Terra Nova: Or 6 Interesting/Handwavy ideas

Adewale Oshineye

Dryad meets Erlang

- Microsoft Dryad (explicit dependency graphs on a superset of map-reduce)
- Erlang (lots of independent processes)
- The combination

DSLs on distributed systems

- Dryad has SQL and Linq interfaces
- Yahoo has implemented a DSL on top of Hadoop (Pig, Hive)

Complex Event Processing

- Applying (business) rules to streams of events in real time

- Disrupting batch-oriented enterprises (algorithmic trading of credit derivatives)

- If only we had cheap large-scale distributed systems (ECS2 + S3 + Hadoop)

- If only they were optimised for low-latency, high-bandwidth and low scaling cost

Sensor networks

- Mobile phones are sensors (MyLocation)
- Sensors in the hands of end-users generate lots of interesting data
- CEP systems could let you process and react to changes in this data

N>1

- Joe Gregorio's blog post:
  - multiple databases
  - datacentres
  - hosting sites
  - heterogenous systems (multiple versions of the same code in the estate at the same time)

N>1 consequences

- Consequences:
  - no central database
  - everything is active/active
  - CAP is a real problem (Yahoo's ZooKeeper project to fix the scaling limits of Hadoop via Paxos-based leader-election algorithms)
  - Everything breaks all the time
  - Dealing with breakages becomes a full-time job (mine)

Provisioning

- Puppet lets you buy X new machines and push out changes to all of them with 1 command

- Puppet – apt-get for everything on the machine (including config files)

- Makes dealing with physical machines (almost) as convenient as dealing with virtual machines

- If only it could do deployment for you

References

- Dryad - http://research.microsoft.com/research/sv/dryad/

- Dryad Tech Talk - http://uk.youtube.com/watch?v=WPhE5JCP2Ak

- Hadooop Pig - http://incubator.apache.org/pig/

- CEP - http://en.wikipedia.org/wiki/Complex_Event_Processing

- Sensor networks - http://en.wikipedia.org/wiki/Wireless_sensor_network

- Joe Gregorio's N>1 blog post - http://bitworking.org/news/218/N-1

- CAP conjecture - http://citeseer.ist.psu.edu/544596.html

- Yahoo ZooKeeper - http://research.yahoo.com/node/1849

- ZooKeeper project - http://zookeeper.sourceforge.net/

- Paxos - http://research.microsoft.com/users/lamport/pubs/paxos-simple.pdf

- Puppet - http://reductivelabs.com/trac/puppet