CSE 222A

Traffic Optimization for Hadoop Based on SDN

Siyuan Liu, Jiaxu Zhu, Xinyi Wang

Achievements

Since we have no access to a real computing cluster, we need to build a virtual Hadoop cluster on a single machine. Specifically, we use *Vagrant* to deploy a virtual cluster on one single machine. Virtual nodes booted created by *Vagrant* is almost as real physical nodes except that they share the same physical memory and network interface. Specially, these virtual nodes have independent IP address allocated by the router. Therefore we deploy the Hadoop platform on those virtual nodes. Cloudera is a popular and powerful manager to help us to do this job.

Now we've already have these actual nodes of Hadoop, we start to think about how to build virtual network over these nodes in the next step. In our project, we deploy *Mininet* as the tool to develop SDN networks. Setting up *Mininet* in VirtualBox, we succeeded in establishing *OpenFlow* switches and building a tree topology over virtual hosts. Virtual hosts can communicate with each other according to the topology and OpenFlow setting.

Plans

Implementation

There is a paper published on Parallel and Distributed Processing Symposium, 2014 IEEE 28th International, *Pythia: Faster Big Data in Motion through Predictive Software-Defined Network Optimization at Runtime* (this paper cites the FlowComb paper we mentioned in proposal). The algorithm and system described in this paper are much similar to what we want to achieve at last.

The main idea of *Pythia* is to run a instrument process on each slave server whose role is to monitor the activity of Hadoop and predict future shuffle transfers. The process integrates these prediction information into a message and sends it to the SDN Events

Collector. Then the SDN Flow Allocator analyses the information collected in the Event Collector, and optimize the networking.

SDN Flow Allocator is the key component to update network and implicitly control data flow on the network. We can use shortest path algorithms, like Dijkstra, to change routing graph on the Allocator and then update routing table for each switch in the network.

To sum up, we will utilize SDN control system which consists of Events Collector and Flow Allocator to optimize the network and data flow for Hadoop where each slave is attached a Instrument Process. Combined with tools and systems we found and utilized, we build the Hadoop Cluster system with *Vagrant* and *Cloudera*; we build the virtual network on the Hadoop system with *Mininet* or some OVS; finally we implement an SDN control system with *OpenFlow*.

Evaluation

Based on what is described in the implementation, our main concern is how the strategies in the SDN Control System influence the job completion time for Hadoop cluster.

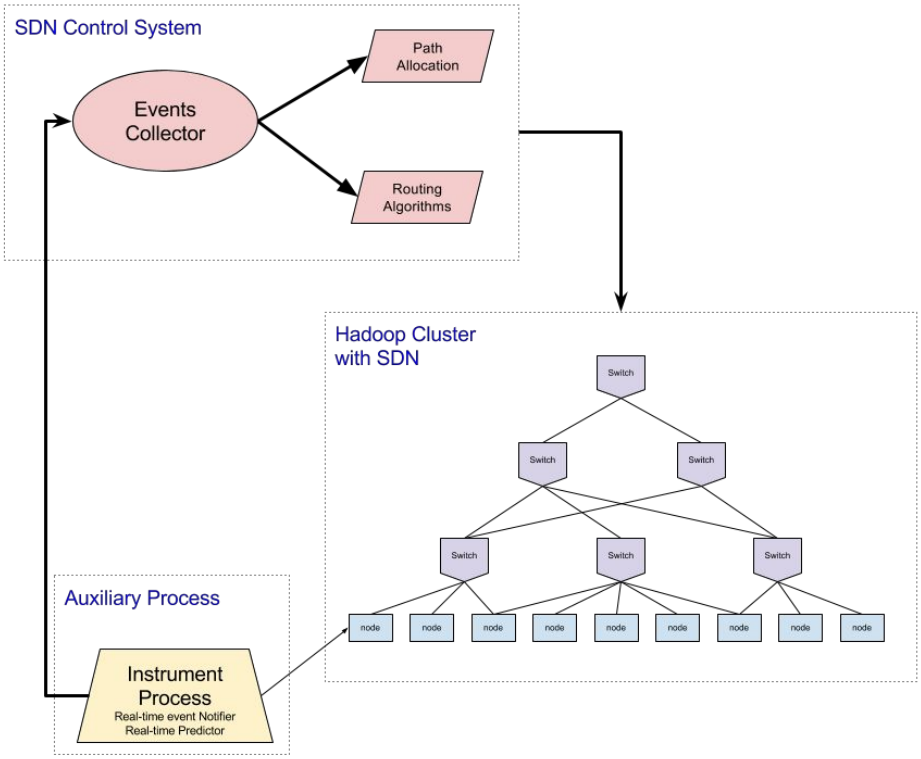
If possible, we also want to get prediction accuracy of Instrument Process. If the accuracy is high, it indicates the prediction is effective. Then, we need to see if the SDN control system takes full use of the prediction information.

During the experiment, we will use some Hadoop benchmarks, such as Sorting and Word Count. With the help of these benchmarks, we can see how the whole system performs with SDN.

Graph

System Graph

Our virtual cluster contains several virtual nodes, including one master (name) node and other slave (data) nodes. These nodes are connected in a multi-level networks, which is built as fat tree like topology. Once a Hadoop job is submitted, it's mapped to data nodes and get reduced on one node, during which there are high network communication demands among these nodes.



With the Hadoop cluster deployed by Vagrant and Cloudera, we get a bunch of nodes (master and slaves). Then we deploy Instrument Process for each node which sends message about event notification and prediction to SDN Control System. SDN control system collect the messages sent from each instrument process, then analyses the information and computes new routing graph for the virtual network which will be re-deployed on switches.

After all these efforts, we try to build a network communication system based on SDN in order to maximize the computation performance of Hadoop.

Aim

First, we try to do experiment to feel the power of SDN and find some optimal strategies of SDN which can lead to a better performance of Hadoop.

Second, we try to build an SDN system independently. The SDN system will not influence how the Hadoop works, which means when a new Hadoop cluster is deployed and jobs are assigned our system is supposed to reduce the job completion time.