Pseudo Code

Docker service scale to 2 replicas.

New ending while loop:

Record start time.
Send request to 10.2.1.38:8000
Record end time

Compute time (end time - start time) and store in array

if(length of array equal or greater than 10):

Compute the average time over last 10 recorded times.

if(average time is greater or equal to 5):

Docker service increment to +1

Else if (averate time is lower than 2.5 and replicas greater than 2):

Docker service decrement by -1

Else:

Do nothing

Empty the compute time array

Wait for 1 second and repeat the while loop.