Implementing Cisco Data Center AI Infrastructure (DCAI)

ID DCAI **Prix** sur demande **Durée** 5 jours

A qui s'adresse cette formation

- Concepteurs Réseau
- Administrateurs Réseau
- Administrateurs de Stockage
- Ingénieurs Réseau
- Ingénieurs Systèmes
- Ingénieurs de Centre de Données
- Ingénieurs Systèmes Conseil
- Architectes de Solutions Techniques
- Intégrateurs / Partenaires Cisco
- Ingénieurs Terrain
- Administrateurs Serveurs
- Responsables Réseau
- Responsables de Programme
- Chefs de Projet

Cette formation prépare à la/aux certifications

Cisco Certified Network Professional Data Center (CCNP DATA CENTER)

Pré-requis

Il n'existe aucun prérequis formel pour cette formation. Cependant, il est recommandé de disposer des connaissances et compétences suivantes avant d'y participer :

- Architecture et opérations de calcul Cisco UCS
- Gamme de commutateurs Cisco Nexus et leurs fonctionnalités
- Technologies cœur des centres de données

Ces compétences peuvent être acquises via les offres de formation Cisco suivantes :

- Implementing Cisco NX-OS Switches and Fabrics in the Data Center (DCNX)
- Cisco Data Center Nexus Dashboard Essentials (DCNDE)
- Implementing and Operating Cisco Data Center Core Technologies (DCCOR)

Objectifs

A l'issue de la formation, vous devrez être en mesure de :

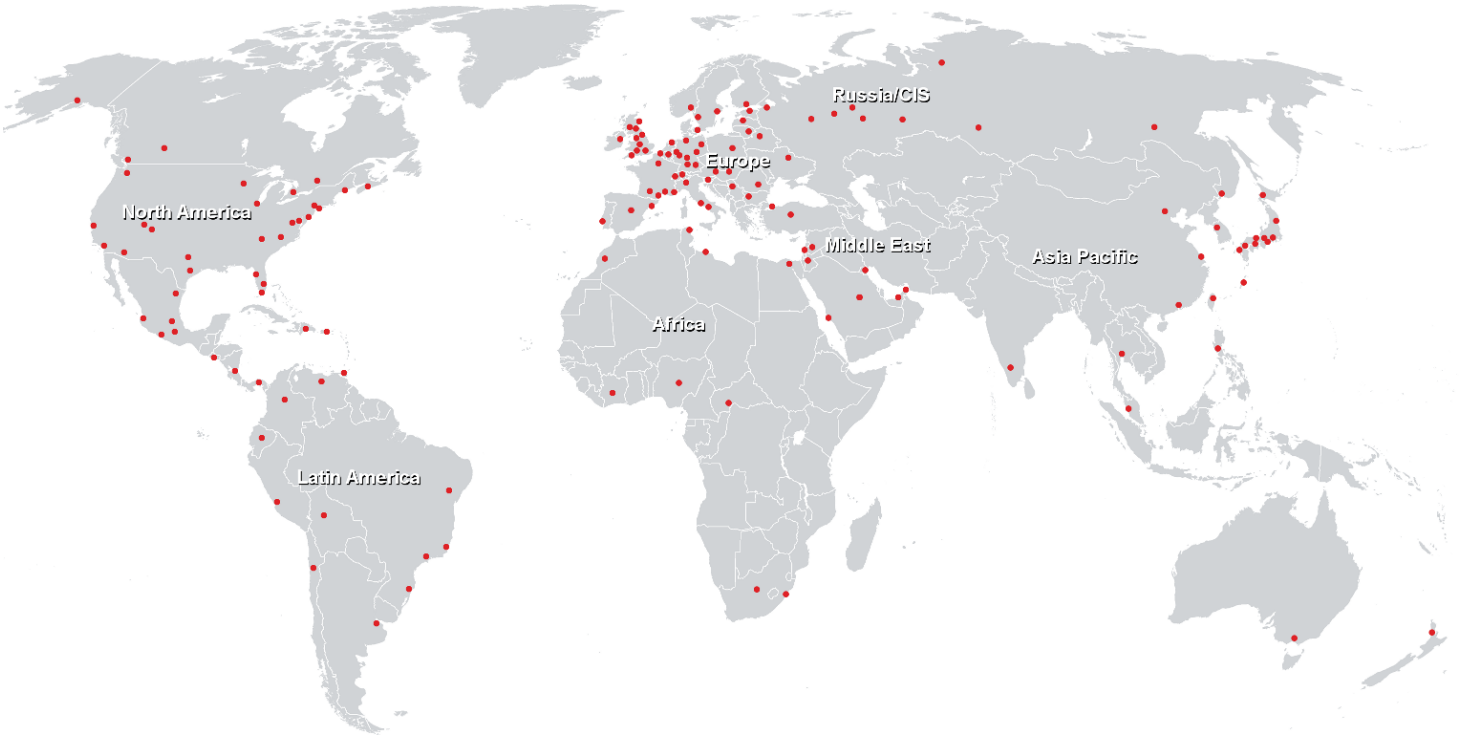
- Décrire les concepts clés de l'intelligence artificielle, en se concentrant sur l'IA traditionnelle, l'apprentissage automatique (machine learning) et l'apprentissage profond (deep learning), ainsi que leurs applications
- Décrire l'IA générative, ses défis et ses tendances futures, tout en examinant les nuances entre les approches traditionnelles et modernes de l'IA
- Expliquer comment l'IA améliore la gestion et la sécurité des réseaux grâce à l'automatisation intelligente, à l'analytique prédictive et à la détection d'anomalies
- Décrire les concepts clés, l'architecture et les principes de gestion de base des clusters IA/ML, ainsi que le processus d'acquisition, d'ajustement fin (fine-tuning), d'optimisation et d'utilisation de modèles ML préentraînés
- Utiliser les capacités de Jupyter Lab et de l'IA générative pour automatiser les opérations réseau, écrire du code Python et exploiter des modèles d'IA afin d'améliorer la productivité
- Décrire les composants essentiels et les considérations nécessaires à la mise en place d'une infrastructure IA robuste
- Évaluer et mettre en œuvre des stratégies efficaces de placement des charges de travail et garantir l'interopérabilité au sein des systèmes IA
- Explorer les normes de conformité, les politiques et les cadres de gouvernance applicables aux systèmes d'IA
- Décrire les pratiques d'infrastructure IA durable, en mettant l'accent sur la durabilité environnementale et économique
- Orienter les décisions relatives à l'infrastructure IA afin d'optimiser l'efficacité et les coûts
- Décrire les principaux défis réseau du point de vue des exigences applicatives IA/ML
- Décrire le rôle des technologies optiques et cuivre dans la prise en charge des charges de travail IA/ML en centre de données
- Décrire les modèles de connectivité réseau et les conceptions réseau
- Décrire les protocoles importants de couche 2 et de couche 3 pour l'IA et le fog computing dans le cadre du traitement distribué de l'IA
- Migrer les charges de travail IA vers un réseau IA dédié
- Expliquer les mécanismes et le fonctionnement des protocoles RDMA et RoCE
- Comprendre l'architecture et les fonctionnalités des fabrics

Ethernet haute performance
- Expliquer les mécanismes réseau et les outils de QoS nécessaires à la construction de réseaux RoCE sans perte et haute performance
- Décrire les mécanismes ECN et PFC, présenter Cisco Nexus Dashboard Insights pour la surveillance de la congestion, explorer l'impact des différentes phases des applications IA/ML sur l'infrastructure des centres de données, et inversement
- Présenter les étapes de base, les défis et les techniques liés au processus de préparation des données
- Utiliser Cisco Nexus Dashboard Insights pour surveiller les flux de trafic IA/ML
- Décrire l'importance du matériel spécifique à l'IA pour réduire les temps d'entraînement et prendre en charge les besoins avancés de traitement des tâches d'IA
- Comprendre le matériel de calcul requis pour exécuter des solutions IA/ML
- Comprendre les solutions d'intelligence et d'IA/ML existantes
- Décrire les options d'infrastructure virtuelle et les considérations associées lors du déploiement
- Expliquer les stratégies de stockage des données, les protocoles de stockage et le stockage défini par logiciel
- Utiliser NDFC pour configurer un fabric optimisé pour les charges de travail IA/ML
- Utiliser des modèles GPT hébergés localement avec le RAG pour des tâches d'ingénierie réseau

Contenu

- Fondamentaux de l'IA
- IA générative
- Cas d'usage de l'IA
- Clusters IA/ML et modèles
- Outils IA — Jupyter Notebook
- Infrastructure IA
- Placement des charges de travail IA et interopérabilité
- Politiques IA
- Durabilité de l'IA
- Conception de l'infrastructure IA
- Défis réseau clés et exigences pour les charges de travail IA
- Transport IA
- Modèles de connectivité
- Réseau IA
- Migration d'architecture vers un réseau IA/ML
- Protocoles au niveau applicatif
- Fabrics convergés à haut débit
- Conception de fabrics sans perte
- Visibilité de la congestion
- Préparation des données pour l'IA
- Performances des données des charges de travail IA/ML
- Matériel d'activation de l'IA
- Ressources de calcul
- Solutions de ressources de calcul
- Ressources virtuelles
- Ressources de stockage
- Mise en place d'un cluster IA
- Déployer et utiliser des modèles GPT open source pour le RAG
- Exploitation et supervision de l'infrastructure IA
- Dépannage de l'infrastructure IA
- Dépannage des problèmes courants des fabrics IA/ML

Centres de formation dans le monde entier





Fast Lane Institute for Knowledge Transfer (Switzerland) AG

Husacherstrasse 3
CH-8304 Wallisellen
Tel. +41 44 832 50 80

info@flane.ch, https://www.flane.ch