

MADOCA DATABASE SYSTEM WITH DUAL SERVERS FOR SACLA

T. Hirono ^{#, A)}, T. Fukui ^{A), B)}, T. Hamano ^{A)}, T. Maruyama ^{B)}, T. Yamaga ^{A), B)}, A. 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 upgraded the database system for the control system of SPring-8 Angstrom Compact Free Electron Laser (SACLA). The upgraded system consists of two database servers to increase server resource such as CPU, memory, disk IOs. The main server handles online data and the archive server handles archived data. Client programs including GUI for operation connected to the main server. The archived data mounted on the main server by the proxy tables, so that the client programs are able to use without any changes. We performed load tests on the new database system. The new system was successfully able to handle data signals of 84,246, while around 80,000 signals are required for the control system in SACLA in a few years.

SACLA 制御系データベースシステムの増強

1. はじめに

SACLA の制御フレームワークには Message and Database Oriented Control Architecture (MADOCA)が採用されており[1]、データベースシステムはデータ収集やアラーム監視などの一部として重要な役割を担っている[2]。しかし、取り扱う信号数やデータベースにアクセスする運転用制御プログラムの増加により CPU 負荷率が常に 50%を超えるなど、高負荷の状態が続いてしまっている。今後、SACLA では加速器やビームラインの増設が予定されているが、現在のデータベースシステムの構成では対応することができない[3]。データベースシステムの増強が必要である。本論文では 2 章と 3 章にデータベースシステムの概略と要求される要件について述べ、4 章に増強後のデータベースシステムのデザインについて議論し、5 章に負荷試験の結果を示し 6 章にまとめる。

2. データベースシステムの概略

MADOCA のフレームワークのうちデータベースシステムは次のような機能を担当している。

- データ収集: データ収集は 2 つのプロセスで構成されている。1 つめのプロセス(Collector Client) は 1 秒から 60 秒の定周期で機器等の制御対象の信号データを収集し、そのデータをデータベースシステムはタイムスタンプつきで時系列にリレーショナルデータベースに保存する。その後 2 つ目のプロセスで 5-10 分おきに定期的にデータを別テーブルにアーカイブし、一定期間後にデータの間引きを行っている。

- アラームの監視: MADOCA のデータベースシステムではあらかじめ信号ごとに監視条件と閾値を設定することができる。データベースシステムでは、リレーショナルデータベースに保存された信号データの最新値と閾値を所定の条件で比較し、アラームとして GUI に表示し、音声通知する機能を有している。また、アラームの発生日時などの履歴もデータベースに保存されている。
- API 関数によるデータ参照機能: MADOCA では GUI などの運転用制御プログラムから SQL やデータベースの構造の知識なくデータベースのデータにアクセスできるよう C 言語の API が提供されている。
- データブラウズ機能: データベースシステムではデータベース内のデータにアクセスできるようウェブサーバと CGI を使ってウェブブラウザへデータを表示する機能を有している。

3. データベースシステムの要件

SACLA の制御系データベースの増強を設計する際、次のような条件を要件とした。

3.1 必要な制御点数の信号が取り扱えること

現在、SACLA のデータベースシステムでは 43,926 点の制御信号のデータ収集と 8,203 点のアラーム監視を行っている。さらに、現在 SACLA の制御系に含まれていない SPring-8 Compact SASE Source が移設され、SACLA の制御系の一部として制御される予定である。また、ビームラインの増設も予定されている。このため、制御信号は 20,000 点程度の増加が予測される。そのほかにも、機器の追加などに従い、制御信号は増加することから、今回

[#] hirono@spring8.or.jp

の増強では 80,000 点以上の信号を取り扱えるデータベースシステムにすることとした。

3.2 ブラウザ機能がほかのプロセスに影響を及ぼさないこと

1 か月以上の長期間の機器の安定性を確認するなど、Web ブラウジングの機能を使って大量のデータを表示させる際、大きな負荷がデータベースにかかり、データベースの応答が遅くなる。一方、データ収集では Collector Client から常に最短で 1 秒周期でデータが送られてくる。これらのデータを欠落させることなく、データベースに保存する必要がある。ブラウザ機能の負荷がほかのプロセスに影響しないシステム構成が必要である。

3.3 API 関数に変更がないこと

SACLA の運転に用いる GUI など多くの制御プログラムがすでに開発されており、API の変更を行うとそれら制御プログラムの再コンパイル作業だけでなく、動作確認など膨大な作業が必要となる。特に運転に使用されている制御プログラムは高度な機能を複雑なアルゴリズムで実装されており、それらすべての動作確認を行うことは大変困難である。そこで、今回の増強では API 関数の変更なしに行えることを要件とした。

3.4 冗長性、高可用性の高いシステムであること

MADUCA のフレームワークではデータベースシステムは必須の構成要素であり、データベースシステムの停止は SACLA の運転も停止につながる。現行のシステムと同等もしくはそれ以上の冗長性、高可用性を持つシステムになるよう設計を行った。

3.5 入れ替え作業の作業時間が短いこと

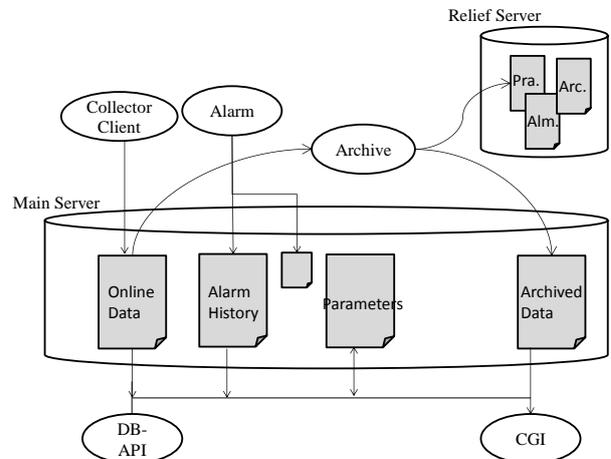
SACLA はすでにユーザーへの供用が開始しており、作業は 1 か月の夏期停止期間に限られる。実環境での試用期間も考慮し、1 週間程度で作業が完了できる必要がある。

4. 構成の設計

図 1 にシステム増強前と増強後の構成を示す。今回の増強では 1 台だったサーバ機を 2 台構成に変更した。これにより、CPU 処理能力、メモリ容量、ディスクの IO などサーバのリソースを 2 倍にすることができる。

2 台のサーバ機にはともに MADUCA で採用されているリレーショナルデータベース Sybase Adaptive Server Enterprise 15.5 Linux 64-bit 版[4]をインストールした。1 台目のメイン機には約 2 時間分の最新のデータとパラメーター等の各種設定値のみを保存し、アーカイブされた大量の過去のデータは 2 台目のアーカイブ機に保存することとした。Collector Client やアラーム監視などはアーカイブされたデータにアクセスする必要がないため、メイン機のみ接続する。このような構成にすることによってブラウザ機能でアーカイブされたデータに大量のアクセス

(a)



(b)

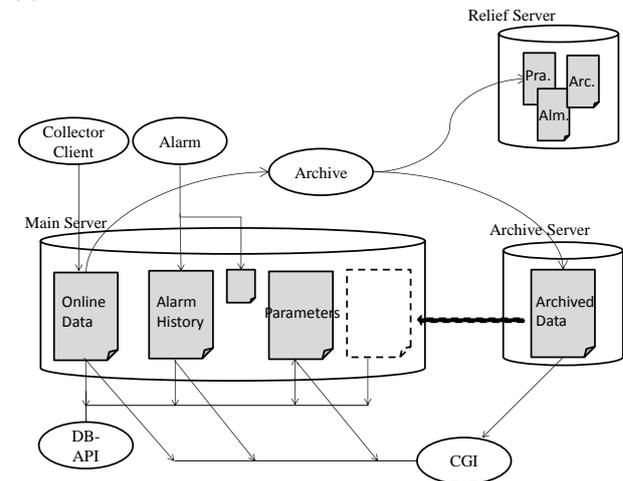


図 1：システム増強前(a)と増強後(b)のデータベースシステムの構成

を行ってもメイン機には影響はなく、Collector Client からのデータは安定してサーバに保存することができる。

GUI など運転用制御プログラムはオンラインデータ、アーカイブデータともにアクセスする必要がある。そこで、Sybase Adaptive Server Enterprise のプロキシテーブル機能[4]を用いた。この機能は、アーカイブ機上のデータをメイン機に接続したプロセスからあたかもメイン機にあるデータであるかのようにアクセスできるものである。これによって運転用制御プログラムに影響することなく構成を変更することができる。

追加するサーバ機には現行機で実績のある Fault Tolerant (FT) サーバを採用した[5]。表 1 に増強後のサーバ機等の諸元を示す。ストレージもオンラインデータ用には 2 重化されている FT サーバのローカルディスクを使用し、大容量の必要なアーカイブ機にはネットワークストレージを採用した。ネットワークストレージのストレージやネットワークイン

ターフェースもすべて二重化されたものを選定し、故障の際も無停止で交換できる構成とした。また、サーバ機の故障やオペレーションミスに備えて従来どおり、リリーフ機も運用することとした。

図 1 (a)の構成から(b)への構成変更はデータを 2 つのサーバに移行する作業とデータをアーカイブするプロセスのアーカイブ先をメイン機からアーカイブ機へ変更するプロセスの改造のみで実現することができる。このため 1 か月の運転停止期間で十分に実現可能である。

5. 負荷試験の結果

本格導入の前に、サーバ機 2 台構成の増強後のデータベースシステムを実環境に試験的に導入し性能を比較した。運転時と同じ状態になるよう、データ収集、アラーム監視は運転に使用しているものを起動した。また、将来の加速器等の増築に対応できることを確認するため、40,000 点のダミーの制御信号を VME に追加し、実際に運転に使用しているプロセスを使って 2 秒周期でデータ収集し、そのうち 5,000 点をアラーム監視に追加した。2 秒周期は現在の SACL A で行われているデータ収集周期のうち一番信号数の多い周期である。運転用制御プログラムは、500 点の制御信号の最新値を 2 秒周期で取得するプログラムを 15 プロセス起動して模擬することとした。試験を行った条件を表 2 に示す。

システムの負荷状況はサーバ機の CPU 負荷、疑似運転プログラムで機器の信号の最新値を取得する際にかかる時間を指標にすることとした。図 2 に CPU の負荷、図 3 に最新値を取得にかかる時間の測定結果を示す。Test 3 場合、CPU の負荷は 71%であったが、増強後のメイン機では取り扱う制御信号数が 83,926 点であっても 60%以下になっている。さらに Test 2 でアーカイブされた過去のデータをブラウザ機能で表示させてもメインサーバ機の CPU 負荷に影響がないことが確認できた。

また、疑似運転プログラムで測定したデータベースのデータ取得時の応答では Test 3 より Test 1 は平均して 2 ms 程度応答が遅いが、1 信号あたり 6 ms 以下であり大きな遅延はない。データベースシステム全体として要件である 80,000 点以上の制御信号を処理できていることがわかった。

6. まとめ

システムの増強のためデータベースシステムの構成の変更を行った。1 台だったデータベースサーバを 2 台に増やし、サーバの能力を 2 倍に増強した。この際、アーカイブされたデータを保存するサーバとデータ収集やアラームで使用されているデータのサーバを分離した。これはより安定したデータ収集等が行える設計である。実際に実環境に試験導入し、負荷試験を行った。80,000 点以上の制御信号データを処理できることを確認した。

本システムは 2012 年度夏季停止期間での導入を予定している。

表 1：データベースシステムのサーバ機の諸元

Main Server	
機種	NEC Express 5800 R320b-M4
CPU	Xeon 2.93 GHz x 12 core (内 8core を DB に使用)
Memory	48GB
Disk	Local Disk: SATA 15000rpm 1TB
Archive Server	
機種	NEC Express 5800 R320a-E4
CPU	Xeon 2.00GHz x 8 core (内 7core を DB に使用)
Memory	48GB
Disk	Local Disk: SATA 177GB NetApp: 8TB
Relief Server	
機種	HP ProLiant DL180 G6
CPU	Xeon 2.67GHz x 12 core (内 8 core を DB に使用)
Memory	48GB
Disk	Local Disk SATA 15000rpm 600GB x 14

表 2：負荷試験の条件

Test 1	
サーバ構成	2 台(増強後)
データ収集	1 秒周期 380 点 2 秒周期 65,972 点 10 秒周期 18,327 点 60 秒周期 247 点
アラーム監視 データ参照	13,203 点 疑似運転プログラム(500 信号 の最新値を 2 秒周期で取得) x 15
Test 2	
サーバ構成	2 台(増強後)
データ収集	1 秒周期 380 点 2 秒周期 65,972 点 10 秒周期 18,327 点 60 秒周期 247 点
アラーム監視 データ参照	13,203 点 疑似運転プログラム x 15 ブラウザ機能で 1 か月分 (43200 データ)を表示
Test 3	
サーバ構成	2 台(増強前)
データ収集	1 秒周期 380 点 2 秒周期 25,972 点 10 秒周期 18,327 点 60 秒周期 247 点
アラーム監視 データ参照	8,203 点 疑似運転プログラム x 15

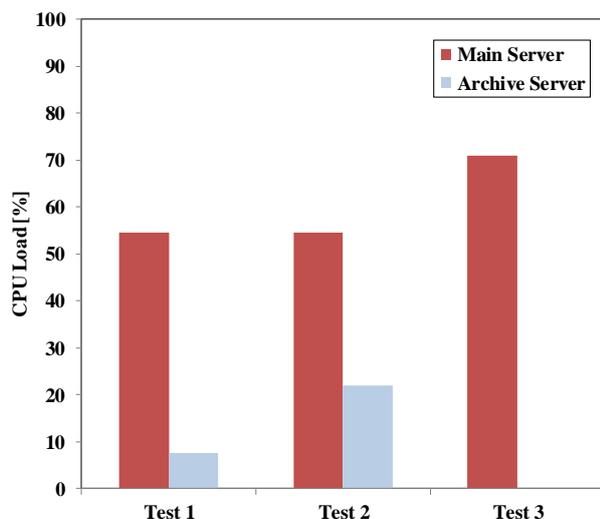


図 2：増強後の 2 台のサーバ機と従来の 1 台構成のサーバ機の CPU 負荷

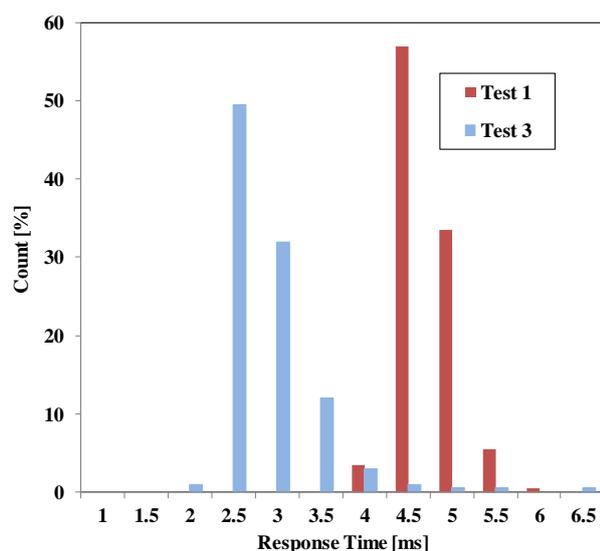


図 3：制御信号の最新値取得にかかる応答時間。疑似運転プログラムでデータベースからデータを 1 信号取得するのにかかる時間の分布。

参考文献

- [1] R. Tanaka, et al., "Inauguration of the XFEL Facility, SACLA, in SPring-8", Proceedings of ICALEPCS 2011, Grenoble, 2011, p585-588
- [2] A. Yamashita et al., "Data Archiving and Retrieval for Spring-8 accelerator complex", Proceedings of ICALEPCS '99, Trieste, Italy, 1999, p434
- [3] T. Hirono et al., "Load Test for the Database System of SACLA", Proceedings of the 8th Annual Meeting of Particle Accelerator Society of Japan, Tsukuba, Aug 1-3, 2011
- [4] <http://www.sybase.jp/products/databasemanagement/adaptiveserverenterprise/>
- [5] <http://www.nec.co.jp/products/pcserver/ft/lineup.shtml>