

Przygotowanie środowiska

Klucz w EC2

1. Logujemy się do konsoli AWS
2. Wyszukujemy moduł EC2
3. Wybieramy odpowiedni region (u góry po prawo)
4. Przechodzimy do sekcji “Key Pairs”
5. “Create key pair”
6. Podajemy własną nazwę
7. Uruchamiamy Puttygen
8. Klikamy “Load na existing private key file”
9. Wybieramy pobrany plik
10. Klikamy “Save private key” i klikamy “Yes”
11. Powstaje plik PPK

Uruchamiamy klaster

1. Przełączamy się na usługę EMR (Services u góry)
2. Wyszukujemy na liście cluster_template
3. Zaznaczamy i klikamy “Clone”
4. Cofamy się do Stage 3
5. Wpisujemy własną nazwę
6. Klikamy Next
7. Wybieramy własny klucz
8. Create cluster

Logujemy się

1. Uruchamiamy Putty
2. Hostname: `hadoop@<master_public_dns>`
3. Connection -> SSH -> Auth -> Browse i wybieramy utworzony PPK
4. Wracamy do Session
5. Wpisujemy swoją nazwę pod Saved Sessions
6. Klikamy Save
7. Klikamy Open

Klonujemy repo

1. Otwieramy GitHub Desktop
2. Add -> Clone Repository

https://bitbucket.org/arekosinski/kompleksowe_bigdata.git