

Hitachi Data Discovery Suite

File & Content Solutions



L'information créée, collectée et multipliée en entreprise atteint un développement exponentiel, sa nature et son contenu est de tous les types, de tous les formats et, parfois, multilingues. Une solution de découverte, d'analyse des contenus, de classification et de recherche est alors un élément primordial d'une gouvernance maîtrisée des données.

La solution Hitachi Data Discovery Suite (HDDS) est la réponse Hitachi, pour aider les entreprises dans ses défis d'accès pertinent à une information précise. Tous les niveaux de l'entreprise sont concernés : l'administrateur système qui a besoin de connaître le niveau de Tiering et de protection des données ; le responsable informatique souhaitant un rapport sur la consommation du stockage ; le responsable sécurité traçant les éventuelles dispersions sur des projets clefs ; le juriste qui a besoin d'isoler rapidement des documents de litige ; le collaborateur qui doit retrouver urgemment les chiffres archivés d'une production, dont le responsable n'est plus en poste ; etc.

Autant de situations où les nécessités de réactivité et de pro activité sont essentielles à la meilleure prise de décision, la transmission d'une connaissance, l'investigation, la maîtrise des risques, le ré-enrichissement d'information, la veille technologique et le respect de réglementations de certains pays, via un processus d'audit précis (e-Discovery).

Hitachi Data Discovery Suite remplit toutes ces tâches, via un moteur d'indexation plein-texte sur des contenus non structurés multilingues et multi-formats. HDDS est aussi et surtout une interface Web utilisateur. C'est-à-dire un moyen simple et clairement efficace, pour adresser tous les niveaux de décisions d'une entreprise.

Conception Innovante



Hitachi Data Discovery Suite est plus qu'une solution logicielle. Hitachi a intégré HDDS dans un package regroupant serveur et stockage, afin de rendre encore plus simple le déploiement et l'évolutivité. Cet ensemble est le résultat d'un constat fort orienté Big Data : indexer des données requiert une puissance de calcul, pour bâtir les dictionnaires et autres thésaurus documentaires ; et une capacité de stockage, pour organiser et conserver les résultats de traitement.

Sans un contrôle complet de la chaîne de recherche, la force de l'indexation est difficilement valorisée. Hitachi a donc bâti une solution matérielle et logicielle évolutive à travers des Nodes autonomes. Ainsi, en fonction du niveau de performance à délivrer et de la quantité d'information à ingérer, l'entreprise acquiert des serveurs HDDS dialoguant entre eux (Search Bus) sur le réseau IP, afin de fédérer leurs traitements respectifs d'indexation, se répartir la charge d'exploration et délivrer l'information appropriée.

Le *Federated Search* consiste donc à mettre en œuvre une architecture distribuée capable d'interagir avec les différentes sources d'acquisition et de les découpler, en consolidant les résultats dans une interface Web, compatible avec les politiques d'accès du système d'information (RADIUS, Active Directory & LDAP).



Les services au sein de la solution HDDS ne se contentent pas uniquement de restituer une information. Dans le cas d'une exploration d'équipements Hitachi, tels qu'Hitachi NAS (HNAS) et Hitachi Content Platform (HCP), HDDS est capable de réaliser des opérations de déplacement. Il s'agit d'effectuer une migration de fichiers, basée sur les contenus et la nature des données, entre le serveur de fichiers HNAS et la solution Cloud-Archivage HCP.

Hitachi Data Discovery Suite, principales innovations

Search Bus & Big Data - Indexation évolutive



Le **Search Bus** est un **réseau de serveurs** HDDS. Chaque serveur est autonome et gère les analyses de son instance et **dialogue** en **IP** avec les autres serveurs HDDS. Le nombre de serveurs dépend de la **profondeur d'indexation**, de la **quantité d'information** et de la répartition des sources dans l'entreprise.

Des **filtres** d'exclusion ou d'inclusion ciblent l'automatisation de l'**indexation**, à la demande ou programmée, à certains types ou taille de fichiers ou des répertoires précis. Les sources de fichiers sont explorées en **CIFS, NFS** ou **HTTPs** (HCP).

Cette architecture **Search Bus** permet à l'entreprise de disposer d'une **puissance évolutive** réellement capable d'adresser des **volumétries** très importantes et des **fichiers** en très grand nombre, telle que celle du **HNAS** (16Po et 4 milliards de fichiers) ou du **HCP** (40Po et 32 milliards de fichiers), et cela dans un **minimum d'effort**, juste en ajoutant des serveurs **tout-en-un** prêt à répondre à de nouvelles sollicitations.

Recherche



Avec un **navigateur Web** sur tous les types de plate-forme cliente, l'administrateur contrôle graphiquement l'activité d'indexation de chaque source, et chaque utilisateur accède en fonction de ses **droits ACLs** (Access Control List) à son champ de **recherche** et de **consultation** des informations.

Les recherches possibles sont de complexité différente. Elles peuvent être basiques avec justes quelques **mot clefs**, évoluées via une requête **multicritères** et exportables pour faciliter l'**audit** par des tiers (E-discovery).

Des **requêtes** préconfigurées sont aussi accessibles en lien *Rich Site Summary* (RSS). Depuis son **interface graphique**, HDDS propose en complément la possibilité de télécharger un **Gadget Windows** qui donne en temps réel les résultats sous forme d'abonnement **RSS**.

Intelligent File Tiering



Dans le cadre d'une **architecture** de **fichiers** et d'**archivage** avec des composants Hitachi Data Systems, la solution HDDS complète son fonctionnel de base, par une capacité de migrer les fichiers entre **HNAS** et **HCP**, de manière transparente, avec des **règles** basées sur l'environnement **IT** (dates, taille et type) et les **contenus** indexés. L'ordre de migration se réalise à partir du résultat d'une recherche, dont la requête peut être enregistrée.

Sans compétence spécifique et via l'interface graphique, l'**administrateur IT** gère donc la programmation et la **classification** de son **stockage** Hitachi au regard de la nature, la **valeur** et l'obsolescence des fichiers.

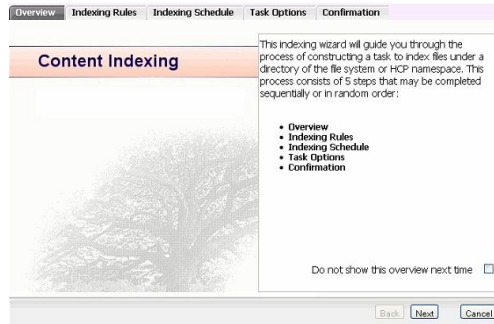
Risk Management



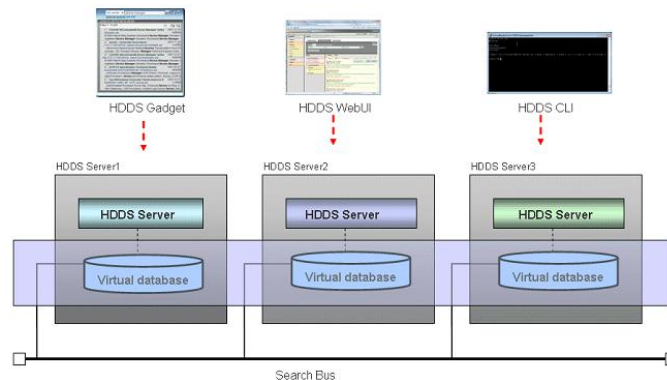
Certaines **informations** nécessitent une attention particulière du fait de leur **valeur** probatoire, technique ou **stratégique**. HDDS propose un **calcul du risque** en fonction du **temps d'accès** et des **lieux de stockage**, incluant une éventuelle copie. Via une **échelle** de 0 à 100, un affichage graphique propose une vue de ce **risque** sur le résultat d'une requête de recherche.

Des copies multiples ne signifient pas nécessairement un **risque** réduit. La **nature** de l'information peut justement impliquer un risque de « **fuite** » à cause d'une **dispersion** du **contenu**, jugée trop importante. Par contre, une copie particulière sur un **WORM**, comme le HCP, peut pondérer un risque, quand il s'agit d'élément de **preuve**. HDDS est donc à même de prendre en compte des **valeurs** de **seuil** (*Threshold : Boundary, Danger, Caution*) configurés et **adaptés** à une recherche.

Hitachi Data Discovery Suite



Accès	Matériel	Service	Extension
Données HTTPs, Object ACL, Access Control List, HCP Tenant/Namespace.	Réseau 4x 1GbE par Serveur. IPv4, IPv6	Recherche Content (Full-text), XML Structural, API, HTTPs, Mode CLI, Windows Gadget, RSS.	Conformité E-discovery, Hash (MD5, SHA, RIPEMD), Export ASCII, CSV & XML.
Administration HTTPs, SNMP (MIB), Mode CLI, RADIUS, DNS, NTP, Email Alert.	Hitachi CR220 2U Serveur autonome avec disques internes, Gateway (Node) connecté à une baie de stockage HDS.	Indexation 77 langues, 400 formats de fichiers, CIFS, NFS.	Solution Hitachi Data Ingestor, Hitachi Content Platform, Hitachi NAS, Hitachi Device Manager.
Contrôle Login/Password, Active Directory, LDAP, SSL, Certificate Authority, Filtre IP (Allow/Deny).		Action Avancée Privileged Delete-Purge, Release Hold. File Migration, Risk Management, Index Schedule.	Support HNAS, HCP, HENP, MS Windows Server, MS SharePoint.



Hitachi Data Systems Corporation

Hitachi est une marque déposée d'Hitachi, Ltd. aux États-Unis et dans d'autres pays. Hitachi Data Systems est une marque déposée et une marque de service d'Hitachi, Ltd. aux États-Unis et dans d'autres pays. Toutes les autres marques de commerce, marques de service et noms de sociétés mentionnés dans le présent document ou site Web appartiennent à leur propriétaire respectif.

Avis : Le présent document est à caractère informatif uniquement et ne stipule aucune garantie, expresse ou implicite, sur aucun équipement ou service offert ou devant être offert par Hitachi Data Systems. Le présent document décrit des fonctionnalités dont les conditions sont définies sur un contrat de maintenance effectif avec Hitachi Data Systems et pouvant dépendre de la configuration, ainsi que des fonctionnalités qui ne sont peut-être pas disponibles actuellement.

Contactez votre revendeur Hitachi Data Systems local pour plus d'informations sur la disponibilité des fonctionnalités et des produits.
 © Hitachi Data Systems Corporation 2012. Tous droits réservés. 2012.

