

amazon.com's Journey to the Cloud

John Rauser - @jrauser
QConSF - November 2011

Amazon Retail != Amazon Web Services

Amazon Retail *is a customer of*
Amazon Web Services

1994 - 2011

July 5, 1994

Cadabra, Inc.

July 16 , 1995

amazon.com launches



Welcome to Amazon.com Books!

*One million titles,
consistently low prices.*

(If you explore just one thing, make it our personal notification service. We think it's very cool!)

SPOTLIGHT! -- AUGUST 16TH

These are the books we love, offered at Amazon.com low prices. The spotlight moves **EVERY** day so please come often.

ONE MILLION TITLES

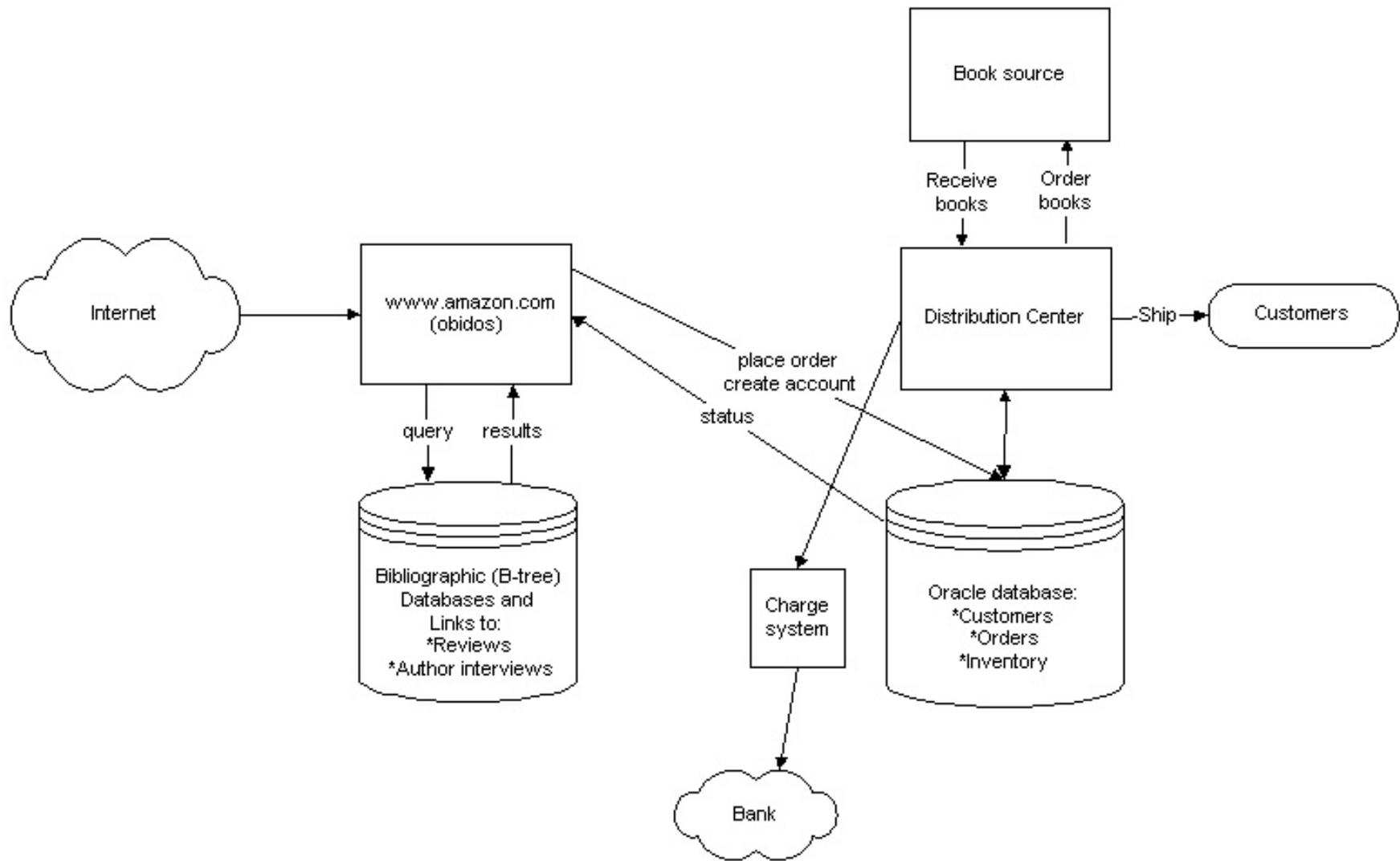
Search Amazon.com's [million title catalog](#) by author, subject, title, keyword, and more... Or take a look at the [books we recommend](#) in over 20 categories... Check out our [customer reviews](#) and the [award winners](#) from the Hugo and Nebula to the Pulitzer and Nobel... and [bestsellers](#) are 30% off the publishers list...

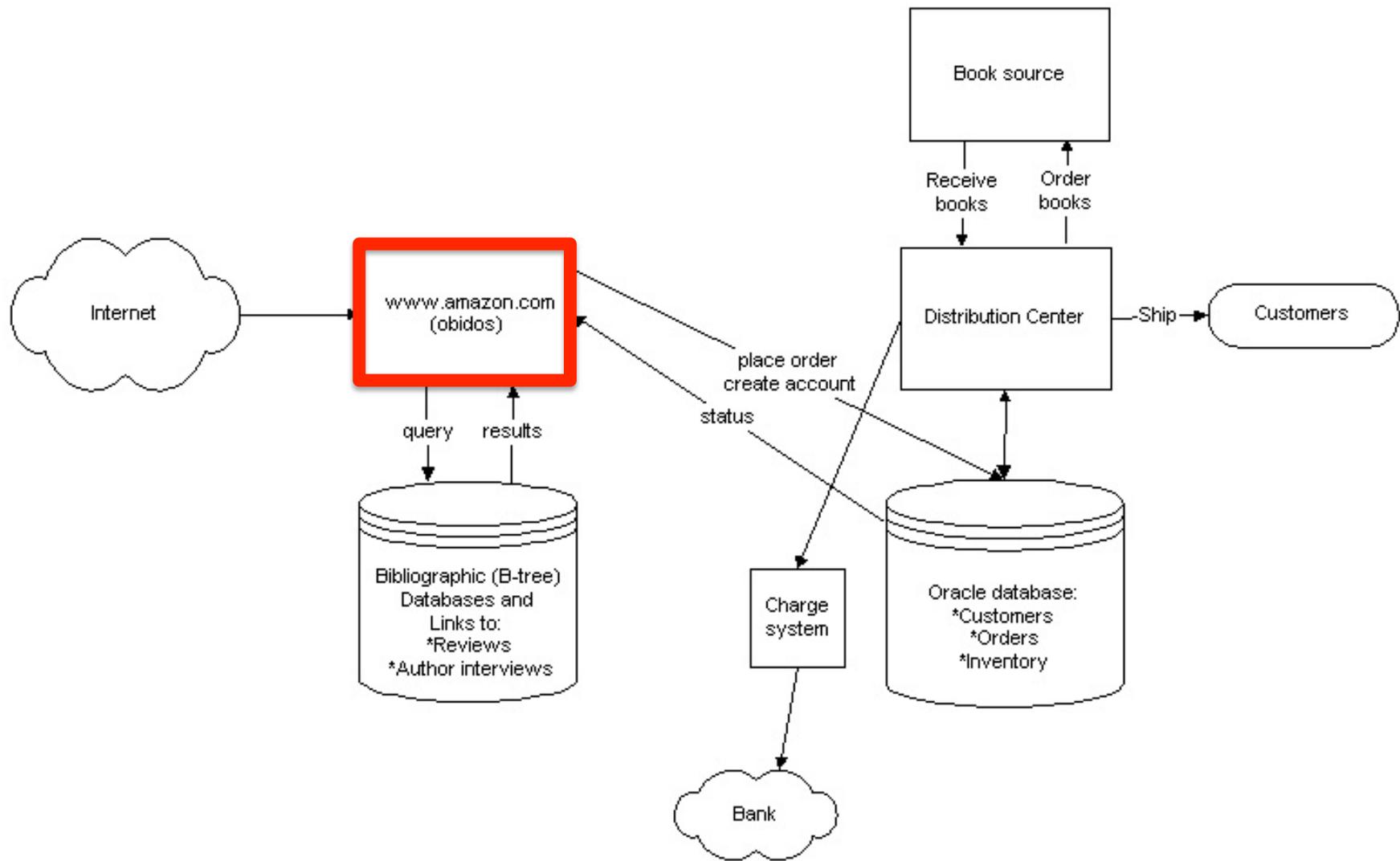
EYES & EDITORS, A PERSONAL NOTIFICATION SERVICE

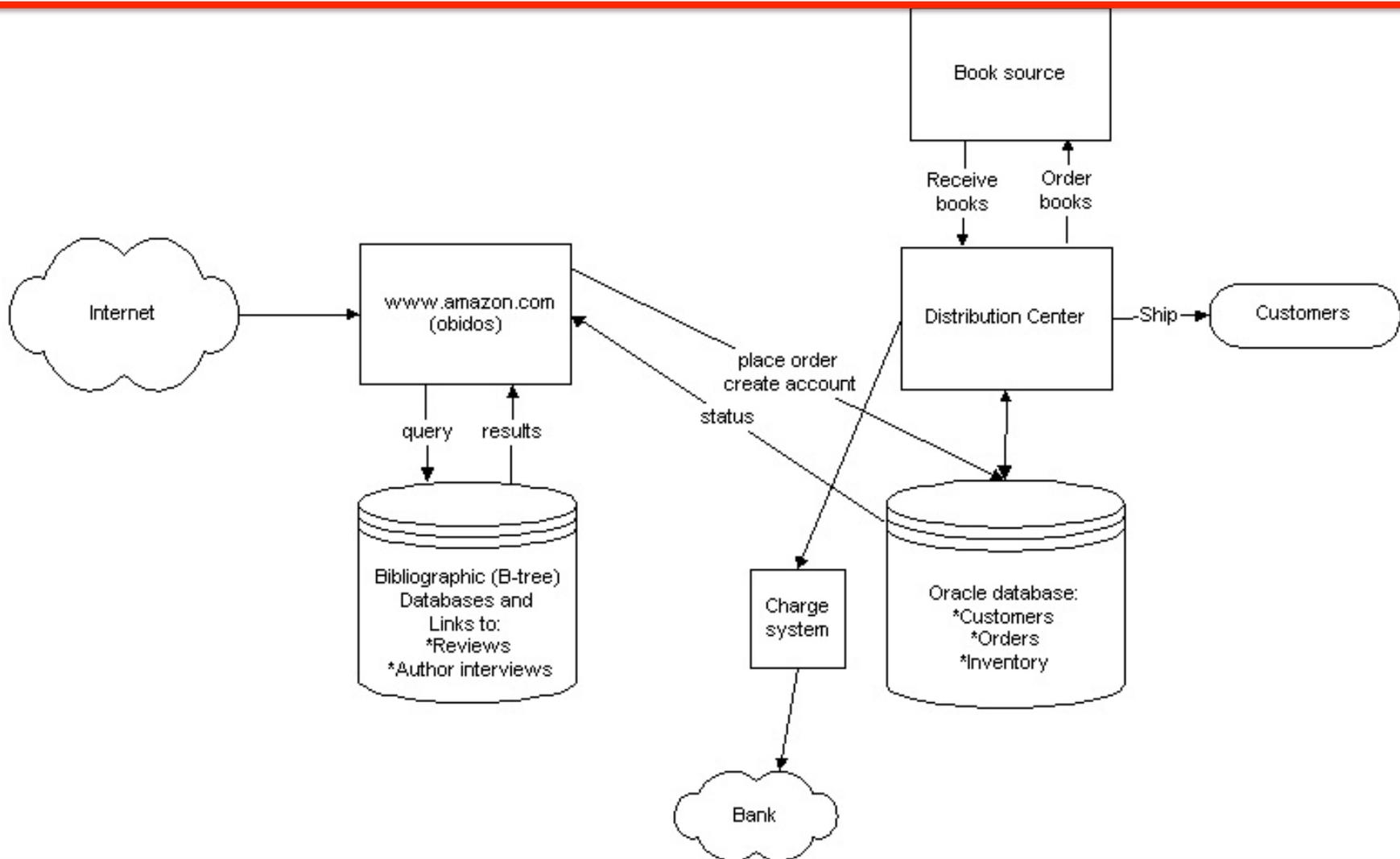
Like to know when that book you want comes out in paperback or when your favorite author releases a new title? Eyes, our tireless, automated search agent, will send you mail. Meanwhile, our human editors are busy previewing galleys and reading advance reviews. They can let you know when especially wonderful works are published in particular genres or subject areas. Come in, [meet Eyes](#), and have it all explained.

YOUR ACCOUNT

Check the status of your orders or change the email address and password you have on file with us. Please note that you **do not** need an account to use the store. The first time you place an order, you will be given the opportunity to create an account.









1997



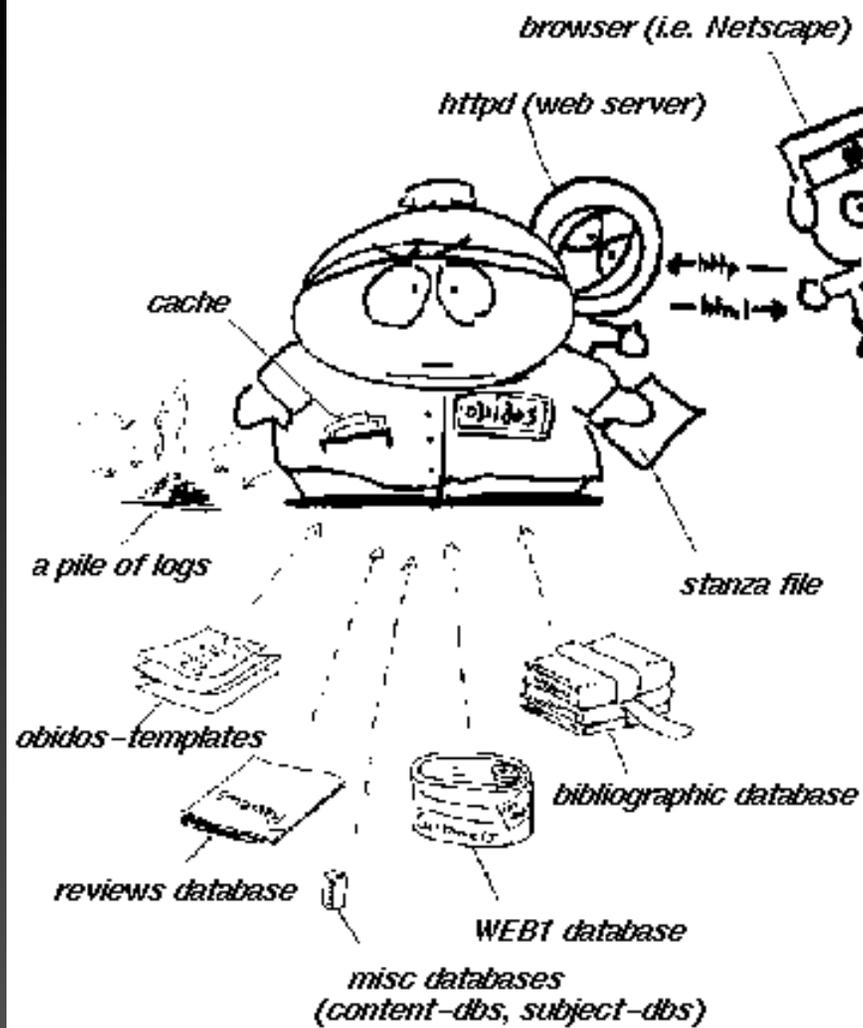
1998

first real data center

1999

Obidos





“The applications that run the business access the database directly and have knowledge of the data model embedded in them. This means that there is a very tight coupling between the applications and the data model, and data model changes have to be accompanied by application changes even if functionality remains the same. This approach does not scale...”

-Amazon.com internal paper, 1998

“The applications that run the business access the database directly and have knowledge of the data model embedded in them. This means that there is a *very tight coupling* between the applications and the data model, and data model changes have to be accompanied by application changes even if functionality remains the same. This approach does not scale...”

-Amazon.com internal paper, 1998

“The applications that run the business access the database directly and have knowledge of the data model embedded in them. This means that there is a very tight coupling between the applications and the data model, and data model changes have to be accompanied by application changes even if functionality remains the same. *This approach does not scale...*”

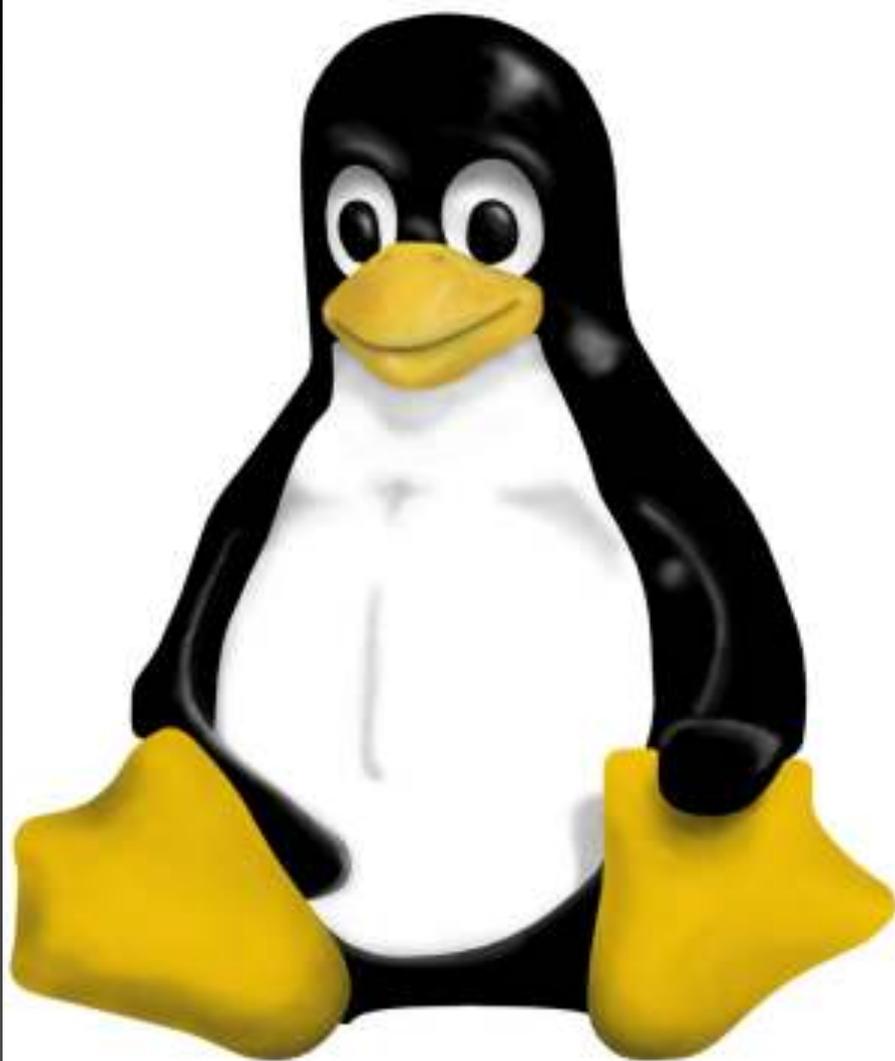
-Amazon.com internal paper, 1998

2000

Distribution Center Isolation

2001

Customer Master Service



```
</prehistory>
```

decouple

service oriented architectures
that scale horizontally
and run on cheap, commodity hardware

distributed computing

increase the speed of execution

develop iteratively





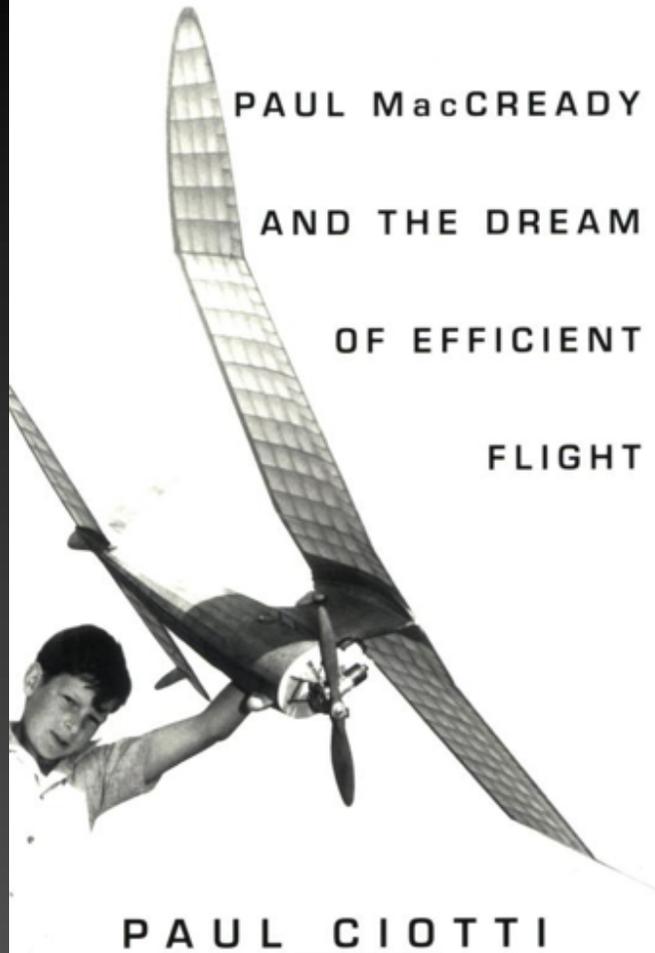
MORE WITH LESS

PAUL MacCREADY

AND THE DREAM

OF EFFICIENT

FLIGHT



PAUL CIOTTI

“They’d get up at four or five in the morning and go flying in the Condor until it got too windy or until they broke the plane, whichever came first. They’d carry the pieces back to the hangar, have breakfast, then work until lunchtime.”

-Paul Ciotti, “More With Less”

“They’d get up at four or five in the morning and go flying in the Condor until it got too windy *or until they broke the plane*, whichever came first. They’d carry the pieces back to the hangar, have breakfast, then work until lunchtime.”

-Paul Ciotti, “More With Less”

idea ratio

The only way to have more good ideas is to have more total ideas.

“Someone would saw off a broom handle to make a splint and tape it to a spar with duct tape, and then minutes later they’d be ready to fly again.”

-Paul Ciotti, “More With Less”

“If something lasted two weeks
before it broke, it was too heavy.”

-Bill Watson

increase the speed of execution

develop iteratively



seek simplicity

“Complicated things break in complicated ways.”

-Kushal Chakrabarti

work backward from the customer



amazon
web services™

March 2006: S3 launches

March 2006: S3 launches

August 2006: EC2 private beta

What could we do with just S3?

Frequently Bought Together



Price For All Three: **\$30.04**

 Add all three to Cart

 Add all three to Wish List

[Show availability and shipping details](#)

- This item:** Spider-Man 3 (Widescreen Edition) ~ Tobey Maguire DVD **\$11.63**
- Spider-Man 2 (Widescreen Special Edition) ~ Tobey Maguire DVD **\$7.75**
- Spider-Man (Widescreen Special Edition) ~ Tobey Maguire DVD **\$10.66**

Theatrical Release Information

US Theatrical Release Date: May 4, 2007

MPAA:  Rated PG-13 for sequences of intense action violence.

Production Company: Columbia Pictures, Marvel Enterprises, Laura Ziskin Productions, Columbia Pictures Industries

USA Box Office: \$337 Million

Worldwide Box Office: \$885 Million

Budget Estimate: \$258 Million

Also Known As: 3 / Spider-Man 3: The IMAX Experience

Filming Locations: 5th Street, Downtown, Los Angeles, California, USA | Alexandria Hotel, Los Angeles, California, USA | Brooklyn, New York City, New York, USA | Bryant Park, Manhattan, New York City, New York, USA | Chicago, Illinois, USA | City Hall Park, Manhattan, New York City, New York, USA | Cleveland, Ohio, USA | Columbia University - Broadway & 116th Street, Manhattan, New York City, New York, USA | Forest Hills, Queens, New York City, New York, USA | Hudson County, New Jersey, USA | Ivar Theatre - 1605 N. Ivar Avenue, Hollywood, Los Angeles, California, USA | Kew Gardens, Queens, New York City, New York, USA | Los Angeles, California, USA | New York City, New York, USA | Queens, New York City, New York, USA

Product Details

Actors: [Tobey Maguire](#), [Kirsten Dunst](#), [James Franco](#), [Thomas Haden Church](#), [Topher Grace](#)

Format: Closed-captioned, Color, Widescreen, NTSC

Language: English, Spanish

Subtitles: Spanish

Region: Region 1 encoding (US and Canada only)

PLEASE NOTE: Some Region 1 DVDs may contain **Regional Coding Enhancement** (RCE). Some, but not all, of our international customers have had



amazon.com
web server

IMDB
service

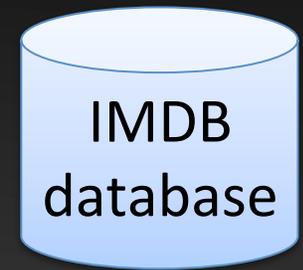
IMDB
database



amazon.com
web server



IMDB
service



IMDB
database



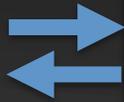
amazon.com
web server



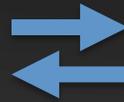
IMDB
service



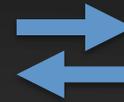
IMDB
database



amazon.com
web server

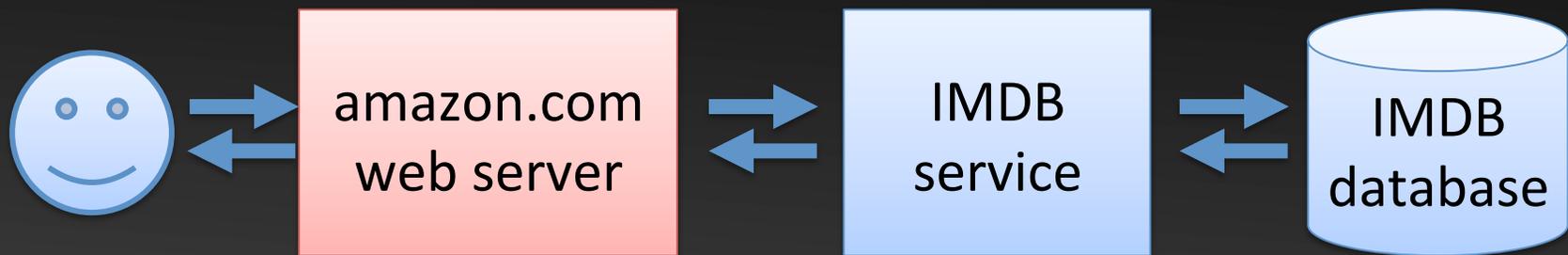


IMDB
service

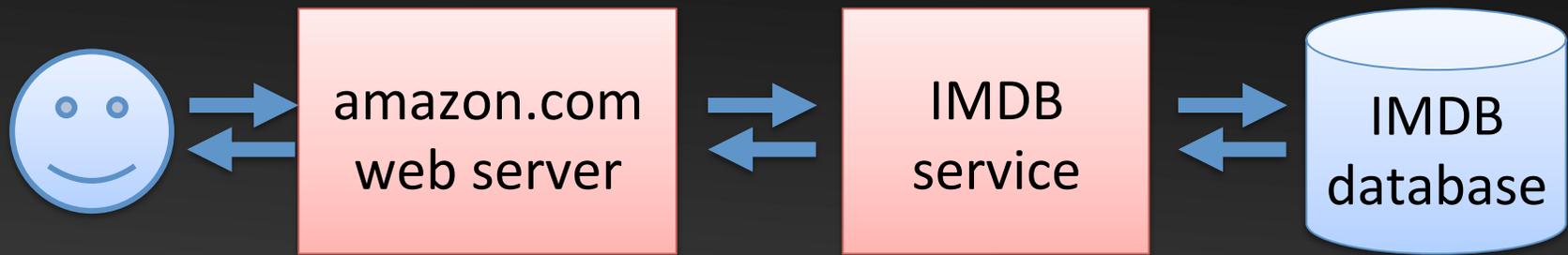


IMDB
database

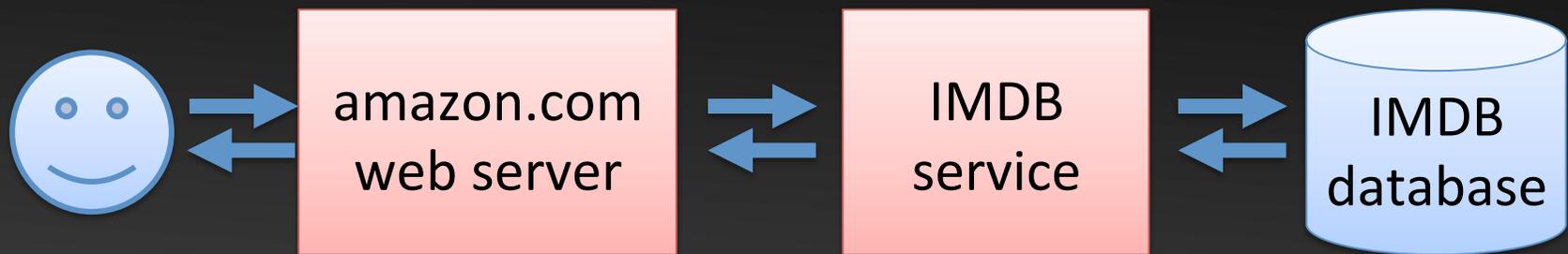
Coupling 1: Code to render IMDB's feature runs in Amazon's web server.



Coupling 2: IMDB must scale in unison with Amazon



Coupling 3: Service interface changes required client updates



Frequently Bought Together



Price For All Three: **\$30.04**

 Add all three to Cart

 Add all three to Wish List

[Show availability and shipping details](#)

- This item:** Spider-Man 3 (Widescreen Edition) ~ Tobey Maguire DVD **\$11.63**
- Spider-Man 2 (Widescreen Special Edition) ~ Tobey Maguire DVD **\$7.75**
- Spider-Man (Widescreen Special Edition) ~ Tobey Maguire DVD **\$10.66**

Theatrical Release Information

US Theatrical Release Date: May 4, 2007

MPAA:  Rated PG-13 for sequences of intense action violence.

Production Company: Columbia Pictures, Marvel Enterprises, Laura Ziskin Productions, Columbia Pictures Industries

USA Box Office: \$337 Million

Worldwide Box Office: \$885 Million

Budget Estimate: \$258 Million

Also Known As: 3 / Spider-Man 3: The IMAX Experience

Filming Locations: 5th Street, Downtown, Los Angeles, California, USA | Alexandria Hotel, Los Angeles, California, USA | Brooklyn, New York City, New York, USA | Bryant Park, Manhattan, New York City, New York, USA | Chicago, Illinois, USA | City Hall Park, Manhattan, New York City, New York, USA | Cleveland, Ohio, USA | Columbia University - Broadway & 116th Street, Manhattan, New York City, New York, USA | Forest Hills, Queens, New York City, New York, USA | Hudson County, New Jersey, USA | Ivar Theatre - 1605 N. Ivar Avenue, Hollywood, Los Angeles, California, USA | Kew Gardens, Queens, New York City, New York, USA | Los Angeles, California, USA | New York City, New York, USA | Queens, New York City, New York, USA

Product Details

Actors: [Tobey Maguire](#), [Kirsten Dunst](#), [James Franco](#), [Thomas Haden Church](#), [Topher Grace](#)

Format: Closed-captioned, Color, Widescreen, NTSC

Language: English, Spanish

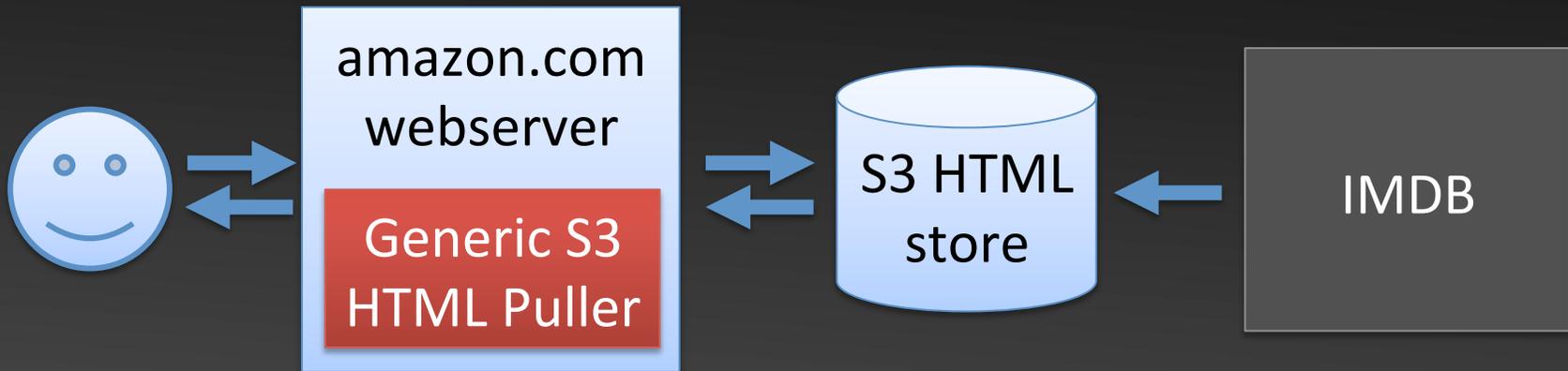
Subtitles: Spanish

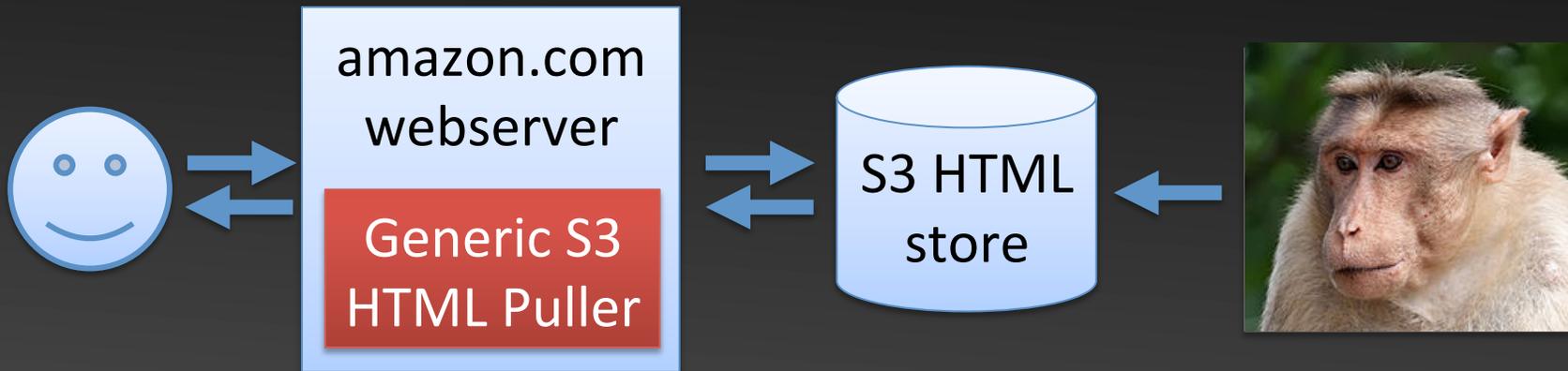
Region: Region 1 encoding (US and Canada only)

PLEASE NOTE: Some Region 1 DVDs may contain **Regional Coding Enhancement** (RCE). Some, but not all, of our international customers have had

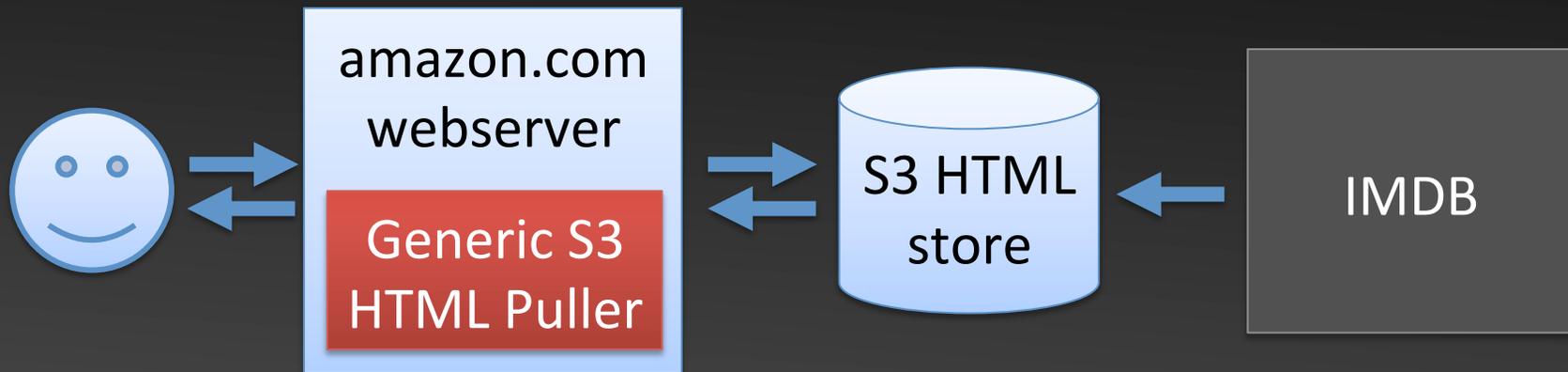
Use S3 as a service, storing pre-rendered HTML



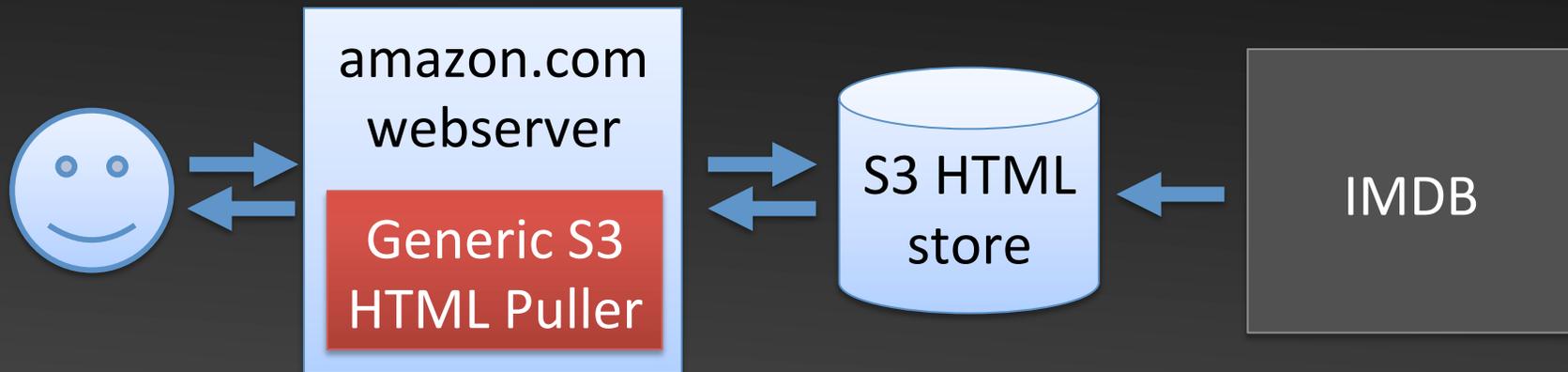




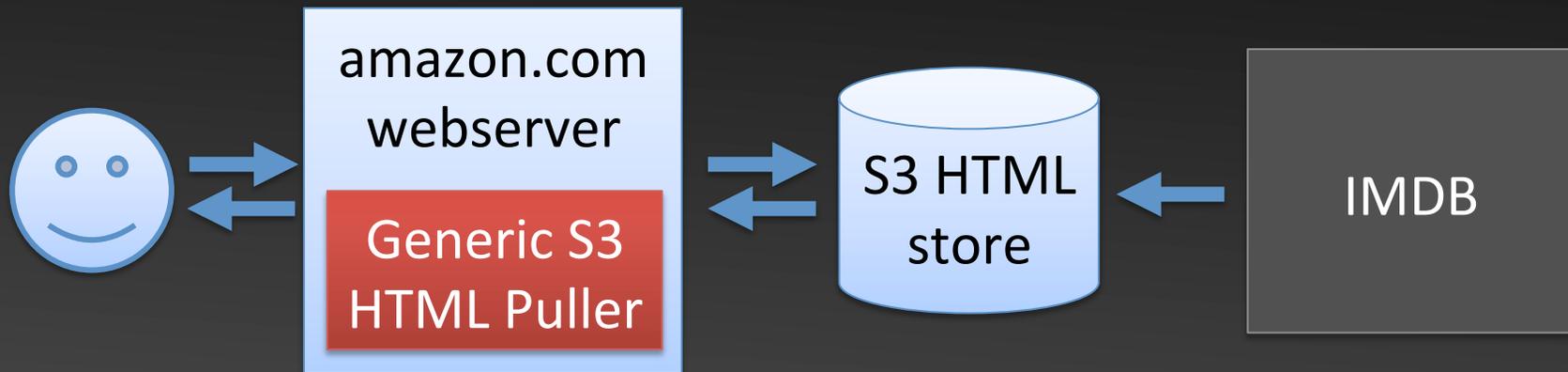
1) Looser coupling between IMDB and Amazon



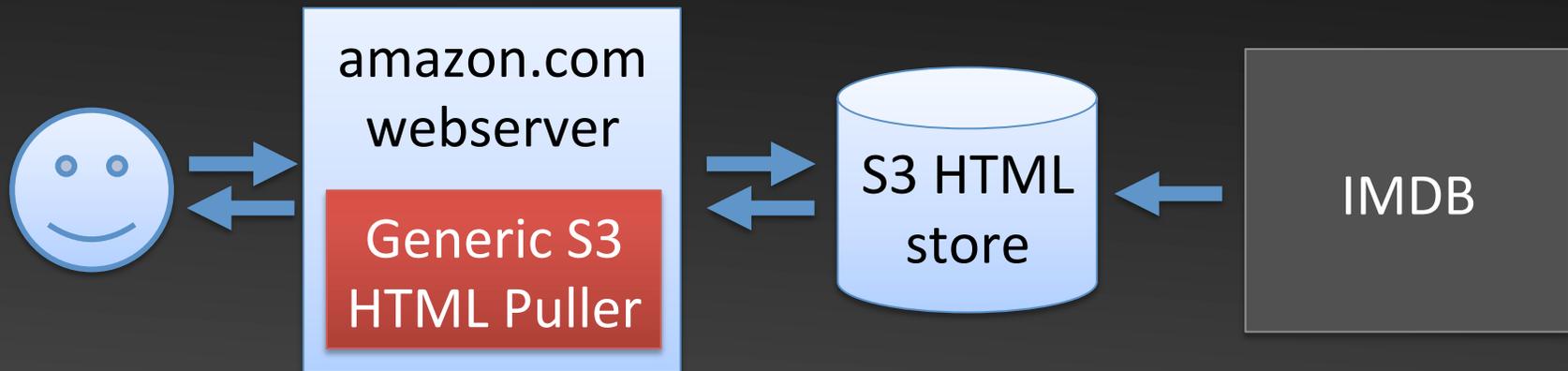
2) S3 handles scaling for IMDB



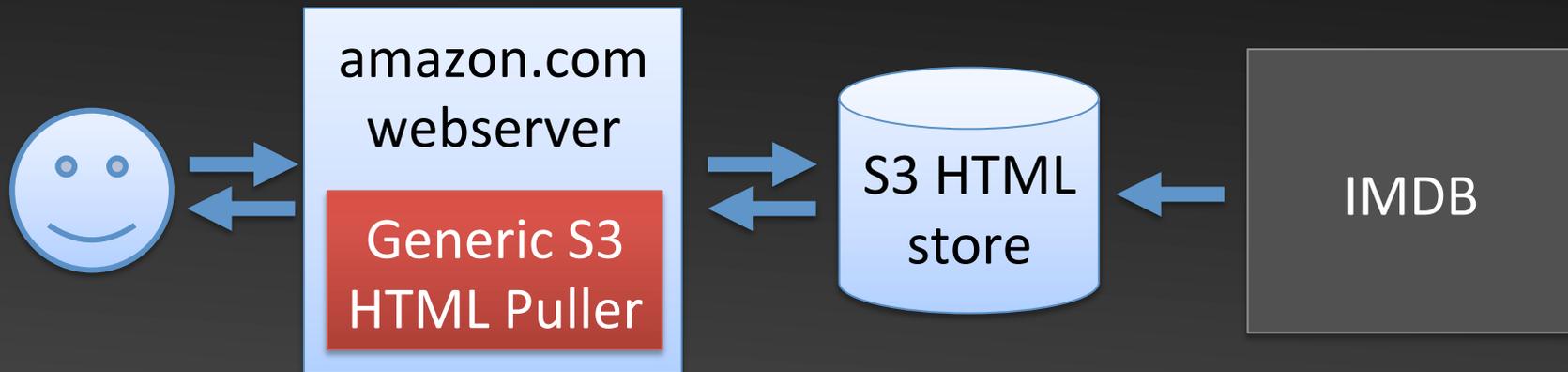
3) Reduced webserver CPU utilization



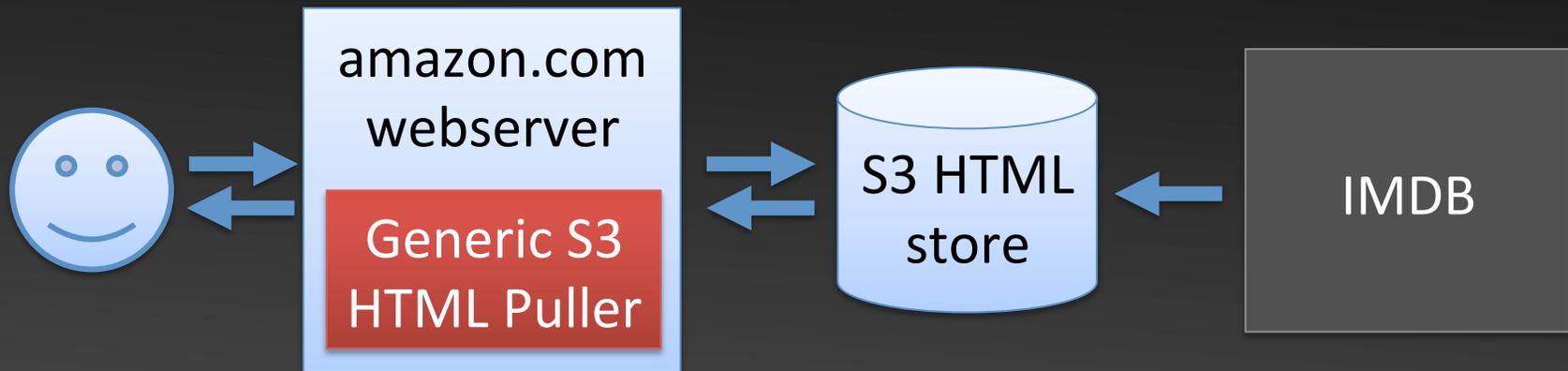
4) Reduced page latency



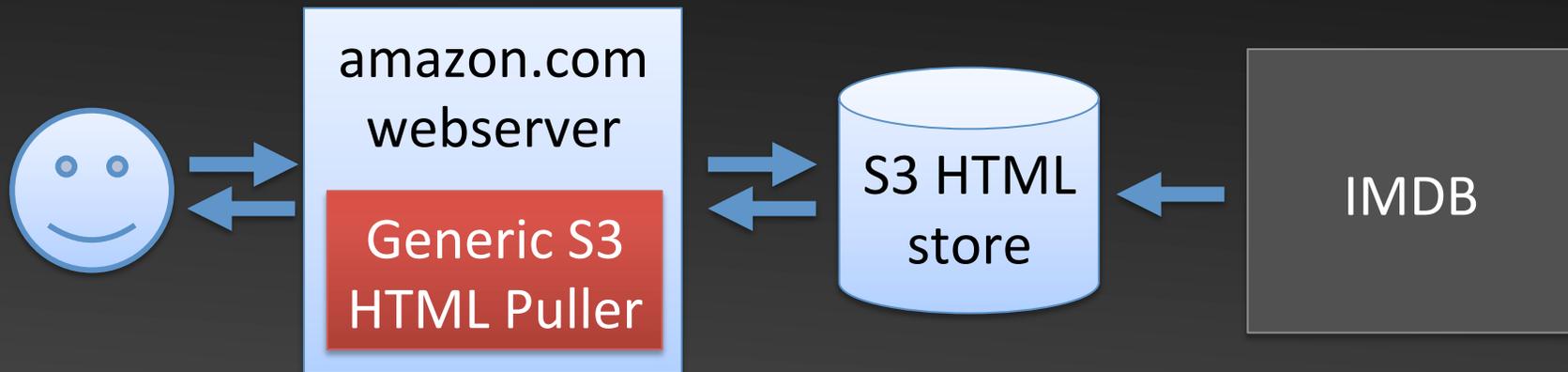
5) Improved availability through reduced dependencies



6) New release model



7) AJAX readiness



What about a more complex use case?

external monitoring

Sample Data > Amazon CA Gateway (Edit | Delete)

www.amazon.ca

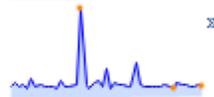
Performance

from in the

First Byte (p50):



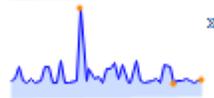
The onLoad Event (p50):



CXA ATF (p50):



Total Load (p50):



Number of runs (count):



Recent Executions (Find more... | Run | 1-click Run)

Start Time	Location	Load Time	Status	Show: <input checked="" type="checkbox"/>	Screenshots <input checked="" type="checkbox"/>	MiniGantt <input checked="" type="checkbox"/>
2011-06-07 09:08:50 GMT	Seattle, USA	2.038s	Successful			
2011-06-07 09:08:49 GMT	Virginia, USA	1.265s	Successful			
2011-06-07 08:39:29 GMT	Virginia, USA	1.325s	Successful			

Your job: Implement a complex system with many moving parts...

Your job: Implement a complex system with many moving parts, **that will run in external data centers...**

Your job: Implement a complex system with many moving parts, that will run in external data centers, **and can scale up quickly.**

Your job: Implement a complex system with many moving parts, that will run in external data centers, and can scale up quickly.

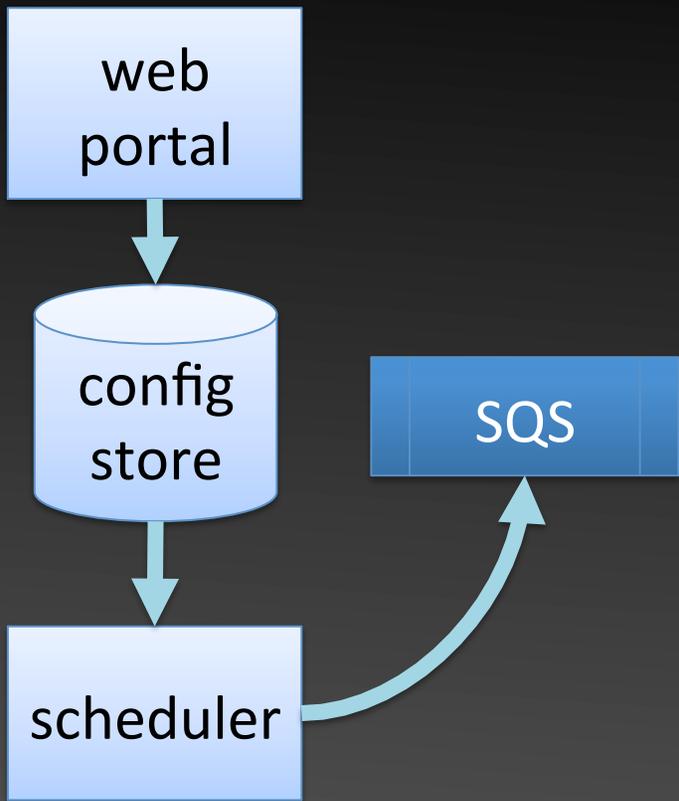
Your team is two people, and you have two months.

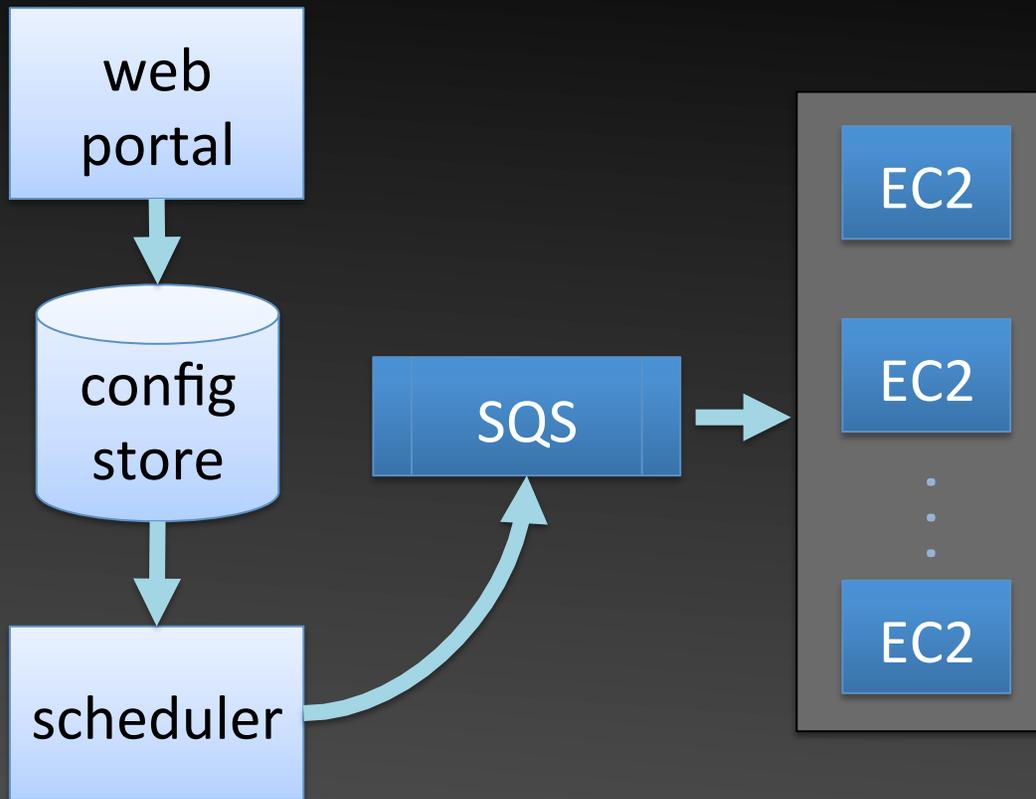
Use as many *AWS* services as possible.

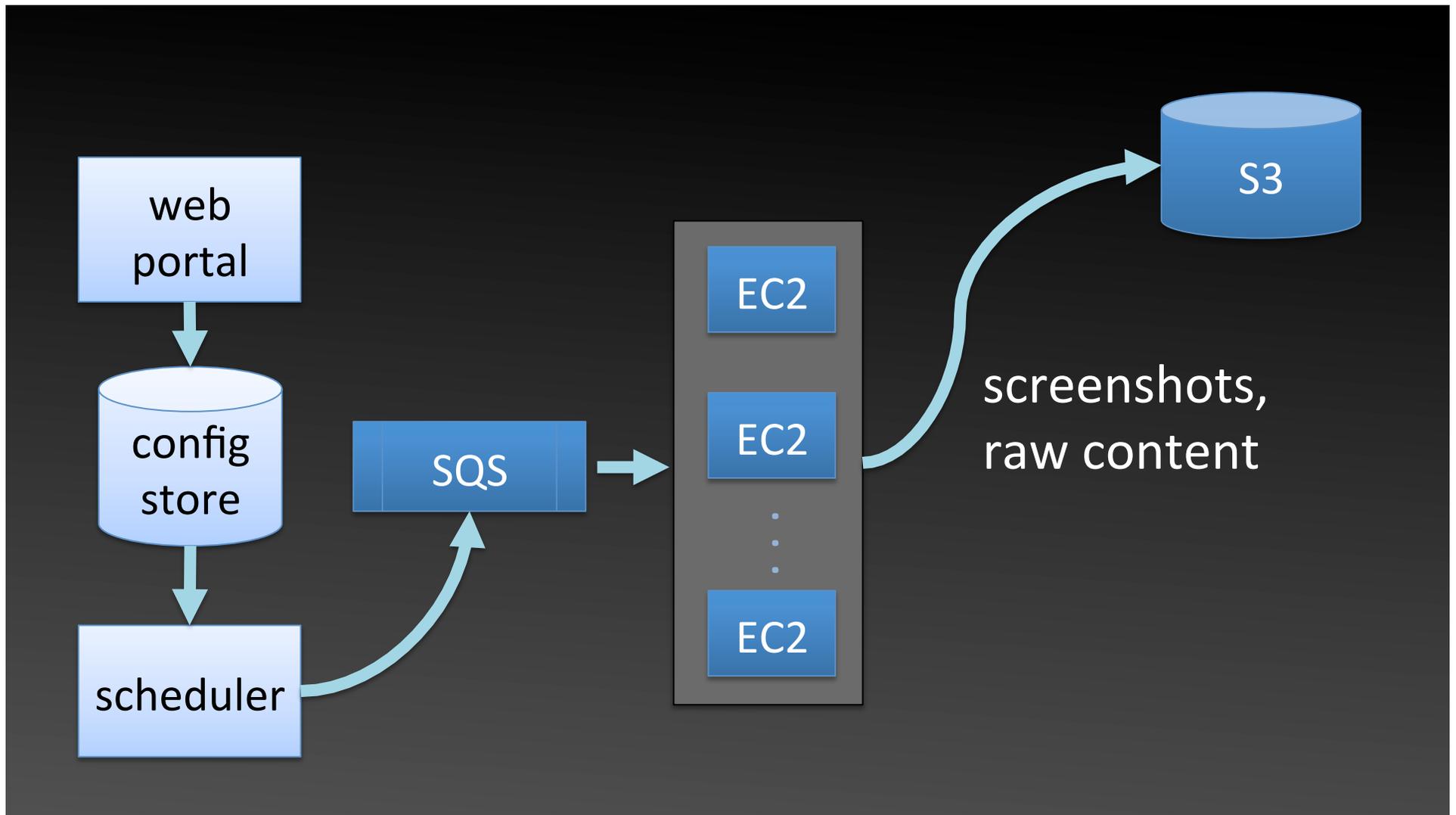
web
portal

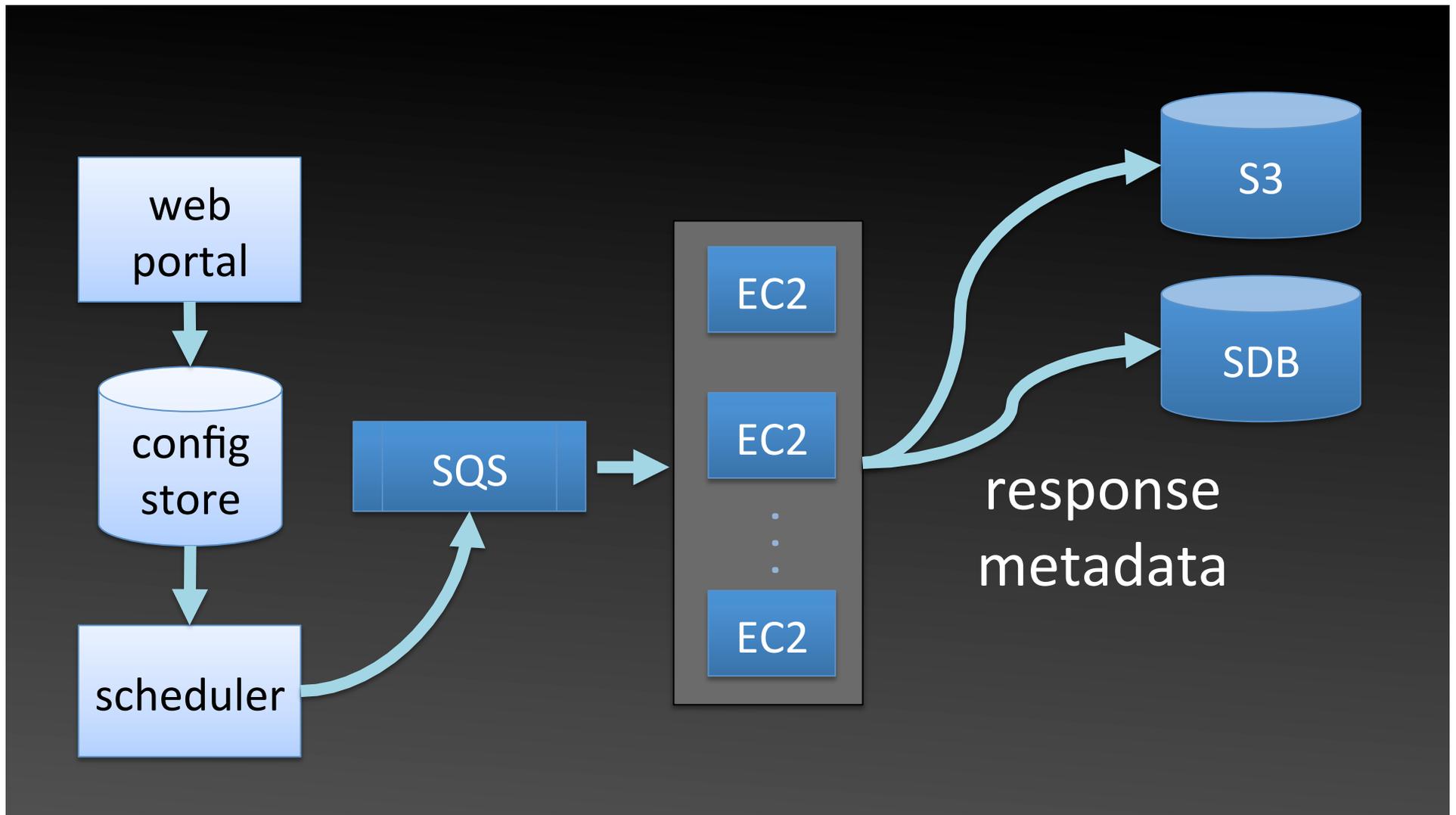


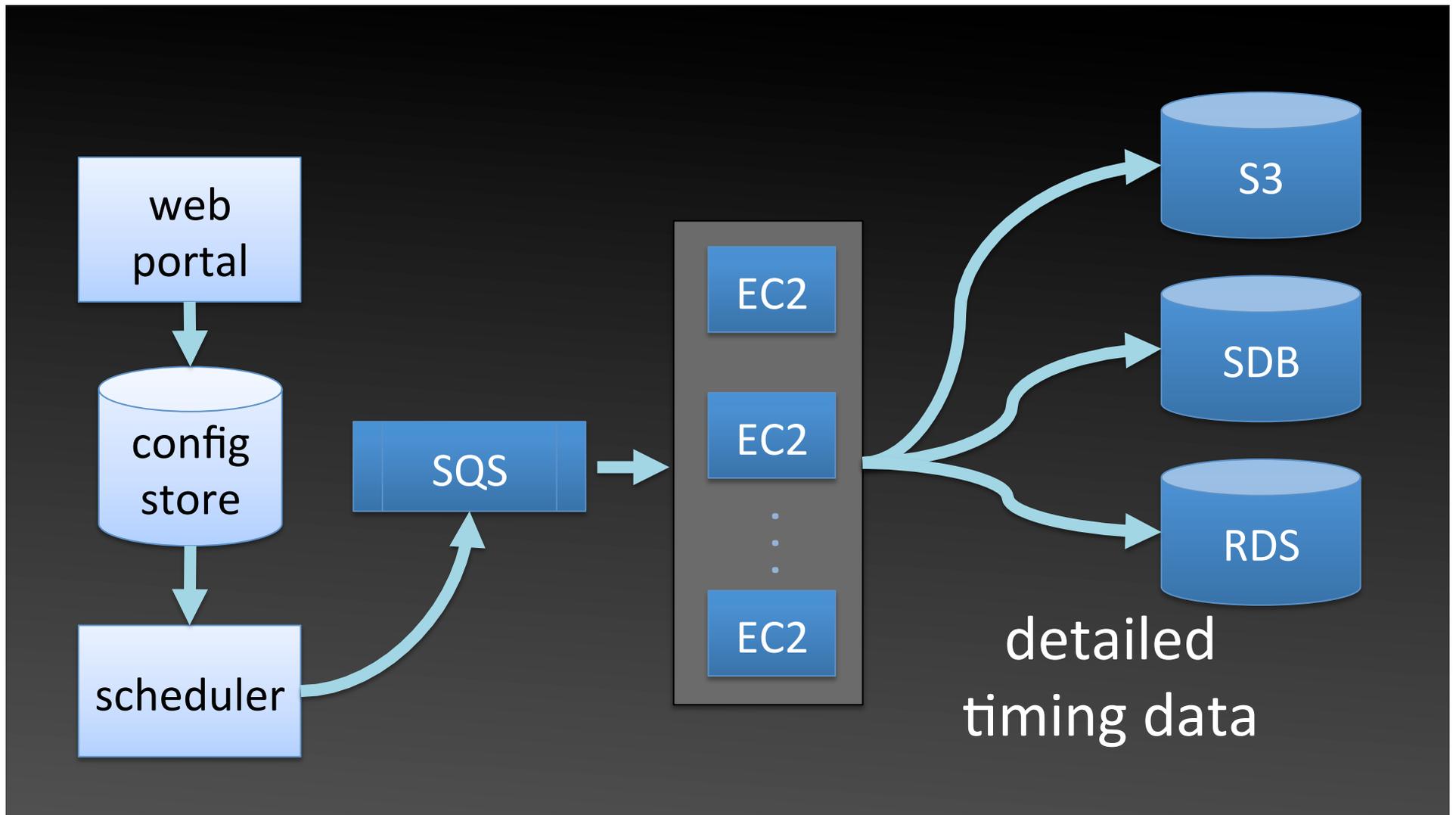
scheduler

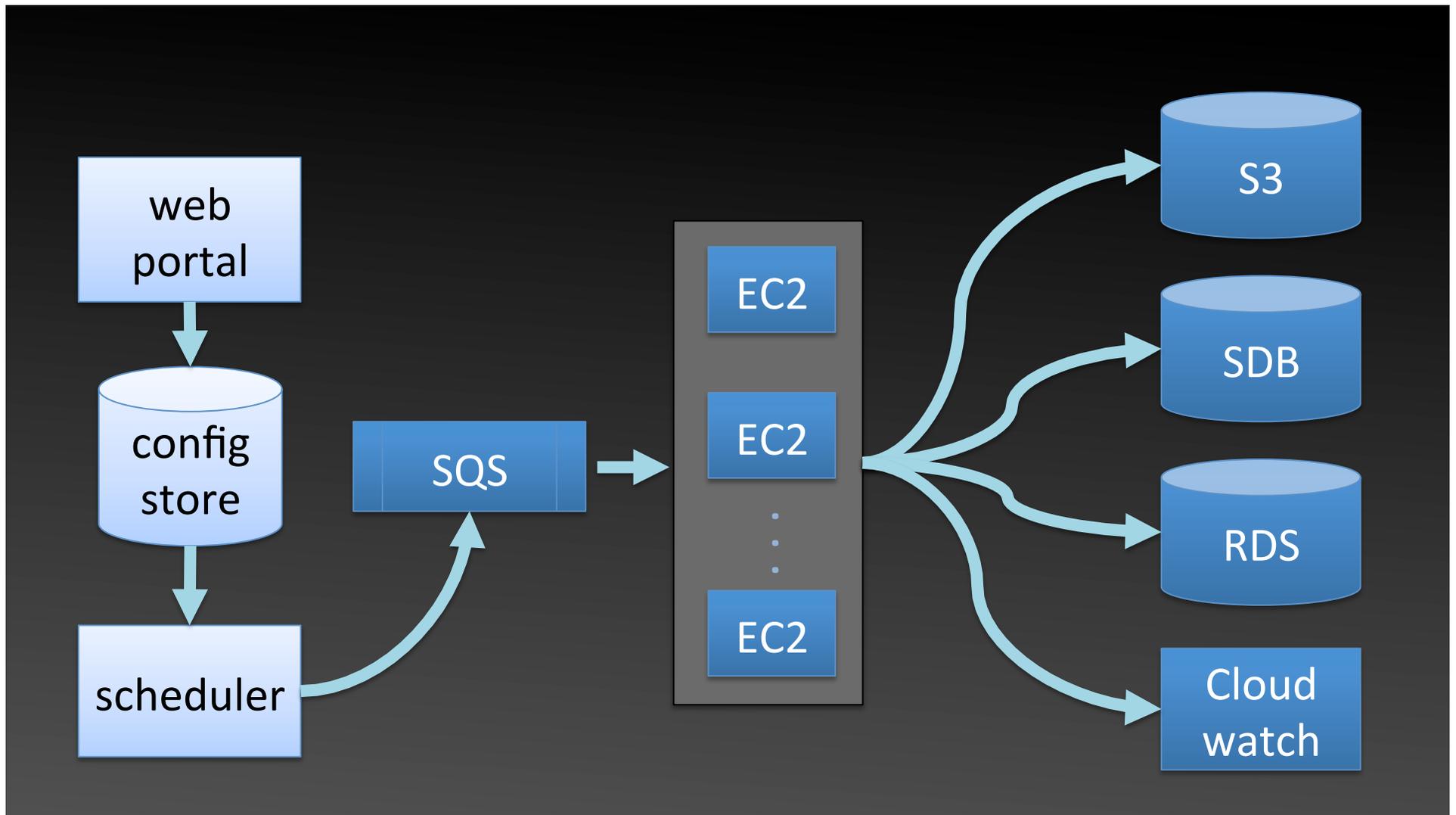


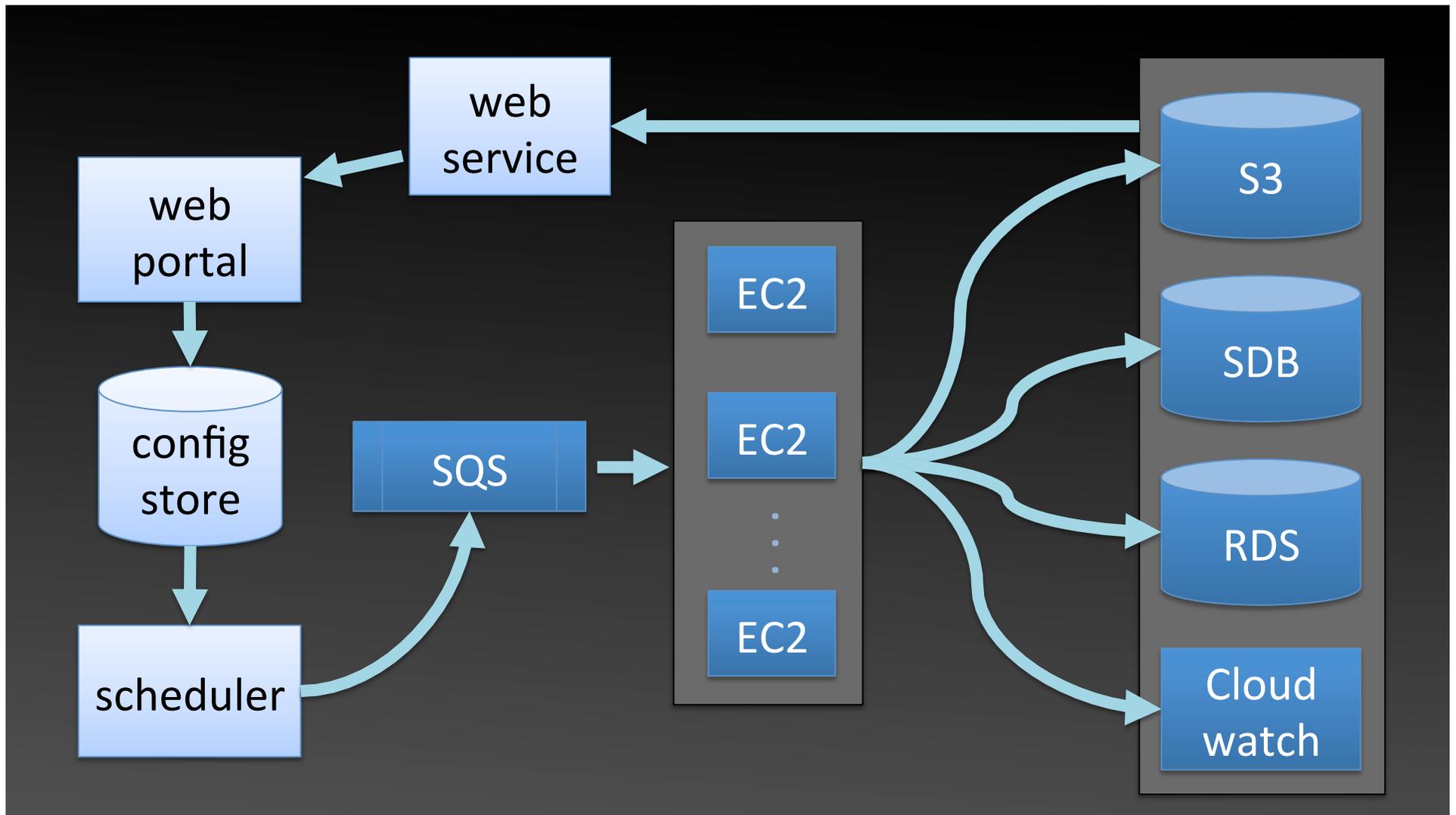












We launched without having to negotiate
any new datacenter co-lo presence.

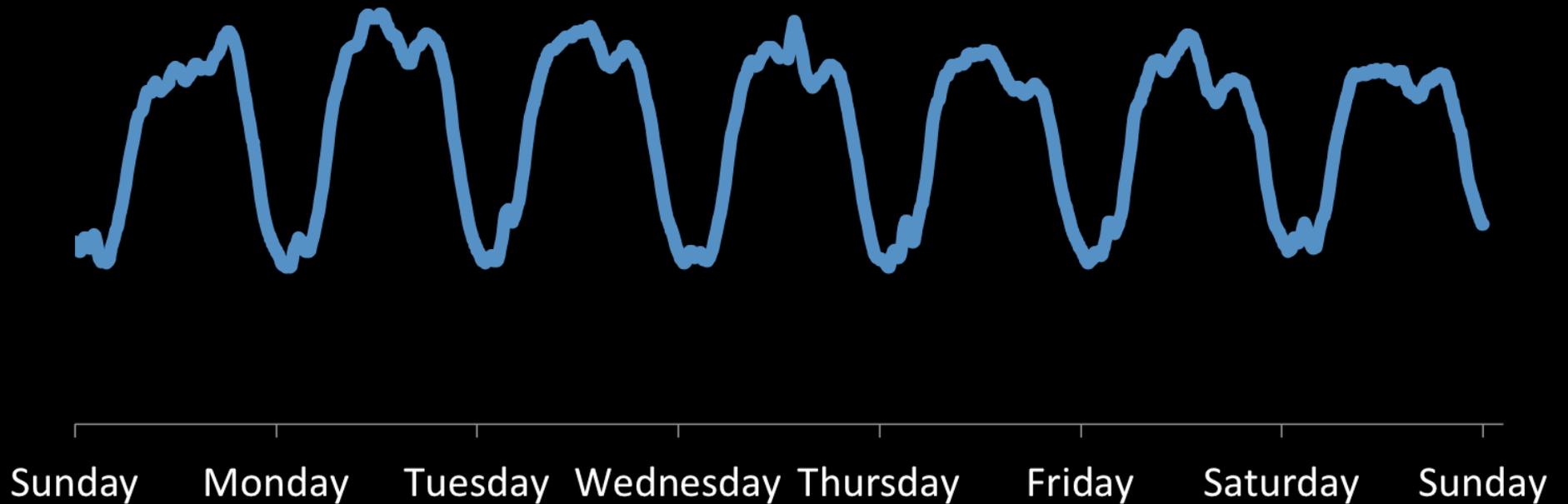
We get true external performance metrics.

We can test features that have
not yet launched.

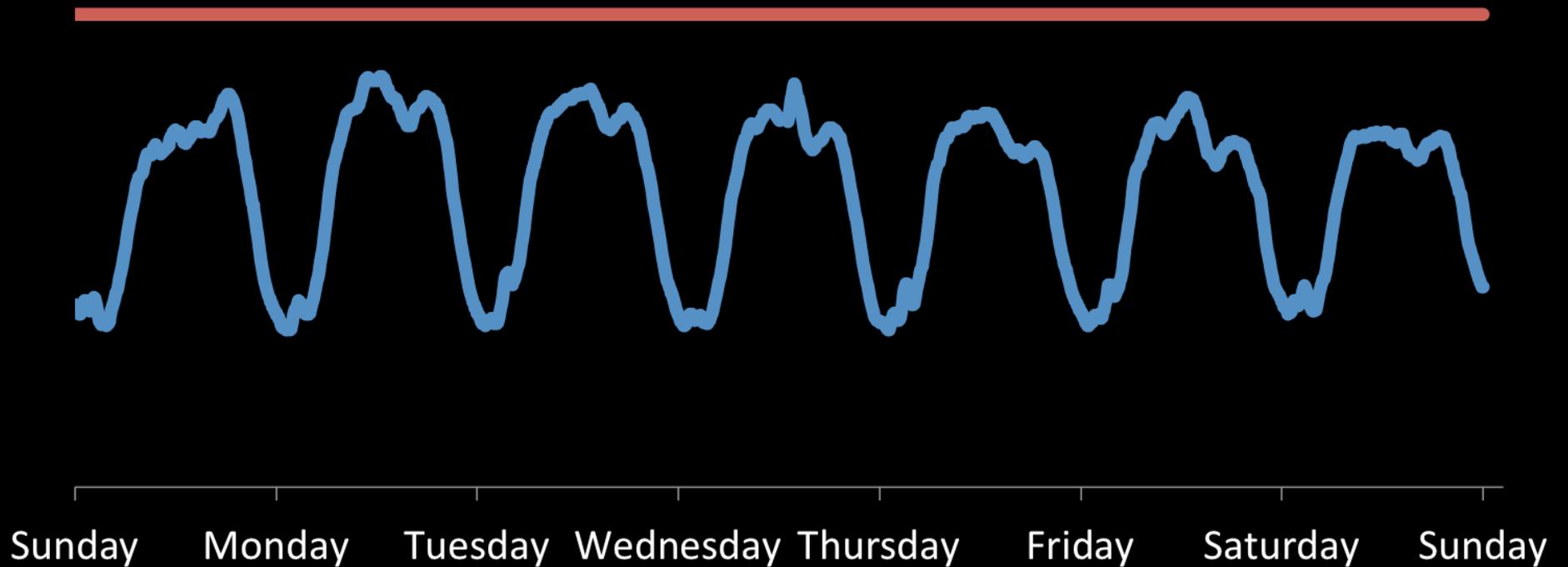
The system scales horizontally to
large amounts of traffic.

What about the amazon.com web servers?

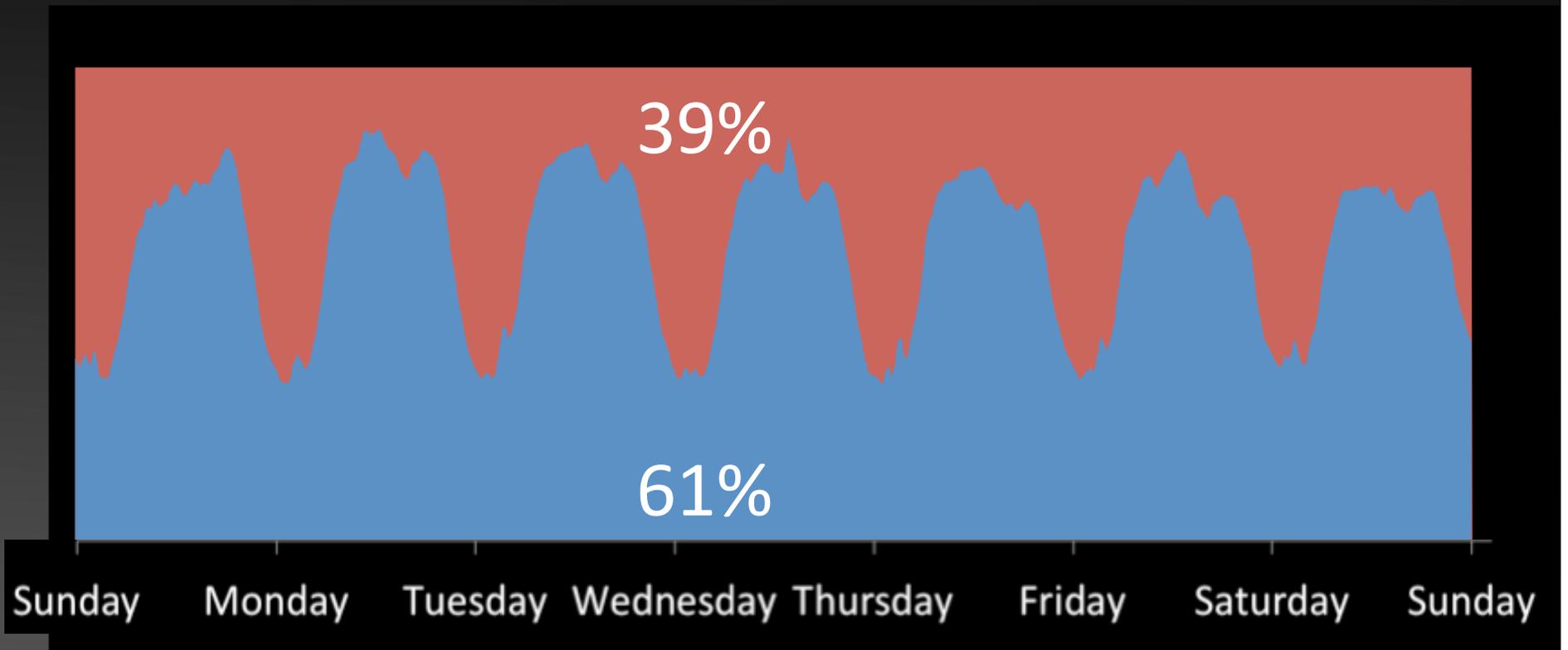
A typical week of traffic at amazon.com



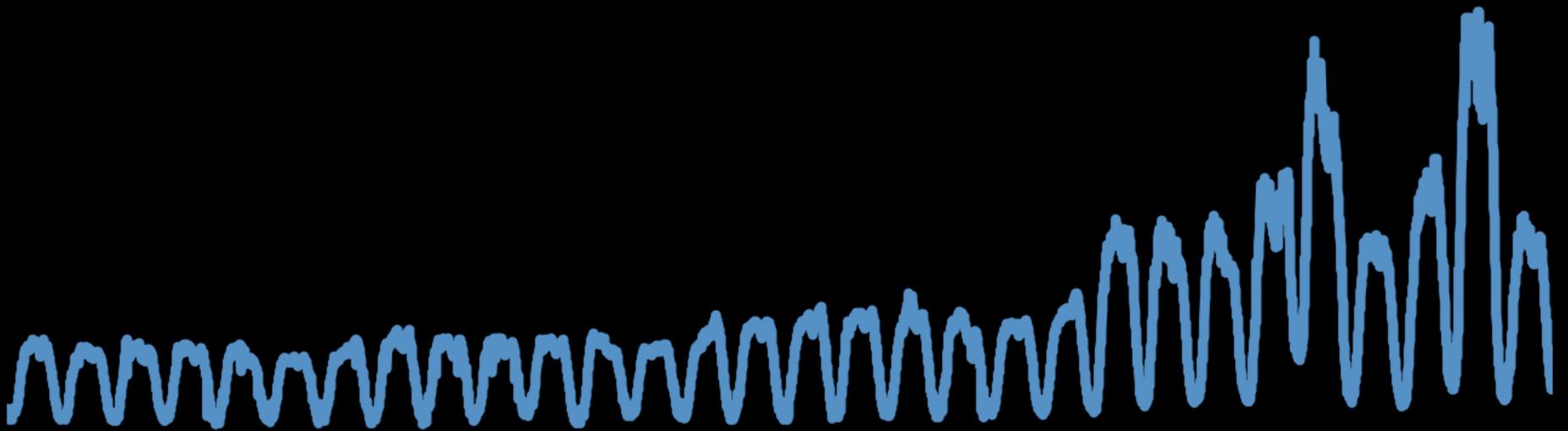
A typical week of traffic at amazon.com



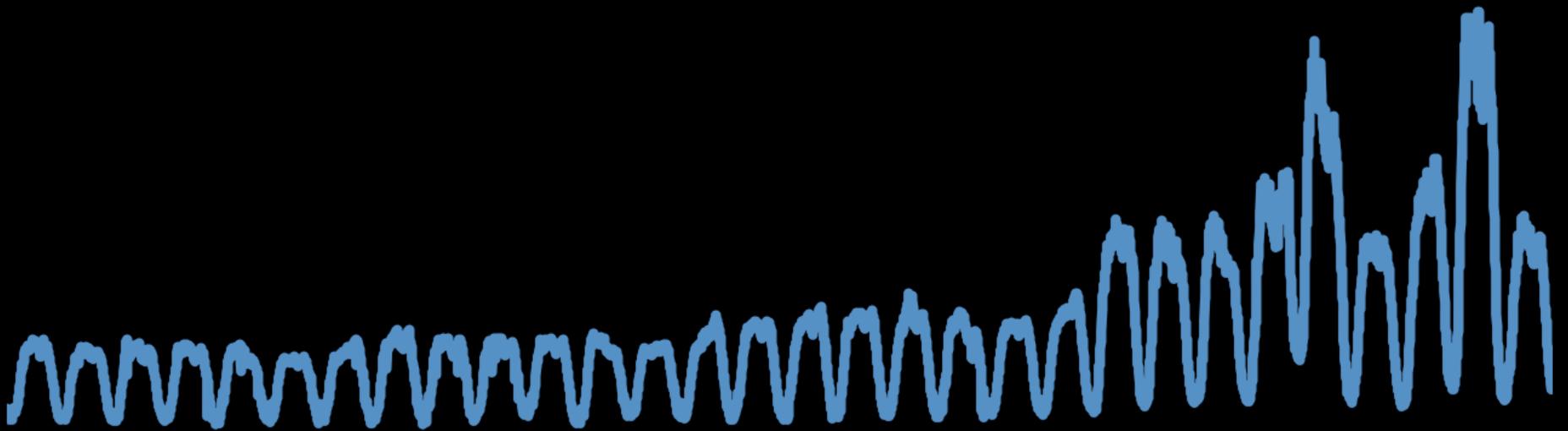
A typical week of traffic at amazon.com



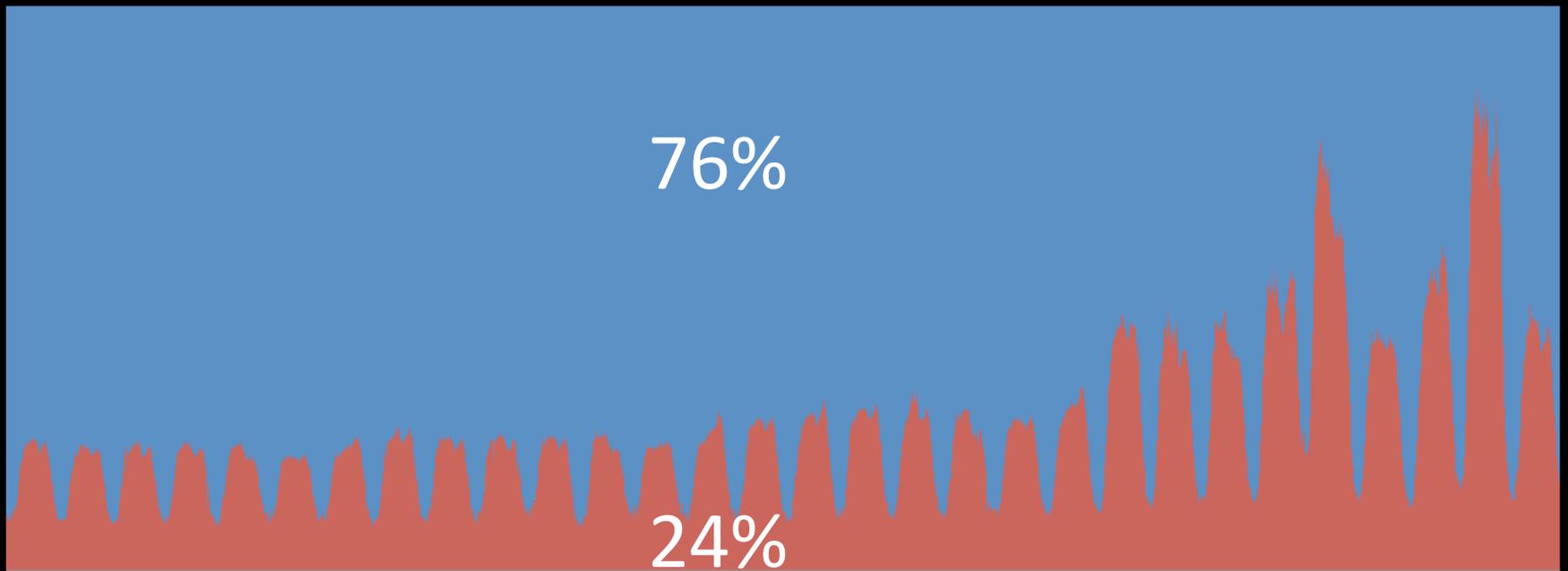
November traffic at amazon.com



November traffic at amazon.com



November traffic at amazon.com



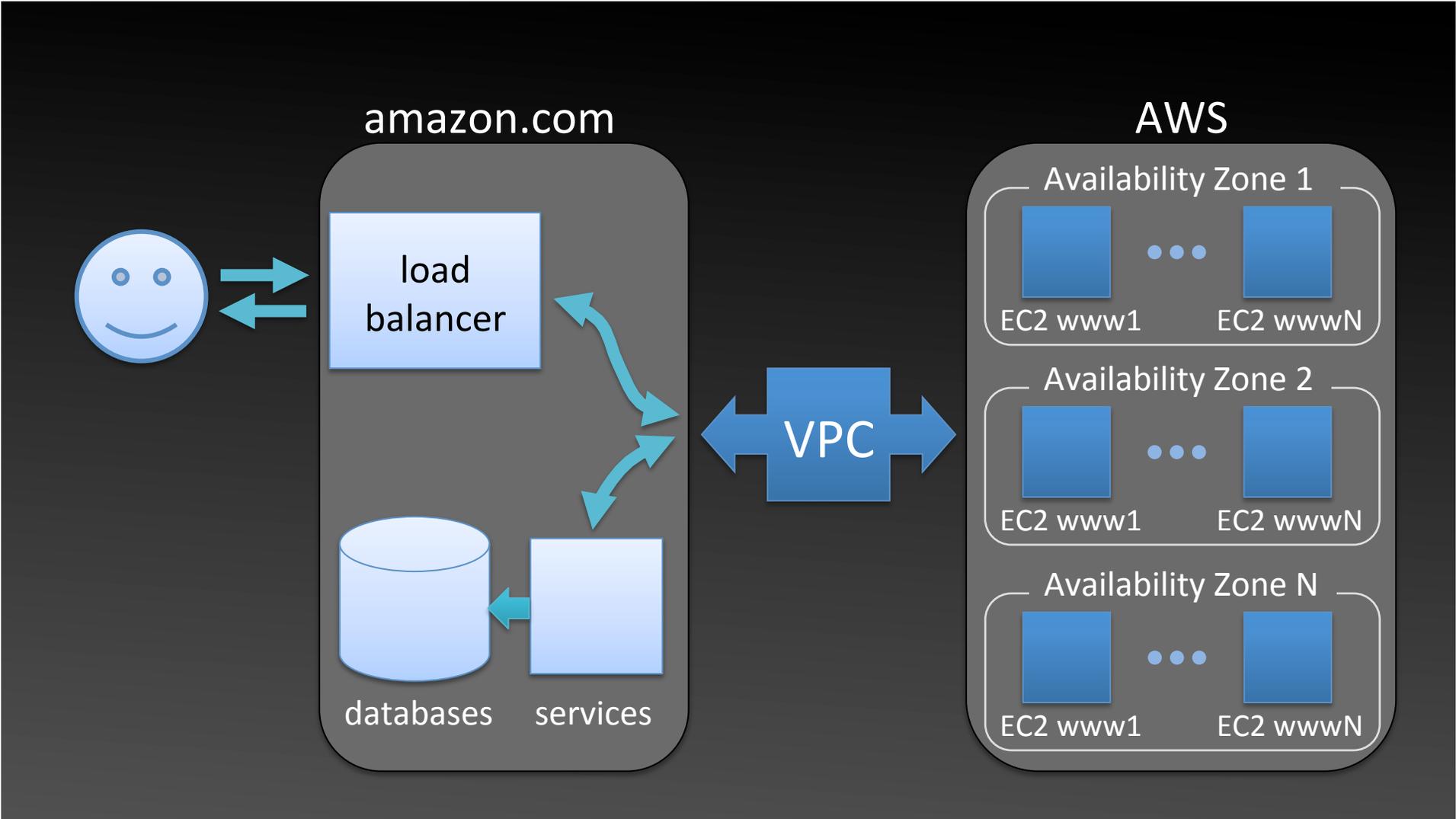
Retail web site hardware is underutilized.

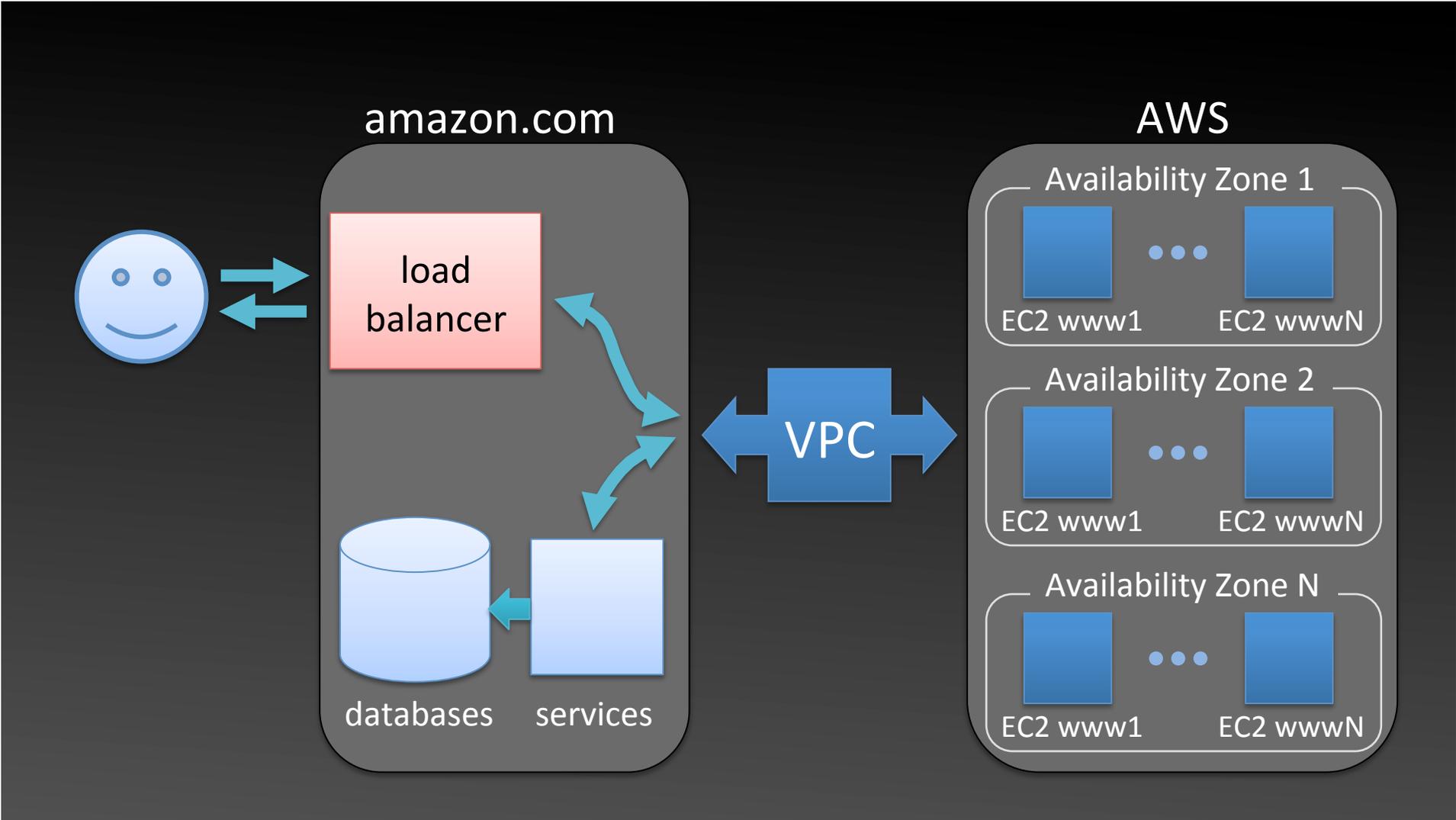
Traffic spikes require heroic effort.

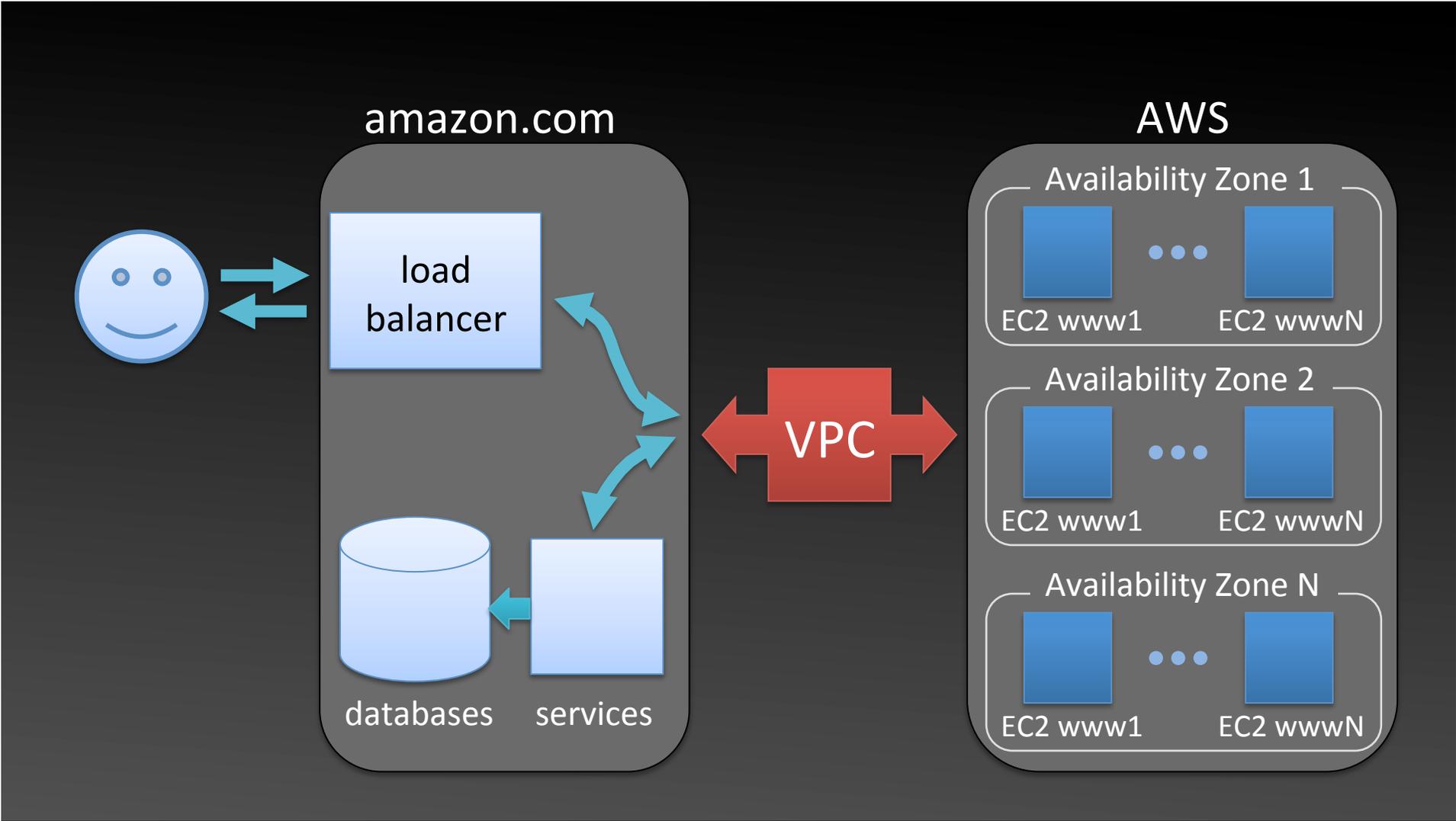
Fleet scaling is discontinuous.

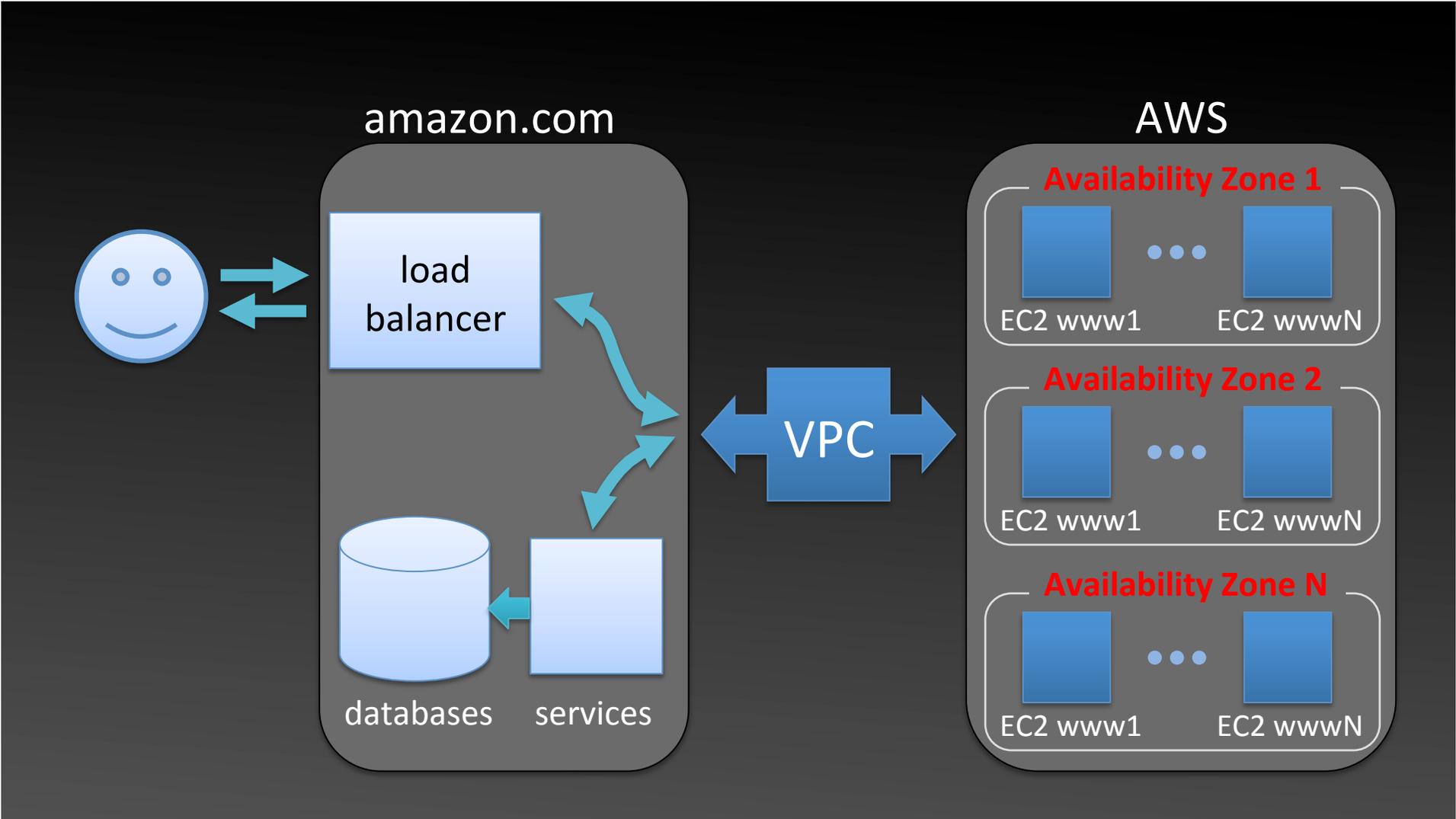
Solution:

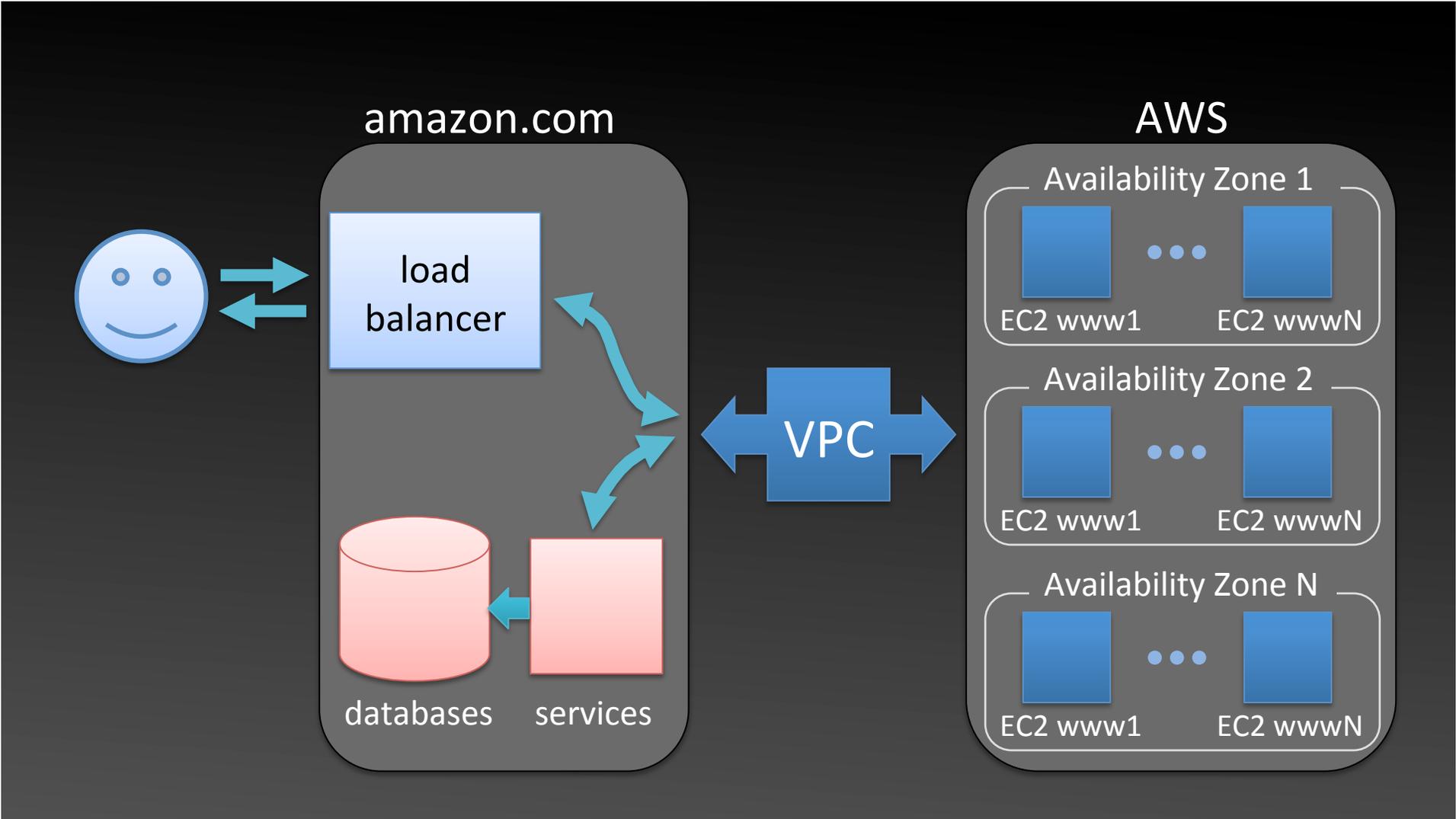
Migrate the entire www.amazon.com
web server fleet to AWS.

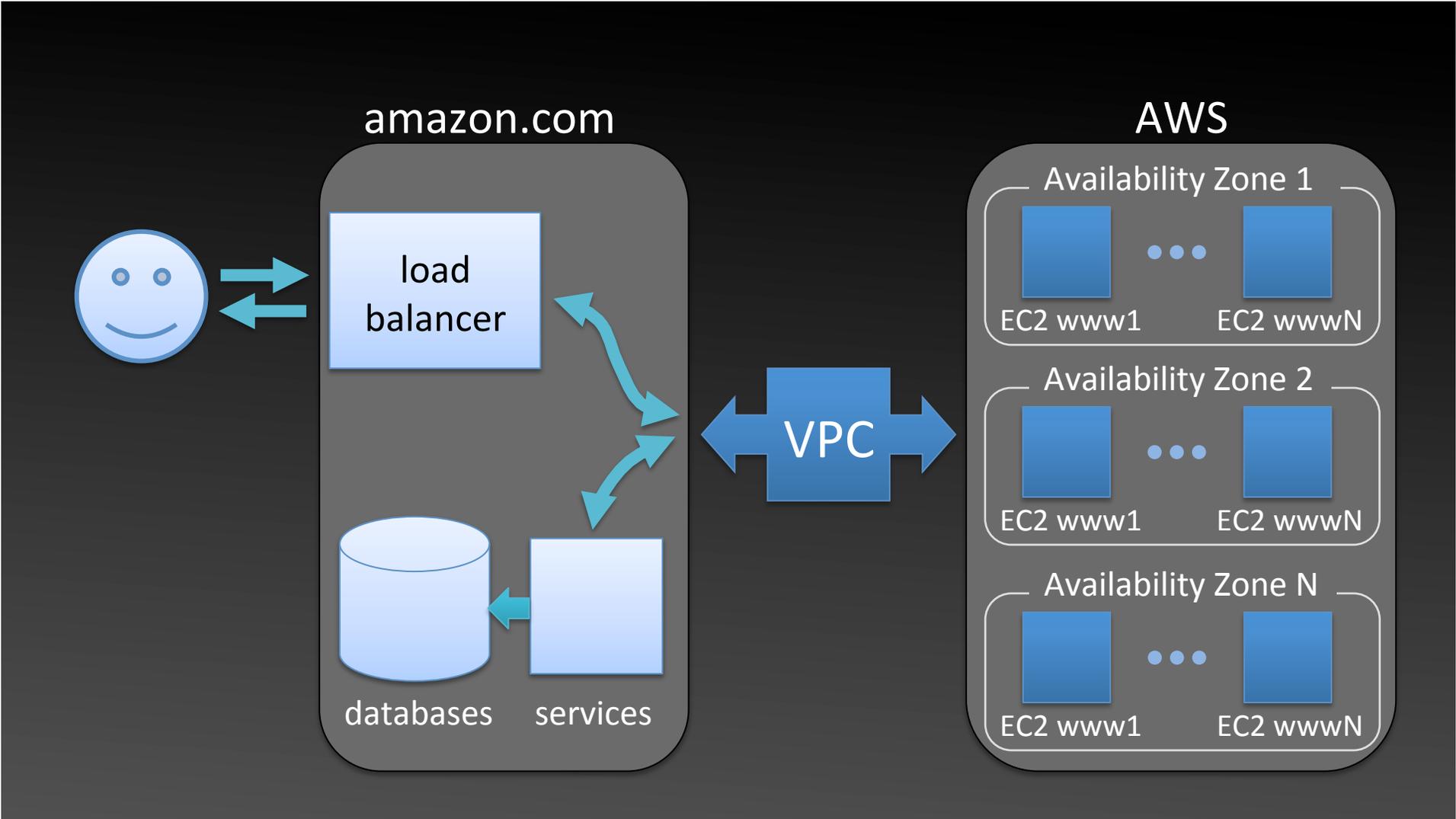












November 10, 2010

We can dynamically scale the fleet in increments as small as a single host.

Traffic spikes require heroic effort.

Traffic spikes ~~require heroic effort~~
can be handled with ease.

What about a database use case?



[Shop All Departments](#)

Search

All Departments

GO



[Your Lists](#)

Your Account > Your Orders

[Orders Listed By Date](#) | [Open Orders](#) | [Digital Orders](#)

Search Your Orders:

Title, Department, Recipient...

Search Orders

Date:

- Select different orders to view-
- Orders placed in the last 30 days
- Orders placed in the past 6 months
- Orders placed in 2011
- Orders placed in 2010
- Orders placed in 2009
- Orders placed in 2008
- Orders placed in 2007
- Orders placed in 2006
- Orders placed in 2005
- Orders placed in 2004
- Orders placed in 2003
- Orders placed in 2002
- Orders placed in 2001
- Orders placed in 2000
- Orders placed in 1999
- Orders placed in 1998

Go

2 orders placed in the last 30 days

Order Placed:

November 8, 2011

[View Order Details](#) | [View Invoice](#)

Order Number: [102-8710903-6493050](#)

Recipient: John Rauser

Order Total: **\$42.31**

SHIPMENT 1 OF 2

Delivery Estimate: November 14, 2011

NOT YET SHIPPED

SHIPPING SOON

SHIPPED



Huggies Diapers Overnite, Step 5
Pack, 52 Count

Sold by: Amazon.com LLC

Available Actions



Shop All Departments

Search

GO



Your Lists

Your Account > Your Orders

[Orders Listed By Date](#) | [Open Orders](#) | [Digital Orders](#)

Search Your Orders:

Search Orders

Date:

- Select different orders to view-
- Orders placed in the last 30 days
- Orders placed in the past 6 months
- Orders placed in 2011
- Orders placed in 2010
- Orders placed in 2009
- Orders placed in 2008
- Orders placed in 2007
- Orders placed in 2006
- Orders placed in 2005
- Orders placed in 2004
- Orders placed in 2003
- Orders placed in 2002
- Orders placed in 2001
- Orders placed in 2000
- Orders placed in 1999
- Orders placed in 1998

Go

2 orders placed in the last 30 days

Order Placed:

November 8, 2011

[View Order Details](#) | [View Invoice](#)

Order Number: [102-8710903-6493050](#)

Recipient: John Rauser

Order Total: **\$42.31**

SHIPMENT 1 OF 2

Delivery Estimate: **November 14, 2011**

NOT

SHIPPING SOON

SHIPPED



Huggies Diapers Overnite, Step 5
Pack, 52 Count

Sold by: Amazon.com LLC

Available Actions





Shop All Departments

Search

GO



Your Lists

Your Account > Your Orders

[Orders Listed By Date](#) | [Open Orders](#) | [Digital Orders](#)

Search Your Orders:

Search Orders

Date:

-Select different orders to view-

- Orders placed in the last 30 days
- Orders placed in the past 6 months
- Orders placed in 2011
- Orders placed in 2010
- Orders placed in 2009
- Orders placed in 2008
- Orders placed in 2007
- Orders placed in 2006
- Orders placed in 2005
- Orders placed in 2004
- Orders placed in 2003
- Orders placed in 2002
- Orders placed in 2001
- Orders placed in 2000
- Orders placed in 1999
- Orders placed in 1998

Go

2 orders placed in the last 30 days

Order Placed:

November 8, 2011

[View Order Details](#) | [View Invoice](#)

Order Number: [102-8710903-6493050](#)

Recipient: John Rauser

Order Total: **\$42.31**

SHIPMENT 1 OF 2

Delivery Estimate: November 14, 2011

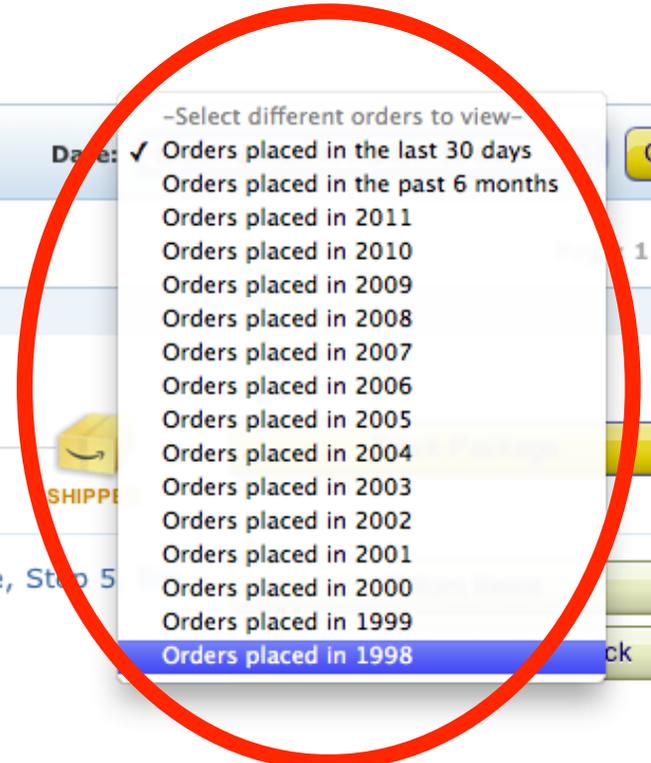
NOT YET SHIPPED SHIPPING SOON



Huggies Diapers Overnite, Step 5
Pack, 52 Count

Sold by: Amazon.com LLC

Available Actions



transactions per second

The *cumulative* amount of data stored predicts long-term hardware spend more accurately than transactions per second.



web
server



orders
service



orders database

web
server



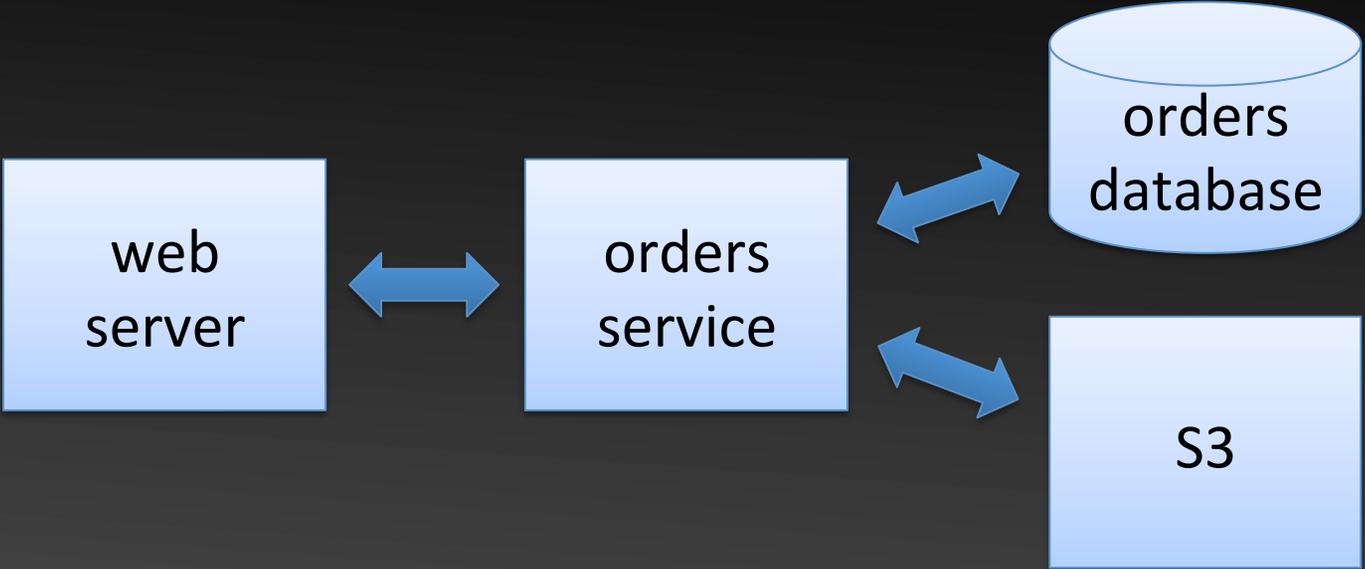
orders
service

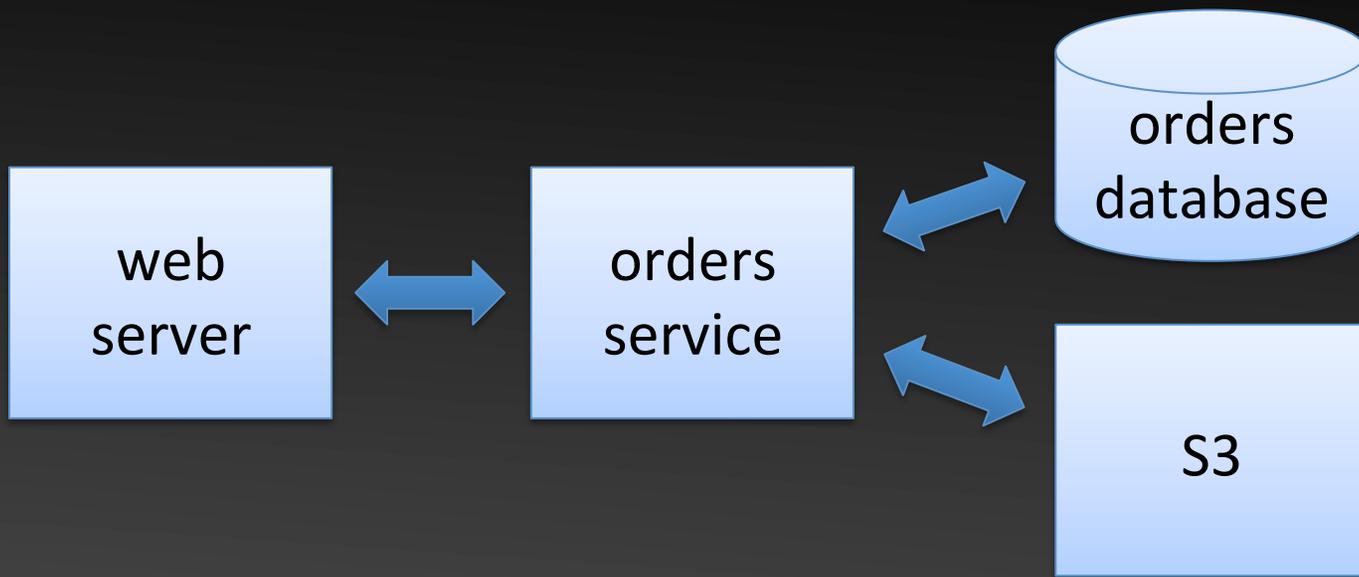


Database infrastructure is expensive.

Two kinds of orders:

- 1) Current, highly dynamic,
frequently accessed orders
- 2) Historical, immutable,
rarely accessed orders





670 million orders (4Tb)

We're spending much less on database hosts.

This sets us up for migration to RDS/SDB.

What about deployment models?

Continuous Deployment



<http://timothyfitz.wordpress.com>

Timothy Fitz

Continuous Deployment

Posted in **Uncategorized** by timothyfitz on February 8, 2009

Alex has just written a refactoring of some website backend code. Since it was a small task, it's committed and Alex moves on to the next feature.

When the code is deployed in production two weeks later it causes the entire site to go down. A one-character typo which was missed by automated tests caused a failure cascade reminiscent of the bad-old-days at twitter. It takes eight hours of downtime to isolate the problem, produce a one character fix, deploy it and bring production back up.

Alex curses luck, blames human infallibility, inevitable cost of software engineering and moves on to the next task.

This story is the day-to-day of most startups I know. It sucks. Alex has a problem and she doesn't even know it. Her development practices are unsustainable. "Stupid mistakes" like the one she made increase as the product grows more complex and as the team gets larger. Alex needs to switch to a scalable solution.

Before I get to the solution, let me tell you about some common non-solutions. While these are solutions to real problems, they aren't the solution to Alex's problem.

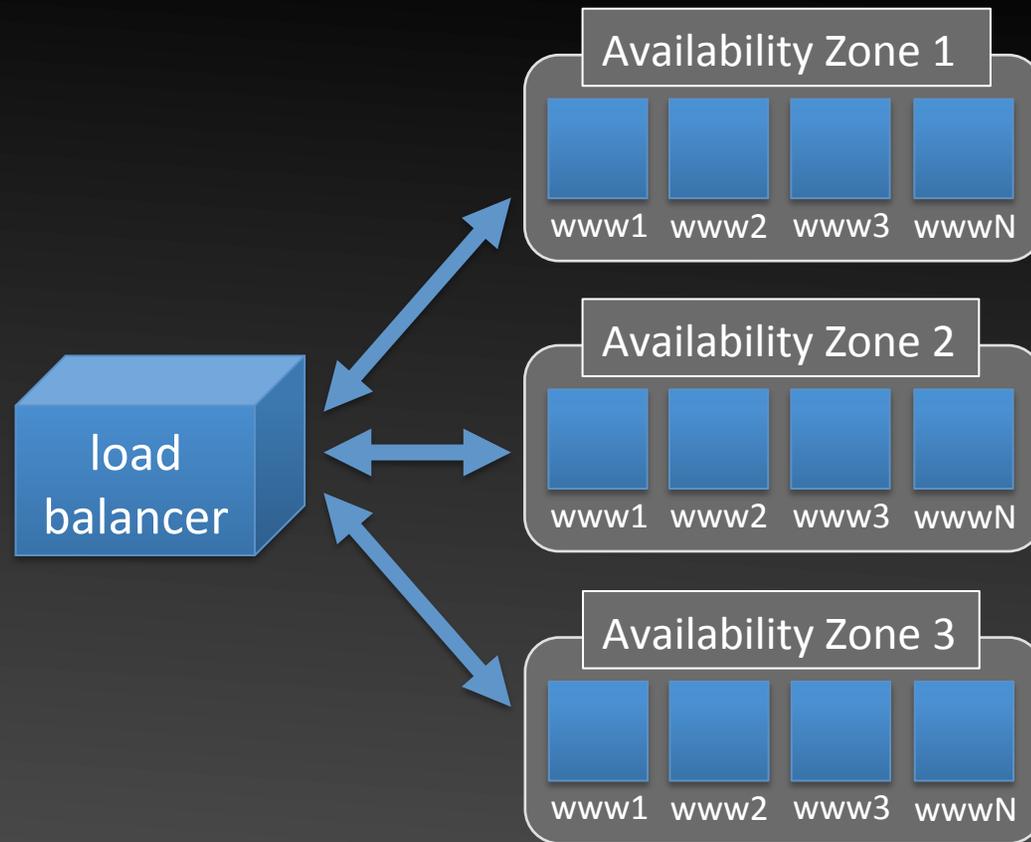


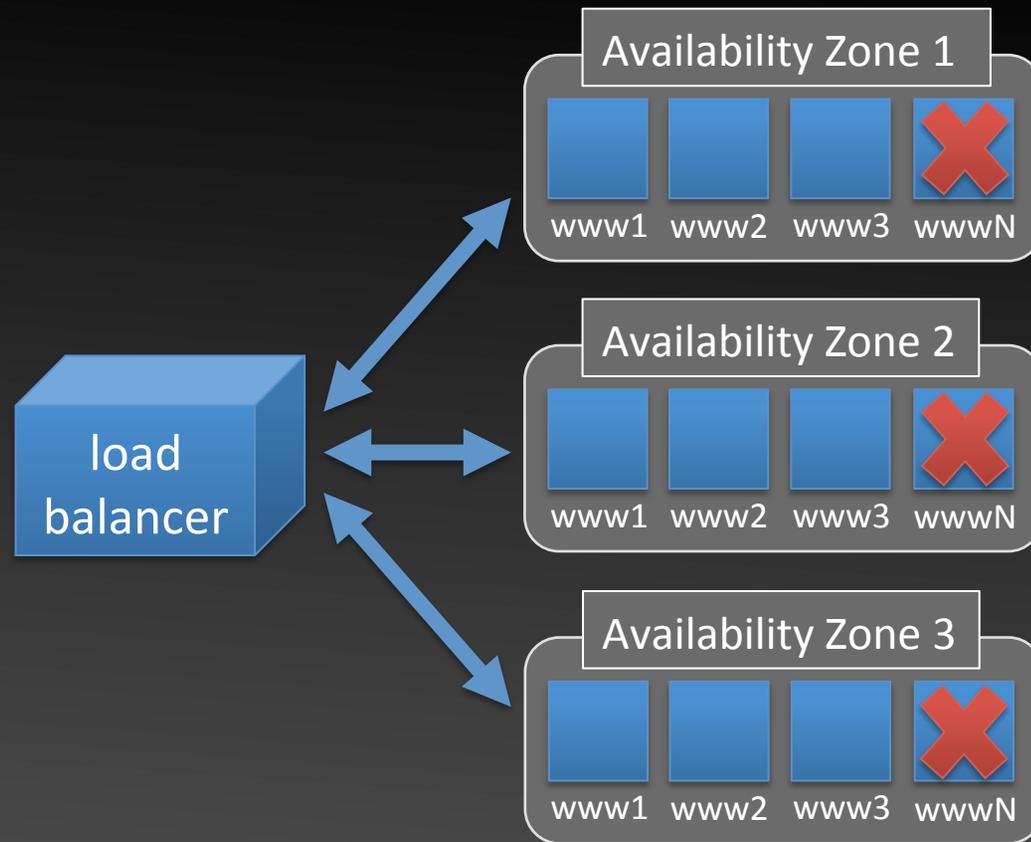
11.6 seconds mean time between deployments
(weekday)

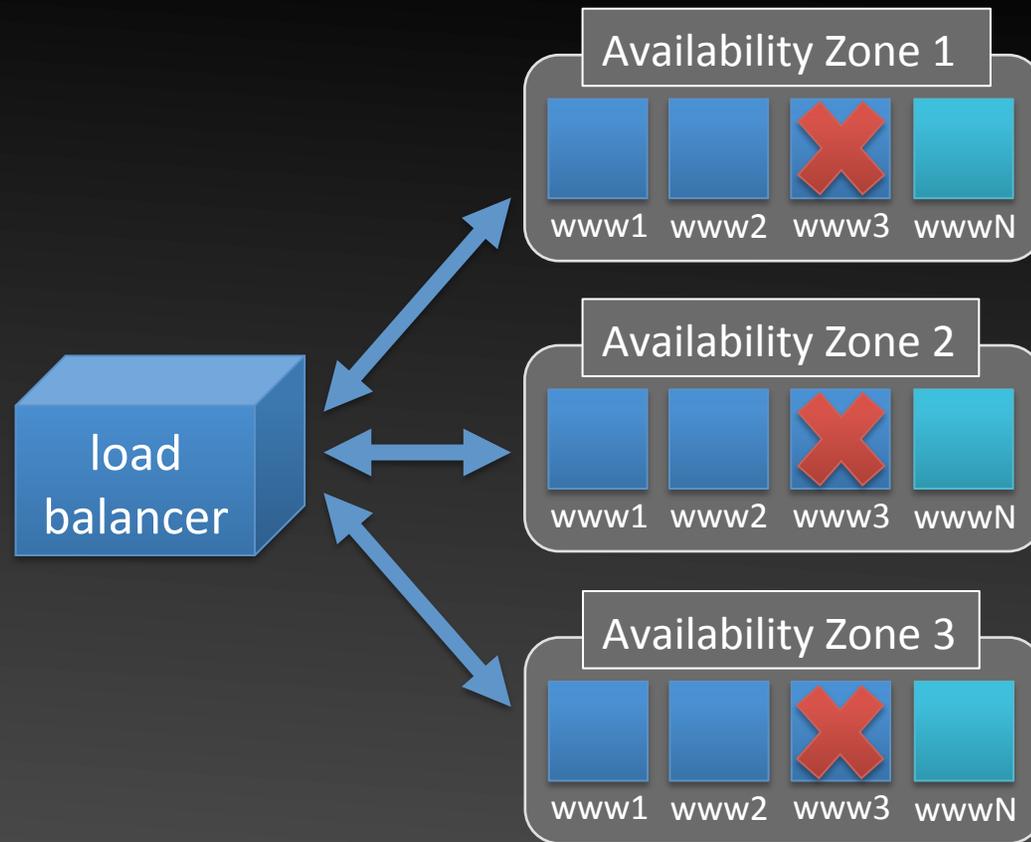
1,079 maximum number of deployments in a
single hour

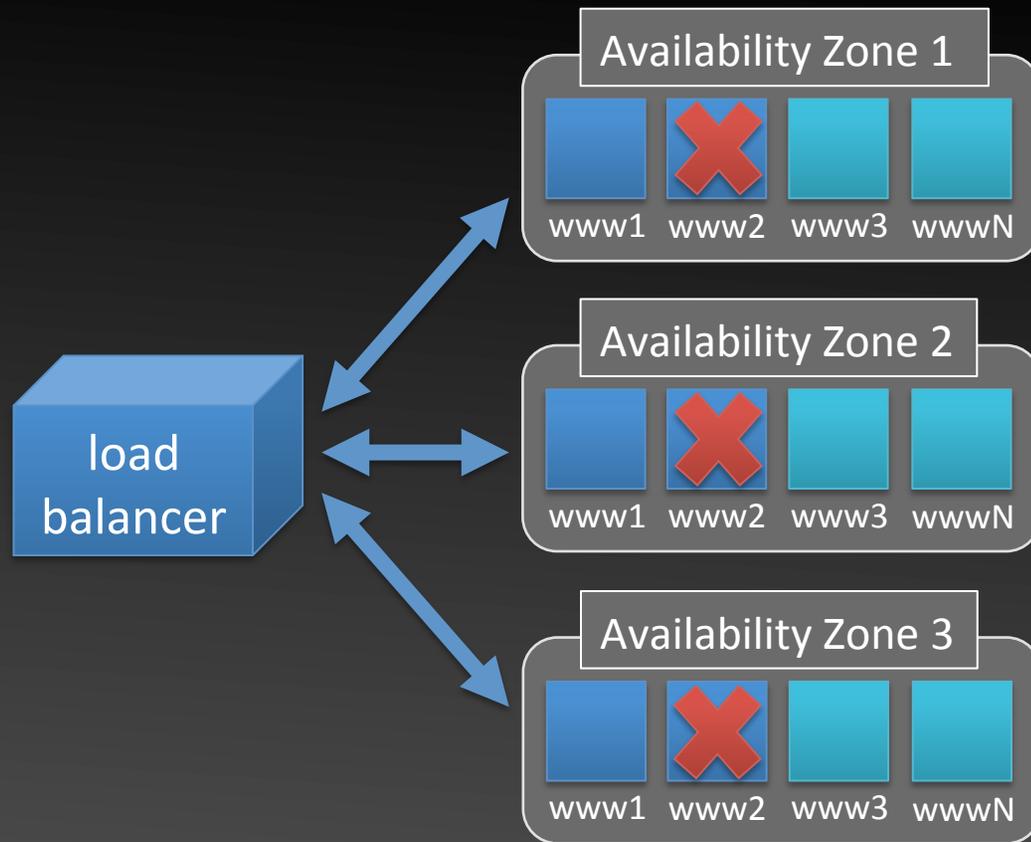
10,000 mean number of hosts simultaneously
receiving a deployment

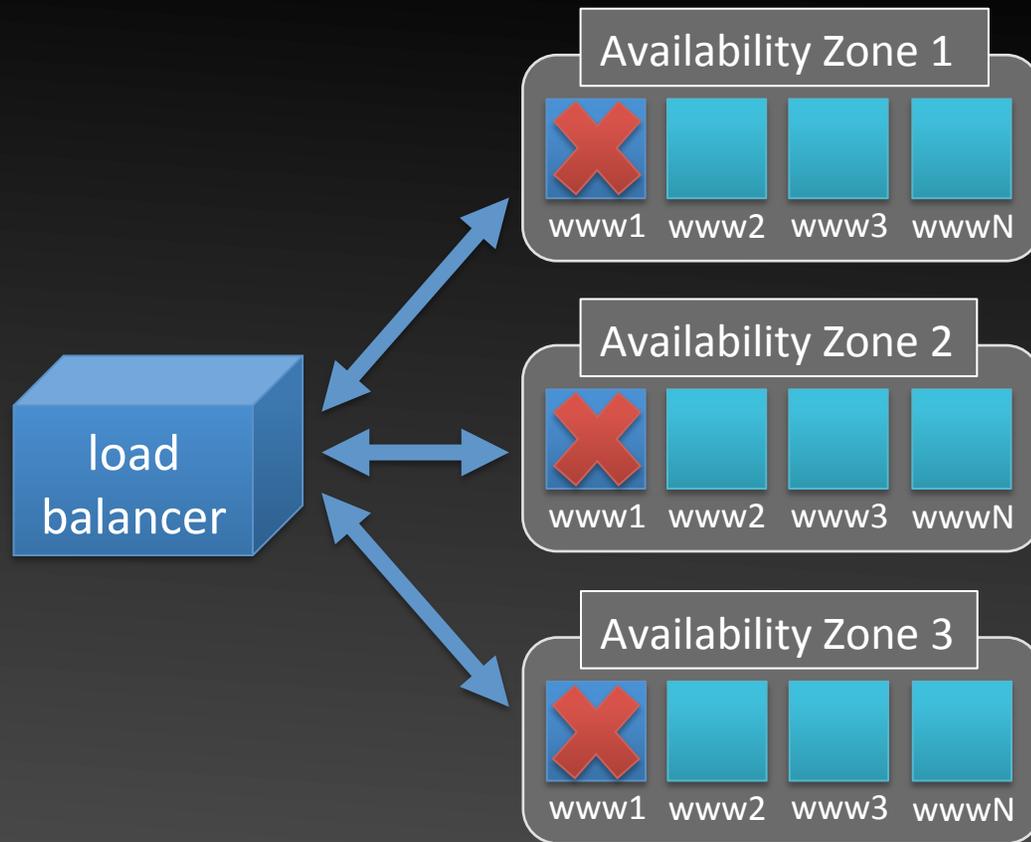
30,000 maximum number of hosts
simultaneously receiving a deployment

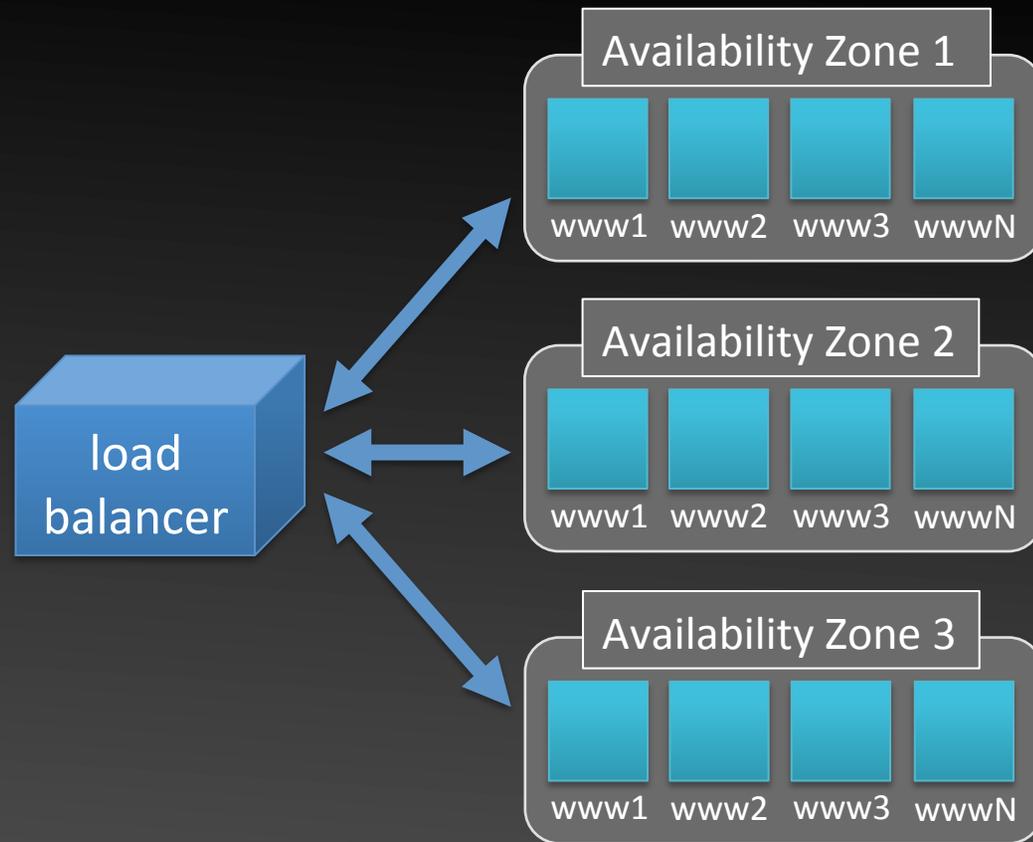










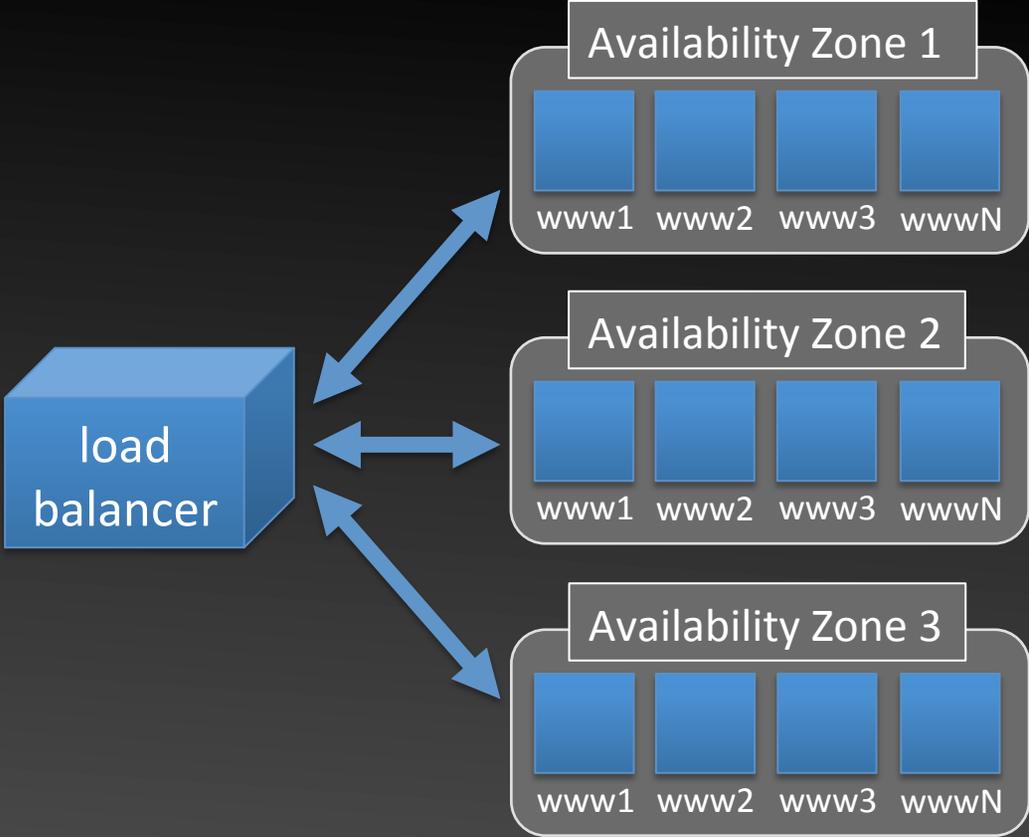


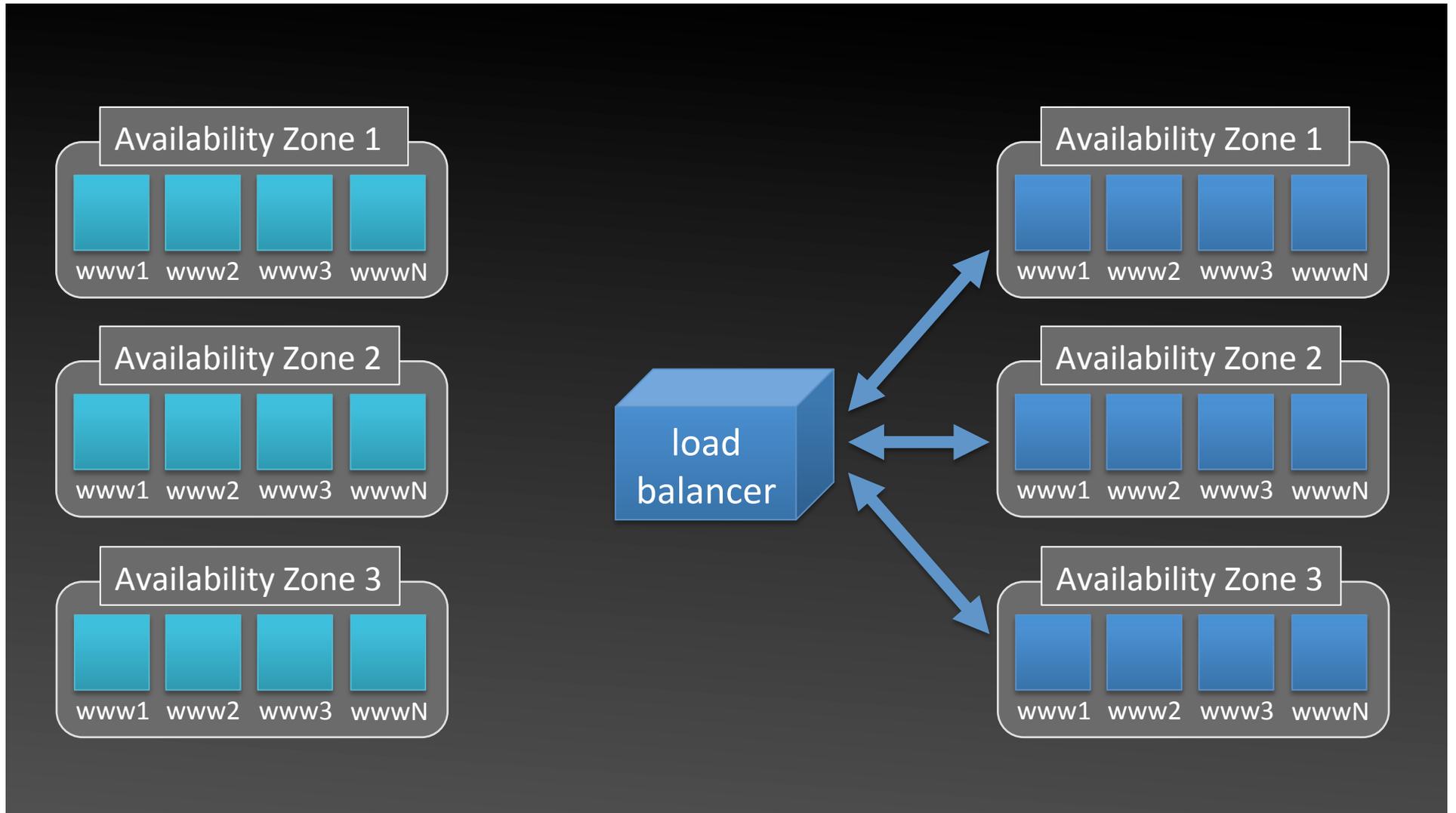
Deploying new software to a fixed fleet requires a complex workflow.

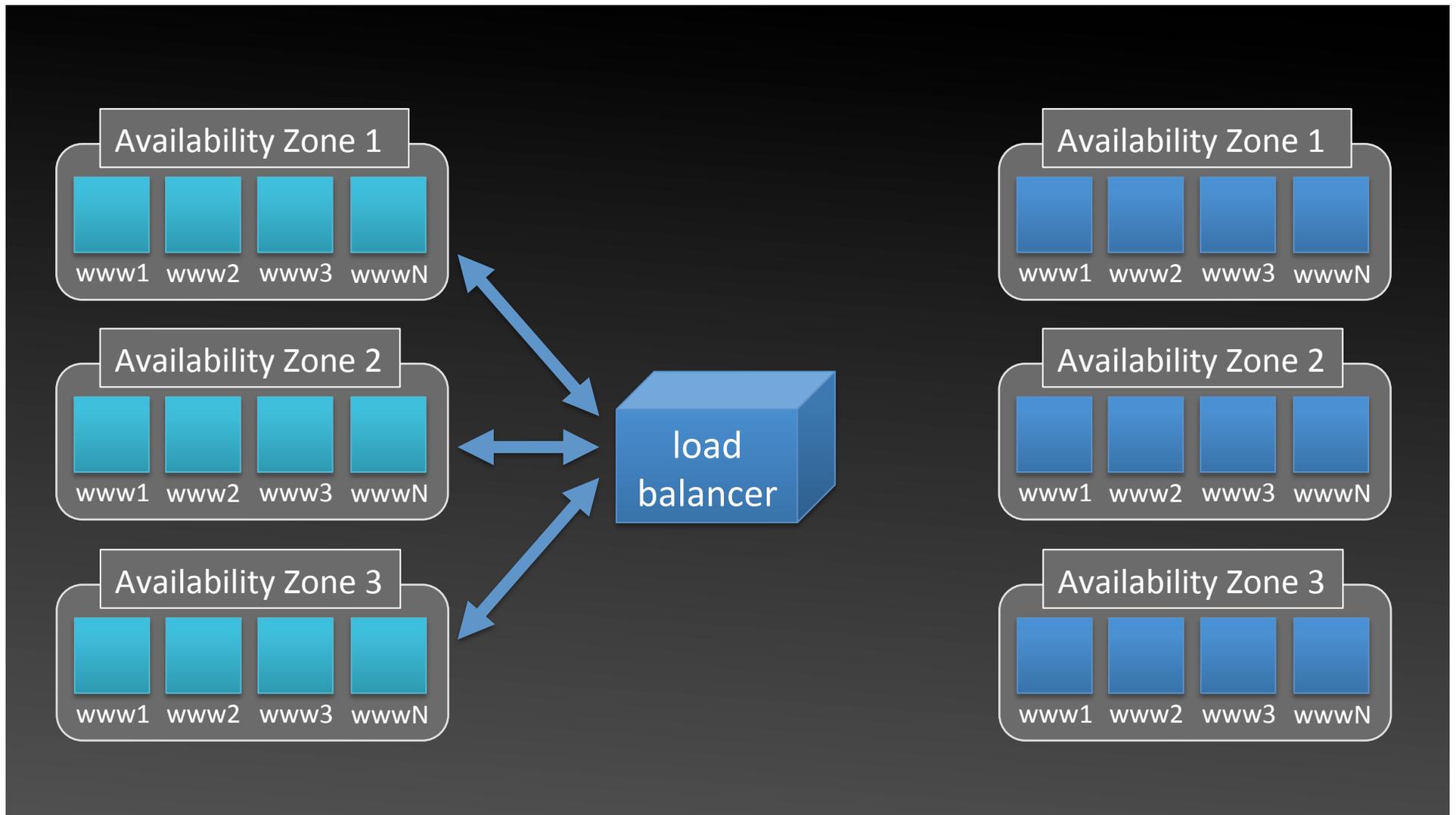
Deploying new software to a fixed
fleet is a slow process.

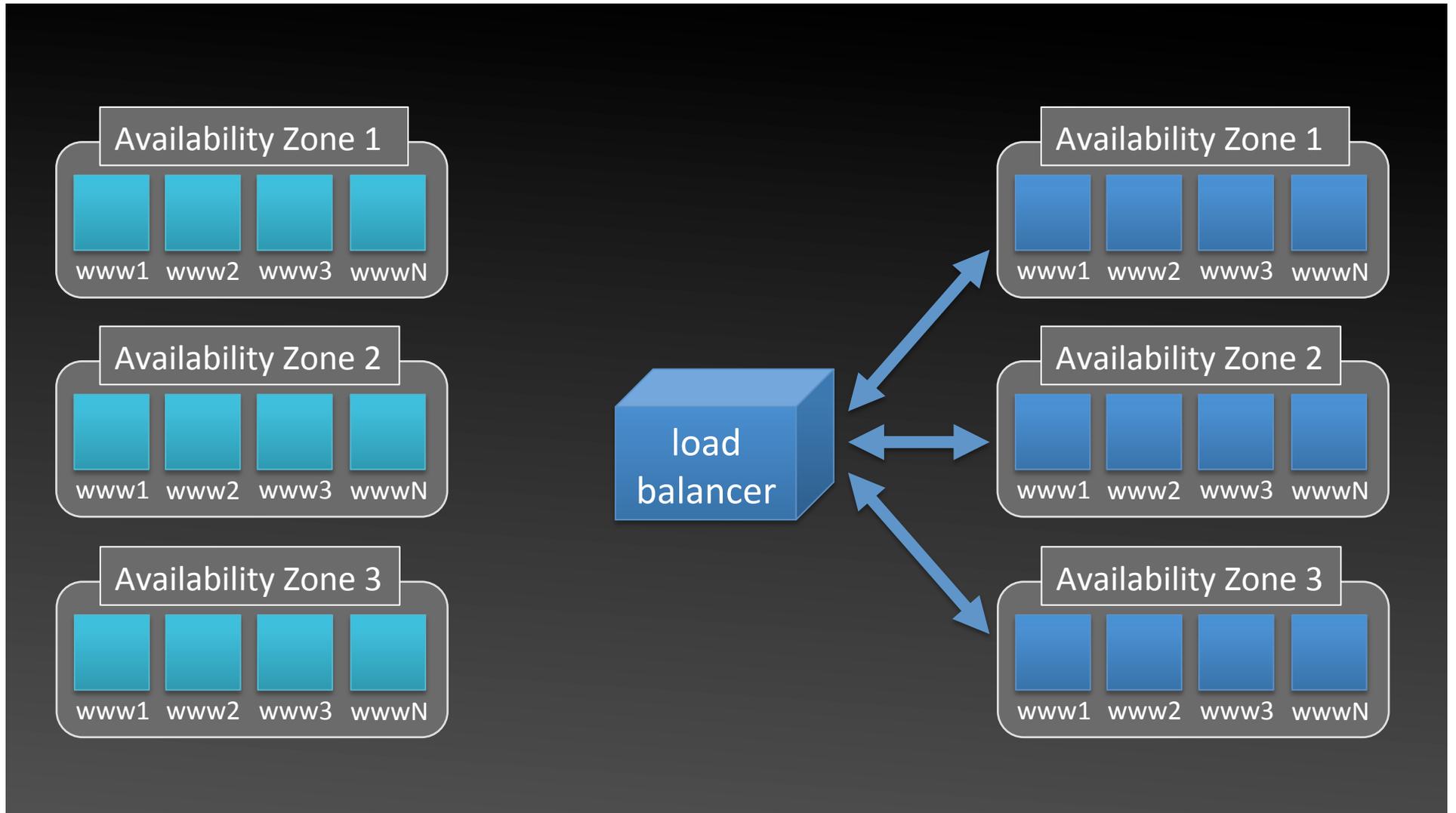
Dealing with failure scenarios requires emergent, high-judgment decisions.

What if you had unlimited capacity?









75% reduction in outages triggered by
software deployments since 2006

90% reduction in outage minutes
triggered by software deployments

~0.001% of software deployments
cause an outage

lessons learned

business lessons

We spend less time on capacity planning.

Fewer conversations with Finance

More innovation, happier developers

I get credit for AWS price reductions.

Be sure to consider compliance issues.

technical lessons

Start with simple applications.

Iterate toward your desired end-state.

Identify reusable components.

Engage security early and treat them as partners.

Migrate to the cloud in concert with your other architectural objectives.

The cloud doesn't cover up sloppy engineering.

end