

GTC 2020 : six annonces à retenir chez NVIDIA

BlueField, Jarvis, Jetson, Merlin, Omniverse, SuperPOD... Autant de marques du portefeuille NVIDIA qui ont fait l'objet d'annonces ce 5 octobre, au premier jour de la GTC 2020.

1 – Maxine : du *deep learning* dans la visio

Il y a aussi de nouvelles venues dans ce portefeuille, comme [Maxine](#). Cette plate-forme de diffusion vidéo exploite des réseaux neuronaux – en particulier [de type antagoniste génératif](#). Ils permettent notamment de simuler un rendu « de face » quand une personne tourne le visage. Mais aussi de réduire la consommation de bande passante. La méthode : ne pas transmettre pas un flux continu d'images, mais des données relatives à la localisation de « points-clés » autour des yeux, du nez et de la bouche.

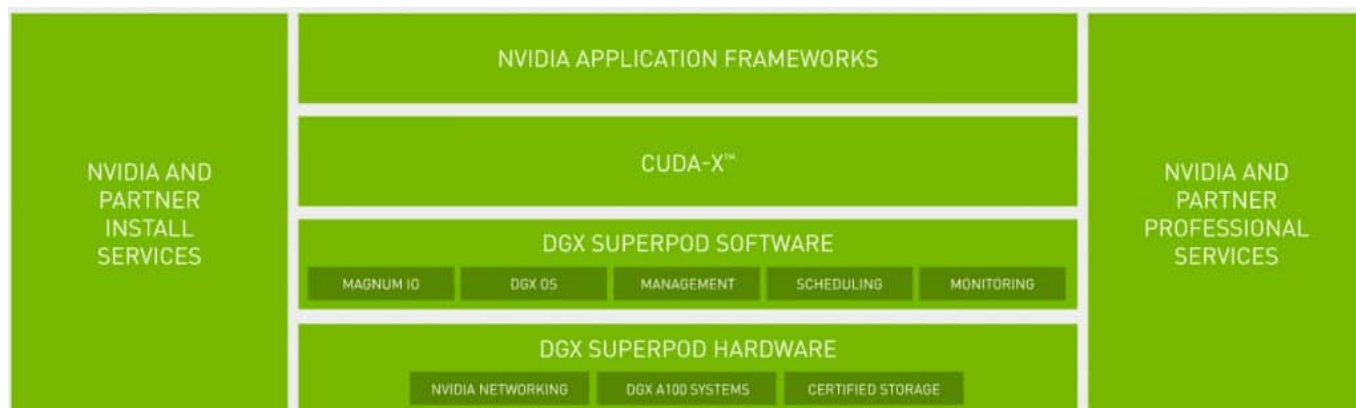
2 – Jetson Nano 2GB : 59 \$ pour expérimenter

D'IA, on parle aussi avec la [Jetson Nano 2GB](#). Il s'agit d'une carte de développement positionnée en entrée de gamme (59 \$; disponibilité fin octobre). Elle embarque 4 cœurs CPU (Cortex-A57), 128 cœurs GPU (Maxwell) et 2 Go de RAM (LPDDR4). La connectique comprend Ethernet (+ Wi-Fi 802.11ac), HDMI, USB 3, USB 2 (trois ports dont un micro-USB), GPIO et MIPI (pour une caméra).

3 – NVIDIA DGX SuperPOD : jusqu'à 700 Pflops

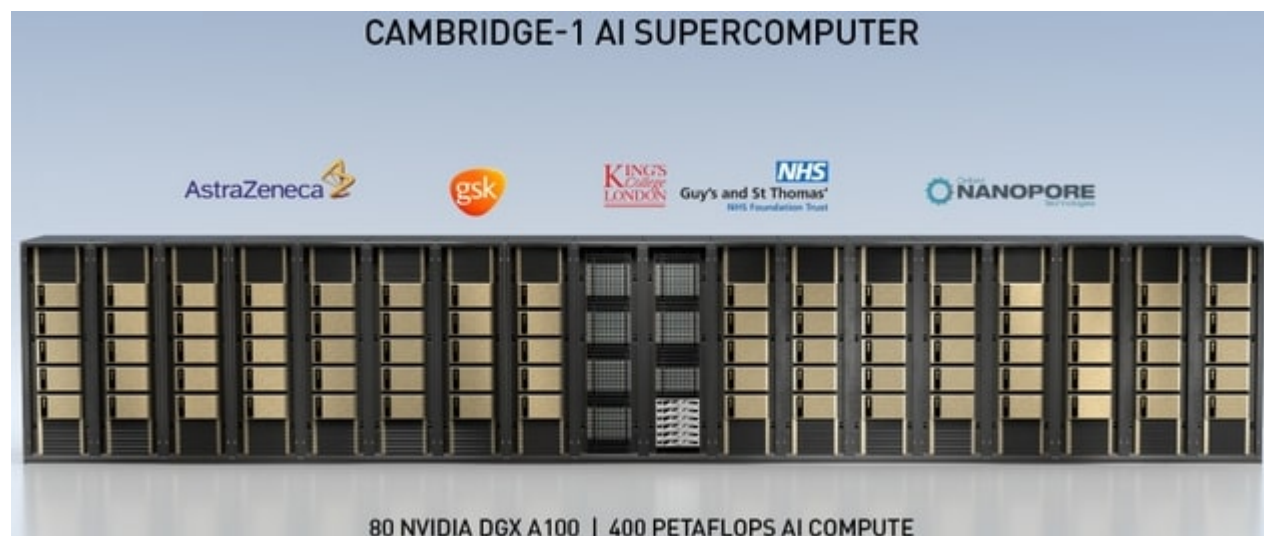
Toujours sur le volet IA, on aura noté le lancement officiel de [DGX SuperPOD Solution for Enterprise](#).

Cette offre d'infrastructure associe calcul, stockage, réseau, outils de gestion et services de data science. Elle constitue l'implémentation « made in NVIDIA » de l'architecture de référence DGX SuperPOD.



L'unité de base comprend 20 nœuds basés sur des GPU A100. On peut en « clusteriser » jusqu'à sept, pour une puissance de 700 Pflops, 10 To de stockage et 200 Gbit/s de bande passante.

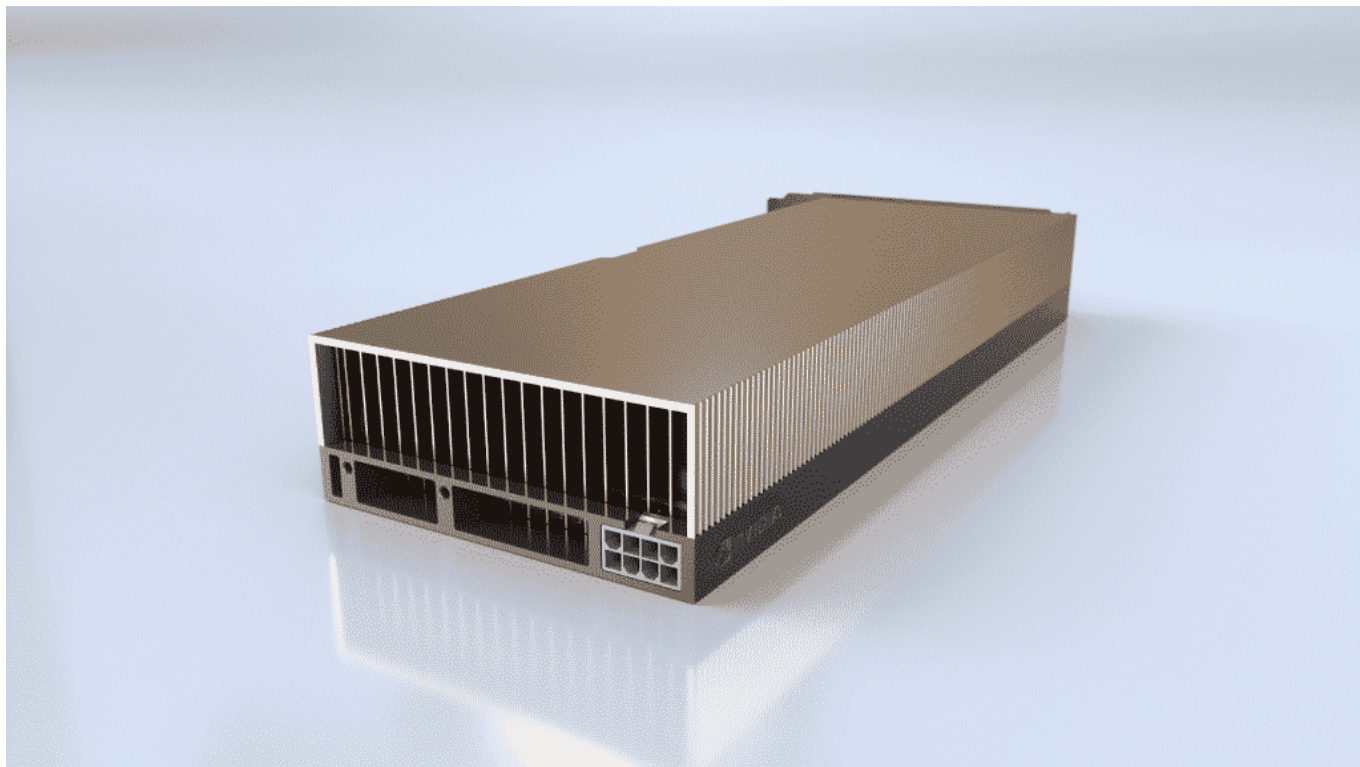
DGX SuperPOD for Enterprise constituera la base du superordinateur [Cambridge-1](#), que NVIDIA entend mettre en place d'ici à la fin de l'année. Puissance annoncée : 400 Pflops.



4 – Ampere s'installe chez NVIDIA

En matière de GPU, on aura noté l'arrivée prochaine de deux modèles : [RTX A6000 et A40](#). Le premier se destine aux stations de travail de bureau. Dell, HP et Lenovo entendent en équiper leur gamme, à partir de fin 2020. Le second vise les serveurs. Cisco, Dell, Fujitsu, HPE et Lenovo sont dans la boucle pour début 2021.

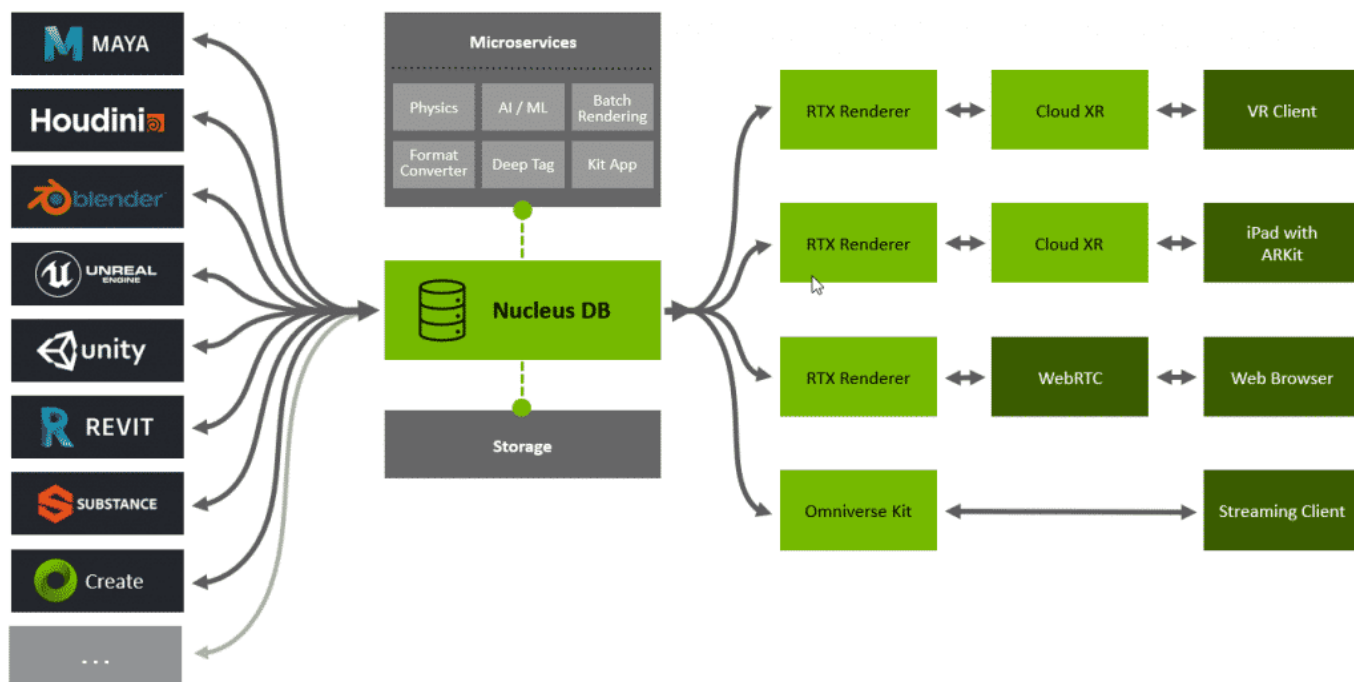
RTX A6000 et A40 (illustré ci-dessous) ont une même base. À savoir : 48 Go de GDDR6, des cœurs RT de 2^e génération et Tensor de 3^e génération, pour une enveloppe thermique de 300 W.



5 – Jarvis, Merlin et Omniverse en bêta ouverte

Les cartes RTX forment le socle d'exécution d'Omniverse, que NVIDIA lance en bêta ouverte à l'occasion de la GTC.

Une quarantaine d'organisations ont participé à la première phase d'expérimentation de cette plate-forme collaborative censée favoriser l'interopérabilité des solutions de création 3D. Elle repose pour cela sur des composants ouvertes, dont le format USD (Universal Scene Description) de Pixar et son adaptation MDL (Material Definition Library) de NVIDIA pour faciliter les échanges entre applications.



Passent aussi en bêta ouverte, [Jarvis et Merlin](#). Ces deux *frameworks* permettent de développer des modèles de deep learning. Respectivement sous la forme d'agents conversationnels et de systèmes de recommandation.

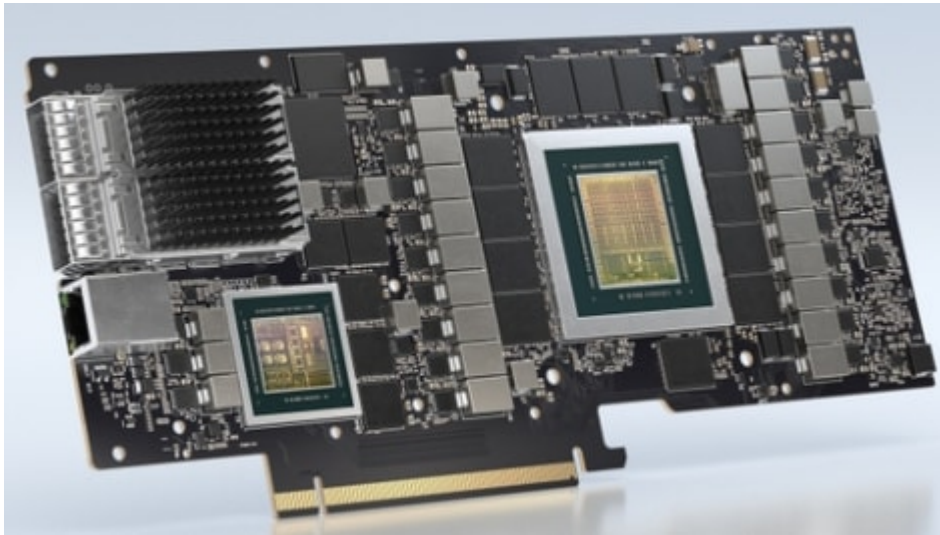
6 – La famille BlueField s'agrandit

La GTC aura aussi été l'occasion d'évoquer les [DPU BlueField de deuxième génération](#). Ils sont d'après NVIDIA, capables de réaliser autant de tâches que 125 cœurs de CPU. Il en existe deux modèles :

- BlueField-2, qui associe cœurs Arm et SmartNIC Mellanox ConnectX-6 Dx
- BlueField-2X, qui ajoute un GPU sur architecture Ampere

Asus, Atos, Dell, Fujitsu, Gigabyte, H3C, Inspur, Lenovo, Quanta/QCT et Supermicro ont l'intention d'intégrer ces DPU dans leur offre, à partir de l'an prochain.

VMware a annoncé sa collaboration avec NVIDIA dans le cadre de l'initiative [Project Monterey](#). [Red Hat](#) et [Canonical](#) se sont exprimés dans le même sens pour leurs distributions Linux respectives. Check Point s'est par ailleurs [engagé](#) à faire la jonction avec son agent Infinity Next Nano.



Illustrations © NVIDIA