DVTS/XCAST with TCP-Friendly Rate Control

Yuji IMAI
FUJITSU LABORATORIES LTD. / WIDE Project
Hiro KISHIMOTO
FUJISTSU LABORATORIES LTD.

Following 3 presentations are collaboration works of the XCAST WG / WIDE Project.

I. DVTS/XCAST with TCP-Friendly Rate Control (Yuji@Fujitsu)

II. Design and Implementation of Sender Initiated Congestion Control (Kazunobu@Panasonic)

III. XCAST on PlanetLab

~ Deploying Overlay Network on PlanetLab ~

(Nobuo@Nagoya Univ.)

Intent of our presentations

- Publication and discussion of individual works in progress.
  1. Clarifying the relation with "**Scalable Adaptive Multicast**" activity and share problem space with researchers of this area.
  2. Call for research collaborations.
     - XCAST on VINI/PlanetLab: The test-bed
     - IRTF SAM Research Group: The community

John & Jeremy will make it in a panel this evening.

Items we will discuss.

1. What's SAM: Scalable Adaptive Multicast. (Yuji)
   - Problem statements
   - Collaborative research activities in IRTF
2. XCAST with TFRC (Yuji & Kazunobu)

   One approach of SAM that realizes P2P flow-control
   - XCAST: for a piece of components of SAM
   - TFRC with XCAST
     - DVTS/XCAST6 with TFRC
     - SICC on XCAST6
3. Demo of the test-bed for novel protocols, include P2P multicast and SAM. (Nobuo)

   One effort to deploy Overlay Network on VINI/PlanetLab.

1. What's SAM: Scalable Adaptive Multicast. (Yuji)
   - Problem statements
   - Collaborative research activities in IRTF
2. XCAST with TFRC obu)

   One approach of SAM tha

   - XCAST: for a piece of c
   - TFRC with XCAST
     - DVTS/XCAST6 with T
     - SICC on XCAST6
3. Demo of the test-bed for novel protocols, include P2P multicast and SAM. (Nobuo)

   One effort to deploy Overlay Network on VINI/PlanetLab.

Just for explanation of remained part, we just mentioned briefly.

In detail, those will be explained by the co-chairs of IRTF SAM RG in a later panel.

SAM: Scalable Adaptive Multicast

i. One of directions that “Multicast” should move forward: **Scalable** especially with respect to number of groups and **adaptive** for dynamics of group membership, network topology and network resource constraints.

ii. Research group of IRTF

- Internet Research Task Force SAM Research Group
  - To explore and research techniques which improve scalability and adaptiveness.
  - Investigate approaches based on ALM(Application Layer Multicast), OM (Overlay Multicast), and native IP multicast, as well as hybrid approaches.
  - http://www.samrg.org
  - http://www.irtf.org/charter?gtype=rg&group=samrg

SAM: Approach.

- Characterizing the problem space
- Comparisons and analysis of existing approaches
- Deployment scenarios which are independent of but can support and evolve with network infrastructure support for native multicast.
- A SAM framework that supports multiple ALM/OM/ native/hybrid protocols
- Etc…

draft-irtf-sam-problem-statement-01.txt

Application Layer Multicast
Overlay Multicast
End System Multicast
P2P Multicast
XCAST: eXplicit Multi-Unicast

SAM on VINI/PlanetLab
test-bed for novel P2P multicasts and SAMs

draft-muramoto-irtf-sam-generic-require-01.txt

SAM: Problem statement & Requirements

- **Problem Statement** (draft-irtf-sam-problem-statement-01.txt)
  - Heterogeneous Multicast Infrastructure
    - Varying Infrastructure by Network Region
    - Regional Transitions
  - Quality of Service
    - Native QOS, No Native Multicast
    - Other Combinations
  - Mobility
    - Multicast Service Selection
    - Transitions between ALM and Native Multicast

A requirement we focused on in our presentations, today.

- **Requirements** (draft-muramoto-irtf-sam-generic-require-01.txt)
  - Multicast capability
  - Minimizing the infrastructure support and fastening the service-connectivity cycle
  - Scalability
    - Routing convergence
    - Dynamic topology
    - Number of group
    - Dynamics of group membership
  - Adaptivity
    - Latency (Delay sensitivity)
    - Redundancy path
    - Data rate control & Congestion avoidance

1. What's SAM: Scalable Adaptive Multicast. (Yuji)
   - Problem statements
   - Collaborative research activities in IRTF
2. XCAST with TFRC (Yuji & Kazunobu)

   One approach of SAM that realizes P2P flow-control
   - XCAST: for a piece of components of SAM
   - TFRC with XCAST
     - DVTS/XCAST6 with TFRC
     - SICC on XCAST6
3. Demo of the test-bed for novel protocols, include P2P multicast and SAM. (Nobuo)

   One effort to deploy Overlay Network on VINI/PlanetLab.

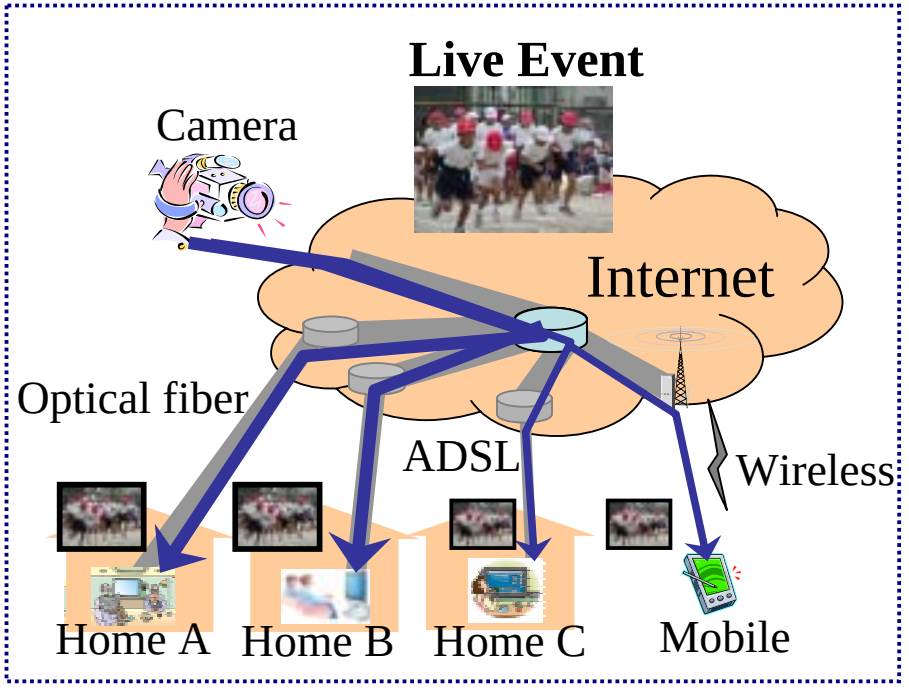
Principle of following XCAST works.

- Go pragmatic.
  - Don’t innovate quite new but find good combination of **“off-the-shelf”** technologies.
    - Unicast routing, TCP flow control, RTP,,,
    - It’s OK looking **boring to die** as academic research because of whole trivial mechanisms.
  - Don’t solve all but firstly focused on the sub-problem area of SAM that we really need.
    - Afterward, think combination or hybrid with other **“off-the-shelves”.**

One Target of SAM

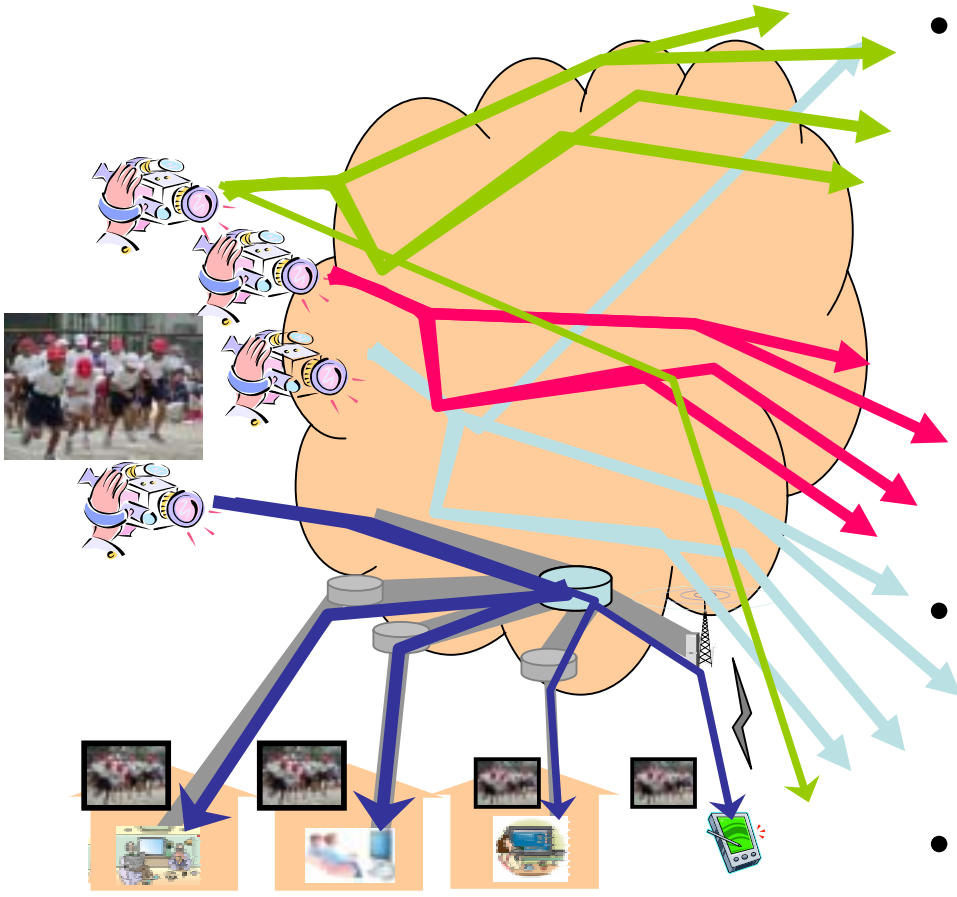
P2P Real-time Group Communication over the Internet

e.g. Personal Broadcasting



- In contrast with professional broadcasting, P2P group communication MUST be
  - **scalable** w.r.t. number of group (= number of stream devices transmitting on the Internet.)
  - **adaptive** for dynamics of network environments.
    - Location of receivers
    - Performance of receiving devices
    - Network condition both on last 1 mile access media and intermediate ISP clouds.

Analysis of our Target Situation of SAM



- Number of group/tree will be enormous.
  - Same as number of transmitting devices.

It means ISPs have to maintain very large multicast routing table if they do it with native IP multicast.

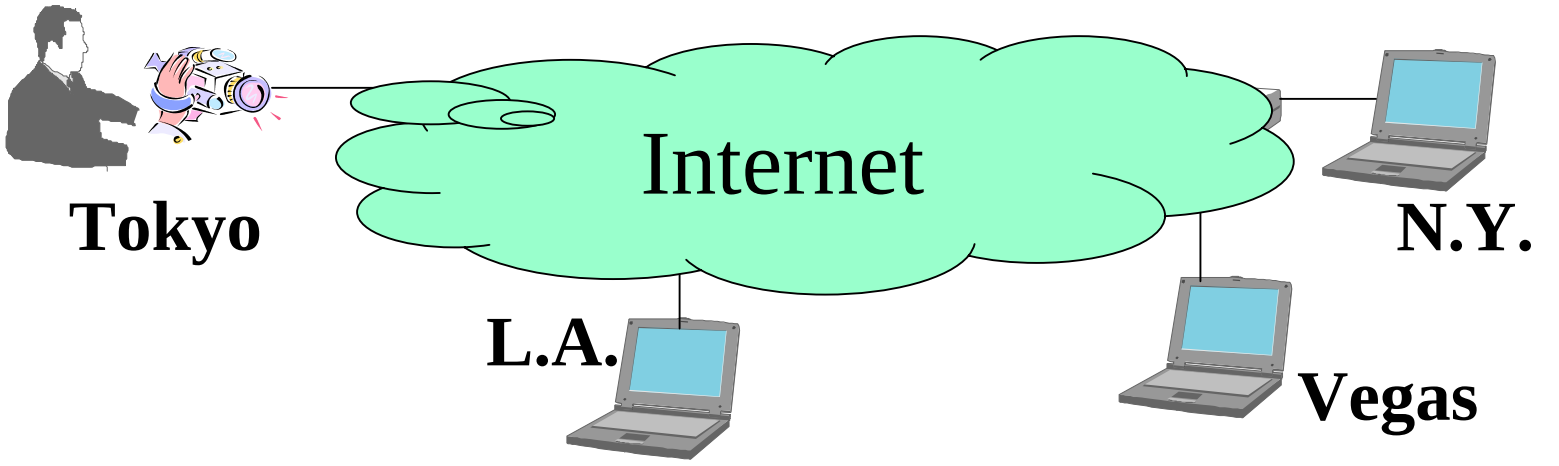
- In major case, number of receivers are relatively small. (upto 16 ~ 32)
- Transmitters can specify all their receivers.

XCAST: eXplicit Multi-Unicast

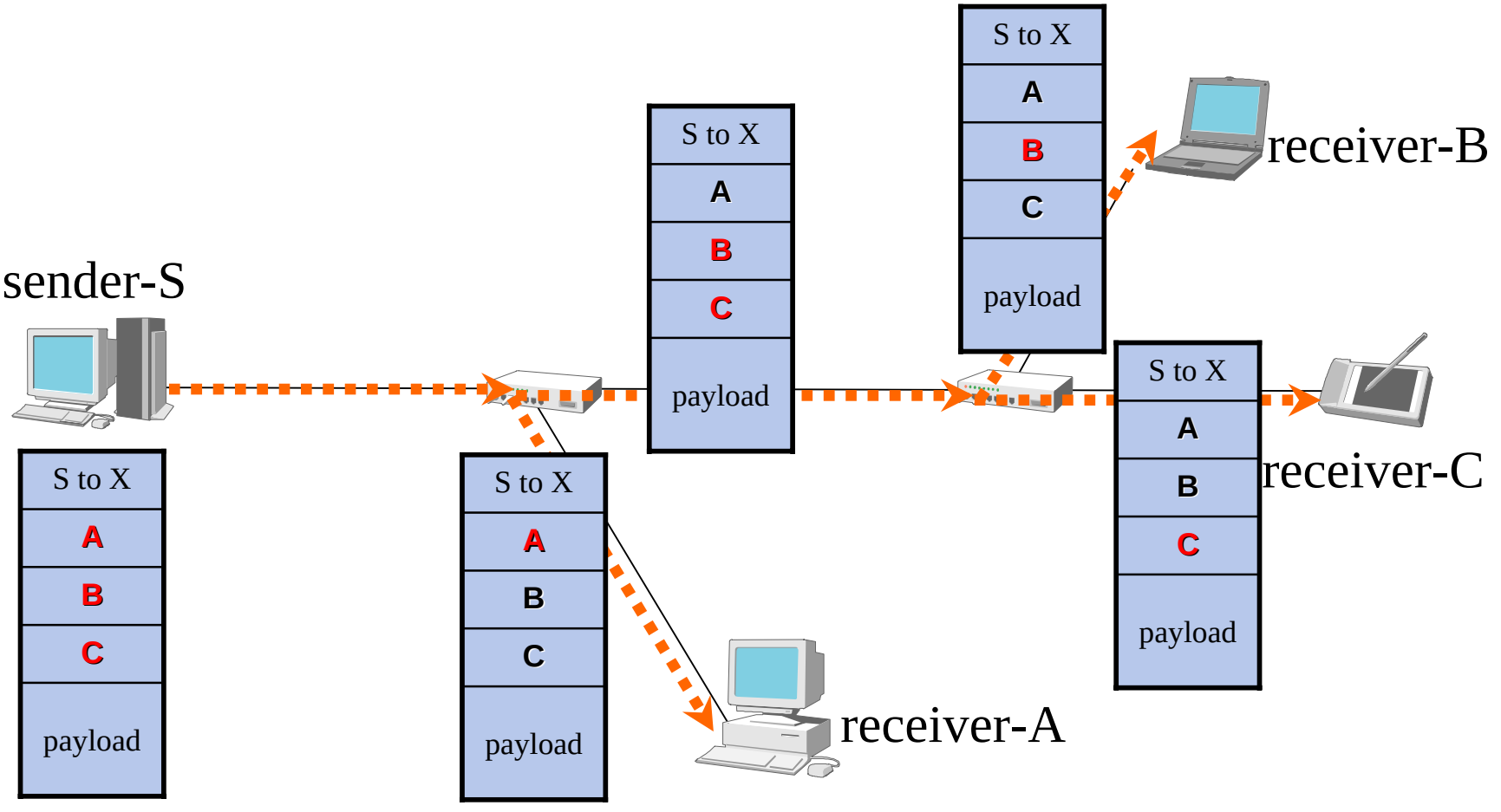
-a piece of candidate components for SAM-

Instead of a group address, use an explicit list of unicast addresses.

IP header SRC=Tokyo DST=XCAST.	XCAST header [**Tokyo, Vegas, N.Y**] [1 , 1 , 0]	UDP Header	Payload



XCAST
- multi-destination routing -



Characteristics of XCAST

- Fit for personal small group
  - Let transmitters keep and write addresses by oneself, explicitly.
  - Very scalable w.r.t. number of groups.
    - Intermediate router don't have to maintain any routing status other than unicast one.
- Trivially simple & stupid
  - Smart edge & stupid network.
- Not do well for groups that have very large receiver
  - Combination scheme is necessary for application area that need seamless scalability w.r.t. number of receivers.

Other features of XCAST

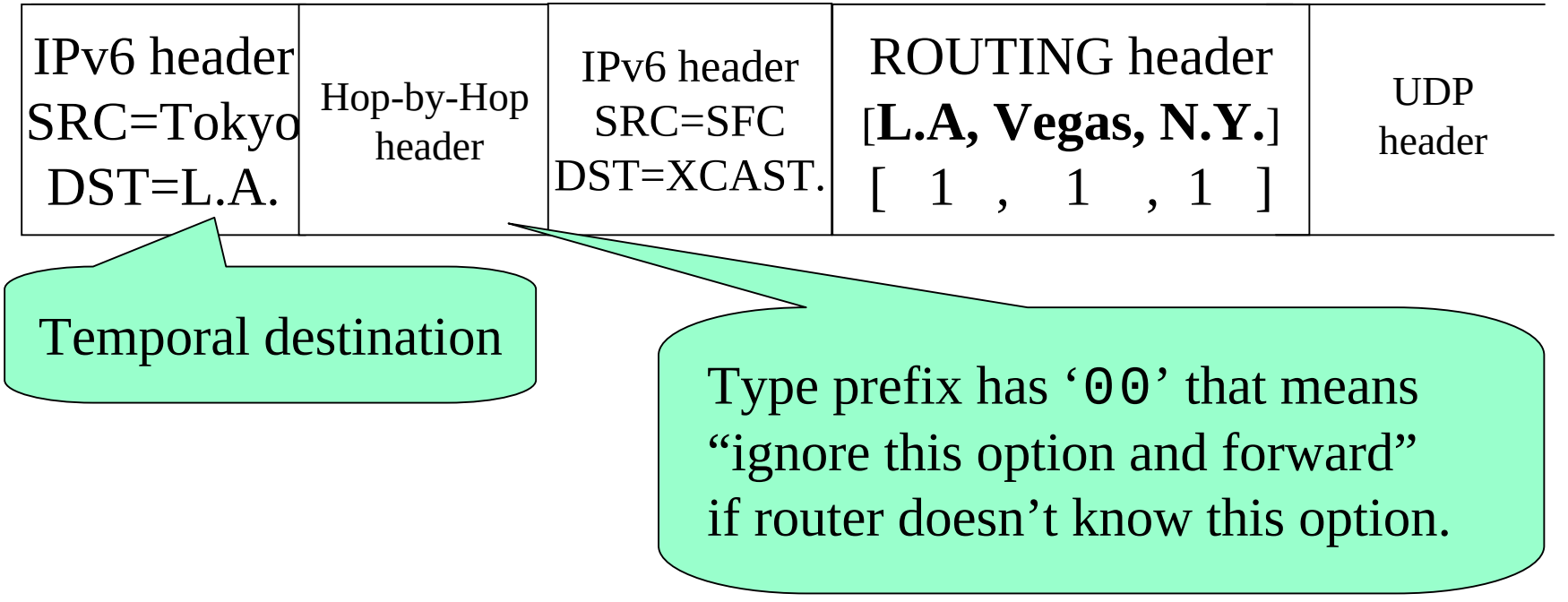
- not mentioned in this presentation -

- Semi-permeable capsulation
  - Method to pass the XCAST datagram through non-XCAST ready network cloud.
  - To ignite **"service-connectivity cycle"** by end-to-end style.
- Combination with signaling & group management systems
  - SIP, XMPP, REST,,,
- Running code
  - For FreeBSD, NetBSD, Linux and Win XP
    - http://xcast6.sourceforge.net
    - Keep stable for 5 years.
- Community
  - IETF efforts
    - "Basic spec" is on the editorial queue for RFC. (Independent submission)
  - OSS development / release
  - Continuous operation of test-bed network.

"Service-connectivity cycle" is a concept represented in the keynote of NOSSDAV '03, "Why Johnny can't multicast", by Mostafa H. Ammar to break chicken-egg situation of Multicast.

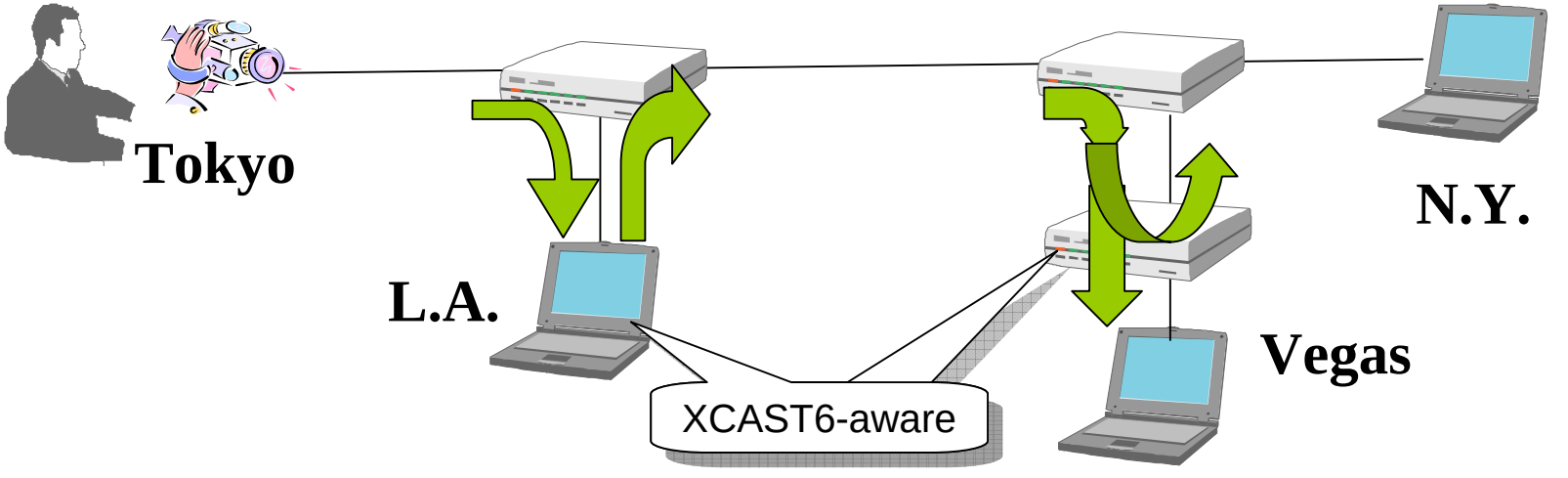
Semi-permeable capsulation

– Passing through the intermediate routers which does not support XCAST6 datagram.



Semi-permeable capsulation

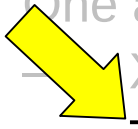
Even if non-XCAST6 routers are on the way, XCAST6 datagram pass through them once and turn back at the next XCAST6 nodes or routers.



i. End node can transmit XCAST6 in any environment and ignite “service-connectivity cycle”.
ii. Delivery path become optimized gradually by installing more XCAST6 routers, afterward.

1. What's SAM: Scalable Adaptive Multicast. (Yuji)
   - Problem statements
   - Collaborative research activities in IRTF
2. XCAST with TFRC (Yuji & Kazunobu)

One approach of SAM that realizes P2P flow-control



   - XCAS: for a piece of components of SAM
   - TFRC with XCAST
     - DVTS/XCAST6 with TFRC
     - SICC on XCAST6
3. Demo of the test-bed for novel protocols, include P2P multicast and SAM. (Nobuo)

One effort to deploy Overlay Network on VINI/PlanetLab.

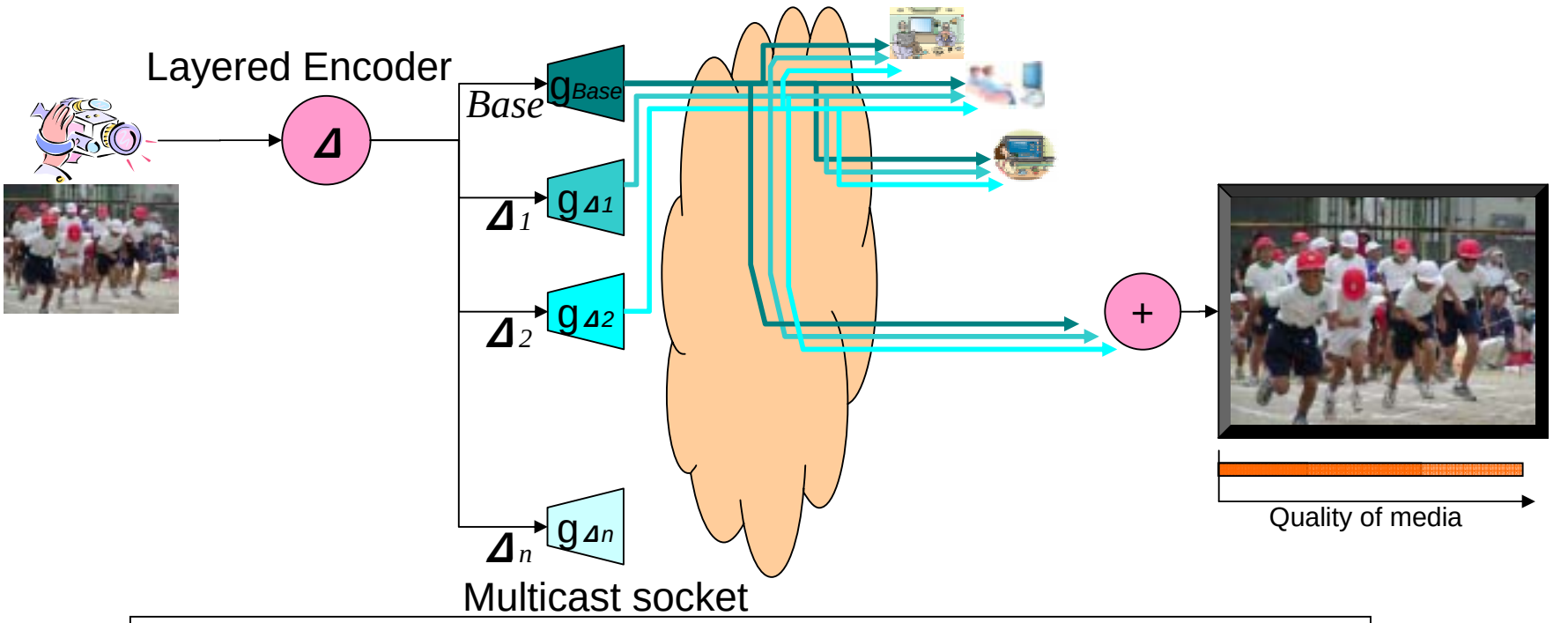
SAM Requirements on adaptiveness for network condition

- Fluctuating factors
  - Topology
  - Routing
    - policy, preference
  - Failure
    - media, SW, router
  - Group creation
  - Membership
  - Amount of flow shared with others
- Required mechanism
  - Reroute the path / tree
    - Rapidness
    - Stability
  - Flow control
    - Avoid congestion
    - Fairness
    - Quick & smooth convergence
    - Fine-adjustment on individual stems of tree for receivers.

Related work: Layered Multicast

Send layered streams of signal by several multicast trees.

S. McCanne, "Receiver-driven layered multicast", in Proc. ACM Sigcomm 1996.



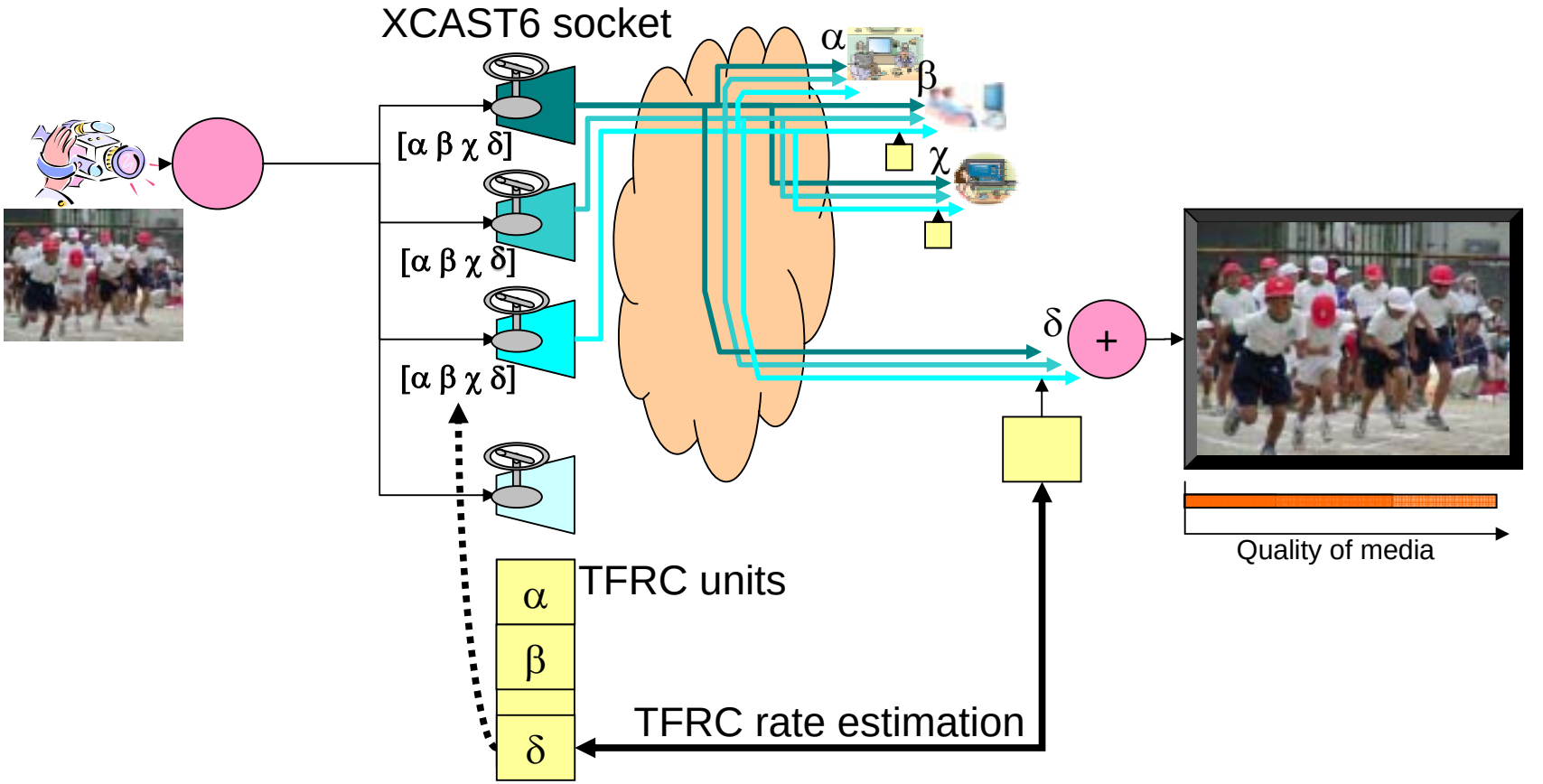
Many papers had issued on improving “Layered Multicast”, but we couldn’t find any “off-the-shelf” variant, yet.

Major improvement metrics among variants of “Layered Multicast”

- Scalability
  - w.r.t. Number of receivers
- Fairness
  - TCP-friendliness
- Stability of rate adjustment
- Quickness of rate adjustment
- Granularity of improvement
  - Progressiveness of quality of stream
- Independence from condition of other receivers

XCAST6 with TFRC

- Apply TCP-Friendly Rate Control simply on individual stem of XCAST tree.



Related works

- End System Multicast
  (@CMU.edu)
  - Target: one-to-many usage
  - TFRC applied on the each link between end system.
    - All intermediate end systems MUST relay datagram adjusting the rate for down stream.
    - Seems difficult to make feed back for VBR encoder of original transmitter.

"A Case for End System Multicast", Yang-hua Chu, Sanjay G. Rao, and Hui Zhang, *Proceedings of ACM SIGMETRICS*, Santa Clara, CA, June 2000.

- Xpand
  (@cs.huji.ac.il)
  - Target: Many-to-many
  - TFRC estimation on between all original transmitting end to last receiving end.
  - Single rate flow by single layer multicast.
    - Rate is limited by most lowest stem of tree.

"Tcp-friendly many-to-many end-to-end congestion control“, I. Shnayderman T. Anker, D. Dolev and I. Sukhov, in Proc. 22st IEEE Symposium on Reliable Distributed Systems. October 2003, IEEE CS

Characteristics of XCAST with TFRC

- For small group end-to-end, not for one-to-many.
- TFRC applied on all end-to-end stems.
  - Transmitter individually estimate the rate for all receiver by TFRC.
  - Enable to control the delivery of packets explicitly.
    - For each destination independently.
    - For each transmission of the packet.
  - Enable to make feed back for VBR encoder in original transmitter.

1. What's SAM: Scalable Adaptive Multicast. (Yuji)
   - Problem statements
   - Collaborative research activities in IRTF
2. XCAST with TFRC (Yuji & Kazunobu)

   One approach of SAM that realizes P2P flow-control
   - XCAST: for a piece of components of SAM
   - TFRC with XCAST
     - DVTS/XCAST6 with TFRC
     - SICC on XCAST6
3. Demo of the test-bed for novel protocols, include P2P multicast and SAM. (Nobuo)

   One effort to deploy Overlay Network on VINI/PlanetLab.

DVTS
Digital Video Transport System

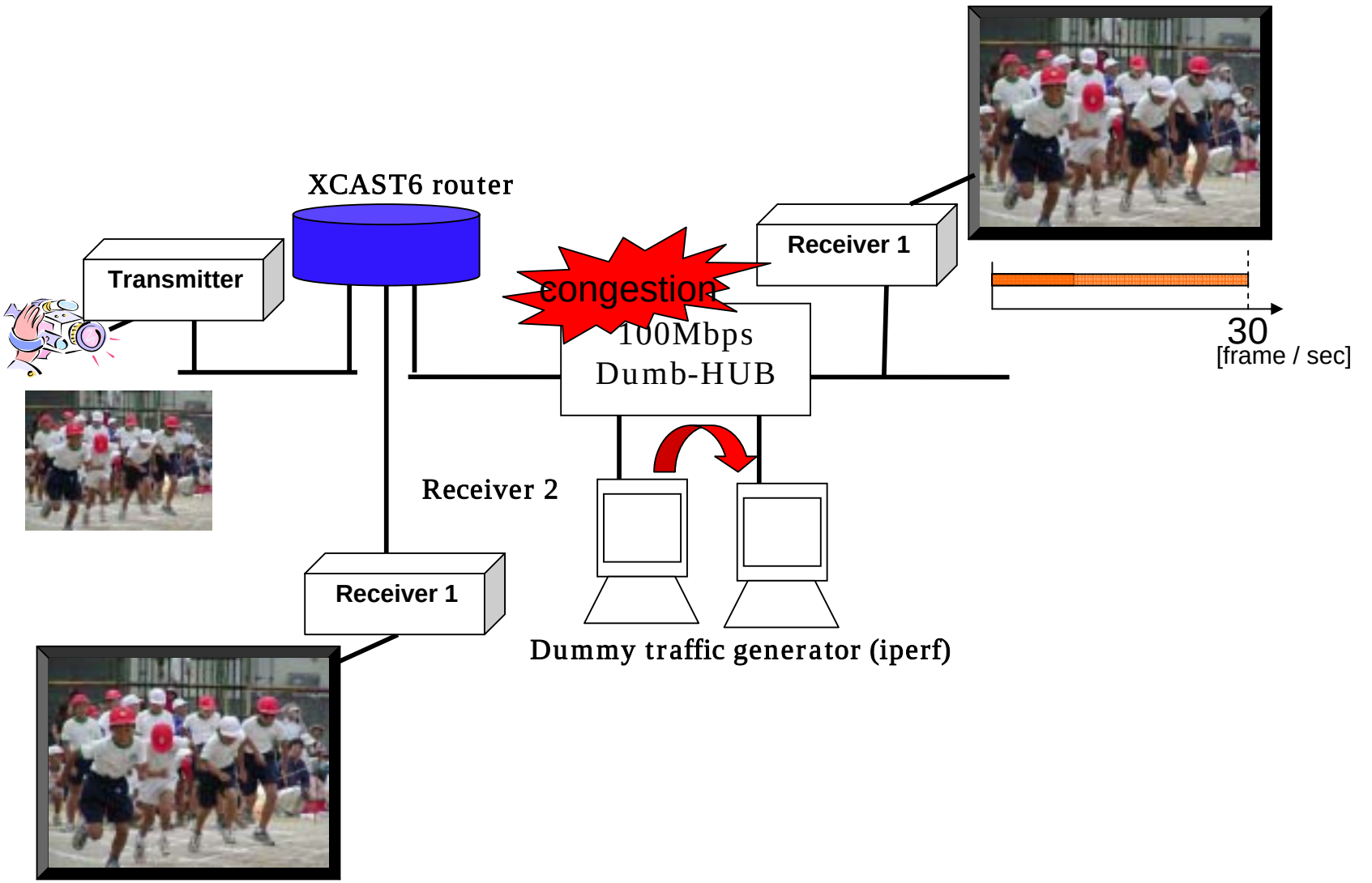
- DV (IEEE1394) over IP
  - RTP Payload Format for DV(IEC 61834) Video [RFC3189]
  - RTP Payload Format for 12-bit DAT Audio and 20- and 24-bit Linear Sampled Audio [RFC3190]
  - Enable to decrease bandwidth by interleaving video frame. (33 ~ 3 Mbps)
  - Both for unicast and multicast, IPv4 and IPv6.
- TFRC applyed on DVTS in unicast case.
  - A. Ogawa, “Implementing TCP-Friendliness in Digital Video over IP”, Feb 2002.

DVTS/TFRC with XCAST6

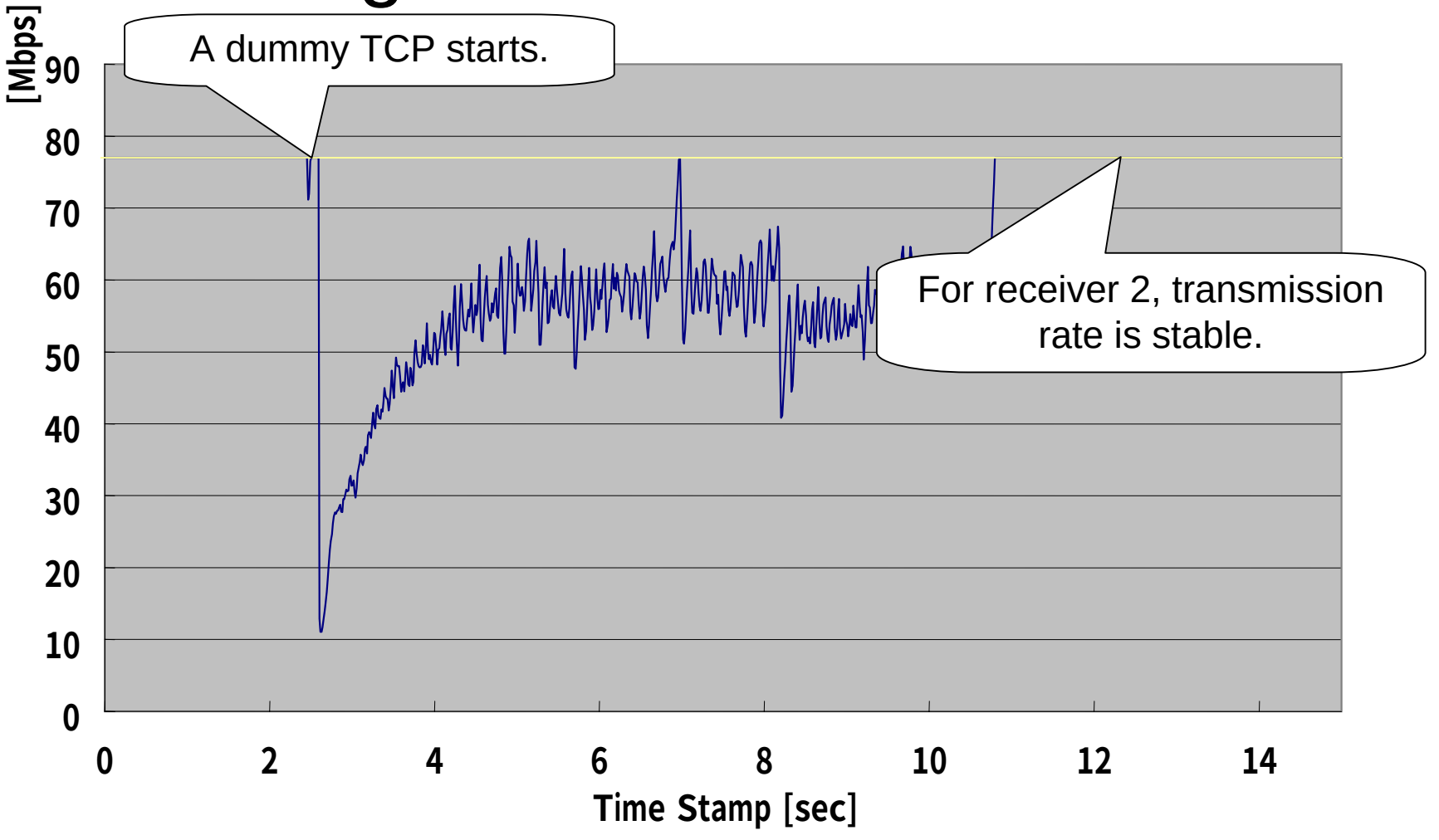
- A conceptual prototype of XCAST with TFRC -

- Code base
  - dvts-0.9b08 of WIDE
  - TFRC sample code of ICIR.org
- RTP/RTCP extension
  - Append a list of RCI (Rate Control Information)
    - RTP/RTCP parameters used by TFRC

Test environment



Transmission Rate
-Congestion with TCP flow-



Result of this part

- TFRC seems to work with XCAST.
  - Confirmed point
    - TFRC transmission adjustment by original TFRC logic.
    - Rate is independently variable from stem to other.
  - Look like but need to confirm more in detail.
    - Keeping TCP-friendliness.
      - Congestion avoided.
      - Fairness
  - Not confirmed at this point
    - Working firmly in real world.

Done in the following Panasonic work

Purpose of XCAST6/PlanetLab