

EAV vs JSONB

...

Сравнение производительности поиска

О чем речь

❖ Гибкий атрибутивный состав

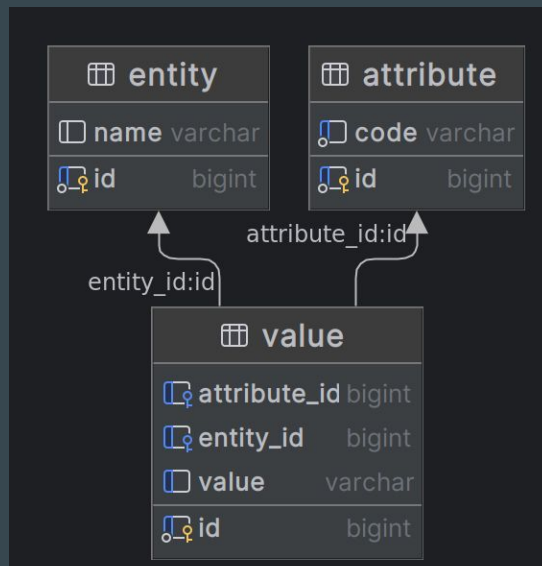
- Интернет-магазины
- Таск-трекеры и тикет-ориентированные системы

❖ Способы достижения в рамках PostgreSQL

- EAV
- JSONB

EAV (Entity-Attribute-Value)

- + Три таблицы, чтоб править всеми
 - + Гибкость
 - + Быстрая вставка / обновление атрибута
 - + Разработка проще
-
- Медленный поиск ?
 - Требуется больше места для хранения

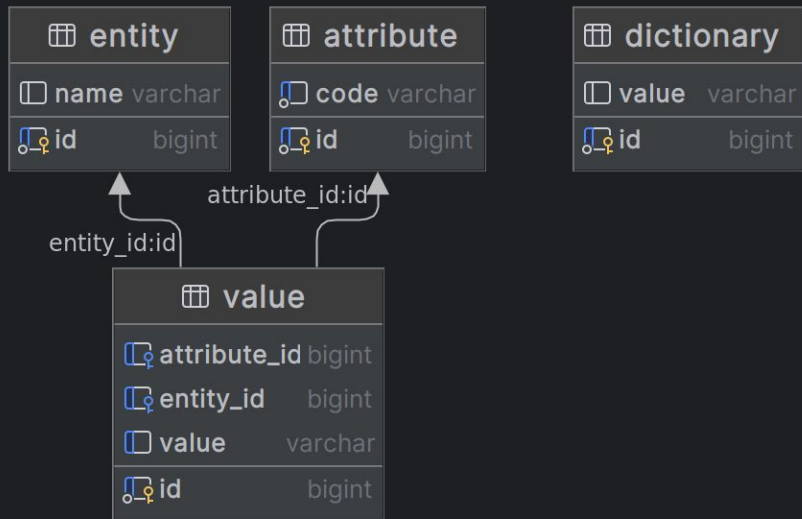


JSONB Column

- + Одна колонка, чтоб править всеми
 - + Гибкость
 - + Отсутствие JOIN / Быстрый поиск ?
 - + Простая и понятная схема данных
 - + Экономия объема хранилища
-
- Обновление медленнее
 - Манипулировать данными немного сложнее

entity	
name	varchar
properties	jsonb
id	bigint

Пререквизиты EAV. Схема данных



Entity	
id	name
1	name1
2	name2
...	


Attribute	
id	code
1	status
2	assignee
...	


Value			
id	attribute_id	entity_id	value
1	1	1	closed
2	1	2	open
1	2	1	user1
2	2	2	user2

Пререквизиты JSONB. Схема данных


entity


 **name** varchar

 **properties** jsonb

 **id** bigint

attribute

 **code** varchar

 **id** bigint

Entity

id	name	properties
1	name1	{ "status": "closed", "assignee": "user1" }
2	name2	{ "status": "open", "assignee": "user2" }

Attribute

id	code
1	status
2	assignee
...	

Наполнение тестовыми данными

EAV

- 10 млн сущностей
- 48 атрибутов (строки)

Итого:

- ❑ 10 млн записей в entity
- ❑ 480 млн записей в value

JSONB

- 10 млн сущностей
- 48 атрибутов (строки)

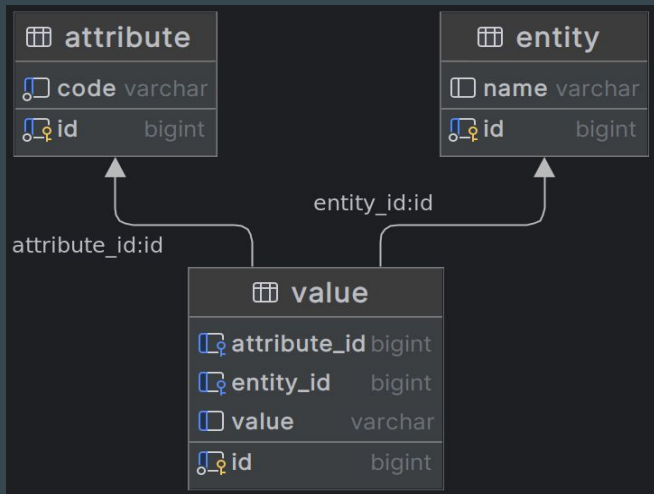
Итого:

- ❑ 10 млн записей в entity

Пререквизиты EAV. Типовой запрос

```
select e.id from entity e
  where (
    exists(select val.id from value val
           join attribute attr on val.attribute_id = attr.id
           where attr.code = 'branch'
                  and val.value = 'develop'
                  and e.id = val.entity_id)
    and exists(...)
    ...
  ) limit 100
```


Индексы EAV



- ❖ value.entity_id
- ❖ value.attribute_id
- ❖ value.value
- ❖ value.attribute id, value.value
- ❖ attribute.code

Прerequisites JSONB. Типовые запросы

```
select e.id
  from entity e
 where (
    e.properties ->> 'system' = 'ppts'
    and e.properties ->> 'application' = 'tasks' and ...
  )
limit 100
```

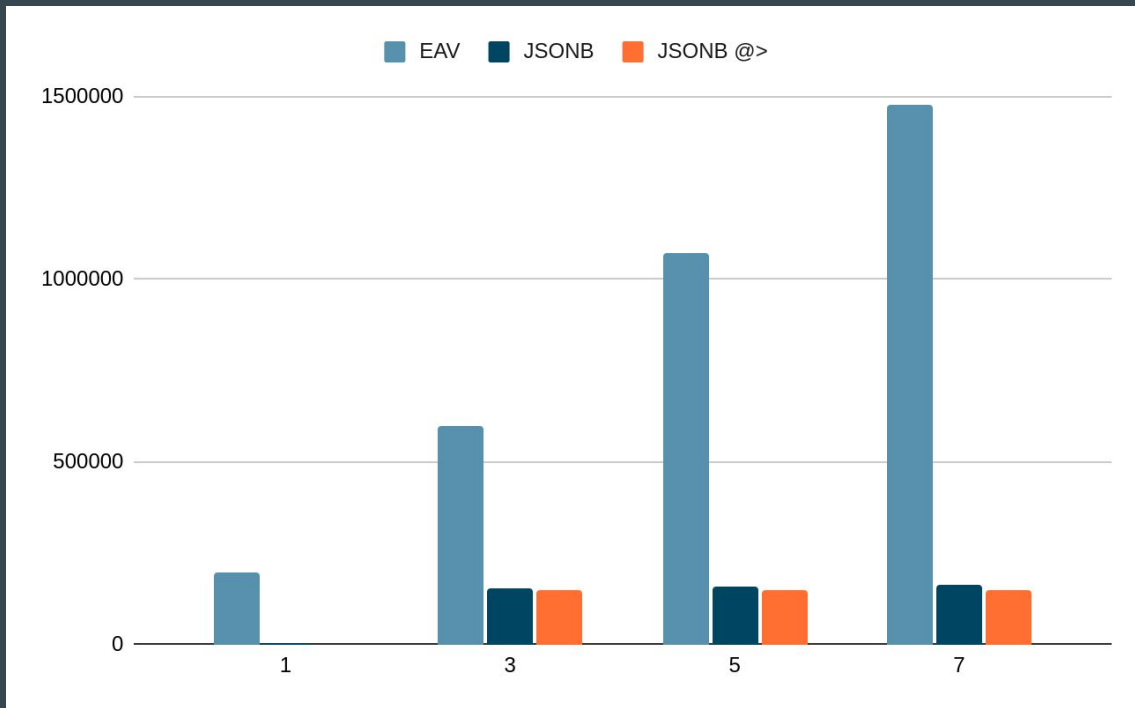
```
select e.id
  from entity e
 where (
    e.properties @> '{"system":"ppts","application":"tasks"}'
  )
limit 100
```

Индексы JSONB

entity	
name	varchar
properties	jsonb
id	bigint

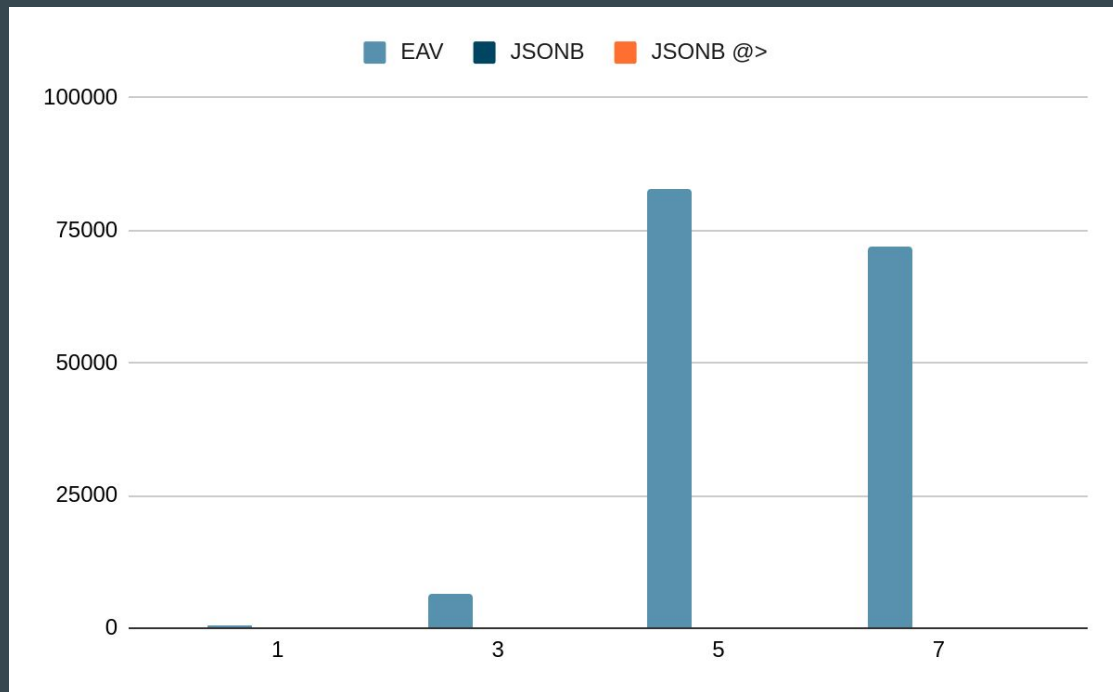
❖ value.properties (GIN)

Поиск по полному соответствию без индексирования 1m (cost)



Атрибуты	EAV	JSONB	JSONB @>
1	199117,61	3157,15	384,52
3	600191,62	154274,53	149076,21
5	1073758,57	158441,1	149076,21
7	1479373,62	162607,67	149076,21

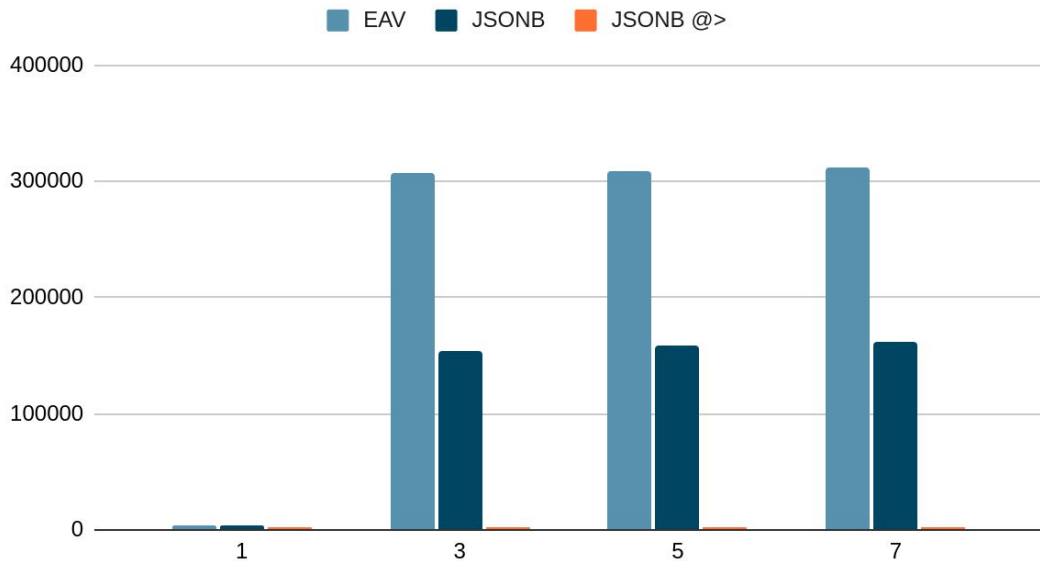
Поиск по полному соответствию без индексирования 1m (ms)



Атрибуты	EAV	JSONB	JSONB @>
1	416,914	0,789	0,767
3	6393,063	191,442	180,685
5	82809,985	213,778	204,627
7	71916,966	204,555	205,152

Поиск по полному соответствию с индексами 1m (cost)

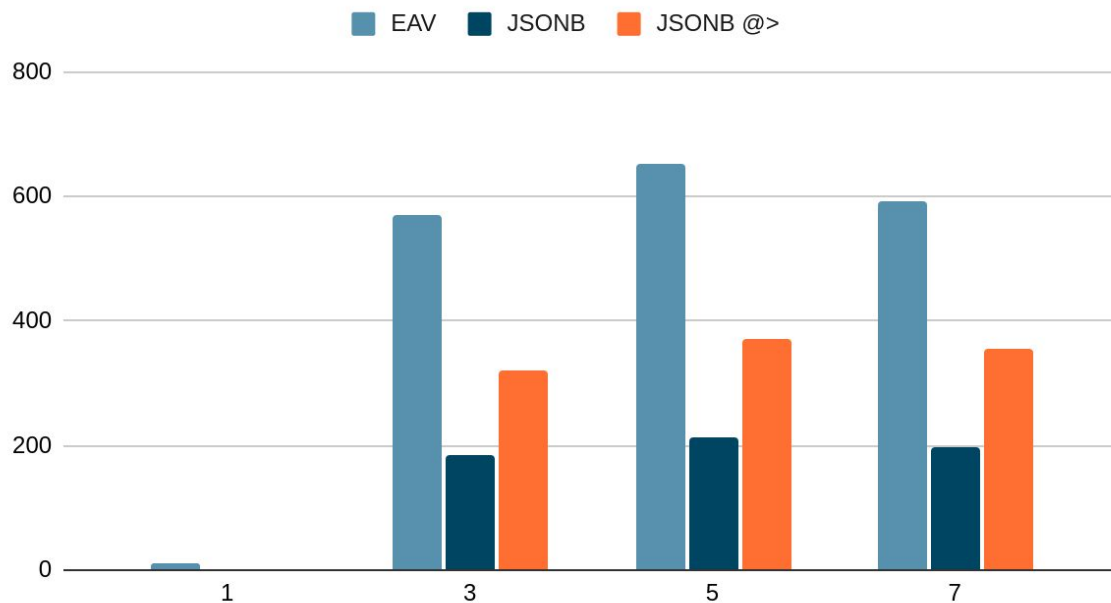
EAV, JSONB и JSONB @>



Атрибуты	EAV	JSONB	JSONB @>
1	3124,1	3157,16	1553,58
3	306644,88	154274,77	1553,58
5	308500,19	158441,43	1553,58
7	311827,53	162608,1	1553,58

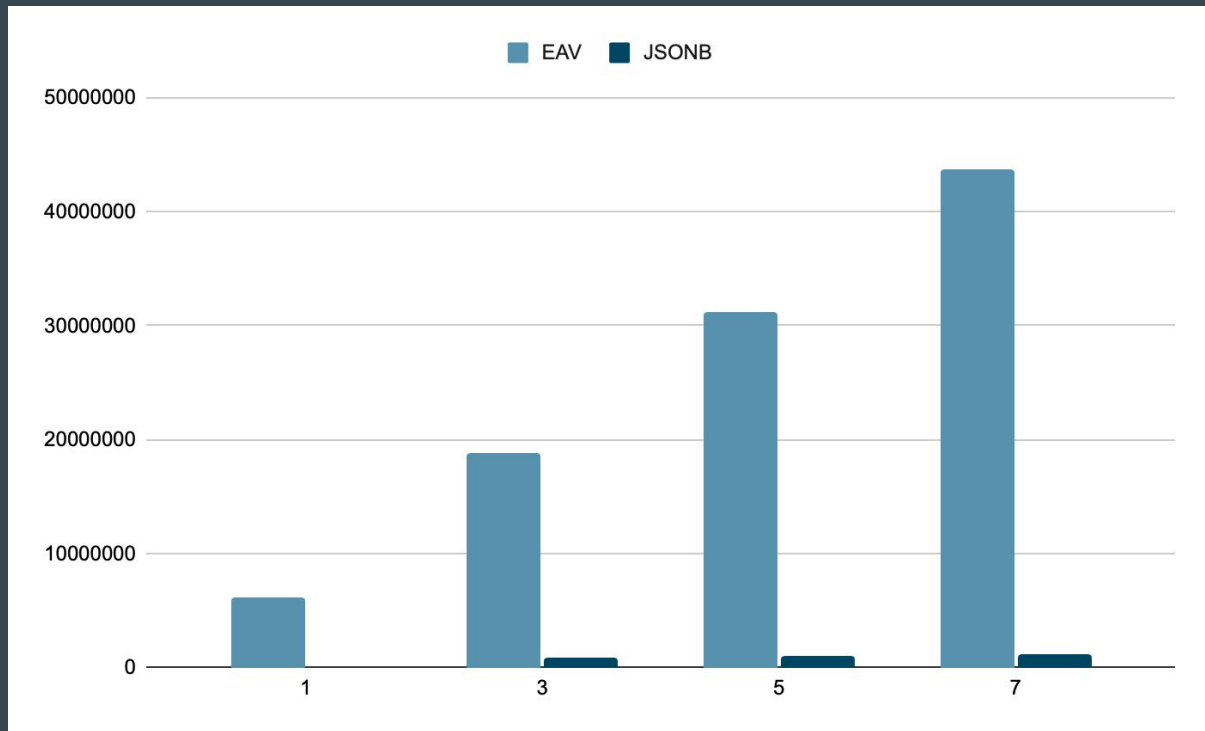
Поиск по полному соответствию с индексами 1m (ms)

EAV, JSONB и JSONB @>



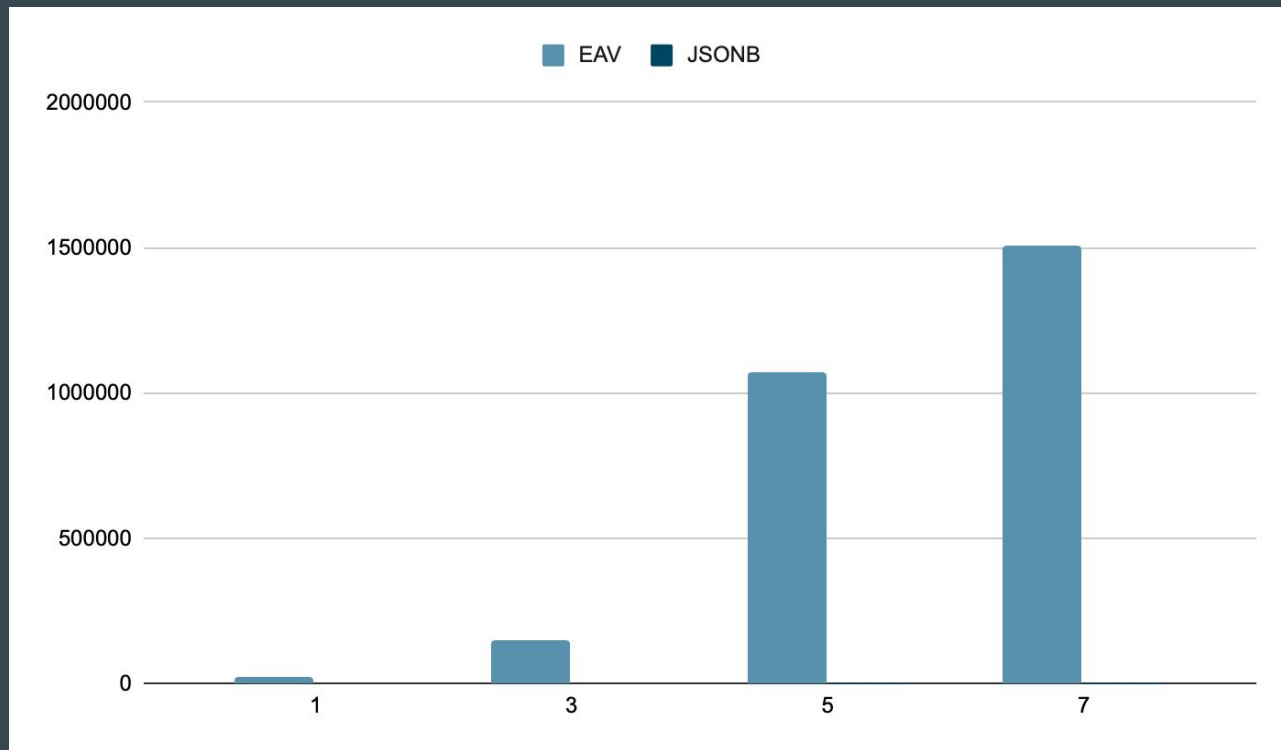
Атрибуты	EAV	JSONB	JSONB @>
1	12,036	0,857	0,764
3	568,651	185,863	320,124
5	652,043	212,772	370,602
7	591,99	196,868	356,617

Поиск по полному соответствию без индексирования 10m (cost)



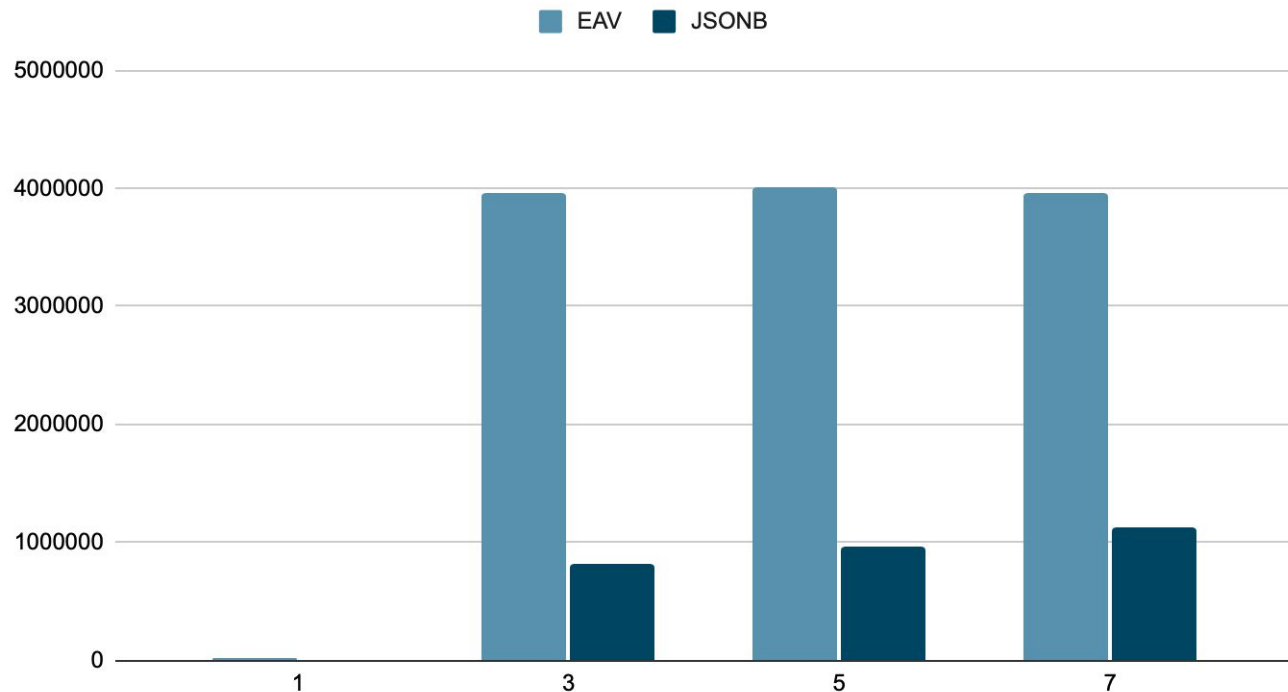
Атрибуты	EAV	JSONB
1	6225496,91	546,91
3	18763580,33	819452,09
5	31256831,97	969270,06
7	43703900,73	1119088,34

Поиск по полному соответствию без индексирования 10m (ms)



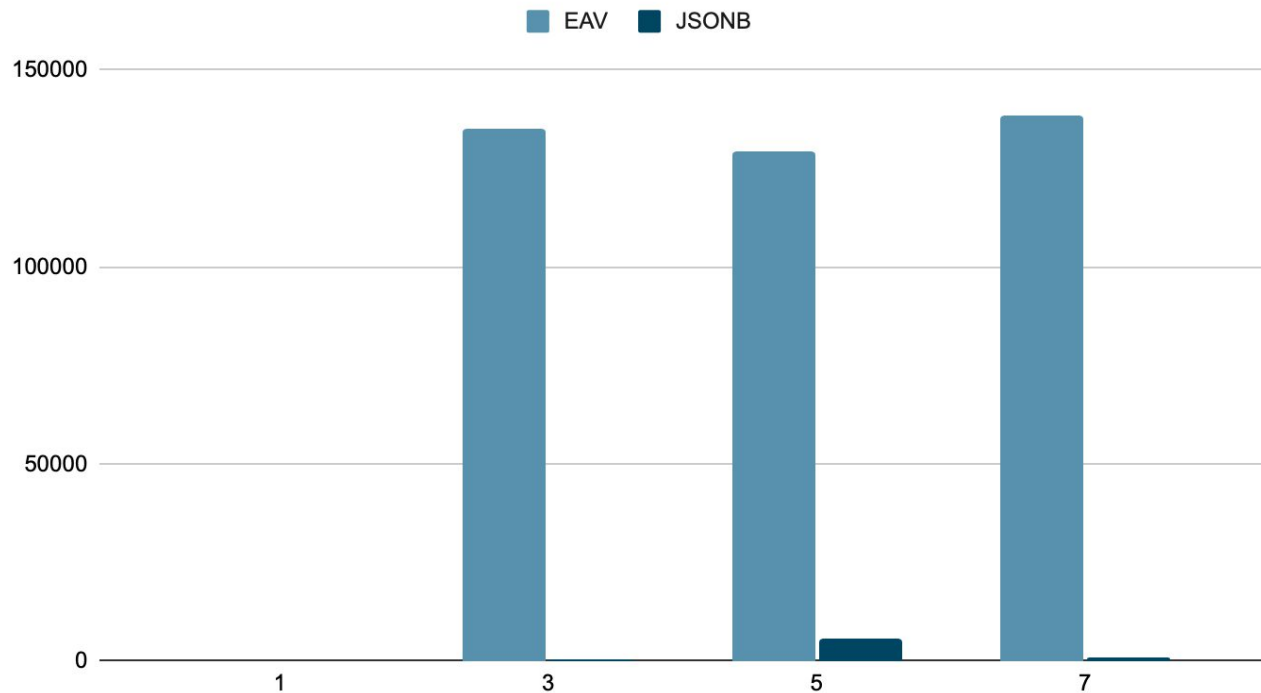
Атрибуты	EAV	JSONB
1	20500,98	0,89
3	152057,14	468,89
5	1072702,68	5564,36
7	1504423,66	1066,64

Поиск по полному соответствию с индексами 10m (cost)



Атрибуты	EAV	JSONB
1	16183,71	546,91
3	3958439,32	819452,09
5	4001084,03	969270,06
7	3951543,53	1119088,34

Поиск по полному соответствию с индексами 10m (ms)



Атрибуты	EAV	JSONB
1	11,50	0,89
3	135179,87	468,89
5	129426,04	5564,36
7	138188,45	1066,64

Выводы

- ❖ В поиске в контексте исследования JSONB показал себя лучше
- ❖ Возможна оптимизация запроса EAV, попробовать другие варианты
- ❖ Read-heavy: JSONB (проверить утверждение)
- ❖ Write-heavy: EAV (проверить утверждение)
- ❖ Провести исследование на document-oriented DB, например MongoDB