

Le cloud IFB et son instance Galaxy



Christophe BLANCHET

Institut Français de Bioinformatique - IFB
French Institute of Bioinformatics - ELIXIR-FR
CNRS UMS3601 - Gif-sur-Yvette - FRANCE

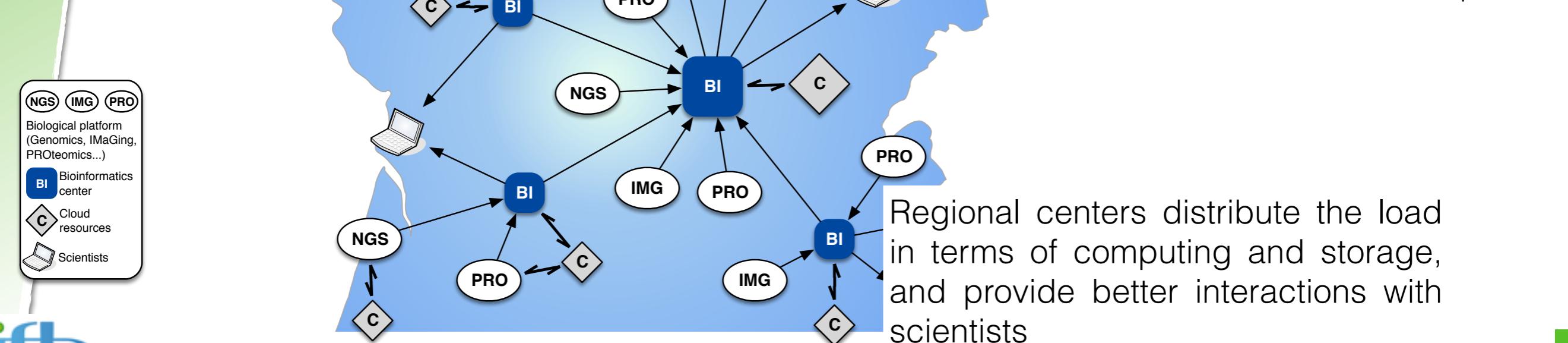
Ecole Bioinformatique Aviesan
28 Septembre 2015, Roscoff

Experimental data in life sciences (FR)

French national platforms (GIS IBISA)	Nb
Cellular imaging	19
Genomic, Transcriptomic	16
Proteomic	13
Structural biology, biophysic	11



Source: omicsmaps.com



A lot of bioinformatics tools



ABYSS 1.3.4

ARIA 2.3

Bioconductor 2.11

biomaj

BLAST+ 2.2.27

Blat 35

Bowtie 0.12.8

Bowtie2 2.0.0-

beta7

BWA 0.6.2

BWA 0.7.10

CAP3

CD-HIT 4.6.1

Clustal Omega
1.0.3

CLUSTALW 2.1

Cufflinks 2.0.2

Cutadapt 1.2.1

E-SURGE 1.9.0

Exonerate 2.2.0

eXpress 1.5.1

FastA 3.6

FastQC 0.10.1

Galaxy portal

GATK 2.3.4

HMMer 3.0

ImageJ 1.48

khmer 1.1

M-SURGE 1.8.5

MEME 4.7

MMSEQ 0.11.2a

Mobyle

MODAL

MultAlin 5.4.1

MUSCLE 3.8.3I

neo4j

Oases 0.2.08

OMSSA 2.1.9

PeptideShaker
0.18.3

phyml 3.1

PREDATOR 2.1.2

proline

python 2.7

R 2.13

R 3.1.1

R 3.1.2

R-studio

Ray 1.3

RSAT

samtools 0.1.18

Samtools 1.1

SearchGUI 1.10.4

SeqClean

Shiny

Stacks

STAR 2.4.0fI

SuMo vI

TGICL

TopHat 2.0.6

trim_galore 0.3.7

Trinity 2.0.4

U-CARE 2.3.2

VCFtools 0.1.11

Velvet 1.2.10

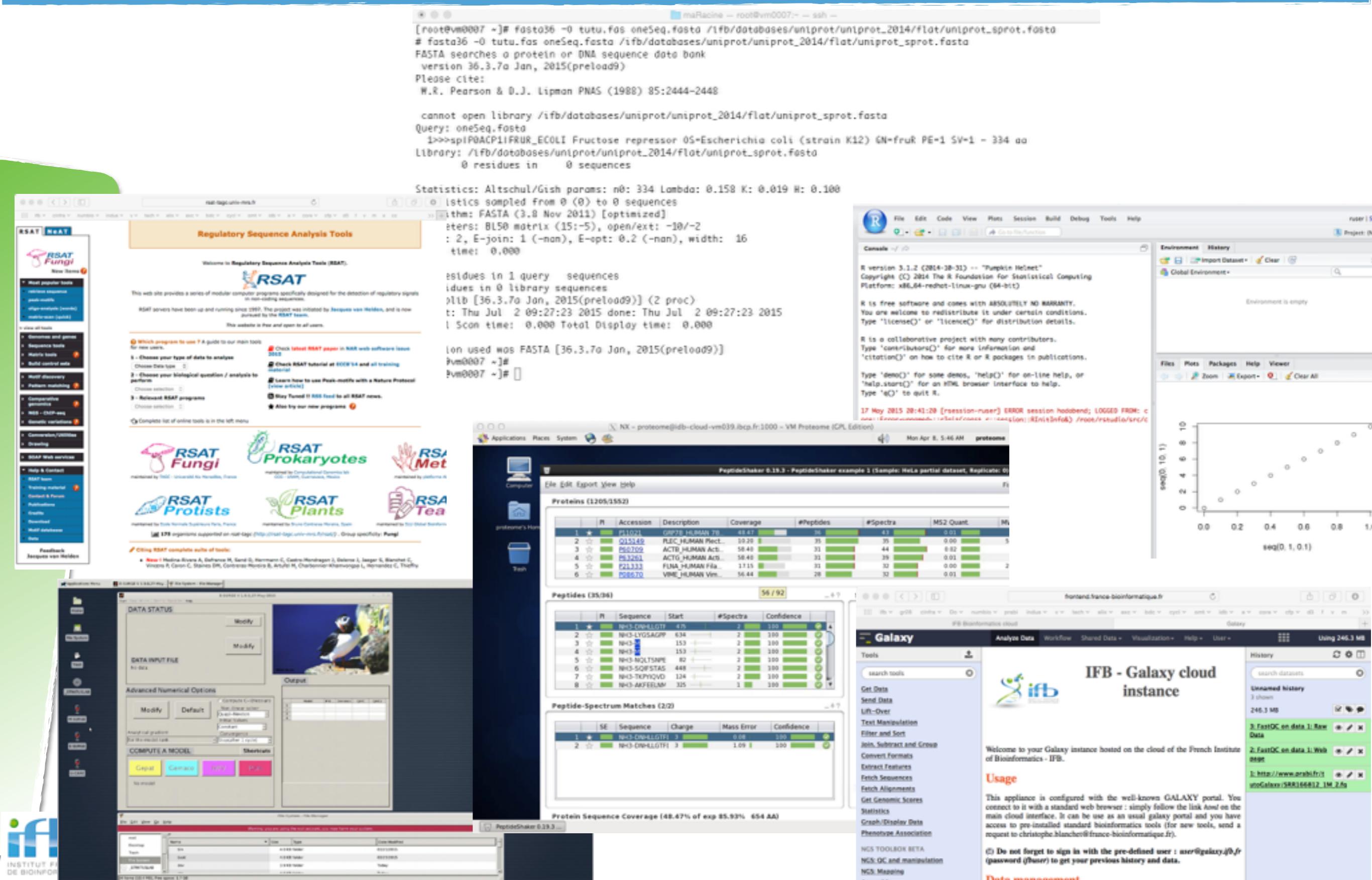
X!tandem

I2-10-01-I

XPLOR-NIH 2.30

...

Many interfaces



The French Institute of Bioinformatics and its e-infrastructure

IFB - Institut Français de Bioinformatique

IFB, the French distributed infrastructure for life-science information

Mission : to make available core bioinformatics resources to the national/international life science research community.

- To provide support for national biology programs
- To provide an IT infrastructure devoted to management and analysis of biological data
- To act as a middleman between the life science community and the bioinformatics/computer science research community



<http://www.france-bioinformatique.fr>

CNRS UMS3601. Avenue de la Terrasse, Bât 21. 91190 Gif-sur-Yvette



ELIXIR French Node

- The European distributed infrastructure for life-science information
- To optimize the interactions and coordination between the national level and ELIXIR and other ESFRI infrastructures in biomedical and environmental field,
- To promote consistency and complementarities between the components offered by the ELIXIR French node and those of other European nodes

IFB e-Infrastructure

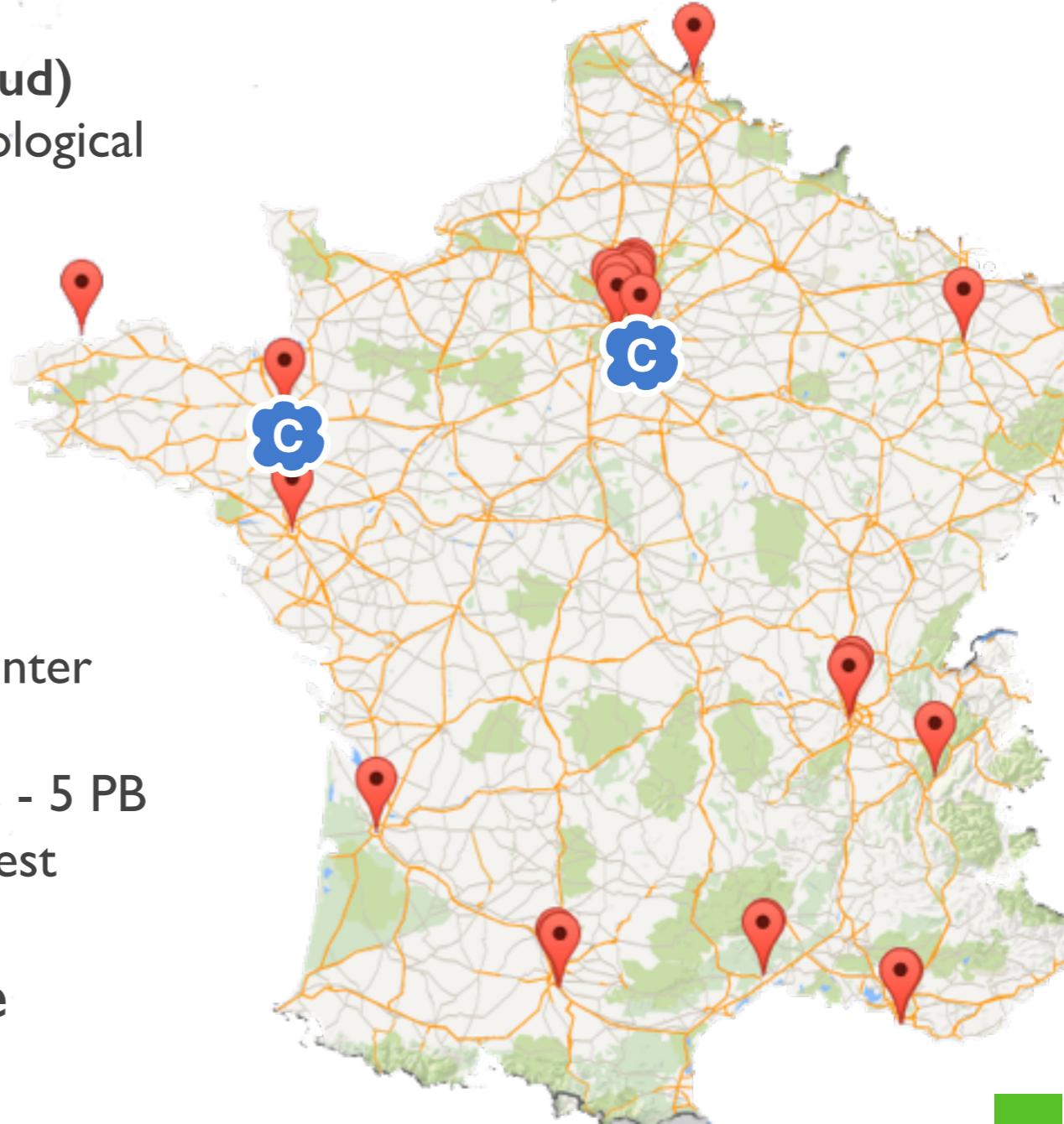
Mission : to provide core bioinformatics resources to the life science research community.

- To set up a **French IT infrastructure (cloud)** devoted to management and analysis of biological data
- To provide hardware, data collections and bioinformatics tools
- To collaborate with international infrastructure (ELIXIR)

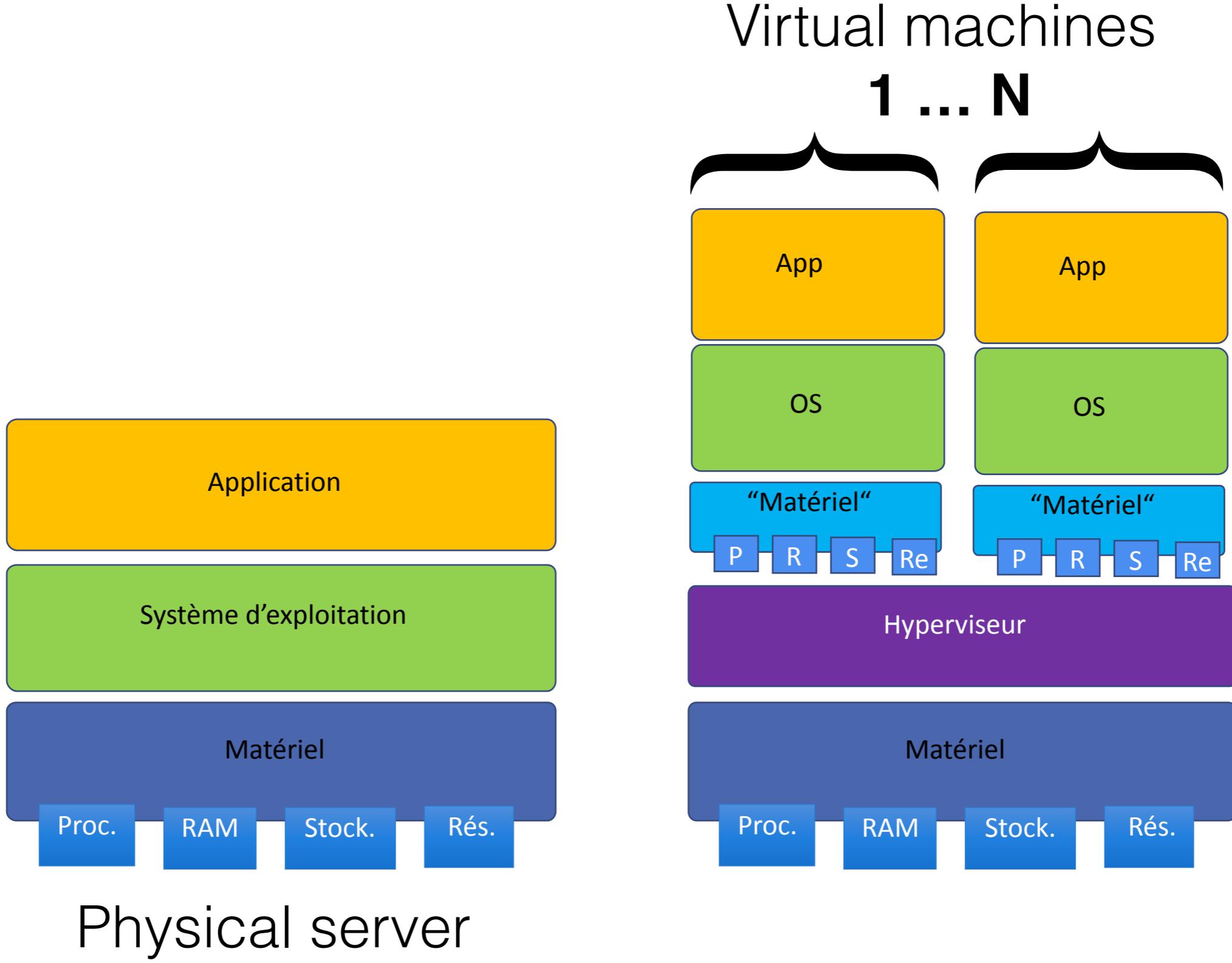
Current resources

- A **national hub : IFB-core**
IT resources hosted at CNRS IDRIS SC center
- A **network of regional centers**
32 bioinformatics platforms - 15,000 cores - 5 PB
- Two running clouds : IFB-core and GenOuest

→ **Create a federation of clouds for life sciences**

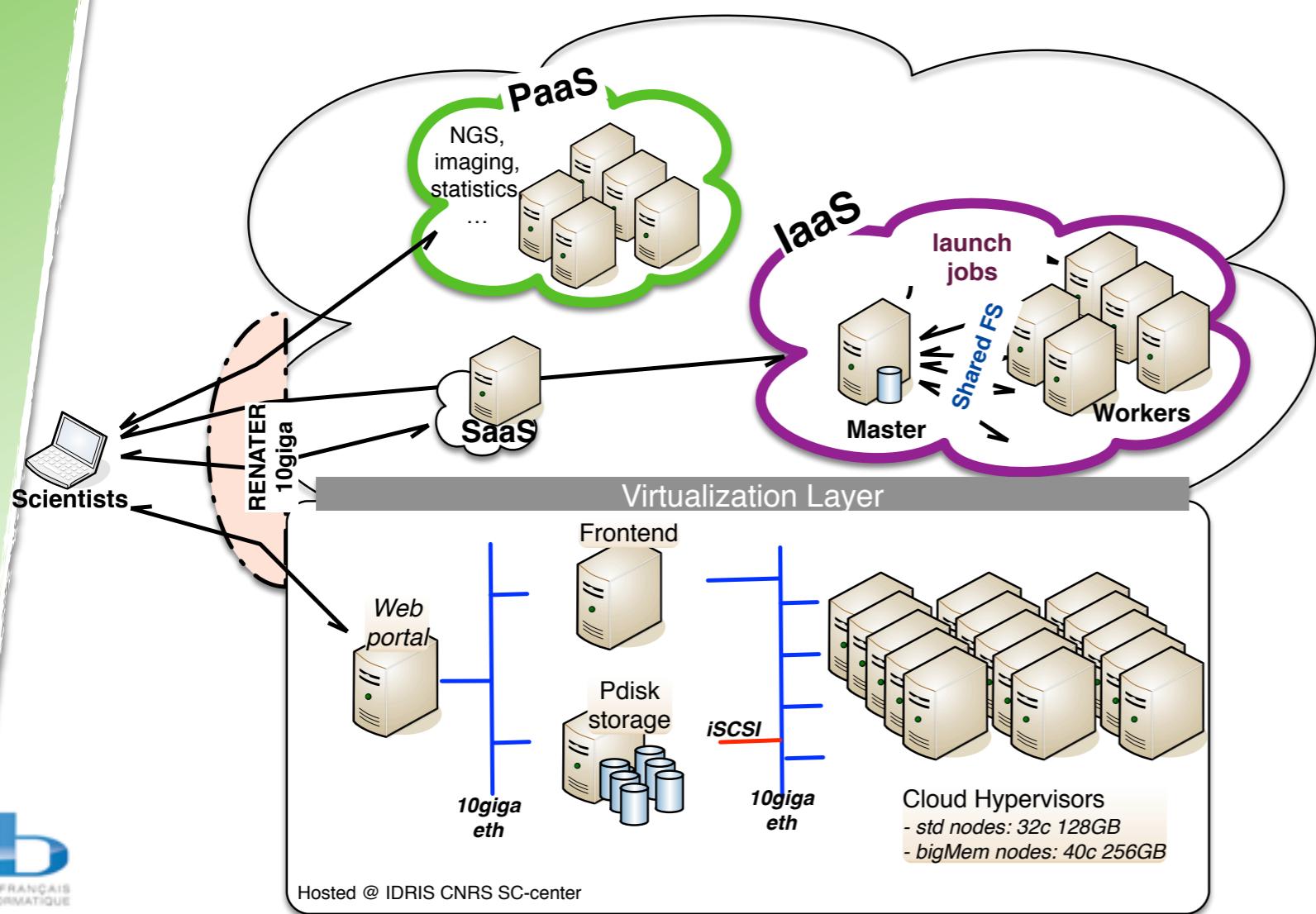


Virtualisation



IFB-core's cloud

IFB-core	# Compute Cores	# TB Storage	# TB RAM	Max VM size	Technology	Location
Pilot	200	50	2	40c 256GB	StratusLab	CNRS-IDRIS, Paris
2016-S1	3,000	500	-	?144c 3TB?	StratusLab	CNRS-IDRIS, Paris
2017	10,000	2,000	-	??	StratusLab	CNRS-IDRIS, Paris

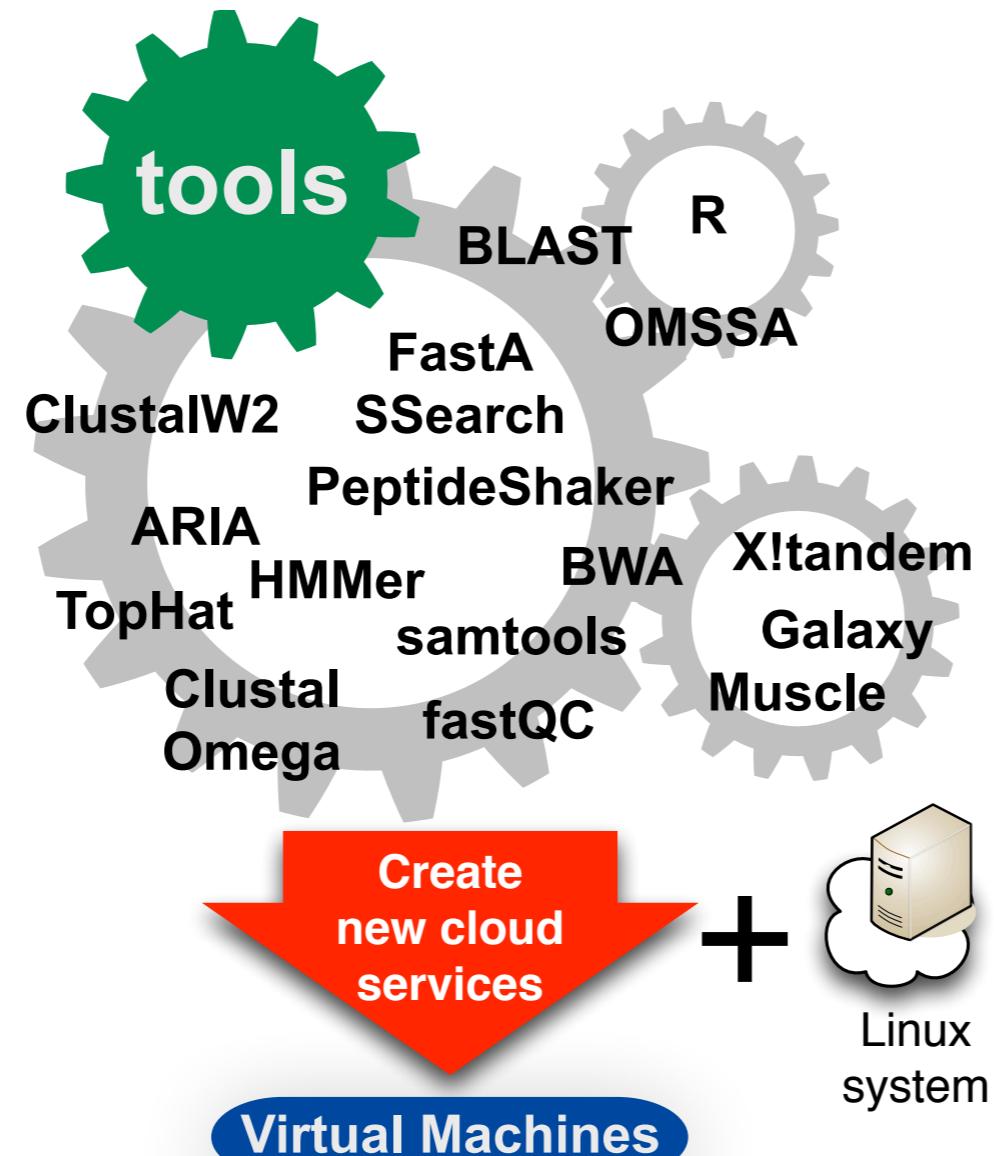
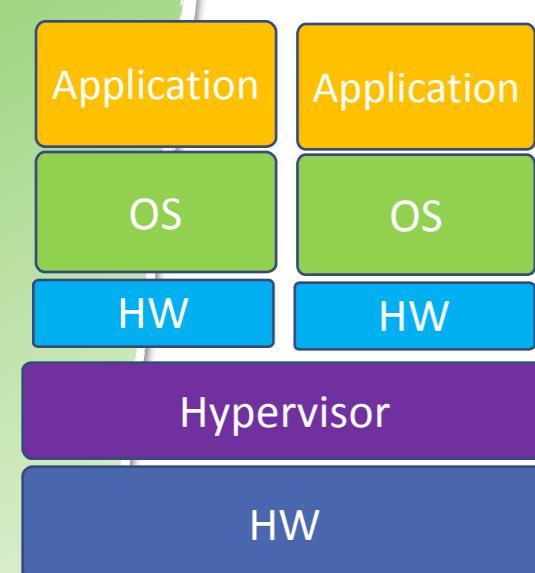
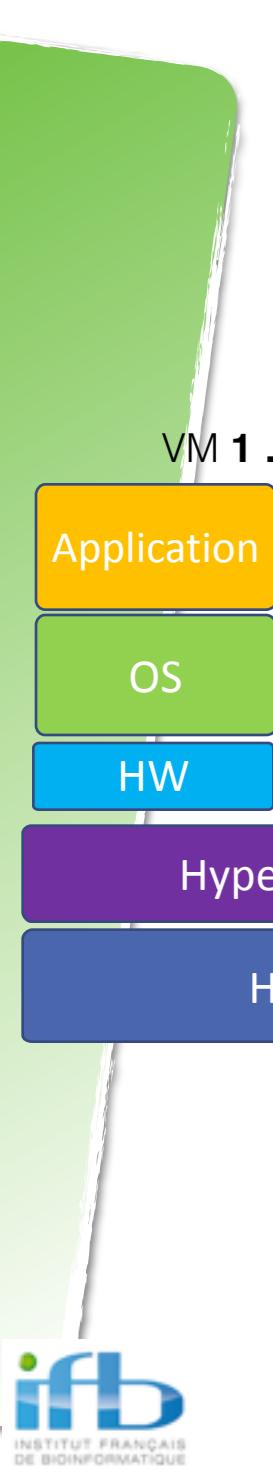


stratuslab



**Provide scientists with
bioinformatics resources**
- data and tools -
as cloud appliances

Create bioinformatics “appliances”

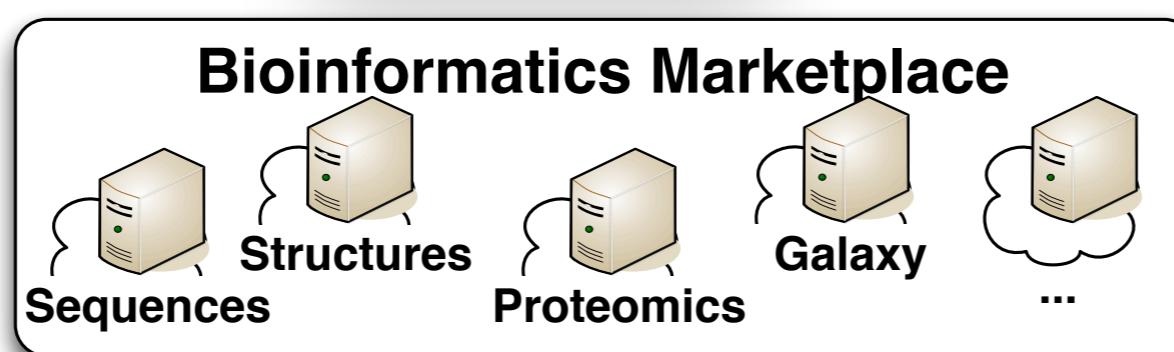


Appliance ?

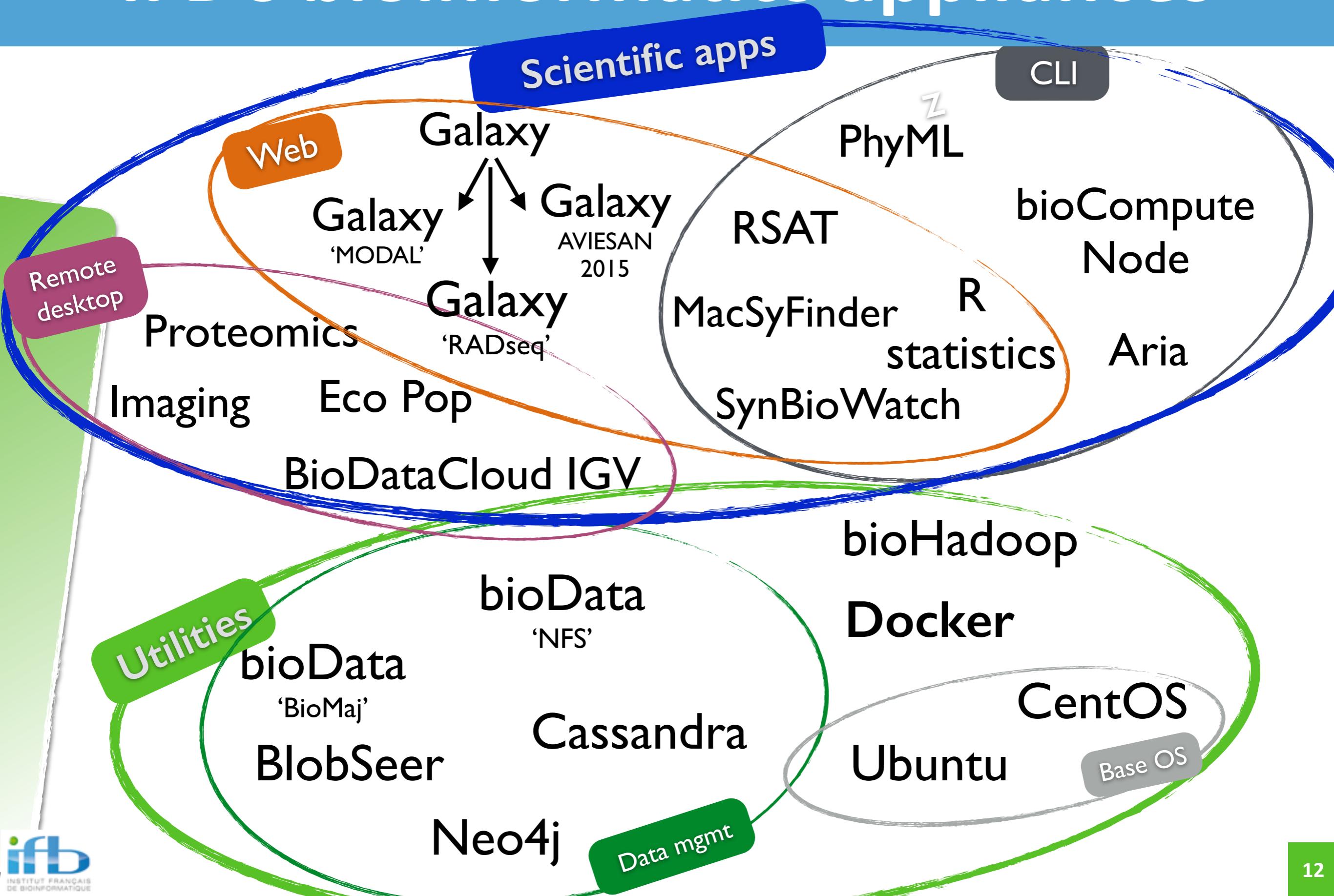
- predefined virtual machine
- including tools, pipeline, recipes...
- Ready to run

Appliance annotation

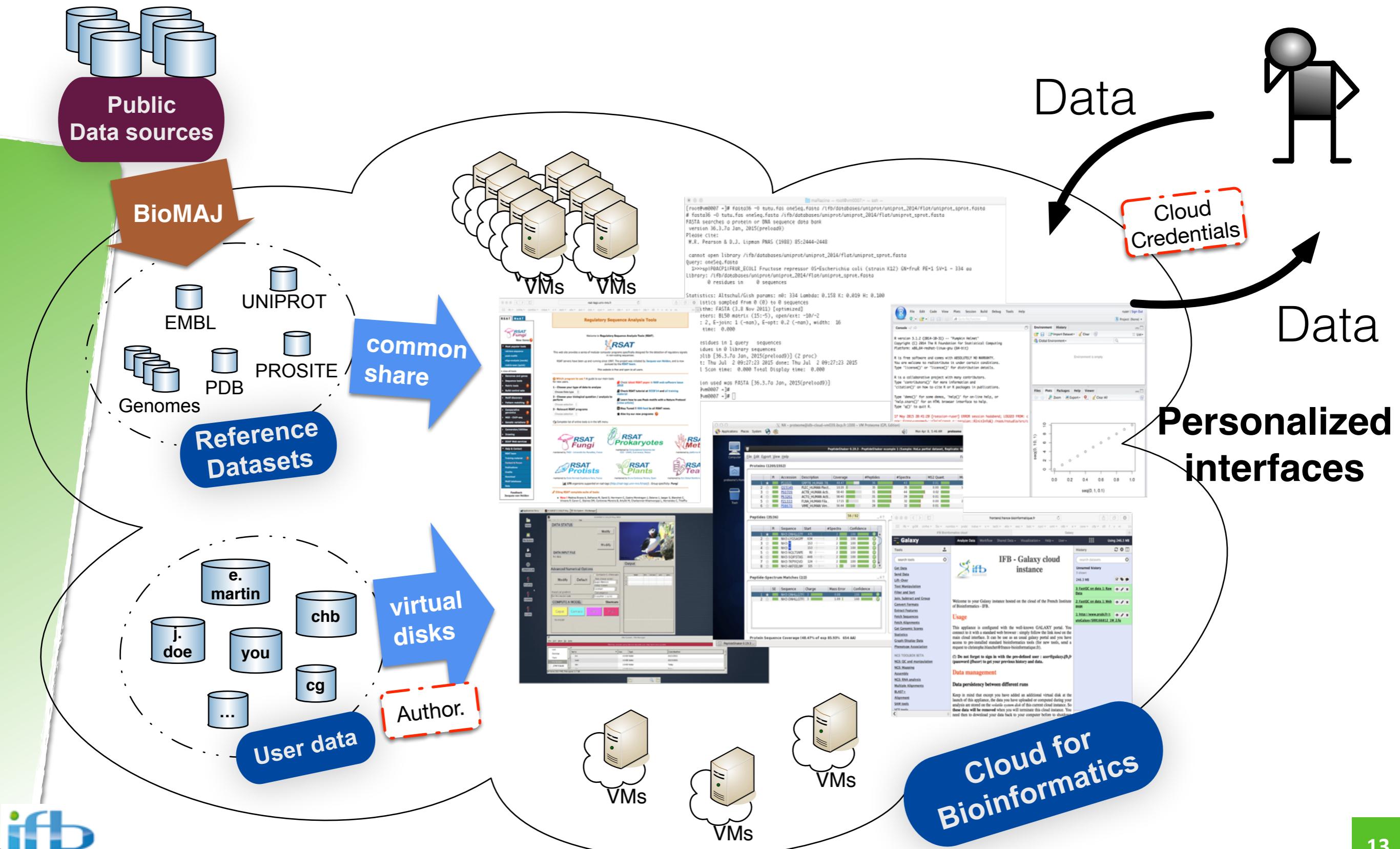
- Title
- Description (w. controlled voc.)
 - ★ Topics
 - ★ Tools
- Contact
- Developer(s) and maintainer(s) !



IFB's bioinformatics appliances



IFB's cloud for Bioinformatics



A cloud driven through a web dashboard

IFB Bioinformatics cloud

Rechercher sur Google ou saisir une adresse

ifb tech elix idb e core bdc cfp d3 m mapred viz r f i g tech f gbio privé v nimbio stra IRT cri 2r a l ngs

Lecteur

IFB BIOINFORMATICS CLOUD

You are signed in as cblanchet

News | Dashboard | MONITOR | SETTINGS | ADMINISTRATION | HELP | SIGN OUT

Hosted at Powered by

News

2014-09-04: The IFB cloud will be demonstrated at ECCB14
2014-08-11: Mise à jour du site web

Rooms

	c2.large	c2.small	c2.xlarge	c3.large	c3.medium	c3.xlarge	c3.xxlarge	m1.medium	m1.xlarge	m1.xxlarge
	18 / 36	78 / 144	8 / 18	18 / 34	39 / 70	8 / 16	4 / 6	15 / 20	2 / 2	0 / 2

Shutdown Go Get IPs Rename

New Instance New Storage Show Instances Show Storages Show Appliances

Search:

STORAGE

Username	ID	Name	Appliance	CPU%	CPU	Mem.	#Storage	Access
cblanchet	502	Public data source	BIO Data	2%	4	8	1	ssh http
cblanchet	515	compute	BIO ComputeNode	0%	4	8	0	ssh
cblanchet	811	ubu 10G	Ubuntu 14.04 (base) - 10GB	0%	4	8	0	ssh
cblanchet	838	RSAT genomes ChB	NFS server	0%	4	8	1	ssh
cblanchet	839	test4	RSAT-ub14 nfs	0%	4	8	0	ssh http
cblanchet	842	testrel	RSAT-ub14-mini	0%	4	8	0	ssh http

1 6 6 24 48 2

Show 25 entries First Previous 1 Next Last

CPU

MEMORY

<http://cloud.france-bioinformatique.fr/cloud>

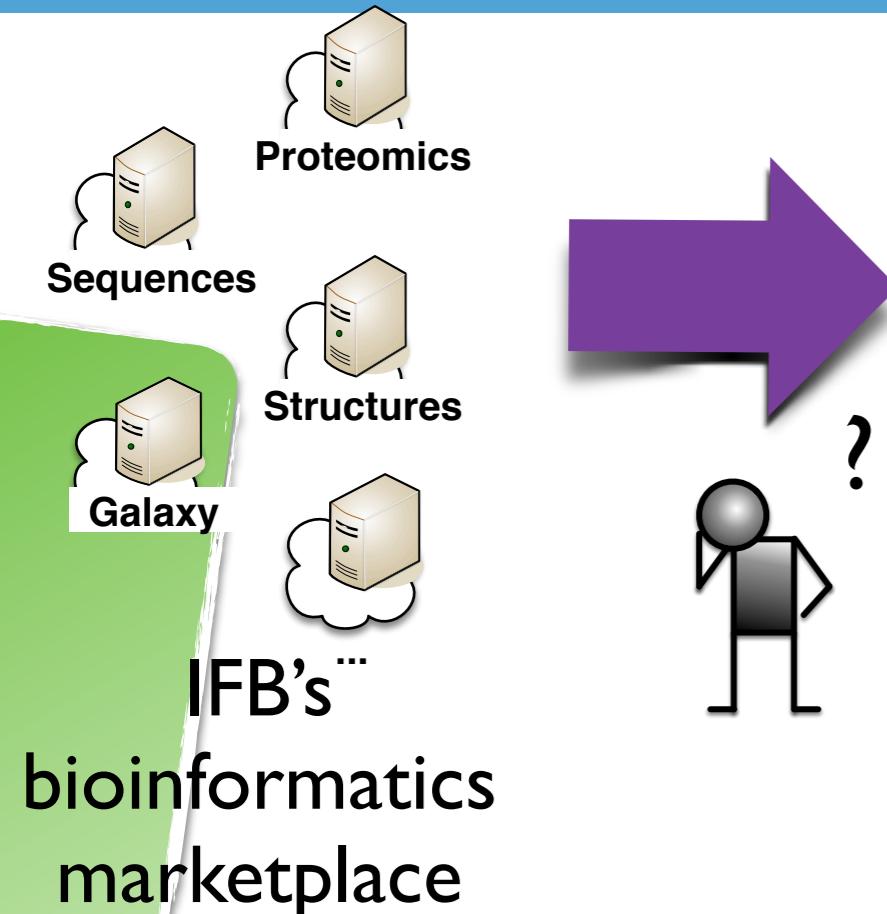
IFB acknowledges funding by the call "Infrastructures in Biology and Health" in the framework of the French "Investments for the Future" Initiative



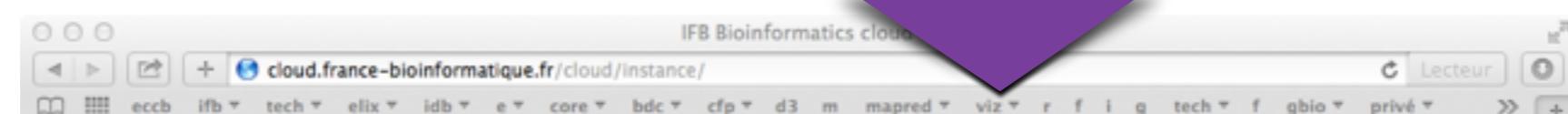
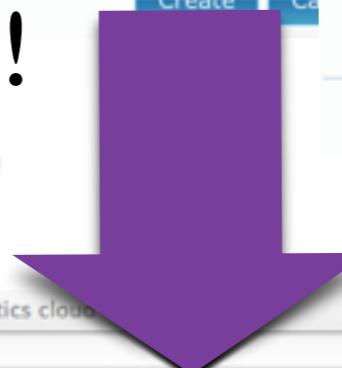
IFB is the French ELIXIR node



Browse the marketplace and run an App !



This screenshot shows two overlapping 'Create Instance' dialog boxes. The top dialog is for 'Galaxy' appliances, and the bottom one is for 'Galaxy MODAL'. Both dialogs include sections for 'Choose The Appliance', 'Configure Your Virtual Machines' (specifying Name, Unique, Type, and Number), and 'Configure Your Storage'. The bottom dialog also features a 'Persistent disk' dropdown menu.



A screenshot of the IFB Bioinformatics Cloud dashboard. The top navigation bar includes 'NEWS', 'Rooms', 'Shutdown', 'Go', 'Get IPs', 'Rename', 'New Instance', 'New Storage', 'Show Instances', 'Show Storages', and a search bar. The 'NEWS' section lists recent updates. The 'Rooms' section shows available instance types: c2.large (11/16), c2.small (44/64), c2.xlarge (5/8), and c3.large (11/15). The main table lists two instances: 'MODAL tools' (Galaxy MODAL) and 'myPortal' (Galaxy). The 'STORAGE' section shows a pie chart with 88.00% free space. The 'CPU' section shows a green progress bar at 100%.

Shutdown	Go	Get IPs	Rename	New Instance	New Storage	Show Instances	Show Storages
Showing 1 to 2 of 2 entries							
ID	Name	Appliance	CPU%	CPU	Mem.	#Storage	Access
968	MODAL tools	Galaxy MODAL	0%	16	64	1	ssh http
969	myPortal	Galaxy	0%	4	8	0	ssh http

Pratique

Connexion au Cloud IFB

<http://cloud.france-bioinformatique.fr/cloud>

The screenshot shows the IFB Bioinformatics Cloud login interface. At the top, there's a navigation bar with various links like 'IFB', 'tech', 'elix', etc. Below it is a main header with the 'IFB BIOINFORMATICS CLOUD' logo, a 'SIGN IN' button, and logos for 'Hosted at oris' and 'Powered by stratuslab'. The central part is a login form with fields for 'Username' (containing 'cblanchet') and 'Password' (containing '*****'). Below the form are links for 'Lost password' and 'Request account'. A note about funding from ANR and INVESTISSEMENTS D'AVENIR is also present. The bottom section shows a 'DASHBOARD' view with news items, a table for 'Rooms', and storage and CPU usage charts.

ID	Name	Appliance	CPU%	CPU	Mem.	#Storage	Access
0			0	0	0	0	0

Rooms

Type	Status
c2.large	16 / 16
c2.small	64 / 64
c2.xlarge	8 / 8
c3.large	15 / 15
c3.medium	31 / 31

STORAGE

CPU

Connectez-vous
au cloud IFB dans
'Sign in'

Connexion au Cloud IFB (2)

Bioinformatics cloud

You are signed in as cblanchet | [Settings](#) | [Instances](#) | [Monitor](#) | [Help](#) | [Sign out](#)

Bioinformatics cloud

Powered by
stratuslab
Hosted at
ifris

Edit profile

Personal Information

Affiliation ? CNRS IFB
City ? Gif-sur-Yvette

Cloud Preferences

Pubkey ?
ssh-dss
AAAAAB3NzaC1kc3MAAACBAMskM5bTOtlo/2EOeyXhbd
FGJuCVZQDssG8NdNs9IfKnRJ6/yVmqEBjouLCy4YY3ylh
YEVMZjVc5wYZgh9aRAzqFac3K+qXJMkZKX+taXaBbWP
PTXM49jo9+/ldr+rkJIN3fbBJT0wN8/ZZ9DFPMDjVyAX1
nOd9VsKnTTvzGs3BPAAAFCQCKdicMAMjafqqKAszABQ
1s/Hs4aQAAIA24WTy7RygmOuRTRcrF9qu40L24rDfg9
R7N839KQH2Pry59Rv/vVByzJMG5IxKSBWZKCScg5d7SJ
3H7TAinlWWxFAsL7I2Z52CiRUHBuGc9dptBGLjP6VwhC
/aAANkKn0rWVaT7BScNzbl7adGTZ8He5tS57XoEnjVfl2

Appliance ? BIO ComputeNode
Instance type ? c2.large

Save Back

Lors de la première connexion

- rubrique ‘Settings’
- complétez vos paramètres

! clé SSH:

- fichier `~/.ssh/dsa.pub`
- attention aux retours à la ligne lors du copier-coller
- la créer avec `ssh-keygen` (ou PuTTYgen)

Le tableau de bord du Cloud IFB

Gérer ses VMs

- Créer / arrêter / renommer

Gérer ses disques virtuels

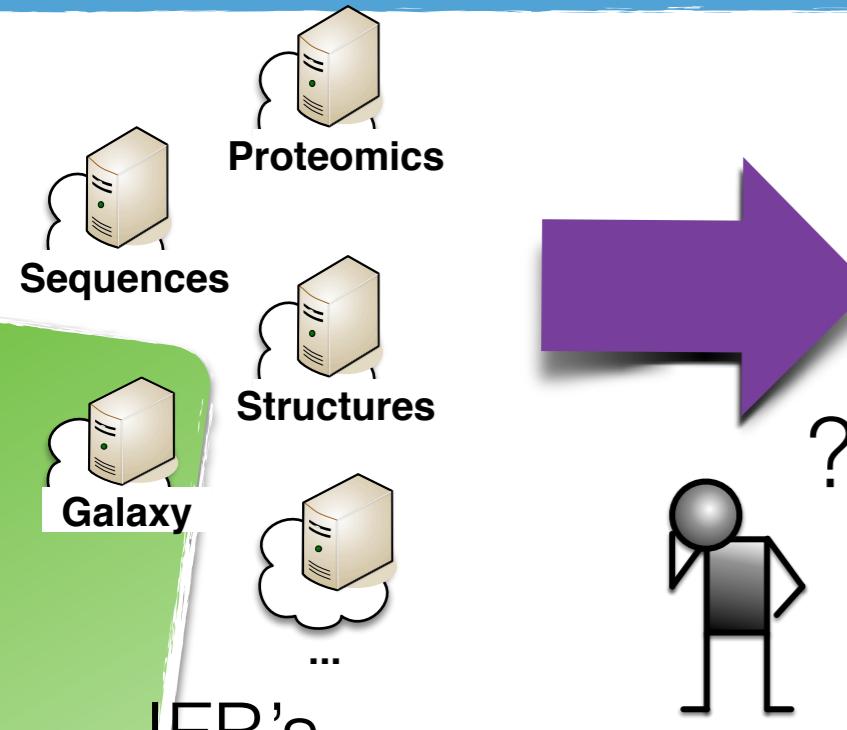
- Créer / supprimer

Visualisez les paramètres

- nom/type/taille
- état/charge CPU
- disque virtuel attaché

Instance	Shutdown		Go	Get IPs	Rename	New Instance	New Storage	Show Instances	Show Storages	Room for VMs
Showing 1 to 10 of 10 entries										
	<input type="checkbox"/>	♦ ID	♦ Name	♦	♦ Appliance	♦ CPU%	♦ CPU	♦ Mem.	♦ #Storage	Access
	<input type="checkbox"/>	3908	13-02		BIO compute node 2.3	0%	2	8	0  	ssh/scp
	<input type="checkbox"/>	3909	prot		Proteomics 2.2	0%	4	16	0  	ssh/scp
	<input type="checkbox"/>	4045	MG#1		MaHT	0%	8	27	0  	ssh/scp
	<input type="checkbox"/>	4046	MG#2		MaHT	0%	8	27	0  	ssh/scp
	<input type="checkbox"/>	4047	bioMR		BIO MapReduce v0.5	1%	8	27	1  	ssh/scp
	<input type="checkbox"/>	4048	bioMR#1		BIO MapReduce v0.5	0%	8	27	0  	ssh/scp
	<input type="checkbox"/>	4049	bioMR#2		BIO MapReduce v0.5	0%	8	27	0  	ssh/scp
	<input type="checkbox"/>	4122	sumo		SuMo v1	0%	2	8	0  	ssh/scp
	<input type="checkbox"/>	4254	fastqc Q		BIO compute node	0%	2	8	0  	ssh/scp
	<input type="checkbox"/>	4261	galaxy		Galaxy	0%	2	8	0  	ssh/scp http
		10		7		52	183	1		
	Show	25	entries							First Previous 1 Next Last

Créer une machine virtuelle



The 'Create Instance' dialog box displays two separate configurations for virtual machines:

- Create Instance (Left):**
 - Choose The Appliance:** Galaxy, Filter by --- THEMATIC FIELDS ---, Galaxy portal
 - Configure Your Virtual Machines:** Name: myPortal, Unique: (unchecked), Type: c2.large (4 CPU, 8GB RAM), Number: 1
 - Configure Your Storage:** Persistent disk: [dropdown menu]
- Create Instance (Right):**
 - Choose The Appliance:** Galaxy MODAL, Filter by --- THEMATIC FIELDS ---, Galaxy portal
 - Configure Your Virtual Machines:** Name: MODAL tools, Unique: (unchecked), Type: c3.xxlarge (16 CPU, 64GB RAM), Number: 1
 - Configure Your Storage:** Persistent disk: ECCB 10

Both dialog boxes have 'Create' and 'Cancel' buttons at the bottom.

IFB's
Marketplace

The screenshot shows the IFB Bioinformatics Cloud dashboard with the following elements:

- Header:** cloud.france-bioinformatique.fr/cloud/instance/ (URL), Lecteur (Reader) button, IFB Bioinformatics Cloud logo, IFB INSTITUT FRANÇAIS DE BIOINFORMATIQUE logo, IFB BIOINFORMATICS CLOUD logo, DASHBOARD tab, Hosted at ORIS logo, Powered by stratuslab logo, YOU ARE SIGNED IN AS IFBTUTO20, News | DASHBOARD | MONITOR | SETTINGS | HELP | SIGN OUT.
- News:** 2014-09-04: The IFB cloud will be demonstrated at ECCB14; 2014-08-11: Mise à jour du site web.
- Rooms:** Shows 2 rooms, 25 entries, and a table with columns: ID, Name, Appliance, CPU%, CPU, Mem., #Storage, Access. Entries include 968 MODAL tools (Galaxy MODAL, 0%, 16, 64, 1, ssh http) and 969 myPortal (Galaxy, 0%, 4, 8, 0, ssh http).
- Storage:** Shows free (88.00%) storage with a pie chart.
- CPU:** Shows a progress bar indicating CPU usage.

Choisir la bonne appliance

Launch a virtual machine

Choose The Appliance

Appliance ? BIO ComputeNode (2015-03)

Filter by ? Genomics tools

--- TOOLS ---

Configure Your Virtual Machine

Name ?

Unique ?

Type ? c3.medium (2 CPU, 8GB RAM)

Number ? 1

Plug Your Additional Storage

Persistent disk ?

Launch a virtual machine

Choose The Appliance

Appliance ? EBA15 Galaxy ChIP-seq 0.8

Filter by ? Genomics tools

deepTools 1.5.9

Configure Your Virtual Machine

Name ?

Unique ?

Type ? c3.medium (2 CPU, 8GB RAM)

Number ? 1

Plug Your Additional Storage

Persistent disk ?

Run

ifb
INSTITUT FRANÇAIS
DE BIOINFORMATIQUE

Caractéristiques d'une VM

The screenshot shows a web-based interface for launching a virtual machine. At the top, there's a navigation bar with links like 'cloud.france-bioinformatique.fr', 'gr28', 'eba15', 'Do', 'prabi', 'indus', etc. Below this is a header with 'BIOINFORMATICS CLOUD' and 'DASHBOARD'. A sub-header says 'YOU ARE SIGNED IN AS DEMO' with links for 'NEWS', 'DASHBOARD', 'MONITOR', 'SETTINGS', 'HELP', and 'SIGN OUT'. The main content area has a title 'Launch a virtual machine' and a sub-section 'Choose The Appliance'. It shows an 'Appliance' dropdown set to 'EBA15 Galaxy ChIP-seq 0.8' and a 'Filter by' dropdown set to 'Genomics tools' with an option 'deepTools 1.5.9' below it. Below this is a section 'Configure Your Virtual Machine' with fields for 'Name' (set to 'maGalaxy'), 'Unique' (unchecked), 'Type' (set to 'c3.medium (2 CPU, 8GB RAM)'), and 'Number' (set to '1'). There's also a section 'Plug Your Additional Storage' with a 'Persistent disk' dropdown set to 'monDDgalaxy'. At the bottom right are 'Run' and 'Cancel' buttons.

Appelée aussi 'Instance'

Définir sa VM

- un nom
- nombre de CPUs
- taille mémoire
- attacher un disque virtuel

Cluster de VMs

- remplir le nombre de VMs
- choix du nom unique

Les disques durs virtuels

Pour stocker ses données

- taille et nombre variable (quota)
- retrouver ses données d'une VM à la suivante.
- ! pas de sauvegarde !

Sur un vdisk

- l'attacher à une (seule) VM
 - à la création de la VM
- ### Partager un vdisk
- mode cluster
 - VM 'NFS'

The screenshot shows the IFB Bioinformatics Cloud dashboard. At the top, it displays the IFB logo, the text "INSTITUT FRANÇAIS DE BIOINFORMATIQUE", and navigation links for "NEWS", "DASHBOARD", "MONITOR", "SETTINGS", "HELP", and "SIGN OUT". It also indicates "YOU ARE SIGNED IN AS DEMO". Below this, it says "Hosted at ORIS" and "Powered by stratuslab".

The main area has tabs for "NEWS", "DASHBOARD", and "STORAGE". The "STORAGE" tab is active, showing a table of virtual disks:

Room for VMs	Name	Size (GiB)	UUID
c2.large	monDisque2	20	383a8a7c-a538-454f-9c50-0c7a310bb50e
c2.small	monDDgalaxy	10	32dfe8f1-fade-4199-8fe5-77131bf53e5b
c2.xlarge		30	

Below the table, there are buttons for "Delete selected storages", "Go", "Rename", "New Instance", "New vDisk", "Show Instances", "Show vDisks", and a search bar. A "ROOM FOR VMs" sidebar lists available instance types with their counts: c2.large (16 / 16), c2.small (64 / 64), c2.xlarge (8 / 8), c3.large (16 / 16), c3.medium (32 / 32), c3.xlarge (8 / 8), c3.xxlarge (4 / 4), and m1.medium (15 / 15). At the bottom, there are links for "First", "Previous", "1", "Next", and "Last".

A green circular progress bar for "CPU" is visible on the right side of the dashboard.

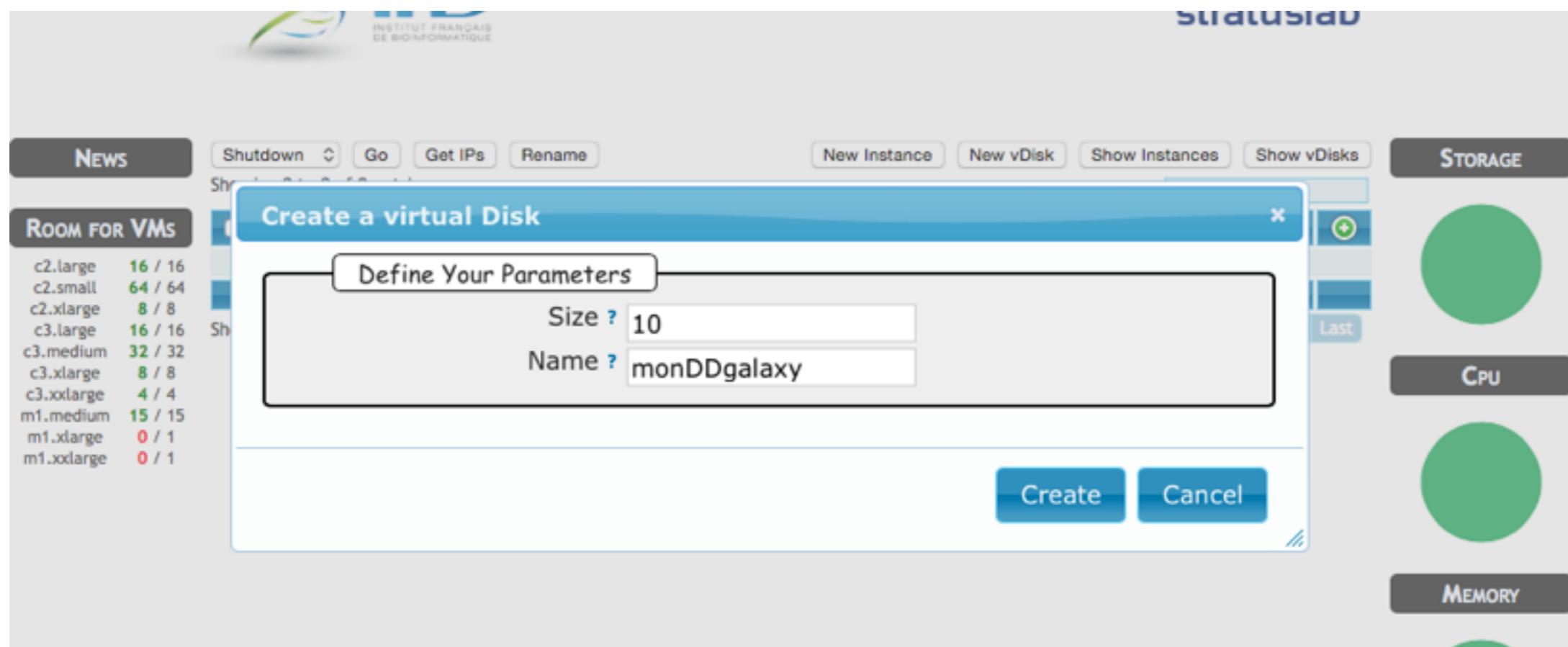
Pratique

Depuis le tableau de bord du cloud

- <http://cloud.france-bioinformatique.fr/cloud/>

Créez votre disque virtuel

- bouton ‘New vDisk’
- monDDgalaxy, 10Go



Pratique

Depuis le tableau de bord du cloud

- <http://cloud.france-bioinformatique.fr/cloud/>

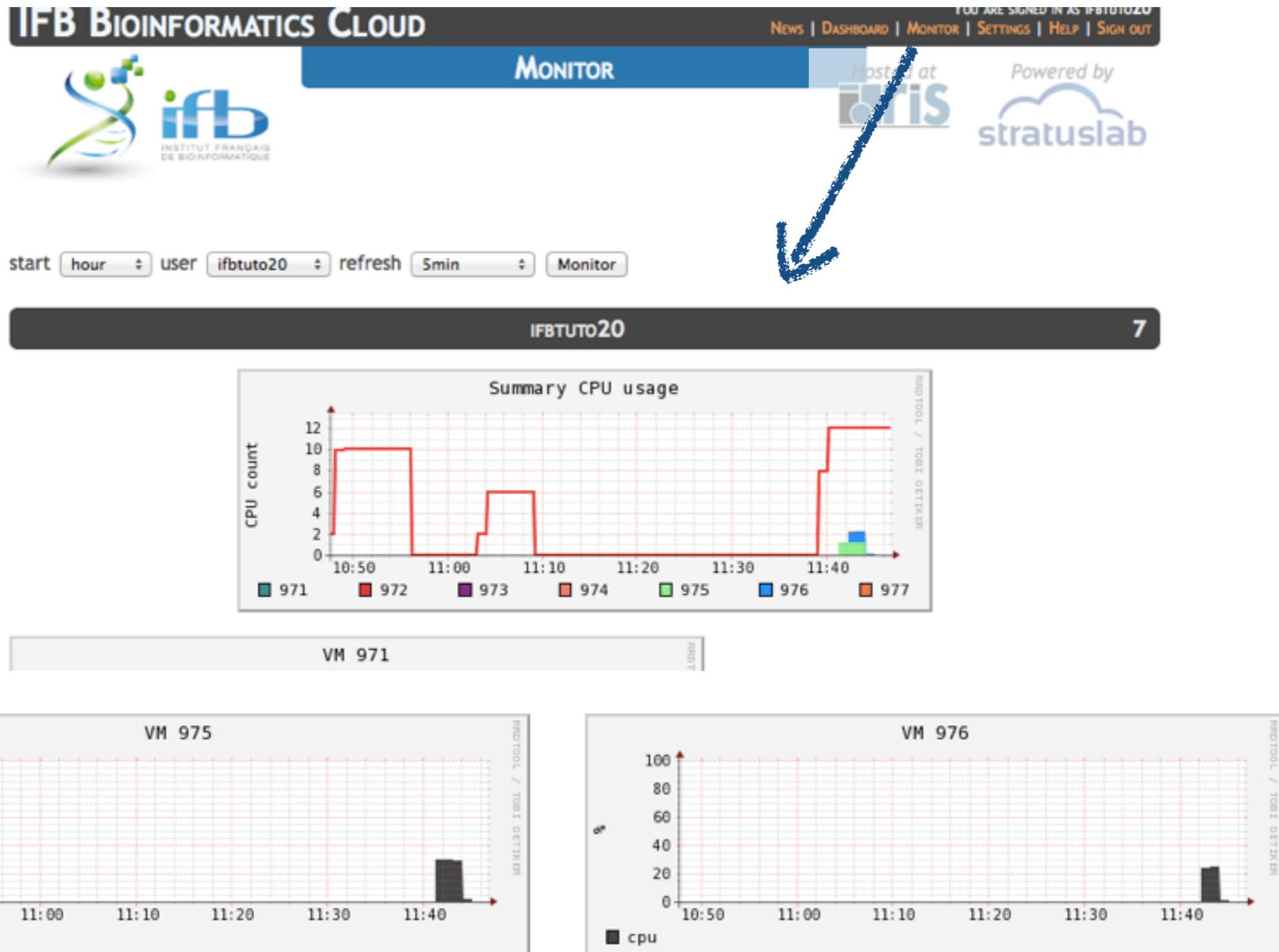
bouton ‘New instance’

- identifier quelle appliance fournit l’outil deepTools ?
- quelle version de l’outil ?

Créer une instance EBAI5 Galaxy ChIP-seq

- ★ nom: galaxy_roscoff_2015
- ★ taille: c3.medium (2 CPU, 8 GB RAM)
- ★ attacher votre disque virtuel “monDDgalaxy”

Monitor your usage



Questions ?

<http://www.france-bioinformatique.fr>

Acknowledgments

- IFB members
 - IFB hub: Patricia, Jean-François, Mohamed, Jonathan, Maxime, Dominique
Alumni : Marie, Quentin
➡ we are hiring !
 - Working group IFB-GRISBI (co-chair with Olivier Collin)
- Appliances developers
 Samuel Blanck (Inria Lille), Jacques van Helden (TAGC), Stéphane Delmotte (PRABI-Doua), Bruno Spataro (PRABI-Doua), Marie-Laure Frachinard (MIGALE), Anis Djari (BioinfoGenoToul), Bertrand Néron (Institut Pasteur), Adrien Josso (MicroScope), Thomas Lacroix (MIGALE), Christian Baudet (CLB), Germain Paimparay & Baptiste Brault (CFB)...
- CNRS IDRIS: R. Medeiros, C. Gauthey and staff
- StratusLab members
- IFB is funded by French programs PIA INBS 2012, BioDataCloud
- EU H2020 projects, CYCLONE (644925) and EGI-Engage (654142)