



ネットワーク遅延とpilot fileを用いた 遠隔観測による最適サーバ選択法の研究

Research of the most suitable server selection method
by the remote observation
using a network delay and pilot file

2007.2.20

工学研究科情報通信制御システム工学専攻

発表者 5ADGM023 中岩 正洋
指導教員 石井 啓之教授



発表の流れ

1. 研究背景
2. 研究目的
3. 既存のサーバ選択法
4. 実環境におけるサーバ選択法の検討-1
5. 実環境におけるサーバ方式法の検討-2
6. 室内環境におけるネットワークサーバ性能の測定
7. まとめ



1. 研究背景

- ADSL, FTTH等のアクセス回線のブロードバンド化が進み, ユーザの数が急速に増加している.
- 映像, 音楽, EC(電子商取引)など様々なサービスが提供されるようになってきているが, それに伴いサーバに対する過度の負荷の集中が起こっている.
- ミラーリング等により, 負荷分散がなされているが, ユーザにとって最適なサーバ選択論理が明確でない.

2. 研究目的

- ミラーリングなどにより複数に分散されたサーバの中から、ユーザがサーバを選択する場合に、出来るだけ短い時間で情報を獲得できるサーバを選ぶ指標について検討する。

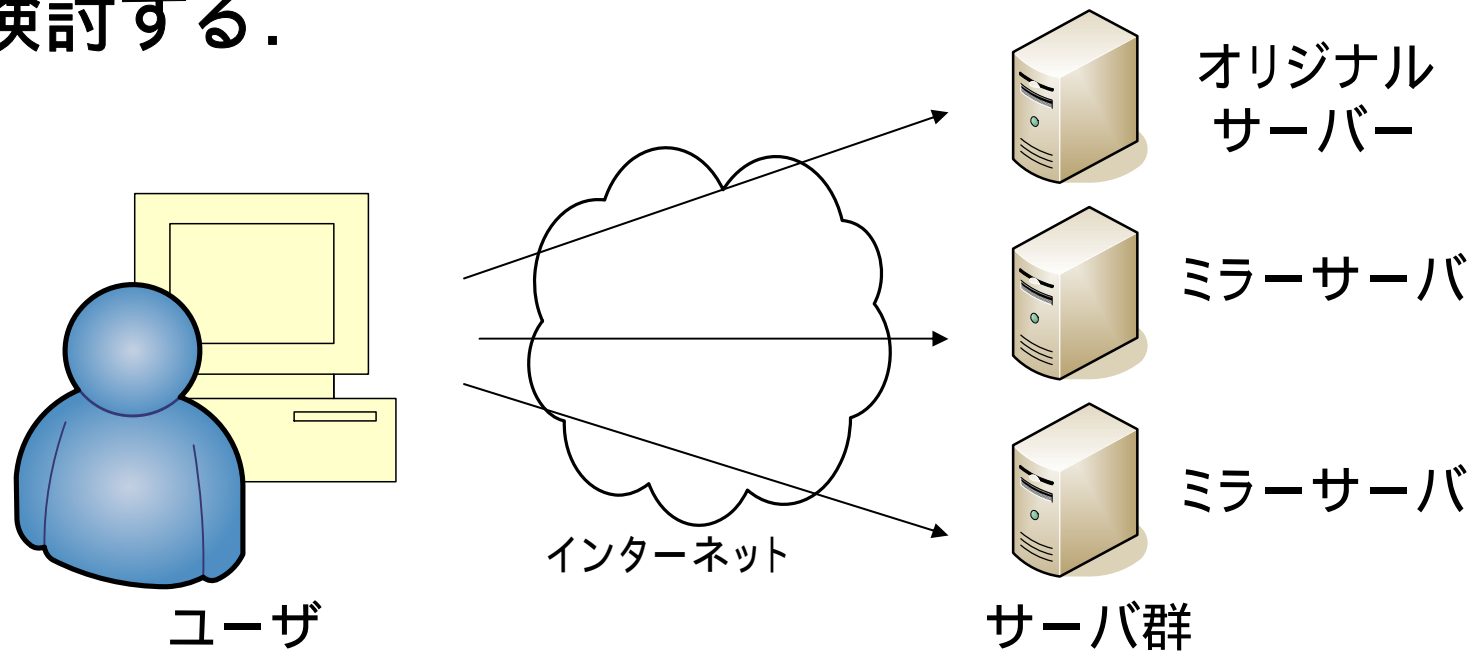


Fig-1 ダウンロード イメージ

3. 既存のサーバ選択法

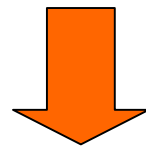
現在ユーザが同一のサービスを提供しているサーバ群から情報をダウンロードする場合, 効率よく情報を得る方法の例として以下の方法がある

1. サービス提供側で, ユーザからの要求の度に一番負荷の少ないサーバを選んで情報を提供する方法
2. ユーザ側で各サーバまでのネットワーク遅延などを測定し遅延の小さいサーバを選択する方法

3. 既存のサーバ選択法

<サーバ側のみで選択する際の問題点>

1. サービス提供側で、ユーザからの要求の度に一番負荷の少ないサーバを選んで情報を提供する際の基準となる代表的な指標として以下のものがある
 - サーバのCPU利用率
 - メモリの利用状況
 - サーバへの接続数



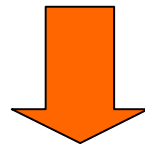
これらの指標となるものはセキュリティ等の問題でサーバ側が公開することは少なくユーザがこれらの情報を知ることは難しい。

4. 実環境におけるサーバ選択法の検討-1

<概要>

2. ユーザ側で各サーバまでのネットワーク遅延を比較し遅延の小さいサーバを選択する方法

- 日本国内で同一のサービスを提供しているサーバ群に対してICMPのping送出行を行い,RTTの測定した.



実験において日本国内のサーバのみを使用したの
で,RTTの値の差は小さくサーバ選択論理には用い
られないことがわかった.

5. 実環境におけるサーバ選択法の検討-2

<検討の概要>

2. ユーザ側で各サーバまでのネットワーク遅延を比較し遅延の小さいサーバを選択する方法
 - ネットワーク遅延 (IPレイヤのRTT) の測定 (海外サーバ)
 - Hop数 (ネットワーク経路上のルータの数) の測定
 - pilot file (ダウンロードがすぐに完了する小さなファイル) という試験的なファイルのDLT (Download Time) の測定
- これらの組み合わせにより最適なサーバ選択の指標について検討する.

5. 実環境におけるサーバ選択法の検討-2

<実験-1>

- 複数の海外のサーバからファイルサイズ5MB, 1MB, 500KB, 100KBのファイルをダウンロードし同時に各サーバに対してICMPのpingを実行しRTT, Hop数を測定.
- 目的のダウンロードしたいファイルサイズを5MBと仮定し、その5MBのDLTとpilot file(1MB, 500KB, 100KB)のDLT, RTT, Hop数との相関係数をもとめる.

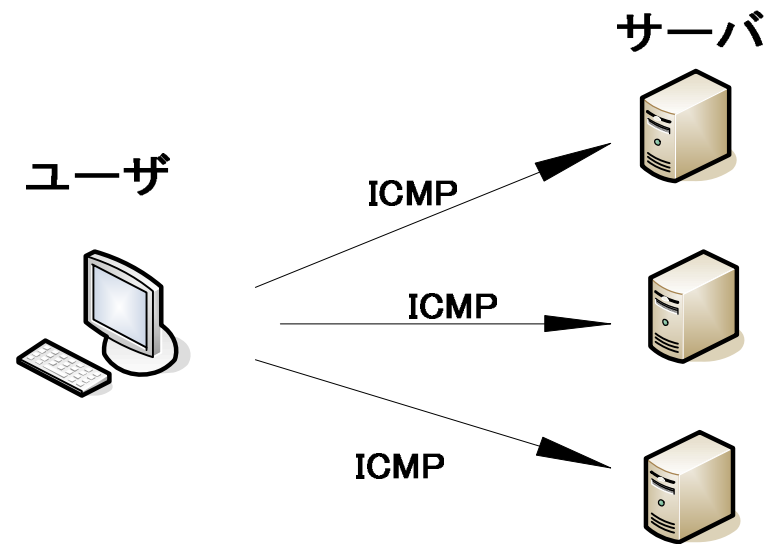


Fig-2 各サーバへの調査

5. 実環境におけるサーバ選択法の検討-2

<相関係数>

- 2つのデータ列 $x = \{x_i\}, y = \{y_i\} (i = 1, 2, \dots, n)$

の間の相関(類似性の度合い)を示す統計学的指標.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad \dots (1)$$

- -1から+1の間の実数値をとり, 1に近いときは2つのデータ列には正の相関があるといい, -1に近ければ負の相関がある. 0に近いときはもとのデータ列の相関は弱い.



5. 実環境におけるサーバ選択法の検討-2

<測定条件>

- 測定期間 : 2006年6月10日 ~ 6月17日
2006年6月24日 ~ 7月1日
- Download file size : 5MB, 1MB, 500KB, 100KB
- 回線 : ADSL回線 24Mb/s
- 測定端末 : Windows XP Professional SP2
Pentium 4 CPU 3.2GHz 1G RAM
- ダウンロード方法 : http

5. 実環境におけるサーバ選択法の検討-2

<結果(1/7)>

測定日 2006年6月10日

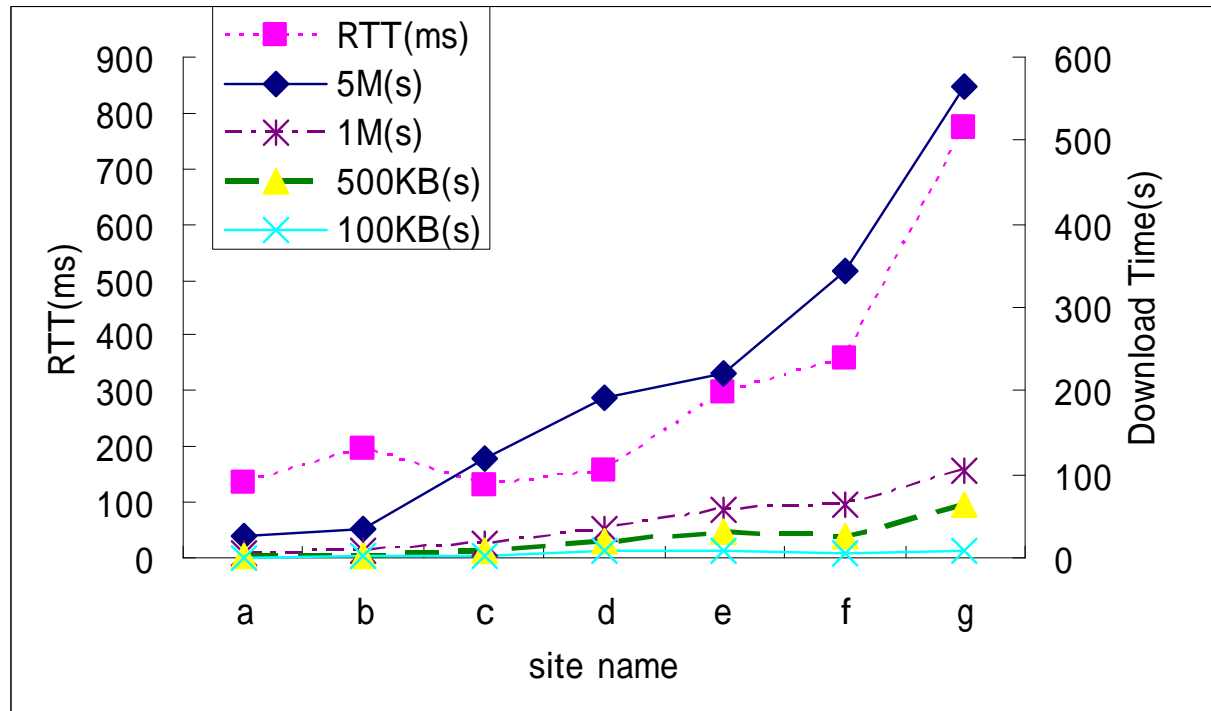


Fig-3 サーバ名とRTTとDLT

- 相関係数 = 0.929(5MBのDLT--RTT)
- 相関係数 1 = 0.981(5MBのDLT--1MBのDLT)
- 相関係数 2 = 0.969(5MBのDLT--500KBのDLT)
- 相関係数 3 = 0.883(5MBのDLT--100KBのDLT)

} pilot fileとの
相関係数



5. 実環境におけるサーバ選択法の検討-2

<結果(2/7)>

測定日 2006年6月24日

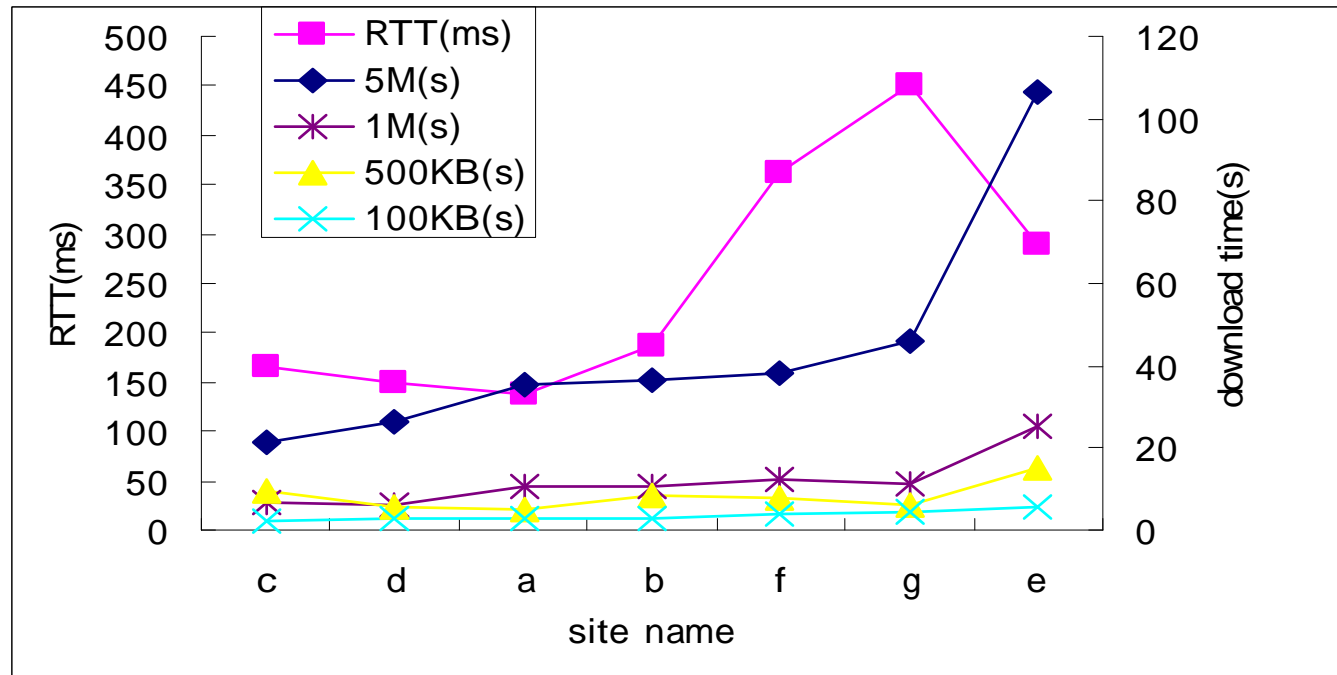


Fig-4サーバ名とRTTとDLT

相関係数 = 0.396(5MBのDLT--RTT)

相関係数 1 = 0.987(5MBのDLT--1MBのDLT)

相関係数 2 = 0.978(5MBのDLT--500KBのDLT)

相関係数 3 = 0.888(5MBのDLT--100KBのDLT)

} pilot fileとの
相関係数

5. 実環境におけるサーバ選択法の検討-2

<結果(3/7)>

- 相関係数 = 0.929(RTT)
- 相関係数 1 = 0.981(1MB)
- 相関係数 2 = 0.969(500KB)
- 相関係数 3 = 0.883(100KB)
- 相関係数 = 0.396(RTT)
- 相関係数 1 = 0.987(1MB)
- 相関係数 2 = 0.978(500KB)
- 相関係数 3 = 0.888(100KB)

Fig-3における相関係数

Fig-4における相関係数

- Fig-3,4より

RTTと5MBのDLTとの相関係数は日時によって、変動があり安定しているとはいえないが、各pilot fileのDLTと5MBとの相関係数は安定しており、1に近い値をとることが多い。また、pilot fileのサイズが小さくなるにつれて相関係数が小さくなっていく。



5. 実環境におけるサーバ選択法の検討-2

<結果(3/7)>

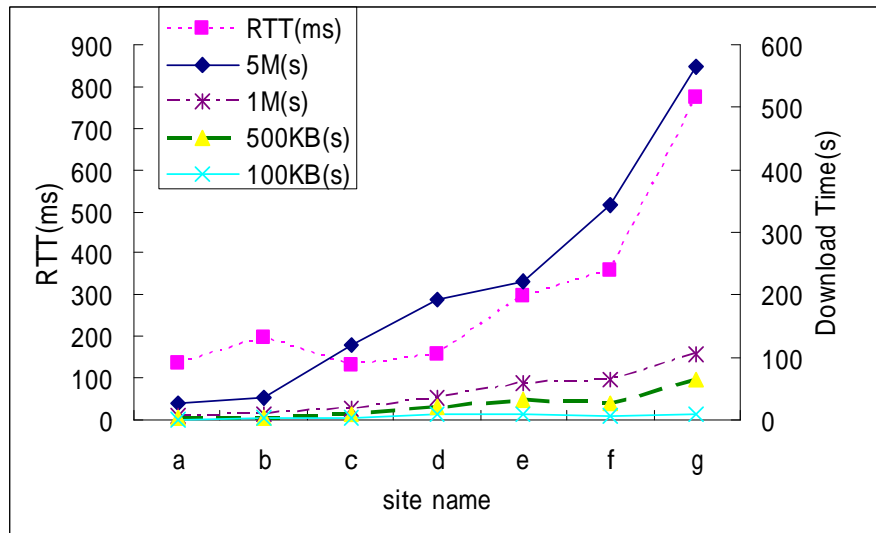


Fig-3サーバ名とRTTとDLT

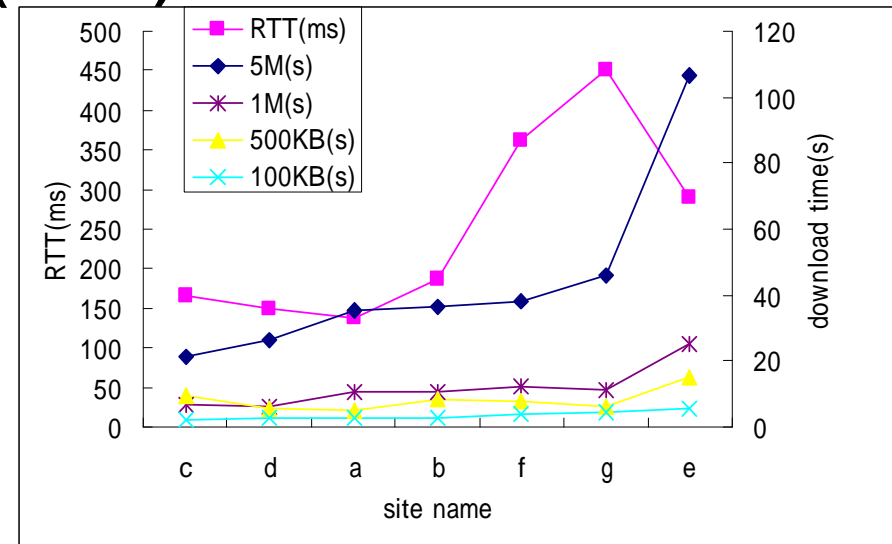


Fig-4サーバ名とRTTとDLT

- Fig-3,4より

RTTと5MBのDLTとの相関係数は日時によって、変動があり安定しているとはいえないが、各pilot fileのDLTと5MBとの相関係数は安定しており、1に近い値をとることが多い。また、pilot fileのサイズが小さくなるにつれて相関係数が小さくなっていく。

5. 実環境におけるサーバ選択法の検討-2

<結果(4/7)>

測定日 2006年6月10日

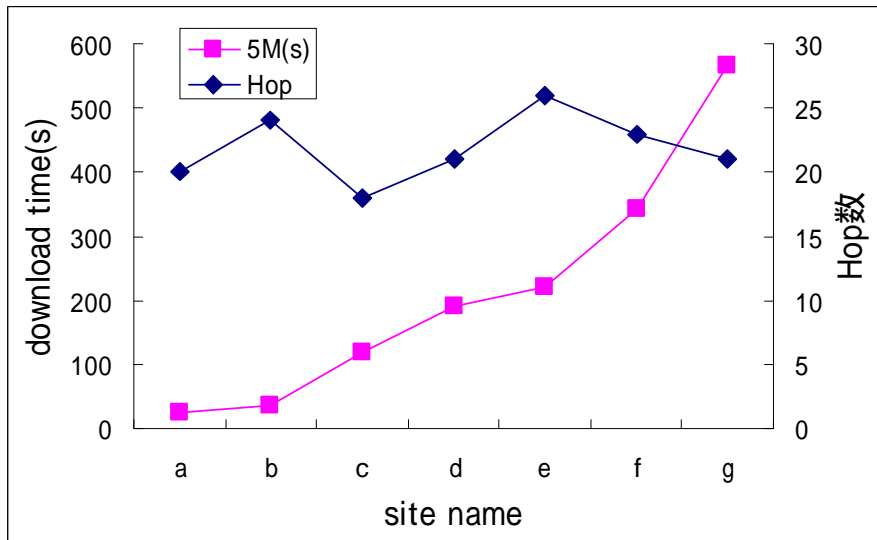


Fig-5 サーバ名とDLTとHop数

測定日 2006年6月24日

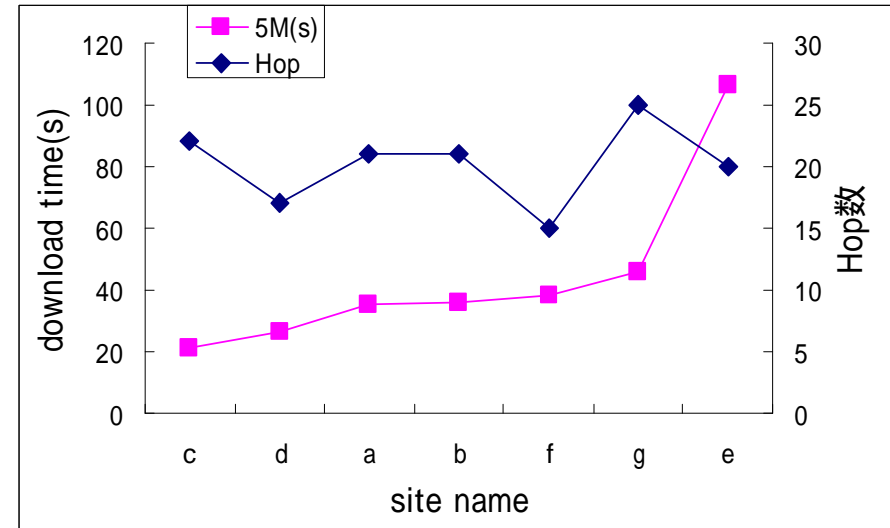


Fig-6 サーバ名とDLTとHop数

相関係数 1 = -0.107 (5MBのDLT--Hop数)—Fig-3

相関係数 2 = 0.051 (5MBのDLT--Hop数)—Fig-4

5. 実環境におけるサーバ選択法の検討-2 <結果(5/7)>

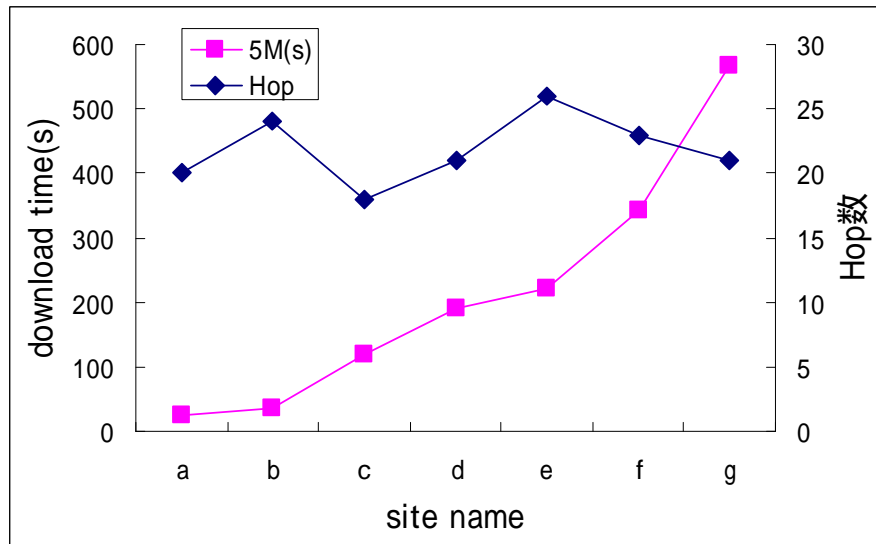


Fig-5サーバ名とDLTとHop数

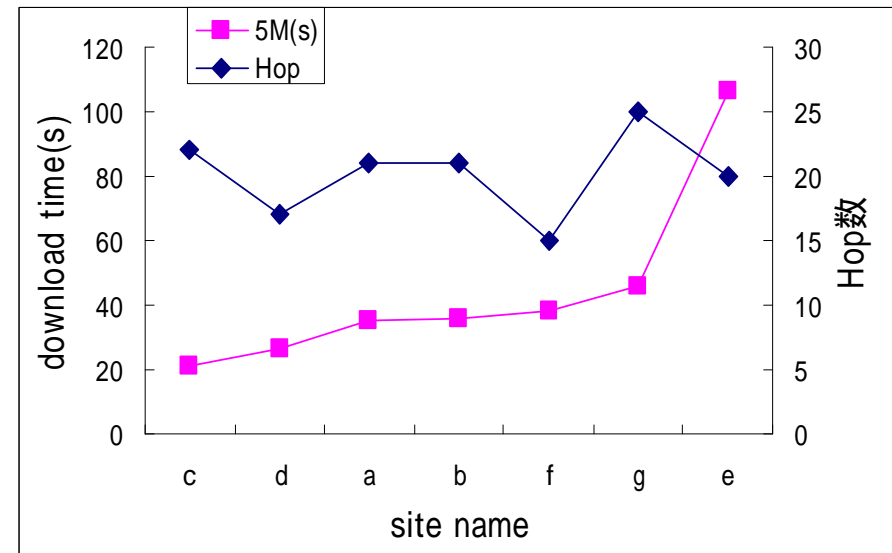


Fig-6サーバ名とDLTとHop数

- Fig5, 6より

Hop数とDLT(5MB)との相関係数は日時に関係なく0に近い値をとることが多く、今回の実験ではHop数の大小はDLT(5MB)にあまり影響していない。

5. 実環境におけるサーバ選択法の検討-2


<考察-1>

- RTTとDLT(5MB)の関係
 - 比較的高い相関係数を得られることが多いが,サーバ側の負荷を考慮していない点から,日時によって変動があると考えられる.
- pilot file(1MB,500KB,100KB)とDLT(5MB)の関係
 - 目的のファイルより小さいpilot fileを事前にダウンロードすることにより,サーバ側の負荷度合いとネットワーク上の遅延を考慮していると考えられる.
- Hop数とDLT(5MB)の関係
 - 今回の実験ではHop数の差があまりなかった為,Hop数の大小がDLT(5MB)の大小を反映していない可能性がある.



6. 室内環境における ネットワークサーバ性能の測定

- サーバ側の負荷状況が開示されていたと仮定し擬似的にサーバ側の負荷状況を変化させて、その負荷がDLTにどれほど影響があるのかを測定を行った。
 1. CPU使用率を変化させた時のDLTとの関係
 2. TCPコネクション数を変化させた時のDLTとの関係



6. 室内環境における ネットワークサーバ性能の測定

<実験-2>

1. サーバのCPU負荷を0% ~ 99%まで20%間隔で変化させて負荷をかけ, 5MBのファイルのDLTを測定
2. サーバに対し, TCPコネクション数を増やすことにより負荷をかけ, 150コネクションづつ変化させて5MBのファイルのDLTを測定

6. 室内環境における ネットワークサーバ性能の測定 <測定構成図>

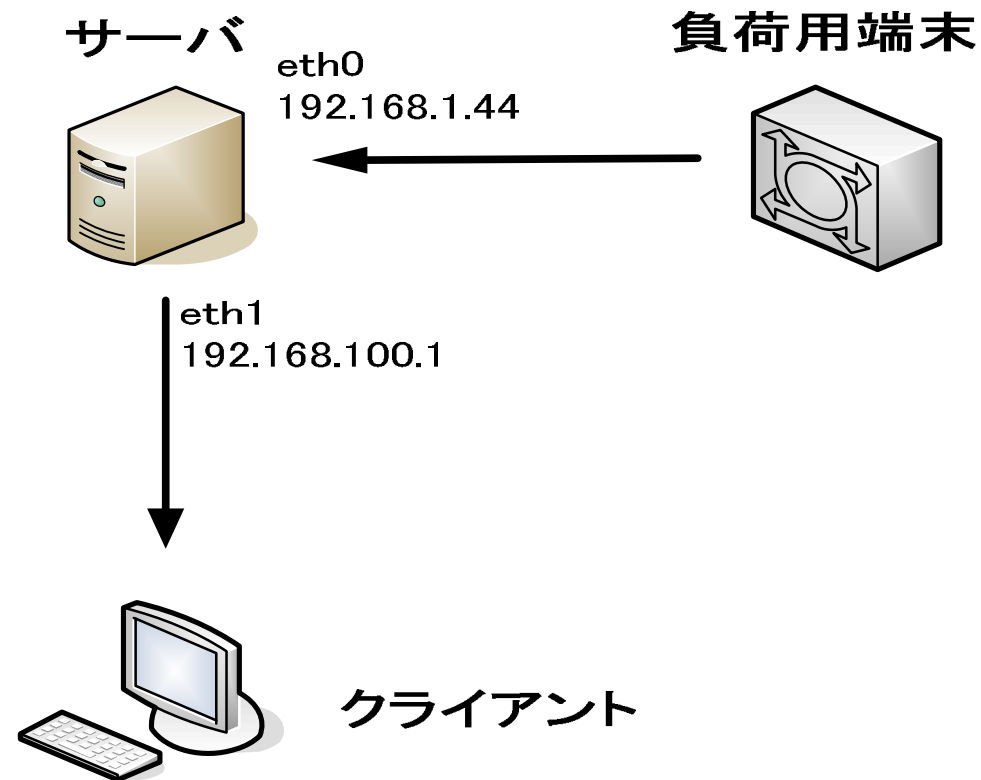


Fig-7 コネクション数の変化による負荷

6. 室内環境における ネットワークサーバ性能の測定 <測定環境>

Table-1 サーバ・クライアント端末

	CPU	memory	NIC	OS
server	celeron 2.4GHz	1GB	100BASE-TX × 2	Fedora core4
client	celeron 2.3GHz	1GB	100BASE-TX × 2	Windows XP Professional SP2

Table-2 負荷用端末

	CPU	memory	NIC	OS	software
PC-1	Pentium 1.2GHz	512MB	100BASE-TX	Windows XP Professional SP2	apache Jmeter
PC-2	PentiumM 1.3GHz	512MB	100BASE-TX	Windows XP Professional SP2	apache Jmeter

6. 室内環境における ネットワークサーバ性能の測定 <結果 - (6/7)>

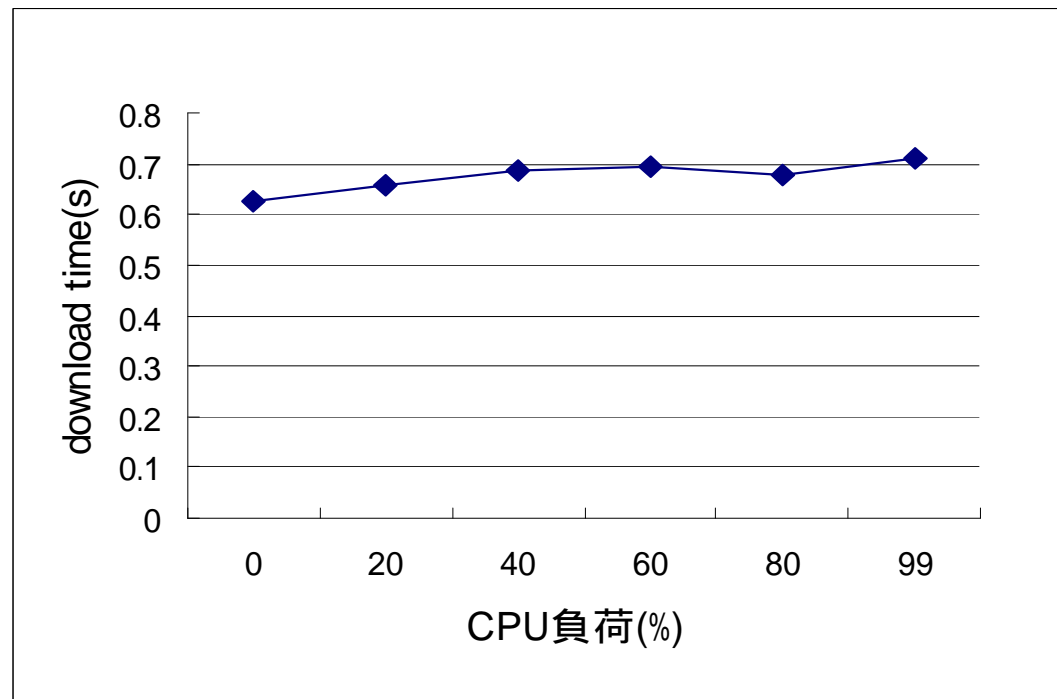


Fig-8 CPU負荷とDLT

Fig-8より

99%ものCPU負荷を発生させているにもかかわらず,何も負荷をかけていない時と比べて,DLTには約0.1(s)ほどの差しか生じなかった.

6. 室内環境における ネットワークサーバ性能の測定 <結果-(7/7)>

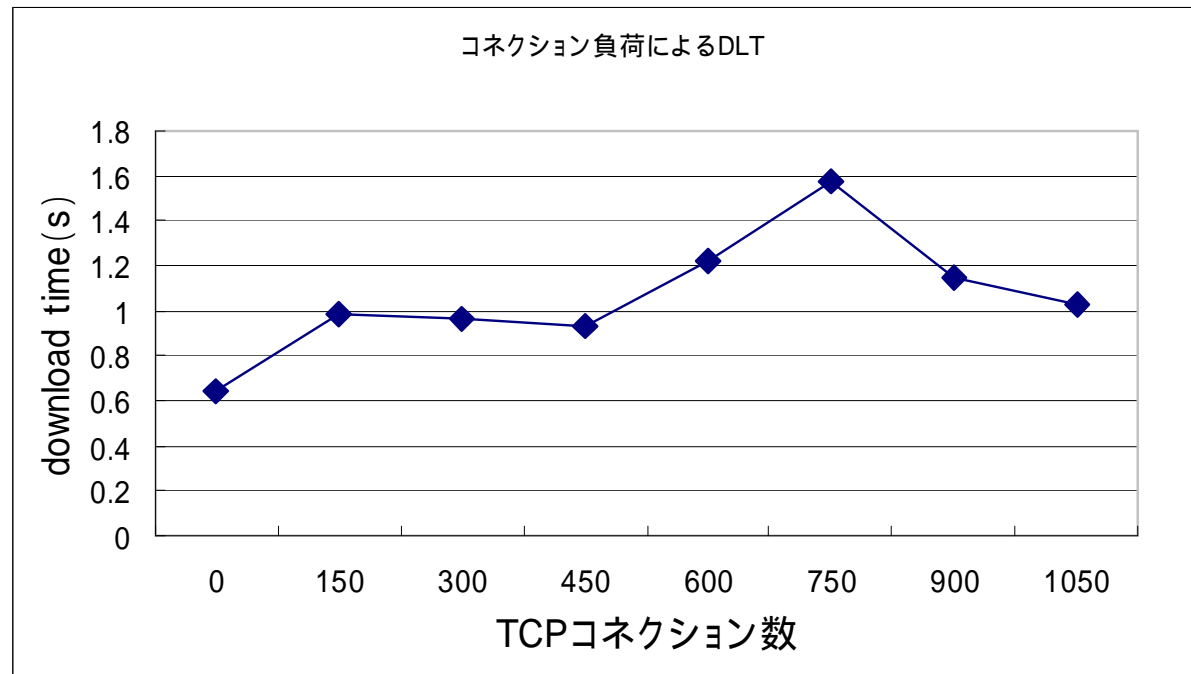


Fig-9 TCPコネクション数とDLT

Fig-9より

何も負荷を加えていない時と比較すると750コネクションの負荷を与えた場合でDLTが倍以上になっている。

6. 室内環境における ネットワークサーバ性能の測定 <考察-2>

- CPU負荷とDLT(5MB)の関係
 - CPUが99%という負荷状況下においても,DLTにほとんど影響を及ぼさないのは,CPUの手助けをしている機能があると考えられ,NICオフロード(CPUに代わってチェックサムの計算、データの分割処理等を行う機能)機能の存在が考えられる.
- コネクション数とDLT(5MB)の関係
 - CPUへの負荷を与えた場合と比べると,サーバのメモリ等のリソースも大きく消費される為,DLTへの影響が大きくなると考えられる.900コネクション以降に関しては,負荷発生装置側の処理が追いつかず,DLTが逆に短くなってしまったと考えられる.



今後の検討

- RTTに関しては比較的高い相関係数を得られることが多いが日時によって変動が大きいためもう少しデータを多めにとる必要がある。
- pilot file sizeの最小値の検討。
- サーバ側開示情報と組合わせた最適サーバ選択法の検討

まとめ

- ユーザが同一のサービスを提供しているサーバ群にアクセスして情報をダウンロードする場合において効率よく情報を得る方法として、pilot fileを事前にダウンロードしそのダウンロード時間が最適なサーバ選択をするうえで有益な指標となることがわかった。
- サーバ側が開示するサーバ選択の情報としてコネクション数が有効であることがわかった。



実績

- **中岩正洋**, 石井啓之 “サーバ負荷の遠隔観測による最適サーバ選択方式の検討” 電子情報通信学会 総合大会
- **中岩正洋**, 石井啓之 “遠隔観測による最適サーバ選択方式の検討” 電子情報通信学会 ソサイエティ大会
- **中岩正洋**, 中園信吾, 石井啓之 “遠隔観測による最適サーバ選択法の検討” 電子情報通信学会 情報ネットワーク研究会



測定で使用したサーバ

- a bonpoo(U.S)
- b Content-Type(U.S)
- c YouSendIt(U.S)
- d Ul taraShare(U.S)
- e Bigupload(Netherland)
- f AxiFile.com(Israel)
- g WebFileHost(unknown)



世界のサーバとのRTT

	RTT (ms)
国内	30 ~ 50
中国	120 ~ 250
米国	160 ~ 250
英国	270
オーストラリア	150 ~ 280
ブラジル	330 ~ 470

研究室から各国の大学に対してPing送信時のRTT



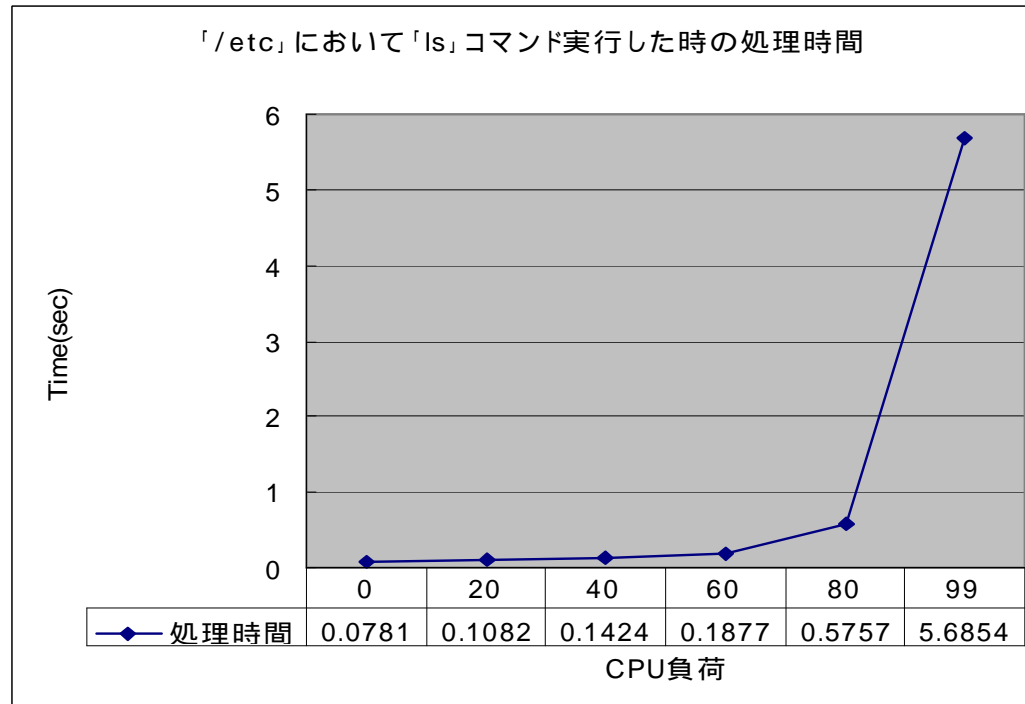
CPU負荷用のJavaプログラム

- CPULoadTester

実験用に作ったCPUに負荷をかけるソフト。
Javaで書いてあるのでOSを問わずに使用可能。
原理は単純で、Initializerである無意味な計算の
ループをさせて、1秒間に何回、回せるかを計測します。
それを元に、CPULoadTesterで負荷をかけます。
Intervalで指定したミリ秒を指定されたパーセンテージ
で分割し、ループ、スリープを繰り返させます。
これで±5%程度の誤差で負荷をかけられる。
環境によって計測できる時間の精度が異なるみたいなの
のでIntervalを指定できるようになっています。



CPU負荷状況の確認



- CPU負荷が大きくなるにつれて「ls」の処理にかかる時間が大きくなっている


NICオフロード機能

- LANのデータ転送速度が、ギガビット・イーサネット (GbE) や10GbEの登場で高速化している。これに伴い、サーバーのCPUにおけるTCP/IPネットワーク処理の負荷が問題になっている。そこで注目されるのがTCP/IP処理をNIC上の専用プロセッサに任せるオフロード機能である。
- ネットワーク・タスクのオフロードは、TCP/IPネットワーク処理のうち、特にCPUに負荷がかかっている処理をNIC (ネットワーク・インターフェース・カード) 上の専用プロセッサに負担させることで、CPUの負荷を軽減させる仕組みである。



NICオフロード機能の確認

- NICのドライバに付属されていたreadmeによると初期値で有効である
- サーバ側で「ethtool」コマンドにより、実際に確認した結果も有効となっていた



```
shin@S1-2:/home/shin
login as: shin
shin@192.168.1.44's password:
Last login: Wed Aug 23 19:52:07 2006 from 192.168.1.45
[shin@S1-2 ~]$ su
Password:
[root@S1-2 shin]# ethtool -k eth0
Offload parameters for eth0:
rx-checksumming: on
tx-checksumming: on
scatter-gather: on
tcp segmentation offload: on
[root@S1-2 shin]#
```

ネットワークアダプタのタスクオフロード機能を有効にする方法

- 1. <スタート>メニュー <ファイル名を指定して実行>と選択。
2. 「名前」に半角で「regedit」と入力し、<OK> ボタンをクリック。
3. 「レジストリエディタ」が起動したら、
HKEY_LOCAL_MACHINE SYSTEM CurrentControlSet Services Tcpip Parametersとキーをたどって開く。
4. <編集>メニュー <新規> <DWORD値>と選択し、名前を「DisableTaskOffload」とする。
5. レジストリエディタを終了し、Windows XPを再起動する。
- これでタスクオフロード機能が有効になりました。
「DisableTaskOffload」のデータ値が「0」で有効に、「1」で無効になりますので、設定後にネットワークの異常が見られる場合は、値を削除するかデータ値を「1」に変更してください。ちなみに、タスクオフロード機能をサポートしていないマシンに設定を行った場合でも、特に問題は起きませんでした。

3. 既存のサーバ選択法

<1.サーバ側で選択する方法の例>

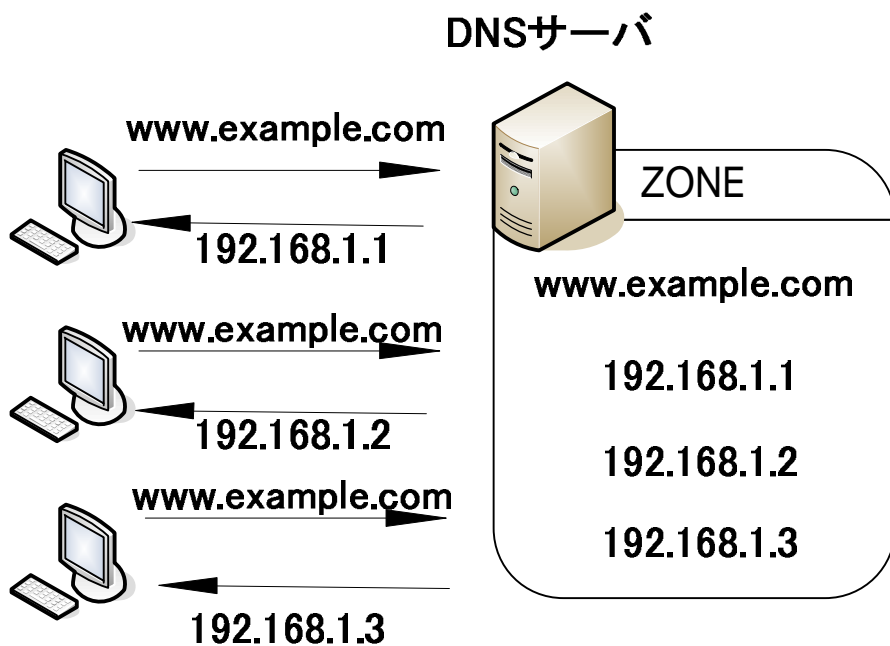


Fig-3.1 DNSラウンドロビン

<2.ユーザ側で選択する方法の例>

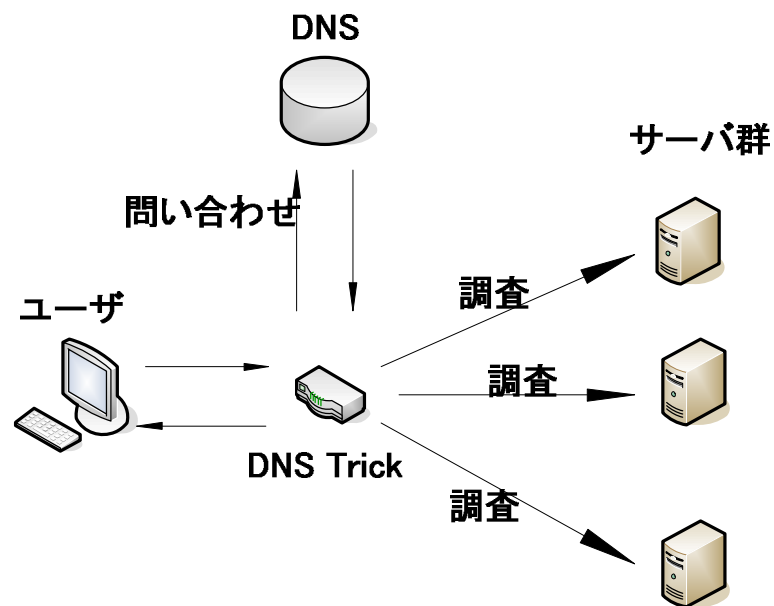


Fig-3.2 DNS Trick

4. 実環境におけるサーバ選択法の検討-1

<結果>

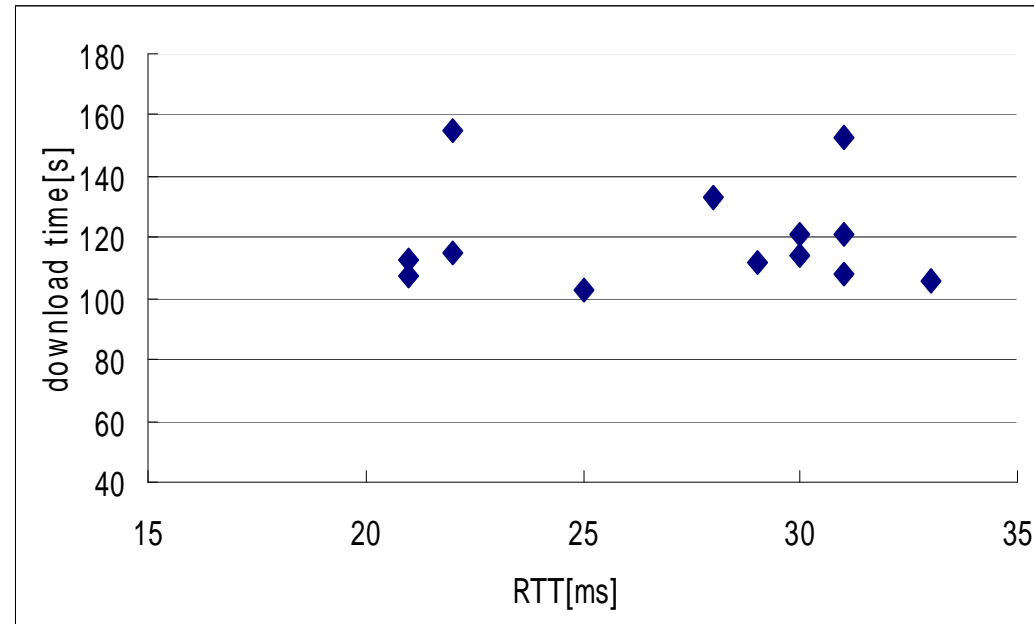


Fig-4.1 download timeとRTTの関係

- Fig-5より

日本国内のサーバのみを使用した測定では、RTTの大小とダウンロード時間の大小は互いに関係が少なく、サーバ選択論理には用いられないことがわかった。