

# オーバーレイ構築ツールキット Overlay Weaver

## オーバーレイ研究と応用ソフトをつなぐ

オーバーレイネットワークの研究基盤として Overlay Weaver を開発・配布している。これを用いると、計算機 1 台上で繰り返し動作試験を行いながら、少ないコード量で structured オーバーレイのアルゴリズムを実装できる。また、アルゴリズムの差し替えや計算機 1 台上での数千ノード規模のエミュレーションが可能であり、アルゴリズム間の公正な比較を行うことができる。また、実装されたアルゴリズムはそのまま実ネットワーク上で動作するため、研究成果が直接アプリケーションに結び付く。

We have been developing an overlay construction toolkit called "Overlay Weaver". Algorithm designers can implement structured overlay algorithms in just hundreds of lines of code with the toolkit and improve them rapidly by iterative testis on a single computer. The toolkit enables designers to make fair and large-scale comparisons between new and existing algorithms. Furthermore, the implemented algorithms can work on a real network in addition to the emulator. The toolkit enables algorithms developed through research to be used in applications directly.

### オーバーレイネットワーク

現在すでに、インターネット上の 300 万台を越えるコンピュータからなる分散システム(例: eDonkey2k)が出現している。今後も、人類が扱う情報の量やネットワークにつながる機器の数は爆発的に増加していくため、数百万台からなる超大規模システムを構築、管理する技術が重要となる。

数万、数百万という台数となっても性能と耐故障性を保つためには、自律的・非集中的にネットワークを構成することが不可欠である。このようなアプリケーションレベルのネットワークはオーバーレイネットワークと呼ばれ、検索やマルチキャストなどの機能を提供する。

**首藤 一幸** しゅどう かずゆき  
shudo@utagoe.com  
ウタゴエ (株)  
取締役・最高技術責任者

スレッド移送、Java JIT コンパイラ、P2P グリッド等の基盤ソフトウェア、特に分散処理ソフトウェアに取り組む。2006 年 3 月まで、産総研グリッド研究センター研究員。2006 年 4 月から、現職。

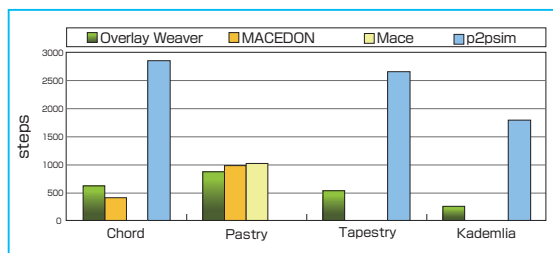


図 1 各アルゴリズムのコード量よく知られた各種の structured オーバーレイアルゴリズムの実装に要したステップ数(コメントや空行は除いた行数)。

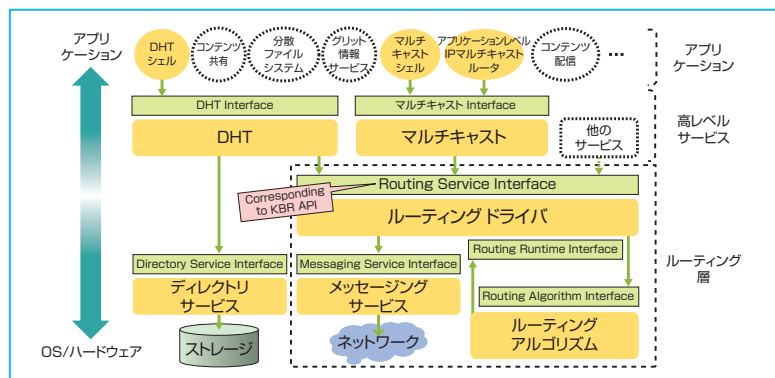


図 2 コンポーネント構成

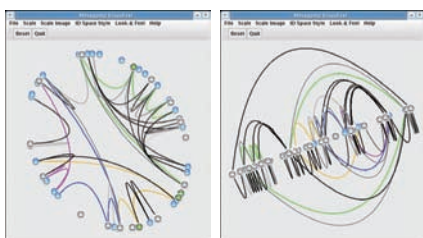


図3 メッセージング可視化ツール  
ノードと通信、また、マルチキャストのための  
配送木が描かれている。

### 研究と応用の距離

オーバーレイの動作アルゴリズムを設計・評価するには、多数のノードを想定した実験が欠かせない。大規模な実験環境、例えば1000台のPCの用意は難しいため、アルゴリズムの動作と有用性の確認はシミュレーションで行われる。その後、インターネット上の実環境で動作させるためには、実環境向けのソフトウェアを、別途開発する必要がある。われわれが開発したオーバーレイ構築ツールキット Overlay Weaver は、この、アルゴリズム研究から応用までの距離を短縮する。このツールキットによりアルゴリズムの実環境での動作と、計算機1台上での大規模エミュレーションの双方が可能となる。そのため、エミュレータを使って大規模な動作試験を繰り返し行いながらアルゴリズムを設計・実装して品質を向上させ、そのまま実環境で動作させることができる。

### ルーティングアルゴリズムの実装が容易に

structured オーバレイは、ID (例えば160ビットの数値)に対応するノードへのルーティングを基本として、その上で、分散ハッシュ表(DHT)やマルチキャスト等の機能を提供する。Overlay Weaver は、ルーティングアルゴリズム固有の処理を、多くのアルゴリズムに共通する処理から分離することに成功した。共通処理の実装はツールキット

が提供するため、アルゴリズムの新規実装が容易になり、Chord、Kademlia、Koorde、Pastry、Tapestry、といったアルゴリズムをそれぞれ数百ステップで実装できた(図1)。同時に、共通処理側の切り替えも可能となり、性質の異なる複数のルーティング手法(iterative/recurisive ルーティング)から应用到適切なものを選択できるようになった。

### ツールキットの構成

Overlay Weaver は、図2中の各コンポーネントと次の各ツールからなる。

- ・分散環境エミュレータ
- ・エミュレーションシナリオ生成器
- ・メッセージカウンタ
- ・メッセージング可視化ツール(図3)

高レベルサービスとしてDHTに加えて、マルチキャスト機能を提供している。サンプルアプリケーションとしては、DHT シェルとマルチキャストシェルを提供している。これらシェルをエミュレータと組み合わせることで、

アルゴリズムの動作試験や比較を行うことができる。

### エミュレーションと実環境での動作

4000 ノードがオーバーレイを構築し、DHT への put、get を行うというシナリオを生成して、PC1 台でエミュレートした。数千ノード規模のエミュレーション、および、アルゴリズムを差し替えての比較が可能であることを確認できた(図4)。また、実機約200台に同様の動作をさせ、実環境での動作も確認した(図5)。

### 今後の展開

今後は、Gnutella に代表される unstructured オーバレイの実装を容易にする方法や、エミュレーションに加えてシミュレーションも可能となるようなアルゴリズム記述方法を検討する。また、Overlay Weaver を用いたテストベッドの構築やアプリケーション開発の支援も進めていく。

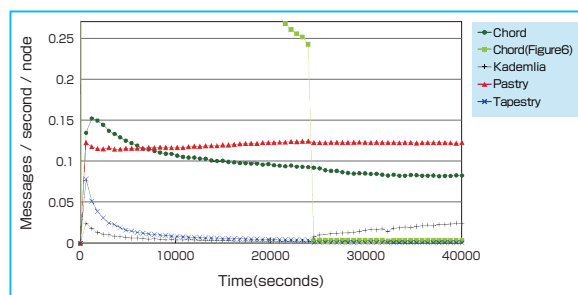


図4 4000 ノードエミュレーション時のメッセージ数

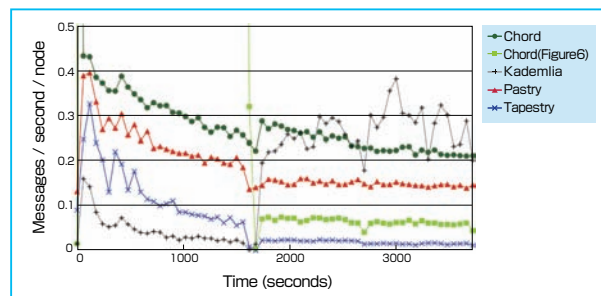


図5 実機約200台でのメッセージ数

#### 関連情報：

- Overlay Weaver の Web サイト：<http://overlayweaver.sf.net/>
- 首藤一幸、田中良夫、関口智嗣：先進的計算基盤システムシンポジウム (SACIS2006, 最優秀論文賞)