

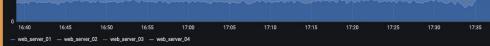
How NOT to use DuckDB



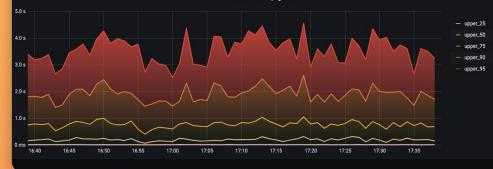
Sam Jewell Staff Software Engineer

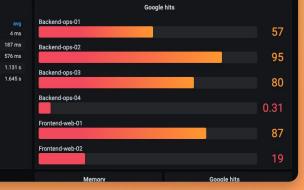
	size	filename	content							
	9050	/etc/passwd	##\x0A# User [Database\x0A#	\x0A# Note that th	nis file				
	🖯 Query	1 ្វ្មី Transt	form data 0							
D	ata source	🖶 testdata-DS	~ ⑦	> Query	MD = auto = 1415	Inter				
	B	(Expression)								
	Operation	SQL								
		1 SELECT size, filename, content								
		ROM read_blob('/e	tc/passwd');							





client side full page load





C-series

D-series

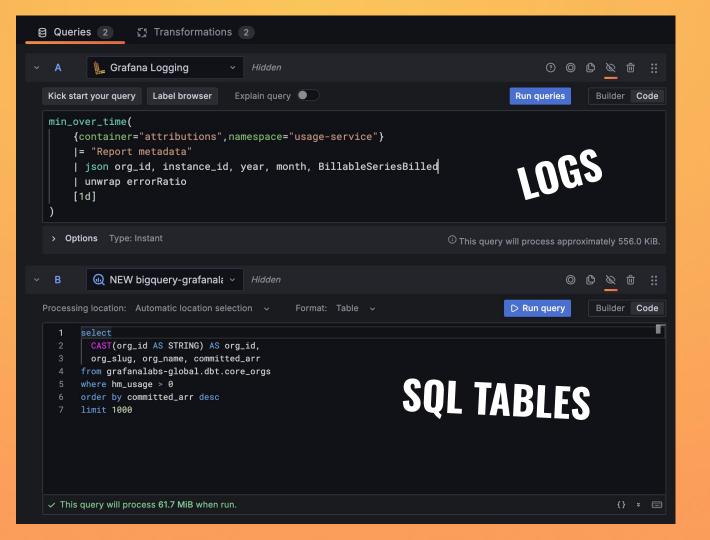
E-series

A-series

B-series



Grafana Labs' big tent strategy



METRICS TRACES JSON & CSV

Ø			Q Search or jump to		œ ೫+k	+ • 0			
	dashboard > Edit panel								
	Add another transformation								
Panel Title 58.5	Search for transformation								
58	View all Combine Calculate new fields Create new visualization Filter Perform spatial operations Reformat Reorder and rename								
57.5 57 58.5	Add field from calculation Use the row values to calculate a new field.	Concatenate fields Combine all fields into a single frame.	Config from query results Beta Set unit, min, max and more.	Convert field type Convert a field to a specified field type.	Create heatmap Alpha Generate heatmap data from source data.	Extract fields Parse fields from content (JSON, labels, etc).	Filter data by query refld Remove rows from the data based on origin query	Filter data by values Remove rows from the query results using user-defined filters.	
56 55.5 55	1 2 1+2		Cani II → Cani II ani Cani I → Cani II ani Cani I → Cani II ani	6 11 mm → 11 mmm 2	→	vote a b (art,b:22) → 11 2 (art) 3 4	+ ● →	a b 3 5 ⊽as2 → a b 2 5	
545 54 535	Filter fields by name Remove parts of the query results using a regex pattern.	Format time Alpha Set the output format of a time field	Group by Group data by a field value and create aggregate data.	Grouping to Matrix Summarize and reorganize data based on three fields.	Histogram Calculate a histogram from input data.	Join by field Combine rows from 2+ tables, based on a related field.	Join by labels Beta Flatten labeled results into a table joined by labels.	Labels to fields Group series by time and return labels or tags as fields.	
53 52.5 52		O yyyy→mn→d yyyy→mn→d yyyy→mn→d	10 2 2 2 2 2 2 2 2 2		•	$\frac{a}{1} + \frac{b}{1} = \frac{b}{2}$	$\begin{array}{c} \bullet & \bullet \\ \bullet & \bullet \\$		
51.5 51 50.5 50	Limit Limit the number of items displayed.	Lookup fields from resource Alpha Use a field value to lookup countries, states, or airports.	Merge series/tables Merge multiple series. Values will be combined into one row.	Organize fields by name Re-order, hide, or rename fields.	Partition by values Alpha Split a one-frame dataset into multiple series.	Prepare time series Stretch data frames from the wide format into the long format.	Reduce Reduce all rows or data points to a single value (ex. max, mean).	Regression analysis	
17:00 17:30 — A-series		Q	+ = → = = ₀	a b c → b d 1 2 3 2 1	□ → □ = + □ =	0 7 8 0 0 0 0 0 12 13 5 0 0 0 0 0 0 12 13 5 0 0 0 0 0 0		→ f(x)	
🖯 Query 🍈 🚺 Transform o									
	Rename fields by regex Rename parts of the query results using a regular expression and replacement pattern.	Rows to fields Beta Convert each row into a field with dynamic config.	Series to rows Merge multiple series. Return time, metric and values as a row.	Sort by Sort fields in a frame.	Spatial operations Alpha Apply spatial operations to query results.	Time series to table Beta Convert time series data to table rows so that they can be viewed in tabular or sparkline format.			
					<u>_</u>	~~~• •			

>	* _	Hi\ min_over_time({container="attributions",namespace="usage-service"} = "Report metadata" jso ?	0	¢	Ø	创	••• ••
>	В	😡 NEW bigquery-grafanal: 🗸 Hidden	0	¢	Ø	创	* * * * * *
	с	Expression ~	0	¢	0	创	•••
	Operati	on Choose ~					
	1	1 select					
	2	BillableSeriesBilled,					
	3	A.org_id,					
	4	org_slug,					
	5	committed_arr,					
	6	instance_id,					
	7	month,					
	8	year,					
	9	abs(Value) AS Error					
	10	from A					
	11	LEFT JOIN B					
	12	ON A.org_id = B.org_id					

Grafana Labs

KALTHER





security % * Customer security questions: create ...



🗨 Messages 🛛 Ə Add canvas 🛛 😂 Files 🕒 Bookmarks 🔗 Pins 🕂

View newer replies

26 September 2024 ~



todd 16:54

while debugging something else, i noticed that the SQL Expression uses the duckdb CLI under the hood and duckdb's CLI has a dot command that allows you to run shell commands. the SQL expression does some checks to prevent running anything that isn't SQL. i attempted to bypass the escaping of single quotes here, but i'm wondering if someone from #security could double check this

👩 🎥 😔 🏩 🧖 21 replies Last reply 4 months ago



•	Q 1.1.3 (stable) ~
.shell CMD ARGS	Run CMD ARGS in a system shell
.show	Show the current values for various settings
.singleline	Set single-line mode
.system CMD ARGS	Run CMD ARGS in a system shell

https://duckdb.org/docs/api/cli/dot_commands.html



read_blob(source)

Description Returns the content from source (a filename, a list of filenames, or a glob pattern) as a BLOB. See the <u>read_blob guide</u> for more details.

Example read_blob('hello.bin')

Result hello\x0A

https://duckdb.org/docs/sql/functions/blob.html#read_blobsource

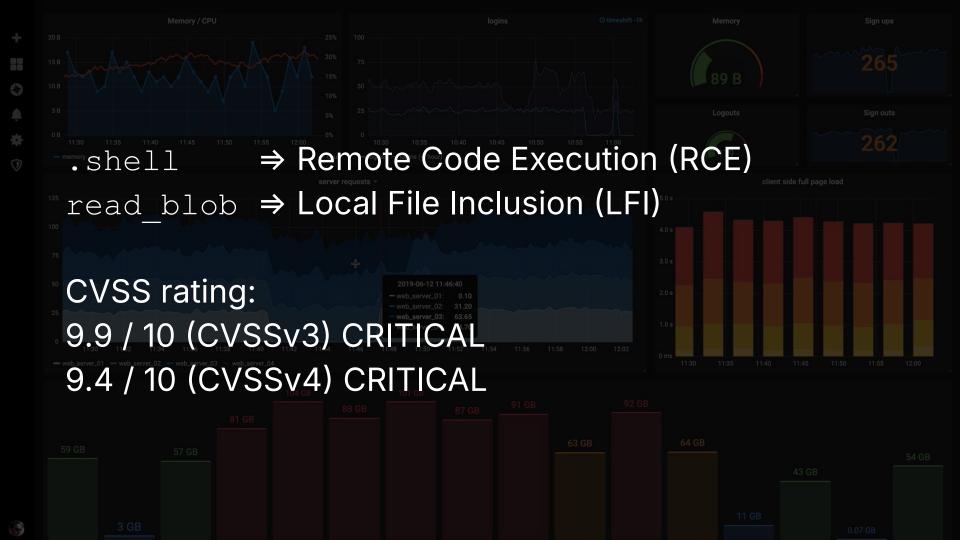


SELECT size, filename, content

FROM read_blob('/etc/passwd');

size	filename	content							
9050	/etc/passwd	##\x0A# User D	atabase\x0A# \	x0A# Note that th	is file is consu	Ited directly	only whe		
Query 1									
Data source	🚻 testdata-DS	~ ⑦	> Query	MD = auto = 1415	Interval = 15s	Query ins	pector		
~ В	(Expression)					© ⊈	₪ ::		
Operation	SQL								
	ELECT size, filen						_		
3	ROM read_blob('/e	tc/passwo');							

Q





- Security impact relatively small
 - Identified internally
 - Behind feature flag only ~2 customers enabled
 - Checked every log-line. No evidence of exploitation

• High effort to remedy

- 2-week SLO for Critical vulnerabilities
- Remove the feature entirely
- Roll to our hosted SAAS
- Rebuild Grafana Enterprise and Grafana OSS
- Disclose and share privately, with 2 week embargo
- We learned a lot
 - Blameless post-mortem



How did we get here?

6

• We implemented on the Grafana backend, which is written in Go

Documentation / Client APIs

Go

The DuckDB Go driver, go-duckdb, allows using DuckDB via the database/sql interface. For examples on how to use this interface, see the <u>official documentation</u> and <u>tutorial</u>.

Note

The go-duckdb project, hosted at <u>https://github.com/marcboeker/go-duckdb</u>, is the official DuckDB Go client.

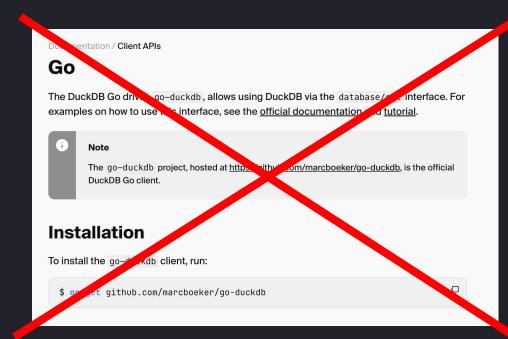
Installation

To install the go-duckdb client, run:

\$ go get github.com/marcboeker/go-duckdb

How did we get here?

- We implemented on the Grafana backend, which is written in Go
- But we can't use CGo





How did we get here?

- We implemented on the Grafana backend, which is written in Go
- But we can't use CGo
- So we used the CLI 😬



Avoid our mistake! Lesson 1

• Embed, don't use the CLI (or parse & validate)

Documentation / Operations Manual

Embedding DuckDB

CLI Client

The <u>Command Line Interface (CLI) client</u> is intended for interactive use cases and <u>not for embedding</u>. As a result, it has more features that could be abused by a malicious actor. For example, the CLI client has the ...sh feature that allows executing arbitrary shell commands. This feature is only present in the CLI client and not in any other DuckDB clients.

.sh ls

• Tip

Calling DuckDB's CLI client via shell commands is **not recommended** for embedding DuckDB. It is recommended to use one of the client libraries, e.g., <u>Python, R, Java</u>, etc.

Avoid our mistake! Lesson 2

• Disable file access

Documentation / Operations Manual / Securing DuckDB

Securing DuckDB

DuckDB is quite powerful, which can be problematic, especially if untrusted SQL queries are run, e.g., from publicfacing user inputs. This page lists some options to restrict the potential fallout from malicious SQL queries.

The approach to securing DuckDB varies depending on your use case, environment, and potential attack models. Therefore, consider the security-related configuration options carefully, especially when working with confidential data sets.

If you plan to embed DuckDB in your application, please consult the "Embedding DuckDB" page.

Reporting Vulnerabilities

If you discover a potential vulnerability, please report it confidentially via GitHub.

Disabling File Access

DuckDB can list directories and read arbitrary files via its CSV parser's <u>read_csv_function</u> or read text via the <u>read_text_function</u>. For example:

SELECT *
FROM read_csv('/etc/passwd', sep = ':');

۵

This can be disabled either by disabling external access altogether (enable_external_access) or disabling individual file systems. For example:





Thank you



Sam Jewell Ƴ@sojewell

Have more questions?

Join us at community.grafana.com

or Grafana public slack #grafana

Get involved: Image: state of the state of t