

J-PARC Main Ring 向け EPICS I/O コントローラとしての小型ファンレスサーバ の運用 10 年

TEN YEARS OF OPERATION OF FANLESS MICRO SERVERS AS EPICS I/O CONTROLLERS FOR THE J-PARC MAIN RING

山田 秀衛^{*,A)}
Shuei Yamada^{*,A)}

A) High Energy Accelerator Research Organization (KEK) / J-PARC Center

Abstract

The J-PARC Main Ring (MR) uses fanless microservers as front-end computers (EPICS I/O controllers; IOCs). In 2007, during the construction phase of the MR, VME single-board computers (VME SBCs) were introduced as IOCs. It was later found that most of the control targets were network devices, and VME buses were not always necessary. In 2014, fanless micro servers were introduced as IOCs on a trial basis. Full-scale deployment began in 2015. By 2023, the transition from VME SBCs to small fanless servers was completed. This article reviews the operational experience with small fanless servers and discusses future prospects.

1. はじめに

J-PARC (Japan Proton Accelerator Research Complex) の Main Ring (MR) 加速器の制御システムは、フレームワークに EPICS [1] を、計算機の OS に Linux を用いて 2007 年に構築され [2]、2008 年のビーム運転開始以来運用されている [3]。EPICS ベースの制御システムのフロントエンド計算機は I/O Controller (IOC) と呼ばれる。

J-PARC MR では、IOC の要件としてふたつ以上のネットワークインタフェース (NIC) を有することを要求している。制御対象の大半は PLC やオシロスコープといったネットワーク接続可能な機器であるが、ネットワーク機器という観点では、必ずしも一般的な PC のように高性能というわけではない。ネットワークケーブルの接続ミスや機器の故障といった原因でブロードキャストストームが発生した場合に制御対象機器がハングアップすることがある。上位制御系のネットワークと制御対象機器のネットワークを分離することで、上位制御系ネットワークの障害の影響が制御対象機器に及ばないようにし耐障害性を高めている。

2. PINON サバ太郎[®] TYPE-P

PiNON サバ太郎[®] Type-P (サバ太郎) はふたつの NIC を有するファンレスの小型サーバである (Fig. 1, [4])。MR 加速器では 2014 年に試験的にサバ太郎を IOC として導入した。2015 年からは本格的にサバ太郎を導入し [5]、建設期の 2007 年に導入した VME SBC [6] からの移行を開始した。2023 年には、VME バスを用いる 37 台の IOC を除いてサバ太郎への移行を完了した (Tab. 1)。2024 年の時点で MR 加速器の IOC として 86 台のサバ太郎を運用中である。VME SBC とサバ太郎の仕様を Tab. 2 に示す。

サバ太郎は小型 (W80.6 mm × D110.6 mm × H34.4 mm) であること、価格が VME SBC の 1/5 程度であることから、VME SBC に比べると導入が容易で



Figure 1: Appearance of PiNON Sabataro[®].

ある。VME SBC を設置できる台数は VME クレータのスロット数で制限されることから、従来は 1 台の IOC に複数の機能を持たせる運用を行っていた。サバ太郎を導入することで機能毎に IOC を分離することが可能になった。また、同じ VME クレータに実装されている複数の VME SBC は電源を共有しているため、特定の IOC の電源を切るためにクレータ全体の電源を切る必要があった。サバ太郎に移行することで、目的とする IOC の電源だけを切ることが可能になった。VME SBC は少しでも導入コストを下げるため、制御対象となる機器の信号点数やデータサイズに応じて CPU やメモリ搭載量の仕様の異なる 4 種類のハードウェアを使い分けていたが、サバ太郎では機種をひとつに絞っている。このように、サバ太郎を導入することで保守性が向上し、管理上のコストも低減している。

IOC にハードウェア障害が発生した場合は、予備機と入れ替えることで速やかに加速器制御システムを復旧する体勢を整えている。が、意外とサバ太郎は故障しなかった。これまでに購入した約 100 台のサバ太郎のうち故障したのは 4 台で、いずれも電源が入らなくなる症状

* shuei@post.kek.jp

Table 1: Migration from VME SBC to Sabataro

| | 2007 | 2008 | 2011 | 2013 | 2014 | 2015 | 2016 | 2021 | 2023 | 2024 |
|----------|------|------|------|------|-------|------|------|------|------|------|
| VME SBC | ~80 | ~80 | ~90 | ~80 | ~80 | ~90 | ~90 | 55 | 37 | 37 |
| Sabataro | | | | | a few | 11 | ~30 | 50 | 80 | 86 |

Table 2: Typical Specifications of IOCs

| Model | CPU | memory | disk |
|--------------------------|---------------------------------|-------------|----------|
| GE V7865 (VME SBC) | CoreDuo T2500 2.0 GHz (2 cores) | 3 GB | diskless |
| GE V7807 (VME SBC) | Pentium M 1.8 GHz (1 core) | 1 GB-1.5 GB | diskless |
| Sanritz SVA041 (VME SBC) | Celeron-M ULV 600 MHz (1 core) | 512 MB | diskless |
| GE V7700 (VME SBC) | Celeron-M ULV 400 MHz (1 core) | 512 MB | diskless |
| Sabataro | Celeron J1900 (4 cores) | 8 GB | SSD |

であった。2013年-2014年に32台導入した制御端末のうち19台が故障して起動しなくなったことに比べると、サバ太郎の故障率は非常に低いと言える。

3. サバ太郎の使用状況

MR 加速器の運転に用いている各 IOC は、自分自身の CPU 使用率やメモリ使用量、ネットワークトラフィックといった使用状況のデータを EPICS レコード化しており、上位制御系でこれらを監視・アーカイブしている。J-PARC MR で運用している 86 台のサバ太郎のうち、84 台の使用状況の 1 時間平均をヒストグラムにして纏めた。のこり 2 台については手違いでアーカイブされていなかったため、ここでは省略した。

個々のサバ太郎の運用を開始した時期は機体によって 2014 年-2024 年と広がりがある。この運用期間の違いを吸収するため、まず機体毎に使用状況のヒストグラムを作り、面積で正規化した。次に、全 84 台のヒストグラムを積み上げて、さらに面積で正規化することで使用状況の分布を作成した。

3.1 CPU 使用率と負荷

サバ太郎の CPU 使用率の分布を Fig. 2 に示す。4 コアある CPU 全てを使いきると CPU 使用率が 100% になる。CPU 使用率が 25% 未満の割合は 99.99% であった。たかだか 1 コアしか使っておらず、CPU の能力にはまだ余裕がある状況である。

サバ太郎には汎用 OS (Alma Linux 9) を用いているので、IOC コアの他にシステムの運用に必要な最低限のサービス (ssh, ntp, syslog など) や定期的に行われるジョブ (logrotate など) も動作している。また、IOC コアはマルチスレッドで動作するので、CPU の能力にはまだ余裕があるといえども、2 コアでは心許ない。

サバ太郎の 1 分間のロードアベレージの分布を Fig. 3 に示す。ロードアベレージは実行待ち状態にあるプロセスの平均数で、コンピュータシステムにおける負荷状況の指標となる値である。一般的には過去 1 分間・5 分間・15 分間の平均値を使うが、ここでは過去 1 分間の値だけを議論する。ロードアベレージが 1 未満の割合は 99.93% であった。

ロードアベレージが 1 を超えた部分は、主電磁石電源

の通電パターンを作成しファイルサーバに書き込む IOC によるものであった。当該 IOC の運用方法を調査したところ、主電磁石電源の調整と MR 加速器の調整の際に、事前に複数の通電パターンを事前に作成・保存していることが判明した。ロードアベレージが大きく CPU 使用率が低い事から、I/O 待ちによる負荷であると考えられる。主電磁石電源の通電パターンの作成は、MR の加速器サイクル (速い取り出し運転では 1.36 s、遅い取り出し運転では 4.24 s) 毎にデータを読み出すようなリアルタイム性は求められないため、負荷が高くても加速器の運転には差支えない。

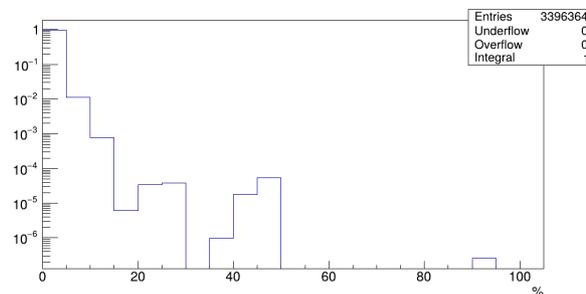


Figure 2: Distribution of CPU usage of 84 Sabataros.

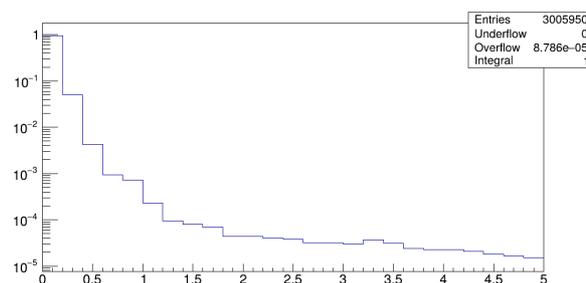


Figure 3: Distribution of load average over 1 minute.

3.2 メモリ使用量

Figure 4 に示すのはサバ太郎のバッファ・キャッシュに使われている分を除いたメモリの使用量の分布である。メモリ使用量が 4 GB 未満の割合は 99.85% であった。4 GB を越えたのは、

- 主電磁石電源の通電パターンを作成・保存する IOC
- MR 加速器全周にある 96 台の BPM データを集約して COD を作る IOC

の 2 台であった。これら以外の IOC に関しては、これまでのところメモリ実装量は 4 GB で充分だった可能性がある。しかしながら、MR 加速器の高度化に伴って IOC が扱うデータサイズが増加することが予想されている。

Linux を用いた計算機では未使用のメモリがあればバッファあるいはキャッシュとして利用される。バッファ・キャッシュを含むメモリ使用量の分布を Fig. 5 に示す。サバ太郎に実装されている 8 GB までメモリを有効に利用していることがわかる。

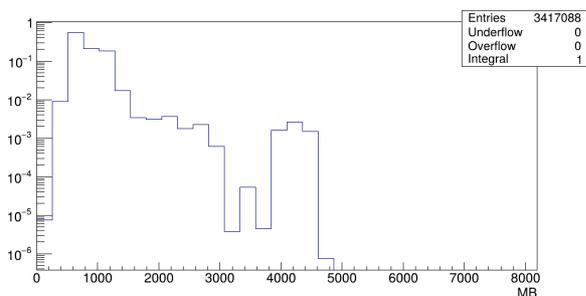


Figure 4: Distribution of memory usage excluding buffers and caches.

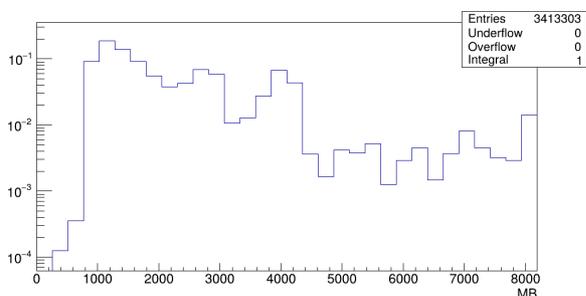


Figure 5: Distribution of memory usage including buffers and caches.

3.3 ネットワークトラフィック

IOC の役割は、制御対象機器ごとに異なる通信プロトコルを EPICS の用いる CA プロトコルに変換することである。通信量の多い制御対象機器はビームモニタの波形データを読み出すためのオシロスコープ等の測定器である。これらを制御する IOC は、MR の加速器サイクル毎にビームの入射から出射まで測定した波形データを機器から読み出し、EPICS レコードに変換して制御

LAN に送信している。これに比べると、制御 LAN 側から IOC を経由して制御対象機器へ向かうトラフィックは無視できるほど少ないので、ここでは割愛する。

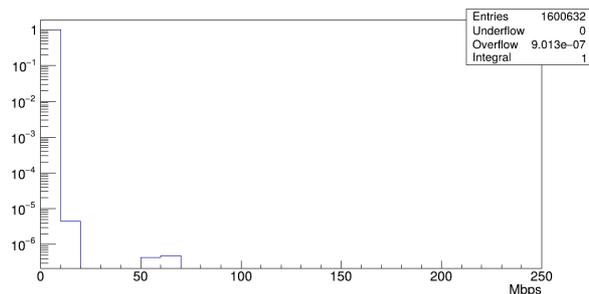


Figure 6: Distribution of inbound network traffic on the device-side port.

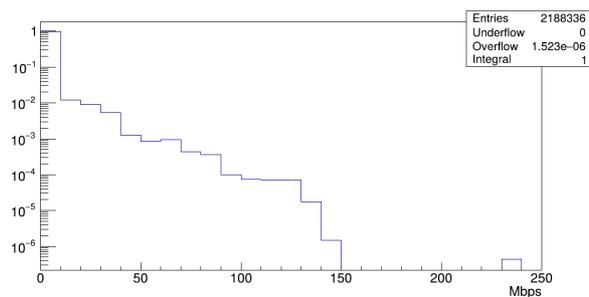


Figure 7: Distribution of outbound network traffic on the control network-side port.

Figure 6 に制御対象機器から IOC へのネットワークトラフィックの分布を示す。10 Mbps 未満のトラフィックの割合が 99.999% を占めている。

一方、Fig. 7 は IOC から上位制御系へのネットワークトラフィックの分布を示したものである。IOC はオシロスコープ等の測定器から読み込んだデータを EPICS レコードに変換し、上位制御系のネットワークに送信している。従って IOC の制御 LAN 側ポートの送信トラフィックは機器側ポートの受信トラフィックと概ね同じ傾向を示すことが期待されるが、実際には期待に反して 10 Mbps を越えている。調査したところ、電磁石電源の通電パターンを作成しファイルサーバに書き込む IOC によるものであった。

4. まとめと今後の展望

IOC としてサバ太郎を選定したのは、性能の面でも故障率の面でも概ね良い選択であった。サバ太郎の CPU 使用率・ロードアベレージ・メモリ使用量・ネットワークトラフィックには、いずれもまだ余裕があることが分かった。ビームモニタの波形データを読み出す IOC では、加速器サイクル (速い取り出し運転では 1.36 s、遅い取り出し運転では 4.24 s) に同期して CPU 使用率・ネットワークトラフィックが変化しているが、本論文では 1 時間平均を用いているため、時間的に均されてしまっ

いる。加速器の1サイクルの間にこれらがどのように変化しているかを明らかにしたいが、1秒を下回る短い時間間隔での調査はシステムの負荷に影響を与えてしまう可能性があって困難なため、今後の課題としたい。

サバ太郎 Type-P はまだ入手可能であるが、次世代 IOC の検討を始めている。MR 加速器は更なる大強度化・高度化を目指す。次期 IOC は現行機と同程度の性能で充分であろう。IOC の能力が足りないのであれば、IOC を増設することで負荷を分散すべきであると考え。しかしながら、より新しいバージョンの・あるいは将来リリースされるバージョンの OS を走らせるために、CPU の世代は新しくしたいところである。また、制御端末ではファンの故障が目だったことから、次世代機もファンレスが望ましい。

参考文献

- [1] EPICS - Experimental Physics and Industrial Control System, <https://epics-controls.org>
- [2] S. Yamada *et al.*, “15 Years of the J-PARC Main Ring control system operation and its future plan”, Proceedings of ICALEPCS 2023, Cape Town, South Africa, pp. 639–644. doi:10.18429/JACoW-ICALEPCS2023-TUPDP049
- [3] T. Koseki *et al.*, “Beam commissioning and operation of the J-PARC main ring synchrotron”, Prog. Theor. Exp. Phys (2012) 02B004. doi:10.1093/ptep/pts071
- [4] Type-P | 製品 | 株式会社ピノー, https://www.pinon-pc.co.jp/hp/?page_id=102
- [5] S. Yamada *et al.*, “Deployment of a tiny fan-less server as IOC in J-PARC Main Ring”, Proceedings of the 13th Annual Meeting of Particle Accelerator Society of Japan, PASJ 2016, pp. 634–636, MOP092.
- [6] N. Kamikubota *et al.*, “J-PARC Control toward Future Reliable Operation”, Proceedings of ICALEPCS 2011, Grenoble, France, pp. 378–381, MOPMS026.