UMONS
Université de Mons

INFORMATIQUE PARALLÈLE ET DISTRIBUÉE

CHAPTER 7 : CLOUD & GPU FOR MULTIMEDIA RETRIEVAL

Sidi Ahmed Mahmoudi

POLYTECH MONS

sidi.mahmoudi@umons.ac.be

19 December 2017

PLAN

Introduction

I. Multimedia retrieval approaches

II. Performance analysis

III. Software and hardware requirement

IV. Cloud-based solution for multimedia retrieval

- Platform architecture
- Webservers and clients management
- Data encryption
- Applications management

V. Experimentations

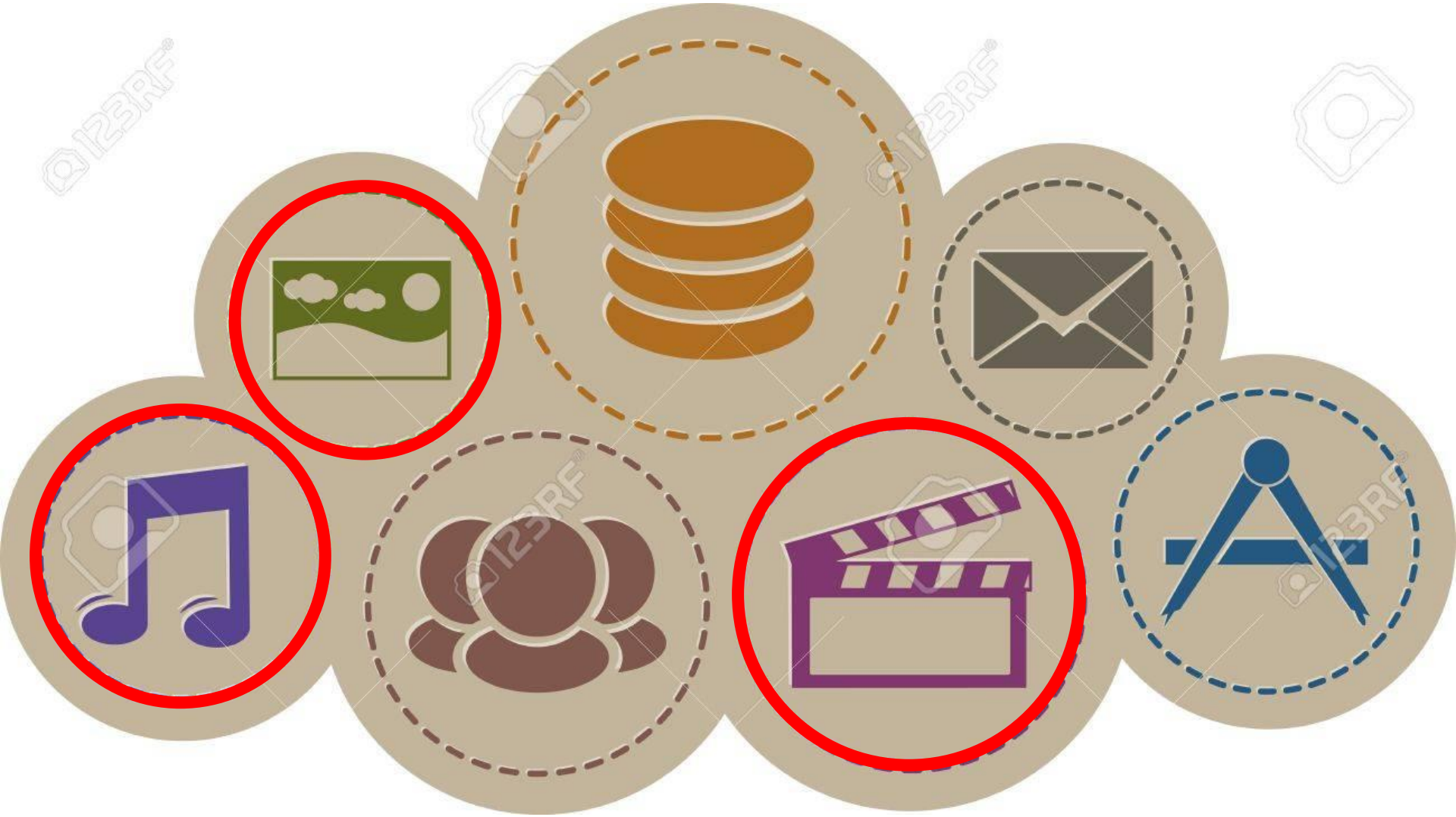
VI. Demonstration

Conclusion

Introduction

- Multimedia data source include contents like text, images, videos, audio or a mixture of them
- Multimedia objects are present everywhere : laptops, computers, smartphones, etc.
- Multimedia Indexing and Retrieval (MIR) manage and facilitate searching for multimedia data

Introduction



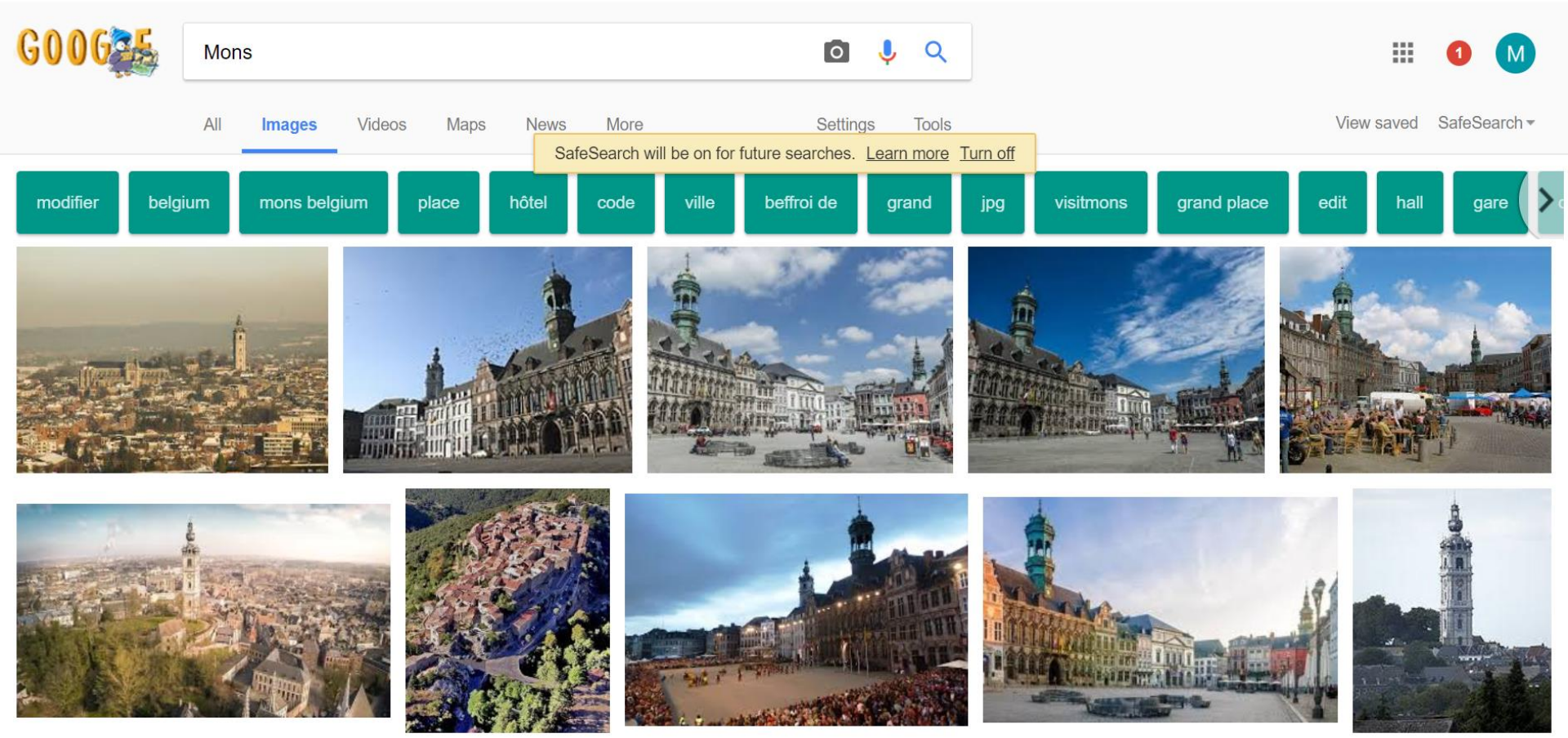
Multimedia retrieval approaches

Main methods of multimedia retrieval and indexation :

- Text based indexation (exp. Google)
- Content based Image Retrieval (CBIR)
    - Syntactic image features
    - Semantic image features

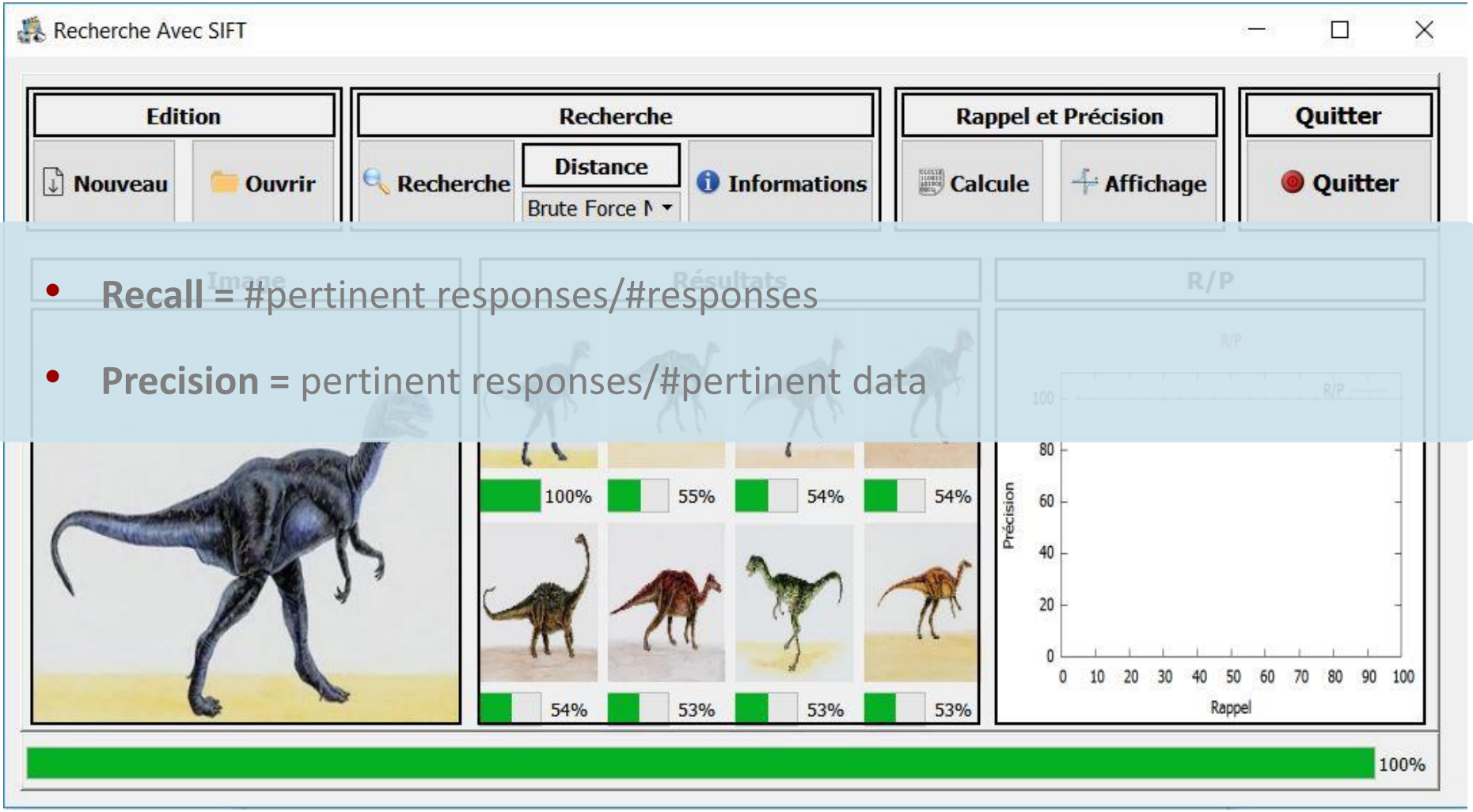
Multimedia retrieval approaches

Text based indexation



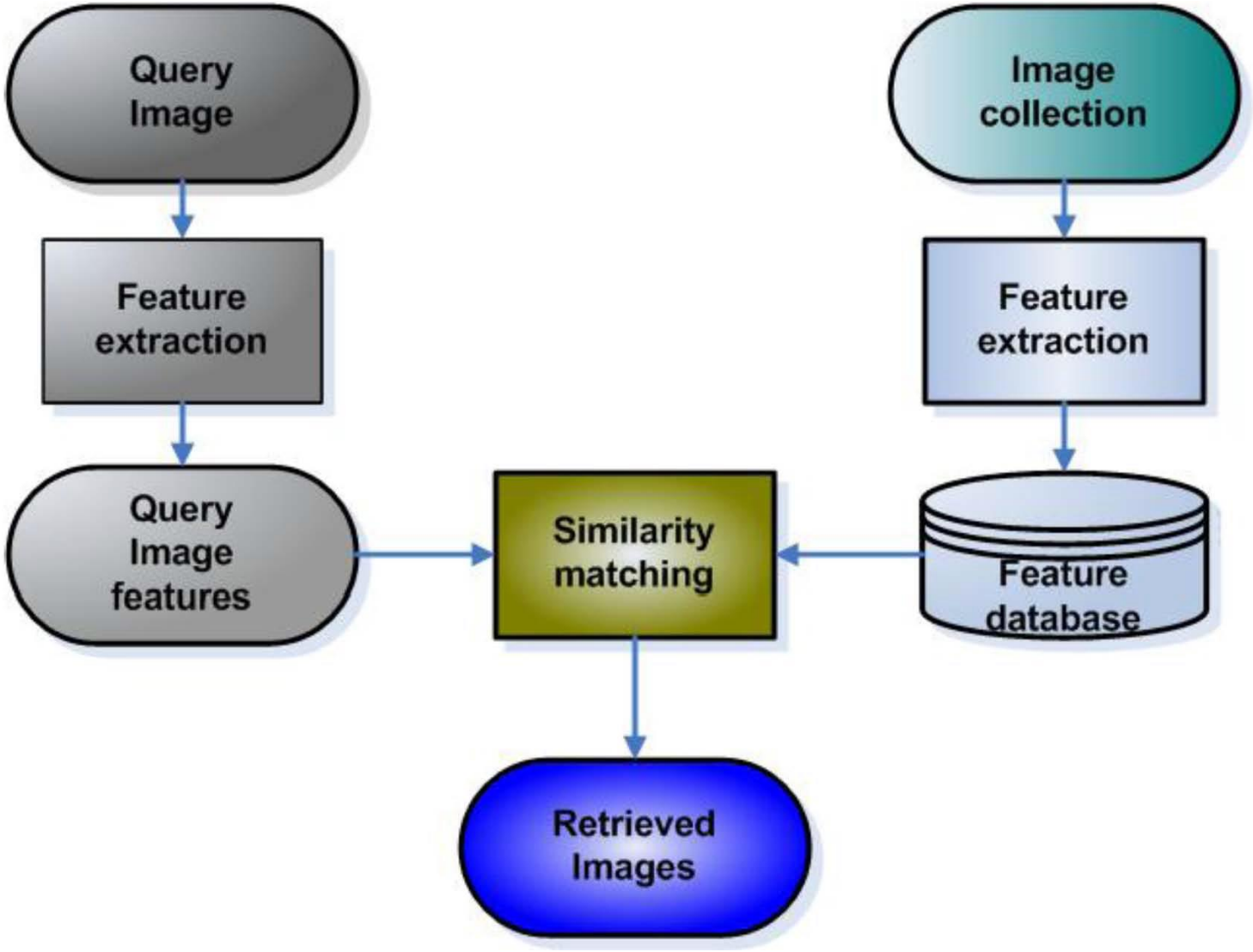
Multimedia retrieval approaches

CBIR



- **Recall** = #pertinent responses/#responses
- **Precision** = pertinent responses/#pertinent data

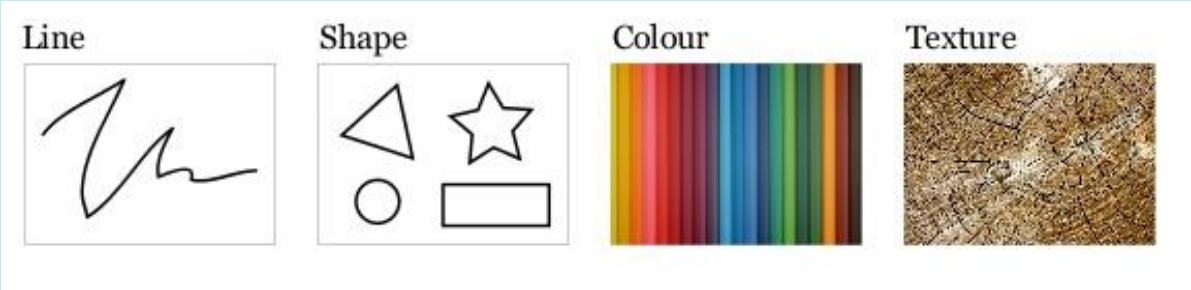
Content Based Image Retrieval (CBIR)



CBIR for Industry

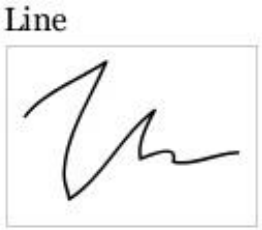
Multimedia retrieval approaches

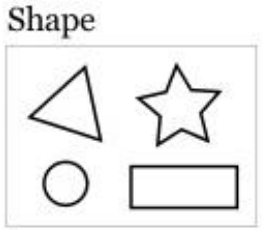
- Syntactic image features
  - Color
  - Texture
  - Shape
  - Points of interest



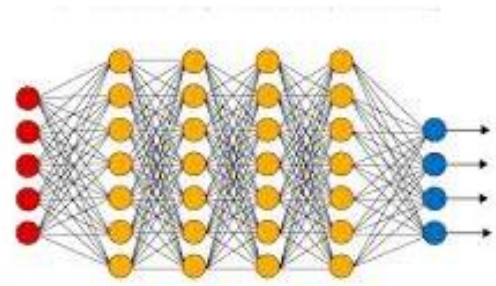
- Semantic image features
  - Objects(human, faces, animals, cars, building, etc.)
  - Topics (fire, pollution, etc.)

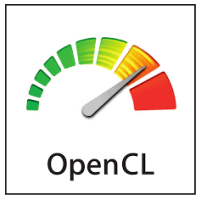
Software requirements



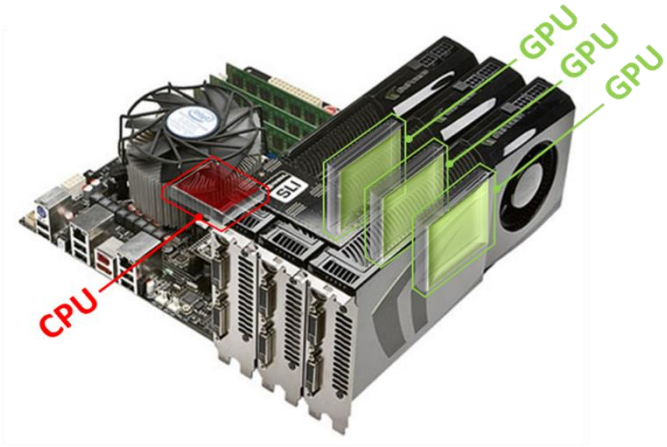


- Syntactic and semantic features extraction
- Similarity measurements





Hardware requirements



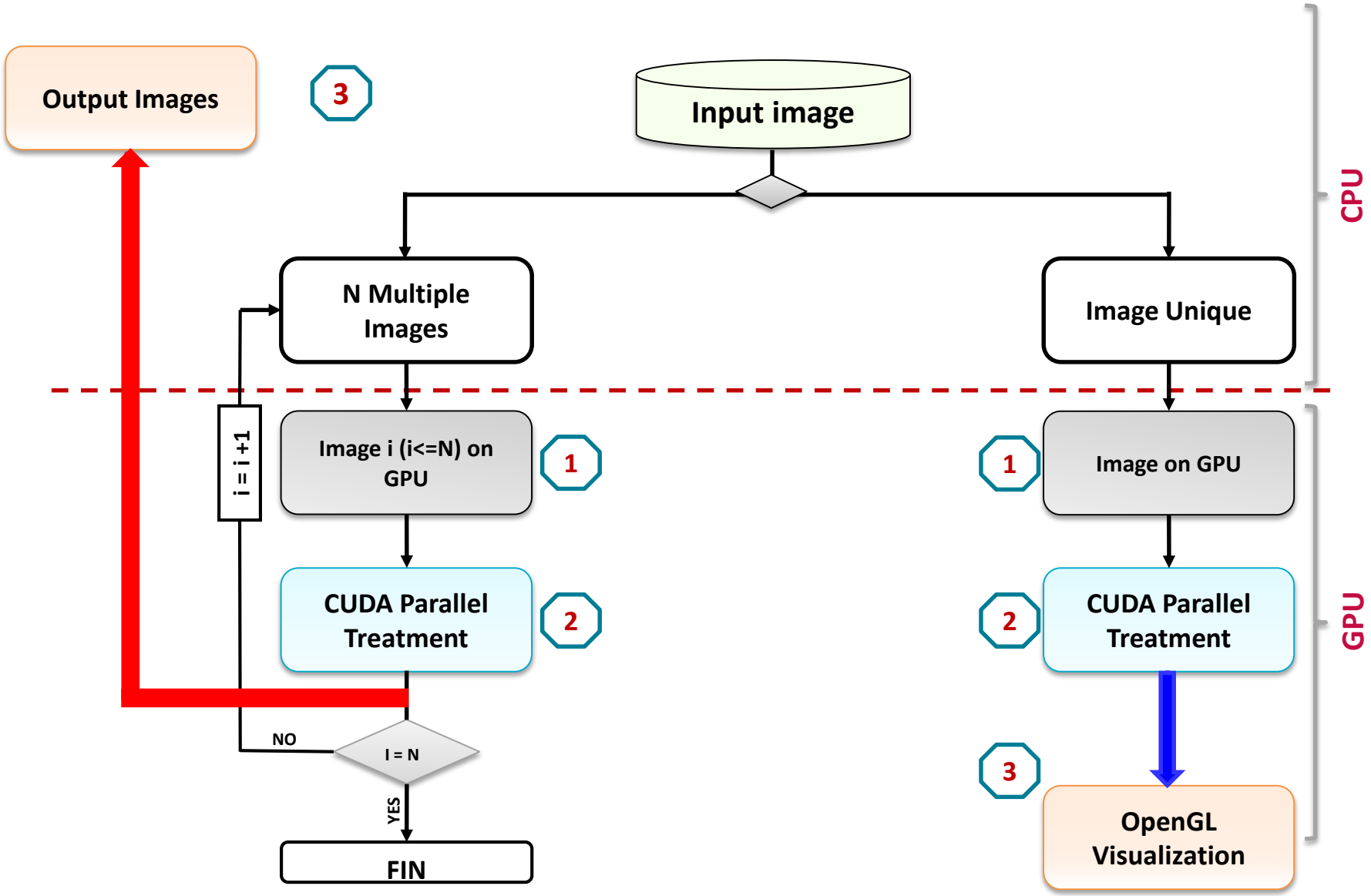
Performance analysis

- The process of multimedia indexation presents a computation that can increase significantly when enlarging the size of databases and the resolution of images (HD, Full HD, 4K, 8K, etc.)

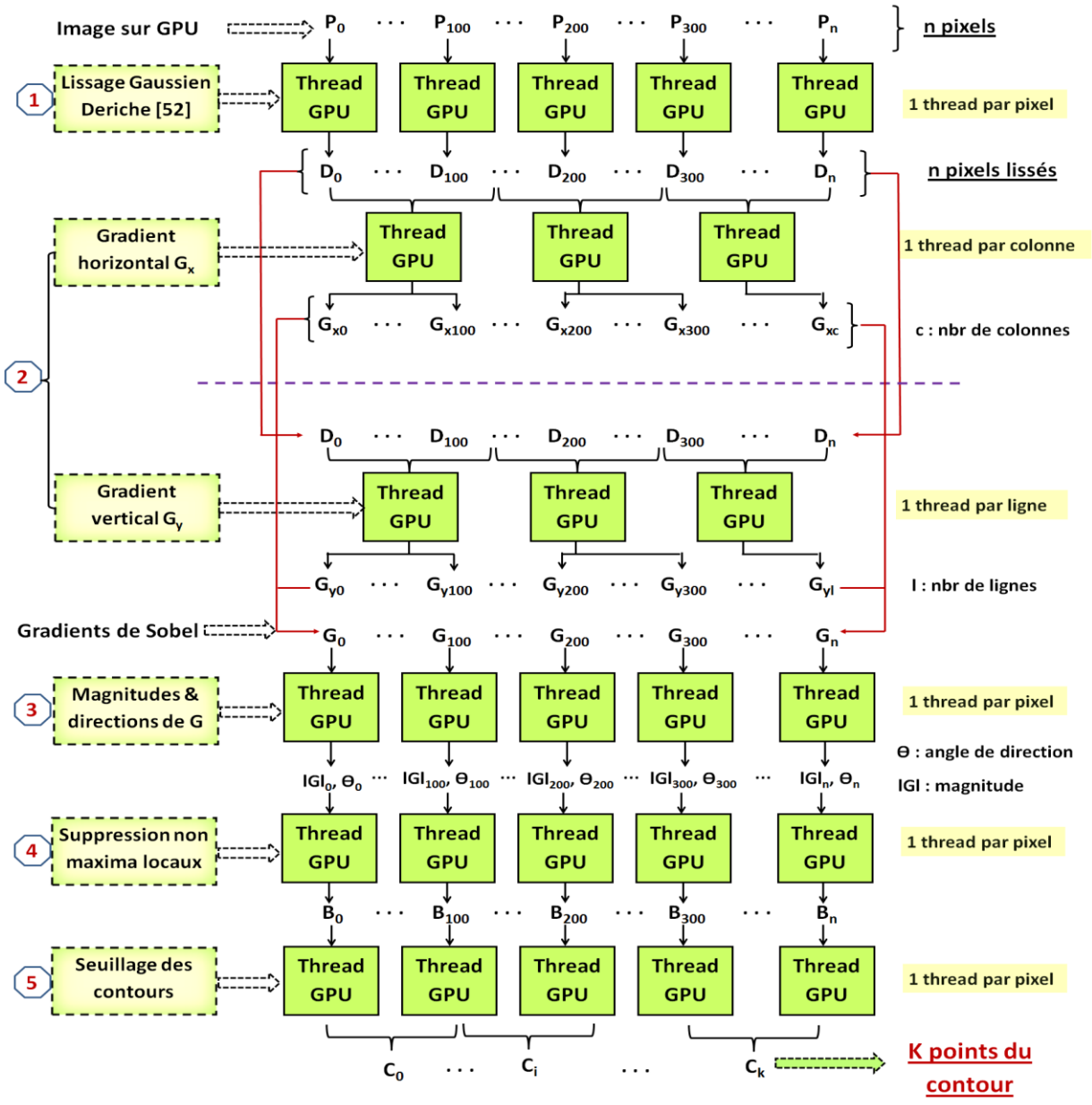
- The process of research is less consuming in times but requires a fast treatment

- To get a good rate of precision, we need to use efficient and complex algorithms and descriptors that require more computation times.

GPU based features extraction



GPU based features extraction



Performance analysis : computation time

Algorithm	Resolution	f	$comp_{pix}$	f_c	$f_c > S$?	Unit	Acc
Gaussian	256×256	0,85	9	$5.0 * 10^5$	No	CPU	00,93 ↘
filtering	512×512	0,88	9	$2.1 * 10^6$	Yes	GPU	11,34 ↗
	1024×1024	0,95	9	$9.0 * 10^6$	Yes	GPU	50,30 ↗
	2048×2048	0,98	9	$3.7 * 10^7$	Yes	GPU	118,9 ↗
Blur	256×256	0,59	9	$3.5 * 10^5$	No	CPU	00,72 ↘
filtering	512×512	0,75	9	$1.8 * 10^6$	Yes	GPU	14,98 ↗
	1024×1024	0,89	9	$8.4 * 10^6$	Yes	GPU	38,99 ↗
	2048×2048	0,94	9	$3.5 * 10^7$	Yes	GPU	84,36 ↗
Median	256×256	0,64	9	$3.8 * 10^5$	No	CPU	00,86 ↘
filtering	512×512	0,75	9	$1.8 * 10^6$	Yes	GPU	49,66 ↗
	1024×1024	0,88	9	$8.3 * 10^6$	Yes	GPU	128,3 ↗
	2048×2048	0,97	9	$3.7 * 10^7$	Yes	GPU	273,1 ↗

Performance analysis: computation time

Algorithm	Resolution	f	$comp_{pix}$	f_c	$f_c > S$?	Unit	Acc
Edge and	256×256	0.55	6	$2.2 * 10^5$	No	CPU	00.87 ↘
Corners	512×512	0.81	6	$1.3 * 10^6$	Yes	GPU	05.88 ↗
detection	1024×1024	0.86	6	$5.5 * 10^6$	Yes	GPU	12.01 ↗
	2048×2048	0.89	6	$2.3 * 10^7$	Yes	GPU	18.53 ↗
Sift	256×256	0,81	>180	$9.5 * 10^6$	Yes	GPU	08,69 ↗
descriptor	512×512	0,85	>180	$4.0 * 10^7$	Yes	GPU	21,99 ↗
	1024×1024	0,88	>180	$1.7 * 10^8$	Yes	GPU	70,77 ↗
	2048×2048	0,93	>180	$7.0 * 10^8$	Yes	GPU	210,6 ↗
Surf	256×256	0,75	<100	$4.9 * 10^6$	Yes	GPU	02,82 ↗
descriptor	512×512	0,79	<100	$2.1 * 10^7$	Yes	GPU	25,18 ↗
	1024×1024	0,84	<100	$8.8 * 10^7$	Yes	GPU	45,39 ↗
	2048×2048	0,89	<100	$3.7 * 10^8$	Yes	GPU	98,08 ↗

Performance analysis : energy consumption

Application	CPU		GPU		Heterogeneous	
	Power (W)	Energy (Wh)	Power	Energy	Power	Energy
Image analysis algorithms						
Color histogram	271,25	0,90	298,12	0,45	…	…
HSV histogram	234,8	0,50	280,77	0,12	…	…
DFT	218	0,58	239	0,22	…	…
Corners detection	216,8	0,60	310,2	0,49	…	…
Contours detection	268,2	4,27	290	1,29	…	…
Image denoising algorithms						
Linear filtering	278,8	0,77	294,6	0,061	…	…
Bilateral filtering	304,6	12,48	390,8	1,32	…	…
Median filtering	241,6	3,88	296,13	0,32	…	…
Gaussian filtering	236,2	1,31	292,33	0,09	…	…
Keypoints detection algorithms						
Sift descriptor	232,8	6,97	305,2	2,98	…	…
Surf descriptor	280,8	1,15	311,6	0,86	…	…

Performance analysis : energy consumption

Application	CPU		GPU		Heterogeneous	
	Power (W)	Energy (Wh)	Power	Energy	Power	Energy
Optical flow computation algorithms						
Lukas canade	223,4	0,60	302,2	0,43	…	…
Farnback	243,6	2,54	282,6	0,76		
tvl1	269,8	1,22	312,2	0,85	…	…
Block matching	220,8	0,53	277,4	0,48	…	…
Use case applications						
Vertebra segmentation	214	294	…	…	232	12
Videos indexation	207	292	…	…	227	10.2
Event detection	155	35.1	248	3.8	…	…
Event localization	152	22.4	246	1.3	…	…
Ventricle tracking	198	51.6	299	4.2	…	…

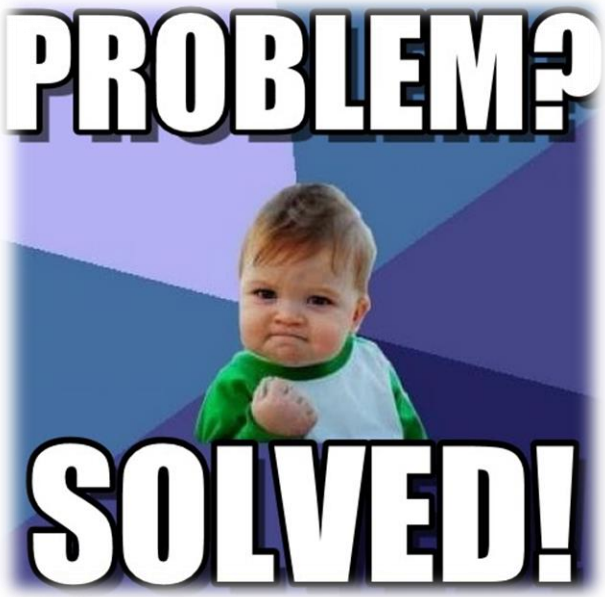
The problem

- Syntactic and semantic features extraction require generally the use of several libraries that should be downloaded, installed and configured
- High intensive applications require a high computing power
- In some situations, it is also necessary to configure the hardware: gpu, multi-cpus, etc.
- In several situations, users are not ready to do these tasks.

- A long process which is not so appreciated by users and even developers

The challenge

- Develop a cloud application for which guests will have access to an efficient method of multimedia retrieval without having to download, install and configure the corresponding software.



Main cloud platforms

- Amazon Web Services (AWS): www.aws.amazon.com

- Google Cloud Platform (GCP): **https://cloud.google.com/**

- Microsoft Azure : www.azure.microsoft.com

Computer vision cloud platforms
CloudCV [CLoudCV]

- CloudCV : Open source cloud Platform for computer vision
- Cloud Platform launched by students of Virginia Lab
- Build, compare and share machine learning algorithms
- Provide and convert deep learning models to web services
- CloudCV provides access to two API (Python and Matlab), and englobes multiple modern components for its backend architecture such as OpenCV, Caffe, Turi (GraphLab).

Computer vision cloud platforms
IPOL [IPOL]

- Image processing On Line (IPOL) proposed a platform for :
  - Basic image processing algorithms
  - Image descriptors
  - Features extraction

- The platform is present within a web interface that allows to upload input values before checking

- More focused on image processing algorithms.

Our cloud based Multimedia Retrieval

Our cloud based multimedia retrieval approach is described within four parts :

- Platform architecture
- Webserver and clients management
- Data encryption
- Applications management

Our cloud based Multimedia Retrieval

Our cloud based multimedia retrieval approach is described within four parts :

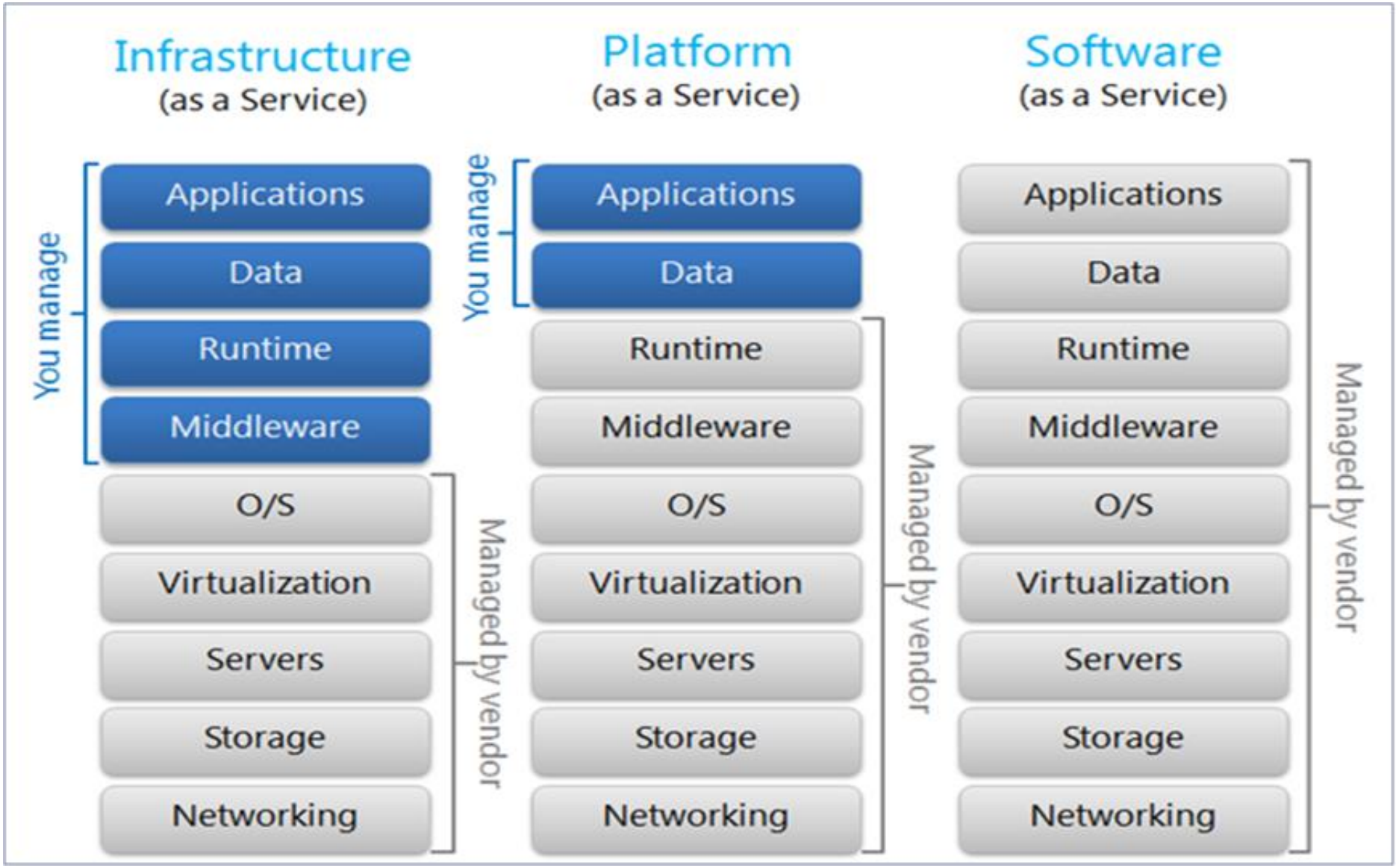
- **Platform architecture**
- Webserver and clients management
- Data encryption
- Applications management

Our cloud based Multimedia Retrieval

I. Platform architecture

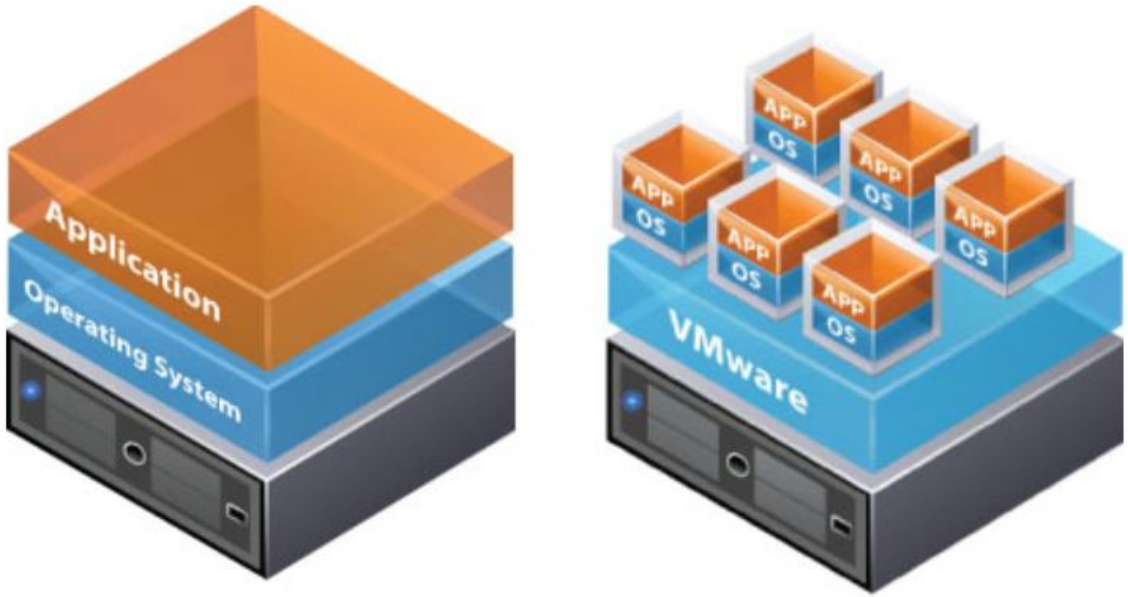
- We use an IaaS architecture in order to provide an SaaS service
- Three virtual machines (VMs) are used in our cloud platform
  - **VM N°1** : server that collect users requests and launch applications
  - **VM N°2** : encrypt data is the user requires more security
  - **VM N° 3:** host and execute our applications
- SFTP protocol is used to communicate between VMs
- SSH protocol is used de run and execute applications

Our cloud based Multimedia Retrieval



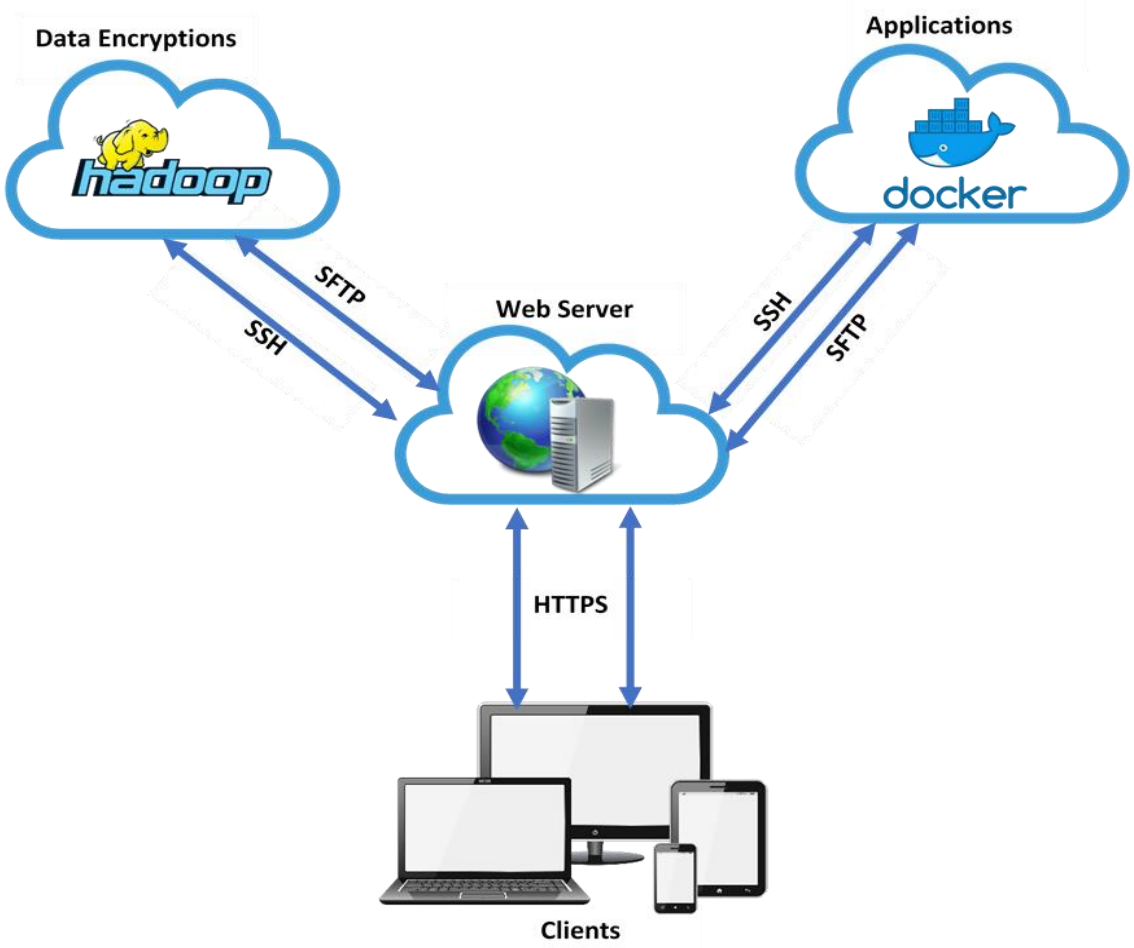
Our cloud based Multimedia Retrieval

The access to the virtual machine is facilitated by a software known as Hypervisor (or monitor). Hypervisor acts as a link between the hardware and the virtual environment and distributes the hardware resources such as CPU usage, memory allotment between the different virtual



Our cloud based Multimedia Retrieval

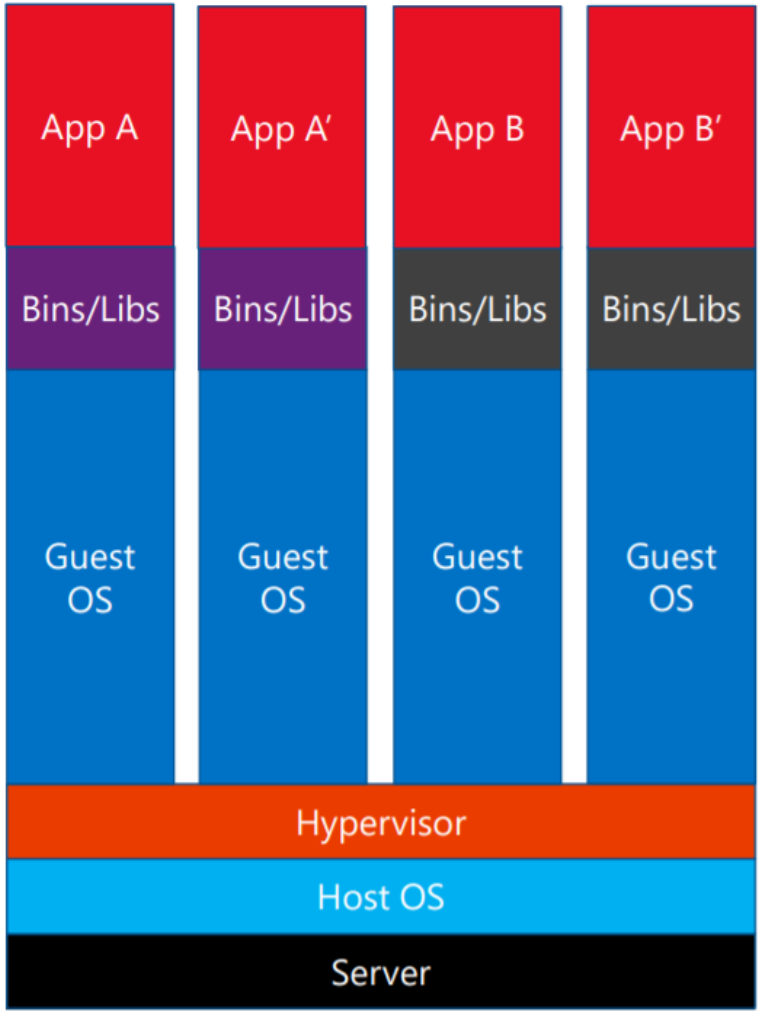
I. Platform architecture

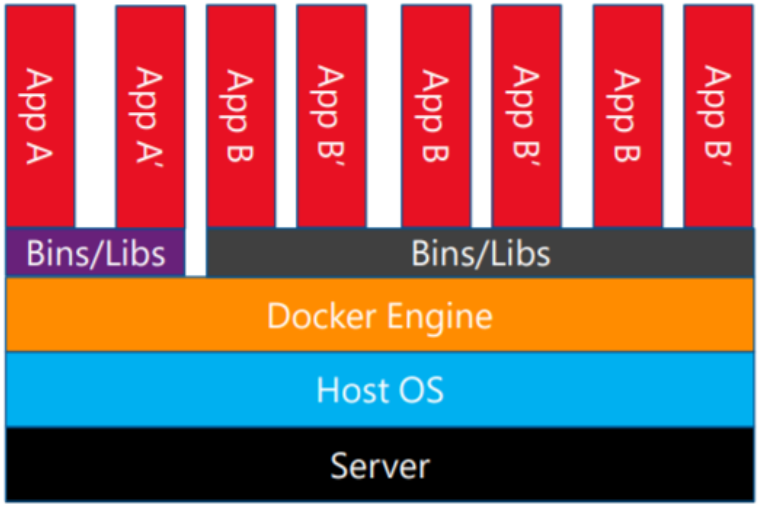


The general architecture of our platform

Our cloud based Multimedia Retrieval

What is Docker ?





Our cloud based Multimedia Retrieval

Docker benefits

- Custom containers
- Less resource requirements
- Multiple instances of the same application
- Portability

Our cloud based Multimedia Retrieval

Our cloud based multimedia retrieval approach is described within four parts :

- Platform architecture
- **Webserver and clients management**
- Data encryption
- Applications management

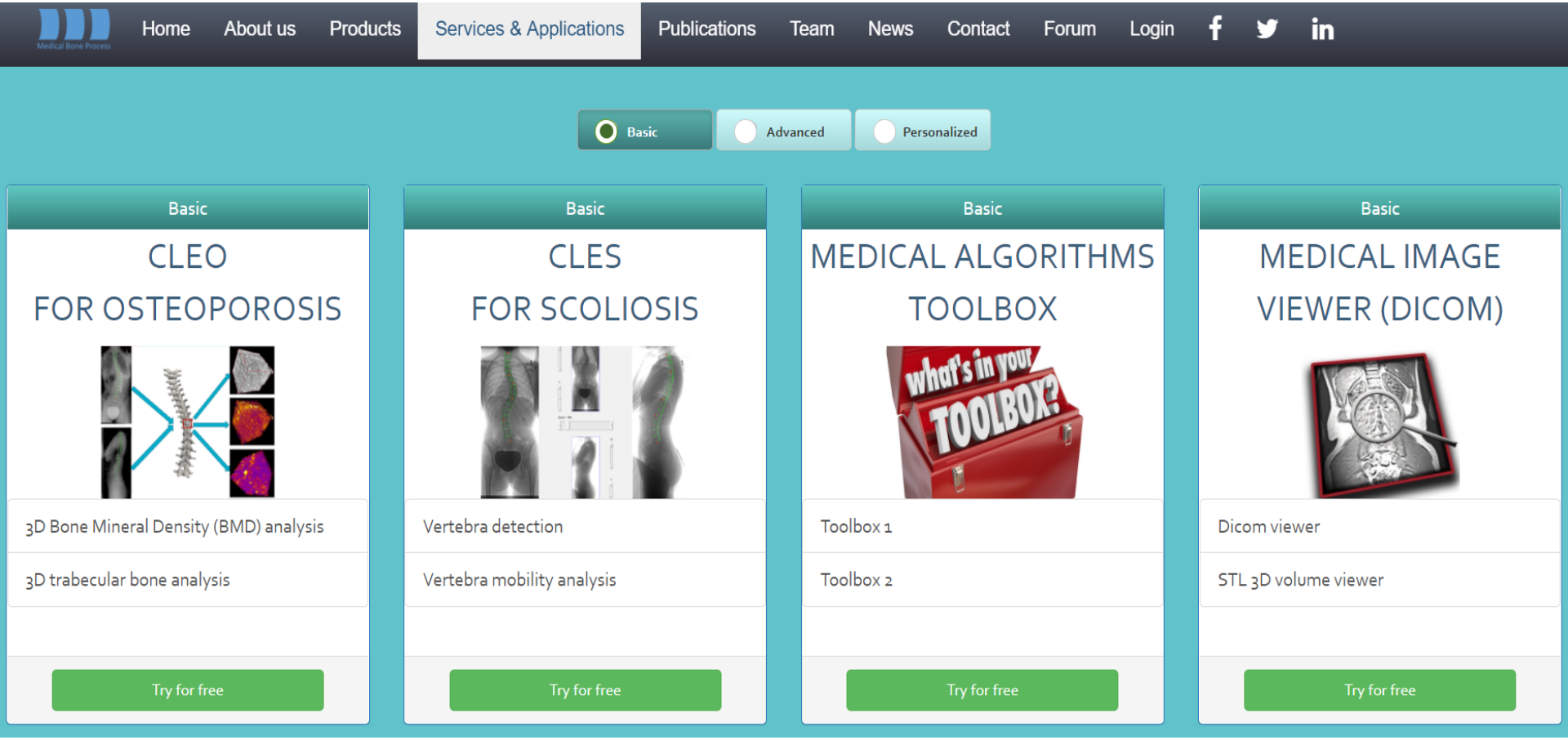
Our cloud based Multimedia Retrieval

II. Web server and clients management

- Website to provide a responsive web application for users
- The website is developed with the Bootstrap framework : multi-platform website that can be run even on mobile devices (smartphone, tablet, etc.)
- The framework Symfony is used to offer more security to our website
- The protocol HTTPS is based on certificate of Let's encrypt, which represents a free, automated, and open certificate authority brought by the non-profit Internet Security Research Group (ISRG).

Our cloud based Multimedia Retrieval

Website

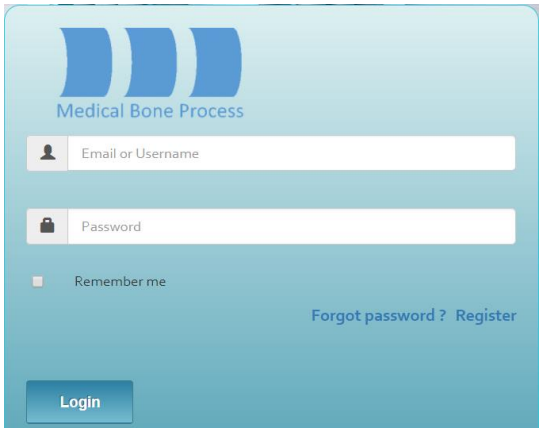


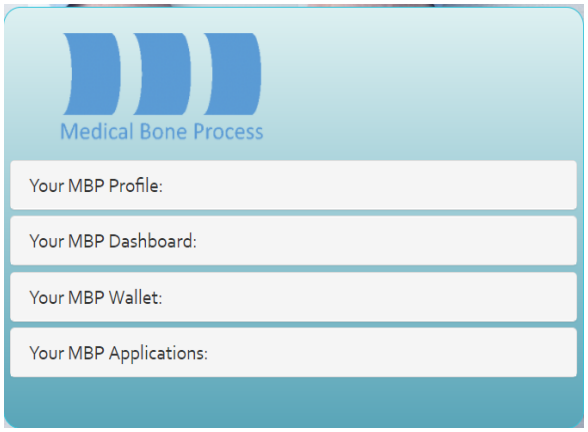
https://bone.media-process.com

Our cloud based Multimedia Retrieval

User session

- **MBP profile :** summarizes the personal data of users
- **MBP Dashboard :** list of saved results by users of the server. Each result can be either consulted or deleted
- **MBP wallet :** user credits that allows to users to launch our applications. Actually, our applications are for free
- **MBP Applications :** List of applications integrated within the medical and video processing platforms.





Our cloud based Multimedia Retrieval

Our cloud based multimedia retrieval approach is described within four parts :

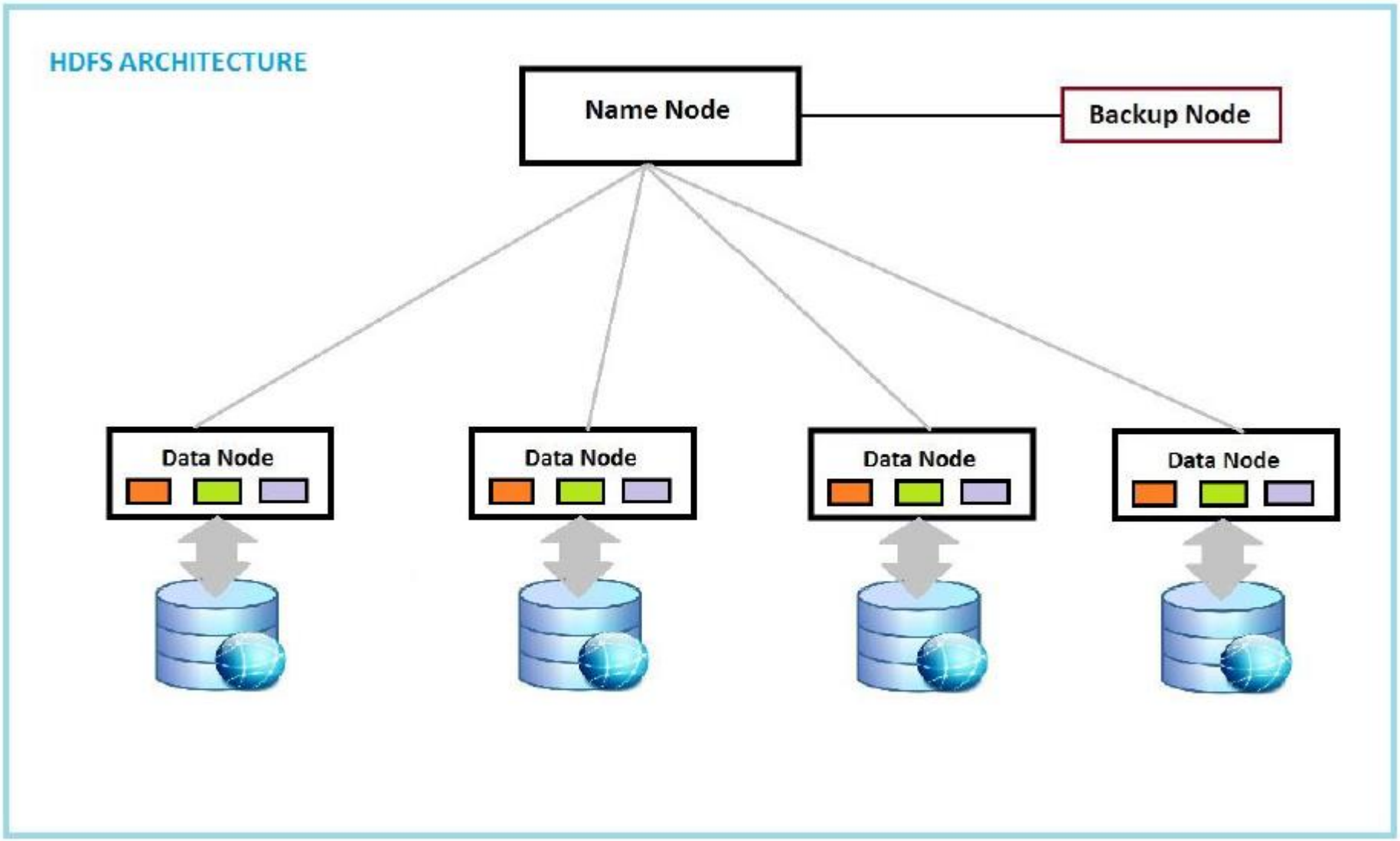
- Platform architecture
- Webserver and clients management
- **Data encryption**
- Applications management

Our cloud based Multimedia Retrieval

III. Data encryption

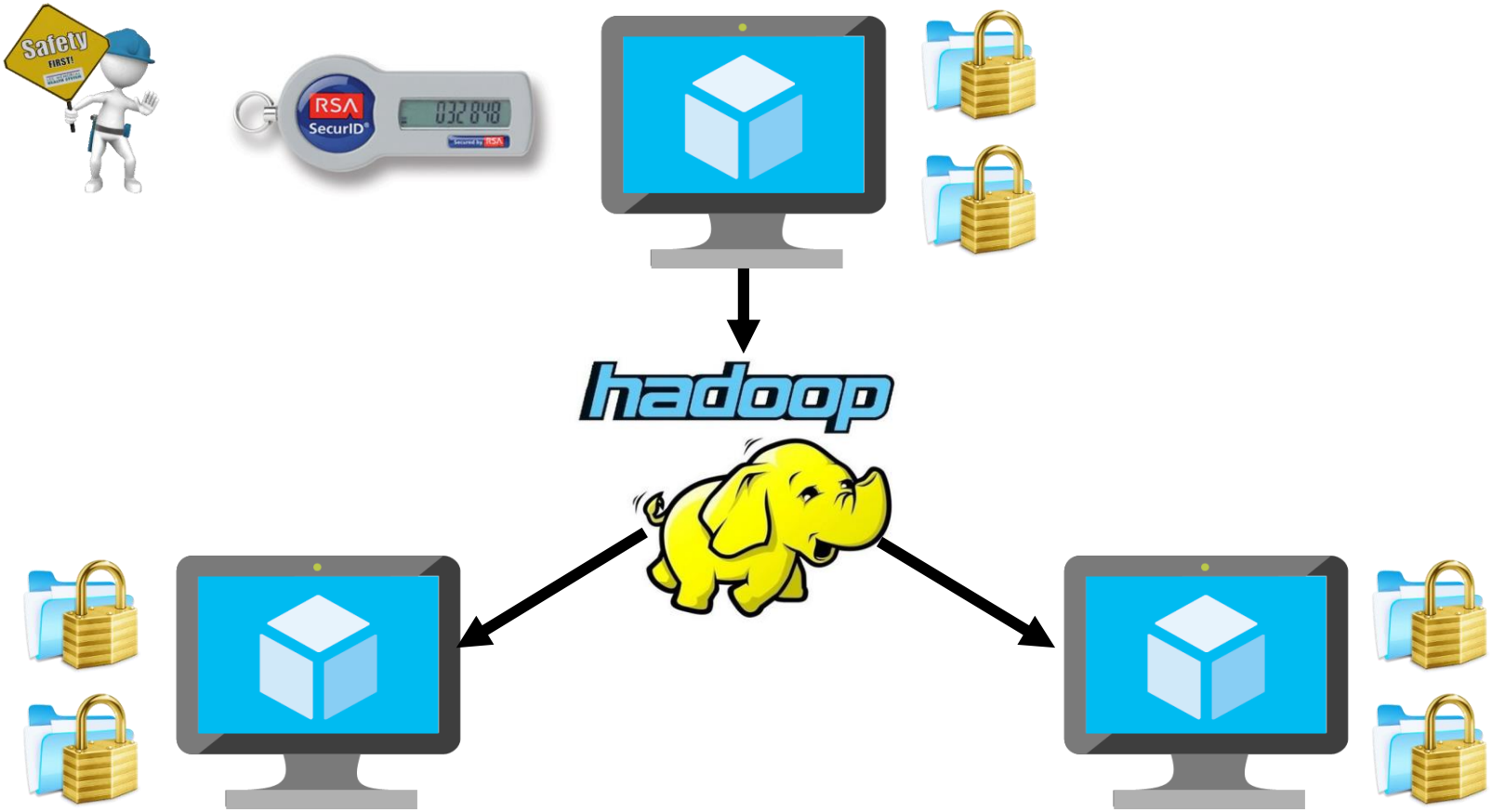
- Securing data is so important for many users
- Data are encrypted with the protocol RSA
- RSA is a cryptosystem for public-key encryption, and is widely used for securing sensitive data, particularly when being sent over an insecure network such as the Internet.
- The framework Hadoop is used to offer a replication of data with the HDFS system. This allows to recover data in case failure.

HDFS (Hadoop Distributed File System)



Our cloud based Multimedia Retrieval

Hadoop?



Our cloud based Multimedia Retrieval

Our cloud based multimedia retrieval approach is described within four parts :

- Platform architecture
- Webserver and clients management
- Data encryption
- **Applications management**

Our cloud based Multimedia Retrieval

IV. Applications management

- The proposed application exploits both CPU and GPU for computations
- The CPU is used for less intensive steps
- The GPU is used for high intensive steps.
- The GPU implementations are using the CUDA API
- The OpenCV library is used for features extraction on CPU

Our cloud based Multimedia Retrieval

Two docker images

1. **Basic-docker-image** : integrates the OpenCV library for images features detection and similarity measurements

2. **Nvidia-docker-image:** integrate the GPU module of OpenCV library and the ncvv compiler that allow to run CUDA programs.

Our cloud based Multimedia Retrieval

The process

1. The user select the application within the website and upload its data
2. The server (VM N°1) receives the request and select the VM N°3 which host the application
3. If users would like to use CPU, the basic-docker-image is selected. Otherwiwe, the nvidia-docker-image is selected
4. The docker container launch the application of multimedia retrieval

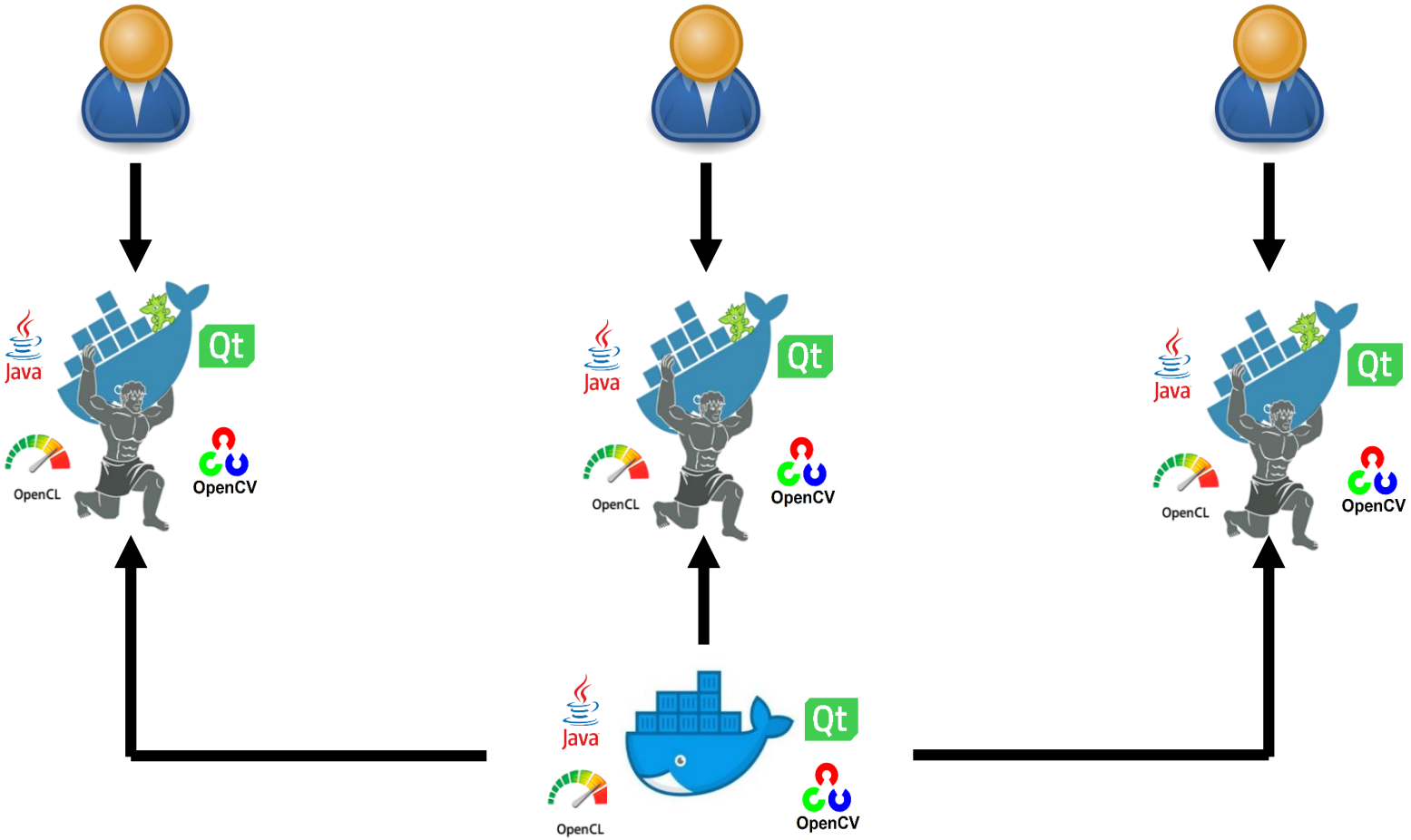
Our cloud based Multimedia Retrieval

The process

5. The results are stored in folder shared between the VM and container
6. The created container is deleted
7. The results are sent to the client within SFTP protocol
8. In case of simultaneous executions of the application, the related VM can create for each user a container, which ensures the simultaneous multi-user access in real time.

Our cloud based Multimedia Retrieval

Why Docker ?



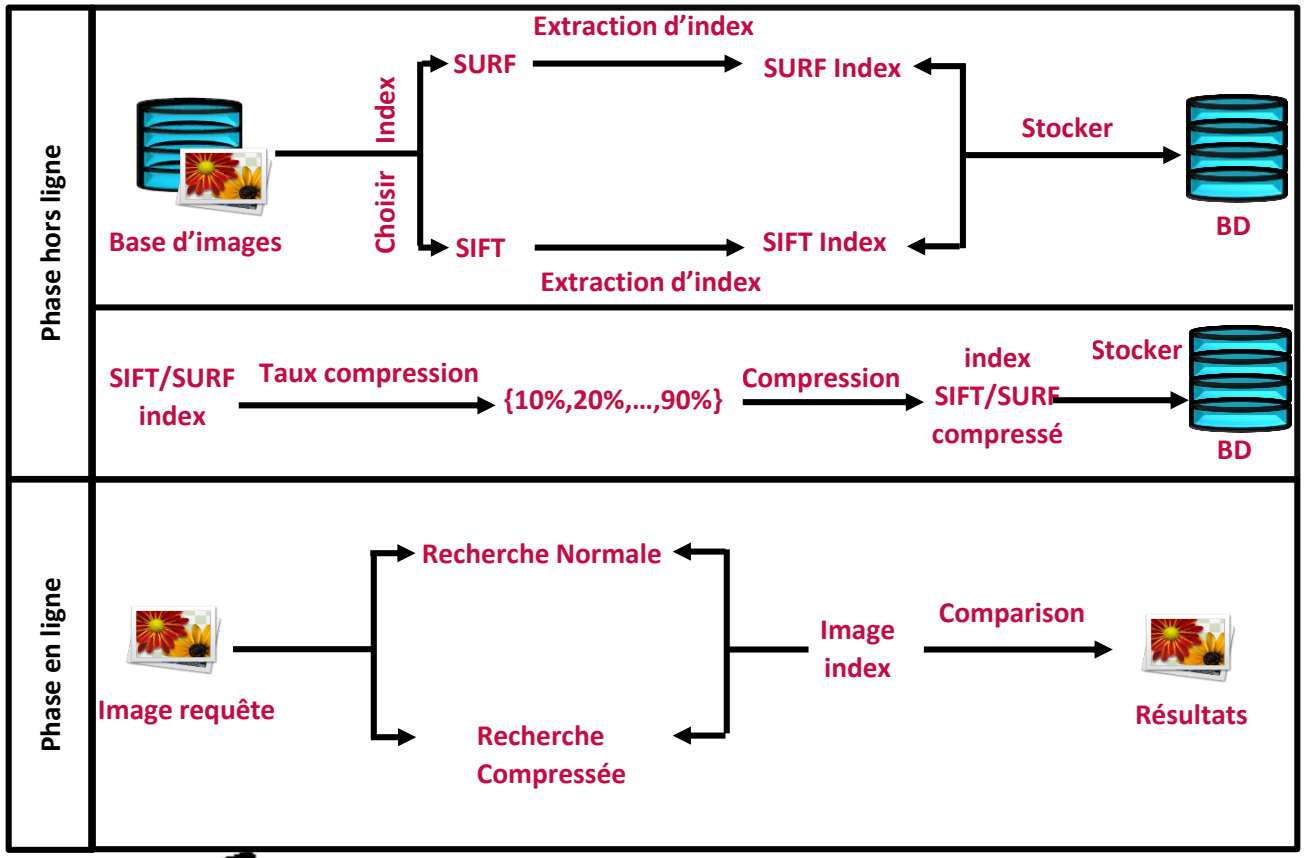
Experimentations

Two main uses cases :

- CPU and GPU based Images indexation and retrieval ([CloudTech])
- CPU and GPU based object detection in real time ([Yolo])

Experimentations

CPU and GPU based multimedia retrieval [CloudTech17]



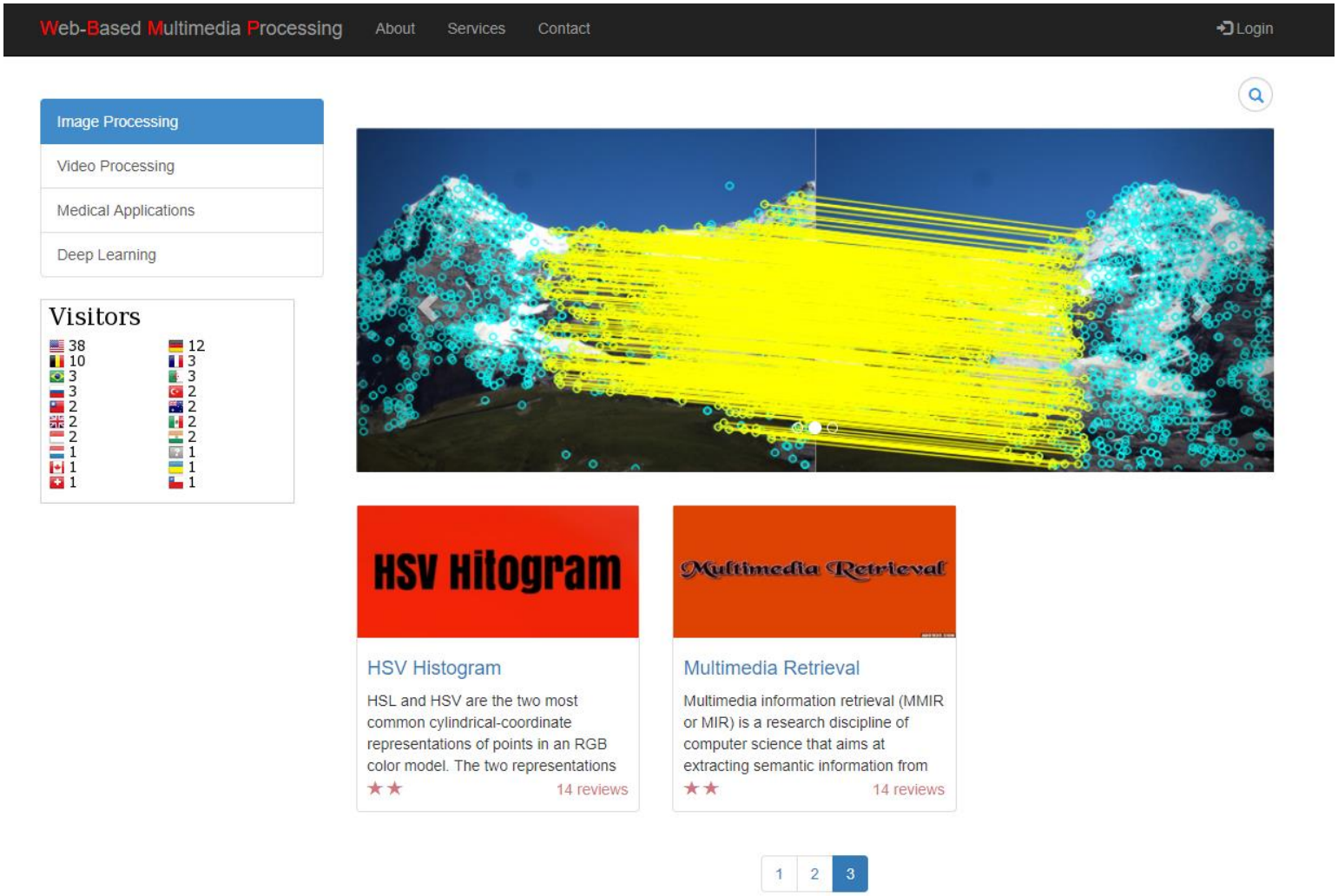
Experimentations

CPU and GPU based multimedia retrieval

- The result of offline phase (indexation) is stored in the VM N° 3
- Each user has to upload its request image within the website
- The connection is required for using applications
- The results can be either displayed or downloaded on users machines

Experimentations

CPU and GPU based multimedia retrieval



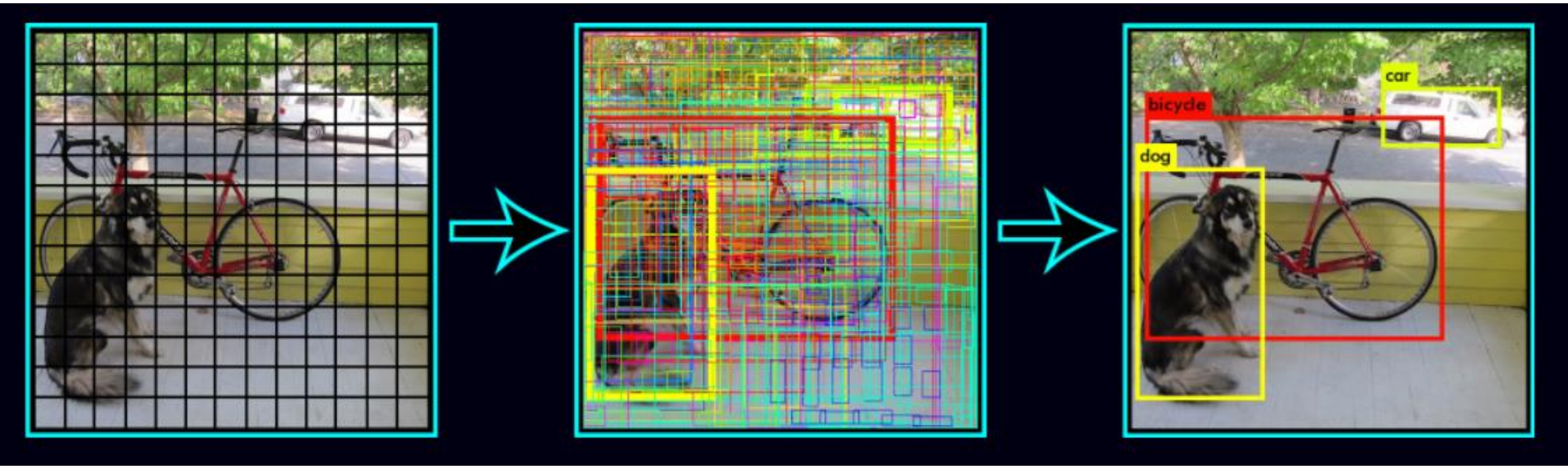
Experimentations

CPU and GPU based object recognition in real time

- The approach of YOLO consists of using a single neural network to the full image
- This network divides the image into regions and predicts bounding boxes and probabilities for each region
- These bounding boxes are weighted by the predicted probabilities.
- As each deep learning method, this method apply a training step before launching the recognition

Experimentations

CPU and GPU based object recognition in real time



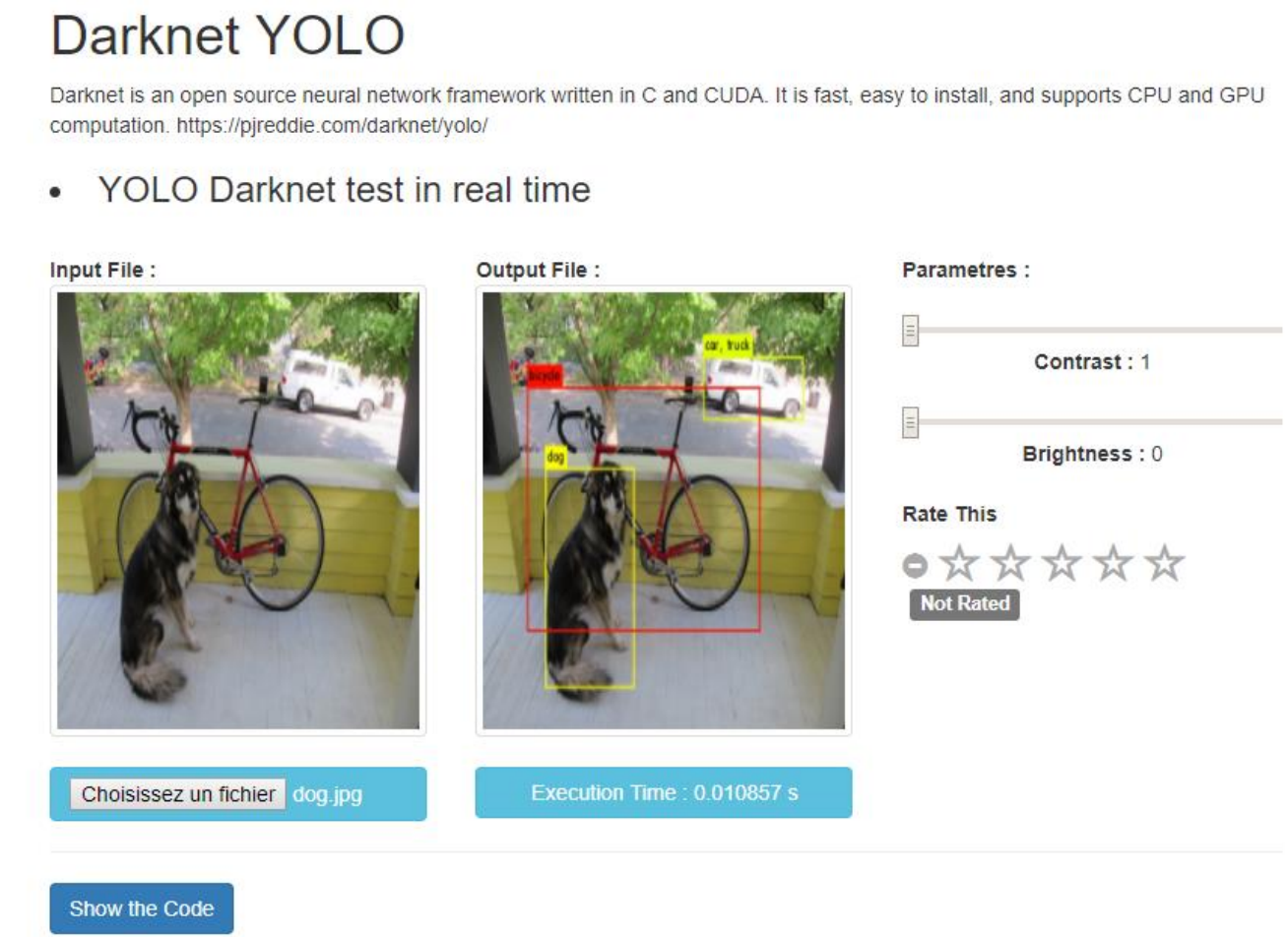
- The training step can exploit CPU, GPU or multi-GPU platforms.
- The real time recognition can exploit GPUs also.

Experimentations

CPU and GPU based object recognition in real time

Experimentations

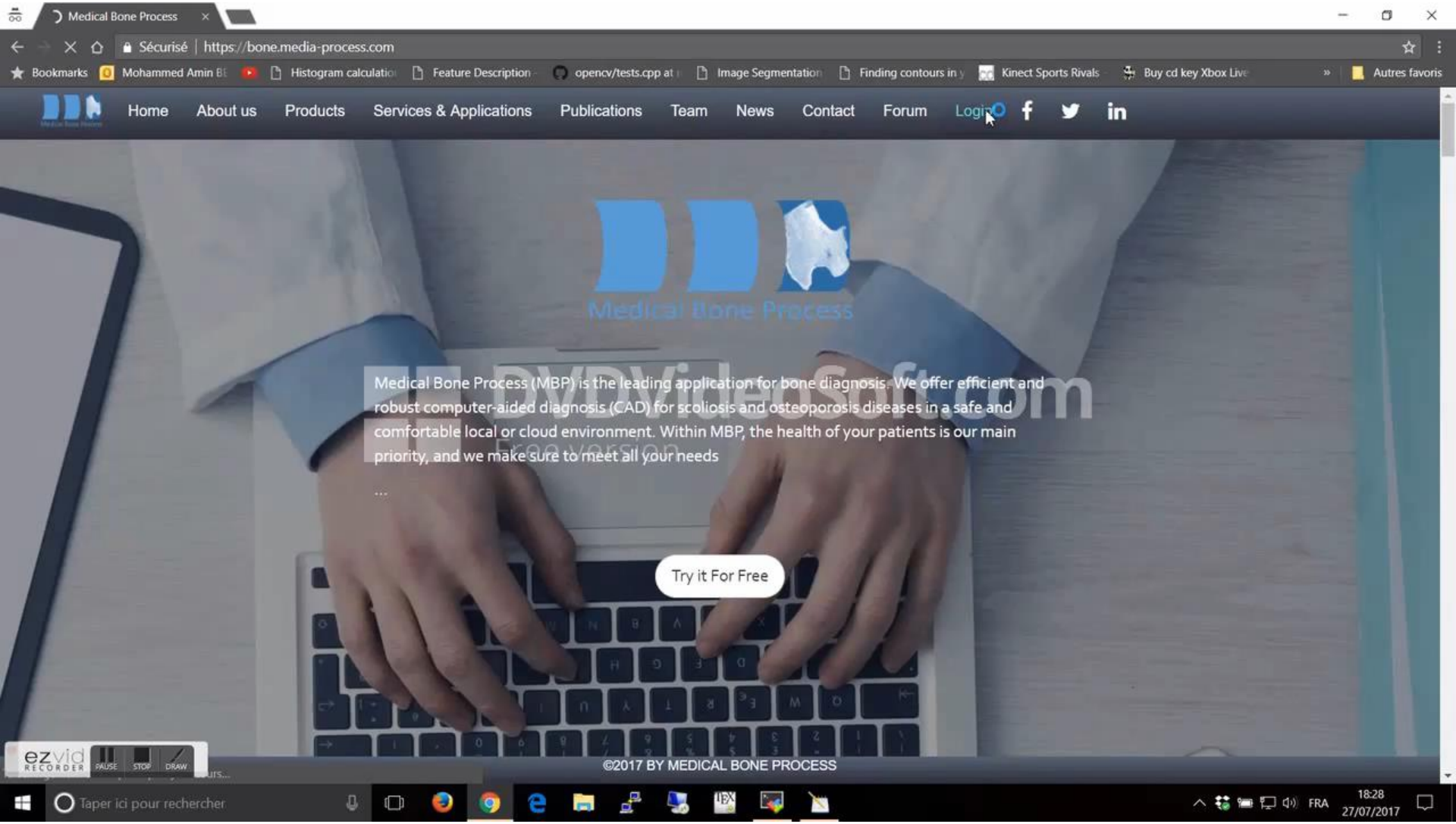
CPU and GPU based object recognition in real time



Demonstrations

Let us try

Demonstrations



Conclusion

- Multimedia indexation and retrieval present a very important tool for various applications.
- High intensive applications that require a high computing power
- So complicated to configure when exploiting heterogeneous platforms
- The use of cloud platforms offers a convenient solution for users and developers

Future works

- Integration of deep learning algorithms of real time object detection
- Integration of the training phase computation in the cloud platform
- Share and provide the source code with collaborators
- Real time object recognition in the cloud using video sequences

References

[1] Agrawal, H. M, "**Cloudcv: Large-scale distributed computer vision as a cloud service**", *In Mobile cloud visual media computing*, pp. 265-290, 2015.

[2] Limare, N. a.-M, "**The IPOL initiative: Publishing and testing algorithms on line for reproducible research in image processing**", *Procedia Computer Science* , pp. 4:716-725, 2011.

[3] Joseph Redmon and Ali Farhadi, " **YOLO9000 : Better, Faster, Stronger** ", *Computer Vision and Pattern Recognition CVPR Conference,* 2017.

[4] Mahmoudi Sidi Ahmed, Belarbi Mohammed Amin, Mahmoudi Said, Belalem Ghalem, "**Towards a Smart Selection of Resources in the Cloud for Low-energy Multimedia Processing**" in Concurrency & Computation : Practice & Experience (2017)

[5] Belarbi Mohammed Amin, Mahmoudi Said, Belalem Ghalem, Mahmoudi Sidi, "**Web-based Multimedia Research and Indexation for Big Data Databases**« , *in CloudTech 2017 : The 3rd International Conference on Cloud Computing Technologies and Applications*, Morocco (2017)

[6] Mahmoudi Sidi, Manneback Pierre, "**Multi-CPU/Multi-GPU Based Framework for Multimedia Processing**« , *in IFIP Advances in Information and Communication Technology. Computer Science and Its Applications,* 456, 2015, 54-65 (2015)

References

[7] Mahmoudi Sidi, Ozkan Erencan, Manneback Pierre, Tosun Souleyman, "**Taking Advantage of Heterogeneous Platforms in Image and Video Processing**" *in Complex HPC book" , Wiley, 978-1-118-71205-4* (2014)

[8] Da Cunha Possa Paulo, Mahmoudi Sidi, Harb Naim, Valderrama Carlos, "**A New Self-Adapting Architecture for Feature Detection**" in *Lecture Notes in Computer Science, 978-1-4673-2257-7, 2012(22) Oslo, 643 - 646, 10.1109/FPL.2012.6339149* (2012)

[9] Mahmoudi Sidi, Manneback Pierre, Augonnet C., Thibault S., "**Traitements d'images sur architectures parallèles et hétérogènes**" *in Technique et Science Informatiques, 31/8-10 - 2012, 8-9-10/2012, 1183-1203, 10.3166/tsi.31.1183-1203* (2012)

[10] Mahmoudi Sidi Ahmed, Manneback P., Augonnet C., Thibault S., "**Détection optimale des coins et contours dans des bases d'images volumineuses sur architectures multicoeurs hétérogènes***", in "20èmes Rencontres Francophones de l'Informatique Parallèle , France* (2011)

[11] Mahmoudi Sidi Ahmed, Manneback Pierre, « **Multi-GPU based Event Detection and Localization using High Definition Videos** » *", in "The 4th International Conference on Multimedia Computing and Systems (ICMCS'14)* , Marrakesch, Morocco (2014)

THANK YOU