

Intelligence artificielle et machine learning: les atouts de la solution Swarm

“

D'un point de vue plus abstrait, le comportement en essaim (swarm behavior, en anglais) est le mouvement collectif d'un grand nombre d'entités automotrices. Du point de vue du modélisateur mathématique, il s'agit d'un comportement émergent résultant de règles simples suivies par des individus et n'impliquant aucune coordination centrale.

Wikipedia

”

LE BIOMIMÉTISME COMME SOURCE D'INSPIRATION

Le biomimétisme ou biométisme consiste à imiter des modèles, des systèmes et des éléments de la nature dans le but de résoudre des problèmes humains complexes. (Wikipedia)

La méthode brevetée de synchronisation de paramètres de cluster partagé de DataCore repose sur un algorithme d'intelligence artificielle (IA) ou de machine learning (apprentissage automatique) (ML) utilisé en production depuis plus de 1,2 milliard d'années, et ce n'est pas une plaisanterie ! Toutefois, notre application du biomimétisme sans doute la plus fondamentale est la conception de notre technologie de cluster massivement parallèle : le comportement de machine learning d'un essaim (swarm, en anglais), composé d'agents quasi-indépendants (les nœuds) ne partageant qu'un ensemble minimal de capteurs, de paramètres et d'objectifs.

Le machine learning ou « apprentissage automatique » est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'« apprendre » à partir de données, c'est-à-dire d'améliorer leurs performances pour résoudre des tâches spécifiques, sans être explicitement programmés pour chacune. (Wikipedia)

Alors que les configurations de cluster d'inspiration géométrique semblent théoriquement former un cercle parfait, en réalité, un cluster en temps réel se révèle souvent un environnement difficile, loin d'être déterministe. Seules des règles et des concepts très simples survivent à ces moments déterminants, tandis que d'autres s'effondrent sous le poids de leur structure. Avec l'architecture DataCore Swarm, même en cas de problème de « split-brain » (deux maîtres) du réseau, les données qui sont accessibles peuvent être lues : aucun effet transitoire de reconfiguration du cluster ne viendra jamais remettre en question ce simple fait. En raison de leur nature

indépendante, les nœuds ne peuvent pas dégrader la structure ni se compromettre les uns les autres. Une formation d'oiseaux migrateurs privée de quelques individus malchanceux reste une formation, tandis qu'un cercle qui perd quelques segments n'est plus un cercle. Cela ne signifie pas que DataCore Swarm est totalement insensible aux pannes

matérielles. Il gère simplement les situations d'urgence de manière raisonnable et autonome. Comme dans la nature, les fonctions essentielles sont privilégiées en local, tandis que d'autres sont temporairement mises au second plan. Le scénario « split-brain » mentionné plus haut n'interfère pas avec les fonctions essentielles de Swarm (c'est-à-dire la lecture et l'écriture d'objets), mais peut entraîner une sur-réplication temporaire. Une fois toutes les connexions correctement rétablies, les fonctions secondaires reviennent en force et éliminent rapidement les répliques en surplus. En fait, les données sont encore mieux protégées pendant ce « moment de crise ».

En tout cas, pour qu'une architecture en essaim fonctionne vraiment bien, il faut surtout que ses nœuds soient simples et agiles. DataCore Swarm répond parfaitement à ces deux exigences

LA SIMPLICITÉ D'ABORD

Tout d'abord, pour définir nos nœuds, nous avons commencé par utiliser du matériel et un environnement de système d'exploitation standard. Nous avons écarté tout ce qui n'était pas indispensable, notamment dans l'aspect logiciel, et ajouté juste le strict nécessaire pour un système physique nu. Nous avons mis la pile logicielle restante, le système d'exploitation et tous les autres aspects dans un livrable constitué d'une seule image, qui est chargé dans la RAM au moment du démarrage. Ce déploiement bare-metal s'effectue en démarrant à partir du réseau PXE ou même à partir d'une simple clé USB. Tous les logiciels sont conservés dans la RAM. Jamais aucune installation de disques, aucun correctif ni aucune mise à niveau n'est nécessaire. L'espace disque ne sert que pour la charge utile