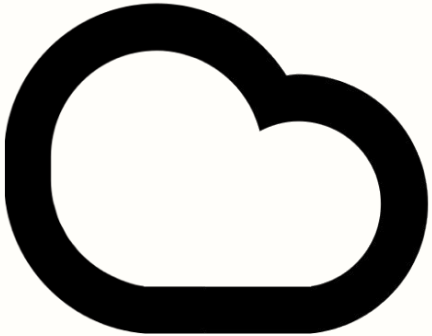


CloudConf

okta

okta
The World's Identity Company

The Scale of Identity

Challenges of Identity Management at Scale

Norberto Leite

Principal Engineer@Okta

#cloudconf24

Safe harbor

This presentation contains "forward-looking statements" within the meaning of the "safe harbor" provisions of the Private Securities Litigation Reform Act of 1995, including but not limited to, statements regarding our financial outlook, long-term financial targets, product development, business strategy and plans, market trends and market size, opportunities, positioning and expected benefits that will be derived from the acquisition of Auth0, Inc. These forward-looking statements are based on current expectations, estimates, forecasts and projections. Words such as "expect," "anticipate," "should," "believe," "hope," "target," "project," "goals," "estimate," "potential," "predict," "may," "will," "might," "could," "intend," "shall" and variations of these terms and similar expressions are intended to identify these forward-looking statements, although not all forward-looking statements contain these identifying words. Forward-looking statements are subject to a number of risks and uncertainties, many of which involve factors or circumstances that are beyond our control. For example, the market for our products may develop more slowly than expected or than it has in the past; there may be significant fluctuations in our results of operations and cash flows related to our revenue recognition or otherwise; we may fail to successfully integrate any new business, including Auth0, Inc.; we may fail to realize anticipated benefits of any combined operations with Auth0, Inc.; we may experience unanticipated costs of integrating Auth0, Inc.; the potential impact of the acquisition on relationships with third parties, including employees, customers, partners and competitors; we may be unable to retain key personnel; global economic conditions could worsen; a network or data security incident that allows unauthorized access to our network or data or our customers' data could damage our reputation and cause us to incur significant costs; we could experience interruptions or performance problems associated with our technology, including a service outage; the impact of COVID-19 and variants of concern, related public health measures and any associated economic downturn on our business and results of operations may be more than we expect; and we may not be able to pay off our convertible senior notes when due. Further information on potential factors that could affect our financial results is included in our most recent Quarterly Report on Form 10-Q and our other filings with the Securities and Exchange Commission. The forward-looking statements included in this presentation represent our views only as of the date of this presentation and we assume no obligation and do not intend to update these forward-looking statements.

Any unreleased products, features or functionality referenced in this presentation are not currently available and may not be delivered on time or at all. Product roadmaps do not represent a commitment, obligation or promise to deliver any product, feature or functionality, and you should not rely on them to make your purchase decisions.

okta

Who here is familiar with Okta?

okta

Who here knows about Auth0?

okta

okta
auth0

Ciao, I'm Norberto!

- Principal Engineer @ Okta
- Databases, that's my thing
- I also like to talk to other people

norberto.leite@okta.com @nleite

Agenda

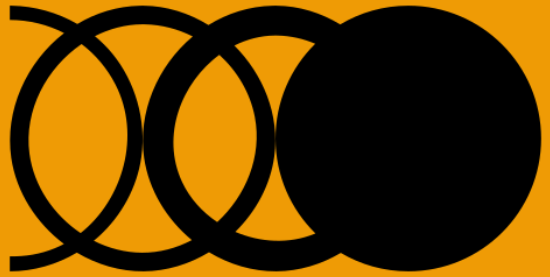
- Challenge of Scaling Identity Management
- Operational Challenges and Solutions
- Service Releases and Infrastructure Operations
- Database Management at Scale

What are the different Degrees of Freedom a SaaS service like CIAM can look like?

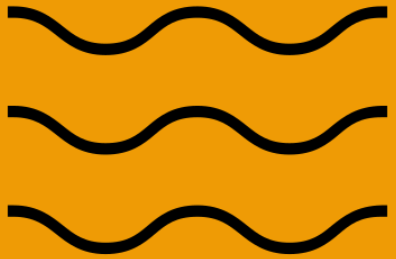
https://en.wikipedia.org/wiki/Degrees_of_freedom_(physics_and_chemistry)

Degrees of Freedom

Translational Motion



Vibration



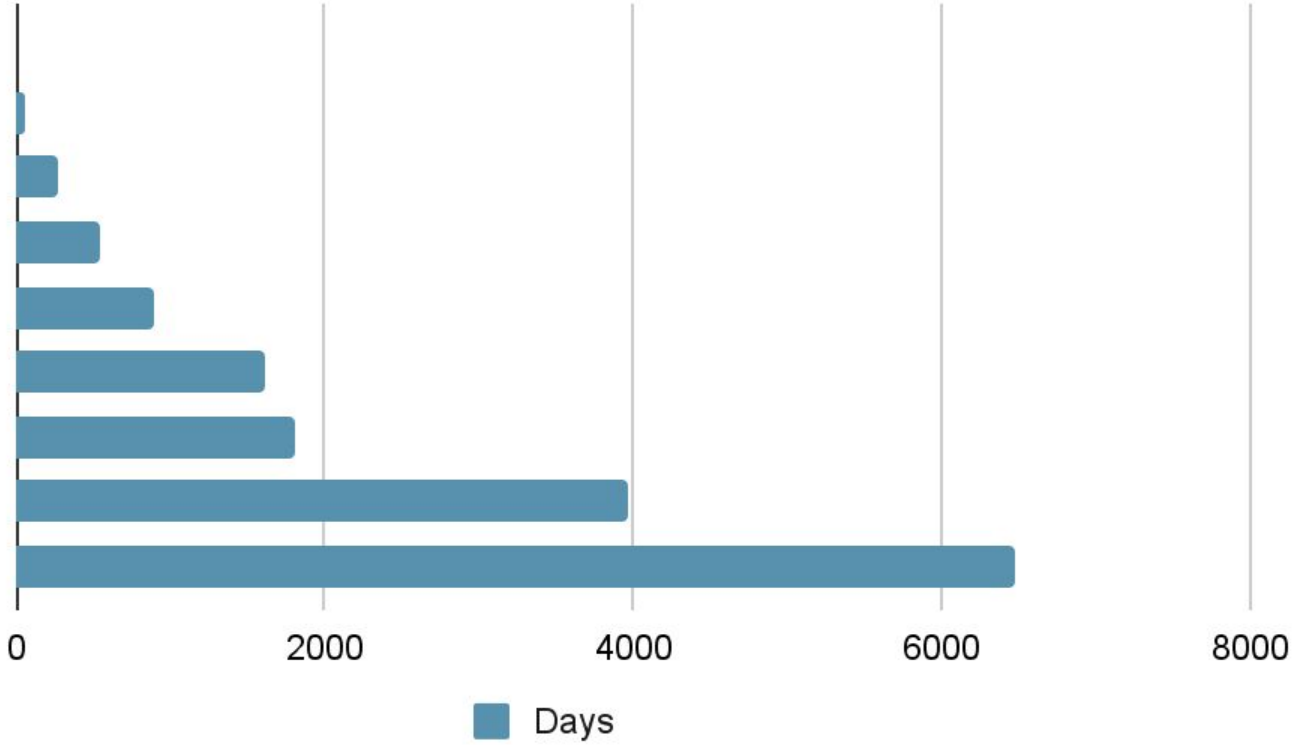
Rotation

Auth0 by Okta

The Challenge of Scaling Identity Management (CIAM)

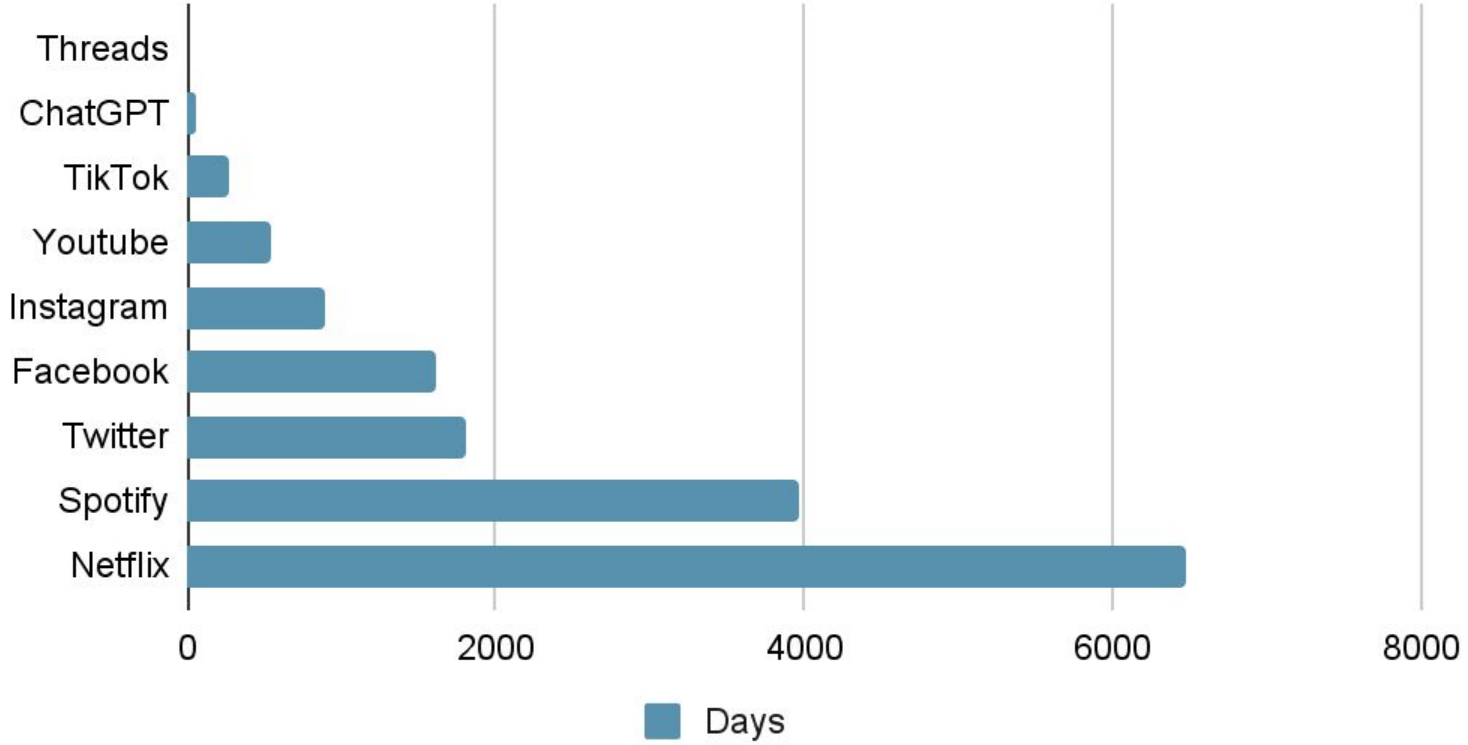
https://en.wikipedia.org/wiki/On_the_Internet,_nobody_knows_you%27re_a_dog

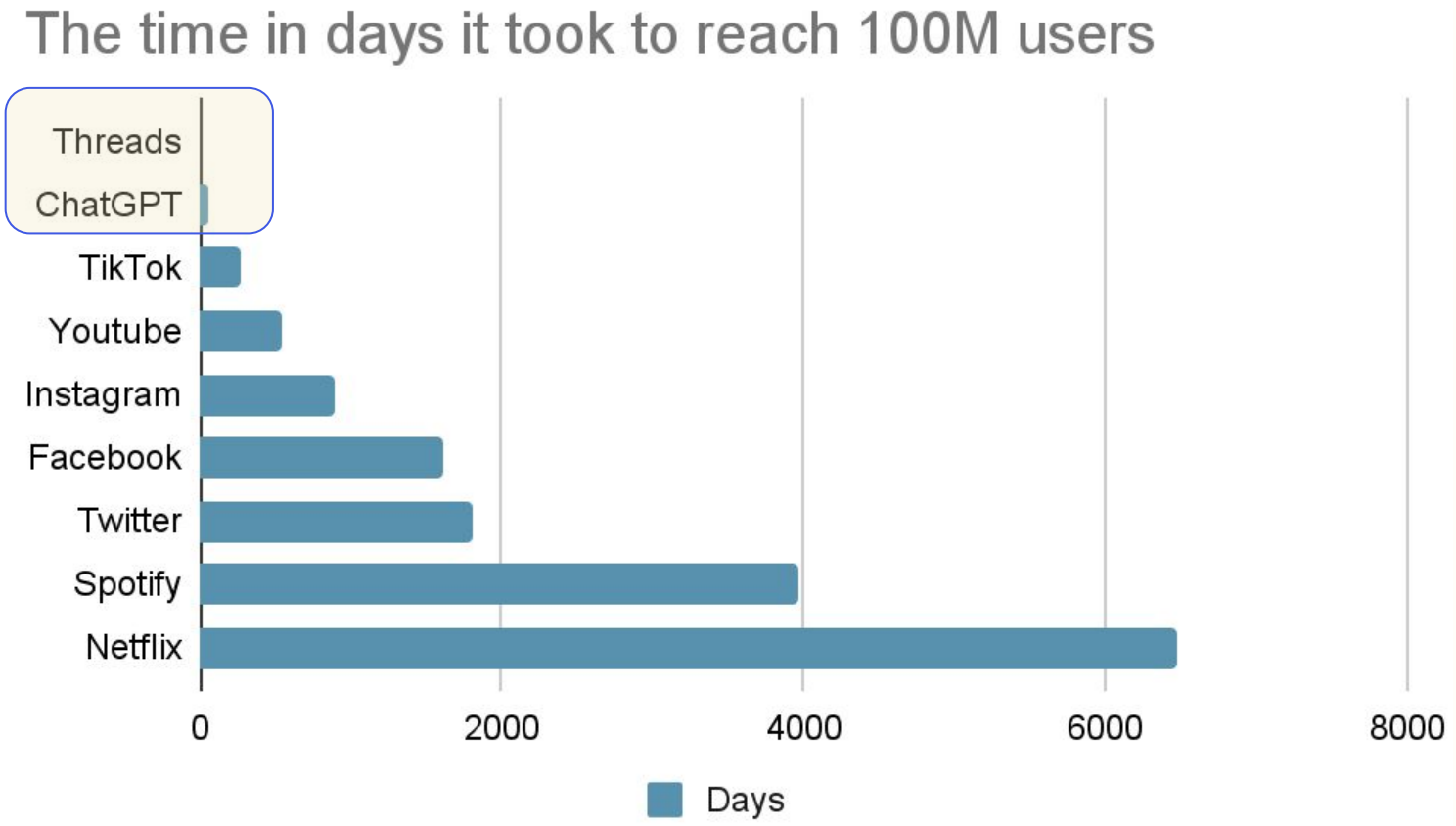
The time in days it took



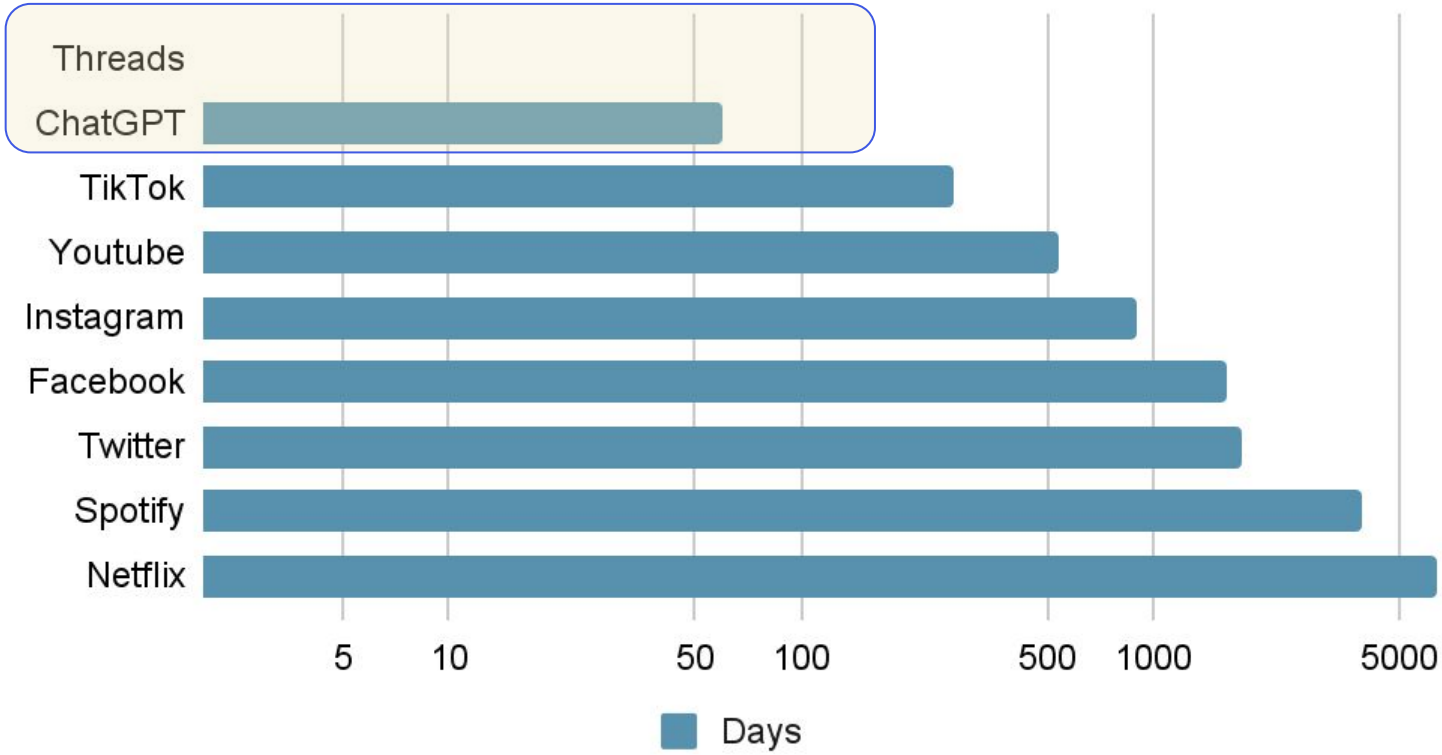
okta

The time in days it took to reach 100M users




The time in days it took to reach 100M users
Threads
ChatGPT
TikTok
Youtube
Instagram
Facebook
Twitter
Spotify
Netflix
0
2000
4000
6000
8000
Days

The time in days it took to reach 100M users (log)

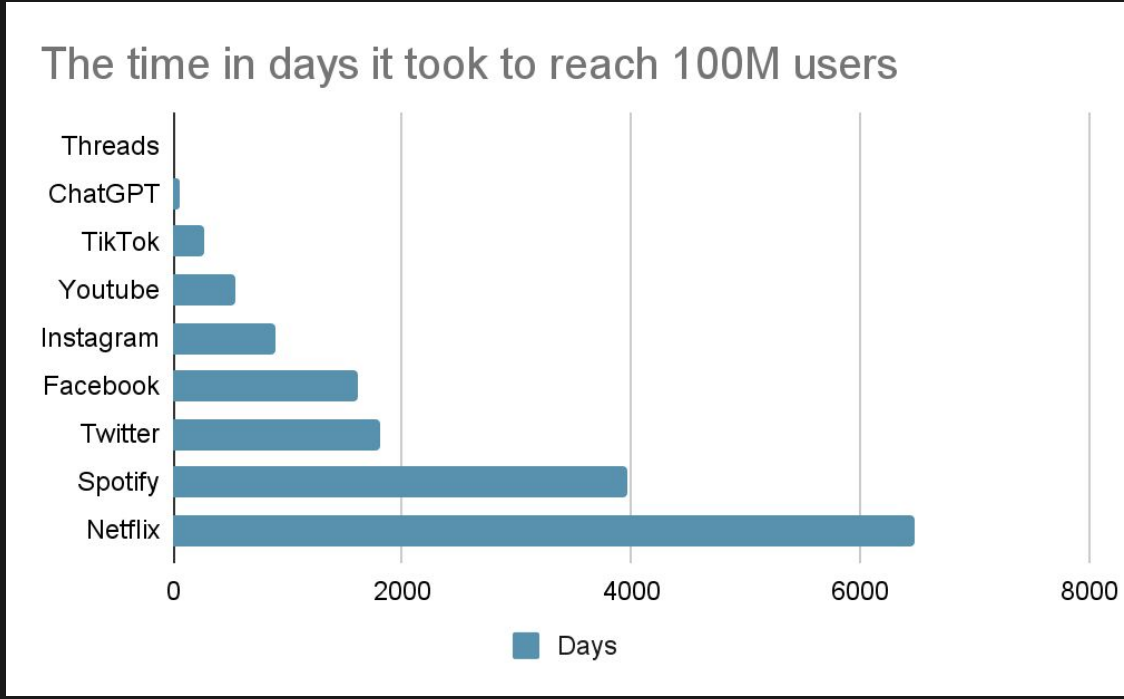


okta

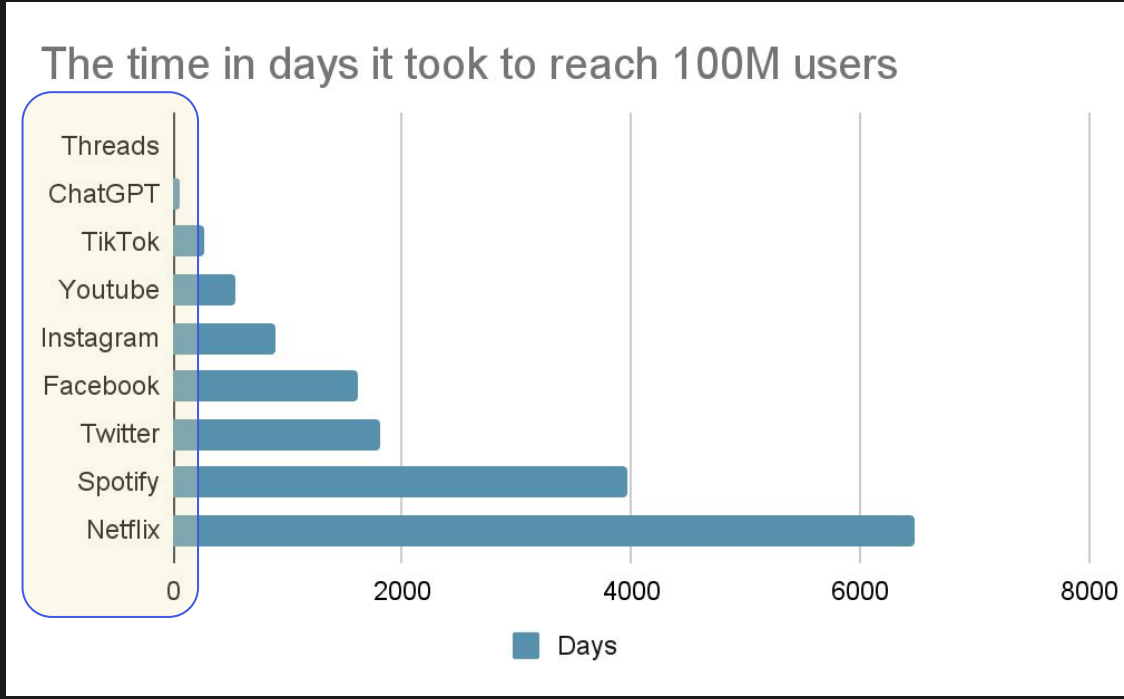
How does one build for 100M users in a week? Or even in 2 days?

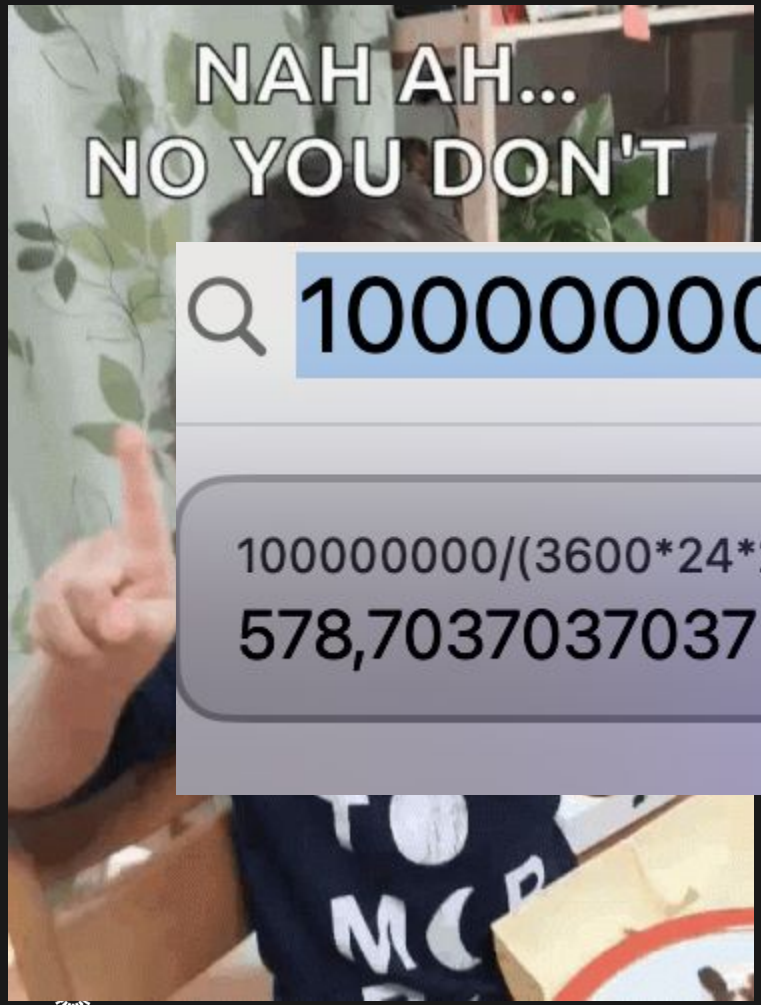
NAH AH...
NO YOU DON'T
TIL
TO

NAH AH...
NO YOU DON'T
TIL
TO


The time in days it took to reach 100M users
Threads
ChatGPT
TikTok
Youtube
Instagram
Facebook
Twitter
Spotify
Netflix
0
2000
4000
6000
8000
Days

NAH AH...
NO YOU DON'T
TIL
TO


The time in days it took to reach 100M users
Threads
ChatGPT
TikTok
Youtube
Instagram
Facebook
Twitter
Spotify
Netflix
0
2000
4000
6000
8000
Days



100000000/(3600*24*2)

100000000/(3600*24*2) =
578,7037037037

The time in days it took to reach 100M users

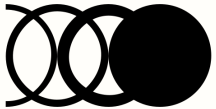
8000

Days

Load Distribution

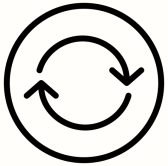
no matter how crazy they may seem

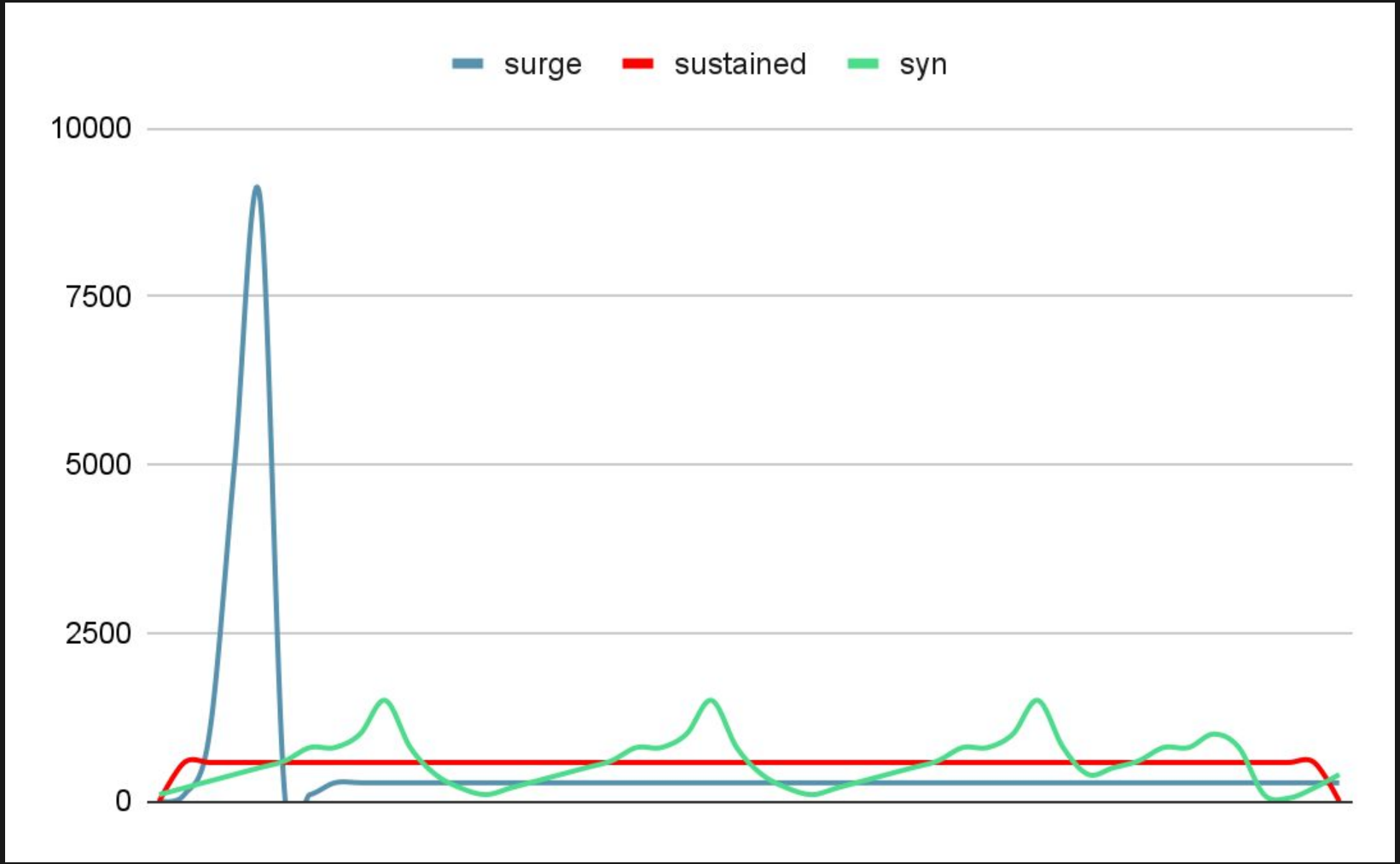
How load evolved over a defined period of time - implies understanding load form an average and/or aggregate perspective



How choppy/variable is your load over short periods of time, where sudden bursts and variations in terms of max and min values

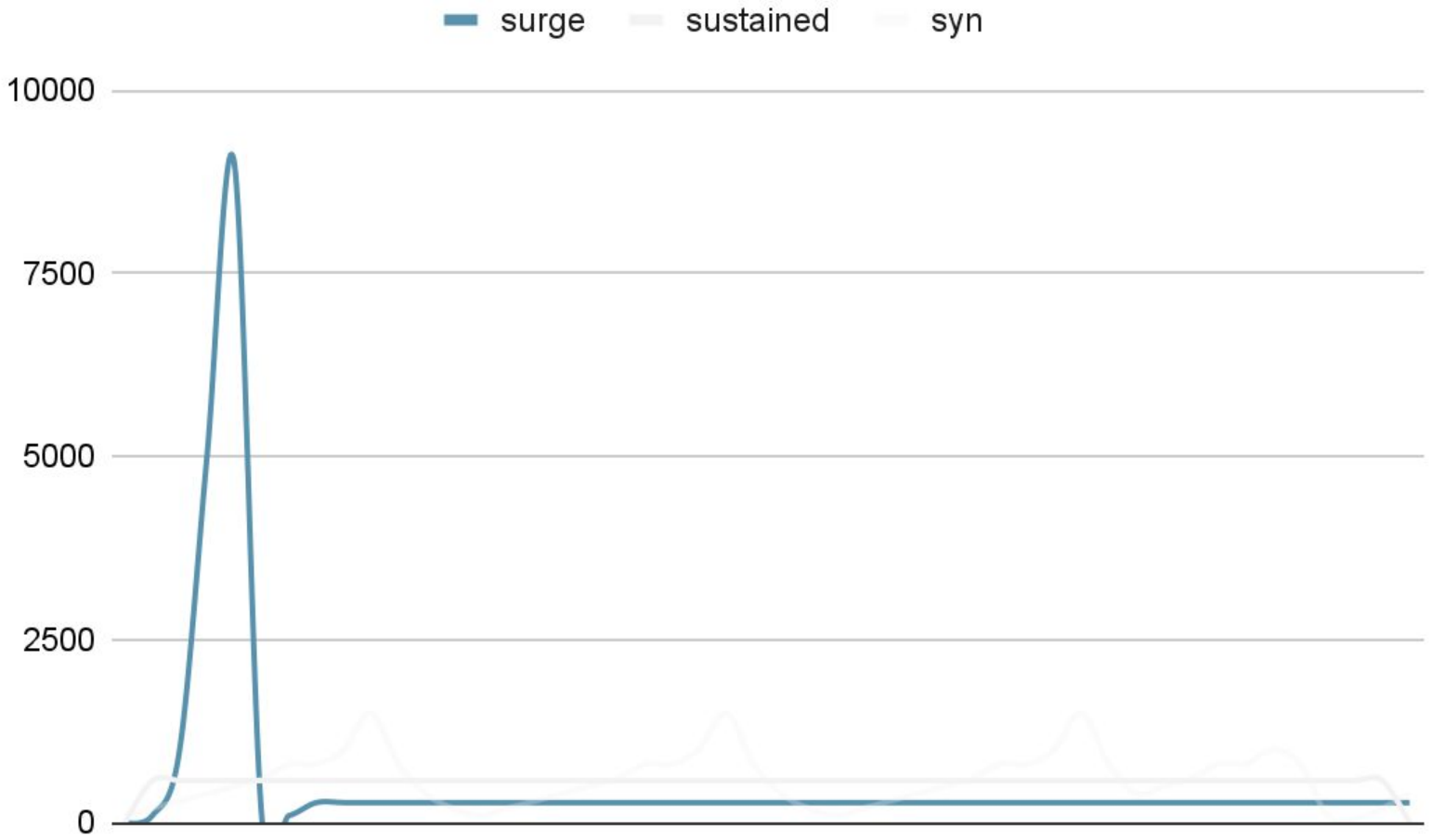
What pattern(s) can we see on our load over recurrent periods of time - how often certain load patterns repeat themselves



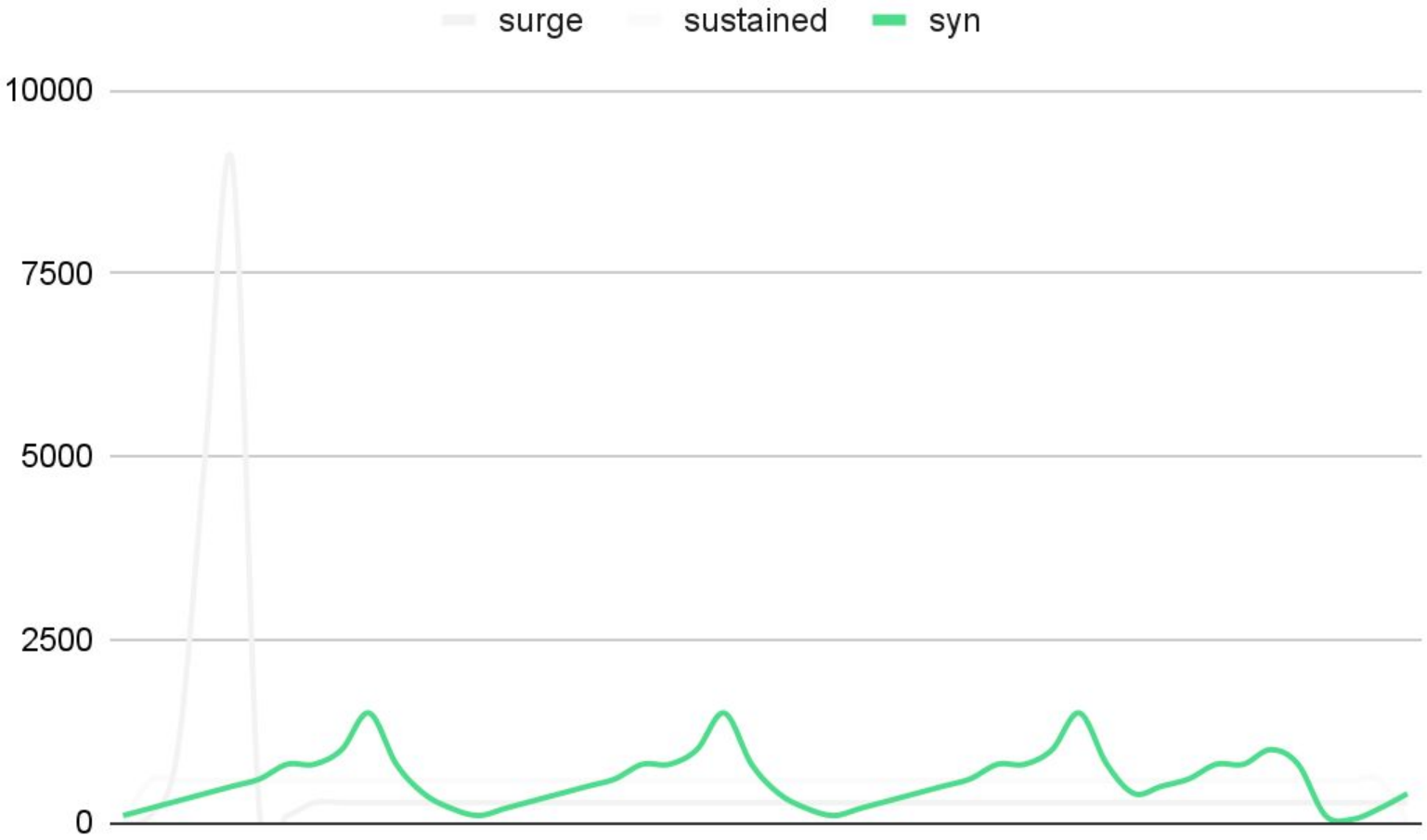


okta

okta



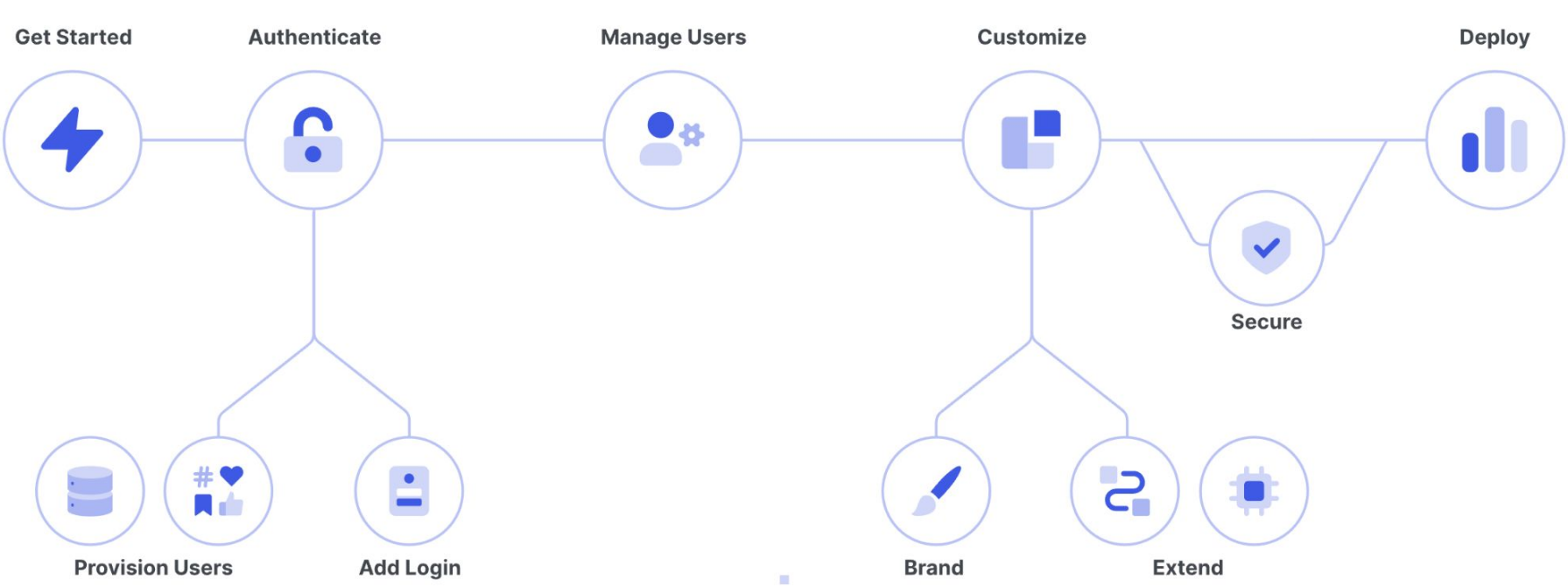
okta


surge
sustained
syn
10000
7500
5000
2500
0

Ok, is it just load that is challenging for an CIAM provider?

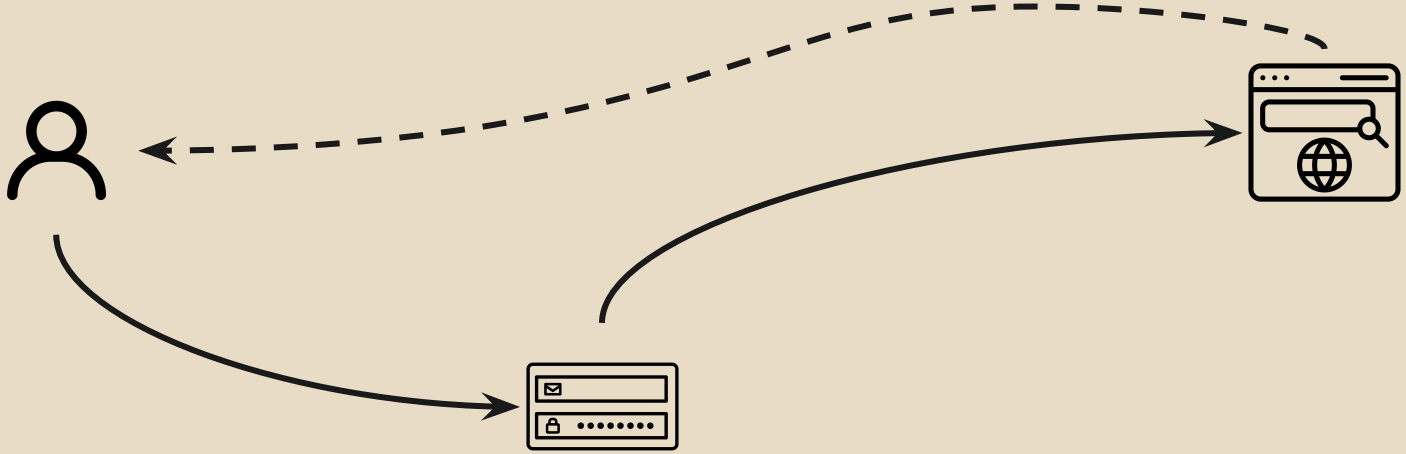
CIAM Platform Features

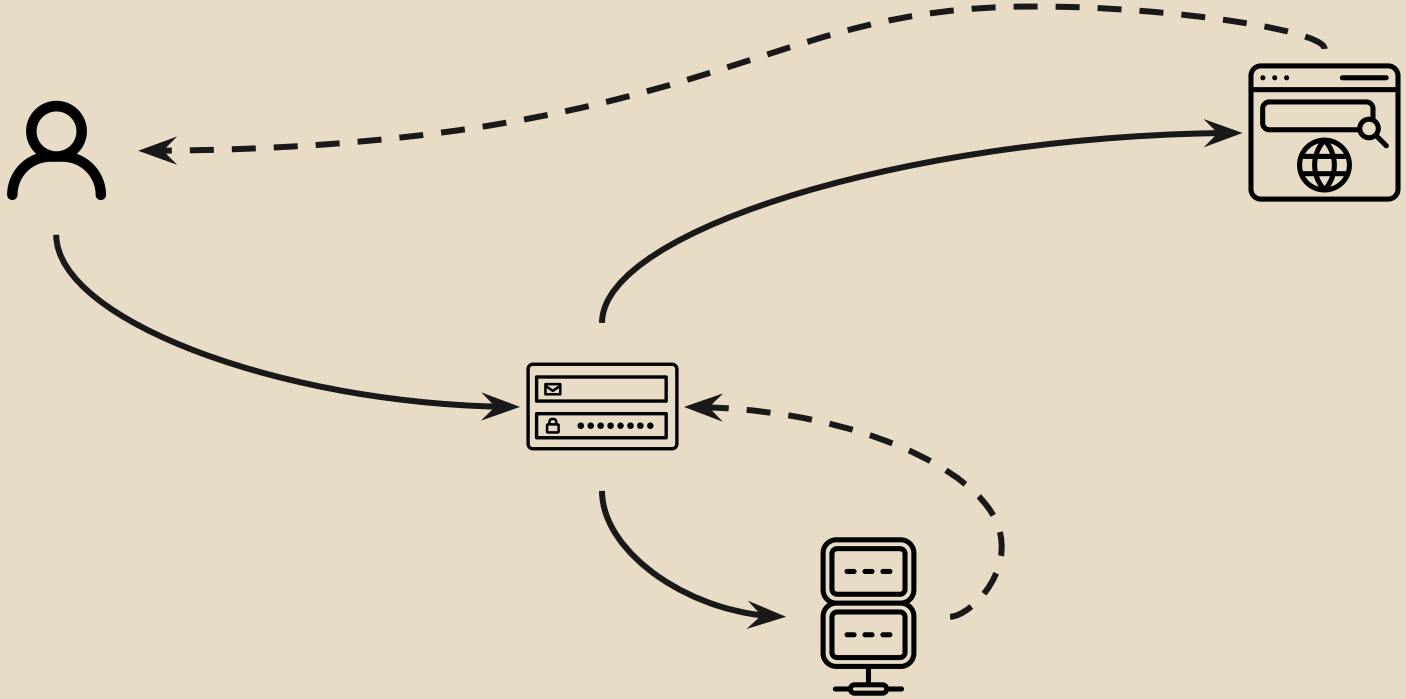
All the goodies

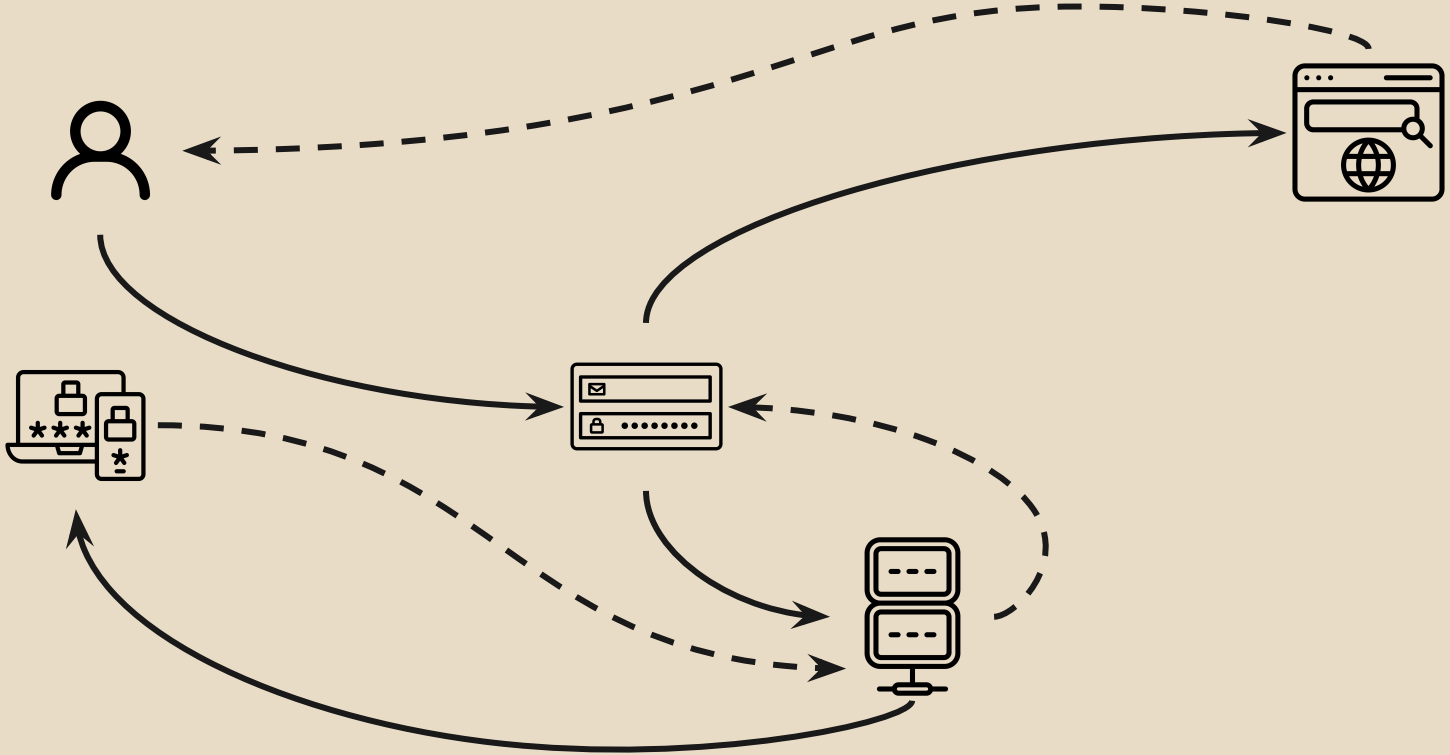


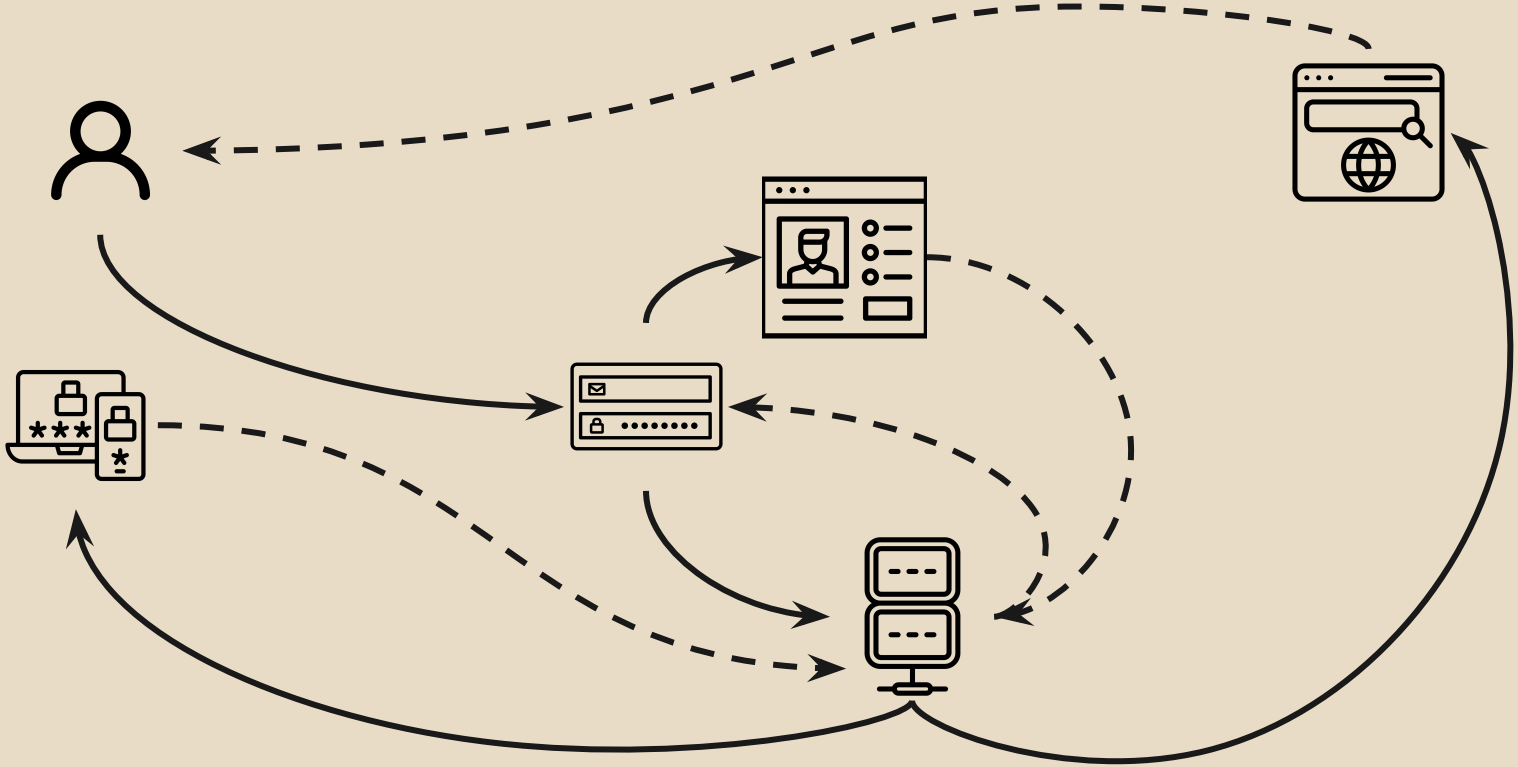
https://auth0.com/docs

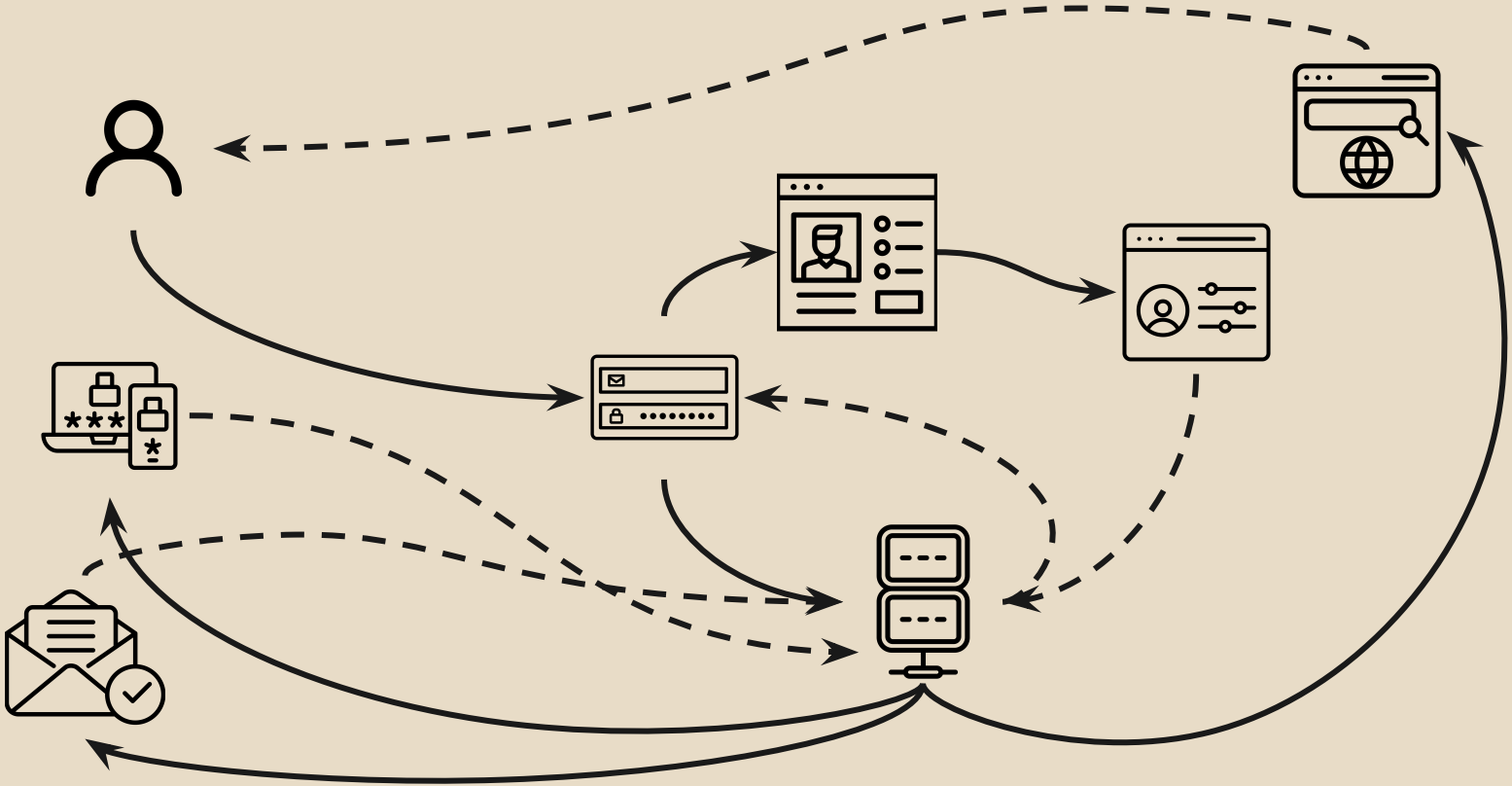
okta

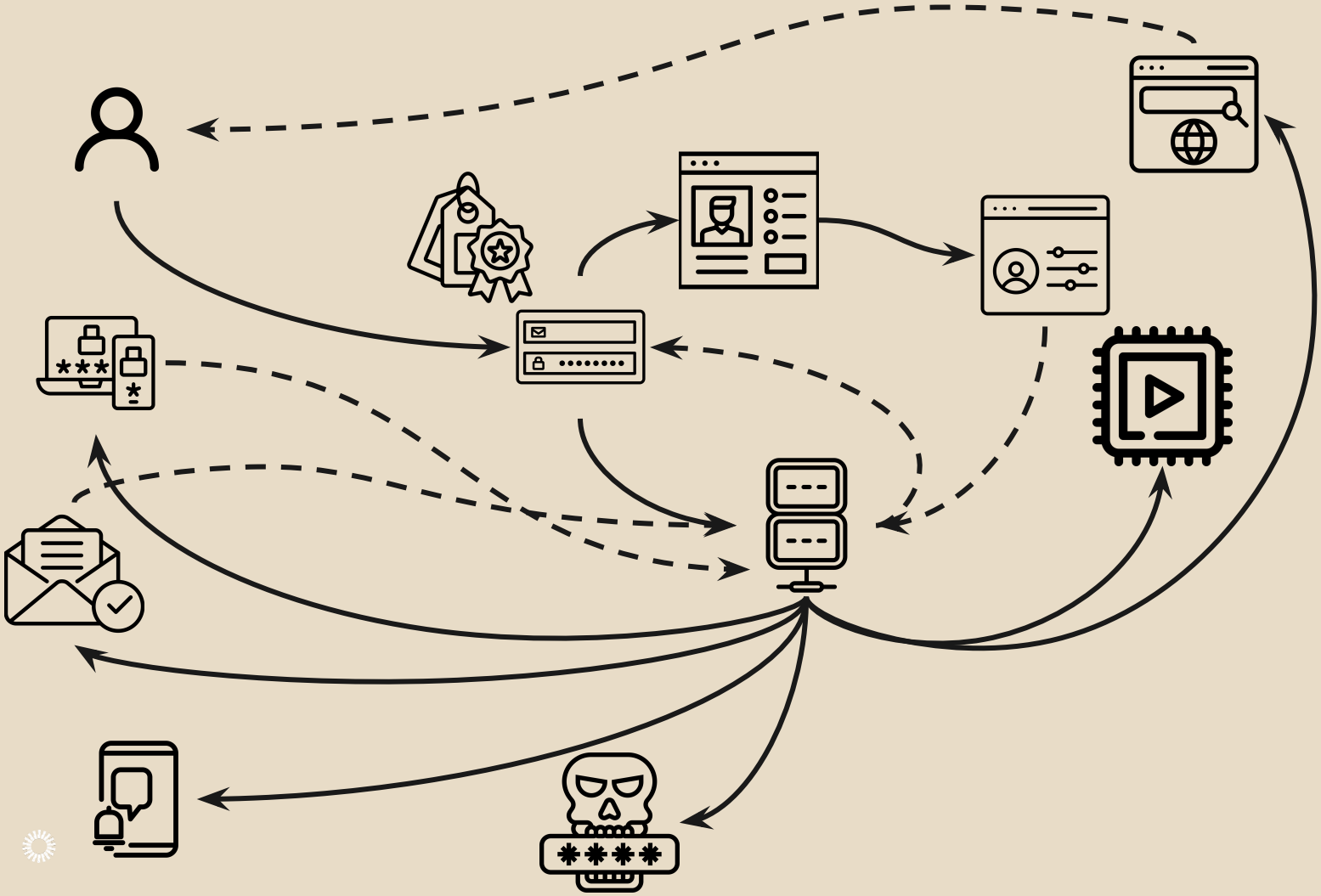




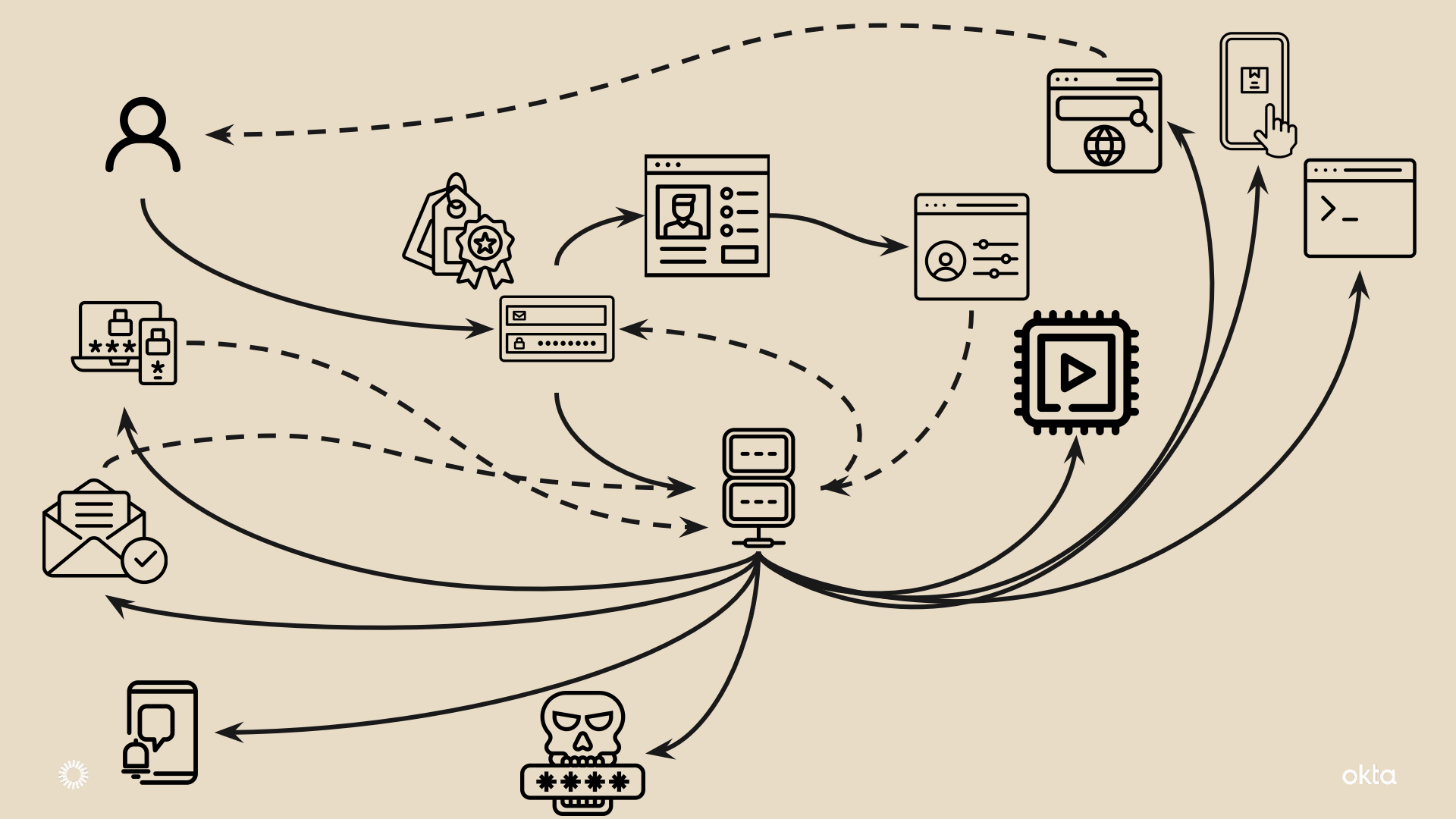


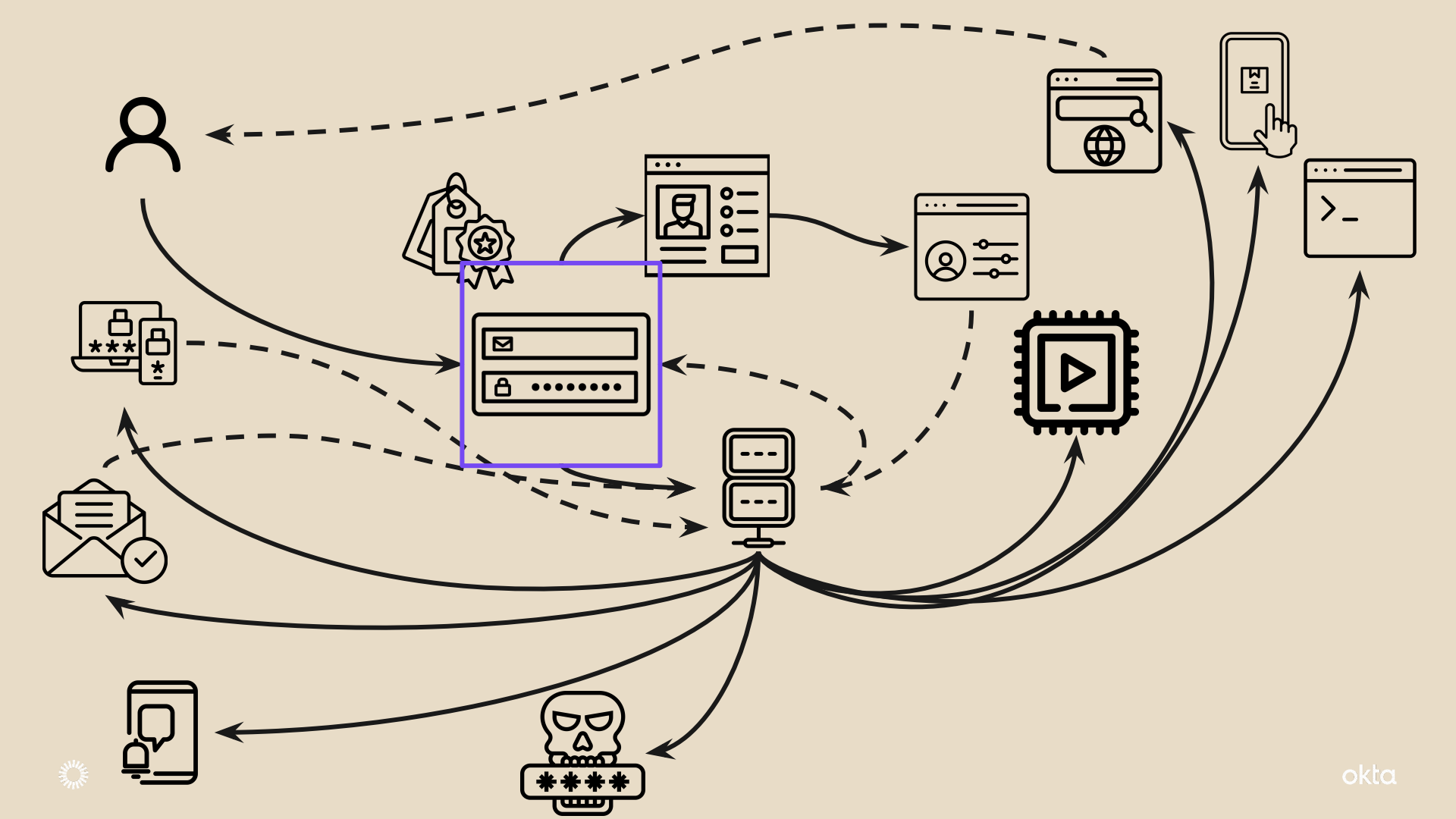


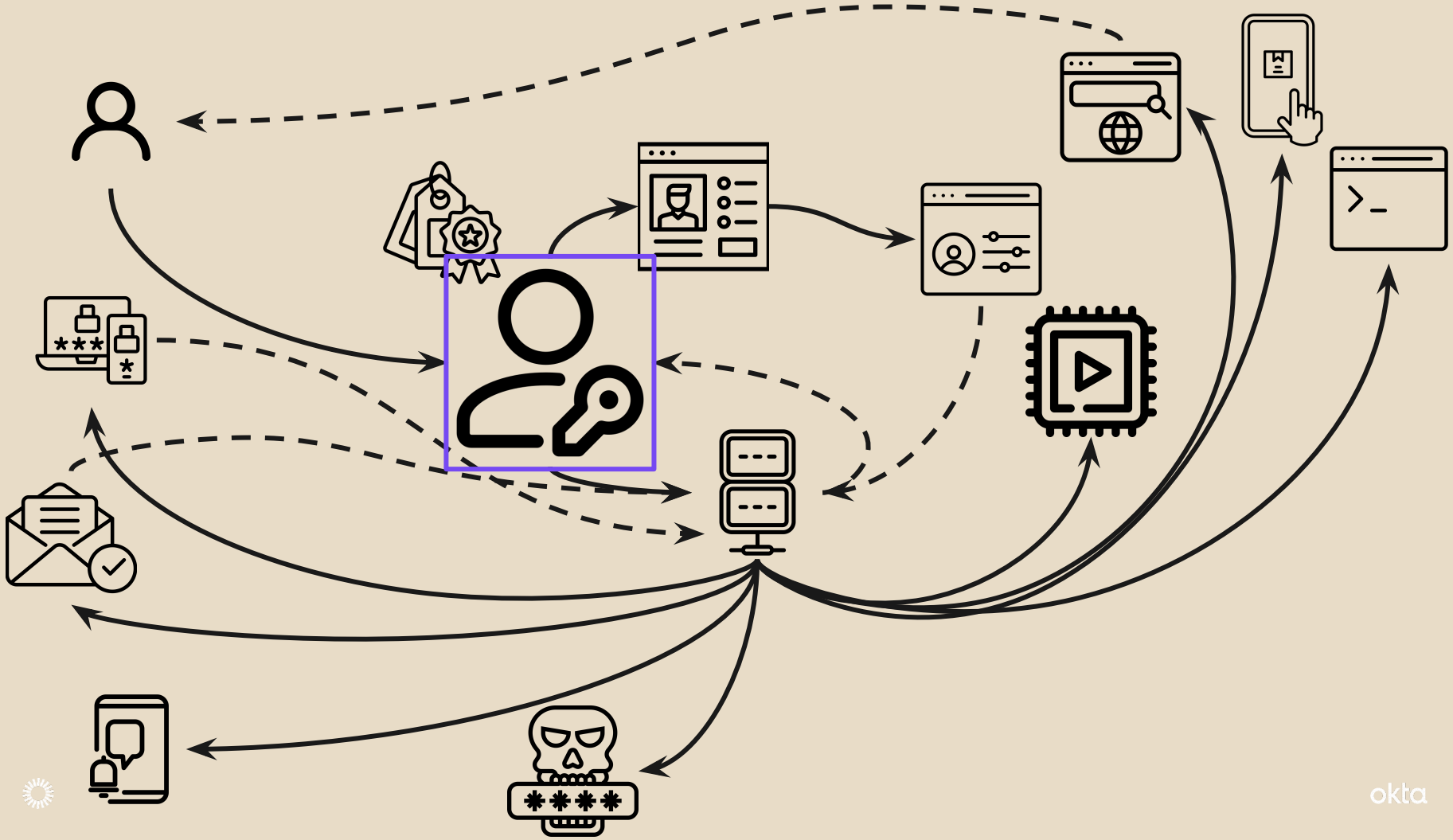


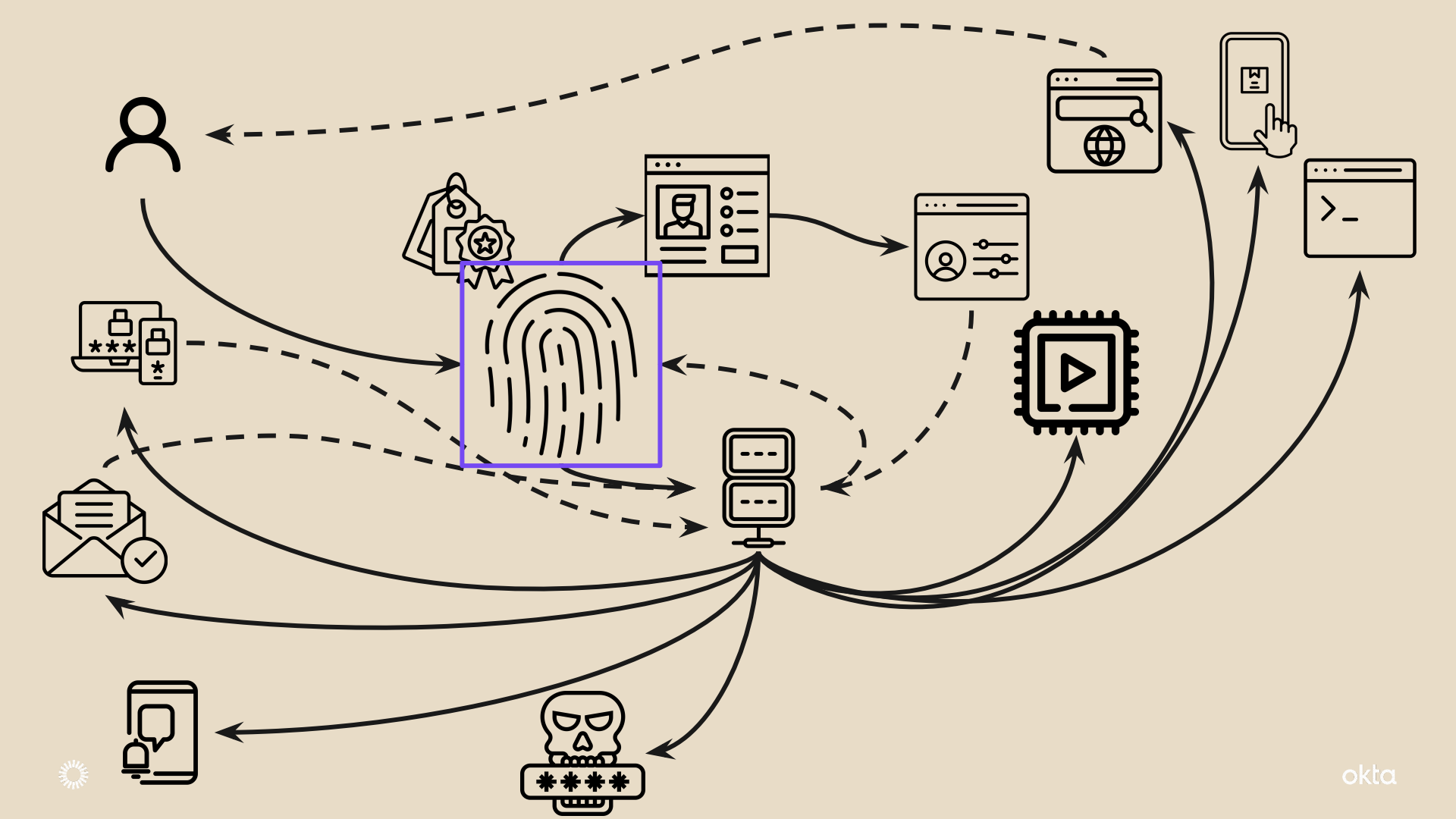


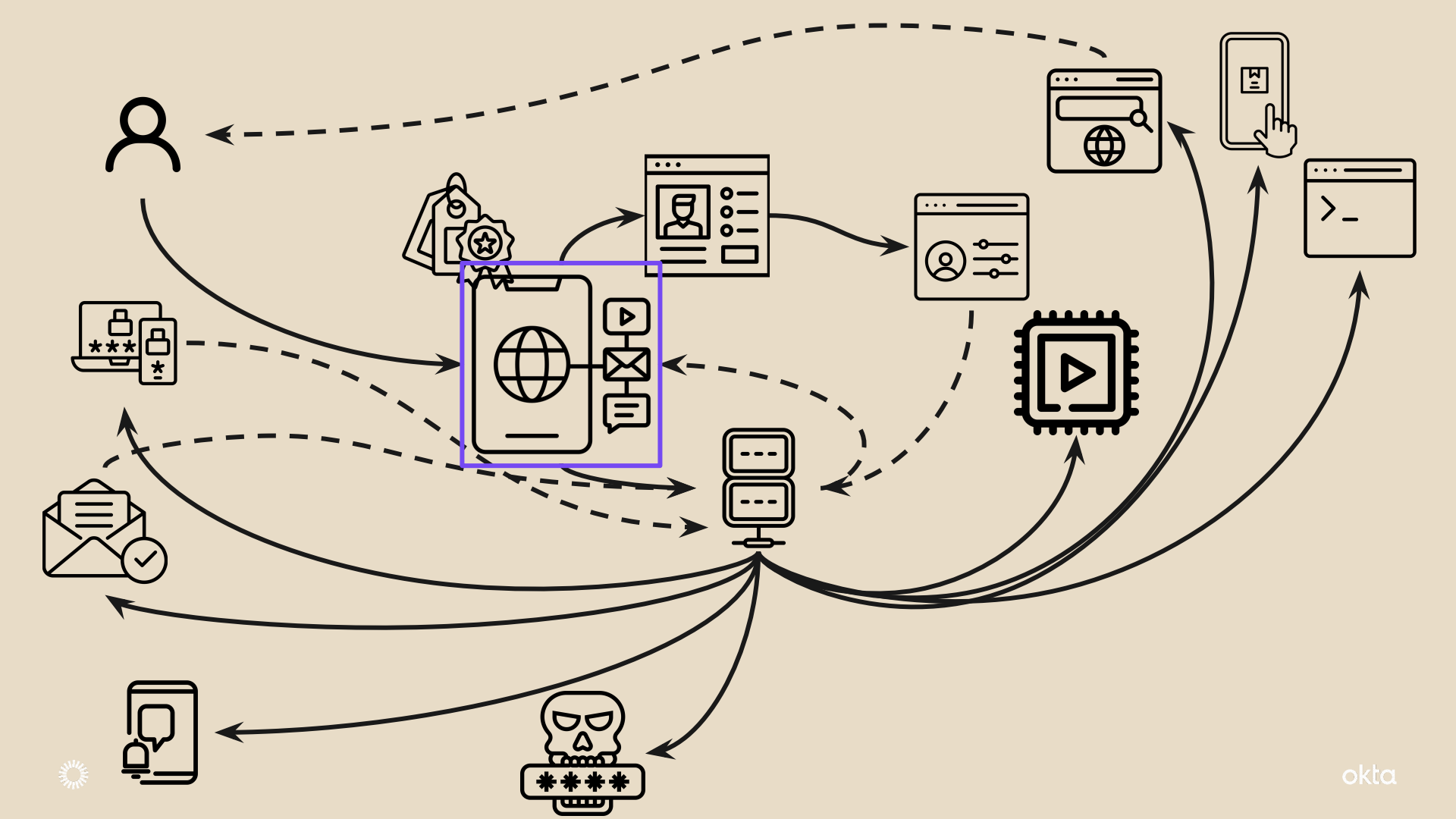
okta

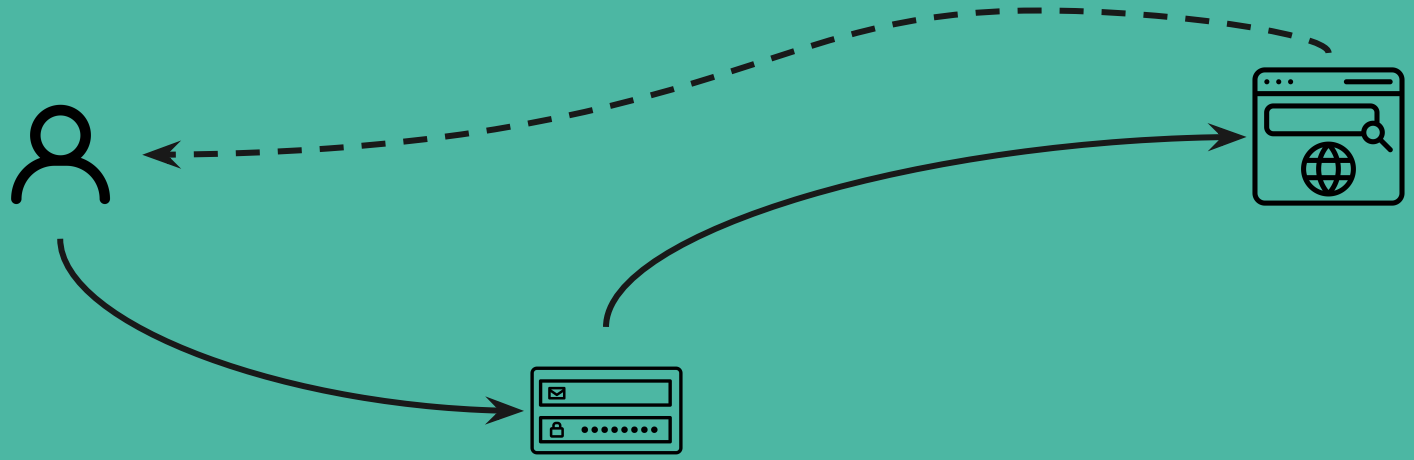

okta


okta


okta


okta

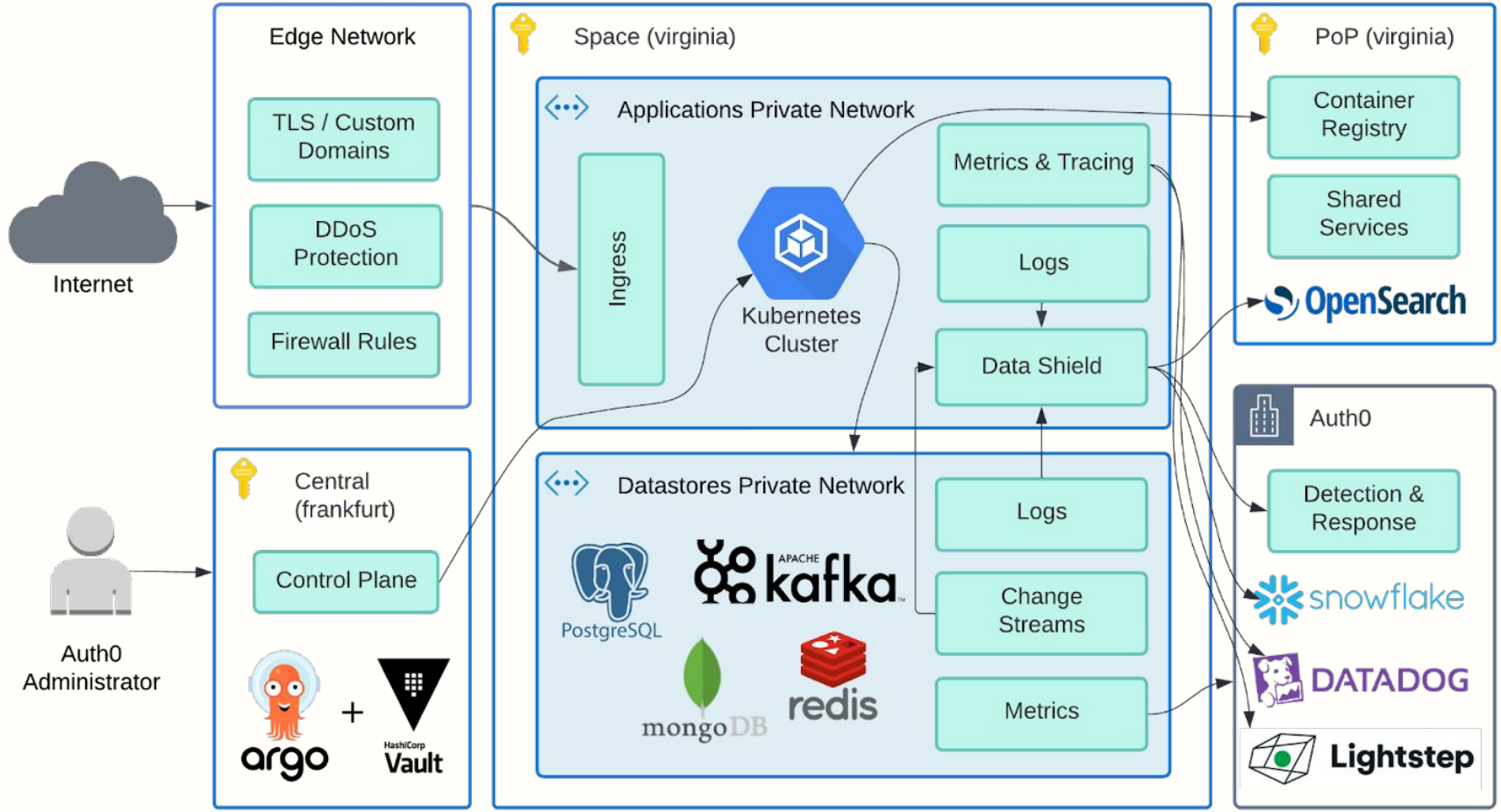

okta



Pretty simple, right?

Auth0 by Okta

Operational Challenges and Solutions

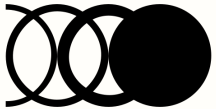


https://auth0.com/blog/the-architect-s-view-of-auth0-s-new-private-cloud-platform/

Platform Complexity

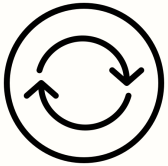
How many toys do we have to play with

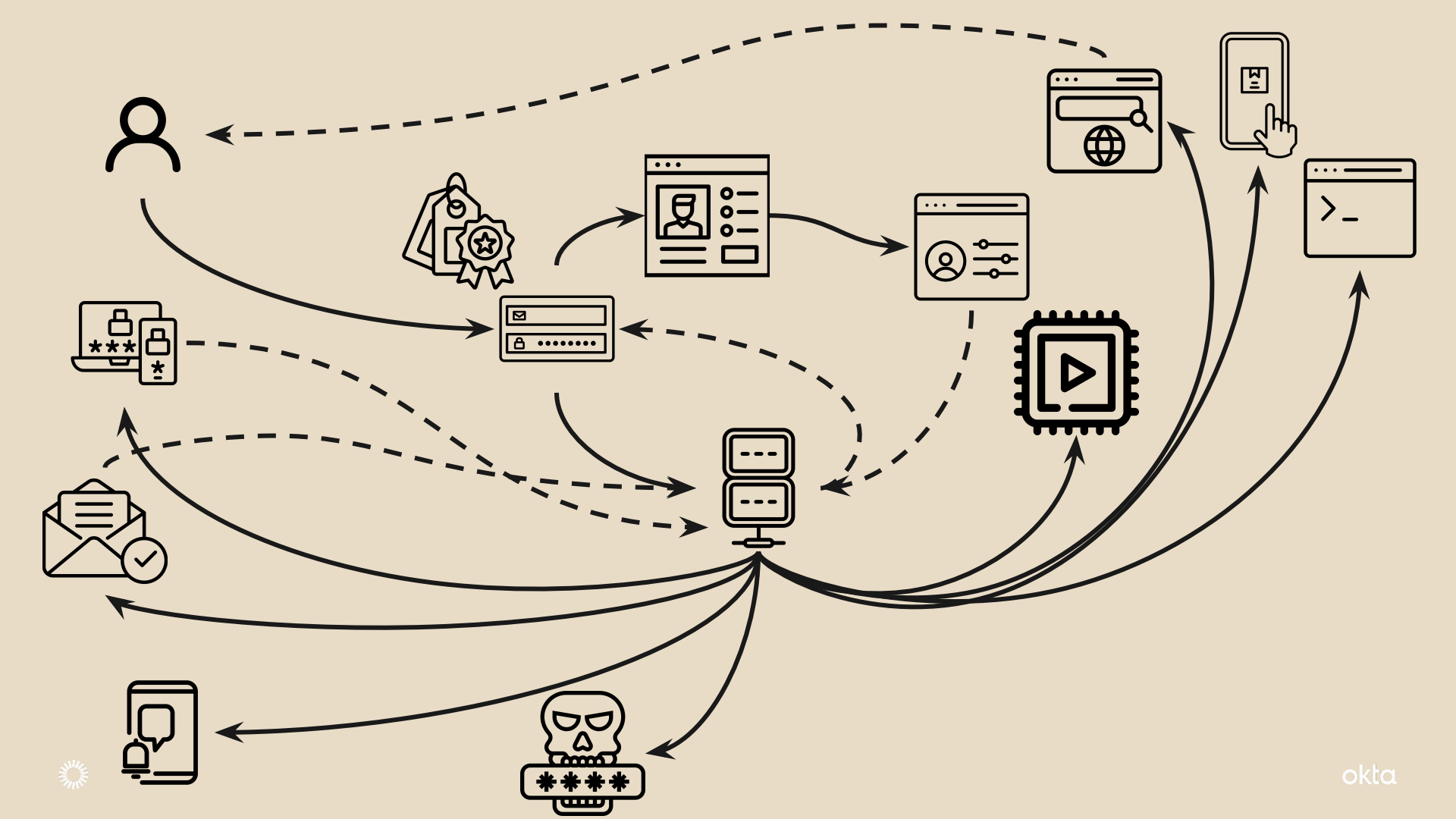
As systems and services evolve new features require different underlying infrastructure - changes on dependencies

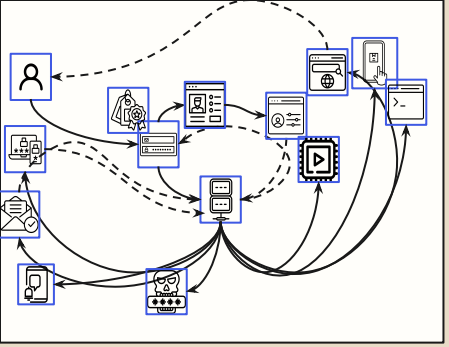


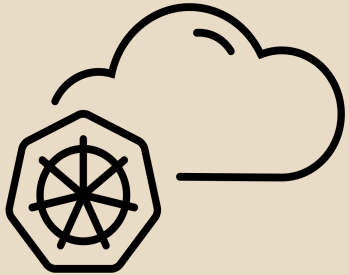
How these different systems interact with each other and what “substrate” systems evolve - monitoring, logs, metrics

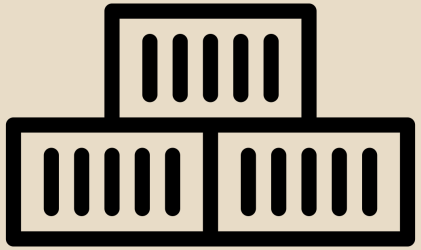
Maintenance, version management, infrastructure releases, and scaling events.

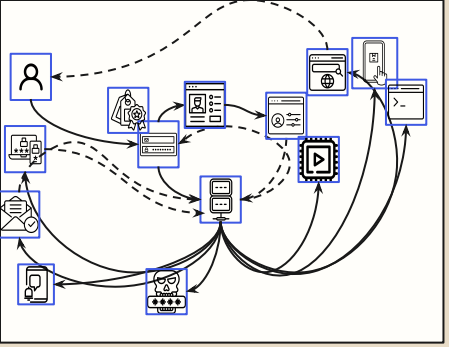


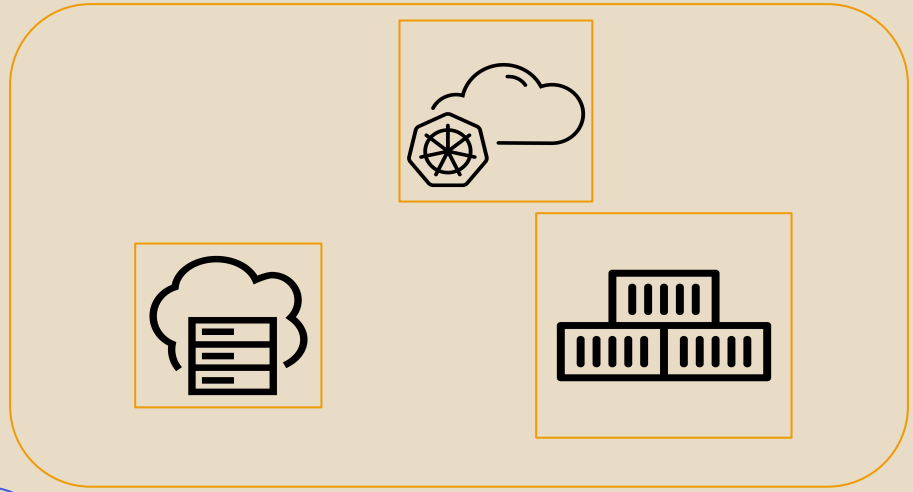

okta

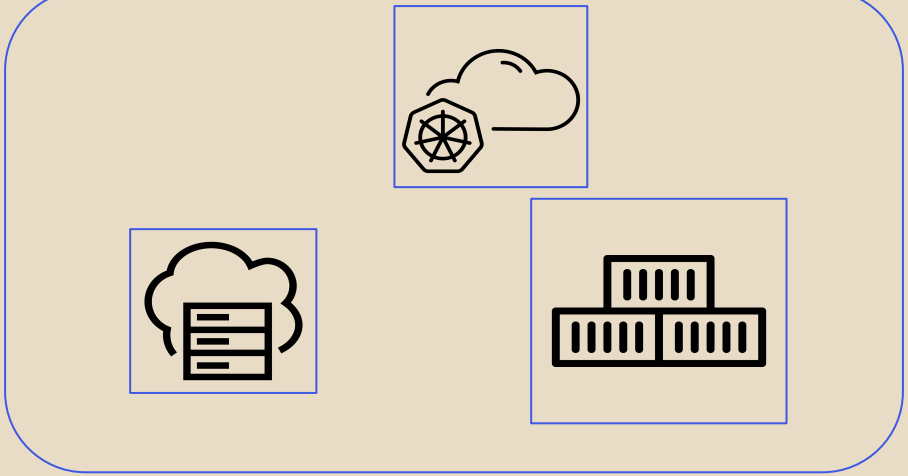


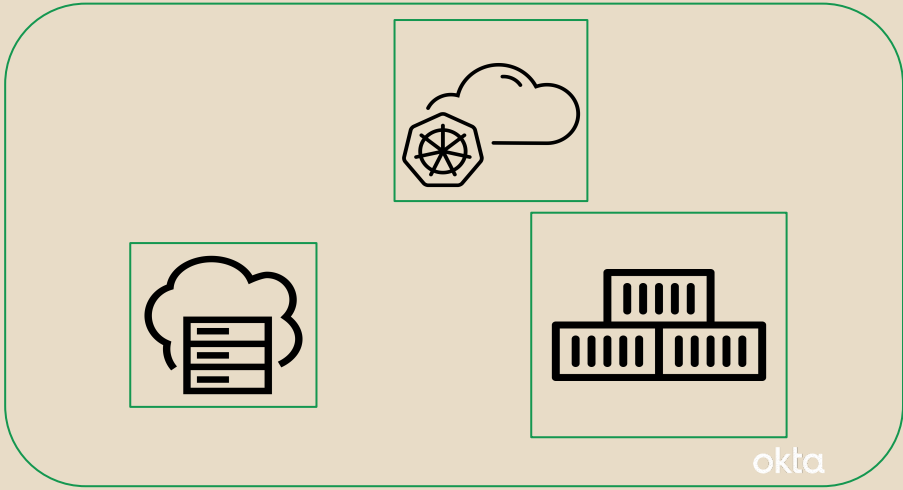












Controlled Operational Challenges

- Logging
- Metrics
- Monitors and Alerts
- Internal Network Configuration
- Release management
- Major version upgrades
- Deployment Failures
- Testing

Not So Under Control Operational Challenges

- Attacks
- Customer configuration creativity
- Outages
- Scale induced miss-calculation
- Cascading failures
- External Network

How do we deal with those challenges?

okta

Solutions for Operational Challenges

We can build that

Predict

System and Load tests

Chaos and **Failure Injection** scenarios

Adjust **auto-scaling** policies

Expect **ridiculous** scenarios

Divert

Rate Limits

Load Shedding

Edge Throttling

Sample logging and metrics

Secondary node reads (eventual consistency)

Protect

No direct access to databases

WAF and good Edge traffic load predictors

Prevent **inefficient queries**

Red-black cluster **deployments**

Scale

Replica sets

Multi-AZ

Vertical Scale

Horizontal Scale

Decouple data layer from compute

okta

Auth0 by Okta

https://xkcd.com/705/

Service Releases and Infrastructure Operations

PLATFORM

A Look Back to Our Platform Resiliency and What's Next

Our commitment to delivering a Tier 0 service for our customers

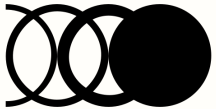
Shiven Ramji
Chief Product Officer

Last Updated On: February 08, 2022

Platform Resiliency

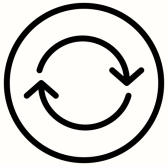
Any weather proof

As we evolve our services we need to keep our fault tolerance and resiliency high



As we grow in terms of customer load and are subjected to more demanding scenarios we need to keep the system stable and reliable

As we deploy and scale up our systems our customers should not be affected by any of such movements



POP QUIZ!

What is the most *common* attribute of a resilient system?

POP QUIZ!

What is the most *common* attribute of a resilient system?

Fault Tolerant

POP QUIZ!

What is the most *common* attribute of a resilient system?

Fault Tolerant

Redundant

okta

POP QUIZ!

What is the most *common* attribute of a resilient system?

Fault Tolerant	Redundant	Scalable

POP QUIZ!

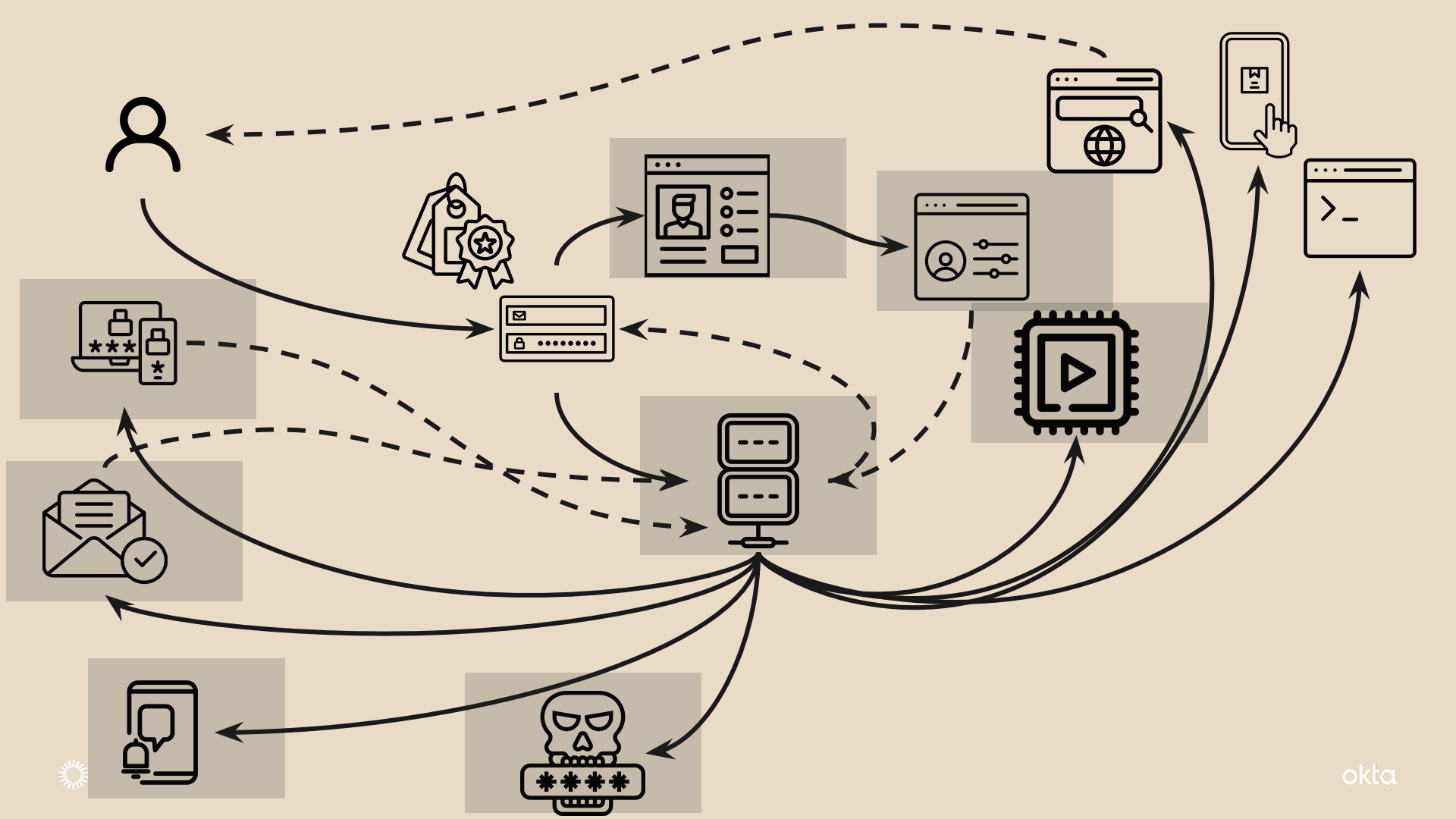
What is the most *common* attribute of a resilient system?

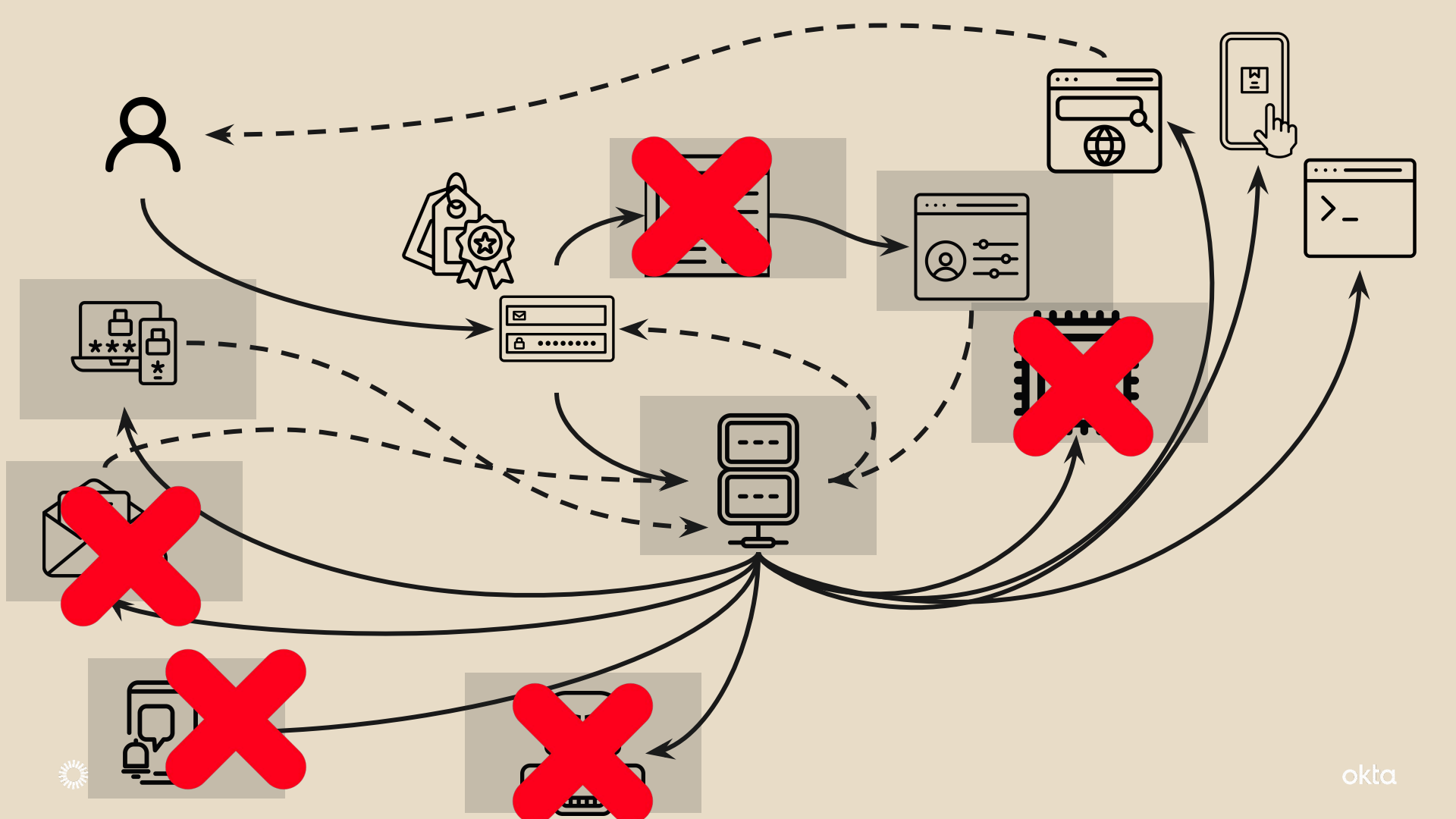
EXPENSIVE $$$

Resilient Platform

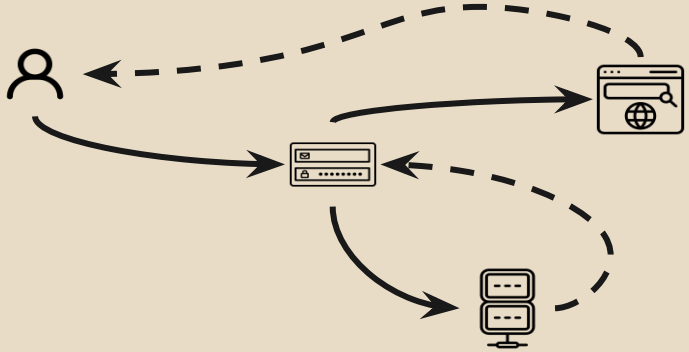
Needs to be all of these things, all at the same time!

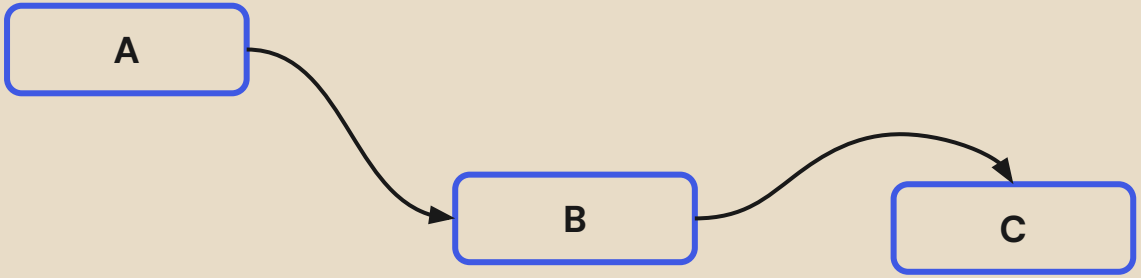
Fault Tolerant	Redundant	Scalable


okta

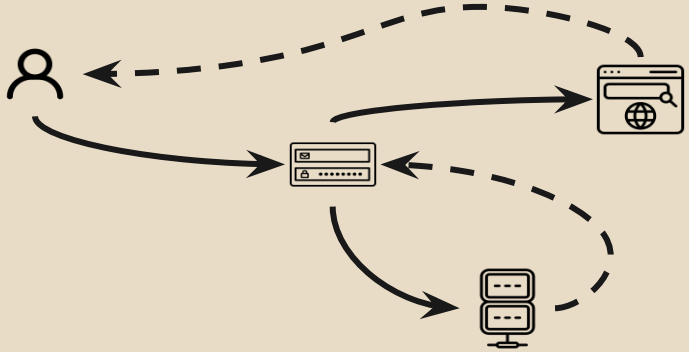

okta

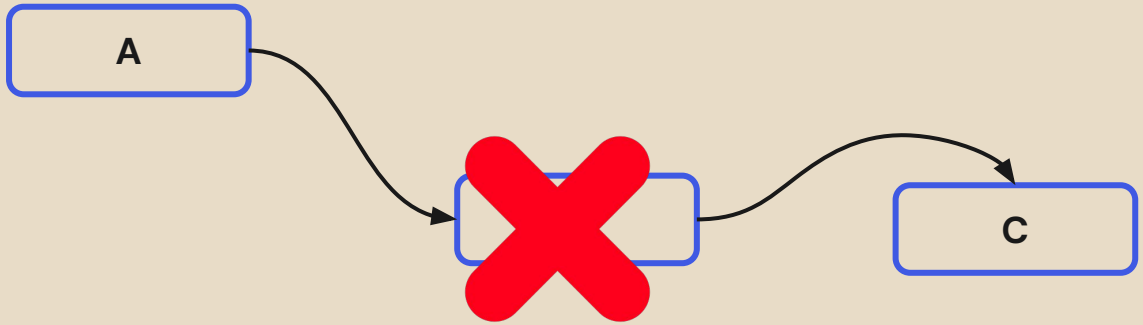
Service Dependencies



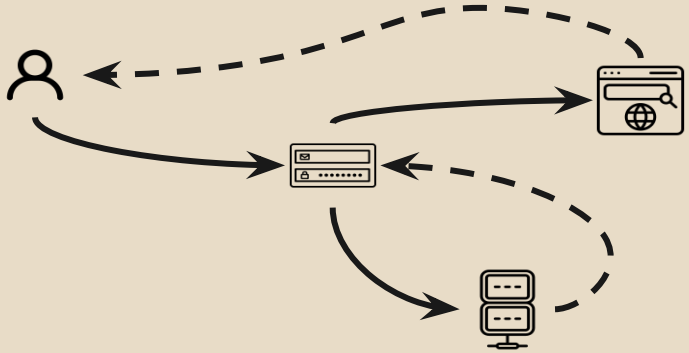


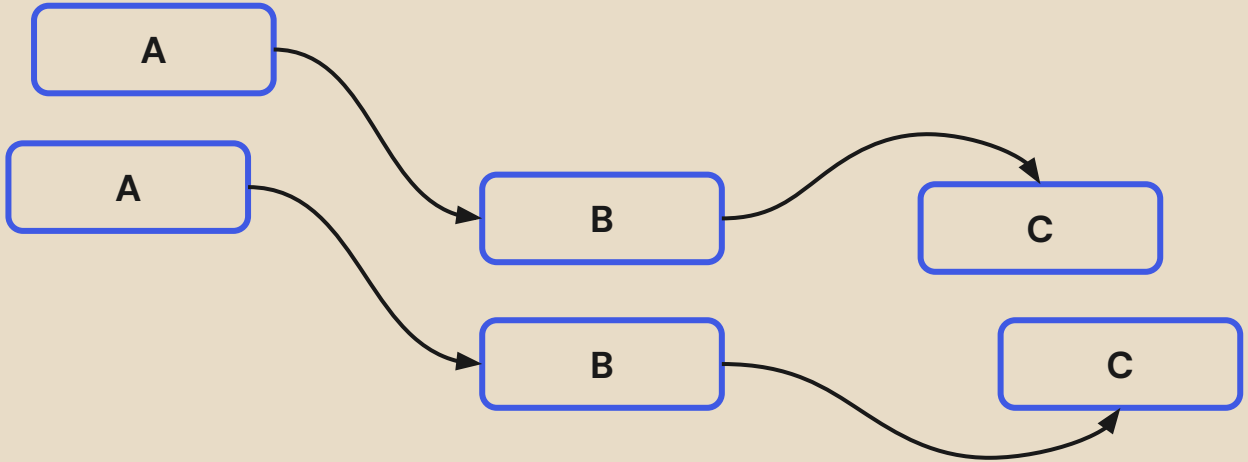
Service Dependencies



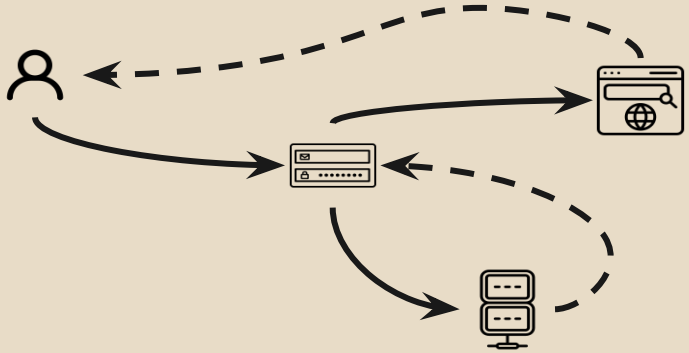


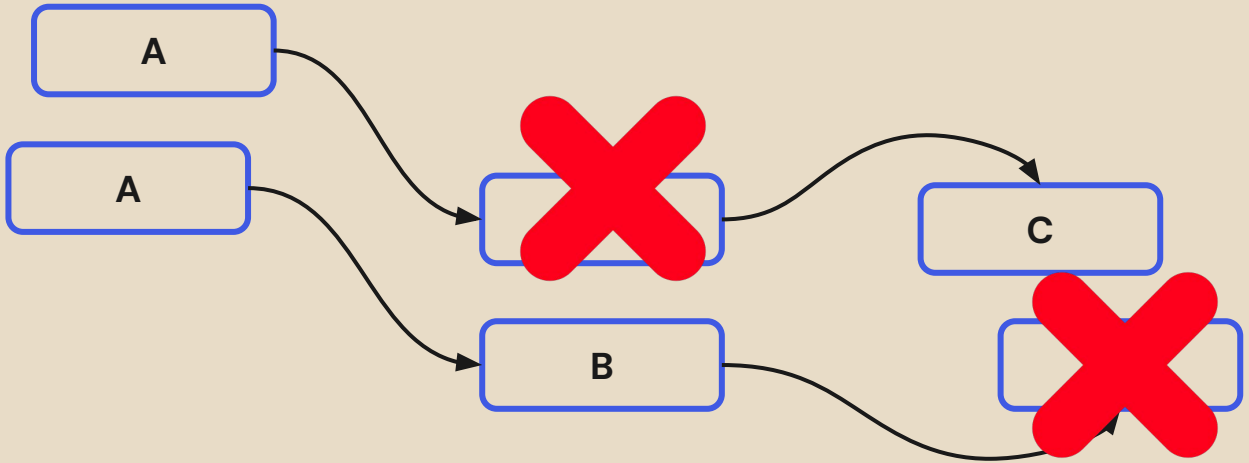
Service Dependencies



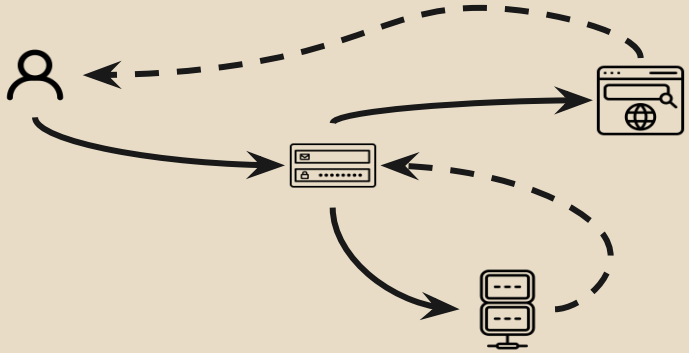


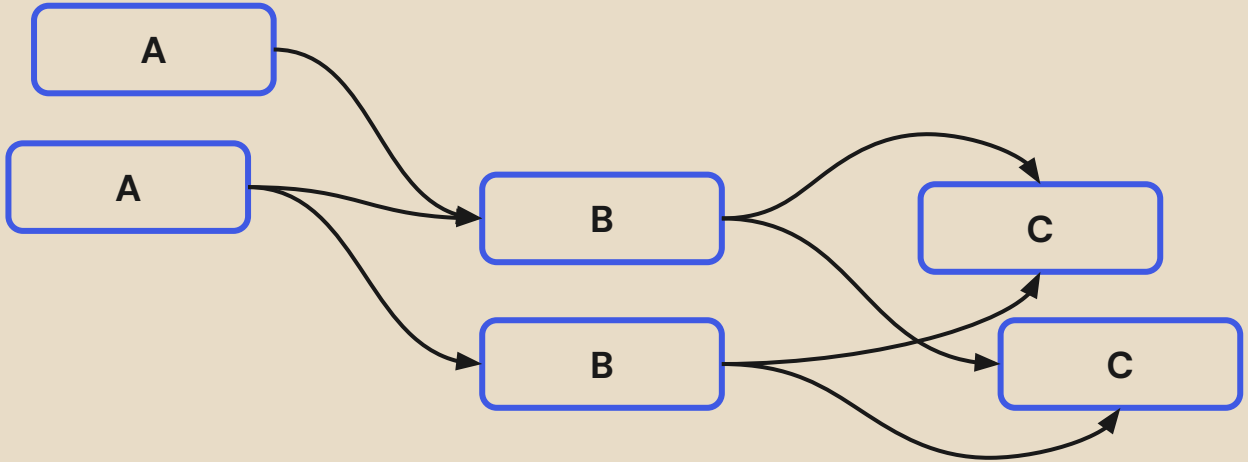
Service Dependencies



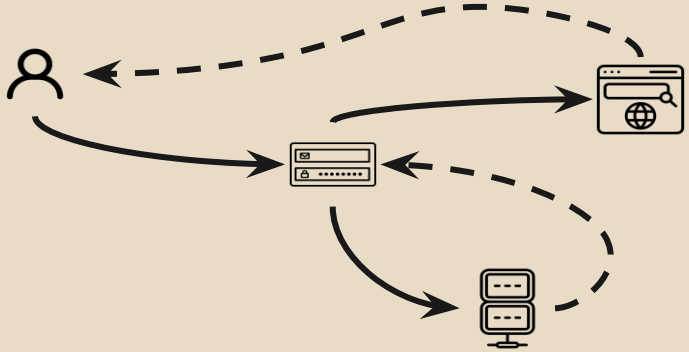


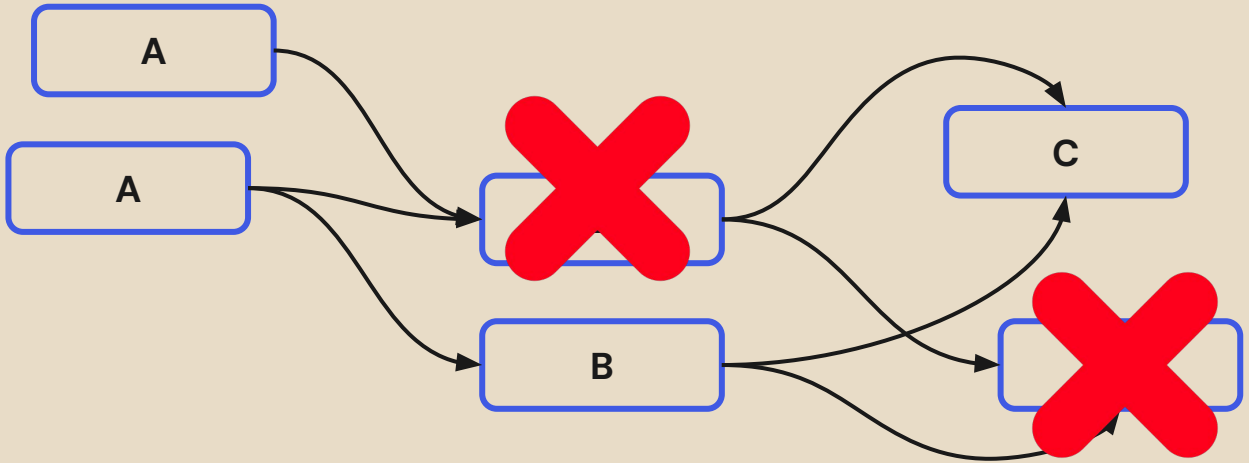
Service Dependencies



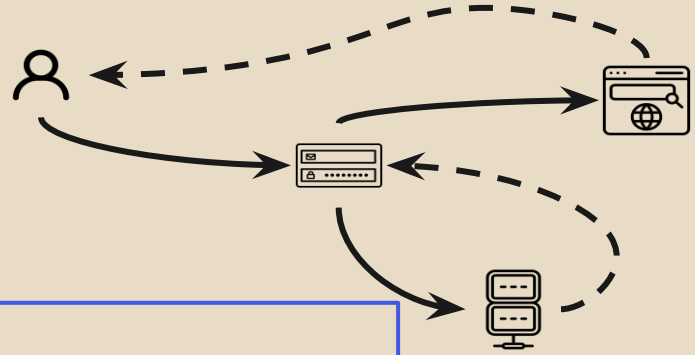


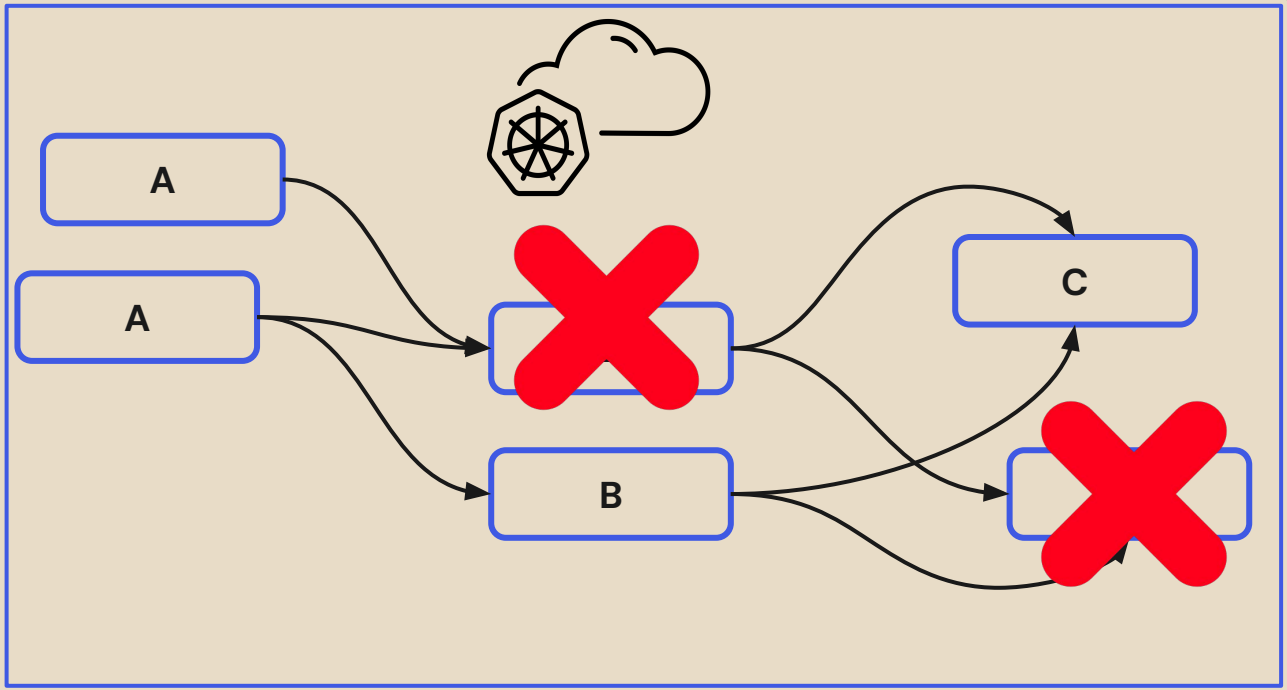
Service Dependencies





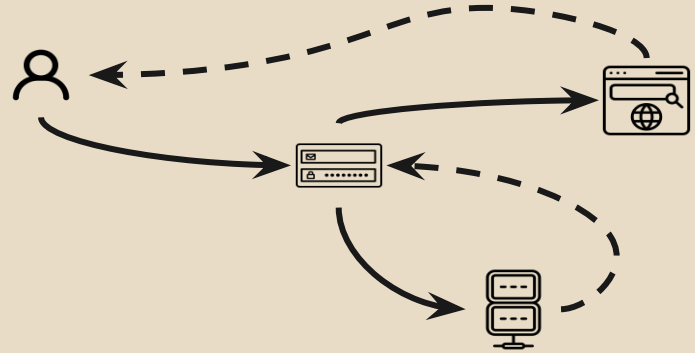
Service Dependencies

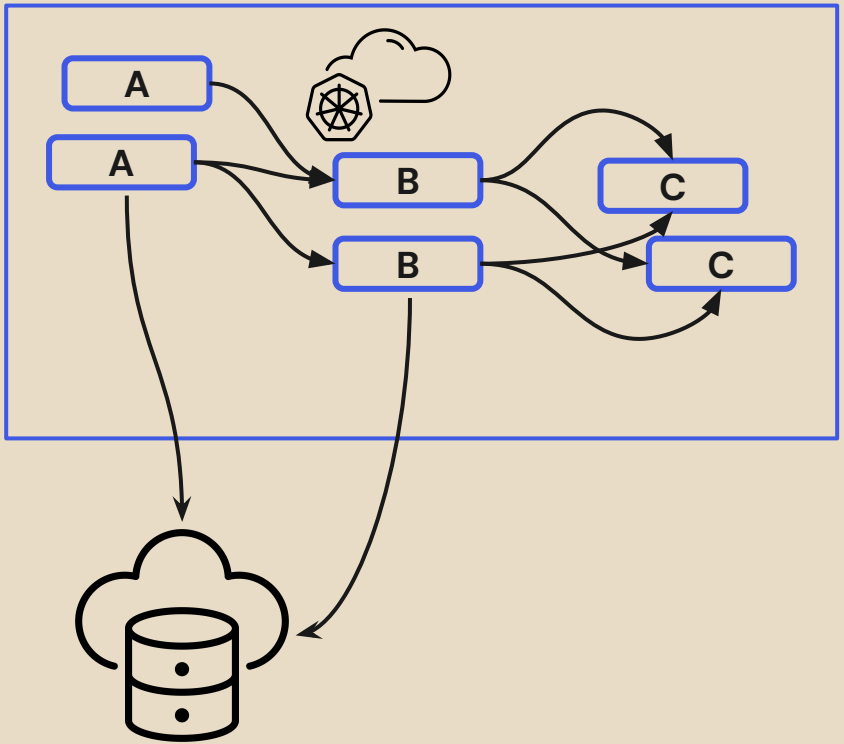




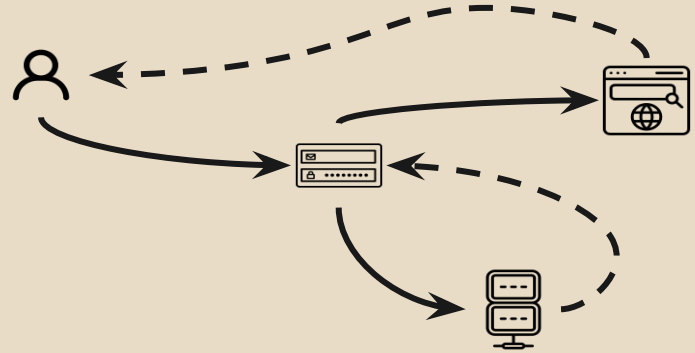
That’s all very nice, but what about external dependencies or IaaS services?

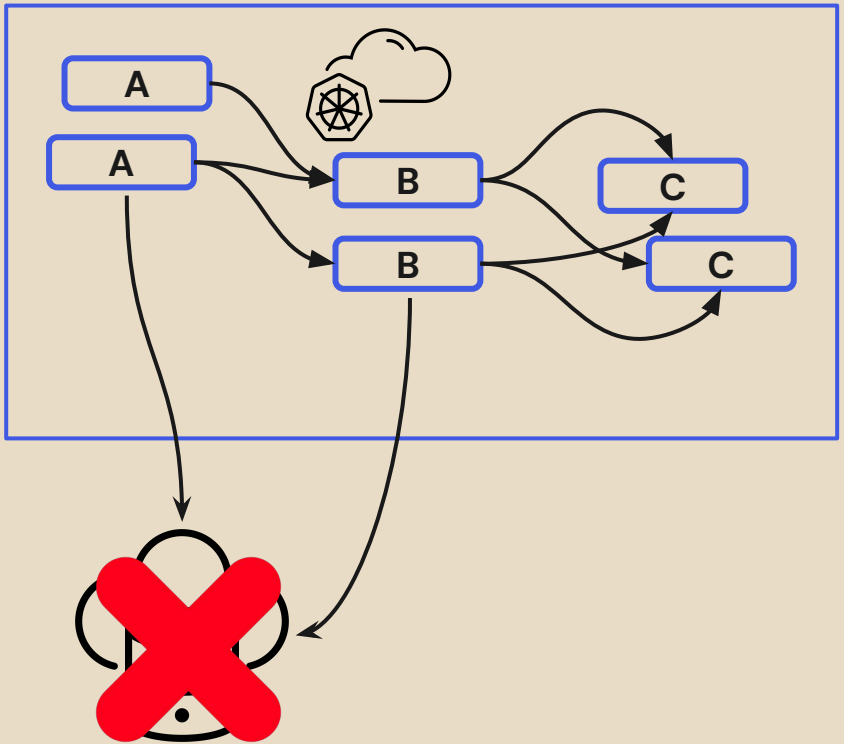
Service Dependencies



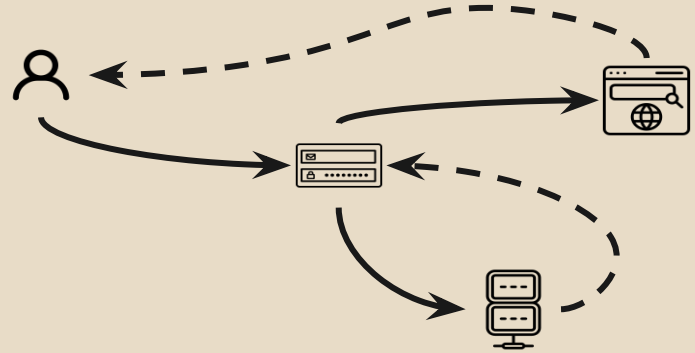


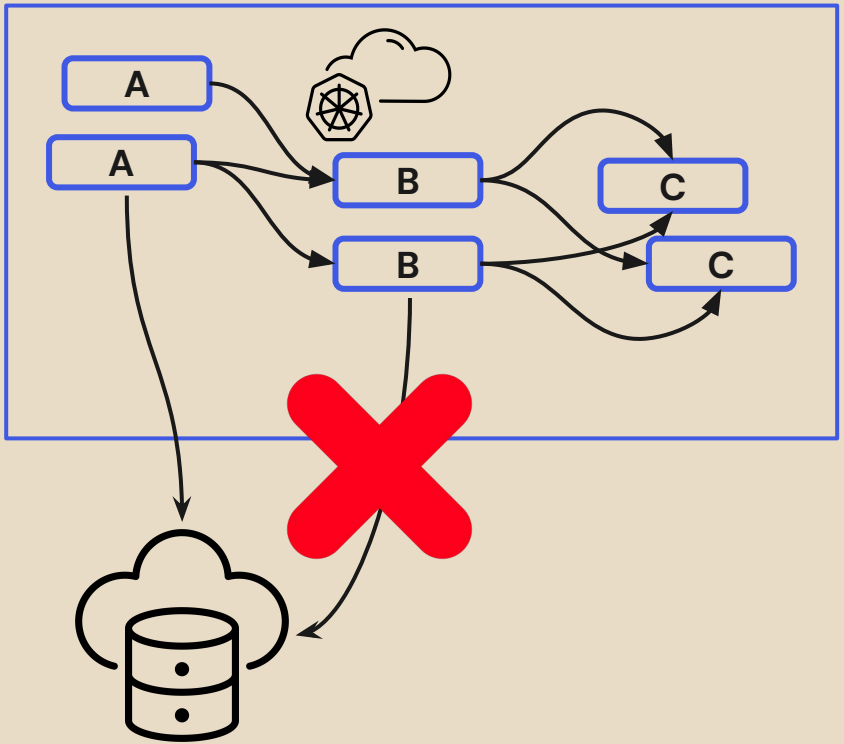
Service Dependencies



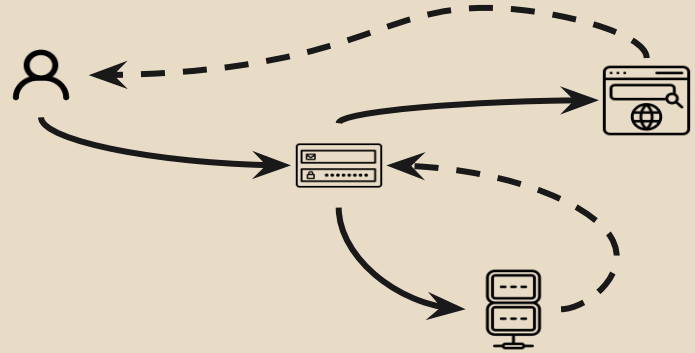


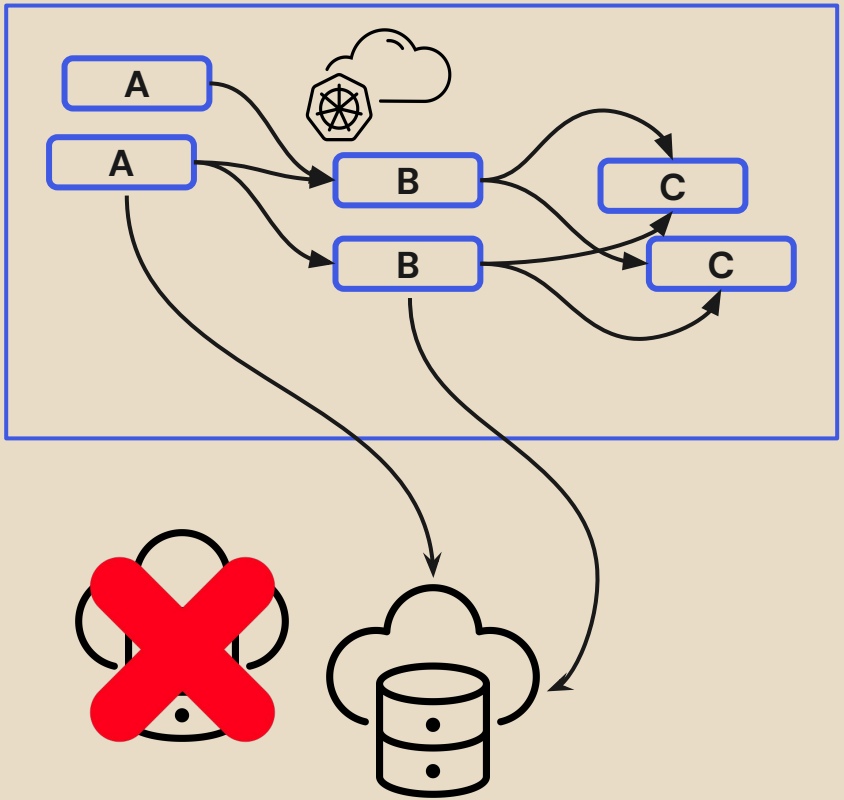
Service Dependencies



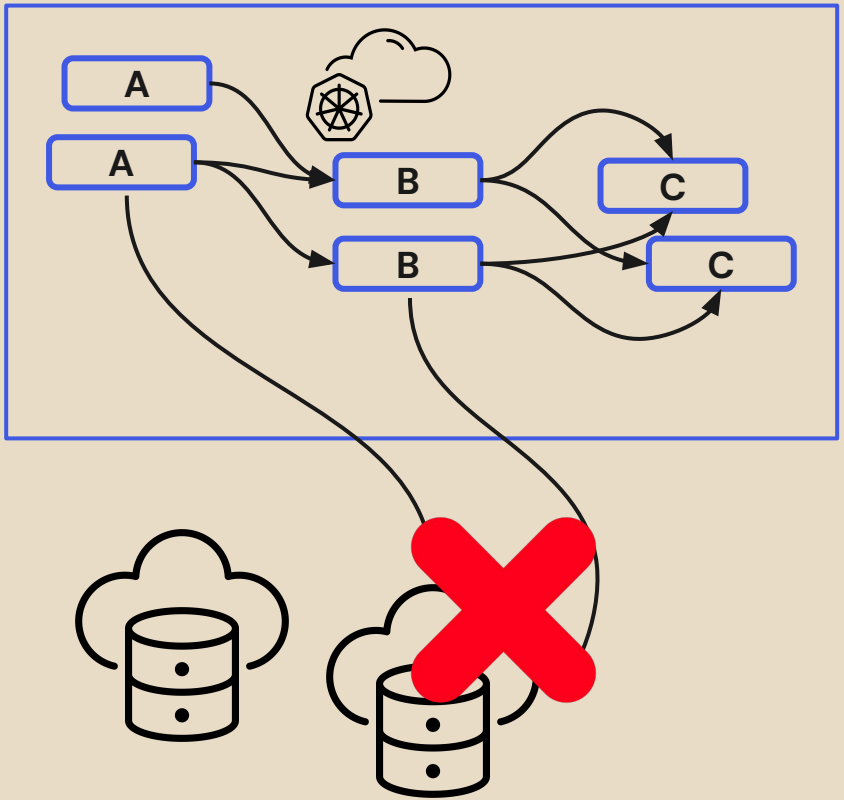


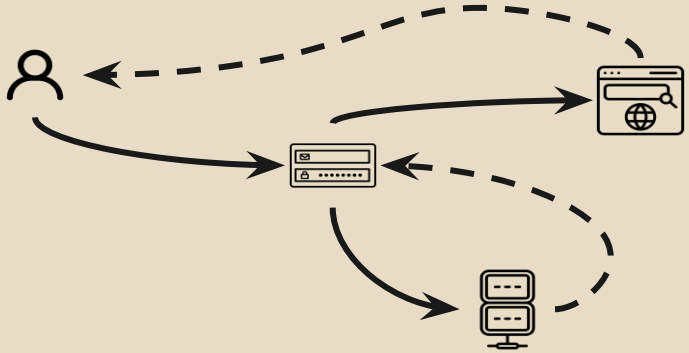
Service Dependencies





Service Dependencies





Degraded Mode

Degraded Mode is the ability your service will have to operate in a reduced capacity

okta

Degraded Mode is the ability your service will have to operate in a reduced capacity

Read Only	Longer extended latency	Reduced set of features

Why building for a **Degraded Mode** is important?

Why building for a **Degraded Mode** is important?

- **Release resources for faster recovery**
- **Faster phased regional failovers**
- **Allow external services to recover**
- **Allow minimal feature set availability**

okta

Auth0 by Okta

Database Management at Scale

Laundry List of Database Problems at Scale

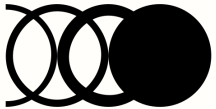
- Bad indexes
- No indexes
- Bad schema migration
- Locking ALTER TABLE
- AUTO-VACUUM
- Extensions OOM
- Manual scripts
- Postgres Major Version Upgrades
- Self-served DDoS
- Cache fallback DoS
- TRIGGERS
- ……

okta

Database Resiliency

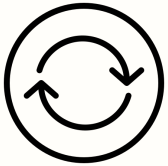
The joy of databases

New versions, new instruction set / api, different topologies and instance sizes vary over time



Databases are not immune to your traffic and need to be protected of it

Database schema evolves: new fields, new indexes, new tables, different cardinality of column values will interact with your systems



okta

Database problems manifest themselves at scale in unexpected ways

Simple Harmless Query

Right?

```
GET http://myservice.com/api/v2/group/hello/users?page=1&limit=50&range=adults
```

```
SELECT a, b, c
FROM users
WHERE a = $1
AND b between $2 AND $3
ORDER BY c DESCENDING
LIMIT $4
SKIP $5
```

Simple Harmless Query

Right??

```
GET http://myservice.com/api/v2/group/hello/users?page=1&limit=50&range=adults
```

```
SELECT a, b, c
FROM users
WHERE a = $1
AND b between $2 AND $3
ORDER BY c DESCENDING
LIMIT $4
SKIP $5
```

```
{
  results: [
    {user: …},
    {user: …}
  ],
  total: 100,
  links: {...}
}
```

Simple Harmless Query

Sure!

```
GET http://myservice.com/api/v2/group/hello/users?page=1&limit=50&range=adults
```

```
SELECT a, b, c
FROM users
WHERE a = $1
AND b between $2 AND $3
ORDER BY c DESCENDING
LIMIT $4
SKIP $5
```

```
{
  results: [
    {user: …},
    {user: …}
  ],
  total: 100,
  links: {...}
}
```

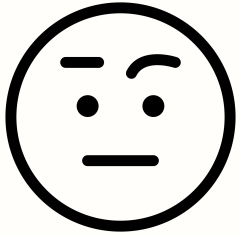
okta

Simple Harmless Query

Wait a second!

```
GET http://myservice.com/api/v2/group/hello/users?page=1&limit=50&range=adults
```

```
SELECT a, b, c
FROM users
WHERE a = $1
AND b between $2 AND $3
ORDER BY c DESCENDING
LIMIT $4
SKIP $5
```



```
{
  results: [
    {user: …},
    {user: …}
  ],
  total: 100,
  links: {...}
}
```

Simple Harmless Query

RUN RUN RUN!!!!

```
GET http://myservice.com/api/v2/group/hello/users?page=1&limit=50&range=adults
```

```
with results AS (
   SELECT count(1) sum
   FROM users
   WHERE a = $1
   AND b between $2 AND $ 3
)
SELECT a, b, c, results.sum
FROM users
WHERE a = $1 AND b between $2 AND $3
ORDER BY c DESCENDING
LIMIT $4
SKIP $5
```

```
{
  results: [
    {user: …},
    {user: …}
  ],
  total: 100,
  links: {...}
}
```

okta

Operation Execution Time ▾	Keys Examined	Docs Returned	In Memory Sort
1069	1175200	100	No
971	1307100	100	No
958	1259900	100	No
950	1282100	100	No
949	1302100	100	No
949	1290800	100	No
947	1126200	100	No
932	1261900	100	No

Important things to look out for at scale

Pagination

Decouple API design from Database Queries

Don't allow **SKIP + LIMIT** queries

At scale exact count will always be eventual

Schema Migrations

Postgres default column values are dangerous

Type sensitive applications will have a hard time adjusting to type changes

MongoDB is great to get started with, can be a problem managing schema later - **schema validation FTW**

New Indexes

ESR Rule FTW - Equality, Sort, Range on index field/column order always!

The more indexes you add, the slower your writes might become

Before **adding** indexes, check if there's one you might be able to **remove**

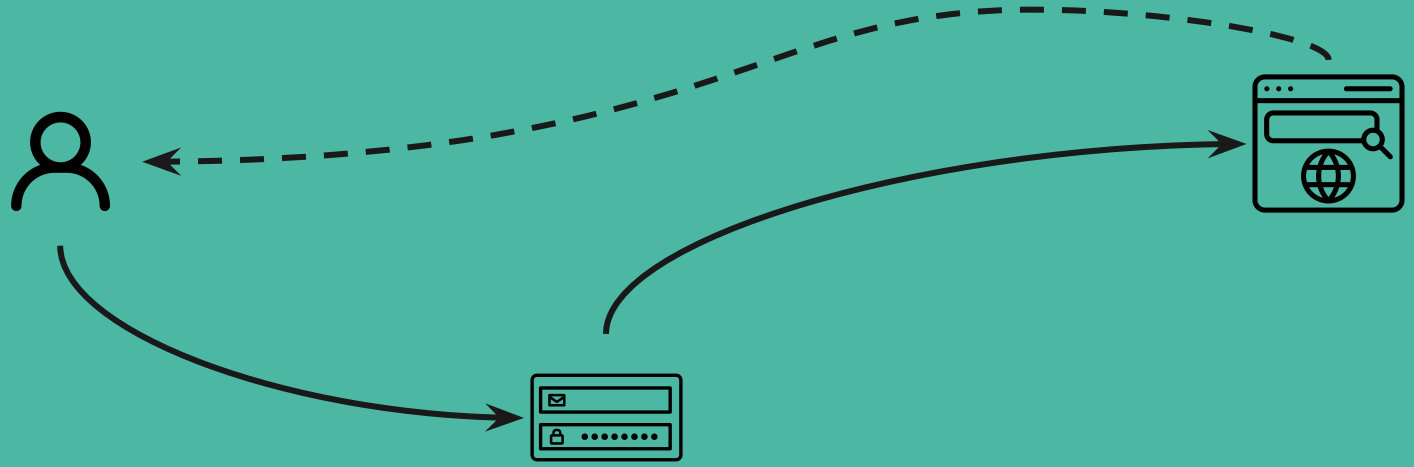
Triggers/ PL-SQL

JUST DON'T DO THEM

okta

okta
The World's Identity Company

Quick Recap



Pretty simple, right?

Resilient Platform

Needs to be all of these things, all at the same time!

Fault Tolerant	Redundant	Scalable

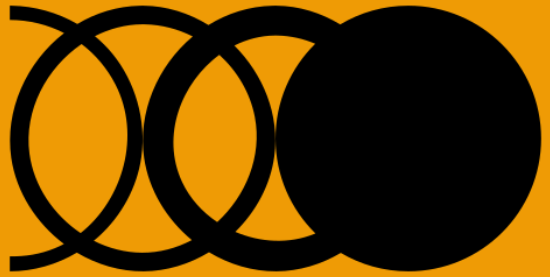
Degraded Mode is the ability your service will have to operate in a reduced capacity

Read Only	**Longer extended latency**	**Reduced set of features**

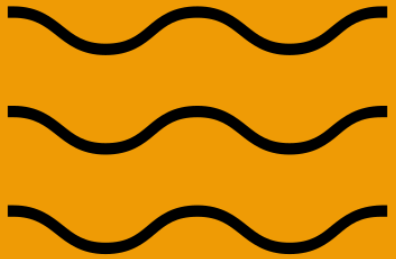
Triggers are devils creation!

Degrees of Freedom

Translational Motion



Vibration



Rotation

Auth0 by Okta

Make login our problem. Not yours.

Authentication

Authorization

Security

Single Sign-On | Adaptive Multi-Factor Authentication | Universal Login | Passwordless | Bot Detection & Prevention | Security Center | Breached Password Detection | Brute Force Protection | FGA

Try if for free

No credit card required.

Special Plans for Startups & Nonprofits

https://a0.to/plg_signup

Plans for Everyone

B2C: your users are consumers

B2B: your users are businesses or a mix of businesses and consumers

Enterprise: Best for production applications that need to scale - **Contact Us**

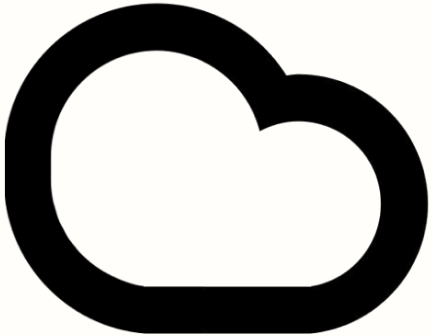
okta

Auth0 by Okta

Q&A

Thank you!

okta



CloudConf

okta