

分散コンピューティング部会の紹介

2021年3月5日

株式会社 ノーチラス・テクノロジーズ
<http://www.nautilus-technologies.com/>
<mailto:contact@nautilus-technologies.com>
Tel: 03-6712-0636

NAUTILUS

NAUTILUS

Copyright © 2021 Nautilus Technologies, Inc. All rights reserved.
Proprietary & Confidential

部会紹介

分散コンピューティング部会の紹介

分散コンピューティング部会の概要

■ 活動目的

- Hadoop/Spark や Asakusa Framework 等、オープンソースの分散コンピューティング技術の導入活用に向けた普及促進活動

■ 活動内容

- ビジネス面・技術面での情報収集や情報交換
- 普及促進に向けた情報展開（Web やセミナー等）
- ユーザコミュニティ／開発コミュニティとの相互協力

■ 参加企業

新日鉄住金ソリューションズ株式会社、東芝デジタルソリューションズ株式会社、株式会社ノーチラス・テクノロジーズ、株式会社日立ソリューションズ、三菱電機インフォメーションシステムズ

■ リーダー

株式会社ノーチラス・テクノロジーズ 目黒 雄一
株式会社日立ソリューションズ 才所 秀明

活動実績

■ 勉強会の主催

- 2018年6月 BigGolliraの紹介
- 2020年11月 Databricks勉強会
- 2021年3月 マイクロソフトデータ分析基盤勉強会（開催予定）

■ コミュニティ活動支援

- Project Tsurugi ユーザ会後援（2019年2月、2020年10月）

■ 部会活動

- Asakusa FrameworkDayの開催（2014年～2016年）
- Asakusa Framework適用判断ポイントのまとめ（2016年公開）

ここ最近の分散界限の話

■ 昨年より一層サービスへの道を歩んでいます

– 分散処理のソリューションがサービス化されている

■ 自分たちで構築する、という前提よりは、サービスを利用することが前提とされてきている

– クラウドが前提としており、OSSを使ったマネージド・サービスという観点が色濃い

» 様々なOSSソリューションを組み合わせるのとはなかなか難易度が高いため、サービス（主にクラウド事業者）側で、ある程度の組み合わせをしてサービス化しているものを使う、という観点になっている

– “自分たちで構築する”という観点が薄くなってきている

» それぞれのミドルウェアやアプリケーションを特徴を理解し、適切に組み合わせで構築し、運用まで実施していく、ということが出来る人が限られている

» レディメイドのサービスを使うことがよりコストメリットがあり、手間がかからないことが浸透してきている

■ 分散のミドルウェアについては、主流がMapReduceからSparkに変わった

– 一見するとSparkをサービス化しているものかどうか分からない（隠蔽されている）

新しい取り組み

Project Tsurugi(劔) ～OSSの分散DB開発～

Project Tsurugiとは

- RDBMSの開発を目的としたプロジェクト
 - DB系Tx系の企業・有志グループ・ユーザ会等により現在開発中です
- NEDO案件(経産省)
 - このProjectの成果は、国立研究開発法人新エネルギー・産業技術総合開発機構(NEDO)の委託業務の結果得られたものになるものがあります

開発の目的

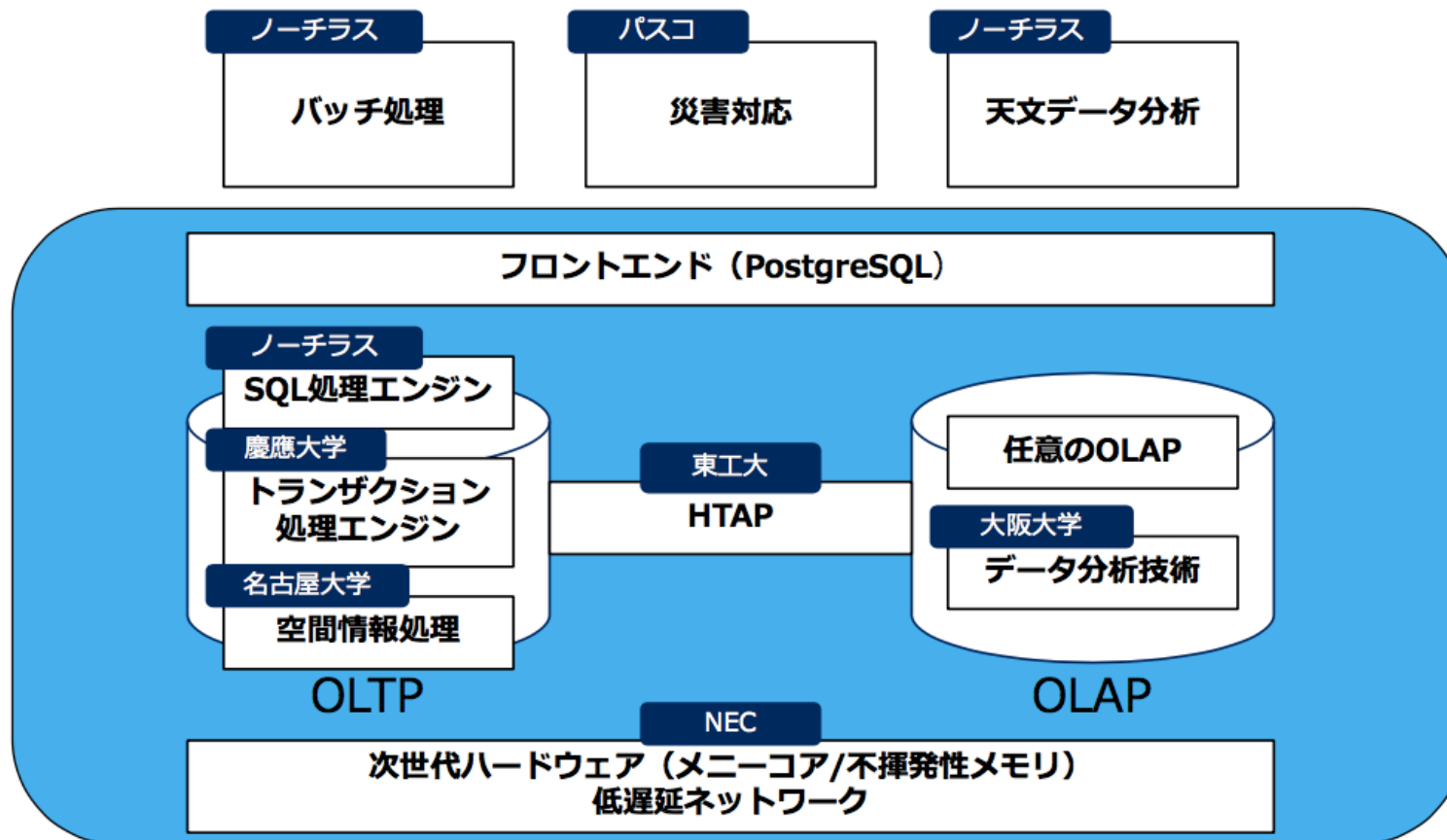
- 選択肢の確保
 - 現状のH/Wアーキテクチャに合致したRDBMSの選択肢をより増やす
- 透明性
 - より広く利用可能であること
 - 実装透明性を確保するためにOSSで開発する
- 受け入れやすいものを
 - エンジンを中心:外側はPostgresそのものとする
 - 既存ユーザに受け入れやすいようにする
- HTAPで利用可能な形態にする
 - 分析と書き込みの両方ニーズを満たす必要がある
- コミュニティベースの開発を企図
 - 可能な限り国内のDB研究者や開発者を集めて共同でつくり上げる
 - すでにユーザ企業を中心としたコミュニティが発足している
 - 協調関係をとりながら加速させる

システム概要

■ 全体像

- PostgreSQLをベースに「新たに開発するOLTPエンジン」と「任意のOLAP」を組み合わせた「**HTAPシステム**」
 - 現時点では、OLAPとしてPostgreSQL(リファレンス)とApache Sparkを想定
 - その他のOLAPについても組み込みのためのフレームワークを提供予定
 - フロントエンド(表見え)はPostgreSQLだが、実際にクエリを処理するエンジンは別になる
 - クエリの入力や結果セットの取得はPostgreSQLのインタフェースから可能
 - 1つの接続先からOLTP、OLAPそれぞれのデータにアクセス可能
 - OLTPに書き込まれたデータはOLAPにも反映される(HTAP)
 - SQLはPostgreSQLの仕様をベースに基本的な命令文から対応予定
 - 運用管理機能(バックアップなど)は独自拡張予定
- ### ■ 基本的にはPostgreSQLに近い使い勝手になることを目指す

システム概要図



今後について

- 研究事業は2022年度まで
- 事業終了段階でa版の提供を予定
 - PoCの実施を可能とする
 - SIとの組み合わせであれば、システムへの組み込みは可能なレベルにする予定
- もっと知りたい方は、以下のページに昨年の10月にユーザ会で説明した資料を公開しているので御覧ください
 - https://www.osscons.jp/distributed_computing/download/



OSS Consortium 分散コンピューティング部会