

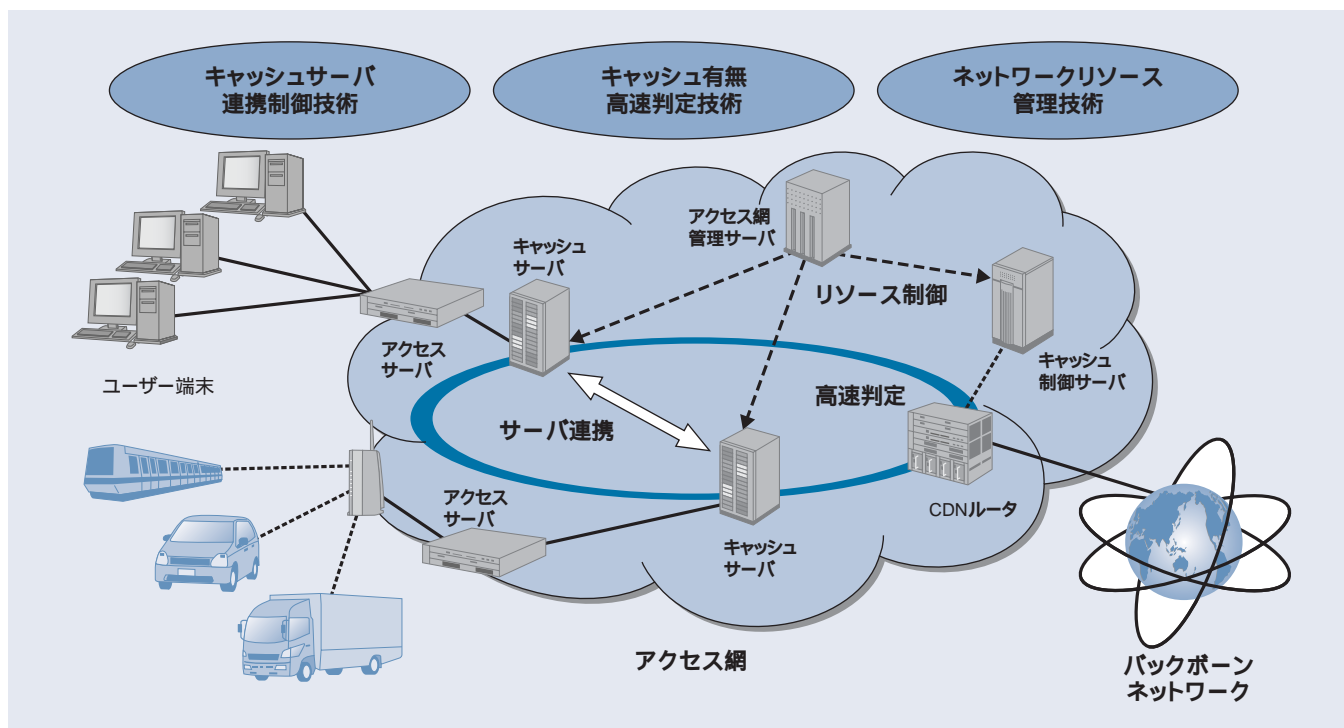
テラビット級スーパーネットワーク時代のアクセス網アーキテクチャ

Access Network Architecture in Tera-Bit Super Network Era

鈴木 敏明 Toshiaki Suzuki

平田 哲彦 Tetsuhiko Hirata

福田 靖 Yasushi Fukuda



注:略語説明 CDN(Contents Distribution Network)

アクセス網アーキテクチャの概要

全国で4,000万人規模の加入者を収容する大規模アクセス網を提供する。ユーザーが要求するコンテンツをアクセス網内のいずれかのキャッシュサーバがキャッシュしている場合、キャッシュサーバ間の連携により、要求コンテンツをキャッシュしているサーバからコンテンツを配信し、コンテンツ受信までの応答時間短縮を実現する。

ネットワークは、電気、ガス、水道のように、日常生活で必要不可欠な存在となりつつある。また、広帯域な常時接続環境の普及により、ネットワークトラフィックは増加の一途をたどり、近い将来は数十から数百テラビットに急増すると予想されている。その場合、バックボーンネットワークを流れるトラフィック量はデータ伝送処理能力を超えるため、輻輳(ふくそう)が発生し、データの廃棄やデータ受信の応答速度低下が発生することになる。

日立製作所は、やがて到来するであろうテラビット時代に向け、バックボーンネットワークへのデータ流入量を

削減し、データ受信の応答速度が速いアクセス網アーキテクチャの開発に取り組んでいる。具体的には、アクセス網内にキャッシュ機能を分散して装備する技術や、ユーザーが要求したデータのキャッシュ有無判定を高速に行うロードアーキテクチャ、アクセス網内のリソースの効率化を実現する網管理技術の開発などである。これらの技術の開発により、バックボーンネットワークでのデータ廃棄が少なく、低遅延でデータ配信を実現するアクセス網を提供することができる。

1 はじめに

2001年1月、内閣にIT(Information Technology)戦略本部が設置され、「e-Japan戦略」が策定された。この戦略で、5年以内に全国で4,000万世帯が高速もしくは超高速回線(30~100Mビット/s)に常時接続できる環

境の整備を目指すことになった。

高速回線ユーザーの増加に伴い、ネットワークのトラフィック量は年々増加の一途をたどっている。近い将来、バックボーン(基幹)ネットワークのトラフィックは、数十から数百テラビットに急増し、処理能力を超過すると予想されている。そのため、テラビット時代の新しいネットワーク

の実現が望まれている。

ここでは、独立行政法人情報通信研究機構の委託研究「テラビット級スーパーネットワークの研究開発」の一環として日立製作所が取り組んでいる、アクセス網アーキテクチャの開発について述べる。

2 テラビット時代のアクセス網

日立製作所は、きたるべきテラビット時代に向け、バックボーンネットワークへのデータ流入量を削減し、データ受信の応答速度が速いアクセス網アーキテクチャの開発に取り組んでいる。

具体的には、(1)アクセス網内のキャッシュヒット率を向上させ、バックボーンネットワークのトラフィックを削減するキャッシュサーバ連携技術、(2)要求されたコンテンツのキャッシュの有無を高速に判定し、低遅延でコンテンツを配信するノードアーキテクチャ、および(3)キャッシュサーバの負荷を平滑化し、サーバリソースを高効率に利用するネットワークリソース管理技術の三つの要素技術により、テラビット時代のアクセス網を提供する。

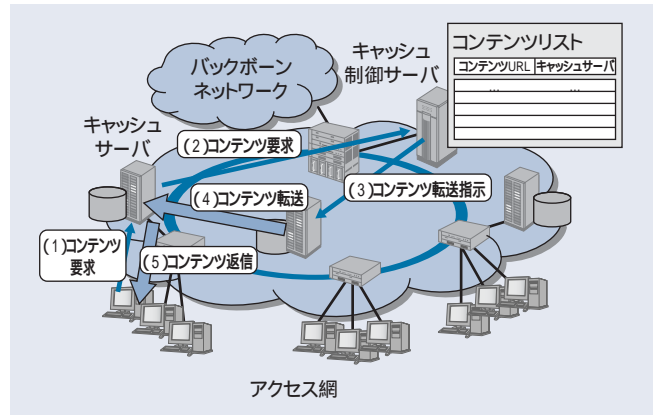
3 アクセス網の要素技術

3.1 分散キャッシュサーバ連携技術

アクセス網内のキャッシュヒット率が低下すると、バックボーンネットワークを介したコンテンツ受信量が増加するため、キャッシュのヒット率を向上させることが重要である。複数のキャッシュサーバでユーザーからの要求を分散処理する場合、1台のキャッシュサーバが保持するコンテンツ数が減少し、キャッシュサーバ当たりのキャッシュヒット率が低下する。そのため、ヒット率の向上をねらいとして、ユーザーが要求したコンテンツが、分散配置したアクセス網内のいずれかのキャッシュサーバに蓄積、保持されている場合は、キャッシュサーバどうしを連携させ、要求コンテンツを配信する。このように、分散配置した複数のキャッシュサーバを連携させてユーザーからのコンテンツ要求にตอบสนองする分散キャッシュシステム“DCCS(Distributed Cache Control System)”を実現した。

DCCSの構成と動作を図1に示す。DCCSは、キャッシュ制御サーバと、ユーザー端末近傍に複数配置するキャッシュサーバから構成する。キャッシュ制御サーバでは、アクセス網内に蓄積、保持された全キャッシュデータの識別子であるURL(Uniform Resource Locator)と、それをキャッシュしているサーバの識別子リストを保持する。

ユーザー端末では、コンテンツをキャッシュサーバに要求する(図1の(1))。要求を受信したキャッシュサーバでは、ユーザーが要求したコンテンツをキャッシュしている



注:略語説明 URL(Uniform Resource Locator)

図1 開発した分散キャッシュシステムの概要

キャッシュ制御サーバでキャッシュシステム内に記憶しているコンテンツの情報を一括管理し、キャッシュサーバ間のコンテンツ送受信を制御する。

場合、要求コンテンツをユーザー端末に返信する。コンテンツをキャッシュしていない場合は、キャッシュ制御サーバにコンテンツを要求する(同図(2))。

キャッシュサーバからコンテンツの要求を受信したキャッシュ制御サーバでは、コンテンツリストを検索する。要求されたコンテンツを保持するキャッシュサーバが存在する場合、キャッシュ制御サーバでは、そのコンテンツを保持しているキャッシュサーバに対して、コンテンツの要求元キャッシュサーバにコンテンツを転送するよう指示する(同図(3))。コンテンツの転送指示を受信したキャッシュサーバでは、要求元キャッシュサーバにコンテンツを送信する(同図(4))。コンテンツを受信した要求元キャッシュサーバでは、コンテンツをキャッシュするとともに、ユーザー端末にコンテンツを返信する(同図(5))。

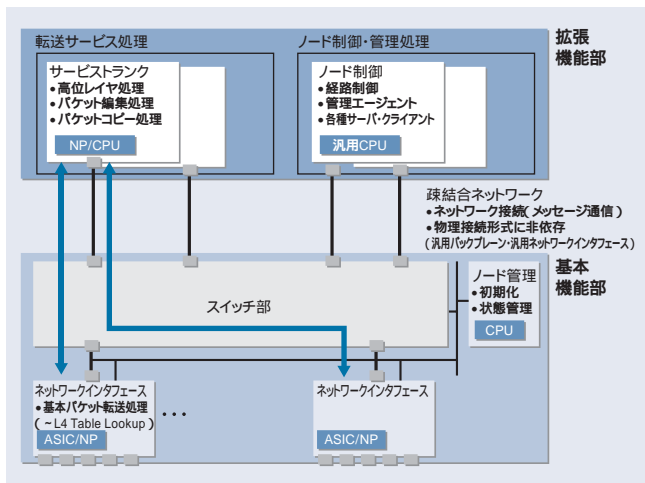
このように、DCCSでは、アクセス網内のいずれかのキャッシュサーバに要求されたコンテンツがキャッシュされている場合、アクセス網内での配信ができる。アクセス網内でのヒット率が向上するため、バックボーンネットワークを介したデータ受信の削減もできる。その結果、バックボーンネットワークでの輻輳を防止し、さらに、ユーザーへの低遅延でのコンテンツ配信を提供できる。

3.2 高機能処理のためのノードアーキテクチャ

DCCSで100万人規模のユーザー要求を処理する場合、キャッシュ制御サーバでの処理を高速化する必要がある。具体的には、アクセス網内にどのようなコンテンツがキャッシュされているかの判定を高速化する。これにより、ユーザーへのコンテンツ配信を低遅延で実現する。キャッシュの有無判定を高速に実現するノードアーキテクチャについて以下に述べる。

3.2.1 モジュール型ノードアーキテクチャ

今回開発したノードアーキテクチャを図2に示す。このノードは、基本機能部と拡張機能部から構成する。基本



注:略語説明 NP(Network Processor), CPU(Central Processing Unit), ASIC(Application Specific IC)

図2 モジュール型ノードのアーキテクチャ
モジュール型ノードでは、提供するサービス機能をモジュールとして複数搭載が可能である。

機能部では、通常のルータで行うIPv4(Internet Protocol Version 4)やIPv6ルーティング処理、および処理が必要なパケットを拡張機能部にある適切なサービスモジュールに転送する処理を行う。拡張機能部は複数のモジュールから構成し、アプリケーションレベル(レイヤ5以上)の処理、およびノードの制御・管理を行う。

このノードアーキテクチャでは、基本機能部と拡張機能部間の接続を疎結合のネットワークとし、拡張機能モジュールとしてネットワークプロセッサ、ASIC(Application Specific IC)、および汎用サーバモジュールまで幅広く利用できるようにした。また、疎結合スイッチ部にネットワークプロセッサを用いることで、拡張機能部へのパケット転送を高速化している。

3.2.2 DCCSのモジュール型ノードでの実現

前項で述べたDCCSでは、要求されたコンテンツをキャッシュしているキャッシュサーバをキャッシュ制御サーバが一括管理する。このため、ユーザーが直接接続するキャッシュサーバ(ローカルキャッシュサーバ)でのキャッシュミスが多発すると、キャッシュ制御サーバへの問い合わせが増大し、その負荷が高くなる。特に、他のキャッシュサーバが要求コンテンツを保持していない場合、キャッシュ制御サーバはバックボーンネットワークを介してオリジンサーバから当該コンテンツの取得を行い、ローカルキャッシュサーバへ転送する必要がある。そのため、このシステムをアクセス網全体といった大規模網に適用する場合、キャッシュ制御サーバの処理が障害になる。

そのため、DCCSのキャッシュ制御サーバの負荷低減を目的に、URLスイッチングの技術を応用した。一般のノードではIP(Internet Protocol)アドレスを基に転送先を決定する。一方、URLスイッチングでは、HTTP(Hypertext Transfer Protocol)リクエストに含まれる

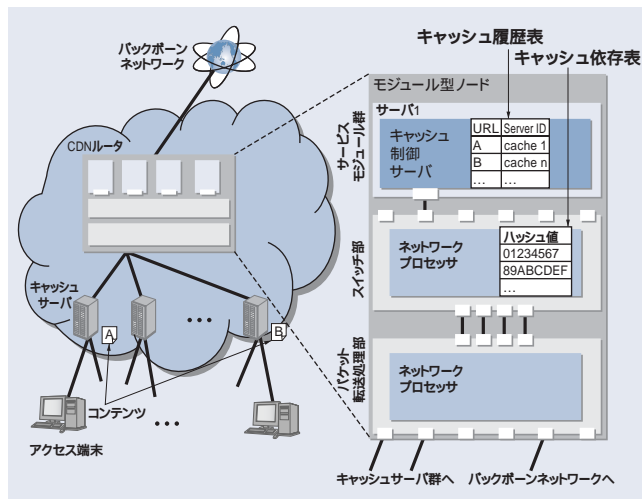


図3 開発ノードの構成
開発したノードには、キャッシュ制御サーバをモジュールとして搭載する。

URL文字列を基に転送先を選択する。開発したノードの構成を図3に示す。この構成では、キャッシュ制御サーバをモジュール型ノードの1モジュールとして実現し、ノードの疎結合スイッチ部にURLスイッチ機能を実装した。

一般に、ネットワークプロセッサで行う処理では、パケットに含まれるヘッダなど一定の位置に固定長で存在するデータを抽出することが多い。そのため、TCP(Transmission Control Protocol)ストリームの先頭に含まれるHTTPヘッダを解析するという、位置や長さが可変となるデータの扱いには大きな負担がかかる。

これを解決するため、URLの特徴を示す固定長の識別子(URLのハッシュ値)をキャッシュサーバ側で要求パケットに埋め込んでいる。スイッチ部では、要求パケットからその識別子を取り出し、テーブル検索を行うことでURL解析の複雑さを回避し、高速な検索を実現している。性能評価の結果、1秒間に2,500リクエストを処理可能であり、ソフトウェアに比較し、3倍の性能達成を確認した。これにより、ユーザーを3倍収容できるメリットが生まれる。

このように、モジュール型アーキテクチャに基づいたネットワークノードとサーバの統合により、キャッシュの有無判定を高速に行い、大規模なアクセス網への適応性を高めた。

3.3 ネットワークリソース管理技術

DCCSでは、多数のキャッシュサーバとキャッシュ制御サーバにより、100万人規模の加入者からのコンテンツ要求を処理する。しかし、コンテンツに対する要求には偏りが生じることから、キャッシュサーバでの負荷を平滑化する必要がある。そのため、ネットワークリソースを管理する運用管理サーバを導入し、リソース利用の最適化を実現する。

アクセス網内の複数のユーザーが、同一コンテンツにすでにアクセスし、かつ各ユーザーが異なるキャッシュサーバに属している場合、これらのキャッシュサーバ上に同一のキャッシュコンテンツが存在する。キャッシュしたコンテンツは同一であるため、加入者近傍のキャッシュサーバへの転送では、すでにキャッシュされたいずれかのコンテンツを配信すればよい。

転送元となるキャッシュサーバを選択するため、アクセス網に運用管理システムを導入する(図4参照)。運用管理システムは、キャッシュサーバのCPU(Central Processing Unit)負荷、経路上の利用可能帯域幅、加入者の接続状態を収集する。また、DCCSの各サーバとノードの管理情報を定期的に収集する。DCCSと運用管理システムとを連携し、動的なリソース制御を実現する。

キャッシュ制御サーバでは、複数のキャッシュサーバがキャッシュの転送元となりうる場合、検索の結果得られた複数の候補サーバのアドレスを、運用管理システムへ通知する。運用管理システムでは、キャッシュ制御サーバから通知されたアドレスをキーとしてキャッシュサーバのCPU負荷情報を検索し、負荷が最小のキャッシュサーバのアドレスをキャッシュ制御サーバに返信する。キャッシュ制御サーバでは、この返信に基づいて転送元のキャッシュサーバを決定し、アクセス網内でのキャッシュの転送を指示する。常に最小負荷のキャッシュサーバを選択することで、複数キャッシュサーバの負荷を平滑化する。

さらに、運用管理システムでは、ノードも含めたDCCSの動的制御により、アクセス網内のキャッシュ転送の通信品質を保証する。キャッシュ制御サーバでは、転送元と転送先のサーバを決定した段階でこれらのサーバのアドレスを運用管理システムへ通知する。運用管理システムでは、この通知を受け、キャッシュ転送元と転送先のサーバ間の経路上で該当フローに優先制御や帯域保証を設定する。

サービスプロバイダーは、DCCSを用いてコンテンツ配信サービスを提供する際、運用管理システムと連携してリソースを管理することで、サーバ負荷の平滑化ができ、必要リソースを最小限に抑えられる。また、ユーザーへのコンテンツ転送時の通信品質を保証することができる。

4 おわりに

ここでは、日立製作所が提供するテラビット級スーパーネットワーク時代のアクセス網アーキテクチャについて述べた。

実現したアクセス網では、ユーザーが要求したコンテ

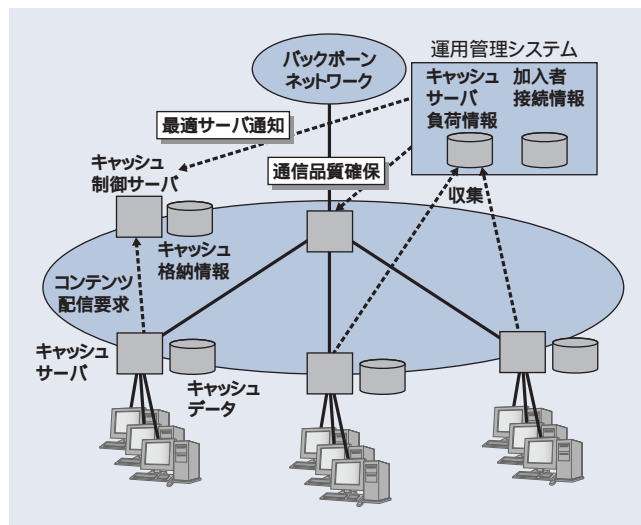


図4 運用管理システムによるリソース管理の仕組み
運用管理システムでは、キャッシュサーバの負荷状態を監視し、最小負荷のキャッシュサーバからコンテンツ配信を行い、負荷の平滑化を図る。

ンツがアクセス網内にキャッシュされている場合は、キャッシュサーバどうしを連携させ、バックボーンネットワークから要求コンテンツを受信することなくユーザーに配信する。これにより、バックボーンネットワークに流入するデータ量の削減が可能であり、輻輳によるデータ廃棄を防止できる。また、キャッシュしたコンテンツを配信することにより、低遅延でのコンテンツ配信を実現した。

日立製作所は、今後も顧客が必要とするネットワークとソリューションの実現に向けた、たゆまない技術革新と「モノづくり」に努めていく考えである。

執筆者紹介



鈴木 敏明

1992年日立製作所入社、中央研究所 ネットワークシステム研究部 所属
現在、モジュール型ノードの研究開発に従事
IEEE会員、電子情報通信学会会員
E-mail: toshiaki@crl.hitachi.co.jp



平田 哲彦

1984年日立製作所入社、中央研究所 ネットワークシステム研究部 所属
現在、情報ネットワークシステムの研究開発に従事
情報処理学会会員、電子情報通信学会会員
E-mail: t-hirata@crl.hitachi.co.jp



福田 靖

1983年日立製作所入社、情報・通信グループ ネットワークソリューション事業部 キャリアネットワークシステム部 所属
現在、ネットワークシステム関連のSE業務に従事
電子情報通信学会会員
E-mail: yasushi.fukuda.me@hitachi.com