

Projet PLE

Univerité de Bordeaux 2020-2021

Alexandre Erard,
Maxime Gresse

December 23, 2020

1 Infrastructure

- Les utilisateurs de twitter produisent environs 504.10^6 tweets par jours. Sachant que 1 tweet fait environ 5Ko.

$$504.10^6 \times 5 \times 3 = 756.10^7 Ko = 7,56To$$

Notre Cluster possède actuellement 18.16To de stockage. Donc on peut stocker:

$$\frac{18,16}{7,56} \approx 2,40$$

un peu plus de 2 jours de tweet.

- Pour 2 jour de tweet on obtient un total de $7,56 \times 2 = 15To$. Un blocs faisant 128Mo on obtient 117187 blocs disponibles.

$$\frac{1,5.10^7}{128} \approx 117187$$

- 5ans de tweets équivaut à $7,56 \times 365 \times 5 = 13797To$ de données. Nous disposons actuellement de 20 machines possédant en moyenne

$$\frac{18.16}{20} = 0.93 = 930Go$$

En gardant un stockage de 930Go par machine il nous faudrait alors 14816 machines.

$$\frac{13797 - 18.6}{0.93} \approx 14815.5$$