

A Study Comparison of Store and Retrieve of DICOM Images using High Performance Data Format

*Douglas D. J. de Macedo, Thiago Coelho, Eros Comunello and M.A.R. Dantas
Post-Graduate Program in Computer Science (PPGCC)
Federal University of Santa Catarina (UFSC) - Brazil*

*Michael A. Bauer
Department of Computer Science
University of Western Ontario
Canada*

Agenda

- Motivation
- DICOM Images
- HDF5
- Experimental Environment and Results
- Conclusions and Future Work

Agenda

- Motivation
- DICOM Images
- HDF5
- Experimental Environment and Results
- Conclusions and Future Work

Motivation

- It is recognized that the *large amount of scientific data* is one *challenge* for *high performance computing environments*.
- In this research work, it is presented a study research in terms of *store* and *retrieve* of *DICOM images*. These images, which are *conventionally stored at relational databases*, can easily *exceed terabytes*.
- Therefore, a study on *how to improve the manipulation* of this data, with a satisfactory level of high performance, is an interesting contribution.

Agenda

- Motivation
- **DICOM Images**
- HDF5
- Experimental Environment and Results
- Conclusions and Future Work

DICOM Images

- The requirement for a standard medical *images exchange among medical equipment*, from different manufacturers, *culminated* in the conception of the *DICOM standard*.
- *DICOM* was *designed* by the American College of Radiology (*ACR*), in cooperation with the National Electrical Manufacturers Association (*NEMA*).
- *After* 1987, when the *standard was conceived*, two *new revisions occurred*. *Finally*, in *1992*, it was published the ACR-NEM Standards Publications PS3, also known as *DICOM 3* (Digital Imaging and Communications in Medicine).

DICOM Images

- In the *figure 1* it is possible to visualize *an example of DICOM image, retrieved from a computed tomography* from the spine and cranium.
- Currently the *DICOM 3* is a *de facto standard* for the Picture Archiving and Communications System (PACS), being supported by the majority of equipment that works with digital medical information. It is *relevant to mention* that a *DICOM image* includes the *picture of the exam*, also *large amount of metadata* related to the patient that is embedded within the image.



Fig. 1. Example of DICOM Image

Agenda

- Motivation
- DICOM Images
- **HDF5**
- Experimental Environment and Results
- Conclusions and Future Work

HDF5

- The Hierarchical Data Format (*HDF*) is a *portable data format* developed at Illinois University. It is a *library for high performance manipulation of scientific data*.
- It was *designed* to enable *storage of objects with large data volume, multidimensional arrays* and *tables with large datasets*. Those elements can be utilized together to support applications.

HDF5

- The *HDF acronym* it is also *refers to a set of software, applications interfaces and utilities* that compound the library and that enable users *to manipulate files inside this format*.
- *Software* packages are been developed since 1988, the *actual version* is the *HDF5*.

Agenda

- Motivation
- DICOM Images
- HDF5
- **Experimental Environment and Results**
- Conclusions and Future Work

Experimental Environment and Results

- In our *experimental storage and recovery tests*, *twenty-five images were sent and retrieved utilizing the DICOM server*. Figure 2 shows our two distinct case studies.
- In *both cases*, it was used *only one medical image using the DICOM standard*, considering two different implementations developed by the wrapper from the HDF5 library.

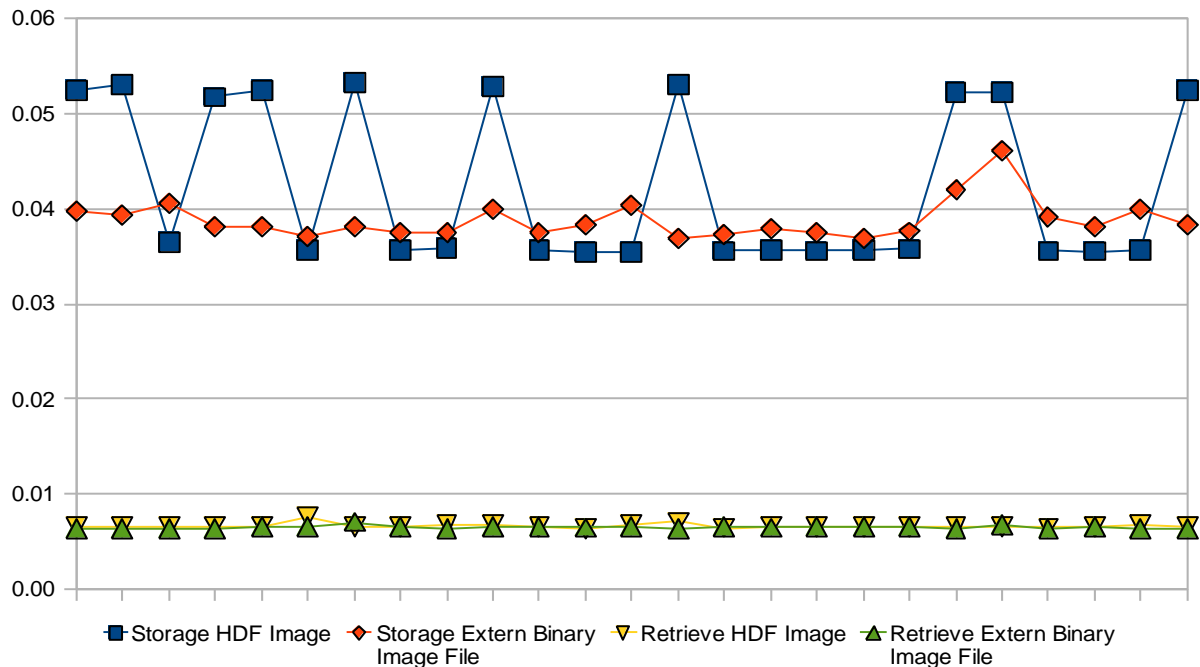


Fig. 2. Graph of storage and recovery tests

Experimental Environment and Results

- It is possible to notice from the comparison highlighted in *figure 2* that *when the H5IM* (the HDF5 high-level library for pictures) *is used there is a large oscillation in the writing time.*
- On the other hand, the *behavior is the opposite* when the *picture storage occurs externally to the HDF5 file.*
- Relating to *the recovery* of medical image operation, it is clear that even utilizing different approaches, *there is no significant difference in terms of elapsed time.*

Agenda

- Motivation
- DICOM Images
- HDF5
- Experimental Environment and Results
- **Conclusions and Future Work**

Conclusions and Future Work

- In this research we have presented *a comparison* between two *different methods* to *save and retrieve large DICOM medical images*, employing the HDF5.
- The *first approach*, considering a picture as an external element for the *HDF5 file*, has a *better performance to save* in comparison to a file compound of metadata and picture.
- However, considering the *retrieve operation* both approaches have reached the *same performance level*.

Conclusions and Future Work

- As a *future work* we are planning to utilize these two approaches through a more *complex distributed environment*, such as *multi-clusters*.
- In addition, draw a comparison with a relational database in the same configuration.

Thank you !

Questions ?

Mario Dantas
mario@inf.ufsc.br