**PostgreSQL DB Recommendation Document**



**Product Description:**

 PostgreSQL is an open source object-relational database system.

* It runs on all major operating systems, including Linux, UNIX (AIX, BSD, HP-UX, SGI IRIX, Mac OS X, Solaris, Tru64), and Windows.
* It is fully ACID compliant, has full support for foreign keys, joins, views, triggers, and stored procedures (in multiple languages).
* It includes most SQL:2008 data types, including INTEGER, NUMERIC, BOOLEAN, CHAR, VARCHAR, DATE, INTERVAL, and TIMESTAMP.
* It also supports storage of binary large objects, including pictures, sounds, or video.
* It has native programming interfaces for C/C++, Java, .Net, Perl, Python, Ruby, Tcl, ODBC, among others.

An enterprise class database, PostgreSQL boasts sophisticated features such as MVCC, point in time recovery, tablespaces, asynchronous replication, nested transactions (savepoints),

Online/hot backups, a sophisticated query planner/optimizer, and write ahead logging for fault tolerance.

It supports international character sets, multibyte character encodings, Unicode, and it is locale-aware for sorting, case-sensitivity, and formatting.

It is highly scalable both in the sheer quantity of data it can manage and in the number of concurrent users it can accommodate.

There are active PostgreSQL systems in production environments that manage in excess of 4 terabytes of data.

**Limit Value**

Maximum Database Size Unlimited

Maximum Table Size 32 TB

Maximum Row Size 1.6 TB

Maximum Field Size 1 GB

Maximum Rows per Table Unlimited

Maximum Columns per Table 250 - 1600 depending on column types

Maximum Indexes per Table Unlimited

**A service-oriented Architecture applied on the Database.**

Cluster solution ( Linux/ EDB)

## Clustering

* [CitusDB](https://www.citusdata.com/citus-products/citusdb-software) - shards and replicates tables across a scalable, high availability cluster of commodity PostgreSQL servers and parallelizes queries for real-time SQL on big data.
* [Greenplum Database](http://www.greenplum.com/index.php?page=greenplum-database) (formerly Bizgres MPP), proprietary. Not so much a replication solution as a way to parallelize queries, and targeted at the data warehousing crowd. Similar to ExtenDB, but tightly integrated with PostgreSQL.
* [GridSQL for EnterpriseDB Advanced Server](http://www.enterprisedb.com/products/gridsql.do) (formerly ExtenDB)
* [HadoopDB](http://db.cs.yale.edu/hadoopdb/hadoopdb.html) - A MapReduce layer put in front of a cluster of postgres back end servers. Shared-nothing clustering.
* [PL/Proxy](https://wiki.postgresql.org/wiki/PL/Proxy) - database partitioning system implemented as PL language.
* [pg\_shard](https://www.citusdata.com/citus-products/pg-shard) - extension that shards and replicates table across many servers, can also [scale out Amazon RDS](https://www.citusdata.com/blog/14-marco/178-scaling-out-postgresql-on-amazon-rds-using-masterless-pg-shard)
* sequoia (jdbc, formerly known as c-jdbc)







Table partition : <http://www.postgresql.org/docs/9.1/static/ddl-partitioning.html>

Cluster solution : [https://wiki.postgresql.org/wiki/Replication,\_Clustering,\_and\_Connection\_Pooling](https://wiki.postgresql.org/wiki/Replication%2C_Clustering%2C_and_Connection_Pooling)

Storage : <http://www.postgresql.org/docs/9.1/static/storage-file-layout.html> (information)

**Licencing :**

PostgreSQL is released under the [PostgreSQL License](http://www.opensource.org/licenses/postgresql), a liberal Open Source license, similar to the BSD or MIT licenses.

PostgreSQL Database Management System (formerly known as Postgres, then as Postgres95)

Portions Copyright (c) 1996-2015, The PostgreSQL Global Development Group
Portions Copyright (c) 1994, The Regents of the University of California

Permission to use, copy, modify, and distribute this software and its documentation for any purpose, without fee, and without a written agreement is hereby granted, provided that the above copyright notice and this paragraph and the following two paragraphs appear in all copies.

IN NO EVENT SHALL THE UNIVERSITY OF CALIFORNIA BE LIABLE TO ANY PARTY FOR DIRECT, INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES, INCLUDING LOST PROFITS, ARISING OUT OF THE USE OF THIS SOFTWARE AND ITS DOCUMENTATION, EVEN IF THE UNIVERSITY OF CALIFORNIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

THE UNIVERSITY OF CALIFORNIA SPECIFICALLY DISCLAIMS ANY WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE SOFTWARE PROVIDED HEREUNDER IS ON AN "AS IS" BASIS, AND THE UNIVERSITY OF CALIFORNIA HAS NO OBLIGATIONS TO PROVIDE MAINTENANCE, SUPPORT, UPDATES, ENHANCEMENTS, OR MODIFICATIONS.

**Why not the GNU General Public License?**

The simple answer is because Postgres like thier license and do not want to change it

**Pros over MSSQL:**

1. CSV Support
2. PostgreSQL runs in Linux, BSD, Windows …etc
3. Complete Procedural language features.
4. Native regular expression support
5. Custom aggregate functions
6. Unicode support
7. Full support of Data types
8. It’s Opensource

More details & source at <http://www.pg-versus-ms.com/>

Over Sybase: [http://www.postgresql.org/message-id/4ED5FD9C020000250004363C@gw.wicourts.gov](http://www.postgresql.org/message-id/4ED5FD9C020000250004363C%40gw.wicourts.gov)

**Cons**:

The length of each lexeme must be less than 2K bytes

The length of a tsvector (lexemes + positions) must be less than 1 megabyte

The number of lexemes must be less than 264

Position values in tsvector must be greater than 0 and no more than 16,383

No more than 256 positions per lexeme

The number of nodes (lexemes + operators) in a tsquery must be less than 32,768

More details at <http://www.postgresql.org/docs/9.3/static/textsearch-limitations.html>

**FAQs:** [**http://www.postgresql.org/docs/faq/**](http://www.postgresql.org/docs/faq/)

**YCSB Benchmarks comparision between PostgreSQL & OracleDB:**

**PostgreSQL VM Details:**

 root @ dbversity : /YCSB-master/logs ] free -g

 total used free shared buffers cached

Mem: 3 3 0 0 0 2

-/+ buffers/cache: 0 3

Swap: 1 0 1

[ root @ dbversity : /YCSB-master/logs ]

[ root @ dbversity : /YCSB-master/logs ] cat /proc/cpuinfo |head

processor : 0

vendor\_id : GenuineIntel

cpu family : 6

model : 44

model name : Intel(R) Xeon(R) CPU X5650 @ 2.67GHz

stepping : 2

microcode : 26

cpu MHz : 2666.761

cache size : 12288 KB

physical id : 0

[ root @ dbversity : /YCSB-master/logs ]

[ root @ dbversity : /YCSB-master/logs ] cat /proc/cpuinfo | grep processor

processor : 0

processor : 1

[ root @ dbversity : /YCSB-master/logs ]

[ root @ dbversity : /YCSB-master/logs ]

**OracleDB VM Details:**

[ root @ dbversity : ~ ] free -g

 total used free shared buffers cached

Mem: 7 1 5 0 0 0

-/+ buffers/cache: 1 6

Swap: 11 0 11

[ root @ dbversity : ~ ]

[ root @ dbversity : ~ ]

[ root @ dbversity : ~ ]

[ root @ dbversity : ~ ] df -h

Filesystem Size Used Avail Use% Mounted on

/dev/mapper/rootvg-lv\_root

 97G 28G 64G 31% /

tmpfs 3.9G 4.0K 3.9G 1% /dev/shm

/dev/sda1 194M 62M 122M 34% /boot

swdclinx1:/xenv 2.5T 1.4T 960G 60% /xenv

[ root @ dbversity : ~ ]

[ root @ dbversity : ~ ]

[ root @ dbversity : ~ ] cat /etc/redhat-release

.bash\_history .pki/

.bash\_logout rpms

.bash\_profile .ssh2/

.bashrc ssh-copy-id.sh

.cshrc .tcshrc

log test.sh

mongodb-linux-i686-3.0.4.tgz .tpm/

.mysql\_history yum.repos.old/

.mysql\_secret

[ root @ dbversity : ~ ] cat /etc/redhat-release

Red Hat Enterprise Linux Server release 6.6 (Santiago)

[ root @ dbversity : ~ ]

[ root @ dbversity : ~ ]

[ root @ dbversity : ~ ] top -c

top - 09:00:31 up 18:37, 2 users, load average: 0.00, 0.00, 0.00

Tasks: 176 total, 1 running, 175 sleeping, 0 stopped, 0 zombie

Cpu0 : 0.8%us, 0.3%sy, 0.0%ni, 98.9%id, 0.0%wa, 0.0%hi, 0.0%si

Cpu1 : 0.8%us, 0.4%sy, 0.0%ni, 98.8%id, 0.1%wa, 0.0%hi, 0.0%si

Cpu2 : 0.8%us, 0.3%sy, 0.0%ni, 98.8%id, 0.0%wa, 0.0%hi, 0.0%si

Cpu3 : 0.7%us, 0.3%sy, 0.0%ni, 98.9%id, 0.0%wa, 0.0%hi, 0.0%si

Mem: 8061328k total, 1928076k used, 6133252k free, 132132k buff

Swap: 12582904k total, 0k used, 12582904k free, 603104k cach

[ root @ dbversity : ~ ] cat /proc/cpuinfo | head

processor : 0

vendor\_id : GenuineIntel

cpu family : 6

model : 47

model name : Intel(R) Xeon(R) CPU E7- 8870 @ 2.40GHz

stepping : 2

microcode : 55

cpu MHz : 2394.030

cache size : 30720 KB

physical id : 0

[ root @ dbversity : ~ ]

[ root @ dbversity : ~ ]

[ root @ dbversity : ~ ] cat /proc/cpuinfo | grep processor

processor : 0

processor : 1

processor : 2

processor : 3

[ root @ dbversity : ~ ]

**PostgreSQL**

**OracleDB**

**PostgreSQL Runtime:**

**OracleDB Runtime:**

**Executive Summary:-**

**Performance:**

Observed very good performance with PostgreSQL when comparing with OracleDB of same set-up in the RDS (Relational Database Service)

**Functionality highlights:**

**Drawbacks:**

No Active Cluster