NWS Concept of Operations
FY2006
Pre-Prototype Laboratory

Final Report

Submitted by
By the
CONOPS Prototype Plan Team
With Support from
ESRL Global Systems Division,
Information Systems Branch

Last Revision:
December 20, 2006

CONOPS Prototype Team:

Gary Foltz - Chairman
James Partain – Lab Sub-Team Leader
Larry Dunn
Darin Figurskey
Rob Hartman
Chip Kasper
Pete Manousos
Rusty Billingsley
Bill Bunting

Major Contributors:

Woody Roberts –Global Systems Division (GSD)
Leigh Cheatwood-Harris - GSD
Tom LeFebvre – GSD
Darien Davis – GSD
Greg Noonan – CR Headquarters
Pete Browning – CR Headquarters
Brian Gockel – OS&T

Facilitators:

Philipia Hillman – SRA-Touchstone
Laura Arcieri - SRA-Touchstone

Table of Contents

1.0 Executive Summary

Under the guidance of the CONOPS Prototype Team, the Earth Systems Research Laboratory (ESRL) Global Systems Division (GSD) developed and implemented a plan for conducting a real-time laboratory exercise from September through early October, 2006. This was a pre-prototype activity which focused on the hardware, software, and infrastructure needed to support a CONOPS prototype with current staffing levels and under the existing AWIPS architecture. The major goals for this laboratory were:

- Determine whether AWIPS software, hardware, communication components, and enhanced features, can support a CONOPS expanded domain capability;
- Provide feedback and recommendations to NWS management and developers on the initial forms and tools of a Resource Allocation Process (ReAP) that will occur daily or even more often.
- Provide feedback and recommendations to NWS management and developers on human factor issues resulting from the expanded domain and ReAP;
- Provide the necessary information for a decision by the NWS Corporate Board to move forward with operational prototypes.

The CONOPS lab was the first time that many of the software components installed on standard AWIPS hardware with expanded domain capabilities had been tested. Testing during the exercise was conducted by field personnel and was designed to load AWIPS and the communications network using possible operational scenarios within CONOPS.

The CONOPS lab was successful. It tested all the necessary components and identified what must be done to make the technology work during the prototype. It provided valuable information for the future. It demonstrated that expanded domain within the enhanced GFE software is viable. The lab also identified the human factors, technology and training issues that must be addressed to support the prototype effort. The field staff that participated in the lab was exposed to the clustered peer concept of operations and felt that it had significant potential. The partnership with GSD was also a resounding success. GSD worked extremely hard to create and support the laboratory. Finally, the lab provided the opportunity to learn by "doing".

<u>**Summary of Critical Findings/Recommendations from the Lab**</u>

The following seven critical findings and recommendations are derived from detailed findings and recommendations found in Section 4 (Evaluation Results). The bracketed numbers refer to the specific items referenced from that section.

C1. Overarching Finding: The lab exercise demonstrated the technology tested could support the CONOPS, but the large domain size (120,000 grid points) and the large number of WFOs in the clusters resulted in issues with system performance and human factors that require mitigation before it is prudent to proceed to implementation of the prototype activities. [Findings 9, 12, and 13]

Recommendation: Design and implement the prototype over domains initially consisting of clusters of two WFOs and eventually clusters of four WFOs with total cluster domain size of approximately 80,000 grid points. [Recommendations 4, 6, and 12]

C2. Finding: The sharing and processing of Intersite Coordination (ISC) grids among cluster WFOs over the large domain of the lab exercise resulted in degradation of performance of the GFE and delays for the forecasters in doing their work. This problem is associated with ISC-related programs within the GFE rather than network bandwidth. [Findings 13, 14, 31, and 36]

Recommendation: In the short term, ISC-related processing should be modified to maximize performance within the existing GFE architecture, smaller domains with fewer WFOs and grid points should be used in the initial prototype activities, and ISC grid status information must be available to forecasters in real-time so they know which ISC grids have and have not been received. In the longer term, other architectural frameworks for more efficient sharing and processing ISC grids must be identified. [Recommendations 9, 10, and 23]

C3. Finding: Participants found creation of legacy text products to be unacceptably time-consuming when required to produce products during exercises that included responsibility for most (all) of the cluster WFOs (e.g. up to an hour just to run the formatters when required to produce text products for 6-8 WFOs). [Finding 15]

Recommendation: The GFE interface for formatting text products should be modified to allow forecasters to format all products for a given WFO with a single action rather than acting on each product separately. Smaller domains with fewer WFOs per cluster are required to mitigate this problem during the prototype. [Recommendation 12]

C4. Finding: Forecasters were generally satisfied with the capabilities of FX-C for the ReAP, but additional capabilities were requested to optimize its usability (several forecasters commented on the desirability of having a single application for chat, drawing, and screen/image sharing). [Findings 25 and 26]

Recommendation: Pre-prototype development should be done in FX-C to accommodate the highest priority requested capabilities (such as changing menus, access to GFE data, etc). [Recommendation 20]

C5. Finding: The expanded-domain capability tested in the lab provides a superior service backup method over the current baseline. [Finding 32]

Recommendation: The expanded-domain CONOPS version of AWIPS should be completed and tested for use in the CONOPS prototype activity, but a parallel effort should begin to target it for national implementation (as early as OB 8.3) to replace the existing service backup capability. [Recommendations 24 and 25]

C6. Finding: Lab participants identified the need for training in three key areas prior to commencement of prototype activities. [Findings 37 and 38]

Recommendation: Training material and a plan for delivery should be developed for (1) use of the new technology (FX-C, enhanced GFE, D2D modifications), (2) collaboration techniques and the culture of collaboration versus coordination, and (3) the cultural shift associated with changing from WFO operations to cluster operations (team building and the overall "concept" of the new CONOPS). [Recommendations 29 and 30]

C7. Finding: The expanded-domain, enhanced AWIPS is not compatible with the legacy AWIPS in some key areas such as the sharing of ISC grids. [Finding 40]

Recommendation: Develop a technical solution for the incompatibility between the CONOPS AWIPS and the legacy AWIPS and require the initial 2-WFO clusters to be comprised of service backup pairs. [Recommendations 32 and 33]

Additional Findings/Recommendations are found in the report dealing with issues at a greater level of granularity. The recommendations delineated above are deemed absolutely critical to the successful accomplishment of the CONOPS Prototype Plan.

If these recommendations are accepted and acted upon, the CONOPS Prototype Team believes that:

- Clustered peer operations for two offices will work with the completion of the additional work described.
- After further software development, clustered peer operations for 4 offices will be possible. A more sophisticated ReAP tracking tool will be necessary.
- GFE Re-architecture work must begin in FY07 if it is to be ready for the Service Oriented Architecture (SAO) of AWIPS-next in late FY09. This is necessary to expand the size of the domain, the number of offices in a cluster, and to determine the national cluster configuration

2.0 Introduction

The NOAA OAR/ESRL Global Systems Division (GSD) Information Systems Branch was tasked by the NWS CONOPS Lab Team to develop an evaluation for conducting a real-time pre-prototype laboratory (subsequently referred to as CONOPS-06) exercise targeted to begin in the last quarter of FY 2006. This exercise focused on demonstrating the required hardware, software, and infrastructure needed to support further field testing of the NWS Concept of Operations (CONOPS) Cluster-Peer proposal in FY2007 (2006 CONOPS Prototype Team, *The National Weather Service Clustered Peer Concept of Operations Prototype Plan*, Fiscal Years 2007-2009) . The major evaluation goals for this first exercise are the following:

- Determine what AWIPS software, hardware, and communication components, and enhanced features can support expanded domain prototype testing in FY2007.

- Provide feedback and recommendations to NWS management and system developers on the various components and sub-components of the system, when used to support an expanded CONOPS domain capability.

This pre-prototype laboratory exercise was an important initial test of the software components installed on standard AWIPS hardware with expanded domain capabilities. Testing conducted during the exercise was designed to cover a subset of possible operational scenarios to be further tested in subsequent prototypes. The primary evaluation components include the AWIPS two-dimensional display and text system (D2D), the AWIPS Graphical Forecast Editor (GFE), FX-Collaborate (FX-C) and other collaboration tools, and the new AWIPS Wide Area Network (NWS-NET). The evaluation efforts were directed toward gathering information on a broad range of topics. These topics include system reliability and performance, user interfaces, basic usability and utility of components, screen resource management, and training.

CONOPS-06 utilized two non-operational AWIPS systems for testing; one located at GSD (referred to as FSLC) and the other at NWS Central Region Headquarters (CRH) in Kansas City (referred to as BCQ). The systems were configured with enhanced operational AWIPS software and other non-AWIPS components including FX-C. NWS communications was used for data ingest and data exchange. The exercise was conducted primarily within the AWIPS firewall, following staging and pre-testing outside the firewall. For additional information on the exercise, see the Pre-Prototype Laboratory Exercise Plan (2006).

Feedback from the exercise will be used as input for future refinements of the CONOPS development plan and to direct development efforts for each component of the system before the next phase of field prototyping. Further, CONOPS-06 will provide the NWS with information relevant to future operational systems.

The following report details evaluation methodologies used during the exercise and the findings and recommendations for each of the major evaluation components of the lab. The appendices provide additional detailed information from the evaluations.

3.0 Evaluation Methodology

A number of evaluation methods were used during the laboratory, including questionnaires, interviews, observations, application logs, system logs, network logs, and exercises. These methods each have various advantages and disadvantages in terms of cost to the evaluators, cost to participants, obtrusiveness in the test environment, and data quality. The complementary nature of the methods helps overcome the deficiencies inherent in each individual method, and as a whole provide a complete picture of laboratory strengths and weaknesses. These techniques are further discussed in Lusk et al. (1999).

Three questionnaires were used during the exercise. An evaluation log (E-Log) was available any time participants wanted to quickly note problems or make other comments. An End-of-Shift questionnaire allowed participants to rate and comment on the system components used during their shift. The final, End-of-Week questionnaire was a more comprehensive list of questions that covered all activities during the week. The questionnaires were designed to be free of evaluator bias, allow respondent anonymity, and provide quantitative information that was summarized with descriptive statistics. The statistics, along with the accompanying questions, appear in Appendix B and Appendix C of this report. Additionally, open-ended questions provided a wealth of anecdotal information and useful suggestions. Responses to the open-ended questions appear in a separate document (Cheatwood-Harris, 2006). All of the questionnaires were administered by the GSD E-Team online via the web.

Application logs from D2D, GFE, and FXC systematically tracked all product selections, tool use, screen manipulations, and some background tasks that the applications ran. The logs were primarily used to document times that heavy network traffic occurred during the exercise. Network logs provided information on network traffic during the exercise.

Members of the CONOPS Team, as well as GSD and CRH staff, observed operations throughout the exercise. They provided additional documentation of activities, problems that occurred, and resolutions to those problems. The CONOPS team member observing at GSD provided daily summaries of lab activities during the exercise.

Interviews were used to supplement questionnaires and provide more in-depth information on different topics during the lab. Unstructured interviews of lab participants were conducted by CONOPS team members when issues arose during the lab that needed additional discussion. Additionally, more formal interviews were conducted at the end of each week (the weekly group debriefings) and were led by the E-Team leader. These debriefings were attended by CONOPS team members, developers, and managers and gave participants an opportunity to comment on the week’s activities directly with those interested parities. This was the last activity of each week, after participants had completed the End-of-Week questionnaire.

Daily exercises conducted during each week were designed to test a range of activities that might occur and to test components of the system to determine whether they could effectively perform different tasks. The exercises followed the same pattern each week, starting with simpler tasks similar to current office operations, and then progressively increased in difficulty throughout the week.

Results from all of these evaluation methods were collected and compiled and form the basis of the findings for this report. When appropriate, specific numerical results are reported along with summaries and quotes from comments collected from questionnaires during the lab.

4.0 Evaluation Results

The following section details all of the findings and recommendations from the lab. Each major evaluation area is covered comprehensively using the feedback gathered from the participants and other metrics. Findings and recommendations that contribute to the overall critical findings and recommendations are identified with an asterisk (*). When appropriate, numerical ratings from the questionnaires are presented along with quotes directly from participants' questionnaire responses. In some cases, the findings and recommendations also include subsequent actions planned to resolve or mitigate problems that were identified. Many of the identified performance problems will be mitigated by plans to reduce cluster sizes initially so smaller grid sizes will be needed for computational purposes (see The National Weather Service Clustered Peer Concept of Operations Prototype Plan Fiscal Years 2007-2009, November 2006). Subsequent lab testing will revisit all of the findings and recommendations from this first lab and document changes that occur.

4.1 AWIPS D2D Workspace

The AWIPS, Display Two Dimensions (D2D) workspace offers display and interactive capabilities for viewing and interacting with operational hydrometeorological data. The evaluation goal for D2D was to determine the extent to which the D2D workstation could support the simulated CONOPS expanded domain environment. For the evaluation, D2D was localized to the two home County Warning Areas (CWAs) in each cluster each week. All of the Satellite Broadcast Network (SBN) data was regularly available for viewing along with a subset of local radar data. Other local datasets were not available. After the first week, an additional map background was made available that highlighted the office's current area of responsibility (AOR) along with the AOR of the other offices in their cluster.

Finding 1: In general, participants found the D2D performance to be better than acceptable (average 3.36 on the daily questionnaire and 3.67 on the weekly questionnaire on a 5-point scale with 3.0 being acceptable) for supporting simulated CONOPS operations.

Finding 2: Participants liked the AOR map backgrounds but wanted an automated process to change backgrounds when the WFO's AOR changed.

D2D system performance, product selection and display, text windows, and cluster support were all rated by participants as acceptable or better. Participants noted that D2D performed as it does in their home offices. They commented on the lack of complete radar data, local procedures, local models and mesonets. These comments were expected due to limitations with the configuration of lab systems and access to these data. Participants also commented on the need to have products available on all scales so product selection would not have to be scale-dependent. Participants noted that the D2D scales and LAPS domain size did not always allow for easy viewing or coverage over the full cluster domain. Furthermore, by design, WarnGen was not fully functional due to the lab configuration.

Finding 3: Participants need access to all locally-generated D2D procedures, data, and model output, for CONOPS operations.

Finding 4: D2D viewing scales and LAPS were not always properly aligned with the CONOPS cluster areas.

D2D screen management and training were rated at slightly less than acceptable. An early AWIPS 7.2 build was used for the lab which was new to the participants. They liked some of the new features, such as longer frame count, but training was needed on the new capabilities. After the first week, release notes were sent to participants so they could learn about the new build features before participating in the lab.

Many participants commented that more AWIPS screen space would be desirable since several applications (e.g. D2D, GFE, FXC and Text) could be running simultaneously on one workstation. Participants noted that many AWIPS users in the field already use multiple desk-tops on each screen to avoid having to reload displays with different products (i.e. they swap desk tops with different products loaded rather than load new products on the one desktop).

Finding 5: Participants need training on D2D changes when upgrades to the system occur.

Finding 6: Screen space became problematic with more applications and larger areas of responsibility.

Recommendation 1: D2D is acceptable for further CONOPS testing with minor enhancements and configuration changes to better accommodate cluster operations and cluster spatial areas.

Recommendation 2: Automate CWA/AOR map background changes to reflect AOR changes as they occur.

Recommendation 3: Consider adding more screens or larger screens to AWIPS to improve screen management characteristics of the system.

4.2 AWIPS Graphical Forecast Editor

The Graphical Forecast Editor (GFE) was a primary focus of evaluations during the CONOPS ’06 lab. The GFE was modified and enhanced prior to the lab to accommodate gridded-forecast preparation and accompanying textual forecast generation for multiple CWAs within each of the four tested Clustered-Peered environments.

Finding 7: Participants successfully generated forecast grids and text products over the cluster CWA areas using the enhanced GFE capabilities.

Only one substantial GFE software problem occurred at the start of the first week which required correction before lab activities could resume. Once this was identified and fixed, the software ran reasonably well, albeit slowly, throughout the rest of the lab period. Occasionally software failures did occur (about 1-3 times per week), but software restarts were typically all that was needed to recover. This recovery procedure is also run in the field.

Finding 8: Participants rated the GFE capabilities as somewhat less than acceptable (average 2.64 on the daily questionnaire and 2.60 on the weekly questionnaire on a 5-point scale with 3.0 being adequate) when used to support CONOPS testing.

Many of the GFE comments were related to the slow performance of the GFE due to the expanded cluster domain size (about 120,000 grid points, on average). The 120k grid was determined before the lab to be the highend of acceptable performance (see Appendix D). This was most noticeable by participants who used much smaller domain sizes, and in some cases lower resolution, at their home offices.

***Finding 9:** Slow GFE performance, running over the 120,000 grid point cluster size domain, was a primary cause of the less than acceptable GFE rating.

***Recommendation 4:** Consider smaller cluster domain sizes for future testing as well as changes to the GFE which would improve performance of the software.

Additional GFE details as well as findings and recommendations are discussed in the following subsections.

4.2.1 GFE Smart Tools and Procedures

GFE Smart Tools allows forecasters to generate forecast parameters from gridded fields previously stored in the local AWIPS database. Several Smart Tools can be bundled and run together as Procedures within the GFE. NWS offices have created a very large number of customized Smart Tools and Procedures that run routinely at individual offices throughout the country. In some cases Smart Tools are shared among offices, but many are modified or otherwise adapted to an office's specific needs or meteorological conditions.

For purposes of the lab, the GFE team at GSD imported Smart Tools and Procedures from each of the participating offices and modified a subset of these for use during each week of the lab. Only a subset of Smart Tools and Procedures could be modified due to several factors. These include: the complexity of Smart Tools changes necessary to run with the GFE CONOPS-configured software, the short amount of time available to make those changes, and the large number of imported Smart Tools (sometimes numbering in the hundreds) and Procedures used by individual offices.

Finding 10: A subset of Smart Tools and Procedures was successfully modified and used to generate forecast grids during the exercise.

Participants noted on several occasions that the lack of a complete set of their offices' Smart Tools and Procedures significantly hampered completion of their tasks during the exercise. As one participant noted, *"The procedures and Smart Tools my office uses were not available. This made the process of creating and populating the grid fields time consuming and tedious."* As the days and weeks progressed, more and more smart tools were modified and made available, thus mitigating some of the concerns voiced by participants.

Finding 11: Participants need a more complete set of their offices' Smart Tools and Procedures to efficiently generate forecast grids for their home and cluster CWAs.

Recommendation 5: Integrate a more complete suite of each office's Smart Tools and Procedures before future CONOPS testing.

Participants also commented on slowness of running Smart Tools over the expanded domain. Comments ranged from, *"Grid calculations that would normally take <30 seconds took 5-10 minutes!"* to *"It is definitely slower that what we have now... considering the increase in resolution and domain, this wasn't as bad as I expected. Some of our tools were pretty slow however - which may be a function of how the tools are written as much as the increased load on the system due to domain and resolution."* On a few occasions the GFE crashed when Smart Tools was running during the exercise. This kind of failure also happens during field operations from time to time, but there are no statistics to determine or compare how often this occurs. Restarting the GFE resolved the problem.

GFE development staff noted that some of the Smart Tools could benefit from a review process that determines whether tools are written to run efficiently. Participants suggested running some of the processes in background (on faster machines) in order to mitigate some of the performance problems. A suggestion was also made to baseline more of the tools and make them available to all offices. GFE development staff had also determined before the exercise that the large cluster-sized forecast domain (120,000 grid points) was likely on the edge of acceptable performance.

***Finding 12:** Smart Tools and Procedures ran significantly slower over the cluster-sized expanded domain.

***Recommendation 6:** Explore and adopt options to improve Smart Tools and Procedures performance by considering Smart Tools software changes, run-time configuration changes, and domain size changes.

Recommendation 7: Proposed cluster offices should determine a common suite of Smart Tools and Procedures that would effectively run over their entire cluster domain.

4.2.2 GFE Intersite Coordination Grids

Intersite Coordination (ISC) grids are locally generated automatically by the GFE for transmission to surrounding offices so they have forecast information to use to improve the consistency of gridded forecasts for the National Digital Forecast Database (NDFD). ISC grids are used to initialize backup operations, and for the purposes of clustered-peer operations. ISC grids can be used to initialize a newly-acquired CWA within an office's AOR. ISC grids are sent every time local forecast fields are modified and saved as "official". The GFE has an internal mechanism that limits the number of grids sent at any particular time (a "throttle") in order to avoid saturating the network.

During the lab, ISC grid size increased as the AOR size increased by a factor equal to the size of each CWA added to the AOR. ISC grids were exchanged only between FSLC and BCQ using the NWSnet MPLS terrestrial network connection (see network performance section). This single site-to-site exchange represented only a small subset of offices that would receive ISC grids operationally. When ISC grids were received, a program ran in the background (ISCMosaic) to update the ISC grids that were displayed for the areas surrounding the local AOR.

***Finding 13:** ISC performance was often noted as problematic during the exercise.

ISC performance appeared to affect lab operations on several occasions and was most noticeable with large AORs and when ISC grids were being exchanged between sites. At times, participants noted ISC applications taking 30 minutes or longer to complete. Participants said GFE performance was sluggish when this occurred. Analysis of network traffic (see network performance section) indicated that there was ample bandwidth when ISC grids were sent and received.

***Finding 14:** Performance of ISC-related programs within the GFE (eg. ISCMosaic) was the likely cause of the ISC performance problems, not network bandwidth.

Recommendation 8: Review ISC-related programs within the GFE to determine what improvements can be made generating ISC mosaic fields.

Participants occasionally performed tasks assuming ISC grids had been received and processed when this was not the case. ISC history information was noted as useful by participants for determining when grids had been modified or sent. However, most of the problems occurred during ISC grid processing, before current ISC history information was available. Participants were also able to successfully send ISC grids based on temporal and element splits between offices, but this too was problematic at times due to insufficient information and safeguards preventing incorrect ISC grid sends.

***Recommendation 9:** Improve ISC Grid status information to include whether grids do or do not need to be sent, or are currently being sent, received, and processed.

Many participants noted that ISC traffic during the lab did not emulate operations. For example, operationally, ISC grids would be transmitted and received from all surrounding offices, not just one as was the case in the lab. Network bandwidth, though not an issue during the lab, may be a problem when more sites are sending and receiving traffic.

***Recommendation 10:** Future testing should better emulate a full load of incoming and outgoing ISC traffic in order to better determine adequate bandwidth requirements.

During the group debriefings, and at other times during the lab, there were discussions about different approaches to ISC capabilities and requirements that could improve the efficiency of the process. These suggestions ranged from simplified grid change information, to ISC grid broadcast via the satellite broadcast network (SBN), to centralized regional or national ISC grid processing. These options may mitigate some of the current or future performance problems.

Recommendation 11: Review the ISC process to determine whether other architectural frameworks can better accommodate ISC requirements and capabilities.

4.2.3 GFE Text Formatters

The IFPS system allows for the generation of text products from grids through the use of formatters. The lab evaluation allowed the testing of the formatters used to generate the baseline set of legacy text products for each office in the cluster under the various CONOPS scenarios. The actual formatters from each WFO participating in the lab were provided to GSD prior to the lab for configuration on the test systems, and these were the formatters tested during the various lab scenarios. Similar to the purpose for the other technical elements of the lab, the purpose of testing the formatters was to ensure that they ran acceptably and produced acceptable output.

Findings from the lab indicate that the performance of the formatters degraded in proportion to the domain expansion. During full-cluster scenarios, the formatters often took upwards of an hour to generate text products. This was deemed unacceptable by the lab participants and observers. Numerous “bugs” were discovered at the beginning of each week, as expected, and these were generally repaired quickly by GSD personnel. Other concerns regarding the formatters in CONOPS scenarios were related to the human factors associated with generating multiple products for multiple WFOs by a single forecaster.

***Finding 15:** In general, participants found the performance of the formatters to be unacceptably slow during simulated CONOPS operations (e.g. when running formatters for multiple WFOs).

***Recommendation 12:** Consider smaller domains, better/faster hardware, and more efficient software as possible solutions to improve text formatter performance. It may also be possible to separate product generation from “assembly”, i.e. by running one formatter containing multiple sites’ products. This would require changes to the product editor.

Finding 16: The domain covered by various formatters (e.g. public vs. fire weather vs. marine) varies in some cases from office to office and product to product, resulting in complexities in editing grids and running the formatters for other WFOs in the cluster.

Finding 17: In particular, formatters for products associated with fire weather were the most problematic.

Recommendation 13: Properly configure domain size and setup of formatters to improve fire-weather product generation during prototype activities.

Finding 18: Better tracking, status monitoring, and product-generation automation capabilities are needed in AWIPS to allow forecasters to manage the numerous products that must be generated for each WFO.

Recommendation 14: Change to the product editor interface (e.g. replace tabs with a matrix) to alleviate tracking and monitoring problems.

In addition to the technical findings from the lab, other findings associated with the formatters concern the paradigm of local vs. centralized generation of text products, and the fact that many formatters often reflect local needs, rather than cluster-wide needs. Individual WFOs often make modifications to a particular formatter, including the required input grids, which are not repeated at other WFOs. Thus, running the same formatter for multiple offices within a cluster would likely make the product appear different than if the “home WFO” were generating the product with their own formatter. Requiring all WFOs to be able to run different formatters for each WFO in the cluster would create great complexities in managing the forecast production process.

Recommendation 15: Issues associated with differences between local WFO formatters vs. baseline or “cluster” formatters must be explored in future prototype activities.

4.2.4 Graphical Hazard Generator

The Graphical Hazard Generator (GHG) is used within the GFE to produce hazard headlines and Valid Time Event codes (VTEC). The GHG was tested during the lab using six test backup exercise scenarios developed for AWIPS build testing. This allowed participants to test GHG functionality for different AORs and to test switching AORs when hazards were in effect.

Finding 19: Participants successfully generated a variety of hazards with associated VTECs and were able to switch between normal operations and service backup using the lab GHG capabilities.

Participants commented that they thought the GHG exercises were a useful part of the lab and the monitoring functions were helpful for monitoring hazards throughout the cluster area. Participants commented that they needed time to learn and understand how to go through the scenarios and shift responsibilities between offices. Testing generally went smoother once they went through the first few exercises but noted that the expanded domains added another level of complexity to the process.

Finding 20: Participants rated the GHG capabilities as slightly less than acceptable (average 2.8 on a 5-point scale with 3.0 being adequate) when used to support CONOPS testing.

Finding 21: The GHG monitor did not always display all hazards issued by an office.

This display problem was identified as an issue with an internal filter that reset when AORs changed. It has since been corrected.

Finding 22: A WFO's Marine and Fire-Weather zones do not always align with the corresponding Public zone or CWA of another office. Furthermore, map boundaries within CWAs and between adjoining CWAs do not always align, all of which caused occasional erroneous hazard identifications.

Some of the WFOs that participated during the lab were responsible for issuing fire weather and/or marine hazards for areas that are outside their normal CWA. The GFE team tried to adapt to as many of these configurations as possible, but hazard-specific differences in CWAs were difficult to accommodate within the current software structure.

Recommendation 16: Correct mapping and display problems noted during the lab, and consider better alignment strategies for program-specific hazard areas of responsibility within each CWA.

Participants noted that there were many tedious tasks associated with the generation of hazard products, especially when their AORs were expanded. They suggested automating and streamlining as many of the tedious tasks as possible to improve the timeliness of product broadcasts.

Recommendation 17: Develop scripts and procedures to streamline and automate as much of the hazard product generation process as possible.

4.3 Collaboration

One of the most important elements to the success of the Cluster-peer Concept of Operations is that of collaboration. In addition to telephones, forecasters in the NWS currently use 12Planet chat software on AWIPS as their primary method of collaborating forecasts between WFOs, RFCs, and with NCEP. 12Planet and the new 12Planet-Whiteboard were both tested in the lab, along with telephones, but the primary collaboration method that was tested was the GSD-developed FX-Collaborate software (FX-C).

FX-C has been used on AWIPS systems for several years to generate graphics, but not to collaborate. Prior to the lab, GSD configured FX-C to operate in the cluster-peer configuration. Forecasters were required to use FX-C multiple times each day in varying scenarios. From the responses of the participants, FX-C was one of the most positive aspects of the entire lab, and generated a lot of enthusiasm for its potential use in collaboration. FX-C was the highest-rated application in the lab, rating 3.55 on average from the end-of-shift questionnaires and 3.67 on average from the end-of-week questionnaires.

During the lab, all of the collaboration methods functioned in the various scenarios. One of the most surprising aspects of the lab was how helpful it turned out to be to have a telephone connection open between the Boulder and Kansas City lab sites on a nearly continuous basis. . The 12Planet-Whiteboard application is primitive, and received an “Inadequate” rating from the participants. 12Planet functioned as expected for chat capabilities, but was much less desired by the participants than FX-C (rated as less than Acceptable).

Finding 23: Having an open line on the telephone between the Boulder and Kansas City lab sites was very helpful in coordination.

Recommendation 18: Cluster Management Teams should investigate the viability of having open telephone lines between peer offices within a cluster. Voice-Over-Internet-Protocol (VoIP) and other web-based technologies may be of use.

Finding 24: The 12Planet-Whiteboard application was provided at the last minute to the lab participants, with little or no configuration or optimization, and had minimal interactive collaboration capabilities. 12Planet-Whiteboard was not deemed useful in the lab.

Recommendation 19: Without the addition of significant interactive collaboration capabilities, 12Planet-Whiteboard is not recommended for prototype activities.

***Finding 25:** Additional capabilities for FX-C were requested to increase its usability (such as changing menus, access to GFE data, etc).

***Finding 26:** Forecasters prefer a single application or mode that includes chat, drawing, and screen/image sharing.

***Recommendation 20:** Pre-prototype development is needed in FX-C to accommodate the highest-priority requested capabilities.

4.4 Resource Allocation Process (ReAP)

An essential aspect of meeting the goals of the future CONOPS is the ability to dynamically allocate resources among peer offices to allow a focus on high-impact events while providing efficiencies in the production of routine environmental information. To accomplish this, a resource allocation process is required. This process is envisioned to include such tools and training that, in concert with collaboration tools, provides a platform for load balancing lower priority duties in response to and in preparation for high impact events.

For the purposes of the lab, a very rudimentary ReAP was created, utilizing a combination of FX-C and GFE/D2D configuration scripts set up by GSD for the purpose of reconfiguring the GFE at each lab site based on the results of the ReAP and collaboration sessions. In general, lab participants found the ReAP concept showed a lot of promise, and it functioned smoothly given the structure of the lab. Participants were cautious regarding how it would work with more WFOs participating, and with more complex weather than what was experienced in the lab, etc.

Finding 27: The ReAP process worked well given the constraints of the lab (only 2 WFOs, generally benign weather).

Finding 28: There did not seem to be technical issues with the ReAP process itself that would prevent it from being used in WFOs during prototype testing.

Recommendation 21: ReAP must be tested under a larger variety of scenarios, and involving more offices, in order to refine the concept.

It is also important to note the success that simple audio technology enjoyed in the lab. As noted in the Collaboration section, a dedicated phone line was often left open between the two lab locations, and forecasters made great use of that simple venue to exchange information and coordinate.

Finding 29: Simple audio technology (a dedicated phone line) proved very valuable to quick and efficient coordination between lab participants.

Recommendation 22: NWS should examine other popular audio-visual technologies for use in collaboration and the ReAP process (e.g. PC-based LiveMeeting/Go-To Meeting, streaming audio/video via web, etc).

Temporal splitting was tested during the exercise as well as limited testing of weather element splits. Participants noted that there needs to be a better way of indicating those responsible for temporal and element sets of gridded information. Furthermore, the grids need to be "lockable" such that whoever has responsibility for that grid cannot have forecast information overwritten by someone else.

Another aspect of the process that needs to be improved is developing the ability for forecasters to monitor the status of ISC grids, e.g. to prevent changes of responsibility while ISC data is still being processed.

Finding 30: A better way of monitoring the status of the division of forecast responsibilities between cluster offices is required.

***Finding 31:** There needs to be a way for forecasters to monitor the current status of ISC grids.

***Recommendation 23:** Developers should work with the lab and prototype participants to develop requirements for a monitoring and grid-status capability**.**

4.5 Service Backup

During the lab there were several scenarios that required the lab participants to operate in a "service backup" mode, where they issued the grids and products on behalf of another WFO as if they were in a traditional service backup mode. While we consider this "service backup" in today's operations, this is considered to be a fundamental, routine operating principle of clustered-peer operations. Thus, it was important to test this several times each week during the lab.

One of the most encouraging findings from the lab was the utility, efficiency, and overall enthusiasm generated on the part of forecasters for this new concept of service backup. The fundamental reason for the new efficiency is that the expanded domain required for clustered-peer operations has the effect of making it much easier for one office to issue the products, including grid and legacy text products, on behalf of another office. The current service backup concept is time consuming, inefficient, and thus difficult to implement, utilize, train, and practice.

Ratings from the lab participants averaged out to 3.6 (out of 5) for Service Backup, the highest rating for anything in the lab aside from FX-C. Some of the specific comments from forecasters regarding expanded domain capabilities to meet current service backup requirements included the following:

"See benefit to larger domain for easier service backup…"

"Service Backup is very straight forward and much much better than the baseline backup offered in the field right now."

"Aside from slow text generation, service backup for grids was easily accomplished."

"This is the way we should do svc bkup. The restore process from ISC needs to be solidified; but will be easily doable."

"The Cluster GFE's greatest utility (out of the box) is going to be simplifying the service back up process. Much, much better than what we have now."

"If the performance, usability, bugs, and methodology issues are addressed, the system would be adequate for performing service backup at WFOs. In fact it would be superior to the current backup methodology and it would take much less time to initiate and return from, service backup."

"…This technique is far superior than the current backup."

"I think one of the best benefits to the larger domain is the ease of service backup. Much easier than current operations in the office. Dynamic domains would make this even easier. This could be used in the field almost immediately."

***Finding 32:** Aside from performance issues (e.g. system slowness due to expanded number of grid points), the expanded-domain enhanced AWIPS as tested in the lab provides a superior service backup capability over current baseline capabilities**.**

***Recommendation 24:** Given developments that support performance enhancements, future prototype testing should test the feasibility of this form of service backup replacing the current baseline capability.

***Recommendation 25:** Given successful testing, this capability for service backup should be targeted for national implementation at the earliest opportunity (e.g. as early as OB8.3 in 2008).

For purposes of the lab, modifications to the GFE grid-data structures were made in order to identify clusters (Cluster ID's) as well as the individual sites (Site ID's). This additional layer of information is not compatible with the operational data structures that don't have Cluster ID's.

Finding 33: Modifications to AWIPS to support expanded domain render it impossible for a non-prototype office to provide service backup to a prototype office.

Recommendation 26: Until national deployment is possible for this service backup capability, prototype clusters must be comprised of primary service backup pairs of offices.

4.6 Other AWIPS Applications

A number of other AWIPS applications such as SCAN, SafeSeas, Fog Monitor, AVNFps, and others were installed for the exercise.   Participants occasionally interacted with these applications, but not systematically since the emphasis of the exercise was primarily in other areas. Participants also noted that the applications were not typically used during quiet weather periods which occurred during the exercise.

Finding 34: Other AWIPS applications were not routinely used or evaluated during the exercise.

Recommendation 27: Other relevant AWIPS applications need to be systematically evaluated in future CONOPS exercises.

4.7 Network

The lab exercise demonstrated clusters of forecast offices exercising the exchange of data for collaboration beyond the current state of the AWIPS system used operationally at WFO sites. Two applications in particular exercised the network bandwidth: GFE and FXC. The FXC application was used for collaborating and defining the AOR. FXC requests a "picture" from a server, compresses it, then sends the results via a socket to a client. The information can be shared among many clients, creating a simultaneous D2D-like visualization of hydromet data. Annotations and drawing primitives can be added to the data and exchanged as well. The GFE application increased the domain of grids exchanged for intersite coordination supporting the execution of each site's forecast-preparation responsibilities as well as the NDFD. IT design of the new cluster concept identified increased network activity as a potential repercussion from the cluster concept.

The current AWIPS WAN uses a frame-relay system to transfer data between sites. A hierarchical hub technology is employed, requiring at least two hops when transferring data between WFOs. There was a desire to have no impact on operations coupled with the need to run the lab exercise on a system internal to the AWIPS WAN. The new NOAANet Multi-Protocol Label Switching (MPLS) system was due to be deployed to pilot sites during the summer prior to the lab exercise. The CONOPS team members were able to accelerate the incorporation of FSLC and BCQ into the MPLS pilot. This was a critical step to enable potentially voluminous exchange of information between the two lab exercise sites (FSLC and BCQ) while minimizing the risk to NWS operations. Further details on the network link, traffic analysis, and information on NOAANet and its relation to CONOPS can be found in Appendix E.

Finding 35: The network capacity for the CONOPS lab was significantly greater than that of the existing/legacy AWIPS WAN and, also, likely greater than the foreseeable-future MPLS WAN capacity by a factor of 30.

The lab exercise was conducted on 45Mbps intersite bandwidth whereas the projected WFO-to-WFO bandwidth will be approximately 1.5Mbps in NOAANet or about 1/30th the tested bandwidth.

***Finding 36:** The exercise showed that the ISC traffic used an average of less than .01% of the bandwidth between FSLC and BCQ, thus network performance did not adversely affect operations during the lab.

As was discussed in the previous finding, the lab was performed on an implementation of MPLS that is roughly 30 times the bandwidth that NOAANet will provide for operational sites. The .01% was skewed by this number.

Also note in the prior ISC discussion, that ISC traffic during the lab was limited to the connection between the two test sites and did not reflect all of the ISC traffic that is generated or received by operational sites. This would likely increase traffic by a factor as much as 6 to 8, depending on the number of surrounding sites sending/receiving ISC data in the current WFO configuration, and likely by an order of magnitude or more in a Cluster-Peer WFO configuration. Additionally, concurrent throughput testing of high-priority messages (e.g. watches and warnings) was not tested during the lab to see if any latency occurred nor was additional high-bandwidth data (e.g. associated radars) sent over the network.

Recommendation 28: Repeat lab exercise with a more realistic WFO scenario adding radar products and other normal WFO functions on a network supporting representative bandwidth and with network monitoring and analysis capabilities in place.

The network link was established only two days before the start of the lab so little time was available to establish all of the network monitoring and analysis capabilities that would be useful for a more complete analysis.

4.8 Training

The participants in the lab were provided several types of training before and during the lab. All participants were provided "CONOPS 101" training via teleconference in the weeks prior to the lab. This allowed them to hear from the CONOPS team the overview of what the CONOPS is, how the lab fit into the prototyping efforts, and provided them with a venue (on multiple occasions) to ask questions of CONOPS team members.

Participants were also provided documentation on FX-C prior to the lab, along with websites with links to the downloadable client software, FAQs, etc.

Finally, on the first morning of the lab, participants were provided with a training session that was intended to provide the "knobology" training for the system used during the lab, including FX-C, the GFE changes for expanded domain, the basic ReAP functionality, and how the lab would be conducted.

Training was evaluated by way of questionnaires during and after the lab. Training was also discussed at the group debriefings at the end of each week. Participants were asked to assess the adequacy of training on D2D, GFE, coordination, and non-GFE applications on AWIPS. While the ratings from the participants generally showed training to be "acceptable", with scores ranging from a low of 2.67 (out of 5) for non-GFE applications to a high of 3.2 for coordination, participant responses and comments from lab observers showed that much could be done to improve training prior to prototyping in field offices.

***Finding 37:** More training is needed on the specific aspects of system changes from baseline AWIPS operations.

***Recommendation 29:** Specific training modules must be developed and delivered to prototype participants on the software tools (i.e. FX-C, D2D and GFE enhancements in support of clustered-peer operations), and on ReAP tools.

***Finding 38:** Training specific to collaboration (beyond methodology, to include culture) is needed.

***Recommendation 30:** Training on collaboration, beyond "knobology", extending to human factors and culture issues, is needed for prototype participants.

Forecasters in the lab also commented on the need to better understand the meteorology, climatology, products, and customer needs of the WFOs in the cluster for which they were forecasting. This includes knowledge of the effects on weather that terrain and other geophysical and local effects have on the forecast elements, as well as the varying factors of product domains and specific customer needs for certain products (e.g. fire weather).

Finding 39: Cluster participants need training on the meteorological and geophysical characteristics for WFOs in the cluster, and on the varying products and customer needs.

Recommendation 31: Cluster Management Teams (CMTs) develop and implement a training program to increase knowledge of local issues at all offices within the cluster prior to full cluster operations.

4.9 Configurability

The lab provided an opportunity for developers and forecast staff to understand the relative complexities of configuring the baseline AWIPS system to perform in a clustered-peer environment. The AWIPS system in the lab was configured with OB7.2, which is the expected release during the beginning of prototype field operations. Enhancements were made to the system to support larger domains, ReAP and interactive collaboration processes, and forecast product generation in the cluster environment.

Many of the most significant changes to the system made for the lab render the AWIPS of cluster offices "incompatible" with AWIPS at non-cluster offices. For example, ISC grid-sharing between a cluster and non-cluster office will not work without further development. In addition, service backup cannot be performed by a non-cluster office on behalf of a cluster office.

***Finding 40:** Many enhancements made to AWIPS to support clustered-peer operations will require potentially significant development to allow cluster and non-cluster offices to interact.

***Recommendation 32:** Require all offices within a prototype cluster to be comprised of primary service backup pairs.

***Recommendation 33:** Development agencies will need to resolve the issues associated with "legacy" and cluster versions of AWIPS capabilities and their interoperability.

Finding 41: Forecasters noted several configuration problems in the lab involving map backgrounds, data mosaics, and domain problems with other AWIPS applications (e.g. LAPS).

Finding 42: Office-to-office differences in GFE weather elements, product coverage areas, edit areas, and other configuration items will need to be resolved prior to field prototyping.

Recommendation 34: Non-GFE configuration issues must be resolved by the CMTs prior to prototyping; GFE issues will require technical solutions by the developers.

Recommendation 35: Offices in the cluster prototypes need to be configured to receive additional radar data (either from the SBN or via the WAN from other ORPGs in the cluster).

5.0 Conclusion

Participants from the CONOPS Pre-Prototype Laboratory rated the conduct of the CONOPS Lab as better than acceptable (3.7 on the 5-point scale). As one participant stated,

"We really learned a lot from this exercise; things that work, things that don't work, things we need to test further. It was also good to get us thinking about all the complex issues that will need to be addressed and all the work that needs to be done before a prototype can proceed".

The findings and recommendations summarized in this report underscore the challenges faced when developing the technology, infrastructure, and human resources necessary to advance this effort. Further, the lab was only possible because of the diligent efforts of the NWS CONOPS team, development organizations, participants, and management who worked together to achieve this significant CONOPS milestone.

Thirty-five recommendations, based on the 42 findings, are provided for consideration by the National Weather Service. Of these 35, the 13 highlighted recommendations that contributed to the overall Critical Recommendations are deemed absolutely critical to the successful accomplishment of the CONOPS Prototype Plan.

If these recommendations are accepted and acted upon, the CONOPS Prototype Team believes that:

- Clustered peer operations for two offices will likely work with some additional work to track ReAP activities.
- After further software development, clustered peer operations for 4 offices will probably work. A more sophisticated ReAP tracking tool will be necessary.
- GFE Re-architecture work must begin in Phase 1 if it is to be ready for the Service Oriented Architecture (SAO) of AWIPS-next in late FY09. This is necessary to expand the size of the domain, the number of offices in a cluster, and to determine the national cluster configuration

6.0 References

CONOPS Prototype Team, September 2006: *The DRAFT National Weather Service Clustered Peer Concept of Operations Prototype Plan*, Fiscal Years 2007-2009, 40pp.

CONOPS Prototype Team, November 2006: *The National Weather Service Clustered Peer Concept of Operations Prototype Plan*, Fiscal Years 2007-2009, 45pp.

ESRL Global Systems Division, Information Systems Branch, August 2006: *FY2006 NWS Concept of Operations Pre-Prototype Laboratory Exercise Plan*, 23pp.

Cheatwood-Harris,2006: CONOPS Lab Exercise System Feedback from On-Line Questionnaires, Catalog of Responses, 60pp.

Lusk, C.M., P. Kucera, W. Roberts, and L. Johnson, 1999: The Process and Methods Use to Evaluate Prototype Operational Hydrometeorological Workstations, Bull. Amer. Meteor. Soc., 80, 57-65.

Appendix A: OSIP Plans

OSIP project number 06-058 documents the overarching CONOPS Statement of Need (SON) and Project Plan. This SON is intended as an "umbrella" SON for other, more specific OSIP projects which provide the foundation for further investigating the CONOPS that was developed and which the Corporate Board approved for testing. Some of these projects are anticipated to include, but are not limited to:

- Expanded domains (e.g. multiple WFO domains: SON #06-059);
- Resource allocation (SON #06-060); and,
- Performance assessment and feedback (SON #06-061).

Additionally, there are existing OSIP projects which will directly support the implementation of the CONOPS. Some of these include, but are not limited to:

- AWIPS evolution Umbrella (SON #04-005) and subordinate projects:
- Data Delivery Paradigm (SON #05-040)
- Visualization Techniques (SON #05-021)
- Information Generation (SON #05-041)
- NWS Collaboration (SON #05-042)
- GFE performance enhancements (SON #06-051)
- Analysis of Record (SON #05-009)
- Accelerate Environmental Modeling & Prediction Capabilities (SON #05-065)
- Verification & Performance Management System (SON #05-032)
- Downscaled NWP Grids of Sensible Hydrometeorological Elements (SON #06-041)

Appendix B: Summary of Numerical Ratings from CONOPS '06 End-*of-Shift Questionnaire*

Listed below are the questions and summary of responses to rating questions from the CONOPS'06 End-Of-Shift Questionnaire. Below each question is a reference code and summary statistics: Mean, Standard Deviation (Std), Minimum rating (Min), Maximum rating (Max), and Number of Responses (N).

Questions and User Response Summary:

3. Rate the performance of D2D from the ConOps Lab Exercise.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3.36, Std=0.58, Min=2, Max=5, N=42

4. Rate the performance of GFE from the ConOps Lab Exercise.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.64, Std=1.01, Min=1, Max=4, N=47

5. Rate the performance of FX-C from the ConOps Lab Exercise.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3.55, Std=0.70, Min=2, Max=5, N=44

6. Rate the performance of 12-Planet from the ConOps Lab Exercise.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.25, Std=1.04, Min=1, Max=4, N=8

7. Rate the performance of COMMS from the ConOps Lab Exercise.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3.16, Std=0.49, Min=1, Max=4, N=37

Appendix C: Summary of Numerical Ratings from CONOPS '06 End-of-Week Questionnaire

Listed below are the questions and summary of responses to rating questions from the CONOPS'06 End-Of-Week Questionnaire. Below each question is a reference code and summary statistics: Mean, Standard Deviation (Std), Minimum rating (Min), Maximum rating (Max), and Number of Responses (N).

<u>Questions and User Response Summary:</u>

1. D2D

1a. Rate the D2D System Performance.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3.67, Std=0.5, Min=3, Max=4, N=9

1b. Rate the D2D Product Selection.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3.11, Std=0.78, Min=2, Max=4, N=9

1c. Rate the D2D Product Use/Display.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3.33, Std=0.71, Min=2, Max=4, N=9

1d. Rate the D2D Cluster Ops Issues.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3, Std=1, Min=1, Max=4, N=9

1e. Rate the D2D Screen Mgt.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.78, Std=0.67, Min=2, Max=4, N=9

1f. Rate the D2D Text Windows.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3, Std=0.71, Min=2, Max=4, N=9

1g. Rate the D2D Training.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.88, Std=0.83, Min=1, Max=4, N=8

2. GFE

2a. Rate the GFE performance.

(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.6, Std=0.84, Min=1, Max=4, N=10

2b. Rate the GFE Usability.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.9, Std=0.99, Min=1, Max=4, N=10

2c. Rate the GFE Service Backup.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3.6, Std=0.52, Min=3, Max=4, N=10

2d. Rate the GFE GHG.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.8, Std=0.79, Min=2, Max=4, N=10

2e. Rate the GFE Training.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3, Std=0.94, Min=2, Max=5, N=10

3. Coordination Tools

3a. Rate the 12-Planet Chat.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=1.67, Std=0.58, Min=1, Max=2, N=3

3b. Rate the 12-Planet White Board.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=1, Std=0, Min=1, Max=1, N=4

3c. Rate the FX-C.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3.67, Std=0.87, Min=2, Max=5, N=9

3d. Rate the Coordination Training.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3.2, Std=1.03, Min=2, Max=5, N=10

4. Communication

4a. Rate the Communications Adequacy.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.8, Std=1.32, Min=1, Max=5, N=10

5. Other (i.e. hydrologic applications and other non-GFE AWIPS applications such as SafeSeas and SCAN)

5a. Rate the Basic Usability/Utility.

(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.5, Std=0.71, Min=2, Max=3, N=2

5b. Rate the User Interface.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.5, Std=0.71, Min=2, Max=3, N=2

5c. Rate the Screen Mgt.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2, Std=0, Min=2, Max=2, N=3

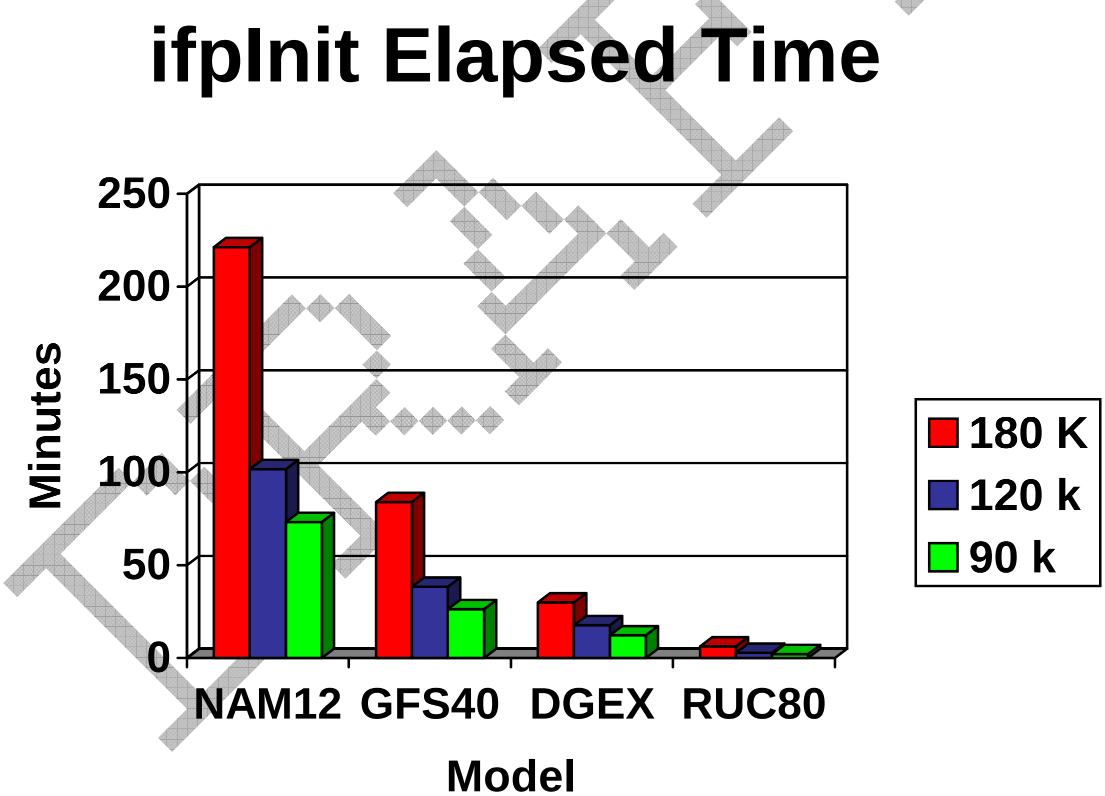
5d. Rate the Training.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=2.67, Std=0.58, Min=2, Max=3, N=3

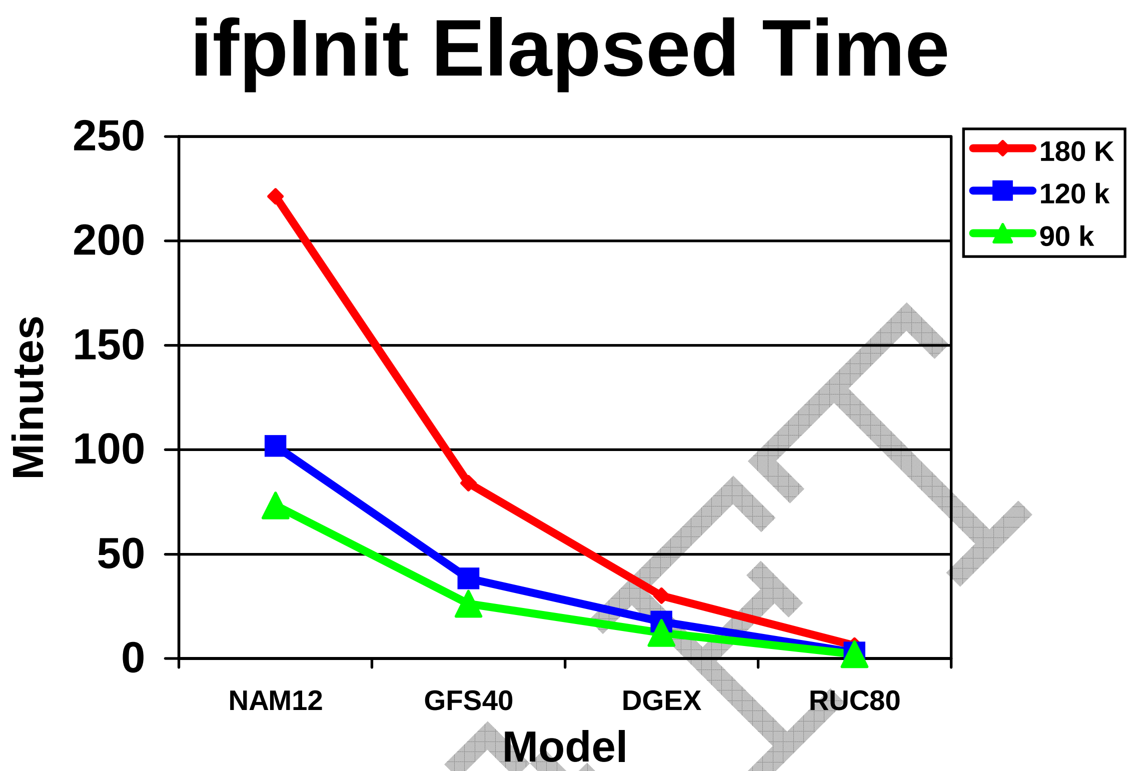
6. Additional Questions

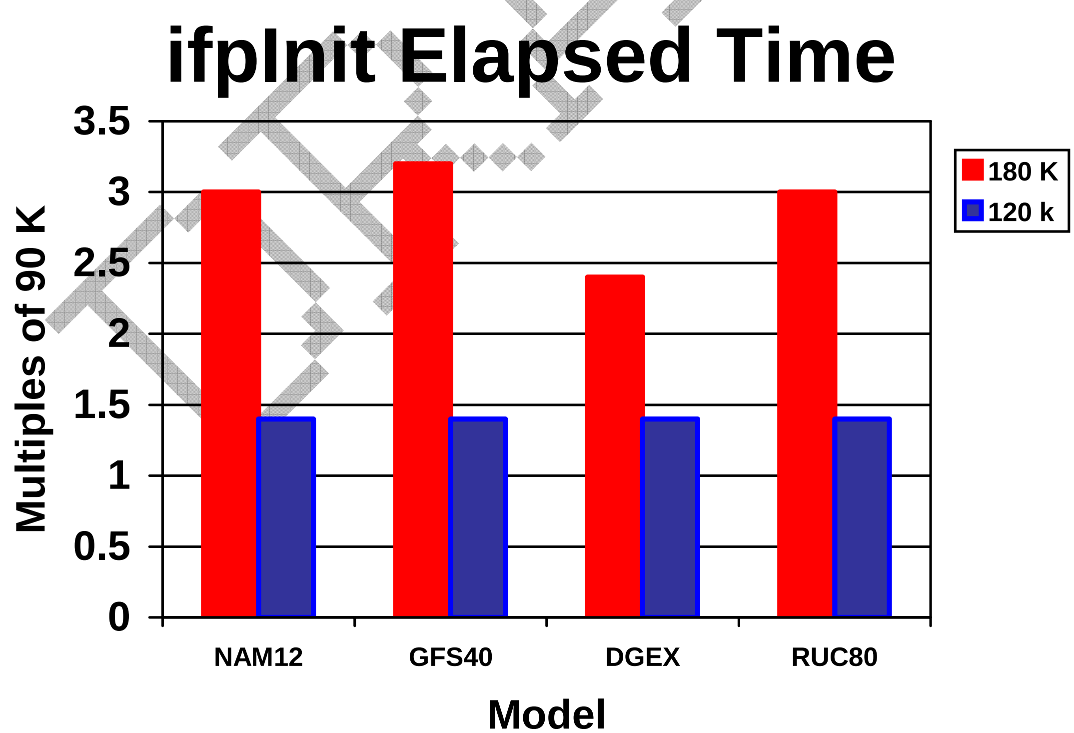
Rate the conduct of the ConOps Lab Exercise.
(1=Inadequate, 3=Acceptable, 5=Exceptional)
Mean=3.7, Std=0.95, Min=2, Max=5, N=10

Appendix D: GFE Grid initialization comparison results from ifpINIT timing test

The following graphs were generated prior to the start of the CONOPS lab in order to determine possible performance differences using different grid sizes ranging from 90,000 to 180,000 grid points. 90,000 grid points represented the high end of current operational areas (750 km X 750 km at 2.5km spacing). The tests were run using ifpINIT programs generating weather elements initialized from different NECP models. The tests showed a non-linear performance change as the number of grid points increased. For example, a doubling of grid points, from 90k to180k, more than tripled ifpINIT times for the GFS40. Based on this evaluation, the CONOPS team decided to set an upper bound of ~120k grid points for this initial lab.




ifpInit Elapsed Time
250
200
150
100
50
0
Minutes
180 K
120 k
90 k
NAM12
GFS40
DGEX
RUC80
Model


ifpInit Elapsed Time
3.5
3
2.5
2
1.5
1
0.5
0
Multiples of 90 K
180 K
120 k
NAM12
GFS40
DGEX
RUC80
Model

Appendix E. Wide Area Network Performance Assessment for Lab Exercise

Introduction and Background

The lab exercise demonstrating clusters of forecast offices exercised the exchange of data for collaboration beyond the current state of the AWIPS system used operationally at WFO sites. Two applications in particular exercised the network bandwidth: GFE and FXC. The FXC application was used for collaborating and defining the current Areas of Responsibility (AOR). FXC requests a “picture” from a server, compresses it, then sends the results via a socket to a client. The information can be shared among many clients, creating a simultaneous D2D-like visualization of hydromet data. Annotations and drawing primitives can be added to the data and exchanged as well. The GFE application increased the domain of grids exchanged for inter-site coordination supporting the execution of each site’s forecast-preparation responsibilities as well as the NDFD. IT design of the new cluster concept identified increased network activity as a potential repercussion from the cluster concept.

The current AWIPS WAN uses a frame-relay system to transfer data between sites. A hierarchical hub technology is employed, requiring at least two hops when transferring data between WFOs. There was a desire to have no impact on operations coupled with the need to run the lab exercise on a system internal to the AWIPS WAN. The new NOAANet MPLS system was due to be deployed to pilot sites during the summer prior to the lab exercise. The CONOPS team members were able to accelerate the incorporation of FSLC and BCQ into the MPLS pilot. This was a critical step to enable potentially voluminous exchange of information between the two lab exercise sites (FSLC and BCQ) while minimizing the risk to NWS operations.

NOAANet and the New CONOPS

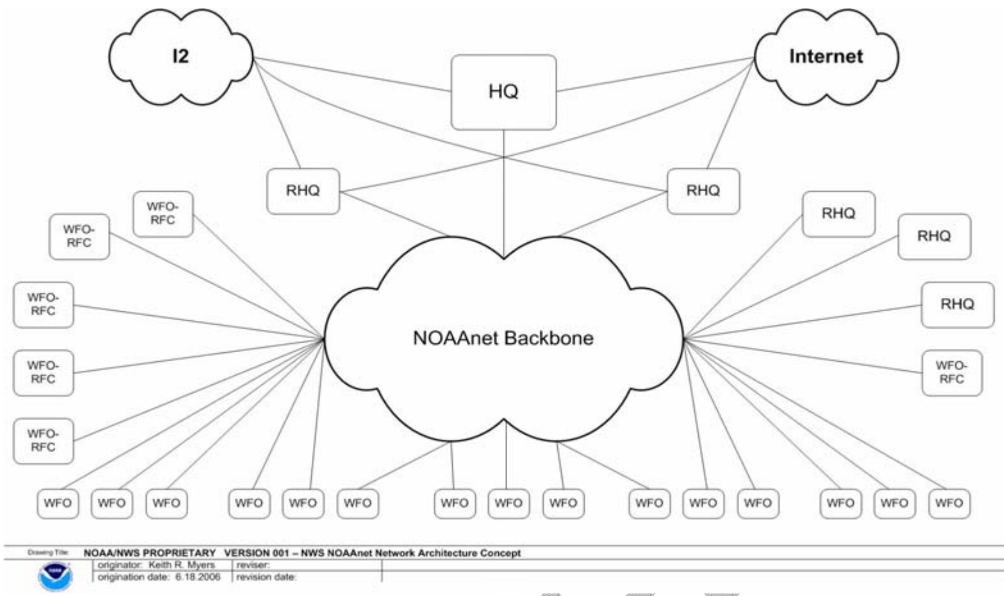
The NOAANet and NWS CONOPS activities, while two separate projects, are not completely independent from each other. The definition and implementation of the new NWS CONOPS will be influenced by the NOAANet project. The NOAANet project, in turn, must accommodate requirements that stem from the evolving CONOPS.

NOAANet will exploit Multi-Protocol Label Switching (MPLS) Virtual Private Network (VPN) technology and it can support multiple VPNs. It is anticipated that NOAANet will establish a new cost-effective, agency-wide network infrastructure by converging and consolidating legacy networks onto a common, centrally-managed, modern system. There are several aspects of the NOAANet transition that have particular relevance to our future CONOPS:

- the replacement of the legacy hub/spoke circuits with potential “any-to-any” connections,

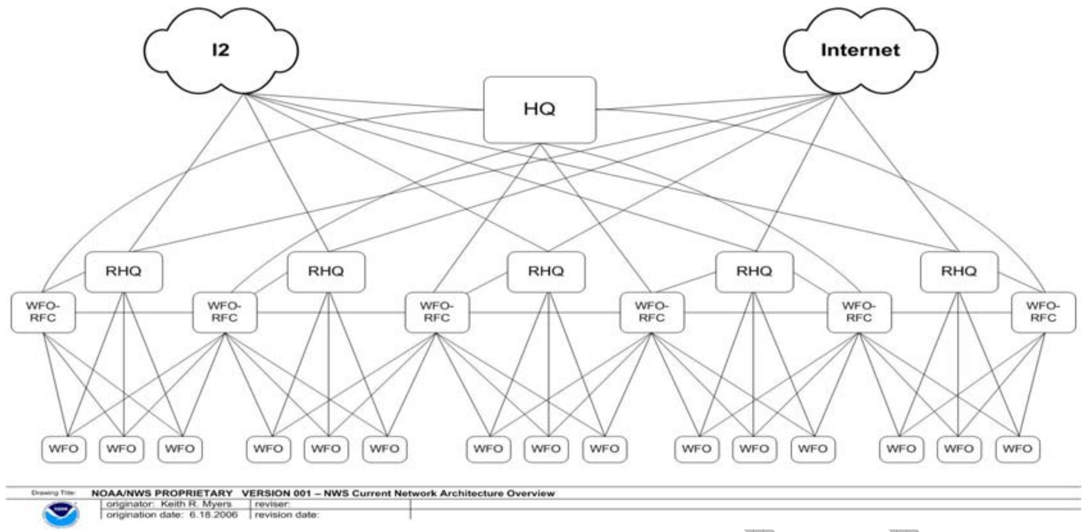
- improved scalability (i.e., potential for increases in network bandwidth to support escalating inter-site product exchange requirements and or to improve product-delivery timeliness), and,
- improved reliability (e.g., fewer critical points of failure).

A depiction of NOAANet is shown below.

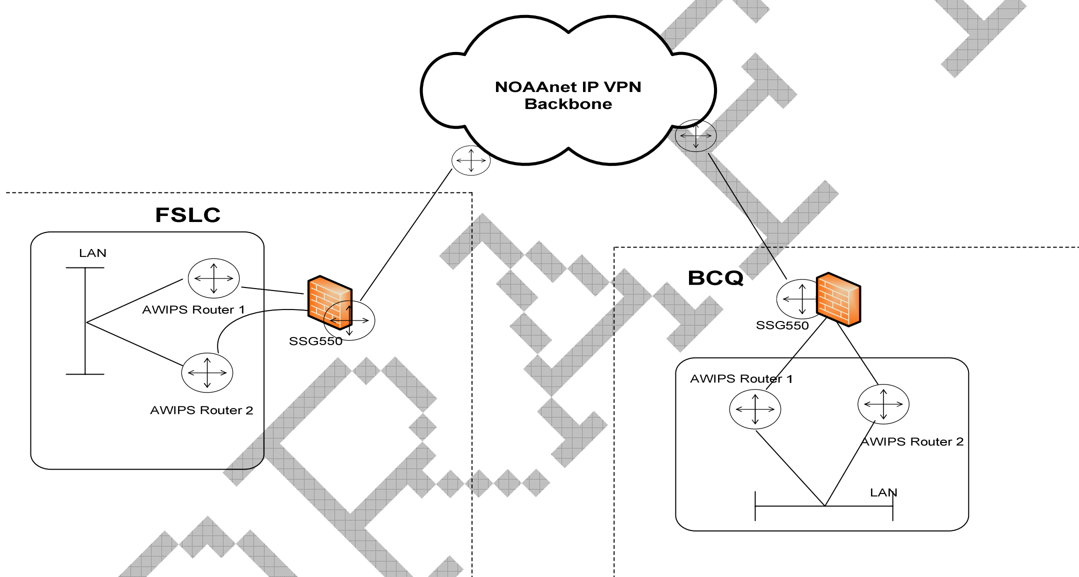


Other benefits of the MPLS will include facilitated continuity of operations for disaster recovery, support of network-based emergency voice communications, and economy-of-scale efficiencies gained by consolidated security, monitoring and management. Advanced features such as multicast addressing may be applied for ISC traffic in order to mitigate major increases in quantity of grids distributed on NOAANet and should be considered in future solutions.

A depiction of the current AWIPS hub-and-spoke WAN is shown below (with the NWS regional WAN also shown). In the AWIPS WAN, Co-located WFO/RFCs serve as regional hubs whereas non-collocated WFOs are spoke-terminating nodes. In the Regional WAN, the Regional Headquarters offices are the hubs.



Legacy AWIPS and NWS Regional WANs – Hub and Spoke Topology



One deficiency of the legacy AWIPS WAN is its limited bandwidth— especially to non-collocated WFOs. Another deficiency is the dissemination in direction inherent in the WAN's hub/spoke architecture (e.g., where WFO-to-WFO communications must pass through one or more intermediary RFC-collocated WFOs in transit). Meanwhile, the "any-to-any" MPLS NOAANet is less constrained. The WAN connectivity for the CONOPS IT lab test was provided in conjunction with the MPLS AWIPS pilot. This connectivity was accomplished by implementing a peer-to-peer or layer-3 VPN model where the customer edge (CE) router exchanged IP routing information with the Sprint router. The WAN IP backbone supported features such as unicast, GRE, IPSec and VRF services. The use of the MPLS pilot in the CONOPS IT lab test is illustrated below.

In order to conduct the lab exercise and not impact the NOAANet pilot test, a separate VPN was established between FSLC and BCQ.

Both AWIPS nodes in the lab test (i.e., FSLC/Boulder and BCQ/Kansas City) were configured with their WAN interfaces (i.e., SSG550) connected to a unique CONOPS VPN backbone for access via a single VPN peering point. Each NOAANet access point was assigned a private autonomous system number that was used for WAN connections to the network-service provider (Sprint). On the WAN connection, the standard routing protocol used on NOAANet was Exterior Border Gateway Protocol (BGP). On the LAN inward-facing connections, Open Shortest Path First (OSPF) was configured on the CE router to establish communications at each site. It is expected that the AWIPS WAN will evolve into a VPN within NOAANet, with an AWIPS/MPLS interface similar to that evaluated during the lab test. However, the precise site-to-site bandwidths to be provided with MPLS are not fully determined at this time. For the lab test, the FSLC/BCQ MPLS link capacity was approximately 45Mbps. It is important to note that this capacity is significantly greater than that of the existing/legacy AWIPS WAN, and also likely greater than the foreseeable-future MPLS WAN capacity. The current, planned link capacities are tabulated below. The network utilization analysis for this lab test factors in the difference between the capacity available during the test (i.e., 45Mbps) and the expected NOAANet/MPLS site-to-site link capacity.

The NOAANet introduces a peerless-IP cloud (PIP). The concept of point-to-point communication is not used.

Link Capacity*

	Legacy AWIPS WAN	MPLS – Planned
Spur (RFC/WFO-to-WFO)	356 Kbps	
Hub (RFC/WFO-to-RFC/WFO)	756 Kbps	
Trunk (NCF-to-RFC/WFO)	1.5 Mbps	
WFO- PIP PIP		>1.5 Mbps
RFC- PIP		>3 Mbps
NC- PIP		>3 Mbps

** Note: Legacy bandwidths are bidirectional, and input/output cannot exceed ½ of link capacity.*
*** Note: Because of the new NOAANet configuration, the current hub topology will no longer apply. National Centers are treated like a WFO in the above table.*

Monitoring/Measuring Tools and Preliminary Results

NetScout *nGenius*® performance manager was the principal tool used to monitor and assess MPLS utilization during the lab test. The tool has a data-collection probe and can produce graphical representations of time-dependent network utilizations. The primary views chosen for inclusion in this report are:

1. Top Ten Applications Over Time
2. Total Link Utilization Over Time

The FSLC graphs (see “Application Breakdown” subtitles) show the top applications all utilizing in the range 0% to 0.2% of the total available MPLS bandwidth. These percentages are based on the tested bandwidth of 40Mbps. Thus, NetScout-reported values in the 0-0.2% utilization range correspond (for a network with a capacity of 1.5Mbps) to utilizations in the 0-6% range.

The ranges above are based on hourly samples, which average periods of heavy and lighter traffic. It is very likely that during the lab test, short-term MPLS utilizations temporarily exceeded the 6% utilizations. It is recommended that the utilization be more thoroughly assessed during future lab tests by configuring NetScout with a morefrequent sampling interval. Furthermore, the NetScout performance measurements should be analyzed and compared to data volumes recorded by application-level logging.

Test Points

The network usage has two impacts on operations. How quickly can a warning get from one site to a neighbor over the terrestrial network? A saturation of the network bandwidth would result in delays of disseminating warnings terrestrially. Also, flooding the MHS at a site may delay disseminating a warning. The current system does not have any message interleaving functionality, so once a large message is started there is no exception handling to interrupt and send a more important, higher priority message.

Network bandwidth was measured using the aforementioned tools. The message delays within the MHS were not measured during this exercise.

The following information was scaled down from the lab exercise conducted on 45Mbps intersite bandwidth. The projected WFO-to-WFO bandwidth will be 1.5Mbps in NOAANet, so all calculations have been scaled by $1/30^{th}$. In conversions from static file sizes (expressed in bytes) to transmission rates (expressed in Mbps), it is assumed that one byte corresponds to 10 bits. This correspondence accounts for some additional network-transmission overhead, beyond the geophysical data elements (e.g. grids) themselves.

- **Network Bandwidth**

  The bandwidth on the FSLC network is 45Mbits/sec. It is anticipated that theWFO sites will have a bandwidth of approximately 1.5Mbps/sec.

  The graphs attached for the 4 weeks of the exercise show that the ISC traffic used an average of less than .01% of the bandwidth between FSLC and BCQ, scaled to 45Mbps. However, a WFO will only have roughly 3% of this bandwidth. A simple analysis showed that roughly 2 ISC grids could be sent per second at 1.5Mbps if the grid size is 68,000 bytes. ISC grid sizes span a wide range, due to compression and variation in data. Also, the GFE software combines grids into one message, sometimes exceeding 10MBytes per message.
- **Message Delay**

  GFE grid files representing a 120,000 point grid compress to roughly 68KBytes.

These compressed grids are used for ISC. At 1.5Mbps, a grid takes roughly 0.45 seconds to transfer. Therefore, it is assumed that a single lower-priority grid could potentially delay a higher-priority warning by up to 0.45 sec.

Recommendations

- The MPLS upgrade was delayed until two days prior to the start of the lab exercise. A longer learning curve of the tools for analysis would have been helpful.

  Action: Repeat lab exercise after analysts have more experience with the network, its troubleshooting procedures, and its performance-monitoring tools.

- Based on the very simple lab tests, the AWIPS MPLS network should easily accommodate the extra traffic incurred by the larger-site grid sizes and an increased quantity of intersite-exchanged grids. However, the MPLS bandwidth available during the lab exercise was much greater than that which will likely be available to sites in an operational setting. Furthermore, during the lab exercise, only a subset of an office's daily activities (and data exchange) were simulated. Thus, the lab exercise's network data volumes and usage patterns have only limited resemblance to the network information exchange that might be seen in a fully-functioning cluster.

  Action: Lab was limited to site-to-site traffic for just ISC and FXC traffic. Repeat lab exercise with more realistic WFO scenario adding radar products and other normal WFO functions on a network supporting representative bandwidth.

- Preliminary results from the lab exercise suggest the AWIPS WAN— especially in the MPLS/VPN era— can support at least some degree of clustering. This initial lab test suggests that clustering would not greatly add to WAN message-dissemination delays. However, as noted previously, the system-performance attributes of this lab test are not readily mapped to those of a full, functioning cluster.

  Action: Repeat lab exercise with additional test points geared toward evaluating latency in disseminating a W/W/A type message. Use a metric such as the "StoreDelay" that is computed weekly by Raytheon's Performance Working Group to measure delays queuing information for SBN transmission. The impact of a 10Mbyte message needs to be determined in a real-life scenario. Hand analysis of a message of this size indicates large latencies in message handling. In depth understanding and application usage of the MHS is critical for this next phase. Understanding of priority queuing is imperative.

- We were unable to thoroughly assess the 12Planet Whiteboard capability during this initial lab test. This tool— and the network traffic it generates-- should be assessed in any subsequent testing.

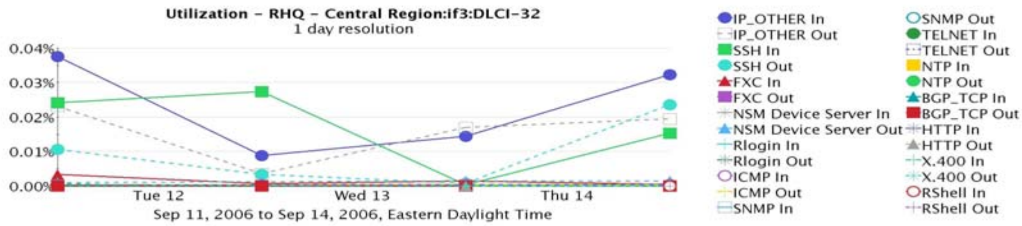
  Action: Install and configure 12Planet's Whiteboard capability before the next lab test.

FSLC/BCQ MPLS Utilizations
NetScout *nGenius*® Performance Manager

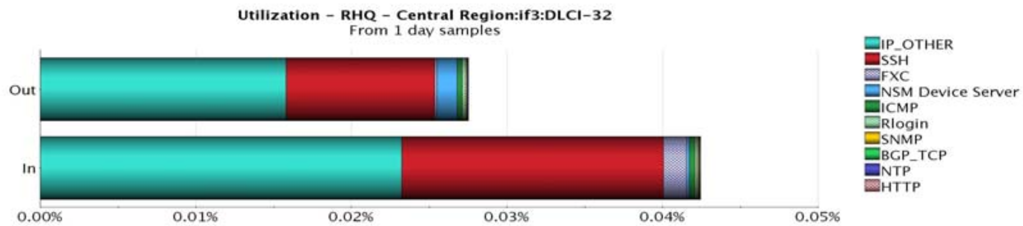
nGenius Performance Manager | **FSLC (Week1)** | NetScout

Source Server: ngenius | Report Time: Oct 16, 2006 11:12:48 AM EDT

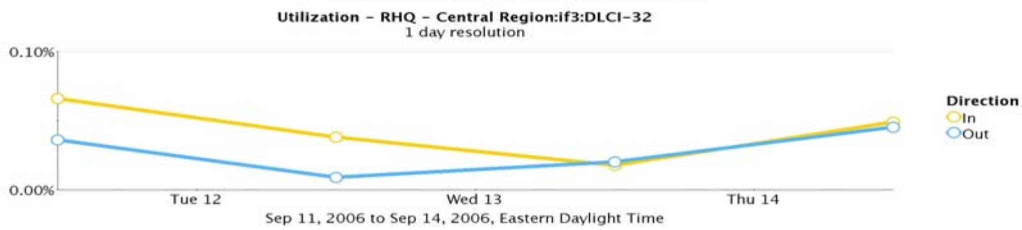
Top 10 Applications over Time - 4 (RHQ - Central Region:if3:DLCI-32)
9/11/06 12:00:00 AM EDT - 9/15/06 12:00:00 AM EDT



Top 10 Applications - 3 (RHQ - Central Region:if3:DLCI-32)
9/11/06 12:00:00 AM EDT - 9/15/06 12:00:00 AM EDT

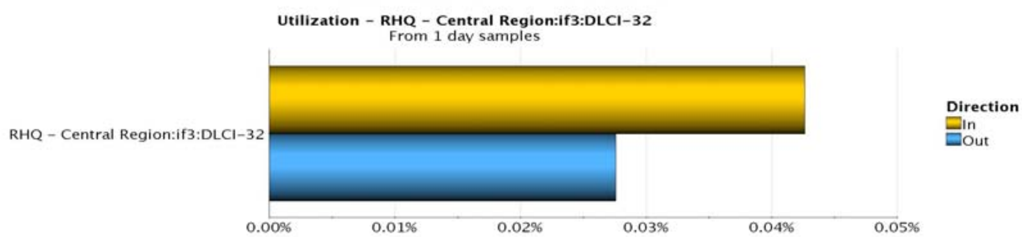


Link Usage over Time - 3 (RHQ - Central Region:if3:DLCI-32)
9/11/06 12:00:00 AM EDT - 9/15/06 12:00:00 AM EDT



FSLC (Week1) | Report Time: Oct 16, 2006 11:12:48 AM EDT

Link Layer Snapshot - 2 (RHQ - Central Region:if3:DLCI-32)
9/11/06 12:00:00 AM EDT - 9/15/06 12:00:00 AM EDT

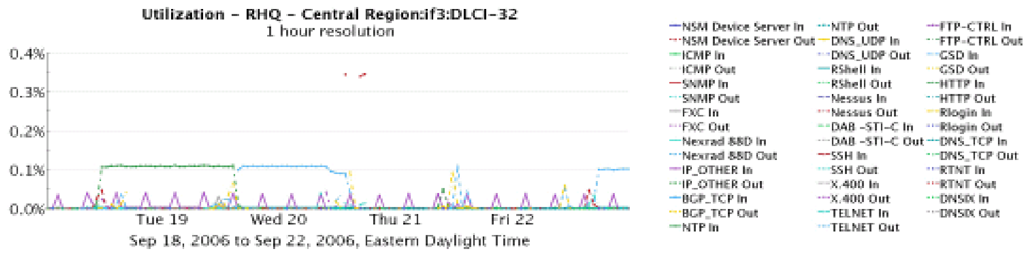


nGenius Performance Manager **FSLC Week2 - Application Breakdown** NetScout

Source Server: ngenius Report Time: Oct 16, 2006 12:20:31 PM EDT

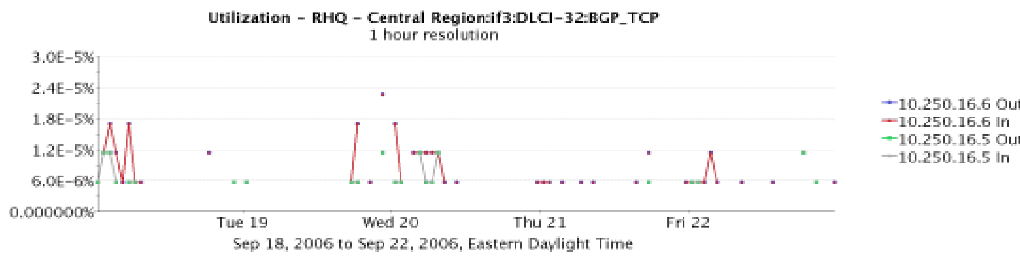
Top 10 Applications over Time - 12 (RHQ - Central Region:if3:DLCI-32)

9/18/06 12:00:00 AM EDT - 9/23/06 12:00:00 AM EDT



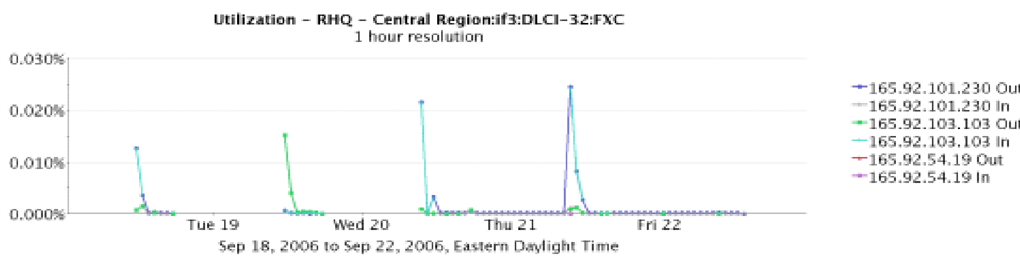
Top 10 AL Hosts over Time - 14 (RHQ - Central Region:if3:DLCI-32) [BGP_TCP]

9/18/06 12:00:00 AM EDT - 9/23/06 12:00:00 AM EDT



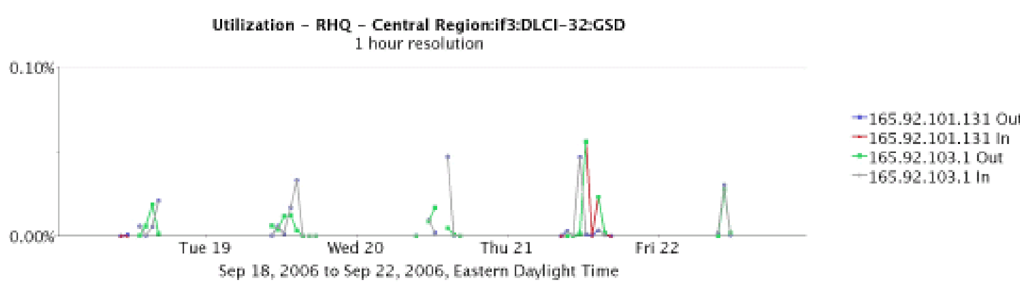
Top 10 AL Hosts over Time - 9 (RHQ - Central Region:if3:DLCI-32) [FXC]

9/18/06 12:00:00 AM EDT - 9/23/06 12:00:00 AM EDT



Top 10 AL Hosts over Time - 8 (RHQ - Central Region:if3:DLCI-32) [GSD]

9/18/06 12:00:00 AM EDT - 9/23/06 12:00:00 AM EDT

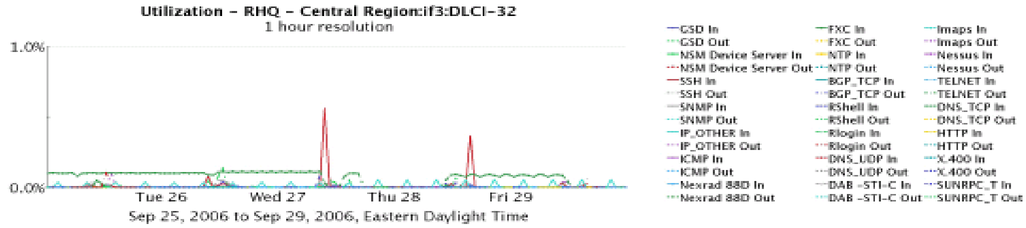


nGenius Performance Manager | **FSLC Week3 - Application Breakdown** | NetScout

Source Server: ngenius | Report Time: Oct 16, 2006 12:18:24 PM EDT

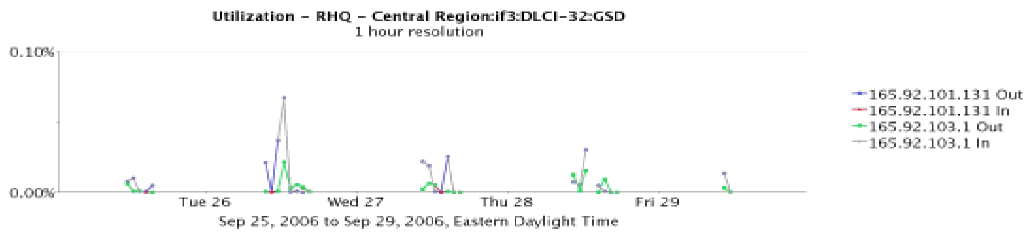
Top 10 Applications over Time - 13 (RHQ - Central Region:if3:DLCI-32)

9/25/06 12:00:00 AM EDT - 9/30/06 12:00:00 AM EDT



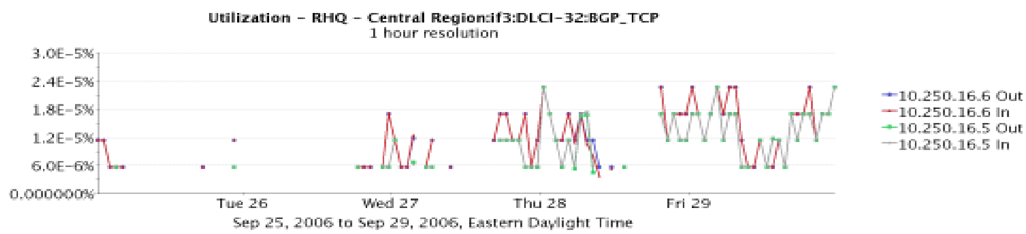
Top 10 AL Hosts over Time - 21 (RHQ - Central Region:if3:DLCI-32) [GSD]

9/25/06 12:00:00 AM EDT - 9/30/06 12:00:00 AM EDT



Top 10 AL Hosts over Time - 20 (RHQ - Central Region:if3:DLCI-32) [BGP_TCP]

9/25/06 12:00:00 AM EDT - 9/30/06 12:00:00 AM EDT

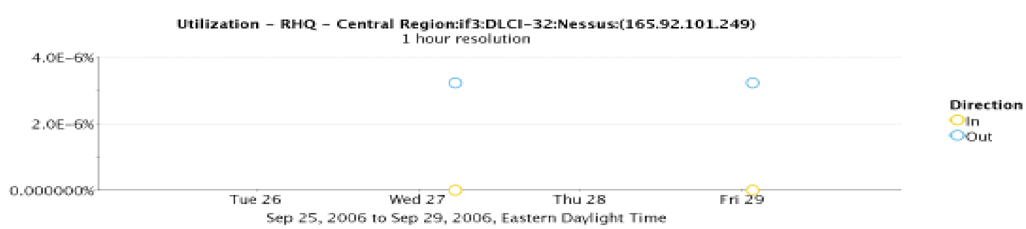


Top 10 AL Hosts over Time - 19 (RHQ - Central Region:if3:DLCI-32) [DAB -STI-C]

9/25/06 12:00:00 AM EDT - 9/30/06 12:00:00 AM EDT

Top 10 AL Hosts over Time - 16 (RHQ - Central Region:if3:DLCI-32) [Nessus]

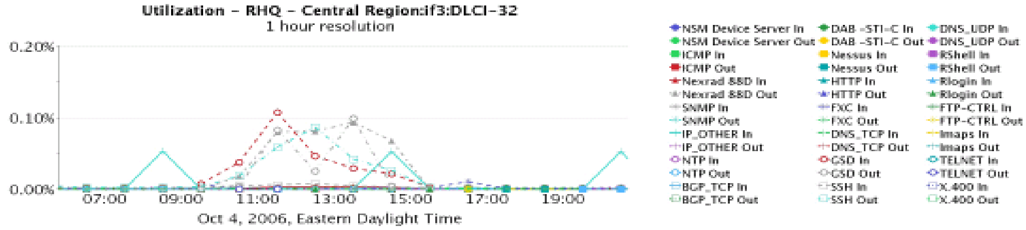
9/25/06 12:00:00 AM EDT - 9/30/06 12:00:00 AM EDT



Source Server: ngenius   Report Time: Oct 16, 2006 12:16:47 PM EDT

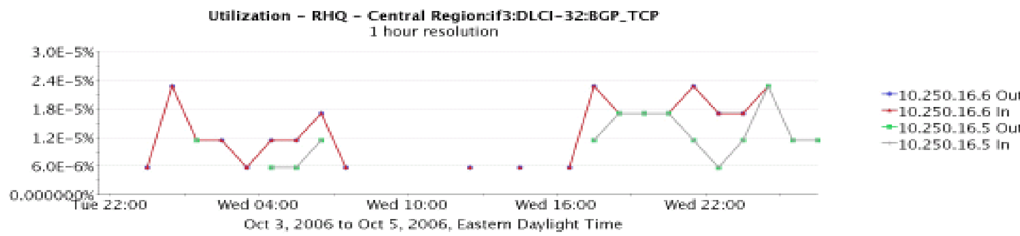
Top 10 Applications over Time - 14 (RHQ - Central Region:if3:DLCI-32)

10/2/06 12:00:00 AM EDT - 10/7/06 12:00:00 AM EDT



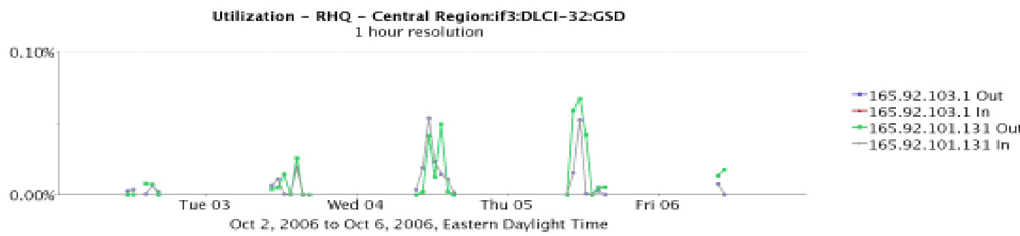
Top 10 AL Hosts over Time - 35 (RHQ - Central Region:if3:DLCI-32) [BGP_TCP]

10/3/06 9:00:00 PM EDT - 10/5/06 3:00:00 AM EDT



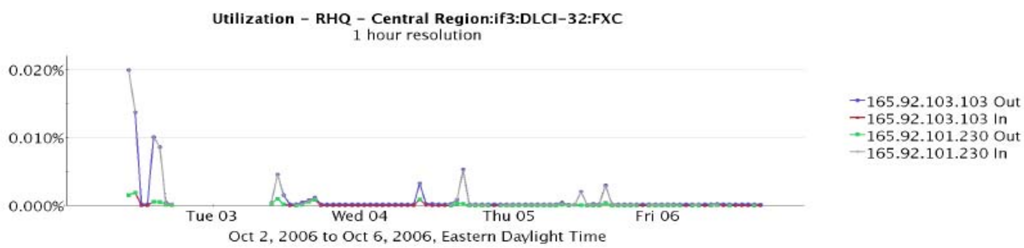
Top 10 AL Hosts over Time - 32 (RHQ - Central Region:if3:DLCI-32) [GSD]

10/2/06 12:00:00 AM EDT - 10/7/06 12:00:00 AM EDT



Top 10 AL Hosts over Time - 31 (RHQ - Central Region:if3:DLCI-32) [FXC]

10/2/06 12:00:00 AM EDT - 10/7/06 12:00:00 AM EDT



Top 10 AL Hosts over Time - 30 (RHQ - Central Region:if3:DLCI-32) [X.400]

10/2/06 12:00:00 AM EDT - 10/7/06 12:00:00 AM EDT

Appendix F: Compiled Findings and Recommendations

[**Findings** and **recommendations** that contribute to the overall critical findings and recommendations are identified with an asterisk (*).]

Finding 1: In general, participants found the D2D performance to be better than acceptable (average 3.36 on the daily questionnaire and 3.67 on the weekly questionnaire on a 5-point scale with 3.0 being acceptable) for supporting simulated CONOPS operations.

Finding 2: Participants liked the AOR map backgrounds but wanted an automated process to change backgrounds when the WFO's AOR changed.

Finding 3: Participants need access to all locally-generated D2D procedures, data, and model output, for CONOPS operations.

Finding 4: D2D viewing scales and LAPS were not always properly aligned with the CONOPS cluster areas.

Finding 5: Participants need training on D2D changes when upgrades to the system occur.

Finding 6: Screen space became problematic with more applications and larger areas of responsibility.

Recommendation 1: D2D is acceptable for further CONOPS testing with minor enhancements and configuration changes to better accommodate cluster operations and cluster spatial areas.

Recommendation 2: Automate CWA/AOR map background changes to reflect AOR changes as they occur.

Recommendation 3: Consider adding more screens or larger screens to AWIPS to improve screen management characteristics of the system.

Finding 7: Participants successfully generated forecast grids and text products over the cluster CWA areas using the enhanced GFE capabilities.

Finding 8: Participants rated the GFE capabilities as somewhat less than acceptable (average 2.64 on the daily questionnaire and 2.60 on the weekly questionnaire on a 5-point scale with 3.0 being adequate) when used to support CONOPS testing.

***Finding 9**: Slow GFE performance, running over the 120,000 grid point cluster size domains, was a primary cause of the less than acceptable GFE rating.

***Recommendation 4**: Consider smaller cluster domain sizes for future testing as well changes to the GFE which would improve performance of the software.

Finding 10: A subset of Smart Tools and Procedures was successfully modified and used to generate forecast grids during the exercise.

Finding 11: Participants need a more complete set of their offices' Smart Tools and Procedures to efficiently generate forecast grids for their home and cluster CWAs.

Recommendation 5: Integrate a more complete suite of each office's Smart Tools and Procedures before future CONOPS testing.

***Finding 12**: Smart Tools and Procedures ran significantly slower over the cluster-sized expanded domain.

***Recommendation 6**: Explore and adopt options to improve Smart Tools and Procedures performance by considering Smart Tools software changes, run-time configuration changes, and domain size changes.

Recommendation 7: Proposed cluster offices should determine a common suite of Smart Tools and Procedures that would effectively run over their entire cluster domain.

***Finding 13**: ISC performance was often noted as problematic during the exercise.

***Finding 14**: Performance of ISC-related programs within the GFE (eg. ISCMosaic) was the likely cause of the ISC performance problems, not network bandwidth.

***Recommendation 8**: Review ISC-related programs within the GFE to determine what improvements can be made generating ISC mosaic fields.

***Recommendation 9**: Improve ISC-Grid status information to include whether grids do or do not need to be sent, or are currently being sent, received, and processed.

Recommendation 10: Future testing should better emulate a full load of incoming and outgoing ISC traffic in order to better determine adequate bandwidth requirements.

Recommendation 11: Review the ISC process to determine whether other architectural frameworks can better accommodate ISC requirements and capabilities.

***Finding 15**: In general, participants found the performance of the formatters to be unacceptably slow during simulated CONOPS operations (e.g. when running formatters for multiple WFOs).

***Recommendation 12**: Consider smaller domains, better/faster hardware, and more efficient software as possible solutions to improve text formatter performance. It may also be possible to separate product generation from "assembly", i.e. by running one formatter containing multiple sites' products. This would require changes to the product editor.

Finding 16: The domain covered by various formatters (e.g. public vs. fire weather vs. marine) varies in some cases from office to office and product to product, resulting in complexities in editing grids and running the formatters for other WFOs in the cluster.

Finding 17: In particular, formatters for products associated with fire weather were the most problematic.

Recommendation 13: Properly configure domain size and setup of formatters to improve fire-weather product generation during prototype activities.

Finding 18: Better tracking, status monitoring, and product-generation automation capabilities are needed in AWIPS to allow forecasters to manage the numerous products that must be generated for each WFO.

Recommendation 14: Change the product editor interface (e.g. replace tabs with a matrix) to alleviate tracking and monitoring problems.

Recommendation 15: Issues associated with differences between local WFO formatters vs. baseline or “cluster” formatters must be explored in future prototype activities.

Finding 19: Participants successfully generated a variety of hazards with associated VTECs and were able to switch between normal operations and service backup using the lab GHG capabilities.

Finding 20: Participants rated the GHG capabilities as slightly less than acceptable (average 2.8 on a 5-point scale with 3.0 being adequate) when used to support CONOPS testing.

Finding 21: The GHG monitor did not always display all hazards issued by an office.

Finding 22: A WFO’s marine and fire weather zones do not always align with the corresponding Public zone or CWA of another office. Furthermore, map boundaries within CWAs and between adjoining CWAs do not always align, all of which caused occasional erroneous hazard identifications.

Recommendation 16: Correct mapping and display problems noted during the lab and consider better alignment strategies for program-specific hazard areas of responsibility within each CWA.

Recommendation 17: Develop scripts and procedures to streamline and automate as much of the hazard product generation process as possible.

Finding 23: Having an open line on the telephone between the Boulder and Kansas City lab sites was very helpful in coordination.

Recommendation 18: Cluster Management Teams should investigate the viability of having open/dedicated telephone lines between peer offices within a cluster. Voice-Over-Internet-Protocol (VoIP) and other web-based technologies may be of use.

Finding 24: The 12Planet-Whiteboard application was provided at the last minute to the lab participants, with little or no configuration or optimization, and had minimal interactive collaboration capabilities. 12Planet-Whiteboard was not deemed useful in the lab.

Recommendation 19: Without the addition of significant interactive collaboration capabilities, 12Planet-Whiteboard is not recommended for prototype activities.

***Finding 25**: Additional capabilities for FX-C were requested to increase its usability (such as changing menus, access to GFE data, etc).

***Finding 26**: Forecasters prefer a single application or mode that includes chat, drawing, and screen/image sharing.

***Recommendation 20**: Pre-prototype development is needed in FX-C to accommodate the highest-priority requested capabilities.

Finding 27: The ReAP process worked well given the constraints of the lab (only 2 WFOs, generally benign weather).

Finding 28: There did not seem to be technical issues with the ReAP process itself that would prevent it from being used in WFOs during prototype testing.

Recommendation 21: ReAP must be tested under a larger variety of scenarios, and involving more offices, in order to refine the concept.

Finding 29: Simple audio technology (a dedicated phone line) proved very valuable to quick and efficient coordination between lab participants.

***Recommendation 22**: NWS should examine other popular audio-visual technologies for use in collaboration and the ReAP process (e.g. PC-based LiveMeeting/GoTo Meeting, streaming audio/video via web, etc).

Finding 30: A better way of monitoring the status of the division of forecast responsibilities between cluster offices is required.

***Finding 31**: There needs to be a way for forecasters to monitor the current status of ISC grids.

Recommendation 23: Developers should work with the lab and prototype participants to develop requirements for a monitoring and grid-status capability.

***Finding 32**: Aside from performance issues (e.g. system slowness due to expanded number of grid points), the expanded-domain, enhanced AWIPS as tested in the lab provides a superior service backup capability over current baseline capabilities.

***Recommendation 24**: Given developments that support performance enhancements, future prototype testing should test the feasibility of this form of service backup replacing the current baseline capability.

***Recommendation 25**: Given successful testing, this capability for service backup should be targeted for national implementation at the earliest opportunity (e.g. as early as OB8.3 in 2008).

Finding 33: Modifications to AWIPS to support expanded domain render it impossible for a non-prototype office to provide service backup to a prototype office.

Recommendation 26: Until national deployment is possible for this service backup capability, prototype clusters must be comprised of primary service backup pairs of offices.

Finding 34: Other AWIPS applications were not routinely used or evaluated during the exercise.

Recommendation 27: Other relevant AWIPS applications need to be systematically evaluated in future CONOPS exercises.

Finding 35: The network capacity for the CONOPS lab was significantly greater than that of the existing/legacy AWIPS WAN and, also, likely greater than the foreseeable-future MPLS WAN capacity by a factor of 30.

***Finding 36**: The exercise showed that the ISC traffic used an average of less than .01% of the bandwidth between FSLC and BCQ, thus network performance did not adversely affect operations during the lab.

Recommendation 28: Repeat lab exercise with a more realistic WFO scenarios adding radar products and other normal WFO functions on a network supporting representative bandwidth and with network monitoring and analysis capabilities in place.

***Finding 37**: More training is needed on the specific aspects of system changes from baseline AWIPS operations.

***Recommendation 29**: Specific training modules must be developed and delivered to prototype participants on the software tools (i.e. FX-C, D2D and GFE enhancements in support of clustered-peer operations), and on ReAP tools.

***Finding 38**: Training specific to collaboration (beyond methodology, to include culture) is needed.

***Recommendation 30**: Training on collaboration, beyond "knobology", extending to human factors and culture issues, is needed for prototype participants.

Finding 39: Cluster participants need training on the meteorological and geophysical characteristics for WFOs in the cluster, and on the varying products and customer needs.

Recommendation 31: CMTs develop and implement a training program to increase knowledge of local issues at all offices within the cluster prior to full cluster operations.

***Finding 40**: Many enhancements made to AWIPS to support clustered-peer operations will require potentially significant development to allow cluster and non-cluster offices to interact.

***Recommendation 32**: Require all offices within a prototype cluster to be comprised of primary service backup pairs.

***Recommendation 33**: Development agencies will need to resolve the issues associated with "legacy" and cluster versions of AWIPS capabilities and their interoperability.

Finding 41: Forecasters noted several configuration problems in the lab involving map backgrounds, data mosaics, and domain problems with other AWIPS applications (e.g. LAPS).

Finding 42: Office-to-office differences in GFE weather elements, product coverage areas, edit areas, and other configuration items will need to be resolved prior to field prototyping.

Recommendation 34: Non-GFE configuration issues must be resolved by the CMTs prior to prototyping; GFE issues will require technical solutions by the developers.

Recommendation 35: Offices in the cluster prototypes need to be configured to receive additional radar data (either from the SBN or via the WAN from other ORPGs in the cluster).