Migrating Internet Archive to Kubernetes

KubeCon 2018 - Seattle, WA

@tracey_pooh
@dvanduzer

2018 December 13
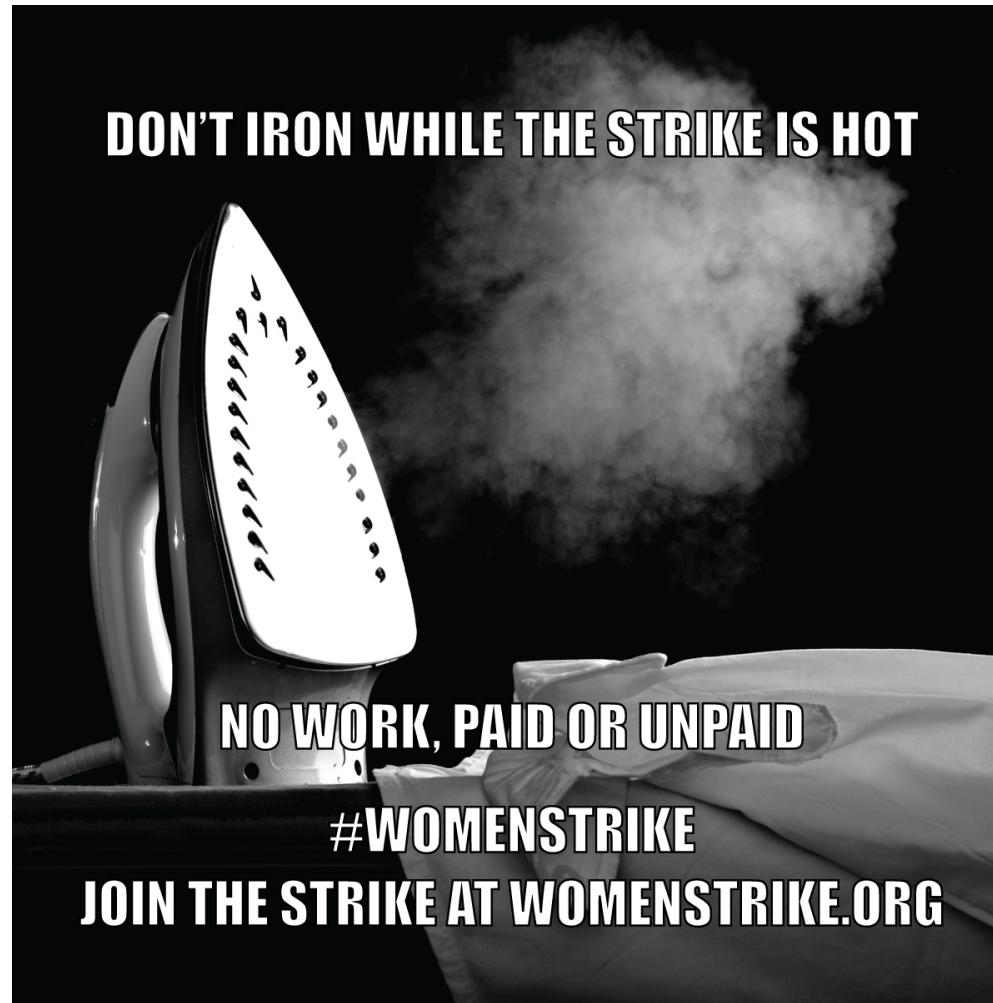
? for keyboard shortcuts

https://archive.org/~tracey/slides

Challenge of Big Changes

DON'T IRON WHILE THE STRIKE IS HOT

NO WORK, PAID OR UNPAID

#WOMENSTRIKE
JOIN THE STRIKE AT WOMENSTRIKE.ORG

What's in it for Us?

- language / platform / OS flexibility of Docker
- model to enforce rebuildable recipes
- more coding; less system building and admin
- avoid cpu excess/starvation from static resource allocation
- fewer "pet" servers

Internet Archive

a Digital Library

Curate

Preserve

Circulate
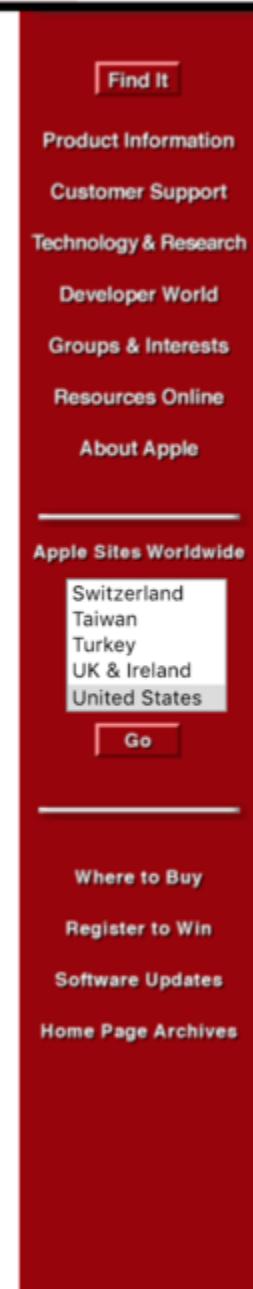
Curate

Curate all the things!

- 300B+ web pages WayBack Machine
- 15M books
- 4M videos & TV News programs
- 4M audio & concerts
- 3M images
- 200K software items & emulation

JULY 14

🌈 Welcome to Apple 1997

Find It

Product Information

Customer Support

Technology & Research

Developer World

Groups & Interests

Resources Online

About Apple

Apple Sites Worldwide

Switzerland
Taiwan
Turkey
UK & Ireland
United States

Go

Where to Buy

Register to Win

Software Updates

Home Page Archives

▽ What's Hot

Preorder Mac OS 8
Now you can preorder Mac OS 8, described by Macworld as "the most comprehensive update to the Mac OS in years, sporting a bold new look, a speedier Finder, more shortcuts and integrated Internet functions."

Want a PowerBook?
Qualify to win a PowerBook 3400/200 by entering this month's Apple Registration Sweepstakes.

Big Help for Small Biz Find out what you need to make your small business a success. Apple Small Business has information about everything from great deals on leasing to

Be the First to Know
Learn about new Macintosh software releases the moment they become available. Check Hot Mac Products to hear about programs like Speed Demon, ReBirth RB-338 and QuickCRC.
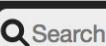
Newton Connects
Newton, Inc., will enhance network connectivity for Newton-based devices this fall via Newton Internet Enabler 2.0. Ethernet capability can connect devices to Local Area Networks.

**Welcome to Mars**
See alien terrain in QuickTime VR movies from

GOODY TWO SHOES.

FARMER MEANWELL was at one time a very rich man. He owned large fields, and had fine flocks of sheep, and plenty of money. But all at once his good fortune seemed to desert him. Year after year his crops failed, his sheep died off, and he was obliged to borrow money to pay his rent and the wages of those who worked on the farm.

At last he had to sell his farm, but even this did not bring him in money enough to pay his debts, and he was worse off than ever.

Among those who had lent money to Farmer Meanwell were Sir Thomas Gripe, and a Farmer named Graspall.

Sir Thomas was a very rich man indeed, and Farmer Graspall had more money than he could possibly use. But they were both very greedy and covetous, and particularly hard on those who owed them anything. Farmer Graspall abused Farmer Meanwell and called him all sorts of dreadful names; but the rich Sir Thomas Gripe was more cruel still, and wanted the poor debtor shut up in jail.

So poor Farmer Meanwell had to hasten from the place where he had lived for so many years, in order to get out of the way of these greedy men.

He went to the next village, taking his wife and his two little children with him. But though he was free from Gripe and Graspall she was not free from trouble and care.

He soon fell ill, and when he found himself unable to get food and clothes for his family, he grew worse and worse and soon died.

His wife could not bear the loss of her husband, whom she loved so dearly, and in a few days she was dead.

The two orphan children seemed to be left entirely alone in the world, with no one to look after them, or care for them, but their Heavenly Father.

They trotted around hand in hand, and the poorer they became the more they clung to each other. Poor, ragged, and hungry enough they were!

Tommy had two shoes, but Margery went barefoot. They had nothing to eat but the berries that grew in the woods, and the scraps they could get from the poor people in the village, and at night they slept in barns or under hay-stacks.

Their rich relations were too proud to notice them. But Mr. Smith, the clergyman of the village where the children were born, was not that sort of a man. A rich relation came to visit him—a kind-hearted gentleman—and the clergyman told him all about Tommy and Margery. The kind gentleman pitied them, and ordered Margery a pair of shoes and gave Mr. Smith money to buy her some clothes, which she needed sadly. As for Tommy he said he would take him off to sea with him and make him a sailor. After a few days, the gentleman said he must go to London and would take Tommy with him, and sad was the parting between the two children.

Poor Margery was very lonely indeed, without her brother, and might have cried herself sick but for the new shoes that were brought home to her.

They turned her thoughts from her grief; and as soon as

Goody Two-Shoes
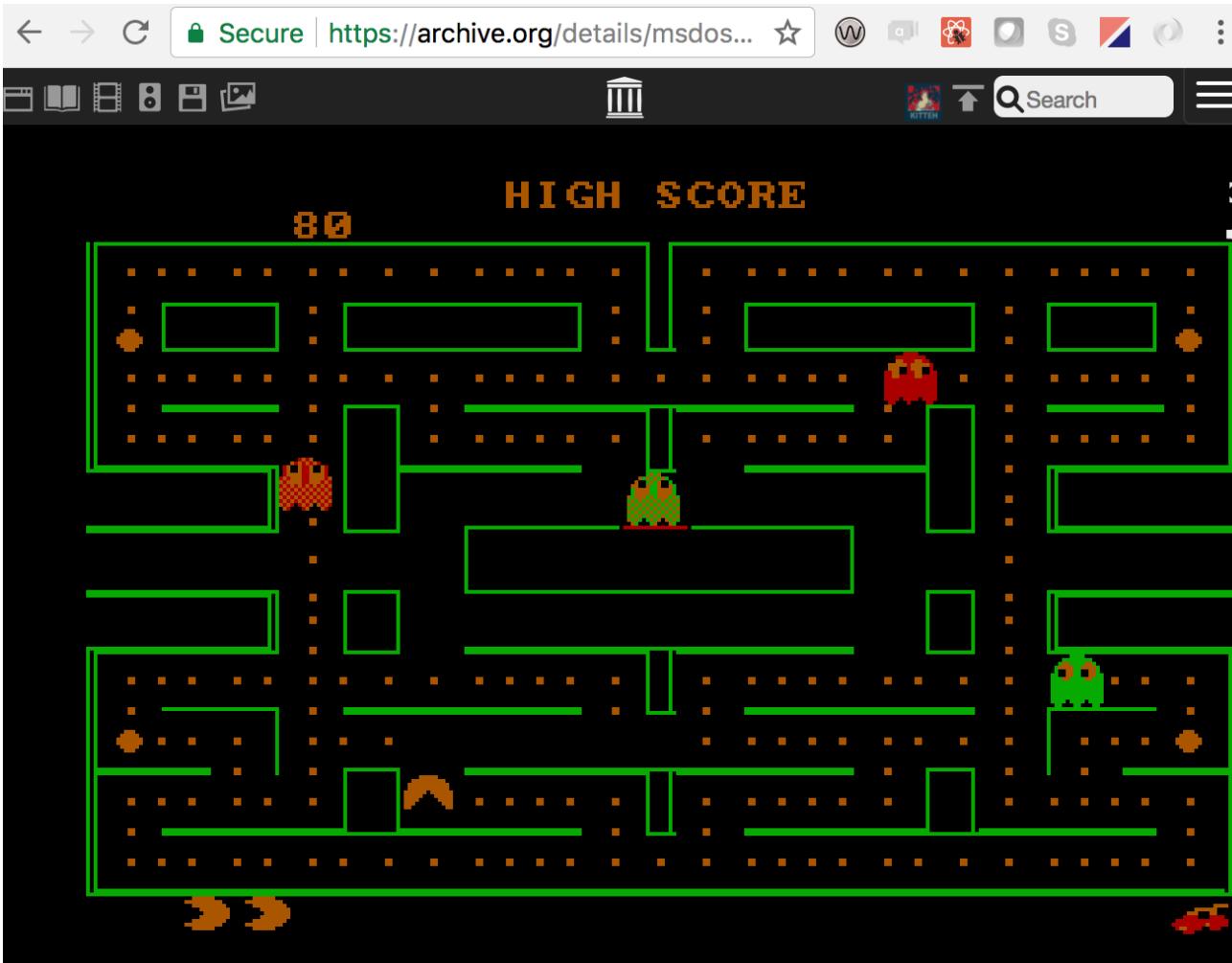
Edit Manage History Book view Barcode

Publication date c1888

Topics Brothers and sisters, Orphans, Conduct of life, Education

Publisher New York : McLoughlin Bro's

HIGH SCORE

80
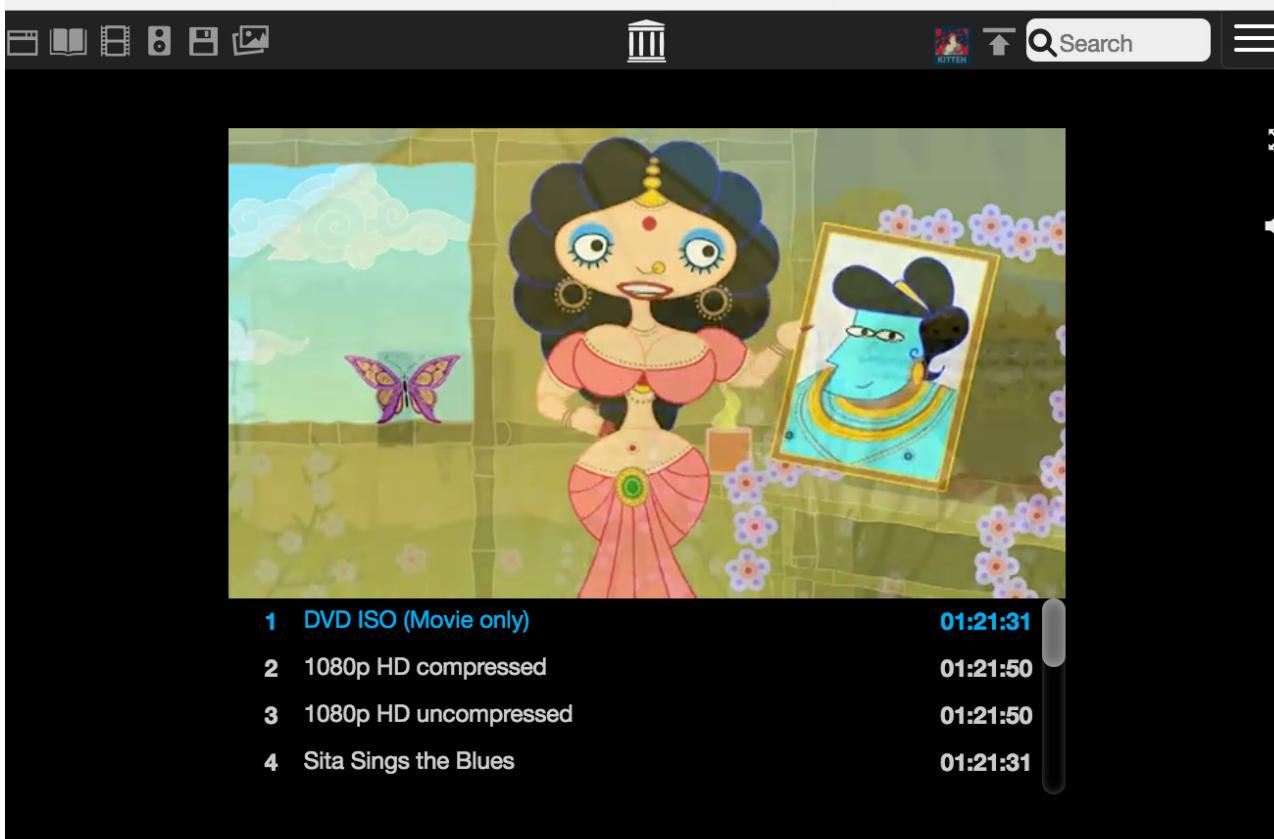
Pac-Man

✎ Edit   ⚙ Manage   🕐 History   ☰ Barcode

by Namco Limited

Publication date 1983

1	DVD ISO (Movie only)	01:21:31
2	1080p HD compressed	01:21:50
3	1080p HD uncompressed	01:21:50
4	Sita Sings the Blues	01:21:31

Sita Sings the Blues

✎ Edit   ⚙ Manage   ⏱ History   ≣ Barcode

by Nina Paley

Publication date 2008
Usage CC0 1.0 Universal (cc)(0)

Source Flash Authoring Files can be found here: http://www.archive.org/details/Sita_Sings_the_Blues

Low Friction Migrations to Kubernetes

- book scanning apps
- crawl frontier management
- finite duration projects
- 'green field' / standalone apps

Preserve

Preserve

- 50+ petabytes over 22 years
- Fourth generation storage system
- Independence by self-hosting
- Multi-datacenter replication

Simplicicty == Longevity

- simple directories of files
- original files with metadata
- multiple checksums
- changes logged forever
- replication: 'just' rsync

```xml
<file name="commute.mp4" source="derivative">
  <title>commute time-lapse</title>
  <format>h.264</format>
  <original>commute.avi</original>
  <md5>ff17ed66e7db5693dd208dd6ac488ff8</md5>
  <mtime>1325973601</mtime>
  <size>11919082</size>
  <crc32>ad1df03a</crc32>
  <sha1>e9f9de8379cd25653d487ab30d198fc61a050091</sha1>
  <length>115.61</length>
  <height>480</height>
  <width>640</width>
</file>
```

- ~~striped RAID~~
- ~~distributed filesystems~~
- boring block devices

Circulate

Serving

- 300B+ web pages WayBack Machine
- 15M books
- 4M videos & TV News programs
- 4M audio & concerts
- 3M images
- 200K software items & emulation

...to ~4M Daily Visitors

Timeline / Evolution

Month	Progress
2014/11	`docker` for audio fingerprints - controversial
2015/10	`MozFest` OpenNews `docker` talk mind blown
2016/8	switched `mp3/mp4` generation to `docker`
2017/6	IA Engr. Presentation
2018/6	IA Engr. Special Topic talk
2018/7	GitLab 'Auto DevOps' Changes Everything - Dev & Ops Harmony - Confessions of a middle child

Timeline / Evolution

Month	Moves into GitLab Auto DevOps Kubernetes
2018/8	archive.org main repo CI/`test` migrated
2018/9	archive.org full pipelines, website 'review apps'
2018/10	bookscan webapps
2018/10	moved to GitLab's docker registry
2018/12	dweb.archive.org (archive.org decentralized)

LogJam BreakThroughs

- TeamUp! (end of 2017)
  - David - ops-heavy history 🍫

  - Tracey - dev-heavy history 🥜
- GitLab advertises `with Kubernetes` July/2018
  - **already using** on premise `GitLab`
  - w/ ssh-key user accounts .org-wide

'on prem' K8 cluster from scratch

- https://gitlab.com/internetarchive/kre8

  - `kubeadm` to create cluster
  - sets up GitLab 'Auto DevOps' CI/CD full pipeline
  - need 1+ VM w/ ssh access and sudo
    - ubuntu based so far
  - coworker can try - low committment
  - `kallr`

Avoid Bad Blood

Work Towards Common Goals

Devs & Ops

www-NAME (bespoke) -v-
Review Apps (GitLab + K8)

webapps quick development at archive.org

WWW-TRACEY.ARCHIVE.ORG

- ssh 'bastion' with homedirs
- NFS /home/ automounts throughout cluster
- edit:
  - emacs/vi on (linux) bastion
  - laptop editors sftp / rsync-over-ssh on `[save]`
- nginx virtual hosts
  - https://archive.org => `/petabox/www/..` (production)
  - https://www-NAME.archive.org => `/home/NAME/petabox/www/..` (dev)
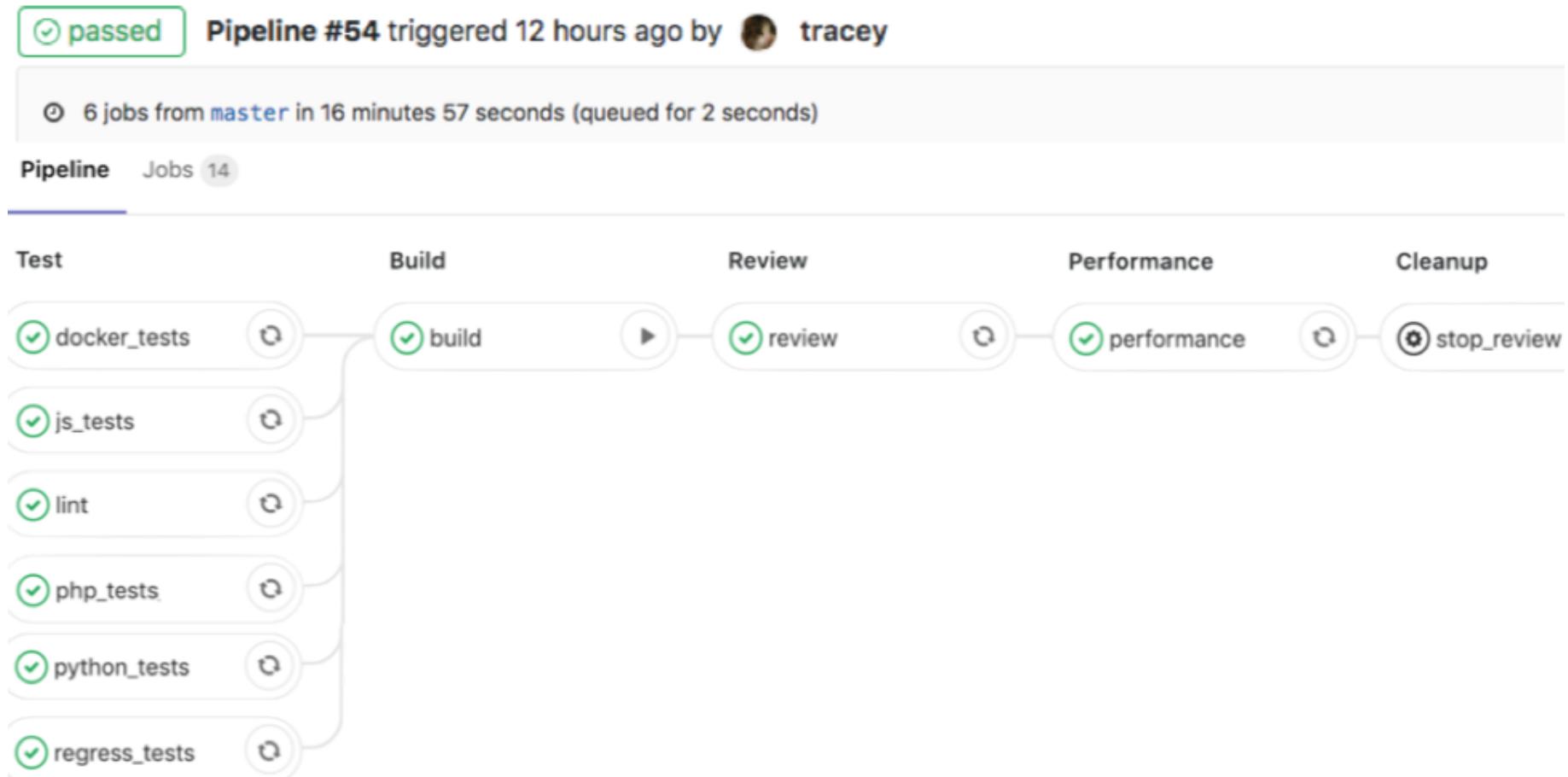
www-NAME.archive.org issues

- single 'dest' (multiple branches tricky)
- ssh keys and acccess
- NFS /home/ dirs

K8 + Auto DevOps 'review apps'

- contractors can work/test full site
  - no ssh keys/access
  - 12h timezone shift OK
    - no staff/approval need - `game changer!`
- git branches => 'review apps' consistent/update over time
- `kre8/branch-sync-edit.sh`
  - 'copy file change to pod on save' fast script
  - webapp instant f/b w/o full push and CI/CD

Auto DevOps Pipelines

Pipeline #54 triggered 12 hours ago by 🖼️ **tracey**

🕐 6 jobs from `master` in 16 minutes 57 seconds (queued for 2 seconds)

Pipeline Jobs 14

Test	Build	Review	Performance	Cleanup
✓ docker_tests	✓ build	✓ review	✓ performance	⚙ stop_review
✓ js_tests				
✓ lint				
✓ php_tests				
✓ python_tests				
✓ regress_tests				

Crawling Case Study

Crawl Discovery Loop

1. Start with a list of websites
2. Fetch all those URLs
3. Extract Hyperlinks from all the fetched content
   - Now you have a list of websites again
   - Which you can Fetch and Extract and...

Heritrix

- open source Java crawler developed at IA
- very good at pipelining these stages
- but, single machine operation

CrawlHQ

- distributed frontier management
- multiple Heritrix instances share single URL deduplication queue
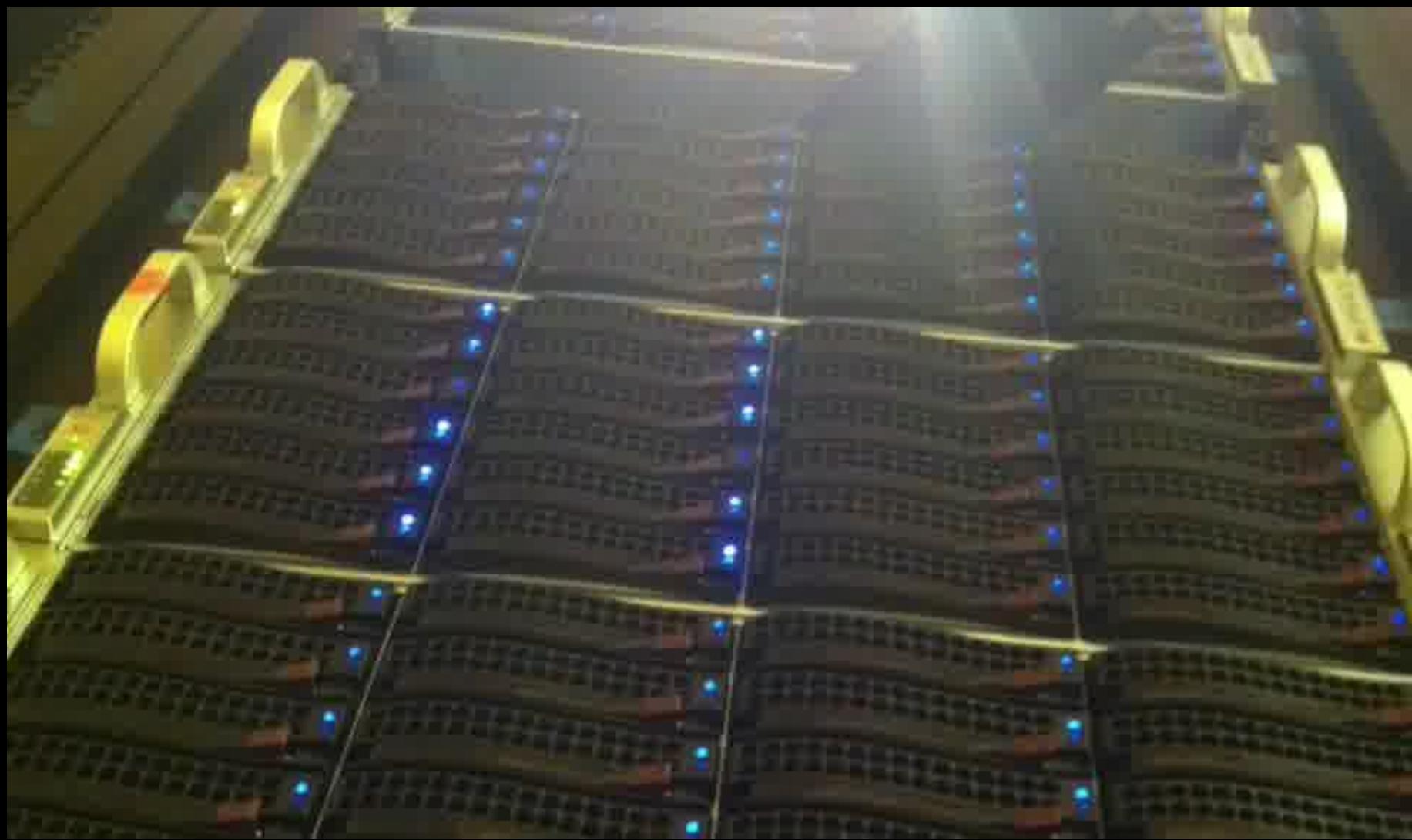
Demo of K8S auto-scaling
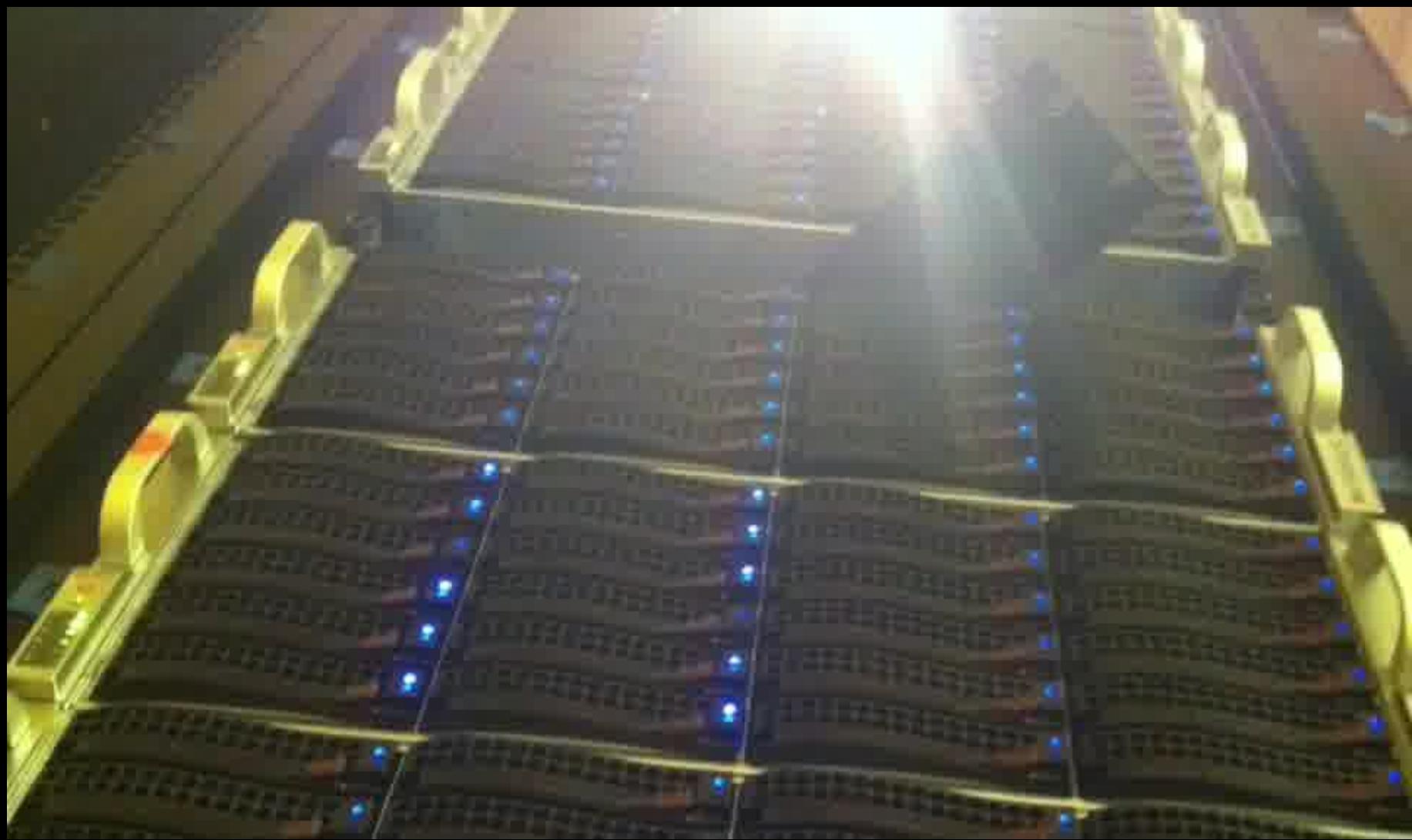
Rewriting CrawlHQ

- Kafka partitions for work queue
- FoundationDB for hashes of seen URLs
- 38,000 lines of code is now 73 lines of Python
- throughput scales automatically with GitLab `REPLICAS` variable
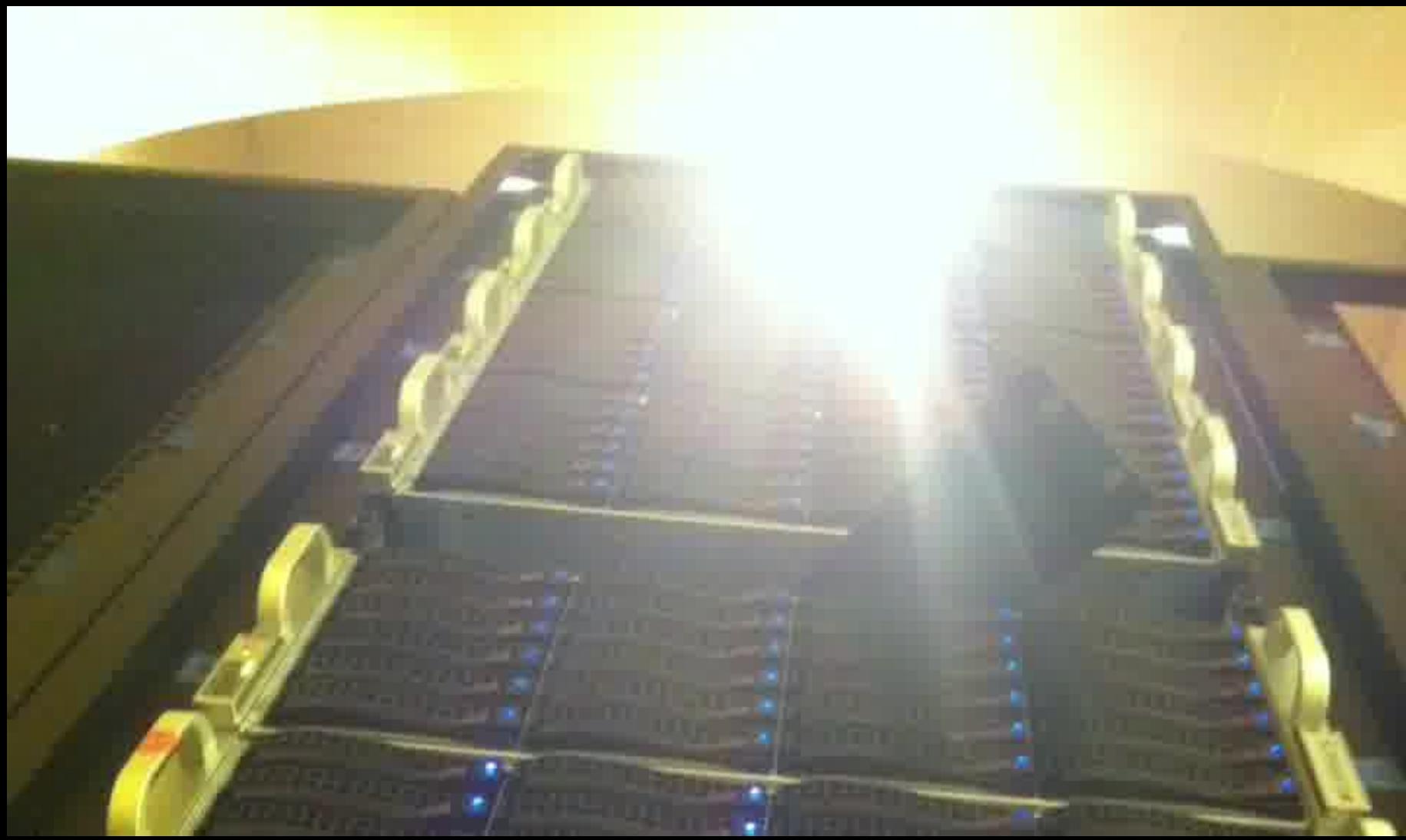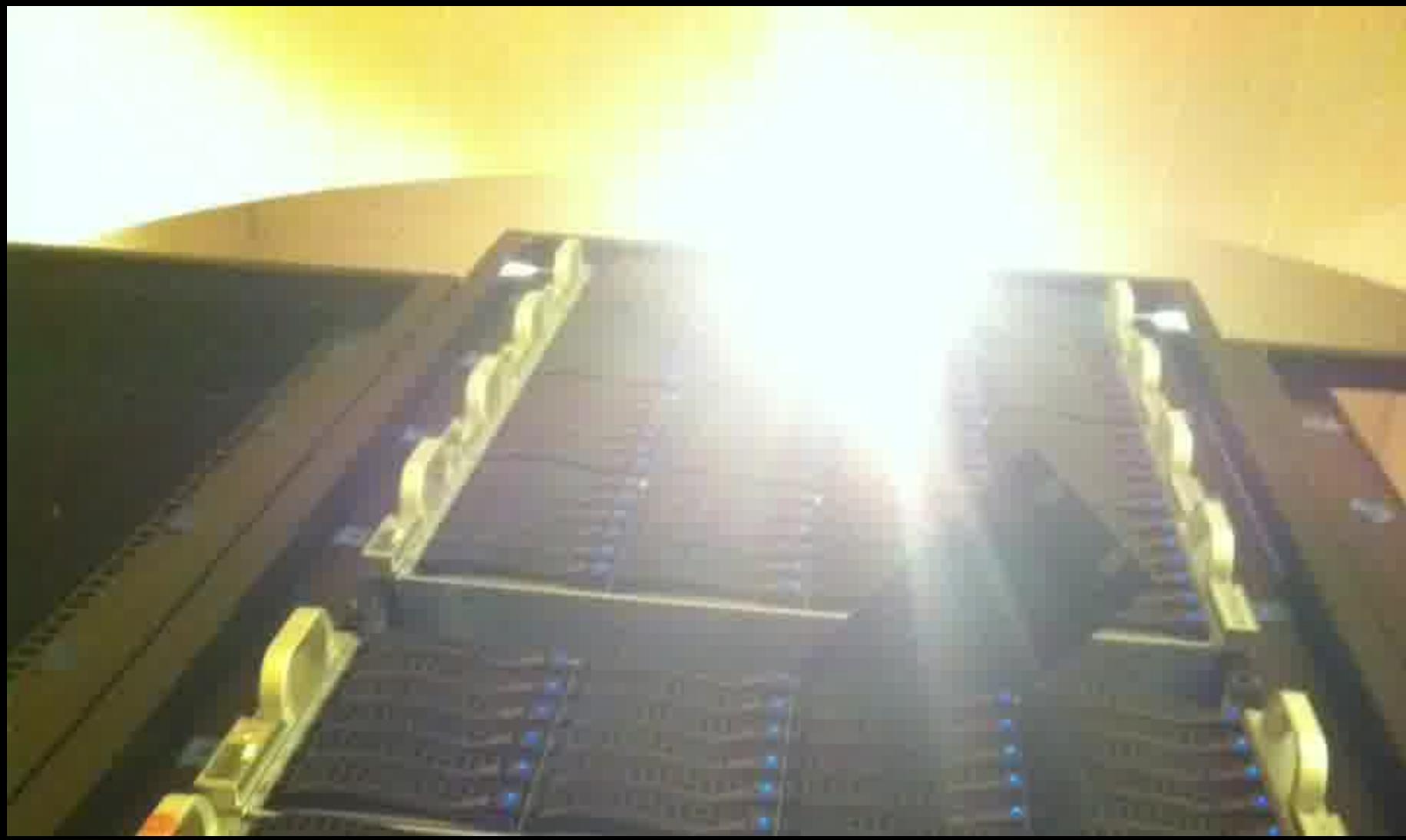
my god ..
it's full of.. petabytes

Containing a Monolith

- 4.8G tree => 6GB .. 9GB docker images
- `test` before `build` = faster dev feedback
- naive `git clone` every CI/CD phase = nonstarter

petabox Repo Docker Images

[petabase]

Baseline docker image
Ubuntu LTS
pkgs
js npm/yarn
php pecl/pear
python pip2/3
petabox repo git SHALLOW clone
CI test after git fast fwd

[petabase]

[workbase]	[webnode]	[dev]	[homenode]
3rd party deriver pkgs + config ffmpeg ..	git fast fwd archive.org website nginx+ supervisor	git fast fwd empty DB/SE daemons/supervisor archive.org website item store/process	git fast fwd misc nod...
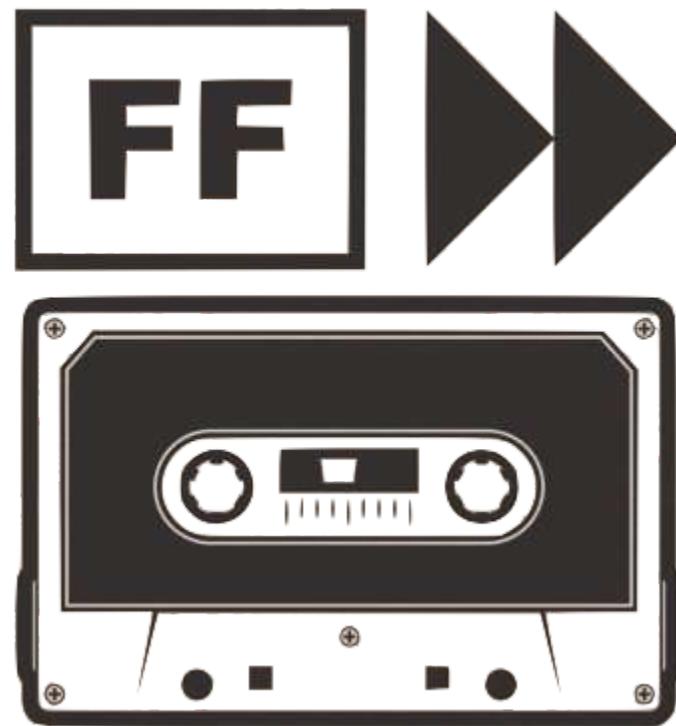
[worker]

git fast fwd
mp3/mp4/etc

Monolith repo speed++

- `[petabase]` baseline image recut every 1-4 weeks
  - git shallow clone (2/3 size)
  - all usages 'fast forward' to branch & commit
- `build` phase uses prior build
  - via registry docker pull for branch

How to Fast Forward Quickly from weeks old Git Shallow Tree?

* Monolith repo Git *

fast forward a git clone depth 1

- `git checkout --detach` *
- `git fetch --depth 1 -f -origin [brnch|ca6745]`

- `git checkout -f [brnch|ca6745]`

* avoid tangle with git tree state

- Workaround 'advertised' commits issue (self-hosted GitLab)

  - `omnibus gitconfig['system'] = {"uploadpack" => ["allowReachableSHA1InWant = true"]}`
  - then able to fetch any commit

BECAUSE WE HAVE THIS GUY (AND YOU DON'T)

Mark Sullivan *Web Archiving Software Engineer*
Mark joined the Archive in August 2016 to help Arc
of the Archive, with new data transfer tools. He gra
Computer Science and a Bachelors of Computer E
classes!). He has worked in different enough enviro
occasionally dreams of expressive computers, and

^ git pro

Top Tips

- can use simple `localstorage` for 1 or 2-node k8 clusters
- `POSTGRES_ENABLED=false` CI/CD var
  - avoids extra (unused!) PV per branch/deploy
- GitLab API for last master `test` phase status
  - 'OK to push live?'
- periodically:
  - remove old replicasets
  - trim registry

LoadBalancer vs Ingress for Alt Ports

- `kubectl expose -ndweb deployment/production --type=LoadBalancer --name=porter`
- `kubectl patch -ndweb svc/porter -p='{"spec":{"externalIPs":["'${K8_INGRESS_IP?}"]}}'`
- allows for non :80 :443 :5000 ports
  - websockets eg:
    - wss://dweb.me:4246

Near Future

archive.org production website => K8

now	static list of VMs + haproxy
new	elastic demand-based scaling

the Internet Archive

to Collect & to Serve

The End