

# Projet Intégrateur

Sujet de Thalès



Duc Hau NGUYEN  
Cindy PONIDJEM  
Anaïs RABARY  
Ezechiél TAPE

**1** Contexte

**2** Choix d'implémentation

**3** Parties Réalisées

**4** Organisation de l'équipe

**5** Démonstration

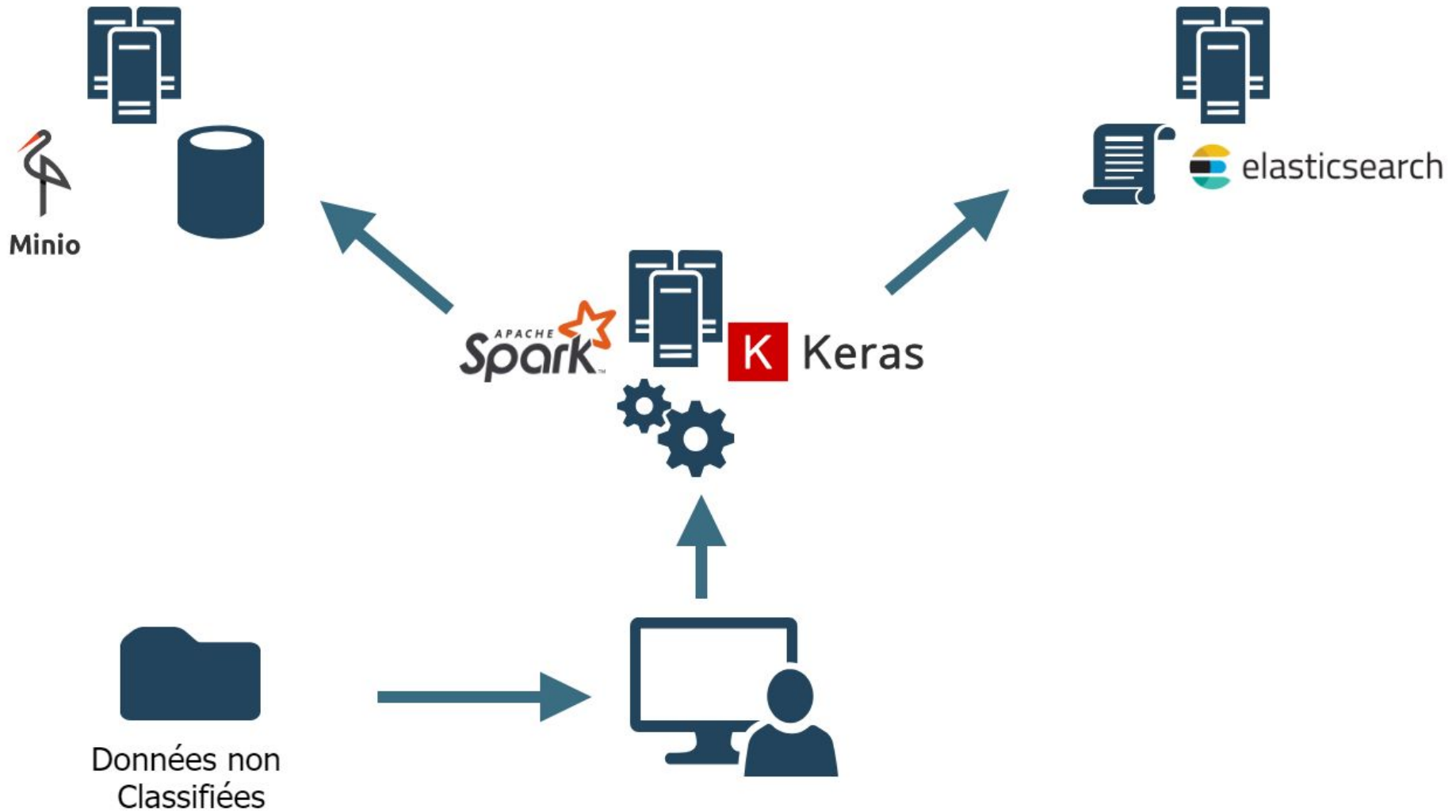
# 1 / Contexte

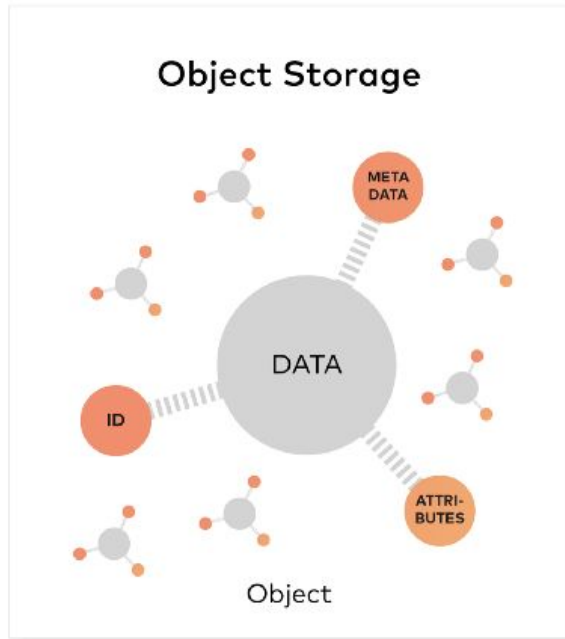
- **Contexte :**
  - Entreprise produisant et gérant plusieurs **images**.
  - Plusieurs traitements appliquées sur ces images.
  - Besoin de **système de stockage distribué**.
- **Objectifs:**
  - Mise en place d'un **système de stockage distribué**.
  - Utilisation de techniques d'apprentissage automatique pour le **traitement des images**.
  - Utilisation d'outil de calcul distribué.



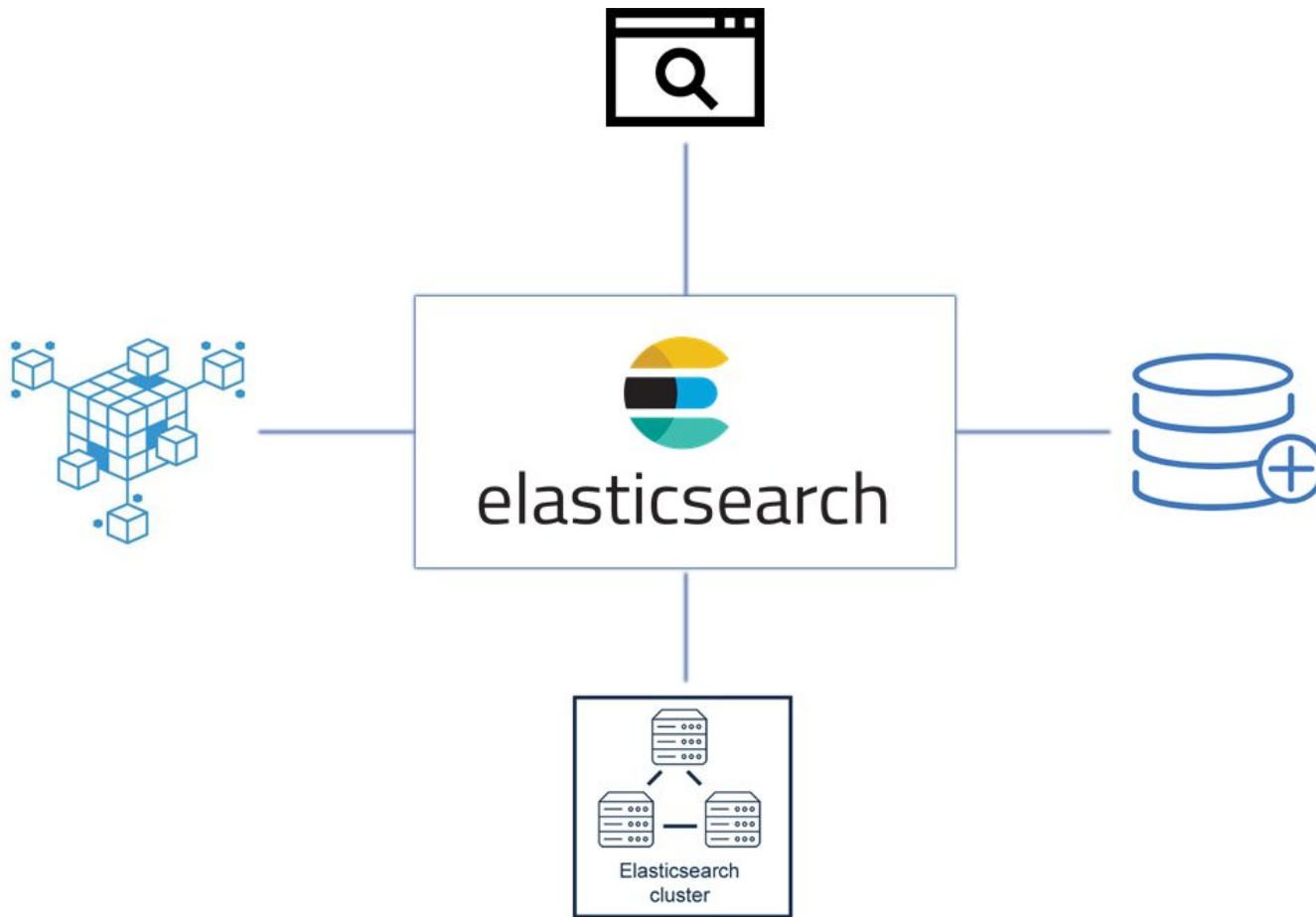
## 2 / Choix d'implémentation

# Choix d'implémentation





- Serveur de stockage orienté **objet**
- **Légèreté**: peut être packager aisément
- Stocker chaque **image prédite** au format .npy





# Zoom sur Spark - Keras

Traitement distribué  
sur des données  
massives

Modèles de  
Réseaux de  
neurones profonds,  
convolutionnels,



Les données dans  
des RDD, partitionés  
pour le traitement  
distribué

Encapsulation du  
modèle Keras dans  
Spark

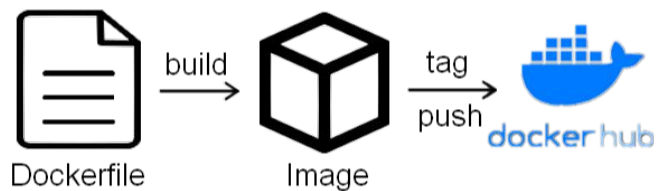


## 3 / Parties Réalisées

## Mise en place de l'infrastructure de base

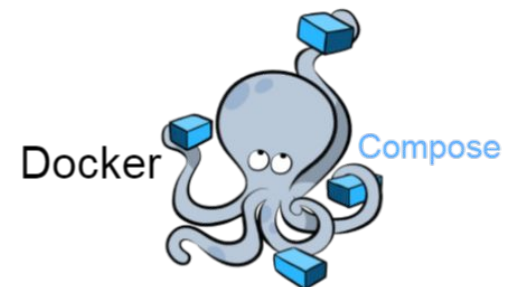
### Etape 1

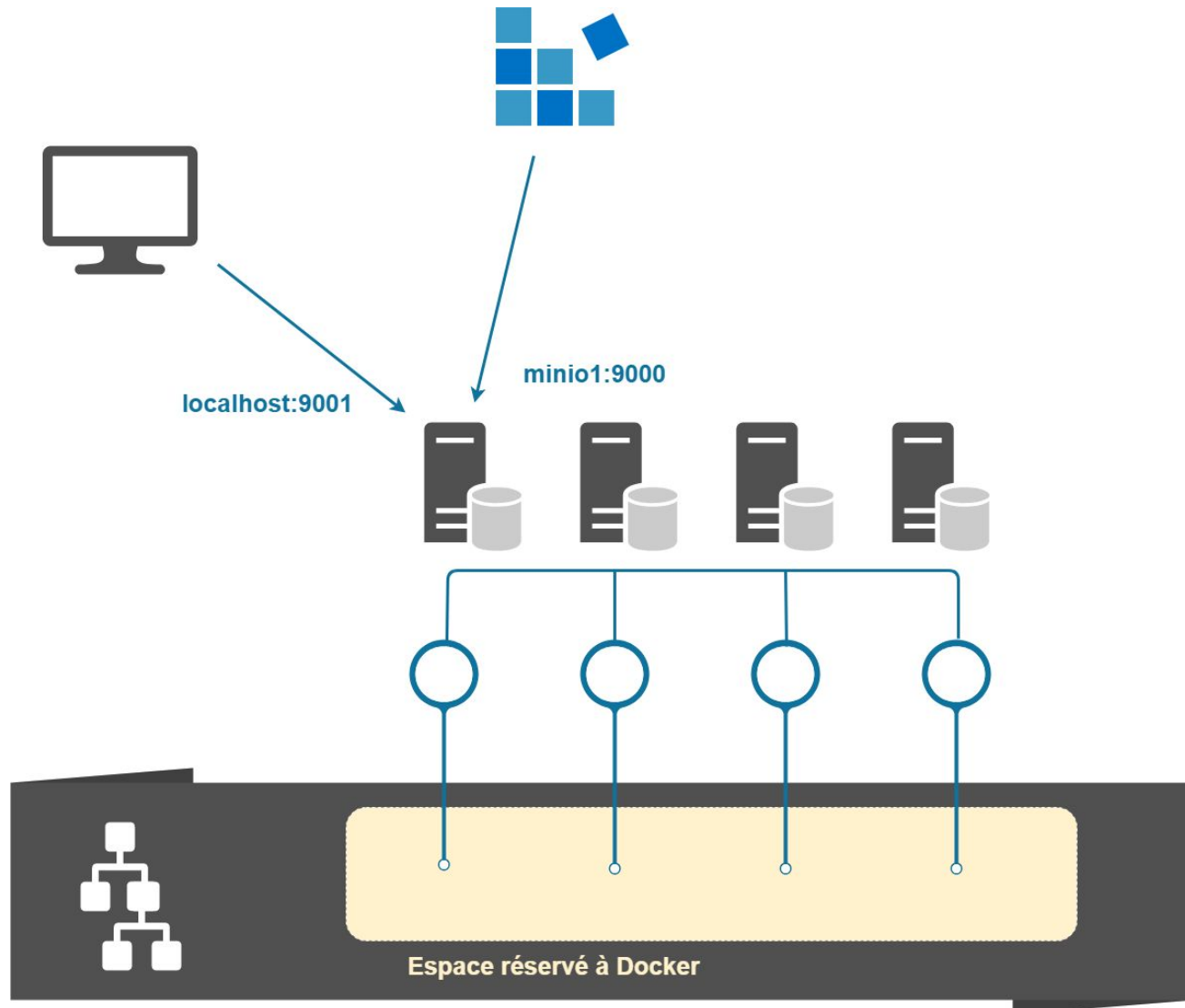
Préparer le service uploader

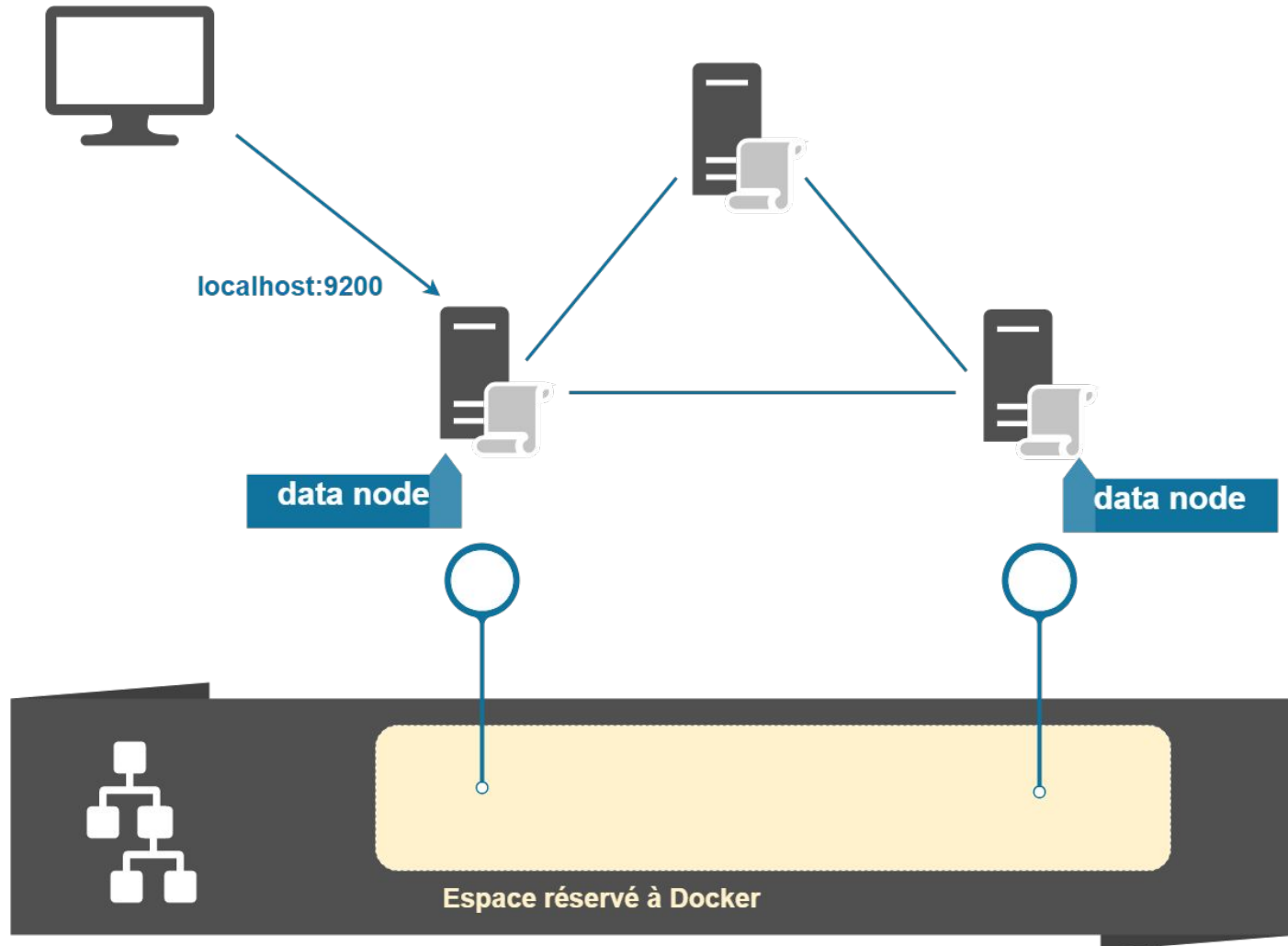


### Etape 2

Déployer les différents services







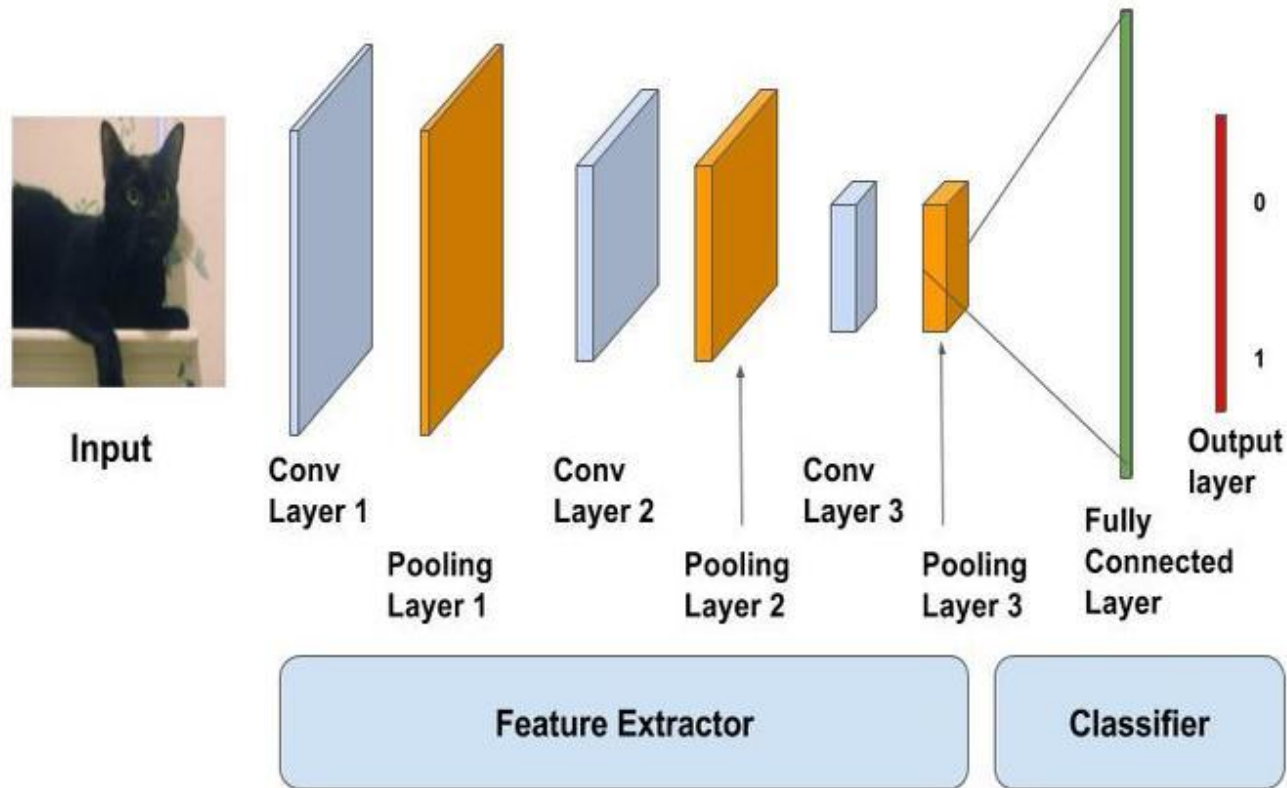
# Réalisation : Classification avec CNN

**Neural Network ou KNN ?**

**Multilayered Perceptron ou CNN ?**

**Architecture du CNN ?**

# Réalisation : Classification avec CNN



# Réalisation : Classification avec CNN

## Transfer learning

### Archi existant

- DenseNet121
- Pré-entraîné sur ImageNet

### Adapter

- Ajouter une couche **Dense** de taille (1024 x 5)
- Fonction d'activation "**softMax**"

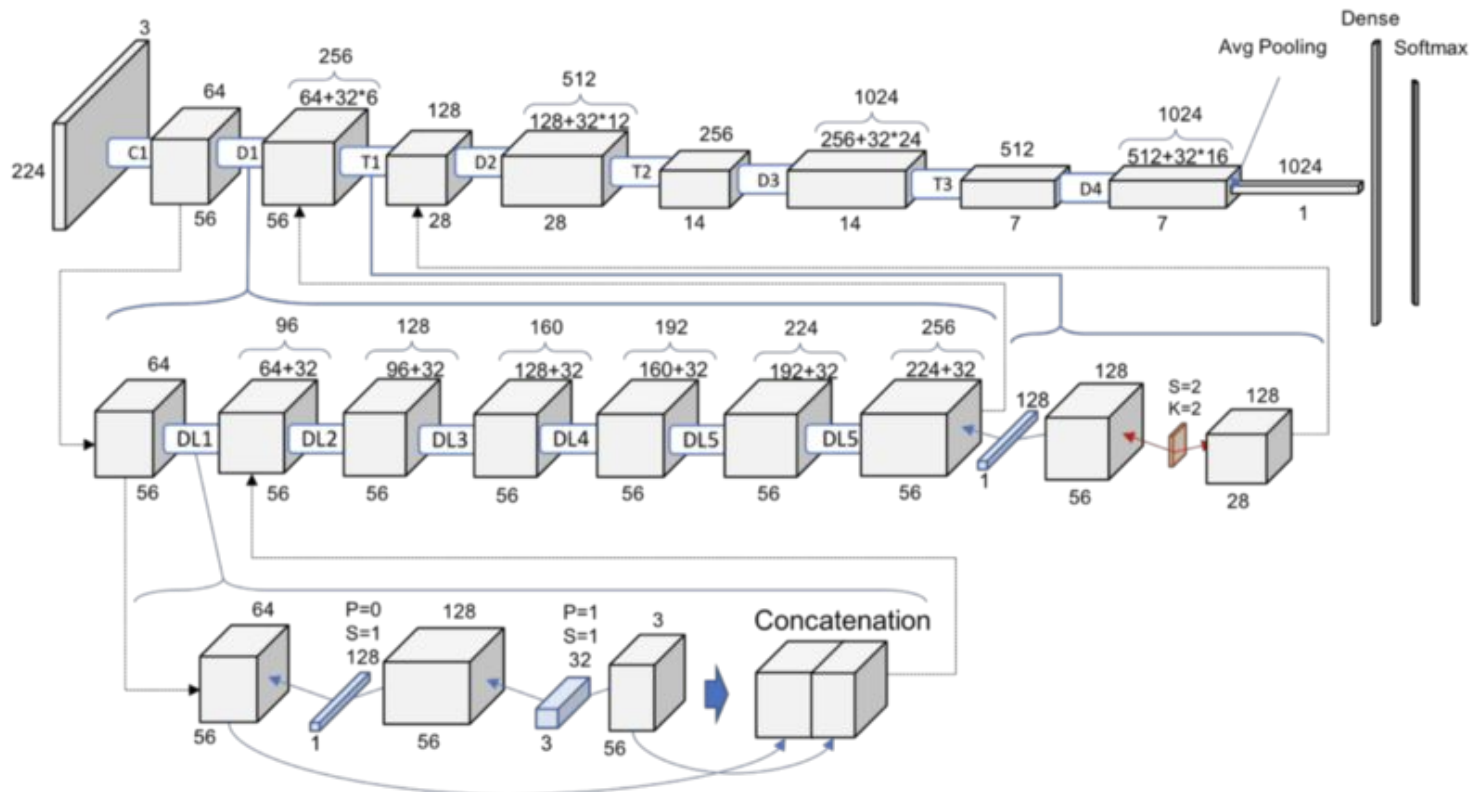
### Ajuster

- Entraîner sur données
- N'apprendre que le sur couche



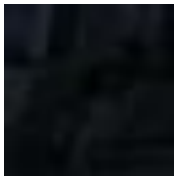
# Réalisation : Classification avec CNN

## DenseNet121



# Réalisation : Classification avec CNN

DenseNet121



32 X 32

32 X 1024

1024 X 5

softmax

# Réalisation : Classification avec CNN

## Résultat

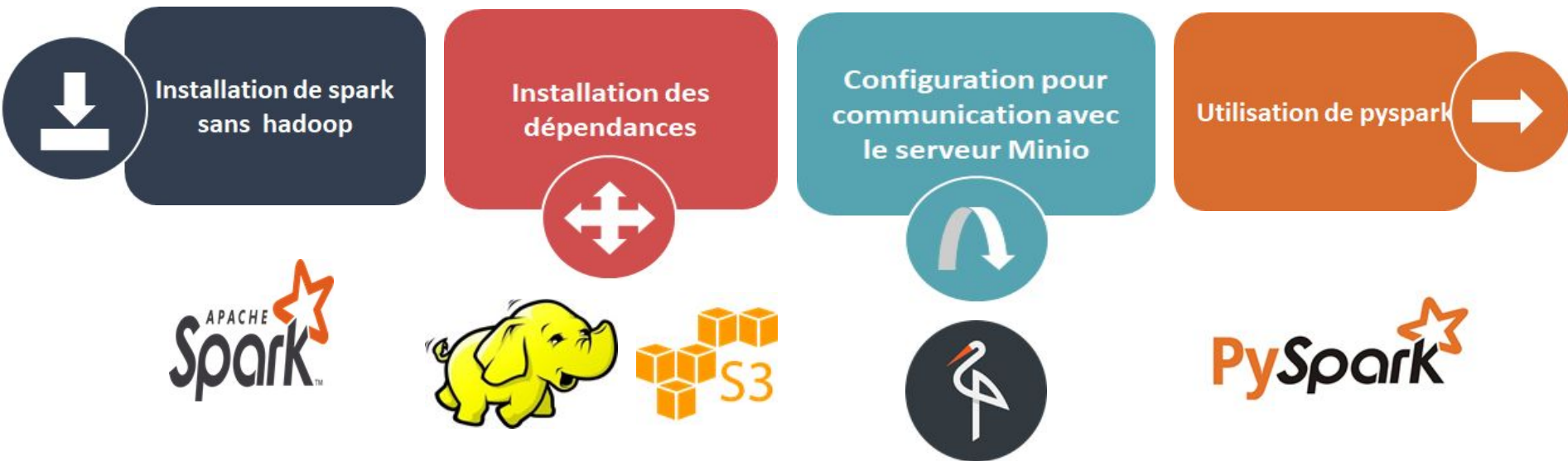
	Modèle basique	Modèle DenseNet
Temps d'entraînement (par epoche)	18 mins 23	11 mins 23
Meilleurs accuracy	72%	76%

**Donnée entraînement:** 163989

**Donnée validation:** 41011

**Environnement:** Ubuntu, 4 coeurs, CPU

# Réalisation : Spark Côté Architecture



# Réalisation : Spark Côté Apprentissage

**Nettoyer les données**

**Entraîner le modèle**

**Evaluer**

**Prédire**

## 4 Organisation de l'équipe

## Répartition de mission



Cindy, Ezechiël



### Infrastructure

Architecture distribuée **Minio**, **Elastic Search**.  
Plateforme **Spark** pour faire tourner l'algorithme  
de classification

**SDCI**



Duc Hau, Anaïs



### Machine Learning

Recherche algorithme classification.  
Implémentation et adaptation avec  
l'environnement distribué

**AP**

## Méthodologie



### Individuel

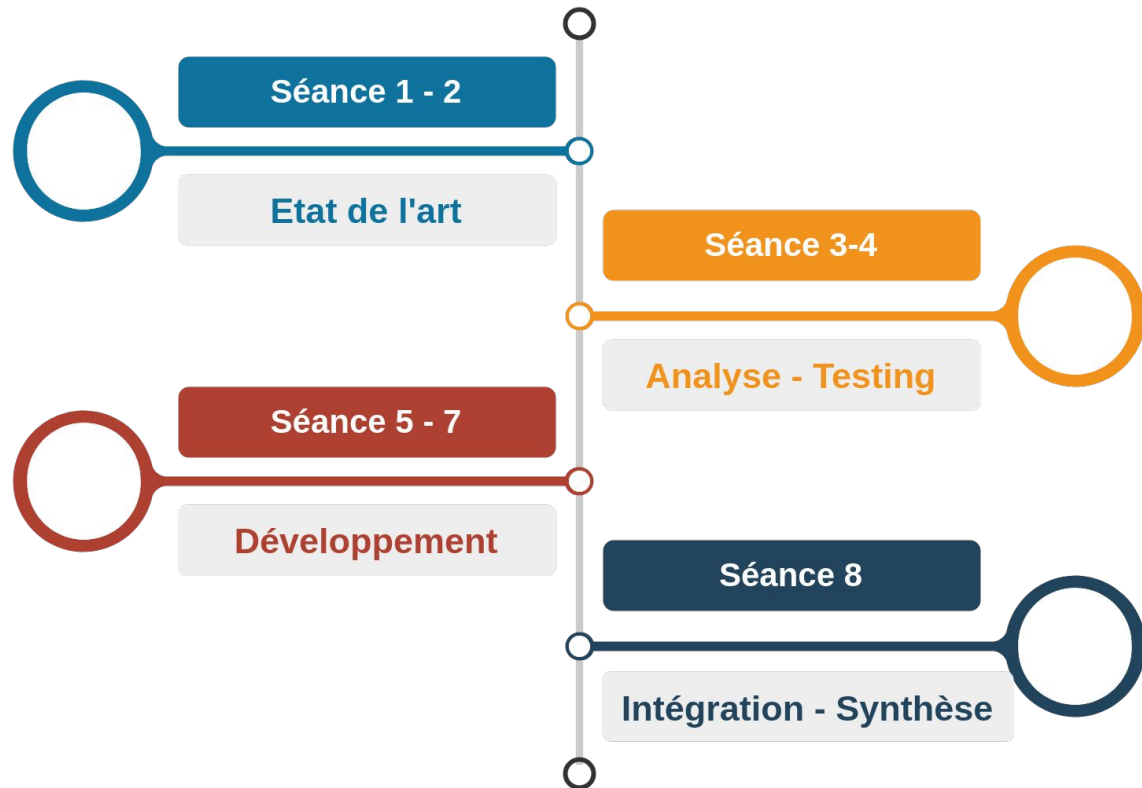
Recherche sur le thème  
Tester en local



### Synchronisation

Reporter l'avancement  
Fixer l'objectif suivant  
Synthétiser questions

## Démarche





## 5 / Démonstration



**Duc Hau NGUYEN**  
**Cindy PONIDJEM**  
**Anaïs RABARY**  
**Ezechiél TAPE**