

オープンソースデータベース GridDB と そのオープンコミュニティ活動



TOSHIBA

東芝デジタルソリューションズ株式会社 GridDBコミュニティ版担当

野々村 克彦

2019.11.23

プロフィール

名前：野々村 克彦

所属：東芝デジタルソリューションズ（株）

ソフトウェア&AIテクノロジーセンター データ基盤技術開発部



Katsuhiko Nonomura

knonomura

Block or report user

Organizations



2019 #2

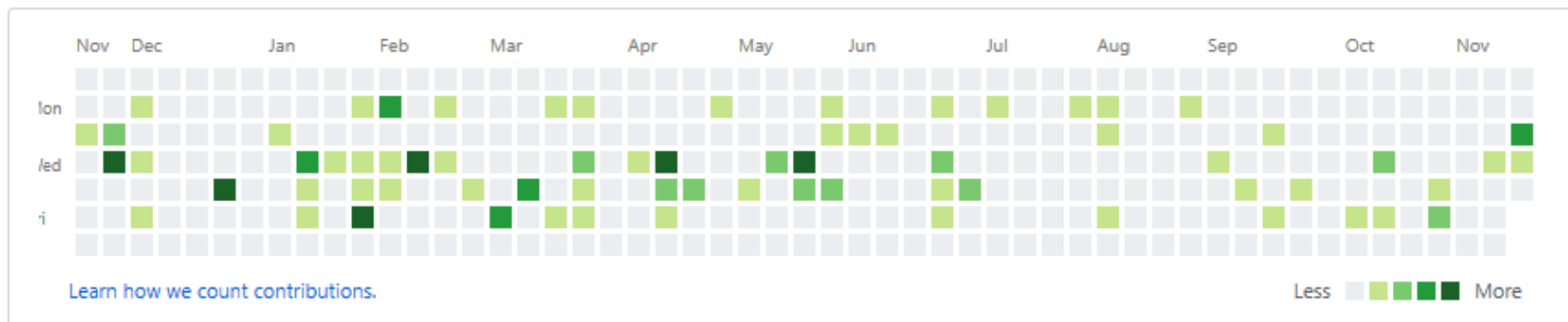
2000年ごろ XMLデータベース開発メンバ

2012年 スケールアウト型DB GridDBの開発メンバ

2015年 GridDBのオープンソースPJ開始

現在 GridDBコミュニティ版の開発・コミッター。GitHub歴 4年

408 contributions in the last year



Contents

01 スケールアウト型データベースGridDBの概要

02 事例紹介

03 OSS活動

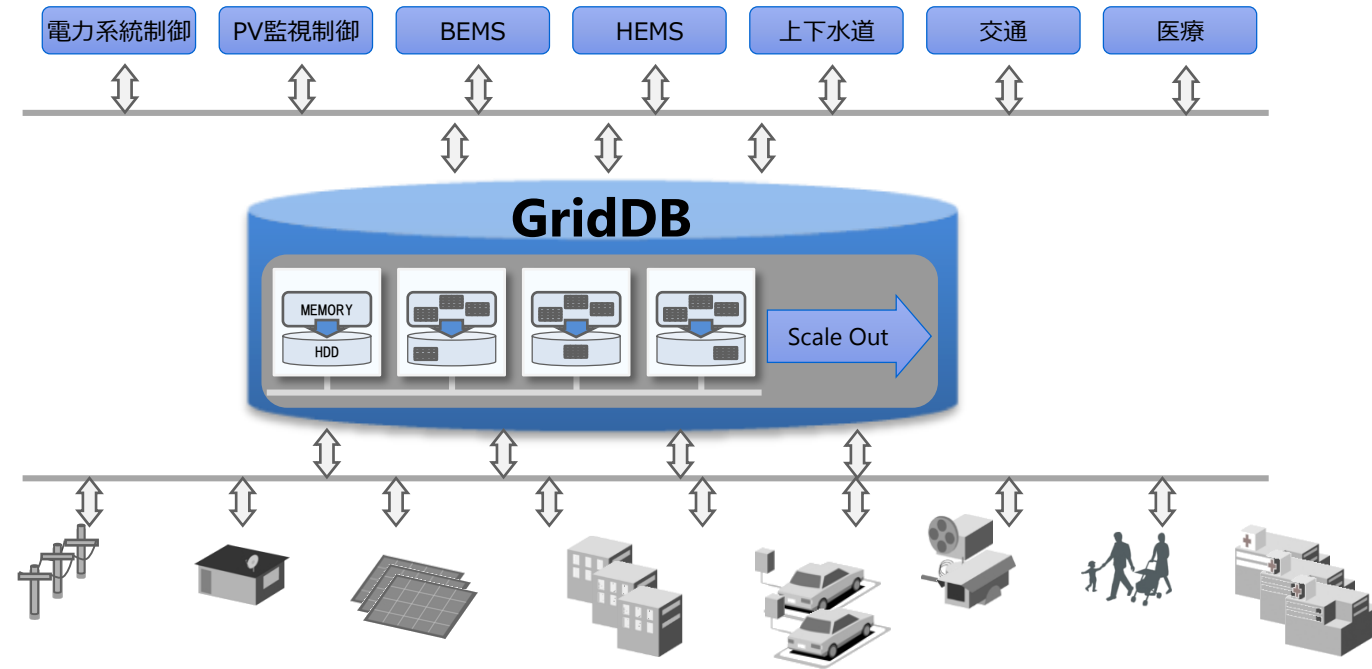
04 まとめ

01

スケールアウト型データベースGridDBの概要

スケールアウト型データベースGridDB

- 日本発のビッグデータ/IoT向けスケールアウト型データベース
- V1.0製品化(2013年)、OSS化(2016年)、V4.2(2019年6月)
- 社会インフラを中心に、高い信頼性・可用性が求められるシステムに適用中



GridDBの特長

IoT指向の データモデル

- データモデルはキー・コンテナ。コンテナ内でのデータ一貫性を保証
- 時系列データ管理する特別な機能
- 過去データをコールド保存する長期アーカイブ機能

高い信頼性と 可用性

- データの複製をノード間で自動的に実行
- ノード障害があってもフェールオーバーによりサービス継続
- 数秒から数十秒の切替え時間

高いスケーラビリティ

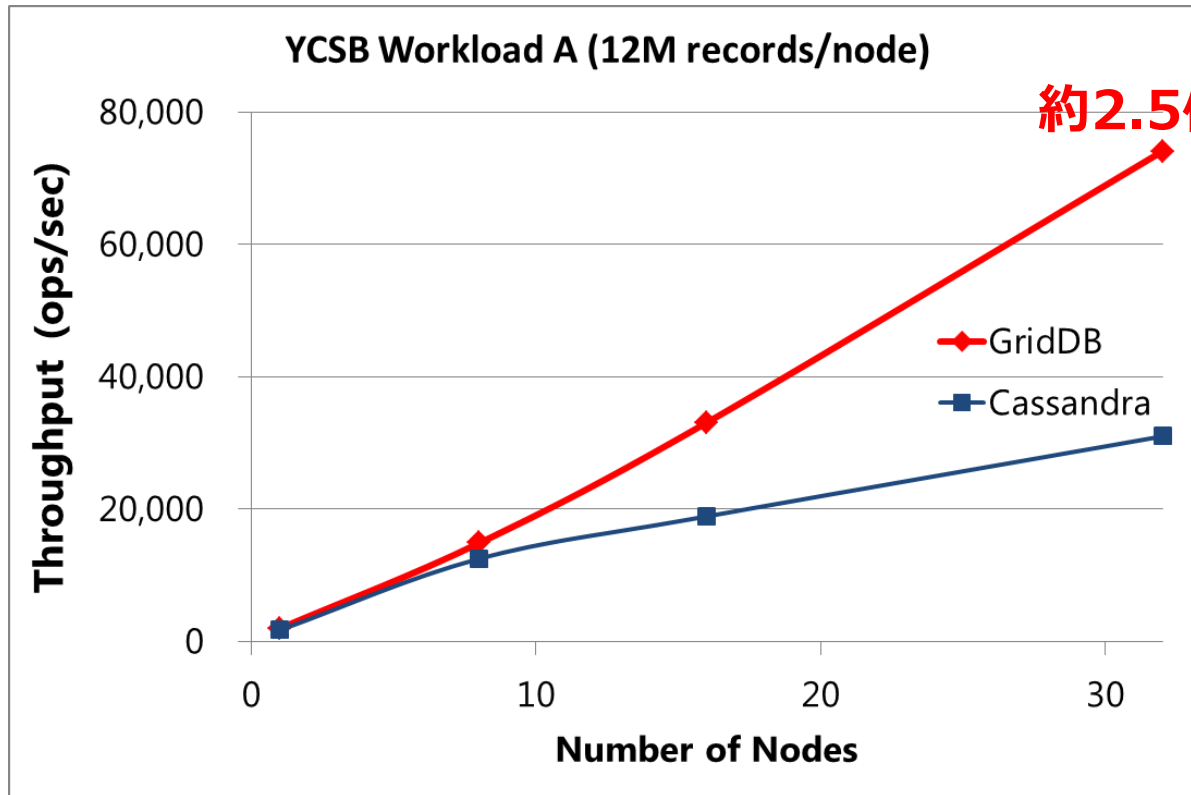
- 少ないノード台数で初期投資を抑制
- 負荷や容量の増大に合わせたノード増設が可能
- 自律データ再配置により、高いスケーラビリティを実現

高性能な NoSQL+SQL

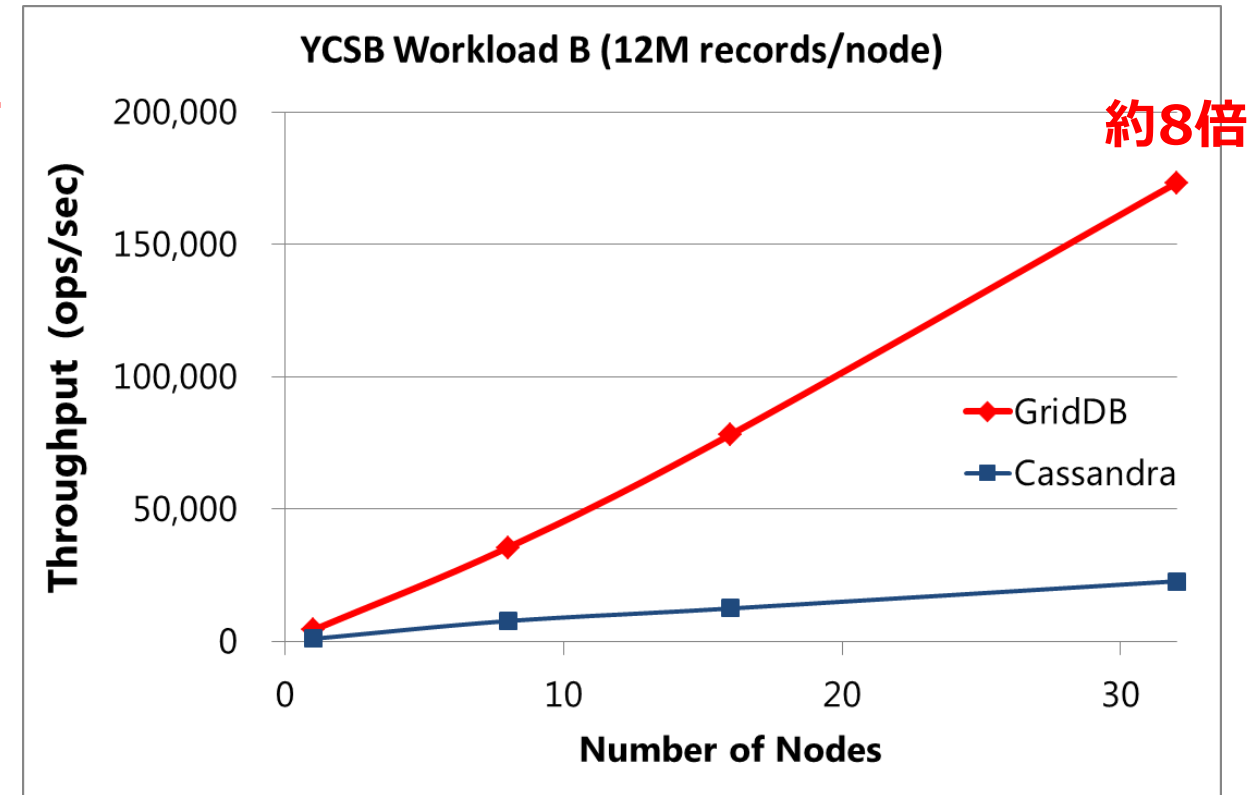
- メモリを主、ストレージを従としたハイブリッド型インメモリDB
- メモリやディスクの排他処理や同期待ちを極力排除
- SQLにおける分散並列処理

YCSB (Yahoo! Cloud Serving Benchmark)

NoSQLの代表的なベンチマーク <https://github.com/brianfrankcooper/YCSB>



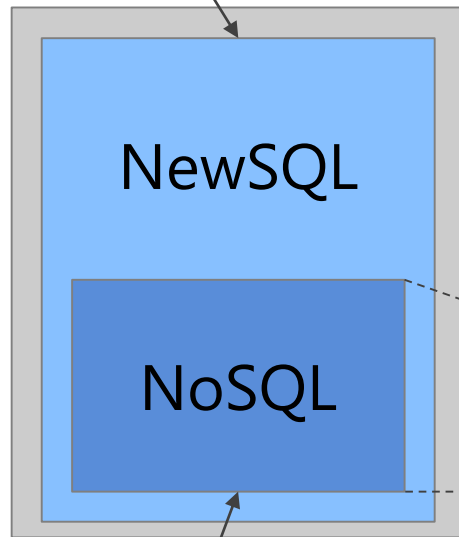
Read 50% + Write 50%



Read 95% + Write 5%

※フィックスターズ社によるYCSBベンチマーク結果

② GridDB AE (Advanced Edition)



① GridDB SE (Standard Edition)

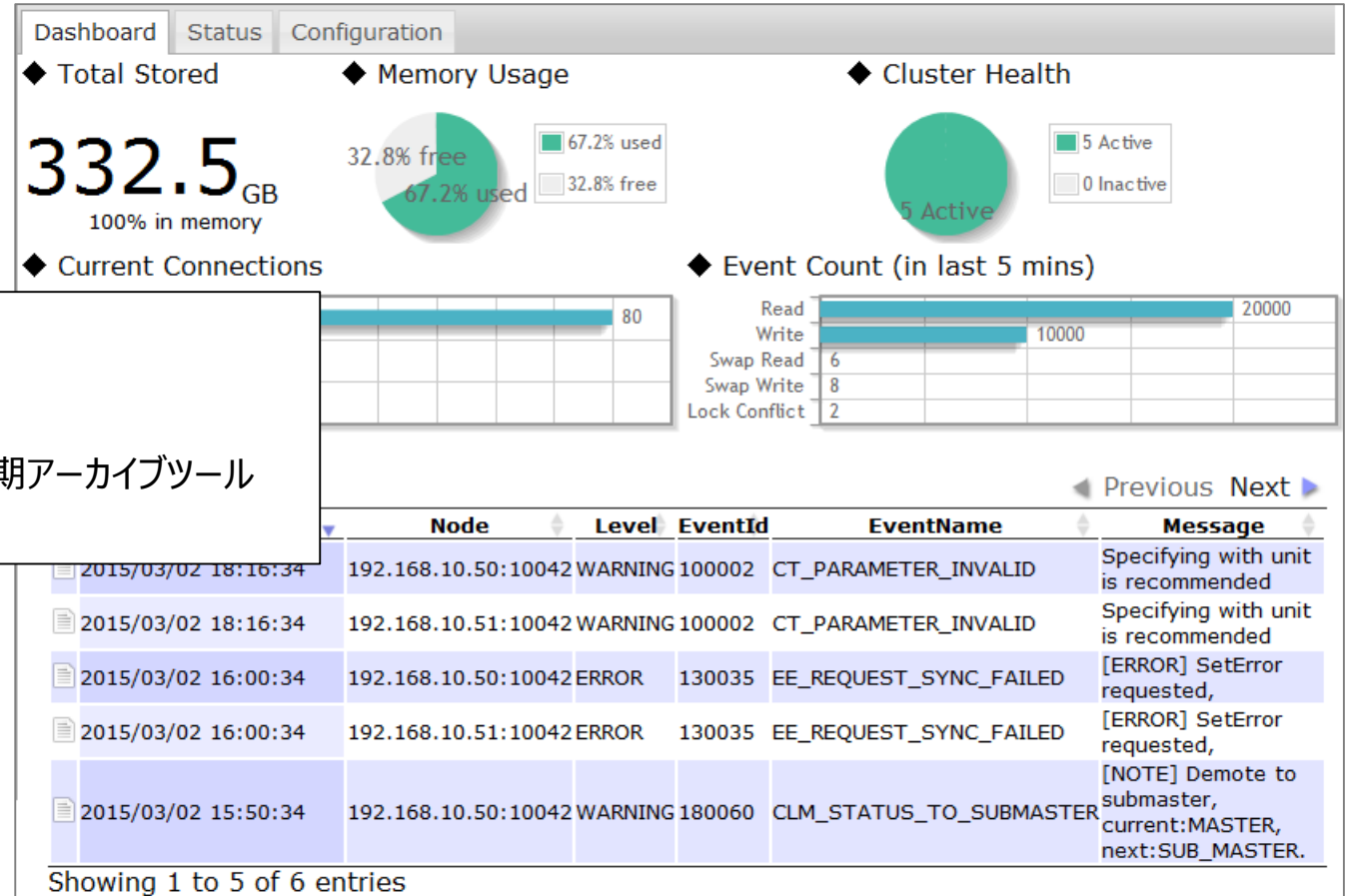
運用ツール

- ・運用管理GUI
- ・コマンドインタプリタ(gs_sh)
- ・バックアップ、Exp/Imp、長期アーカイブツールなど



③ GridDB CE (Community Edition)

Monitoring Dashboard



02

事例紹介

主な適用事例

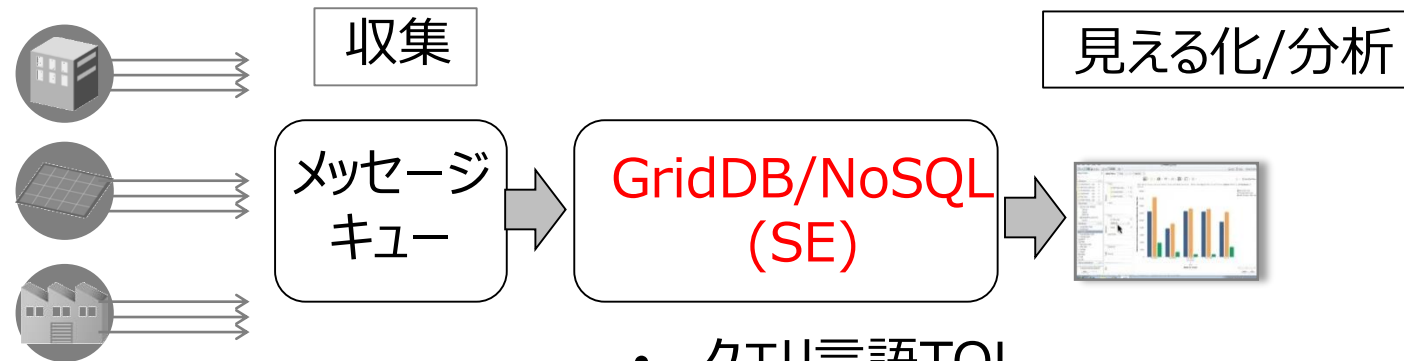
- 社会インフラを中心に、高い信頼性・可用性が求められるシステムに適用中

- フランス リヨン 太陽光発電 監視・診断システム
発電量の遠隔監視、発電パネルの性能劣化を診断
- クラウドBEMS
ビルに設置された各種メータの情報の収集、蓄積、分析
- 石巻スマートコミュニティ プロジェクト
地域全体のエネルギーのメータ情報の収集、蓄積、分析
- 電力会社 低圧託送業務システム
スマートメータから収集される電力使用量を集計し、需要量と発電量のバランスを調整
- 神戸製鋼所 産業用コンプレッサ稼働監視システム
グローバルに販売した産業用コンプレッサをクラウドを利用して稼働監視
- 東芝機械 IoTプラットフォーム
工作機器、射出成形、ダイカストマシン、など膨大な製造データを管理
- デンソー ファクトリー IoT
工場の生産性向上、世界130工場に展開予定
- DENSO International Americaの次世代の車両管理システム
車両の各種センサーデータを用いる車両管理システムのPoC
-

主な適用パターン

1. 時系列データの管理（見える化など）
2. Hadoop、Sparkによる分散処理
3. OSS連携によるPoC(Proof of Concept)

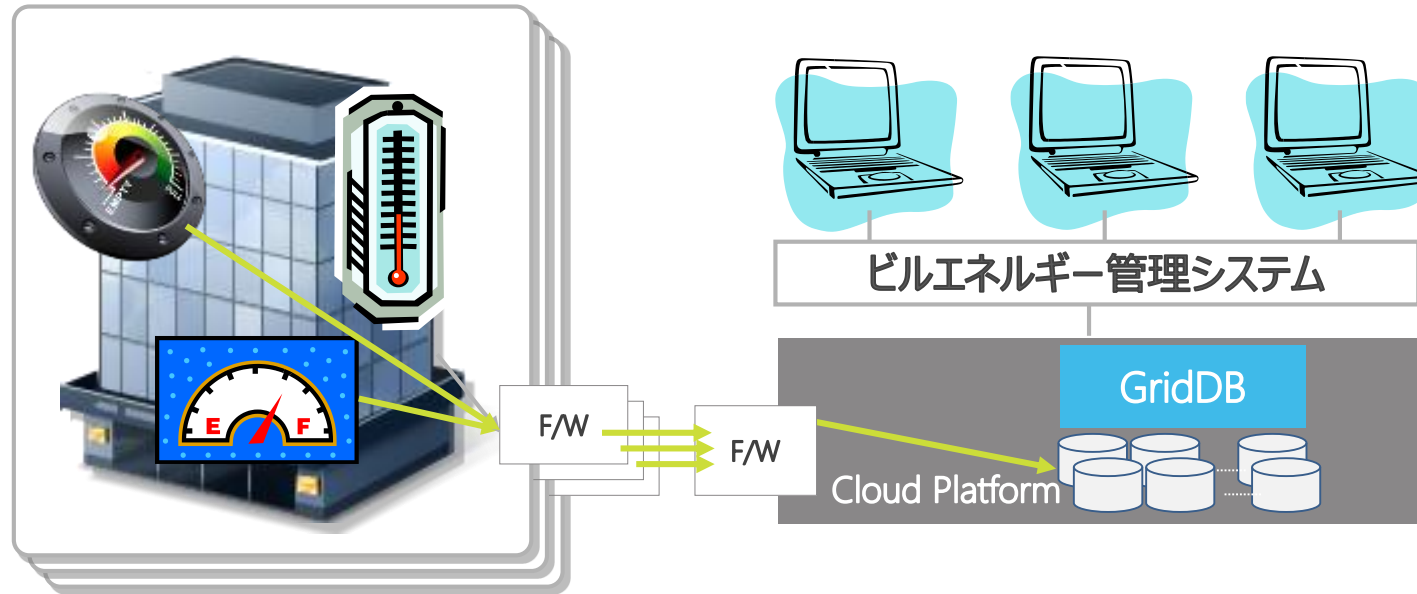
1. 時系列データの管理



- クエリ言語TQL
- 集計、サンプリングなど時系列専用関数
- バッチ処理(MultiPut/Get/Query)
- 期限解放
- アフィニティ
- ……
- 自律データ再配置技術 (ADDA)
- ……

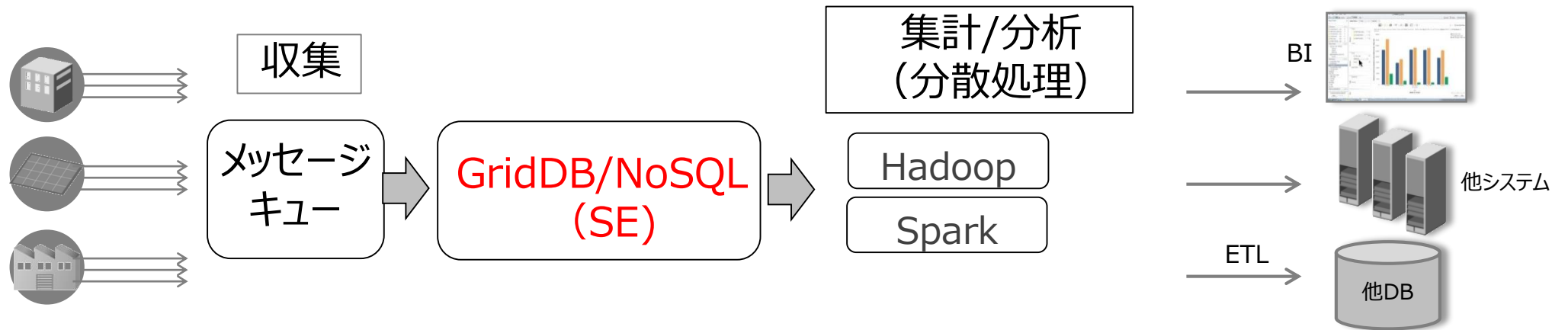
- 高速、かつ高い信頼性と可用性のあるシステムを実現

事例：ビルエネルギー管理システム（BEMS）



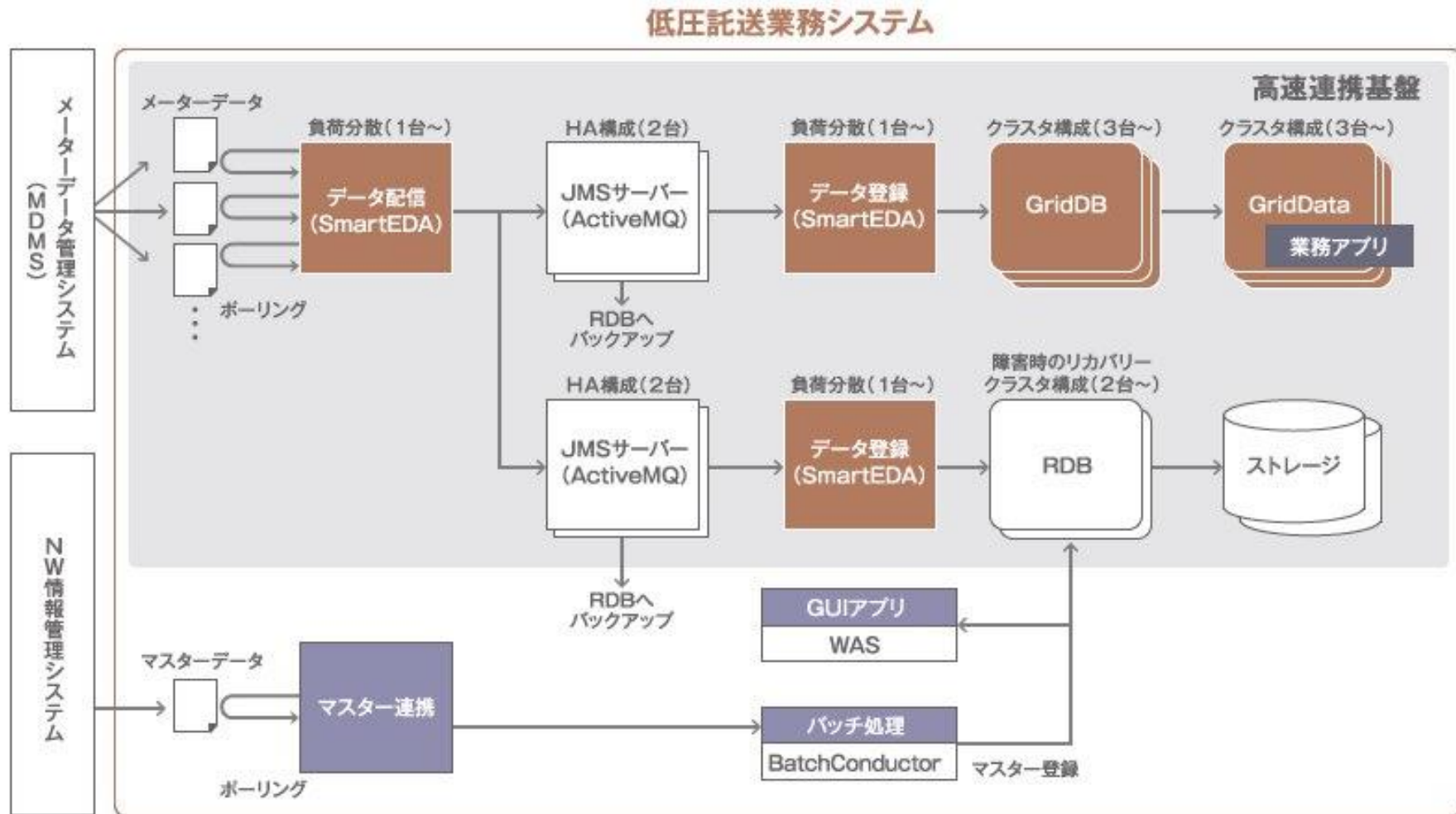
- 2015年からBEMSサービスを提供。
- 数百の建物から収集された2TBを超えるデータを格納。
- 各建物には約50個のセンサー。センサーデータは1分間隔で収集。
- 毎秒1,000件以上のデータの読み込みと書き込みが必要。

2. Hadoop、Sparkによる分散処理



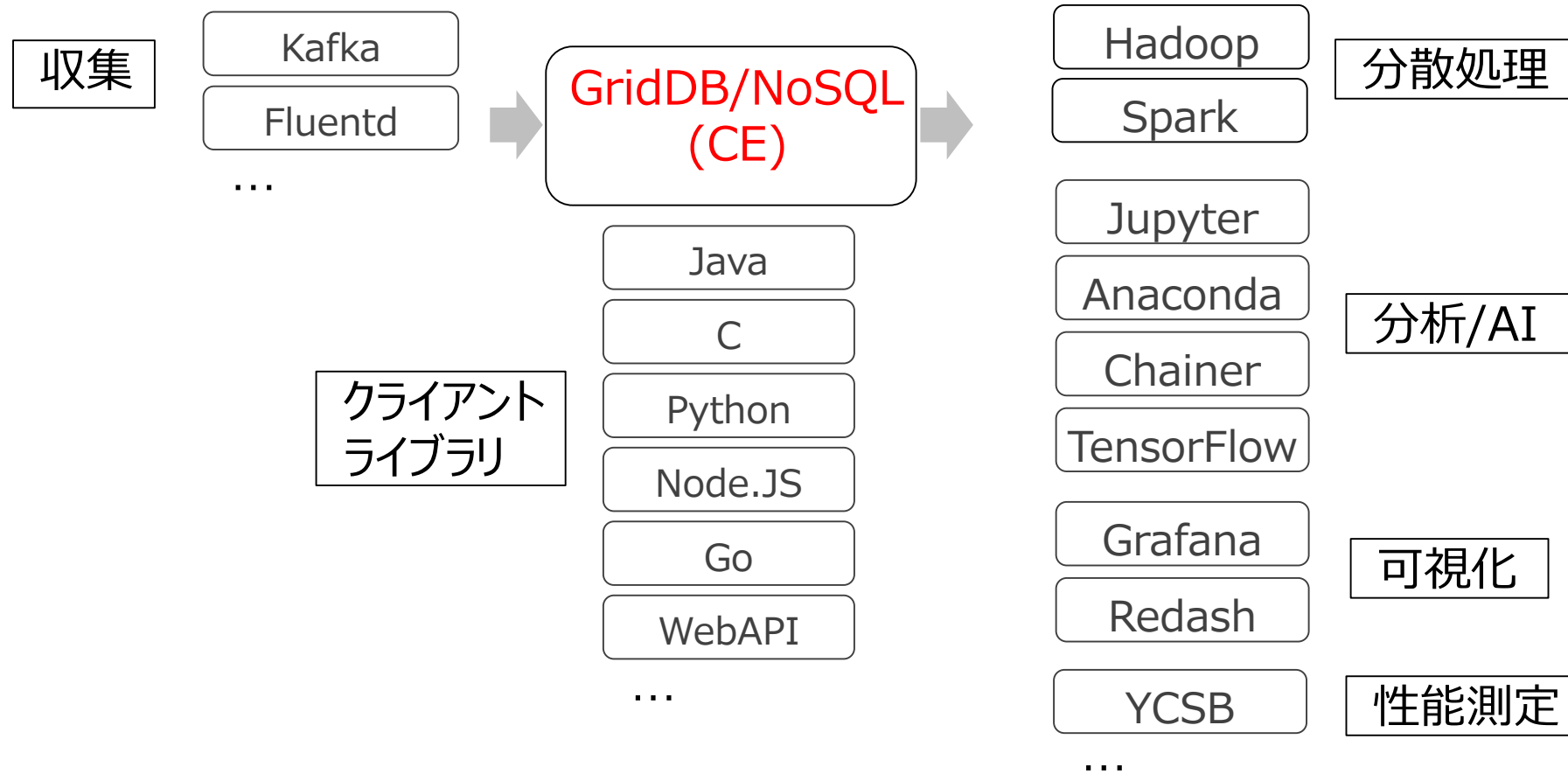
- Hadoop、Sparkを使って集計・分析を大規模分散処理
⇒ GridDB/NoSQLの特長を最大限に活かす

事例：電力会社（システム構成）



- 数百万台のスマートメータから30分おきに送られてくるメータデータ3ヶ月分を蓄積（数百億レコード、数TB）
- 2016年4月の運用開始以来、安定稼働

3. OSS連携によるPoC

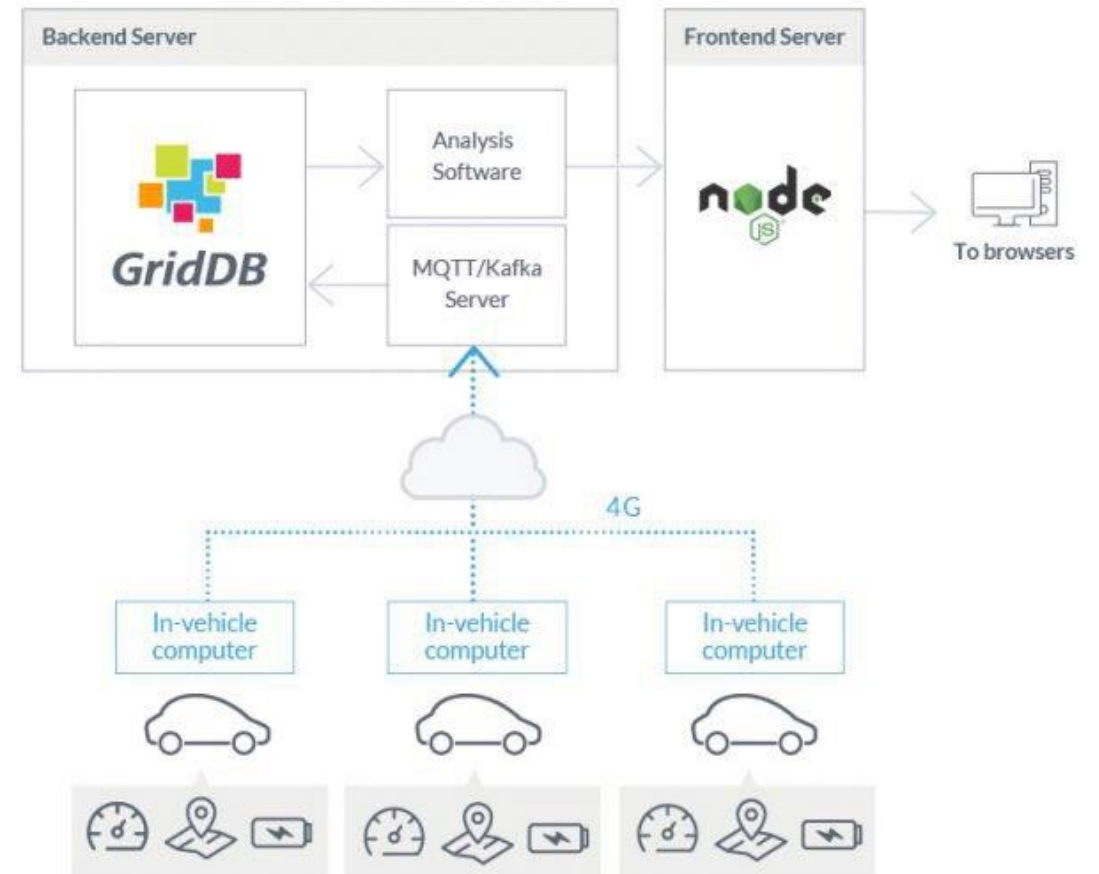


- 様々な開発言語のクライアントライブラリ
- 主要OSSのコネクタ

事例：自動車産業

- 次世代の車両管理システムの構築のPoC。
- エッジコンピューター側やデータ表示のWeb フロントエンド側も含めて、2, 3ヶ月という極めて短時間で実現。

<https://griddb.net/ja/blog/griddb-automotive/>



03

GridDBのOSS活動

- **GridDBのNoSQL機能をGitHub上にソース公開(2016/2)**



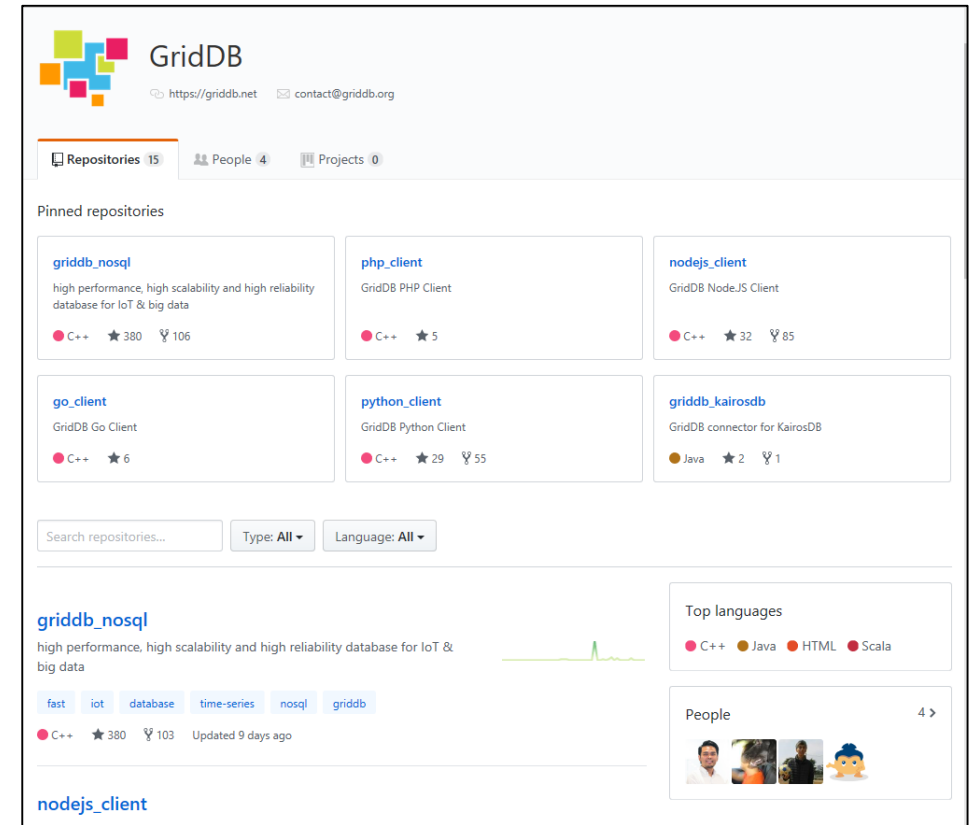
https://github.com/griddb/griddb_nosql

- **目的**

- ビッグデータ技術の普及促進
 - 多くの人に知ってもらいたい、使ってもらいたい。
 - いろんなニーズをつかみたい。
- 他のオープンソースソフトウェア、システムとの連携強化

- **ライセンス**

- サーバはAGPL-3.0
- 各種開発言語のクライアント、OSSとのコネクタはApache-2.0



最近の活動（前回のOSC報告）

2018年

- 10月 **GridDB Perlクライアントのソース公開(GitHub)**

- 11月 **YCSBリポジトリ本家へのPullRequest**

※YCSBのメインコミッターからお誘いがあった

2019年

- 1月 **GridDB V4.1 Community Editionのソース公開(GitHub)**

- 2月 **Javaクライアントのパッケージ公開(Maven Central Repository)**

GridDB V4.1 CEのUbuntu向けパッケージ公開(GitHub)

Node.JSクライアントのインタフェース改良 v0.8

最近の主な活動（前回のOSCC以降）

2019年

- 3月 WebAPIのソース公開(GitHub)
- 4月 Grafanaプラグインのソース公開(GitHub)
- 5月 GridDB V4.2 CEのソース公開(GitHub)
- 6月 **YCSBリポジトリ本家へのPullRequest採用・マージ**
CクライアントのCentOS用RPMパッケージ公開(OpenSUSE Build Service)
- 7月 Grafanaリポジトリ本家へのPullRequest (活動中)
- 8月 CクライアントのUbuntu用DEBパッケージ公開(OpenSUSE Build Service)
Pythonクライアント(Ubuntu/Windows版)のソース公開(GitHub)
- 9月 Node.JSクライアント(Ubuntu/Windows版)のソース公開(GitHub)
- 10月 Redashプラグイン、Fluentdプラグインのソース公開(GitHub)
Goクライアント(Ubuntu/Windows版)のソース公開(GitHub)
- 11月 CクライアントのWindows用MSIファイル公開(GitHub)
GridDB V4.3 CEのソース公開(GitHub)

主なOSS活動

① GridDB本体の機能強化

② 主要OSSとの連携強化

③ APIの拡充

④ GitHub以外のサイトからの情報発信

- パッケージ
- デベロッパーズサイト（WP、ブログなど）・・・フィックスターズ社
- SNS・・・フィックスターズ社

⑤ 主要OSSリポジトリへのコントリビュート

- YCSB、Grafanaなど

⑥ プラットフォームの拡充

⑦ その他

- OSCなどカンファレンス参加
- ハンズオン無料セミナー・・・（株）アイ・ティ・イノベーション

OSS活動の全体イメージ

性能測定

収集

分散処理

可視化

Webアプリ

分析

AI/機械学習 ...

④ GitHub以外のサイトからの情報発信

PyPI/npm/Maven/...

② 主要OSSとの連携強化

Spark
コネクタ

Fluentd/Grafana/Redash
プラグイン

YCSB
コネクタ

Kafka
コネクタ

Hadoop
MapReduce
コネクタ

WebAPI

Python/Node.JS/Go/PHP/Ruby/Perlクライアント

③ APIの拡充

Javaクライアント

Cクライアント

① GridDB本体の機能強化

GridDB V4.3 CE(Community Edition)

⑤ 主要OSSリポジトリへのコントリビュート (YCSB、Grafanaなど)

GitHub



(スケールアウトだけでなく) スケールアップの強化 :

1. ブロックサイズ拡大

- データベース初期作成時に選択できるブロックサイズを最大32MBまで拡張可能に。

2. チェックポイントファイルの分割配置

- チェックポイントファイルを分割し、複数ディレクトリへ分散配置可能に。

3. 外部と内部の通信を分離するネットワーク構成

- クライアント-ノード間の外部通信とノード間の内部通信に別々のネットワークインタフェース割り当てを可能に。

その他 :

4. 複合索引 [性能向上]

- 複数カラムを指定した索引を作成可能に。

5. 複合ロウキー [開発機能]

- 複数カラムを指定してロウキーを設定可能に。

6. タイムゾーン指定 [開発機能]

④ GitHub以外のサイトからの情報発信 パッケージ

パッケージ公開サイト	概要	公開済のGridDB関連ソフト
PyPI (https://pypi.org/)	Python言語用	Pythonクライアント (CentOS)
npm (https://www.npmjs.com/)	Node.js言語用	Node.JSクライアント (CentOS)
Maven Central Repository (https://search.maven.org/)	Java言語用。Mavenがデフォルトで利用するパブリックリポジトリ	Javaクライアント
OpenSUSE Build Service (https://build.opensuse.org/)	OpenSUSEの他、Ubuntu、CentOSなどのディストリビューションを公開するサイト。yumコマンドなどが使える	Cクライアント (CentOS/Ubuntu)
DockerHub (https://hub.docker.com/)	Dockerイメージ用	サーバ、Java/Pythonクライアント (CentOS)
RubyGems (https://rubygems.org/)	Ruby言語用	
GitHub Packages, now out of beta (https://github.blog/2019-11-13-universe-day-one/#github-packages)	GitHub上でソースと一緒にパッケージを共有できるサービス。先日正式版に。JavaScript (npm)、Java (Maven)、Ruby (RubyGems)、.NET (NuGet)、Dockerイメージなどに対応	

⑤ 主要OSSリポジトリへのコントリビュート(1/5) 例：YCSBコネクタ

• YCSBとは

- Yahoo! Cloud Serving Benchmark

<https://research.yahoo.com/news/yahoo-cloud-serving-benchmark/>

- NoSQLの代表的なベンチマーク。但し、必ずしもIoT向けではない
- Load/Runの2フェーズ、Runは6種のworkloadから成る

work load	type	insert	read	update	scan
A	Update heavy		50%	50%	
B	Read mostly		95%	5%	
C	Read only		100%		
D	Read latest	5%	95%		
E	Short ranges	5%			95%
F	Read-modify-write		50%	50% ※read-modify	

対応しているDB一覧：

HBase
Hypertable
Cassandra
Couchbase
Voldemort
MongoDB
OrientDB
Infinispan
Redis
GemFire
DynamoDB
Tarantool
Memcached
など約40種

- 最新版0.17(2019/9)にてCrail、CosmosDB、Postgres NoSQL、Table Store、**GridDB**、VoltDBが追加された

⑤ 主要OSSリポジトリへのコントリビュート(2/5) 経緯その1

- 2016/10 GridDB用のYCSBコネクタをソース公開(GitHub)

...

- 2018/9 YCSBリポジトリのメインコミッターからPullRequestのお誘いあり

https://github.com/griddb/griddb_ycsb/issues/1

⇒快諾。しかし、いくつかの課題があった

※緑色：YCSBリポジトリのメインコミッターのアクション

⑤ 主要OSSリポジトリへのコントリビュート(3/5) 課題と対策

(課題)

- YCSBの（ソースからでなく）バイナリからビルドするビルド環境。Mavenに対応していない
- 依存関係にあるJavaクライアント(jar)をローカルに置く必要がある

(対策)

- 3ステップで進めることにした。

Step1. ソースからビルドするMavenビルド環境への対応

Step2. Javaクライアントのパッケージ(Jar)をMaven Central Repositoryから公開

Step3. Maven Centoral Repository上のJavaクライアントを使うビルド環境への対応

⑤ 主要OSSリポジトリへのコントリビュート(4/5) 経緯その2

- 2018/11 ソースからビルドするMavenビルド環境への対応(step1)

YCSBリポジトリにPullRequest

<https://github.com/brianfrankcooper/YCSB/pull/1258>

しかし、自動ビルドができなため、CI Travisでエラー状態

※メインコミッタがHBaseなどApacheプロジェクトのコミッタでもあり、多忙なため応答なし

- 2019/1 次のステップを活動中と報告したところ、「いいね」だけ反応あり



- 2019/2 JavaクライアントのパッケージをMaven Central Repositoryに公開(step2)

YCSBリポジトリのPullRequestを修正(step3)

⇒CI Travisをパス

```
<name>GridDB Binding</name>
<packaging>jar</packaging>

<dependencies>
  <dependency>
    <groupId>com.github.griddb</groupId>
    <artifactId>gridstore</artifactId>
    <version>4.0.0</version>
  </dependency>
  pom.xml
```

⑤ 主要OSSリポジトリへのコントリビュート(5/5) 経緯その3

- 2019/4 **YCSBメインコミットによるレビュー開始。コメントがポストされる**
- 2019/5 **指摘事項を反映**
 - ・テーブル名をコンフィグ設定可能に
 - ・printStackTrace()をloggerに変更、不要なログ出力をカット
 - ・READMEにライセンスヘッダ追加、パッケージ名の変更
 - ・テストプログラムの追加
- 2019/6 **トランクにマージ**
- ...
- 2019/9 **0.17のリリースに伴うテスト依頼あり。⇒テスト実施し報告**
<https://github.com/brianfrankcooper/YCSB/issues/1355>
0.17が正式リリース

Tested Datastores

Changed in this release and verified to work

- **NEW** Apache Crail (incubating) (#1266)
- Apache HBase on HDP 2.6.5 via `hbase098`, `hbase10`, `hbase12`, `hbase14`, and `asynchbase` bindings and CDH6 via `hbase20` binding
 - add documentation for using client side buffering (#1283)
 - changed API for table checks (#1315)
 - fixed doc for skipping WAL writes (#1324)
- Apache Ignite (incubating) 2.6.0 (via `ignite` and `ignite-sql`)
 - client version bump (#1270)
- **NEW** Azure Cosmos DB (#1264 #1298 #1300)
 - and Azure Cosmos DB Cassandra API via the `cassandra-cql` binding with SSL (#1294)
- JDBC checked with `MsSQL` and `CockroachDB 2.0.3`
 - **NEW** now works with `mssql` (#1350)
 - adds multi-row inserts (#1220)
 - **!** uses semicolon for sharding (#1233)
- MongoDB versions 4.2.0, 4.0.12, 3.6.14, 3.4.23, 3.2.22, 3.0.15, 2.6.12
 - added compatibility for MongoDB v4 (#1217)
- **NEW** Postgres NoSQL with Postgres 9.5 and 9.6 via the `postgresnosql` binding (#1242)
- **NEW** Table Store from Alibaba Cloud (#1286)
- **NEW** Toshiba GridDB (#1258)
- **NEW** VoltDB v9.1.1 via the `voltdb` binding (#1319)

「Toshiba GridDB」がリリースノートに記載された

- アプリケーション開発者向けのサイト
- 様々なコンテンツを公開
 - ホワイトペーパー
 - ブログなど
- 2019/5にリニューアル



最近のブログ

- Golangクライアント
- GridDBサーバーとクライアントDockerコンテナを使ってDevOpsを快適に使う方法
- TQLの例 (GridDBクエリ言語)
- Pythonクライアント
- NodeJSクライアント
- GridDB 4.2 Community Editionがリリースされました
- 3ノードGridDBクラスタをすばやくデプロイする方法

※最近にはビデオ付きのブログあり



Golangクライアント

By Israel In Blog Posted September 30, 2019

昨年ブログ記事で、GridDB Go Clientを使って起動し実行する方法について説明しました。その後プロセスとインストールの方法にいくつか変更があるので、このブログでアップデートします。

Tags: Golang

MORE



GridDBサーバーとクライアントDockerコンテナを使ってDevOpsを快適に使う方法

By Owen In Blog Posted September 30, 2019

このブログでは、以前投稿したDockerに関するブログ記事に基づいてGridDBサーバー、Pythonクライアント、Javaクライアントを構築し、開発者が簡単にカスタマイズ可能なコンテナをデプロイするためのDockerfileと起動スクリプトを紹介します。

Tags: DevOps, Docker

MORE

- GridDBに関するリリース、イベント、などをお知らせします。
(日本国内向け)



主催 株式会社アイ・ティ・イノベーション



無料セミナー

DB-Engines時系列DB部門👑トップ10入り

<https://db-engines.com/en/ranking/time+series+dbms>

第3回

GridDB Community Edition(CE) 概要とプログラミング基礎

～Java APIを利用したアプリケーション開発～

2019年

11/27 水 10:30～16:30 (10:00 受付開始)

■ 会場：アイ・ティ・イノベーション

■ カリキュラム

〒108-0075
東京都港区港南4-1-8 リバージュ品川5階
TEL : 03-5783-2811

講義	オープニング
	GridDBとは
	アーキテクチャ

05

まとめ

まとめ

- GridDBはビッグデータ・IoT向けのスケールアウト型データベースです。
- 主な適用事例とオープンソース活動をご紹介しました。
 - 今後も様々な拡張、拡充を進めて参ります。

GridDBのオープンソース版(GridDB CE)を是非とも使ってみてください。

<https://github.com/griddb/>

また、GridDB SEの評価版もありますので、ダウンロードしてお試してください。

https://ict-toshiba.jp/download_form_griddb/

ご参考 : GridDBに関する情報

- **GridDB GitHubサイト**

griddb github	検索
---------------	----

- https://github.com/griddb/griddb_nosql/

- **GridDB デベロッパーズサイト**

griddb net	検索
------------	----

- <https://griddb.net/>

- **Twitter: GridDB (日本)**

griddb jp	検索
-----------	----

- https://twitter.com/griddb_jp

- **Twitter: GridDB Community**

- <https://twitter.com/GridDBCommunity>

- **Facebook: GridDB Community**

- <https://www.facebook.com/griddbcommunity/>

- **Wiki**

- <https://ja.wikipedia.org/wiki/GridDB>

- **GridDB お問い合わせ**

- OSS版のプログラミング関連 : Stackoverflow(<https://ja.stackoverflow.com/search?q=griddb>)もしくはGitHubサイトの各リポジトリのIssueをご利用ください

- プログラミング関連以外 : contact@griddb.netもしくはcontact@griddb.orgをご利用ください



TOSHIBA