1. **Data modeling**

Model is abstraction of real world. Model 作用 P6

Model包括：entity/ relationship/ constraint

Model type：hierarchical modeling(P13)/network modeling (P15)/ relational modeling (P21-22)/E-R(P27)

各种modeling的优缺点P16-18

Relational DBMS: P23. SQL-RDBMS: P25

O-ODB 优缺点: P30 / O-RDB 优缺点P32

Data Model 历年的变化：from hierarchical to NoSQL (P33)

Layered Data abstractions: External Model(P35-36)🡪 Conceptual Model(P37-38)🡪

🡪 Internal Model(P39-40)🡪 Physical Model(P41)

1. **Relationship modeling**

Keys: P4. Full functional dependence(good thing) P6

Entity Integrity & PK type (P8)

Reference Integrity (P9)

Operations on tables (select/project/union/intersection/difference/product/join/divide) P16-24

(P25) Join type: Nature Join(P28-31). / Equijoin / Theta Join

(P26) Join type: Inner Join/ Outer Join: including left join(P32) and right join(P33)

Data dictionaries: P35

1. **E-R diagram**

Attributes定义以及分类： P3-P7

P10: Relationships: Participants / Connectivity / Cardinality(P11-12)

Existence Dependence (P13-14)

Weak vs Strong relationship: P15-17

Weak Entity 的两个条件：exist dependency & strong relationship P18-21

Relationship Participation 包括: optional(P24) / Mandatory(P25)

P26: Relationship Degree: Unary/ Binary / Ternary

Bridge Entity (P29). Developing E-R diagram: P32

1. **Extended E-R(EER)**

Supertype & subtype 含义 P4

Specialization Hierarchy: P5-6. Type: disjoint/ overlapping: P9

Completeness Constraint: P11. Type: Partial completeness/ Total completeness

1. **Normalization --** A process to create clean design

Stop at 3NF, over normalization will slow: P5

Normalization 用处： P6

P15: Dependency type: Partial/ Transitive

0NF🡪1NF: P16-19 remove redundant

1NF🡪2NF: P20-21 No partial dependency

2NF🡪3NF: P22-23 No transitive dependency

Denormalization: P30-31

1. **SQL**
2. **More SQL – views/ triggers/embedded SQL**

Join: Outer/Inner/Full

Subqueries: select (P10) / where (P11) / IN (P14) /HAVING(P15) / ALL/ANY(P18) / FROM (P19)/

Attribute list subqueries P21/ Correlated subqueries P25-27

SQL Function: date and time/ numeric/ string/ conversion

P31-33: Relational set operators: UNION/ INTERSECTION/ DIFFERENCE

Views: create views P34

Procedural languages SQL: block of code / triggers P42-46

Stored procedures P47-49 / Stored functions P50-53

1. **Transaction Management**

ACID+S: P4

Transaction Log: P6

P7：Concurrency control(Problems): 包括：lost update P9-11 / uncommit data P12-14 / inconsistent retrievals P15

Locking method: exclusive lock/ shared lock P17

Lock level: 18-19

P20: type of locks: Binary lock/ exclusive lock/ shared lock(P21)

3 lock states: unlocked/ shared(read) / exclusive(write)

2 Phase-locking (2PL): P22-25

Deadlocks: why occurs (P27-28) / 解决：P29-30 / prevention： P31

Wait/Die and Wound/Wait: P32

1. **Distributed Databases**

影响centralized DB的因素：P3

Distributed and centralized DB compare: P4-6

Distributed 包括： distributed processing(P8)/ distributed databases P9 / distributed fragment

DDBMS including: Transaction processor(TP/TM ) and Data processor(DP/DM)--- P11

User与TM交互， TM与DM交互

3种DDBMS: SDSP (P13)/MPSD (P14) /MPMD (P15-16)

DDBMS restrictions: P18

Distributed DB concurrency control P19

Two Phase commit protocol(2PC). P21

Transparent Features. P26-27

DDBMS Design: Fragmentation (P31)/ Replication (P32) / Allocation (P33)

CAP theorem : consistency/ availability/ partition (P34-35)

1. **Database connectivity**

Options: 1. P12: OOBC (P8) , DAO+Jet (P9) , RDO (P10)

2. OLE-DB (P13-15)

3. ADO.NET (P16-18)

4. Native SQL (P7)

5. JDBC (P19) 优点：P20 example and 架构： P21-22

Web-to-DB: Middleware P24, architecture: P26-27

Client-side-extension: P28

Web application server: uses (P29)/ Features (P30)

1. **Performance tuning**

DBMS architecture -- P5

Optimization: 1. Operator optimize (P10)

2. Time optimize (P11)

3. Information used to optimize P12

P14-15 Query processing: 1. Parsing. P16-17

2. Execution P18

3. Fetching P19

Bottleneck P20

Index P21-23

B-tree Index: P25-26

Index 使用场景：P27

Optimizer: rule-based / cost-based P28

P31 : SQL performance tuning 包括： conditional expressions (p32-33)

DBMS performance tuning: RAID P35-36

1. **Business Intelligence—**Data warehouse

Framework. --P3

BI benefits: --P4

Operational Data vs. Decision Support Data --P6, P9

Decision support data requirement --P8

ETL process (extract🡪 transform 🡪 loading) --P10

Data Marts --P11

Data warehouse rules --P12

P13: Star schema components: facts / dimensions/ attributes/ attributes hierarchy (P14-15)

Snowflake schema --P20

Optimize data warehouse & structure data warehouse --P21-23

Data analytics & tools --P24

Online Analytical processing(OLAP) --P25-27

Relational OLAP. –P28

Multidimensional OLAP. --P29

ROLAP vs MOLAP. -- P30

SQL extension for OLAP. –P31

Data lake. P33