



PIDR : Etude comparative SSH – Tatktuk – UDPCast

Bellegarde Loïc
Coelho Matthieu

Encadrant : Tomasz Buchert
Nussbaum Lucas

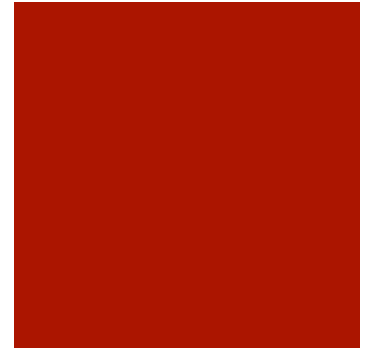


Sommaire



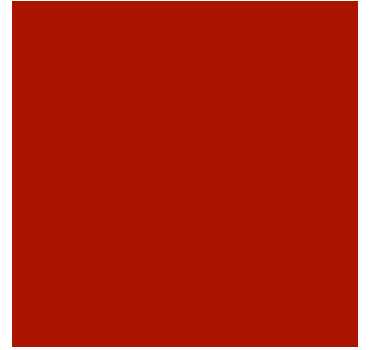
- Introduction
- Objectifs
- Plateforme Grid5000
- Protocole SSH
- Taktuk
- UDPCast
- Comparatif
- Suite du projet
- Conclusion

Introduction



- La distribution de données dans des clusters de calcul est un problème difficile mais très important
 - Diverses approches ont été considérées (distribution en chaîne, hiérarchique, P2P, etc.)
- ⇒ Manque de fiabilité, inefficacité ou difficultés avec passage à grande échelle

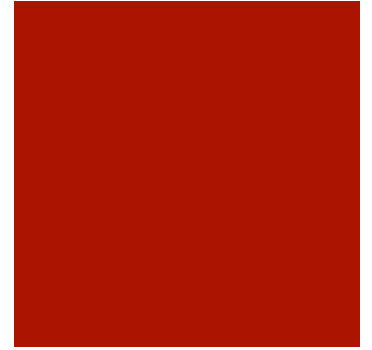
Objectifs



Evaluer différentes stratégies basées sur le broadcast fiable de données

- SSH
 - Taktuk
 - UDPCast
-
- Optimiser UDPCast en modifiant les paquets de correction d'erreurs (*fountain codes*)

Grid5000



Plateforme de calcul expérimentale

Réseau de plusieurs ordinateurs (Cluster)

Répartit sur 10 sites dans toute la France (Bordeaux, Lyon, Nancy, Lille...)

1171 Noeuds

2218 Processeurs

Protocole SSH



- Secure Shell (SSH)
- Créé en 1995
- Connexion à distance à un ordinateur pour obtenir un shell
- Exécuter des commandes depuis la frontale

Exemple :

Copier le fichier « nomfichier » vers « nomNode »/tmp

```
scp -r nomfichier root@${nomNode}:/tmp/
```

Le temps de transfert, évolue linéairement suivant le nombre de nœud !

Taktuk



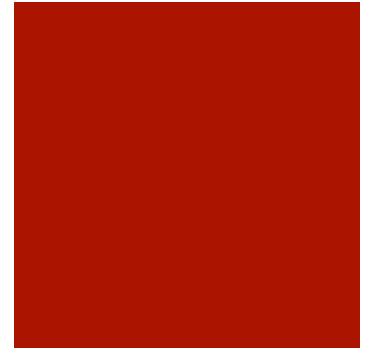
- Outil permettant le déploiement en parallèle d'exécution de commande
- Créé en 2004 par **Guillaume Huard, Cyrille Martin**

Exemple :

Copie le fichier « nomfichier » vers « nomNode »/tmp

```
taktuk -s -l root -f nomNode broadcast put [« nomfichier » ] [ /tmp]
```

UDPCast



Outil de transfert de fichier

Distribue les données à plusieurs machines dans le réseau simultanément

Le fonctionnement repose sur la propriété de multicast du protocole UDP

Séparé en deux programmes distincts :

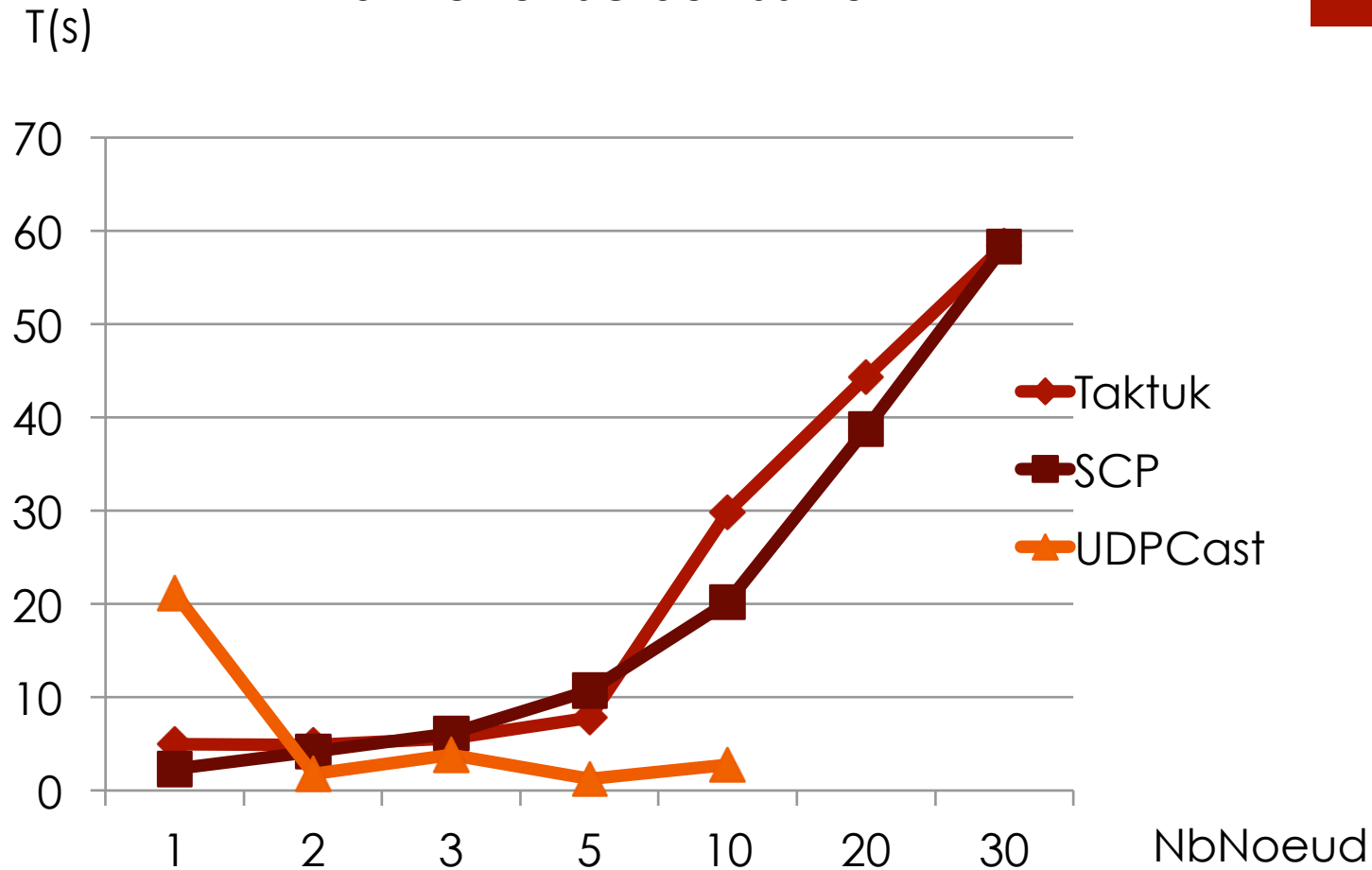
- * udp-receiver : retrouve (via broadcast) un udp-sender sur le réseau et s'y connecte.

- * udp-sender : attend la connexion des udp-receiver et quand une touche est pressée (d'un côté ou de l'autre) le transfert est lancé.

Il propose également différentes options qui seront manipulées par la suite du projet.

Comparatif

Un fichier ISO de 100mo

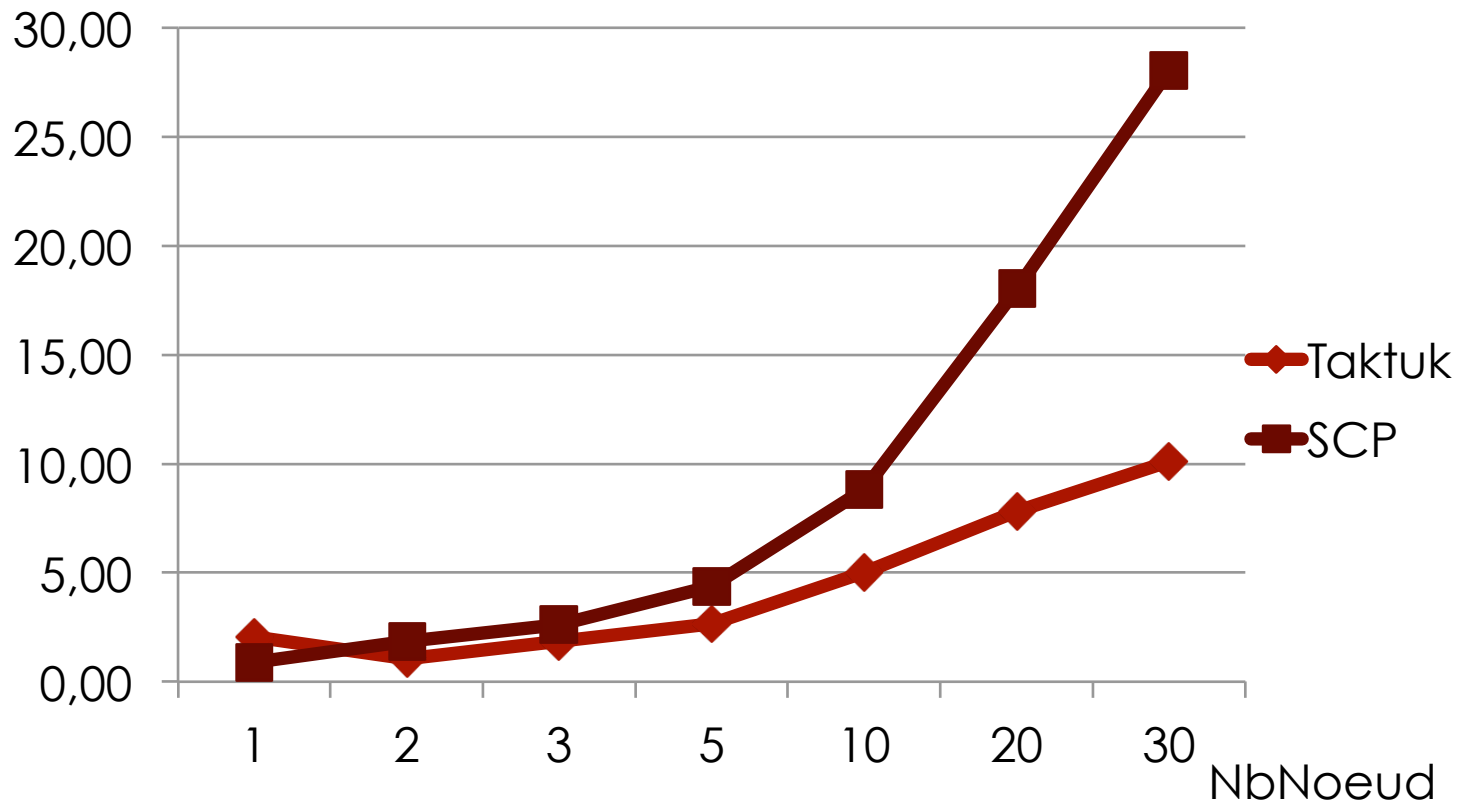


Comparatif

Un fichier ISO de 16,4mo contenant 434 éléments



T(s)



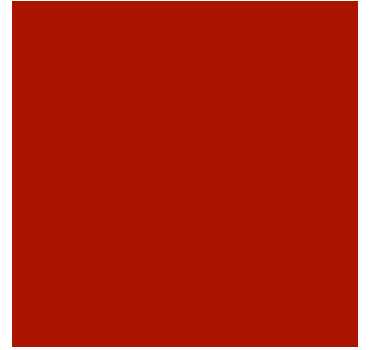
Suite du projet



- Effectuer différents jeux de tests
 - en augmentant le nombre de noeuds
 - en changeant le nombre de fichier et leur taille
- Améliorer UDPCast en implémentant d'autres stratégies de diffusions :

Fountain codes (broadcast classique avec ajout de paquets de correction d'erreurs)

Conclusion



- UDPCast semble être plus performant
- Implantation facile dans Grid5000