

Méthodes d'accès aux données

Domaine	Ingénierie et Architecture
Filière	Informatique
Orientation	Logiciel (IL)
Mode de formation	Plein temps

Informations générales

Nom:	:	Méthodes d'accès aux données
Identifiant:	:	MAC
Années académiques	:	2018-2019, 2019-2020, 2020-2021
Responsable:	:	Nastaran Fatemi
Charge de travail:	:	120 heures d'études
Périodes encadrées:	:	64 (= 48 heures)

Semestre	E1	S1	S2	E2	S3	S4	E3	S5	S6
Cours								32	
Laboratoire								32	

Connaissances préalables recommandées

L'étudiant doit connaître et savoir utiliser les notions suivantes:

- concepts de base du modèle relationnel (modélisation, langage SQL, formes normales);
- algorithmes et structures de données (graphes, tables, B-Arbres);
- langage Java (éléments de base du langage).

Les unités d'enseignement BDR (bases de données relationnelles), ASD1 & 2 (algorithmes et structures de données) et POO1 (programmation orientée objet, Java) permettent d'acquérir ces connaissances.

Objectifs

A l'issue de cette unité d'enseignement, l'étudiant-e sera capable de :

- expliquer les principaux protocoles et algorithmes mis en œuvre par les SGBD pour la gestion de la concurrence (gestion des transactions).
- décrire la problématique liée à l'adaptation du modèle Objet au modèle Relationnel (« Impedance mismatch »).
- maîtriser les différentes catégories de modèles de données NoSQL.
- utiliser les langages de requêtes NoSQL.
- décrire les principaux modèles de recherche d'information (RI).
- utiliser le modèle vectoriel pour l'indexation et la recherche d'information textuelle.
- utiliser les modèles récents d'analyse de données "big data" tel que le map-reduce et le RDD de Spark.

A l'issue des travaux pratiques en laboratoire, l'étudiant-e sera en outre capable de :

- réaliser une application interagissant avec une base de données NoSQL.

Méthodes d'accès aux données

- réaliser une application d'indexation et de recherche des données textuelles via la librairie Apache Lucene.
- réaliser une application d'analyse de données big data avec map-reduce ou Apache Spark.

Contenu et formes d'enseignement

Répartition des périodes indiquée à titre informatif.

Cours: 32 périodes

- Introduction	2
- Gestion des transactions : Notion de transaction, mécanismes de reprise, mécanismes et algorithmes de gestion de la concurrence, traitement des verrous mortels.	2
- Principes de systèmes distribués et le théorème CAP	2
- Modèle de données et langages de requêtes NoSQL	8
- Indexation et recherche d'information textuelle avec le modèle vectoriel	6
- Analyse des données big data avec map-reduce	2
- Analyse des données big data avec Spark	8
- Travaux écrits	2

Laboratoire: 32 périodes

- Phase 1: Etude spécifique d'une nouvelle technologie de base de données NoSQL et sa présentation	8
- Phase 2: Réalisation d'une application de base de données avec la technologie étudiée et présentée en phase I	12
- Indexation et recherche d'information vectorielle avec la librairie Apache Lucene	6
- Analyse de données big data avec map-reduce ou Spark	6

Bibliographie

[1] "Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems", Martin Kleppmann, Mars 2017.

[2] "Next Generation Databases NoSQL and Big Data", Guy Harrison, 2015.

[3] "Database Systems, The Complete Book", 2nd Edition, Hector Garcia-Molina, Jeffrey D. Ullman, and Jennifer Widom, 2008.

Contrôle de connaissances

Cours:

L'acquisition des matières de cet enseignement sera contrôlée au fur et à mesure par des tests et des travaux personnels tout au long de son déroulement. Il y aura au moins 2 test d'une durée totale d'au moins 1 période.

Laboratoire:

ils seront évalués sur la base des rapports de manipulation, à 3 reprises au minimum.

Examen:

L'atteinte de l'ensemble des objectifs de formation sera vérifiée lors d'un contrôle final commun écrit d'une durée de 90 minutes.

Matériel autorisé:

- Information communiquée directement par l'enseignant.

Calcul de la note finale

Note finale = moyenne cours x 0.3 + moyenne laboratoire x 0.2 + moyenne examen x 0.5

Fiche validée le 13.09.2018 par Donini Pier