

# MapReduce program fejlesztése és végrehajtása

Rusznák Attila

SZTAKI





ELKH Cloud

# Tartalomjegyzék

1. A fejlesztőkörnyezet beállítása
2. Az alkalmazás tesztelése
3. A JAR fájl build-elése
4. A JAR fájl felmásolása a HDFS-re
5. A MapReduce job monitorozása
6. A YARN monitorozása



# A fejlesztőkörnyezet beállítása

# A Hadoop letöltése

A fejlesztés előtt le kell töltenünk a fejlesztéshez szükséges függőségeket:

- ▶ <https://archive.apache.org/dist/hadoop/common/hadoop-3.3.0/>

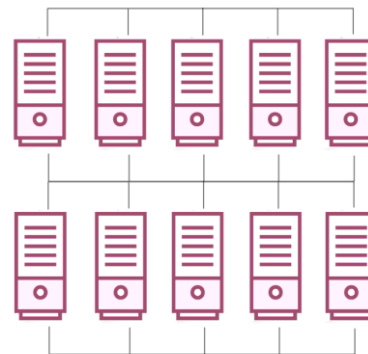
A letöltendő fájl:

- ▶ `hadoop-3.3.0.tar.gz`



Saját számítógépünk  
(fejlesztés, tesztelés)

Éles környezet



Hadoop klaszter az  
ELKH Cloud-on



Index of /dist/hadoop/common/hadoop-3.3.0

Name	Last modified	Size	Description
<a href="#">Parent Directory</a>	-	-	-
<a href="#">CHANGELOG.md</a>	2020-07-15 17:05	376K	
<a href="#">CHANGELOG.md.asc</a>	2020-07-15 17:05	819	
<a href="#">CHANGELOG.md.sha512</a>	2020-07-15 17:05	153	
<a href="#">RELEASENOTES.md</a>	2020-07-15 17:05	26K	
<a href="#">RELEASENOTES.md.asc</a>	2020-07-15 17:05	819	
<a href="#">RELEASENOTES.md.sha512</a>	2020-07-15 17:05	156	
<a href="#">hadoop-3.3.0-aarch64.tar.gz</a>	2020-07-15 17:19	478M	
<a href="#">hadoop-3.3.0-aarch64.tar.gz.asc</a>	2020-07-15 17:19	819	
<a href="#">hadoop-3.3.0-aarch64.tar.gz.sha512</a>	2020-07-15 17:19	168	
<a href="#">hadoop-3.3.0-rat.txt</a>	2020-07-15 17:05	2.0M	
<a href="#">hadoop-3.3.0-rat.txt.asc</a>	2020-07-15 17:05	819	
<a href="#">hadoop-3.3.0-rat.txt.sha512</a>	2020-07-15 17:05	161	
<a href="#">hadoop-3.3.0-site.tar.gz</a>	2020-07-15 17:33	40M	
<a href="#">hadoop-3.3.0-site.tar.gz.asc</a>	2020-07-15 17:33	819	
<a href="#">hadoop-3.3.0-site.tar.gz.sha512</a>	2020-07-15 17:33	165	
<a href="#">hadoop-3.3.0-src.tar.gz</a>	2020-07-15 17:05	32M	
<a href="#">hadoop-3.3.0-src.tar.gz.asc</a>	2020-07-15 17:05	819	
<a href="#">hadoop-3.3.0-src.tar.gz.sha512</a>	2020-07-15 17:05	164	
<a href="#">hadoop-3.3.0.tar.gz</a>	2020-07-15 17:30	478M	
<a href="#">hadoop-3.3.0.tar.gz.asc</a>	2020-07-15 17:30	819	
<a href="#">hadoop-3.3.0.tar.gz.sha512</a>	2020-07-15 17:30	160	

# Az IntelliJ letöltése és telepítése

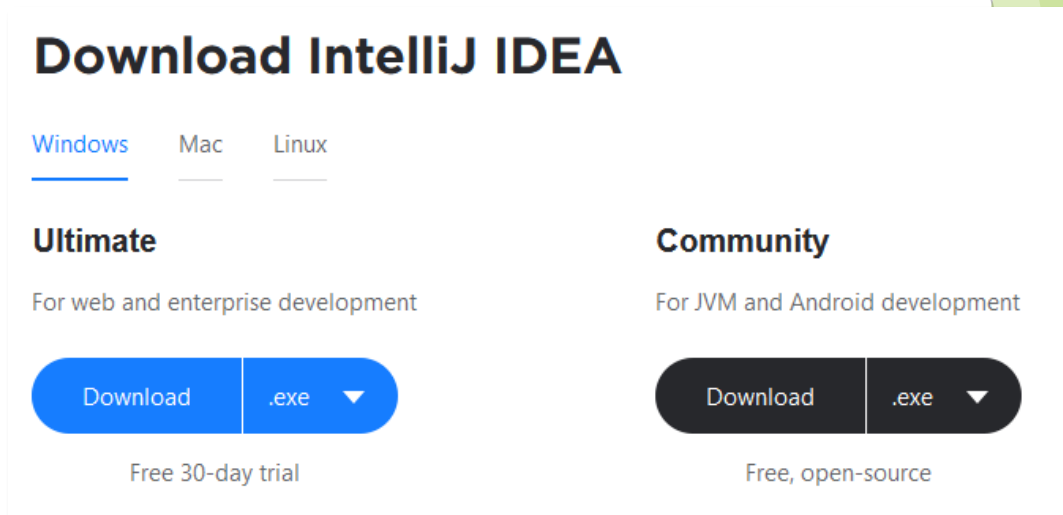

Számos fejlesztőkörnyezet létezik a piacon Java-ra, mi most az IntelliJ-ben nézzük meg hogyan lehet beállítani a fejlesztőkörnyezetet.

Az alábbi címen tudjuk letölteni a fejlesztőkörnyezetet:

- ▶ <https://www.jetbrains.com/idea/download/>

Kétféle kiadása van:

- ▶ Ultimate
- ▶ Community (nyílt forráskódú)



**Download IntelliJ IDEA**

[Windows](#) [Mac](#) [Linux](#)

**Ultimate**  
For web and enterprise development

[Download](#) [.exe](#) ▼

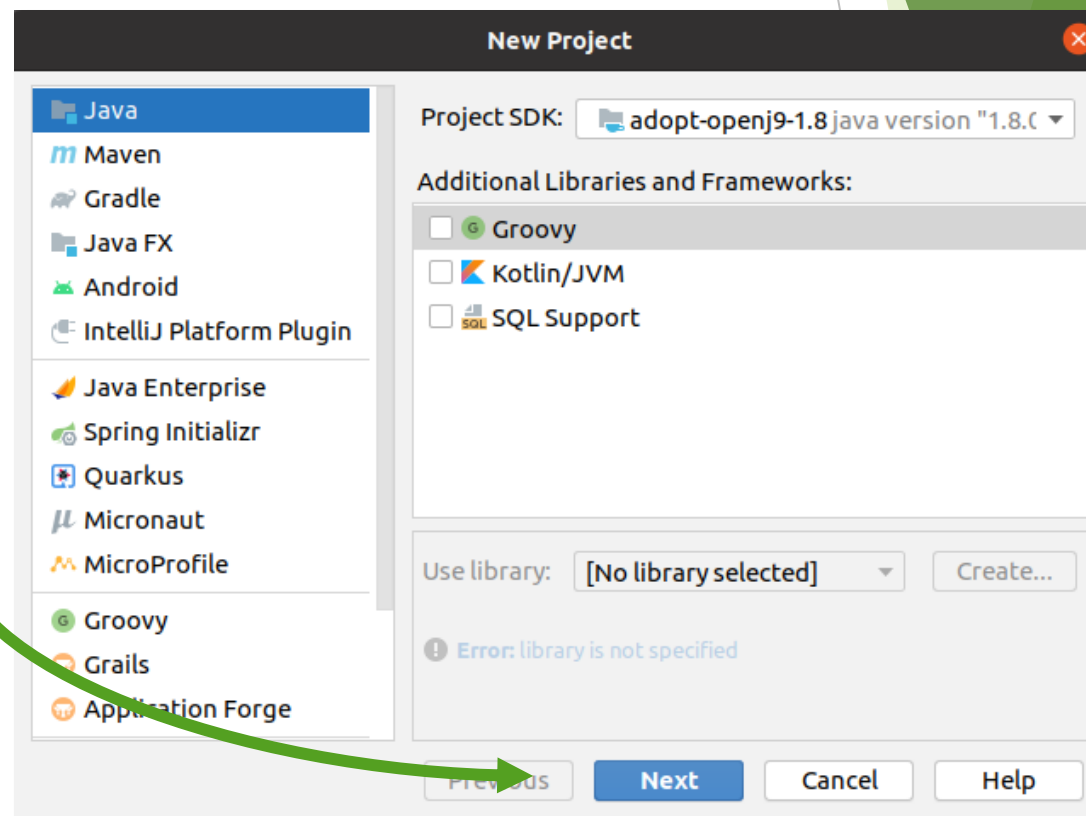
Free 30-day trial

**Community**  
For JVM and Android development

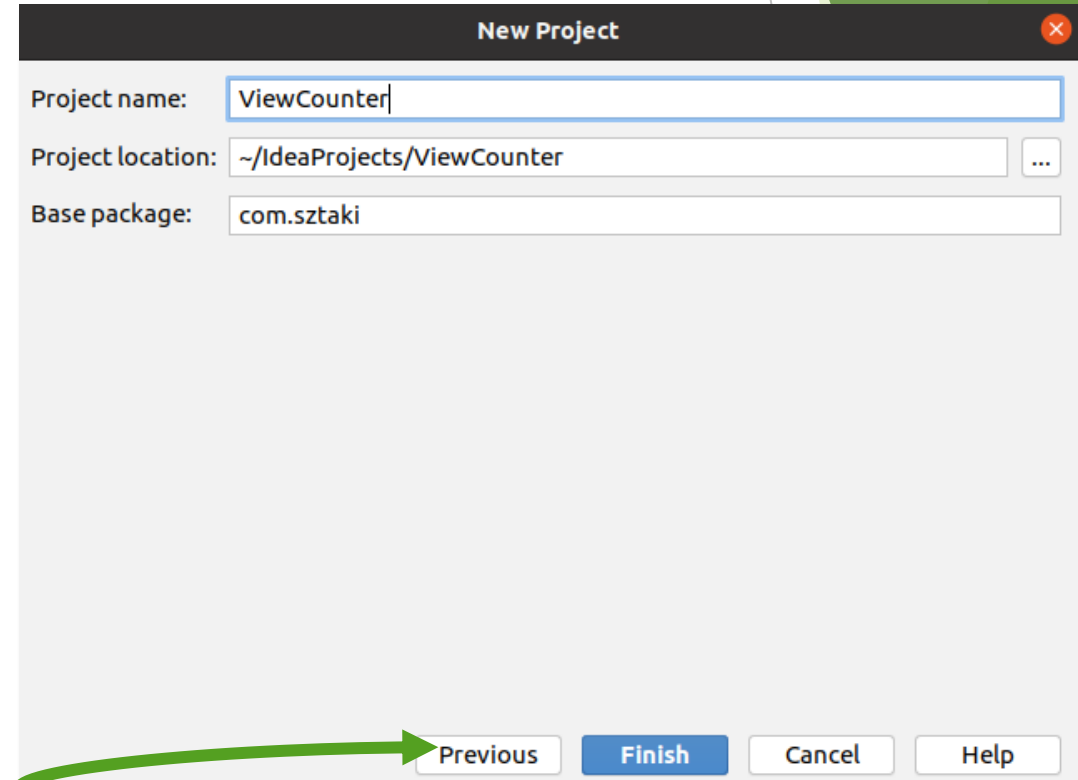
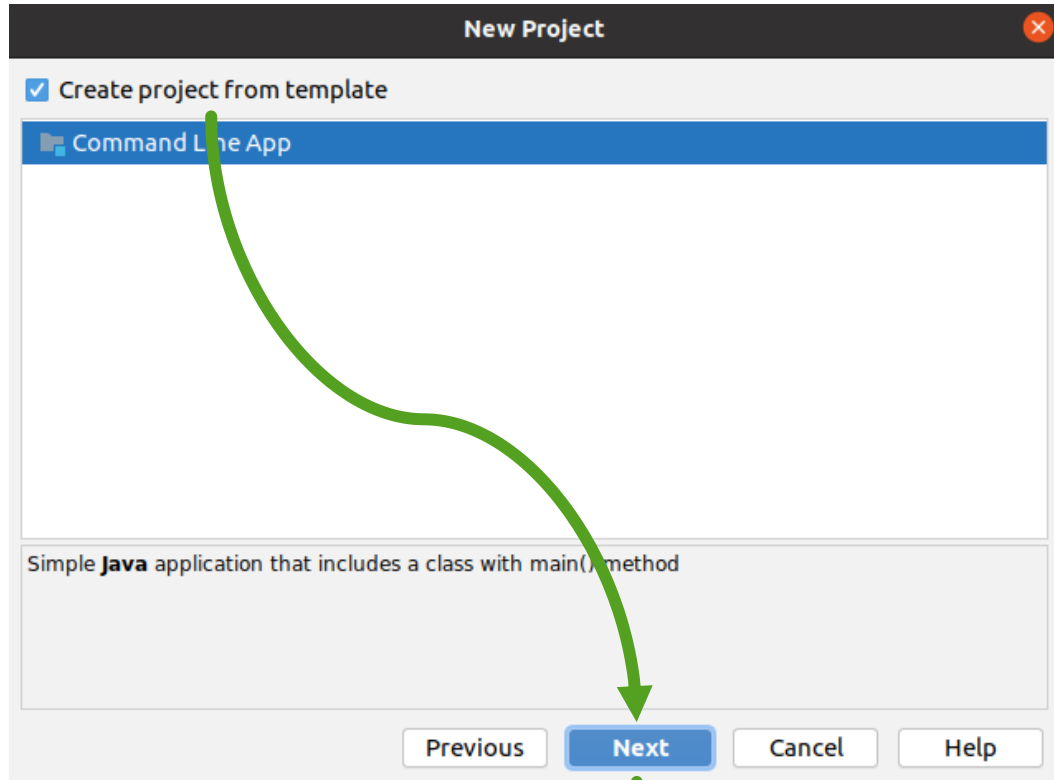
[Download](#) [.exe](#) ▼

Free, open-source

# Java konzolos alkalmazás létrehozása



# Java konzolos alkalmazás létrehozása



# Létrejön az üres projekt template



The screenshot shows an IDE window for a project named 'ViewCounter'. The left sidebar displays the project structure with the following items:

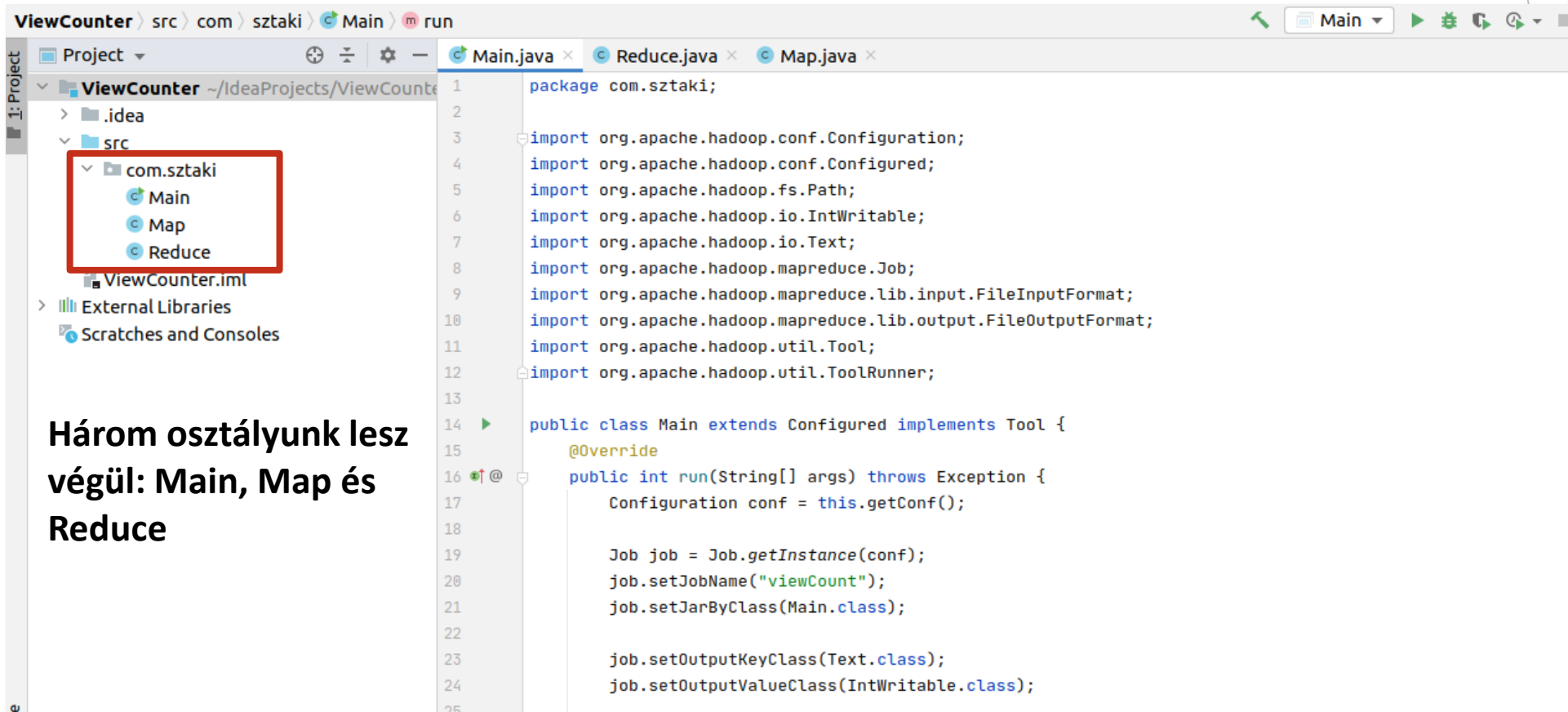
- Project
- ViewCounter ~/IdeaProjects/ViewCounter
  - .idea
  - src
    - com.sztaki
  - ViewCounter.iml
- External Libraries
- Scratches and Consoles

The main editor area shows the 'Main.java' file with the following code:

```
1 package com.sztaki;
2
3 public class Main {
4
5     public static void main(String[] args) {
6
7     }
8 }
9
```



# Megírjuk az alkalmazást

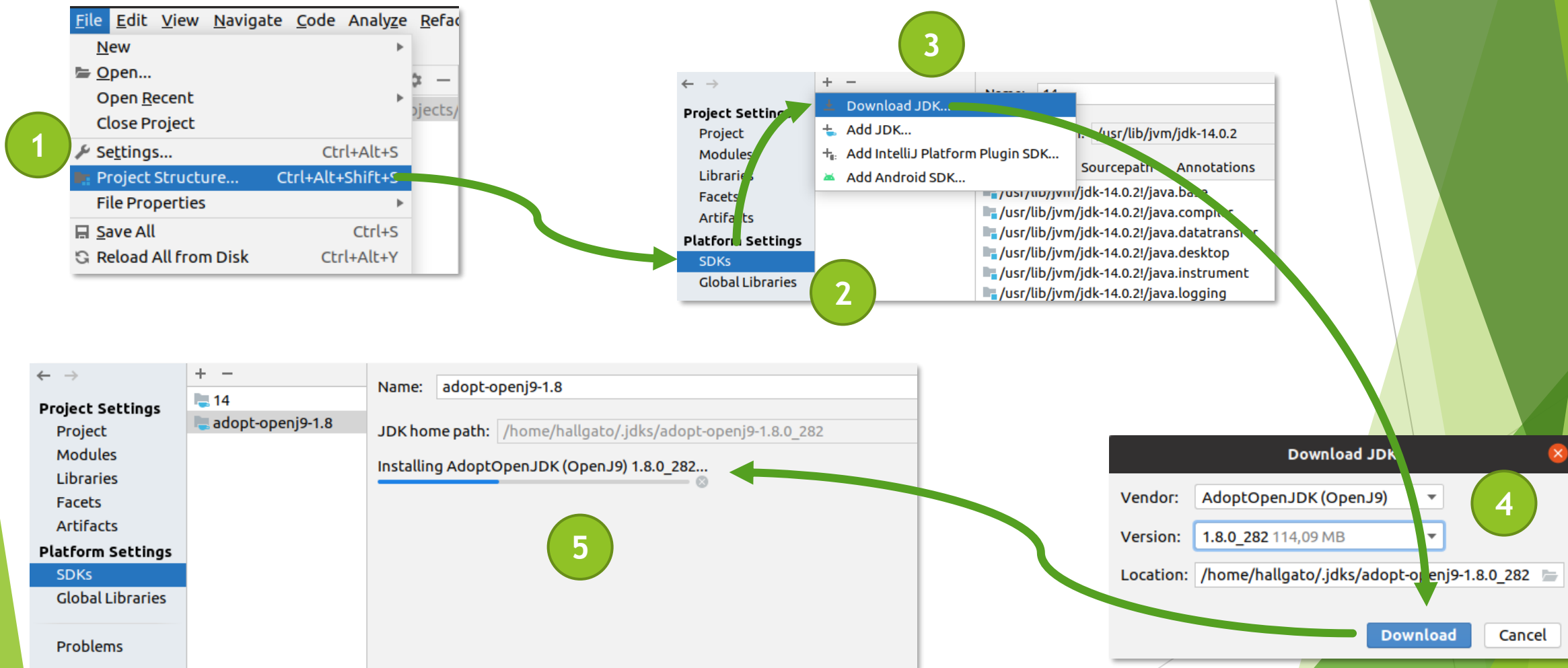


The screenshot shows an IDE window for a project named 'ViewCounter'. The project structure on the left shows a package 'com.sztaki' containing three classes: 'Main', 'Map', and 'Reduce'. The main editor displays the code for 'Main.java', which is a Hadoop MapReduce application. The code includes imports for Hadoop configuration and file system classes, and a 'Main' class that implements the 'Tool' interface. The 'run' method of 'Main' sets up a Hadoop job named 'viewCount' and configures it to output text files.

```
1 package com.sztaki;
2
3 import org.apache.hadoop.conf.Configuration;
4 import org.apache.hadoop.conf.Configured;
5 import org.apache.hadoop.fs.Path;
6 import org.apache.hadoop.io.IntWritable;
7 import org.apache.hadoop.io.Text;
8 import org.apache.hadoop.mapreduce.Job;
9 import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
10 import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
11 import org.apache.hadoop.util.Tool;
12 import org.apache.hadoop.util.ToolRunner;
13
14 public class Main extends Configured implements Tool {
15     @Override
16     public int run(String[] args) throws Exception {
17         Configuration conf = this.getConf();
18
19         Job job = Job.getInstance(conf);
20         job.setJobName("viewCount");
21         job.setJarByClass(Main.class);
22
23         job.setOutputKeyClass(Text.class);
24         job.setOutputValueClass(IntWritable.class);
25     }
26 }
```

Három osztályunk lesz  
végül: Main, Map és  
Reduce

# Az 1.8-as JDK letöltése



1. File > Project Structure... (Ctrl+Alt+Shift+S)

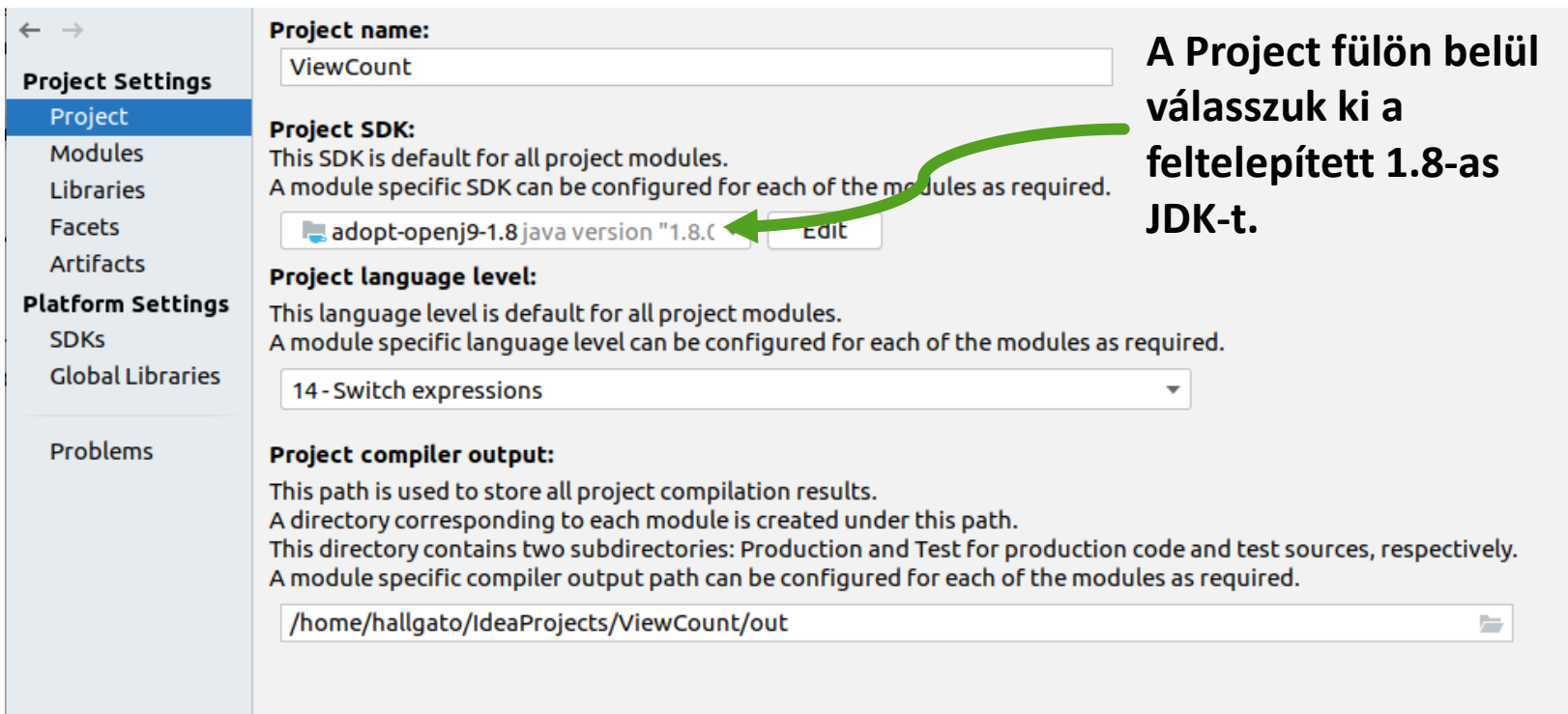
2. Project Settings > SDKs

3. Download JDK...

4. Download JDK dialog box:  
Vendor: AdoptOpenJDK (OpenJ9)  
Version: 1.8.0\_282 114,09 MB  
Location: /home/hallgato/.jdk/adopt-openj9-1.8.0\_282

5. Installing AdoptOpenJDK (OpenJ9) 1.8.0\_282...

# Állítsuk át a projekt szinten is a JDK-t



The screenshot shows the 'Project SDK' configuration in IntelliJ IDEA. The 'Project SDK' section is highlighted, and a green arrow points to the 'adopt-openj9-1.8 java version "1.8.0'" entry. To the right of the screenshot, there is a text box explaining the action.

**Project name:** ViewCount

**Project SDK:** This SDK is default for all project modules. A module specific SDK can be configured for each of the modules as required.

adopt-openj9-1.8 java version "1.8.0"

**Project language level:** This language level is default for all project modules. A module specific language level can be configured for each of the modules as required.

14 - Switch expressions

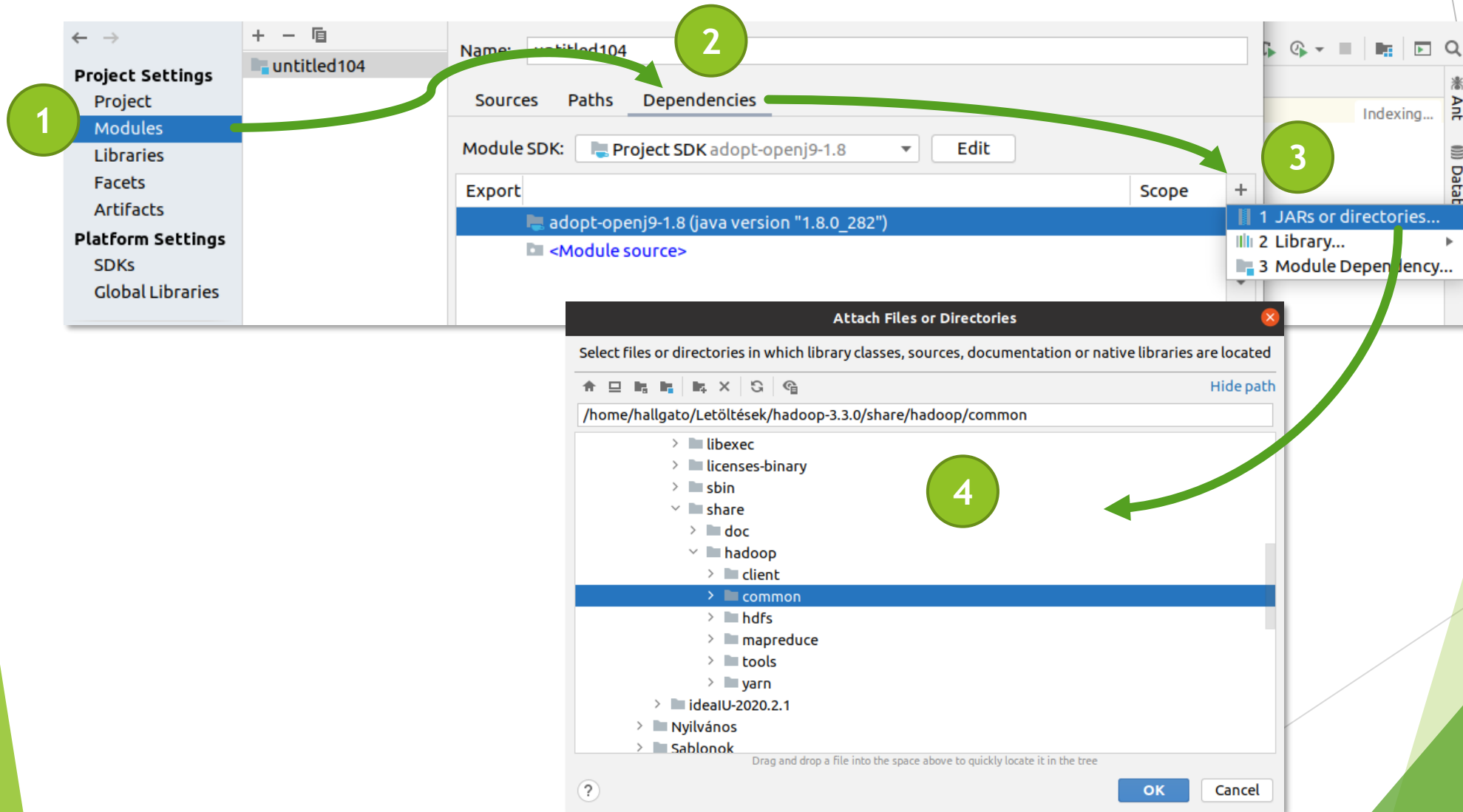
**Project compiler output:** This path is used to store all project compilation results. A directory corresponding to each module is created under this path. This directory contains two subdirectories: Production and Test for production code and test sources, respectively. A module specific compiler output path can be configured for each of the modules as required.

/home/hallgato/IdeaProjects/ViewCount/out

**A Project fülön belül  
válasszuk ki a  
feltelepített 1.8-as  
JDK-t.**

**Ha újabb Java verzióban fejlesztünk, mint ami a Hadoop leíróban található, akkor nem fogjuk tudni futtatni a MapReduce alkalmazást.**

# A Hadoop függőségeinek betöltése



1

2

3

4

Project Settings

- Project
- Modules**
- Libraries
- Facets
- Artifacts

Platform Settings

- SDKs
- Global Libraries

Name: untitled104

Sources Paths **Dependencies**

Module SDK: Project SDK adopt-openj9-1.8 Edit

Export Scope +

- adopt-openj9-1.8 (java version "1.8.0\_282")
- <Module source>

1 JARs or directories...

2 Library...

3 Module Dependency...

Attach Files or Directories

Select files or directories in which library classes, sources, documentation or native libraries are located

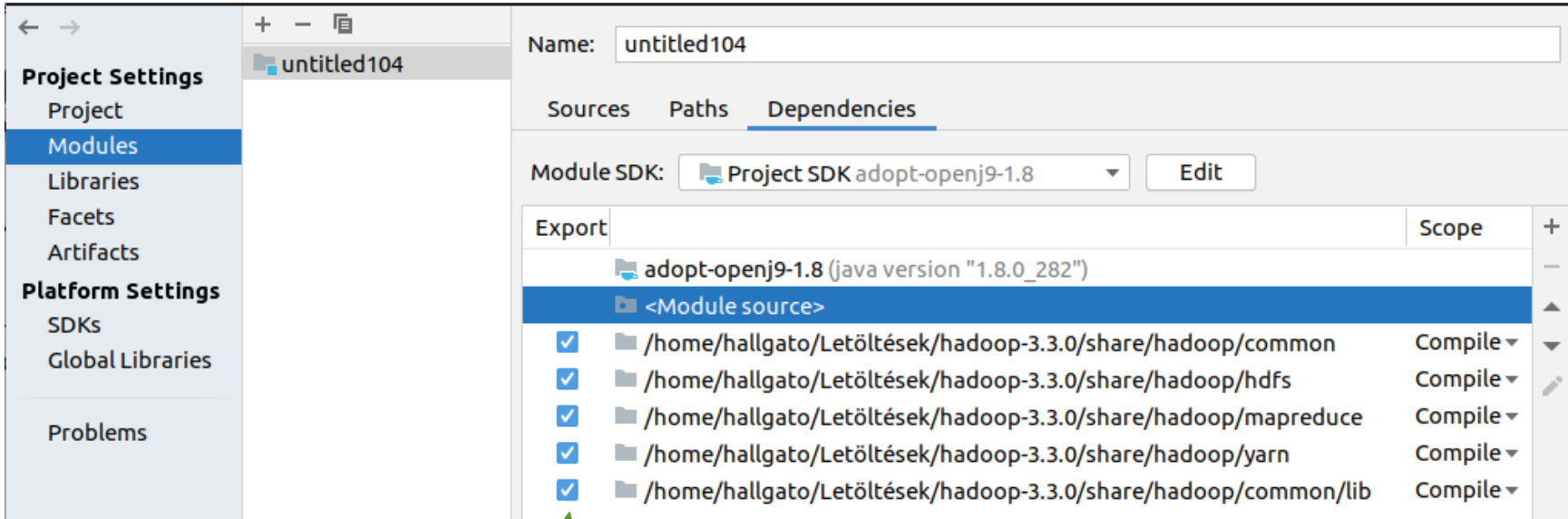
/home/hallgato/Letöltések/hadoop-3.3.0/share/hadoop/common

- libexec
- licenses-binary
- sbin
- share
  - doc
  - hadoop
    - client
    - common**
    - hdfs
    - mapreduce
    - tools
    - yarn
- idealU-2020.2.1
- Nyilvános
- Sablonok

Drag and drop a file into the space above to quickly locate it in the tree

OK Cancel

# A Hadoop függőségeinek betöltése



The screenshot shows the 'Dependencies' tab in the 'Project SDK adopt-openj9-1.8' settings. The 'Export' column is checked for the following paths, and the 'Scope' is set to 'Compile':

Export	Scope
adopt-openj9-1.8 (java version "1.8.0_282")	
<Module source>	
<input checked="" type="checkbox"/> /home/hallgato/Letöltések/hadoop-3.3.0/share/hadoop/common	Compile
<input checked="" type="checkbox"/> /home/hallgato/Letöltések/hadoop-3.3.0/share/hadoop/hdfs	Compile
<input checked="" type="checkbox"/> /home/hallgato/Letöltések/hadoop-3.3.0/share/hadoop/mapreduce	Compile
<input checked="" type="checkbox"/> /home/hallgato/Letöltések/hadoop-3.3.0/share/hadoop/yarn	Compile
<input checked="" type="checkbox"/> /home/hallgato/Letöltések/hadoop-3.3.0/share/hadoop/common/lib	Compile

Jelöljük be a betöltött modulok melletti jelölőnégyzeteket.

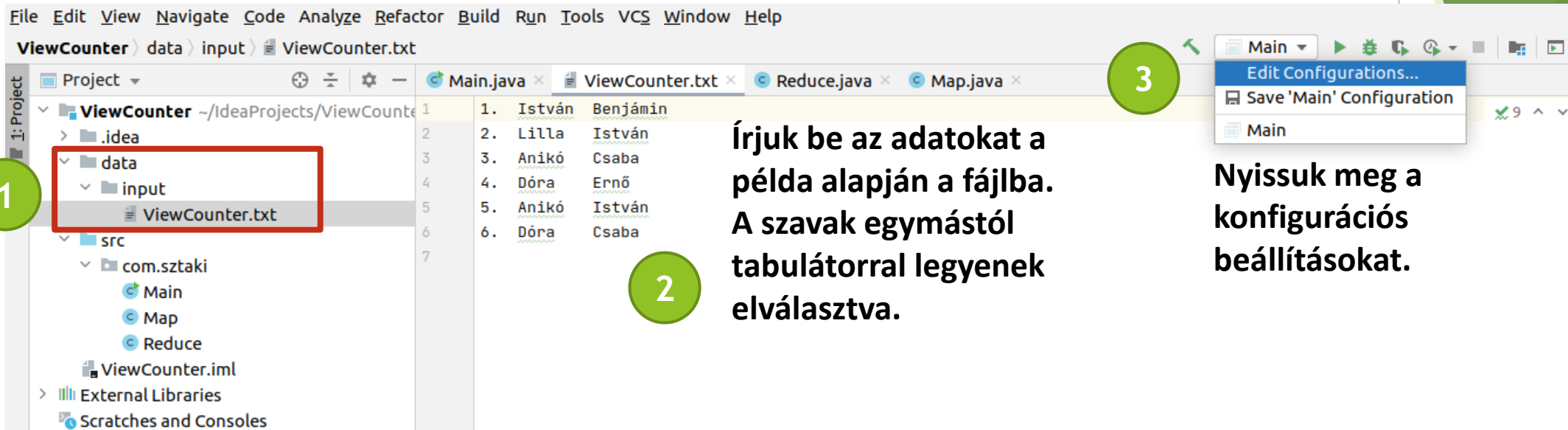
**Az alábbi modulokat kell betölteni:**

- hadoop-3.3.0/share/hadoop/common
- hadoop-3.3.0/share/hadoop/hdfs
- hadoop-3.3.0/share/hadoop/mapreduce
- hadoop-3.3.0/share/hadoop/yarn
- hadoop-3.3.0/share/hadoop/common/lib



# Az alkalmazás tesztelése

# Az alkalmazás tesztelése



1

1.	István	Benjámín
2.	Lilla	István
3.	Anikó	Csaba
4.	Dóra	Ernő
5.	Anikó	István
6.	Dóra	Csaba

2

3

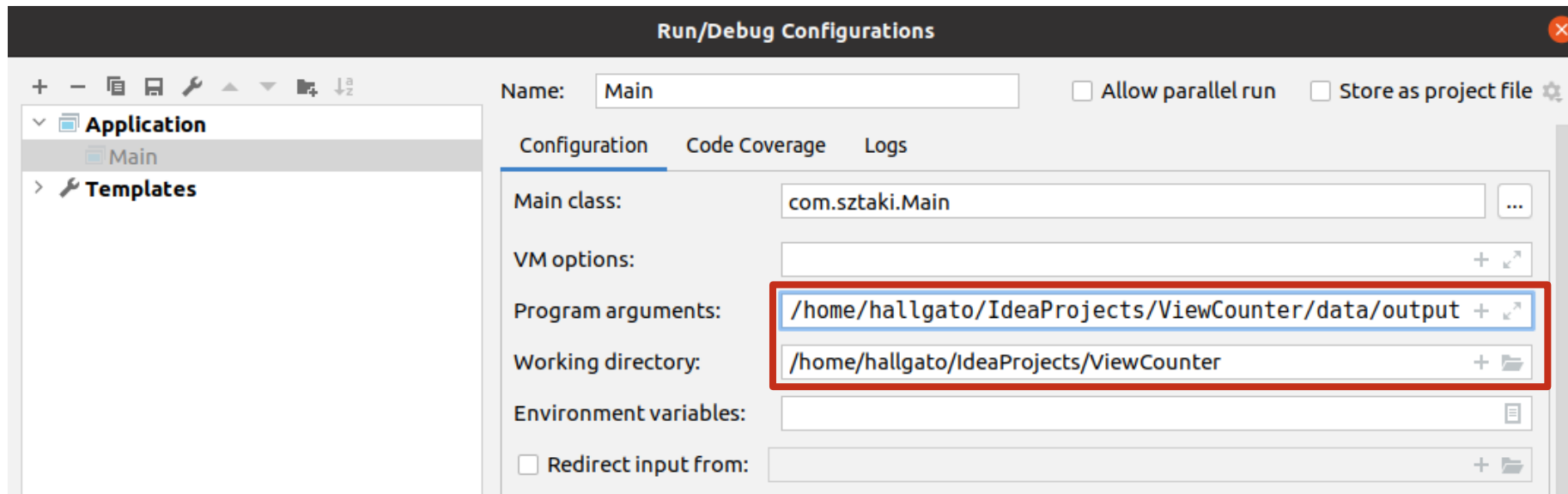
Edit Configurations...  
Save 'Main' Configuration  
Main

Írjuk be az adatokat a példa alapján a fájlba. A szavak egymástól tabulátorral legyenek elválasztva.

Nyissuk meg a konfigurációs beállításokat.

Hozzuk létre a data/input/ útvonalat, majd hozzunk létre benne egy üres ViewCounter.txt üres fájlt.

# Az alkalmazás tesztelése



A Working directory alapján adjuk meg a programnak a két szükséges bemeneti és kimeneti argumentumot. A példában a hallgató felhasználóval vagyunk belépve, tehát:

Bemeneti könyvtár: `/home/hallgato/IdeaProjects/ViewCounter/data/input`

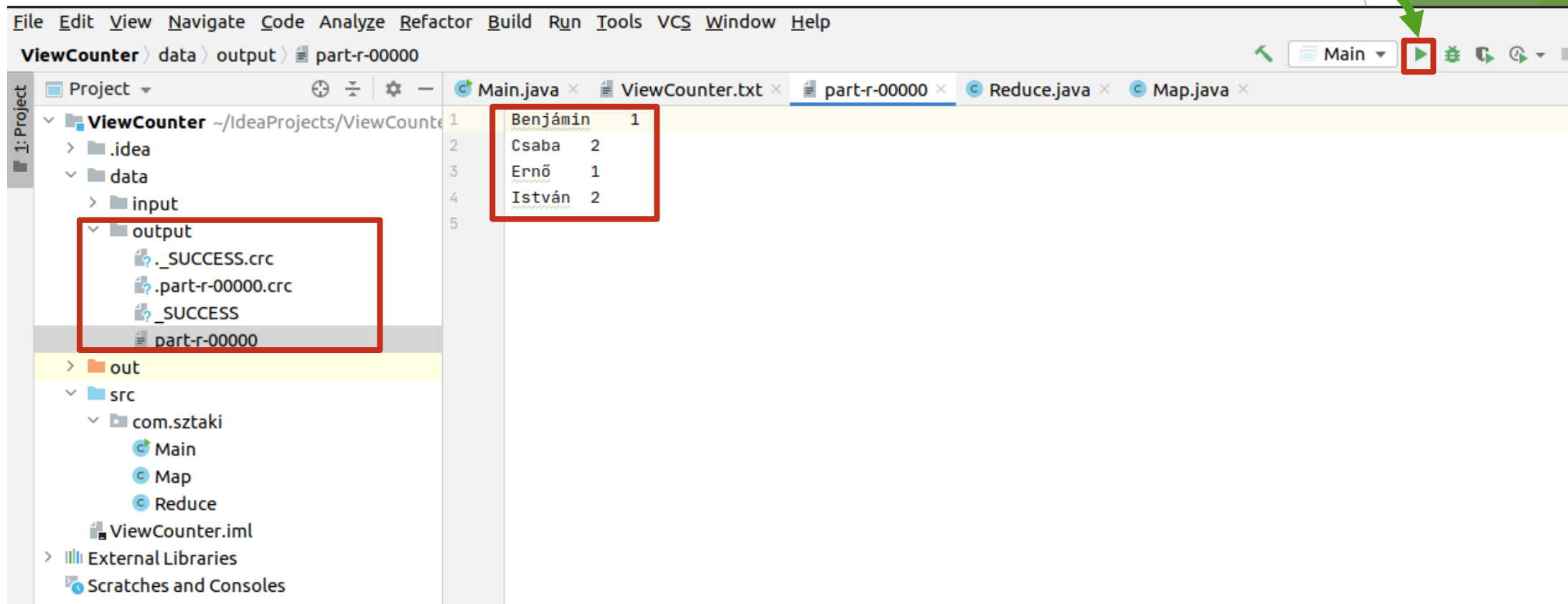
Kimeneti könyvtár: `/home/hallgato/IdeaProjects/ViewCounter/data/output`

**Az output mappát ne hozzuk létre különben nem fogjuk tudni futtatni a programot!**



# A futtatás eredménye

Futtassuk le a programot.



The screenshot shows an IDE window with the following elements:

- Run Button:** A green play icon in the top right toolbar, highlighted with a red box and a green arrow pointing to it from the text "Futtassuk le a programot."
- File Explorer:** Located on the left, it shows the project structure. The 'output' directory is expanded, showing files: `._SUCCESS.crc`, `.part-r-00000.crc`, `._SUCCESS`, and `part-r-00000`. This area is also highlighted with a red box.
- Code Editor:** The main window displays the output of the program:
 

1	Benjámín	1
2	Csaba	2
3	Ernő	1
4	István	2
5		

A program elkészíti az output mappát, majd lesz benne egy part-r-00000 fájl. Azt megnyitva láthatjuk a MapReduce eredményét.



ELKH Cloud

# A JAR fájl build-elése



ELKH Cloud

# Mi az a JAR?

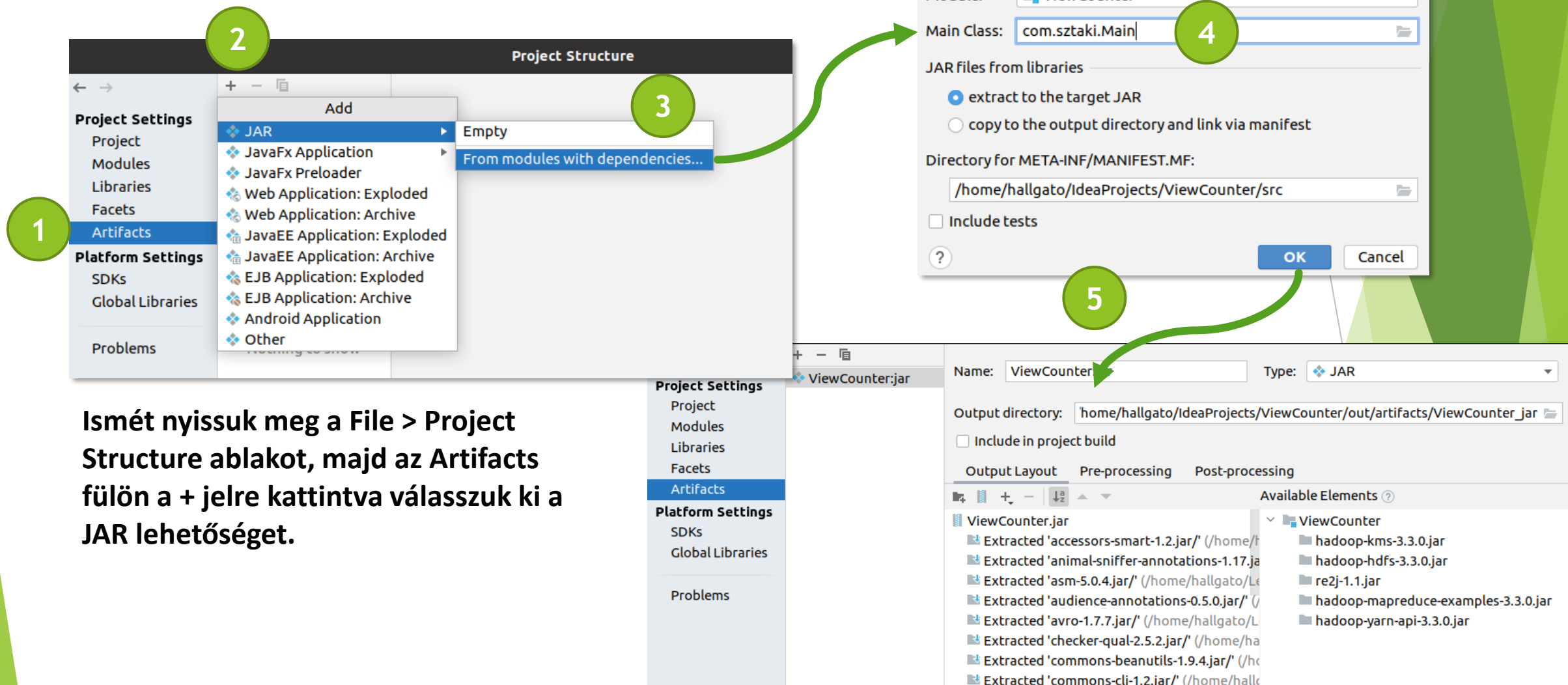
A **JAR** fájl olyan ZIP fájlformátumú fájl

- ▶ Java osztályokat és hozzájuk tartozó metaadatokat tartalmaz
- ▶ A teljes alkalmazás így egyetlen fájlba kerül sűritésre
- ▶ Maga a Java alkalmazás végrehajthatóvá válik



# A JAR fájl létrehozása

Válasszuk ki a Main class-t.



The image illustrates the process of creating a JAR file in IntelliJ IDEA through five numbered steps:

1. Open the **Project Structure** window and select the **Artifacts** tab.
2. Click the **+** button to add a new artifact.
3. In the **Add** dialog, select **JAR** and then **From modules with dependencies...**.
4. In the **Create JAR from Modules** dialog, select the **ViewCounter** module and specify **com.sztaki.Main** as the **Main Class**.
5. Click **OK** to create the JAR.

The final screenshot shows the **ViewCounter.jar** artifact in the **Artifacts** view, with its **Output directory** set to `home/hallgato/IdeaProjects/ViewCounter/out/artifacts/ViewCounter_jar`. The **Available Elements** list includes the **ViewCounter** module and various dependencies such as `hadoop-kms-3.3.0.jar`, `hadoop-hdfs-3.3.0.jar`, `re2j-1.1.jar`, `hadoop-mapreduce-examples-3.3.0.jar`, and `hadoop-yarn-api-3.3.0.jar`.

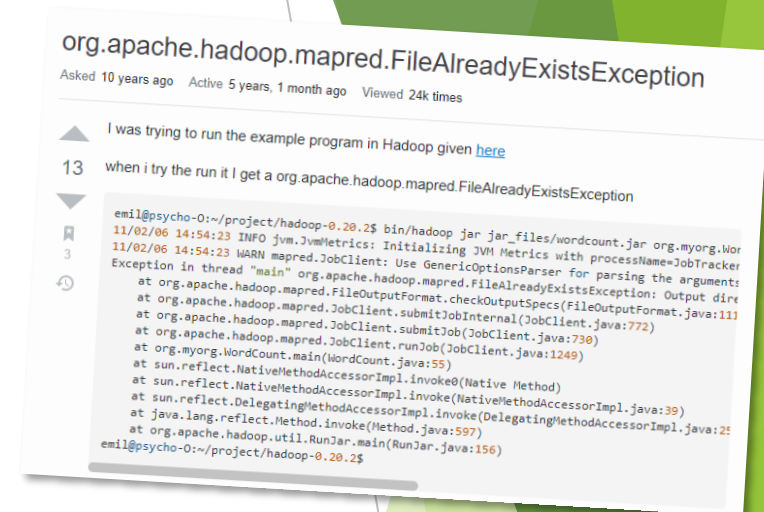
Ismét nyissuk meg a **File > Project Structure** ablakot, majd az **Artifacts** fülön a **+** jelre kattintva válasszuk ki a **JAR** lehetőséget.

Betöltést követően megjelennek a **JAR** fájlok.

# A JAR fájl build-elése előtt

Néhány fontos dolog a fejlesztéssel kapcsolatban:

- ▶ A JAR generálása előtt a **Main osztályban** figyeljünk az args tömbre:
  - ▶ Path inputFilePath = new Path(args[1]);
  - ▶ Path outputFilePath = new Path(args[2]);
  - ▶ Ne 0. és 1. indexet használjunk, ez ugyanis hibát fog okozni!



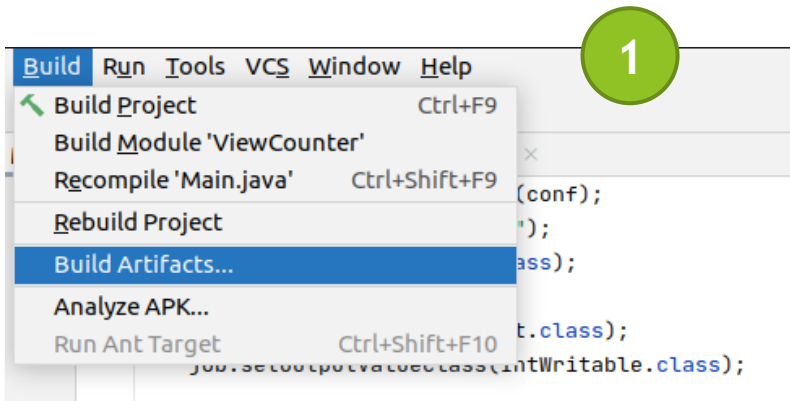
```

29     Path inputFilePath = new Path(args[1]);
30     Path outputFilePath = new Path(args[2]);
31     FileInputFormat.addInputPath(job, inputFilePath);
32     FileOutputFormat.setOutputPath(job, outputFilePath);
33
34     return job.waitForCompletion(verbose: true) ? 0 : 1;
35 }
36
37
38 public static void main(String[] args) throws Exception {
39     int exitCode = ToolRunner.run(new Main(), args);
40     System.exit(exitCode);
41 }

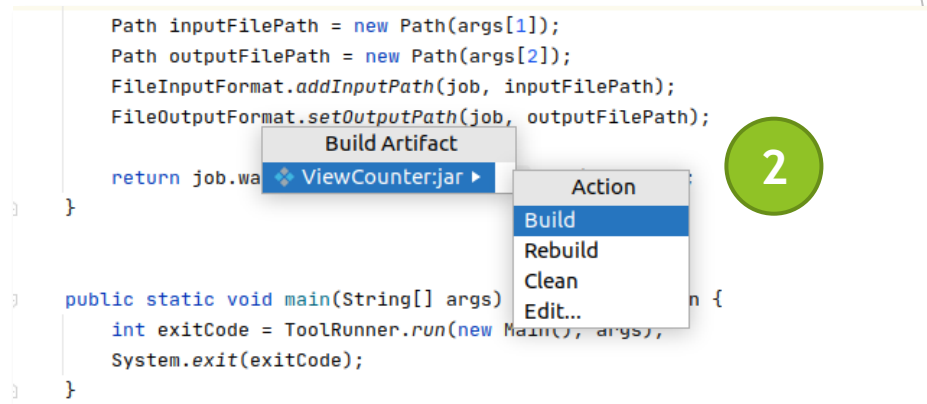
```

A 0. és 1. indexek tesztelésnél működnek, a Hadoop-ban csak akkor fog helyesen lefutni az alkalmazás ha átírjuk őket 1. és 2. indexekre.

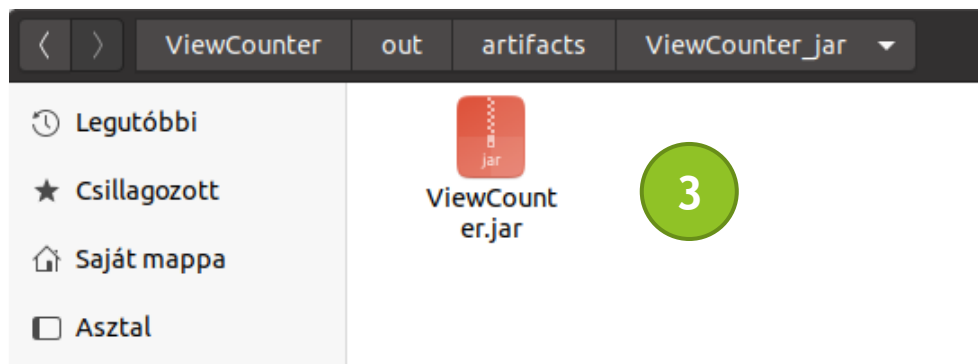
# A JAR build-elése



Válasszuk ki a **Build > Build Artifacts...** menüpontot.



A háttérben megjelenik egy helyi menü, ott válasszuk a **Build** opciót.



A programot lefuttatva létrejön a JAR fájl a **/ViewCounter/out/artifacts/ViewCounter.jar/** mappában.



ELKH Cloud

# A JAR fájl felmásolása a HDFS-re

# A JAR fájl NameNode-ra másolása

- ▶ A Windows-os gépünkön helyezük el egy mappában a JAR fájlt:
  - ▶ C:\pscp\WordCounter.jar
- ▶ Engedélyezzük a jelszó alapú hitelesítést a masteren:
  - ▶ `sudo nano /etc/ssh/sshd_config`
- ▶ Írjuk át a fájlban a következő sort **yes**-re:
  - ▶ PasswordAuthentication **no**
- ▶ Indítsuk újra az SSH szolgáltatást:
  - ▶ `sudo service ssh restart`
- ▶ A Windows parancssorból (cmd) adjuk ki a másolás parancsot:
  - ▶ `scp C:\pscp\ViewCounter.jar hadoopuser@<kulso-ip-cim>:/home/hadoopuser`

```
# To disable tunneled clear text passwords, change to no here!  
PasswordAuthentication yes  
#PermitEmptyPasswords no  
  
# Change to yes to enable challenge-response passwords (beware i  
  
^G Get Help   ^O Write Out  ^W Where Is   ^K Cut Text    ^J Justify  
^X Exit       ^R Read File  ^\ Replace    ^U Uncut Text ^T To Spell
```

```
C:\Users\RusznakA>scp C:\pscp\ViewCounter.jar hadoopuser@193.224.59.182:/home/hadoopuser  
hadoopuser@193.224.59.182's password:   
ViewCounter.jar
```

```
28% 39MB 7.2MB/s 00:13 ETA
```



# Futtatás előtti teendők a NameNode-on

- ▶ Hozzuk létre az input.txt fájlt, amin végrehajtjuk a job-ot:
  - ▶ `nano input.txt`
- ▶ Ez a txt fájl a NameNode-on kerül tárolásr
- ▶ Hozzunk létre néhány könyvtárat a HDFS-en:
  - ▶ `hadoop fs -mkdir /sztaki`
  - ▶ `hadoop fs -mkdir /sztaki/input`
  - ▶ **Output könyvtárat ne készítsünk, azt a Hadoop fogja létrehozni!**
- ▶ Másoljuk át a feldolgozandó fájlt vagy fájlokat az input mappába:
  - ▶ `hadoop fs -copyFromLocal input.txt /sztaki/input`

```
hadoopuser@hadoop-master: ~  
GNU nano 2.9.3 input.txt  
1. István Benjámín  
2. Lilla István  
3. Anikó Csaba  
4. Dóra Ernő  
5. Anikó István  
6. Dóra Csaba
```

```
hadoopuser@hadoop-master:~$ hadoop fs -ls /sztaki/input  
Found 1 items  
-rw-r--r-- 2 hadoopuser supergroup 102 2021-02-24 11:42 /sztaki/input/input.txt  
hadoopuser@hadoop-master:~$
```

# A JAR fájl futtatása a NameNode-on

- ▶ Futtassuk le a JAR fájlt a helyi fájlrendszerben (ne másoljuk át a HDFS-re):

- ▶ `hadoop jar ViewCounter.jar com.sztaki.Main /sztaki/input /sztaki/output`

Az args[1] eleme

Az args[2] eleme

```
2021-02-24 11:43:39,488 INFO mapred.LocalJobRunner: Finishing task: attempt_local1862087498_0001_r_000000_0
2021-02-24 11:43:39,488 INFO mapred.LocalJobRunner: reduce task executor complete.
2021-02-24 11:43:40,271 INFO mapreduce.Job: map 100% reduce 100%
2021-02-24 11:43:41,273 INFO mapreduce.Job: Job job_local1862087498_0001 completed successfully
2021-02-24 11:43:41,294 INFO mapreduce.Job: Counters: 36
```

- ▶ Nézzük meg a generált output fájl tartalmát:

- ▶ `hadoop fs -cat /sztaki/output/part-r-00000`

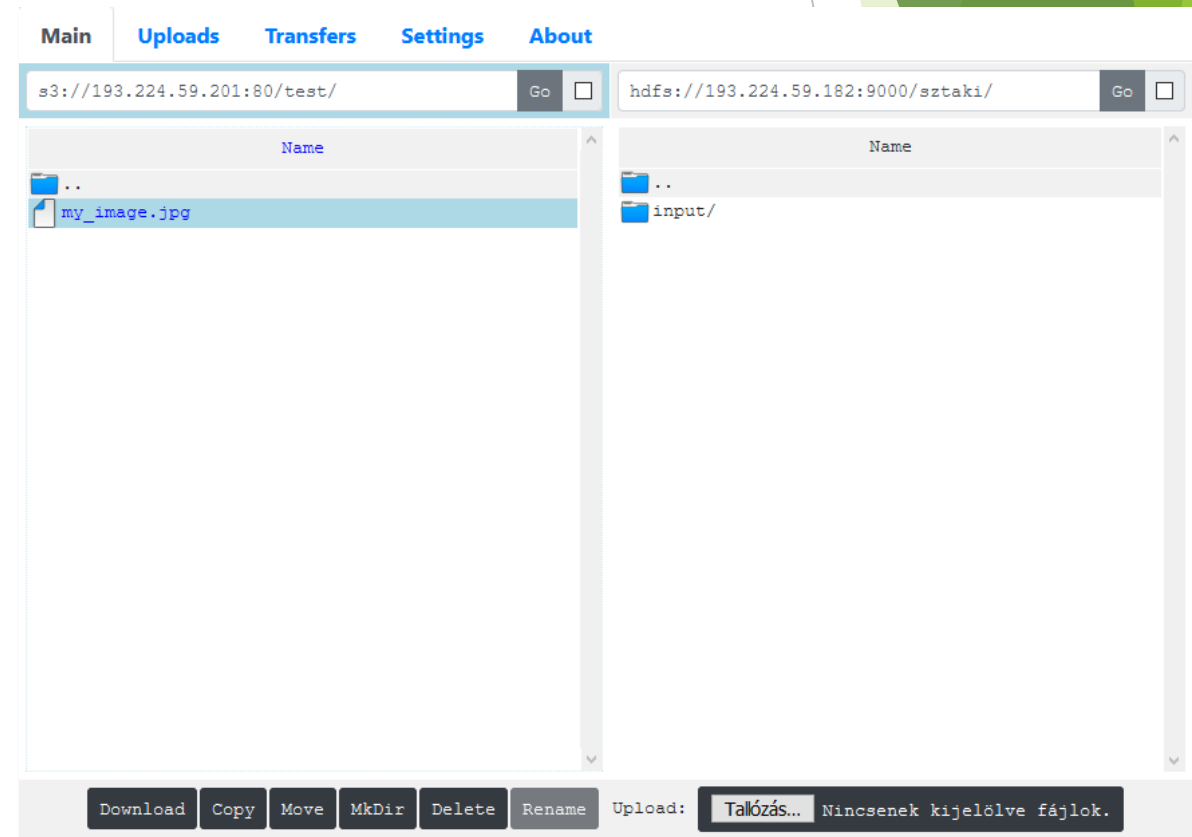
```
hadoopuser@hadoop-master:~$ hadoop fs -cat /sztaki/output/part-r-00000
Benjámín      1
Csaba        2
Ernő         1
István       2
hadoopuser@hadoop-master:~$
```

# Nagyméretű fájlok a HDFS-re másolása

Egy valódi projekt esetében erre nyújt megoldást a Data Avenue

- ▶ Segítségével méretkorlát nélkül tudunk fájlokat másolni a HDFS-re
- ▶ Grafikus felülettel rendelkezik (böngészős)
- ▶ Könnyű csatlakozni a távoli tárhelyre
- ▶ Számos protokollt támogat

Így nincs szükségünk parancssorban történő fájlmozgatásra, ahogyan azt az előző diákon láttuk.





ELKH Cloud

# A MapReduce job monitorozása

# A MapReduce job monitorozása



## All Applications

Cluster

- About
- Nodes
- Node Labels
- Applications
  - NEW
  - NEW SAVING
  - SUBMITTED
  - ACCEPTED
  - RUNNING
  - FINISHED
  - FAILED
  - KILLED
- Scheduler

Tools

### Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes
2	0	0	2	0	0 B	16 GB	0 B	0	16 GB	0	1

### Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation
Capacity Scheduler	[MEMORY]	<memory:1024, vCores:1>

Show 20 entries

ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	Final
<a href="#">application 1476694929592 0004</a>		wordcount	MAPREDUCE	default	Mon Febr 22 16:24:11 +0550 2021	Mon Febr 22 16:24:17 +0550 2021	FINISHED	SU

A webes felület elérhető a <http://<public-ip>:8088> címen

# A MapReduce job monitorozása

## All Applications

Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
1	2	1	2 GB	16 GB	0 B	1	8	0	1	0	0	0	0

App	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
[MEMORY]		<memory:1024, vCores:1>	<memory:8192, vCores:8>

User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI	Blacklisted Nodes
1476694929592_0005	wordcount	MAPREDUCE	default	Mon Feb 22 16:24:11	N/A	RUNNING	UNDEFINED	<input type="text"/>	ApplicationMaster	0



## MapReduce Application application\_1476694929592\_0005

- Cluster
- Application
  - About Jobs
- Tools

### Active Jobs

Job ID	Name	State	Map Progress	Maps Total	Maps Completed	Reduce Progress	Reduces Total	Reduces Completed
job_1476694929592_0005	wordcount	RUNNING	<input type="text"/>	30	0	<input type="text"/>	1	0

Logged in as:

# A MapReduce job monitorozása



## MapReduce Application application\_1476694929592\_0005

Logged in as: d

Logged in as:

- ▶ Cluster
- ▼ Application
  - About
  - Jobs
- ▶ Tools

### Active Jobs

Show 20 entries Search:

Job ID	Name	State	Map Progress	Maps Total	Maps Completed	Reduce Progress	Reduces Total	Reduces Completed
job_1476694929592_0005	wordcount	RUNNING		30	0		1	0

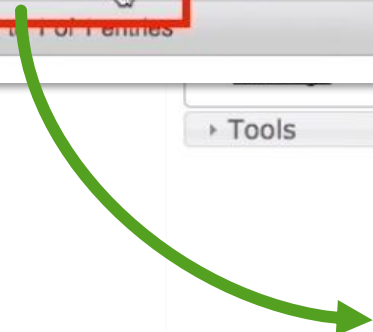
Showing 1 of 1 entries

### Job Overview

Node	Logs
192.168.0.100:8042	logs

- ▶ Tools

Task Type	Progress	Total	Pending	Running	Complete
<b>Map</b>	<div style="width: 100%;"></div>	30	3	5	22
<b>Reduce</b>	<div style="width: 100%;"></div>	1	0	1	0
Attempt Type	New	Running	Failed	Killed	Successful
<b>Maps</b>	3	5	0	0	22
<b>Reduces</b>	0	1	0	0	0







ELKH Cloud

# A YARN monitorozása



# A Capacity ütemező finomhangolása

- ▶ Nyissuk meg a kapacitás ütemező beállításait:

- ▶ `sudo nano /usr/local/hadoop/etc/hadoop/capacity-scheduler.xml`

- ▶ Állítsuk be a következő tulajdonságokat:

```
<property>
  <name>yarn.scheduler.capacity.root.queues</name>
  <value>dev, prod</value>
</property>
<property>
  <name>yarn.scheduler.capacity.root.dev.capacity</name>
  <value>30</value>
</property>
<property>
  <name>yarn.scheduler.capacity.root.prod.capacity</name>
  <value>70</value>
</property>
```

# Az ütemező monitorozása

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes
1	0	1	0	6	7 GB	16 GB	0 B	6	8	0	1	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation
Capacity Scheduler	[MEMORY]	<memory:1024, vCores:1>

Show 20 entries


ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress
application_14767783022510001		wordcount	MAPREDUCE	prod	Mon Febr 22 17:42:20 +0550 2021	N/A	RUNNING	UNDEFINED	

Showing 1 to 1 of 1 entries

A példában a Capacity ütemező két részre került felbontásra:

- dev: 30%
- prod: 70%

# Az ütemező monitorozása



Logged in as: dr.who

## Application application\_1476778302251\_0001

Cluster

- About
- Nodes
- Node Labels
- Applications
  - NEW
  - NEW SAVING
  - SUBMITTED
  - ACCEPTED
  - RUNNING
  - FINISHED
  - FAILED
  - KILLED
- Scheduler

Kill Application

Application Overview

User:	
Name:	wordcount
Application Type:	MAPREDUCE
Application Tags:	
YarnApplicationState:	RUNNING: AM has registered with RM and started running.
Queue:	prod
FinalStatus Reported by AM:	Application has not completed yet.
Started:	Mon Febr 22 17:42:20 +0550 2021
Elapsed:	53sec
Tracking URL:	ApplicationMaster
Diagnostics:	

# Az ütemező monitorozása

## Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes
1	0	1	0	6	7 GB	16 GB	0 B	6	8	0	1	0

## Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Capacity Scheduler	[MEMORY]	<memory:1024, vCores:1>	<memory:8192, vCores:8>

## Application Queues

### Legend: Capacity

- root
- + Queue: dev
- + Queue: prod

## Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes
1	0	1	0	1	2 GB	16 GB	0 B	1	8	0	1	0	0	0

## Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Capacity Scheduler	[MEMORY]	<memory:1024, vCores:1>	<memory:8192, vCores:8>

## Application Queues

### Legend: Capacity Used Used (over capacity) Max Capacity



# A Hadoop klaszter újraindítása

- ▶ A JPS segítségével tudjuk lekérni az aktuálisan futó Hadoop folyamatokat:
  - ▶ `jps`
- ▶ A teljes klaszter leállítása:
  - ▶ `stop-all.sh`
- ▶ A teljes klaszter újraindítása:
  - ▶ `start-all.sh`

```
hadoopuser@hadoop-master: ~  
hadoopuser@hadoop-master:~$ jps  
2704 SecondaryNameNode  
2424 NameNode  
2953 ResourceManager  
3277 Jps  
hadoopuser@hadoop-master:~$ █
```



Köszönöm a figyelmet!