PF AuBI, 7 décembre 2022, Clermont-Ferrand

Atelier Biosphère

Christophe Blanchet

Institut Français de Bioinformatique - IFB
CNRS UAR3601 & PRABI-AMSB, LYON, FRANCE

Agenda

7 décembre 2022, 9h00-13h00

Ateliers Biosphère

- Introduction
- Créer son compte
- Déployer une app
- Se connecter à sa VM
- Modifier l'environnement
  - ★ python, conda, pip
  - ★ docker
- Déployer son propre environnement
  - ★ avec un conteneur : Rstudio, RShiny et Jupyter
  - ★ avec un dépôt git : R Shiny
  - ★ Avec un dépôt git (install/deploy)
- Préparer son environnement Rstudio et Jupyter
- Outils de développement : git, VScode

Institut Français de Bioinformatique

- 36 plateformes and teams, 1 coordinating hub
- Missions
  - Provide **core infrastructure of bioinformatics services** to the French science community
  - Provide **training** in bioinformatics for **life scientists and bioinformaticians**
  - Provide **support to communities in biology, health, agronomy and environment** research programs, and an access to **high-level expertise and mutualized competences**
  - **Develop a strategic vision** to maintain French research at the highest level of expertise in the analysis of biological data
  - Be a **lever to ambitious national research projects** in the life sciences and their applications
  - Ensure the **international representation of the French bioinformatics community**

Infrastructure Cloud de l’IFB

Une infrastructure informatique pour déployer des environnements bioinformatiques dans des machines virtuelles et conteneurs pour le traitement des données biologiques.

- Un **portail unifié Biosphère** pour déployer tous les environnements bioinformatiques sur tous les clouds.
- Des **environnement de traitement pré-définis,** prêts à être **déployés en un clic** depuis le **catalogue RAINBio**
- Un **compte unique** pour **tous les clouds,** avec les **identifiants officiels** des utilisateurs dans leur organisme (fédération d’identités)
- **8 sites cloud** avec plus de 10 000 vCPU et 40 To RAM, ressources à large mémoire, haute fréquence CPU et grand nombre de coeurs par CPU.
- Haute-disponibilité avec les différents sites utilisables de la même façon.
- Accès aux bases de données publiques de référence**.**

SE CONNECTER

Utilisez vos identifiants de votre organisme (CNRS, INRAE, Inserm, Universités...)

Nous utilisons la fédération d'identité européenne eduGAIN. Si votre organisme académique n'est pas dans cette fédération, vous pouvez utiliser un compte local avec votre adresse professionnelle.

Assistance Biosphère

Doc en ligne

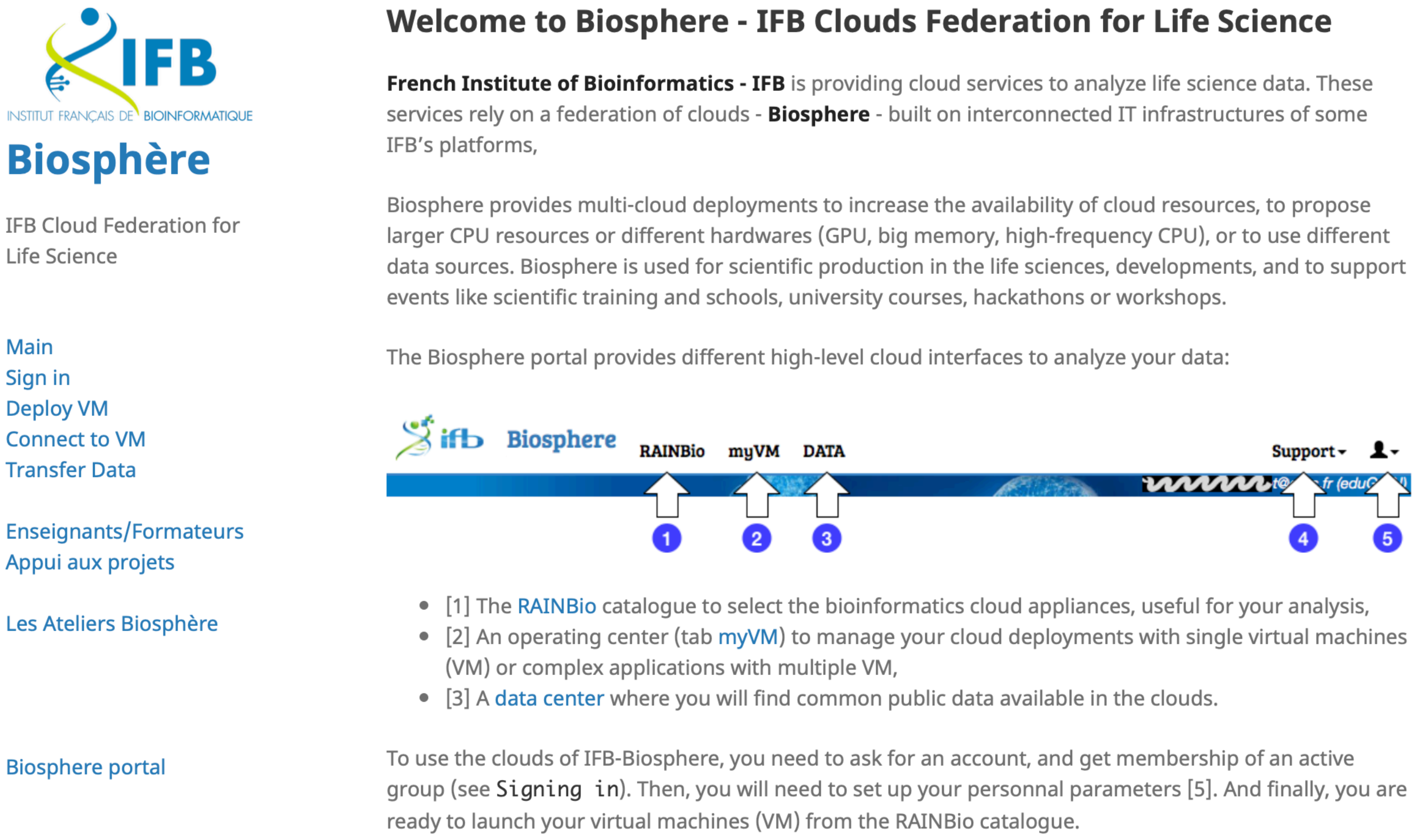
- https://ifb-elixirfr.github.io/biosphere/

Adresse mail support

- biosphere-support@genouest.org

Vos PFs

- Etat du Système



INSTITUT FRANÇAIS DE BIOINFORMATIQUE

Biosphère

IFB Cloud Federation for Life Science

Main
Sign in
Deploy VM
Connect to VM
Transfer Data

Enseignants/Formateurs
Appui aux projets

Les Ateliers Biosphère

Biosphere portal

Welcome to Biosphere - IFB Clouds Federation for Life Science

French Institute of Bioinformatics - IFB is providing cloud services to analyze life science data. These services rely on a federation of clouds - **Biosphere** - built on interconnected IT infrastructures of some IFB's platforms,

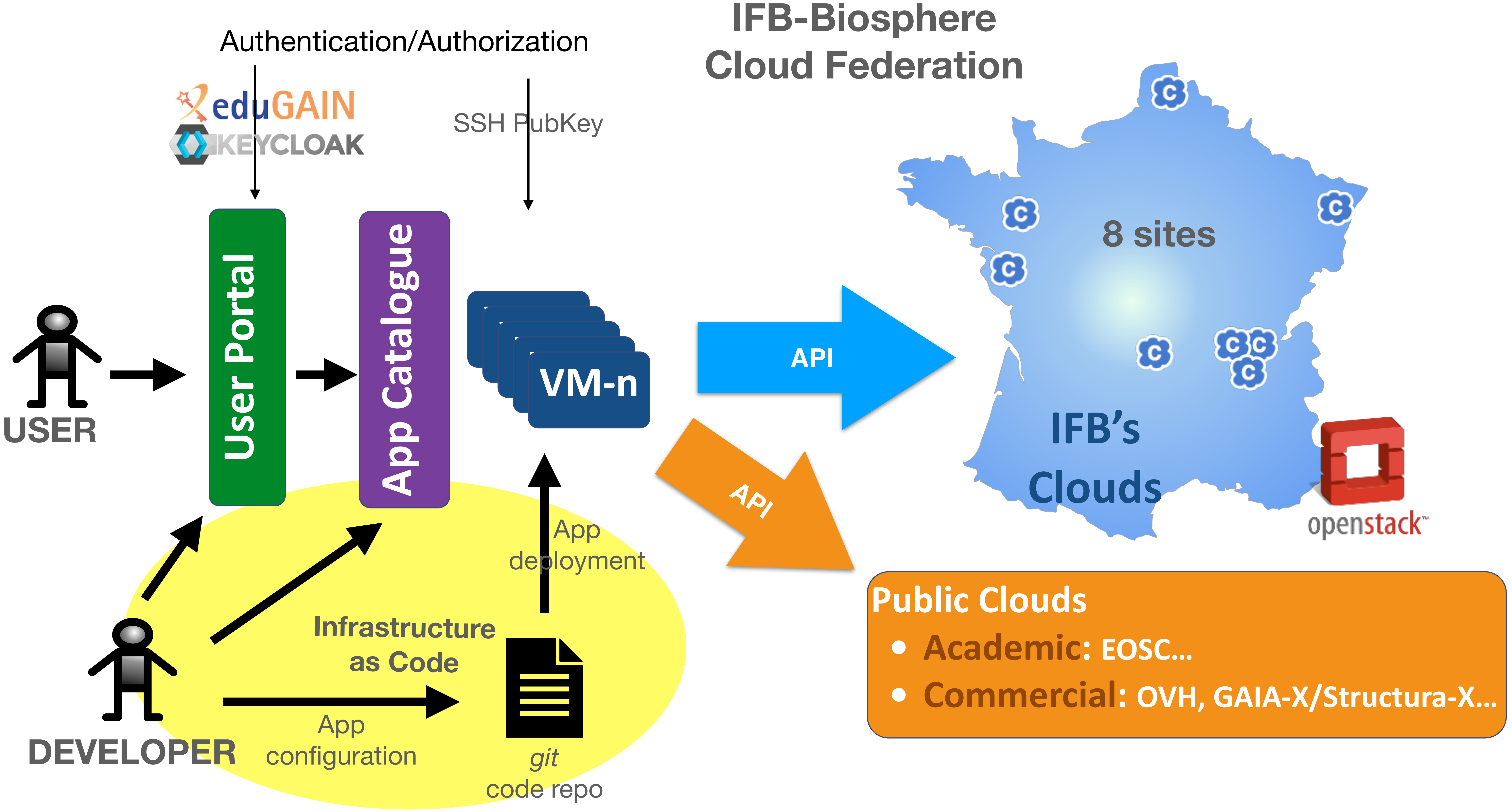
Biosphere provides multi-cloud deployments to increase the availability of cloud resources, to propose larger CPU resources or different hardwares (GPU, big memory, high-frequency CPU), or to use different data sources. Biosphere is used for scientific production in the life sciences, developments, and to support events like scientific training and schools, university courses, hackathons or workshops.

The Biosphere portal provides different high-level cloud interfaces to analyze your data:

- [1] The RAINBio catalogue to select the bioinformatics cloud appliances, useful for your analysis,
- [2] An operating center (tab myVM) to manage your cloud deployments with single virtual machines (VM) or complex applications with multiple VM,
- [3] A data center where you will find common public data available in the clouds.

To use the clouds of IFB-Biosphere, you need to ask for an account, and get membership of an active group (see `Signing in`). Then, you will need to set up your personnal parameters [5]. And finally, you are ready to launch your virtual machines (VM) from the RAINBio catalogue.

Architecture Clouds Biosphère



Public Clouds

- **Academic**: EOSC…
- **Commercial**: OVH, GAIA-X/Structura-X…

Se connecter au cloud Biosphère

1. S'identifier avec son identifiant de son organisme, université…
2. Choisir son unité d'affiliation, un projet ou une formation

SE CONNECTER

Utilisez vos identifiants de votre organisme (CNRS, INRAE, Inserm, Universités…)

Nous utilisons la fédération d'identité européenne eduGAIN. Si votre organisme académique n'est pas dans cette fédération, vous pouvez utiliser un compte local avec votre adresse professionnelle.

Rechercher le groupe « Atelier Biosphère »

LISTE DES GROUPES

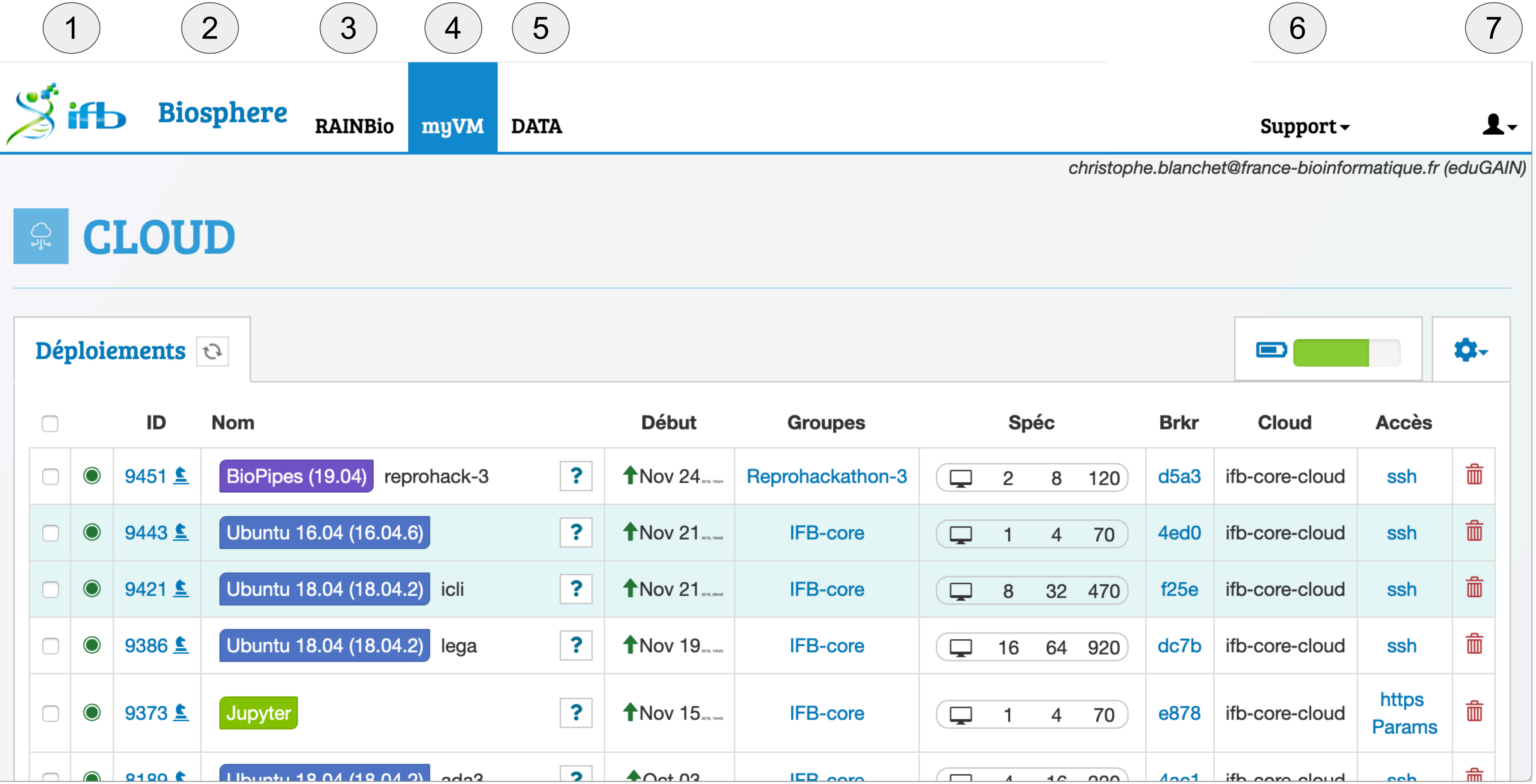
Afficher 10 éléments

Rechercher : inserm

	Nom	Site web	Actions
👥	BioCampus (UAR3426 BioCampus Montpellier), CNRS-UAR3426, INSERM-US009	↗	+
👥	BISCEm (Biologie Intégrative Santé Chimie Environnement), US 42 INSERM / UMS 2015 CNRS	↗	+
👥	iBV (Institute of Biology Valrose), CNRS UMR7277, Inserm U1091, UNS	↗	+
👥	INSERM-DSI (Département des Systèmes d'Information de l'INSERM), DSI	↗	+
👥	Institut Cochin (Institut Cochin U1016 Inserm/UMR8104 CNRS/UMR-S8104), U1016, UMR8104	↗	+
👥	IRSET (Institut de recherche en santé, environnement et travail - Inserm U1085), U1085	↗	+
👥	MitoVasc (MitoVasc - Université d'Angers, UMR CNRS 6015/INSERM 1083), UMR CNRS 6015/INSERM 1083	↗	+

https://ifb-elixirfr.github.io/biosphere/signin

Portail Multi-Cloud



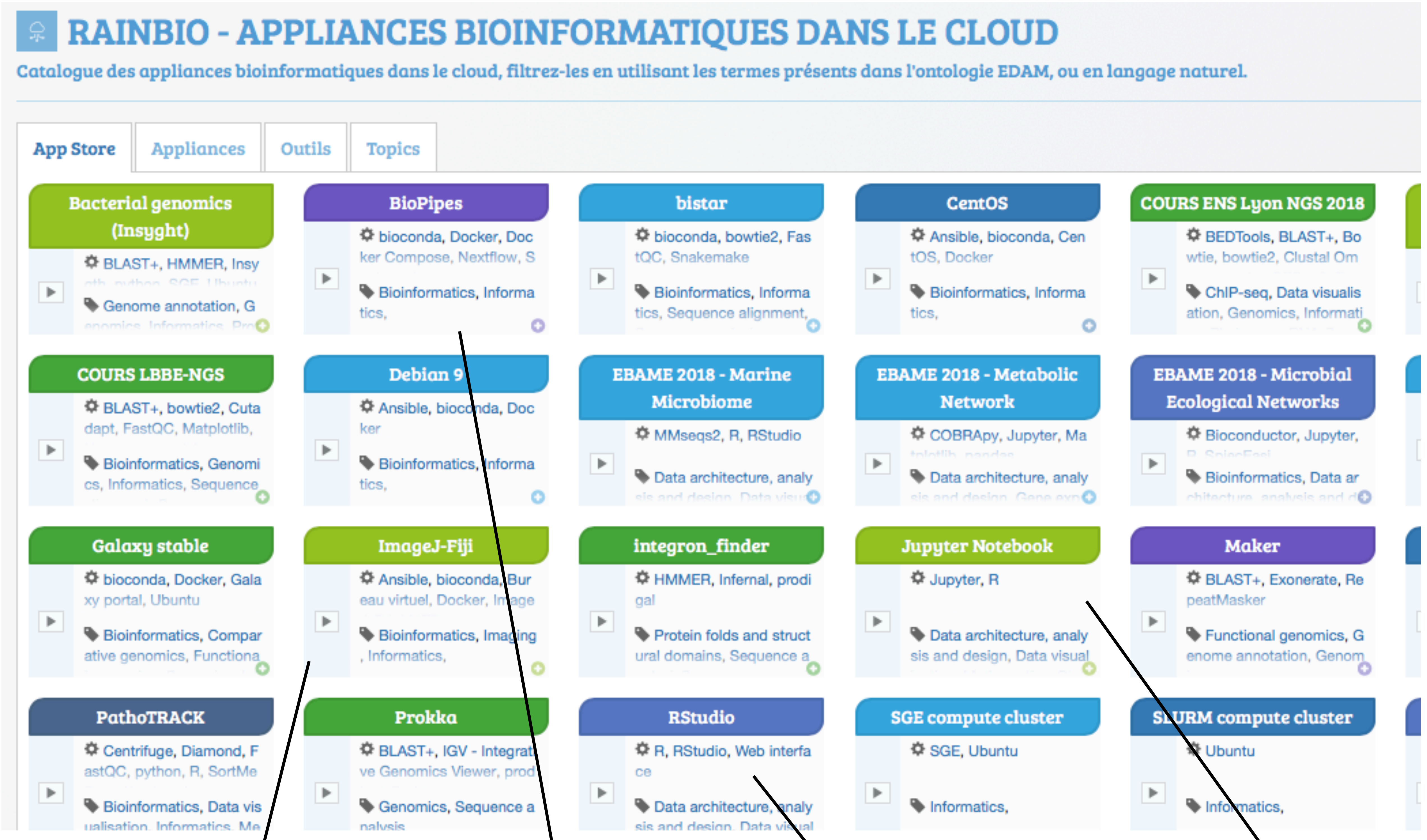
christophe.blanchet@france-bioinformatique.fr (eduGAIN)

☐		ID	Nom		Début	Groupes	Spéc				Brkr	Cloud	Accès	
☐	●	9451	BioPipes (19.04) reprohack-3	?	↑Nov 24 2019, 15h24	Reprohackathon-3		2	8	120	d5a3	ifb-core-cloud	ssh	
☐	●	9443	Ubuntu 16.04 (16.04.6)	?	↑Nov 21 2019, 19h05	IFB-core		1	4	70	4ed0	ifb-core-cloud	ssh	
☐	●	9421	Ubuntu 18.04 (18.04.2) icli	?	↑Nov 21 2019, 09h49	IFB-core		8	32	470	f25e	ifb-core-cloud	ssh	
☐	●	9386	Ubuntu 18.04 (18.04.2) lega	?	↑Nov 19 2019, 10h25	IFB-core		16	64	920	dc7b	ifb-core-cloud	ssh	
☐	●	9373	Jupyter	?	↑Nov 15 2019, 14h45	IFB-core		1	4	70	e878	ifb-core-cloud	https Params	

1. Site IFB
2. Accueil Biosphère
3. Catalogue des appliances
4. Tableau de bord
5. Données
6. Informations utiles
7. Paramètres utilisateur

RAINBio - Catalogue des Apps Bioinformatiques

- Appliances, des environnements prêts-à-l'emploi
  - Déploiement en 1 clic
  - 1 à N machines virtuelles, multi-cloud
  - Interfaces usuelles: SSH, HTTP, Desktop
  - Définie par les développeurs
  - Adaptables par les utilisateurs
- Apps de base
  - Outils génériques : linux, workflows, Rstudio/Shiny, Jupyter…
- Apps thématiques
  - Génomique (ChIP-seq, RNA-Seq…), annotation de génome, génomique comparative, bioimagerie, chemins métaboliques, écologie microbienne, métagénomique, transcriptomique…



RAINBIO - APPLIANCES BIOINFORMATIQUES DANS LE CLOUD

Catalogue des appliances bioinformatiques dans le cloud, filtrez-les en utilisant les termes présents dans l'ontologie EDAM, ou en langage naturel.

App Store | Appliances | Outils | Topics

Bacterial genomics (Insyght)
BLAST+, HMMER, Insy
Genome annotation, G

BioPipes
bioconda, Docker, Docker Compose, Nextflow, S
Bioinformatics, Informatics,

bistar
bioconda, bowtie2, FastQC, Snakemake
Bioinformatics, Informatics, Sequence alignment,

CentOS
Ansible, bioconda, CentOS, Docker
Bioinformatics, Informatics,

COURS ENS Lyon NGS 2018
BEDTools, BLAST+, Bowtie, bowtie2, Clustal Om
ChIP-seq, Data visualisation, Genomics, Informati

COURS LBBE-NGS
BLAST+, bowtie2, Cutadapt, FastQC, Matplotlib,
Bioinformatics, Genomics, Informatics, Sequence

Debian 9
Ansible, bioconda, Docker
Bioinformatics, Informatics,

EBAME 2018 - Marine Microbiome
MMseqs2, R, RStudio
Data architecture, analy

EBAME 2018 - Metabolic Network
COBRApy, Jupyter, Ma
Data architecture, analy

EBAME 2018 - Microbial Ecological Networks
Bioconductor, Jupyter,
Bioinformatics, Data ar

Galaxy stable
bioconda, Docker, Galaxy portal, Ubuntu
Bioinformatics, Comparative genomics, Functiona

ImageJ-Fiji
Ansible, bioconda, Bureau virtuel, Docker, Image
Bioinformatics, Imaging, Informatics,

integron_finder
HMMER, Infernal, prodigal
Protein folds and structural domains, Sequence a

Jupyter Notebook
Jupyter, R
Data architecture, analysis and design, Data visual

Maker
BLAST+, Exonerate, RepeatMasker
Functional genomics, Genome annotation, Genom

PathoTRACK
Centrifuge, Diamond, FastQC, python, R, SortMe
Bioinformatics, Data vis

Prokka
BLAST+, IGV - Integrative Genomics Viewer, prod
Genomics, Sequence a

RStudio
R, RStudio, Web interface
Data architecture, analy

SGE compute cluster
SGE, Ubuntu
Informatics,

SLURM compute cluster
Ubuntu
Informatics,

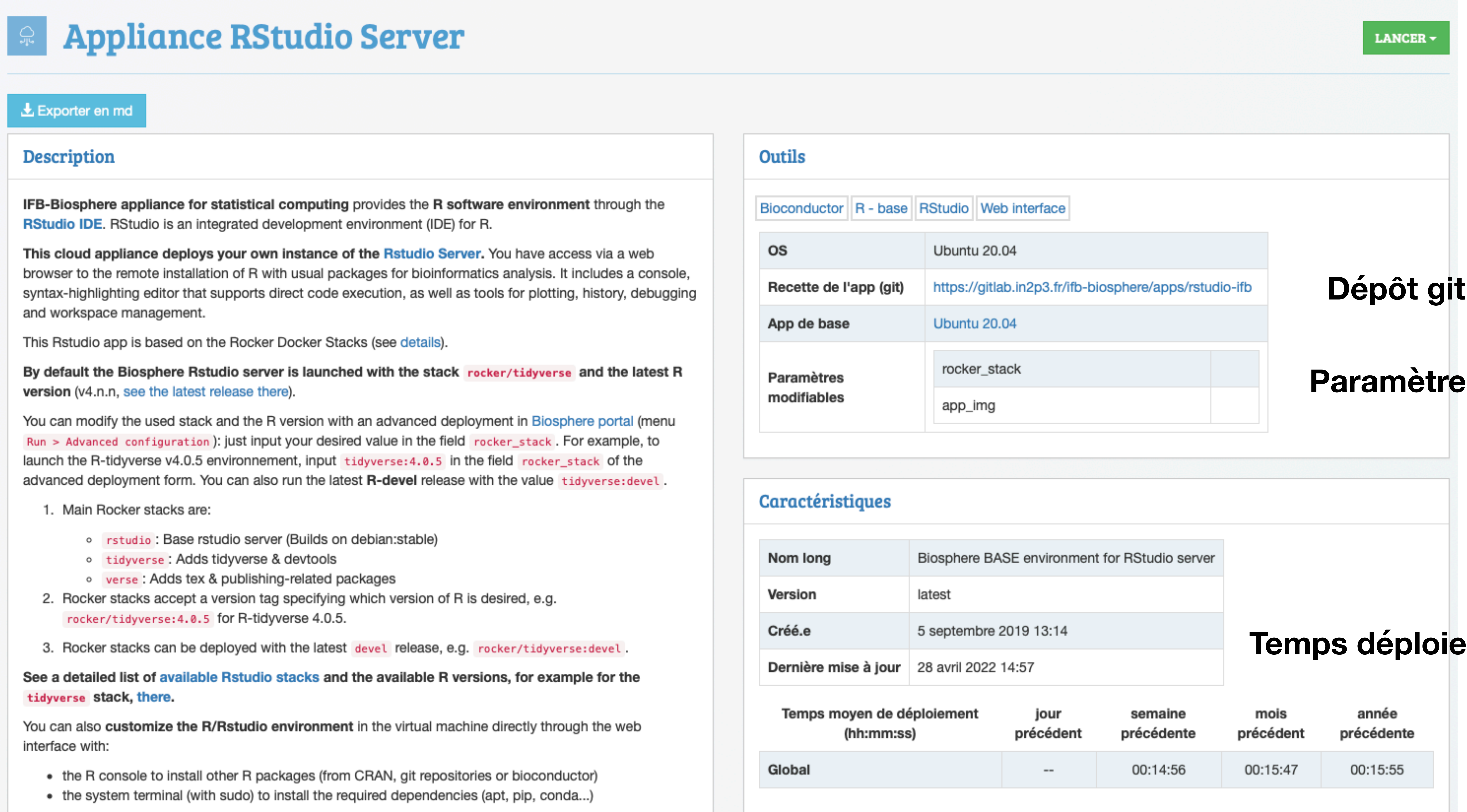
Workflows

IFB Cloud - Apps de base

- **Ubuntu/Debian, CentOS/RockyLinux**
  - ★ Docker +Compose, Ansible(roles), Bioconda
  - ★ Desktop (XFCE)
- **BioPipes**
  - ★ Nextflow, SnakeMake, CWLtool, Conda + Bioconda, Docker + Compose
- **R, Rstudio, Shiny**
  - ★ stacks tidyverse, verse, ML/cuda, geospatial…
- **Jupyter Notebook**
  - ★ kernels Python, R, Julia
  - ★ stacks notebook, R, scipy, datascience, tensor flow…
- **Bioimage**
  - ★ ImageJ-Fiji, Icy
  - ★ connecteur OMERO
- **Cytoscape**
  - ★ connecteurs NDexBio, PSICQUIC
- **Clusters** SGE, Slurm…

Maintenues par IFB-core
Infrastructure as Code
https://gitlab.in2p3.fr/ifb-biosphere/apps

App Rstudio



Appliance RStudio Server

Description

IFB-Biosphere appliance for statistical computing provides the **R software environment** through the **RStudio IDE**. RStudio is an integrated development environment (IDE) for R.

This cloud appliance deploys your own instance of the Rstudio Server. You have access via a web browser to the remote installation of R with usual packages for bioinformatics analysis. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management.

This Rstudio app is based on the Rocker Docker Stacks (see details).

By default the Biosphere Rstudio server is launched with the stack `rocker/tidyverse` and the latest R version (v4.n.n, see the latest release there).

You can modify the used stack and the R version with an advanced deployment in Biosphere portal (menu `Run > Advanced configuration`): just input your desired value in the field `rocker_stack`. For example, to launch the R-tidyverse v4.0.5 environnement, input `tidyverse:4.0.5` in the field `rocker_stack` of the advanced deployment form. You can also run the latest **R-devel** release with the value `tidyverse:devel`.

1. Main Rocker stacks are:
    - `rstudio` : Base rstudio server (Builds on debian:stable)
    - `tidyverse` : Adds tidyverse & devtools
    - `verse` : Adds tex & publishing-related packages
2. Rocker stacks accept a version tag specifying which version of R is desired, e.g. `rocker/tidyverse:4.0.5` for R-tidyverse 4.0.5.
3. Rocker stacks can be deployed with the latest `devel` release, e.g. `rocker/tidyverse:devel`.

See a detailed list of available Rstudio stacks and the available R versions, for example for the `tidyverse` stack, there.

You can also **customize the R/Rstudio environment** in the virtual machine directly through the web interface with:

- the R console to install other R packages (from CRAN, git repositories or bioconductor)
- the system terminal (with sudo) to install the required dependencies (apt, pip, conda...)

Outils

Bioconductor | R - base | RStudio | Web interface

OS	Ubuntu 20.04
Recette de l'app (git)	https://gitlab.in2p3.fr/ifb-biosphere/apps/rstudio-ifb
App de base	Ubuntu 20.04
Paramètres modifiables	rocker_stack app_img

Caractéristiques

Nom long	Biosphere BASE environment for RStudio server
Version	latest
Créé.e	5 septembre 2019 13:14
Dernière mise à jour	28 avril 2022 14:57

Temps moyen de déploiement (hh:mm:ss)	jour précédent	semaine précédente	mois précédent	année précédente
Global	--	00:14:56	00:15:47	00:15:55

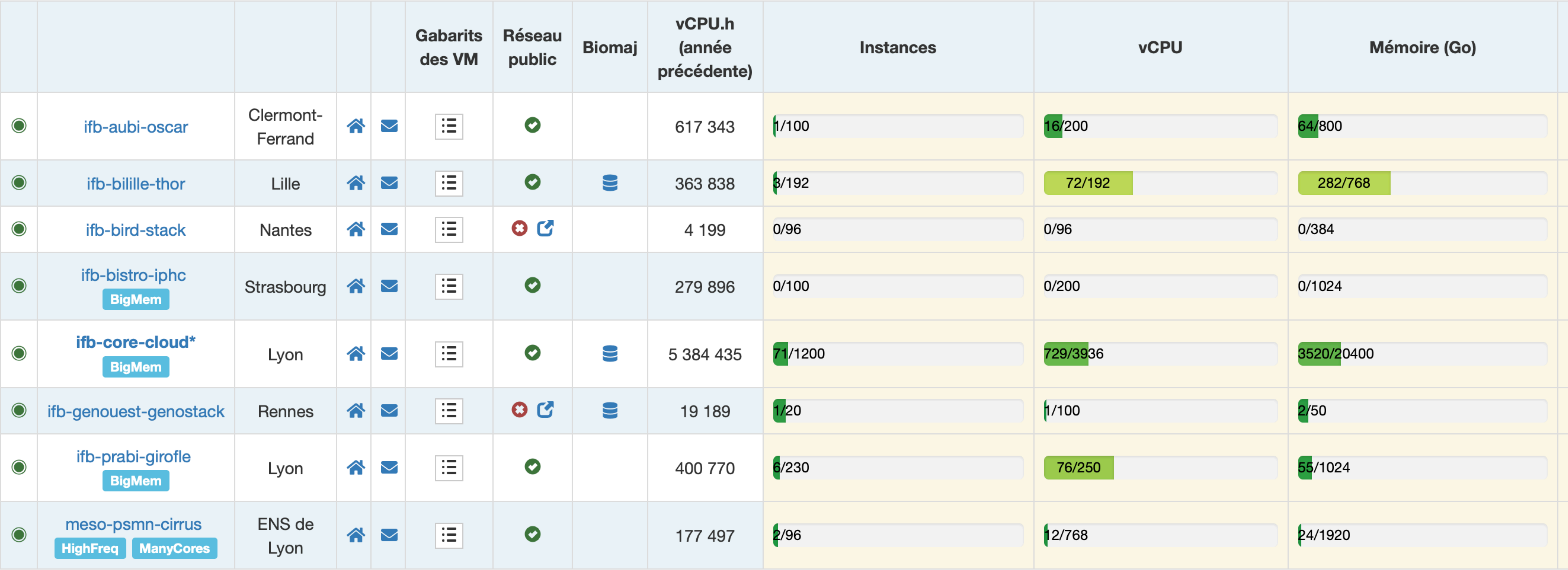
Dépôt git

Paramètres

Stacks

Temps déploiement

Ressources informatiques du cloud IFB



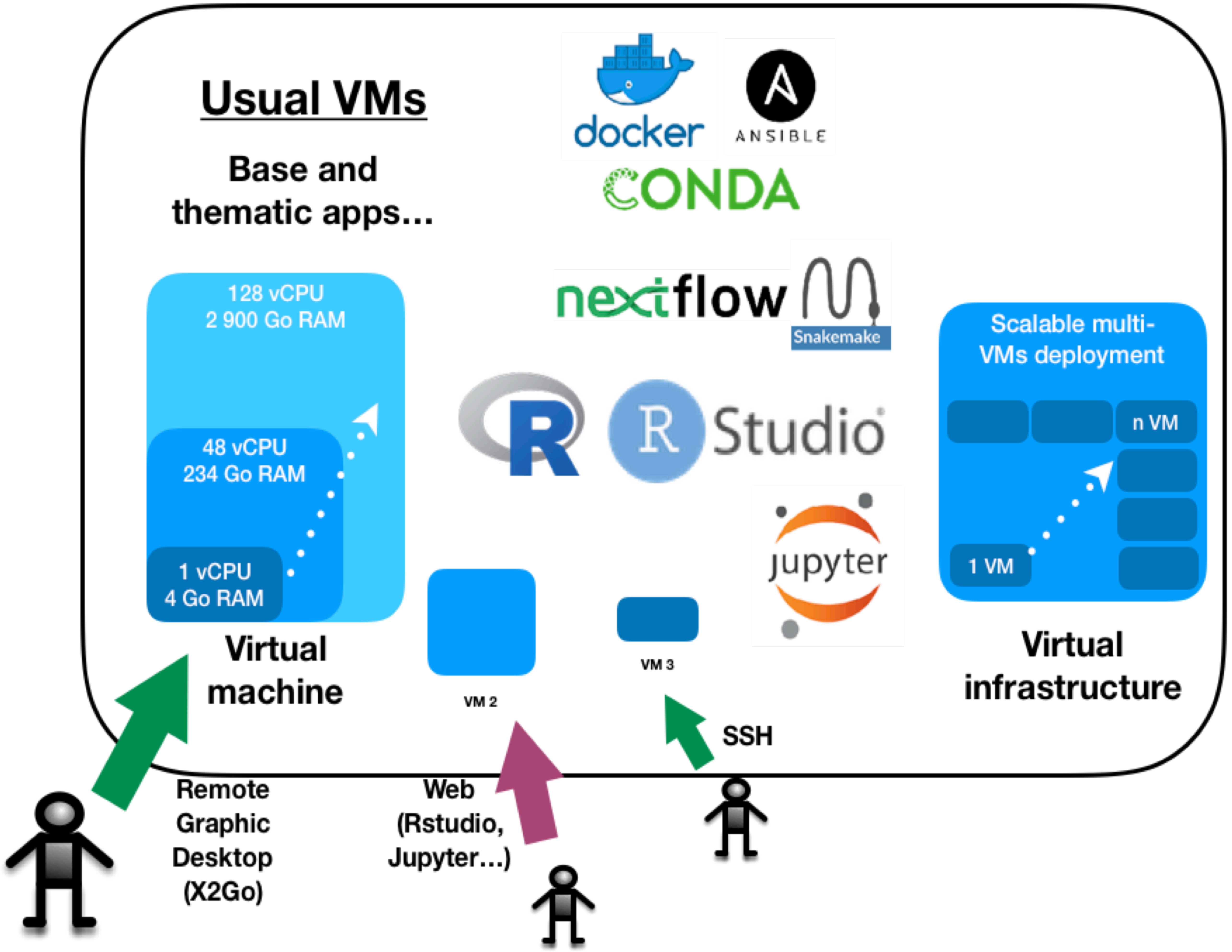
					Gabarits des VM	Réseau public	Biomaj	vCPU.h (année précédente)	Instances	vCPU	Mémoire (Go)
◉	ifb-aubi-oscar	Clermont-Ferrand	🏠	✉	☰	✓		617 343	1/100	16/200	64/800
◉	ifb-bilille-thor	Lille	🏠	✉	☰	✓	🛢	363 838	3/192	72/192	282/768
◉	ifb-bird-stack	Nantes	🏠	✉	☰	✗ ↗		4 199	0/96	0/96	0/384
◉	ifb-bistro-iphc BigMem	Strasbourg	🏠	✉	☰	✓		279 896	0/100	0/200	0/1024
◉	**ifb-core-cloud*** BigMem	Lyon	🏠	✉	☰	✓	🛢	5 384 435	71/1200	729/3936	3520/20400
◉	ifb-genouest-genostack	Rennes	🏠	✉	☰	✗ ↗	🛢	19 189	1/20	1/100	2/50
◉	ifb-prabi-girofle BigMem	Lyon	🏠	✉	☰	✓		400 770	6/230	76/250	55/1024
◉	meso-psmn-cirrus HighFreq ManyCores	ENS de Lyon	🏠	✉	☰	✓		177 497	2/96	12/768	24/1920

*Cloud par défaut

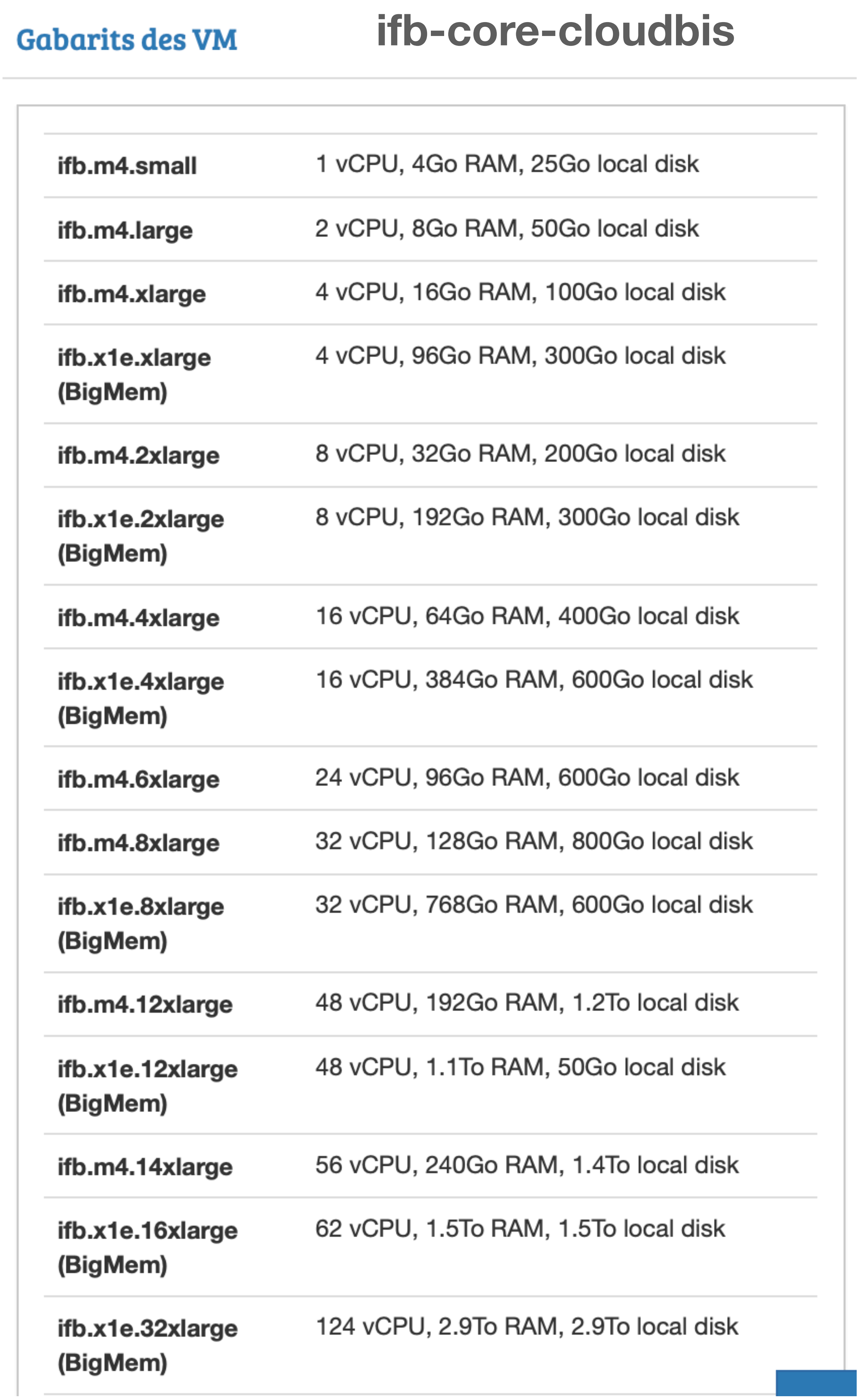
https://biosphere.france-bioinformatique.fr/cloud/system_status

Cloud Modular Environments

- **Predefined environments**
  - Base apps and thematic ones
- **Standard virtual machine**
  - Single VM to bunch of VMs.
  - up to 64 vCPUs-250 GB RAM.
- **Specific VM**
  - **BigMemory**: up to 3 TB RAM.
  - **HighFrequency** : up to 3.8GHz
  - **ManyCores**: up to 255 cores in a VM
  - **GPU** (to come)
- **Bring-Your-Own-Tools (BYOT)**:
  - Admin rights in the VM,
  - Deploy with your own container image,
  - Configure the environment with your own git repository (Infrastructure as Code).

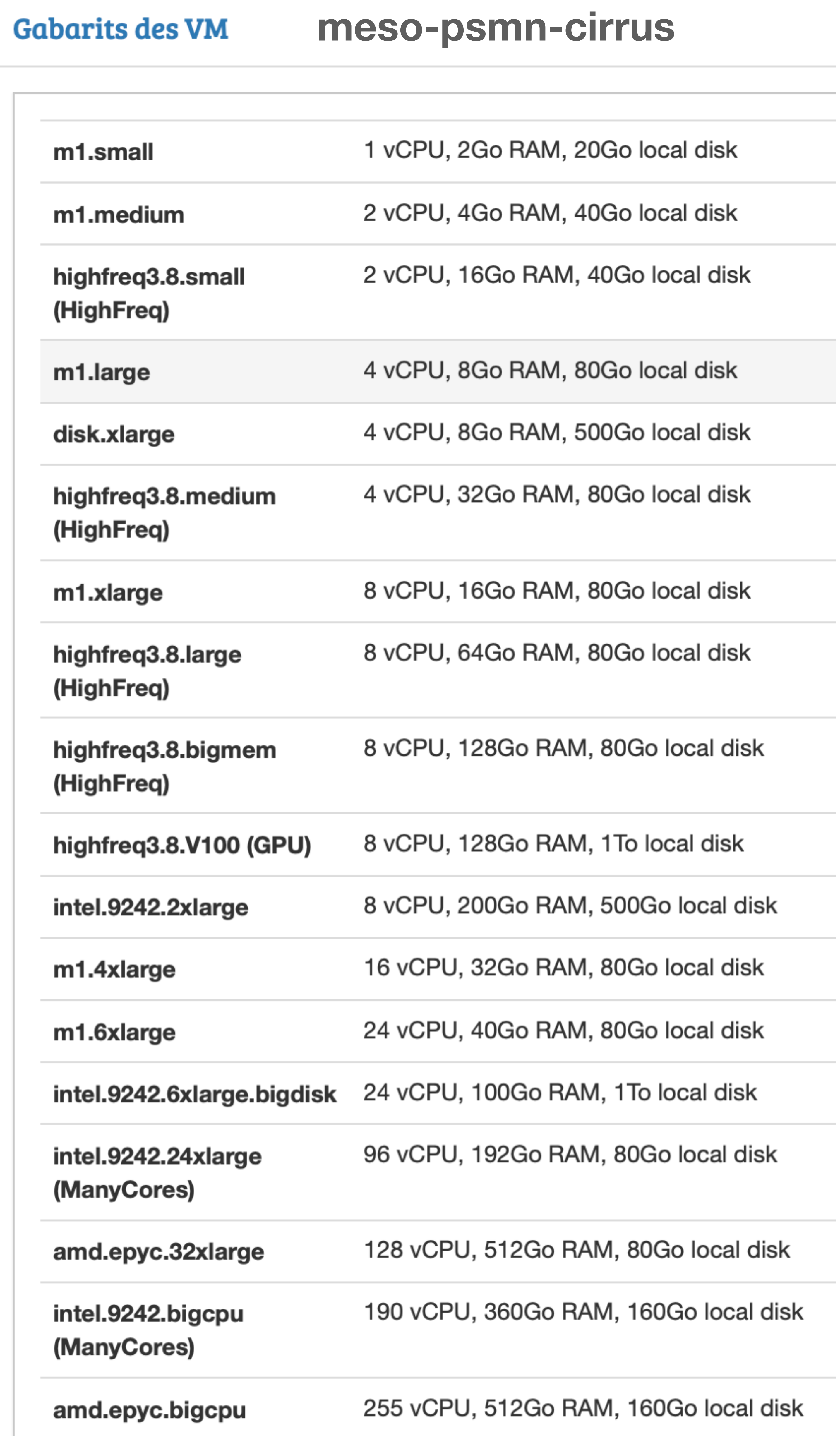


Exemples de Gabarits



Gabarits des VM ifb-core-cloudbis

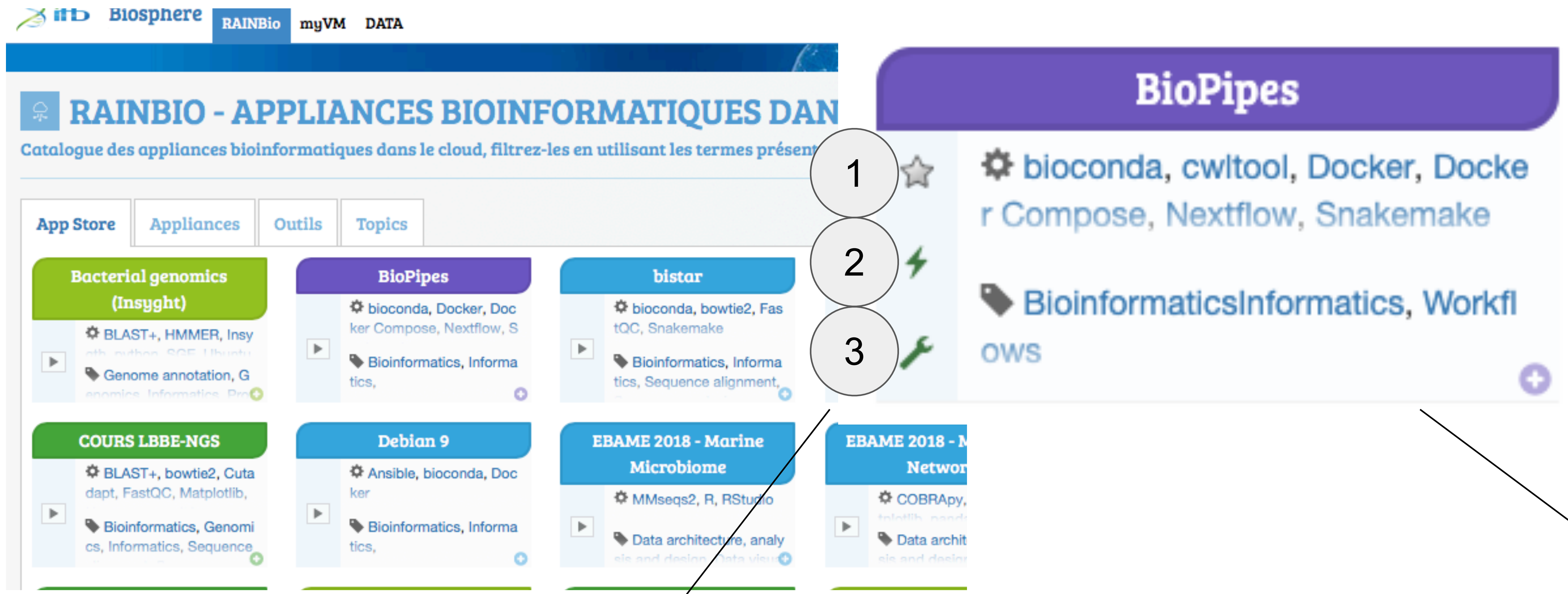
ifb.m4.small	1 vCPU, 4Go RAM, 25Go local disk
ifb.m4.large	2 vCPU, 8Go RAM, 50Go local disk
ifb.m4.xlarge	4 vCPU, 16Go RAM, 100Go local disk
ifb.x1e.xlarge (BigMem)	4 vCPU, 96Go RAM, 300Go local disk
ifb.m4.2xlarge	8 vCPU, 32Go RAM, 200Go local disk
ifb.x1e.2xlarge (BigMem)	8 vCPU, 192Go RAM, 300Go local disk
ifb.m4.4xlarge	16 vCPU, 64Go RAM, 400Go local disk
ifb.x1e.4xlarge (BigMem)	16 vCPU, 384Go RAM, 600Go local disk
ifb.m4.6xlarge	24 vCPU, 96Go RAM, 600Go local disk
ifb.m4.8xlarge	32 vCPU, 128Go RAM, 800Go local disk
ifb.x1e.8xlarge (BigMem)	32 vCPU, 768Go RAM, 600Go local disk
ifb.m4.12xlarge	48 vCPU, 192Go RAM, 1.2To local disk
ifb.x1e.12xlarge (BigMem)	48 vCPU, 1.1To RAM, 50Go local disk
ifb.m4.14xlarge	56 vCPU, 240Go RAM, 1.4To local disk
ifb.x1e.16xlarge (BigMem)	62 vCPU, 1.5To RAM, 1.5To local disk
ifb.x1e.32xlarge (BigMem)	124 vCPU, 2.9To RAM, 2.9To local disk



Gabarits des VM meso-psmn-cirrus

m1.small	1 vCPU, 2Go RAM, 20Go local disk
m1.medium	2 vCPU, 4Go RAM, 40Go local disk
highfreq3.8.small (HighFreq)	2 vCPU, 16Go RAM, 40Go local disk
m1.large	4 vCPU, 8Go RAM, 80Go local disk
disk.xlarge	4 vCPU, 8Go RAM, 500Go local disk
highfreq3.8.medium (HighFreq)	4 vCPU, 32Go RAM, 80Go local disk
m1.xlarge	8 vCPU, 16Go RAM, 80Go local disk
highfreq3.8.large (HighFreq)	8 vCPU, 64Go RAM, 80Go local disk
highfreq3.8.bigmem (HighFreq)	8 vCPU, 128Go RAM, 80Go local disk
highfreq3.8.V100 (GPU)	8 vCPU, 128Go RAM, 1To local disk
intel.9242.2xlarge	8 vCPU, 200Go RAM, 500Go local disk
m1.4xlarge	16 vCPU, 32Go RAM, 80Go local disk
m1.6xlarge	24 vCPU, 40Go RAM, 80Go local disk
intel.9242.6xlarge.bigdisk	24 vCPU, 100Go RAM, 1To local disk
intel.9242.24xlarge (ManyCores)	96 vCPU, 192Go RAM, 80Go local disk
amd.epyc.32xlarge	128 vCPU, 512Go RAM, 80Go local disk
intel.9242.bigcpu (ManyCores)	190 vCPU, 360Go RAM, 160Go local disk
amd.epyc.bigcpu	255 vCPU, 512Go RAM, 160Go local disk

Déployer une App



RAINBIO - APPLIANCES BIOINFORMATIQUES DAN

Catalogue des appliances bioinformatiques dans le cloud, filtrez-les en utilisant les termes présent

Bacterial genomics (Insyght) – BLAST+, HMMER, Insy – Genome annotation, G

BioPipes – bioconda, Docker, Docker Compose, Nextflow, S – Bioinformatics, Informatics,

bistar – bioconda, bowtie2, FastQC, Snakemake – Bioinformatics, Informatics, Sequence alignment,

COURS LBBE-NGS – BLAST+, bowtie2, Cutadapt, FastQC, Matplotlib, – Bioinformatics, Genomics, Informatics, Sequence

Debian 9 – Ansible, bioconda, Docker – Bioinformatics, Informatics,

EBAME 2018 - Marine Microbiome – MMseqs2, R, RStudio – Data architecture, analy

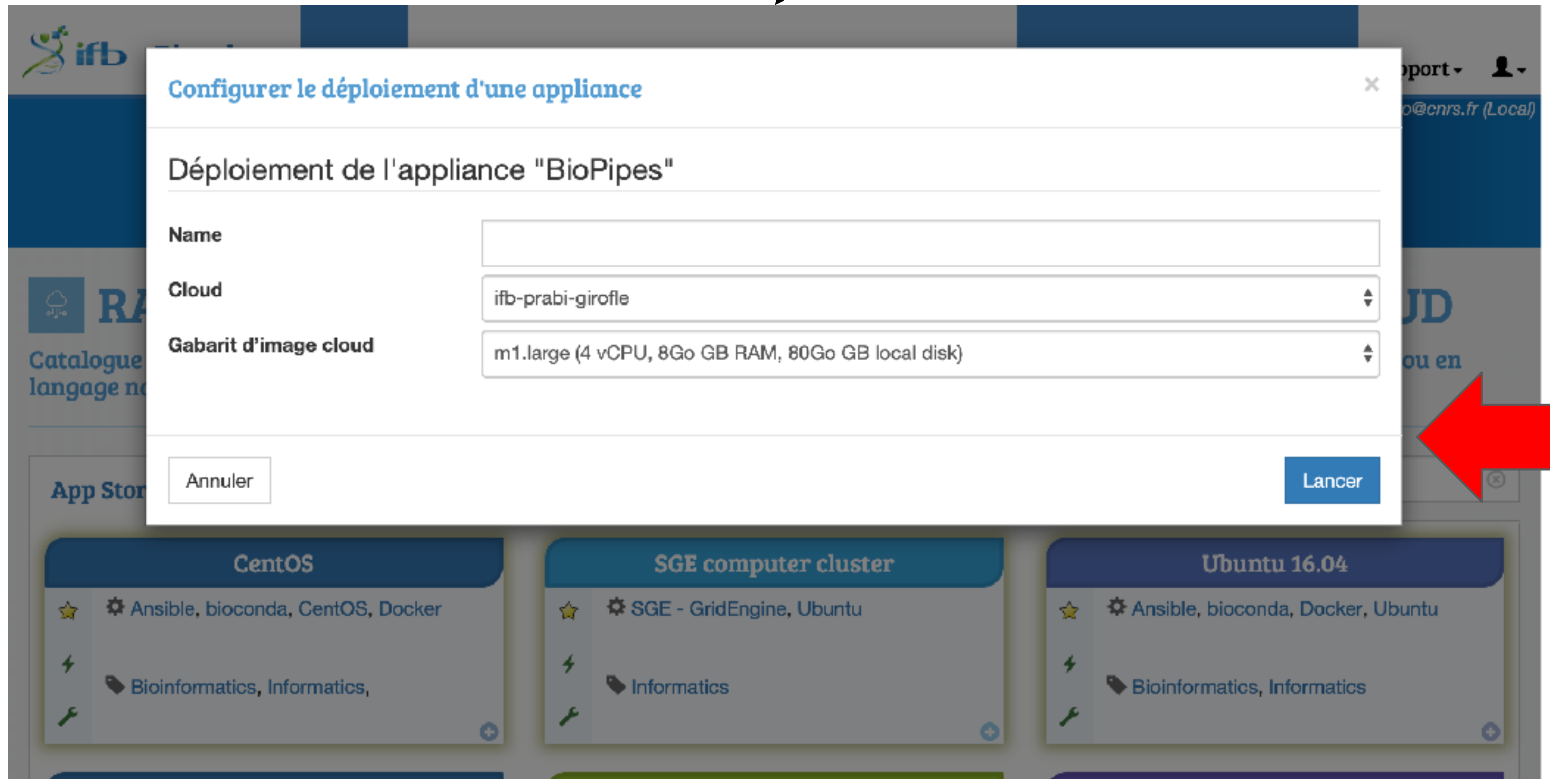
EBAME 2018 - Networ – COBRApy, – Data archit

BioPipes

bioconda, cwltool, Docker, Docker Compose, Nextflow, Snakemake

BioinformaticsInformatics, Workflows

1. Mise en favori
2. Lancement rapide
3. Déploiement avancé



Configurer le déploiement d'une appliance

Déploiement de l'appliance "BioPipes"

Name:

Cloud: ifb-prabi-girofle

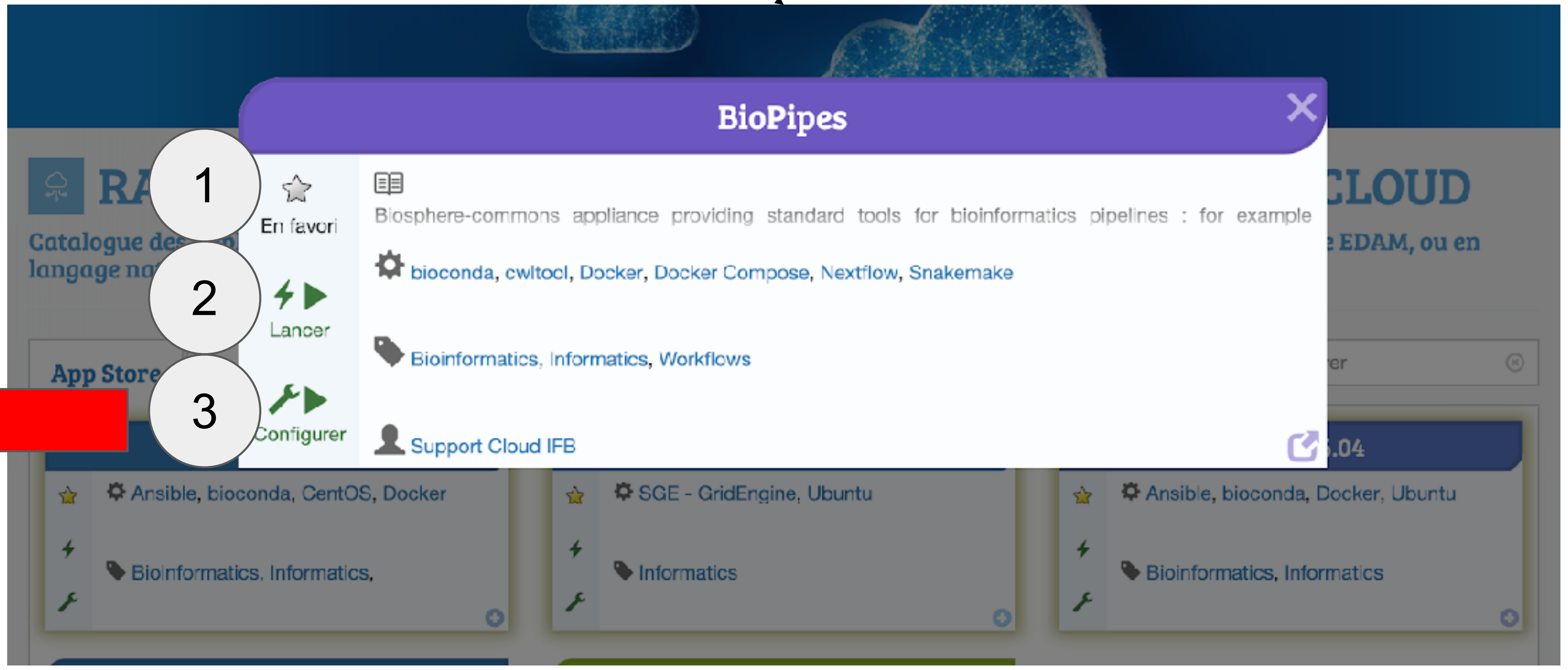
Gabarit d'image cloud: m1.large (4 vCPU, 8Go GB RAM, 80Go GB local disk)

Annuler | Lancer

CentOS – Ansible, bioconda, CentOS, Docker – Bioinformatics, Informatics,

SGE computer cluster – SGE - GridEngine, Ubuntu – Informatics

Ubuntu 16.04 – Ansible, bioconda, Docker, Ubuntu – Bioinformatics, Informatics



BioPipes

Biosphere-commons appliance providing standard tools for bioinformatics pipelines : for example

bioconda, cwltool, Docker, Docker Compose, Nextflow, Snakemake

Bioinformatics, Informatics, Workflows

Support Cloud IFB

Ansible, bioconda, CentOS, Docker – Bioinformatics, Informatics,

SGE - GridEngine, Ubuntu – Informatics

Ansible, bioconda, Docker, Ubuntu – Bioinformatics, Informatics

Pratique

Déploiement d'une VM

- choix taille vm
- choix site > page statut système
- temps déploiement moyen app > page de l'app
- choix stack rstudio/jupyter (galaxy)
- choix conteneur externe (rstudio, shiny, jupyter)

Se Connecter à une VM

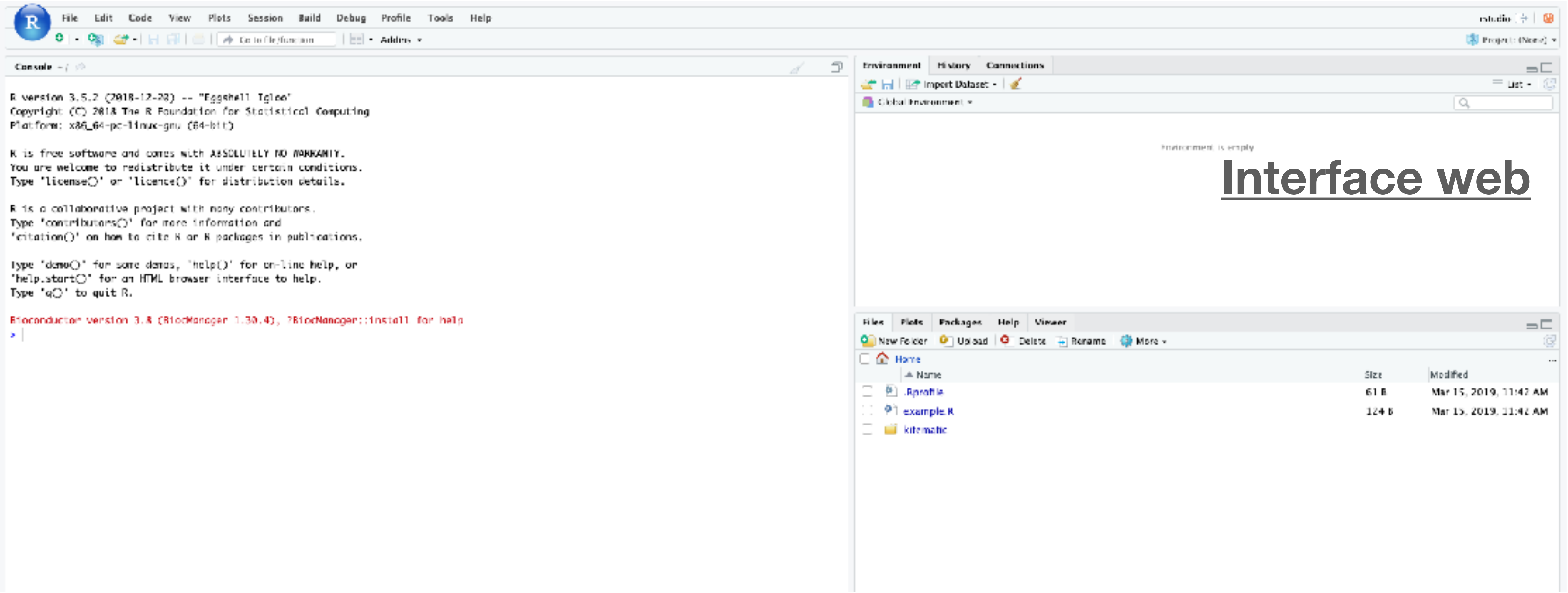
Accès

ssh Params

ssh

ssh https

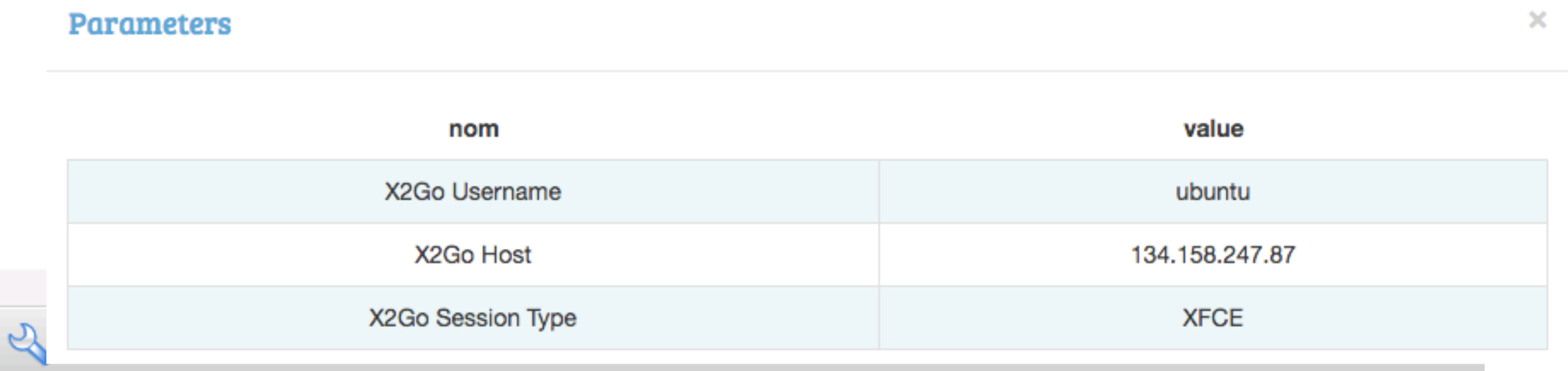
HTTP/S



```
R version 3.5.2 (2018-12-20) -- "Eggshell Igloo"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Bioconductor version 3.8 (BiocManager 1.30.4), ?BiocManager::install for help
>
```

Interface web

Bureau X2Go



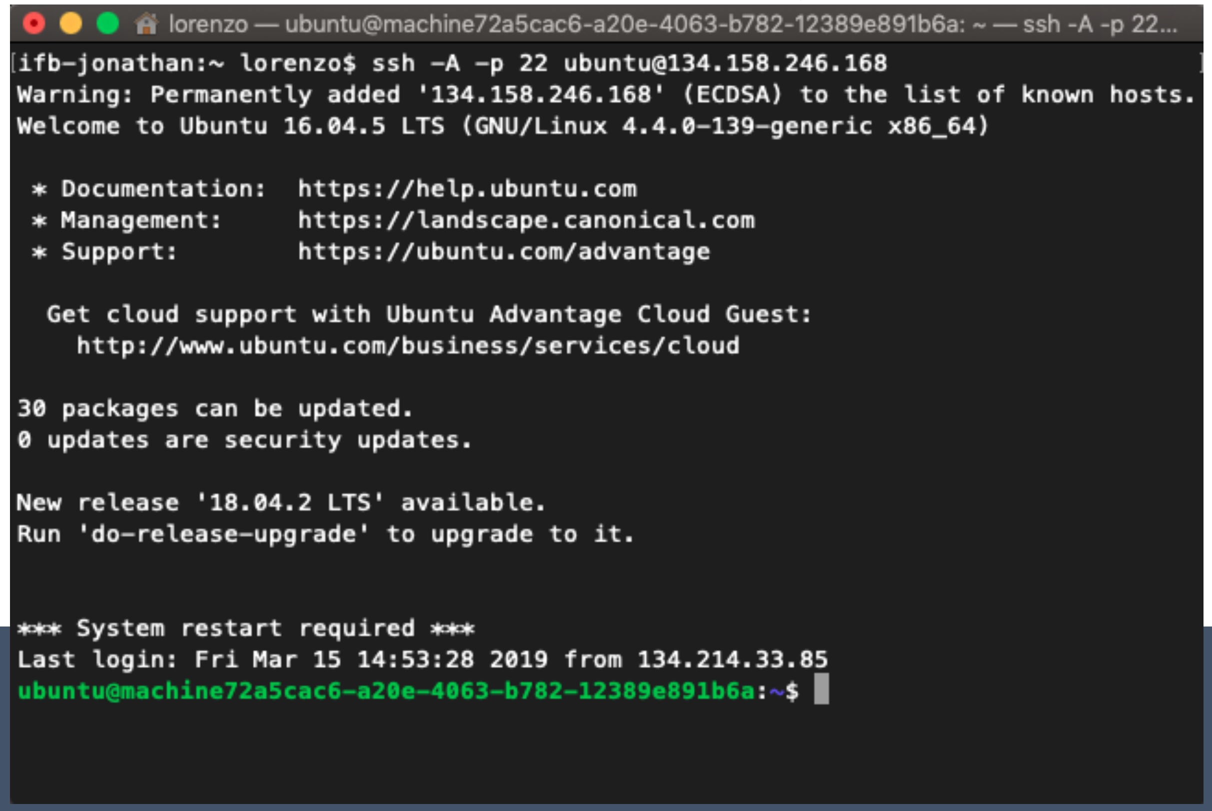
Parameters

nom	value
X2Go Username	ubuntu
X2Go Host	134.158.247.87
X2Go Session Type	XFCE

SSH

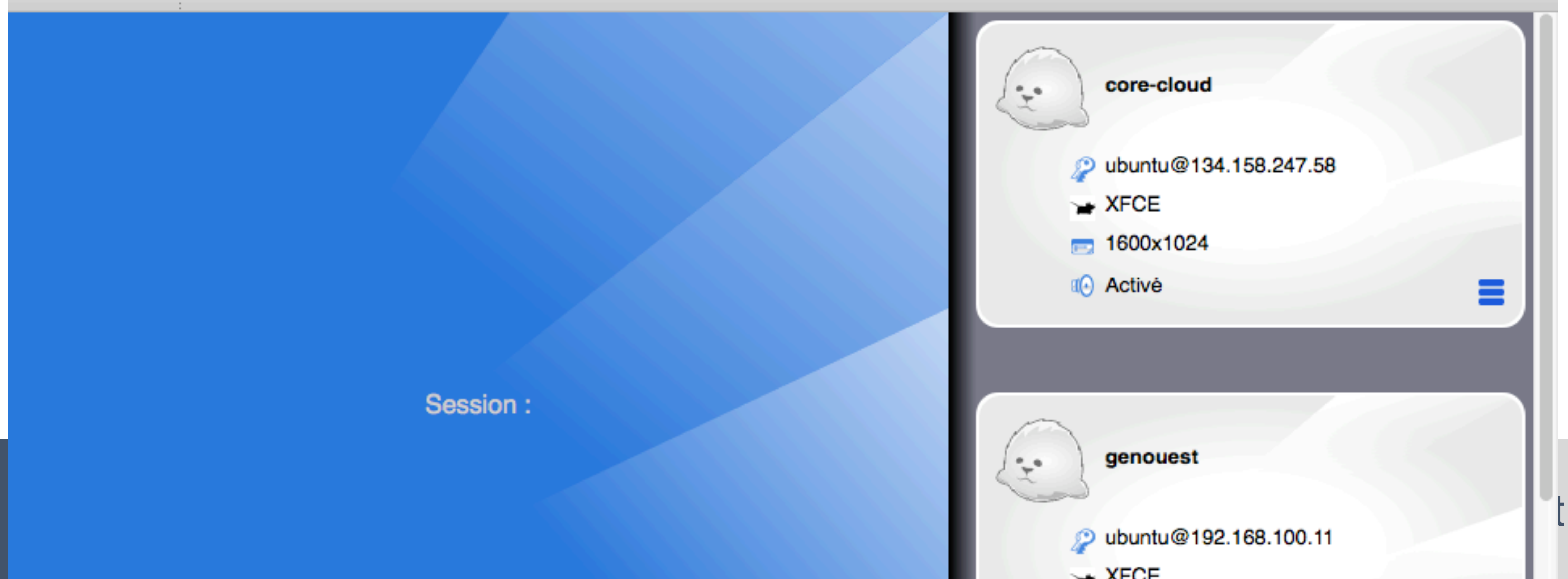
+ SSH PubKey

Terminal



lorenzo — ubuntu@machine72a5cac6-a20e-4063-b782-12389e891b6a: ~ — ssh -A -p 22...

```
ifb-jonathan:~ lorenzo$ ssh -A -p 22 ubuntu@134.158.246.168
Warning: Permanently added '134.158.246.168' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 16.04.5 LTS (GNU/Linux 4.4.0-139-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/advantage

  Get cloud support with Ubuntu Advantage Cloud Guest:
    http://www.ubuntu.com/business/services/cloud

30 packages can be updated.
0 updates are security updates.

New release '18.04.2 LTS' available.
Run 'do-release-upgrade' to upgrade to it.


*** System restart required ***
Last login: Fri Mar 15 14:53:28 2019 from 134.214.33.85
ubuntu@machine72a5cac6-a20e-4063-b782-12389e891b6a:~$ 
```



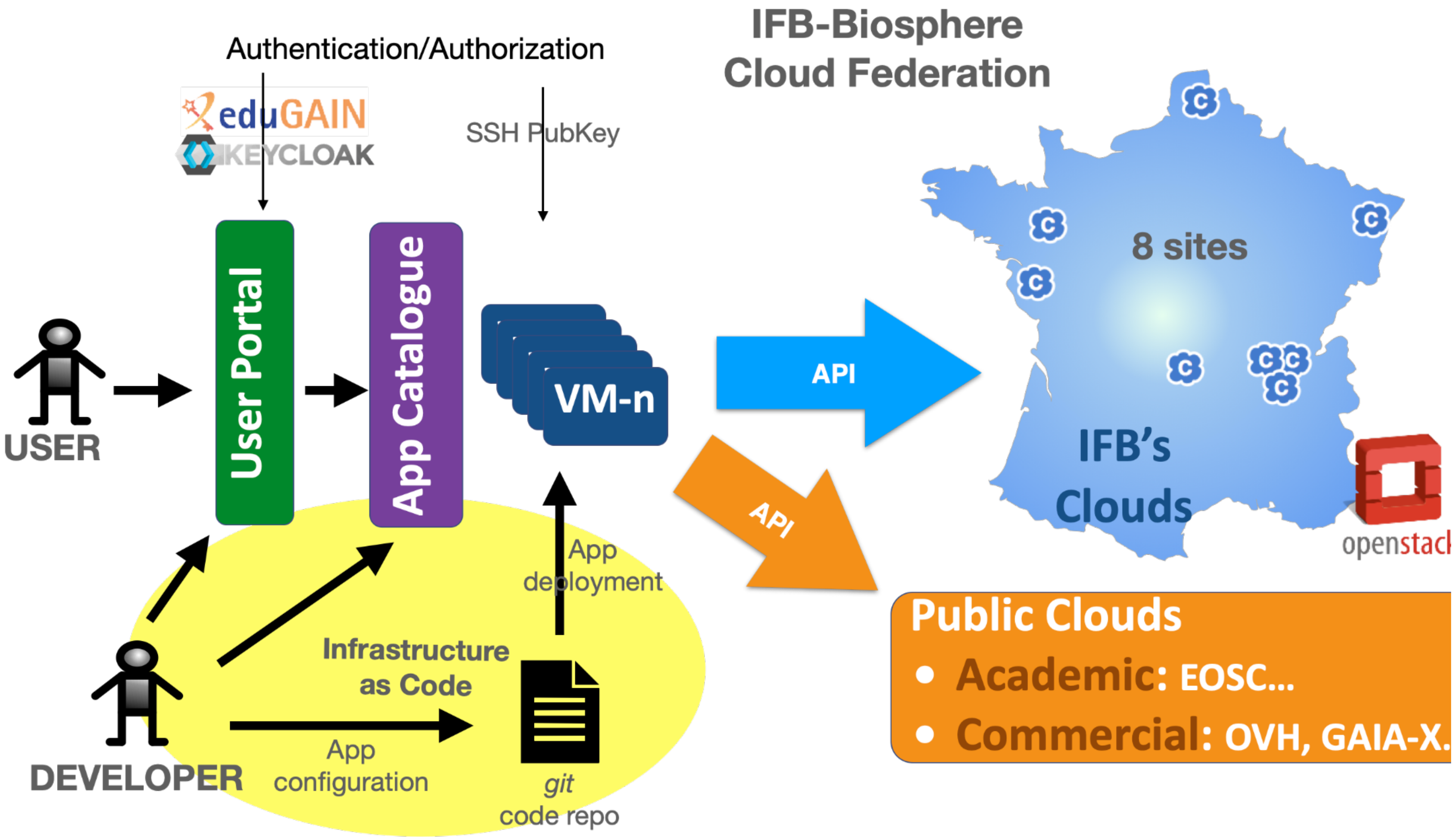
Session :

core-cloud
ubuntu@134.158.247.58
XFCE
1600x1024
Activé

genouest
ubuntu@192.168.100.11
XFCE

Bureau graphique à distance

Configuration des apps clouds



Authentication/Authorization

eduGAIN

KEYCLOAK

SSH PubKey

IFB-Biosphere
Cloud Federation

USER

User Portal

App Catalogue

VM-n

API

API

8 sites

IFB's
Clouds

openstack

App
deployment

Infrastructure
as Code

DEVELOPER

App
configuration

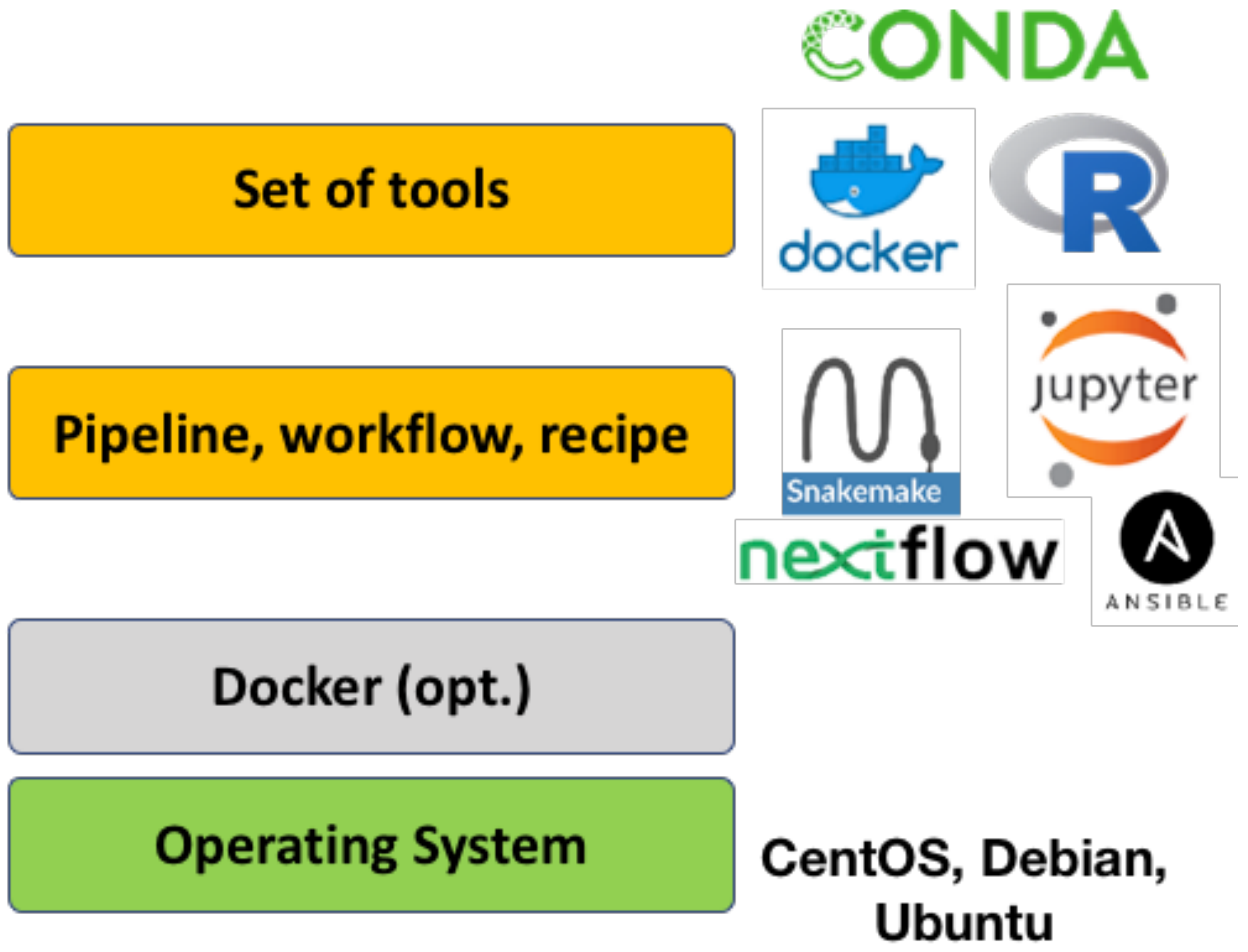
git
code repo

Public Clouds

- **Academic:** EOSC…
- **Commercial:** OVH, GAIA-X.

Recettes de déploiement

- Dépôts git, integration continue
- Conteneurs docker, référencés dans des catalogues
- Workflows, avec docker
- Jupyter Notebooks
- Scripts R…



CONDA

Set of tools

docker

R

Pipeline, workflow, recipe

Snakemake

jupyter

nextflow

ANSIBLE

Docker (opt.)

Operating System

CentOS, Debian,
Ubuntu

Virtualisation

Virtual machines

1 … N

Web

SSH

Application

Operating System

Hardware

CPU RAM Stor. Netw.

Physical server

App **App** Web

OS **OS** SSH

Virtual HW **Virtual HW**

Hypervisor

Hardware

CPU RAM Stor. Netw.

VM-based Model

App Web

App **Docker**

OS **OS** SSH

Virtual HW **Virtual HW**

Hypervisor

Hardware

CPU RAM Stor. Netw.

Hybrid VM-Container Model

Configuration d’une app

Configuration dans un dépôt ‘git’

- script shell
- environnement conda
- environnement docker compose
- playbook ansible
- script R

Conteneur externe (Docker Hub ou autres registres)

App Rstudio

- config R : ‘install.R’
- dépendances : ‘apt.txt’

App Shiny

- SHINY_REPO (in): Public Git URL of the shiny app repository
- SHINY_REPODIR (in): Directory in the git repo containing the app files (server.R, ui.R)

App workflow

- Jupyter notebook
- nextflow, snakemake, CWL (+docker/conda)

App basée sur un dépôt ‘git’

- Description
  - ★ `README.md`
- Installation
  - ★ `install.sh`
- Déploiement
  - ★ `deploy.sh`
- Intégration continue
  - ★ `Dockerfile`
  - ★ `.git-ci.yaml`
- Utilitaires : configuration et templates
  - ★ `ansible playbook & role`
  - ★ `miniconda-install.sh`
  - ★ `profile-ifb.sh`
  - ★ `docker compose`
  - ★ `nginx/traefik`
  - ★ `apt.txt`
  - ★ `install.R`...

Exemples de dépôts

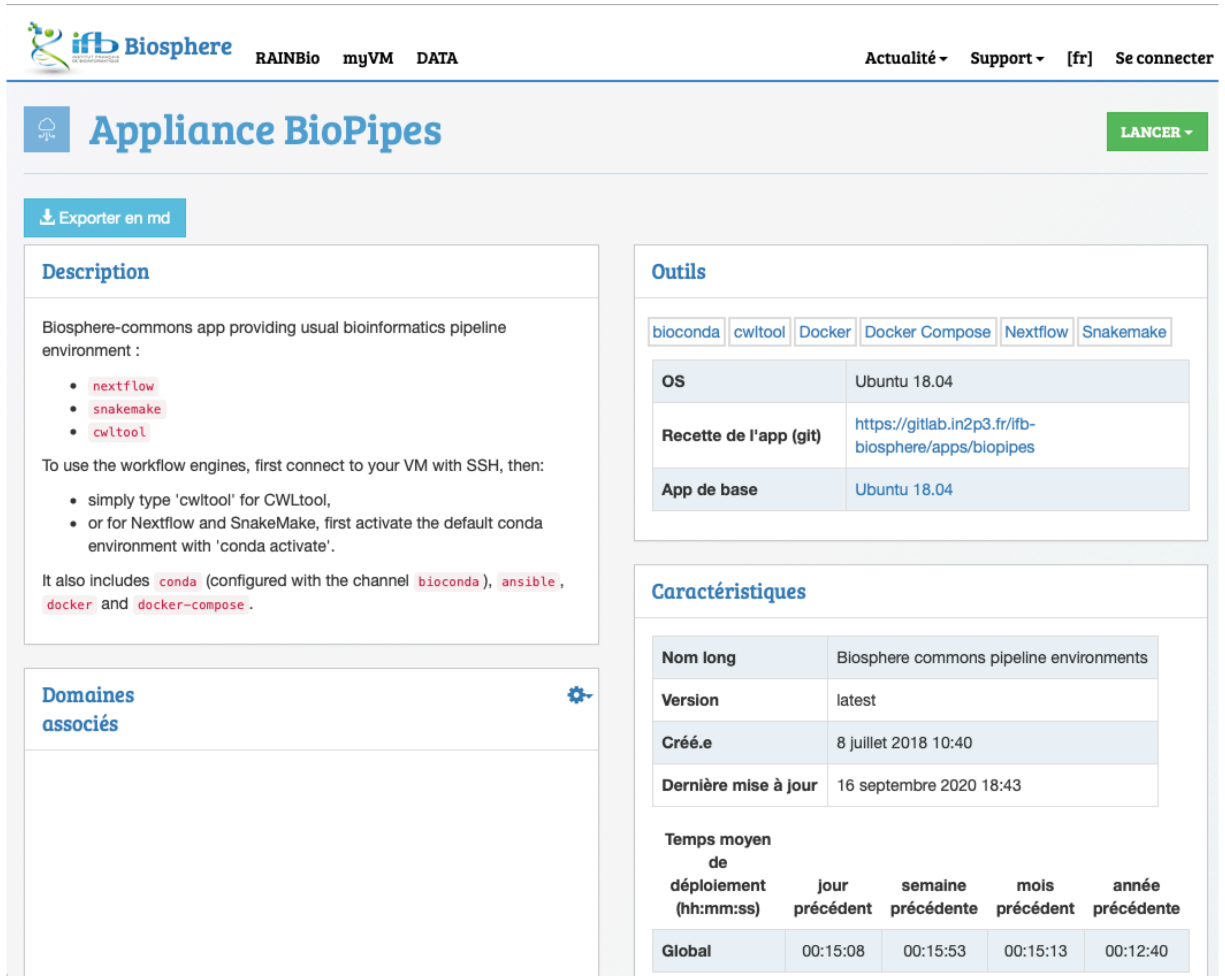
- **Biosphere Apps**
  - ★ https://gitlab.in2p3.fr/ifb-biosphere/apps
- ANF Metabiodiv
  - ★ https://gitlab.in2p3.fr/christophe.blanchet/app-metabiodiv
- Ecole EBAME
- Genopole 2019
  - ★ https://github.com/IFB-ElixirFr/biosphere-genopole-genome-variation
- Integron Finder
  - ★ https://github.com/C3BI-pasteur-fr/IFB-playbook

```
1) git clone …
   /ifb/apprepo/app_name
2) install.sh
3) deploy.sh
```

Apps Biosphère - Infastructures as Code

BioPipes

- La configuration de la VM est basée sur une recette stockée dans un dépôt *git*



Appliance BioPipes

Description

Biosphere-commons app providing usual bioinformatics pipeline environment :

- `nextflow`
- `snakemake`
- `cwltool`

To use the workflow engines, first connect to your VM with SSH, then:

- simply type 'cwltool' for CWLtool,
- or for Nextflow and SnakeMake, first activate the default conda environment with 'conda activate'.

It also includes `conda` (configured with the channel `bioconda`), `ansible`, `docker` and `docker-compose`.

Domaines associés

Outils

bioconda | cwltool | Docker | Docker Compose | Nextflow | Snakemake

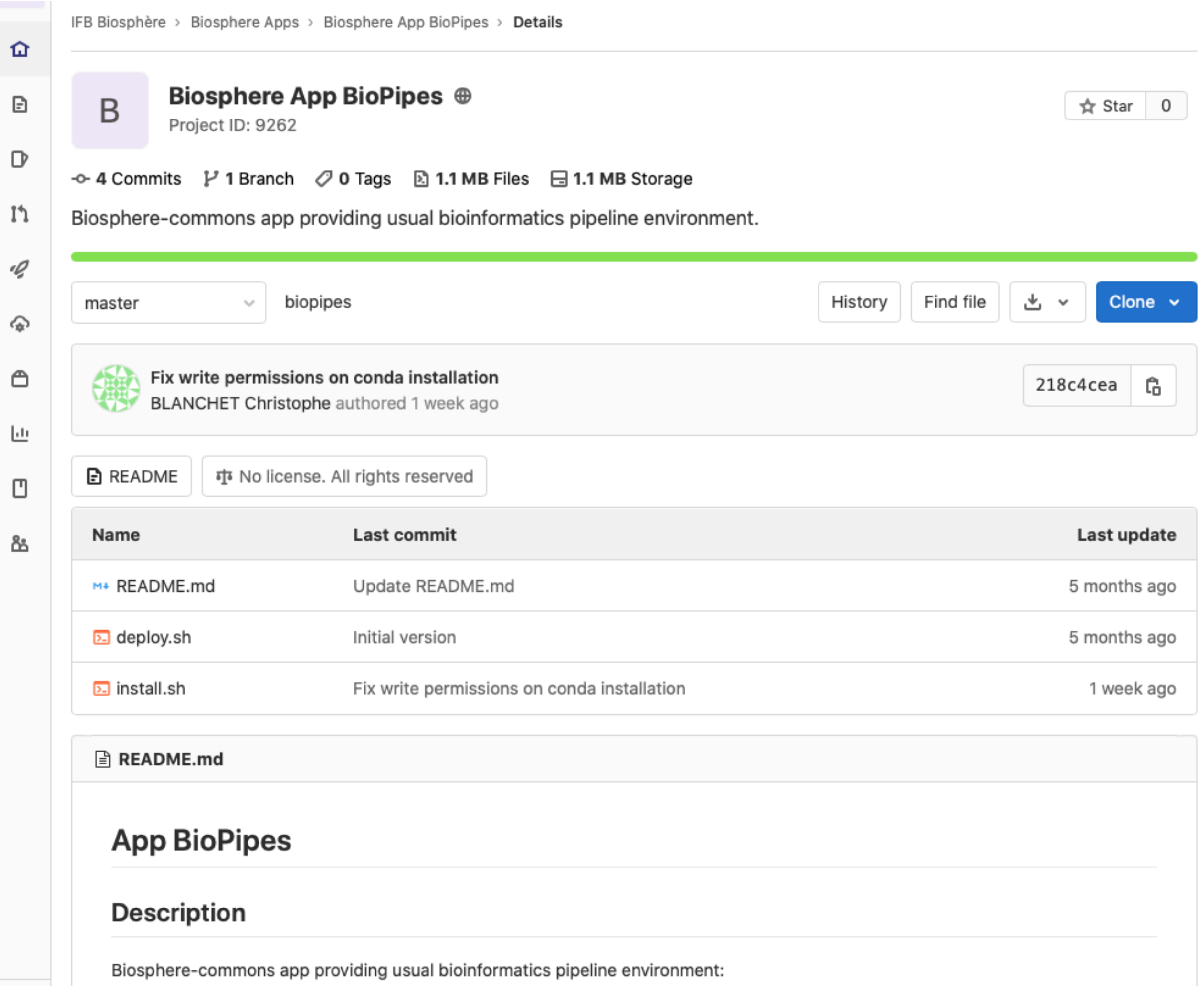
OS	Ubuntu 18.04
Recette de l'app (git)	https://gitlab.in2p3.fr/ifb-biosphere/apps/biopipes
App de base	Ubuntu 18.04

Caractéristiques

Nom long	Biosphere commons pipeline environments
Version	latest
Créé.e	8 juillet 2018 10:40
Dernière mise à jour	16 septembre 2020 18:43

Temps moyen de déploiement (hh:mm:ss)	jour précédent	semaine précédente	mois précédent	année précédente
Global	00:15:08	00:15:53	00:15:13	00:12:40

https://biosphere.france-bioinformatique.fr/catalogue/appliance/119



Biosphere App BioPipes

Project ID: 9262

4 Commits · 1 Branch · 0 Tags · 1.1 MB Files · 1.1 MB Storage

Biosphere-commons app providing usual bioinformatics pipeline environment.

Fix write permissions on conda installation

BLANCHET Christophe authored 1 week ago

Name	Last commit	Last update
README.md	Update README.md	5 months ago
deploy.sh	Initial version	5 months ago
install.sh	Fix write permissions on conda installation	1 week ago

README.md

App BioPipes

Description

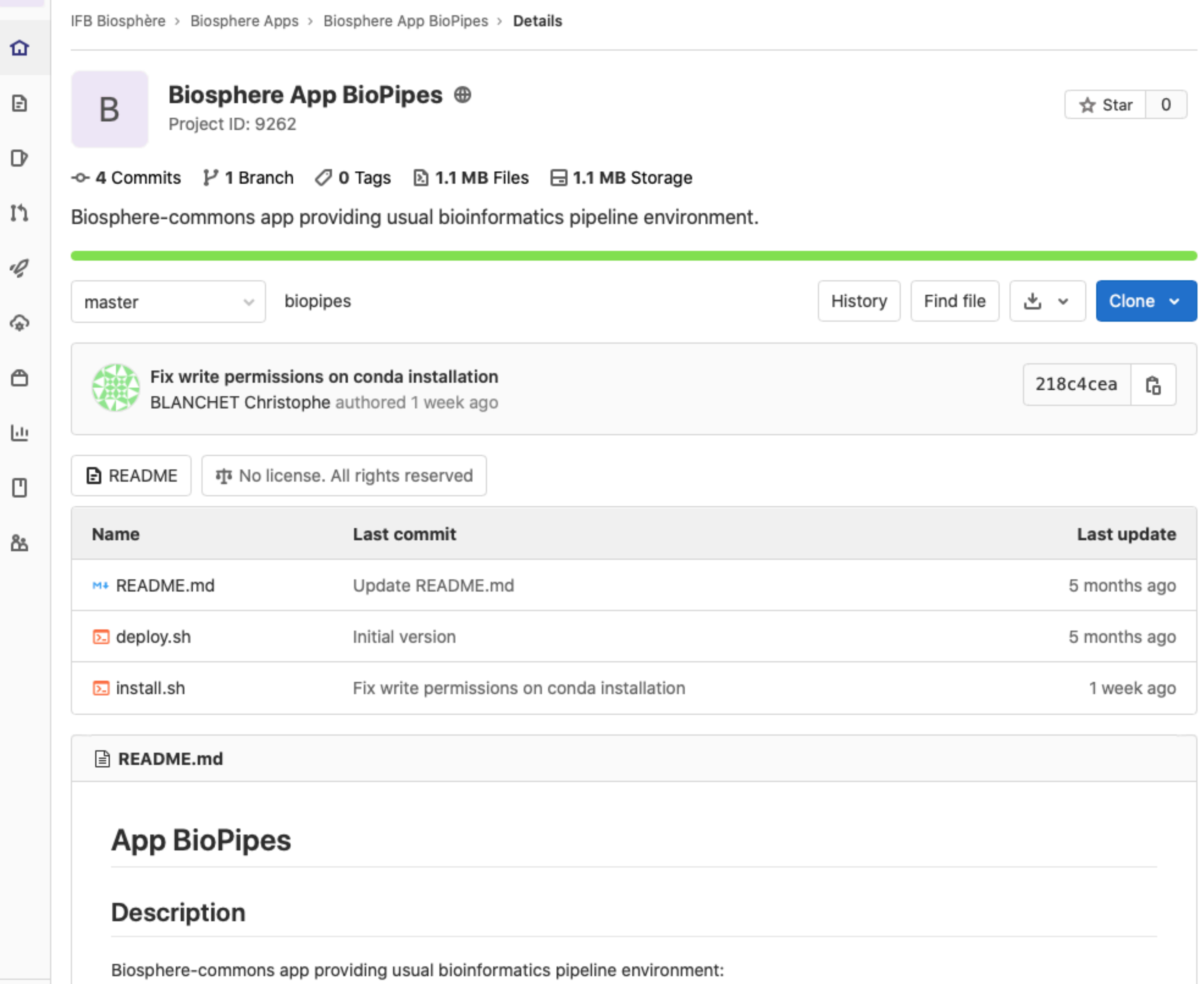
Biosphere-commons app providing usual bioinformatics pipeline environment:

https://gitlab.in2p3.fr/ifb-biosphere/apps/biopipes

Apps Biosphère - Héritage de Configurations

BioPipes < Ubuntu-IFB

- La configuration BioPipes hérite de la configuration Ubuntu-IFB
  - ★ Ubuntu-IFB : bioconda, docker, ansible
  - ★ BioPipes : Nextflow, Snakemake, CWLtool



IFB Biosphère › Biosphere Apps › Biosphere App BioPipes › **Details**

Biosphere App BioPipes
Project ID: 9262

4 Commits · 1 Branch · 0 Tags · 1.1 MB Files · 1.1 MB Storage

Biosphere-commons app providing usual bioinformatics pipeline environment.

master · biopipes

Fix write permissions on conda installation
BLANCHET Christophe authored 1 week ago · 218c4cea

README · No license. All rights reserved

Name	Last commit	Last update
README.md	Update README.md	5 months ago
deploy.sh	Initial version	5 months ago
install.sh	Fix write permissions on conda installation	1 week ago

README.md

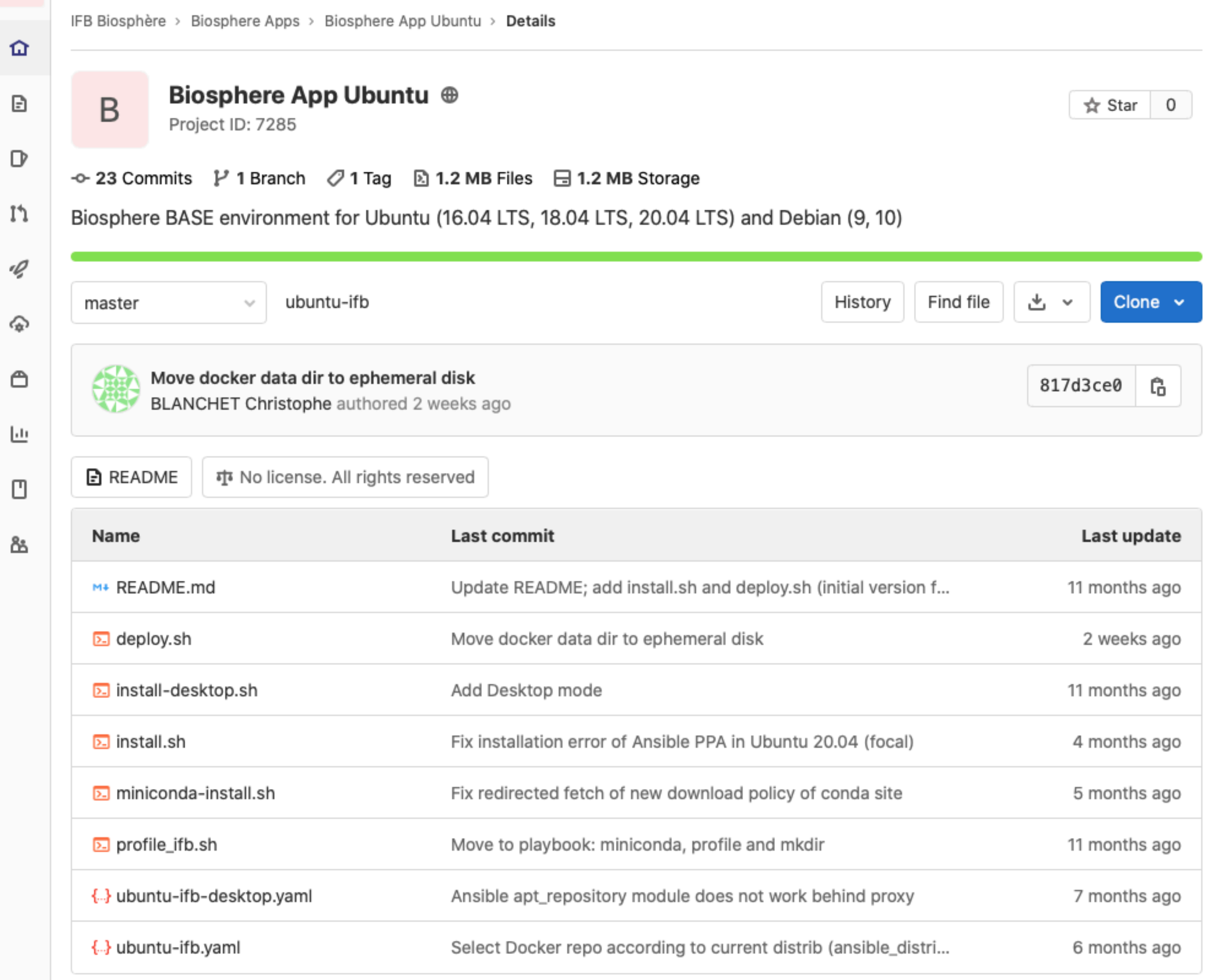
App BioPipes

Description

Biosphere-commons app providing usual bioinformatics pipeline environment:

https://biosphere.france-bioinformatique.fr/catalogue/appliance/119

https://gitlab.in2p3.fr/ifb-biosphere/apps/biopipes



IFB Biosphère › Biosphere Apps › Biosphere App Ubuntu › **Details**

Biosphere App Ubuntu
Project ID: 7285

23 Commits · 1 Branch · 1 Tag · 1.2 MB Files · 1.2 MB Storage

Biosphere BASE environment for Ubuntu (16.04 LTS, 18.04 LTS, 20.04 LTS) and Debian (9, 10)

master · ubuntu-ifb

Move docker data dir to ephemeral disk
BLANCHET Christophe authored 2 weeks ago · 817d3ce0

README · No license. All rights reserved

Name	Last commit	Last update
README.md	Update README; add install.sh and deploy.sh (initial version f...	11 months ago
deploy.sh	Move docker data dir to ephemeral disk	2 weeks ago
install-desktop.sh	Add Desktop mode	11 months ago
install.sh	Fix installation error of Ansible PPA in Ubuntu 20.04 (focal)	4 months ago
miniconda-install.sh	Fix redirected fetch of new download policy of conda site	5 months ago
profile_ifb.sh	Move to playbook: miniconda, profile and mkdir	11 months ago
ubuntu-ifb-desktop.yaml	Ansible apt_repository module does not work behind proxy	7 months ago
ubuntu-ifb.yaml	Select Docker repo according to current distrib (ansible_distri...	6 months ago

https://biosphere.france-bioinformatique.fr/catalogue/appliance/120

https://gitlab.in2p3.fr/ifb-biosphere/apps/ubuntu-ifb

Déployer son propre environnement

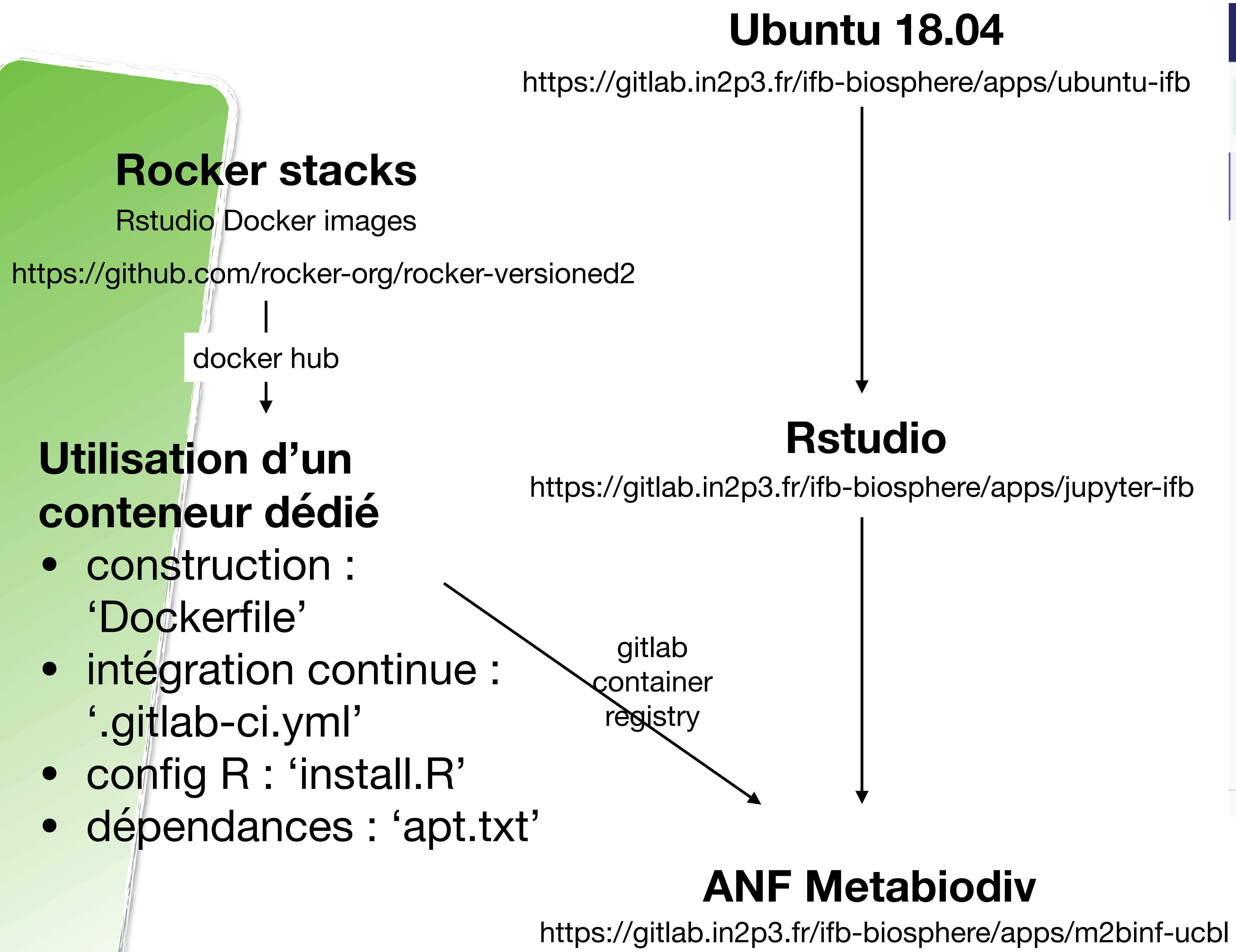
Outils

- code source (git)
- paquet linux (apt/yum) : ‘apt.txt'
- paquet pip
- environnement conda (bioconda)
- conteneur (docker hub, registres locaux, biocontainer)
- paquet R (CRAN, git) : ‘install.R’

Pipelines

- netxflow
- snakemake
- CWL
- jupyter notebook

App ANF Metabiodiv à base de Rstudio



Ubuntu 18.04

https://gitlab.in2p3.fr/ifb-biosphere/apps/ubuntu-ifb

Rocker stacks

Rstudio Docker images

https://github.com/rocker-org/rocker-versioned2

docker hub

Rstudio

https://gitlab.in2p3.fr/ifb-biosphere/apps/jupyter-ifb

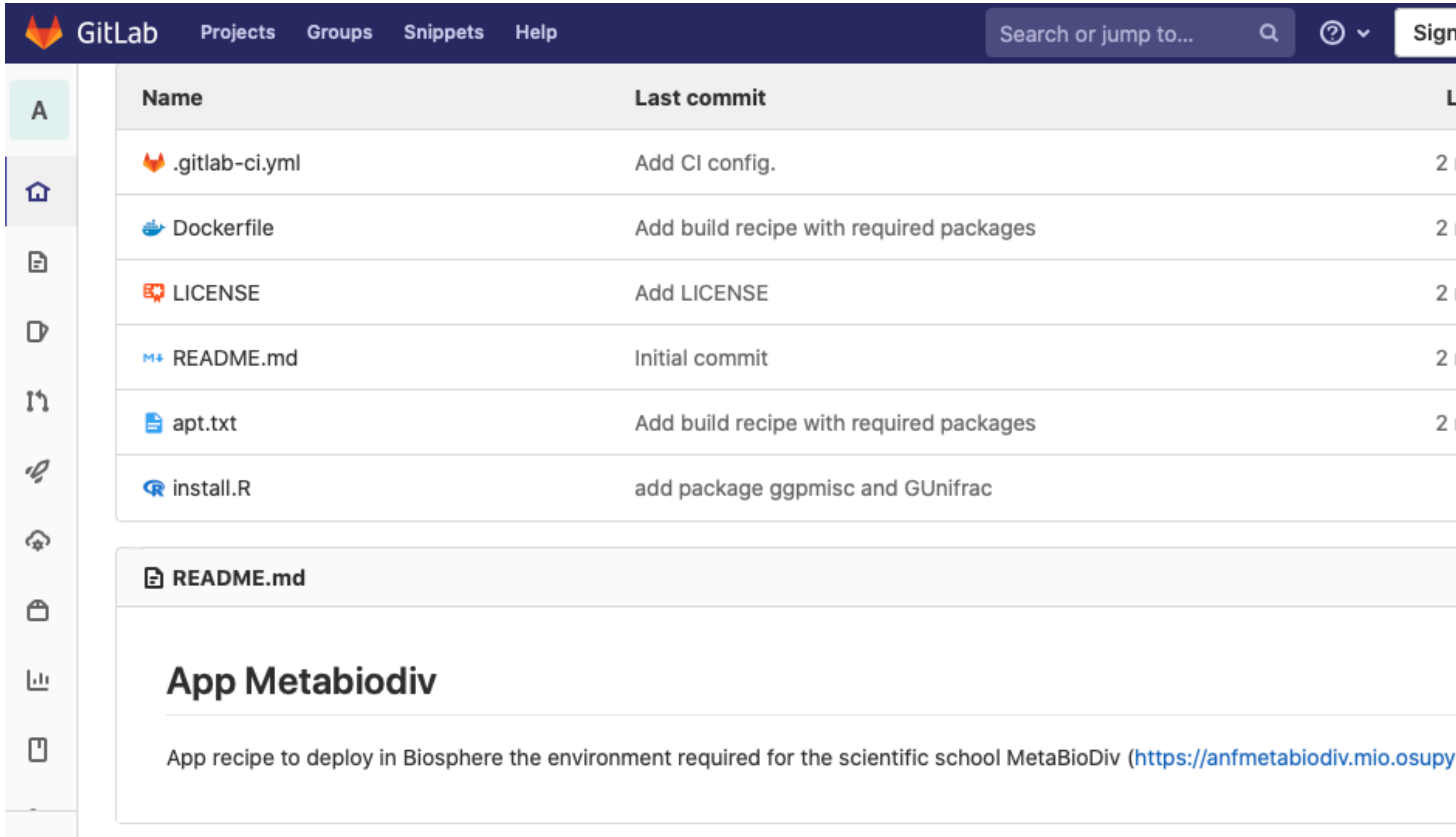
Utilisation d’un conteneur dédié

- construction : ‘Dockerfile’
- intégration continue : ‘.gitlab-ci.yml’
- config R : ‘install.R’
- dépendances : ‘apt.txt’

gitlab container registry

ANF Metabiodiv

https://gitlab.in2p3.fr/ifb-biosphere/apps/m2binf-ucbl



Name	Last commit
.gitlab-ci.yml	Add CI config.
Dockerfile	Add build recipe with required packages
LICENSE	Add LICENSE
README.md	Initial commit
apt.txt	Add build recipe with required packages
install.R	add package ggpmisc and GUnifrac

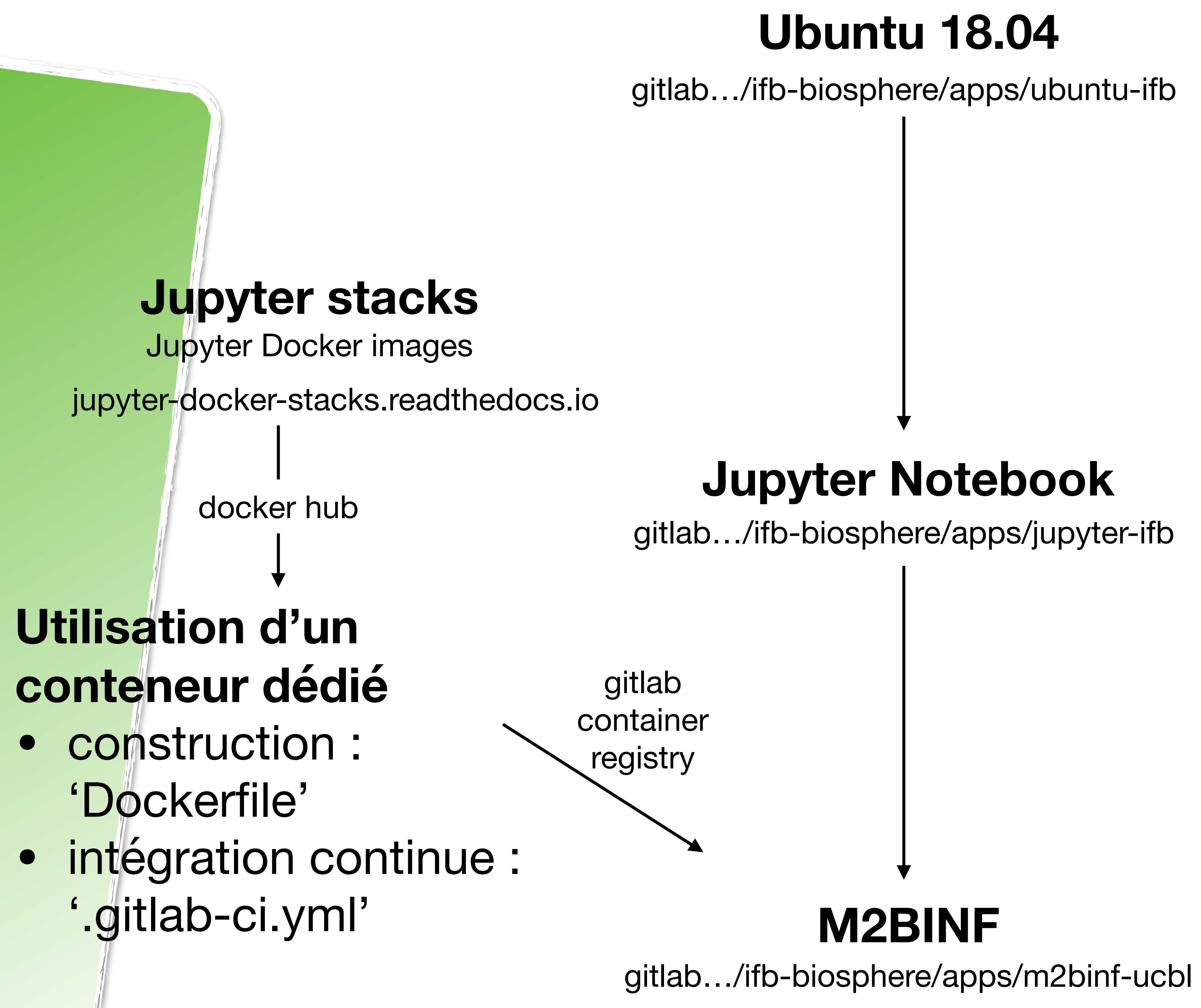
README.md

App Metabiodiv

App recipe to deploy in Biosphere the environment required for the scientific school MetaBioDiv (https://anfmetabiodiv.mio.osupy

https://gitlab.in2p3.fr/christophe.blanchet/app-metabiodiv

App M2BINF à base de Jupyter



Ubuntu 18.04
gitlab…/ifb-biosphere/apps/ubuntu-ifb

Jupyter stacks
Jupyter Docker images
jupyter-docker-stacks.readthedocs.io

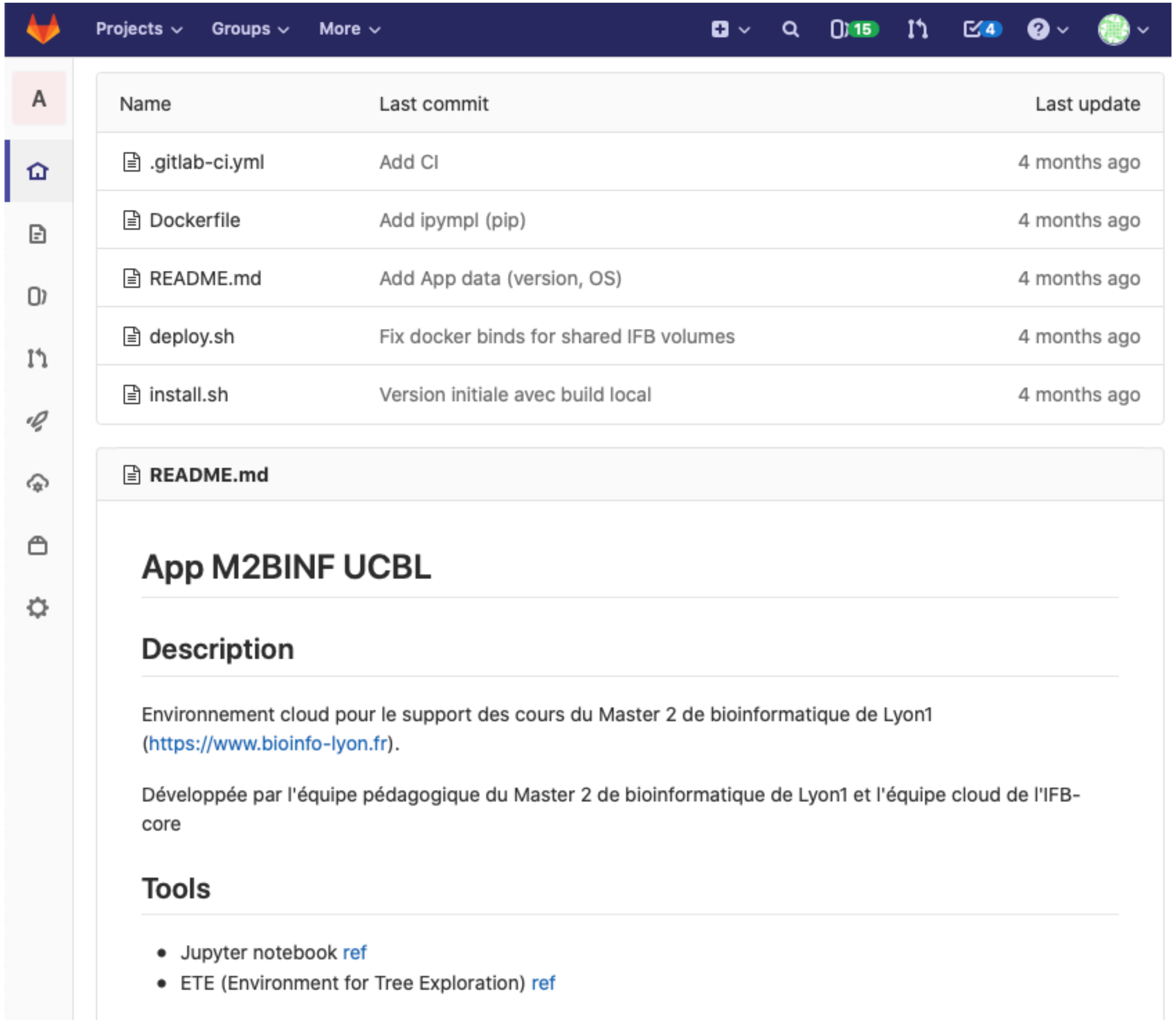
docker hub

Jupyter Notebook
gitlab…/ifb-biosphere/apps/jupyter-ifb

Utilisation d'un conteneur dédié
- construction : 'Dockerfile'
- intégration continue : '.gitlab-ci.yml'

gitlab
container
registry

M2BINF
gitlab…/ifb-biosphere/apps/m2binf-ucbl



Name	Last commit	Last update
.gitlab-ci.yml	Add CI	4 months ago
Dockerfile	Add ipympl (pip)	4 months ago
README.md	Add App data (version, OS)	4 months ago
deploy.sh	Fix docker binds for shared IFB volumes	4 months ago
install.sh	Version initiale avec build local	4 months ago

README.md

App M2BINF UCBL

Description

Environnement cloud pour le support des cours du Master 2 de bioinformatique de Lyon1 (https://www.bioinfo-lyon.fr).

Développée par l'équipe pédagogique du Master 2 de bioinformatique de Lyon1 et l'équipe cloud de l'IFB-core

Tools

- Jupyter notebook ref
- ETE (Environment for Tree Exploration) ref

https://gitlab.in2p3.fr/ifb-biosphere/apps/m2binf-ucbl