

コンピュータ ネットワーク

トランスポート層 / NAT
2011/12/7

前回の復習

経路制御プロトコル

- ・概念と仕組み
- ・アルゴリズム

プロトコル例

- ・RIP
- ・OSPF
- ・BGP

経路制御の階層化

本日の流れ

トランスポート層 プロトコル

- ・ポート番号
- ・UDP
- ・TCP

NAT 技術

ドメイン名

2011/12/07

2

トランスポート層

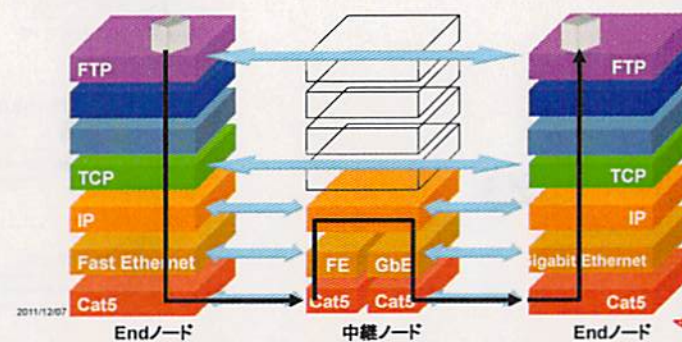
2011/12/07

3

トランスポート層

Layer 4 と呼ばれる

データを運ぶ際の取り決め



2011/12/07

データフォーマット

Ethernet ヘッダ	IP ヘッダ	TCP or UDP ヘッダ	データ	FCS
-----------------	-----------	-------------------	-----	-----

UDP ヘッダ

送信元 ポート番号	宛先 ポート番号
長さ	Checksum

TCP ヘッダ

送信元 ポート番号	宛先 ポート番号
シーケンス番号	
Acknowledge 番号	
Flag	ウィンドウサイズ
Checksum	緊急ポインタ

2011/12/07

ポート番号

IP アドレス

- データの宛先ホストを示す

ポート番号

- ホスト上での宛先アプリケーション (サービス) を示す
- ホストは受信したデータをどのアプリケーションに渡すかを宛先ポートにて判断

ポート番号の定義

- 0~1023 : Well known port
- 1024~49151 : Registered port
- 49152~65535 : Dynamic port

2011/12/07

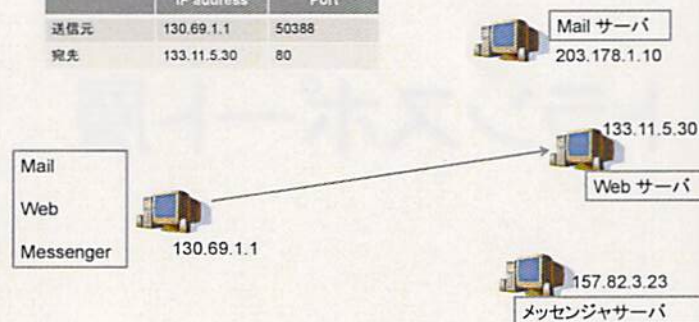
IPアドレスとポート番号



2011/12/07

IPアドレスとポート番号

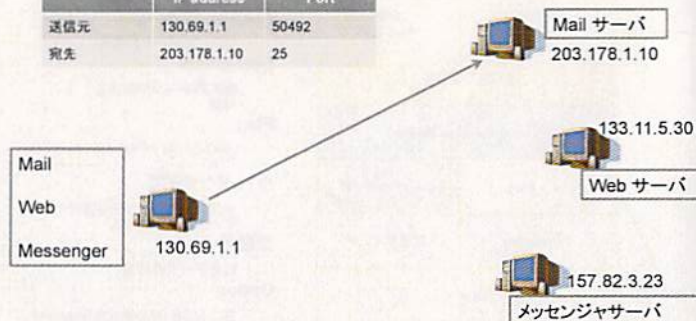
	IP address	Port
送信元	130.69.1.1	50388
宛先	133.11.5.30	80



2011/12/07

IPアドレスとポート番号

	IP address	Port
送信元	130.69.1.1	50492
宛先	203.178.1.10	25

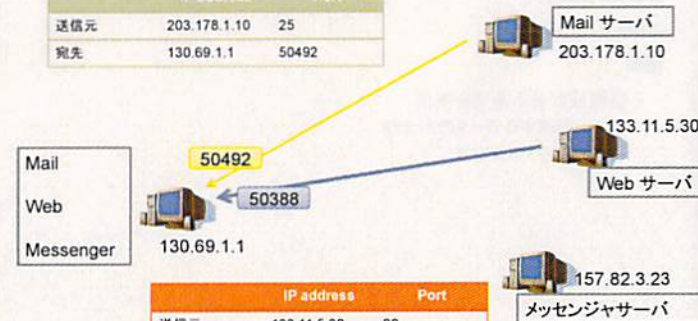


2011/12/07

9

IPアドレスとポート番号

	IP address	Port
送信元	203.178.1.10	25
宛先	130.69.1.1	50492



2011/12/07

10

ポート番号とサービス

UNIX 系 OS

- /etc/services

Windows

- C:\Windows\system32\drivers\services

ftp	21/udp	# File Transfer [Control]
ftp	21/tcp	# File Transfer [Control]
ssh	22/udp	# SSH Remote Login Protocol
ssh	22/tcp	# SSH Remote Login Protocol
telnet	23/udp	# Telnet
telnet	23/tcp	# Telnet
smtp	25/udp	# Simple Mail Transfer
smtp	25/tcp	# Simple Mail Transfer

2011/12/07

11

UDP

User Datagram Protocol

RFC768

信頼性の無い通信

UDP ヘッダ

- 長さ: UDP ヘッダ + データのバイト数
- Checksum: 整合性のチェックに利用

16bit	16bit
送信元 ポート番号	宛先 ポート番号
長さ	Checksum

2011/12/07

12

TCP

コネクション指向
ストリーム型通信

特徴

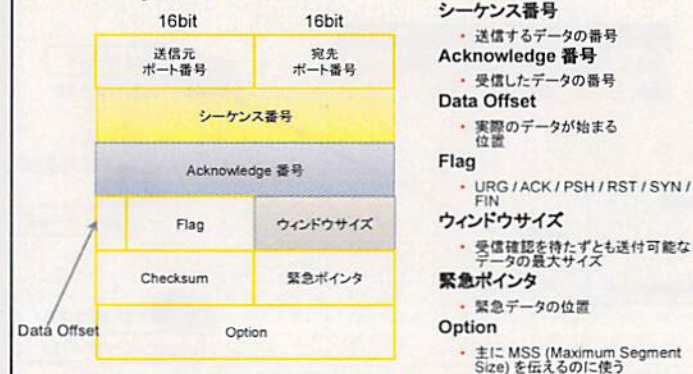
- 信頼性がある通信を提供
 - 到達するデータの完全性
 - データの逐次性

2011/12/07

13

TCP ヘッダ

14

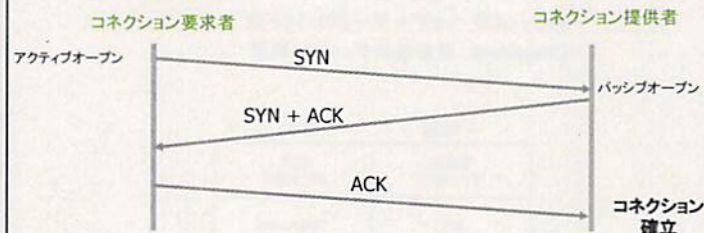


2011/12/07

TCPの仕組み (コネクションの確立方法)

3 way handshake

- SYN: コネクションを初期化する
- ACK: 確認応答



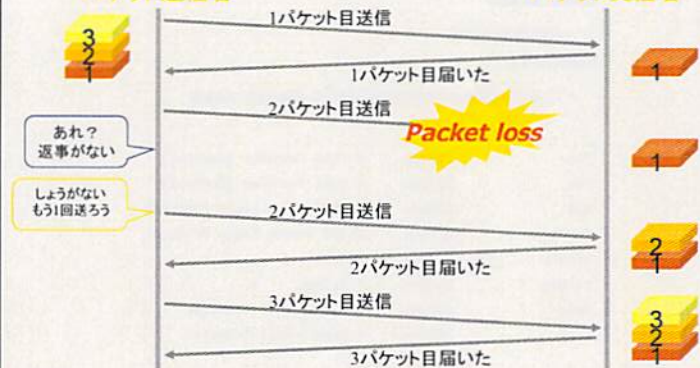
2011/12/07

15

TCPの仕組み(再転送)

パケット送信者

パケット受信者



2011/12/07

16

TCPの仕組み (送受信データ量の制御)

フローコントロール

- 相手のバッファ要領にあわせたフロー制御

コンジェスチョンコントロール

- 途中経路でのパケットの損失に対処

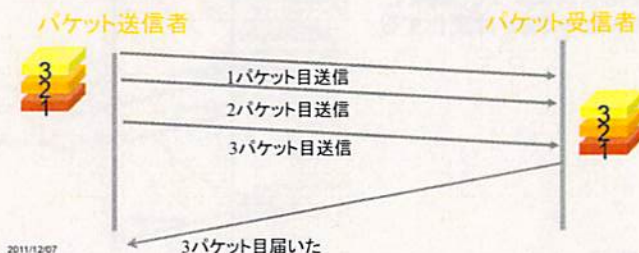
2011/12/07

TCPの仕組み(まとめ送り)

Ack を待たずに送り出す

まとめて送れるデータ量 = Window Size

1パケットの最大サイズ = MSS (Maximum Segment Size)



2011/12/07

シーケンス番号

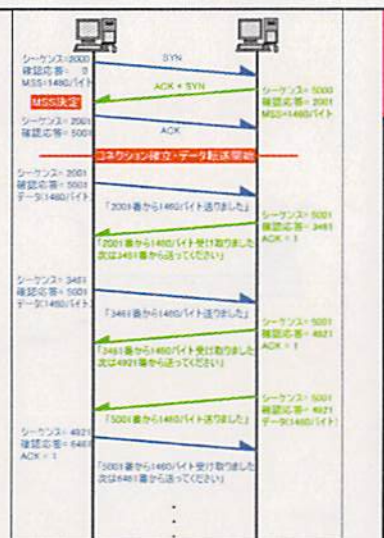
データ列の番号

初期値はランダム

- 右の場合は 2000 と 5000

送った byte 分だけ
番号を増加させる

Ack は受け取った
番号 + 1 の番号で
返す

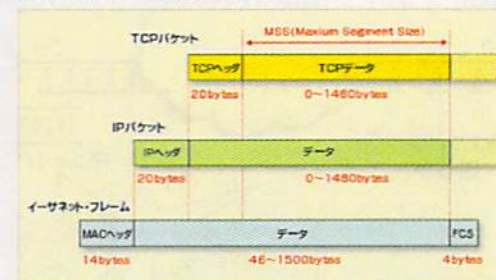


2011/12/07

MSS (MAXIMUM SEGMENT SIZE)

1. パケットで送れる最大実データ量

TCP のオプションとして通知可能



2011/12/07

出典：http://www.atmarkit.co.jp/iten2k/network/baswinian016/baswinian016_02.html

Window Size

-
- Figure 1 is a network diagram illustrating a client-server interaction. On the left, a box labeled 'クライアント' (Client) is connected to a computer icon. On the right, a box labeled 'サーバー' (Server) is connected to a computer icon. The client sends a message 'SEND0001' to the server. The server responds with a message 'SEND0001'. The diagram shows the flow of data between the client and the server, with the client's request and the server's response being the primary focus.

2011/12/07

出處：<http://atnetwork.info/cpio/cpio104.htm>

相手があふれたら、送信者が加減する

-

2011/12/07

②

③ 送るか。

①

早すぎてパツファがあふれる

2011/12/07

23

②

③ じゃ早くくるか

送信者

受信者

キュー(バッファ)

③ じゃ早くおくるか

①

遅いから余裕があるな

2011/12/07

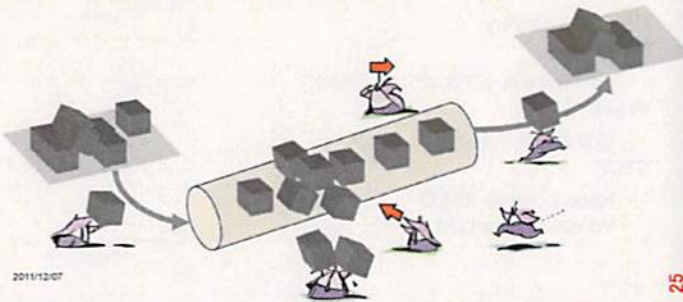
10

コンジェスチョンコントロール

congestion = 輻輳

ネットワークが混んでたら、やっぱり送信者が加減する

- 電話や放送にはないコントロール



25

コンジェスチョンコントロール

- ② もうちょっとゆっくり送ろう。



- ① 受信者からしばらく応答がない
輻輳(コンジェスチョン)が発生して
パケットがとどいてなさそうだ

2011/12/07

26

スロースタート

Window Size に関わらず、最初はず 1 パケット送信する

- => 無事 Ack が届く
今度は 2 パケット連続で送る
- => 無事 Ack が届く
次は 4 パケット連続で...

閾値まで繰り返す

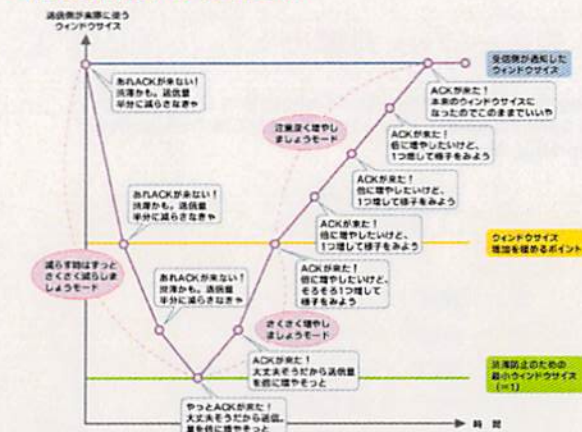
Ack が返ってこなかった場合

- 同時送信パケット数を「半分」にする

2011/12/07

27

輻輳制御の仕組み



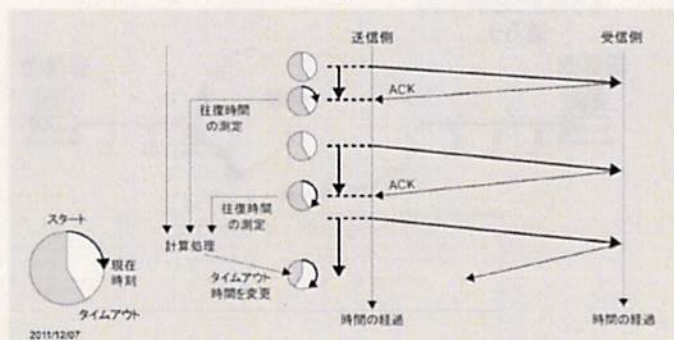
2011/12/07

出典: <http://www.atmarkit.co.jp/network/kenai/tcp11/02.html>

28

タイムアウト

Ack 待ち時間を決めるために RTT (Round Trip Time) を利用



29

TCP アルゴリズムの種類

Tahoe

- ・スロースタート

Reno

- ・Fast Recovery

NewReno

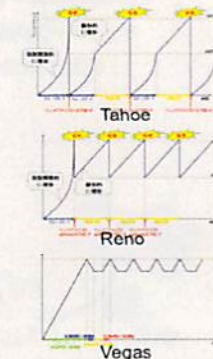
- ・Fast Recovery アルゴリズムの修正

Vegas

- ・輻輳の発生予想に RTT を利用

CTCP

- ・Reno と Vegas の複合
- ・Windows Vista 以降



30

TCP スループットの算出

Promoting the Use of End-to-End Congestion Control in the Internet, Sally Floyd and Kevin Fall, IEEE/ACM Transactions on Networking, May 1999

$$T < \frac{1.5\sqrt{2/3} \times B}{R \times \sqrt{p}}$$

- T: 評価するトラフィック
B: 接続リンクでの MTU
R: RTT (Round Trip Time)
p: パケット喪失率

2011/12/07

31

TCPの仕組み (コネクション終了時)



32

NAT

2011/12/07

33

NAT とは ?

NAT

- Network Address Translation

NAPT

- Network Address and Port Translation
- 別名 IP Masquerade

Global IP address と Private IP address を変換する仕組み

- 家庭用のブロードバンドルータ
- 研究室内ネットワーク

2011/12/07

34

IPv4アドレスの不足

インターネットユーザ数の増加に伴い
IPv4アドレスが枯渇

各ISPは厳しい審査を経てIPアドレスを取得
限られたグローバルアドレスを有効に活用

- NAT/NAPT
 - 弊害もあるが現状で主に使用されている

対策技術

- IPv6
- 長期的な解決方法

2011/12/07

35

プライベートIPアドレス

- 限られた範囲内で一意になるアドレス

- そのままでは外と通信できない。
- 自宅やオフィスのLAN、ファイヤウォールの内側などで利用される

□ RFC1918

- 192.168.0.0/16 (192.168.0.0 – 192.168.255.255)
- 172.16.0.0/12 (172.16.0.0 – 172.31.255.255)
- 10.0.0.0/8 (10.0.0.0 – 10.255.255.255)



2011/12/07

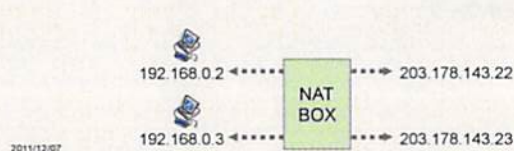
36

NAT

RFC1631: The IP Network Address Translation

プライベートIPアドレスをグローバルIPアドレスに変換

- 本来の意味ではIPアドレスの対応は1対1
 - 同時接続分のグローバルIPアドレスを消費する
- ヘッダのIPアドレス部分のみを変換
- 今日ではNAPT(後述)をNATと称するケースが多い



37

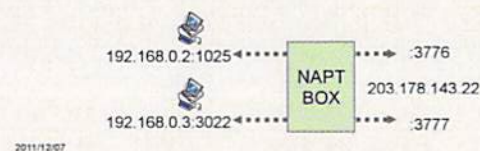
NAPT

Network Address Port Translation

- RFC2391: Load Sharing using IP Network Address Translation (LSNAT) 内に記述

IPアドレスに加え、ポート番号も変換する

- 1つのグローバルアドレスで複数のプライベートIPアドレスを持つ端末を接続できる



38

IPアドレスとポート番号 (NAT)

	IP address	Port
送信元	10.0.1.2	50872
宛先	133.11.5.30	80

Mail サーバ
203.178.1.10



	IP address	Port
送信元	130.69.2.2	3098
宛先	133.11.5.30	80

2011/12/07

39

IPアドレスとポート番号 (NAT)

	IP address	Port
送信元	133.11.5.30	80
宛先	10.0.1.2	50872

Mail サーバ
203.178.1.10



	IP address	Port
送信元	133.11.5.30	80
宛先	130.69.2.2	3098

2011/12/07

40

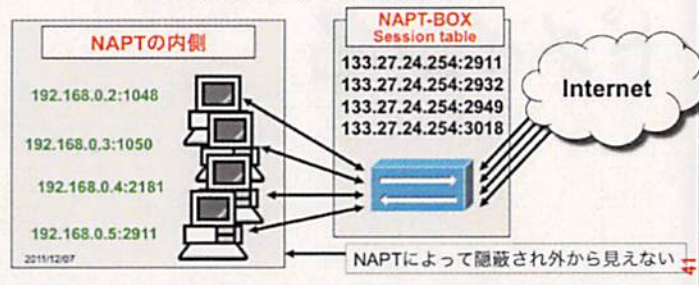
NAPTのメリット

グローバルIPアドレスの節約が可能

- 一つのグローバルアドレスを使って、複数のマシンが外部と通信できる

アクセス制御

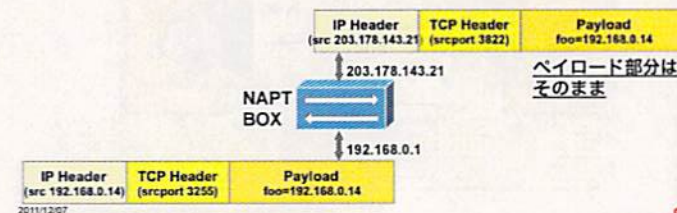
- 外部から内部ネットワークが隠蔽される



NAPTのデメリット(1)

NAT/NAPTを通過できないプロトコルがある

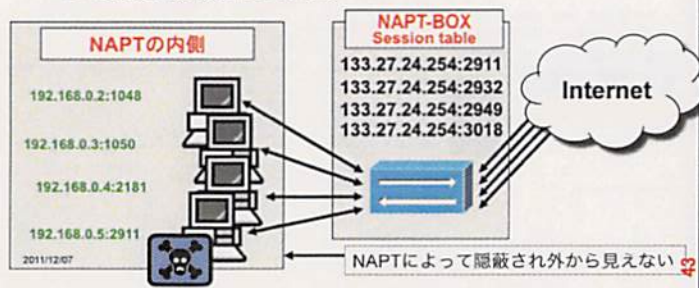
- ペイロード内にIPアドレスを含むプロトコルを利用するアプリケーション
 - 例: FTP, VoIP, NetMeeting, P2P
 - 特定のNAT実装のみで利用可能



NAT/NAPTのデメリット (2)

セキュリティ上の問題に対処しにくい

- NAPTの内側に悪意のあるホストがあっても、インターネットから見ると隠蔽されてしまう



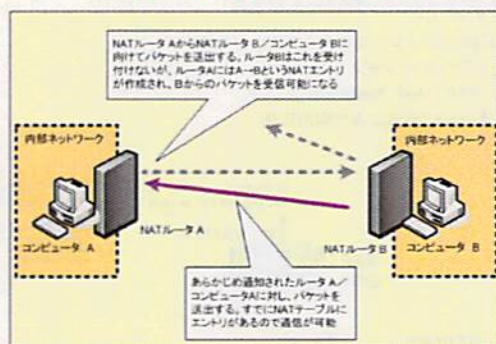
NAT TRAVERSAL

NATの欠点(仕様)である、NAT外 => NAT内の通信を実現する技術

- TURN (Traversal Using Relay NAT)
- ICE (Interactive Connectivity Establishment)
- STUN (Simple Traversal of UDP through NAT)
- UPnP (Universal Plug and Play)
- UDP Punching Hole

UPnPを除く他の技術は、一度 NAT 内部から外部のサーバに対してパケットを通過させることで、NAT 外 => 中の通信を可能とする

UDP パンチングホールの例



2011/12/07

出典: http://www.stimarket.co.jp/ser/2/experiments/udp/udp02_03.html

45

ドメイン名

46

ドメイン名と IP アドレス

ドメイン名とは

- www.u-tokyo.ac.jp
- www.yahoo.co.jp
- www.google.co.jp

IP アドレスとは

- 133.11.128.254
- 2001:200:180:299:217:f2ff:fe0e:d6d0

47

ドメイン名はここに



48

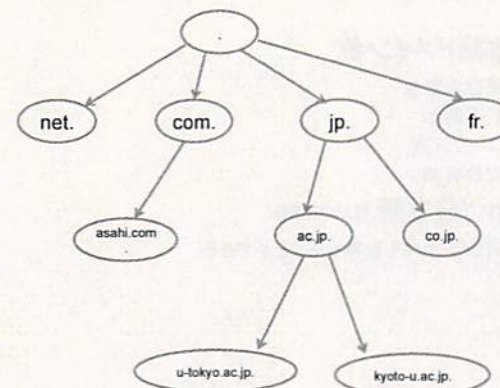
名前(ドメイン名)の仕組み

階層構造になっている

- www.itc.u-tokyo.ac.jp
 - jp 日本
 - ac 大学組織
 - u-tokyo 東京大学
 - itc 情報基盤センター
 - www ホストの名前

49

名前構造概念図



50

TLD (TOP LEVEL DOMAIN)

gTLD (Global Top Level Domain)

- com, net, org, edu, gov, mil, int, info, biz, name, pro, museum, aero, coop, jobs, travel, mobi, ...

ccTLD (Country Code Top Level Domain)

- 国(地域)別のトップレベルドメイン
 - jp, kr, cn, fr, nz, au, ...

51

日本のドメイン名

属性ドメイン

- co.jp, ac.jp, go.jp, ad.jp, ne.jp, or.jp, gr.jp, ed.jp, lg.jp

JP直下ドメイン

- u-tokyo.jp
- asahi.jp
- 誰でも購入することができる
- JPRS(日本レジストリサービス) が管理

52

日本語も使えるように

日本語ドメイン名

- 東京大学.jp
- 新宿駅.jp
- ローソン.jp
- 総務省.jp

<http://日本語.jp/case/>

- 同じくだれでも取得することができる

53

多言語ドメイン (IDN)

他の言語も当然存在する

- 한국어.kr
- 中文.cn
- Русский.ru

.com, .org, .net でも登録可能

- 스포츠조선.org
- com. מתנותלגבר

54

完全多言語化

完全多国語化

- 東京大学.日本
- 総務省.日本
- Web を見るには多言語ドメイン対応ブラウザが必要
 - IE8, Safari 3, Firefox 3 はどれも対応済み

メールも

- 関谷@情報基盤センター.東京大学.日本
- 対応したメールソフトウェアが必要
- まだこれから

55