

Intoduction à l'architecture des systèmes distribués

Joseph Allemandou

Octobre 2016

Ce cours a pour objectif de présenter de façon pragmatique les notions nécessaires à la compréhension des architectures de systèmes distribués. Les exemples utilisés sont de complexité croissante :

- Parallélisation d'un algorithme
- Résilience et distribution de charge d'un site web
- Stockage et calcul distribué — HADOOP.

1 Jour 1 — Première approche

1.1 Calculer plus vite - Distribuer les traitements

- Exécution parallèle [1] [2] [3]
- Threads & Processus [4] [5]
- Synchronisation [6] [7]
- Approfondissement [8] [9] [10] [11] [12] [13] [14] [15]

1.2 La résilience — L'autre bénéfice

- Architecture distribuée [16] — L'exemple d'un site web
- Bénéfices apportés par la réplication [17] [18] [19]
- Modèles de réplication [20] [21]
- Approfondissement [22] [23] [24] [25] [26]

2 Jour 2 — HADOOP

2.1 Architecture d'un système de stockage

- Réplication des données (fiabilité, tolérance aux pannes, disponibilité)
- Représentation de la donnée (blocks)
- Briques logicielles nécessaires (namenode, datanode)
- Lecture / écriture sur un système distribué (description des interactions entre briques logicielles lors de la lecture ou l'écriture)

2.2 Architecture d'un système de calcul

- Localité des données (déplacement du code plutôt que des données)
- Problèmes associés à la parallélisation du calcul
- Paradigme map-reduce

- Briques logicielles nécessaires (resource manager, application master, worker)
- Problèmes sous-jacents à map-reduce (shuffle, RAM limits, sort)

3 Jour 3 — Big Data & Évaluation

4 Vous avez dit BIG DATA [27] ?

En vrac

- Blah

Références

- [1] Wikipédia, L'encyclopédie libre. Parallélisme (informatique). [https://fr.wikipedia.org/wiki/Parall%C3%A9lisme_\(informatique\)](https://fr.wikipedia.org/wiki/Parall%C3%A9lisme_(informatique)), 2004. [En ligne; accédé le 21 Septembre 2016].
- [2] Wikipédia, L'encyclopédie libre. Programmation concurrente. https://fr.wikipedia.org/wiki/Programmation_concurrente, 2005. [En ligne; accédé le 21 Septembre 2016].
- [3] Wikipédia, L'encyclopédie libre. Multitâche. <https://fr.wikipedia.org/wiki/Multit%C3%A2che>, 2003. [En ligne; accédé le 21 Septembre 2016].
- [4] Wikipédia, L'encyclopédie libre. Thread (informatique). [https://fr.wikipedia.org/wiki/Thread_\(informatique\)](https://fr.wikipedia.org/wiki/Thread_(informatique)), 2004. [En ligne; accédé le 21 Septembre 2016].
- [5] Wikipédia, L'encyclopédie libre. Processus (informatique). [https://fr.wikipedia.org/wiki/Processus_\(informatique\)](https://fr.wikipedia.org/wiki/Processus_(informatique)), 2004. [En ligne; accédé le 21 Septembre 2016].
- [6] Wikipédia, L'encyclopédie libre. Synchronisation (multitâches). [https://fr.wikipedia.org/wiki/Synchronisation_\(multit%C3%A2ches\)](https://fr.wikipedia.org/wiki/Synchronisation_(multit%C3%A2ches)), 2006. [En ligne; accédé le 21 Septembre 2016].
- [7] Wikipédia, L'encyclopédie libre. Interblocage. <https://fr.wikipedia.org/wiki/Interblocage>, 2005. [En ligne; accédé le 21 Septembre 2016].
- [8] Wikipédia, L'encyclopédie libre. Ordonnancement dans les systèmes d'exploitation. https://fr.wikipedia.org/wiki/Ordonnancement_dans_les_syst%C3%A8mes_d%27exploitation, 2002. [En ligne; accédé le 21 Septembre 2016].
- [9] Wikipédia, L'encyclopédie libre. Communication inter-processus. https://fr.wikipedia.org/wiki/Communication_inter-processus, 2004. [En ligne; accédé le 21 Septembre 2016].
- [10] Wikipédia, L'encyclopédie libre. Verrou (informatique). [https://fr.wikipedia.org/wiki/Verrou_\(informatique\)](https://fr.wikipedia.org/wiki/Verrou_(informatique)), 2003. [En ligne; accédé le 21 Septembre 2016].
- [11] Wikipédia, L'encyclopédie libre. Sémaphore (informatique). [https://fr.wikipedia.org/wiki/S%C3%A9maphore_\(informatique\)](https://fr.wikipedia.org/wiki/S%C3%A9maphore_(informatique)), 2005. [En ligne; accédé le 21 Septembre 2016].
- [12] Wikipédia, L'encyclopédie libre. Exclusion mutuelle. https://fr.wikipedia.org/wiki/Exclusion_mutuelle, 2005. [En ligne; accédé le 21 Septembre 2016].
- [13] Wikipédia, L'encyclopédie libre. Section critique. https://fr.wikipedia.org/wiki/Section_critique, 2005. [En ligne; accédé le 21 Septembre 2016].

- [14] Wikipédia, L'encyclopédie libre. Problème des producteurs et des consommateurs. https://fr.wikipedia.org/wiki/Probl%C3%A8me_des_producteurs_et_des_consommateurs, 2006. [En ligne; accédé le 21 Septembre 2016].
- [15] Wikipédia, L'encyclopédie libre. Problème des lecteurs et des rédacteurs. https://fr.wikipedia.org/wiki/Probl%C3%A8me_des_lecteurs_et_des_r%C3%A9dacteurs, 2005. [En ligne; accédé le 21 Septembre 2016].
- [16] Wikipédia, L'encyclopédie libre. Architecture distribuée. https://fr.wikipedia.org/wiki/Architecture_distribu%C3%A9e, 2005. [En ligne; accédé le 21 Septembre 2016].
- [17] Wikipédia, L'encyclopédie libre. Réplication (informatique). [https://fr.wikipedia.org/wiki/R%C3%A9plication_\(informatique\)](https://fr.wikipedia.org/wiki/R%C3%A9plication_(informatique)), 2008. [En ligne; accédé le 21 Septembre 2016].
- [18] Wikipédia, L'encyclopédie libre. Point individuel de défaillance. https://fr.wikipedia.org/wiki/Point_individuel_de_d%C3%A9faillance, 2006. [En ligne; accédé le 21 Septembre 2016].
- [19] Wikipédia, L'encyclopédie libre. Répartition de charge. https://fr.wikipedia.org/wiki/R%C3%A9partition_de_charge, 2004. [En ligne; accédé le 21 Septembre 2016].
- [20] Wikipédia, L'encyclopédie libre. Maître-esclave. <https://fr.wikipedia.org/wiki/Ma%C3%AEtre-esclave>, 2008. [En ligne; accédé le 21 Septembre 2016].
- [21] Wikipédia, L'encyclopédie libre. Réplication multi-maîtres. https://fr.wikipedia.org/wiki/R%C3%A9plication_multi-ma%C3%AEtres, 2013. [En ligne; accédé le 21 Septembre 2016].
- [22] Wikipédia, L'encyclopédie libre. Théorème cap. https://fr.wikipedia.org/wiki/Th%C3%A9or%C3%A8me_CAP, 2010. [En ligne; accédé le 21 Septembre 2016].
- [23] Wikipédia, L'encyclopédie libre. Consensus (informatique). [https://fr.wikipedia.org/wiki/Consensus_\(informatique\)](https://fr.wikipedia.org/wiki/Consensus_(informatique)), 2012. [En ligne; accédé le 21 Septembre 2016].
- [24] Wikipédia, L'encyclopédie libre. Paxos (informatique). [https://fr.wikipedia.org/wiki/Paxos_\(informatique\)](https://fr.wikipedia.org/wiki/Paxos_(informatique)), 2013. [En ligne; accédé le 21 Septembre 2016].
- [25] Wikipédia, L'encyclopédie libre. Split-brain. <https://fr.wikipedia.org/wiki/Split-brain>, 2011. [En ligne; accédé le 21 Septembre 2016].
- [26] Wikipédia, L'encyclopédie libre. Illusions de l'informatique distribuée. https://fr.wikipedia.org/wiki/Illusions_de_l%27informatique_distribu%C3%A9e, 2008. [En ligne; accédé le 21 Septembre 2016].
- [27] Wikipédia, L'encyclopédie libre. Big data. https://fr.wikipedia.org/wiki/Big_data, 2010. [En ligne; accédé le 21 Septembre 2016].