

Rapport

Sur l'article : Arbitration Policies for
On-Demand User-Level I/O Forwarding on
HPC Platforms

GOEDEFROIT Charles

30 octobre 2022

1 L'objectif de l'article

L'article s'intéresse à différentes politiques d'accès aux nœuds de données en fonction des schémas d'accès aux données par l'application.

L'article propose une politique basée sur le problème du sac à dos (Knapsack problem) à choix multiple. Ce problème cherche à maximiser la bande passante globale en donnant plus de nœuds I/O aux applications qui en ont le plus besoin.

Il propose aussi une solution qui permet d'utiliser, à la demande, au niveau utilisateur et pendant l'exécution, différentes politiques de *I/O forwarding*.

Il montre que leur technique augmente de façon transparente l'utilisation de la bande passante globale jusqu'à 85% par rapport à la politique statique utilisée par défaut. Il le montre avec de nombreuses expérimentations et directement sur une infrastructure (live setup).

2 Le contexte de l'article

Le *I/O forwarding* est une technique très utilisée dans le HPC. Cette technique permet d'augmenter les performances, en diminuant la contention, d'accès aux serveurs de stockage de données tout en étant transparent pour l'utilisateur.

Évite l'accès direct à la machine de stockage parallèle

Le *I/O forwarding* consiste à un groupe de nœuds qui résolve les requêtes de l'application et les transmet au système de fichiers parallèle. Permet aussi de changer la forme des requêtes (Ex : au lieu de plusieurs petites requêtes on en fait une seule grande).

La technique habituelle est d'assigner statiquement un nœud de stockage à un nœud de calcul. Ces liens ne correspondent pas tous le temps au transfert de données d'un nœud à l'application à besoin. Cette technique amène à une mauvaise utilisation des ressources.

3 The problem

4 Explication des algo

5 Le perfs

6 Les autre publications

7 Definition *

La contention : forte tentions sur un serveur (beaucoup de demandes).