

# Stockage froid : la réponse de Facebook à l'explosion de ses données

Quotidiennement, les utilisateurs de **Facebook** déposent 300 millions de photos sur leurs comptes. Et certains jours exceptionnels, comme le prochain Halloween, les téléchargements sur le réseau peuvent dépasser le milliard de photos. Au final, Facebook se retrouve dans l'obligation d'ajouter tous les mois 7 Po de capacité de stockage...

Pour **Jay Parikh**, vice-président Infrastructure Engineering de Facebook, qui a présenté la problématique lors de la conférence Structure Europe qui vient de se tenir à Amsterdam, il est temps de réagir. Car le réseau social se trouve confronté à une double problématique : les volumes de photos stockées augmentent, mais il ne peut les supprimer.

## Revoir l'architecture matérielle du datacenter

Une solution pourrait consister à adopter la même stratégie qu'Amazon, qui, avec son offre d'archivage Glacier, propose à ses clients de placer les données les moins chaudes sur une infrastructure de stockage moins réactive (jusqu'à plusieurs heures pour obtenir un accès à une donnée), mais en contrepartie moins chère. Intéressant, pour peu que l'utilisateur en accepte les contraintes.

Facebook, qui veut réduire ses coûts de stockage, ne peut adopter cette approche. Ni supprimer des photos, même s'il s'agit de documents qui n'ont pas été consultés depuis 10 ans. Le risque de mécontenter ses utilisateurs est trop grand, l'accès à un photo doit pouvoir se faire à n'importe quel moment et très rapidement.

Pour Jay Parikh, la solution consiste tout d'abord à revoir l'architecture des datacenters de Facebook. Il faut non plus raisonner en optimisation de la consommation en fonction des tâches à accomplir, mais à l'inverse en espaces de plus en plus importants, sans y associer autant de puissance énergétique.

## Open Compute Project

C'est pourquoi Facebook a lancé un projet qui vise à traiter cette problématique en s'appuyant sur un logiciel qui va gérer la migration des données d'un modèle d'architecture de datacenter à un autre en fonction de l'évolution de la données du 'chaud' vers le 'froid'.

Ce projet s'appuie sur Open Compute Project (OCP), lancé en avril 2011, qui ambitionne de créer des designs de matériels open source dédiés au datacenter.

En mais dernier, Facebook a ouvert son projet et a été rejoint par AMD, Canonical, Fidelity, HP, Quanta, Salesforce.com, Supermicro, Tencent, et VMware. Les serveurs (et donc les serveurs de stockage) répondent également à la spécification Open Rack originaire de HP et de Dell.

L'ensemble du projet OCP, avec le design optimisé de certains matériels, vise à optimiser l'architecture du datacenter afin de réduire sa consommation énergétique.

## [À suivre en page 2 : Open Compute Project et le logiciel façon Facebook](#)

### Open Compute Project

Mis en application par Facebook sur son datacenter de Prineville, celui qui se trouve en face du futur datacenter d'Apple (lire « [Apple lance la première phase de son nouveau datacenter à énergie 100 % renouvelable](#) »), l'approche OCP a permis de réduire de 38 % la consommation énergétique, et ainsi de 24 % le coût du datacenter par rapport aux datacenters existants du réseau social.



Il va de soi qu'OCP est destiné en priorité aux datacenters de très grande envergure, consommant des centaines de milliers de serveurs, les seuls à pouvoir imposer aux fabricants de concevoir le design de leur architecture selon leurs directives.

Pour autant, la présence d'acteurs comme AMD, HP ou Supermicro dans OCP est une indication forte sur l'avenir de l'architecture open source développée selon les standards de Facebook. Surtout que le projet n'est pas propriétaire : il reste accessible à tous.

### Du hardware au software OCP

Est-ce qu'il en sera de même des travaux de Facebook sur le logiciel ? Car s'il est vrai que le projet OCP porte sur l'optimisation du matériel, il lui faut un logiciel de qualité pour piloter l'ensemble.

Un sujet encore plus sensible lorsqu'il s'agit pour Facebook de gérer et de faire migrer automatiquement les données sur différentes couches matérielles de stockage, jusqu'au stockage froid pris dans le sens du plus économique possible, mais sur lequel les données devront rester accessibles très rapidement.

Facebook semble vouloir jouer le jeu. En tout cas, Jay Parikh a donné une indication forte quant aux délais qu'il s'impose : le stockage froid devra être implémenté dans l'infrastructure de Facebook dans les une à deux prochaines années.

Pour accéder à l'Open Compute Project et à la spécification Open Rack 1.0 : [opencompute.org](https://opencompute.org)