



PostgreSQL – Neues und Besonderes der führenden Open-Source-Datenbank

Harald Armin Massa

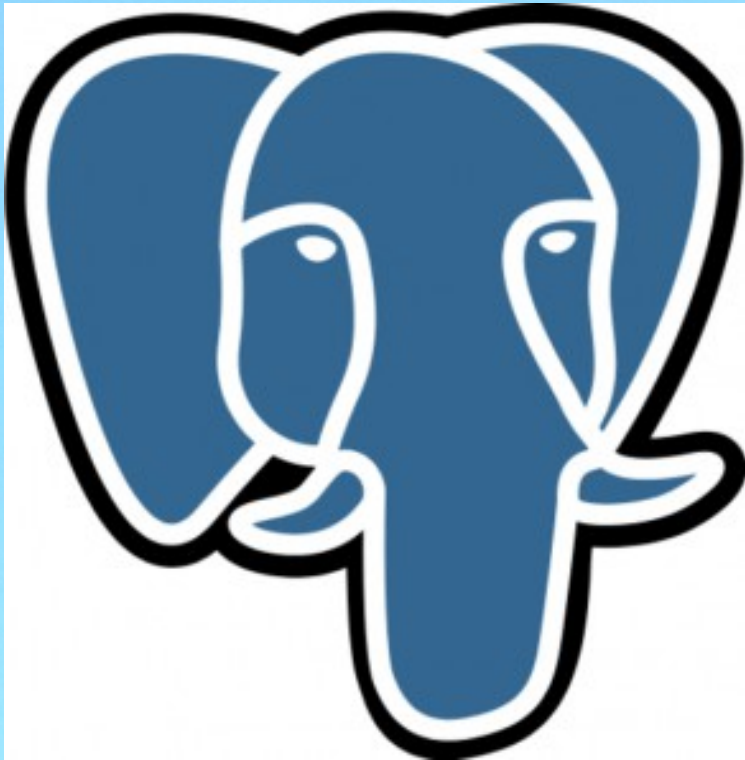


Linux Tag Augsburg, 23.März 2013





PostgreSQL





Wer ist das? Warum macht der Werbung für PostgreSQL? Was versucht er zu verkaufen?



2ndQuadrant

Professional PostgreSQL

- 24/7 Support
- Remote DBA
- Hochverfügbarkeit, Hochperformance
- Backup & Disaster Recovery
- Training
- Weiterentwicklung PostgreSQL im Kundenauftrag

- Personalwerbung





Harald Armin Massa
2ndQuadrant

erste Datenbankerfahrungen ca. 1984

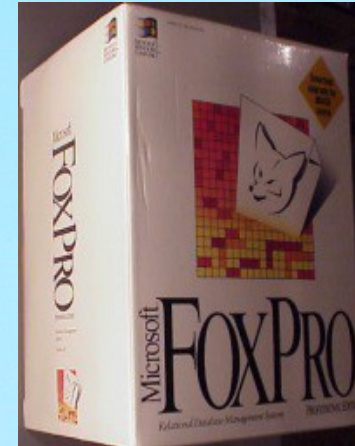
NoSQL Datenbanken auf C64 / VC1541





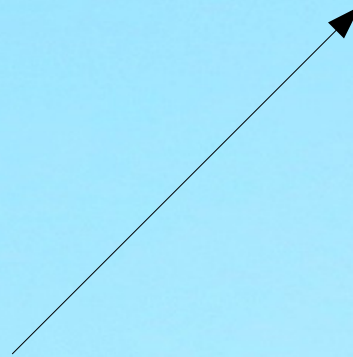
IBM SNA
DOS/VSE
MVS/ESA

ABAP/4
COBOL
CICS





ca. 1997





warum – 1



2002-12-11 Bruce Momjian

Reorder psql \? help into groupings, idea from Harald...
...help into groupings, idea from Harald Armin Massa.

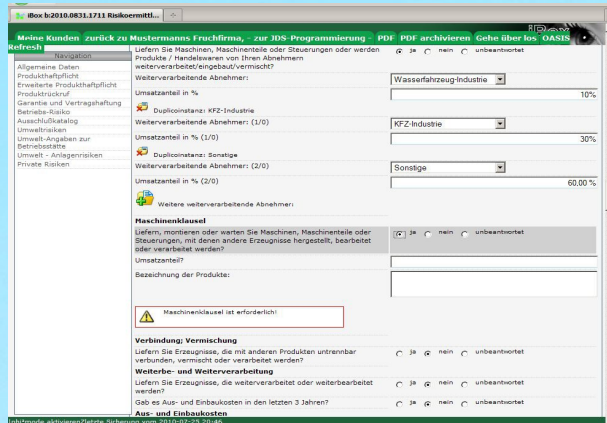
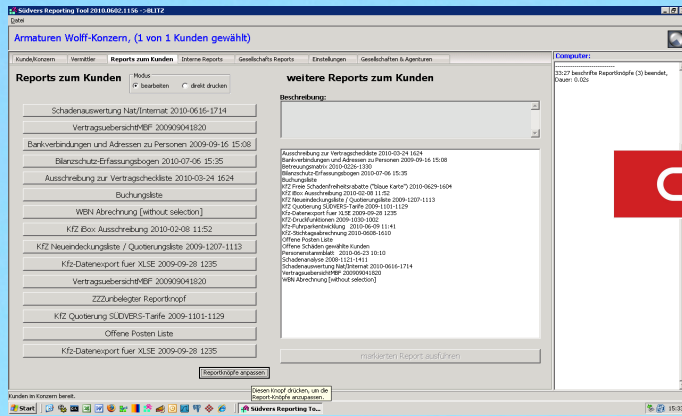
[commit](#) | [commitdiff](#) | [tree](#)

```
ws2ndq=> \?
General
  \copyright          show PostgreSQL usage and distribution terms
  \g [FILE] or ;     execute query (and send results to file or |pipe)
  \h [NAME]          help on syntax of SQL commands, * for all commands
  \q                 quit psql

Query Buffer
  \e [FILE]          edit the query buffer (or file) with external editor
  \ef [FUNCNAME]    edit function definition with external editor
  \p                 show the contents of the query buffer
  \r                 reset (clear) the query buffer
  \s [FILE]          display history or save it to file
  \w FILE            write query buffer to file

Input/Output
  \copy ...          perform SQL COPY with data stream to the client host
  \echo [STRING]    write string to standard output
  \i FILE            execute commands from file
  \o [FILE]          send all query results to file or |pipe
  \qecho [STRING]   write string to query output stream (see \o)

Informational
  (options: S = show system objects, + = additional detail)
  \d[S+]             list tables, views, and sequences
  \d[S+] NAME        describe table, view, sequence, or index
```

EUROPYPYTHON

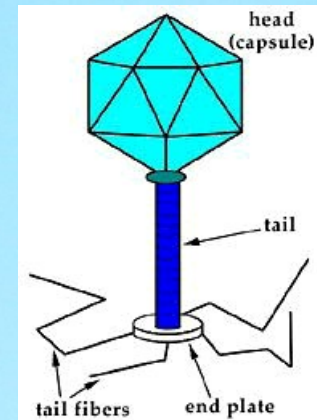
member of
Python Software Foundation



wer



Sternberg Astronomical Institute
Catalog Access Services





ibox b:2010.0831.1711 Risikoermittl...

Meine Kunden zurück zu Mustermanns Fruchfirma, - zur JDS-Programmierung - PDF PDF archivieren Gehe über los OASIS

Refresh

Navigation	Liefere Sie Maschinen, Maschinenteile oder Steuerungen oder werden Produkte / Handelswaren von Ihren Abnehmern weiterverarbeitet/eingebaut/vermischt?	<input checked="" type="radio"/> ja <input type="radio"/> nein <input type="radio"/> unbeantwortet
Allgemeine Daten	Weiterverarbeitende Abnehmer:	Wasserfahrzeug-Industrie
Produkthaftpflicht	Umsatzanteil in %	10%
Erweiterte Produkthaftpflicht	Duplicoinstanz: KFZ-Industrie	
Produktrückruf	Weiterverarbeitende Abnehmer: (1/0)	KFZ-Industrie
Garantie und Vertragshaftung	Umsatzanteil in % (1/0)	30%
Betriebs-Risiko	Duplicoinstanz: Sonstige	
Ausschlusskatalog	Weiterverarbeitende Abnehmer: (2/0)	Sonstige
Umweltrisiken	Umsatzanteil in % (2/0)	60,00 %
Umwelt-Angaben zur Betriebsstätte	Weitere weiterverarbeitende Abnehmer:	
Umwelt - Anlagenrisiken		
Private Risiken		

Maschinenklausel

Liefere, montieren oder warten Sie Maschinen, Maschinenteile oder Steuerungen, mit denen andere Erzeugnisse hergestellt, bearbeitet oder verarbeitet werden?

ja nein unbeantwortet

Umsatzanteil?

Bezeichnung der Produkte:

Maschinenklausel ist erforderlich!

Verbindung; Vermischung

Liefere Sie Erzeugnisse, die mit anderen Produkten untrennbar verbunden, vermischt oder verarbeitet werden?

ja nein unbeantwortet

Weiterbe- und Weiterverarbeitung

Liefere Sie Erzeugnisse, die weiterverarbeitet oder weiterbearbeitet werden?

ja nein unbeantwortet

Gab es Aus- und Einbaukosten in den letzten 3 Jahren?

ja nein unbeantwortet

Aus- und Einbaukosten

phfmode aktivierenZuletzt Sicherheit vom 2010-07-25 20:46

PostgreSQL

erste version PostgreSQL on Windows

started development with PostgreSQL 7.4 binaries from an anonymous FTP server

pilot rollout: 8.0beta1



```
*** STOP: 0x0000000A (0xFFBDF004, 0x00000010, 0x00000000, 0x800079D3)
IRQL_NOT_LESS_OR_EQUAL*** Address 800079d3 has base at 80007000 - SCSIPTORT.SYS

CPUID:GenuineIntel 5.2.4 irq:1:f DPC SYSVER 0xf0000565

Dll Base DateStmp - Name Dll Base DateStmp - Name
80100000 36224cda - ntoskrnl.exe 80010000 35e72341 - hal.dll
80001000 35ca19d3 - symc810.sys 80007000 35e5c313 - SCSIPTORT.SYS
802fa000 353e319e - Disk.sys 8038c000 36269e3f - CLASS2.SYS
80390000 36238303 - Mfs.sys f8ea8000 31ec6c8d - Floppy.SYS
f8eb8000 353e319c - Cdrom.SYS f9214000 00000000 - Null.SYS
f9074000 35eb144b - KSecDD.SYS f9215000 00000000 - Beep.SYS
f8ee8000 353e3184 - i8042prt.sys f907c000 353e318a - mouclass.sys
f9084000 31ec6c94 - kbdc1ass.sys f8f00000 35648e19 - VIDEOPRT.SYS
f8c10000 353e3155 - S3.SYS f9094000 360ea154 - vga.sys
f8f30000 353e31df - Msfs.SYS f8c30000 353e31d5 - Npfs.SYS
fe4bb000 362043ba - NDIS.SYS f909c000 35fe17b4 - ndistapi.sys
a0000000 36248f4f - win32k.sys fe482000 35d9fd5f - S3.dll
f8c70000 353e3626 - Cdfs.sys fe437000 35dde3d6 - Fastfat.SYS
f9132000 35fe29fd - rasacd.sys fe4a7000 31ec6e6c - TDI.SYS
fe3ea000 36243c12 - Tcpip.sys fe3cc000 36129a8d - netbt.sys
fe49f000 35e5e7f3 - rasarp.sys f8cf0000 3548ba13 - asynmac.sys
fe493000 31ec6e15 - elnk3.sys f8d00000 35fe1816 - ndiswan.sys
fe3bb000 3610249c - afd.sys f8fbf000 353e35d4 - netbios.sys
fe47e000 31ec6c9b - Parport.SYS fe476000 353e318f - Parallel.SYS
f913c000 31ec6c9d - ParVdm.SYS f8d40000 35ef29c4 - Serial.SYS
fe352000 35f03aa8 - rdr.sys fe2ef000 35b7f615 - srv.sys
fe2de000 353e362c - mup.sys

Address dword dump Build [1381] - Name
801499f8 800079d3 800079d3 805f58c8 8047b000 805f5a06 805f1400 - SCSIPTORT.SYS
80149a0c 80149a3c 80140a3c 800079bf 805f58c8 8047b000 805f5a06 - ntoskrnl.exe
80149a10 800079bf 800079bf 805f58c8 8047b000 805f5a06 805f1400 - SCSIPTORT.SYS
80149a5c 80149a84 80149a84 00000000 800079d3 00000008 00010296 - ntoskrnl.exe
80149a64 800079d3 800079d3 00000008 00010296 00000000 805f1530 - SCSIPTORT.SYS
80149a80 80149ab8 80149ab8 80149ac0 800039b6 805f5a48 fe26dd3c - ntoskrnl.exe
80149a84 80149ac0 80149ac0 800039b6 805f5a48 fe26dd3c 805732c8 - ntoskrnl.exe
80149a88 800039b6 800039b6 805f5a48 fe26dd3c 805732c8 80149ab8 - symc810.sys
80149a98 80149ab8 80149ab8 805f5a48 fe26dd3c 805f14f0 fe26dd3c - ntoskrnl.exe
80149ac0 80149af8 80149af8 8000442b 805f5a48 805f14f0 fe26dd3c - ntoskrnl.exe
80149a54 8000442b 8000442b 805f5a48 805f14c0 805f14c0 fe26dd3c - symc810.sys
80149af8 8000442b 8000442b fe26dd3c 805f5a48 fe26dd3c 805e8d08 - symc810.sys
80149afc 800023fd 800023fd fe26dd3c 805f5a48 fe26dd3c 805e8d08 - symc810.sys
80149b14 800084eb 800084eb 805f5a48 fe26dd3c 80149b44 fe26dd4c - SCSIPTORT.SYS

Beginning dump of physical memory
Restart and set the recovery options in the system control panel
or the /CRASHDEBUG system start option. If this message reappears,
contact your system administrator or technical support group.
```



warum - 2



<http://www.crowdrise.com>





Professor Michael Stonebraker

POSTGRES – 1986

Successor to the INGRES relational database system

(quoting & rephrasing „The design of POSTGRES“, initial paper by M. Stonebraker):

- user extensibility for data types, operators and access methods
- facilities for active databases (i.e., alerters and triggers)
- simplify the DBMS code for crash recovery
- take advantage of [...]multiple tightly-coupled processors



theoretical limits as of now

Limit	Value
Maximum Database Size	unlimited
Maximum Table Size	32 TB
Maximum Row Size	1.6 TB
Maximum Field Size	1GB
Maximum Rows per Table	unlimited
Maximum Columns per table	250-1600 (depending on column types)
Maximum Indexes per Table	unlimited



Database

Feature

	A	B	C	D
1	X	X	X	X
2	X	X	X	X
3	X	X	X	X
4			X	
5	X	X		X
6	X		X	



license

PostgreSQL is released under the PostgreSQL License, 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) 1996-2013, 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.



license

Portions Copyright (c) 1996-2013, The
**PostgreSQL Global Development
Group**

un-buy-able. No danger of anyone
taking over and changing license terms.

Portions Copyright (c) 1994, The Regents
of the University of California



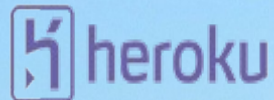
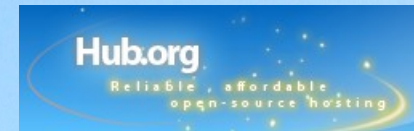
license

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, [...]





some of the sponsors





Konzepte & Konsequenzen

- Meritokratie für commit-Rechte
 - Die, die sich durch Patches & Beiträge als würdig erweisen
 - NICHT erreichbar durch Position in Firma





Konzepte & Konsequenzen

- Funktionen des Betriebssystems nicht erneut programmieren
 - Betriebssystem für Plattenzugriff
 - Betriebssystem-Cache nutzen
 - physical backup via filecopy
 - no „certified for PostgreSQL“ needed





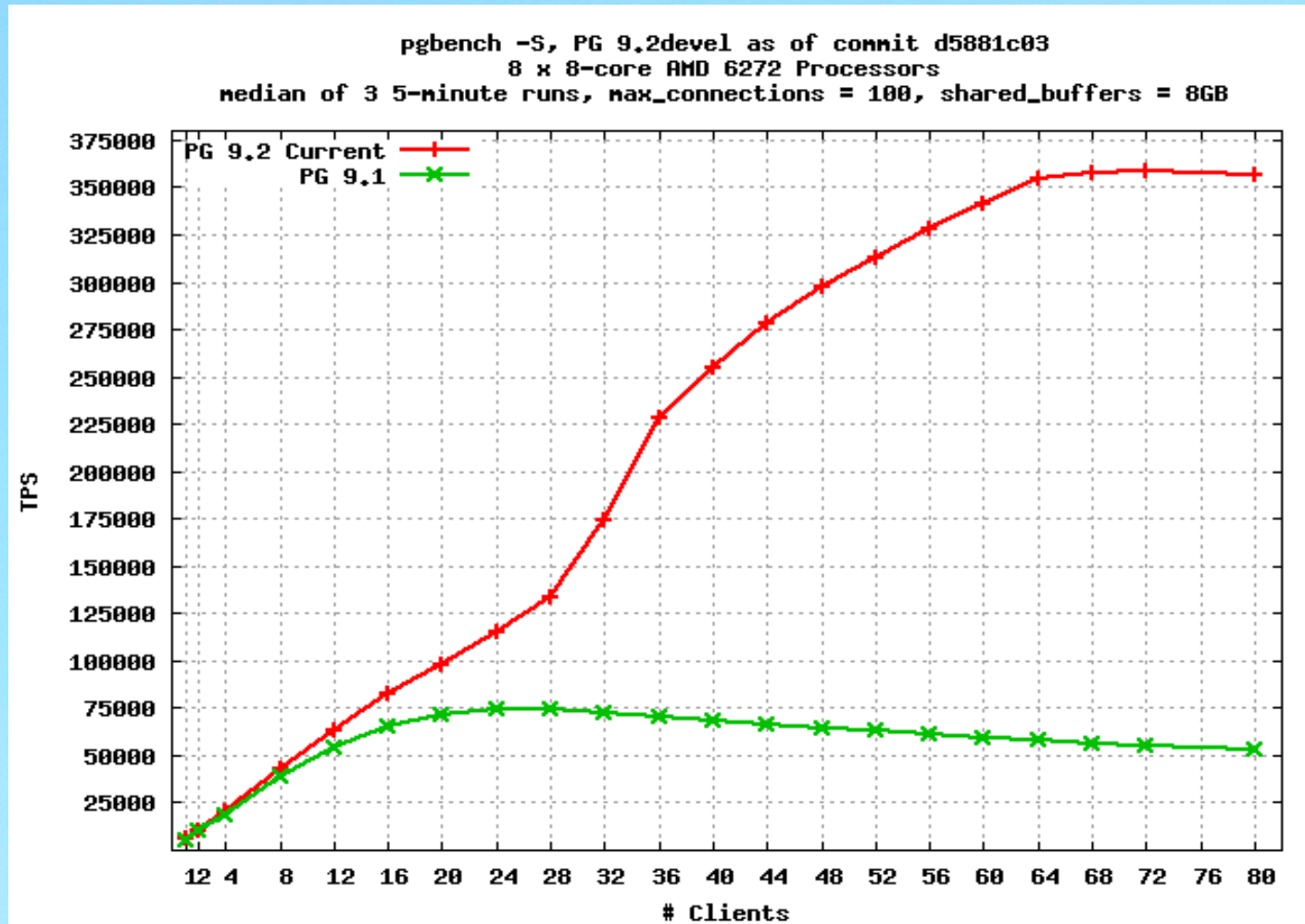
Konzepte und Konsequenzen

- Knapper Zugriff = DEFAULT
- Von Anfang an für mehrere Prozessoren designed
 - Preis der interprocess communication
 - Neue Prozesse teuer auf älteren Betriebssystemen
 - Kosten auch bei nur einem Prozessor...
 - JETZT mit Multicore in allen Maschinen ...





Konzepte und Konsequenzen





Konzepte und Konsequenzen

- open engeneering
 - Offene Diskussionen
 - publicly archived & docmented discussions
 - public flame wars





„user driven development“

exhibit 1: the psql utility

- interactive shell
- tab completion
- line editing / history
- available since the beginning



development by database users

exhibit 2: datatype text

- up to 1 Gigabyte per field
- better performance than char, varchar
- one solution for every length
- no need for „are 80 chars wide enough for x“





development by database users

exhibit 3: COPY FROM / TO <csv table>

- included in psql command line tool & server
- no extra utility for import/export

exhibit 4: pg_dump / pg_restore

backup running system into 1 file

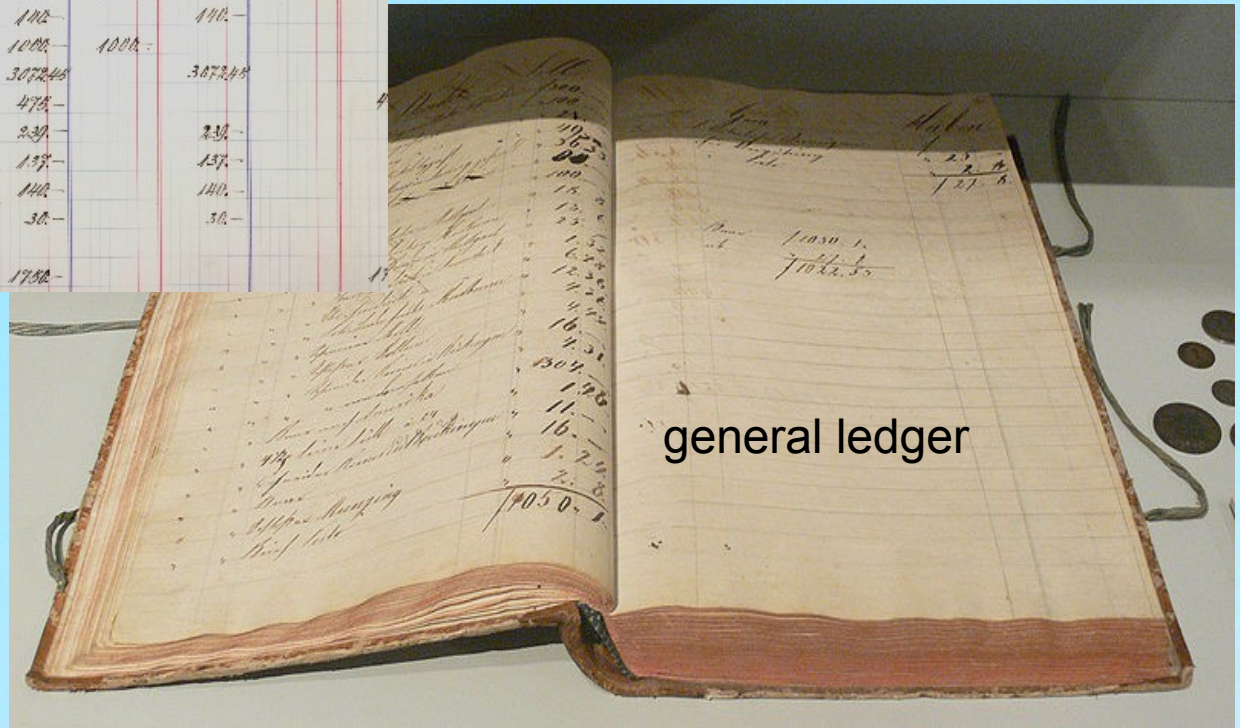




Mehrfachverwendung von Konzepten & Code

Seite	Tag	Name, Ort und Gegenstand	BETRAG	Kasse		Kassa	
				Soll	Haben	Soll	Haben
		Übertrag					
	8.	1 amerikan. Journal mit Abdruck (Amerikaner)	208		208		
	7.	Lohnzahlung an Arbeiter	140		140		
433	7.	K. Grotzsch, Leubach, gelbe Bau	1368	1368			
"	7.	do. Kaufpreis 4 R. 4 1/2 1/2	140			140	
"	8.	Karl Finsel, Seibitzau für 1 gelb. Bau	100				
"	9.	do. gelbe Bau	1000	1000			
	14.	Lohnzahlung an Arbeiter	140		140		
"	3.	13. Reichs-Schlichter, Krieger gelbe Bau	1000	1000			
"	10.	do. Kaufpreis 4 R. 4 1/2 1/2	3072 1/2		3072 1/2		
"	10.	do. Kaufpreis 4 R. 4 1/2 1/2	475				
"	12.	Ordn. Ranzel, Krieger	230		230		
	21.	für gelbe Bau	137		137		
"		Lohnzahlung an Arbeiter	140		140		
"		Einzahlung an Bauarbeiter f. Lohn 7 1/2 - 24/24	30		30		
"	12.	Levy, Krieger, Leipzig - Krieger für gelbe Bau	1750				

ca. 1340, Genua



general ledger

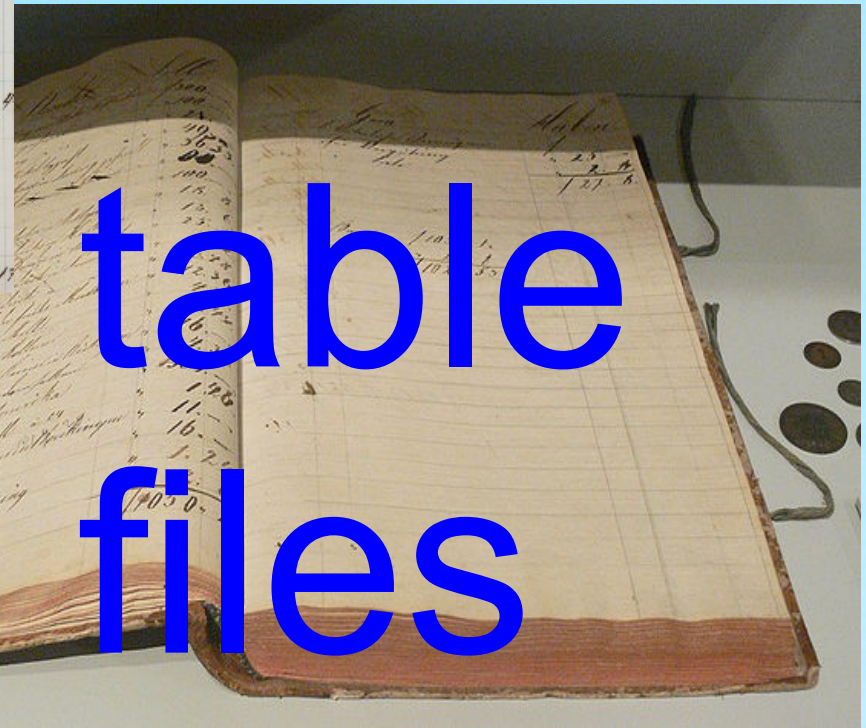




Mehrfachverwendung von Konzepten & Code

Seite	Tag	Name, Ort und Gegenstand	BETRAG	Kasse		Kassa	
				Soll	Haben	Soll	Hab
		Übertrag					
8.		1 amerikan. Journal mit Abdruck (Amerikaner)	208		208		
7.		1. Abdruck von Bremer P.	140		140		
4. 11.		K. Schacht & Co. Hamburg	1368	1368			
" 7.		" " " " " " " " " " " "	140			140	
" 5.	14.	Karl Schacht & Co. Hamburg	100				
" 9.		" " " " " " " " " " " "	1000	1000			
14.		Erkennung von Aktien Brief	140		140		
" 3.	18.	Franken-Schlachten, Frankfurt	1000	1000			
" 10.		" " " " " " " " " " " "	307246	307246			
" 10.		" " " " " " " " " " " "	275				
" 13.	20.	Erkennung von Aktien Brief	30		30		
21.		für die Kupfer Wappenstein	137		137		
"		Erkennung von Aktien Brief	140		140		
"		Erkennung von Aktien Brief	30		30		
" 11.	22.	Erkennung von Aktien Brief	1750				

WAL-
records
table
files





Mehrfachverwendung von Konzepten & Code

WAL-record

WAL-record

WAL-record

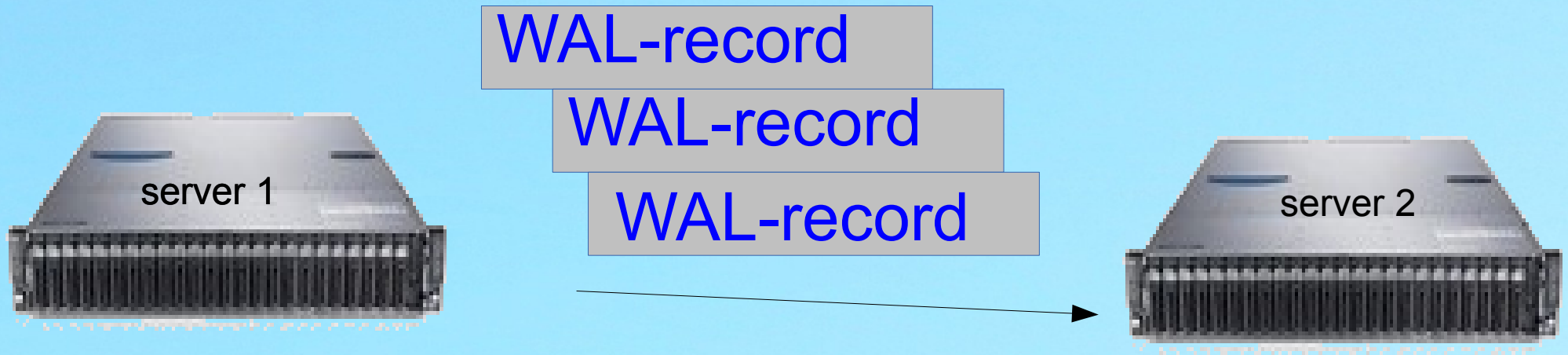


online physical backup

Point In Time Recovery



Mehrfachverwendung von Konzepten & Code



(streaming) replication





server-side-programming: Procedural Languages

- a wealth of languages is available
 - PL/Python
 - PL/Java
 - PL/Perl
 - PL/Tcl
 - PL/R
 - PL/Mono
 - PL/V8
 - PL/scheme

```
CREATE OR REPLACE FUNCTION pyoav(jsondata
text, feldname text)
  RETURNS text AS
$BODY$
  import json
  jsonparsed=json.loads(jsondata)

  return jsonparsed.get(feldname,'<nicht vorhanden>')
$BODY$
LANGUAGE plpythonu VOLATILE
COST 100;
```



server-side-programming: Procedural Languages

- Neue Prozedurale Sprachen evtl. zu einfach anzulegen:

```
CREATE FUNCTION LOL_MAIN_TEST(TEXT)
RETURNS BOOLEAN
LANGUAGE PLLOLCODE
AS $$
HAI
    VISIBLE INFO LOL1
    FOUND YR WIN
KTHXBYE
$$;
```

```
SELECT LOL_MAIN_TEST('IM IN YR DATABUKKIT');
```

```
INFO: IM IN YR DATABUKKIT
```

```
lol_main_test
```

```
-----
```

```
t
```

```
(1 row)
```



server-side-programming

- Nutzer kann Funktionen entwickeln, die im Server laufen
- call nach PostgreSQL via Server Programming Interface
- advantage open source
 - detaillierte documentation
 - debugging in den servercode



Extending: new data types

- PostgreSQL operation catalog-driven
- including information about data types
- new data types = new catalog entries
- including operations on them





Extending: data type examples

- PostBIS - Bioinformatics Booster for PostgreSQL
 - DNA_SEQUENCE
 - RNA_SEQUENCE
 - AA_SEQUENCE
 - ALIGNED_DNA_SEQUENCE
 - ALIGNED_RNA_SEQUENCE
 - ALIGNED_AA_SEQUENCE





Extending: data type examples

- hstore
 - stores key => value pairs
 - foo => bar, baz => whatever
- range types
 - part of PostgreSQL 9.2
 - range of some element type
 - full index support for range-overlapping queries
- json (new in 9.2)



json

```
CREATE TABLE tuwienosql  
(  
  id_t serial NOT NULL,  
  dokument json,  
  CONSTRAINT tuwienosql_pkey PRIMARY KEY (id_t)  
);
```

```
INSERT INTO tuwienosql(  
  dokument)  
VALUES ('{"tiername":"Goofy", "ohren":"lang"}');
```

[...]

```
CREATE FUNCTION xtraktname (mydokument json) returns text [...];  
CREATE INDEX tuwieno_idx ON tuwienosql(xtraktname (dokument));
```

[...]

```
ALTER TABLE tuwienosql ADD COLUMN tiername text;
```




Extending: entirely new index types

quoting PostgreSQL documentation, Chapter 52:

„Index Access Method Interface Definition“

...The core system knows nothing about indexes beyond what is specified here, so it is possible to develop entirely new index types by writing add-on code...





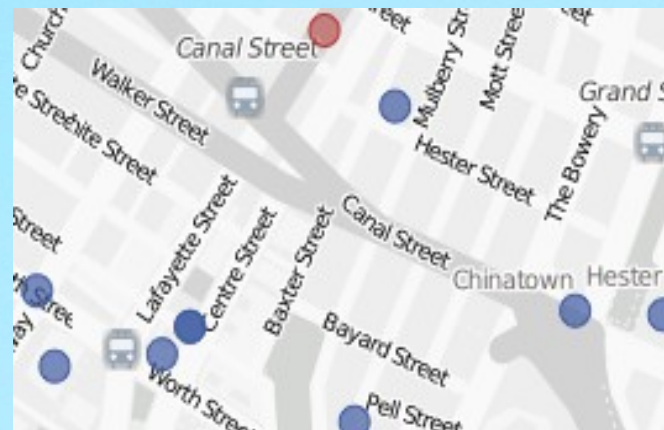
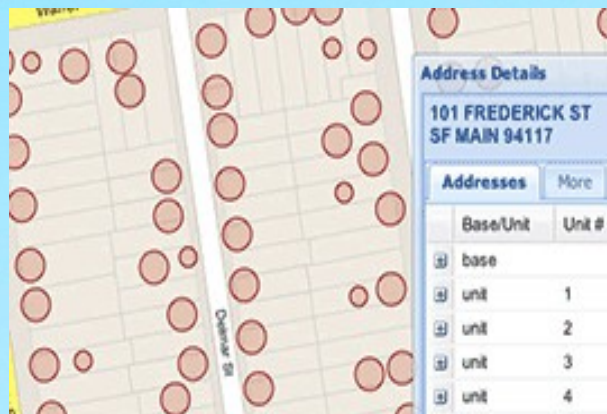
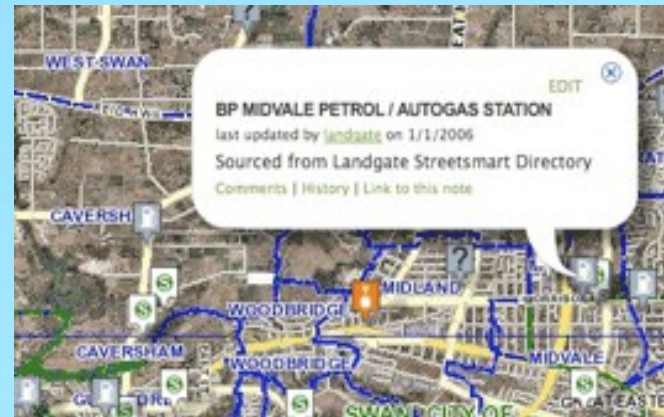
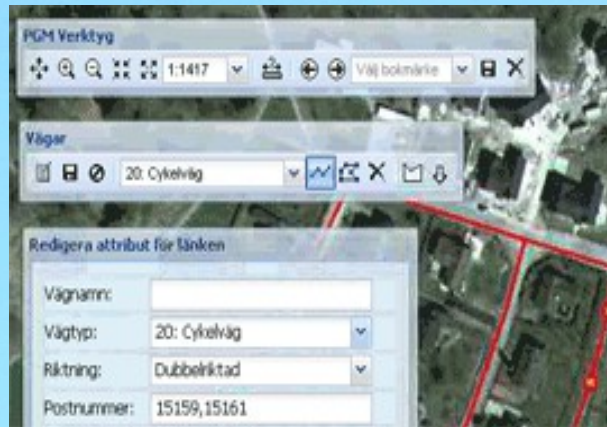
Advanced Index Types

- GIST
 - Generalized Search Tree
 - may be lossy
 - used for example to speed up full text search
 - used for k-nearest-neighbour indexing (return sorted rows if „distance“ has a meaning for the data type)
- GIN
 - Generalized INverted Index
 - used to speed up full text search





PostGIS





extending the extensibility

version.myparticleclass
myparticleclass.so
myparticleclass.sql
myparticleclass.plpgsql
myparticleclass.howtoinstall





extending the extensibility

```
CREATE EXTENSION <extension>;
```

PostgreSQL 9.1



Dimitri Fontaine,
2ndQuadrant France



Elephants at Amboseli national park against Mount Kilimanjaro
Date: 11 August 2012
Author: Amoghavarsha amoghavarsha.com



Danke.

Zeit für Fragen.

Harald Armin Massa
2ndQuadrant