

CPU LOADS AND NETWORK TRAFFIC IN J-PARC MR CONTROL SYSTEM

Makoto Takagi ^{#,A)}, Norihiko Kamikubota^{B)}, Noboru Yamamoto^{B)}, Susumu Yoshida^{A)}, Shigenobu Motohashi^{A)}, Takao Iitsuka^{A)}, Daisuke Takahashi^{A)}, Hiroyuki Nemoto^{C)}

^{A)} Kanto Information Service(KIS)

8-21 Bunkyo, Tsutiura, Ibaraki, 300-0045

^{B)} High Energy Accelerator Research Organization(KEK)

1-1 Oho, Tsukuba, Ibaraki, 305-0801

^{C)} ACMOS INC.

2713-7 Muramatsu, Tokai, Naka, Ibaraki, 319-1112

Abstract

Beam commissioning of J-PARC Main Ring (MR) has been started in May, 2008. A lot of applications were developed in two years, and it is used for the control of the accelerator and the beam commissioning. In J-PARC MR, Blade type Server computer was maintained because of the application for beam commissioning. CPU load of Blade type Server computer has increased as the application increases. And the load of the MR control network increases for the increase of observed EPICS record. In this report, the change in CPU load and the change in the amount of the network traffic in J-PARC MR Control System are described.

J-PARC MR 制御システムの CPU 負荷とネットワークトラフィック

1. はじめに

J-PARC MR の運転は 2008 年 5 月より開始された^[1]。この 2 年間で様々なアプリケーションが開発され、加速器の制御やビームコミッショニングに利用されている。アプリケーションの増加に伴い、運転用アプリケーションの為に整備された Blade 型 Server 計算機の CPU 負荷は徐々に増加している。また、監視レコード数の増加やサイズの大きな波形データのモニタ・ディスクへの記録などによって、MR 制御ネットワークの負荷も徐々に増加している。

CPU 負荷やネットワークトラフィックなどの監視には Cacti を利用している^[2]。Cacti を利用する事で、SNMP エージェントが取得した値やスクリプトの出力結果をグラフ化することが可能である。監視機器・項目の追加や監視データのグラフ表示は付属の Web インタフェースを利用する事によって簡単に行う事が可能となっている。

2. Blade 型 Server 計算機

MR 制御システムでは、運転用アプリケーションを実行する為の計算機として Blade 型 Server 計算機を利用している。MR 運転開始時には 4 枚の Blade(Intel Xeon Dual Core 3GHz, 2GB Memory)が MR 運転用アプリケーションの為に整備された。OS には Scientific Linux 4 が採用されている。2009 年 3 月には、Blade の CPU 負荷増加に伴い、開発用として利用していた同スペックの 2 枚の Blade が運転用として追加された。2010 年には更に 5 枚の Blade(Intel Xeon 4core 2cpu 2GHz, 20GB Memory)が追加された。OS には Scientific Linux 5 が採用されている^[3]。

アプリケーションの起動は、J-PARC Task Launcher(jkl)から行う。アプリケーションの起動の際にどの Blade を利用するかをメニューから選択する事が可能となっており、ユーザーは比較的負荷の低い Blade を選んでアプリケーションを実行する(運転中常時起動しているアプリケーションや Blade に高い負荷の掛かるアプリケーションに関しては、CPU を固定している場合もある)。図 1 は J-PARC Task Launcher のメニューからアプリケーションを実行する Blade を選択しているところである。Blade の横に Load の状態が表示され、Load の低い Blade を選択する際の目安となる。

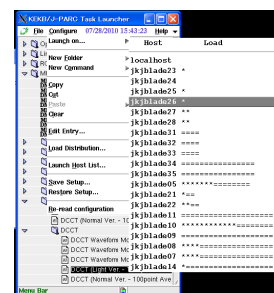


図 1 : J-PARC Task Launcher(jkl)

図 2 は、MR 運転開始時から稼動している 4 枚の Blade と 2009 年 3 月に追加された 2 枚の Blade の CPU 負荷(Load Average)の平均値の変化のグラフである。

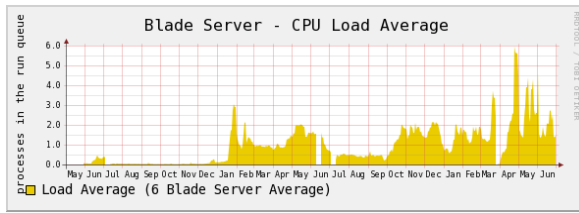


図 2 : MR 運転開始時からの Blade の CPU 負荷

Load Average は 1CPU で処理可能な量を 1 とした場合の CPU リソースの要求量を反映した値である。Blade は dual core の CPU を搭載している為、Load Average の値が 2.0 を超えると CPU の負荷が高い状態であると言える。現状では何らかの理由で一時的に大きな値になる事はあっても、2.0 を超える状態が続く事はあまりないようである。今後は、新たに追加された 5 枚の Blade へのアプリケーションの移行が進むと思われる為、負荷は下がっていくと考えられる。

図 3 は、新たに追加された 6 枚の Blade の CPU 負荷の平均値の変化のグラフである。

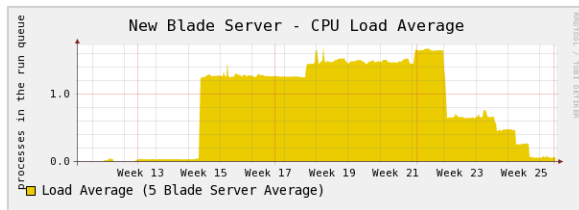


図 3 : 新たに追加された Blade の CPU 負荷

新たに追加された Blade は 4 core の CPU を 2 つ搭載している為、Load Average の値が 8.0 を超えると CPU の負荷が高い状態と言える。現状ではピーク時でも 8.0 を大きく下回っており、処理能力的にはまだまだ余裕のある状態と考えられる。

図 4 は、新しく追加した Blade の内の 1 枚の CPU 負荷の変化のグラフである。

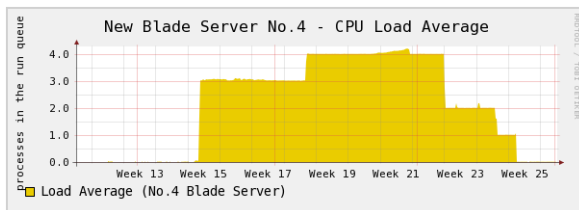


図 4 : 新たに追加された Blade の内 1 枚の CPU 負荷

この Blade は Load Average の値が長時間 4.0 をキープしている。これは、Blade で運転用の高負荷アプリケーションが 4 つ稼動しており、8 つの CPU の内 4 つの CPU を占有していた為である。このケースでは、高負荷のアプリケーションと実行される Blade の管理が出来ていた為、特に問題は無い。しかし、原因不明の CPU 負荷の増加には注意する必要があり、監視が必要である。

3. ネットワークトラフィック

図 5 は、MR ネットワークの概要図である。MR 制御システムでは、MR の各電源棟(D1,D2,D3)およびニュートリノ棟(NU),ハドロン棟(HD)のネットワークトラフィックは、全て D3 電源棟の 1 台のネットワークスイッチに集約され、中央制御棟(CCB)へ送られる。この D3 電源棟のネットワークスイッチの各ポートを監視する事により、MR 制御システムのネットワークトラフィックを大まかに把握する事が可能である。

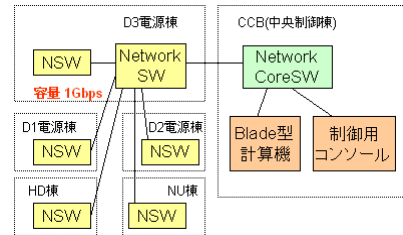


図 5 : MR ネットワーク概要

図 6 は、CCB のトラフィックの変化のグラフである。

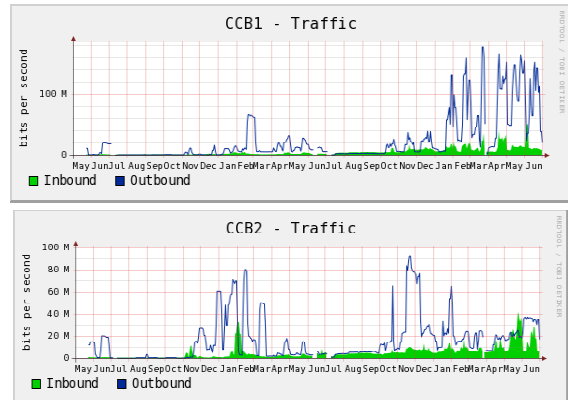
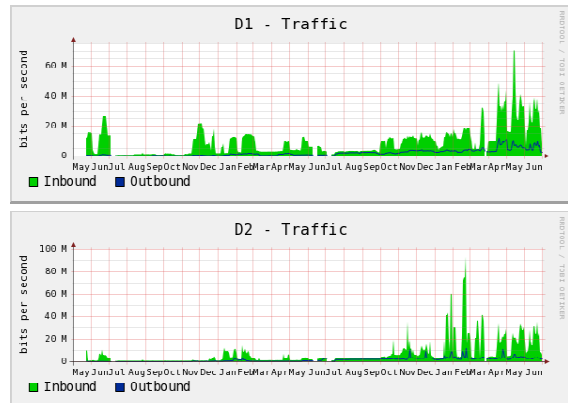


図 6 : CCB(中央制御棟)へのトラフィック

2010 年に入ってからトラフィックが大幅に増えており、100Mbps 以上のトラフィックが流れている事が分かる。

図 7 は、D1,D2,D3 電源棟のトラフィックの変化のグラフである。



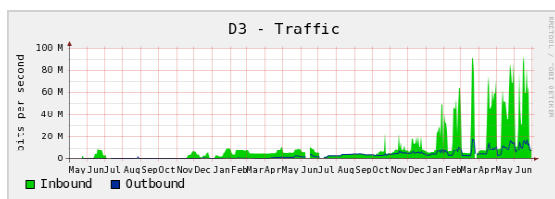


図 7 : D1,D2,D3(MR 電源棟)へのトラフィック

CCB と同様に入ってからトラフィックが大幅に増えている。どのポートでトラフィックが増加しているのか、どのアプリケーションがトラフィック増加の原因なのか、今後調査をしていく必要がある。

4. Data Archive

MR 制御システムでは、運転状態の記録に EPICS 標準の Data Archive System である Channel Archiver を利用している^[4]。Channel Archiver では、データ収集に ArchiveEngine というアプリケーションを利用する。MR 運転開始時点では 1U のラックマウント型計算機にて ArchiveEngine を動かしていたが、2009 年 1 月より、Blade 型 Server 計算機(Dual Core Intel Xeon 3GHz CPU, 2GB Memory)に移行した。

図 8 は、2009 年 1 月以降の Blade の CPU 負荷とネットワークトラフィックの変化のグラフである。

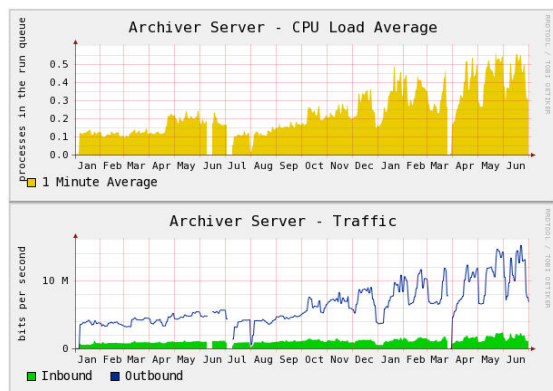


図 8 : Archiver 用 Blade の CPU/ネットワーク負荷

MR 加速器の運転が進むにつれて監視する EPICS レコードの点数が増加した影響で、2009 年 1 月と比べて CPU 負荷が約 4 倍、ネットワークトラフィックは約 3 倍に増加している。しかし、Load Average は現時点でも 0.5 程度であり、Dual Core の CPU を搭載している(Load Average が 2.0 を超えると高負荷)事を考えればかなり余裕のある状態と言える。また、ネットワークトラフィックは現状でも 15Mbps 程度であり、1Gbps の性能を考慮すると全く問題無い数字であることが分かる。今後、監視対象 EPICS レコードが 2 倍程度に増えたとしても、十分に対応可能であると考えられる。

5. Disk Server

MR 加速器制御システムでは、運転データ保存の

為の Disk Server を整備している。Disk Server はネットワーク経由で NFS mount される事によって利用され、Channel Archiver のデータや運転中の制御画面の Screen Shot の保存先として利用されている。現状では 8.7TB が利用可能である。

図 9 は、2008 年 7 月以降の Disk Server のネットワークトラフィックの変化のグラフである。

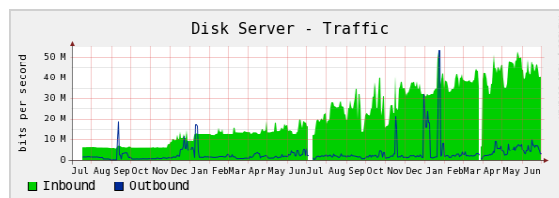


図 9 : Disk Server のネットワーク負荷

MR 加速器の運転が進むにつれてネットワークトラフィックは約 8 倍程度に増加している。ネットワークトラフィックの約 1/3 は、Channel Archiver によるものである。それ以外では、DCCT や BPM などの波形データをショット毎にファイルに保存するプログラムが常時動作しており、そのデータ保存によるトラフィックが大部分であると考えられる。ネットワークトラフィックは現状でも 50Mbps 程度であり、1Gbps の性能を考慮すると全く問題無いと言える。

4. まとめ

MR の運転開始当初と比べると、CPU 負荷・ネットワークトラフィック共に増加傾向にある。

新規 Blade の追加などもあり、CPU ついては現状でも余裕がある状態であるが、特定の Blade に負荷が集中したりしないよう、アプリケーションの配置を含めた監視は必要であろう。

1Gbps というインタフェースの性能を考えると、ネットワークトラフィックに関しても現状では問題無いと考えられるが、トラフィックの変化は常に監視する必要がある。

今後も引き続き CPU 負荷やネットワークトラフィックを監視を行っていき、今後の CPU・ネットワークの増強の目安としたい。

参考文献

- [1] T.Koseki, et al., "Status of J-PARC Main Ring Synchrotron", PAC07, p736-738
- [2] <http://www.cacti.net/>
- [3] N.Kamikubota, et al., "J-PARC MR 運転用計算機の利用環境整備", in this meeting
- [4] N.Kamikubota, et al., "Data Archive System for MR J-PARC Main Ring", IPAC10, Kyoto, Japan, <http://ipac10.org/>