



## Rapport de Similitude Turnitin

Cours Frameworks de stockage et de traitement distribués par Khaled Jouini

De Quick Submit (Quick Submit)

Traité le 14-déc.-2024 12:55 PM CET

Numéro : 2552189813

Nombre de mots : 6705

Indice de Similitude	Similarité par source
2%	Internet Sources: 2% Publications : 1% Copies de l'étudiant : 1%

## Sources :

- 1% match (Internet depuis le 02-déc.-2020)  
<http://www.wikiwai.com/2019/11/25/modelisation-big-data/>
- 1% match (Internet depuis le 05-déc.-2020)  
[https://inba.info/redhat-9-0-guide-de-personnalisation-fr\\_585db4b4b6d87f9da58b6c46.html](https://inba.info/redhat-9-0-guide-de-personnalisation-fr_585db4b4b6d87f9da58b6c46.html)
- 1% match (Travaux des étudiants depuis le 09-mai-2023)  
[Submitted to Unizin, LLC on 2023-05-09](#)

## Texte de la copie :

tiun0Te'hc0RgGee4es.h:ioDIIseyfei,mmtfIhiSanpeBiwltNiif6viae:teht9dIS7GSDe.8yu-ami1dt-pa4eo9P-s1riSu-o9tmco0er1aos6sgn3ine-O2ga.ponJeundrlayLAat2innr0gag1ely5Ss.CyissltuaestmtelnsrstDe..renseigntScaanlde.Implementation h a d @ g m Disponible sur ce Lien. Documentation officielle d'Apache Hadoop Apache Hadoop, quésaco? Hadoop HDFS (Hadoop Distributed File System) Hadoop : Framework open source pour le stockage et le traitement (MapReduce) distribués. ÉPcrinritceipnaJuaxvsau/pCpo1r,tisn/scpoinétrdibeutGeourosq:leYaFhSoeot,sFoancMeBaopoRke,dIBuiMcle,.Hortonworks, etc. a c o m Composantes clés HDFS pour le stockage distribué Écojs.yMHZPsoitBkgaèa/pmKHSReEiveee:edpBue:DcLrae:DnOpisgtorarciublhgrueeeléessterdtasreateihtuearm,uCtenonitnvfedigaisuttrial g m h a d @ YARN (Hadoop 2.0) : Traitements autres que MapReduce Sqoop (chargement de données relationnelles), Oozie, Storm (flux de données), etc. 1Mais utilisable avec d'autres langages Chapitre 2 - Systèmes de Fichiers Distribués HDFS Architecture Master/Slaves Journalisation et failover o m Haute disponibilité Fédération Formats de fichiers Commandes Hadoop m a il. c Google File System (GFS) g Qujiz.k le h a d @

## 2Système de fichiers Fonctionnalités d'un système de fichiers

, entre autres :

Mmcfcraaéhennieirpp/r.duulélaatttriuiooinnedddeesssffificicchhiieersrrssp,iandrseélesresorpp,rseourgaprtaiominmmseserosendttm'd: dans un a c o m Allocation de la place sur mémoires secondaires : les fichiers étant de taille différente et cette taille pouvant être dynamique, le SGF alloue à chaque fichier g m uLddonoecnjnsncano.rléiimpskeatbsitvri;oeepnsvo(adnureoriasmcbe,ilcelaahld,decereerhssgas:requai.lune.eus.ltef)incsréehdcigeeeromspuse a d @ h etc. Principe du Système de Fichiers Distribué Hadoop (HDFS) Objectif de HDFS : HLeDsFFSicehsitecsosnoçnut dpiovisrésstoecnkberloectsg(éérgearldeemsednotnanpéeeslémsasspsiliivltse,soudechmuannksiè).reLedsisbtrliobcusee. c o m d'un même fichier sont typiquement stockés sur des serveurs différents. a Chaque bloc est répliqué (3 fois par défaut) pour assurer la tolérance aux pannes. HICLeeDsojCFnmSç.oounpdkepiferiomepurtertdtdiedreescloctecercaméltleuiterérn,etsd.suespéspqdriuomenennrtéeieeltsleceso,spteitmr h a d @ g m sur ou près du stockage physique afin de réduire la transmission des données. Namenode et DataNodes HDFS : repose sur 2 types de noeuds, les NameNodes et les DataNodes DcNhaautmanNeksNo)doede(s(esrevvuerudrededmonéntéaedso)n:nsétoecs)ke et restitue les ibllo.cs de données (les c o m Stocke le répertoire, l'espace des noms, l'arborescence et les métadonnées des fichiers. j.CLpLcoeahhkersaasnsDqctdaruaaenentltailesahbNe)éllooeilbacdce,telroulgs'craeeelad.ndlrevieseoasssiteebionldnotcupdséDedraio'utb d @ g m a h a Si le NameNode constate qu'un bloc n'est pas suffisamment répliqué, il initie une réplication sur d'autres DataNodes Hadoop HDFS (Hadoop Distributed File System) m a il. c o j,k le h a d @ g m Figure: Figure reprise de "IBM Skills Academy: Big Data Engineer - Student Guide - 2019" Chapitre 2 - Systèmes de Fichiers Distribués HDFS Architecture Master/Slaves Journalisation et failover o m Haute disponibilité Fédération Formats de fichiers Commandes Hadoop m a il. c Google File System (GFS) g Qujiz.k le h a d @ File System Image (fsImage) Lm(eeé.Ngta.adrmeonenoNméodemseacgtaeepndtru'eurgenisptfircéehrilioeedsr,iqmmuoiegdmriafectionatntidoden'susnisntcaehprmusnt edit Log a c o m fsImage Le fsImage est un fichier qui contient une capture instantanée (snapshot) de toutes les j.mlrClérkhépeatpearrgdtéaoésiernennsnté,enmelos'éémdmtuaslotdsierydeesutfaèiscumyhsdeiteéérdmsm,eepaferidcrraehmgiefeiscrrsihp des @ g m h a d Fichier de journalisation Edits Log editsLLOegeditsLog est un fichier journal dans lequel toutes les mioldi.fications apportées au a c o m système de fichiers (ajout, suppression, renommage de fichiers, etc.) sont enregistrées de manière séquentielle. j.CqeLceudakhipaetpstdLueoiutrrsegemLs.omedgtoudedfessiftlrmrceaalascstgeoeioennn.sttieerulstipreeonlu'réretagasitsutarrécetrualdaa @ g m h a d Failover et Démarrage du NameNode Processus de démarrage du NameNode LfEiocnrhssuieditreus,ddleeépMNUaaismrrlaeegNfeico,hdleieerNlfitasllmemeeadNgioetsd.Leogchpaorugreaepnplmiqéuemroitoireutl'ilms.lea de c o m survvenues après la capture du dernier fsImage. Le NameNode combine ces deux fichiers pour