An IBIS White Paper
July 2011

IBISimg Release 2.1
Medical Imaging Performance Benchmark on Enterprise Hardware

Table of Contents

1.0 Executive Summary

Over a two-week period in June 2011, IBIS worked with an Enterprise Hardware provider to complete a performance benchmark for IBISimg, the IBIS medical image data management solution, on Oracle WebLogic Server 11g with Oracle Database 11g Release2 on Enterprise Hardware.

Using a representative sample set of medical image data from five modalities – MR, PET, CT, Ultrasound and Mammography – we measured the performance of five primary IBISimg SOA-based Web Services functions – upload/insert, extract/normalize, validate, download and view – across 25, 50 and 100 client connections on Enterprise Hardware configurations.

In summary, our findings were as follows:

- As Enterprise Hardware is scaled up from 1 to 4 instances and the number of concurrent user connections is increased from 25 to 100, IBISimg scales accordingly, and data throughput is, on average, over 2.5 times faster
- As user connections are increased, user response times improve consistently
- IBISimg has been verified to scale up to at least 800 concurrent user connections.
- Over 1TB of medical image data can be uploaded to the system in one hour.
- Time to first image on DICOM viewer is sub-second for MR, PET and CT

2.0 Background

Over the past several years, while imaging technologies have become extremely sophisticated and continue to become ever more so, the PACS systems used to manage and store image data have evolved much more slowly and continue to utilize technologies and data storage solutions that belong to another era.

As a result, health care facilities have had to struggle with some very serious challenges: data storage for vendor PACS systems is file-based, with considerable on-going maintenance and migration issues and costs; vendor systems tend to be proprietary, so even smaller facilities may contain several heterogeneous PACS systems; there is little or no interoperability between PACS systems. While some vendors offer a Vendor Neutral Archive (VNA) to address some of these issues, even the better VNA's are little more than revamped PACS systems with some limited additional capability and they are all severely bounded by their use of file-based storage.

The medical imaging community in clinical research has also faced considerable challenges. In their case, medical image data are typically scattered across multiple repositories at external partner Core Labs, providing little or no opportunity for cross-study analysis or, more generally, for data mining or re-purposing. Lack of consistency in image data quality has long been recognized as a major issue. The data management process is typically not streamlined – email, ftp, DVDs, hard disk transfers and similar technologies are still widely used – and is prone to error. And, finally, sponsors' access to data is far from optimal, and this has led to sometimes costly delays in data review and in the timing of Go/No Go decisions.

Recognizing the issues, the needs and the opportunities in this space, IBIS has developed IBISimg, a comprehensive, state-of-the-art, enterprise-class image data management platform, which offers a clearly differentiated alternative to traditional PACS systems. A core suite of data upload, normalization, validation, download and viewing capabilities are delivered via a secure, scalable Services Oriented Architecture (SOA) and these capabilities are rounded out with a comprehensive set of supporting, client-driven functionality. Image data are stored directly in the database and industry standards – HIPAA, HL7, CDISC ODM and DICOM – are integral to the system.

3.0 Purpose

The purpose of this benchmark is to measure data throughput on IBISimg deployed under Oracle WebLogic Server, using Oracle Database 11g deployed on Enterprise Hardware and to demonstrate the system's scalability across increased client connections and increased Enterprise Hardware instances.

We used a representative sample of medical image data from five modalities – Ultrasound, CT, MR, PET and Mammography – and recorded throughput and performance measurements across the five major processes in IBISimg – upload/insert, extract/normalize, validate, download and view. To measure scalability, we grew both the number of concurrent client connections and the Enterprise Hardware configuration.

4.0 Benchmark Design

This benchmark is made up of four major components: the IBISimg system, installed on Oracle WebLogic Server; the Oracle Database 11g, including Oracle 11g Multimedia DICOM, deployed on Enterprise Hardware; the test data, which are a representative sample of medical image data from the field; and the client test suite, which uses IBISimg Web Services to process the test data sets.

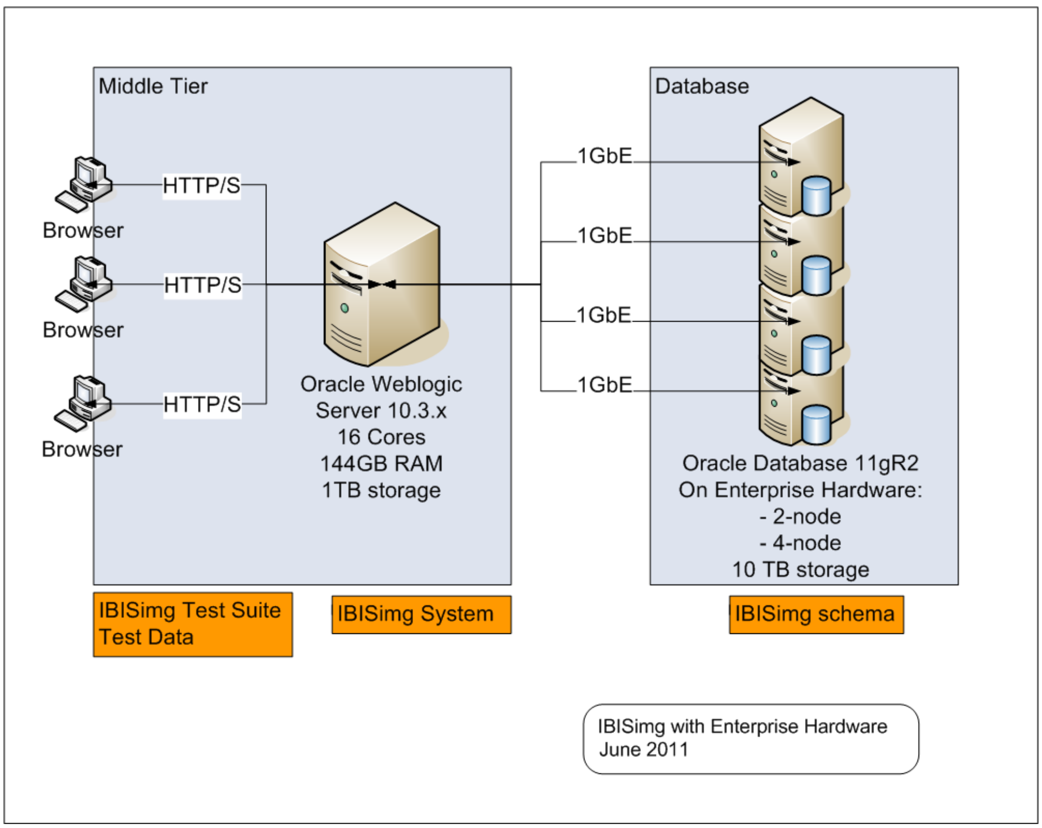

Figure 1 System Environment

4.1 IBISimg on Oracle WebLogic Server

IBISimg is Java EE compliant and is deployed under Oracle WebLogic Server.

Item	Description
Software	Oracle Enterprise Linux Oracle WebLogic Server IBISimg v2.1.1
Server	16 core, 144GB RAM, 1TB storage

Table 1 Application Environment

Major functional components of IBISimg are implemented using a Services Oriented Architecture (SOA) and the system includes a suite of Web Services through which system functionality may be invoked.

The Web Services layer is stateless in order to optimize performance and scalability.

The test suite for this benchmark uses the SOA interface.

4.2 Oracle Database 11g on Enterprise Hardware

IBISimg uses Oracle Database 11g. All medical images are stored directly in the database using Oracle Multimedia DICOM functionality.

The database software is deployed on Enterprise Hardware, and is configured with 1, 2 and 4 instances.

Item	Description
Database Software	Oracle Database Enterprise Edition 11.2.0.2
Database Schema	IBISIMS
Total Space	10TB
Enterprise Hardware	4 x Database Servers, with ➢ 2 Quad-Core Intel® Xeon® E5540 Processors (2.53 GHz) ➢ 72 GB memory ➢ Disk Controller HBA with 512MB Battery Backed Write Cache ➢ 4 x 146 GB SAS 10,000 RPM disks ➢ Dual-Port QDR InfiniBand Host Channel Adapter ➢ 4 Embedded Gigabit Ethernet Ports

Table 2 Database Environment

4.3 Benchmark Data

The output of a diagnostic medical imaging device is usually referred to as a study or a scan. A study is a set of related images, often referred to as 'slices', each stored in a DICOM file.

Studies can vary greatly from each other – a study can be large or small and can be made up of many slices or a few slices.

We have used a representative sample of medical image data for this benchmark in order to present a variety of real-world challenges to the system. Two of the datasets – Ultrasound and Mammography – are relatively large in overall size and each contains a relatively small number of large files. The other three datasets – CT, MR and PET – are smaller in size and each contains a relatively large number of small files.

The data used for this benchmark is taken from publicly available medical imaging repositories.

Modality	# Studies	# DICOM Files	Average File Size (KB)	Total Size (GB)
Mammography	5,000	5,000	33,019	157.45
Ultrasound	300	6,000	14,410	82.45
CT	500	230,000	71	15.68
MR	2,006	200,590	56	10.79
PET	1,000	227,000	11	2.40

Table 3 Test Data

4.4 Benchmark Test Suite

IBISimg provides a rich set of functionality through a set of SOA interfaces. For the purposes of this benchmark, five major capabilities of the system are measured.

IBISimg Function	Description
Upload/Insert	Secure, authenticated data upload. The client connects to the IBISimg system on Oracle WebLogic Server to initiate the upload. Then the SOA interface reads the DICOM files on the client, transfers these files from the client to the application tier, which stores them into the DICOM objects in the database.
Extract/Normalize	The system reads the set of DICOM files for each study from the database, processes the DICOM header metadata and stores back a normalized representation of the DICOM header metadata per study in the database
Validate	The system validates the DICOM header metadata against a data quality specification ("conformance document") for each study

IBISimg Function	Description
Download	Secure, authenticated data download. The client connects to the IBISimg system on Oracle WebLogic Server then the SOA interface reads the DICOM files out of the database and writes these files on the client, along with a CDISC ODM manifest file that describes all studies/files contained in the download
View	A virtual viewing operation consisting of a combination of DICOM CFind and CMove operations initiated at the client and processed through the middle tier out of the database

Table 4 Test Suite

Each test is executed as a set of “headless browser” client connections through which the IBISimg Web Services are invoked. All calls to the system are first authenticated and then processed by the system.

All benchmark measurements are stored in the database.

The number of concurrent client connections is configurable for each test and this is used to measure scalability.

5.0 Benchmark Measurements

Four configurations of the environment were used, as follows:

- 25 IBISimg client connections / 1 instance on Enterprise Hardware
- 50 IBISimg client connections / 2 instances on Enterprise Hardware
- 50 IBISimg client connections / 4 instances on Enterprise Hardware
- 100 IBISimg client connections / 4 instances on Enterprise Hardware

Each of the datasets was processed through each of these configurations and the results captured in the database. Following are the results:

5.1 Throughput using 1 instance on Enterprise Hardware

In order to establish a baseline, each of the five data sets were first processed on a single instance on Enterprise Hardware using 25 concurrent client connections, which yielded the following results:

Modality	Total Volume (GB)	# DICOM Files	# Clients	Upload/ Insert	Extract/ Normalize	Validate	Down-load
Mammography	157.45	5,000	25	00:36:38	00:24:59	00:00:55	00:24:11
Ultrasound	82.45	6,000	25	00:13:05	00:06:24	00:00:03	00:13:03
CT	15.68	230,000	25	00:14:34	00:05:46	00:14:09	00:13:22

MR	10.79	200,590	25	00:19:09	00:03:35	00:04:09	00:15:34
PET	2.40	227,000	25	00:17:44	00:03:17	00:08:28	00:22:05

Table 5 Throughput on 1 instance

5.2 Throughput using 2 instances on Enterprise Hardware

The same five sets of data were then processed on a 2-instance Enterprise Hardware with 50 concurrent client connections, which yielded the following results:

Modality	Total Volume (GB)	# DICOM Files	# Clients	Upload/ Insert	Extract/ Normalize	Validate	Down-load
Mammography	157.45	5,000	50	00:19:19	00:12:52	00:00:45	00:12:34
Ultrasound	82.45	6,000	50	00:10:30	00:03:15	00:00:05	00:06:53
CT	15.68	230,000	50	00:12:42	00:05:32	00:08:24	00:10:11
MR	10.79	200,590	50	00:17:11	00:02:41	00:04:20	00:08:03
PET	2.40	227,000	50	00:10:08	00:02:38	00:07:51	00:10:30

Table 6 Throughput with 2 instances on Enterprise Hardware / 50 Clients on IBISimg

5.3 Throughput using 4 instances on Enterprise Hardware – Run #1

The same five sets of data were then processed on a 4-instance Enterprise Hardware with 50 concurrent client connections, which yielded the following results:

Modality	Total Volume (GB)	# DICOM Files	# Clients	Upload/ Insert	Extract/ Normalize	Validate	Down-load
Mammography	157.45	5,000	50	00:10:57	00:12:03	00:00:47	00:06:51
Ultrasound	82.45	6,000	50	00:06:09	00:01:48	00:00:06	00:03:49
CT	15.68	230,000	50	00:10:43	00:03:13	00:05:02	00:07:10
MR	10.79	200,590	50	00:16:17	00:01:51	00:03:54	00:05:44
PET	2.40	227,000	50	00:08:00	00:02:40	00:05:05	00:06:25

Table 7 Throughput with 4 instances on Enterprise Hardware / 50 Clients on IBISimg

5.4 Throughput using 4 instances on Enterprise Hardware – Run #2

The same five sets of data were then processed on a 4- instance Enterprise Hardware with 100 concurrent client connections, which yielded the following results:

Modality	Total Volume	# DICOM	# Clients	Upload/ Insert	Extract/ Normalize	Validate	Down-load

	(GB)	Files					
Mammography	157.45	5,000	100	00:10:07	00:11:05	00:00:41	00:07:38
Ultrasound	82.45	6,000	100	00:05:40	00:01:49	00:00:04	00:03:55
CT	15.68	230,000	100	00:07:13	00:02:38	00:03:50	00:06:23
MR	10.79	200,590	100	00:12:04	00:01:38	00:03:27	00:06:00
PET	2.40	227,000	100	00:07:21	00:02:11	00:03:38	00:05:28

Table 8 Throughput with 4 instances on Enterprise Hardware / 100 Clients on IBISimg

5.5 DICOM Viewer

The image viewing operation is different from the other operations in some key aspects. Firstly, in order to match a real world scenario, we had to simulate the DICOM clients. (The alternative would have been to deploy up to 100 separate DICOM client workstations, which was not feasible). Secondly, in this run, image data must be transferred over the DICOM protocol, using DICOM CFIND and CMOVE commands, as opposed to using the HTTP protocol. The DICOM protocol is “chatty” and is less efficient than the HTTP protocol, however, as of this time, it is the only option available for communication with a DICOM viewing device. Finally, and most importantly, whereas for all other runs in this benchmark, client connections are used to perform pieces of the same job, in this case, as we add DICOM client connections, each DICOM client initiates its own study retrieval process, in other words, as the number of client connections grows, the volume of data transferred grows equivalently. Therefore, scalability will be demonstrated if we can come close to sustaining throughput levels as we raise the number of client connections.

The key differentiating factors for this run are the size of the DICOM files and the number of files per study. For example, in the case of Mammography, the study is a single file with, on average, a size of about 32MB, whereas in the case of PET, the study is made up of, on average, 227 files, with a total size of only 2.5MB.

We have measured the “time to first image” by which we mean the time it takes to complete the transfer of the first image file to the DICOM viewer from the time of the CMOVE request.

Modality	# DICOM Files	Average File Size (KB)	1-instance/ 25 clients	2-instance/ 50 clients	4-instance/ 50 clients	4-instance/ 100 clients
			Time in seconds			
Mammography	1	33,020	2.2705750	2.2121390	1.8298640	1.611606
Ultrasound	20	14,409	1.076121	0.337373	0.299864	0.263294
CT	460	71	0.602292	0.597066	0.565201	0.534238
MR	100	56	0.5018590	0.494019	0.4704650	0.4411280
PET	227	11	0.3480620	0.2924960	0.2273020	0.1852850

Table 9 DICOM Viewer

5.6 Throughput using 4 instances on Enterprise Hardware – Run #3

Finally, as a separate exercise, we wanted to demonstrate that since the Web Services Layer is stateless, it scales to any number of clients. To demonstrate its scalability, we processed a dataset of 1 TB of Ultrasound data using 800 connections. This required 2 Oracle WebLogic Server instances, and was also run on the 4-instance Enterprise Hardware, and yielded the following results:

Modality	Total Volume	# DICOM Files	# Clients	Upload/ Insert	Extract/ Normalize	Validate	Download
Ultrasound	1 TB	72,770	800	00:54:42	00:13:19	00:00:14	00:39:26

Table 10 Throughput using 4 instances on Enterprise Hardware / 800 Clients in IBISimg

6.0 Benchmark Analysis

The datasets we have used were selected with a view to providing as broad a range of variability in the data as possible, which, we believe, when coupled with the system configuration changes as we scale up, provide a representative measurement of the system's performance and scalability.

The following sections detail the analysis of the results and the basis for our conclusions.

6.1 Process #1. IBISimg Upload/Insert: Throughput and Scalability Analysis on Enterprise Hardware

This is a two-part operation involving the physical file transfer from the client and the insert of the file into the DICOM repository.

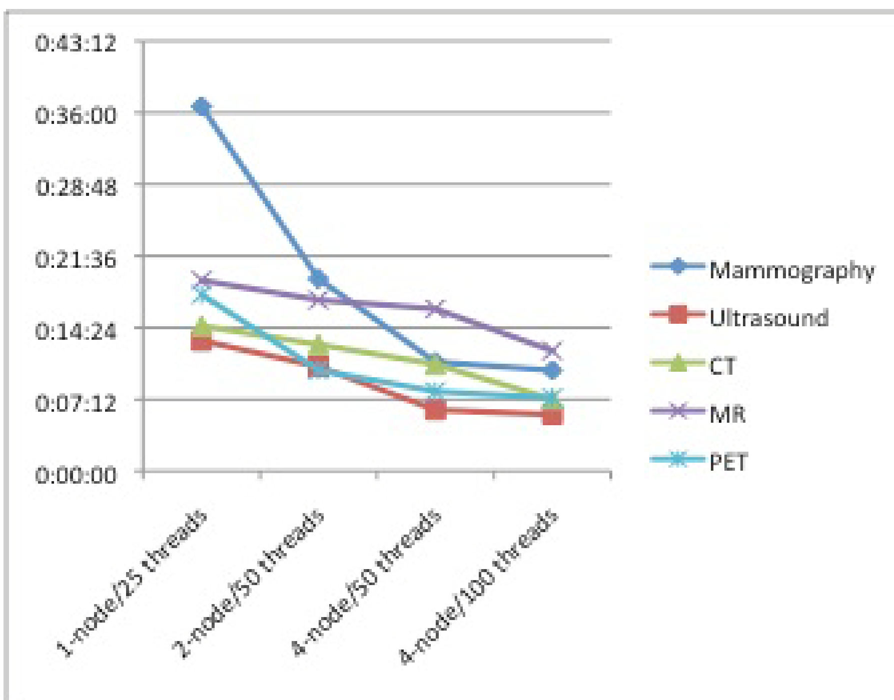

Figure 2 IBISimg Upload/Insert

For each of the five datasets, data throughput increased consistently for the upload/insert process as the system configuration was scaled up. Overall, dataset throughput was between 1.6 and 3.6 times faster, for MR and Mammography respectively, on the 4-instance/100 connections configuration than on the 1-instance/25 connections configuration.

Observation #1: The IBISimg data upload/insert process performs consistently across diverse datasets, improves throughput and scales with increased client connections/Enterprise Hardware instances.

6.2 Process #2. IBISimg Extract/Normalize: Throughput and Scalability Analysis on Enterprise Hardware

This is a three-part operation involving the extraction of the DICOM files from the database to the IBISimg system on Oracle WebLogic Server, the processing of these DICOM files by study to define a normalized representation of their DICOM header data, and the storage of the resulting normalized data into the database.

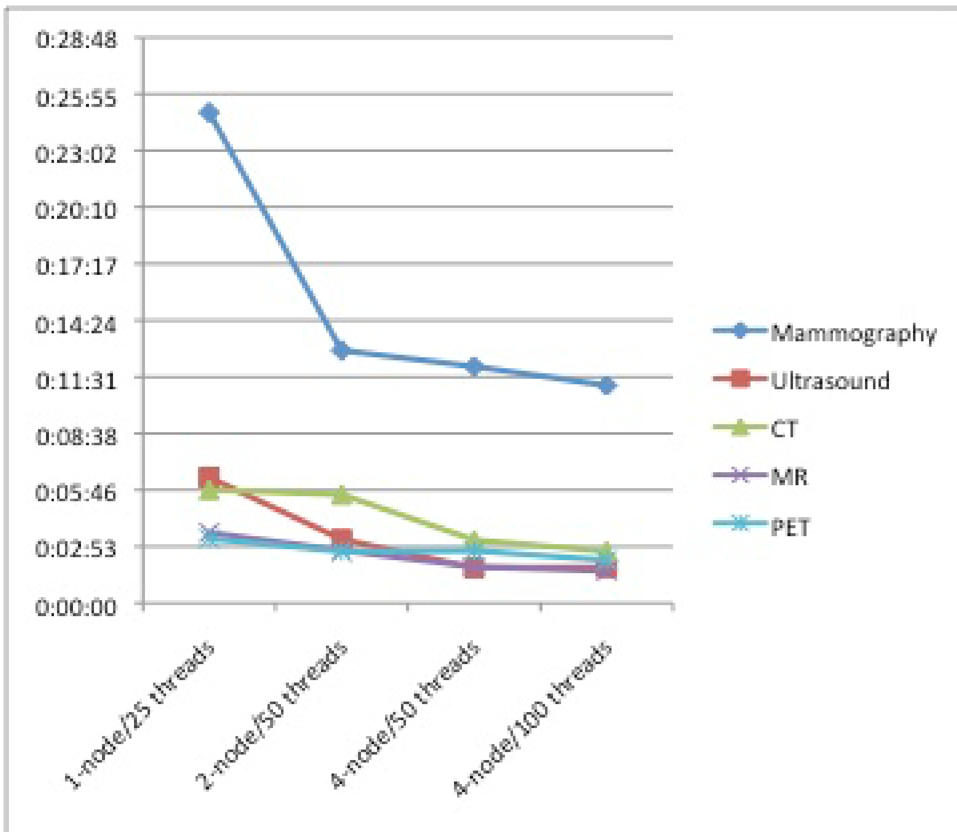

Figure 3 IBISimg Extract/Normalize

In the case of four of the five datasets, data throughput increased at all stages as the system configuration was scaled up – there was a slight aberration for PET between 2-instance/50 connections and 4-instance/50 connections, which may be an artifact of the network. Across the four configurations, however, all dataset throughput times were improved – by between 1.5 times for PET and 3.5 times for Mammography.

Observation #2: The IBISimg data extract/normalize process performs consistently across diverse datasets, improves throughput and scales with increased client connections/Enterprise Hardware instances.

6.3 Process #3. IBISimg Data Validation: Throughput and Scalability Analysis on Enterprise Hardware

This operation is performed primarily on the database. Each study is validated against a data quality specification or 'conformance document'. Neither the client nor the Oracle WebLogic Server tier is involved more than peripherally.

The key determinants for the amount of processing required, in this case, is the number of studies and the relative complexity of their header data. Based on the known characteristics of the different data sets, it is to be expected that CT, MR and PET will require considerably more processing than either of the other two datasets.

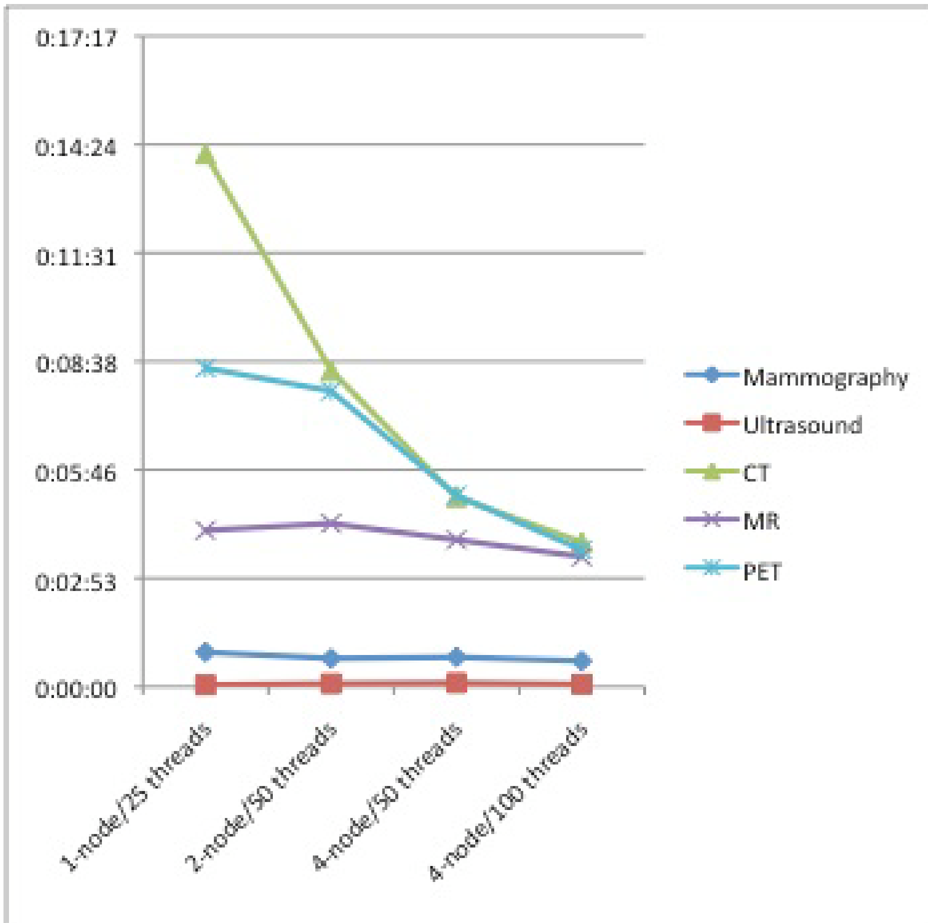

Figure 4 IBISimg Data Validation

In the case of the three datasets with larger numbers of files per study, data throughput increased at all stages as the system configuration was scaled up. The improvement in the Mammography dataset was minimal, as expected, and the numbers for Ultrasound were statistically insignificant.

Across the four configurations, dataset throughput times were 3.7 times faster for CT, 2.3 times for PET and 1.2 times for MR.

Observation #3: The data validation process performs consistently across diverse datasets, and improves materially for more complex validations with increased Enterprise Hardware instances and increased connections

6.4 Process #4. IBISimg Download: Throughput and Scalability Analysis on Enterprise Hardware

This operation is a combination of the retrieval of the DICOM data from the database and the physical file transfer function from the Oracle WebLogic Server to the client. The Download process also writes a CDISC ODM manifest file to the client that describes all studies/files contained in the download.

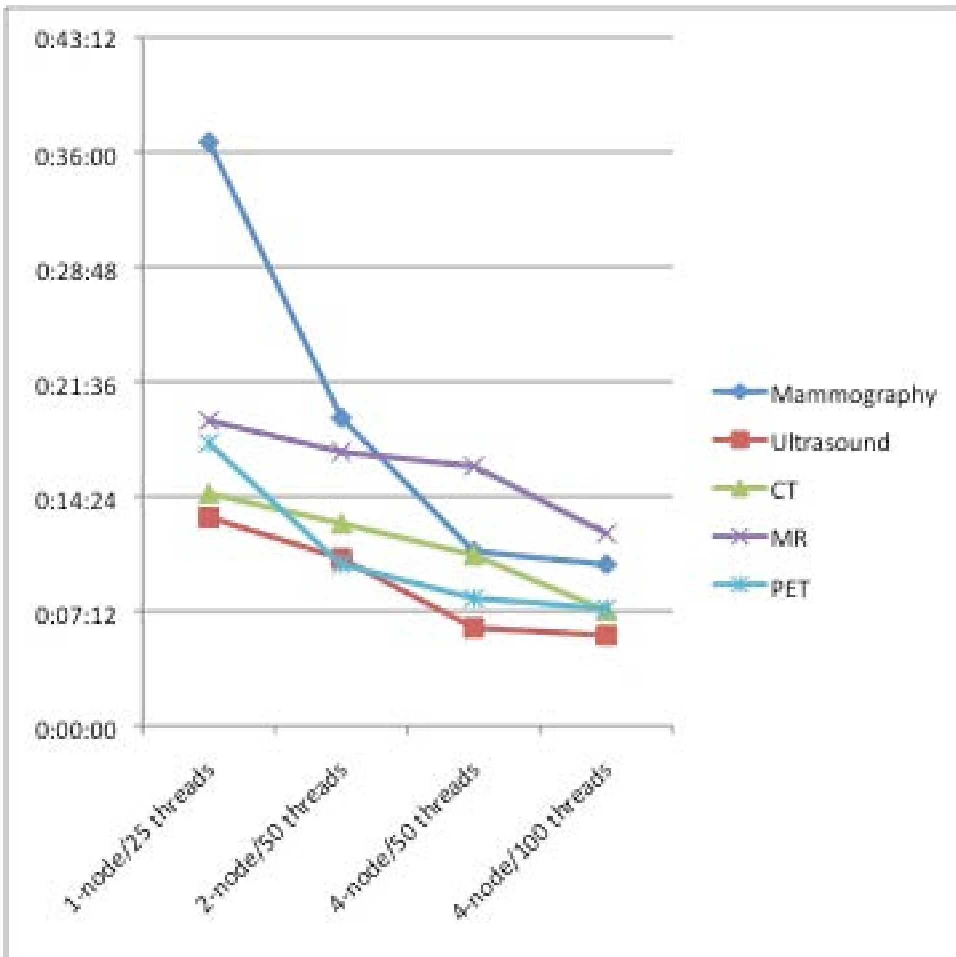

Figure 5 IBISimg Data Download

For each of the five datasets, data throughput increased consistently for the download process as the system configuration was scaled up. Overall, dataset throughput was between 2.1 and 4.0 times faster, for CT and Mammography respectively, on the 4-instance Enterprise Hardware/100 connections configuration than on the 1-instance Enterprise Hardware/25 connections configuration.

Observation #4: The IBISimg data download process performs consistently across diverse datasets, improves throughput and scales with increased client connections/Enterprise Hardware instances.

6.5 IBISimg Combined Processes 1 - 4 Totals: Throughput and Scalability Analysis on Enterprise Hardware

The total processing times for the combined four processes of upload/insert, extract/normalize, validate and download across each of the four configurations is displayed below.

Modality	Total Volume (GB)	# DICOM Files	1-instance/ 25 clients	2-instance/ 50 clients	4-instance/ 50 clients	4-instance/ 100 clients
Mammography	157.45	5,000	1:26:43	0:45:30	0:30:38	0:30:03
Ultrasound	82.45	6,000	0:32:35	0:20:43	0:11:52	0:11:28
CT	15.68	230,000	0:47:51	0:36:49	0:26:08	0:20:04
MR	10.79	200,590	0:42:27	0:32:15	0:27:46	0:23:09
PET	2.40	227,000	0:51:34	0:31:07	0:22:10	0:18:38

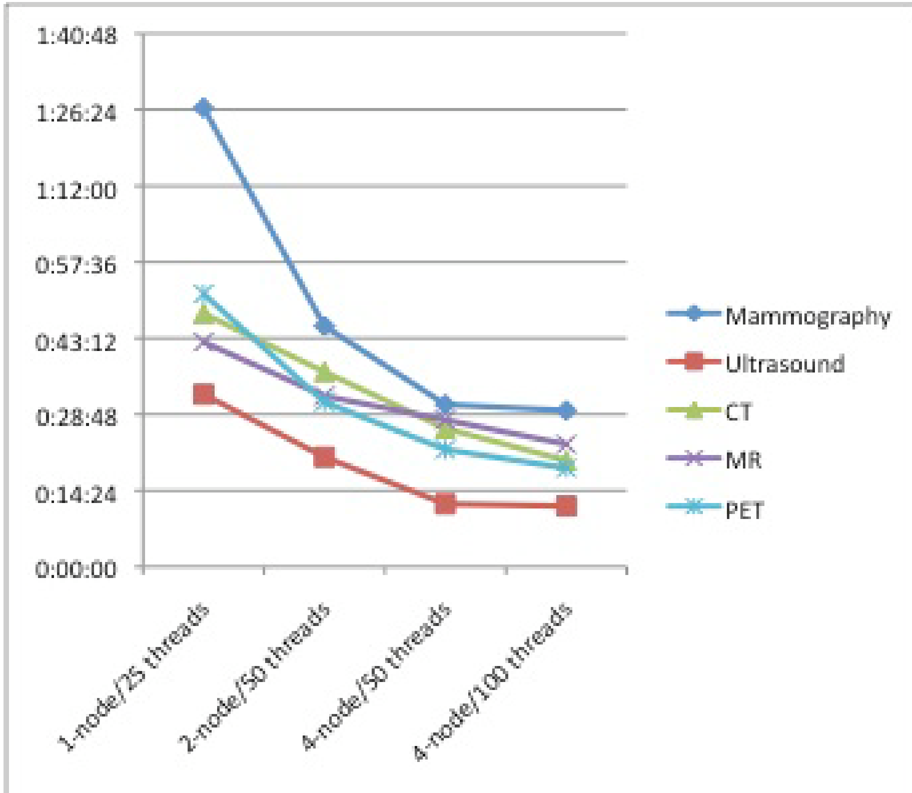

Figure 6 Combined Process Totals

On average, total dataset throughput times were between 1.8 and 2.9 times faster on the 4-instance ***Enterprise Hardware***/100 connections configuration than on the 1-instance/25 connections configuration.

6.6 Process #5. IBISimg DICOM View: Throughput and Scalability Analysis on Enterprise Hardware

The image viewing operation is different from the other operations in some key aspects. Firstly, in order to match a real world scenario, we have had to simulate the DICOM clients. (The alternative would have been to deploy up to 100 separate DICOM client workstations, which was not feasible). Secondly, in this run, we transfer image data over the DICOM protocol, using DICOM CFIND and CMOVE commands as opposed to using the HTTP protocol. The DICOM protocol is “chatty” and is less efficient than the HTTP protocol however, as of this time, it is the only option available for communication with a DICOM viewing device. Finally, and most importantly, whereas for all other runs in this benchmark, client connections (“threads”) are used to perform parts of the same job, in this case, as we add DICOM client connections, each DICOM client initiates its own new study retrieval process. In other words, as the number of client connections grows, the volume of data transferred grows equivalently. Therefore, scalability will be demonstrated if we can come close to retaining throughput levels as we raise the number of client connections.

We have measured the “time to first image” by which we mean the time it takes to complete the transfer of the first image to the DICOM viewer.

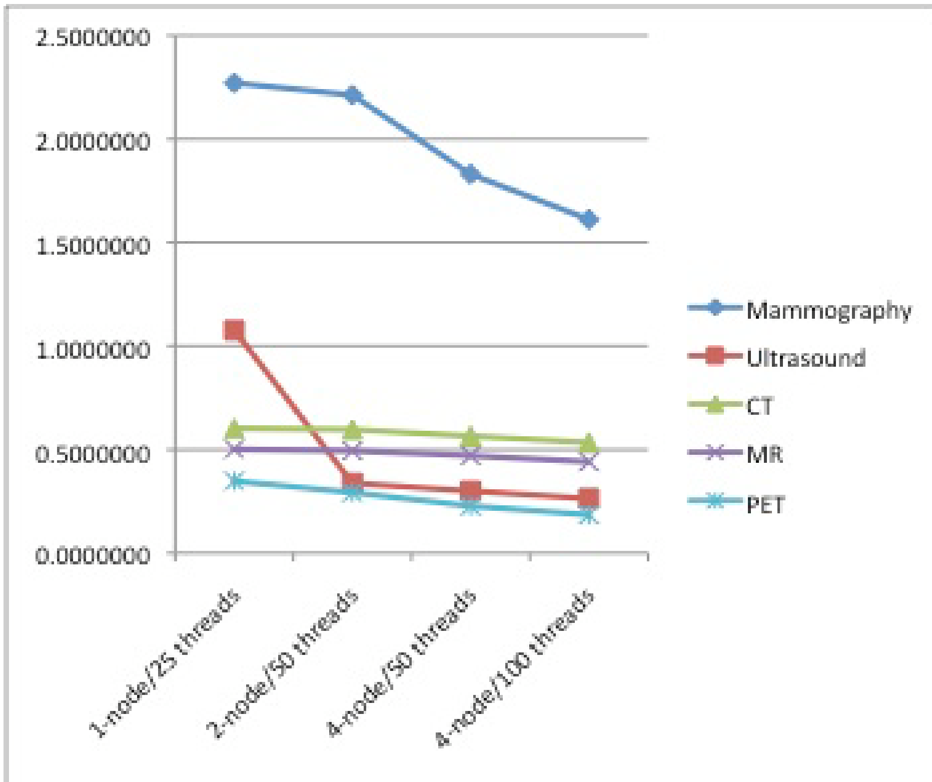

Figure 7 DICOM Viewer

The key differentiating factors for the DICOM Viewer are the size of the DICOM files and the number of files per study. For example, in the case of Mammography, the study is a single file with, on average, a size of about 32MB, whereas in the case of PET, the study is made up of, on average, 227 files, with a total size of only 2.5MB.

Observation #6: The DICOM Viewer performs consistently across diverse datasets, and sustains or improves throughput with increased connections/ Enterprise Hardware instances

Observation #7: For larger datasets, the data download operation (over the HTTP protocol) performed consistently at a much higher rate than the DICOM protocol

7.0 Conclusions

Five major conclusions are supported by the data:

- All major IBISimg functionality on Enterprise Hardware scales as the number of user connections is increased
- All major IBISimg functionality on Enterprise Hardware scales as the number of Enterprise Hardware instances is increased
- IBISimg can process 1TB of data in under an hour
- Time to first image on DICOM viewer is sub-second for MR, PET and CT

- IBISimg sustained user response times as it scaled up to 800 concurrent user connections.