

LOAD TEST FOR THE DATABASE SYSTEM OF SACLA

Toko Hirono ^{#, A)}, Toru Fukui ^{A), B)}, Takashi Hamano ^{A)}, Takemasa Masuda ^{A)}, Akihiro Yamashita ^{A)}

^{A)} Japan Synchrotron Radiation Research Institute

1-1-1 Sayo-cho Sayo-gun, Hyogo, 679-5198

^{B)} RIKEN

1-1-1 Sayo-cho Sayo-gun, Hyogo, 679-5198

Abstract

We performed load test on the database system of MADOCA for SACLA. The database system was consisted of many components, such as server computers, storage system, network, client workstations. Each component was selected to meet the requirements from SACLA operation. However the system has not tested as one system. The actual servers were used at the load test to ensure the system performance. The bottle-neck of the system was also investigated. The results show the system has appropriate ability and the memory of primary database server was the bottle-neck of the system.

SACLA 制御系データベースシステムの負荷試験

1. はじめに

SPring-8 Angstrom Compact Free Electron Laser (SACLA) [1]の制御系には Message and Database Oriented Control Architecture (MADOCA) [2]が採用されている[3]。MADOCA の機能の一部であるデータ収集やアラーム監視には、リレーショナルデータベースを中心としたデータベースシステムが構築されている。このデータベースシステムの個々の構成要素には、これまで MADOCA が導入されている SPring-8 での実績[4]を基に必要な性能をもつ機器が選定されている。しかし、これらを組み合わせてシステム全体としての性能評価を行ったことがなかった。MADOCA では収集された全データを定期的にリレーショナルデータベースに保存しているため、特にデータベースの性能が重要になる。

一方、SACLA の制御系では、コミッショニング時から 40000 点以上の信号の収集を行う必要があり、MADOCA が導入された制御系のなかでも最大の規模になる見込みであった。そのため、SACLA の制御系データベースシステムがコミッショニング時に必要なデータ収集やアラーム監視が行えること、およびシステムのボトルネックの特定をする必要があった。そこで、SACLA 制御系に導入したデータベースシステムの負荷試験を行った。

本稿では、まず、第 2 章で SACLA の制御系に導入されているデータベースシステムの構成について説明し、第 3 章に負荷試験の方法、第 4 章にその結果を示す。最後に、第 5 章で負荷試験についてまとめるとともに、今後のデータベースシステムの課題について議論する。

2. データベースシステムの構成

MADOCA の機能のうち、データ収集とアラーム監視の構成を図 1 に示す。データ収集とアラーム監視は 2 台のデータベースサーバと Collector Client 用

計算機、アラーム監視用計算機とそれに接続される複数の機器やオペレーターコンソールから構成されている。SACLA の制御系に導入されているそれぞれの計算機の仕様を表 1 に示す。

表 1: ハードウェア構成

プライマリデータベースサーバ	CPU: Intel Xeon E5504 2GHz x 8 core OS: RedHat Enterprise Linux 5.5 64bit Memory: 16GB Network: 1G x 2 port ストレージ: NetApp
セカンダリデータベースサーバ	CPU: Xeon X5407 3.3GHz x 8 core OS: CentOS 5.5 64bit Memory: 16MB Network: 1G x 1 port Webserver: Apache 2.0 ストレージ: 7.5krpm SAS
Collector-Clientサーバ	CPU: Opteron 2216HE 1GHz x 2 core OS: SuSE Linux 12 Memory: 4GB Network: 1G
アラーム監視サーバ	CPU: Opteron 2216HE 1GHz x 2 core OS: SuSE Linux 12 Memory: 4GB Network: 1G
オペレーターコンソール(6台)	CPU: Opteron 2216HE 1GHz x 2 core OS: SuSE Linux 12 Memory: 4GB Network: 1G

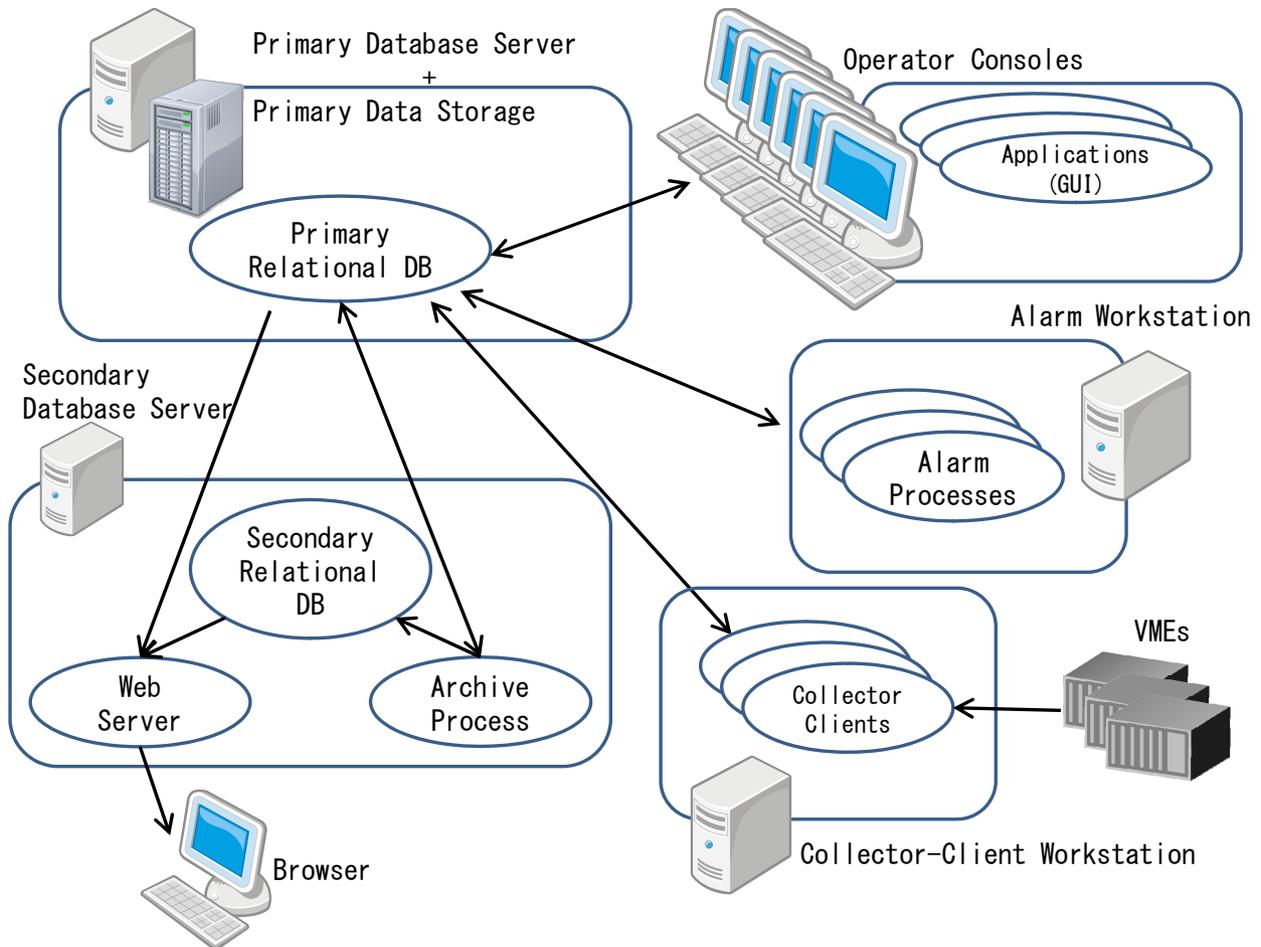


図1: データ収集系の構成図。リレーショナルデータベースを中心に複数のサーバ計算機から構成されている。

データベースシステムではプライマリデータベースとセカンダリデータベースの2台が使われている。これらのデータベースに接続するクライアントは次の5種類である。

- **Collector Client:** 機器等の制御対象からデータ収集を行うプロセスである。VME 計算機などの機器制御から決められた周期でデータを集め、主データベースへタイムスタンプとともに時系列のデータとして保存する。
- **アーカイブプロセス:** Collector Client によって収集、保存されたデータを主データベースの別領域とセカンダリデータベースの2か所にコピーする。また、アーカイブプロセスはデータの間引きも行っている。
- **GUI などの加速器運転用アプリケーション:** データベースへのアクセスには専用の API 関数が準備されており、GUI など SACLA の運転に使用されるアプリケーションはこの API を使用してデータベースのデータを利用している。機器から収集された時系列データの最新値にアクセスするものが多い。

- **アラーム監視プロセス:** 機器から収集されたデータの値とあらかじめ設定されている閾値を比較し、アラームの条件が成立すると GUI と音声によって通知するプロセスである。
- **ウェブサーバの CGI:** 機器から収集されたデータはウェブサーバの CGI を使ってグラフ表示することができる。データはアーカイブプロセスがコピーしたセカンダリデータベースから取得している。

3. データベースシステムの負荷試験

SACLA の制御系では、加速器のエージングなどのために既にデータ収集やアラーム監視を開始しているデータに加えて、約 20000 点のデータが増加する見込みであった。そこで、データベースシステムの負荷試験では既存のデータにダミーデータを追加してデータ収集やアラーム監視を行った。また、現在のデータシステムで扱えるデータ数の最大値を得るため、29000 点追加した条件でも負荷試験を行った。追加したデータは次のとおりである。

- 条件 1:
 - データ収集:

- 既存データ 23675 点
(周期 1sec, 2sec 14757 点、10sec 以上 8918 点)
- ダミーデータ 19000 点 x 2 sec
- アラーム監視:
 - 既存データ 260 点
 - ダミーデータ 5000 点
- 条件 2:
 - データ収集:
 - 既存データ 23675 点
 - ダミーデータ 29000 点 x 2 sec
 - アラーム監視:
 - 既存データ 260 点
 - ダミーデータ 5000 点

負荷試験ではデータ収集、アラーム監視がシステム全体として必要な性能を持っていることを実証する必要がある。そこで、負荷試験時に未設置であった VME を除き、すべて実機を用いて行った。VME の代わりに仮想サーバを条件 1 では 3 つ、条件 2 では 4 つ準備し、その仮想サーバでダミーのデータを作り、周期を 2 秒に設定して Collector Client で収集した。

上記条件下で Collector Client、アーカイブプロセス、アラーム監視プロセスを起動し、プライマリデータベースサーバの CPU 負荷率とネットワークトラフィック量を Linux OS 標準の System Admin Reporter 機能で測定した。また、Collector Client についてはデータベースに保存された時系列データのタイムスタンプからデータ収集の遅延量についても調べた。また、アプリケーションからのアクセスを評価するため、時系列データの最新値の取得にかかる時間を測定した。前章で述べた 5 種類のデータベースの接続のうちウェブサーバの CGI はアクセス数が少ないことおよび、そのほとんどがセカンダリデータベースへのアクセスであることから、省略した。

4. 結果

各クライアントを起動させたときのプライマリデータベースの CPU 負荷とネットワークトラフィック量を図 2 に示す。

条件 1 の場合、アーカイブプロセス、アラーム監視プロセスが問題なく動作した。また、時系列データの最新値を 2 秒周期で 12500 点取得した場合でも CPU、ネットワークトラフィックともに余裕があることが分かった。図 3 に条件 1 の場合のデータ収集の遅延量と示す。時系列データの最新値を 2500 点行った場合、Collector Client は 99% のデータを遅延なくデータ収集することができた。また、応答時間は平均で 1 点当たり 146msec であった。5000 点以上のデータ取得を行うと、データ収集の遅延が顕著になることが分かった。2500 点の時系列データの最新値を取得することはデータ収集を行っているデータのうち約 8 点に 1 点を GUI などの加速器運転に使用するアプリケーションで常時使用することと同義である。しかし、実運用では、データ収集を行う予定

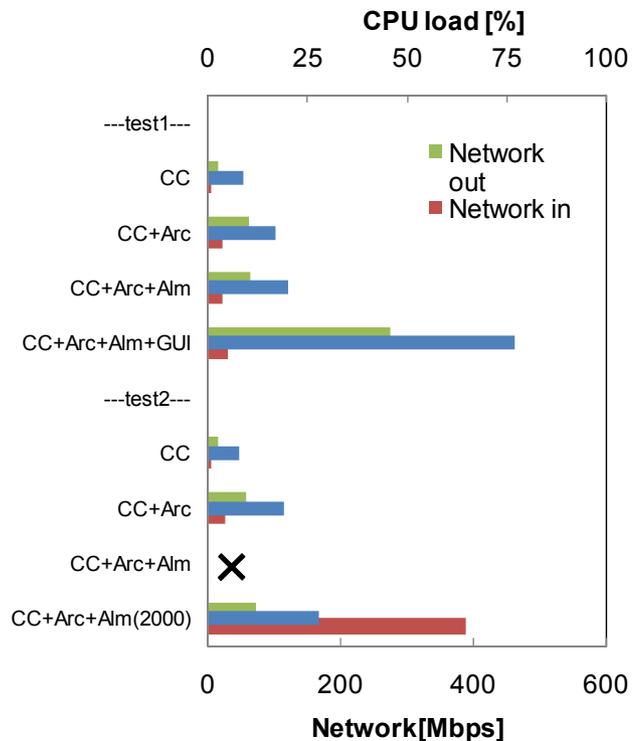


図 2: 条件 1 と条件 2 の場合のプライマリデータベースサーバにおける CPU 負荷率とネットワークトラフィック量。Collector Client(CC)、アーカイブプロセス(Arc)、アラーム監視(Alm)、最新値取得アプリケーション(GUI)を順次起動して測定を行った。ただし、CC+Arc+Alm(2000)は条件 2 のうちアラーム監視するデータ点を 2000 点として測定した結果である。

のデータのうち多くがデータのアーカイブやアラーム監視に使用するのみで、加速器運転に使用するアプリケーションで常時使用するデータは 2500 点以下となる見込みである。よって、本結果からコミッショニング時に必要な性能を有していることができる。

条件 2 の場合、Collector Client とアーカイブプロセスとアラーム監視を同時に起動することができなかった。そこで、アラーム監視を行うデータ点を 2000 点に減らして、どこにボトルネックがあるかの調査を行った。CPU およびネットワークトラフィックはどちらも計算機の性能と比較すると余裕があるが、ネットワークトラフィックのうち受信データ量が条件 1 に比べて 10 倍以上増えていることがわかった。プライマリデータベースのストレージはネットワーク接続型であるため、ネットワークトラフィックはリレーショナルデータベースとクライアント間の通信とストレージとの和である。そこで、ストレージとリレーショナルデータベース間、クライアントとリレーショナルデータベースと間の通信量、およびリレーショナルデータベース内のメモリヒット率を測定した。測定の結果を表 2 に示す。

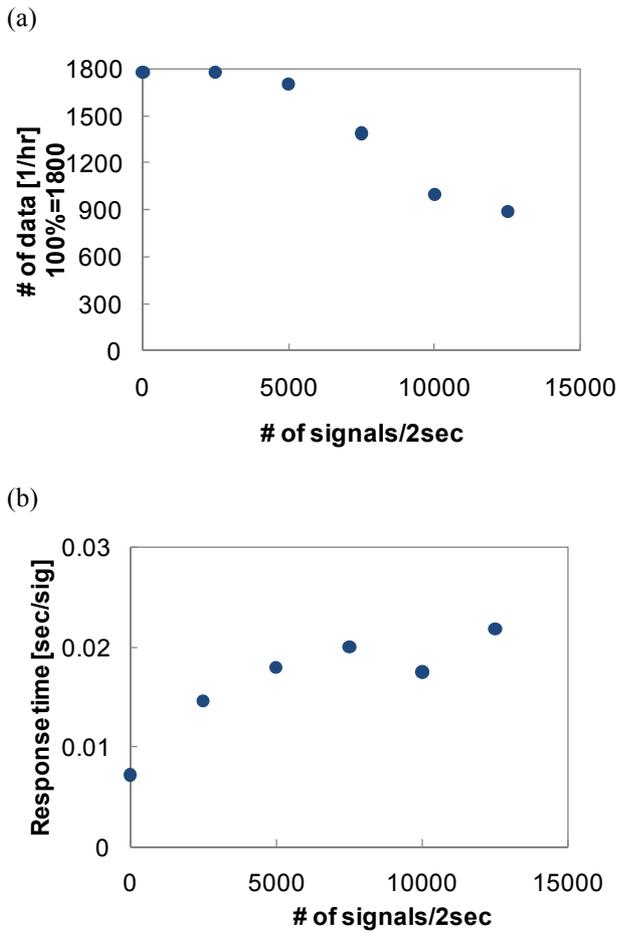


図 3: 条件 1 の場合のオペレーターコンソールから最新値のデータを取得した場合のデータ収集系が保存できたデータ数(a)とオペレーターコンソールで 1 点の時系列データの最新値を取得するのにかかった時間(b)。横軸は取得するデータ点数である。

表 2: プライマリデータベースのデータ受信量

	条件 1	条件 2
ストレージアクセス	420times/sec	22595times/sec
クライアント	914442Mbps	321788Mbps
メモリヒット率	92.8%	56.8%

表 2 より、受信データ量の増加はリレーショナルデータベースがストレージからデータを読み込むことが原因であることがわかる。また、データベースのメモリヒット率も低下している。リレーショナルデータベースは、条件 1 ではキャッシュ内のデータを用いて処理ができていたが、条件 2 ではキャッシュの容量を超えてしまい、処理のつどデータをストレージから読み込んでいると考えられる。クライアントからの受信量が条件 1 より条件 2 の方が大きいのは、条件 2 の方がデータベースの応答が遅く、クライアントからの処理命令のデータ量が少ないためである。以上のことから、データベースシステムのボトルネックはプライマリデータベースのメモリ量であることが分かった。

5. まとめと今後の課題

実環境を模擬した負荷試験の結果、SACLA の制御系に導入されているデータベースシステムは次のとおりまとめることができる。

- 42000 点までのデータを収集し、そのうち 5000 点のデータに対してアラーム監視をおこない、GUI などのアプリケーションで利用することができる。
- 52000 点のデータをデータ収集し、アラーム監視やアプリケーションで利用するためにはプライマリデータベースのメモリ増強が必要である。

今後の課題として、メモリを増強した場合の次のボトルネックとデータ収集可能なデータ点数の把握が必要であるが、今回の負荷試験を行った際の手法を用いれば、これらの情報を得ることができる。

参考文献

- [1] <http://xfel.riken.jp>
- [2] R. Tanaka et al., "Control System of the SPring-8 Storage Ring", ICALEPCS'95, Chicago, USA, 1995, p.201.
- [3] 福井さんの ICALEPCS(2007? 2009?)
- [4] A.Yamashita et al., "Data Archiving and Retrieval for Spring-8 accelerator complex", Trieste, Italy, 1999, p434.