Storm : flux de données

Lien : https://innov-systems.com/formation/storm-flux-de-donnees

DURÉE
3 jours (21h)

RÉFÉRENCE
BSI18

CATÉGORIE
Big Data : développement, langages et NOSQL

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Mettre en oeuvre Storm pour le traitement de flux de données
- Comprendre les fondamentaux du développement d'applications Big Data en temps réel
- Assurer le traitement des grosses quantités de données en temps réel
- Disposer des connaissances nécessaires pour configurer un environnement de développement

POUR QUI ?

- Chefs de projets
- Data scientists
- Développeurs

Innov Systems

Programme détaillé

1 / Introduction à l'architecture temps réel

- Traitements temps réel
- Les architectures LAMDA
- Les architectures KAPPA
- Les architectures SMACK

2 / Introduction à Storm

- Problématiques du Big Data
- Le Cloud Computing et le Big Data
- Présentation de Storm : fonctionnalités, architecture, langages supportés
- Définitions : spout, bolt, topology
- Apache Storm dans le projet Hadoop
- Storm et Spark Streaming : différences des modes de traitement de données
- Du Map Reduce aux topologies de Storm
- Développement et déploiement de topologies
- Modèle de données (clé, valeur)

4 / Architecture apache Storm

- Étude des composants d'un cluster Storm : master node "nimbus" et worker nodes
- Positionnement par rapport à un cluster Hadoop
- Le modèle de données
- Différents types de flux
- Rôles des Nimbus et ZooKeeper
- La relation entre Nimbus et Supervisors
- Le cluster intermédiaire de Zookeepers et leurs relations avec les Nimbus et Supervisors
- Définition de l'environnement de développement

Innov Systems

- Création de projets basés sur Storm
- Les composants primitifs : Spouts et Bolts
- Comprendre le cycle de vie du déploiement d’une topologie
- Paramétrer la communication interne de Storm

5 / Premiers pas avec Storm

- Configuration d'un environnement de développement
- Installation d'un cluster Storm
- Le projet storm-starter

6 / Gestion des flux de données

- Définition des flux
- Les différentes stratégies de distribution de données
- Implémenter des stratégies personnalisées
- Réaliser des jointures sur les flux
- Création de topologies regroupants des flux entre différents noeuds
- Communication entre flux en JSON
- Lecture de flux d'origines diverses (JMS, Kafka, ...)

7 / Haute disponibilité

- Tolérance aux pannes : principe de fiabilisation des master node, workers node, nimbus
- Garantie de traitement des flux : principe, paramètres TOPOLOGY_MESSAGE_TIMEOUT_SECS, TOPOLOGY_ACKERS
- Traitements temps réel avec Trident
- Scalabilité horizontale : Notion de taux de disponibilité, Ajouter des serveurs d’application web , Scalabilité des bases de données

8 / Parallélisme des topologies

- Schéma d’exécution d’une topologie (Worker, Executor, Task)
- Maîtriser et provisionner la parallélisation d’une topologie

Inov Systems

- Techniques de « Rebalancing »

9 / Fiabilité de traitement des messages

- Le mécanisme de suivi des messages
- Le cycle de vie d’un message
- Traiter les messages avec l’API Java
- Définir la fiabilité avec l’API Storm
- Assurer la fiabilité d’une application : méthodes et stratégies

10 / Gérer l’exécution des topologies

- Utiliser les Storm Metrics
- Supervision avec les commandes Storm CLI
- Storm UI

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Innov Systems

Prochaines dates programmées

10 au 12 Fév. 2027	Présentiel - Casablanca
07 au 09 Avr. 2027	Présentiel - Casablanca
02 au 04 Juin 2027	Présentiel - Casablanca
04 au 06 Août 2027	Présentiel - Casablanca
06 au 08 Oct. 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés

Innov Systems