

# Mesos-Based Big Data Infrastructures @ Douban

[tianzhongbo@douban.com](mailto:tianzhongbo@douban.com)

2016/11/15

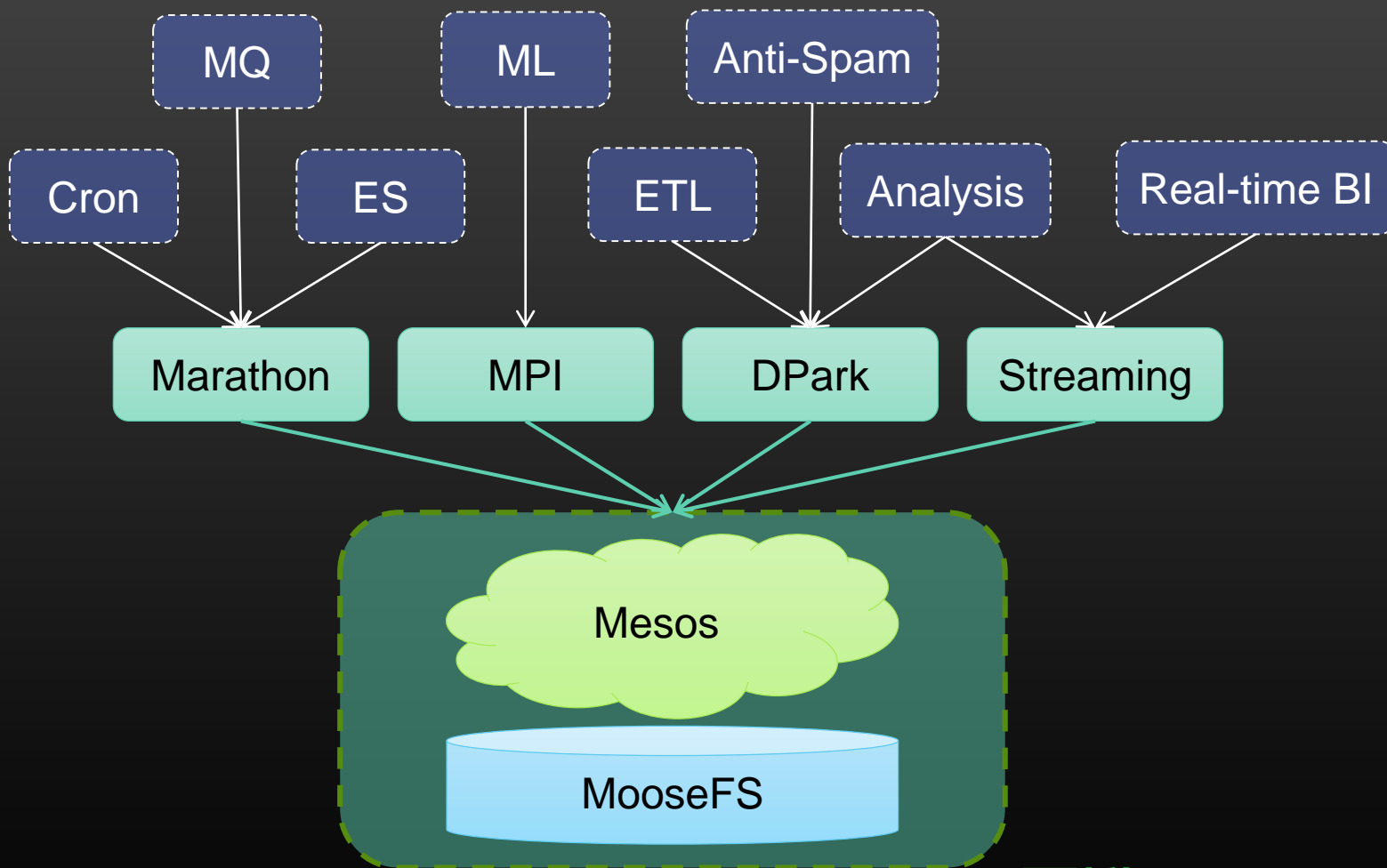
# Who Are We

- [www.douban.com](http://www.douban.com)
- Online Book/Movie/TV/Music Database & Recommender System
- SNS for People with Similar Interests
- Python as Main Language
- Founded in 2005
- ~120M Register Users
- ~240M PV per Day

# Mesos @ Douban

- 2 Clusters (0.27.2, 0.28.2)
- ~5000 CPUs, ~5TB Mem
- ~3M Tasks/Day
- ~70% Resource Utility
- Since 2011
  - Experimental: 0.3, 0.4, 0.5, 0.9
  - Intensively Patched: 0.12, 0.14, 0.15, 0.18, 0.20
  - Stable: 0.26, 0.27, 0.28

# Big Picture



# DPark

- <https://github.com/douban/dpark>
- Spark-like Framework in Python on Mesos
- Faster, Less Overhead
- Born with Multi-Tenant Support
- Higher Resource Utilization
- Integration with Python Data Analytics Stack
- Local-Aware Scheduler for MooseFS

# PyMesos

- <https://github.com/douban/pymesos>
- Pure Python Scheduler/Executor Driver Implementation for Mesos
- HTTP API (Since 0.2.0)
- Support CPython 2.7/3.5 & PyPy
- Mesos Launcher with “*subprocess*” Interface

# TFMesos

- <https://github.com/douban/tfmesos>
- Distributed Tensorflow on Mesos
- In-Graph/Between-Graph Training
- Dynamic Cluster Management
- Docker & GPU Support
- Support Mesos Containerizer

# And More...

- DStream:
  - Streaming Framework for Real-time Analytics
- MPI:
  - Hydra Based MPI Launcher on Mesos
- Paracel
  - <https://github.com/douban/paracel>
  - Parameter Server Implementation for Large Scale Distributed Machine Learning



# Issues with Mesos

- 100% Master CPU Usage
- Task Stuck in “*Staging*” Forever
- Large Task Starvation
- Role Priority?
- Mesos Preemptive Resource Allocator?
- Filter with Resource Limit?
- Backward Compatibility

# Future with Mesos

- Upgrade to Mesos 1.0.x
- Chronos ?
- Mesos-DNS ?
- Custom Preemptive Allocator ?
- Integration with In-House Image Registry
- Job Dependency/Workflow System Integration

# Thank You

