

# SRBシステムの導入

飯田好美<sup>A)</sup>、佐々木節<sup>A)</sup>

<sup>A)</sup>高エネルギー加速器研究機構 計算科学センター

## 概要

KEK では多くの実験データを他の研究所と共有するために、SDSC(San Diego Supercomputer Center)で開発された SRB(Storage Resource Broker)システムの導入を検討している。SRB システムではインターネット上のマシンのリソースや、ユーザの情報、データセットなどを DB で管理し、システム全体があたかも 1 台のマシンであるかのように扱うことができる。

ここでは、SRB システムと KEK での導入試験について報告する。

## 1 はじめに

現在 KEK で行われている BELLE 実験のデータは、二重のファイアウォールによりネットワーク上セキュリティの保たれた場所に保管されている。そのため、他の外国研究施設などとデータの共有を行うことが非常に難しく、限られた一部のネットワークからしか BELLE のデータにアクセスすることはできない。また、これらの大量のデータは 1 箇所のテープライブラリに集中して保管されている。

今後これらのデータは、他の外国研究施設などと共有や、不慮の事故などによるデータの消失が無いよう異なる環境でのバックアップが必要であると考えられる。そのため、データの共有、バックアップなどを行うシステムとして SRB の導入試験を行うことになった。

## 2 SRB

SRB とはインターネット上の様々なデータリソースに接続し、複製されたデータセットにアクセスするためのインターフェイスを提供するミドルウェアである。SRB には MCAT と呼ばれる Metadata Catalog があり、ここで SRB システム内のユーザ・グループ、リソース、アクセス方法、データ情報などを管理する。リソースとして定義できるのはストレージシステムや周辺機器などのハードウェアシステムのみではなく、DBMS やファイルシステムなどのソフトウェアシステムなども含まれる。また、これらのリソースへのアクセス方法も MCAT により定義されているため、データ名や物理的な位置を気にすることなく、データの属性に基づいて SRB に登録されているデータセットやリソースにアクセスできる。

SRB は大きく分けて 4 つの側面を持つ。まず一つ目として、階層型フォルダを持ち、ファイル操作や所有権、アクセス制御などの点からファイルシステムの機能。次に、透過性のある複製、異種ストレージの接続、集合データの移動や巨大データの摂取などを行うのでデータグリッド管理システムとしての機能。それから、メタデータの連結と問い合わせ、収集レベルのメタデータ、ユーザやアプリケーションレベルのメタデータ、データとメタデータの大量摂取などデジタルライブラリとしての機能。最後に、識別子の維持とグローバルネームスペース、シングルサインオン、メタデータを通しての意味のある連結、属性を元にしたデータ検索、Web サービスアーキテクチャなど有意義な Web としての機能である。

### 3 システム構成

試験用のシステムには DELL サーバー2台と IBM サーバー1台の合計3台のマシンを使用し、図1のような構成とした。また、それぞれのマシンは表1のような仕様になっている。

これら全てに SRB2.1.0 と Globus Toolkit 2.2.4 をインストールした。また、このうち DELL マシン1台には DB2 v8.1 をインストールし MCAT-enabled SRB サーバーとし、IBM マシンには HPSS-client-Lib 4.3 をインストールし HPSS-enabled SRB サーバーにした。

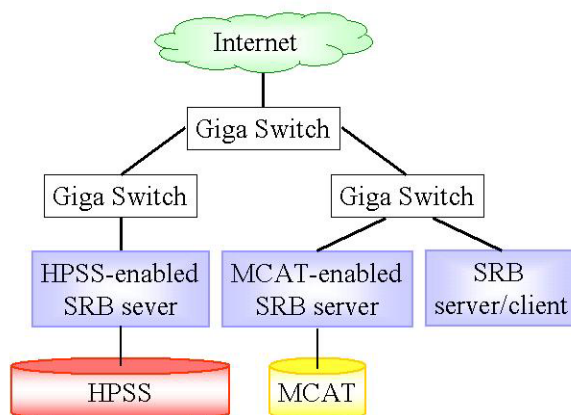


図 1. システム構成

表 1. マシンスペック

	DELL PowerEdge2650	IBM eServer xSeries 205 model 5GJ
CPU	Intel Xeon 2.8GHz × 2	Intel Pentium4 2.8GHz
Memory	512MB	512MB
Disk	36GB	40GB
NIC	Intel Pro/1000XF	Giga-Ether
OS	Linux RedHat 7.2	Linux RedHat 7.3

### 4 インストール・設定

#### 4.1 MCAT-enabled SRB server

SRB システムを構築する際、最初に MCAT-enabled SRB サーバーを立ち上げなければいけない。また、MCAT-enabled SRB サーバーを立ち上げるには、MCAT をインストールする必要がある。MCAT をインストールできる DBMS としてはフリーのソフトウェアである PostgreSQL と商用の Oracle, DB2 の 3 種類がある。今回の導入試験では最初は PostgreSQL 7.2.3 を使用して MCAT のインストールを行った。しかし PostgreSQL はフリーソフトであることから、DB へのアクセス制御などを充分に行えずセキュリティ面に不安がある、DB 自体にバックアップを取る機能が充実していないためバックアップ時に性能を落とす可能性があるなどの恐れがあった。そのため、DBMS を PostgreSQL から DB2 に変更し、MCAT のインストールを行った。

#### 4.2 HPSS-enabled SRB server

今回、HPSS-enabled SRB サーバーは Linux Redhat7.3 上に構築された。通常 HPSS の認証方法として、AIX 上では DCE 認証、Solaris 上では No\_DCE 認証が使用されている。しかし、今回構築した Linux 上では Kerberos 認証が使用されており、DCE 認証とも No\_DCE 認証とも違うため SRB を作成した SDSC ではそのような認証方式を考慮していなかった。そのため、配布されたソースをそのままコンパイルしただけでは Kerberos 認証を利用できず HPSS に接続できないため、Kerberos ライブラリーなどを認識するように合計 3 つのファイルを編集してコンパイルした。

### 4.3 ポート設定

SRB システムは KEK 外部とも通信を行うことを想定しているため、開けるポートの設定を行う必要がある。SRB で使用するポートは 2 種類あり、一つは認証を行うために使用するポート、もう一つはデータ転送などに使用するポートである。これらのポート番号の設定は最初のコンパイル時に行わなければならない、変更する際は再コンパイルが必要になる。また、認証を行うためのポートは一つで良いが、データ転送を行うためのポートはクライアントの接続数や、パラレル転送などの数によって必要なポート数は変わるため、どの程度の範囲を開けるかを考慮する必要がある。現在の設定ではデータ転送用のポートとして 200 ポートをオープンにしている。

### 4.4 認証

SRB システムでは様々な認証方式が使用できるようになっているが、今回テストを行ったのは暗号化パスワード認証と GSI(Grid Security Infrastructure)認証である。

暗号化パスワード認証はクライアントユーザーのホームディレクトリに平分のパスワードを記入したファイルを保存しておき、認証を行うときにこのパスワードが暗号化されてサーバーで認証を行うという方式である。この方式では SRB の他に特に設定するものもなく、また SRB システムにアクセスする際もパスワードなどを打つ必要がないため SRB をとりあえず使ってみたいと思う場合には適している。しかし、平分パスワードをファイルに書いて置いておくという方式はセキュリティ的には良くないため、運用を考える場合には適しているとは言えない。

GSI 認証は Globus Toolkit が使用している認証方式で、SSL(Secure Sockets Layer)がベースとなっている。この方式では暗号化された証明書ファイルを保存しておき、その証明書の DN(Distinguish Name)を MCAT に登録することで GSI 認証を行うことができるようになる。また、そのままでは SRB システムに接続する際にパスワードを打つ必要があるが、SSL によって暗号化したパスワードを作成して保存し、それを環境変数として設定することによってパスワードを打つ必要をなくすることも可能である。クライアントユーザーは Globus Toolkit のインストール、証明書の要求と発行などを行わなければならない、最初の設定では多少手間が必要となるが、実運用の際にはこちらの方式を行いたいと思う。

## 5 まとめ

SRB システムの導入により複数のマシンに接続されているストレージなどに対して、一度の認証のみで接続可能になるというのはとても魅力的である。また、SRB にアカウントを持つことによって、SRB サーバーのローカルユーザーとしてアカウントを持っていなくても、それらのストレージを利用することが可能になるということも長所である。また SRB の Srepligate を使用すると、ファイルを異なるリソースに保存し、それを自動的に相互バックアップすることができ、Sbload を使用すると、大量のファイルをまとめて一度に転送できるため、それぞれのファイルごとに転送するよりも転送効率が良い。

今回の試験は SRB2.1.0 で行ったものだが、現在は SRB3.0.1 がリリースされており、すでに SRB3.0.0 までは導入済みでそのテストを行っている。SRB2 と SRB3 の大きな違いは ZONE と呼ばれる概念である。SRB2 までは一つの SRB システムの中に一つの MCAT のみが存在し、そのみでしか SRB 内の管理はできなかった。しかし SRB3 では ZONE の中にそれぞれ一つの MCAT が存在し、その ZONE 同士がお互いに通信することができる仕様になっている。

今後はこの SRB3 を使用して、転送速度の測定や実際に外国研究所とのデータ転送などを行いたい。また、現在は DB2 にインストールした MCAT を使用してシステムを構築しているが、Oracle を使用したシステムの

構築を行い、それぞれの性能試験もしたいと考えている。また、現在 HPSS のパラレル転送ができていないため、その設定も行う予定である。

## 参考文献

- [1] “SDSC Storage Resource Broker”, SRB Homepage (<http://www.npaci.edu/DICE/SRB/>)
- [2] “The Globus Alliance”, Globus Homepage (<http://www.globus.org/>)