

## **BigData Concepts et Enjeux**

Pour comprendre les concepts et les apports du BigData, les impacts sur sur l'entreprise au niveau du traitement des données, des architectures, de l'organisation, Pythagore F.D. organise une formation «BigData Concepts et enjeux» **le 28 mai 2018**

Une des caractéristiques techniques des technologies BigData est l'aspect distribué, qu'il s'agisse du stockage des données ou du calcul.

Afin de mettre en évidence les possibilités offertes par une architecture distribuée, mais également les difficultés liées à la mise en oeuvre, Pythagore F.D. met à disposition, lors des ateliers, une infrastructure dédiée 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 Cassandra, accès Amazon pour la gestion de ressources dans le cloud, etc ...)

**Cette infrastructure est également mise à disposition des participants pour des tests ultérieurs, et la réalisation de POC lors de la validation d'architectures BigData.**

Le programme de la formation est le suivant :

### **Introduction**

A l'origine du BigData : traitement de volumes importants de données non structurées, traitements optimisés de flux de données au fil de l'eau, liés aux nouvelles technologies et aux nouveaux usages.

Domaines concernés : recherche scientifique, médical, e-commerce, sécurité, prédictif, ...

Exemples : lutte contre la criminalité, fraude, santé, ressources énergétiques

Apports des évolutions techniques sur différents aspects : stockage, indexation/recherche, calcul.

Concepts clés : ETL, Extract Transform Load, CAP, 3V, 4V, données non structurées, prédictif, Machine Learning.

Quelques applications : Watson (IBM), Amazon Rekognition

Le positionnement des technologies de cloud, BigData et noSQL, de data-mining. Eléments d'architecture.

Gouvernance des données :

importance de la qualité des données, fiabilité, durée de validité, sécurité des données

Aspects législatifs : sur le stockage, la conservation de données, etc ...

sur les traitements, la commercialisation des données, des résultats

### **Stockage distribué**

Caractéristiques NoSQL

Les différents modes et formats de stockage. Besoin de distribution.

Définition de la notion d'élasticité.

Principe du stockage réparti :

Définitions : réplication, sharding, gossip protocol, hachage,

Systèmes de fichiers distribués : GFS, HDFS, Ceph

Les bases de données : Cassandra, HBase, MongoDB, CouchBase, Riak, BigTable, ..

### **Calcul et restitution.**

Apport des outils de calculs statistiques

Langages adaptés aux statistiques, liens avec les outils BigData.

Outils de calcul et visualisation :

R, SAS, Spark, Tableau, QlikView, ...

Caractéristiques et points forts des différentes solutions.

### **Evolutions**

Liens vers les nouveaux métiers : Hadoop scientists, Data scientists, CDO.

Analyse des données au service de l'entreprise

Rôle de la DSI dans la démarche BigData.

Ouverture sur l'OpenData : principe, la démarche publique, les licences.

Exemple : portail data.gouv.fr

Les offres Saas BigData comme Google BigQuery.

Les limites. Les nouveautés annoncées.

---

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](mailto:pdf@pythagore-fd.fr), ou sur le site [www.pythagore-fd.fr](http://www.pythagore-fd.fr)