



Automating Schema Migrations

with GitHub Actions, skeema & gh-ost

Shlomi Noach

GitHub

FOSDEM 2020



About me



@github/database-infrastructure

Author of **orchestrator**, **gh-ost**, **freno**, **ccql** and others.

Blog at <http://openark.org>

github.com/shlomi-noach

@ShlomiNoach



GitHub

Built for developers

Busy and growing

40M+ developers

3M organizations

44M repositories created in the past year

Actions, Packages,
Security Advisories & Updates,
Code Navigation & Search,
Notifications, Sponsors,
Mobile, ...



Incentive



With MySQL as the backend, new features imply schema changes: new tables, new columns, index changes, iterative schema changes, experiments.



What's in a migration?

More than **CREATE**, **ALTER** or **DROP TABLE**



What's in a migration?



Designing, coding, local testing, review process, queueing, scheduling, executing, controlling, auditing, versioning...



Ownership



Developer

Design code + schema change

Developer

Publish

Peer

Review

DBA

Review

DBA

Formalize statement/command

DBA

Locate

DBA

Schedule

DBA

Run

DBA/SRE

Audit/monitor/control

DBA

Cut-over/complete migration

DBA

Cleanup

DBA

Notify

Developer

Deploy

Developer

Merge



Ownership, our previous state



Developer	Design code + schema change
Developer	Publish
Peer	Review
DBA	Review
DBA	Formalize statement/command
DBA	Locate
DBA	Schedule
DBA	Run
gh-ost/chatops	Audit/monitor/control
DBA	Cut-over/complete migration
DBA	Cleanup
DBA	Notify
DBA	Deploy
DBA	Merge



A complex flow



Multiple domains (code, MySQL, production, communication).

Multiple environments (dev, production).

Multiple owners (devs, DBAs, SREs).





A combinatory solution

Loosely coupled, independent components.

Each solves an aspect of the problem.

Orchestrated to create an automated flow.





Code

A schema change should be presented as code.

Coupled with application code.

Versioned.



Code



At GitHub, we use git.

We also happen to author GitHub for code hosting, versioning and management.



Pull Request



The change

Review

CI

Discussion



Code: New PR

Schema change demo for FOSDEM 2020 #64

Edit

 Open shlomi-noach wants to merge 1 commit into `master` from `fosdem-2020-demo`

 Conversation 0

 Commits 1

 Checks 0

 Files changed 1

+2 -2 



shlomi-noach commented now

+  ...

Demonstrating `skeefree` for FOSDEM 2020.

In a schema change PR we make no code changes; it's just the schema. The developer should create a preceding/followup PR such that the code supports the change.

  Schema change demo for FOSDEM 2020

 1a4b4b4

Add more commits by pushing to the `fosdem-2020-demo` branch on `github/skeefree-test`.



Review required

At least 1 approving review is required by reviewers with write access. [Learn more](#).

Some checks haven't completed yet

2 queued checks

[Hide all checks](#)

  skeema-diff / skeema-labels-validation (pull_request) *Queued — Waitin...* [Details](#)

  skeema-diff / skeema-diff (pull_request) *Queued — Waiting to run this c...* **Required** [Details](#)

Merging is blocked

Merging can be performed automatically with 1 approving review.

Reviewers

No reviews—at least 1 approving review is required.

Assignees

No one—assign yourself

Labels

None yet

Projects

None yet

Milestone

No milestone

Linked issues

None yet

Notifications

[Customize](#)

 **Unsubscribe**

You're receiving notifications because you authored the thread.



Code: New PR

Changes from all commits ▾ File filter... ▾ Jump to... ▾ ⚙ ▾

0 / 1 files viewed ⓘ

Review changes ▾

4 schema/skeefree_test/some_table.sql 📄

Viewed ...

@@ -5,6 +5,6 @@ CREATE TABLE `some_table` (5 `time_created` datetime NOT NULL DEFAULT CURRENT_TIMESTAMP, 6 `time_updated` datetime DEFAULT NULL, 7 `random_hash` char(40) CHARACTER SET ascii DEFAULT NULL, 8 - PRIMARY KEY (`id`), 9 - UNIQUE KEY `random_hash_key` (`random_hash`) 10) ENGINE=InnoDB DEFAULT CHARSET=utf8;	5 `time_created` datetime NOT NULL DEFAULT CURRENT_TIMESTAMP, 6 `time_updated` datetime DEFAULT NULL, 7 `random_hash` char(40) CHARACTER SET ascii DEFAULT NULL, 8 + `fosdem_year` INT UNSIGNED NOT NULL DEFAULT '2020', 9 + PRIMARY KEY (`id`) 10) ENGINE=InnoDB DEFAULT CHARSET=utf8;
---	--



What's the migration?

From code to SQL statement





skeema

<https://www.skeema.io/>

<https://github.com/skeema/skeema>

Open source

Developed by Evan Elias





```
schema/  
  .skeema  
  my_schema1/  
    .skeema  
    some_table.sql  
    another_table.sql  
  my_schema2/  
    .skeema  
    foo.sql  
    bar.sql
```





```
$ cat .skeema
```

```
[skeema-diff-ci]  
host=127.0.0.1  
port=3306  
user=root
```





```
$ cat .my_schema1/some_table.sql
```

```
CREATE TABLE `some_table` (  
  `id` int(10) unsigned NOT NULL AUTO_INCREMENT,  
  `hostname` varchar(128) NOT NULL,  
  `time_created` datetime NOT NULL DEFAULT CURRENT_TIMESTAMP,  
  `time_updated` datetime DEFAULT NULL,  
  `random_hash` char(40) CHARACTER SET ascii DEFAULT NULL,  
  PRIMARY KEY (`id`)  
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```





```
$ skeema push env-name
```

```
# Connects to MySQL server, applies filesystem schema onto  
database
```

```
$ skeema pull env-name
```

```
# Imports schema definition from MySQL onto filesystem
```

```
$ skeema diff env-name
```

```
# Prints schema migrations changing the state of MySQL server to  
filesystem
```



skeema

Where?



GitHub Actions



Kick off workflows with GitHub events like push, issue creation, or a new release.

An action runs in a container on GitHub's infrastructure (default).

Action has repository's context and can operate on the repository.



GitHub Actions



Run skeema from within Action.

Fetch skeema as part of Action flow.



GitHub Action: skeema-diff, simplified

```
skeema-diff:  
  runs-on: ubuntu-latest  
  steps:  
  - uses: actions/checkout@v2  
    with:  
      ref: master  
  - name: push master schema to MySQL  
    env:  
      MYSQL_PWD: root  
    run: |  
      skeema push skeema-diff-ci  
  - uses: actions/checkout@v2  
  - name: skeema diff  
    skeema diff skeema-diff-ci --allow-unsafe
```



GitHub Action: skeema-diff

```
- uses: actions/checkout@v2
  with:
    ref: ${{ github.event.pull_request.base.sha }}
- name: push master schema to MySQL
  ...
- uses: actions/checkout@v2
  with:
    ref: ${{ github.event.pull_request.head.sha }}
- name: skeema diff
  /tmp/skeema-ci/skeema push skeema-diff-ci --allow-unsafe --ddl-
wrapper='echo "\n-- skeema:ddl:begin\n"{DDL}";\n-- skeema:ddl:end"' | sed
-e 's/^USE /-- skeema:ddl:use /g' | sed -n '/^-- skeema:ddl:use /p;/^--
skeema:ddl:begin/,/^-- skeema:ddl:end/p' | tee /tmp/skeema-ci/skeema-
diff.sql
```



Action/ skeema

Schema change demo for FOSDEM 2020 #64

 **Open** shlomi-noach wants to merge 1 commit into `master` from `fosdem-2020-demo` 

 Conversation **0**  Commits **1**  Checks **2**  Files changed **1**

 Schema change demo for FOSDEM 2020 1a4b4b4 ▾

▼  **skeema-diff**
on: pull_request

✓ [skeema-labels-validation](#)

● **skeema-diff**

skeema-diff / skeema-diff
started 10s ago

- ▶ ✓ **Set up job**
- ▶ ✓ **Run actions/checkout@v1**
- ▶ ✓ **Find skeema file**
- ▶ ✓ **Fetch skeema**
- ▶ ✓ **Fetch hub**
- ▶ ✓ **push master schema to MySQL**
- ▶ ✓ **Run actions/checkout@v1**
- ▶ ✓ **skeema diff**
- ▶ ✓ **validate diff**
- ▶ ✓ **check label**
- ▶ ✓ **validate migration:skeefree:queued status**
- ▶ ✓ **label PR**
- ▼  **comment PR**
- ▼  **comment suggestion for review**



Action/ skeema

Schema change demo for FOSDEM 2020 #64

[Edit](#)

[Open](#) shlomi-noach wants to merge 1 commit into `master` from `fosdem-2020-demo`

Conversation 0

Commits 1

Checks 0

Files changed 1

+2 -2



shlomi-noach commented 33 seconds ago

+ 😊 ...

Demonstrating `skeefree` for FOSDEM 2020.

In a schema change PR we make no code changes; it's just the schema. The developer should create a preceding/followup PR such that the code supports the change.

Schema change demo for FOSDEM 2020

✓ 1a4b4b4

github-actions bot added the `migration:skeema:diff` label now



github-actions bot commented now

+ 😊 ...

```
-- skeema:diff
-- skeema:ddl:use skeefree_test ;
-- skeema:ddl:begin
ALTER TABLE some_table ADD COLUMN fosdem_year int(10) unsigned NOT NULL DEFAULT '2020',
DROP KEY random_hash_key ;
-- skeema:ddl:end
```



github-actions bot commented now

+ 😊 ...

@shlomi-noach it looks like you are making schema changes. When ready for production review & migration, please add the label `migration:for:review`.

Reviewers

No reviews—at least 1 approving review is required.

Assignees

No one—assign yourself

Labels

`migration:skeema:diff`

Projects

None yet

Milestone

No milestone

Linked issues

None yet

Notifications

Customize

🔊 Unsubscribe

You're receiving notifications because you were mentioned.

1 participant





gh-ost

GitHub's online schema migration tool.

Low-impact (nearly no-impact) in production.

Auditable, configurable, controllable.

Open source

<https://github.com/github/gh-ost/>

<https://github.blog/2016-08-01-gh-ost-github-s-online-migration-tool-for-mysql/>

<https://speakerdeck.com/shlominoach/githubs-online-schema-migrations-for-mysql>





skeefree

Name coined by Tom Krouper

A service to orchestrate the flow: PR, Actions/skeema, gh-ost.

Developed internally at GitHub



skeefree



Probes and detects schema change PRs

Analyzes skeema changes

Initiates and follows up on review/approval status

Schedules the migration

Runs the migration (gh-ost/direct)

Follows up and reports on PR



Flow: New PR

Schema change demo for FOSDEM 2020 #64

[Edit](#)

[Open](#) shlomi-noach wants to merge 1 commit into `master` from `fosdem-2020-demo`

Conversation 0

Commits 1

Checks 0

Files changed 1

+2 -2



shlomi-noach commented now

+ 😊 ...

Demonstrating `skeefree` for FOSDEM 2020.

In a schema change PR we make no code changes; it's just the schema. The developer should create a preceding/followup PR such that the code supports the change.

Schema change demo for FOSDEM 2020

1a4b4b4

Add more commits by pushing to the `fosdem-2020-demo` branch on `github/skeefree-test`.



Review required

At least 1 approving review is required by reviewers with write access. [Learn more](#).

Some checks haven't completed yet

2 queued checks

[Hide all checks](#)

skeema-diff / skeema-labels-validation (pull_request) Queued — Waitin... [Details](#)

skeema-diff / skeema-diff (pull_request) Queued — Waiting to run this c... **Required** [Details](#)

Merging is blocked

Merging can be performed automatically with 1 approving review.

Reviewers



No reviews—at least 1 approving review is required.

Assignees



No one—assign yourself

Labels



None yet

Projects



None yet

Milestone



No milestone

Linked issues



None yet

Notifications

[Customize](#)

Unsubscribe

You're receiving notifications because you authored the thread.



Flow: New PR

Changes from all commits ▾ File filter... ▾ Jump to... ▾ ⚙ ▾ 0 / 1 files viewed ⓘ [Review changes ▾](#)

4 ■ ■ ■ schema/skeefree_test/some_table.sql

Viewed ...

5 @@ -5,6 +5,6 @@ CREATE TABLE `some_table` (6 `time_created` datetime NOT NULL DEFAULT CURRENT_TIMESTAMP, 7 `time_updated` datetime DEFAULT NULL, 8 `random_hash` char(40) CHARACTER SET ascii DEFAULT NULL, 9 - PRIMARY KEY (`id`), 10 - UNIQUE KEY `random_hash_key` (`random_hash`) 11) ENGINE=InnoDB DEFAULT CHARSET=utf8;	5 `time_created` datetime NOT NULL DEFAULT CURRENT_TIMESTAMP, 6 `time_updated` datetime DEFAULT NULL, 7 `random_hash` char(40) CHARACTER SET ascii DEFAULT NULL, 8 + `fosdem_year` INT UNSIGNED NOT NULL DEFAULT '2020', 9 + PRIMARY KEY (`id`) 10) ENGINE=InnoDB DEFAULT CHARSET=utf8;
---	---



Flow: CI build

Schema change demo for FOSDEM 2020 #64

 Open shlomi-noach wants to merge 1 commit into `master` from `fosdem-2020-demo` 

 Conversation **0**  Commits **1**  Checks **2**  Files changed **1**

 Schema change demo for FOSDEM 2020 `1a4b4b4` ▾

▾  **skeema-diff**
on: pull_request

✓ [skeema-labels-validation](#)

● **skeema-diff**

skeema-diff / skeema-diff
started 10s ago

- ▶ ✓ Set up job
- ▶ ✓ Run actions/checkout@v1
- ▶ ✓ Find skeema file
- ▶ ✓ Fetch skeema
- ▶ ✓ Fetch hub
- ▶ ✓ push master schema to MySQL
- ▶ ✓ Run actions/checkout@v1
- ▶ ✓ skeema diff
- ▶ ✓ validate diff
- ▶ ✓ check label
- ▶ ✓ validate migration:skeefree:queued status
- ▶ ✓ label PR
- ▾ ● comment PR
- ▾ ● comment suggestion for review



Flow: skeema analysis

Schema change demo for FOSDEM 2020 #64

Edit

[Open](#) shlomi-noach wants to merge 1 commit into `master` from `fosdem-2020-demo`

Conversation 0 Commits 1 Checks 0 Files changed 1

+2 -2



shlomi-noach commented 33 seconds ago

+ 😊 ...

Demonstrating `skeefree` for FOSDEM 2020.

In a schema change PR we make no code changes; it's just the schema. The developer should create a preceding/followup PR such that the code supports the change.

Schema change demo for FOSDEM 2020

✓ 1a4b4b4

github-actions bot added the `migration:skeema:diff` label now



github-actions bot commented now

+ 😊 ...

```
-- skeema:diff
-- skeema:ddl:use skeefree_test ;
-- skeema:ddl:begin
ALTER TABLE some_table ADD COLUMN fosdem_year int(10) unsigned NOT NULL DEFAULT '2020',
DROP KEY random_hash_key ;
-- skeema:ddl:end
```



github-actions bot commented now

+ 😊 ...

@shlomi-noach it looks like you are making schema changes. When ready for production review & migration, please add the label `migration:for:review`.

Reviewers

No reviews—at least 1 approving review is required.

Assignees

No one—assign yourself

Labels

`migration:skeema:diff`

Projects

None yet

Milestone

No milestone

Linked issues

None yet

Notifications

Customize

🔊 Unsubscribe

You're receiving notifications because you were mentioned.

1 participant



Flow: review & label

+2 -2

Reviewers

hubot

Assignees

No one—assign yourself

Labels

migration:for:review

migration:skeema:diff

github-actions bot commented 2 minutes ago + 😊 ...

@shlomi-noach it looks like you are making schema changes. When ready for production review & migration, please add the label `migration:for:review`.

shlomi-noach added the `migration:for:review` label 1 minute ago

hubot approved these changes now [View changes](#)



Flow: skeefree analysis

A screenshot of a Hubot chat interface. The header shows 'hubot commented now' on the left and '+ 😊 ...' on the right. The main content area contains the text 'Migration instructions for @github/database-infrastructure:' followed by a code block containing a database migration command.

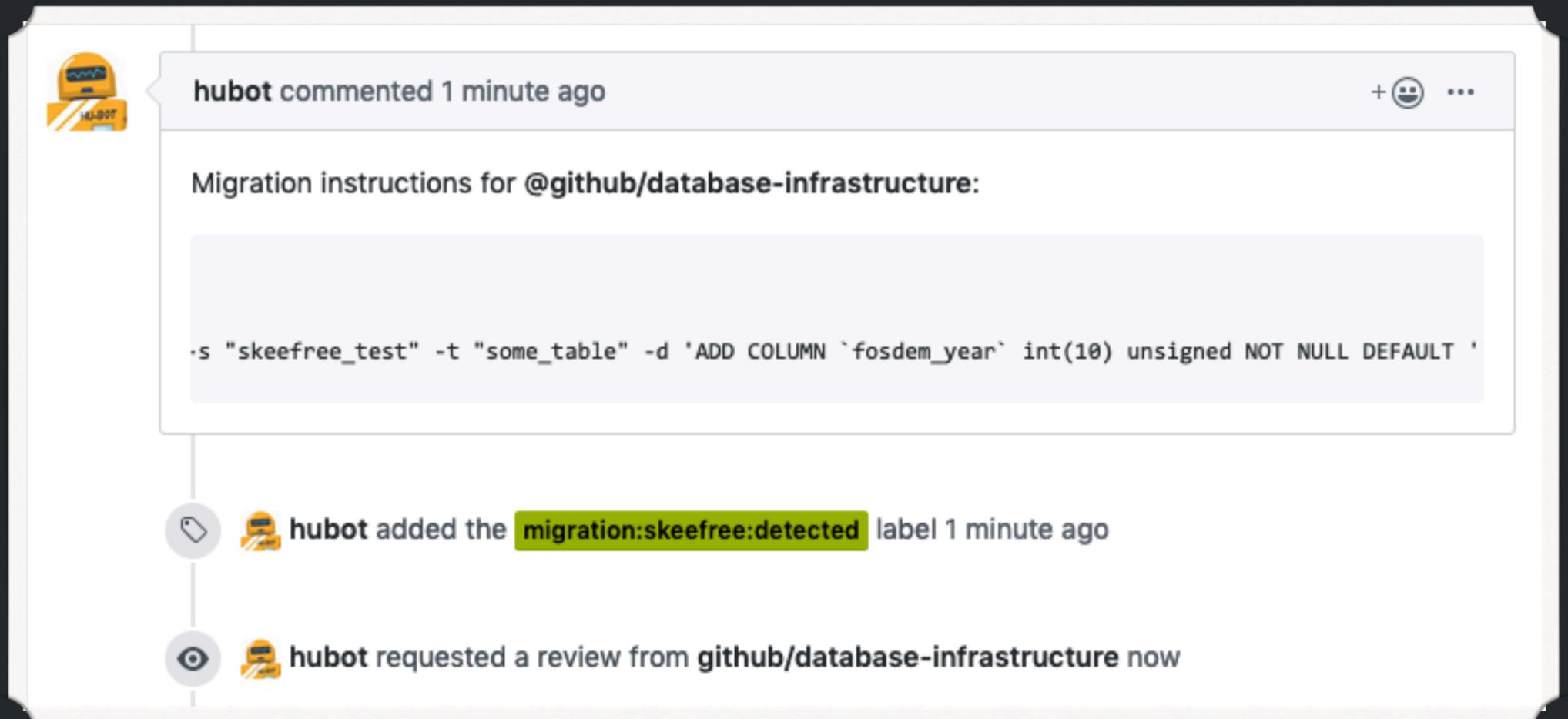
hubot commented now + 😊 ...

Migration instructions for @github/database-infrastructure:

```
-s "skeefree_test" -t "some_table" -d 'ADD COLUMN `fosdem_year` int(10) unsigned NOT NULL DEFAULT '
```



Flow: Review request



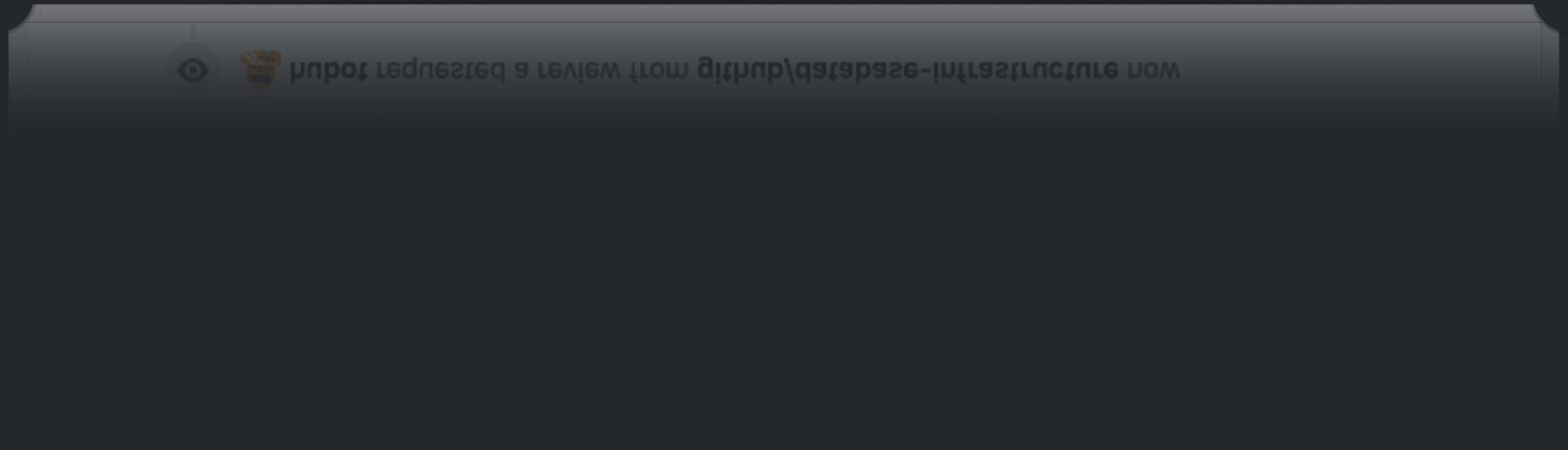
hubot commented 1 minute ago

Migration instructions for @github/database-infrastructure:

```
-s "skeefree_test" -t "some_table" -d 'ADD COLUMN `fosdem_year` int(10) unsigned NOT NULL DEFAULT '
```

hubot added the **migration:skeefree:detected** label 1 minute ago

hubot requested a review from github/database-infrastructure now



hubot requested a review from github/database-infrastructure now

Reviewers 

 hubot  

 Database Infrastruct... 

Assignees 

No one—assign yourself

Labels 

migration:for:review

migration:skeefree:detected

migration:skeema:diff



Flow: migration execution



hubot commented 34 seconds ago



ALTER TABLE some_table migration has been started by skeefree 🙌



hubot commented 31 seconds ago



ALTER TABLE some_table **noop** migration has been executed by skeefree successfully 🙌



hubot commented 29 seconds ago



ALTER TABLE some_table migration has been started by skeefree 🙌



hubot commented 24 seconds ago



ALTER TABLE some_table migration has been executed by skeefree 🎉



hubot removed the migration:skeefree:queued label now



hubot commented now



@shlomi-noach All migrations in this PR are in complete status. Please go ahead and follow your standard deploy/merge flow.



Flow: deploy & merge

Schema change demo for FOSDEM 2020 #64

 **Merged** shlomi-noach merged 1 commit into `master` from `fosdem-2020-demo`  now

 Conversation **0**

 Commits **1**

 Checks **0**

 Files changed **1**



shlomi-noach commented 10 minutes ago



Demonstrating `skeefree` for FOSDEM 2020.

In a schema change PR we make no code changes; it's just the schema. The developer should create a preceding/followup PR such that the code supports the change.



Ownership: skeefree



Developer	Design code + schema change
Developer	Publish
Peer	Review
DBA	Review
skeema/CI	Formalize statement/command
skeefree	Locate
skeefree	Schedule
skeefree	Run
gh-ost/chatops	Audit/monitor/control
skeefree	Cut-over/complete migration
GC	Cleanup
skeefree	Notify
Developer	Deploy
Developer	Merge





Impact

Database team work reduced to *minutes per week*.

Developers have visibility into status. Get notified on their PR.

Better time utilization; migrations start executing as soon as possible, not based on a human availability.



skeefree



Uses internal services in GitHub's infrastructure

- Inventory service
- MySQL discovery service
- Chatops integration
- Internal libraries (e.g. logging)



Open Source



skeefree is coupled with GitHub's infrastructure:

- Inventory service
- MySQL discovery
- Chat/chatops

We nonetheless hope that the community finds it useful and are releasing it in partial state.

Release to be announced.



Thank you!



Questions?

github.com/shlomi-noach

@ShlomiNoach

