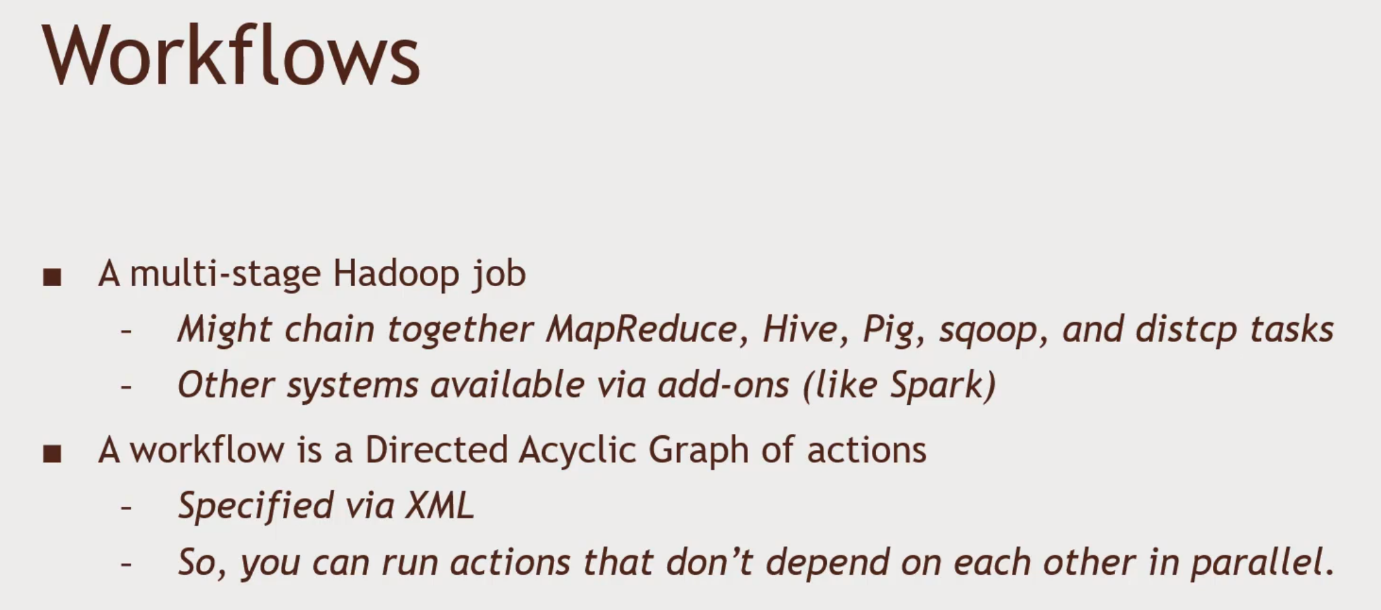
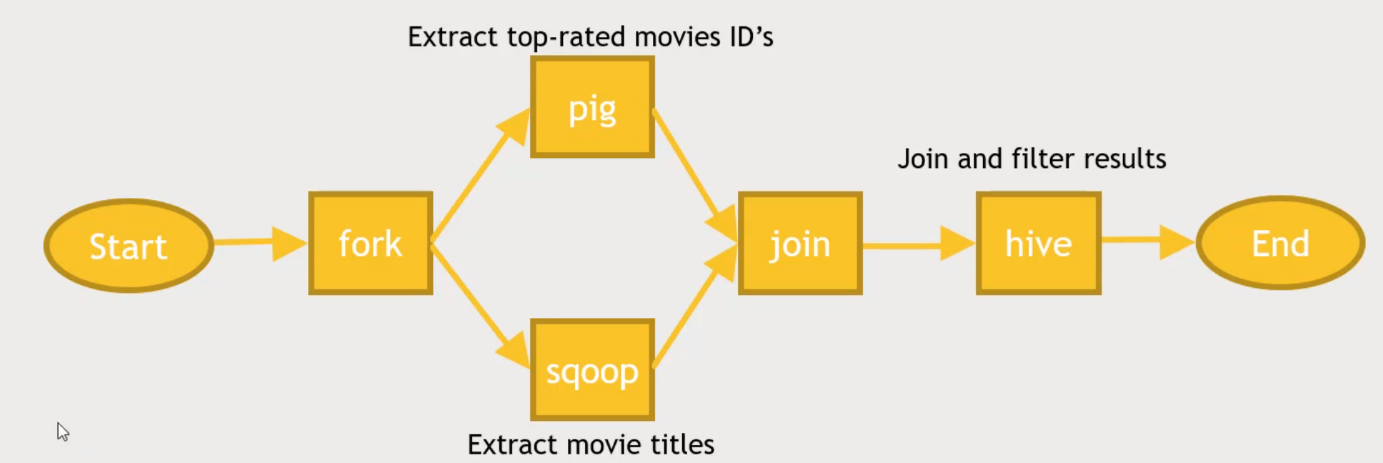
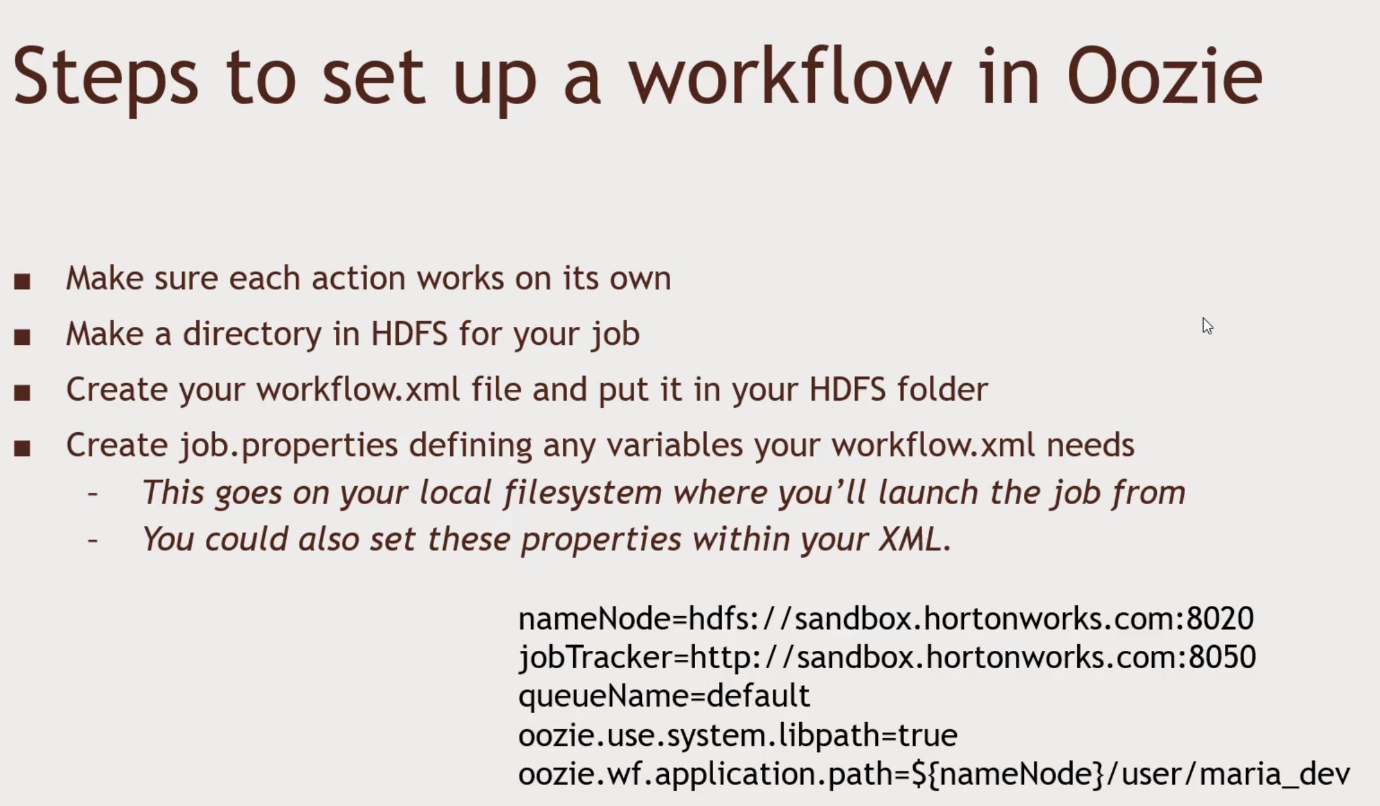
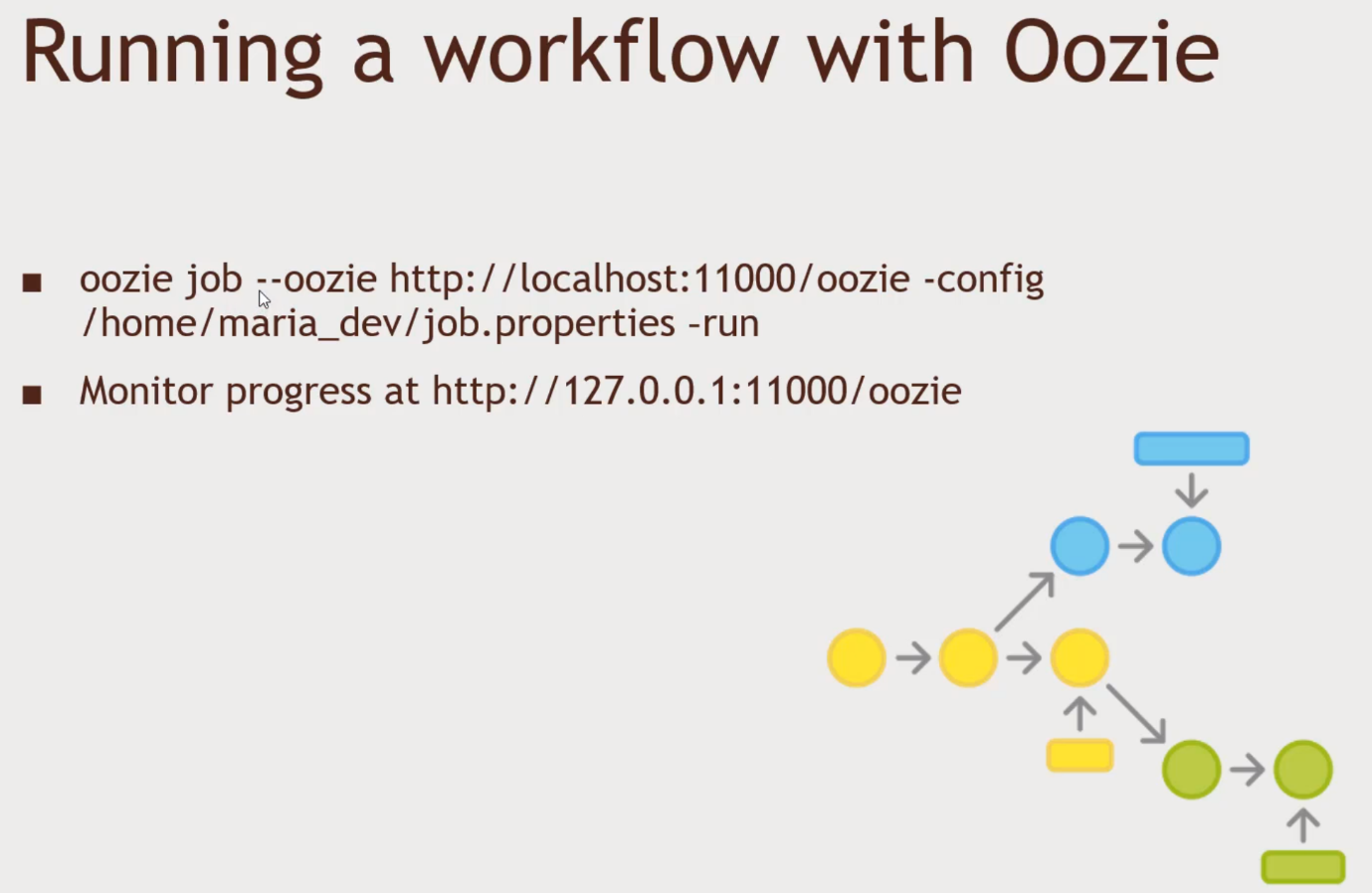
Oozie

* Burmese for ‘Elephant Keeper’
* A system for running and scheduling Hadoop tasks.



Typical XML workflow of Oozie



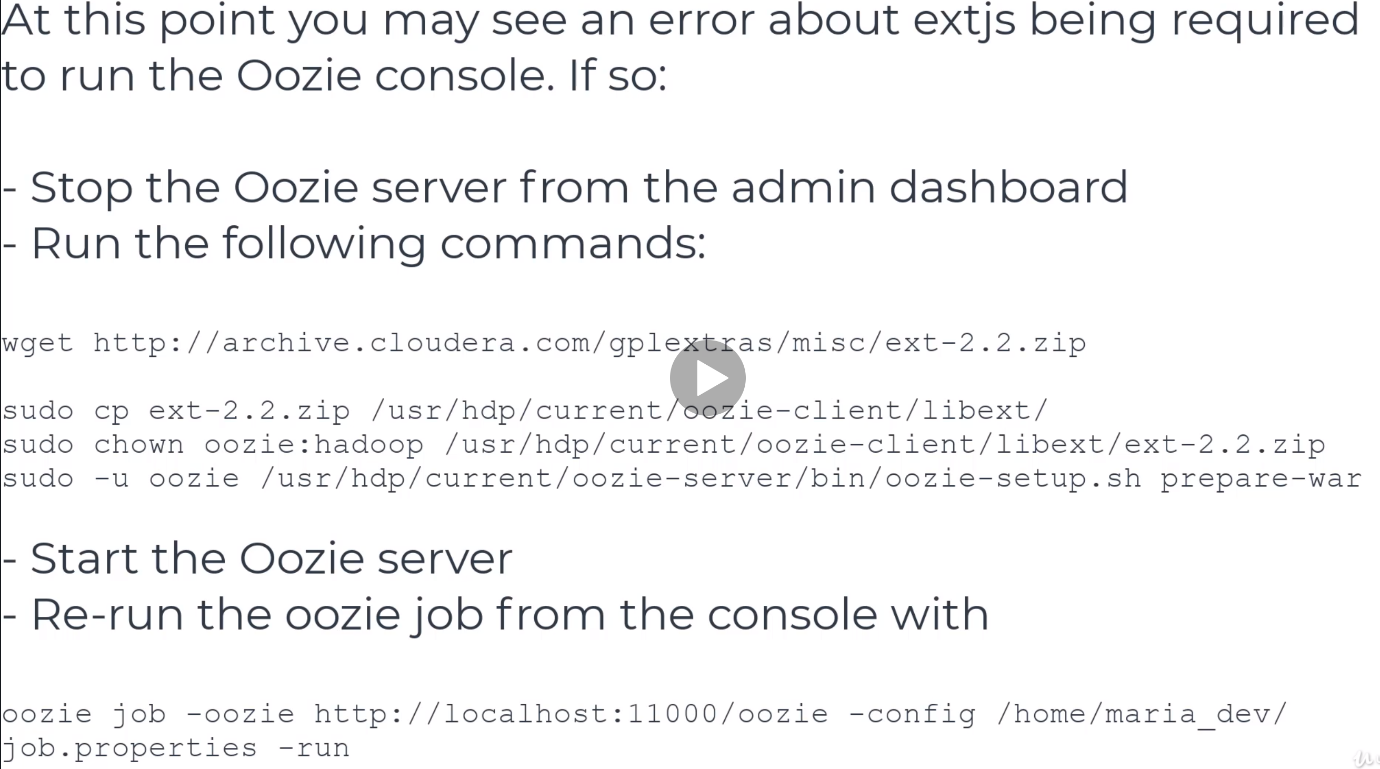
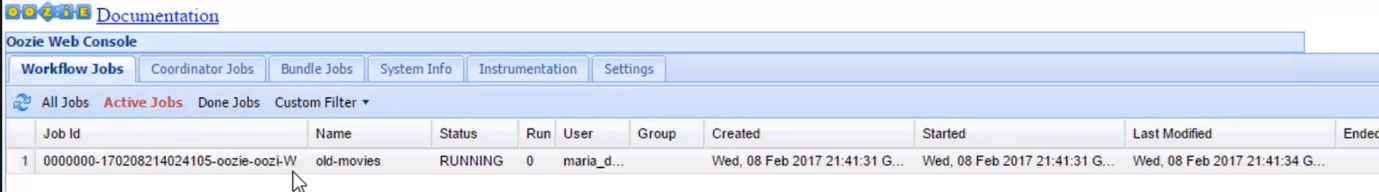
Oozie Coordinators

* Schedules workflow execution
* Launches workflows based on a given start time and frequency
* Will also wait for required input data to become available
* Run in exactly the same way as a workflow

Oozie bundles

* New in Oozie 3.0
* A bundle is a collection of coordinators that can be managed together
* You may have a bunch of coordinators for processing log data in various ways
  + By grouping them in a bundle, you could suspend them all if they were some problem with log collection

Oozie practice

1. Login to Hortonworks and Login to Putty
2. Ensure that MySQL has some data
   1. mysql -u root – p
   2. show databases;
3. If not, do following in the command prompt
   1. wget <http://media.sundog-soft.com/hadoop/movielens.sql>
   2. mysql -u root -p
   3. set names ‘utf8’;
   4. set character set utf8;
   5. create database movielens;
   6. use movielens;
   7. source movielens.sql;
   8. show tables;
   9. select \* from movies limit 10;
4. Permissions
   1. grant all privileges on movielens.\* to ‘’@’localhost’;
5. Get hive script
   1. wget <http://media.sundog-soft.com/hadoop/oldmovies.sql>
   2. less oldmovies.sql
6. Get Oozie setup
   1. wget <http://media.sundog-soft.com/hadoop/workflow.xml>
   2. wget <http://media.sundog-soft.com/hadoop/job.properties>
   3. hadoop fs -put workflow.xml /user/maria\_dev
   4. hadoop fs -put oldmovies.sql /user/maria\_dev
   5. mySQL connector: hadoop fs -put /user/share/java/mysql-connector-java.jar /user/oozie/share/lib/lib\_20161025075203/sqoop
   6. Go to Services Ambari, Oozie, restart all for Oozie to pick up the mySQL connector.
   7. oozie job -oozie [http://localhost:11000/oozie -config /home/maria\_dev/job.properties](http://localhost:11000/oozie%20-config%20/home/maria_dev/job.properties) – run
   8. Go to browser and see
   9. You will see status
   10. Go to files view and find the output under user/maria\_dev and find it out.