

## 地震データ大規模解析サーバーの利用について

平成 27 年 11 月 30 日  
東京大学地震研究所

### 目的

長期間にわたる連続地震波形データ等を解析する際に共通して利用する計算資源を集中させた研究支援基盤を設置運用することにより、準備段階も含めた解析に要する時間の短縮や管理コストの低減を図り、災害の軽減に貢献するための地震火山観測研究計画におけるデータ解析研究の量的質的发展に資する。

### 背景

昨今の計算機やストレージ、ネットワークの性能向上に伴い、長期間の地震波形データを用いた解析を行うことが現実的となった。例えば、テンプレート波形を使って連続地震波形データから地震を検出し地震活動の推移を考察する研究 (Kato et al., 2012 など) は、その好例である。しかし、技術的に可能となった解析ではあるが、その研究基盤 (インフラ) の整備は遅れている。特に、過去の連続波形データの多くはベアドライブやテープ等に記録後、オフラインの状態で保管されている。このように保管されているデータを利用する際には、一時的にデータを展開する解析用ストレージ上で、データの読み出しを行う必要があり、解析に至るまでの準備時間が相当に必要となっている。そこで「長期間地震波形データ等解析システム」を構築し、長期間の地震波形データ等を容易に解析するための研究基盤として整備を進めることを計画した。

### ハードウェア構成

解析サーバー 4 台 (Intel Xeon® E5-2660v2 2 基、メモリ 128GB)  
大容量ストレージ 実効約 1PB (52TB の筐体が 24 基)  
データ受信サーバー 1 台 (Intel Xeon® E5-2660v2 2 基、メモリ 128GB)  
高速ストレージ 1 台 (48TB)

### 主なソフトウェア環境

CentOS 6.6  
WIN, win32tools, GMT, SAC, Seismic Unix, GNU Science Library, Intel Compiler  
その他 OS と同時にインストールされる標準的なツール類 (Perl 等)

### 設置場所

地震研究所テレメータ室

## 対象となる主なデータ（ストレージにコピーされる予定のデータ）

データ流通網に乗って地震研に届いたデータ（2002年4月1日～、テレメータデータ）

Hi-netは2007年まではビットシフトしている。

パケット欠落等で地震研に届いていないデータもある。

グループ内で合意のとれた臨時合同観測データ

1989年～2002年4月1日までのデータ（グループ内利用）もこちらに含む。

データ流通網経由で取得されたグループ内公開のオンライン観測点も含む。

メディアの劣化や管理者退職等に伴うデータの散逸を防ぐ意識も裏にはある。

非公開オンライン観測点データ

1989年～2002年4月1日までのデータ（非公開）もこちらに含む。

## 利用方法

データ流通協定の細目協定別紙にある大学の地震データ利用ガイドライン（研究発表する際には申請と許可の手続きを経る）に従って相互にデータ利用を行う。

※ガイドライン [http://eoc.eri.u-tokyo.ac.jp/eisei\\_system/riyou/guideline.htm](http://eoc.eri.u-tokyo.ac.jp/eisei_system/riyou/guideline.htm)

具体的な利用手続きは、現行の HARVEST システム（大学のデータ利用システム）と同等のユーザインタフェースを使って利用申し込みと関係機関への周知、利用の諾否、利用者への回答を行う。この際、Hi-net や F-net 等防災科研のデータの利用を希望する者には、防災科研のユーザ ID の入力を求め、当該 ID を防災科研に通知する。防災科研のユーザ ID は、防災科研がデータ利用許可手続きを進める際に利用される。

データ利用が許可されれば、アカウントを発行（eic の利用申請が未然の場合はそれも行う）して、サーバーの利用を開始する。アカウントの有効期限は、年度末までとする。

なお、グループ内利用や非公開データについては、各データの管理者から個別に利用許可を得てチャンネル情報を入手して利用していただく。

また、セキュリティ対策のコストを低減させるため、所外からは eic を経由して解析サーバーにログインする。eic の利用促進し解析の高度化を進めるという副次的効果も狙っている。

## 成果の公開とその報告

成果の公開に際しては、データ流通協定の細目協定の別表 4-1 のデータ提供機関に関するクレジット及び文部科学省による災害の軽減に貢献するための地震火山観測研究計画と地震研究所共同研究プログラムについて、謝辞等にもれなく記す。

また、データ提供機関に当該成果の報告を行う。ただし、多数の機関に対して個別に報告を行うのは煩雑であり、確実な報告を継続させるためにも、ワンストップサービスによる成果の共有等の仕組みが必要であると考えられる。

以上

(公社)日本地震学会 秋季大会  
2015年10月28日(水)A会場 午前 S21. 地殻活動ビッグデータ  
が拓く新世界[コンビーナ:中川茂樹・加藤愛太郎・汐見勝彦]

S21-09

## 長期間連続地震波形データ等解析システムの開発

# 中川茂樹・鶴岡弘(東大地震研)・  
加藤愛太郎(名古屋大学)・平田直  
(東大地震研)

2015/11/30

平成27年度第2回地震・火山噴火予知研  
究協議会資料

1

## 長期間地震波形データを 利用した研究

- 連続波形データを「なめる」研究
  - ambient noise
  - matched filter technique
- 過去のイベント波形を使った研究
  - デジタル化以降の全データ
- 計算情報環境の進展で可能になった
- 世界と戦うためには研究のスピードアップが必須

2015/11/30

平成27年度第2回地震・火山噴火予知研  
究協議会資料

2

## 現在の問題点

- 過去の連続データへのアクセスが困難、且つ、  
時間がかかる  
⇒研究進展に対する障害
  - オフラインメディアになっている(特にテープ)
  - 一時的に連続データを展開する場所が必要
- 各大学で独立して同じようなデータを所有
  - ある意味で冗長化(この点は長所)
  - 管理コストは×9になっている
- 臨時観測データの適切な管理?
  - テープが読めなくなる
- 予算、人員の縮減

2015/11/30

平成27年度第2回地震・火山噴火予知研  
究協議会資料

3

## 解析システム

- 利用頻度は低いが、大学がまとめればそれなりに利用されるであろう
- 新たな発想を思いついた時に即解析できる、という仕組みは研究者にとっても魅力的である
- 各大学で独立に同じ事をやるよりは、資源を集中した方が効率的だろう

**連続データの解析基盤を構築**

2015/11/30

平成27年度第2回地震・火山噴火予知研  
究協議会資料

4

## 本システムの対象となる主な 長期間地震波形データ

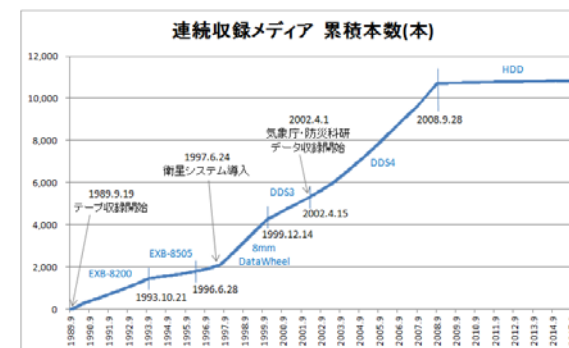
- 迅速な研究には、連続データがオンライン化されていて、すぐに利用できることが必要
- 主な対象データ
  - デジタル化以降の大学のデータ
  - 一元化以降の公開データ
  - 臨時観測のデータ
- 利用頻度は、それ程高くないと考えられる
- (問題点)膨大な資源(設備、金銭、時間)

2015/11/30

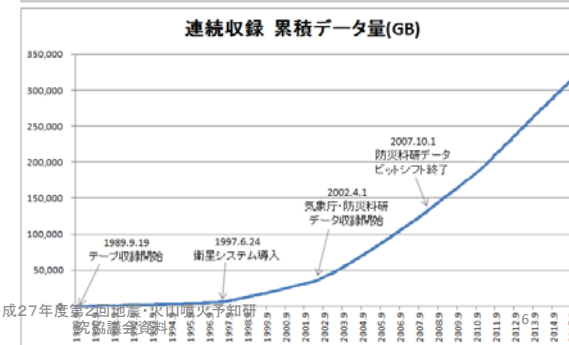
平成27年度第2回地震・火山噴火予知研究協議会資料

5

地震研究所で  
オフライン媒体に  
格納されている  
全データ容量：  
~320 TB



全てのデータが  
テープからHDDに  
複写



## 解析システムの実現方法と機能

- 商用クラウド利用
- オンプレミス
- 大容量、高速、安定
- ストレージと共有用のサーバ(冗長性と保守)
- データは55MB/分=2.4TB/月で単調増加
- データの適切なアクセス制御、セキュリティ
- 適切な維持管理、継続性
- 経費(初期費用、保守運営費用)

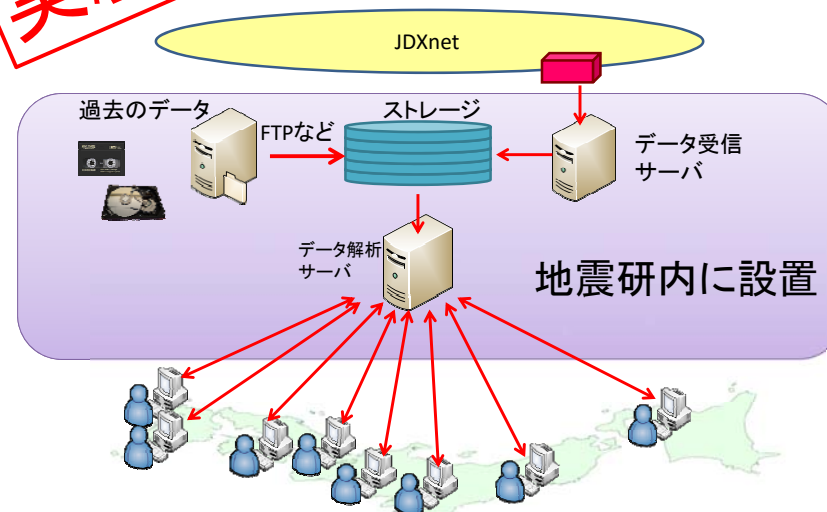
2015/11/30

平成27年度第2回地震・火山噴火予知研究協議会資料

7

**実験**

## オンプレミス案



2015/11/30

平成27年度第2回地震・火山噴火予知研究協議会資料

8

## ハードウェア構成

- ストレージ 52TB×24台
  - 過去から10年後程度までの連続波形データ
  - うち1台はユーザのホーム(解析)領域
  - 1989~2010のデータは年内にオンライン
- 解析サーバ 4台
  - Intel Xeon E5-2660 v2 (2.2GHz) 2基
  - メモリ128GB
  - ストレージとは10GBase-Tで接続
- 高速ストレージサーバ 約48TB
  - 約1年分の連続波形データを収録可能

2015/11/30

平成27年度第2回地震・火山噴火予知研究協議会資料

9

## ソフトウェア環境

- CentOS 6.6
- WINシステム
- win32tools
- GMT
- SAC
- Seismic Un\*x
- GNU Science Library
- Intel Compiler
- Python(ObsPy)やR等は検討中

2015/11/30

平成27年度第2回地震・火山噴火予知研究協議会資料

10

## 利用方法の案

- 利用者は各観測データ提供機関から個別に利用許可を得る。
- 大学の地震データ利用ガイドラインに従って、HARVESTシステムを使って、利用申し込みと関係機関への周知、利用の諾否、利用者への回答を行う。この際、防災科研のユーザIDも活用する。
- 謝辞にデータ提供機関をもれなく書く。

(技術面では)

- UNIXのグループを使ったデータアクセス制御
- チャネル表による利用制限

2015/11/30

平成27年度第2回地震・火山噴火予知研究協議会資料

11

## 利用申請(案)の例

\*\*\*注意事項\*\*\*  
E-MAILアドレスは、半角英数字で正確に記入してください。

氏名:	
役職名/身分:	
所属機関(研究機関名):	
郵便番号:	
連絡先住所:	
電話番号(伊勢):	
FAX:	
E-MAIL:	

学生の場合には、以下の事項にも記入してください。

学年:	
指導教員名:	
指導教員E-MAIL:	

利用希望のデータ  
必要事項はもれなく書いてください。フォーマットは自由です。

研究テーマ:	
研究内容:	
利用するデータの時間またはイベント名:	

利用する観測点名:  
 震源データ利用 (波形データを利用されない方はこちらを選択)  
 波形データ利用 (以下の観測機関リストから選択する)  
 北海道大学  札幌大学  東北大学  東京大学  名古屋大学  
 京都大学  高知大学  九州大学  鹿児島大学  
 防災科研Hi-net  防災科研F-net  防災科研のユーザID  
 気象庁  産総研  国土地理院  海洋研究開発機構  地震予知総合研究機構  
 青森県  東京都  静岡県  神奈川県  温泉地学研究所

利用するデータの成分名:  
 全成分  UD成分  NS成分  EW成分  水平成分

確認コード:

申請する 取消

2015/11/30

平成27年度第2回地震・火山噴火予知研究協議会資料

12

## まとめ

- 商用クラウドはまだ高い！（特にストレージ）
- 地震研に**解析基盤**を**実験的**に設置  
解析実験への参加者を募集
- 利用について
  - 大学の地震データ利用ガイドラインに従う
  - データ利用申請等はHARVESTシステムを活用
  - データ利用の成果を共有する仕組み
- 大規模並列計算機等との連携？
- 世の中のビッグデータ解析ツールの利用？