



PostgreSQL ecosystem

Speaker: 林宗禧 @ COSCUP x RubyConf TW 2021

Taiwan PostgreSQL User Group



林宗禧

- PostgreSQL 愛好者 (2012-)
- PostgreSQL 推廣者 (2017-)

How to use ?

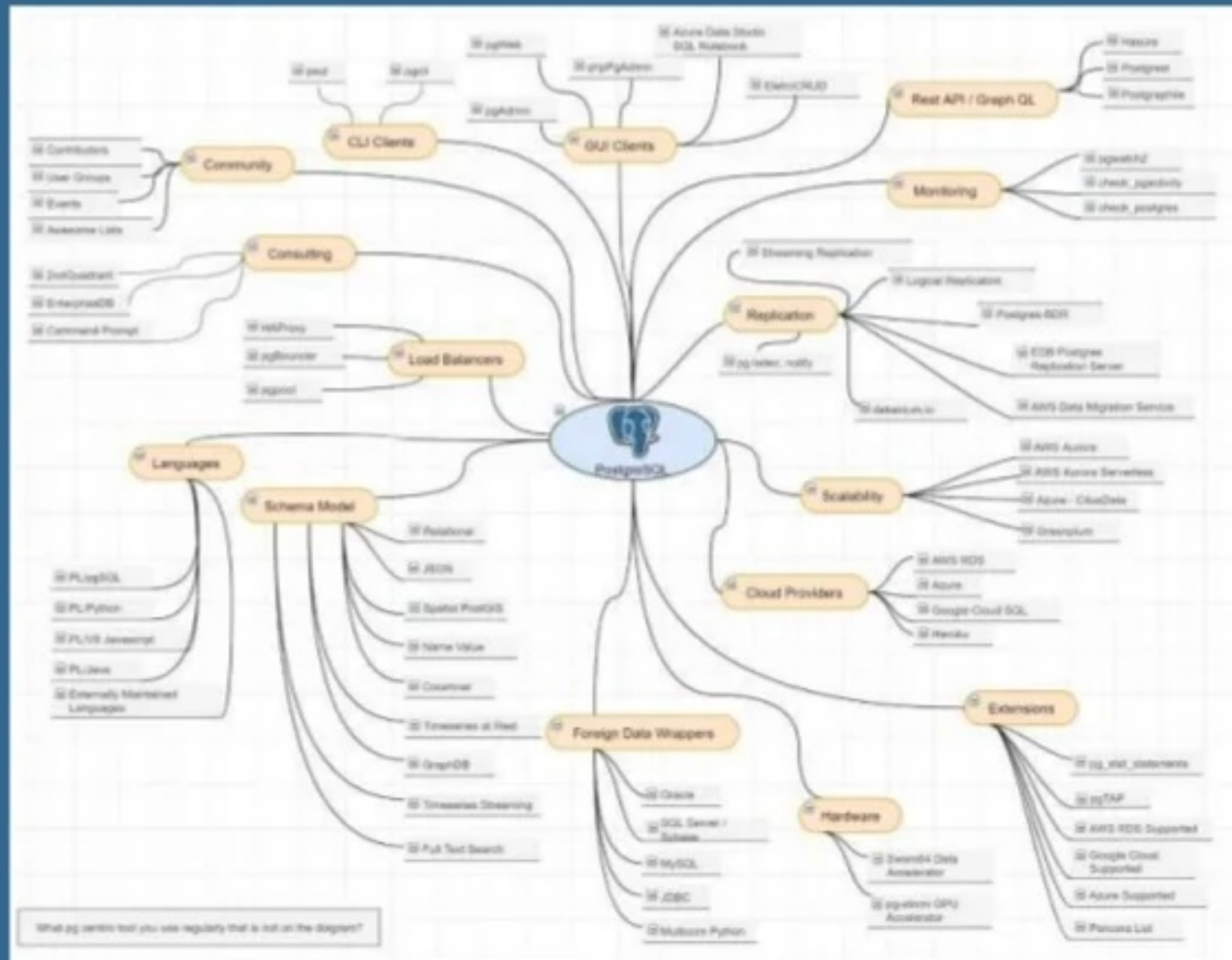
- 以前: 開發 FDW 套件 (C, Python 都有)
- 後來: 到處整合 PG 的應用
 - 推 Industry 4.0, 讓業主不經意的導入 PG
 - 推 Smart City Solutions, 拿 PG 做基礎





Overview – PostgreSQL ecosystem (1/3)

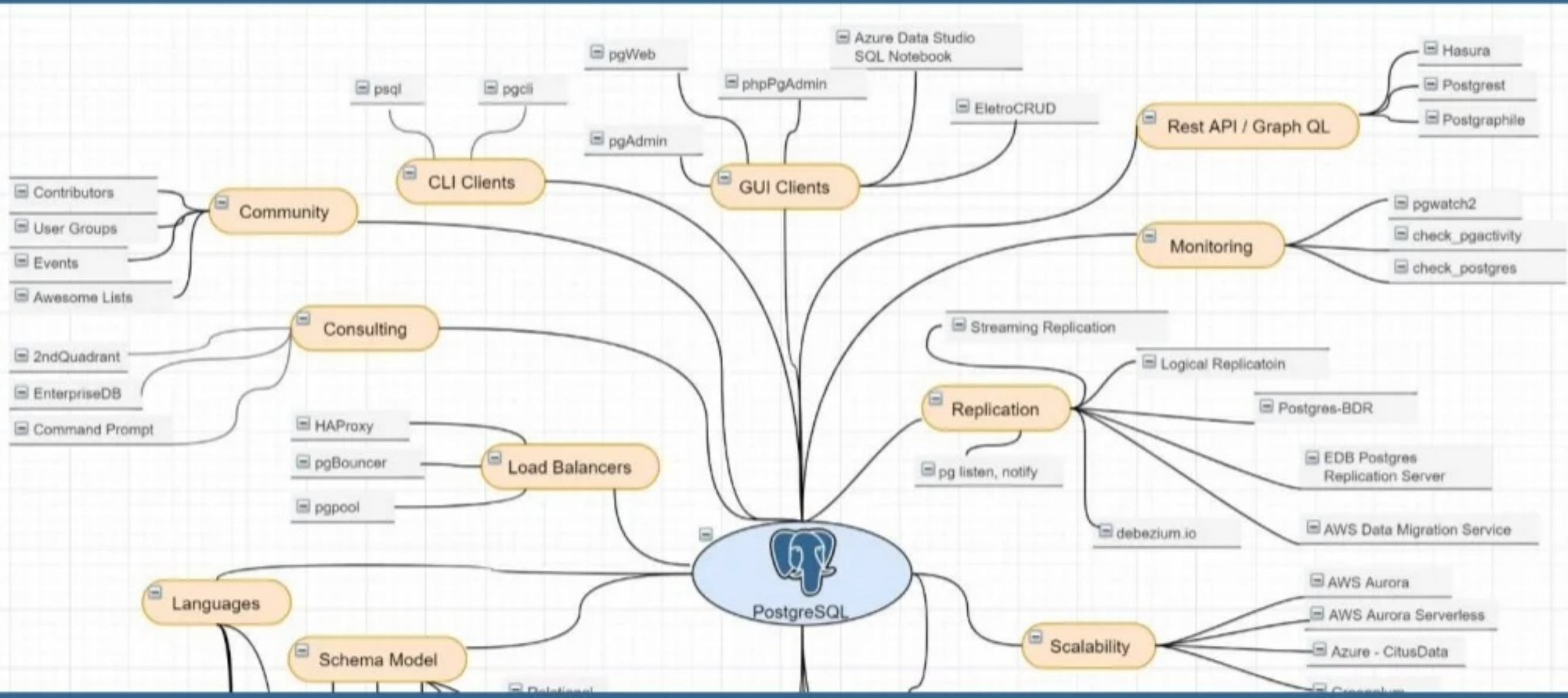
- <https://code.efficiencygeek.com/>





Overview – PostgreSQL ecosystem (2/3)

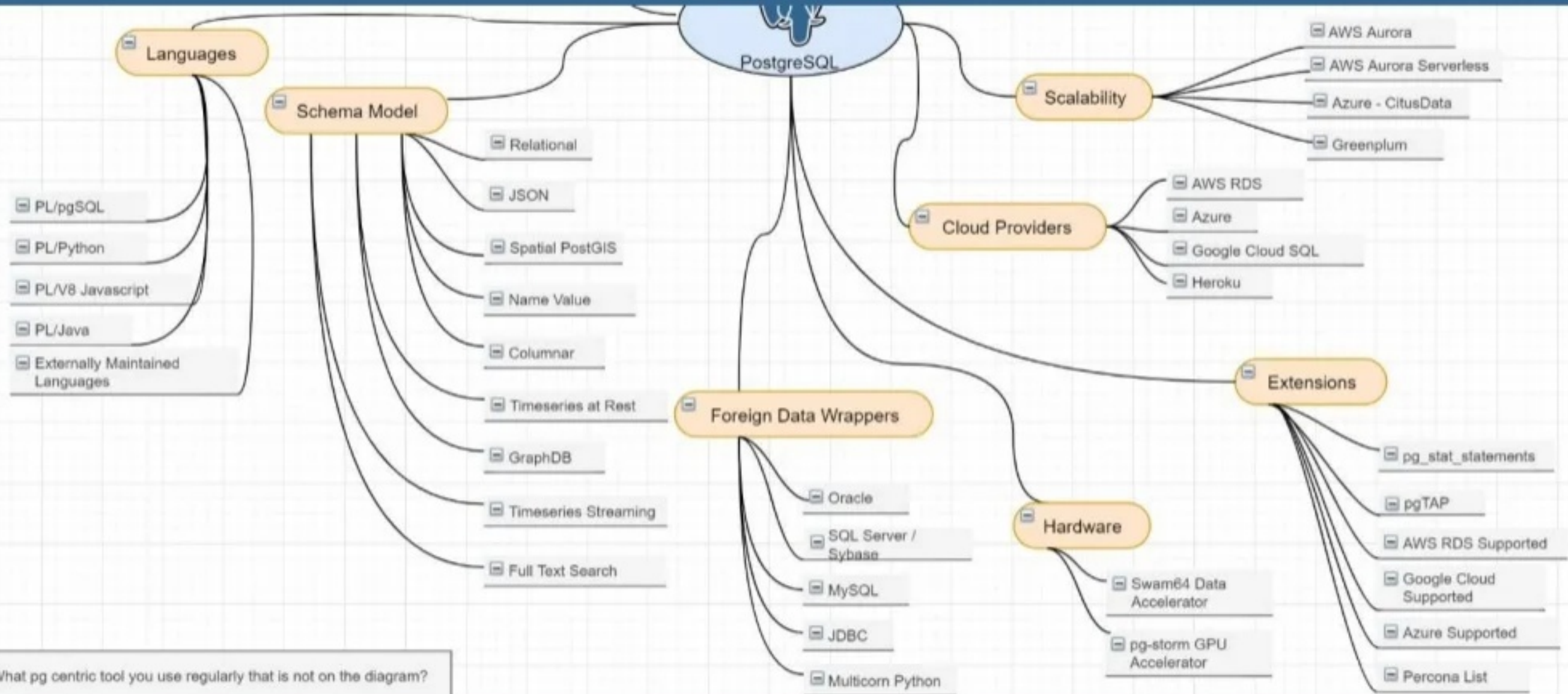
- <https://code.efficiencygeek.com/>





Overview – PostgreSQL ecosystem (3/3)

- <https://code.efficiencygeek.com/>





Agenda

01. Load Balancers
02. Replication
03. Rest API
04. Schema Model
05. Scalability
06. Foreign Data Wrappers
07. Others & support



Agenda

01. Load Balancers

02. Replication

03. Rest API

04. Schema Model

05. Scalability

06. Foreign Data Wrappers

07. Others & support



01. Load Balancers

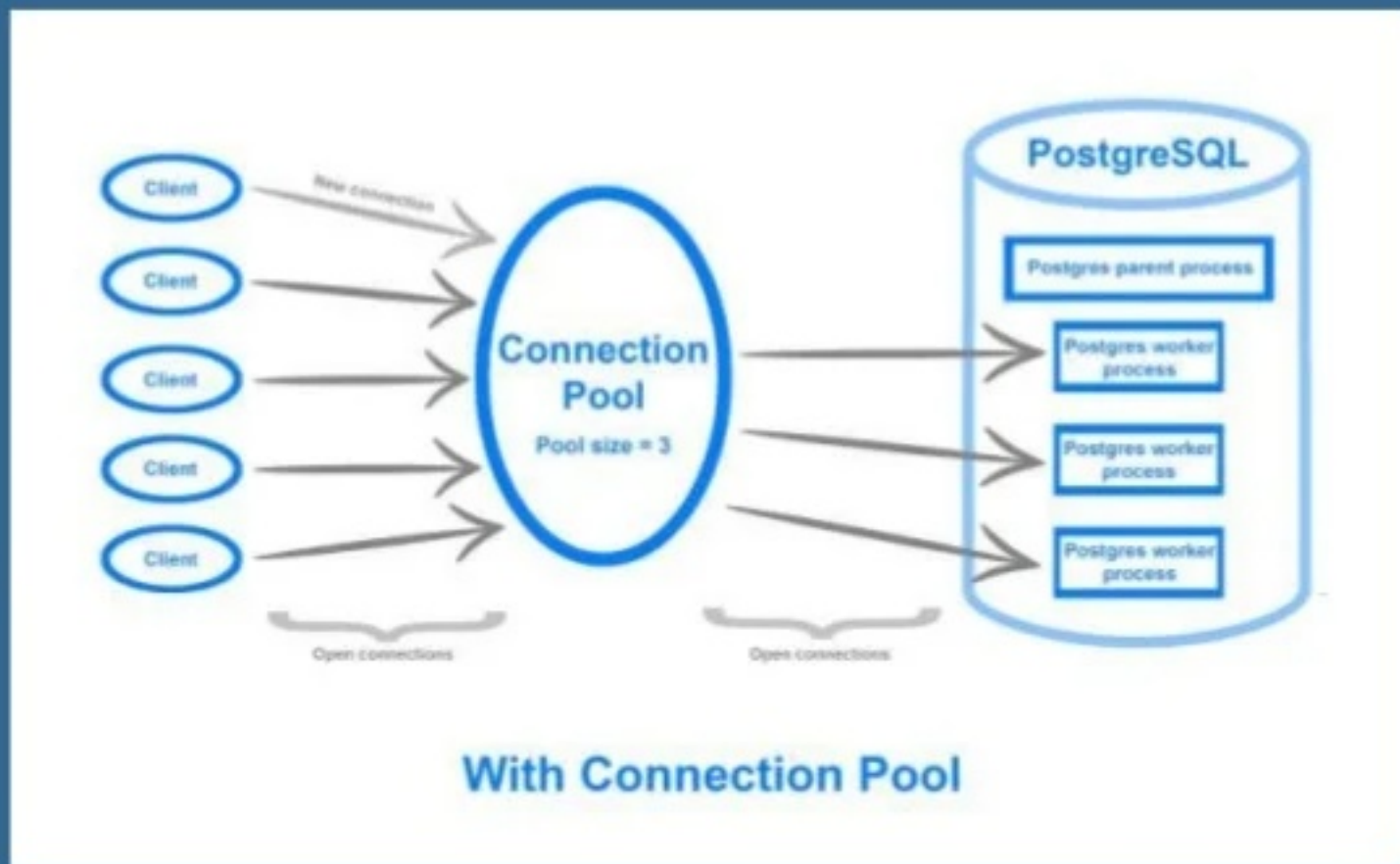
PgBouncer

- <https://www.pgouncer.org/>
- 1.15 - Nov 19, 2020
- 輕量級的 connection pool
- 提供重複使用 session 機制
 - PG 9.5 沒有自動於時斷線機制
 - PG 9.6 Transaction Timeout 參數

```
# vim /etc/pgbouncer/pgbouncer.ini

[database]
Db1 = port=5432 dbname=demo_db1
db2 = host=127.0.0.1 port=5432 dbname=demo_db2

[pgbouncer]
listen_addr = *
listen_port = 6432
auth_type = md5
auth_file = /etc/pgbouncer/userlist.txt
pool_mode = session
```



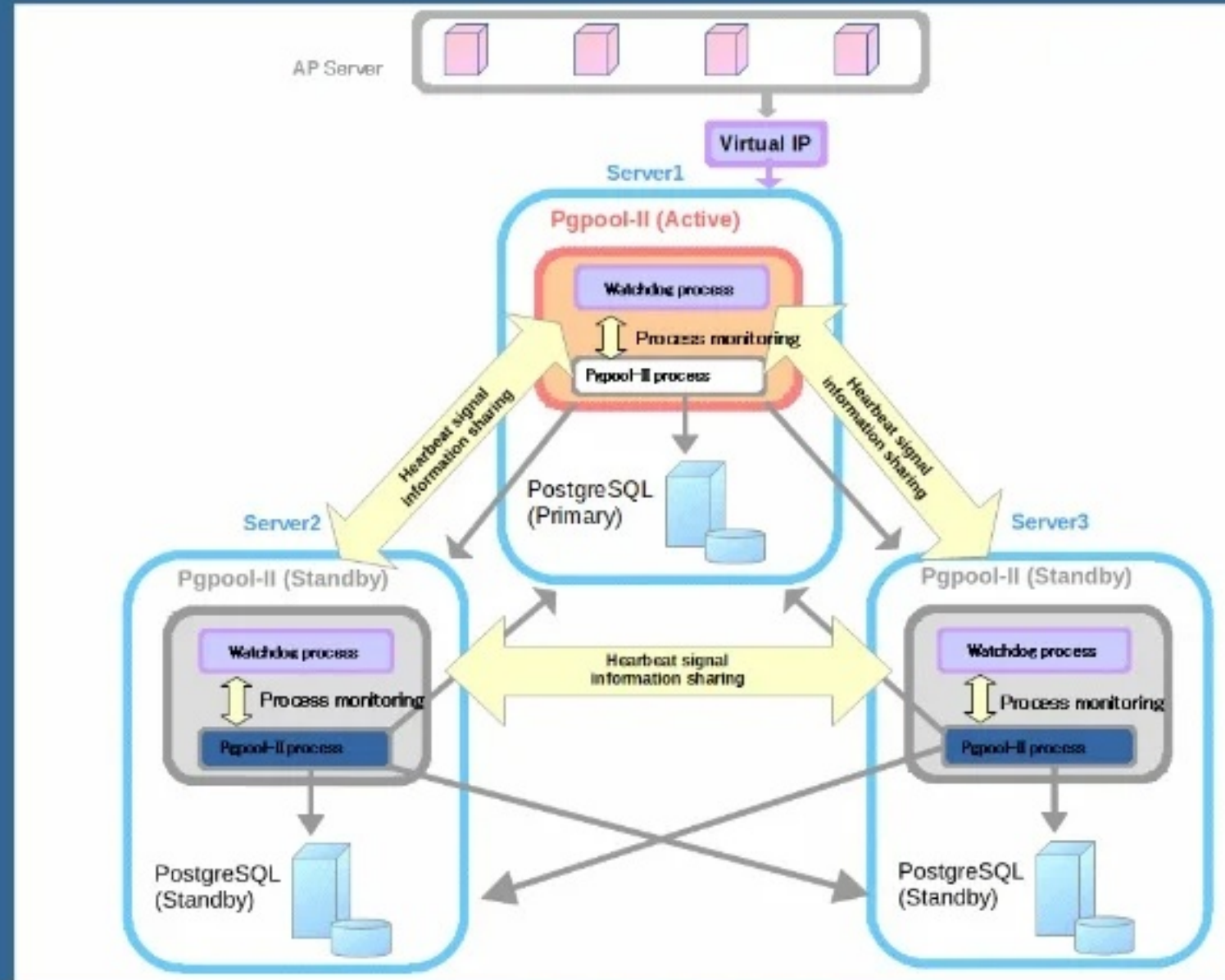
<https://mlog.club/article/2987957>



01. Load Balancers

Pgpool-II

- <https://pgpool.net>
- 4.2.3 - May 20, 2021
- Connection Pooling
- Load Balance
- Replication
- Parallel Query
- pgpoolAdmin



<https://www.pgpool.net/docs/pgpool-II-4.1.0/en/html/example-cluster.html>



Agenda

01. Load Balancers

02. Replication

03. Rest API

04. Schema Model

05. Scalability

06. Foreign Data Wrappers

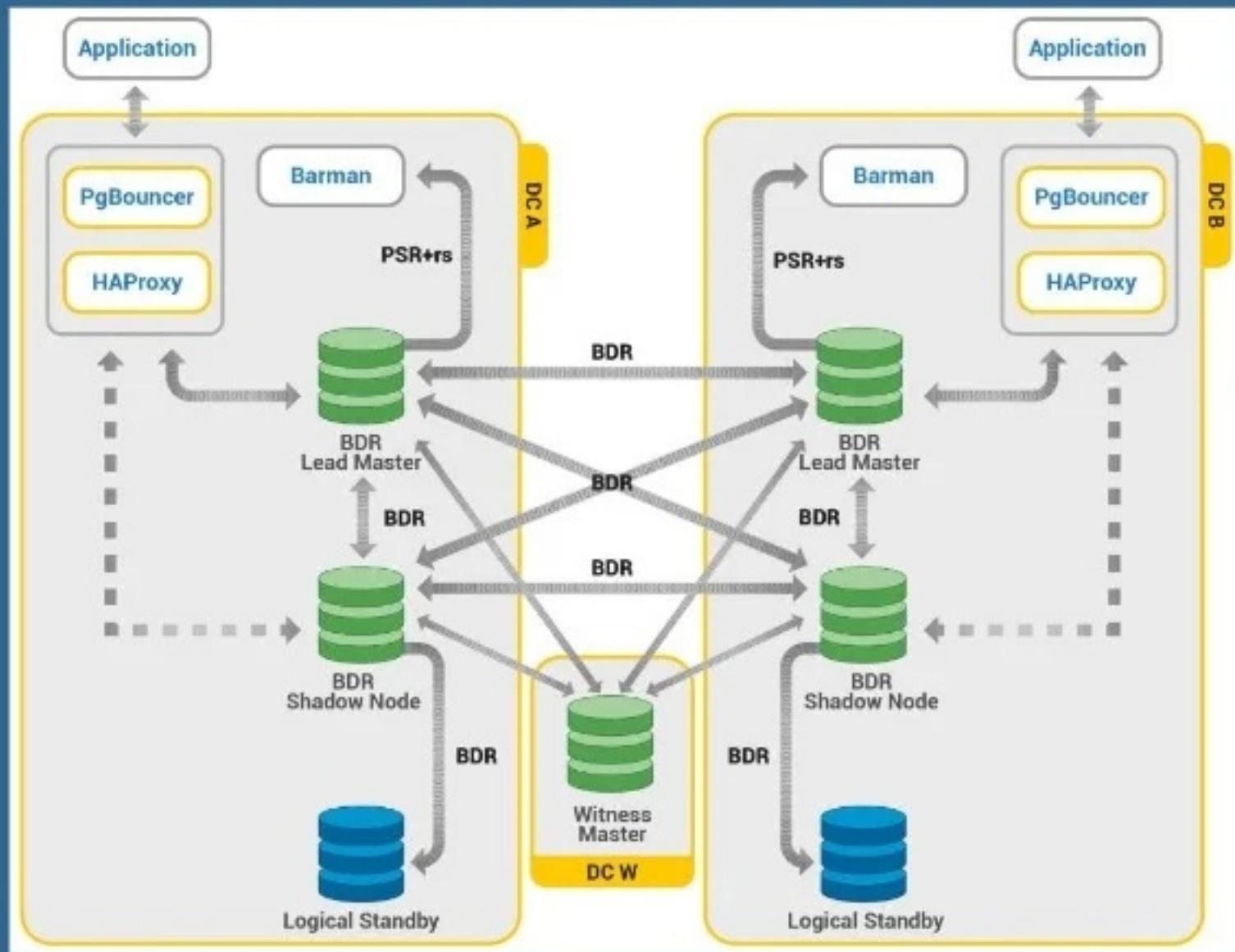
07. Others & support



02. Replication

Postgres BDR

- 2ndQuadrant
- github.com/2ndQuadrant/bdr
- 1.15 - Nov 19, 2020
- Bi-Directional Replication
- Async multi-master replication
- 最終一致性
- 默認以最後的更新為準
- BDR使用場景
 - 有利於: 單一節點修改, 減少衝突
 - 不利於: 同一時間查詢各節點之修改



<https://www.2ndquadrant.com/en/resources/highly-available-postgresql-clusters/>



Agenda

01. Load Balancers

02. Replication

03. Rest API

04. Schema Model

05. Scalability

06. Foreign Data Wrappers

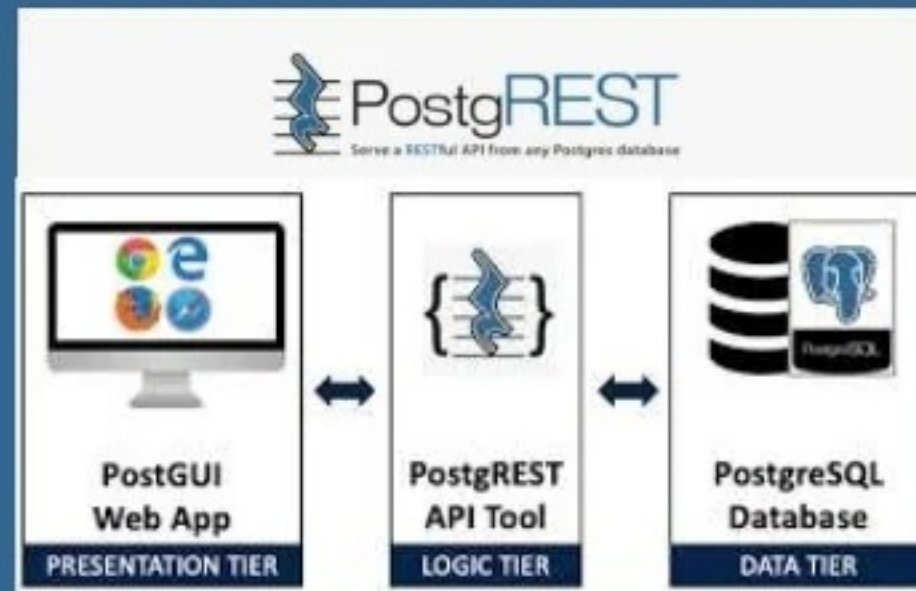
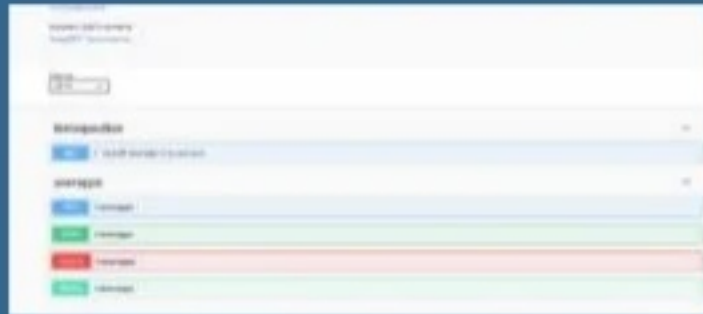
07. Others & support



03. Rest API

PostgREST

- <https://github.com/PostgREST>
- 7.0.1 - May 19, 2020
- 連接資料表直接提供RESTful API
- 可提供 Swagger ui



Purohit_Priyank_Kamlesh_202011_MSc_thesis.pdf

Movie listing

Search

Id	title	description	Actors
1	Academy Dinosaur	A Epis Drama of a Feminist And a Mad Scientist who must Battle a Teacher in The Canadian Rockies	Penelope Guinness, Christian Gable and 8 others
2	Ace Goldfinger	A Astounding Epistle of a Database Administrator And a Explorer who must Find a Car in Ancient China	Bob Fawcett, Minnie Zellweger and 2 others
3	Adaptation Holes	A Astounding Reflection of a Lumberjack And a Car who must Sink a Lumberjack in A Baloon Factory	Nick Wahlberg, Bob Fawcett and 3 others
4	Affair Prejudice	A Fanciful Documentary of a Frisbee And a Lumberjack who must Chase a Monkey in A Shark Tank	Jodie Degeneres, Scarlett Damon and 3 others
5	African Egg	A Fast-Paced Documentary of a Pastry Chef And a Dentist who must Pursue a Forensic Psychologist in The Gulf of Mexico	Gary Phoenix, Dustin Tautou and 3 others

Movie listing

Search

Id	title	description	Actors
1	Academy Dinosaur	A Epic Drama of a Feminist And a Mad Scientist who must Battle a Teacher in The Canadian Rockies	Penelope Guinness, Christian Gable and 8 others
2	Ace Goldfinger	A Astounding Epistle of a Database Administrator And a Explorer who must Find a Car in Ancient China	Bob Fawcett, Minnie Zellweger and 2 others
3	Adaptation Holes	A Astounding Reflection of a Lumberjack And a Car who must Sink a Lumberjack in A Baloon Factory	Nick Wahlberg, Bob Fawcett and 3 others
4	Affair Prejudice	A Fanciful Documentary of a Frisbee And a Lumberjack who must Chase a Monkey in A Shark	Jodie Degeneres, Scarlett

<https://github.com/tatut/postgrest-ui>



Agenda

- 01. Load Balancers
- 02. Replication
- 03. Rest API
- 04. Schema Model**
- 05. Scalability
- 06. Foreign Data Wrappers
- 07. Others & support

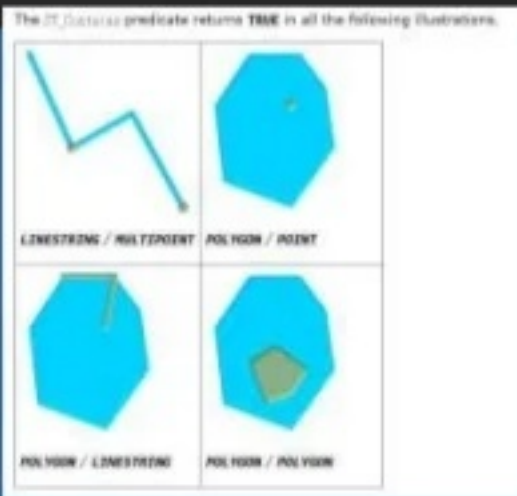


04. Schema Model

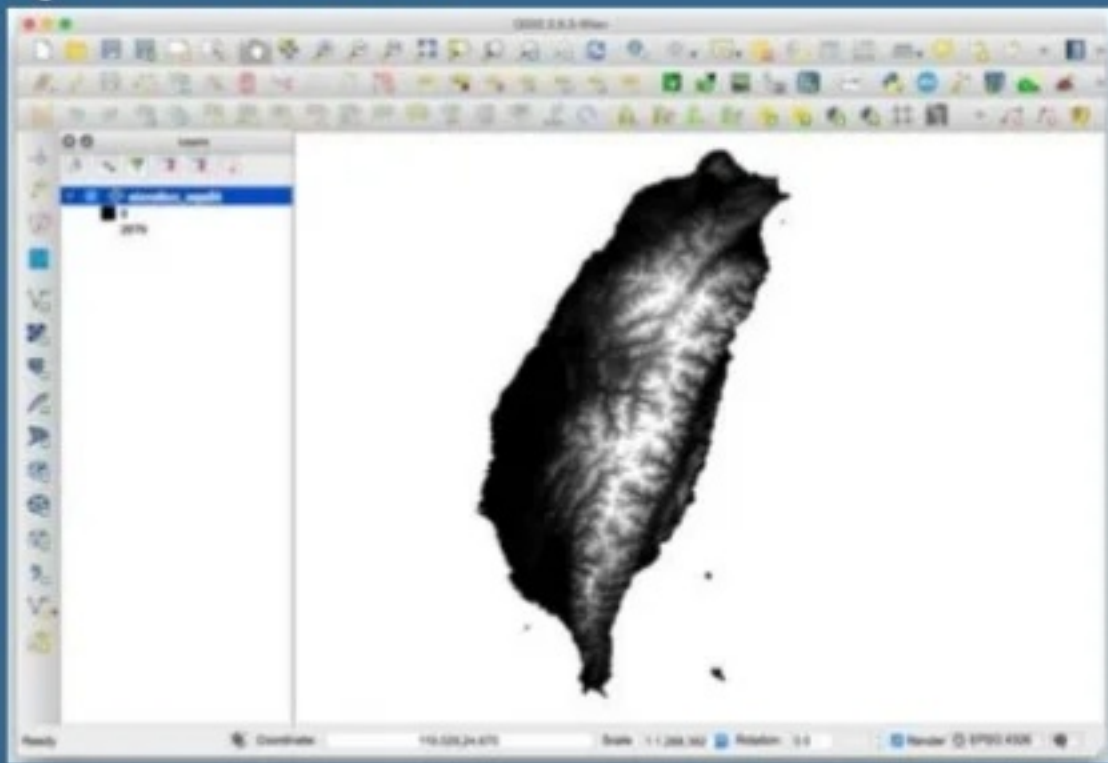
PostGIS

- <https://postgis.net/>
- 3.1.3 – July 20, 2020
- 地理圖資資料庫模組
(Spatial Database)
- 商用GIS的最佳替代方案

```
SELECT superhero.name  
FROM city, superhero  
WHERE ST_Contains(city.geom, superhero.geom)  
AND city.name = 'Gotham';
```



QGIS 讀取 PostGIS



<https://www.slideshare.net/mutolisp/postgis-57051746>



04. Schema Model

PipelineDB

- <https://github.com/pipelinedb/>
- 1.0 – Feb 20, 2019
- 基於PG的時序型資料庫

```
pipeline -c "SELECT sum(count) FROM test_view"
sum
```

```
-----
100001
(1 row)
```

```
pipeline -c "SELECT * FROM test_view ORDER BY
count DESC limit 10"
```

```
key | count
```

```
-----+-----
```

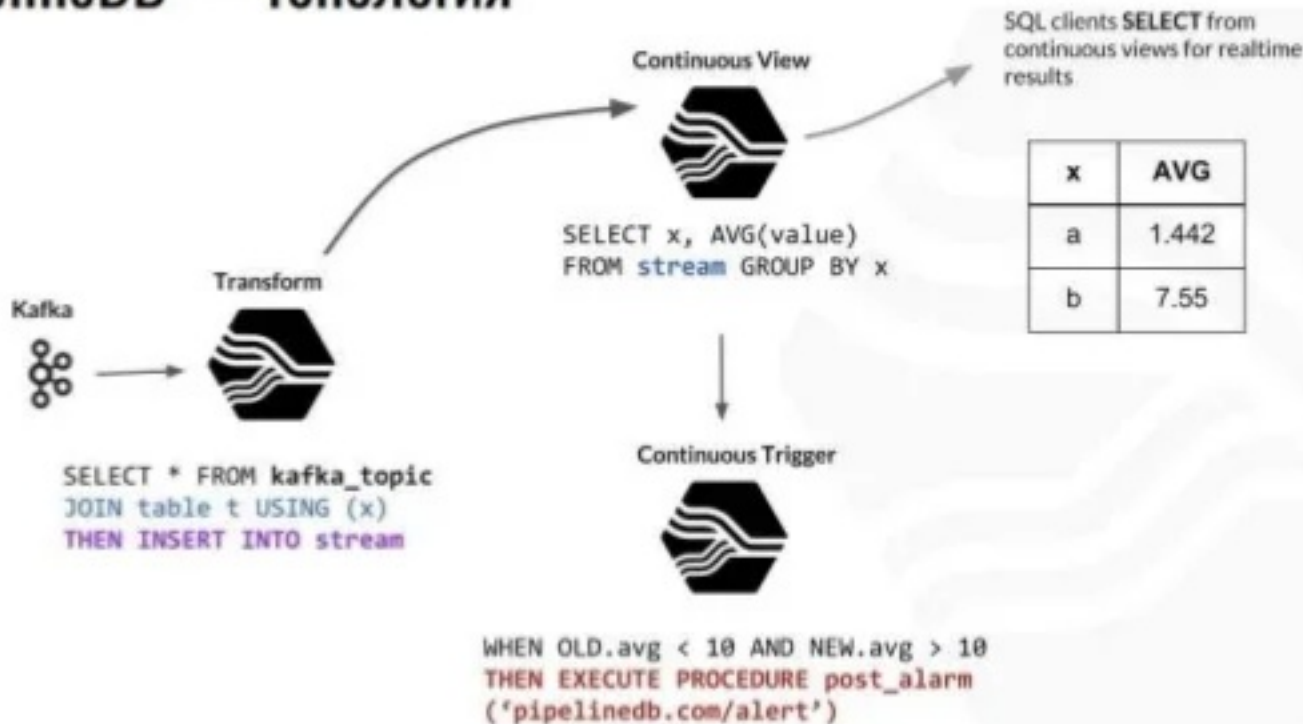
```
2 | 10124
8 | 10100
1 | 10042
7 | 9996
4 | 9991
5 | 9977
3 | 9963
6 | 9927
9 | 9915
10 | 4997
0 | 4969
```

```
(11 rows)
```



PIPELINE^{DB}

PipelineDB — топология

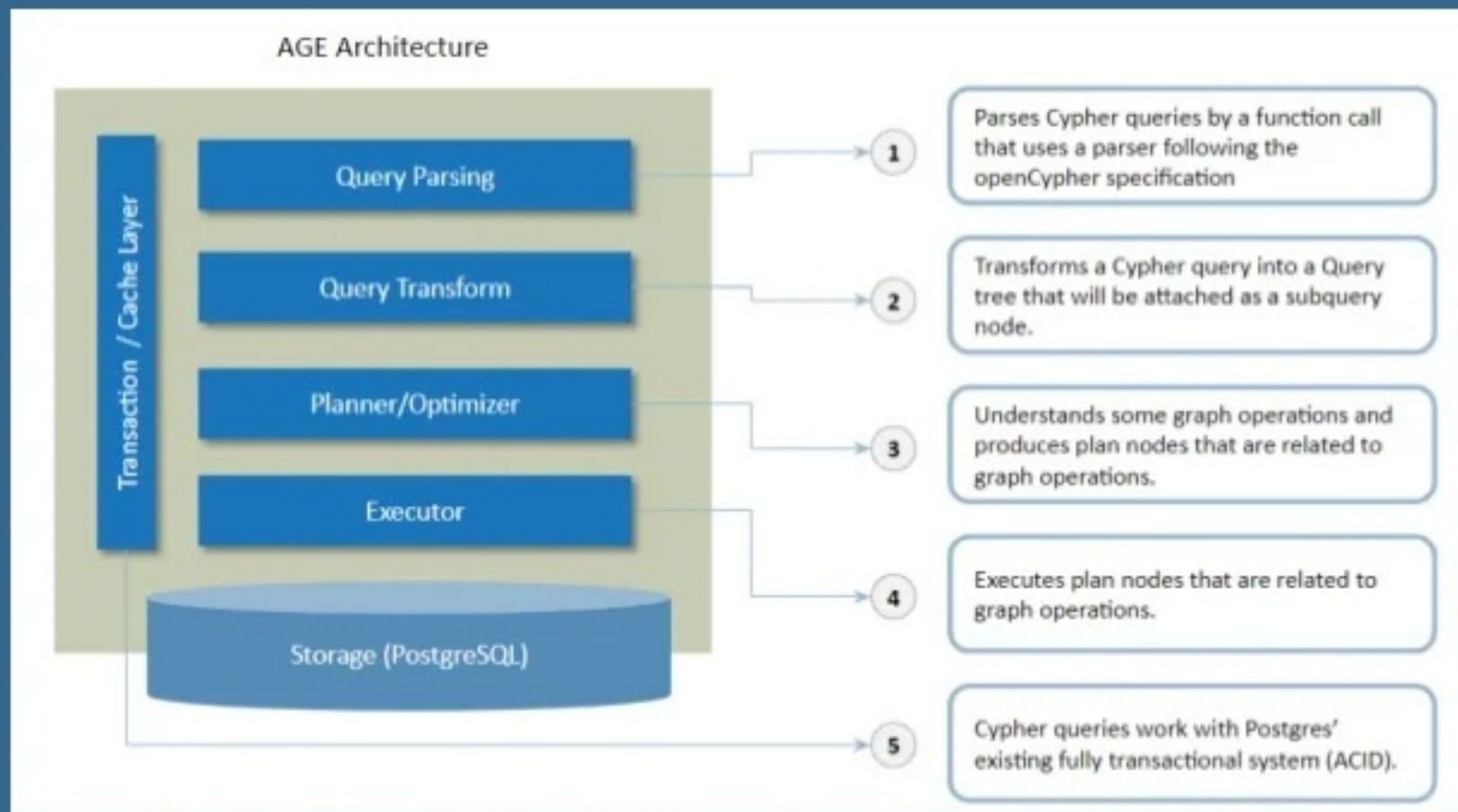




04. Schema Model

Apache AGE (Incubating)

- <https://age.incubator.apache.org/>
- v0.5.0 – Jun 30, 2021
- A Graph Extension for PG
提供 graph function
- 從PG10開始開發，
目前基於PG11開發中





Agenda

01. Load Balancers

02. Replication

03. Rest API

04. Schema Model

05. Scalability

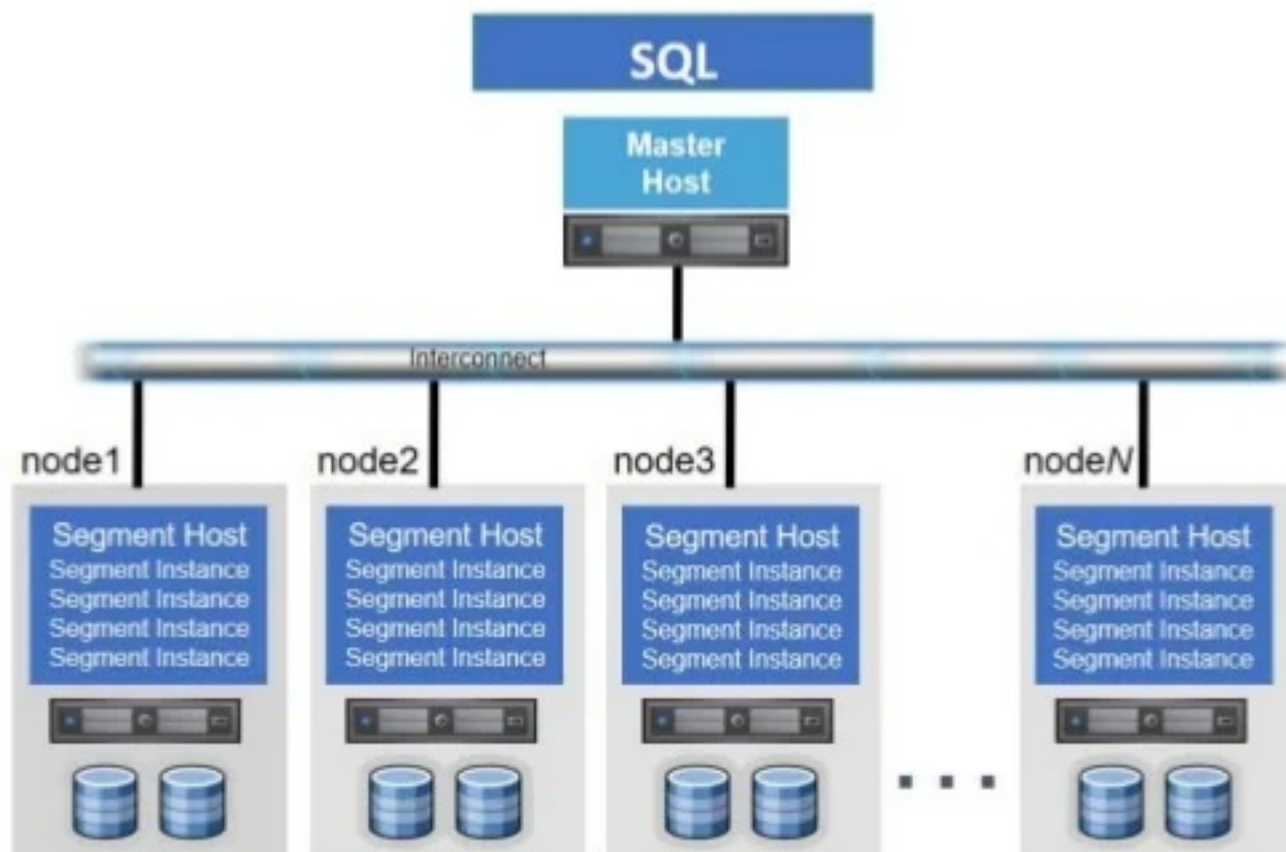
06. Foreign Data Wrappers

07. Others & support



Greenplum DB

- <https://greenplum.org/>
- 6.7.1 – April 2020
- based on PG 9.4
- 架構採用MPP (Massively Parallel Processor)
- 在 MPP 系統中， 每個 SMP 節點也可以運作自己的作業系統、資料庫等
- Data distribution
- Shared nothing
- MapReduce



<https://kknews.cc/tech/5meq9l3.html>



05. Scalability

Postgres-XC / XC2 / XL

- XC <https://sourceforge.net/projects/postgres-xc/>
- XC2 <https://github.com/postgres-x2/>
- XL <https://www.postgres-xl.org/>
- Data distribution
- Shared nothing
- GTM
- Coordinator
- Global Catalog

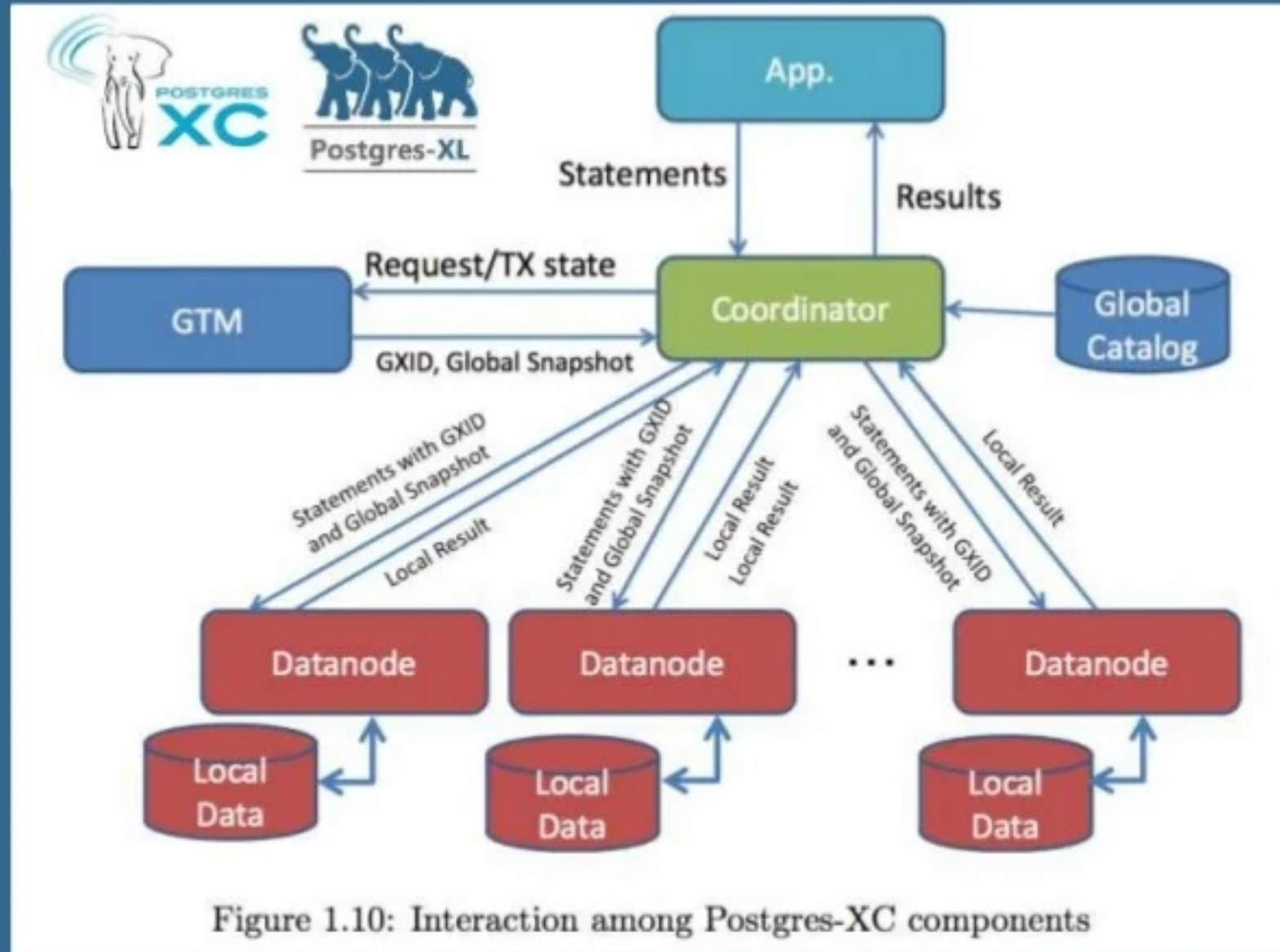


Figure 1.10: Interaction among Postgres-XC components

<https://zhuanlan.zhihu.com/p/26391728>



Agenda

01. Load Balancers

02. Replication

03. Rest API

04. Schema Model

05. Scalability

06. Foreign Data Wrappers

07. Others & support



06. Foreign Data Wrappers

PG9.1 FDW - SELECT

PG9.3 FDW -INSERT / UPDATE / DELETE

https://wiki.postgresql.org/wiki/Foreign_data_wrappers

- Generic/Specific SQL Database Wrappers
- NoSQL Database Wrappers
- File / Geo / LDAP Wrappers
- Generic/Specific Web Wrappers
- Big Data Wrappers
- Column-Oriented Wrappers
- Scientific Wrappers
- Operating System Wrappers
- Exotic Wrappers





06. Foreign Data Wrappers

Generic / Specific SQL Database Wrappers

Generic SQL Database Wrappers

Data Source	Type	License	Code	Install	Doc	Notes
ODBC	Native		github			CartoDB took over active development of the ODBC FDW for PG 9.5+
JDBC	Native		github			Not maintained ?
JDBC2	Native		github			
SQLAlchemy	Multicorn	PostgreSQL	GitHub	PGXN	documentation	Can be used to access data stored in any database supported by the sqlalchemy python toolkit.
GDAL/OGR	Native	MIT	GitHub	yum.postgresql.org , apt.postgresql.org , and part of PostGIS windows bundle (application stackbuilder)		Can access many kinds of data sources (Relational databases, spreadsheets, CSV files, web feature services, etc). Uses the GDAL library which supports hundreds of formats to access the data. Exposes vector data as PostGIS geometry columns if you have PostGIS installed. Works great with both spatial and non-spatial data.
VirtDB	Native	GPL	GitHub			A generic FDW to access VirtDB data sources (SAP ERP, Oracle RDBMS)

Specific SQL Database Wrappers

Data Source	Type	License	Code	Install	Doc	Notes
PostgreSQL	Native	PostgreSQL	git postgresql.org		documentation	
Oracle	Native	PostgreSQL	github	PGXN	website	
MySQL	Native		github	PGXN	example	FDW for MySQL
Informix	Native	PostgreSQL	github			
Firebird	Native	PostgreSQL	github	PGXN	README	version 1.2.0 released (2020-10)
SQLite	Native	PostgreSQL	github	PGXN	README	An FDW for SQLite3 (write support and several pushdown optimization)
Sybase / MS SQL Server	Native		github	PGXN		An FDW for Sybase and Microsoft SQL server
MonetDB	Native		github			



06. Foreign Data Wrappers

NoSQL Database Wrappers

Data Source	Type	License	Code
BigTable or HBase ↗	Native Rust Binding (RPGFFI) ↗	MIT	Github ↗
Cassandra ↗	Multicorn ↗	MIT	Github ↗
Cassandra2	Native	MIT	Github ↗
Cassandra ↗	Multicorn ↗	PostgreSQL	Github ↗
ClickHouse ↗	Multicorn ↗	BSD	Github ↗
ClickHouse ↗	Native	Apache	Github ↗
CouchDB ↗	Native	PostgreSQL	Github ↗
CouchDB ↗	Native	PostgreSQL	Github ↗
GridDB ↗	Native	PostgreSQL	Github ↗
InfluxDB	Native	PostgreSQL	Github ↗
Kafka ↗	Native	PostgreSQL	Git Hub ↗
Kyoto Tycoon ↗	Native	MIT	Github ↗

Data Source	Type	License	Code
MongoDB ↗	Native	GPL3+	Github ↗
MongoDB ↗	Multicorn ↗	MIT	Github ↗
MongoDB ↗	Multicorn ↗		Github ↗
Neo4j ↗	Multicorn ↗	GPLv3	Github ↗
Neo4j ↗	Native	?	Github ↗
Quasar ↗	Native	Apache	Github ↗
Redis ↗	Native	PostgreSQL	Github ↗
Redis ↗	Native	BSD	Github ↗
RethinkDB ↗	Multicorn ↗	MIT	Github ↗
Riak ↗	Multicorn ↗	PostgreSQL	Github ↗
WhiteDB ↗	Native	MIT	Github ↗
RocksDB ↗	Native	Apache	Github ↗



06. Foreign Data Wrappers

File/Geo/LDAP Database Wrappers

File Wrappers

Data Source	Type	License	Code	Install	Doc
CSV	Native	PostgreSQL	git.postgresql.org		documentation
CSV	Multicorn	PostgreSQL	GitHub	PGXN	documentation
CSV / Text Array	Native		GitHub		How to
CSV / Fixed-length	Native		GitHub		
CSV / gzipped	Multicorn		GitHub		
Compressed File	Native		GitHub		
Document Collection	Native	PostgreSQL	GitHub		wiki
JSON	Native	GPL3	GitHub		Example
Multi-File	Multicorn	PostgreSQL	GitHub	PGXN	doc
Multi CDR	Native	PostgreSQL	GitHub	PGXN	
Parquet	Native	PostgreSQL	GitHub		
pg_dump	Native	New BSD	GitHub		
TAR Files	Native		GitHub		
XML	Multicorn	PostgreSQL	GitHub	PGXN	
ZIP Files	Native		GitHub		

Geo Wrappers

Data Source	Type	License	Code	Install	Doc
GDAL/OGR	Native	MIT	GitHub		
Geocode / GeoJSON	Multicorn	GPL	GitHub		
Open Street Map PBF	Native	PostgreSQL	GitHub		

LDAP Wrappers

Data Source	Type	License	Code	Install	Doc
LDAP	Native		GitHub	PGXN	
LDAP	Multicorn	PostgreSQL	GitHub	PGXN	documentation



06. Foreign Data Wrappers

Generic / Specific Web Database Wrappers

Generic Web Wrappers

Data Source	Type	License	Code	Install	Doc
Git	Multicorn	PostgreSQL	GitHub	PGXN	
Git	Native	MIT	GitHub		
ICAL	Multicorn	PostgreSQL	GitHub		pdf
IMAP	Multicorn	PostgreSQL	GitHub	PGXN	documentation
RSS	Multicorn	PostgreSQL	GitHub	PGXN	documentation
www	Native	PostgreSQL	GitHub	PGXN	wiki
pgsql-http	Native	PostgreSQL	GitHub	Compile	

Specific Web Wrappers

Data Source	Type	License	Code	Install	Doc
Database.com	Multicorn	BSD	GitHub		
Dun & Badstreet	Multicorn	PostgreSQL	GitHub		
DynamoDB	Multicorn	GPL	GitHub		
Facebook	Multicorn		GitHub		
Fixer.io	based on www_fdw		GitHub		
Google	Multicorn	PostgreSQL	GitHub	PGXN	
Heroku dataclips	Native	PostgreSQL	GitHub		
Keycloak	Multicorn	MIT	GitHub	PGXN	README
Mailchimp	Multicorn	PostgreSQL	GitHub		
Parse	Multicorn	MIT	GitHub		
S3	Native	PostgreSQL	GitHub	PGXN	
S3CSV	Multicorn	GPL 3	GitHub		
Telegram	Multicorn	PostgreSQL	GitHub		
Twitter	Native	PostgreSQL	GitHub	PGXN	
Treasure Data	Native	Apache	GitHub	PGXN	
Treasure Data	Multicorn	Apache	GitHub		
Google Spreadsheets	Multicorn	MIT	GitHub		
Open Weather Map	Multicorn	MIT	GitHub		



Generic / Specific Web Database Wrappers

Big Data Wrappers

Data Source	Type	License	Code	Install
Elasticsearch	Multicorn	MIT	GitHub	
Google BigQuery	Multicorn	MIT	GitHub	
file_fdw-gds (Hadoop)	Native		GitHub	
Hadoop	Native	PostgreSQL	Bitbucket	
HDFS	Native	Apache	GitHub	
Hive	Multicorn		GitHub	
Hive / ORC File	Native		GitHub	
Impala	Native	BSD	GitHub	
Apache Arrow	Native	GPLv2	GitHub	

Column-Oriented Wrappers

Data Source	Type	License	Code	Install	Doc
Columnar Store	Native		github	example	
MonetDB	Native		github		
GPU Memory Store	Native	GPL v2	github		

Scientific Wrappers

Data Source	Type	License	Code	Install
Ambry	Multicorn		GitHub	
ROOT files	Native		GitHub	
VCF files (Genotype)	Multicorn		GitHub	

Operating System Wrappers

Data Source	Type	License	Code
Docker	Multicorn	Expat	GitHub
Log files	Multicorn	PostgreSQL	GitHub
OpenStack / Telemetry	Multicorn	PostgreSQL	GitHub
OS Query	Multicorn	PostgreSQL	GitHub
Passwd	Native	PostgreSQL	GitHub
Process	Multicorn	PostgreSQL	GitHub
Environment Variables	Multicorn	MIT	GitHub



Agenda

- 01. Load Balancers
- 02. Replication
- 03. Rest API
- 04. Schema Model
- 05. Scalability
- 06. Foreign Data Wrappers
- 07. Others & support



07. Others & support

Apache MADlib



The screenshot shows the Apache MADlib website. At the top is a navigation bar with the MADlib logo on the left and links for Home, Product, Documentation, Community, and a red Download button on the right. Below the navigation bar is a large teal banner with the text "Apache MADlib: Big Data Machine Learning in SQL". Underneath this banner are three columns of text: "Open source, commercially friendly Apache license", "For PostgreSQL and Greenplum Database®", and "Powerful machine learning, graph, statistics and analytics for data scientists". To the right of these columns is a "Read More" link. Below the banner is a section titled "Getting Started with Apache MADlib using Jupyter Notebooks" with a paragraph of text. This is followed by a section titled "MADlib 1.18.0 Release" with a paragraph of text and a "New features include:" heading.

Apache MADlib

Home Product Documentation Community [Download](#)

Apache MADlib: Big Data Machine Learning in SQL

- Open source, commercially friendly Apache license
- For PostgreSQL and Greenplum Database®
- Powerful machine learning, graph, statistics and analytics for data scientists

[Read More](#)

Getting Started with Apache MADlib using Jupyter Notebooks

We have created a [library of Jupyter Notebooks](#) to help you get started quickly with MADlib. It includes many commonly used algorithms by data scientists.

MADlib 1.18.0 Release

On April 5, 2021, MADlib completed its eighth release as an Apache Software Foundation Top Level Project.

New features include:



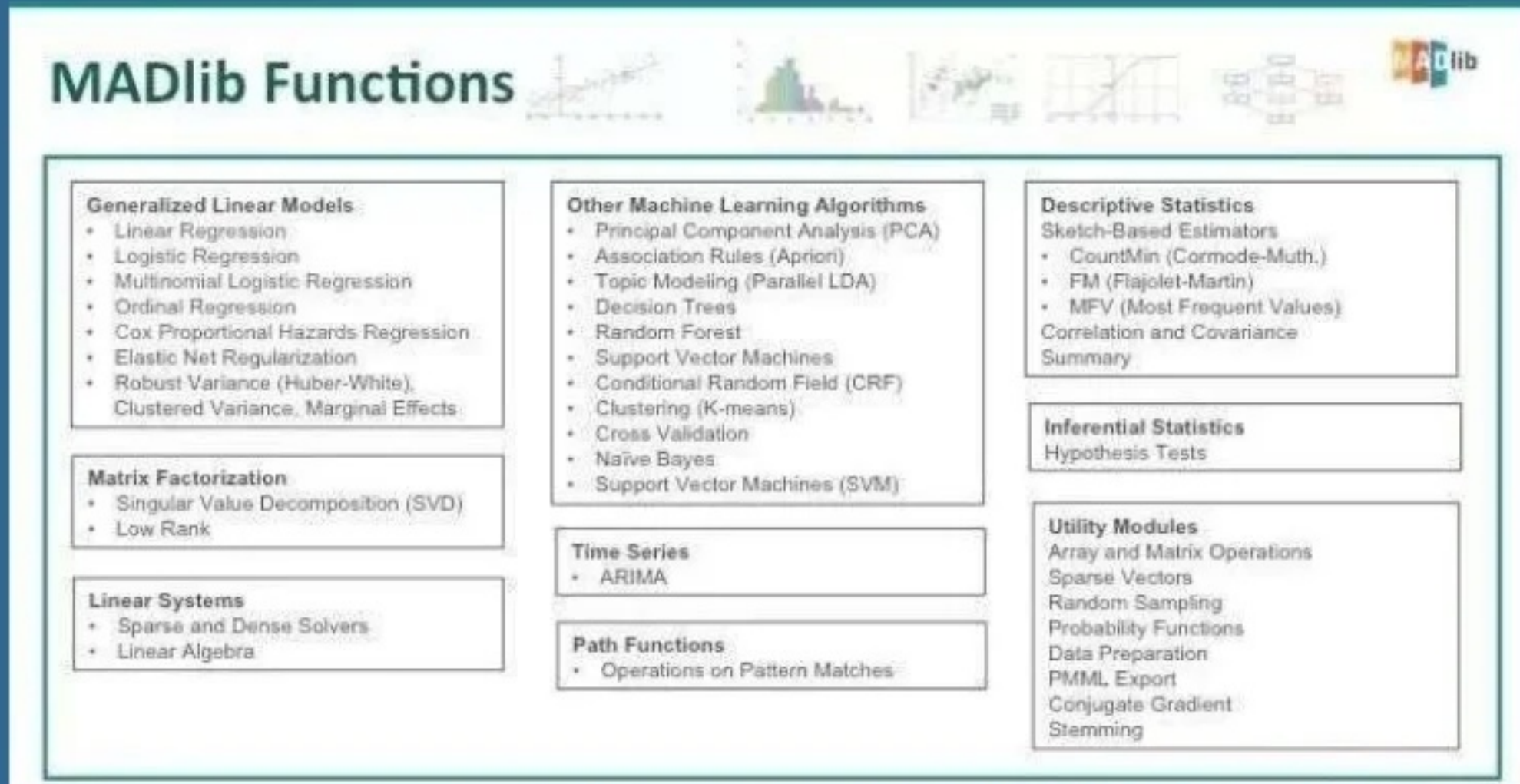
Apache MADlib

Functions

- 機器學習
- 矩陣分析
- 統計分析
- ...

支援

- PostgreSQL
- Greenplum



<https://www.xuehua.us/a/5ebf13a880cf9e4954ac1981>

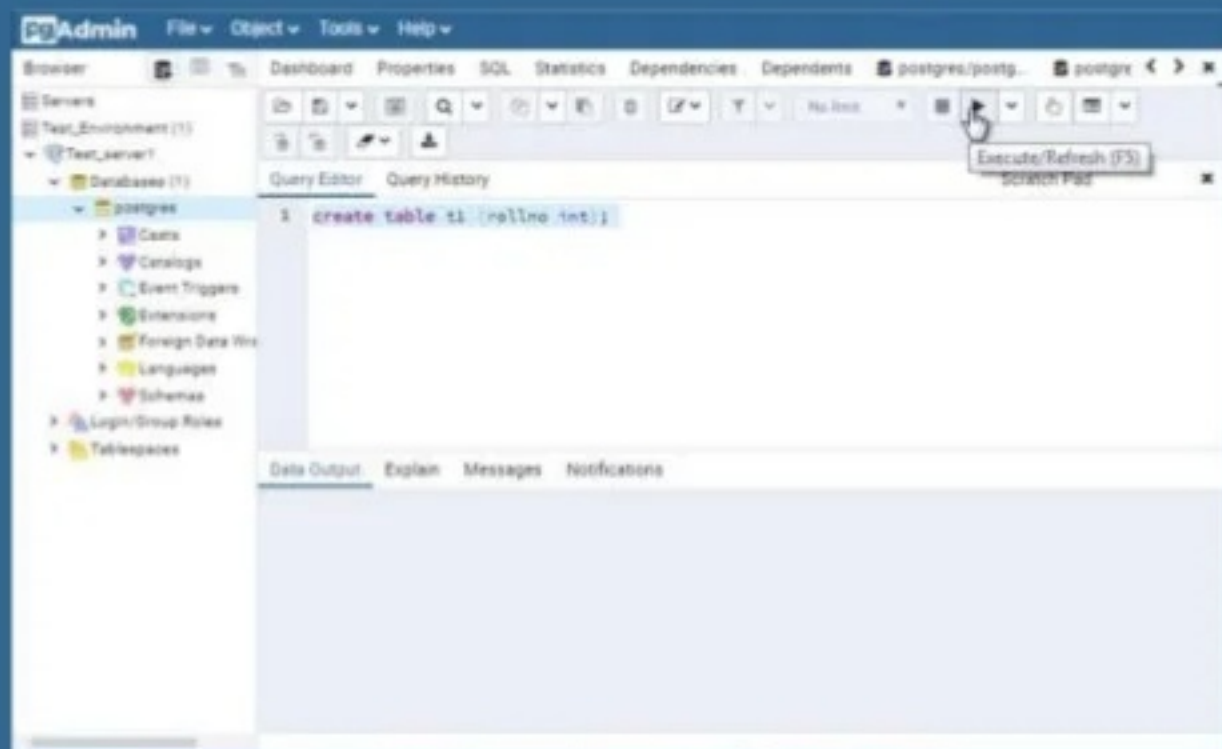


07. Others & support

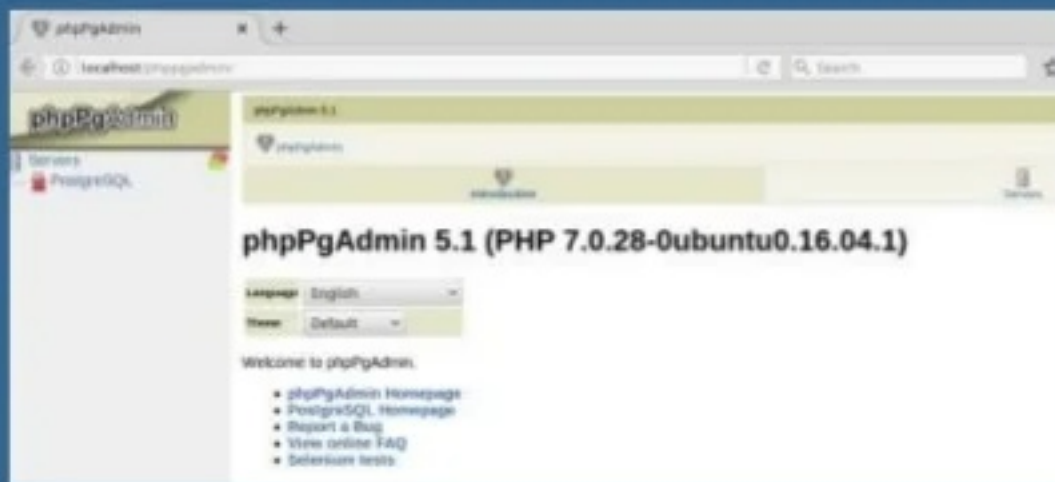
GUI / WebUI



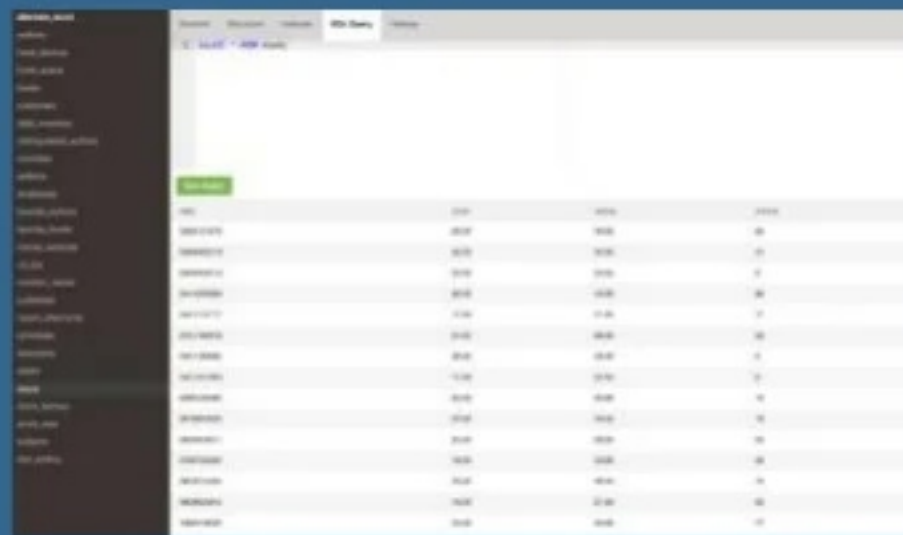
pgAdmin Version 4
Management Tools for PostgreSQL



PHPPGADMIN



 **pgweb**





有很多不得不推的理由

最開源、最穩定、最先進...

- Data Types
- Data Integrity
- Concurrency and Performance
- Reliability, Disaster Recovery
- Security
- Extensibility
- Internationalization, Text Search
- Not relational data (JSON, Hstore, Cstore)
- GIS (Geographic Information System)

Why should I use PostgreSQL as Database in my Startup/Company



Lucio Chiessi

Follow

Sep 21, 2018 · 4 min read





持續貢獻

DBMS of the Year: PostgreSQL



DB-engine ranking

- <https://db-engines.com/en/ranking>

373 systems in ranking, July 2021

Rank			DBMS	Database Model	Score		
Jul 2021	Jun 2021	Jul 2020			Jul 2021	Jun 2021	Jul 2020
1.	1.	1.	Oracle +	Relational, Multi-model ⓘ	1262.66	-8.28	-77.59
2.	2.	2.	MySQL +	Relational, Multi-model ⓘ	1228.38	+0.52	-40.13
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model ⓘ	981.95	-9.12	-77.77
4.	4.	4.	PostgreSQL +	Relational, Multi-model ⓘ	577.15	+8.64	+50.15
5.	5.	5.	MongoDB +	Document, Multi-model ⓘ	496.16	+7.95	+52.68
6.	↑ 7.	↑ 8.	Redis +	Key-value, Multi-model ⓘ	168.31	+3.06	+18.26
7.	↓ 6.	↓ 6.	IBM Db2	Relational, Multi-model ⓘ	165.15	-1.88	+1.99
8.	8.	↓ 7.	Elasticsearch +	Search engine, Multi-model ⓘ	155.76	+1.05	+4.17
9.	9.	9.	SQLite +	Relational	130.20	-0.33	+2.75
10.	↑ 11.	10.	Cassandra +	Wide column	114.00	-0.11	-7.08



期許...

- 持續為推廣而奔走，期許...
 - 國內使用Oracle的辛苦業者，改用 PG
 - 鼓勵新創在初期直接採用 PG
 - 資料庫教育直接使用 PG 做為教材
 - 與國內 SI 共同合作發展台版 PG



Thank you.

PostgreSQL ecosystem
@ COSCUP x RubyConf TW 2021

Taiwan PostgreSQL User Group
林宗禧 linjose@postgresql.tw

