AFRL-IF-RS-TR-2007-151
Final Technical Report
June 2007

EMPIRICAL KNOWLEDGE TRANSFER AND COLLABORATION WITH SELF-REGENERATIVE SYSTEMS

Raytheon Company

Sponsored by
Defense Advanced Research Projects Agency
DARPA Order No. T120

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the Air Force Research Laboratory Rome Research Site Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-IF-RS-TR-2007-151 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/s/ /s/

ROBERT J. VAETH
Work Unit Manager

WARREN H. DEBANY, Jr.
Technical Advisor, Information Grid Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
JUN 2007	Final	Sep 04 – Dec 06

4. TITLE AND SUBTITLE

EMPIRICAL KNOWLEDGE TRANSFER AND COLLABORATION WITH SELF-REGENERATIVE SYSTEMS

5a. CONTRACT NUMBER
FA8750-04-C-0286

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62702F

6. AUTHOR(S)

Thomas Bracewell

5d. PROJECT NUMBER
T120

5e. TASK NUMBER
00

5f. WORK UNIT NUMBER
01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Raytheon Company
1847 W. Main Rd.
Portsmouth RI 02871-1087

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington VA 22203-1714

AFRL/IFGB
525 Brooks Rd
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-IF-RS-TR-2007-151

12. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA# 07-298

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Raytheon collaborated with Cornell on the DARPA Self-Regenerative Systems program to develop new technologies supporting granular scalable redundancy. The key focus of Raytheon's effort was to research the operational hardening and application of these technologies in military information systems that can support joint combat environments and Network Centric Warfare (NCW), and to advise Cornell regarding the communications architectures and needs of major applications in these settings.

15. SUBJECT TERMS
Self-regenerative systems, technology transfer, event filter, middleware

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Robert Vaeth
U	U	U	UL	28	19b. TELEPHONE NUMBER (*Include area code*)

TABLE OF CONTENTS

1. Introduction

1.1 Overview

The main goal of DARPA's Self-Regenerative Systems (SRS) program is to develop technology for building military computing systems that provide critical functionality at all times, in spite of damage caused by errors or attacks. The SRS program addresses four key technology areas: biologically-inspired diversity, cognitive immunity and regeneration, granular scalable redundancy, and reasoning about insider threats.

Raytheon collaborated with Cornell on the DARPA Self-Regenerative Systems program to develop new technologies that support granular scalable redundancy. The key focus of Raytheon's effort was to research the operational hardening and application of these technologies in military information systems that can support joint combat environments and Network Centric Warfare (NCW), and to advise Cornell regarding the communications architectures and needs of major applications in these settings.

Our collaborative effort addressed the challenges facing military information systems which, as they become increasingly net-centric, distributed and service oriented, also need to be scalable, high-performance and survivable. In order for these systems to operate through attack, they must be intrusion tolerant – which requires redundancy. It is fundamentally difficult to support all of these needs at once. For example, scalability, high performance and redundancy give rise to the basic conflict between real-time and fault-tolerant determinism. The challenge is in developing technologies that enable us to develop systems that can better support all of these needs. The solution lies in the development of new middleware and infrastructure technologies.

1.2 Purpose

This document summarizes the findings of Raytheon's SRS effort entitled Empirical Knowledge Transfer and Collaboration with Self-Regenerative Systems. Raytheon performed this work as an associate contractor teamed with Cornell in their SRS initiative entitled Quicksilver: a Middleware for Scalable Self-Regenerative Systems.

1.3 Document Outline

Raytheon's role and approach to this effort is summarized in section 2 of this document. Section 3 summarizes Cornell's technologies and their potential applications in defense systems. Section 4 contains conclusions and recommendations for moving forward with these technologies. A novel application for the Cayuga event filter is described in Appendix A.

2. Approach

Raytheon's approach to this effort was twofold: to serve in an advisory capacity to Cornell, to help their investigators gain a better understanding of the architectures and needs of major applications in Air Force and Navy settings, and to help assess the Cornell technologies and their application in these settings.

To achieve program objectives, Raytheon applied its knowledge of military systems, combat environments and net-centric warfare scenarios to assess the strengths and limitations of Cornell's technologies, and to identify target applications for these technologies and explore technology transition.

3. Results and Discussion

3.1 Technology Overview

For the SRS program, Cornell proposed to develop Quicksilver, a technology whose component technologies support state capture, event monitoring and large-scale distribution of applications.

State capture focused on capturing the state of a platform for quick detection of changes that may require a self-repair or some other response. Event monitoring dealt with the problem of supporting complex queries in scalable high-performance systems.

The large-scale distribution component addressed the need to distribute large data objects and streams to a large number of recipients in real or near-real time. The state capture component could support basic intrusion tolerance capabilities such as graceful degradation and recovery in basically any of our candidate military systems; our focus was on investigating the application of the event monitoring and large-scale distribution components.

3.1.1 Scalable Reliable Multicast

Military datacenters use Service Oriented Architecture (SOA) to implement massively distributed applications. A scalable, reliable multicast that has timeliness guarantees are needed to enable replicated services to meet performance requirements while enabling systems to meet their scalability requirements.

Multicast is the fastest way to update a replicated service, but multicast protocols generally have limited scalability or reliability. Guarantees of timeliness or reliability can be readily achieved, but providing both guarantees is difficult. Cornell's Ricochet technology addresses this problem and offers a highly scalable, reliable multicast.

3.1.2 Scalable High-Speed Event Filtering

Publish/subscribe is a popular paradigm in which to express users' interests ("subscriptions") in certain kinds of events ("publications"). Traditional topic-based and

content-based pub/sub systems allow users to express stateless subscriptions that are evaluated over each event that arrives at the system. These systems can scale to millions of registered subscriptions at very high event rates, but have limited expressive power. However, many applications require the ability to handle stateful subscriptions that involve more than a single event, and users want to be notified as soon as one of their stateful subscriptions is satisfied. In Cayuga, however, subscriptions can span multiple events involving parameterization and aggregation, while maintaining scalability in the number of subscriptions and event rate.

3.2 Application Space

3.2.1 Intended Application Space

The target application space for Ricochet and Cayuga is distributed systems that must provide scalability, time-critical services, fault tolerance or intrusion tolerance, or any combination of these capabilities.

Raytheon has determined that Ricochet would substantially improve the performance of its Total Shipboard Computing Environment Infrastructure (TSCEI) middleware if Ricochet is integrated into the DDS standard middleware that is required by the Navy in its future naval combat systems. The same holds for other scalable time-critical military systems that could use a faster, more reliable multicast than is currently available.

The Cayuga event filter would help large distributed systems detect and diagnose failures and intrusions in real time. It would also enable system applications to detect stateful sequences of real-time events. This will make it possible to efficiently analyze data on the fly with minimal effort on the part of applications or application developers.

3.2.2 Extending the Application Space

In addition to exploring potential uses in the target application space, Raytheon has determined that these technologies have applications in other domains as well. These domains were not anticipated at the time of the proposal. The application space can be extended either by extending the technology or by discovering new ways to apply the technology. We found ways to extend both technologies and broaden their application space. For example:

The Ricochet technology could be integrated into new reliable protocols for real-time communication over very high speed links. Today's communication links have high latency and sometimes experience drop in and out; round trip times are terrible. Net-centric warfare will demand faster, more reliable wide area networks (WANs) than can be achieved with today's protocols. Ricochet could help us solve this problem.

Cayuga could be used to provide individual war fighters with timely situation awareness information relevant to their mission and current position. Information would be delivered to the warfighter via the Future Combat System's Communications and Computers network and a personal digital assistant. This information would be tailored

for the war fighter's current location using the Global Positioning System, and would include but not be limited to information to which the warfighter subscribes. A detailed description of this application is provided in Appendix A. This hypothetical and novel application was reviewed and endorsed by retired military officers who judged the application to be quite useful.

3.3 Technology Assessment

3.3.1 Ricochet

3.3.1.1 Application

The key benefit Ricochet offers is the ability to overcome communication challenges that limit the performance of GIG/NCES platforms. The GIG depends on reliable communication protocols, including publish-subscribe. Systems need low latency, dependable message delivery even when systems sustain damage or are under attack. Existing pub-sub technology does not address these challenges, and they scale poorly. This forces DoD system developers to create brittle workarounds in the application layer, or to make undesirable compromises in overall system performance.

DoD pub-sub protocols have not evolved rapidly enough to meet these challenges. For example, peer-to-peer epidemic (gossip) protocols such as Ricochet offers have never been applied in DoD systems. Existing protocols take an all-or-nothing approach to delivery. When a system is under stress, middleware is more likely to fail, compounding the likelihood that a system will be unable to perform its mission. Ricochet's probabilistic approach to delivery enables better solutions and more resiliency under stress.

Ricochet addresses these problems by providing reliable, large scale real-time message delivery using gossip protocols and forward error correction (FEC). It significantly improves delivery latency compared to state of the art protocols such as Scalable Reliable Multicast (SRM) which was developed for DARPA's Fault Tolerant Networks program.

Cornell used SRM as the baseline protocol against which to compare Ricochet. SRM is not available in products, but it is probably the most scalable protocol developed prior to SRS. SRM provides a self-repairing communication protocol that eliminates retransmission of lost packets It includes features for handling packet loss that are based on Negative AcKnowledgement (NAK). Discovery of packet loss is relative to the time between messages sent..

Baseline numbers indicate that Ricochet can achieve 2 to 3 orders of magnitude latency reduction compared to SRM as the number of groups increases. As groups increase, throughput is sustainable. When messages don't get through, message recovery time is greatly reduced. This reduces the impact of scaling and failure on system performance.

3.3.1.2 Extending the Technology

There are several ways to extend the Ricochet technology that would speed technology transition and extend its application space into new areas where it could have significant impact.

3.3.1.2.1 Integration with DDS

Raytheon has successfully applied the Data Distribution Service (DDS) message bus in its DDG-1000 TSCEI architecture. This pub-sub communication protocol is standards-based and is a Navy requirement for their Navy Open Architecture (NOA). DDS provides the overall quality of service required by DDG-1000 system services, which include time-critical services for C2, weapons control and sensor systems.

The DDG-1000 team has expressed great interest in seeing Ricochet integrated into DDS. Several factors lead us to want to advance the state of DDS. For example, DDG-1000 has failover requirements for system failures caused by faults or ship damage. These system requirements impact middleware requirements. When errors, time-critical traffic or large volumes of activity occur in DDG-1000 systems, network performance is stressed.

While customer requirements are currently being met, future applications may place more severe requirements on middleware. TSCE and NOA will be with us for a long time – at least through the deployment cycle of the entire family of ships including DDG-1000, LCS, and CVN-21. Independent of any desire to make TSCE intrusion tolerant, we will need advances in high-performance scalable reliable multicast to support future NOA and TSCEI systems.

DDS would be made more scalable, reliable and fast by integrating Ricochet with the RTI message bus (a product of Real Time Innovations, a supplier of DDS middleware). Raytheon has discussed this concept with Real Time Innovations' lead engineers and management, with Cornell, and with the DDG-1000 development team. RTI and Cornell have confirmed that such an enhancement to DDS is readily achievable and that it would be transparent to the applications that use DDS.

In addition to improving TSCEI, the benefits of successfully combining DDS and Ricochet include making available a standards-based multicast that will meet the demand that the services have for more high performance middleware.

3.3.1.2.2 High speed links for Net-Centric Warfare

Ricochet's gossip-based protocol has potential use in wide area networks. The GIG needs reliable scalable protocols that support real-time communication over very high speed links. Today's communication links have high latencies and will sometimes experience drop in and out; e.g. some links show occasional bursts of 40 Gb capacity, but also get disruptions, and round trip time can be unacceptable. Protocols such as TCP fail in such settings. If they run at all, these protocols have serious throughput and latency

properties. One might see 25Gb averaged over time, and yet only get a trickle of data through if TCP is used as is.

3.3.1.2.3 Support for Byzantine Fault Tolerance

Byzantine fault tolerance is particularly important to intrusion tolerant systems, since intrusions can at any time in any replica. Several SRS projects focused on the need to make Byzantine fault tolerance faster and more scalable.

Our team discussed the idea of integrating Ricochet and Byzantine fault tolerant protocols to create a more highly scalable, low overhead Byzantine protocol. The main obstacle to using Byzantine protocols in military systems is their high overhead and lack of scalability. Since Ricochet and Byzantine protocols operate on different levels of the middleware stack, integration should be feasible and should lead to improvements in the time required to achieve replica consistency and support failover.

3.3.2 Cayuga

3.3.2.1 Application

Intrusion detection and the ability to assess tactical situations with reusable, data-driven pattern recognition engines are two hard problems that we face in building next generation combat systems. Cornell's event filter Cayuga would enable systems to address these problems.

For example, real-time analysis of the deluge of event log entries required for information assurance auditing is currently unachievable. Without efficient scanning and processing tools, available data are essentially ignored. This greatly limits the potential for failure and intrusion detection and diagnosis during system operation. Cayuga can analyze this data on the fly.

Cayuga can assess tactical situations with data-driven patterns that are reusable, generic, and easily expressed. It can scan an application's patterns quickly enough to meet tactical engagement timelines, and it allows us to specify policies in a high-level language. Leveraging these features will allow us to write more compact applications, and make more applications reusable.

3.3.2.2 Extending the Technology

We have identified two ways to extend the power of Cayuga to support stateful XML filtering, and to handle queries that involve multiple messages.

3.3.2.2.1 XML

XML is becoming a standard way to encapsulate information. Extending Cayuga so that it can handle data that is encapsulated as XML would be a logical next step. Current XML filtering technology is stateless; it lacks the important capability to see messages in the context of information from prior messages, and thus lacks a necessary capability for combat systems. Yet SOA systems require complex message routing and policy

enforcement functionality for objects associated with XML-encoded metadata. Scalable message routing, filtering and transformation lie at the core of this infrastructure.

3.3.2.2.2 Message Relationships

Extending Cayuga to support intra- and inter-message relationships would make it possible for us to apply Cayuga in scenarios that require systems to detect and filter sequences of events. Intra-message relationships permit queries to "refer back" to data bound to variables in the same message. Inter-document relationships enable queries across messages such as "notify us when you see the same vehicle crossing the intersection twice", or "notify us when you see two identical warning messages from the dual band radar."

3.4 Target Programs

3.4.1 Military Systems

Cornell's SRS technologies will benefit a broad range of systems too numerous to list here. The application space for Ricochet and Cayuga includes distributed systems that must provide scalability, time-critical services, fault tolerance or intrusion tolerance, or any combination of these capabilities. This describes many if not most of the combat systems, intelligence systems, C4ISR systems, weapon control systems, sensor systems, missile defense systems and precision engagement systems and networks that are planned or are now in development.

3.4.2 Critical Infrastructures

While the SRS program is focused primarily on the military sector, critical US infrastructures and assets also need to be made survivable. Such assets exist in aviation, non-DoD government, banking and finance, telecommunications, public health, emergency services, postal and shipping, ecommerce, energy, rail and mass transit, pipelines, the chemical industry and water and sewer services. Cornell technologies can help us build intrusion tolerant systems in these critical infrastructures.

Many sophisticated attacks against these critical infrastructures have already occurred. Our dependence on these systems, and their vulnerability to terrorism and information warfare, warrant a large investment in advancing the state of security in these systems. Critical infrastructures need to be intrusion tolerant. This will require the adoption of new architectures and technologies such as Cornell's.

3.5 Technology Transition

3.5.1 Seeking Out Opportunities

As part of the SRS initiative, Raytheon has sought out opportunities for technology transition.

The DDG-1000 development community has shown considerable interest in the Cornell technologies. We have established dialogue with Cornell, Real Time Innovations Inc. and the DDG-1000 team and plan to continue pursuing technology transition after the SRS initiative is completed.

Demonstrating and benchmarking SRS technologies under conditions of interest to programs of record such as DDG-1000 will be an important step in technology transition. Programs of record generally cannot risk adopting new technologies with low Technology Readiness Levels until demonstration has occurred in collaboration with the research team that created the technology. Outside the SRS program, Raytheon and Cornell have submitted white papers to DARPA which identify opportunities to speed technology transition through such experiments. Success can be gradual as technologies mature; we do not expect immature technologies to meet all program needs.

Key considerations in finding the right opportunities for technology transition include:

- The technology enables significant system capabilities
- System complexity is non-trivial but not overwhelming
- Technology insertion is feasible given legacy system, platform, compiler and tool issues
- Program management supports technology transition
- Technology has tactical (task force to task element) applicability
- An evolutionary development path can be found
- The technology's Technology Readiness Level (TRL) is 4 or higher
- Adoption is well timed for technology transition

3.5.2 A Specific Opportunity

Ricochet addresses several critical DDG-1000 needs. While multicast is the fastest way to update a replicated service, current protocols have limited scalability or reliability. Ricochet offers scalability, reliability and timeliness not found in any available technology. Both DDG-1000 and the military SOA programs that the common architecture supports need these features.

The ability to manage node failures is critical to DDG-1000. Its 400-node, time-critical, fault tolerant system must promptly detect and handle node failures. COTS middleware cannot support the timeliness and scalability needed. During system startup, using COTS middleware, the false positive detection rate for node failures is 15% with only 40 nodes online and a sampling rate one quarter as fast as it needs to be. The high false positive rate has a domino effect as the system struggles to reconfigure itself. We expect Ricochet will eliminate this problem.

Time-critical service-oriented military datacenters need a scalable reliable multicast with timeliness guarantees. These datacenters must handle replicated, massively distributed services and microservices. No currently available multicast can handle the scalability, reliability and timeliness that these systems need. Ricochet addresses this problem.

4. Conclusions

4.1 Summary

Based on analysis and the results of Cornell's experiments with its SRS technologies, Raytheon has determined that the functions and performance offered by Ricochet and Cayuga offer significant advantages over current technologies in supporting the needs of current and future military programs.

The success of Cornell's effort on SRS can be seen by the continued interest in their technologies at Raytheon and on the DDG-1000 program. Technology transition will however depend on many factors such as achieving increasingly successful demonstrations of these technologies in real-world environments as the technologies mature.

The Cornell technologies bring us closer to making intrusion tolerant systems practical. They are among the many technologies and methods that will be needed in order to produce intrusion tolerant systems.

4.2 Recommendations

Technology transition should be pursued through further collaborations among Cornell, defense contractors including Raytheon, middleware suppliers, military programs and defense agencies including DARPA. Opportunistic experimental evaluation of the SRS technologies will show how the technologies will enable significant new capabilities or dramatically improve the performance of defense systems.

Technology extensions are recommended to help broaden the technologies' application space and to speed technology transition. Integrating Ricochet into DDS will certainly speed that transition, as will finding new application space for both Ricochet and Cayuga. Extending the technologies will require further research and collaboration among research teams (e.g. to support more scalable Byzantine fault tolerance) and between researchers and industry (e.g .to support integration of Ricochet into DDS). Finding new application space will require us to broaden technology awareness among potential users, through presentations and publications such as those which were given or produced in the course of our research.

5. Terms and Acronyms

C4ISR	Command, Control, Communication, Computer, Information, Surveillance ans Reconnaissance
CVN-21	The Navy's next-generation aircraft carrier program
DDG-1000	The Navy's next-generation destroyer program
DDS	Data Distribution Service
GIG	Global Information Grid
LCS	Littoral Combat Ship
NCES	Net-Centric Enterprise Services
NOA	Navy Open Architecture
RTI	Real Time Innovations Inc.
SRM	Scalable Reliable Multicast (developed for DARPA FTN program)
TCP	Transmission Control Protocol
TSCE	Total Shipboard Computing Environment. The distributed computing system which will be used in the Navy's next generation Family of Ships (FOS) including DDG-1000, LCS and CVN-21.
TSCEI	Total Shipboard Computing Environment Infrastructure. The TSCE's system and software infrastructures.
XML	Extensible Markup Language

Appendix A. A Military Application for Cayuga

1. Introduction

This study describes a novel application for the Cayuga event filter which Cornell has developed for the DARPA Self-Regenerative Systems program. It shows how Cayuga could be used to deliver timely mission critical situation awareness (SA) information to an army of warfighters.

1.1 Battlefield Situation Awareness

One of the important technological improvements pursuant to the military is the ability to provide the war fighter with better information. More accurate and timely information will increase the probability that the war fighter can successfully carry out missions, as well as improve their capability for self-protection. Equally important is to provide the war fighter with as complete a situational awareness (SA) scenario as possible. This vision of improving the accuracy and immediacy of information in today's environment is difficult as many of the current C2 systems were developed in a stovepipe fashion and are not able to communicate directly. To counteract this limitation, the Air Force has commissioned the creation of the JBI, which provides the substrate for C2 subsystems to be able to integrate effectively.

In the future, information-based systems of many varieties will be JBI network compliant. Data will be posted and queried by using a publish-subscribe mechanism. The large amounts of data that will be accessible to the JBI network will be significant. It will be far larger than an analyst could possibly absorb, interpret and disseminate in real-time. An automated data processing mechanism is clearly needed to maximize the usefulness of the information.

1.2 The Warfighter's PDA

In this application, Cayuga is used to deliver timely mission critical situation awareness (SA) information to an army of warfighters. The information delivered to each warfighter is tailored to their particular needs and their current position. The information is delivered to warfighters outfitted with rugged Personal Digital Assistants (PDAs) connected to the Joint Battlespace Infosphere (JBI).

1.3 Applicable Technologies

The JBI community has named its information processing mechanism "fuselets". According to www.fuselet.org:

> Fuselets are a light-weight, special-purpose Joint Battlespace Infosphere (JBI) client program that provides value-added information processing functions that are under the control of the JBI platform. The information processing functions take existing information objects as input and manipulate them in some way to produce new information objects.

It is anticipated that the new objects created by the fuselet processing are often higher-level events or, a better knowledge of situation awareness. Fuselets create knowledge that would not have been realized from distinct "un-fused" lower-level events.

Cornell University has developed a fuselet processor called, *Cayuga*, which seamlessly fits into the JBI infrastructure of distributed-networked systems. Cayuga produces new information objects by allowing for queries that poll and interpret independent data over widely distributed network systems. This is done by filtering events and attributes of interest either immediately or over a period of time. It is a flexible, high-speed processing engine which supports parameterization as well as aggregation and is therefore pertinent to a large class of applications.

Cayuga works using the principles of *Events* and *Subscriptions,* where:

- an Event is a set of attribute - value pairs
- a Subscription is a set of predicates on event attributes

The Cayuga implementation is fast and scalable and can sort out complex events for many subscribers.

2. Operational Scenario

In the proposed application, war fighters are outfitted with a Personal Digital Assistant (PDA) that contains a Global Positioning System (GPS). The PDA is wirelessly connected to the JBI so that up-to-date positioning information is transmitted. In addition, a database stores dangerous areas in the Iraqi theatre that are considered hazardous. Perhaps they are the locations were there is a high incidence of Improvised Explosive Devices (IEDs), or areas like Fullujah that have been under the control of insurgents.

A subscription could be generated by one of the troops that could be stated as:

Query 1. *Notify me when I'm within 1 km of a dangerous area.*

The corresponding events (attribute, values) would consist of:

- The subscribing troops (Serial number, Location)
- All the hazardous areas (Name, Location)

The events would be the Serial number and Location of the troop who made the request as well as all the hazardous areas. The subscription is that Cayuga would be searching for any dangerous area within 1 km of the troop that made the query. Perhaps a platoon commander would submit a more sophisticated query:

Query 2. *Notify me when any member of my platoon is within 1 kilometer of any hazardous area.*

This query could be found to be so effective, that the chairman of the joint chiefs issues a query as follows:

Query 3. *Notify any troop world-wide that is located within 1 km of a hazardous area.*

Query 3 shows how without much difficulty it would become impossible for an analyst, or group of analysts to manually satisfy the query in a timely fashion.

Besides being quantitatively computationally rigorous, queries can become more qualitatively complex. In the Iraqi theatre, nearly 30% of the fatalities are caused by Improvised Explosive Devices (IEDs). It could be determined that in some areas, the probability of an IED explosion is much higher wherever an insurgent has been observed. This could result in the following query:

Query 4. *Notify any troop located in Iraq who is within 5 km of any location where an insurgent has been observed within the past 24 hours.*

Note that the earlier queries are memory-less – they only compare attribute values to constant values. For Query 4 the system has to remember where every insurgent had been over the past 24 hours.

Query 4 also demonstrates that as the qualitative query becomes more multifaceted, the magnitude of the processing intensity grows. As more systems are involved, the sophistication and distributed nature of the subscriptions become more complex and the number of events is increased.

The next query involves informing troops in a timely fashion so that they can consume an antidote with enough time for the agent to act. The following query involves a GPS system, chemical sensor systems and weather systems.

Query 5. *Notify any troop downwind from a chemical attack that the attack has occurred.*

Note the criticality of the timeliness response to this query. For example, the timeliness of this information may determine whether the war fighter has no choice but to evacuate immediately, or may allow him to take an antidote, hold his position and wait for updated information.

Additional distributed systems simply add to the variation of subscriptions, limited only to one's imagination:

Query 6. *Notify all troops along 5 km of the coastline when a tsunami is detected in the corresponding body of water.*

Cayuga provides an effective and very efficient mechanism for advanced queries.

Query 7. *Notify me when the distance to an insurgent has been reduced to within ½ the distance since the start of the query*.

Note that in this example, the distance parameter has not been specified and the query requires information from prior events. This act of referring to values of prior events in a multi-event query is called parameterization. The efficiency of Cayuga's parameterization implementation is one of its most powerful capabilities.

Cayuga can sift through complex events for many subscribers. It has many potential applications for providing deciphering complex C2 information for armies of warfighters. Cayuga scales much more efficiently than traditional, publisher-subscribe approaches to event queries providing an immediate situation awareness picture to each war fighter. An example follows.

3. Chemical Attack Scenario

During a major battle in Fallujah, Iraq a chemical weapon is detonated near the Al-Samarai mosque. US troops in the battlefield and surrounding areas, need immediate instructions after the detonation has occurred. The instructions are one or a combination of the following options:

- Don protective equipment and clothing
- Evacuate immediately
- Take the appropriate antidote

Some extra information may be required. For example, a war fighter may need to know where to evacuate to and which direction to head, as he would not want to run into an insurgent stronghold. Or, if an antidote is to be taken, perhaps it needs to be taken in the

future, based on when the war fighter will be immersed in the contaminated area and the length of time the chemical agent is effective.

Note that troops don't necessarily have to be in the direct path of the chemical plume when the detonation occurs. Some troops may be in transit and will shortly be in the path of the plume.

3.1 Systems and Databases

Participating in this scenario, there are three sensor systems that contain one or more unique databases.

- Chemical Detection Sensor System. This sensor detects chemical explosions and determines the current location of the plume.

- Weather Sensor System. This sensor detects and forecasts the location of the chemical plume cloud. Plume forecasting can be predicted by evaluating its current location and applying winds and wind forecasts to the plume.

- War Fighter Registry System. This system keeps track of the location of troops who are carrying a Personal Digital Assistant (PDA) that is GPS enabled. The PDA transfers the troop location over a wireless network to a distributed database.

3.1.1 Chemical Detection Sensor System Database

The Chemical Detection Sensor System database consists of three tables. The three tables and their relationship to each other are shown in Figure 1.

A. Chemical Explosion Table

This table stores an entry for each detected chemical explosion. The explosion is stored with an ID number and with the name of the chemical agent.

B. Grid

This table contains the points in four-dimensional space, which are the corner-points of a quadrangle chemical plume, an altitude component and time.

For the purposes of the Cayuga example, the four grid points represent the four corners of the plume. The $\text{Altitude}_{\text{minimum}}$ and $\text{Altitude}_{\text{maximum}}$ are the high and low points of the z-component of the plume. This representation is a straightforward way to correspond to a

three-dimensional object in space, so that one can compute whether a latitude / longitude / altitude point, like the location of a soldier, overlaps with a plume.

The joint labeled with a "1" between the Chemical Explosion table and next to the Grid table indicates that one grid table entry represents the shape of the plume for each explosion. In a more complicated scenario, the grid could consist of multiple grid-point entries, each representing the center of a small 2-dimensional space. This would allow for more sophisticate plume shapes, beyond just a quadrangle.

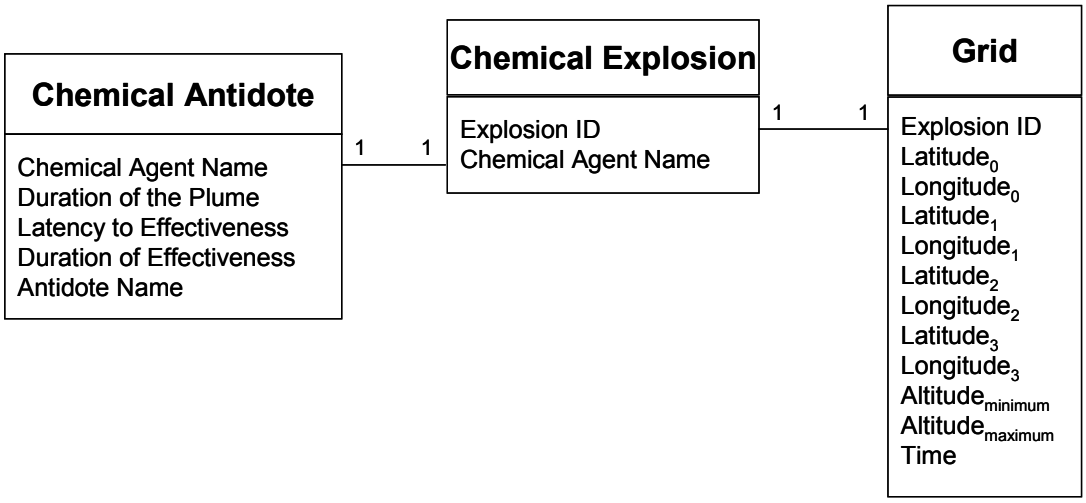

Figure 1 - Chemical Detection Sensor Relationship Diagram

C. Chemical Antidote

This table contains information corresponding to the antidote that a war fighter would ingest to counter a chemical agent. Notice, as shown in Figure 1, that there is one and only one antidote for each chemical agent. And, each antidote is only associated with one chemical agent. A more sophisticated scenario could indicate that there may be multiple antidotes for a chemical agent and/or an antidote could be effective against more than one chemical agent.

Duration of the Plume indicates how long the war fighter needs to be protected from the chemical agent.

Each chemical agent has two effectiveness attributes. The *Latency to Effectiveness* attribute is the time required for the antidote to take effect after the war fighter has ingested it. The *Duration of Effectiveness* is the duration that the antidote is effective after the *Latency to Effectiveness* period has elapsed. The timeline is graphically represented in Figure 2, which shows that only during the *Duration of Effectiveness* is the war fighter protected from the chemical agent.

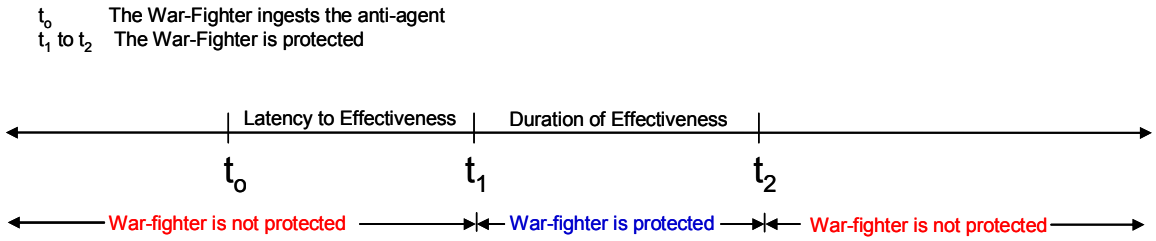

Figure 2 - Antidote Timeline

As a point of interest, note that some antidotes are effective immediately. In fact, some antidotes are considered harmful if ingested when no exposure has occurred. These antidotes are to be administered only when symptoms of chemical exposure are present.

3.1.2 Weather Sensor Databases

The Weather Sensor system has two databases. The first database consists of two tables. The two tables and their relationship to each other are shown in Figure 3.

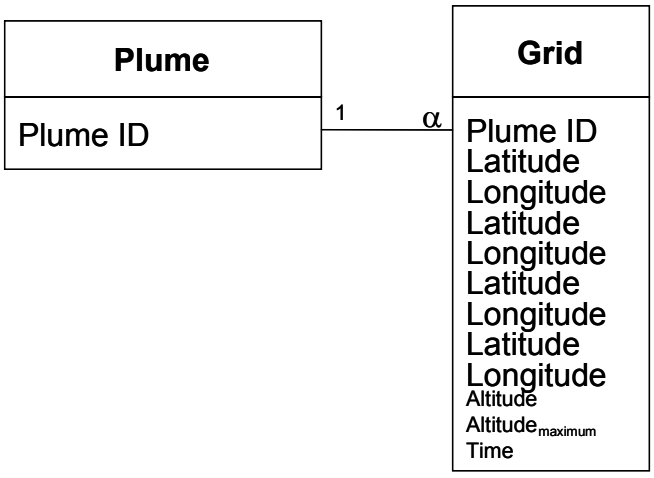

Figure 3 - Weather System Sensor Relationship Diagram, Database I

A. Plume Table

This table stores an entry for each detected plume. Each plume has an ID.

B. Grid Table

This table contains the points in four-dimensional space, which are the corner-points of a quadrangle chemical plume and an altitude component. This table is similar to the grid table in Figure 1. The similarity is that the four grid points represent the four corners of the plume and an altitude component. The difference is that there may be multiple entries based on a time component. If the time is the current time, then it represents the current location of the plume. Times in the future represent a plume forecast. Because of the forecasts there is more than one grid entry per plume. This is denoted by the "α" between the Plume table and the Grid table, which simply represents in this notation, ">=1".

The second Weather Sensor database consists of one table. It is simply a table of prevailing winds within the sensor area. The table is shown in Figure 4. The direction of a prevailing wind is measured in terms of where the air is coming from. A northerly wind blows air from north to south. A southwesterly wind blows air from the southwest to the northeast.

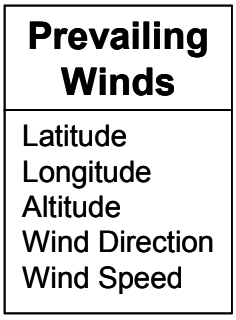

Figure 4 – Weather System Sensor, Database II

3.1.3 War Fighter Registry System Database

The War Fighter Registry System database consists of two tables. The two tables and their relationship to each other are shown in Figure 5. The database tables are:

A. Soldier Table

This table stores an entry for each war fighter. The initial information is the soldier's serial number and the list of subscriptions that are applicable to him. Over time, additional entries include what alarms the soldier has acknowledged and when that acknowledgement occurred. For example, if a soldier is instructed to take an antidote, this would be acknowledged by the press of a button on a PDA, which would also record the acknowledgement time. This acknowledgement could effect what alarms the soldier receives in the future.

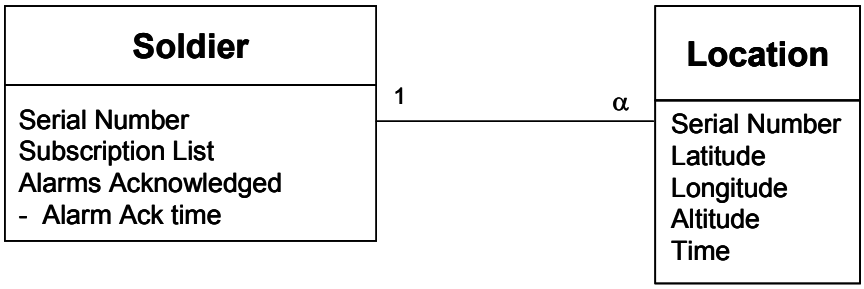

Figure 5 - War Fighter Registry Relationship Diagram

B. Future Locations Table

This table contains the specific point of the war fighter's location in four-dimensional space, which is his location in latitude and longitude with an altitude component. This table is similar to the grid table in Figure 3, in that it has multiple time entries where, based on the history of the soldier's movement a forecast can be made on where the soldier is expected in the future. If the time is the current time, then it represents the current location of the war fighter. The difference from the table in Figure 3 is that the war fighter is a point in space, containing a latitude/longitude/altitude component but not a grid.

It is anticipated that a soldier with a PDA with GPS capabilities would have line-of-sight to enough satellites to report all three positional latitude/longitude/altitude dimensions. The War Fighter Registry system would take these reports over time for each war fighter and be able to compute velocity and forecast of the war fighter position.

3.2 Situation Awareness Provided

We now consider the subscriptions that would be made and the events that would occur to mitigate a chemical attack at the Al-Samarai mosque during a major battle in Fallujah, Iraq. The C2 systems available are those that are described in Section III.

Consider that the Iraqi forces anticipated a strong US attack from the south and placed most of their troops there. In the ensuing attack, US forces quickly overrun the Iraqis in the north and begin their track southward, where most of the severe fighting is taking place. To protect the ground troops fighting in the south and those troops moving in from the north, four subscriptions have been generated and are processed by Cayuga to mitigate casualties in case of a chemical attack. These queries are described in the remainder of the section.

Query 1. *When a chemical explosion occurs, notify any soldier in the area when they are within 1 km from the center of the blast.*

With a chemical detonation occurring at the Al-Samaria mosque, all troops located within 0.5 km of the detonation would get an alarm.

The affected war fighters would get an alarm with the following information:

1. A graphic with:
    a. the location of the explosion
    b. the location of war fighter
2. The direction of the prevailing winds where the explosion occurred
3. The direction of the prevailing winds where the war fighter is located
4. The antidote for the chemical weapon and the corresponding time attributes
5. A procedural recommendation

Because of the close proximity to the explosion, the war fighter would more than likely receive a recommendation to evacuate the area.

The war fighter would then have the option to acknowledge the recommendation. Depending on the recommendation and whether or not an acknowledgement is made may impact on how Cayuga processes future events, as will be demonstrated in the next query.

Query 2. *Notify any soldier in the area if his projected position will bring him into contact with a chemical plume within the Latency to Effectiveness time*.

This query means that a war fighter's forecasted position and the forecasted position of a plume will intersect. This intersection, however, will occur too soon for the war fighter to ingest an antidote such that the antidote will be effective. This query is represented in Figure 7.

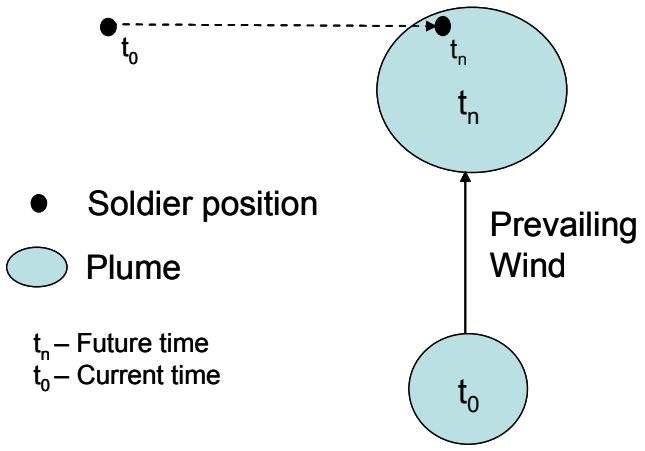

Figure 6 - Plume and Soldier Current and Projected Positions

As this case is defined as the soldier engulfed by the plume within the *Latency to Effectiveness* time, equation (1) applies:

(1) $(t_n - t_0) <$ (Latency to Effectiveness time)

The affected war fighters would get an alarm with the following information:

1. A graphic with
    a. The current location of the plume
    b. The forecasted location of the plume
    c. The current location of the war fighter
    d. The forecasted location of the war fighter
2. The direction of the prevailing winds where the explosion occurred
3. The direction of the prevailing winds where the war fighter is located
4. The antidote for the chemical weapon and the corresponding time attributes
5. A procedural recommendation

Because the war fighter does not have time to ingest the antidote and have it take effect, he would more than likely receive a recommendation to evacuate the area. Given the forecast of the plume is known, the recommendation would include a recommended direction of retreat.

The war fighter would then have the option to acknowledge the recommendation. If the recommendation is acknowledged, the war fighter would not get any further alarms unless there was some major change in the forecast that would alter the recommendation. If the war fighter did not follow the recommendation, Cayuga would assume that the

alarm was not received by the soldier and future alarms would continue to be sent. One option the soldier could have is to stop his movement and ingest the antidote and wait the recommended time. Then he could proceed forward.

Query 3. *Notify any soldier in the area if his projected position will bring him into contact with a chemical plume later than the Latency to Effectiveness time.*

This query means that a war fighter's forecasted position and the forecasted position of a plume will intersect. This intersection will occur after the war fighter has had time to ingest an antidote such that the antidote will be effective. This query is represented in Figure 7.

As this case is defined as the soldier engulfed by the plume after *Latency to Effectiveness* time, equation (2) applies:

(2) $(t_n - t_0) >=$ (Latency to Effectiveness time)

Because the war fighter does have time to ingest the antidote and have it take effect, he would more than likely receive a recommendation to take the antidote.

The war fighter would then have the option to acknowledge the recommendation and Cayuga would halt further alarms per the acknowledgement.

Affected war fighters would get an alarm with the same information as **Query 2.**

Query 4. *Notify any soldier in the area if his projected position will bring him into contact with a chemical plume but much later than the Latency to Effectiveness plus the Duration of Effectiveness time.*

This query means that a war fighter's forecasted position and the forecasted position of a plume will intersect. This intersection, however, will occur so far in the future that the antidote will lose its effect earlier than desired; perhaps before he has been engulfed by the plume.

As this case is defined as the soldier engulfed by the plume after the effect of the antidote has worn off, equation (3) applies:

(3) $(t_n - t_0) >>$ (Latency to Effectiveness time) + (Duration of Effectiveness time)

The affected war fighters would get an alarm with the same information as **Query 2** with an additional piece of information with a time in the future instructing him when to take the antidote.

3.3 Sample Database Values

This section presents some examples of database numbers for the Chemical Detection and the War Fighter Registry Systems.

A. Chemical Detection Sensor System

Table 1 lists example entries for a variety of values in the Chemical Detection Sensor Database.

Table Entry	Value	Units
Explosion ID	1	N/A
Chemical Agent Name	Agent00	N/A
Duration of Plume	8	Hours
Latency to Effectiveness	1	Hours
Duration of Effectiveness	4	Hours
Antidote Name	Antidote00	N/A
Explosion Center	Al-Samari Mosque	N/A
Explosion Altitude minimum	0	Meters
Explosion Altitude maximum	50	Meters

Table 1 - Chemical Detection Sensor Database Example Numbers

B. War Fighter Registry System

Table 2 lists the number of US soldiers and their proximity to the Al-Samarai mosque at the time of the explosion. The soldiers are assumed to be evenly dispersed.

Number of Troops	Distance from the Explosion Center
100	<= 1 km
200	> 1 km and <= 2 km
500	> 2 km and <= 3 km
4000	> 3 km and <= 20 km

Table 2 - Number of US Soldiers

Soldiers at any given moment are traveling in any given direction. However, soldiers who are north of the Al-Samarai mosque or are within 1 km south of the mosque have a propensity for heading south towards Al-Shohada.