



DIGITAL  
TALENT  
SCHOLARSHIP

# DIGITAL TALENT SCHOLARSHIP 2019

## Big Data Analytics





DIGITAL  
TALENT  
SCHOLARSHIP

# Hadoop Multi-Node Cluster





DIGITAL  
TALENT  
SCHOLARSHIP

# Membuat Instance untuk Master & Slave



# Membuat Master (Namenode)

- Beri nama instance yang akan menjadi master (*namenode*) untuk memudahkan pengelolaan

- Saat hover, klik tombol pensil

- Catat (*copy*) public DNS:

- ec-2-x-x-x.compute1...**

- Catat Private IPs

The screenshot shows the AWS Management Console interface. At the top, there's a navigation bar with 'Services', 'Resource Groups', and a user profile. Below this is a table of instances. The first instance is named 'namenode' with ID 'i-03b40f69dfc1b4cae', type 't2.micro', and availability zone 'us-east-1c'. It is in a 'running' state. A red arrow points from the 'namenode' name in the table to the 'Description' tab of the instance details. The 'Description' tab shows various instance details. A red arrow points from the 'Public DNS (IPv4)' field, which contains 'ec2-54-88-40-230.compute-1.amazonaws.com', to the red text 'ec-2-x-x-x.compute1...'. Another red arrow points from the 'Private DNS' field, which contains 'ip-172-31-84-70.ec2.internal', to the red text '70.ec2.internal'. A third red arrow points from the 'Private IPs' field, which contains '172.31.84.70', to the red text '172.31.84.70'.

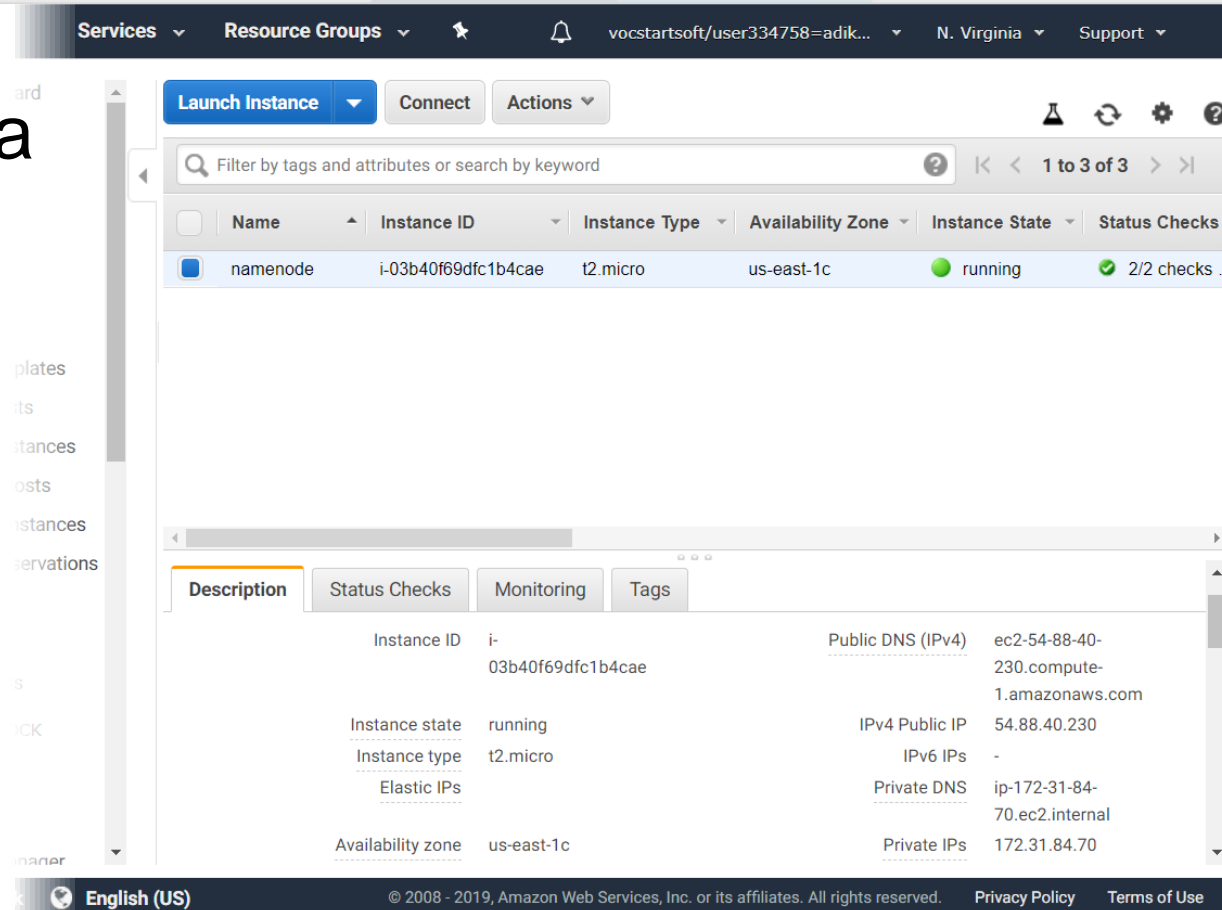
Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks
namenode	i-03b40f69dfc1b4cae	t2.micro	us-east-1c	running	2/2 checks ..

Field	Value
Instance ID	i-03b40f69dfc1b4cae
Instance state	running
Instance type	t2.micro
Elastic IPs	
Availability zone	us-east-1c
Public DNS (IPv4)	ec2-54-88-40-230.compute-1.amazonaws.com
IPv4 Public IP	54.88.40.230
IPv6 IPs	-
Private DNS	ip-172-31-84-70.ec2.internal
Private IPs	172.31.84.70

# Membuat Slave (Datanode) (1)

- Dari instance yang sudah ada yang menjadi master/ namenode maka buat Image baru untuk menjadi datanode



The screenshot displays the AWS Management Console interface. At the top, there's a navigation bar with 'Services', 'Resource Groups', and a search bar. Below this, a table lists EC2 instances. One instance, named 'namenode', is highlighted. It has an Instance ID of 'i-03b40f69dfc1b4cae', is of type 't2.micro', located in 'us-east-1c', and is in a 'running' state. The status checks show '2/2 checks passed'.

Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks
namenode	i-03b40f69dfc1b4cae	t2.micro	us-east-1c	running	2/2 checks passed

Below the table, the 'Description' tab is selected, showing details for the instance 'i-03b40f69dfc1b4cae'.

Property	Value	Property	Value
Instance ID	i-03b40f69dfc1b4cae	Public DNS (IPv4)	ec2-54-88-40-230.compute-1.amazonaws.com
Instance state	running	IPv4 Public IP	54.88.40.230
Instance type	t2.micro	IPv6 IPs	-
Elastic IPs	-	Private DNS	ip-172-31-84-70.ec2.internal
Availability zone	us-east-1c	Private IPs	172.31.84.70

# Membuat Slave (Datanode) (2)

- Klik kanan pada instance namenode pilih menu **Image** > **Create Image**

The screenshot shows the AWS Management Console interface. At the top, there are navigation tabs for 'Services', 'Resource Groups', and a search bar. Below this, there are buttons for 'Launch Instance', 'Connect', and 'Actions'. A search bar with the text 'Filter by tags and attributes or search by keyword' is present. The main table lists instances, with one instance named 'namenode' selected. A right-click context menu is open over this instance, showing options like 'Connect', 'Get Windows Password', 'Create Template From Instance', 'Launch More Like This', 'Instance State', 'Instance Settings', 'Image', 'Networking', and 'CloudWatch Monitoring'. The 'Image' option is highlighted, and a sub-menu is shown with 'Create Image' selected. Below the menu, the instance details are visible, including Instance ID, Instance state (running), Instance type (t2.micro), and Availability zone (us-east-1c).

Description		Status Checks		Monitoring		Tags	
Instance ID	i-03b40f69dfc1b4cae	Public DNS (IPv4)	ec2-54-88-40-230.compute-1.amazonaws.com	IPv4 Public IP	54.88.40.230	IPv6 IPs	-
Instance state	running	Private DNS	ip-172-31-84-70.ec2.internal	Private IPs	172.31.84.70		
Instance type	t2.micro						
Elastic IPs							
Availability zone	us-east-1c						



DIGITAL  
TALENT  
SCHOLARSHIP

# Membuat Slave (Datanode) (3)

## Create Image

Instance ID ⓘ

i-03b40f69dfc1b4cae

Image name ⓘ

slave

Image description ⓘ

slave for hadoop

No reboot ⓘ

☐

### Instance Volumes

Volume Type ⓘ	Device ⓘ	Snapshot ⓘ	Size (GiB) ⓘ	Volume Type ⓘ	IOPS ⓘ	Throughput (MB/s) ⓘ	Delete on Termination ⓘ	Encrypted ⓘ
Root	/dev/sda1	snap-0e384451033d3767e	8	General Purpose S3	100 / 3000	N/A	<input checked="" type="checkbox"/>	Not Encrypted

Add New Volume

Total size of EBS Volumes: 8 GiB  
When you create an EBS image, an EBS snapshot will also be created for each of the above volumes.

Cancel Create Image



DIGITAL  
TALENT  
SCHOLARSHIP

# Membuat Slave (Datanode) (4)

aws Services Resource Groups vocstartsoft/user334758=adik... N. Virginia Support

EC2 Dashboard Launch Instance Connect Actions

Events

Tags Filter by tags and attributes or search by keyword 1 to 1 of 1

### Create Image

✓ Create Image request received.  
View pending image [ami-0e305be327d273501](#)

Any snapshots backing your new EBS image can be managed on the [snapshots screen](#) after successful image creation.

Close

IMAGES AMIs Bundle Tasks

Elastic IPs

Private DNS ip-172-31-84-70.ec2.internal

Availability zone us-east-1c

Private IPs 172.31.84.70

Security groups launch-wizard-1.

Secondary private IPs

Feedback English (US) © 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use





DIGITAL  
TALENT  
SCHOLARSHIP

# Membuat Slave (Datanode) (5)

The screenshot shows the AWS Management Console interface. The left sidebar contains navigation links for EC2 Dashboard, Events, Tags, Reports, Limits, INSTANCES, and IMAGES. The main content area displays a table of AMIs. The table has columns for Name, AMI Name, AMI ID, Source, Owner, and Visibility. A single row is visible with the name 'slave', AMI ID 'ami-0e305be327d273501', and Source '151341626195/...'. Below the table, the details for the selected AMI are shown, including its ID, Name, Owner, and Source.

Name	AMI Name	AMI ID	Source	Owner	Visibility
slave	ami-0e305be327d273501	151341626195/...	151341626195	Private	

Image: ami-0e305be327d273501

AMI ID	AMI Name
ami-0e305be327d273501	slave

Owner	Source
151341626195	151341626195/slave

# Membuat Slave (Datanode) (6)

The screenshot shows the AWS Management Console interface. On the left sidebar, the 'INSTANCES' section is expanded, and 'Launch Templates' is selected. The main content area shows a table of AMIs. A context menu is open over the AMI named 'slave', showing options like 'Launch', 'Spot Request', 'Deregister', 'Register New AMI', 'Copy AMI', 'Modify Image Permissions', 'Add/Edit Tags', and 'Modify Boot Volume Setting'. Below the table, the 'Details' tab is selected, showing the AMI ID 'ami-0e305be327d273501', the Owner '151341626195', and the Source '151341626195/slave'.

Name	AMI Name	AMI ID	Source	Owner	Visibility
slave			151341626195/...	151341626195	Private

Image: ami-0e305be327d273501

AMI ID	AMI Name
ami-0e305be327d273501	slave

Owner	Source
151341626195	151341626195/slave



DIGITAL  
TALENT  
SCHOLARSHIP

# Membuat Slave (Datanode) (7)

**aws** Services ▾ Resource Groups ▾ ☆ 🔔 vocstartsoft/user334758=adik... ▾ N. Virginia ▾ Support ▾

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

## Step 2: Choose an Instance Type

your applications. [Learn more](#) about instance types and how they can meet your computing needs.

Filter by: All instance types ▾ Current generation ▾ [Show/Hide Columns](#)

**Currently selected:** t2.micro (Variable ECUs, 1 vCPUs, 2.5 GHz, Intel Xeon Family, 1 GiB memory, EBS only)

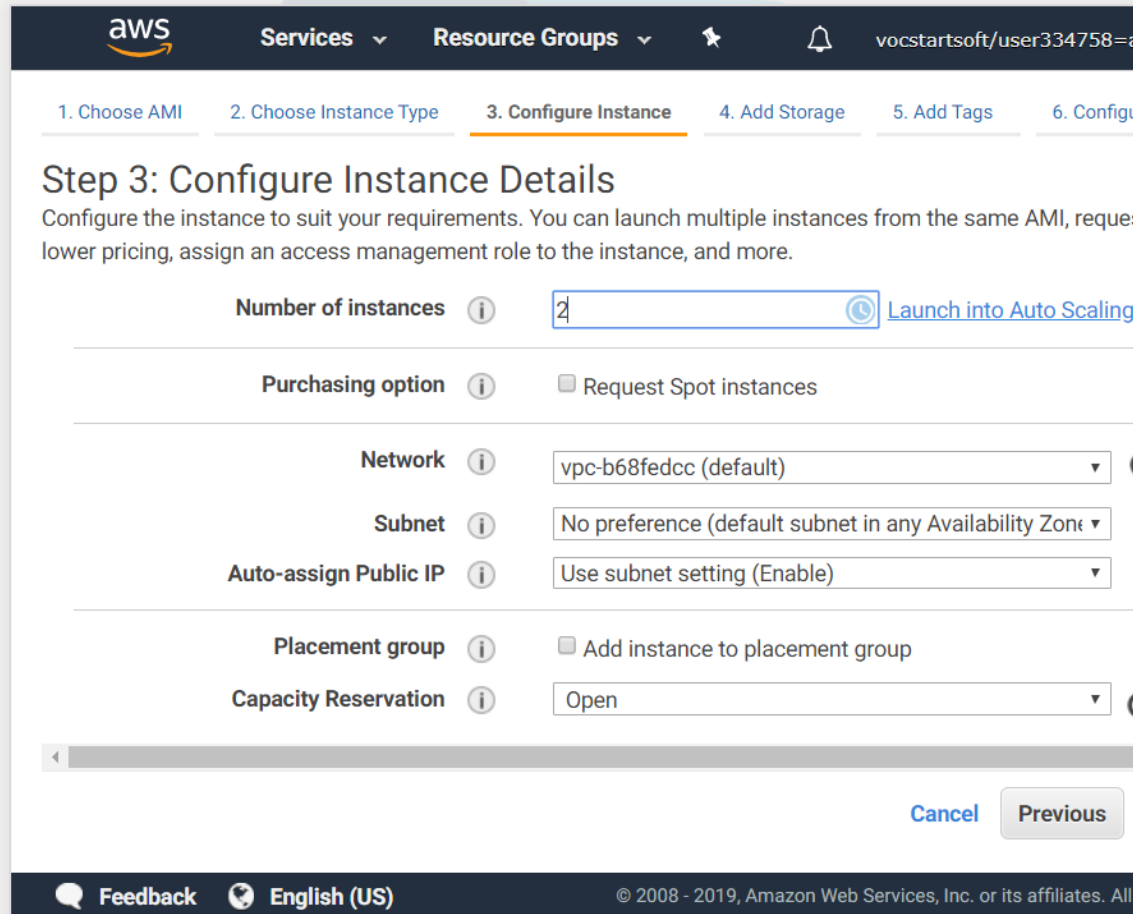
	Family ▾	Type ▾	vCPUs ⓘ ▾	Memory (GiB) ▾	Instance Storage (GB) ⓘ ▾	EBS-Optimized Available ⓘ ▾	Network Performance ⓘ ▾	IPv6 Support ⓘ
<input type="checkbox"/>	General purpose	t2.nano	1	0.5	EBS only	-	Low to Moderate	Yes
<input checked="" type="checkbox"/>	General purpose	t2.micro Free tier eligible	1	1	EBS only	-	Low to Moderate	Yes
<input type="checkbox"/>	General purpose	t2.small	1	2	EBS only	-	Low to Moderate	Yes

[Cancel](#) [Previous](#) [Review and Launch](#) [Next: Configure Instance Details](#)

[Feedback](#) [English \(US\)](#) © 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)

# Membuat Slave (Datanode) (8)

- Isi Number of instance sebanyak datanode yang ingin dibuat
- Misalnya dalam contoh buat 2 datanode



aws Services Resource Groups vocstartsoft/user334758=a

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure

### Step 3: Configure Instance Details

Configure the instance to suit your requirements. You can launch multiple instances from the same AMI, request lower pricing, assign an access management role to the instance, and more.

Number of instances ⓘ 2 [Launch into Auto Scaling](#)

Purchasing option ⓘ ☐ Request Spot instances

Network ⓘ vpc-b68fedcc (default) ▼

Subnet ⓘ No preference (default subnet in any Availability Zone) ▼

Auto-assign Public IP ⓘ Use subnet setting (Enable) ▼

Placement group ⓘ ☐ Add instance to placement group

Capacity Reservation ⓘ Open ▼

[Cancel](#) [Previous](#)

Feedback English (US) © 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.



DIGITAL  
TALENT  
SCHOLARSHIP

# Membuat Slave (Datanode) (9)

**aws** Services ▾ Resource Groups ▾ ★ 🔔 vocstartsoft/user334758=adik... ▾ N. Virginia ▾ Support ▾

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

## Step 7: Review Instance Launch


Please review your instance launch details. You can go back to edit changes for each section. Click **Launch** to assign a key pair to your instance and complete the launch process.

**⚠ Improve your instances' security. Your security group, launch-wizard-3, is open to the world.**

Your instances may be accessible from any IP address. We recommend that you update your security group rules to allow access from known IP addresses only.

You can also open additional ports in your security group to facilitate access to the application or service you're running, e.g., HTTP (80) for web servers. [Edit security groups](#)



▼ AMI Details [Edit AMI](#)

 **slave - ami-0e305be327d273501**

slave for hadoop

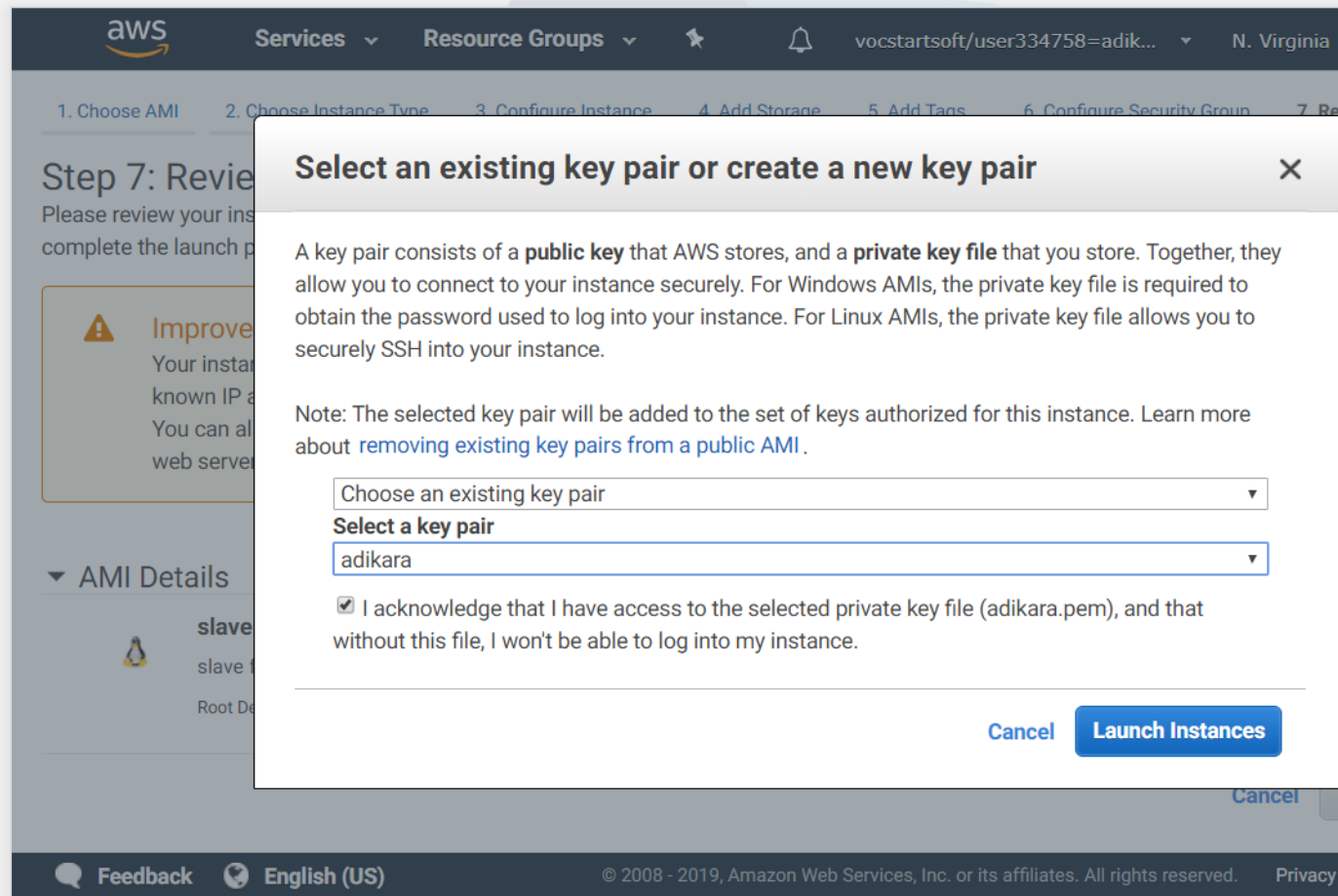
Root Device Type: ebs Virtualization type: hvm

[Cancel](#) [Previous](#) [Launch](#)

 [Feedback](#)  [English \(US\)](#) © 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)

# Membuat Slave (Datanode) (10)

- Gunakan key pair yang sama dengan yang dibuat untuk namenode



aws Services Resource Groups vocstartsoft/user334758=adik... N. Virginia

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

Step 7: Review

Please review your instance configuration and complete the launch process.

**Improve Your instance**

Your instance is not yet optimized for performance. You can always improve your instance performance by using the Amazon EC2 Instance Scheduler.

AMI Details

slave

slave

Root Device

**Select an existing key pair or create a new key pair**

A key pair consists of a **public key** that AWS stores, and a **private key file** that you store. Together, they allow you to connect to your instance securely. For Windows AMIs, the private key file is required to obtain the password used to log into your instance. For Linux AMIs, the private key file allows you to securely SSH into your instance.

Note: The selected key pair will be added to the set of keys authorized for this instance. Learn more about [removing existing key pairs from a public AMI](#).

Choose an existing key pair

Select a key pair

adikara

☒ I acknowledge that I have access to the selected private key file (adikara.pem), and that without this file, I won't be able to log into my instance.

Cancel Launch Instances

Feedback English (US)

© 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy



DIGITAL  
TALENT  
SCHOLARSHIP

# Membuat Slave (Datanode) (12)

**aws** Services ▾ Resource Groups ▾ ☆ 🔔 vocstartsoft/user334758=adik... ▾ N. Virginia ▾ Support ▾

## Launch Status

**Your instances are now launching**  
The following instance launches have been initiated: [i-073039aec610c5bbd](#), [i-088c9a66b1696a281](#) [View launch log](#)

**Get notified of estimated charges**  
[Create billing alerts](#) to get an email notification when estimated charges on your AWS bill exceed an amount you define (for example, if you exceed the free usage tier).

### How to connect to your instances

Your instances are launching, and it may take a few minutes until they are in the **running** state, when they will be ready for you to use. Usage hours on your new instances will start immediately and continue to accrue until you stop or terminate your instances.

Click **View Instances** to monitor your instances' status. Once your instances are in the **running** state, you can **connect** to them from the Instances screen. [Find out](#) how to connect to your instances.

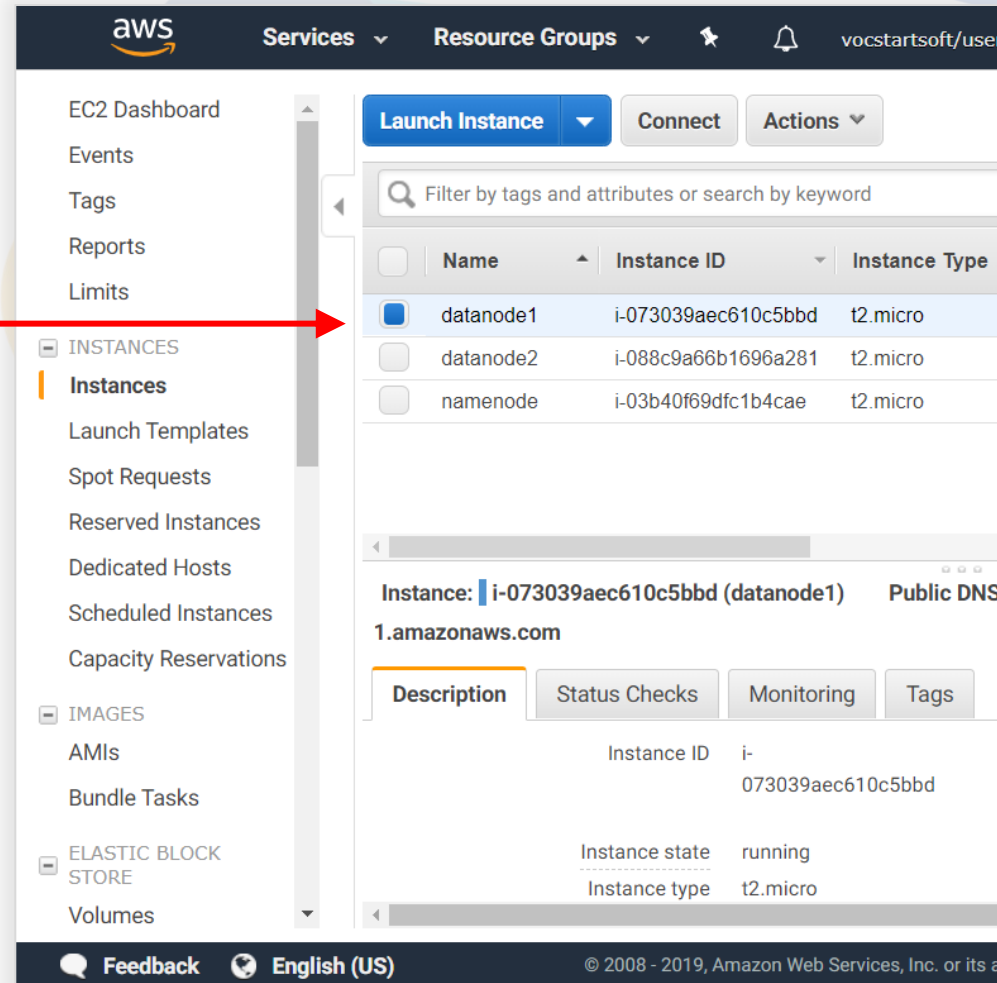
▼ Here are some helpful resources to get you started

Feedback English (US) © 2008 - 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#)



# Membuat Slave (Datanode) (13)

- Ubah nama masing-masing instance dari image yang baru dibuat menjadi **datanode1**, **datanode2**, dst.



aws Services Resource Groups vocstartsoft/user

EC2 Dashboard

Events

Tags

Reports

Limits

INSTANCES

Instances

Launch Templates

Spot Requests

Reserved Instances

Dedicated Hosts

Scheduled Instances

Capacity Reservations

IMAGES

AMIs

Bundle Tasks

ELASTIC BLOCK STORE

Volumes

Launch Instance Connect Actions

Filter by tags and attributes or search by keyword

	Name	Instance ID	Instance Type
<input checked="" type="checkbox"/>	datanode1	i-073039aec610c5bbd	t2.micro
<input type="checkbox"/>	datanode2	i-088c9a66b1696a281	t2.micro
<input type="checkbox"/>	namenode	i-03b40f69dfc1b4cae	t2.micro

Instance: i-073039aec610c5bbd (datanode1) Public DNS 1.amazonaws.com

Description Status Checks Monitoring Tags

Instance ID i-073039aec610c5bbd

Instance state running

Instance type t2.micro

Feedback English (US) © 2008 - 2019, Amazon Web Services, Inc. or its a





DIGITAL  
TALENT  
SCHOLARSHIP

TERBUKA  
UNTUK  
DISABILITAS

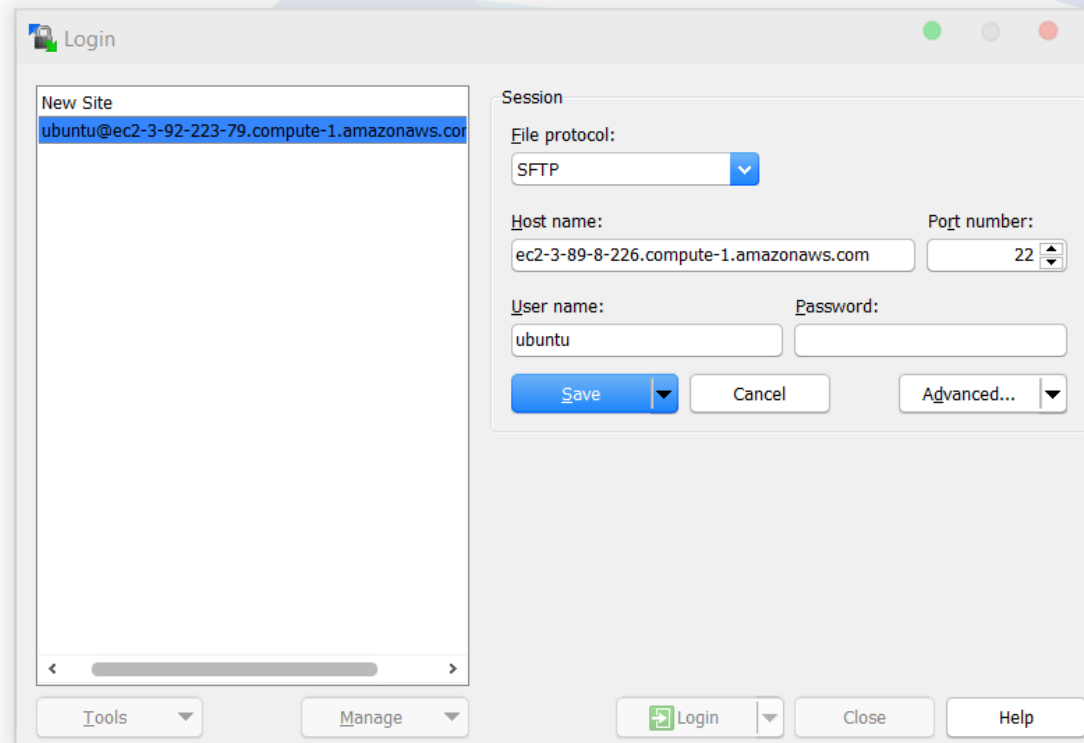
BREAK  
YOUR  
LIMITS!

# Buat Koneksi SSH



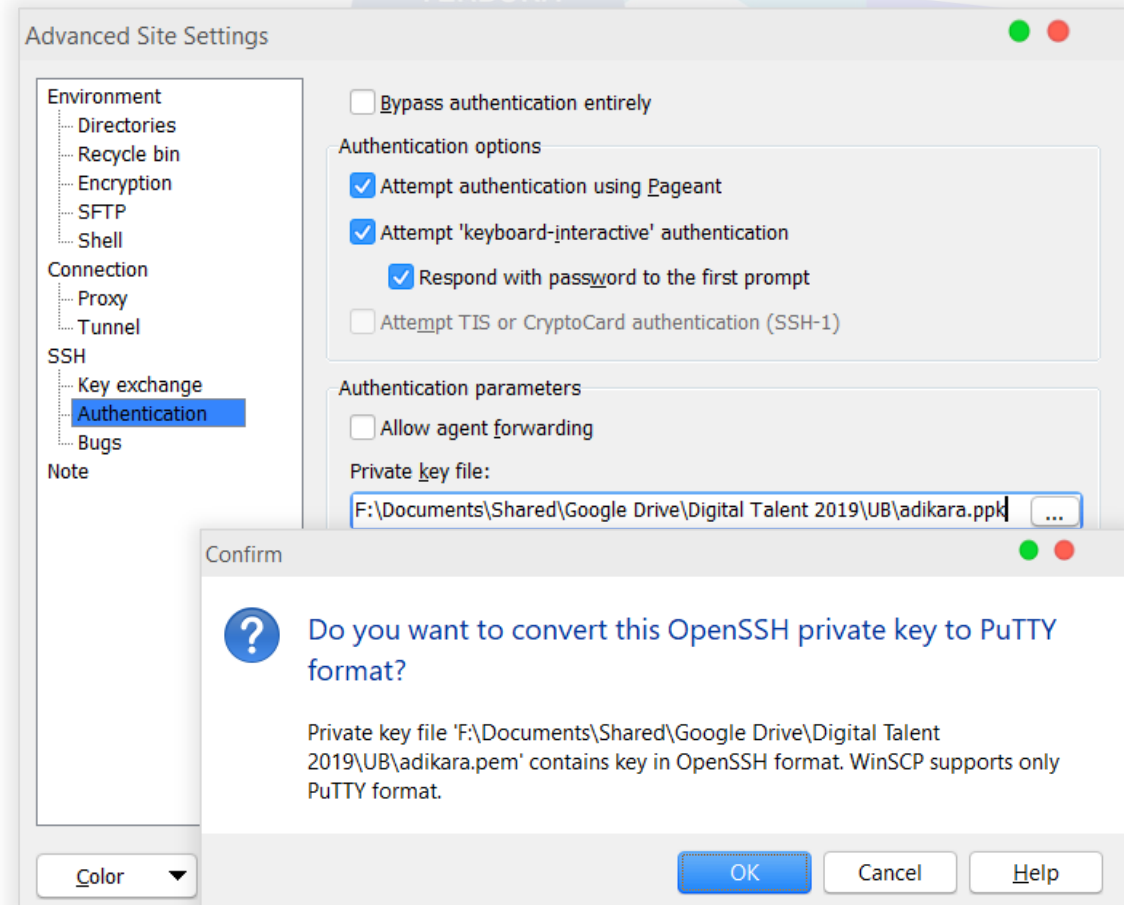
# Buat Koneksi dengan WinSCP (1)

- Buka WinSCP
- Isikan hostname sesuai Public DNS di EC2
- Isikan User name = "ubuntu"
- Klik tombol Advanced...



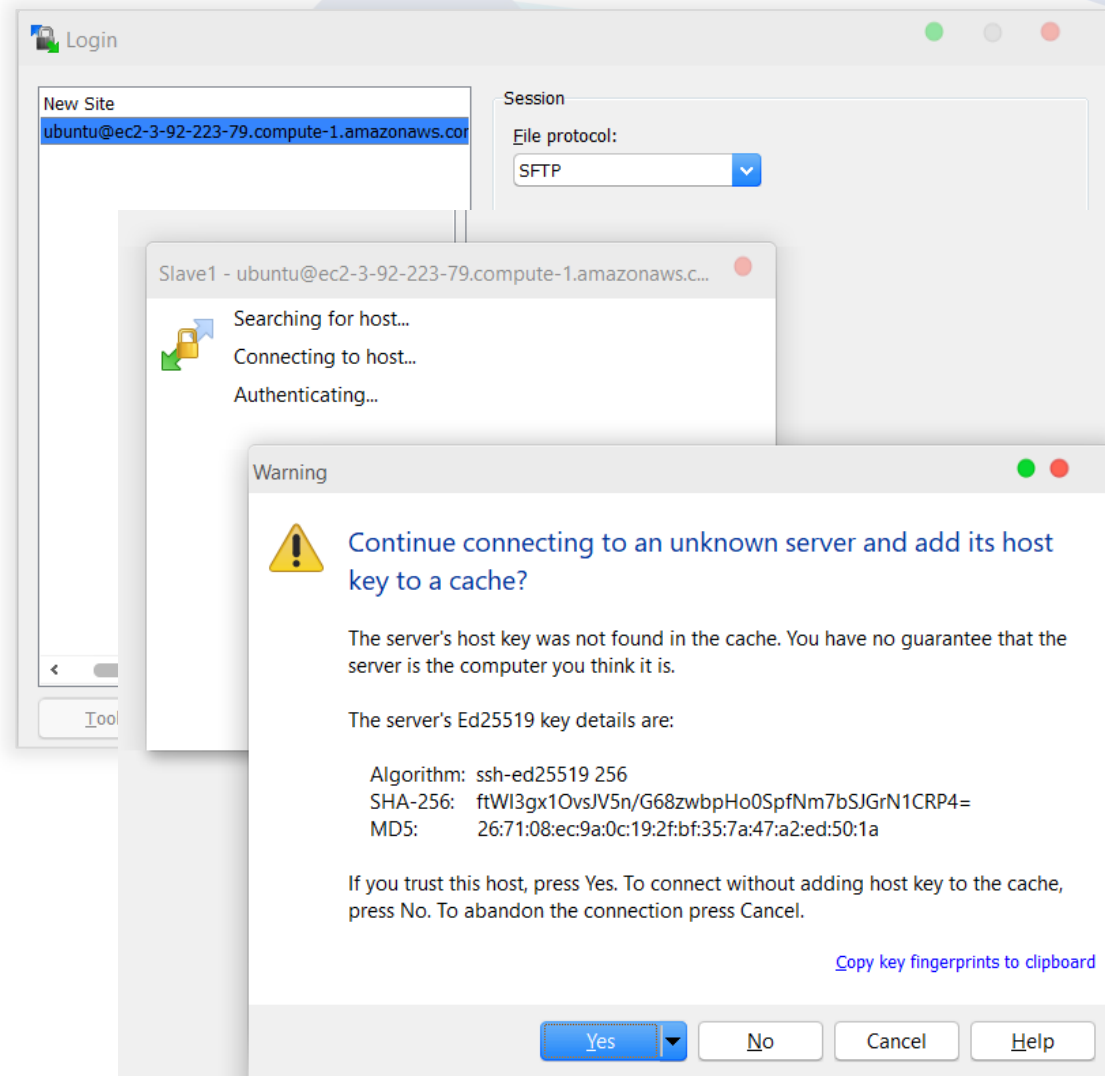
# Buat Koneksi dengan WinSCP (2)

- Pilih SSH > Authentication
- Pilih fail .pem yang sudah diunduh
- Ketika diminta mengubah dari fail OpenSSH (.pem) ke PuTTY (.ppk) pilih OK
- Tutup jendela Advanced dengan pilih OK



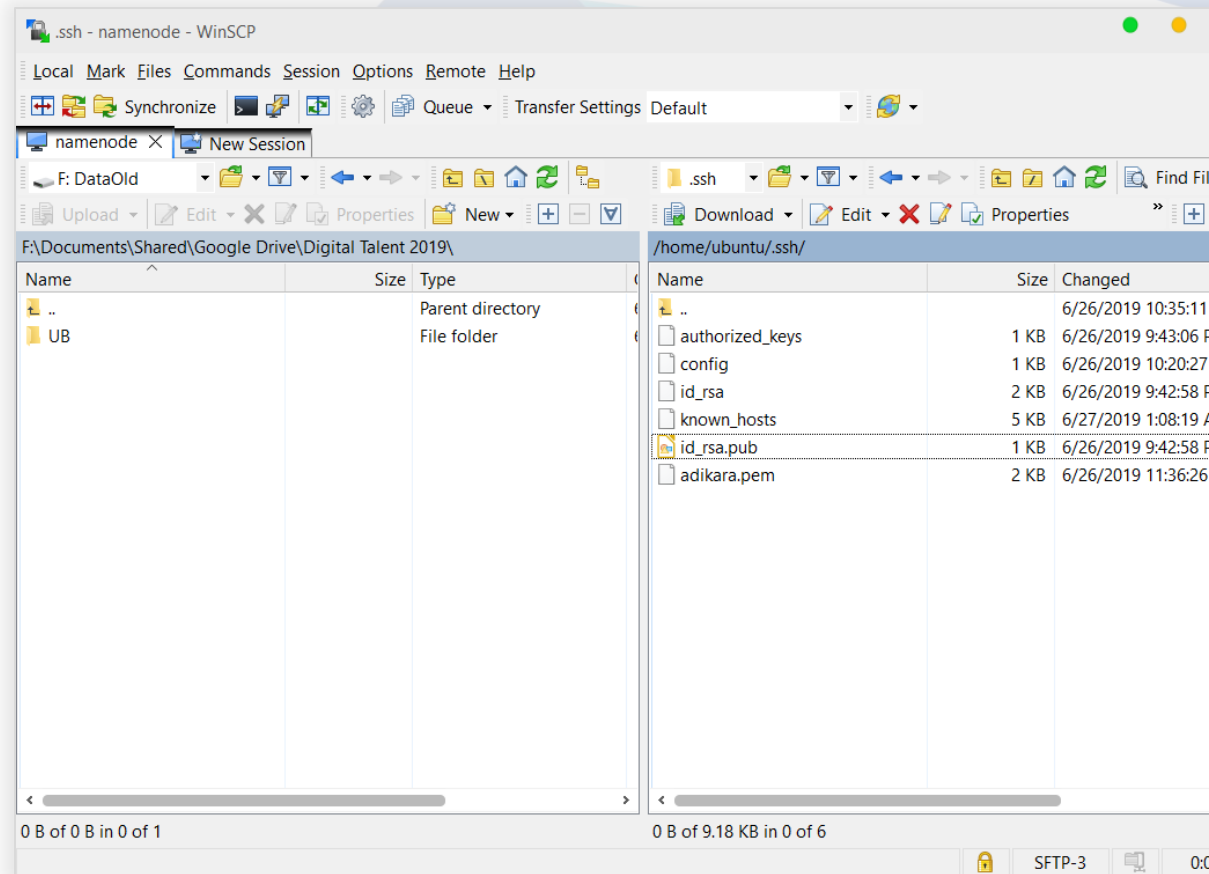
# Buat Koneksi dengan WinSCP (3)

- Klik Login
- Apabila muncul jendela Warning pilih Yes



# Buat Koneksi dengan WinSCP (4)

- Di panel sebelah kanan pilih folder `~/ .ssh`
- Misalnya bila nama `user=ubuntu`, maka masuk ke folder `/home/ubuntu/ .ssh`

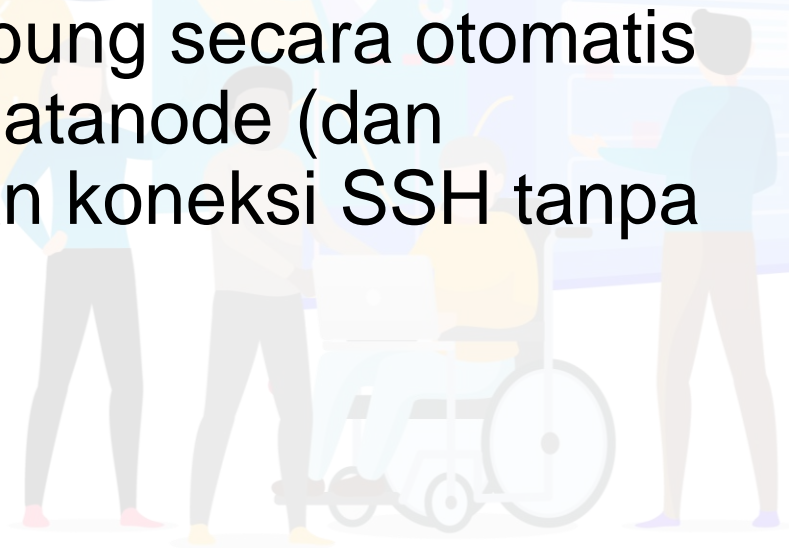




DIGITAL  
TALENT  
SCHOLARSHIP

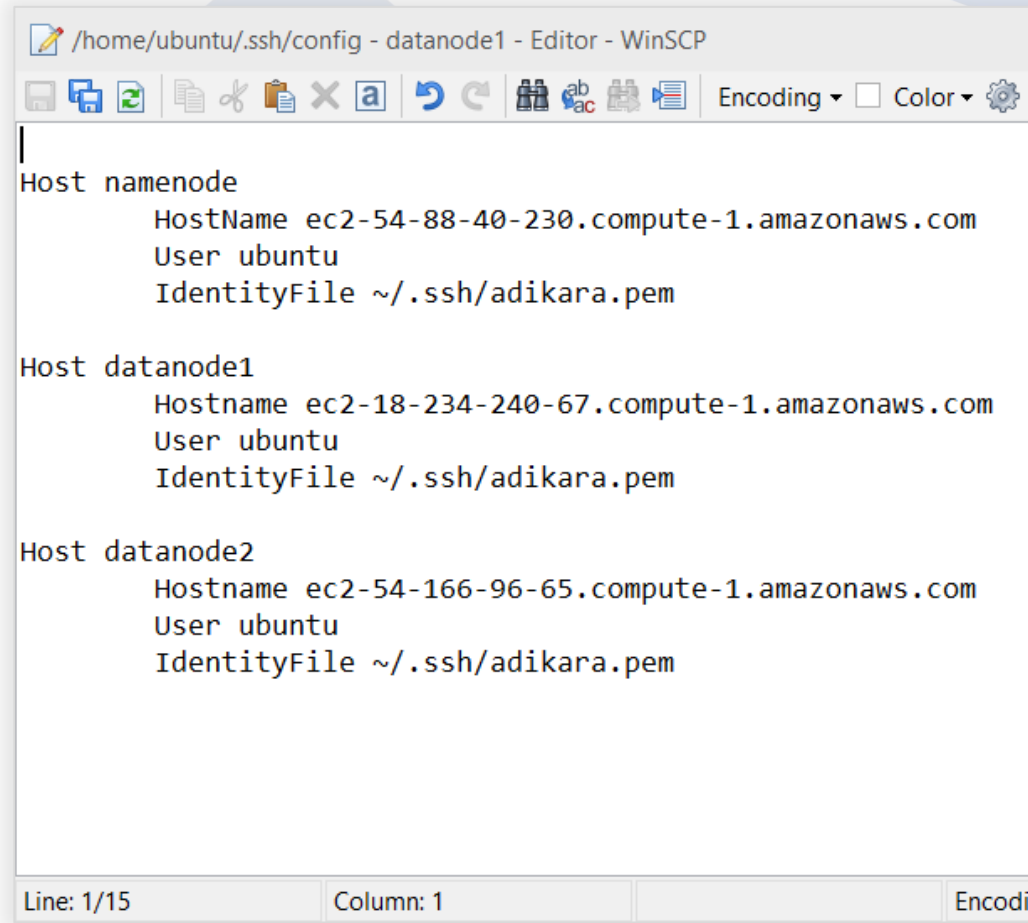
# Setting Password-less SSH (1)

- Melalui WinSCP (atau PuTTY) kita menggunakan koneksi dengan password (melalui fail .pem)
- Supaya Hadoop bisa terhubung secara otomatis dari namenode ke semua datanode (dan sebaliknya) maka diperlukan koneksi SSH tanpa password (*password-less*)



# Setting SSH di Semua Node

- Buat atau ubah fail **config** di `~/.ssh/config`
  - Di panel kanan klik kanan, **New > File**
- Buat entry untuk semua node (**namenode & datanode**)
  - Beri nama **Host** sesuai fungsinya
  - Nama **Hostname** masing-masing host adalah public DNS
  - **User=ubuntu**
  - Tentukan lokasi **IdentityFile (.pem)**



```
/home/ubuntu/.ssh/config - datanode1 - Editor - WinSCP

Host namenode
    HostName ec2-54-88-40-230.compute-1.amazonaws.com
    User ubuntu
    IdentityFile ~/.ssh/adikara.pem

Host datanode1
    Hostname ec2-18-234-240-67.compute-1.amazonaws.com
    User ubuntu
    IdentityFile ~/.ssh/adikara.pem

Host datanode2
    Hostname ec2-54-166-96-65.compute-1.amazonaws.com
    User ubuntu
    IdentityFile ~/.ssh/adikara.pem
```

Line: 1/15      Column: 1      Encodi



DIGITAL  
TALENT  
SCHOLARSHIP

# Kopikan Public Key ke Semua Node

- Login dengan PuTTY ke namenode atau lewat SSH
  - `ssh -i "key_pair.pem" ubuntu@namenode` (atau)
  - `ssh -i "key_pair.pem" ubuntu@ec2-xx-xx-xx`
- **Buat public key dengan `ssh-keygen` di namenode dan kopi ke `authorized_keys` dengan perintah:**
  - `ssh-keygen -f ~/.ssh/id_rsa -t rsa -P ""`
  - `~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys`
- Untuk koneksi password-less maka copy `authorized_keys` ke semua node dengan perintah:
  - `ssh datanode1 'cat >> ~/.ssh/authorized_keys' < ~/.ssh/id_rsa.pub`
  - `ssh datanode2 'cat >> ~/.ssh/authorized_keys' < ~/.ssh/id_rsa.pub`
  - `ssh datanode3 'cat >> ~/.ssh/authorized_keys' < ~/.ssh/id_rsa.pub`

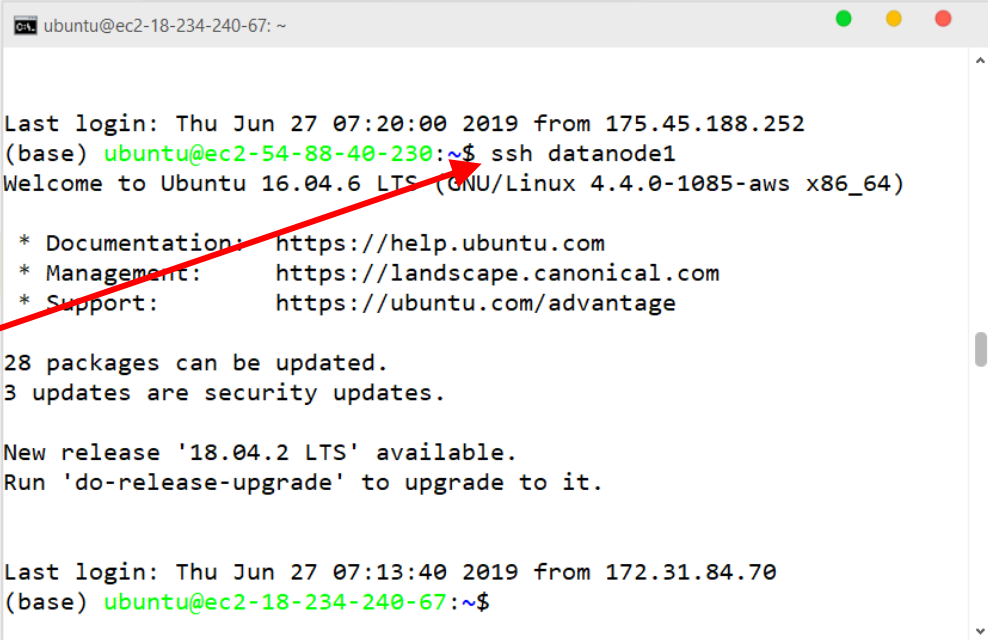




DIGITAL  
TALENT  
SCHOLARSHIP

# Contoh Koneksi SSH ke Semua Node

- Coba koneksi dari master/namenode ke datanode1 dan datanode2
  - ssh datanode1
  - ssh datanode2
- Begitu juga coba koneksi dari datanode1/ datanode2 ke namenode
  - ssh namenode



```
ubuntu@ec2-18-234-240-67: ~  
  
Last login: Thu Jun 27 07:20:00 2019 from 175.45.188.252  
(base) ubuntu@ec2-54-88-40-230:~$ ssh datanode1  
Welcome to Ubuntu 16.04.6 LTS (GNU/Linux 4.4.0-1085-aws x86_64)  
  
* Documentation:  https://help.ubuntu.com  
* Management:    https://landscape.canonical.com  
* Support:        https://ubuntu.com/advantage  
  
28 packages can be updated.  
3 updates are security updates.  
  
New release '18.04.2 LTS' available.  
Run 'do-release-upgrade' to upgrade to it.  
  
Last login: Thu Jun 27 07:13:40 2019 from 172.31.84.70  
(base) ubuntu@ec2-18-234-240-67:~$
```

A red arrow points from the text 'ssh datanode1' in the list to the terminal window.



DIGITAL  
TALENT  
SCHOLARSHIP

# Konfigurasi NameNode dan DataNode

TERBUKA  
UNTUK  
DISABILITAS

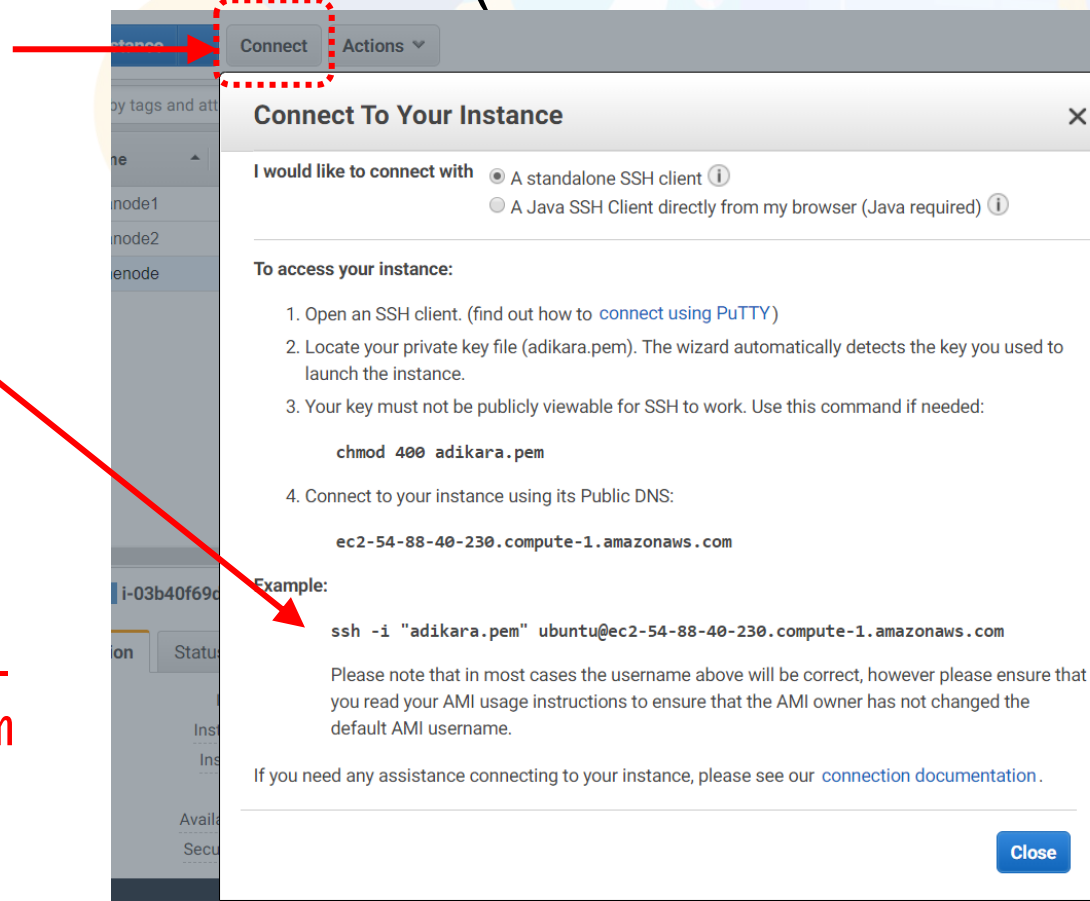
BREAK  
YOUR  
LIMIT



# Menyiapkan Master (Namenode)

- Buat koneksi ke instance yang akan menjadi master melalui PuTTY atau SSH (atau SSH di Bash di WSL)
- Klik tombol Connect untuk mengetahui cara Connect
- Di terminal gunakan perintah ssh, misalnya:

- `ssh -i "nama_ke.pem" ubuntu@ec2-xx-xx-xxx.amazonaws.com`



# Menyiapkan Master (Namenode)

1. Login ke ssh (slide sebelumnya)

2. Ubah nama hostname

- Nilai awalnya biasanya `ip-XX-XX-XX-XX`

- Ganti ke public DNS dengan perintah

```
sudo hostname  
<public_dns>
```

- Misalnya

```
sudo hostname ec2-54-  
88-40-230.compute-  
1.amazonaws.com
```

```
ubuntu@ec2-54-88-40-230: ~  
(base) ubuntu@ec2-54-88-40-230:~$ sudo hostname ec2-54-88-40-230.co  
mpute-1.amazonaws.com  
(base) ubuntu@ec2-54-88-40-230:~$ hostname  
ec2-54-88-40-230.compute-1.amazonaws.com  
(base) ubuntu@ec2-54-88-40-230:~$
```

- Cek perubahan nama host dengan perintah `hostname`

# Menyiapkan Slave (Datanode)

## 1. Login ke datanode, mis

- ssh datanode1

## 2. Ubah nama hostname untuk datanode

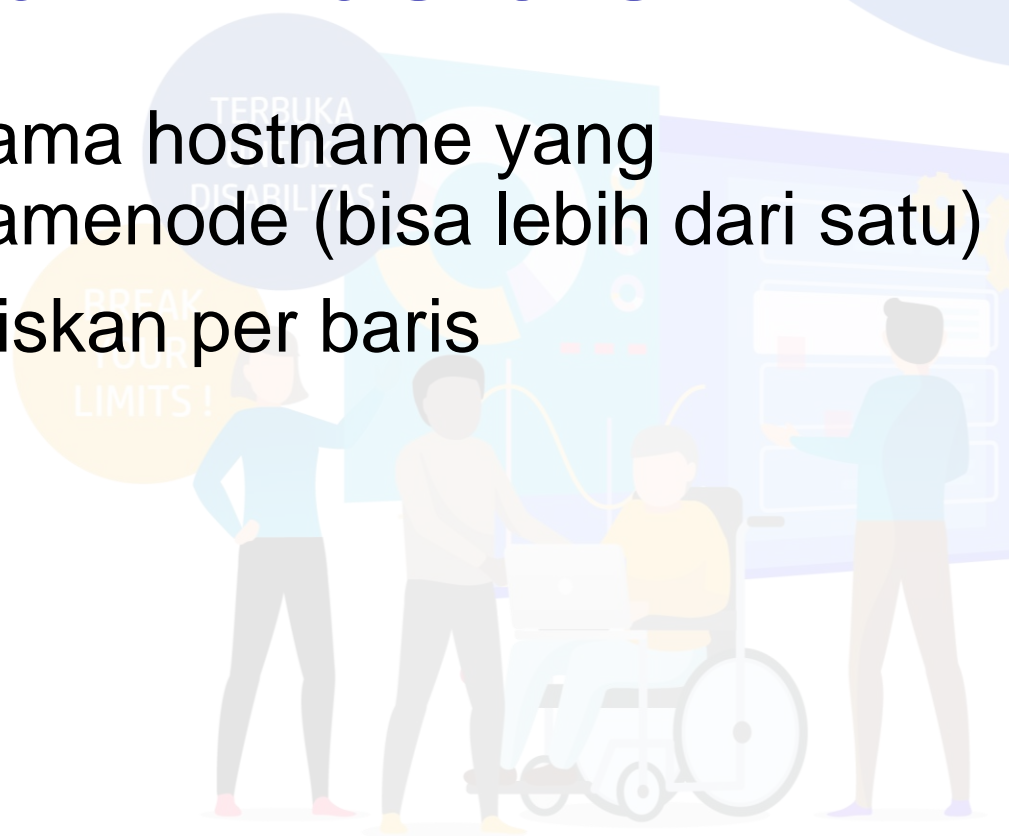
- Nilai awalnya biasanya ip-xx-xx-xx-xx
- Ganti ke public DNS dengan perintah  
    sudo hostname  
    <public\_dns>
- Misalnya  
    sudo hostname ec2-54-88-40-230.compute-1.amazonaws.com

```
ubuntu@ec2-54-88-40-230: ~  
(base) ubuntu@ec2-54-88-40-230:~$ sudo hostname ec2-54-88-40-230.compute-1.amazonaws.com  
(base) ubuntu@ec2-54-88-40-230:~$ hostname  
ec2-54-88-40-230.compute-1.amazonaws.com  
(base) ubuntu@ec2-54-88-40-230:~$
```

- Cek perubahan nama host dengan perintah  
    hostname

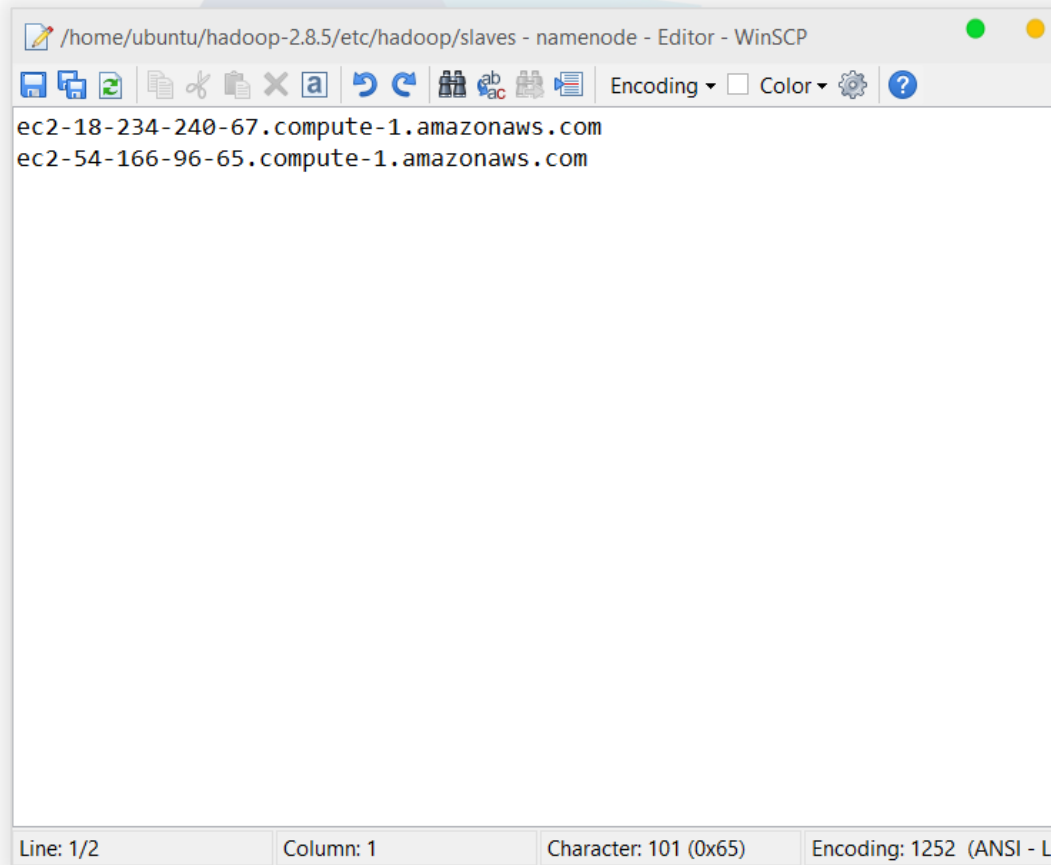
# Konfigurasi Fail "masters"

- File ini berisi nama-nama hostname yang digunakan sebagai namenode (bisa lebih dari satu)
- Nama hostname dituliskan per baris



# Konfigurasi Fail "slaves"

- File ini berisi nama-nama hostname yang digunakan sebagai datanode (bisa lebih dari satu)
- Nama hostname dituliskan per baris



```
/home/ubuntu/hadoop-2.8.5/etc/hadoop/slaves - namenode - Editor - WinSCP  
ec2-18-234-240-67.compute-1.amazonaws.com  
ec2-54-166-96-65.compute-1.amazonaws.com  
  
Line: 1/2      Column: 1      Character: 101 (0x65)      Encoding: 1252 (ANSI - L)
```



DIGITAL  
TALENT  
SCHOLARSHIP

# Konfigurasi Fail "core-site.xml"

- Ganti localhost pada <value> dengan public DNS dari namenode

```
/home/ubuntu/hadoop-2.8.5/etc/hadoop/core-site.xml - namenode - Editor - WinSCP

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

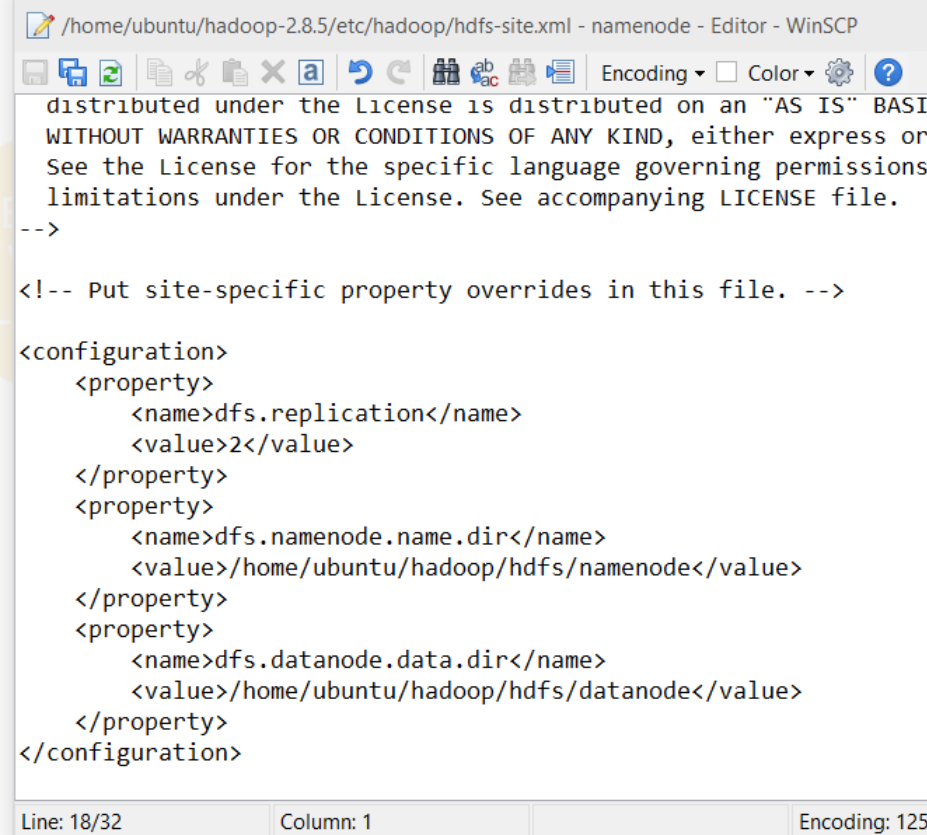
<configuration>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://ec2-54-88-40-230.compute-1.amazonaws.com:9000</value>
  </property>
</configuration>
```

Line: 22/24      Column: 59      Character: 58 (0x3A)      Encoding: 1252 (ANSI - L)



# Konfigurasi Fail "hdfs-site.xml"

- Pada property **dfs.replication**, isi `<value>` sebanyak datanode yang digunakan
  - contoh di sini ada 2, **datanode1** dan **datanode2**
- Pada **dfs.namenode.name.dir**, isi `<value>` dengan lokasi path nantinya metadata dari namenode disimpan
- Pada **dfs.datanode.data.dir**, isi `<value>` dengan lokasi path nantinya data sebenarnya dari datanode disimpan
- Lokasi ini harus ada (buat dulu) dan mendapat permission untuk read write



```
/home/ubuntu/hadoop-2.8.5/etc/hadoop/hdfs-site.xml - namenode - Editor - WinSCP
distributed under the License is distributed on an "AS IS" BASIS
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or
See the License for the specific language governing permissions
limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

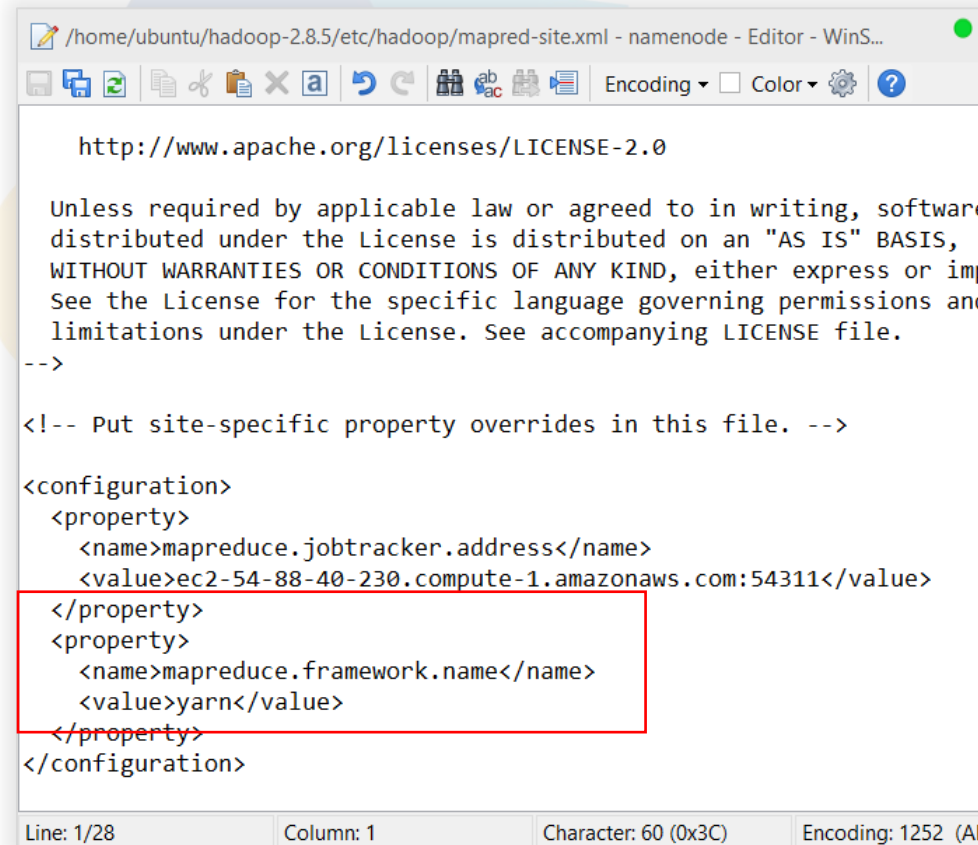
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>2</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>/home/ubuntu/hadoop/hdfs/namenode</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>/home/ubuntu/hadoop/hdfs/datanode</value>
  </property>
</configuration>

Line: 18/32      Column: 1      Encoding: 125
```



# Konfigurasi Fail "mapred-site.xml"

- Untuk Hadoop 1, isi value pada **mapreduce.jobtracker.address**, dengan public DNS dari namenode dan port default 54311
- Untuk > Hadoop 2 tidak perlu diisikan, yang penting mengisi **mapreduce.framework.name** dengan **yarn**
  - Tugas JobTracker digantikan oleh YARN
  - Sebagai gantinya isi di **yarn.resourcemanager.address** di **yarn-site.xml**



```
/home/ubuntu/hadoop-2.8.5/etc/hadoop/mapred-site.xml - namenode - Editor - WinS...
Encoding Color
http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or imp
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->

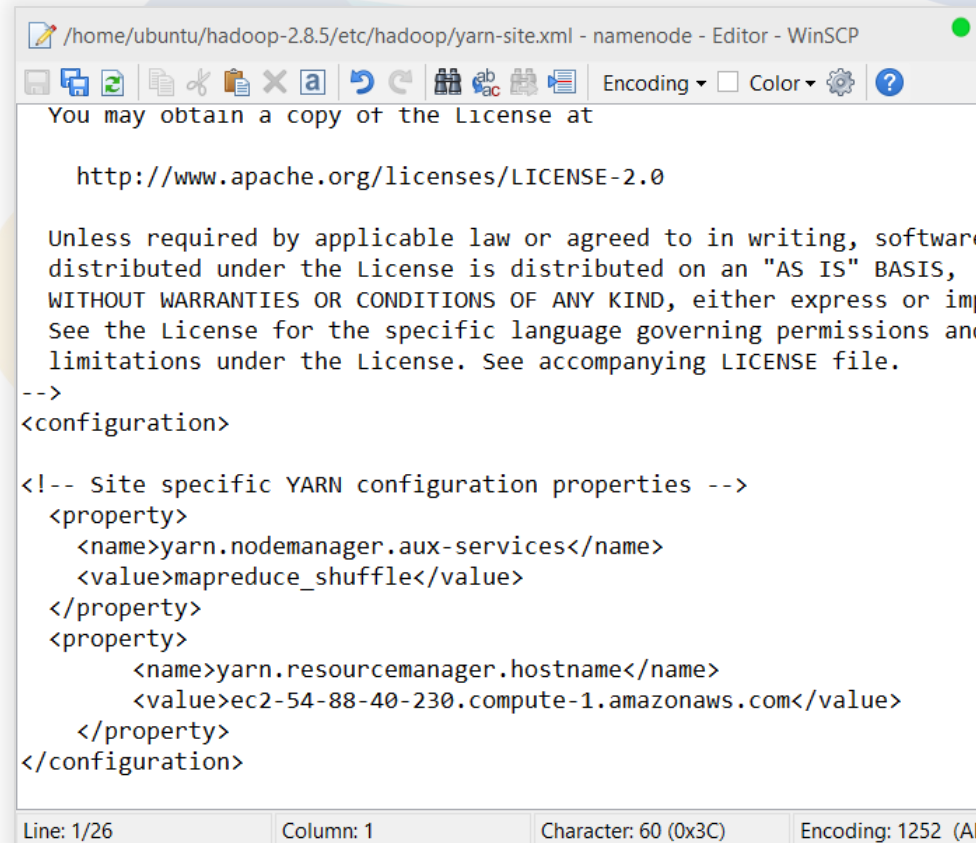
<!-- Put site-specific property overrides in this file. -->

<configuration>
  <property>
    <name>mapreduce.jobtracker.address</name>
    <value>ec2-54-88-40-230.compute-1.amazonaws.com:54311</value>
  </property>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>

Line: 1/28 Column: 1 Character: 60 (0x3C) Encoding: 1252 (A)
```

# Konfigurasi Fail "yarn-site.xml"

- Buat dan isi property `yarn.resourcemanager.hostname` dengan "public DNS" saja dari namenode
  - dengan port default
- *atau*
- Buat dan isi property `yarn.resourcemanager.address` dengan "public DNS:port"

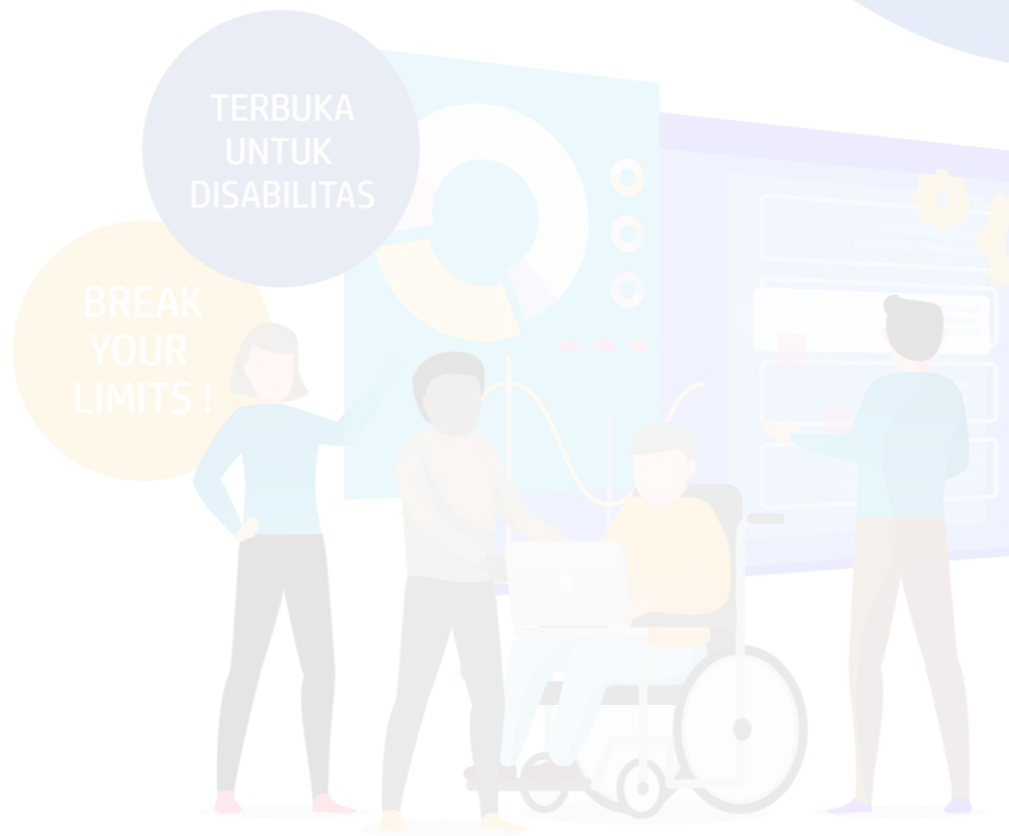


```
/home/ubuntu/hadoop-2.8.5/etc/hadoop/yarn-site.xml - namenode - Editor - WinSCP
You may obtain a copy of the License at
http://www.apache.org/licenses/LICENSE-2.0
Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->
<configuration>
<!-- Site specific YARN configuration properties -->
<property>
  <name>yarn.nodemanager.aux-services</name>
  <value>mapreduce_shuffle</value>
</property>
<property>
  <name>yarn.resourcemanager.hostname</name>
  <value>ec2-54-88-40-230.compute-1.amazonaws.com</value>
</property>
</configuration>
Line: 1/26      Column: 1      Character: 60 (0x3C)      Encoding: 1252 (A)
```



DIGITAL  
TALENT  
SCHOLARSHIP

# Sudo install SCP





DIGITAL  
TALENT  
SCHOLARSHIP

# Kopikan Fail Konfigurasi ke Datanodes

- Melalui PuTTY atau SSH, pindah current directory ke lokasi konfigurasi Hadoop
  - Contoh lokasi: `/home/ubuntu/hadoop-2.8.5/etc/hadoop/`
  - `cd /home/ubuntu/hadoop-2.8.5/etc/hadoop/`
- Kopikan konfigurasi Namenode ke semua Datanode
  - `scp *.xml datanode1:/home/ubuntu/hadoop-2.8.5/etc/hadoop/`
  - `scp *.xml datanode2:/home/ubuntu/hadoop-2.8.5/etc/hadoop/`



DIGITAL  
TALENT  
SCHOLARSHIP

# Menjalankan Hadoop





DIGITAL  
TALENT  
SCHOLARSHIP

# Menjalankan Hadoop

- Masuk ke folder **hadoop/sbin**
  - Kecuali bila sudah masuk di PATH di ~/.bashrc, maka bisa panggil perintah dari path/folder apa saja
- Ketikkan perintah
  - start-dfs.sh
  - start-yarn.sh
  - **atau start-all.sh** (akan deprecated/dihapus di versi baru)
- Bila gagal cek di fail log

TERBUKA  
UNTUK  
DISABILITAS

```
ubuntu@ec2-54-88-40-230: ~  
(base) ubuntu@ec2-54-88-40-230:~$ start-dfs.sh  
Starting namenodes on [ec2-54-88-40-230.compute-1.amazonaws.com]  
  
ec2-54-88-40-230.compute-1.amazonaws.com: starting namenode, log  
ging to /home/ubuntu/hadoop-2.8.5/logs/hadoop-ubuntu-namenode-ec  
  
ubuntu@ec2-54-88-40-230: ~  
(base) ubuntu@ec2-54-88-40-230:~$ start-yarn.sh  
starting yarn daemons  
starting resourcemanager, logging to /home/ubuntu/hadoop-2.8.5/1  
ogs/yarn-ubuntu-resourcemanager-ec2-54-88-40-230.compute-1.amazo  
naws.com.out  
ec2-54-166-96-65.compute-1.amazonaws.com: starting nodemanager,  
logging to /home/ubuntu/hadoop-2.8.5/logs/yarn-ubuntu-nodemange  
r-ec2-54-166-96-65.compute-1.amazonaws.com.out  
ec2-18-234-240-67.compute-1.amazonaws.com: starting nodemanager,  
logging to /home/ubuntu/hadoop-2.8.5/logs/yarn-ubuntu-nodemang  
er-ec2-18-234-240-67.compute-1.amazonaws.com.out  
(base) ubuntu@ec2-54-88-40-230:~$
```

# Cek Proses Hadoop yang Berjalan

- Apabila Hadoop berhasil berjalan di namenode dan datanode maka akan muncul beberapa proses.
- Login ke namenode dan datanode lewat PuTTY/SSH
- Cek dengan **jps**
  - Namenode
  - Datanode

ubuntu@ec2-54-88-40-230: ~

```
(base) ubuntu@ec2-54-88-40-230:~$ jps
1905 SecondaryNameNode
1682 NameNode
2323 Jps
2060 ResourceManager
(base) ubuntu@ec2-54-88-40-230:~$
```

ubuntu@ec2-18-234-240-67: ~

```
(base) ubuntu@ec2-18-234-240-67:~$ jps
1536 DataNode
1680 NodeManager
2210 Jps
(base) ubuntu@ec2-18-234-240-67:~$
```





DIGITAL  
TALENT  
SCHOLARSHIP

# Menghentikan Hadoop

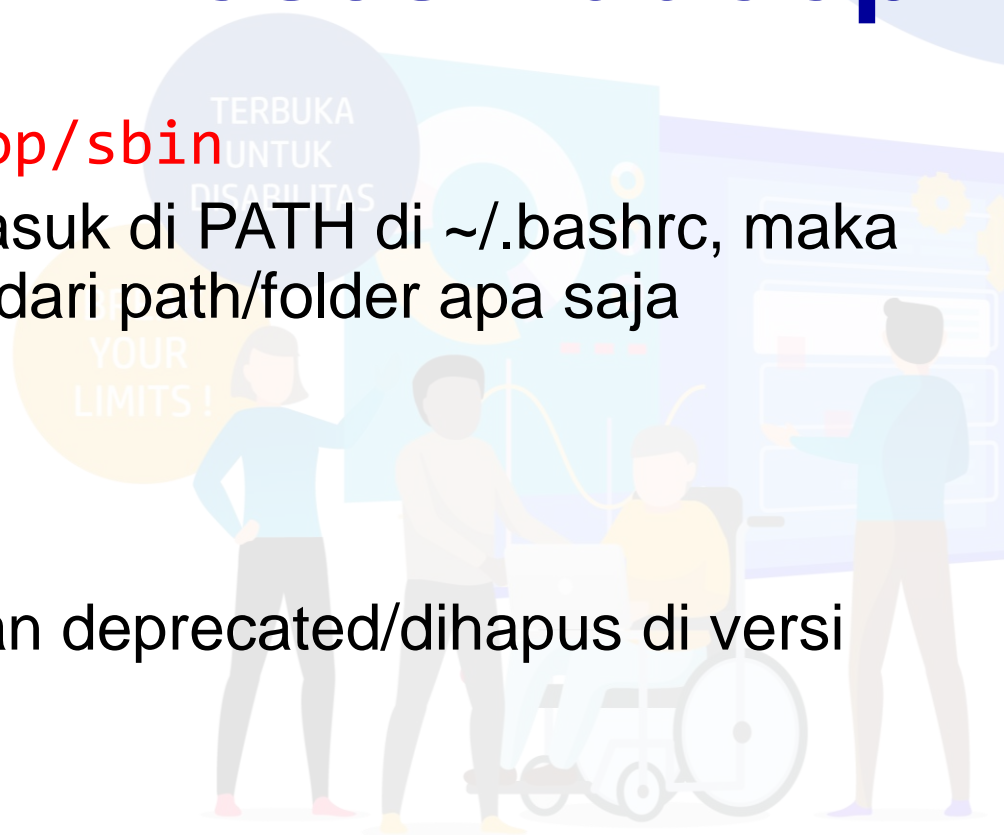
TERBUKA  
UNTUK  
DISABILITAS

BREAK  
YOUR  
WAY



# Menghentikan Proses Hadoop

- Masuk ke folder **hadoop/sbin**
  - Kecuali bila sudah masuk di PATH di ~/.bashrc, maka bisa panggil perintah dari path/folder apa saja
- Ketikkan perintah
  - **stop-dfs.sh**
  - **stop-yarn.sh**
  - **atau stop-all.sh** (akan deprecated/dihapus di versi baru)
- Cek dengan **jps**
  - Bila proses diberhentikan semua, maka tinggal proses **jps** saja yang muncul





DIGITAL  
TALENT  
SCHOLARSHIP

## IKUTI KAMI



DIGITAL  
TALENT  
SCHOLARSHIP

- digitalent.kominfo
- digitalent.kominfo
- DTS\_kominfo
- Digital Talent Scholarship 2019

Pusat Pengembangan Profesi dan Sertifikasi  
Badan Penelitian dan Pengembangan SDM  
Kementerian Komunikasi dan Informatika  
Jl. Medan Merdeka Barat No. 9  
(Gd. Belakang Lt. 4 - 5)  
Jakarta Pusat, 10110



digitalent.kominfo.go.id

