Please

Ask questions through the app

Rate Session

Thank you!





@GOTOber

gotober.com



centro

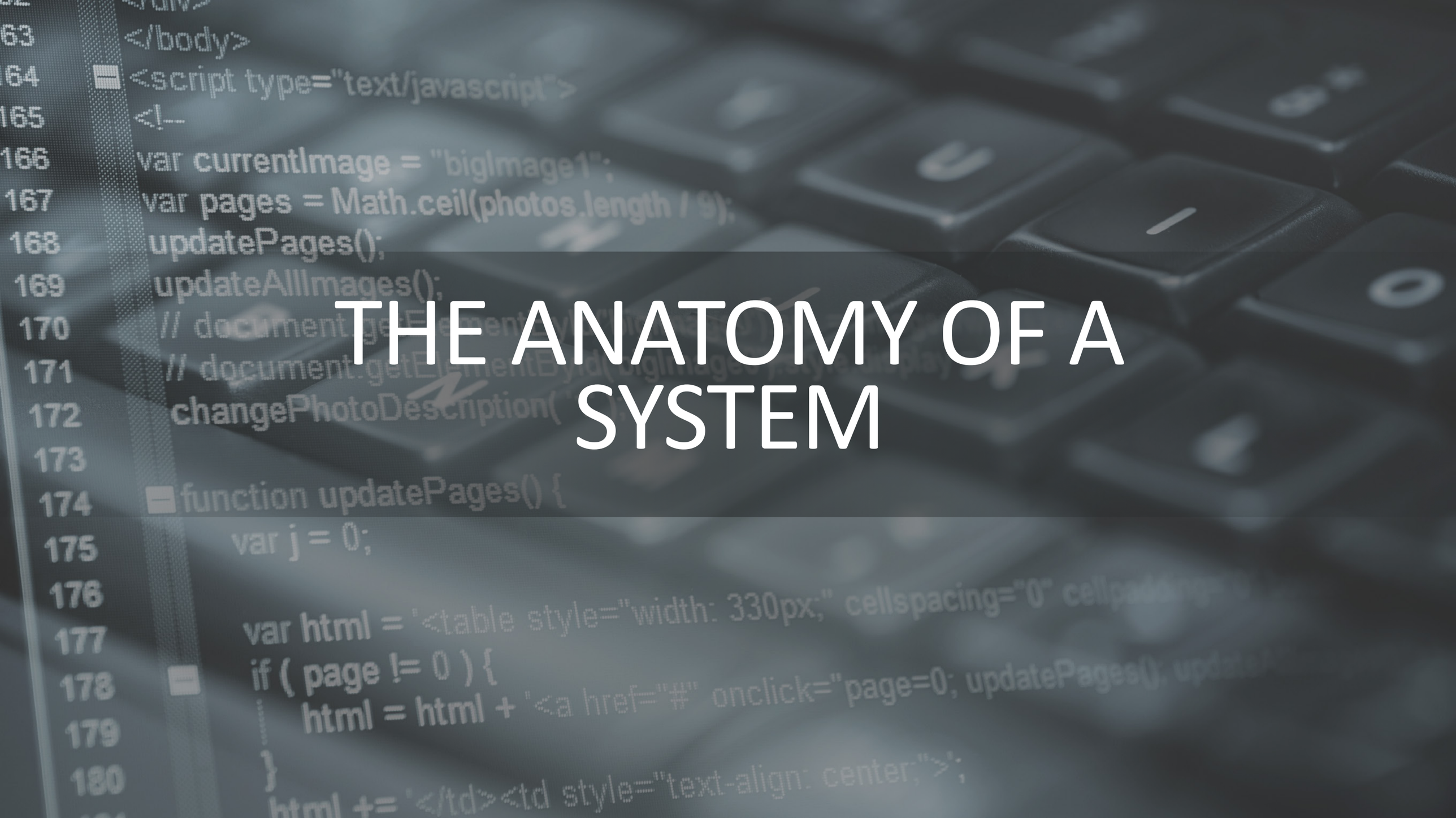
TIERED TRAGEDY: A PEEK INTO FAILURE

A walkthrough of an actual incident and the lessons learned

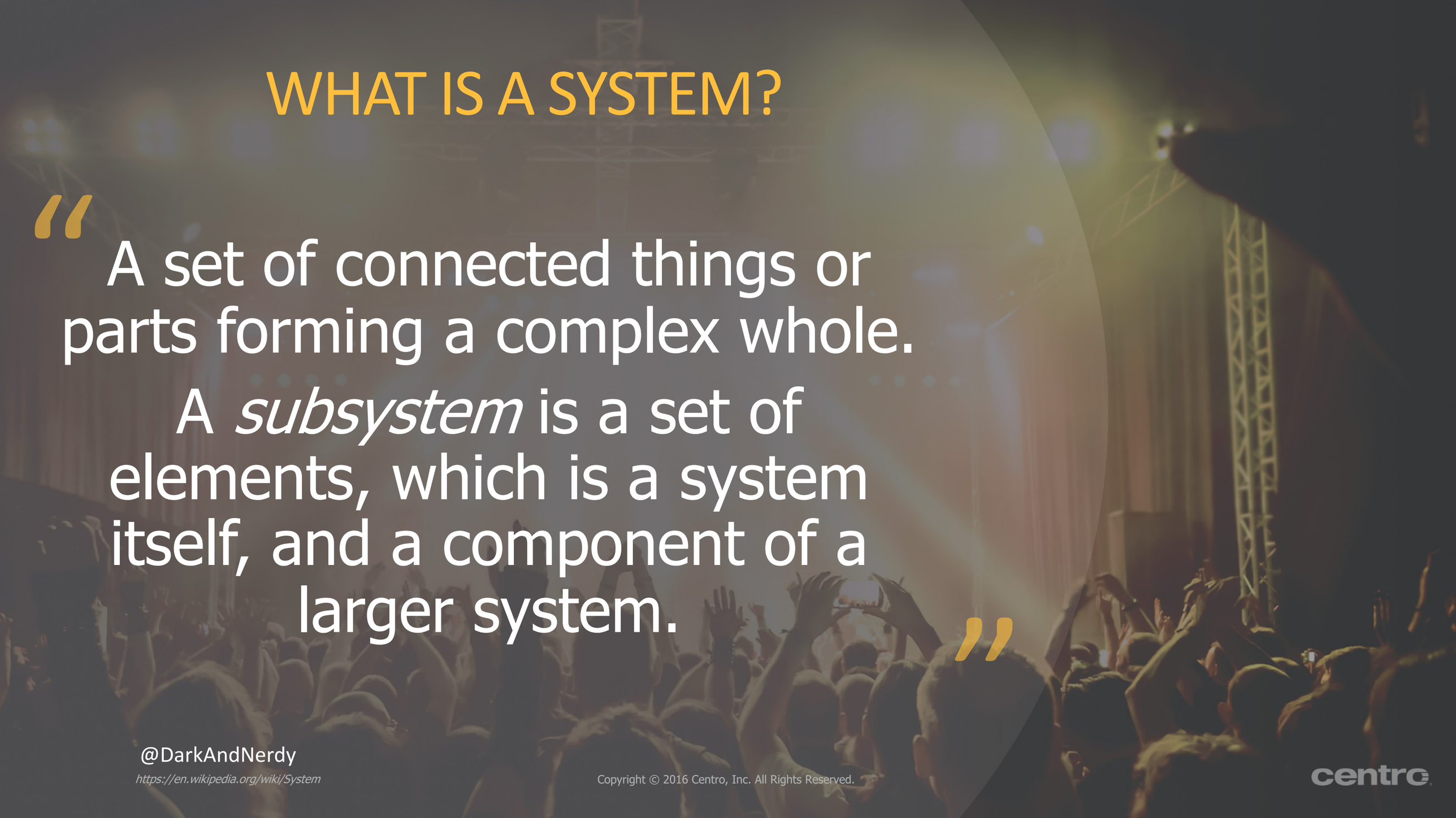


Copyright © 2016 Centro, Inc. All Rights Reserved.

centro


THE ANATOMY OF A SYSTEM

WHAT IS A SYSTEM?



> A set of connected things or parts forming a complex whole.
>
> A *subsystem* is a set of elements, which is a system itself, and a component of a larger system.

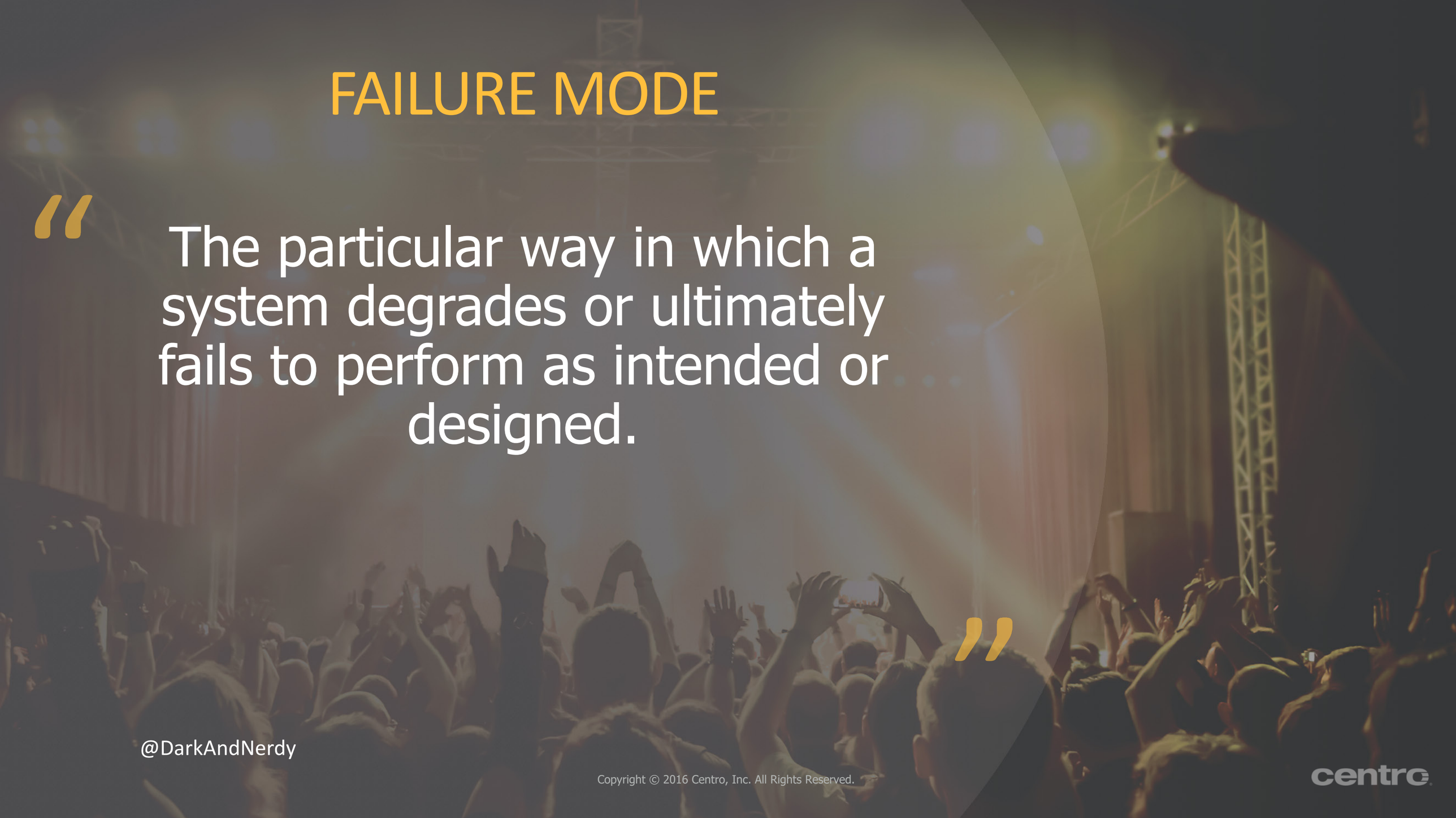
@DarkAndNerdy

https://en.wikipedia.org/wiki/System

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro

FAILURE MODE



> The particular way in which a system degrades or ultimately fails to perform as intended or designed.

@DarkAndNerdy

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro

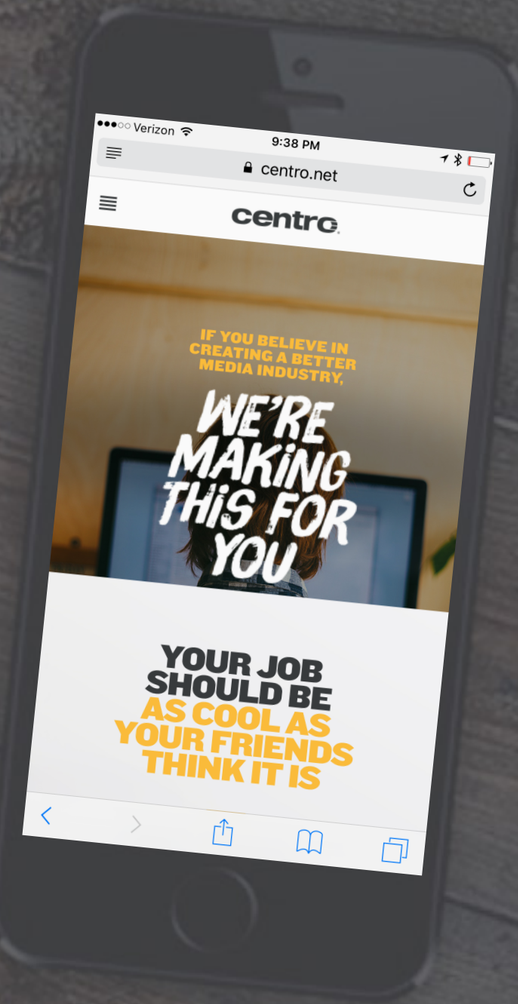


Jeff Smith

Centro

Director, Production Operations

- @DarkAndNerdy
- jeff.smith@centro.net
- www.allthingsdork.com



centro

centro

Founded in 2001

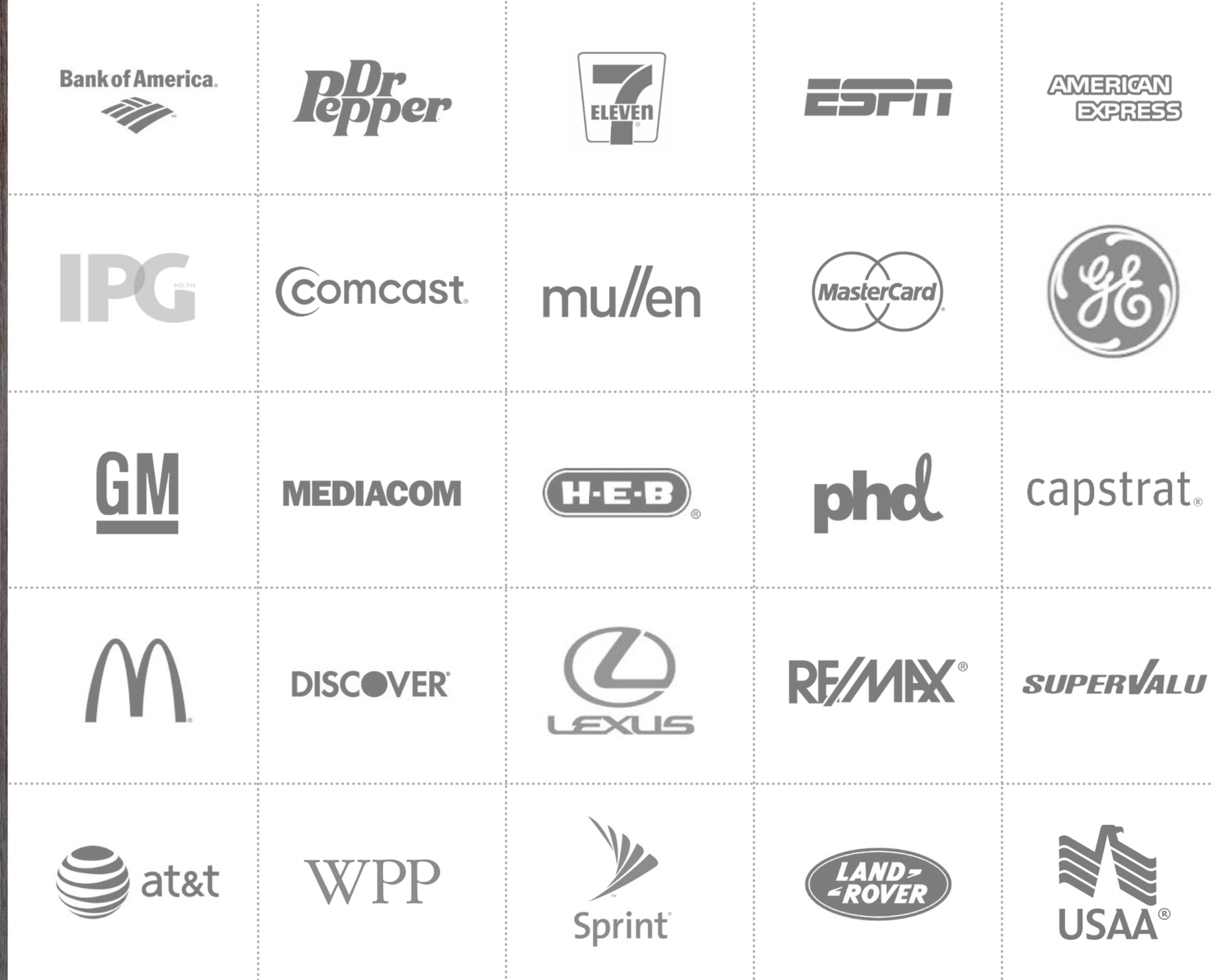
Headquartered in Chicago

600+ Employees

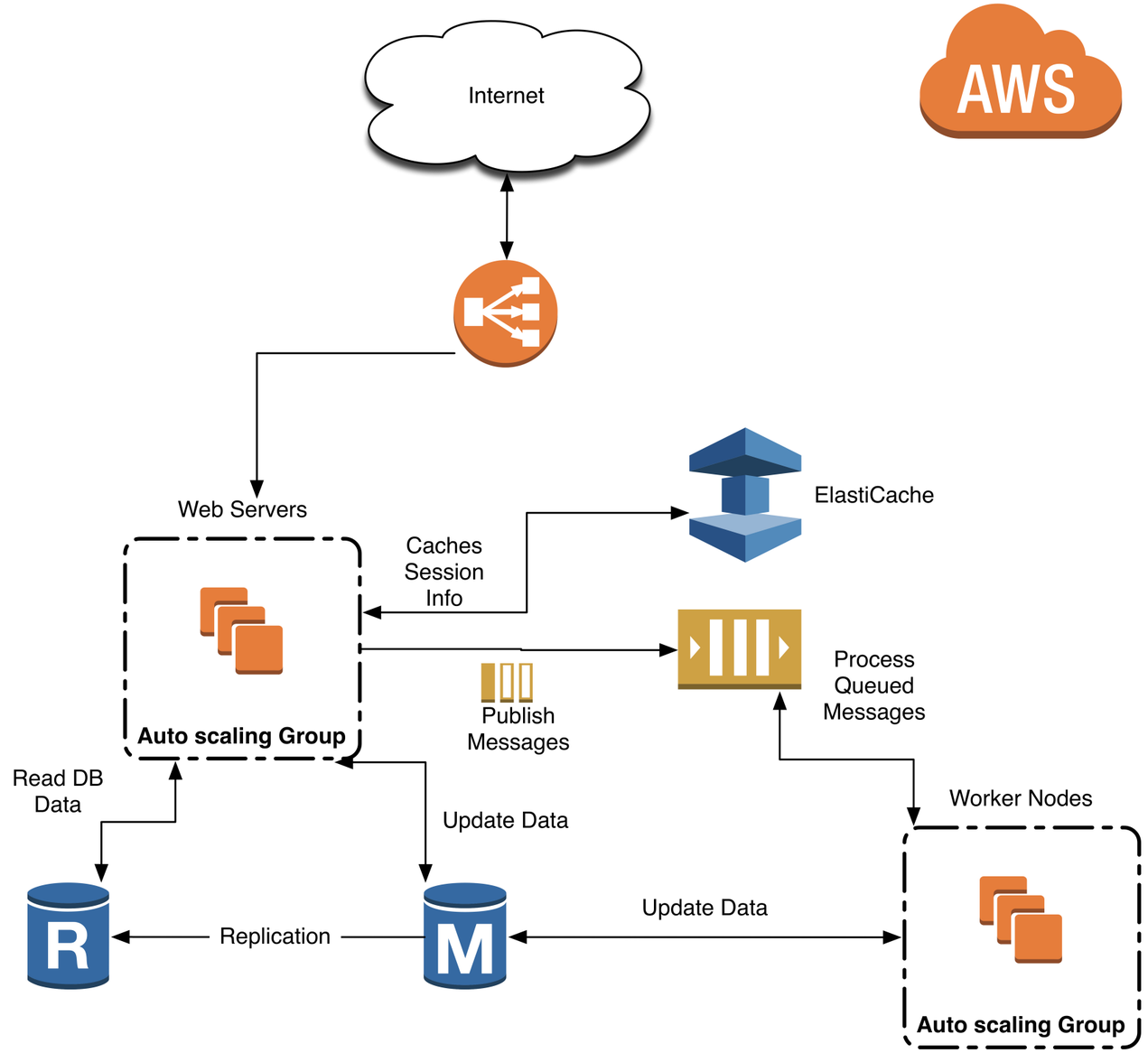
2000+
Agencies

6000+
Brands

200,000+
Campaigns




A FAILURE WALK THROUGH



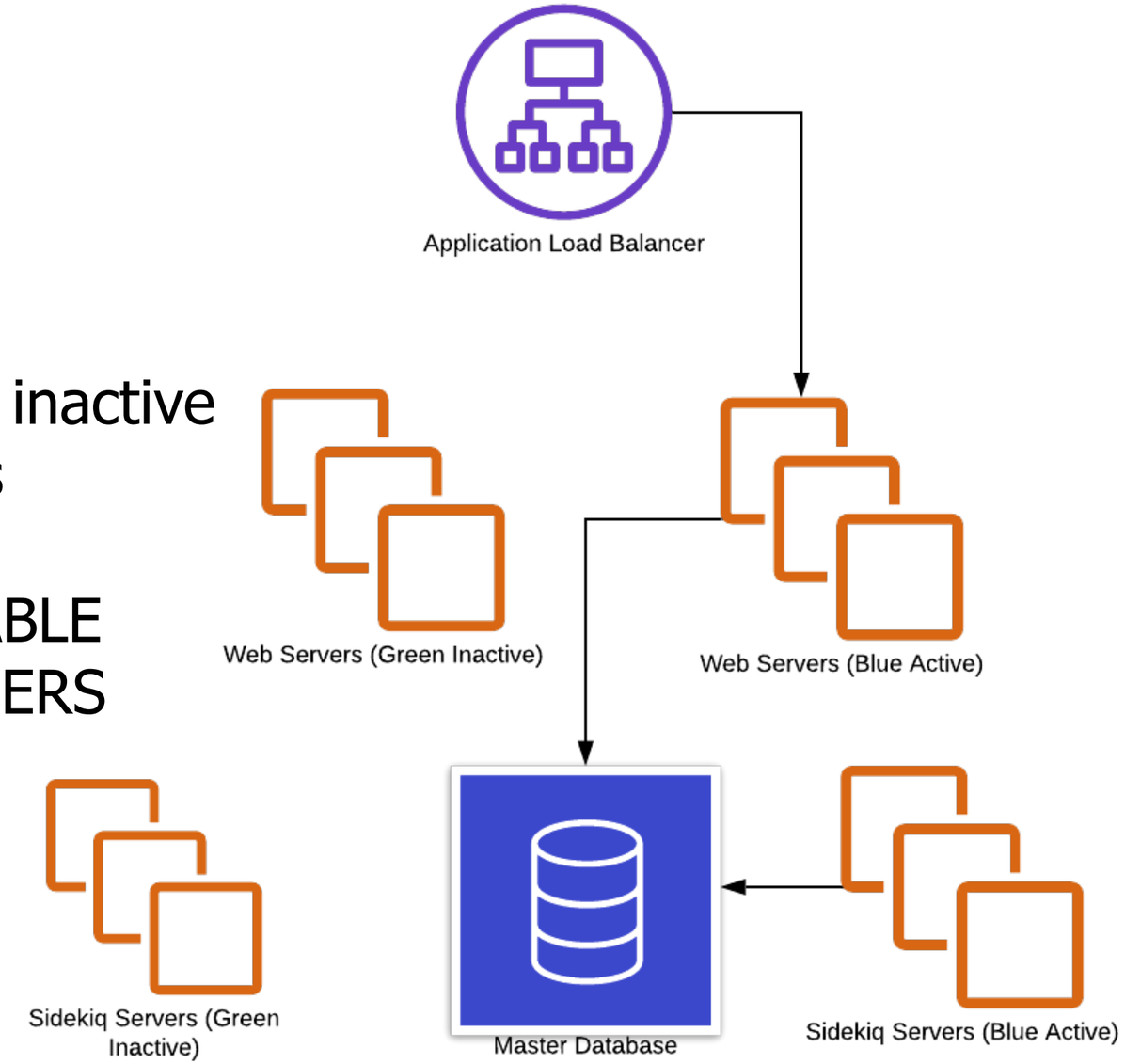
High Level System Diagram

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro

Migrations on inactive nodes

ALTER TABLE AUTH_USERS



User generated action causing lots of long running transactions

Deployment Process

Copyright © 2016 Centro, Inc. All Rights Reserved.
centro

POSTGRES LOCKING PRIMER



EVERY QUERY GENERATES A LOCK OF SOME TYPE



SOME LOCK TYPES CONFLICT WITH EACH OTHER



CONFLICTING LOCKS GO INTO A QUEUE AND EXECUTE IN ORDER



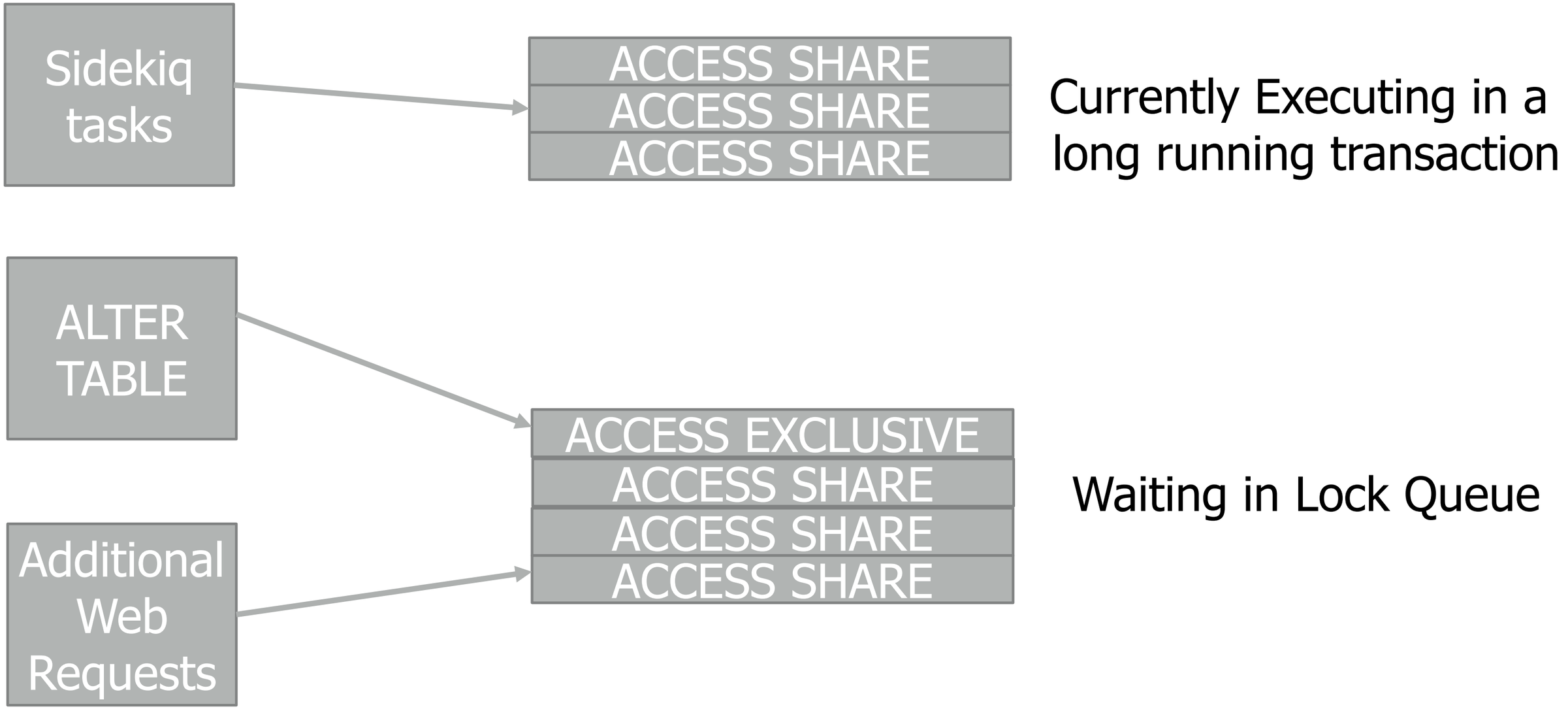
BLOCKING CONFLICTING LOCKS BLOCK SUBSEQUENT LOCKS



LOCKS ARE HELD UNTIL COMMIT OR ROLLBACK

Copyright © 2016 Centro, Inc. All Rights Reserved.

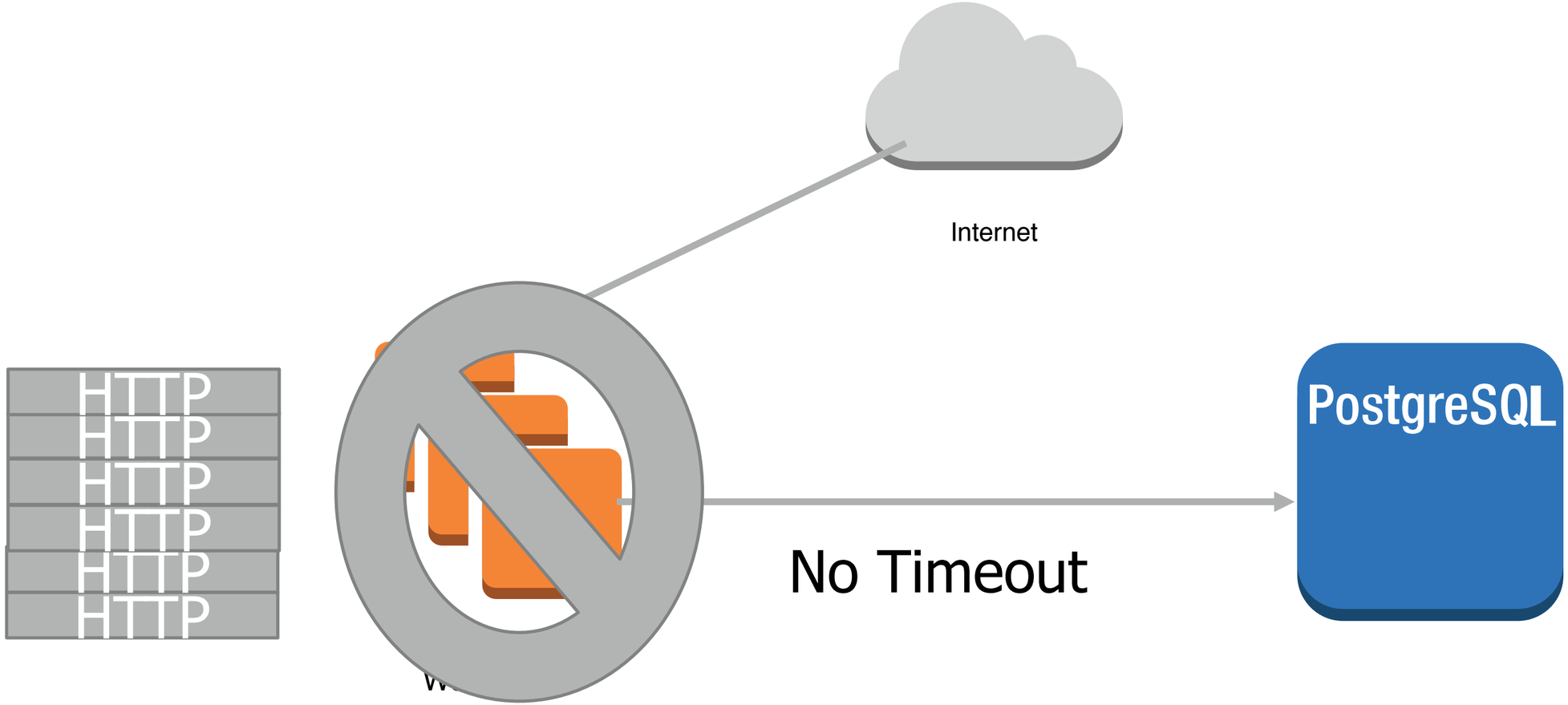
centro



Troubled System

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro



Troubled System

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro

TIMELINE OF EVENTS



Copyright © 2016 Centro, Inc. All Rights Reserved.

centro


MAKING
IMPROVEMENTS

BRIEF RECAP OF ISSUES

Contributing factors to failure

1

Alter Table

DDL Statement on a very busy table

2

Long running transactions

4+ hour long transaction times

3

No DB Timeouts on web tier

Web nodes wait indefinitely

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro

POSSIBLE ALERTS MISSED

Would not have prevented failure, but would give better troubleshooting guidance

1

Long running transaction

Soft alerts when locks are being held for a long time or are ‘idle in transaction’

2

Database Lock Threshold

Alerting on number of DB locks would have clued in deploy team something was wrong

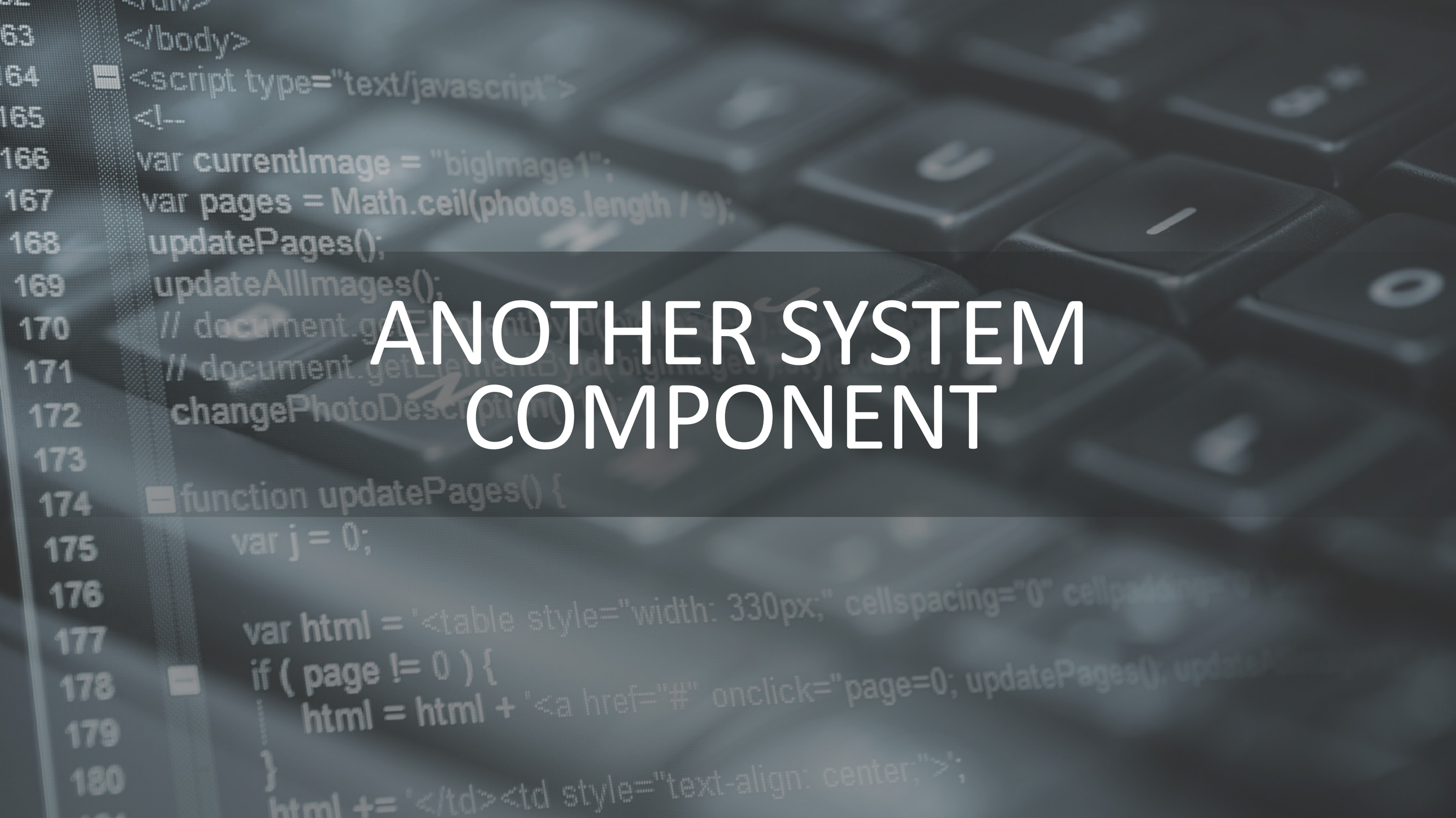
3

Blocked queries

Alert when a large number of queries are blocked

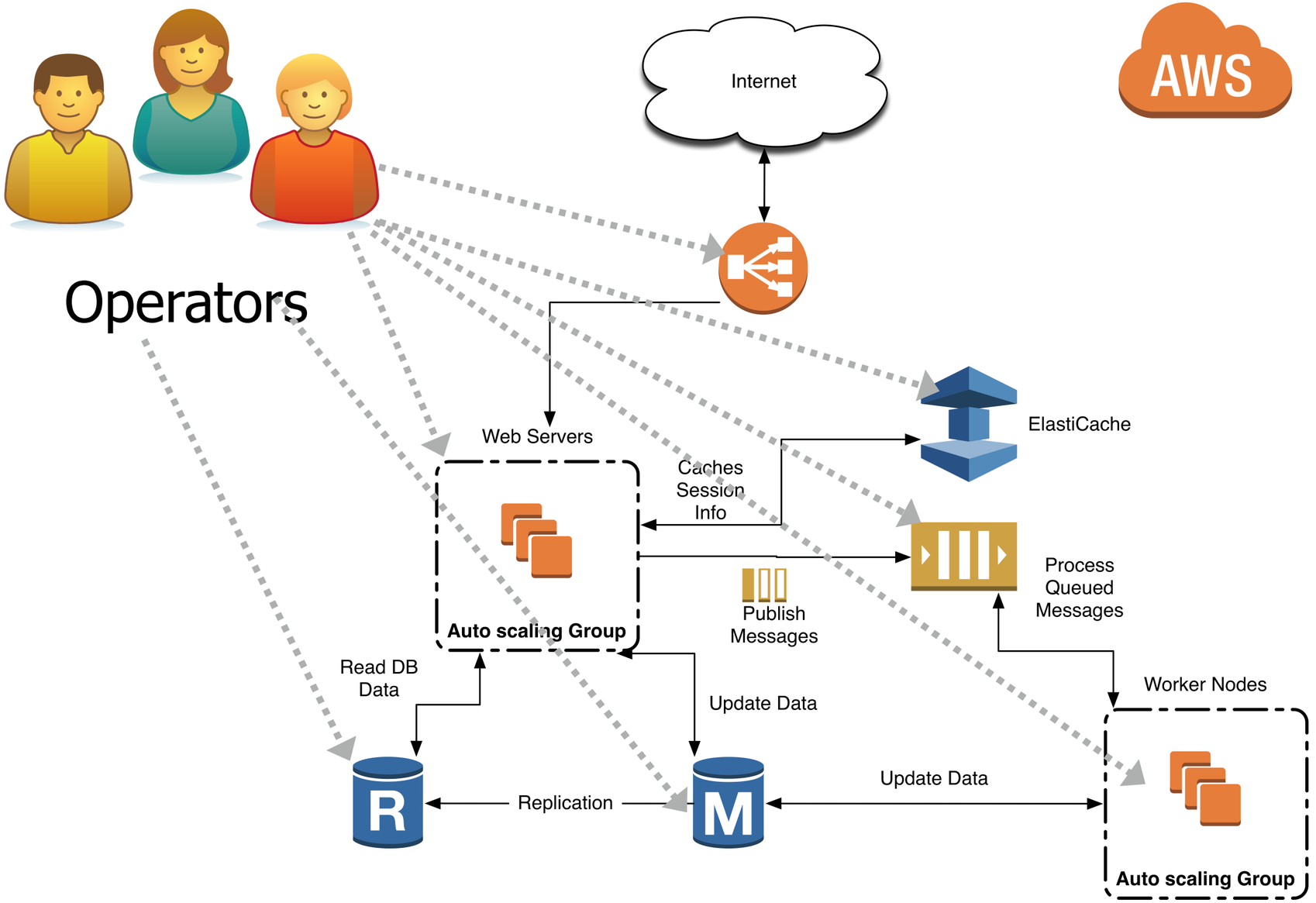
Copyright © 2016 Centro, Inc. All Rights Reserved.

centro


ANOTHER SYSTEM
COMPONENT



Copyright © 2016 Centro, Inc. All Rights Reserved.



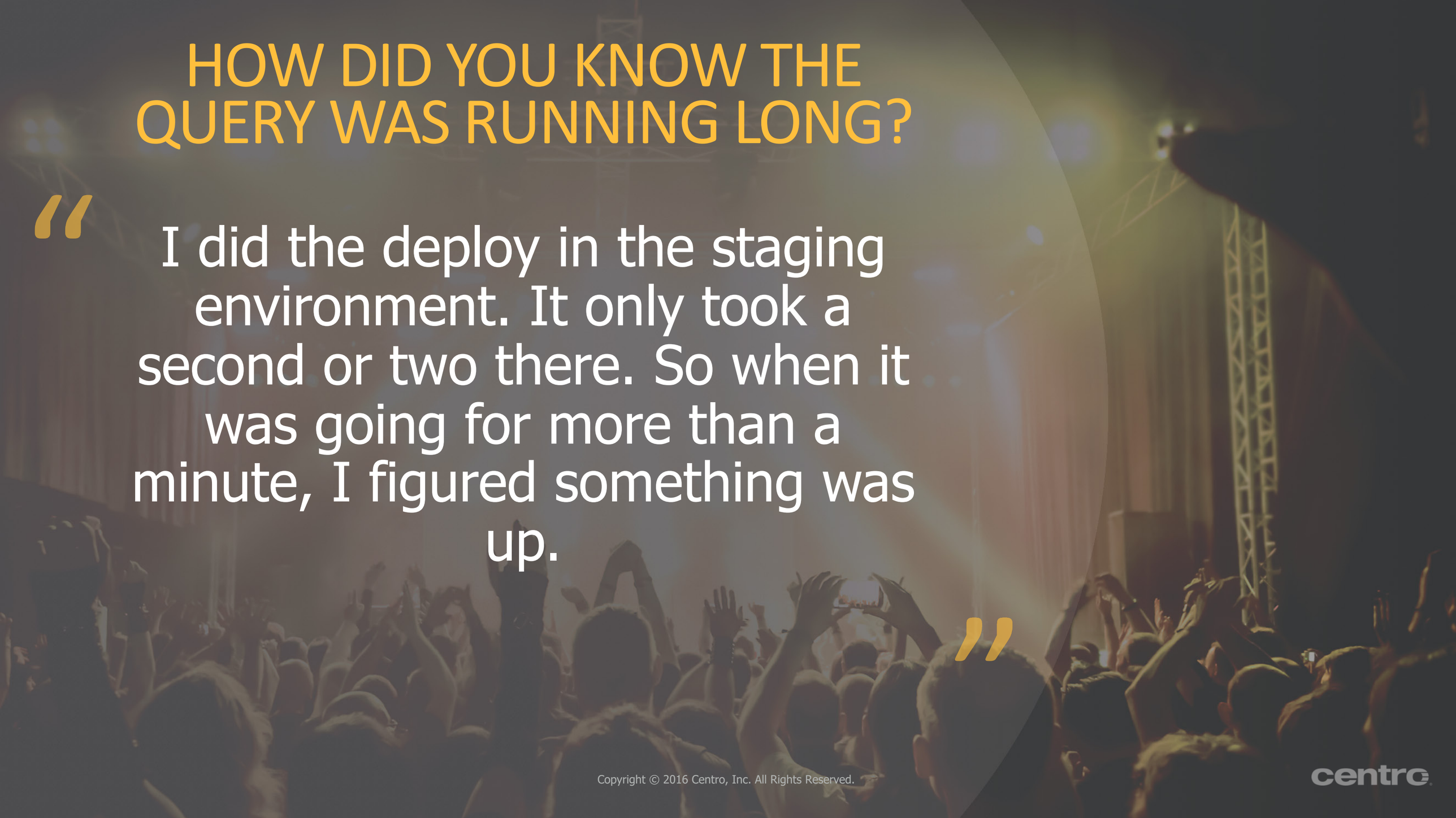
High Level System Diagram

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro


Backspace
Ins
Del
Home
Pg Up
Pg Dn
Num Lock
Pause Break
PrtSc SysRq
Ctrl
Shift
7 Home
8
9 Pg Up
4
5
6
1 End
2

HOW DID YOU KNOW THE QUERY WAS RUNNING LONG?



> “I did the deploy in the staging environment. It only took a second or two there. So when it was going for more than a minute, I figured something was up.”

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro



kerry.borden 2:46 AM

@marvin run platform migrations for production on inactive deployment



marvin APP 2:46 AM

Running DB migrations for environment production



marvin APP 2:46 AM

[centro.toggle_salt | production | 5da7b9f5f92a0b23357eb5f5 | kerry.borden] - Setting salt to disable in environment production

[centro.toggle_salt | production | 5da7b9f5f92a0b23357eb5f5 | kerry.borden] - Salt was disabled on production.

[platform.run_migrations | production | 5da7b9e1f92a0b23357eb5e6 | kerry.borden] - Stopping sidekiq for production

[platform.restart_service | production | 5da7ba07f92a0b23357eb601 | kerry.borden] - Success! stop-ed sidekiq_default consumer_daemon sidekiq_auxiliary sitescout_workers bi_workers report_workers campaign_refresh_workers in production.

[platform.run_migrations | production | 5da7b9e1f92a0b23357eb5e6 | kerry.borden] -
Found 1 staging migrations run time for version 4.91.0:
11 seconds | on staging02 @ 10-10-2019 20:02, package: 4.91.0-705

[platform.run_migrations | production | 5da7b9e1f92a0b23357eb5e6 | kerry.borden] - DB Migrations complete!

WHY DIDN'T WE CATCH THIS IN STAGING?



"We don't' have users in Staging. Also the Sidekiq job that was running in production creating the long running transaction was initiated by a user and was kind of an unusual use case. It was a perfect storm."

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro

WHY DID WE NEED TO PAGE OPS?

> Nobody else has access to the DB at that layer. Plus we don't have the Postgres knowledge to find blocking queries. I wouldn't even know where to start finding out that info.

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro



Norman Chan 2:36 PM

@marvin list locked queries for production



marvin APP 2:36 PM

Checking for locked queries for production...



marvin APP 2:36 PM

[platform.locked_queries | production | 5d71013bf92a0b204e3b537c | norman.chan] - 175 blocked queries.

##########################

[40579] `SELECT *, pg_size_pretty(total_bytes) AS total

, pg_size_pretty(index_bytes) AS INDEX

, pg_size_pretty(toast_bytes) AS toast

, pg_size_pretty(table_bytes) AS TABLE

FROM (

SELECT *, total_bytes-index_bytes-COALESCE(toast_bytes,0) AS table_bytes FROM (

SELECT c.oid,nspname AS table_schema, relname AS TABLE_NAME

, c.reltuples AS row_estimate

, pg_total_relation_size(c.oid) AS total_bytes

, pg_indexes_size(c.oid) AS index_bytes

, pg_total_relation_size(reltoastrelid) AS toast_bytes

FROM pg_class c

LEFT JOIN pg_namespace n ON n.oid = c.relnamespace

WHERE relkind = 'r'

) a

) a;`

is blocked by:

[34040] - `ALTER TABLE "dsp_line_items" ADD "user_input_media_cost" decimal(16,8)`

Running for 2487 seconds from cmm@10.40.11.63. Status: `active`.

##########################

centro

WHY WASN'T THIS MIGRATION FLAGGED FOR A MAINTENANCE RELEASE?



"This was the developer's first time really working in the platform. He wasn't aware of the history or the access contention around this table. And it went smoothly in staging."

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro

Workzone Settings

Properties | Reviewers | Approvals | Merge

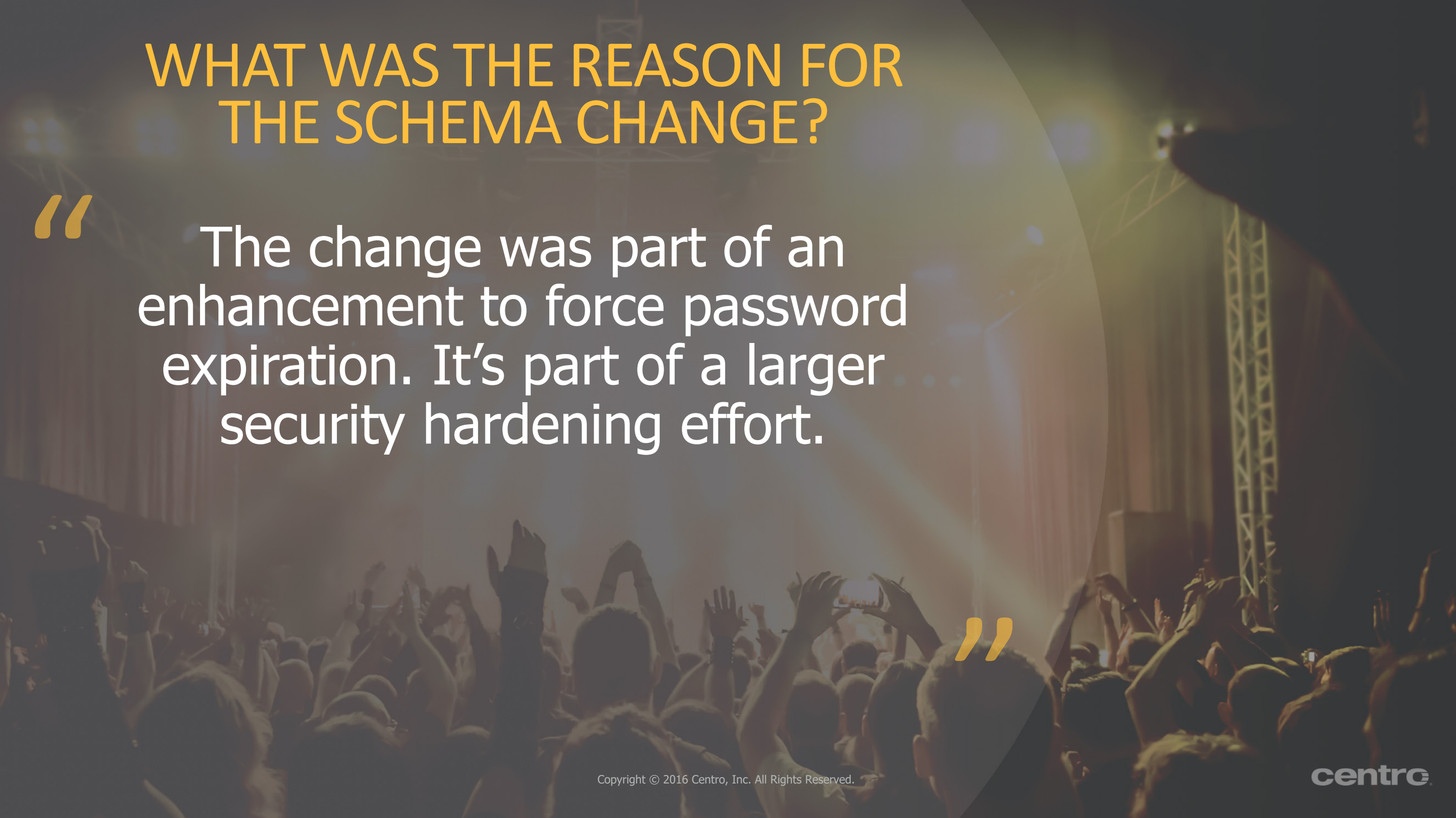
Pull Request Reviewers

Add Reviewers to a Pull Request's target branch automatically. Marks mandatory reviewers and groups.

Branch	Users	Groups	Suggested
refs/heads/dev Files & Paths ...	*No users have been selected*	*No groups have been selected*	0

Path / File	Users	Groups
docker/**/*	Ken Schultz	*No groups have been selected*
.ruby-version	jeff smith Ken Schultz	*No groups have been selected*
doc/rfcs/**/*	*No users have been selected*	platform-dev
.rubocop*	Ken Schultz	*No groups have been selected*
db/**/*	Josh Davison Ken Schultz	*No groups have been selected*

WHAT WAS THE REASON FOR THE SCHEMA CHANGE?



"The change was part of an enhancement to force password expiration. It's part of a larger security hardening effort."

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro



NIST SP800-63B SECURITY GUIDELINES

Verifiers SHOULD NOT require memorized secrets to be changed arbitrarily (e.g., periodically). However, verifiers SHALL force a change if there is evidence of compromise of the authenticator.

Copyright © 2016 Centro, Inc. All Rights Reserved.

WAS IT A TECHNICAL FAILURE?

- A change that wasn’t needed
- Was missed by a review process that was disabled
- The change couldn’t be simulated in staging
- Interacted poorly with an unusual user action
- Which led to a system failure

Copyright © 2016 Centro, Inc. All Rights Reserved.



LESSONS LEARNED



Work in a vacuum

Talk with all stakeholders

Use automation to free expertise

Automating common questions

Staging is not production

Making it even close is a lot of work

Humans are part of the system

Our processes can have technical impacts

RECAP

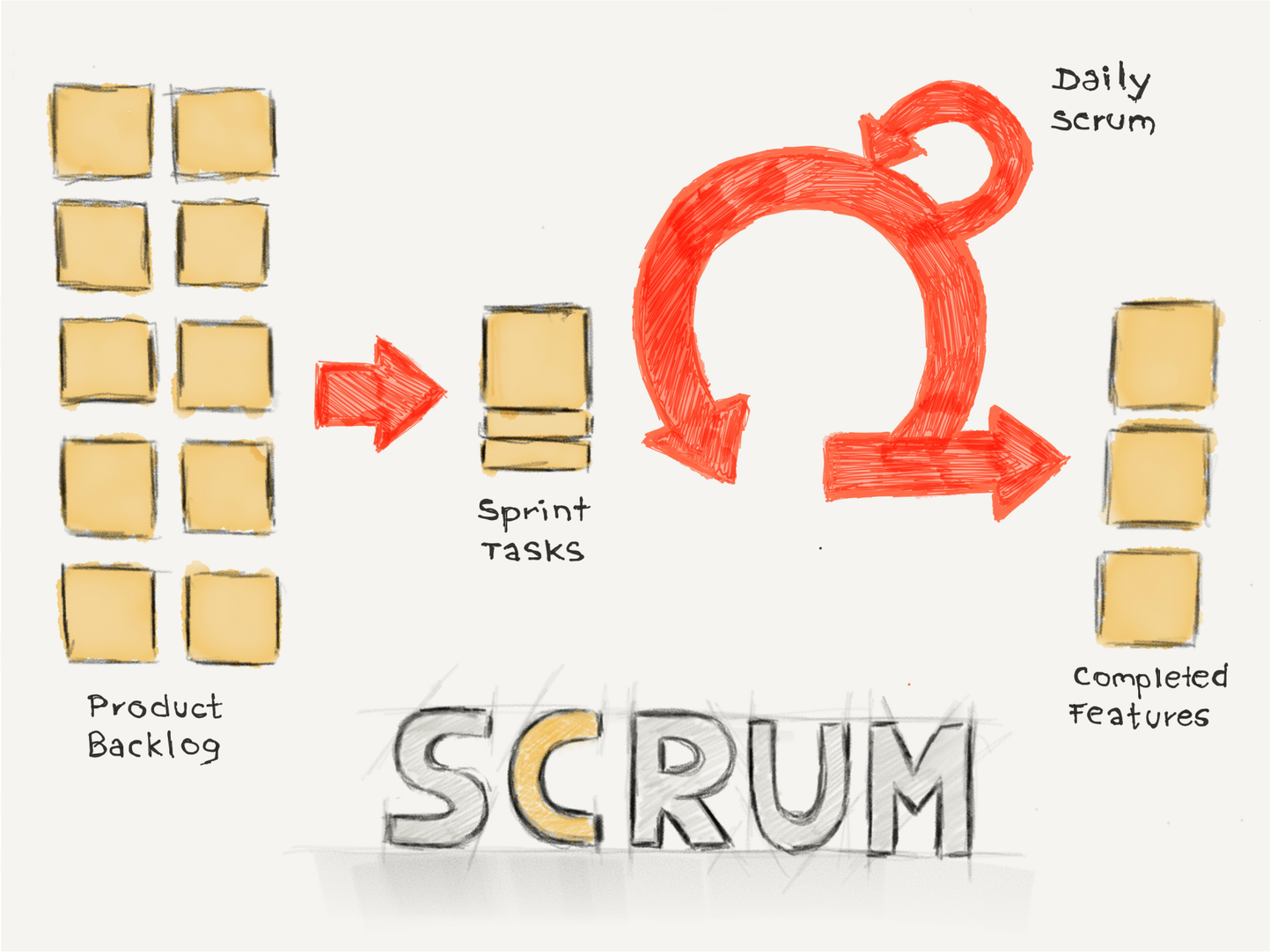
No single root cause

A combination of failures triggers tragedy

Copyright © 2016 Centro, Inc. All Rights Reserved.

centro

UPTIME AND STABILITY ARE A FEATURE



Copyright © 2016 Centro, Inc. All Rights Reserved.

centro

THANK YOU.

Jeff Smith
jeff.smith@centro.net
@darkandnerdy

centro

Please

Remember to rate this session

Thank you!

