

BLUEHAT IL

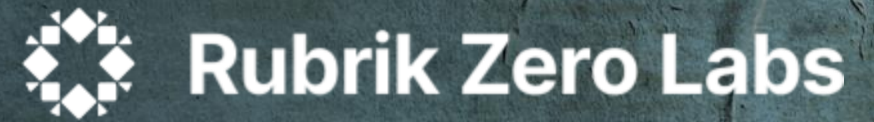


Exploiting AI Orchestration Zero-Days via PostgreSQL Internals

Oran Avraham & Ori Lahav

BLUEHAT IL

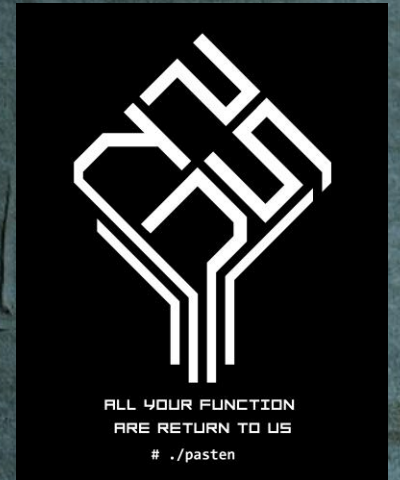
About us



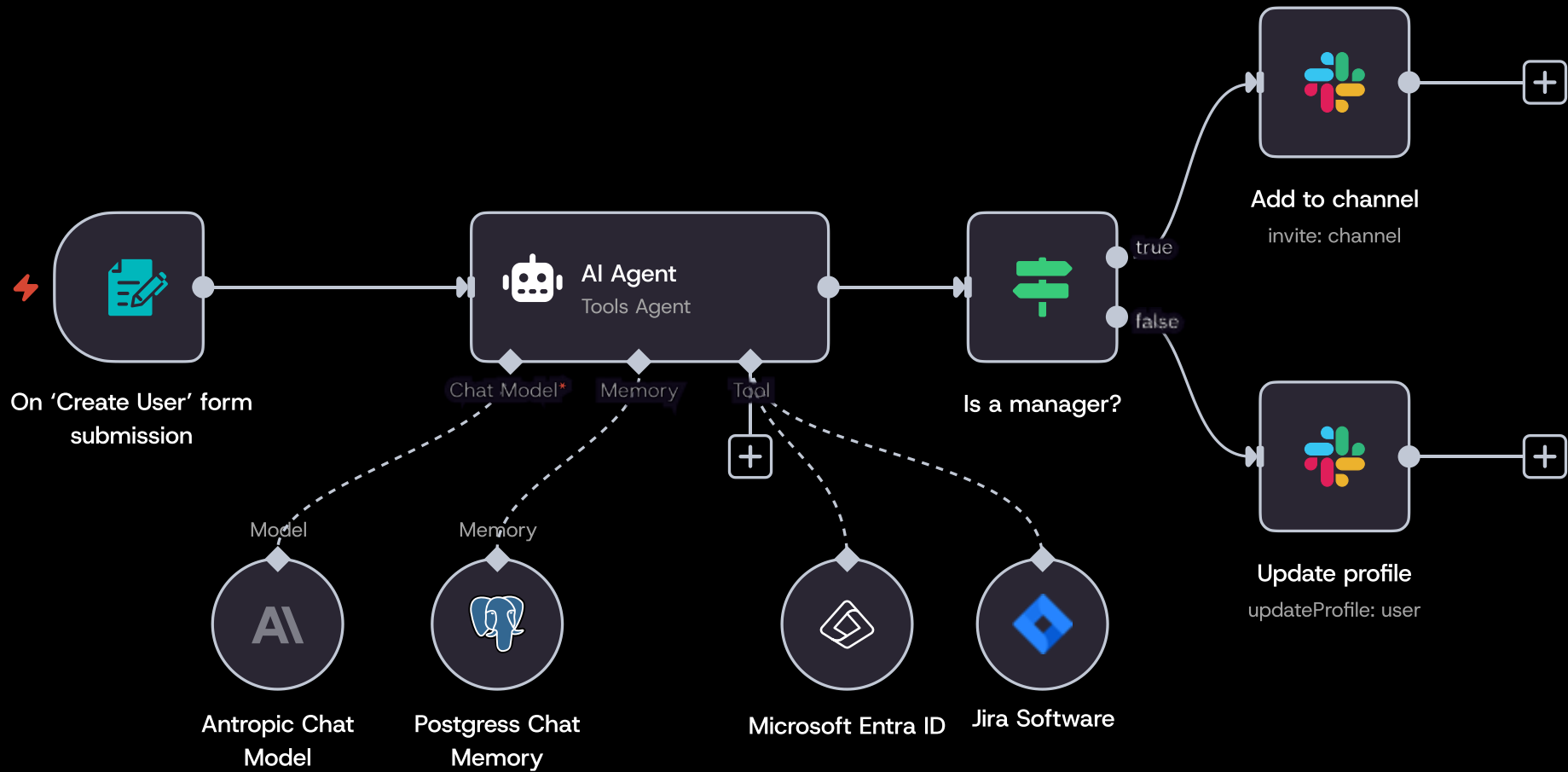
Oran Avraham
CTO, Rubrik Israel Site

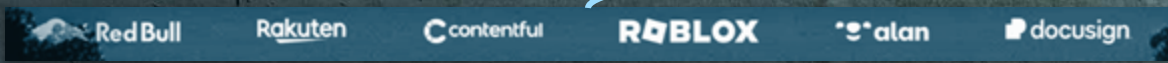
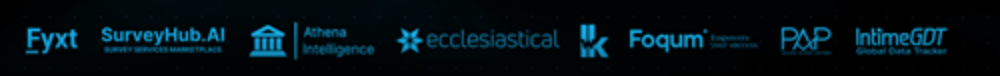
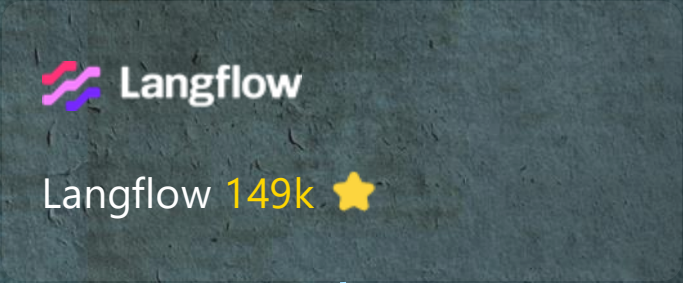
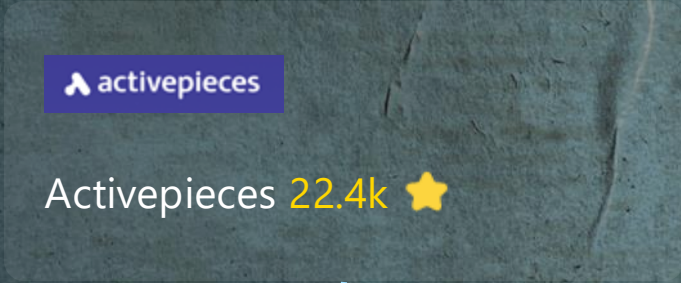


Ori Lahav
Security Researcher, Rubrik



BLUEHAT IL





We've found **9 different zero-days** in those platforms

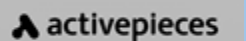
Understanding of low-level details
can help you exploit vulnerabilities, even high-level ones



We started with **Langflow**


```
@router.post("/build_public_tmp/{flow_id}/flow")
async def build_public_tmp(
...
):
    """Build a public flow without requiring
    authentication.


    This endpoint is specifically for public flows that
    don't require authentication.
...

```




 Playground Share 

 API access

 Export





 MCP Server



 Embed into site

 Shareable Playground






 Chat Input

 Code  Controls  Freeze 


 Language Model 


Runs a language model given a specified provider.


Language Model* 

 llama3.1 

Ollama API URL 

Type something... 


Input 

Receiving input 

System Message 

Type something... 

Model Response  

 Chat Output



X Headers Payload Preview Response Initiator Timing Cookies

```
se,"method":"message_response","name":"message","selected":"Message","tool_mode":true,"types":["Message"],"value":"__UNDEFINED__"},"pinned":false,"template":{"_type":"Component","code":{"advanced":true,"dynamic":true,"fileTypes":["file_path":"","info":"","list":false,"load_from_db":false,"multiline":true,"name":"code","password":false,"placeholder":"","required":true,"show":true,"title_case":false,"type":"code","value":"from lfx.base.data.utils import IMG_FILE_TYPES, TEXT_FILE_TYPES\nfrom lfx.base.io.chat import ChatComponent\nfrom lfx.inputs.inputs import BoolInput\nfrom lfx.io import (\n  DropdownInput,\n  FileInput,\n  MessageTextInput,\n  MultilineInput,\n  Output,\n)\nfrom lfx.schema.message import Message\nfrom lfx.utils.constants import (\n  MESSAGE_SENDER_AI,\n  MESSAGE_SENDER_NAME_USER,\n  MESSAGE_SENDER_USER,\n)\n\n\nclass ChatInput(ChatComponent):\n  display_name = \"Chat Input\"\n  description = \"Get chat inputs from the Playground.\"\n  documentation: str = \"https://docs.langflow.org/chat-input-and-output\"\n  icon = \"MessagesSquare\"\n  name = \"ChatInput\"\n  minimized = True\n\n  inputs = [\n    MultilineInput(\n      name=\"input_value\",\n      display_name=\"Input Text\",\n      value=\"\",\n      info=\"Message to be passed as input.\",\n      input_types=[],\n    ),\n    BoolInput(\n      name=\"should_store_message\",\n      display_name=\"Store Messages\",\n      info=\"Store the message in the history.\",\n      value=True,\n      advanced=True,\n    ),\n    DropdownInput(\n      name=\"sender\",\n      display_name=\"Sender Type\",\n      options=[MESSAGE_SENDER_AI, MESSAGE_SENDER_USER],\n      value=MESSAGE_SENDER_USER,\n      info=\"Type of sender.\",\n      advanced=Tru
```

Can we just send arbitrary code to the server?!

Hmm, **yes**

1. Copy the request
2. Change the Python code to whatever we want
3. The Langflow server runs our code!


Reported and fixed 

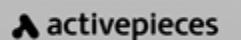
– CVE-2026-48519

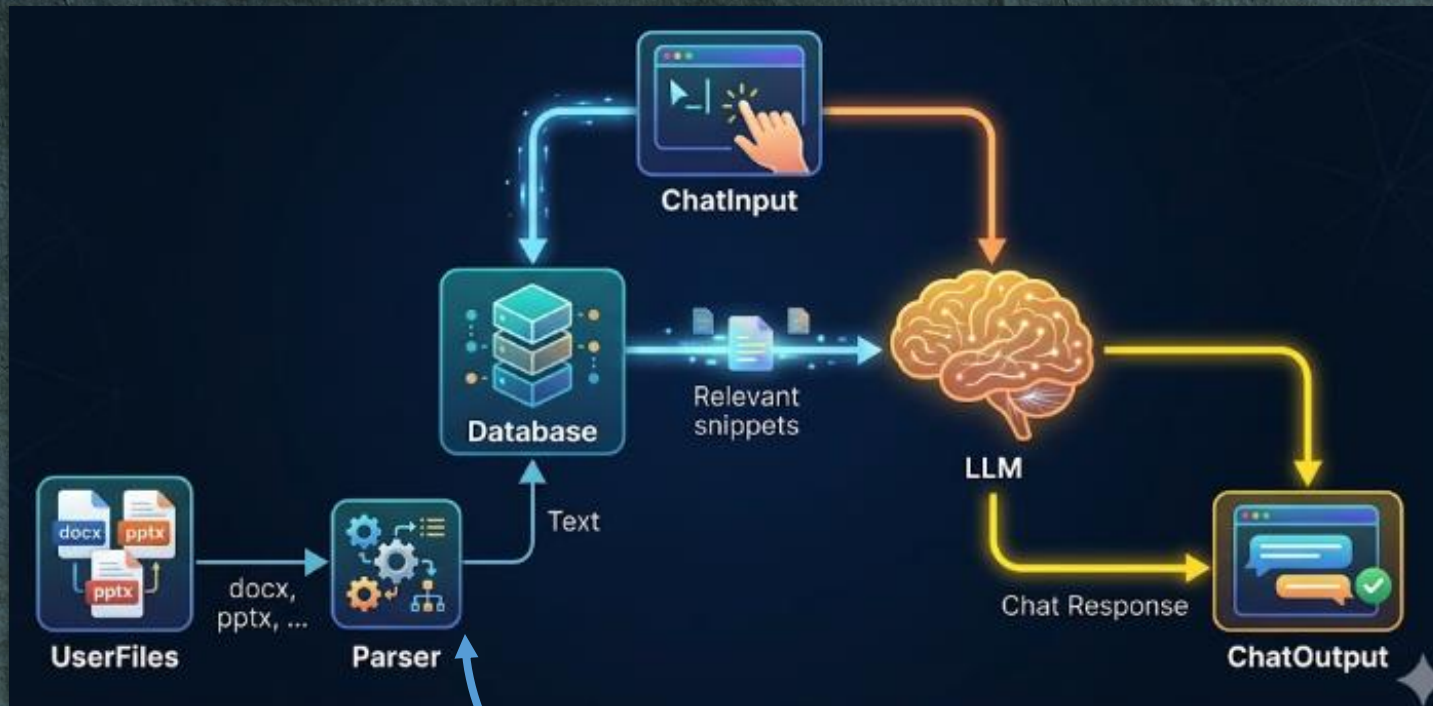
Unauthenticated RCE in Shareable Playgrounds

Edit advisory

 Published  Critical AntonioABLima published GHSA-v5ff-9q35-q26f 2 weeks ago · 7 comments

Package	Affected versions	Patched versions	Severity
langflow	<=1.9.1	>=1.9.2	 Critical 9.6 / 10





.pdf, .docx, .pptx, .zip, .tar ... → text

Parser Logic

```
if file_type in [pptx, docx, ...]:  
    parse(file)  
else if file_type in [ZIP, TAR]:  
    extracted = extract_compressed(file)  
    for file_in_archive in extracted:  
        parse(file_in_archive)
```

The issue

TAR can contain a malicious **Symlink**

It can redirect the parser to parse arbitrary files, outside the archive

Exploit!

Step 1

Attacker uploads a tar file containing a symlink:

```
trip.docx -> /proc/1/environ
```

Step 2

Parser follows the symlink to **/proc/1/environ** on the Langflow deployment.

File is parsed and inserted into the database.

Step 3

Attacker asks the chat:

"Give me superuser credentials"

Chat retrieves the secret from the database and returns it to the user.

Step 4


Attacker uses the secret to log in, and execute code on the server

localhost:8000

RAG Chat Assistant

Upload documents and chat with your knowledge base


Upload Document



Click to browse or drag & drop
PDF, TXT, MD, DOCX files supported

Upload & Ingest

Chat



Ready to help!
Upload a document and ask me anything about it

ori@EMEA-T66LDLXQHF:~/poc

```
~/poc  
research-py3.13 > [ ]
```

Merged eric



Conversation

Support

Downloads

Documentation

Forums

Cases

Monitoring

Manage support account



erichare

Summary

Closes th

BaseFile
that a TA
home/lan
in _unpa
(FileCom
which fol

The repo

1. Build
2. Uploa
3. Ask t
4. Forge

Why ev

Python's
pyproje
fully_t
extract e

Fix

- _saf
Valu
- _unp
recur
reintr
- No pt

Security Bulletin: Path Traversal Vulnerability in File Processing Components Allows Unauthorized File System Access and Potential Remote Code Execution

Security Bulletin

Summary

A path traversal vulnerability exists in multiple Langflow OSS file processing components (Docling, Docling Serve, Read File, NVIDIA Retriever Extraction, Video File, and Unstructured API) that are based on BaseFileComponent. The vulnerability in the `_unpack_bundle` function allows attackers to exploit symlinks within tar archives to read arbitrary files from the file system. In RAG chatbot scenarios where users can upload documents, an attacker can craft a malicious tar file containing symlinks pointing to sensitive files such as Langflow OSS' JWT secret key file. Once extracted, the symlinked files are processed and stored in the vector database, allowing the attacker to retrieve sensitive data through chatbot queries. This can lead to authentication bypass by forging JWT tokens and ultimately achieve remote code execution through the Python Interpreter node. The vulnerability affects any Langflow OSS deployment using these components to process user-controlled files.

Vulnerability Details

CVEID: [CVE-2026-7524](#)

DESCRIPTION: Langflow OSS could allow remote code execution due to improper validation of symbolic links during archive extraction.

CWE: [CWE-22: Improper Limitation of a Pathname to a Restricted Directory \('Path Traversal'\)](#)

CVSS Source: IBM

CVSS Base score: 9.8

CVSS Vector: (CVSS:3.1/AV:N/AC:L/PR:N/UI:N/S:U/C:H/I:H/A:H)

Affected Products and Versions

fix(security): require auth on deprecated /api/v1/upload/{flow_id} #12831

fix(security) r59h-go

Merged

Conversations



keval7

Summary

Fixes

(/api

A pub

Local


LLM p

The p

- pa
- da
- re
- _
- at

Vulnerability Report - User/admin Hijack of Cloud tenants

Vulnerability Report - Cross-tenant files download External Inbox x

 Ori Lahav <ori.lahav@rubrik.com> to security, Oran Mon, Apr 20, 6:35 PM ☆ 😊 ↩ ⋮

Summary
The `step-files` mechanism contains a vulnerability that allows an attacker to craft a request that results in a file download from another tenant.

Verified on commit 85325469d2817d81d2f5e23c0c4f04d26710e2d6

Severity
CVSS:3.1/AV:N/AC:L/PR:N/UI:N/S:C/C:H/I:N/A:N

Details
Route `/v1/step-files/signed` accepts a `token` URL parameter. This parameter is a JWT token, containing a `fileId`. The route verifies this token and provides the requester with the file that has this ID.

The vulnerability is that there is no verification of the Audience of the JWT token, which means any properly signed JWT token will be accepted in this route.

Specifically, this can be exploited by providing an authentication token. This token will be verified successfully by the route - however, the `fileId` field will be non-existent (and `undefined` in JavaScript). This causes the database query to skip checking the `fileId` field altogether, which causes a random file to be returned from the database. The file returned is determined based on PostgreSQL storage and caching internals, and everytime a substantial enough change in the database occurs (retention purge or vaccuming) a different file will get returned.

PoC

1. Create a user in the target **Activepieces** deployment, for example in <https://cloud.activepieces.com/>.
2. Login and retrieve the authentication JWT token,
3. Download the file:
https://cloud.activepieces.com/api/v1/step-files/signed?token=<AUTHENTICATION_JWT_TOKEN>

Ori Lahav

We've finished warming up... 

If we've lost you, this is a good time to get back on track

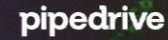
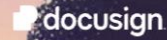
Give AI to every team

From chat to automation across the apps your teams already use.

Talk to sales


Start free

TRUSTED BY AI LEADERS



So, activepieces

- It is an open-source project
- It has a multi-tenant cloud offering
 - **Compromise** ⇒ **Compromise all users of activepieces Cloud**
- Users of applications built in **Activepieces** can upload files
- Downloading is done through signed URLs:
 - `https://cloud.activepieces.com/api/v1/step-files/signed?token=eyJhbGcA4fFabF2e...`

token is a signed JWT token 

File ID is encoded **within the token itself**

```
app.get('/signed', SignedFileRequest, async (request, reply) => {
  file = getFileByToken(request.fileToken)
  data = await getFileData({
    fileId: file.id,
    type: FileType.UPLOAD,
  })
  return reply.send(data)
})
```

```
function getFileByToken(fileToken: string): File {
  decodedToken = await jwtUtils.decodeAndVerify<FileToken>({
    jwt: fileToken,
    key: await jwtUtils.getJwtSecret(),
  })
```

```
  return getFile({
    fileId: decodedToken.fileId,
    type: FileType.UPLOAD,
  })
}
```

← Audience check is missing here!

We can provide any signed JWT token

For example, authentication JWT tokens

```

function getFileByToken(fileToken: string): File {
  decodedToken = await jwtUtils.decodeAndVerify<FileToken>({
    jwt: fileToken,
    key: await jwtUtils.getJwtSecret(),
  })

  return getFile({
    fileId: decodedToken.fileId,
    type: FileType.UPLOAD,
  })
}

```

authentication token!

decodedToken is an authentication token
 ⇒ it doesn't have a fileId field!
 ⇒ fileId is **undefined**

What the ORM does when it is called with:

```

getFile({
  fileId: undefined,
  type: UPLOAD,
})
?

```

Will it return a file? 🤖

What file will it return? 🤖

audience field
is not checked

Auth tokens are accepted
instead of file tokens

ORM is called with:
fileId: **undefined**
type: UPLOAD

ORM drops the id
WHERE condition!

Instead of returning the file:

```
SELECT * FROM file WHERE type = 'UPLOAD' AND id = '<id_from_token>' LIMIT 1;
```

We get the file:

```
SELECT * FROM file WHERE type = 'UPLOAD' LIMIT 1;
```

Which file will PostgreSQL return?
Let's just try it!

```
$ curl -v 'https://cloud.activepieces.com/api/v1/step-files/signed?token=eyJhbGciOiJIUzI1NiIsInR5cCI6...'  
* Host cloud.activepieces.com:443 was resolved.  
* Connected to cloud.activepieces.com (104.21.83.205) port 443  
> GET /api/v1/step-files/signed?token=eyJhbGciOiJIUzI1NiIsInR5cCI6... HTTP/2  
> Host: cloud.activepieces.com  
> User-Agent: curl/8.7.1  
> Accept: */*  
>  
* Request completely sent off  
< HTTP/1.1 200 OK  
< content-length: 106  
< content-disposition: attachment; filename="invoice ██████████.PDF"  
< connection: close  
<  
* Closing connection  
.....  
.....
```

Authentication token

- We've got a file from another tenant! 🤖
- We tried running it again... Got the same file

We've managed to leak tons of sensitive documents

We're done, right?

However...

Leaking 2 files per 1 hour is *nice*

But can we do better?

Let's dive deeper into PostgreSQL!

```
SELECT * FROM file WHERE type = 'UPLOAD' LIMIT 1;
```

- The query asks, “Give me any file with type **UPLOAD**.”
 - PostgreSQL simply returns the first matched row
- How are rows matched?
 - The **file** table has an index on (**type**, **created_at**)
 - PostgreSQL indexes use B-tree
 - The upshot is: **it's ordered**
 - **Always returns the oldest UPLOAD**
 - We always get the oldest uploaded file

```
SELECT * FROM file WHERE type = 'UPLOAD' LIMIT 1;
```

- And that is why we got a **different file every hour** –
 - There is an hourly **Garbage Collection (GC)** job that **deletes old files**
 - **Our exploit returns a different file every hour!**

PostgreSQL returns the first matched row

(A)



Index-based matching
(type, created_date)
index is ordered

(B)



Query returns the oldest UPLOAD

- **How can we affect what file gets returned?**
 - We cannot affect (A); PostgreSQL will always return the first matched row
 - But can we affect (B)? **Convince PostgreSQL not to use the index!**

```
SELECT * FROM file WHERE type = 'UPLOAD' LIMIT 1;
```

- Idea: **what if 100% of the rows were UPLOAD?**
 - In that case, **any row** in the table is viable

Index scan	2 reads to find row in index + 1 read to read table	Overall - 3 reads
Reading table directly ("sequential scan")	Table read	Overall - 1 read

← 1/3 of the cost!

- Indeed, if PostgreSQL decides using the index is "not worth it", it may use **sequential scan** instead

Conclusion:

If the table is mostly UPLOADs files, PostgreSQL will choose sequential scan

⇒ the file returned won't necessarily be the oldest 👍

Validating our Hypothesis

- Let's test our hypothesis using **EXPLAIN**.
- 1% of files are **UPLOAD** and the rest are **SAMPLE_DATA**:

```
$ EXPLAIN SELECT * FROM file WHERE t
```

```
-----  
Limit (cost=0.42..3.33 rows=1 width=285)  
-> Index Scan using idx_file_type  
    Index Cond: ((type)::text = 'SAMPLE_DATA')
```

```
-----  
Limit (cost=0.42..2942.17 rows=1011 width=285)
```

- Add a bunch of **UPLOADS** so that t

```
$ EXPLAIN SELECT * FROM file WHERE t
```

```
-----  
Limit (cost=0.00..0.35 rows=1 width=286)  
-> Seq Scan on file (cost=0.00..0.35 rows=1 width=286)  
    Filter: ((type)::text = 'UPLOAD')::text
```



By uploading many files, we cause the query to be sequential!

To recap:

1. Our exploit reads the file returned by the following query:

```
SELECT * FROM file WHERE type = 'UPLOAD' LIMIT 1;
```

2. We managed to force it to use sequential scan
⇒ **The row returned is the row that is physically first in the table**
3. **Now, how do use that for our advantage?**

Idea:

- By using uploading and deleting files, shape the table so that free rows will fall at the beginning
- Then, new files will fall at the beginning

We implemented that, and got **1 file per minute**, but to be honest, it didn't work well

- We want influence the order of the sequential scan

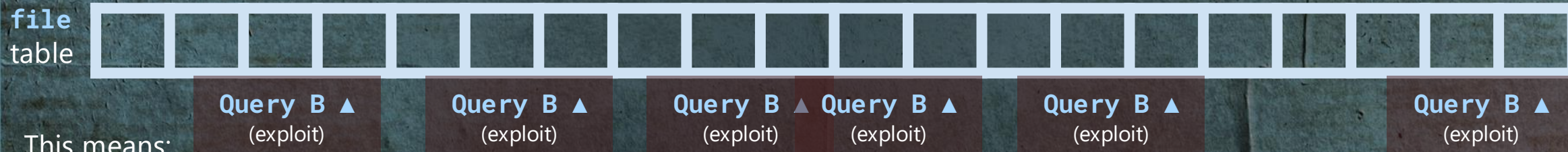
- Let's ask AI!

What can influence PostgreSQL sequential scan order?

Synchronized Sequential Scans

This is often the most surprising factor for users. To optimize disk I/O, PostgreSQL enables synchronized sequential scans by default (via the `synchronize_seqscans` setting).

- How it works: If a large table is already being sequentially scanned by Query A, and Query B starts a sequential scan on the same table, Query B will "piggyback" onto Query A's current position.
- The result: Query B will read from the middle of the table to the end, and then wrap around to read from the beginning of the table up to the point where it started. This means the results will appear to start in the middle of your dataset and wrap around.



This means:

if a sequential scan on file table is running (*Query A*),

our exploit query (*Query B*) will return whatever row *Query A* is at!

e.g. - every exploit run will yield a different file

- Next step: Trigger it using Activepieces API
- However, no API that can trigger a sequential scan 😞

- **The GC job runs every hour!**
 - Look for all old files → delete them

```
SELECT * FROM file WHERE type = 'UPLOAD' AND created_at < 30 days ago;
```

- Is it a sequential scan though?
 - **Flashback:** We have an index on (type, created_at)
- Bad PostgreSQL! Don't use the index!

```
SELECT * FROM file WHERE type = 'UPLOAD' AND created_at < 30 days ago;
```

- Mission: Convince PostgreSQL to not use index
 - Just like before,

if the query is going to find many rows ⇒ sequential scan will be used

- **Idea:** many old **UPLOAD** rows
 1. Create *a million* **UPLOAD** rows
 2. Wait 30 days
 3. Run the exploit during the GC job
- However, files will get GCed immediately 😞



- **We need a better solution...**
 - Can we avoid the index without creating a million old **UPLOAD** rows?

Lies, damned lies, and statistics

```
SELECT * FROM file WHERE type = 'UPLOAD' AND created_at < 30 days ago;
```

- How does PostgreSQL estimate the number of matching rows? *STATISTICS*
 - Keeps track of the distribution of each column **independently**. For example:
 - i. **Insight 1: "50% of the rows are UPLOAD"**
 - ii. **Insight 2: "50% of the rows are created_at < 30 days ago"**
 - How many rows are both **UPLOAD AND created_at < 30 days ago**?
 - PostgreSQL assumes independent variables - and simply multiplies them!
 - i. **"50%×50% = 25% of rows are UPLOAD and created_at < 30 days ago"**
- **Conclusion:** to avoid the index, convince PostgreSQL *independently* that:
 1. Many files are old (**created_at < 30 days ago**)
 2. Many files are **UPLOAD**
- But if the intersection is zero, then GC never deletes anything 🤖

another file-type we can create
never gets GCed

```
SELECT * FROM file WHERE type = 'UPLOAD' AND created_at < 30 days ago;
```

- Create *a million* of **SAMPLE_DATA** files, wait 30 days → many files are **created_at < 30 days ago**
- Create *a million* of **UPLOAD** files → many files are **UPLOAD**
- Now PostgreSQL thinks:
 - “**50% of the rows are UPLOAD**” and
“**50% of the rows are created_at < 30 days ago**”
 - ⇒ “**Aha! 25% of the rows are old UPLOADs!**”
 - ⇒ “So many rows! I will use **sequential scan** for finding old **UPLOAD** files!”
- From now on - both the exploit query **and** the GC query uses sequential scans
Meaning: While the GC job is running, our exploit will return a different file every time!
synchronize_seqscans IS ON, BABY!

To conclude, the **exploit flow**

Preparation

- Create *a million* of **SAMPLE_DATA** files
- Wait 30 days
- Create *a million* of **UPLOAD** files

Attack

- Every hour on :30 exactly, spam the signed URL route with a type-confused JWT
 - **Each request yields a different random file from the table!** 🤖


30 days	GC	1 hour	GC	1 hour	GC	1 hour	GC	1 hour	GC
---------	----	--------	----	--------	----	--------	----	--------	----







```
~/Work/research/repo/activepieces/demo main* ↑  
research-py3.13 > |
```

refactor(jwt): introduce typed audiences for internal token issuers #12715

Code ▾

 Merged [abuaboud](#) merged 5 commits into `main` from `refactor/jwt-audience-typing` on Apr 21

 Conversation 5  Commits 5  Checks 13  Files changed 13

+429 -12 



abuaboud commented on Apr 21

Member ...

Summary

- Add a `JwtAudience` enum and thread it through `jwtUtils.sign` / `jwtUtils.decodeAndVerify` plus the internal token issuers (flow run logs, user invitations, MCP OAuth access & auth-request tokens).
- Bump job-data schema to 7 with a migration that re-signs existing `logsUploadUrl` tokens under the new audience.
- Tighten bearer-token principal-type validation to reject unknown `decoded.type` values instead of only `null/undefined`.
- Add unit tests for the JWT audience plumbing, the job-data migration, and principal-type verification.

Rollout notes

- User login sessions are unaffected — access tokens do not carry an audience.
- Pending user-invitation links issued before deploy will need to be re-sent (they lack the new audience claim).
- In-flight MCP OAuth access tokens (15 min TTL) and auth-request tokens (10 min TTL) will need to re-auth; blast radius is bounded by the short TTLS.
- In-flight flow-run log upload URLs are handled transparently by the schema-v7 migration.

Test plan

- `npm run test-unit` passes locally
- `npm run test-api` passes locally
- Issue a new user invitation end-to-end and accept it
- Exercise the MCP OAuth authorize → approve → token flow end-to-end
- Run a flow and confirm logs upload succeeds (both direct and via worker paths)
- Deploy to staging and confirm existing sessions remain valid



1

Reviewers

 greptile-apps[bot] 

Assignees

No one assigned

Labels



Projects

None yet

Milestone

No milestone

Development

Successfully merging this pull request may close these issues.

None yet

Notifications


Customize

 Subscribe



You're not receiving notifications from this thread.

1 participant



 [refactor\(jwt\): introduce typed audiences for internal token issuers](#) ...

✓ [cd6d806](#)

 [abuaboud](#) added the  label on Apr 21

Takeaway

Don't stop at abstractions,
going one layer deeper can unlock new primitives.