

KEK 電子陽電子入射器ビーム同期データ収集、解析システム開発 DEVELOPMENT OF SYSTEMS FOR ARCHIVING AND ANALYZING BEAM SYNCHRONOUS DATA FOR KEK e+/e- LINAC

宮原房史^{#,A)}, 早乙女秀樹^{B)}, 大房拓也^{B)}, 工藤拓弥^{C)},
榎本嘉範^{A)}, 佐藤政則^{A)}, 佐武いつか^{A)}, 王迪^{A)}
Fusashi Miyahara^{#,A)}, Hideki Saotome^{B)}, Takuya Ofusa^{B)}, Takuya Kudou^{C)},
Yoshinori Enomoto^{A)}, Masanori Satoh^{A)}, Itsuka Satake^{A)}, Di Wang^{A)}

^{A)} High energy Accelerator Research Organization (KEK)

^{B)} Kanto Information Service (KIS)

^{C)} Mitsubishi Electric System & Service Co.,Ltd

Abstract

The injector linac is operated in multiple beam modes that can be switched every 20 ms for beam injection to different circular accelerators (SuperKEKB HER/LER, PF, PF-AR). Since RF and pulsed magnet system are controlled on a pulse-by-pulse basis, the data for these devices is updated at 50 Hz. The beam position monitor and all of these instrument data are archived, allowing synchronized data extraction. The data archiving system and a synchronous data extraction system will be reported.

1. はじめに

KEK 電子陽電子入射器(入射器)は最大 50 Hz で SuperKEKB HER/LER, PF, PF-AR へ電荷量、エネルギーの異なる電子/陽電子ビームを供給している。入射器では入射する円形加速器に対応したビームモードを用意し、モードごとのビーム制御を行っている。ビームモードはイベントシステムを用いて光信号でショット番号とともに入射器の各機器に配信されている。イベントシステムを用いた測定機器は測定データと共にショット ID を記録する。現在、50 Hz でデータを測定し EPICS waveform で配信しているシステムはビーム位置モニター(BPM)、RF モニター、パルス電磁石電源システムである。BPM、RF モニター、パルス電磁石電源のデータは合計で約 2 千点あり、50 Hz でデータを保存するとデータ量は膨大となるが、ビーム軌道のフィードバックや異常な軌道、比較的早い位置のドリフト等が発生した場合の原因調査にはビームと同期したこれらのデータのアーカイビングが欠かせない。またアーカイビングしたデータから任意のデータを高速に抽出するシステムも要求される。入射器のビーム同期データアーカイビングシステムとその実用例、高速データ抽出システムの開発を報告する。

2. ビーム同期データアーカイビング

入射器では様々な機器を EPICS IOC [1]で制御しており、全てのデータ(測定値、設定値等)は Archiver Appliance に記録されている[2]。現在、約 14.5 万点の PV が登録されている。BPM などパルスごとに値が変化するデータも全データの記録は可能であるが、50 Hz で全データを記録すると、ファイルサイズが膨大となりストレージを圧迫することと、通常の更新が遅い PV のデータ抽出の速度にも影響してしまう。そこで Archiver

Appliance は BPM など更新の早いデータに関しては最短で 1 秒ごとにサンプリングしたデータを記録している。また Archiver Appliance はデータの取りこぼしの確認や同時に大量の PV の値を取得したり、ビームと同期したデータを抽出することは容易ではない。

そこで入射器では 50 Hz でデータの取得、保存が必要なデータに関しては Archiver Appliance とは別に EPICS waveform を ASCII 形式で保存している。50 Hz でデータを取得しているシステムの概略を Fig. 1 に示す。BPM IOC はビーム位置モニター読み出しシステム[3]の制御とデータ収集、配信を行っており、計 22 台の BPM IOC で 107 台のストリップライン型 BPM を管理している。BPM は接続された各 BPM の位置(x, y)と電荷量を表す個々の EPICS PV のほかに、これらの BPM の位置、電荷量と ADC 値とショット ID をまとめた waveform (配列)

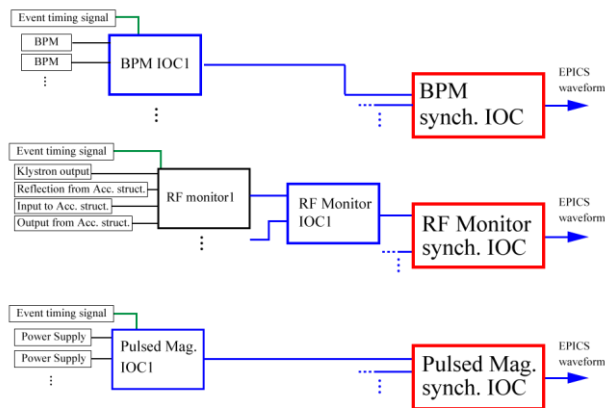


Figure 1: Schematic of the synchronous data generation system.

をビームモードごとに更新している。ショット ID はビーム 1 パルスごとにイベントシステムから送信される 32 bit 整

[#] fusashi.miyahara@kek.jp

数で、ビームモード等と共にイベントシステムで配信される。個々の IOC で更新される waveform は非同期であるため、ビームの軌道解析等の際は、ショット ID が同じデータを解析する必要がある。そこで、各 BPM IOC waveform についてショット ID が同じデータをそろえ、まとめて一つの waveform にするために BPM の同期 IOC (BPM sync. IOC) 開発した。同期 IOC は全 BPM の情報を記録した waveform の更新と同時にビームモードごとに waveform 形式のまま ASCII データを保存している。ASCII データは 5000 イベントごとに更新されるが、1 ファイルあたり 46 MB でデータ量が膨大となることから、一定時間経過後にビーム無しのデータを除いて gzip で圧縮している。BPM の同期 IOC の waveform はビーム軌道のフィードバックにも利用されている。RF モニター[4]はクライストロンの出力、SLED/加速管への入力と反射の RF を測定しており、BPM と同様に個々の IOC から取得した全 64 台の RF モニターのデータを RF モニター用同期 IOC で一つの waveform にまとめている。入射器はビームモードに応じてエネルギーが異なるため、パルス電磁石を用いてビームモードごとの軌道制御を行っている。パルス電磁石の制御はパルス電磁石 IOC で行っており[5]、パルス偏向電磁石電源 4 台、四極電磁石電源 34 台、ステアリング電磁石電源 63 台を計 16 台のパルス電磁石 IOC で管理している。パルス電磁石の設定電流値、測定値はパルス電磁石同期 IOC で一つに waveform にまとめられる。RF モニター、パルス電磁石の同期データも BPM 同様に 5000 ショット単位で ASCII で保存され、一定時間経過後に gzip で圧縮されるが、ビームの有無と関係なく全ショットを保存している。またデータファイルと測定時刻の対応をとるため、各同期 IOC はビームモードごとにファイル名、最初のショット ID とタイムスタンプ(日時、UNIX time)を記したインデックスファイルを更新する。

3. ビーム同期データの活用

3.1 BPM, RF モニター, パルス電磁石データ同期

BPM, RF モニター、パルス電磁石の同期データはそれぞれ別に保存されているため、これらの相関を調べるためには、ショット ID を比較し、そろえる必要がある。任意のモニターに関して指定したビームモード、開始/終了時刻の同期データを抽出する ruby ベースのプログラムを開発した。このプログラムは指定時刻とインデックスファイルから解析対象のファイルを絞り込み、BPM, RF モニター、パルス電磁石のデータを一度バッファに代入し、ショット ID が同じデータを出力する。実行速度はほぼデータの読み出し速度で決まり、25 Hz 更新のデータから 5 分間の同期データを抽出するのに要する時間は 1 分半程度で、短時間のビーム変動を調査する目的には十分である。

3.2 同期データの活用

全てのビームショットが記録されていることから数秒から数分程度の短時間のビーム軌道の変動や変動幅の推移などビームの安定性の評価や他の機器の突発的な異常により引き起こされたビーム軌道の変化なども確認で

きるため BPM 単体のデータだけでも有用である。

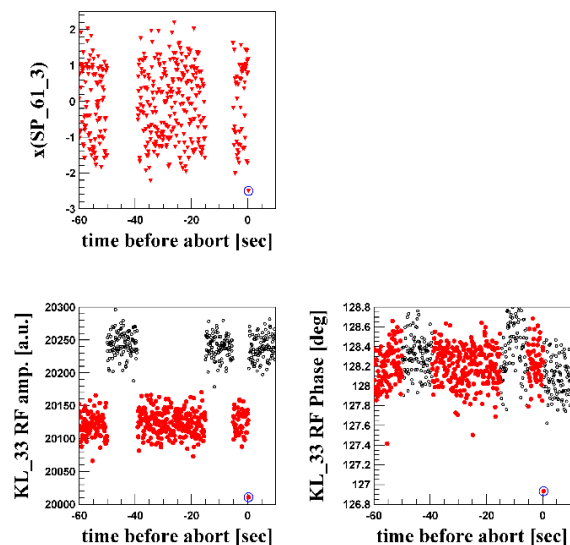


Figure 2: Changes in beam position and an RF just before LER abort at e+ beam injection. The upper shows the change in horizontal position at BPM SP_61_3 with large dispersion. Lower left and right figure show RF amplitudes and phases for KL_33. Red circle and black open circle correspond to beam on and off, respectively.

RF モニターやパルス電磁石のデータとビーム軌道の同期データは SuperKEKB HER/LER 入射に同期したビームアボートの原因特定や非同期測定では判別可能なパルス機器のわずかな変動とビーム軌道の相関の確認などに利用されている。入射器 3 セクターの加速ユニットへの入力 RF の異常が原因でビーム軌道が変化し、LER アボートに至った事象の例を Fig. 2 に示す。アボートに対応した時刻 (t=0) で RF モニターで測定した振幅がアボート前と比べて 0.6 %程度下がっており、位相も変化していることがわかる。このため、エネルギーが変化しエネルギー圧縮システムのシケイン中央に置かれた BPM で軌道も低エネルギー側にシフトしていることが分かる。

4. 同期データ抽出の高速化

4.1 抽出プログラムの改良

同期データ解析で使用頻度が高い用途は入射器内での異常なビーム軌道の原因調査やビームとパルス機器の相関の解析である。この場合は解析するデータの時間区間は 1 時間以内であることが多く、BPM, RF モニター、パルス電磁石の同期データからその都度同期データを抽出する方法でもそれほど問題にはならない。しかし、不特定の期間から任意の条件の事象を探したり、長期間のデータを解析する場合はデータ抽出の時間が問題となる。そこで、あらかじめ BPM, RF モニター、パルス電磁石の同期データを生成し、高速でデータ抽出可能なバイナリフォーマットに変換することを検討した。

4.2 HDF5 フォーマット

同期データ用のバイナリフォーマットとして HDF5 フォーマット[6]の評価を行った。HDF5 フォーマットはディレクトリとファイルで構成される様なフォーマットで、グループ(ディレクトリに相当)と呼ばれる構造とデータセットで構成される。グループには他のグループやデータセットが格納可能で、データセットはデータ列(多次元も可能)に対応する。大きなファイルでも高速にデータ抽出が可能(インデックス管理が必要ない)で、gzip などを用いた圧縮も利用可能である。また、C、C++、Python、Fortran、MATLAB などサポートしている言語、処理系が多い。

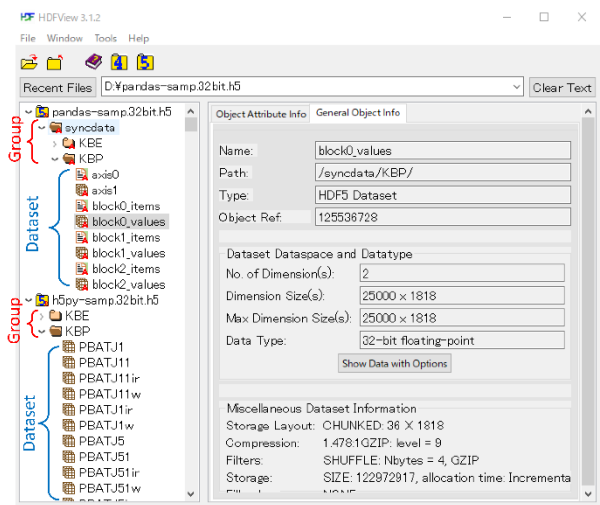


Figure 3: Example of HDF5 structure.

HER/LER 入射用のビームモード KBE/KBP で、それぞれ 25000 ショット、計 50000 ショットの ASCII データを python の Pandas (Python でデータ解析を支援するライブラリ) と h5py の異なる2つのライブラリを用いて HDF5 にしたときのデータ構造を Fig. 3 に示す。図中に syncdata, KBE, KBP と表されているものが、グループ、block0_values や PBATJ1 で表されている変数がデータセットである。Pandas で生成したファイルのデータセット block0_values は 25000 行 x 1818 列の単精度浮動小数点型の多次元のデータセット、block0_items には各列のインデックス(PBATJ1 等)が記載されており、Pandas のデータフレームとして直接扱うことが可能である。一方、h5py で生成したファイルではビームモードを表すグループの下に各変数が 1 次元のデータセットとして保存されている。各ライブラリで圧縮有り、無しで HDF5 化したファイル、ASCII データを圧縮したファイルのファイルサイズを Table 1 にまとめる。どちらのライブラリも gzip による圧縮でファイルサイズが 1/4 程度になっており、テキストデータをそのまま圧縮するよりも効率的である(ただし、テキストデータの中身で物理的に意味の無い桁を丸めた場合は 13%程度まで圧縮が可能である)。Python を用いた解析では Pandas で HDF5 化したファイルの読み出しのほうが h5py を用いた場合よりも約 6 倍高速であったことと、データ解析の容易さから Pandas を用いた HDF5 フォーマット同期データ生成システムを開発した。

Table 1: Data Format and Data Size for 50000 Events

Format	File size [kB]	Ratio
Raw Data (ASCII)	952040	100.00
HDF5 (h5py, 32bit, w/o comp.)	389272	40.89
HDF5 (h5py, 32bit, gzip Lv=7)	217608	22.86
HDF5 (pandas, w/o comp.)	1305916	137.17
HDF5 (pandas, 32bit, gzip Lv=9)	243872	25.62
HDF5 (pandas, 32bit, gzip Lv=7)	244120	25.64
gzip Lv=9	409596	43.48
gzip Lv=7 (optimized format)	123636	12.99

4.3 データ抽出速度

データ抽出速度を BPM, RF モニター、パルスマグネットの各同期データ(圧縮済み)から生成する従来の方法と、同期済みのデータ HDF5 化したファイル、ASCII 形式を圧縮したファイルと比較した。使用したプログラムとデータのフォーマットごとに 25 Hz 運転時のデータから 10 分間分のデータを読み出すのに要した時間を Table 2 にまとめた。Pandas を使う場合は、あらかじめ 1 時間分の同期データを gzip で圧縮したものと、HDF5 化したデータを利用した。あらかじめ同期データを生成しておけばこれまでの一から同期データを生成する方法の約 1/5 の時間でデータ抽出が可能である。HDF5 に変換したデータを使えば、これよりもさらに高速化が可能で、これまでの方法の 1/25 の時間でデータ抽出が可能である。また HDF5 の同期データを用いれば、指定した時間の区間の抽出だけでなく、様々な条件(例えばある BPM でのビーム位置の範囲を指定)を重ね合わせたデータ抽出も容易に行えるため、データ解析の時間を大幅に削減できる。

Table 2: Data Format and Time Required for 10 Minutes of Synchronous Data Extraction

Program and format	Data size [MB]	Time [s]
Ruby: Compressed raw data (gzip) and generate synchronous data	124	~240
Pandas: Compressed synchronous data (gzip) for KBE 25 Hz, 1hour	1286	46
Pandas: HDF5 for KBE 25 Hz + KBP 12.5 Hz, 1hour	608	9

4.4 HDF5 データ生成の課題

HDF5 フォーマットは大きなファイルサイズでも高速にデータ処理を行うことが可能であるが、1 日単位など大量のデータをひとつにまとめるとファイルサイズが大きくなりすぎてしまい、通常の方法でデータを読み出すと、読み出し時に必要なメモリがシステムのメモリサイズを越えてしまう。例えば、Table 2 に示した 1 時間分の HDF5 ファイル(KBE 25Hz, KBP 12.5 Hz で 608 MB)を Pandas のデータフレームに代入する場合、1.7 GB のメモリを使用

した。したがって、入射器では 1 時間程度で一つのファイルにする必要があり、長時間のデータを解析する場合は解析プログラムで対応する必要がある。また、一度 HDF5 のファイルを生成すると同じ構造でデータの追記が出来ないため、データの追加等の場合は一から同期データを作り直す必要がある。

入射器では 2022 年秋から HDF5 化した同期データの生成を行い、任意のバイナリ(ROOT 等)に変換するプログラムや Archiver Appliance のデータとも関係可能な WEB ベースの誰もが容易に解析を行えるツールの開発を行い、試験的な運用を行う予定である。

参考文献

- [1] Experimental Physics and Industrial Control System (EPICS); <https://epics.anl.gov/>
- [2] I. Satake *et al.*, “KEK 電子・陽電子入射器における Archiver Appliance の運用状況”, Proceedings of the 17th Annual Meeting of Particle Accelerator Society of Japan, Online, Sep. 2-4, 2020, pp. 735-738.
- [3] R. Ichimiya *et al.*, “SuperKEKB に向けた高位置分解能 LinacBPM 読出しシステムの開発”, Proceedings of the 11th Annual Meeting of Particle Accelerator Society of Japan, Aomori, Japan, Aug. 9-11, 2014, pp. 427-431.
- [4] H. Katagiri *et al.* “RF Monitor System for SuperKEKB Injector Linac”, Proceedings of 9th International Particle Accelerator Conference, Vancouver, Canada, April 29 – May 4, 2018, pp. 2128-2130.
- [5] Y. Enomoto *et al.* “A New Pulsed Magnet Control System in the KEK Electron Positron LINAC”, Proceedings of 9th International Particle Accelerator Conference, Vancouver, Canada, April 29 – May 4, 2018, pp. 2121-2123.
- [6] HDF5; <https://www.hdfgroup.org/solutions/hdf5/>