

SQL Anti Patterns

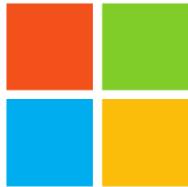
SQL, performance and bad design practices

Kennie Nybo Pontoppidan

Cambridge, September 12th 2015



Sponsors, sponsors, sponsors



Microsoft

sqlcloud

redgate

ingeniously simple

Axioworks
Practical Innovation



PYRAMID
ANALYTICS

Profisee®

ap

idera®

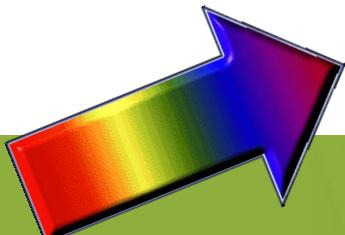
BIG BANG
DATA CO

DELL Software

SanDisk®

effektor
SELF SERVICE DATA WAREHOUSE & BI

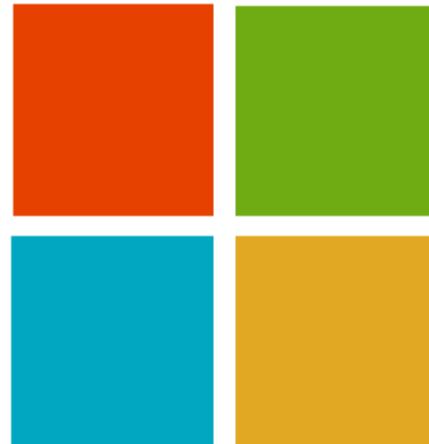
DATA
IDOLS.



PASS SQL
saturday

Stick around for SWAG!

- (Prize drawing at 5.00 p.m)
- Be there (in person) or be



Prizes

SQLCloud — XBox One!!! (make sure you know how to play the Jet Set Willy vendor competition!)

SQLCloud — Dark Side Of The Moon, The Wall and Wish You Were Here

Microsoft — MCP Voucher

Data Idols — Supercar Driving Experience

Profisee — Signed MDS Book and £50 Amazon Voucher

SQLSentry — Plan Explorer PRO License

Pyramid Analytics — Coffee Machine

Effektor — Lego Big Ben

Rencore — SPCAF Professional License (valued at \$2099!)

Axioworks — Amazon Fire TV and £75 Amazon Voucher

DELL — Beats by Dr Dre Headphones

Redgate — DLM Workshop

Pluralsight — Annual Plus subscriptions

All Your Base Conference — 3 x Conference Tickets

VisualISP — Annual training subscription

Loads of books and much much more!!!



Volunteers

- All our volunteers and organisers do not get paid for running SPS SQLSaturday Cambridge and do it because they believe in the power of community (or are just plain nuts).
- Please show your gratitude for making this possible by:
 - Giving them a hug
 - Shaking their hand
 - Saying thank you
 - Coming back next year
 - Consider getting your company to pay for a precon next year
 - Spreading the word
 - Getting involved yourself
 - But most of all, by enjoying the day!

Speakers

- Thank the speakers for donating their time, energy and expenses

Who am I?



Kennie Nybo Pontoppidan

@ knp@rehfeld.dk

 @KennieNP

 dk.linkedin.com/in/kennie/

<http://Imgify.com/?q=kennie+Pontoppidan>

$$\frac{dx}{\sqrt{x^3+x}} = \frac{dx}{\sqrt{x^3+\sqrt{x^2}}} = \left[\begin{array}{l} \sqrt{x} = E \\ x = E^2 \\ dx = 2E^2 dt \end{array} \right] = \frac{2E^2 dt}{\sqrt{E^6+E^2}} = \frac{2E^2 dt}{E\sqrt{E^4+1}} = \frac{2dt}{\sqrt{E^4+1}}$$

$$\int \left(\frac{\frac{t^5+1}{5}}{t-1} - \frac{1}{t+1} \right) dt = 6 \left(t^2 - t + 1 - \frac{1}{t+1} \right)$$

$$+ E - Cn |E| + C =$$

$$\left[\frac{x}{2} \right]^2 + \sqrt{x} \cdot (n \sqrt[6]{x^5+1}) + C$$





HOW SAFE IS YOUR COMPUTER?





The University of Queensland, St. Lucia campus in the foreground, Brisbane River and downtown Brisbane City in the background





©2007 Jim M. Goldstein, All Rights Reserved

THE BELL CURVE

JIM M. GOLDSTEIN • JMG-GALLERIES.COM





CONSULTING

IF YOU'RE NOT A PART OF THE SOLUTION,
THERE'S GOOD MONEY TO BE MADE IN PROLONGING THE PROBLEM.

EFFEKTOR

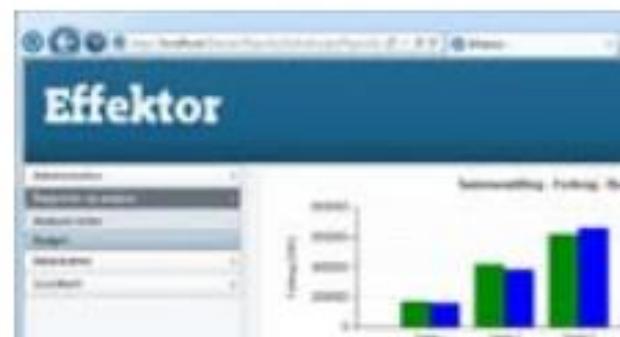
Self Service Datawarehousing



Rehfeld Effektor delivers a platform for a full BI-solution supporting all relevant needs importing source system data over data warehousing to reporting both on relational cubes. The platform is built around the concepts of ease of use: working with your data changes to structure and/or data should not require deep technical expertise.

A HEALTHY ARCHITECTURE FOR THE DATA WAREHOUSE

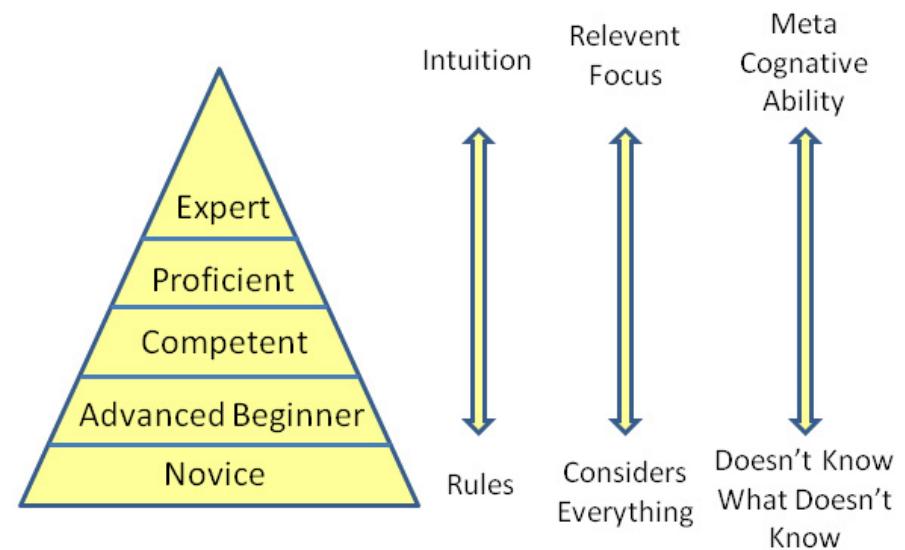
The core in Effektor is a data warehouse architecture modeled after industry standards. This ensures reuse of code and stability of implementations on Effektor. The enterprise data warehouse (EDW) in Effektor looks like how you would do it in a hand



Program

1. Anti patterns and the Dreyfuss expert model
2. Logical anti patterns
3. Physical anti patterns
4. Query anti patterns
5. Reporting anti patterns

Anti patterns and the Dreyfuss expert model



What is a (design) pattern?

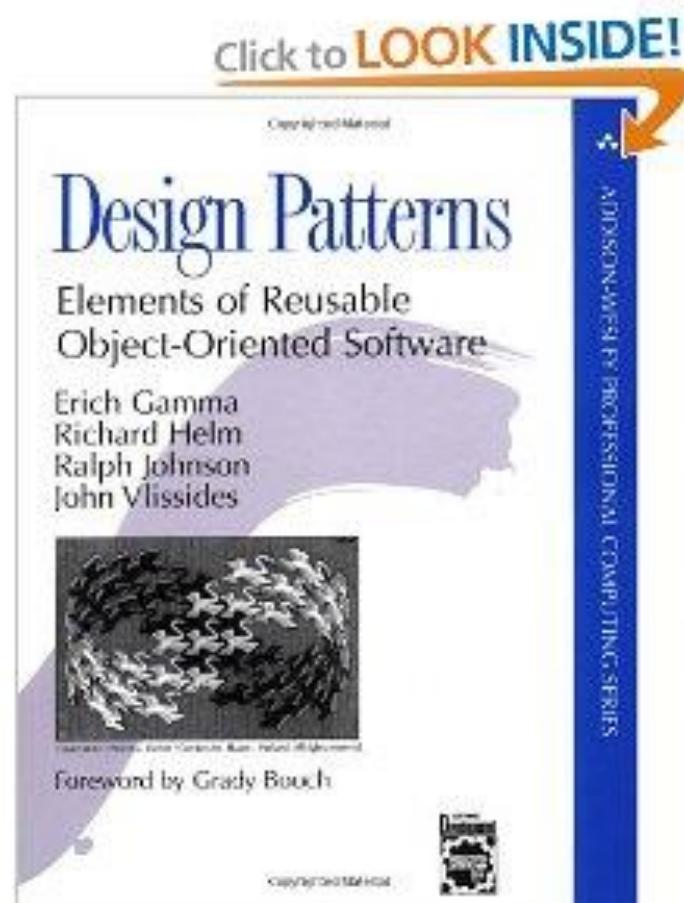
a pattern describes a problem that occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice.

- Christopher Alexander, A Pattern Language: Towns, Buildings, Construction (1977)

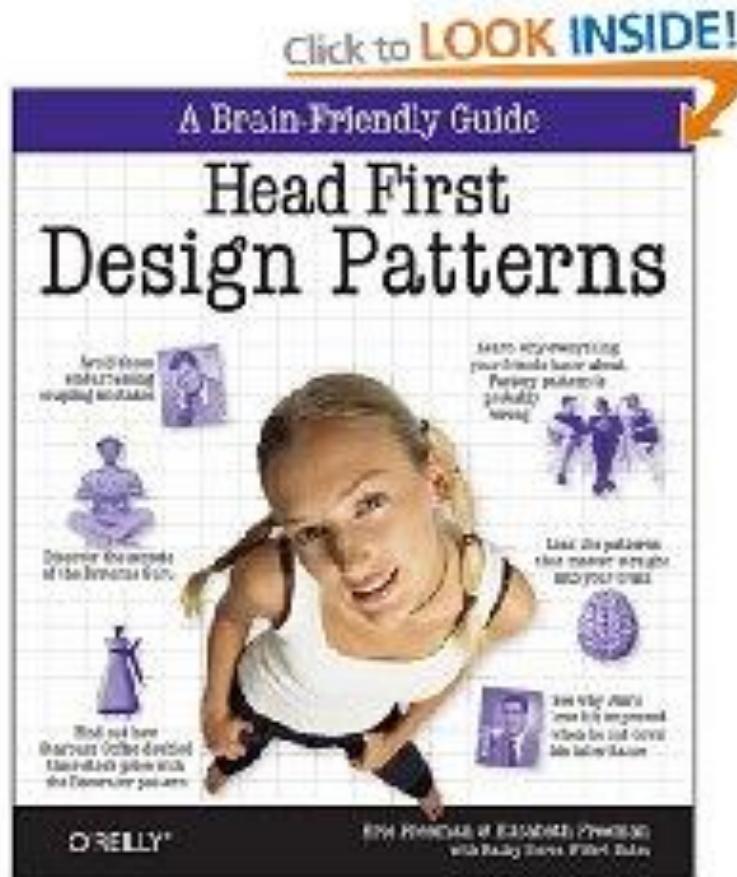
Christopher Alexander



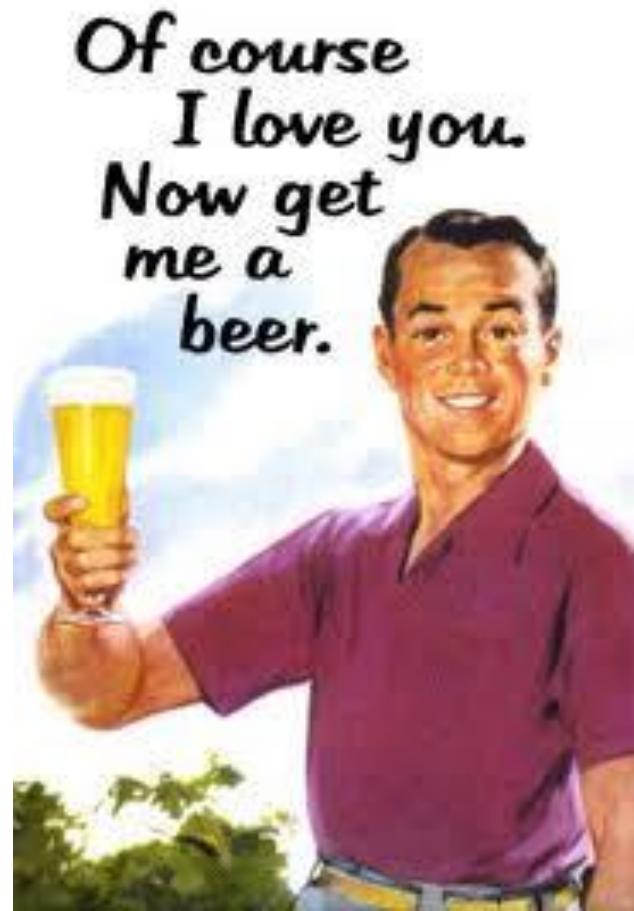
GoF (1994)



Rather this!



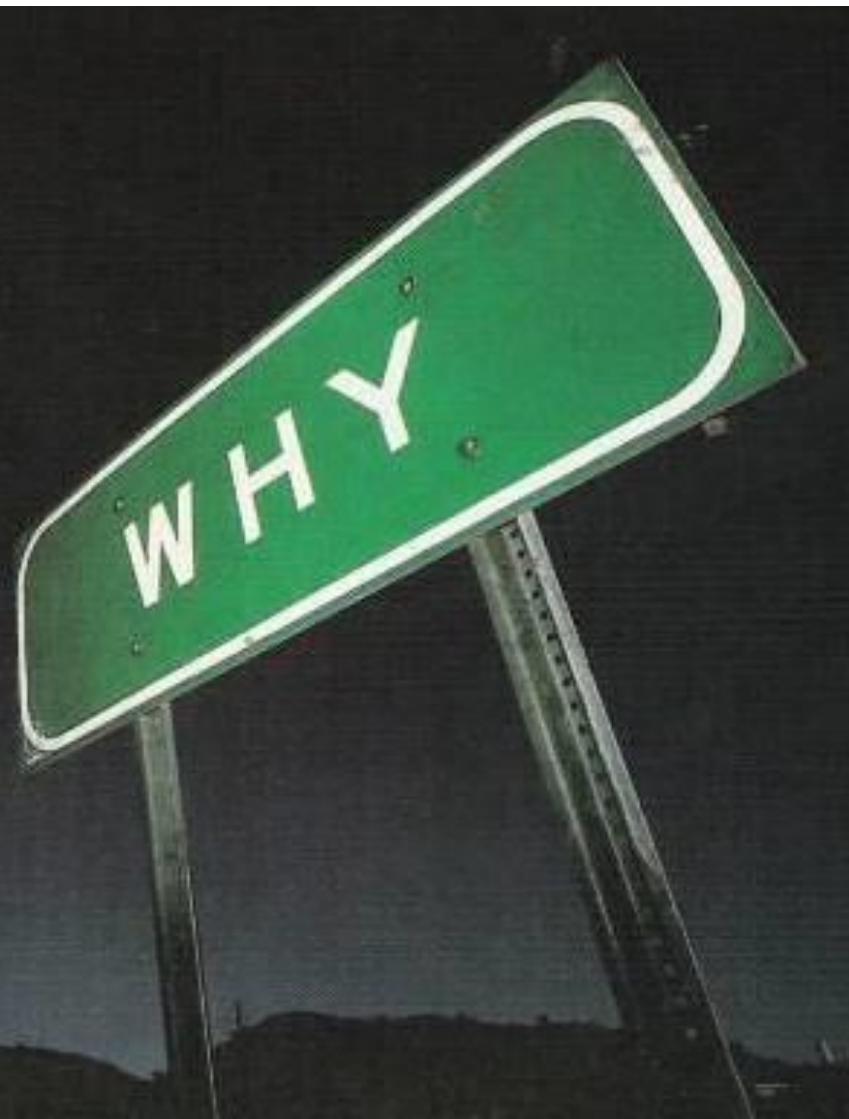
Loved by thousands...



hated by even more...

<http://www.youtube.com/watch?v=8-6VyYrMnlw>





ORC 79-2
FEBRUARY 1979

(12)

THE SCOPE, LIMITS, AND TRAINING IMPLICATIONS OF THREE
MODELS OF AIRCRAFT PILOT EMERGENCY RESPONSE BEHAVIOR

by

STUART E. DREYFUS

and

HUBERT L. DREYFUS

DA071320



DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

OPERATIONS
RESEARCH
CENTER

UNIVERSITY OF CALIFORNIA · BERKELEY

PASS SQL
saturday

The Dreyfus^{^2} expert model



Dreyfus level 1 - novice

Fresh from school – no experience

Needs rules to make decisions

Without experience these rules are context free

Example

A Mac user calls a support center, and is asked to press the Windows button



Dreyfus level 2 - advanced beginner

Recognizes features in a situation

Expands the set of context free rules with rules based on elements from the situation

Example

Asks for the user OS in the beginning of the support call. Then follows OS specific guide



Dreyfus level 3 - competent

Uses a hierarchical decision procedure.

Chooses a number of relevant elements
(makes a plan) and follow it

Example

In a trouble shooting situation where a changed configuration now gives an error, the set of elements are "what did we change?"

The plan is then: revert all changes, and then for each thing changed, enable it until you see the error again

Not considering if the things changed could have caused the symptoms

Dreyfus level 4 - proficient

The proficient uses maxims to guide him/her

Maxim:

- An expression of a general truth or principle
- a principle or rule of conduct

- Dictionary.com

Examples of maximes

From eXtreme Programming:

- ``simplest thing that could possibly work''
- ``test everything that could break.''

In IT operations

- ``automate!''

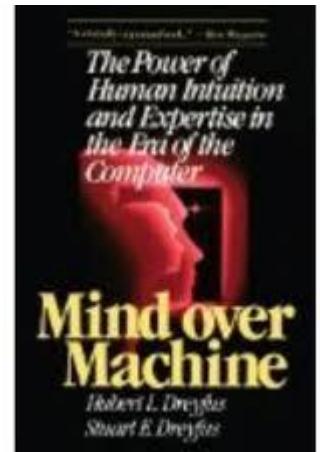
Dreyfus level 4 - proficient (cont.)

The proficient uses maxims to organize and understand the task at hand in an *intuitive* way

Then uses analysis to find out how to solve it

Intuition (Dreyfus^2 definition)

... [intuition is] the understanding that
effortlessly occurs upon seeing similarities with
previous experiences



Hubert L. Dreyfus og Stuart E. Dreyfus, Mind Over Machine: the Power of Human Intuition and Expertise in the Era of the Computer

Dreyfus level 4 - proficient (cont.)

In contrast to the use of rules on the first Dreyfus levels maxims are *rules interpreted in the given context*

Example

Take the maxim

- ``test everything that could break''

A beginner would ask: "but what exactly do I test?

The proficient will look at the situation, draw on his experience and then be able to make a list of relevant tests

Dreyfus level 5 - expert/master

Read my work on expertise in IT operations
(in Danish only)



<http://itu.dk/people/kennie/masterprojekt/>

Why all this philosophy?

Anti patterns are maxims
for bad design



An anti pattern

"... is a pattern that may be commonly used but is ineffective and/or counterproductive in practice"

- Some repeated pattern of action, process or structure that initially appears to be beneficial, but ultimately produces more bad consequences than beneficial results, and
- A refactored solution exists that is clearly documented, proven in actual practice and repeatable.



FOR YOU

ImpImages.com



And now to something completely different...



Cat Logic



Full

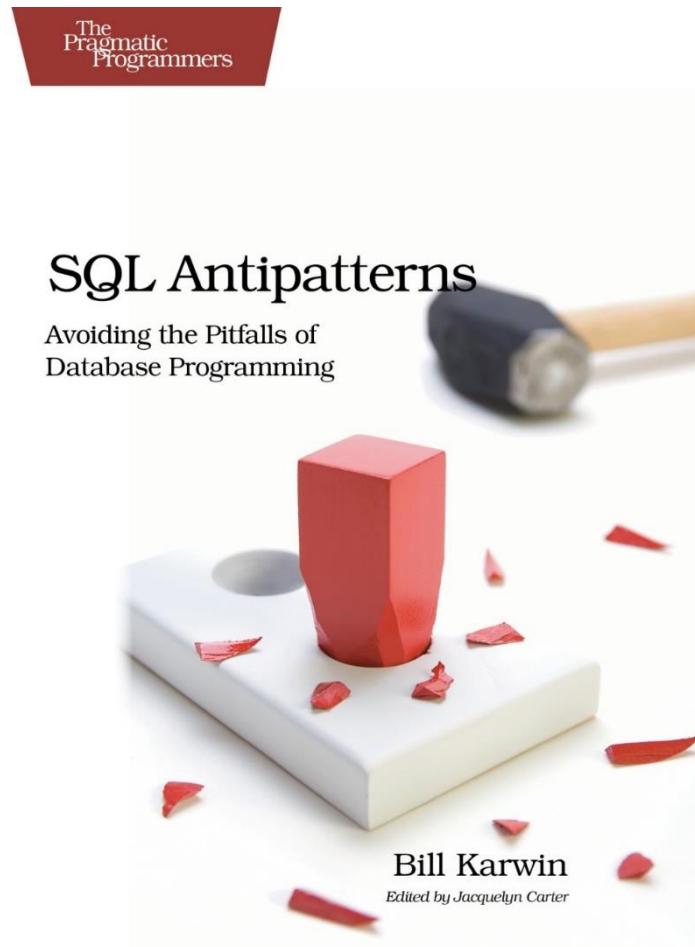


Empty

Logical anti patterns

- ID required (4)
- Fear of the unknown (14)
- Entity-Attribute-Value (6)

Buy this book!



ID required

- Anti pattern: Every table has an id column
- Story: This is how we do it in the Microsoft world or the story on how the OR-mapper tried to control the database schema
- Example: relation table with id

Fear of the unknown

- Anti pattern: don't use null
- Nulls are great:
 - null means "unknown"
 - Works well in aggregation functions
 - Can cause problems in joins – left outer
- Example: The 9999-12-31 date problem
- Example: null as status in printer job model

Entity-Attribute-Value

- Anti pattern: schemaless model
- Story: The big mother type table
- Story: scs_texts and scs_text_values
- Story: Effektor versions 1-4
 - 60 times better load time
- Story: Organizable entity – the datamodel from hell (when querying it)



Physical anti patterns

- Rounding errors (10)
- 31 flavors (11)
- Phantom files (12)
- Index shotgun (13)

Rounding errors

Theorem 6

Let p be the floating-point precision, with the restriction that p is even when $\beta > 2$, and assume that floating-point operations are exactly rounded. Then if $k = \lceil p/2 \rceil$ is half the precision (rounded up) and $m = \beta^k + 1$, x can be split as $x = x_h + x_l$, where $x_h = (m \otimes x) \ominus (m \otimes x \ominus x)$, $x_l = x \ominus x_h$, and each x_i is representable using $\lfloor p/2 \rfloor$ bits of precision.

Rounding errors

Anti pattern: use float data type

59.95 cannot be finitely represented in IEEE 754
base 2

```
create table real_demo(x binary_float);
insert into real_demo(x) values (59.95);
select * from real_demo where x = 59.95;
```

David Goldberg: what every ...

Rounding errors (cont.)

```
declare @tReal table(col real)
declare @tFloat table(col float)

insert into @tReal(col) values (59.95)
insert into @tFloat select col from @tReal

select r.col, f.col
from @tReal as r
cross join @tFloat as f
```

100 %



Results



Messages

	col	col
1	59,95	59,9500007629395

31 flavors





31 flavors

Anti pattern: Business logic in a check constraint

=> Maintenance nightmare

When is it ok?

Story: the SKMS project

Phantom files

Anti pattern: file references in the database, files on the file system

Problems:

- Delete from ...
- Transaktioner
- Rollback
- Backup
- Database privileges

Use: Filetable/Filestream, BFILE

Index shotgun



Index shotgun

Anti pattern:

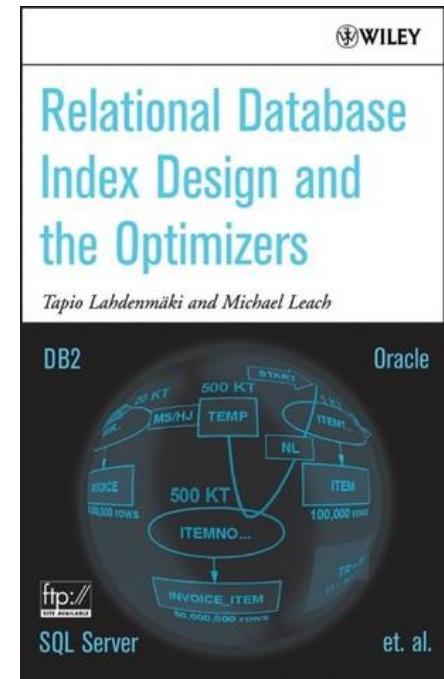
```
while(query performance problem) {  
    create new index  
}
```

Index design:

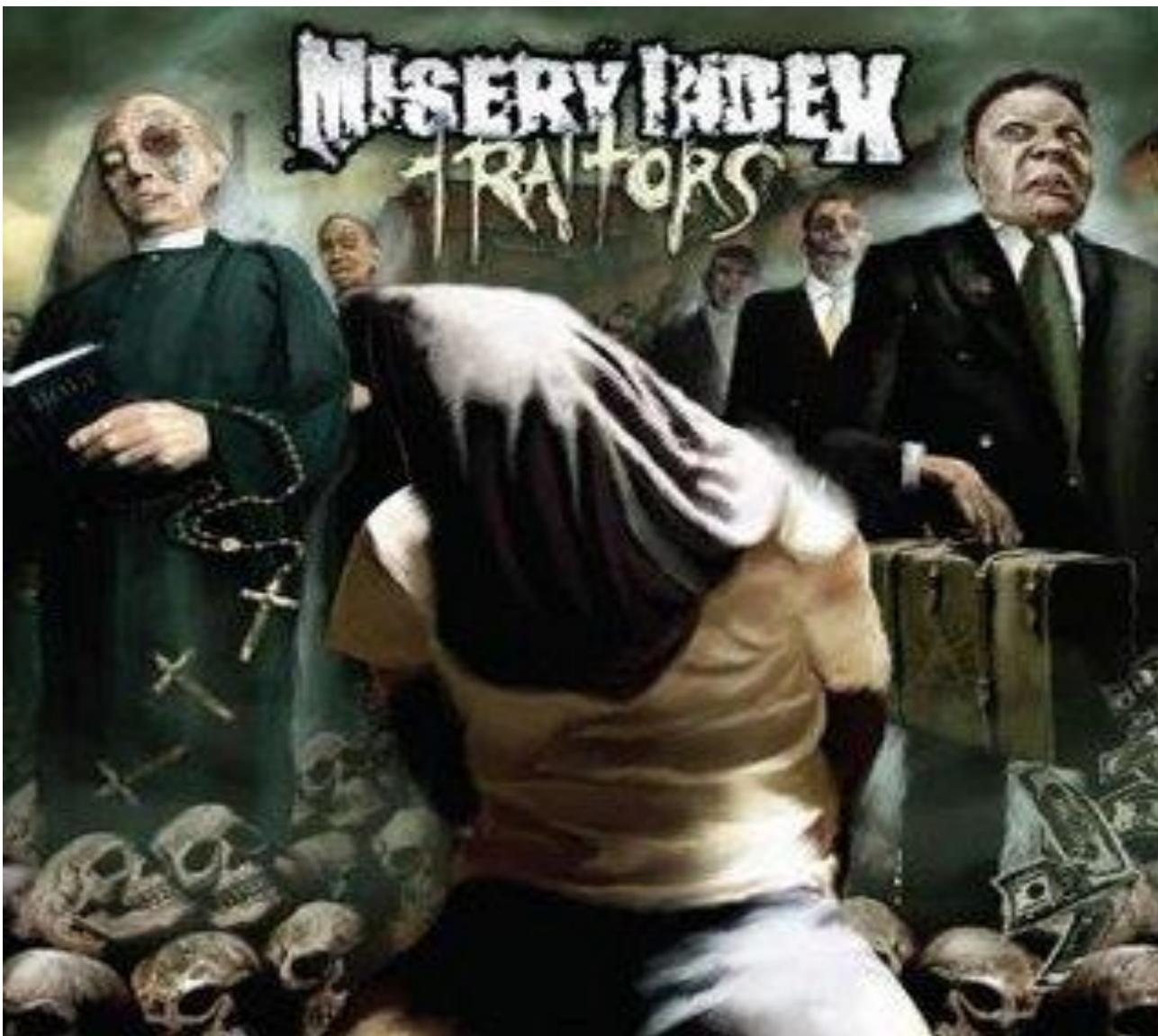
not my problem, your problem

Better:

- Know your optimizer
- DBA's and developers work together
- Learn the white art of index design



<http://www.youtube.com/watch?v=hB8JQ1kHHbo>



The SQL body used to be naked...

```
ORACLE://HR.ORCL/Query ORACLE://SYS.ORCL/Query Start Page

END "IS_PRIME";
-- pl/sql procedure to add new job to jobs table
PROCEDURE "ADD_NEW_JOB" ("P_JOB_ID" IN JOBS.JOB_ID%TYPE,
                          "P_JOB_TITLE" IN JOBS.JOB_TITLE%TYPE,
                          "P_MIN_SALARY" IN JOBS.MIN_SALARY%TYPE,
                          "P_MAX_SALARY" IN JOBS.MAX_SALARY%TYPE) IS
BEGIN
  -- use the package variable JOBS_REC to create new record
  jobs_rec.job_id := p_job_id;
  jobs_rec.job_title := p_job_title;
  jobs_rec.min_salary := p_min_salary;
  jobs_rec.max_salary := p_max_salary;
  -- insert the job record into the table
  insert into jobs (job_id, job_title, min_salary, max_salary)
  values (jobs_rec.job_id, jobs_rec.job_title,
          jobs_rec.min_salary, jobs_rec.max_salary);
END "ADD_NEW_JOB";
END "OBE";
```

The command(s) successfully completed.

Google this (at home)

sql

naked

woman

Query anti patterns

- Random selection (16)
- Poor mans search engine (17)
- Implicit columns (19)

Random selection

Anti pattern: get a single random record (and a full table scan for free)

Examples:

```
select * from some_table order by rand() limit 1
```

Random selection (Oracle)

Better:

```
select * from (
    select *
        from some_table sample(1)
        order by dbms_random.value
)
where rownum = 1
```

Random selection (sqlserver)

Better:

```
select * from some_table  
tablesample(1 rows)
```

GOOGOTH

[Web](#) [Images](#) [News](#) [Maps](#) [Products](#) [Groups](#) [Gothic Den](#) [more>>](#)

Search Google web pages

- [Advanced Search](#)
- [Preferences](#)
- [Language Tools](#)

[Googoth Search](#) [I'm Feeling Depressed](#)

[Google Online Jobs](#) [Gizoogle](#)

Poor mans search engine

Anti pattern: implement searching with like operator:

```
select *
  from some_table
 where upper(text_field) like '%upper(search_pattern)%'
```

Problems:

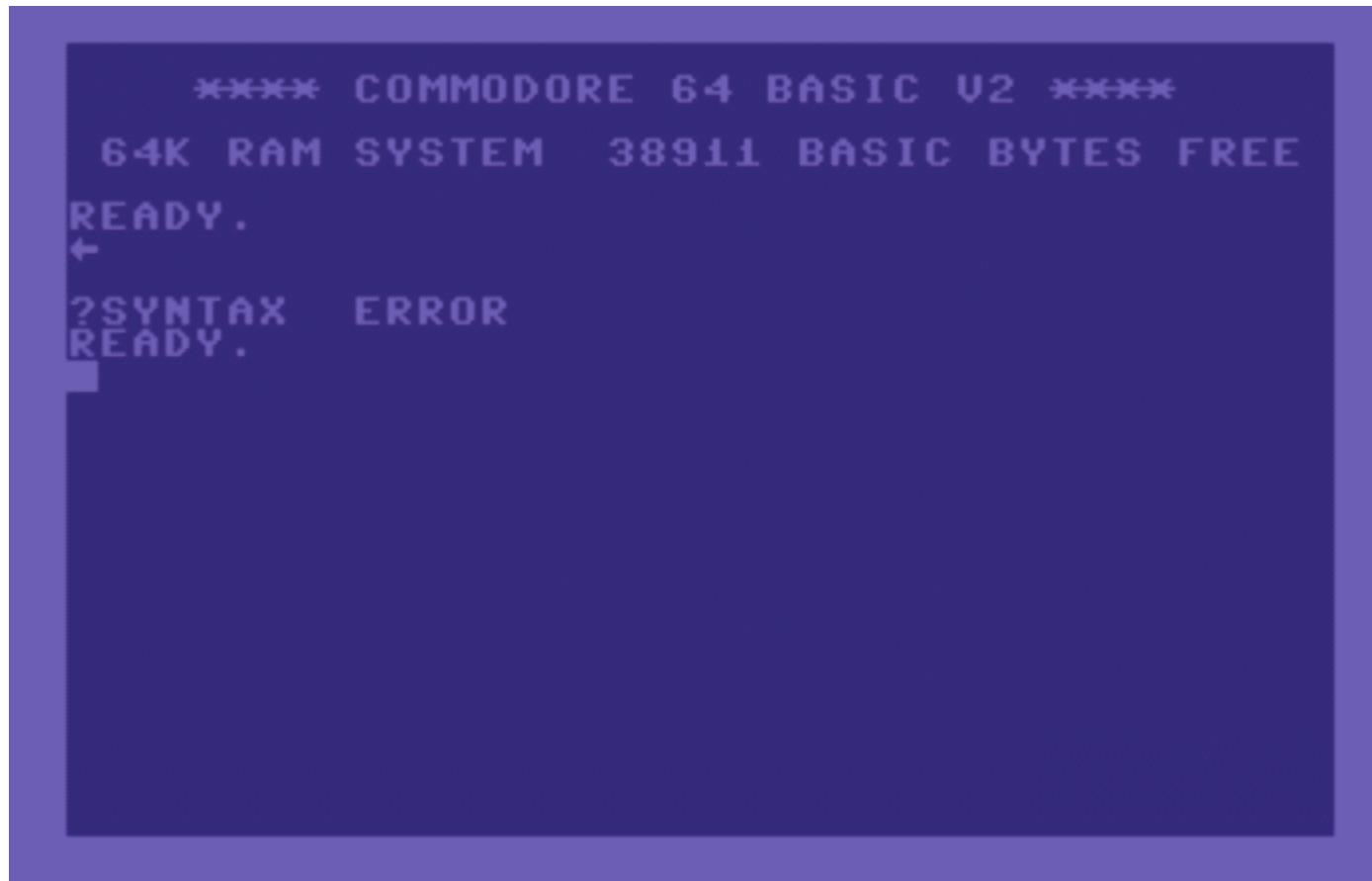
- Index usage?
- like '%search_string'
- like 'search_string%'

Better: use build in functionality:

OracleText/FullTextSearch/lucene.net



Did you spot the syntax error?



Implicit columns

Anti pattern:

semantics based on current table design.

Implicit semantics is (almost always) bad!

Examples

- insert into tableA select *
- Insert into tableB values (1,2,3)

Story: DEF views

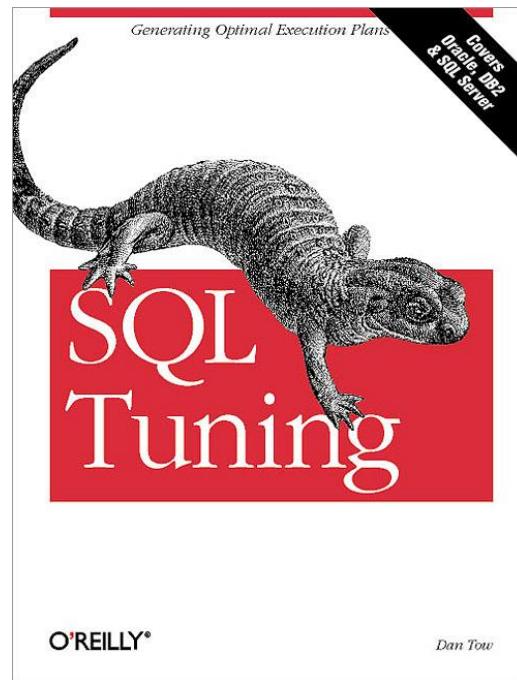
Better:

- `insert into tableA(col1, col2) select 1 as col1, 2 as col2`



Reporting anti patterns

- How do you fix report performance problems
- 4 bad habits



How do you fix report performance problems?

"Beginning with the assumption that no one person will ever read a huge application report from cover to cover, why should applications ever need huge report queries?"

Dan Tow: "SQL Tuning", kapitel 10



Reasons for large reports - 1

A report has many readers, each of whom is interested in a different subset of the data. No one reads the report from cover to cover, but any given part of it might be interesting to at least one reader

Action: Split it!

Reasons for large reports - 2

All details of a report are potentially interesting at the time the report is requested, but end users will read only a small part of the report, based on which questions happen to arise that the end users must answer.

Action: BI - Drill down



Reasons for large reports - 3

Only a subset of the query data is ever used

Action: Trim it

Reasons for large reports - 4

A report is required only for legal reasons, not because anyone will ever read it

Action: Kill it or live with it!

Two queries walk into a bar. They overhear a query next to them say, 'Shut the @#!@#% up you dumb son of a #\$_@%&!@\$'). The first query leans over to the second and whispers, 'hmmm... she must be using a CURSOR'...

http://www.reddit.com/r/programming/comments/7v978/a_sql_query_walks_into_a_bar_and_sees_two_tables/





ES QUEUE EL