

Gestion de Ressources d'Hadoop



Hadoop MapReduce: Limites

73

- Le **Job Tracker** s'exécute sur une seule machine, et fait plusieurs tâches (gestion de ressources, ordonnancement et monitoring des tâches...);
 - Problème de scalabilité et de goulot d'étranglement : les nombreux datanodes existants ne sont pas exploités: max de 5000 noeuds et 40,000 tasks s'exécutant simultanément (Yahoo).
- Si le **Job Tracker** tombe en panne, tous les jobs doivent redémarrer;
 - Problème de disponibilité: SPoF (*single point of failure*)
- Le nombre de map slots et de reduce slots est prédéfini:
 - Problème d'exploitation: si on a plusieurs map jobs à exécuter, et que les **map slots** sont pleins, les **reduce slots** ne peuvent pas être utilisés, et vice-versa.
- Le **Job Tracker** est fortement intégré à Map Reduce!
 - Problème d'interopérabilité: impossible d'exécuter des applications non-MR sur HDFS

Solution: Hadoop 2.0

76



Hadoop v1.0

MapReduce
Data Processing
& Resource Management

HDFS
Distributed File Storage



Hadoop v2.0

MapReduce

**Other Data
Processing
Frameworks**

YARN

Resource Management

HDFS

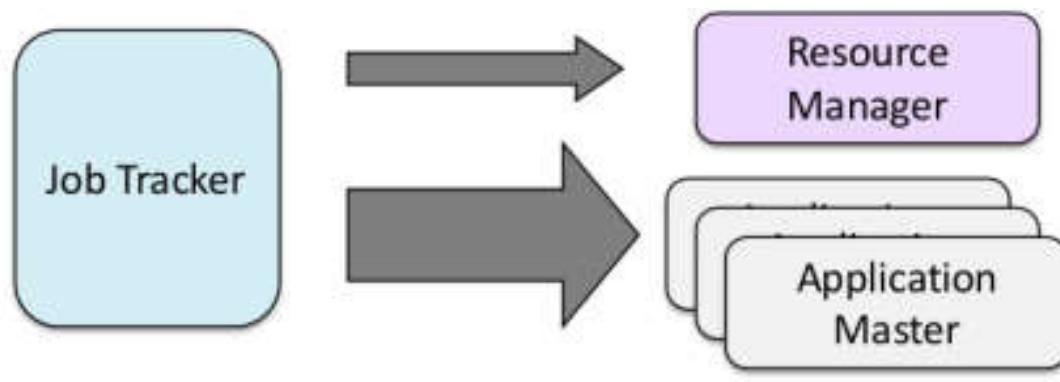
Distributed File Storage

BigData

YARN: C'est quoi?

77

- **Yarn: Yet Another Resource Negotiator.**
- Un framework permettant d'exécuter n'importe quel type d'application distribuée sur un cluster Hadoop, pas uniquement les applications MapReduce.
- Pas de notion de slots: Les noeuds ont des ressources (CPU, mémoire..) allouées aux applications à la demande.
- **Idée clé:** Séparer la gestion des ressources de celle des tâches MR.



BigData

YARN- Les démons

78

□ Resource Manager (RM)



- Tourne sur un nœud master
- Ordonnanceur global des ressources
- Permet l'arbitrage des ressources entre plusieurs applications.

□ Node Manager (NM)

- S'exécute sur les nœuds slaves
- Communique avec RM

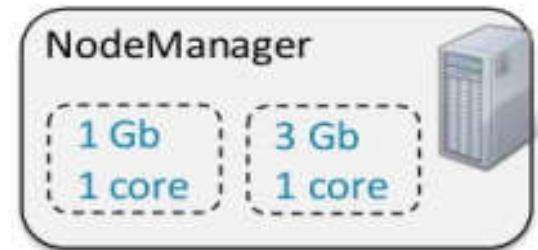


YARN- Les démons

79

❑ Containers

- ❑ Crées par le RM à la demande
- ❑ Se voit allouer des ressources (mémoire, CPU) sur le noeud slave: Il n'y a plus de slots prédéfinis.
- ❑ Les applications tournent sur un ou plusieurs containers.

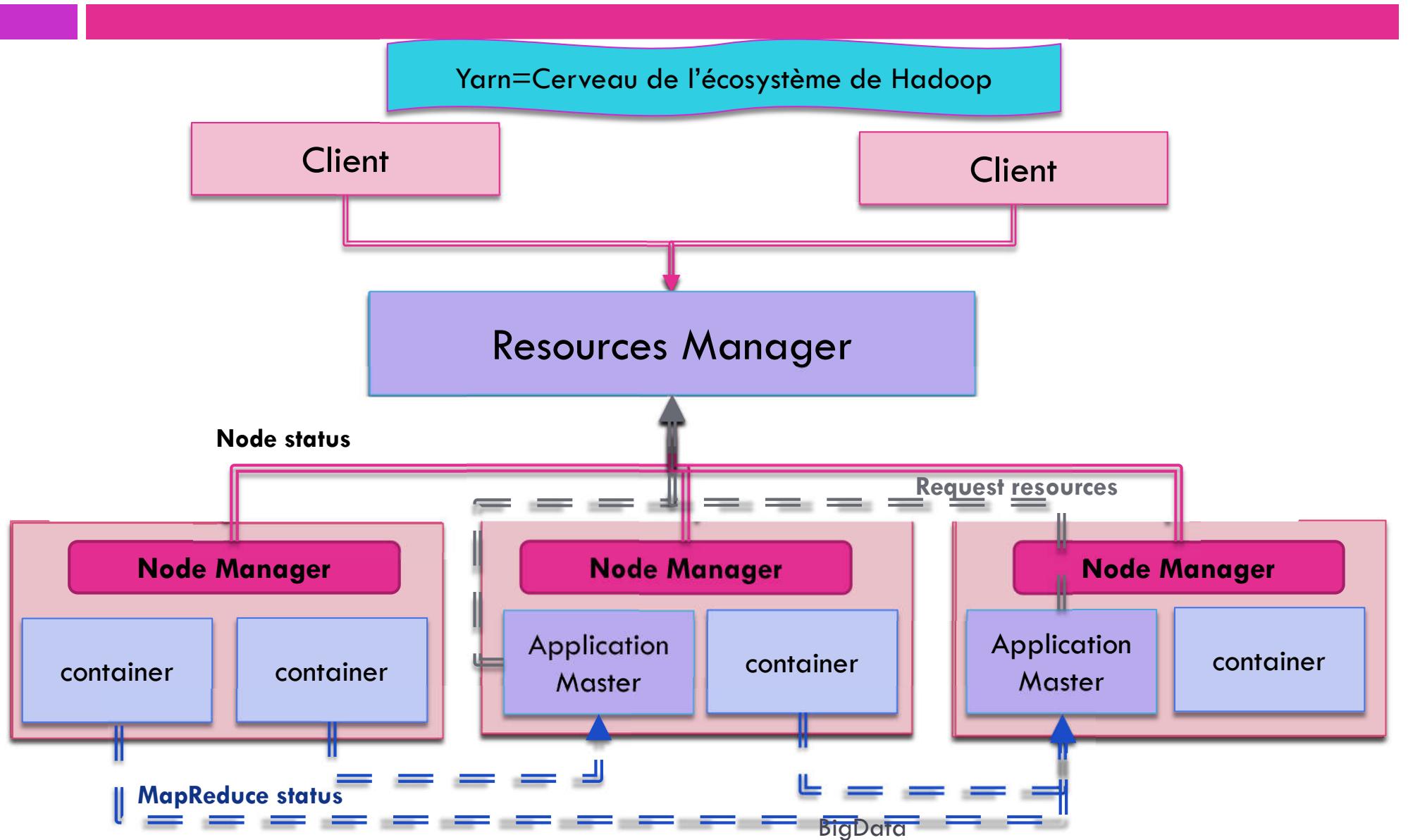


❑ Application Master (AM)

- ❑ Un seul par application
- ❑ S'exécute sur un container
- ❑ Demande plusieurs containers pour exécuter les tâches d'applications.

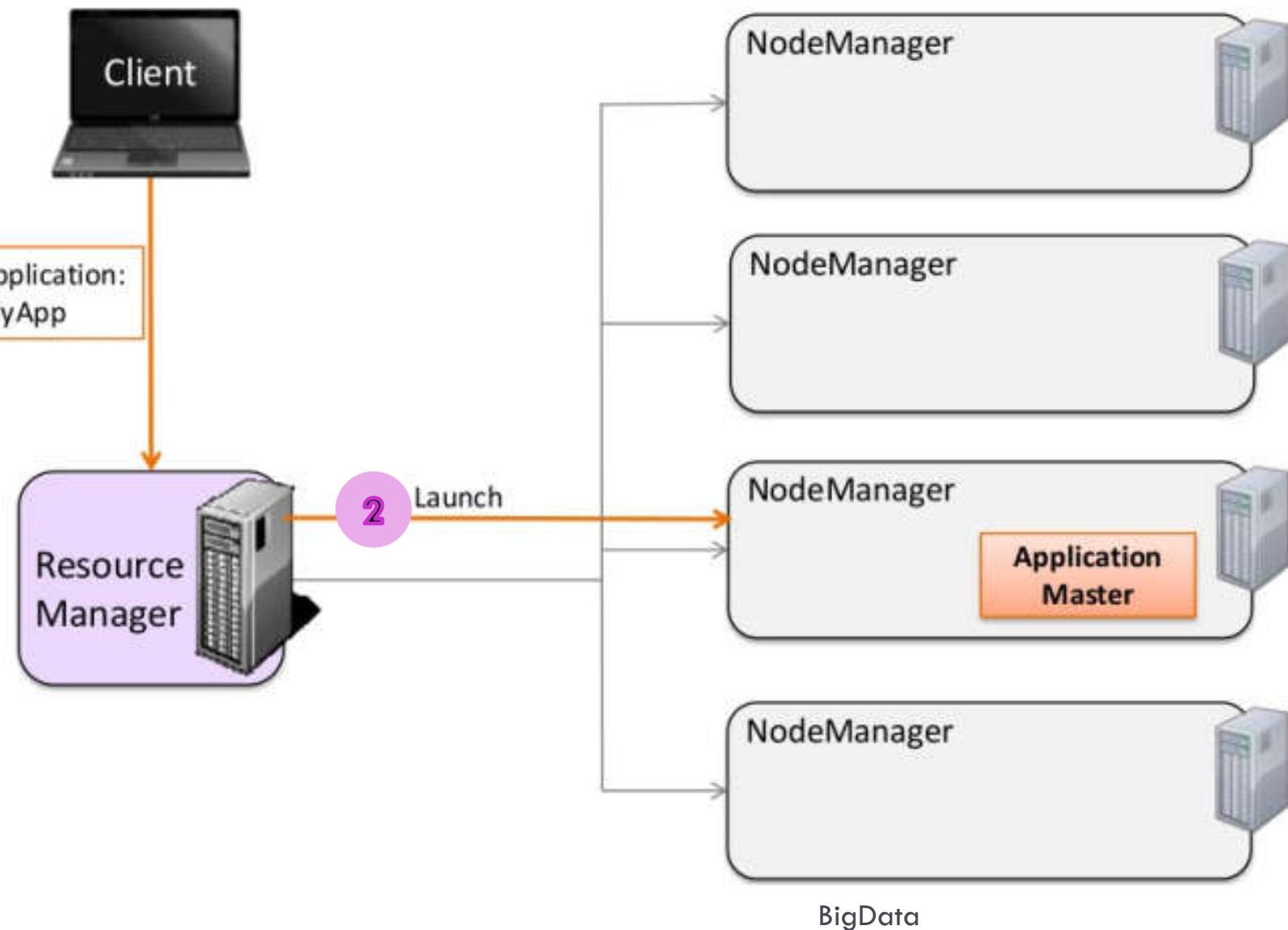


YARN- Fonctionnement



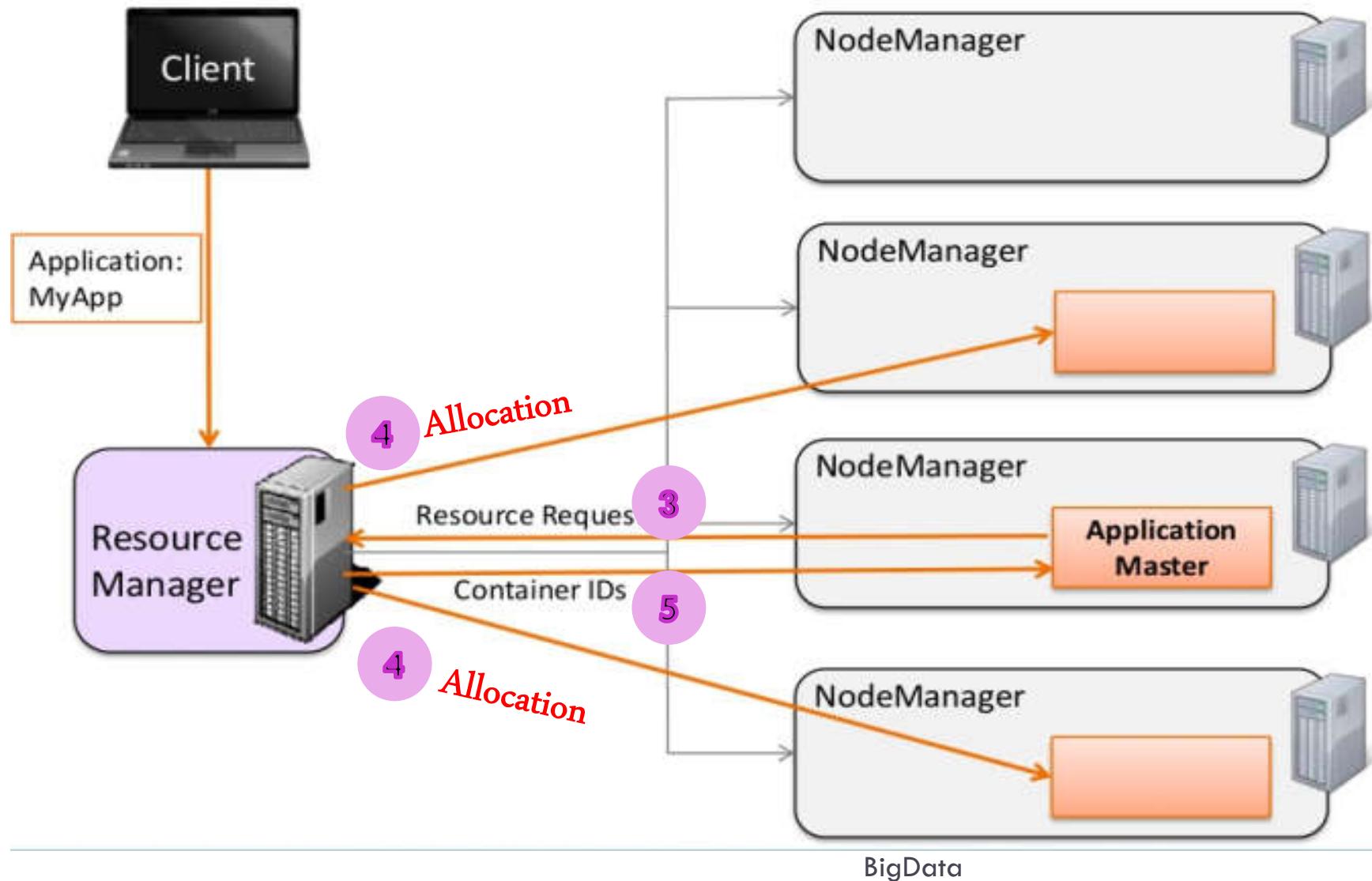
Cluster YARN: Exécution d'une application

81



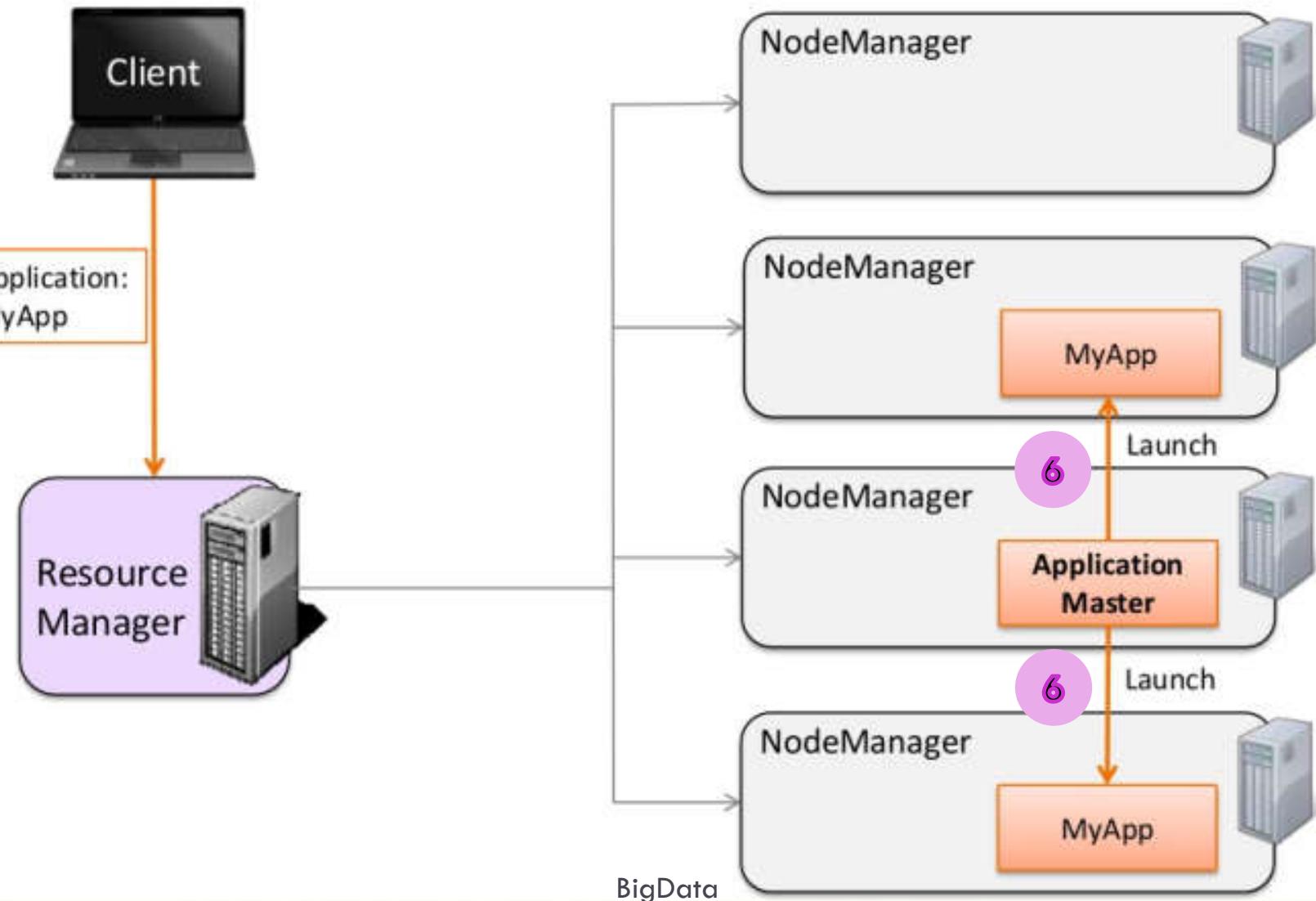
Cluster YARN: Exécution d'une application

82



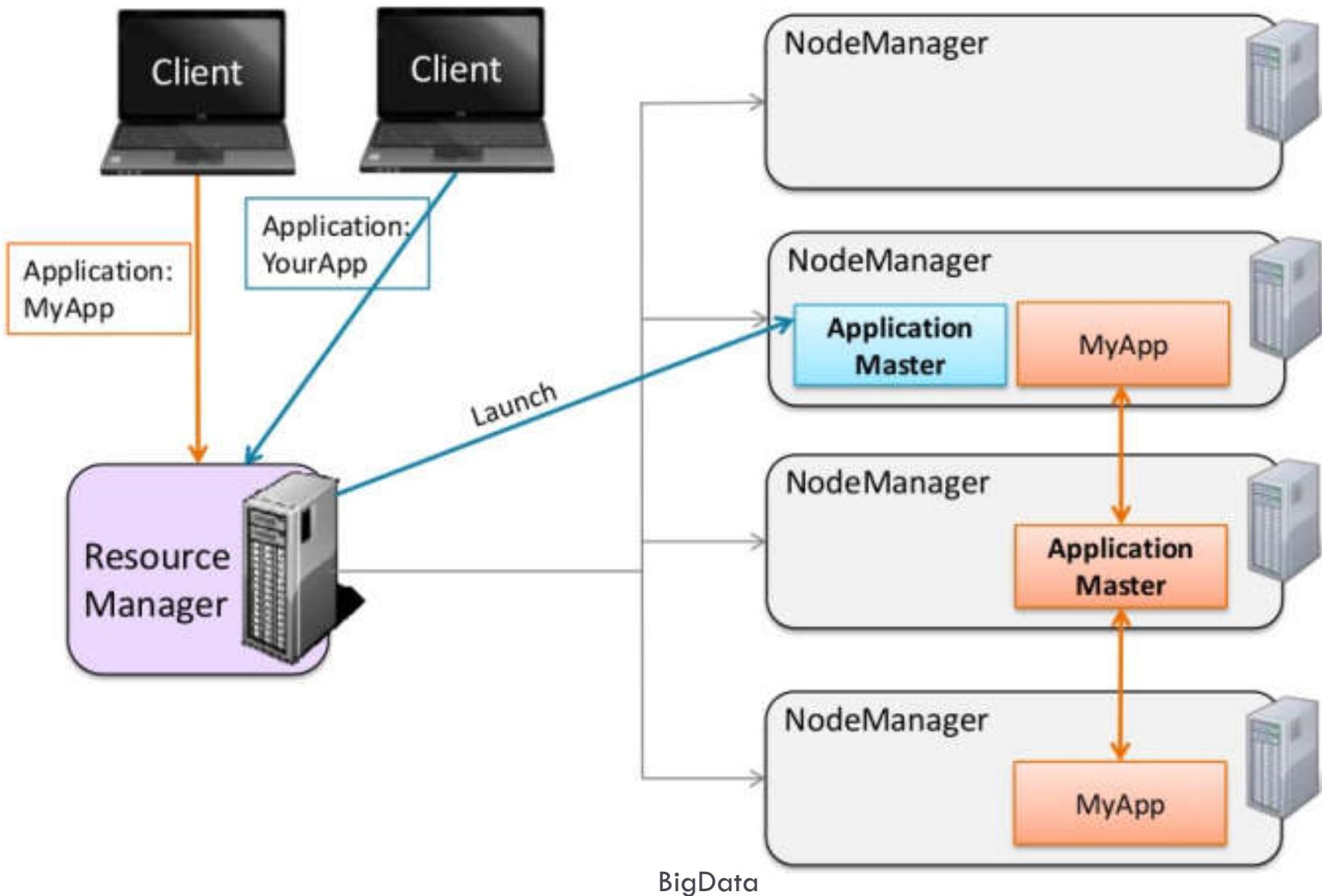
Cluster YARN: Exécution d'une application

83



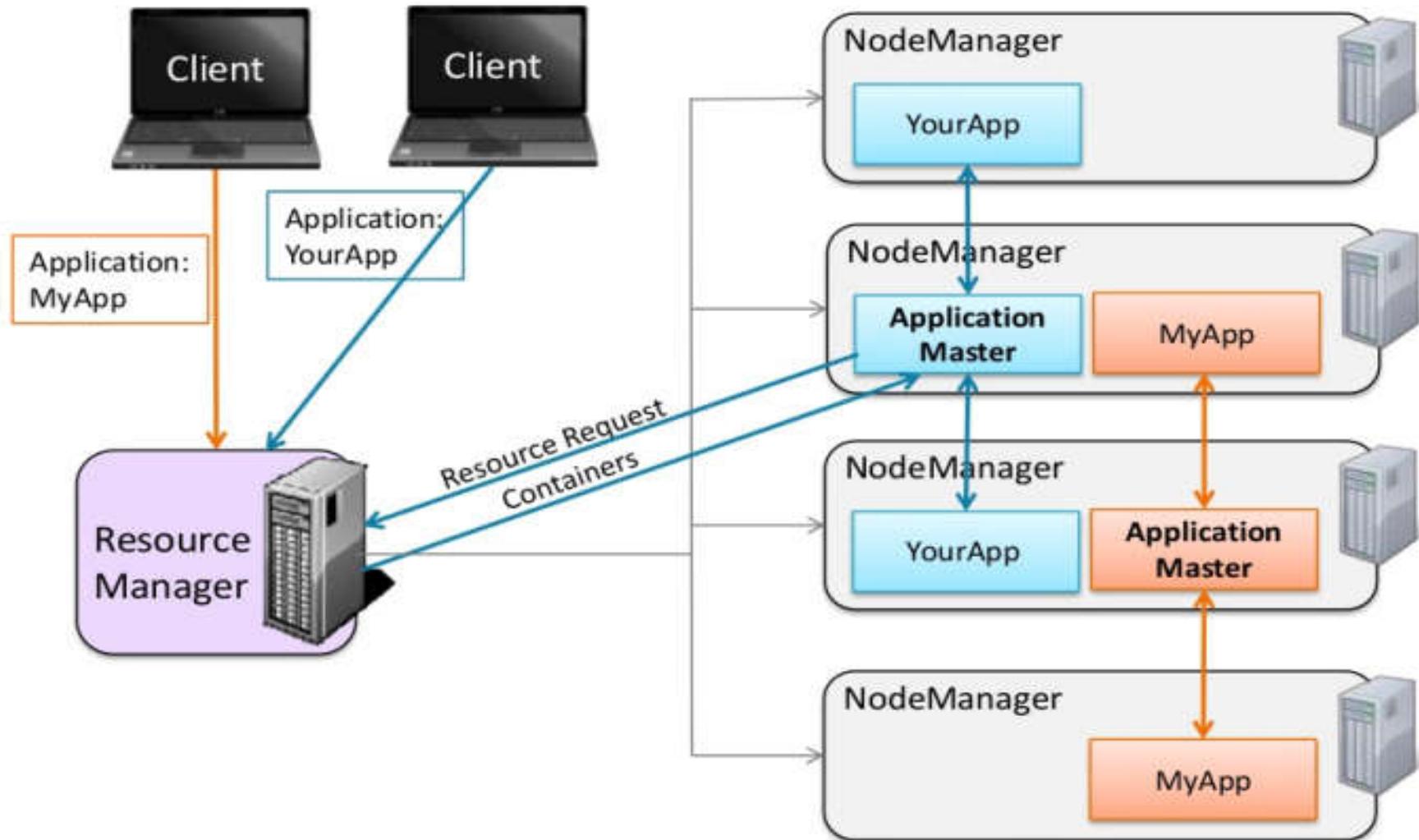
Cluster YARN: Exécution d'une application

84



Cluster YARN: Exécution d'une application

85



YARN et Map-Reduce v2

86

- Yarn ne sait pas quel type d'application s'exécute: peut être MR ou autres applications non-MR comme:
 - ▣ Distributed Shell
 - ▣ Impala
 - ▣ Apache Giraph
 - ▣ Spark
 - ▣ Autres : <https://cwiki.apache.org/confluence/display/HADOOP2/PoweredByYarn>



Quelques applications pouvant s'exécuter de manière native sur Hadoop (Source :
HortonWorks)

YARN et Map-Reduce v2

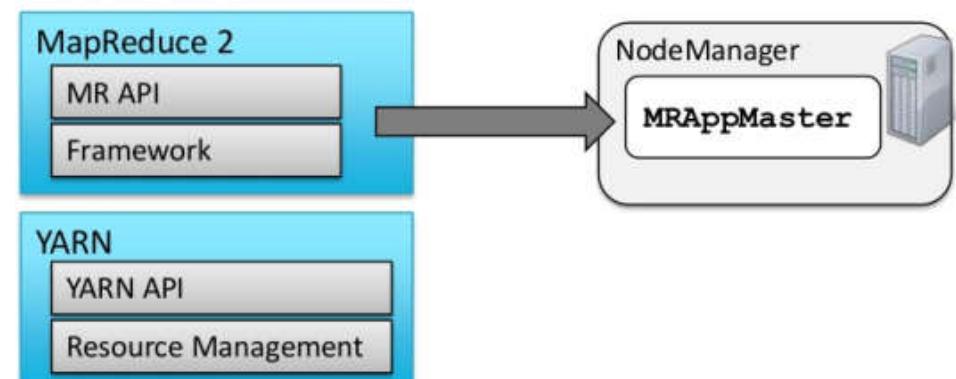
87

□ MR2: C'est quoi?

- Avec Yarn, il n'y a plus de l'unique JobTracker pour exécuter les jobs et des taskTracker pour exécuter les tasks des jobs.
- L'ancienne version MRv1 a été écrite pour s'adapter avec YARN et a été appelée MRv2.

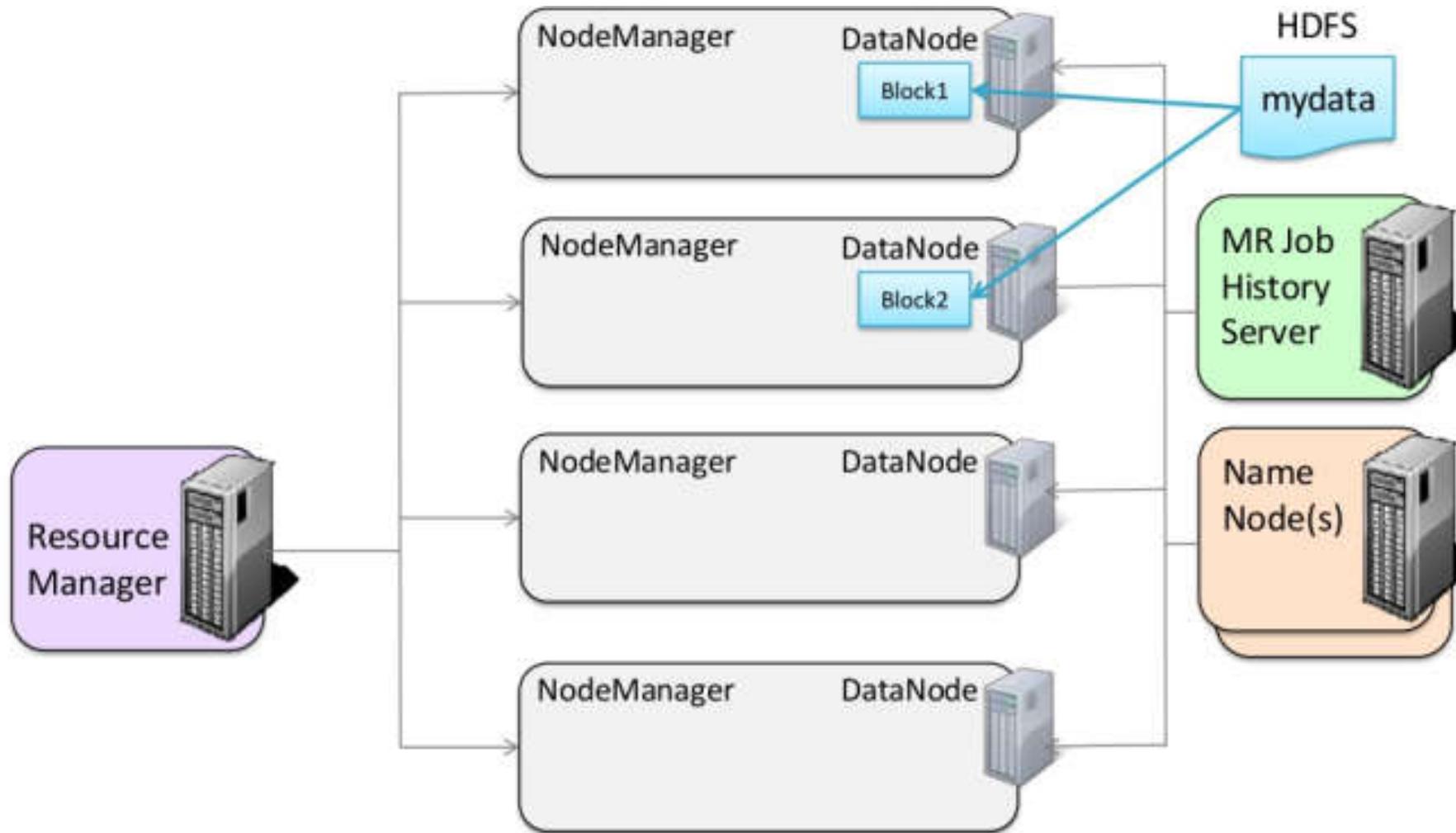
□ MR2 utilise Hadoop:

- Hadoop inclus un MRAppMaster (MapReduce ApplicationMaster) pour gérer les jobs MR.



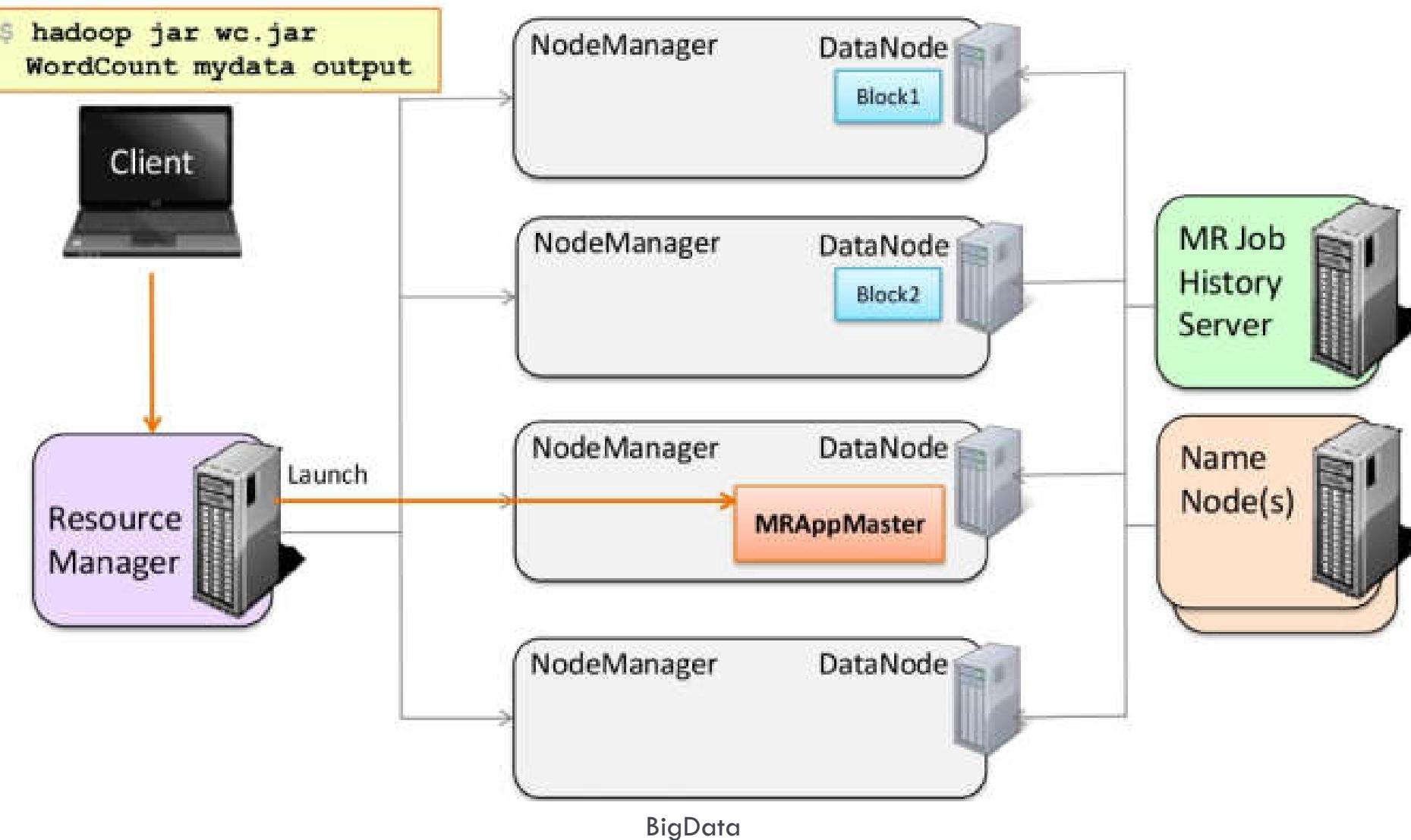
Cluster YARN: Exécution d'un job Map-Reduce

88



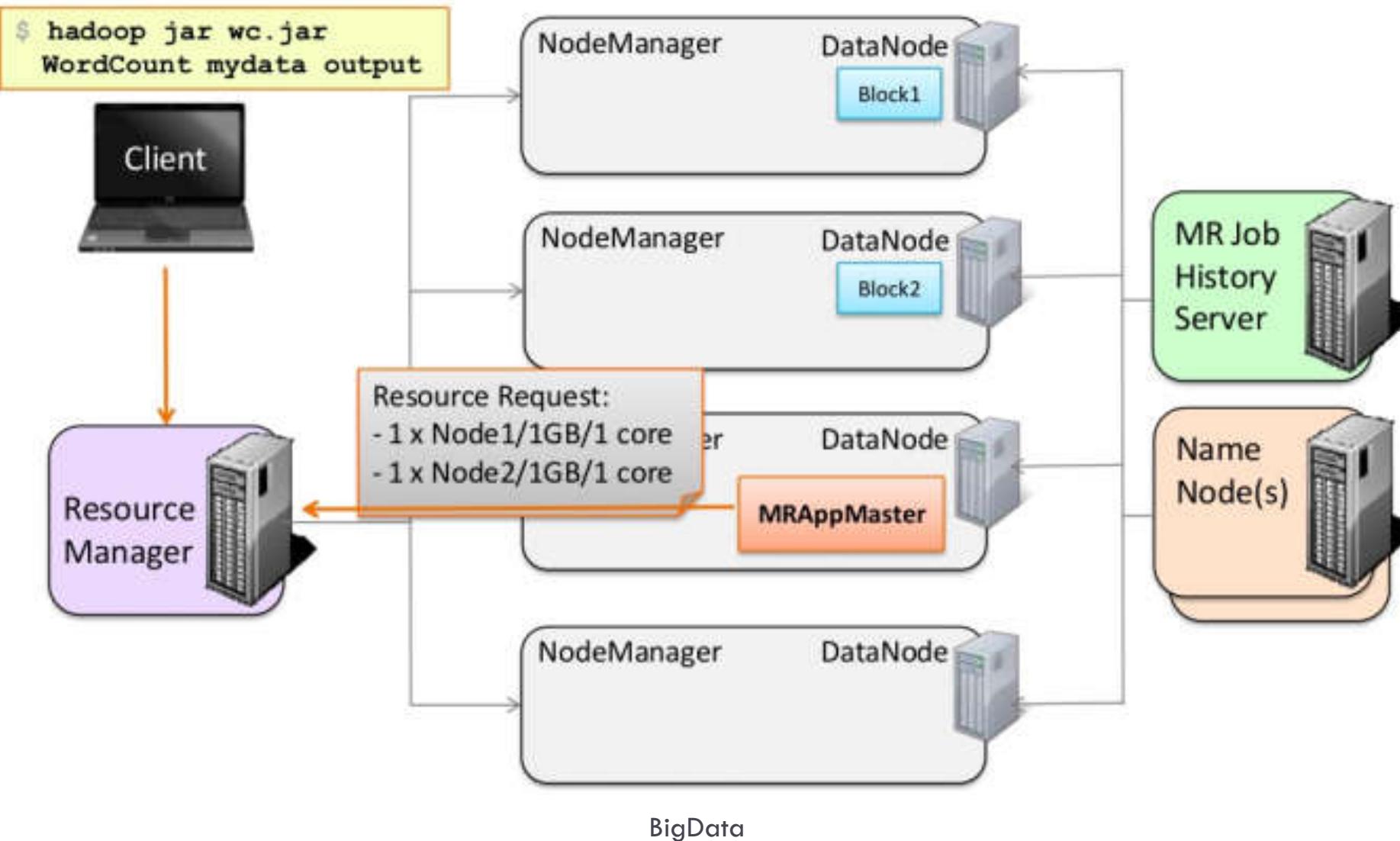
Cluster YARN: Exécution d'un job Map-Reduce

89



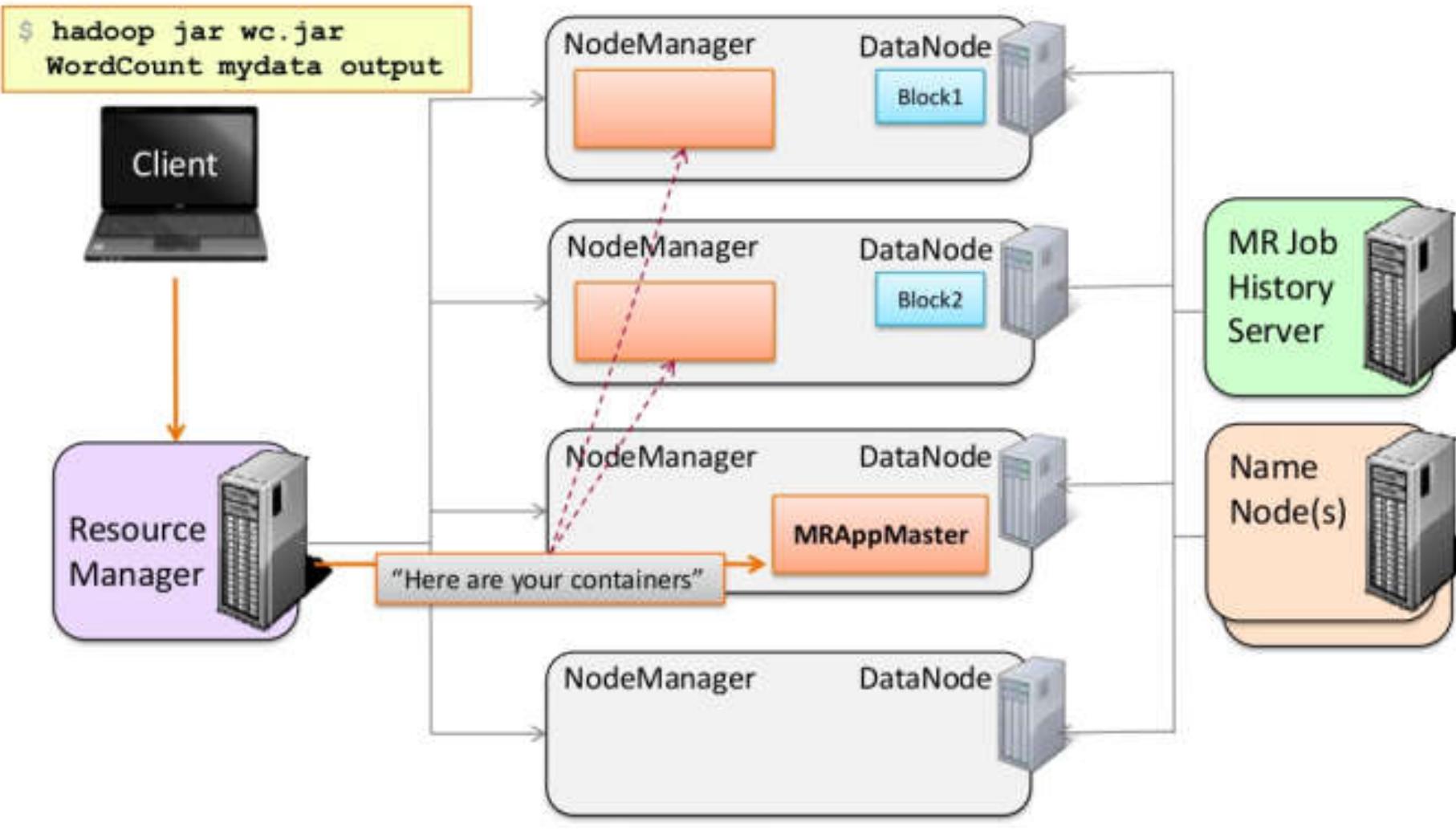
Cluster YARN: Exécution d'un job Map-Reduce

90



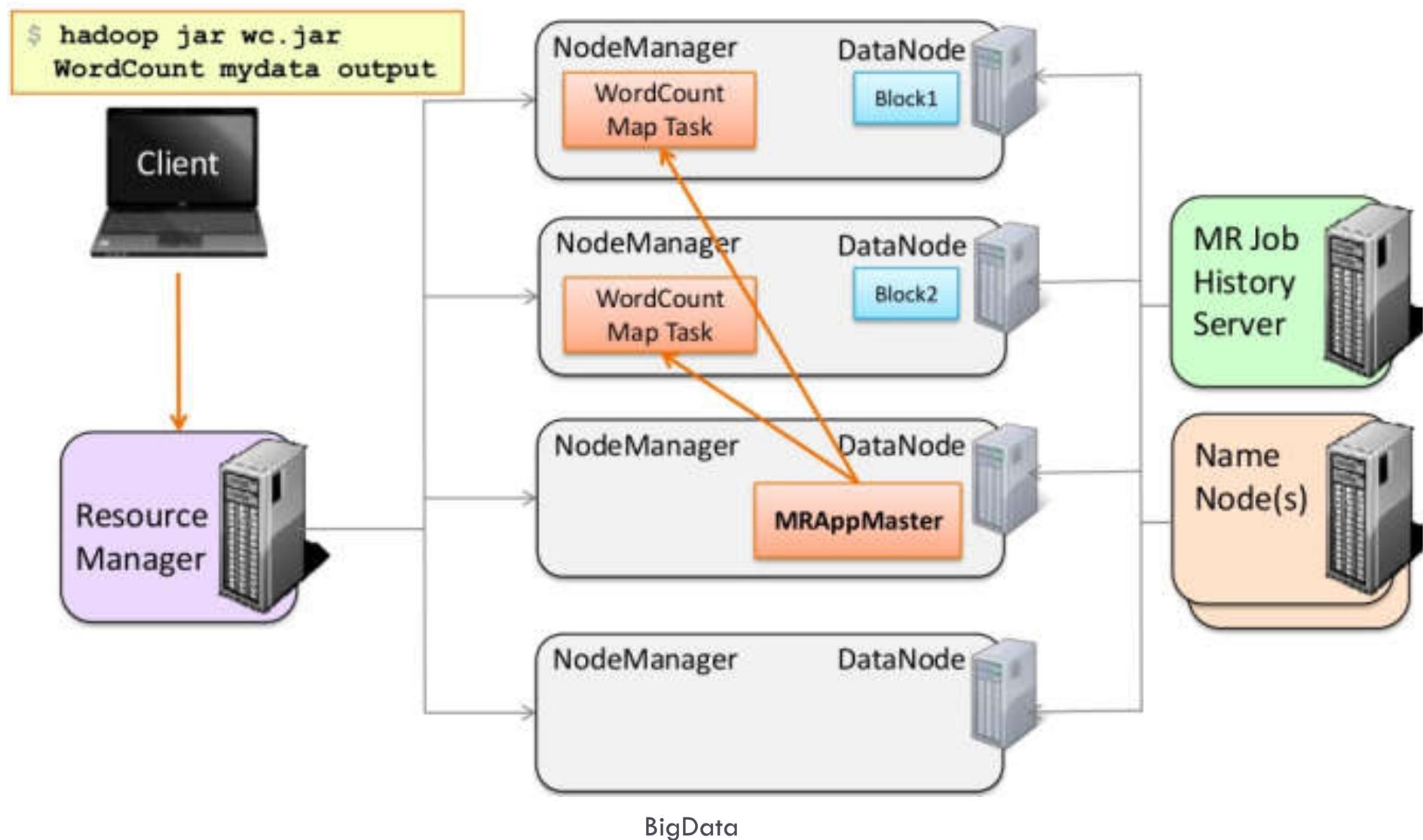
Cluster YARN: Exécution d'un job Map-Reduce

91



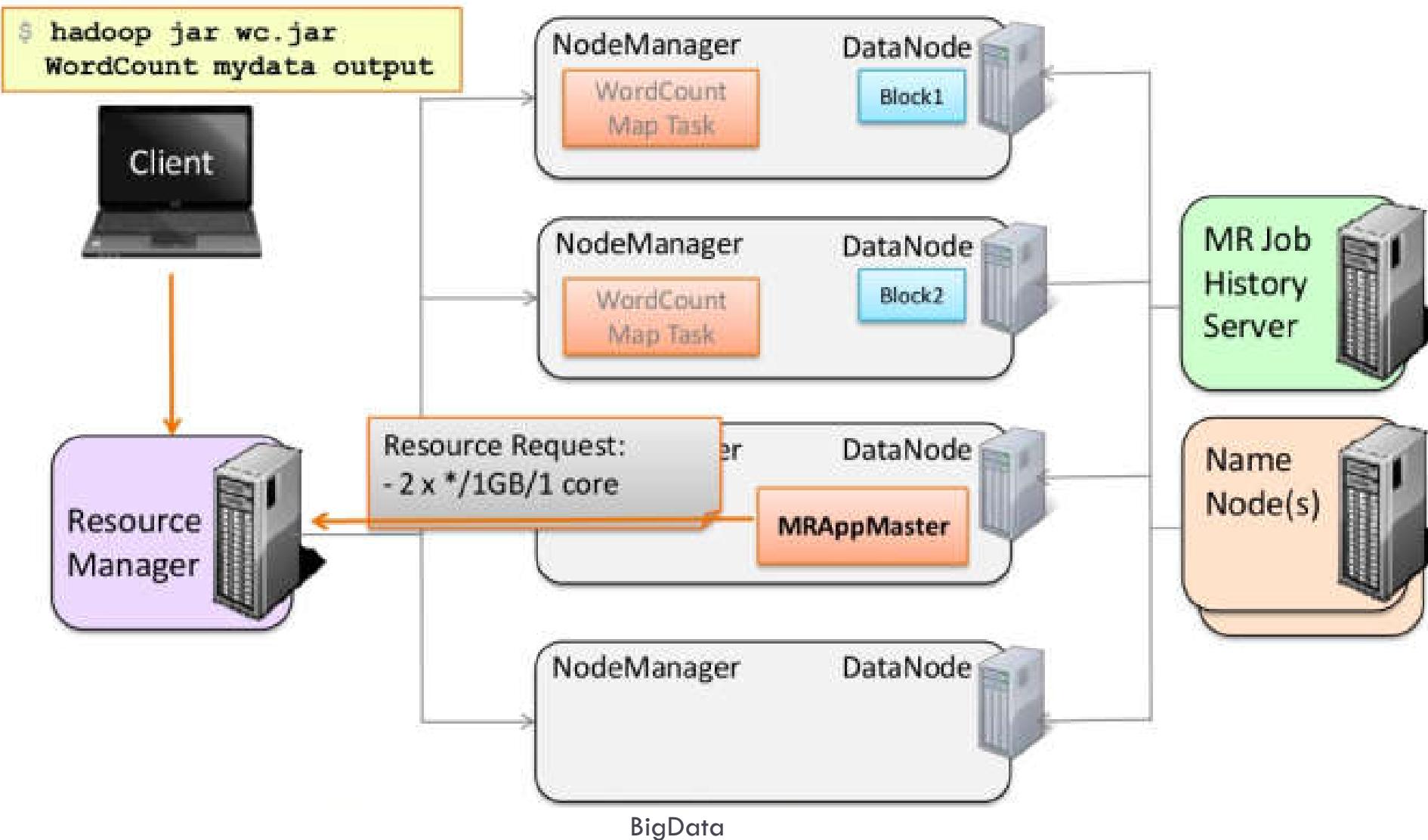
Cluster YARN: Exécution d'un job Map-Reduce

92



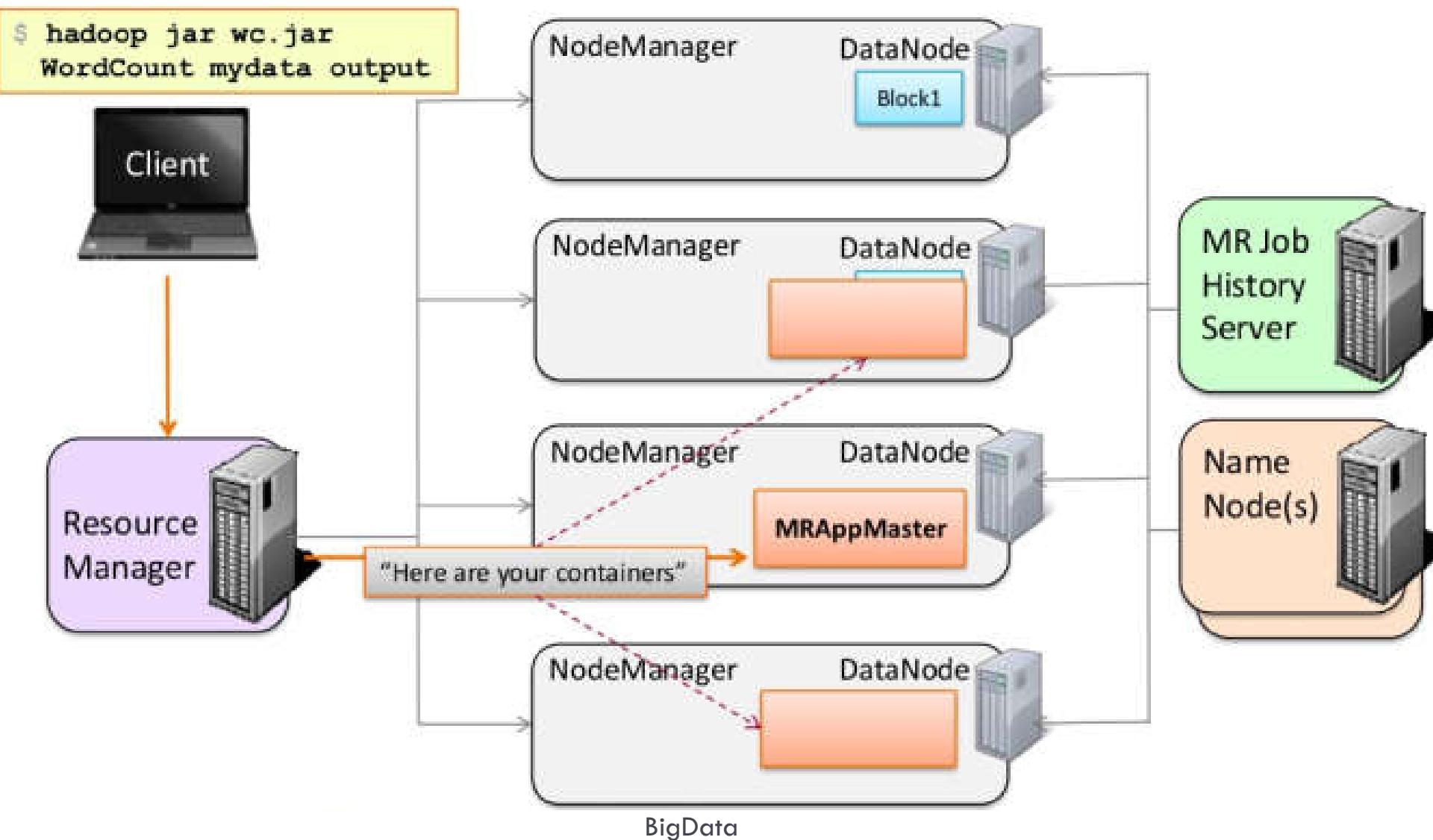
Cluster YARN: Exécution d'un job Map-Reduce

93



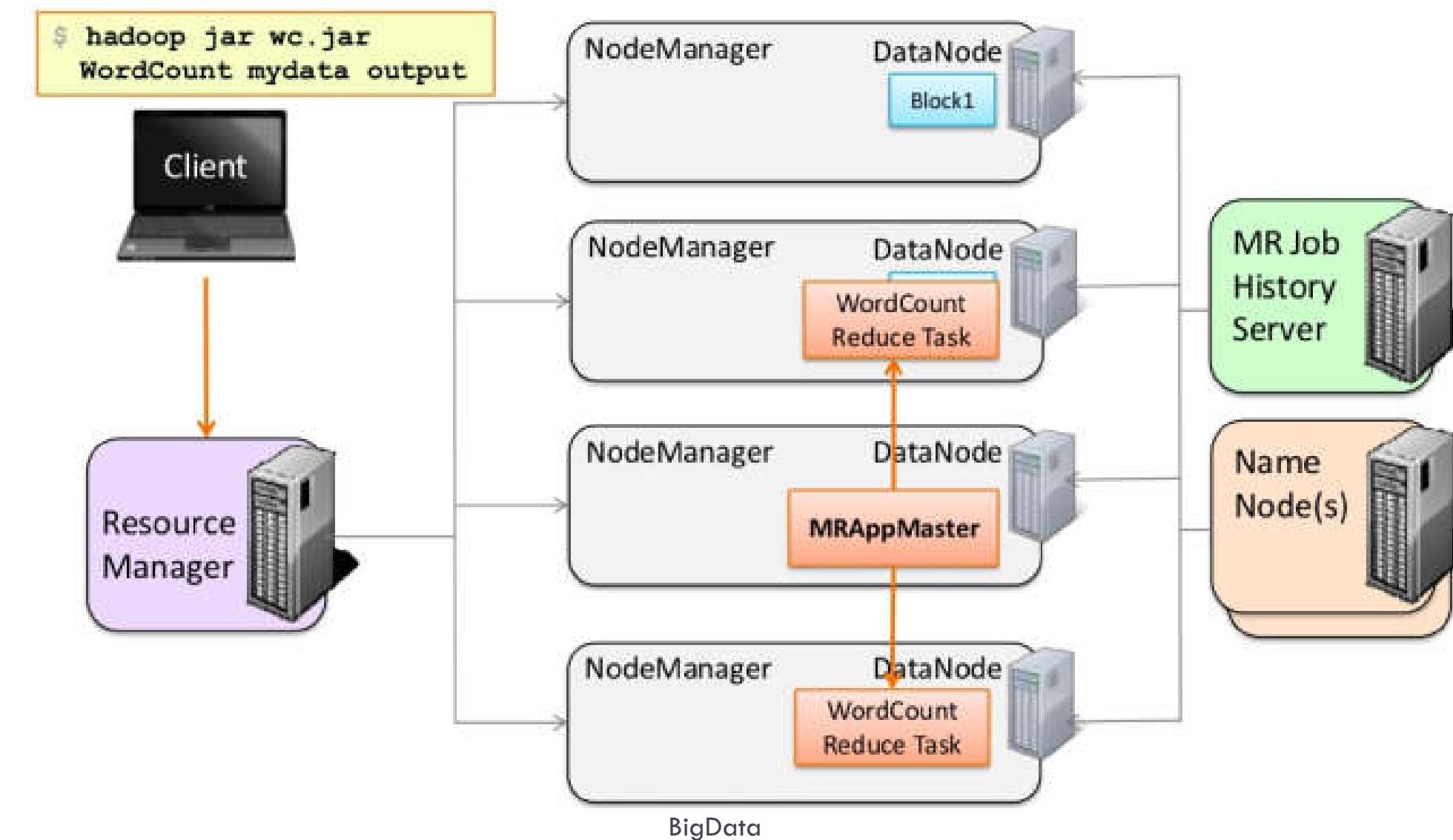
Cluster YARN: Exécution d'un job Map-Reduce

94



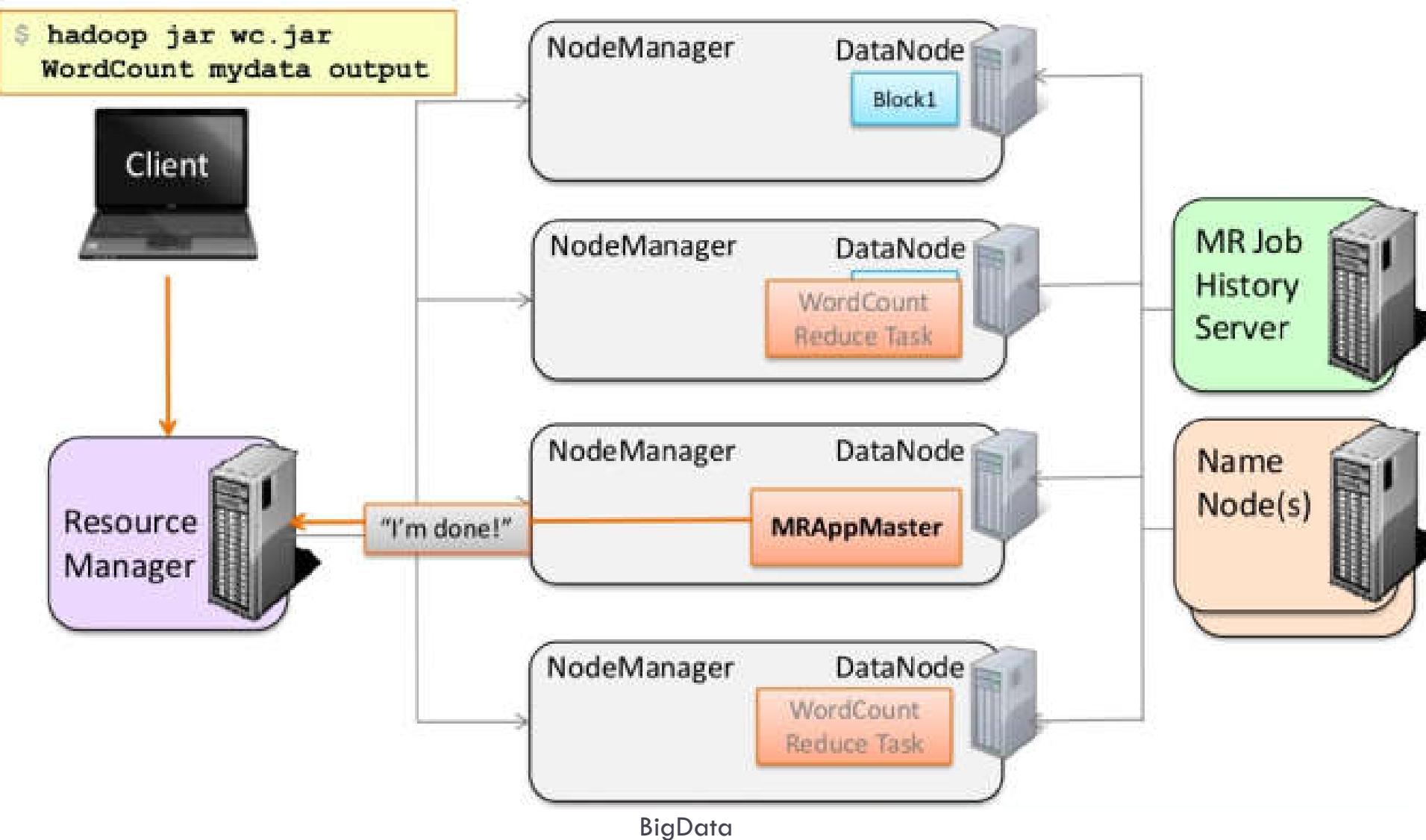
Cluster YARN: Exécution d'un job Map-Reduce

95



Cluster YARN: Exécution d'un job Map-Reduce

96



Cluster YARN: Exécution d'un job Map-Reduce

97

- dans Yarn, Shuffle s'exécute comme un service auxiliaire
- Il s'exécute dans le NodeManager JVM comme un service persistant.

