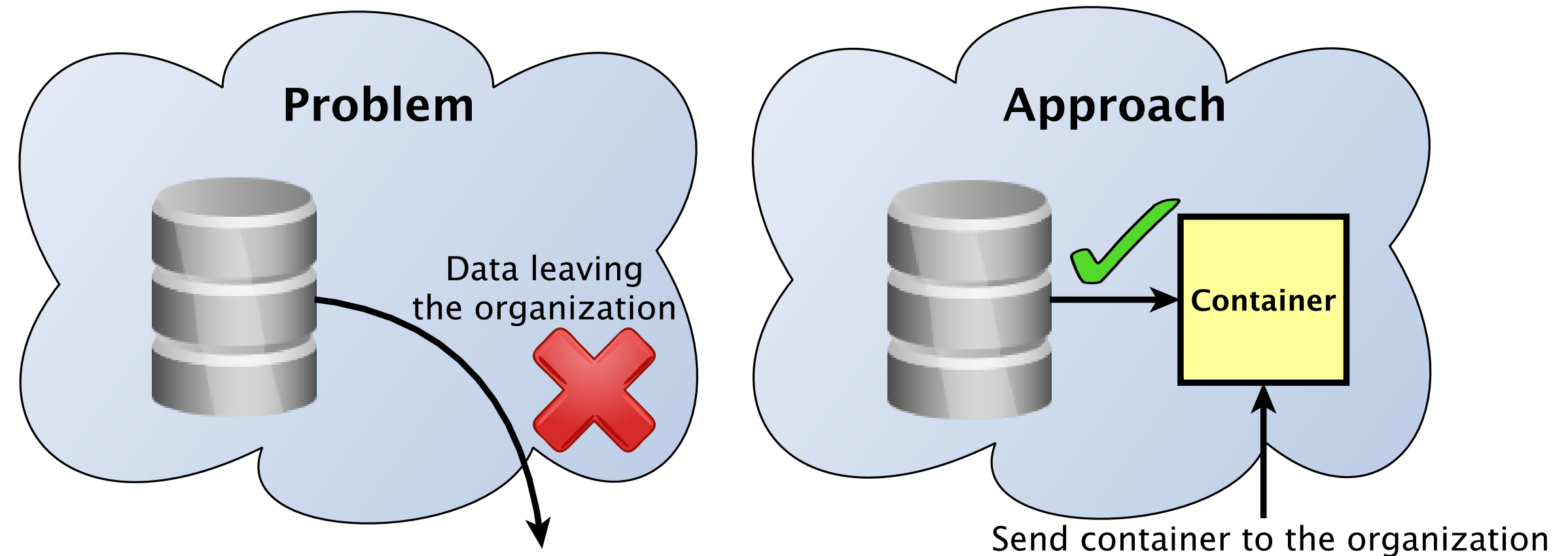


# Container-based remote data processing

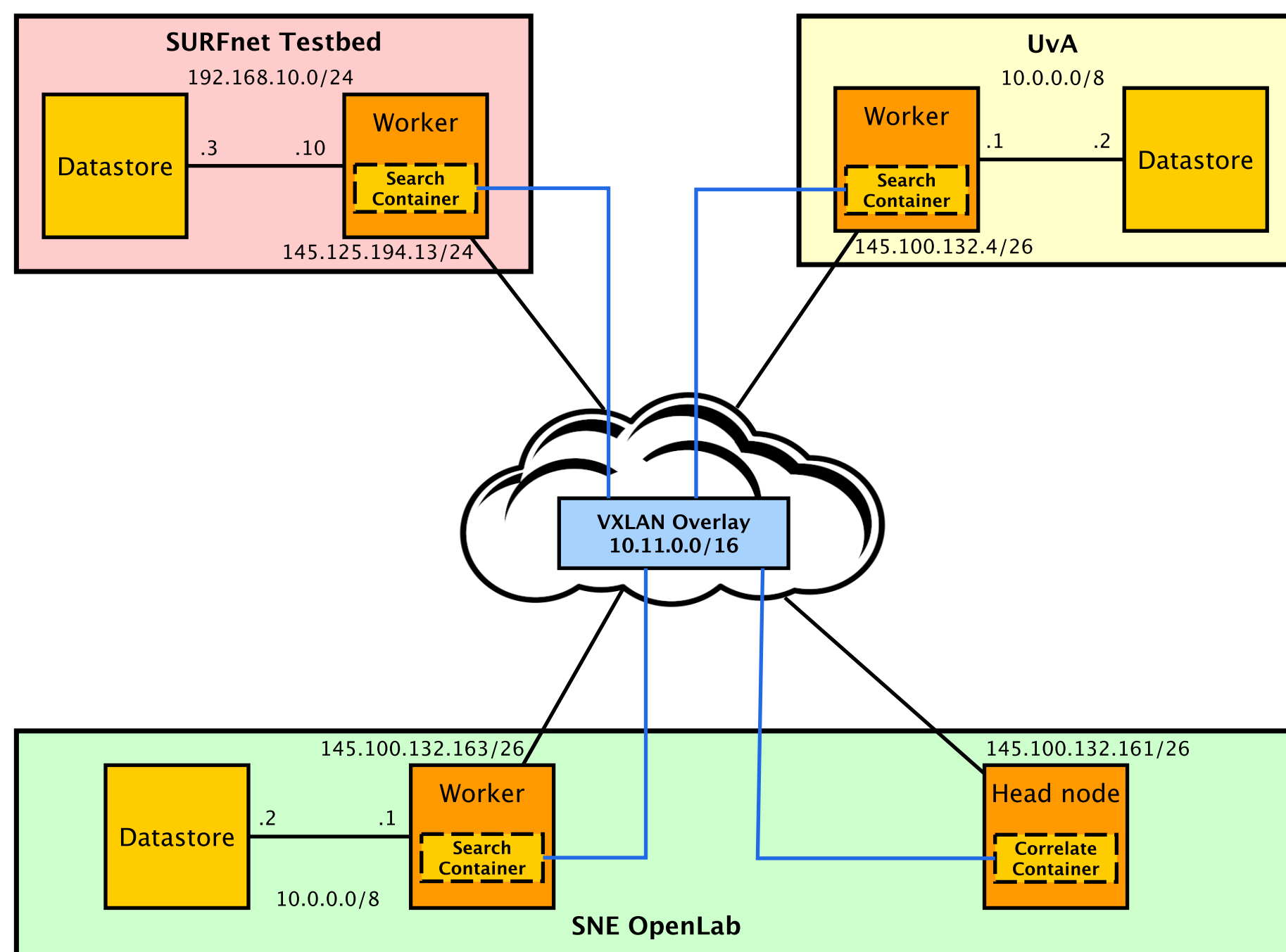


## Problem Description

- Scientific datasets are usually made publicly available  
....but data cannot always leave the organization premises
- On-site data processing can be challenging because of incompatibility of systems or lack of manpower
- Can a container-based system perform remote on-site data processing efficiently?
- What are the networking issues to solve?



## Underlay and Overlay



### Main features:

- Networked containers
- VXLAN overlay
- Containers that perform data retrieval and computation
- Containers built on-demand
- On-site data processing
- Distributed data source
- Multiple sites with datasets

## The Game

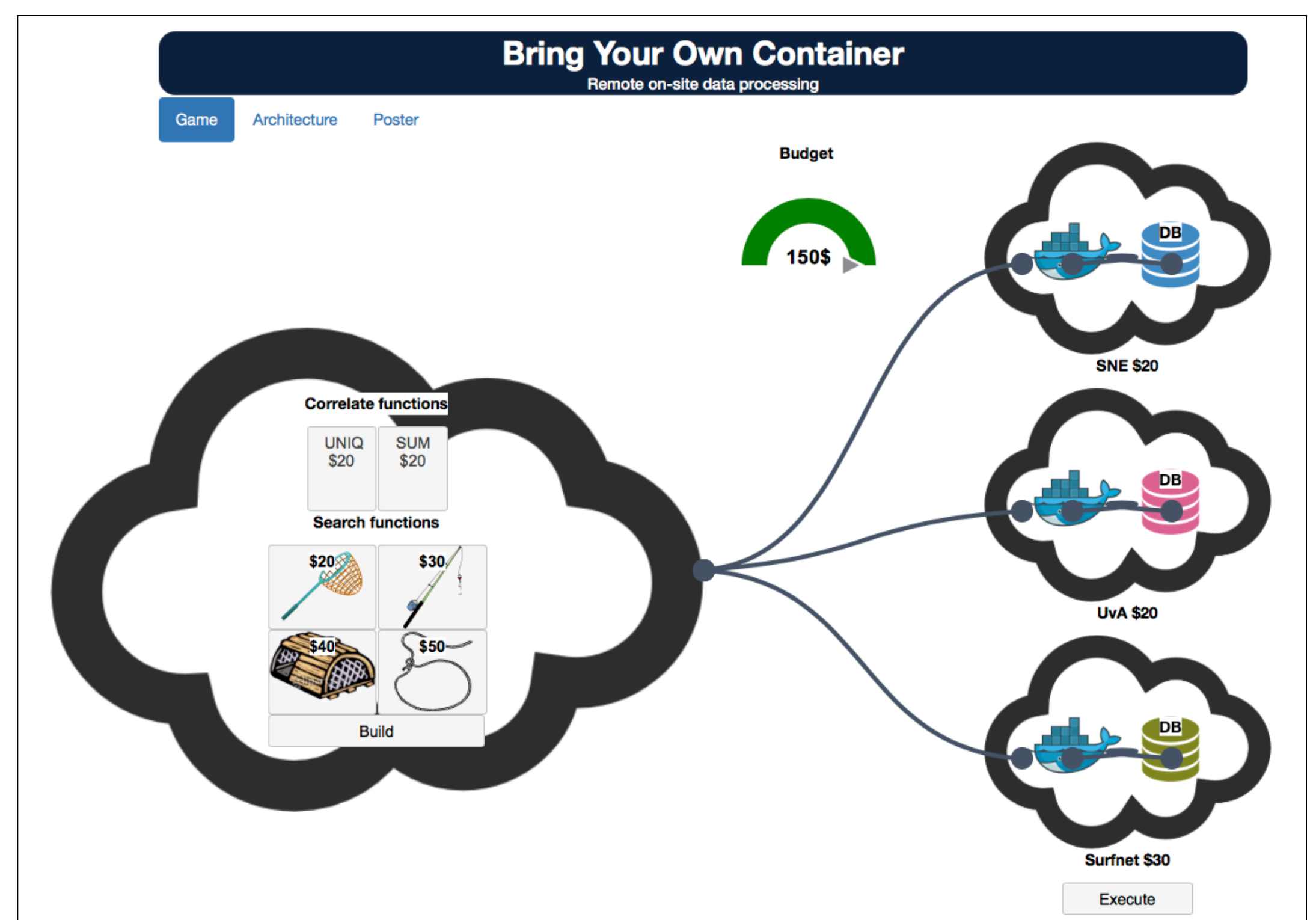
Our SC16 demo is a gamification of the remote dataset processing architecture.

How many different animal species can you find? You have a fixed budget and each function and processing will cost you money!

In our game you will:

- Select a correlate function to combine the results of the different sites.
- Pick different search functions, represented as tools, to find animals in the remote datasets.
- Build containers with the search and correlate functions.
- Execute the containers on the sites of your choice.

Will you have the best score?



### More information:

- <http://byoc.lab.uvalight.net/info>
- <http://sne.science.uva.nl/sne/gigaport3>
- <http://delaat.net/sc>