



Architettura di storage

Architettura di storage

- Cluster
- Databases
- Tabelle
- Indici
- Blocchi

Cluster

- Un database cluster è un insieme di basi di dati che sono gestiti da una singola istanza del processo server
- Creare un cluster database consiste in:
 - Creare le directory in cui il database depositerà i dati
 - Creare le tabelle comuni di catalogo
 - Creare i template dei db
- Un processo postmaster per ogni cluster
- Un cluster può essere interrogato attraverso una sola directory data che può avere più tablespaces

Cluster Layout - overview

- Tutte le informazioni riguardante il cluster sono nella directory data
 - Global → system tables e control files
 - Base → database tables e temp tables
 - pg_tblspc → links tablespace
 - control/pg_contro → checkpoint master file
 - pg_xlog → transactiona recovery log
 - pg_clog, pg_subtrans → commit status
 - pg_multixact → row locking
 - pg_twophase → transazioni nello stato prepared

Cluster Layout - overview

- Conf
 - postgresql.conf → file di configurazione
- Logs
 - Serverlog → log
- Files di stato
 - postmaster.pid
 - PG_VERSION

Struttura delle directory

- PostgreSQL\8.x\
 - Bin
 - **Data → dove vengono memorizzati i dati**
 - Doc
 - Include
 - Lib
 - Man
 - Share

Organizzazione dei dati

- Ogni database viene memorizzato in una directory
 - Con uno o più files per ogni relazione
 - Ogni file se > 1 Gb viene splittato
- Ogni relazione è memorizzata in un solo tablespace
 - Si possono gestire link ad altre directory nel filesystem
 - Per defaults gli indici vengono memorizzati nello stesso tablespace

OID - database

- OID: Object identifier : è l'identificatore di un oggetto (database, tabella...)
 - Es : testdb=# SELECT datname, oid from pg_database

where datname='testdb';

datname		oid
testdb		16384

- Al database testdb è associato l'oid 16384

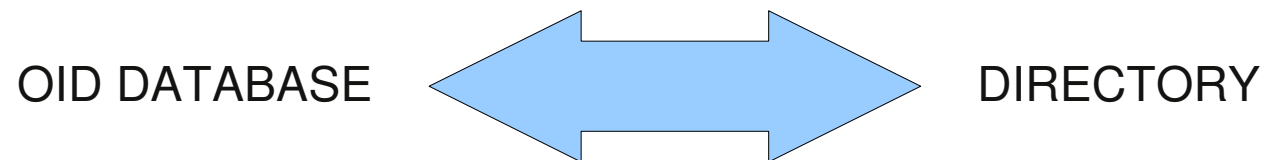
OID - database

- Andiamo a vedere ora all'interno della cartella data/base

- `ls -l | grep 16384`

```
drwx----- 2 postgres postgres 4096 2008-09-10 09:58 16384
```

- Troviamo una directory dal nome 16384



OID - Tables

- testdb=# create table testtb (id serial not null primary key, nome char(20));

NOTICE: CREATE TABLE will create implicit sequence "testtb_id_seq" for serial column "testtb.id"

NOTICE: CREATE TABLE / PRIMARY KEY will create implicit index "testtb_pkey" for table "testtb"

CREATE TABLE

OID - Tables

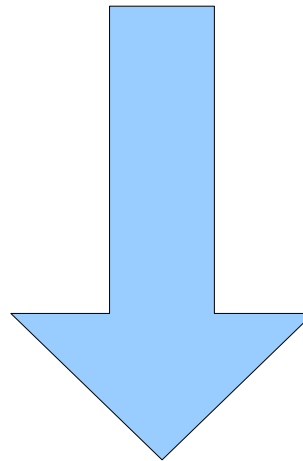
- testdb=# SELECT relname,oid from pg_class where relname='testtb';

```
relname | oid
-----+-----
testtb  | 16387
(1 row)
```

OID - Tables

- `data/base/16384# ls -l | grep 16387`

```
-rw----- 1 postgres postgres    0 2008-09-10 11:05 16387
```



**Una tabella diventa un file il cui nome
è l'object identifier**

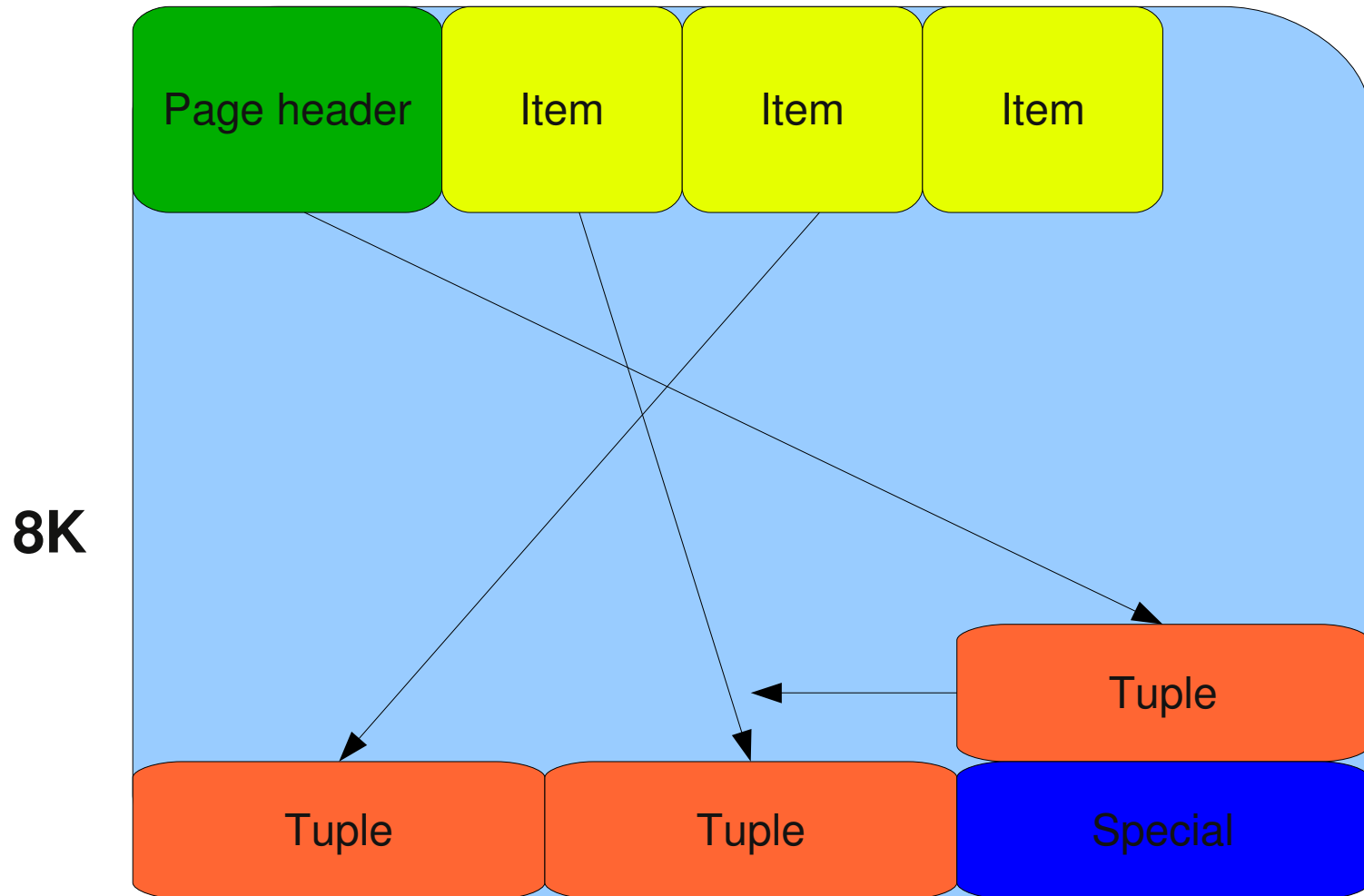
OID - Tables

- Le tabelle e gli indici vengono memorizzati in files separati
 - Il nome del file è uguale all'oid dell'oggetto
- Le tabelle o gli indici (Relazioni) che sono più grandi di 1 Gb vengono divisi in segmenti di dimensione inferiore al Gb
- Il primo segmento di chiama con il nome del file uguale all'oid agli altri si aggiunge 1,2 etc..

Layout di pagina

- Pagine di 8k → vengono caricate nello shared buffer
- Page header (20 bytes):
 - Informazioni generali riguardo la pagina
 - Puntatori allo spazio libero
- Puntatori alle tuple
- Spazio libero non allocato
- Row/Index Entry : tupla corrente
- Informazioni di sistema: metodo di accesso indice,

Layout di pagina



Layout di pagina

```
testdb=# \d
```

```
List of relations
```

```
Schema | Name | Type | Owner
```

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

```
public | testtb | table | postgres
```

```
public | testtb_id_seq | sequence | postgres
```

```
(2 rows)
```

```
testdb=# INSERT INTO testtb (nome) values ('Enrico');
```

```
INSERT 0 1
```

```
testdb=# SELECT * from testtb;
```

```
id | nome
```

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

```
1 | Enrico
```

Layout di pagina

```
testdb=#
```

```
SELECT relname,oid,relpages,reltuples from pg_class where relname  
  ilike 'test%';
```

relname	oid	relpages	reltuples
testtb_id_seq	16385	1	1
testtb	16387	0	0
testtb_pkey	16391	1	0 → viene creato un indice unico

(3 rows)

Architettura di storage

- Cluster
- Databases
- Tabelle
- Indici
- Blocchi

