SAS 8.2 における OLAPソリューションの紹介

SAS Institute Japan
プロフェッショナルサービス第2部

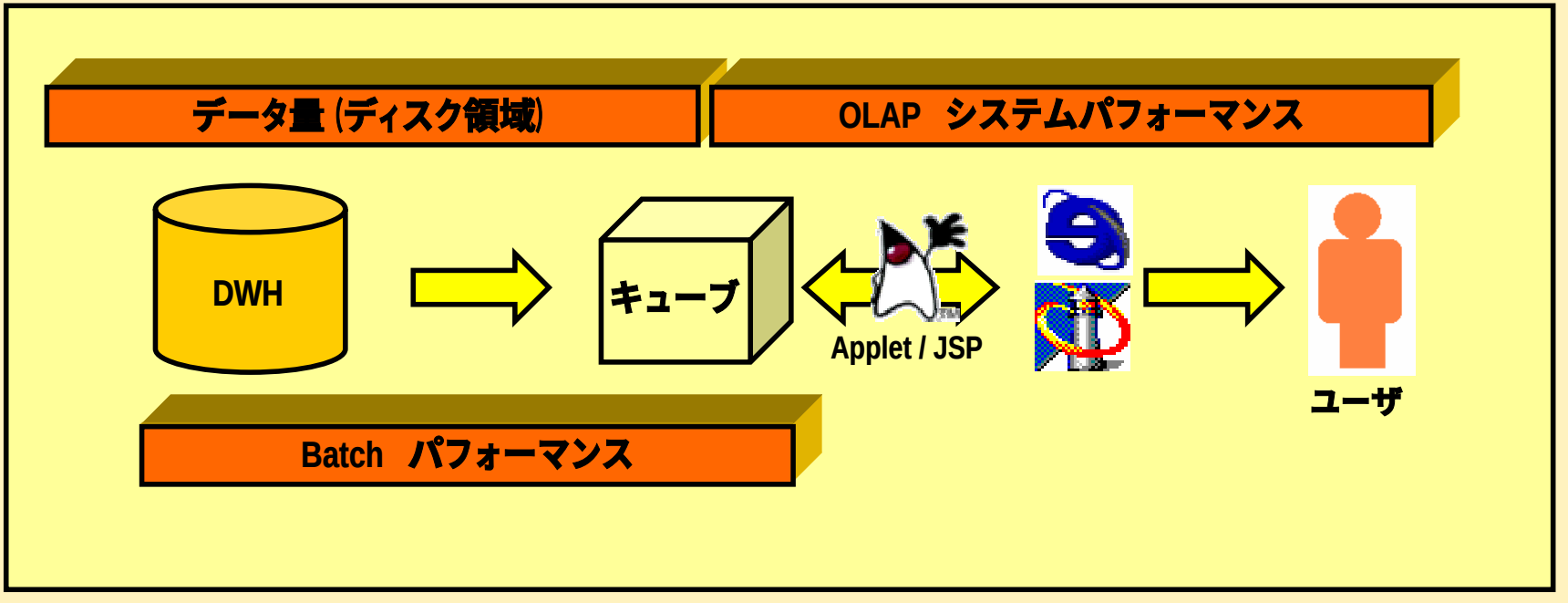
Agenda

1. OLAPシステムの要件
2. 大容量データの弊害
3. SASが提供するOLAP
4. HOLAPの特徴
5. HOLAPの位置付け
6. HOLAPの　キューブ構成手法
7. HOLAPの　キューブ構成手法　その2
8. デモンストレーション（環境）
9. まとめ　HOLAPを適用する時のポイント

Copyright © 2003, SAS Institute Inc. All rights reserved.

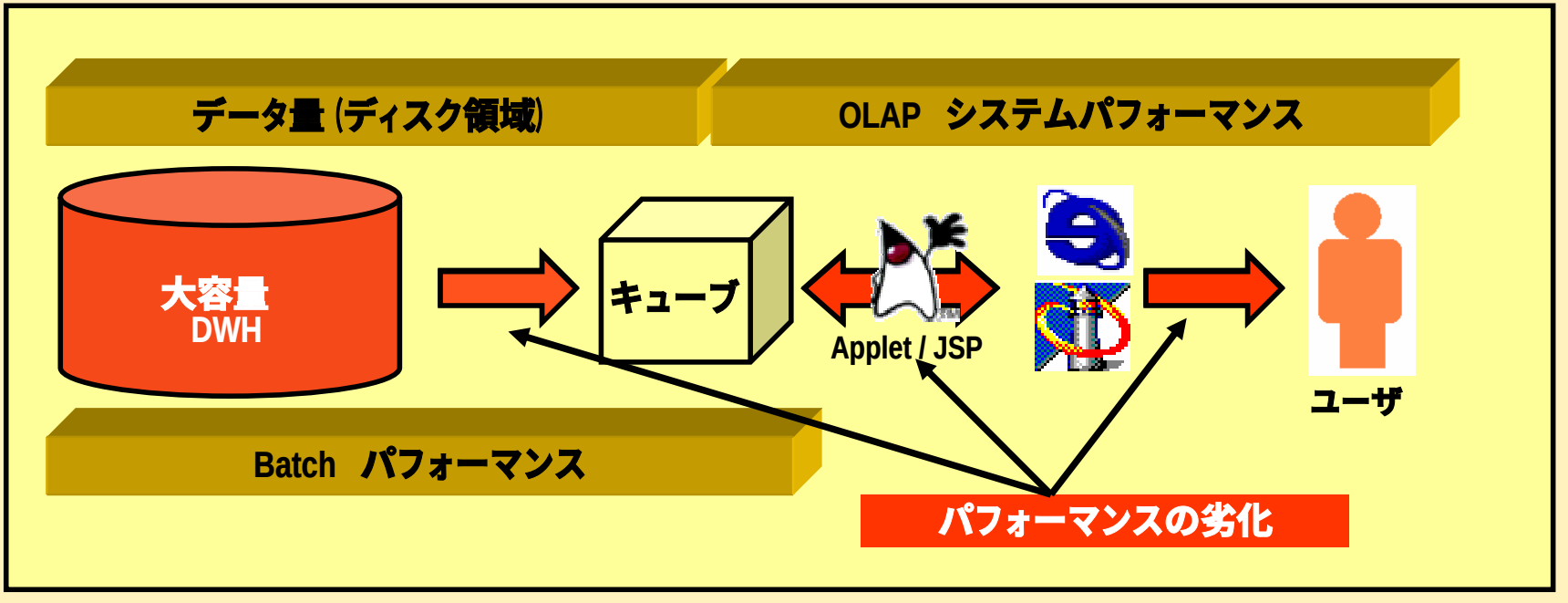
1. OLAPシステムの要件

- モニタリングすべき情報が適切に提供される （情報の網羅性･データ量）
- 適切なアプリケーション パフォーマンス （レスポンス）
- データ抽出からキューブを作成するパフォーマンス （Batch パフォーマンス）

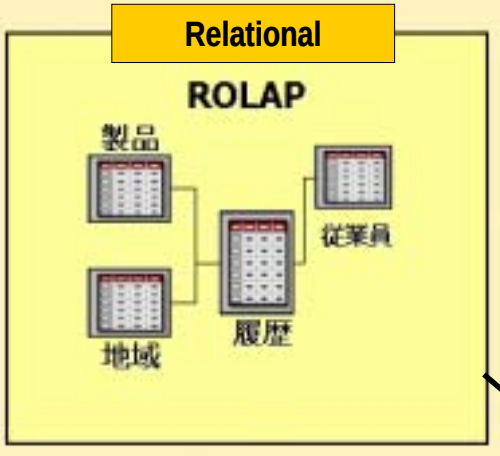


2. 大容量データの弊害

- 網羅すべき情報 (DWH) が大容量の場合はパフォーマンスに弊害あり。
- ディスク領域、システムパフォーマンス、Batchパフォーマンスの劣化。
- 結果として抽出されたキューブを制限せざるをえない。



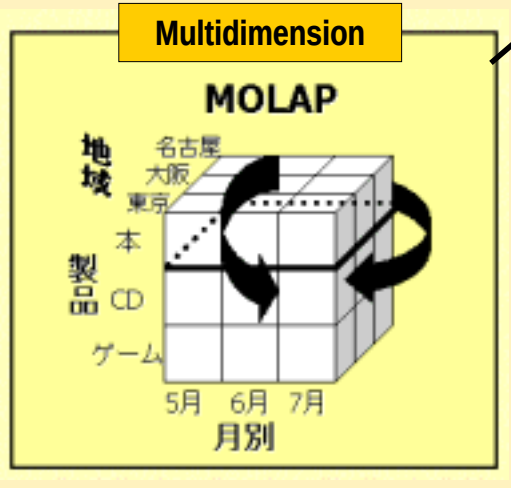
3．SASが提供するOLAP



ROLAPの特徴

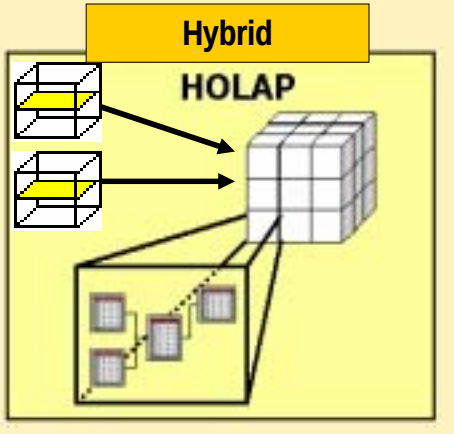
- データの網羅性　：　○
- レスポンス　：　×
- バッチパフォーマンス　：　○

ROLAP　と　MOLAPの
弱点を補完したHOLAP



MOLAPの特徴

- データの網羅性　：　×
- レスポンス　：　○
- バッチパフォーマンス　：　×



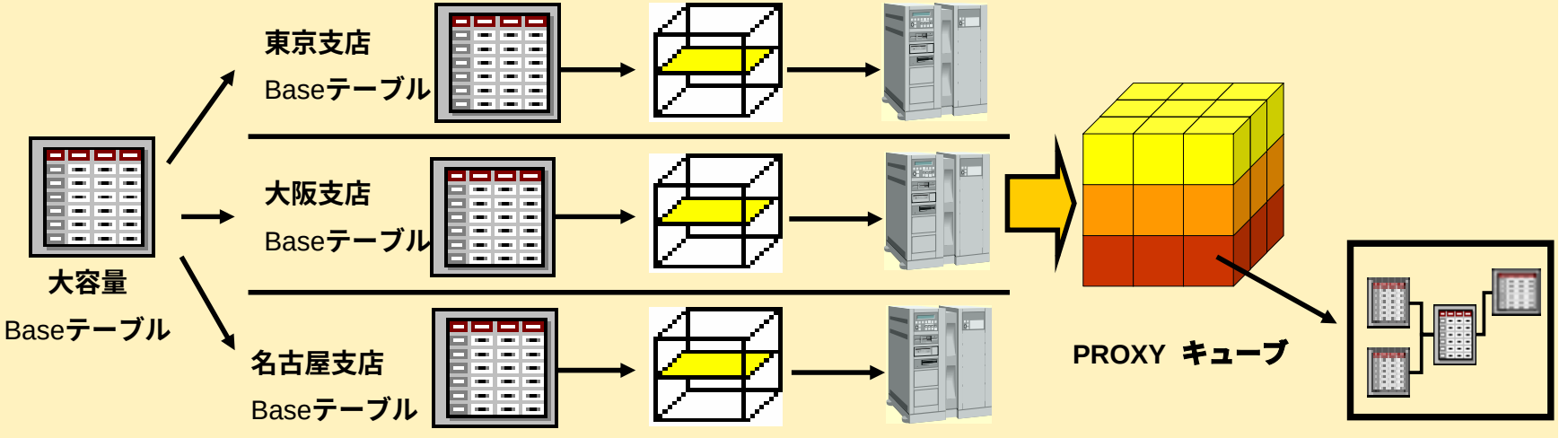
Copyright © 2003, SAS Institute Inc. All rights reserved.

4. SASが提供するHOLAPの特徴

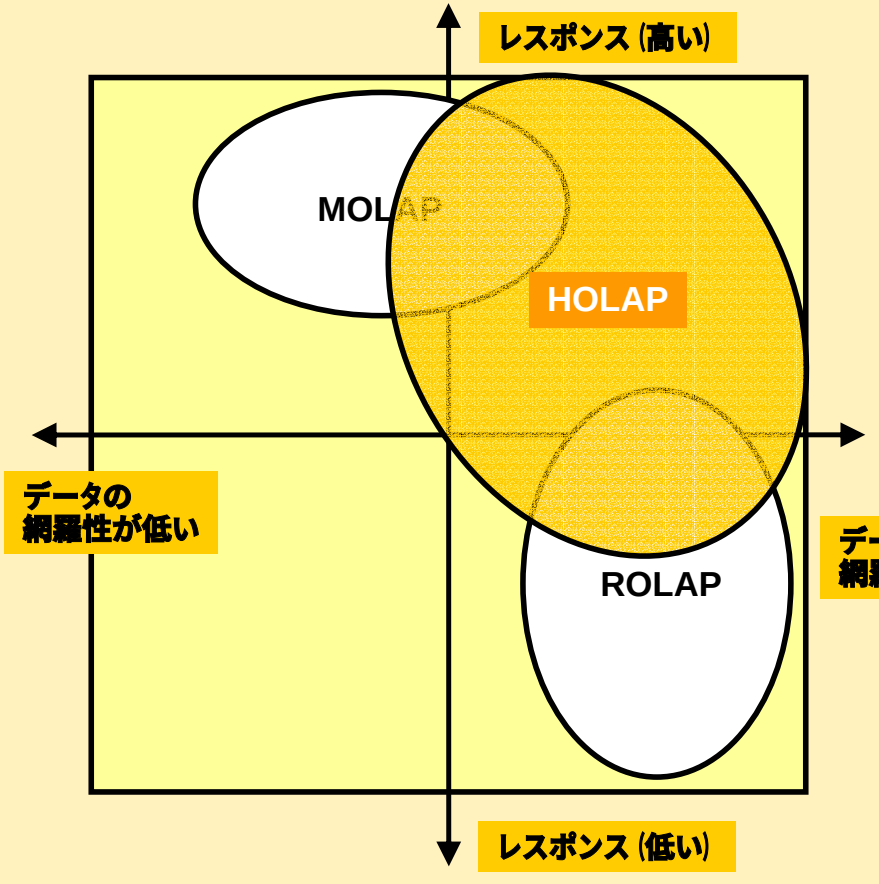
- HOLAP ＝ ROLAP ＋ MOLAP
- 複数のキューブ、テーブルを用いて、仮想的なキューブを作成する (PROXY キューブ)

適用例

- 参照頻度が高く、レスポンスが要求されるような情報はMOLAPで対応する
- レスポンスよりもデータの網羅性が必要とされるものはROLAPで対応する
- 複数のキューブの格納場所 (サーバー) を物理的に分けて、負荷分散させる事が可能
- Racking Stacking 手法によりキューブを分割する事で、より柔軟なHOLAPが提供可能



5. HOLAPの位置付け



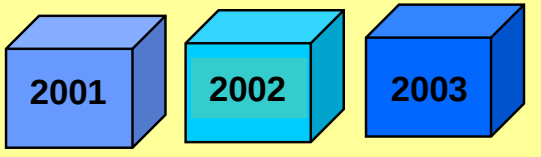
- HOLAPはパフォーマンスとデータ量との適切なバランスを取りうるアーキテクチャを提供
- 要求されるシステムに合致した適用範囲を提供する。
- OLAPシステムの拡張性を意識できる為、ハードウェアの容量や、パフォーマンス向上、階層情報の変更など、OLAPシステムとしての変化（向上）に柔軟に対応

6. HOLAPの キューブ構成手法

- **Racking 手法 1**

  値毎にキューブを細分化する

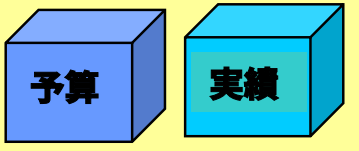
  EX) 2001年 – 2002年 – 2003年



- **Racking 手法 2**

  分析変数毎にキューブを細分化する

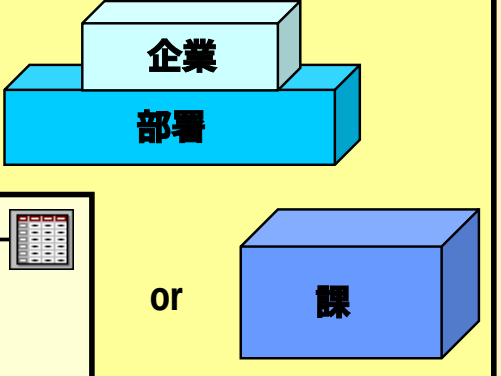
  EX) 予算 – 実績



- **Stacking 手法**

  サブテーブル毎にキューブを細分化

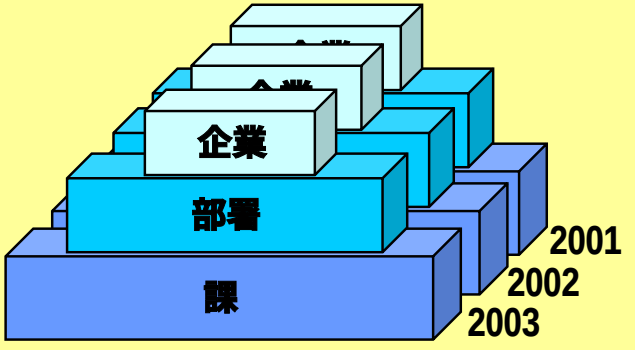
  EX) 企業 – 部署 – 課



7. HOLAPの キューブ構成手法 その2

- **Racking ＋ Stacking**

  より汎用的な軸でモニタリングする必要がある場合はRacking と Stackingを組み合わせて、MDDBを細分化することが可能となる



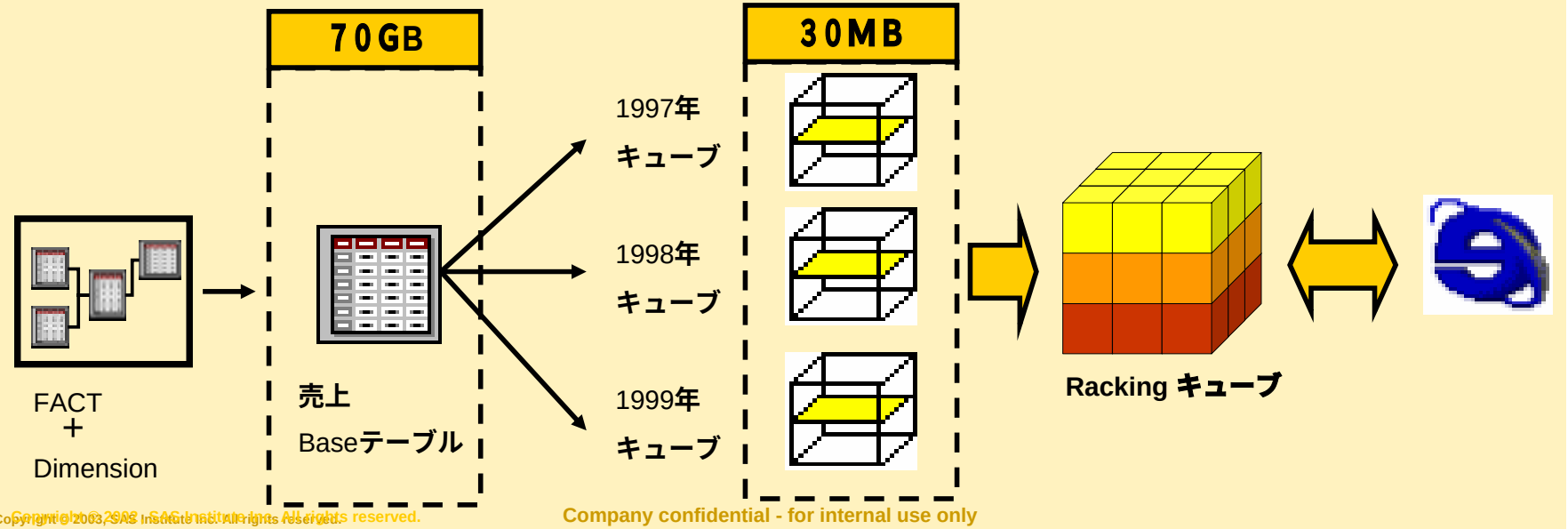
Copyright © 2003, SAS Institute Inc. All rights reserved.

8. デモンストレーション （環境）

・デモシナリオ ： 全国チェーンのスーパーマーケットの売上データ OLAPシステム
・ Baseテーブルサイズ **約5億オブザベーション （約70 GB）**

OS ：Windows2000 Professional　　メモリ: 512 M
CPU : 850 MHz　　SAS: Base SAS 8.2

・Factテーブル＋Dimensionテーブル → Baseテーブル作成時間 **約 3．5時間**
・Baseテーブル → 複数のキューブ作成時間 **約 3時間** （直列処理）



9. まとめ　HOLAPを適用する際のポイント

- OLAPシステムとしての拡張性を考慮すること
- 一度作成して終わりではなく、コストとサービスのトレードオフを常に意識し、トライアンドエラーでアプリケーションの精度を高める事が重要

Copyright © 2003, SAS Institute Inc. All rights reserved.