



Elephant on Solaris

Zdeněk Kotala

Solaris Sustaining Engineer

Sun Microsystems



Agenda

- About PostgreSQL
- Integration with OpenSolaris
 - > Packages
 - > SMF
 - > RBAC
- Tuning
- DTrace
- Sun cluster

PostgreSQL

- One of the most advanced opensource SQL databases
- Fully ACID compliant
- Full support for foreign keys, joins, views, triggers, and stored procedures (in multiple languages)
- Support most of SQL92, SQL99
- www.postgresql.org

PostgreSQL history

- Postgres started in 1986
 - > Split of Ingres
 - > Project led by Michael Stonebracker
 - > Ended in 1993
- Postgres95
 - > Released in 1994 by Andrew Yu and Jolly Chen
 - > New SQL parser
- PostgreSQL
 - > All sources released on internet in 1996

PostgreSQL on Solaris

- Announced on December 2005
- Released as part of Solaris 10 6/06
- DTrace support in 8.2
- Niagara T2000 donated to community

Solaris packages

- PostgreSQL 8.1 from S10 6/06
 - > + located in standard path
 - > - does not support multiversion installation
 - > - has misplaced files
 - > - patch are growing in size
- PostgreSQL 8.2 from S10 6/07
 - > + supports multiversion
 - > + SMF, RBAC
 - > + considerate the patch size
 - > - located in non standard path

Service Management Facilities - SMF

- Replacement of inet.d scripts
- Two instances for PostgreSQL
 - > `svc:/application/database/postgresql:version_81`
 - > `svc:/application/database/postgresql:version_82`

RBAC

- Colorized view of unix security model
- Similar to sudo, has finer granularity
- New profile - Postgres Administration
 - > Allows user to commands like initdb, pg_ctl with postgres uid
- New role - postgres
 - > Role is assumed by “su – postgres” command
 - > Necessary for data dir access (e.g. postgresql.conf)

Compiling

- Use SunStudio 11 (especially on SPARC)
- 64-bit binary
 - > Access to more than 4GB memory
 - > Increased performance on AMD64
 - > Slower on SPARC
- DO NOT USE -fast !

PostgreSQL configuration

- `wal_sync_method = opendata_sync / fsync`
- `full_page_writes = off`

Solaris configuration

- IPC SysV resources are enough by default since Solaris 10
- Use forcedirectio on disk with WAL files
- Mount filesystem with noatime option
- Check if write cache is enable
 - > format -e
 - > set ata:ata_write_cache = 1

UFS vs. ZFS

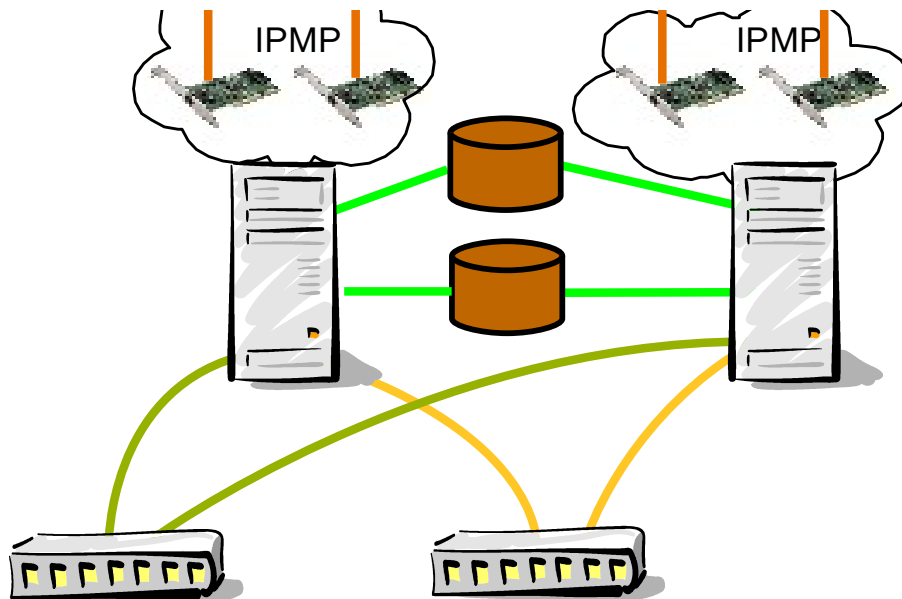
- UFS
 - > Metadata logging
 - > `opendata_sync`
 - > OLTP
- ZFS
 - > Copy on write
 - > `fsync`
 - > Data Warehouse

DTrace

- Dynamic trace framework
- 42 = DTrace
- Eleven probes in PostgreSQL 8.2
- DTrace scripts on www.pgfoundry.org
- Useful for application bottleneck analysis

Sun Cluster

- PostgreSQL dataservice since version 3.2
- Only as a failover service





Elephant on Solaris

Zdeněk Kotala

zdenek.kotala@sun.com