

文部科学省  
「革新的実行原理に基づく超高性能データベース  
基盤ソフトウェアの開発」  
の活動進捗状況

平成21年12月1日

文部科学省 研究振興局 情報課

# 「情報の巨大集積化と利活用基盤技術開発」連携施策群における位置づけ

## 目標

Web上及び非Web上にある大量かつ多様な情報を、個人が簡便、的確、かつ安心して収集、分析できる

次世代の情報検索・情報解析技術基盤を構築する。

2012年までに、情報爆発時代に対応した超巨大データの戦略的活用を可能とする。

非Web空間(実空間)

Web(サイバー空間)

[2009年度からの連携枠組み]

### 必要性

- Web情報の価値の拡大
- Web情報の学術研究、経済活動への高度活用に対する期待

- Web情報分析のための技術の発展が必要

### 情報の活用

ビジネス、生活、教育・研究、医療等

### Web社会分析プロジェクト

- Web情報(テキスト、動画、画像、音声等)を効率よく収集してアーカイブを形成
- 蓄積したWeb情報の高度な分析

過去の情報を含めた高度な分析

### 情報の分析

新たな分析手段

連携を拡充

情報の信憑性を検証

[2008年度までの連携枠組み]

### 情報大航海プロジェクト

- 巨大集積化した情報を再利用できるように仕分け
- 整理整頓して器に格納

分析手段

情報大航海の一部成果も活用し、研究開発を効果的に推進

### センサ情報の社会利用コンテンツ化プロジェクト

- プライバシー情報を管理
- センサ情報のコンテンツ化

### 超高性能データベースプロジェクト

処理の効率を格段に向上

### 情報信憑性検証技術プロジェクト

- 格納された情報の純度を上げるべく精錬
- 純度の保証付きでユーザーに提供

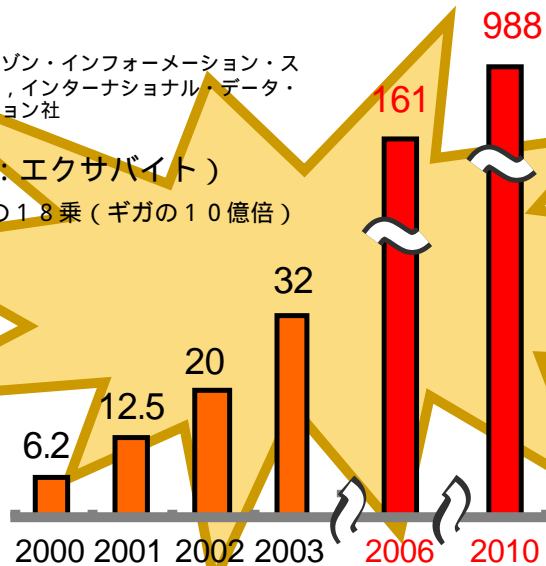
# 非順序型実行原理に基づく超高性能データベースエンジンの開発

## 背景と目的

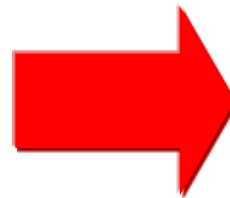
出典：ホライゾン・インフォメーション・ストラテジー社、インターナショナル・データ・コーポレーション社

(単位：エクサバイト)

エクサ：10の18乗（ギガの10億倍）



IT利用の急速な進展により、社会の中で生み出される情報量は従前と比べ飛躍的に増大し続けている



**情報爆発時代の到来**

我が国が新規産業・高付加価値産業を継続的に創出し続け、国際競争力を維持・強化していくとともに、国民生活の安全及び安心を確保するためにも、巨大なデータを高速に処理し、戦略的に活用することが不可欠。

# 研究開発の概要

## 課題

膨大な（エクサバイト級）データを高速に処理できるデータベースシステムが必要

プロセッサやディスクドライブなどのハードウェア技術による高性能化のみではデータベースシステムの性能を飛躍的に向上させることが困難

è データベース実行原理の革新による性能ブレークスルーが必要

## 研究開発内容

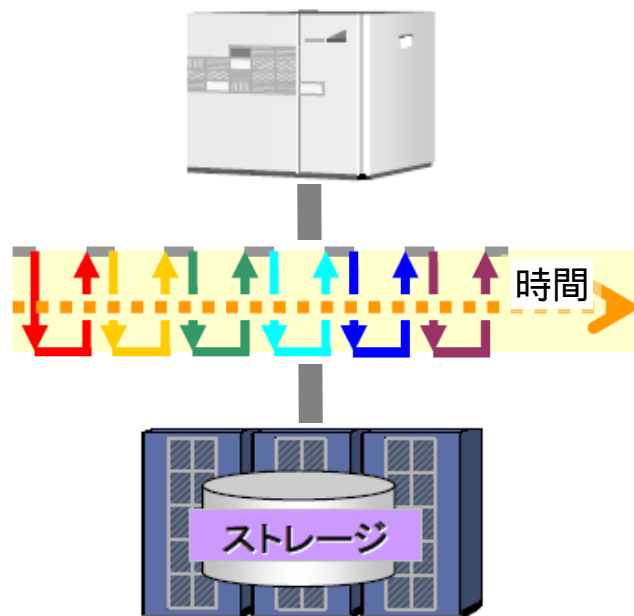
- 革新的な非順序（アウトオブオーダー）型実行原理に基づく超高性能データベースエンジンを開発するため、以下の研究開発を実施する。

- (1) 革新的実行原理である「非順序型データベース実行原理」の確立
- (2) 上記(1)に基づくデータベース基盤ソフトウェアの設計・実装
- (3) 実アプリケーションによる有効性の実証

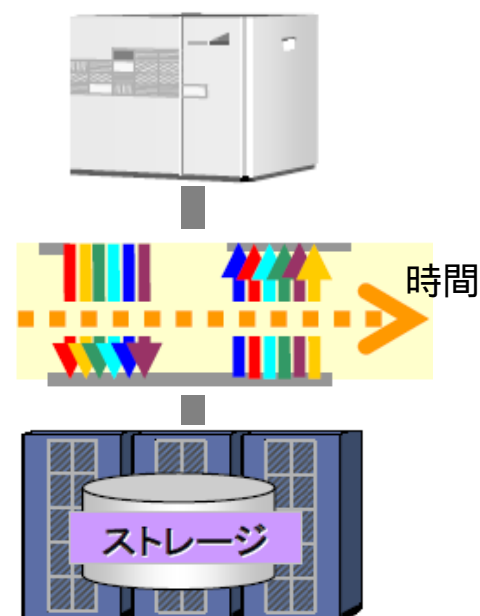
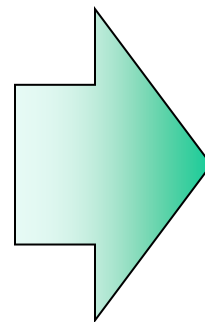
# 非順序型データベースエンジンとは

## 特 徴

- 超大量非同期IO発行機構 (大量のデータ要求を同時に発行できる機能)
- 実行時動的IOスケジューリング機構 (大量に発行されたデータ要求を最適に処理するための機能)
- ストレージ駆動型アウトオブオーダ実行機構 (到着した順序でデータの解析処理をすることができる機能)



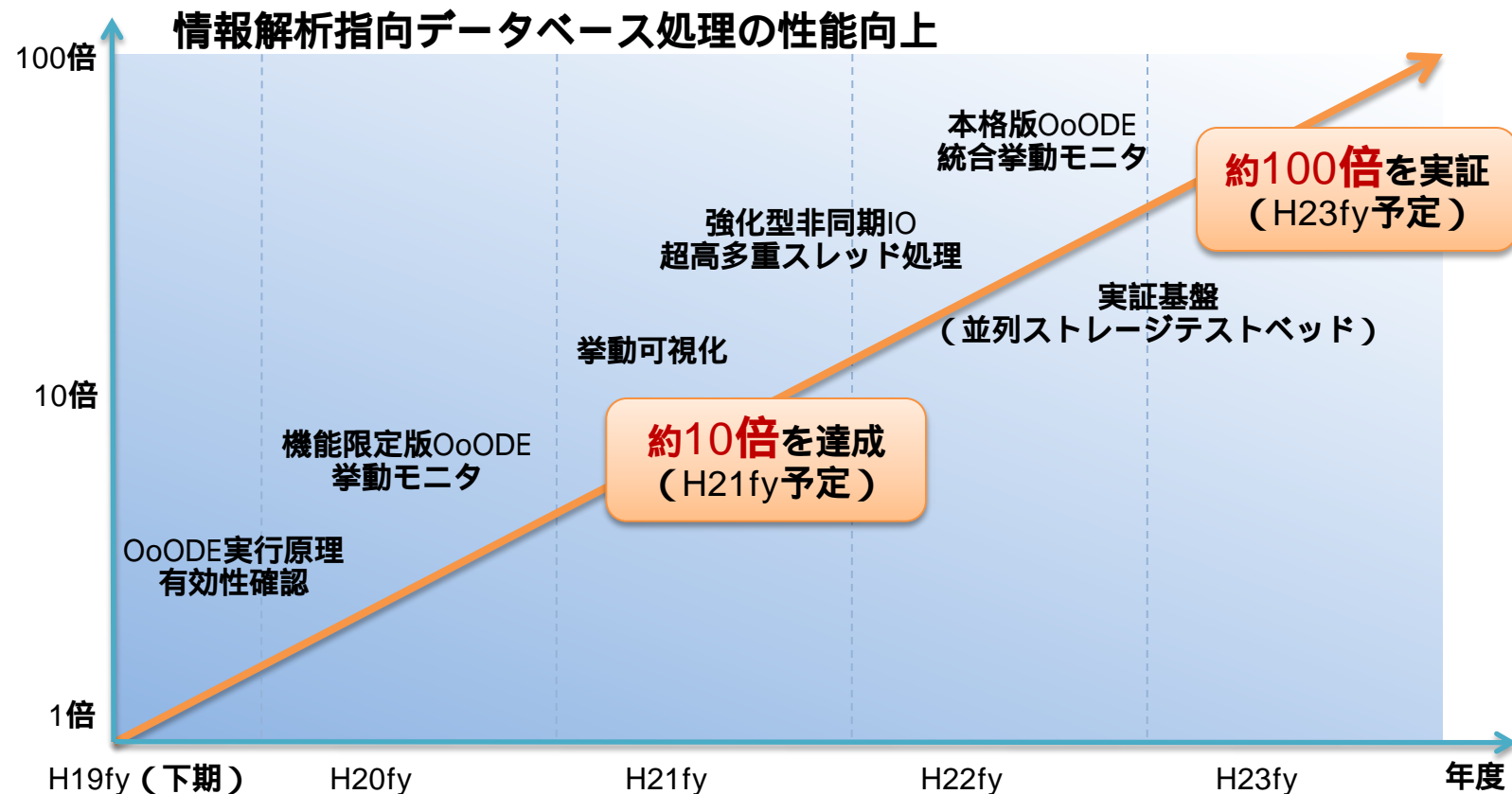
従来型データベースエンジン



新型データベースエンジン (OoODE)

# 研究開発の成果目標

- 平成21年度中には、非順序型実行原理を一部のデータベース演算に適応したデータベースエンジンを開発し、従来の技術の約10倍のデータベース解析処理性能を達成する。
- 平成23年度中には、非順序型実行原理を本格的に適応したデータベースエンジンを開発し、従来の技術の約100倍のデータベース解析処理性能を達成する。



# 期待される効果

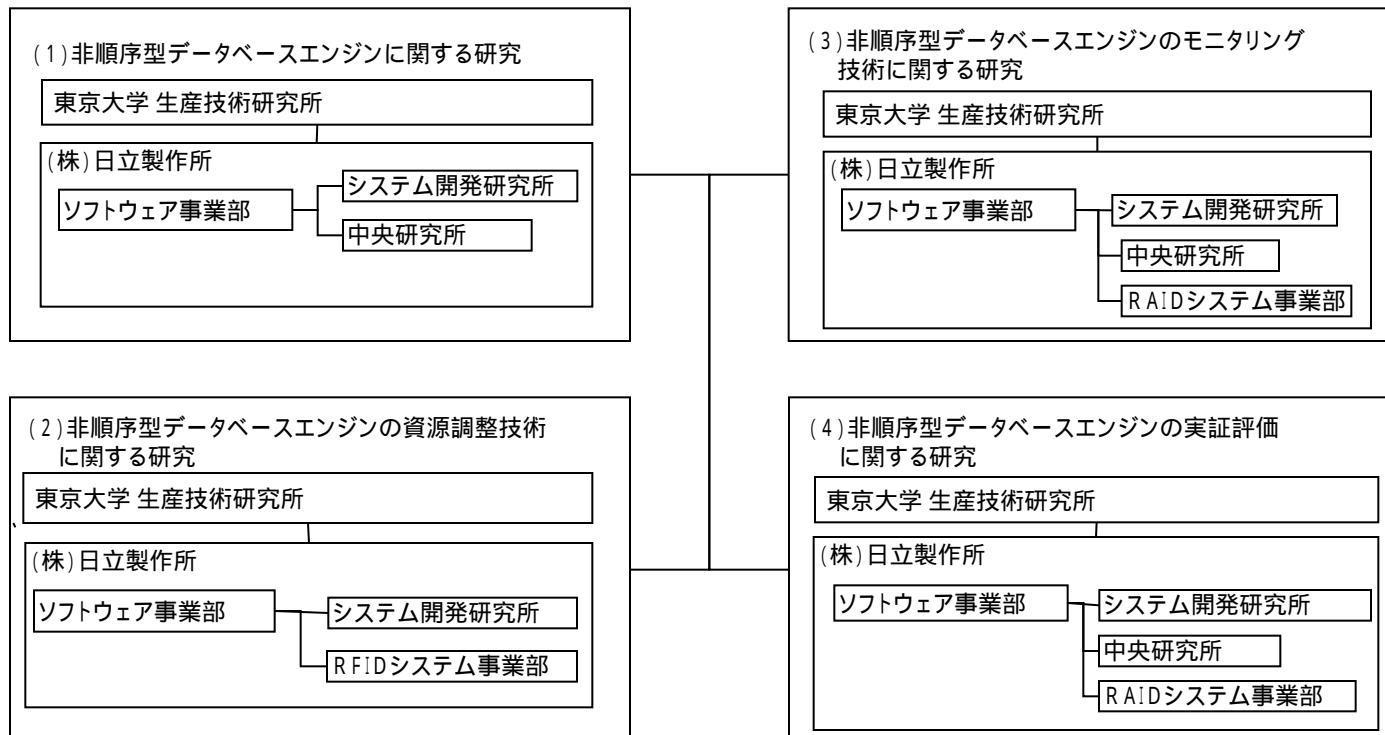
- データベースを活用している様々な分野(特に金融や小売等のビジネス分野)において、本施策で開発した超高性能なデータベースエンジンを活用することにより、これまで以上にデータベースを戦略的に活用することが可能となり、我が国産業の国際競争力の向上が期待できる。
- また、例えばトレーサビリティや、センサ技術等の分野で利用することにより、国民生活の安全・安心の向上にも寄与することが期待される。
- 更に、世界に先駆けて非順序型データベース実行原理を確立することにより、データベース管理システムなどの基盤ソフトウェア産業の国際競争力の向上が期待できる。

# 研究開発体制

東京大学生産技術研究所を研究拠点として、東京大学と日立製作所の産学連携体制により実施する。

研究代表者：喜連川 優 教授

(東京大学生産技術研究所)





# これまでの主な進捗状況

## 【平成19年】

7月20日(金) 採択課題(事業実施機関)の決定・公表

## 【平成20年】

1月30日(水) 連携施策群「情報の巨大集積化と利活用  
基盤技術開発連携群の活動」シンポジウム

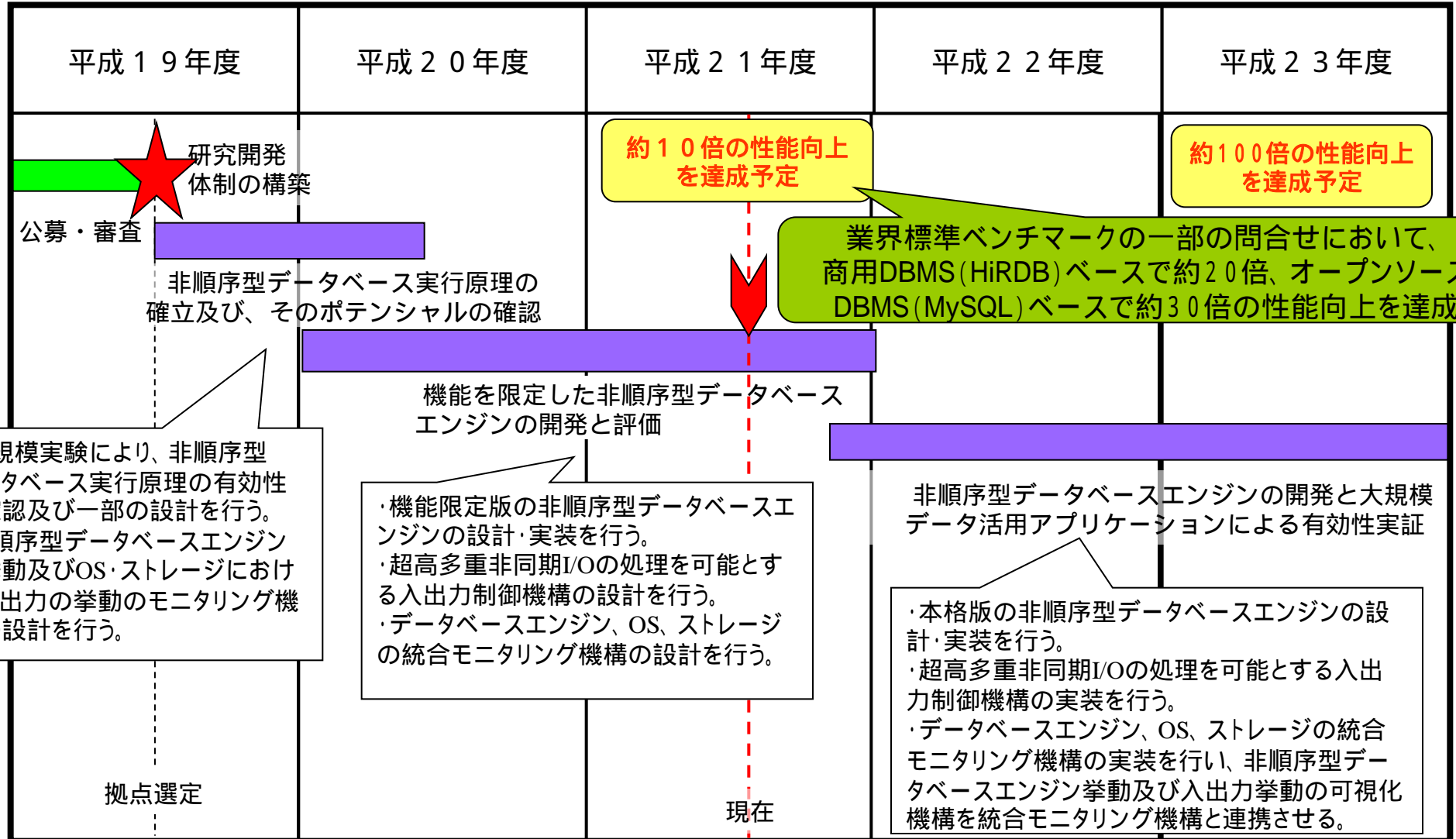
## 【平成21年】

2月 4日(水) 連携施策群「情報の巨大集積化と利活用  
基盤技術開発連携群の活動」シンポジウム

6月15日(月) 情報科学技術委員会 中間評価 審議

6月16日(火) 研究計画・評価分科会 中間評価 報告

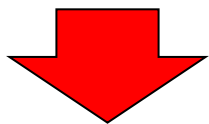
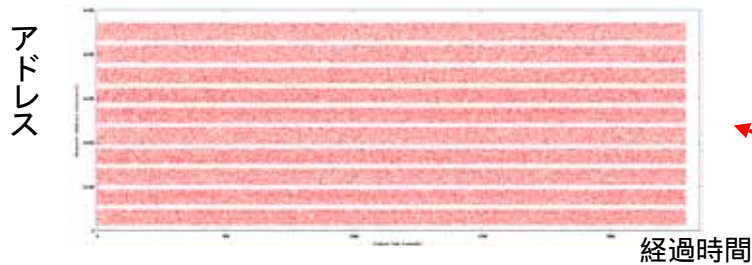
# 研究開発計画



# オープンソースDBMS (MySQL) をベースとする 非順序型データベースエンジンの開発

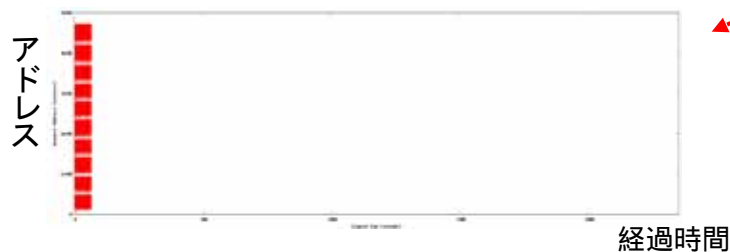
オープンソースDBMS (MySQL) をベースとした非順序型データベースエンジンの開発を進め、部分的ではあるがマルチコア環境への対応させた。その性能を評価したところ、業界標準ベンチマークの一部の問合せにおいて、**約32倍の高速化を達成した**。

順序型データベースエンジン (MySQL)

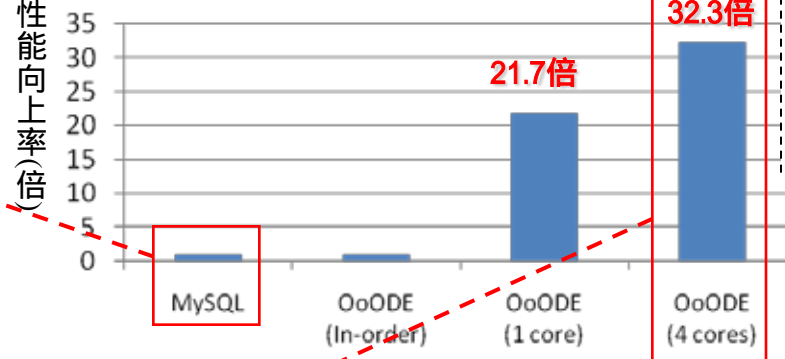


約1/32に短縮

非順序型データベースエンジン (OoODE(4cores))

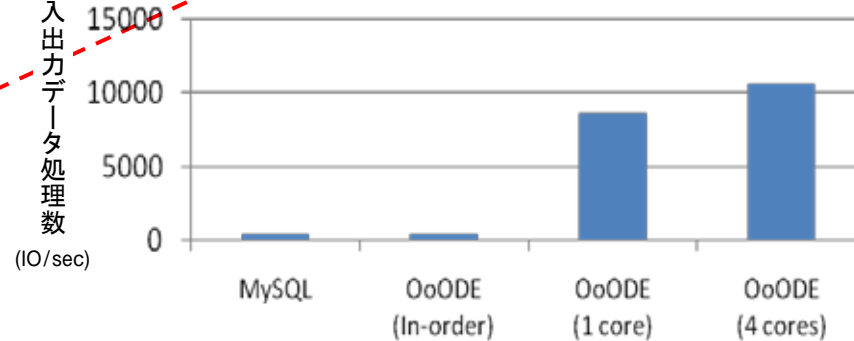


性能向上率(倍)



[評価環境]  
 CPU: Quad-core Xeon Processors x 4  
 メモリ: 32GB  
 HDD台数: 20台  
 (10000rpm 146GB FC(RAID-0))  
 OS: RedHat Enterprise Linux 5.3  
 データ: TPC-H SF=100 (約100GB)  
 処理: TPC-H Q.14相当

入出力データ処理数  
(IO/sec)



# 商用DBMS (HiRDB) をベースとする 非順序型データベースエンジンの開発

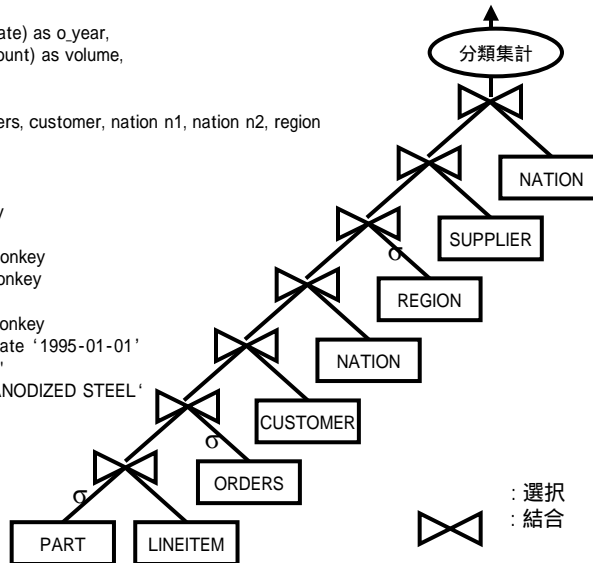
商用DBMS (HiRDB) をベースとした非順序型データベースエンジンの開発を進め、業界標準ベンチマークの複雑度の高い問合せにおいて、約20倍の高速化を達成した。年度内にベンチマークのうち半数の問合せをサポートする予定である。

## 業界標準ベンチマーク TPC-H Query 8

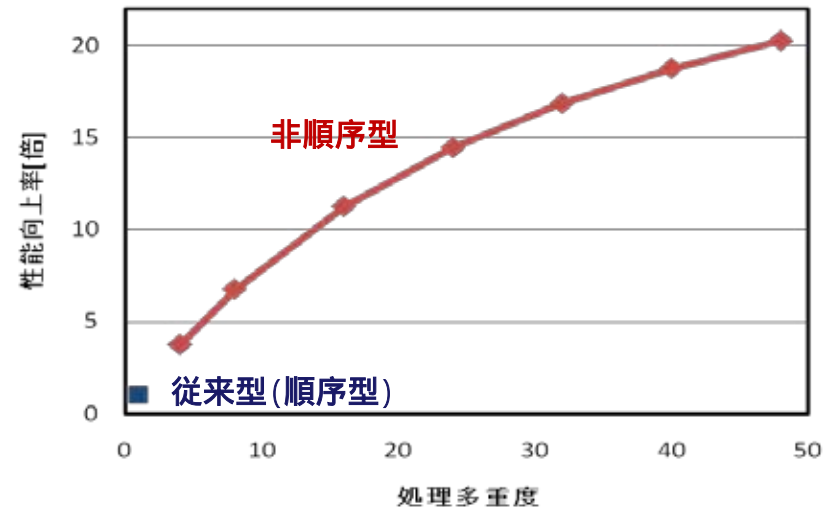
指定地域における指定商品種別の2年間の市場シェア変化を分析

```

SELECT o_year,
       sum(case when nation = 'BRAZIL' then volume else 0 end) / sum(volume) as
mkt_share
FROM (
  SELECT
    extract(year from o_orderdate) as o_year,
    l_extendedprice * (1-l_discount) as volume,
    n2.n_name as nation
  FROM
    part, supplier, lineitem, orders, customer, nation n1, nation n2, region
  WHERE
    p_partkey = l_partkey
    and s_suppkey = l_suppkey
    and l_orderkey = o_orderkey
    and o_custkey = c_custkey
    and c_nationkey = n1.n_nationkey
    and n1.n_regionkey = r_regionkey
    and r_name = 'AMERICA'
    and s_nationkey = n2.n_nationkey
    and o_orderdate between date '1995-01-01'
    and date '1996-12-31'
    and p_type = 'ECONOMY ANODIZED STEEL'
    and p_size < 3
  ) as all_nations
GROUP BY o_year
ORDER BY o_year;
  
```



## 従来型 (順序型) エンジンに対する 非順序型エンジンの性能向上率

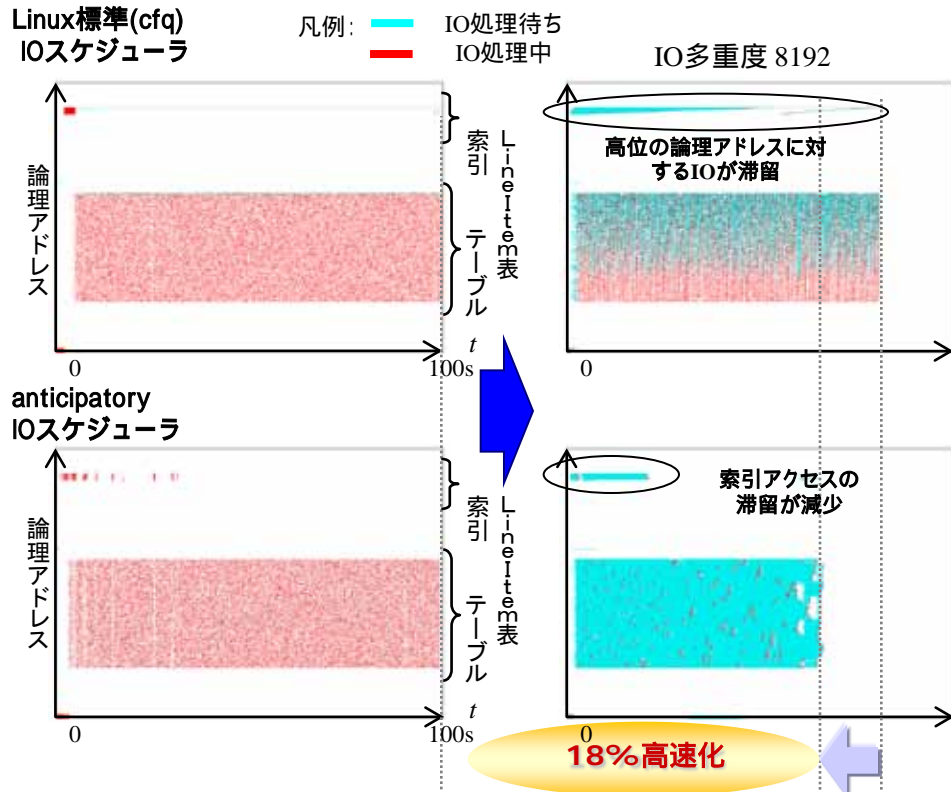


プロセッサ Intel Xeon 2.66GHz  
メモリ 32GB  
ストレージ HDD (100Krpm) 20台  
データセット TPC-H SF=1000 (1TB相当)

# モニタリング機構の設計・開発

非順序型データベースエンジンならびにOSの微細挙動を観測するモニタリング機構の開発を進めている。OSにおける入出力処理機構の精緻な解析によって従来知られていないLinuxスケジューラにおける非同期入出力に対する特異挙動を解明したほか、データベースエンジン・OSに渡る多層的な挙動観測によってタスク間の長い排他待ちの解析を可能とするなど、データベースエンジン開発の大幅な効率化を実現している。

## 入出力スケジューラの特異挙動の解析



## タスク間排他待ちの解析

