ROOT2002 Workshop
CERN, October, 16th, 2002

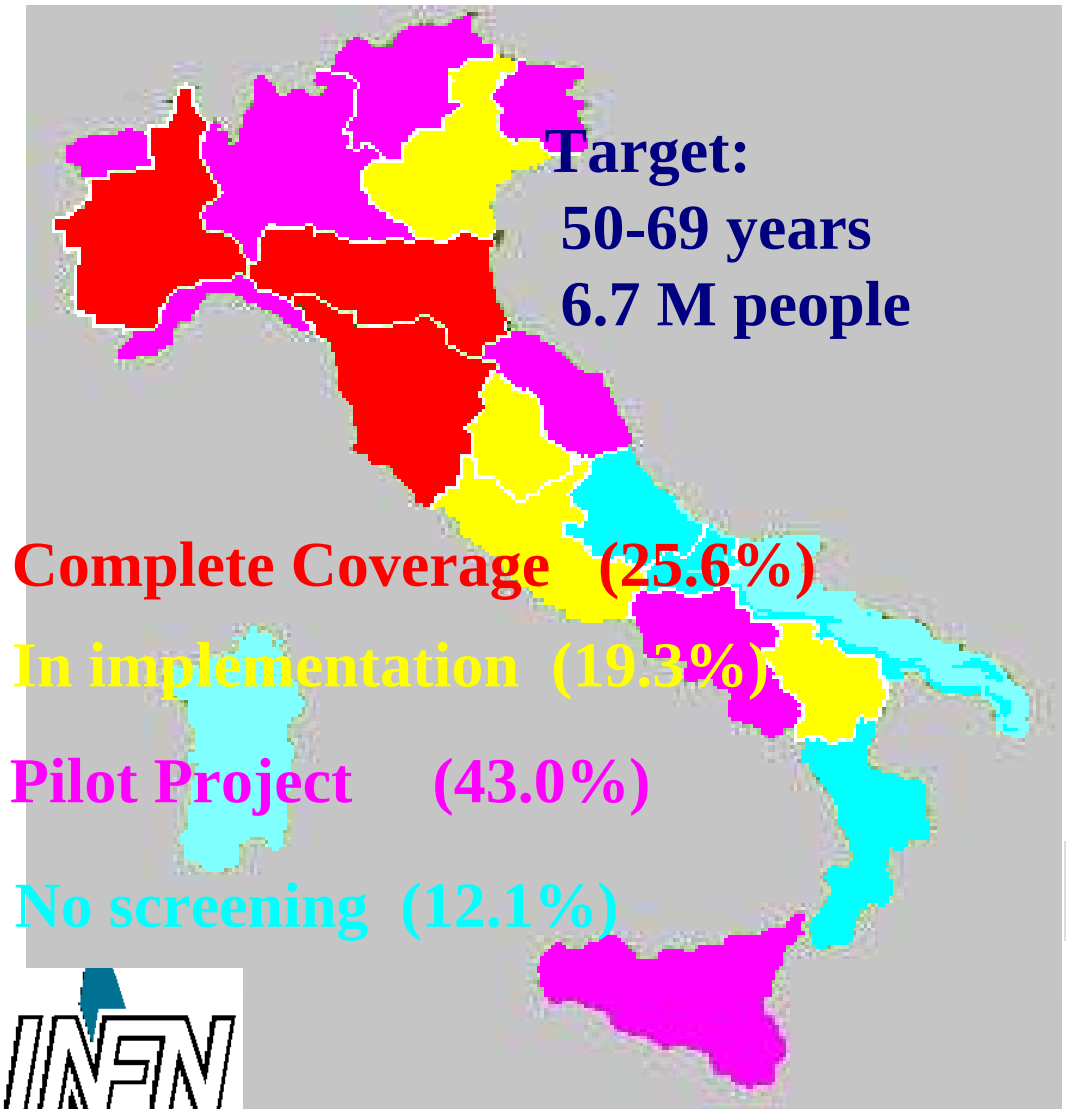
Computer Assisted Diagnosis in Mammography with PROOF

Ernesto López Torres
INFN, Sezione di Torino

- **The CALMA project**
  + GRID + PROOF
- **Image Processing**
- **GPCALMA: present situation and prospects**

GPCALMA

Mammographic Screening in Italy
(October 1999)



Grid Platform for a Computer Assisted Library for Mammography

Participants

INFN & Universities:

Cagliari, Catania, Napoli, Palermo, Pisa, Sassari, Torino

Hospitals:

Bari, Livorno, Napoli, Palermo, Sassari, Torino, Udine

Mammographic Screening

Sensitivity → 73% - 88%

(classified positives/true positives)

Specificity → 83% - 92%

(classified negatives/true negatives)

2% - 10% increase with ***cross-check, but:***

Time & Cost problems

→ ***Computer Aided Detection***

The CALMA Project (1998-2001)

Mammographic Database: about 5,000 images, 12 MB each

85μm step, 4096 colours

The largest in Europe

- Computer Aided Detection:
  - tissues
  - damages
  - microcalcification clusters

Tested as cross-check diagnosis

- WorkStation for: digitazation
  - archiving
  - human diagnosis
  - computer assisted diagnosis
  - statistical analysis
  - training

In Use in several hospitals

INFN

Computer Aided Detection Results

Tissue	DENSE (CAD)	ADIPOSE (CAD)	GLANDULAR (CAD)
DENSE (true)	100%	0%	0%
ADIPOSE (true)	15%	68%	17%
GLANDULAR (true)	4%	3%	93%

Damages

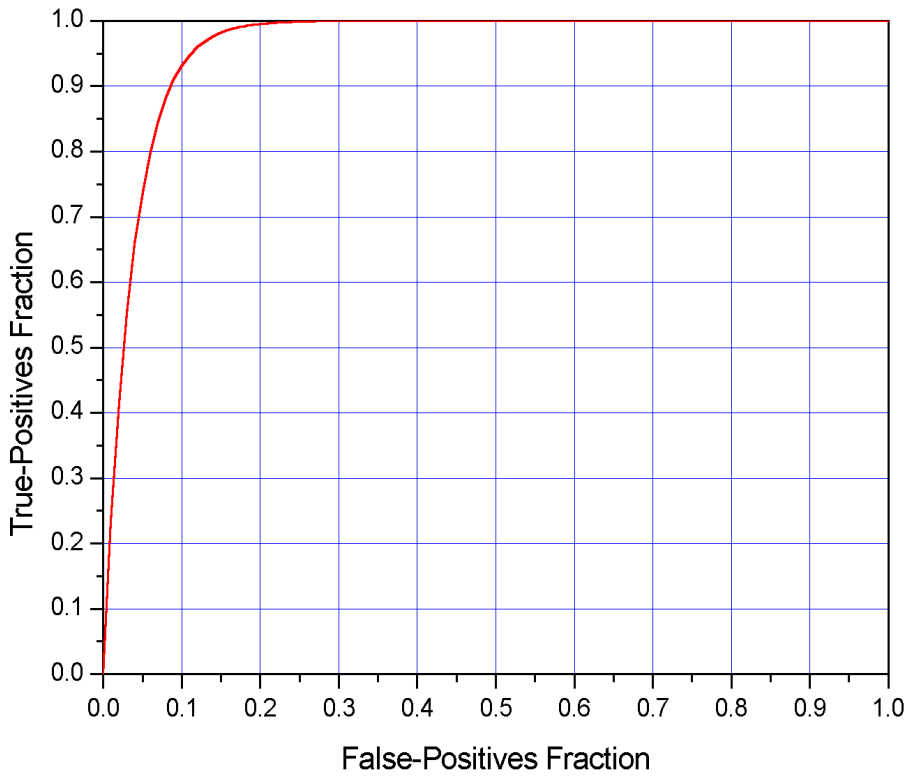
SENSITIVITY > 90%

AVERAGE NUMBER OF FALSEPOSITIVES/IMAGE: 1.4

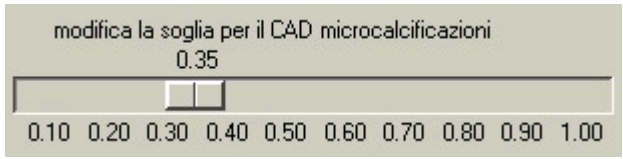
AVERAGE AREA OF THE SELECTED Region Of Interest: 35cm^2

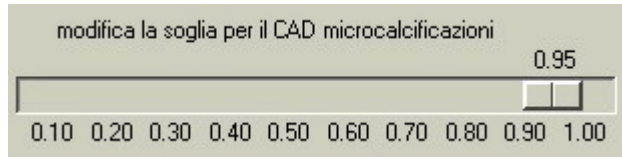
INFORMATION REDUCTION FACTOR: 12

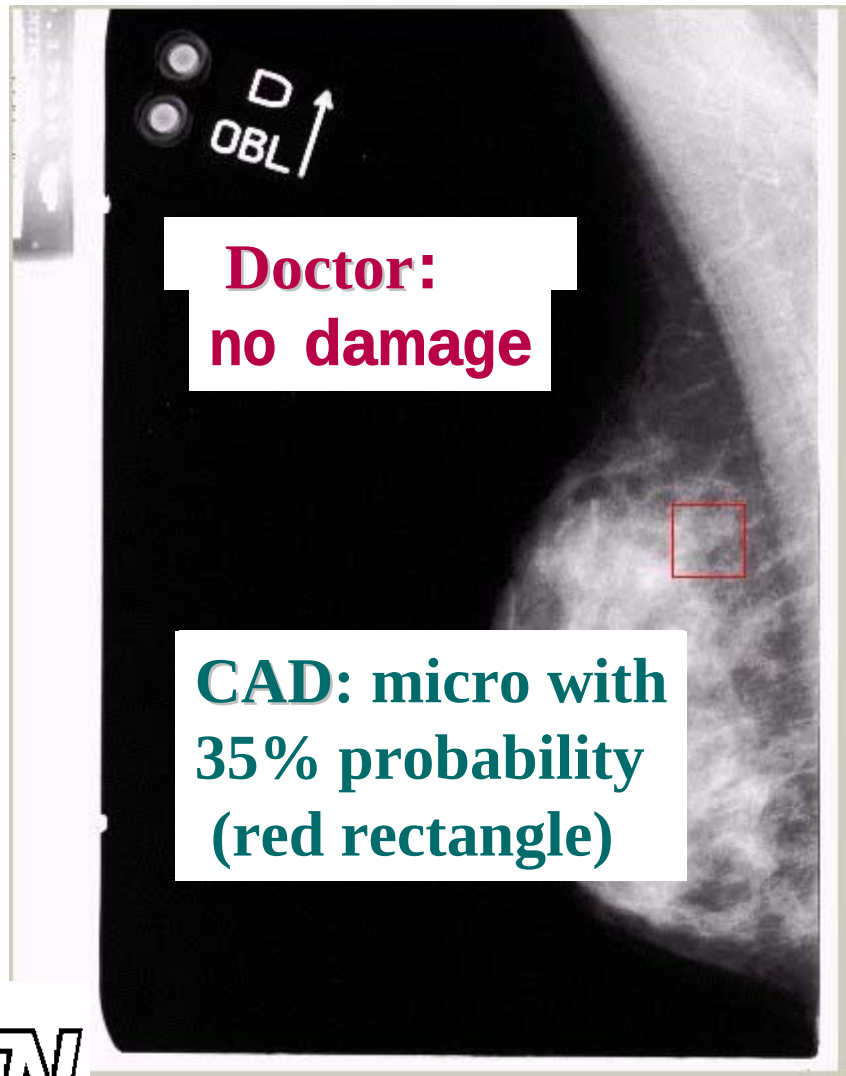
865 mammograms (370 *with* 495 *without* microcalcification clusters)



CALMA GUI: example



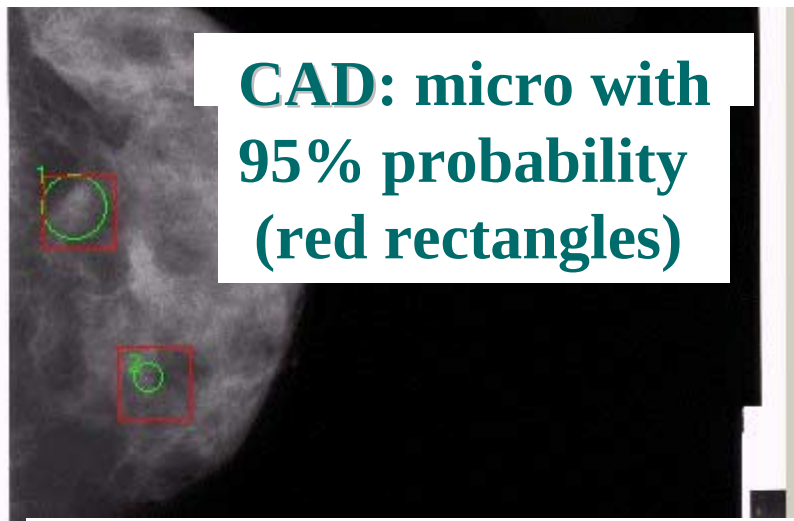




Doctor: no damage

CAD: micro with 35% probability (red rectangle)

Doctor: granular micros, radiologically suspicious (green circles)

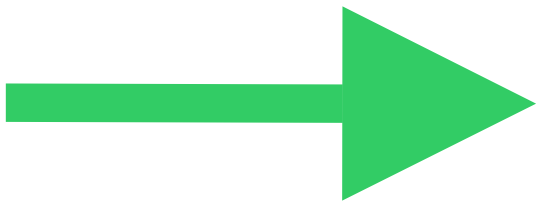


CAD: micro with 95% probability (red rectangles)

Histology: granular micros (breast cancer)

CALMA: Mammographic Screening

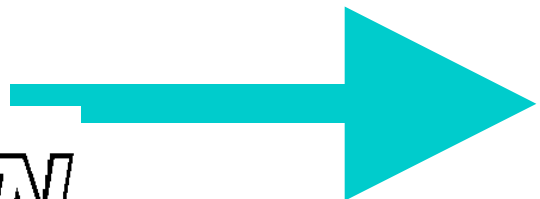
- Rapidly increasing database (virtually unlimited)
- Access required to all the images
- **Teleradiology**
  - Telediagnosis and Teletraining



the "GRID philosophy" in mammographic CAD

. **Intrinsically Distributed Database**

. **Network conditions**



Move code rather than data

Which technology?

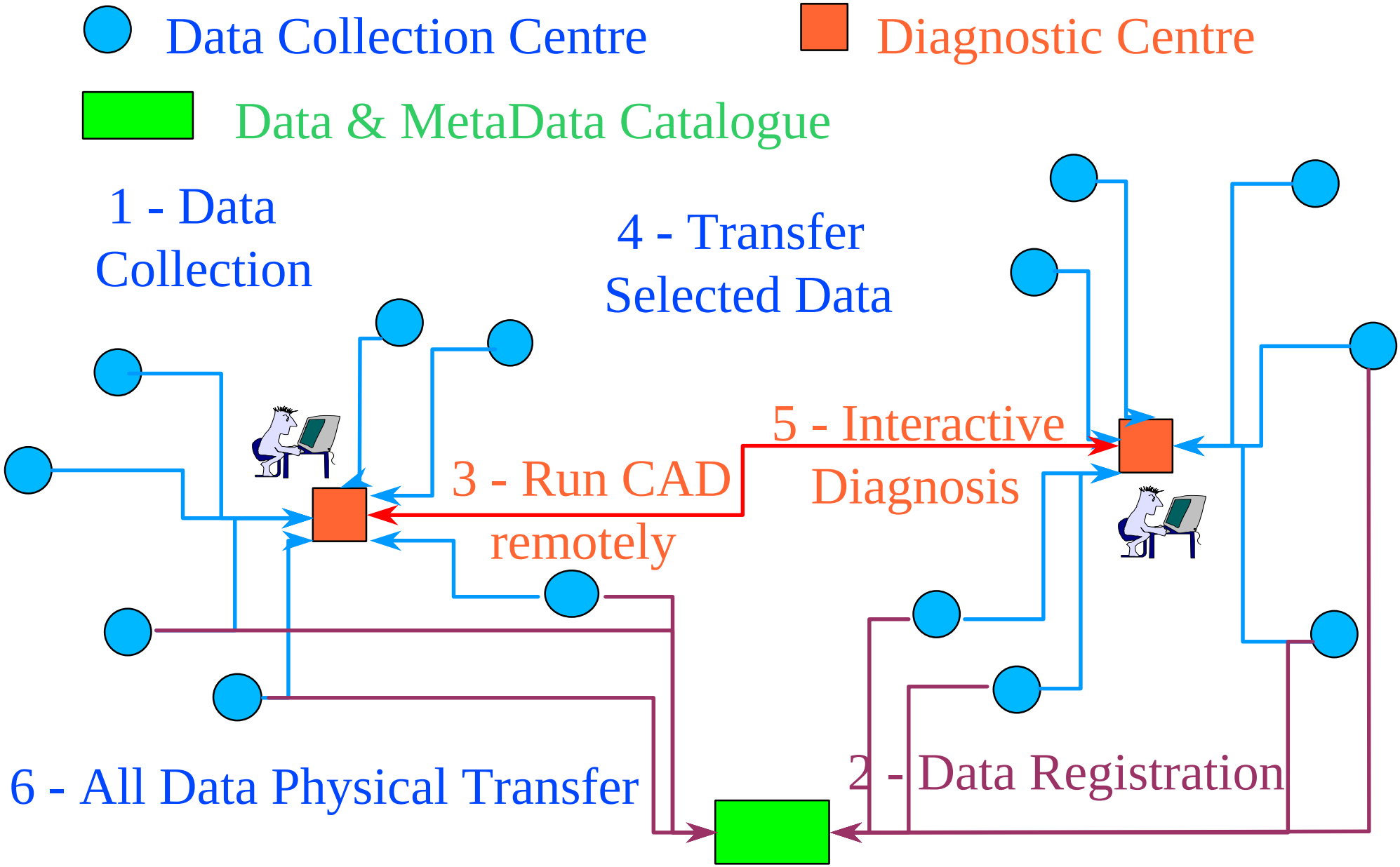
ROOT + PROOF

- CINT
- Interactive development
- Move the code rather than the images!
- Graphic User Interface
- Integration/Interface with GRID-like Services

Goal

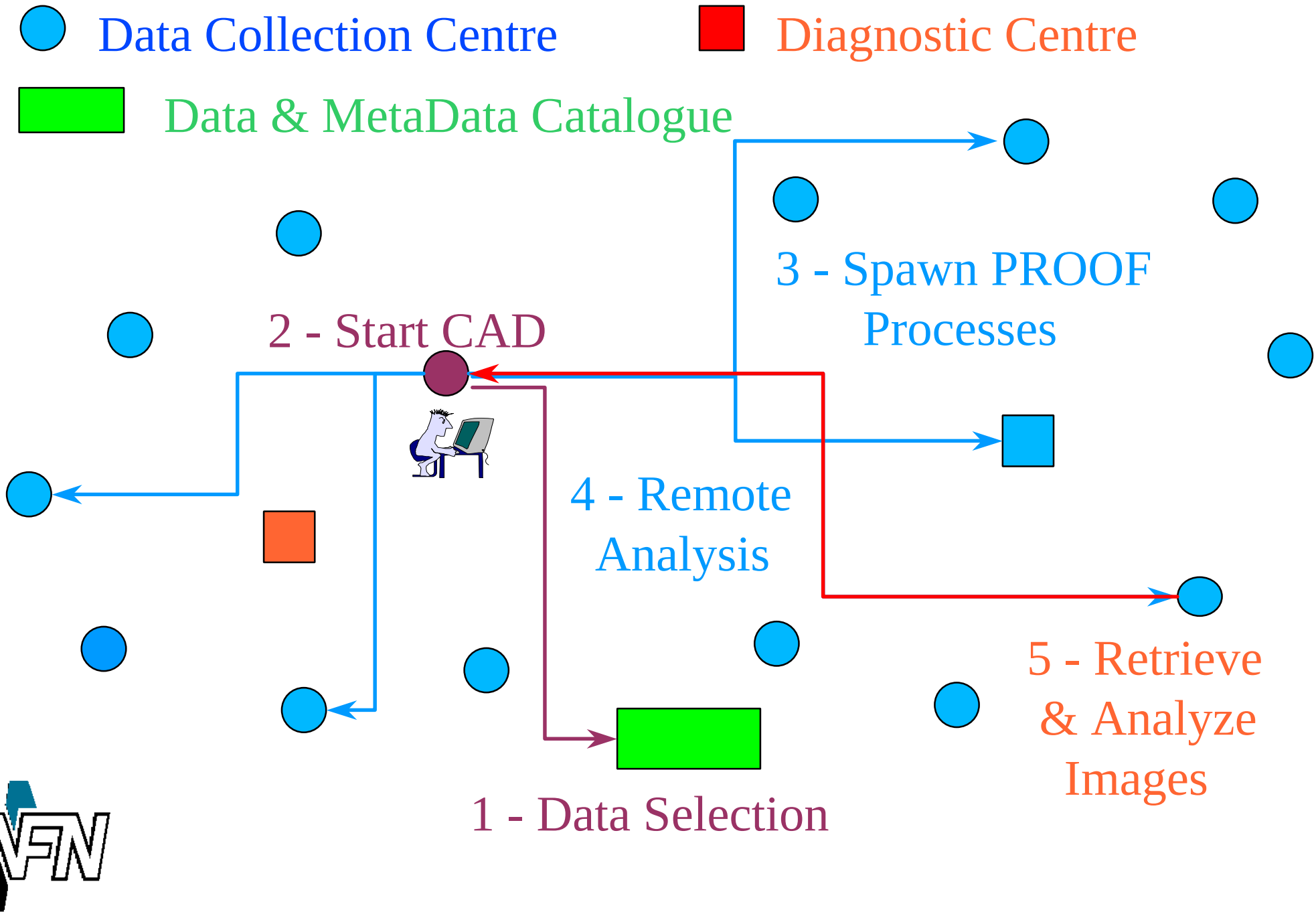
- Set up a prototype of the
  GPCALMA Virtual Organisation
- Hardware in n nodes
  Hospitals for the image storage
  INFN sites for the configuration of services
  (PROOF + GRID)

GPCALMA: telediagnosis for screening



CAD selection to minimize data transfers for quick diagnosis

GPCALMA: teletraining, epidemiology



GPCALMA Ongoing Activities

- Microcalcification and Masses Algorithms available
- Image treatment (see next slide)
- Image format: CALMA/DICOM/ROOT Interface
- Source code structure "a la AliRoot"
- CVS Server to public the source code being set up
- MetaData Catalogue Design

Workplan

- Configuration of the User Server (AliEn?)
- Configuration of the MetaData Catalogue (MySQL, through AliEn?)
- Implementation of the API to register/access the Data

ROOT developments on Image Treatment

New features added to TImage (TASImage) class:

- continuous zoom
- navigation in the image
- contrast and brigthness control
- reversed image colors
- center image
- selection between contextmenu and floating toolbar (new class TASImageToolBar)

Could they be integrated in the ROOT release?

PROGRAM DEMO...

Summary

GPCALMA: Management of a Distributed Database of Mammographic Images

- *Telediagnosis for Screening*
- *Teletraining*

Original ROOT application outside HEP
Reduced-functionality GRID system
PROOF testing
ROOT-AliEn interface development