# The Impact of Residential Broadband Traffic on Japanese ISP Backbones

-- SRCCS Workshop on Internet Measurement, Modeling, and Analysis --

Kenjiro Cho (IIJ / WIDE)
Kensuke Fukuda (NTT / WIDE)
Hiroshi Esaki (U.Tokyo / WIDE)
Akira Kato (U.Tokyo / WIDE)

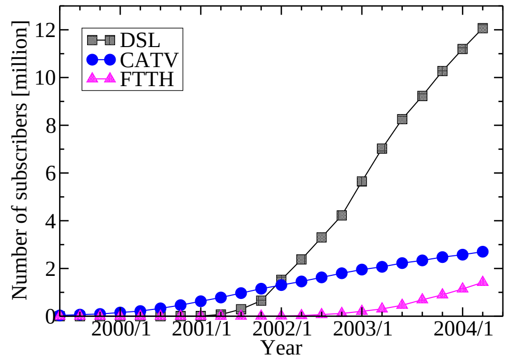


## major findings in our study

- our data is considered to cover 41% of total Japanese traffic
  - total RBB traffic in Japan is estimated to be about 250Gbps
- 70% of RBB traffic is constant, peak in the evening hours
- RBB traffic is much larger than office traffic, so backbone traffic is dominated by RBB traffic
- traffic volume exchanged via private peering is comparable with volume exchanged via major IXes
- within external traffic, international traffic is about 23% for inbound and about 17% for outbound
- regional RBB traffic is roughly proportional to regional population

## introduction

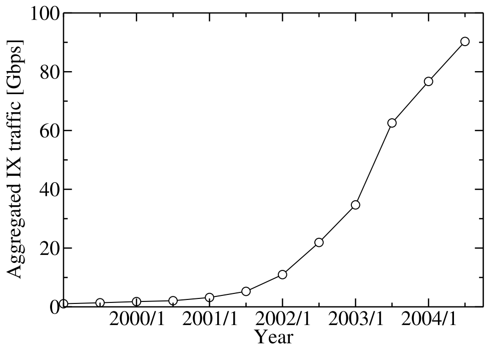
- rapidly growing residential broadband access
  - low-cost high-speed services, especially in Korea and Japan
- total RBB subscribers in Japan as of Feb 2004: 14.5 million
  - DSL:11 million, CATV:2.5million, FTTH:1 million






## unprecedented traffic increase in backbone

- traffic growth at the major Japanese IXes
  - how much is contributed by residential broadband traffic?



## background

- concerns about rapid growth of RBB traffic
  - backbone technologies will not keep up with RBB traffic
  - ISPs cannot invest in backbone simply for low-profit RBB
- ISPs and policy makers need to understand the effects of RBB
  - although most ISPs internally measure their traffic
    - data are seldom made available to others
    - measurement methods and policies differ from ISP to ISP

- to identify the macro-level impact of RBB traffic on ISP backbones
  - an unofficial study group was formed with specialists
    - members from 7 major Japanese ISPs and government
- goals: traffic measurement across multiple ISPs, to identify
  - ratio of RBB traffic to other traffic
  - changes in traffic patterns
  - regional differences



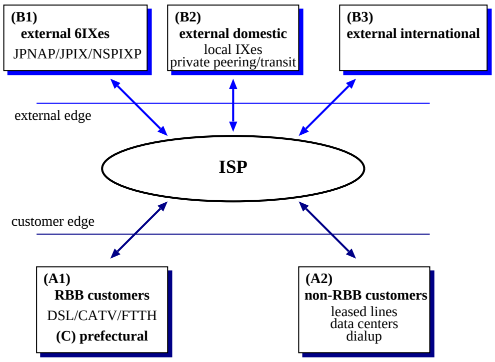
## traffic data collection across multiple ISPs

- requirements
  - find a common data set for all participating ISPs
    - focus on operational aspects
  - workload and investment for ISPs should not be high
  - data should be coarse not to reveal sensitive information but meaningful enough to analyze traffic
- challenges: mostly political or social, not technical

- we found that most ISPs use MRTG/RRDtool to monitor SNMP inOctet/outOctet of almost all routers in their service networks
  - if we can classify traffic into a common set, ISPs can provide aggregated traffic info

## traffic groups at ISP boundary for data collection

- focus on traffic crossing ISP boundaries (customer and external)
  - customer traffic is summable
  - external traffic could have double-counts (but small in our results since participating ISPs are peering with each other)






## descriptions of traffic groups

- 5 traffic groups selected by existing operational practice of the ISPs
  - it is not possile to draw a strict line on global Internet
    - e.g., residential/business, domestic/international

| traffic group | description | notes |
|---|---|---|
| **(A1) RBB customers** | residential broadband customer lines | includes small business customers using RBB |
| **(A2) non-RBB customers** | includes leased lines, data centers, dialup lines | may include RBB customers behind leased lines |
| **(B1) external 6IXes** | links for 6 major IXes (JP-NAP/JPIX/NSPIXP in Tokyo/Osaka) | |
| **(B2) external domestic** | external domestic links other than the 6IXes (regional IXes, private peering, transit) | domestic: both link-ends in Japan. includes domestic peering with global ASes |
| **(B3) external international** | external international links | |
| **(C) prefectural** | RBB links divided into 47 prefectures in Japan | prefectural links from 2 RBB carriers |

## methodology

- month-long traffic logs for the 5 traffic groups with 2-hour resolution
  - MRTG's resolution for monthly log
- a script to read and aggregate a list of MRTG/RRDtool logs
  - each ISP creates log lists and makes aggreagated logs by themselves without disclosing details
- another script to make graphs from the results using RRDtool
- biggest workload for ISP
  - creating lists by classifying large number of per-interface logs
    - some ISPs have more than 100,000 logs!
  - maintaining the lists
    - frequent planned and unplanned configuration changes

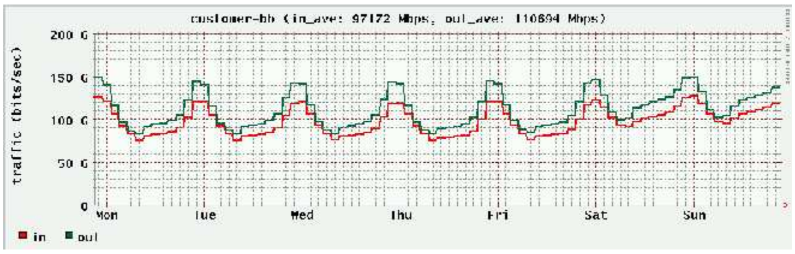


## results

- 2-hour resolution traffic logs for September and October 2004
  - by re-aggregating logs provided by 7 ISPs
- in weekly analysis, holidays are excluded
  - holiday traffic is closer to weekend traffic
- IN/OUT from ISPs' view

## RBB customer weekly traffic
### in September 2004

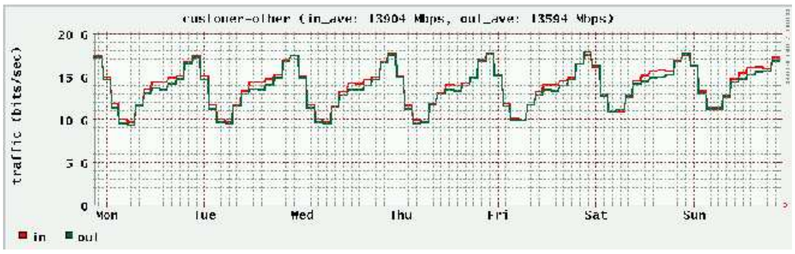
- DSL/CATV/FTTH customer traffic of the 7 ISPs
  - inbound and outbound are almost equal
  - 100Gbps on average!
  - 70Gbps is constant, probably due to p2p applications
  - daily fluctuations: peak from 21:00 to 23:00





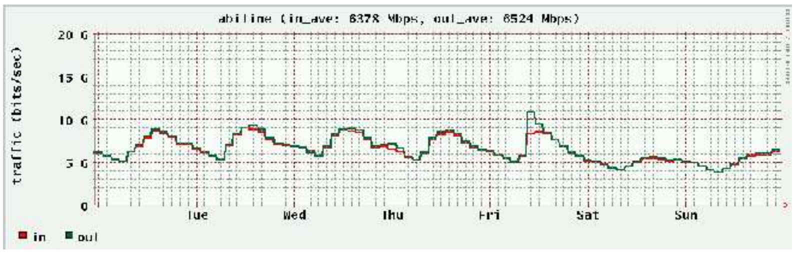

## non-RBB customer weekly traffic
### in September 2004

- leased lines/data center/other customers
  - home user traffic is still dominant (by peak hours)
    - because leased lines include 2nd/3rd level ISPs
  - larger office hour traffic than RBB customer traffic
- only 4 ISPs provided data for this group
  - some ISPs have too many routers, historycally mixed up settings



## ABILENE weekly traffic
### in October 2004

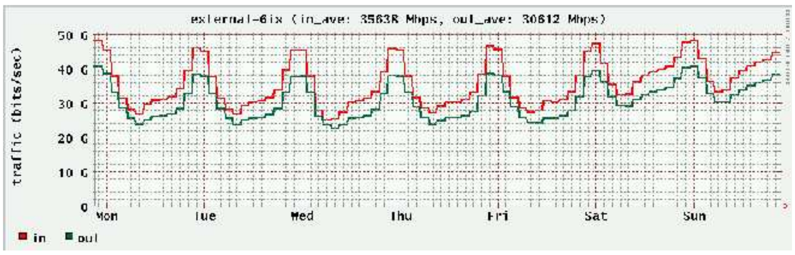
- an example of well-known academic or business usage pattern
  - peak hours around noon
  - weekdays have more traffic than weekend
- our results considerably deviate from traditional usage pattern!





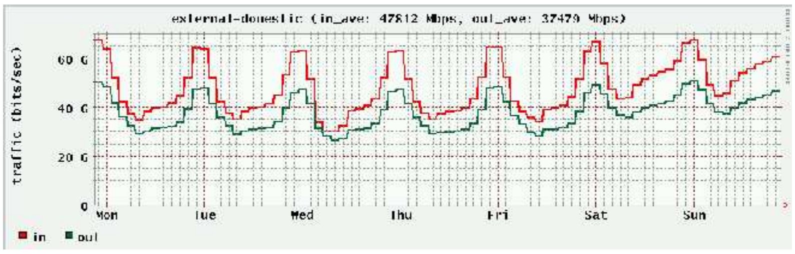

## weekly external traffic to/from 6 major IXes
### in September 2004

- IX traffic is also strongly affected by residential traffic



## weekly other domestic external traffic
### in September 2004

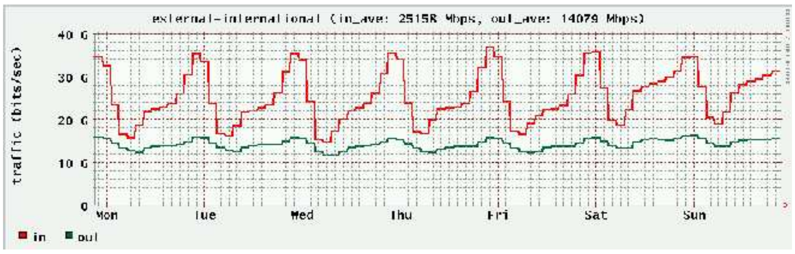
◦ private peering/transit, regional IXes (mainly private peering)
  - traffic volume and pattern are similar to IX traffic





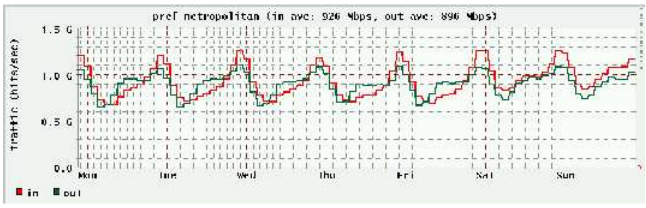

## weekly international external traffic
### in September 2004

◦ international traffic
  - inbound much larger than outbound
    ▫ traditional content downloading seems still dominant

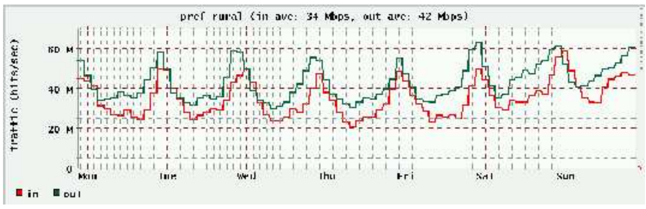


## prefectural traffic

- similar temporal traffic pattern across different prefectures
  - e.g., peak in evening, 70% is constant, regardless the volume
    - one metropolitan prefecture (with larger office hour traffic)



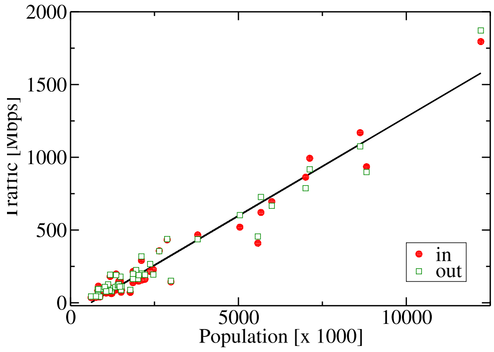

    - one rural prefecture





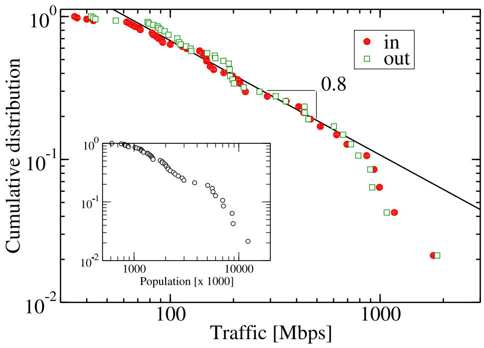

## prefectural population and traffic

- a scatter plot of population and traffic volume
  - traffic is roughly linear to population!
  - similar result with the number of Internet users



## scaling property of prefectural traffic volume

- cumulative distribution of prefectural traffic on a log-log scale
  - power law distribution with a cutoff point at 700Mbps
    - no typical size of prefectural traffic volume!
- sub-plot: cumulative distribution of prefectural population
  - power law is directly derived from population distribution!






## looking at numbers

- customer traffic and external traffic
- monthly average in bits/second
  - September and October data

## average rates of aggregated customer traffic

- only 4 ISPs provided (A2), so when estimated by these 4 ISPs
  - (A1)/(A1 + A2) = 65% for inbound, 67% for outbound

| | (A1)customer-RBB (7 ISPs) | | (A2)customer-non-RBB (4 ISPs) | |
|---|---|---|---|---|
| | inbound | outbound | inbound | outbound |
| Sep | 98.1G | 111.8G | 14.0G | 13.6G |
| Oct | 108.3G | 124.9G | 15.0G | 14.9G |



## average rates of aggregated external traffic

- (B2), mainly private peering, exceeds (B1), major IXes
  - a large amount of traffic is exchanged via private peering
  - IX data may not be a good index of nation-wide traffic volume
    - ratio of (B2) could be overestimated, since private peering is usually only between large ISPs

| | (B1)ext-6ix (7 ISPs) | | (B2)ext-dom (7 ISPs) | | (B3)ext-intl (7 ISPs) | |
|---|---|---|---|---|---|---|
| | in | out | in | out | in | out |
| Sep | 35.9G | 30.9G | 48.2G | 37.8G | 25.3G | 14.1G |
| Oct | 36.3G | 31.8G | 53.1G | 41.6G | 27.7G | 15.4G |

## average rates of total customer and total external

- if we assume all customer traffic is external (no ISP internal traffic)
  - inbound of (A) should be close to outbound of (B)
  - outbound of (A) should be close to inbound of (B)
- ISP internal traffic can be derived from the differences
- but, in our data, (A2) is from only 4ISP

| | (A)customer(A1+A2) | | (B)external(B1+B2+B3) | |
|---|---|---|---|---|
| | inbound | outbound | inbound | outbound |
| Sep | 112.1G | 125.4G | 109.4G | 82.8G |
| Oct | 123.3G | 139.8G | 117.1G | 88.8G |



## IX traffic

- (B1) compared with one obtained directly from the IXes
  - our share is 41% of the total IX traffic
    - if we assume this is the traffic share of the 7 ISPs, the total RBB traffic in Japan is about 250Gbps

| | (B1)ext-6ix outbound | traffic observed by IXes inbound |
|---|---|---|
| Sep | 30.9G | 74.5G |
| Oct | 31.8G | 77.1G |

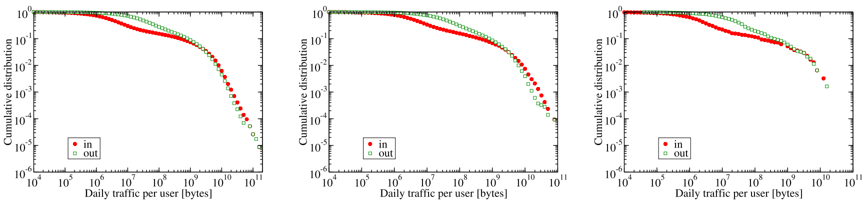
## distribution of per-customer traffic in one ISP

- one of the ISPs provided per-customer traffic info for October 2004
  - by sampled NetFlow and matching customer ID with assigned IP addresses
- we used average daily traffic volume per customer for analysis
- results are consistent with the aggregated traffic



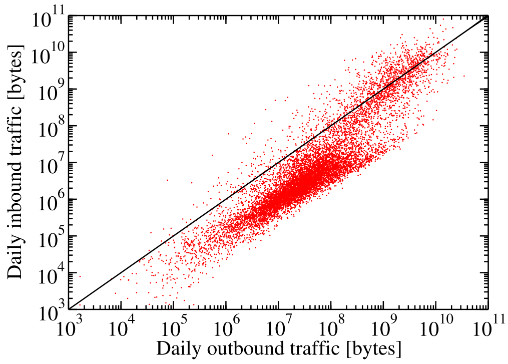
## cumulative distribution of daily traffic per user

- all prefectures (left), metropolitan (middle) and rural prefecture (right)
  - complementary cumulative distribution on a log-log scale
  - distribution similar in all prefectures, differences only in tail length
  - knee point: 4% of customers use more than 2.5GB/day (230kb/s)
  - outbound is dominant for most customers but not for heavy hitters



## correlation of inbound and outbound per customer

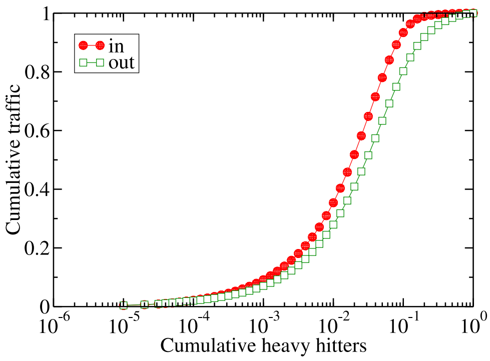
- high density cluster below and parallel to the unity line
  - outbound is 10 times larger than inbound
- in higher volume region, another cluster around the unity line
  - file-sharing over FTTH?






## cumulative distribution of heavy hitters

- cumulative distribution of traffic volume of all of the prefectures
  - with heavy hitters in decreasing order of volume
- top N% of heavy hitters use X% of the total traffic
  - e.g., top 4% uses 75% of total inbound traffic, 60% of the outbound



## discussions

- it is essential for ISPs to prepare for the future to accommodate innovations brought by empowered end-users
- RBB traffic accounts for 2/3 of ISP backbone traffic
  - a significant impact on pricing and cost structures of ISP business
- properties of RBB traffic differ considerably from academic or office traffic often seen in literature
  - research results from academic networks may no longer apply to commercial traffic
- inbound/outbound rates are roughly equal throughout our data sets
  - it affects the design of asymmetric access technologies
- a large amount of traffic is exchanged by private peering
  - IXes data may not be a good index of nation-wide traffic volume
- traffic volume is roughly proportional to regional population
  - it affects the design of capacity planning for the future Internet



## conclusion

- our study to understand residential broadband traffic in Japan
  - cooperation with major ISPs and government
- details on a paper (ACM SIGCOMM CCR special issue)

- future work
  - we will continue collecting aggregated traffic logs from ISPs
  - plans to do more detailed analysis of RBB traffic by sampling


- acknowledgments
  - support and assistance with data collection
    - IIJ, Japan Telecom, K-Opticom, KDDI, NTT Communications, POWEREDCOM, SOFTBANK BB
  - support in coordinating our study
    - Ministry of Internal Affairs and Communications of Japan