Benchmark Testing the Digital Imaging Network–Picture Archiving and Communications System Proposal of The Department of Defense

Philip D. Shelton, David K. Lyche, Gary S. Norton, John Romlein, David P. Lawrence, Mike A. Cawthon, Jerry A. Thomas, and Nancy Richardson

The Department of Defense issued a Request for Proposal (RFP) for its next generation Picture Archiving and Communications System in January of 1997. The RFP was titled Digital Imaging Network—Picture Archiving and Communications System (DIN-PACS). Benchmark testing of the proposed vendors' systems occurred during the summer of 1997. This article highlights the methods for test material and test system organization, the major areas tested, and conduct of actual testing. Department of Defense and contract personnel wrote test procedures for benchmark testing based on the important features of the DIN-PACS Request for Proposal. Identical testing was performed with each vendor's system. The Digital Imaging and Communications in Medicine (DICOM) standard images used for the Benchmark Testing included all modalities. The images were verified as being DICOM standard compliant by the Mallinckrodt Institute of Radiology, Electronic Radiology Laboratory. The Johns Hopkins University Applied Physics Laboratory prepared the Unix-based server for the DICOM images and operated it during testing. The server was loaded with the images and shipped to each vendor's facility for on-site testing. The Defense Supply Center, Philadelphia (DSCP), the Department of Defense agency managing the DIN-PACS contract, provided representatives at each vendor site to ensure all tests were performed equitably and without bias. Each vendor's system was evaluated in the following nine major areas: DICOM Compliance; System Storage and Archive of Images; Network Performance; Workstation Performance; Radiology Information System Performance; Composite Health Care System/Health Level 7 communications standard Interface Performance; Teleradiology Performance; Quality Control; and Failover Functionality. These major sections were subdivided into workable test procedures and

were then scored. A combined score for each section was compiled from this data. The names of the involved vendors and the scoring for each is contract sensitive and therefore can not be discussed. All of the vendors that underwent the benchmark testing did well. There was no one vendor that was markedly superior or inferior. There was a typical bell shaped curve of abilities. Each vendor had their own strong points and weaknesses. A standardized benchmark protocol and testing system for PACS architectures would be of great value to all agencies planning to purchase a PACS. This added information would assure the purchased system meets the needed functional requirements as outlined by the purchasers PACS Request for Proposal.

Copyright © 1999 by W.B. Saunders Company

KEY WORDS: PACS, DICOM, benchmark testing, RIS, Health Level 7.

THE DEPARTMENT OF DEFENSE (DoD) has been actively involved in Picture Archiving and Communications Systems (PACS) since the early 1990s. PACS and teleradiology have been deployed into seven medical centers and fourteen smaller facilities from 1991 to 1996 using the Medical Diagnostic Imaging Support System (MDIS) contract. The Government began work in 1996 on a Request For Proposal (RFP) titled Digital Imaging Network–Picture Archiving and Communications System (DIN-PACS) to provide the Government with a second generation PACS contracting vehicle. The RFP was released in January of 1997. The Government team decided that benchmark testing of the vendors proposals would provide valuable information into the nature and functionality of each system. During the spring of 1997 development of the benchmark testing procedures and test equipment was completed. Testing of the vendors' proposed DIN-PACS systems took place during the summer of 1997. The following describes the benchmark testing process used in these evaluations.

MATERIALS AND METHODS

The benchmark testing team consisted of the core members who completed the DIN-PACS RFP. Members included active duty military and civilian contract employees. This group

Editor's Note: Unfortunately, the contract firm that did the testing will not reveal the minimum and maximum values comprising the range of answers across all vendors.

From the Joint Imaging Technology Project Office, Fort Detrick, MD; Informatech Inc, and the Defense Supply Center Philadelphia; the Radiology Department, National Naval Medical Center, Bethesda, MD; Radiology Department, Brooke Army Medical Center, San Antonio TX; and the Uniformed Services University of the Health Sciences, Department of Radiology, Bethesda, MD.

Address reprint requests to Lt. Col. Philip D. Shelton, MD, USAF, MC, Joint Imaging Technology Project Office, Bldg 1056, Fort Detrick, MD 21702-5012.

Copyright © 1999 by W.B. Saunders Company
0897-1889/99/1202-0007\$10.00/0

included three radiologists, two medical physicists, four biomedical engineers, two clinical engineers and one radiology technologist/RIS expert. Communications/Network experts were actively involved in the writing of the DIN-PACS RFP requirements, however they were not available for the benchmark testing.

PACS/teleradiology consultants from Informatech, Inc. (Frederick, MD) and personnel from Johns Hopkins University Applied Physics Laboratory (JHU-APL) developed test procedures based on RFP requirements. The test procedures essentially included every testable function in the RFP. All team members then reviewed these test procedures and changed them as needed before final approval.

The image server for the testing was developed by staff at JHU-APL. The server consisted of a Sun Sparc 5 (Sun Microsystems Inc, Palo Alto CA). The server was controlled by two PCs using hypertext markup language (HTML) script.

The Maryland Telemedicine Consortium gathered the images for benchmark testing. The image types included computed radiography (CR), computed tomography (CT), magnetic resonance imaging (MRI), single and multi-frame ultrasound, nuclear medicine, digital fluoroscopy, angiography and secondary capture. The Mallinckrodt Institute of Radiology, Electronic Radiology Laboratory verified that the images were Digital Imaging and Communications in Medicine (DICOM) standard version 3.0 (1) compliant.

Testing was performed on the vendor's equipment at a site of their choosing. This was typically done in a corporation's development laboratory. The testing team spent two 12 hour days with each vendor performing the testing. The test procedures were organized into nine major categories as outlined in Table 1.

These major categories were further subdivided into numerous smaller evaluations. Each test item was given a score of excellent (exceeds RFP requirements), good (meets RFP requirements), marginal (meets RFP requirements with limitations), or poor (does not meet RFP requirements). All the test scores for a section were compiled and an overall score using the same criteria was given for each section. The exact test procedures used is beyond the scope of this article. The following paragraphs will provide the reader with the basic functionality tested in each section.

DICOM Conformance testing included the DICOM standard, version 3.0, functionality and service classes listed in Table 2. DICOM verification is used by the PACS or other DICOM device to determine what DICOM functionality another PACS or DICOM device can perform. DICOM Storage Service class is a standard image format that allows two DICOM Storage Class capable systems to transfer images and header information.

Table 1. Benchmark Test Categories

1. DICOM Conformance
2. System Storage and Archive of Images
3. Network Performance
4. Workstation Performance
5. Radiology Information System (RIS) Performance
6. CHCS/HL7 Interface
7. Teleradiology Performance
8. Quality Control Functionality
9. Crisis Management

Table 2. DICOM Categories

1. Verification
2. Storage as a Service Class User (SCU) and Service Class Provider (SCP)
3. Query and Retrieve SCU and SCP
4. Storage Commitment
5. Worklist Management
6. Print

Service Class User's role is to initiate/request the required DICOM functionality. As an example, an acquisition modality requesting to send images to a PACS initiates the DICOM Storage Service Class functionality as the SCU. Service Class Provider's role is to acknowledge the DICOM request from the SCU and to perform the requested DICOM function. As an example, a PACS receiving a DICOM Storage Service Class request from an acquisition modality (SCU) which acknowledges the request and performs the DICOM functionality is the SCP. DICOM Service Class Query allows separate DICOM systems to view the exam listing of the other system and then to move the exams via DICOM Service Class Retrieve. DICOM Service Class Storage Commitment protects the images on the SCU (sending device) from being deleted until the SCP (receiving device) sends a message of validation that all the images have been safely received.

DICOM Worklist Management is designed to automatically pass radiology information system patient demographic information on scheduled radiology procedures to the PACS/acquisition device to seamlessly create a worklist. This eliminates the need for hand entry of the patient's demographic information into the PACS/acquisition device.

DICOM Print Class provides the ability to print any DICOM acquisition modality image format to any DICOM Print capable printer. Readers who are not well versed in DICOM nomenclature are referred to Dr Steven Horii's excellent paper, "A Nontechnical Introduction to DICOM."²

System Storage and Archive of Images included the major areas in Table 3.

Image storage and archive management was tested by sending images into the system from the server and tracking the pathway the images followed throughout the system. Several key questions that were asked included the following: At what time and where in the process was a permanent copy of the image stored within the system? For systems using compression of images in long-term storage, was an uncompressed image held in the system until a diagnostic reading had been performed? If exams were stored on local workstations was a copy of the exam also stored elsewhere in the system? Was it possible and who had the capability to delete or add images or exams?

Database management evaluation included answering the following questions: What device was being used as the database manager (eg, specific database manager or RIS)? Did the database manager provide full integration of the system? In particular, were the images and reports "married" together in seamless fashion even though they were stored in separate

Table 3. System Storage and Archive of Images

1. Image Storage and Archive Management
2. Database Management

areas? Was the database arranged so that all exams and reports for any patient could be viewed in a single folder? Was it possible to develop multiple worklists based on multiple selection criteria? Did the system have the capability for the development of consultation and teaching files?

Network Performance evaluation included the performance criteria in Table 4.

Each of these three test areas are critical to assure the system will be able to perform and support the workflow requirements need to support both the radiology department as well as the entire hospital.

Transit time into the system was evaluated by sending images into the system from the benchmark server and determining how long it took for the images to arrive at different points along the path. Timed tests were performed to determine when the images became available on the quality control workstation or other initial entry point, intermediate storage, and the long-term archive.

Display times of different exam types (eg, CR, CT, and MR) stored at the workstation and the intermediate storage device were evaluated. Time to de-archive different types of exams from the long-term archive was evaluated. These tests were performed with the proposed DIN-PAC systems in both loaded (maximum image throughput and network traffic) and unloaded (minimal image and network traffic) conditions.

Workstation Performance included the major areas in Table 5.

The functions tested for CR/single type images included image flip and rotate, image pan, fit image to monitor, masking capability (ability to mask off unexposed bright white areas around the image), default display protocols (eg, on which monitors the old and new images will be placed), were the default display protocols user-specific and stored centrally within the system so that users have universal access to their specific default displays at any workstation, variable zoom capability by interpolation and replication, gray-scale inversion, window and leveling by user specific presets and manually, one-to-one pixel match with original data, focal interrogation tool with zoom, gray-scale inversion and window and leveling, ability to view images in stack mode.

The functions tested for multiple images exams included default display protocols (DDP) for each modality type as well as the capability to have multiple DDP's for different exams types within a modality (eg, MR brain versus MR knee), viewing exams in stack or tile mode, synchronized scrolling of displayed images within an exam (eg, different image sequences of an MRI) and between new and historical exams, window and leveling presets for each modality as well as for multiple exam types within that modality (eg, CT head, chest, abdomen), manual window and leveling for a single image or the entire exam. The Utility functions included many of the typical utilities found on much of the software on the market today.

Consultation capability evaluated if it was possible to interactively operate separate pointing devices and use annotation so that providers on separate workstations could communicate back and forth about the same displayed exam.

Table 4. Network Performance

1. Image and Examination Transit Time into the System
2. Image and Examination Display Times
3. Image and Examination Retrieval Times from the Long-term Archive

Table 5. Workstation Performance

1. Display and manipulation of CR and other single image exams
2. Display and manipulation of multiple image exams such as CT and MR
3. Annotations and measurements
4. Utility functions
5. Consultation capability
6. Access privileges
7. Hardcopy generation from the workstation
8. Emulation window for the Composite Health Care System (CHCS)
9. Functionality of government-furnished equipment (GFE) personal computer software

Access Privilege assured that a user specific log on and password was required and that different users could be given different levels of access to the system. Examples to the levels of access include: (1) only a radiologist would have the privilege to mark an exam as being dictated; (2) only the System Administrator would have access to system wide functions.

Hardcopy generation of single images, entire examinations or selected images from an examination at a workstation to a network printer was evaluated. Also could this functionality be limited to only users with the appropriate log-on privilege.

The emulation window on the workstation of the Composite Health Care System (CHCS), which is the DoD wide Hospital Information System (HIS), is required to assure that the users may have access to the HIS without needing a second terminal.

The government-furnished PC was designed so that images and reports could be viewed on a desktop PC with communications access to the system. The PCs have less functionality and image interrogation capability than a workstation. The network connectivity and the software functionality as outlined in the RFP was evaluated by loading the vendors' software onto a PC that was part of the benchmark testing equipment.

Radiology Information System (RIS) Functionality testing evaluated the required elements of the RIS as outlined in Table 6.

The functionality in Table 6 includes the typical functions required of any RIS. Close integration of the RIS with the DIN-PACS database is required as noted in the Database management portion of the System Storage and Archive of

Table 6. Radiology Information System Functionality

1. Patient Registration
2. Order Entry
3. Order Scheduling
4. Patient Arrival
5. Radiology Encounter
6. Results Entry
7. Results Distribution
8. Workload Management Reporting
9. File and Table Building
10. Prefetch and Autorouting Support
11. Patient Tracking
12. Unscheduled Patient Functions
13. Report Generation
14. Crisis Management

Images section. The linkage of reports and images together within the system is a key concept of the DIN-PACS contract. The need for this close integration is also made evident by the fact that information for prefetching and autorouting will be coming from the RIS.

CHCS/DIN-PACS RIS Interface testing evaluated the system's ability to send and receive information to and from the DoD-wide hospital information system, Composite Health Care System (CHCS), using a Health Level 7 (HL7)³ standards based interface.

Teleradiology Performance categories evaluated are listed in Table 7.

Teleradiology DICOM conformance requirements included the same functions as required for the entire DIN-PACS. This portion of the testing was adversely affected by the team's inability to test DICOM Results Management, which is crucial for remote exam interpretation and reporting. The DIN-PACS teleradiology spoke was evaluated to ensure that the teleradiology spoke systems had the same local functionality and performance, with the exception of fetch times from the long-term archive, as outlined for all systems in the RFP.

The prefetch and fetch operations from remote archives were tested to assure that these functions provided the same level and ease of function to providers working at a remote sites.

The take home teleradiology functionality included the same functionality that was required of the GFE PC.

Quality Control Functionality testing evaluated the items outlined in Table 8.

The display monitor quality control testing used digital physics images that were sent to the system from the benchmark server. Multiple criteria were evaluated, including luminance, grayscale display function based on the ACR-NEMA Grayscale Standard,⁴ spatial resolution, geometric and linear distortion, contrast detectability, image smearing and edge detection.

The display monitor calibration testing determined if there was vendor provided software within the workstation capable of self-calibration of that workstation's monitors to a uniform luminance readings for specified pixel values.

The physics test images were used to evaluate the system's ability to accurately manage and display images. Specific areas tested included evaluating the systems capability of displaying a full image of known matrix size (2K × 2K) without minification and assessing the system's ability to provide the expected luminance for known pixel values.

Database image integrity was evaluated to determine if the DIN-PACS accurately maintained all images, exams, and reports. The system was evaluated to ensure that the correct number of images in an exam was mapped to the correct report and that the correct number of exams was mapped to the correct patient folder.

The diagnostic image evaluation entailed viewing images sent to the system from the server with known abnormalities to assure these findings were faithfully reproduced with the proposed system.

Table 7. Teleradiology Performance

1. DICOM Conformance
2. Teleradiology Spoke Functionality
3. Prefetch and Fetch Operations from Remote Archives
4. Take Home Teleradiology Functionality

Table 8. Quality Control

1. Workstation Display Monitor Quality
2. Monitor Calibration Software
3. Physics Test Images
4. Database Image Integrity
5. Diagnostic Image Evaluation

Crisis Management categories evaluated are outlined in Table 9.

System recovery testing was performed by disconnecting and/or powering down key components of the system such as the QC workstation or other entry point into the system, intermediate storage device, long-term archive, and RIS. This was performed while the remaining components were active with new exams and reports being sent to the system. Questions that were answered included the following: Was there a graceful recovery requiring minimal operator input? Was any information old or new lost during downtime?

Removable media functionality was tested to determine if it was possible to manually transport images between system components when a component or a portion of the network was not functional.

RESULTS

All of the vendors that underwent the benchmark testing did well. There was no one vendor that was markedly superior or inferior. There was a typical bell shaped curve of abilities. Each vendor had their own strong points and weaknesses. All the vendors did well on DICOM conformance, Radiology Information System Performance, CHCS/HL7 Interface, Teleradiology Performance, Quality Control Functionality and Crisis Management. In the System Storage and Archive of Images section several vendors did have problems with integration between devices. The RIS of all but one vendor was not well integrated with the PACS. This limited integration included the reports not being closely matched with the images and in some cases the reports were not viewable on the PACS. Also the prefetch information from the RIS was either not passed to the PACS or not used by the PACS so that no prefetching was performed. Network performance varied from good to poor. The workstations all worked well and the interfaces for the most part were easily used. There were considerable differences in the capabilities for viewing multi-image studies and single image.

Table 9. Crisis Management

1. System Recovery Capability
2. Removable Media Functionality

DISCUSSION

The final selection of the vendors was based on four criteria: written proposal, benchmark testing, product improvement plan, and pricing. The Written Proposal was the vendors response to the RFP. This went through several rounds of clarification before final acceptance. The product improvement plan allowed vendors to include functionality that they presently lacked, but was planned for commercial delivery within the first year after contract award. Pricing was based upon comparison of five generic sites with specific requirements, maintenance costs were also included.

The vendors who responded to the RFP include many of the leaders in the PACS arena as well as leading Integrators. The vendors ranged from a non-partnered vendor to vendors with multiple partners. Each respondent was considered to be qualified and able to produce a PACS. There were no organizations involved that were in any way considered to be suspect. The major advantages to the benchmark testing was to allow the evaluation team to put there hands on the equipment and see first hand how it functioned. The radiologists were able to get a very good understanding and feel for the workstations. How the entire system functioned as a whole was an invaluable peace of information.

The uninitiated PACS evaluator can easily be fooled by a "slick" workstation with an easy to operate interface, but if the remainder of the system does not operate smoothly and efficiently, that great workstation will be of little value.

A standardized benchmark testing system for the PACS industry with regularly timed evaluations and published reports would be of great value to all agencies planning to purchase a PACS. Presently, there is little independent performance data available, which leaves the prospective PACS buyers to their own devices to make an informed decision. A good source of information for the needed functionality is the DIN-PACS contract, SP020097-R-8002 DINPACS.⁵ Another good source of information would be a knowledgeable PACS consulting agency.

CONCLUSION

PAC systems have evolved into effective devices which are changing the practice of radiology and medicine as a whole. The complexity of these systems encompasses many areas of expertise. The cost of these systems is not trivial. All organizations considering the purchase of a PACS must arm themselves with as much information as possible before making a final decision.

REFERENCES

1. American College of Radiology, National Electronics Manufacturers Association: Digital Imaging and Communications in Medicine Version 3.0

2. Horii SC: A Nontechnical Introduction to DICOM. Radio-Graphics 17:1297-1309, 1997

3. Health Industry Level 7 Interface Standard Version 2.0

4. American College of Radiology, National Electronics

Manufacturers Association: Standard on Grayscale Display Function, Working Draft 0.8, October 1996

5. Defense Supply Center, Philadelphia: Digital Imaging Network-Picture Archive and Communications System Contract, Contract Number SP 020097-R-8002 DINPACS. Internet address: www.dpssc.dla.mil/medical/medequip/what.htm