- Our internal client produce currently ~3 billion requests per day
- Sine curve load patterns with scattered random load spikes
- Different kind of services with different load patterns and runtime behavior involved, e.g. workers
Auto Scaling for async processing

- Workers process jobs from a queue and scale based on the queue size
- Services can have a high and low priority queue
- If auto scaling does not work right we get called at night

![Diagram of Auto Scaling for async processing](image)

Oscillations

Tasks → Message Queue → Async Processing → Service

<table>
<thead>
<tr>
<th>- Queue size</th>
<th>- Oscillations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>5</td>
</tr>
<tr>
<td>High</td>
<td>20</td>
</tr>
</tbody>
</table>

Drawbacks of traditional scaling algorithms

- Manual parameter “fine-tuning” for each type of service
  - Simulation
  - Load test
  - Production
- Another manual fine-tuning necessary if the load pattern on production changes
Reinforcement learning can play Space Invaders, maybe also Auto Scaling?

Scaling is a control problem

- Deep Q Learning showed great results in playing Atari Games, see “Human-level control through deep reinforcement learning” [1]
- If this algorithm can learn to play different Atari games, why not scaling?

Analogies between Scaling & Playing Space Invaders

- **Optimise**
  - shoot aliens
  - don't die
- **Delayed feedback** (after next move)
  - Time between firing and hitting the aliens
- **Control Problem**

- **Optimise**
  - good customer experience
  - minimum compute costs
- **Delayed feedback** (after scaling)
  - Time to boot up compute instances before work can be done (and queue size decreases)
- **Control Problem**
  - Historical data is insufficient
Deep Reinforcement Learning Space Invaders

- Rules of the environment are unknown
- Agent’s actions affect the subsequent data it receives
First iteration, scaling using DeepQL

The Simulation
- Influx as simple sine curve
- Overflow queue
- Instance count
- Possible Agent Actions: +1, 0, -1
- Actions have delayed effect

Observation
- Load
- Queue size
- Capacity

Rewards
- Positive reward for load between 40-60%
- -1 * (queue size)
- Positive reward for not scaling -> action 0
- Negative reward for the count of compute instances
Two years of disappointment
DeepQ after improvements

fps = 0.0
frame = 1
avg reward = -1.00000
instance cost = 0 $
load = 100
instances = 2
Influx = 2197
influx_derivative = 0.00
actions q = a: 1
avg queue size = 0.000
avg instances = 2.000
avg load = 0.000
Make it easy

- Training on fixed input signals over 10k frames
- Smaller NN, 20x20

Reward signal

**Compute Cost**
Weighted number of instances by the current load

\[-(1 - \frac{\text{load}}{100}) \times (\frac{\#\text{instances}}{\text{max\_instances}})\]

**Queue size**
Negative normalized number of messages in the queue using inverse odds ratio function

\[\frac{p}{1 + p}\]
Production data

fps = 0.0
frame = 101
avg reward = -0.70284
instance cost = 192 $
load = 20
instances = 58
influx= 1000
influx_deriviate= 0.00
actions q = a: -1
avg queue size = 15047.198
avg instances = 40.000
avg load = 92.614
Still a lot of work to do
• Scale more than one instance per step
• Must converge more stable
• Maybe PPO works better?
OpenAI Scaling Gym

- Original experiment was ported to OpenAI Gym
- The Scaling Gym is Open Source! Contributions welcome
  - [https://github.com/adobe/gym-scaling/](https://github.com/adobe/gym-scaling/)
Rate today’s session

Cyberconflict: A new era of war, sabotage, and fear

9:55am-10:10am Wednesday, March 27, 2019
Location: Ballroom
Secondary Topics: Security and Privacy

Rate This Session

We’re living in a new era of constant sabotage, misinformation, and fear, in which everyone is a target, and you’re often the collateral damage in a growing conflict among states. From crippling infrastructure to sowing discord and doubt, cyber is now the weapon of choice for democracies, dictators, and terrorists.

David Sanger explains how the rise of cyberweapons has transformed geopolitics like nothing since the invention of the atomic bomb. Moving from the White House Situation Room to the dens of Chinese, Russian, North Korean, and Iranian hackers to the boardrooms of Silicon Valley, David reveals a world coming face-to-face with the perils of technological revolution—a conflict that the United States helped start when it began using cyberweapons against Iranian nuclear plants and North Korean missile launches. But now we find ourselves in a conflict we’re uncertain how to control, as our adversaries exploit vulnerabilities in our hyperconnected nation and we struggle to figure out how to deter these complex, short-of-war attacks.

David Sanger
The New York Times

David E. Sanger is the national security correspondent for The New York Times as well as a national security and political contributor for CNN and a frequent guest on CBS This Morning, Face the Nation, and many PBS shows.

Session page on conference website

O’Reilly Events App
And yes, we are hiring

Stefanie Grunwald, Sr Data and platform engineer
Michael Friedrich, Sr Computer Scientist

MAKE IT AN EXPERIENCE
And yes, we are hiring