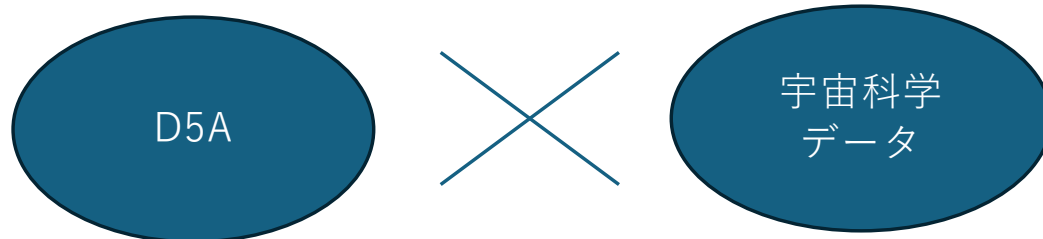


Apollo月震データへの高速 データ表示へのアプローチと その発展

山本幸生¹ 古庄晋二² 飯沢篤志 手塚宏史 長尾正
松久孝志³ 飯田学³ 岡島和裕³ 光内章³ 高木亮治¹

1. 宇宙航空研究開発機構 2. NNIテクノロジーズ株式会社 3. 株式会社セック

増え続ける宇宙科学資産の利活用



D5Aの持つCOLD DATAへの
高速なアクセス

更新が止まった
膨大な過去の宇宙科学資産

D5Aフォーマットの特徴

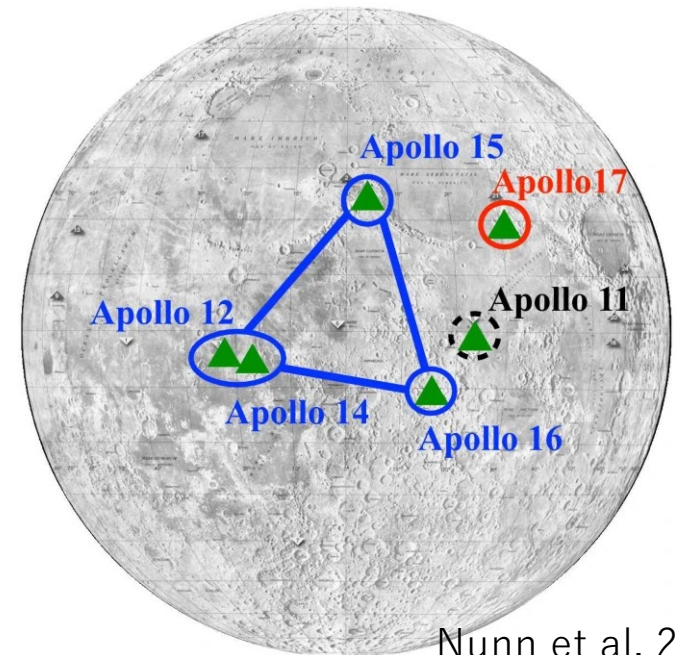
- 更新のないデータに対して高速に動作
- 自然数インデックス(Natural Number Index; NNI)を利用した「検索」「集計」「ソート」等を高速化するインデックス
- UNIONやLatchなどのデータの組合せを行ってもインデックスが失われない。
- 細かいチューニングが不要
- 全てのカラム／カラムの組合せ／部分集合にインデックスが自動的に付与される
- カラム数に対する制約の緩さ（最大10万カラムまで実験済み）

Apollo月震データのD5A化

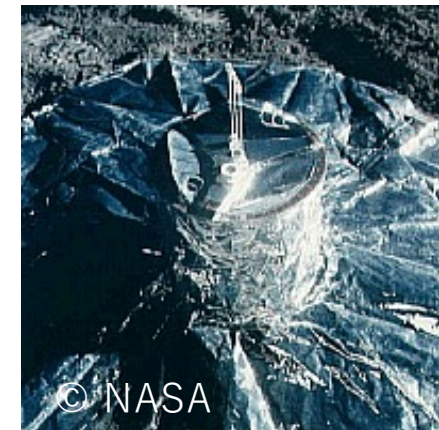
月惑星探査アーカイブサイエンス拠点集会2024

Apollo月地震データとは？

- Apolloミッションのサイエンスパッケージ(ALSEP)の一つである地震計等による
人類が唯一保有する月震データ
- テキサス大中村教授によりアーカイブされ、その後ISASに送付
- 独自フォーマットで総計110GBのデータ
- 生データがJAXAのDARTSウェブサイトからダウンロード可能
- CSV形式でデータをダウンロード可能なビューワMoon Seismic Monitorを開発(2012~2020)



Nunn et al. 2020



2012年に開発したシステムの問題点

- User Interfaceの問題
 - Flashを用いていた → 後に脆弱性のためアプリごと終了
 - 画面上で波形をポイントしても値表示が困難だった
- Databaseの問題
 - データベースに汎用のPostgreSQLを用いておりオーバーヘッド大
 - 月震データの区間を3時間とするとデータ転送が10分以上経っても始まらない
 - さらに伸ばすとハングアップ
 - 高速化のために専門家を雇用し多大なチューニングを実施
 - Array型とPL/pgSQLによる展開

たかだか110GBでなんでこんなことになるのか？

Apollo月地震データのデータ構造

ALSEP word 1 CONT. 905	ALSEP word 2 CONT. 59	ALSEP word 3 CONT. 380	ALSEP word 4 SPZ 0	ALSEP word 5 LSM 1023	ALSEP word 6 SPZ 0	ALSEP word 7 SWS -	ALSEP word 8 SPZ 0
ALSEP word 9 LPX 524	ALSEP word 10 SPZ 0	ALSEP word 11 LPY 495	ALSEP word 12 SPZ 0	ALSEP word 13 LPZ 497	ALSEP word 14 SPZ 0	ALSEP word 15 SIDE -	ALSEP word 16 SPZ 0
ALSEP word 17 LSM -	ALSEP word 18 SPZ 0	ALSEP word 19 LSM -	ALSEP word 20 SPZ 0	ALSEP word 21 LSM -	ALSEP word 22 SPZ 0	ALSEP word 23 SWS -	ALSEP word 24 SPZ 0
ALSEP word 25 LPX 525	ALSEP word 26 SPZ 0	ALSEP word 27 LPY 494	ALSEP word 28 SPZ 0	ALSEP word 29 LPZ 497	ALSEP word 30 SPZ 0	ALSEP word 31 SIDE -	ALSEP word 32 SPZ 0
ALSEP word 33 HK 35	ALSEP word 34 SPZ 0	ALSEP word 35 TidX 331	ALSEP word 36 SPZ 0	ALSEP word 37 TidY 234	ALSEP word 38 SPZ 0	ALSEP word 39 SWS -	ALSEP word 40 SPZ 0
ALSEP word 41 LPX 525	ALSEP word 42 SPZ 0	ALSEP word 43 LPY 495	ALSEP word 44 SPZ 0	ALSEP word 45 LPZ 496	ALSEP word 46 CV 0	ALSEP word 47 SIDE -	ALSEP word 48 SPZ 0
ALSEP word 49 LSM -	ALSEP word 50 SPZ 0	ALSEP word 51 LSM -	ALSEP word 52 SPZ 0	ALSEP word 53 LSM -	ALSEP word 54 SPZ 0	ALSEP word 55 SWS -	ALSEP word 56 SIDE -
ALSEP word 57 LPX 525	ALSEP word 58 SPZ 0	ALSEP word 59 LPY 495	ALSEP word 60 SPZ 0	ALSEP word 61 LPZ 496	ALSEP word 62 SPZ 0	ALSEP word 63 SIDE -	ALSEP word 64 SPZ 0

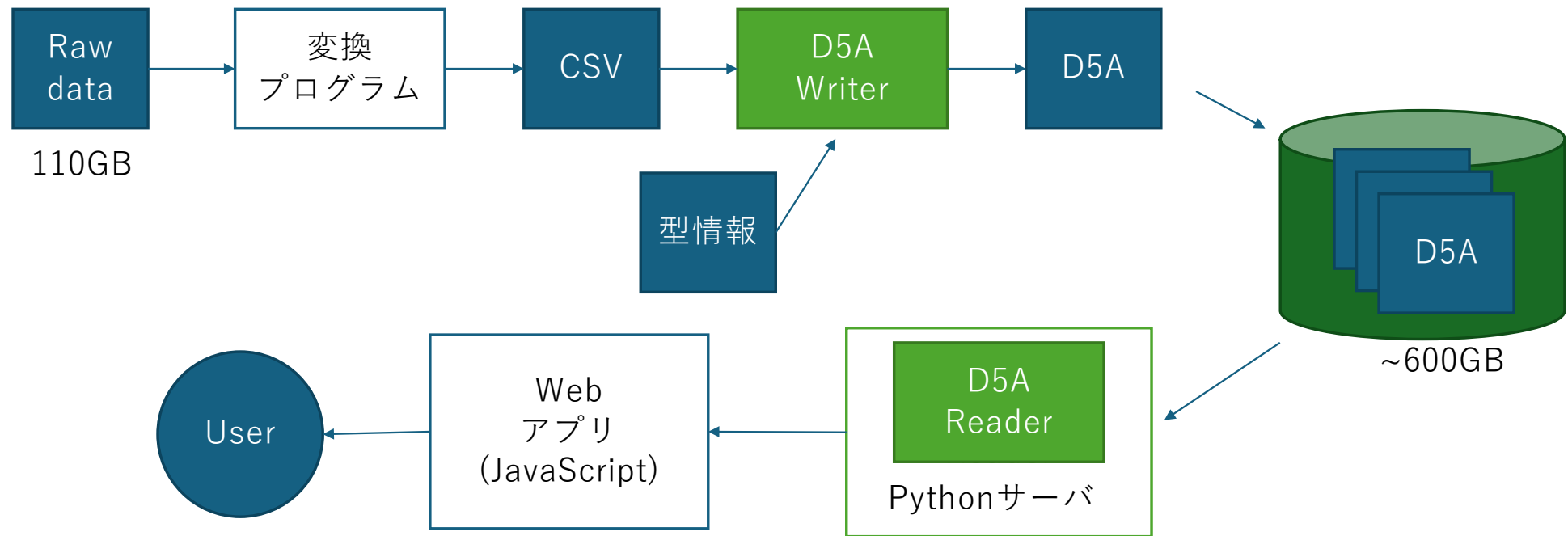
- 1フレームを0.604秒で伝送
- SPZが31個 ~ 53Hz
- LPX/LPY/LPZが4個 ~ 6.625Hz

53Hzを1年間運用すると
 $53 * 86400 * 365 \sim 16.7$ 億

実際の運用期間は7.5年
 さらに号数Apollo12,14,15,16,17

~10億レコードをほどほどのマシンで
 高速に扱いたい

D5Aフォーマットへの変換と利用



新しいApollo月震システム

<https://darts.isas.jaxa.jp/app/apollo/>



- ウェブアプリとして十分な速さ
- 表示範囲を1日にしても軽快
- CSVでダウンロード可
- D5Aでもダウンロード可
- マウスオーバーで値表示

当時のD5Aフォーマットの課題

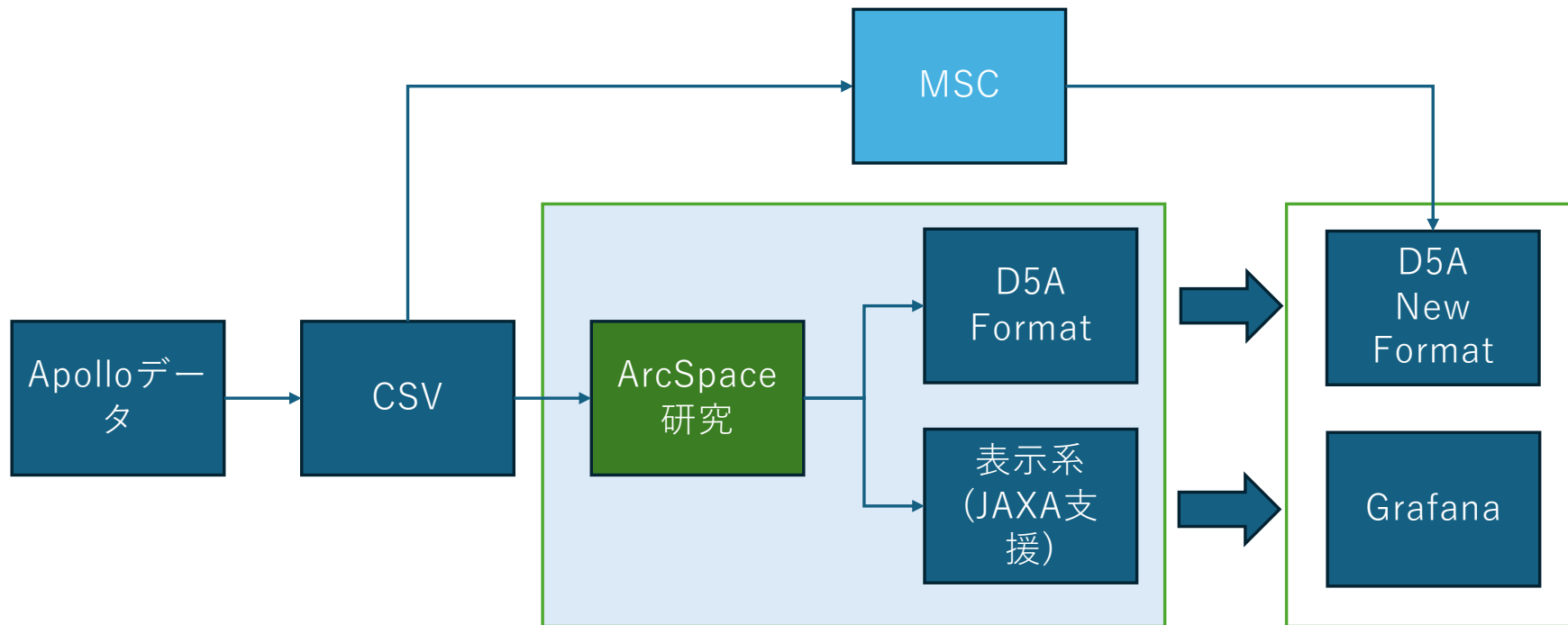
- 支援を受けて開発したApolloデータのD5A化(旧D5A)の課題
 - CSVをD5A変換するWriterがPASCALで書かれていた → コンパイルに難あり、将来性に不安
 - D5Aファイルに圧縮機能がなくデータサイズが肥大化した(~600GB)
→ 巨大データを考慮した場合に、データサイズが数倍になると、そのまま金銭的成本に跳ね返る
- 周辺ツールとの連携が弱い
 - Apollo月震データのViewerは専用のツールを作成

D5Aの進化

新たなD5Aへの発展

- 目的別データベース
 - D5Aの持つ「複数の表形式データを低コストで組み合わせ可能」な特徴を元にMulti Source Composer (MSC)へと発展
 - 複数の表形式データを仮想的に結合して大きな表形式データとして表現可能な「D5AVU」を構築
 - 複数拠点に置かれたD5AファイルをD5AVUを通して一つのテーブルとして取り扱い可能
- CSVをD5AにするWriter/ReaderがPascalからC++へ
- D5Aが新しくなり圧縮に対応
- D5Aを閲覧するために専用アプリケーションを開発する必要があったが、汎用グラフソフトウェアのGrafanaのバックエンドが動作

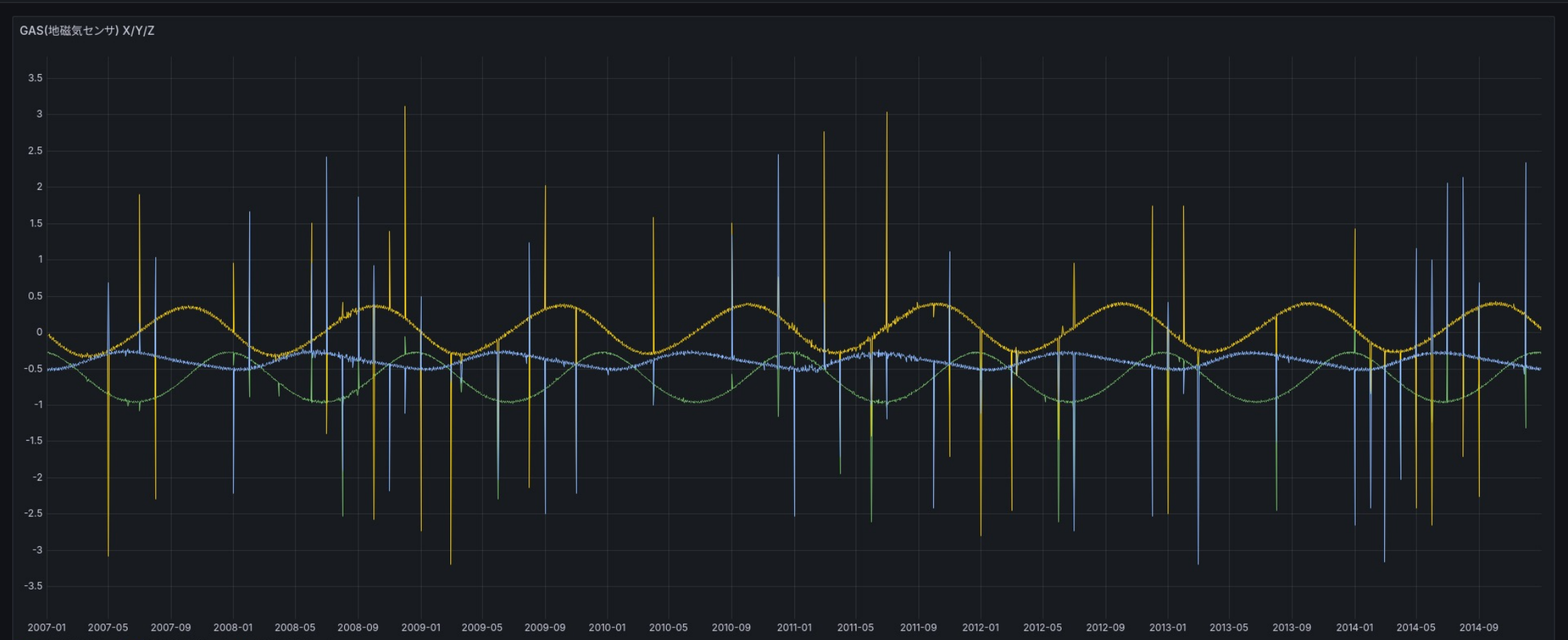
ArcSpace研究と新たなD5Aの関係



ArcSpace研究によるApolloデータのD5Aフォーマットによるグラフ表示を実証したことによりD5Aフォーマットの有用性が証明できた。本実証の経験は、新たなD5A開発への足掛かりとして重要な実証実験だった。

MSCを用いた高速アクセスのデモ

- データ
 - Apollo月震計SPZ
 - 太陽観測衛星ひので(SOLAR-B)
- デモ内容
 - 長大期間におけるアクセス性能
 - 短期間データの特徴を逃さないピラミッディング
 - 汎用グラフツールのGrafanaで表示



月惑星探査アーカイブサイエンス拠点集会2024