

Pour son Cloud, Microsoft met ARM en concurrence avec le x86

Microsoft a profité du Open Computer Summit de Santa Clara pour annoncer l'arrivée de puces ARM sur son projet Olympus, le design de serveur Open Source que Redmond a présenté en octobre dernier. La conséquence ? Les datacenters de nouvelle génération du n°1 mondial du SaaS et du n°2 mondial du IaaS (derrière AWS) devraient prochainement accueillir des serveurs motorisés par des SoC (System on a chip) fournis par les industriels du monde ARM. Dont Qualcomm et Cavium. Une inflexion nette par rapport au tout x86 qui domine chez les grands noms du Cloud.

Qualcomm livre à Microsoft sa puce Centriq 2400, gravée en 10 nm et offrant 48 cœurs (ci-dessus, le serveur qui en découle). Le ThunderX2 ARMv8-A de Cavium monte, lui, à 54 cœurs. Ces deux industriels ont développé une carte-mère compatible avec le projet Olympus, la garantie d'une intégration simple dans l'architecture des datacenters de Microsoft.

Déjà des plans de déploiement

Pour l'instant toutefois, Redmond se réserve la version de Windows Server pour ARM qu'il emploie sur ces architectures. Le premier éditeur mondial indique que les serveurs ARM sont actuellement déployés en tests, aux côtés des architectures x86 classiques, sur des tâches liées à la recherche, à l'indexation, au stockage, au Big Data ou au Machine Learning, là où les caractéristiques d'ARM se révèlent les plus intéressantes. « *Ce que nous avons constaté est assez incroyable* », écrit Leendert van Doorn, un ingénieur d'Azure dans un [billet de blog](#), qui ajoute que certains services Cloud maison ont d'ores et déjà des « *plans de déploiement futurs* » basés sur ARM. Redmond ajoute que cette architecture est aussi intéressante en raison des capacités d'extension du jeu d'instruction. En creux, une critique du x86, qui est jusqu'à aujourd'hui la référence dans les datacenters du Cloud et dans ceux des entreprises.

Le pied que Microsoft vient ainsi de mettre dans les architectures ARM est un signe concret de l'intérêt grandissant des industriels du Cloud pour cette famille de composants, qui a conquis le monde des terminaux mobiles. Une façon aussi de mettre sous pression Intel, l'acteur ultra-dominant du x86. L'intérêt de Microsoft pour ARM peut s'expliquer tant par les gains de performances que ces processeurs devraient procurer sur certaines tâches, que pour des questions de coût. La concurrence entre industriels est plus affirmée dans le monde ARM que dans le x86 où AMD peine à réellement bousculer le géant Intel. Et ces puces consomment également moins d'énergie. Or, pour prendre des parts de marché dans le Cloud, les géants du Cloud ont besoin de bâtir au meilleur coût leurs infrastructures à l'échelle du globe. Récemment, Redmond aurait ainsi brutalement [réduit ses commandes auprès de HPE](#), faute d'un accord sur les tarifs du matériel vendu par le leader mondial du marché des serveurs.

Adapter ARM au Cloud de Microsoft ?

Mais, pour les grands opérateurs de Cloud comme Microsoft, la réflexion peut aller au-delà de cet aspect concurrentiel. « *En raison de l'échelle nécessaire à certains services Cloud, donc au nombre de machines nécessaires à leur fonctionnement, il devient économiquement envisageable d'optimiser le hardware pour les tâches à effectuer et non plus l'inverse, même si cela signifie un changement de l'architecture du jeu d'instruction (Instruction Set Architecture, ISA)* », écrit Leendert van Doorn. Autrement dit, Redmond pourrait se servir de ses volumes de commandes pour pousser les fabricants de composants à bâtir des puces adaptées à son Cloud. Et il semblerait que, pour l'instant en tout cas, les industriels du monde ARM soient les plus réceptifs à ce changement de paradigme. La société dirigée par Satya Nadella explore d'ailleurs depuis déjà longtemps cette piste de l'adaptation du silicium aux workloads du Cloud. Azure exploite par exemple déjà [des puces reprogrammables](#).

Notons que des versions du projet Olympus bâties sur les puces Intel sont évidemment également au programme. Elles reposent sur des processeurs de la génération Skylake, dont la déclinaison pour serveurs (Skylake-EP) intègre le jeu d'instructions AVX-512 permettant de travailler en une seule opération sur un ensemble de données de 512 bits. Microsoft travaille également avec AMD, sur une déclinaison de son design de référence intégrant [les futurs processeurs Naples](#) de la firme de Sunnyvale.

A lire aussi :

[Cloud : comment Microsoft repense l'architecture de ses datacenters Azure](#)

[Processeurs ARM pour serveurs : AMD n'y croit \(presque\) plus](#)

[Online atomise les prix de ses serveurs ARM Scaleway](#)