JDEV 2015

T3 - Données massives scientifiques (Big data), recherche par les données

Mots clés :

Données ouvertes (open data)

Masse de données (big data), base de données NoSQL, persistance des données,structuration des données,

map-reduce, R, Julia, Python

Présentation

- http://devlog.cnrs.fr/jdev2015/t3
- **Plénière:**
  - Introduction au big data
  - Les bases NoSQL : Exemple : MongoDB, Cassandra, Oracle
  - Bases de données NoSQL orienté graphe : Neo4j
  - Elasticsearch : un moteur de recherche open source pour implémenter vos services de recherche big data et d'analyse
  - Hadoop, MapReduce et Spark pour vos développements de service
- **Ateliers :**
  - Bases noSQL : Oracle, Neo4J, Mongo DB,Cassandra
  - Prise en main d'elasticsearch et de Kibana
  - Atelier Hadoop : Map-Reduce et Spark pour le calcul scientifique + Traitement des données en parallèle avec Map-Reduce et Spark
  - Python, apprentissage statistique et analyse de données pour la modélisation prédictive avec scikit-learn
  - Julia pour vos calculs scientifiques intensifs

Présentation

- **GT:**
  - Foire aux bibliothèques thématiques scientifiques python : Echange et retour d'expérience. Extension des bibliothèques de références numpy, scipy, matplotlib à sa thématique scientifique. Ouverture à d'autres langages (fortran, C, R, julia, …) , intégration de l'existant par le côté glue du langage.

  - Les différentes méthodes d'optimisation d'un code Python (Cython, Numba, programmation parallèle ou utilisation de GPU avec CUDA)