

J-PARC Main Ring の制御計算機の刷新 RENOVATION OF CONTROL COMPUTERS FOR J-PARC MAIN RING

山田秀衛 ^{*A)}、高橋大輔 ^{B)}、吉田奨 ^{B)}

Shuei YAMADA ^{*A)}, Daisuke TAKAHASHI ^{B)}, Susumu YOSHIDA ^{B)}

^{A)}High Energy Accelerator Research Organization (KEK) / SOKENDAI / J-PARC Center

^{B)}Kanto Information Service (KIS)

Abstract

Console system for J-PARC Main Ring (MR) was designed in 2007 and had been used for accelerator commissioning and operation since then. It was composed of 20 diskless thin clients and 10 terminal servers. Both of them are PC-based computers running 32-bit Scientific Linux (SL) as their operating system. In 2013, upgrade to a new console system was triggered by update from SL4 to SL6. Based on use experiences of those thin clients, the new system consist of ordinary fat clients and supplemental blade servers for high-load applications. 64-bit SL6 is employed to fully exploit hardware performance. Evaluation of fat clients is carried successfully out during commissioning of MR. Presently replacement of thin clients is almost finished and 21 fat clients have been installed. Migration scenario and prospects are discussed.

1. はじめに

近年の計算機の性能向上と低価格化には目覚ましいものがある。掌に乗るような小型の PC の性能が 5 年前のサーバ機を凌駕するようになった。

J-PARC Main Ring (MR) の加速器制御システムは、制御システムのフレームワークと OS に EPICS^[1] と Scientific Linux^[2] 4 (SL4) を用いて 2007 年に構築され^[3]、2008 年 5 月に MR のビーム運転を開始^[4]して以来運用されてきた。制御システムを構築したときに EPICS の GUI アプリケーションを開発するのに標準的に使われていた EDM や MEDM といったツールは、X Window System を必要としている。そこで、端末用計算機には Thin Client を GUI の描画部分とマウス・キーボードからの入力部分に特化した X 端末として使用し、アプリケーションは Blade 計算機で実行することにして負荷分散を図った。Thin Client は管理の手間とコストを抑えるべく、SL4 をネットワークブートするディスクレス機とした。Blade 計算機には IBM Blade Center HS20 および HS21 が採用され、端末用の Thin Client には HP Compaq t5720 および t5730 が採用されていた。

本論文では、今後 5 年間の展望を見据えた新しい環境の構築について述べる。

2. 制御計算機の刷新

6 年間のビーム運転の間に、MR 加速器は着実に性能を向上させてきた。それとともに加速器の制御に用いられるアプリケーションも高度になり、CPU パワーを補うべく Blade 計算機を増強してきた^[5]。しかし、GUI の表示に Thin Client を使用し、CPU には Blade 計算機を用いる運用形態は 2007 年以来不変であった。

2.1 端末用計算機

2012 年に SL の配布元による SL4 のサポートが終了したことを契機に、今後 5 年間の運用を見据えて次期端末用計算機を選定した。次期端末の候補として、SSD に 64 ビット版 SL6 をインストールした Intel NUC DC53427HYE を 1 台導入し、2013 年 12 月から 2014 年



Figure 1: Photograph of HP Compaq t5730 (left), Intel NUC (right), and a D Battery for Size Comparison (center). T5720 is not shown but its appearance is almost same as t5730.

1 月にかけて MR のコミッショニングで実際に用いて評価した。これまでのビーム運転での運用経験から、選定には以下のような点を考慮した。

- 十分な CPU パワー、メモリ容量があること。
制御計算機で Web ブラウザや PDF ビューア、オフィススイートを使いたいという需要が増え、Blade 計算機に負荷が集中するようになった。また、EDM や MEDM に変わって主流になりつつある Control System Studio (CSS)^[6] を用いたアプリケーションの導入が今後進み、一層 CPU パワーとメモリが必要になることが予測された。そこで、端末上でアプリケーションを実行する可能性を考慮した。
- SL6 が動作すること。
Blade 計算機や、制御対象となる機器のフロントエンド計算機である EPICS I/O Controller (IOC) の OS は SL4 から SL6 へ移行している^[7]。端末用計算機でもこれらと同じ SL6 を用いることで、円滑な移行と管理の簡略化を図った。

* shuei@post.kek.jp

Table 1: Specifications of Console Computers

	Intel NUC DC53427HYE	HP Compaq t5730	HP Compaq t5720
CPU	Core i5-3427U	Sempron 2100+	Geode NX 1500
CPU Frequency	1.8 - 2.8 GHz	1 GHz	1 GHz
Numer of Cores	2 (4threads)	1	1
Amount of Memory	8 GB	512 MB	512 MB
GPU	Intel HD Graphics 4000	Matrox EpicA TC2 / TC4	Matrox EpicA TC2 / TC4
Maximum Number of Monitors	3	2 or 4	2 or 4
Boot Media	SSD	Network (PXE)	Network (PXE)

- OS が GPU をサポートしていること。
既存の端末である t5720 / t5730 の GPU (Matrox EpicA TC2 / TC4) のデバイスドライバは SL4 には含まれていなかった。メーカーが提供するデバイスドライバを使用していたが、2009 年で更新が終了しており SL6 では動作しないことが判明した。次期端末は今後数年にわたって加速器の運転に用いられることから、搭載している GPU のデバイスドライバが OS で標準的にサポートされている機種を選定した。
- 端末のネットワークブート / ディスクレスには拘らない。
SL6 の標準的な手法でネットワークブートする場合、ブートイメージの 2 倍程度の容量のメモリを消費する [8]。一方、SL6 で NFS を介して root ファイルシステムをマウントした場合、端末に負荷がかかると動作が極端に遅くなり現象がみられた。また、複数の端末で NFS root を用いた場合、ネットワークに負担がかかることが予測された。
- 省スペースで複数のモニタへの出力可能であること。
J-PARC 中央制御室の限られたスペースを有効に活用するため、各 Thin Client を 2 台ないし 4 台のマルチモニタ構成で運用していた。少なくとも 2 台以上のモニタに出力できる機種を選定した。
- Thin Client の原因不明のハングアップ。
20 台の Thin Client を加速器の運転に用いていたが、毎月 1 台程度の頻度でいずれかの Thin Client がハングアップしていた。致命的とは言えないまでも加速器の運転に支障を来す問題であったが、原因が特定できるほど頻繁にハングアップするわけではなかった。

NUC と Blade 計算機の OS は同じ SL6 で、ネットワーク透過な X Window System が採用されている。従って、当然予想されることであるが、従来 Blade 計算機で実行し Thin Client で表示していた加速器制御アプリケーションを NUC で実行・表示しても操作上の違いは見られなかった。また、評価中に NUC がハングアップするような現象もみられなかった。

評価は満足のいくもので、端末用計算機の後継として NUC を採択した。MR コミッショニングの際にしばしば端末用計算機が不足する事態に陥ったので、20 台の Thin Client を 22 台の NUC で置き換えることにした。Table 1 に新旧端末用計算機の仕様を、Fig. 1 に t5730 と

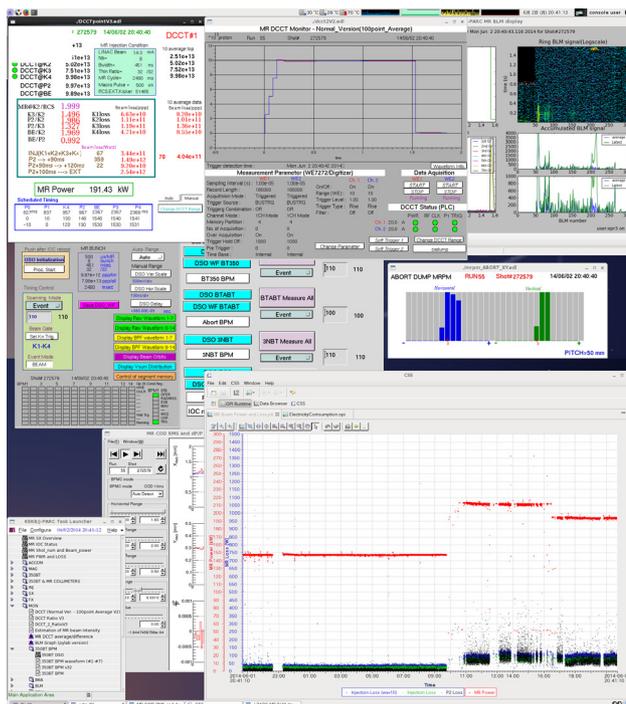


Figure 2: A Screenshot of Intel NUC During Commissioning of MR.

NUC 外観を示す。

20 台を超える NUC にあらかじめ OS をインストールしておき、一度に Thin Client を置き換えるには準備期間を必要とする。また、加速器の停止期間中に全ての端末を一度に置き換えた場合、評価の際に顕れなかった問題が発生すると加速器の運転再開に支障を来すことになる。そこで、MR 加速器を運転をしながら毎週 1~2 台程度の頻度で新しい端末に置き換えていき、仮に問題が発生したとしても短期間に集中しないような手法をとった。2014 年 1 月から端末計算機の機種更新を開始し、2014 年 7 月までに 1 台の Thin Client を残して更新を完了した。図 2 に、MR のコミッショニング中の NUC のスクリーンショットを示す。2014 年 1 月から 6 月までの MR の運転中に、NUC が原因不明のハングアップをするなどの重篤な障害は発生していない。

Table 2: Specifications of Blade Servers

Type	HS20	HS21	HS21a	HS22	HS22a	HS23e
CPU	Xeon	Xeon 5110, 5160	Xeon E5405	Xeon E5504	Xeon X5690	Xeon E5 2470
CPU Frequency	2.8 GHz	1.6, 2.0 GHz	2.0 GHz	2.0 GHz	3.46 - 3.74 GHz	2.3 - 3.1 GHz
Numer of Cores in Total	2	4	8	4	12 (24 threads)	8 (16 threads)
Amount of Memory	1 GB	2 GB	14 GB	20 GB	24 GB	48 GB

2.2 Blade 計算機

MR のビーム運転を開始した 2008 年には、サーバ用とアプリケーション実行用の Blade 計算機を合計 15 枚用いていた。2011 年には Blade 計算機上に仮想マシンを導入し、仮想マシン上で動作する IOC を MR の運転に投入した^{[9][10]}。2012 年に Blade 計算機は 24 枚にまで増強された^[5]。また、これまで 8 枚の Blade 計算機で実行していた主要なサーバ (RDB, ldap, dhcp, Archiver など) を仮想マシンに移行し、Blade 計算機としては 3 台に集約することで計算機資源の有効利用を図った。2014 年 7 月現在はサーバ用・MR 運転用・仮想マシンのホスト用の HS20 / HS21 / HS21a 型を、HS23e 型に更新中である。Table 2 にこれまでに導入してきた Blade 計算機の仕様を、Table 3 に 2008 年の MR 運転開始時と 2014 年 7 月現在の Blade 計算機の構成 (予備を除く) を示す。機種更新は 2014 年 10 月までに完了し、総数 24 枚 (HS23e 型 10 枚、HS22a 型 1 枚、HS22 型 10 枚、HS21 型 3 枚) となる見通しである。

3. まとめと今後の展望

MR 加速器の制御に用いる端末用計算機と Blade 計算機を刷新した。端末用計算機は MR のコミッショニングで実際に用いて評価し、加速器の運転中に新しい端末に置き換えていった。アプリケーションを Blade 計算機で実行し端末用計算機で表示する分離型の形態から、端末用計算機上で実行し表示する形態へ運用方法は変わったが、X Window System を用いることでその違いが端末用計算機を操作するユーザーからは隠蔽されている。

Blade 計算機上に構築した仮想マシンを MR の運転に投入して以来、仮想マシンの重要性が年々増している。仮想マシン上で動作する IOC の数は 2013 年には 31 台であったのが 2014 年には 38 台にまで増えており、今後も増加していくことが見込まれている。MR の運転に用いる制御アプリケーションを端末用計算機の上で実行できるようになったため、制御アプリケーションの実行に割り当てられていた Blade 計算機を仮想マシンのホストに転換していくことが予想される。

参考文献

- [1] EPICS - Experimental Physics and Industrial Control System, <http://www.aps.anl.gov/epics/>
- [2] <http://www.scientificlinux.org>
- [3] S. Yoshida, *et al.*, "Console System Using Thin Client for the J-PARC Accelerators", Proceedings of ICALEPCS 2007, p383 (2007).
- [4] T. Koseki, *et al.*, "J-PARC MR の運転状況", 加速器学会誌 2012 年 9 巻 1 号, p30-40 (2012).

Table 3: Changes of Numbers of Blade Servers

	purpose	type and number
2008	server (web, RDB, ldap, ...)	HS20 × 4
	MR operation	HS21 × 9
	RCS	HS21 × 2
2014	server (web, login)	HS23e × 1 HS21 × 1
	MR operation	HS23e × 4 HS22a × 1 HS22 × 3 HS21 × 2
	VM Host (server)	HS22 × 2
	VM Host (MR IOC)	HS23e × 2 HS22 × 3
	RCS	HS22 × 2 HS21 × 2

- [5] N. Kamikubota, *et al.*, "Improvement of Computer Systems for J-MARC MR Control", Proceedings of the 9th Annual Meeting of Particle Accelerator Society of Japan, p.741, WEPS117, (2012).
- [6] CSS - Control System Studio, <http://controlsystemstudio.org/>
- [7] S. Yamada, "Upgrade of software toolkits for EPICS Input Output Controllers in J-PARC Main Ring", Proceedings of the 10th Annual Meeting of Particle Accelerator Society of Japan, p.1106, SUP080 (2013).
- [8] <http://www.livecd.ethz.ch/diskless.html>
- [9] N. Kamikubota, *et al.*, "Virtual IO Controllers at J-PARC MR using Xen", Proceedings of the ICALEPCS 2011, Grenoble, France Oct.2011 p.1165.
- [10] N. Kamikubota, "J-PARC MR 制御での仮想マシンの応用", Proceedings of the 10th Annual Meeting of Particle Accelerator Society of Japan, SAP092 (2013).