

MySQLベースのカラム指向DB

InfiniDB® でビッグデータを活用してみる

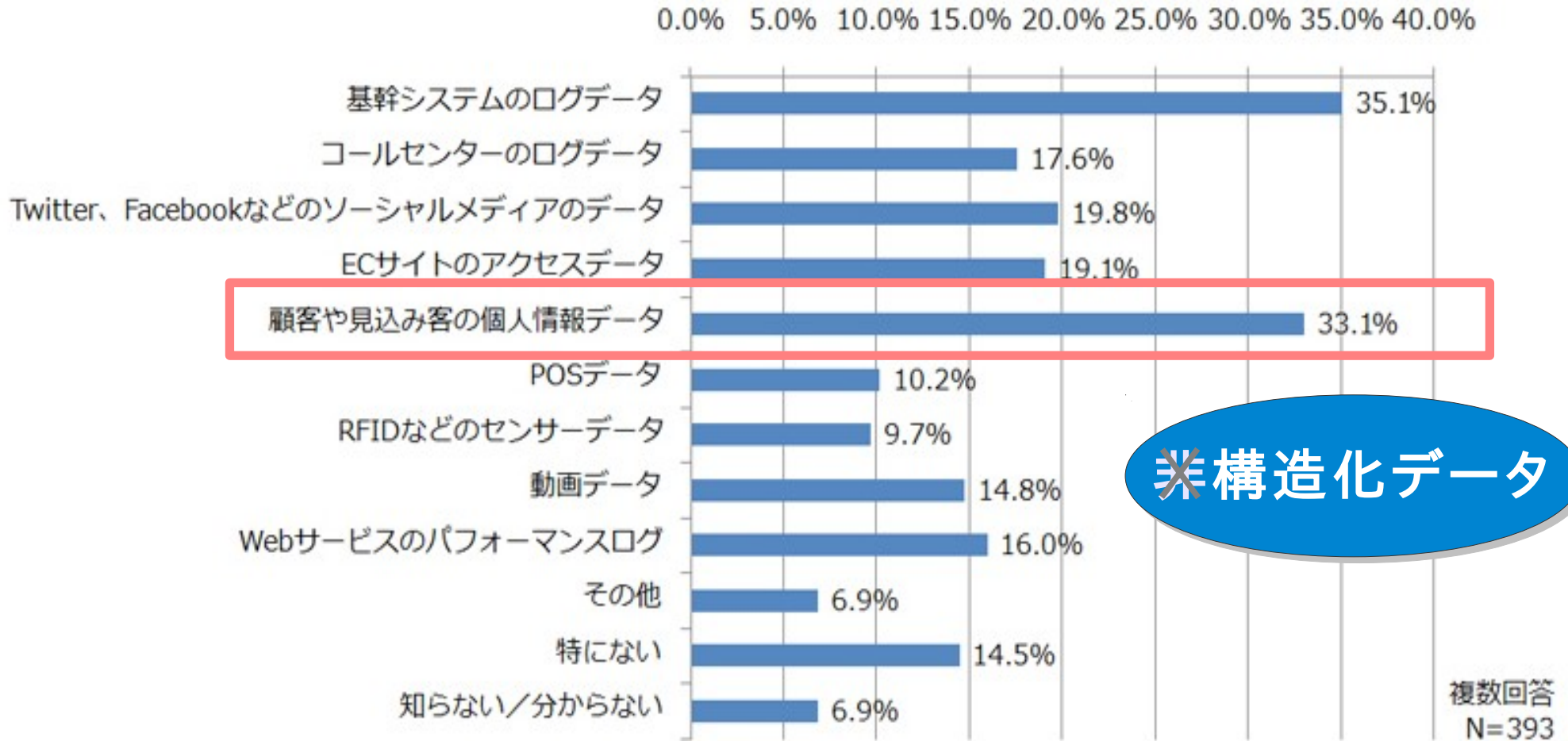
Scalable. Fast. Simple.

株式会社アシスト
花谷 俊英

Fast

Simple

Scalable



出典: TechTargetジャパン「大規模データ(ビッグデータ)活用」に関する調査レポート
<http://techtarget.itmedia.co.jp/tt/news/1201/31/news03.html> (有効回答: 474)

- 汎用DBMSの限界
 - 決められた時間以内に処理が終わらない
 - DWH系向けに最適化されていない
- 手が届かないDWHアプライアンス
 - 数千万円～数億円の世界
 - 使いこなすためのノウハウがない、
技術者がいない



汎用DBMSでDWH系の 処理を実行してみる

スター・スキーマ

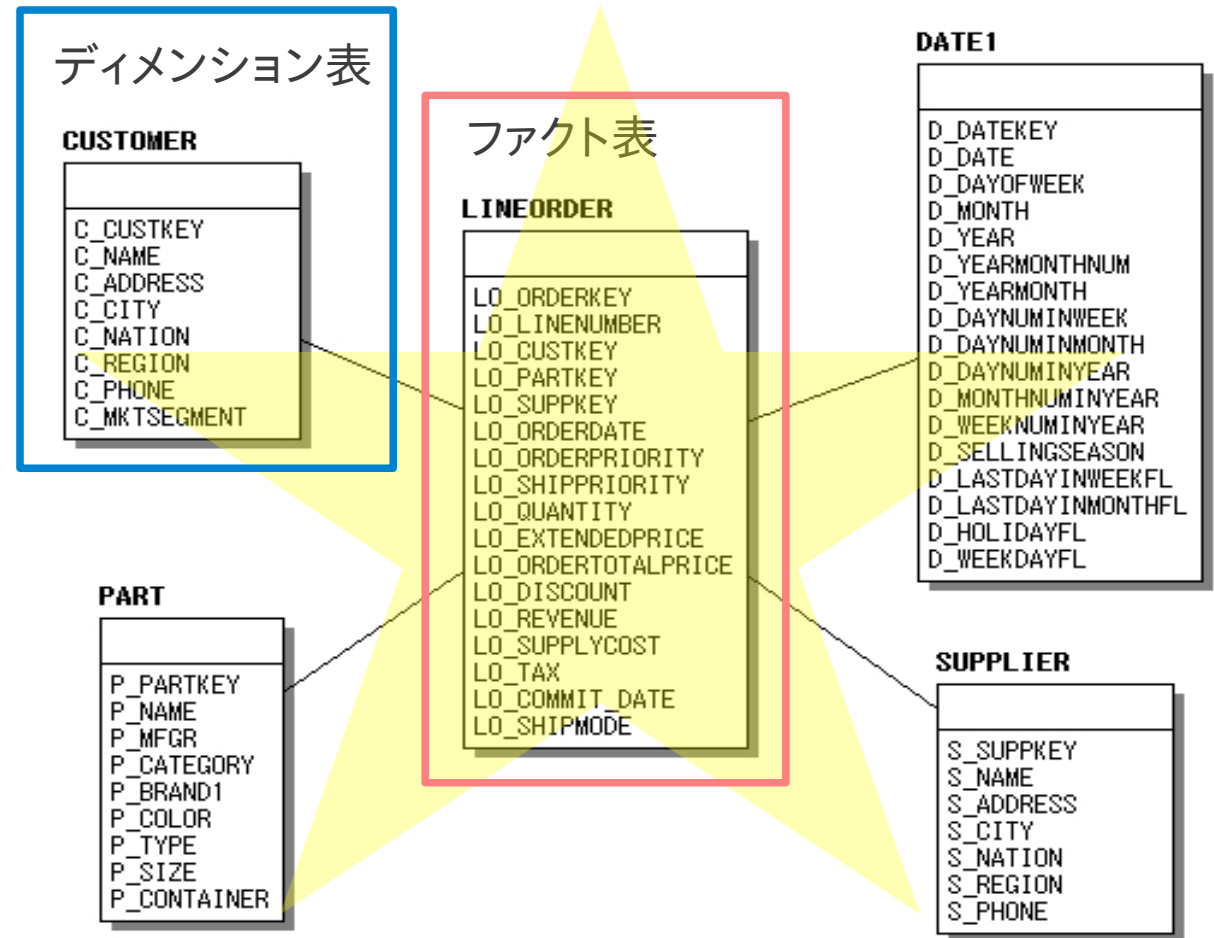
- ファクト表

- 分析対象になる
売上データ
- 件数は億単位

- デイメンション表

- マスタ
- 分析の軸になる
製品名などのデータ
- 件数は少ない

Star Schema BenchmarkのER図



- 多数の分析軸

- ファクト表と複数の
ディメンション表を結合
- 分析軸のグループ処理

- 広範囲のアクセス

- 返される結果は数件だが、
大半のデータにアクセス
(例)売上高の集計など
- ファクト表の列数は多いが、
SQLで指定されるのはごく一部

Star Schema BenchmarkのSQL

```
select sum(lo_extendedprice *
lo_discount) as revenue
from lineorder, date
where lo_orderdate = d_datekey
and d_year = 1993
and lo_discount between 1 and 3
and lo_quantity < 25;
```

```
select sum(lo_revenue), d_year, p_brand1
from lineorder, date, part, supplier
where lo_orderdate = d_datekey
and lo_partkey = p_partkey
and lo_suppkey = s_suppkey
and p_brand1 = 'MFGR#2221'
and s_region = 'EUROPE'
group by d_year, p_brand1
order by d_year, p_brand1;
```

- 行単位での処理

- 処理対象ではない列にもアクセス



- 行単位でブロックに格納
 - 不要な列でも、ブロック内にあれば強制的にI/O



1ブロック(ページ) : 8KB~32KB

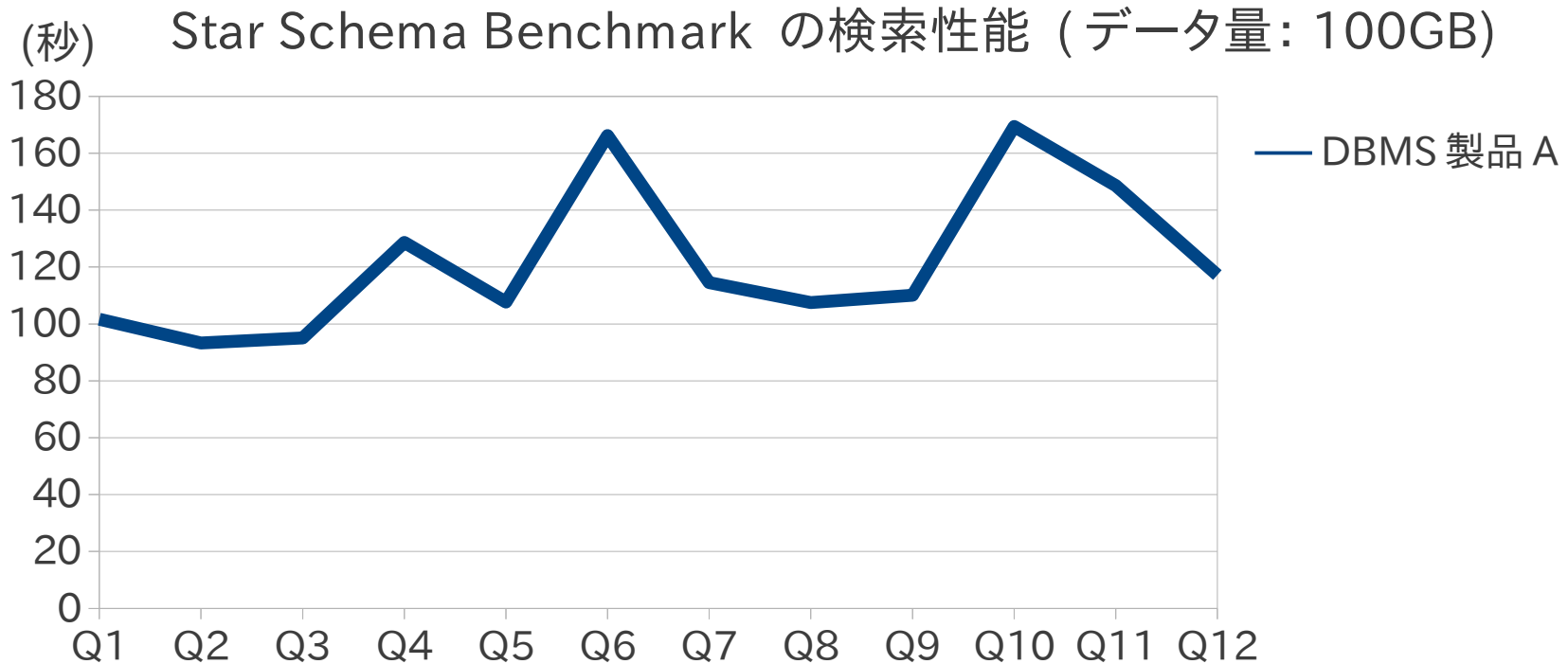
1ブロックに格納できる行数 (ヘッダを除く概算)

ブロックサイズ ÷ 行サイズ

ブロックサイズが8KBの場合
20バイトの行を400行格納できる

1億行のデータにアクセスするには
25万ブロックのI/Oが必要

- 各クエリの応答時間が長く、バラつきがある



CPU : Xeon 5675 (3.06GHz 6コア)×2
RAM : 96GB
DISK : 600GB x 20 (FC)



汎用DBMSでDWH系の 処理をチューニングしてみる

- 索引

- 何を条件に検索するか決まっている場合には有効
- × 分析軸が多数あるため、検索条件を予測できない

- SQLヒント

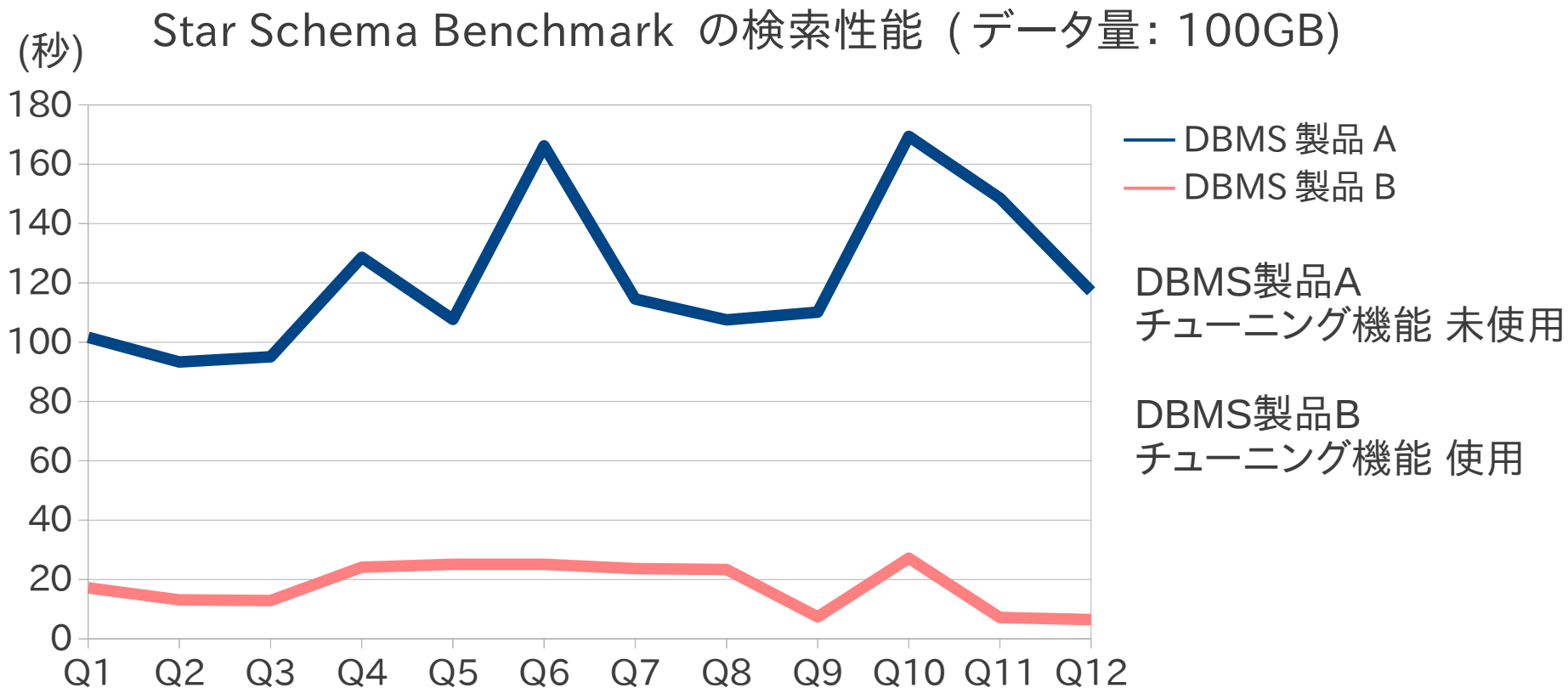
- 実行計画を強制的に変更することが可能
- × 大量I/Oが前提なので実行計画を変更しても効果が薄い

- データ圧縮


- 読み込みブロック数を減らすことが可能
- × 全ての列を解凍するため、CPU使用率が上昇

- パーティショニング
 - 読み込みブロック数を減らすことが可能
 - ✗ キー設計が難しく、分布に偏りがあると効果がでない
- パラレル処理
 - 複数プロセスでの並列処理が可能
 - ✗ ディスクI/Oがボトルネックになると効果がでない

- チューニング機能の有無で5倍以上の性能差



- アーキテクチャ上、DWH系処理に強くない
- チューニング機能の活用が必須
 - 分析軸を考慮した索引設計
 - メモリサイズを考慮した圧縮設定
 - 分布を考慮したパーティション設計
 - コア数に合わせたパラレル度設定

- 
- 効果を出すためには綿密な設計とテストが必要
 - 性能を維持するため、運用中のメンテナンスが必要
 - DBMSごとにチューニング機能の有無、費用が異なる

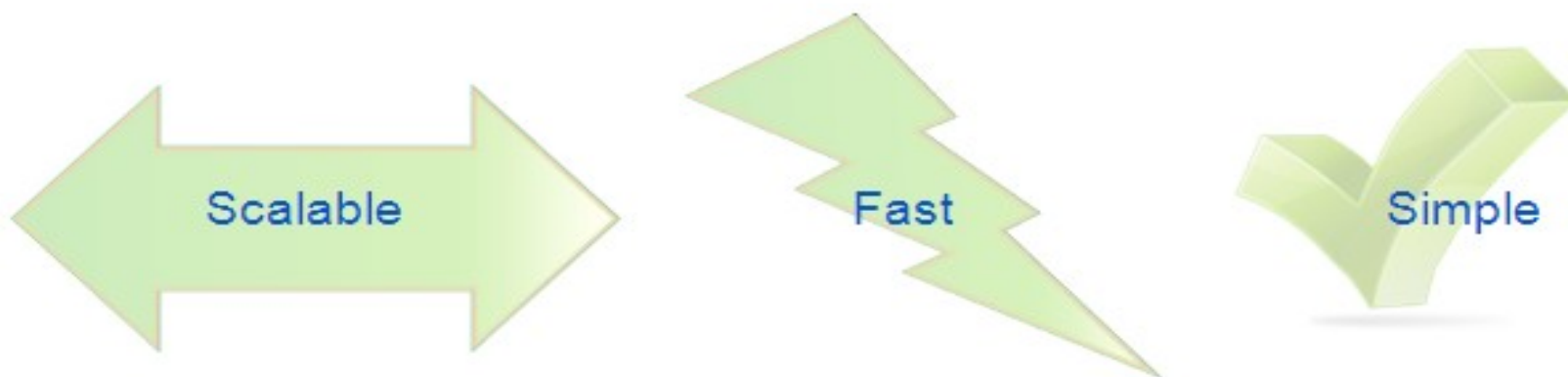
InfiniDB® で

Scalable. Fast. Simple.

DWH系の処理を実行してみる

その力 無限大

- DWH系処理に特化した列指向DBMS
 - MySQLベースのインスタンス + 独自のストレージエンジン



スケーラブル

1台のサーバからスタート可能。スケールアウトにより性能がリニアに向上

高速

汎用DBMSと異なる仕組みでデータを効率的に処理し、H/W性能を限界まで引き出す

シンプル

汎用IAサーバで構築可能。チューニングが不要なため運用負荷を大幅に軽減

- 列単位での処理
 - 処理対象の列にのみアクセス

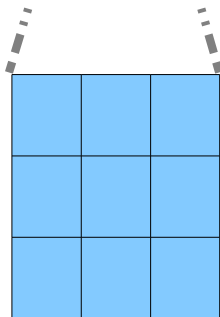
SQLの対象列

注文番号	顧客ID	製品ID	価格	数量	注文日
1	501	1001	¥90,000	50	2001/1/1
2	200	1001	¥90,000	20	2001/1/1
3	350	1003	¥70,000	10	2001/1/1
4	100	1001	¥90,000	5	2001/1/1
5	50	1002	¥65,000	30	2001/1/2
.....					
100000000	254	1001	¥90,000	20	2011/12/31

- 列単位でブロックに格納
 - 不要な列を含むブロックはI/Oされない

注文番号	顧客ID	製品ID	価格	数量	注文日
1	501	1001	9000	50	2001/1/1
2	200	1001	9000	20	2001/1/1
3	350	1003	7000	10	2001/1/1
4	100	1001	9000	5	2001/1/2
5	50	1002	6000	30	2001/1/2
6	125	1004	5500	10	2001/1/3
7	300	1005	5000	1	2001/1/3
8	144	1010	1000	10	2001/1/3
9	198	1001	9000	12	2001/1/3

エクステント
(8MB~64MB)



■ データブロック (8KB)

- 各エクステントにデータの最大値と最小値を格納
- 索引の作成、メンテナンス、統計情報収集が不要

(例) ある特定の顧客ID (例: 930) を参照する場合



- a列、b列を対象とした参照処理

```
SELECT sum(a) , avg(b) FROM 表A WHERE a < 条件値 and b > 条件値;
```

エクステント	a列	b列	c列	d列
1
2
3
4
5
6
7
8
9
10
11
12
13















- 参照対象ではないc列、d列を除外  の部分

```
SELECT sum(a) , avg(b) FROM 表A WHERE a < 条件値 and b > 条件値;
```

エクステント	a列	b列	c列	d列
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		

- a列の条件値の範囲外を除外  の部分

```
SELECT sum(a) , avg(b) FROM 表A WHERE a < 条件値 and b > 条件値;
```

エクステント	a列	b列	c列	d列
1		...		
2		...		
3		...		
4		
5		
6		...		
7		...		
8		
9		...		
10		...		
11		...		
12		
13		

- b列の条件値の範囲外を除外  の部分

```
SELECT sum(a) , avg(b) FROM 表A WHERE a < 条件値 and b > 条件値;
```

エクステント	a列	b列	c列	d列
1		...		
2		...		
3				
4	...			
5	...			
6				
7		...		
8	...			
9				
10				
11				
12		
13		

- 重複した範囲を除外  の部分

```
SELECT sum(a) , avg(b) FROM 表A WHERE a < 条件値 and b > 条件値;
```

エクステント	a列	b列	c列	d列
1	緑	オレンジ	紫	紫
2	緑	オレンジ	紫	紫
3	緑	青	紫	紫
4	オレンジ	青	紫	紫
5	オレンジ	青	紫	紫
6	緑	青	紫	紫
7	緑	オレンジ	紫	紫
8	オレンジ	青	紫	紫
9	緑	青	紫	紫
10	緑	青	紫	紫
11	緑	青	紫	紫
12	紫	紫
13	紫	紫

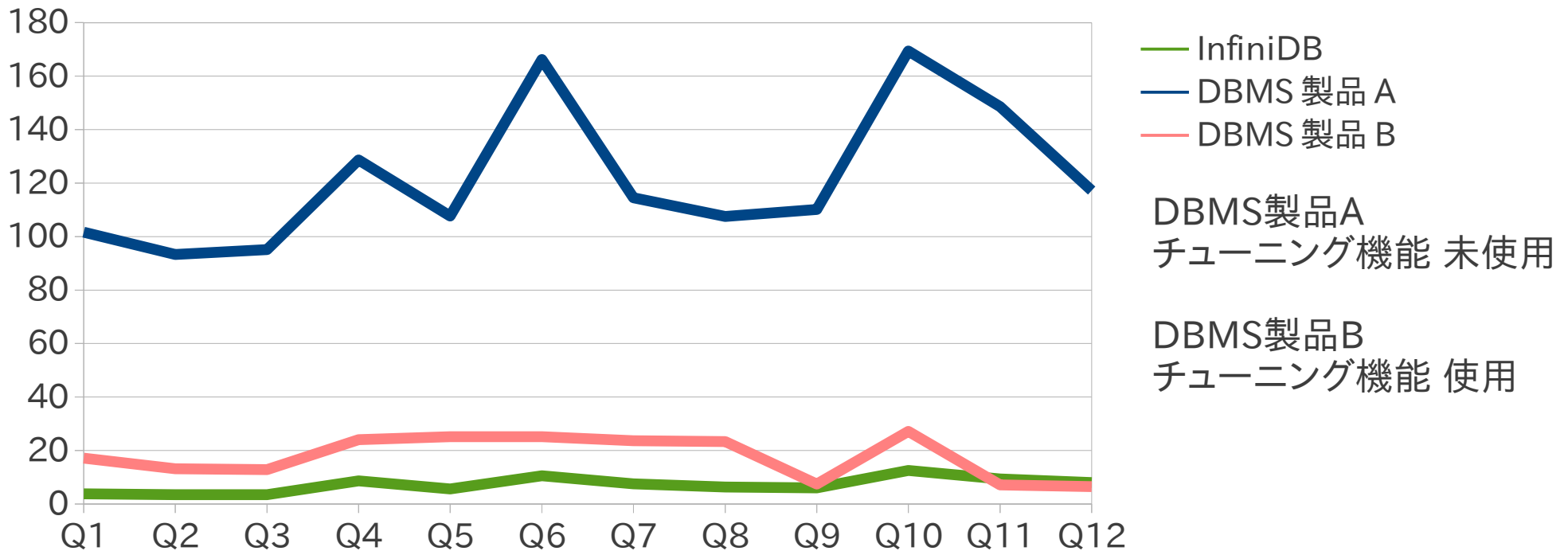
- 絞り込まれたエクステントのみアクセス

```
SELECT sum(a) , avg(b) FROM 表A WHERE a < 条件値 and b > 条件値;
```

エクステント	a列	b列	c列	d列
1	緑	オレンジ	紫	紫
2	緑	オレンジ	紫	紫
3	緑	青	紫	紫
4	オレンジ	青	紫	紫
5	オレンジ	青	紫	紫
6	緑	青	紫	紫
7	緑	オレンジ	紫	紫
8	オレンジ	青	紫	紫
9	緑	青	紫	紫
10	緑	青	紫	紫
11	緑	青	紫	紫
12	アクセス対象		紫	紫
13	アクセス対象		紫	紫

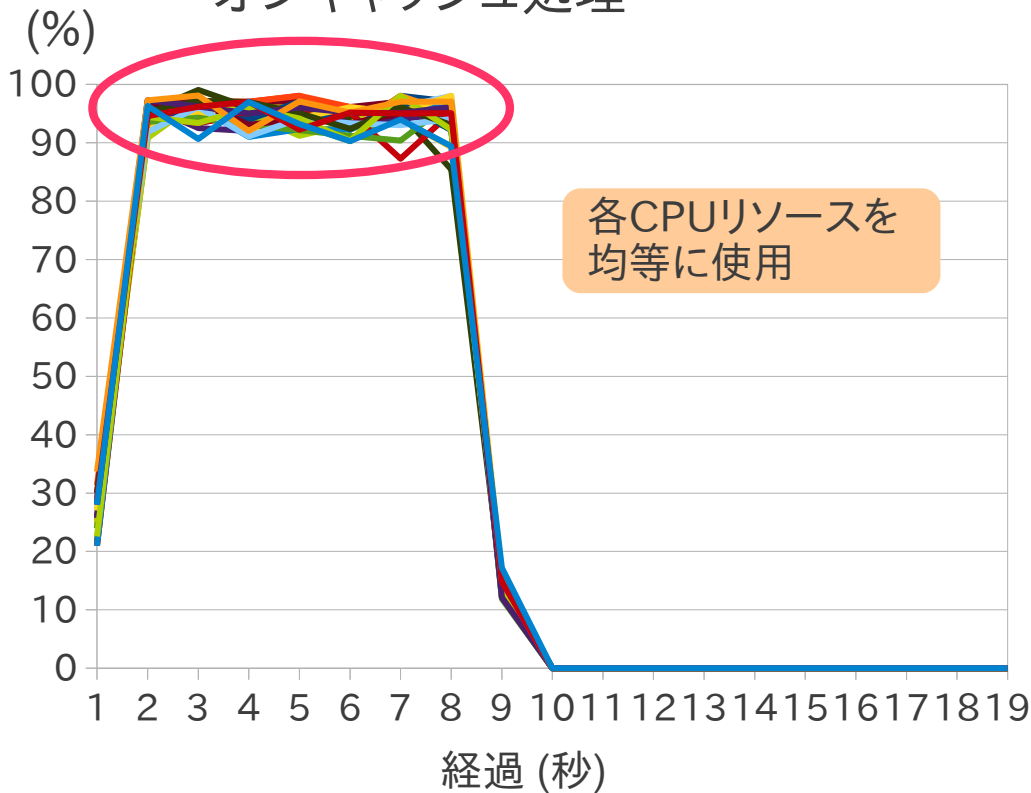
- チューニング不要で汎用DBMSより高性能
 - 列指向 + 索引を使用しない絞り込みの効果

(秒) Star Schema Benchmark の検索性能 (データ量: 100GB)

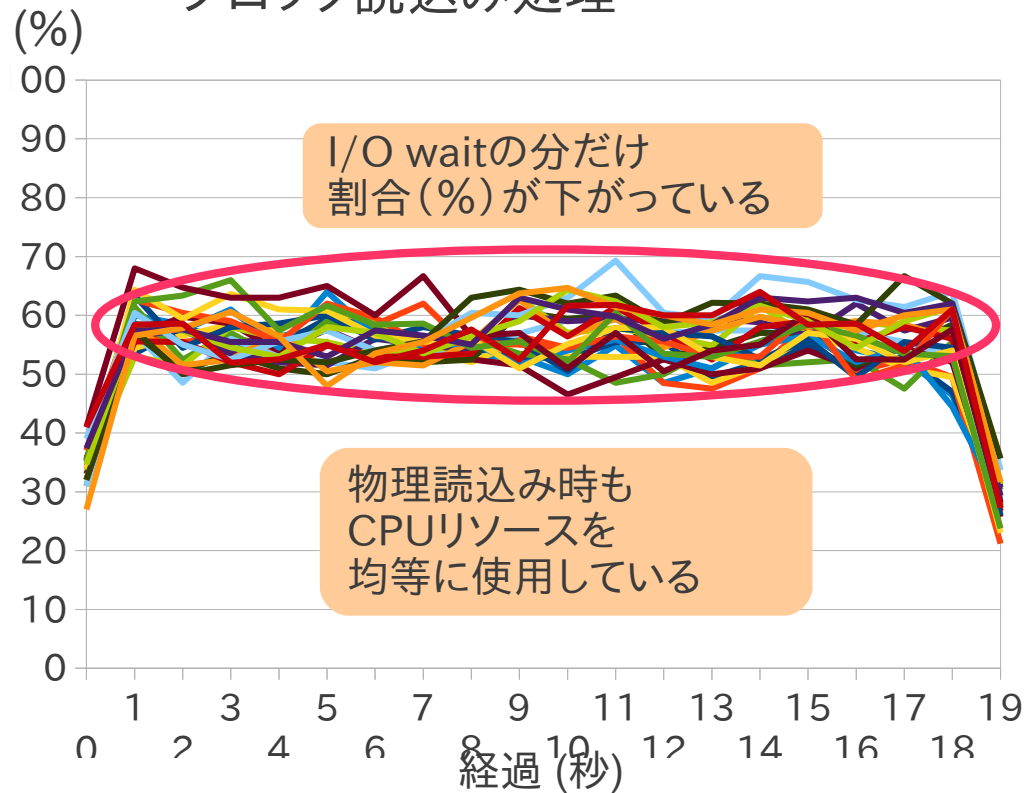


- 全てのCPUリソースを均等にフル活用

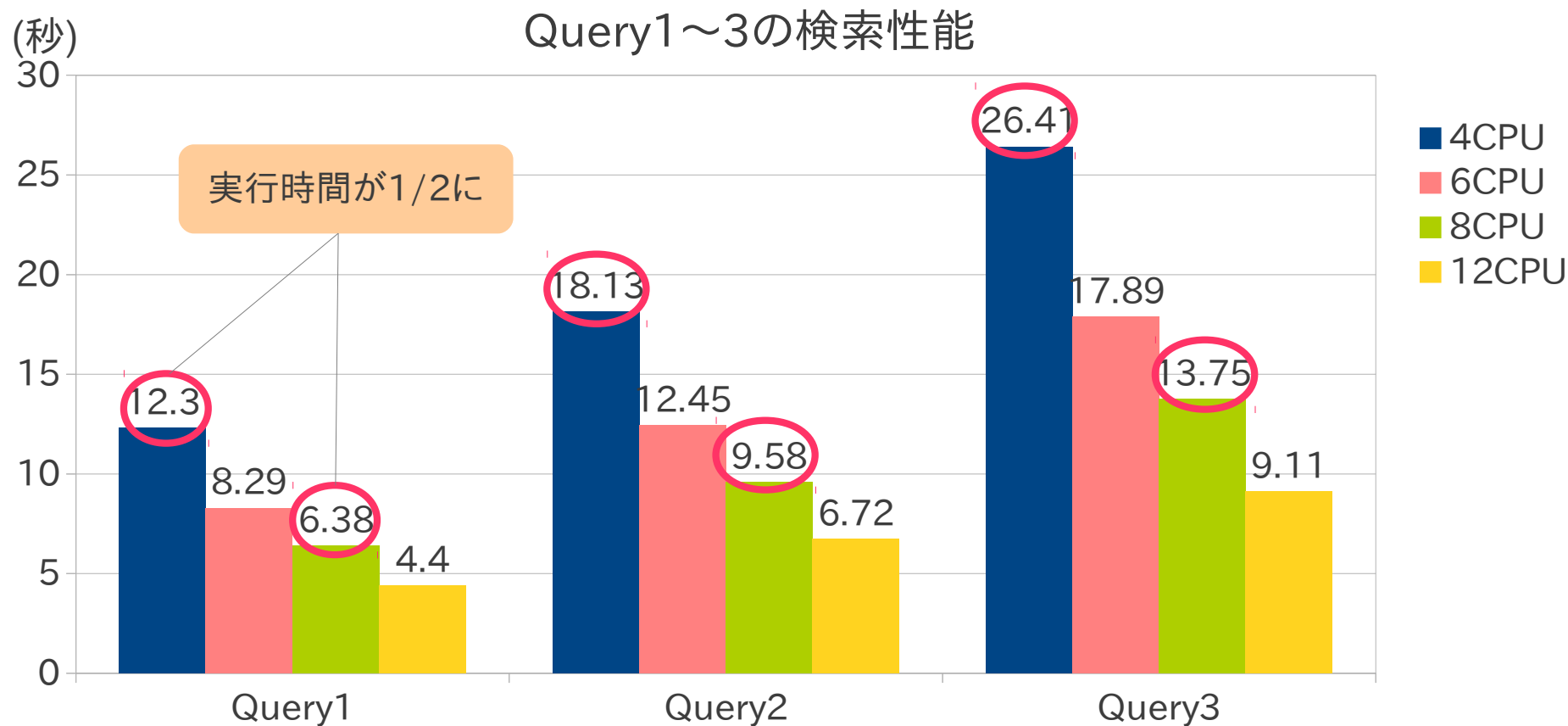
InfiniDBのCPU使用率 (%usr)
オンキャッシュ処理



InfiniDBのCPU使用率 (%usr)
ブロック読み込み処理



- CPUリソースに比例して、応答性能が向上



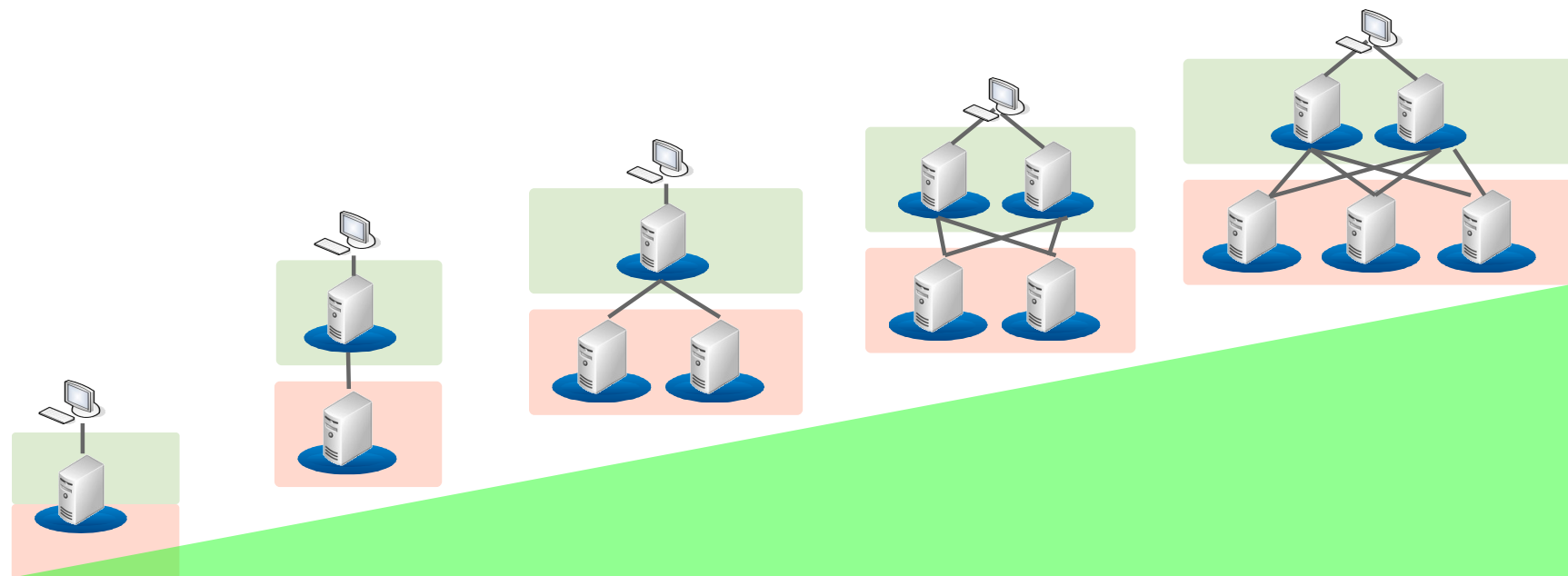
InfiniDB® で

Scalable. Fast. Simple.

DWH系処理を更に高速化してみる

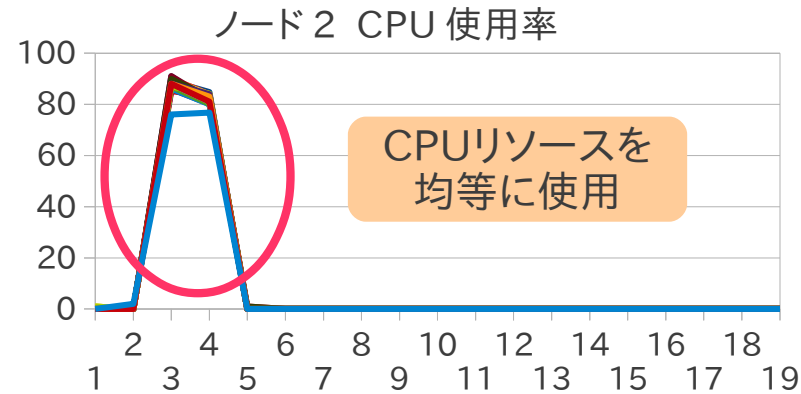
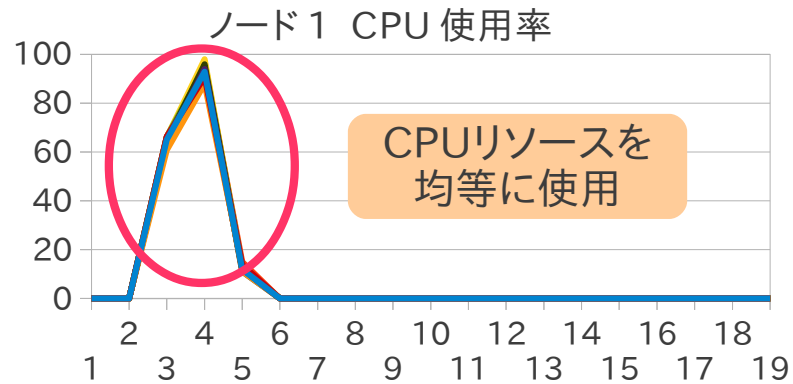
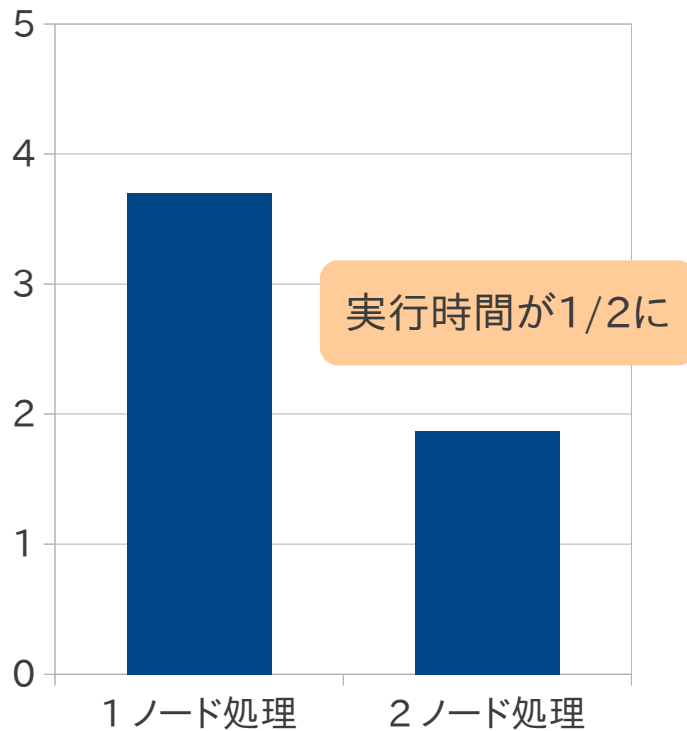
その力 無限大

- サーバの性能が限界を迎えたら、スケール可能
 - ユーザ数の増加に合わせて縦に拡張
 - データ量の増加に合わせて横に拡張




- 複数ノードによる並列処理で応答性能が向上

検索性能の比較



- InfiniDBはDWH系の処理に特化したDBMS
- チューニング不要で、最初から速い
 - 列指向 + 索引を使用しない絞り込みによるI/Oの極小化
 - パラレル処理によるCPUリソースのフル活用
- スケールアウトによるリニアな性能向上



大量の構造化データをSQLで処理したいケースに
最適なデータベース
(BIツールからのアドホック検索、レポート作成、集計など)

InfiniDB[®] の

Scalable. Fast. Simple.

ダウンロード & インストール方法

その力 無限大

- InfiniDBの開発元
 - 社名: Calpont Corporation
 - 本社: 米テキサス州フリスコ
- InfiniDBの製品体系
 - コミュニティ版 (GPL)
 - 商用版

calpont



名称		Community Edition	Enterprise Edition
ライセンス		無償 (GNU GPL)	有償
サポート		無	有
拡張方式	スケールアップ (CPU制限)	○ ※利用可能なコア数4	○
	スケールアウト	×	○
主な機能差		圧縮機能なし	圧縮機能あり
マニュアル		英語	日本語

- InfiniDBコミュニティサイト(infinidb.org)にアクセス
- ユーザ登録後、ダウンロード可能

<http://infinidb.org/>

The screenshot shows the top navigation bar of the InfiniDB Community Edition website. The navigation menu includes: Home, Community, Downloads, Forums, Blogs, Resources, Support, and About. Below the navigation bar, the slogan "SCALABLE. FAST. SIMPLE." is displayed in large, bold, orange letters. To the right of the slogan, two buttons are highlighted with red circles and boxes: "JOIN NOW" (marked with a red circle containing the number 1) and "DOWNLOAD" (marked with a red circle containing the number 2). Below the slogan, a paragraph of text describes the product: "Contributed by [Calpont](#), InfiniDB Community Edition is an open source, scale-up analytics database engine for your data warehousing, business intelligence and read-intensive application needs. Enabled via MySQL™ and purpose-built for an analytical workload with column-oriented technology at its core, the multi-threaded capabilities of InfiniDB Community Edition fully encompass query, transactional support and bulk load operations. So come on in, grab a download and [get started](#)." At the bottom left, there is a "LOGOUT" button. At the bottom right, there is a search bar with the text "Search" and a "SEARCH" button.

- Calpont社のダウンロードサイトにアクセス
- ユーザ登録なしでダウンロード可能

<http://www.calpont.com/products/tryinfinidb>

calpont
ACCELERATING DATA INSIGHTS

InfiniDB Enterprise - 64 Bit RPMs

Learn first hand about the capabilities of InfiniDB, the columnar database from Calpont, or learn how to size an analytic database to meet your business needs by downloading our software and documentation to get you up and running in hours.

Please register to download

License Agreement

CALPONT CORPORATION

BY CHECKING THE "I AGREE TO THE TERMS LISTED ABOVE" CHECKBOX, DOWNLOADING OR INSTALLING INFINIDB ENTERPRISE EDITION, YOU AGREE TO BE BOUND BY AND BECOME A PARTY TO THIS AGREEMENT. IF YOU DO NOT AGREE TO ALL OF THE TERMS OF THIS AGREEMENT, YOU MUST CLICK THE BUTTON THAT INDICATES THAT YOU DO NOT ACCEPT THE TERMS OF THIS AGREEMENT AND YOU MAY NOT DOWNLOAD, INSTALL, OR USE INFINIDB

First Name: *
Last Name: *
Company Name: *
Email Address: *
Job Title: *
Phone Number: *
State: *
Number of Employees: *
Industry: *
Department: *
I Agree To License Agreement: *

Download Now

- はじめてのInfiniDB
 - InfiniDBのインストール
 - InfiniDBの起動・停止
 - データベースの作成
 - データロード
 - データ確認

本セッションにて配布!



- InfiniDB情報発信
 - 検証記事執筆
 - Webセミナー
 - OSC出展
- 新バージョン検証
- 他製品との連携検証

InfiniDBキャラクター

インディ



InfiniDB[®]

Scalable. Fast. Simple.

database@ashisuto.co.jp

その力 無限大

*InfiniDBは米国Calpont Corporationの商標または登録商標です。
*その他記載されている会社名、製品名は各社の商標または登録商標です。