

#	Experiment	Configuration
1	Table Initializer: Optimized Policy: Greedy	Executor Granularity: 1 Executor Strategy: Static Target Latency: 30 Seconds Decision Interval: 1 Minute History Window: 2 Minutes Latency Granularity: 10 Seconds Learning Factor: 0.7 Discount Factor: 0.9 Reward: Prefer Scale-In When Load Is Decreasing
2	Table Initializer: Zero Policy: One Minus Epsilon Epsilon: 0.2	
3	Table Initializer: Random Policy: Greedy	