The Practice of Chaos Engineering

Ana Medina
Chaos Engineer at Gremlin
@ana_m_medina
Ana Medina

@ana_m_medina

Chaos Engineer @ Gremlin

Previously Software Engineer / SRE @ Uber, Also worked/interned @ SFEFCU, Google, Quicken Loans, Stanford University and Miami Dade College.

College dropout. Self taught engineer.
How many of you have heard of Chaos Engineering?
How many of you have run a Chaos Engineering experiment?
What is Chaos Engineering?
Chaos Engineering

Thoughtful, planned experiments designed to reveal the weakness in our systems.
Chaos Engineering

Inject something harmful to build an immunity.

-@KoltonAndrus
Gremlin Founder and CEO
Why?

- Microservices
- Systems are scaling fast
- Downtime is really expensive
- Our dependencies will fail
- Pager fatigue and burnout really hurts
Use Cases:

- Outage reproduction
- On-call training
- Strengthen new products
- Battle test new infrastructure and services

Amazon AWS S3 outage is breaking things for a lot of websites and apps

Darrell Etherington  etherington  2 years ago
What do you need before doing Chaos Engineering?

- Monitoring/Observability
- On-Call and Incident Management
- Cost of Downtime Per Hour
Chaos Engineering is not

- Unexpected or unmonitored experiments
- Creating outages
“Chaos Engineering Without Observability ... Is Just Chaos”

-@mipsytipsy
Charity Majors
CEO of honeycomb
Minimize the Blast radius

1. Form a hypothesis.
2. Run an experiment.
3. Abort Conditions
   - Failure
   - Success
4. Find and fix issues.
5. Scale up and repeat.
THE BEGINNING

Chaos Monkey

Level 0

VALUE PROVIDED
Prepare for host failures in the cloud

APPROACH TAKEN
Random

MATURITY REQUIRED
Low
THE FIRST STEP

Infrastructure Failures

Level 1

VALUE PROVIDED
Prepare for host-level failures

APPROACH TAKEN
Disciplined

MATURITY REQUIRED
Basic Operations
Network Failures

Level 1.5

**VALUE PROVIDED**
Prepare for high impact events

**APPROACH TAKEN**
Gameday

**MATURITY REQUIRED**
Networking expertise
Application Failures

Level 2

VALUE PROVIDED
Safely validate the user experience

APPROACH TAKEN
Precision Experiments

MATURITY REQUIRED
Advanced
Latency added to 50% of Android traffic
Exceptions - 50% of android traffic failed
You can and should inject chaos at every layer of your stack

- Application
- API
- Caching
- Database
- Hardware
- Cloud Infrastructure / Bare metal
Top places to inject chaos

cassandra

Amazon ECS

elasticsearch

kafka
# Our Awesome Test Application

## Real World Scenario
What happens when the message queue becomes 100ms slower due to load?

## Hypothesis
The application layer will see slower commits, memory utilization will increase due to a larger queue depth, and eventually the queue will overrun and the application will lose data.

## Monitoring Tools
**Humans / native linux tools / Data Dog**

## Experiment
- **Tool:** Network Gremlin
- **Attack:** 100ms Latency
- **Scope:** Single Producer Node <-> Message Queue, TCP Port 6667
- **Duration:** 10 minutes

## Abort Conditions
- Data Loss
- 500 errors for users or consuming services
- Compound latency beyond 1000ms in any consuming service.

---

**THE RESULTS**

You, the reader, should run the test and record your results, as they will be unique to your application and environment.
WHAT I LEARNED RUNNING THE CHAOS LAB:

Kafka Breaks

Getting Started:

- Identify top 5 critical systems
- Choose system
- Whiteboard the system
- Determine what experiment you want to run: (resource, state, network)
- Determine Blast Radius
Companies doing Chaos Engineering

- Netflix
- Twilio
- Under Armour
- Dropbox
- Expedia
- Uber
- Jet
- Amazon
- Gremlin
Chaos Days
Chaos Days: Dedicated day for your entire company to focus on building resilience instead of new products.

https://www.gremlin.com/community/tutorials/planning-your-own-chaos-day/
“What could go wrong?”

“Do we know what will happen if this breaks?”
Chaos Day Crew:

VP Engineering / CTO / COO
Executive Assistant
Engineering Director / Manager
Senior Engineer
New Grad / Intern Engineer
What experiments can you run?

- Reproduce outage conditions
- Unpredictable circumstances
- Large traffic spikes
- Race conditions
- Datacenter failure
- Time travel - system clocks to be out of sync
- Network errors
- CPU overloads
What tools can you use?

Chaos Monkey

Chaos Monkey is a resiliency tool that helps applications tolerate random instance failures.

github.com/Netflix/chaos-monkey
What tools can you use?

Simian Army

Tools for keeping your cloud operating in top form. Chaos Monkey is a resiliency tool that helps applications tolerate random instance failures.

github.com/Netflix/Simian Army
What tools can you use?

Litmus

Litmus is chaos engineering for stateful workloads on Kubernetes, hopefully without learning curves

github.com/openebs/litmus
What tools can you use?

Powerful Seal

A powerful testing tool for Kubernetes clusters.

github.com/bloomberg/powerfulseal
What tools can you use?

Gremlin

Failure as a Service.


Schedule experiments using UI, API and CLI. Provides 11 gremlins out of the box.
Learn more:
Join the Chaos, Join Slack:

@ana_m_medina
#reactive18

Join the Chaos, Join Slack:

# slack

bit.ly/chaos-eng-slack
1,900+ members across the world
THANKS!

ana@gremlin.com

@ana_m_medina