ESS - Backup & Restore Guide

Introduction

Welcome, ESS Administrators. This guide is crafted for your role, focusing on the pragmatic aspects of securing crucial data within the Element Server Suite (ESS). ESS integrates with external PostgreSQL databases and persistent volumes and is deployable in standalone or Kubernetes mode. To ensure data integrity, we recommend including valuable, though not strictly consistent, data in backups. The guide also addresses data restoration and a straightforward disaster recovery plan.

Software Overview

ESS provides Synapse and Integrations which require an external PostgreSQL and persistent volumes. It offers standalone or Kubernetes deployment.

- If you are using Standalone deployment, please refer to Single-node Storage & Backup Guidelines.
- If you are using Kubernetes deployment, we strongly recommend to leverage your own cluster backup solutions for effective data protection.

You'll find below a description of the content of each component data and db backup.

Synapse

1. Synapse deployments creates a PVC named `<element deployment cr name>-synapse-media`. It contains all users medias (avatar, photos, videos, etc). It does not need strict consistency with database content, but the more in sync they are, the more medias can be correctly synced with rooms state in case of restore.
2. Synapse requires an external postgressql database which contains all the server state.

Adminbot

1. Adminbot integration creates a PVC named `<element deployment cr name>-adminbot`. It contains the bot decryption keys, and a cache of the adminbot logins.

Auditbot

1. Auditbot integration creates a PVC named `<element deployment cr name>-auditbot`. It contains the bot decryption keys, and a cache of the adminbot logins.
2. Auditbot store the room logs of your organization either in an S3 Bucket or the aforementioned PVC. Depending on the critical nature of being able to provide room logs for audit, you need to properly backup your S3 Bucket or the PVC.

Matrix Authentication Service

1. Matrix Authentication Service requires an external postgresql database. It contains the homeserver users, their access tokens and their Sessions/Devices.

Sliding Sync

1. Sliding Sync requires an external postgresql database. It contains Sliding Sync running state, and data cache. The database backup needs to be properly secured. This database needs to be backed-up to be able to avoid UTDs and initial-syncs on a disaster recovery.

Sydent

1. Sydent integration creates a PVC named -sydent. It contains the integration SQLite database.

Integrator

1. Integrator requires an external postgresql database. It contains information about which integration was added to each room.

Bridges (XMPP, IRC, Whatsapp, SIP, Telegram)

1. The bridges require each an external postgresql database. It contains mapping data between Matrix Rooms and Channels on the other bridge side.

Backup Policy & Backup Procedure

There is no particular prerequisite to do before executing an ESS backup. Only Synapse and MAS Databases should be backed up in sync and stay consistent. All other individual components can be backed up on it's own lifecycle.

Backups frequency and retention periods must be defined according to your own SLAs and SLIs.

Data restoration

The following ESS components should be restored first in case of complete restoration. Other components can be restore on their distinctively, on their own time :

1. Synapse Postgresql database
2. Synapse medias
3. Matrix Authentication Service database (if installed)
4. Restart Synapse & MAS (if installed)
5. Restore and restart each individual component

Disaster Recovery Plan

In case of disaster recovery, the following components are critical for your system recovery :

1. **Synapse Postgresql database** is critical for Synapse to send consistent data to other servers, integrations and clients.
2. **Synapse keys** configured in ESS configuration (Signing Key, etc) are critical for Synapse to start and identify itself as the same server as before.

3. **Matrix Authentication Service Postgresql database** is critical for your system to recover your user accounts, their devices and sessions.

The following systems will recover features subsets, and might involve reset & data loss if not recovered :

1. **Synapse media storage** : Users will loose their Avatars, and all photos, videos, files uploaded to the rooms wont be available anymore
2. **Adminbot & Auditbot data** : The bots will need to be renamed for them to start joining all rooms and logging events again
3. **Sliding Sync** : Users will have to do an initial-sync again, and their encrypted messages will display as "Unable to decrypt" if its database cannot be recovered
4. **Integrator** : Integrations will have to be added back to the rooms where they were configured. Their configuration will be desynced from integrator, and they might need to be reconfigured from scratch to have them synced with integrator.

Security Considerations

Some backups will contain sensitive data, Here is a description of the type of data and the risks associated to it. When available, make sure to enable encryption for your stored backups. You should use appropriate access controls and authentication for your backup processes.

Synapse

Synapse media and db backups should be considered sensitive.

Synapse media backups will contain all user medias (avatar, photos, video, files). If your organization is enforcing encrypted rooms, these medias will be stored encrypted with each user e2ee keys. If you are not enforcing encryption, you might have media stored in cleartext here, and appropriate measures should be taken to ensure that the backups are safely secured.

Synapse postgresql backups will contain all user key backup storage, where their keys are stored safely encrypted with each user passphrase. Synapse DB will also store room states and events. If your organization is enforcing encrypted rooms, these will be stored encrypted with each user e2ee keys.

Adminbot

Adminbot PV backup should be considered sensitive.

Any user accessing it could read the content of your organization rooms. Would such an event occur, revoking the bot tokens would prevent logging in as the adminbot and stop any pulling of the room messages content.

Auditbot

Auditbot PV backup should be considered sensitive.

Any user accessing it could read the content of your organization rooms. Would such an event occur, revoking the bot tokens would prevent logging in as the auditbot and stop any pulling of the room messages content.

Logs stored by the auditbot for audit capabilities are not encrypted, so any user able to access it will be able to read any logged room content.

Sliding Sync

Sliding-Sync DB Backups should be considered sensitive.

Sliding-Sync database backups will contain Users Access tokens, which are encrypted with Sliding Sync Secret Key. The tokens are only refreshed regularly if you are using Matrix Authentication Services. These tokens give access to user messages-sending capabilities, but cannot read encrypted messages without user keys.

Sydent

Sydent DB Backups should be considered sensitive.

Sydent DB Backups contain association between user matrix accounts and their external identifiers (mails, phone numbers, external social networks, etc).

Matrix Authentication Service

Matrix Authentication Service DB Backups should be considered sensitive.

Matrix Authentication Service database backups will contain user access tokens, so they give access to user accounts. It will also contain the OIDC providers and confidential OAuth 2.0 Clients configuration, with secrets stored encrypted using MAS encryption key.

IRC Bridge

IRC Bridge DB Backups should be considered sensitive.

IRC Bridge DB Backups contain user IRC passwords. These passwords give access to users IRC account, and should be reinitialized in case of incident.

Revision #10
Created 26 February 2024 09:57:58 by Gaël Goinvic
Updated 6 November 2024 13:20:31 by Kieran Mitchell Lane