CASSANDRA SUMMIT 2016

Develop scalable applications with DataStax Drivers for Apache Cassandra and DataStax Enterprise

Alex Popescu & Bulat Shakirzyanov

CASSANDRA SUMMIT 2016

CASSANDRA SUMMIT 2016

DataStax Drivers

Smart clients for Apache Cassandra and DataStax Enterprise

DataStax Drivers

Java

node JS

php

python

Microsoft .NET

Goals of DataStax Drivers

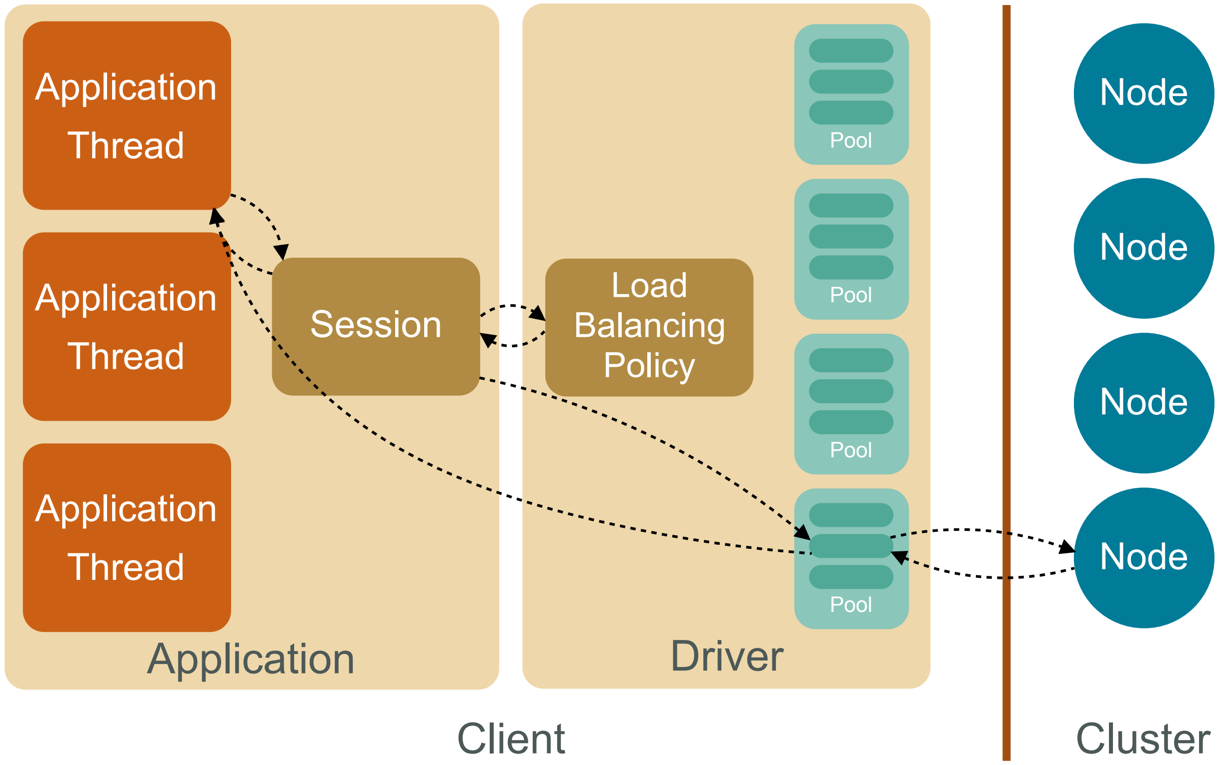
- Consistent set of features across languages
  - Asynchronous execution
  - Automatic cluster discovery
  - Connection pools and automatic reconnection
  - Load balancing
  - Fault tolerant
  - Address resolution
- Flexible to the core
- Consistent terminology
  - Cluster -> Session -> PreparedStatement & Statement -> Future or ResultSet

CASSANDRA SUMMIT 2016

Load Balancing

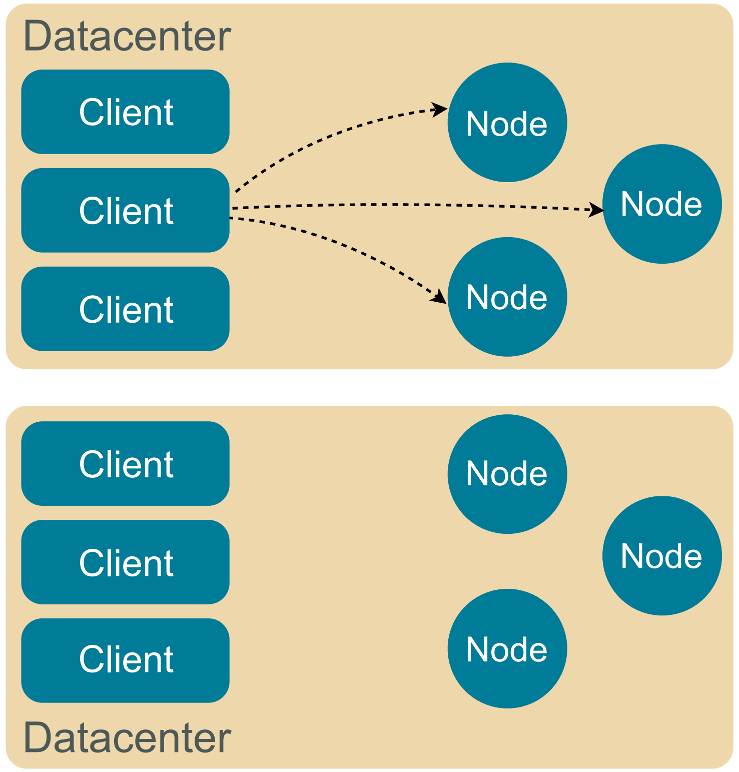
Principles and Implementations

Load Balancing



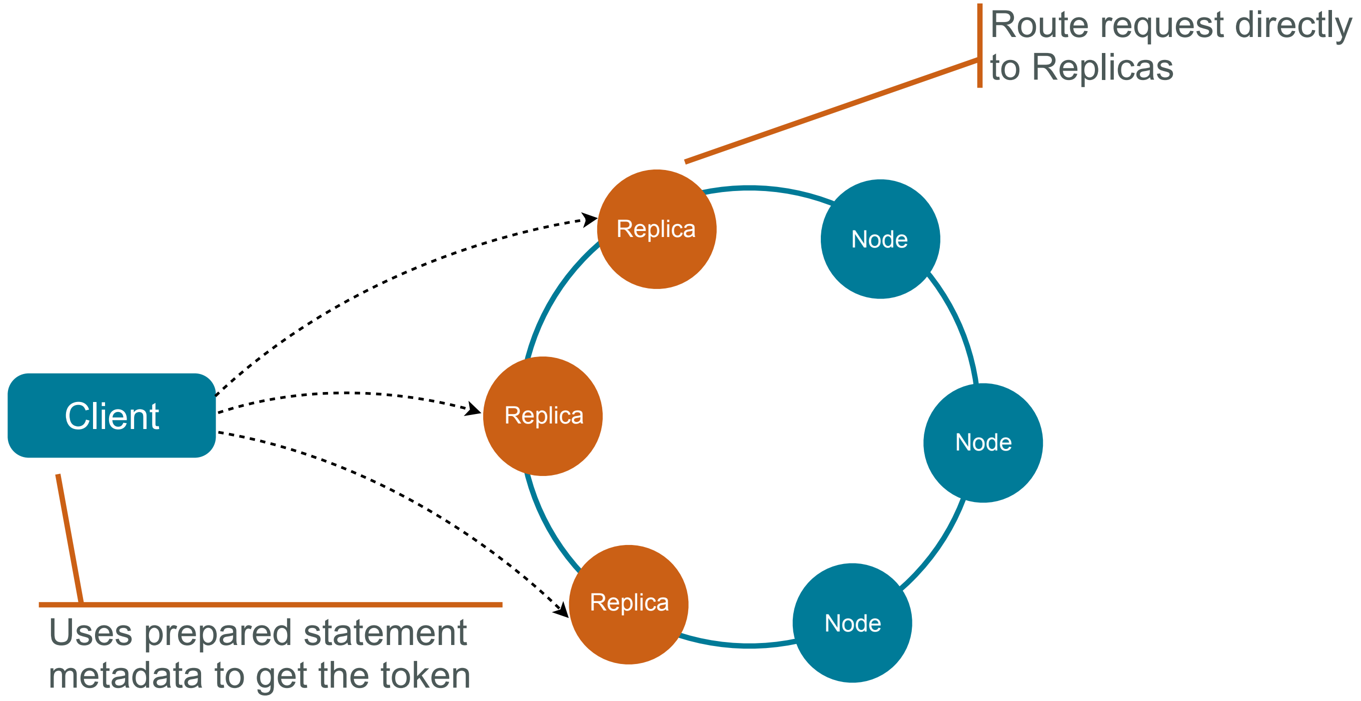
CASSANDRA SUMMIT 2016

DataCenter Aware Balancing



Local nodes are queried first, if none are available, the request could be sent to a remote node.

Token Aware Balancing



Route request directly to Replicas

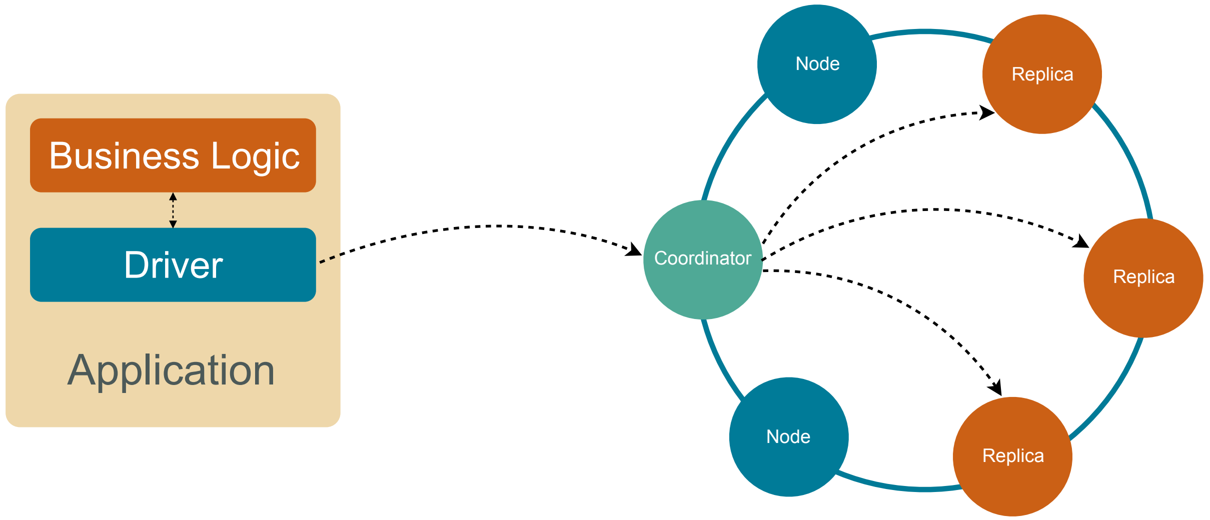
Uses prepared statement metadata to get the token

CASSANDRA SUMMIT 2016

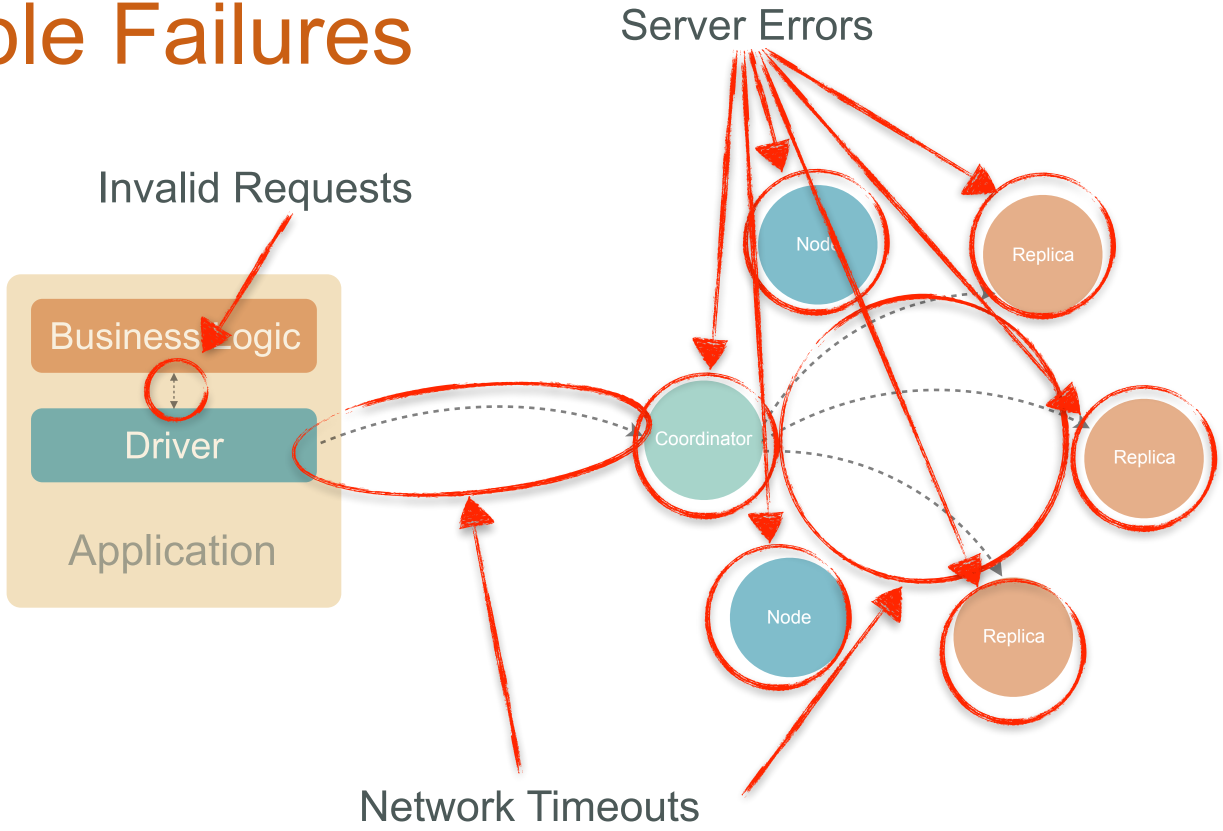
Fault Tolerance

Sources of Failure and Error Handling

Fault Tolerance

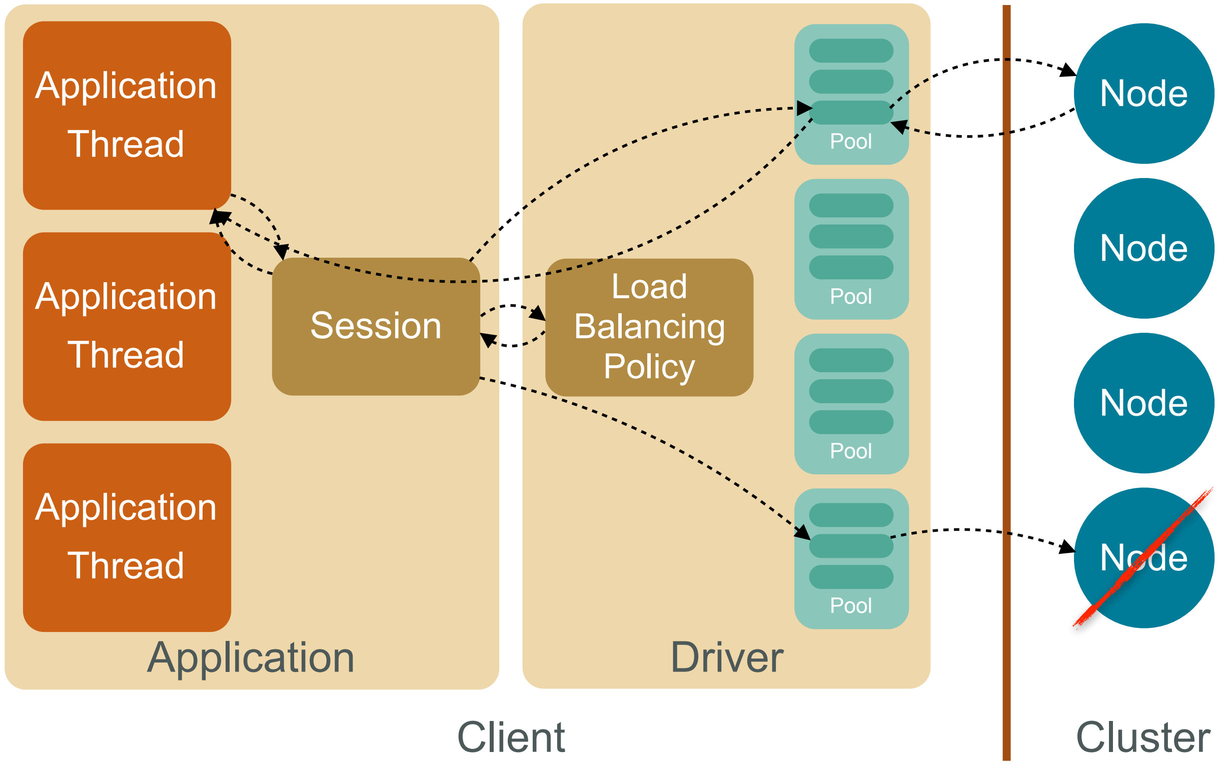


Possible Failures



CASSANDRA SUMMIT 2016

Automatic Retry of Server Errors



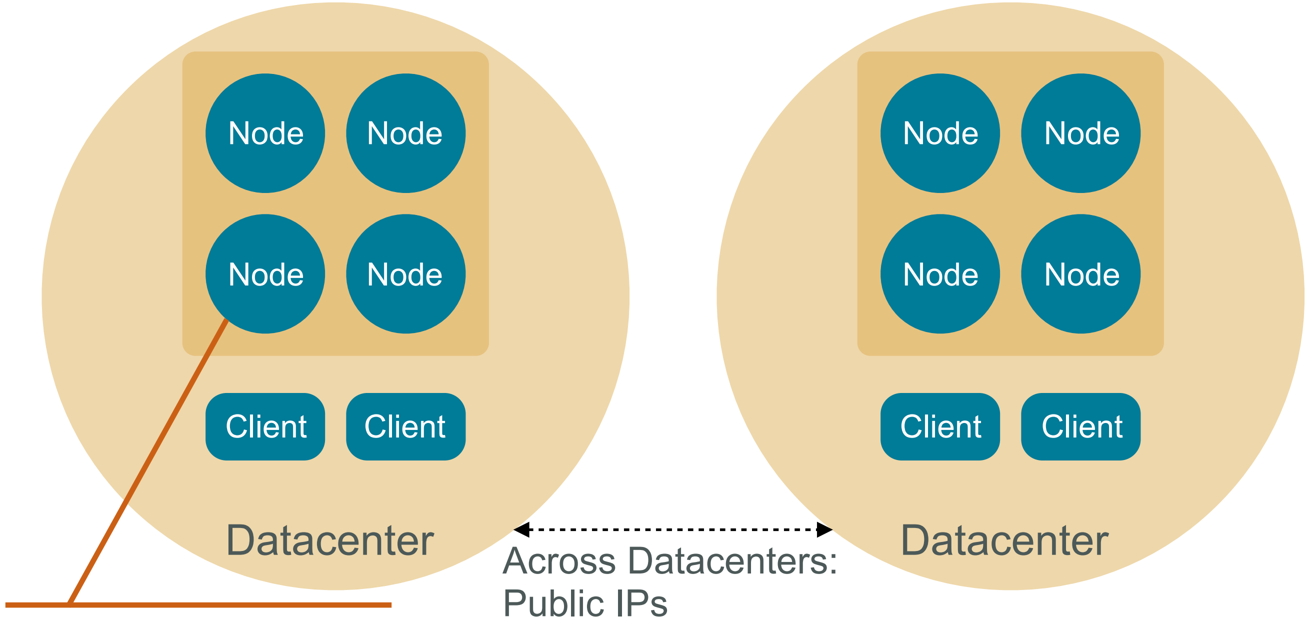
CASSANDRA SUMMIT 2016

CASSANDRA SUMMIT 2016

Address Resolution

Topology Aware Client

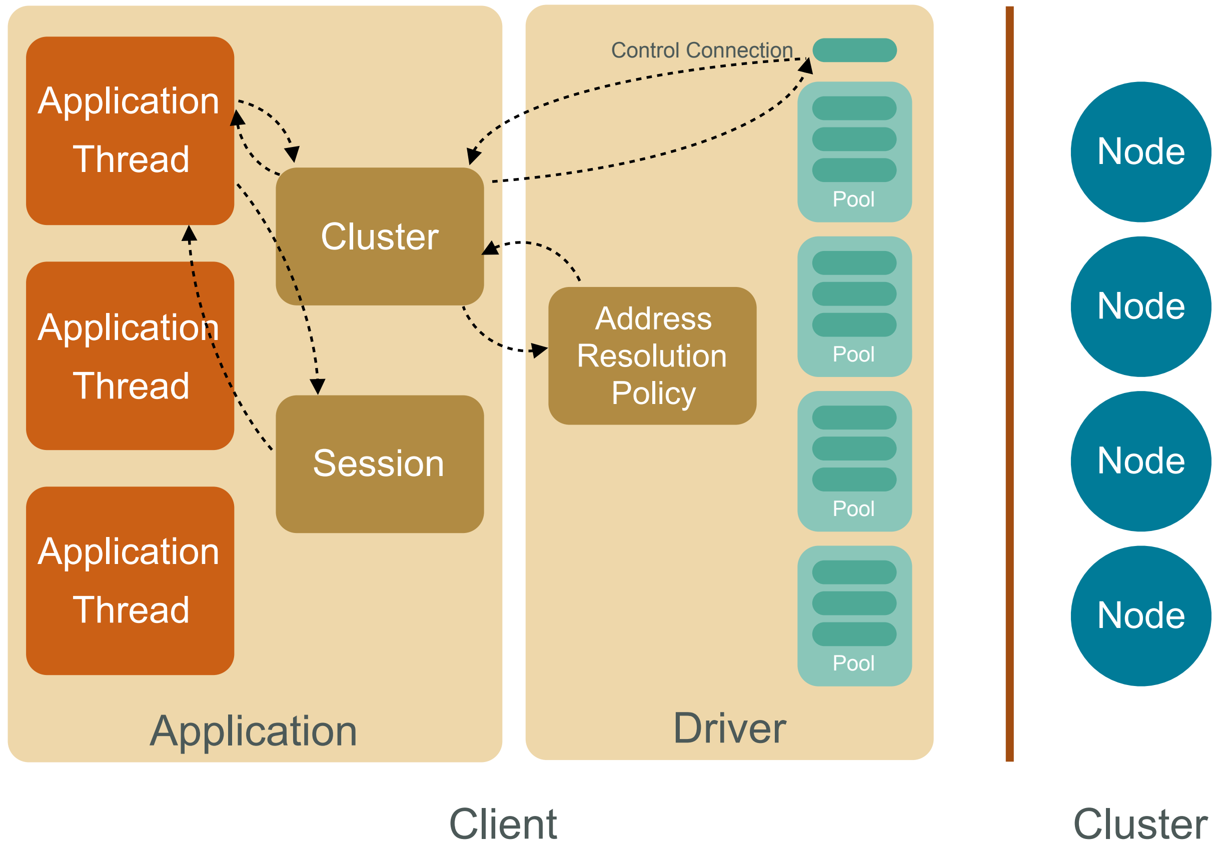
Multiple Addresses



Across Datacenters:
Public IPs

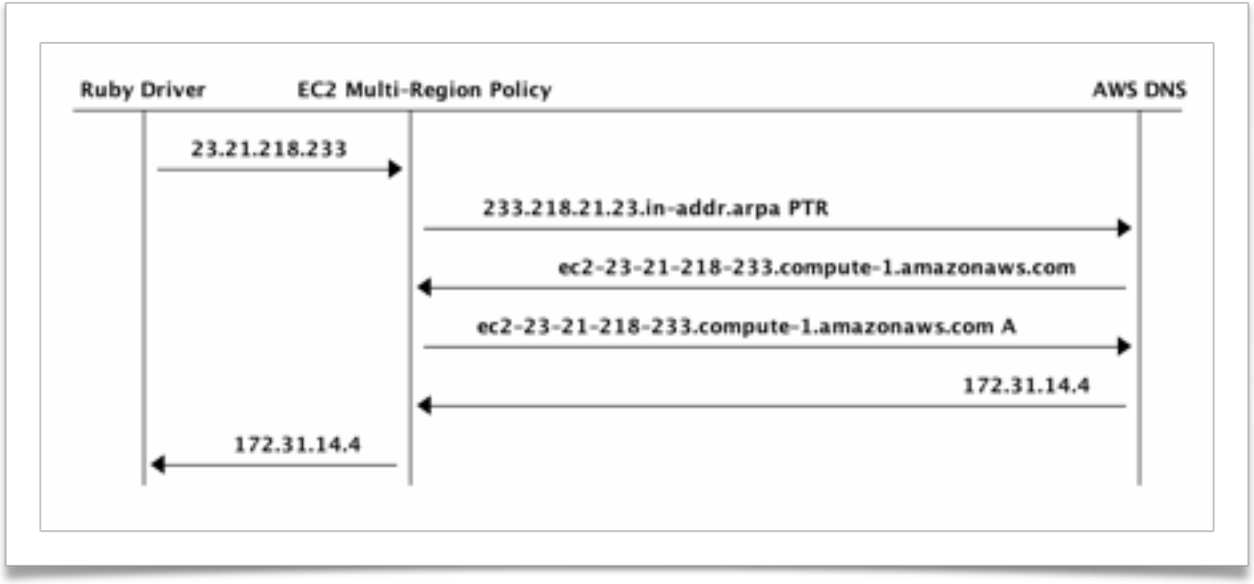
Within Datacenter:
Private IPs

Address Resolution



CASSANDRA SUMMIT 2016

EC2 Multi-Region Address Resolution


Ruby Driver
EC2 Multi-Region Policy
AWS DNS
23.21.218.233
233.218.21.23.in-addr.arpa PTR
ec2-23-21-218-233.compute-1.amazonaws.com
ec2-23-21-218-233.compute-1.amazonaws.com A
172.31.14.4
172.31.14.4

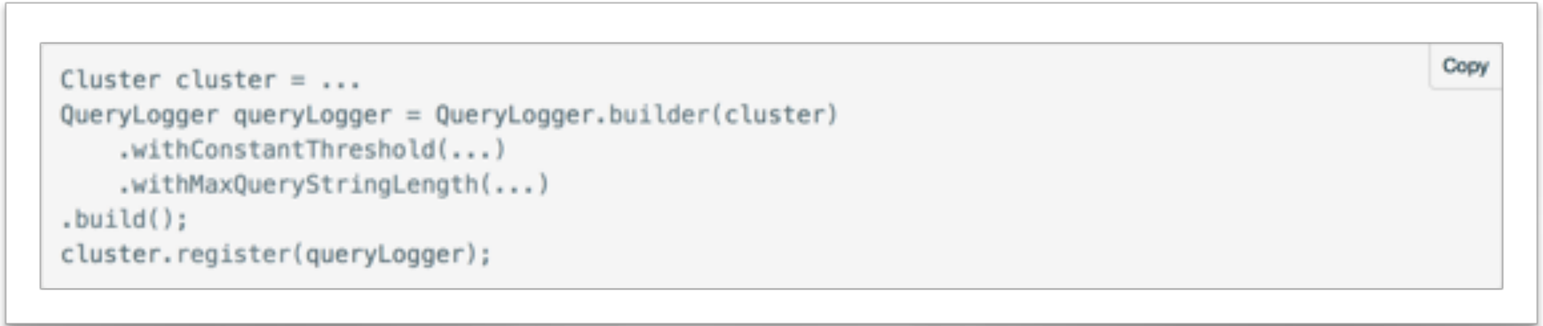
CASSANDRA SUMMIT 2016

CASSANDRA SUMMIT 2016

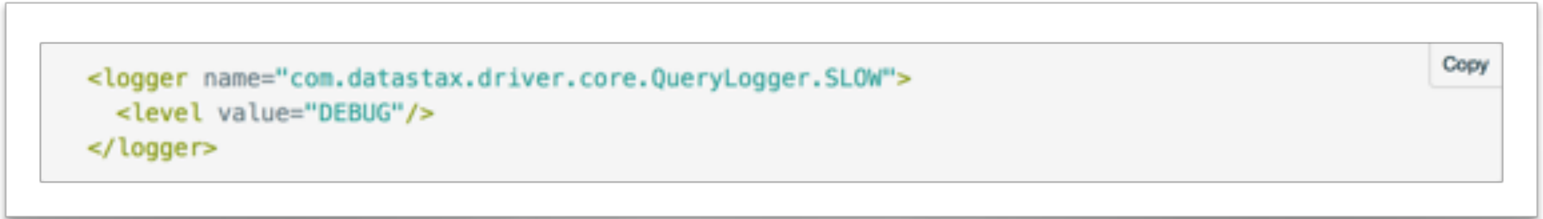
One more thing…

Sneak peek of new and upcoming driver features

Sneak peek: Slow query log



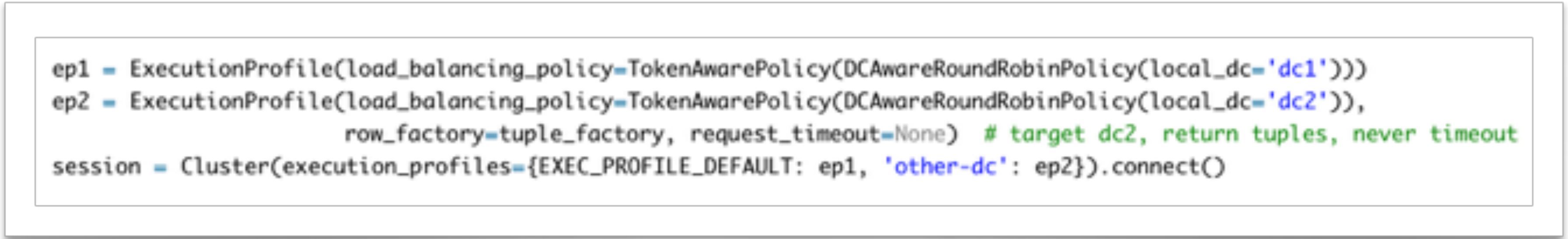
```java
Cluster cluster = ...
QueryLogger queryLogger = QueryLogger.builder(cluster)
    .withConstantThreshold(...)
    .withMaxQueryStringLength(...)
.build();
cluster.register(queryLogger);
```



```xml
  <logger name="com.datastax.driver.core.QueryLogger.SLOW">
    <level value="DEBUG"/>
  </logger>
```

CASSANDRA SUMMIT 2016

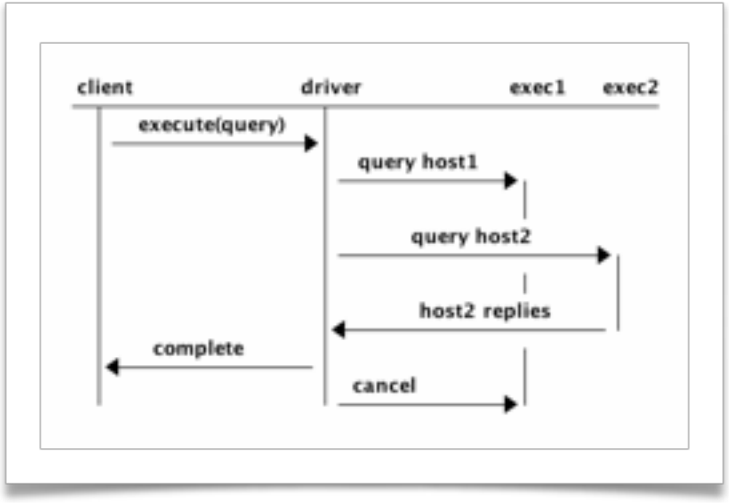
Sneak peek: Execution profiles

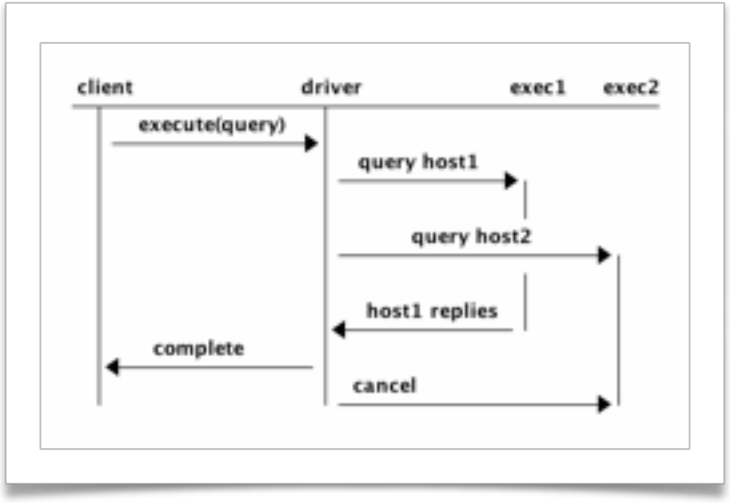


```
ep1 = ExecutionProfile(load_balancing_policy=TokenAwarePolicy(DCAwareRoundRobinPolicy(local_dc='dc1')))
ep2 = ExecutionProfile(load_balancing_policy=TokenAwarePolicy(DCAwareRoundRobinPolicy(local_dc='dc2')),
                       row_factory=tuple_factory, request_timeout=None)  # target dc2, return tuples, never timeout
session = Cluster(execution_profiles={EXEC_PROFILE_DEFAULT: ep1, 'other-dc': ep2}).connect()
```

CASSANDRA SUMMIT 2016

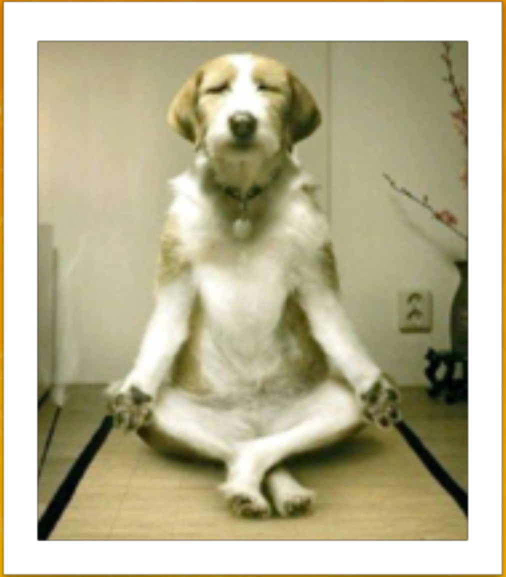
Sneak peak: Speculative retries





CASSANDRA SUMMIT 2016

Q&A



Thank you!