Proc. IEEE Globecom Workshop on Radar and Sonar Networks, Anaheim, CA, Dec. 2012.

Radar Networking in Collaborative Adaptive Sensing of Atmosphere: State of the Art and Research Challenges

H. M. N. Dilum Bandara and Anura P. Jayasumana
Department of Electrical and Computer Engineering
Colorado State University
Fort Collins, CO 80523, USA.
{dilumb, anura}@engr.ColoState.edu

Michael Zink
Department of Electrical and Computer Engineering,
University of Massachusetts,
Amherst, MA, USA.
zink@cs.umass.edu

***Abstract*— Collaborative Adaptive Sensing of the Atmosphere (CASA), in which a network of small weather radars collaborates in real time, has emerged as the preferred alternative for accurately detecting and tracking hazardous weather phenomena such as tornados. We address data transport and multi-sensor data fusion solutions proposed in context of CASA, focusing on their ability to support streaming and data pull modes of radar data access. How CASA can leverage emerging technologies, such as Named Data Networking (NDN), virtualization, and cloud computing to meet the performance, quality of service, and resource utilization requirements in a scalable manner are addressed. Open issues and research challenges for realizing the true potential of large-scale CASA networks are considered.**

Keywords*— *multi-sensor data fusion; named data; overlay networks; radar networks; streaming

I. INTRODUCTION

Emerging Distributed Collaborative Adaptive Sensing (DCAS) systems [1] sense the physical world at a far greater spatial and temporal resolution that has not been hitherto possible. DCAS systems rely on a multitude of heterogeneous and distributed sensors, ranging from mote-based, resource-limited, low-power, and task-specific wireless sensor nodes to resource-rich, high-power, and multipurpose sensors such as radars. Data generated by these sensors are processed using distributed groups of computational, storage, and bandwidth resources [2]. A key defining characteristic of DCAS systems is *data pull* where end-user information needs determine how and what group(s) of system resources are utilized to generate and process the required data [3]. Collaborative Adaptive Sensing of the Atmosphere (CASA) [1-3] is a DCAS system based on a dense network of low-cost weather radars that collaborate in real time to detect and track hazardous, localized weather phenomena such as tornados and severe storms.

Because of its flexibility and global accessibility, the Internet has emerged as the preferred medium of communication among sensors, actuators, processing nodes, and end-users in DCAS systems such as CASA and arrays of radio telescopes that try to emulate an Earth-size telescope [4]. A networking researcher may be inclined to abstract these systems as having large amounts of bits that need to be moved from one location to another hence not to be any different from well-established research domains such as bulk transfers and streaming. While this abstraction is valid in a limited sense, better performance at application level and Quality of Service (QoS) for data transfer can be gained by paying attention to the specific characteristics of sensor data and their relevance to end users. For example, though radar applications cannot tolerate burst losses, they can tolerate much higher packet losses compared to streaming video, if most relevant packets for a given application are delivered. Therefore, while developing networking solutions, it is important to take into account the specific characteristics of data and its relevance to end users.

We review data transport and multi-sensor Data Fusion (DF) solutions proposed in the context of CASA, which is a multi-sensor, multi-application, and multi-user DCAS system. High data rates (ranging from Mbps to Gbps), large volumes of data, real-time data transmission and processing requirements, and collaboration among large heterogeneous groups of sensors, applications, and end users not only differentiate CASA from conventional sensor networks but also make it more challenging to achieve the desired performance, QoS, and scalability. CASA supports three modes of data delivery: streaming, pull, and archive. We discuss state-of-the-art solutions that are proposed and currently being used in these modes of operation. CASA also leverages several other sensors and emerging technologies such as Named Data Networking (NDN), virtualization, and cloud computing to enhance the detection and forecasting accuracy, performance, QoS, and resource utilization. Preliminary work related to these attempts is also presented. We conclude with a discussion on research opportunities and challenges to realize future large-scale radar networks.

Section II presents a detailed background on CASA. One-to-one and one-to-many radar data streaming solutions and estimation of multi-sensor DF latency are presented in Section III. Section IV discusses proposed Peer-to-Peer (P2P) and cloud-based approaches for multi-sensor DF. Pull-based data access mechanisms using sensor and data dependent content names are discussed in Section V. Research challenges and opportunities are presented in Section VI.

II. BACKGROUND

Current highly-expensive, widely-spaced, and long-range weather radars are unable to sense the lower 3 km of the atmosphere due to the Earth's curvature and terrain blockage [1]. Ability to sense the lower 3 km of the atmosphere is important

This research is supported in part by the Engineering Research Center program of the National Science Foundation under award number 0313747.

for accurate detection and forecasting of localized weather events such as tornados and flash floods. Central to the CASA research effort is the use of large numbers of low-cost small radars, spaced close enough to see the lower 3 km of the atmosphere in spite of Earth's curvature and to avoid resolution degradation caused by radar beam spreading [1]. Such a dense network of radars enables the same weather event to be sensed from multiple angles consequently increasing the accuracy of sensing, detection, and prediction. CASA radars and processing nodes running DF algorithms communicate with each other to adjust their sensing, data processing, and communication strategies in direct response to the evolving weather and to changing end user needs [1, 3]. CASA IP1 test bed in Oklahoma consisted of four radars placed on a rhombus with inter-node spacing of 30 km. IP1 is currently being relocated to Dallas, TX and will be expanded into an eight-radar network. IP1 radars covered an area of ~7,000 km^2 using a transmission range of 40 km and were connected to the Internet. IP1 is controlled through Meteorological Command and Control (MC&C) which closes the loop between sensing and radar tasking [1]. MC&C ingests data from radars, identifies meteorological features in data, reports features to end users, and determines future scan strategies of radars based on the detected weather features and users' information needs. To satisfy CASA's goal of detecting severe weather events within 60 s, a closed loop is executed every 30 s. A potential nationwide CASA radar network deployment in the U.S. is estimated to require 10,000 radars [1].

CASA supports a diverse set of meteorological algorithms (referred to as *applications*) and end users [3, 5]. Each application pulls one or more types of data from one or more radars. For example, radar images that we see on TV newscasts are drawn using reflectivity data from clouds that are typically generated by a single radar. More accurate reflectivity images can be generated using the Network-Based Reflectivity Retrieval (NBRR) algorithm that pulls reflectivity data from three or more radars that sense the same region in atmosphere within an acceptable time window [2]. Both Doppler velocity and reflectivity data from two to three radars are needed to estimate the wind velocity accurately. Same data are used for tornado-tracking and low-flyer surveillance [6] applications. Therefore, the same data type may be accessed by multiple applications. These applications are utilized by a diverse set of end users such as the National Weather Service (NWS), Emergency Managers (EMs), scientists, media, and commercial entities. Users may issue queries periodically for surveillance or when an interesting weather event is detected within their Area Of Interest (AOI). For example, a NWS forecast office sends a separate query for each of the applications of interest for counties under their jurisdiction. For surveillance purposes, they may pull data from reflectivity and velocity applications every five minutes. However, when an active weather event is detected, reflectivity, velocity, NBRR, nowcasting, and Quantitative Precipitation Estimation (QPE) applications are also queried at a higher sampling rate. A researcher trying to understand the physical properties of a tornado may use velocity and tornado-tracking applications every 30 s to acquire frequent samples.

While a dense network of radars can substantially improve the detection, forecasting, and warning time, it creates many challenges due to the sheer number of sensors involved, the heterogeneous network and communication infrastructure, and the volume of data generated. IP1 radars generate raw data at rates up to 800 Mbps, which reduces to 3.3 Mbps with preprocessing. In some cases, e.g., to preserve the accuracy or for archiving purposes, it is preferable to transfer raw data. The next generation of solid-state CASA radars is expected to generate raw data at several Gbps. Even though the system is mission critical, CASA uses the Internet as the preferred medium of communication because of its flexibility, global accessibility, and low cost. Therefore, new transport and multi-sensor DF solutions are needed to timely transmit and process large volumes of radar data while overcoming dynamic network conditions. These solutions need to support streaming, pull, and archive modes of data delivery. Streaming is typically used by surveillance applications. In the pull mode, end users issue queries to pull the desired data for their AOIs. Archive mode is used to access raw data archived at radar sites or MC&C. Due to the lack of bandwidth, raw data are currently stored at radar sites and physically transferred to researches using hard disks. With the expansion of bandwidth, we believe that it will also become feasible to transfer raw data through the Internet using data streaming and pull modes.

III. RADAR DATA STREAMING

Radars stream the generated time-series data to DF applications that detect meteorological features and generate forecasts. Two solutions for transferring data to fusion applications or end users are discussed first. Then an analytical model for multi-sensor DF latency under bursty cross-traffic is presented.

A. *TCP-Friendly Rate Adaptation Based on Loss (TRABOL) [7]*

It is desirable to stream all the relevant radar data to the DF node or the end user as it increases the utility of meteorological applications. However, when the bandwidth is scarce, the utility of an application can be still maintained at an acceptable level by transferring the most relevant data. The rate at which all the relevant data are transferred is called the *target rate* while the rate required to transfer the most relevant data is called the *minimum rate*. TCP is unable to meet the minimum rate requirement of high-bandwidth raw-data streams. Alternatively, UDP does meet the minimum rate but is prone to random packet losses that are not desirable in radar DF. TRABOL is an application-layer, UDP-based end-to-end transport protocol that adapts the sender's transmission strategy based on the available bandwidth and application requirements [7]. When the packet-loss rate at the receiver exceeds a given threshold, it informs the sender. The sender then aggressively reduces its transmission rate and with time tries to gradually increase it until another packet-loss message arrives from the receiver or target rate is reached. It has been shown that an Additive Increase and Multiplicative Decrease (AIMD) rate control policy similar to that in TCP is desirable for TRABOL. It also makes TRABOL TCP friendly. Once the rate is decreased, the sender uses the knowledge about the application receiving the radar data to prioritize packets to be sent at the new rate. Suppose the sender was streaming at rate r. Once a negative acknowledgement is received, rate is reduced to $r/2$. Rather than dropping ½ of the packets randomly, the sender uses the application information to decide what packets to transmit. While radars take multiple samples of a given location in the atmosphere, not all samples are equally important. For example, if an application is interested in reflectivity, it is sufficient to send intermediate samples. However, adjacent samples are needed to calculate the wind velocity and next two samples may be skipped. TRABOL utilizes such application-specific sample selection schemes to enhance the utility of the delivered data. However, the rate does not drop below the minimum rate hoping most of the packets will go through. Simulation-based analysis showed that TRABOL is more effective than TCP and UDP, and data quality (measured as the increase in standard deviation (STD) of received data) reduced by only 10% when packet loss is 50%.

B. *Application-Aware Overlay Networks (AWON)[8]*

Multiple applications and end users often tend to access the same data streamed from a radar. TRABOL is not suitable for such multicasting scenarios as it is designed to be end-to-end. Moreover, multicasting is further complicated when available bandwidths and desired data rates of end users are different. For example, NWS users may prefer to receive all the packets while an EM (using a mobile device) deployed in the field may prefer to receive only the most important packets. Given the lack of network-level support for multicasting and application-aware data selection, overlay networks are a viable alternative for CASA-like large networks of sensors and processing nodes.

Application-Aware Overlay Networks (AWON) [8] is an architecture for deploying application-aware services in overlay networks. Application-specific Service Plug-ins (ASPs) are used to inject application-specific functionality into the overlay nodes. AWON allows applications to regulate the flow of data through the overlay nodes by extracting, selecting, fusing, and repacking data while considering application-specific constraints. For example, an ASP may be deployed to filter packets in an application-aware manner when the bandwidth of intermediate overlay links is insufficient to stream all the packets from a radar or a user is interested in receiving data at a lower rate. For example, while node *b* in Fig. 1 receives four packets from the radar, the link between *b* and *c* can accommodate only two packets. Rather than dropping two random packets at *b*, an ASP can be deployed at *b* to drop packets in an application-aware manner. Radars can mark packets based on their importance for applications under different transfer rates. For example, packet p_1 marked with r_1 in Fig. 1 indicates that the packet is important for any application that can receive data at rate r_1. Similarly, p_1 is also important for applications that can receive data at r_2, r_3, and r_4 ($r_1 > r_2 > r_3 > r_4$). Assuming radar is streaming reflectivity data, p_2 is not essential for a user that can receive data only at r_2. Hence, p_2 is marked as only being essential for a user with r_1. As alternate samples are important to calculate reflectivity, p_3 is marked as both r_1 and r_2. Once the packets arrive at *b*, it can drop p_2 and instead forward p_3 while increasing the utility of the delivered data. All four packets are forwarded to *d* as it can receive data at r_1. ASPs may be also used to fuse data from multiple radars within the network [8]. Experimentations on PlanetLab show that AWON can improve the quality of data delivered to the end users under varying network conditions. For example, Fig. 2 shows that the STD of received reflectivity data increased by only 5-15% even when the intermediate link capacity is reduced by 60%.

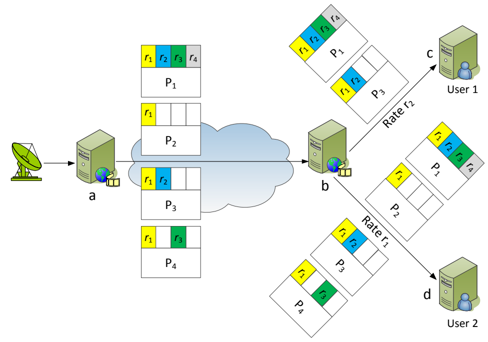

Figure 1. Application-aware packet marking and dynamic data selection.

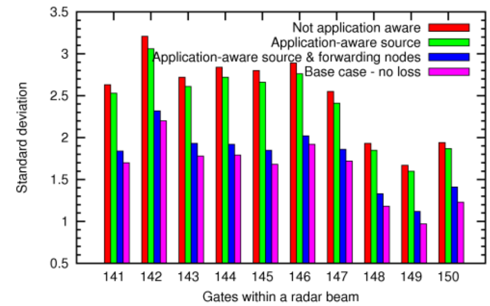

Figure 2. Quality of data delivered using AWON [8].

C. *Modeling Multi-Sensor Data Fusion Latency[9]*

Applications such as Doppler velocity and NBRR need to aggregate data from multiple radars. These applications need all the data of overlapping radar beams before DF can begin. However, due to the variability in transmission and queueing delays on different links, data from different radars reach the processing node at different times. Thus, the processing node needs an extra data synchronization time before the fusion processing can begin. Therefore, overall DF delay is dependent on both the amount of data to be processed and self-similar network traffic conditions. Given the real-time nature of CASA, it is important to design the network to satisfy the given deadlines, which are typically defined as a set of probabilistic bounds. Lee et al. [9], presented an analytical model to predict the DF latency in Internet-based multi-sensor and multi-hop DF systems. The model takes into account the periodic nature of sensor data generation and burstiness in cross traffic, and proposes a novel technique to estimate the periodic backlog created at the queues of intermediate nodes probabilistically. Information about the backlog and link utilization characteristics is then used to derive the DF latency. A comparison of the model and simulation results confirms that the model is capable of estimating the multi-sensor, DF latency accurately (with relative error ≤ 10%). The proposed model can be used to provision network bandwidth while deploying DCAS systems and developing data synchronization strategies for distributed DF.

IV. Distributed Multi-Sensor Data Fusion

Distributed and collaborative DF provides an attractive implementation choice for CASA because data are constantly being generated, processed, pushed, and pulled among groups of sensors, storage, and processing nodes. First, two multi-sensor DF strategies based on P2P and cloud computing are presented. We then discuss how CASA is leveraging other sensors to enhance the accuracy of detection and prediction.

A. *Peer-to-Peer Collaboration Framework for Data Fusion [2]*

A distributed framework in the form of a collaborative P2P system is an attractive alternative paradigm for DCAS systems. P2P architectures enable applications to aggregate and utilize unused resources from anywhere in the system to achieve better performance and QoS. P2P systems alleviate single points of failure and are robust under random node failures, and therefore, suitable for distributed DF under hostile conditions. A P2P collaboration solution for CASA multi-sensor DF algorithms is presented in [2]. Fig. 3 illustrates the multi-radar DF procedure in which the data acquired or received from remote radars are used as supplementary information for correcting sensing errors in the data collected by the local radar. Radar *a* requires its local data, as well as data from radars *b* and *c* to correct its sensing errors such as those due to the attenuation of the radar signal (e.g., in NBRR). Similarly, node *b* needs the data collected by *a* and *c*. As illustrated in Fig. 3, multi-radar DF algorithms generally consist of two steps, *preprocessing* and *DF processing*. In the first step, the collected raw radar measurements are preprocessed, e.g., sub-sampled, to meet the particular requirements of the application. Due to the large volume of data collected by each of the radars, even this simple quality control of data takes considerable computations and thus time. In step two, the algorithm detects the significant features by integrating the preprocessed data.

A radar needs to know which subset of its neighboring radars will generate the data for the same volume in atmosphere and at what time. For example, once *a* generates its own data, it needs to determine whether recent data from radars *b* and *c* are already available. If so, *a* may pull each data item from *b* and *c* separately or both items from one of them. If *b* recently used *c*'s data for processing, it may still have a copy of that data. If *c* has only the raw data (which need extra time for preprocessing) or its link is congested, it may be better to pull both data items from *b* as it reduces the latency. Such latency-aware data selection is achieved using a dynamic peer selection protocol, namely Best Peer Selection (BPS). BPS allows the multi-sensor DF algorithms to obtain the desired data in a timely, efficient, scalable, and distributed manner. Upon receiving a query, a BPS-enabled peer estimates the time required to provide the desired data considering its computation and communication overhead, and responds to the query-initiator with its estimate. Based on the responses, BPS finds a subset of the peers that can collectively provide the given set of data within the time constraint. If the desired data are not already generated by *b* or *c*, *a* has to wait until they are generated. A P2P-based publisher/subscriber mechanism is provided to enable radars to subscribe to other overlapping radars indicating their interest for a particular data item. Once the radar generates the particular data item, a notification is sent to all the subscribers. After receiving a notification, the radar uses BPS to determine a suitable subset of peers to download the data while minimizing latency. Fig. 4 shows the distribution of DF time with 40% network cross traffic. BPS outperforms random and fixed peer selection strategies and ~80% of the time DF latency is less than 15 s.

B. *Data Fusion on Data-Intensive Clouds [10-11]*

Weather surveillance applications that are executed once in every few minutes require moderate amounts of resources. However, once a rare but severe weather event is detected large amounts of computing, storage, and bandwidth resources are needed to acquire and process relevant data. Though it is mission critical, provisioning a CASA network for such infrequent peak demands is neither economically feasible nor practical

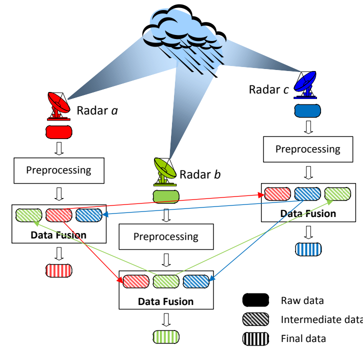

Figure 3. Radar data preprocessing and multi-radar data fusion.

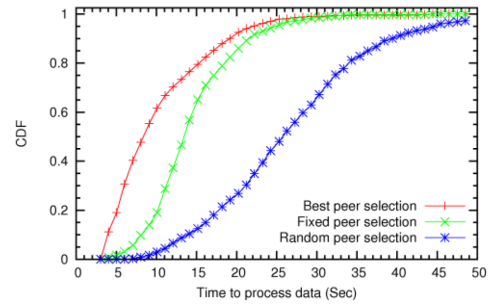

Figure 4. Data fusion time with 40% cross traffic [2].

due to the spatial and temporal locality of hazardous atmospheric events. Cloud computing provides a viable alternative where large amounts of computing, storage, and bandwidth resources can be added to the CASA network as-and-when needed. CASA, in collaboration with the GENI project, is developing a Data-Intensive Cloud (DICloud) computing environment for researchers [10-11]. The DICloud test bed can be used to conduct data-intensive experiments from start (data collection) to finish (processing, forecasting, and archiving). It allows users to specify workflows, indicating the subset of sensors to use as well as how to process and store the data generated by those sensors using computing and storage resources in the cloud. Currently, DICloud users have access to radars, weather stations, and pan-tilt-zoom cameras that are interconnected using long-range 802.11 wireless links. These sensors are developed as part of the GENI ViSE project [11], and users are granted exclusive access to these resources through virtualization. Data processing and storage services are provided using Amazon Web Service compute and storage resources.

C. *Integrating Heterogeneous Sensors to Enhance Detection [12]*

Even the CASA radars have several fundamental limitations in terms of temporal sampling rate and low-altitude blockage that can hinder their effectiveness [12]. It has been demonstrated that tornadoes and their precursors produce very low-frequency infrasound (< 20 Hz) [12]. Therefore, infrasound can be used to increase the accuracy of detection, warning-time, and localization of developing tornados because of its potential to overcome coverage limitations of radars. Once a tornado or its precursor is detected, scanning strategy of radars can be adapted to sense and track the area of developing tornados. In 2011, CASA deployed infrasound sensors at two of the IP1 sites [12-13]. Each of the sites had four sensors, as multiple sensors are required for the localization of the infrasound source. Each site also had a micro-weather station, as wind speed, pressure, and temperature corrections are need for accurate detection of infrasound sources. Sensor readings were forwarded to the MC&C through the Internet connections used by IP1 radars. Data was processed centrally and was corrected for effects from wind noise and atmospheric pressure. Infrasound sensors were able to detect one tornado in Oklahoma during the 2011 tornado season [13]. More experiments have to be carried out in upcoming tornado seasons with better wind filters as the ones used in the 2011 experiment hindered the detection under high wind conditions [13]. Automated adaptation of CASA radars based on infrasound detection will be integrated once the detection accuracy of infrasound sensors is improved.

V. PULL-BASED RADAR DATA ACCESS

In the data-pull mode, the end users query for radar data. Based on these queries, radars and associated computing facilities are tasked. Two data pull mechanisms based on radar-specific and data-specific names are discussed next.

A. Multi-User Data Sharing [14]

In the pull mode users specify an AOI for the radar to scan, an application type (e.g., NBRR), and a deadline. While most queries overlap spatially and temporally (as most users are interested in analyzing the same weather event), some queries may specify conflicting AOIs. Overlapping AOIs can be aggregated hence radars need to scan those areas only once. However, when multiple distinct AOIs are given, radars need to decide on which subset of the areas to scan within a given epoch while increasing the end-use utility. A multi-user data sharing system for a single radar is presented in [14]. End users connect to the CASA network using a proxy that issues aggregated queries to radars on behalf of the users. Each query specifies an AOI as a *sector* to be scanned by the given radar and a deadline to receive the data. Proxy aggregates queries and specifies a *weight* for each sector. Weights are determined based on the priority of users (e.g., NWS may have the highest priority) and deadlines (a query that is nearing the deadline has a higher weight). Weighted sectors are sent to the radar at each epoch enabling it to decide the best scanning strategy given the weights and mechanical properties of the radar. A wavelet-based progressive compression scheme is also proposed in [14] to encode the most relevant samples and radar beams first. It enables the delivery of most important data when the link(s) between the proxy and radar is congested. Simulation-based study using IP1 radar data show this solution to be able to improve the accuracy of queries by an order of magnitude compared to a utility-agnostic solution, and the utility reduced only by 15% when the bandwidth is dropped by a factor of 15.

B. Named Data Networking for Multi-Sensor Data Fusion [5]

End user queries in [14] specify the AOI as a sector in relation to the given radar. However, in practice, it is natural for a user to specify an AOI as a rectangular geographic area using latitudes and longitudes. As the users are not interested in the underlying mechanisms of CASA, they would expect the system to translate their AOIs into a set of radars that cover the given AOIs. Moreover, radar-specific queries hinder the possibility of benefitting from the high density of CASA radars, as queries have to be send to the given radar even through other overlapping radars may be idle. By enabling users to specify AOIs as geographic areas, CASA can use its knowledge about the underlying system to decide the best scanning strategies for different radars during times of heavy usage and partial system failures. Alternatively, radar data can be named based on their geographic location and data type. This meshes well with Named Data Networking (NDN) [15], designed to address a similar conflict in the Internet, where users value the ability to access content irrespective of its location whereas the Internet was designed to facilitate end-to-end resource access. NDN accesses contents using their names and enables in-network caching, multicasting, duplicate message suppression, enhanced security, and mobility.

In [5], we present a distributed multi-sensor DF solution for CASA implemented as an overlaid NDN network. We first developed the following naming convention:

$$/x_1/y_1/x_2/y_2/\texttt{application}/\texttt{time}$$

It enables users to specify the desired AOI (rectangular area specified using lower-left (x_1, y_1) and upper-right (x_2, y_2) corners), application *type*, and a *time* (e.g., to indicate user is looking for the most recent data). Fig. 5 illustrates how a set of interests (query equivalent of NDN) from users searching data for application type A_5 and AOI_1 is forwarded using the overlay network. First, user U_1 sends its query to proxy P_1. P_1 then creates an interest packet with a name indicating AOI_1 and A_5. Then an overlay network is used to route the interest to a node capable of processing the desired data which is near the desired AOI (this enables processing data close to the source). For example, U_1's interest packet is forwarded from P_1 to A_5, which is capable of running application A_5. Once the interest reaches A_5, the set of radars covering AOI_1 is found using the overlay network. A subscription interest packet is then broadcast to all the radars responsible for covering the given AOI. Once a subscription interest is received, each radar first checks its cache for the requested data. If matching data are not already cached, it negotiates with neighboring radars to decide what areas to scan based on the given AOI, application, and time. If a radar already has the data or decides to generate data, it then responds with a data packet indicating what data are or will be available. Once a data packet with a list of available data items arrives from a radar, the application then sends another interest packet to pull the desired data. When data for the entire AOI are received from the required number of radars, the application then processes the data and sends the aggregated data back to the user(s). Similarly, U_2's interest packet follows the path $P_2 \rightarrow A_3 \rightarrow P_1 \rightarrow A_5$. As the paths from U_1 and U_2 overlap at P_1, only one interest packet will be delivered to the node running A_5. If a second interest arrives while the pre-

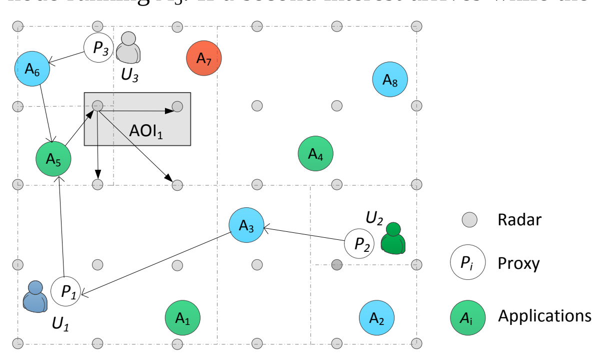

Figure 5. Named interest packet routing in a radar data fusion network.

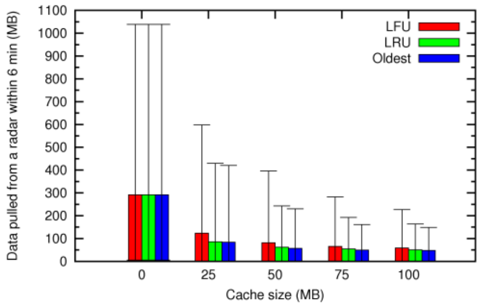

Figure 6. Data pulled from a radar under varying cache sizes and policies.

vious interest is waiting for the data, NDN enabled nodes suppress the duplicate interest. This prevents retrieval of duplicate data from processing nodes and radars. If U_2's interest arrives after the interest from U_1's was already answered, data cached at P_1 will be sent to U_2 without pulling new data from processing nodes or radars. Therefore, due to the high spatial and temporal locality in user queries, NDN-enabled CASA network can achieve substantial bandwidth saving while distributing the load across radars. Simulation-based analysis using reflectivity data from an actual weather event showed 87% reduction in average bandwidth consumption of radars (see Fig. 6) and latency in answering queries was reduced by 95%.

VI. Research Opportunities and Challenges

Distributed and collaborative DF provides an attractive implementation choice for CASA real-time weather monitoring because large volumes of data are constantly being generated, processed, pushed and pulled among groups of radars, storage, and processing nodes. While we have demonstrated the suitability of P2P-based DF for CASA, a lot more work is needed to aggregate groups of heterogeneous, distributed, dynamic, and multi-attribute resources as-and-when needed. Though many P2P-based solutions have been proposed for multi-attribute resource discovery, more work is needed on key phases of resource aggregation such as resource matching, binding, and compensation which are largely unaddressed [5]. We envision future CASA networks consisting of a mixture of radars such as IP1, solid state, long-range (still useful for detecting weak echoes and sensing over the ocean), and many special purpose ones. Collaboration and adaptation of these radars will be challenging as radars may employ different frequencies, waveforms, resolutions, and scanning strategies. Recent CASA work illustrated the benefits of integrating other sensors. Thus, future CASA networks will consist of a multitude of heterogeneous and distributed sensors to supplement the radar data and enhance radar-scanning strategies. Given the diversity of data types, data generation patterns, and processing and bandwidth requirements of these sensors, novel data-specific transport and distributed DF algorithms need to be developed. While cloud computing is a natural fit to the infrequent peak resource demands in CASA, lack of bandwidth to transfer large amounts of radar data in and out of the cloud is a major limiting factor. While it is becoming possible to aggregate and allocate computing and storage resources across multiple cloud sites, dynamically allocating high bandwidth across multiple ISPs servicing different geographic regions is not straightforward. Emerging network virtualization solutions could be useful in building virtual networks that span multiple ISPs [11]. Moreover, while cloud computing provides rapid scalability, it still takes few minutes to deploy a new virtual machine and get it up and running. This delay is significant as we may lose the ability to analyze the initial formation of a weather event once it is detected. Integration of sensors such as infrasound will not only increase the warning time but also provide a buffer time to allocate cloud resources. Exploiting new cloud-based processing strategies such as MapReduce for meteorological algorithms is also of interest. Analytical studies on quantifying the benefits of using cloud computing for weather monitoring compared to allocating dedicated resources is also useful. We are planning to perform such a study using IP1 data. NDN-based DF solution supports the most common naming convention of querying data using an application type and AOI. However, it is also important to support the radar-specific (for calibrating radars) and event-specific (to support queries such as "*find all location with hail"*) naming conventions. Perhaps the most rewarding but most challenging task is the integration of all these solutions to build a unified CASA network.

References

[1] D. McLaughlin et al., "Short-wavelength technology and the potential for distributed networks of small radar systems," Bull. Amer. Meteor. Soc., vol. 90, Dec. 2009.

[2] P. Lee, A. P. Jayasumana, H. M. N. D. Bandara, S. Lim, and V. Chandrasekar, "A peer-to-peer collaboration framework for multi-sensor data fusion," J. of Network and Computer Applications, May 2012.

[3] J. Kurose et al., "An end-user-responsive sensor network architecture for hazardous weather detection, prediction, and response," In Proc. Asian Internet Conference (AINTEC), Nov. 2006.

[4] X. Zhang and J. Li, "VLBI data transfer over fast long-distance hybrid network," In Proc. 3rd IEEE Int. Conf. on Computer Science and Information Technology (ICCSIT), July 2010.

[5] H. M. N. D. Bandara, "Enhancing collaborative peer-to-peer systems using resource aggregation and caching: A real-world, multi-attribute resource and query aware approach," PhD Dissertation, Colorado State University, Fall 2012.

[6] D. Pepyne et al., "Dense radar networks for low-flyer surveillance," In Proc. IEEE Conf. on Technologies for Homeland Security, Nov. 2011.

[7] T. Banka, A. Maroo, A. P. Jayasumana, V. Chandrasekar, N. Bharadwaj, and S. Chittibabu, "Radar networking: Considerations for data transfer protocols and network characteristics," In Proc. 21st Int. Conf. on IIPS for Meteorology, Oceanography, and Hydrology, Jan. 2005.

[8] T. Banka, P. Lee, A. P. Jayasumana, and J. Kurose, "An architecture and a programming interface for application-aware data dissemination using overlay networks," In Proc. 2nd IEEE/Create-Net/ICST COMSWARE '07, Jan. 2007.

[9] P. Lee, A. P. Jayasumana, S. Doshi, and V. Chandrasekar, "Data fusion latency in Internet-based sensor networks," In Proc. 33rd IEEE LCN '09, Oct. 2009.

[10] Data Intensive Cloud Control, Available: http://geni.cs.umass.edu/vise/dicloud.php

[11] D. Irwin, P. Shenoy, E. Cecchet, and M. Zink, "Resource management in data-intensive clouds: Opportunities and challenges," In Proc. 17th IEEE Work. on Local and Metropolitan Area Networks, May 2010.

[12] D. Pepyne et al., "An integrated radar-infrasound network for meteorological infrasound detection and analysis," In Proc. 91st American Meteorological Society Annual Meeting, Jan. 2011.

[13] D. Pepyne and S. Klaiber, "Highlights from the 2011 CASA Infrasound field experiment," In Proc. 92nd American Meteorological Society Annual Meeting, Jan. 2012.

[14] M. Li et al., "Multi-user data sharing in radar sensor networks," In Proc. 5th ACM Conf. on Embedded Networked Sensor Systems, Nov. 2007.

[15] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking named content," In Proc. CoNEXT '09, Dec. 2009.