

# Real World Big Data

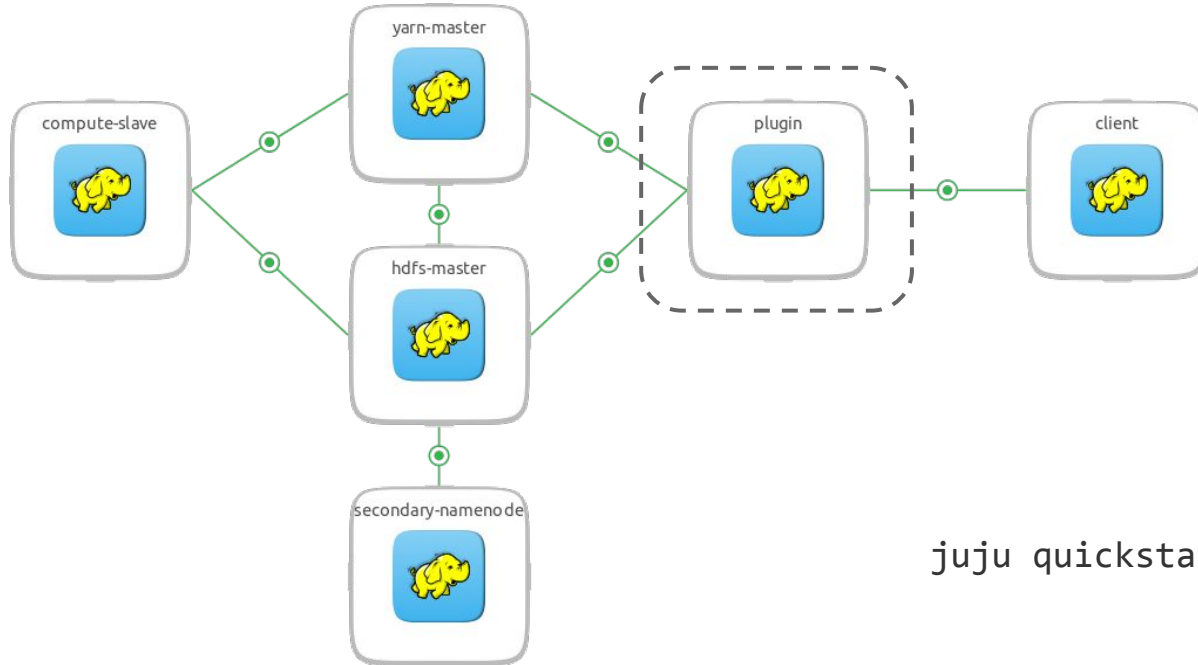
Jorge Castro, Kevin Monroe

# Juju & Big Data

- Modeling language for service oriented environments
  - Deploy, relate, scale
  - Reliable
  - Repeatable
  - Observable

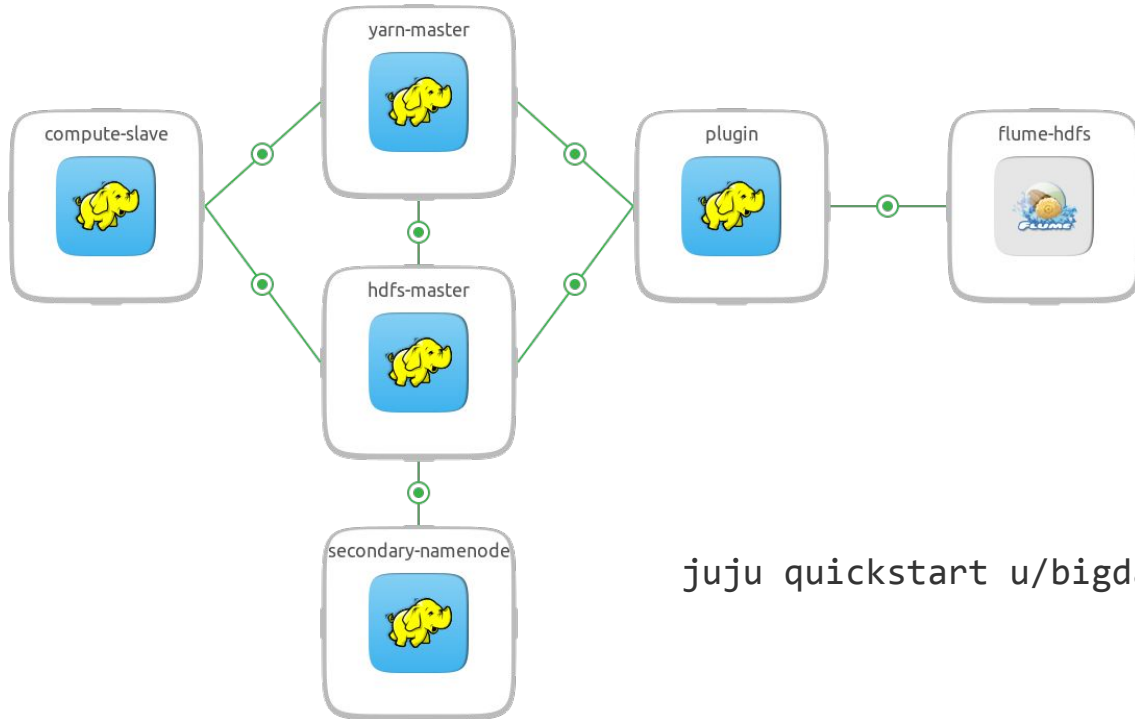


# Plugging in to Hadoop



juju quickstart apache-core-batch-processing

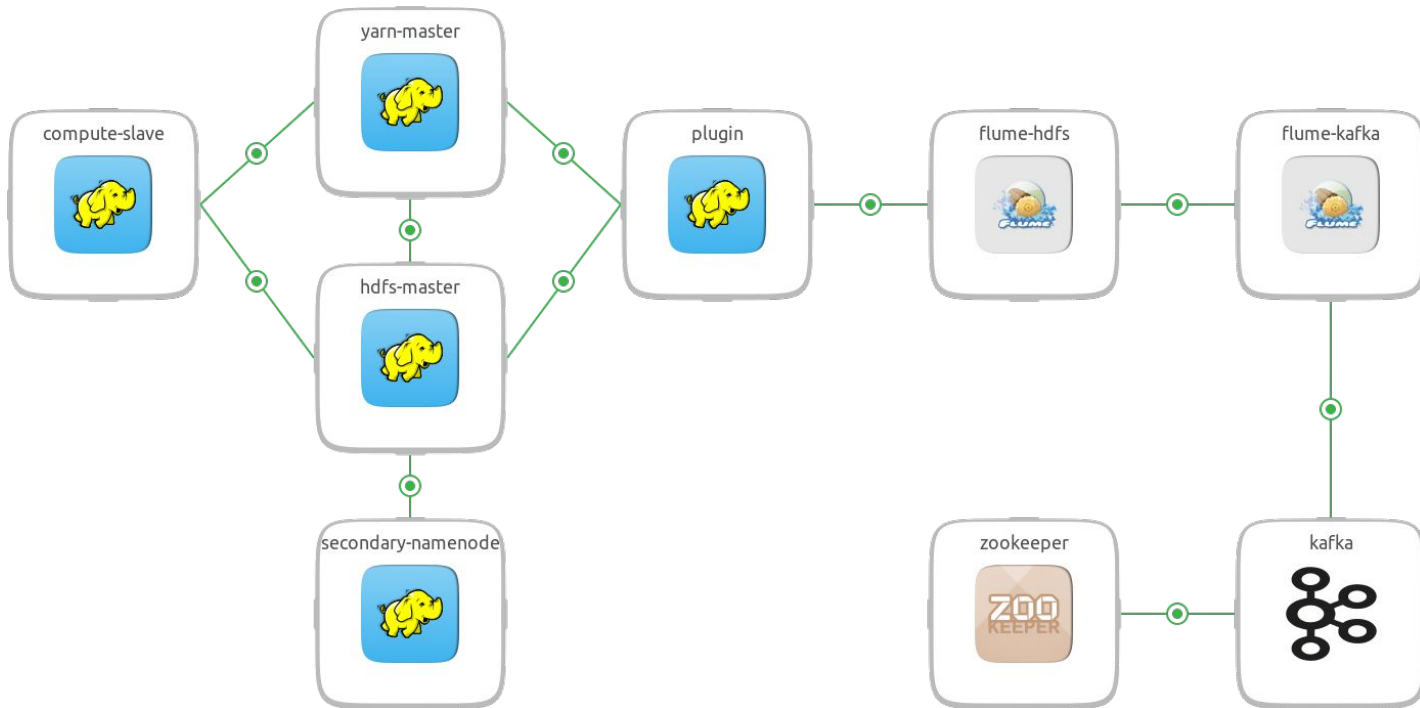
# Add Apache Flume



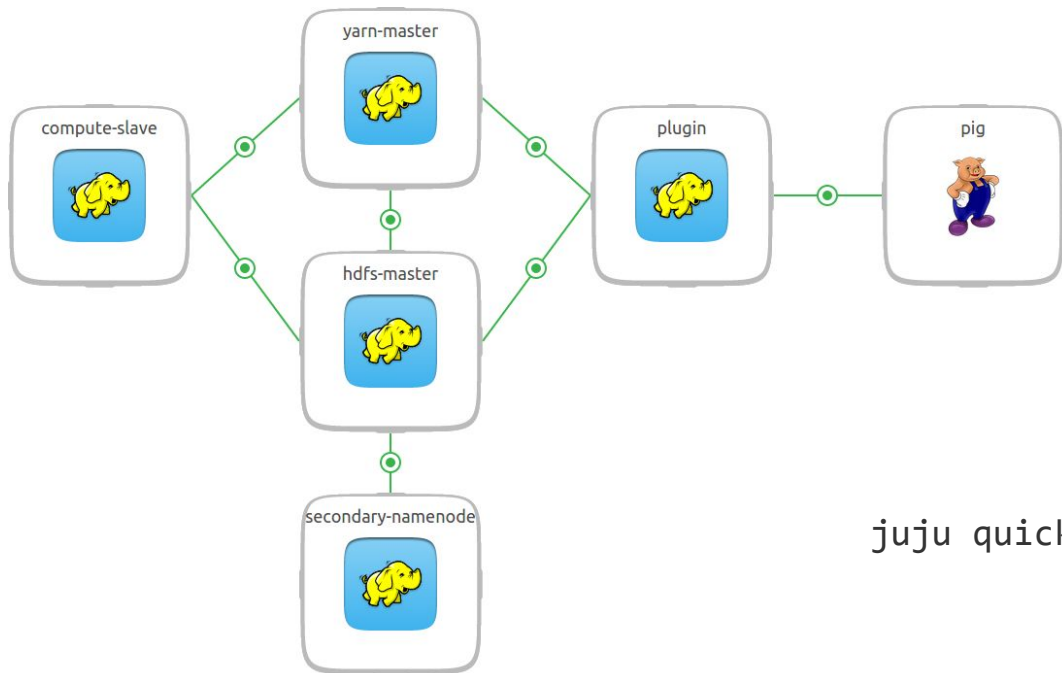
`juju quickstart u/bigdata-dev/apache-ingestion-flume`

# Add Apache Kafka

`juju quickstart u/bigdata-dev/apache-flume-ingestion-kafka`



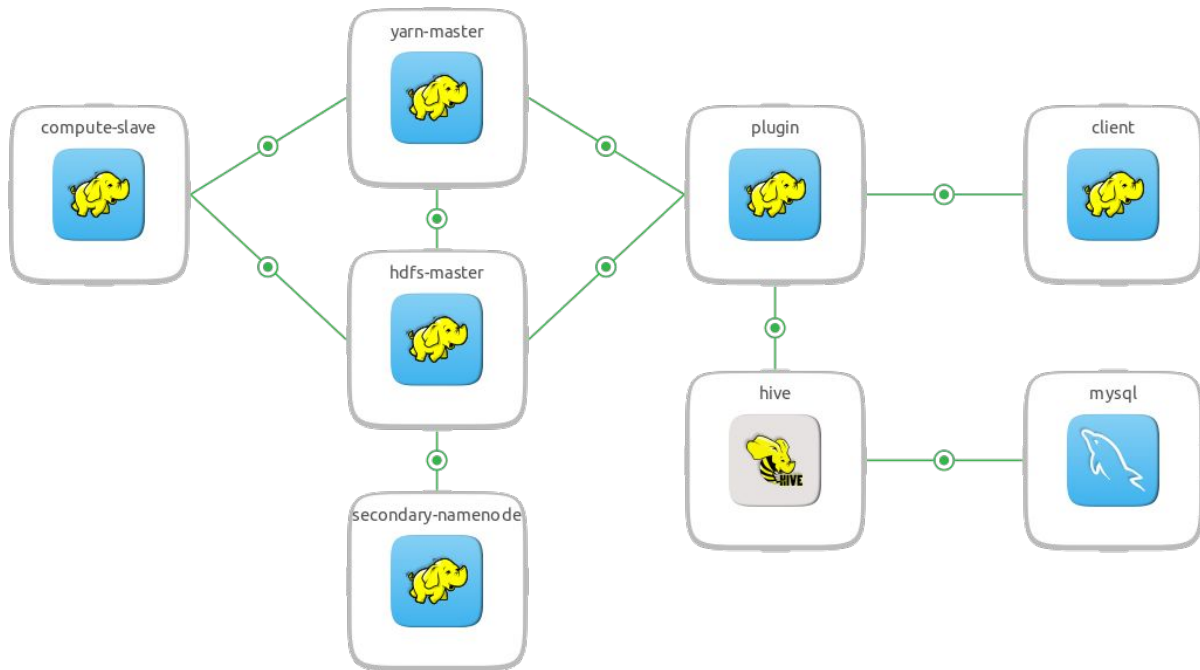
# Add Apache Pig



```
juju quickstart apache-analytics-pig
```

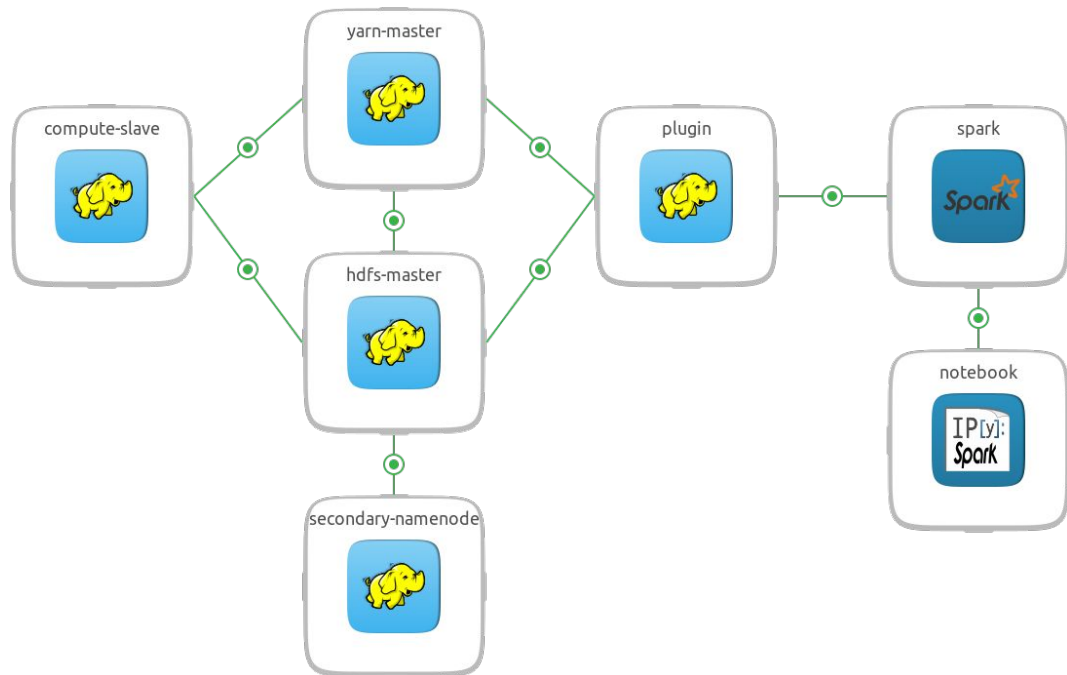
# Add Apache Hive

```
juju quickstart apache-analytics-sql
```



# Add Spark and iPy Notebook

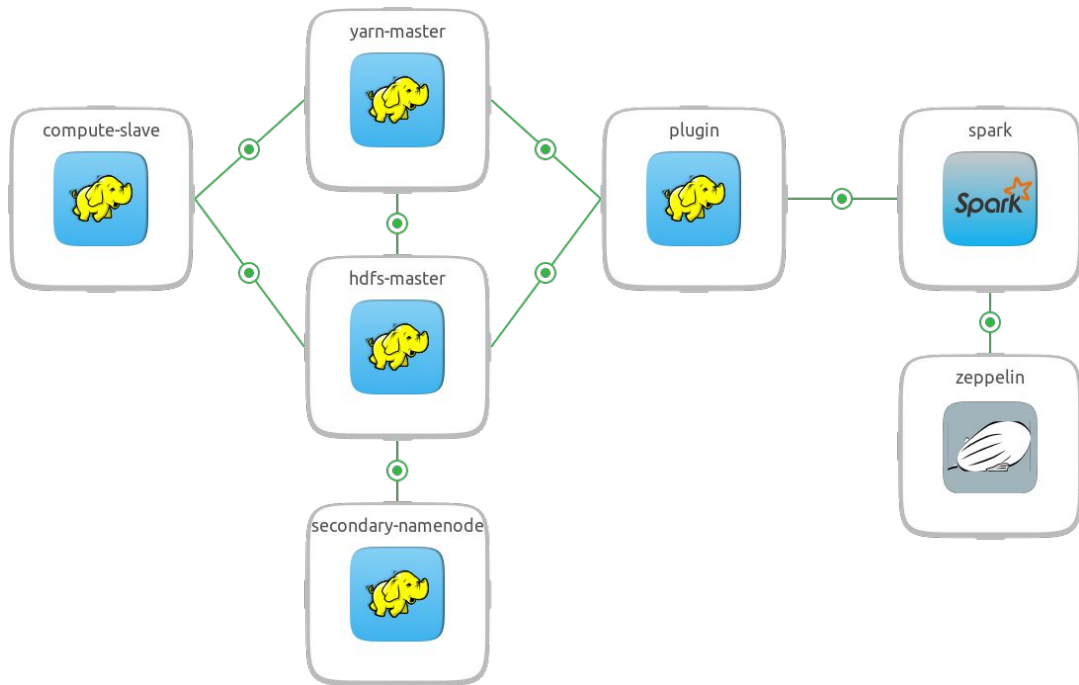
`juju quickstart apache-hadoop-spark-notebook`





# Add Apache Zeppelin

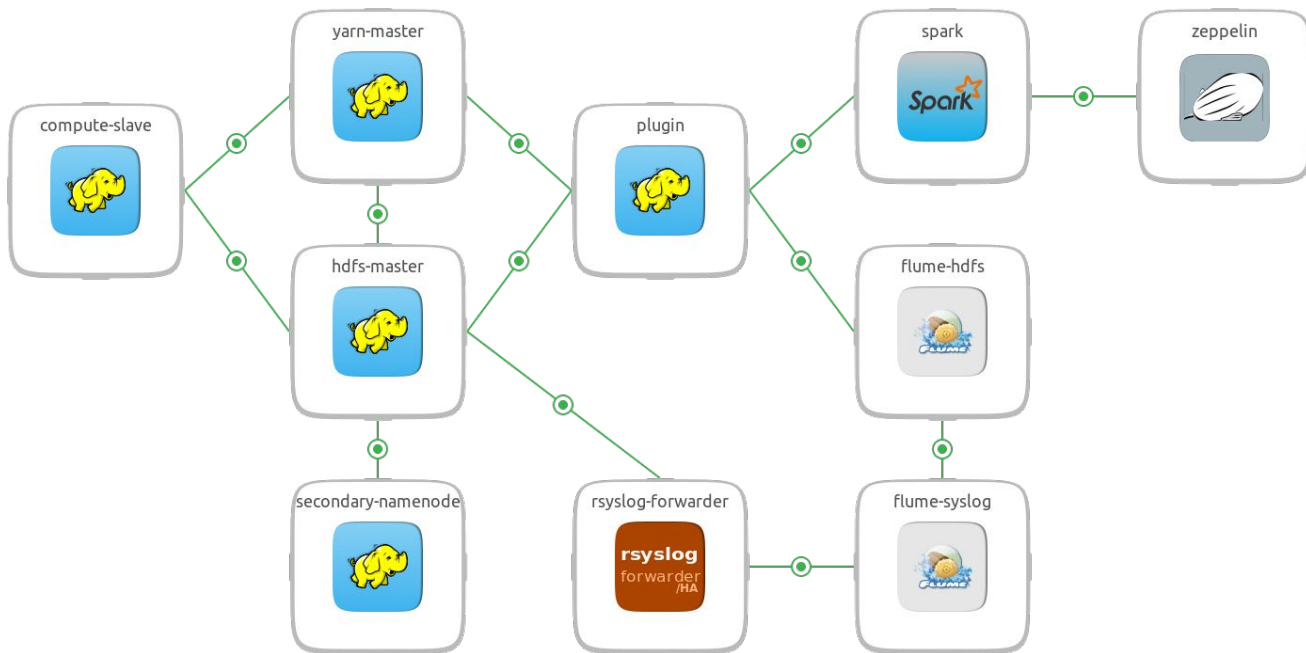
```
juju quickstart apache-hadoop-spark-zeppelin
```



Build *Your* Solution

# Syslog Analytics

juju quickstart realtime-syslog-analytics



<http://developer.juju.solutions>

# Resources

- Mailing Lists
  - Juju Big Data: <https://lists.ubuntu.com/mailman/listinfo/bigdata>
  - Juju: <https://lists.ubuntu.com/mailman/listinfo/juju>
- IRC (Freenode)
  - #juju
- Web
  - [bigdata.juju.solutions](http://bigdata.juju.solutions)
  - [developer.juju.solutions](http://developer.juju.solutions)
    - **Free AWS creds!!**
  - [jujucharms.com/docs](http://jujucharms.com/docs)

**Thanks!**