0. 1. Physical Database Design in Relational Databases (1)

- **Factors that Influence Physical Database Design:**
  A. Analyzing the database queries and transactions
- **For each query, the following information is needed.**
  1. **The *files* that will be accessed by the query;**
  2. **The *attributes* on which any *selection* conditions for the query are specified;**
  3. **The *attributes* on which any *join* conditions or conditions to link multiple tables or objects for the query are specified;**
  4. **The *attributes* whose values will be *retrieved* by the query.**
- **Note: the attributes listed in items 2 and 3 above are candidates for definition of access structures.**

- **العوامل التي تؤثر علي تصميم قاعده البيانات المادية:**
  أ- تحليل الاستعلامات والمعاملات الخاصة بقاعده البيانات
- **لكل استعلام ، المعلومات التالية مطلوبه.**
  1. الملفات التي سيتم الوصول اليها من قبل الاستعلام ؛
  2. -2السمات التي تحدد عليها شروط الاختيار للاستعلام ؛
  3. السمات علي اي شروط أو شروط الانضمام لربط جداول متعددة أو كائنات للاستعلام محدده ؛
  4. السمات التي سيتم استرداد القيم من قبل الاستعلام.
- **ملاحظه:** السمات المذكورة في البندين 2 و 3 أعلاه هي المرشحين لتعريف هياكل الوصول.

Physical Database Design in Relational Databases (2)

- Factors that Influence Physical Database Design (cont.):
  A. **Analyzing the database queries and transactions** (cont.)
  - For each **update** transaction or operation, the following information is needed.
    1. The files that will be updated;
    2. The type of operation on each file (insert, update or delete);
    3. The attributes on which selection conditions for a delete or update operation are specified;
    4. The attributes whose values will be changed by an update operation.
  - **Note:** the attributes listed in items 3 above are candidates for definition of access structures. However, the attributes listed in item 4 are candidates for avoiding an access structure.

- **العوامل التي تؤثر علي تصميم قاعده البيانات المادية (تابع):**
  ✓ تحليل الاستفسارات والمعاملات الخاصة بقاعده البيانات (تابع)
- **بالنسبة لكل حركه أو عمليه تحديث ، تكون المعلومات التالية مطلوبه.**
  1. الملفات التي سيتم تحديثها;
  2. نوع العملية علي كل ملف (ادراج أو تحديث أو حذف);
  3. السمات التي يتم تحديد شروط التحديد لها لعمليه حذف أو تحديث;
  4. السمات التي سيتم تغيير قيمها بواسطة عمليه تحديث.

ملاحظه: السمات المسرودة في العناصر 3 أعلاه هي المرشحين لتعريف هياكل الوصول. ومع ذلك ، السمات المسرودة في البند 4 هي المرشحين لتجنب بنيه الوصول.

Physical Database Design in Relational Databases (3)

- Factors that Influence Physical Database Design (cont.):

B. **Analyzing the expected frequency of invocation of queries and transactions**

- The expected frequency information, along with the attribute information collected on each query and transaction, is used to compile a cumulative list of expected frequency of use for all the queries and transactions.
- It is expressed as the expected frequency of using each attribute in each file as a selection attribute or join attribute, over all the queries and transactions.
- **80-20 rule**
  - 20% of the data is accessed 80% of the time

- العوامل التي تؤثر علي تصميم قاعده البيانات المادية (تابع):

Physical Database Design in Relational Databases (4)

- Factors that Influence Physical Database Design (cont.)

C. **Analyzing the time constraints of queries and transactions**

- Performance constraints place further priorities on the attributes that are candidates for access paths.
- The selection attributes used by queries and transactions with time constraints become higher-priority candidates for primary access structure.

- Factors that Influence Physical Database Design (cont.)

D. **Analyzing the expected frequencies of update operations**

- A minimum number of access paths should be specified for a file that is updated frequently.

- العوامل التي تؤثر علي تصميم قاعده البيانات المادية (تابع) جيم-تحليل القيود الزمنيه للاستعلامات والمعاملات
- وتضع قيود الأداء أولويات اضافيه علي السمات المرشحة لمسارات الوصول.
- تصبح سمات التحديد المستخدمة من قبل الاستعلامات والحركات مع قيود الوقت المرشحين ذوي الاولويه الأعلى للبنية الاساسيه للوصول
- العوامل التي تؤثر علي تصميم قاعده البيانات المادية (تابع) دال-تحليل الترددات المتوقعة لعمليات التحديث
- يجب تحديد عدد ادني من مسارات الوصول للملف الذي يتم تحديثه بشكل متكرر.
- .

- Factors that Influence Physical Database Design (cont.)

E. **Analyzing the uniqueness constraints on attributes**

- Access paths should be specified on all candidate key attributes — or set of attributes — that are either the primary key or constrained to be unique

- العوامل التي تؤثر علي تصميم قاعده البيانات المادية (تابع)

ه-تحليل قيود التفرد علي السمات

- يجب تحديد مسارات الوصول علي كافة سمات المفتاح المرشح-أو مجموعه من السمات-التي تكون اما المفتاح الأساسي أو مقيده لتكون فريدة.

Physical Database Design in Relational Databases (5)

- **Physical Database Design Decisions**
- **Design decisions about indexing**
  - ✓ Whether to index an attribute?
  - ✓ What attribute or attributes to index on?
  - ✓ Whether to set up a clustered index?
  - ✓ Whether to use a hash index over a tree index?
  - ✓ Whether to use dynamic hashing for the file?

Physical Database Design in Relational Databases (6)

- **قرارات تصميم قاعده البيانات المادية**
- **تصميم القرارات حول الفهرسة**
  - ✓ ما إذا كان سيتم فهرسه سمه ؟
  - ✓ ما سمه أو سمات لفهرسه علي ؟
  - ✓ ما إذا كان سيتم اعداد فهرس متفاوت المسافات ؟
  - ✓ ما إذا كان سيتم استخدام فهرس تجزئه فوق فهرس شجره ؟
  - ✓ ما إذا كنت تريد استخدام التجزئة الديناميكية للملف ؟

- **Physical Database Design Decisions (cont.)**
- **Denormalization as a design decision for speeding up queries**
  - The goal of normalization is to separate the logically related attributes into tables to minimize redundancy and thereby avoid the update anomalies that cause an extra processing overheard to maintain consistency of the database.
  - The goal of denormalization is to improve the performance of frequently occurring queries and transactions. (Typically the designer adds to a table attributes that are needed for answering queries or producing reports so that a join with another table is avoided.)
  - Trade off between update and query performance

- **التجريد من المال كقرار تصميم لتسريع الاستعلامات**
- الهدف من التطبيع هو فصل السمات المرتبطة منطقيا إلى الجداول لتقليل التكرار التالي تجنب الشذوذ التحديث الذي يسبب المعالجة الاضافيه سمع للمحافظة علي تناسق قاعده البيانات.
- والهدف من التجريد من الدين هو تحسين أداء الاستفسارات والمعاملات التي يكثر حدوثها. (عاده ما يضيف المصمم إلى سمات الجدول المطلوبة للاجابه عن الاستعلامات أو إنتاج التقارير بحيث يتم تجنب صله مع جدول آخر.)
- التداول بين التحديث وأداء الاستعلام

---

2. An Overview of Database Tuning in Relational Systems (1)

- **Tuning**:
  - The process of continuing to revise/adjust the physical database design by monitoring resource utilization as well as internal DBMS processing to reveal bottlenecks such as contention for the same data or devices.
- **Goal:**
  - To make application run faster
  - To lower the response time of queries/transactions
  - To improve the overall throughput of transactions

- ضبط:
- عمليه الاستمرار في تنقيح/تعديل تصميم قاعده البيانات المادية عن طريق رصد استخدام الموارد وكذلك معالجه نظم المعلومات الداخلية للكشف عن اختناقات مثل الخلاف لنفس البيانات أو الاجهزه.

- هدف:
- لجعل تشغيل التطبيق أسرع
- لتقليل وقت الاستجابة للاستعلامات/المعاملات
- لتحسين الانتاجيه الاجماليه للمعاملات

An Overview of Database Tuning in Relational Systems (2)

- **Statistics internally collected in DBMSs:**
  - Size of individual tables
  - Number of distinct values in a column
  - The number of times a particular query or transaction is submitted/executed in an interval of time
  - The times required for different phases of query and transaction processing
- **Statistics obtained from monitoring:**
  - Storage statistics
  - I/O and device performance statistics
  - Query/transaction processing statistics
  - Locking/logging related statistics
  - Index statistic

- **إحصاءات المجمعة داخليا في الاداره الاتحادية للإحصاء:**
- حجم الجداول الفردية
- عدد القيم المميزة في عمود
- عدد المرات التي يتم فيها إرسال/تنفيذ استعلام أو معامله معينه في فتره زمنيه
- الأوقات المطلوبة لمراحل مختلفه من الاستعلام ومعالجه المعاملات
- **الإحصاءات التي تم الحصول عليها من الرصد:**
- إحصاءات التخزين
- الإدخال/الإخراج وإحصاءات أداء الجهاز
- الاستعلام/إحصائيات معالجه المعاملات
- تامين/تسجيل الإحصاءات ذات الصلة
- احصائيه مؤشر

An Overview of Database Tuning in Relational Systems (3)

- **Problems to be considered in tuning:**
  - How to avoid excessive lock contention?
  - How to minimize overhead of logging and unnecessary dumping of data?
  - How to optimize buffer size and scheduling of processes?
  - How to allocate resources such as disks, RAM and processes for most efficient utilization?

المشاكل التي يجب مراعاتها في التوليف:

- كيفيه تجنب الإفراط في قفل الخلاف ؟
- كيفيه تقليل النفقات العامة للتسجيل وإلقاء البيانات غير الضرورية ؟
- كيفيه تحسين حجم المخزن المؤقت وجدوله العمليات ؟
- كيفيه تخصيص الموارد مثل الاقراص وذاكره الوصول العشوائي والعمليات للاستخدام الأكثر كفاءه ؟

An Overview of Database Tuning in Relational Systems (4)

- **Tuning Indexes**
  - Reasons to tuning indexes
    - Certain queries may take too long to run for lack of an index;
    - Certain indexes may not get utilized at all;
    - Certain indexes may be causing excessive overhead because the index is on an attribute that undergoes frequent changes
  - **Options to tuning indexes**
    - Drop or/and build new indexes
    - Change a non-clustered index to a clustered index (and vice versa)
    - Rebuilding the index

- **توليف الفهارس**
- أسباب لضبط الفهارس
- قد تستغرق بعض الاستعلامات وقتا طويلا ليتم تشغيلها بسبب عدم وجود فهرس;
- قد لا يتم استخدام بعض الفهارس علي الإطلاق;
- قد تتسبب بعض الفهارس في الحمل الزائد لان الفهرس موجود علي سمه تمر بتغييرات متكررة

- **خيارات لضبط الفهارس**
- إسقاط أو/وبناء فهارس جديده
- تغيير فهرس غير متفاوت المسافات إلى فهرس متفاوت المسافات (والعكس)
- أعاده بناء الفهرس

An Overview of Database Tuning in Relational Systems (5)

- **Tuning the Database Design**
  - Dynamically changed processing requirements need to be addressed by making changes to the conceptual schema if necessary and to reflect those changes into the logical schema and physical design.

- ضبط تصميم قاعده البيانات
- ديناميكيا تغيير متطلبات المعالجة تحتاج إلى ان تعالج عن طريق اجراء تغييرات علي المخطط المفاهيمي إذا لزم الأمر ، وتعكس تلك التغييرات في المخطط المنطقي والتصميم المادي.

An Overview of Database Tuning in Relational Systems (6)

- **Tuning the Database Design (cont.)**
  - **Possible changes to the database design**
    - Existing tables may be joined (denormalized) because certain attributes from two or more tables are frequently needed together.
    - For the given set of tables, there may be alternative design choices, all of which achieve 3NF or BCNF. One may be replaced by the other.
    - A relation of the form R(K, A, B, C, D, …) that is in BCNF can be stored into multiple tables that are also in BCNF by replicating the key K in each table.
    - Attribute(s) from one table may be repeated in another even though this creates redundancy and potential anomalies.
    - Apply **horizontal partitioning** as well as **vertical partitioning** if necessary.

- **ضبط تصميم قاعده البيانات (تابع)**
- **التغييرات المحتملة في تصميم قاعده البيانات**
- قد يتم ربط الجداول الموجودة (التي تمت الاضافه اليها) لان هناك حاجه متكررة مع بعض السمات من جدولين أو أكثر.
- النسبة لمجموعه الجداول المعينة ، قد تكون هناك خيارات بديله للتصميم ، وكلها تحقق الهدف 3 أو الصندوق الوطني للمراه. ويمكن الاستعاضة عن أحدهما بالأخر.
- ويمكن تخزين العلاقة بين النموذج R (k ، A، B، C، D، (...الموجود في الصندوق الوطني للمراه في جداول متعددة موجودة أيضا في المكتب الوطني للمراه بتكرار المفتاح K في كل جدول.
- السمة , من جدول واحد قد تتكرر في آخر حتى وان كان هذا يخلق التكرار والشذوذ المحتملة.
- تطبيق التقسيم الأفقي وكذلك التقسيم الراسي إذا لزم الأمر.

An Overview of Database Tuning in Relational Systems (7)

- **Tuning Queries**
  - **Indications for tuning queries**
    - A query issues too many disk accesses
    - The query plan shows that relevant indexes are not being used.

- **توليف الاستعلامات**
- **مؤشرات لتوليف الاستعلامات**
- المسائل الاستعلام عدد كبير جدا من الوصول إلى القرص
- تظهر خطه الاستعلام انه لا يتم استخدام الفهارس ذات الصلة.

An Overview of Database Tuning in Relational Systems (8)

- **Tuning Queries (cont.): Typical instances for query tuning**
  - In some situations involving using of correlated queries, temporaries are useful.
  - If multiple options for join condition are possible, choose one that uses a clustering index and avoid those that contain string comparisons.
  - The order of tables in the FROM clause may affect the join processing.
  - Some query optimizers perform worse on nested queries compared to their equivalent un-nested counterparts.

- **توليف الاستعلامات (تابع): المثيلات النموذجية لتوليف الاستعلام**
- وفي بعض الحالات التي تنطوي علي استخدام الاستفسارات المترابطة ، تكون المؤقتات مفيده.
- إذا كانت هناك خيارات متعددة لشرط الصلة ممكنة ، اختر أحد يستخدم فهرس تجميع وتجنب تلك التي تحتوي علي مقارنات السلاسل.
- ترتيب الجداول في جمله from قد يؤثر علي معالجه الصلة.
- بعض أمثليه الاستعلام أداء اسوا علي الاستعلامات المتداخلة بالمقارنة مع نظيراتها المكافئة للأمم المتحدة المتداخلة.
- تستند العديد من التطبيقات إلى طرق العرض التي تحدد البيانات التي تهم تلك التطبيقات. أحيانا تصبح هذه الآراء مبالغه.

An Overview of Database Tuning in Relational Systems (10)

- **Additional Query Tuning Guidelines**
  - A query with multiple selection conditions that are connected via OR may not be prompting the query optimizer to use any index. Such a query may be split up and expressed as a union of queries, each with a condition on an attribute that causes an index to be used.
  - Apply the following transformations
    - NOT condition may be transformed into a positive expression.
    - Embedded SELECT blocks may be replaced by joins.
    - If an equality join is set up between two tables, the range predicate on the joining attribute set up in one table may be repeated for the other table
  - WHERE conditions may be rewritten to utilize the indexes on multiple columns.

- إرشادات اضافيه لتوليف الاستعلام
- استعلام مع شروط التحديد المتعددة المتصلة عبر أو قد لا يكون المطالبة محسن الاستعلام لاستخدام اي فهرس. قد يتم تقسيم مثل هذا الاستعلام ويتم التعبير عنه كاتحاد من الاستعلامات ، لكل منها شرط علي سمه تؤدي إلى استخدام فهرس.
- تطبيق التحويلات التالية
- ولا يجوز تحويل الشرط إلى تعبير إيجابي.
- قد يتم استبدال كتل SELECT المضمنة بالصلات.
- إذا تم اعداد صله المساواة بين جدولين ، فقد يتم تكرار داله التقييم النطاق علي السمة join التي تم اعدادها في جدول واحد للجدول الآخر
- (where) قد تتم أعاده كتابه الشروط لاستخدام الفهارس علي أعمده متعددة.

Summary

- **Physical Database Design in Relational Databases**
- **Database Tuning in Relational Systems**

- تصميم قاعده البيانات المادية في قواعد البيانات العلائقية
- ضبط قاعده البيانات في النظم العلائقية