EKONOMICKÁ UNIVERZITA

Fakulta hospodárskej informatiky

Katedra aplikovanej informatiky

Inštalácia hadoop 3.3.3 na windows 10

**Meno študentov:** Martin Jankech, Patrik Hajdučík

**Stupeň/ročník:** druhý/2.ročník

**Školský rok:** 2022/2023

**Meno prednášajúceho:** Ing. Jaroslav Kultán PhD.

**Meno cvičiaceho:** SCHMIDT, Peter, Ing. Mgr., PhD.

Obsah

[Úvod 3](#_Toc128482051)

[1.0 Postup inštalácie 4](#_Toc128482052)

[Záver 10](#_Toc128482053)

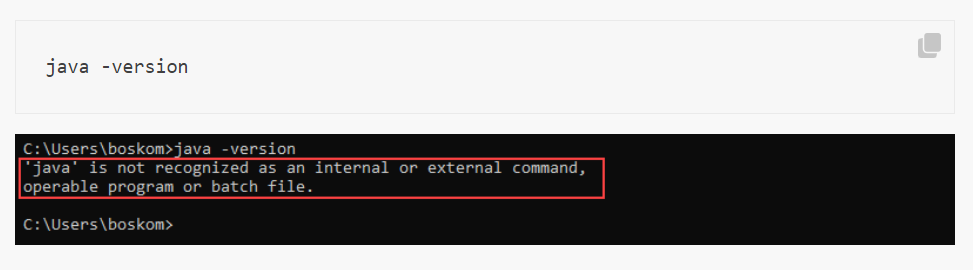
[Bibliografia 11](#_Toc128482054)

# Úvod

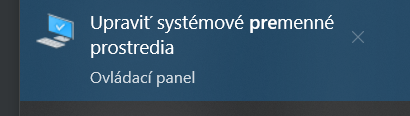
V tomto zadaní si ukážeme a popíšeme postup akým sa inštaluje hadoop 3.3.3 na operačný systém Windows 10. Pred inštaláciou si je potrebné nainštalovať javu. My použijeme verziu 1.8.0\_271, ktorá je zaručene touto verziou hadoopu podporovaná.

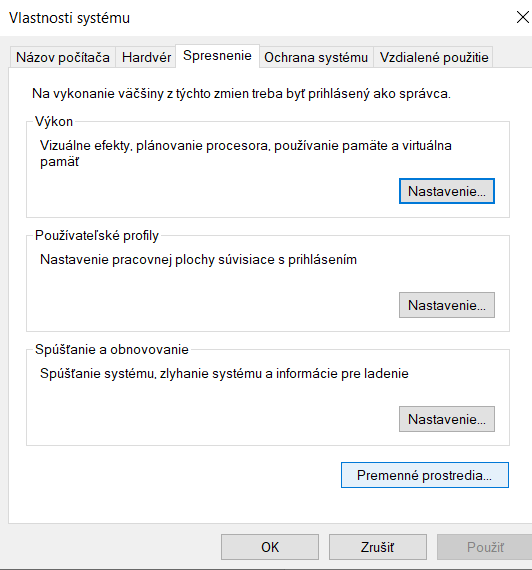
# 1.0 Postup inštalácie

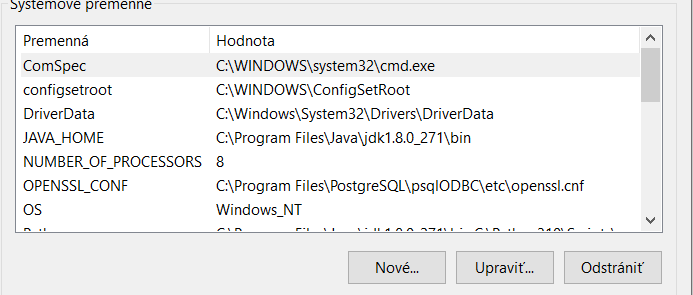
1. Pre inštaláciou hadoopu je potrebne skontrolovať či je Java nainštalovaná v systéme. Vykonať to môžeme napísaním príkazu java -version v príkazovom riadku. Pokiaľ dostaneme takýto výstup, tak Java nie je v našom systéme nainštalovaná.



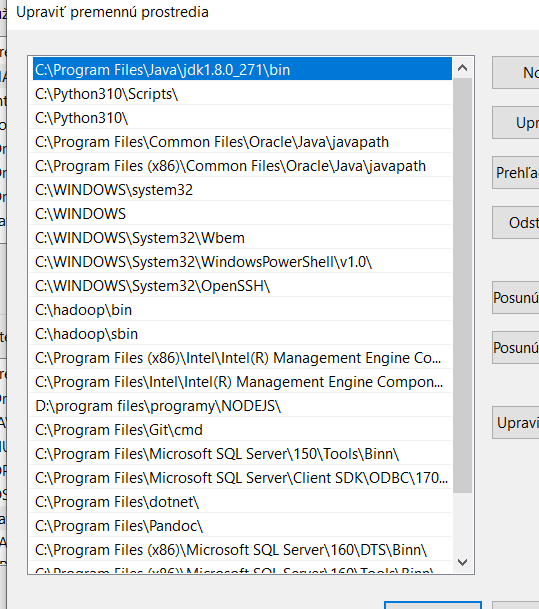
1. Pokiaľ nemáme javu nainštalovanú tak ju treba stiahnuť napr. so stránky Oracle a nainštalovať. Odporúčame do koreňového adresára C: (spôsobuje najmenej problémov ).
2. Po inštalácii je potrebné nastaviť systémové premenné prostredia.
3. Vytvoríme premennú JAVA\_HOME a pridáme cestu k nainštalovanej jave.



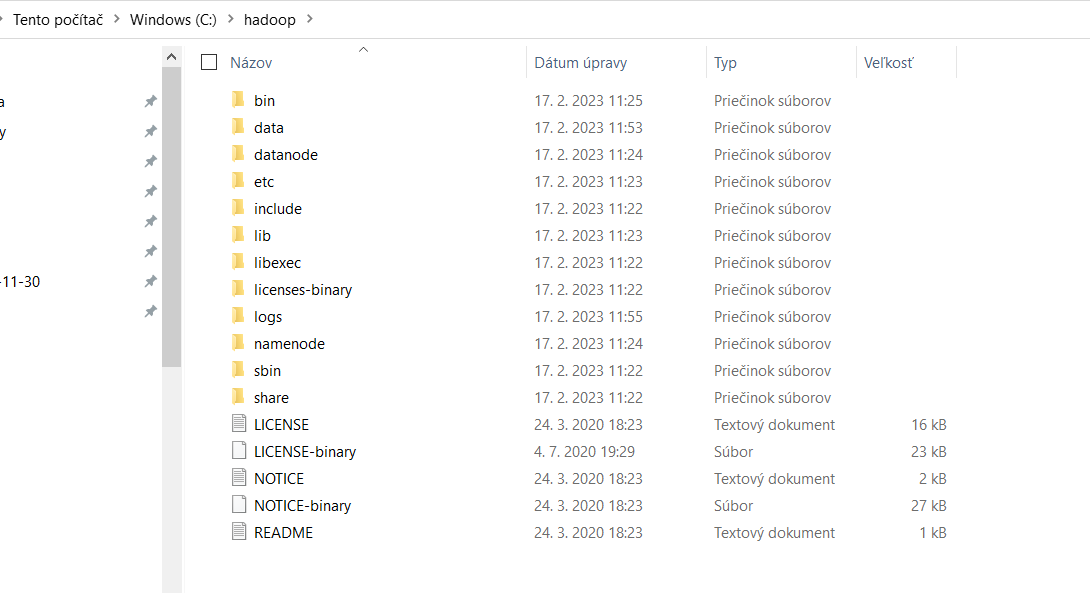




1. Taktiež pridáme cestu k jave v premennej path rovnako ako na obrázku.



1. Po inštalácii javy sa môžem pustiť na inštaláciu hadoopu.
2. Najprv si stiahneme hadoop so stránky Apache hadoop. My použijeme verziu 3.3.3
3. Rozbalený adresár nahráme na cieľové miesto. V princípe môže byť v inom adresári alebo na inom mieste, ale najmenej problémov sa vyskytne keď je v roote na bootovacom disku v našom prípade disk C. Názov adresára Hadoop 3.3.0 sme premenovali na Hadoop. Pozn. Je to dobré urobiť ešte pred tým než nastavíme cesty k súborom.
4. V adresári data vytvoríme 2 adresáre datanode a namenode



1. Stiahneme si konfiguračný súbor pre Hadoop 3.3.0 - 3.3.0 Configuration Files Apache Hadoop.zip, ktorý obsahuje jeden adresár bin. Obsah tohto adresára nakopírujeme do pôvodného adresára bin, a pôvodné súbory prepíšeme. Stiahnuť to možeme napr. odtiaľto

<https://github.com/brainmentorspvtltd/BigData_RDE/blob/master/Hadoop%20Configuration.zip>

alebo

<https://kai.ponuky.info/index.php?menu=6>

1. Ďalej si otvoríme adresár etc kde budeme konfigurovať 1cmd a 4 xml súbory
2. core-site.xml b. hadoop-env.cmd c. hdfs-site.xml d. mapred-site.xml e. yarn-site.xml
3. **core-site.xml**

doplníme konfiguráciu na

<configuration>

<property>

<name>fs.defaultFS</name>

<value>hdfs://localhost:9000</value>

</property>

</configuration>

Nastavíme hadoop-env.cmd set JAVA\_HOME=C:\Java\jdk1.8.0\_341

My sme museli cestu upraviť takto



13. Nasleduje hdfs-site.xml, kde nastavíme konfiguráciu na

<configuration>

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

<property>

<name>dfs.namenode.name.dir</name>

<value>file:///C:/hadoop/data/namenode</value>

</property>

<property>

<name>dfs.datanode.data.dir</name>

<value> file:///C:/hadoop/data/datanode</value>

</property>

</configuration>

14. Mapred-site.xml

<configuration>

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

</configuration>

15. Yarn-site.xml upravíme na

<configuration>

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

</property>

<property>

<name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</name>

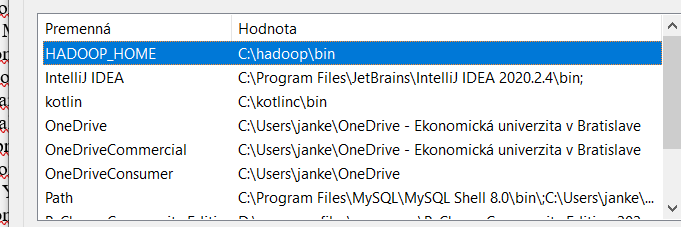
<value>org.apache.hadoop.mapred.ShuffleHandler</value>

</property>

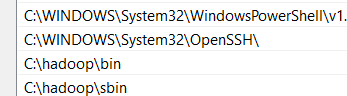
</configuration>

Nezabudnite všetky konfiguračné zmeny uložiť

16. dalej pridáme cestu k hadoopu ako používateľskú premennú



17. A tak isto aj systémovú premennú v path

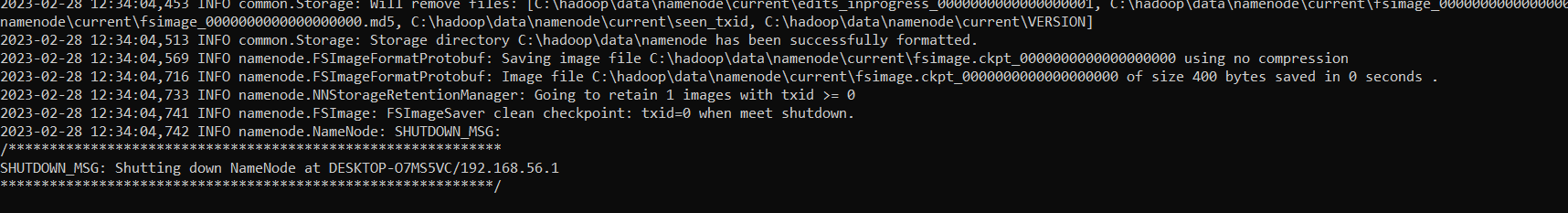


18. Ďalej naformátujeme namenode



19. Potvrdíme Y



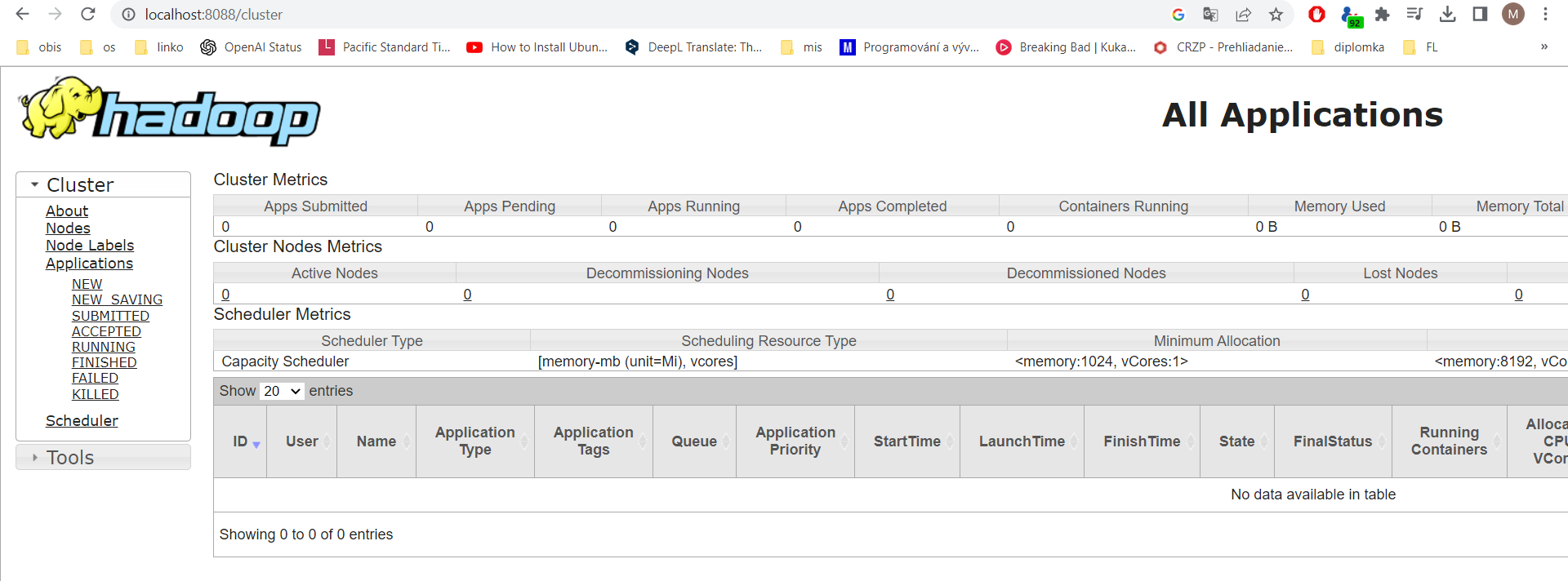


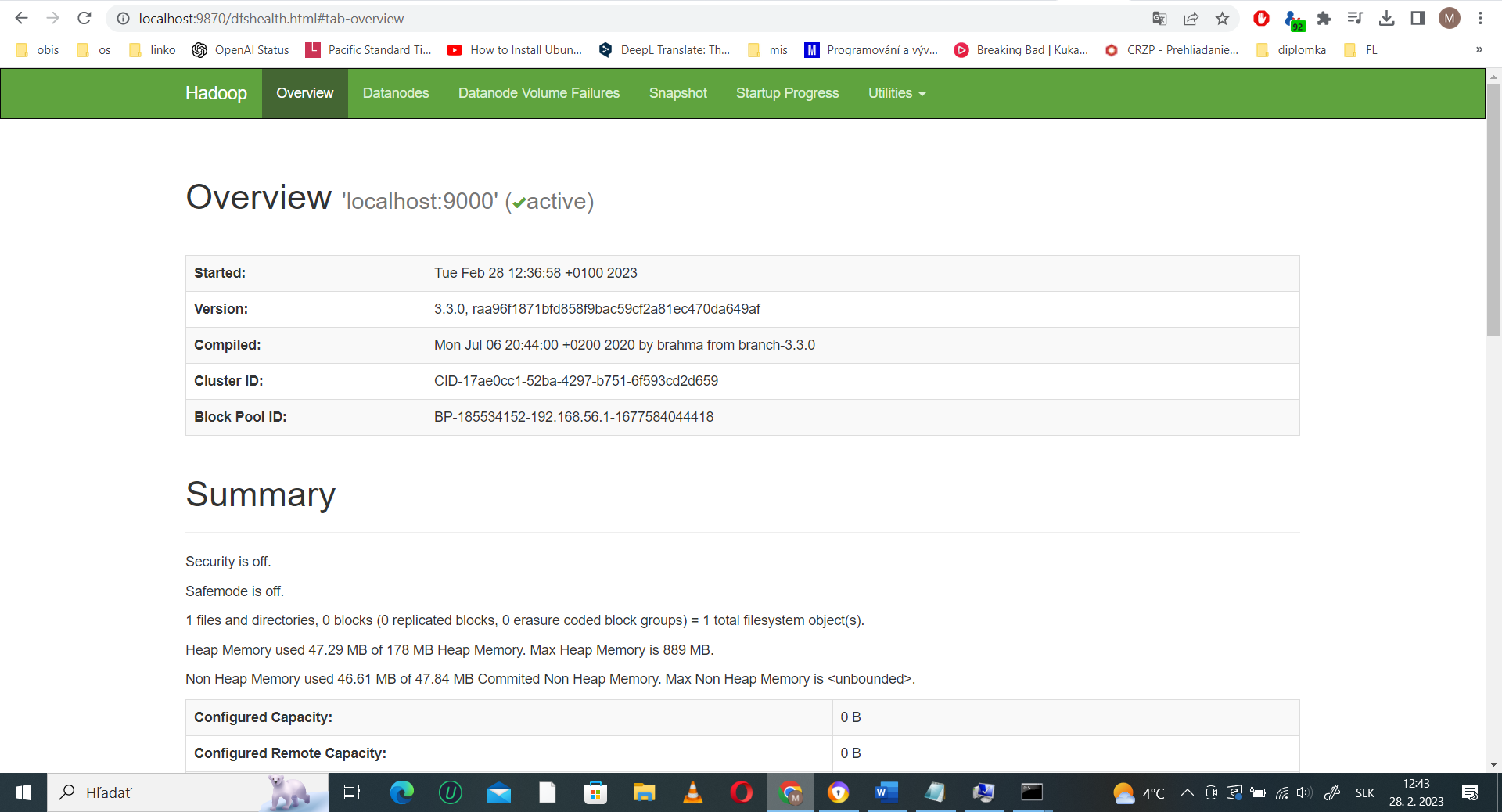
20. Príkazom C:\Users\EU>start-dfs.cmd spúšťame namenode a datanode a príkazom C:\Users\EU> start-yarn.cmd spúšťame yarn

Príkazom JPS môžeme otestovať či sa vykonávajú všetky služby

Či nám služby fungujú môžeme overiť aj ich otvorením v prehliadači na localhoste a portoch 8088 a 9870

(P., 2023)





# Záver

V tomto zadaní sme si ukázali a popísali ako nainštalovať Hadoop na Windows. Až na jednu zmenu, ktorú sme museli spraviť pri zadávaní cesty k jave sme postupovali rovnako ako v návode, ktorý nám bol poskytnutí. S potešením môžeme prehlásiť že inštalácia podľa návodu bola úspešná v všetko funguje ako má.

# Bibliografia

P., S. (28. 2 2023). *Ako nainštalovať Hadoop*. Dostupné na Internete: https://kai.ponuky.info/index.php?menu=6