Zowe High Availability

2020PI3

Hills

- As a z/OS system programmer I can configure Zowe z/OS Components to deliver its services through a Highly Available endpoint, using familiar practices and technology.

- As a z/OS system programmer I an able to use familiar console automation offering to run, configure, and monitor Zowe z/OS components.

Current Understanding

- There's currently more questions than answers re HA for Zowe
- There is currently no clear statement around HA for Zowe across all components
- Different interpretations of what HA means, topics are too broad
- Some components talked about HA but unclear in practice and what does it mean to the end users
- Solution from one squad may not be consistent with other squads'
- Unclear what HA means to Zowe extenders for the different types of plugins

Goals for 2020PI3

- Collect opinions and get agreement on what HA means for Zowe
  - What should be included, either in short term or long run
  - What should not be included and should be resolved as separated issue
- Discuss all aspects of HA we defined and propose solutions, setup POC tests when necessary
- Create Zowe HA draft of configuration guidance
- Identify implementation work items and prepare for PI4

What will be discussed through PI3

- Definition
- Stateful Components
- Load balancing
- Failover
  - Health check and Monitoring
  - Lifecycle / Recovery failed address space
  - Backup
- Extra
  - Rolling update
  - HA on Zowe pre-reqs including z/OSMF and node.js
  - CLI + z/OSMF under SYSPLEX
  - Test method

- Output:
  - Solutions for each topic

HA - Definition

- What should be included?
  - Load balancing
  - Failover
  - Backup and recovery admin suggestions
  - High Scalability
- What configurations we will support?
  - Single LPAR
  - SYSPLEX
  - Containerized

HA - Stateful Components

- This will also affect containerization and high scalability
- If we should convert all stateful components to stateless
- Which Zowe components are stateful
- What states we are storing in instance level
- Should we introduce central db/cache to store shared states
- How do we validate

HA - Load balancing

- Choices
  - APIML Ribbon
  - SYSPLEX Distributor
  - Third-party load balancer, hardware
- Will the load balancer support session based rules
- Will the load balancer support long HTTP connections and websocket
- Other protocols other than HTTP/HTTPS
  - TCP generic
  - UDP

Investigation Matrix

	APIML Ribbon	SYSPLEX Distributor
Gateway		
Discovery Service		
Other Components		

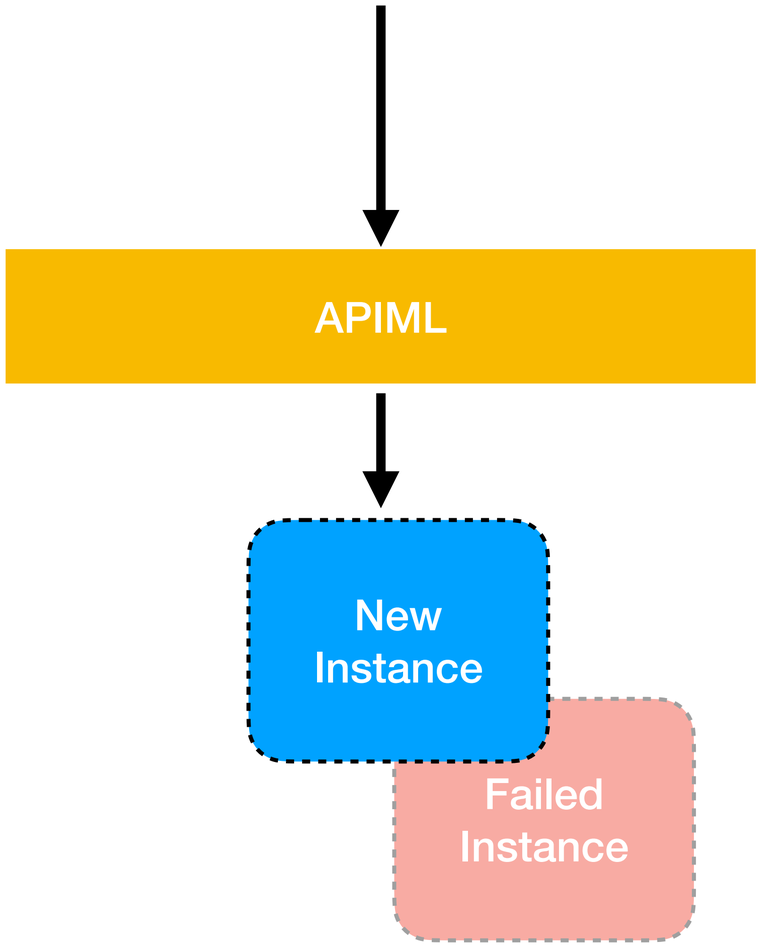
HA - Failover

- How to identify failed instance / component
- Who should be responsible for health check and monitoring the failures
- Can Eureka handle active health check
  - Dynamic vs static service registration
- How to tell load balancer not to direct traffic to failed instance and switch to healthy instances
- How to restart from failure
  - What about port, should be dynamic assigned
- What if it fails again
- Can we always start 2+ instances for a component
- We should output reasonable test to log for admin to track
- Should we suggest System Automation and what’s suggested policy
- Should we suggest physical backup / recovery suggestions

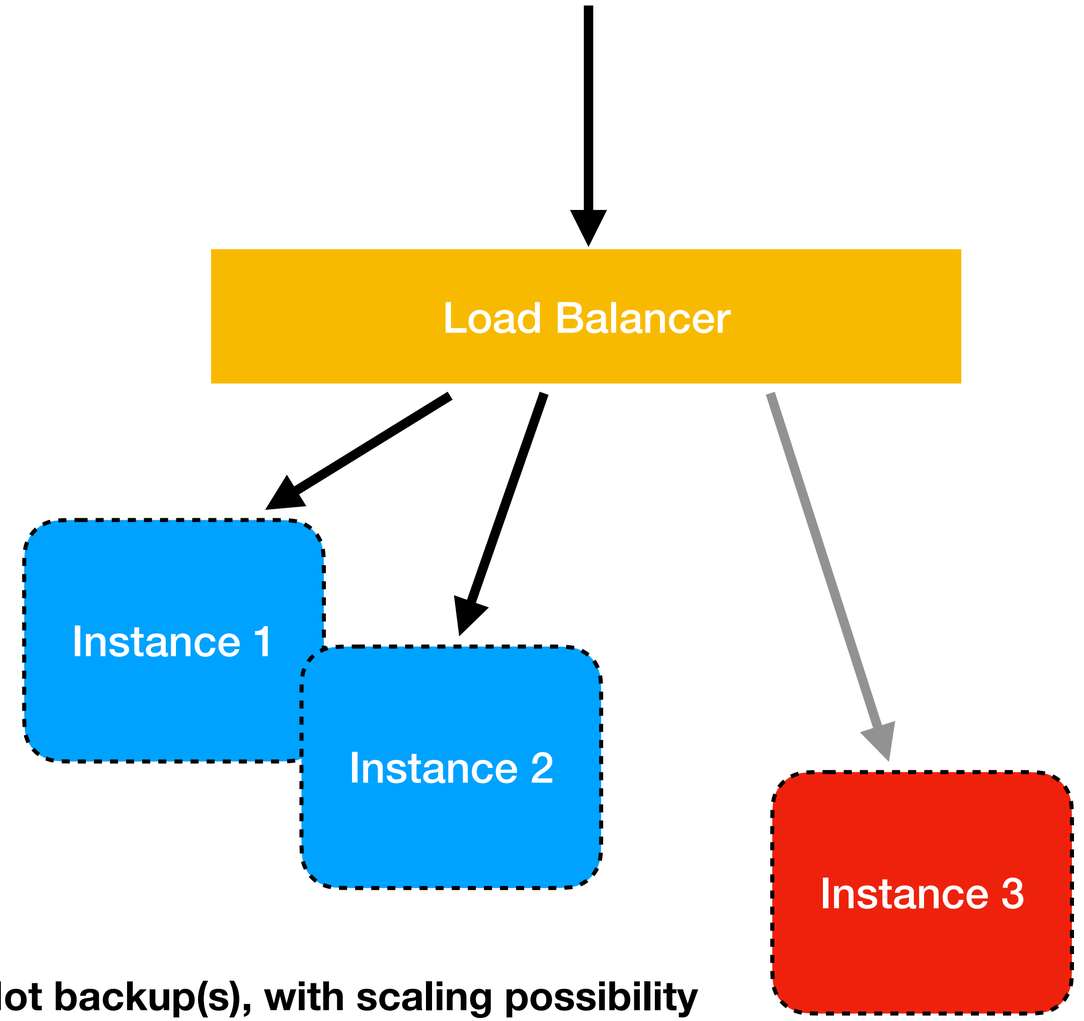
HA - Other topics

- Should we support rolling update to achieve 0-downtime
- We heavily depends on z/OSMF, where should we point the customer to find HA solution for z/OSMF
- Will CLI be able to handle z/OSMF running on SYSPLEX
- How do we test our HA solutions
- What we can tell extenders related to HA

Use Cases



Single instance, restart once it fails



Hot backup(s), with scaling possibility

Goals for 2020PI3S1

- Epic: https://github.com/zowe/zowe-install-packaging/issues/1467
- Sprint 1 item: Research user expectations https://github.com/zowe/zowe-install-packaging/issues/1468
- Output:
  - A clearly defined scope for what high availability means for Zowe. This scope includes the qualities of High Availability to target (i.e., reliability, scalability, disaster recovery) as well as some deployment attributes (sysplex, multi-sysplex, container+sysplex, etc.).
- Present results to the team on a Zowe Architecture call