



# Linux in the nuclear industry

19 Mars 2018

Mehdi DOGGUY  
EDF



Who am I?

# Who am I?

- ▶ 2006 : M1 Informatique, Paris 11

# Who am I?

- ▶ 2006 : M1 Informatique, Paris 11
- ▶ 2007 : M2 MPRI, Paris 7

# Who am I?

- ▶ 2006 : M1 Informatique, Paris 11
- ▶ 2007 : M2 MPRI, Paris 7
- ▶ 2008 - 2011 : PhD thesis in Computer Science, Paris 7  
*"Determinacy and confluence in synchronous process calculi"*

# Who am I?

- ▶ 2006 : M1 Informatique, Paris 11
- ▶ 2007 : M2 MPRI, Paris 7
- ▶ 2008 - 2011 : PhD thesis in Computer Science, Paris 7  
*"Determinacy and confluence in synchronous process calculi"*
- ▶ 2011 - 2012 : Post-Doc, ENS Ulm  
Affine typing in ReactiveML

# Who am I?

- ▶ 2006 : M1 Informatique, Paris 11
- ▶ 2007 : M2 MPRI, Paris 7
- ▶ 2008 - 2011 : PhD thesis in Computer Science, Paris 7  
*"Determinacy and confluence in synchronous process calculi"*
- ▶ 2011 - 2012 : Post-Doc, ENS Ulm  
Affine typing in ReactiveML
- ▶ 2008 - present : Debian developer  
Started with small packages in the OCaml team...

# Who am I?

- ▶ 2006 : M1 Informatique, Paris 11
- ▶ 2007 : M2 MPRI, Paris 7
- ▶ 2008 - 2011 : PhD thesis in Computer Science, Paris 7  
*"Determinacy and confluence in synchronous process calculi"*
- ▶ 2011 - 2012 : Post-Doc, ENS Ulm  
Affine typing in ReactiveML
- ▶ 2008 - present : Debian developer  
Started with small packages in the OCaml team...
- ▶ 2012 - present : HPC Expert, EDF



# About EDF

- ▶ World nuclear energy leader
- ▶ Europe leader in hydro power
- ▶ Industrial operator in Asia and United States
  
- ▶ Some key figures
  - ▶ 71.2 billion euros annual revenue
  - ▶ 37.6 M clients worldwide
  - ▶ 584.7 TWh produced annually
  - ▶ 136 GW production capacity
  - ▶ 73 nuclear reactors, 78% of production
  - ▶ 154 845 employees worldwide
  
- ▶ Large R&D and engineering divisions



<https://www.edf.fr/en/the-edf-group/who-we-are/edf-at-a-glance>

# Intro

Since 2003, EDF has been using a custom distro on scientific workstations, HPC clusters and servers for industrial R&D and engineering.

EDF decided to make its distro publicly available and turn it into an Open Source community-driven project.

Outline of this presentation:

- ▶ business context
- ▶ scientific and engineering IT needs
- ▶ our solution

# 1 Scientific computing needs

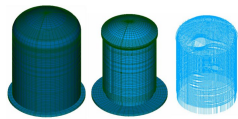
# Scientific computing needs at EDF

- ▶ R&D
  - ▶ Conception
  - ▶ Information technology
  - ▶ Renewable energies
  - ▶ Electrical networks
  - ▶ ...
- ▶ Engineering
  - ▶ Safety studies
  - ▶ Simulations
- ▶ Energy management
  - ▶ Reduce downtime on existing reactors
  - ▶ Planning consumption and production weeks in advance

# Scientific computing at a glance

## ► Modeling

- Approximate reality with a model
- Often need for a modeler



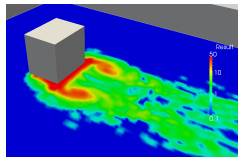
## ► Simulation

- Execution of a numerical code computing the behavior of the model system
- A whole area of software development
- Need for the fastest hardware to work on large arrays of floats



## ► Visualization

- Results exploration and analysis
- Need for the best graphics hardware and displays



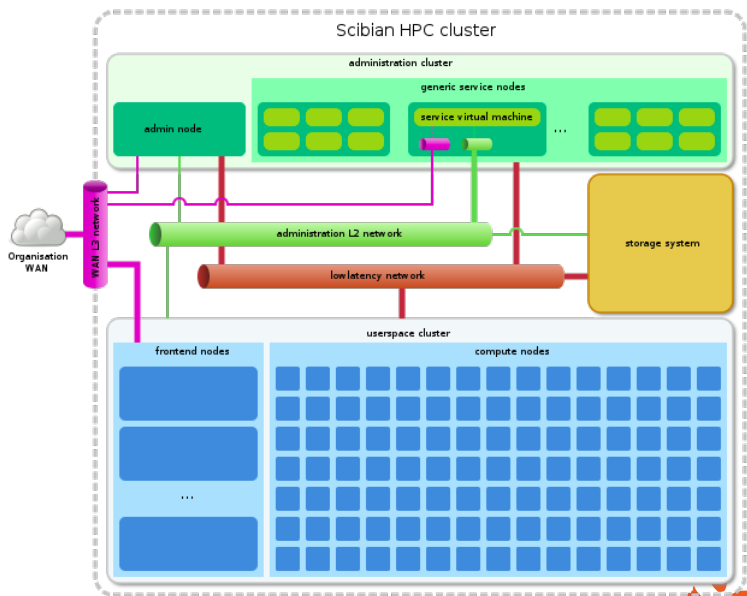
# Open Source scientific tools, by EDF

- ▶ Code\_Saturne
  - ▶ general-purpose computational fluid dynamics software
  - ▶ <https://www.code-saturne.org/>
  
- ▶ Code\_Aster
  - ▶ Structures and Thermomechanics Analysis for Studies and Research
  - ▶ <https://www.code-aster.org/>
  
- ▶ Salomé
  - ▶ Integration platform for numerical simulation
  - ▶ <http://www.salome-platform.org/>
  
- ▶ ...

# Typical IT user needs

- ▶ Scientific workstation / laptop
  - ▶ Modeling and visualizing
  - ▶ Developing scientific simulation codes
  - ▶ Accessing the HPC clusters
- ▶ High performance clusters
  - ▶ Thousands of compute nodes connected using fast network
  - ▶ More or less specialized depending on applications
- ▶ Computing chains
  - ▶ Servers or small clusters
  - ▶ Regular execution of the same code
  - ▶ Coupling with other components

# An HPC<sup>1</sup> cluster



<sup>1</sup>High Performance Computing



## 2 Scibian

# Our solution: Scibian



# What's Debian?

# What's Debian?

- ▶ A **technical project**
  - ▶ Building a successful distribution
  - ▶ With a real impact on the world



# What's Debian?

- ▶ A **technical project**
  - ▶ Building a successful distribution
  - ▶ With a real impact on the world
- ▶ A **philosophical & political project**
  - ▶ Promoting and defending Free Software
  - ▶ With an impact on many users



# What's Debian?

- ▶ A **technical project**
  - ▶ Building a successful distribution
  - ▶ With a real impact on the world
- ▶ A **philosophical & political project**
  - ▶ Promoting and defending Free Software
  - ▶ With an impact on many users
- ▶ A **social experiment**
  - ▶ Thousands of volunteer contributors all over the world



# What's Debian?

- ▶ A **technical project**
  - ▶ Building a successful distribution
  - ▶ With a real impact on the world
- ▶ A **philosophical & political project**
  - ▶ Promoting and defending Free Software
  - ▶ With an impact on many users
- ▶ A **social experiment**
  - ▶ Thousands of volunteer contributors all over the world

## Part of the landscape of Free Software:

- ▶ 1983 – GNU
- ▶ 1991 – Linux
- ▶ **1993** – Slackware, **Debian**, Red Hat
- ▶ 1998 – OSI created by former DPL; definition based on DFSG
- ▶ 2004 – Ubuntu

# Free Software without distributions

Many different:

- ▶ **project hosting solutions** (sourceforge, github, author's homepage, etc.)
- ▶ **build systems** (autotools, CMake, Rubygems, custom-made)
- ▶ **causes of problems** (dependencies, conflicting libraries, etc.)

1. download foo-1.0.tar.gz  
*checksum mismatch, missing public key, etc.*
2. ./configure  
*error: missing bar, baz,...*
3. foreach (bar, baz, ...) go to 1 until success
4. make  
*error: symbol not found*
5. make install  
*error: cp: cannot create regular file /some/weird/path*



# A Free Software distribution



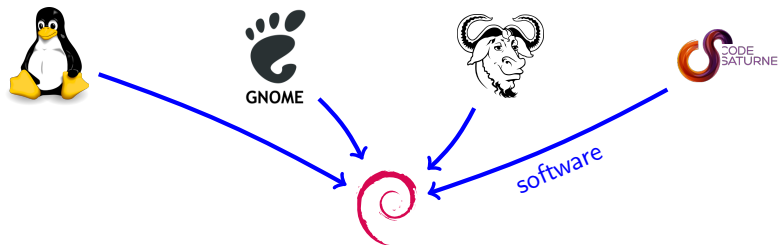
Alice

Bob

Carol

Dave

# A Free Software distribution



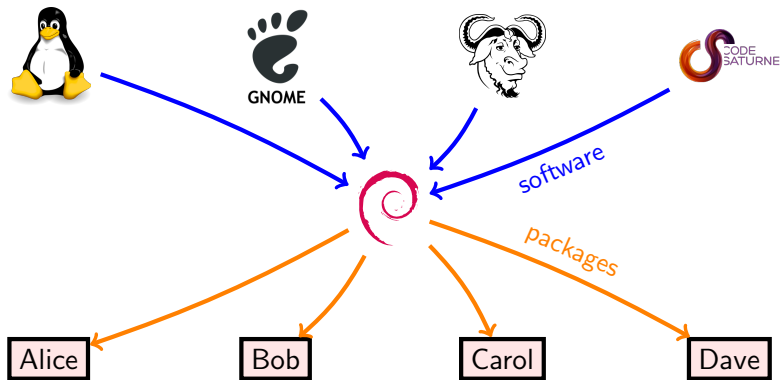
Alice

Bob

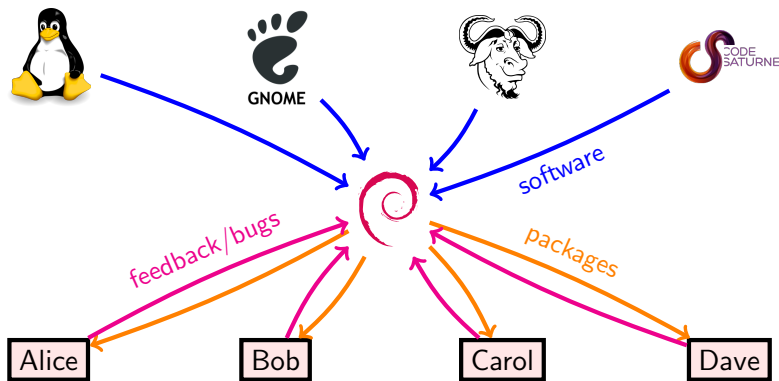
Carol

Dave

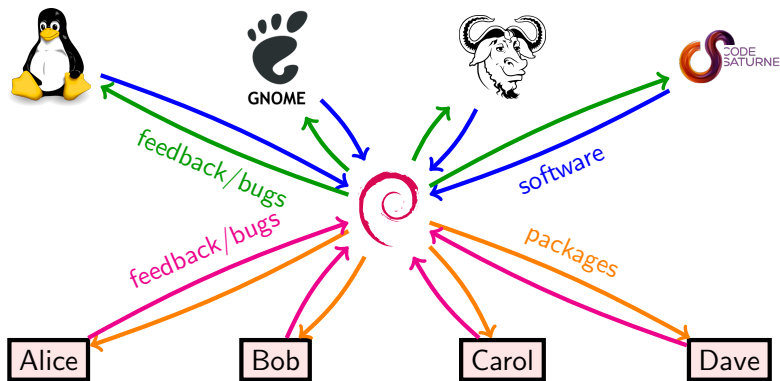
# A Free Software distribution



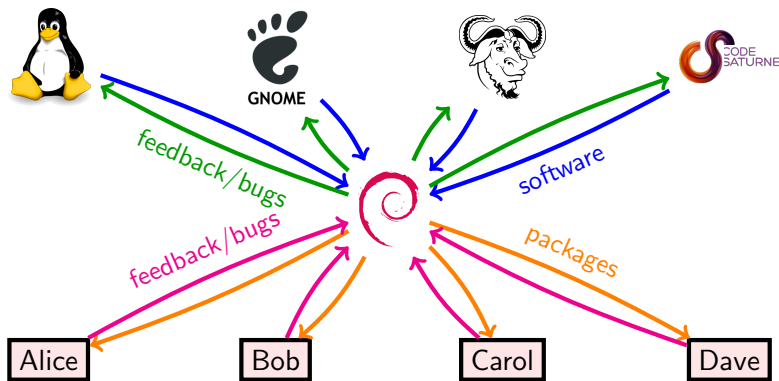
# A Free Software distribution



# A Free Software distribution



# A Free Software distribution



1. Integration of upstream projects (avoid conflicting versions)
2. Efficient infrastructure to distribute software to users (mirrors)
3. Unified interface for software installation, upgrade and removal
4. Intermediate, unified support layer

## A very successful project

- ▶ **+50 000 packages**, The largest Free Software archive
- ▶ **a dozen of ports** (with 3 different kernels – Linux, Hurd, FreeBSD)
- ▶ 100% Free Software (including infrastructure)
- ▶ Known for stability, robustness, expertise of package maintainers
  - ▶ Huge focus on QA tests, including for packages in the long tail
- ▶ Stable releases every 2 years (+/- few months)

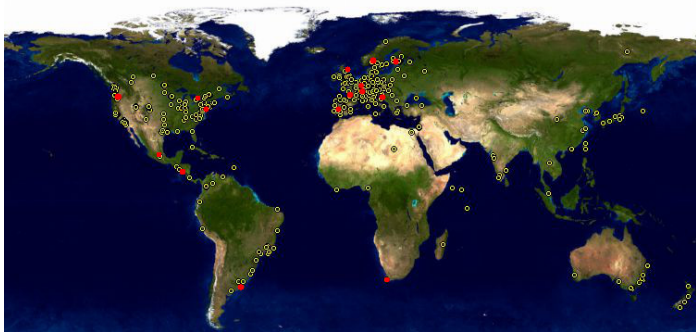
# Who is doing Debian?

- ▶ **Thousands of volunteers** (including approx. 1 000 Debian Developers)
- ▶ Organized in **teams**, focusing on:
  - ▶ **Packaging tasks**: Ruby, KDE, scientific, etc.  
and also: porters, security, release management, blends, translation, etc.
  - ▶ **Distribution infrastructure**: system administration, packages archive, build daemons, mirrors, quality assurance, etc.
  - ▶ **Project support & communication**: press, documentation, trademark, auditors & accounting, events, videos, etc.
- ▶ Real **experts** of their packages, and of their tasks in general



## A distributed and independent organization

- ▶ Developers coming from 63 countries (us: 18%, de: 17%, **fr: 10%**)
- ▶ Distributed infrastructure
- ▶ Myriad of loosely-connected services  $\rightsquigarrow$  easy to add your own
- ▶ Assets held by several Trusted Organizations
- ▶ **No company with a major influence on Debian**



# Why choosing Debian for Scibian?

- ▶ **One OS for workstations, servers and clusters**
  - ▶ Full binary compatibility
- ▶ Appropriate release cycle
  - ▶ One major version approx. every 3 years
  - ▶ Frequent updates (for critical bugs and security fixes)
- ▶ Largest scientific software offering
  - ▶ Only Ubuntu matches, by following Debian repositories
- ▶ Designed for customization
  - ▶ Custom repositories, easy deployment
  - ▶ Modular and hookable installer
- ▶ Community openness
  - ▶ Easy to get interesting changes into the distribution
- ▶ Easy to integrate applications
  - ▶ Cool packaging helpers
  - ▶ Abundant documentation

# Why choosing Debian for Scibian?

- ▶ One OS for workstations, servers and clusters
  - ▶ Full binary compatibility
- ▶ **Appropriate release cycle**
  - ▶ One major version approx. every 3 years
  - ▶ Frequent updates (for critical bugs and security fixes)
- ▶ Largest scientific software offering
  - ▶ Only Ubuntu matches, by following Debian repositories
- ▶ Designed for customization
  - ▶ Custom repositories, easy deployment
  - ▶ Modular and hookable installer
- ▶ Community openness
  - ▶ Easy to get interesting changes into the distribution
- ▶ Easy to integrate applications
  - ▶ Cool packaging helpers
  - ▶ Abundant documentation

# Why choosing Debian for Scibian?

- ▶ One OS for workstations, servers and clusters
  - ▶ Full binary compatibility
- ▶ Appropriate release cycle
  - ▶ One major version approx. every 3 years
  - ▶ Frequent updates (for critical bugs and security fixes)
- ▶ **Largest scientific software offering**
  - ▶ Only Ubuntu matches, by following Debian repositories
- ▶ Designed for customization
  - ▶ Custom repositories, easy deployment
  - ▶ Modular and hookable installer
- ▶ Community openness
  - ▶ Easy to get interesting changes into the distribution
- ▶ Easy to integrate applications
  - ▶ Cool packaging helpers
  - ▶ Abundant documentation

# Why choosing Debian for Scibian?

- ▶ One OS for workstations, servers and clusters
  - ▶ Full binary compatibility
- ▶ Appropriate release cycle
  - ▶ One major version approx. every 3 years
  - ▶ Frequent updates (for critical bugs and security fixes)
- ▶ Largest scientific software offering
  - ▶ Only Ubuntu matches, by following Debian repositories
- ▶ **Designed for customization**
  - ▶ Custom repositories, easy deployment
  - ▶ Modular and hookable installer
- ▶ Community openness
  - ▶ Easy to get interesting changes into the distribution
- ▶ Easy to integrate applications
  - ▶ Cool packaging helpers
  - ▶ Abundant documentation

# Why choosing Debian for Scibian?

- ▶ One OS for workstations, servers and clusters
  - ▶ Full binary compatibility
- ▶ Appropriate release cycle
  - ▶ One major version approx. every 3 years
  - ▶ Frequent updates (for critical bugs and security fixes)
- ▶ Largest scientific software offering
  - ▶ Only Ubuntu matches, by following Debian repositories
- ▶ Designed for customization
  - ▶ Custom repositories, easy deployment
  - ▶ Modular and hookable installer
- ▶ **Community openness**
  - ▶ Easy to get interesting changes into the distribution
- ▶ Easy to integrate applications
  - ▶ Cool packaging helpers
  - ▶ Abundant documentation

# Why choosing Debian for Scibian?

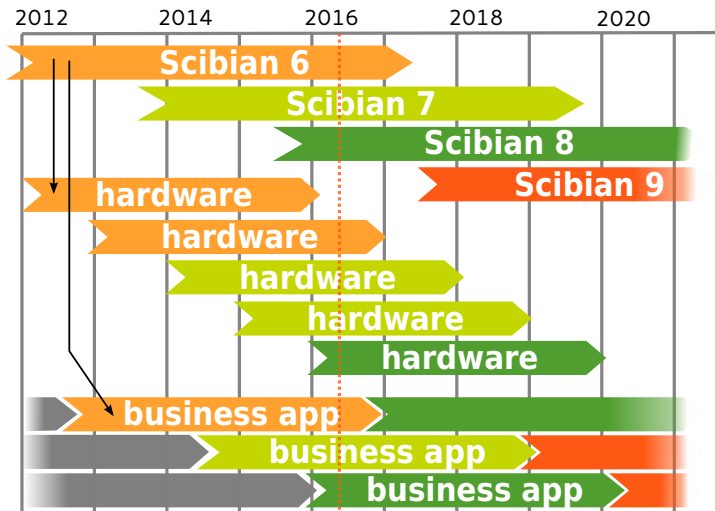
- ▶ One OS for workstations, servers and clusters
  - ▶ Full binary compatibility
- ▶ Appropriate release cycle
  - ▶ One major version approx. every 3 years
  - ▶ Frequent updates (for critical bugs and security fixes)
- ▶ Largest scientific software offering
  - ▶ Only Ubuntu matches, by following Debian repositories
- ▶ Designed for customization
  - ▶ Custom repositories, easy deployment
  - ▶ Modular and hookable installer
- ▶ Community openness
  - ▶ Easy to get interesting changes into the distribution
- ▶ **Easy to integrate applications**
  - ▶ Cool packaging helpers
  - ▶ Abundant documentation

# Our solution: Scibian





# Life cycle



# Why a derivative?

- ▶ 6+ year support
- ▶ Custom security support
- ▶ Workstation and HPC hardware support backports (mainly for Infiniband, OmniPath, nVidia GPUs and newer Intel micro-architectures)
- ▶ Support of some libs removed from Debian
- ▶ Upgrade to newer major versions of some software

# Business Applications

Packaging of business applications follows a few rules:

- ▶ No maintainer scripts and no services
- ▶ No files outside `/opt/$name-$version`, except:
  - ▶ a script in `/usr/bin`
  - ▶ manpages
  - ▶ copyright file
  - ▶ icons and a desktop file
- ▶ Pre-Depends must be empty
- ▶ No alternative dependencies and no Provides
- ▶ All files are owned by root and writable only by root
- ▶ No setuid/setgid binaries (or other means to escalate privileges)
- ▶ ...

Each application is made available on each published version of Scibian.

# 3 Tools integrated into Scibian

# Puppet HPC

The main goal of Puppet-HPC is to provide a common generic configuration management system that can be used effortlessly across multiple HPC clusters and organizations.

The Puppet-HPC software stack notably provides:

- ▶ Many generic Puppet modules (>80) for all technical components required on a HPC cluster
- ▶ Defined data model for representing the description of an HPC cluster based on Hieradata
- ▶ Tools to easily deploy and manage the configuration with high-scalability requirements

It is heavily tested on Debian and used in production on thousands of Debian machines.

# Goals behind Puppet HPC

- ▶ The code base can be re-used and the development effort is shared.
- ▶ The same code is run on many different environments, it is therefore more tested and more reliable.
- ▶ The code can be easily tested on a small testing environment even if the data is different from the production environment.

More details available at:

- ▶ [https://edf-hpc.github.io/puppet-hpc/puppet\\_hpc\\_reference-0.1.html](https://edf-hpc.github.io/puppet-hpc/puppet_hpc_reference-0.1.html)
- ▶ <https://github.com/edf-hpc/puppet-hpc>

# Scibian HPC Installation Guide

Standardize Scibian cluster deployments by:

- ▶ Defining a general architecture for an HPC Cluster
- ▶ Describing how to install it using tools packaged in Scibian
- ▶ Showing how to configure it using our Puppet modules

Document available at:

<https://scibian.github.io/scibian-hpc-install-guide/>

## 4 Why should I use Scibian?



# Why should I use Scibian?

- ▶ Still use Debian (We don't even change the kernel!)
- ▶ More relaxed update rules
- ▶ A set of integrated tools for deployment and scientific computing infrastructures
- ▶ Meet and work with other industrial users
- ▶ Well tested solution in a high demanding production environment
- ▶ Integrate specific proprietary software (that is not possible to integrate in Debian)
- ▶ ...



Merci!

Useful links:

- ▶ <https://twitter.com/ScibianLinux>
- ▶ <https://github.com/scibian/>
- ▶ <https://github.com/edf-hpc>
- ▶ <https://lists.debian.org/debian-hpc/>