Distributed Collaborative Adaptive Sensing for Hazardous Weather Detection, Tracking, and Predicting[*]

J. Brotzge[1], V. Chandresakar[2], K. Droegemeier[1], J. Kurose[3],
D. McLaughlin[4], B. Philips[4], M. Preston[4], and S. Sekelsky[4]

[1] Center for Analysis and Prediction of Storms, University of Oklahoma,
100 East Boyd Norman, OK 73019-1012,
{jbrotzge,kkd}@ou.edu
[2] Dept. Electrical & Computer Engineering Colorado State University,
Fort Collins, CO 80523-1373,
chandra@engr.colostate.edu
[3] Dept. Computer Science, University Massachusetts,
Amherst MA 01003
kurose@cs.umass.edu
[4] Dept. Electrical and Computer Engineering, University Massachusetts,
Amherst MA 01003
{mclaughlin,bphilips,mpreston,sekelsky}@ecs.umass.edu

Abstract. A new data-driven approach to atmospheric sensing and detecting/predicting hazardous atmospheric phenomena is presented. Dense networks of small high-resolution radars are deployed with sufficient density to spatially resolve tornadoes and other dangerous storm events and overcome the earth curvature-induced blockage that limits today's ground-radar networks. A distributed computation infrastructure manages both the scanning of the radar beams and the flow of data processing by dynamically optimizing system resources in response to multiple, conflicting end-user needs. In this paper, we provide a high-level overview of a system architecture embodying this new approach towards sensing, detection and prediction. We describe the system's data rates, and overview various modes in which the system can operate.

1 Introduction

Current approaches for observing the atmosphere are based upon a paradigm of widely separated, geographically fixed sensors that, by necessity, operate independent of the phenomena being observed and of the sometimes disparate needs of multiple end users. This is true of the WSR-88D Next Generation Doppler Radar (NEXRAD) system, which consists of 141 ground-based radars and serves as the cornerstone of the weather and storm-sensing system in the United States. This system has

tremendous capabilities to observe many types of weather phenomena, yet it remains *fundamentally constrained* in sensitivity and spatial resolution owing to the very long operating range (hundreds of km) over which the radars operate. Systems such as this are unable to view much of the lower atmosphere because the radar beam overshoots approximately the lowest 1 km due to the curvature of the Earth. These systems are also unable to spatially resolve many tornadoes and other damage-causing phenomena that have spatial scales smaller than the 2-4 km resolution achieved by the radar beams at long ranges.

This paper introduces a new, transforming paradigm for atmospheric sensing based on Distributed Collaborative Adaptive Sensing (DCAS) networks designed to overcome the fundamental limitations of current approaches to sensing and predicting atmospheric hazards. *Distributed* refers to the use of large numbers of small, solid-state radars, spaced appropriately, to overcome blockage due to the Earth's curvature, resolution degradation caused by spreading of radar beams at long ranges, and the large temporal sampling intervals that result from today's use of mechanically scanned antennas. In addition to providing high-resolution sampling throughout the entire troposphere, this distributed concept lends itself to the efficient utilization of low-power solid-state radars. These radars operate *collaboratively*, via coordinated targeting of multiple radar beams, based on atmospheric and hydrologic analysis tools (detection, tracking, and predicting algorithms). This enables the critical resources of the sensing system, such as radiated power, antenna beam positions, and data communications bandwidth, to be allocated to enable sampling and data acquisition in specific regions of the atmosphere where particular threats exist. *Adaptive* refers to the ability of these radars and the associated computing and communications infrastructure to rapidly reconfigure in response to changing conditions in a manner that optimizes response to competing end user demands. For example, such a system could track tornadoes for public warning while simultaneously collecting information on the parent storm and providing quantitative precipitation estimates for input to hydrologic prediction models. The system is thus driven by the data needs of the end-users as shown in Figure 1. The very first test bed of the project employing four radars will not use phased arrays: it will use off-the-shelf dishes moved around using motorized pedestals, and that system will be up and running in Spring 2005. There is underway a technology push for the project to make a phased array version of the system. This version will replace the individual dishes with phased array panels. Since there won't be any moving parts, the beams will be able to move around much quicker in response to data requests. That system will be fielded in the Spring of 2006 and will have 9 phased array radars arranged in a cooperative network.

In October of 2003, the National Science Foundation created an Engineering Research Center (ERC) among the University of Massachusetts (lead university), University of Oklahoma, Colorado State University, and the University of Puerto Rico at Mayaguez, and a consortium of industrial partners to lay the fundamental and technological foundations for DCAS and to investigate the practicability of this new paradigm for storm sensing and prediction. Called the Center for Collaborative Adaptive Sensing of the Atmosphere (CASA) [2], this ERC will create a series of test-beds that enable exploration of the design-space of DCAS using fielded hardware and software and enable proof-of-concept demonstration experiments involving specific end-users. This paper describes the architecture and data requirements for NetRad, which is the first of a series of system-level test-beds to be created and deployed by this center.
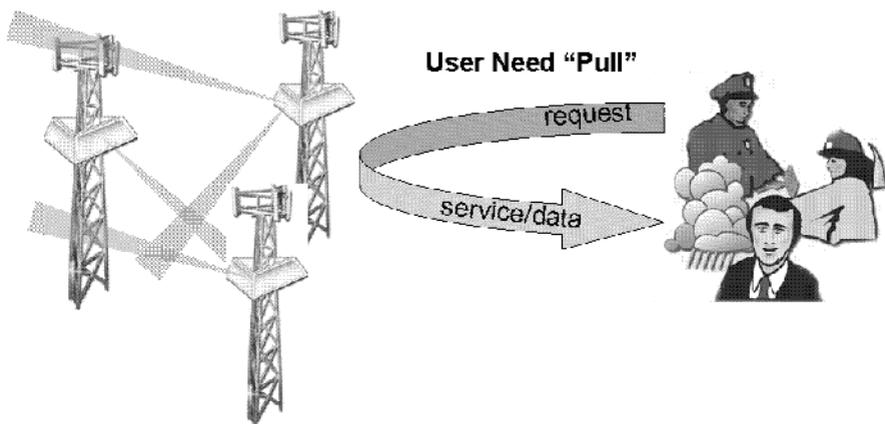
Fig. 1. Distributed Collaborative Adaptive Sensing System (DCAS) Driven by End-User Data Needs
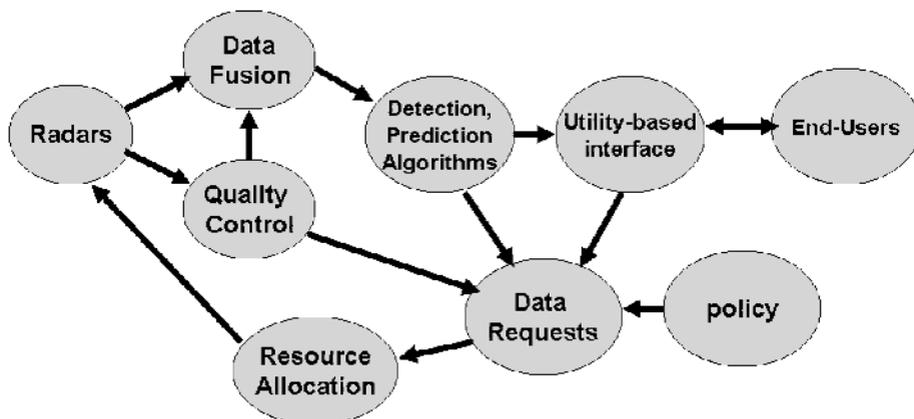
Fig. 2. DCAS System Functional View

2   DCAS Data Handling Architecture

High-level functional views of the NetRad architecture are shown in Figures 2 and 3. Users, either directly, or through meteorological algorithms, request stored and live radar data from the system. Competing resource requests (e.g., for radar beam scheduling) are mediated by the resource control algorithms in order to optimize overall utility. Data is returned from the radars through the data store, and to the requesting algorithms and users.

Figure 3 shows that data is at the very heart of the NetRad system. A **distributed data storage** facility is currently being designed that supports *(i)* querying, *(ii)* reading, and *(iii)* writing of data. The data store acts as an intermediary between data producers and consumers. The data producers include the radars and QC algorithms
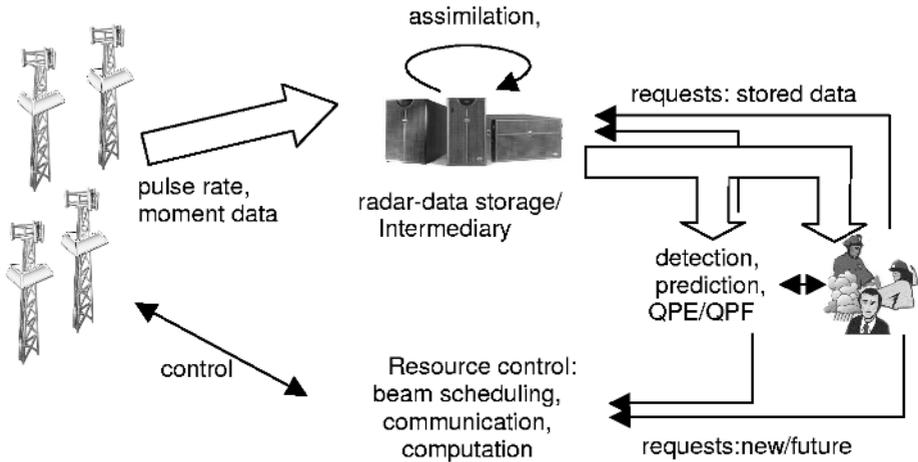
Fig. 3. Overview of NetRad architecture

that create QC-enhanced data; the data consumers include end-users who want direct access to retrieved data, hazardous weather detection and prediction algorithms, quantitative precipitation estimation and prediction algorithms, and quality control algorithms that read data in order to create new QC-enhanced data. The use of a data storage facility to provide a level of indirection [4] (i.e., act as an intermediary) between data producers and data consumers provides many benefits in terms of system structure. For example, with this level of indirection, it is easy for multiple data consumers to receive the same data, and to dynamically add new data consumers for the data stream. Similarly, a late arriving request for data that is currently being streamed to an existing consumer can be satisfied by adding the new consumer to the ongoing multicast data stream, and initiating a second patching stream that allows the new consumer to receive the missed data.

The interface to the data store will provide for read, write, and query [1] operations.

- **Data store write interface.** Radar observations are written into the data store. Similarly outputs from the QC (Quality Control) algorithms (e.g., to perform dealiasing) are written into the data store. Each stored data set will have a data provenance indicating the data source characteristics, and a history of all processing that has been performed to produce the data set.
- **Data store query interface**. The data store will provide a query interface over the globally stored data. An individual query will return a "handle" to data matching the query, as well as the provenance information stored with the data. The handle can then be used directly to read the data, if desired. Two basic types of queries that will be supported. The first type of query specifies a spatial region and a time interval; a handle to all matched data is returned. The second type of query is one that specifies a condition on the *value of data*; a handle to all matched data is again returned.
- **Data store read interface.** A handle is presented to the data store to read data associated with the query. The data store provides a pipe abstraction that allows data to be read as they are being written. This allows a data consumer

to receive data in real-time, i.e., without waiting for it to first be stored in the data store.

We note that while the actual storage location of data may be contained in the data pedigree, data readers and writers will typically read/write/query data in a storage-location-blind manner.

3   Per-radar Data Rates

In this section we describe the radar data being produced, as well as the overall rates associated with this data. A NetRad radar operating in "surveillance mode" (surveying the volume of atmosphere above the radar and out to a radius of 30 km) will sense roughly over an area with a radius of 30km, to a height of 3 km. This sensed volume is divided into unit volumes (referred to as *voxels*) of approximately 500m by 500m by 100m. These dimensions are set by the combination of the radar's two-degree antenna beamwidth and the 1 uSEC width of the transmitted pulse. A NetRad radar will thus sample approximately 350K voxels every 30 seconds.

The following six quantities are estimated per voxel every 30 seconds. These values are often referred to as ***moment data*** (referring to moments of the Doppler Spectrum of the received signals), as they are averaged over a number of radar pulses transmitted in a given beam position.

- **Reflectivity (Z):** This is a measure of the amount of the backscattered signal returned to a radar due to scatterers (e.g., raindrops and hail) in the voxel being sampled. It is proportional to the volumetric radar cross section of the observation volume and proportional to the sixth power of the diameter of the ensemble of hydrometeors in the observation volume.
- **Mean Doppler Velocity (V).** This is the mean of the Doppler velocity spectrum, indicating the average radial velocity of scatterers in the voxel.
- **Doppler Spectrum Width (W).** This is a measure of the spectral spread and often approximated by the standard deviation of a Gaussian shaped model for the Doppler spectrum.
- **Differential reflectivity (Z_{dr}).** NetRad will be deploying polarimetric radars, transmitting a signal that has equal powers at horizontal and vertical states Z_{dr} is the ratio of the power returned by the horizontally and vertically polarized pulses and is a measure of the deviation from a spherical shape, of the scatterers.
- **Correlation coefficient (ρ_{hv}).** This is a measure of the correlation between the horizontally and vertically polarized returns.
-   **Differential phase (Φ_{DP}).** This is a measure of the difference in phase between the horizontally and vertically polarized returns, dominated by the propagation phase difference between two polarization states.

Given that the 350K voxels each produce six four-byte moment data (Z,V,W, Z_{dr}, ρ_{hv},Φ_{DP}) every 30 seconds, the moment data rate is slightly over 2 Mbps. A small amount of additional overhead (including timing, radar operating parameters, and packetization overhead) will increase this value slightly.

Certain meteorological algorithms will require "raw" (unaveraged) radar data – per-pulse data corresponding to the received signals (amplitude and phase in the horizontal and vertical polarization directions, for 8 bytes total per point) received in response to each of the N pulses for a given voxel. The data rate for this ***pulse-rate data*** is thus approximately 100 Mbps.

We note that these requirements are for uncompressed data. Studies of compression of WSR-88D NEXRAD moment data (which is of much coarser scale than NetRad data) indicates that compression ratios between 1:6 and 1:10 can be achieved [3]. The compression ratios of NetRad moment and pulse data, which result from sensing the atmosphere at a much finer scale than NEXRAD, remain an open question.

In summary then, the per-radar data rates are approximately 2 Mbps for moment data, and 100 Mbps for pulse rate data. Each radar in the network will contribute this amount of data.

4   Adaptive Sensing through Data Mining

Once the moment data have been collected in the distributed data storage facility, the data are ready to be 'mined'. Data mining describes the process by which observations and/or gridded analyses/forecasts are interrogated to extract useful information.

A number of algorithms are included within the NetRad architecture for the detection of hazardous weather phenomena. Which algorithm is applied and when is dependent upon the scanning mode of the radar. Five scanning modes are established: Severe Storm Anticipation, Tornado Acquisition, Tornado Pinpointing, Quantitative Precipitation Estimation, and Data Assimilation and Prediction. Thus, through data mining processes, the radar scanning strategy adapts based upon the development of hazardous weather events.

During the *Severe Storm Anticipation* mode, the radar scanning strategy consists of general survey volume scans in search of any hazardous weather development. A Storm-Cell algorithm determines storm-cell boundaries and other radar observed thunderstorm characteristics. A Boundary Identification algorithm determines discontinuities that may trigger further storm growth and/or lead to tornadogenesis. Other algorithms search for shear and mid-level convergence as possible precursors for tornado development.

If a tornado precursor is identified, one or more nearby radars will switch to *Tornado Acquisition* mode. During this mode, the radar(s) will only focus on the areas of interest while collaborative radars continue to survey for general thunderstorm development. Meanwhile, several additional detection algorithms now begin to search for actual tornado signatures. Reflectivity and vorticity features, such as hook echoes and regions of circulation, are determined. From these features, once a tornado is identified, then *Tornado Pinpointing* mode is activated.

Tornado Pinpointing mode takes advantage of the collaborative radar system by triangulation of the beams to isolate the location of a tornado to within 100 m. Tornado triangulation merges data from multiple nodes to pinpoint the exact location (and possible trajectory) of the tornado(s).

Heavy rainfall often is an overlooked hazard in thunderstorms. Once storm cells are detected, the *Quantitative Precipitation Estimation* (QPE) mode is activated.

Scanning strategies will commence that allow rainfall fields to be mapped in real-time. The *Data Assimilation and Prediction* mode is also activated as storm cells are identified.

Assimilation of the NetRad data into a gridded 3-D volume analysis provides a critical advantage for improved detection and forecast performance: a means to combine NetRad data with external data sources such as NEXRAD Doppler radar, geostationary satellite, surface and upper-air observations, and operational numerical analysis and prediction grids. Assimilation describes the process by which atmospheric observations are combined with prior forecast information to yield a physically complete, dynamically consistent 3-dimensional state of the atmosphere, and includes such techniques as 3D- and 4D-variational analysis, ensemble Kalman filtering, single-Doppler velocity retrieval, and simple adjoints.

Such real-time assimilation of NetRad data provides the analyses for forecasts of hydrometeors, 3-D winds, and thermodynamic quantities, on spatial scales of order 1 km, for use in nowcasting and forecasting severe storm threats including but not limited to strong surface winds, hail, heavy precipitation, and tornadic potential. Nowcasting involves the use of meteorological information to generate 0-1 hour forecasts which rely heavily upon statistical techniques. One- to three-hour forecasts are made using much more advanced numerical weather prediction models, such as the Advanced Regional Prediction System (ARPS) and the Weather Research Forecast (WRF) models.

As a tornado or other important weather feature is identified, the corresponding coordinates (latitude and longitude) are sent to the **Distributed Resource Control Component** (DRCC). The DRCC converts the tornado coordinates into the appropriate polar coordinates (elevation and azimuth). The DRCC then adjusts scan strategies among the nodes. The exact scan strategy is based, in part, on the prioritization needs as defined by the End Users. These scan strategy commands are then sent back to the individual nodes.

5   Summary

The DCAS paradigm described in this paper is a dynamic data-driven system that samples the atmosphere when and where end-user needs are greatest. The concept has the potential to dramatically improve our ability to observe the lower troposphere, and to vastly improve our ability to observe, understand, and predict severe storms, floods, and other atmospheric and airborne hazards. Many knowledge and technology barriers need to be overcome prior to creating practicable DCAS systems, however. Among these barriers are the need to define an architecture and an appropriate interface to end-users, the lack of small, low-cost solid-state radars, the need to define optimization approaches and policies to adaptively allocate resources in a quantifiably optimal fashion. The NetRad system described at high level in this paper is but the first of a series of system-level test-beds that will be created by the CASA Engineering Research Center to explore the DCAS design space and instantiate the concept in the field.

References

1. J. Annis, Y. Zhao, J. Voeckler, M. Wilde, S. Kent, I. Foster, "Applying Chimera Virtual Data Concepts to Cluster Finding in the Sloan Sky Survey." *Supercomputing 2002* (SC2002), November 2002.
2. Center for Collaborative Adaptive Sensing of the Atmosphere, www.casa.umass.edu
3. S.D. Smith, K. Kelleher, Lakshmivarahan, "Compression of NEXRAD (WSR-88D) Radar Data Using Burrows-Wheeler Algorithm," *Proc.18th IIPS,* Orlando FL, Amer. Meteor. Soc., 133-135.
4. I. Stoica, D. Adkins, S. Zhuang, S. Shenker, S. Surana, "Internet Indirection Infrastructure," *Proc. ACM Sigcomm 2002.*