

# MySQL Cluster導入のケーススタディ 【公開用バージョン】

住商情報システム株式会社, 廣濱 顕司

2009/09/09

# 公開用バージョンについて

---

- ▶ 本ドキュメントは、2009/9/9にSUN主催で開催された「MySQL Clusterセミナー」において、「MySQL Cluster導入のケーススタディ」として発表されたセッションから、固有の数値などを除去したものとなります
- ▶ セミナーのタイトル
  - ▶ タイトル: マルチコアシステムを最大限に活かすMySQLのスケラビリティと高可用性実現セミナー～ 機能向上したMySQL Cluster 7.0最新版の詳細と高可用性システムのユーザ導入事例 & 実構築例を一挙ご紹介 ～
  - ▶ URL: <http://jp.sun.com/company/events/2009/000387.html>

# 自己紹介

- ▶ 2003年よりMySQL関連の業務を担当
  - ▶ 2003/8/4にLarry Stefonic氏とサンフランシスコでMTGしたのが全ての始まり
  - ▶ 立ち上げ時はMySQL社とのパートナーリングその他雑用全般を担当
  - ▶ 現在は以下のMySQL関連業務
    - ▶ コンサルティング
    - ▶ トレーニング講師
    - ▶ サポート
- ▶ 趣味
  - ▶ 旅行
  - ▶ ガジェット



# アジェンダ

---

- ▶ MySQL ClusterとSCS
- ▶ MySQL Cluster導入の流れ
- ▶ 楽天証券様 MySQL Cluster導入のケーススタディ
- ▶ さいごに

# MySQL Cluster と SCS

# MySQL Clusterとは

---

## ▶ 特徴

- ▶ 非共有ディスク型
  - ▶ 特殊なHWを必要としない
- ▶ アクティブ・アクティブ型
  - ▶ フェールオーバーの時間が非常に短い
- ▶ インメモリデータベース (5.1以降はディスクテーブルもサポート)
  - ▶ 高い性能
- ▶ 高い可用性
  - ▶ 各ノードの冗長構成が可能

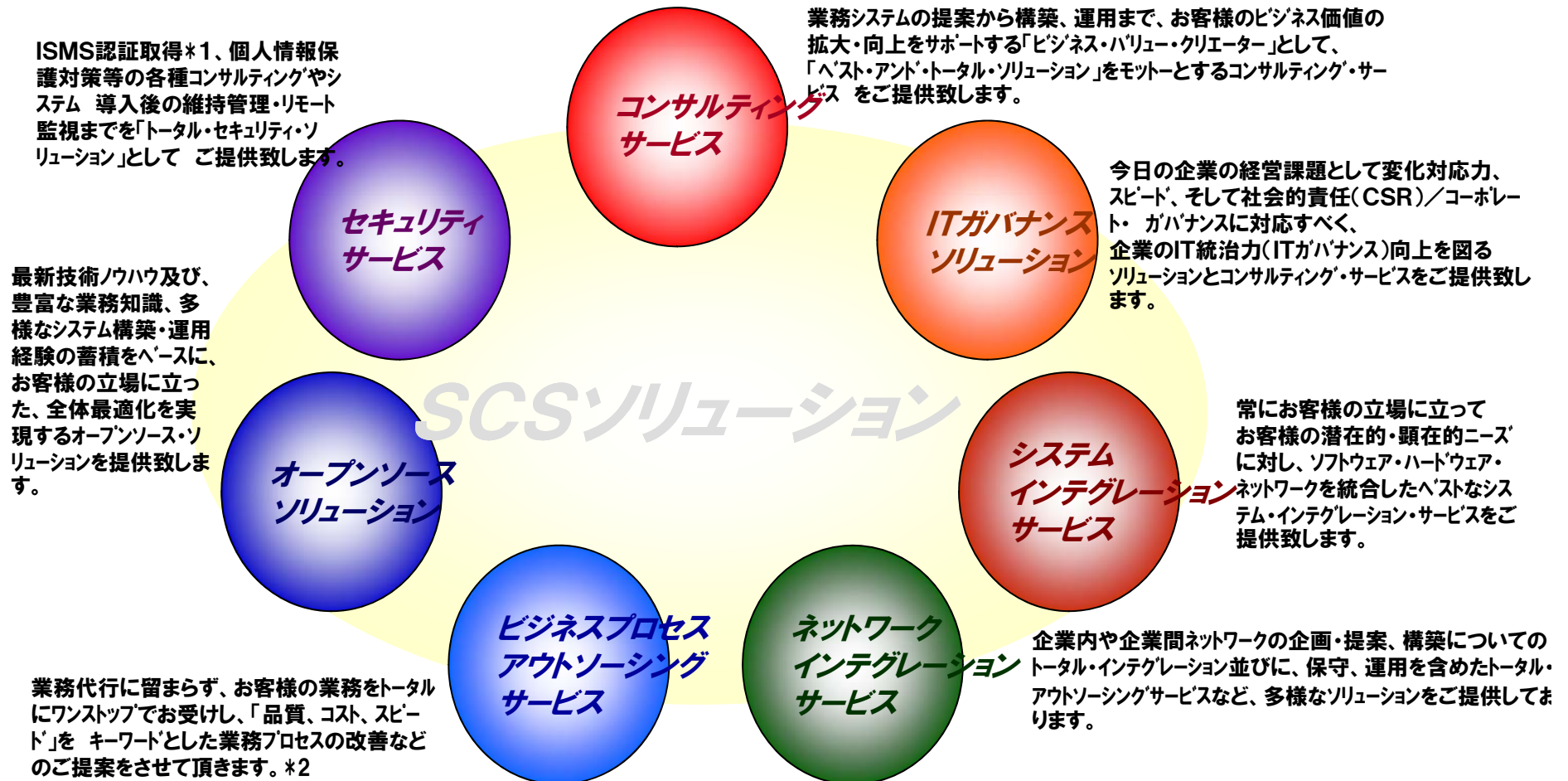
# SCSとは

---

- ▶ Sumisho Computer Systems Corporation
- ▶ 住商情報システム(旧名:住商コンピューターサービス株式会社)の略称
- ▶ 中堅システムインテグレータ
- ▶ 従業員数:3,415名(2009年3月末現在 連結ベース)
- ▶ 設立:1969年

# SCSのご紹介 ～ご提供するワンストップサービス～

ご提案から運用まで一体のワンストップなトータルソリューションをご提供致します



\*1 適正なセキュリティレベルを保持しているかどうかを認定する制度「ISMS適合性評価制度」に基づく評価認定。現在、日本情報処理開発協会(JIPDEC)を中心に2002年より正式運用されています

\*2 弊社は、1997年9月、日本初のアウトソーシング分野におけるISO9001及びTickITの認証を取得しております。

# MySQLへのSCS取り組み

---

- ▶ 国内初の「プラチナ・パートナー」
  - ▶ 前身のストラテジックアライアンスパートナーは2003年から
- ▶ オフィシャル・トレーニングの提供(2005年から)
- ▶ 過去MySQL Inc.への技術者派遣実績
- ▶ 日本語環境特有のノウハウ
  - ▶ 日本語処理機能をSCS技術者が改善支援
- ▶ IPAによるOSS性能・信頼性評価への参加
  - ▶ OSDL DBT-IのMySQL対応とMySQL性能検証
  - ▶ MySQL Clusterの可用性と性能検証
- ▶ 各種大手ベンダーへプロフェッショナルサービスなどを提供している実績
- ▶ 国内トップクラスのMySQL Cluster技術
- ▶ 日本語全文検索:MySQL Enterprise + Sennaの開発とサポート(2006年から)
  - ▶ Tritonnプロジェクトの運営

# MySQL ClusterへのSCSの取り組み

---

- ▶ 2004年
  - ▶ MySQL Clusterの検証を開始
- ▶ IPAプロジェクト (2004-2005)
  - ▶ オープンソースDBMSの評価プロジェクトにMySQL Cluster担当として参加
- ▶ Linux World 2005
  - ▶ 日立ブレードシンフォニーでMySQL Clusterをデモンストレーション
- ▶ Linux World 2006
  - ▶ 「MySQL Clusterの最適構成」セッション担当
- ▶ 2007年～
  - ▶ 実案件における問い合わせ/サービス提供

# IPAプロジェクトの概要

---

- ▶ 正式名称
  - ▶ 2005年度上期オープンソースソフトウェア活用基盤整備事業「OSS性能・信頼性評価 / 障害解析ツール開発」DB層
- ▶ SCSはMySQL Cluster担当として参加
- ▶ 評価レポートなどはIPAのサイトで公開中
  - ▶ <http://www.ipa.go.jp/software/open/forum/development/index.html>

# IPAプロジェクトの結果

---

- ▶ MySQL ClusterのHA機能で不可解な挙動は無い
  - ▶ ノード障害、ネットワーク障害、サーバー障害などにも対応
  - ▶ フェールオーバーは非常に高速：数秒
- ▶ インターコネクต์にSCI (Dolphin Interconnect Solutions社)を利用することで性能は最大50%向上した
- ▶ MySQL Cluster 4.1 / 5.0には高負荷時の安定性で若干の課題あり
  - ▶ 安定性は6.2で解決
  - ▶ 性能はCGEおよび6.2で改善

# DB Magazine 2008年8月号「徹底検証： MySQL Cluster」ベンチマーク結果のまとめ

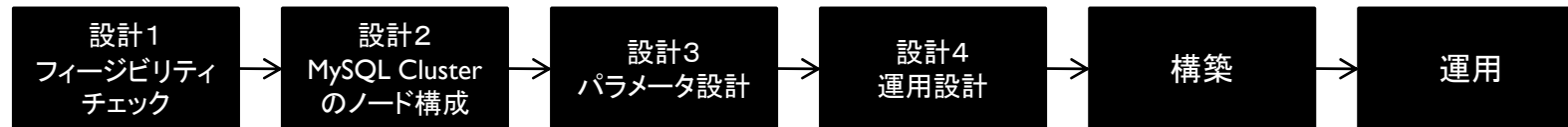
---

- ▶ MySQL Cluster 6.2は安定している
  - ▶ MySQL Cluster 4.1, 5.0, 5.1を利用している場合はアップグレードを強く推奨
- ▶ ディスクテーブルの全面的な採用は未だ早い
  - ▶ 性能面および安定面で懸念あり
- ▶ Data Node数は少なく、SQL Node数は多く、レプリカ数は少なく、同時接続数は50くらい、が初期構成の目安
  - ▶ MySQL Cluster 5.0と傾向は変わらない
- ▶ Data Nodeはシングルスレッドでの動作となるので、CPUなどを増やしても性能はスケールしない
  - ▶ 7.0で対応済

# MySQL Cluster導入の流れ

# MySQL Cluster導入の流れ

---



- ▶ MySQL Cluster導入にあたっては、導入前のフェージビリティのチェックとノード構成が最も重要
- ▶ ベストプラクティスを研究し、アンチパターンに陥らない設計が重要

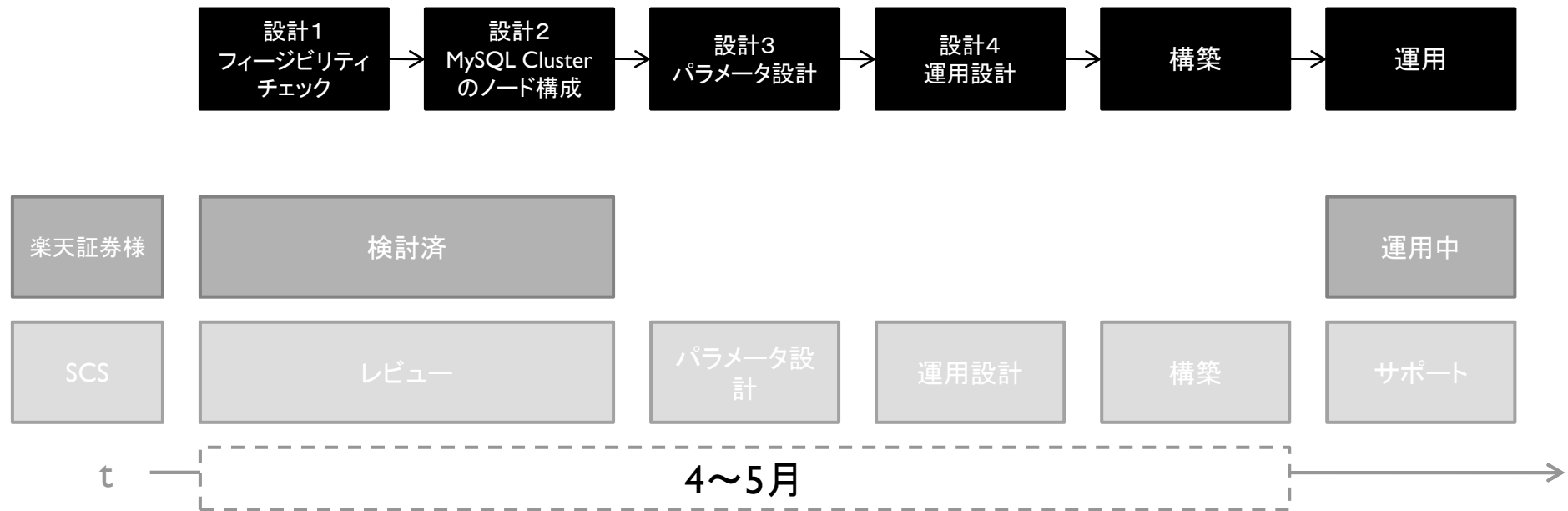
# 楽天証券様MySQL Cluster導入 のケーススタディ

# 楽天証券様のMySQL Cluster導入のタイムフレームとSCS

	楽天証券様の当初想定スケジュール	SCS
4月	(購入前)初めてのMTG	
5月	中旬～下旬:本番環境構築	下旬:MySQL Cluster構築
6月	テスト 末日:システム・リリース	ドキュメント作成および検収

- ▶ ファーストコンタクトから2ヶ月で構築完了
- ▶ ドキュメント作成および検収に1ヶ月
- ▶ 今回は、既にMySQL Cluster適用の可否を判断頂いていたので比較的短時間で構築できた

# MySQL Cluster導入の流れ



- ▶ SCSはフィージビリティチェックを含む設計・構築から運用支援およびサポートを提供

# 設計 1：フェージビリティの確認

---

- ▶ 代表的なヒアリング項目
  - ▶ MySQL Clusterを適用できるか
    - ▶ データ容量
      - メモリテーブルの利用を推奨
    - ▶ クエリ
      - サブクエリ、結合など
    - ▶ 価格
      - ライセンスおよびサポート
  - ▶ 求められる可用性のレベル
  - ▶ アクティブ・アクティブを生かせるか
  - ▶ アンチ・パターン診断
    - ▶ HDDはデータ容量 \* 7が目安(7.0)
    - ▶ 一度に大量にデータを削除するのではなく、分割して削除

## 設計 2 : MySQL Clusterのノード構成

---

- ▶ 過去の実績・経験から最適なノード構成を提案
  - ▶ IPAプロジェクトでのベンチマーク、社内検証環境でのベンチマークなどからベストプラクティスを把握
  - ▶ 最適なノード数のバランス
  - ▶ ノード間通信
  - ▶ SQL Nodeの配置(アプリケーションとの同居)
- ▶ 代表的なヒアリング項目
  - ▶ MySQL Clusterに要求されるスループット (TPS)
  - ▶ クエリの概要
  - ▶ Management Nodeは負荷も軽いので、別サーバーで他のサーバと同居、冗長構成の検討
  - ▶ Geographical Replicationの要・不要

# 設計 3 : パラメータ設計 (1)

---

- ▶ テーブルサイジング
  - ▶ DDLおよび想定データ件数から算出
    - ▶ Data Memory: x GB
    - ▶ Index Memory: y GB
- ▶ 設定ファイル作成
  - ▶ テーブル数z程度
  - ▶ ディスクテーブルの利用有無
  - ▶ 割り当て可能な物理メモリ
    - ▶ Data Memory: xx GB
    - ▶ Index Memory: yy GB
  - ▶ LCP
  - ▶ Redo Log
  - ▶ バックアップ世代
- ▶ テーブル数が多い場合、トランザクションが多い場合、Geographical Replication構成を採用する際などは、別途チューニングが必要

## 設計 3 : パラメータ設計 (2)

---

- ▶ クエリ実行計画
  - ▶ EXPLAINなどでクエリの確認
  - ▶ PKアクセスが望ましい
  - ▶ 結合、サブクエリの有無
  - ▶ インデックス過不足の確認
  - ▶ condition pushdownの確認
- ▶ テーブル定義などのレビュー
  - ▶ 文字コードはlatin1が推奨
  - ▶ varcharは現在のバージョンでは可変長
  - ▶ 一行の最大長は8K (text / blobはのぞく)
- ▶ PKアクセス以外は注意
- ▶ Data Nodeを増やす場合はDistribution Awarenessを利用

```
ROWS: 10  
Extra: Using where with pushed condition  
w in set (0.00 sec)
```

# 設計 4 : 運用設計

---

- ▶ 運用設計と試験
  - ▶ 起動・停止手順
  - ▶ バックアップ・リストア
  - ▶ ログファイルなどのメンテナンス
  - ▶ 監視項目
  - ▶ オンラインでの設定変更手順と変更可能な項目の整理

# 構築

---

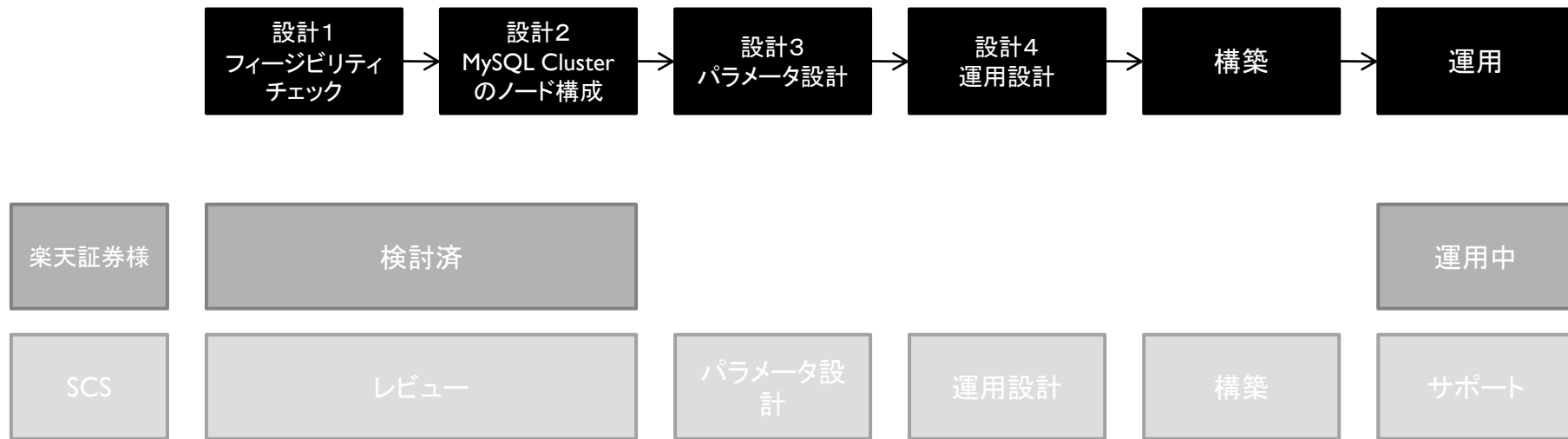
- ▶ インストールにあたって
  - ▶ パッケージまたはtar.gzの選択
  - ▶ ndbdまたはndbmtdの選択
- ▶ 単体試験の実施とドキュメント化
  - ▶ MySQL Cluster単体試験
  - ▶ NDBテーブルの挙動確認
  - ▶ MySQL ClusterのHA試験

# 運用

---

- ▶ 楽天証券様で実施
- ▶ SCSサポートにて不具合発生時は対応

# 楽天証券様のMySQL Cluster導入の流れ



- ▶ 購入頂いたSCSサービス
  - ▶ 設計・構築
  - ▶ サポート
  - ▶ ドキュメント完備
- ▶ サポート開始前のQAも鋭意対応
- ▶ Geographical Replicationは利用していないので比較的シンプルな構成
- ▶ 楽天証券様がMySQL Clusterの特徴を正しく理解されていたため、大きな問題には遭遇しなかった

さいごに

## SCSの「MySQLサービス」を利用して頂く利点

---

- ▶ プリセールスの段階からテクニカルなアドバイスが可能
  - ▶ 通常は概要設計に相当
- ▶ SCSの「MySQLサービス」を利用することでTCOの削減
  - ▶ SCSは設計・構築時のノウハウが豊富
  - ▶ SCSのアドバイスでアンチ・パターンに陥らない
  - ▶ SCSはベストプラクティスを熟知している
- ▶ システムインテグレータとしてニュートラルな立場でMySQL Clusterを提案
  - ▶ システムによってはLifeKeeper, VCS, DRBDなどを提案
  - ▶ 無闇に新機能、新製品は提案しません

# SCSではそのほかにも

---

- ▶ MySQLトレーニング
  - ▶ MySQL 5.1 for DBA
  - ▶ MySQL High Availability
  - ▶ MySQL Performance Tuning
- ▶ MySQLサポート
  - ▶ 2005/4より開始
  - ▶ 60社以上へのSCSサポートプラス導入実績
- ▶ MySQLプロフェッショナルサービス(コンサルテーション)
- ▶ システム・インテグレーション
  - ▶ MySQLを利用したシステム構築
  - ▶ MySQLとOracleを利用するシステム構築(同じチームにOracleエンジニアが在籍)

# 参考資料

---

mysql scs

検索



- ▶ 今回の資料
- ▶ 過去セミナーでの資料
  - ▶ MySQL Cluster最適構成
  - ▶ MySQL日本語全文検索
  - ▶ MySQL日本語処理完全解説
  - ▶ DBT-1ベンチマーク結果
- ▶ トレーニング情報

## さいごに

---

- ▶ MySQL Clusterは「使え」ます
  - ▶ Oracle RACと比較されるものではありません
  - ▶ 独特の「クセ」があります
  - ▶ うまく使うと高可用・高速なDBをリーズナブルな価格で利用できます
  - ▶ お困りの点がありましたら、SCSへお気軽にご相談下さい
- ▶ ご清聴ありがとうございました