DE LA RECHERCHE À L'INDUSTRIE

www.cea.fr

PathoTRACK platform on IFB Cloud

IFB Cyclone Hackathon

Adrien JOSSO

IFB-Core - Genoscope/LABGeM

November 15, 2016

CBRNE Research and Innovation Project

- Partners
  - CEA - CNS / LABGeM
  - CEA - CNG / Sequencing Platform
  - IFB-Core
  - DGA - Maîtrise NRBC
- started since 2015
- born from CBRNE R&I project named *SynBioWatch* (2012-2014)

PathoTRACK project
Aims

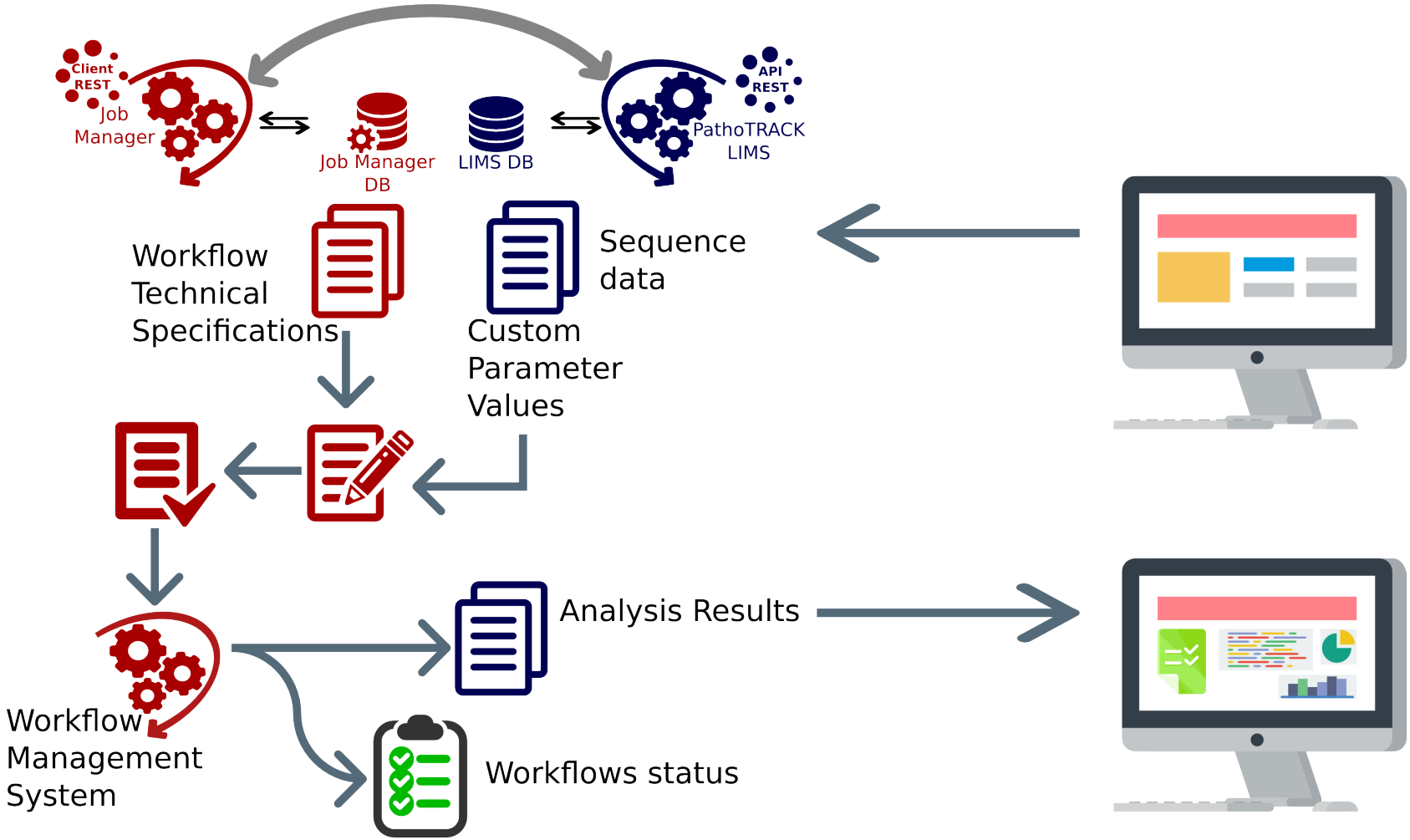
Wet Lab / Sequencing

- Improve DNA extraction protocol (complex matrix: soil, petrol, ...)
- Benchmarck **"portable"** sequencing technologies
  - Long reads: Oxford Nanopore Minion, IonTorrent PGM
  - *versus*
  - Short reads: Illumina MiSeq

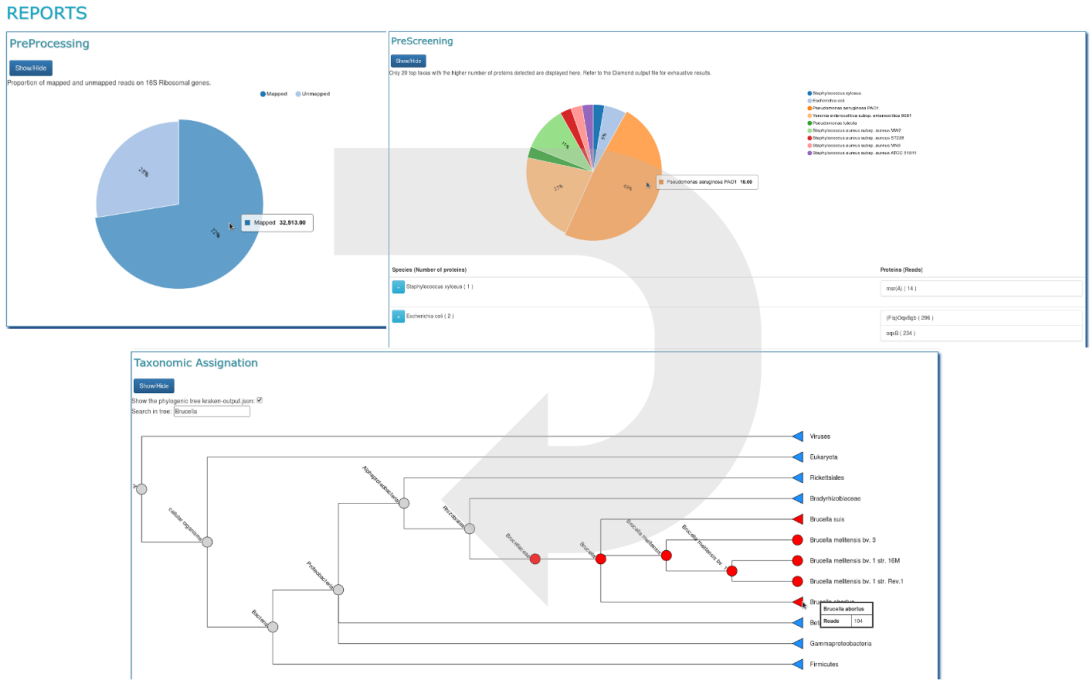
BioInformatics

- **Quick identification** and **characterization** of pathogens from environmental and clinical samples
- Develop reference databases (Taxonomique Assignation, ...)
- Bioinformatic tools integration (SortMeRna, Diamond, Kraken, Centrifuge)
- Workflow management system: SnakeMake

PathoTRACK platform
Workflow Management System



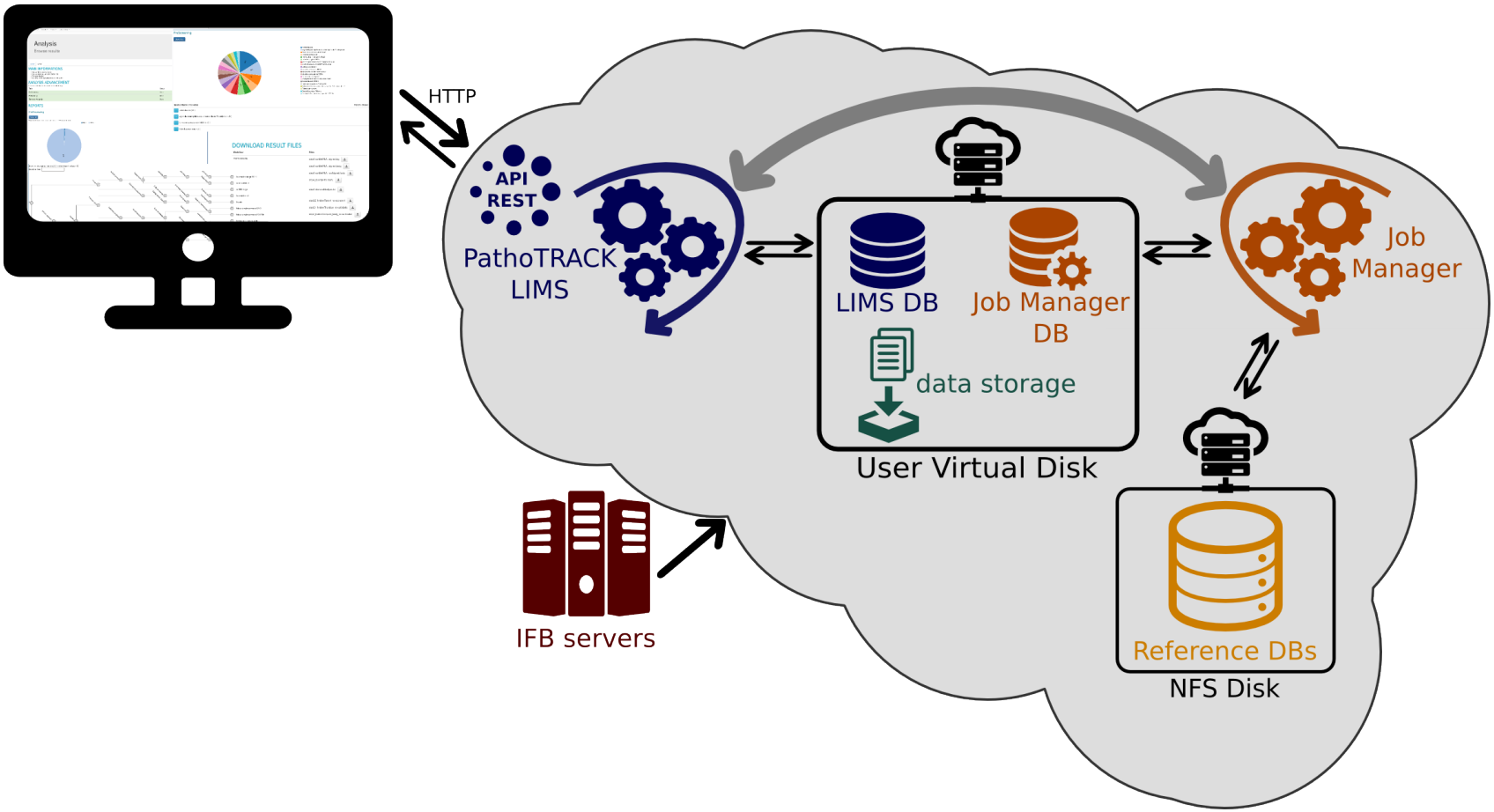
PathoTRACK platform
Analysis



3 workflows

- Preprocessing sequencing data
- Functional annotation (genes involved in drug resistance and virulance)
- Taxonomic assignation

PathoTRACK platform
On IFB Cloud



Micro-Services Architecture

- 4 services
  - Web User Interface
  - LIMS
  - Workflow Manager
  - MySQL server
- Each service is elastic, resilient, composable, minimal, and complete

- OS: Linux CentOS 6 (7 in progress) (64 bits)
- Can be deployed on Desktop, VM, servers, HPC clusters, Cloud (IFB)
- Start and stop the platform using a daemon

Installation

- The platform is packaged in an archive (tar.gz)
- Bash script (need root privileges)
- Update the system
- Install all needed packages using yum (they will be packaged with the platform to perform an offline installation)
- Install all services of PathoTRACK platform
- Configure the automatic launch of the platform (init.d / chkconfig)
- Configure Apache httpd server in ProxyPass mode (ports redirecting)
- Initialize the platform

Use the CYCLONE components to deploy the platform on the new version of IFB cloud

- Use SlipStream to deploy the platform ? each service on a dedicated VM ?
- Configure VM access (Authentication)
- Configure a dedicated cluster on IFB Cloud ?

Thank for your attention

PathoTRACK platform
Development Team

LABGeM

Project Leader: Stephane Cruveiller, Ph D.

- Johann Beghain
- Marine Séjourné

CNG Sequencing platform

- Robert Olaso, Ph D.
- Céline Bigot, Ph D.

IFB-Core

- Adrien Josso