

### THÈMES – CONSIGNES ET CONSEILS

|          |                                      |          |
|----------|--------------------------------------|----------|
| <b>1</b> | <b>Storage</b>                       | <b>2</b> |
| <b>2</b> | <b>DataRouting</b>                   | <b>3</b> |
| <b>3</b> | <b>Présentation</b>                  | <b>4</b> |
| <b>4</b> | <b>Principe d'API</b>                | <b>5</b> |
| <b>5</b> | <b>API Management et exemples</b>    | <b>6</b> |
| <b>6</b> | <b>Batch</b>                         | <b>7</b> |
| <b>7</b> | <b>Stream</b>                        | <b>8</b> |
| <b>8</b> | <b>Device Management et Sécurité</b> | <b>9</b> |

### Introduction et consignes générales

Dans ce document, vous trouverez, en plus des **slides présentés** lors du premier cours, quelques mots-clés, indications et conseils pour réaliser votre cours sur le thème sélectionné.

Vous trouverez notamment chaque fois :

**Format** Les outils pédagogiques (présentation, exercices, démonstrations, etc) que je vous conseille d'utiliser pour que ce cours soit le plus pertinent possible.

**Sujets** La liste des sujets à *au moins* évoquer durant votre présentation. Ces sujets servent à la fois de guide pour votre présentation, mais aussi de contrainte pour m'assurer que vous aborderez des notions qui me semblent importantes.

Cependant, si, en préparant votre cours, l'un des sujets ne vous semble pas pertinent, n'hésitez pas à m'en parler !

**Mots-clés** Les mots à taper sur votre moteur de recherche préféré (**Qwant**, **Ecosia**, **Lilo...**) afin de trouver des contenus, des références, des articles, des sujets de travaux pratiques, etc.

En termes d'organisation, je vous rappelle que :

- Vous devez préparer ce cours avec **l'ensemble de l'équipe**.

Vous pouvez vous séparer le travail, les parties, ou tout faire ensemble. Comme bon vous semble. Vous devez cependant tous être capables de présenter n'importe quelle partie du cours.

Si vous constatez, ou si je constate, une maîtrise trop hétérogène du sujet entre les membres de l'équipe, une évaluation individuelle pourrait être attribuée.

- Vous devez me présenter, **au moins une semaine avant** :

1. Le plan détaillé de votre cours
2. Le déroulé envisagé (présentation, questions, démonstrations, petits exercices d'animation, etc)
3. Les sources principales et votre compréhension de ces sources
4. 2 questions QCM qui vous sembleraient les plus pertinentes pour votre cours (qui doivent révéler les notions les plus importantes pour vous).

- Vous pouvez **prendre rendez-vous avec moi autant de fois que vous le souhaitez** (et que je peux l'assurer... en fonction du délai et de votre organisation pour me proposer ces rencontres...).

## PARTIE 1: STORAGE

### 1.1 Format

Présentation, éventuellement démonstration (notamment pour bases Graph par exemple).

### 1.2 Sujets

- Contexte du stockage de données, petit historique.
- Différents types de bases de données (SQL et NoSQL) et leurs contraintes/avantages.
- Focus minimum sur les BDD relationnelles, Documents et Graphs avec exemples d'usages différents.
- Overview des fournisseurs de solutions libres, en ligne, etc.
- Expliciter les architectures (répartition, redondance, backup, etc).
- Mettre en avant les BDD TimeSeries et leurs intérêts.
- Focus spécifique obligatoire sur MongoDB (architectures, fonctionnement, replica, sharding, principe de clé de répartition, équilibrage, backup, etc).

### 1.3 Mots-clés

BDD, SQL, NoSQL, MySQL, MongoDB, Cassandra, Neo4j, HDFS, HBase, OpenTSDB, InfluxDB, Redondance, Clé de répartition, Sharding,

## PARTIE 2: DATAROUTING

### 2.1 Format

Présentation et démonstration de **Nifi** au travers de divers exemples.

### 2.2 Sujets

- Principe du Data Routing et des outils de gestion/orientation de flux de données.
- Présentation de Nifi, principes, histoire, architecture, déploiement, limites, etc.
- Différents exemples d'usage de Nifi avec orientation de flux conditionnel. Utiliser notamment les processeurs HTTP, FTP, PutFile.
- Exemple d'erreur de flux, gestion d'alerte, envoi de mails.
- Exemple avec une queue de messages.
- Mettre en avant la gestion des droits d'accès, le principe de templates, l'usage de l'API, le backup d'architecture de flux.
- Expliquer les différences avec MiNifi.

### 2.3 Mots-clés

Nifi, MiNifi, Data Routing, Tolérance aux fautes, gestion de flux de données, dataflow management, orientation de flux de données,

## PARTIE 3: PRÉSENTATION

### 3.1 Format

Présentation, démonstration d'une solution.

### 3.2 Sujets

- Intérêts des dashboard, exemples d'usages et limites.
- Principe de personnalisation de l'affichage au contexte, au client, à l'expérience et aux compétences.
- Principe de Data Visualisation (différents types de représentations de données pour différents objectifs).
- Présentation des différents acteurs de leurs avantages/contraintes.
- Démonstration d'un acteur avec différents types de visualisations, de filtres, de navigation, etc.

### 3.3 Mots-clés

Data Visualisation, DataViz, Dashboard, Data Storytelling, PowerBI, QuickSight, Data Studio, Grafana, Kibana, D3JS, Time Series, données temporelles, données spatiales, compteur, historique de données, principes de design et visualisation, actionneurs, filtres, temps de réponse, droit d'accès, croisement de données

## PARTIE 4: PRINCIPE D'API

### 4.1 Format

Présentation avec démonstration de l'usage de Swagger/OpenAPI.

### 4.2 Sujets

- Intérêts et histoire du principe d'API.
- Principes de RESTful et détail des requêtes HTTP (GET/PUT/POST/DELETE).
- Principe de Stateless / Stateful.
- Normalisation avec OpenAPI et génération automatique de code.
- Importance de la documentation pour un usage répandu de son API.
- Gestion des versions d'une API.
- Démonstration de la création d'une API par définition d'une OpenAPI.
- Démonstration rapide de la création d'un serveur à partir de la génération de code basée sur une OpenAPI
- Différents frameworks pour la création d'API.

### 4.3 Mots-clés

Swagger, OpenAPI, CodeGen, Flask, NodeJS, Java, Documentation, RESTful, GET, PUT, POST, DELETE, JSON

## PARTIE 5: API MANAGEMENT ET EXEMPLES

### 5.1 Format

Présentation.

### 5.2 Sujets

- Principes et intérêts du management d'API.
- Gestion d'accès à une API (droits, quotas, etc).
- Principe de token/clé, révocation, authentification (et exemple d'usage avec éventuellement diagramme de séquence, etc).
- Importance de l'orchestration et de la répartition de charge.
- Monitoring d'usage, de performance, etc.
- Différents acteurs et options (Cloud, OpenSource, etc).

### 5.3 Mots-clés

API management, quotas, Kong, API Gateway, Kong, Apigee, Tyk.io, tokens, api key, orchestration, répartition de charge

## PARTIE 6: BATCH

### 6.1 Format

Présentation avec exemples concrets (voir code et démonstration).

### 6.2 Sujets

- Intérêt du traitement de données en Batch et limites.
- Plateformes proposant la fonctionnalité, langages de programmation, avantages et limitations.
- Cas d'usage précis avec exemples (éventuellement de code et démonstration).

### 6.3 Mots-clés

Hadoop Map Reduce, Spark, Spark Streaming (Micro-Batch), historique de données

## PARTIE 7: STREAM

### 7.1 Format

Présentation avec exemples concrets (voir code et démonstration).

### 7.2 Sujets

- Intérêt du traitement de données en Stream et limites.
- Plateformes proposant la fonctionnalité, langages de programmation, avantages et limitations.
- Cas d'usage précis avec exemples (éventuellement de code et démonstration).

### 7.3 Mots-clés

Stream, Storm, Flink, Spark Streaming, Alerting



## PARTIE 8: DEVICE MANAGEMENT ET SÉCURITÉ

### 8.1 Format

Présentation.

### 8.2 Sujets

- Principe de Cloud Privé/Public/Hybride.
- Problématique de la gestion de la vie d'un objet connecté (déploiement, monitoring, mise à jour, etc).
- Sécurité des objets connectés et de leur accès à distance (différents types d'attaques, etc).
- Sécurité du récepteur de données (authentification de l'objet, différents types d'attaques, etc).
- Droit sur les données, leur stockage, leur juridiction, RGPD, etc.
- Principe d'anonymisation et de pseudonymisation, limites et contraintes.

### 8.3 Mots-clés

Sécurité, anonymisation, pseudonymisation, RGPD, juridiction, authentification, mise à jour, cycle de vie, monitoring, DDoS