

# Hacking the Pachyderm

## Scaling Servers and People

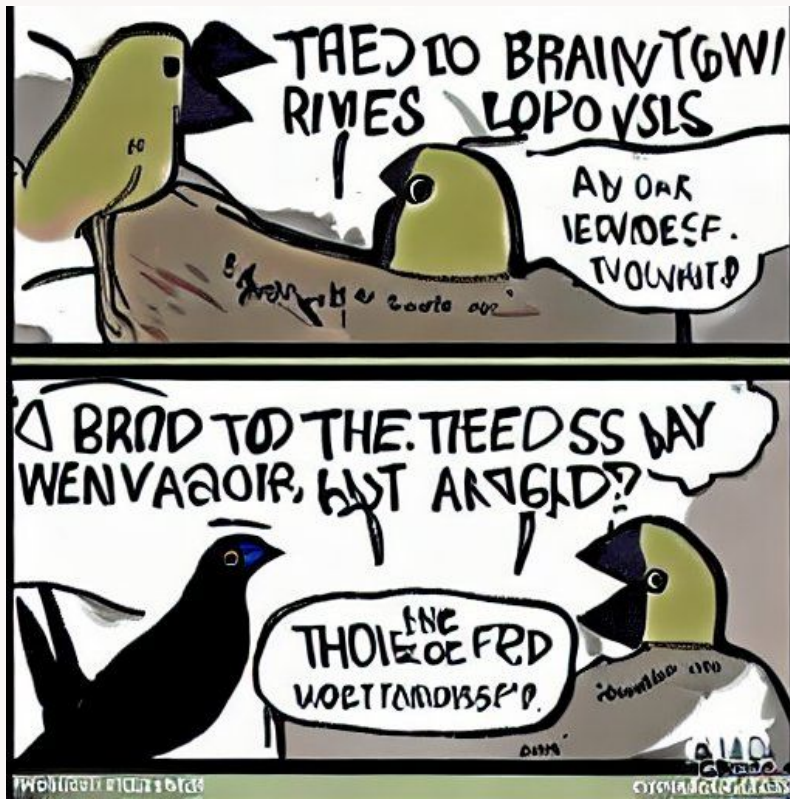
Hazel Weakly, Hachyderm

Preston Doster, Hachyderm

@hazelweakly@hachyderm.io

@esk@hachyderm.io





Art courtesy of AI magic

@hazelweakly@hachyderm.io  
@esk@hachyderm.io




Art courtesy of AI magic



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

```
curl -X GET https://agenda
```

 Hachyderm becoming one of the largest elephant watching sanctuaries

 Challenges we faced in undertaking the Great Elephant Migration

 What's next



# Your RCA is Not My RCA

*But Your RCA is Also Useless*



Root cause: 2 failing  
HDDs

@hazelweakly@hachyderm.io  
@esk@hachyderm.io

# Your RCA is Not My RCA

*But Your RCA is Also Useless*



## **Perfect hindsight solution:**

1. Upload media to object storage from the start
2. Migrate a week earlier

✨ done ✨

# Time Loop Paradox of Distributed Systems



Any problem that is  
prevented before it occurs  
will spontaneously  
reincarnate elsewhere

...

but *weirder*  
and more *fucked up*

“Root Causes”  
are Individual  
Insights



Systems Resilience  
requires Systemic  
Adaptation

@hazelweakly@hachyderm.io  
@esk@hachyderm.io



Make  
Change  
Easy



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

Making change  
easy to handle

≠

Making change  
easy to do



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

# The Great Elephant Migration



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

# The Old Habitat



Experimental ZFS build



Debug linux kernel for testing root exploits



Mastodon serves media off of NFS



Main host has: Mastodon, Nginx, Kubernetes,  
Postgres, and Redis.



# The Timeline



Nov. 19: Building v2 hachyderm with NixOS



Nov. 27: NFS is unusable, we attempt to fix



@hazelweakly@hachyderm.io

@esk@hachyderm.io

# Feat: nixos is the future!

Merged krisnova merged 21 commits into `main` from `nixos-hosts` on Nov 22, 2022

Conversation 7

Commits 21

Checks 0

Files changed 69



hazelweakly commented on Nov 20, 2022

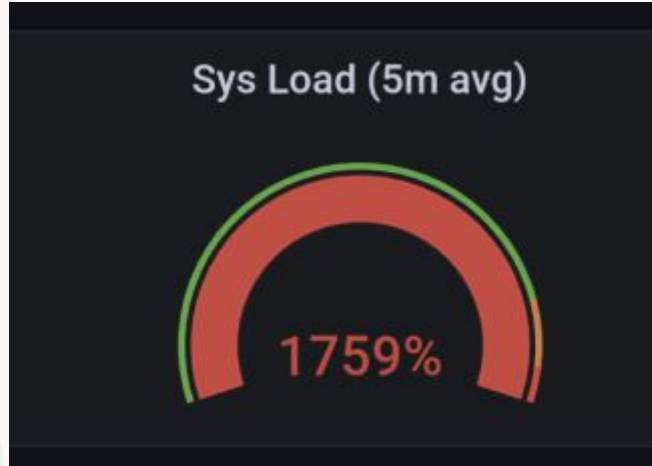


Bye arch!



@hazelweakly@hachyderm.io

@esk@hachyderm.io



Hi! Sys Load expert here!  
This is not funny,  
computers only do this  
when they're in extreme  
distress



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

# I accidentally

the largest and most comprehensive  
collection of notes on scaling mastodon

*Whoops*

@hazelweakly@hachyderm.io

@esk@hachyderm.io





# The Timeline



Nov. 19: Building v2 hachyderm with NixOS



Nov. 27: NFS is unusable, we attempt to fix



Nov. 28: Unable to process queues fast enough



Nov 28: Run Hachyderm in both EU and US



@hazelweakly@hachyderm.io

@esk@hachyderm.io

# The Cycle of Toil



EU wakes up. Balances  
sidekiq queues  
manually



US wakes up. I tell  
everyone the optimal  
numbers via  
calculation



We're both wrong



GOTO :weed-leaf:

@hazelweakly@hachyderm.io

@esk@hachyderm.io

# teaf: revert nixos is the future! Arch Linux is the future!

# #2

Merged krisnova merged 21 commits into `main` from `nixos-hosts` on Nov 22, 2022

Conversation 7

Commits 21

Checks 0

Files changed 69



hazelweakly commented on Nov 20, 2022



The `yay`s will continue until stability improves



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

# Breaking the Toil Cycle

One of the highest leverage things done was Nova introducing these two things at the same time:



All changes needed to be logged via PR



Scripts which explicitly listed the services to modify



@hazelweakly@hachyderm.io

@esk@hachyderm.io

# Breaking the Toil Cycle

"I thought you said to make change easy"



Redistributing friction in key areas balanced the human workflows and streamlined communication



Humans need backpressure and circuit breakers too



@hazelweakly@hachyderm.io

@esk@hachyderm.io

# The Timeline



Nov. 28: Unable to process queues fast enough



Nov 28: Run Hachyderm in both EU and US



Nov 29: Start moving to object storage + tmpfs



Nov 30: prepare to move the DB to the EU



@hazelweakly@hachyderm.io

@esk@hachyderm.io

# How 2 make postgres h8 u

High number of connections 2354



"stuff is slow, add more workers"



(this adds connections to postgres and slows it down)



postgres: \*doing more IO\*



realize queries take too long



you run out of connections



GOTO PAIN

@hazelweakly@hachyderm.io

@esk@hachyderm.io

# NFS: Not Fucking Scaleable



Tani Aura  11/29/2022 3:22 PM

The rails app is effectively doing this:

1. Start db transaction
2. Upload image to NFS
3. INSERT or UPDATE statement
4. Commit the transaction





How many  
IOPS does it  
take to...





300 IOPS  
12 fsync/s



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

“NFS is too reliable  
and performant  
for me, how do  
I fix that?”



Answer:  
run it over the atlantic  
ocean and store all the files  
on a system with two failing  
SSDs

@hazelweakly@hachyderm.io  
@esk@hachyderm.io

# Why didn't we notice this earlier?

`we_had_no_idea_what_we_were_doing.jpg.mp3`



Metrics are data



Data is only as good as your understanding





No(thing) was ever obviously wrong

@hazelweakly@hachyderm.io

@esk@hachyderm.io

# Why did we take the approach we did?

 We wanted to not be beholden to corporate ulterior motives (see: Twitter)

 We wanted to build a service that was meaningfully owned by the community it served.



We should have  
burned out



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

# Why didn't we burnout?



Intrinsic vs extrinsic motivation



Aligned incentives throughout the process



Psychological safety



We chose our own solution and its downsides



We continually supported each other



You'll never be  
prepared for  
exponential  
scale



... But you *can* be prepared  
to support each other  
through the challenges  
you'll face

@hazelweakly@hachyderm.io  
@esk@hachyderm.io



# Operating a mastodon instance is great

It's all the reliability and ease of maintenance of a rails application, combined with the user experience of hosting a mail server, and the developer workflow of day drinking on the Linux kernel mailing list.



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

Soo... what's next?



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

# Today



Production service, ~**20k MAU**



**Stateful:** people care about their toots!



**Real-time:** people care about their toots!



Key constraint: **volunteer** time is limited



# Infrastructure values

 Brian Cantrill's [Platform as a Reflection of Values](#) - watch it!

 Shape our technology **choices**

 Inform how we **prioritize** our time






What do we **value most** as  
the Hachyderm Infra team?



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

# Approachability

-  Low barrier to entry
-  Open to all backgrounds & levels
-  Let volunteers use their limited time **effectively**



# Safety

 Encourage experimentation & growth

 **It's OK to break prod**, so let's **learn** from it when we do!





 We're in it together



@hazelweakly@hachyderm.io

@esk@hachyderm.io

# Transparency

-  No hidden agendas
-  Docs, docs, docs
-  Clear team backlog
-  Public incident reviews - share always!





# Key Choices



@hazelweakly@hachyderm.io  
@esk@hachyderm.io

# NixOS

 Base OS & operating environment

 Self-described: **Reproducible, Declarative, and Reliable**

 Safe and Approachable\*

2 v2/nixos/nixos/hosts/pgsql/default.nix

@@ -7,7 +7,7 @@

7 system.stateVersion = "23.05"; # Do not change lightly!

8

9 services.postgresql.enable = true;

10 - services.postgresql.package = pkgs.postgresql\_11;

7 system.stateVersion = "23.05"; # Do not change lightly!

8

9 services.postgresql.enable = true;

10 + services.postgresql.package = pkgs.postgresql\_14;

# Multi-cloud



From Day 1



Hetzner, Linode, DigitalOcean, Nova's  
Basement



Paradox: many clouds yield simplicity



# lots of stuff we don't need vs. the stuff we do

All services

Services by category

- Compute**
  - EC2
  - LightSail
  - Lambda
  - Batch
  - Elastic Beanstalk
  - Serverless Application Repository
  - AWS Outposts
  - EC2 Image Builder
  - AWS App Runner
  - AWS Sinspace Weaver
- Containers**
  - Elastic Container Registry
  - Elastic Container Service
  - Elastic Kubernetes Service
  - K8s for OpenShift Service on AWS
- Storage**
  - S3
  - EFS
  - FSx
  - S3 Glacier
  - Storage Gateway
  - AWS Backup
  - AWS Elastic Disaster Recovery
- Database**
  - RDS
  - ElastiCache
  - Neptune
  - Amazon QLDB
  - Amazon DocumentDB
  - Amazon Keyspaces
  - Amazon Timestream
  - DynamoDB
  - Amazon MemoryDB for Redis
- Migration & Transfer**
  - AWS Migration Hub
  - AWS Application Migration Service
  - Application Discovery Service
  - Database Migration Service
  - AWS Transfer Family
- Developer Tools**
  - CodeStar
  - CodeCommit
  - CodeArtifact
  - CodeBuild
  - CodeDeploy
  - CodePipeline
  - Cloud9
  - CloudShell
  - X-Ray
  - AWS FIS
  - Amazon CodeCatalyst
  - AWS AppConfig
  - Amazon CodeWhisperer
  - Application Composer
- Customer Enablement**
  - AWS IQ
  - Managed Services
  - Activate for Startups
  - Support
- Robotics**
  - AWS RoboMaker
- Blockchain**
  - Amazon Managed Blockchain
- Satellite**
  - Ground Station
- Quantum Technologies**
  - Amazon Braket
- Management & Governance**
  - AWS Organizations
  - CloudWatch
  - AWS Auto Scaling
  - CloudFormation
  - Config
  - OpsWorks
  - Service Catalog
  - Systems Manager
- Machine Learning**
  - Amazon SageMaker
  - Amazon Augmented AI
  - Amazon CodeGuru
  - Amazon DevOps Guru
  - Amazon Comprehend
  - Amazon Rekognition
- AWS Cost Management**
  - AWS Cost Explorer
  - AWS Budgets
  - AWS Marketplace Subscriptions
  - AWS Application Cost Profiler
  - AWS Billing Conductor

Google Cloud

Management

Compute

Storage

Analytics

Networking

Distributed Cloud

Serverless

Databases

Operations

Security

CI/CD

Artificial Intelligence

Integration Services

Tools

Other Google products

Support

**Security**  
Meet your policy and compliance objectives

Name	Description
<a href="#">Security</a>	Security management controls and capabilities
<a href="#">Compliance</a>	Tools that help support data regulations

**CI/CD**  
Integrate and deliver continuously

Name	Description
<a href="#">Cloud Build</a>	Continuous integration delivery platform
<a href="#">Container Registry</a>	Private container registry storage
<a href="#">Source Repositories</a>	Hosted private git repos
<a href="#">Artifact Registry</a>	Universal build artifact management
<a href="#">Cloud Deploy</a>	Managed continuous delivery to GKE

**Artificial Intelligence**  
Leverage machine learning products on a trusted platform

Name	Description
<a href="#">Vertex AI</a>	One AI platform, every ML tool you need
<a href="#">Natural Language</a>	Custom text models
<a href="#">Tables</a>	Custom data models, now also in Vertex AI
<a href="#">Translation</a>	Language detection and translation
<a href="#">Document AI</a>	Document analysis, classification, and searches
<a href="#">Recommendations AI</a>	Custom recommendations for products
<a href="#">Retail</a>	Data-driven solutions for retailers
<a href="#">Talent Solution</a>	Job search with ML
<a href="#">DocAI Warehouse</a>	AI-integrated cloud-native document management platform
<a href="#">Discovery Engine</a>	Cross-industry discovery solutions

**HETZNER**

- Servers
- Volumes
- Load Balancers
- Floating IPs
- Networks
- Firewalls
- Security

**Akamai**

- Linodes
- Volumes
- NodeBalancers
- Firewalls
- StackScripts
- Images
- Domains
- Databases
- Kubernetes
- Object Storage
- Longview
- Marketplace

# Terraform Cloud

 Declarative, Reproducible, and **Reliable**

 **Lowers barrier** to manage our infra

 **Predictable** workflow - PR, plan, apply



@hazelweakly@hachyderm.io

@esk@hachyderm.io

PR-driven workflow (familiar, comfortable, approachable)

Preskton commented last week • edited

Destroying my pgsqL sandbox host now that I've tested the restore in #285 .

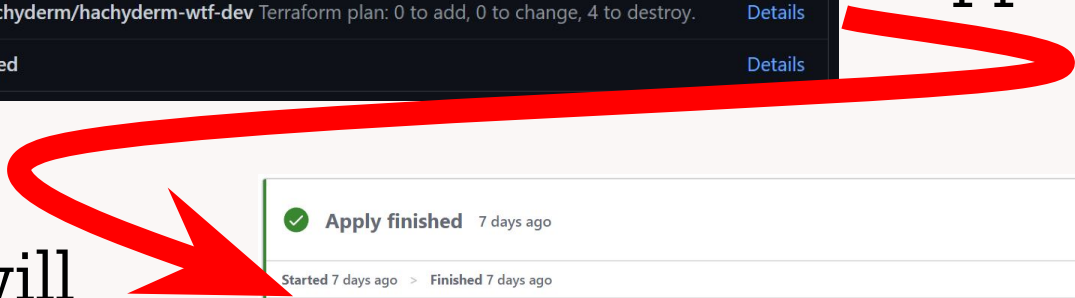
Destroying pgsqL sandbox Verified ✓ 0566bf6

cla-bot bot added the **Access Granted** label last week

Preskton merged commit `bff3d38` into `main` last week Hide details Revert

2 checks passed

- ✓ Terraform Cloud/hachyderm/hachyderm-wtf-dev Terraform plan: 0 to add, 0 to change, 4 to destroy. Details
- ✓ verification/cla-signed Details



Know what will happen before merge (safety, confidence)

✓ **Apply finished** 7 days ago Resources: 0 added, 0 changed, 4 destroyed  
Run Triggers: 1 runs queued

Started 7 days ago > Finished 7 days ago

4 destroyed

Filter resources by address... Filter by action Terraform 1.3.7 Download raw log

> - aws	module.hetzner_esk_pgsqL_sandbox2.aws_route53_record.host...	✓ Deleted	id=z03419831fnokffi7pyrk_pgsqL-big-es...
> - aws	module.hetzner_esk_pgsqL_sandbox2.aws_route53_record.host...	✓ Deleted	id=z03419831fnokffi7pyrk_pgsqL-big-es...
> -	module.hetzner_esk_pgsqL_sandbox2.hcloud_placement_group...	✓ Deleted	id=136512
> -	module.hetzner_esk_pgsqL_sandbox2.hcloud_server.servers[0]	✓ Deleted	id=29821905

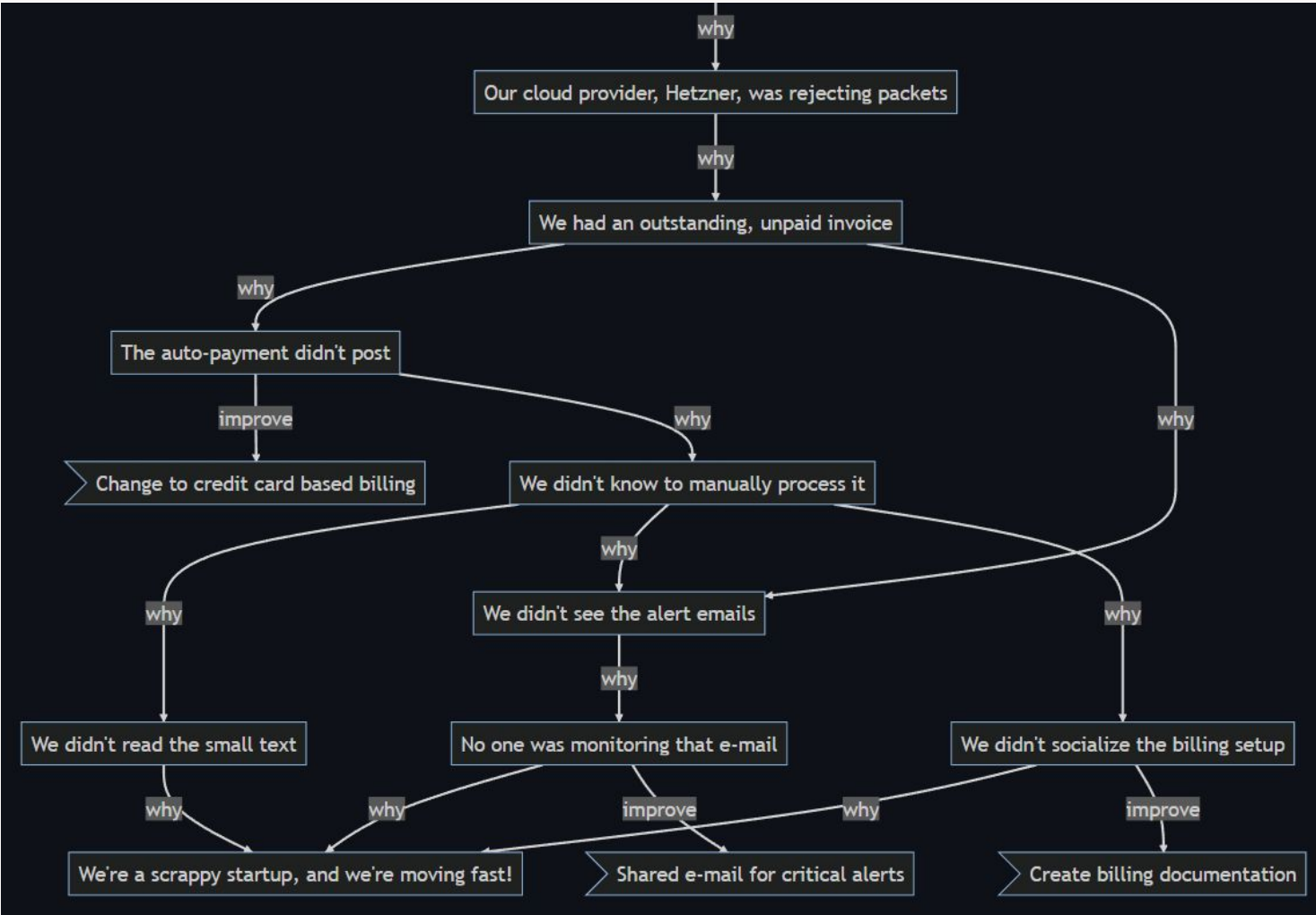
# Write it Down

 Something is better than nothing

 PR comments, good issue descriptions,  
Markdown docs in repo

 **Anything** the next human can use







Q + A



@hazelweakly@hachyderm.io  
@esk@hachyderm.io