

How to import projects between Gerrit Servers



IMPORT →
← **EXPORT**

Luca Milanesio

Gerrit Maintainer
Member of the Gerrit ESC
Gerrit Release Manager

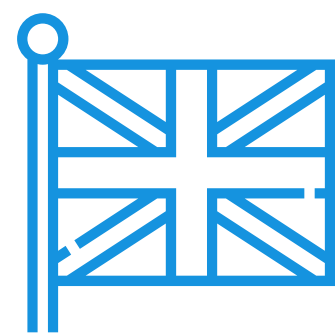
Join at
slido.com
#3462 670





How many Gerrit servers you have?

About GerritForge Inc.



US Corporation
Coming from
UK roots



HQ in Sunnyvale CA
since 2019



Committed to
OpenSource and
to Gerrit Code
Review since 2009



Agenda

1. Importing projects with Gerrit 2.x / ReviewDb
2. The Gerrit importer plugin (by SAP)
3. The promise of NoteDb with Gerrit 2.15
4. Problems with importing projects in the Gerrit 3.x era
5. Current situation with Gerrit 3.7
6. Ongoing work for Gerrit 3.8 (master branch)
7. Q&A

Gerrit 2.x

```
docker run -ti -p 80:8080 -h localhost gerritcodereview/gerrit:2.13.14
```

The screenshot shows the Gerrit web interface. At the top, there is a navigation bar with tabs for 'All', 'My', 'Projects', 'People', 'Plugins', and 'Documentation'. Below this, there are sub-tabs for 'List', 'General', 'Branches', 'Tags', 'Access', 'Dashboards', and 'Create New Project'. The main content area displays a change titled 'Change 1 - Needs Code-Review Label'. The description of the change is 'Insert the description of the change.' and the change ID is 'Ia89b5427559a7259224321e4bc9503bdd905c5b2'. On the right side, there is a metadata panel with the following information: Owner: Administrator, Reviewers: (empty), Project: test-repo, Branch: master, Topic: (empty), Strategy: Merge if Necessary, and Updated: 0 seconds ago. Below the metadata panel, there are buttons for 'Cherry Pick', 'Rebase', 'Abandon', and 'Follow-Up'. At the bottom of the page, there is a 'Code-Review' section.

Gerrit 2.x - architecture

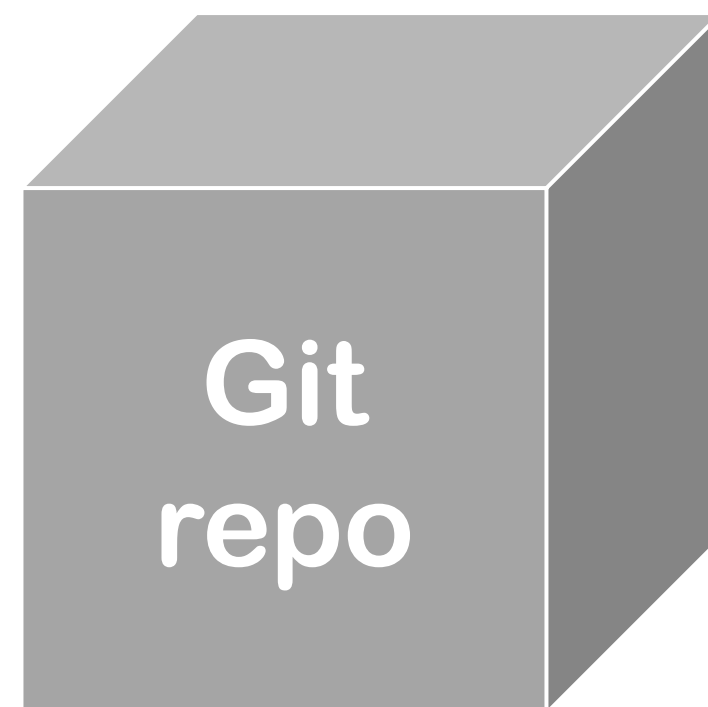
GWT UI

http://localhost/#/c/1/

Gerrit backend

http://localhost/c/1/

a79cc06e518 refs/changes/01/1/1



changes

CHANGE_KEY	DEST_PROJECT_NAME	DEST_BRANCH_NAME	CHANGE_ID	CURRENT_PATCH_SET_ID
I7fc4c285d6.	test-repo	refs/heads/master	1	1

patch_sets

REVISION	CHANGE_ID	PATCH_SET_ID
a79cc06e518	1	1

❖ Importing projects in Gerrit 2.x

Only repository without reviews “out of the box”

1. Copy the bare repository to the new Gerrit server
2. Restart Gerrit
3. Remove the ‘refs/changes/*’ because of risk of conflict of the change numbers

Importer plugin (by SAP)

Importing projects with Gerrit 2.x / ReviewDb

- **Import the repository (only refs/heads, refs/tags)**
- **Replay the review on the new server**
(can be done online, can be resumed)

Gerrit source and destination must be compatible at API level

Purpose of the SAP's importer plugin

Consolidating Gerrit Servers

- Same Gerrit version
- Smooth transition, online with current traffic

❖ HOW people used the SAP's importer plugin

Live migration between releases

- Different Gerrit versions
- Replicating production data, reviews
- Canary deployments and early adoption



How would you use the importer?

NoteDb released in Gerrit 2.15

Gerrit User Summit 2017 in London



The image shows a man in a red Google t-shirt standing at a podium on a stage. Behind him is a large screen displaying a presentation slide. The slide is titled 'NoteDb Timeline' and shows a timeline from 2014 to 2018. The timeline includes three key commits: `I87db8a25` (Jun '14), `I428978a9` (Dec '13), and `I0771f5e8` (Oct '15). The commit `I428978a9` is described as 'Store PatchSetApprovals in a git-notes-based database'. The commit `I0771f5e8` is described as 'BatchUpdate'. The commit `I87db8a25` is described as 'Store PatchLine Comments in Git notes'. The screen also shows a snippet of commit history for 2014, including the commit `I87db8a25` and the text 'Store PatchLine Comments in Git notes'.

NoteDb promises

- **Simplicity**
- **Consistency**
- **Auditability**
- **Extensibility**
- **New features**
- **Distributed code-reviews**
 - **Shawn always said “*With NoteDb, I should be able to do reviews on a plane and then just push the NoteDb meta-data refs when I land*”**

Steps towards project-scoping of NoteDb refs

<https://gerrit-review.googlesource.com/c/gerrit/+108592>

[...] changes will be served at `/c/project-name/+123` instead of `/p/project-name/+c/1234`. This is to stay more consistent with the REST API which serves all changes at `/changes` instead of `/projects`.

If a user uses a URL without the project (`/c/123`), we retrieve the project upon the first API call and rewrite the URL.

All instances of `Project.NameKey` and `String project` are marked `@Nullable` where they can actually be null. In the API layer they are consistently marked `@Nullable` to reflect the current status of the backend. That is, a project can optionally be provided. This might be changed to a more strict policy at a later stage.

Gerrit 2.16/3.0 with NoteDb

Can I import a project?

```
Caused by: org.eclipse.jgit.errors.ConfigInvalidException: Change 1: invalid identity,  
expected <id>@576acc63-c810-4599-b0a5-f7174ab3b459: 1000000@e332c9c3-45ef-4200-ae44-39bd691bf017  
  at com.google.gerrit.server.notedb.ChangeNotes.parseException(ChangeNotes.java:100)  
  at com.google.gerrit.server.notedb.LegacyChangeNoteRead.lambda$parseIdent$0(LegacyChangeNoteRead.java:57)  
  at java.base/java.util.Optional.orElseThrow(Optional.java:408)  
  at com.google.gerrit.server.notedb.LegacyChangeNoteRead.parseIdent(LegacyChangeNoteRead.java:55)  
  at com.google.gerrit.server.notedb.ChangeNotesParser.parseIdent(ChangeNotesParser.java:907)  
  at com.google.gerrit.server.notedb.ChangeNotesParser.parse(ChangeNotesParser.java:340)  
  at com.google.gerrit.server.notedb.ChangeNotesParser.parseAll(ChangeNotesParser.java:215)  
  at com.google.gerrit.server.notedb.ChangeNotesCache$Loader.call(ChangeNotesCache.java:361)  
  at com.google.gerrit.server.notedb.ChangeNotesCache$Loader.call(ChangeNotesCache.java:338)
```


Gerrit 2.16/3.0 with NoteDb

Let's have a closer look

```
commit 72dae710cf3f423b86cab05bf0ed9620f4ef20b1
Author: Gerrit User 1000000 <1000000@ce44561d-70ad-44ea-b187-18656b195c51>
Date: Thu Nov 10 22:41:59 2022 +0000
```

Create change

Uploaded patch set 1.

```
Patch-set: 1
Change-id: Ie87dfce5d0e84f1b23248abada77168767adc0f0
Subject: Change on gerrit-01
Branch: refs/heads/master
Status: new
Topic:
Commit: b376175edac71f478073ccbba6c0d5c29c00e2e
Tag: autogenerated:gerrit:newWipPatchSet
Groups: b376175edac71f478073ccbba6c0d5c29c00e2e
Private: false
Work-in-progress: true
```

Is NoteDb a broken promise? (1)

Closer look:

Account Number

```
commit 72dae710cf3f423b86cab05bf0cd9620f4cf20b1
Author: Gerrit User 1000000 <1000000@ce44561d-70ad-44ea-b187-18656b195c51>
Date: Thu Nov 10 22:41:59 2022 +0000
```

ServerId

Create change

Uploaded patch set 1.

Patch-set: 1

Change-id: Ie87dfce5d0e84f1b23248abada77168767adc0f0

Subject: Change on gerrit-01

Branch: refs/heads/master

Status: new

Topic:

Commit: b376175edac71f478073ccbba6c0d5c29c00e2e

Tag: autogenerated:gerrit:newWipPatchSet

Groups: b376175edac71f478073ccbba6c0d5c29c00e2e

Private: false

Work-in-progress: true

❖ Is NoteDb a broken promise? (2)

- **ServerId**
Identifies where the commit was first created
- **Account Number**
Account incremental ID to be used as Git identity

Is NoteDb a broken promise? (3)

Gerrit identity parsing during reindex

```
/**
 * Returns an AccountId for the given email address. Returns empty if the address isn't on this
 * server.
 */
public static Optional<Account.Id> parseIdent(PersonIdent ident, String serverId) {
    String email = ident.getEmailAddress();
    int at = email.indexOf('@');
    if (at >= 0) {
        String host = email.substring(at + 1, email.length());
        if (host.equals(serverId)) {
            Integer id = Ints.tryParse(email.substring(0, at));
            if (id != null) {
                return Optional.of((new Account.Id(id));)
            }
        }
    }
    return Optional.empty();
}
```

Is NoteDb a broken promise? (4)

<https://gerrit-review.googlesource.com/c/gerrit/+/-/75610>

```
commit f367b5d3aee7dc84e9be799646f4076d9db9a0d5
Author: Dave Borowitz <dborowitz@google.com>
Date: Tue Mar 8 20:45:44 2016 -0500
```

Include a UUID portion in NoteDb author identities

Author identities include per-server specific account IDs, so it is not safe to mix IDs from different servers. Ensure **each server only ever produces identities with one ID** during its lifetime, by writing out a random UUID to `gerrit.config` as `gerrit.serverId`. This happens during `init`, and optionally lazily during startup.

For now NoteDb changes can be migrated between servers as long as this file is kept intact. Eventually, when federating changes between servers, **we will need come up with some mechanism for coalescing various per-server identities into a single account, like the current AccountExternalId mapping** (except not exactly that because Shawn regrets it). Such a mechanism will simply need to know how to handle this kind of UUID format.

Eventually ... 6 years later

<https://gerrit-documentation.storage.googleapis.com/Documentation/3.7.0/config-gerrit.html#gerrit.importedServerId>

[gerrit.importedServerId](#)

ServerId of the repositories imported from other Gerrit servers. Changes coming associated with the imported serverIds are indexed and displayed in the UI.

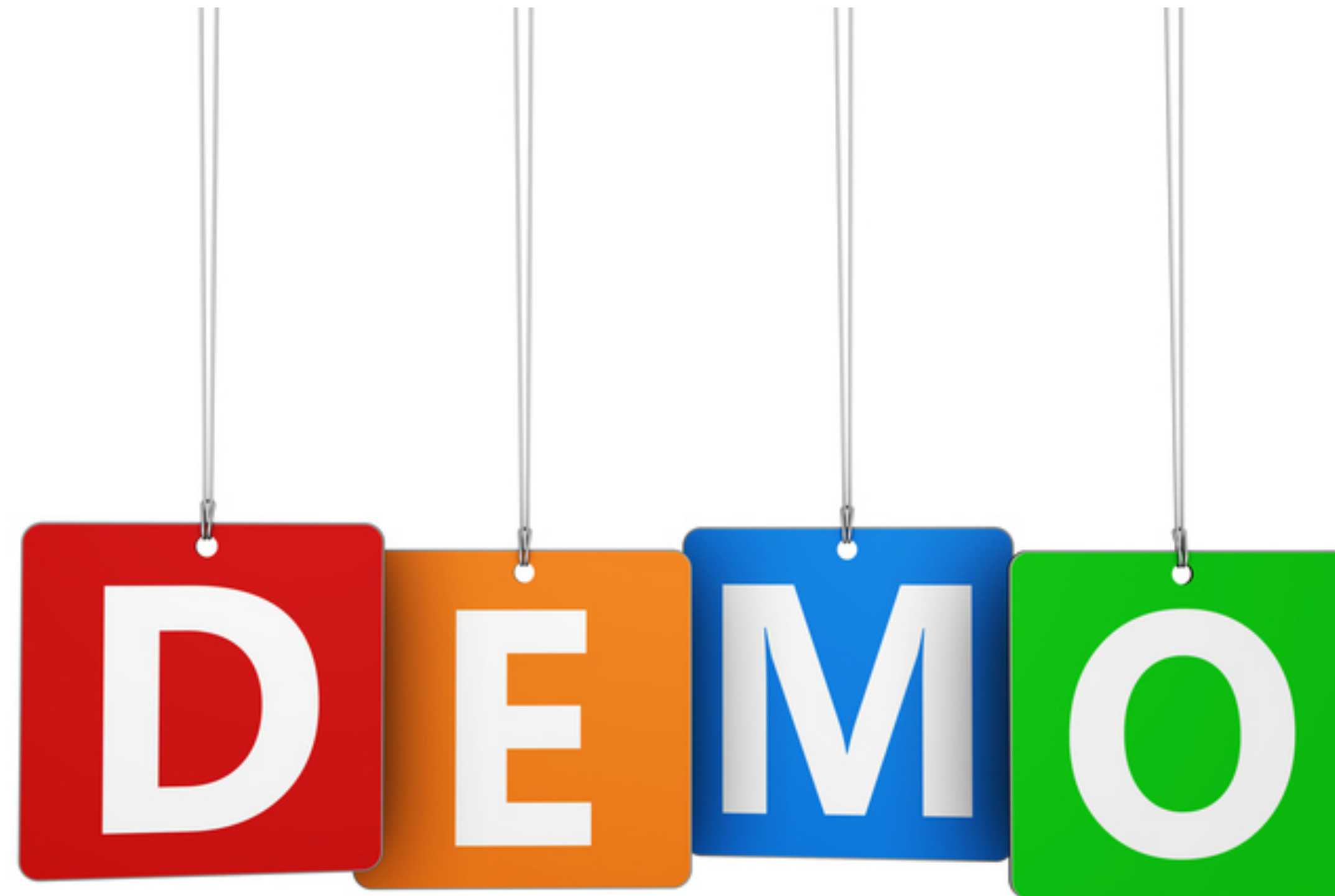
Specify multiple `gerrit.importedServerId` for allowing the import from multiple Gerrit servers with different serverIds.

NOTE

The account-ids referenced in the imported changes are used for looking up the associated account-id locally, using the `imported: external-id`. Example: the account-id 1000 from the imported server-id 59a4964e-6376-4ed9-beef will be looked up in the local accounts using the `imported:1000@59a4964e-6376-4ed9-beef` external-id.

If this value is not set, all changes imported from other Gerrit servers will be ignored.

Zero-downtime upgrade to v3.6 and v3.7



Imported external-ids in Gerrit v3.7

Identities		
Status	Email Address	Identity
		uuid:f85d97ab-46c7-4158-8063-a55883c547b1 DELETE
		imported:1000002@b48c0221-df9d-48a1-a151-e37a13586100 DELETE
		imported:1000001@e27a6e07-c2e0-4be5-b867-1cbbe8641423 DELETE

Imported projects in Gerrit v3.7

The screenshot shows the Gerrit v3.7 web interface. At the top, the browser address bar displays `localhost:8080/c/test-repo-1/+1`, which is highlighted with a red rectangle. Below the address bar is the Gerrit navigation bar with links for **Gerrit**, **CHANGES**, **YOUR**, **DOCUMENTATION**, and **BROWSE**. The main content area shows an **Active** change on `test-repo-1`. On the left, the **Change Info** sidebar includes: **Owner** (Administrator), **Reviewers** (empty), **CC** (john), **Repo | Branch** (test-repo-1 | master), and **Submit Requirements** (Code-Review No votes). The main content area displays the change title `change on test-repo-1`, the **Change-Id** `I9fdf746748bf2c4e9ff525124e741ecad2a205c0`, and an **EDIT** button. At the bottom, it shows **Comments** (No comments).

Imported projects in Gerrit v3.7

The screenshot shows the Gerrit web interface. At the top, the browser address bar contains the URL `localhost:8080/c/test-repo-2/+1`, which is highlighted with a red rectangle. Below the address bar is the Gerrit navigation bar with links for **Gerrit**, **CHANGES**, **YOUR**, **DOCUMENTATION**, and **BROWSE**. The main content area shows a change titled "test of a change from a second round of reviews" with a status of "Active" and 1 star. On the left, the "Change Info" section lists: Owner (john), Author (John Doe), Committer (John Doe), Reviewers (empty), and CC (Matt). The "Repo | Branch" is `test-repo-2 | master`. At the bottom left, "Submit Requirements" shows "Code-Review" is disabled and there are "No votes". On the right, a "REPLY" button is visible above a text input field containing the change title and ID: `Change-Id: I75d60e396544d3dddbec634df38fe5e836b8ee83`. An "EDIT" button is located at the bottom right of the input field. Below the input field, it says "Comments No comments".

Limitations in Gerrit v3.7

1. Account federation mapping tool is missing
Input: All-Users.git of another Gerrit server
Output: Add the external-ids to the local accounts
2. Limited to 64 other importedServerId
Change number is used as primary key in Lucene: we use the 3 MSB of the number with the importedServerId position
3. Manually bump the refs/sequences/changes manually to the max of the imported projects
4. Still some leftover bugs in the GUI for flakiness of reverse mapping of Change number to project name.

Way forward for Gerrit v3.8 and beyond

1. Introduce an account federation mapping tool
Input: All-Users.git of another Gerrit server
Output: Add the external-ids to the local accounts

Online? Offline?

Under discussion, pending a new design proposal

Way forward for Gerrit v3.8 and beyond

2. Make the Change number project-specific

Introduce a refs/sequences/changes inside the project

Use project hierarchy to fallback, eventually to All-Projects.git

Under discussion, pending a new design proposal

Way forward for Gerrit v3.8 and beyond

3. Introduce the project-name in the Lucene index document-id
Migration path?
Backward compatibility for zero-down upgrades?

Under discussion, pending a new design proposal

Q&A: excited about the future of Gerrit?



Image from: <http://cypp.rutgers.edu/ru-voting/political-information/public-opinion-polls/>

Wants to know more?

GerritForge.com/contact

