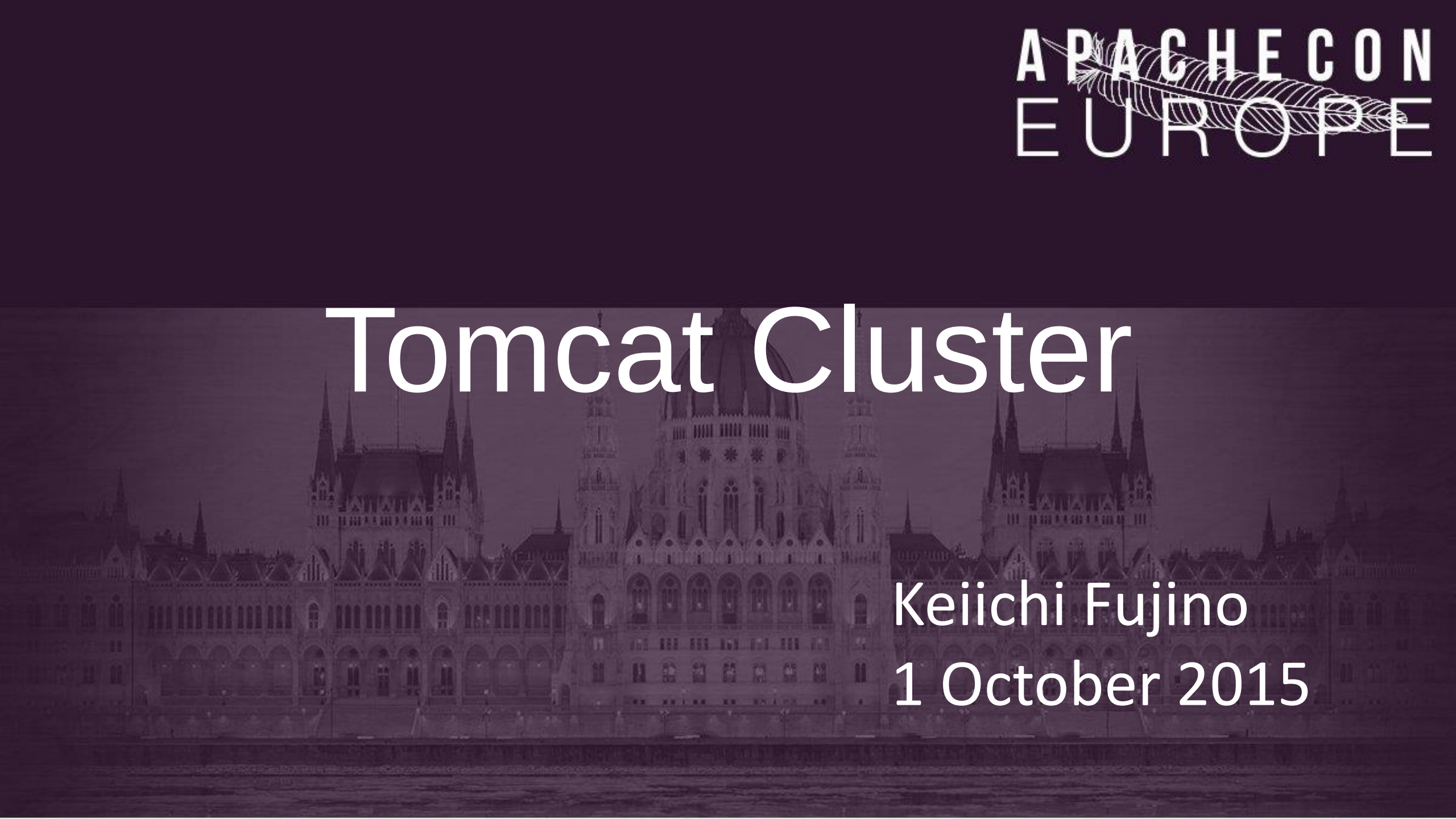


Tomcat Cluster

Keiichi Fujino

1 October 2015

Agenda

- About me
- Tomcat Clustering Overview
- Session Replication
- Channel Component
- Other Cluster features(If time remains)

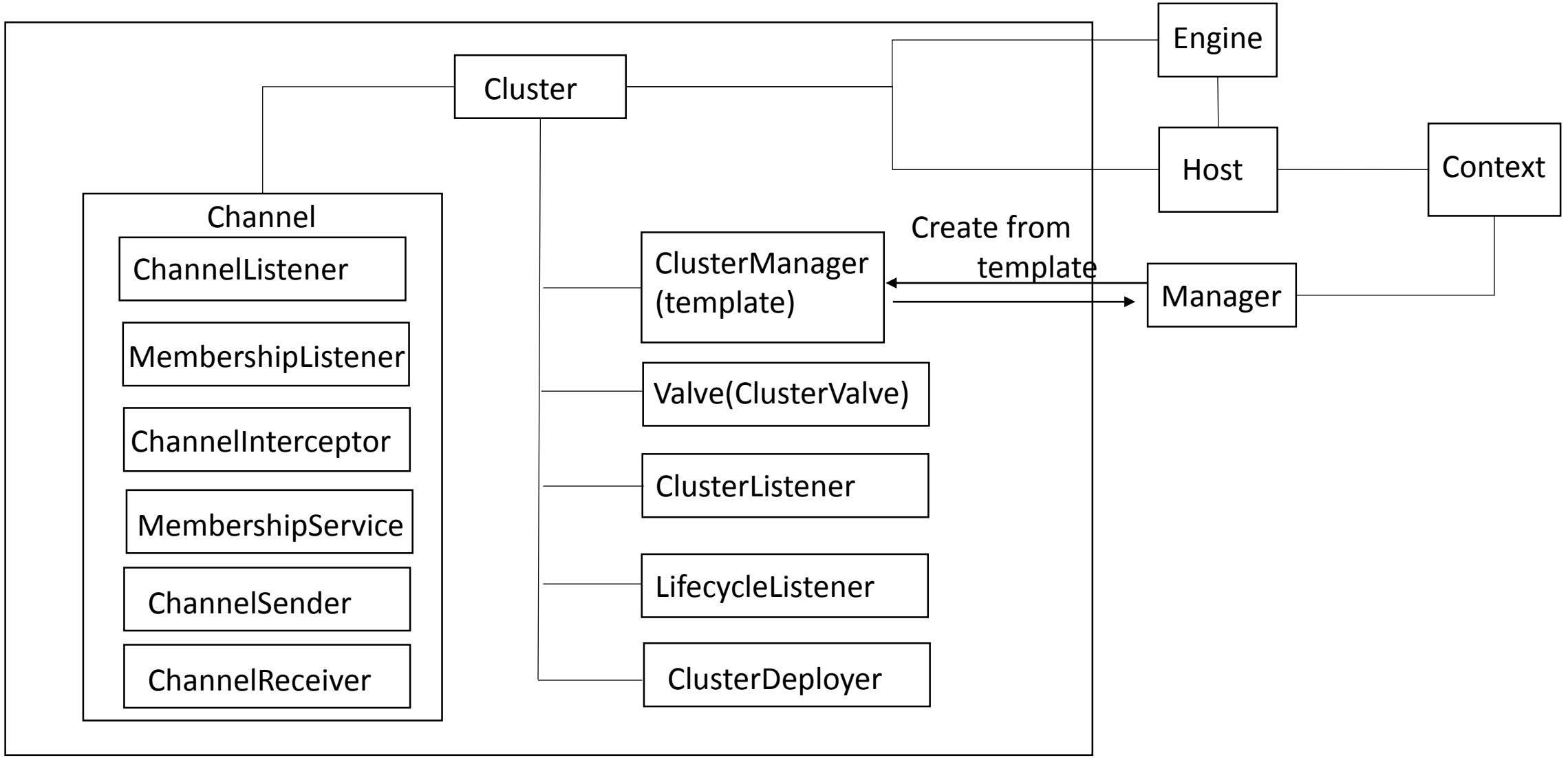
About me

- Kanagawa/Japan
- Software Developer since 2002
- Apache Tomcat committer since 2010
  - kfujino@apache.org

Clustering Overview

- What is Cluster?
  - Performance improvement
  - High availability
- Tomcat Clustering
  - Cluster membership
  - Session Replication
- Load balancing is not a Tomcat feature
  - Use mod_jk / mod_proxy_balancer

Cluster Architecture



Cluster Architecture

- Cluster
  - The main component of Tomcat Cluster
- Cluster Manager
  - The session manager for the session replication
- Valve(Cluster Valve)
  - The same as usual Valve
  - Added to the request processing pipeline automatically

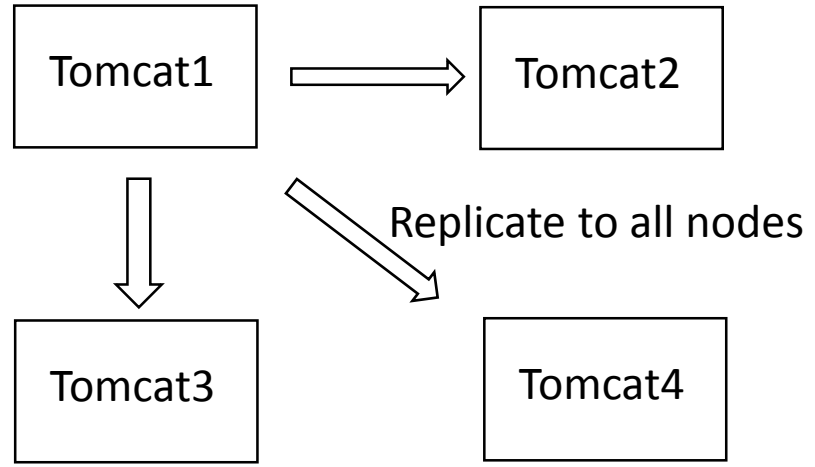
Cluster Architecture

- Cluster Deployer
  - Sharing of WAR files among cluster nodes
- ClusterListener, LifecycleListener
  - Listen cluster messages and events
- Channel
  - Performs messaging and grouping among the cluster nodes

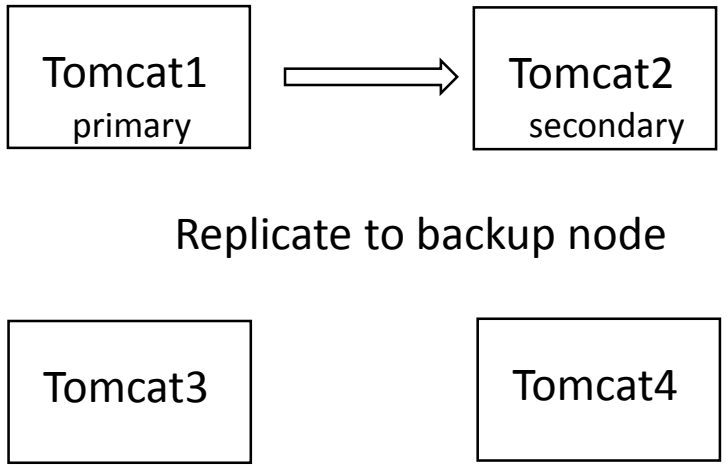
Session Replication

Session replication implementations

- All-to-All session replication
  - DeltaManager (Default)



- Primary-Secondary session replication
  - BackupManager



Use constraints

- sticky session
  - If you use the BackupManager, This is required
- Triggered session create/expire & set/removeAttribute methods
- Make sure that your web.xml describe the <distributable/> element
- Session attributes must implement java.io.Serializable
  - If some of session attributes does not implement java.io.Serializable, you should use the sessionAttributeFilter attribute

How to configure

- Configure Cluster Manager
  - DeltaManager or BackupManager
- Configure Channel components
- Enable org.apache.catalina.ha.tcp.ReplicationValve
- Enable org.apache.catalina.ha.session.ClusterSessionListener
  - DeltaManager only

Delta Replication

Delta Replication

- Replicate only the changes of session
  - Not all session data
- Replicate all changes of session at the time of end of request
  - Not replication per change of session

Delta Replication

Register Delta Info

- *Add ATTRIBUTE(*Attr_A, Value_A)
- *Add ATTRIBUTE(*Attr_B, Value_B)
- *Remove ATTRIBUTE(*Attr_A)
- *Add ATTRIBUTE(*Attr_B, Value_BB)

Default : recordAllActions=false

TYPE	ACTION	NAME	VALUE
~~*ATTRIBUTE*~~	~~*SET*~~	~~Attr_A~~	~~Value_A~~
~~*ATTRIBUTE*~~	~~*SET*~~	~~Attr_B~~	~~Value_B~~
ATTRIBUTE	*REMOVE*	Attr_A	null
ATTRIBUTE	*SET*	Attr_B	Value_BB

recordAllActions=true

TYPE	ACTION	NAME	VALUE
ATTRIBUTE	*SET*	Attr_A	Value_A
ATTRIBUTE	*SET*	Attr_B	Value_B
ATTRIBUTE	*REMOVE*	Attr_A	null
ATTRIBUTE	*SET*	Attr_B	Value_BB

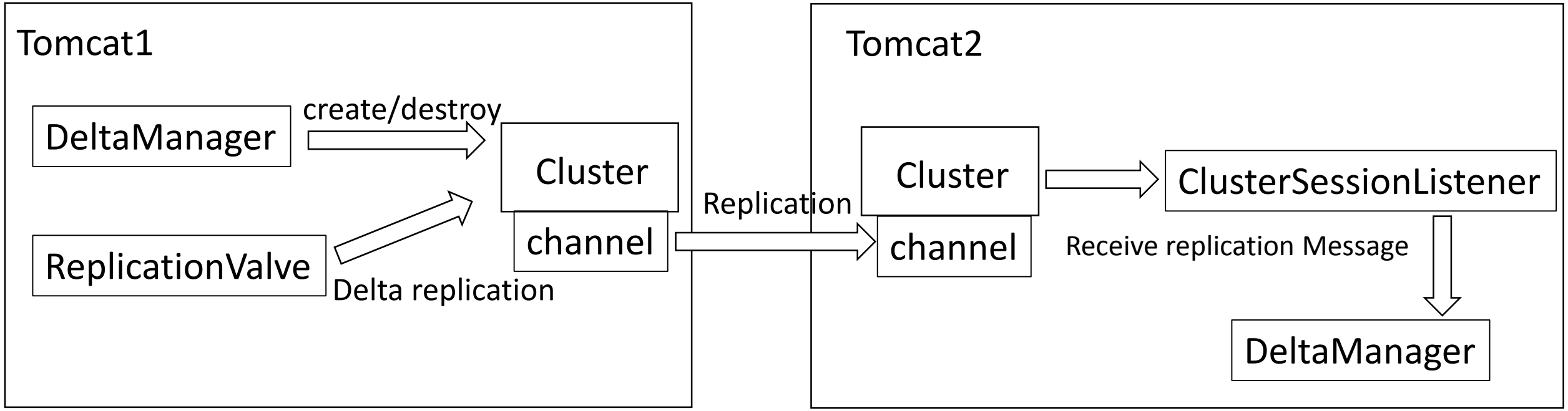
Session Replication

DeltaManager

- All-to-All session replication
- Default Session Manager in Cluster environment
- For small cluster

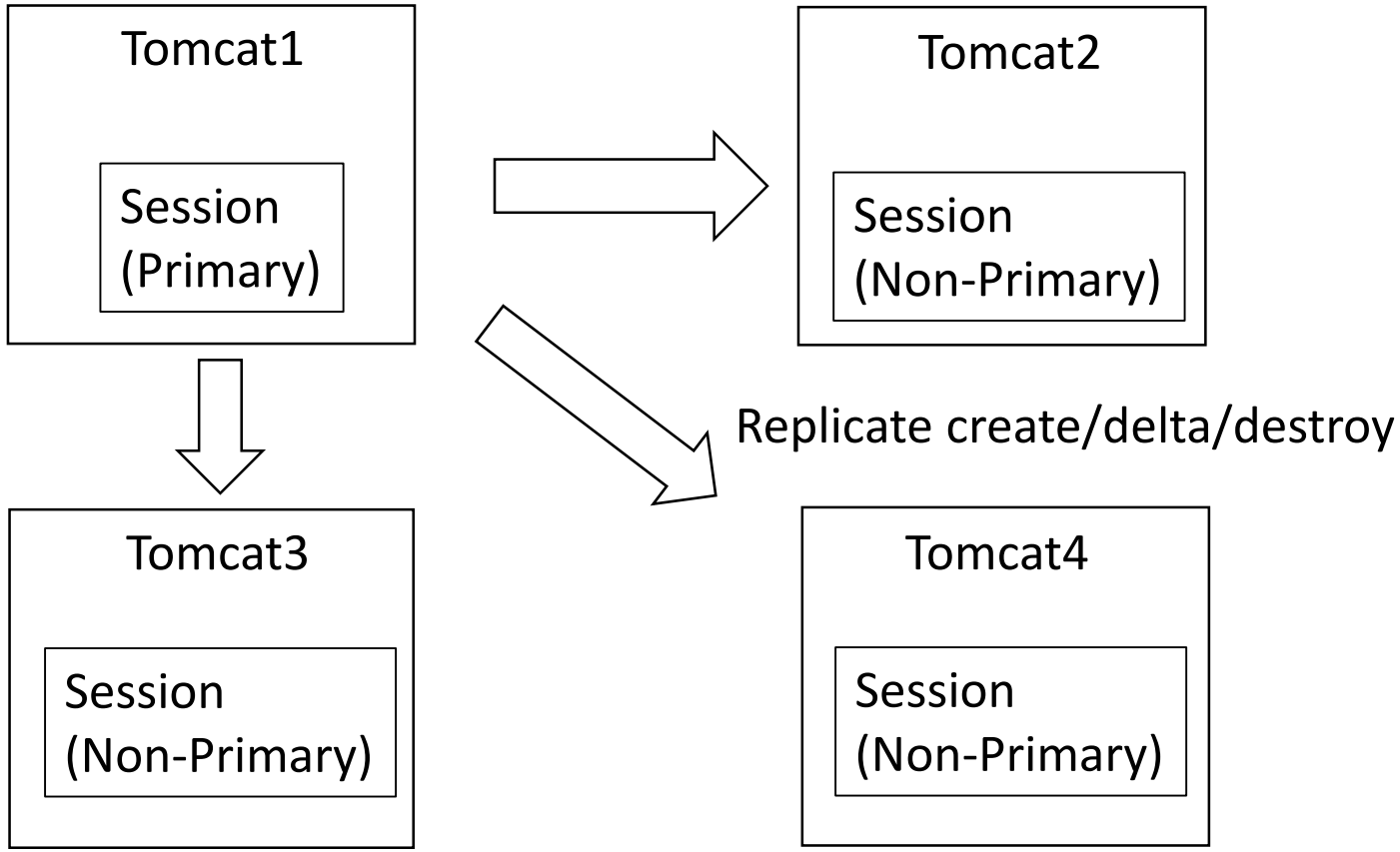
DeltaManager

Architecture



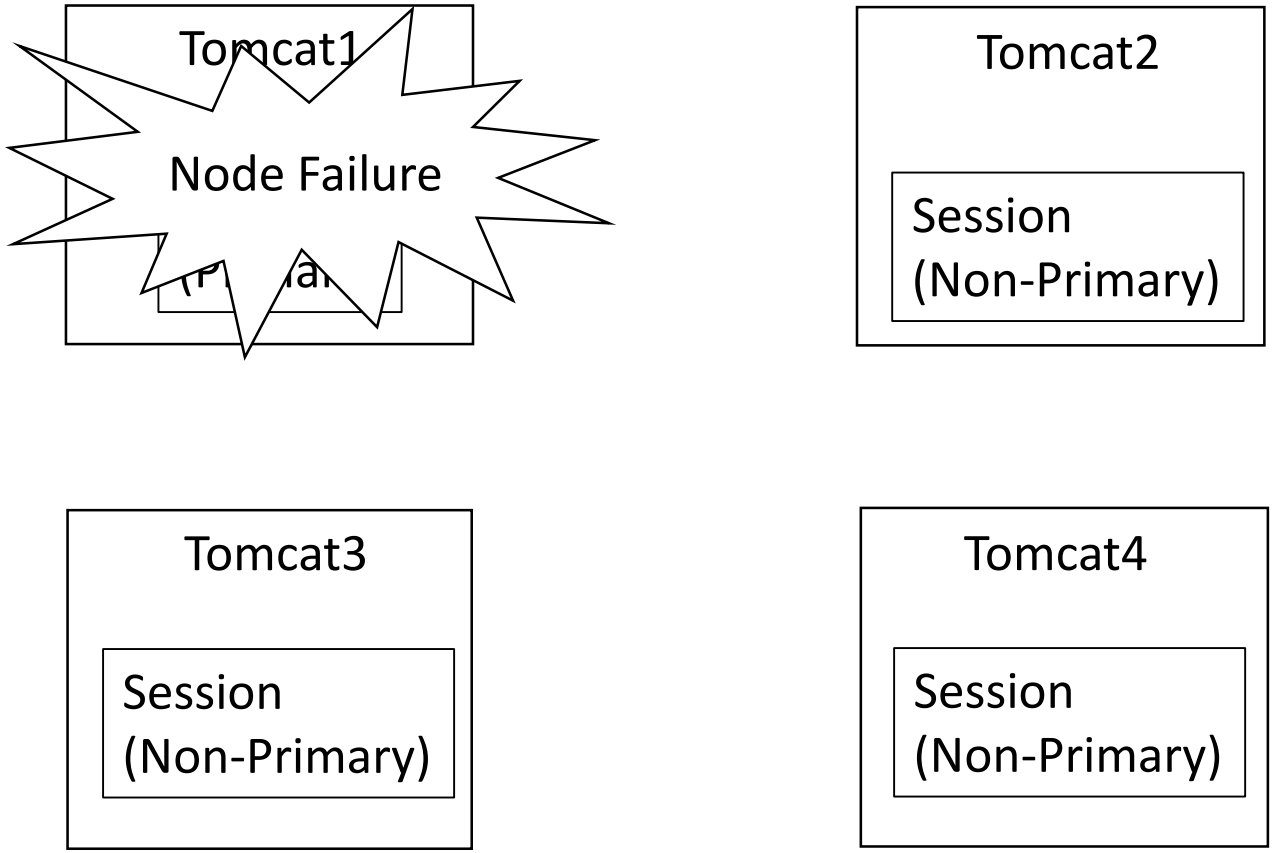
DeltaManager

Behavior



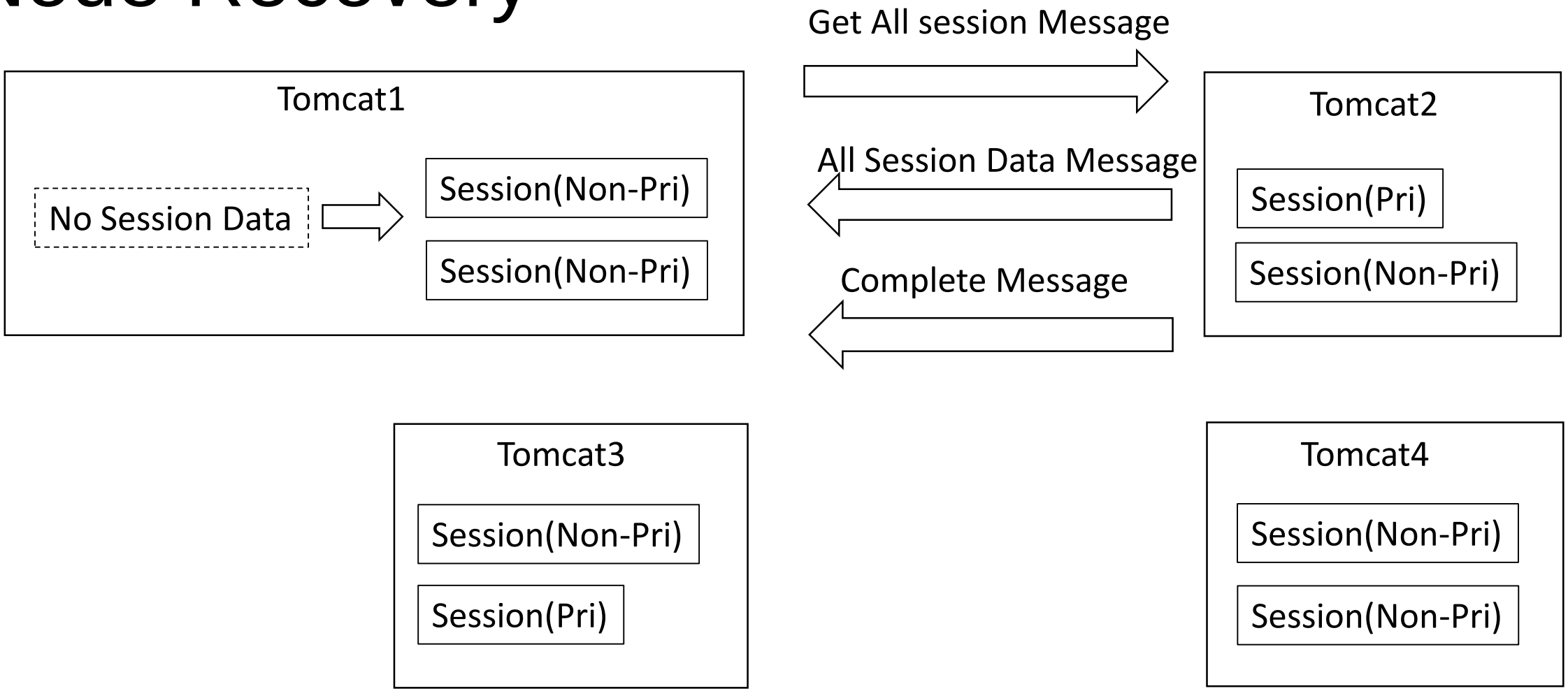
DeltaManager

Node Failure



DeltaManager

Node Recovery



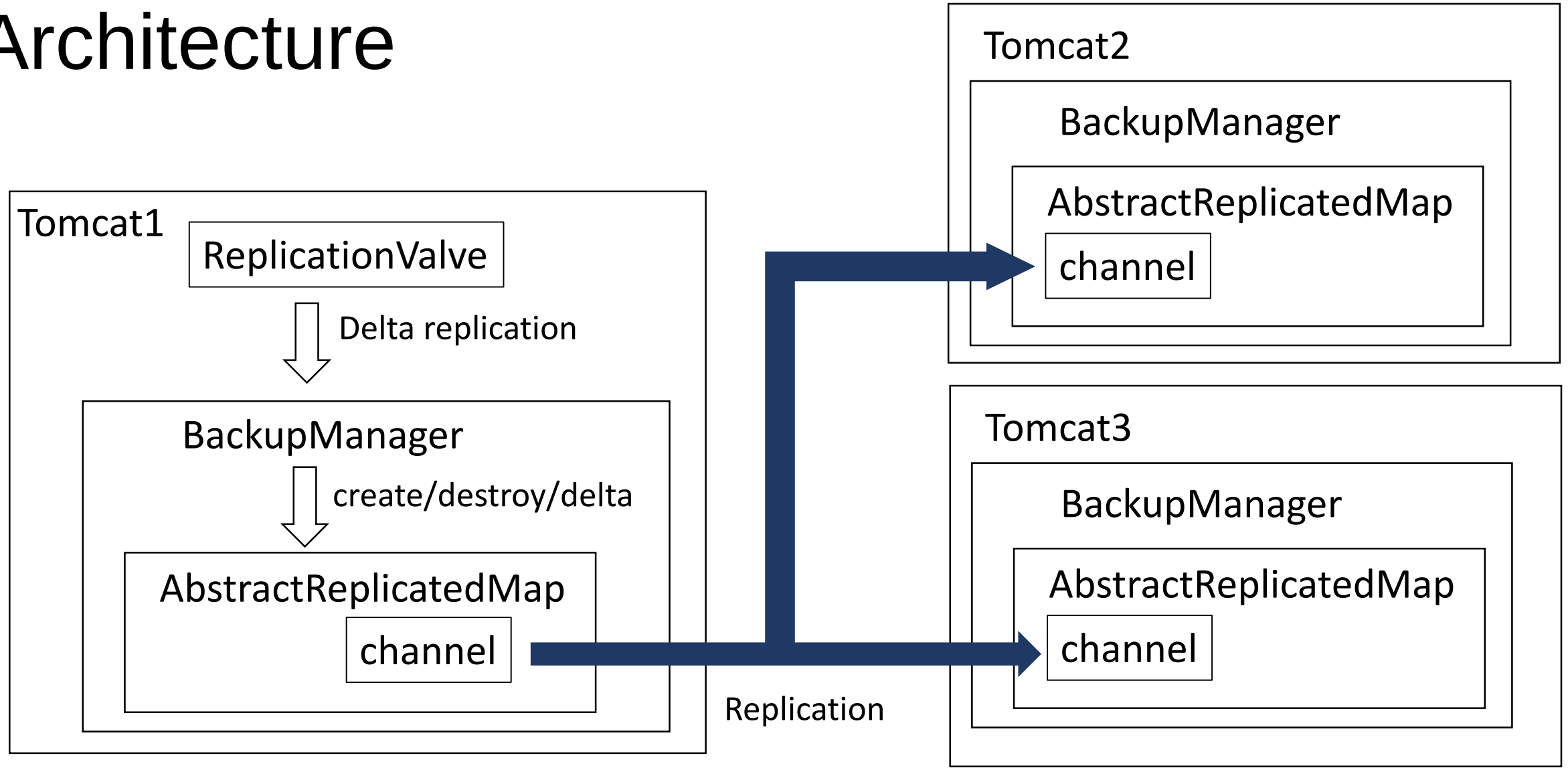
Session Replication

BackupManager

- Primary-Secondary session replication
- For large cluster

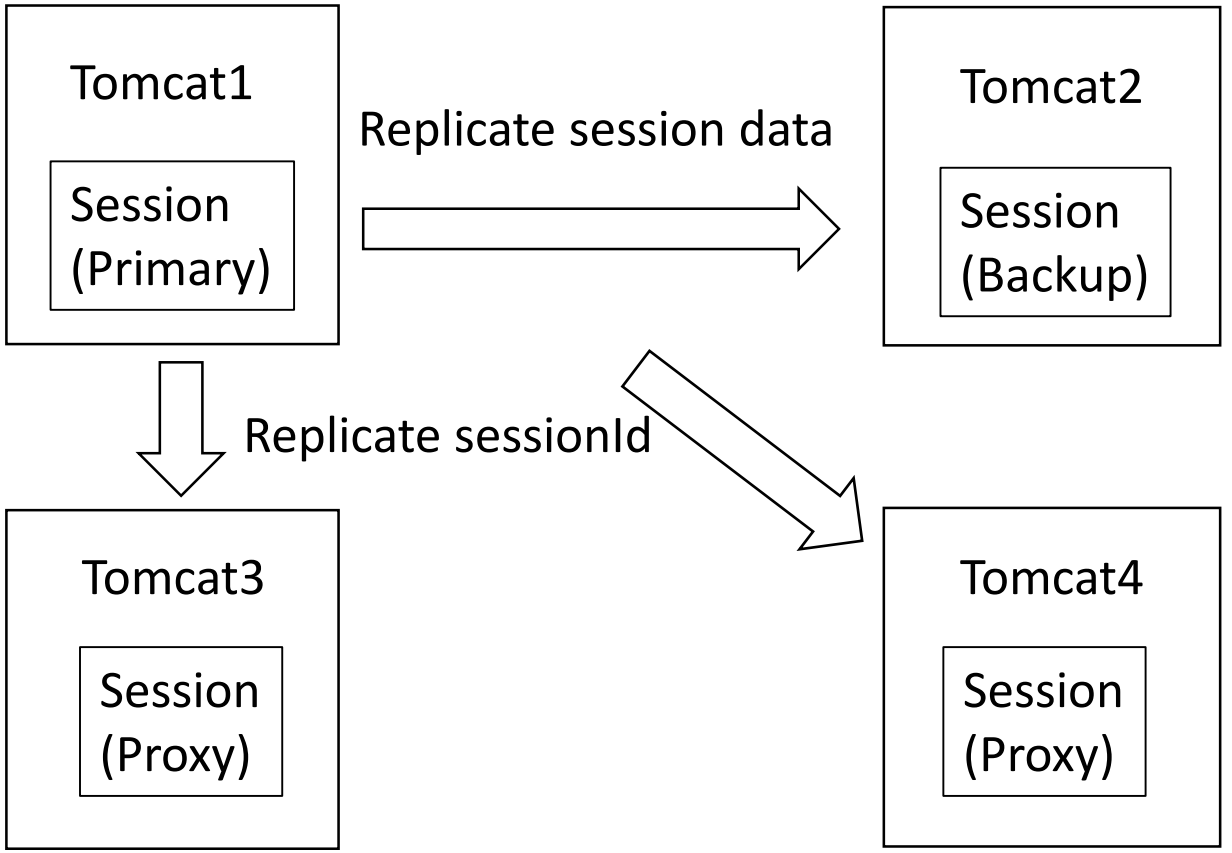
BackupManager

Architecture



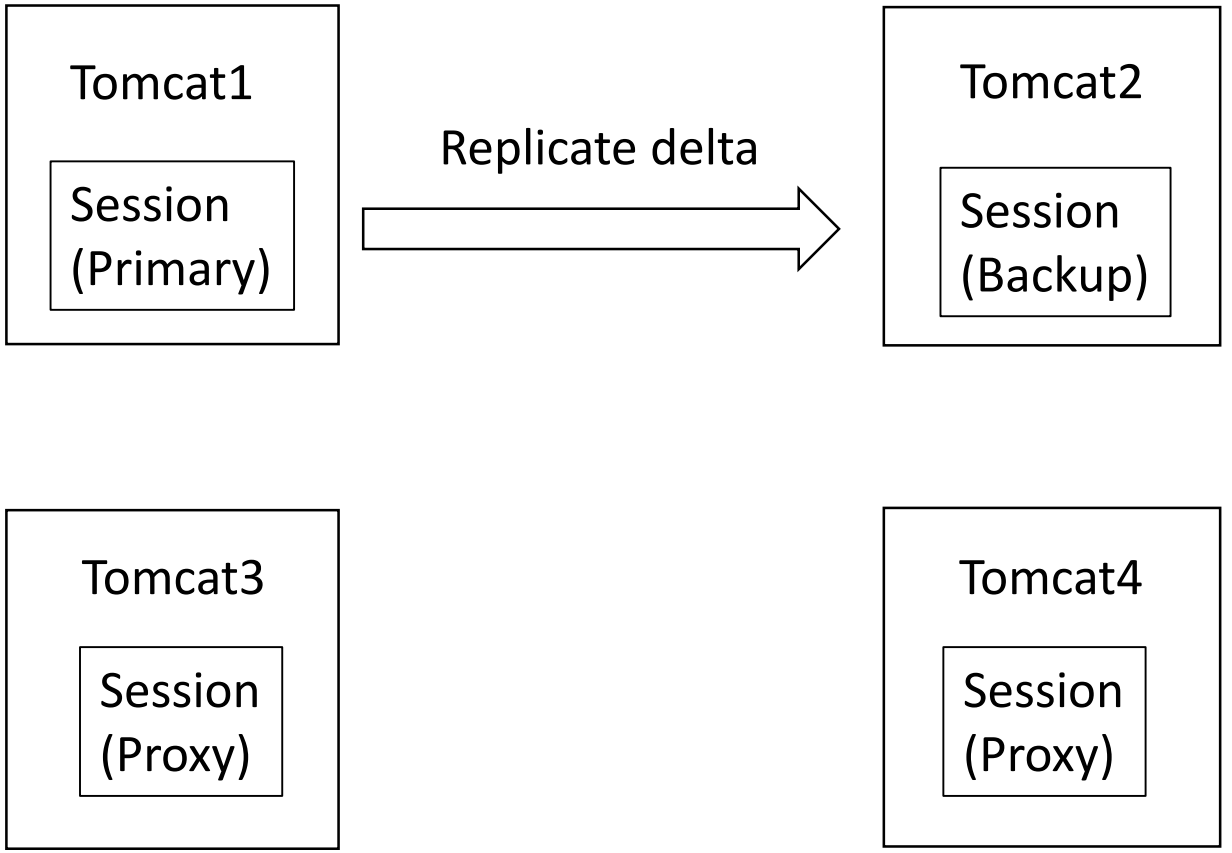
BackupManager

Behavior-create



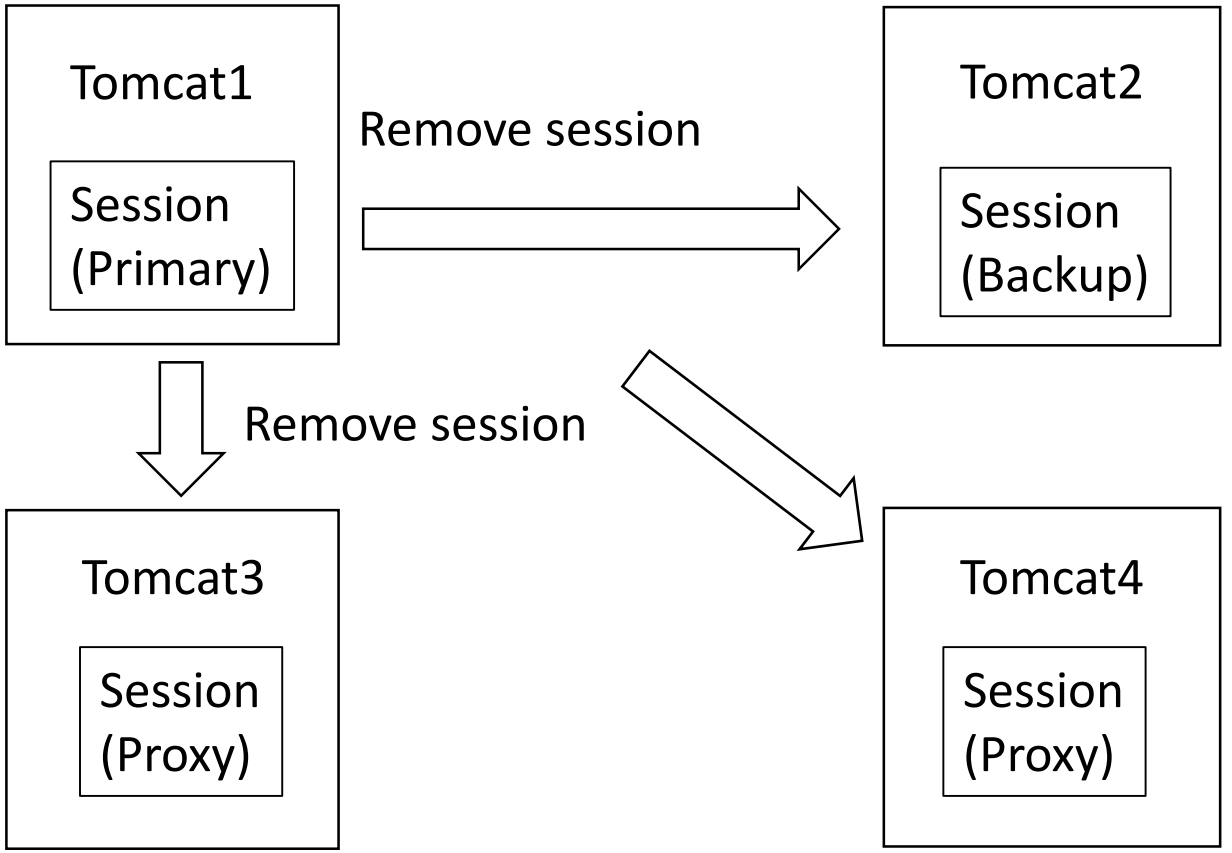
BackupManager

Behavior-delta



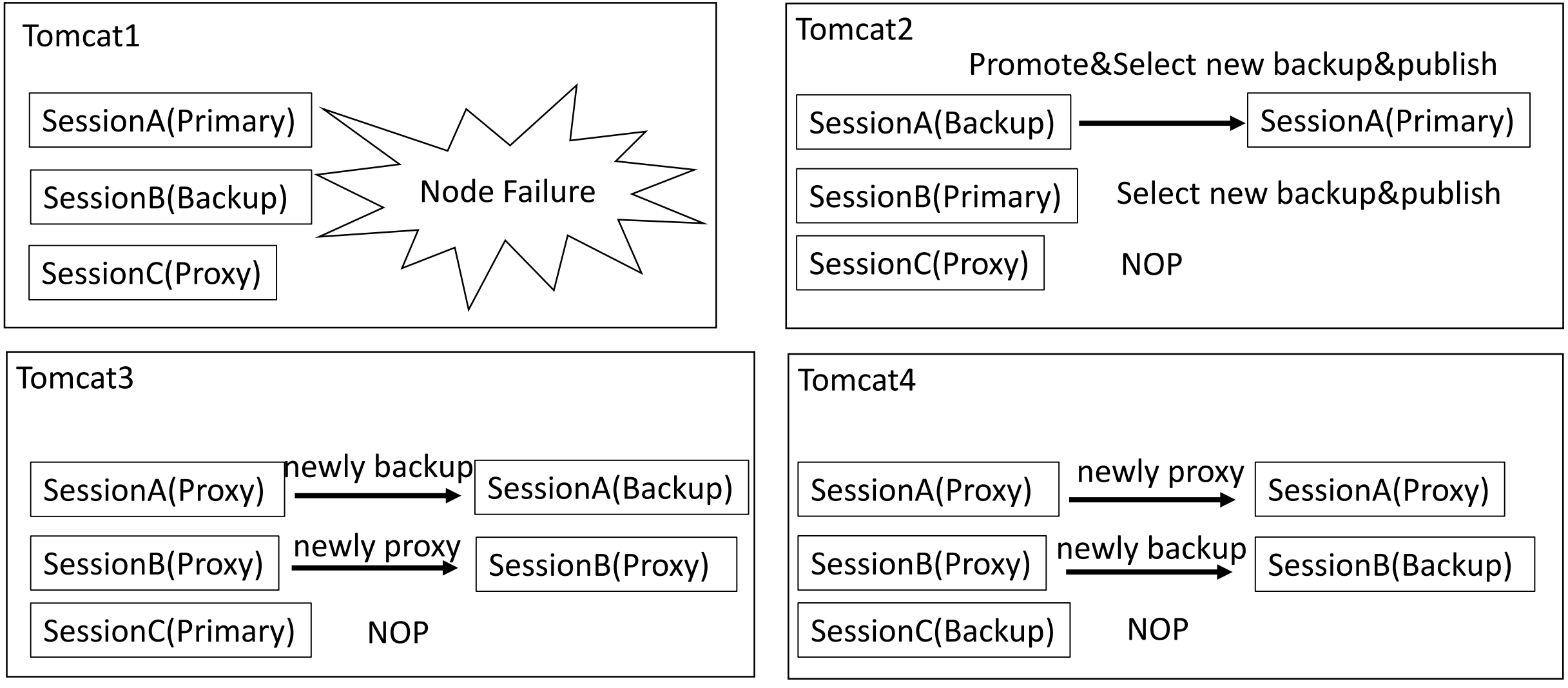
BackupManager

Behavior-destory



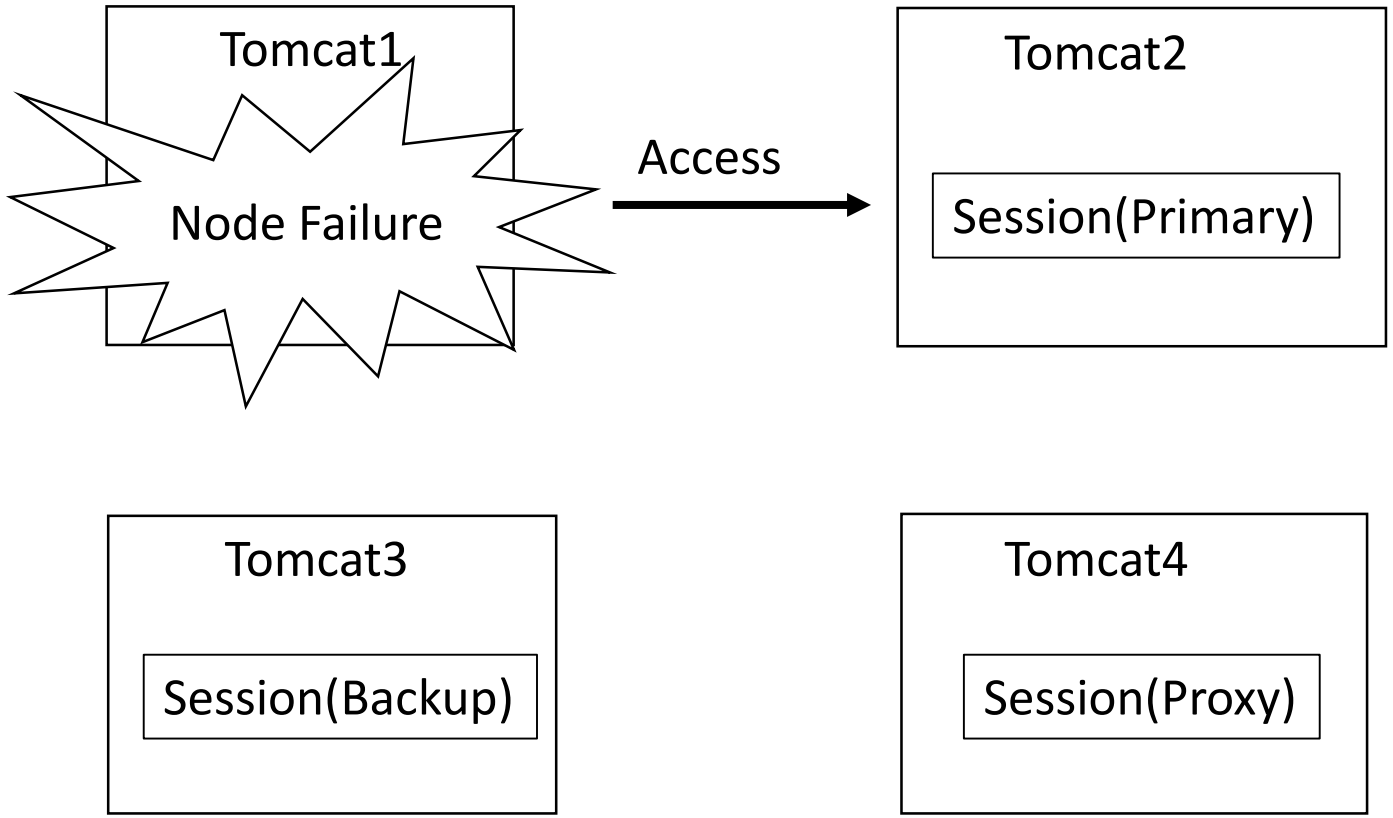
BackupManager

Node Failure



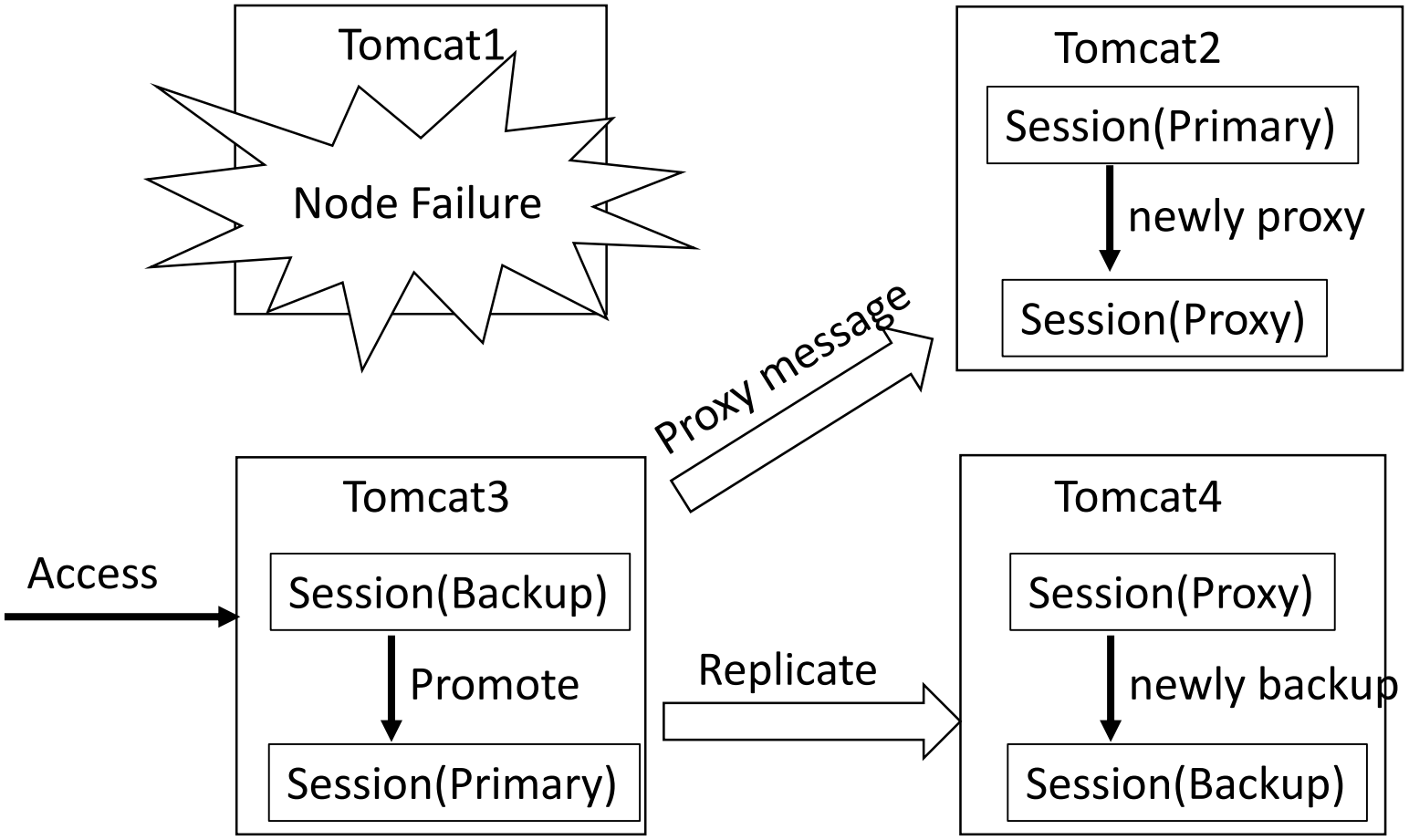
BackupManager

Access to Primary node



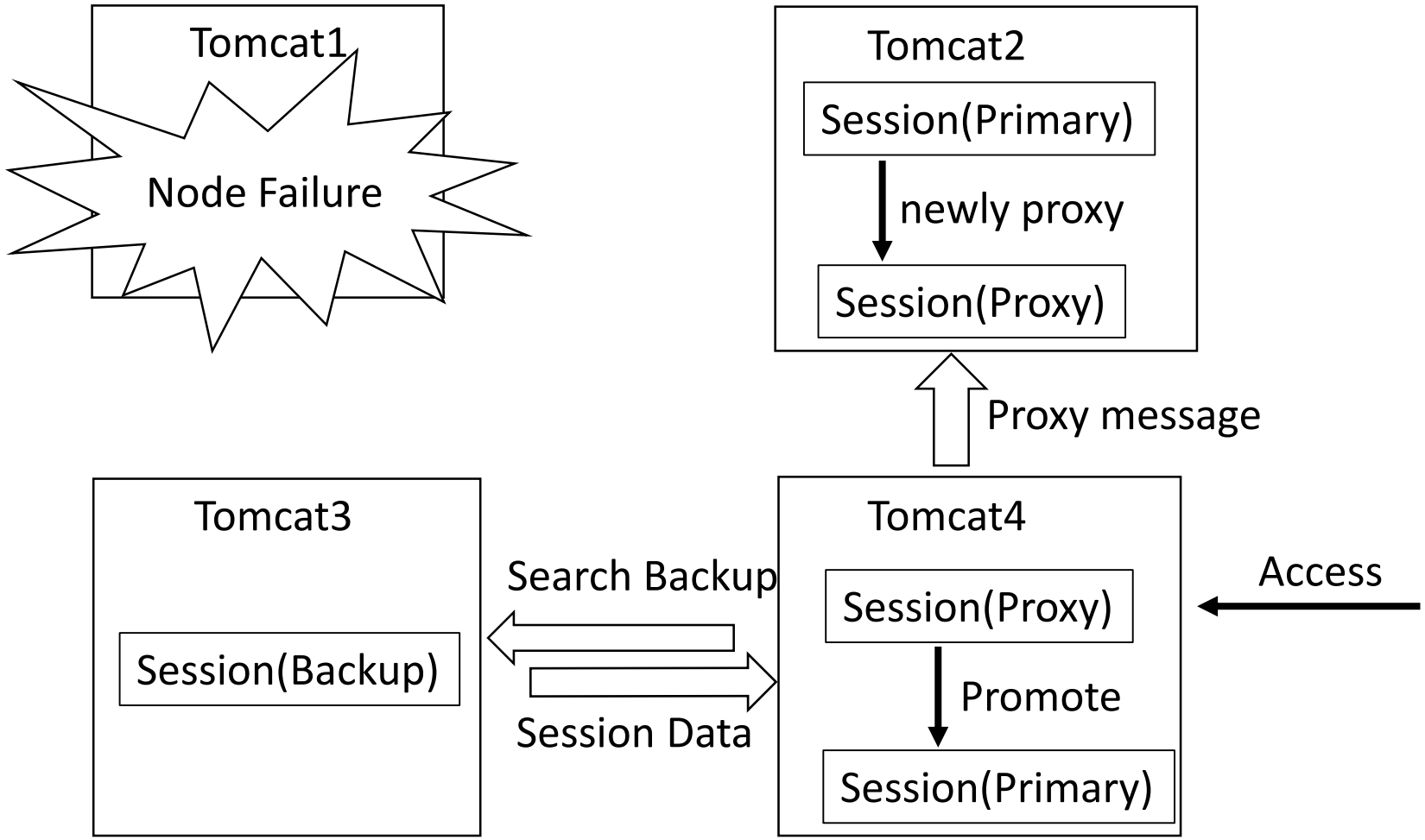
BackupManager

Access to Backup node



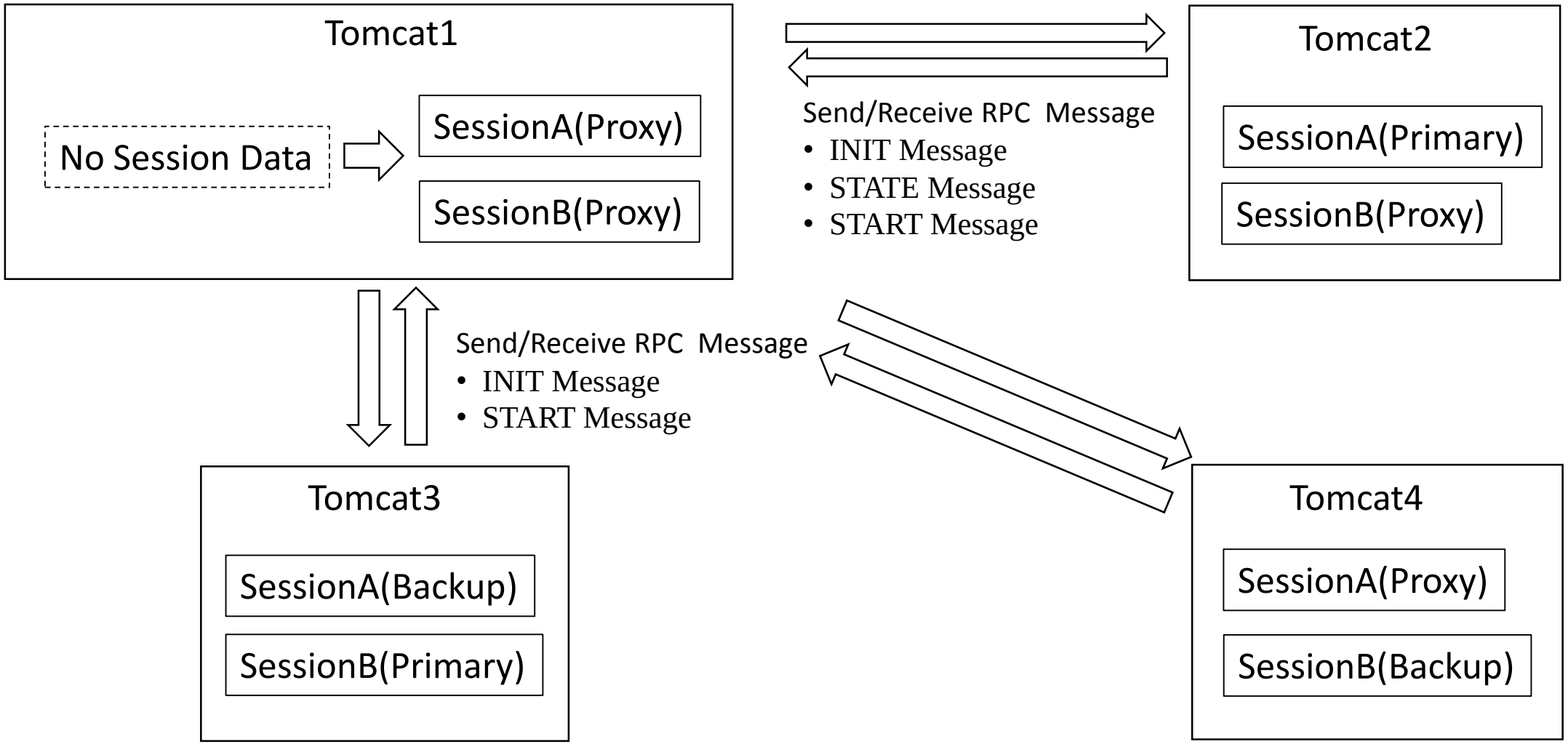
BackupManager

Access to Proxy node



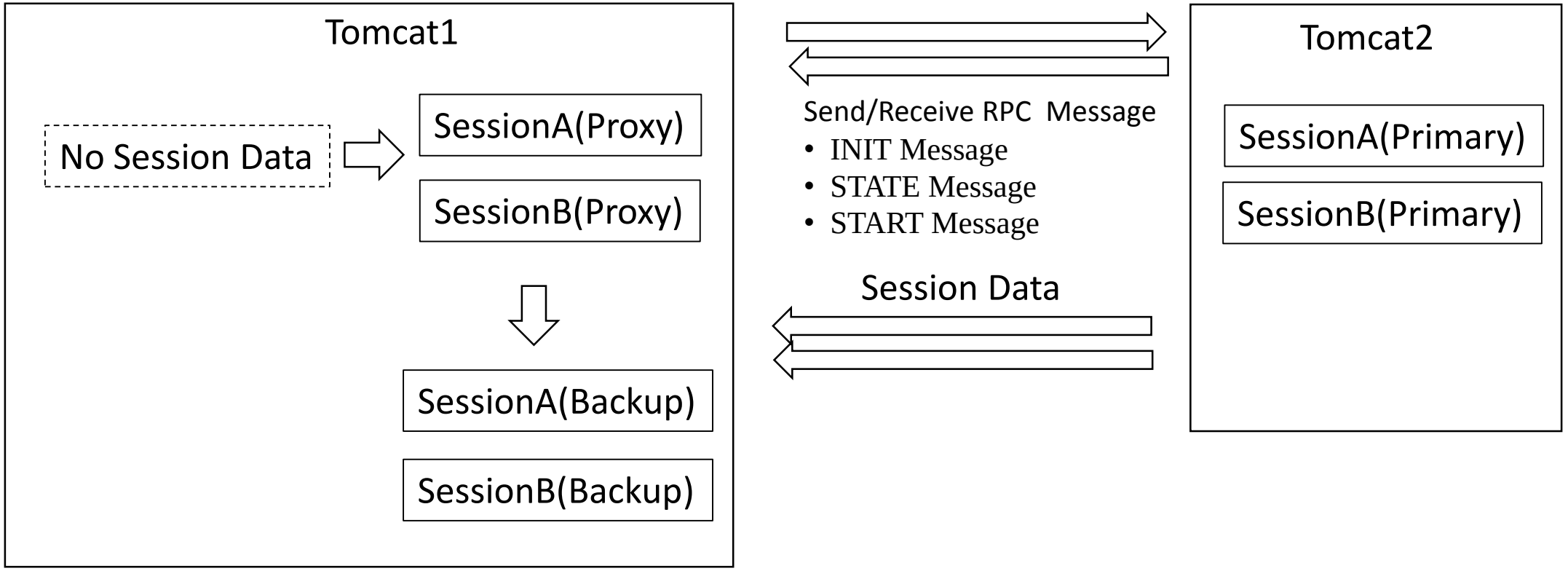
BackupManager

Node Recovery



BackupManager

Node Recovery

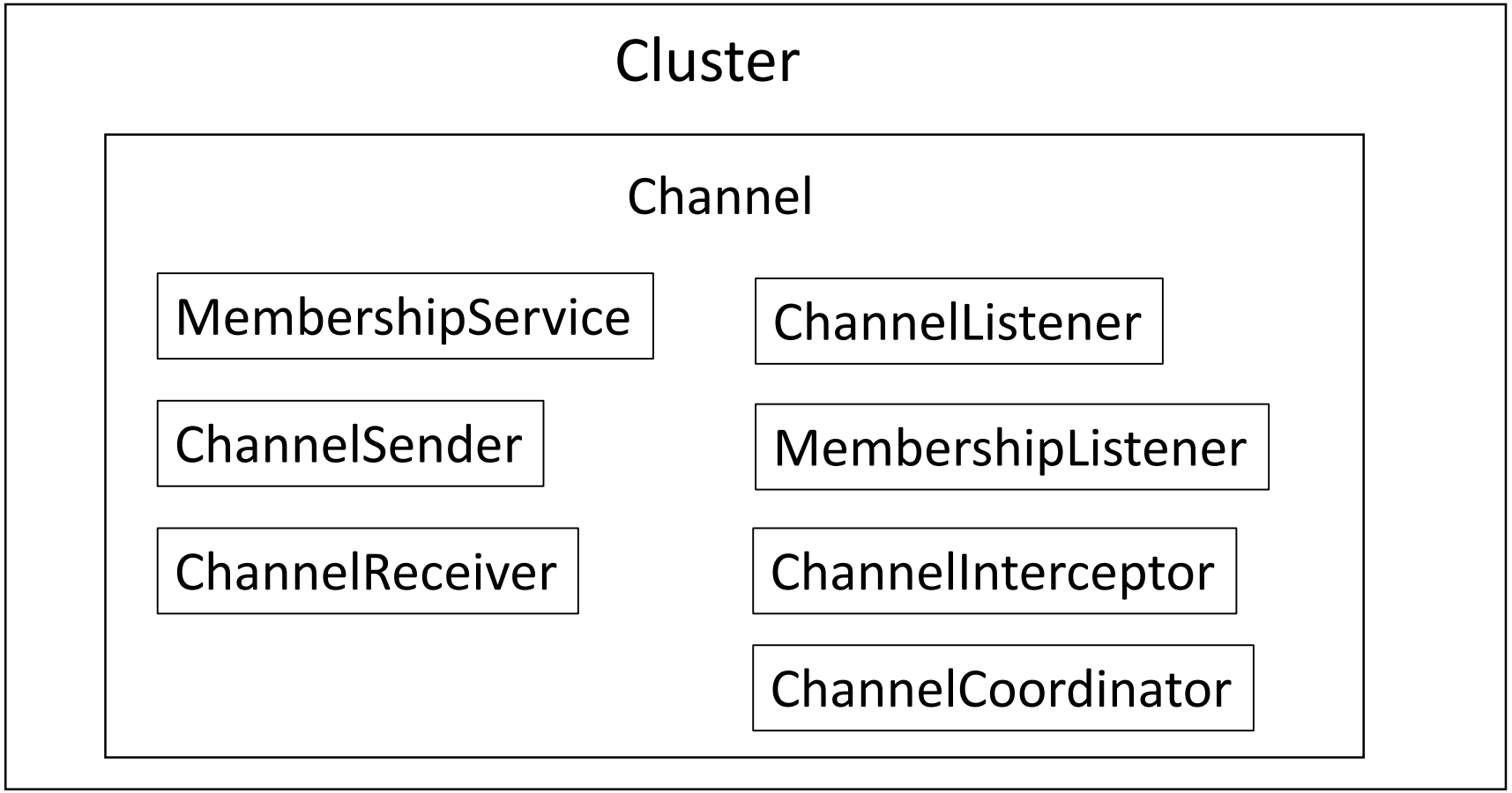


Channel

Channel

- What is Channel ?
  - Messaging & Grouping component
- Responsibility
  - Membership
  - Send channel message
  - Receive channel message

Channel Components



Channel Components

- MembershipService
  - The component which build a cluster group
  - Start a multicast receivier thread and a multicast sender thread

Channel Components

- ChannelSender
  - Send the channel message to another nodes in the cluster group
  - Sender Queue size is specified in poolSize attribute
- ChannelReceiver
  - Receiving a channel message from other nodes in the cluster group
  - Tuning of the maxThreads attribute depends on send option of channel message

Channel Components

- ChannelListener
  - Listen received channel messages
- MembershipListener
  - Listen add/remove of cluster members

Channel Components

- ChannelInterceptor
  - Intercept a channel message and a member detection
- ChannelCoordinator
  - Coordinates the ChannelInterceptors

Sample config

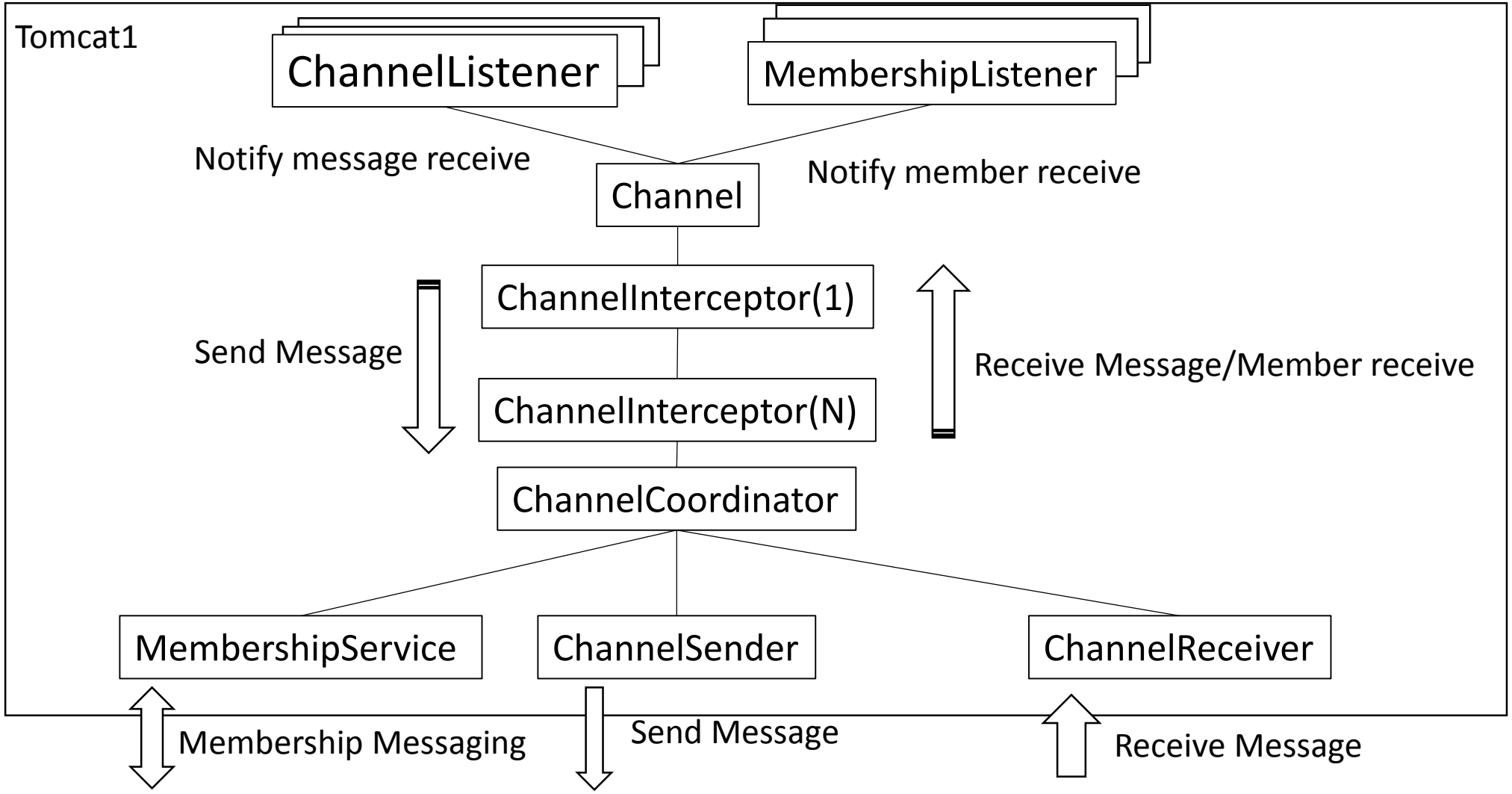

```
<Cluster className="org.apache.catalina.ha.tcp.SimpleTcpCluster">
 ... ...
  <Channel className="org.apache.catalina.tribes.group.GroupChannel">
    <Membership className="org.apache.catalina.tribes.membership.McastService"
       address="229.0.0.61“ port="45564“ frequency="500“ dropTime="4000" />

    <Receiver className="org.apache.catalina.tribes.transport.nio.NioReceiver"
       address=“auto“ port=”4004“ autoBind=”100” maxThreads=“25”/>

    <Sender className="org.apache.catalina.tribes.transport.ReplicationTransmitter">
      <Transport className="org.apache.catalina.tribes.transport.nio.PooledParallelSender"
         timeout="5000“  poolSize="25"/>
    </Sender>

    <Interceptor className="org.apache.catalina.tribes.group.interceptors.MessageDispatch15Interceptor" />
    <Interceptor className="org.apache.catalina.tribes.group.interceptors.TcpFailureDetector"/>
  </Channel>
 ... ...
</Cluster>
```

Channel Architecture



ChannelInterceptor

TCPFailureDetector

- Main Features
  - Intercept memberDisappeared events.
  - Member check at the time of send errors.
- Other Feature
  - Manage membership

ChannelInterceptor

MessageDispatch15Interceptor

- Asynchronous send message
- Use Thread Pool
  - maxThreads
  - maxSpareThreads
- Make sure
  - Cluster’s channelSendOptions attribute = 8 or 10 (8+2)
  - BackupManager’s mapSendOptions attribute = 8 or 10 (8+2)

ChannelInterceptor

Notes on use with TCPFailureDetector

- The Order is important

Session Message

- Intercept memberDisappeared
- Member check at send errors

Session Message

- Intercept memberDisappeared
- ~~Member check at send errors~~

ChannelInterceptor

StaticMembershipInterceptor

- The static membership instead of multicast

- Make sure
  - Disable multicast membership
    - Cluster’s channelStartOptions attribute = 3
  - Enable TcpPingInterceptor for nodes failure detection
  - Enable TcpFailureDetector for membership management
  - The order is

    TcpPingInterceptor->TcpFailureDetector->StaticMembershipInterceptor

ChannelInterceptor

Sample config

```
<Interceptor className="org.apache.catalina.tribes.group.interceptors.TcpPingInterceptor"/>
<Interceptor className="org.apache.catalina.tribes.group.interceptors.TcpFailureDetector"/>
<Interceptor className=
      "org.apache.catalina.tribes.group.interceptors.StaticMembershipInterceptor">
 <Member className="org.apache.catalina.tribes.membership.StaticMember"
    port="4010" host="hostA"
    uniqueId="{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,2}" />
 <Member className="org.apache.catalina.tribes.membership.StaticMember"
    port="4010" host="hostB"
    uniqueId="{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,3}" />
</Interceptor>
```

ChannelInterceptor

ThroughputInterceptor

- Measuring the send and receive of channel messages

DomainFilterInterceptor

- Filter the members that join cluster group by the domain

Other Cluster features

JvmRouteBinderValve

- Change jvmRoute of session ID to JvmRoute of the current node
- Enable sticky session again

ClusterDeployer

- Replicate the WAR among clusters

ClusterSingleSignOn

- Replicate SSO information in cluster

Questions?

Thank You