إدارة قواعد البيانات الموزّعة

د. حمدي داود سلمان
جامعة الفاتح

المقدمة.

شهدت العقود الأخيرة نموا هائلا في تطبيقات قواعد البيانات واستخداماتها في تخزين البيانات، ومعالجتها، واسترجاعها في مختلف المنظمات تقريبا بما في ذلك الأعمال، والرعَاية الصحية، والتعليم، والأعمال الحكومية، والمكتبات، واستخدم الأفراد قواعد البيانات على الحاسبات الشخصية، كما استخدمتها مجموعات العمل التي تتصل بقواعد البيانات الموجودة على خادم الشبكات واستخدمها كل العاملين المستخدمين للتطبيقات الموزّعة على مستويات الإدارة المختلفة.

وبعد هذه الفترة من النموَ السريع، صار من المتوقع ازدياد الطلب على تقنيات قواعد البيانات، وكل الأدلة تشير إلى أن هذا الطلب سوف يحظى بأهمية متزايدة.

إن استخدام قواعد البيانات DB في المنظمات يعني وجود بياناتها في صورة مهيكلة ومتكاملة ومتعلقة ببعضها مما يسهل على مستخدمي تلك القواعد في تخزين واسترجاع ومعالجة البيانات، كما أن تلك البيانات تخدم وتصف دوائر العمل ومناطقه والعلاقات التي تربط بينهما بحيث تحقق التكامل المطلوب في استخدام البيانات، وعدم التكرار والتضارب بين البيانات المتاحة لكل نطاق أو دائرة، وإن هؤلاء المستخدمين يمكنهم أن يستخدموا البيانات في الإجابة على أسئلة خاصَة بالنطاق وتكون المعلومات

بيانات وضعت في إطار أو أجري عليها تشغيل وتم تقديمها فـــي صـــورة مناسبة للتفسير البشري [1].

لأي قاعدة بيانات نظام إدارة DBMS ، وتتعدد نظم الإدارة وتتميز فيما بينها من عدة نواحي ومعايير ومنتجين. وصار التنافس بين الـــشركات المنتجة كبيرا لغرض كسب أكبر عدد من المستخدمين مما أدى إلى تطورا واسع في هذا المجال وبروز أنظمة حديثة تـــضم العديـــد مـــن الميـــزات والتسهيلات التي ساعدت في انتشار استخدام أنظمة قواعد البيانات خاصـــة مع السهولة المتزايدة في الاستخدام حتـــى صـــارت فـــي متنـــاول غيـــر المتخصصين [2].

ويمكن أن توَصف تطبيقات قواعد البيانات، فـــي قواعـــد بيانـــات للحاسب الشخصي وقواعد بيانات مجموعة العمل،وقواعد بيانات الأقـــسام، وقواعد بيانات المنظمة أو المنشأة، وهي الأكثر أهميه في وقتنـــا الحاضـــر لأنها قاعدة بيانات متكاملة لدعم القرار وتستخلص محتوياتها مـــن قواعـــد بيانات تشغيل متنوعة.

وبالرغم من هذه التطورات إلا أن الزيادة في اســـتخدام نظـــم إدارة قواعد البيانات وازدياد الاعتماد علها في تطبيقات واسعة ومختلفـــة، أبـــرز وجود عدد من التقنيات ساعدت في توفير اتصال أســـرع، وأســـهل لكـــل المستخدمين، وهي قواعد البيانات العلائقية الـــشيئية (اســـتعمال الـــصوت والصور الثابتة والمتحركة) لتحقيق هذه الأهداف.

إدارة قواعد البيانات الموزَعة (DDBMS) وهي عبارة عن قاعـــدة بيانات متكاملة لكنها موزعة إلى أقسام ومخزنة على حاسبة واحدة أو عـــدة حاسبات، وتعمل تحت سيطرة DBMS، وأن هذه الحواسب مربوطة بشبكة Network.

---

[1] الترا جيفري، بيرسكون ب مارى، مكفادن آر فريد; تعريب : سرور علي إبراهيم سرور: إدارة قواعد البيانات الحديثة الرياض : مكتبة دار المريخ، 2003م ص ص 72 - 83

[2] نبيل عبد الله قمصاني "الاتجاهات السلوكية لمستخدمي قواعد المعلومات والمنتجين لها"، القاهرة عالم الكتب مج 21، 6 ، 2000م ص 36.

عندما تكون المنظمة موزّعة يمكن أن تخزن قواعد بياناتها على حاسب مركزي أو توزعها على حاسبات مركزية (أو خليط من الإثنين) وتكون قاعدة البيانات الموزّعة Distributed Databases قاعدة بيانات منطقية فردية موزّعة طبيعيا عبر الحاسبات الموجودة في المواقع المتعددة، وترتبط بشبكة اتصالات بيانات. وتحتاج قاعدة البيانات الموزّعة إلى نظم إدارة قواعد بيانات متعددة، وتعمل على كل المواقع البعيدة. وتساهم درجة تعاون نظم DBMSs ، وعملها مع بعضها البعض، أو عن طريق موقع رئيسي ينسق الطلبات التي تشمل بيانات من المواقع المتعدّدة، في تمييز الأنواع المختلفة من بيئات قواعد البيانات الموزّعة.

وتشجع ظروف أعمال مختلفة على استخدام قاعدة البيانات الموزّعة للأسباب التالية [1]:

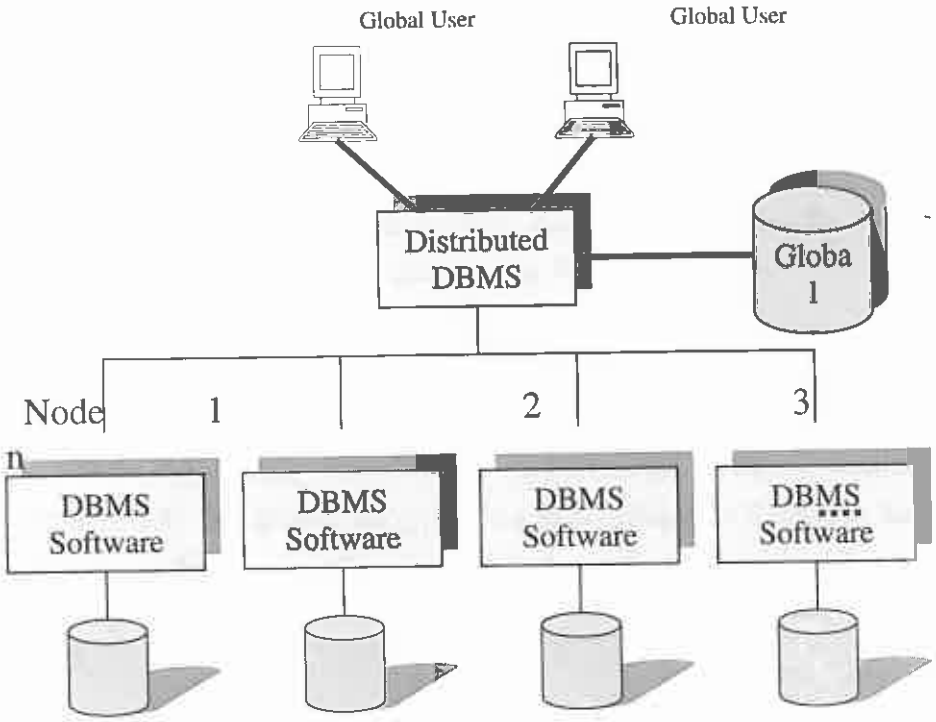
1. انتشار واستقلال وحدات الأعمال distribution and autonomy of business عادة ما تكون أجزاء وأقسام وتسهيلات المنظمات الحديثة موزّعة جغرافيا (وربما عالميا). وعادة يكون لكل وحدة سلطة أنتاج نظم معلوماتها الخاصة بها وعادة ما تريد أن تتحكم هذه الوحدات في بياناتها محلية.

2. اقتسام البيانات، Data sharing حتى قرارات الأعمال متواضعة التعقيد تحتاج إلى اقتسام البيانات عبر وحدات الأعمال لذلك يجب أن يكون دمج البيانات عبر قواعد البيانات المحلية مريحا عند الطلب.

3. تكاليف اتصالات البيانات ومرونتها data communication cost and reliability يمكن أن تكون تكاليف إرسال كميات بيانات كبيرة عبر شبكة الاتصالات أو التعامل مع حجم كبير للعمليات الجارية من مواقع بعيدة مرتفعة التكاليف. و اقتصادية أكثر، في العادة تحدد موقع البيانات والتطبيقات لتكون قريبة من مواقع الحاجة لها. كما أن الاعتماد على الاتصالات بالبيانات يمكن أن يكون محفوفا بالمخاطر. لذلك يكون الاحتفاظ

[1] مرجع سابق ، إدارة قواعد البيانات الحديثة الرياض : مكتبة دار المريخ، 2003 ص 143

بنسخ محلية أو أجزاء من بيانات طريقة فيها مرونة وثقه، لدعم الحاجة إلى الاتصال السريع بالبيانات عبر المنظمة.



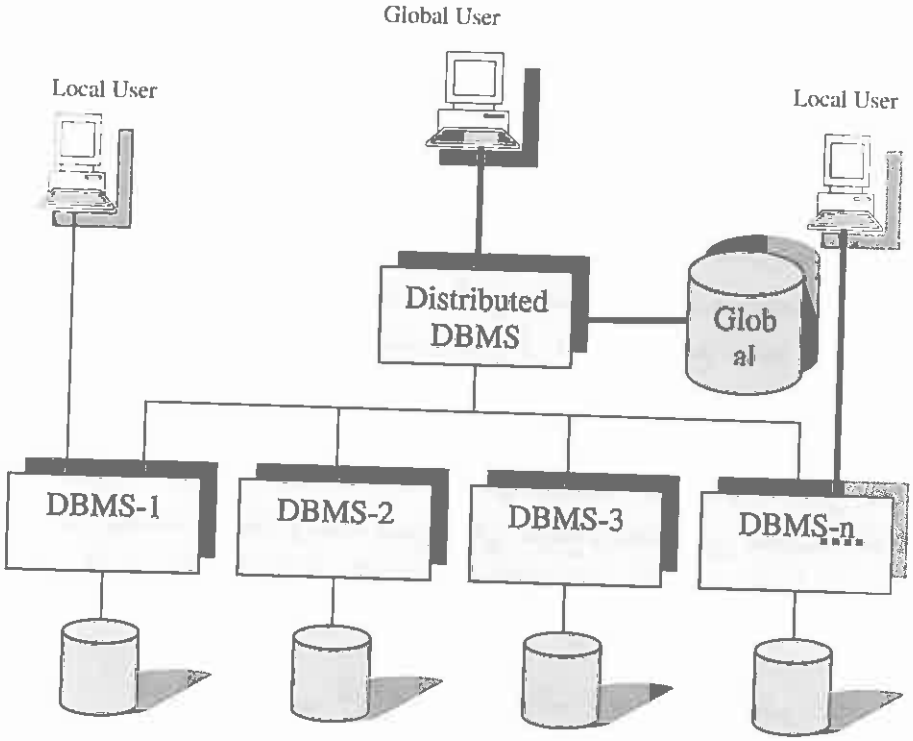
شكل (1) بيئة قاعدة البيانات موزّعة متجانسة[1]

ويمثل شكل(1) بيئة قاعدة بيانات موزّعة متجانسة. وتعرف هذه البيئـــة تقليديا بالخواص التالية:-

- توزيع البيانات على كل العقد.
- يستخدم نفس النظام في كل موقع.
- يدير النظام المنتشر كل البيانات (لذلك لا يوجد أي موقـــع غيـــر مشمول).

[1] Bell .D. , and Grimson .1992 **Distributed Database System** : Addison – Wesley pp 24 - 39

- يتصل كل المستخدمين بقاعدة البيانات من خلال مخطط شامل واحد أو تعريف قاعدة بيانات واحدة شاملة.
- يكون المخطط الشامل اتحاد كل مخططات قواعد البيانات المحليـة ببساطة. ومن الصعب فرض بيئة متجانسة في معظم المنظمات كما أن إدارة البيئات غير متجانسة تكون أكثر صعوبة.



شكل (2) بيئة قاعدة البيانات موزّعة غير متجانسة[1]

وكما هو مبين في أعلاه توجد الكثير من البيئات الموزّعة غير متجانسة بالسمات التالية كما يمثلها شكل (2):

- تكون البيانات موزعة على كل عقد.
- يمكن أن تستخدم نظم DBMSs مختلفة عند كل عقدة.

[1] Bell, D., and Grimson , 1992 Distributed Database System , Addison- Wesley

- يحتاج بعض المستخدمين إلى اتصال محلي فقط بقاعدة البيانات والذي يمكن أن يتحقق باستخدام نظام DBMS ومخطط محليين فقط.
- يوجد مخطط شامل والذي يسمح للمستخدمين المحليين بالاتصال بالبيانات البعيدة.

أهداف قواعد البيانات الموزّعة The Objective of Distributed Database

أحد الأهداف الرئيسية لقاعدة البيانات الموزّعة هو توفير اتصال سهل للمستخدمين الموجودين في مواقع مختلفة بالبيانات. ولتحقيق هذا الهدف يجب أن يوفر النظام قاعدة بيانات الموزّعة، ما يسمى بشفافية الموقع location transparency ، والذي يعني أن المستخدم (أو برنامج المستخدم) الذي يستخدم البيانات يحتاج أن يعرف موقع تواجد هذه البيانات. ويقوم النظام بتمرير أي طلب لاسترجاع البيانات، أو تجديدها من أي موقع تلقائيا إلى الموقع نفسه، أو المواقع المرتبطة بالتشغيل المطلوب. ولا يهتم المستخدم، بتوزيع البيانات، وتظهر كل البيانات الموجودة في الشبكة كقاعدة بيانات منطقية واحدة مخزنة في موقع واحد. وفي الحالة المثالية يمكن أن يرتبط الاستفسار الفردي بالبيانات من جدأول موجودة في مواقع متعددة؛ كما لو كانت البيانات كلها موجودة في موقع واحد.

الهدف الثاني لقواعد البيانات الموزّعة هي الاستقلال المحليLocal autonomy والذي يعني أن الموقع يمكن أن يدير، ويجري تشغيل لقاعدة بياناته بصورة مستقلة عندما تفشل الروابط مع العقد الآخرى[1]. ومع الاستقلال المحلي يكون لكل موقع المقدرة على مراقبة البيانات المحلية وإدارة الأمن، وتسجيل العمليات الجارية.

وبالمقارنة مع قواعد البيانات المركزية، توجد العديد من المميّزات لأي صيغة من صيغ قاعدة البيانات الموزّعة، وفيما يلي أهمها:

---

1 Date, C.j. 1995 An Introduction to Database System ,6th ed Addison-Wesley

1. زيادة المرونة والثقة والإتاحة increased reliability and availability، عندما تفشل النظم المركزية لا تصبح قاعدة البيانات متاحة لكل المستخدمين. ويستمر النظام المنتشر في العمل على مستوى منخفض بعض الشيء عندما تفشل مكوناته.وتعتمد المرونة والثقة والإتاحة (من ضمن أشياء آخرى) على كيفية توزيع البيانات.

2. المراقبة المحليةlocal control ، يشجع توزيع البيانات المجموعة المحلية على ممارسة مراقبة أكبر على بياناتهم، والذي يروج تحسين سلامة البيانات، وإدارتها. ويستطيع المستخدمون، في نفس الوقت الاتصال بالبيانات غير محلية عندما يلزم ذلك. ويمكن اختيار نظم المكونات للموقع المحلي لتتفق مع عمل تشغيل البيانات المحلي وليس الشامل.

3. النمو على صورة مقاطع modular growth ، افرض أن المنظمة تتوسع في موقع جديد، أو تضيف مجموعة عمل جديدة. من الأسهل أو الاقتصادي أكثر في العادة إضافة حاسب محلي، وأن إضافة البيانات المصاحبة إلى الشبكة الموزّعة أسهل من التوسع في حاسب مركزي كبير. كما تكون هناك فرص اقل أيضا لإزعاج المستخدمين الموجودين عما يكون الحال عليه عندما يتم تعديل الحاسب المركزي أو توسيعه.

4. تكاليف اتصالات أقلlower communication costs ، مع النظام المنتشر يمكن وضع بيانات أقرب إلى النقطة المراد استخدامها ويمكن أن يقلل هذا تكاليف الاتصالات مقارنة مع النظام المركزي.

5. استجابة أسرع faster response، اعتمادا على كيفية توزيع البيانات يمكن تحقيق معظم طلبات بيانات من المستخدمين الموجودين في موقع معين عن طريق البيانات المخزنة في هذا الموقع.ويتسبب هذا في الإسراع بعملية تشغيل الاستفسار نظراً لانخفاض الاتصالات، وتأخير الحاسب المركزي. كما يمكن أيضا

تقسيم التفسيرات المعقدة إلى استفسارات فرعية يمكن تشغيلها علـــى التوازي في مواقع متعددة يوفر استجابة أكثر سرعة.

تواجه قاعدة البيانات الموزّعة بعض التكاليف والعيوب أيضا:[1]

1. تكلفـــة نظـــم البـــرامج وتعقيـــدها software cost and complexity، تلزم نظم بـــرامج أكثـــر تعقيـــدا (خاصـــة نظـــام DBMS) لبيئة قاعدة البيانات الموزّعة.

2. مصاريف إضافية للتشغيل processing overhead ، يجب أن تتبادل المواقع المختلفة بالرسائل، وتنفذ حسابات إضـــافية لـــضمان التنسيق المناسب عبر البيانات عند المواقع المختلفة.

3. سلامة البياناتdata integrity ، سوف يتعرض المنتج الثـــانوي(العقد node) إلى زيادة في التعقيد، وفي حالة التجديد،سيتطلب ذلك الحاجة إلى التنسيق غير المناسب والمشاكل الأخـــرى الخاصـــة بسلامة البيانات.

4. استجابة بطيئة show response ، إذا لم تكن البيانات موزعة بصورة مناسبة طبقا لاستخدامها، أو إذا كانت الاستفـــسارات غيـــر مشكلة بصورة صحيحة، يمكن أن تكون الاستجابة علـــى البيانـــات المطلوبة بطيئة للغاية.

خيارات نشر قواعد البيانات

كيف يجب أن تنتشر (توُزع) قاعدة البيانات عبر المواقع (أو عُقد) الشبكة ويلاحظ وجود أربع استراتيجيات لتوزيع أو نشر قواعد البيانات [2]:

---

[1] özsu, M. T., and p. Valduriez. 1992."**Distributed Database System**:Where Wre We" Database Programming &Disgn 5 (April): pp 49-55

[2] Barquin , R . 1996. ' On the First Issue of The Journal of Data Warehousing . ' The Journal of Data Warehousing 1 (July)

1-تكرار البيانات.
2- التجزئة الأفقية.
3- التجزئة الراسية.
4-مجموعة خليط من مما سبق.

وسوف يتم التعرض لكل هذه الاستراتيجيات ونوضحها باستخدام قواعد البيانات العلاقية وتسري نفس المفاهيم (مع بعض التعديلات) على نماذج بيانات الأخرى مثل الهرمية والشبكية.

وللإيضاح نفترض أن أحد البنوك له العديد من الفروع الموزّعة في المحافظات ولبلد مثل العراق، علما أن إحدى العلاقات الأساسية في قاعدة بيانات البنك هي علاقة العميل Customer، و يبين الجدول (1) الصيغة المختصرة لهذه العلاقة، وللتبسيط، تطبق عيّنة البيانات الموجودة في العلاقة على فرعين اثنين فقط (محافظة بغداد ومحافظة البصرة) والمفتاح الأول في هذه العلاقة هو رقم الحساب (Acc_Number) والمفتاح الثاني يكون اسم الفرع Branch_Name الذي فتّح لديه العملاء حساباتهم (وبالتالي فإنهم ينفذون معظم عملياتهم الجارية هناك).

جدول رقم (1)

علاقة العميل Customer للبنك

Acc_Number	Customer_Name	Branch_Name	Balance
200	Hamdi	Baghdad	1000
302	Samer	Basrah	250
260	Aws	Basrah	38
482	Osama	Baghdad	796
621	Ali	Basrah	168
310	Salman	Baghdad	1500
222	Dawood	Baghdad	330

1. تكرار البيانات Data Replication

الخيار المتزايد الشعبية هو توزيع البيانات و تخزين نسخة مستقلة من قاعدة البيانات عند كل موقع، اثنين، أو أكثر. ويمكن أن تسمح عملية التكرار المنظمة بنقل قاعدة البيانات من الحاسب الكبير المركزي إلى خدَم أقسام اقل تكلفة، أو خدم مواقع محددة، قريبة من المستخدمين النهائيين (كوب Koop)[1]. ويمكن أن يستخدم التكرار تقنيات قاعدة بيانات متزامنة، أو غير متزامنة؛ رغم أن التقنيات غير المتزامنة التقليدية أكثر في البيئة المتكررة. ويمكن أن تخزن علاقة Customer المبيّنة في الجدول أعلاه في كلا من محافظة بغداد و محافظة البصرة، على سبيل المثال. فإذا خزنت نسخة في كل موقع يصبح لدينا حالة التكرار الكامل (والتي يمكن ألا تكون عملية باستثناء إذا كانت قواعد البيانات صغيرة نسبيا).

وتوجد خمس مميزات لتكرار البيانات:

أ- والوثوق والمرونة Reliability ، إذا فشل احد المواقع المحتوي على العلاقة (أو قاعدة البيانات)، يمكن أن توجد نسخة دائما في موقع آخر دون أي تأخيرا في مرور الشبكة. كما أن النسخة المتاحة يمكن تجديدها كلها بمجرد حدوث العمليات الجارية وتجدّد العقد غير المتاحة بمجرد عودتها للخدمة.

ب- استجابة سريعة Fast Response، يستطيع كل موقع توجد فيه النسخة كاملة أن يجري تشغيلا للاستفسارات محليا، ويتسبب هذا في تشغيل الاستفسارات بسرعة.

ج- إمكانية تجنب مقاطع سلامة العمليات الجارية الموزّعة المعقدة Possible Avoidance of Complicated Distributed Transaction Integrity Routines عادة تنشط قواعد البيانات المتكرّرة على فترات زمنية مجدولة، لذلك تستخدم معظم صيغ التكرار عندما تكون بعض العقد غير مشغولة في التزامن، وبالتالي من الممكن نسخ قاعدة البيانات.

[1] Koop , P .1995. "Replication at Work" . DBMS 8 (March) : 54-60

د- إلغاء ازدواج العقدة Node Decoupling، يمكن أن تستمر كل عملية جارية دون تنسيق عبر الشبكة. لذلك، إذا كانت العقد لا تعمل أو مشغولة أو غير موصلة (مثال ذلك حاله الحاسبات الشخصية المحمولة) يتم التعامل مع العملية الجارية عندما يرغب المستخدم في ذلك. وفي حالة تزامن الوقت الحقيقي للتجديدات تنسق عملية تقع خلف الشاشة (وراء الستار) لكي يتم نسخ بيانات.

ه- تقليل مرور الشبكة عند الوقت الأولي Reduce Network Traffic at Prime Time يحدث تجديد البيانات أثناء ساعات الأعمال الأولية عندما يكون المرور للشبكة أعلى ما يمكن، ويكون الطلب على الاستجابة السريعة أكبر ما يمكن. ويتم النسخ المتأخر عندما يقل تكرار المرور للشبكة لإرسال التجديدات إلى عقد آخرى في غير ساعات الوقت الأولية.

وللتكرار عيبين أساسيين :

- متطلبات التخزين storage requirements، يجب أن يكون لكل موقع نسخة كاملة من DBMS، لها نفس سعة التخزين اللازمة إذا كانت البيانات مخزنة مركزيا تماما. وتتطلب كل نسخة مكان تخزين (تنخفض تكلفته بصفة مستمرة) ويلزم وقت تشغيل لتجديد كل نسخة موجودة على كل عقدة.

- تعقيد التجديد، وتكلفته complexity and cost of updating، كلما تم تجديد العلاقة، فيجب أن تجدّد عند كل موقع يمسك بإحدى النسخ ويمكن أن يطلب التجديد المتزامن بالقرب من الوقت الحقيقي والذي تكثر فيها حركة الشبكة، كما سيتضح ذلك فيما بعد تحت موضوع بروتوكول الالتزام.

ولهذه الأسباب يفضل تكرار البيانات عندما تكون معظم طلبات التشغيل للقراءة فقط، وعندما تكون البيانات استاتيكية نسبيا، كما في حالة الكتالوجات، وأدلة الهاتف، وجدول القطارات، وما إلى ذلك وتعتبر تقنية تخزين CD-ROM، وDVD، وسطا اقتصاديا لقواعد البيانات المتكررة. ويستخدم التكرار في حالة البيانات غير المشتركة حيث لا يحتاج أحد

المواقع إلى تجديد الوقت الحقيقي للبيانات المحفوظة في المواقـــع الآخـــرى (ثي The, 1994).[1] وفي هذه التطبيقات، تحتاج البيانات إلـــى أن تكـــون متزامنة بأسرع ما يمكن عمليا. ولا يعد التكرار طريقة حيويـــة لتطبيقـــات الخط المفتوح، مثل حجوزات مقاعد الطيران، والعمليـــات الجاريـــة لآلات الصرافة التلقائية، والأنشطة المالية الآخرى، وهناك أنواع من التكرار :

أ- تكرار اللقطة السريعة

توجد مخططات مختلفة لتجديد نسخ البيانات. وتدعم لقطات ســـريعة دورية، أو نسخ الجدأول البسيطة لبعض التطبيقات، مثل تلك الخاصة بدعم القرار وعمل مستودعات البيانات التي لا تحتاج إلى بيانات حالية [2] والبديل، يمكن إرسال تلك الصفحات التي تغيرت فقط بعد آخر لقطة سريعة، والذي يسمى إنعاشا مميزا.

ب- تكرار قريب من الوقت الحقيقي

بالنسبة إلى المتطلبات القريبة من التنفيذ المباشر (Real Time) ، يمكن تخزين وتمرير الرسائل لكل عملية جارية مكتملـــة عبـــر معلومـــات الشبكة لكل العقد لتجديد البيانات بأسرع ما يمكن، دون فرض التأكيد علـــى عقدة المنبع في تجديد قاعدة البيانات.وإحدى طرق إنتاج مثل هذا الرســـائل تكون عن طريق استخدام المنشطات. والمنشط هو مجموعة من عبـــارات SQL تأخذ إسما معيناً وتخزن في قاعدة البيانات،وينفـــذ (ينـــشط) عنـــدما يحـــدث تعـــديل فـــي البيانـــات (INSERT، و UPDATE، وDELETE).

ت- تكرار السحب

تعد كل المنشطات المـــشروحة أعـــلاه للتكـــرار المتـــزامن أمثلـــة لاستراتيجيات الدفع. وتوجد استراتيجيات للسحب أيضا. وفي استراتيجيات السحب تتحكم عقدة الهدف، وليست المصدر، في تجـــدد قاعـــدة البيانـــات المحلية و متى يلزم التجديد، ولاستراتيجيات السحب ميزة مراقبة الموقـــع المحلي عند الحاجة لذلك، كما يمكن أن يتعامل مع التجديدات.لـــذلك يكـــون

[1] The , L . 1994 , "Distribute Data Without Choking the Net " Datamation (January 7) : pp 35-38

[2] Edelstien ,H. 1995 , "The Challenge of Replication, Part 1." DBMS 8 (March) : pp 46-52

التزامن أقل تعطيلا، ويحدث عند الحاجة كل موقع له فقط، وليس عندما يعتقد موقع رئيسي أنه من الأفضل عمل التجديد.

ث- متى يستخدم التكرار

يعتمد كون التكرار تصميما بديلا حيويا لقاعدة البيانات الموزّعة على عدد من العوامل [1]:

1- خطوط وقت البيانات Data Timeline، التطبيقات التي يمكن أن تتعامل مع بيانات قليلة الطلب (سواء كان هذا لبعض ثواني، أو بعض ساعات) تكون أفضل في التكرار.

2- إمكانيات نظام إدارة قاعدة البيانات DBMS Capabilities، للنظام إمكانية هامة في تحديد التكرار.فإذا كان سيدعم استفسارات تشير إلى بيانات في أكثر من عقدة واحدة أم لا.. فإذا لم يكن هذا هو الحال، فيكون التكرار مرشحا أفضل من مخططات التجزئة.

3- تضمينات الأداء Performance Implications، يمكن أن يحث التكرار في كل عقدة تنشط دوريا.وعند حدوث هذا التنشيط، يمكن أن تكون العقدة الموزّعة مشغولة جدا للتعامل مع حجم كبير من التجديدات. فإذا حدث التنشيط بواسطة منشطات الأحداث (مثال ذلك، عندما يتراكم حجم معين من التغييرات)، فيمكن أن يحدث التنشيط عند الوقت الذي تكون العقدة البعيدة مشغولة في أداء عمل محلي.

4-عدم التجانس في الشبكة Heterogeneity in the Network يمكن أن يتحقق التكرار إذا استخدمت العقدة المختلفة نظم تشغيل مختلفة، ونظم DBMS مختلفة، أو من الأكثر شيوعا استخدام تصميمات قواعد بيانات مختلفة. ويمكن أن تعني التغييرات من أحد المواقع إلى مواقع أخرى مقطعا مختلفا لترجمة التغييرات.

5- إمكانية شبكة الاتصالات Communications Network Capabilities، يمكن أن تمنع سرعة النقل، وسعته في شبكة اتصالات

[1] Froemming, G. 1996. "Design and Replication :Issues with Mobile Applications-Part I." DBMS 9(March) : pp 48-56

بيانات الإنعاش الكامل للجداول الكبيرة جدا. وبالتالي يسمح بالتكرار لتلافي التأخير في عملية تحديث البيانات.

2-التجزئة الأفقية

مع التجزئة الأفقية، توضح بعض الصفوف الجدول (أو العلاقة) في علاقة أساس عند احد المواقع، وتوضع الصفوف الآخرى في علاقة أساس مع موقع آخر. وبعمومية أكبر، توزع صفوف العلاقة على العديد من المواقع، ويبين الجدول (2) نتيجة عمل التجزئة الأفقية لعلاقة Customer، حيث يوجد كل صف (Data) الآن في فرعه المحلي، فإذا أجري العملاء معظم عملياتهم الجارية من فرعهم المحلي فعلا، فيتم تشغيل العمليات الجارية محليا، وتكون أوقات الاستجابة أقل ما يمكن. وعندما يبدأ العميل عملية جارية في فرع آخر يجب أن تنقل العملية الجارية إلى الفرع المحلي للتشغيل، ويعاد نقل الاستجابة إلى الفرع الذي بدأ العملية الجارية (هذا هو النمط الطبيعي للأفراد الذين يستخدمون آلات الصرافة التلقائية ATM). فإذا تغير نمط استخدام العميل (ربما بسبب النقل)، يمكن للنظام أن يكشف هذا التغيير، وينقل السجل ديناميكيا إلى المواقع التي تبدأ فيه معظم العمليات الجارية. وبصورة موجزة، للتجزئة الأفقية و لقاعدة البيانات الموزّعة أربع مميزات رئيسة :

أ- الكفاءة efficiency، تخزن البيانات بالقرب من المكان الذي تستخدم فيه، ومعزولة عن البيانات الآخرى التي يستخدمها المستخدمون، أو برامج آخرى.

ب- الأمثلية المحلية local optimization ، يمكن أن تخزن البيانات بغرض تحقيق أمثلية للأداء للاتصال المحلي.

ت- الأمن security، لا تكون البيانات متاحة لكافة المستخدمين في الموقع، وإنما للذين لديهم الصلاحيات فقط.

ث- سهولة عمل الاستفسارات ease of querying، يكون دمج البيانات عبر الأجزاء الأفقية سهلا نظر لسهولة دمج الصفوف باستخدام الاتحادات عبر الأجزاء.

جدول رقم (2) يبين التجزئة الأفقية

أ- فرع BAGHDAD

ACC-Number	Customer-Name	Branch-Name	Balance
150	Hamdi	Baghdad	1000
185	Samer	Baghdad	520
203	Aws	Baghdad	860
520	Osama	Baghdad	750

ب- فرع AL-BASRAH

ACC-Number	Customer-Name	Branch-Name	Balance
150	Ali	AL-basrah	1000
189	Salman	AL-basrah	520

لذلك، عادة ما تستخدم التجزئة الأفقية عندما تكون الوظيفة التنظيمية موزّعة، مع اهتمام كل موقع بفئة جزئية فقط من لحظات الكينونة(غالبا ما تكون معتمدة على المواقع الجغرافية).

و للتجزئة الأفقية عيبين أساسيين أيضا:

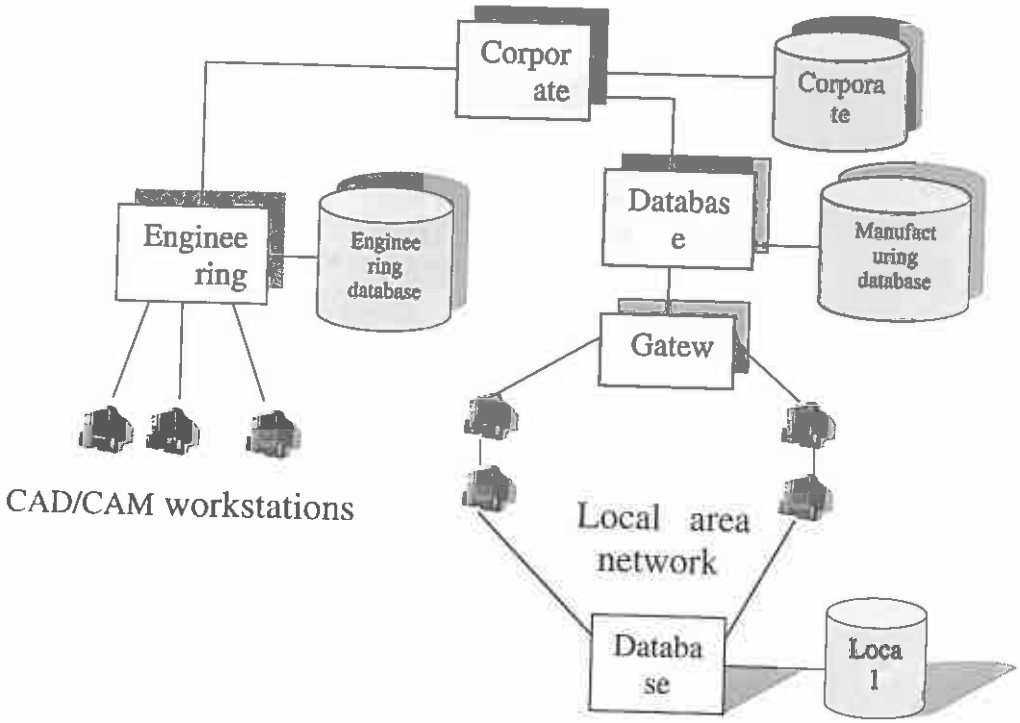
أ- سرعة الاتصال غير متسقة inconsistent access speed ، عندما تكون هناك حاجة إلى أجزاء متعددة، يكون هناك اختلاف في وقت الاتصال بالبيانات، وخصوصاً عندما تكون مخزنة محليا فقط.

ب- القابلية لتلف الاحتياطي backup vulnerability، نظرا لعدم تكرار البيانات عندما يصبح الاتصال بالبيانات الموجودة في احد المواقع غير ممكن، لا يمكن تحويل الاستخدام إلى مواقع أخرى توجد فيها نسخة من البيانات، ويمكن أن تفقد البيانات إذا لم يتم تنفيذ عمل احتياطي مناسب عند كل موقع.

3-التجزئة الرأسية

مع طريق التجزئة الرأسية يتم إسقاط بعض أعمدة العلاقة، في علاقة أساس في أحد المواقع، ويتم إسقاط الأعمدة الآخرى في علاقة أساس في موقع آخر (وبعمومية أكثر، يمكن إسقاط الأعمدة في العديد من المواقع). ويجب أن تقتسم العلاقات الموجودة عند كل موقع من المواقع نطاقاً مشتركا، وذلك حتى يمكن إعادة تشييد الجدول الأصلي.

لتوضيح التجزئة الرأسية، فإننا نستخدم تطبيق شركة التصنيع المبين في الشكل (3) ويبين الجدول (3) علاقة الجزء Part مع رقم الجزء Part_Number كمفتاح أولي. ويستخدم قسم التصنيع بعض هذه البيانات أساسا، بينما يستخدم قسم الهندسة بعضها البعض في اغلب الأحوال. وتكون البيانات موزعة على حاسبات الأقسام المناظرة لهذا باستخدام التجزئة الرئيسة كما هو مبين في الشكل، ويتم الحصول على كل جزء من الأجزاء المبينة في الجدول(4) عن طريق أخذ الإسقاطات (أي الأعمدة المختارة) للعلاقة الأصلية. ويمكن الحصول على العلاقة الأصلية بدورها عن طريق أخذ الروابط للأجزاء الناتجة.



شكل رقم (3) يبين نظام التشغيل منتشر لإحدى شركات التصنيع

(المصدر: Database Programming & Design، Vol. 2, April 1989)

جدول رقم (3) يبين علاقة الجزء Part

Part-Number	Name	Cost	Drawing-No	Qty-on-Hand
P2	Ahmad	100	123-7	20
P7	Hamdi	550	621-0	100
P3	Samer	48	174-3	0
P1	Aus	220	416-2	16
P8	Nabeel	16	321-0	50

جدول رقم (4) يبين التجزئة الرأسية

أ- قسم الهندسة

Part-Number	Drawing-No
P2	123-7
P7	621-0
P3	174-3
P1	416-2
P8	321-0

ت-قسم التصنيع

Part-Number	Name	Cost	Qty-on-Hand
P2	Ahmad	100	20
P7	Hamdi	550	100
P3	Samer	48	0
P1	Aus	220	16
P8	Nabeel	16	50

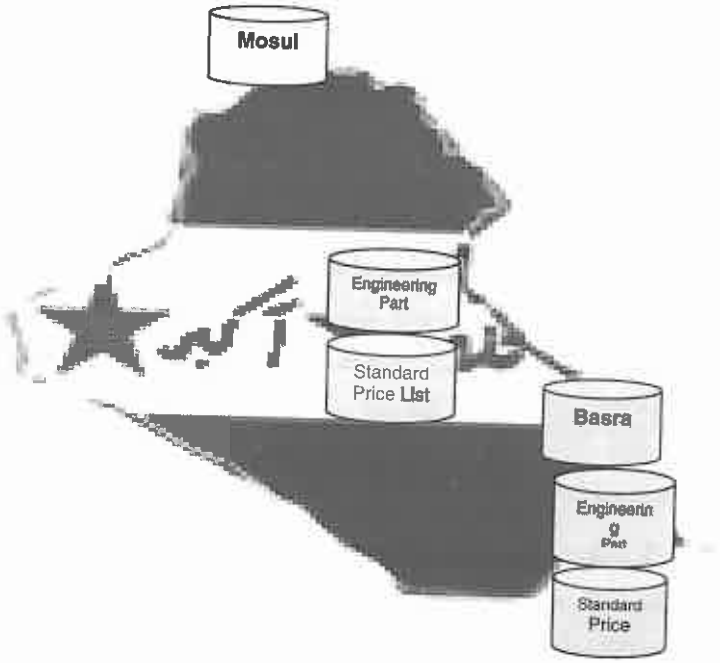
4-خليط العمليات

وهي مجموعة خليط غير محددة تقريبا للاستراتيجيات السابقة. فيمكن أن تخزن بعض البيانات مركزيا، بينما يتكرر بعضها الآخر في مواقع مختلفة. كما يمكن أن يكون من المرغوب فيه وجود أجزاء أفقية راسية، لعلاقة معينة، للبيانات الموزعة ويوضح الشكل التالي مثالا لإستراتيجية الخلط. ويلاحظ ما يلي :

1- تكون كل من البيانات Engineering Part ، و Accounting ، وCustomer مركزية في مواقع مختلفة.
2- تتجزأ بيانات الأجزاء النمطية Standard (أفقيا) عبر الثلاثة مواقع.
3- تتكرر قائمة الأسعار النمطية Standard Price List عند ثلاثة مواقع.

والقاعدة الغالبة في تصميم قاعدة البيانات الموزّعة هي أن البيانات يجب أن تخزن عند المواقع التي سيتم الاتصال بها و في أغلب الحالات (رغم أن بعض الاعتبارات الآخرى مثل الأمن، والتكلفة يمكن أن تكون

مهمة أيضا). ويلعب إداريو البيانات دورا مركزيا وهاما في تنظيم قاعـدة البيانات الموزّعة بغرض جعلها موزّعة، وليست (لا مركزية).



شكل (4) استراتيجيات توزيع قواعد البيانات

اختيار الإستراتيجية الصحيحة لقواعد البيانات الموزّعة

اعتماداً على الحوارات السابقة يمكن تنظيم قاعدة البيانات الموزّعـة في خمس طرق فريدة وهي :

1- مركزية كلية عند موقع واحد، ويتم الاتصال به من الكثير من المواقـع الموزّعة جغرافيا.
2- متكررة جزئيا، أو كليا عبر مواقع موزّعة جغرافيا، مع التجديد الدوري لكل نسخة بلقطات سريعة.
3- متكررة جزئيا، أو كليا عبر المواقع موزّعة جغرافيا، مع تزامن قريب من الوقت الحقيقي Real time للتجديدات.

4- مجزأة إلى قطاعات في مواقع مختلفة موزّعة جغرافيا، وتظل في قاعدة بيانات منطقية واحدة، ونظام DBMS منتشر واحد.
5- مجزأة في قطاعات مستقلة وغير متكاملة تنتشر على حاسبات، ونظم برامج قاعدة بيانات متعدد.

ولا يمكن المفاضلة بين هذه الطرق الخمسة بدون أن يربط الاختيار بطبيعة عمل المنظّمة. وأن المقارنة بهذه الطرق الخمسة يتم الأخذ بالاعتبارات التالية: قابلية التوسع بإضافة عقد جديدة، والتحميل الزائد للاتصالات والطلب على شبكة الاتصالات، وقابلية الإدارة، واتساق البيانات، ويحتاج مصمّم قاعدة البيانات الموزّعة أن يزن هذه العوامل لاختيار الإستراتيجية الجيدة لبيئة قاعدة بيانات موزّعة معينة. ويعتمد اختيار الإستراتيجية الأفضل على العديد من العوامل.

نظام إدارة قاعدة البيانات الموزّعة Distributed DBMS

حتى توجد قاعدة بيانات موزّعة يجب أن يكون هناك نظام إدارة قاعدة بيانات موزّعة، والذي ينسق الاتصال بالبيانات عند عقد مختلفة. وتسمى بنظام إدارة قاعدة بيانات موزّعةDistributed DBMS . ورغم أنه يمكن أن يوجد نظام DBMS في كل موقع لإدارة قاعدة البيانات المحلية الموجودة في هذا الموقع، فيلزم نظام DBMS أيضا بتنفيذ الوظائف التالية [1]:

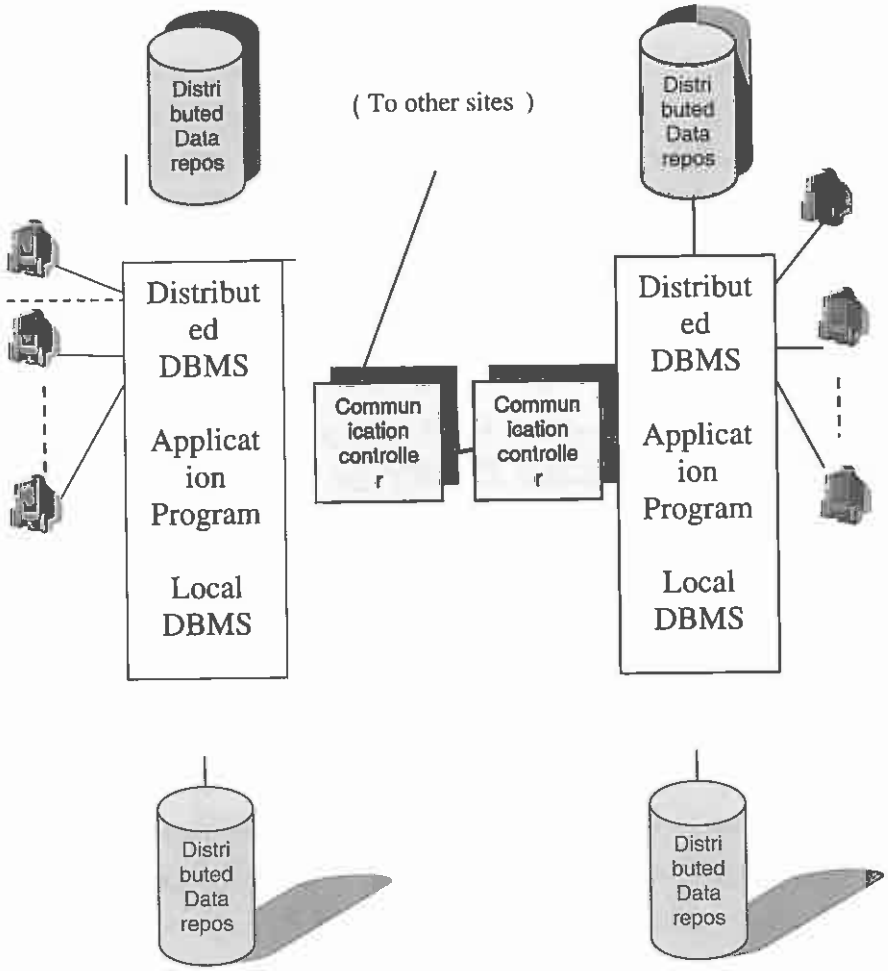
1-حفظ تتبع مكان تواجد البيانات في قاموس البيانات الموزّعة.
2- تحديد الموقع الذي تسترجع منه البيانات المطلوبة، والموقع الذي يتم تشغيل كل جزء من الاستفسار المنتشر.
3- إذا لزم الأمر، ترجمة الطلب عند العقد باستخدام DBMS محلي في طلب مناسب إلى عقدة آخرى باستخدام DBMS مختلفة، ونموذج بيانات مختلفة، وإعادة البيانات إلى العقدة الطالبة في صورة تقبلها هذه العقدة.
4- توفير وظائف إدارة البيانات مثل الأمن، والتزامن ومراقبة الإغلاق التاتل، وأمثلة الاستفسار، والاستعادة من الفشل.

[1] Elmasri, R. and S. B. Navathe. 1994 . Fundamentals of Database System. Menlo Park, Ca:Benjamin/Cummings

5- توفير الاتساق عبر نسخ البيانات في المواقع البعيدة.

ويمكن أن توجد نظم DBMSs مختلفة تعمل عند كل موقع محلي، مع نظام DBMS واحد رئيسي يراقب التداخلات بين أجزاء قاعدة البيانات. وتسمي مثل هذه البيئة قاعدة بيانات موزّعة غير متجانسة. وعمليا فلا يكون عدم التجانس الكامل في وقتنا الحالي، وتوجد إمكانيات محدودة مع بعض المنتجات عندما يتبع الكل نفس معمارية البيانات (مثال ذلك، العلاقية).

يبين شكل(5) معمارية شائعة لنظام الحاسب مع إمكانية DBMS منتشر. لكل موقع DBMS محلي يدير قاعدة البيانات المخزنة في هذا الموقع. كما أن لكل موقع نسخة من DBMS المنتشر وقاموس/دليل Data Dictionary/Directory (DD/D) للبيانات الموزّعة المصاحبة له. ويحتوي (DD/D) الموزّعة على مواقع كل البيانات في الشبكة، وذلك مع تعريف البيانات. ويقوم DBMS المنتشر بتشغيل طلبات المستخدمين، أو برامج التطبيقات للبيانات أولا، والذي يحدد إذا كانت العملية الجارية محلية، أو شاملة وتكون العملية الجارية المحلية Transaction Local تلك التي تخزن فيها البيانات المطلوبة كلها في المواقع المحلية. وتتطلب العملية الجارية الشاملة Global Transaction الإشارة إلى البيانات الموجودة في موقع واحد، أو أكثر غير محلي لتحقيق الطلب. وبالنسبة إلى العمليات الجارية المحلية، يمرّرَ DBMS المنتشرَ الطلب إلى DBMS المحلي، وبالنسبة إلى العملية الجارية الشاملة، يوجه DBMS المنتشر الطلب إلى مواقع آخرى طبقا لما يلزم عمله. وتتبادل نظم DBMSs الموزّعة الرسائل عند المواقع المشاركة وعند الحاجة ، لذلك يجرى تشغيل العملية الجارية حتى تكتمل (أو تتوقف، إذا لزم الأمر ذلك). ويمكن أن تكون هذه العملية معقدة للغاية.



شكل (5) معمارية نظام DBMS منتشر

في مناقشتنا لمعمارية النظام المنتشر شكل (5) افتراضنا أن نسخ DBMS الموزّعة، و (DD/D) وجود عند كل موقع (لذلك فان (DD/D)يكون بنفسه مثالا لتكرار البيانات). والبديل هو وضع DBMS المنتشر، و DD/D في موقع مركزي، كما أن هناك استراتيجيات أخرى ممكنة أيضا، إلا أن الحل المركزي يكون عرضه للفشل، وبالتالي تقل الرغبة فيه.

ويجب أن يعزل DBMS المنتشَر المستخدمين بأكبر قدر ممكن عن تعقيدات إدارة قاعدة البيانات الموزّعة. كذلك يجب أن يجعل DBMS المنتشر مواقع البيانات شفافاً في الشبكة، كذلك يعمل على أن تكون سمات آخرى لقاعدة البيانات الموزّعة تمتاز بالشفافية أيضا.

وهناك أربعة أهداف لنظام DBMS المنتشَر تسهل عند تحقيقها، في تشَييد البرامج، واستخراج البيانات، وهذه الأهداف هي :

1. شفافية الموقع

رغم انتشار البيانات جغرافيا، ويمكن أن تنتقل من مكان لآخر، فمع شفافية الموقع يمكن العمل (بما في ذلك المبرمجين) كما لو كانت كل البيانات موجودة عند عقدة واحدة.

2. شفافية التكرار

رغم أن نفس عنصر البيانات يمكن أن يتكرر في العديد من العقد في الشبكة، فمع شفافية التكرار (وتسمي في بعض الأحيان شفافية التجزئة) يمكن أن يعامل المبرمج أو المستخدم عنصر البيانات كما لو كان العنصر بيانات واحدة في عقدة فردية.

3. شفافية الفشل

يكون كل موقع (أو عقدة) في النظام المنتشر عرضة لنفس أنواع الفشل التي يتعرض لها النظام المركزي(بيانات خطأ، وكسر في رأس القرص، وغيرها). ألا أن هناك مخاطر فشل إضافي في ربط الاتصالات (أو فقدان الرسائل). ولكي يكون النظام صلبا يجب أن يكون قادرا على اكتشاف Detect الفشل، وإعادة التشكيل Reconfigure للنظام حتى يمكن أن تستمر الحسابات، وتستعاد عندما يعاد إصلاح المشغل، أو الرابط.

4. بروتوكول الالتزام

لضمان سلامة البيانات لعمليات التجديد، ينفذ مدير العمليات الجارية بروتوكول التزام، وهو إجراء جيد التعريف(يشمل تبادل الرسائل) لضمان أن العملية الجارية الشاملة أكملت بنجاح في كل موقع أو أنها أهملت [1].

[1] Date ,D . J. , 1995 , An Introduction to Database Systems. 6th ed .Reading , MA :Addison-Wesley

4. شفافية التزامن

عندما يتصل العديد من المستخدمين بقاعدة البيانات ويجددوها يمكن أن تفقد سلامة البيانات، إلا إذا استخدمت آلية للإغلاق لحماية البيانات من تأثيرات التجديدات المتزامنة. وتكون مشكلة مراقبة التزامن أكثر تعقيدا في قاعدة البيانات الموزّعة نظرا لانتشار المستخدمين المتعددين عبر المواقع المتعددة، وعادة ما تكون البيانات متكررة في العديد من المواقع أيضا [1].

تطور نظم إدارة قواعد البيانات الموزّعة Evolution of Distributed DBMS

لا زالت إدارة قواعد البيانات الموزّعة تقنية متطورة بدلا من كونها تقنية محددة. ولا توفر الإصدارات الحالية لمنتجات DBMS الموزّعة كل السمات التي سبق وصفها في الأقسام السابقة. مثال ذلك توفر بعض المنتجات شفافية الموقع للعمليات الجارية للقراءة فقط، دون أن تدعم التجديدات الشاملة، ولتوضيح تطور منتجات DBMS الموزّعة، فإننا نصف بإيجاز ثلاثة مراحل في هذا التطور، وحدة العمل البعيدة، ووحدة العمل الموزّعة، والطلبات الموزّعة.

1. وحدة عمل البعيدة

تسمح المرحلة الأولى ببدء عبارات SQL متعددة عند احد المواقع، وتنفيذها كوحدة واحدة من العمل على DBMS بعيد فردي. ويجب أن يستخدم كل من الحاسب المبتدئ والحاسب المستقبل نفس DBMS. ولا يستشير الحاسب المبتدئ دليل بيانات لتحديد الموقع الذي يحتوي على الجدأول المختارة في وحدة العمل البعيدة، وبدلا من ذلك، يجب أن يعرف التطبيق المبتدئ أين توجد البيانات، ويتصل بنظام DBMS البعيد قبل كل وحدة عمل بعيدة. لذلك لا يدعم مفهوم وحدة العمل البعيدة شفافية الموقع.

وتسمح وحدة العمل البعيدة بالتجديدات عند حاسب بعيد فردي. وتكون كل التجديدات في وحدة العمل مؤقتة حتى تجعلها عملية الالتزام

[1] المرجع السابق ، إدارة قواعد البيانات الحديثة ; تعريب : سرور علي إبراهيم سرور : - الرياض: مكتبة دار المريخ، 2003.ص ص 232 - 245

دائمة، أو يلغيها الدوران للخلف. لذلك يحتفظ بسلامة البيانات لموقع واحد بعيد. إلا أن التطبيق لا يمكن أن يتأكد من سلامة العملية الجارية عندما يكون مشمولا بأكثر من موقع واحد بعيد. وبالإشارة إلى قاعدة البيانات المبينة في الشكل (5) يمكن أن يجدد التطبيق في محافظة بغداد الملف الموجود في محافظة البصرة مع الاحتفاظ بسلامة العملية الجارية. إلا أن التطبيق لا يمكنه أن يجدد الملف الموجود في الموقعين في نفس الوقت ويظل متأكدا من الاحتفاظ بسلامة العملية الجارية. لذلك لا توفر وحدة العمل البعيدة شفافية الفشل أيضا.

2. وحدة العمل الموزّعة

تسمح وحدة العمل الموزّعة بإشارة عبارات مختلفة في وحدة العمل إلى مواقع DBMS بعيدة متعددة. وتدعم هذه الطريقة بعضا من الشفافية للموقع نظرا لاستشارة دليل البيانات لتحديد DBMS المحتوى على الجدول المختار في كل عبارة. إلا أنه يجب أن توجد كل الجداول في عبارة SQLفردية في نفس الموقع. لذلك لا تسمح وحدة العمل الموزّعة بالاستفسار التالي المصمّم لتجميع أجزاء المعلومات لكل من المواقع الثلاث في الشكل(5).

```
SELECT      PART_ NUMBER, PART_NAME
FROM        PART
WERE        COLOR-'ORANGE'
ORDER BY    PART_NUMBER
```

و بالمثل، لا تسمح وحدة العمل الموزّعة بعبارة واحدة أن تحاول تجديد البيانات وفي أكثر من موقع واحد. مثال ذلك، تهدف العبارة التالية إلى تجديد ملف الأجزاء الموجود في المواقع الثلاثة :

```
UPDATE      PART
SET         UNIT-PRICE -127.49
WHERE       PART_NUMBER -12345
```

3. الطلب المنتشر

يسمح الطلب المنتشر بأن تشير عبارة SQL واحـــدة إلـــى جـــداول موجودة في أكثر من DBMS واحد بعيد، مع التغلب على القيد الرئيـــسي الخاص بوحدة العمل الموّزَعة. ويدعم الطلب المنتشر شفافية الموقع الحقيقي نظرا لأن عبارة SQL واحدة يمكن أن تشير إلى جداول موجودة في مواقع متعددة. إلا أن طلب المنتشر يمكن، أو لا يمكن، أن يدعم شفافية التكـــرار، أو شفافية الفشل. وربما ينقضي بعض الوقت قبل أن يظهر النظام DBMS منتشراً حقيقياً يدعم كل سمات الشفافية التي سبقت مناقشتها.

منتجات نظام إدارة قاعدة البيانات الموّزَعة

لدي معظم موردي نظم إدارة قواعد البيانات صيغة موّزَعة. وفـــي أغلب الحالات لاستغلال إمكانيات قاعدة البيانات الموّزَعة يجب أن يعمـــل DBMS لأحد الموردين عند كل عقدة (بيئـــة قاعـــدة بيانـــات موّزَعـــة متجانسة). وتعد صيغ الخادم والعميل لقاعدة البيانـــات الموّزَعـــة الـــصيغة الأكثر شيوعا، مع الجدال في هذا، الموجود في وقتنا الحالي. ففـــي بيئـــة الخادم والعميل من السهل جدا تعريف قاعدة البيانات بجدأول موجودة عنـــد عقد متعددة في شبكة منطقة محلية، أو شبكة منطقة عريضة. وبمجـــرد أن يبني برنامج المستخدم الرابط مع كل موقع بعيد ويتم تحميل نظم وســـيطة مناسبة لقاعدة البيانات، تتحقق شفافية المواقع الكاملة لذلك، فـــي صـــيغة قاعدة بيانات الخادم والعميل تتاح قواعد بيانات الموّزَعة بالفعل لأي مطور نظم معلومات، كما يتاح نظام DBMS غير متجانس أيضا.

ورغم استمرار هذه الطرق، فهي توضح الغرض العام لكيفية تنـــاول الموردين المختلفين إدارة قاعدة البيانات الموّزَعة. وهنـــاك صـــعوبة فـــي اختيار احد المنتجات الموّزَعة نظرا لأن الإمكانيات الدقيقة يجب أن تتفـــق بعناية مع حاجات المنظمة. كما أنه مع مثل هذه الخيارات الكثيـــرة، ومـــع تعامل كل منتج مع البيانات الموّزَعة بصورة مختلفة، فمن المستحيل تقريبا تخطيط الأسس العامة لإدارة قاعدة البيانات الموّزَعة. ويحتاج تـــصميم أي قاعدة بيانات موّزَعة إلى تحليل دقيق لكل من احتياجات الأعمال وتعقيدات

النظام. ومن أكثر منتجات نظام إدارة قاعدة البيانات الموزّعة استخداما في الوقت الحاضر هي "منتجات ميكروسوفت، وخادم لغة الاستفسار المهيكلة".

حيث يستخدم SQL Server طريقة الموقع الأولي، والتي فيها قاعدة البيانات الموجودة في أحد المواقع الوحيدة التي يمكن أن تستقبل تجديدات. كما يستخدم SQL Server طريقة النشر والاشتراك المستخدمة في العديد من المنتجات DBMS الموزّعة الآخرى أيضا. والمفهوم الهام مع SQL Server هو ما يستغرقه تماثل النشر والاشتراك.

وبالنسبة إلى قاعدة البيانات المحمّلة (قاعدة بيانات موجودة على حاسب رئيسي أو خادم حاسب شخصي مربوط على شبكة)، يعمل SQL Server بالاتصال مع مشغلات ODBC، وآلة Jet المتاحة مع Access والبيسك المرئي.

يتم الوصول إلى ***مصدر بيانات ODBC*** (قاعدة بيانات مفتوحة Open Database Connectivity). هو مصدر للبيانات ومعلومات الاتصال المطلوبة للوصول إلى هذه البيانات. من أمثلة مصادر البيانات هي Microsoft Access وMicrosoft SQL Server و Oracle RDBMS وجدول بيانات والملف النصي. تتضمن أمثلة معلومات الاتصال موقع الملقم واسم قاعدة البيانات ومعرف تسجيل الدخول وكلمة المرور وخيارات برنامج تشغيل ODBC المختلفة التي تصف كيفية الاتصال بمصدر البيانات.

في معمارية ODBC، يتصل أحد التطبيقات (مثل Access أو برنامج Visual Basic Microsoft)، بإدارة برنامج تشغيل ODBC ، الذي يستخدم بدوره برنامج تشغيل ODBC معين (مثل، برنامج تشغيل ODBC Microsoft SQL) للاتصال بمصدر بيانات (في هذه الحالة، ***قاعدة بيانات Microsoft SQL Server***).
للاتصال بمصادر البيانات هذه، يجب القيام بما يلي:

- تثبيت برنامج تشغيل ODBC الملائم على الكمبيوتر الذي يحتوي على مصدر البيانات.

- تعريف اسم مصدر بيانات (DSN) باستخدام **إدارة مصادر بيانات ODBC** لتخزين معلومات الاتصال في تسجيل Microsoft Windows أو ملف DSN أو سلسلة اتصال في التعليمات البرمجية Visual Basic لتمرير معلومات الاتصال مباشرة إلى "إدارة برامج تشغيل ODBC".

- ***مشغل قاعدة بيانات Microsoft Jet*** :"ذلك الجزء من نظام قواعد بيانات Access الذي يستعيد البيانات من قواعد بيانات المستخدم والنظام ويخزنها فيها. ويمكن اعتبار مشغل قواعد بيانات Microsoft Jet على أنه إدارة البيانات التي تأسست عليها أنظمة قواعد بيانات مثل "Access".[1]

المـــراجـــع

1. آلتر أ جيفري، بيرسكون ب مآرى، مكفادن آر فريد; تعريب : سرور علي إبراهيم سرور: إدارة قواعد البيانات الحديثة.- الرياض : مكتبة دار المريخ، 2003م.
2. م. هاني عبد النبي و عزب محمد عزب /الطريق إلى احتراف SQL /دار الكتب العلمية للنشر والتوزيع – القاهرة، 2004م.
3. نبيل عبد الله قمصاني. "الاتجاهات السلوكية لمستخدمي قواعد المعلومات والمنتجين لها". القاهرة عالم الكتب مج 6،21. 2000م.
4. Bell.D.and Grimson.1992 Distributed Database System. Addison-Wesley.
5. Date, C‹j. 1995 An Introduction to Database System.6th ed Addison-Wesley.
6. Özsu‹ M. T., and p. Valduriez. 1992."Distributed Database System :Where Were We" Database Programming &Design 5 (April).
7. Koop, P.1995. "Replication at Work," DBMS 8 (March)
8. The‹ L. 1994" Distribute Data Without Choking the Net." Datamotion (January 7)
9. Froemming‹ G. 1996. "Design and Replication :Issues with Mobile Applications-Part 1."DBMS 9 (March)

[1] مركز تعليمات Microsoft Access المجهز مع نسخة Office XP

10. Edelstien،H. 1995 "The Challenge of Replication, Part 1 "DBMS 8(March)
11. Elmasri, R. and S. B. Navathe. 1994. Fundamentals of Database System. Menlo Park, Benjamin/Cummings
12. مركز تعليمات Microsoft Access المجهز مع نسخة Office XP