



ID de Contribution: 45

Type: **Présentation standard**

## Utilisation d'une plateforme Hadoop/Spark pour des données astrophysiques

*mardi 27 septembre 2016 12:00 (20 minutes)*

Une expérimentation est menée dans l'environnement VirtualData/OpenStack au LAL.  
Plusieurs caractéristiques de cette technologie sont en cours d'évaluation:

- mise en place d'une infrastructure Hadoop/Spark/scipy sur une cloud OpenStack
- mesures de performance (configuration du cluster, suivi des facteurs d'échelle)
- apprentissage des modules fonctionnels de la plateforme Spark (Scala, Python)
- exploitation des données binaires structurées dans Spark à travers le module Hadoop/avro
- application aux fichiers FITS
- quelques exemples de cas d'usage dans le contexte de l'expérience LSST

**Christian Arnault**

Axel Chevarin

Réza Ansari

Johann Cohen Tanoudji

**Auteur principal:** M. ARNAULT, Christian (CNRS)

**Co-auteurs:** M. CHEVARIN, Axel (LAL); M. COHEN-TANUGI, JOHANN (LUPM, Université de Montpellier); Prof. ANSARI, Réza (LAL)

**Orateur:** M. ARNAULT, Christian (CNRS)

**Classification de Session:** Offline

**Classification de thématique:** Offline