Architecture of distributed picture archiving and communication systems for storing and processing high resolution medical images

Victoria Tokareva

Laboratory of Nuclear Problems

Joint Institute for Nuclear Research

tokareva@jinr.ru

October 4, 2017

AYSS'17

1.
BIG DATA AND ANALYTICS TOOLS

At modern medical imaging

Medical imaging

is the technique and process of creating visual representations of the interior of a body for clinical analysis and medical intervention, as well as visual representation of the function of some organs or tissues.

Including:

- Computer tomography (X-ray imaging);
- Magnetic resonance imaging (MRI),
- Medical ultrasonography
- Positron emission tomography (PET)
- Single-photon emission computed tomography (SPECT)
- etc.

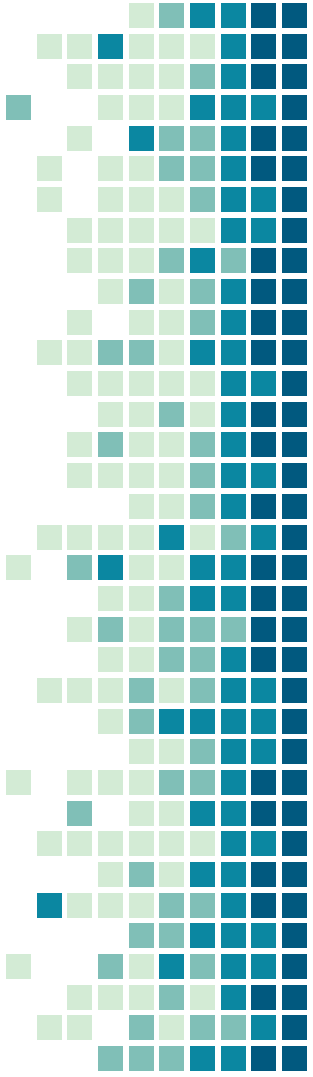
Computed tomography (CT) is a technique for imaging cross-sections of an object using a series of X-ray measurements taken from different angles around the object.

The art of medical computed tomography is constantly evolving and the last years have seen new ground breaking systems with multi-row detectors. These tomographs are able to increase both scanning speed and image quality compared to the single-row systems more commonly found in hospitals today.

Henrik Turbell

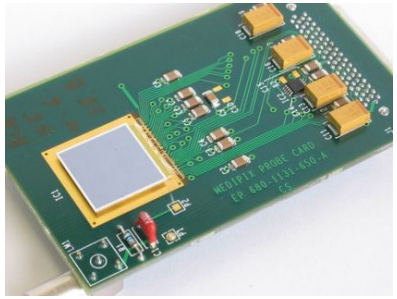
Trends and challenges

- Hardware evolution
- Data growth
- Algorithms improving
- Growing system complexity



MEDIPIX DETECTOR TECHNOLOGY

Medipix detector chip is a new generation x-ray detector, developed by the Medipix3 Collaboration.



Important features:

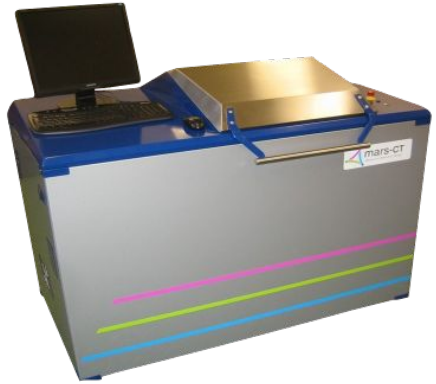
- energy resolving
- photon counting

Perspectives:

- spectral molecular imaging technology;
- allows to make radiology procedures faster and at lower cost;
- significantly decreases radiation dose during the study.

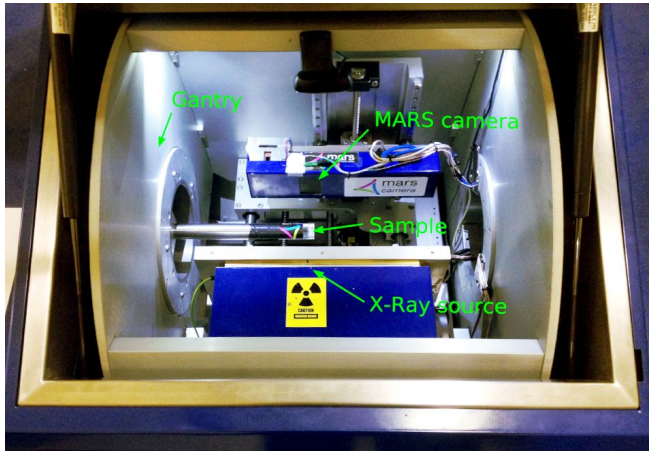
A **Medipix All Resolution system (MARS)** CT scanner obtained by DLNP JINR incorporates the Medipix detector chip.

It is used for developing new generation detectors based on Medipix chips with silicon or GaAs sensors.



Scanning on MARS microCT

The gantry (X-ray source and camera) with the scanning equipment is attached to the MARS tomograph and rotated around the scanned sample. The gantry rotation axis is horizontal. A test sample (up to 100 mm in diameter and 300 mm length) is placed in the center and can be moved along the rotation axis.



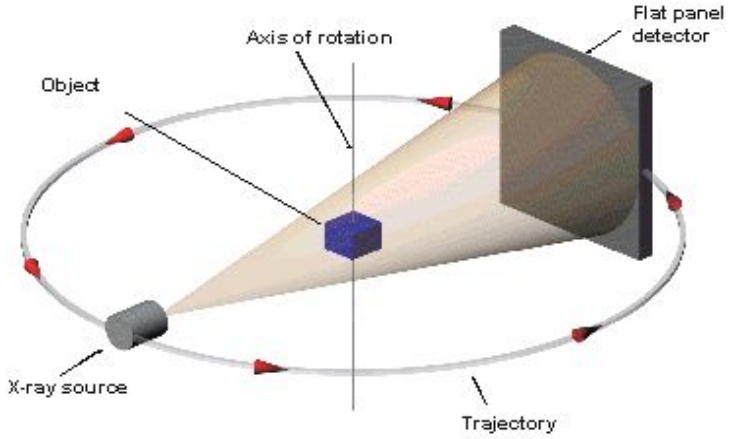
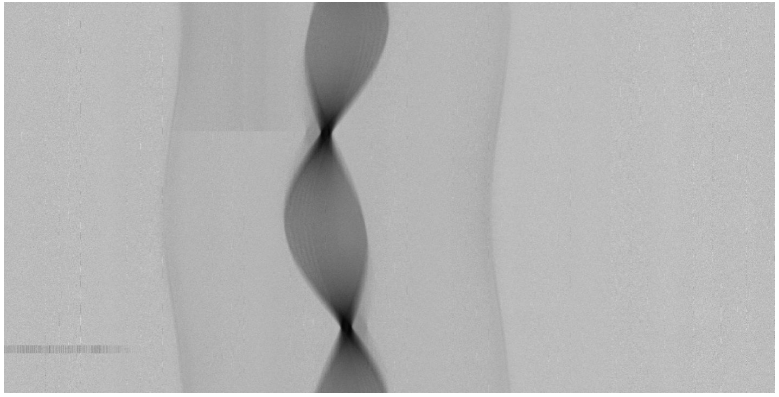
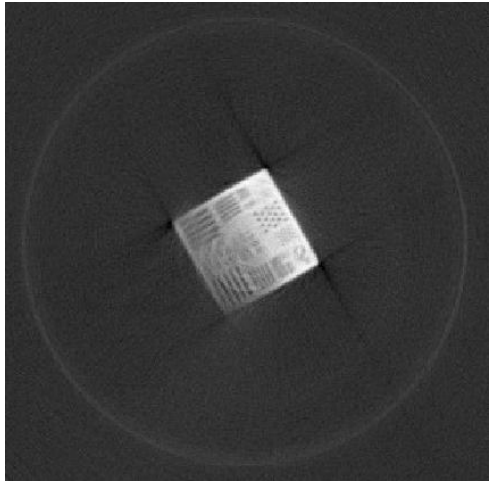


Image reconstruction and analysis

The result of scanning is a set of shadow projections obtained for different angles. The input data after preprocessing is given to reconstruction program as a set of filtered synograms. The program reconstructs the slices of the scanned object.





An unique certainty

- The x-ray source spectrum is 120 kVp;
- Micro-focus with a focal spot size of <100 μm;
- Available energy range (20–140 keV);
- High spatial resolution (~55 microns);
- Up to 720 angles of scan.

...And data amounts

Subject	Size, cm	Angles	Size of data (uncompressed)
Stone	1 * 1 * 3	360	12 Gb
Arteria (piece of)	3 * 5 * 5	360	60 Gb
Thorax	≈ 20 * 20 * 30	360	≈ 2,5 Tb
Thorax	≈ 20 * 20 * 30	720	≈ 5 Tb
Full body	≈ 80 × 80 × 210	720	≈ 150 Tb

Computational complexity

Image reconstruction is the most computationally expensive part.

For example, for the FDK (Feldkamp, Davis & Kress, 1984) algorithm, that is the most widely used for cone-beam geometry, we have the computational complexity of $O(N^4)$.

General scheme components

- Medical device connected to the PC (an assistant workstation or modality);
- Data storage;
- Data base;
- Server for image reconstruction;
- Viewer of reconstructed samples.

2. INDUSTRY STANDARDS

And technical requirements

DICOM Standard

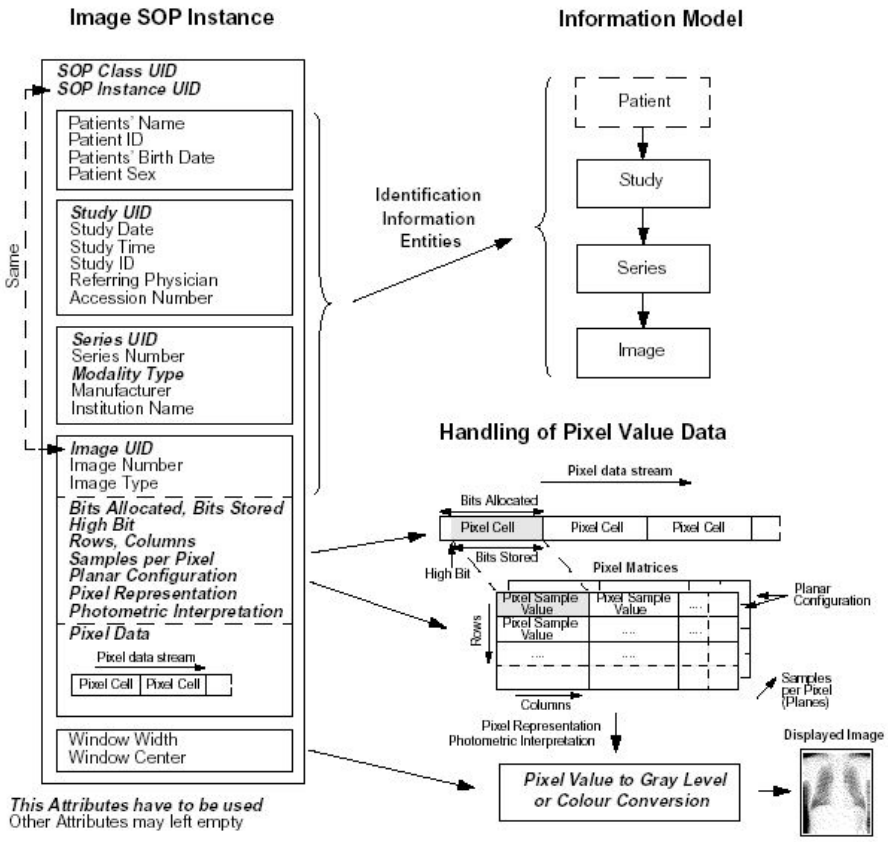
Digital Imaging and Communications in Medicine (DICOM) — is the international standard for medical images and related information (ISO 12052).

DICOM defines:

- Formats for medical images
- Necessary quality
- Ability of system to:
  - Produce;
  - Store;
  - Display;
  - Process;
  - Send;
  - Retrieve;
  - Query;
  - Print medical images and related documents.

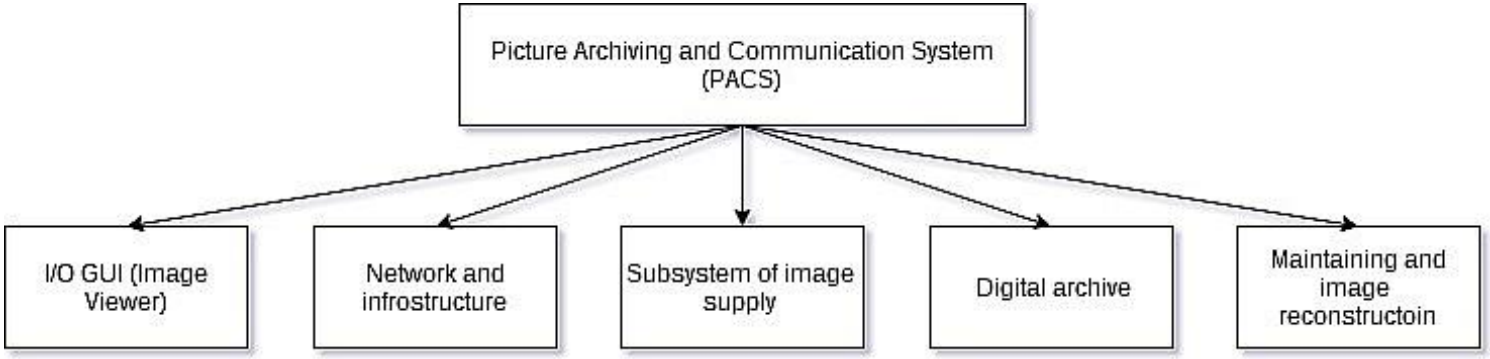
http://dicom.nema.org/Dicom/about-DICOM.html

DICOM INFO MODEL

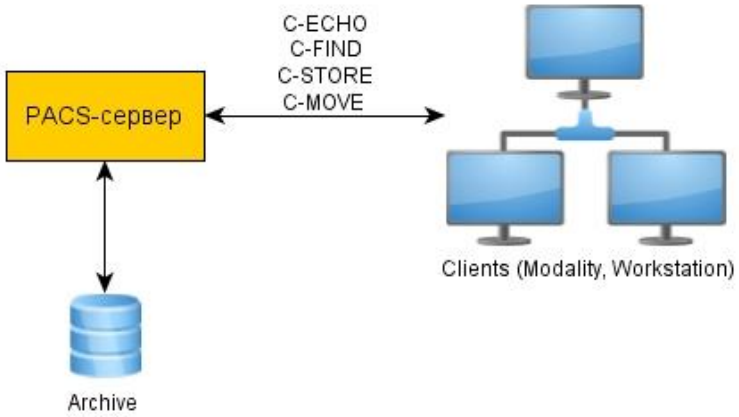


Picture Archiving and Communication System

(PACS) is a system based on the universal DICOM standard, which uses a server to store and allow facile access to high-quality radiologic images, including conventional films, CT, MRI, PET scans and other medical images over a network.

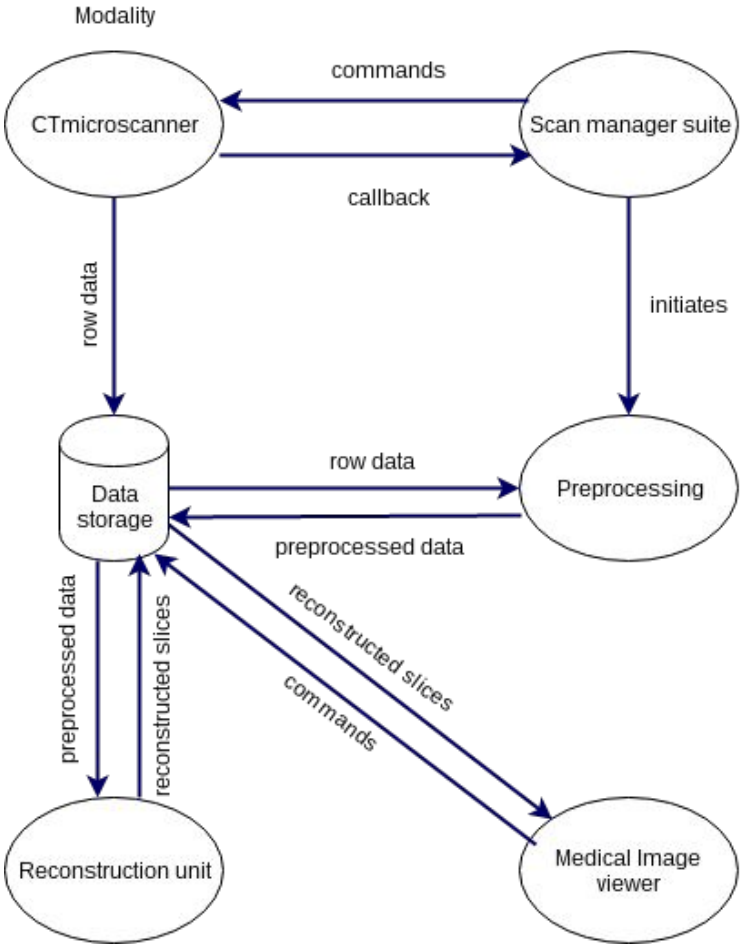


General PACS scheme

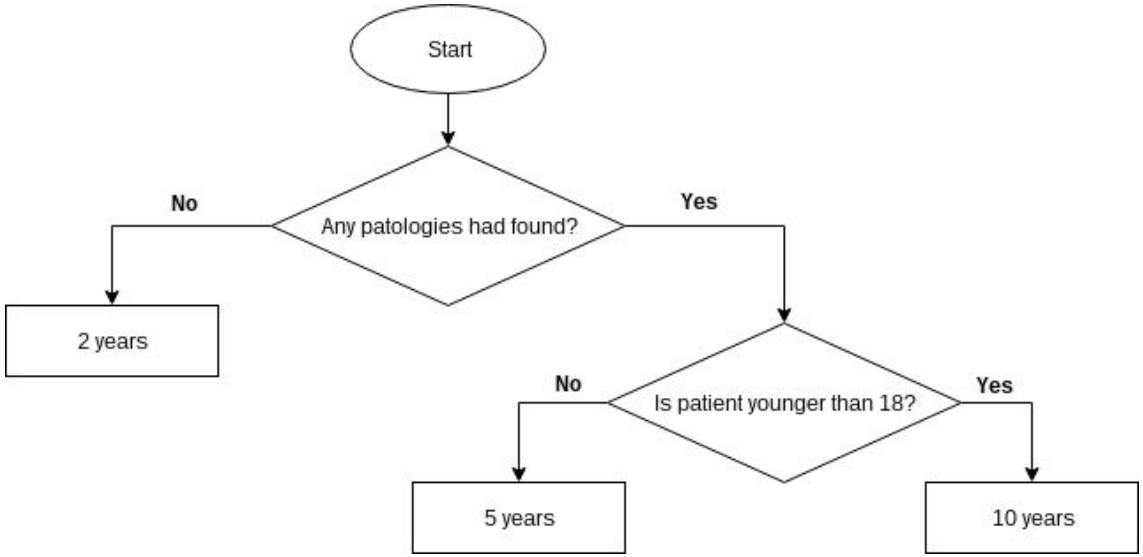


Source: https://habrahabr.ru/post/193134/

Simple PACS Scheme



Storage requirements for medical images in Russian Federation



According p. 12 Application 24 of direct №132 of Ministry of Health

3. MAKING PACS DISTRIBUTED

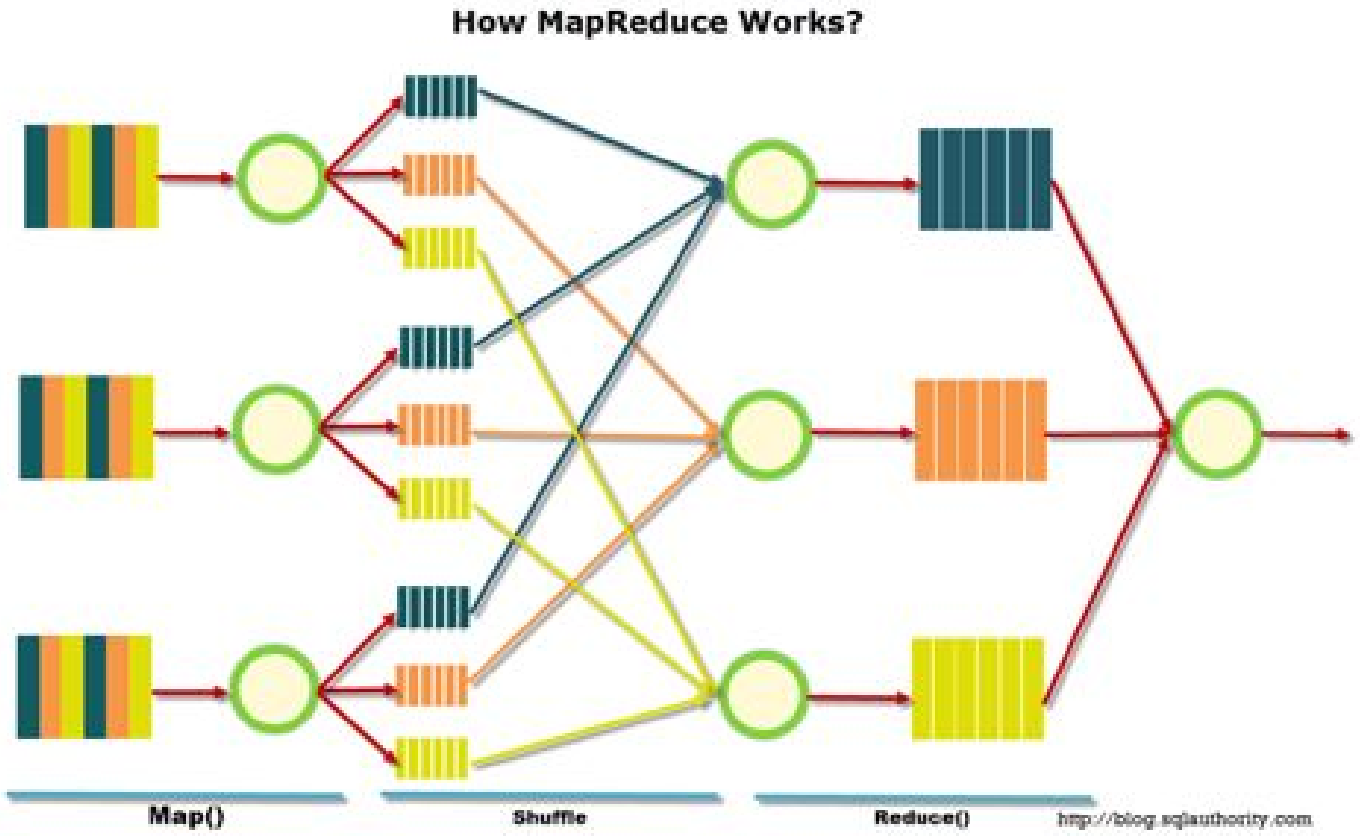
Problems and solutions

Hadoop

is an open-source software framework used for distributed storage and processing of dataset of big data on computer clusters.

The core of Apache Hadoop consists of a storage part, known as Hadoop Distributed File System **(HDFS)**, and a processing part which is a **MapReduce** programming model. Hadoop distributes files across nodes in a cluster, taking advantage of data locality.

Hadoop can be deployed in a traditional onsite data center as well as in the cloud.



https://blog.sqlauthority.com/2013/10/09/big-data-buzz-words-what-is-mapreduce-day-7-of-21/

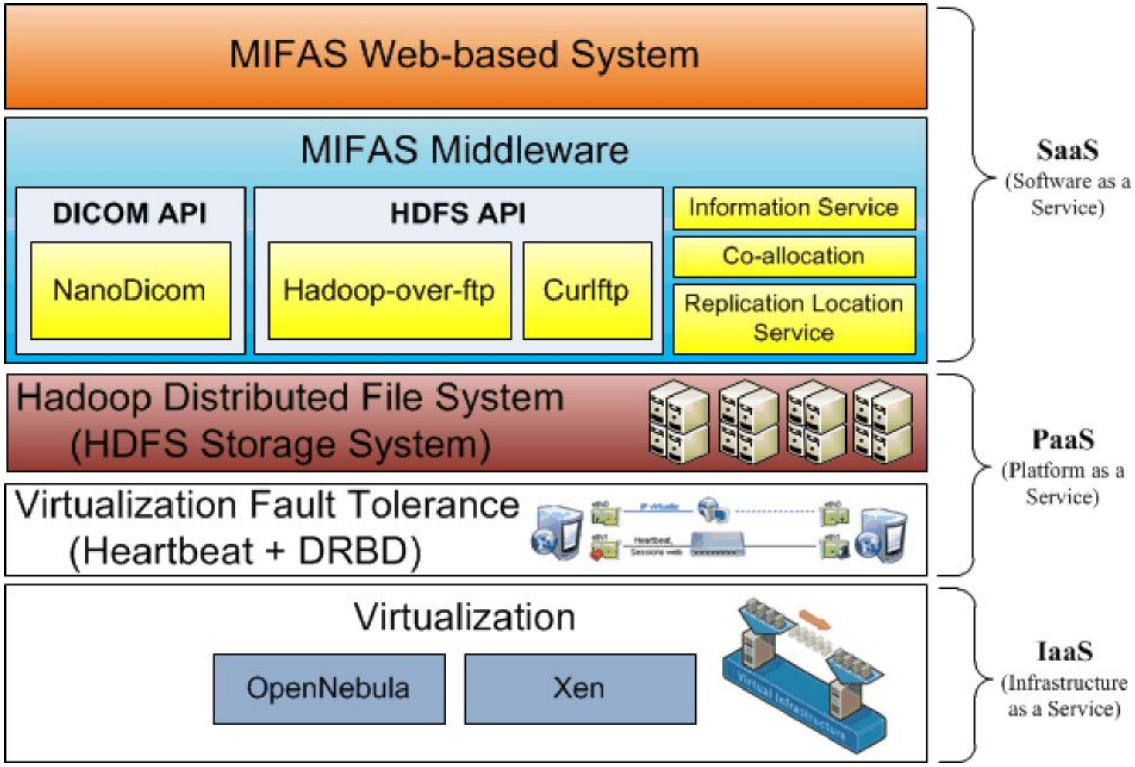
Making PACS distributed

Advantages

- Scalability
- Cost effective
- Replication

Disadvantages

- Image retrieval time



C.-T. Yang et al. A Medical Image File Accessing System with Virtualization Fault Tolerance on Cloud, GPC 2012, LNCS 7296, pp. 338–349, 2012.

New challenges here?

Of course!

- Shifting focus from storage-oriented PACS to processing-oriented PACS;
- Distributed data preprocessing and reconstruction;
- Looking for other ways to speed up the system

THANKS!

Any questions?

Contact me: tokareva@jinr.ru