NetNewsを使った信頼性のあるデータ通信の技法

井澤 志充

北陸先端科学技術大学院大学
情報科学研究科

NetNewsシステムは、多人数での情報のやり取りを行う手段として広く一般に用いられている。本稿ではNetNewsシステムのもつ広域性・同報性・蓄積性に着目し、分散環境におけるデータ通信路として運用した経験に基づき、その問題点や必要とされる機能などについて述べる。また、本稿で提案するシステムを、1998年1月17,18日にWIDEプロジェクトによって行われた第3回インターネット災害訓練に用いられたIAAシステムの一部として用いられた方法と、その結果について述べる。

A technics of reliable data transportation using NetNews

Yukimitsu Izawa

School of Information Science,
Japan Advanced Institute of Science and Technology, Hokuriku

NetNews system have used all over the world to provide communication environment between many peoples. We perceive NetNews system have ability to support wide area, and it can broadcast to many people in a time. So, We use NetNews system as a data communication line. In the 3rd Internet disaster drill, We have implemented a widely distributed reliable database system. We describe how we implement the system and result of it.

1　はじめに

本稿では、インターネット上での信頼性のあるデータ通信網としてNetNewsシステムを用いた手法を提案する。またWIDEプロジェクトのライフライン・ワーキンググループが1998年1月17,18日に行った「第3回インターネット災害訓練」において運用された「生存者情報データベース」IAA(I Am Alive)システムにおける、データトランスポート部として本手法で述べるNetNewsシステムベースのデータ通信網を用いた時に、必要とされた機能や問題点等について述べる。

2　システム概要

本稿で提案する手法のねらいは、以下のような要求をもつデータ通信における信頼性を向上させることである。

- 広域に分散したシステム間での同報データ通信
- 即時性は問わない
- 耐規模性に優れる
- 配送されるデータの機密保持が要求される

NetNewsシステムは、これらのうちはじめの3条件は満たしていることに加え、広く一般に用

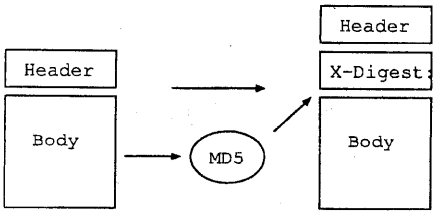


図 1: Message digesting header

いられていることを鑑みると信頼するに足る情報配送系であると思われる。また、本システムで構築するNetNewsのフィード網は、一般に開かれたものではなく、同様な目的を持ったクラスタ間にのみ張られた閉じた網である。

配送データの機密保持機構については、第2.2節を参照されたい。

2.1 データ破損検出

NetNewsシステムは、誤り訂正を行うTCPを用いた通信を行っているため、非常にデータの信頼性は高い。しかし、配送途中で何らかの事故が起こった場合には、データが破損する可能性を持っている。たとえば配送経路上のデータスプールでのディスク破損等がこれにあたる。そこで本稿では、NetNewsのアーティクルのbody部のMD5[1]を計算し、これを`X-Digest:`というヘッダにその値を付加することを提案する(図1)。

このヘッダを用いることで、データ配送を受けた時にこのヘッダが保持するbody部のMD5を再度計算することで配送されたデータの完全性をチェックすることができる。もし、配送されたデータが破損していることが判明すれば、NetNewsフィード網における上流のサイトに、再度データの転送を要求することができる。

2.2 データの機密保持機構

本稿で用いるNetNewsシステムベースのデータ配送機構は、閉じたフィード網を用いるため、第三者へデータが配送されることはない。しかし、各クラスタ間のネットワーク上を流れる記事データに対して、IPスプーフィングアタック(覗き見

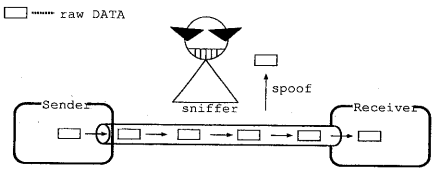


図 2: NetNews Data spoofing

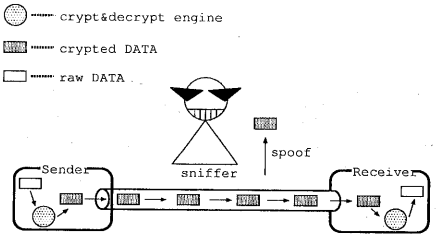


図 3: Data spoofing avoidance

攻撃)[2]を行うことで、データ内容を覗き見することができる(図2)。

そこで本手法では、DESアルゴリズムによるデータの暗号化を行い、暗号化したデータをUUencodingしてテキストデータに変換したものを配送するようにした(図3)。

このように配送するデータを暗号化することで、配送するデータが段三者により覗き見されてもその内容の機密は保持される。この場合の安全性は、使用する暗号化アルゴリズムにのみ依存する。

2.3 入力系

本システムの入力系について述べる。入力系に用いるプログラムは、NNTP[3]のクライアントとして動作する。このクライアントプログラムの動作概要を以下に示す。

1. 配送するデータを受け付ける
2. データに対してDES暗号化アルゴリズムを用いた暗号化を施す
3. 暗号化したデータをUUencodingし、テキストデータに変換する
4. 変換したテキストデータのMD5値を計算し、`X-Digest`ヘッダを生成する

5. 以上の作業で生成されたヘッダおよびボディをNNTPサーバに投稿する

全ての入力系と出力系の間で、あらかじめ安全な方法でDES暗号化アルゴリズムに用いる共有鍵を共有しておく必要がある。

2.4 配送系

配送系に関して考慮すべき点は、配送が確実に行われるようにしなければならない点である。そこで、配送系に冗長性をもたせるというアプローチとして、別のリンクを用いたバックアップのリンクを用いる方法を提案する。

また、数分から数十分後に、同じデータを同じクラスタに配送するihave/sendme機構を用いたり、フィードされる側から数十分から数時間に一回の割合で配送されたデータのリストを配送する側に通知するihave/sendme control機構を用いることにより、配送の信頼性を向上させることができる。

2.4.1 フィード網のバックアップ

フィード網を補助するフィードリンクをバックアップリンクと呼ぶ。バックアップリンクを構築する際に考慮すべき点は、

- 本来のフィード網に依存しない
  つまり、バックアップリンクのみで稼働した場合にも、全てのクラスタにデータを配送できなければならない。
- 本来のフィード網とは別のリンクを用いる
  本来のフィード網と物理的に同じリンクを用いたのではバックアップの意味がないため、物理的に別のリンクを用いるべきである。

以上のような事柄を考慮してバックアップリンクを構築しなければならない。

2.5 出力系

本システムの出力系について述べる。出力系に用いるプログラムは使用目的によって違うが、どのプログラムにも共通する点は、標準入力から配送されたデータを受取り、処理する点である。このプログラムの動作概要を以下に示す。

1. 配送されてきたテキストデータを受け付ける
2. `X-Digest`ヘッダに基づき、body部のMD5値を計算する
   計算結果と`X-Digest`の値が違う場合には、配送経路においてデータが破損したと見なし、データ復旧のキューに本記事の`Message-Id`を書き出して次のデータの処理に移る
3. MD5値が一致した場合には、body部のテキストデータに対してUUdecodingを施し、暗号化されたバイナリデータを取り出す
4. 取り出されたデータに対して、DES暗号を用いた復号を行う
5. 以上の作業で取り出されたデータを次の処理系に渡す

破損したデータは、その`Message-Id`をデータ復旧プログラムが用いるキュー・ファイルに書き込まれる。

2.6 データ復旧系

データ復旧に用いられるプログラムは、出力系で用いられたプログラムが書き出した破損データの`Message-Id`リストを、その入力として用いる。データ復旧プログラムは、NNTPを用いてNetNewsフィード網における上流のサイトに、データの転送を再度要求する。要求して送られてきたデータ記事をあらためてローカルのニュースサーバに投稿する。

通常、NetNewsにおいて記事の一意性は`Message-Id`によって保証され、ニュースサーバにおける記事の管理はこの`Message-Id`を用いて行われる。したがって、再度投稿されたデータ記事は、ニュースプールに保存されている破損した以前のデータ記事を上書きする。

再度投稿されたデータ記事は、通常にフィードされてきた記事と同様に処理される。

3 IAAシステムにおける実装

阪神淡路大震災における情報網の混乱を鑑み、WIDEプロジェクトはライフラインワーキンググ

ループを設立した。ライフラインワーキンググループでは、ライフラインとしてのインターネットという観点から、いろいろな可能性を追求している。

阪神淡路大震災では、安否に関する情報として死亡者情報というネガティブなものが殆んどであった。死亡者情報は生存者と比べて圧倒的に少数である。従来の情報伝達手段では、少数を伝達する方が効率的であるため、死亡者情報を伝達する方が情報効率が良い。しかし、インターネットの様々な技術は、生存者を確認するというような大規模な情報を扱えるだけの十分なポテンシャルをもっていると考えられる。

そこでライフラインワーキンググループではその研究の一つとして、災害時における生存者情報データベースシステム (IAA システム) を提案し実装している。

3.1 IAA システム

この生存者情報データベースは、IAA システムと呼ばれている。IAA は、"I Am Alive" の略で、生存者情報を意味する。IAA システムの目的は、コンピュータとコンピュータネットワークを利用して大規模な情報収集と検索サービスを提供することである。

IAA システムの設計にあたっては、さまざまな議論がなされてきたが、IAA システムが実際の災害時に有効に働くための要件として、

- 冗長性
- 安定性
- 信頼性
- 耐規模性

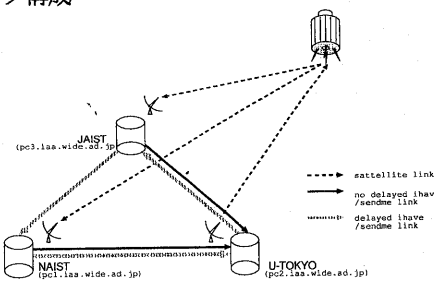
が挙げられ、トランスポート部もこの要件を満たさなければならない。

IAA システムは、

- インターフェース部
- パーサ部
- トランスポート部
- データベース部

から成り立っており、日本各地に IAA クラスタと呼ばれるトランスポート部によって結ばれたコンピュータ群によって IAA システムは構成されている。

図 4: 第 3 回インターネット災害訓練のネットワーク構成



3.2 トランスポート部概要

第 3 回インターネット災害訓練に用いられた IAA システムのネットワーク構成を、図4に示す。

IAA クラスタは、次の 3 箇所に設置した。

- 東京大学
- 北陸先端科学技術大学院大学
- 奈良先端科学技術大学院大学

各サイトは WIDE バックボーン (インターネット地上網) に接続されていると同時に、衛星回線によるマルチキャスト網でも接続された。通信衛星には、WIDE と共同研究している「(株) 日本サテライトシステムズ」の協力により JCSAT-1 号を利用し、東京大学のクラスタを打ち上げ局、その他の 2 つのクラスタを受信局とした。

また、衛星による回線をトランスポート部のメインフィードにし、地上網をバックアップリンクとした。

3.2.1 トランスポート部の実装

トランスポート部における実装について、入出力系、配送系にわけて説明する。詳細については [4] を参照されたい。

入力系 入力系には、`transport.pl` という perl のライブラリを用意した。この perl ライブラリは、*open(), put(), close()* の 3 つの関数を提供する。このライブラリを利用することで、利用者は NetNews の配送系に依存した手続きを意識することなく、本データ通信システム

図 5: transport.pl を用いた入力系

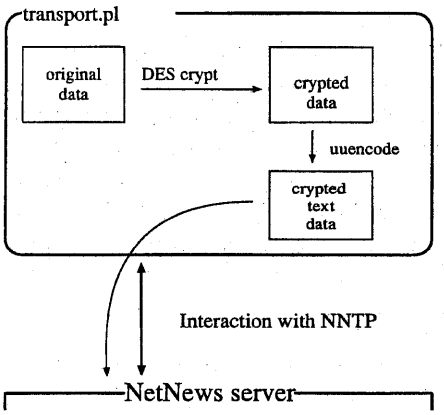
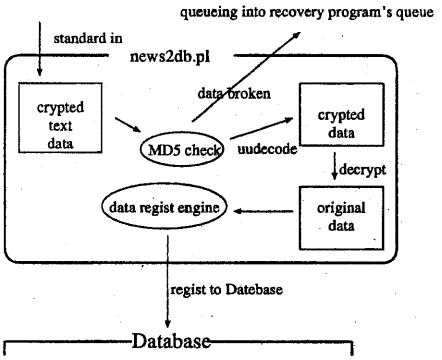


図 6: news2db.pl を用いた出力系



を利用することができる。transport.pl を用いた入力系については図5を参照されたい。

配送系 配送系にはNetNews システムとして inn-1.7.2 を用いた。

出力系 出力系には、news2db という、配送されたデータをデータベースに登録するプログラムを用いた。news2db は inn から、exploder として起動されるように設定した。この方法を用いることで、単位時間当たり大量のデータがやってきても、大量のプロセスを起動せずに処理することができる (図6)。

4 評価

4.1 transport.pl にかかった時間

transport.pl の *put()* に要した時間を計測した。*put()* の処理は次の通りである。

1. 配送するデータを DES 暗号化アルゴリズムを使って暗号化する (DES CBC モード 128bit 長)
2. UUencoding する
3. MD5 を計算する
4. NNTP を用いて、ニュースサーバに記事を投稿する

このうち、データを受け取ってから MD5 の計算終了までの時間を計測した。全ての IAA クラスタで、1 秒以内に処理されていることが分かった。

4.2 news2db にかかった時間

news2db.pl の記事をニュースプールから取り出して、データベースに登録するのに要した時間を計測した。news2db の処理は次の通りである。

1. ニュースプールの記事を open する
2. MD5 を計算する
3. UUdecoding する
4. DES 暗号化アルゴリズムを使って復号化する
5. 以上で得られたデータをデータベースに登録する

このうち、ニュースプールの記事を open してから、DES 暗号化アルゴリズムを使って復号化するまでの時間を計測した。その結果、全ての IAA クラスタで 1 秒以内に処理されていることが分かった。

以上より、本システムで提案する手法における計算コストは、実用に十分耐える程小さいことがわかった。

4.3 バックアップリンクによって配送された記事数

衛星リンクの受信局になった奈良と北陸のサイトに配送された記事のうち、本リンクを使って配送された記事数とバックアップリンクをと使って配送された記事数を以下の表にまとめた。

クラスタ	本リンク	バックアップリンク
奈良	420	18
北陸	654	3

奈良のクラスタのバックアップリンク利用率が高いのは、実験中に30分ほどクラスタマシンが不具合によって停止したためである。これらの不具合が発生した場合にも、バックアップリンク網を構築する事によって、データ配送が行われることを確認した。

以上より、本稿で提案する手法が有効である事が分かった。

5 今後の課題

今後の課題は、サーバ間で使用するリンクの状態によって、動的にフィード網を構築・変更できるしくみを考案し、データ通信のフィードリンクに容易にクラスタを追加できるようにする必要がある。

また、衛星リンクを用いて多数のクラスタ間の通信網を効率よく構築するための手法を考案する必要がある。

参考文献

[1] Ronald L. Rivest. The MD5 Message-Digest Algorithm. RFC 1321, April 1992.

[2] Computer Emergency Response Team (CERT). Ip spoofing and hijacked terminal connections. Available via anonymous ftp from info.cert.org in /pub/cert_advisories, January 1995.

[3] Brian Kantor and Phil Lapsley. Network News Transfer Protocol. RFC 977, February 1986.

[4] Yukimitsu Izawa, Shuji Ishii, Nobuhiko Tada, and Masaya Nakayama. Implementation and evaluation of widely distributed database system using satellite based multicast and netnews system for the transport mechanism. In *Proceedings of Internet Workshop '98(IWS'98)*, pp. 75–83. IEICE and ETL, 1998年3月.