System	Cloud Service Availability				FMEA Number	1	
Subsystem	Software Stack		Potential Failure Modes and Effects		Prepared By	Adrian Cockcroft	
Component					FMEA Date	12/5/2018	
Design Lead		Key Date	N/A		Revision Date	11/10/2019	
Core Team					Page	1 of 1	

Item / Function	Potential Failure Mode(s)	Potential Effect(s) of Failure	Sev	Potential Cause(s)/ Mechanism(s) of Failure	Prob	Current Design Controls	Det	RPN	Recommended Action(s)	Responsibility & Target Completion Date	Action Results: Actions Taken	ew Sev	ew Occ	ew Det	New RPN
Authentication to cloud services	Client can't authenticate	Can't connect application	5	Certificate timeout, version mismatch, account not setup, credential changed	3	Log and alert on authentication failures	3	45							0
	Slow or unreliable authentication	Slow start for application	4	Auth service overloaded, high error and retry rate	3	Log and alert on high authentication latency and errors	4	48							
								0							0
Client request to cloud service endpoint	Service unknown, address un-resolvable	Delay while discovery or DNS times out, slow fallback response	5	DNS configuration error, denial of service attack, or provider failure	1			0							0
	Service unreachable, request undeliverable	Fast fail, no response	4	Network route down or no service instances running	1			0							0
	Service reachable, request undeliverable	Connect timeout, slow fail, no response	4	Service frozen/not accepting connection	1			0							0
	Request delivered, no response - stall	Application request timeout, slow fail, no response	4	Broken service code, overloaded CPU or slow dependencies	1			0							0
	Response undeliverable	Application request timeout, slow fail, no response	4	Network return route failure, dropped packets	1			0							0
	Response received in time but empty or unintelligible	Fast fail, no response	3	Version mismatch or exception in service code	2			0							0
	Request delivered, response delayed beyond spec	Degraded response arrives too late, slow fallback response	6	Service overloaded or GC hit, dependent services responding slowly	2			0							0
	Request delivered, degraded response delivered in time	Degraded timely response	2	Service overloaded or GC hit, dependent services responding slowly	2			0							0
EC2 Control Plane	Instance request refused, direct or via autoscaler	Capacity limited or control plane failure		Limit reached, or Insufficient Capacity Exception				0	Service call for increased limit. Try a different instance type, different zone, or different region						
	Instance created but fails to start	Bad instance hardware						0	Retry via autoscaler						0
	Instance slow to start							0							
EC2 Network Control Plane	Configuration request refused	Capacity limited or control plane failure		Limit reached, or Insufficient Capacity Exception				0	Service call for increased limit. Try a different zone, or different region						
	Network creation started but operation fails							0	Pre-allocate all network structures in all regions						0
Database Service (Aurora Postgres)	Control plane configuration request refused							0	Service call for increased limit. Try a different zone, or different region						0
	Database table creation started but operation fails							0	Pre-allocate all database tables in all regions						
	Master Instance failure														

Master instance reboot														
Cross region read replica failure														
Excessive cross region replication latency														
Cross region replication source node failure														
Cross region replication target node failure														
Read replica instance failure														
Excessive in-region replication latency														
Excessive commit latency														
Excessive select latency														
Excessive provisioning latency for secondary instatnce														
Failure to detatch secondary instance from the cluster														
Failure to scale storage before it fills up														
Overload: Transaction rate higher than cluster capacity.														
CPU utilization spikes on database instances causing timeouts														
High memory utilization on database instance														
Maximum number of database connections limit reached														
Errors during JDBC network request operations														
Latency during JDBC operations														
Failure in correlated 2-phase commit														
Master and replica in region failure														
Regional failure recovery impacted by flood of backed up requests														
Failure to distinguish between instance failure, zone failure, and regional failure														

							0								0
Message Queue	High queue latency	Messages backed up in queue		High network latency											
				Network packet loss											
				Network packet corruption											
				Instance storage overloaded											
				Root volume full											
				Processes dying											