


Module Code	School of Computing University of Leeds	
COMP5850M	Coursework 1 - Report	 UNIVERSITY OF LEEDS

Full Name: Fanhui Meng
Coursework Title: VIM

Username: sc19fm
Deadline Date: 13/03/2020

Part 1: Java OpenNebula Cloud API (OCA) (10 marks)

Provide an explanation of the implementation of this task. The inclusion of the entire code is not required but you may include snippets if you wish.

VM template (1 mark)

```
Virtual Machine Template:
CPU="0.1"
SCHED_DS_REQUIREMENTS="ID = 104"
NIC=[
    NETWORK_UNAME="oneadmin",
    NETWORK="vnet1" ]
LOGO="images/logos/debian.png"
DESCRIPTION="A ttylinux instance with VNC and network context scripts, available
for testing purposes. In raw format."
DISK=[
    IMAGE_UNAME="oneadmin",
    IMAGE="ttylinux Base" ]
SUNSTONE_NETWORK_SELECT="YES"
SUNSTONE_CAPACITY_SELECT="YES"
MEMORY="128"
HYPERVISOR="kvm"
GRAPHICS=[
    LISTEN="0.0.0.0",
    TYPE="vnc" ]
```

Information OpenNebula provides about the VM (1 mark)

This is the information OpenNebula stores for the new VM:

```
<VM><ID>36431</ID><UID>366</UID><GID>1</GID><UNAME>sc19fm</UNAME><GNAME>users</GNAME><NAME>one-36431</NAME><PERMISSIONS><OW
NER_U>1</OWNER_U><OWNER_M>1</OWNER_M><OWNER_A>0</OWNER_A><GROUP_U>0</GROUP_U><GROUP_M>0</GROUP_M><GROUP_A>0</GROUP_A><OTHER
_U>0</OTHER_U><OTHER_M>0</OTHER_M><OTHER_A>0</OTHER_A><PERMISSIONS><LAST_POLL>0</LAST_POLL><STATE>3</STATE><LCM_STATE>3</LCM_STATE><PREV_STATE>3</PREV_STATE><PREV_LCM_STATE>3</PREV_LCM_STATE><RESCHED>0</RESCHED><STIME>1583599711</STIME><ETIME>0</
ETIME><DEPLOY_ID>one-36431</DEPLOY_ID><MEMORY>0</MEMORY><CPU>0</CPU><NET_TX>0</NET_TX><NET_RX>0</NET_RX><TEMPLATE><AUTOMAT
IC_REQUIREMENTS><![CDATA[CLUSTER_ID = 100 & !(PUBLIC_CLOUD = YES)]]></AUTOMATIC_REQUIREMENTS><CPU><![CDATA[0.1]]></CPU><DIS
K><CLONE><![CDATA[YES]]></CLONE><CLONE_TARGET><![CDATA[SYSTEM]]></CLONE_TARGET><CLUSTER_ID><![CDATA[100]]></CLUSTER_ID><DATA
STORE><![CDATA[default]]></DATASTORE><DATASTORE_ID><![CDATA[1]]></DATASTORE_ID><DEV_PREFIX><![CDATA[hd]]></DEV_PREFIX><DIS
K_ID><![CDATA[0]]></DISK_ID><DRIVER><![CDATA[raw]]></DRIVER><IMAGE><![CDATA[ttylinux Base]]></IMAGE><IMAGE_ID><![CDATA[6]]>
</IMAGE_ID><IMAGE_UNAME><![CDATA[oneadmin]]></IMAGE_UNAME><LN_TARGET><![CDATA[NONE]]></LN_TARGET><READONLY><![CDATA[NO]]></
READONLY><SAVE><![CDATA[NO]]></SAVE><SIZE><![CDATA[40]]></SIZE><SOURCE><![CDATA[/var/lib/one/datastores/1/dadbd6245eeb9a56
7ac47195da41831a]]></SOURCE><TARGET><![CDATA[hda]]></TARGET><TM_MAD><![CDATA[shared]]></TM_MAD><TYPE><![CDATA[FILE]]></TYPE
></DISK><GRAPHICS><LISTEN><![CDATA[0.0.0.0]]></LISTEN><PORT><![CDATA[42331]]></PORT><TYPE><![CDATA[vnc]]></TYPE></GRAPHICS>
<MEMORY><![CDATA[128]]></MEMORY><NIC><AR_ID><![CDATA[0]]></AR_ID><BRIDGE><![CDATA[br0]]></BRIDGE><CLUSTER_ID><![CDATA[100]]
></CLUSTER_ID><IP><![CDATA[10.1.7.77]]></IP><MAC><![CDATA[02:00:0a:01:07:4d]]></MAC><NETWORK><![CDATA[vnet1]]></NETWORK><NE
TWORK_ID><![CDATA[60]]></NETWORK_ID><NETWORK_UNAME><![CDATA[oneadmin]]></NETWORK_UNAME><NIC_ID><![CDATA[0]]></NIC_ID><PHYDE
V><![CDATA[bond0]]></PHYDEV><SECURITY_GROUPS><![CDATA[0]]></SECURITY_GROUPS><VLAN><![CDATA[YES]]></VLAN><VLAN_ID><![CDATA[2
]]></VLAN_ID><NIC><SECURITY_GROUP_RULE><PROTOCOL><![CDATA[ALL]]></PROTOCOL><RULE_TYPE><![CDATA[OUTBOUND]]></RULE_TYPE><SEC
URITY_GROUP_ID><![CDATA[0]]></SECURITY_GROUP_ID><SECURITY_GROUP_NAME><![CDATA[default]]></SECURITY_GROUP_NAME><SECURITY_GR
oup_RULE><SECURITY_GROUP_RULE><PROTOCOL><![CDATA[ALL]]></PROTOCOL><RULE_TYPE><![CDATA[INBOUND]]></RULE_TYPE><SECURITY_GROUP
_ID><![CDATA[0]]></SECURITY_GROUP_ID><SECURITY_GROUP_NAME><![CDATA[default]]></SECURITY_GROUP_NAME><SECURITY_GROUP_RULE><V
MID><![CDATA[36431]]></VMID></TEMPLATE><USER_TEMPLATE><DESCRIPTION><![CDATA[A ttylinux instance with VNC and network contex
t scripts, available for testing purposes. In raw format.]]></DESCRIPTION><HYPERVISOR><![CDATA[kvm]]></HYPERVISOR><LOGO><![
CDATA[images/logos/debian.png]]></LOGO><SCHED_DS_REQUIREMENTS><![CDATA[ID = 104]]></SCHED_DS_REQUIREMENTS><SUNSTONE_CAPACIT
Y_SELECT><![CDATA[YES]]></SUNSTONE_CAPACITY_SELECT><SUNSTONE_NETWORK_SELECT><![CDATA[YES]]></SUNSTONE_NETWORK_SELECT></USER
_TEMPLATE><HISTORY_RECORDS><HISTORY><OID>36431</OID><SEQ>0</SEQ><HOSTNAME>cscloudin6.cloud.comp.leeds.ac.uk</HOSTNAME><HID>
9</HID><CID>100</CID><STIME>1583599718</STIME><ETIME>0</ETIME><VMMAD>kvm</VMMAD><VNMAD>dummy</VNMAD><TMMAD>shared</TMMAD
><DS_LOCATION>/var/lib/one/datastores</DS_LOCATION><DS_ID>104</DS_ID><PSTIME>1583599718</PSTIME><PETIME>1583599719</PETIM
E><RSTIME>1583599719</RSTIME><RETIME>0</RETIME><ESTIME>0</ESTIME><EETIME>0</EETIME><REASON>0</REASON><ACTION>0</ACTION></HI
STORY></HISTORY_RECORDS></VM>
```

Run No.	VM instantiation time	VM deletion time
1	8934 ms	157 ms
2	8486 ms	128 ms
3	9998 ms	126 ms
4	11203 ms	135 ms
5	8111 ms	167 ms
Average	9346.4 ms	142.6 ms
Standard Deviation	1123.22	16.43

Use `currentTimeMillis()` method to get the current system time (`startTime`) before instantiation a VM. Then I write a loop, keep checking the VM's status. Until the VM is up running, break the loop. Then use the same method to get the time (`endTime`) after instantiation. And the instantiation time is the `endTime` minus the `startTime`. Then use the same way to get the VM deletion time. Set `startTimeDelete` before delete. Set `endTimeDelete` after `finalizeVM()` method and make sure there is nothing wrong. Finally, use these two numbers to calculate the deletion time.

[illegible]

Part 2: VM Migration (15 marks)

Provide an explanation of the implementation of this task. The inclusion of the entire code is not required but you may include snippets if you wish.

Requirements (2 marks)

Migrate existing VM to a better host. Considering the host's load balance, CPU usage, memory usage and disk usage. Ideally, the host with lower load balance, CPU usage, etc. should be better.

Solution Design (2 marks)

Check all of the hosts' information in the host pool. Assign each factor with a weight. As far as I concerned, the number of VM shouldn't be the main factor. Because some of the VMs require more CPU and memory, and some of them are not. The more important consideration should be the CPU, memory and disk usage. So I assign 0.5 to the number of VM, assign 1 to the other three factors.

As for these usages, I use the maximum usage divided by the current usage, then multiply 100 to represent these usages.

Finally, I add up the four parameters with each weight for each host. Then the host with the lowest number is precisely the host I want the VM to migrate.

Implementation (2 marks)

Get the host id, CPU usage, memory usage and disk usage from the host pool. Then put each host with its properties into an array. Sum up these numbers with each weight. Then compare with each other, find the lowest and it's host id. Live migrate the VM to the target host. Wait until the VM is up running again. Finally, check the target host information, make sure the VM is already in the target host.

Measure the time it takes to migrate the VM. (2 marks)

Run No.	VM migration time
1	15 ms
2	24 ms
3	21 ms
4	19 ms
5	16 ms
Average	19 ms
Standard Deviation	3.29

Evidence of successful run, e.g. screenshot (3 marks)

```
170.9930616450781
169.18557389709287
190.60436548827036
186.8285870777604
186.8285870777604
186.8285870777604
186.8285870777604
186.8285870777604
186.8285870777604
170.9930616450781      0
169.18557389709287    1
I'm gonna migrate VM to 7
```



```
Start to migrate vm...
migrate ok.
Wait till it run again
Time for migration 6141ms
ok, check host info
<HOST><ID>7</ID><NAME>csccloudln8.cloud.comp.leeds.ac.uk</NAME><STATE>2</STATE><I
M MAD><![CDATA[kvm]]></IM MAD><VM MAD><![CDATA[kvm]]></VM MAD><VN MAD><![CDATA[d
ummy]]></VN MAD><LAST MON TIME>1584025840</LAST MON TIME><CLUSTER ID>100</CLUSTE
R ID><CLUSTER>csccloud1</CLUSTER><HOST SHARE><DISK USAGE>0</DISK USAGE><MEM USAGE
>45219840</MEM USAGE><CPU USAGE>3440</CPU USAGE><MAX DISK>16513820</MAX DISK><MA
X MEM>127541472</MAX MEM><MAX CPU>3000</MAX CPU><FREE DISK>4367891</FREE DISK><F
REE MEM>0</FREE MEM><FREE CPU>3155</FREE CPU><USED DISK>11307068</USED DISK><USE
D MEM>0</USED MEM><USED CPU>44</USED CPU><RUNNING VMS>40</RUNNING VMS><DATASTORE
S><DS><FREE MB><![CDATA[4367891]]></FREE MB><ID><![CDATA[105]]></ID><TOTAL MB><![
CDATA[16513820]]></TOTAL MB><USED MB><![CDATA[11307068]]></USED MB><DS><DATAS
TORES><HOST SHARE><VMS><ID>19106</ID><ID>19230</ID><ID>19372</ID><ID>19373</ID><
ID>19384</ID><ID>19419</ID><ID>19441</ID><ID>22004</ID><ID>22005</ID><ID>22006<
/ID><ID>22007</ID><ID>22008</ID><ID>22009</ID><ID>22010</ID><ID>22011</ID><ID>28
179</ID><ID>29767</ID><ID>30406</ID><ID>31095</ID><ID>31314</ID><ID>31318</ID><I
D>31362</ID><ID>31363</ID><ID>31364</ID><ID>31616</ID><ID>31644</ID><ID>31697</I
D><ID>31734</ID><ID>31764</ID><ID>32004</ID><ID>32041</ID><ID>32633</ID><ID>3495
5</ID><ID>35265</ID><ID>35450</ID><ID>37143</ID><ID>37388</ID><ID>37402</ID><ID>
38141</ID><ID>39276</ID><VMS><TEMPLATE><ARCH><![CDATA[x86_64]]></ARCH><CPUSPEED
><![CDATA[1496]]></CPUSPEED><HOSTNAME><![CDATA[csccloudln8.cloud.comp.leeds.ac.uk
]]></HOSTNAME><HYPERVISOR><![CDATA[kvm]]></HYPERVISOR><MODELNAME><![CDATA[Intel(
R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz]]></MODELNAME><NETRX><![CDATA[86179710430]]>
</NETRX><NETTX><![CDATA[183952953316]]></NETTX><RESERVED_CPU><![CDATA[]]></RESER
VED_CPU><RESERVED_MEM><![CDATA[]]></RESERVED_MEM><VERSION><![CDATA[4.12.1]]></VE
RSION><TEMPLATE><HOST>Trying to hold the new VM... ok.
```

```
This is the information OpenNebula stores for the new VM:
<VM><ID>39276</ID><UID>366</UID><GID>1</GID><UNAME>sc19fm</UNAME><GNAME>users</G
NAME><NAME>one-39276</NAME><PERMISSIONS><OWNER_U>1</OWNER_U><OWNER_M>1</OWNER_M>
<OWNER_A>0</OWNER_A><GROUP_U>0</GROUP_U><GROUP_M>0</GROUP_M><GROUP_A>0</GROUP_A>
<OTHER_U>0</OTHER_U><OTHER_M>0</OTHER_M><OTHER_A>0</OTHER_A></PERMISSIONS><LAST
POLL>0</LAST_POLL><STATE>3</STATE><LCM_STATE>3</LCM_STATE><PREV_STATE>3</PREV_ST
ATE><PREV_LCM_STATE>3</PREV_LCM_STATE><RESCHED>0</RESCHED><STIME>1584025795</STI
ME><ETIME>0</ETIME><DEPLOY_ID>one-39276</DEPLOY_ID><MEMORY>0</MEMORY><CPU>0</CPU
><NET_TX>0</NET_TX><NET_RX>0</NET_RX><TEMPLATE><AUTOMATIC_REQUIREMENTS><![CDATA[
CLUSTER_ID = 100 & !(PUBLIC_CLOUD = YES)]]></AUTOMATIC_REQUIREMENTS><CPU><![CDAT
A[0.1]]></CPU><DISK><CLONE><![CDATA[YES]]></CLONE><CLONE_TARGET><![CDATA[SYSTEM]
]></CLONE_TARGET><CLUSTER_ID><![CDATA[100]]></CLUSTER_ID><DATASTORE><![CDATA[def
ault]]></DATASTORE><DATASTORE_ID><![CDATA[1]]></DATASTORE_ID><DEV_PREFIX><![CDAT
A[hdi]]></DEV_PREFIX><DISK_ID><![CDATA[0]]></DISK_ID><DRIVER><![CDATA[raw]]></DRI
VER><IMAGE><![CDATA[ttylinux Base]]></IMAGE><IMAGE_ID><![CDATA[6]]></IMAGE_ID><I
MAGE_UNAME><![CDATA[oneadmin]]></IMAGE_UNAME><LN_TARGET><![CDATA[NONE]]></LN_TAR
GET><READONLY><![CDATA[NO]]></READONLY><SAVE><![CDATA[NO]]></SAVE><SIZE><![CDATA
[40]]></SIZE><SOURCE><![CDATA[/var/lib/one//datastores/1/dadbd6245eeb9a567ac4719
```

```
<VM><ID>39276</ID><UID>366</UID><GID>1</GID><UNAME>sc19fm</UNAME><GNAME>users</G
NAME><NAME>one-39276</NAME><PERMISSIONS><OWNER_U>1</OWNER_U><OWNER_M>1</OWNER_M>
<OWNER_A>0</OWNER_A><GROUP_U>0</GROUP_U><GROUP_M>0</GROUP_M><GROUP_A>0</GROUP_A>
<OTHER_U>0</OTHER_U><OTHER_M>0</OTHER_M><OTHER_A>0</OTHER_A></PERMISSIONS><LAST
POLL>0</LAST_POLL><STATE>3</STATE><LCM_STATE>3</LCM_STATE><PREV_STATE>3</PREV_ST
ATE><PREV_LCM_STATE>3</PREV_LCM_STATE><RESCHED>0</RESCHED><STIME>1584025795</STI
ME><ETIME>0</ETIME><DEPLOY_ID>one-39276</DEPLOY_ID><MEMORY>0</MEMORY><CPU>0</CPU
><NET_TX>0</NET_TX><NET_RX>0</NET_RX><TEMPLATE><AUTOMATIC_REQUIREMENTS><![CDATA[
CLUSTER_ID = 100 & !(PUBLIC_CLOUD = YES)]]></AUTOMATIC_REQUIREMENTS><CPU><![CDAT
A[0.1]]></CPU><DISK><CLONE><![CDATA[YES]]></CLONE><CLONE_TARGET><![CDATA[SYSTEM]
]></CLONE_TARGET><CLUSTER_ID><![CDATA[100]]></CLUSTER_ID><DATASTORE><![CDATA[def
ault]]></DATASTORE><DATASTORE_ID><![CDATA[1]]></DATASTORE_ID><DEV_PREFIX><![CDAT
A[hdi]]></DEV_PREFIX><DISK_ID><![CDATA[0]]></DISK_ID><DRIVER><![CDATA[raw]]></DRI
VER><IMAGE><![CDATA[ttylinux Base]]></IMAGE><IMAGE_ID><![CDATA[6]]></IMAGE_ID><I
MAGE_UNAME><![CDATA[oneadmin]]></IMAGE_UNAME><LN_TARGET><![CDATA[NONE]]></LN_TAR
GET><READONLY><![CDATA[NO]]></READONLY><SAVE><![CDATA[NO]]></SAVE><SIZE><![CDATA
[40]]></SIZE><SOURCE><![CDATA[/var/lib/one//datastores/1/dadbd6245eeb9a567ac4719
5da41831a]]></SOURCE><TARGET><![CDATA[hda]]></TARGET><TM_MAD><![CDATA[shared]]><
/TM_MAD><TYPE><![CDATA[FILE]]></TYPE></DISK><GRAPHICS><LISTEN><![CDATA[0.0.0.0]]
></LISTEN><PORT><![CDATA[45176]]></PORT><TYPE><![CDATA[vnc]]></TYPE></GRAPHICS>
MEMORY><![CDATA[128]]></MEMORY><NIC><AR_ID><![CDATA[0]]></AR_ID><BRIDGE><![CDATA
[br0]]></BRIDGE><CLUSTER_ID><![CDATA[100]]></CLUSTER_ID><IP><![CDATA[10.1.2.154]
]></IP><MAC><![CDATA[02:00:0a:01:02:9a]]></MAC><NETWORK><![CDATA[vnet1]]></NETWO
RK><NETWORK_ID><![CDATA[60]]></NETWORK_ID><NETWORK_UNAME><![CDATA[oneadmin]]></N
ETWORK_UNAME><NIC_ID><![CDATA[0]]></NIC_ID><PHYDEV><![CDATA[bond0]]></PHYDEV><SE
CURITY_GROUPS><![CDATA[0]]></SECURITY_GROUPS><VLAN><![CDATA[YES]]></VLAN><VLAN_I
D><![CDATA[2]]></VLAN_ID></NIC><SECURITY_GROUP_RULE><PROTOCOL><![CDATA[ALL]]></P
ROTOCOL><RULE_TYPE><![CDATA[OUTBOUND]]></RULE_TYPE><SECURITY_GROUP_ID><![CDATA[0
]]></SECURITY_GROUP_ID><SECURITY_GROUP_NAME><![CDATA[default]]></SECURITY_GROUP
_NAME></SECURITY_GROUP_RULE><SECURITY_GROUP_RULE><PROTOCOL><![CDATA[ALL]]></PROTO
COL><RULE_TYPE><![CDATA[INBOUND]]></RULE_TYPE><SECURITY_GROUP_ID><![CDATA[0]]></
SECURITY_GROUP_ID><SECURITY_GROUP_NAME><![CDATA[default]]></SECURITY_GROUP_NAME>
</SECURITY_GROUP_RULE><VMID><![CDATA[39276]]></VMID></TEMPLATE><USER_TEMPLATE><D
ESCRIPTION><![CDATA[A ttylinux instance with VNC and network context scripts, av
ailable for testing purposes. In raw format.]]></DESCRIPTION><HYPERVISOR><![CDAT
A[kvm]]></HYPERVISOR><LOGO><![CDATA[images/logos/debian.png]]></LOGO><SCHED_DS_R
EQUIREMENTS><![CDATA[ID = 104]]></SCHED_DS_REQUIREMENTS><SUNSTONE_CAPACITY_SELEC
T><![CDATA[YES]]></SUNSTONE_CAPACITY_SELECT><SUNSTONE_NETWORK_SELECT><![CDATA[YE
S]]></SUNSTONE_NETWORK_SELECT></USER_TEMPLATE><HISTORY_RECORDS><HISTORY><OID>392
76</OID><SEQ>0</SEQ><HOSTNAME>csccloudln1.cloud.comp.leeds.ac.uk</HOSTNAME><HID>2
0</HID><CID>100</CID><STIME>1584025838</STIME><ETIME>1584025855</ETIME><VMMAD>k
vm</VMMAD><VNMAD>dummy</VNMAD><TMMAD>shared</TMMAD><DS_LOCATION>/var/lib/one/
/datastores</DS_LOCATION><DS_ID>104</DS_ID><PSTIME>1584025838</PSTIME><PETIME>15
84025839</PETIME><RSTIME>1584025839</RSTIME><RETIME>1584025855</RETIME><ESTIME>0
</ESTIME><EETIME>0</EETIME><REASON>2</REASON><ACTION>2</ACTION></HISTORY><HISTOR
Y><OID>39276</OID><SEQ>1</SEQ><HOSTNAME>csccloudln8.cloud.comp.leeds.ac.uk</HOSTN
AME><HID>7</HID><CID>100</CID><STIME>1584025852</STIME><ETIME>0</ETIME><VMMAD>k
vm</VMMAD><VNMAD>dummy</VNMAD><TMMAD>shared</TMMAD><DS_LOCATION>/var/lib/one/
/datastores</DS_LOCATION><DS_ID>104</DS_ID><PSTIME>0</PSTIME><PETIME>0</PETIME><
RSTIME>1584025855</RSTIME><RETIME>0</RETIME><ESTIME>0</ESTIME><EETIME>0</EETIME>
<REASON>0</REASON><ACTION>0</ACTION></HISTORY></HISTORY_RECORDS></VM>
```

The new VM one-39276 has status: runn

Discussion of the results (4 marks)

The host with more number of VM doesn't mean its CPU and memory are occupied more. To consider if a host is suitable for migration, we should focus on its CPU, memory and disk usage. More basically, we should see the host's status, whether it's on or off, and its allocated CPU, memory is available or not. And if the hosts are in the same cluster, which means within the same LAN, it should take a shorter time to migrate. I use live migrate the VM to a new host, which means transfer running VM between hosts without disconnecting the client or application. And this way only take a short time to migrate a VM.

Part 3: Resource Scaling and Performance/Energy Consumption Trade-Off (10 - 25 marks, depending on application and challenge)

Details of the application considered (stress, MPI, Hadoop, other) (1-3 marks)

Use Hadoop data-parallel processing job, run MapReduce application on up to 4 Vms and four hosts.

Design of the experiments (1-4 marks)

I design six experiments:

1. 1 MapReduce application run on 1 VM and 1 physical host
2. 1 MapReduce application run on 2 VMs and 1 physical host
3. 1 MapReduce application run on 2 VMs and 2 physical hosts
4. 1 MapReduce application run on 4 VMs and 1 physical host
5. 1 MapReduce application run on 4 VMs and 2 physical hosts (2 VMs for each host)
6. 1 MapReduce application run on 4 VMs and 4 physical hosts

Then use Zabbix-based monitoring infrastructure to keep an eye on the power consumption, CPU usage and runtime for each experiment.

Implementation of the experiments (1-4 marks)

Expand the .txt file size for the application to 167MB. (I did this by copy and pasted the content of the book several times) In this way can make the application run for a longer time, and it's better for the observation.

Follow the instruction, run the MapReduce application on a single node. Use "time" command record the runtime for a single node as well as the application start time and end time. Then go to the Zabbix website, find the CPU usage and power consumption of the host within the period.

Then do the same thing for multi VMs and hosts. Create another VM and follow the instruction, make it as the slave node. For different experiments, live migrate the VM to the same/different hosts with the master node. Then use the same method to record the CPU, power and runtime.

Discussion of results (3-10 marks))

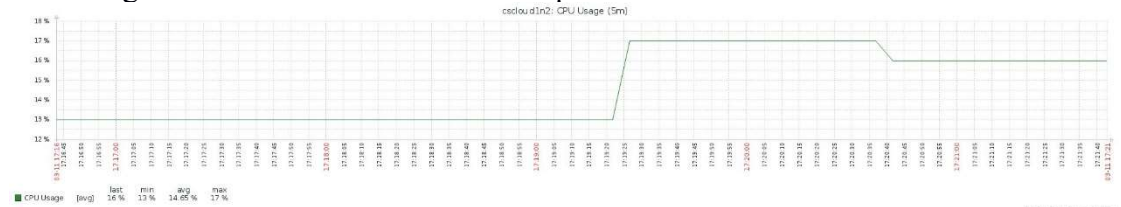
1. 1 MapReduce application run on 1 VM and 1 physical host

<input checked="" type="checkbox"/>	38567	sc19fm	users	test	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.11.70	
-------------------------------------	-------	--------	-------	------	---------	-----------------------------------	------------	--

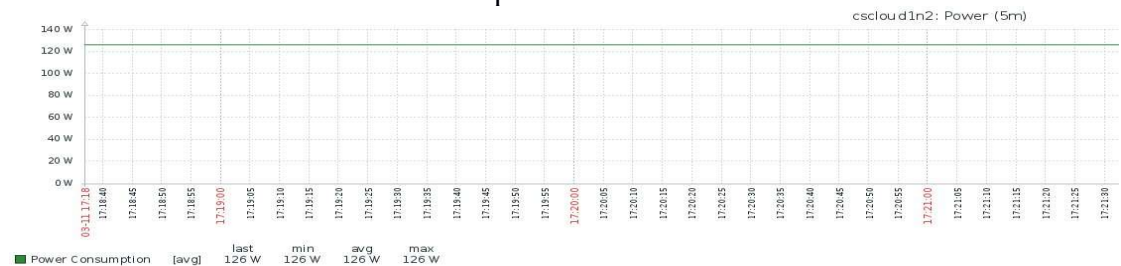
Runtime as below: 1:51.92

```
yarn jar OurMapReduceJob.jar org.myorg.WordCount hdfsbook.txt output6 97.72s user 5.64s system 92% cpu 1:51.92 total
```

CPU usage for host cscloud1n2.cloud.comp.leeds.ac.uk



Power for host cscloud1n2.cloud.comp.leeds.ac.uk



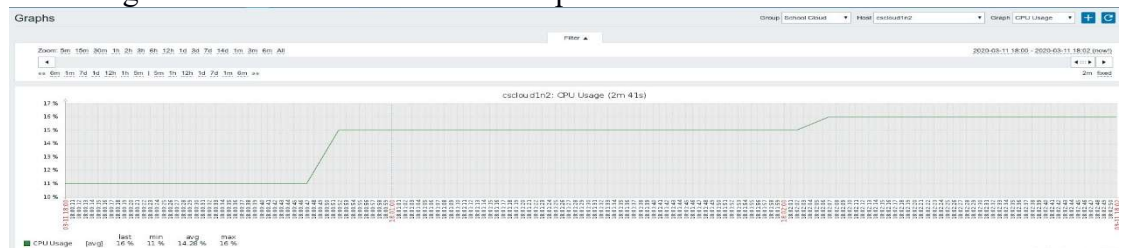
2. 1 MapReduce application run on 2 VMs and 1 physical host

<input type="checkbox"/>	ID	Owner	Group	Name	Status	Host	IPs	
<input checked="" type="checkbox"/>	38574	sc19fm	users	slave1	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.11.76	
<input checked="" type="checkbox"/>	38567	sc19fm	users	test	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.11.70	

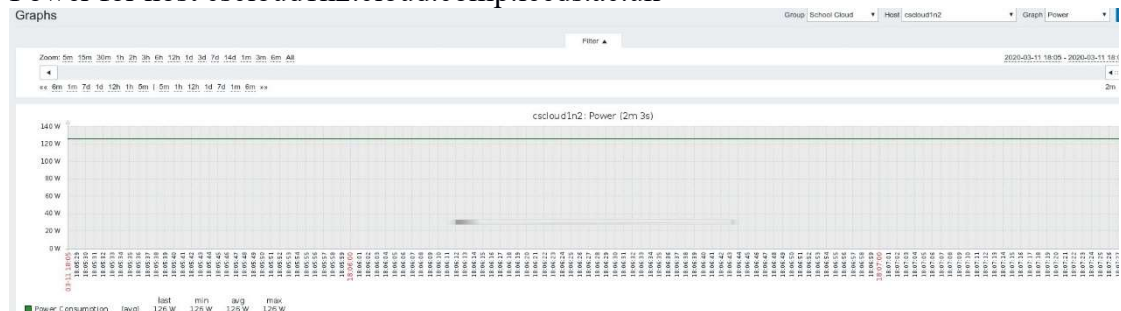
Runtime as below: 1:52.36

```
yarn jar OurMapReduceJob.jar org.myorg.WordCount hdfsbook.txt output9 99.00s user 5.62s system 93% cpu 1:52.36 total
```

CPU usage for host cscloud1n2.cloud.comp.leeds.ac.uk



Power for host cscloud1n2.cloud.comp.leeds.ac.uk



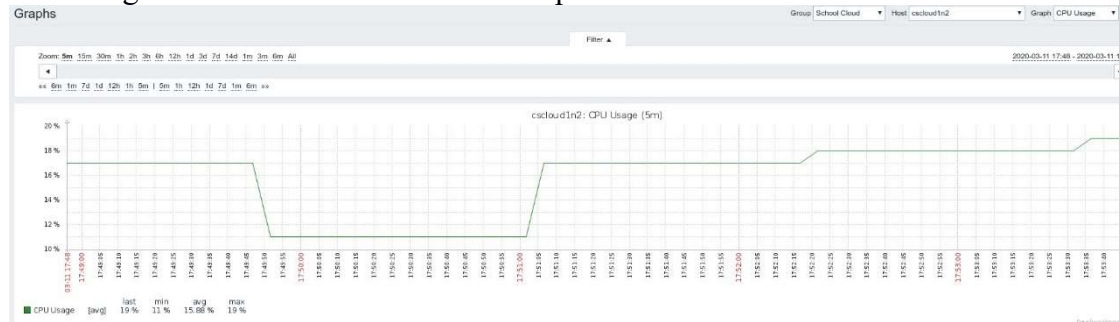
3. 1 MapReduce application run on 2 VMs and 2 physical hosts

ID	Owner	Group	Name	Status	Host	IPs
38574	sc19fm	users	slave1	RUNNING	cscloud1n9.cloud.comp.leeds.ac.uk	10.1.11.76
38567	sc19fm	users	test	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.11.70

Runtime as below: 1:57.31

yarn jar OurMapReduceJob.jar org.myorg.WordCount hdfsbook.txt output8 98.22s user 5.46s system 88% cpu 1:57.31 total

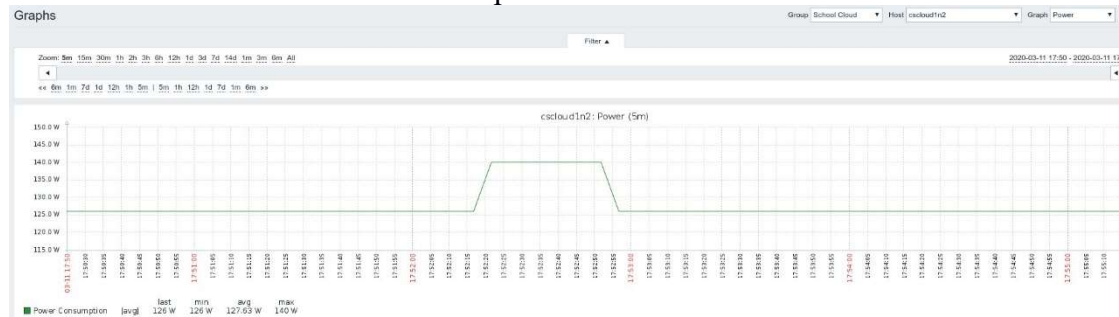
CPU usage for host cscloud1n2.cloud.comp.leeds.ac.uk



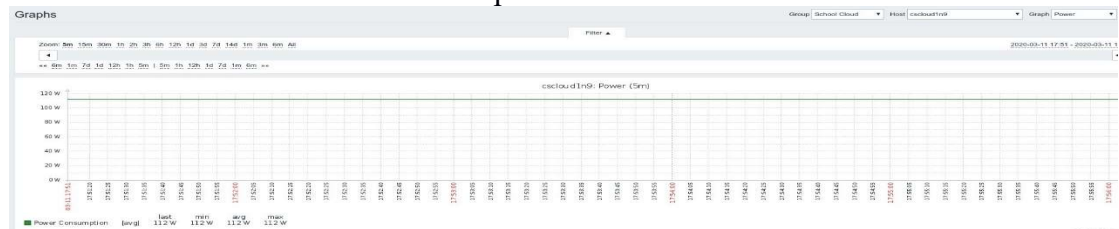
CPU usage for host cscloud1n9.cloud.comp.leeds.ac.uk



Power for host cscloud1n2.cloud.comp.leeds.ac.uk



Power for host cscloud1n9.cloud.comp.leeds.ac.uk



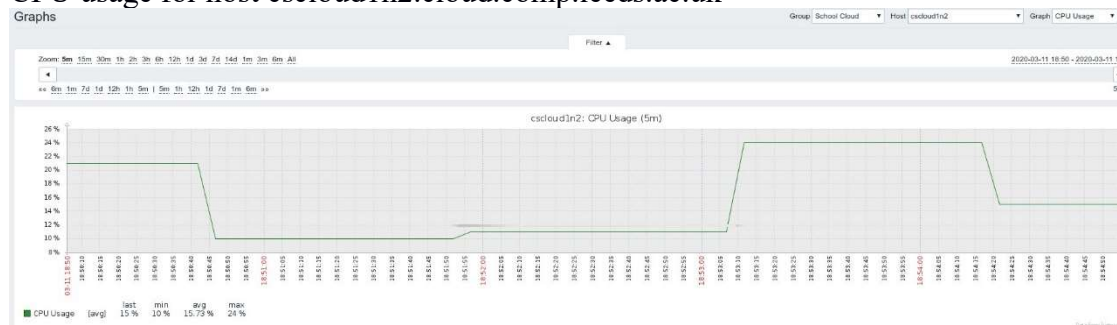
4. 1 MapReduce application run on 4 VMs and 1 physical host

<input checked="" type="checkbox"/>	38729	sc19fm	users	slave3	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.0.143	
<input checked="" type="checkbox"/>	38728	sc19fm	users	slave2	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.0.137	
<input checked="" type="checkbox"/>	38574	sc19fm	users	slave1	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.11.76	
<input checked="" type="checkbox"/>	38567	sc19fm	users	test	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.11.70	

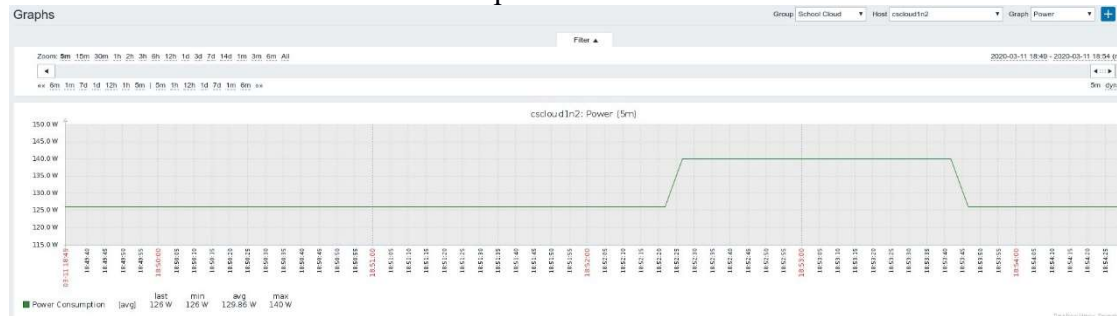
Runtime as below: 1:58.50

```
yarn jar OurMapReduceJob.jar org.myorg.WordCount hdfsbook.txt output10
ser 6.17s system 88% cpu 1:58.50 total
```

CPU usage for host cscloud1n2.cloud.comp.leeds.ac.uk



Power for host cscloud1n2.cloud.comp.leeds.ac.uk



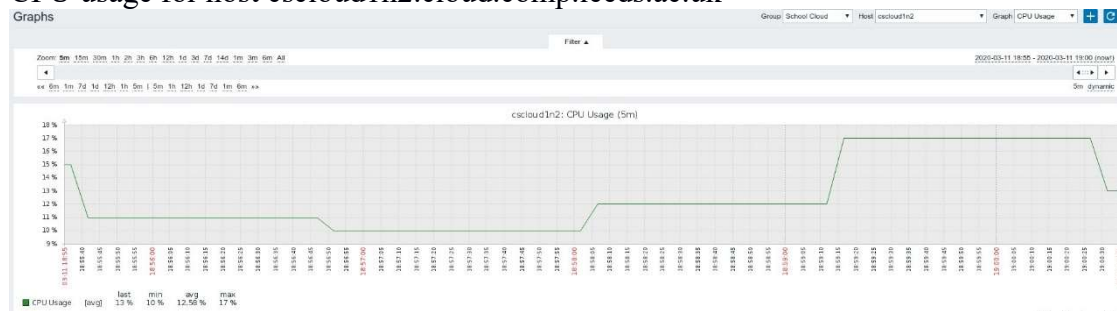
5. 1 MapReduce application run on 4 VMs and 2 physical hosts (2 VMs for each host)

<input type="checkbox"/>	ID	Owner	Group	Name	Status	Host	IPs	
<input checked="" type="checkbox"/>	38729	sc19fm	users	slave3	RUNNING	cscloud1n1.cloud.comp.leeds.ac.uk	10.1.0.143	
<input checked="" type="checkbox"/>	38728	sc19fm	users	slave2	RUNNING	cscloud1n1.cloud.comp.leeds.ac.uk	10.1.0.137	
<input checked="" type="checkbox"/>	38574	sc19fm	users	slave1	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.11.76	
<input checked="" type="checkbox"/>	38567	sc19fm	users	test	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.11.70	

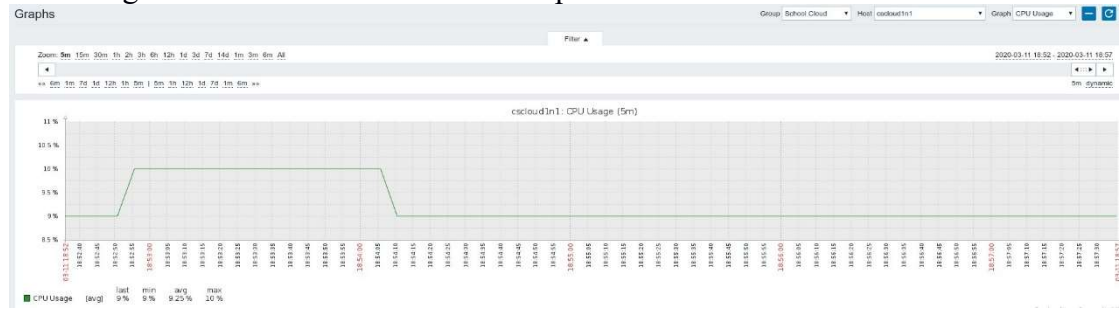
Runtime as below: 1:56.07

```
yarn jar OurMapReduceJob.jar org.myorg.WordCount hdfsbook.txt output11 97.60s u
ser 6.20s system 89% cpu 1:56.07 total
```

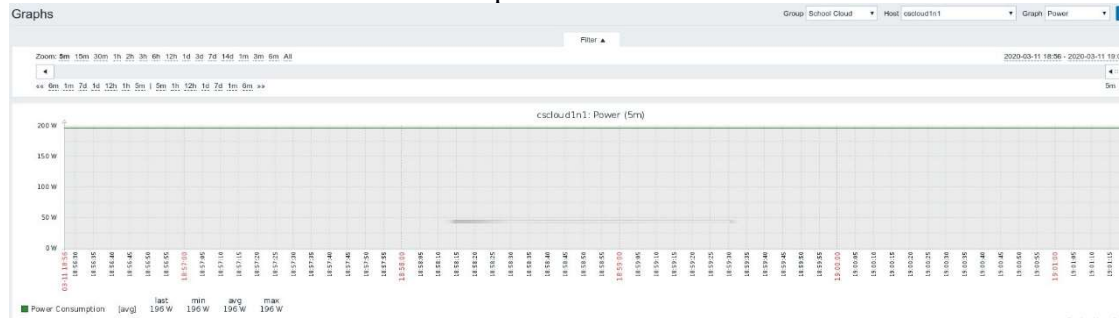
CPU usage for host cscloud1n2.cloud.comp.leeds.ac.uk



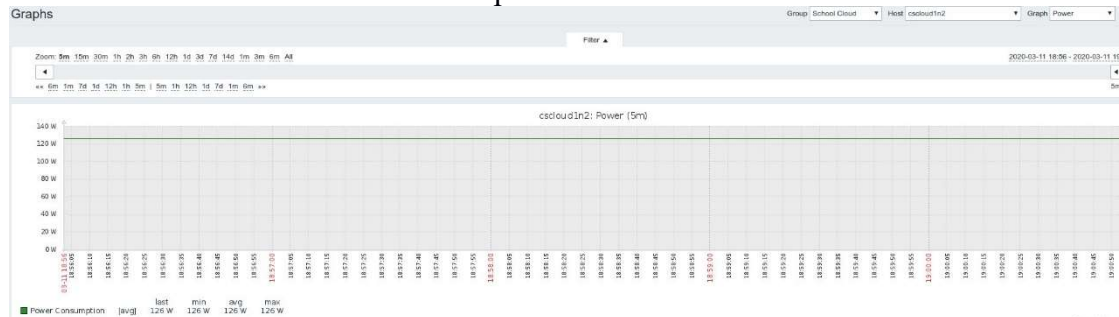
CPU usage for host cscloud1n1.cloud.comp.leeds.ac.uk



Power for host cscloud1n2.cloud.comp.leeds.ac.uk



Power for host cscloud1n1.cloud.comp.leeds.ac.uk



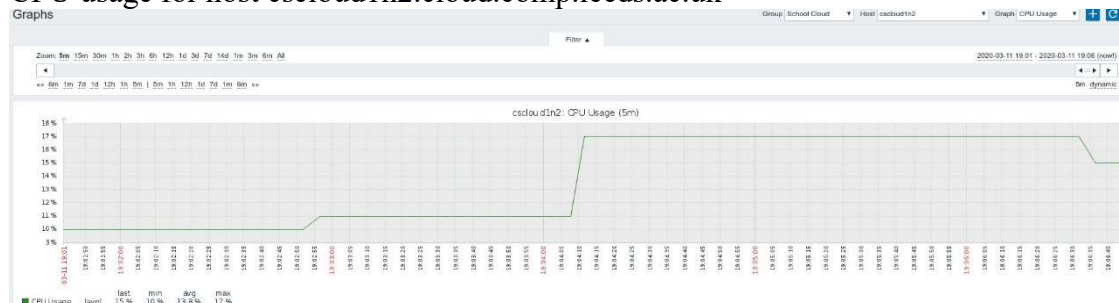
6. 1 MapReduce application run on 4 VMs and 4 physical hosts

<input type="checkbox"/>	ID	Owner	Group	Name	Status	Host	IPs
<input checked="" type="checkbox"/>	38729	sc19fm	users	slave3	RUNNING	cscloud1n1.cloud.comp.leeds.ac.uk	10.1.0.143
<input checked="" type="checkbox"/>	38728	sc19fm	users	slave2	RUNNING	cscloud1n3.cloud.comp.leeds.ac.uk	10.1.0.137
<input checked="" type="checkbox"/>	38574	sc19fm	users	slave1	RUNNING	cscloud1n12.cloud.comp.leeds.ac.uk	10.1.11.76
<input checked="" type="checkbox"/>	38567	sc19fm	users	test	RUNNING	cscloud1n2.cloud.comp.leeds.ac.uk	10.1.11.70

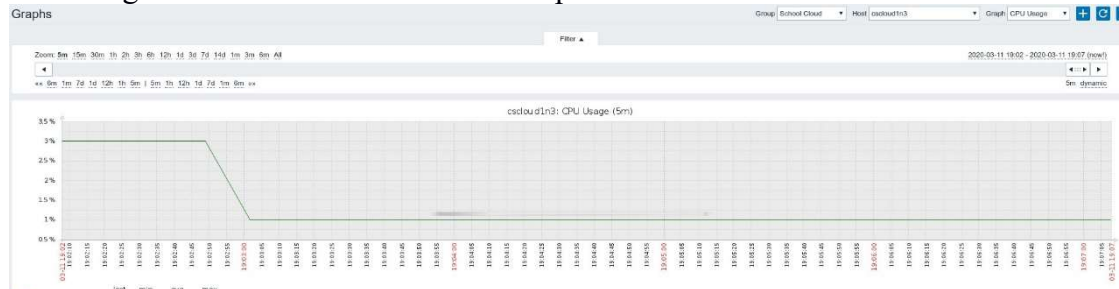
Runtime as below: 1:53.46

```
yarn jar OurMapReduceJob.jar org.myorg.WordCount hdfsbook.txt output12 97.80s user 5.70s system 91% cpu 1:53.46 total
```

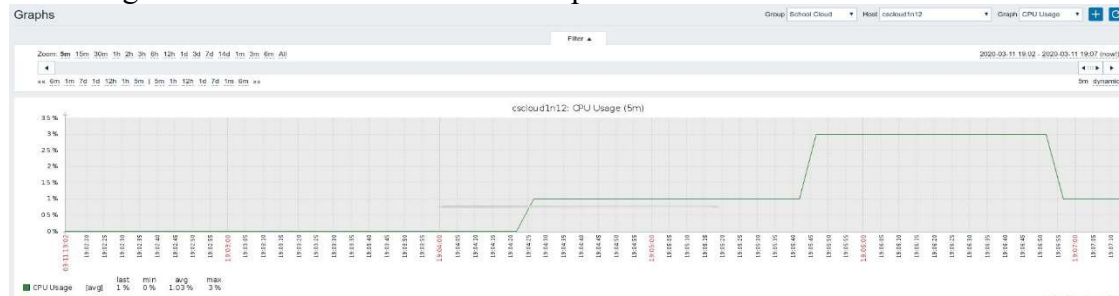
CPU usage for host cscloud1n2.cloud.comp.leeds.ac.uk



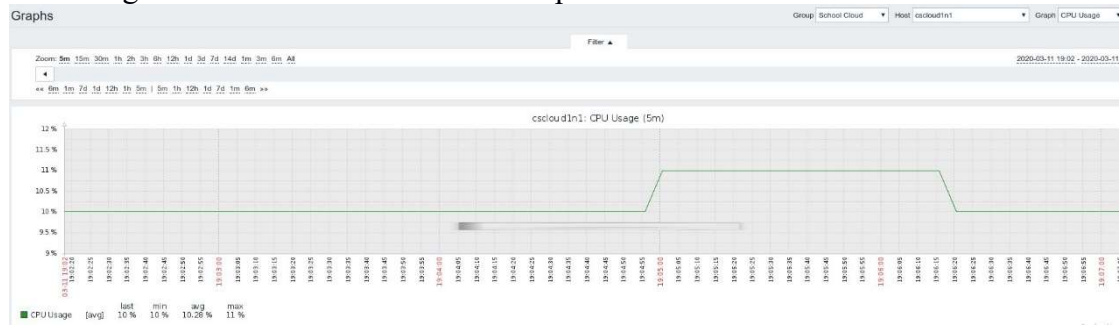
CPU usage for host cscloud1n3.cloud.comp.leeds.ac.uk



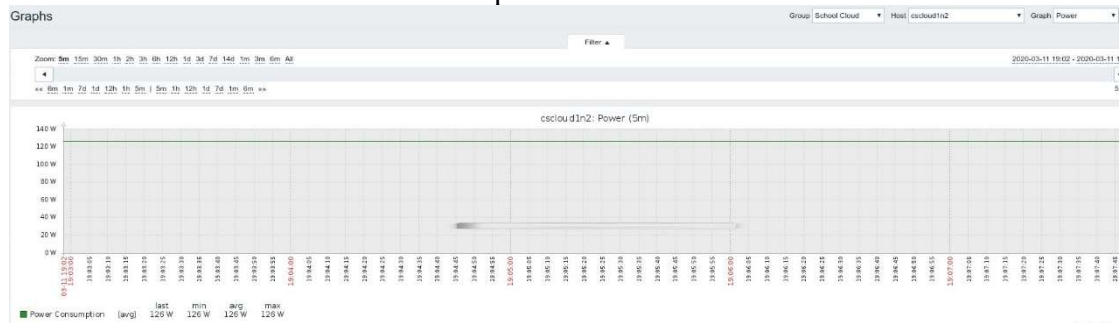
CPU usage for host cscloud1n2.cloud.comp.leeds.ac.uk



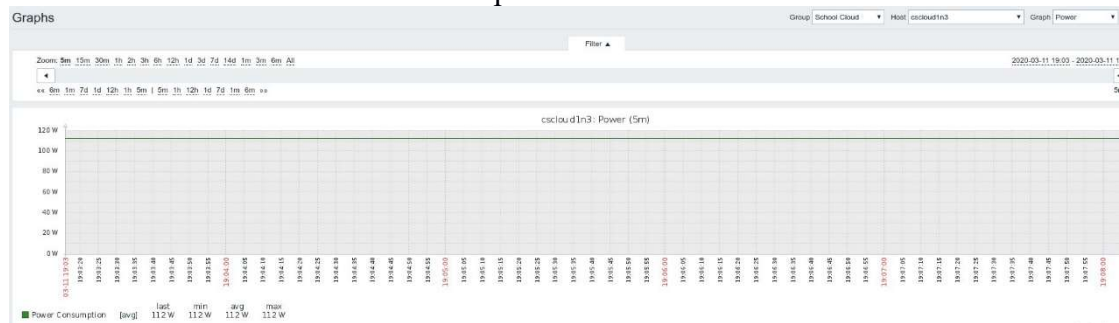
CPU usage for host cscloud1n1.cloud.comp.leeds.ac.uk



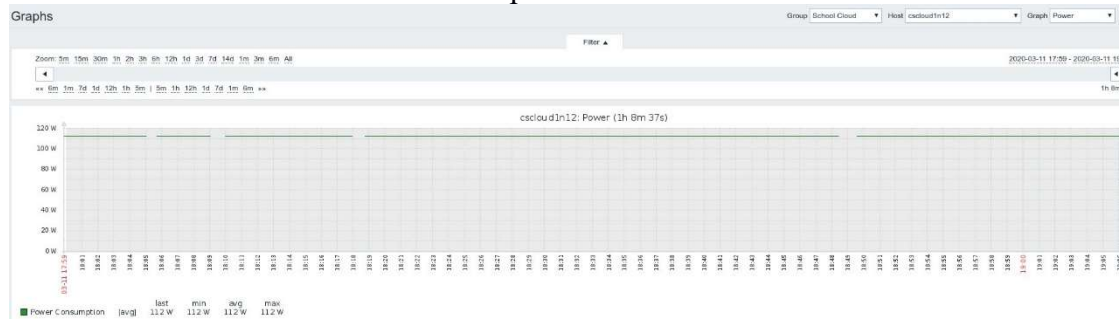
Power for host cscloud1n2.cloud.comp.leeds.ac.uk



Power for host cscloud1n3.cloud.comp.leeds.ac.uk



Power for host cscloud1n12.cloud.comp.leeds.ac.uk



Power for host cscloud1n1.cloud.comp.leeds.ac.uk



Comments follow the screenshots show above.

For the runtime, different experiments only show a small difference. And this is because VM is independent, they not be affected by other VM even if they are running on the same physical host. But these VMs have a master and slave relationship, they can communicate with each other. And this communication can happen in LAN or WAN, and it depends on if these VMs run on the same host. Communication in LAN can be faster than WAN. And this could cause different runtime.

For CPU usage, the usage increased 2 or 3 seconds after I input the execution command; this is probably because of the network delay. With the same number of VM that runs on a different number of hosts, the CPU usage can be changed. If the VMs run on a separate physical machine, the CPU for each host is lower. And this is because hosts share the load balance—for example one VM for each host, rather than 4 VMs in the same host.

For power consumption, this is quite strange cause the power barely changes when I run the application. And this probably because about 80 VMs are running on each host. And each time when running a form on one of the VM, this can be a reasonably small change for the host. So the power is not likely to be affected by a tiny application. Another thinking is because the host's CPU is not able for turbo boost, which means it cannot change it's frequency and always provide it's best performance.

Evidence of successful run, e.g. screenshot (4 marks)

```
2020-03-11 19:48:16,063 INFO mapred.Task: Task 'attempt_local1142230462_0001_r_000000_0' done.
2020-03-11 19:48:16,074 INFO mapred.Task: Final Counters for attempt_local1142230462_0001_r_000000_0: Count$
  File System Counters
    FILE: Number of bytes read=1041239444
    FILE: Number of bytes written=1041732362
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=174700491
    HDFS: Number of bytes written=546391
    HDFS: Number of read operations=12
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=3
  Map-Reduce Framework
    Combine input records=0
    Combine output records=0
    Reduce input groups=41992
    Reduce shuffle bytes=347078276
    Reduce input records=29448120
    Reduce output records=41992
    Spilled Records=29448120
    Shuffled Maps =2
    Failed Shuffles=0
    Merged Map outputs=2
    GC time elapsed (ms)=7
    Total committed heap usage (bytes)=170020864
  Shuffle Errors
    BAD_ID=0
    CONNECTION=0
    IO_ERROR=0
  File Output Format Counters
    Bytes Written=546391
2020-03-11 19:48:16,079 INFO mapred.LocalJobRunner: Finishing task: attempt_local1142230462_0001_r_000000_0
2020-03-11 19:48:16,079 INFO mapred.LocalJobRunner: reduce task executor complete.
2020-03-11 19:48:16,384 INFO mapreduce.Job: map 100% reduce 100%
2020-03-11 19:48:16,385 INFO mapreduce.Job: Job job_1142230462_0001 completed successfully
2020-03-11 19:48:17,710 INFO mapreduce.Job: Counters: 35
  File System Counters
    FILE: Number of bytes read=1654985926
    FILE: Number of bytes written=2270202652
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=483622806
    HDFS: Number of bytes written=546391
    HDFS: Number of read operations=24
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=5
  Map-Reduce Framework
    Map input records=3434860
    Map output records=29448120
    Map output bytes=288182024
    Map output materialized bytes=347078276
    Input split bytes=224
    Combine input records=0
    Combine output records=0
    Reduce input groups=41992
    Reduce shuffle bytes=347078276
    Reduce input records=29448120

    Spilled Records=88344360
    Shuffled Maps =2
    Failed Shuffles=0
    Merged Map outputs=2
    GC time elapsed (ms)=1014
    Total committed heap usage (bytes)=462876672
  Shuffle Errors
    BAD_ID=0
    CONNECTION=0
    IO_ERROR=0
    WRONG_LENGTH=0
    WRONG_MAP=0
    WRONG_REDUCE=0
  File Input Format Counters
    Bytes Read=174700491
  File Output Format Counters
    Bytes Written=546391
yarn jar OurMapReduceJob.jar org.myorg.WordCount hdfsbook.txt outputttt 104.18s user 6.97s system 87% cpu 2:06.57 total
```

CPU

And the output file as follows:

```
hduser@VM10-1-11-70 /tmp | 80x24 | pts/2
File Edit View Search Terminal Help
GNU nano 2.7.4 File: output
#2600] 1
$5,000) 1
($1 1
(1) 17
(2) 27
(3) 12
(4) 1
(5) 1
(801) 1
(At 1
(Barclay 1
(Berg 3
(Borodinó) 1
(Boris 1
(By 1
(Daniel 1
(Davout). 1
(Denísov's 0
(Ermólov 1
[ Read 41992 lines ]
```

Any other comments: