



Apache Whirr

Tom White, Cloudera, @tom_e_white
jclouds Meetup, Skype
20 January 2010

What is Whirr?

- An Apache Incubator project for running services in the cloud
- A community for sharing code and configuration for running clusters

- It is used for
 - Testing
 - Evaluation and proof of concept
 - Production (future)

Services

- Hadoop
- ZooKeeper
- Cassandra
- HBase

- Under development
 - Mesos, MongoDB, Kafka, Flume...

```
whirr.cluster-name=hadoop
```

```
whirr.instance-templates=1 nn+jt,5 dn+tt
```

```
whirr.provider=ec2
```

```
whirr.identity=${env:AWS_ACCESS_KEY_ID}
```

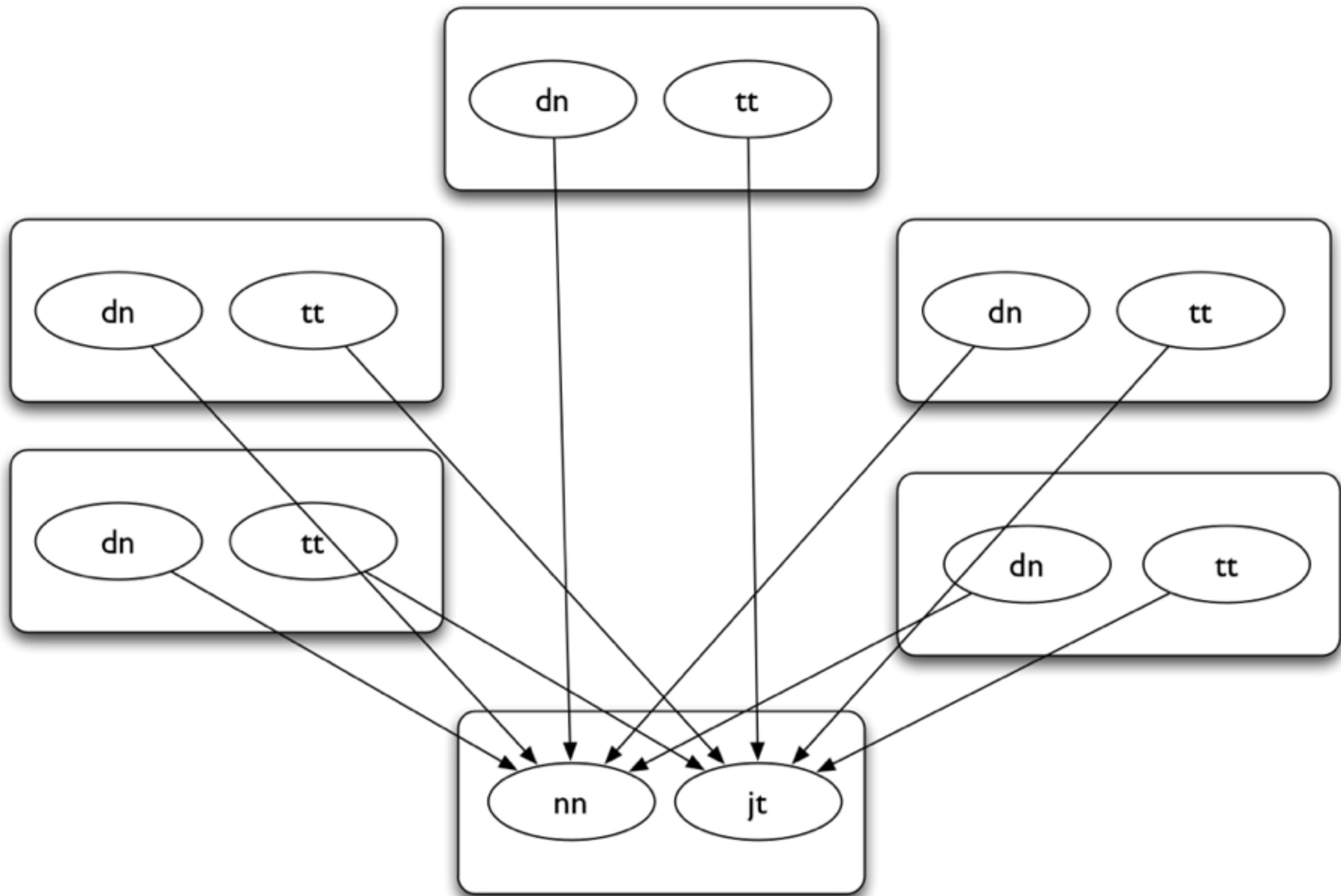
```
whirr.credential=${env:AWS_SECRET_ACCESS_KEY}
```

```
whirr.hardware-id=c1.xlarge
```

```
whirr.image-id=us-east-1/ami-da0cf8b3
```

```
whirr.location-id=us-east-1
```

```
% whirr launch-cluster --config hadoop-ec2.properties
```



Java API

```
import org.apache.commons.configuration.PropertiesConfiguration;  
import org.apache.hadoop.conf.Configuration;  
import org.apache.whirr.service.ClusterSpec;  
import org.apache.whirr.service.Service;  
import org.apache.whirr.service.hadoop.HadoopCluster;
```

```
Configuration conf =  
    new PropertiesConfiguration("hadoop-ec2.properties");  
ClusterSpec spec = new ClusterSpec(conf);  
Service s = new Service();  
Cluster cluster = s.launchCluster(spec);  
// interact with cluster  
s.destroyCluster(spec);
```

Whirr is built on jclouds

- jclouds is cloud-neutral
 - Whirr currently supports EC2 and Rackspace – more planned
- jclouds provisions nodes
 - Compute API
- jclouds provides scriptbuilder facilities
 - runurl
 - general scripts
- jclouds adds new features fast
 - local VMs (libvirt)
 - byon (“bring your own nodes”)

Join us

- Release 0.3.0 coming soon
 - HBase service
 - Composable services
- Two new committers in the last month
- Find out more at
 - <http://incubator.apache.org/whirr>
- Twitter: @tom_e_white