

クラウドコンピューティング

基礎論

第9回

創造情報・小林克志

ikob@acm.org

Course Outline

- Administrivia
- Cloud computing
- Service reliability
- Scale-up / Scale-out
- Distributed data stores
- Global services
- Datacenter networkings (1)
- Datacenter networkings (2)
- Network performance
- User experiences
- Network latencies
- Advanced topics

For hands-on exercise :

Install two softwares

1.Wireshark : A packet capture and analyzer

- Just install package from <http://www.wireshark.org>

2.NS2 : Network simulator.

Three options are there:

A. Docker

- Install docker software and container:
 - <https://github.com/ekiourk/docker-ns2>
 - X-window server is required. It depends on OS.

B. Native applicationIf you are using Linux, use this option.

- Install ns-allinone-2.35 from source because NS-2 package.
Note that some distribution may not work.
- X-window, perl, gnuplot are also required.

C. Virtual Machine (VM)

- Install Hypervisor Software.
 - Oracle VirtualBoX is free.
vmware or others are also welcome.
- Linux VM image with NS2 software will be available from the course Web.

演習に向けて：2つのソフトウェアをインストールする

1. パケットアナライザ wireshark

- <http://www.wireshark.org> を参考にインストールすること。

2. ネットワークシミュレータ NS2

A. Docker

- Docker をインストールし、コンテナを使用する。
 - <https://github.com/ekiourk/docker-ns2>
 - X-window の設定は OS に依存する。

B. Native

- Linux を利用している場合は、この方法を使う。
- ns-allinone-2.35 を install すること。
 - X-window, perl, gnuplot なども必要となる。

C. VM で動作

- ハイパーバイザの導入
 - Oracle VirtualBox であれば無償
vmware 他でもかまわない
- NS2 付きの Linux 仮想マシンイメージを講義ページで配布する

Final report

- Choose one network latency topic surveyed in:
Briscoe, Bob, et al. "Reducing Internet Latency: a survey of techniques and their merits." IEEE Communications Surveys & Tutorials.
- Write review of **all of the cited papers** in the chosen topic, but excluding [202].
The review must include both strong and weak points of the papers.
- Submit via course Web.
 - Submission due date is July 31.
 - The review must be 2-4 pages long and written either in English or Japanese.

最終レポート

- 以下のサーベイ論文からネットワーク遅延に関するトピックを一つ選択せよ:

Briscoe, Bob, et al. "Reducing Internet Latency: a survey of techniques and their merits." IEEE Communications Surveys & Tutorials.

- 選択したトピックで引用されている**全ての論文**をレビューすること。

レビューには論文の良い点、悪い点両方を含むこと。

- 講義 Web ページから提出すること。

- 締め切りは 7 月 31 日 (金)

- レビューは 2-4 ページにまとめ、英語あるいは日本語で記述すること。

銀行システム、外部技術に死角 りそなや新生、同時障害 安全強化策、管理が複雑に

2018/6/28付 | 日本経済新聞 朝刊

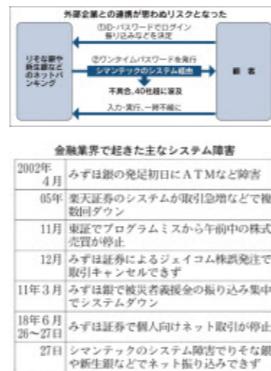
りそなグループや新生銀行など複数の金融機関で27日、インターネットバンキングでの振り込みができない障害が発生した。個人認証の安全強化のために採用したシマンテック社のシステムで不具合が起きたのが原因だ。銀行本体ではなく、IT（情報技術）企業のサービスを起点に広がった今回の障害。外部との連携が増えるなかで、銀行システムの新たな死角が浮き彫りになった。

「ワンタイムパスワードを利用したお取引がエラーになっています」。27日午前、りそなグループのホームページで顧客に対する告知が掲載された。行内でシステムを監視する部署が、ネットでの振り込みができない状態に気づいたという。

りそな他にも新生銀行やセブン銀行、常陽銀行、山口銀行、複数の信用金庫などで同様の障害が起きた。その数は40社超にのぼるとみられる。各社は、顧客がネットバンキングで振り込みする際の認証を強化するため、一定時間ごとに更新されるパスワードを入力する仕組みを導入している。このシステムを提供するシマンテックで不具合が生じ、影響は全国に広がった。

シマンテック関係者によると、パスワードは「データセンター経由でクラウドで発行している」という。日本では金融機関が多くこのシステムを使っている。海外では社内ネットへの接続に使われることが大半で、事業会社などでも同様な問題が起きたとみられる。シマンテック社は障害発生後、ホームページで約30分ごとに情報を更新。午後1時すぎ、問題が解消したと公表した。「原因は究明中」としている。

金融機関のシステム障害はこれまで、経営統合に伴うシステムの見直し時など、内部で起きるのが通例だった。他社のサービスを起点に顧客に不具合が生じる今回のようなケースは「これまでほぼ例がない」（大手銀）。



今回のような認証システムが広がったのは約5年前。従来はネットバンクの利用者に乱数表を別途送付し、振り込みの際に手元の乱数表の番号を打ち込んでもらう方法を探っていた。だが振り込め詐欺が広がり、犯人が乱数表の番号まで聞き出してしまう事例も増えた。事態を重くみた金融機関は乱数表をやめ、より安全性の高いワンタイムパスワードの導入に動いた。

例えばりそなでは、銀行本体のネットバンキングとシマンテックのシステムを「API（アプリケーション・プログラミング・インターフェース）」と呼ばれる技術仕様によりつながっている。

今年施行した改正銀行法では金融機関に対し、IT企業が銀行システムに接続できるよう仕様を公開する「オープンAPI」の整備を求めている。オープンAPIでIT企業との連携が進めば、家計簿や会計ソフト上から直接振り込みができるといったサービスも提供しやすくなる。

金融機関はネットバンキングの利用客そのものも増やそうとしている。実際、三菱UFJ銀行のネットバンキング利用者は5年間で約4割増。店舗への来客数が減るなか、費用を削減するためにもネットへのシフトは続く見通しだ。

外部企業との連携で顧客の利便性は高まる半面、自社で管理しきれないシステムのリスクも増える——。今回のネット障害は新たな教訓を残した。不具合が起きた銀行の幹部は「一つのサービスに集約するのは危険。代替手段を残しておくことが大事だ」と話した。

本サービスに関する知的財産権その他一切の権利は、日本経済新聞社またはその情報提供者に帰属します。また、本サービスに掲載の記事・写真等の無断複製・転載を禁じます。

Nikkei Inc. No reproduction without permission.

Today's Assignment

- University people are frustrated with the academic affair system, such as UTAS, due to poor usability. A smartphone application named Oraio is released, which improves the usability of the official service. Oraio requires an user to register own ID/password. Then, Oraio retrieves sign-upped classes, and builds own schedule and attendance records automatically, i.e. Web scraping or crawling.
Some universities officially discouraged their students using such application because of the security concern.
- Whether should the university administrator of information technology allow or disallow such application ?
 - If your student ID is even: Discuss the above from the viewpoint of CSO.
 - Otherwise: Discuss the above from the viewpoint of CIO.
- Submit your answers in Japanese or in English via the course web.

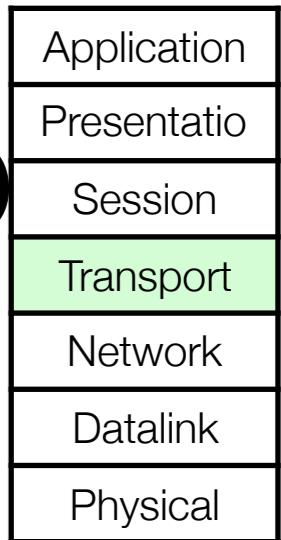
Note: CIO: Chief Information Officer / CSO : Chief Security Officer

本日の課題

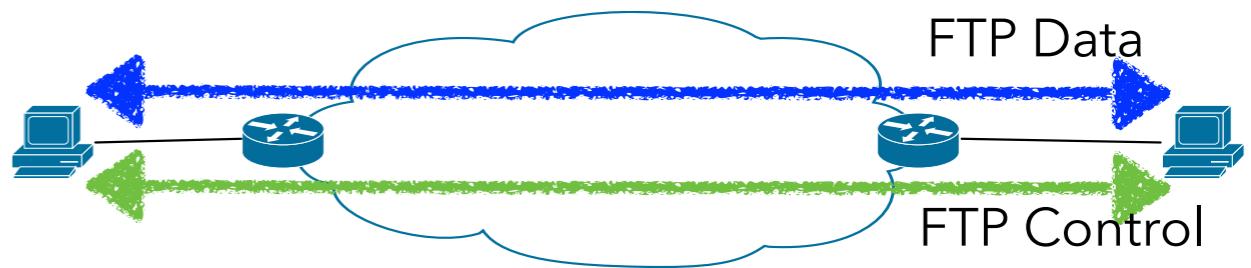
- 大学の学務システムについて使い勝手などで不満は多い。学務システムを向上させるスマートホンアプリケーションとして Oraio が登場した。Oraio はマッシュアップアプリケーションの一種で学生の ID/Password を利用して学務情報を取得、スケジュール、出席履歴などが生成される。
いくつかの大学はこのようなアプリケーションの利用に対してセキュリティの懸念などから注意喚起をおこなった。
- 大学の情報技術責任者としてこのようなアプリケーションを許容すべきか禁止すべきか?
 - 学籍番号が偶数 : CSO の視点で議論せよ
 - それ以外 : CIO の視点で議論せよ
- 講義 Web ページから回答すること。

CIO: Chief Information Officer / CSO : Chief Security Officer

Five tuple : flow/connection identification



1. Source address

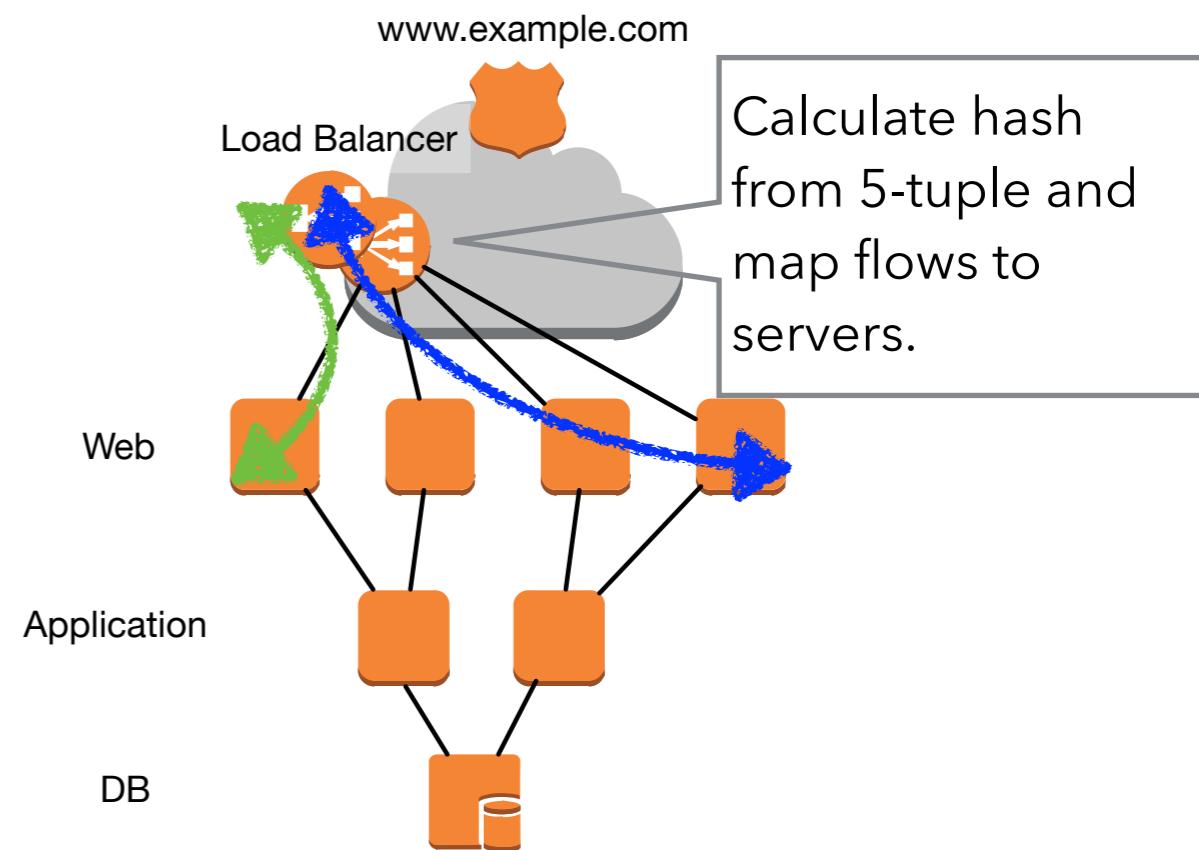


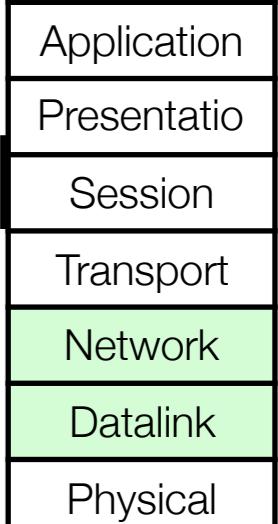
2. Destination address

3. Upper (Transport) Protocol
{UDP, TCP}

4. Source port

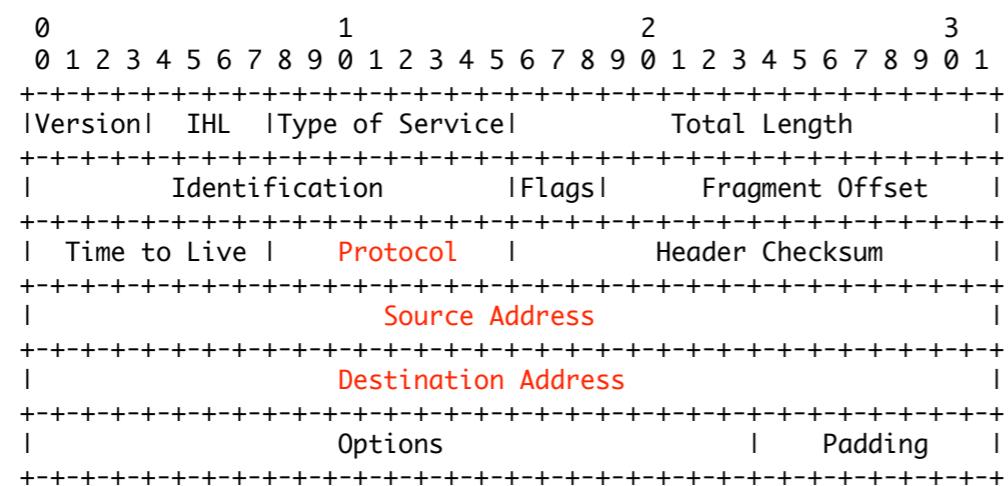
5. Destination port



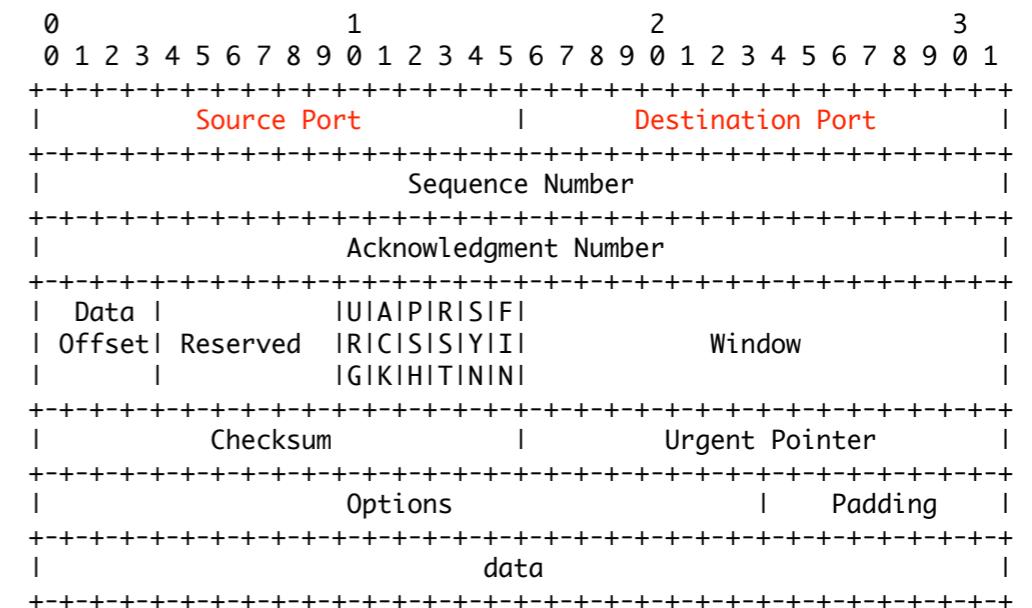


Five tuple : flow/connection identification(cont'd)

IP header:

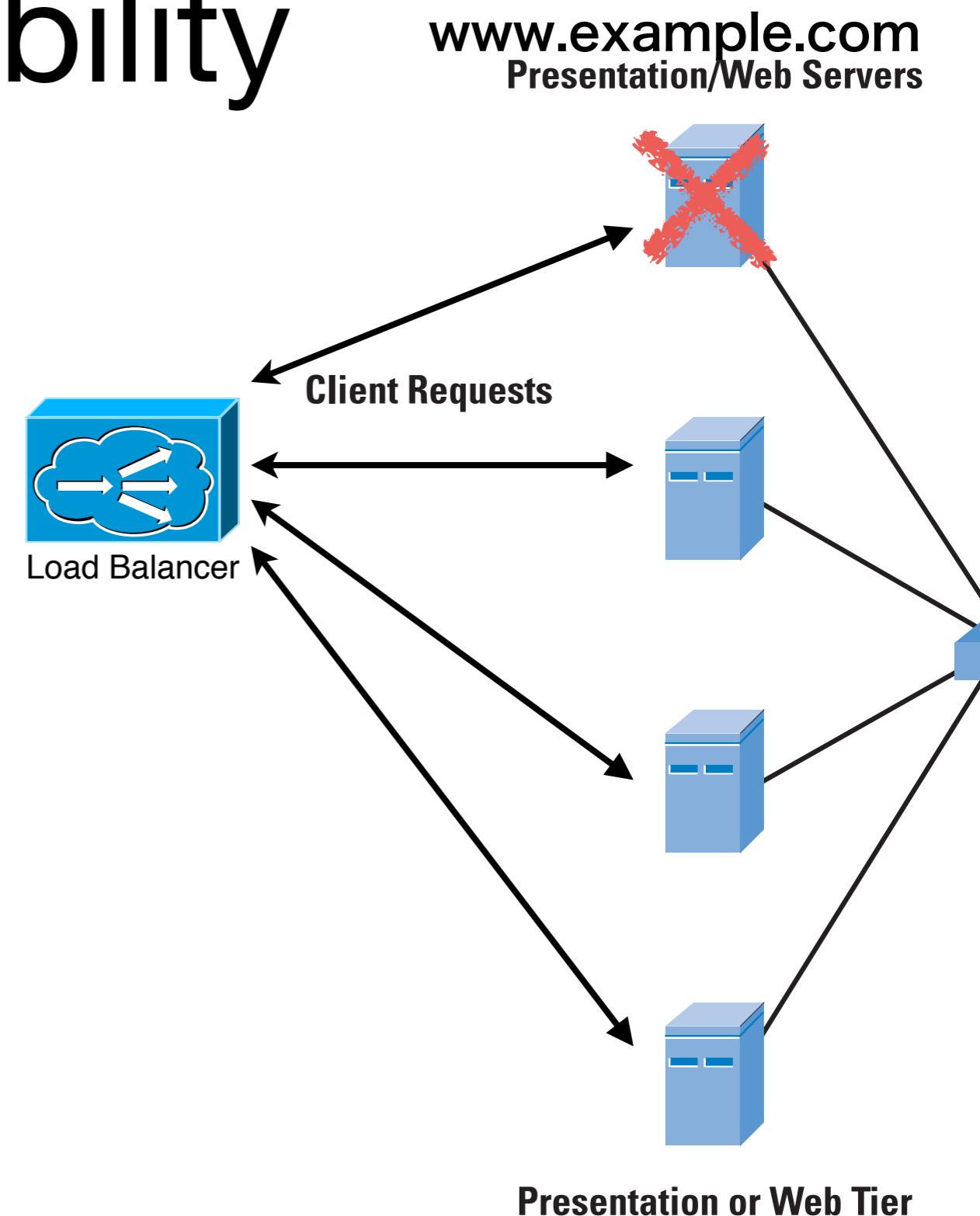


TCP header:



Load balancing policy and availability

- Policy
 - Hash: based IP address or other client-specific info.
 - Least connections: assign most least connection server
 - Round robin: new connection to the next server with RR
 - Weighted ver. of above: considering server condition both static and dynamic
- Health check
 - monitoring servers and update list in several levels (layers), e.g., ICMP(ping), application layer polling, server load, manual...



Outline

- Administravia
- Homework review
- User eXperience on Web service
- NS2 simulation

Today's Quiz

1. Show two or more Contents Delivery Network (CDN) hosted sites.
2. Tell the reasons why such sites look like hosted on CDN than own Web server.
3. Submit your answers either in Japanese or in English via the course web.

Today's Quiz

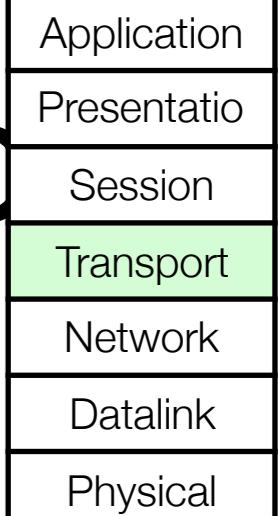
1. Contents Delivery Network (CDN)でホストされている Web サービスを 2つ以上示せ。
- 2.これらのサービス が自身の CDN でホストされている理由を示せ。
3. Submit your answers either in Japanese or in English via the course web.

Today's Assignment

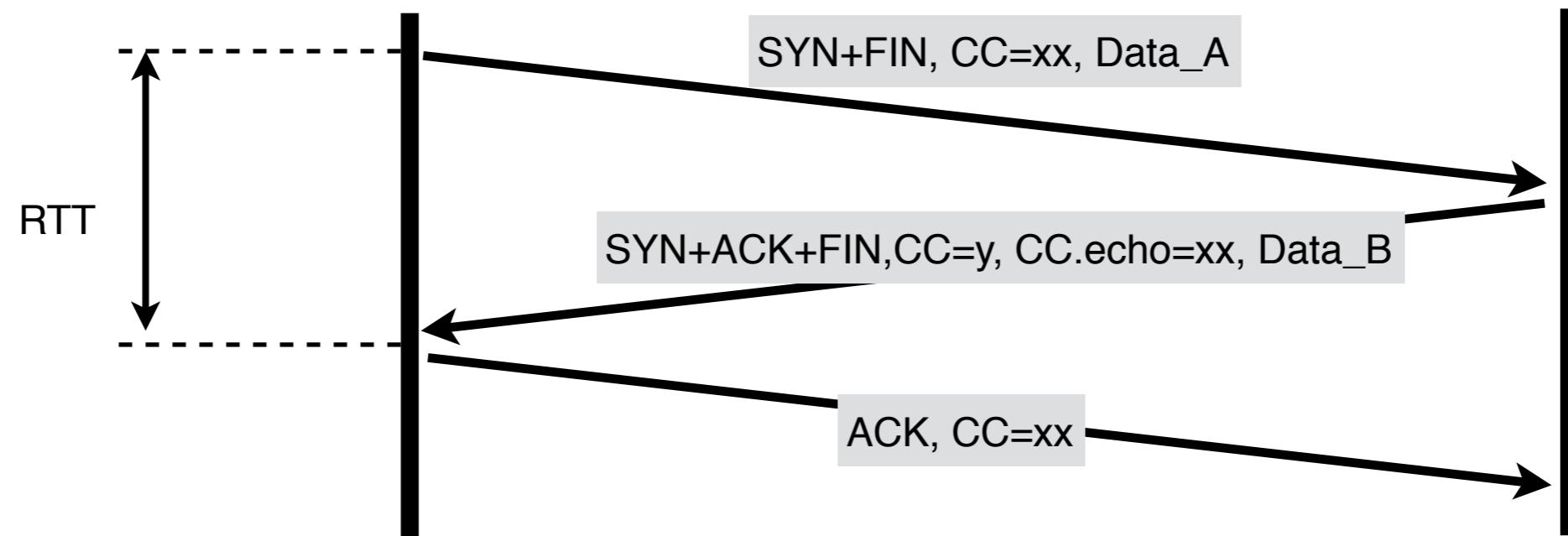
- In the class, TCP Fast Open (RFC7413) was introduced as an approach to reduce the latency of TCP connection set-up. Another option T/TCP - TCP extensions for Transactions (RFC1644) which shares the same goal had been standardized for 20 years. In addition, T/TCP reduces the TCP set-up latency not only from the second or later connections, but from the first one. However, now T/TCP standard has been obsoleted.
- Read RFC7413, RFC1644 and related documents. Discuss why T/TCP has been obsoleted.
- Submit your answers either in Japanese or in English via the course web.

本日の課題

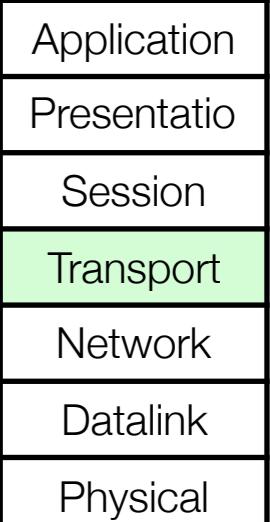
- 講義では、TCP コネクションセットアップの遅延を抑える手法として TCP Fast Open (RFC7413) を取り上げた。TCP Fast Open と同じ目的で - TCP extensions for Transactions (RFC1644) が 20 年前に標準化されている。さらに、T/TCP ではセットアップ遅延を 2 つめ以降のコネクションだけではなく、最初のコネクションから抑えることができる。しかしながら、T/TCP 標準は廃止された。
- RFC7413, RFC1644 および関連文書を読み、T/TCP が廃止された理由を考察せよ。
- 講義 Web ページから回答すること。



T/TCP - TCP Extensions for Transactions (RFC1644)

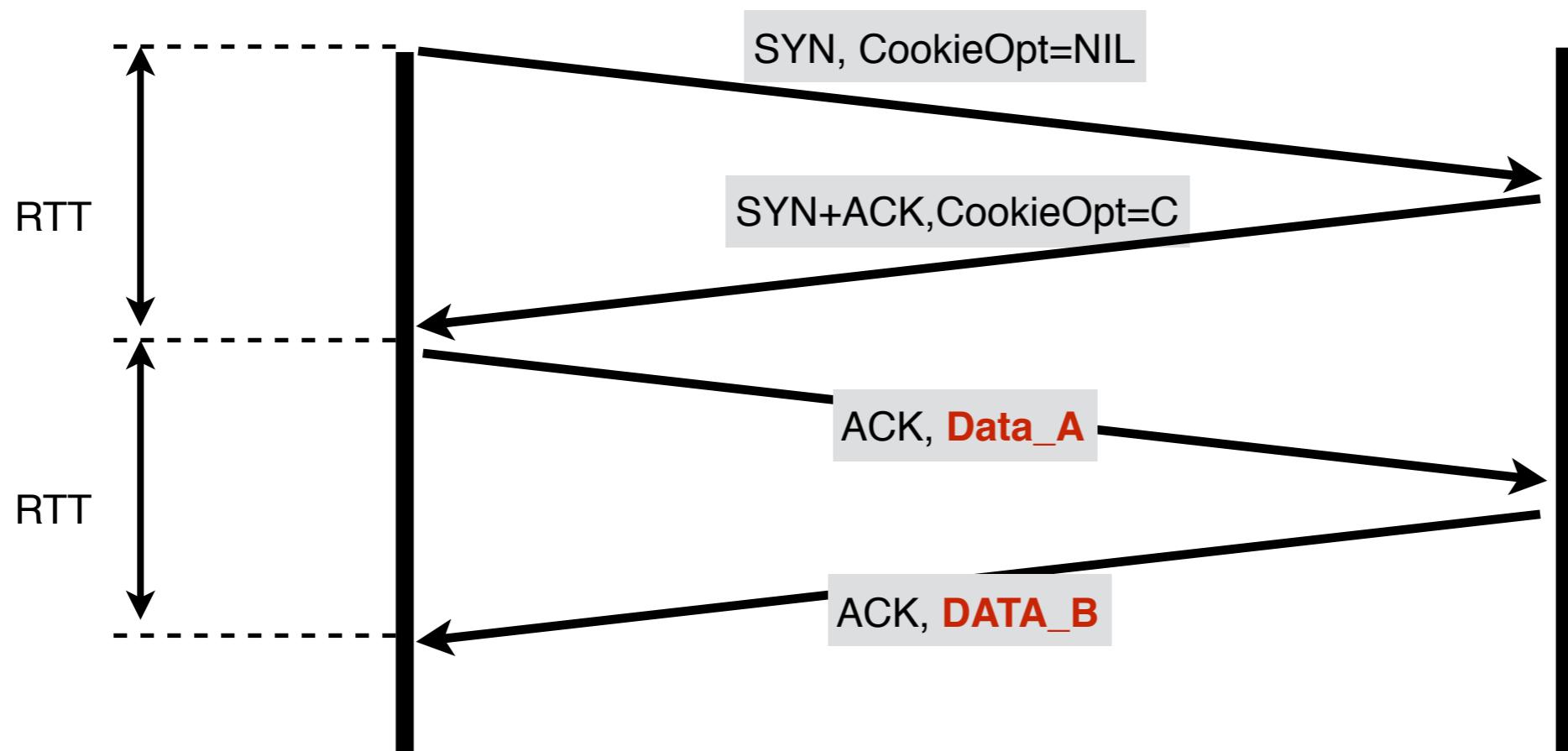


Incompatible with TCP SYN cookies which mitigate risks from TCP SYN flood attack.
 TCP SYN flood with spoofing source address are used in DDoS attacks.

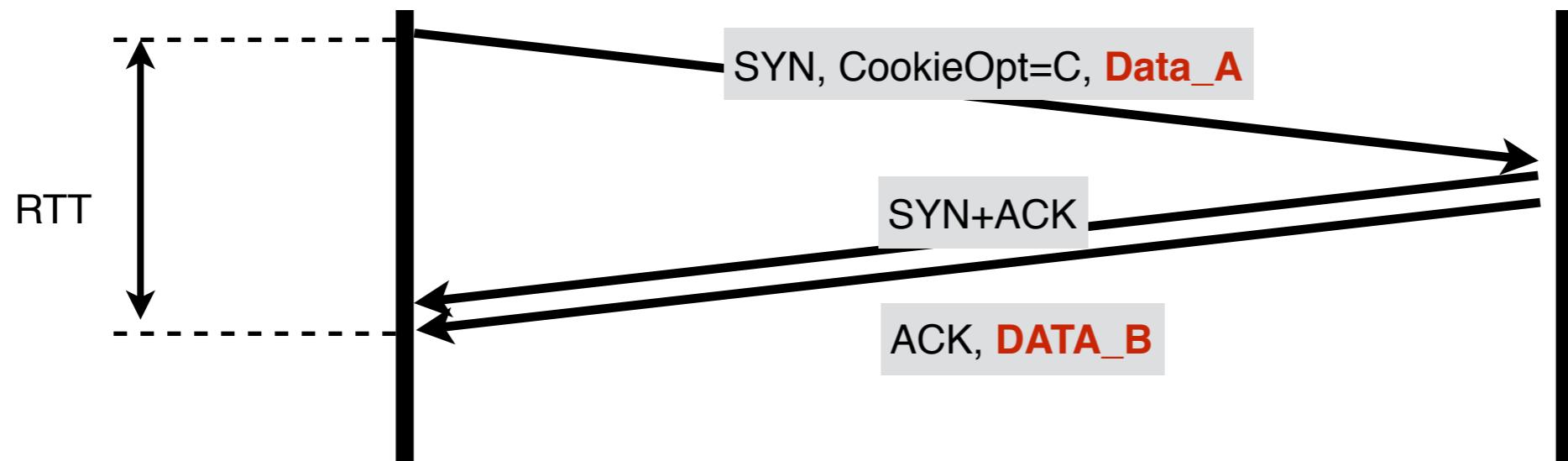


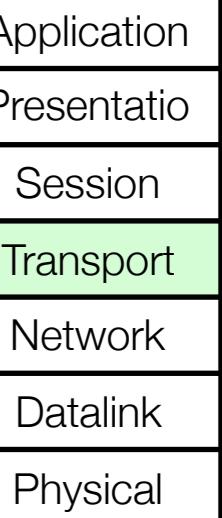
TCP Fast Open (RFC7413)

1st connection



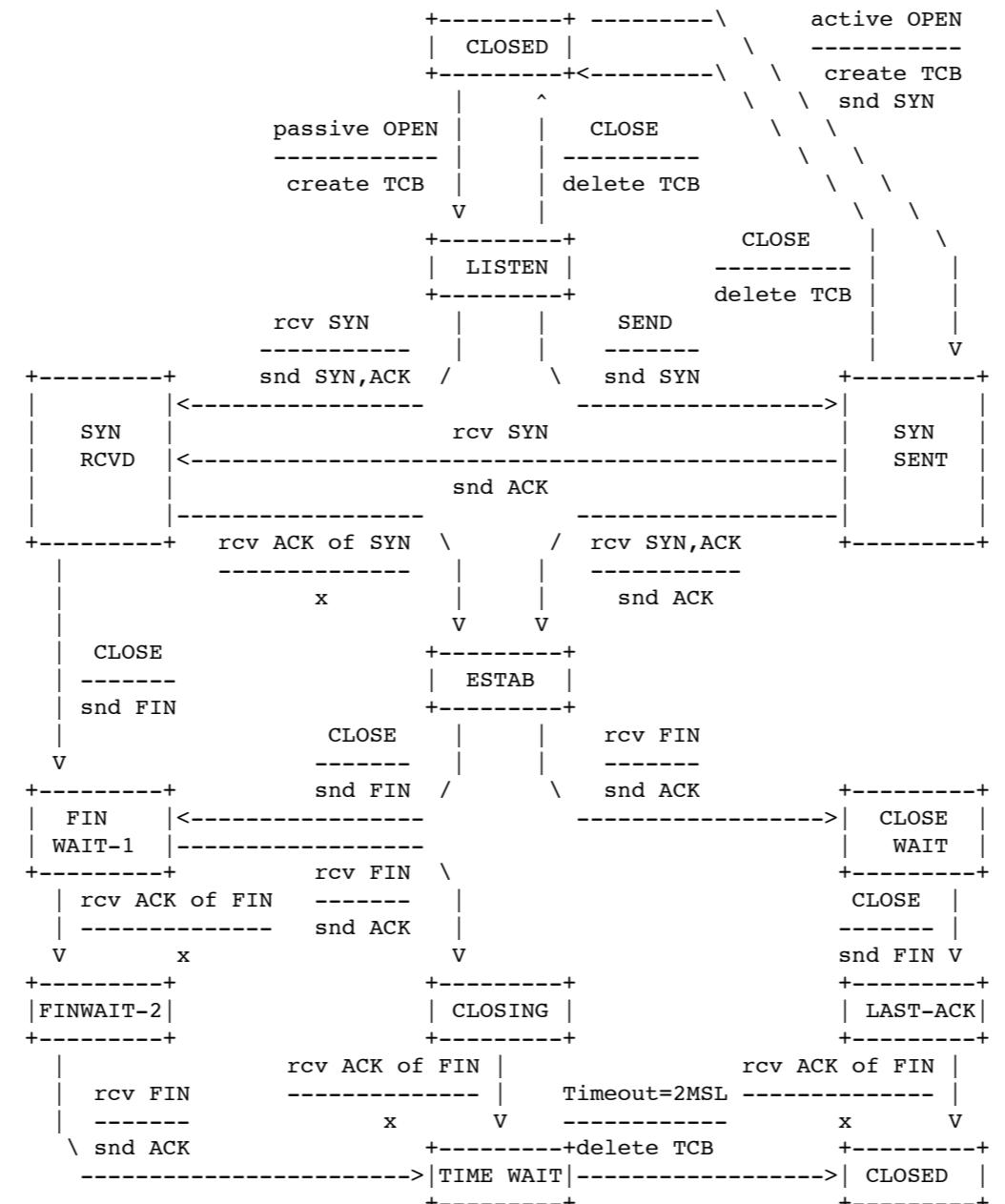
2nd connection





September 1981

Transmission Control Protocol
Functional Specification

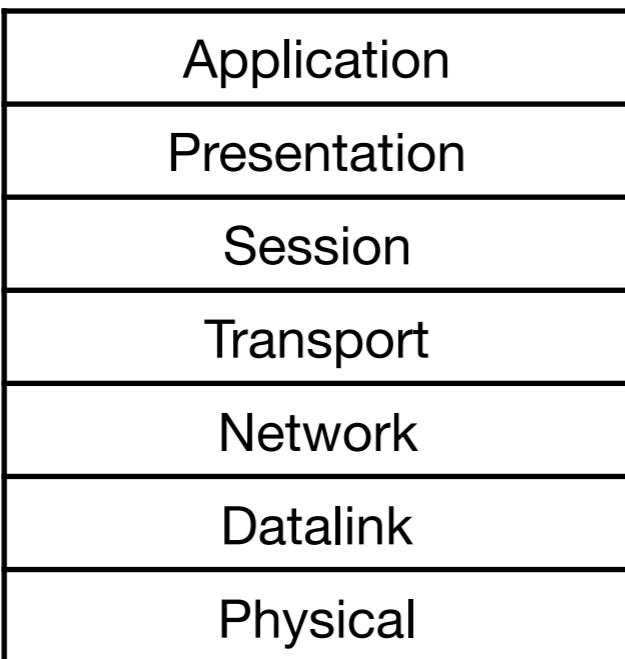


TCP Connection State Diagram
Figure 6.

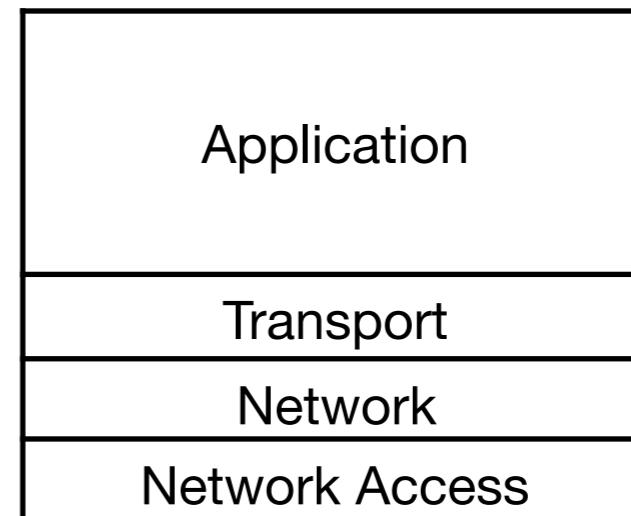
[Page 23]

Data communication and layer models

- OSI was standardized by ISO and ITU with spending a lot of efforts. However, Internet Protocol by IETF which competes OSI are widely accepted.
 - Only 7 layer reference model is survived, which is referred a lot of text books.
- Note : IETF uses DARPA model instead.



OSI reference model (7 layers)
(OSI: Open Systems Interconnection)



DARPA model (4 layer)
not including Physical layer

X.500 by OSI
and
Light Directory Access Protocol(LDAP) by IETF

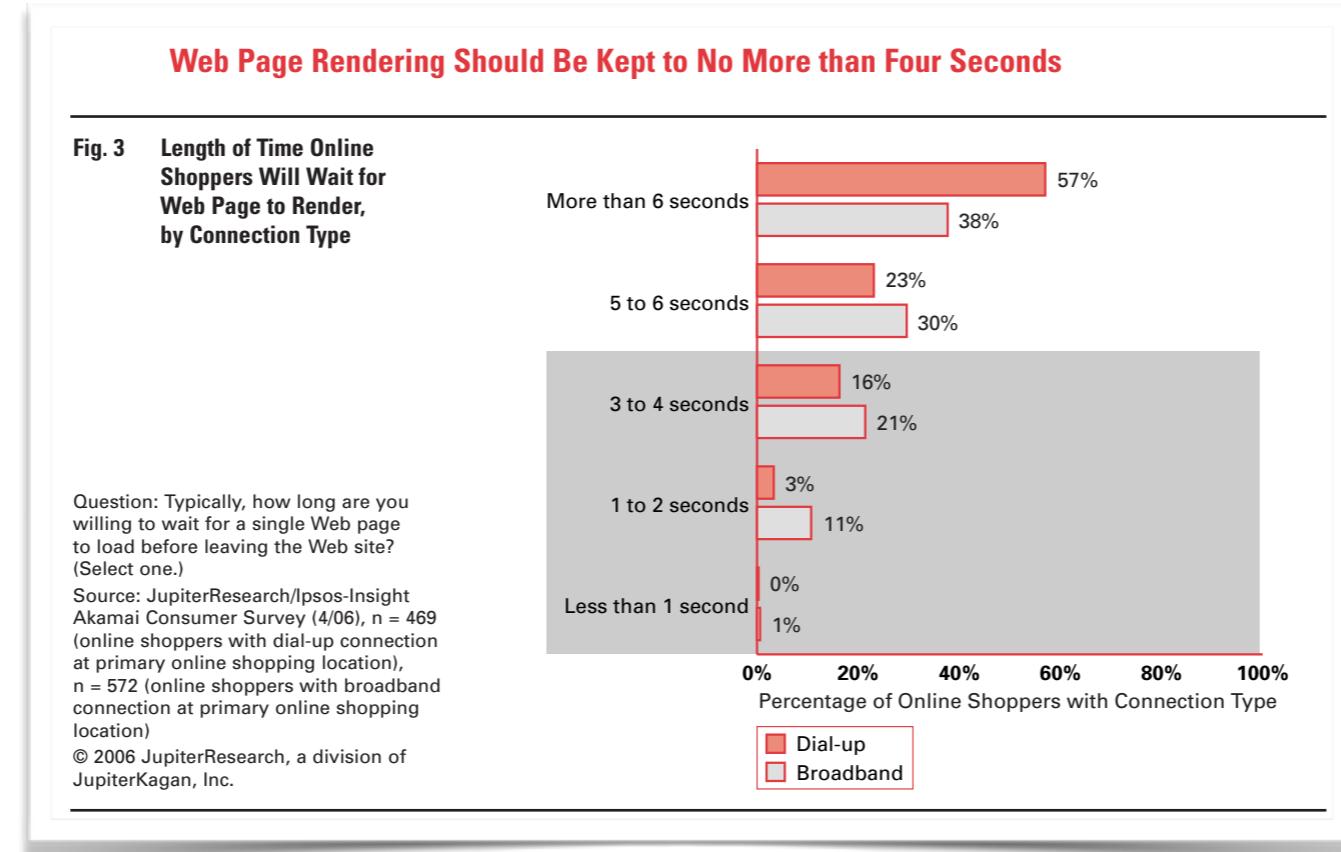
Outline

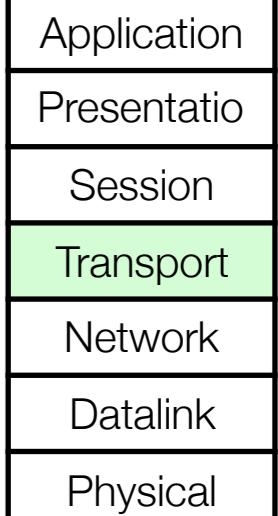
- Administravia
- Homework review
- User eXperience on Web service
- NS2 simulation

Application
Presentation
Session
Transport
Network
Datalink
Physical

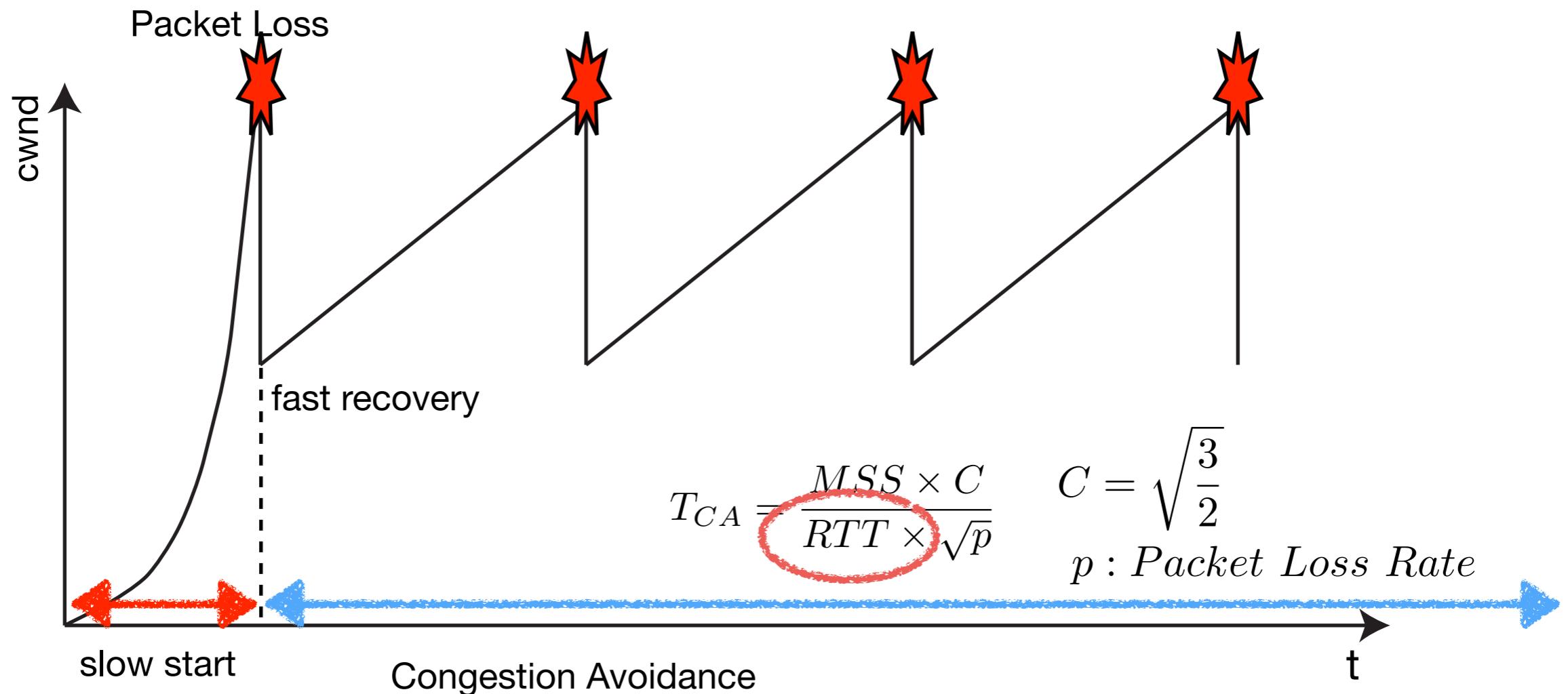
Four/Two second rule in Web services

- < 4 sec. render completion limit to keep customers' attention @ 2006
 - The limit is decreasing, e.g., 2 sec. @2013
- Poor satisfaction decreases customer loyalty and revisiting.
- To miss opportunity, and to lost revenue on e-Commerce.





TCP window behavior on NewReno



Mathis, Matthew, et al. "The macroscopic behavior of the TCP congestion avoidance algorithm." ACM SIGCOMM Computer Communication Review 27.3 (1997): 67-82.

Latency in network applications

- Legacy:
 - Processing, Propagation, Queueing, De-jitter, Retransmission, Flow-control,
- Today:
 - Entire application process from the perspective of entire UX
 - Smooth interactive UI with Ajax

Application
Presentation
Session
Transport
Network
Datalink
Physical

Latency of network service caused by ...

Intra-node (server, client) :

- Can be reduced by improving node performance and/or software, but not discussed in this class.

Client - Server communication :

- Global network, or the Internet

Inter-node in server cluster :

- Depending on communication in DC network

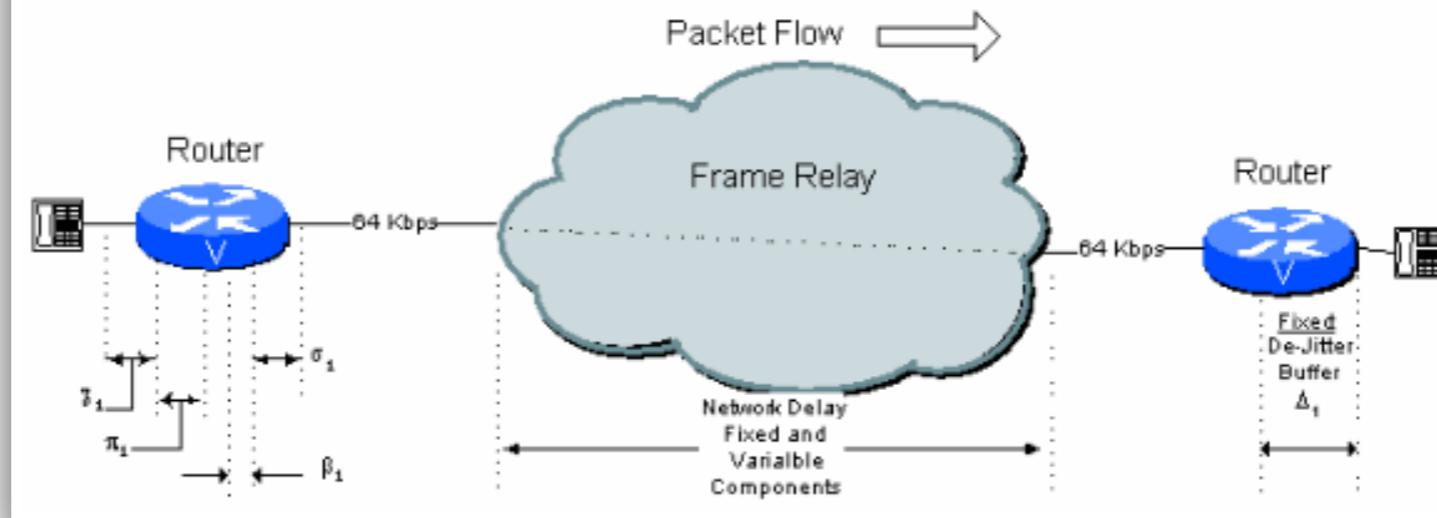
Packet level latency in Networks

- Cause of latency, in case of VoIP:
 - Processing (Coder)
 - Packetization
 - Serialization
 - Queueing / Buffering
 - Network Switching (Propagation + Switching)
 - De-jitter

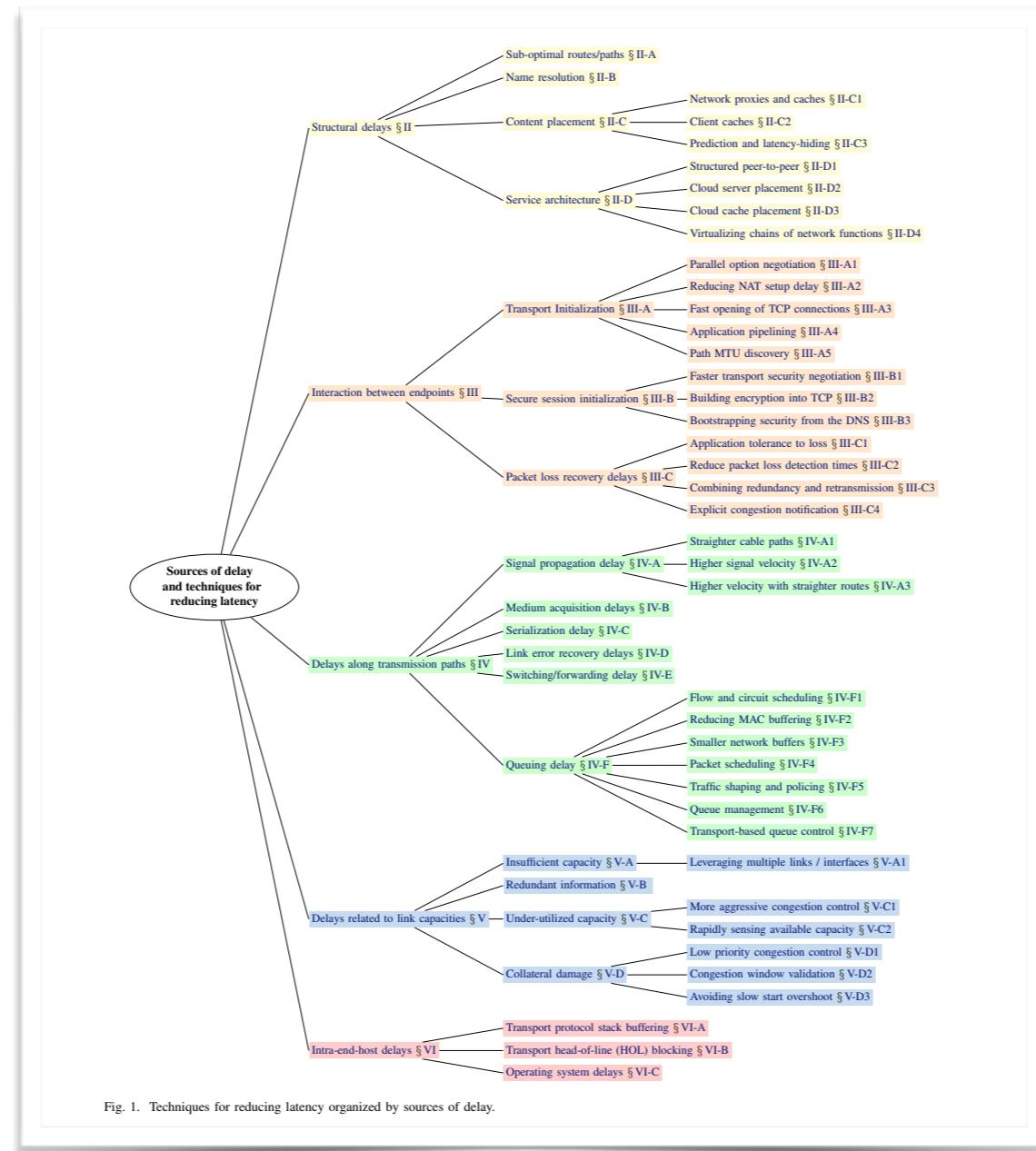
Table 6 .1: Single Hop Delay Calculation

Delay Type	Fixed (ms)	Variable (ms)
Coder Delay, χ_1	18	
Packetization Delay, π_1	30	
Queuing/Buffering, β_1		8
Serialization Delay (64 kbps), σ_1	5	
Network Delay (Public Frame), ω_1	40	25
De-jitter Buffer Delay, Δ_1	45	
Totals	138	33

Figure 6 - 1: Single Hop Example Connection



Source of Delay in Network applications



Briscoe, B., Brunstrom, A., Petlund, A., Hayes, D., Ros, D., Tsang, I.-J., Gjessing, S., Fairhurst, G., Griwodz, C. & Welzl, M., "Reducing Internet Latency: A Survey of Techniques and their Merits," IEEE Communications Surveys & Tutorials (2014)

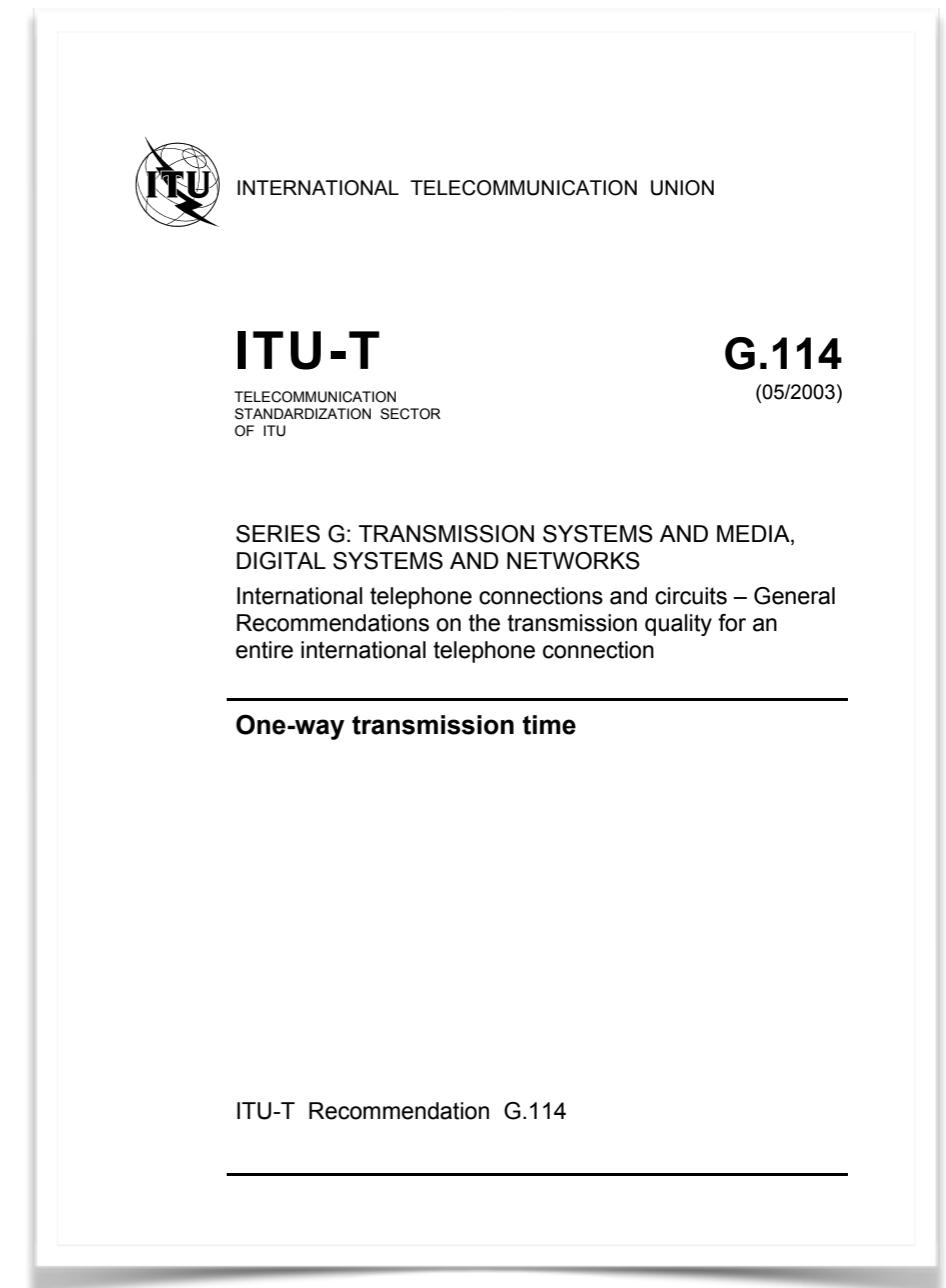
Regional : Network connectivity to applications

- End-to-end network latencies (,or delays) between cloud services and end-systems give significant impacts to user-experiences (UX) and to application performances.
 - Latencies to end applications depend on the location of service.
 - Distance is the biggest factor, such as, 10ms of Tokyo - Osaka, 100ms of Tokyo - LAX, in round-trip-time(RTT).
 - Re-locating server is the most effective approach to reduce latencies.
 - Cloud service providers deploy their services to major cities.
 - CDN providers deploy own edge systems to many Point of Presence (PoP)s

Application
Presentation
Session
Transport
Network
Datalink
Physical

Latency at Voice over IP (VoIP)

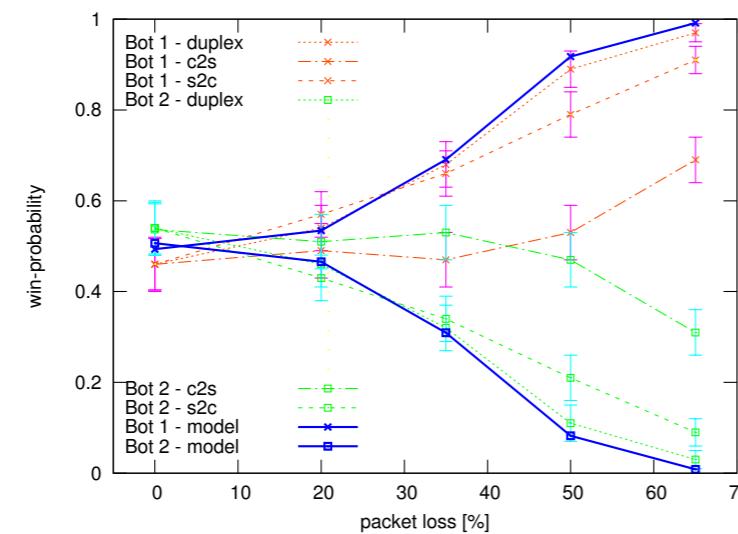
- < 150msec. for interactive communication by ITU-T G.114.
 - One-way, Mouth-to-ear.
- VoIP Deployed 16M@2013 by NTT.
 - Prioritized class in on Diffserv CoS (Class of Service, RFC2475)
 - NGN (Next Generation Network)
 - NTT announces ISDN (INS) will terminated at 2025.



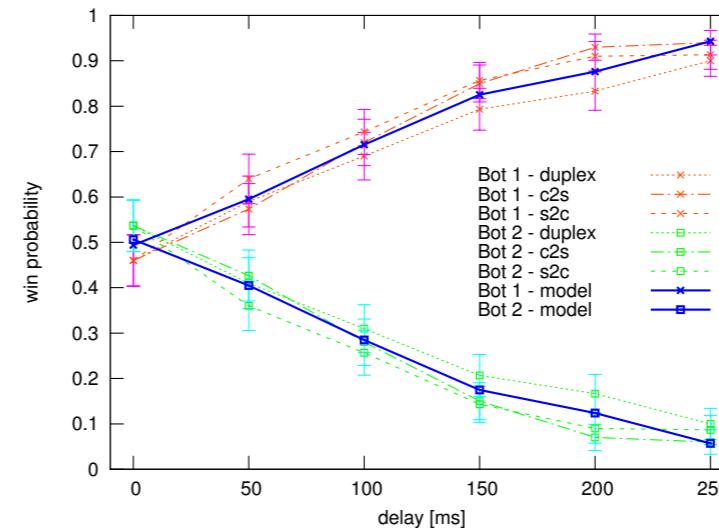
Application
Presentation
Session
Transport
Network
Datalink
Physical

Latency at On-line Game

- > 50 msec latency behind is disadvantage in First Person Shooting (FPS), e.g. Quake3
 - But packet loss is not significant
 - (Short size) UDP packets sent to update player's status.
 - If a packet is lost, able to recover by using subsequents packets information.
 - Resilient to packet losses rather than latency.



(c) win probability vs. loss

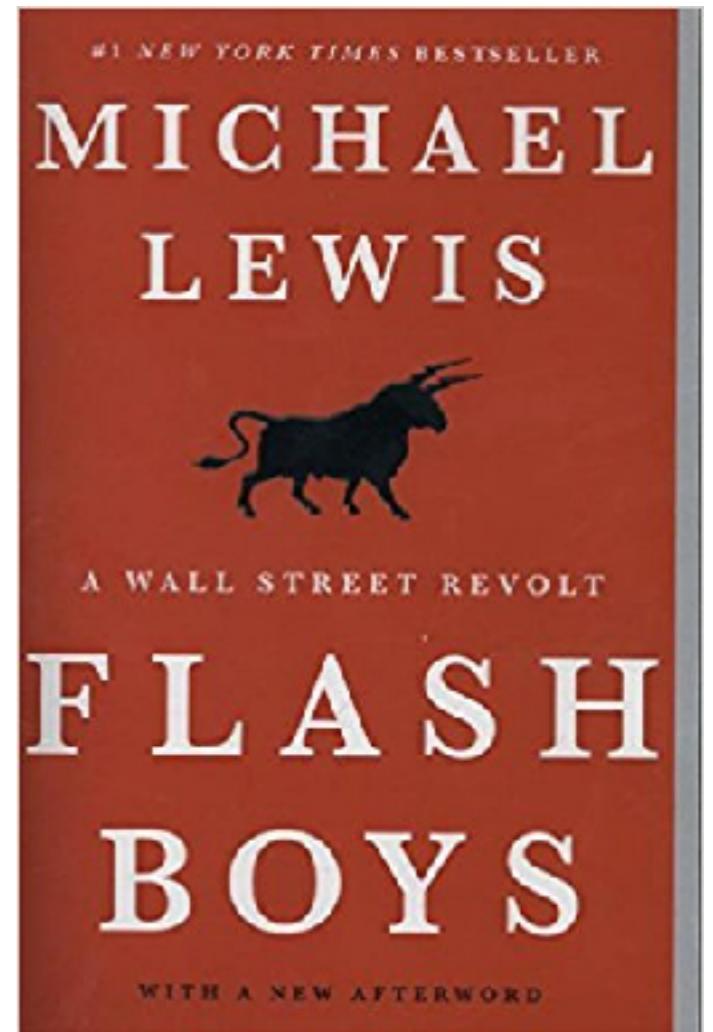


(c) win probability vs. delay

Application
Presentation
Session
Transport
Network
Datalink
Physical

Network Latency and HFT

- High Frequency Trading (HFT) is an algorithmic trading strategy using computer which carries out sell and buy within a second or less.
- HFT represents large share in today.
- The trading system in Tokyo Stock Exchange (TSE) has been ready for HFT since 2011.
- Dealers prefer to install the server at the physically nearest location to the markets.
- Markets offer co-location space, or datacenter (DC) to dealers. For fairness among dealers, NYSE DC ensures the same network latencies between servers with optical fiber spools.

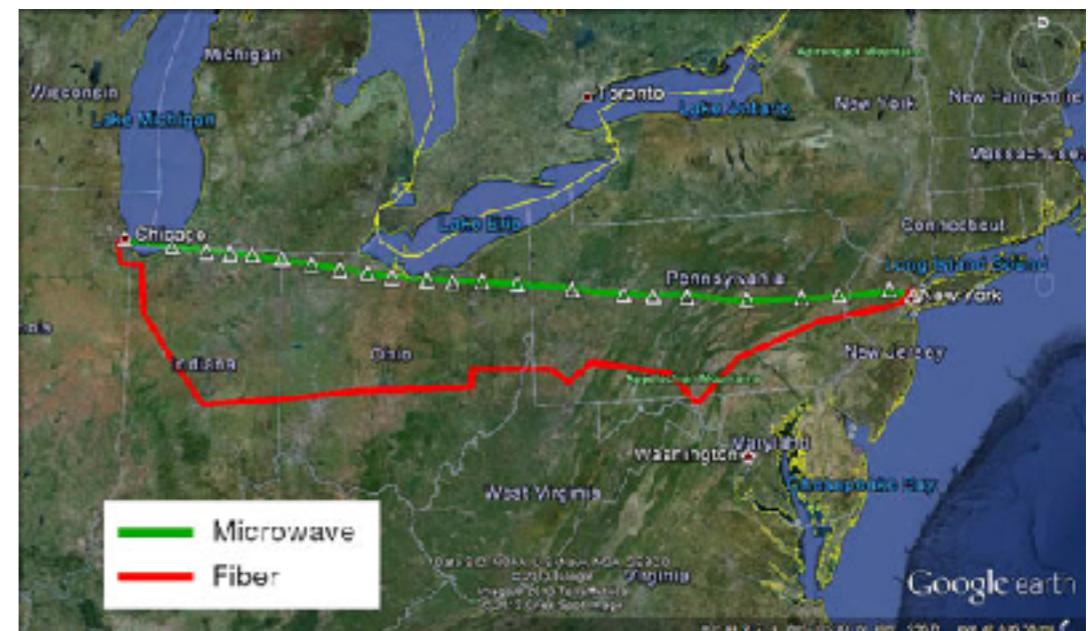


Michael Lewis, "Flash Boys: A Wall Street Revolt", 2015

Application
Presentation
Session
Transport
Network
Datalink
Physical

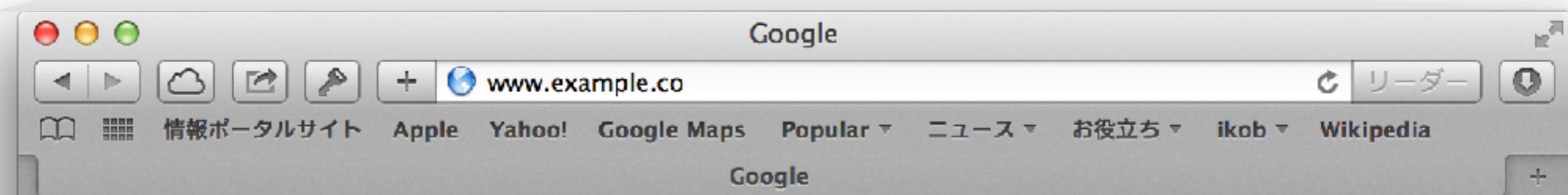
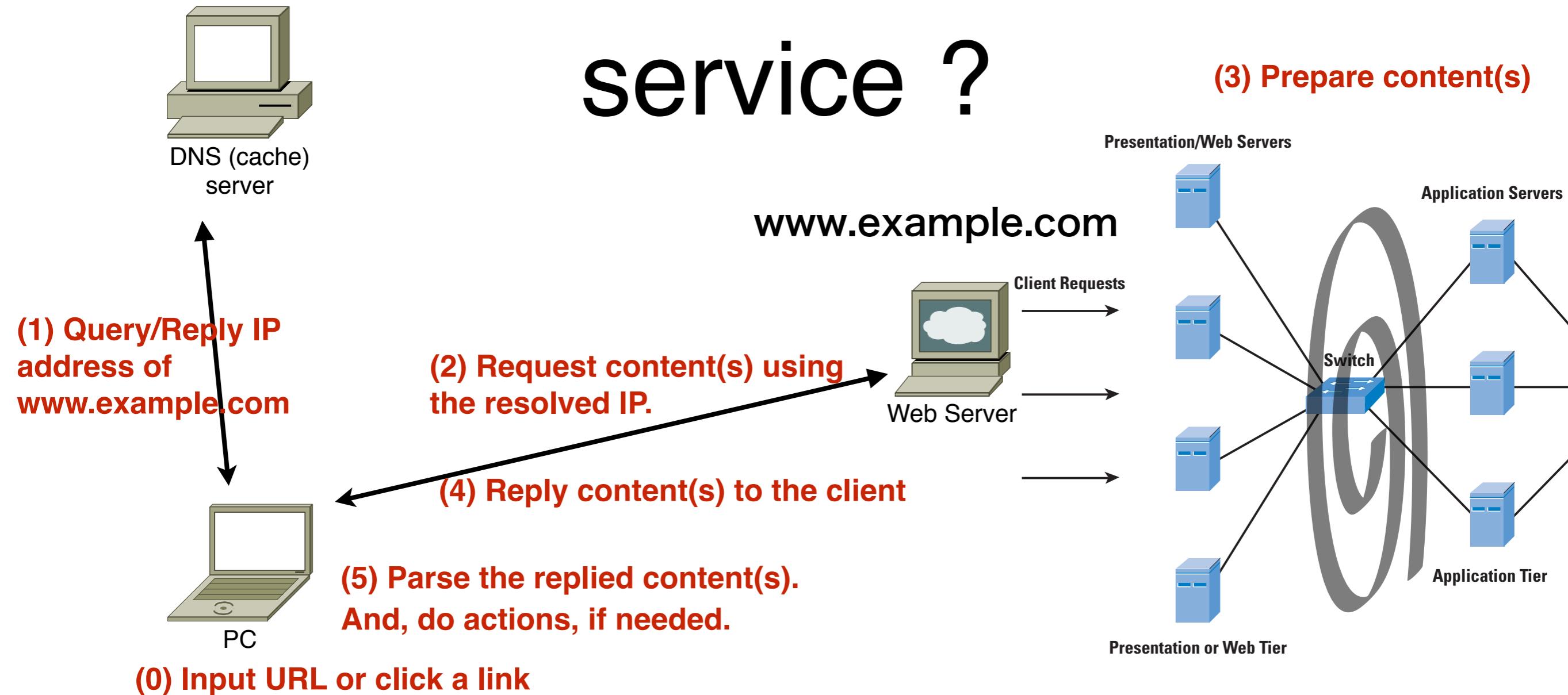
Network Latency and HFT(cont'd)

- Microwave wireless Local-Loop (WLL) comes back
 - WLL provides point-to-point leased-circuit services.
 - In terms of bandwidth capacity, WLL, < 150Mbps, has disadvantage against today's fibre capacities, > 10Gbps.
 - But, of latency, WLL has advantages in
 - Network Path Length: WLL can take straight path between connecting points but fiber paths depend on
 - Speed of light : WLL is faster than optical fiber.
in vacuum : $c = 3 \times 10^8 \text{ m/sec.}$
in optical fiber ($N = 1.5$) : $2 \times 10^8 \text{ m/sec.}$

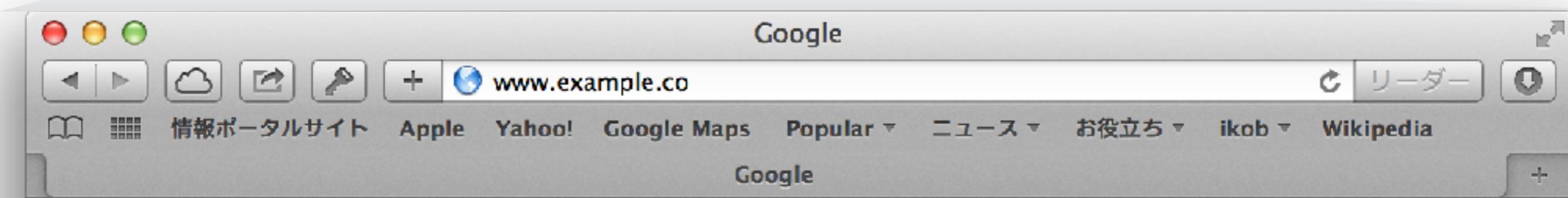
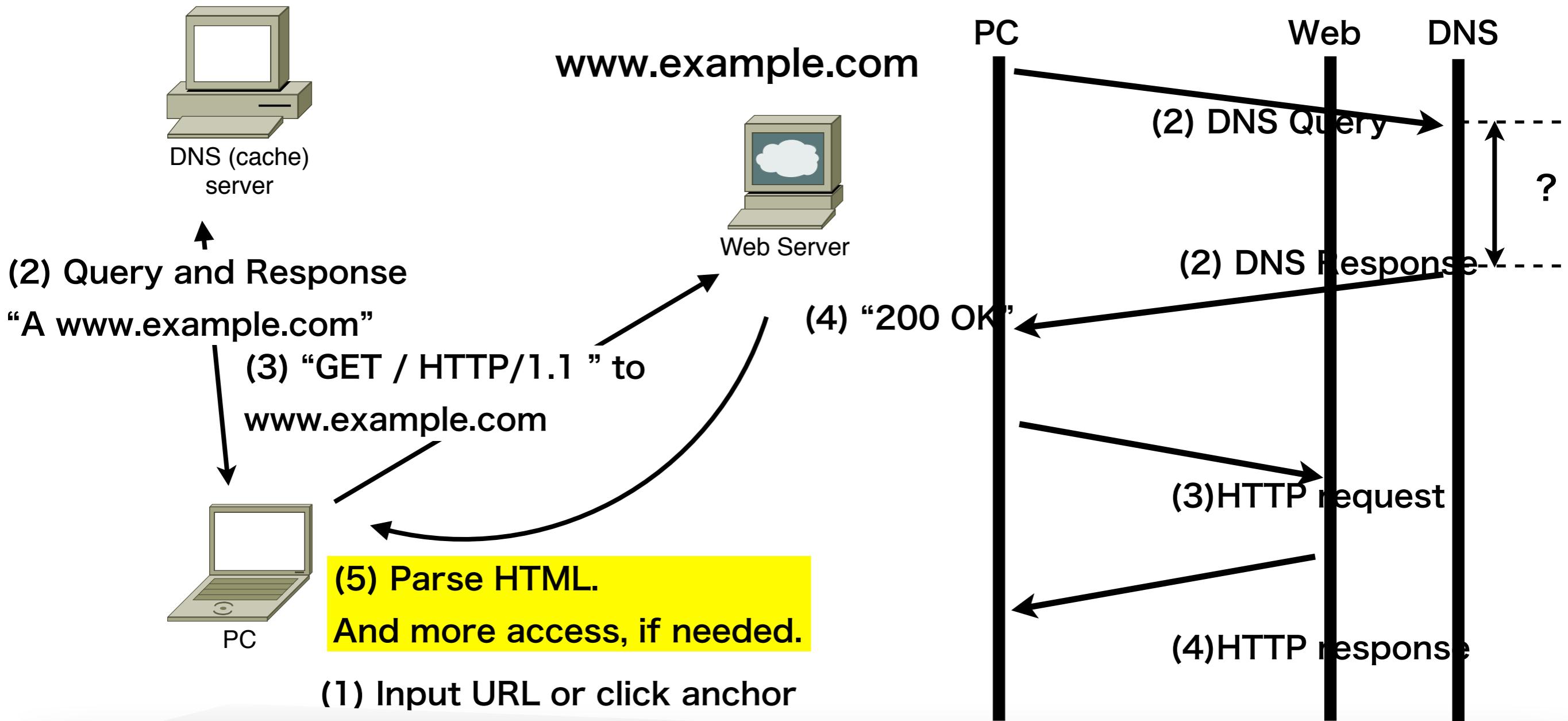


<http://www.exaltcom.com/HFT.aspx>

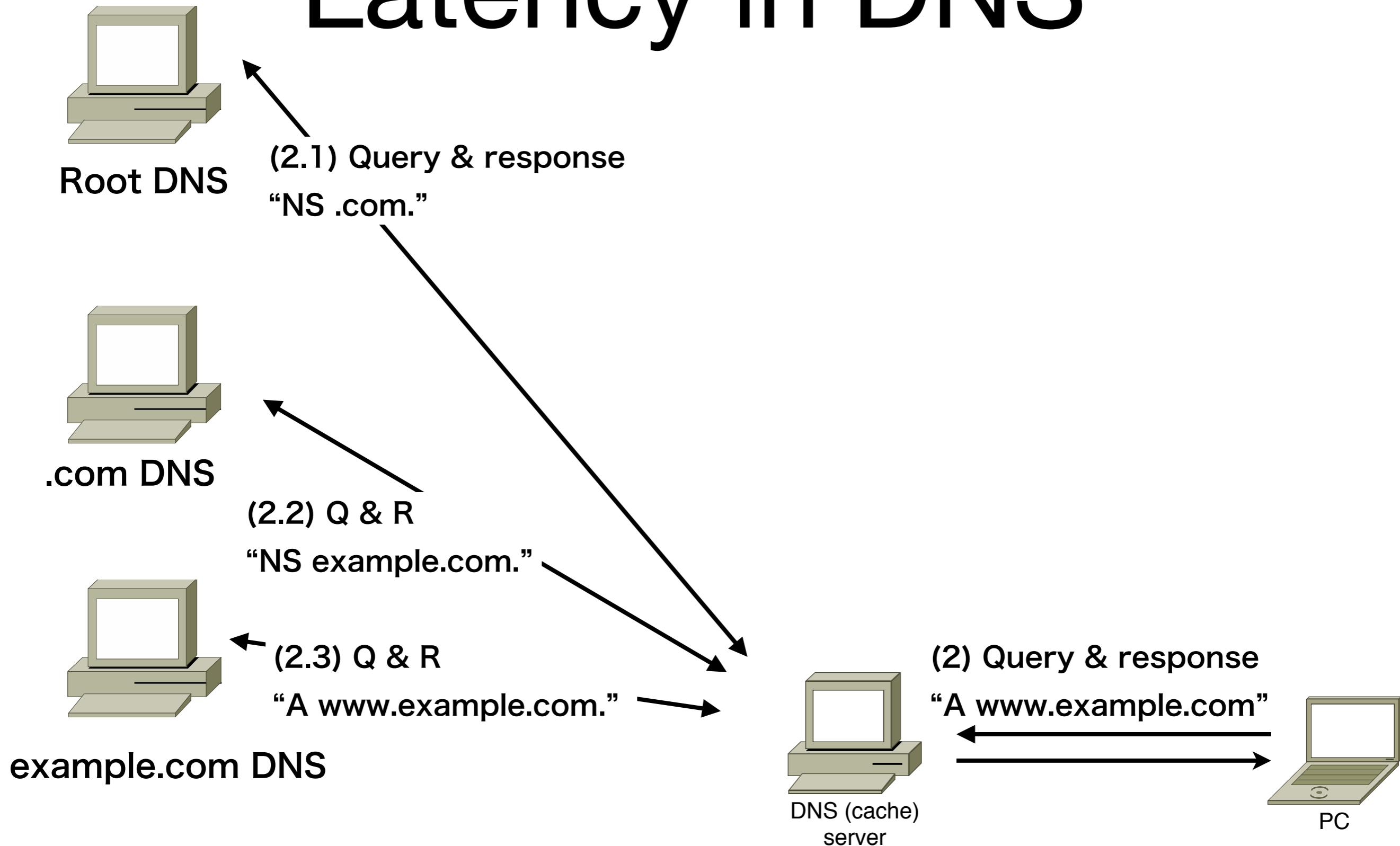
How client accesses Web service ?

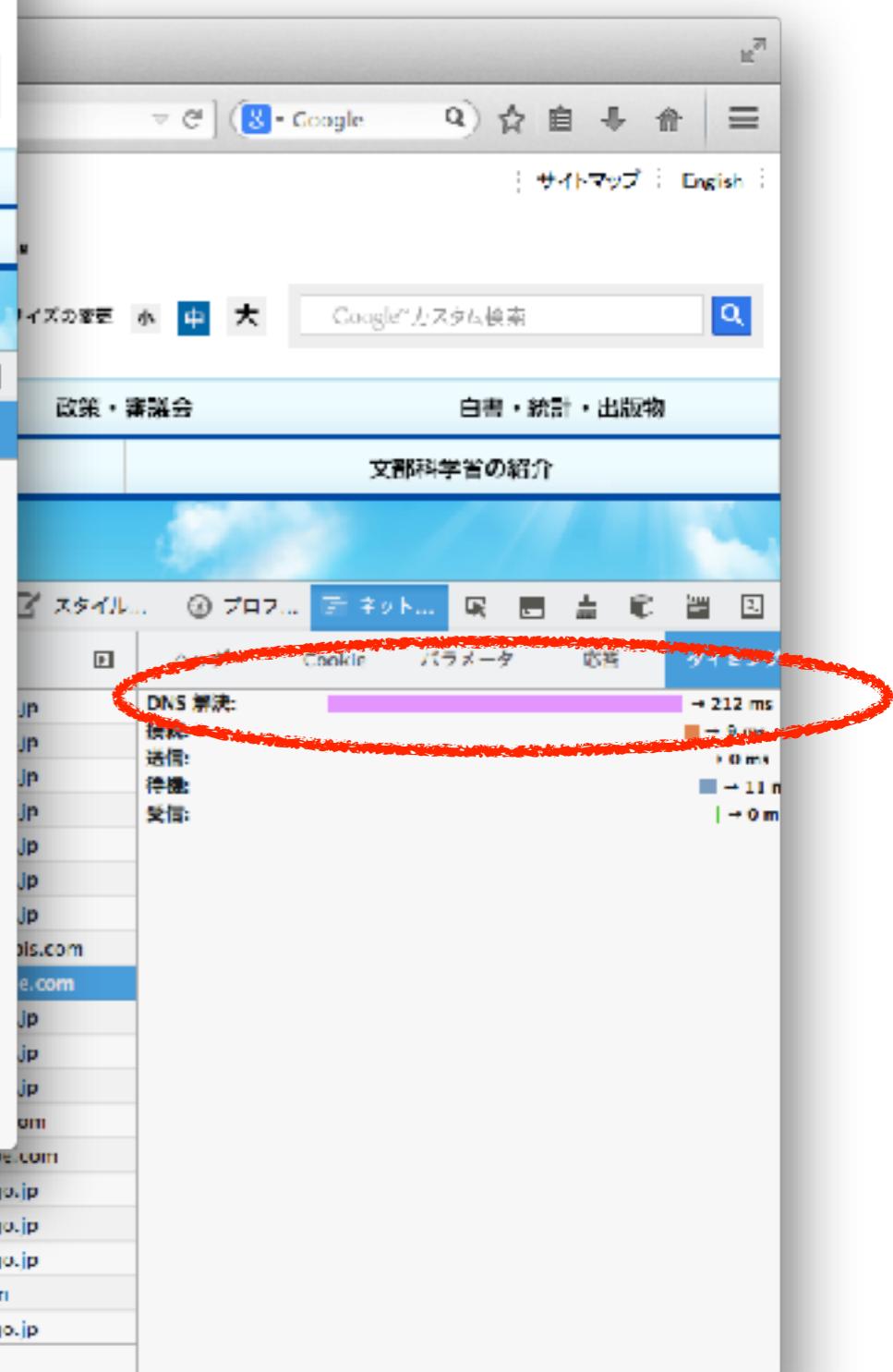
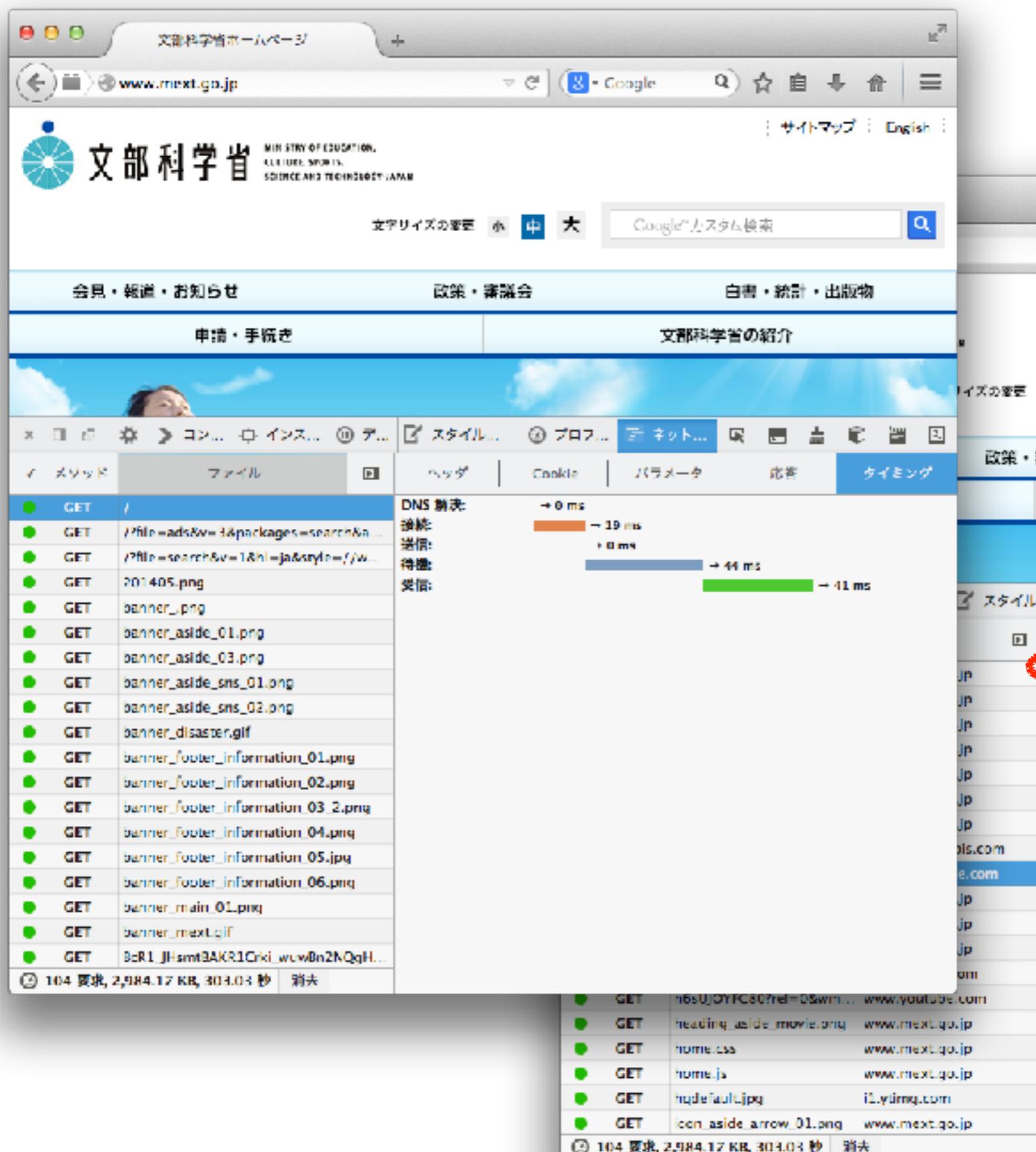


How client accesses Web service ?



Latency in DNS





Google Public DNS

- Client just sets 8.8.8.8, 8.8.4.4 as DNS servers provided by Google.
- Improve query response by mitigating cache miss:
 - Provisioning serving clusters
 - Reliable service
 - Load-balancing for shared caching
 - Distributing serving clusters for wide geographical coverage
 - Anycast : same IP address shared by more than one ends.
 - Better security

Timelines on Web inspector

The screenshot shows the MEXT homepage with a large banner featuring a woman looking up at the sky and the text "未来に向かって行く力を伸ばしたい". To the right is a "Topics" sidebar with a blue header and a search bar. The main content area has a light blue background with white clouds.

The Firefox Web Inspector is overlaid on the page, specifically the Network tab. The table lists 21 network requests:

メソッド	ファイル	ドメイン	タイプ	サイズ	0 ms	120 ms	640 ms	960 ms	1.28 秒	1.60 秒	1.82 秒	2.24 秒
GET	/	www.mext.go.jp	html	55.29 KB	→ 10 ms							
GET	common.css	www.mext.go.jp	css	15.61 KB		→ 13 ms						
GET	home.css	www.mext.go.jp	css	9.65 KB		→ 35 ms						
GET	smp_top.css	www.mext.go.jp	css	12.49 KB		→ 14 ms						
GET	cookieTop.js	www.mext.go.jp	js	3.24 KB		→ 15 ms						
GET	jquery-1.8.3.min.js	www.mext.go.jp	js	91.44 KB		→ 31 ms						
GET	search.js	www.mext.go.jp	js	4.08 KB		→ 31 ms						
GET	jsapi	www.google.com	js	24.10 KB		→ 41 ms						
GET	jquery.megamenu-2.1.js	www.mext.go.jp	js	7.15 KB		→ 42 ms						
GET	jquery.carouFredSel-6.2.0-packed.js	www.mext.go.jp	js	35.23 KB		→ 49 ms						
GET	jquery.cycle.all.js	www.mext.go.jp	js	51.97 KB		→ 49 ms						
GET	jquery.closeFooter.js	www.mext.go.jp	js	0.59 KB		→ 49 ms						
GET	home.js	www.mext.go.jp	js	1.52 KB		→ 51 ms						
GET	logo.png	www.mext.go.jp	png	7.98 KB		→ 6 ms						
GET	?file=search&v=1&hl=ja&style=/www.goo...	www.google.com	js	0.76 KB		→ 43 ms						
GET	g_search_design.css	www.mext.go.jp	css	3.96 KB		→ 6 ms						
GET	text_size-adjust_01.nnn	www.mext.go.jp	nnn	0.59 KB		→ 6 ms						

At the bottom, there are tabs for HTML, CSS, JS, XHR, Font, Image, Media, and Flash, with JS selected. A status bar at the bottom right shows "81 要求, 2,878.90 KB, 2.43 秒 消云".

Connection View

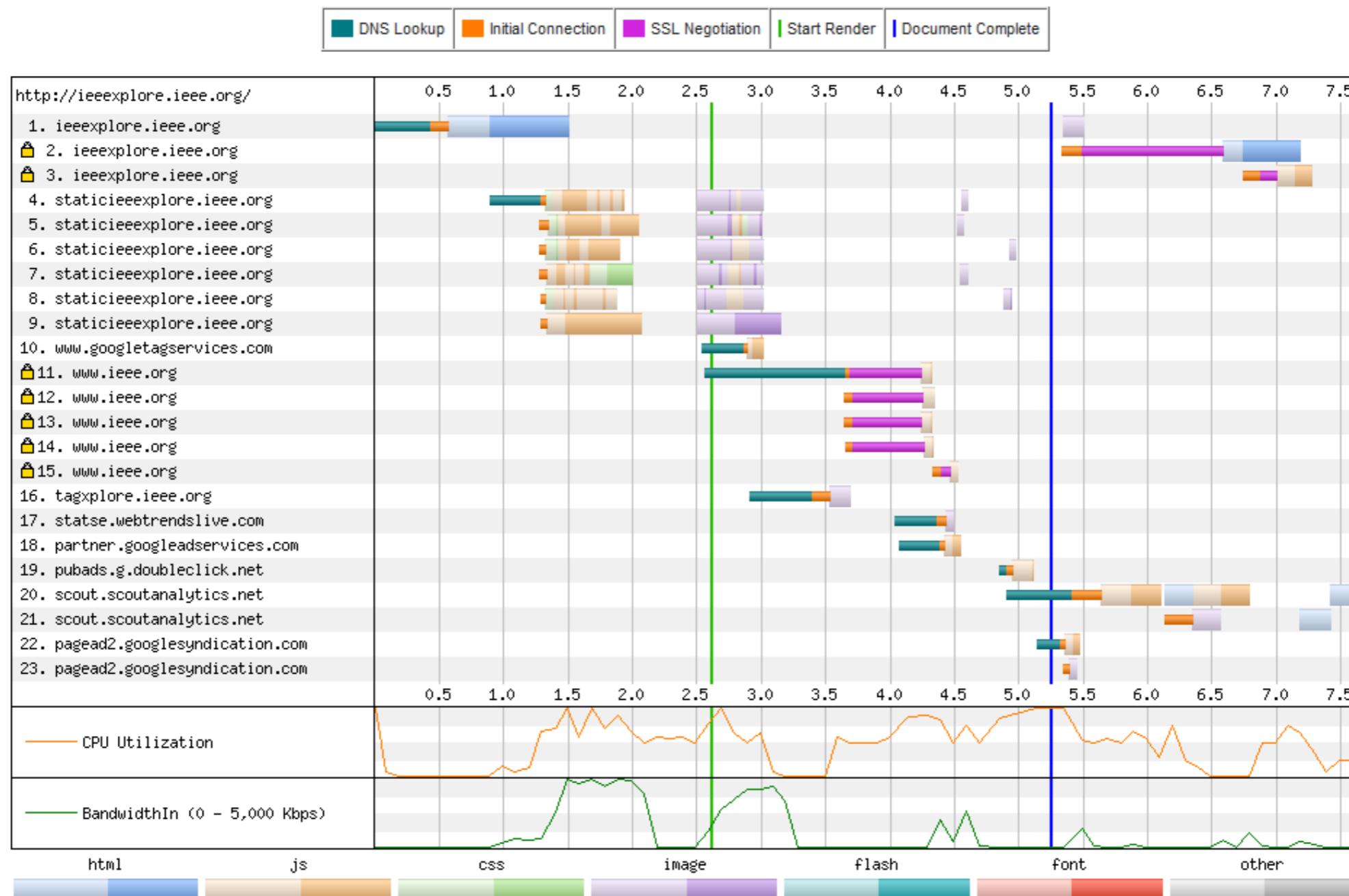
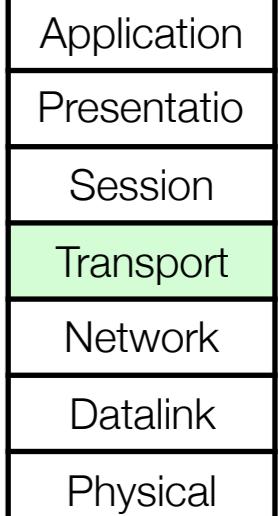


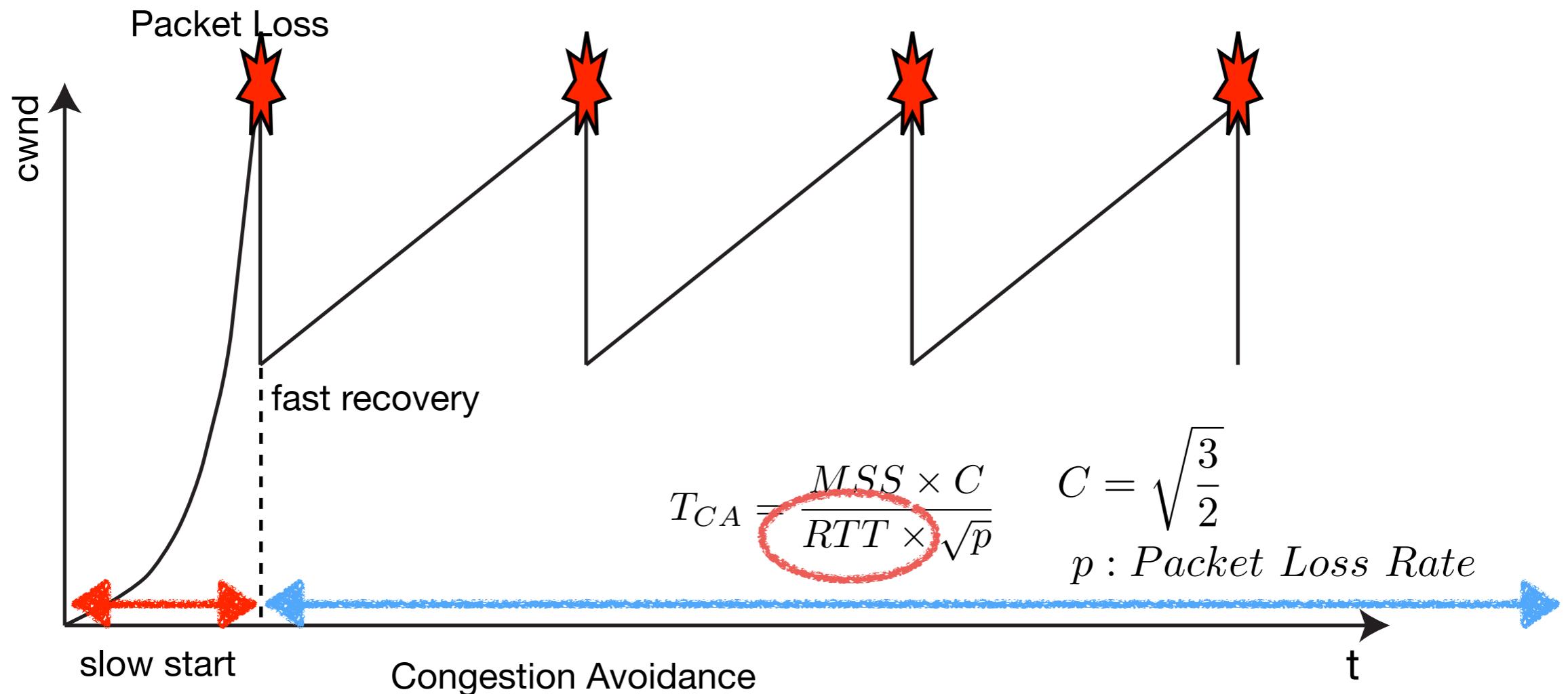
Fig. 1. Waterfall diagram showing the timing of download of an apparently uncluttered example Web page (ieeexplore.ieee.org), actually comprising over one hundred objects, transferred over 23 connections needing 10 different DNS look-ups. The horizontal scale is in seconds. This access was from Stockholm, Sweden, over a 28ms RTT 5 Mb/s down 1 Mb/s up cable access link, using Internet Explorer v8 without any prior cache warming. Source: www.webpagetest.org

Outline

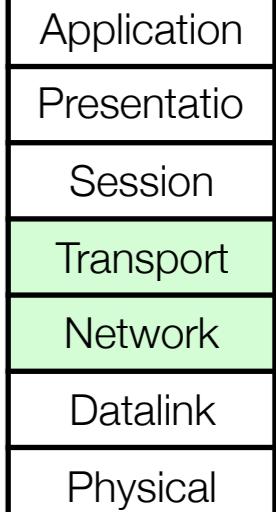
- Administravia
- Homework review
- User eXperience on Web service
- NS2 simulation



TCP window behavior on NewReno



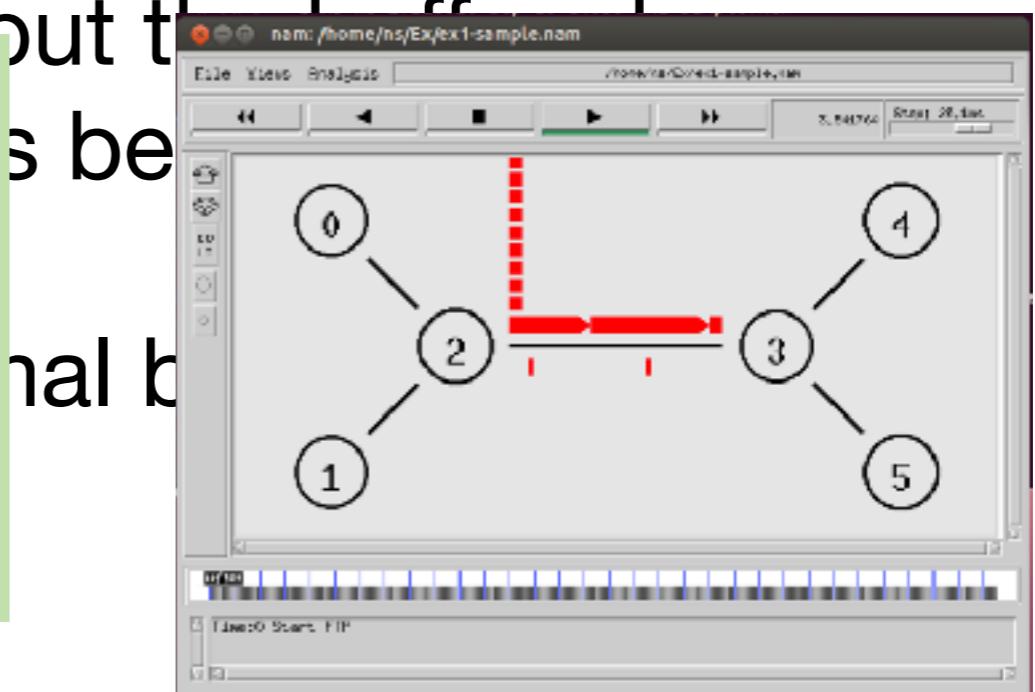
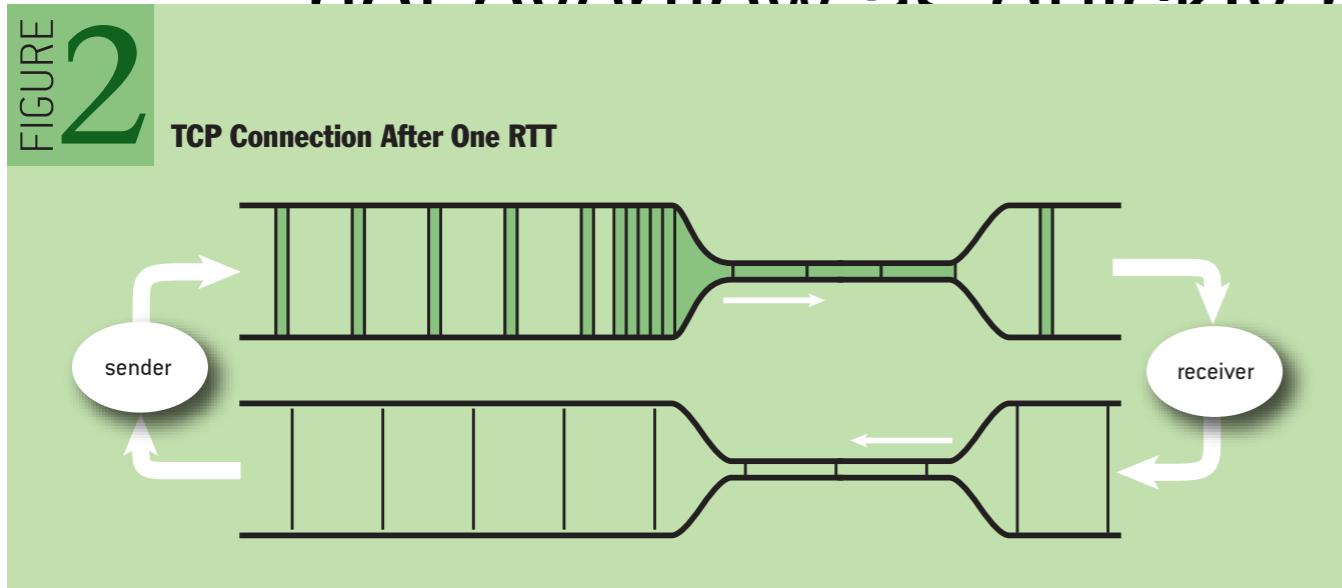
Mathis, Matthew, et al. "The macroscopic behavior of the TCP congestion avoidance algorithm." ACM SIGCOMM Computer Communication Review 27.3 (1997): 67-82.



Queue / packet buffer at router interface

- Absorb burst traffic caused by different rate links.
- In case of small sizes: Unable to absorb large burst.
- Large: longer queuing delays. Buffer space does not overflow as quickly but the loss rate is higher.

FIGURE
2

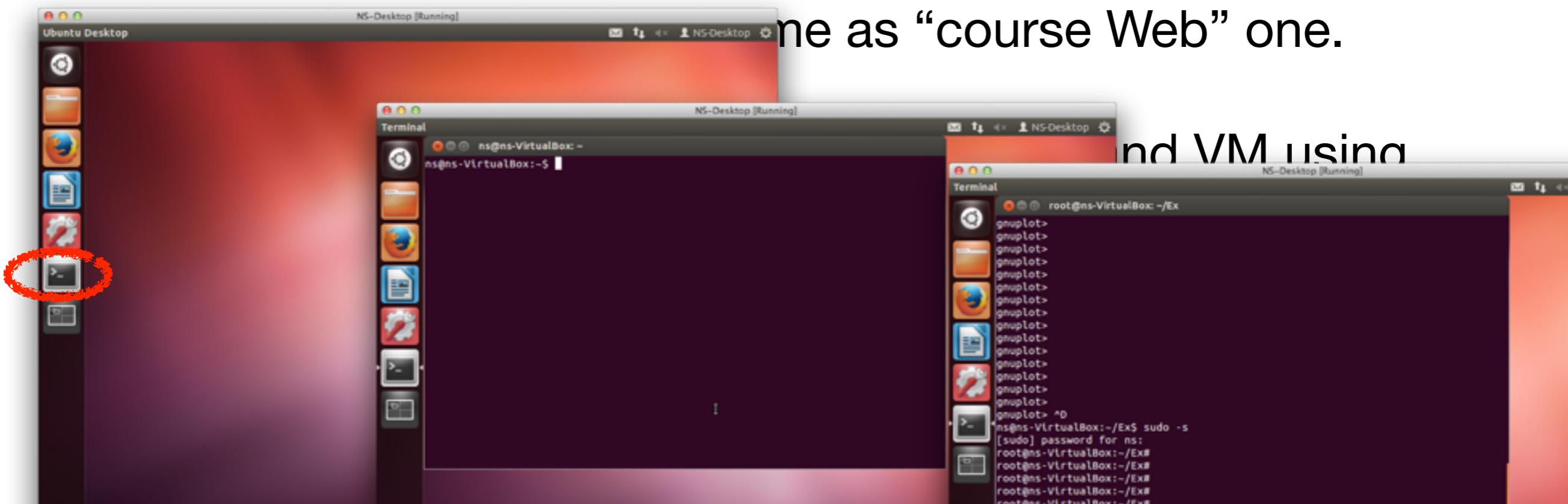


Installing NS2 simulator with Oracle VirtualBox

- Download and install “VirtualBox Hypervisor” application package from
<https://www.virtualbox.org>
- Download VM package (OVA1.0) from the course Web site
- Start “VirtualBox”
- Import VM package
- Start VM
 - user/password is the same as “course Web” one.

Installing NS2 simulator with Oracle VirtualBox (cont'd)

- Start VM
 - Ubuntu Desktop will be appeared.
 - Start console



Booting NS2 simulator with Docker

- Install Docker platform from docker.com
- Prepare X11 server application, such as:
XQuartz for MacOSX, MobaXterm for Windows.
- Find and set the IP address of Docker plathome as:

```
$ ip=$(ifconfig en0 | grep inet | awk '$1=="inet" {print $2}')
```

- Start X11 server and allow remote access as:

```
$ xhost +
```

- Create a directory to share between the container and server. Then copy sample script as:

```
$ mkdir ~/ex-1
$ cp <Download file> ~/ex-1/
```

- Start docker container as:

```
$ docker run -it -v ~/ex-1:/opt \
-e DISPLAY=${ip}:0 ekiourk/ns2 bash
root@e2e1dbbc0950:/ns2/ns-2.35#
```

- Verify shared directory contents as:

```
root@e2e1dbbc0950:/ns2/ns-2.35# ls /opt
ex1-sample.tcl
root@e2e1dbbc0950:/ns2/ns-2.35#
```

Hello NS2 !!

- Test “tcl” command as:

```
ns@ns-VirtualBox:~/Ex$ ns
% puts "Hello World!"
Hello World!
% ^D
ns@ns-VirtualBox:~/Ex$
```

NS2 (Network Simulator 2)

- Has been developed by VINT (Virtual InterNetwork Testbed) Project funded by US DARPA(DefenseAdvanced Research Projects Agency)
- Simulates packet switched network with discrete event driven simulation approach.
 - Transport, Routing, Wireless link, mobile
 - A lot of remarkable outputs from NS2
 - esp. to improve transport behavior, such as, TCP.

NS2 (Network Simulator 2) (cont'd)

- Written by oTCL (object oriented extension for TCL) and C++
 - Easy to implement and to evaluate new ideas.
 - Simulation scenarios are described by oTCL program.
 - A network specified as a graph which consists of nodes and links.
 - Transport stacks and applications are agents attached to nodes .
- Though NS3 project successor of NS2 has been started, this lecture uses NS2 due to a visualization issue in NS3.

NS2: Run simulation with script

- Download the simulation script, ex1-sample.tcl, from the course Web.
- Run NS2 with the script.
- Run a packet animation application, “nam”.

```
ns@ns-VirtualBox:~/Ex$ ls  
ex1-sample.tcl  
ns@ns-VirtualBox:~/Ex$ ns ex1.sample.tcl  
ns@ns-VirtualBox:~/Ex$ ls  
ex1-sample.nam  ex1-sample.tcl  ex1-sample.tcp  ex1-sample.tr  
ns@ns-VirtualBox:~/Ex$  
ns@ns-VirtualBox:~/Ex$  
ns@ns-VirtualBox:~/Ex$ nam &
```

Saved trace and animation files.

Today's quiz

- After download “ex1-sample.tcl” script from the course Web, Run it using “ns”
- Display Network AniMation (NAM) animation with the generated file as “ex1-sample.nam”.
- Upload captured snapshot via the course web **during this class.**

本日のクイズ

- “ex1-sample.tcl” スクリプトを講義 Web よりダウンロードした後、NS でシミュレーションを実行せよ。
- 生成された “ex1-sample.nam” を Network AniMation (NAM) で表示せよ。
- 画面のスナップショットを クラス時間内 に講義 Web からアップロードせよ。