

IBP により生成されるメタデータの解析

荒久田 博士

Hiroshi ARAKUTA

1 はじめに

対象問題の大規模化、複雑化により、生成される測定データは増加の一途を辿っている。そういった背景より近年、データインテンシブコンピューティングが注目されている。データインテンシブコンピューティングとは、大規模、大量のデータ処理を行うコンピューティング技術である。大規模、大量のデータインテンシブコンピューティングでは、様々なリソースが広域に散在する。そのため、リソース間的高速接続、効率アクセスを実現するための Grid 技術が重要な鍵となっている。大規模、大量のデータを扱う Grid 環境は Data Grid と呼ばれ、Data Grid 環境における大規模なデータインテンシブコンピューティングの実現に向け、分散ファイルシステムや並列ファイルシステムが開発されている。

本稿では、分散ファイルシステムの一つである IBP¹⁾ について述べる。また、IBP を利用したシステムの提案、およびメタデータに関して述べる。

2 IBP を利用したシステムの提案

2.1 既存の Grid 環境における問題点

対象問題の大規模化、複雑化により、生成される計測データは増加の一途を辿り、Grid 環境における大規模データインテンシブコンピューティングへの需要が高まっている。Grid 環境を効率的に利用するため、様々な Grid ミドルウェアが開発されているが、それらは効率よく Grid 環境を利用することを目的としており、ファイルアクセスや利用に関しては考慮されていない。一方、ユーザが容易にファイルアクセスをするための手段として分散ファイルシステムがあげられる。しかし一般的に利用されている分散ファイルシステムを利用したさいには次のような問題がある。

- ローカルストレージ容量の限界
ローカルホストのストレージ容量が Grid 環境で提供されるストレージ容量と比べて小さい場合、生成される計測データ全てをローカルホストへ保存することはできない。
- ファイルアクセス
リモートストレージ環境に生成される大規模、大量のファイルから必要なものをネットワーク越しに探索、処理することは困難である。

- サーバー間のファイル転送

サーバー間が連携されていない場合、あるサーバーから異なるサーバーへのファイル転送を行うさいには別のサーバー、もしくはローカルホストを介さなければならない。

これらの問題は、Grid 環境の拡大、および生成されるデータやファイルの増大に伴って顕著に現れる。そこで、これらの問題点を考慮したシステムの提案を行う。

2.2 提案システムの目的

提案システムでは、メタデータを用いてファイル処理を行う。ユーザはファイルやデータのメタデータのみをローカルホストで管理する。実データのファイルサイズは大きなものになると数 GB から数 PB となるため、ユーザがこれらを管理することは非常に困難となる。しかしメタデータのファイルサイズは数 KB であるため、ユーザの負担は少なくなる。つまり、メタデータのみをユーザが管理することにより、Grid 環境において生成される大規模、大量のファイルやデータをローカルホストの乏しいストレージ環境で管理することが可能となる。またメタデータをローカルで管理するため、リモートに存在するファイルの探索をローカルで行うことができる。これにより、必要でないファイルやデータを全てローカルホストへ転送することなく、必要なファイルのみを処理することが可能となる。さらに、メタデータの転送により、広域に散在するサーバー間でのファイル連携を実現する。

本システムの実装には IBP を用いる。また、実装にあたり IBP において生成されるメタデータについて理解する必要がある。3 章で IBP について述べ、4 章において IBP 利用のさいに生成されるメタデータについて述べる。

3 IBP

Internet Backplane Protocol (以下 IBP) は、Tennessee 大学で開発された遠隔のストレージを管理、使用するためのミドルウェアである。IBP は大規模な分散システムおよび分散アプリケーションを支えるストレージシステムとして開発された。現在、世界中で約 450 の IBP サーバが稼働しており、ストレージの総容量は 30TB となっている。

3.1 IBP の構成要素

IBP 自体は物理層の上位に位置するプロトコルであり、アプリケーションから IBP を使用するには LoRS²⁾ と呼ばれる API を介する。また IBP を利用するには、LoRS だけではなく LBone, exNode を利用する。IBP, LBone, exNode, LoRS の関係を Fig. 1 に示す。

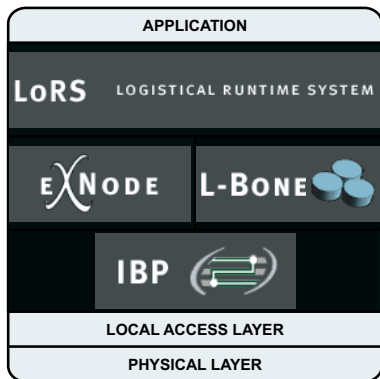


Fig. 1 IBP の概要

- IBP
IBP において利用する物理ディスクの管理を行う。要求される資源の割り当て、解放を行う。
- LBone
分散している IBP サーバのディスク容量やサーバの位置、ネットワークの状態などの情報を管理し、上位のレイヤに必要な情報を提供する。
- exNode
IBP を利用したさいに生成されるメタデータを exNode と呼ぶ。ファイルの送受信、更新は exNode を介してデータにアクセスする。
- LoRS
LoRS は、実際にデータへアクセスするさいに利用する API である。LBone と通信を行い、ネットワーク的な距離が近い IBP サーバを探し出し、IBP サーバとデータの送受信、更新を行う。

3.2 IBP の特徴

NFS や AFS など一般的な分散ファイルシステムと異なり、IBP は以下のような特徴を有する。

- メタデータのローカル管理
一般的な分散ファイルシステムでは、リモートもしくはローカルのストレージ環境に実データとメタデータを保存している。しかし IBP では、実データはリモートのストレージで管理を行い、メタデータはローカルで管理を行う。そのため、リモートのストレージに存在するデータの検索をローカルで行うことができる。

- スレッド分割によるデータの管理
IBP では、一つのデータを複数のスレッドに分割、複製を行いストレージへの保存を行う。ネットワーク的な距離の近いストレージからスレッド単位でデータをダウンロードするため、特定のネットワークにかかる負荷の低減を可能とする。また、ファイル転送時間の縮小を可能とする。

4 IBP において生成されるメタデータ

IBP においてファイルの転送を行い、生成されるメタデータについて調査を行った。生成されるメタデータを Fig. 2 に示す。Fig. 2 より、メタデータは XML で記述されており、ファイル名や API のバージョン、ファイル転送のさいに分割された全てのスレッドの情報などが記述されていることが分かる。スレッドの情報として、分割されたスレッドのファイルサイズや保存されている IBP サーバに関する情報、スレッドを扱うさいに認証を行うための情報が記述されている。

```
<?xml version="1.0"?>
<exnode xmlns:exnode="http://loci.cs.utk.edu/exnode">
  <exnode:metadata name="Version" type="string">3.0</exnode:metadata>
  <exnode:metadata name="filename" type="string">LoRS-Sample</exnode:metadata>
  <exnode:metadata name="lorsversion" type="double">0.810000</exnode:metadata>
  <exnode:mapping>
    <exnode:metadata name="alloc_length" type="integer">39</exnode:metadata>
    <exnode:metadata name="alloc_offset" type="integer">0</exnode:metadata>
    <exnode:metadata name="e2e_blocksize" type="integer">524288</exnode:metadata>
    <exnode:metadata name="exnode_offset" type="integer">60</exnode:metadata>
    <exnode:metadata name="logical_length" type="integer">6</exnode:metadata>
    <exnode:function name="xor_encrypt">
      <exnode:argument name="KEY" type="string">W(~712q</exnode:argument>
      <exnode:argument name="blocksize" type="integer">524321</exnode:argument>
    </exnode:function>
    <exnode:argument name="checksum">
      <exnode:argument name="blocksize" type="integer">524321</exnode:argument>
    </exnode:argument>
  </exnode:mapping>
  <exnode:read>ibp://hostname:port/key/WRMKey/READ</exnode:read>
  <exnode:write>ibp://hostname:port/key/WRMKey/WRITE</exnode:write>
  <exnode:manage>ibp://hostname:port/key/WRMKey/MANAGE</exnode:manage>
  </exnode:mapping>
  .
</exnode>
```

Fig. 2 IBP により生成されるメタデータの一部

5 まとめ

本報告では、現在開発されている Grid ミドルウェアでは考慮されていないファイル処理を行うためのシステムを提案した。提案システムではメタデータを利用する。そこでまず IBP 利用により生成されるメタデータの調査を行った。しかし本報告においてメタデータの全てが分かったわけではない。不明部分に関しては、現在引き続き検討中である。メタデータの更なる理解と、提案システムの構築を今後の課題とする。

参考文献

- 1) The Logistical Computing and Internetworking Laboratory. <http://loci.cs.utk.edu/>.
- 2) Tang Ming, Yeo Chai Kiat, Lim Teck Meng. Data Management for the IBP-based Data Grid.