Description du module, disponible en: FR

Advanced Databases

Informations générales

Nombre de crédits ECTS

3

Code du module

TSM_AdvDaBa

Valable pour l'année académique

2026-27

Dernière modification

2025-10-07

Coordinateur/coordinatrice du module

Marcelo Pasin (HES-SO, marcelo.pasin@he-arc.ch)

Explications concernant les langues d'enseignement par site :

- L'enseignement est dispensé dans la langue indiquée ci-dessous pour chaque site et chaque exécution du module.
- Les supports de cours sont disponibles dans les langues indiquées ci-dessous pour chaque site et chaque exécution du module. Lorsque plusieurs langues sont utilisées, la proportion de contenu disponible dans chaque langue est précisée (100 % = ensemble des supports de cours).
- Les examens (questions et réponses) sont entièrement rédigés dans la langue indiquée ci-dessous pour le site et l'exécution du module concernés. Ils se déroulent en présentiel.

	Lausanne			Lugano	Zurich		
Leçons		**X** F 100%					
Documentation		**X** F 10%	**X** E 90%				
Examen		**X** F 100%	**X** E 100%				

Catégorie de module

TSM approfondissement technico-scientifique

Recommandations de profil

Ce module est recommandé pour les profils suivants :

	Lausanne
Profil	CS

Leçons

2 leçons et 1 leçon de pratique par semaine

Compétences préalables

Connaissances préalables, compétences initiales

- Modèle relationnel, algèbre relationnelle et normalisation
- SQL : modélisation des données, langage de requête, transactions et droits d'accès
- Optimisation des requêtes, index de base de données
- Langages de programmation orientée objet

Brève description du contenu et des objectifs

Les applications modernes reposent sur des systèmes de données capables de traiter des volumes importants d'informations tout en garantissant la fiabilité, la cohérence et la performance. Ce cours propose une exploration approfondie des fondations des systèmes de données et des architectures distribuées.

L'objectif est de comprendre les principes fondamentaux qui sous-tendent les systèmes de données intensifs : fiabilité, passage à l'échelle, tolérance aux pannes, cohérence et consensus. Les étudiant-e-s apprendront à analyser les compromis entre cohérence, disponibilité et latence, et à concevoir des architectures adaptées aux besoins réels des applications distribuées.

Objectifs, contenus, méthodes

Objectifs d'apprentissage, compétences à acquérir

À la fin du module, les étudiant-e-s seront capables de :

- Expliquer les principes de **fiabilité**, **scalabilité** et **maintenabilité** des systèmes de données ;
- Comparer différents **modèles de données et langages de requête** (relationnel, document, graphe, clé-valeur) et comprendre leurs usages ;
- Décrire les **mécanismes internes de stockage et d'indexation** (B-trees, LSM-trees, colonnes, sérialisation, compatibilité ascendante) ;
- Comprendre les **modèles de réplication** (leader-follower, multi-leader, leaderless) et leurs effets sur la cohérence ;
- Mettre en œuvre des **stratégies de partitionnement (sharding)** et de répartition de charge ;
- Expliquer le fonctionnement et les limites des **transactions distribuées**, des protocoles de **verrouillage** et de **commit en deux phases** ;
- Identifier les **défaillances partielles** dans les systèmes distribués et leurs conséquences ;
- Analyser et comparer les **modèles de cohérence** (forte, causale, éventuelle) et les **protocoles de consensus** (Paxos, Raft) ;
- Évaluer les compromis entre **cohérence, disponibilité et tolérance aux pannes** (CAP theorem) ;
- Utiliser ces connaissances pour **concevoir, choisir et configurer** des systèmes de données adaptés à divers contextes.

Contenu des modules avec pondération du contenu des cours

Contenu du cours et pondération

Fondations des systèmes de données (40 %)

- Fiabilité, tolérance aux pannes, passage à l'échelle
- Modèles de données et langages de requête
- Stockage, indexation et sérialisation des données
- Compatibilité et évolution des schémas

Systèmes de données distribués (60 %)

- Réplication et partitionnement
- Transactions distribuées et isolation
- Défaillances partielles, détection et récupération
- Cohérence, consensus et compromis de conception

Méthodes d'enseignement et d'apprentissage

Ce cours se base sur des présentations théoriques et des exercices pratiques, des laboratoires ou des projets de groupe.

Bibliographie

Diapositives des cours, références à des ressources Internet et à des livres.

Martin Kleppmann, *Designing Data-Intensive Applications*, O'Reilly.

Evaluation

Évaluation supplémentaire pendant le semestre

Le module ne comprend pas d'évaluation supplémentaire pendant le semestre

Principe pour les examens

En règle générale, tous les examens réguliers de fin de module se déroulent sous forme écrite. Concernant les examens de répétition, leur format (écrit ou oral) sera communiqué par l'enseignant-e en même temps que le calendrier des examens.

Examen de fin de module régulier et examen écrit de répétition

Type de l'examen

Examen écrit

Durée de l'examen

120 minutes

Aides autorisées

Les aides suivantes sont autorisées:

Aides électroniques autorisées

Aucune aide électronique autorisée

Autres aides autorisées

Résumé écrit à la main, 2 pages A4 maximum.

Exception : En cas d'examen électronique sur Moodle, des modifications des aides autorisées peuvent survenir. Dans ce cas, les aides autorisées seront annoncées par les enseignant-e-s avant l'examen.

Cas spécial: examen de répétition oral

Type de l'examen

Examen oral

Durée de l'examen

30 minutes

Aides autorisées

Sans aides