



ネットワーク効率的利用実証研究WGの成果発表

インフラP2P連携の取り組み紹介と
ヒントサーバ実験の詳細・考察

P2Pネットワーク実験協議会 第2回シンポジウム
2009.2.19@東京大学本郷キャンパス
NTT サービスインテグレーション基盤研究所
亀井聡

P2Pとインフラを巡って
～効率的利用を目指した取り組み～

各国におけるP2Pとインフラを巡る動き

U.S.

Comcast/BitTorrent 問題における中立性問題を受け, DCIA P4P が発足. 同動きを受け IETF72 で ALTO/TINA BOF 開催, **IETF73で ALTO/LEDBAT WG 発足.**

EU

BBCの放送再配信にP2Pを利用 (**Kontiki/iPlayer**), ISPへの負担が大きく問題に. 国の支援 (**FP7**) で**P2P-Next/NAPA-WINE**といった取り組み実施.

JP

WinMX/Winnyによるインフラ逼迫・中立性懸念の動きを受けて**P2Pネットワーク実験協議会発足**. 総務省がオブザーバ.

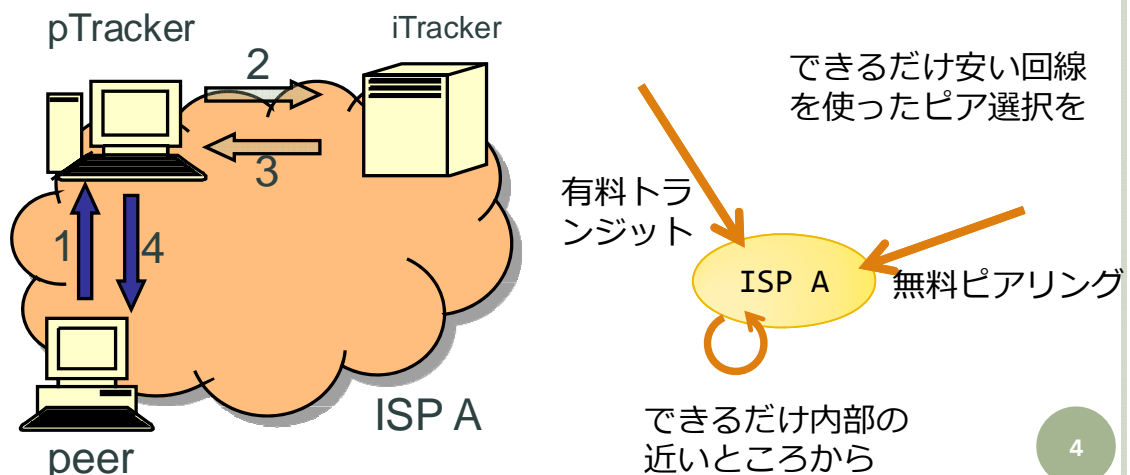
2007年後半から2008年にかけて同時多発的に発生

3

DCIAによるP4P

P4P=Proactive network Provider Participation for P2P

- ISPのトポロジ情報を持ったiTrackerをアプリケーションのTracker (Bittorrent等) に外付けすることにより, ネットワーク的に適切なピアを選択させ, ネットワークの利用効率を上げるための技術.

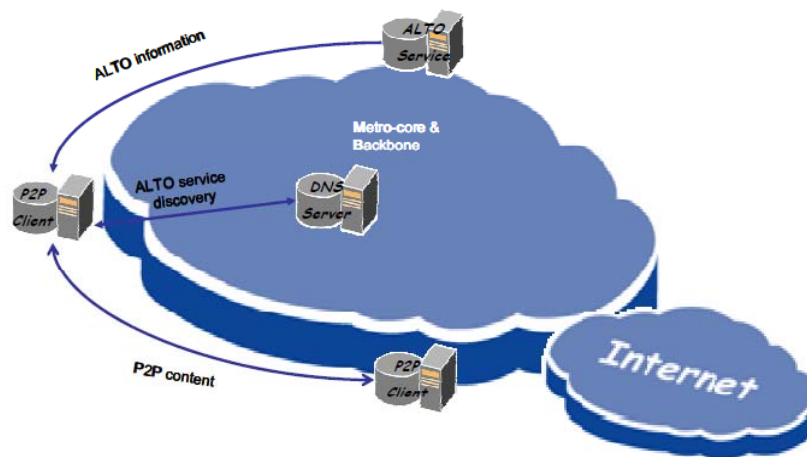


4

IETFにおけるALTO WG

ALTO BoF (Application Layer Traffic Optimization)

- DCIA P4Pにて議論されている技術をIETFへ展開しようとする動き。
- 技術的にもP4Pと同様で、ネットワーク情報を用いた最適なピア選択モデルを提案。

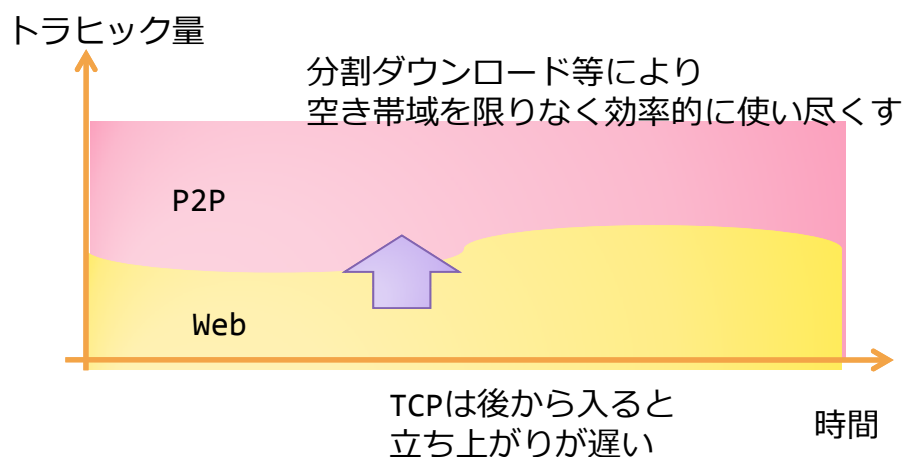


5

IETFにおけるLEDBAT WG

TANA BoF ⇒Ledbat (Low Extra Delay Background Transport) WG

- P2Pトラヒックによるボトルネックがアクセスにあるような場合を想定し、トランスポート層の挙動を変えることにより問題解決を目指す。
- 帯域を出来る限り埋めようとするTCPの挙動を見直すことや、less-than-best effort クラスの利用・複数セッションを用いた転送方式などが課題として挙げられている

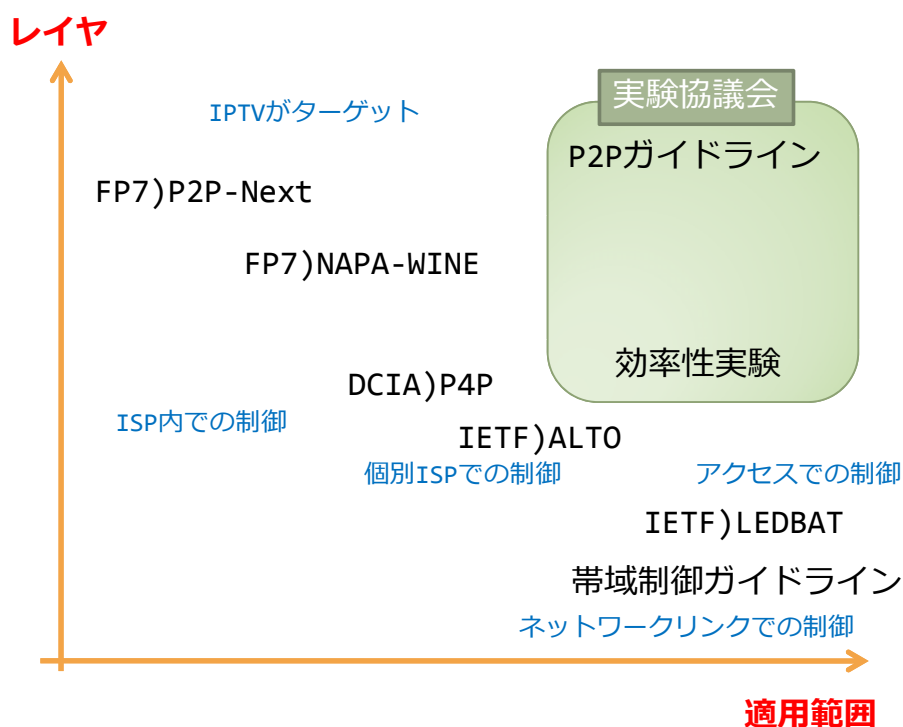


6

各種取り組みと比較した実験協議会の目的

- 最終目標であるインフラとアプリの連携はほぼ各取り組み共通.
- 効率化目標を**バックボーン**トラヒックに置いている.
 - ALTO/P4Pは個々のISPでのトラヒックが対象
 - アクセスボトルネックなアメリカの事情が反映か
- **国内ネットワーク構造**への適用時のカスタマイズ
 - 全国面ISP, アクセスISP, 地域ISP等の間での利益背反等.
- **汎用的**に使えるゆるやかな枠組みの提供を目指す.
 - BitTorrentがアメリカでは強い.
- **レイヤ間連携**の強化
 - アプリケーションガイドライン・ネットワークガイドラインとの連携, 各層の関係者の参加により実質的な連携と議論が深まる.

P2Pインフラ連携技術の俯瞰図



ヒントサーバ検証概要

実験の方針

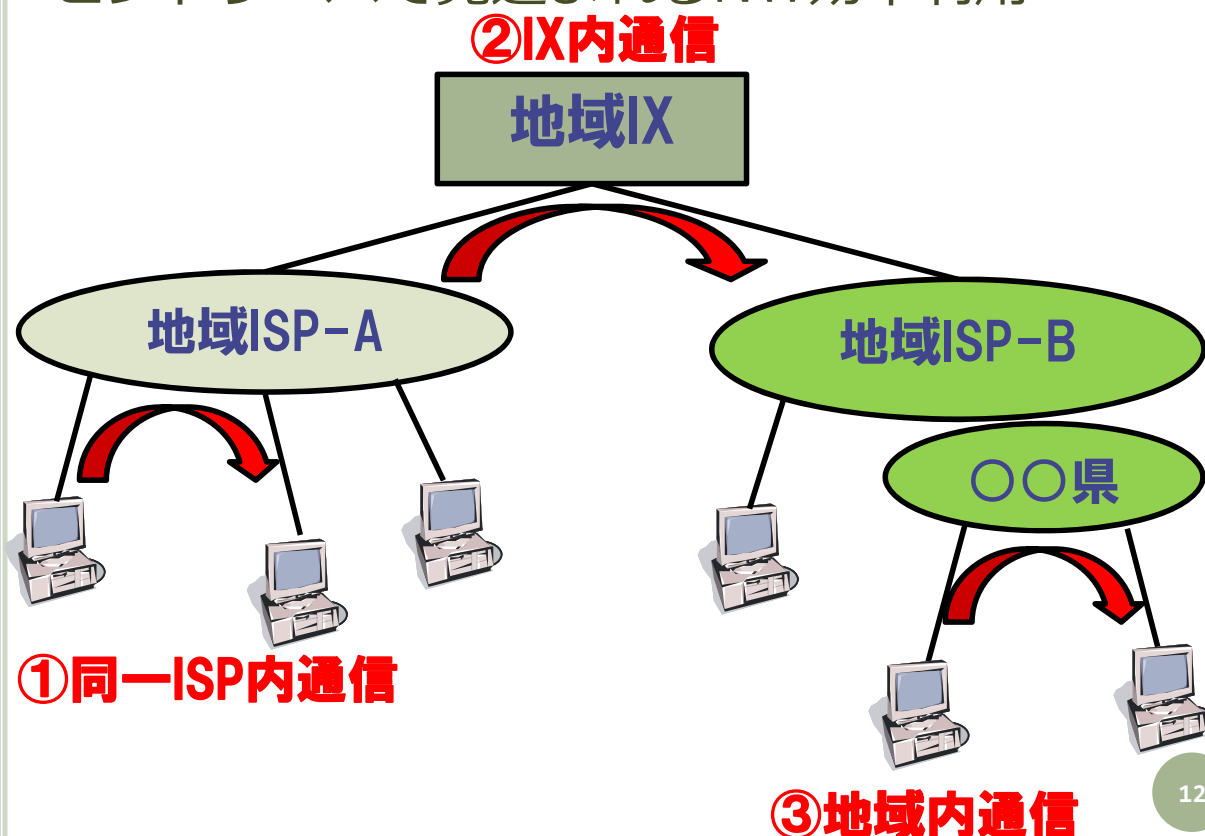
- ISP間バックボーンにおける最適化を目指す.
- 特定方式に依存しない, ピュアP2P機構でも使える汎用的な方式を.
- ゆるやかな情報共有による制御を目指したい.
- とは言え, まずは制御の仕組みをできるだけ共通で作ることが急務.
- まずはスモールスタートとして, ASを見た制御, 地域性を見た制御を可能とする仕組みを考える.
- 制御ポイントはP4P/ALTOと同じくピア選択

ネットワーク効率化を行う手段と方針

- ピアの所属ネットワーク情報を基に最適な接続先決定を「支援する」システム
 - ・ 提示されるのはあくまでもヒント情報
- 実施したいこと
 - ・ なるべくネットワーク的に近いピアを接続したい
 - 同一ISP・近接ISP・同一地域・近接地域 etc...
- ネットワーク的な距離って？
 - ・ おそらくは「コスト」が最も効くが、公表は難しい。
 - ・ まずは使えそうなデータから開始して効果を示し、徐々に粒度を上げていく。

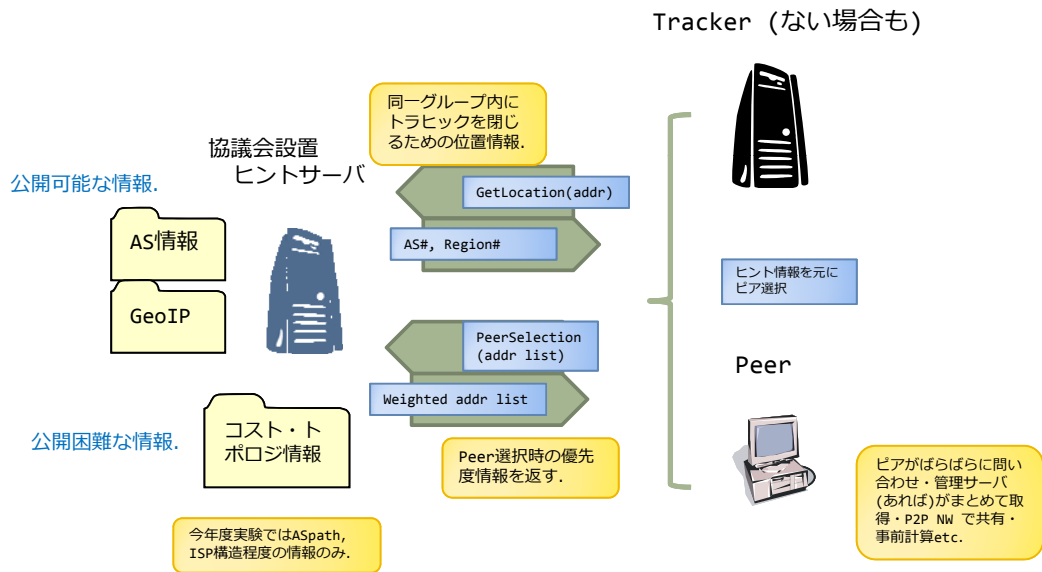
11

ヒントサーバで見込まれるNW効率利用



12

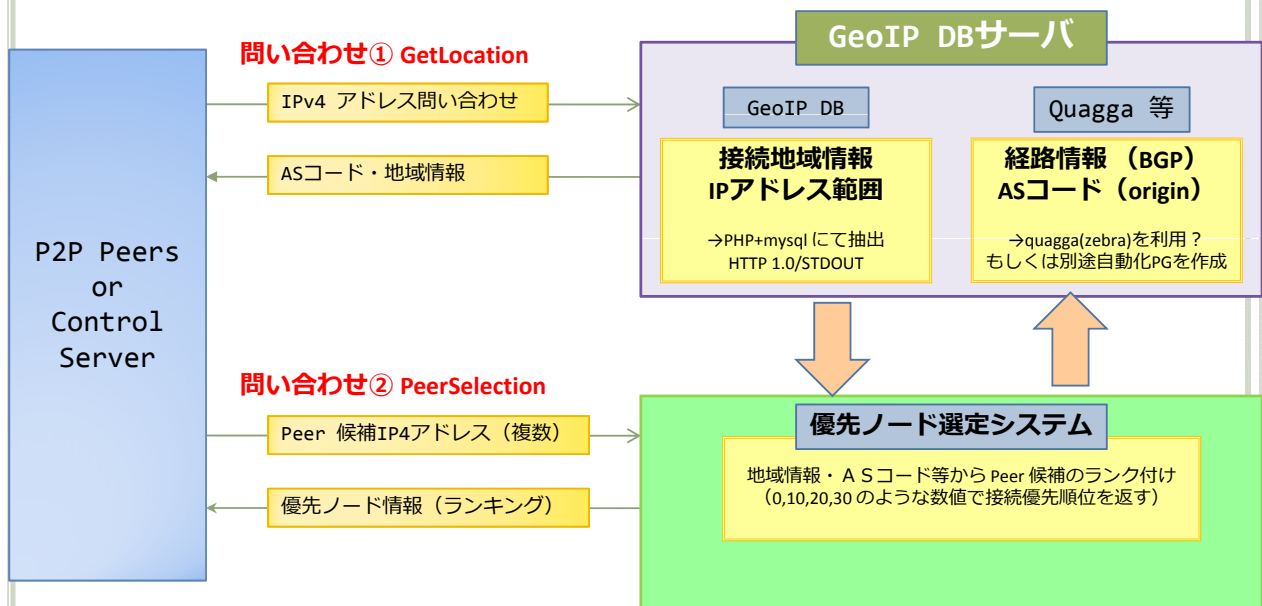
協議会実験アーキテクチャ



13

ヒントサーバ実装(今年度版)

GeoIP DB/ヒントサーバを設置することにより、P2Pノードへ近接エリア内での相互接続を促し、ネットワーク利用効率の向上を図る。



Peer Selection Hint Server

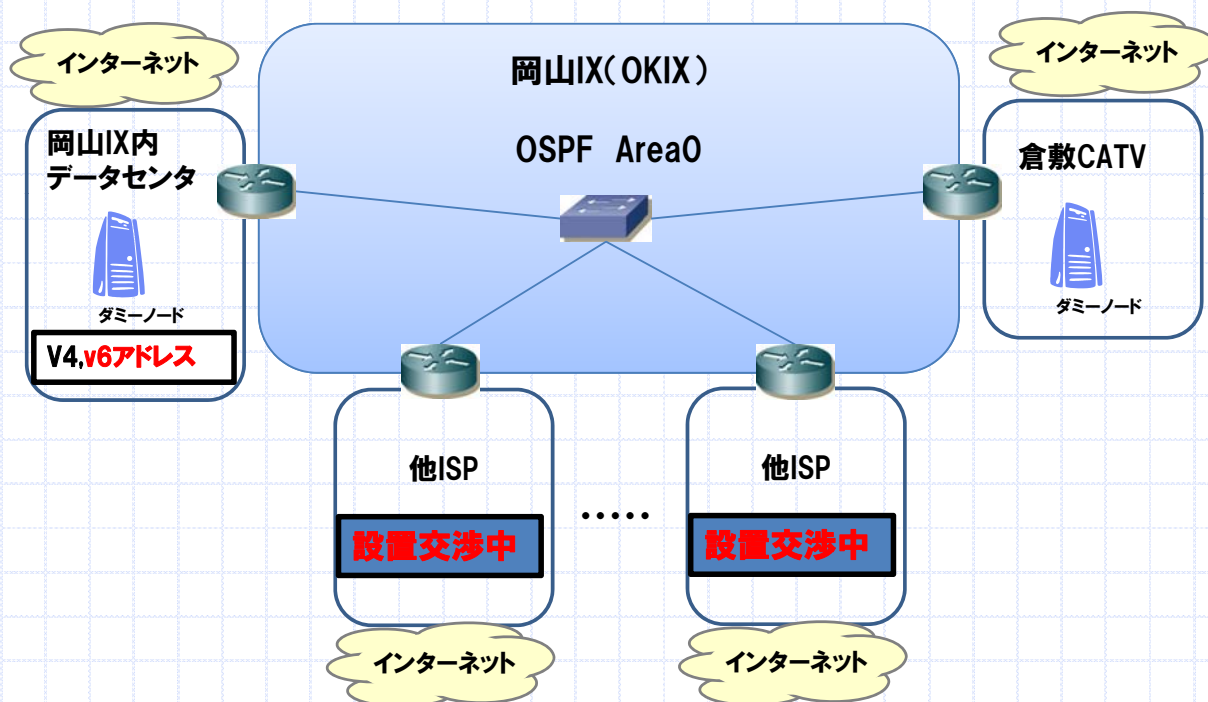
14

ヒントサーバ距離の計算方法

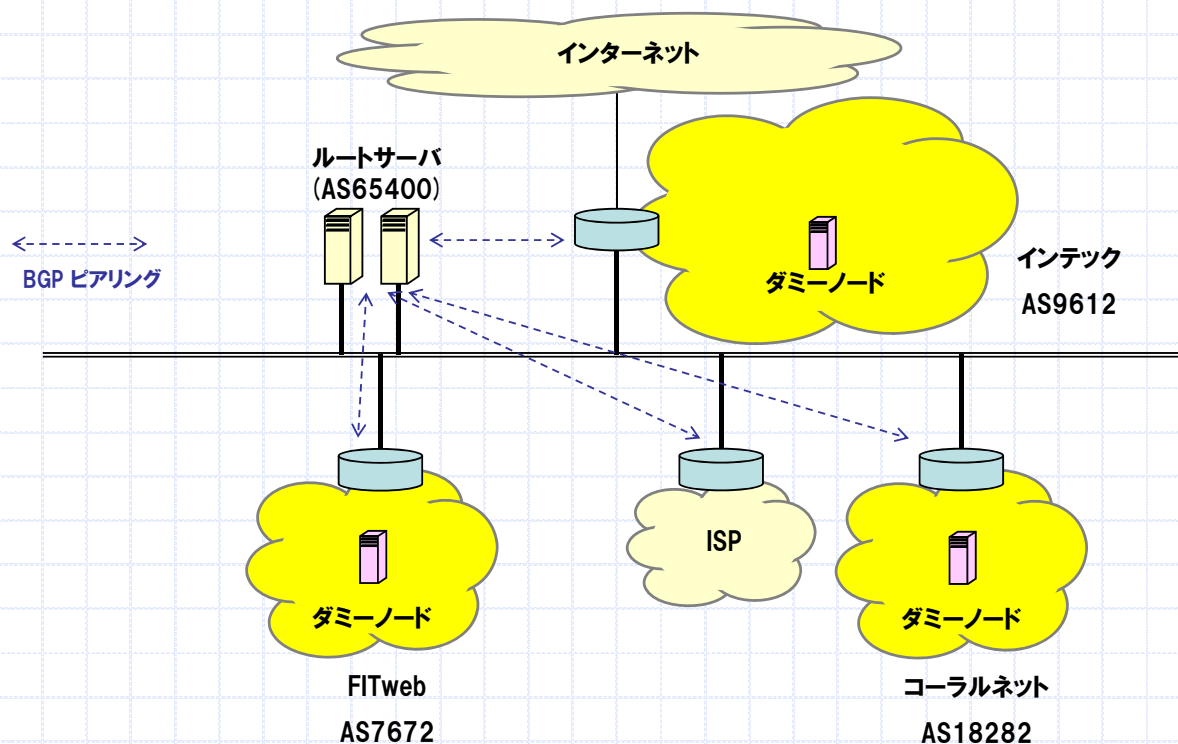
- 2ピア間の距離(ヒントサーバ距離と定義)を元にできるだけ近いピアへとリンクを張り替えることで効率化を目指す.
- ヒントサーバ距離計算には2つの尺度を用いる.
 - ASパス距離(ISP同士の距離)
 - ASパス(今年度はOCNからのみ測定)一致度(A-B-CとA-B-D : 4/6, A-B-C-DとA-B-C-E : 6/8)を計算
 - $X = \text{int}((1.0 - \text{ASパス一致度}) * 15)$
 - どちらかが不定なら15, 完全一致なら0.
 - 物理距離
 - 県庁所在地間距離を利用
 - 県間距離を昇順にソートし、ほぼ同数ずつ 1~14 を割り当て.
 - どちらかが不定なら15, 一致なら0.

主要な配信実験とその結果

ダミーノードの設置状況(岡山IX)



ダミーノードの設置状況(富山IX)



実験内容

- ・ヒントサーバとP2Pソフトの連携をまずは確認。
- ・ヒントの結果によりNWを効率的に利用した配信が行われている事を確認

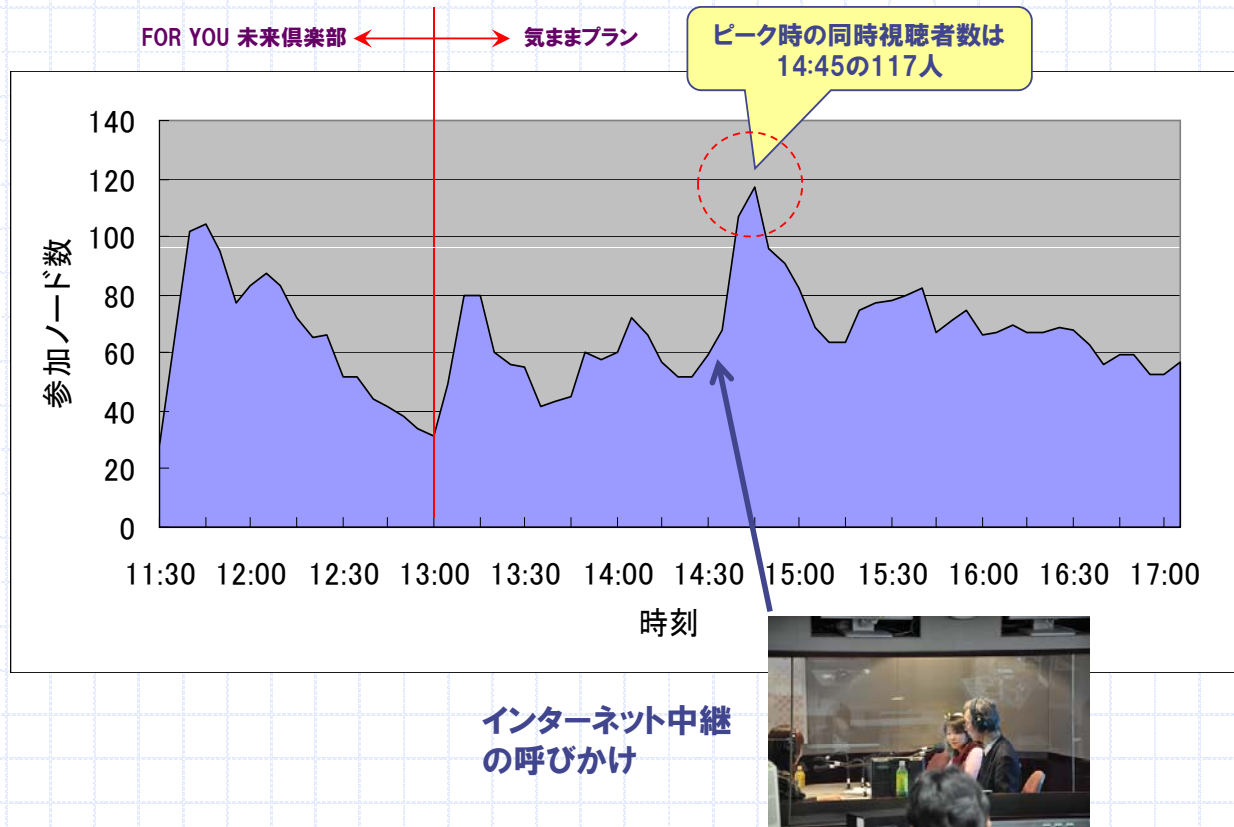
	Live①	Live②	Live③	Vod①
コンテンツ	FMとやま ラジオネット中継	H2Aいぶき 打ち上げ中継	JFL中継	京都府 学習映像配信
期間	約5H	約2H	約2H	3日間 (金～日)
ユーザ形態	全国/一般	全国/一般	全国/一般	全国/モニター
ヒントサーバ 利用の有無	有/無で実験	有	無	有/無で実験
解析対象	ISP、地域、接続 時間、制御サー バlog、ヒントサー バlog	ISP、地域、接続 時間、制御サー バlog、ヒントサー バlog	ISP、地域、接 続時間、制御 サーバlog	ISP、地域、接 続時間、ヒント サーバlog
備考				

Live中継実験①

1. FMとやまのお昼の人気番組をインターネットlive中継
2. ヒントサーバ利用の有無で2回実験
3. 富山ユーザが多くヒント利用で富山IXで閉じたトラフィックを想定



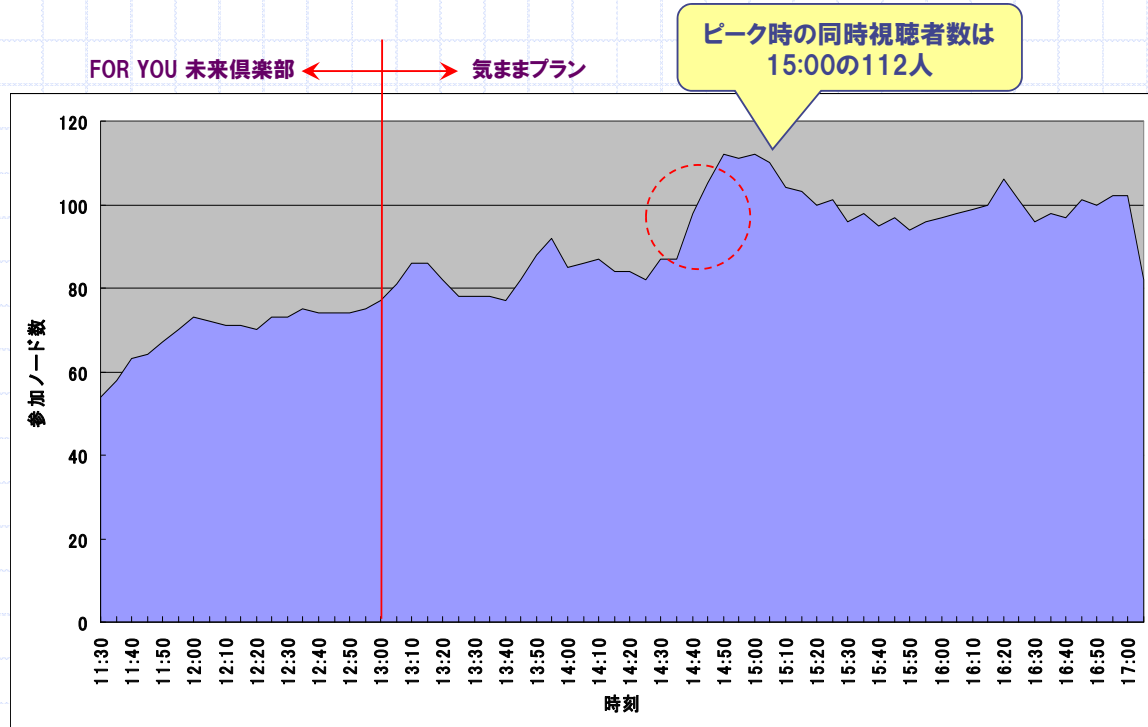
Live中継実験① ヒント無し アクセス推移



Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation, All rights reserved

21

Live中継実験① ヒント有り アクセス推移



Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation, All rights reserved

22

Live中継実験②

1. H2A15号(いぶきを搭載)の打ち上げをインターネットlive中継
2. **ヒントサーバ利用**で実験
3. 全国の広い地域のユーザが視聴



Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation, All rights reserved

23

Live中継③

1. 11/30にJFLの流通経済大学vsガイナレ鳥取を中継
2. 試合は茨木県で開催されたアウェー、鳥取県ではCATV中継
3. ヒントサーバ未利用、今後継続的に中継実施し、データ解析予定

The screenshot displays a live stream interface. On the left, a video player shows a soccer match between流通経済大学 (Ryutsu Keikai University) and ガイナーレ鳥取 (Gainare Tottori). The score is 0-2. Below the video player, there's a section for '配信チャンネル情報' (Broadcast Channel Information) with details: ID: 808612090597, Title: ガイナーレ最終戦 (Gainare Final Match), and Summary: 最終戦 (Final Match). A '配信停止' (Stop Broadcast) button is also present. On the right, a 'トポロジー図を表示' (Show Topology Diagram) button is visible, along with a 'チャンネルID: 808612090597' and buttons for '手動更新' (Manual Update) and '更新' (Update). Below these, a network topology diagram is shown, illustrating a hierarchical structure with nodes labeled 'ncom1', 'applet', and 'VAMP'.

Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation, All rights reserved

24

VOD配信実験①

1. 京都府実験によりVOD配信実験
2. **ヒントサーバ利用の有無**で2回実験
3. 全国でモニターを募集し、データ取得
・北海道～鹿児島まで広く募集



P2Pネットワーク実験 公式ホームページ
実験期間 2008年12月31日水曜 まで



実験の趣旨

インターネットによる動画配信が一般化する中、自治体においても、広報、教育（学校教育、生涯学習、e-ラーニング）、文化情報発信（デジタルアーカイブ等）をはじめとして、その活用は様々な分野に広がっています。ただ、利用者側、配信側双方の制約から小さな画面で配信せざるを得ないのが実状です。

一方、ブロードバンドゼロ地域の解消、地上デジタル放送や大画面テレビの普及に伴い、インターネットにおいても高画質配信を求めるニーズが確実に高まると思われることから、安価にかつ安全に高画質配信を実現する技術が望まれます。

今回の「**P2Pネットワーク**を活用した源氏物語関連の生涯学習情報配信**実験**」では、そのような状況を踏まえ、自治体の立場から、生涯学習等の高画質番組配信における**P2Pネットワーク**技術の有効性を検証しようとするものです。

Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation, All rights reserved

25

実験結果と考察・課題

実験における検証目標(将来のものも含め)

- ピア選択のためのヒントサーバの動作検証
- ピアがヒントサーバと連携して制御機構をうまく動作させることが可能か
- 連携動作がネットワーク利用効率に及ぼす影響
- 連携結果が利用者に及ぼす影響
- 制御に用いた尺度は適切か

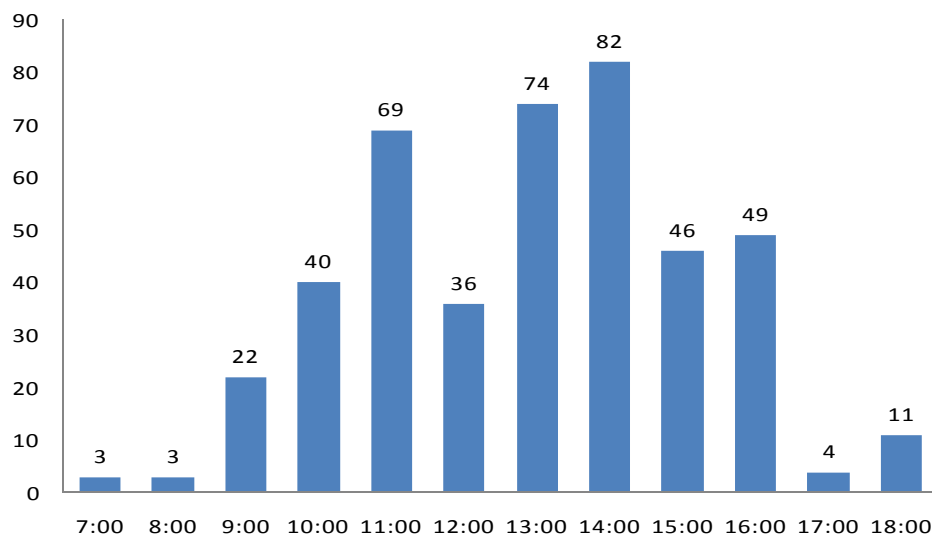
連携動作とその効果の観測

- 測定点
 - ダミーノード
 - 論理リンクの両端IPアドレス
 - 接続時間, 転送量
 - ヒントサーバ
 - リクエスト内容・結果・時間
 - 各P2P事業者の協力
 - Live 論理ツリー構造とそのIPアドレス(両端)
 - VoD 切り替えタイミングとそのIPアドレス(元・前・後)

ヒントサーバの動作状況

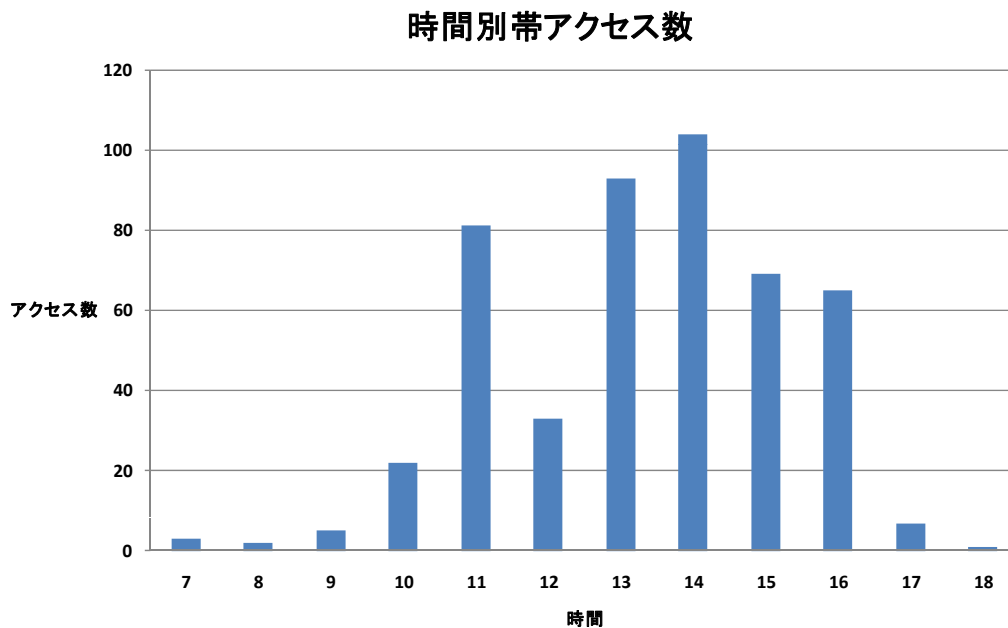
LIVE中継実験① 時間帯別アクセス数 (GETLOCATION)

● 時間帯別アクセス数



LIVE中継実験①

時間帯別アクセス数 (PEERSELECTION)



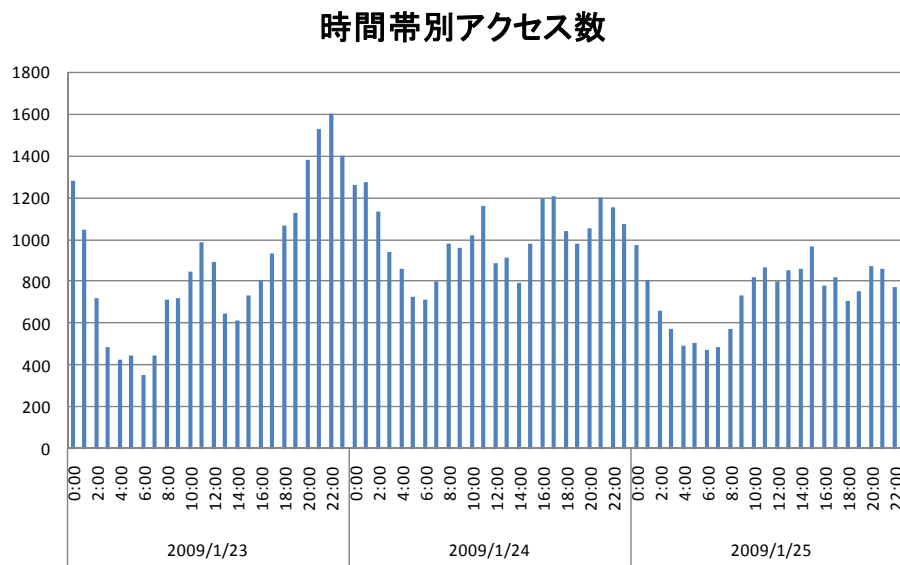
Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation, All rights reserved

31

VOD配信実験①

時間帯別アクセス数 (PEERSELECTION)

● 時間帯別アクセス数

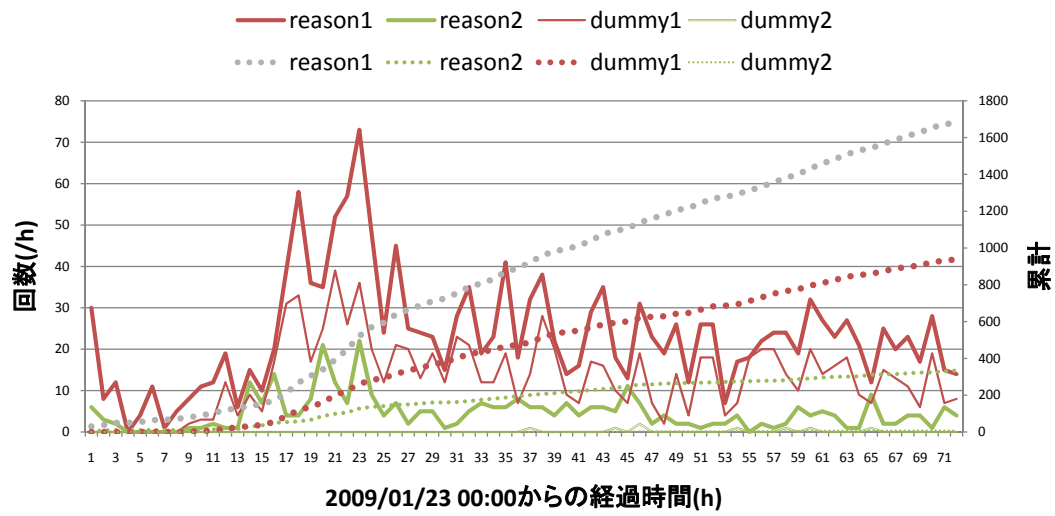


Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation, All rights reserved

32

VOD配信実験① ヒント有り

切り替えの発生頻度(理由別)



Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation, All rights reserved

33

ピアの選択状況

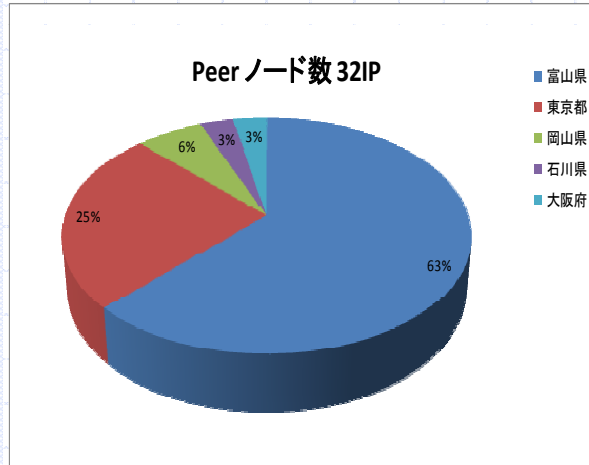
Copyright(C) 2009 NTT Service Integration Laboratories. All right reserved.

34

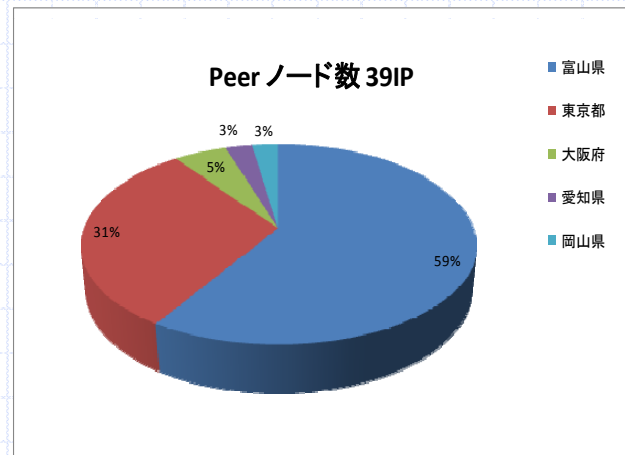
Live中継実験① ヒント無し アクセス地域

富山のユーザが富山ではなく東京へアクセスしている

富山のユーザ環境(ダミーノード)におけるアクセス地域



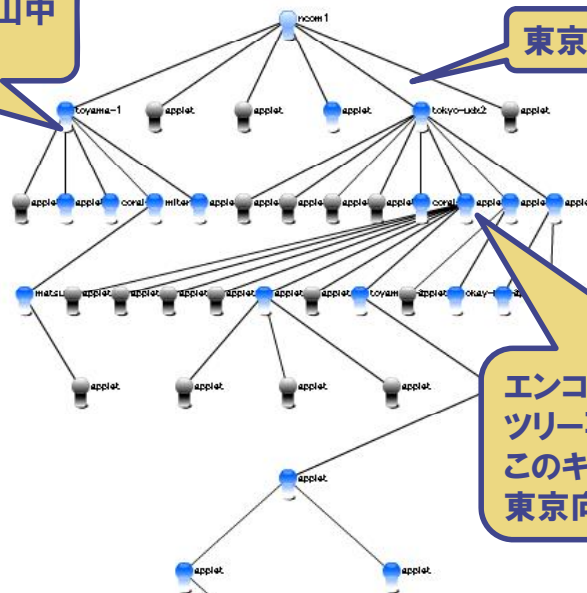
東京のユーザ環境(ダミーノード)におけるアクセス地域



Live中継実験① ヒント無し ツリー状況

できあがったツリーにAS、地域などの関連性がみられない

富山ユーザが多いのに富山中継ノードへ接続が少ない



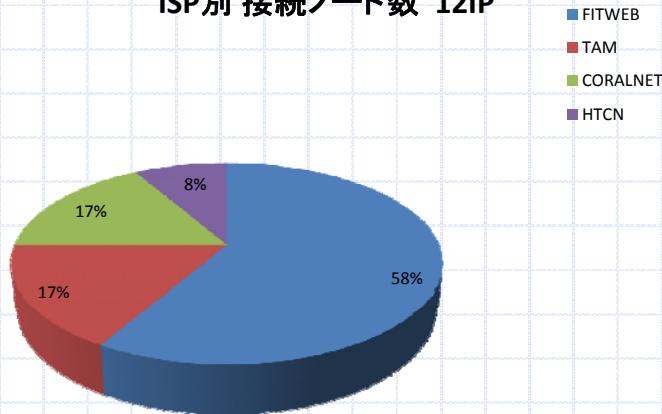
東京の中継ノードへ接続

エンコーダ再起動時に
ツリー再構築したため
このキャプチャー時には
東京向けが特に多かった

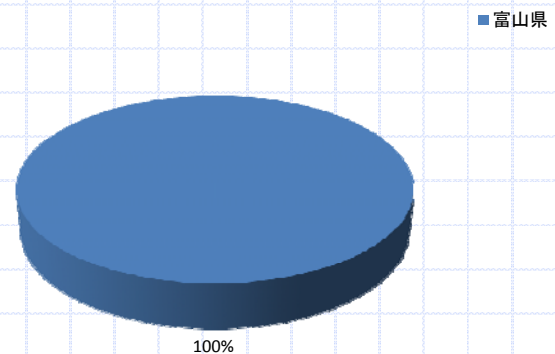
Live中継実験① ヒント有り FITweb設置

FITWEBへ設置したノードでは100%富山
顕著な効果が得られた
接続ISPは富山IXに所属

ISP別 接続ノード数 12IP



地域別 接続ノード数 12IP



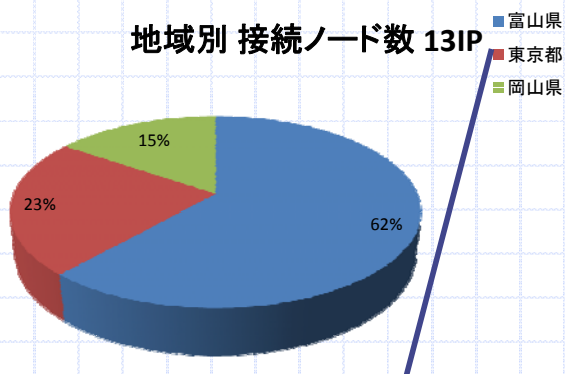
Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation. All rights reserved

37

Live中継実験① ヒント有り アクセス地域

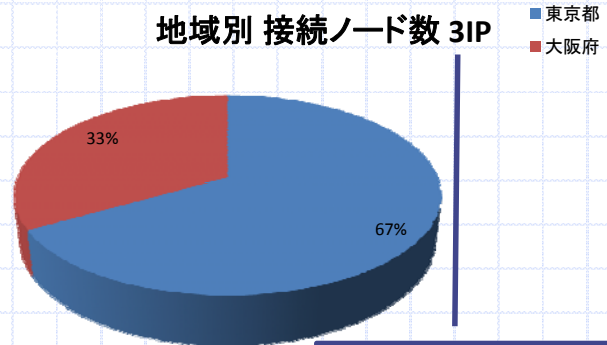
- ・富山へは富山地域、東京へは東京地域という特性がみられた
- ・多数が富山ユーザだが、そのユーザが東京に接続することはなかった

富山のユーザ環境(ダミーノード)におけるアクセス地域



実はEMOBILE
AS特性が優先か？

東京にある独自ASのユーザ環境(ダミーノード)におけるアクセスISP



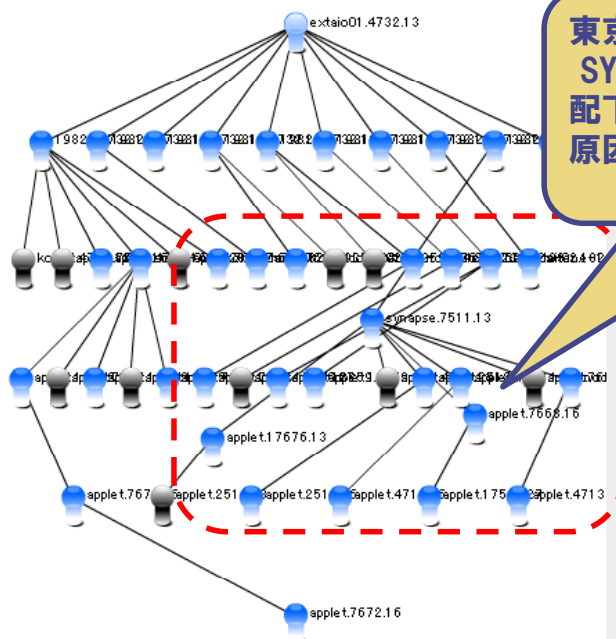
ツリーのため
ダミーノード収集IP
数が少ない場合も

Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation. All rights reserved

38

Live中継実験① ヒント有り ツリー状況

GeolPの検索結果が地域と違う地域になり、異なる地域の接続が一部発生
IPの地域情報をISPにもらう仕組みが必要。

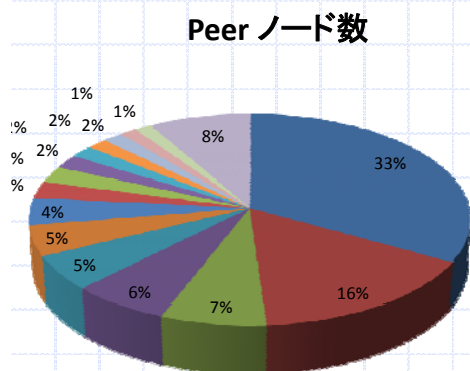


Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation. All rights reserved

41

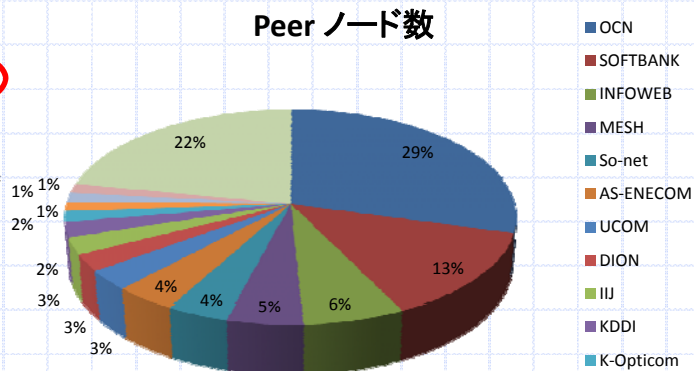
Live中継③地域、ISP分布

地元ではTV中継を視聴、その他の地域の人がインターネットを視聴？
鳥取県のユーザ以外のアクセスが比較的多い。
地元ISPのアクセスは少ない



431ユニークIPの地域内訳

■ 東京都
■ 大阪府
■ 鳥取県
■ 広島県
■ 島根県
■ 神奈川県
■ 岡山県
■ 千葉県
■ 埼玉県
■ 愛知県
■ 北海道
■ 福岡県
■ 静岡県
■ 富山県
■ 兵庫県



431ユニークIPのISP内訳

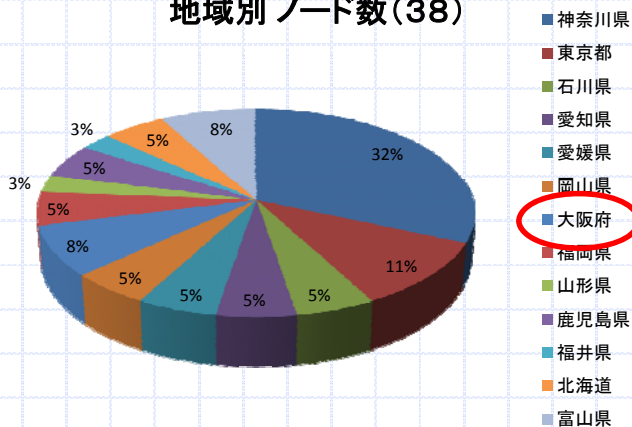
Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation. All rights reserved

VOD配信実験① アクセス地域

ヒントサーバを利用する事により一部地域で多少の地域優先状況がみられた
※データはダミーノード含む

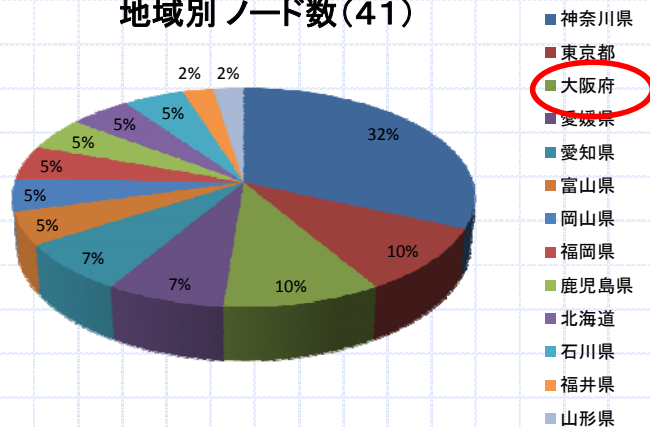
大阪設置のOCNダミーノード
(ヒント無し)

地域別 ノード数(38)



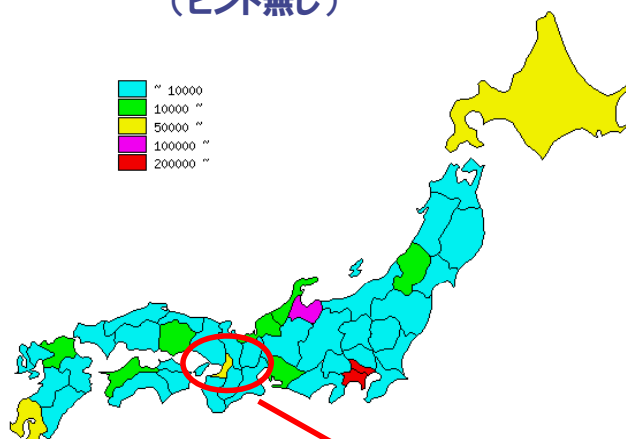
大阪設置のOCNダミーノード
(ヒント有り)

地域別 ノード数(41)

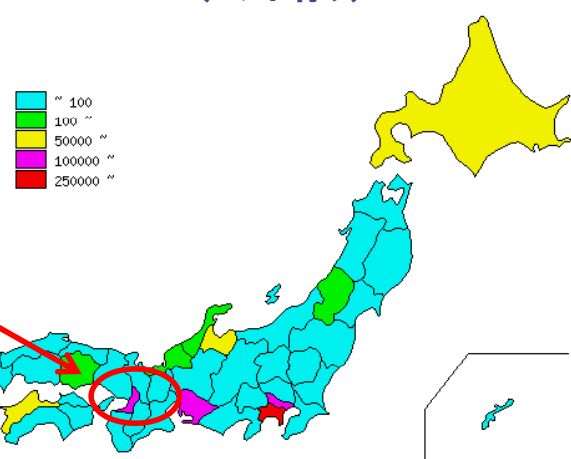


VOD配信実験① 接続時間による地域分布

大阪設置のOCNダミーノード
(ヒント無し)



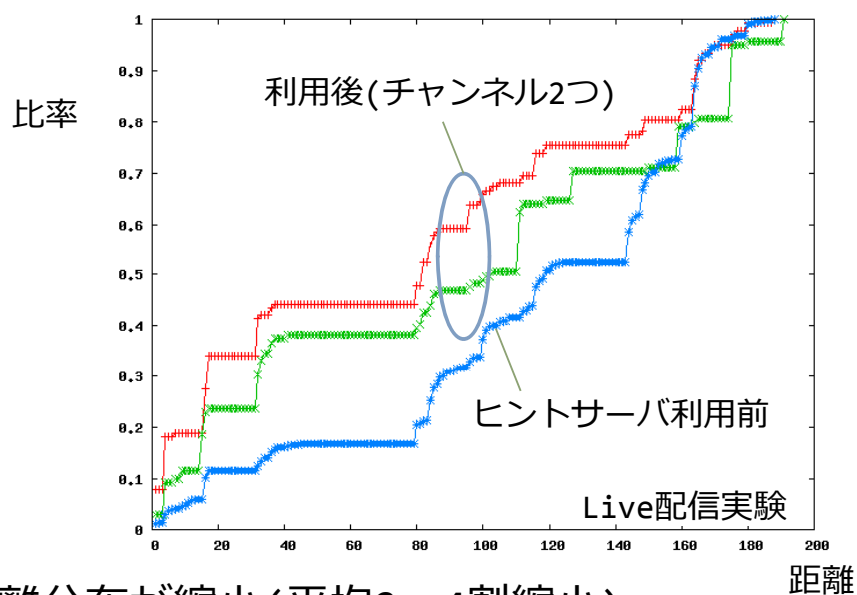
大阪設置のOCNダミーノード
(ヒント有り)



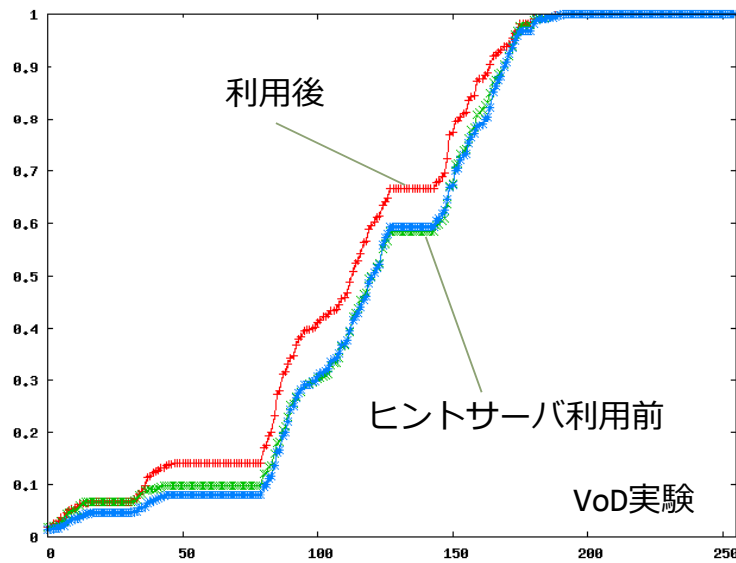
全体としての距離分布

ネットワーク利用効率性は高まっているのか？

○ 距離の累積頻度分布の変化



○ 距離分布が縮小(平均2~4割縮小)

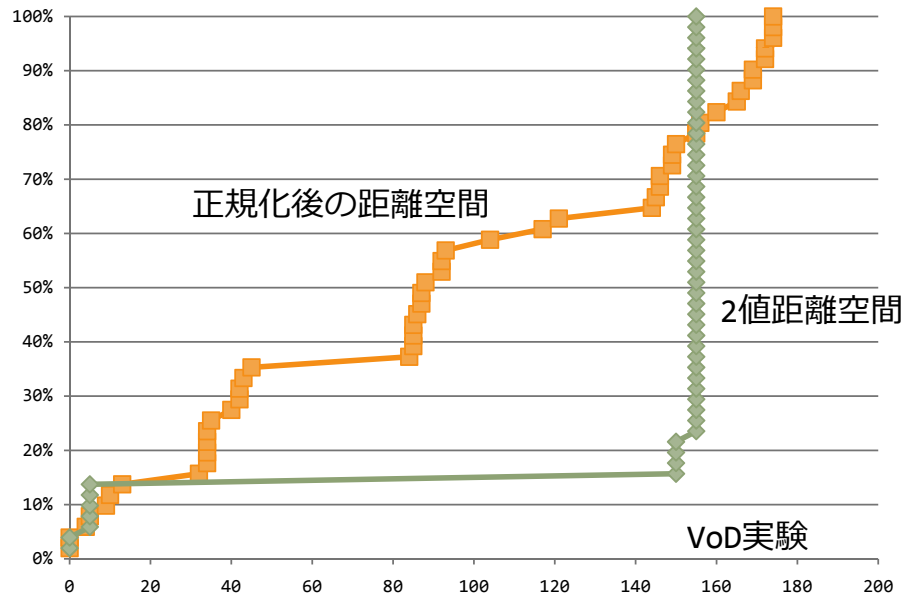


ダミーノード測定の限界？

- 地域分散をしているために評価値が悪く出がち。
 - 制御がうまくいくと真ん中で固まる？
- 正規化手法(人口での重み付け等)の見直しが必要か。
- アプリ側にも統一された統計化機能実装が必要。

ヒントサーバ距離について

- 一致・不一致のみの0/1で計算するとほとんどヒットしないという結果になるので、適切な空間定義は必要



49

検証できたこと

- ピア選択のためのヒントサーバの動作検証
→良好
- ピアがヒントサーバと連携して制御機構をうまく動作させることが可能か
→良好
- 連携動作がネットワーク利用効率に及ぼす影響
→測定方法に課題
- 連携結果が利用者に及ぼす影響
- 制御に用いた尺度
→今後の課題

50

今後の課題

- 効果検証は引き続き
 - ・ シミュレーションもあわせて.
- ダミーノードの運用とデータ分析のシステム化
 - ・ 処理のルール(正規化等)の検討も必要.
- P2Pアプリケーションにおけるデータ収集.
 - ・ スケーラビリティ等を犠牲にしない方法の検討.
- ヒントサーバアーキテクチャ
 - ・ C/Sでいくのか, HTTP GET/POST でいいのか.
 - ・ データ収集部も統合?
- ヒントサーバ距離算出アルゴリズム
 - ・ 距離空間の定義.
 - ・ 複数事業者により算出されたASパスの統合?
- 事業者個別事情に合せたカスタマイズ余地
 - ・ その他の重み・情報(キャッシュサーバ等)の反映
 - ・ ネットワーク構造(v6, NAT, アクセス構造等)

51

実験を予定しているコンテンツ例

- JAXAのイベント中継(全国中継)
 - 近々には若田宇宙飛行士搭乗STS-119(15A)ミッション
- 天体中継(全国中継)
 - 7月22日の日食中継 今年是世界天文年
- 地域スポーツ ホーム&アウェイ中継
 - From 大都市 To 地域
 - ホームに元気なメディア(CATV等)があればホームの中継はある
 - アウェイの中継はコスト的に厳しい場合がほとんど
 - 現在、JFLの複数のチームと共同配信を計画中
- インターネット市民塾の授業コンテンツ(動画)の配信
 - 同一地域内
 - 富山インターネット市民塾等
- 有事の際の防災情報等の配信
 - From 大都市 To 地域
 - 有事の際の一斉アクセス集中に有効
 - ・ このためにも常時にP2Pソフトを利用してもらう

今後の取り組みについて

(1) ネットワーク構造を考慮したP2P配信方式の実証

- InterDomain (ISP間)、IntraDomain (ISP内)
- 大手ISP内のトラフィックエンジニアリング
- NGN、IX(広域、地域)
- CATV型アクセスネットワークのP2P対策、多段NAT対策

(2) ヒントサーバ方式の継続実証

- NWの効率的なアルゴリズムの改良
- 各P2Pソフトウェアへの実装促進

(3) IPv4アドレス枯渇対策

- IPv4→IPv6移行時におけるP2P通信の対応

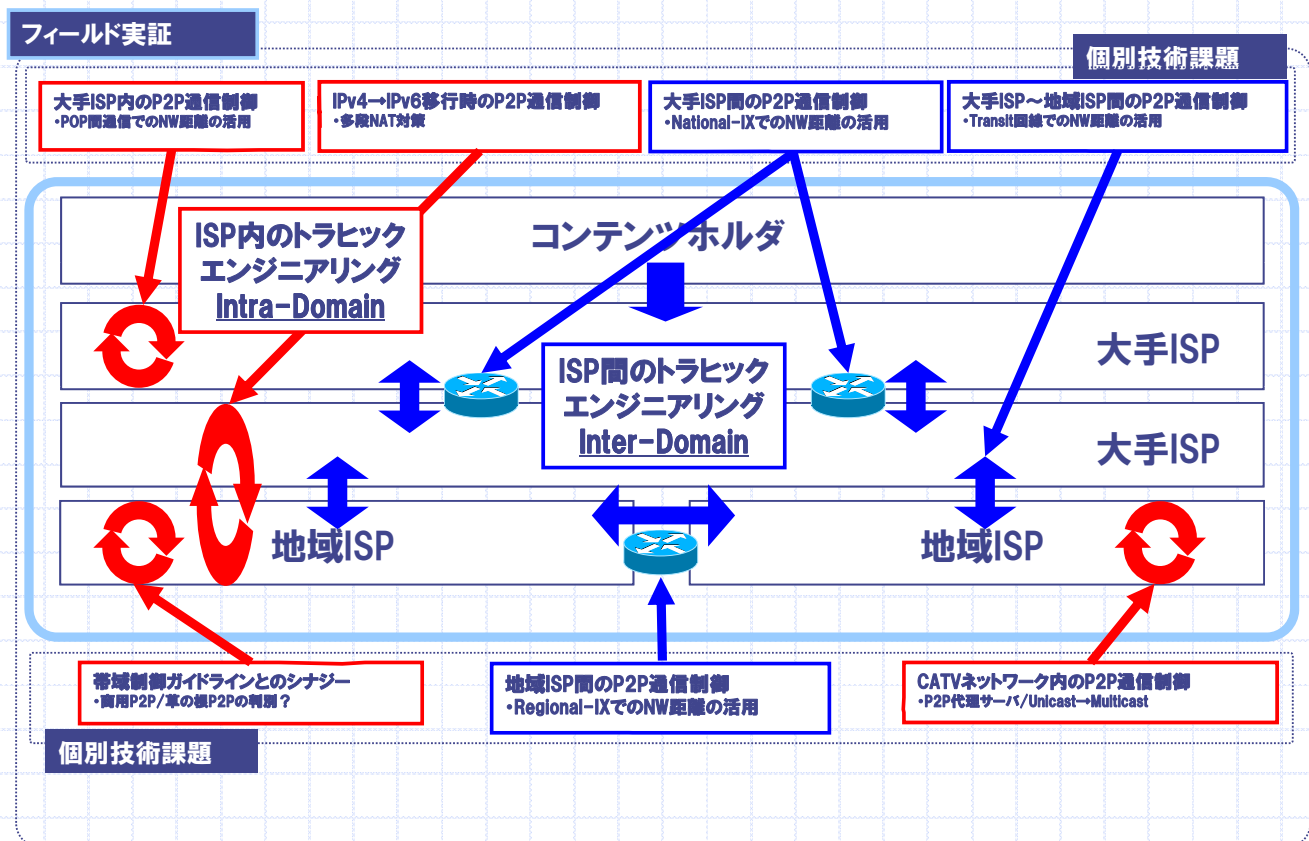
(4) 関連組織との連携、活動のグローバル化

- 国内:「帯域制御の運用基準に関するガイドライン作成委員会」
- 海外:US P4P/IETF、Euro P2P-Next etc...

Copyright (C) 2009, Tatsuya YAMASHITA, NTT Communications Corporation. All rights reserved

53

P2Pネットワーク実証実験 俯瞰図



P2Pネットワーク実験協議会のゴール

インフラに優しいP2P, P2Pに優しいインフラを目指して…

