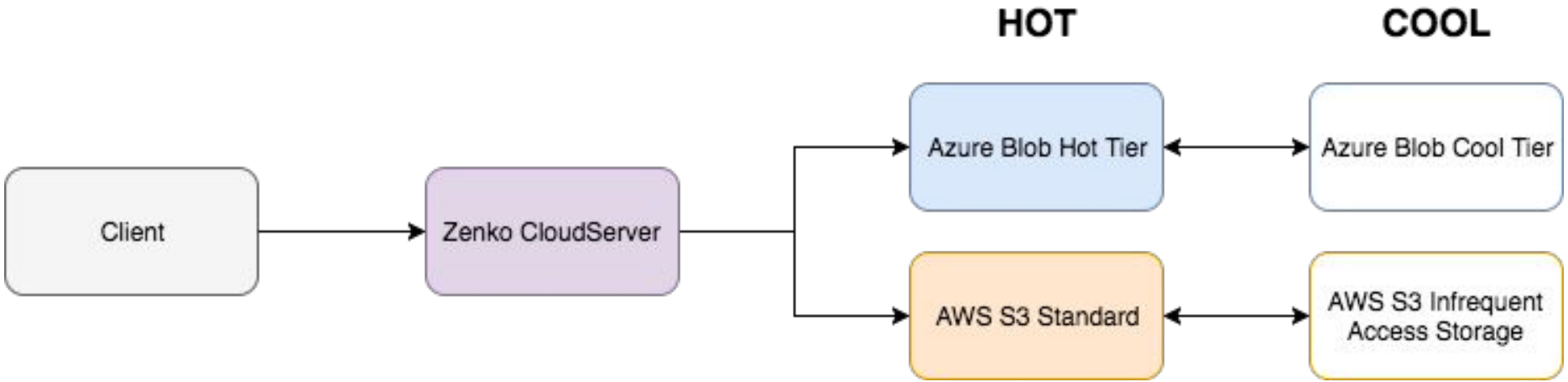
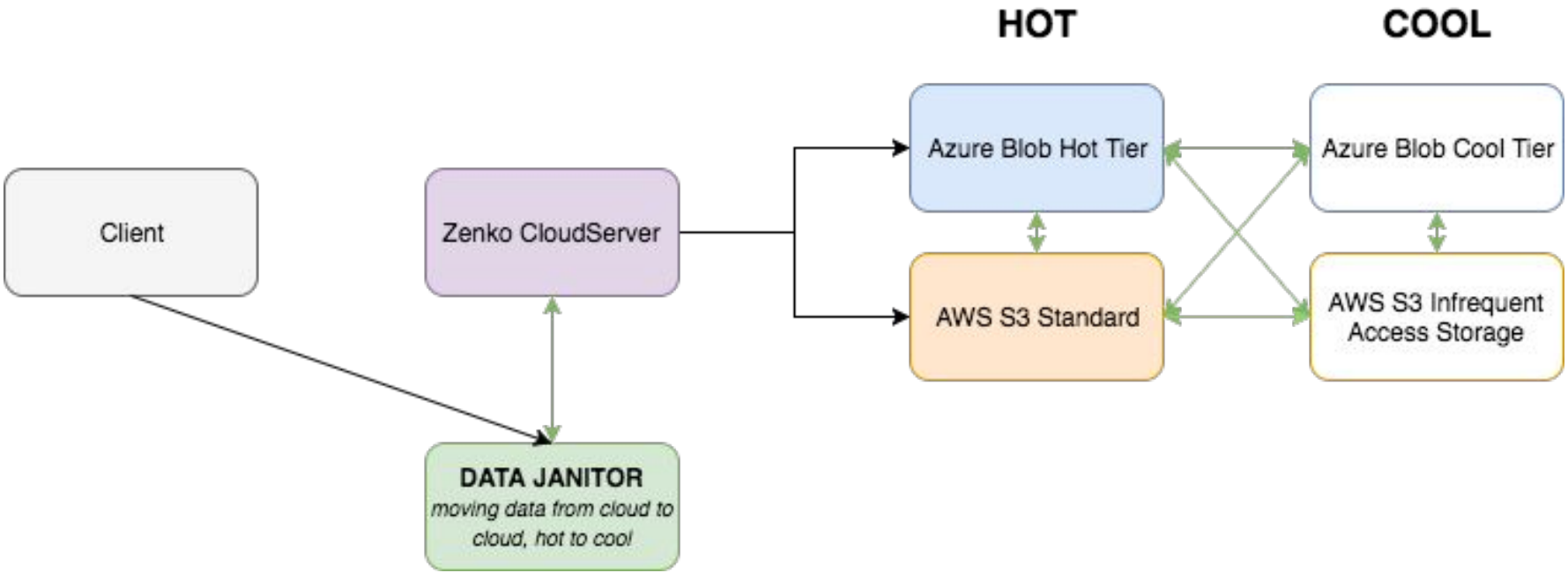
Data-Janitor

Ze Janitorz

The situation :



What we are building :



Transfer your data for a better concurrent offer, is it worth it ?

Algo in three steps :

➔ compare access, transfer and storage prices of each part

➔ calculate the cost of transfer :

$$size_{object} * \underbrace{price_{download}}_{\text{free upload}} + (price_{GET} + price_{PUT}) * \overbrace{\frac{size_{object}}{size_{chunk}}}^{nb_{chunks}}$$

➔ how long would it take to absorb the cost of transfer ?

Transfer your data for a better concurrent offer, is it worth it ?

Region : EU-West	AWS Hot	Azure Hot
Stockage /month and /Gb	0,024$	0,0196$
Download /Gb	0,09$	0,087$
Request PUT (for 10 000)	0,0053$	0,0054$
Request GET (for 10 000)	0,0042$	0,0043$
Other request (for 10 000)	0,0053$	0,0054$

Example : **AWS to Azure**

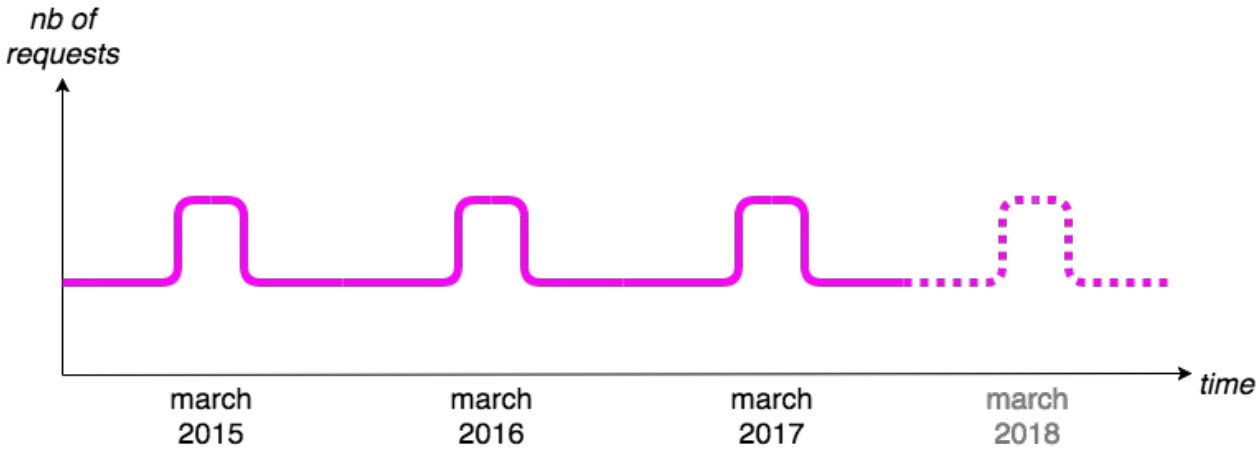
Size : 1 Gb
Chunks number : 100
Cost of transfer : about 9 cents

Benefit per month thanks to lower Azure's stockage price : 0,44 cent/Gb

Rentabilty in 20 months

Transfer your data due to your usage, is it worth it ?

- ➔ **A.I. project based on objects metadatas** : which requests ? How many ?
- ➔ **Anticipate** the trend of your usage and **switch between cool and hot storage** :



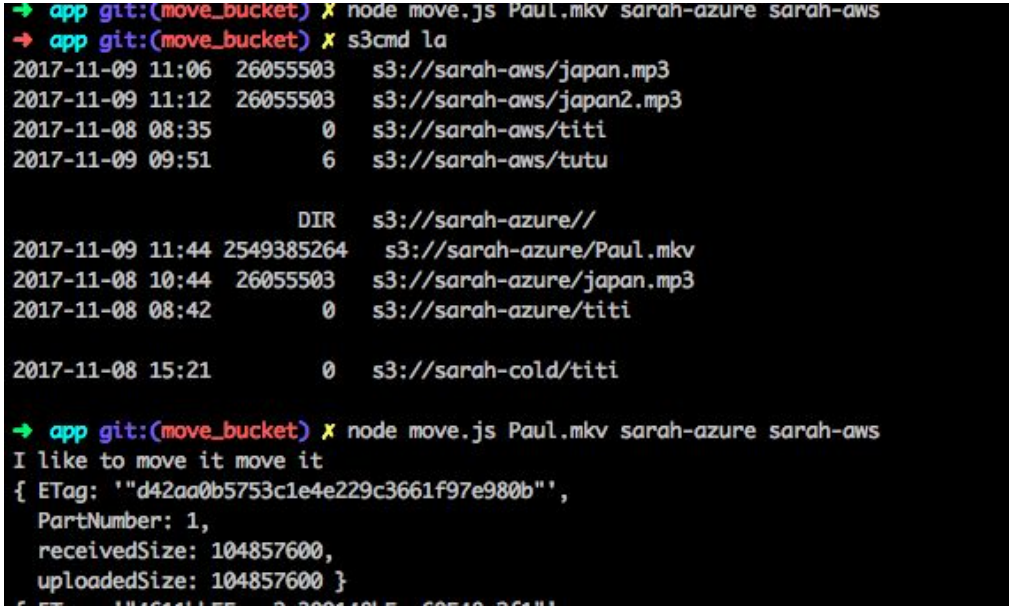
Example :
Go to a hot storage in march, go back to cool for the rest of the year.

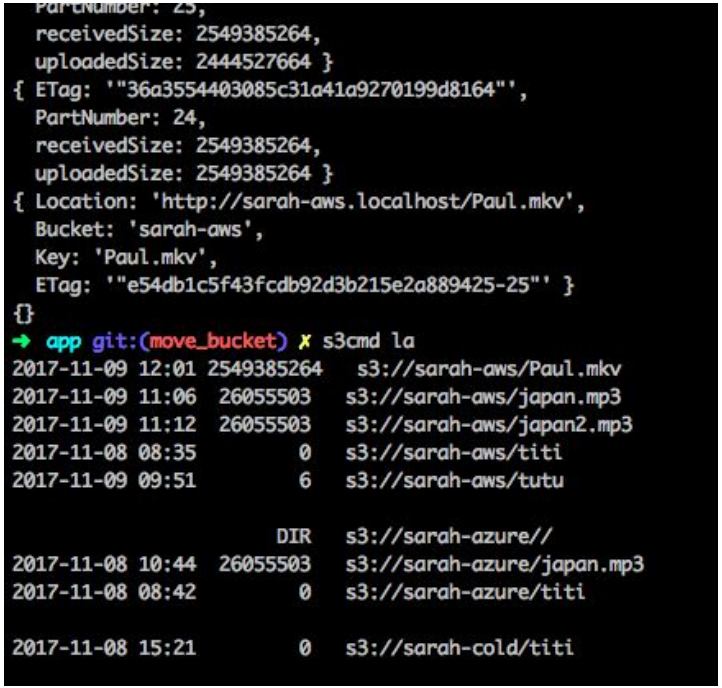
- ➔ **Move** the data to the cheaper cool/hot storage (Azure or AWS) thanks to Zenko CloudServer used by the Data-Janitor

Demo Time : Data Transfer !

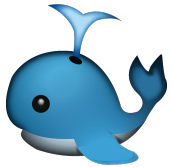
Demo Time : Data Transfer !

Test with a 2.4Gb file:





What we've learned...?



How to use Docker

How to use Zenko

How to navigate through clouds

How to code an app in node.js !
How to put it into a container

How cloud providers make you pay…!

How 'streams' work and how to transfer big files by using chunks of data.

The future...

- ➔ Implement ourself the **library of stream uploads** that we used

- ➔ Implement the **A.I based on the datas usage**.
  Keep in mind the complexity of :
  - getting all the metadata before Zenko's usage
  - mathematic model to anticipate the trend
  - the Azure's and AWS's fees (many particular conditions)

- ➔ Add a **feature in the A.I** : move the datas due to the economical conjectures and the security

Thanks for your attention