Predictive Replica Placement for Distributed Data Stores in Fog Environments with Mobile Clients

Master's Thesis Exposé

Malte Mustermann

Context (short)

Fog computing approaches and applications have been developed in recent years. A problem with fog computing is data storage and distribution. One platform that tries to solve this problem is FogStore. This platform allows applications to declaratively describe which data should be stored where.

An application, on a mobile phone, for example, communicates with the FogStore nodes via a middleware on the client. The application declares to the middleware where to store which data, and FogStore then handles data replication and distribution. This allows for the developer of a fog application to focus on the core product.

Problem Statement

However, the application developer still has to decide which data to replicate to which node. In many cases, this decision can be handled by a standard component as often the application prefers the data on the closest node based on latency. Therefore a program can be developed that continuously monitors the latency to the nodes to decide where to replicate that data. This approach however still has some problems. For example, the data transfer from one fog node to another or from the cloud to a fog node can be relatively slow. When for example a large machine learning model has to be transferred to a remote location with poor network connectivity, this could take minutes to hours. While an agent is connected to a node that does not store some needed data the application might have to connect to a further away node with higher latency which might or might not be tolerable, depending on the application.

Idea

If however the application, a middleware, the fog storage framework, or some other component could predict the next locations of an agent and the route to this location, the needed data could be preloaded and would be ready for the agent when the agent connects to the node. In this thesis, we will answer the question of how a prediction of an agent's next location can improve the performance of a fog storage application. We will explore how different algorithms and parameters can improve different metrics. A main problem is to balance

between the cost of the closest node not storing the desired data and the cost of excess data storage.

Approach + Evaluation

In this thesis, we will use existing approaches of next location prediction and develop new algorithms. Next location prediction uses historical location data of one or multiple users to predict the next location and in some cases also the times of arrival and departure. Our approaches will use the data of a single user, not of multiple users, as this could lead to privacy concerns. This allows us to make this prediction in the middleware, as all needed data is available there. We will implement different machine learning approaches for this problem of next place prediction / next node prediction with the characteristics described above. Additionally, we will develop multiple metrics for measuring and comparing our algorithms.

We will evaluate our approaches using a simulation. In our simulation, we will generate nodes in different ways. We then can analyze how the algorithms work with different node densities.