



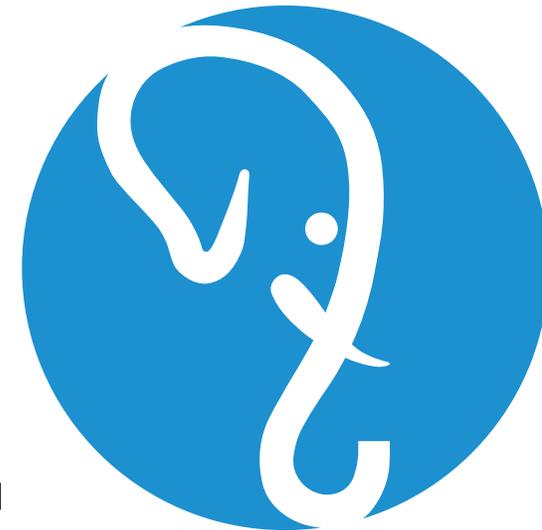
# Postgres Pro - отечественная коммерческая СУБД

Марк Ривкин  
Руководитель отдела технической  
поддержки продаж



# Что такое Postgres Pro Enterprise

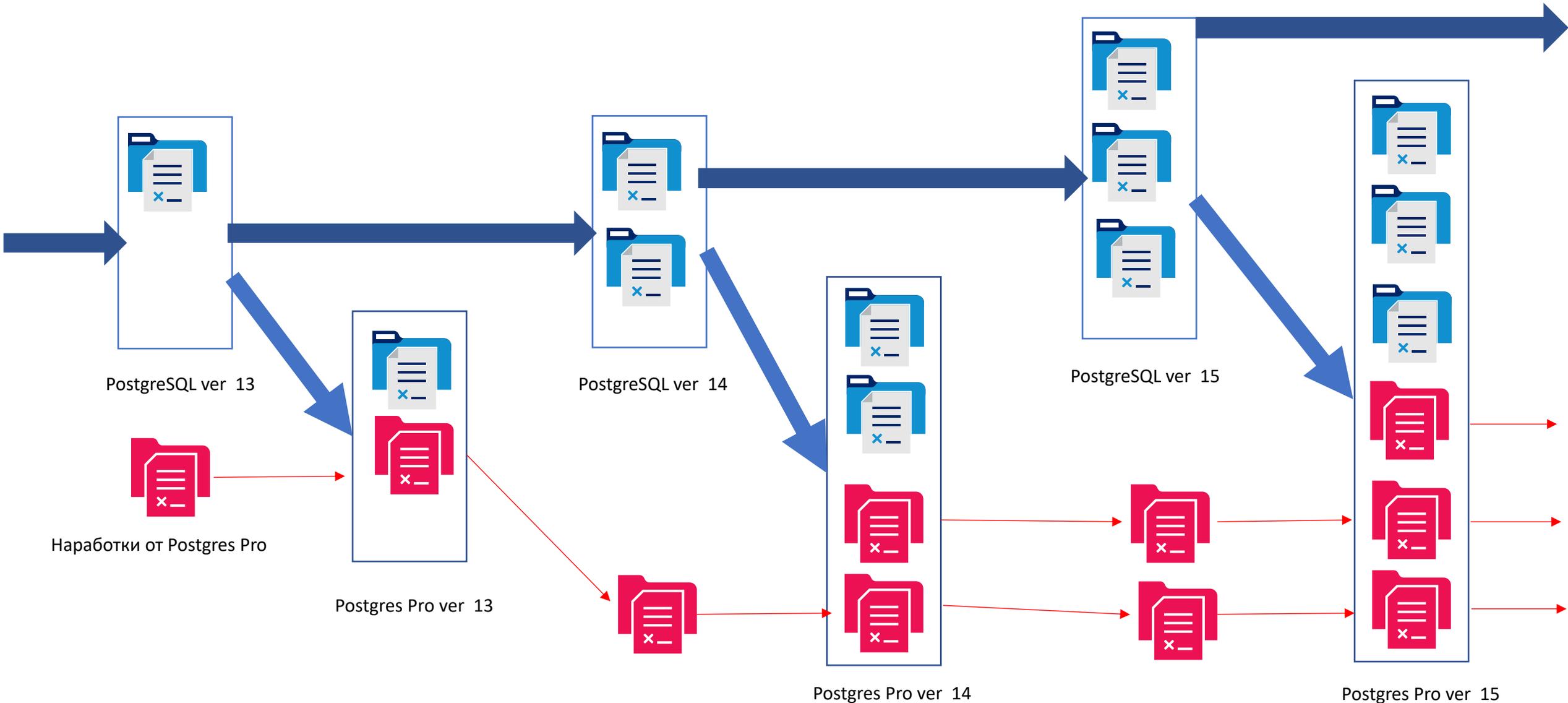
- Отечественная **коммерческая** объектно-реляционная СУБД для высоконагруженных систем крупных предприятий
- Богатая функциональность (таблицы, индексы, MV, view, последовательности, функции, процедуры, триггеры, ACID, роли/привилегии, ограничения целостности, BLOB и т д)
- Пакеты (packages), ORA2PGPRO, ORAFCE
- Тех поддержка 24x7, хорошая документация
- Разработчики и техническая поддержка в России
- Сертификат ФСТЭК уровень доверия УД 4 (защита конфиденциальной информации, у Oracle и MS - УД ниже и = 6)
- В реестре российских программных продуктов
- Расширяемый код
- Регулярное слияние с ванильной СУБД PostgreSQL



# Слияние версий PostgreSQL и Postgres Pro Enterprise



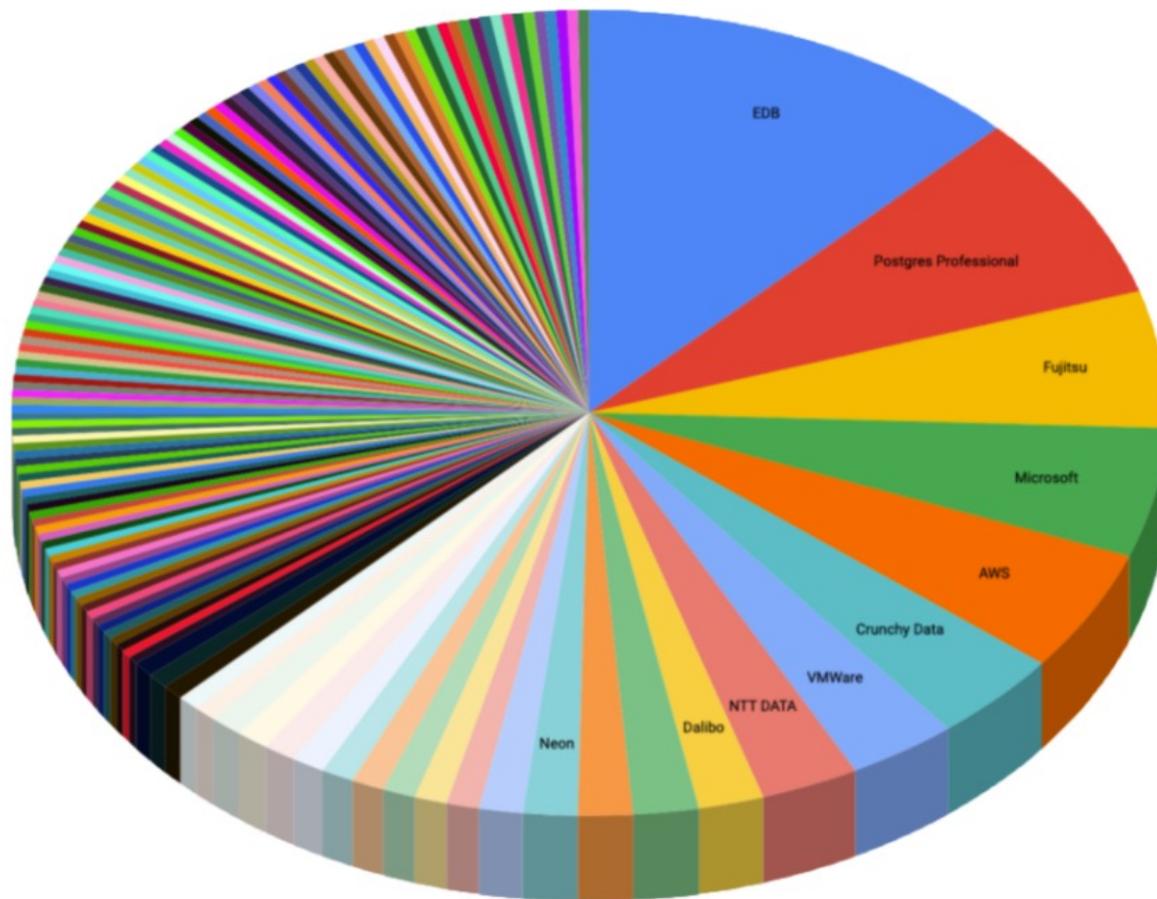
(версии 9, 10, 11, 12 ....)



# Postgres Pro Enterprise

- Многоплатформенность
- Расширяемость кода
- Различные типы данных
- Высокая надежность и доступность (24 x 7)
- Безопасность
- Масштабируемость (БД 130+ Тб, десятки тысяч пользователей, блокировки MVCC, 64 bit номер транзакции)
- Производительность и эффективность использования ресурсов
- Управляемость

# Отчет EDB о контрибьюторах в PostgreSQL 15.

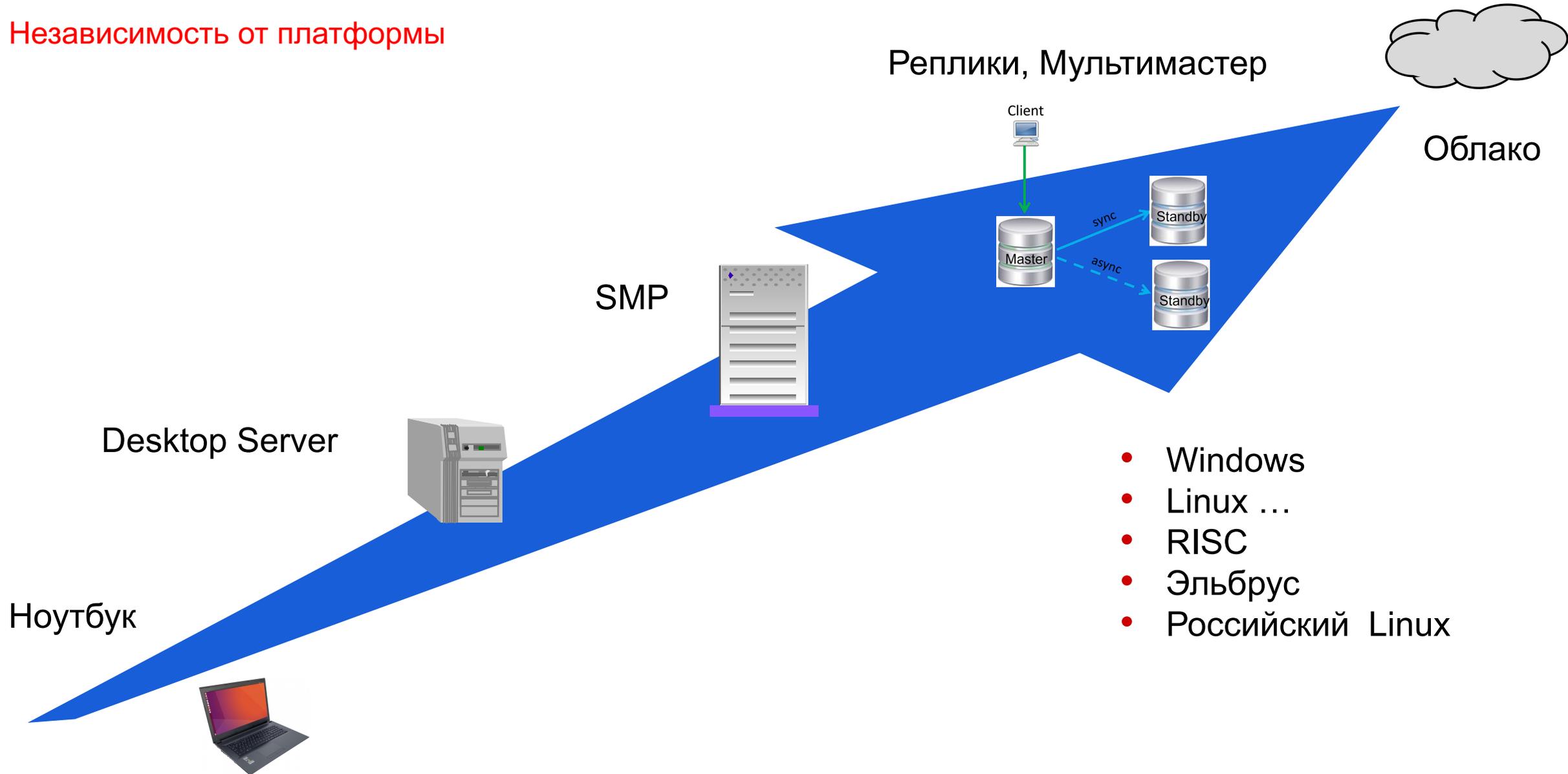


Code contributors to Postgres by company (PostgreSQL.org + PG 15 Release Notes) without personal or freelance

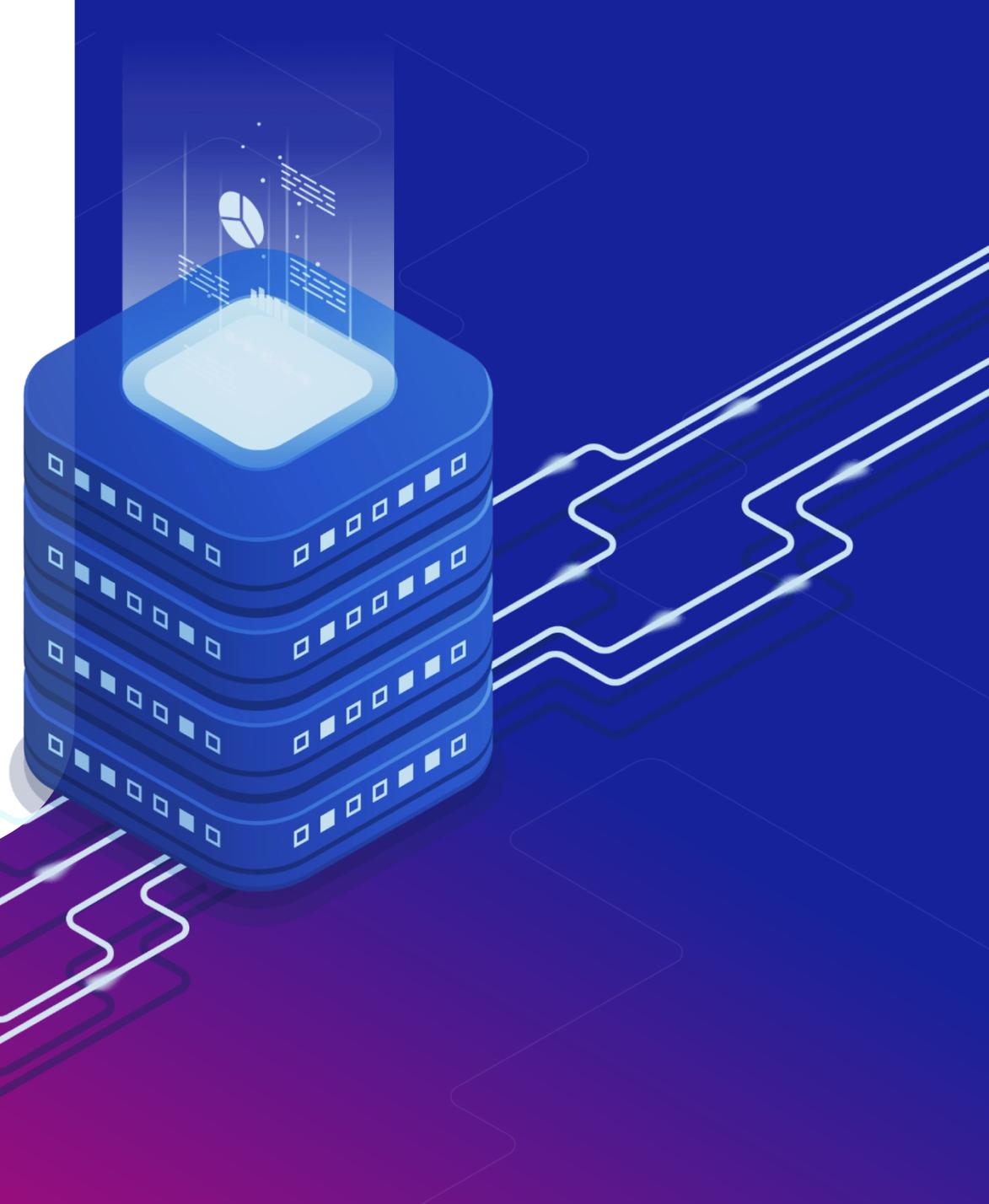
- EDB Postgres Professional Fujitsu Microsoft AWS Crunchy Data VMWare NTT DATA Dalibo Timescale HighGo Neon Cybertec Adjust
- Credativ Google Kontur Materialize NTT Red Hat SRA OSS Yugabyte Aiven Instaclustr Loxodata pganalyze PostgreSQL Experts
- AD Parts Analytics Engines Anastigmatix Apple Arcion Labs Arenadata Atos Avaya Axians NL Bank of China BCL Betsys Betterment
- Bigbank Blacksmith Applications Braintree Caesars Digital Capital Rx Capsico Health CdC Citus Data Clearco Cockroach Labs Code Synthesis Tools
- Codice Lieve Cofano Software Solutions Coinbase Conova Communications GmbH CrateDB CRSCube Data Egret dbi services Deutsche Telecom Dext
- DockYard Doctolib DuckDB Labs EdgeDB End Point Corp Entelect EPAM Systems Fivetran Forest Management Institute Garner Gentoo
- GLS Bank GTT HeteroDB HP IBM ILande Illuminated Computing Index Instructure Intel Intellasoft Intezer JackDB Jampp 58 more

# Разработав однажды, запускайте везде - 36 платформ

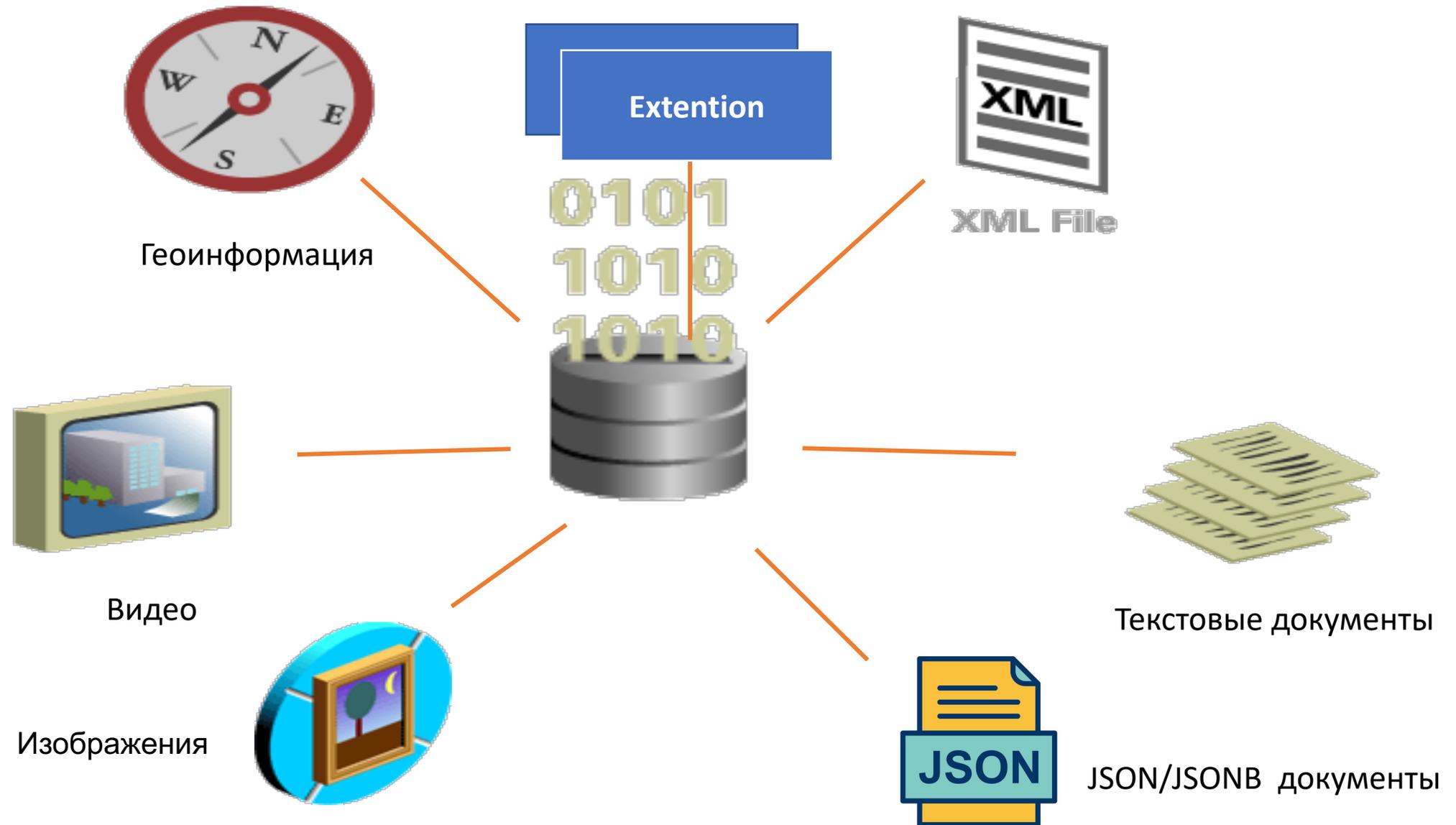
Независимость от платформы

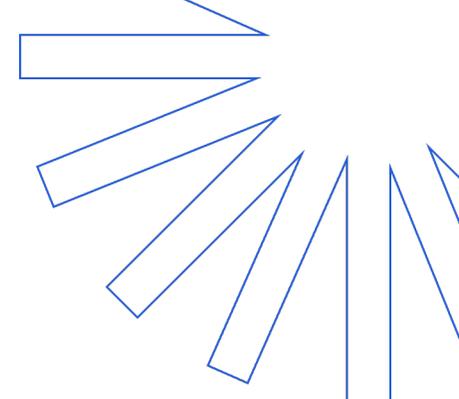


# Различные типы данных, расширяемость



# Работа с неструктурированными данными



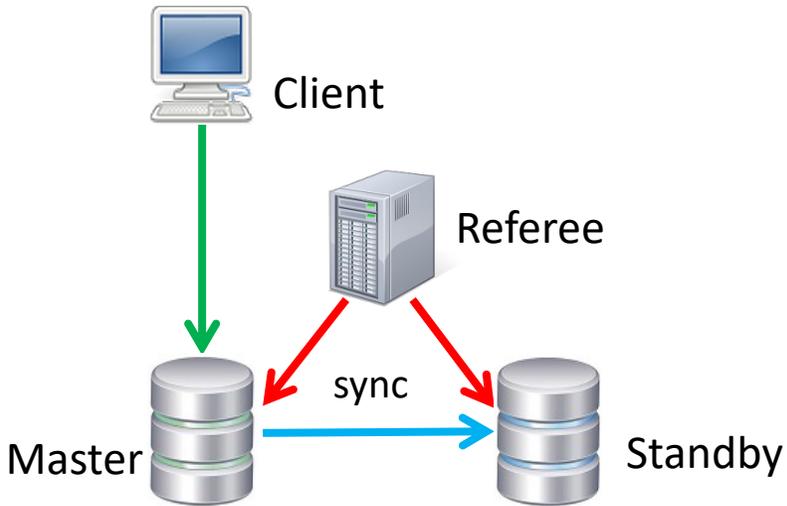


# Высокая надежность и доступность

24 x 7, 5 сек задержки при переключении узлов кластера

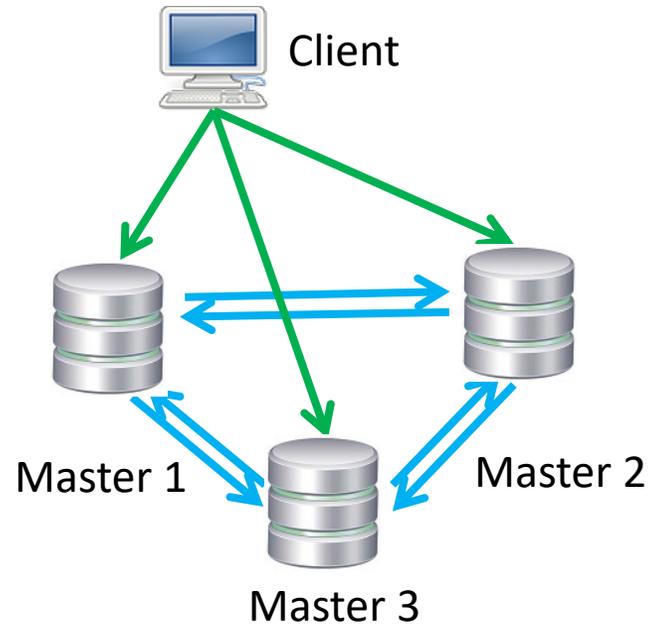


# Надежность и масштабирование



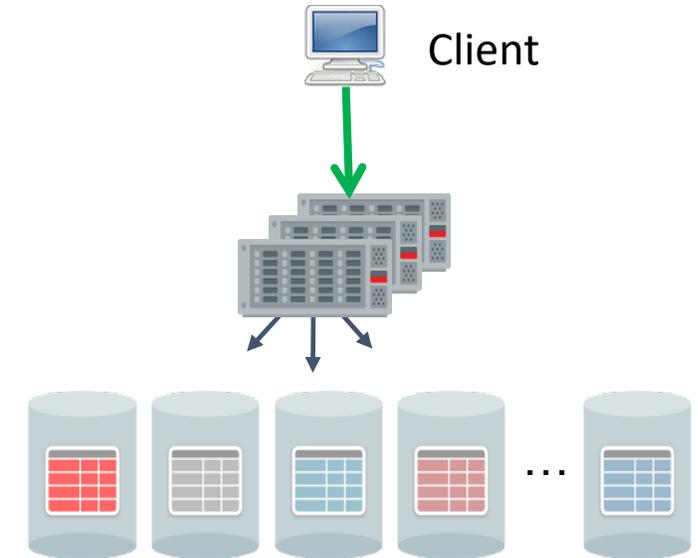
## Реплика (Standby)

Катастрофоустойчивость,  
надежность, масштабирование



## Мультимастер

Быстрое переключение,  
надежность, масштабирование



## Шардинг

Снимает ограничения на размер БД,  
масштабирование

# Шардинг

Шардирование по хэш функции снимает ограничения на размер БД

Таблицы БД в схеме

## Customers

Customer	Name
123	Mary
456	John
999	Peter

## Orders

Order	Customer
4001	123
4002	456
4003	999
4004	456
4005	456

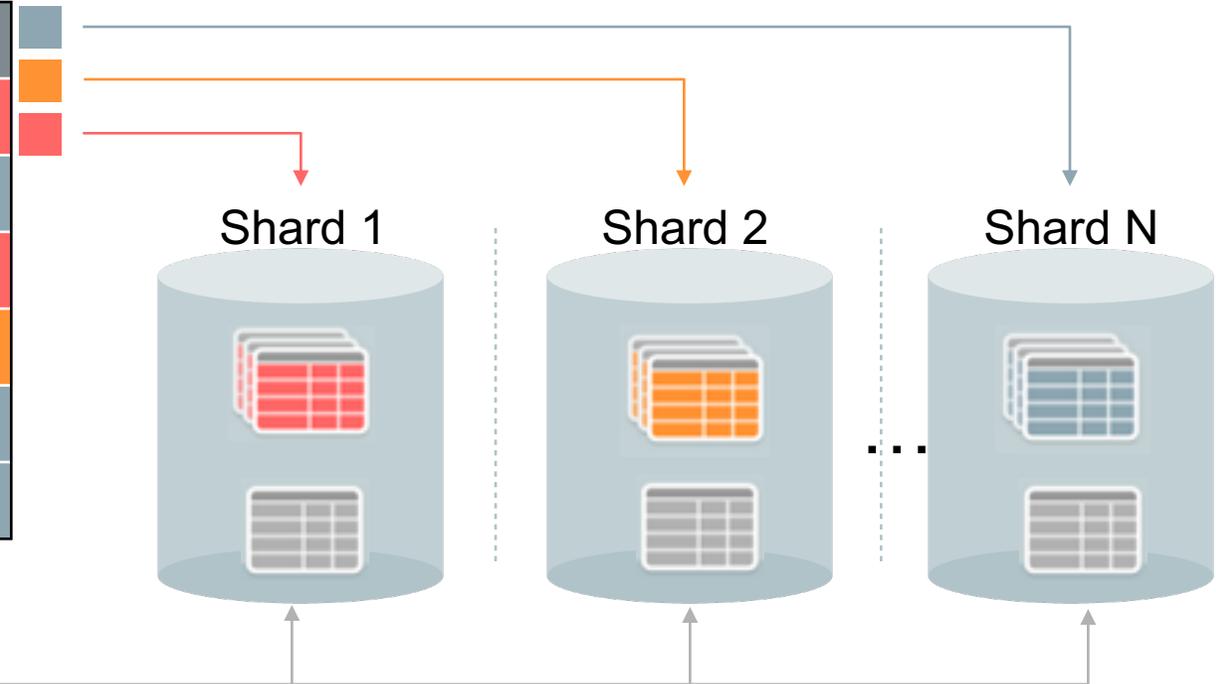
## LineItems

Customer	Order	Line
123	4001	40011
999	4003	40012
123	4001	40013
456	4004	40014
999	4003	40015
999	4003	40016

## Products

SKU	Product
100	Coil
101	Piston
102	Belt

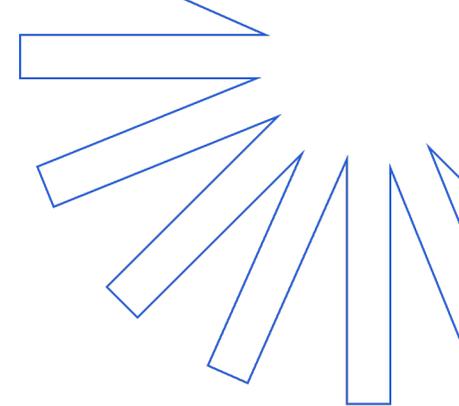
Sharded Tables



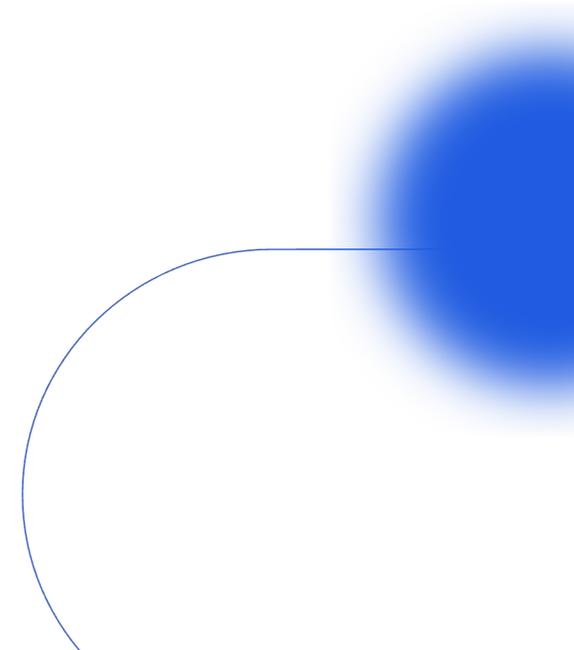
Таблицы справочников дублируются по всем БД

# Еще о надежности

- Онлайн операции:
  - Изменение параметров
  - Перестройка индексов
  - Изменение структуры таблиц
  - Работа с партициями
  - Патчи
- Backup/Восстановление БД, dump
  - Инкрементальный (3 варианта) и полный Backup, слияние полного и инкрементальных backup
  - Параллельность резервного копирования и восстановления
  - Сжатие backup
  - Восстановление БД на момент времени (PITR)
  - Пробное восстановление
  - Политики хранения
  - [Backup на объектное хранилище S3]
- Dump – создание скриптов экспорта
- Автономные транзакции



**Безопасность**



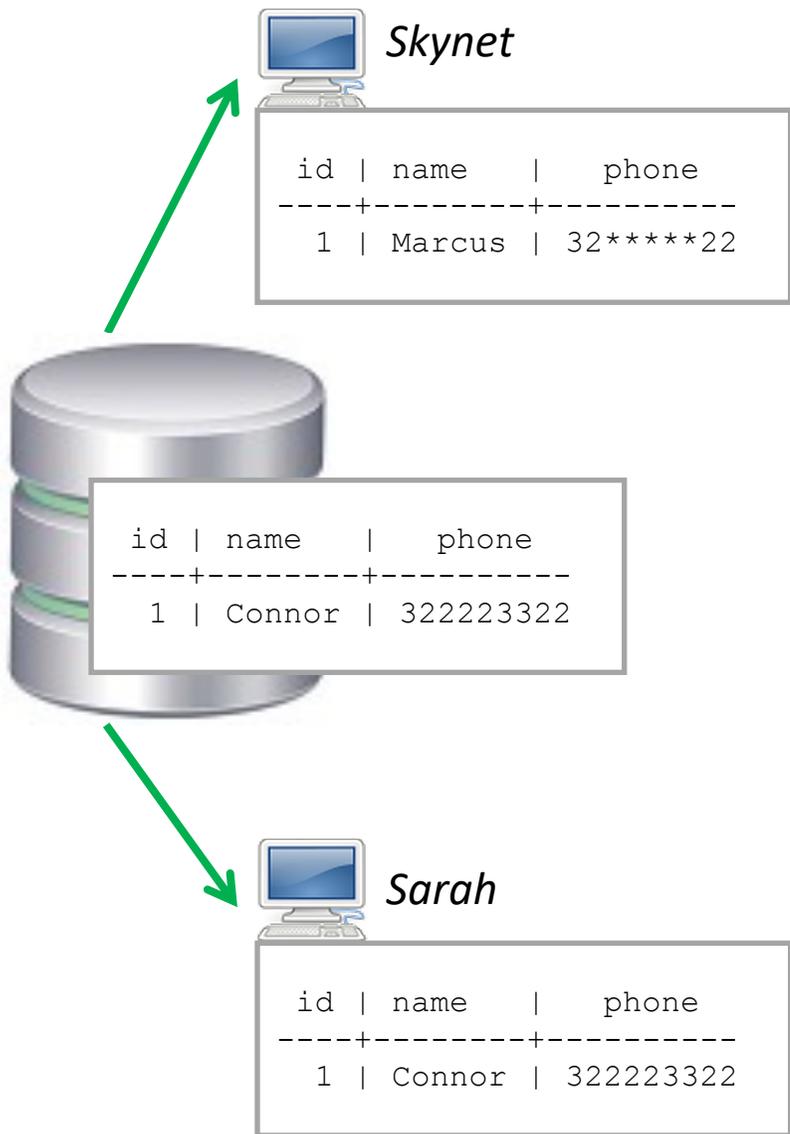


# Авторизация и аутентификация

- Имя/пароль
- Внешние системы
  - Kerberos
  - LDAP
  - Active Directory
  - Radius
  - PAM
  - peer/ident
- Авторизация ОС
- Авторизация БД
- Шифрование пароля
  - scram-sha-256
  - md5

# Аудит действий

- Поддерживается ответственным администратором...
  - Гибкость и дискретность аудита
    - Аудит пользователей, операций, обращений к объектам, привилегий ...
    - Множество событий, подлежащих аудиту
    - Уровни логирования в журнал: DEBUG5, DEBUG4, DEBUG3, DEBUG2, DEBUG1, INFO, NOTICE, WARNING, ERROR, LOG, FATAL и PANIC
  - Журналы: stderr, csvlog, syslog, eventlog(wind), журнал аудита
  - pg\_proaudit, Аладдин



**pgpro\_anonymizer** — расширение для маскирования или замены конфиденциальных данных непосредственно внутри экземпляра Postgres Pro:

- Динамическое маскирование: сокрытие данных только от недоверенных пользователей.
- Статическое маскирование: преобразование данных в БД в соответствии с правилами.
- Экспортирование замаскированных данных в файл SQL.

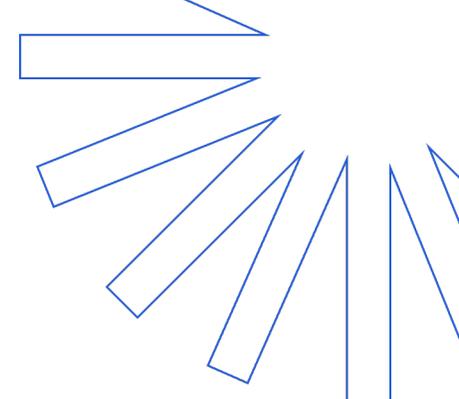
Правила маскирования задаются метками безопасности:

```
SELECT anon.start_dynamic_masking();
```

```
SECURITY LABEL FOR anon ON COLUMN people.name  
IS 'MASKED WITH FUNCTION anon.fake_first_name()';
```

```
SECURITY LABEL FOR anon ON COLUMN people.phone  
IS 'MASKED WITH FUNCTION anon.partial(phone,2,$$*****$$,2)';
```

```
SECURITY LABEL FOR anon ON ROLE Skynet, Mark IS 'MASKED';
```



# Масштабируемость

Использование всех ресурсов компьютера, поддержка больших БД, больших нагрузок



# Работа с большими БД и под большой нагрузкой

- **Разнесение нагрузки** – реплика, мультимастер, шардинг, встроенный load balancer
- **Секционирование**
- **Параллельная обработка**
- **Сжатие**
- **Управление ресурсами (resource manager)** – приоритизация для сеанса на использование `cpu`, `io_read`, `io_write` (1,2,4,8) 4-default, 8 - макс
- Сегодня у заказчиков есть БД 130+ Tb (300 tb яндекс почта, 140 tb ВРС), поддержка 10000+ пользователей

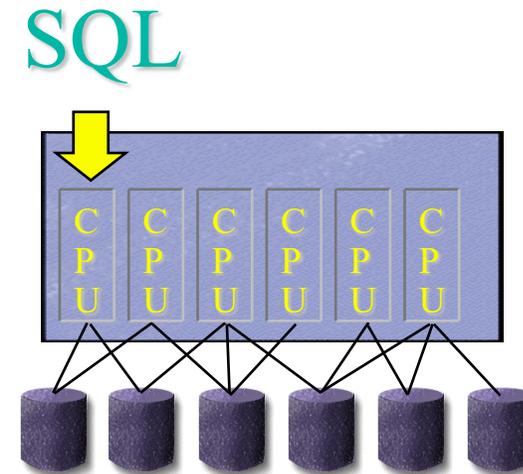
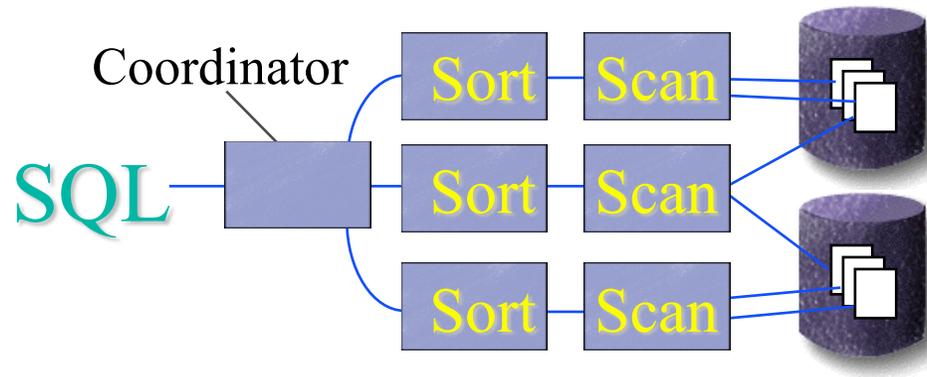
# Секционирование таблиц и индексов

*Улучшает доступность, управляемость*

- Секционирование упрощает управление VLDB
- Секционирование значительно повышает производительность (partition wise)
- Разделы прозрачны для пользователей и приложений
- Данные разбиваются в зависимости от приложения
- Каждый раздел может иметь свой tablespace



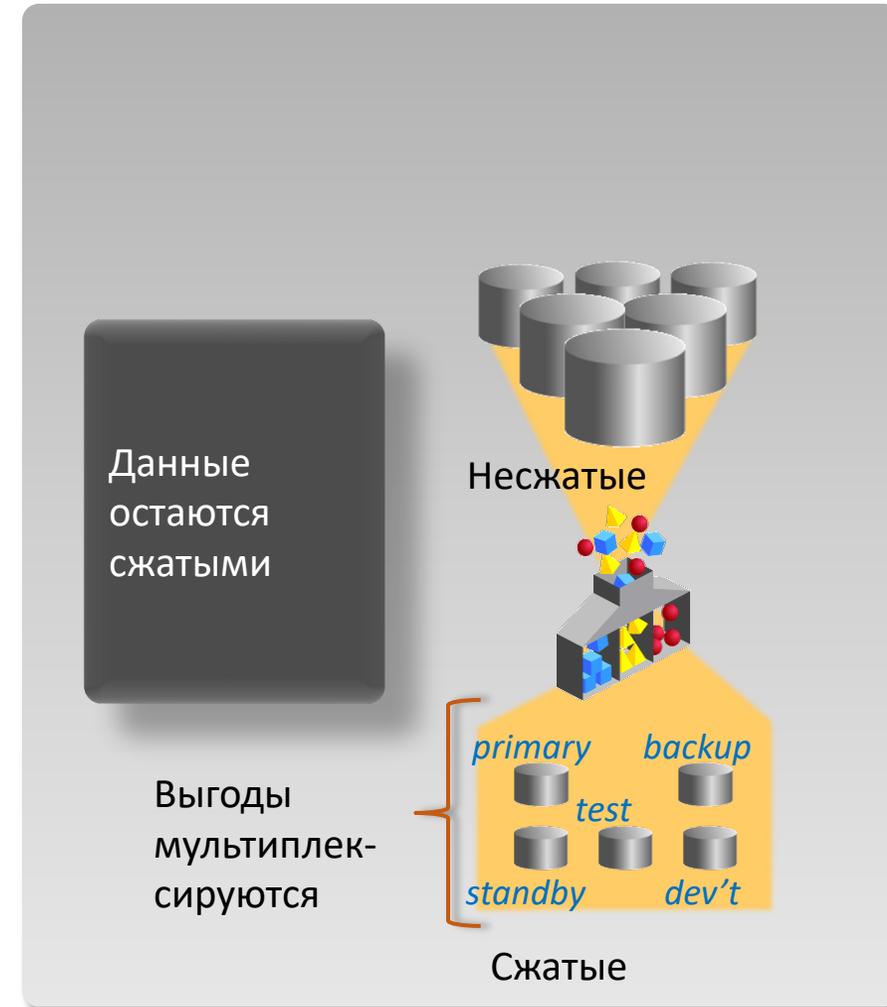
# Parallel Query



- max\_parallel\_workers\_per\_gather > 0
- Параллельное сканирование
- Параллельное сканирование индекса или по индексу
- Параллельное соединение, агрегирование
- Optimizer Hints (Администратор имеет контроль)
- Parallel-Aware Query Optimizer

# Сжатие

- Можно сжимать:
  - Структурированные данные (таблицы, индексы)
  - TOAST таблицы
  - Неструктурированные данные (текст, XML и т д)
  - Backup – утилита `pg_probackup`
  - Dump - утилита `pg_dump`
- CFS – Compression File System - встроена в Postgres Pro Enterprise, задается для Tablespace
- Json → JsonB
- Выбор алгоритма и степени сжатия (`lz4`, `zstd`, `zlib`, `pglz...`)
- Выгоды сжатия
  - Уменьшает использование диска – обычно в 2 – 3 и до 5 раз
  - Уменьшает ввод/вывод



- Множество индексов, MV
- Инкрементальный backup
- MVCC – читатели не блокируют писателей, писатели - читателей
  
- Ограничения Postgres Pro Enterprise
  - Размер БД - не ограничен
  - Максимальный размер таблицы - 32 Tb
  - Количество записей в таблице – не ограничено
  - Количество индексов – не ограничено
  - Максимальный размер первичного ключа – 32 колонки
  - Максимальная длина записи - 400 Гб
  - Максимальная длина атрибута TOAST - 1 Гб
  - Максимальная длина BLOB – 4 Tb

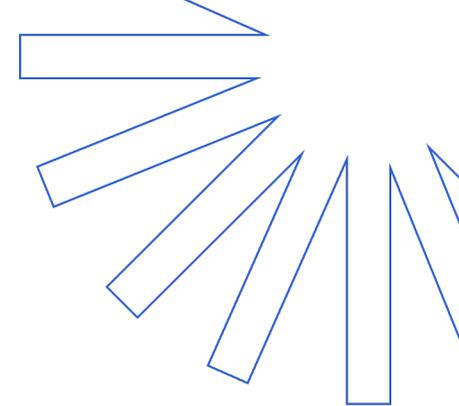
# Производительность

- Хороший оптимизатор (планировщик) запросов
- AQO – адаптивная оптимизации запросов (учет ошибок статистики при следующем выполнении запроса)
- Хинты + PG\_hint\_plan (SQL patch), заморозка планов
- Множество различных типов индексов
- Секции и материализованные представления, wise join
- Перенос табличных пространств на быстрые диски





# Производительность



# Производительность

- Хороший оптимизатор (планировщик) запросов
- AQO – адаптивная оптимизации запросов (учет ошибок статистики при следующем выполнении запроса)
- Хинты + PG\_hint\_plan (SQL patch), заморозка планов
- Множество различных типов индексов
- Секции и материализованные представления, wise join
- Перенос табличных пространств на быстрые диски

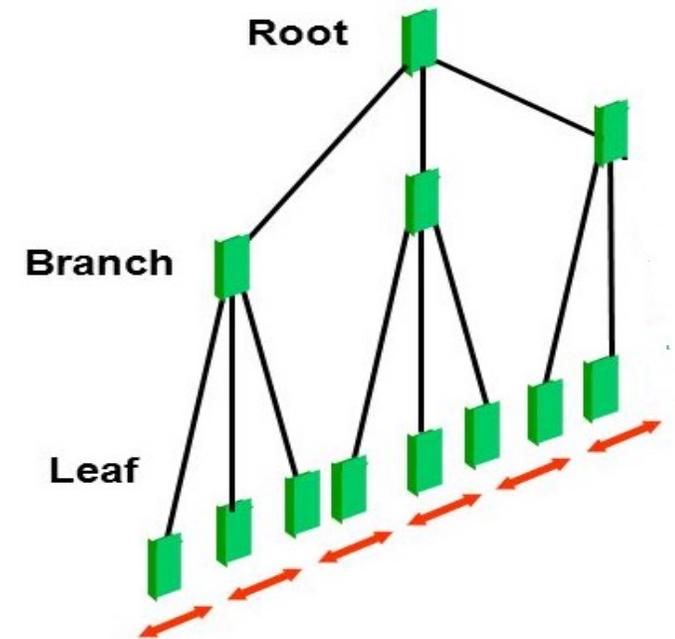


# Производительность

- JSONB
- Prepared SQL запросы
- VACUUM ANALYZE – сбор статистики для оптимизатора
- 64 – битный счетчик транзакций
- Вынос нагрузки на реплику, мультимастер, шарды

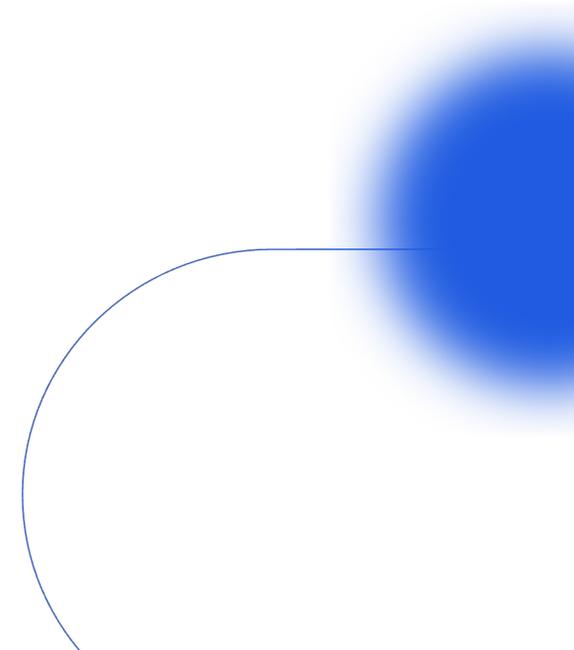
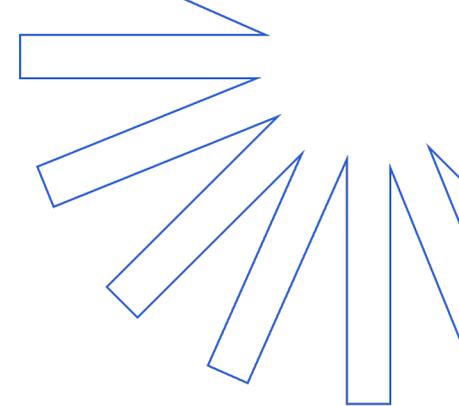
# Индексы на все случаи жизни

- B-деревья ( $>$ ,  $>=$ ,  $<$ ,  $<=$ ,  $=$ )
- Хеш ( $=$ )
- GIST - (generalized search tree) – для **гео**, **текстов**, **рисунков** – разбивка на квадраты по ф-ии соответствия (близости, вхождению, пересечению, левее-правее, KNN и т д)
- SP-GIST – space partitioning GIST (разбиение на непересекающиеся квадраты)
- GIN – (Generalized Inverted Index - обратный индекс) , например, **индекс в конце** книги – где встречается элемент). Хорош для **полнотекстового** поиска
- RUM – **развитие GIN**, поиск релевантных документов, **полнотекстовый** поиск, **знает ПОЗИЦИЮ** слова на странице, слова отстоящие на  $x$  слов, сортировка по схожести
- BRIN - (Block Range Index) - аналог **storage index** (исключение “блоков”)
- Триграммы для GIST и GIN
- Покрывающие и частичные индексы, индексы по выражению





# Управляемость



# Что должна делать система управления

- Управление объектами БД и кластера ( просматривать иерархию и характеристики и администрировать)
  - базы, TS, пользователи, схемы, роли, таблицы, MV .....
  - PgAdmin, Dbeaver, Valentina Studio, Navicat
- Мониторинг поведения БД (метрики, графики по времени, дашборды, богатые наборы шаблонов, тренды .....
- Zabbix+Mamonsu, Prometheus
- Конфигурирование (реплик, узлов, FDW, баз .....
- Диагностика (поиск, выявление, предсказание проблем)
- Настройка БД и запросов

PostgresPro  
ENTERPRISE MANAGER

УПРАВЛЕНИЕ

- Дашборд
- Инстансы
- Базы данных
- Лог событий
- Расписание
- Кластеры
- Бэкап
- Профайлер

СИСТЕМА

- Управление пользователями
- Настройки
- Аудит

<> PSQL >

## Дашборд

ИНСТАНСЫ <b>2</b>	СЕРВЕРЫ <b>2</b>	КЛАСТЕРЫ <b>0</b>	ПОСЛЕДНИЕ УВЕДОМЛЕНИЯ
БАЗЫ ДАННЫХ <b>4</b>	ОБЩИЙ РАЗМЕР БД <b>1002 MB</b>	Похоже здесь пусто	

## ЛОГ СОБЫТИЙ

Дата, время	Инстанс	База данных	Тип	Log
12.04.2023, 11:47:04.43	PgPro Manager repository	pgpro_manager_repo	LOG	duration: 555.846 ms rows: 1 size: 7 bytes statement: WITH cfs_write_speed_lates FROM monitoring_history mh JOIN metrics m ON mh.metric_id = m.id WHERE m.metric_ mh.value_set DESC LIMIT 1), cfs_scan_speed_latest AS ( SELECT mh.metric_value FF
12.04.2023, 11:47:03.16	PgPro Manager repository	pgpro_manager_repo	LOG	duration: 554.918 ms rows: 1 size: 7 bytes statement: WITH cfs_write_speed_lates FROM monitoring_history mh JOIN metrics m ON mh.metric_id = m.id WHERE m.metric_ mh.value_set DESC LIMIT 1), cfs_scan_speed_latest AS ( SELECT mh.metric_value FF
12.04.2023, 11:46:33.38	PgPro Manager repository	pgpro_manager_repo	LOG	duration: 1076.063 ms rows: 105 size: 15074 bytes statement: SELECT m.metric_key mh.metric_value, m.datatype, m.units, m.delta, mh.value_set::timestamp(0), EXTRA metrics m JOIN monitoring_history mh ON (m.id = mh.metric_id) WHERE lower(m.metri mh.value_set >= '2023-04-12 08:31:28' AND mh.value_set <= '2023-04-12...
12.04.2023, 11:46:33.34	PgPro Manager repository	pgpro_manager_repo	LOG	duration: 1463.161 ms rows: 105 size: 14580 bytes statement: SELECT m.metric_key mh.metric_value, m.datatype, m.units, m.delta, mh.value_set::timestamp(0), EXTRA metrics m JOIN monitoring_history mh ON (m.id = mh.metric_id) WHERE lower(m.metri mh.value_set >= '2023-04-12 08:31:28' AND mh.value_set <= '2023-04-12...

## Инстансы

5 инстансов | 3 сервера | 0 кластеров

ДОБАВИТЬ ИНСТАНС

Наименование	Сервер	Чексуммы	Сбор логов	Роль	БД	
<b>14/main</b> Порт: 5432 Директория: /usr/lib/postgresql/14/bin/ Версия Postgres: 14.7	619DDD95962F agent-postgres	off	<input type="checkbox"/>	primary	Базы данных: 2 Транзакций в секунду: 1.03 Соединения: 1 Средняя нагрузка: 0.10 / 0.15 / 0.16	
<b>PgPro Manager repository</b> Порт: 5432 Директория: /opt/pgpro/ent-14/bin/ Версия Postgres: 14.6 (enterprise)	BAFA824CE87C 172.26.0.2	on	<input checked="" type="checkbox"/>	primary	Базы данных: 2 Транзакций в секунду: 12.63 Соединения: 1 Средняя нагрузка: 0.10 / 0.15 / 0.16	
<b>PgPro-EE</b> Порт: 5432 Директория: /opt/pgpro/ent-14/bin/ Версия Postgres: unknown	N/A agent-pgproeeqqq	-	<input checked="" type="checkbox"/>	n/a	Базы данных: - Транзакций в секунду: - Соединения: - Средняя нагрузка: - / - / -	
<b>rew</b> Порт: 5432 Директория: / Версия Postgres: -	N/A agent-postgres	-	<input type="checkbox"/>	n/a	Базы данных: - Транзакций в секунду: - Соединения: - Средняя нагрузка: - / - / -	
<b>test</b> Порт: 5432 Директория: / Версия Postgres: unknown	N/A agent-pgproeeqqq	-	<input type="checkbox"/>	n/a	Базы данных: - Транзакций в секунду: - Соединения: - Средняя нагрузка: - / - / -	

# Мониторинг

- Командная строка или графические инструменты
- Средства ОС (ps, iostat, vmstat, sar, top, df, du, quota ...)
- Statistics Collector – сбор статистики в БД (таблицы pg\_stat ...), таблицы/view для отслеживания выполнения ANALYZE, CREATE INDEX, VACUUM, CLUSTER, COPY, backup, блокировок.
- PGPRO\_STATS – мониторинг выполнения SQL
- Журнал сообщений сервера - ОС или утилита PGBADGER с GITHUB (only english)
- EXECUTION PLAN
  - EXPLAIN
  - EXPLAIN ANALYZE – реальное выполнение
- Расширение PGPRO\_PWR

- [Top tables by blocks vacuum fetched](#)
- [Top indexes by blocks vacuum fetched](#)
- [Top tables by blocks vacuum read](#)
- [Top indexes by blocks vacuum read](#)
- [Top tables by WAL size generated by vacuum](#)
- [Top tables by vacuum operations](#)
- [Top tables by analyze operations](#)
- [Cluster settings during the report interval](#)

## Server statistics

### Database statistics

Database	Transactions			Block statistics			Block I/O times			Tuples				Temp files		Size	Growth
	Commits	Rollbacks	Deadlocks	Hit(%)	Read	Hit	Read	Write	Ret	Fet	Ins	Upd	Del	Size	Files		
db5	3377			100.00		185864			2129743	39349						9317 kB	
demo	17183			100.00		193323			2298257	43985						707 MB	
mamonsu_database	61593			100.00		359089			6689577	48937		1416		6998 MB	1337	9517 kB	
postgres	3764	5		99.93	1096	1554516	0.10		12364299	647192	21928	12584	14362	2306 kB	9	86 MB	4744 kB
Total	85917	5		99.95	1096	2292792	0.10		23481876	779463	21928	14000	14362	7000 MB	1346	811 MB	4744 kB

### Session statistics by database

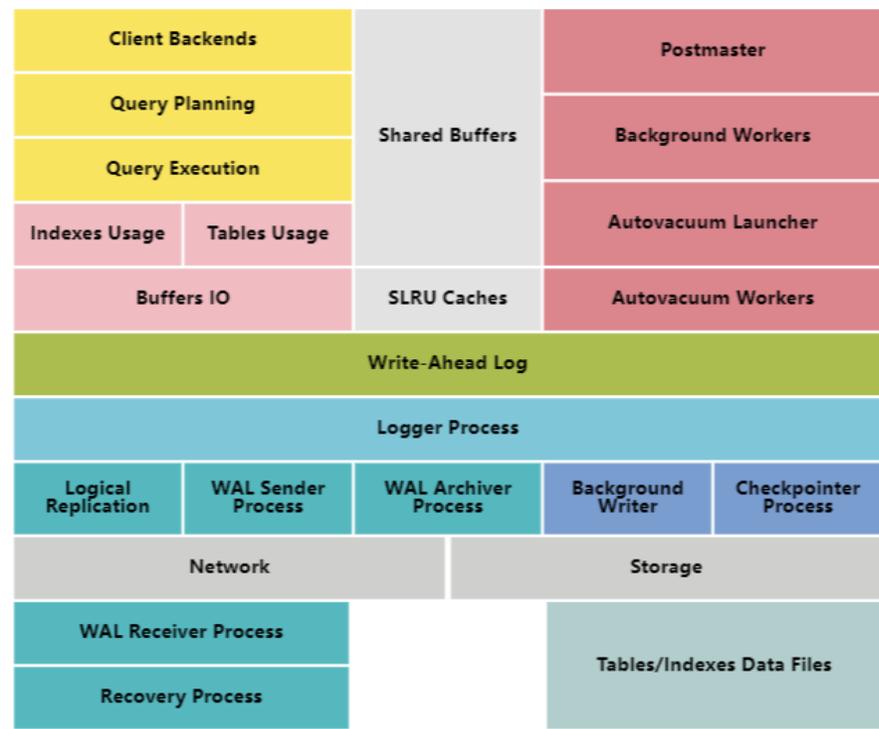
Database	Timings (s)			Sessions			
	Total	Active	Idle(T)	Established	Abandoned	Fatal	Killed
db5	80463.05	4.36	0.36	8			
demo	123699.00	225.80	0.33	11	3		
mamonsu_database	80411.25	192.43	0.34	8			
postgres	126511.75	10.96	4.11	25	1		
Total	411085.04	433.54	5.13	52	4		

# Postgres Observability - pgstats.dev

15

## Postgres Observability

- 14 `pg_stat_ssl`
- 13 `pg_stat_activity`
- 12 `pg_backend_memory_contexts()`
- 11 `pg_stat_monitor`
- 11 `pg_stat_kcache`
- 11 `pg_wait_sampling`
- 11 `pg_blocking_pids()`
- 10 `pg_stat_progress_cluster`
- 10 `pg_stat_progress_copy`
- 9.6 `pg_stat_all_indexes`
- 9.6 `pg_stat_all_tables`
- 9.6 `pg_statio_all_indexes`
- 9.6 `pg_statio_all_tables`
- 9.5 `pg_statio_all_sequences`
- `pg_get_wal_replay_pause_state()`
- `pg_is_wal_replay_paused()`
- `pg_current_wal_lsn()`
- `pg_wal_lsn_diff()`
- `pg_ls_logdir()`
- `pg_current_logfile()`
- `pg_replication_slots`
- `pg_stat_replication_slots`
- `pg_stat_replication`
- `pg_stat_subscription`
- `pg_stat_wal_receiver`
- `pg_stat_archiver`
- `pg_ls_archive_statusdir()`
- `pg_stat_database_conflicts`
- `pg_stat_recovery_prefetch`



- `pg_buffercache`
- `pg_shmem_allocations`
- `pg_stat_slru`
- `pg_stat_activity`
- `pg_stat_database`
- `pg_stat_progress_vacuum`
- `pg_stat_progress_analyze`
- `pg_stat_all_tables`
- `pg_stat_wal`
- `pg_ls_waldir()`
- `pg_walfile_name()`
- `pg_current_wal_insert_lsn()`
- `pg_walfile_name_offset()`
- `pg_current_wal_flush_lsn()`
- `pg_last_wal_receive_lsn()`
- `pg_last_wal_replay_lsn()`
- `pg_last_xact_replay_timestamp()`
- `pg_stat_bgwriter`
- `pg_stat_progress_basebackup`
- `pgstattuple`
- `pg_tablespace_size()`
- `pg_database_size()`
- `pg_total_relation_size()`
- `pg_relation_size()`
- `pg_table_size()`
- `pg_indexes_size()`
- `pg_ls_tmpdir()`
- `pg_ls_dir()`
- `pg_relation_filenode()`
- `pg_relation_filepath()`
- `pg_filenode_relation()`

nicstat

iostat

# Сравнение версий

Характеристика/ возможность	PostgreSQL	Postgres Pro Enterprise
Сертификат ФСТЭК		Postgres Pro Enterprise Certified
Обновление версии без остановки сервера		•
Отказоустойчивая конфигурация: мастер-мастер (мультимастер)		•
Встроенный пул соединений		•
64-битовый счетчик транзакций		•
Сжатие данных		•
Инкрементальное блочное резервное копирование		•
Эффективное секционирование (10-ки тысяч секций)		•
Автономные транзакции		•
Встроенный планировщик заданий		•
Оптимизация запросов методами машинного обучения - AQO		•

# Сравнение версий

Характеристика/ возможность	PostgreSQL	Postgres Pro Enterprise
Шардман		•
JsonB		•
Mamonsu – агент мониторинга		•
PG_Probackup		•
PGPRO_PWR		•
Хинты		•
Покрывающие индексы		•
KNN: Индексный поиск ближайших соседей		•
Комбинирование полнотекстового индекса с индексом релевантности (RUM)		•
Мониторинг производительности и контроль ожидания при выполнении запросов		•

## Характеристика/ возможность

## PostgreSQL

## Postgres Pro Enterprise

Триггеры событий входа

•

Приоритизация ресурсов

•

Online перемещение таблиц

•

Автоматическое исправление  
страниц из реплики

•

Исправление WAL из буферов  
памяти

•

Работа мастера при выходе из  
строя одной из реплик

•

Изменение конфигурации  
других сеансов

•

Отложенный сброс временных  
таблиц на диск

•

Pg\_wait\_sampling – история и  
профиль ожиданий

•

Режим автоподготовки  
операторов

•

## Характеристика/ возможность

## PostgreSQL

## Postgres Pro Enterprise

Пакеты (как в Oracle)



Orafce дополнительные системные пакеты



Ora2pg EE



Статическое маскирование



Динамическое маскирование



SR\_PLAN – заморозка плана



SQL/JSON



Поддержка платформ ARM и Эльбрус



Backup в S3



