The Philosophy of Benchmark Testing a Standards-Based Picture Archiving and Communications System

Nancy E. Richardson, Jerry A. Thomas, David K. Lyche, John Romlein, Gary S. Norton,
and Quentin E. Dolecek

The Department of Defense issued its requirements for a Digital Imaging Network–Picture Archiving and Communications System (DIN-PACS) in a Request for Proposals (RFP) to industry in January 1997, with subsequent contracts being awarded in November 1997 to the Agfa Division of Bayer and IBM Global Government Industry. The Government's technical evaluation process consisted of evaluating a written technical proposal as well as conducting a benchmark test of each proposed system at the vendor's test facility. The purpose of benchmark testing was to evaluate the performance of the fully integrated system in a simulated operational environment. The benchmark test procedures and test equipment were developed through a joint effort between the Government, academic institutions, and private consultants. Herein the authors discuss the resources required and the methods used to benchmark test a standards-based PACS.

Copyright © 1999 by W.B. Saunders Company

KEY WORDS: DIN-PACS, DICOM, benchmark testing, teleradiology, digital imaging, image management.

THE GOVERNMENT released its requirements to industry for an open architecture, standards-based Digital Imaging Network–Picture Archiving and Communications System (DIN-PACS) in a Request for Proposals (RFP)¹ in January 1997, with subsequent contracts being awarded in November 1997 to the Agfa Division of Bayer and IBM Global Government Industry. The goal of the DIN-PACS project is to support business process changes throughout the Military Health System (MHS) and within the practice of military radiology. The vision of this effort is to create a "virtual" radiology department by eliminating multiple "place" constraints that arise both within and between diagnostic medical centers. Within facilities, the DIN-PACS is intended to eliminate the necessity of creating film and allow access to images and corresponding radiographic results by multiple users at any place and any time. Between facilities, the system is intended to create opportunities to dynamically shift workload at any time and to any location where clinical expertise is available.²

SYSTEM REQUIREMENTS

Overview

The DIN-PACS is structured to be an open-system, standards-based network of digital devices designed for effective acquisition, transmission, display, and management of diagnostic imaging studies both within and between medical treatment facilities. As a minimum, the system is required to utilize Digital Imaging and Communications in Medicine (DICOM) 3.0 and Health Level 7 (HL7) standard formats in its communication with external systems and components. The Government identified compliance with these standards as well as the overall clinical and technical performance of the fully integrated system to be the key technical criteria in the source selection process.

DIN-PACS Model

Figure 1 presents the DIN-PACS model, which includes the functional components of the PACS as well as the communication pathways between the DIN-PACS and external components and systems. The shaded areas represent the components that comprise the DIN-PACS itself and are required to be provided by the DIN-PACS vendor. The unshaded areas represent components that interface to the DIN-PACS and are provided via separate procurement channels.

At the center of the model is the archive, which is the permanent, long-term, safe storage area for all diagnostic imaging studies acquired into the DIN-PACS. The image storage layer of the model represents a combination of short-term and intermediate storage implementations. This storage layer

---

From the Defense Supply Center Philadelphia, Ft Detrick, MD; Uniformed Services University of Health Sciences, Bethesda, MD; Informatech, Inc, Frederick, MD; and The Johns Hopkins University Applied Physics Laboratory, Laurel, MD.

Address reprint requests to Nancy E. Richardson, MSE, Defense Supply Center Philadelphia Liaison Office, US Army Medical Materiel Agency, 1423 Sultan Dr, Suite 100, Ft Detrick, MD 21702-5001.

*Copyright © 1999 by W.B. Saunders Company
0897-1889/99/1202-0006\$10.00/0*

Fig 1. DIN-PACS model.

differs functionally from the archive in terms of time depth of studies resident in the storage layer, the retrieval speed of studies from storage, and the fact that the storage media is volatile as opposed to permanent. The heart of the central system is the database and control subsystem responsible for storage of system and examination data as well as overall integration of system function.

The standard hospital information system (HIS) in the Department of Defense is the Composite Health Care System (CHCS). The DIN-PACS will communicate in a bidirectional manner with CHCS via an HL7 compliant interface. The DIN-PACS RFP requires that a radiology information system (RIS) be integrated with the DIN-PACS. This integrated RIS provides additional functionality that CHCS does not currently support as well as a level of "failover" operations to allow clinical operations to continue in the event of failure of connection to CHCS. ("Failover" operations are defined as those methods incorporated into the system which allow for uninterrupted clinical processing despite any single point of failure within the system.)

The output and display component of the DIN-PACS includes diagnostic and review workstations

as well as quality control workstations. The diagnostic and review workstations are identical in functionality and differ only in the resolution of their respective monitors. The quality control (QC) workstation is a specialized form of a review workstation used as an interface device between the DIN-PACS and acquisition modalities that currently do not provide all the DICOM services required by the Government (eg, modality worklist management, storage commitment). The QC workstation also provides QC-type manipulation functionality (eg, flip, rotate, mirror, invert gray scale, window, and level). Laser imagers are provided for instances in which printed films are required.

The model includes representations for the DICOM and HL7 standards, which are used for communication between the DIN-PACS and all external components and systems. The Government was explicit in its requirements for DICOM conformance required across these interfaces, addressing the specific DICOM service classes to be supported at each input and output device and system interface (eg, acquisition modalities, teleradiology, personal computers [PCs], laser printers). The DIN-PACS specification required that external communications be accomplished using DICOM

and HL-7. However, the internal communications protocol within the PACS was left to the vendor's discretion so as to allow the vendors to use proprietary protocols where necessary to meet the diagnostic imaging and reporting throughput requirements of the RFP.

BENCHMARK TESTING

The goal of the Government's technical evaluation process was to determine the level of compliance of each proposed DIN-PACS with the published requirements of the RFP. To best achieve this goal, the Government's source selection process consisted of evaluating a written technical proposal as well as conducting a benchmark test of the proposed DIN-PACS at each vendor's integrated test facility. A written proposal provides a vehicle for vendors to document the particular technical specifications of each component of their proposed system as well as to explain, in detail, the proposed concept of clinical operations of the fully integrated system. In contrast, the benchmark test allows the user to evaluate the actual performance of the proposed, fully integrated system in a simulated operational environment. Such testing allows the user to collect both qualitative and quantitative information that cannot be adequately evaluated based on information submitted in a written proposal. The results of benchmark testing are beneficial to the user even after contract award by providing a test platform and objective performance data against which installed DIN-PACS networks can be evaluated during acceptance inspection testing.

Benchmark Evaluation Areas

The DIN-PACS statement of work was subdivided into nine functional areas: DICOM Conformance, System Storage and Archive, Workstation Performance, Network Performance, RIS Functionality, CHCS Interface, Teleradiology, Quality Control, and Reliability. For classification purposes, each specification of the statement of work was assigned to one of these functional areas. In the overall technical evaluation schema, these nine areas were used as subfactors under the overall benchmark test evaluation factor. Upon completion of benchmark testing, each subfactor was assigned a rating and the subfactor ratings were combined to obtain an overall rating for the benchmark test.

Benchmark Test Hierarchy

The Government took a systematic bottom-up approach to structuring the benchmark test by defining a testing hierarchy. This bottom-up approach looks at the DIN-PACS from the perspective of a number of individual components and subsystems that function unto themselves as well as in conjunction with each other to comprise an overall integrated system. Given this perspective, the benchmark test structure was divided into three phases, defined in order of least to most complex in the testing hierarchy: the Component and Subsystem Demonstration (CSD), the System Integration Test (SIT), and the Clinical Loading Test (CLT).

The purpose of the CSD phase was to test the isolated functional performance of each component and subsystem of the DIN-PACS. Figure 2 represents a generic input/output (I/O) model for a CSD test.

The evaluator tests each component/subsystem by providing a controlled input and measuring the resulting output. Specific examples of a component/subsystem are each of the shaded blocks in Fig 1 (ie, diagnostic/review workstation, DICOM IN/OUT interface, RIS).

The SIT phase expands on the CSD phase by testing clinically relevant groups of components. The purpose of the SIT is to evaluate the throughput and performance of physically and logically integrated components. Figure 3 represents a generic input/output model for an SIT test.

Each group of components is tested by providing a controlled input to the first component and measuring the resulting output of one or more of the following components. The combination of components and controlled data sequences may be thought of as "threads" through the system. Specific examples of such threads would be specific clinical scenarios to be accomplished within the DIN-PACS (eg, examination ordering, previous examination de-archiving, new image acquisition,

Fig 2. Generic I/O model: Component and subsystem demonstration test.

Fig 3. Generic I/O model: System integration test.

soft-copy display and diagnosis, radiographic reporting).

The CLT phase expands on the SIT phase a step further by evaluating the system at both the component and integrated system level in the presence of background system and network traffic. The purpose of the CLT is to evaluate the true performance of the system in the most clinically relevant environment, when the system is being used simultaneously by multiple users. Figure 4 represents a generic input/output model for a CLT test.

In this case, each component is subject to multiple near simultaneous inputs and outputs, shown by dashed lines. Concurrently a controlled input thread, shown by the solid line, is tracked through the system, and resulting outputs are measured. This controlled thread is performed under both load and no-load conditions. The outputs under both conditions are compared to determine the impact of system and network loading on resulting throughput and performance.

Specific examples of CSD, SIT, and CLT tests will be described in detail later in this report.

Benchmark Test System

The Government contracted with The Johns Hopkins University Applied Physics Laboratory (JHU/APL) to develop a test system to be used during the benchmark test.³ Assisting in this prepa-

ration were Mallinckrodt Institute of Radiology, MIS Labs, Inc, and the Maryland Consortium for Evaluation of Telemedicine. The test system simulated the expected inputs and outputs to a DIN-PACS (eg, acquisition devices, DICOM workstations) as well as network loading conditions typical of a hospital network. Design and construction of the benchmark test system consisted of three tasks: compilation of a suitable set of test images with corresponding patient demographic and report information; acquisition and integration of computer hardware and network interfaces; and acquisition and design of software to perform required benchmark scripts.

The first step in constructing the benchmark test system was compilation of a set of valid DICOM test images. Images from various modalities (eg, computed radiography [CR], computed tomography [CT], ultrasonography, magnetic resonance imaging [MRI], film digitized images) were used. The compiled image test set represented a variety of patients, modalities, diagnostic imaging equipment manufacturers, and types of examinations (ie, chest CR v extremity CR). To protect patient confidentiality, all patient name and identification data were modified on the original image sets. These imaging examinations were subdivided into those to be preloaded to the DIN-PACS under test to serve as "historical" examinations and those to

Fig 4. Generic I/O model: Clinical loading test.

Fig 5. Benchmark test system.

be stored into the DIN-PACS “live” during the benchmark test itself. A series of physics images consisting of standard test patterns also were constructed in a DICOM format for use in evaluating display monitor quality. All images were then validated for DICOM compliance using the Mallinckrodt Institute of Radiology (MIR) Electronic Radiology Laboratory (ERL) Central Test Node (CTN) DICOM demonstration software. The MIR CTN software was implemented for this validation so as to have a vendor-independent platform for validation of DICOM conformance. RIS data were generated corresponding to each imaging examination. The RIS data included patient demographics, radiology orders, appointment information, and, for those examinations to be used as preloaded historical examinations, radiology reports.

Preloading of historical examinations to the DIN-PACS archive before testing was accomplished using an IBM PC-compatible laptop computer running Microsoft Windows NT operating system. The laptop contained a PCMCIA SCSI card that connected the laptop to an Iomega 1 Gbyte Jaz drive, which stored the historical examinations for preloading, and a second PCMCIA Ethernet card that allowed the laptop to be connected to the vendor’s Transmission Control Protocol/Internet Protocol (TCP/IP) network. Associated radiographic examination orders and reports were

input to the vendor system by transmission from a supporting CHCS system. After transfer to the vendor’s system, the images and supporting data and reports were migrated, by the vendor, to the system archive layer so they would react as historical examinations during the test.

The benchmark test system hardware consisted of a SUN ULTRASPARC workstation running a UNIX platform and two Pentium PCs. Benchmark test software running on the UNIX platform included MIR Central Test Node (CTN) software (publicly available by ftp at http://www.erl.wustl.edu/ftp_server.html), Apache Web Server (publicly available by ftp at <http://www.apache.org/dist/>), and Perl version 5.004 (publicly available by ftp at <http://www.perl.com/CPAN/src>). A Hypertext Markup Language (HTML) system was used to create a graphic user interface (GUI) on one of the PCs to run the actual test software on the remote UNIX system using predefined scripts. CHCS messaging was accomplished via an Internet connection to a CHCS platform located at Brooks Air Force Base, San Antonio, TX.

The MIR CTN software was used to establish DICOM associations to validate the Service Object Pair (SOP) classes defined in the Government RFP. SOP classes tested included both Service Class User (SCU) and Service Class Provider (SCP) messaging for DICOM Verification, Storage, and Query/Retrieve. Additional DICOM compliance

BOOKLET 1 (CSD) DICOM CONFORMANCE		BOOKLET 3 (CSD) RIS FUNCTIONALITY	
Test No.	Test Title	Test No.	Test Title
1	DIN-PACS DICOM Communication (SCU)	29	CHCS/DIN-PACS Interface
2	QC Workstation DICOM Communication (SCU)	30	Minimum RIS Functionality
3	Teleradiology Workstation DICOM Communication (SCU)	31	Patient Registration
4	GFE PC Workstation DICOM Communication (SCU)	32	Order Entry
5	GFE PC Workstation DICOM Communication (SCP)	33	Order Scheduling
6	DIN-PACS Workstation DICOM Communication (SCP)	34	Patient Arrival
7	Teleradiology Workstation DICOM Communication (SCP)	35	Radiology Encounter
8	Database Archive and Display	36	Results Entry
9	Telecommunication Network	37	Results Distribution
10	Hardcopy Device Interface	38	Workload/Management Reporting
11	QC Workstation Modalities	39	File and Table Building
12	GFE PC Module Attributes	40	Exam Procedure Database
13	Image Storage and Archive	41	Prefetch Support
14	Exam Retrieval and Display Speed	42	Report Generation
		43	Fetch Capabilities
BOOKLET 2 (CSD) WORKSTATION PERFORMANCE		BOOKLET 4 (CSD) TELERADIOLOGY	
15	GFE PC Display	44	DIN-PACS Teleradiology Hub
16	DIN-PACS Display	45	Spoke Teleradiology
17	Display Monitor Image Quality	46	Prefetch from Remote Archives
18	Display Calibration		
19	CR Image Display and Manipulation		
20	CT and MR Image Display and Manipulation		
21	Default Display Protocols and Prefetch Function		
22	Annotations and Measurements		
23	Folder Systems		
24	Utility Functions		
25	Image Deletion and Command Reversal		
26	Hardcopy		
27	CHCS Window		
28	CHCS Reports		

Fig 6. Benchmark test procedures.

testing was performed between the Benchmark test site and MIR ERL via the Internet. These included testing of DICOM Worklist Management functionality.

Loading of the DIN-PACS network during the CLT phase was accomplished using two methods. Background generic network traffic was generated by connecting the two PCs to two points on the DIN-PACS local area network (LAN) and sending and receiving packets between the two points. The Government operator controlled the initiation and termination of loading as well as the rate of message traffic and the packet sizes. The second type of loading consisted of DICOM message traffic. This method stored DICOM images, of various modality types, at up to 1 gigabit per second to the DIN-PACS to create a load on the internal DIN-PACS components (eg, storage and archive devices, workstations, RIS/database sub-system).

Figure 5 shows the relationship between the benchmark test system and the DIN-PACS under test using the DIN-PACS model. The benchmark test system effectively acted as the DICOM input and output devices (eg, DICOM modalities, DICOM workstations, etc), and the Brooks Air Force Base platform acted as the CHCS for purposes of testing.

Benchmark Test Procedures

Given time constraints of the testing period and the resources required to accomplish the testing, the Government selected a subset of the RFP requirements to benchmark-test. Each RFP specification was analyzed for feasibility in testing and assigned to one or more of the appropriate benchmark test phases (ie, CSD, SIT, and CLT). Selected RFP specifications were grouped by phase, and explicit test procedures were written for each phase. Each test procedure was derived from its

BOOKLET 5 (SIT) DICOM CONFORMANCE SYSTEM STORAGE AND ARCHIVE NETWORK PERFORMANCE WORKSTATION PERFORMANCE QUALITY CONTROL		BOOKLET 7 (CLT) DICOM CONFORMANCE NETWORK PERFORMANCE WORKSTATION PERFORMANCE	
Test No.	Test Title	Test No.	Test Title
47	System Architecture	69	Image Transfer Performance
48	DIN-PACS Support for DICOM Workstation	70	PC Software
49	DIN-PACS DICOM Support for GFE PC	71	Consultation Capability
50	Database Management	BOOKLET 8 (CLT) RIS FUNCTIONALITY CHCS COMPATIBILITY TELERRADIOLOGY	
51	Workstation Storage and Archive Operation	72	Prefetch Requirements
52	System Crisis Management	73	Phase I Clinical Scenario using CHCS with System Loading
53	Communications Network Interface	74	Phase II Clinical Scenario with System Loading
54	Worklist Functions	75	MTF/MTF Teleradiology with System Loading
55	Access Privileges	BOOKLET 9 DIAGNOSTIC AND IMAGE PHYSICS TESTS	
56	Quality Control Test Images	76	Diagnostic Test 1
57	Database Image Integrity	77	Diagnostic Test 2
		78	Diagnostic Test 3
		79	Image Physics
BOOKLET 6 (SIT) RIS FUNCTIONALITY CHCS COMPATIBILITY TELERRADIOLOGY			
58	Phase II CHCS RIS		
59	Patient Tracking		
60	Operation Concepts		
61	Acquisition of Prescheduled Patient Record		
62	Unscheduled Patient Functions		
63	Radiologist Image Interpretation		
64	Phase I Clinical Scenario		
65	Phase II Clinical Scenario		
66	DIN-PACS/MTF Teleradiology Architecture		
67	MTF-to-MTF Teleradiology		
68	Take-home Teleradiology		

Fig 6. (Cont'd).

corresponding RFP specification(s). A total of 79 test procedures were compiled and organized into nine test booklets aligned with the benchmark test areas and benchmark test phases. A list of the final test procedures is provided in Fig 6.

For more information on the testing, see the report in this issue by Dr Philip Shelton et al.

CONCLUSION

Benchmark testing was found to be an effective tool for determining compliance of a standards-based PACS with contract requirements. Results of the benchmark test confirmed that a simulated clinical

environment can be used to evaluate the performance of a truly integrated PACS. The benchmark test proved that standards-based PACS are feasible and available in the commercial marketplace today.

REFERENCES

1. Defense Personnel Support Center, Philadelphia, PA, Solicitation No. SP0200-97-R-8002, January 1997
2. Richardson NE, Thomas JA, O'Riley PS, et al: Highlights of the digital imaging network—Picture archiving and communications system project. *J Digit Imaging* 10:44-46, 1997
3. The Johns Hopkins University Applied Physics Laboratory: DIN-PACS Benchmark Testing Using the Medical Information System Testbench (MIST) Final Report, November 1997