Efficient distribution of crypto object and cipher-text in cluster environments

Milestone 3
Project recap

- Apache Spark and Homomorphic encryption
- SparkFHE plugin
- Problems with clustering

Dot-Product of two vectors of encrypted numbers

Test case(s):

\[
\begin{align*}
\text{vec}_a & : (0,1,2,3,4) \\
\text{vec}_b & : (4,3,2,1,0)
\end{align*}
\]

SparkFHE outputs:

Dot product: 10

Source: https://github.com/SpiRITlab/SparkFHE-Examples/wiki
<table>
<thead>
<tr>
<th>Frameworks</th>
<th>Agents</th>
<th>Roles</th>
<th>Offers</th>
<th>Maintenance</th>
<th>HadoopFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>5678c6558d-0000</td>
<td>driver-2020006152250-0011</td>
<td>Driver for sparkhe_dot_product_examples</td>
<td>*</td>
<td>FAILED</td>
</tr>
<tr>
<td>...</td>
<td>5678c6558d-0000</td>
<td>driver-2020006152246-0010</td>
<td>Driver for sparkhe_basic_OPs_examples</td>
<td>*</td>
<td>FINISHED</td>
</tr>
<tr>
<td>...</td>
<td>5678c6558d-0000</td>
<td>driver-2020006152153-0009</td>
<td>Driver for sparkhe_encryption_decryption</td>
<td>*</td>
<td>FINISHED</td>
</tr>
<tr>
<td>...</td>
<td>5678c6558d-0000</td>
<td>driver-2020006152114-0008</td>
<td>Driver for sparkhe_keygen</td>
<td>*</td>
<td>FINISHED</td>
</tr>
<tr>
<td>...</td>
<td>5678c6558d-0000</td>
<td>driver-2020006152037-0007</td>
<td>Driver for TestConnectionToSharedLibrary</td>
<td>*</td>
<td>FINISHED</td>
</tr>
<tr>
<td>...</td>
<td>5678c6558d-0000</td>
<td>driver-2020006154740-0006</td>
<td>Driver for sparkhe_dot_product_examples</td>
<td>*</td>
<td>FAILED</td>
</tr>
<tr>
<td>...</td>
<td>5678c6558d-0000</td>
<td>driver-2020006154736-0005</td>
<td>Driver for sparkhe_basic_OPs_examples</td>
<td>*</td>
<td>FINISHED</td>
</tr>
<tr>
<td>...</td>
<td>5678c6558d-0000</td>
<td>driver-2020006154607-0004</td>
<td>Driver for sparkhe_encryption_decryption</td>
<td>*</td>
<td>FINISHED</td>
</tr>
<tr>
<td>...</td>
<td>5678c6558d-0000</td>
<td>driver-2020006154533-0003</td>
<td>Driver for sparkhe_keygen</td>
<td>*</td>
<td>FINISHED</td>
</tr>
<tr>
<td>...</td>
<td>5678c6558d-0000</td>
<td>driver-2020006154528-0002</td>
<td>Driver for TestConnectionToSharedLibrary</td>
<td>*</td>
<td>FINISHED</td>
</tr>
</tbody>
</table>

Using Apache Mesos
Accomplished Work

Proposed deliverables

- Successful execution of SparkFHE vector operations on a cluster
- Backend development of SparkFHE demo application
Accomplished Work

Actual deliverables

- Root cause analysis of problem with cluster computing
- Automation of multiple steps such as update the cluster code, changes to existing scripts, adding scripts to call the demo application backend
- Backend development of SparkFHE demo application
Results and evaluation

- Successful demo of application
  - Cluster based development
  - HDFS automation
  - API calls
- Evaluation is tricky for vector operations
Demo: SparkFHE backend application
Conclusion

• Homomorphic encryption has the potential to address multiple privacy related issues given efficient usage using cluster computing.

• Cluster computing can have breakdown at multiple ends, knowledge of these steps are important.

• Apache Spark’s scheduling is a very interesting study.

• The capstone project hasn’t been a complete success.
Future work

- Use of dynamic allocation to possibly fix the cluster issue
- Similar backend application to run on Kubernetes
- Using homomorphic encryption for Private set intersection
Thank you