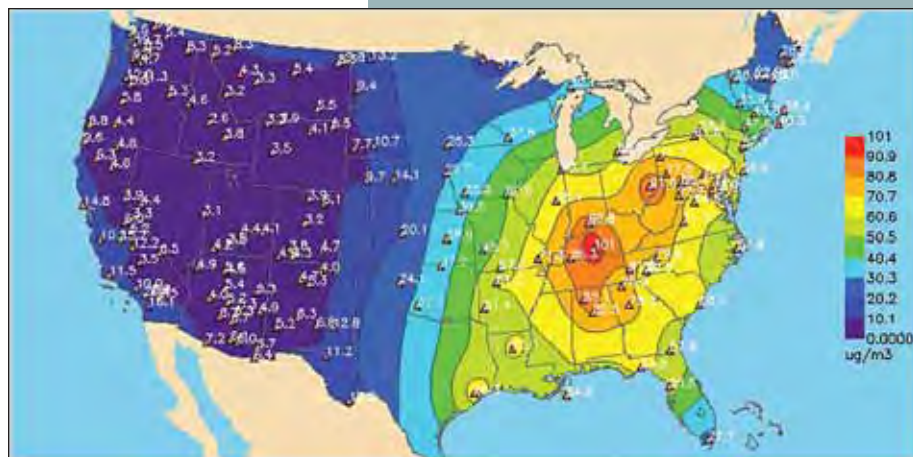
Figure 4. A variety of mapping tools can be used to view the geospatial distribution of the data.

modeling specialists at UNC are using the satellite data to enhance the inputs and boundary conditions of the Community Multi-scale Air Quality Modeling System (CMAQ). Once the model has been appropriately augmented with the satellite data, the results will be added to the raw data already in VIEWS to provide a unique comparison between the original model results and those that have been augmented with the satellite data, an effort that is expected to improve the predictive capabilities of the air quality models and significantly increase understanding of both the observed and simulated systems as changes in anthropogenic and natural emissions occur over time. The availability of both the raw and model-integrated satellite data will complement the existing inventory of ground-based, modeled, and emissions data in VIEWS to provide end users with a uniquely comprehensive collection of air quality data that can be visualized and compared using an integrated set of tools.

The VIEWS team also won a proposal in 2008 to incorporate results from the Navy Aerosol Analysis and Prediction System (NAAPS) in collaboration with the U.S. Naval Research Laboratory (NRL). NAAPS, with data assimilation of NASA Earth Sciences data, produces accurate and timely quantitative forecasts and analyses of the distributions of dust and smoke that originate outside the U.S. The VIEWS team will work with

NRL to add NAAPS data to VIEWS in order to provide better quantification of the influx and impact of international aerosol plumes. The use of these data will result in more thorough identification of exceptional events, long range transport of aerosols, and better estimates of the impact of these sources on U.S. air quality. And because the transport of international aerosol plumes to the U.S. is a complicated, three-dimensional process, the combination of NAAPS data with the satellite aerosol optical depth data (being added to VIEWS as part of the 2007 ROSES project) will provide a unique opportunity to study the long range transport of aerosols and may also help yield improvements to air quality model inputs and boundary conditions.

From the outset, a primary goal of the VIEWS development team has been to create a foundational software architecture that can be leveraged and extended easily by both the VIEWS team and outside developers that are implementing air quality decision support systems. To facilitate this "interoperability", an emphasis has been placed upon designing VIEWS with a Service-Oriented Architecture (SOA), a design principle that separates functionality into distinct units, or services, which are provided over the Internet for users to combine and reuse in the development of their own applications. At present, many of the data retrieval mechanisms of VIEWS already make use of these services internally, and VIEWS developers are working to finalize these services so that they can be made directly available to other developers over standard Internet protocols. When complete, this collection of services will be offered to end users and other developers through an application programming interface that will allow the creation of "mashups" using VIEWS data and services.

A critical component of this effort is collaboration with developers of other air quality systems to establish flexible and comprehensive metadata standards. To facilitate this, the VIEWS team has been participating in the development of the Global Earth Observation System of Systems (GEOSS), a long term project of the Group on

Earth Observations (GEO) to collect and disseminate improved data, information, and models to stakeholders and decision makers. In 2008, VIEWS was selected to participate in the GEOSS Architectural Implementation Pilot as a "persistent operational exemplar" of a decision support system that is developing services that support the GEOSS Societal Benefit Areas (SBAs) using a consistent architecture and suite of standardized data exchange protocols. VIEWS will participate as an air quality data node in the proposed "system of systems", and will help facilitate the development of data exchange and interoperability standards to be utilized in the project.

The VIEWS team also participated in the EPA Data Summit held at Research Triangle Park, NC. The purpose of the summit was 1) to convene organizations and individuals with key roles in retrieving, storing, disseminating, and analyzing air quality data in order to learn about and explore efficient means of leveraging the numerous individual efforts underway; 2) to examine the mechanisms and potential opportunities for "interoperability" between existing systems, 3) to assist EPA/OAQPS in honing its role in the larger air quality data community; and 4) to begin to establish a community-wide strategy for responding to user defined needs. Subsequently, the EPA provided funding to help VIEWS become more of a multiple-pollutant data system and expand its scope beyond regional haze issues. Multiple datasets were considered, such as Ozone, CO, SO2, NH3, etc, and it was eventually decided that the incorporation of AQS Ozone data would be the optimal pilot project by which to begin this effort. The VIEWS team has imported the hourly EPA Ozone data into the integrated database and will be working with the EPA to design and develop tools for exploring, visualizing, and analyzing the data.

In conjunction with these collaborations, the VIEWS team has been working over the past two years to develop an updated version of the website that will be initially launched in the fall of 2009 and further enhanced with additional incremental releases over the following year. The new website incorporates a variety of new technologies and techniques, and is deployed upon newly upgraded hardware. The update introduces a new "look and feel" to the site and will offer several new features, including a user forum, a new database query wizard, dynamic contour maps, air quality index maps, and daily images from NASA's Terra, Aqua, Aura, and CALIPSO satellites. Various new tools also make use of Google maps for the display of monitoring site locations and geospatial datasets, and a searchable "Dataset Index" allows the user to more easily explore the complete inventory of the integrated database and search for data by keyword. Photographs of monitoring site locations are available for some networks, and visibility and weather webcams are also offered. Raw data can be downloaded as files or visualized through dynamic charts, graphs, and maps. The VIEWS website is publicly accessible and does not require a username and password, but users are encouraged to register for an account so that they can take advantage of certain personalization features and help build an online community of air quality data specialists and consumers.

The VIEWS development team consists of principal investigator Shawn McClure, a software engineer who developed the database and

Figure 5. The locations of known emissions sources and power generation facilities are shown on a map using data from the EPA's eGRID database.
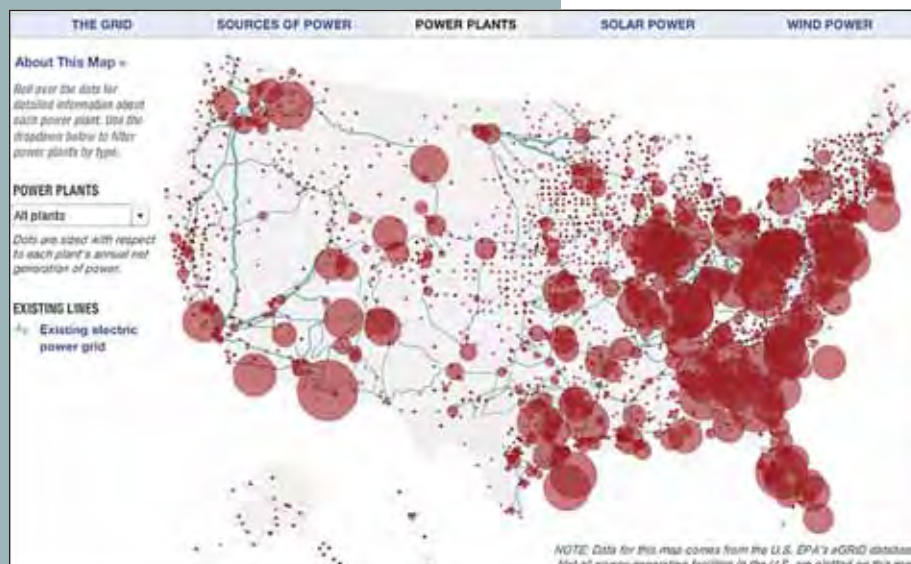
VIEWS
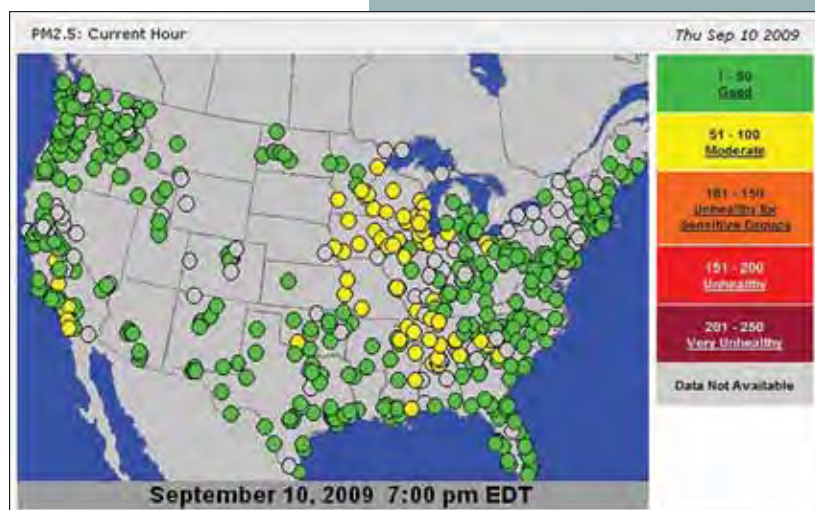
Figure 6. Air Quality Index (AQI) maps from AIRNow.gov are made available in a dynamic tool that allows quick date range selection and display customization.

State University, and researcher Dr. Duli Chand, a scientist from the University of Washington in Seattle who holds a PhD in Atmospheric Trace Gases and did his post doctorate at the Max Planck Institute for Chemistry in Mainz, Germany. Providing administrative leadership from the WRAP and the Western Governors' Association is Tom Moore, and providing scientific guidance from the National Park Service is Dr. Bret Schichtel. In addition, Helene Bennett of CIRA contributes invaluable administrative and operational support on a daily basis, and Dr. Bill Malm of the National Park Service has also contributed instrumental scientific guidance and support. The VIEWS team is committed to making the system as useful and relevant as possible to the evolving needs of air quality researchers worldwide, so visitors to the website are encouraged to provide comments, suggestions, and requests for additional new data and functionality.

software infrastructure on which VIEWS is built, developer Dr. John Huddleston, a retired federal employee and registered professional engineer who holds a PhD in Geophysics from Colorado

Job Opportunities in Atmospheric Science and Related Research

If you are interested in employment opportunities at CIRA, please visit http://www.cira.colostate.edu/personnel/employment_opportunities and enter your e-mail address. Then, when an open position is posted on the CIRA website, you will receive an e-mail prompt to view the position announcement and apply if interested/qualified. Topical areas relevant to CIRA include:

Research and Postdoctoral Research Fellowships:
• Satellite Algorithm Development, Training, and Education: development of algorithms and applications for weather forecasting, emphasizing regional and mesoscale meteorological phenomena. Development of environmental parameter retrieval techniques based on current and future satellites, potentially in concert with in situ, or other remote sensing observations. Development of training/educational materials for operational forecasters using distance learning methods and web-based demonstrations.
• Regional to Global Scale Modeling Systems: providing improvements to weather and climate numerical prediction. Topics include atmospheric and ocean dynamics, radiative forcing, clouds and moist convection, land surface modeling, hydrology, and coupled modeling of the Earth system.
• Data Assimilation: development and improvement of techniques to assimilate environmental observations (satellite, terrestrial, oceanic, and biological) to characterize the environmental state for use in analysis, modeling, and prediction of weather/climate.
• Climate/Weather Processes: use numerical models and environmental data, including satellite observations, to understand processes that are important to predicting weather and short-term climate and understanding the coupling between weather and regional climate.

• Data Distribution: identify effective and efficient methods for quickly distributing/displaying large environmental datasets (observational and model fields) using data networks, web map services, data compression algorithms, and other techniques.

Information Technology Support:
• Software Engineering, Electrical Engineering, Linux Cluster Administration, Linux and Windows IT, Data Acquisition, Large Data Set Processing and Archive, Satellite Operations, Infrastructure Management, Web Development, Database Development and Maintenance, Network Management

Administrative Support:
• Customer service, grants management and administration, and general office duties

CIRA is jointly sponsored by Colorado State University and the National Oceanic and Atmospheric Administration. Colorado State University is an equal opportunity/affirmative action employer and complies with all Federal and Colorado State laws, regulations, and executive orders regarding affirmative action requirements in all programs. The Office of Equal Opportunity and Diversity is located in 101 Student Services Building. In order to assist Colorado State University in meeting its affirmative action responsibilities, ethnic minorities, women and other protected class members are encouraged to apply and so identify themselves. Colorado State University is committed to providing a safe and productive learning and living community. To achieve that goal, background investigations for all final candidates being considered for employment are conducted. Background checks may include, but are not limited to, criminal history, national sex offender search, and motor vehicle history.

CIRA COMMUNIQUÉ

Please congratulate the following employees who received the 2009 CIRA Research and Service Initiative Award:

Linn Barrett … for outstanding performance, dedication to the people of CIRA in successfully managing the complex Human Resources function, and contributions to the administrative team

Jacques Middlecoff … for outstanding performance as an integral part of the development of GSD's state-of-the-art global weather and climate predication model

Sher Schranz … for outstanding leadership of researchers and software developers in the design, development, testing, deployment and operational support for the FX-Net, Gridded FX-Net and Fire Weather Projects

Ning Wang … for outstanding performance as an integral part of the development of GSD's state-of-the-art global weather and climate predication model

2008 Clean Air Excellence Awards Recipients

EPA's Clean Air Excellence Awards Program, established at the recommendation of the Clean Air Act Advisory Committee, annually recognizes and honors outstanding, innovative efforts that help to make progress in achieving cleaner air. The following recipient received a Clean Air Excellence Award in May 2009.

Dr. William Malm
National Park Service
Thomas W. Zosel Outstanding Individual Achievement Award

While some call him Dr. Visibility in the United States, Dr. William Malm of the National Park Service is more formally recognized as the leading scientist behind the visibility protection provisions of the Clean Air Act. His science-driven policies are a testament to his dedication to the environment and his perseverance in bringing science to the issue of air quality.

Since making some of the first visibility and air quality measurements in the National Park Service system at the Grand Canyon in 1972, he has designed and built instrumentation to measure the effects of atmospheric aerosols on the scenic qualities of landscape features, as well as their optical and chemical properties. By linking visibility impairment to specific sources, Dr. Malm's studies have lead to requirements for pollution reduction at major power plants in the Southwest. Through his formulation of radiation transfer algorithms, his pioneering of visibility perception studies, and his leadership in collaborative efforts, Dr. Malm has also played an integral role in improving air quality by significantly reducing sulfur emissions.

In addition to his technical achievements, Dr. Malm serves as the intellectual leader responsible for the Interagency Monitoring of PROtected Visual Environments (IMPROVE) Network. From the establishment of the IMPROVE monitoring network, to the development of the IMPROVE equation (which is the basis for EPA regional haze regulations), to the very metrics used to characterize visibility, he has applied sound science to protecting our nation's most treasured vistas.

Dr. Malm has demonstrated leadership, outstanding achievement, and lasting commitment to promoting clean air and helping to achieve better air quality for 30 years. The steadily improving visibility we enjoy in many parts of the U.S. is largely due to the research and advocacy of Dr. William Malm.

CIRA retiree, **Jim Purdom**, recently participated as a member of the US Delegation to CBS-XVI in Dubrovnik, Croatia. Highlights included:

1) served as a member of the U.S. Delegation to TECO WIGOS (March 23-24) and CBS (March 25- April 1).

2) presented accomplishments for the CBS Open Program Area Group on Integrated Observing Systems (OPAG-IOS).

3) was awarded the Cosmonautics Federation of Russia Yuri Gagarin Medal.

Mark DeMaria received the Richard H. Hagemeyer Award at the 63rd Interdepartmental Hurricane Conference, which is given annually to honor people who have made sustained contributions to the U.S. Hurricane Program. The primary basis of the award was a long history of successful development and operational transition of tropical cyclone forecast products, including the SHIPS and LGEM intensity models, the rapid intensity index (co-developed with HRD), the NHC wind speed probability program and the NESDIS tropical cyclone formation probability product. Previous winners of the award include Joanne Simpson, Max Mayfield and Chris Velden.

The paper "Stratospheric impact of the Chisholm pyro-cumulonimbus eruption: Part 1. Earth-viewing satellite perspective," by M. Fromm (NRL) et al., for which **Dan Lindsey** is a co-author, has been selected to receive an Alan Berman Research Publication Award at the Naval Research Laboratory. *Congratulations to Mark and Dan!*

Please also join me in congratulating **Michael Hiatt**. Michael was honored at the "Celebrate Colorado State!" awards ceremony yesterday afternoon for the 2009 Distinguished Administrative Professional Award. The award is presented to Administrative Professional staff for continuing meritorious and outstanding achievement in the areas of outreach, teaching, administration

Please congratulate the following employees for reaching service milestones:

Travis Andersen – 10 years of service

Linn Barrett – 10 years of service

Daniel Bikos – 10 years of service

Kevin Brundage – 15 years of service

Scott Copeland – 15 years of service

Laura Fowler – 20 years of service

Paul Hamer – 10 years of service

Jeffrey Lemke – 15 years of service

Chungu Lu – 20 years of service

Mary McInnis-Efaw – 10 years of service

Evan Poster – 10 years of service

Dale Reinke – 20 years of service

Daniel Schaffer – 10 years of service

David Watson – 20 years of service

Loretta Wilson – 25 years of service

and/or research and is awarded by the AP Council.

The citation notes: "As Infrastructure Manager, Michael leads CIRA's technology support team and works with project leaders to strategize and incorporate technology in CIRA's research projects. His ability to assess technology needs and execute efficient and timely solutions is a valued skill to the organization. He accomplishes goals within the project budget with a history of successful projects."
Congratulations, Michael!

It is my pleasure to pass along the news that **Don Reinke** headed to NASA/JPL in early June to accept his recently announced "NASA Exceptional Public Service Medal" for his leadership and tireless efforts on behalf of the CIRA CloudSat Data Processing Center!

The citation notes: "Don's contributions to the CloudSat mission have been truly exceptional and have led to a cost-effective, robust, and flexible data processing operation that has been praised by several NASA review boards. This award recognizes your exemplary performance as the manager of the CloudSat data processing center during system design, development, and operations."
This is a prestigious honor, so please join me in offering our congratulations to Don!

June 2009 GSD Team Member of the Month – Kevin Brundage

The following nomination comes from Assimilation and Modeling Branch Chief Stan Benjamin.

"Kevin, as a computer scientist, has been a critical member not only of the Assimilation and Modeling Branch, but also of GSD, ESRL, and NOAA. For RUC, Rapid Refresh,

HRRR, and FIM development, Kevin has repeatedly found more efficient designs for scripts, code, post-processing, and grid transfer that have been very important for reliability of these complicated model/assimilation systems. Kevin is also a critical member of the ESRL High-Performance Computing (HPC) (jet) management team, including developing benchmarks for overall NOAA HPC requests for proposals."

August 2009 GSD Team Member of the Month – Randy Collander

Randy was part of a team that was awarded the August GSD Team Member of the Month award for their work on MADIS (Meteorological Assimilation Data Ingest System). FSL's MADIS Project makes integrated, quality-controlled observations available to the meteorological community. More information on MADIS can be found here:
http://www-sdd.fsl.noaa.gov/ MADIS_Overview/MADIS_Overview.html

US-Norway Research Collaborations Strengthen

A recent initiative from U.S. Secretary of State Hillary Clinton and Norwegian Minister of Foreign Affairs Jonas Gahr Støre, aims to strengthen U.S.-Norwegian collaboration in polar research, with a special focus on Svalbard, Norway. Their initiative was presented and discussed at the 50-year anniversary meeting of the joint Arctic Council and Antarctic Treaty Consultative Meeting in Washington, D.C., in April 2009, and identified key scientific questions that should be given priority.

Among the list of suggested Norway-US research collaborations was a merging of CIRA snow, weather, and climate modeling tools with Norway's field observations in Svalbard, Norway (the high Arctic). For the past 15 years, under the leadership of **Dr. Glen Liston**, scientists at CIRA/Colorado State University have been developing regional climate modeling tools to simulate the evolution and changes occurring within Arctic systems. These tools include atmospheric models, hydrologic models, ecosystem models, and data assimilation models. In addition, they have developed a nested modeling structure that is able to simulate the feedbacks and interactions among the key climate systems and the associated physical and biological climate system variables. Development of a regional climate model for Svalbard would be an excellent tool for climate-change researchers and managers interested in understanding an Arctic "hot spot," i.e., an Arctic region where the climate is changing very fast.

CIRA Student, Darren Van Cleave, Honored at the AMS Numerical Weather Prediction Conference

At the conference in Omaha there were student awards given in four categories among 42 competitors:
(1) PhD Student Oral Presentation (18 competitors),
(2) PhD Student Poster Presentation (7 competitors),
(3) M.S./B.S. Student Oral Presentation (8 competitors), and
(4) M.S./B.S. Student Poster Presentation (9 competitors).

Darren tied for first place in category 4 with his poster "Relat-

ing Snowfall Patterns over the Central and Eastern US to Infrared Imagery of Extratropical Cyclone Comma Heads."

All students who presented at the conference have much to be proud of; they and their faculty advisors have done a superb job and helped to advance our science and nurture our future research pioneers.

Louie Grasso Visits Third Graders

"Last spring, April 2009, Mrs. Weisman's third grade Bauder Elementary class studied outerspace. She invited me to visit her class to discuss space, gravity, and show pictures of deep-space objects. During my visit I asked the class to, 'jump up.' Everyone was unable to 'just' jump-up; they all came back down. When I asked them why they all came back down, I was overwhelmed with all kinds of answers. After their 'gravity' experiment, I explained how gravity helps to form a star. From here, I introduced them to two stars rotating around each other: A binary pair.

"The class was more than happy to break up into groups of two and then rotate around each other like a binary pair of stars. I then asked if they would like to see some deep-space objects. They all replied with an enthusiastic yes! We proceeded onto a viewing room. They saw actual images of open star clusters, globular star clusters, gaseous nebula, planetary nebula, and galaxies. I finished up by telling them that all deep-space objects are held together by the same thing that holds us all to the ground: Gravity."

SOS

Science On a Sphere® Update – New Sites and New Visualizations

Mike Biere and Steve Albers

The Science On a Sphere® (SOS) project continues to gain momentum. The exhibit is now installed at nearly 50 museums and science centers, including a number of foreign sites. Foreign sites with an SOS exhibit now include *National Museum of Natural Science* in Taiwan, *Gwacheon National Science Museum* in South Korea, *Climate Institute* in Mexico, *Cité des Sciences et de l'Industrie* in France (see Fig. 1), and most recently *Finnish Science Center* in Finland. In addition to these new permanent exhibits, the French space agency, CNES, used their SOS as the centerpiece of their pavilion at the 2009 Paris Air Show (see Fig. 2).

Domestic installations of SOS have also been accelerating this year. SOS exhibits in the U.S. added so far in 2009 are *Harsco Science Center*, Harrisburg, PA; *North Carolina Aquarium*, Manteo, NC; *National Zoo*, Washington D.C.; *Alaska State Museum*, Juneau, AK; *Stennis Space Center*, MS; *Wildlife Experience*, Parker, CO; and *Oregon Museum of Science and Industry*, Portland, OR.

Dataset and Visualization Update

CIRA staff continue to produce an expanding variety of datasets, many of which take the form of animations, often in real-time. Locally gathered real-time earthquake data from the Taiwan Central Weather Bureau has been integrated with the global database from USGS to give greater earthquake accuracy over Taiwan (see Fig. 3). A new animation of the NASA "Blue Marble" image with nighttime lights was also created (see Fig. 4).

We made an animation showing the establishment of GLOBE schools around the world over the past 14 years. We worked on setting up a quasi real-time animation from the STEREO pair of satellites orbiting the sun (Fig. 5) to show more than half the solar disk in extreme ultraviolet light. This involves adapting our re-projection software so it can work with the spacecraft images provided by the Goddard Space Flight Center. An

*Figure 1. Exhibit Installation in progress at* **Cité des Sciences et de l'Industrie** *in Paris, June 2009. CIRA staff on the SOS project were on site to assist at this first permanent European installation of Science On a Sphere.*

Figure 2. SOS was the centerpiece of the CNES pavilion at the 2009 Paris Air Show. CIRA personnel aided CNES in setting up their traveling SOS exhibit for the air show.
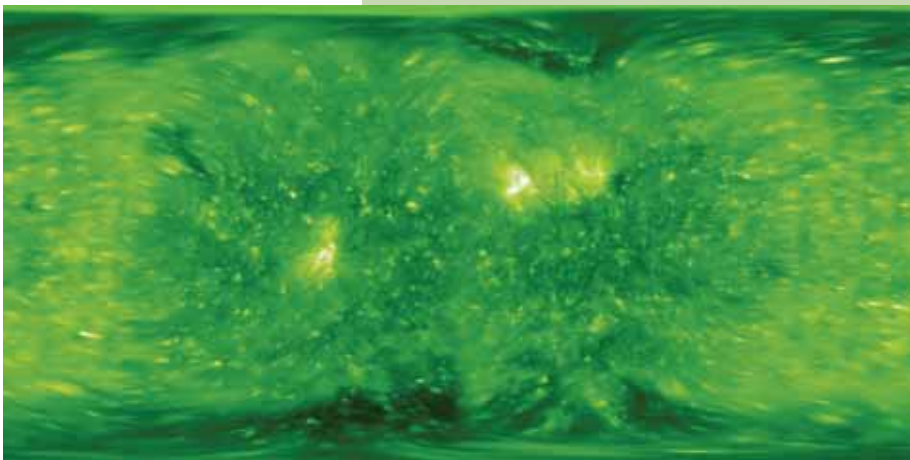
SOS
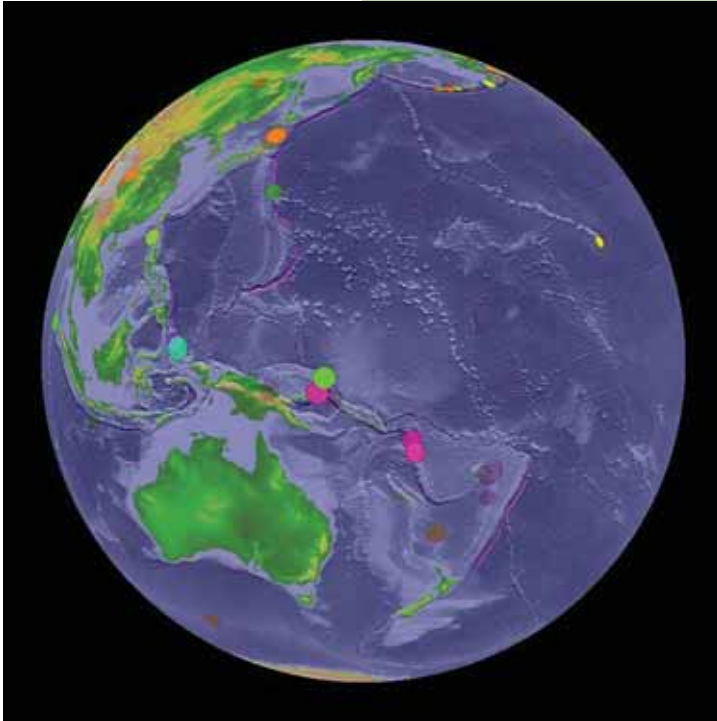
Figure 3. Real-time earthquake data from the merged databases of the USGS and the Taiwan Central Weather Bureau.

Figure 4. A new animation of the NASA "Blue Marble" image with nighttime lights.

Figure 5. Cylindrical map of the Sun in extreme ultra-violet light as mapped using a pair of views about 90 degrees apart from NASA's STEREO satellite. Note that more than half of the Sun's surface is visible in this view on May 17, 2009 at 2340UTC. The remaining portion of the sun is filled in with "persistence" information from when the far side had last rotated in view. These are displayed in near real-time with Science On a Sphere. Animations are shown using frames at 10-minute intervals and help to provide advance warning of solar storms.

algorithm was developed to use persistence imagery to fill in the part of the sun that has rotated out of view.

Looking elsewhere in the solar system, the map of Mercury was updated with improved use of older Mariner imagery along with improved navigation of the newer images from Messenger. Circling Saturn are many interesting satellites; the map of the moon Mimas was reworked with improved navigation information for the individual images. The Enceladus map now has color imagery more widely used, as well as improved navigation in the vicinity of the South Pole.

Some new Cassini imagery was added to the Tethys map. Maps were updated for Rhea, Titan and Iapetus using the latest available Cassini spacecraft imagery. The Iapetus map is also slated to be used in a paper for the journal *Science*.

Looking ahead to the future, we are currently working on some lightning displays as well as combined earthquake/tsunami animations.

LDM

A High Performance LDM Data Cluster

Bob Lipschutz

Within the NOAA Earth System Research Laboratory (ESRL), the Global Systems Division (GSD) develops weather information systems, weather forecast models, and other applications in support of the National Weather Service, the Federal Aviation Administration, and other agencies. Well-known GSD products include the Rapid Update Cycle (RUC) model, the Local Analysis and Prediction System (LAPS), the Meteorological Assimilation Data Ingest System (MADIS), and Science on a Sphere (SOS). A common feature of these and other GSD projects is that they require observational and model data provided by acquisition systems running within GSD's Central Facility (CF). The CF systems, created by CIRA developers (Chris MacDermaid – Lead, Bob Lipschutz, Paul Hamer, Amenda Stanley, Leslie Ewy, Patrick Hildreth, Richard Ryan, and Glen Pankow) in the Data Systems Group (DSG) of the Information and Technology Services (ITS) branch handle some 500 GB of incoming data per day as they acquire, decode, store, and distribute the needed datasets for GSD scientists and their collaborators.

Recently, to substantially improve GSD's CF data systems, the DSG CIRA team worked with ITS systems administrators to assemble and configure a new six-host clustered data system. This new system, dubbed the DSG Cluster (DC), was commissioned on September 1, 2009. Replacing a collection of aging Linux High-Availability (HA) pairs and stand-alone platforms, the DC is a scalable Linux cluster that offers high throughput performance as well as excellent reliability, resource utilization and configurability.

Limitations of the HA Architecture

While the CF's HA pairs have operated successfully for a number of years (Lipschutz and MacDermaid, 2005), and provided robust availability through their automated fail-over mechanism, a major drawback has been the need for two machines, one of which is idle, to support a single processing configuration. Thus, system expansion or refreshment necessitated acquiring and supporting two new hosts at a time, adding to systems administration effort, power and cooling burden, and cost. In addition, while HA pairs may initially have shared common system configurations, the individual nature of the systems resulted in drift in the configurations over time that became problematic to maintain.

At the application level, needing to individually configure each host's functionality was tedious. It often forced choices for where to locate processing based on where prerequisite data existed, rather than where system resources were available. This led to unbalanced processing loads, or alternatively, more complicated data flows. Further, limited local disk space on the HA systems often hampered processing of larger datasets. Documenting the processing for a dataset and troubleshooting data outages was complicated by the need to identify the specific platforms where processing elements were performed.

LDM

Goals for the Cluster

In planning for the cluster, the intent was to address the HA systems' limitations with these benefits:

- Scalability – general processing capacity can be added one machine at a time, rather than two.
- Throughput – cluster load balancing mechanisms allow for much greater aggregate throughput than was achieved on hand-configured systems. Of course, throughput also improves using newer processors and fast networking and storage devices.
- Utilization – cluster hosts are more effectively used than HA pairs in that half the hosts are not virtually idle backups. Rather, all cluster nodes are active, and load balancing methods ensure that jobs are distributed sensibly. In addition, the cluster's shared storage arrangement provides space to all cluster applications, whereas in the HA architecture, spare disk space on one host is unavailable to another host that might need space.
- Maintainability – cluster hosts share a common system configuration, substantially simplifying ongoing system administration. Host reboots and failovers for patching also are easier than with the HA arrangement. On the application side, multiple configurations are replaced by a single, common set of configuration files. In addition to reducing the number of files to maintain, this reduces the total number of configuration lines by eliminating duplication of many configuration items across hosts.
- Supportability – application data and log files reside on the shared storage, eliminating the need to know which specific host ran the job.
- Energy efficiency – the cluster allows us to perform data processing more efficiently in terms of energy consumption. Higher efficiency in CPU utilization means less heat generation, and lower cooling requirements compared to the HA computing infrastructure.

In addition to these general goals, we strongly desired to retain as much as possible of the long-established CF data processing application software, and thus minimize code and configuration changes. We further recognized a continuing need, well served by the HA pairs, to be able to split onto separate platforms such high volume data streams as those from the NWS Satellite Broadcast Network (SBN – NOAAPORT) and the WSR-88D Level-II radars, and thereby use independent processing resources to minimize resource contention.

Software Components

To achieve these goals, we implemented the DC with these key open source software components:

- the **CentOS** Enterprise Linux operating system,
- the **Red Hat Cluster Suite** for managing cluster-wide application services and failovers,
- **Sun Grid Engine (SGE)** for job activation and load balancing,
- **fcron** for cluster-wide time-based job triggering,
- Unidata's **Local Data Manager (LDM)** for data transport and event-based job triggering, and
- **Open-E** storage software for controlling the NFS-mounted back-end Data Storage Server (DSS) RAID disk system.

Within the framework provided by these software elements, the Object Data System (ODS) applications developed by CIRA's Paul Hamer (Hamer, 2005) are responsible for such data processing tasks as converting GOES GVAR satellite data, Gridded Binary (GRIB) model data, WSR-88D Level-II radar data, and a variety of point observation data types into the netCDF formats needed by GSD user applications. Only minimal changes to the existing ODS software were needed to accommodate the cluster architecture. In addition, two new scripts were developed to provide common methods for submitting jobs to the SGE job queue from LDM (*pqact*) and *fcron*, while several old scripts were extended to include a cluster-wide locking mechanism to avoid concurrent instances of some jobs.

LDM and ODS

The Unidata Local Data Manager has long served as the key middleware in the CF for client-server data transport and event-driven data processing services, running on numerous hosts for a variety of data processing tasks. Extending LDM capabilities for GSD, a number of ODS LDM client applications attach directly to the LDM data queue to find, extract, and process desired data messages. For example, the *LdmGrib2Flat* client captures GRIB records in the NOAAPORT stream and writes them into forecast hour GRIB files in dynamically determined directories using the center, sub-center, model, and grid parameters provided within the data. Likewise, the ODS LdmPoint2Tar client writes selected point data messages into hourly tar files by data type.

On completion of a GRIB file, *LdmGrib2Flat* emits a data arrival "notification" message indicating the location on disk of the newly available file (Fig. 1). The notification identifiers are specified in a *pqact* configuration file to determine subsequent processing. For instance, given a matching notification key, a GFS model GRIB file may be converted to netCDF using the ODS *Grib2NetCDF* application. In particular, *pqact* pipes the notification to a script which extracts the GRIB file path, then submits the appropriate *Grib2NetCDF* command to the SGE job queue for execution.

Point data type tar files can be subsequently read by the *Point2NetCDF* program to decode the raw data and create netCDF files, for example containing METAR or maritime observations. These jobs typically are run on a schedule submitted to the SGE queue by *fcron*.

Notification message identifier string:

> GSD.TEXT.MMXX53.KBOU.082144.LdmGrib2Flat_NOAAPORT.NCEP.GFS.A211.024Hour.200909081800.0925118000024.*

Notification message contents:

> MMXX53 KBOU 082144
>
> DATA AVAILABILITY NOTIFICATION MESSAGE 09-251
>
> NOAA/ESRL GLOBAL SYSTEMS DIVISION, BOULDER, CO
>
> Tue Sep 8 21:44:27 2009
>
> file: /public/data/grib/noaaport/7/0/96/211/0925118000024
>
> END - SENT - GSD/ITS

Figure 1. A sample ODS data arrival notification message.

LDM Data Services

A key cluster design decision was to differentiate the LDM data flows on the cluster from the notification-based event handling. To handle incoming and outgoing data, independently defined LDM instances are configured under Cluster Suite to run as relocatable cluster services. These services normally run on different hosts, but can run concurrently in a failover situation. The specific hosts on which each service may run is determined in Cluster Suite by defining so-called Failover Domains.

Based on the known data flows and the desire to logically group some of the processing and configuration items, the following services have been established:

- ldm-noaaport – ingest service for all NOAA-PORT data
- ldm-radar – ingest service for WSR-88D Level-II and WSI NOWrad radar data
- ldm-other – ingest service for GOES GVAR and miscellaneous data types

Figure 2. DSG Cluster logical flow.

DSS disk array. As files are stored, the various data writing methods emit notification messages. And, rather than inserting the notifications back into their own service queues, the writers instead put the notification messages into a local "Notify LDM" queue.

The Notify LDM and SGE

The Notify LDM (*ldm-notify*) operates independently on all the cluster hosts to initiate processing on receipt of data arrival notification messages. Using identical *ldmd.conf* and *pqact.conf* configuration files on each host, the Notify LDM instances start up at host boot time using the local host IP address. While requesting no external data, the Notify LDMs accept the notification messages that are inserted into their notify queues by the various ingest and processing jobs running on the cluster. The Notify *pqact* processes can then determine the next processing step(s) to be taken for each newly arrived data file. For example, some files simply need to be copied from the DSS to GSD's public file server. Other files may require additional format conversion or other processing. After each processing step, a new notification message is inserted into the Notify queue to provide a trigger for yet another step.

The key to the DC architecture is this: rather than simply running jobs on the host on which a notification was emitted, jobs are submitted to SGE, which finds the least loaded host for the job. Since data are visible from each host and Notify LDM is configured identically on all hosts, processing is thus dynamically distributed across the cluster members. If a member leaves the cluster, say for OS patching, the remaining hosts pick the work that would have been assigned to the missing host.
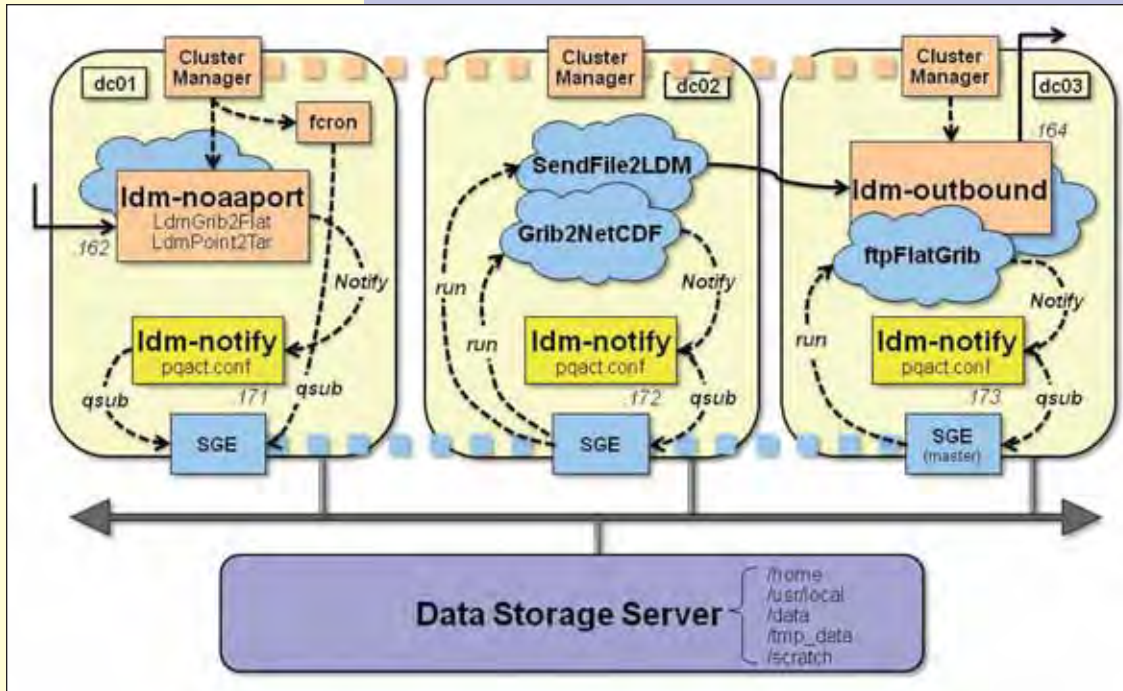
- ldm-outbound – distribution service for sending DC data to other CF hosts
- ldm-fdr – Facility Data Repository service for saving data to GSD's Mass Store System

To facilitate the management of these LDM instances in a shared environment, we slightly modified the *ldmadmin* driver script to allow us to relocate the standard configuration and data queue files into instance-specific directories. Each instance establishes a unique IP address in addition to its distinct configuration and queue, enabling multiple *rpc.ldmd* services to exist on a host. Thus, an external LDM may request data from the ldm-outbound service without knowing the particular host it happens to be on. At startup, an LDM service will first create a new memory resident (*tmpfs*) file system for its queue and construct a new data queue. This scheme provides the highest possible throughput performance and guarantees that on failover an LDM queue is not corrupted.

As an LDM ingest service receives requested data, its associated *pqact* and ODS client processes are responsible for writing the data to the shared

Logical Flow

Typical processing flow is depicted in Fig. 2. While there are numerous scenarios for processing the many types of data handled by the cluster, the following steps describe a typical flow:

- On system startup, the ldm-noaaport service is started by Cluster Manager on dc01 and starts receiving NOAAPORT data from its upstream source.
- The *LdmGrib2Flat* client, which is started by ldm-noaaport on dc01, reads from the ldm-noaaport queue and writes GRIB data into time-based filenames in product-specific directories under /data/grib/noaaport.
- On completion of a GRIB data file, LdmGrib-2Flat inserts a LdmGrib2Flat_NOAAPORT_Local notification into the local ldm-notify queue.
- *pqact* on dc01 identifies the notification message as one that requires a *Grib2NetCDF* action, and submits the configured job into the SGE queue.
- SGE finds the least loaded host in its list of available cluster hosts and runs the *Grib-2NetCDF* job on that host, say dc02.
- *Grib2NetCDF* on dc02 reads the specified GRIB file and creates a NetCDF file in the configured directory under /tmp_data/grids. When finished, *Grib2NetCDF* inserts a Grib-2NetCDF_Local notification message into the local ldm-notify queue on dc02.
- *pqact* on dc02 identifies the notification message as one that requires a *moveTmpDataFiles* action to copy the file from the /tmp_data location to the desired public server location, and so submits the configured job to SGE.
- *moveTmpDataFiles* is started by SGE on the least loaded node, and on completion inserts a new notification message in the ldm-notify queue on that node.
- If necessary, another job may then be activated, for example to send the netCDF file into the ldm-outbound queue for delivery to an external customer.

As the activity driven by the LDM data ingest services proceeds, another large set of actions is initiated by time-based data acquisition, mainly ftp fetches, triggered by the cluster's *fcron* service. Similar to the *pqact* jobs, the *fcron* jobs are also submitted to SGE, and so distributed evenly across the cluster. In addition to acquisition tasks, *fcron* jobs also perform a variety of housekeeping tasks such as file purging and monitoring.

Concluding Remarks

The DSG Cluster system now operating in GSD's Central Facility establishes a reliable, capable, and highly scalable new architecture for meeting substantial and ever-growing data ingest and processing requirements. While the learning curve traversed to successfully implement this system was relatively steep and, at times, arduous, the resulting system has been well worth the effort.

Currently, nearly 200 *pqact* entries are configured to handle specific data arrival events and *fcron* manages some 140 different time-based jobs. Driven by the data flow and the work to be performed, we are presently logging nearly 160,000 jobs submitted to the cluster's SGE queue each day. We expect these numbers to only grow as new requirements are added by our GSD data consumers.

References

Hamer, P., 2005: "FSL Central Facility data systems concepts." *21st Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, 9-13 January 2005, Amer. Meteor. Soc, San Diego, CA.

Lipschutz, R. and C. MacDermaid, 2005: "Recent advances in the FSL Central Facility data systems." *21st Int. Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, 9-13 January 2005, Amer. Meteor. Soc, San Diego, CA.

CIRA Mission

The Cooperative Institute for Research in the Atmosphere (CIRA) is a research institute of Colorado State University.

The overarching Vision for CIRA 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.

Expanding on this Vision, our Mission is:
To serve as a nexus for multi-disciplinary cooperation among CI and NOAA research scientists, University faculty, staff and students in the context of NOAA-specified research theme areas in satellite applications for weather/climate fore-casting. Important bridging elements of the CI include the communication of research findings to the international scientific community, transition of applications and capabilities to NOAA operational users, education and training programs for operational user proficiency, outreach programs to K-12 education and the general public for environmental literacy, and understanding and quantifying the societal impacts of NOAA research.

Cooperative Institute for Research in the Atmosphere

College of Engineering – Foothills Campus, Colorado State University
1375 Campus Delivery, Fort Collins, CO 80523-1375
(970) 491-8448
www.cira.colostate.edu

If you know of someone who would also like to receive the *CIRA Magazine*, or if there are corrections to your address, please notify us.

Colorado State University