DirX – Query Optimization

Frank Huber

VDE/ITG-FG 5.2.4
"Netzdatenbanken"

Outline

- DirX / LDAP / X.500
- Query processing in general
- Query processing and optimization in DirX
  - Rewriting
  - Optimization
  - Execution
- Results
- Relation to network databases
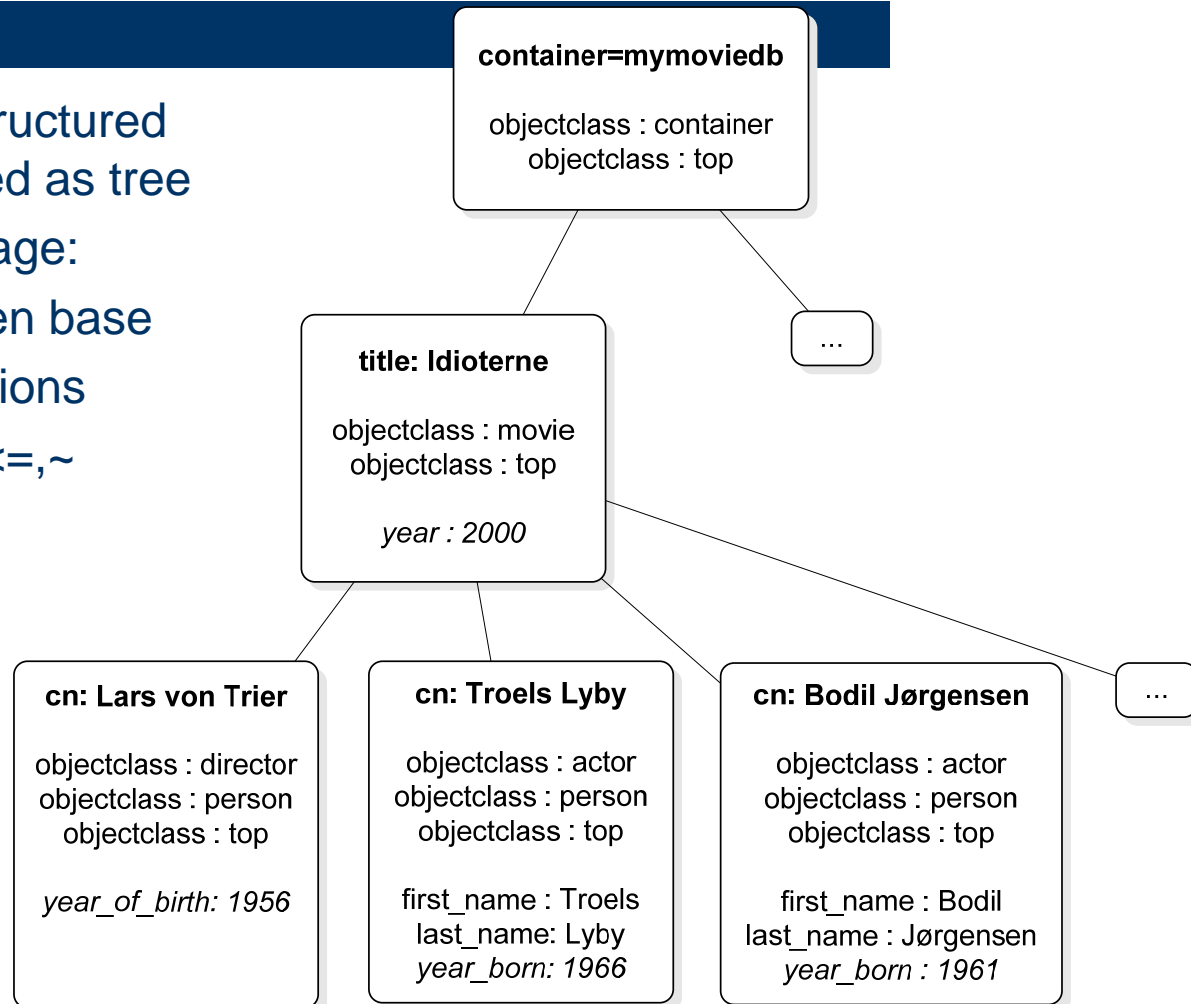
Outline

- ***DirX / LDAP / X.500***
- Query processing in general
- Query processing and optimization in DirX
  - Rewriting
  - Optimization
  - Execution
- Results
- Relation to network databases

DirX / LDAP / X.500

- DirX – LDAPv3, DSMLv2 and X.500 directory server
- Base for Identity Management Systems
  - Identities like persons, organizations, applications, …
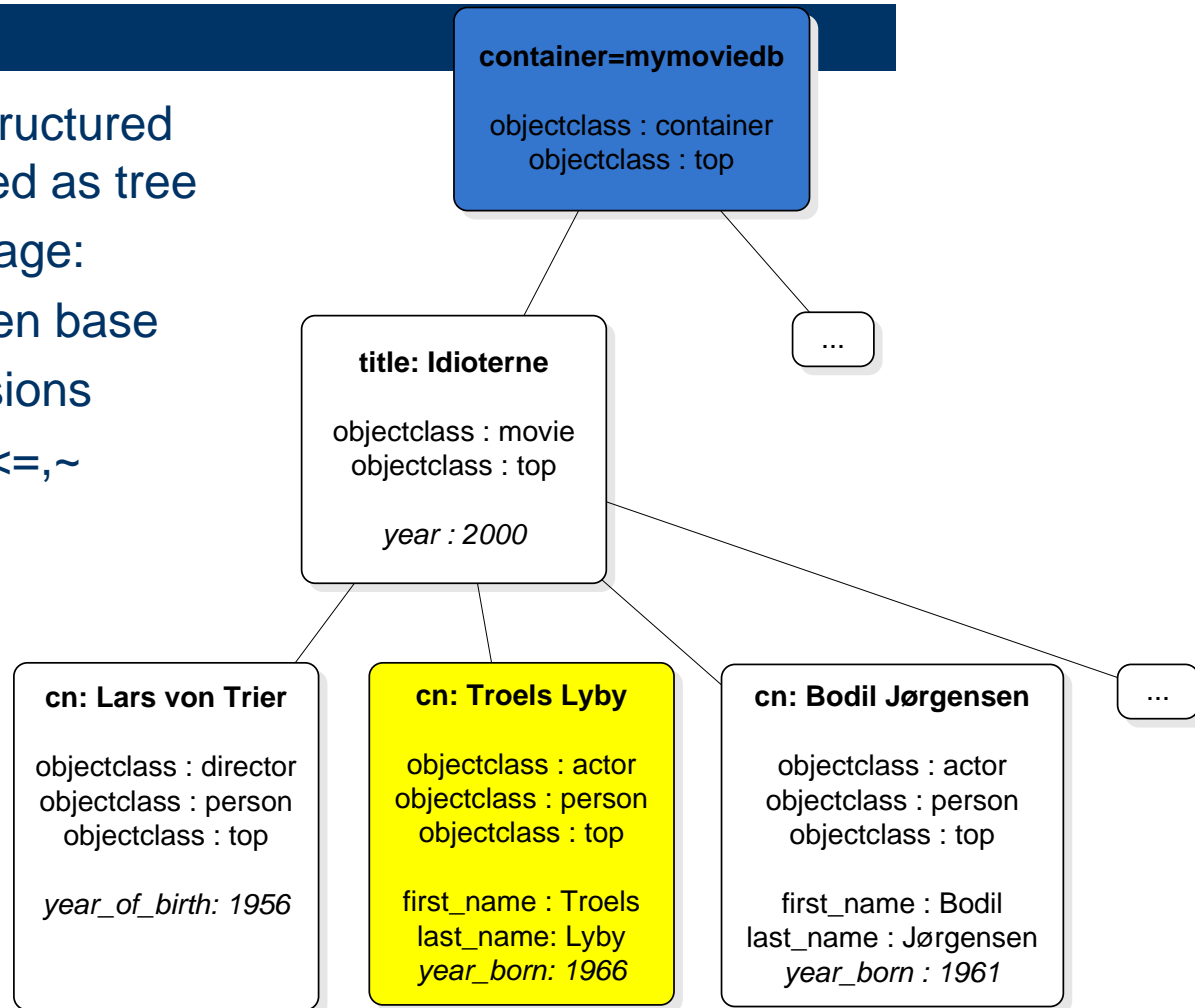- Meta Directories

LDAP / X.500

- Hierarchical, semi-structured data model, organized as tree
- LDAP – query language:
  – Relative to a given base
  – Boolean expressions
  – Operators =,>=,<=,~

  and wildcards

container=mymoviedb

objectclass : container
objectclass : top

...

title: Idioterne

objectclass : movie
objectclass : top

year : 2000

...

cn: Lars von Trier

objectclass : director
objectclass : person
objectclass : top

year_of_birth: 1956

cn: Troels Lyby

objectclass : actor
objectclass : person
objectclass : top

first_name : Troels
last_name: Lyby
year_born: 1966

cn: Bodil Jørgensen

objectclass : actor
objectclass : person
objectclass : top

first_name : Bodil
last_name : Jørgensen
year_born : 1961

LDAP / X.500

- Hierarchical, semi-structured data model, organized as tree
- LDAP – query language:
  - Relative to a given base
  - Boolean expressions
  - Operators =,>=,<=,~
  and wildcards
- Example query:
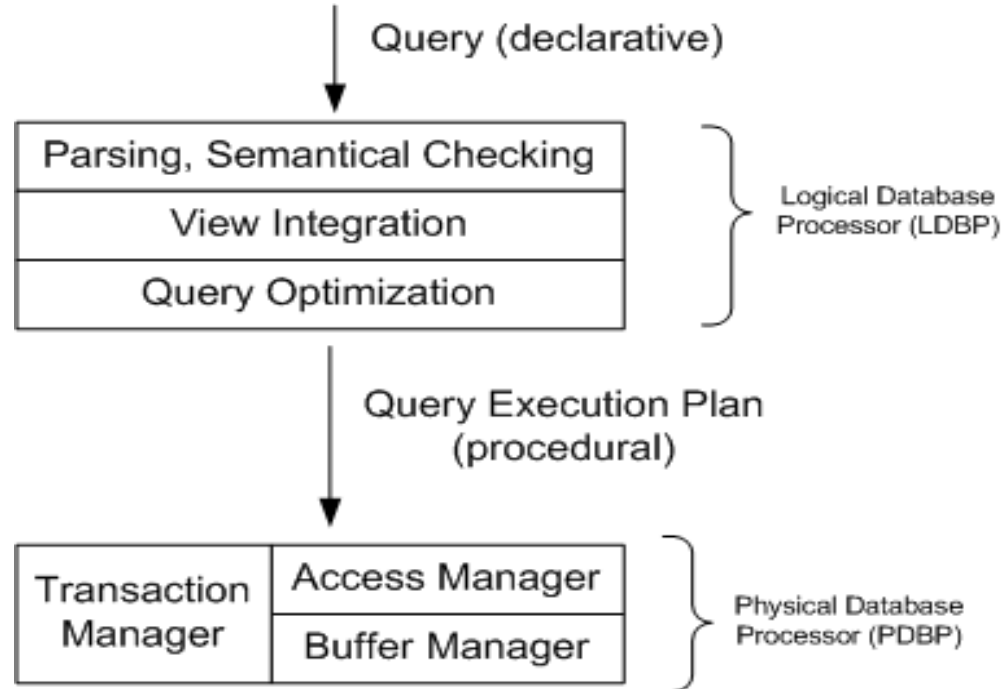
/container=mymoviedb,
SUB,
&((lastName=Lyby)
   (year_born=1966))

container=mymoviedb

objectclass : container
objectclass : top

...

title: Idioterne

objectclass : movie
objectclass : top

year : 2000

...

cn: Lars von Trier

objectclass : director
objectclass : person
objectclass : top

year_of_birth: 1956

cn: Troels Lyby

objectclass : actor
objectclass : person
objectclass : top

first_name : Troels
last_name: Lyby
year_born: 1966

cn: Bodil Jørgensen

objectclass : actor
objectclass : person
objectclass : top

first_name : Bodil
last_name : Jørgensen
year_born : 1961

Outline

- DirX / LDAP / X.500
- ***Query processing in general***
- Query processing and optimization in DirX
  - Rewriting
  - Optimization
  - Execution
- Results
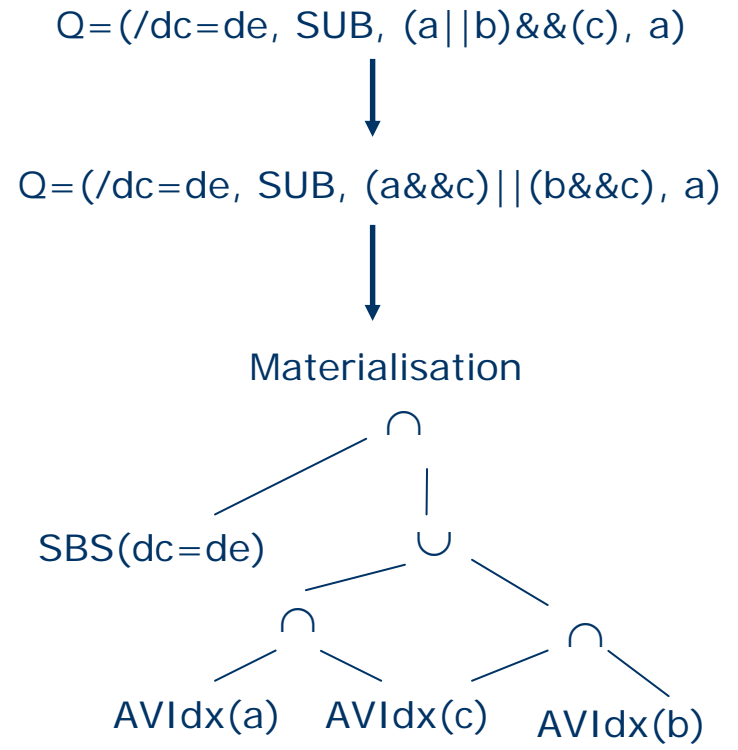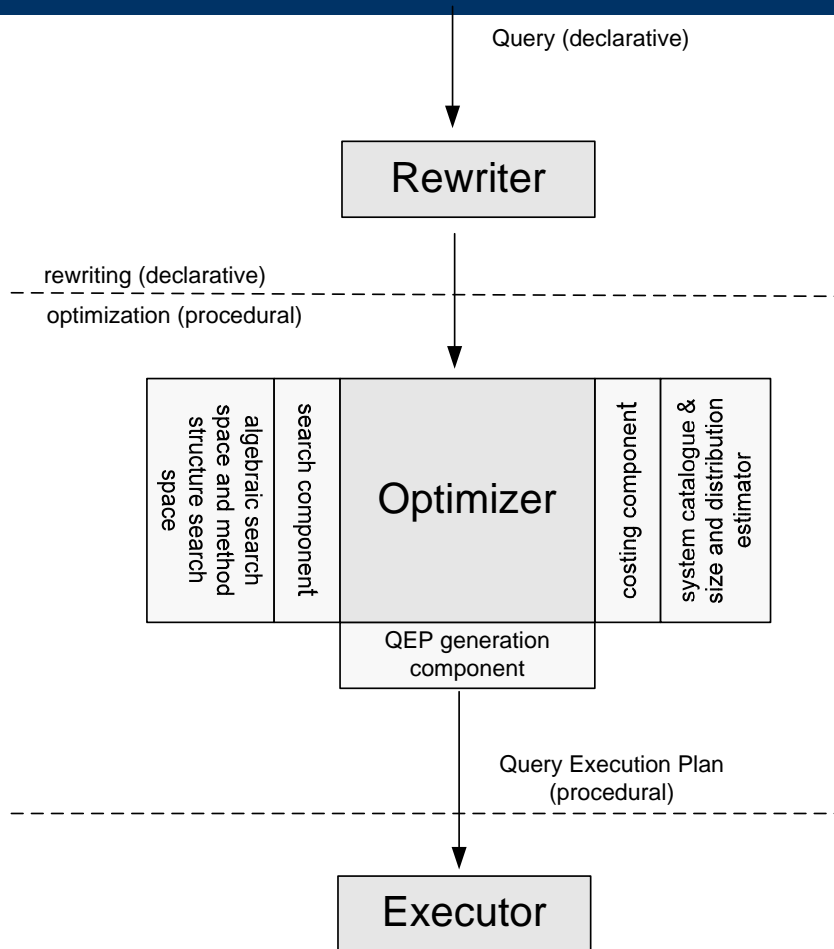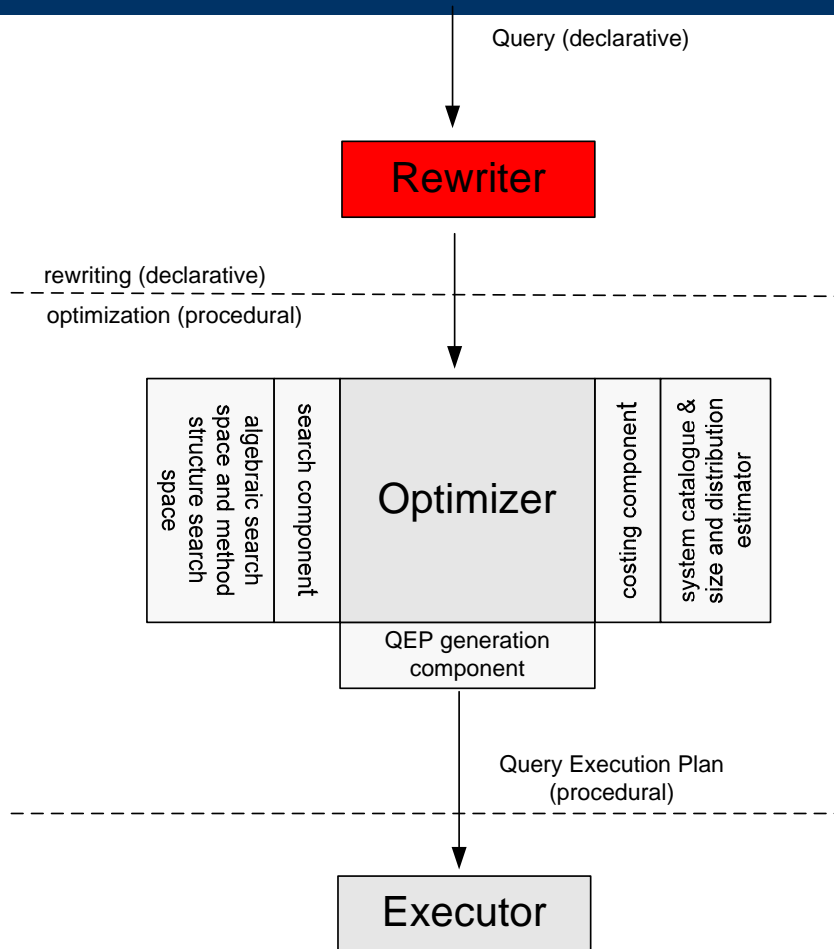- Relation to network databases

Query processing

Outline

- DirX / LDAP / X.500
- Query processing in general
- ***Query processing and optimization in DirX***
  - Rewriting
  - Optimization
  - Execution
- Results
- Relation to network databases

Query processing in DirX

Query (declarative)

Rewriter

rewriting (declarative)

optimization (procedural)

algebraic search space and method structure search space	search component	Optimizer	costing component	system catalogue & size and distribution estimator

QEP generation component

Query Execution Plan (procedural)

Executor

Q=(/dc=de, SUB, (a||b)&&(c), a)

Q=(/dc=de, SUB, (a&&c)||(b&&c), a)

Materialisation

SBS(dc=de)

AVIdx(a)   AVIdx(c)   AVIdx(b)

Rewriter

Query (declarative)

Rewriter

rewriting (declarative)

optimization (procedural)

Optimizer

algebraic search
space and method
structure search
space

search component

costing component

system catalogue &
size and distribution
estimator

QEP generation
component

Query Execution Plan
(procedural)

Executor

$Q=(/dc=de, \; SUB, \; (a||b)\&\&(c), \; a)$

$Q=(/dc=de, \; SUB, \; (a\&\&c)||(b\&\&c), \; a)$

Rewriter

- Filter transformation into disjunctive normal form (DNF)
- Checking satisfiability against schema definition
- Finding tautologies, contradictions, redundancies, ranges
- Grouping disjunctions with same attribute type with IN-operators

Optimizer

Query (declarative)

Rewriter

rewriting (declarative)

optimization (procedural)

Optimizer

algebraic search space and method structure search space

search component

costing component

system catalogue & size and distribution estimator

QEP generation component

Query Execution Plan (procedural)

Executor

Q=(/dc=de, SUB, (a&&c)||(b&&c), a)

Materialisation

SBS(dc=de)

AVIdx(a)   AVIdx(c)   AVIdx(b)
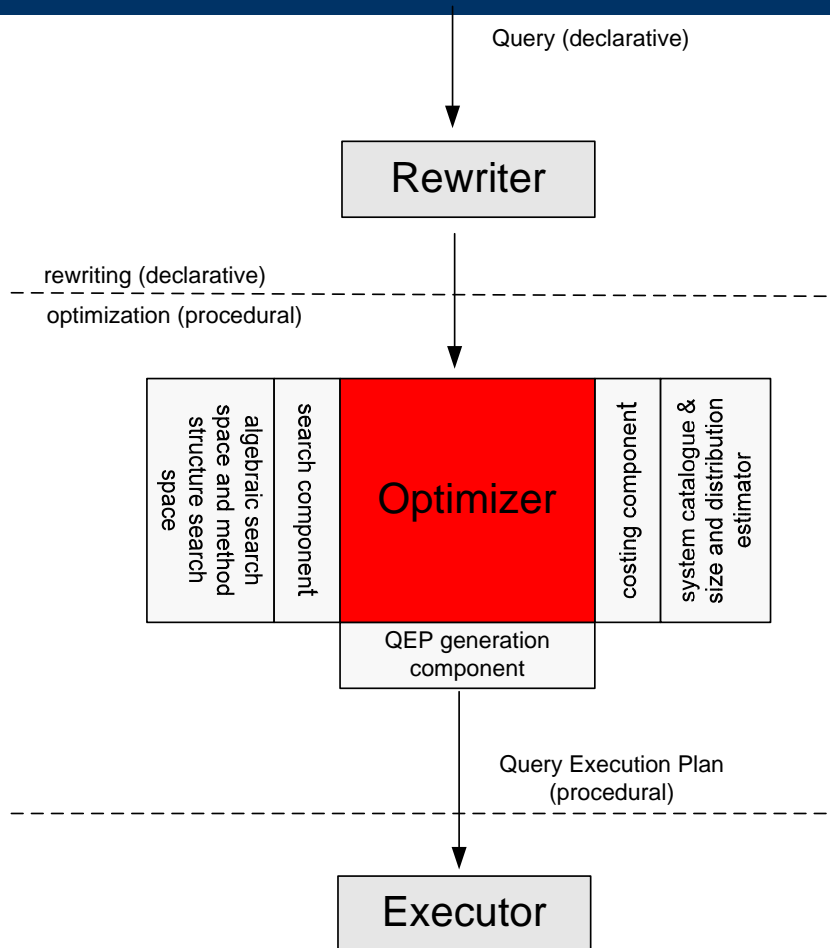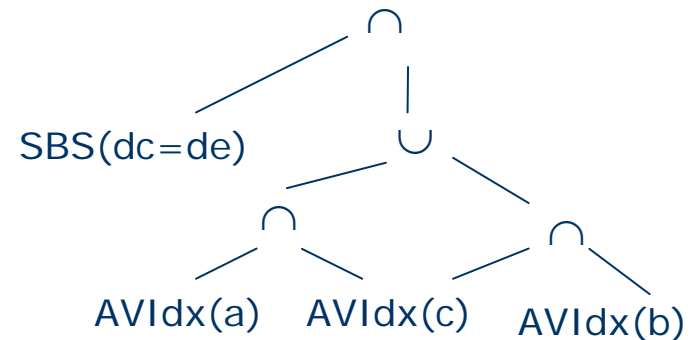
Optimizer

- Creates „optimal" query execution plan (QEP)
- Cost model defines optimality
- Output: procedural QEP
  - Directed, acyclic graph
    - edges ➜ flow of data
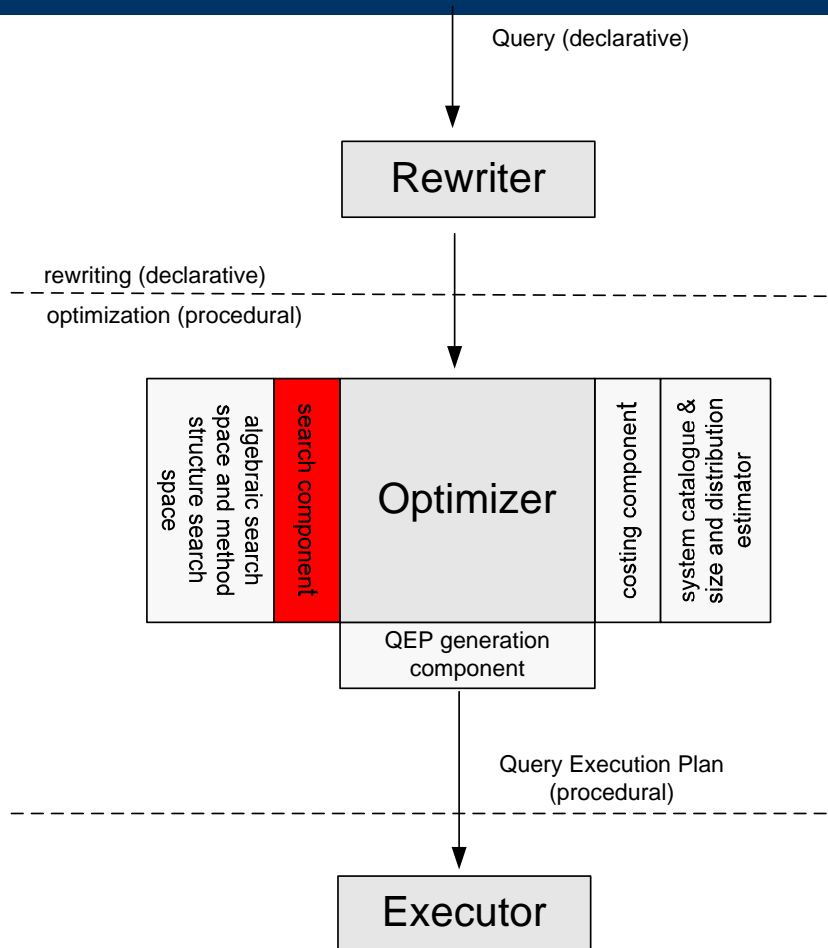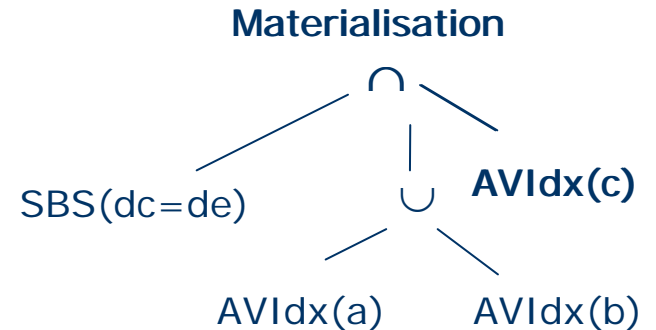   - nodes ➜ operators

Search component

Query (declarative)

Rewriter

rewriting (declarative)
optimization (procedural)

algebraic search space and method structure search space

search component

Optimizer

costing component

system catalogue & size and distribution estimator
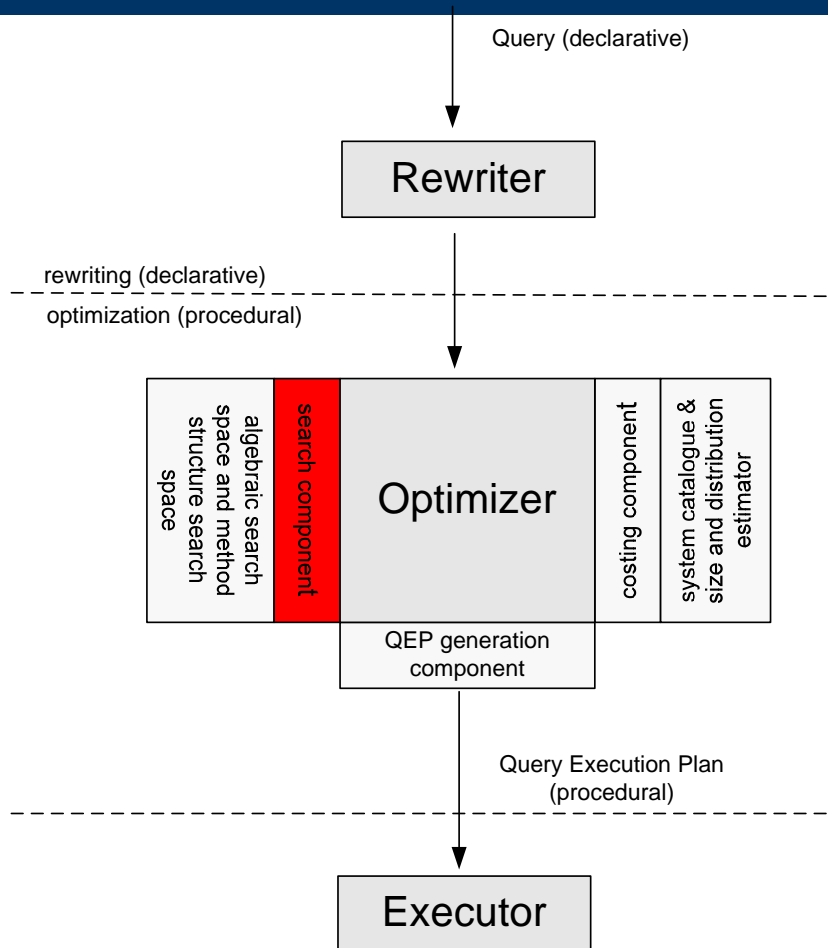
QEP generation component

Query Execution Plan (procedural)

Executor
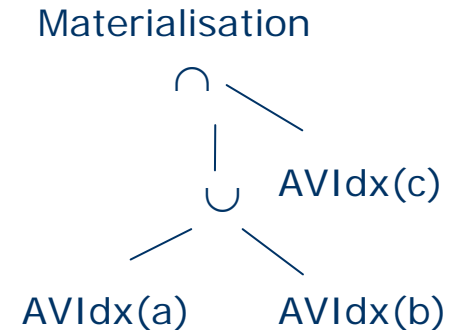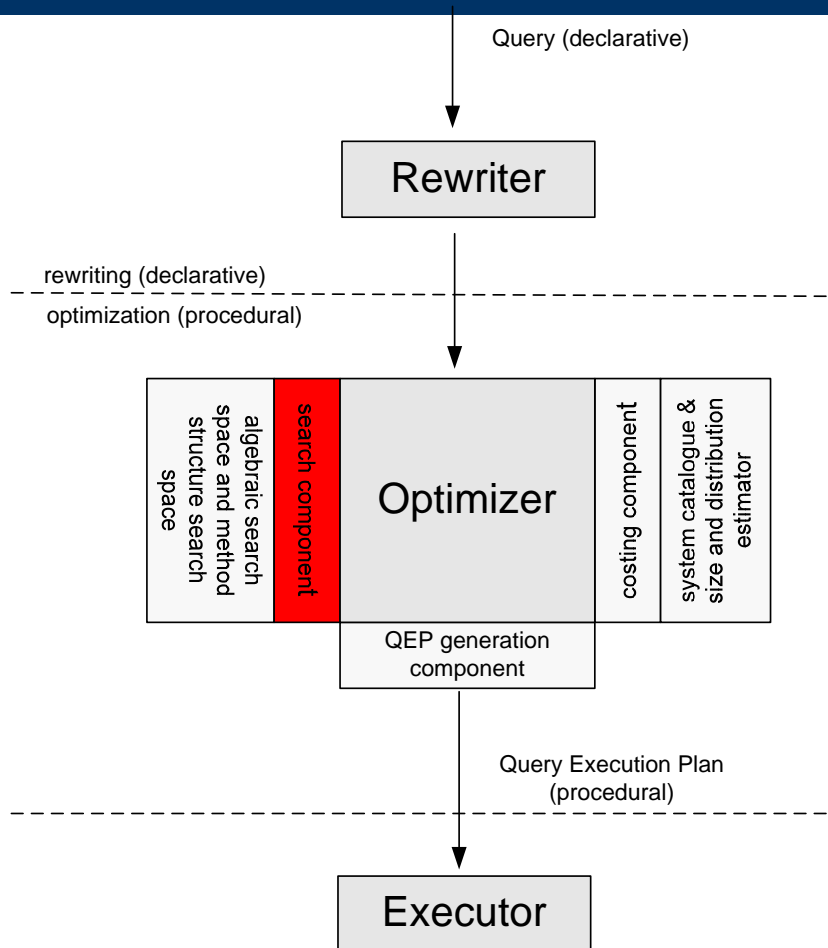
- Explores the search space
- Uses methods of factorization and pruning

Materialisation

SBS(dc=de)

AVIdx(a)   AVIdx(c)   AVIdx(b)

Search component

Query (declarative)

Rewriter

rewriting (declarative)

optimization (procedural)

| algebraic search space and method structure search space | search component | Optimizer | costing component | system catalogue & size and distribution estimator |

QEP generation component

Query Execution Plan (procedural)

Executor

- Explores the search space
- Uses methods of factorization and pruning

Materialisation

SBS(dc=de)

AVIdx(c)

AVIdx(a)   AVIdx(b)

Search component

Query (declarative)

Rewriter

rewriting (declarative)

optimization (procedural)

| algebraic search space and method structure search space | search component | Optimizer | costing component | system catalogue & size and distribution estimator |

QEP generation component

Query Execution Plan (procedural)

Executor

- Explores the search space
- Uses methods of factorization and pruning

Materialisation

AVIdx(c)

AVIdx(a)   AVIdx(b)

Search component

- Explores the search space
- Uses methods of factorization and pruning

Costing Component

Query (declarative)

Rewriter

rewriting (declarative)

optimization (procedural)

algebraic search space and method structure search space

search component

Optimizer

costing component

system catalogue & size and distribution estimator

QEP generation component

Query Execution Plan (procedural)

Executor

C=1250
Materialisation

C=1000
SBS(dc=de)

C=250

C=30

C=220

AVIdx(a)

AVIdx(c)

AVIdx(b)

C=10

C=20

C=200
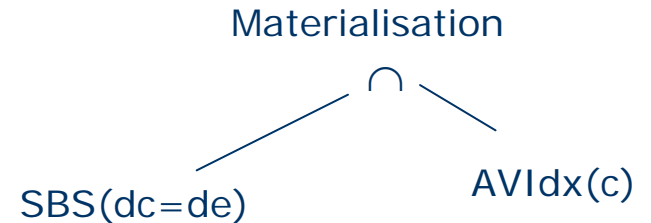
Costing component

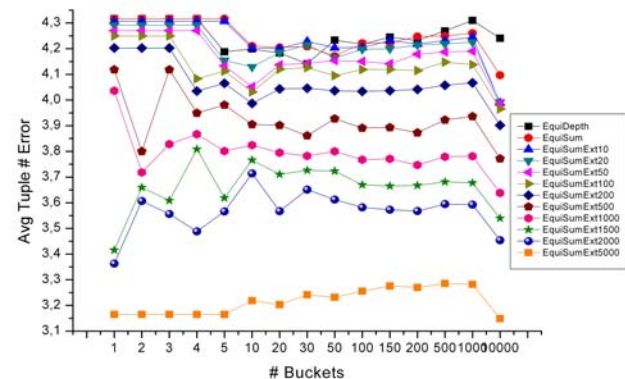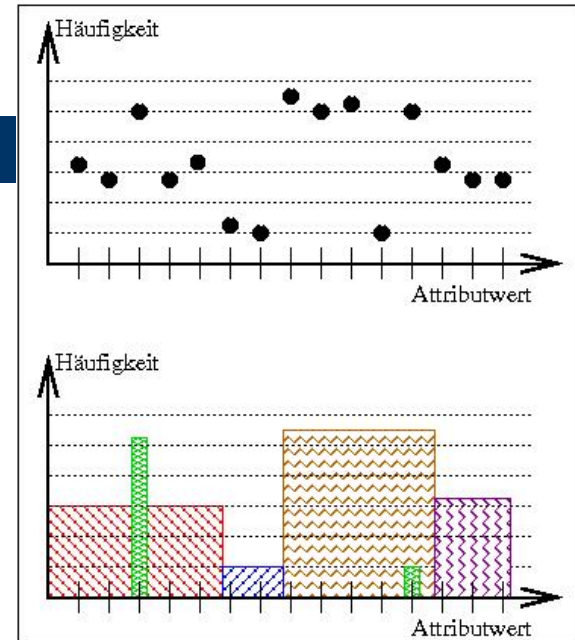– Cost model for each operator in the physical layer

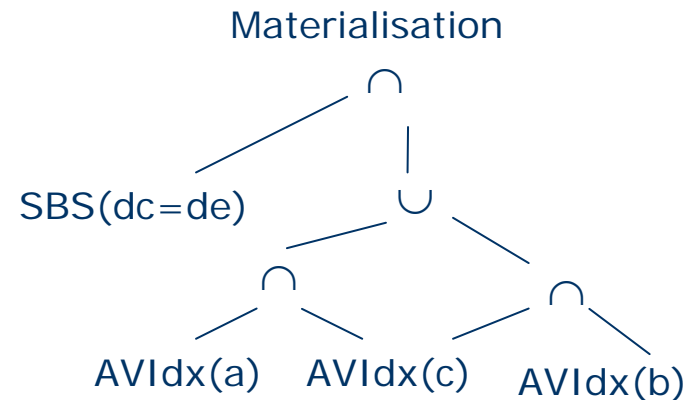– Statistical model for bit strings
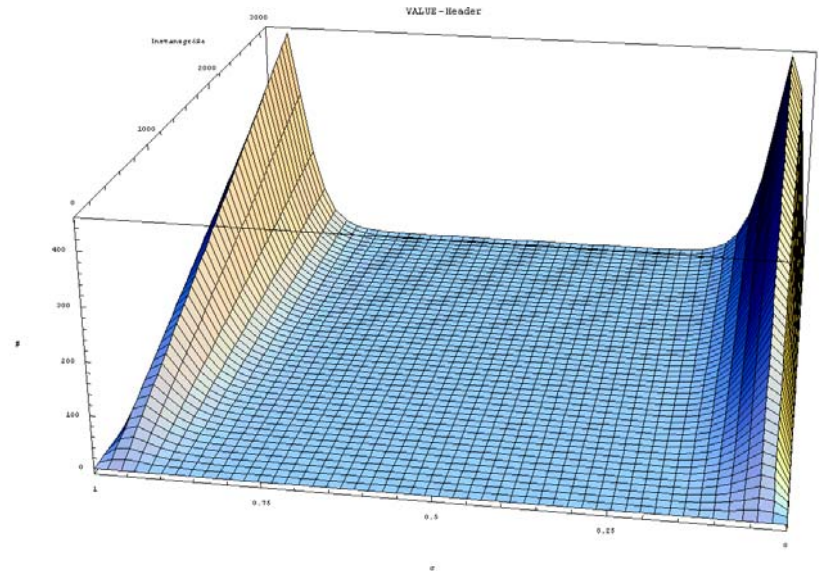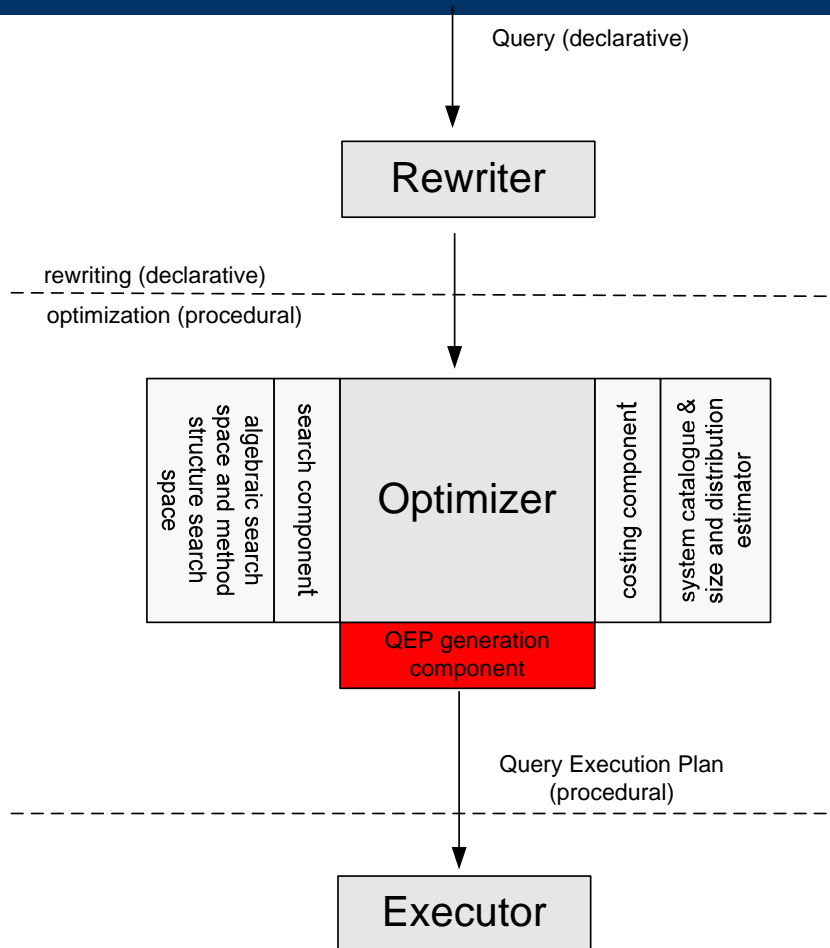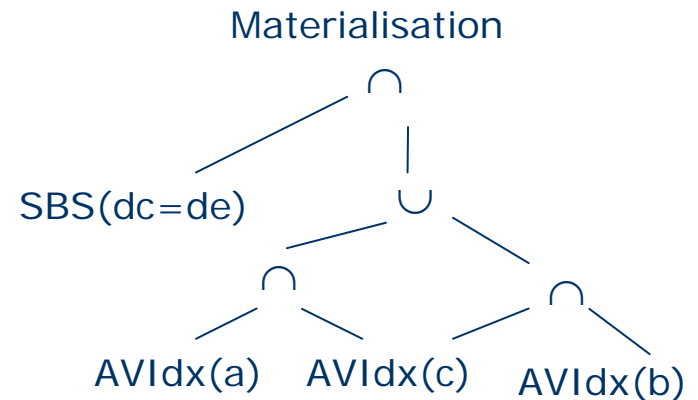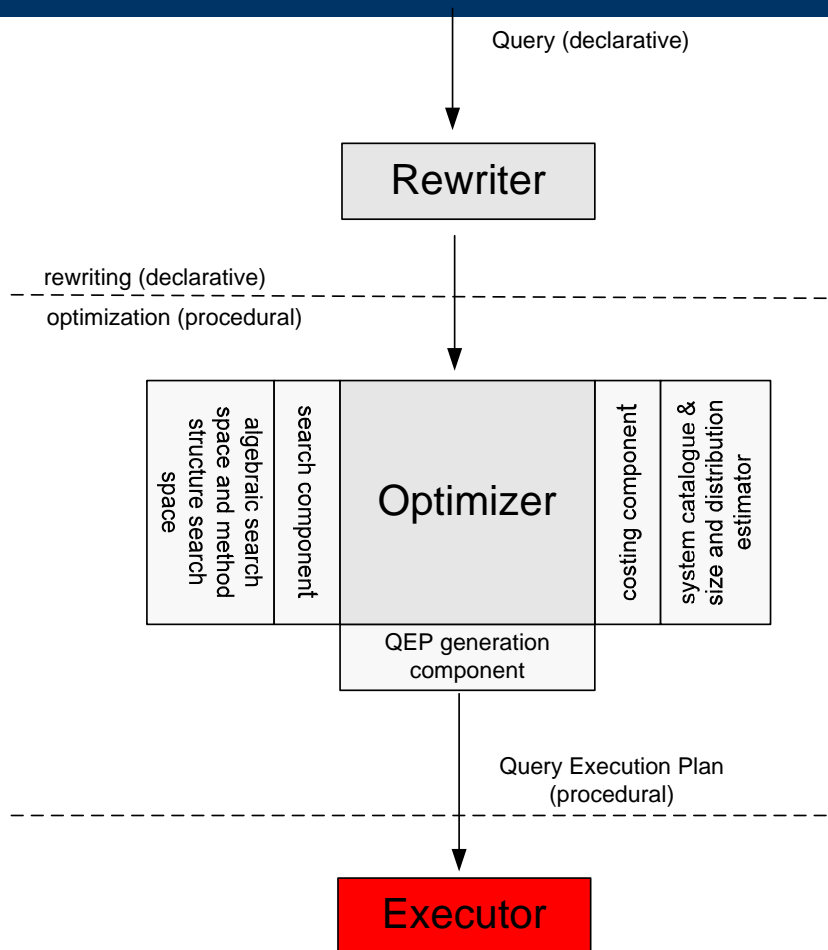
Statistics component
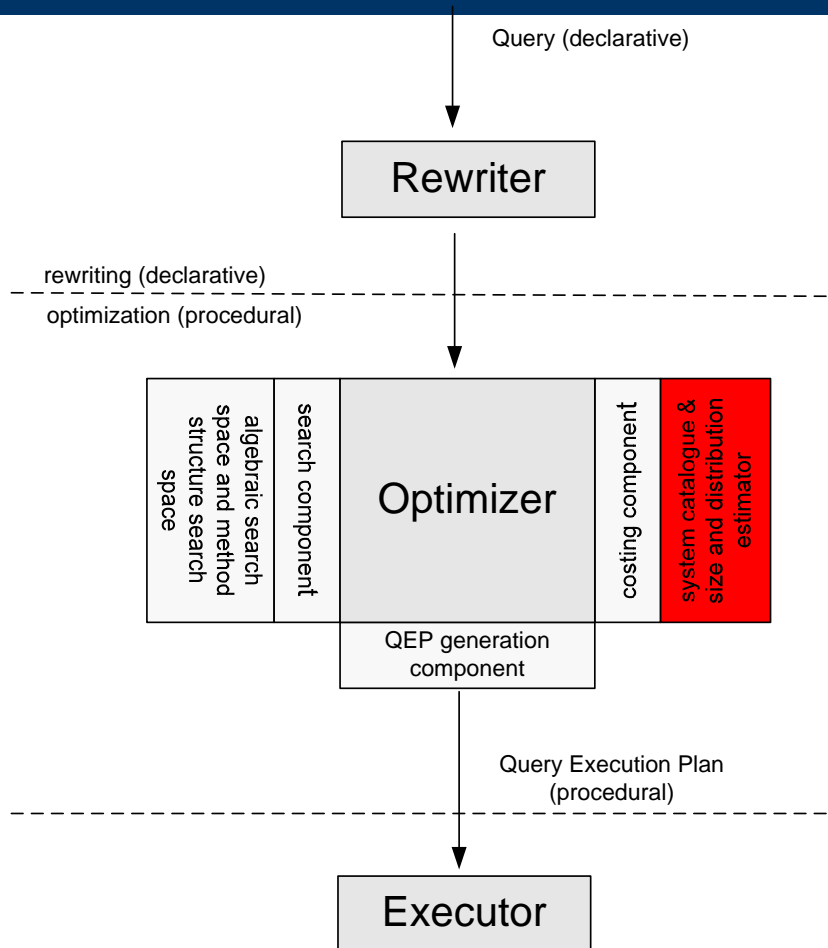
Statistics component

– Annotates terms with selectivity and other statistical parameter

– Gives statistical information about the whole directory and sub trees

– Uses a specialized histogram type

Query generation component

Query (declarative)

Rewriter

rewriting (declarative)

optimization (procedural)

algebraic search space and method structure search space	search component	Optimizer	costing component	system catalogue & size and distribution estimator

QEP generation component

Query Execution Plan (procedural)

Executor

Materialisation

SBS(dc=de)

AVIdx(a)   AVIdx(c)   AVIdx(b)

Query execution engine

Query (declarative)

Rewriter

rewriting (declarative)

optimization (procedural)

Optimizer

- algebraic search space and method structure search space
- search component
- costing component
- system catalogue & size and distribution estimator

QEP generation component

Query Execution Plan (procedural)

Executor

Materialisation

SBS(dc=de)

AVIdx(a)   AVIdx(c)   AVIdx(b)

Query execution engine

- Implements different operators:
  - SBSIndex()
  - AVIndex(), AVIndexRange(), AVIndexIn()
  - Union(), Intersection(), Complement()

- Evaluates the generated QEP

- Open/Next/Close-Paradigm

- Uses „Sideways Information Passing"

Outline

- DirX / LDAP / X.500
- Query processing in general
- Query processing and optimization in DirX
  - Rewriting
  - Optimization
  - Execution
- ***Results***
- Relation to network databases

Results
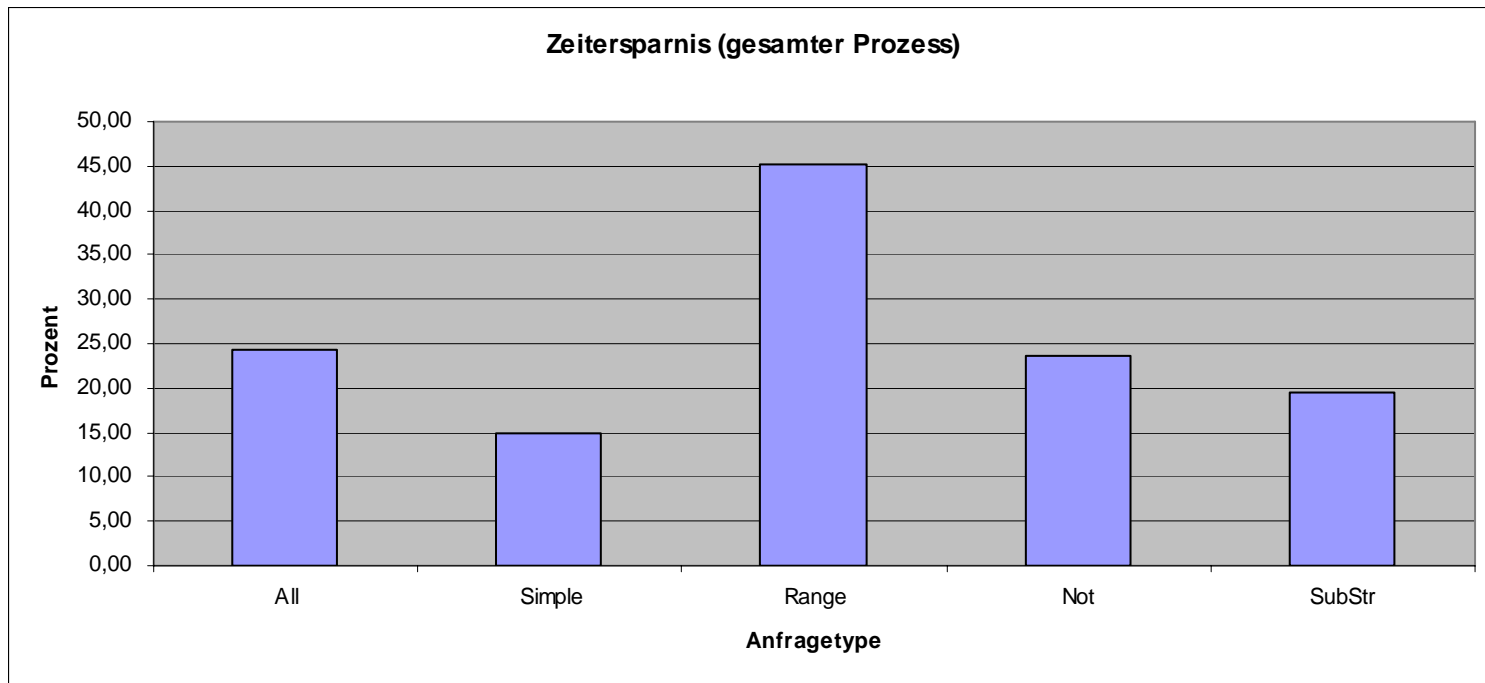
Zeitersparnis (gesamter Prozess)

Outline

- DirX / LDAP / X.500
- Query processing in general
- Query processing and optimization in DirX
  – Rewriting
  – Optimization
  – Execution
- Results
- ***Relation to network databases***

Identity and Access Management @ Mobile Access - Vodafone

Die Herausforderung

Corporate White / Yellow Pages für die gesamte Vodafone Group, bei Bedarf auch für mobile Mitarbeiter
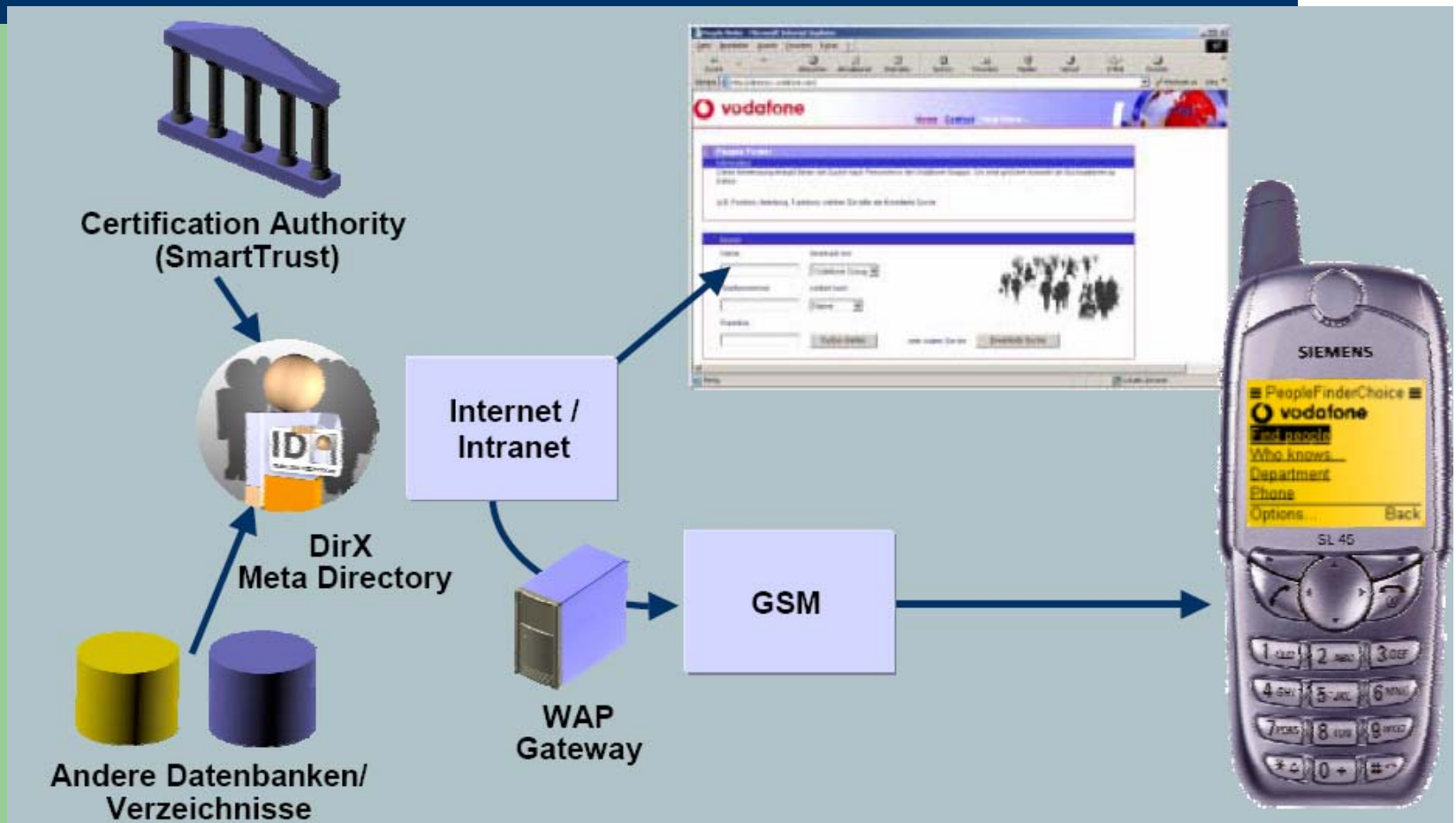
Die Lösung

- DirX Meta Directory für White / Yellow Pages und für die Verwaltung von Public Keys (PKI)

- Die Integration in WAP-Portale ermöglicht mobilen Zugang über ein WAP-fähiges Mobiltelefon

- Integration mit Mobiltelefon-Funktionen, z.B. Anrufen, e-Mail

Der Kundennutzen

- Verbesserte Kommunikationsabläufe:
- Bessere Datenqualität durch:
- Datensynchronisation
- Direkter Zugriff für mobile Benutzer
- Verbesserter Kundenservice durch integriertes Knowledge Management
- Verfügbarkeit von Public Keys für höhere Sicherheit

Identity and Access Management @ Mobile Access - Vodafone

Questions?