

# 特別研究報告

題目

## 物理網構成を考慮した ハイブリッド型 P2P 動画像ストリーミング配信機構の評価

指導教官

村田 正幸 教授

報告者

末次 信介

平成 16 年 2 月 19 日

大阪大学 基礎工学部 情報科学科

物理網構成を考慮したハイブリッド型 P2P 動画像ストリーミング配信機構の評価

末次 信介

内容梗概

インターネット利用の普及，アクセス回線の高速化，パーソナルコンピュータの高性能化を背景に，多様な動画像コンテンツのインターネットを介した視聴が一般的になりつつある．現在のインターネットにおいて，多数の利用者に動画像配信サービスを提供するためには，効率的な動画像ストリーミング配信機構が必要不可欠である．マルチキャスト技術を用いれば，サーバに負荷を与えることなく同時に多数の利用者に動画像データを送信できる．そのためには，ネットワークに十分な数のマルチキャストに対応したルータがあること，および配信開始時において，動画像に対する十分な数の視聴要求があることが必要である．しかしながら，現在のインターネットでは一般に利用可能なマルチキャストルータは存在せず，また，利用者の動画像配信に対する要求は多様で，時間的にも分散している．

そこで，本報告では，P2P 通信にもとづくアプリケーションレベルのマルチキャストと，多少の待ち時間を許容することで多数の利用者へ効率よく動画像データを配信できるブロードキャストスケジューリングアルゴリズムにより，サーバに大きな負荷を与えることなく，効率的な動画像ストリーミング配信を実現する．しかしながら，そのためには，ピアが適切に配信ツリーを構築しなければならず，特に，階層的な物理構成を持つネットワークにおいては，その構成や階層内，階層間の通信に対する制約を考慮しなければならない．本報告では，物理網構成を考慮した P2P マルチキャスト配信ツリーを構築するための機構をシミュレーションにより評価した．その結果，提案機構を用いることにより，ネットワークの規模や利用者数によらず，物理網構成にあった適切な配信ツリーが構築され，障害発生時にも動画像再生の途切れが少ない動画像ストリーミング配信が行えることを示した．

## 主な用語

動画像ストリーミング, P2P, マルチキャスト, ブロードキャストスケジューリング

# 目次

|     |   |    |
|-----|---|----|
| 1   | はじめに  | 8  |
| 2   | P2P 通信におけるブロードキャストスケジューリングによる<br>動画像ストリーミング配信 | 11 |
| 2.1 | P2P ネットワーク                                    | 11 |
| 2.2 | ブロードキャストスケジューリングアルゴリズム                        | 12 |
| 3   | 物理網構成を考慮した動画像ストリーミング配信機構                      | 18 |
| 3.1 | ハイブリッド型 P2P 動画像ストリーミング配信機構の概要                 | 18 |
| 3.2 | 物理網構成を考慮した配信ネットワーク階層化                         | 19 |
| 3.3 | セグメント受信スケジュール決定アルゴリズム                         | 20 |
| 3.4 | 配信ツリー構築機構                                     | 21 |
| 3.5 | 障害回復機構  | 26 |
| 4   | シミュレーションによる評価                                 | 30 |
| 4.1 | 評価尺度  | 30 |
| 4.2 | シミュレーション条件                                    | 30 |
| 4.3 | 基本特性の評価                                       | 32 |
| 4.4 | ピア数の変化の影響                                     | 37 |
| 4.5 | 物理網構成の影響                                      | 38 |
| 4.6 | 障害発生率の影響                                      | 40 |
| 4.7 | ファンアウト数の影響                                    | 40 |
| 4.8 | 動画像の長さの影響                                     | 43 |
| 4.9 | シミュレーション評価のまとめ                                | 45 |
| 5   | 物理ネットポロジを考慮した動画像ストリーミング配信機構の評価                | 47 |
| 5.1 | シミュレーション条件                                    | 48 |
| 5.2 | 物理ネットポロジを考慮したシミュレーション評価                       | 48 |

|        |    |
|--------|----|
| 6 おわりに | 53 |
| 謝辞     | 54 |
| 参考文献   | 55 |

## 目 次

|    |  |    |
|----|--|----|
| 1  | 物理ネットワークと P2P 論理ネットワーク . . . . .                   | 12 |
| 2  | スター型 P2P 論理ネットワーク . . . . .                        | 13 |
| 3  | マルチキャストによる動画配信 . . . . .                           | 14 |
| 4  | マルチキャストによる動画配信の必要帯域 . . . . .                      | 14 |
| 5  | ピラミッドブロードキャストの時間帯域マップの例 ( $\alpha = 2$ ) . . . . . | 15 |
| 6  | ピラミッドブロードキャストにおける受信バッファ内データ量の変化 . . . . .          | 16 |
| 7  | ネットワークの階層化 . . . . .                               | 19 |
| 8  | 要求受付時間と要求受付予備時間 . . . . .                          | 20 |
| 9  | 第 2 セグメント以降のスケジュール . . . . .                       | 22 |
| 10 | 提案機構における受信バッファ内データ量の変化 . . . . .                   | 22 |
| 11 | 配信ツリー参加の様子 . . . . .                               | 23 |
| 12 | 障害回復の手順の例 . . . . .                                | 29 |
| 13 | 動画のセグメント分割 . . . . .                               | 31 |
| 14 | GTS の負荷の変動 . . . . .                               | 33 |
| 15 | 再生開始までの待ち時間の分布 . . . . .                           | 33 |
| 16 | 再生の途切れ時間の分布 . . . . .                              | 35 |
| 17 | 最大バッファ内データ量の分布 . . . . .                           | 35 |
| 18 | バッファ内データ量の変化 . . . . .                             | 36 |
| 19 | 新規ピア到着率の影響 . . . . .                               | 39 |
| 20 | 拠点数, 拠点内サブネット数の影響 . . . . .                        | 41 |
| 21 | 障害発生率の影響 . . . . .                                 | 42 |
| 22 | ファンアウト数の影響 . . . . .                               | 44 |
| 23 | 動画の長さの影響 . . . . .                                 | 46 |
| 24 | 物理ネットワークと論理ネットワーク . . . . .                        | 47 |
| 25 | 物理ネットワークにおけるリンクの負荷 . . . . .                       | 50 |
| 26 | 物理ネットワークを考慮した場合の新規ピア到着率の影響 . . . . .               | 51 |

27 物理トポロジを考慮した場合の障害発生率の影響 . . . . . 52

## 表 目 次

|   |                                     |    |
|---|-------------------------------------|----|
| 1 | アドレステーブルの例 . . . . .                | 25 |
| 2 | 標準的なパラメータ設定でのシミュレーション結果 . . . . .   | 32 |
| 3 | 物理ネットポロジを与えた場合のシミュレーション結果 . . . . . | 49 |

## 1 はじめに

インターネット利用の普及，アクセス回線の高速化，パーソナルコンピュータの高性能化を背景に，音楽のビデオクリップ，映画の予告編，ニュース映像やアニメーション，娯楽番組，TVドラマまで幅広い動画像コンテンツのインターネットを介した視聴が一般的になりつつある．従来の World Wide Web を基礎としたダウンロード型の配信サービスでは，利用者は，いったんサーバから動画像データの全てを取得した後，これを再生しなければならない．そのため，視聴開始までに相当の待ち時間が発生する．一方，ストリーミング型の動画像配信サービスでは，利用者は，動画像全体の取得完了を待つことなく，動画像データのうち受信した部分から順次復号化し，再生，視聴するため，待ち時間を抑えることができる．クライアント端末では，配信サーバからのデータ転送遅延や遅延の変動に備え，いくらかの動画像データを自らのバッファに蓄積した後，再生を開始する．

動画像のストリーミング配信により，利用者はいつでも好きな時に望む動画像を見ることが出来る．しかしながら，World Wide Web，ファイル共有システム，高速計算システムなど従来のシステムやサービスがいずれも，Webサーバ，ファイルサーバ，CPUサーバといった全ての情報やデータ，演算を集中管理，制御するサーバに依存しているように，現在の動画像ストリーミング配信機構も，動画像配信サーバを用いたサーバクライアント型のアーキテクチャにもとづいている．そのため，利用者数が増大すると，サーバへの負荷が増加するとともにサーバ近傍のネットワークの輻輳が発生するため，動画像データの配送に遅延やデータ損失が生じ，動画像再生の途切れや乱れを招く．さらに，負荷増大や機器障害などによりサーバが停止すると，システム全体が機能しなくなる．サーバの機器を高性能で耐障害性の高いものに置き換える，またアクセス回線を含むサーバ近傍のネットワークをより高速なものに変更することにより，ある程度問題を軽減することができるが，負荷の集中を避けられるわけではない．また，CDN (Contents Delivery Network) 技術により，ミラーサーバ，プロキシサーバ，キャッシュサーバなどをネットワークの様々な場所に導入し，負荷分散を図ることができる．しかしながら，これらはいずれも固定的に配置されるため，ネットワークにおける利用者分布の偏りや変化，また，サーバに対する要求の変化に十分対応することができない．ネットワーク内でルータがデータを複製して複数の送信先へデータを転

送するマルチキャスト技術を用いれば、サーバへの負荷集中を防ぐとともに同時に多数の利用者へ効率的に動画画像が配信できる。しかしながら、現在のインターネットでは一般に利用可能なルータはマルチキャストに対応しておらず、また、利用者の動画画像配信に対する要求は多様で、時間的にも分散しているため、十分な効果が期待できない。

そこで本報告では、P2P (Peer to Peer) 通信によるアプリケーションレベルのマルチキャストと、多少の待ち時間を許容することで多数の利用者へ効率よく動画画像データを配信できるブロードキャストスケジューリングアルゴリズムにより、サーバに大きな負荷を与えることなく動画画像ストリーミング配信を実現する。P2P 通信は、個々のホストが、サーバを介さず直接情報やデータをやりとりする新しい通信アーキテクチャである。P2P ネットワークを構成するホストはピアと呼ばれ、ピア間の隣接関係によって論理ネットワークを構成する。P2P ネットワークでは、それぞれのピアが情報やデータを持ち合い、互いに交換することで、また、それぞれの計算能力の範囲内で計算処理の一部を担うことにより、特定のピアやネットワークに負荷をかけることなく、多数の利用者間でファイル共有や高速計算処理を実現することができる。P2P 通信を利用し、動画画像を視聴しているクライアントが、受信した動画画像データをさらに複数のクライアントに転送することにより、アプリケーションレベルのマルチキャストが可能となる。しかしながら、P2P 通信におけるピア間の論理ネットワークは、必ずしもリンクやルータなどからなる物理的なネットワークの構造を考慮して構築されるわけではない。そのため、一部のリンクやルータに負荷が集中し、配信サービス全体の性能が低下してしまう可能性がある。

一方、ブロードキャストスケジューリングアルゴリズムは、動画画像の配信要求から視聴開始までのある程度の待ち時間を前提とした動画画像配信の手法である [1]。動画画像データはセグメントと呼ばれる塊に分割され、それぞれ異なるマルチキャストグループを用いて繰り返し配信される。利用者は、先頭のセグメントから順に、マルチキャストグループに参加し、最も近い時刻に開始される配信を受信する。先頭のセグメントを小さくし、残るセグメントの大きさや配信タイミングを適切に設定すれば、視聴開始までの待ち時間を短くするとともに、途切れのない動画画像再生が可能となる。

本報告では、ブロードキャストスケジューリングアルゴリズムを利用した P2P 通信にも

とづく効率的な動画ストリーミング配信機構の有効性を評価する。なお、本報告における特別研究は、日本電気株式会社との共同研究にもとづくものであり、本機構は日本電気株式会社の提案するものである。この機構は、数千人単位の拠点がいくつか、あるいは数十人単位の拠点が数百あるネットワークにおいて、セグメントごとに、物理網構成やシステムの条件を考慮した P2P 論理ネットワーク上の階層的な配信ツリーを構築、管理する。また、配信遅延を考慮したスケジューリング機構、ピアの離脱などに対処する障害回復機構を有する。本報告では、ネットワークの規模や利用者数、また、障害の頻度などの、ピアへの負荷、待ち時間や動画再生の途切れ、障害回復時間などへの影響をシミュレーションにより評価する。また、物理ネットワークにおけるリンク負荷についてもあわせて評価する。

以下、第 2 章では、P2P ネットワークと、ブロードキャストスケジューリングアルゴリズムについて述べる。第 3 章では、ハイブリッド型 P2P 動画ストリーミング配信機構について述べ、第 4 章において、計算機シミュレーションによりこれを評価する。また、第 5 章においては、物理トポロジを考慮したシミュレーション評価を行う。最後に第 6 章において、本報告のまとめと今後の課題について述べる。

## 2 P2P 通信におけるブロードキャストスケジューリングによる 動画像ストリーミング配信

### 2.1 P2P ネットワーク

P2P 通信では、ピアがピア間の接続による隣接関係を確立することにより論理的な P2P ネットワークを構築する。新たに P2P ネットワークに参加するピアは、既にどのようなピアが P2P ネットワークに存在し、それらのうちどれと隣接関係を設定するかを知らなければならない。Napster [2] や KaZaA [3] といった代表的なファイル共有アプリケーションで用いられている P2P 通信プロトコルでは、P2P ネットワークに参加するピアは、まずピアやピアの持つ資源を管理するサーバとの間で接続を確立する。次に、他のピアと共有したいファイルの情報をピア自身の情報とあわせてサーバに登録する。P2P ネットワークからファイルを取得する際には、キーワードなど望むファイルの情報をサーバに送信する。サーバは、登録情報から該当するファイルを検索し、ファイルの所有者やファイル名などの情報をピアに回答する。ピアは望むファイルを所有しているピアとの直接の接続を確立し、ファイルを要求、取得する。このように、ピアやコンテンツに関する情報を蓄積、管理するサーバにより、ファイルなどの情報を交換、共有するためのピアの紹介を受けるものを、ハイブリッド型 P2P と呼ぶ。ネットワークに分散する多数のピアを有機的に結合し、それらの持つ計算能力、蓄積容量などを統合的に利用することにより、大規模計算やファイル分散管理を行うグリッドコンピューティングでは、ピアやピアの資源を管理するため、多くの場合、ハイブリッド型 P2P プロトコルを用いている [4]。

一方、Gnutella [5] や Winny [6] といった P2P 通信プロトコルは、ハイブリッド型 P2P にみられるようなコンテンツ検索のための集中管理サーバを持たないため、ピア型 P2P と呼ばれる。ピア型 P2P では、ピアは、ブートストラッピングノード (bootstrapping node) と呼ばれる、P2P ネットワークに参加しているピアに関する情報を管理するサーバから紹介を受け、あるいはメッセージをブロードキャストすることにより、他のピアの存在を知り、それらのピアとの接続を確立することにより P2P ネットワークに参加する。コンテンツの検索に際しては、ピアはまず、接続を確立した隣接ピアに検索メッセージを送信する。検索

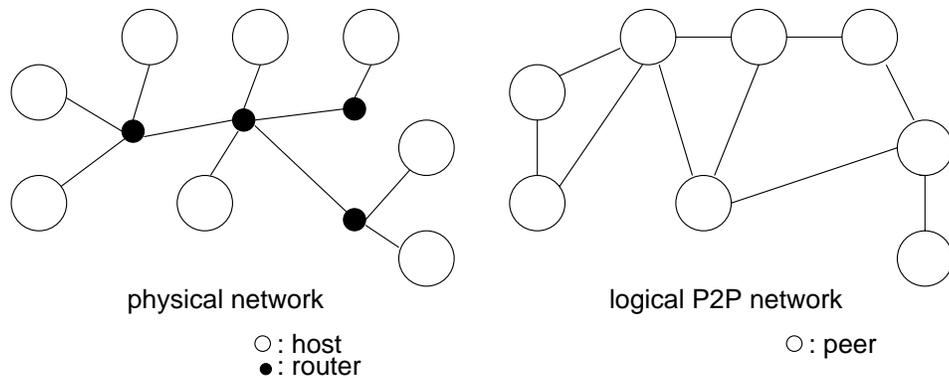


図 1: 物理ネットワークと P2P 論理ネットワーク

メッセージは隣接ピアにより複製され、さらにそれぞれの隣接ピアへと転送されていく。このような、ピアによるメッセージの複製、中継をフラッディングと呼ぶ。コンテンツを所有するピアは、応答メッセージを、検索メッセージを転送して来た隣接ピアへと送信する。応答メッセージは、検索メッセージと同じ経路を遡ることによりもとのピアへと到達する。フラッディングにより P2P ネットワークから所望のコンテンツを発見したピアは、所有者のピアから直接コンテンツを取得する。

このようにして、ホストとルータ、リンクからなる物理ネットワーク上に、ピアとピア間接続によるリンクからなる論理ネットワークが構築される (図 1)。ハイブリッド型 P2P ではサーバを中心としたスター型の (図 2)、ピア型 P2P では隣接関係にもとづくトポロジのネットワークとなる。ピア型 P2P における、隣接関係を確立するピアの選択にあたっては、接続ピア数や遅延などを考慮するが、構築された論理ネットワークは必ずしも物理網構成を十分反映したものとは限らない。

## 2.2 ブロードキャストスケジューリングアルゴリズム

マルチキャストによる動画ストリーミング配信においては、効率的なマルチキャストツリー構築のため、同一動画像に対する多数の配信要求が同時にサーバに届かなければならない (図 3 (a))。配信要求が散発的にサーバに到着した場合に、それらの個々の要求に対する

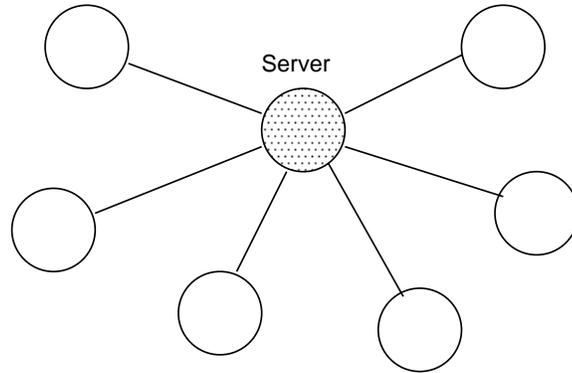


図 2: スター型 P2P 論理ネットワーク

ツリー構築はユニキャスト配信に他ならず，多くの帯域を必要とし，非常に効率が悪い (図 3 (b))．そのため，配信開始までの多少の遅延を許すことにより，ある一定の要求受付時間を設けてその間に到着した要求をまとめて処理する手法 (図 3 (c))，あるいはある一定数の要求受信後にマルチキャストツリーを構築する手法 (図 3 (d)) などが提案されている [7]．しかしながら，例えば一定時間ごとの配信を行う方式において，符号化レート  $b$ ，長さ  $L$  の動画を最大待ち時間  $D$  で多数の利用者に配信するためには，図 4 に示すように，最大  $bL/D$  の帯域が必要になる．

ブロードキャストスケジューリングアルゴリズムは，動画をセグメントに分割し，セグメントごとに異なるチャンネルで繰り返して配信することで，マルチキャストによる効率的な動画ストリーミング配信を実現する [1, 8, 9]．ブロードキャストスケジューリングにおける動画の受信，再生，および配信の様子は，図 5 に示されるような時間帯域マップ (Temporal-Bandwidth Map) によって表すことができる [1]．それぞれの四角はセグメントに対応しており，幅は時間を，高さは帯域の大きさを表す．図では，符号化レート  $b$  の動画が 3 つのセグメント  $S_1 \sim S_3$  に分割されている．図の上部をプレイアウト・エリアと呼び，クライアントの動画の視聴状況を表す．ここでは，クライアントが符号化レート  $b$  にしたがってセグメントを先頭から順に途切れなく再生，視聴していることが示されている．また，図の下部をブロードキャストイング・エリアと呼び，動画データの配信状況を表す．図で

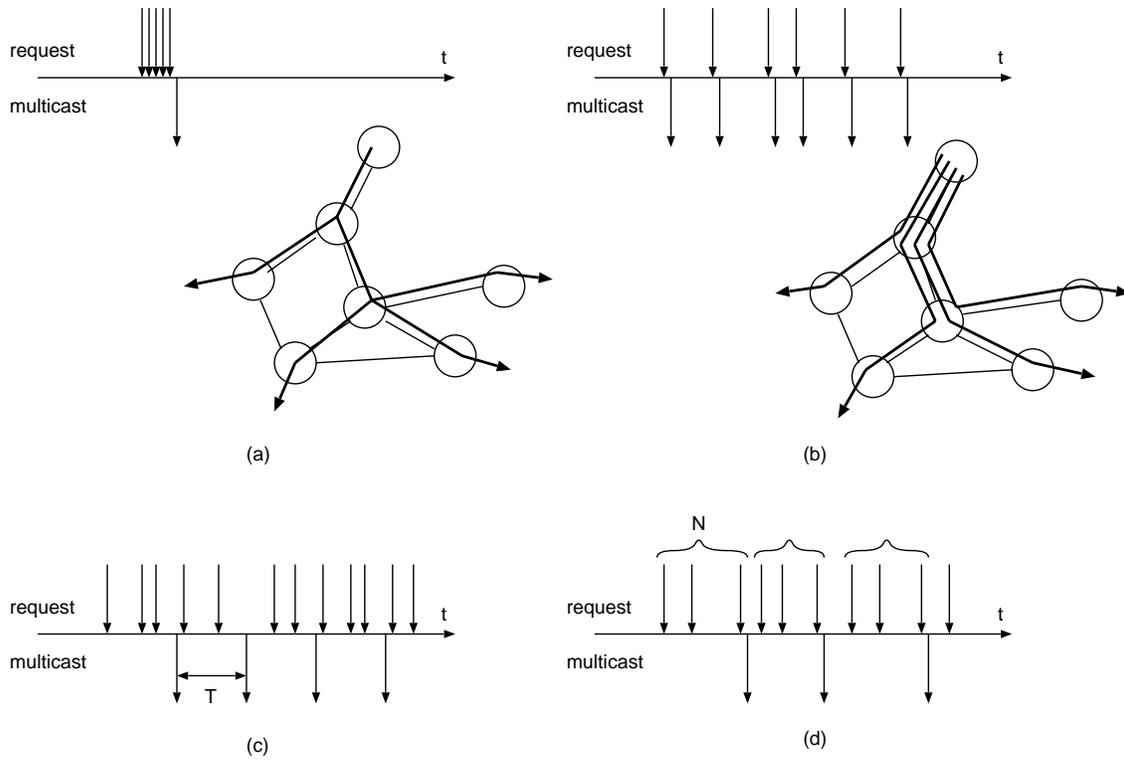


図 3: マルチキャストによる動画配信

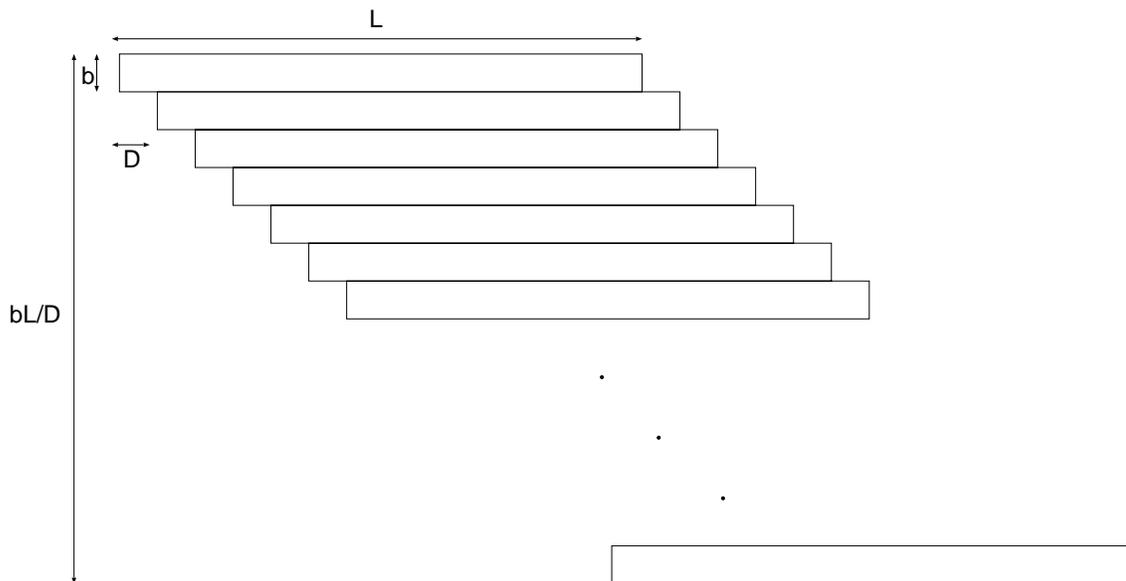


図 4: マルチキャストによる動画配信の必要帯域

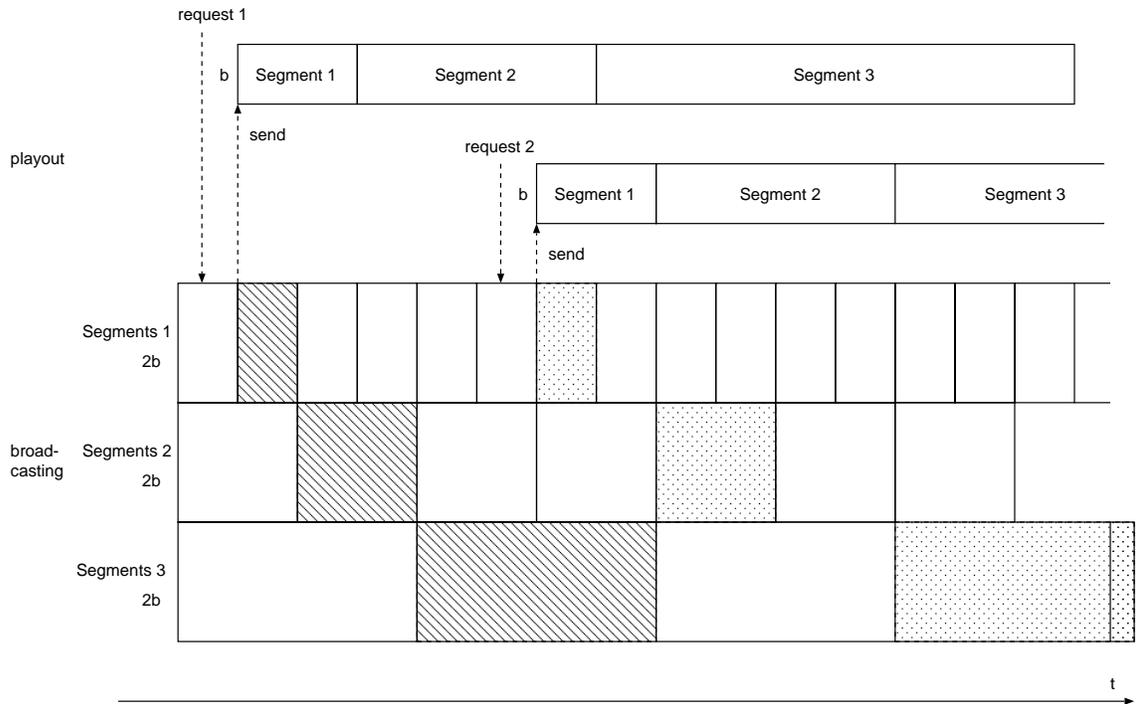


図 5: ピラミッドブロードキャストの時間帯域マップの例 ( $\alpha = 2$ )

は、セグメントごとに異なる 3 つのチャンネルで、それぞれの符号化レートの倍のレートでの配信が繰り返されスケジュールされていることが示されている。配信レートが倍であることから、セグメントあたりの配信時間が再生時間の半分になっていることが分かる。なお、あるチャンネルにおけるセグメント配信の繰り返しをスロットと呼ぶ。セグメントごと、あるいはスロットごとにマルチキャストグループを設定し、クライアントが適切なマルチキャストグループに参加して動画セグメントを受信することにより、動画ストリーミング配信が達成される。

ブロードキャストスケジューリングには、セグメントの分割法や配信レートなどの異なるいくつかのアルゴリズムが提案されている。例えば、ピラミッドブロードキャスト (Pyramid Broadcasting) では、動画はセグメントの長さが等比数列となるように分割される [9]。一方、スカイスクレーパブロードキャスト (Skyscraper Broadcasting) では、幾何級数となるようなセグメント分割を行う [8]。本節では、本報告で用いるピラミッドブロードキャストについて説明する。なお、以降では、ピラミッドブロードキャストを PB と表記する。

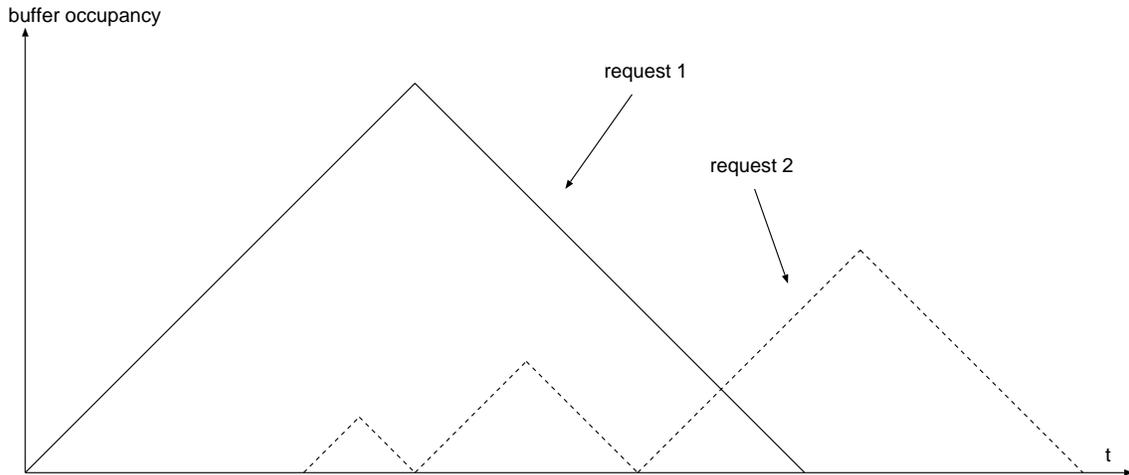


図 6: ピラミッドブロードキャストにおける受信バッファ内データ量の変化

PB では、セグメント長が公比  $\alpha$  の等比数列となるように動画像を分割し、それぞれを符号化レートの  $\alpha$  倍のレートで配信する。ただし、 $\alpha > 1$  である。図 5 は、PB における  $\alpha = 2$  の場合の時間帯域マップである。

利用者からの動画像配信要求に対し、要求受信のタイミングに応じて、それぞれのセグメントを受信するスロットが決定される。第 1 セグメントについては、最も近い時間に配信が開始されるスロットが割り当てられる。第 2 セグメント以降においては、前セグメントの配信直後から、最も近い時間に配信が開始されるスロットでセグメントを受信する。例えば、図中、request 1 と書かれたタイミングで動画像配信を要求した利用者は、斜線で表されたスロットのセグメントを順に受信することになる。また、request 2 のタイミングで動画像配信を要求した利用者は、網かけされたスロットのセグメントを順に受信する。

利用者は、動画像セグメントを受信し始めると、バッファに蓄積するとともにすぐに再生を開始する。再生の完了した動画像セグメントはバッファから削除される。図 6 は、request 1, request 2 を発行した利用者端末におけるバッファ内データ量の変化を表している。なお、直線の傾きは  $b$  または  $-b$  である。文献 [1] では、遅延のない理想的なネットワークにおいて、PB ではバッファが枯渇することなく途切れのない動画像の視聴が実現できることが示されている。 $\alpha = 2$  の時、配信レートは符号化レートの 2 倍であるので、PB における動画

像再生開始までの最大待ち時間は、第1セグメントの長さの半分に等しい。第1セグメントの長さを  $W$  とした時、図5のように動画像が3つのセグメントに分割されたとすると、動画像の長さは  $7W$  となる。一定時間ごとに配信を繰り返すマルチキャストによる動画像ストリーミング配信において、PB と同じく最大待ち時間  $W/2$  を実現するためには、配信開始を  $W/2$  ずつずらしたマルチキャストを行わなければならない。図4において、 $L = 7W$ 、 $D = W/2$  となるため、チャンネル数は14であり、また、それぞれのチャンネルの配信レートは  $b$  であることから、配信に必要な帯域は  $14b$  となる。一方、PB では、図5に示されており、必要帯域は  $6b$  であり、効率的な動画像ストリーミング配信が行えることが分かる。

### 3 物理網構成を考慮した動画像ストリーミング配信機構

本章では、P2P 通信によるブロードキャストスケジューリングアルゴリズムを利用した動画像ストリーミング配信機構について述べる。なお、本報告における研究は日本電気株式会社との共同研究にもとづくものであり、本章で述べる機構は日本電気株式会社の提案するものである。以降では、動画像を配信、受信する主体をピアと呼ぶ。本機構は、P2P ネットワーク全体の管理するサーバを有し、ハイブリッド型の P2P アーキテクチャである。ただし、配信ツリーはピアとピアの直接の接続によって構築されるため、ピア型 P2P のトポロジとなる。

#### 3.1 ハイブリッド型 P2P 動画像ストリーミング配信機構の概要

動画像データ配信のスケジューリングアルゴリズムとしては、 $\alpha = 2$  の PB を用いる。したがって、動画像データは第 1 セグメントの長さを  $W$  とすると、順に  $W, 2W, 4W, 8W, \dots, 2^{i-1}W$  という大きさに分割され、それぞれ異なるチャネルを用いて、符号化レートの 2 倍の配信レートで送出される。動画像セグメントは、P2P 通信によるアプリケーションレベルマルチキャストにより配信される。動画像セグメントの配信ツリーはスロットごとに構築されるが、多くの企業、組織においては、拠点や部署といった組織的構造にもとづいて物理ネットワークが階層化されていることを考慮し、3 階層に構造化される。階層化については、3.2 節において詳細を述べる。

動画像の配信を受けたいピアはまず、各セグメントをどのスロットで受信すればいいのかを、スケジュール管理サーバ (SS; Schedule Server) に問い合わせる。SS は、動画像データの配信スケジュールを管理するサーバであり、3.3 節で述べるスケジュール決定アルゴリズムにもとづいてピアの動画像セグメント受信をスケジューリングし、ピアに通知する。

ピアは、それぞれのセグメントについて、定められたタイミングに、配信ツリー管理サーバ (GTS; Global Tree Server) へ配信ツリー参加要求を送信する。3.4 節に述べる配信ツリー構築機構にもとづいて配信ツリーが設定され、セグメントの配信が開始される。

ピアの離脱や回線の故障などにより、配信が正常に動作しなくなる現象を障害と呼ぶ。障

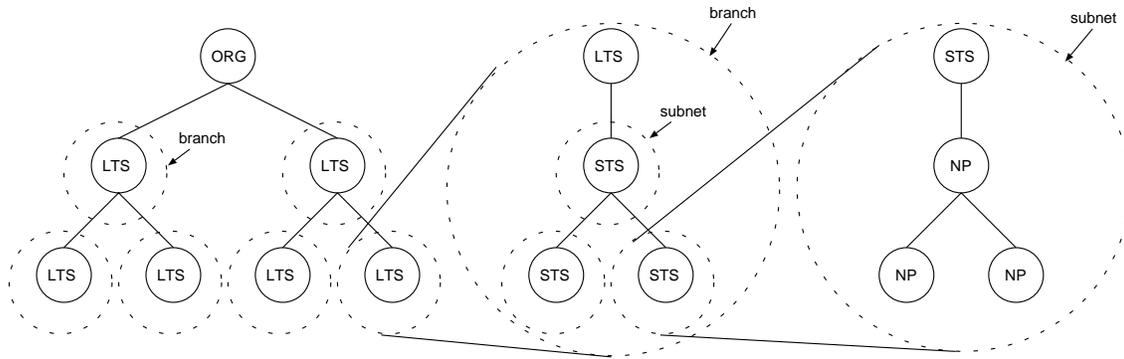


図 7: ネットワークの階層化

害が発生すると、3.5 節において述べる障害回復機構により、配信ツリーの再構築を行う。

### 3.2 物理網構成を考慮した配信ネットワーク階層化

セグメント配信のための P2P 論理ネットワークを、拠点 (branch)、拠点内サブネット (subnet) といった物理ネットワークの構成にもとづいて構造化する。それぞれの拠点には 1 つの拠点代表 (LTS; Local Tree Server) が、拠点内のそれぞれのサブネットには 1 つのサブネット代表 (STS; Subnet Tree Server) が、後述のアルゴリズムにより選出される。また、LTS、STS 以外のピアを、NP (Normal Peer) と呼ぶ。

これらを、図 7 のような、3 階層の配信ツリーに構成する。配信ツリーは、動画サーバ (ORG; Origin Server) を根とした LTS 間ツリー、LTS を根とした STS 間ツリー、STS を根とした NP 間ツリー、から構成される。なお、LTS、STS にとって、同一階層での配信ツリーの子を単に子、また、LTS に接続した STS、STS に接続した NP を直下の子と呼ぶ。

それぞれの階層において、各 LTS、STS、NP に接続可能な子の数をファンアウト数と呼ぶ。図では、ファンアウト数は 2 である。ファンアウト数は、利用可能な帯域や処理能力により決定される。

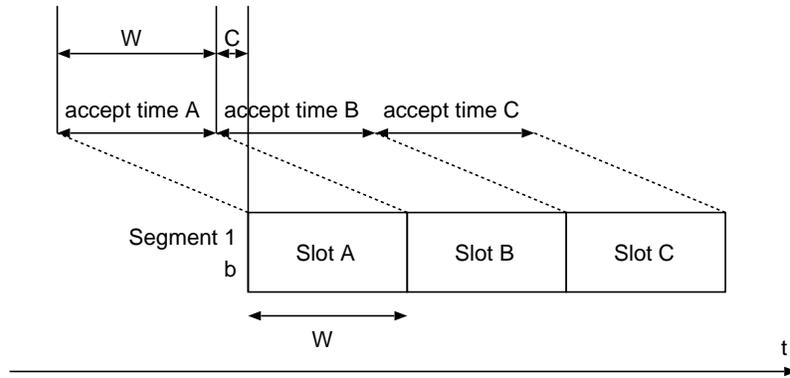


図 8: 要求受付時間と要求受付予備時間

### 3.3 セグメント受信スケジュール決定アルゴリズム

新たに動画像ストリーミング配信サービスを受けるピアは、すべてのセグメントについて受信するスロットを決定するため、SS (Scheduling Server) にスケジュール決定要求を送信する。スケジュール決定要求には、配信を要求する動画像の識別子が記述されている。SS は、ピラミッドブロードキャストイングアルゴリズムにもとづいてスロットを決定し、各セグメントのスロット割り当てと、それらのスロットに対する配信ツリー参加要求の送信時間を、ピアに通知する。

2.2 節で述べた通り、PB では、要求を受信した時刻に一番近く配信が開始される第 1 セグメントのスロットをクライアントに割り当てる。しかしながら、本報告の動画像ストリーミング配信機構では、P2P 論理ネットワーク上にセグメント配信のための配信ツリーを構築するため、実際にセグメントが配信可能になるまでに時間がかかる。そこで SS は、第 1 セグメントのそれぞれのスロットに対して要求受付時間を定める。図 8 に示すように、あるスロットの要求受付時間はそのスロットの配信開始の  $W + C$  秒前から  $W$  秒間とする。この期間に SS がスケジュール決定要求を受信したピアに対しては、対応するスロットでの第 1 セグメント配信が行われる。なお、 $C$  を要求受付予備時間と呼び、 $C$  秒前に要求を受け付けたピアを含む配信ツリーが構築されるのに十分な時間に設定される。

第 2 セグメント以降については、 $\alpha = 2$  の PB にもとづいてスロットが決定される。ただし、第 1 セグメントと第 2 セグメントは同時に受信が可能とする。例えば、図 9 において、

request 1 のタイミングでスケジュール決定を要求したピアは、網かけされたスロットを順次受信することになるが、request 2 のタイミングで要求したピアは、斜線のように、第 1 セグメントと同時に第 2 セグメントを受信する。第 1 セグメントと第 2 セグメントの同時受信により、図 10 に示されるように、動画像ストリーミング配信サービスへの参加直後に多くの動画像データをバッファに蓄積することができる。そのため、再生開始間もない期間にピアの離脱やリンク故障など配信ツリーの障害が発生しても、障害回復完了までのバッファ枯渇による動画像再生の途切れを避けることができる。

### 3.4 配信ツリー構築機構

ピアは、セグメントを受信するよう指定されたスロットについて、スロットの配信開始時間から  $W + C$  秒前から  $W$  秒間の要求受付時間内に、配信ツリー参加要求を送信する。配信ツリー参加要求には、動画像の識別子、セグメント識別子、およびスロット識別子が含まれる。要求の送信先は、直前のセグメントにおけるピアの役割 (ピアタイプと呼ぶ) によって異なる。

ピアの配信ツリーへの参加は、GTS および配信ツリー内のピアが、接続先となるピアの候補を順次紹介することにより行われる。図 11 に、ピアの配信ツリー参加までのメッセージのやりとりの様子を示す。まず、ピアは、第 1 セグメントの配信ツリーに参加するため、GTS (Global Tree Sexrver) に参加要求を送出する。GTS は、ピアの所属する拠点の拠点サーバ (LTS) が存在すれば LTS を、さらにサブネットサーバ (STS) が存在すれば STS を紹介する。図では、GTS はピアに STS を紹介している。ピアは、紹介されたピアに、接続要求を送信する。接続要求を受信した STS は、直下の子がない場合は、ピアを直下の子とし、直接接続する。STS に直下の子がすでにいる場合には、ピアに直下の子である NP を紹介する。ピアは紹介されたピアに再度接続要求を送信する。接続要求を受信した NP は、子の数がファンアウト数未満ならピアとの接続を確立するが、ファンアウト数に達している場合にはさらにその子を紹介する。このように順次紹介を受けることにより、配信ツリーに参加することができる。以下、配信ツリー構築の詳細を述べる。

新たにサービスに参加したピアは、まず SS からスケジュールを受信する。スケジュール

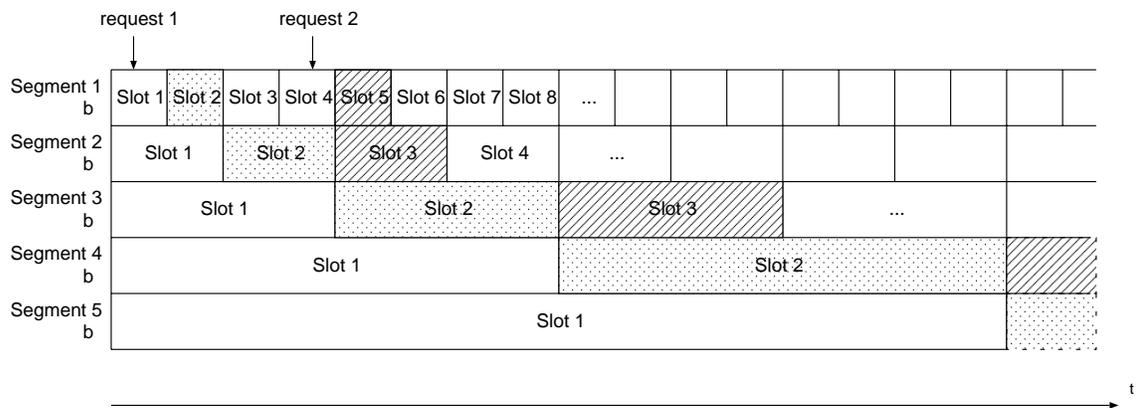


図 9: 第 2 セグメント以降のスケジュール

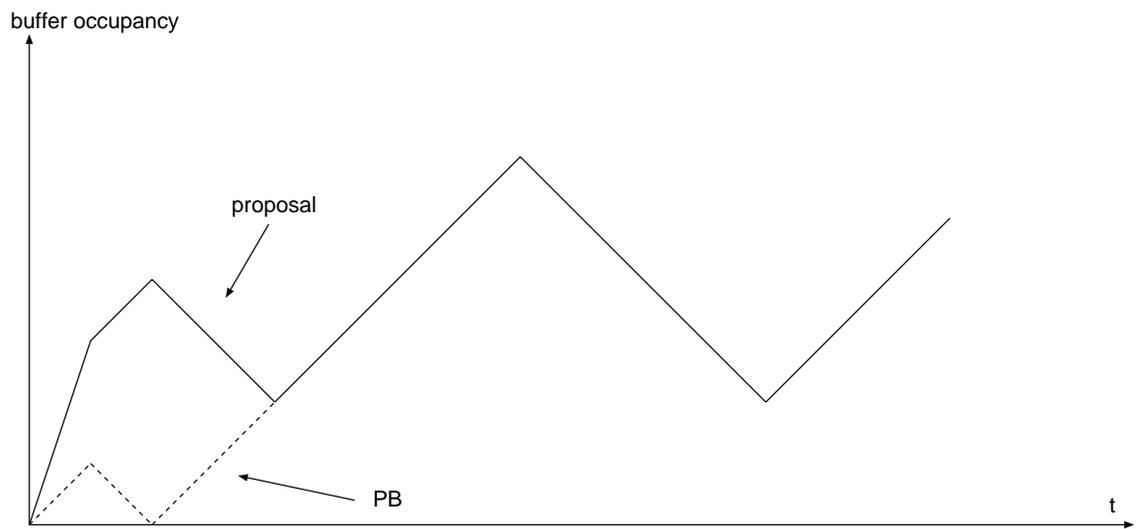


図 10: 提案機構における受信バッファ内データ量の変化

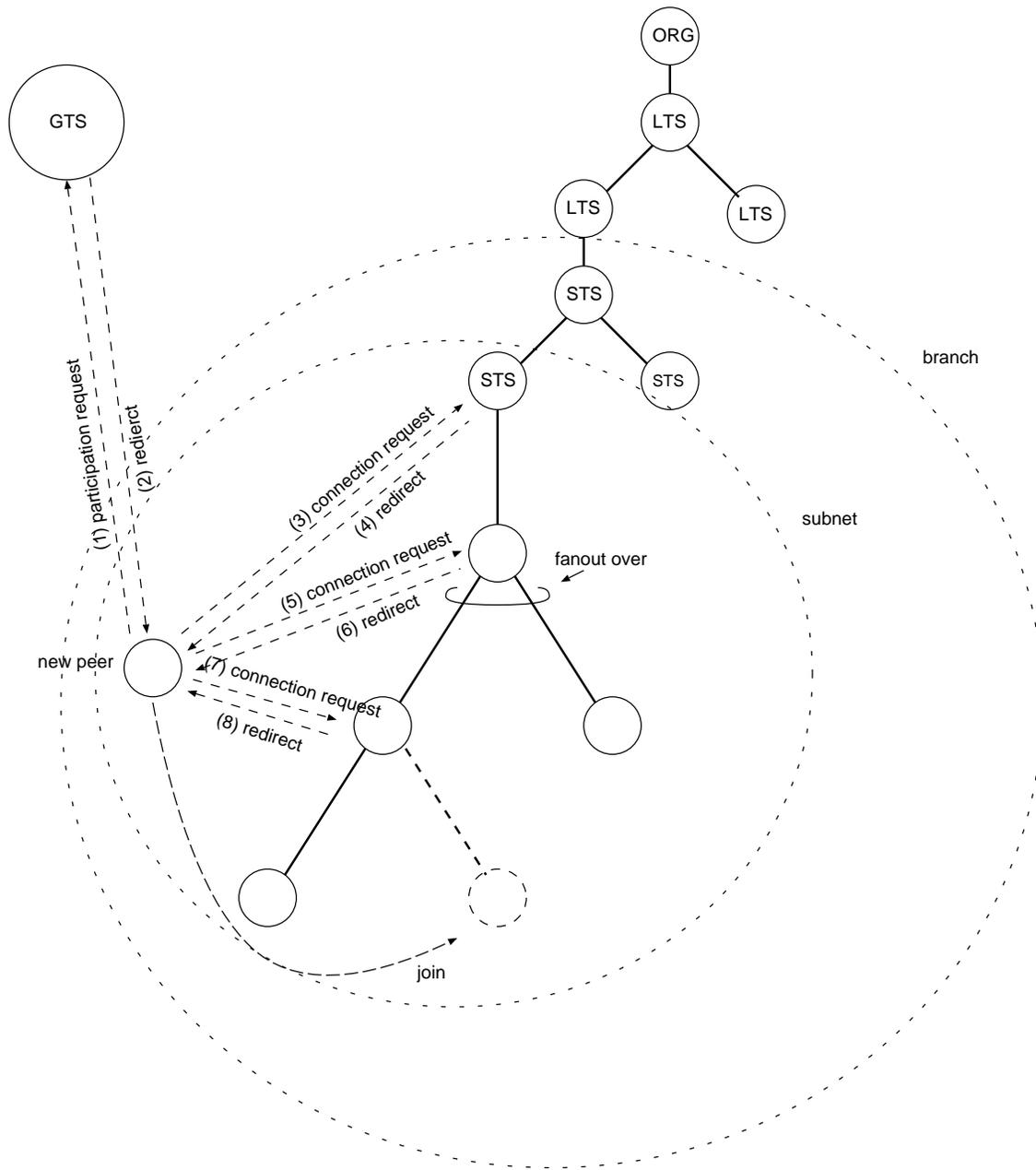


図 11: 配信ツリー参加の様子

決定要求の受付時間と第 1 セグメントに対する配信ツリー参加要求の受付時間は重なっているため、ピアはスケジュールを受信するとすぐに、GTS に配信ツリー参加要求を送信する。ピアからの配信ツリー参加要求に対し、GTS はピアのピアタイプを定め、接続先のピアを紹介する。なお、ピアの配信ツリー参加要求に対して紹介されるピアを仮親と呼ぶ。ピアタイプの決定と仮親紹介のため、GTS はアドレステーブルを管理している。アドレステーブルの例を表 1 に示す。アドレステーブルは、拠点、サブネットを識別するためのネットワークアドレス、拠点ごとに拠点名と LTS の IP アドレス、および拠点内の STS の IP アドレスからなる。ただし、拠点、サブネットのネットワークアドレスと拠点名は、事前にシステム管理者によって与えられる。LTS や STS の IP アドレスは、LTS、STS が未指名の場合、いずれも 0.0.0.0 に初期化されている。GTS は配信ツリー参加要求を受信すると、アドレステーブルとピアの IP アドレスから、ピアを LTS、STS、NP のいずれかに任命し、LTS、STS の場合にはアドレステーブルにピアの IP アドレスを登録する。まず、ピアの属する拠点の LTS がない場合には、ピアを LTS に任命し、仮親として動画像サーバの IP アドレスを通知する。また、ピアの IP アドレスをアドレステーブルに登録する。ピアの属する拠点の LTS は既にあるがサブネットに STS がない場合には、ピアを STS に任命し、仮親として LTS の IP アドレスを通知するとともにピアを STS として登録する。STS がすでにある場合には、ピアタイプを NP とし、STS を仮親に指定する。

ピアは GTS から、自身のピアタイプと仮親の IP アドレスを受け取り、記憶する。ピアは、指定された仮親に接続要求を送信し、接続を試みる。同時に、自身のピアタイプを仮親に知らせる。ピア A がピア B の接続要求を受信したとする。ピア A は、ピア B のピアタイプと自身のピアタイプを比較する。両者が同じで子数がファンアウト数未満であれば、ピア A はピア B を子と定め、自身の管理する子のアドレスリストにピアの IP アドレスを登録し、接続を確立する。ピアタイプが一致しているがすでにファンアウト数だけの子を持つ場合には、子として接続した順に、子を新しい仮親としてピア B に通知する。したがって、配信ツリーは幅優先で構築されることになり、動画像サーバからピアへの動画像セグメントの配信遅延を抑えることができる。一方、ピアタイプが異なる場合には、ピア B はピア A の下の階層の配信ツリーに参加することになる。ピア A が直下の子を持たない場合

表 1: アドレステーブルの例

| 拠点名 | ネットワークアドレス   | ピアタイプ | LTS , STS の IP アドレス |
|-----|--------------|-------|---------------------|
| 札幌  | 192.1.0.0/16 | LTS   | 0.0.0.0             |
|     | 192.1.1.0/24 | STS   | 0.0.0.0             |
|     | 192.1.2.0/24 | STS   | 0.0.0.0             |
| 東京  | 192.2.0.0/16 | LTS   | 192.2.1.1           |
|     | 192.2.1.0/24 | STS   | 0.0.0.0             |
|     | 192.2.2.0/24 | STS   | 0.0.0.0             |
| 名古屋 | 192.3.0.0/16 | LTS   | 192.3.1.1           |
|     | 192.3.1.0/24 | STS   | 192.3.1.2           |
|     | 192.3.2.0/24 | STS   | 192.3.2.1           |
| 大阪  | 192.4.0.0/16 | LTS   | 192.4.1.1           |
|     | 192.4.1.0/24 | STS   | 0.0.0.0             |
|     | 192.4.2.0/24 | STS   | 192.4.2.1           |
| 福岡  | 192.5.0.0/16 | LTS   | 192.5.2.1           |
|     | 192.5.1.0/24 | STS   | 192.5.1.1           |
|     | 192.5.2.0/24 | STS   | 192.5.2.2           |

には、ピア B を直下の子として登録、接続するが、そうでない場合には、直下の子を新たな仮親として紹介する。ピア B は、紹介された仮親を仮親のアドレスリストに追加し、改めて新しい仮親への接続を試みる。したがって、配信ツリーへの参加に成功した時、ピア B はその祖先ピアの IP アドレスのリストを持つことになる。

続くセグメントに対する配信ツリーへの参加は、前セグメントでのピアタイプによって異なる。なお、配信ツリーへの参加はセグメント順とし、第 1 セグメントと第 2 セグメントの同時受信を指示されている場合にも、第 1 セグメントの配信ツリーの参加後に、第 2 セグメントの配信ツリー参加要求または接続要求を送信する。前セグメントで NP であったピア

は、前回の STS を仮親とし、要求受付が既に開始されている場合はすぐに、そうでない場合には、受付時間のうち一様分布に従うランダムな時刻に、接続要求を送信する。以後、第 1 セグメントと同様に、仮親の紹介、接続を繰り返すことで配信ツリーに参加できる。前セグメントで STS であったピアのピアタイプが LTS または STS に変わった場合、ピアはそのむねの通知を受け、新しい STS の紹介を受けるため、改めて GTS に配信ツリー参加要求を送信する。

一方、前セグメントで LTS または STS であったピアは、NP であったピアの接続要求を受信、処理できるよう、いち早く上位階層の配信ツリーを構築しなければならない。そのため、前セグメントで LTS または STS であったピアは、要求受付の開始と同時に GTS に配信ツリー参加要求を送信する。要求受付が既に始まっている場合には、ただちに要求を送信する。

### 3.5 障害回復機構

配信ツリーの構築中に指定された仮親にアクセスできない、動画像データの受信中に親が離脱したなど、動画像データの配信に支障が生じることを、障害と呼ぶ。障害が発生した場合には、障害回復により配信ツリーを再構築する。配信ツリー再構築の手順は、アクセスできなかった仮親や離脱した親と自身のピアタイプによって異なる。

仮親または親と自身のピアタイプが同じであった場合には、配信ツリー参加時に生成された祖先ピアのリストにもとづき、祖父母のピアに接続要求を送信することで、配信ツリーの再構築を試みる。祖父母のピアは接続の切断によって、子、すなわち障害回復を行うピアにとっては仮親または親のピアの離脱、障害を検知し、子のリストから当該ピアを取り除く。ピアは、接続に成功すれば、障害回復を完了する。ファンアウト数の制限により接続に失敗した場合には、祖父母より仮親の紹介を受け、配信ツリー構築時と同様の手順でいずれ配信ツリーに参加する。祖父母の障害によりアクセスできなかった場合は、さらにその親へ、と順に祖先ピアをさかのぼっていく。祖先の情報がない場合には、再び GTS に配信ツリー参加要求を送信する。なお、アクセスできなかった仮親または離脱した親が LTS、STS だった場合には、GTS にピアの離脱を通知する。通知を受け取った GTS はアドレステーブル

から対応する拠点，またはサブネットの LTS または STS のアドレスを初期化する．

仮親または親と自身のピアタイプが異なる場合は，ピアタイプの組合せは LTS と STS ，または STS と NP である．ピアは GTS に親の離脱を通知するとともに，GTS に配信ツリー参加要求を送信する．その結果，ピアのタイプが NP から STS ，あるいは STS から LTS に変化すると，このピアに接続していたピアにとっては親の離脱となり，同様に障害回復を行うことになる．

障害回復手順において紹介された仮親と自身のピアタイプが，LTS と NP など正常な組合せでない場合には，GTS へ配信ツリー参加要求を送信する．

図 12 に障害回復の様子を例示する．図中，ORG は動画像サーバである．LTS ，STS ，NP はピアタイプを表し，それぞれ拠点サーバ，サブネットサーバ，その他のピアである．また，数字はピアの識別子を示す．図の左上部に示す配信ツリー (a) において，ピア 2 が離脱した場合を考える．配信ツリー (b) に示すように，LTS 間配信ツリーでピア 2 の子であったピア 4 は，親であるピア 2 の離脱を認識するとただちに，祖父母のピア 1 へ接続を試みる (1)．ピア 1 は，ピア 2 の離脱により，ファンアウトに空きができたので，ピア 4 を子として受け入れる (2)．ピア 4 は，ピア 1 と接続を確立し，配信ツリー (c) が構成される．ピア 5 は，ピア 2 の直下の子であり，ピアタイプが異なるため，ピア 2 の離脱を認識すると，GTS へ配信ツリー参加要求を送信する (3)．同じ拠点内に LTS がいないため，ピア 5 は，LTS に任命され (4)，ピア 5 は LTS に昇格する．ピア 5 は，LTS に任命されると同時に，仮親として ORG の IP アドレスを受け取る．ピア 5 は ORG に接続を試みるが (5)，ORG のファンアウト数の制限により，ピア 1 を紹介される (6)．同様に，ピア 1 への接続要求 (7) に対しては，仮親としてピア 4 を紹介され (8)，ピア 4 に接続することができる (9)(10)．ピア 5 の昇格により，親を失ったピア 6，ピア 7，ピア 8 も障害回復を行う．ピア 6 は，祖先ピアリストにもとづき，祖父母であったピア 2 に接続を試みる (11)．しかしながら，ピア 2 は既に離脱しているため，アクセスできない．ピア 6 はピア 2 の親のアドレスを所持していないため，GTS に配信ツリー参加要求を送信する (12)．その結果，同じ拠点内にあるピア 5 を紹介され (13)，ピア 5 に接続し (14)(15)，配信ツリーは (e) のようになる．また，ピア 7 は，ピア 6 と同様に，ピア 2 に接続を試みた後 (16)，GTS に配

信ツリー参加要求を送信し (17), ピア 5 を紹介され (18), 接続を試みる (19). しかしながら, ピア 5 には既に直下の子としてピア 6 が接続しているため, ピア 6 を紹介され (20), ピア 6 に接続する (21)(22). このようにして構築された配信ツリー (f) において, 親であったピア 5 とピアタイプの異なるピア 8 は, GTS に配信ツリー参加要求を送信する (23). サブネットには STS がいないため, STS に任命され, 拠点サーバであるピア 5 を紹介される (24). ピア 5 に直下の子であるピア 6 の紹介を受け (25)(26), ピア 6 に接続を試みる (27)(28). その結果, 配信ツリーは (g) のようになる. 続いて, ピア 8 の昇格により, 親を失ったピア 9, ピア 10 も同様に障害回復を行う. ピア 9 はピア 5 に接続を試みるが (29), ピア 5 は既に LTS に昇格しているため (30), 再度 GTS に配信ツリー参加要求を提出し (31), 同じサブネット内にあるピア 8 を紹介され (32), ピア 8 に接続する (33)(34). ピア 10 もピア 9 と同様, ピア 5 に接続を試みた後 (35)(36), 再度 GTS に配信ツリー参加要求を提出し (37), 同じサブネット内にあるピア 8 を紹介され (38), ピア 8 にピア 9 を紹介され (39)(40), ピア 9 に接続する (41)(42), このようにして障害回復が完了し, 全てのピアが配信ツリーに再度組み込まれる. なお, ピアの障害回復時手順をピアごとに記述したが, 実際には, 例えばピア 5 の昇格と同時にピア 6, ピア 7, ピア 8 は同時に障害回復を開始する.

障害回復が完了すると, 新たな親から動画像セグメントの続きを受信する. しかしながら, 当該セグメントの視聴終了により親に動画像データが残っていない場合は, 動画像再生に途切れが生じる.

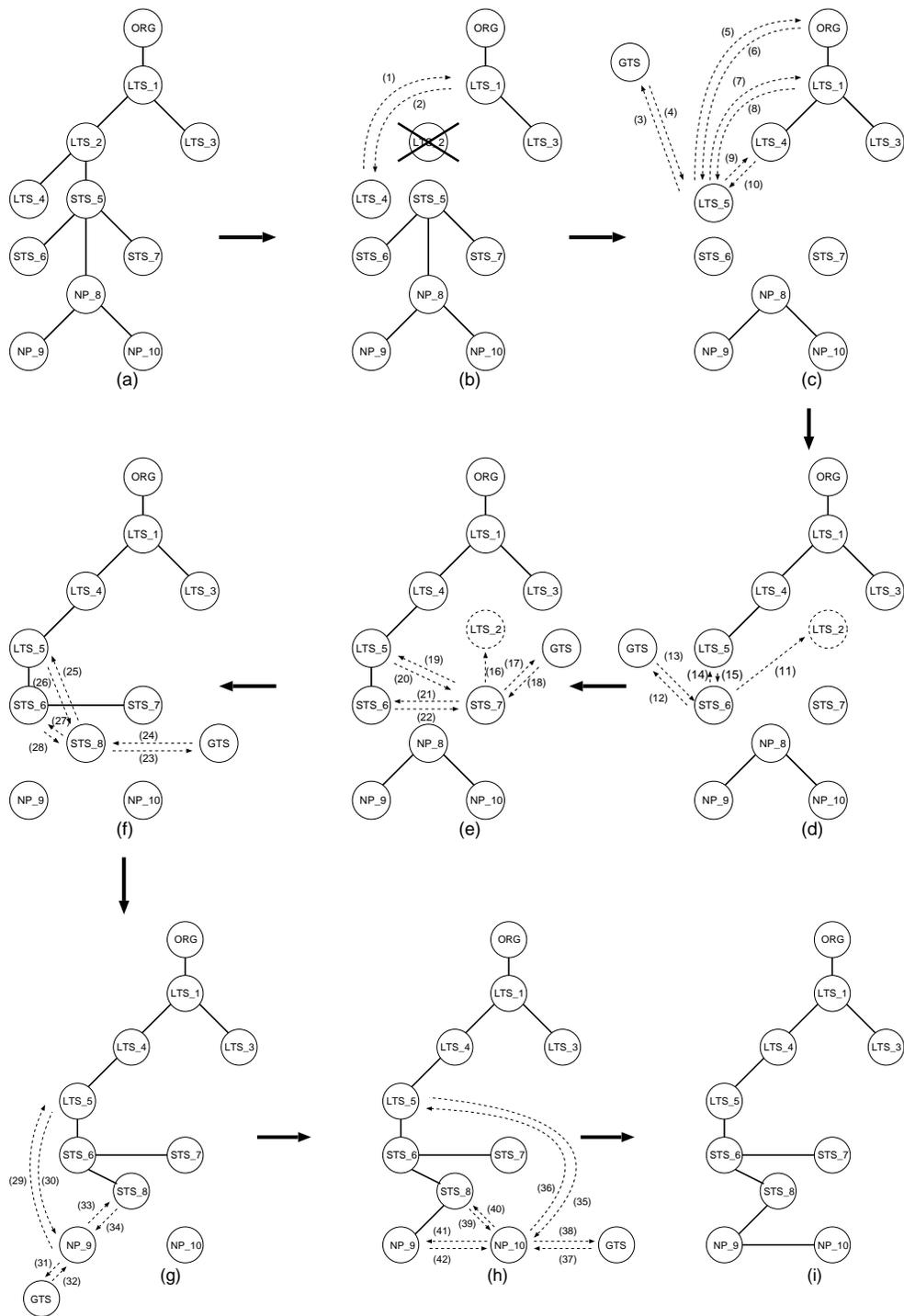


図 12: 障害回復の手順の例

## 4 シミュレーションによる評価

シミュレーションにより，3章で述べた P2P ネットワークにおける動画ストリーミング配信機構を評価する．ただし，本章では，拠点，サブネットといった物理構成を考慮した配信ツリーを構築するが，ピア間の接続については論理的なものを扱い，物理的なネットワークにおけるリンクとは異なる．

### 4.1 評価尺度

シミュレーションにおいては，GTS やピアの負荷，障害回復時間，再生開始までの待ち時間，再生の途切れ時間，バッファ内データ量について，動画配信要求の到着率，拠点数やサブネット数，および障害発生率やファンアウト数の制限，ビデオの長さの影響を評価する．

1 秒間あたりに GTS が受信する配信ツリー参加要求と LTS や STS の離脱通知の総数を GTS の負荷とする．また，ピアごとの 1 秒間あたりに受信する接続要求の最大数をピアの負荷とする．

仮親へのアクセス失敗や，親の離脱を認識してから，新たに親への接続を確立するまでの時間を障害回復時間とする，また，SS にスケジュール決定要求を送信してから第 1 セグメントの受信が開始されるまでの時間，障害などによる動画配信の途切れのためバッファが枯渇して動画の再生が止まった時間の合計についても評価する．

### 4.2 シミュレーション条件

評価に際しては，以下を標準的なパラメータ設定とし，これらのうち，ピア数や拠点数などを変化させてシミュレーションを行った．

拠点数，拠点内のサブネット数は，ともに 5 とした．また，サブネットや拠点，ピアによらずピアと GTS 間，ピアと SS 間およびピアと ORG 間の往復伝搬遅延を 400 ミリ秒とした．また，拠点内，サブネット内，および拠点間，サブネット間でのピアとピア間の往復伝搬遅延はいずれも共通で 20 ミリ秒とした．ORG，ピアのファンアウト数は 3 とした．

動画ストリーミング配信では，符号化レート 1 Mbps で 186 秒の動画を図 13 のよ

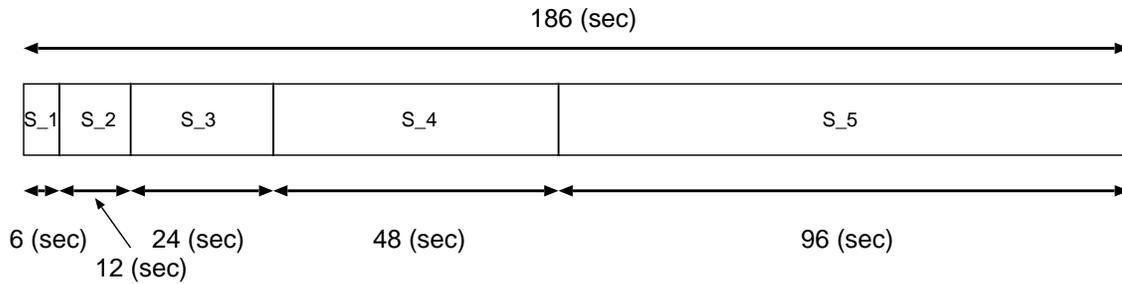


図 13: 動画像のセグメント分割

うに5つのセグメントに分割して配信する．動画像は全体で 23.25 Mbyte である．セグメント長は，第1セグメントから順に 6, 12, 24, 48, 96 秒である．ブロードキャストスケジューリングアルゴリズムとしては  $\alpha = 2$  の PB を用いる．したがって，セグメントの転送時間は，第1セグメントから順に 3, 6, 12, 24, 48 秒である．よって，要求受付時間  $W$  は 3 秒となる．また，要求受付予備時間  $C$  を 1 秒とした．

1 秒間あたりに新しく動画像配信サービスに参加し，動画像の配信を要求するピアの数を平均 30 とした．これを新規ピア到着率と呼ぶ．新しいピアはまず，受信スケジュールの決定要求を SS に送信するため，SS へのスケジュール決定要求の到着率は平均 30 となる．要求受付時間は 3 秒であるため，第1セグメントのあるスロットの配信ツリーに参加するピア数の平均は  $30 \times 3 = 90$  となる． $\alpha = 2$  であることから，第2セグメントのスロットは第1セグメントの 2 スロットに相当する．したがって，第2セグメントの配信ツリーには第1セグメントの 2 スロット分のピアが参加するため，平均ピア数は 180 となる．同様に，第3, 4, 5 セグメントの配信ツリーに参加する平均ピア数は 360, 720, 1440 となり，ある瞬間に配信ツリーに参加しているピア数の合計は平均 2790 となる．

いずれのピアも，毎秒  $1/250 = 0.004$  の確率で，動画像視聴を完了せずにサービスから離脱する．動画像配信サービスを受けるピア数の平均は 2790 であるから，毎秒  $2790 \times (1/250)$  11 ピアがサービスから離脱していくことになる．

シミュレーションは，ORG, GTS, SS のみの状態から開始した．ピア数の少ない初期状態の影響を除外するため，以降のシミュレーション結果においては，シミュレーション開始

表 2: 標準的なパラメータ設定でのシミュレーション結果

|                     | 最大   | 最小     | 平均     |
|---------------------|------|--------|--------|
| GTS の負荷             | 611  | 13     | 87.3   |
| ピアの負荷               | 49   | 0      | 4.81   |
| 障害回復時間 (秒)          | 1.82 | 0.0200 | 0.147  |
| 再生開始までの待ち時間 (秒)     | 5.35 | 1.44   | 2.94   |
| 再生の途切れ時間 (秒)        | 0.98 | 0      | 0.0019 |
| 最大バッファ内データ量 (MByte) | 12.1 | 6.01   | 9.14   |

から 144 秒後を 0 とし, 1440 秒間のふるまいについて示す。なお, 再生開始までの待ち時間や再生の途切れ時間については, 配信途中で離脱したピアや, シミュレーション終了時に視聴途中であったピアを除き, 動画を最後まで視聴したピアのみを対象とした。また, 同じ条件のもとでシミュレーションは 10 回ずつ行い, シミュレーション結果は, それらの平均値を用いた。

#### 4.3 基本特性の評価

前節で述べたパラメータ設定でのシミュレーションを行った結果を, 表 2 にまとめる。

図 14 に, GTS の負荷の変化を示す。図より, 48 秒周期の大きな負荷や 24 秒周期, 12 秒周期など周期的な負荷増大が見られる。これは, 全セグメントの配信開始時刻が一致し, 要求を提出するべき時間の開始が重なるのが 48 秒周期であること, 第 1~第 4 セグメント間での同様の同期が 24 秒周期であること, などによるものである。特に, 第 3.4 節で述べた通り, 第 2 セグメント以降については, 先のセグメントでの LTS や STS であったピアが要求受付開始と同時に配信ツリー参加要求を送信するため, 瞬間的な負荷増大を招く。

図 15 に, ピアごとの再生開始までの待ち時間の分布を示す。ピアの配信ツリーへの参加要求時刻は, 要求受付時間内でランダムに決定されるため, 待ち時間分布は一様になる。要求受付開始と同時に SS に要求が届いたピアについて考える。ピアの送出したスケジュール

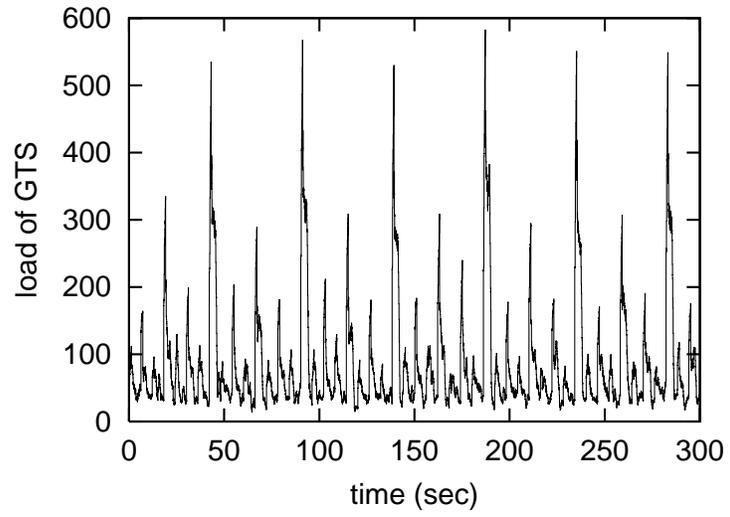


図 14: GTS の負荷の変動

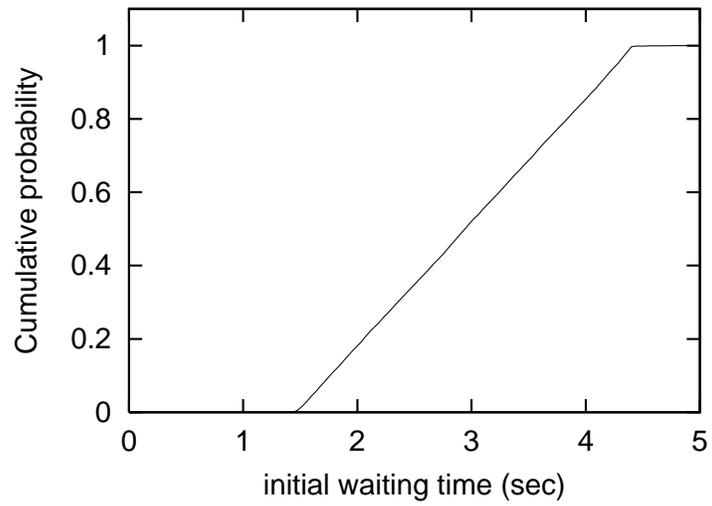


図 15: 再生開始までの待ち時間の分布

決定要求が、SS に届くまでには、SS-ピア間の往復伝搬遅延の半分である 0.2 秒の遅延が発生する。その後、動画像の配信が開始されるまでに  $W + C = 4$  秒の待ち時間が生じる。なお、この間に GTS へのツリー参加要求や仮親の紹介、接続が行われ、配信ツリーが構築される。要求受付開始と同時に動画像の配信を要求することから、このピアは、LTS となり、ORG の直接の子となると考えられる。したがって、ORG からの配信による遅延は、0.2 秒である。これらを合計すると 4.4 秒となり、これが正常動作における最大値である。表 2 において、再生開始までの最大待ち時間が 4.4 秒以上であるのは、動画像受信直前の親 LTS の離脱により、LTS が障害回復を行い、動画像の受信が遅れたためである。一方、要求受付終了直前に SS にスケジュール決定要求が届いたピアにおいては、SS にスケジュール決定要求が届く 0.2 秒の遅延、および要求受付予備時間の 1 秒の待ち時間が生じる。また、このピアは NP になり、配信ツリーの葉となると考えられるため、第 1 セグメントの配信遅延が、ORG-LTS 間で 0.2 秒、加えて他のピア間の遅延が数十ミリ秒発生する。再生開始までの待ち時間の最小値は 1.44 秒となっており、おおよそ計算と一致していることが分かる。

図 16 に、ピアごとの再生の途切れ時間の分布を示す。1440 秒間のシミュレーションにおけるサービスに参加したピアの数は、 $1440 \times 30 = 43200$  である。一部のピアは途中で離脱したため、計測対象は、半分弱のおよそ 20000 ピアであった。図より、再生の途切れの発生したピアは約 0.3%、すなわち約 60 ピアである。バッファ内データ量が 0 である LTS が離脱した場合、同様にバッファにデータが蓄積されていない直下の子の STS は、障害回復によりおよそ 0.8 秒の再生の途切れ時間が生じる。また、STS が障害回復を行っている間、その子孫ピアは動画像を受信することができないため、同様に 0.8 秒の再生の途切れが発生する。そのため、再生の途切れが 0.8 秒付近であるピアが多く存在している。

図 17 に、ピアごとの最大バッファ内データ量の分布を示す。バッファ内データ量の変化は、どのスロットでセグメントを受信したかに大きく依存するため、値は離散的になる。配信スケジュールは周期的であり、その周期は第 5 セグメントの送信時間に等しく 48 秒である。第 1 セグメントの送信時間は 3 秒であることから、受信スケジュールは 16 通りになる。図 18 に、最大バッファ内データ量が最大および最小となるスケジュールをそれぞれ示す。網かけされたスロットのセグメントを受け取った場合に、最大バッファ内データ量は最

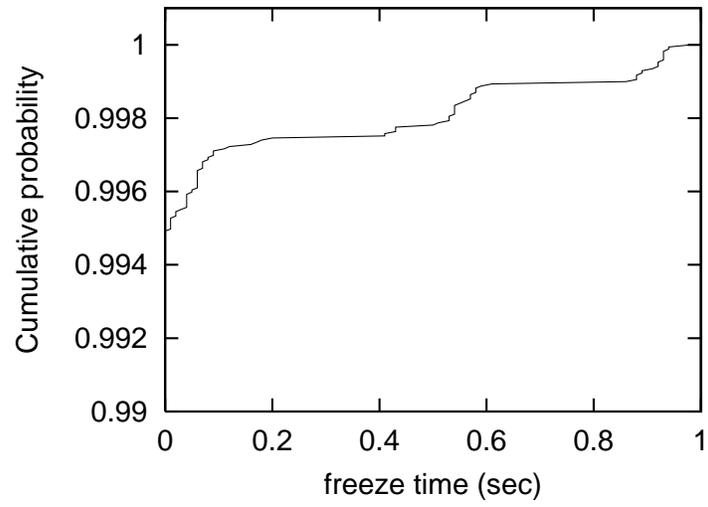


図 16: 再生の途切れ時間の分布

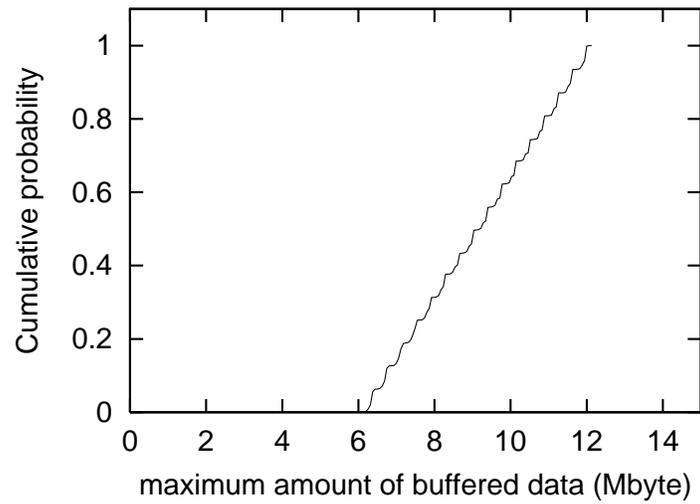


図 17: 最大バッファ内データ量の分布

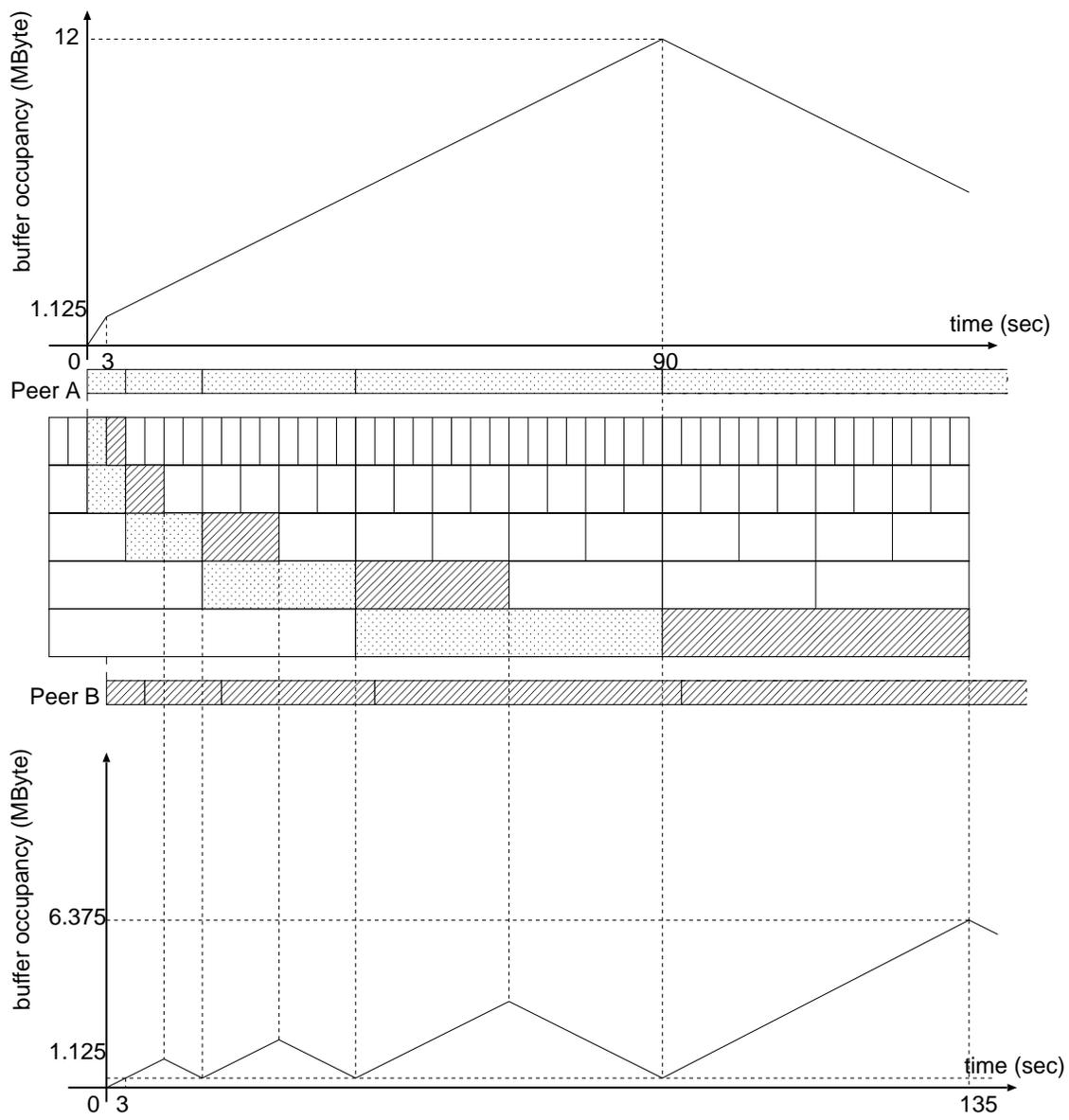


図 18: バッファ内データ量の変化

大となり、最大値は 12 Mbyte となる。一方、斜線のスロットのセグメントを受け取った場合に、最大バッファ内データ量は最小となり、最小値は 6.375 MByte となる。動画像データの大きさは 23.25 Mbyte であることから、動画像データの半分の容量のバッファが必要であることが分かる。

#### 4.4 ピア数の変化の影響

新規ピア到着率を、毎秒 10 から 70 まで変化させた場合のシミュレーション結果を図 19 に示す。ある時刻における全配信ツリーの参加ピア数は、930 から 6510 となる。

ピアのサービス参加は、第 1 セグメントに対する要求受付時間内で一様分布のため、GTS、ピアの負荷は、新規ピア到着率にほぼ比例して増加する。一方、平均障害回復時間はピア数の増加にしたがって減少する。本シミュレーションでは、ピア数によらず、拠点数、拠点あたりサブネット数はいずれも 5 としているため、ピア数の増加によりそれぞれのセグメント内の NP の数が増える。LTS や STS に障害が発生すると、その直下の子である STS または NP は障害回復のために GTS への問い合わせが必要となり、障害回復までに最短で 400 ミリ秒かかる。一方、NP に障害が発生した場合には、その子が祖父母に接続できれば 20 ミリ秒での障害回復が可能である。したがって、平均障害回復時間は減少する。

ピア数の増加により配信ツリーが深くなると、配信遅延が大きくなり、再生開始までにより多くの時間が必要となる。しかしながら、本シミュレーションではファンアウト数が 3 であることから、深さが 1 大きくなるためにはピア数が 3 倍増加しなくてはならず、また、遅延の増分は 10 ミリ秒である。よって、再生開始までの待ち時間の平均値は、ピア数によらずほぼ一定である。一方、ピア数の増加にともない、動画像配信直前の親の離脱が生じやすくなるため、障害回復のために最大遅延は大きくなっている。

再生の途切れに関する評価結果は、2 つのグラフに示す。図 16 に示したとおり、ほとんどのピアは再生の途切れを経験しない。平均途切れ時間は 3 ミリ秒よりも小さく、知覚できない程度である。一方、最大の途切れ時間は、新規ピア到着率の増加にともなって大きくなる。障害が発生しても、ピアのバッファに十分な量の動画像データが蓄積されていれば、動画像再生の途切れは起こらない。図 18 に示したとおり、第 1 セグメントの受信中が最も

バッファに動画データが蓄積されておらず、障害発生時に途切れが発生しやすいと考えられる。3.5 節で述べたとおり、LTS や STS に障害が発生すると、GTS へのツリー参加要求の送信が必要となる。直下の子、およびその子は障害回復に時間がかかる。新規ピア到着率が 10 のとき、第 1 セグメントの配信ツリーの参加ピア数の平均は 30 である。拠点数が 5、また、拠点内サブネット数は 5 であるから、LTS は平均 5、STS は平均 25 となり、ほとんどのサブネットには NP が存在せず、したがって多くの STS には直下の子が存在しない。新規ピア到着率の増加にともない、直下の子およびその子の存在するサブネットが増え、再生の途切れ時間が大きくなる。新規ピア到着率が 30 になると、全てのサブネットに STS の直下の子とさらにその子が存在するので、それ以上の途切れの増加がいったんおさえられる。ただし、ピア数が増えると障害回復中にさらなる障害が発生しやすくなるため、再生の途切れが大きくなる。

最大バッファ内データ量は、ピア数の変化の影響を受けない。

#### 4.5 物理網構成の影響

拠点数、拠点内サブネット数を 5~14 に変え、シミュレーションを行った結果を図 20 に示す。拠点数、拠点内サブネット数の 2 つは同じ値とし、横軸はその値を示す。

4.3 節で述べたとおり、GTS の負荷は全てのセグメントのツリー参加要求受付が同時に開始される時に最も高くなる。第 2 セグメント以降については、先のセグメントで LTS または STS であったピアがツリー参加要求を送信するため、GTS の負荷の最大値は、拠点数、拠点内サブネット数の到着率の増加にともなって大きくなる。

ピアの負荷の最大値は、LTS、STS の数が 11 のときに最小となっている。STS の増加にともない、LTS の負荷は増大する。一方、総ピア数は変わらないため、サブネットあたりの NP の数が減少し、STS の負荷は低下する。LTS、STS の数が 11 までは STS の負荷が、さらに数が増えると LTS の負荷が、それぞれピアの最大負荷として現れている。

また、前章で述べた通り、LTS、STS の直下の子、およびその子の障害回復には時間がかかるため、LTS、STS の増加にしたがい障害回復時間の平均値は大きくなる。最大値は、障害の発生の仕方によって異なるため不安定であるが、おおむね LTS、STS の増加にとも

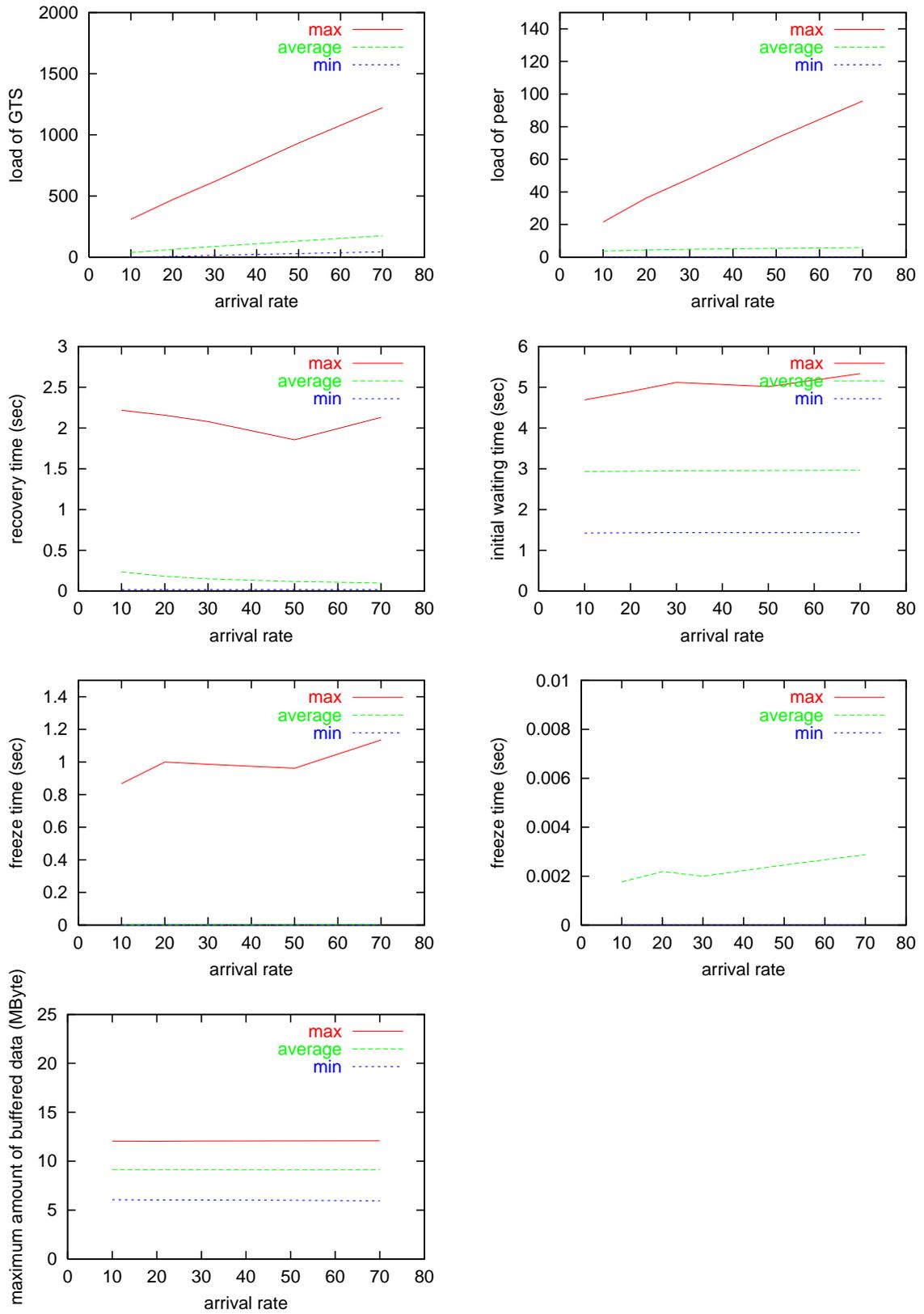


図 19: 新規ピア到着率の影響

なって大きくなる。

LTS, STS の増加により, 配信ツリーが浅くなるため, 再生開始までの待ち時間は減少する。

前節でも述べた通り, 第 1 セグメントの受信中に LTS または STS の障害が発生したとき, 直下の子およびその子において, 最も動画像の途切れが発生しやすい。LTS, STS の増加にともない, それら直下の子やその子が減るため, 再生の途切れは減少する。

最大バッファ内データ量は, ピア数の変化の影響を受けない。

#### 4.6 障害発生率の影響

ピアあたりの毎秒の障害発生率を 0 ~ 0.005 まで変化させ, シミュレーションを行った結果を図 21 に示す。障害発生率 0.005 の場合, 最後まで動画像を視聴するピアの割合はおよそ 39% である。

障害の増加にともない, LTS, STS の離脱の通知を受ける回数が増えるため, GTS の負荷は増加する。障害の影響は, 離脱したピアの周辺ピアにしか及ばず, また, ピア数は非常に多いため, ピアの負荷はほぼ変わらない。

障害回復中に再び障害の影響を受けると, 障害回復に要する時間は長くなる。そのため, 障害発生率の増加にともない, 障害回復時間の最大値は増加する。ただし, その確率は低いいため, 平均値はほとんど変わらない。また, 再生開始までの待ち時間, 再生の途切れ時間ともに, 障害回復時間の増加によって大きくなる。

最大バッファ内データ量は, 障害発生率の変化の影響を受けない。

#### 4.7 ファンアウト数の影響

ファンアウト数を 2 ~ 10 に変えて, シミュレーションを行った結果を図 22 に示す。

ファンアウト数が大きくなると, あるピアの障害に対して障害回復を行うピアの数が増える。そのため, ピアの負荷はファンアウト数の増加にともない増加する。また, 障害回復中に再び障害に出会う確率も高くなるため, 障害回復時間, 再生の途切れ時間も増加する。

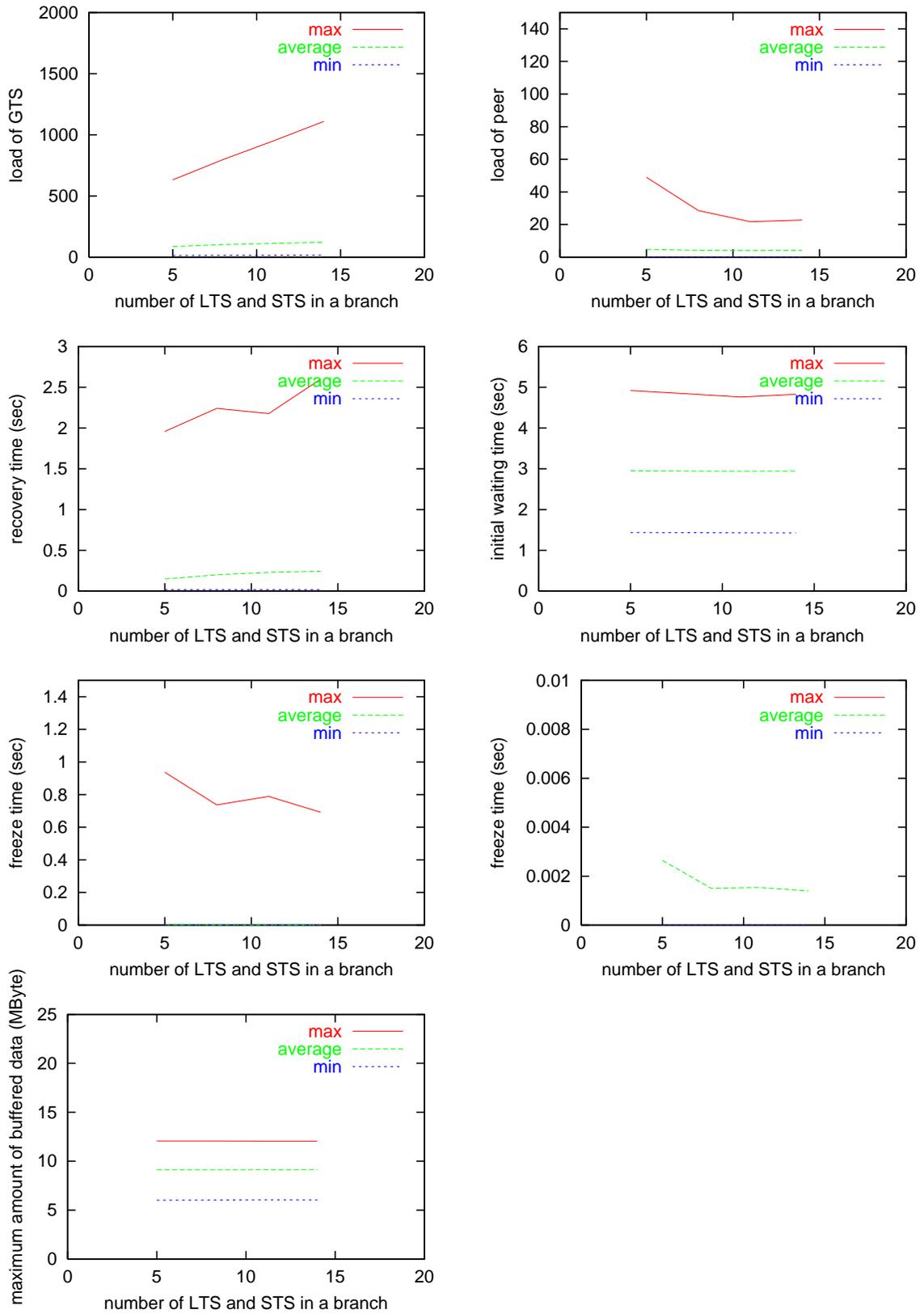


図 20: 拠点数, 拠点内サブネット数の影響

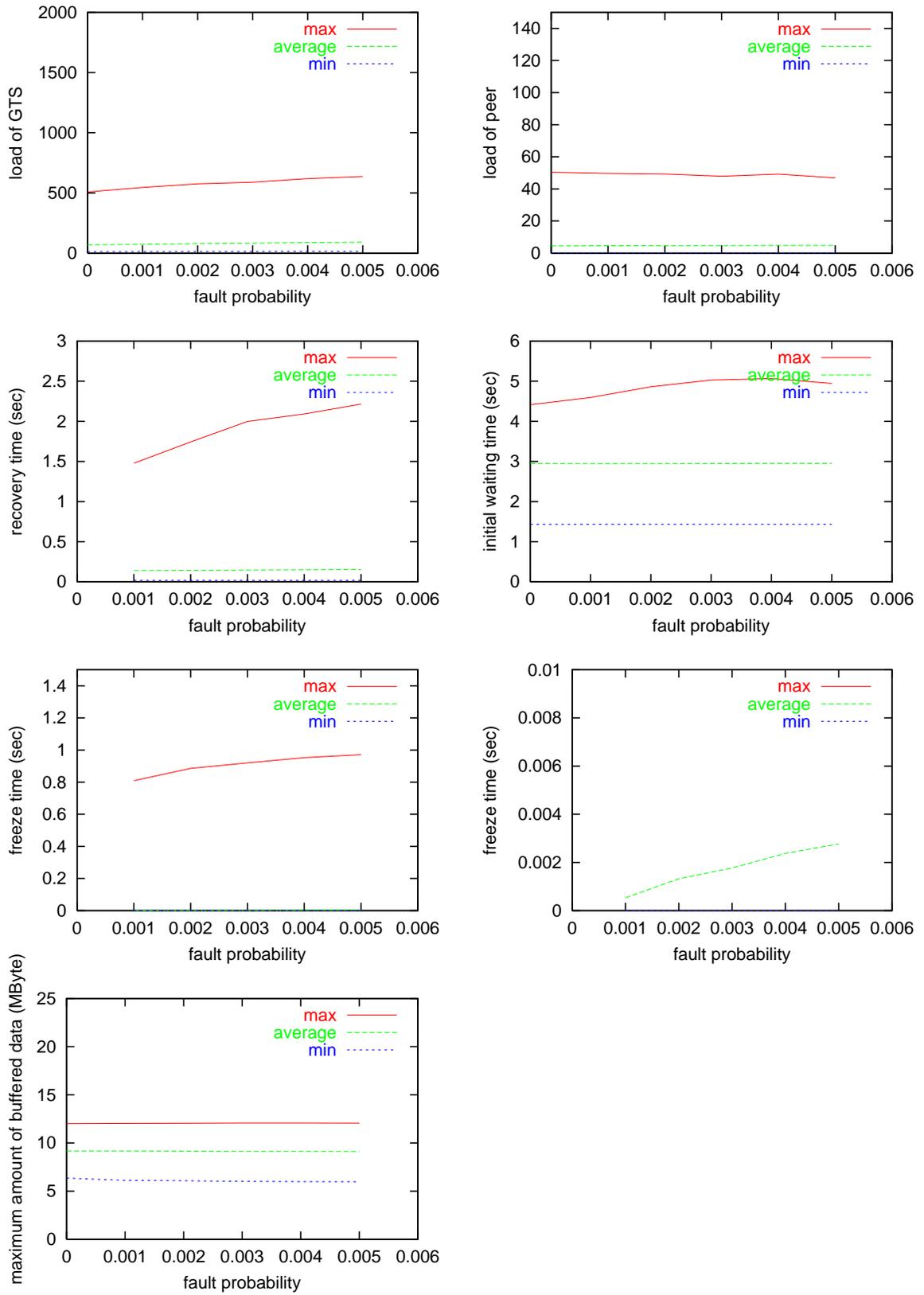


図 21: 障害発生率の影響

ファンアウト数の増加にともない、配信ツリーが浅くなるため、配信による遅延は減少する。しかしながら、ファンアウト数が 3 のとき、ピア数が 1000 であったとしても LTS 間の配信ツリーの深さは平均 2、STS 間の配信ツリーの深さは平均 3、NP 間の配信ツリーの深さは平均 4、深さの合計の平均は 9 であり、ファンアウト数が 10 の場合は 6 であるので、その差は小さく、再生開始までの待ち時間はほとんど変化しない。また、拠点数やサブネット数に変化はなく、障害回復のため GTS に配信ツリー参加要求を送信する LTS や STS の直下の子の数は変わらないため、GTS の負荷は変わらない。

最大バッファ内データ量は、ファンアウト数の変化の影響を受けない。

#### 4.8 動画像の長さの影響

動画像の長さを 90 ~ 762 秒に変え、シミュレーションを行った結果を図 23 に示す。ただし、第 1 セグメントの長さは 6 秒で固定であり、動画像の長さの増加にともない、セグメント数を増やした。動画像の長さがおよそ 2 倍になると、セグメント数は 1 増える。セグメント数の増加にともない、配信ツリー参加要求も増えるため、GTS の最大負荷は増加する。

第 1 セグメントの長さは固定のため、セグメント数の増加にともない最終セグメントは長くなり、最終セグメントの配信ツリーは深くなる。ただし、拠点および拠点あたりのサブネットの数は変わらず 5 であるため、NP の数のみが増える。その結果、STS の負荷の増加によって、ピアあたりの負荷も大きくなる。障害回復にそれほど時間を要しない NP の数が増えるため、障害回復時間の平均値は減少する。ただし、動画像が長くなると、ピアがシステム内に滞在する時間が長くなり、障害件数も多くなる。そのため、障害回復時間の最大値は増加している。

セグメント数が変化しても、第 1 セグメント受信までの手順は変わらない。よって、再生開始までの平均待ち時間は変化しない。また再生途切れ時間についても、3.3 節でも述べた通り第 1 セグメントの影響が大きいので、セグメント数の影響は受けない。図において、セグメント数の増加にともない再生開始までの待ち時間の最大値、再生の途切れ時間の最大値が減少しているが、これは動画像を最後まで視聴したピアに対してのみ評価を行っており、動画像が長くなったことにより、評価対象となったピアの数が減少したことによる。

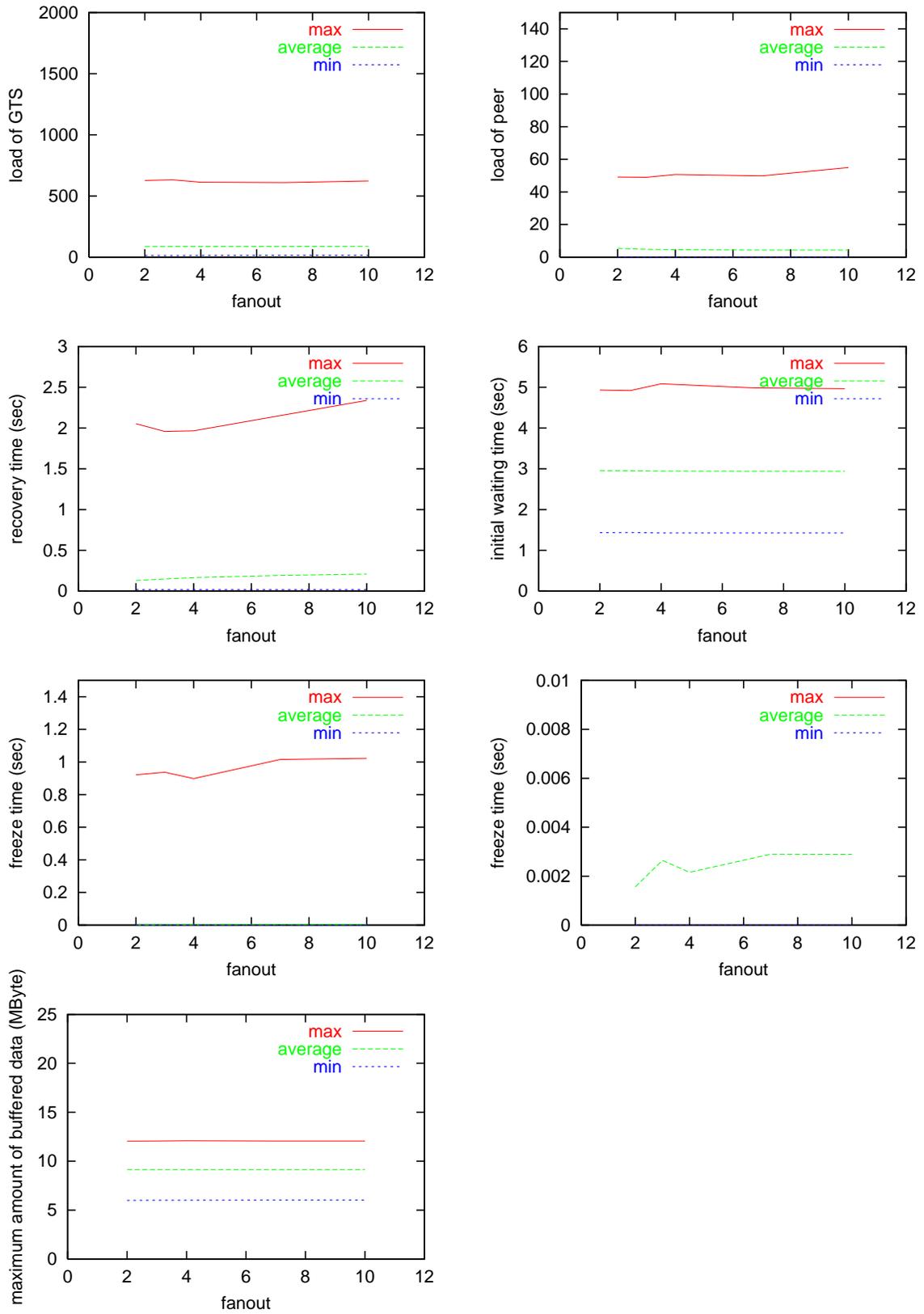


図 22: ファンアウト数の影響

最大バッファ内データ量は、ビデオの長さに比例して増加する。

#### 4.9 シミュレーション評価のまとめ

以上のシミュレーションにより、提案機構を用いることで、ピア数や障害発生率、ファンアウト数、動画像の長さの変化によらず利用者に再生の途切れをほとんど知覚させない動画像データの配信が行えることが分かった。また、ピラミッドスケジューリングアルゴリズムにより、再生開始までの待ち時間は平均 3 秒以下、最大でも約 5 秒となり、ピア数などの影響を大きく受けないことが示された。

しかしながら一方で、GTS やピアの負荷はピア数や障害発生率に対して線形に増加している。特にピア数増加に対して GTS やピアの最大負荷は急激に増加しており、十分なスケーラビリティを達成しているとは言えない。これは、本機構が動画像配信には P2P 通信技術を用いていながらも配信ツリーの構築や配信制御をサーバに依存するハイブリッド型アーキテクチャにもとづいていることによる。したがって今後は、これらの負荷を軽減する分散型の機構を検討する必要があると考えられる。

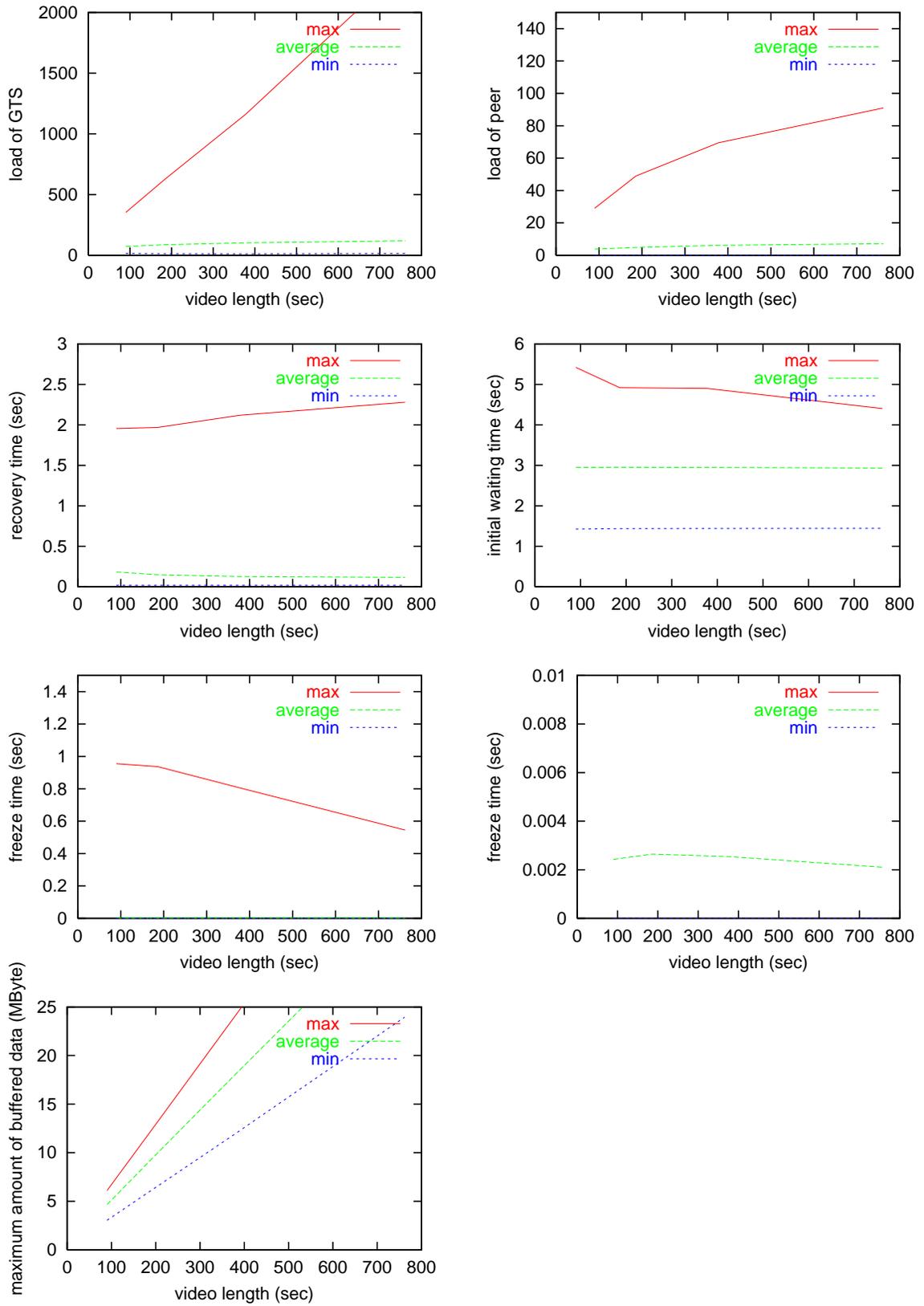


図 23: 動画の長さの影響

## 5 物理ネットワークを考慮した動画ストリーミング配信機構の評価

4章でのシミュレーションにおいては、拠点やサブネットといった物理的なネットワーク構成にもとづく制御を行う一方で、いずれのピア間でも往復伝搬遅延が 20 ミリ秒であるなど、論理的なネットワークでの評価を行っており、実際の物理的なネットワークトポロジについては考慮しておらず、また、動画配信がネットワークに与える影響についても評価していない。

ある組織において、広域 VPN サービスを利用して図 24 (a) のようなスター型（ハブ-スポーク型）のネットワークで拠点間を接続しているものとする。このネットワーク上に、図 24 (b) のような配信ツリーが構築されたものとする。図 24 (b) の論理的な配信ツリーでは、動画データは  $ORG \rightarrow LTS A \rightarrow LTS B$  と順次転送されて行く。一方、図 24 (a) に示す物理網においては、 $ORG \rightarrow HUB \rightarrow LTS A \rightarrow HUB \rightarrow LTS B$  と、動画データはハブと LTS A 間のリンクを 2 度通る。拠点内、サブネット内の帯域と比較して、多くの場合ネットワーク事業者との契約により接続されている拠点間のリンクは帯域が狭く、また、高価である。したがって、配信ツリーの構築にあたっては拠点、サブネットといった物理的なネットワーク構成だけでなく、物理ネットワークについても考慮しなければならない。

本章では、物理ネットワークを考慮したシミュレーション評価によりハイブリッド型 P2P 動画ストリーミング配信機構の問題点を明らかにする。

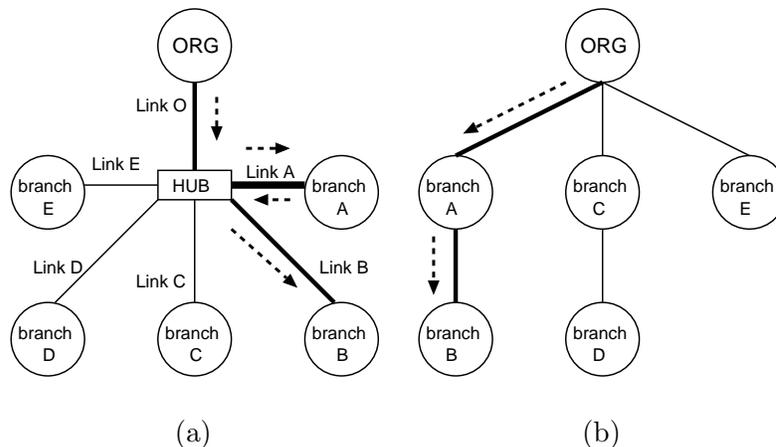


図 24: 物理ネットワークと論理ネットワーク

## 5.1 シミュレーション条件

拠点間の物理ネットワークは、図 24 (a) に示すスター型とする。ORG とハブ間のリンクをリンク O、拠点 A~E とハブ間のリンクをリンク A~E と表記する。それぞれのリンクの片方向伝搬遅延は 100 ミリ秒である。また、拠点内、サブネット内では帯域が十分あることから、メッシュ型のネットワークであるものとする。ピア間の往復伝搬遅延は、前章と同じく 20 ミリ秒とする。また、前章同様、符号化レート 1 Mbps で 186 秒の動画を、 $\alpha = 2$  のピラミッドスケジューリングアルゴリズムに従い 5 つのセグメントに分割、配信するものとする。新規ピア到着率を 30、障害発生率を 0.004、要求受付時間を 3 秒、要求受付待ち時間を 1 秒とした。

## 5.2 物理ネットワークを考慮したシミュレーション評価

4.1 節の評価尺度に加え、リンクの負荷として 1 秒間あたりに拠点間のリンクに流れる動画データの量について評価した結果を、表 3 にまとめる。

前章では全てのピア間での片方向伝搬遅延は 10 ミリ秒であったが、物理ネットワークを考慮することにより LTS 間の遅延が 0.19 秒長くなったため、図 24 (b) における拠点 B、D 内のピアの再生までの待ち時間が 0.19 秒増加した。その結果、再生開始までの平均待ち時間は  $0.19 \times (2/5) = 0.076$  長くなっている。また、LTS 間ツリーの障害回復に時間がかかるため、障害回復時間、再生の途切れ時間も増加している。ただし、GTS やピアの負荷、最大バッファ内データ量は、LTS 間の遅延の影響を受けない。

図 25 に、リンクの負荷の変化の様子を示す。符号化レート 1 Mbps、 $\alpha = 2$  であることから、セグメントあたりの配信レートは 2 Mbps である。動画は 5 セグメントに分割されているため、子を持たない拠点においてはリンクの負荷は受信セグメント数によって 2~10 Mbps となる。一方、子を持つ拠点では、リンクを動画データが往復するため、4~20 Mbps となる。全ての拠点において全てのセグメントを受信しているため、拠点ごとのリンクの負荷は 10~20 Mbps の間を変動している。また、ORG のファンアウト数は 3 であるため、ORG のリンクの負荷は、おおよそ 30 Mbps で一定である。

表 3: 物理網トポロジを与えた場合のシミュレーション結果

|                      | 最大   | 最小     | 平均     |
|----------------------|------|--------|--------|
| GTS の負荷              | 616  | 14     | 86.6   |
| ピアの最大負荷              | 51   | 0      | 4.79   |
| 障害回復時間 (秒)           | 2.42 | 0.0200 | 0.157  |
| 再生開始までの待ち時間 (秒)      | 5.24 | 1.44   | 3.03   |
| 再生の途切れ時間 (秒)         | 1.47 | 0      | 0.0041 |
| 最大バッファ内データ量 (Mbyte)  | 12.2 | 5.91   | 9.12   |
| リンク O の負荷 (Mbit/sec) | 35.0 | 25.7   | 29.7   |
| リンク A の負荷 (Mbit/sec) | 21.8 | 7.88   | 14.0   |
| リンク B の負荷 (Mbit/sec) | 21.9 | 7.81   | 14.0   |
| リンク C の負荷 (Mbit/sec) | 21.5 | 8.07   | 13.8   |
| リンク D の負荷 (Mbit/sec) | 21.7 | 7.74   | 13.8   |
| リンク E の負荷 (Mbit/sec) | 21.0 | 7.97   | 14.1   |

新規ピア到着率を 10 ~ 70 に変え、シミュレーションを行った結果を図 26 に示す。新規ピア到着率の影響は、物理トポロジを考慮しない場合 (図 19) と同様であるが、拠点間の伝搬遅延の増加により、障害回復時間、再生開始までの待ち時間、再生の途切れが大きくなっている。リンクの負荷は、ピア数の増加の影響を受けず、変わらない。これは、アプリケーションレベルマルチキャストを用いるとともに物理網構成を考慮した配信ツリーの階層化によるものである。サーバクライアントモデルでは、利用者の増加にともなってサーバ近傍のリンクの負荷が増えるが、P2P 通信を利用したマルチキャストにより、利用者数に対する拡張性が得られることが分かる。

次に、障害発生率を 0 ~ 0.005 に変化させ、シミュレーションを行った結果を図 27 に示す。物理網トポロジを考慮しない場合 (図 21) と同様の傾向を示しているが、障害回復時間、再生開始までの待ち時間、再生の途切れ時間ともに大きくなっている。また、拠点間配信ツ

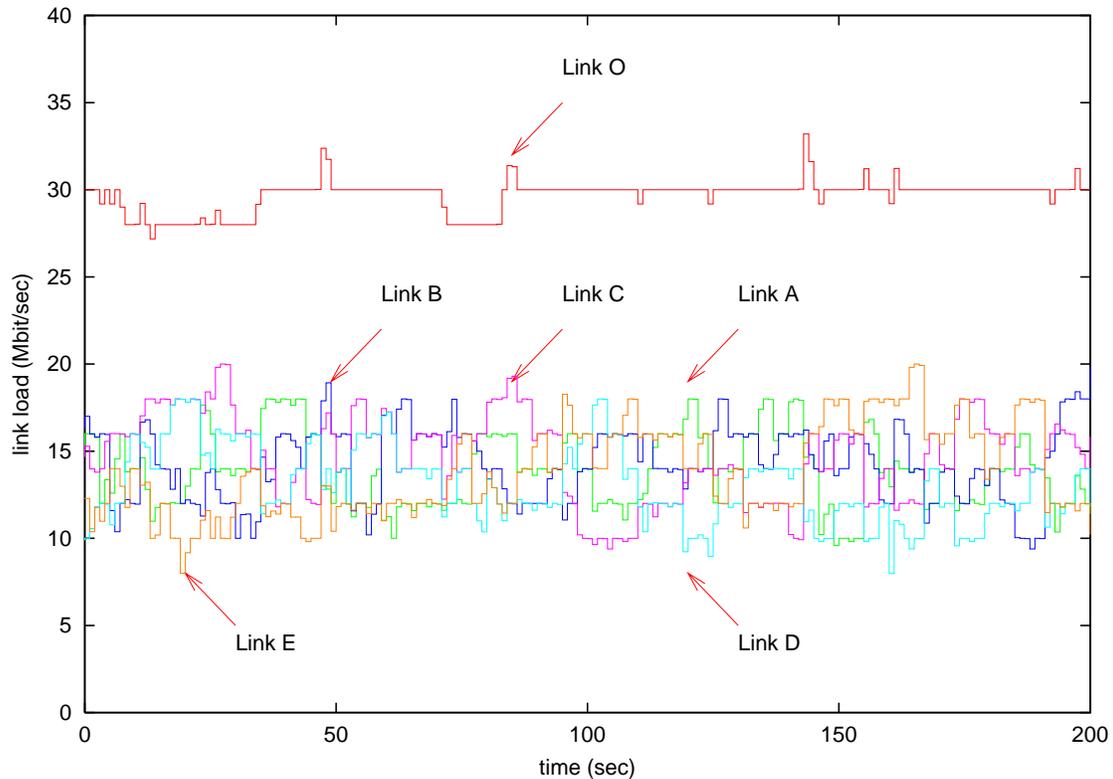


図 25: 物理ネットワークにおけるリンクの負荷

リーノ葉の拠点で障害が発生した場合と、節で発生した場合には遅延が異なるため、障害の発生仕方により、再生の途切れ時間の最大値に関しては、図 21 に比べてさらに不安定となっている。リンク負荷の平均値は、障害発生率に影響されない。しかしながら、障害の増加ともない、セグメントの配信時間にずれが生じるため、時間ごとの負荷の変動が大きくなり、最大値と最小値の差が増加する。

以上の結果より、物理ネットワークを考慮した場合にも、ピア数や障害発生率の影響をそれほど受けない動画ストリーミング配信が行えることがわかった。しかしながら、拠点間リンクの負荷の変動が大きく、また、同じ動画が往復するなど冗長な配信ツリーが構築されている。再生開始までの待ち時間や再生の途切れ時間も配信ツリーの構成の影響を受けるため、物理ネットワークを考慮した効率的な動画ストリーミング配信を行うためのツリー構築機構が必要であると考えられる。

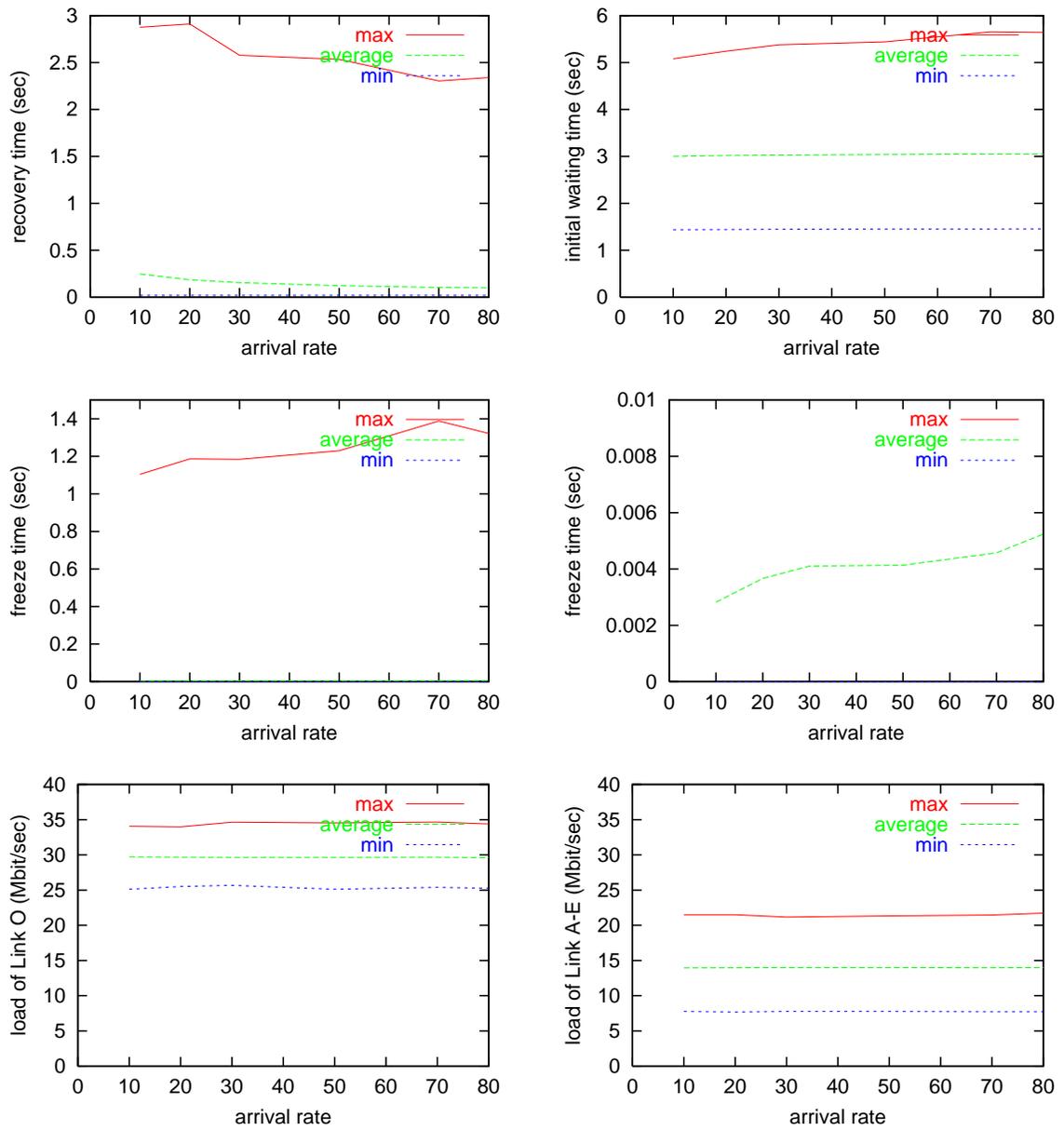


図 26: 物理ネットワークを考慮した場合の新規ピア到着率の影響

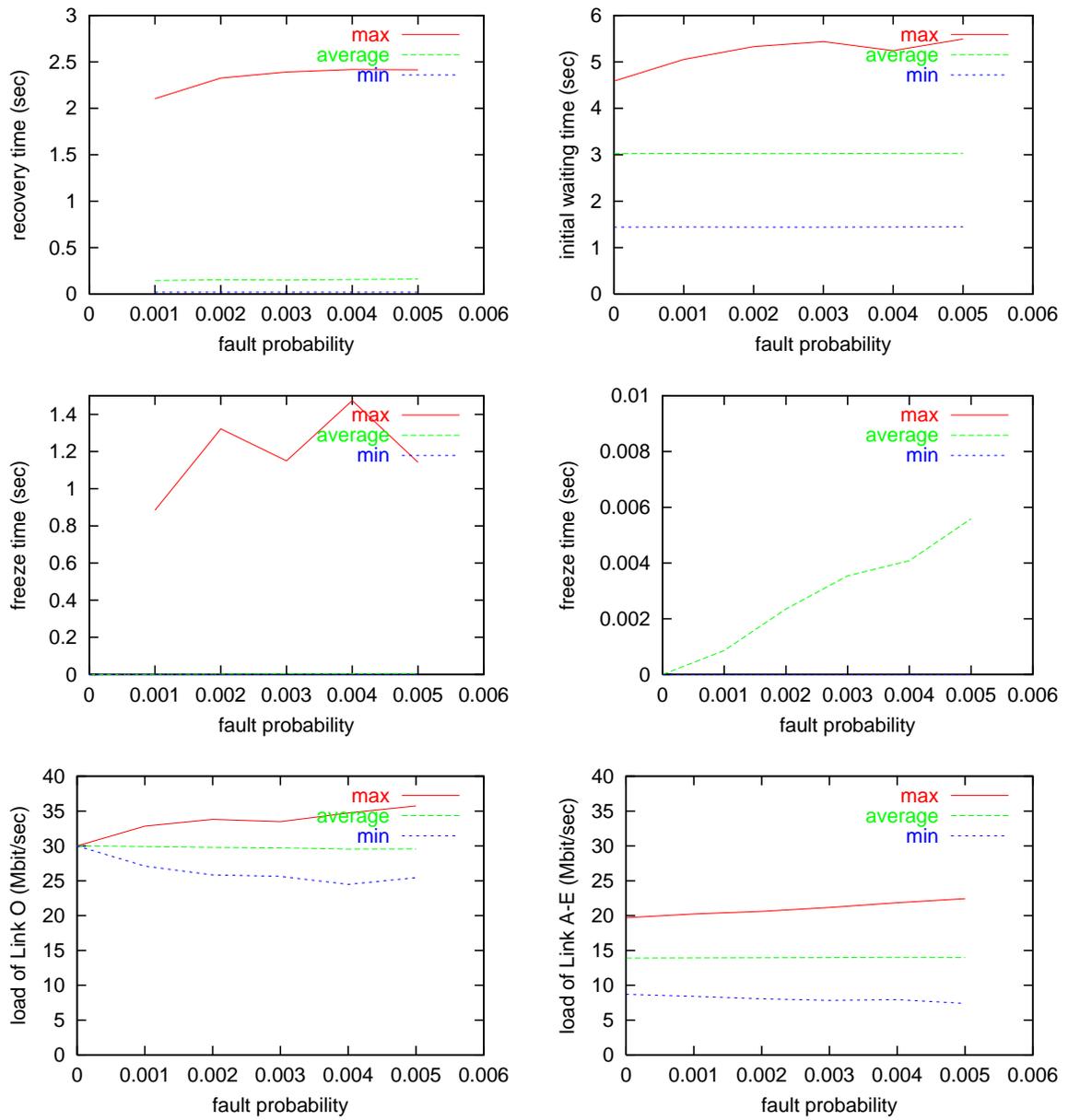


図 27: 物理トポロジを考慮した場合の障害発生率の影響

## 6 おわりに

本報告では、インターネットにおける動画像ストリーミング配信サービスにおいて、途切れない動画像の配信と、ネットワークの負荷軽減を図るための、P2P ネットワークにおける物理網構成を考慮した動画像ストリーミング配信機構について、シミュレーションによりその有効性を評価した。その結果、障害の多い環境でも、多数のピアに、途切れの小さく応答性の高い動画像ストリーミング配信を提供することが可能であることを示した。また一方で、サーバやピア、リンクの負荷が大きく、さらなる効率的な機構の開発が必要であることを明らかにした。

今後は、より複雑なネットワークでの効率の良い動画像ストリーミング配信機構について検討したい。そのため、サーバやピアの負荷を軽減、分散するメッセージ交換手順や、リンクの負荷を軽減、分散する配信ツリーの構築アルゴリズムの研究に取り組みたい。

## 謝辞

本報告を終えるにあたりまして、御指導、ご教授を頂きました大阪大学 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 村田正幸教授に深く感謝します。また、本報告において、終始御指導頂きました大阪大学 大学院情報科学研究科 情報ネットワーク学専攻 マルチメディアネットワーク講座 宮原研究室の若宮直紀助教授に心から感謝致します。

並びに、日頃より適切な助言を頂きました大阪府看護大学の菅野正嗣助教授、大阪大学大学院情報科学研究科助教授の馬場健一助教授、今瀬研究室の大崎博之助教授、大阪市立大学の阿多信吾講師、村田研究室の長谷川剛助教授、大阪大学経済学部の荒川伸一助手、宮原研究室の牧一之進助手に心から感謝致します。

さらに、共同研究をとおして多大なご支援を頂きました日本電気株式会社 村瀬勉様、谷口邦弘様に、深く感謝します。

最後に、本報告において適切な御指導と助言を頂いた笹部昌弘氏、侍建港氏、山田和広氏、谷口義明氏、吉田剛氏を始めとする村田研究室および宮原研究室の皆様方に心より御礼申し上げます。

## 参考文献

- [1] A. Hu, “Video-on-Demand Broadcasting Protocols: a Comprehensive Study,” in *Proceedings of INFOCOM 2001*, pp. 508–517, Apr. 2001.
- [2] “Napster.com.” available at <http://www.napster.com/>.
- [3] “Kazaa Media Desktop.” available at <http://www.kazaa.com/>.
- [4] I. Foster and C. Kesselman, *The Grid 2*. Morgan Kaufmann Pub, Nov. 2003.
- [5] “Gnutella.com.” available at <http://www.gnutella.com>.
- [6] “Winny Tips.” available at <http://www.nan.sakura.ne.jp/winny/>.
- [7] E. L. Abram-Profeta and K. G. Shin, “Scheduling Video Program in Near Video-on-Demand Systems,” in *Proceedings of the fifth ACM International Conference on Multimedia*, pp. 359–369, ACM Press, Nov. 1997.
- [8] K. A. Hua and S. Sheu, “Skyscraper Broadcasting: A new Broadcasting Scheme for Metropolitan Video-on-Demand System,” in *Proceedings of Special Interest Group on Data Communication (SIGCOMM '97)*, pp. 89–100, June 1997.
- [9] S. Viswanathan and T. Imilelinski, “Pyramid Broadcasting for Video on Demand Service,” in *Proceedings of the SPIE Multimedia Computing and Networking Conference*, vol. 2417, pp. 66–67, Feb. 1995.