



Advanced
Network
Architecture
Research

P2Pネットワークキング

大阪大学サイバーメディアセンター
先端ネットワーク環境研究部門
(兼 大阪大学大学院情報科学研究科)

村田正幸

e-mail: murata@cmc.osaka-u.ac.jp

http://www.anarg.jp/



Osaka University



Advanced
Network
Architecture
Research



Osaka University

ピアツーピア (Peer-to-Peer: P2P) コンピューティングモデル

- コンピュータ同士で直接情報をやりとりすることにより、コンピュータ資源や情報、各種サービスを共有する
 - 分散配置されたCPU資源、ディスク資源、情報資源の共有
 - デバイス、センサーの共有
 - ユーザ間の対等かつ直接的なコミュニケーション（情報交換、オンラインゲーム、協調作業...）
- 電子メールは人のレベルではP2P型コミュニケーションであるが、情報はメールサーバを経由

P2Pサービスの例

- ファイル共有・交換
 - Napster：情報資源の分散・共有
 - Gnutella：メタ情報資源も分散化
- CPU共有
 - インテル社：分散コンピューティング；支社のコンピュータを用いた負荷分散
 - SETI@home
 - ・ UCB 宇宙科学研究所によるSETI (Search for Extraterrestrial Intelligence)プロジェクト；電波で連絡を取ろうとしている異星人の探査
 - ・ スクリーンセーブソフトを起動させることによりデータ分析の一部を負担，結果をSETIサーバに報告
- ディスク共有：PAST（マイクロソフト社）
- グループワーク：JXTA（SUN）

M. Murata

3

情報通信インフラとしてのP2P

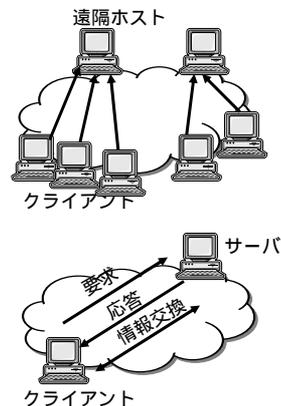
- 分散配置されたCPU資源，ディスク資源，情報資源の共有
 - グリッドコンピューティング
- デバイス，センサーの共有
 - センサーネットワーク
- ユーザ間の対等かつ直接的なコミュニケーション
 - アドホックネットワーク
- 情報通信インフラ，アプリケーションプラットフォームのP2P化，すなわち，P2P型ネットワークアーキテクチャの構築

M. Murata

4

P2Pネットワークの原型

- 広域網では閉じたピアリング
 - 第3層 (IPアドレス) を情報識別子とした「情報発見」
 - telnet, ftpによる計算資源、情報・知識の共有
 - 閉じたコミュニティを形成
- LANでは、
 - 第2層プロトコルを用いることによるオープンなピアリング
 - ・ MACアドレスに基づくARP
 - サーバクライアントモデルに基づくP2P型通信
 - ・ 各ホストはサーバにもクライアントにもなる
 - ・ 例: ファイル共有 (NFS)
 - しかし、広域ネットワーク (インターネット) ではトラフィックの増大を招くため、実現されていなかった

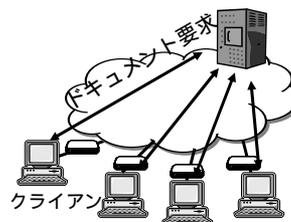


M. Murata

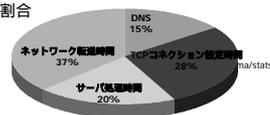
5

クライアントサーバモデルに基づく Webコンピューティングモデルの問題点

- Webシステムの発展によるサーバへの負荷集中
 - サーバ/クライアントの固定化
 - ・ クライアントはモデム経由で必要な時だけインターネットに接続
 - サーバボトルネック (CPU, ディスク)
 - ・ プロキシキャッシング
 - ・ サーバの並列化
- 新しいコンピューティングパラダイムの可能性
 - ユーザの常時接続化
 - ・ 世界中の総資源量: 100億MHz, 10Pbytes (パソコン1億台を仮定)
 - 通信容量の飛躍的な増大



Webドキュメントダウンロード時間の時間的割合



M. Murata

6

P2P型コンピューティング モデルによる解決策

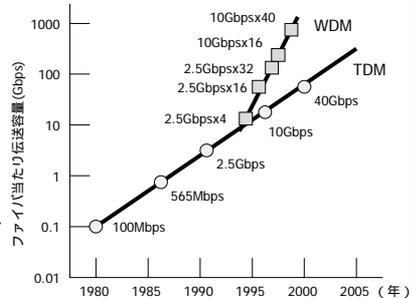
■ 自律分散型コミュニケーションによる ボトルネックの解消

- ムーアの法則：CPUのコストパフォーマンスは18ヶ月で2倍に向上（10年で100倍）
- ビルジョイの法則(?)：回線容量は9ヶ月～1年で2倍に向上（10年で1,000倍）
- 通信量が増大したとしてもネットワークのフラット化のメリットは大きい

■ サーバ主体のWebシステムからの脱却

- ロバストネス、スケーラビリティの確保
- 導入コスト、管理コストの削減
- 情報化時代における自立・分散・協調による主体的活動
- 新しいビジネスモデルの構築（サーバを介さないことによる中抜き）

M. Murata



7

P2Pのメリット

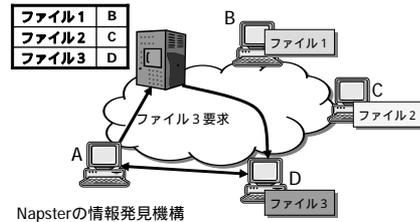
- 情報システムの特定の運営者、管理者不要
- Webサーバの巨大化
 - プローカの中抜きによるサーバ、ネットワークへの初期投資の回避
 - Scalable、growableな情報システムの構築
- 縦割り組織からプロジェクトベースのダイナミックな組織へ
 - VPNの場合
 - ・ サブネット化実現のための管理オーバーヘッド
 - ・ 多重帰属を決定する交換ノードが容易にボトルネックになる
 - ネットワークの多重化、ユーザの多重帰属（サイバー社会）への対応

M. Murata

8

情報探索・発見の方法 ハイブリッド型

- Napstar
- ピア情報（メタ情報）のみサーバで管理
- メリット
 - 探索が速い
 - 探索、情報転送の記録
- デメリット
 - 一極集中による負荷増大
 - 対故障性の脆弱さ

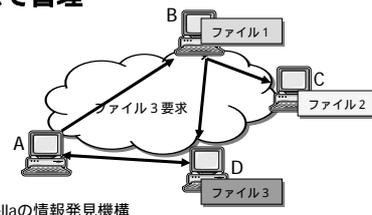


M. Murata

9

情報探索・発見の方法 ピュア型

- Gnutella
- ピア情報（メタ情報）のみサーバで管理
- メリット
 - サーバ不要
 - 対故障性
 - 匿名性
- デメリット
 - 対故障性
 - スケーラビリティ
 - ・ TTL=10の場合、すべてのノードが6ピアにブロードキャストするとすると 6^{10} メッセージの発生

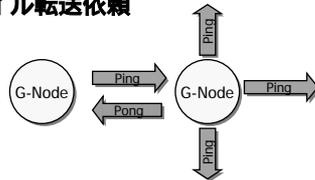
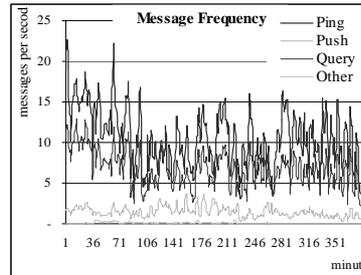


M. Murata

10

Gnutellaのトラフィック内訳

- Ping: 相手Peerを探索・発見
- Pong: Pingに対する応答
- Query: ファイル探索・発見
- QueryHit: Queryに対する応答。Peer自身がキャッシュに持つ情報を返し、その後、TTLがゼロになるまでリレー
- Push: ファイアウォール内へのファイル転送依頼



M. Murata

11

JXTA

- “open network computing platform for peer-to-peer computing” ; P2Pインフラストラクチャ
- プロトコルの規定
 - Peer Discovery Protocol
 - Peer Resolver Protocol
 - Peer Information Protocol
 - Peer Membership Protocol
 - Pipe Binding Protocol
 - Endpoint Routing Protocol
- Peer Group
 - 特長
 - Peerの集合
 - メッセージ交換はPeer Group内に限定
 - Peerは複数のGroupに所属可能
 - 初期状態ですべてのPeerが属するグループが一つ存在する
 - グループの利用方法
 - セキュア環境
 - モニタ環境

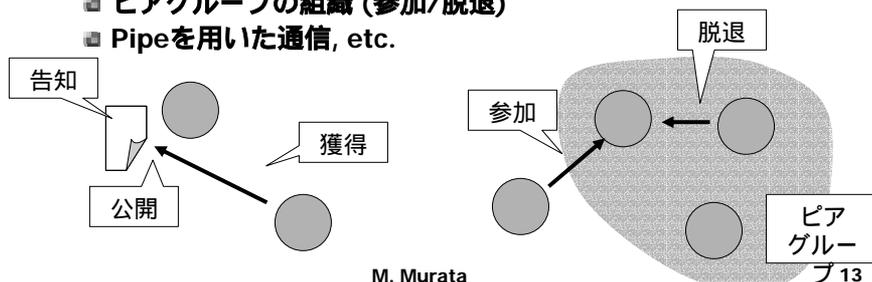
M. Murata

12

JXTAの動作例

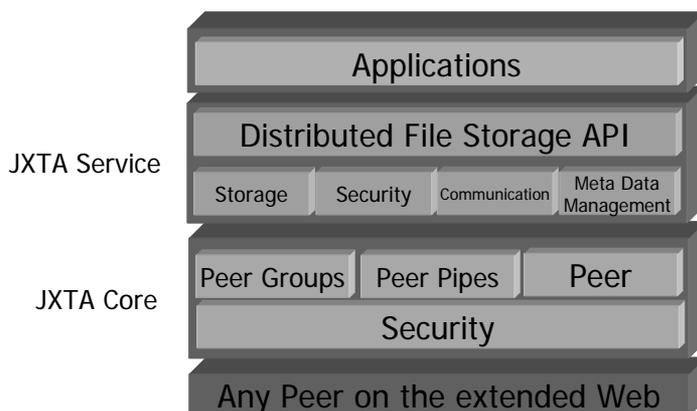
■ 動作例

- 資源の告知(advertisement)の公開(publish)
 - ・ 資源 = Peer, Contentなど
- 公開された資源(=告知)の獲得(discover)
- ピアグループの組織 (参加/脱退)
- Pipeを用いた通信, etc.



M. Murata

JXTAモデル



M. Murata

14

JXTASearch

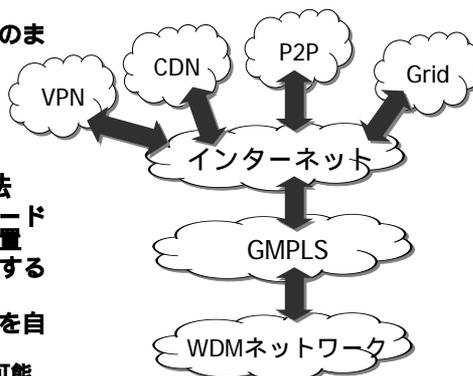
- 動的なネットワークへの対応
- 登録・探索に要する遅延の短縮
- 集中型（セキュリティ、メンバーシップ、アカウントニング）



15

オーバーレイネットワークとしてのP2P

- 論理網と物理網
 - データ転送は物理網をそのまま利用 (Gnutella)
- 効率的な論理網の構成手法
 - 論理網を構成する管理ノード（集中型、分散型）の設置
 - 物理網のQoS機能を利用する
 - ・ IntServ, DiffServ
 - 論理ノードが物理網特性を自律的に把握
 - ・ 計測（ホップ数、利用可能帯域、、、）



M. Murata

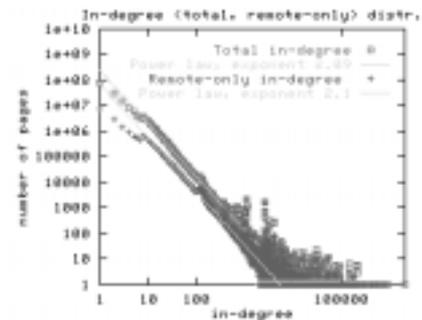
16

次世代ネットワークのキーワード

- Scalability
 - IntServへの反省
- Adaptability
 - 帯域、ノード
 - IntServ、DiffServへの反省
- Mobility
 - 端末、ノード、ネットワーク

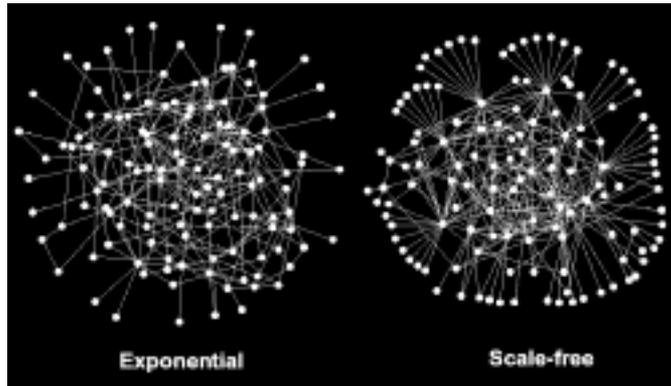
Power Law NetworkとしてのP2P

- Power Law分布
 - 事象Xの確率密度
 $P[X=x]=x^{-k}$
- 多くの事例
 - 人のネットワーク (Small World)
 - 文献引用ネットワーク
 - インターネットのASレベルの接続リンク数
 - HTMLページのリンク数



Power Lawネットワークの視覚的表現

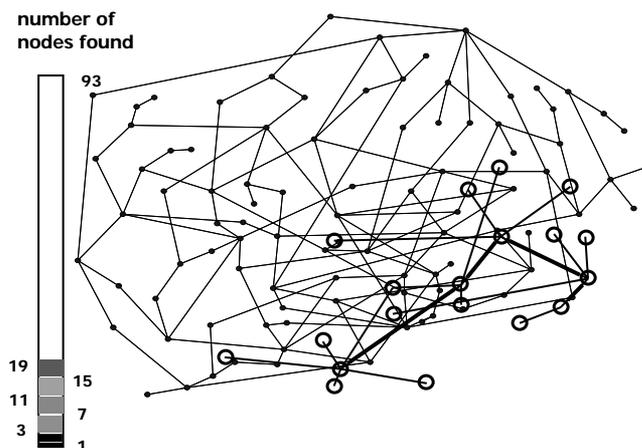
- ランダムネットワーク
- Power Lawネットワーク



19

ランダムグラフの例

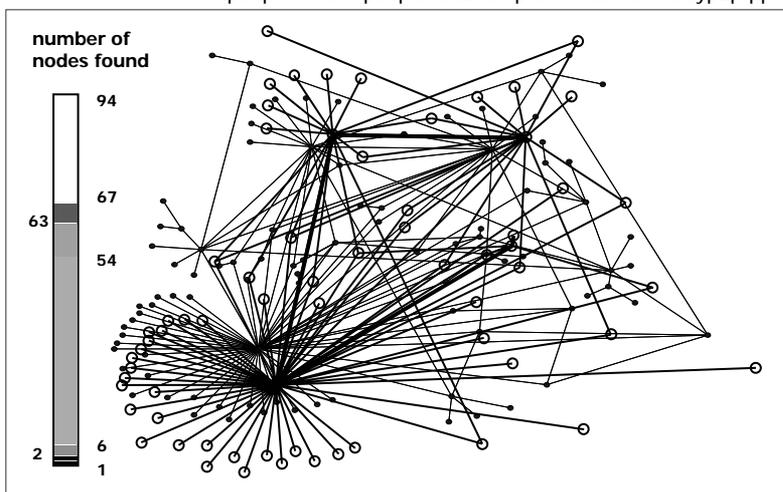
www.hpl.hp.com/shl/people/ladamic/presentations/oreillyp2p.ppt



20

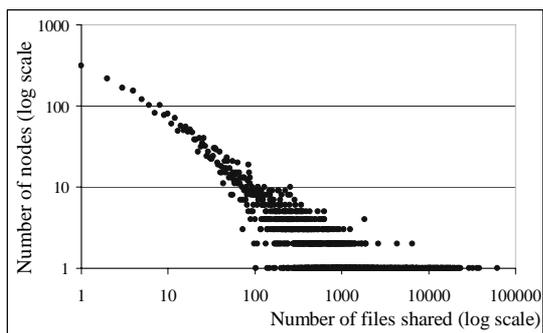
Power Law ネットワークの例

www.hpl.hp.com/shl/people/ladamic/presentations/oreillyp2p.ppt



Power Lawに従うことの意味 —ノードあたりの共有ファイル数—

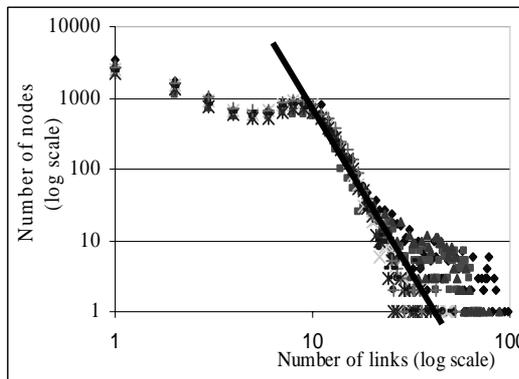
- キャッシュによる
 - トラフィック量抑制
 - Queryの遅延抑制
- TTLの適正な値の存在



Power Lawに従うことの意味 ーノードあたりのリンク数ー

■ 特徴

- ランダムなノード故障には強い
- リンクの集中するノードの故障（アタック）に弱い
- ランダム性を高める論理網の構築



M. Murata

23

グリッド計算との関係

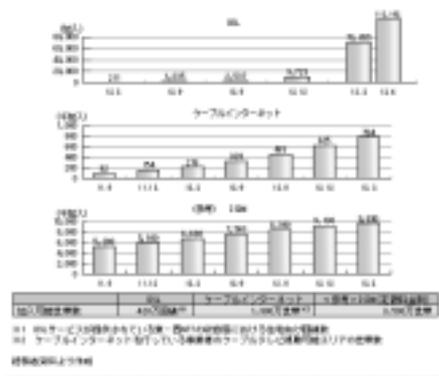
- グリッド計算（計算グリッド、データグリッド、アクセスグリッド）
 - 仮想的な組織（個人、組織、資源の動的な集合）の間で柔軟かつ安全に協調して資源共有を行う
 - 対故障性、インターネットのトラヒック特性に対する補償
- P2Pの利用
 - 資源・発見機構の利用
 - 柔軟なネットワークインフラの提供

M. Murata

24

P2Pの発展を阻害する要因？

- 情報の「質」を誰が保証するか？
 - 著作権、セキュリティ、課金
- ビジネスモデル？
 - 現状、資源を「管理」することによってビジネスが成立している
 - データセンター、サーバ、CDNの否定
 - ISPの料金制度
- 「ブロードバンドアクセス」の通信インフラ
 - 「非平衡」DSL
- モバイル化に対応できるか？



モバイル環境への適応

- 情報取得・探索から情報発信、さらには、分散コンピューティングへ
- 従来のトップダウン型手法によるネットワーク構築では
 - 階層構造によるボトルネックの発生
 - 例：Mobile IP、i-Mode、携帯電話、無線ネットワーク、...
- (サーバを経由しない) フラットなピア・ツー・ピア型通信、情報交換の実現へ

モバイル環境への適用のための課題

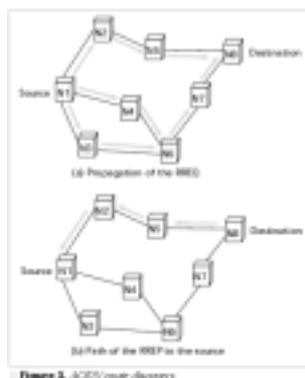
- P2P型通信においては、モバイルネットワーク利用者が直接互いに情報交換を行う
 - 情報資源を発信、伝送するためのアップリンクの通信速度が情報交換の効率に大きな影響を与える
 - 情報資源の出現、消失にともなう動的かつ急激なトラフィックの集中が生じる
 - 有線ネットワークと無線ネットワーク間で透過的に高速なデータ通信を行えるネットワークアーキテクチャ、および情報を効率よく配置、管理し、それらを配信するためのアプリケーションインフラストラクチャの構築の必要性
- 情報の受信者が提供者となり、かつ動的にその所在が変化する
 - サーバ中心型の情報資源発見機構やデータ配信機構は意味をなさなくなる
 - 多数のユーザが対等にデータをやりとりする
 - ネットワーク上に分散した情報資源の共有、および分散した情報提供者間でのデータ交換を効率的に行うための機構の必要性
- P2P型通信の応用システムとしてのウェアラブルコンピューティングにおいては、
 - いつでもどこでも、他の作業をしながらでもアプリケーションを使うことができる
 - 情報交換を継続しながら利用者が移動する
 - ウェアラブルコンピューティング環境を想定した情報の利用と管理手法の必要性

M. Murata

27

アドホックネットワーク技術との統合

- 高速データ転送のための要素技術
 - マルチプルアクセス技術
 - ・ CSMA/CA (IEEE 802.11)
 - ルーティングプロトコル
 - ・ オンデマンド (ソースルーティング型; AODV, DSR, ...)
 - ・ プロアクティブ (テーブル駆動型; DSDV, ...)
 - データ転送プロトコル
 - ・ TCPベース
 - ・ エンドツーエンド; 許されるのはモバイル側の改変まで



M. Murata

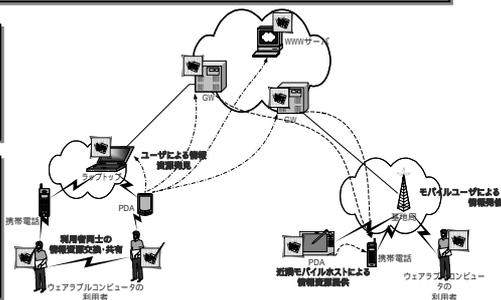
28

現状の課題

- モバイル環境におけるP2P型通信においては
 - 情報資源を発信、伝送するためのアップリンクの通信速度が情報交換の効率に大きな影響を与える
 - 情報資源の出現、消失に伴って動的かつ急激なトラヒックの集中が生じる

情報を効率よく配置、管理し、それらを配信するためのアプリケーションインフラストラクチャの構築

有線ネットワークと無線ネットワーク間で透過的かつ高速なデータ通信を行えるネットワークアーキテクチャの確立

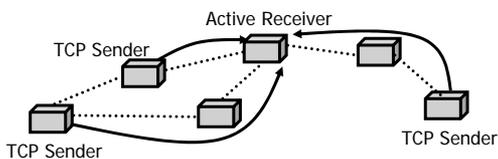


M. Murata

31

Short-Lived TCPコネクションに適したルーティング方式

- アドホックネットワークにおける遅延劣化の要因
 - ルート探索による遅延
 - TCPコネクション確立のための遅延
 - リンク切断による新たなルート探索のための遅延
- エンド間プロトコルであるTCPの改良は不適切
- 提案する手法
 - プロアクティブルート探索とオンデマンドルート探索の併用
 - リンク切断に対処するため、複数ルートを管理
 - TCPコネクション確立パケットをルート探索と同時に送信

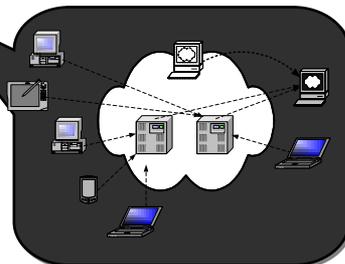
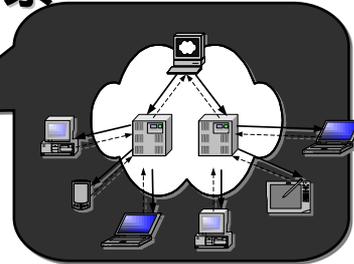


M. Murata

32

研究の背景

- プロキシサーバ、ミラーサーバなど複数サーバ配置による負荷分散
 - モバイルP2P型通信におけるホスト（情報発信源，受信者）の移動，発生，消失
 - 固定的な配置のためトラヒック，システム構成の変化に柔軟に対応できない
- モバイルホストが他のホストに代わって情報を取得，発信するなど動的な情報資源配置，管理，配信機構の必要性

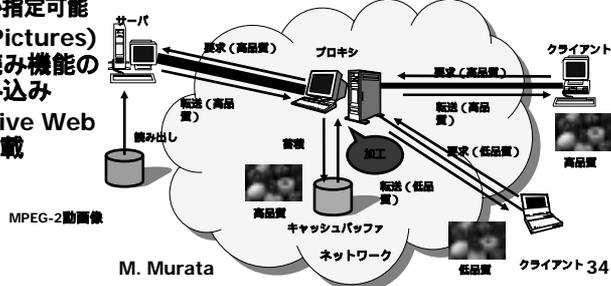


M. Murata

3

ストリーム型メディア向け P2P向けキャッシュ技術

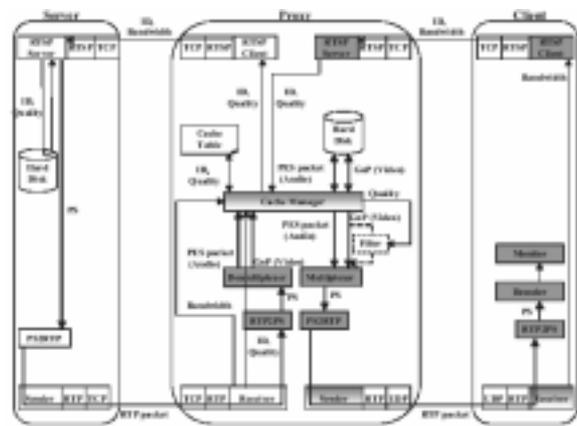
- 動画品質調整機構のプロキシへの組み込み
 - 動画データ通信のレート制御としてTFRC (TCP Friendly Rate Control) の採用
 - クライアントの要求品質のレート制御への反映
 - サーバ/プロキシ間の利用可能な帯域が小さく、要求品質を満たせるデータを時間内に取得できない場合には、クライアントが画質と実時間性のどちらを優先するのか指定可能
- GoP (Group of Pictures) を単位とした先読み機能のプロキシへの組み込み
- 日立社AWG (Active Web Gateway) への搭載



M. Murata

クライアント 34

システム構成図



M. Murata