Jason Reader, Ling Xiang

Dr. Davis

CSE 485 Section B

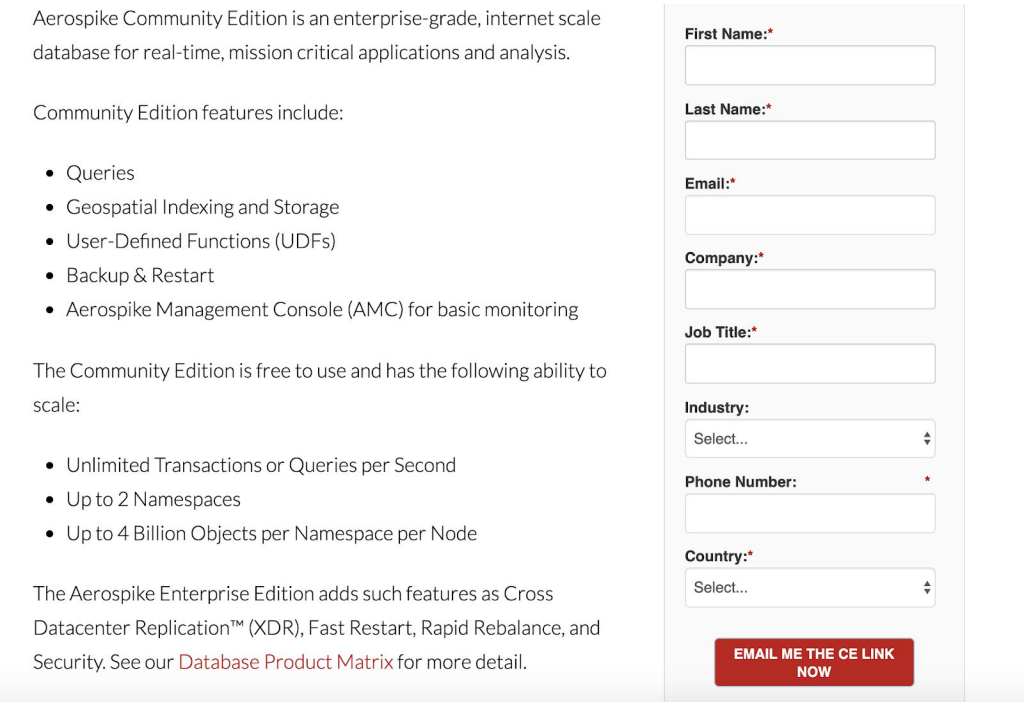
9 April 2019

CSE 485 Final Project

A. Instructions to install and configure the system (with screenshots) for one platform

- Go to https://www.aerospike.com/
- Click on Try Now in the top right corner

- After that, you want to scroll down and fill out the required information and then click the “Email me the CE link now”.



Aerospike Community Edition is an enterprise-grade, internet scale database for real-time, mission critical applications and analysis.

Community Edition features include:

- Queries
- Geospatial Indexing and Storage
- User-Defined Functions (UDFs)
- Backup & Restart
- Aerospike Management Console (AMC) for basic monitoring

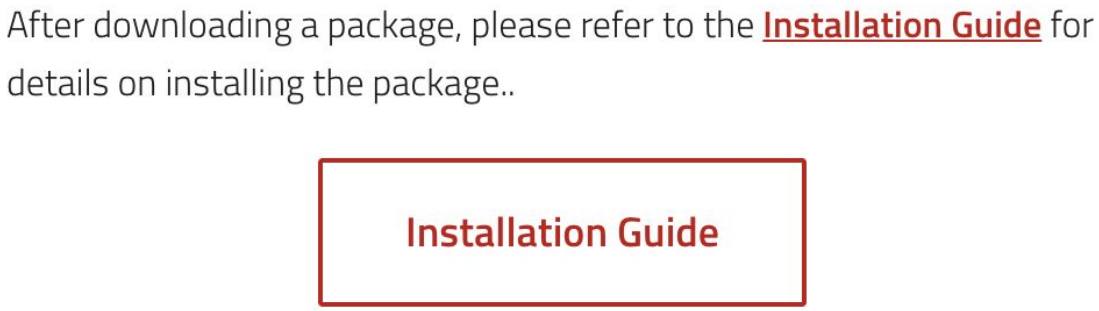
The Community Edition is free to use and has the following ability to scale:

- Unlimited Transactions or Queries per Second
- Up to 2 Namespaces
- Up to 4 Billion Objects per Namespace per Node

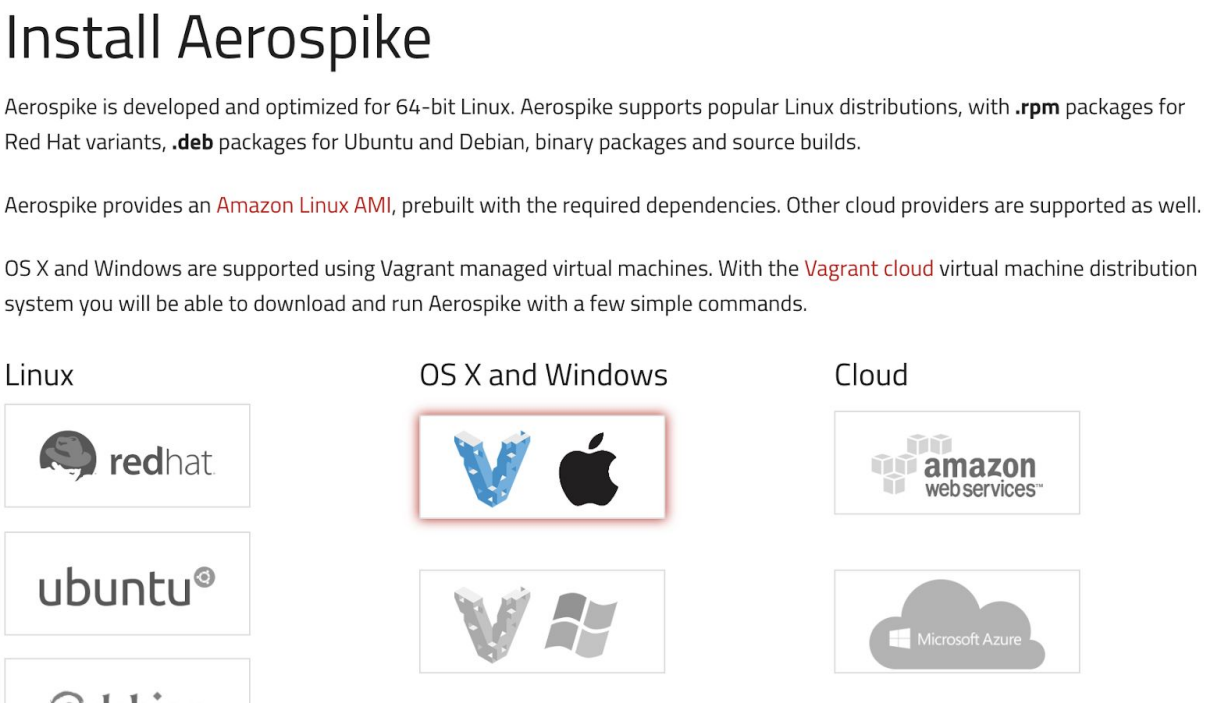
The Aerospike Enterprise Edition adds such features as Cross Datacenter Replication™ (XDR), Fast Restart, Rapid Rebalance, and Security. See our Database Product Matrix for more detail.

- Then, you go to the email you entered and click on the “click here” hyperlink.

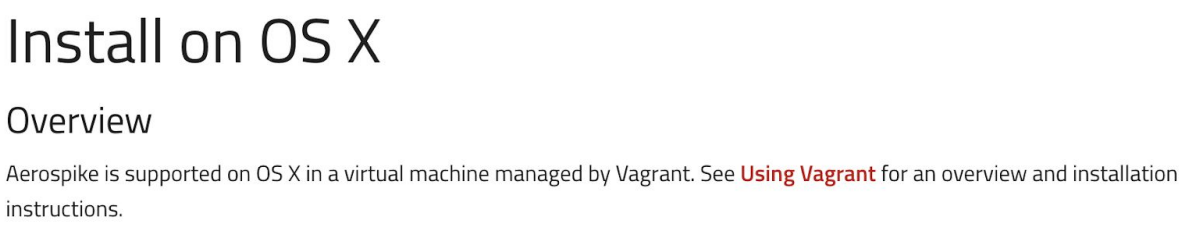
- This link will take you to an installation page where you want to click on the tab that says Installation Guide.



- This will take you to a page where you will select the OS that you are looking to download this software on.



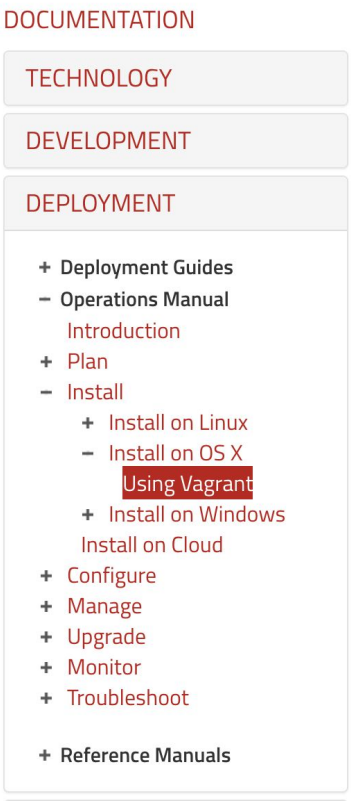
- If you don't have vagrant on your computer, you need to install vagrant first. Click on "Using Vagrant". It will lead you to new a page, follow the instructions on the page.



- You will also need to download VirtualBox, by following this link and clicking on the proper OS.

https://www.oracle.com/technetwork/server-storage/virtualbox/downloads/index.html

- After downloading both Vagrant and VirtualBox, and following the proper instructions on the Using Vagrant page, you will want to click back on the “Install on OS X” (or whatever OS you have.



- After clicking on that page, you want to follow the instructions given.
- Finally, to enter the Aerospike workspace you want to enter the following commands below.

- From there you should be able to begin exploring the Aerospike language!

5. Description of a larger data source to be used in the assignment (source, contents, size)

- Source: https://www.kaggle.com/mehdidag/black-friday/version/1
- Black Friday data set
    - Contents: User_ID, Product_ID, Gender, Age, Occupation, City_Category, Marital_Status, Product_Category_1, Product_Category_2, Product_Category_3, Purchase
    - Size: 5 MB (550,000 documents)

How to populate the database:

- Manually insert the data with the following commands when you are within the aql(aerospike query language) in terminal

create index gender on test.BF (gender) String

create index age on test.BF (age) Numeric

create index marital_Status on test.BF (marital_status) Numeric

create index purchase on test.BF (purchase) Numeric

insert into test.BF (PK, gender, age, marital_status, purchase) values (1, 'F', 17, 0, 8370)

insert into test.BF (PK, gender, age, marital_status, purchase) values (2, 'M', 55, 0, 7969)

insert into test.BF (PK, gender, age, marital_status, purchase) values (3, 'M', 35, 0, 15227)

insert into test.BF (PK, gender, age, marital_status, purchase) values (4, 'M', 50, 1, 19215)

insert into test.BF (PK, gender, age, marital_status, purchase) values (5, 'M', 35, 0, 5254)

insert into test.BF (PK, gender, age, marital_status, purchase) values (6, 'F', 45, 1, 11755)

insert into test.BF (PK, gender, age, marital_status, purchase) values (7, 'M', 45, 1, 11788)

insert into test.BF (PK, gender, age, marital_status, purchase) values (8, 'M', 35, 1, 19614)

insert into test.BF (PK, gender, age, marital_status, purchase) values (9, 'F', 35, 0, 16662)

insert into test.BF (PK, gender, age, marital_status, purchase) values (10, 'F', 45, 1, 16352)

insert into test.BF (PK, gender, age, marital_status, purchase) values (11, 'F', 17, 0, 19172)

insert into test.BF (PK, gender, age, marital_status, purchase) values (12, 'M', 17, 0, 7746)

Data Demands:

1. List the gender, age, marital status, and purchase amount of customers who spent more than $7500 but no more than $20000.
2. List marital status and purchase amount for people who are at least 40 but no older than 60.
3. List all the purchase for all female customer.

Answer:

1. select gender, age, marital_status, purchase from test.BF where purchase BETWEEN 7500 and 20000

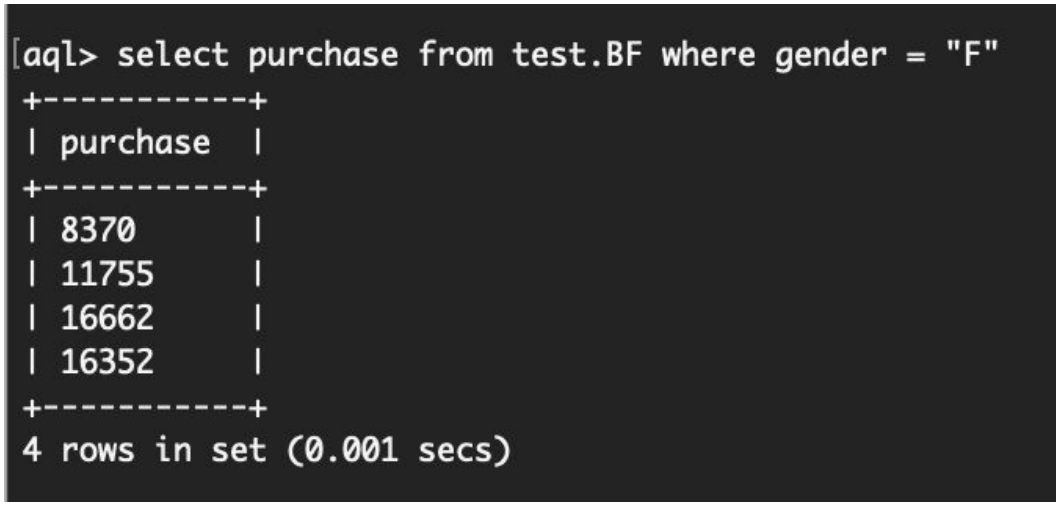
```
aql> select gender, age, marital_status, purchase from test.BF where purchase BETWEEN 7500 and 20000
+---------+------+----------------+-----------+
| gender  | age  | marital_status | purchase  |
+---------+------+----------------+-----------+
| "F"     | 45   | 1              | 16352     |
| "M"     | 55   | 0              | 7969      |
| "M"     | 17   | 0              | 7746      |
| "F"     | 45   | 1              | 11755     |
| "M"     | 45   | 1              | 11788     |
| "M"     | 50   | 1              | 19215     |
| "F"     | 17   | 0              | 8370      |
| "F"     | 35   | 0              | 16662     |
| "M"     | 35   | 0              | 15227     |
| "M"     | 35   | 1              | 19614     |
+---------+------+----------------+-----------+
10 rows in set (0.009 secs)

OK
```

2. select marital_status, purchase from test.BF where age between 40 and 60

```
aql> select marital_status, purchase from test.BF where age between 40 and 60
+----------------+-----------+
| marital_status | purchase  |
+----------------+-----------+
| 1              | 11755     |
| 1              | 11788     |
| 1              | 16352     |
| 1              | 19215     |
| 0              | 7969      |
+----------------+-----------+
5 rows in set (0.010 secs)
```

3. select purchase from test.BF where gender = "F"



Time Log:

Jason Reader	Ling Xiang							
Date	Start Time	End Time	Location	Duration	Topic of Work Session	Action Item	Total Time	Who
4/8	20:00	21:40	Benton	1:40	During the session, we completed the necessary work to create the install instructions for our database and to describe the data source that is to be used during our final project.	Begin to write queries and work within our database to complete the necessary tasks for the final project.	1:40	Both
4/15	20:00	21:20	Benton	1:20	Work on populating database and providing instructions for classmates. Also begin working on presentation	Ask questions concerning populating database and continue to work on presentation and documentation	3:00	Both
4/17	16:00	18:00	Benton	2:00	Created a set named BF in test namespace. Uploaded some data into test.BF. Wrote a filtering query about it.	Asking questions about aggregate queries. Making powerpoint about everything.	5:00	Both
4/18	15:05	16:10	Benton	1:05	Finished some powerpoint slides and asked questions about aggregation	Come up with one more complex query and finishing up powerpoint.	6:05	Both
4/19	13:00	13:30	Home	0:30	Finished up some slides and added pictures for visualization. Came up with a data demand.	Polish powerpoint slides and finish up data demands.	6:35	Jason
4/19	16:10	17:00	King Library	0:50	Finished data demand, final check the documents and slides.		7:25	Ling