

令和4年度 卒業研究発表会



気象庁GRIB2フォーマットのパーサーの開発

山形県立産業技術短期大学校庄内校 情報通信システム科
伊藤 肇、鈴木章祐、高橋 あい

チーム名 : Météo



説明内容

1. 背景
2. 現状
3. 課題と解決目標
4. システム概要
5. 運用
6. アピールポイント
7. まとめ



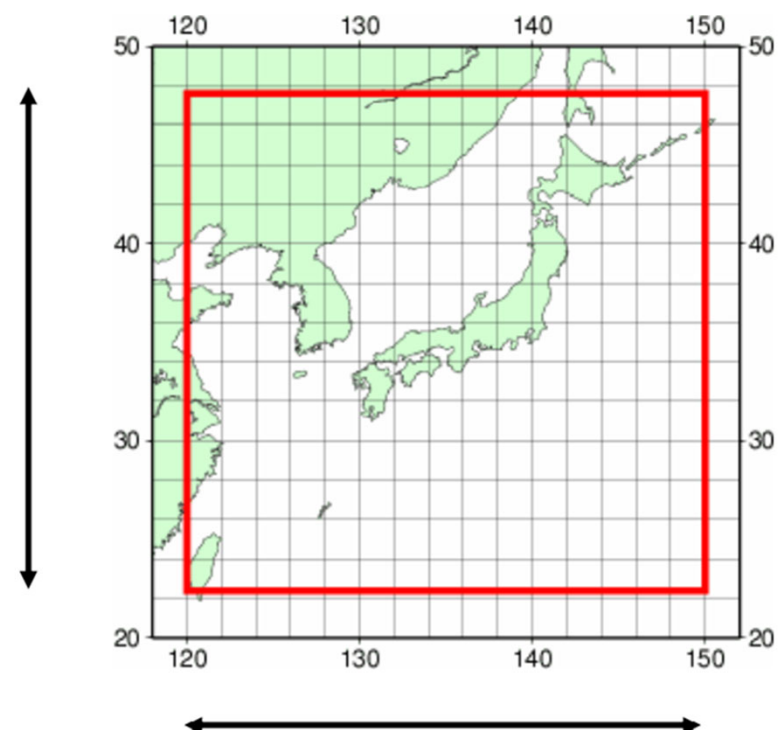
1.背景

昨今の激甚気象災害の被害を
最小限にしたい



重要

数値予報モデルGPVなどの
格子点形式の気象ファイル



1.背景

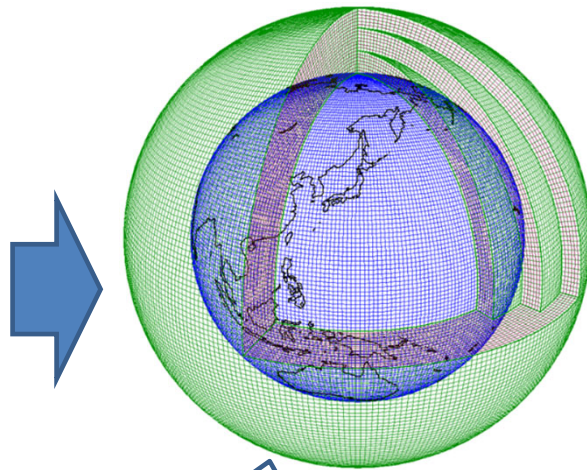
数値予報とは

気象レーダー

気象衛星

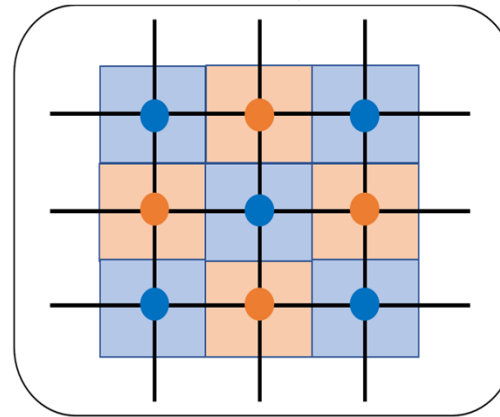
アメダス

直接観測

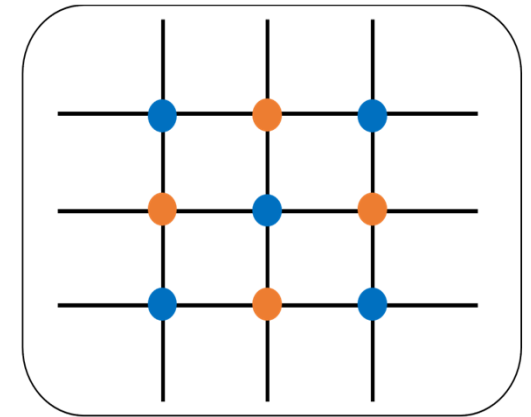


分割した格子点に気温・風などの
気象要素の値を割り当てる

GPV(Grid Point Value)



○格子点領域の代表値

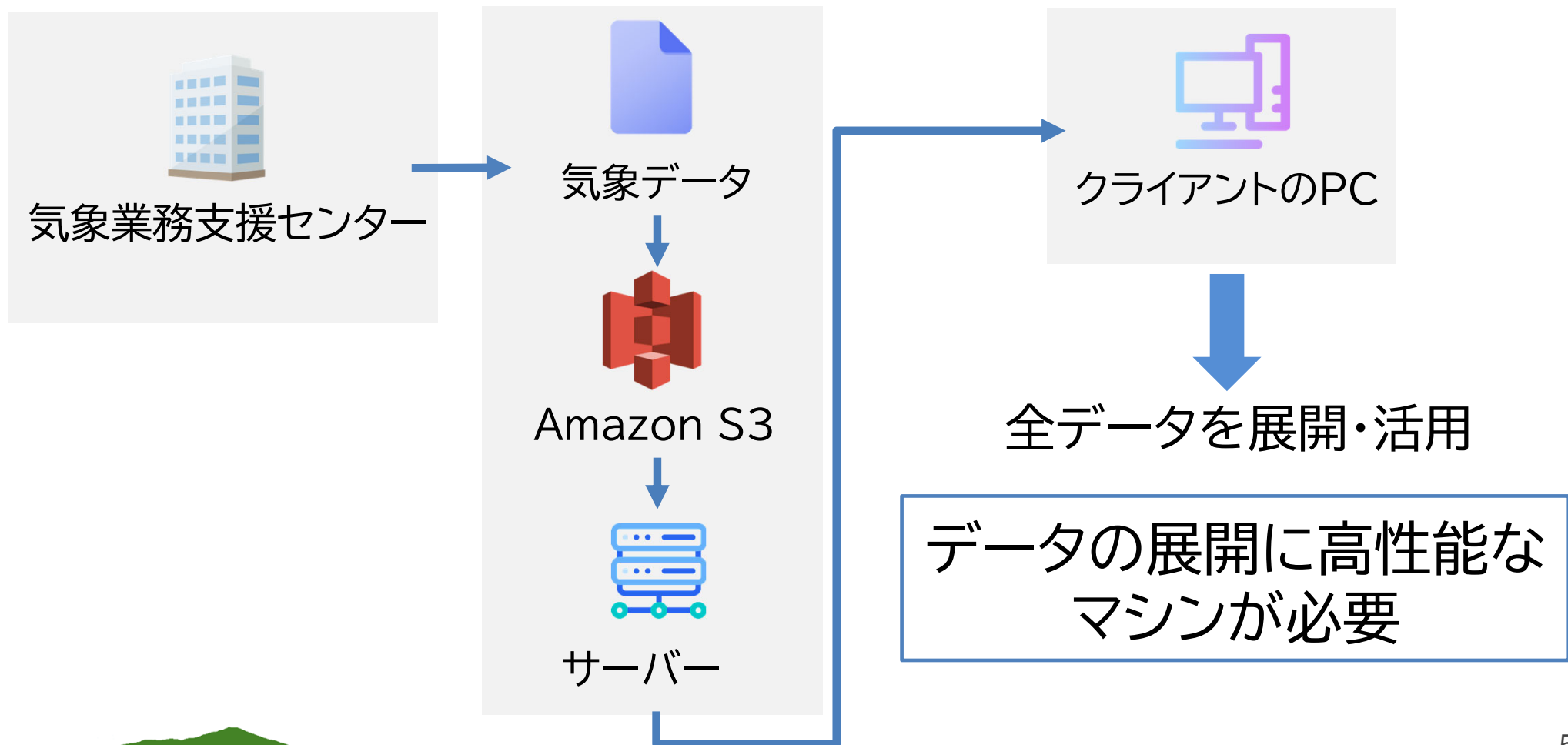


×格子点ピンポイントの地点の値



1.背景

天気予報のシステム



1.背景

気象データはGRIB2と呼ばれる高圧縮な形式で配信されている

CSVデータ(男女別人口 国勢調査)

1	"都道府県コード","都道府県名","元号","和暦(年)","西暦(年)","注","人口(総数)","人口(男)","人口(女)"↓
2	"00","全国","大正",9,1920,"",55963053,28044185,27918868↓
3	"01","北海道","大正",9,1920,"",2359183,1244322,1114861↓
4	"02","青森県","大正",9,1920,"",756454,381293,375161↓
5	"03","岩手県","大正",9,1920,"",845540,421069,424471↓
6	"04","宮城県","大正",9,1920,"",961768,485309,476459↓

GRIB2データ(気象庁サンプルファイル)

	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	Decoded Text
	52	49	42	FF	FF	00	02	00	00	00	00	04	20	8A	31		G R I B 1
	00	00	15	01	00	22	00	00	02	01	01	07	E1	0C	05		 "
00000020	00	00	00	00	01	00	00	00	48	03	00	00	03	B4	D9	00	 H
00000030	00	00	00	06	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	FF	

簡単に活用
できない

1.背景

GRIB2形式なので気象庁のデータフォーマットを知らなければ扱いにくい

地上物理量に用いるGRIB2のフォーマットおよびテンプレートの詳細

オクテット	内容	表	値	備考
1~4	GRIB		"GRIB"	国際アルファベットNo.5(CCITT IA5)
5~6	保留		missing	
7	資料分野	符号表0. 0	0	気象分野
8	GRIB版番号		2	
9~16	GRIB報全体の長さ		*****	69,241,393(0-15時間予報) 78,716,561(16-33時間予報) 26,238,929(34-39時間予報) 52,477,745(40-51時間予報) 118,074,785(52-78時間予報)
1~4	節の長さ		21	
5	節番号		1	

気象庁GRIB2データフォーマット



デジタルインフラとしてどうなのか？



誰もが扱いやすい形で配信したらよいのではないか

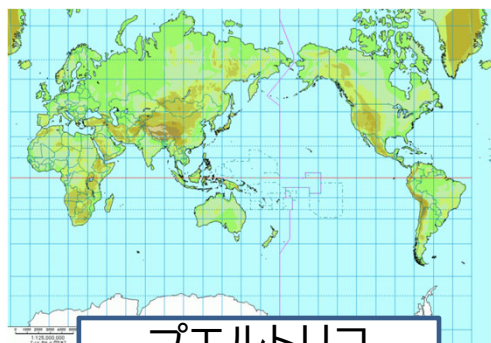


2.現状

気象庁仕様のGRIB2ファイル

座標系の違い

メルカトル図法



プエルトリコ
ハワイ
オセアニア

ユニバーサル 極心平射図法



アラスカ

ランベルト 正角円錐図法



アメリカ本土

NOAA
(アメリカ国立気象局)

気象庁

緯度・経度



2.現状

気象庁仕様のGRIB2ファイル

データフォーマット

13	背景作成処理識別符	JMA定義	*****	31=メソ数值予報（数值予報モデルの改良により変更される場合がある）
----	-----------	-------	-------	------------------------------------

気象庁独自の定義(JMA定義)が使われている

3.課題と解決目標

課題

- 気象庁の独自仕様(GRIB2から気象庁がカスタマイズし、一部をそぎ落としたデータ)のため、NOAA(米国海洋大気庁)の処理プログラムでは一部しかデータを取り出せない
- 高精度なビッグデータである気象データを利活用するための基盤がない
- GRIB2形式のままでは、AI等による処理・データの可視化・データの蓄積等ができない

3.課題と解決目標

解決目標

- 気象庁の仕様に合わせてGRIB2ファイルをデコードできる
- デコードしたデータを活用する基盤を作る
- デコードしたデータをデータベースに蓄積する

4.システム概要

システムの機能

- 京都大学生存圏研究所が有償のGRIB2データを研究用に配布しているので、それを自動取得する
- GRIB2データの解析を行い、JSONファイルで出力する
- CSVファイル等を元にデータベースへ気象データを格納する
※今年度はメソ数値予報モデルGPV(15時間予報)を対象とする
- データベースから古いデータを削除する



4.システム概要

前提条件・制約条件

- クラウド環境に構築するためコンテナ化する
- パーサー(GRIB2の読み出し)部分は高速性を求める
処理の早いプログラミング言語を使うこと

4.システム概要

本研究で使用するソフトウェア

OS



Ubuntu Server 22.04 LTS

仮想化システム



コンテナ型のアプリケーション
実行環境

フレームワーク



マイクロサービスの開発を
容易にするフレームワーク

ソフトウェア開発
プラットフォーム



ソースコードを管理するツール



4.システム概要

本研究で使用するソフトウェア

プログラミング言語



ライブラリが豊富で可読性の高い言語



実行速度が高く、汎用性の高い言語

データベース
管理システム



豊富な機能、拡張性を備えたリレーショナルデータベース
※アクセスにはpgAdmin4を使う

データベース
拡張モジュール

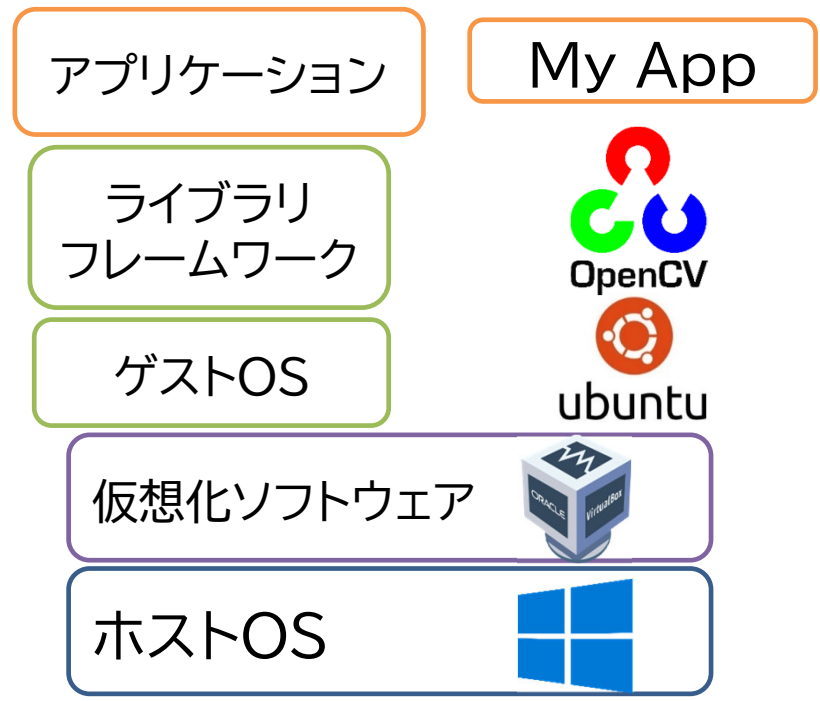
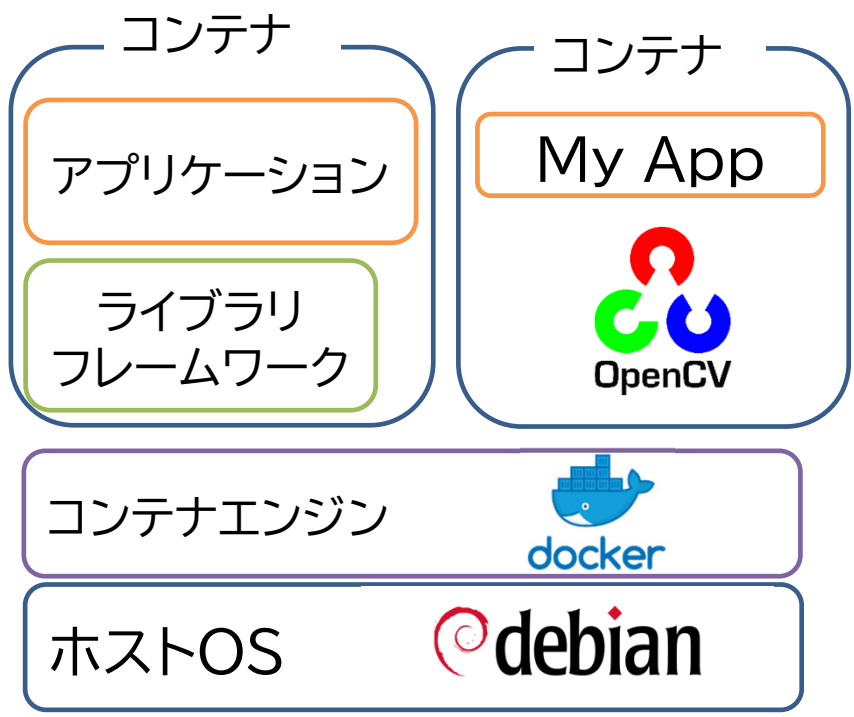


PostgreSQLで地理空間情報を扱うための拡張



4.システム概要

コンテナ化とは アプリケーション、ライブラリなどをひとまとめでした
仮想化技術

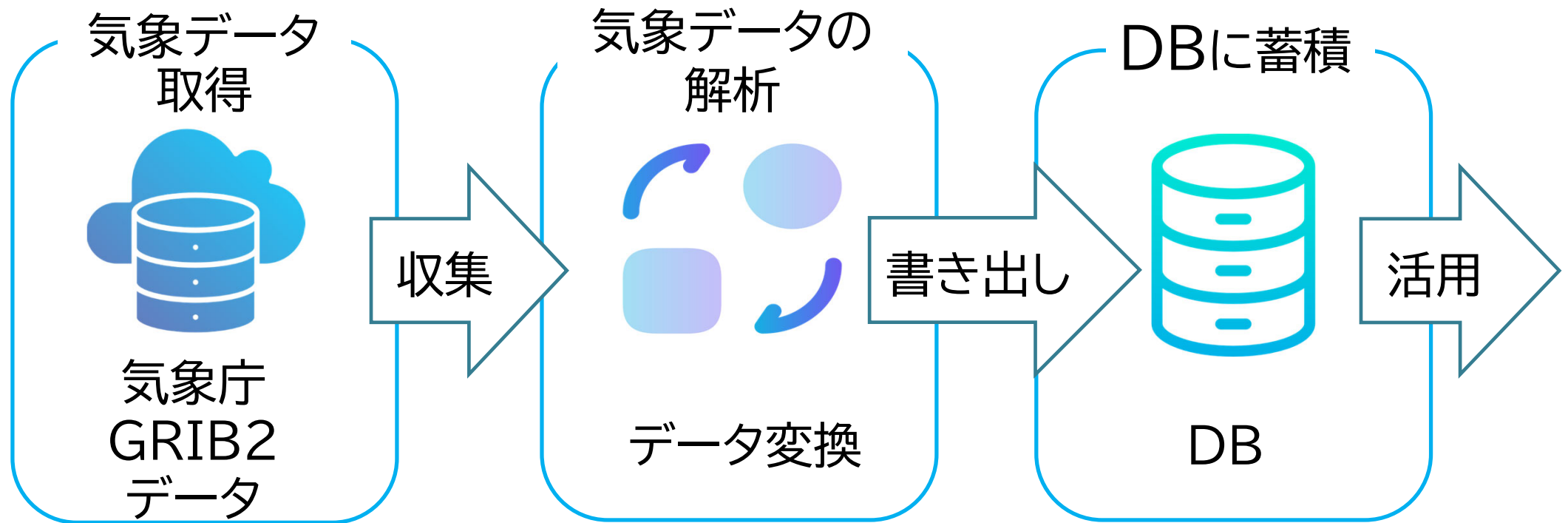


コンテナ化したシステム

従来の仮想化

4. システム概要

ETLとは



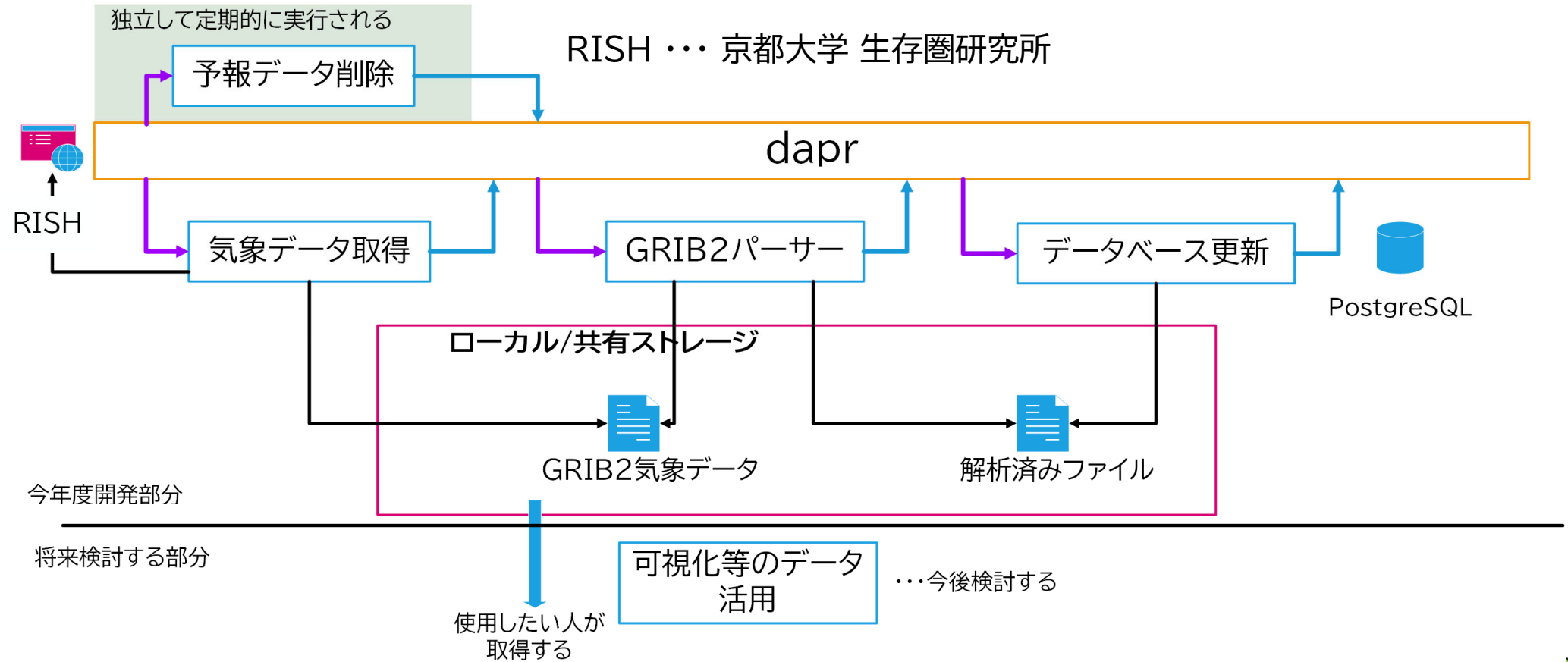
E = Extract:抽出 T = Transform:変換 L = Load:書き出し



4.システム概要

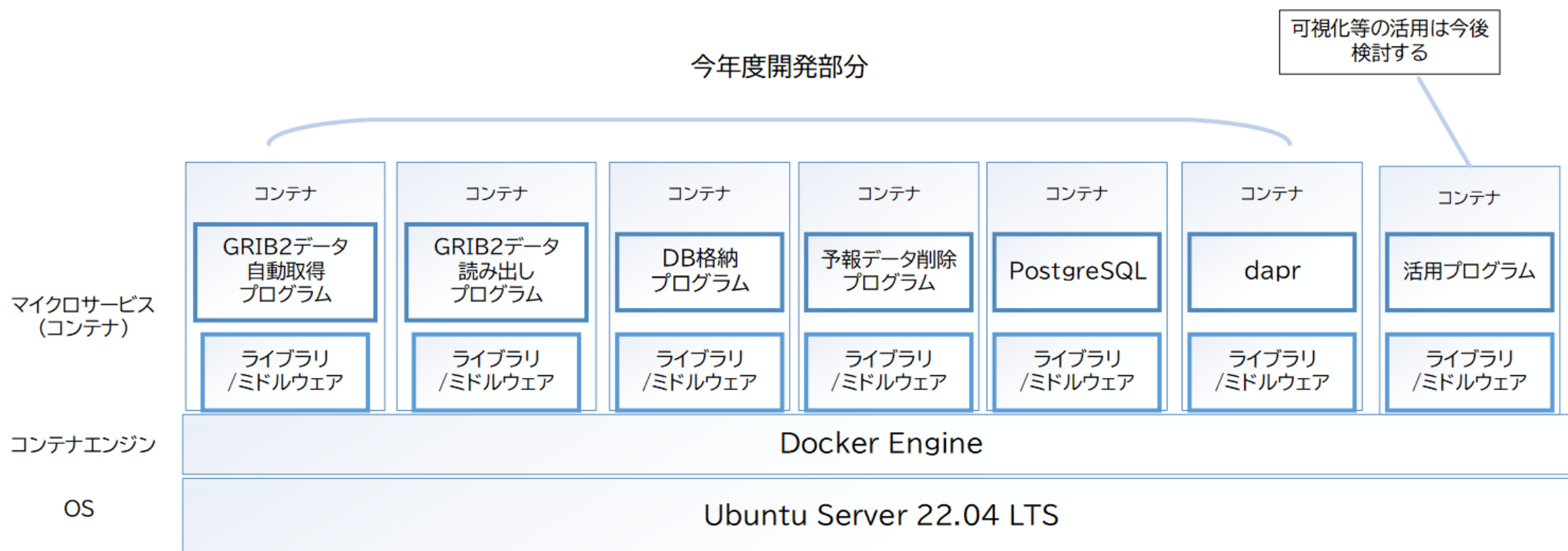
システム構成図

□ …マイクロサービス → …daprからの起動命令
→ …ファイルの読み取り/書き込み → …HTTP通信



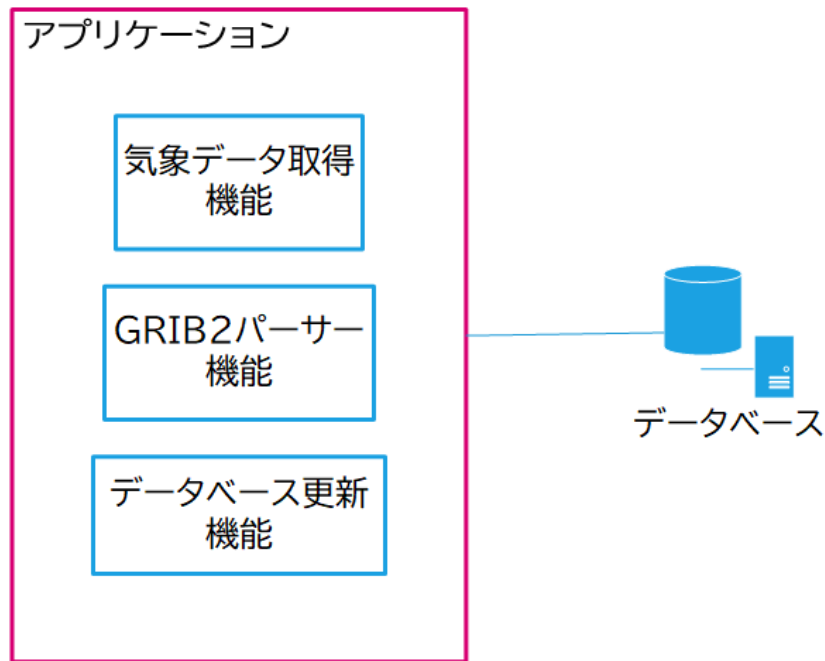
4.システム概要

ソフトウェア構成



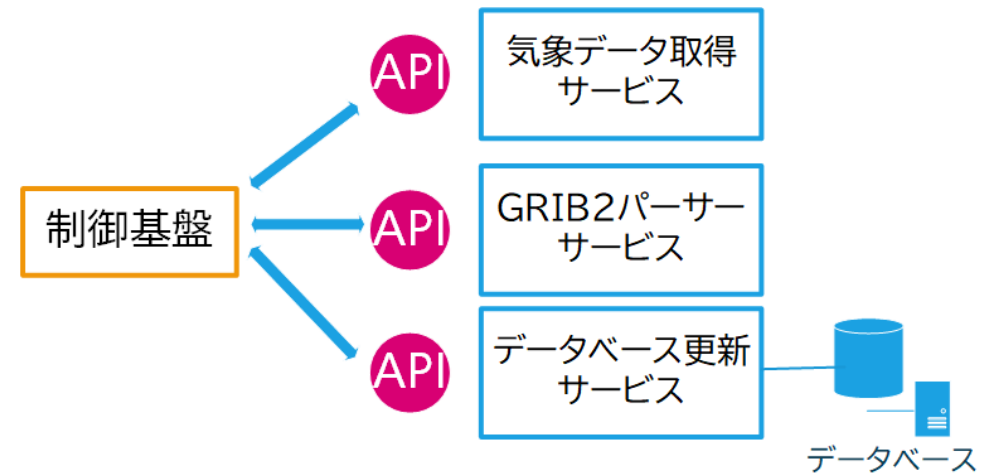
4.システム概要

モノリシックアーキテクチャー



1つのアプリケーション内に
様々な機能

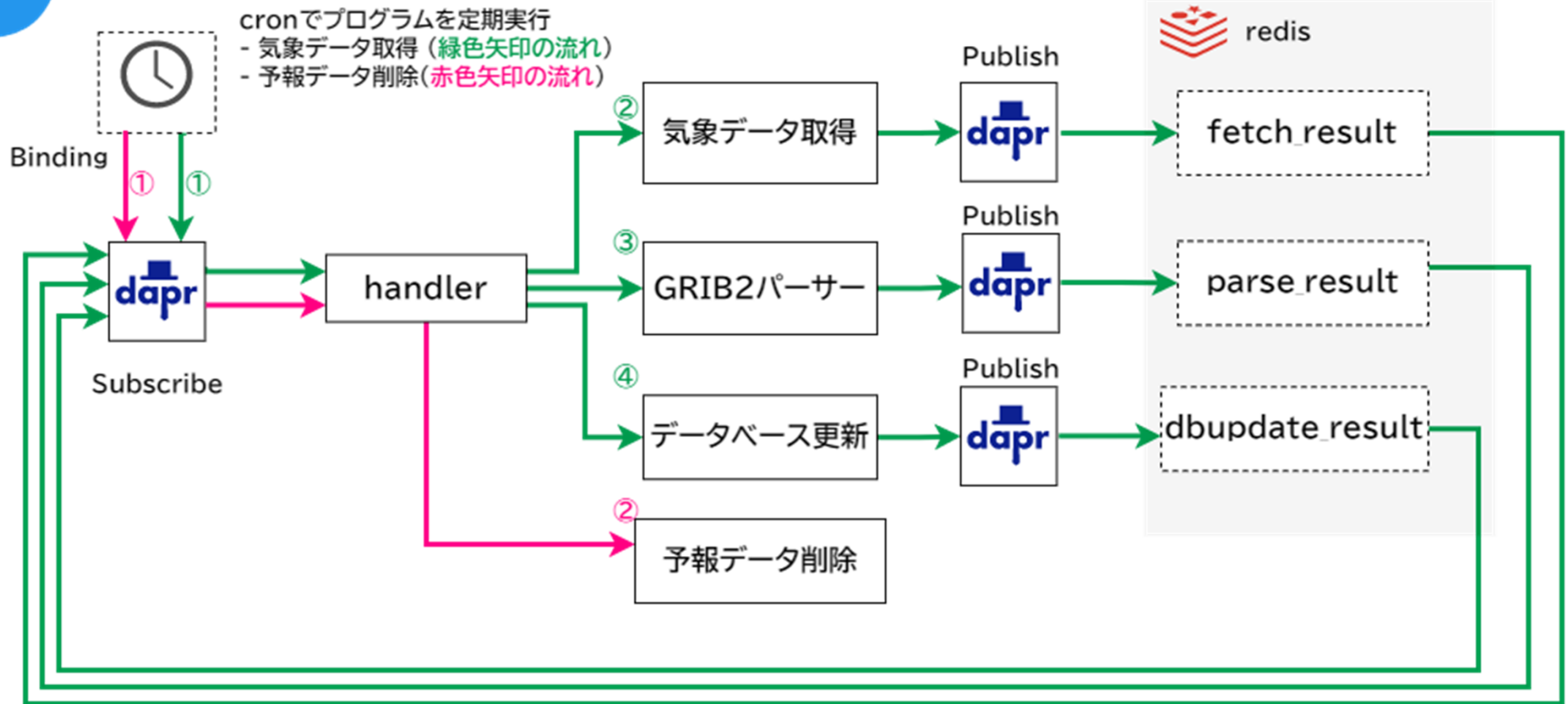
マイクロサービス アーキテクチャー



小さなプログラムを組み合わせ
大きな機能を実現



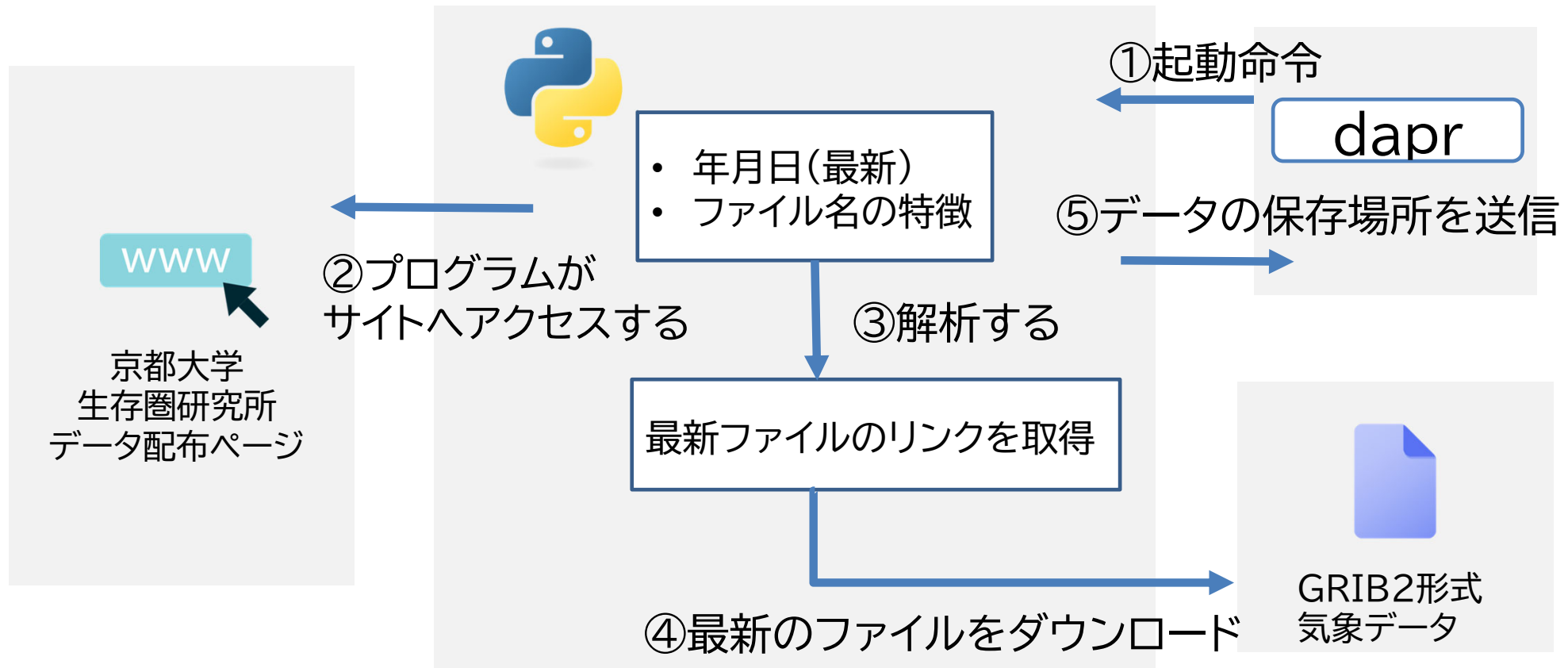
4.システム概要



4.システム概要

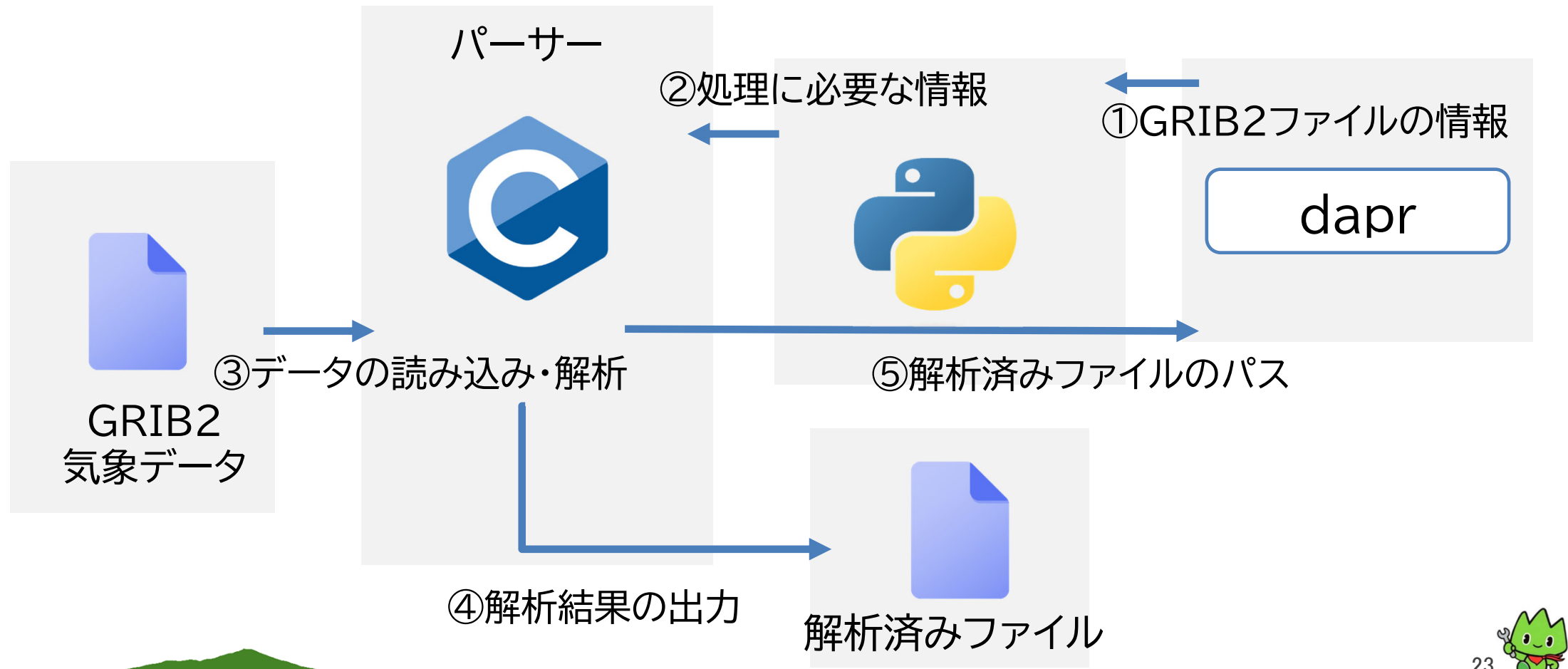
気象データ取得サービス

気象データ取得プログラム



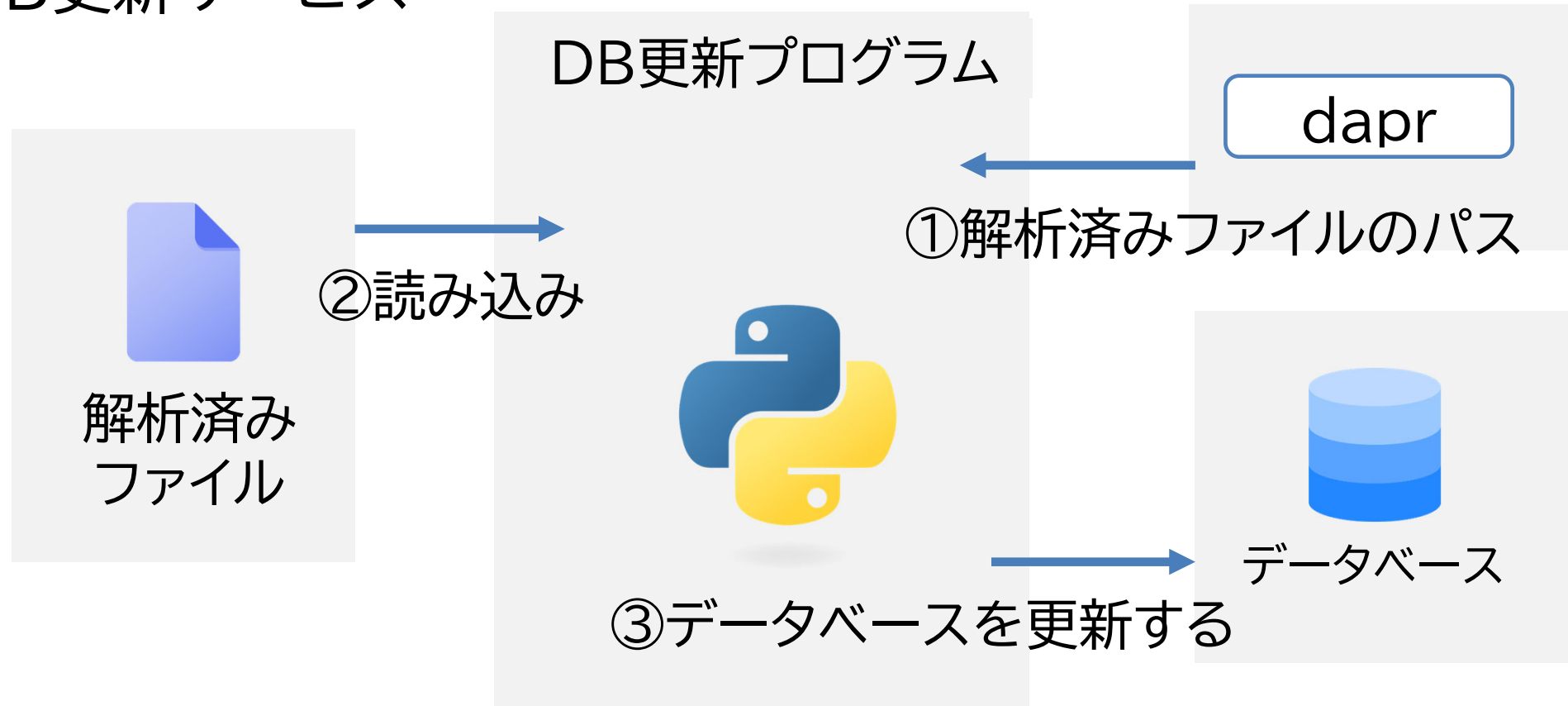
4. システム概要

GRIB2パーサーサービス



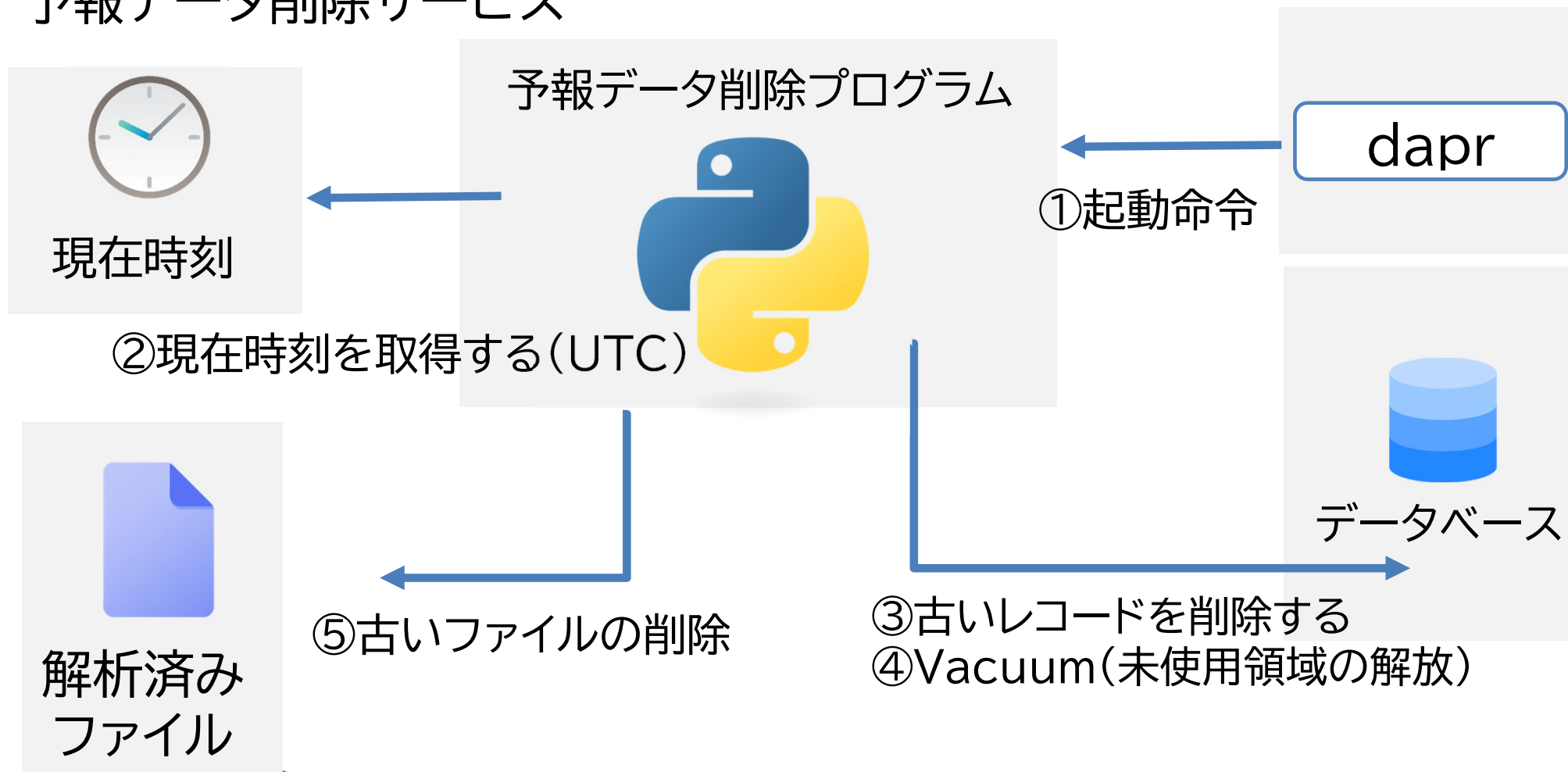
4. システム概要

DB更新サービス



4. システム概要

予報データ削除サービス



5.運用

解析済みデータの内容

例:メソ数値予報モデルGPV地上面 日射量

JSONファイル形式で出力

Valueは配列形式で格納

Key	数値予報モデル	msm_gpv_surf
Value	基準時刻	2023-02-03 00:00:00
	データ時刻	2023-02-03 00:00:00
	開始時刻(統計値のデータのみ)	2023-02-03 00:00:00
	終了時刻(統計値のデータのみ)	2023-02-03 01:00:00
	データの種類	Solar_Radiation
	緯度	22.400000
	経度	150.00000
	格子点値	586.934204
	データの単位	W・m ⁻²

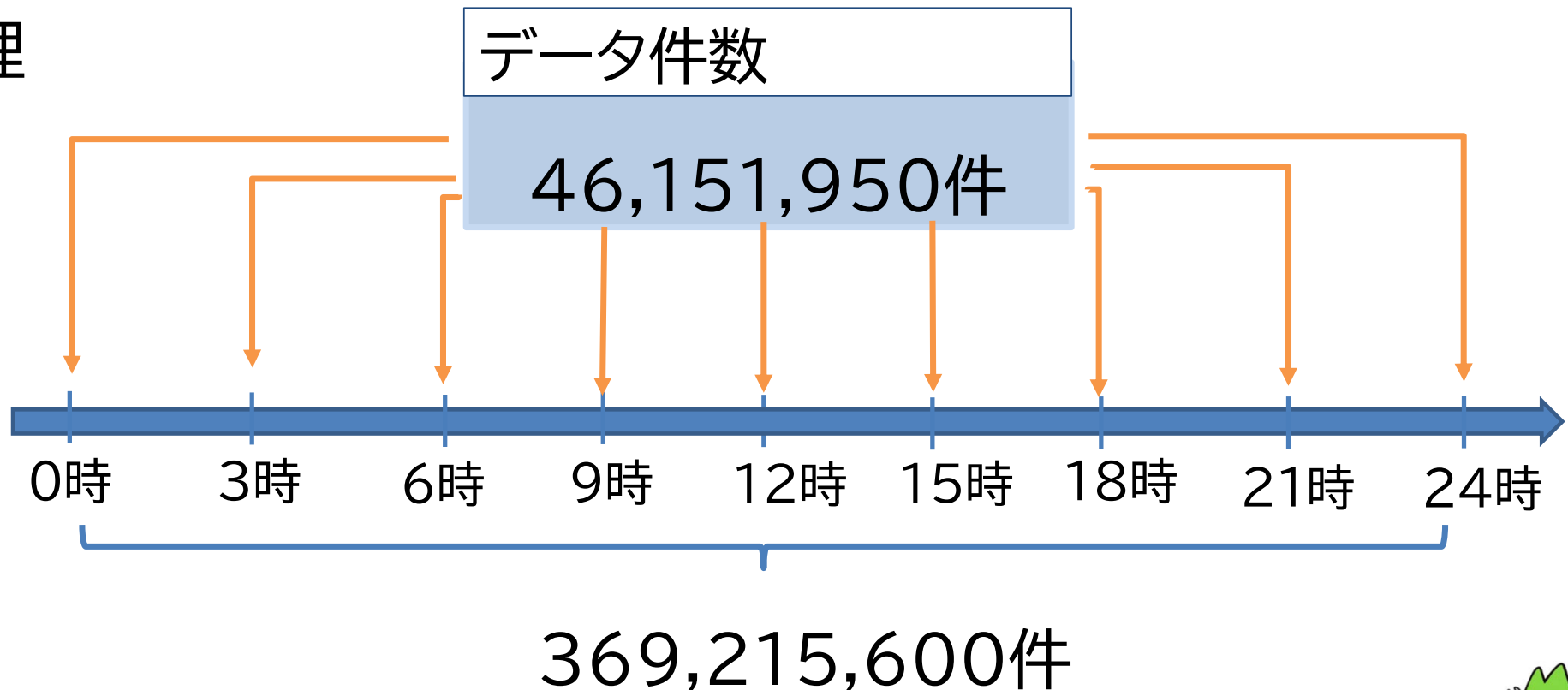
5.運用

予測開始から15時間分のデータのみ実行した

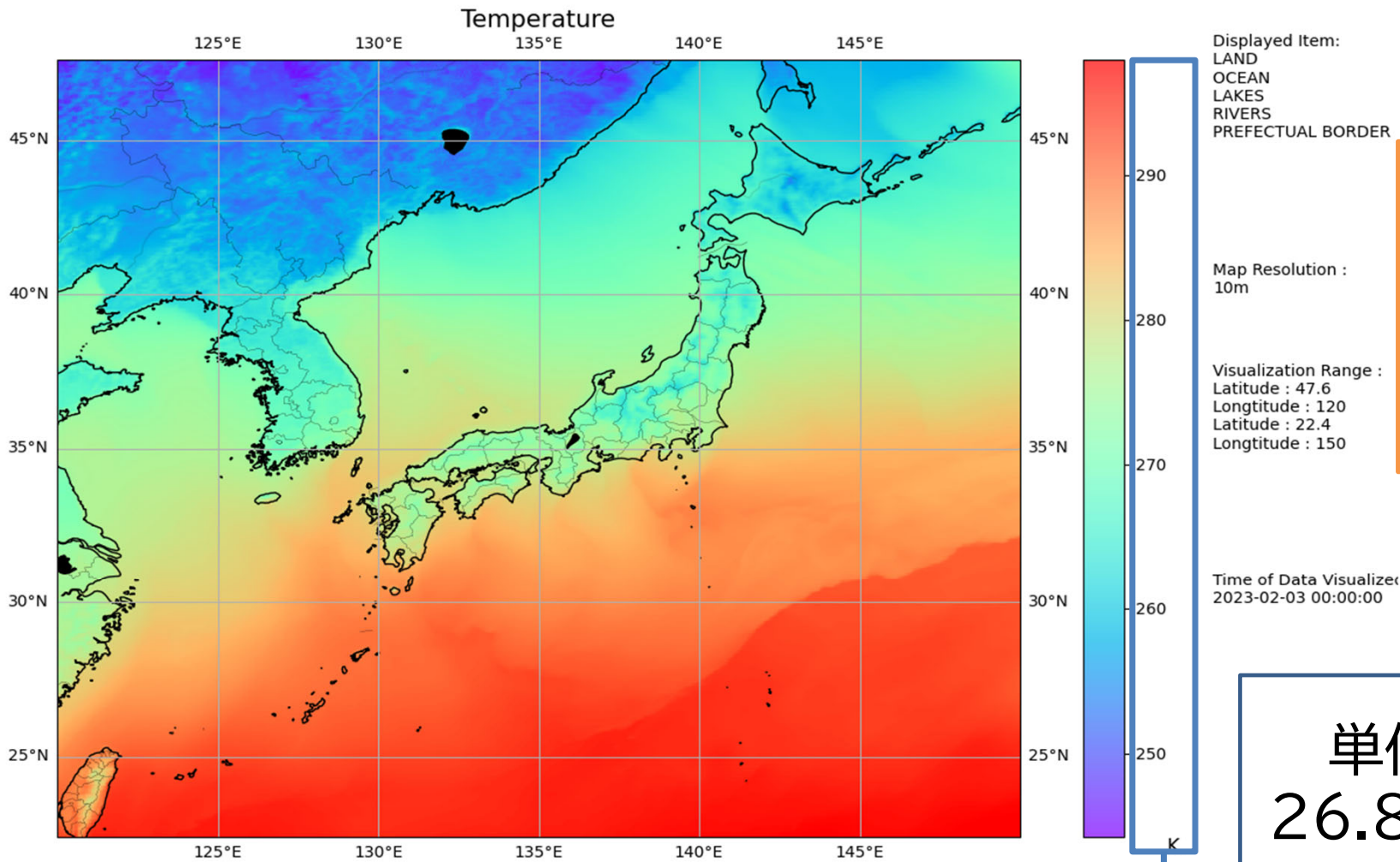
3時間ごとに新しい
データが更新される

1回の処理

1日運用



5.運用



使用したデータ

メソ数値予報モデル
2023-02-03
00:00:00
気温
(Temperature)

単位を摂氏に直すと
26.85℃から-33.15℃



6.アピールポイント

- パーサーサービスから、JSON形式で出力するので他のプログラムで活用しやすい
- データベースに格納するので蓄積したデータを活用しやすい
- 気象庁より細かい粒度での気象情報の取得が可能
- 取得できるデータの量が多い



6.アピールポイント

1時間当たりのデータ量の比較

本システム(JSON形式)

407MB

気象庁(CSV形式)

1.117MB

信頼性の高いデータを
取得できる

GRIB2でのみ取得できるデータ

相対温度、風の東西成分、風の南北成分、上昇流、地上気圧、海面更生気圧、高度、日射量、全雲量、下層雲量、中層雲量、上層雲量

7.まとめ

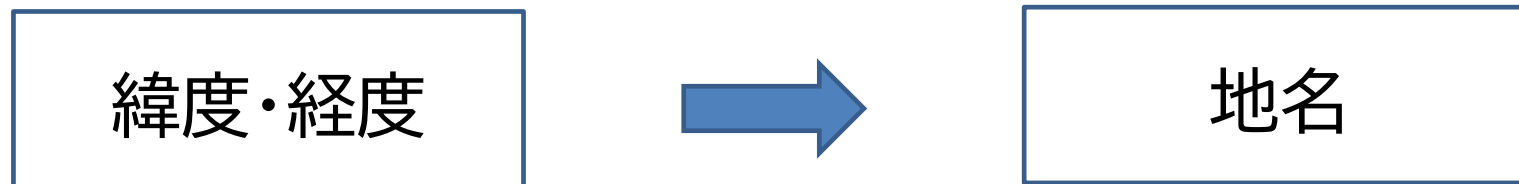
成果

- GRIB2の解析を行うシステムを作ることができた
- どのような環境でも利用しやすく、扱いやすいJSONファイル形式で出力することができた
- データベースに解析後のデータを蓄積し、活用できるETL基盤を作ることができた

7.まとめ

今後の展望

- エリア(市町村など)別にピンポイントでデータを利用できるようにする



- 解析できるGRIB2のファイルを増やす
- 圧縮形式が異なるデータでも解析できるようにする