

On Server 1, Run this Script every <time>:

Check memory (or CPU)

- if memory > 90% call Webhook1
- if memory < 60% call Webhook2

* Script needs to run reliably at all times

Policy1 = Scale UP by X Webhook1 invokes Policy1

Policy2 = Scale DOWN by Y Webhook2 invokes Policy2

*Note:

- X and Y values need to be tuned based on the ramp up characteristics and time taken to spin up an image
- Min = 0 means that Autoscaling Group with all policies and webhooks can exist without any server running