ADVANCES IN MECHANISMS SUPPORTING DATA COLLECTION ON FUTURE FORCE NETWORKS: PRODUCT MANAGER C4ISR ON-THE-MOVE

M. J. Amabile*, J. Sypniewski, M. R. Griffis
Product Manager C4ISR On-The-Move
Fort Monmouth, NJ 07703

ABSTRACT

Product Manager C4ISR On-The-Move (PM C4ISR OTM) has a mandate to conduct integrated lab- and field-based activities for the purpose of demonstrating, evaluating and assessing emerging C4ISR technologies. The conduct of such activities has led to the development of a suite of tools and techniques for the collection of data across large, multi-component, multi-enclave system-of-systems. Within this paper, we discuss the motivation, design and implementation of a practical suite of tools employed during field exercises conducted by PM C4ISR OTM in support of the Research & Development (R&D), Acquisition, and Test communities.

1. INTRODUCTION

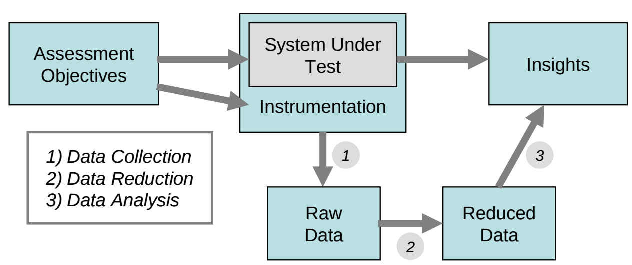
A key component within any test or assessment lies in the ability to accurately and precisely record the state of a system as it is exposed to varying conditions. The measurement of the state of a system can become increasingly challenging as the complexity of the system under test grows. This is particularly true for the tactical networks now under development by the Army and its Sister Services. The Future Force architecture calls for networks of mobile ad-hoc networks (MANETs), which require tailorable, ubiquitous and seamless connectivity between warfighting elements. Furthermore, "the network" is no longer limited to just the transport layer – rather, the general definition of the Future Force network now includes the systems and applications that employ the transport layer as a communications mechanism. In this view, the network is the Command and Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance (C4ISR) support to the Warfighter.

As the C4ISR network and its discrete parts are being developed, they need to be assessed. These assessments must be performed at the component level, at the integrated systems level and at the end-user level. Because of the complexity of the C4ISR network as an aggregate, the intricacy and maturity of its components and the varied manner in which the network needs to be assessed during its lifecycle, the requirements for quantifying the state of the network are nearly as complex as the network itself. Over the past seven years, PM C4ISR OTM has made significant advances in addressing these challenges.

As part of its mandate to provide a relevant environment for the assessment of emerging technologies in a C4ISR System-of-Systems (SoS) configuration, PM C4ISR OTM has sought to design, develop and/or adapt tools and techniques to measure the state of C4ISR networks across the range of their operation. These tools and techniques are termed the Future Force Instrumentation, Data Collection and Reduction (FF-IDCR) suite. Within this paper, we seek to define terms for instrumentation, data collection and reduction, and draw their relationship to the process of data analysis. Additionally, we describe the FF-IDCR suite in terms of its functionality and capability.

2. INSTRUMENTATION, DATA COLLECTION & REDUCTION

The terms "instrumentation", "data collection", "data reduction" and "data analysis" are used across many disciplines and have varying definitions. For the purposes of this paper, the following figure attempts to provide a framework for their definition:



In this simplified model of a test/assessment, the fundamental objectives (i.e., what is to be learned) drives the design of both the System Under Test (SUT) as well as the systems that "instrument" the SUT. Here, *instrumentation* is defined as hardware or software systems that augment the SUT in such a way as to enable monitoring of SUT components during operation. In a sense, the instrumentation systems "wrap" the SUT, and provide a harness within which the SUT components can operate such that information quantifying that operation can be acquired and recorded.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **01 DEC 2008**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **Advances In Mechanisms Supporting Data Collection On Future Force Networks: Product Manager C4isr On-The-Move**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Product Manager C4ISR On-The-Move Fort Monmouth, NJ 07703**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002187. Proceedings of the Army Science Conference (26th) Held in Orlando, Florida on 1-4 December 2008, The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**7**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

During test execution, the instrumentation systems produce data relative to SUT operation; the process of employing instrumentation systems to perform this task is termed *data collection*. Data collection also includes the process of aggregating and organizing, or *harvesting*, raw products if the SUT is distributed in nature.

The raw data products created by the instrumentation systems are generally not readily usable. The process of *data reduction* converts raw data into a more meaningful form, ranging from tables and graphs to textual summaries. It is important to note that data reduction is not data analysis. In contrast, *data analysis* requires a human subject matter expert, who is skilled in the art of the SUT, to review, compare and draw conclusions from reduced data products, and often from experience, to form insights that inform the fundamental objectives. These processes have dependencies and linkages over the lifecycle of a test event, but are distinct activities that each require time, effort and specific expertise.

3. THE FF-IDCR SUITE

The FF-IDCR suite can be categorized into several main systems:

- The Ground Truth system provides a flexible mechanism for capturing, recording, and optionally beaconing, position and navigation information from entities participating in a test.
- The Network Data Collection system acts to capture both the streams of data transiting the communications systems, as well as the operational status and configuration of network devices.
- The Player Interactions system is designed to record salient aspects of the digital displays acting as the end-user interface, as well as the inputs provided by role players during system operation.
- Finally, the Log File & Harvesting system enables the aggregation of application and system log files, as well as raw data files from the other FF-IDCR systems, into discrete archives tagged with relevant contextual information.

The following subsections describe the design & implementation approaches employed for each of these systems.

3.1 The Ground Truth System

Ground Truth, or Time-Space-Position-Information (TSPI), is collected via a suite of configurable components, optionally linked via a wireless backhaul system to one of the PM's field test centers. Four basic configurations are generally employed:

Basic Dismount	• Differential GPS device • Local archive
Beaconing Dismount	• Differential GPS device • Local archive • Wireless beacon
Basic Vehicle	• Differential GPS device • Local archive • Access to backhaul network
Relay Vehicle	• Differential GPS device • Local archive • Relay for Beaconing Dismounts • Access to backhaul network

These configurations support a range of collection requirements, spanning from the need to capture nodal locations in communications evaluations to full-scale Soldier-in-the-loop technical assessments. In each case, GPS data is archived locally, generally at the rate of one sample per second. Beacons are sent less frequently, based on the number of elements in the network.

A backhaul network can be implemented using a set of fixed structures and quick-erect towers. The PM's 2004-2008 activities employed such a network to varying degrees, which allowed coverage of the eastern portion of the Fort Dix range area. When available, the backhaul provides data to support:

- An integrated view of all instrumented entities in a manner that is "out-of-band" to any SUT networks; this integrated view can also be archived for rapid replay of a test event
- Real-time input to simulation environments, where such data can be employed to "ghost" the live entities into the simulation; this capability enables simulated sensor systems to "see" the live entities

Data is generated and collected via a lightweight application, termed *tget*, that is tightly integrated with an attached commercial Global Positioning System (GPS) receiver via a custom interface box. This interface box is used to provide power to the GPS receiver, buffer the receiver's one pulse-per-second (1PPS) output signal, and provide connections between the 1PPS and serial data lines from the receiver to a 9-pin serial interface common on most host computers. A custom software driver within *tget* enables the precise capture of the leading edge of the 1PPS signal and ties this event (and the associated GPS time-position tuple) with a tick of the host computer's CPU. The resulting data stream is archived locally and sent via interprocess communication to a local control application which can subsequently beacon this data to a centralized station for monitoring or insertion into a simulation environment.

The linkage with the host computer's CPU timer enables other collection applications running on the same computer to obtain very precise time tagging of given events. If these applications also associate events of interest with the CPU timer, then a post-processing step can map the CPU time to GPS time via the data collected by *tget*. The Network Data Collection System relies on this information as described in the next section.

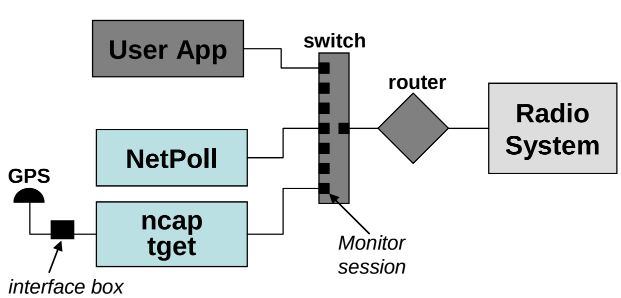
Once data files are harvested into a common archive, the FF-IDCR post-processing capability enables export into a relational database, forming an integrated set of position records for an activity. Since real-world data tends to have gaps caused from intermittent GPS outages (especially in operations under heavy foliage), various smoothing techniques are generally applied to create a regularly sampled data set. This integrated and refined data set is termed the Ground Truth Database, and includes regularly sampled tuples of:

(time, entityId, locationVector, errorMetric)

3.2 The Network Data Collection System

A significant amount of research and development has been invested into tools and techniques to understand network behavior under field conditions. This is a non-trivial problem, in part due to the variety of tactical communications systems and their theory of operation, and in part due to the fact that nodes are distributed over significant areas. The former challenge is partially addressed by focusing collection at the Open Systems Interconnect (OSI) model's Layer 3 via widely used packet capture techniques; variations among communications systems below Layer 3 must be addressed in a more customized manner. The latter challenge is addressed via distributed time synchronization techniques.

The following diagram illustrates the general manner in which a network node is instrumented for Layer 3 data capture:



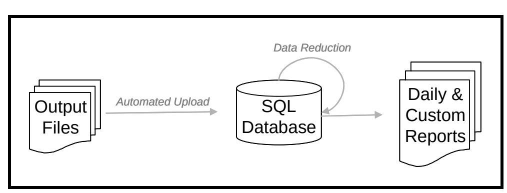
This model applies to most fixed and vehicular-mounted C4ISR systems, and can be applied to dismounted systems (though these have many variants which will not be discussed here).

Most modern switch devices provide a mechanism to set up a *monitor session*, in which frames sent from or received by a specified group of ports can be duplicated and sent to a designated monitor port. This only impacts overall system performance if the data rate across the switch approaches the capabilities of the device's backplane. Generally, switches can handle Gigabits of data per second, while radio systems can only support Megabits of data per second (or less). Thus, this technique provides no measurable impact on the SUT.

Once the monitor session is properly configured, a host attached to the switch's monitor port can then observe all traffic within the monitor session, and a packet capture application can archive this traffic. PM C4ISR OTM has employed several commercial, freeware and custom packet capture tools over the past several years, and has also developed a lightweight, custom implementation, termed *ncap*. As described in Section 3.1, the Ground Truth System provides a linkage between host computer CPU time and GPS time, and *ncap* leverages this to perform highly precise (<1msec) time tagging of offered and received packets. Such precision is critical to the quantification of delay and jitter metrics across nodes in a tactical network. All data is archived locally in the industry-standard PCAP format.

An associated reduction tool, termed *parse*, enables "parsing" of the packets into the higher-level protocols that may be embedded within or across Layer 3 packets. *parse* currently supports User Datagram Protocol (UDP), Transmission Control Protocol (TCP), Internet Control Message Protocol (ICMP), Open Shortest Path First (OSPF), Distributed Interactive Simulation (DIS), Joint Variable Message Format (JVMF), selected Cursor on Target (COT) messages, Voice over Internet Protocol (VoIP) and Cisco NetFlow. A variety of other protocols/messages are supported, and additional parsers can be written and integrated into the application in a straightforward manner. Any PCAP-compliant tool can also be used to review, process or visualize *ncap*-collected data.

The *parse* application exports data to a relational database, which facilitates the creation of formatted and ad-hoc reports:



Standard reduced products include:

- Offered & Received Load – time histories of aggregate observed load
- Top Talkers – ordered list of observed load between host pairs and application ports
- Pairwise Statistics – time histories of packet or message completion rates, latency and jitter

The combination of packet capture and associated reduction tools provides insight into Layer 3 and above performance. An additional lightweight application developed by PM C4ISR OTM, termed *NetPoll*, provides a view of alternate Layer 3 products and selected Layer 1 and Layer 2 performance statistics as they are made available via each network device. *NetPoll* employs the Simple Network Management Protocol (SNMP) to periodically probe the status of the Management Information Base (MIB) of SNMP-capable devices. Most routers support this capability, and *NetPoll* uses this to collect input/output statistics on selected interfaces and the devices' routing tables. Most radios and modems also support SNMP, and custom tools are created for each system. For example, during 2007 & 2008, polling scripts for L3's MPM-1000 and the Harris Highband Networking Radio (HNR) were developed and employed. Simple post-processing filters convert the raw output of *NetPoll* into tabular form suitable for upload into relational databases or use within common spreadsheet packages. Standard reduced products from *NetPoll* include:

- Nodal Status – time histories of representative reported quantities per device (e.g., CPU utilization, uptime)
- Next Hop – time history of the node used by a sending node as the next hop for Layer 3 routing to a given destination node; these graphs are useful in evaluating the performance of active routing protocols (e.g., OSPF)
- Link Visualization – a dynamic portrayal of nodal position and link state over time; generally exported as "KML" files and visualized within GoogleEarth

3.3 The Player Interactions System

Complementing the automated systems described in the preceding sections, several independent methods are used to record the manner in which human operators or role players interact with the C4ISR systems under evaluation. These methods support requirements to quantify the actual output of C4ISR system user interfaces (e.g., to facilitate comparison with "ground truth" data), as well as requirements to assess the manner in which role players interacted with the systems themselves (e.g., to capture what a user saw or heard during a particular activity).

Two primary forms of data collection are employed in this regard:

- Screen Capture. Periodic digital captures are taken of the displays presented to each key experimental participant; this is accomplished with a mix of commercial and custom tools, depending on the host computer's capabilities and operating system; in certain instances these tools also enable corresponding capture of player input via mouse and keyboard actions. Periodic screen captures at high frequencies (greater than once per minute) tend to use significant local storage, but can be invaluable for rapid post-exercise review.
- Audio Capture. Digital and analog recording of tactical voice traffic at selected locations can be performed via a mix of commercial and custom tools. Audio captures also use significant local storage and are difficult to easily "reduce" into simpler forms. Little direct audio capture was performed as part of the 2007 activity, though all tactical Voice over Internet Protocol (VoIP) packets were collected in 2008.

3.4 The Log File & Harvesting System

The final class of FF-IDCR collection systems focuses at the file level. Most software applications that drive C4ISR systems support some type of logging capability, and these logs can augment standard data collection systems or even eliminate the need for additional collection in certain circumstances. Similarly, operating system and application configuration information is critical to collect in order to properly portray the context within which systems were assessed. As well, the collection systems described above each produce data files that represent their raw output.

Based on this need to consistently and accurately aggregate files on host systems, a custom lightweight application termed *DC2P* (Data Collection Control Panel) was developed. This application reads from a configuration file and aggregates and compresses selected files from the local host into an archive on a target folder (generally a network location). It can be configured to either move or copy source files, and this enables the creation of a single configuration file that can be employed across all hosts involved in an exercise, thereby minimizing the potential for human error. The compressed file archives from each host are organized by experimental run on a central server and are mirrored on a backup data server.

Following the end-of-day harvesting, a series of automated and manual processes are conducted to verify the existence of primary data products, perform basic checks on file validity and transform the directory hierarchy of the collected products into one that better supports technical analytic efforts. Data is generally organized by experimental run, data class and collection point.

As its name suggests, *DC2P* also serves to control and monitor the other collection applications running on the same computer and/or on a local area network. It supports a remote monitoring and control capability and maintains a separate log that records operational status of each monitored application.

4. USES OF THE FF-IDCR SUITE

The FF-IDCR suite has been developed, employed, refined and vetted during multiple integrated C4ISR SoS assessments from 2004 through 2008. Within this section, we describe two case studies: the Program Manager Warfigher Information Network – Tactical (PM WIN-T) Engineering Field Test (EFT), conducted in the fall of 2007, and C4ISR OTM Event 08 (E08), conducted in the spring/summer of 2008.

4.1 2007 WIN-T Engineering Field Test

As part of the 2007 WIN-T EFT, CERDEC's PM C4ISR OTM partnered with PEO C3T's PM WIN-T in order to design, construct and assess an integrated transmission architecture based on the WIN-T Increment 2 specification. PM C4ISR OTM's role in these assessments included test design support, instrumentation, data collection and data reduction. These activities enabled PM WIN-T to comply with a June 2007 Acquisition Decision Memorandum (ADM) and obtain an early look at portions of the Increment 2 architecture within a network of over 20 nodes. The goal of the EFT was to build a body of evidence supporting critical technology readiness for key architecture components. Nearly 750 Gigabytes of raw data were collected over 20 days of testing in October 2007, with interim products supplied to WIN-T as the test progressed.

Initial insights regarding detailed system performance, based on collected data, were presented during the event's Presentation Days in early November, and then refined into a Technology Maturity Assessment (TMA) document provided to the Director, Defense Research & Engineering (DDRE). As a direct result of the EFT and the TMA, DDRE subsequently approved several WIN-T Critical Technology Elements as Technology Readiness Level (TRL) 6, a requirement of the June 2007 ADM. As a follow-on effort, PM C4ISR OTM is in the process of supporting WIN-T's upcoming Increment 2 Developmental Test (DT), Engineering Field Test (EFT) and Limited User Test (LUT), scheduled from November 2008 through March 2009.

4.2 C4ISR OTM Event '08

C4ISR OTM Event '08 (E08) was the core activity of PM C4ISR OTM during FY2008. Its purpose was to:

- Mitigate risk for and enable C4ISR technology development
- Explore engineering challenges associated with C4ISR systems integration
- Define and mature metrics that quantify the technical performance of C4ISR systems and systems-of-systems
- Study cognitive impacts of the employment of integrated C4ISR systems
- Utilize and assess varying solutions in support of Future Force C4ISR instrumentation, data collection & reduction.

The E08 architecture consisted of over 100 live systems across the C4ISR technology base, including elements of WIN-T Increments 1, 2 and 3, the Joint Tactical Radio System (JTRS) Soldier Radio Waveform (SRW) and Wideband Networking Waveform (WNW), the Enhanced Position Location Reporting System (EPLRS) in alternative configurations, Tactical Common Data Links (TCDL), the A160T Unmanned Aerial System (UAS), the USAF Paul Revere testbed, and other airborne and ground-based sensor systems, the Distributed Common Ground System – Army (DCGS-A), the Army Battle Command System (ABCS) and elements of Future Combat Systems (FCS) Battle Command including the System of Systems Common Operating Environment (SoSCOE).

Emerging results reflecting findings across the scope of E08 were presented during C4ISR OTM E08 Presentation Days in August 2008. The complete Final Report is planned for publication in November 2008.

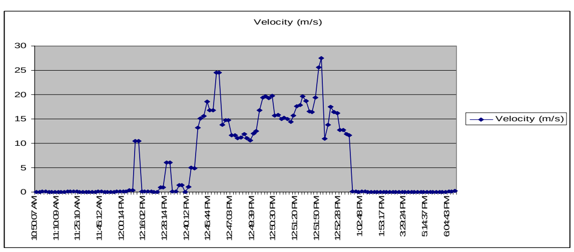
5. SAMPLE REDUCED PRODUCTS

This section provides a graphical summary of the varying types of products created by the systems described above. Analysts use products such as these in combination with other reduced products or in combination with ancillary data (e.g., observer notes) in order to draw conclusions and gain insight. "Data" cannot be properly analyzed without knowledge both of the context in which that data was gathered and of the conditions under which the SUT was stimulated.

5.1 Ground Truth Examples

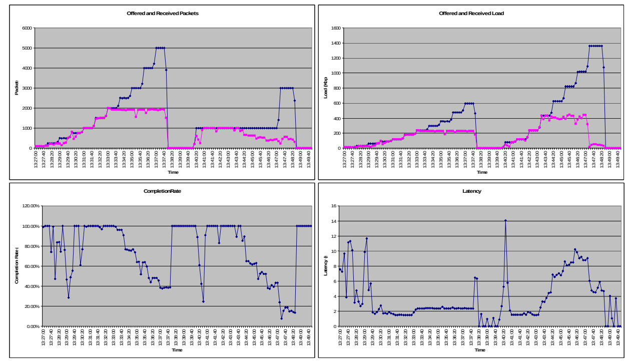
The figure below illustrates the Ground Truth picture during the pre-mission preparation phase of a Soldier-in-the-loop technical assessment conducted in 2006. The blue icons represent Friendly Forces, while the red icons represent the Threat. This data represents the "integrated" Ground Truth picture – in other words, the combination of all individually-collected position records – and is displayed here via a visualization application:

Position data can also be used to derive approximations of heading and velocity. The line graph below illustrates the velocity of one of the tactical vehicles over time during a communications evaluation conducted during E07:

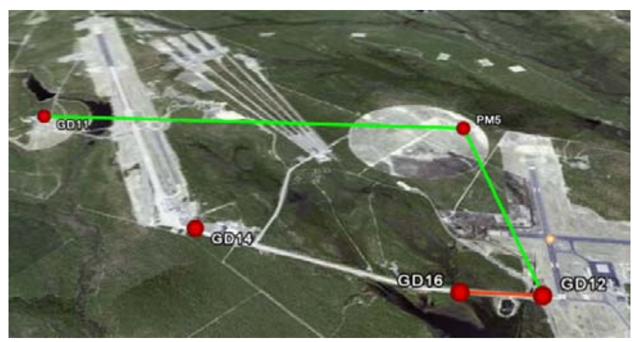


5.2 Network Data Examples

The set of four thumbnail plots shown below illustrate standard products produced by the Network Data Collection System. These tools treat data transfers between pairs of nodes as "flows", and report aggregated statistics relevant to those flows. These statistics include: (1) offered load, representing the traffic sent from one node to another, measured in packets or bits per second; (2) received load, representing the traffic actually received from the sending node, measured in packets or bits per second; (3) completion rate, or the percentage of packets sent that were successfully received; and (4) latency, representing the average transit time for successfully received packets. These values are all indicative of the response of a communications system to varying stimuli (e.g., load, mobility, line-of-sight obstruction, etc.):

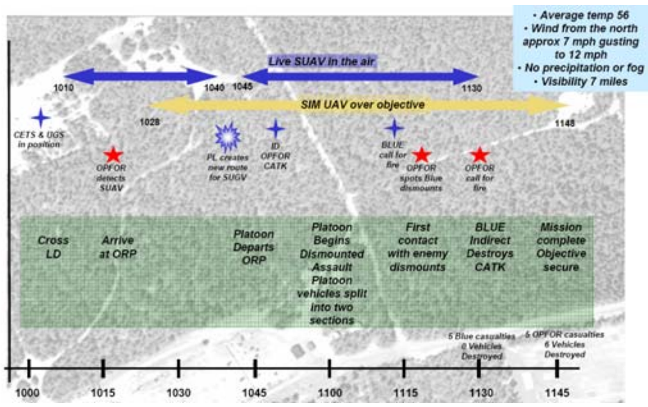


Network data products can be combined with the Ground Truth data products, and visualized within a geospatial context as shown below. Here, the results of a communications evaluation are portrayed graphically, with node locations shown by the red circles and link status illustrated by the presence and color of lines between the nodes. This particular data product is formatted in Keyhole Markup Language (KML), which is the native input format for the GoogleEarth application. Many of the integrated data products from E07, the WIN-T 2007 EFT and E08 have been cast into KML, since it leverages the capability and ubiquity of the freeware GoogleEarth product.

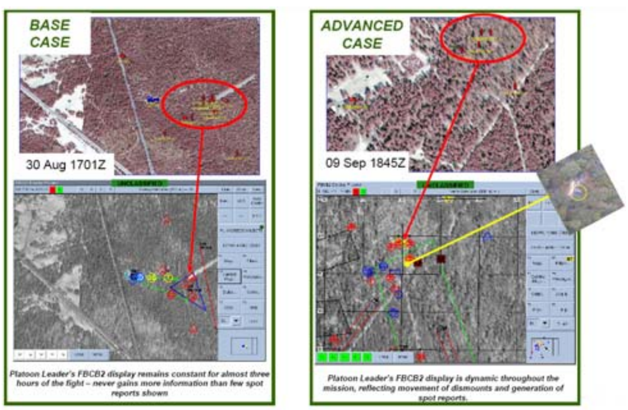


5.3 Integrated Examples

Data reduction is not a solely automated process. Manual integration may be required in order to represent criteria such as observer notes or environmental data. The figure below, created in support of the first C4ISR activity at Fort Dix in 2005, illustrates the timeline of a Soldier-in-the-loop technical assessment. Both the "operational" timeline of Soldier maneuver and combat operations, as well as the "technical" timeline of C4ISR asset availability and employment are displayed:



Extending this concept, the graphic below represents a comparison between the performance of varying system-of-systems configurations across two technical assessments in 2005



This product integrates the Ground Truth picture, screen captures of the battle command system, observer notes, and a saved image "chip" captured by an unmanned aerial sensor.

CONCLUSIONS

The development of the FF-IDCR tool suite is a Science & Technology (S&T) effort, designed to support assessment events that provide quantifiable feedback to technology developers and program managers. This initiative assists in the articulation of requirements which can be leveraged by the Department of Defense (DoD) Test community, which must develop and employ such tools during the developmental and operational tests of the Future Force network.

ACKNOWLEDGEMENTS

Each of the integrated C4ISR events described in this paper has resulted from the efforts of numerous people across several dozen organizations. The activities to design, architect, integrate, test, execute and subsequently reduce and analyze data for such efforts were large and complex undertakings. It would be very difficult to properly allocate credit among the many participants responsible for the aggregate findings articulated and visualized within this report. PM C4ISR OTM, and the authors, wish to thank those who silently, but actively, supported this work.