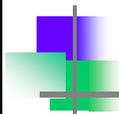




XML Consortium

XMLDBベンダー訪問 総覧



XMLDB勉強会

リーダー 加藤哲義
(株式会社サイバーテック)
katou@cybertech.co.jp

2007.5.18

本年1月、XQuery 1.0が正式勧告

XMLDBが続々と登場



着手意向が高い技術トップ10

順位	技術名	昨年順位
1	IP v6	2
2	生体認証	3
3	情報漏洩防止ツール	1
4	ブレード・サーバー	
5	RFID	6
6	テスト・品質管理ツール	7
7	デジタル署名	5
8	XML DBMS	
9	プロジェクトマネジメント	14
10	SOA	10

XMLDB勉強会活動

2006. 8.10. 第1回設立準備委員会

9. 5. 第2回設立準備委員会 設立趣意書提示

10.4. 第1回XMLDB勉強会 (XMLDBセミナー併催)

11. 2. 第2回XMLDB勉強会

12. 6. 第3回XMLDB勉強会・マーケティング・サブグループ

12. 7. 第3回XMLDB勉強会・技術・サブグループ

⋮
⋮
⋮

各サブグループ、2007年5月まで毎月1回会合

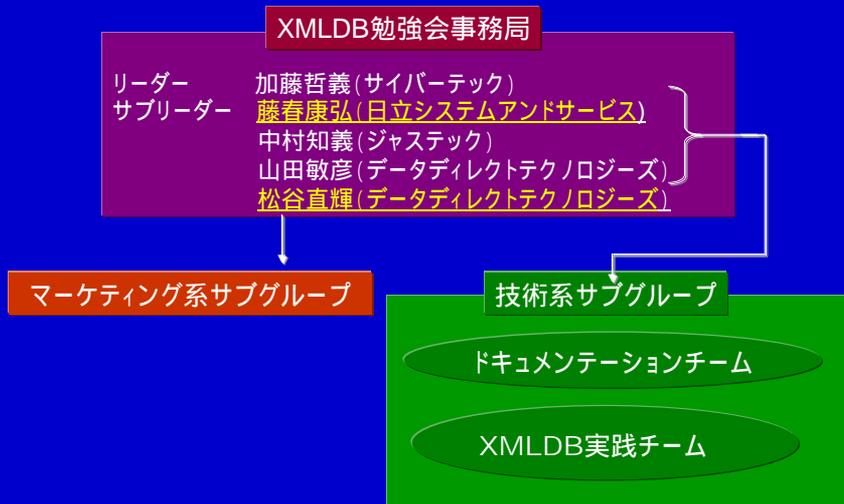
XMLコンソーシアムイベントへの参加

- XMLコンソーシアム Day (2006.12.11)
「XMLDB勉強会、活動の意義と目標」発表

XMLデー・タベース事例セミナー参加 (2007.1.23)

XMLコンソーシアムWeekにて活動成果発表 (2007.5.18)

3. 体制



活動概要

マーケティング系サブグループ

ベンダー訪問を行い、

その製品の特長とユーザ事例や、実ユーザの生の声を聞く。
それをもとにディスカッション。
訪問記を残す。



訪問先

マーケティング系サブグループ

日本IBM 訪問 (2006.12.2)



サイバーテック 訪問 (2007.1.10)



日本オラクル 訪問 (2007.2.5)



三井物産セキュアディレクション 訪問 (2007.4.19)



東芝・メディアフュージョン 訪問 (2007.5.24)



活動概要

技術系サブグループ

ドキュメンテーションチーム

XMLDB(Xpriori)実践チーム

XMLDB(Cyber Luxeon)実践チーム

XMLDB(DB2)実践チーム

W3C 「XQuery Use Cases」 翻訳
「W3Cの歩き方」みたいなものもつくれたら

実際にXMLDBを使って勉強しながらXMLDB導入ステップの考察、利用方法を考えていく。

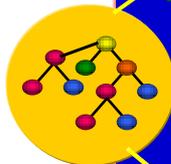


XMLDBの位置づけ

意味の構造からシステムの構造への変換

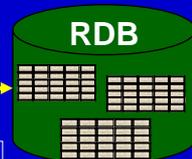
XML=意味の構造

```
<?xml version="1.0" encoding="Shift_JIS" ?>
<健康診断データ>
  <健康者>
    <個人データ>
      <名前>加藤大介</名前>
      <性別>男</性別>
      <年齢>50</年齢>
      <身長>174</身長>
    </個人データ>
    <体重>63</体重>
  </健康者>
  <上>110</上>
  <下>80</下>
</血圧>
<病>正常</病>
<心>正常</心>
<視力>50歳以上かつ非高眼圧検査必要</視力>
<遠視医 遠視>高眼圧検査</遠視医 遠視>
<遠視医 正常>正常</遠視医 正常>
<遠視医 正常>正常</遠視医 正常>
<心電図>正常</心電図>
</健康者>
</健康診断データ>
```



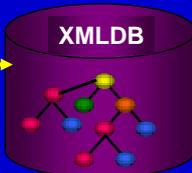
人為的変換

テーブル設計と
マッピング

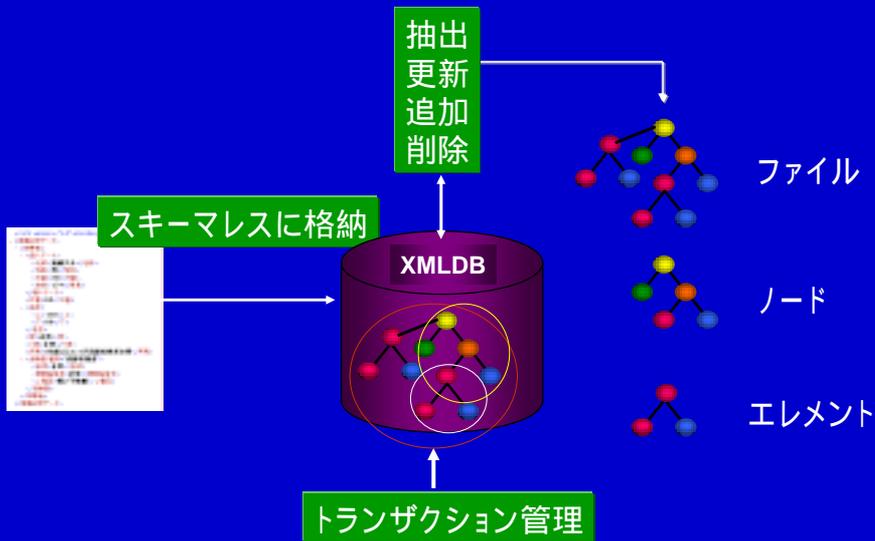


機械的変換

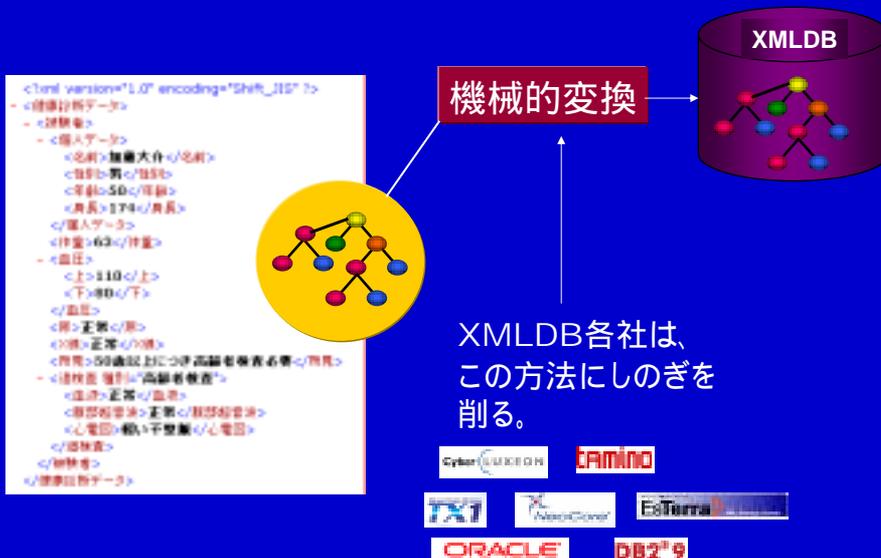
||
スキーマレス



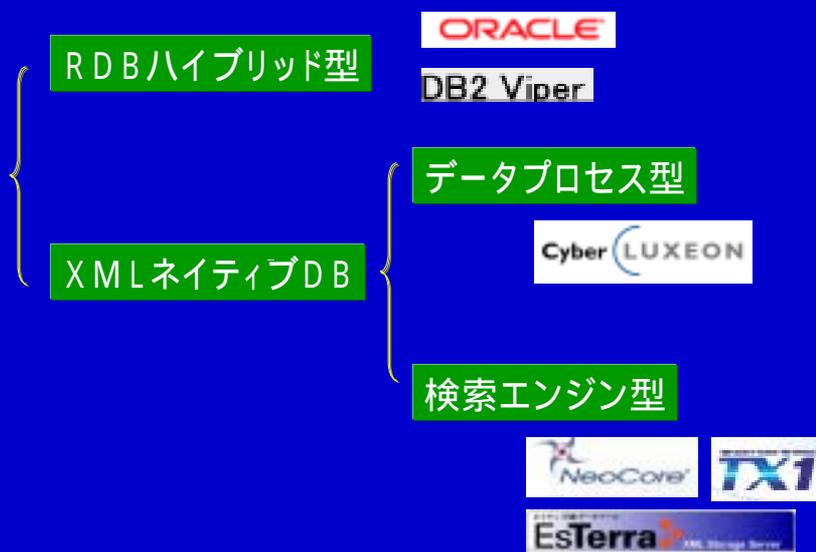
XMLDBの特性



機械的変換はユーザは楽な分、製品は苦勞する



XML DBのカテゴリ



日本IBM 訪問

2006.12.2.

以下、当日の日本IBM様からの資料を引用しています。

DB2 9のベンチマークサマリー

Information Management software

© 2004 IBM Corporation

従来のXMLデータ・サポート: 2つの方法

DB2 Viper Star Festival

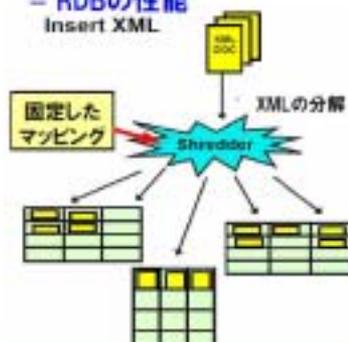
- 1列へそのまま格納
 - スキーマの柔軟性あり
 - 検索時の翻訳が重い
索引付けに限界

```

order date="2004-11-18"
  <customer>Tocompover/customer</customer>
  <part key="02"> ... </part>
  <part key="03"> ... </part>
</order>
    
```

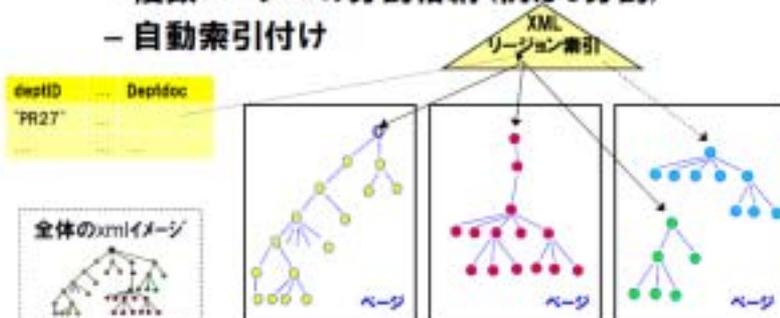
Order (CLOB)	ID
order date="2004-11-18" <customer>Tocompover/customer</customer> <part key="02"> ... </part> <part key="03"> ... </part> </order>	1

- XMLを一定フォーマットへ分解して格納
 - スキーマの柔軟性を失う
 - RDBの性能
Insert XML



・ 自動的な格納形式

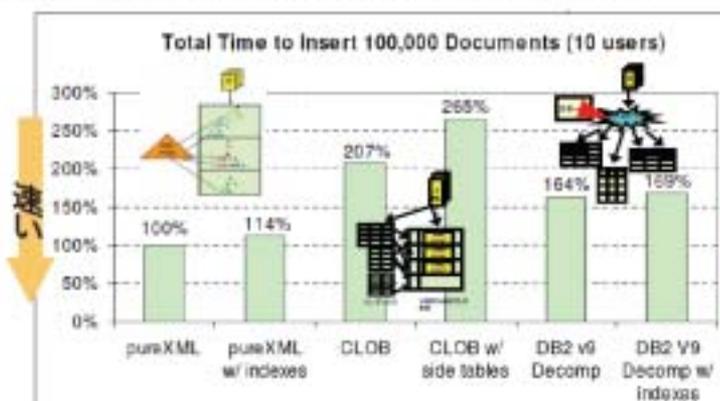
- 複数ページへの分割格納 (例は3分割)
- 自動索引付け



DB2プリフェッチ機能による同時並列読取り

IBM

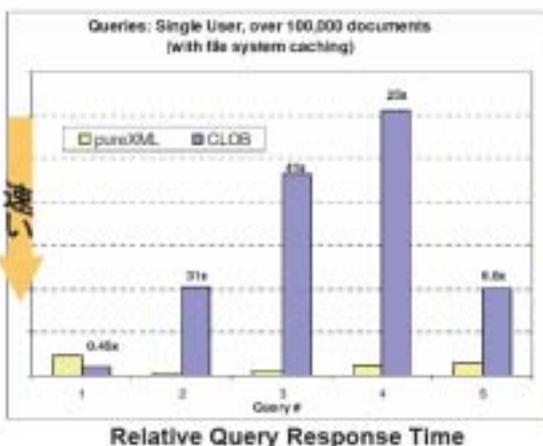
pureXML vs CLOB vs シュレッド型: Insertのパフォーマンス



・複数ユーザーによる同時実行

- ▶ CLOBはpureXMLに比べて2倍時間がかかる
- ▶ シュレディング (XMLを分解してリレーショナル表にマッピング) はpureXMLに比べて1.5~1.7倍時間がかかる
- ▶ XML文書 2KB~20KB

pureXML vs CLOB Query Performance



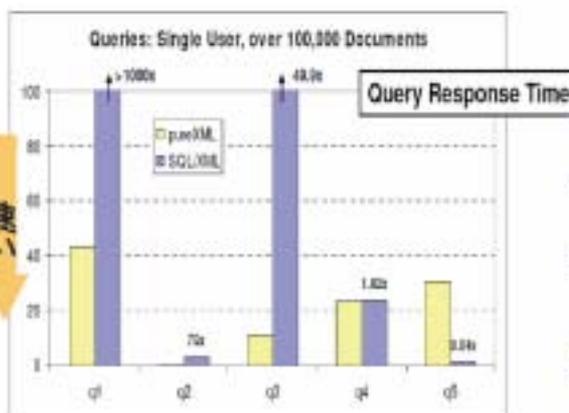
- XML列へのクエリはCLOB列へのクエリと比較して20倍~40倍高速
- CLOBは検索時の構文解析が高負荷



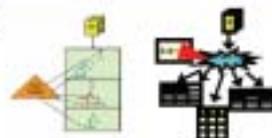
Query

- 1. 文書全体の検索、全文書指定
- 2. 文書全体の検索、1つの述部(custid)で1文書指定
- 3. 文書全体の検索、5つの述部で複数の文書指定
- 4. 全ての文書への部分検索
- 5. いくつかの文書への部分検索、1つの述部

pureXML XQuery vs. シュレド型に対するSQL/XML



- XML列へのクエリはシュレド型へのクエリと比較して50倍以上高速
- q1、q3はRDBからXML文書を組み立てるために高負荷



Query

- 1. 文書全体の検索、全文書指定
- 2. 文書全体の検索、1つの述部で1文書指定
- 3. 文書全体の検索、5つの述部で複数の文書指定
- 4. 全ての文書への部分検索
- 5. いくつかの文書への部分検索、1つの述部

DB2 9 pureXMLの活用事例

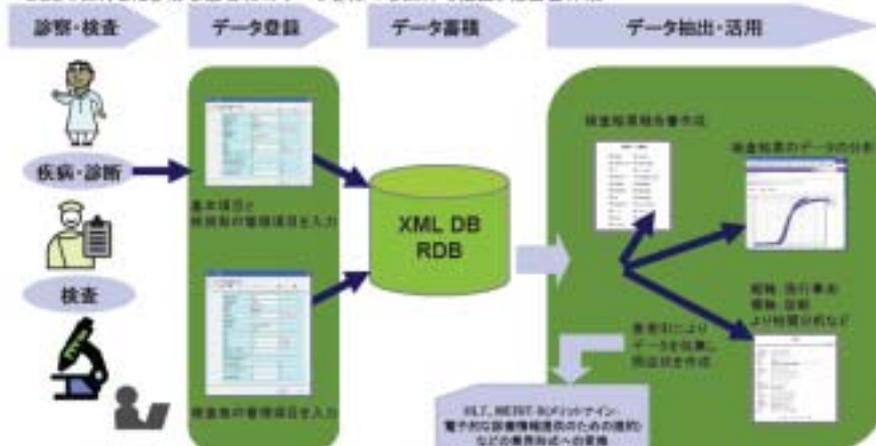
日本アイ・ピー・エム株式会社
ソフトウェア事業
Info.Mgmt.テクニカル・セールス

Information Management software

© 2008 IBM Corporation

活用事例1.電子カルテシステム

- 電子カルテの追加機能としてこれまで入力していなかった臨床データをXMLで設計し入力
- 「退院サマリー」「検査レポート」等をXMLで設計し入力。ペーパーレス。業務の効率改善、患者サービスの向上。入力したデータは「ドクターの臨床研究」「学会資料などに採用」
- DB2 9保存した多彩な患者様のデータを様々な条件で抽出。報告書作成



医療データ/XMLによる表現

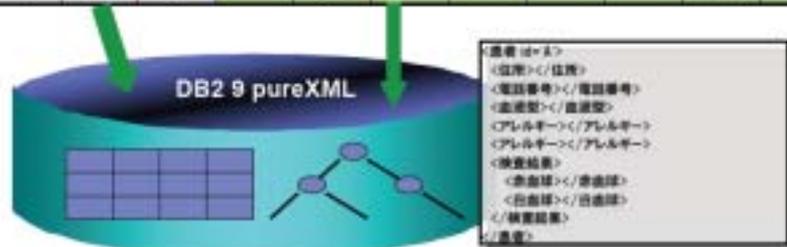
- 属性がまばらなデータをXMLで表現
 - 値の入っている属性のみXMLのタグとして表現
 - 1表でも実現可能
 - システム化が困難だった分野へも適用が現実化
- 柔軟なデータ変更にも対応
 - 変更が他のデータに影響を与えない



医療データ/RDBとXMLのハイブリッド・デザイン

- キーや必須属性はリレーショナル列へ、まばらで変更の多い属性はXML列へ

患者ID	氏名	性別	生年月日	住所	電話番号	血液型	保険種	アレルギー 1	アレルギー 2	検査結果 1	検査結果 2
A	レ	レ	レ	レ	レ	レ		レ	レ	レ	レ
B	レ	レ	レ	レ			レ	レ	レ	レ	レ
C	レ	レ	レ		レ	レ					
D	レ	レ	レ	レ		レ	レ			レ	レ
E	レ	レ	レ	レ			レ				
F	レ	レ	レ		レ					レ	



例1) ハイブリッド形式のデザイン

- 2表で自然な従属を反映
 - 誰が何々を購入
 - 顧客情報や言語の拡張性大

注文 - 戦略的施策を早期に実現可能

注文番号	顧客番号	注文
987564331	A6789	
...	...	

商品

商品番号	価格	課税	説明
A54	200	Y	
985	115	Y	
...

```

<注文 partNo="987564331">
  <顧客 custid="A6789">
    <name>.....</name>
    <address>.....</address>
  </顧客>
  <item>
    <item>
      <商品番号>A54</商品番号>
      <数量>12</数量>
    </item>
    <item>
      <商品番号>985</商品番号>
      <数量>1</数量>
    </item>
  </item>
</注文>

<商品 partNo="A54">
  <言語 id="1">
    <説明>鉛筆</説明>
  </言語>
  <言語 id="2">
    <説明>Pencils</説明>
  </言語>
</商品>
    
```

例1) 純リレーショナル形式のデザイン

- 注文伝票: 4表のジョイン

注文

注文番号	顧客番号
987564331	A6789
...	...

商品

商品番号	価格	課税
A54	200	Y
985	115	Y
...

注文明細

注文番号	商品番号	数量
987564331	A54	12
987564331	985	1
...

商品カタログ

商品番号	言語	説明
A54	1	鉛筆
A54	2	Pencils
...

日本オラクル 訪問

2007.2.5.

以下、当日の日本オラクル様からの資料を引用しています。

Oracle XML DB 事例のご紹介

当日使われた日本オラクル社の

某医療薬品研究所 (1)

- 会社概要
 - 遺伝医学研究、人ゲノム解析、医薬品開発
- 遺伝調査サーバ
 - 大規模研究データを格納
 - 紙媒体から電子データへ変換
 - データ収集とWeb サポート
 - XML DB を採用による、システムのパフォーマンス向上、コストの削減化を実現
- リレーショナルシステムからの移行
 - 利用テーブル数を 1/8 に削減
 - アプリケーションコード 1/4 に削減

ORACLE

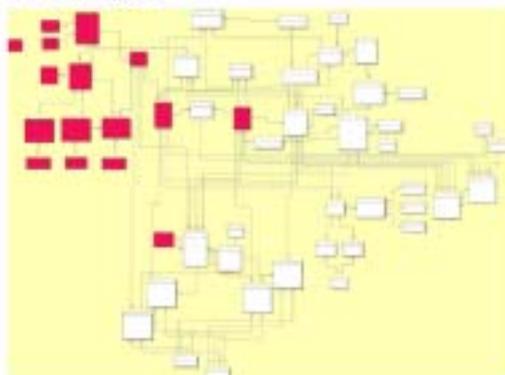
Copyright © 2007, Oracle. All rights reserved.

31

某医療薬品研究所 (2)

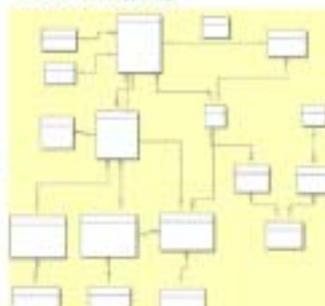
XML DB 導入によるデータモデルのシンプル化

RDBMS構造



導入前

XML DB構造



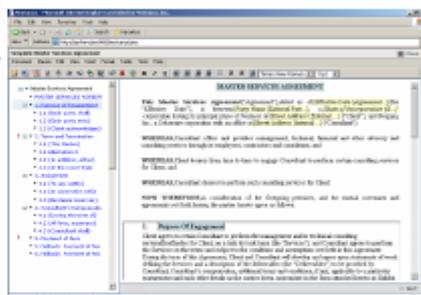
導入後

ORACLE

Copyright © 2007, Oracle. All rights reserved.

32

- 2000年設立
- 契約管理(ECM)ソリューションを提供
- 実績
 - 40社以上の顧客
 - 大部分がGlobal1000の企業
 - 40カ国以上でグローバル展開
 - ユーザ数8万
 - 20万契約の最適化



某製造業様 (1)

- 事業内容
 - 情報関連機器、電子デバイスの開発メーカー
- 要望
 - 社内文書をXML化し、ナレッジ・システムを構築したい
 - 構造変化の多い製品マニュアルをXMLで管理したい
- Oracle XML DBの提案
 - 社内文書共有ファイリングシステムとして利用
 - 製品マニュアルを全てXML文書化して管理

某出版会社様 (1)

- 事業内容
 - 出版物に関連したサイトをWebで公開
- 課題
 - 出版物ごとにフォーマットが異なる。
 - RDB構造だと設計変更による企画からカットオーバーまでの時間がかかる
- 要望
 - サイトへの短期公開の実現



XMLDBを用いて柔軟なデータ構造を実現

ORACLE

Copyright© 2007, Oracle. All rights reserved.

30

サイバーテック 訪問

2007.1.10.

以下、当日のサイバーテックからの資料を引用しています。

アジェンダ

第1部 13:30 ~ 15:00

- ・Cyber Luxeonに観るXMLDBのフェロモン
- ・ユーザ事例から観るXMLDB実践の足跡

第2部 15:15 ~ 17:00

・ユーザ直言

石川島播磨重工株式会社・宇宙航空事業本部
生産管理部 宮下幹夫様

・ディスカッション

「XMLDBフェロモンの嗅ぎ方とそのユーザ像」

データ管理プロセス

```
<?xml version="1.0" encoding="Shift_JIS"?>
<実験DATA管理システム>
<実験情報>
  <実験名>高圧抗菌実験</実験名>
  <GROUP名>基礎解析第3グループ</GROUP名>
  <報告者>加藤哲義</報告者>
</実験情報>
<実験結果内容>
<試験条件>
  <溶媒>生理的食塩水</溶媒>
  <放置時間>24</放置時間>
  <メッシュサイズ></メッシュサイズ>
  <サンプル量>500CC</サンプル量>
</試験条件>
<測定項目>
  <ストレッチ量>10%</ストレッチ量>
  <アクリロニトリル量>10%</アクリロニトリル量>
  <塩分量>10%</塩分量>
  <エポキシ量>ephr</エポキシ量>
</測定項目>
</実験結果内容>
</実験DATA管理システム>
```

Cyber LUXEON

XMLをパーシングして
階層的なDOMオブ
ジェクトに変換

DOMオブジェクトのまま
内蔵のOODBに格納

メモリ

OBJECT DB

検索 (XPath・XQuery)
追加更新 (XUG)

検索、更新の際はメモリ上に
DOMオブジェクトを展開して
メモリスピードで処理 (XMLキャッシュ)

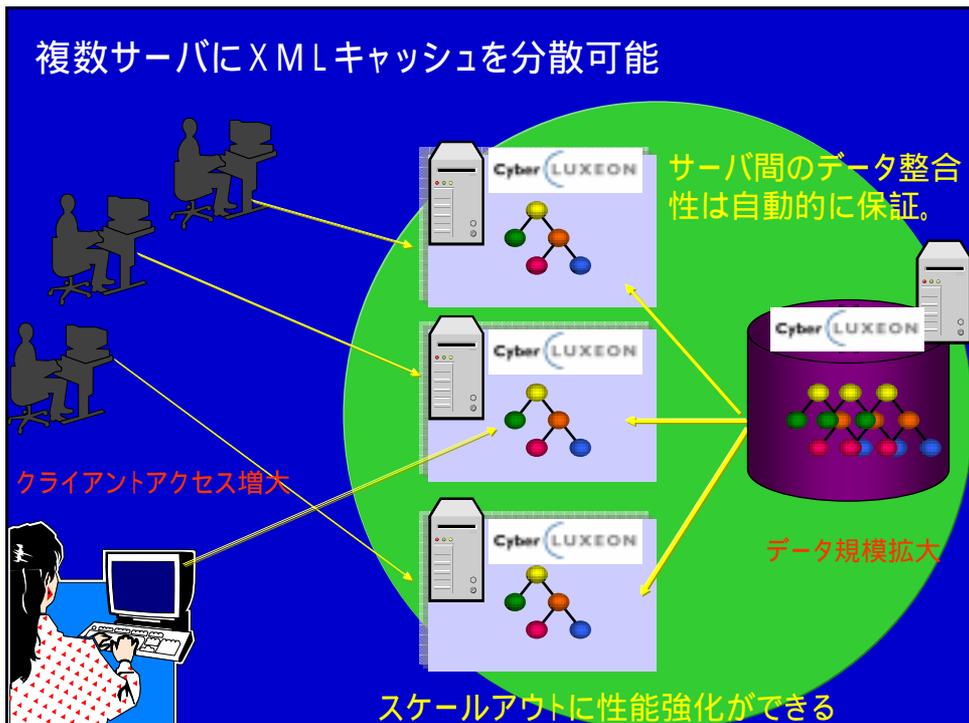


特長2：分散キャッシュ機構

XMLデータはメモリ上に配置され、メモリアクセスでデータ処理が行われる（XMLキャッシュ機構）。

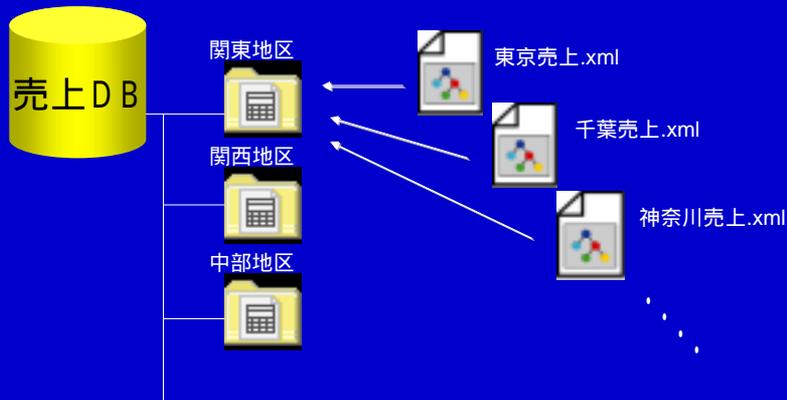
このXMLキャッシュはサーバーを増設することで、パラレル化でき、スケールアウトによる段階的な性能強化が可能となる（分散キャッシュ機構）。

分散キャッシュにおいても、各サーバ間のデータの整合性をとる機構がとられており、複数サーバによる巨大な論理DB空間を形成することが可能。



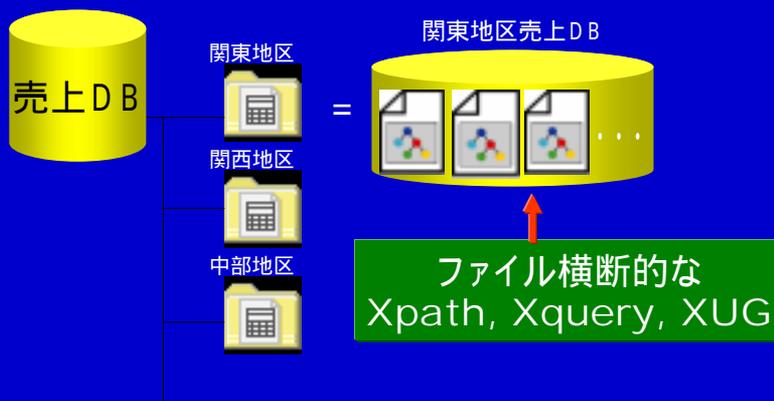
特長3: フォルダDB機構

フォルダーに投入されたxmlファイルを横断してタグ検索、更新、削除できる。これにより階層的なフォルダー一つ一つを仮想的なDBとして扱える。



特長3: フォルダDB機構

フォルダーに投入されたxmlファイルを横断してタグ検索、更新、削除できる。これにより階層的なフォルダー一つ一つを仮想的なDBとして扱える。



ユーザ直言

エンジン整備作業記録の電子化

石川島播磨重工業殿

宇宙航空事業本部 生産管理部
宮下 幹夫様による「XMLDB勉強会
(2007年1月10日)講演より抜粋

システム化の背景

整備作業、指示記録の電子化ニーズ

- ・顧客である航空会社及びその所轄機関へ管理精度向上に向けた取り組みが必要。
- ・管理制度向上のため文章をデータとして扱えるシステムが必要とされた。
- ・従来からEXCEL(VBマクロ)で管理していたが、データ量の増加に対応出来なくなった。

システム化の要件

作業指示記録の電子化ニーズ

階層型DB

整備(指示と記録)情報はエンジンを構成する各階層レベルに対応することが必要であり、必然的に『DBの階層構造化が可能なこと』が条件。

関連システムとのリンク

重複入力を防止するために、現在稼働中システムとのリンクがリアルタイムに可能であること。



XMLDB導入成果

1. 整備情報管理
エンジン構成を表す階層構造は最小で3、多い場合は数十階層に細分化
XMLデータベースは理論的には無制限であり、且つ容易に構造変更が可能
当初は40階層まで検討したが、管理コストを考慮して現在は5階層程度
2. スキーマレスによる項目変更へのフレキシビリティの高さ
短期間での開発においては、確定済み仕様と検討中の仕様を混在させながら進めざるを得ないケースが多い。
今回も同様の状況下での開発を進めたが、項目増減(DB変更)が日程に影響を与えることなく開発を可能にした。
3. リソースの有効活用
既に既存システムの端末が数多く設置されていることから、現有機器を使用することを前提にWEB系システムとし、これと親和性の高いXMLデータベースを採用することでリソースの有効活用を実現。
4. 電子帳票化対応
XML+FO&PDFを使っでの帳票設計・出力・保管が容易に、且つ多種多様な整備記録や作業記録の電子帳票化要求への対応が可能。

今後の計画

1. レスpons向上への取り組み
データの増加と共に、レスponsが悪くなるのはどんなデータベースでも同じ。
構造変更が簡単に出来る強みを生かし、レスpons向上に向けた構造化を検討。
2. 手順書 / 記録の電子化推進
作業現場には多くの機種と、それらに対応した数多くの手順書が存在している。
手順書自体は構造化された文書であり、その内容は大きく「指示」と「記録」。
ドキュメント(指示)とデータ(記録)の統合はXMLの主機能であり、今後発展性のある課題。
3. 既存システムとの再編
稼働中システムは汎用系～サーバー系(RDB/XML)と多岐に亘っており、それぞれの強みを生かした再編(棲み分け)とリンクの推進。
(統合は理想的な姿ではあるが、データの特徴により必ずしもベストではない)

まとめ Cyber Luxeonのポジション

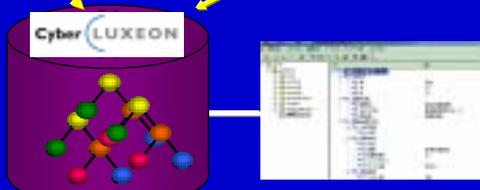


基幹



データ統合の柔軟さ

統合情報ハブ



XML

クライアント

XSL
電子フォーム

データの出し入れの柔軟さ



XMLDBを使うこと

ファイル管理のアバウトさとデータ処理の利便性

の両立

使い途は後から出てくる、みたいな**ブログ**のノリ

ベンダー訪問

参加者とのディスカッション拾い読み

まとめ

XMLDB勉強会の特質と期待

マーケティング系サブグループ

ユーザ事例

帰納

XMLDBの
適用領域

XML自体の
適用領域

演繹

XQuery, XMLDBの特徴

技術系サブグループ

ゆくゆくは

XMLDBを使うシステム造りのメソロジー

をつかっていきたい。

完