Data Center Networking pour l'IAML : GPU networking, RoCEv2, InfiniBand et UEC

Lien : https://innov-systems.com/formation/data-center-networking-pour-liaml-gpu-networking-rocev2-infiniband-et-uec

DURÉE
5 jours (35h)

RÉFÉRENCE
RT70

CATÉGORIE
Data Center Networking et Fabrics EVPN-VXLAN

Innov Systems

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Comprendre les exigences réseau des workloads IA/ML
- Maîtriser les architectures GPU networking modernes
- Différencier InfiniBand, RoCEv2 et Ethernet avancé
- Concevoir des clusters IA haute performance
- Optimiser les communications distribuées GPU
- Réduire la latence des entraînements IA
- Améliorer la scalabilité des infrastructures ML
- Comprendre les standards émergents comme UEC
- Superviser les performances réseau IA en production
- Concevoir des architectures data center optimisées IA
- Intégrer stockage et compute dans les pipelines IA
- Étudier des cas concrets de data centers IA industriels

POUR QUI ?

- Architectes data center
- Ingénieurs réseau
- Ingénieurs IA/ML
- Machine learning engineers
- HPC engineers
- Cloud architects
- DevOps engineers
- SRE (Site Reliability Engineers)
- Data engineers
- Ingénieurs infrastructure IT
- Chercheurs en IA et calcul haute performance
- Toute personne impliquée dans les infrastructures IA à haute performance

Innov Systems

Programme détaillé

1 / Introduction au networking pour l'IA/ML en data center

- Explosion des workloads IA et besoins réseau spécifiques
- Contraintes des entraînements distribués
- Rôle critique du réseau dans les performances IA

2 / Architecture des clusters IA/ML

- GPU clusters et interconnexion haute performance
- Topologies réseau (fat-tree, spine-leaf)
- Séparation compute, storage et data pipelines

3 / GPU networking et communication distribuée

- Communication inter-GPU (all-reduce, NCCL)
- Optimisation des échanges entre nœuds
- Réduction de la latence inter-nœuds

4 / InfiniBand pour l'IA haute performance

- Principes et architecture InfiniBand
- Latence ultra-faible et débit élevé
- Cas d'usage en supercomputing IA

5 / RoCEv2 (RDMA over Converged Ethernet)

- Fonctionnement du RDMA sur Ethernet
- Comparaison avec InfiniBand
- Cas d'usage dans les data centers modernes

Innov Systems

6 / UEC (Ultra Ethernet Consortium)

- Objectifs de l’Ultra Ethernet
- Évolution des standards Ethernet pour l’IA
- Interopérabilité et scalabilité future

7 / Optimisation du trafic IA distribué

- All-reduce et gradients distribués
- Réduction de la congestion réseau
- Scheduling des communications GPU

8 / Storage et data pipelines haute performance

- Accès aux datasets massifs
- Parallel file systems (Lustre, Ceph)
- Optimisation I/O pour training IA

9 / Monitoring et observabilité réseau IA

- Mesure de latence et jitter
- Détection des goulots d’étranglement
- Telemetry réseau en temps réel

10 / Études de cas et architectures industrielles

- Supercalculateurs IA
- Clusters cloud GPU (AWS, Azure, GCP)
- Bonnes pratiques de design réseau IA

Innov Systems

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

01 au 05 Fév. 2027	Présentiel - Casablanca
29 Mars au 02 Avr. 2027	Présentiel - Casablanca
24 au 28 Mai 2027	Présentiel - Casablanca
19 au 23 Juil. 2027	Présentiel - Casablanca
13 au 17 Sep. 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés