

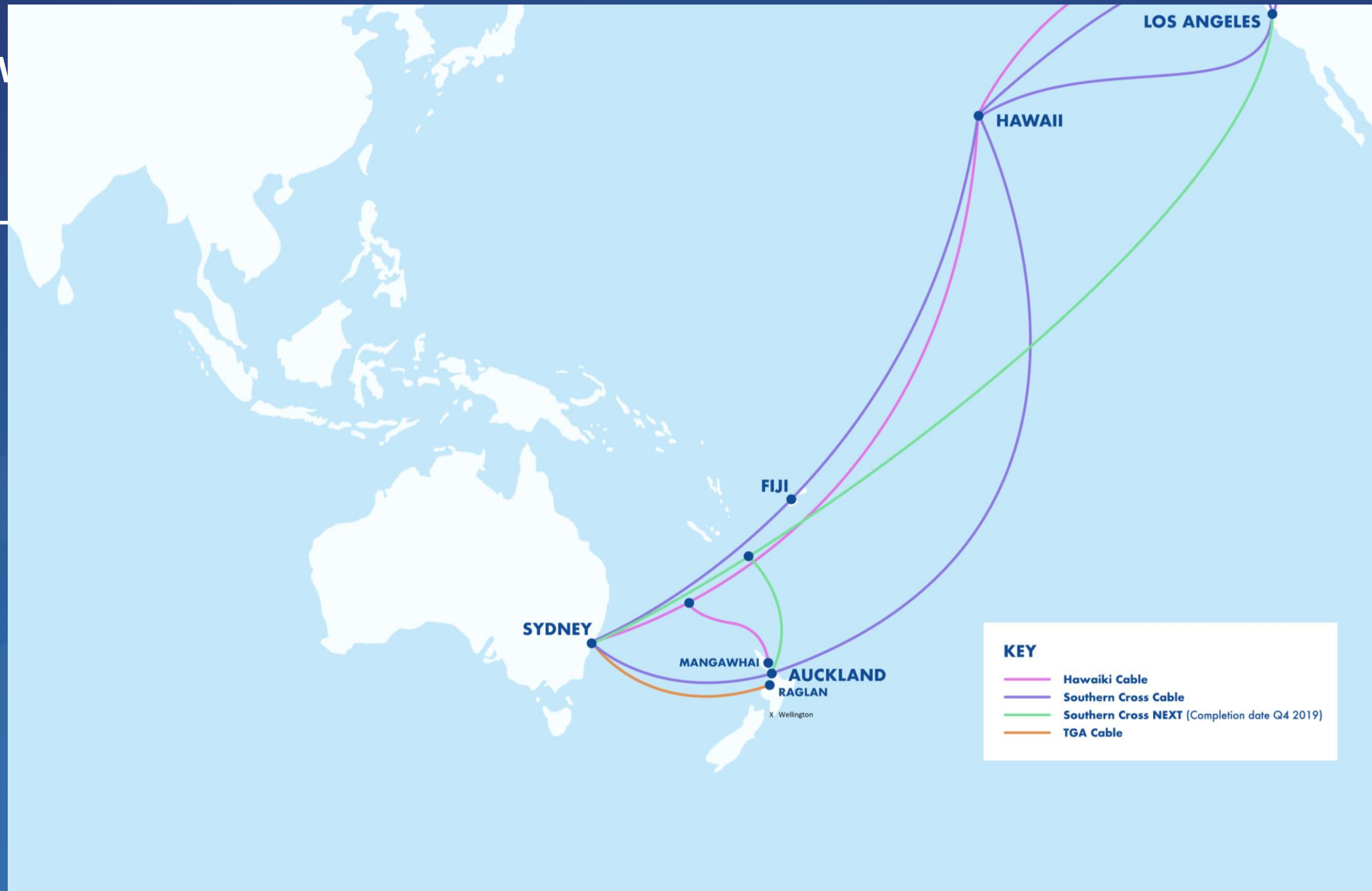


A journey into the long white Cloud

Andy Ziegler

W

L



Why did NZ MetService decide to go into the Cloud

Leading up to 2012 – the main emerging drivers

- Cost
 - Data centre at capacity and huge investment coming up if
 1. HQ data centre brought to required resilience (2nd power line, 2nd internet, higher earthquake rating, ...)
 2. Redundant capacity in other NZ data centre established (still prone to volcano or earthquake)
- 50% of revenue from clients overseas

Legacy design for single data centre and conventional infrastructure
Challenge and Opportunity!

Strategic Decision: Amazon Cloud best path forward!

But no Amazon Cloud in NZ!!



Why did NZ MetService continue to go into the Cloud

Exploring the Cloud and seeing “Sunlit uplands”

Unlimited
resources

- Most expensive (modelling) infrastructure has peaky demand
- No waiting for data in science at no extra cost, no queue
- Global footprint for resilience at controllable risk and cost
- No compromises for a robust release process
- Self-healing infrastructure “farming”

Cost

- Improve innovation agility turning infrastructure into reusable code
- Region agnostic deployment – all Clouds look the same
- Cheap spot-price market
 - looking for today’s cost downdraft in the Cloud



Fun

- It makes for great puns



Reproducible Workflows for Scientists

Where are the benefits in this?

Aim: Increase speed of agility of ideas
into inventions
into innovations.

Make infrastructure easy to use

- Infra virtually identical in research, test and prod
- Provide same building blocks and workflows in each



Reproducible Workflows for Scientists

Aim: Increase speed of agility of ideas
into inventions
into innovations.

Infrastructure easy to use

Cost optimised

- Reduce waiting for data (Labour \$\$\$ > Compute \$)
- Stick to one live version in use (own and 3rd party)
- One convention of names, deploys, OS, compilers, ...



Reproducible Workflows for Scientists

Aim: Increase speed of agility of ideas
into inventions
into innovations.

Infrastructure easy to use
Cost optimised

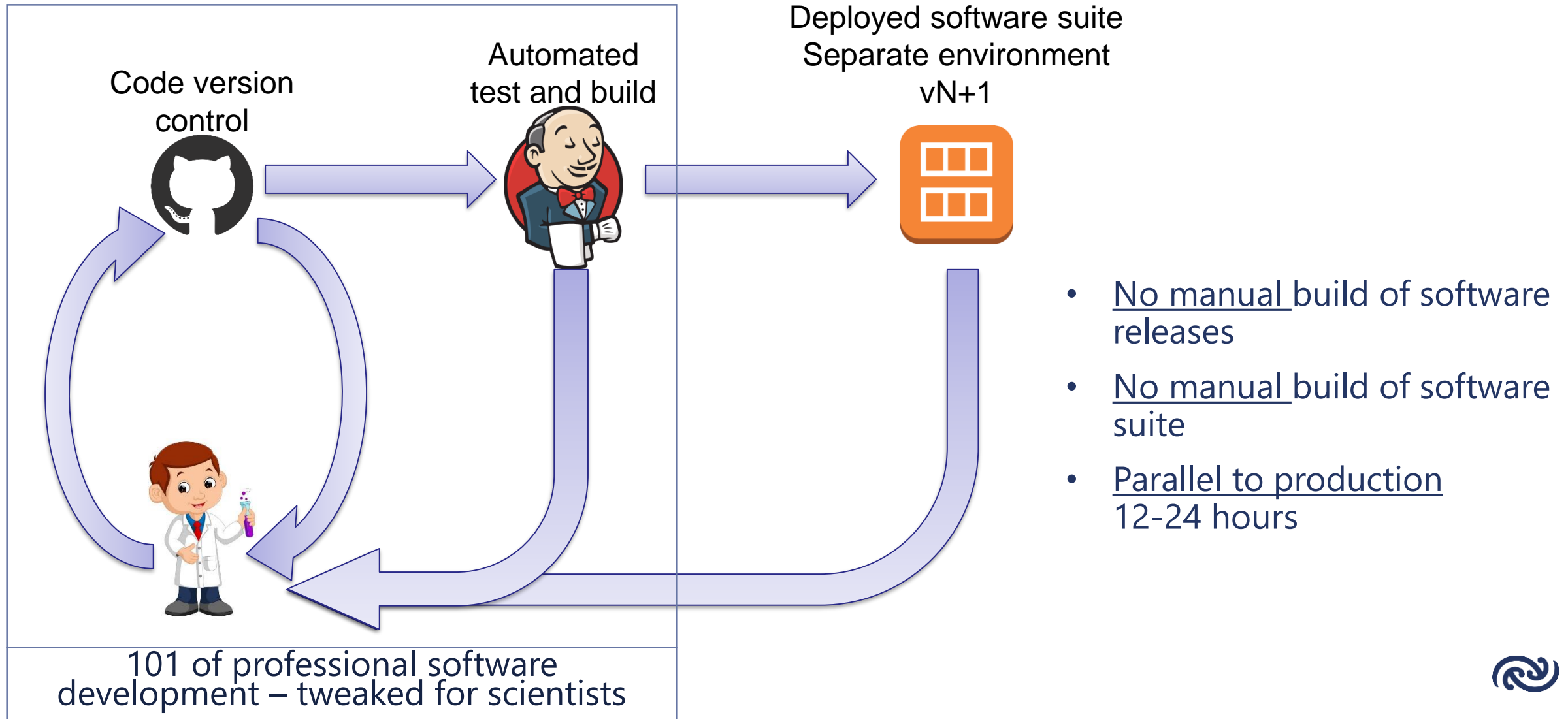
Robust path to production with predictability of

- Performance of promised solutions
- Cost of promised solutions
- Output of promised solutions



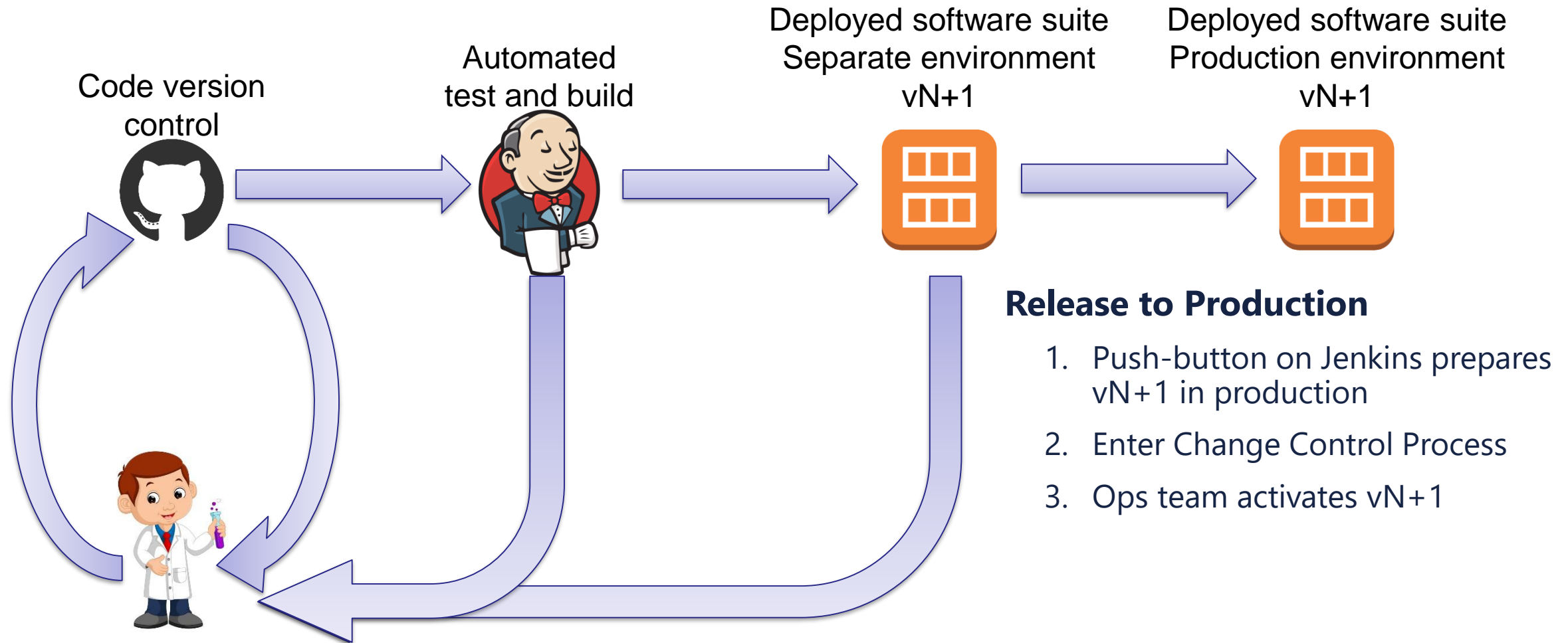
Enablers for Research and Development

From Idea to Invention



Enablers for Research and Development

Release to Production

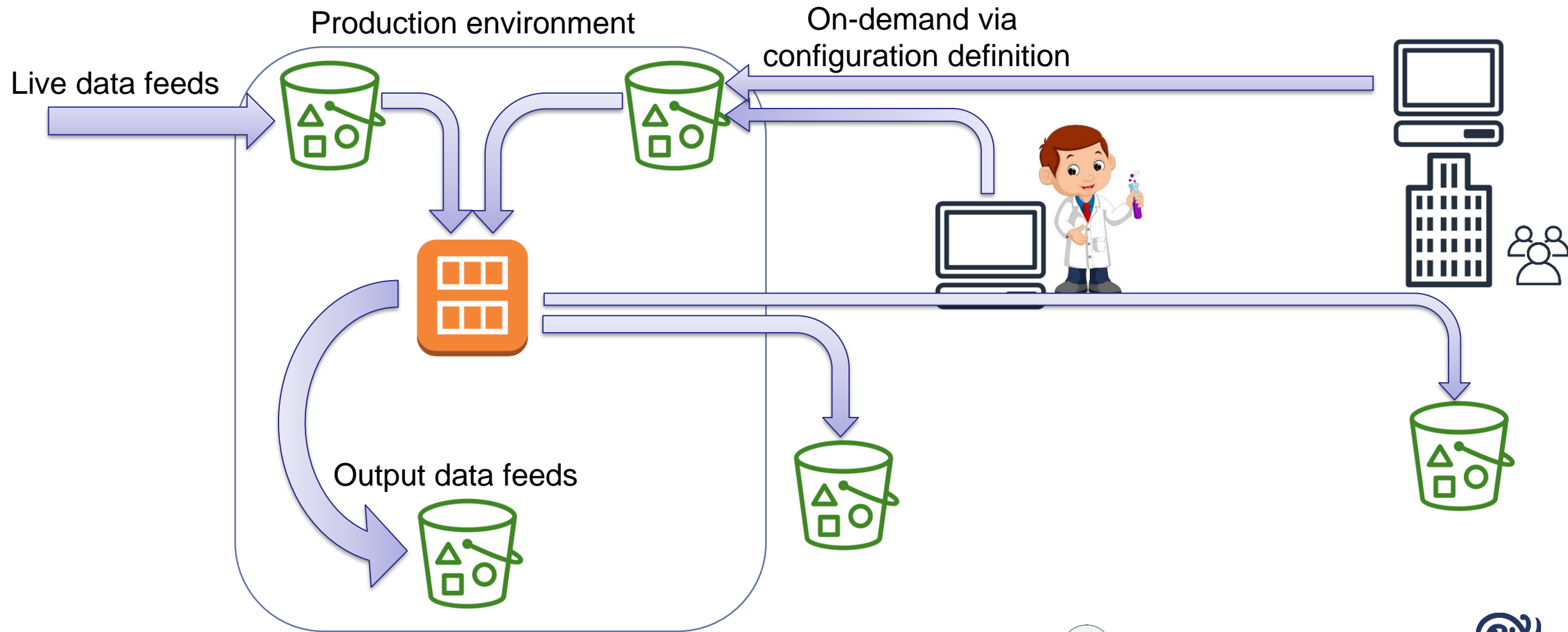


Breaking down the Cloud



Reproducible workflows

Make it once for use in different ways



What are the Cloud particles?



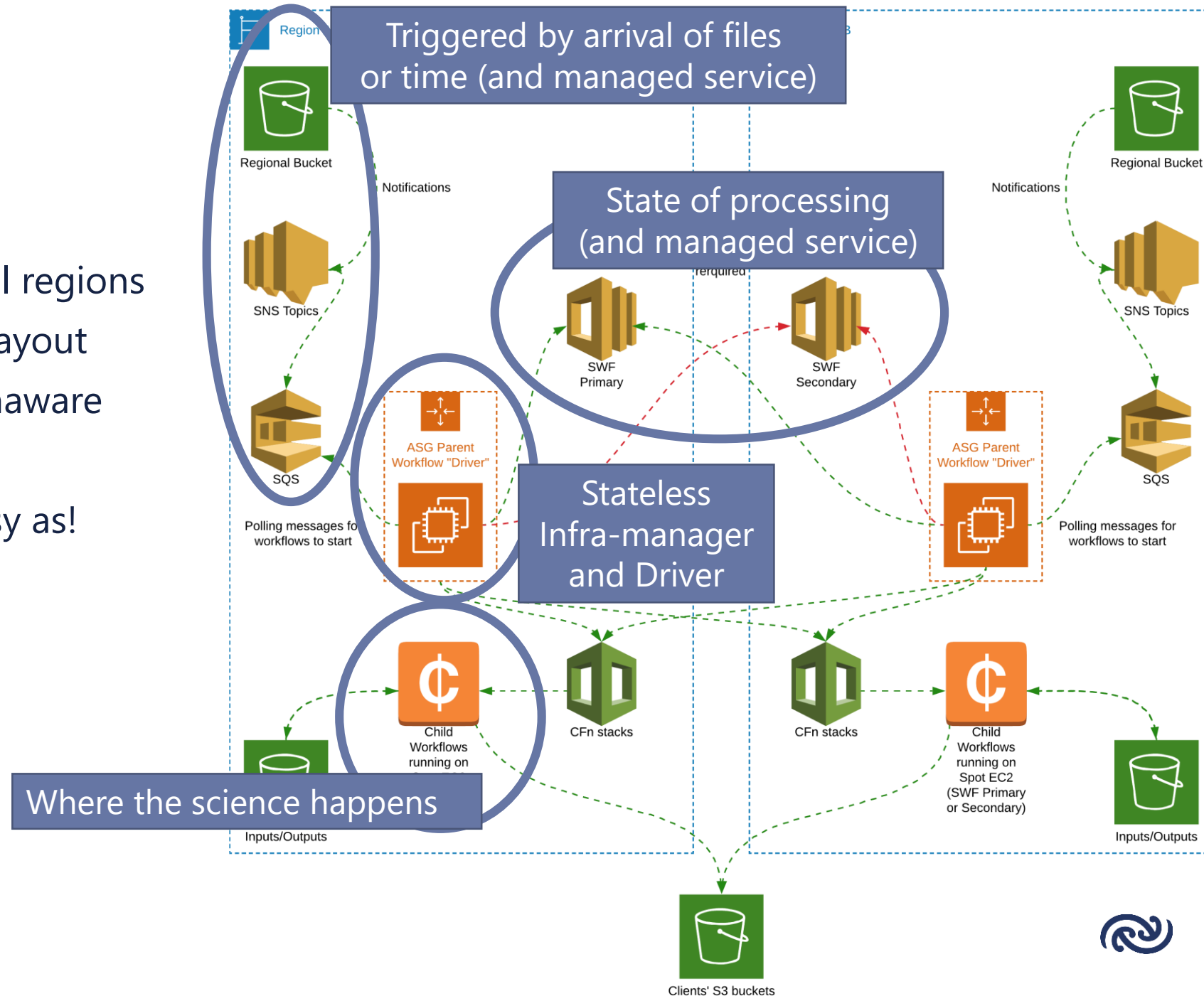
The Cloud particles

What keeps the Cloud moving?

Agnostic design:

- Identical input data in all regions
- Identical infrastructure layout
- Location and "mode" unaware

Add or remove region – easy as!



Expect thunderstorms in the Cloud



- Conflict between Access Management vs. Agnostic design
- Be paranoid about cost
 - Monitor for protection, not for blame
- There are limits
 - And some you will only know about once you hit them
- Think SciDevOps during design – learning curve is huge
- Break with conventional datacenter thinking
 - Think out of the box – literally
 - Expect resistance!



There will be resistance



Frodo



Gandalf



Sam



Gollum

Become a Farmer



And once they've done their work





What's the difference between wild and farmed cattle?



Just because it's cattle doesn't mean it's without owner.

Learned this the hard way!

Easy come YES, easy go **NOT SO** – keep the Cloud white and shiny



Are you on cloud nine yet?

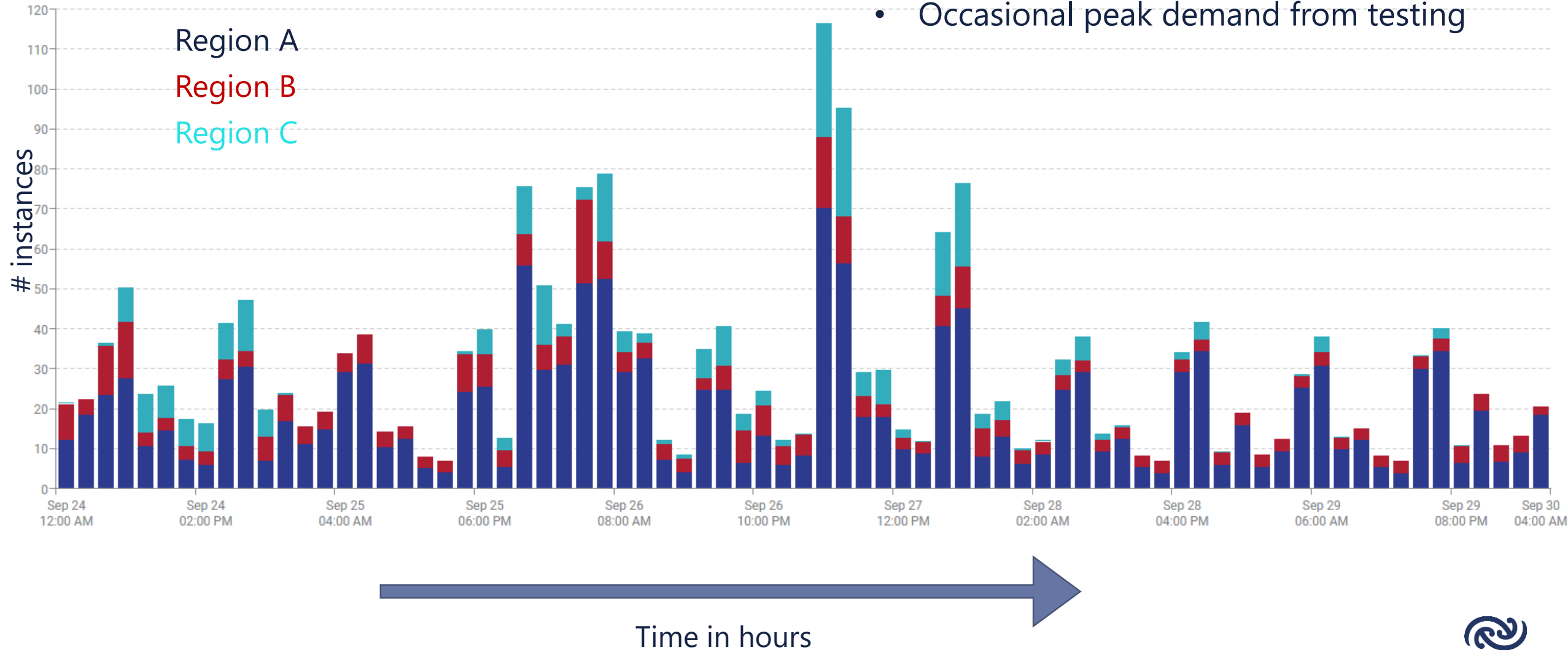
- All cattle farming
- All NWP models run asap
 - Some now finish before they started in HQ data centre
- Faster R&D results (LSM, ECMWF sigma, ...) and roll out to production
- Himawari 8 volcanic ash detection every 10 min
 - If eruption detected => HYSPLIT dispersion ensemble with eruption params
- Significant changes to "Infra-manager and Driver" robustly tested and released
- Spot price usage provides about 60% savings
- 75% cheaper than owning infrastructure with compromises
- It was the right strategic decision!



Hourly # instances

Hourly # instances

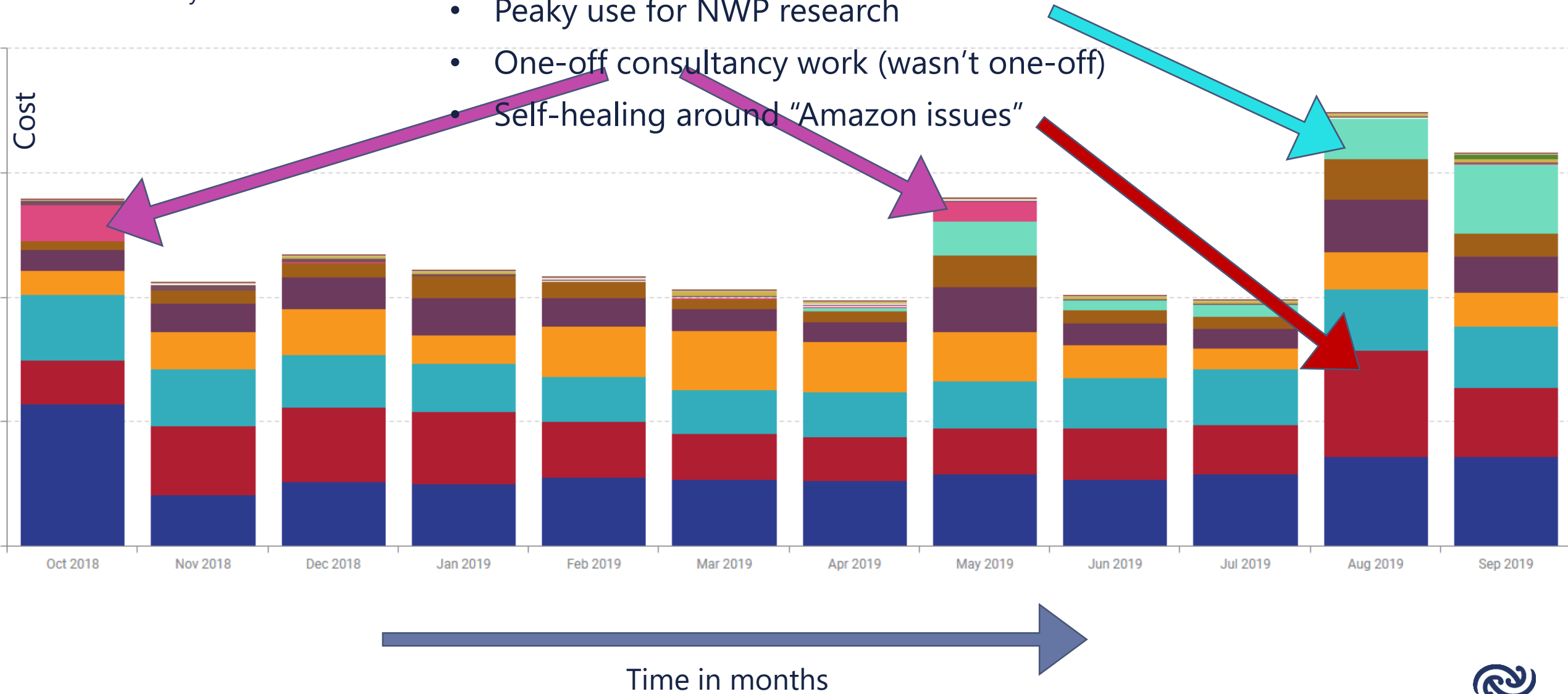
- Preferred region changes flexibly
- 12-hourly pattern
- Occasional peak demand from testing



Is the cost as elusive as the Clouds?

Cost analysis

- Peaky use for NWP research
- One-off consultancy work (wasn't one-off)
- Self-healing around "Amazon issues"



Tipps & Tricks - condensation of cloud experiences

- Automate, automate, automate
 - There is no one-off experiment/deployment/build!
 - Automation drives conventions, simple conventions drive agnostic design
 - It's a path that requires patience but pays off hugely later
 - 1. Explore, 2. Confirm, 3. Automate
- Establish professional software development tools and practices
- Infrastructure as Code with a location and "mode" agnostic design
- It's totally worth it!





Thank you 