

Atelier « Hadoop : l'écosystème »

Afin de comprendre quels sont les différents éléments de l'écosystème Hadoop et leurs rôles respectifs, comprendre l'architecture des applicatifs hadoop et savoir quels sont les apports et les cas d'usage des solutions hadoop. Pythagore F.D. organise une formation « **Hadoop, l'écosystème** » **le 19 février 2018**

Le point essentiel des formations Pythagore F.D. est l'utilisation d'une infrastructure adaptée au BigData : le matériel pédagogique mis à disposition des participants permet de mettre en pratique tous les concepts abordés (clusters de calcul pour MapReduce, baies de stockage distribué, jeux de données pour les tests de performances). Des démonstrations illustrent chaque concept important, et les participants disposent chacun d'un cluster hadoop afin de tester les possibilités, l'architecture, et de bien comprendre les points forts d'une architecture hadoop, mais également les difficultés de mise en oeuvre et les précautions à prendre.

Le programme de la formation est le suivant :

Introduction

Présentation de NoSQL. Le théorème CAP.

Historique du projet hadoop

Les fonctionnalités : stockage, outils d'extraction, de conversion, ETL, analyse, ...

Exemples de cas d'utilisation sur des grands projets.

Les principaux composants :

HDFS pour le stockage et YARN pour les calculs.

Les distributions et leurs caractéristiques (HortonWorks, Cloudera, MapR, GreenPlum, Apache, ...)

L'architecture

Terminologie : NameNode, DataNode, ResourceManager

Rôle et interactions des différents composants

Présentation des outils d'infrastructure : ambari, avro, zookeeper;

de gestion des données : pig, oozie, tez, falcon, pentaho, sqoop, flume;

d'interfaçage avec les applications GIS;

de restitution et requêtage : webhdfs, hive, hawq, impala, drill, stinger, tajo, mahout, lucene, elasticSearch, Kibana

Les architectures connexes : spark, cassandra

Exemples interactifs

Démonstrations sur une architecture Hadoop multi-noeuds.

Mise à disposition d'un environnement pour des exemples de calcul

Travaux pratiques :

Recherches dans des données complexes non structurées.

Applications

Cas d'usages de hadoop.

Calculs distribués sur des clusters hadoop

Pythagore F.D. est un centre de formation, dans les domaines suivants :

- les bases de données et le BigData avec NoSQL, Cassandra, MongoDB, Hadoop, Spark, Storm ...
- la virtualisation et l'orchestration avec xen, kvm, lxc, Docker, et le cloud : cloudstack et openstack, openNebula, cobbler, etc ...

- TCP/IP (IPv6, snmp, Architecture, Sécurité, Administration de réseaux IP, VoIP, ...)
- Unix et Linux, et les applicatifs Apache, Openldap, Squid, Nagios, Zabbix, OCS/GLPI, puppet , chef...
- Développement (langage C, Java, Jee, technologies Jee, JBoss, WebServices, PHP, Perl, Python , ...)

Les formations sont dispensées soit dans les locaux de la société à Paris, soit sur site client, ou à distance en classes virtuelles.

Lieu : centre de formation Pythagore F.D., Paris9è

Pour toute demande d'information, contactez nos conseillers au 33 (0)1 55 33 52 10, ou par mail à l'adresse pdf@pythagore-fd.fr, ou sur le site www.pythagore-fd.fr