Road space rationing on the Internet: Pay more to get better throughput in best-effort networks

Katsushi Kobayashi
ikob@riken.jp
RIKEN AICS, Japan

AICS

K computer
RIKEN Advanced Institute for Computational Science

Internet collapse by BW growth ?

- Exponential traffic growth.
  - Growth not only of # of user, but of per-user traffic.
- ISP $$$ does not increase enough for BW growth.
  - # of user will not increase more in higher penetration.
  - ISP subscription $$ remains lower.
    - Additional service, or BW upgrade required to raise $$.
  - Small portion of users occupy almost of all traffic.
    - Top 4% consumes 2/3 of total BW by 2.5GB/day, while typical by 0.1GB/day in Japan [6].
    - Difficult to raise $$ with remaining imbalance.

Better pricing

- Traffic meter of access appropriate than "flat-rate" ?
- Pricing should reflects how much share of "bottleneck" of congestion and/or of investment.
  - Where is "bottleneck" ? At our side or his/her side ?
  - "bottleneck" stays at access even deployed FTTH ?
- flat-rate is preferred by user due to predictable expense.
  - Does ISP prefers meter ? Total network cost in short period does NOT depend on traffic volume.

Service differentiation

- Service differentiation maximizes also user satisfactions, not only ISP revenues.
- Differentiation mechanism should
  - be effective through Internet.
  - be able to apply data traffic as Web service consuming much traffic.
  - supports large # of levels to cover broader type of users from residential to large company.

Road space rationing in cities.

- Reduce peak travel demand in urban with small cost in Athens, Mexico City, and Sao Paulo.
- Restricting access based on last digit of license No.
  - e.g., odd and even No. can access in alternate days.
  - 20% reduction expected.
- Can avoid to buy two or more vehicles :
  - realize “pay more for get better” that is expected in Internet QoS.

Road space rationing in Internet

- RSR: allocates partitioned supply to partitioned demand
- Time slot partitioning used in City is unfeasible.
  - Alternative day access is bad experience.
  - Only 2 classes is not enough for 100+ traffic difference.
  - Synchronization mechanism, if granular time slot.
- Network partitioning other than time slot.

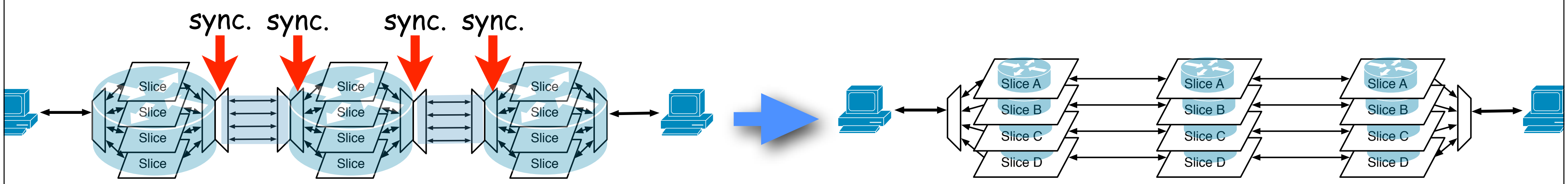
	Demand	Supply
City	Vehicles	Road space
Internet	Packet rate	Bandwidth

Multi-slice network architecture

- Partitioning into multiple virtual network slice/overlay.
  - Mux-DeMux on Host/Subscriber edge.
- Benefit from more than one slice/overlay network.
  - DynaBone : DDoS resistance by detaching attacked slices [19].
  - FAIN (Flexible Arrays of Inexpensive Nodes) : Scalable BW by eliminating "synchronization" point in paralleled router and link [11].

FAIN:e2e virtual slice (1)

- Goal : network architecture for “parallel-in-global”
  - Eliminate synchronization point from router and link.
- FAIN : network sliced into dedicated large # of virtual slices.
  - Physical router and Link also sliced into virtual router and link.
  - Packet cannot across multiple slices.
  - All slices have same network topology.
- Virtual slices are allocated to a parallel component of physical device.
- Prevent reorder within a virtual slice for transport.

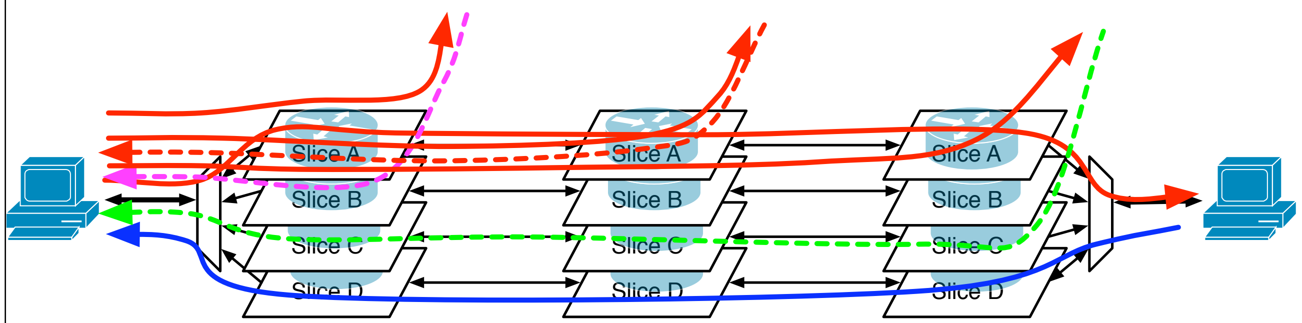


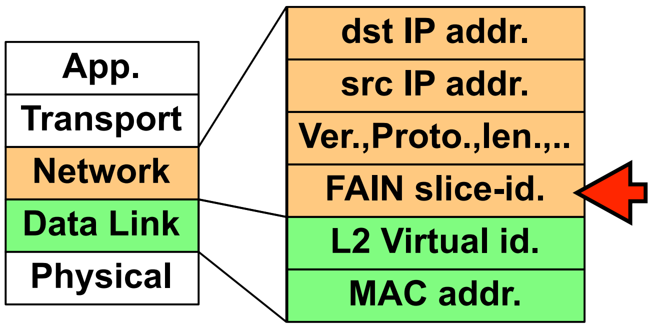
FAIN:e2e virtual slice (2)

- Slice ID at L3
  - FAIN slice-id: assigned/checked by User-ISP edge
  - IPv6 20-bits flow-label is candidate

Note: slice-id can be mapped to lower-layer virtualization ID. as 802.1q tagged VLAN.

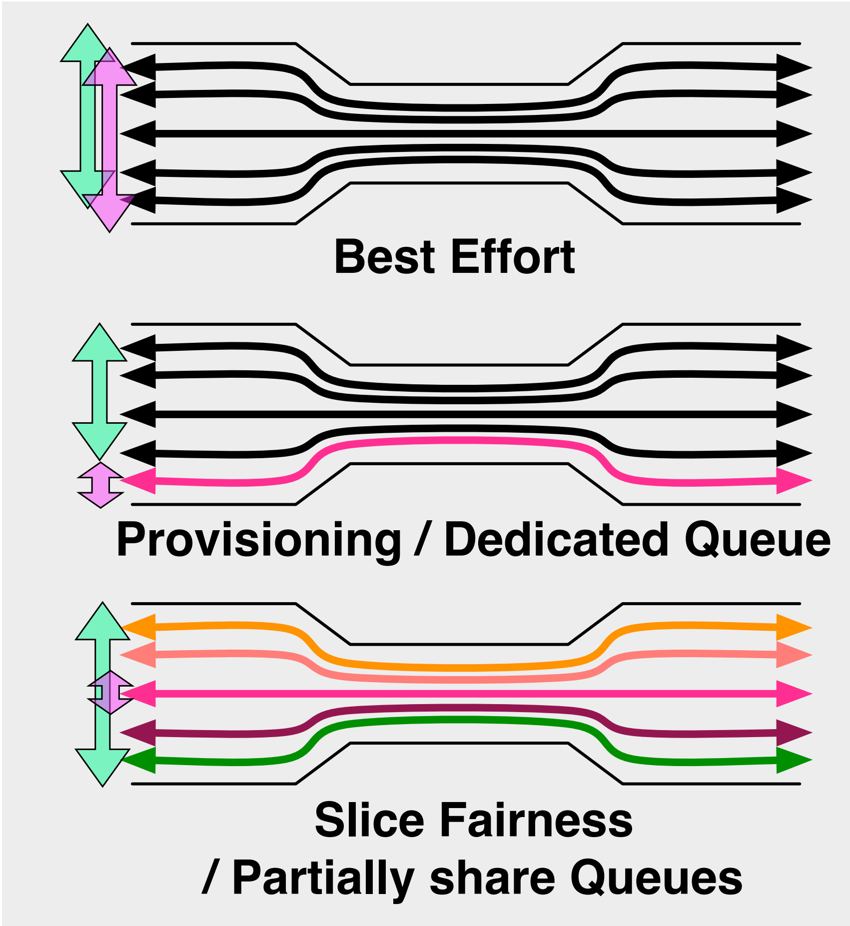
- How edge (Mux-Demux) work:
  - Send: subscriber edge assigns or checks virtual slice in packets according subscriber-ISP contract.
  - Receive: Accept all packets from any virtual slice.
    - Not require end-to-end slice negotiation.





Fairness among slices

- Every slice has same priority through Internet.
  - Service differentiation made with # of accessible slices.
  - "Pay more, get more slices"
    - Up to 1M classes by 20bits.
    - Any traffic, incl. best effort is possible.



Can fix traffic imbalance ?

- Only to achieve instantaneous rate differentiation. Imbalance dispute is based on amount of data. How amount of data ?
- Differentiation is preserved in amount of data too, as:
  - Regular user : $f(t)$ rate capacity of each slice.
  - Premier own N slices : $g(t) = N \times f(t)$
  - always : $$\int_{t_0}^{t_0+T} g(t)\,dt = N \times \int_{t_0}^{t_0+T} f(t)\,dt$$
- Note: $f(t), g(t)$ represent CAPACITY, not real traffic rate. N times differentiation is not true, when not sharing same "bottleneck" However, most users will prefer higher CAPACITY even with best-effort as to choose higher access BW menu.

RSR for Internet

- Router: per-slice fairness
  - Assign slice(s) to each router queue.
  - Insufficient # of queues on exiting router, as 8 - 10,000.
    - Aggregate 1+ slices, e.g., only refer upper bits.
- ISP-user edge: applying contract
  - Filter or rewrite ingress packets by slice ID.
- End system:
  - Regular = 1 slice: unnecessary to change.
  - Premium = 1+ slice: some efforts, and much design space.

How ends handle 1+ slice

- Assign a destination to one of accessible slices.
  - For large Web service. FaceBook, Google, ... etc.
  - Differentiation made of managing # of users in each room.
- A connection uses 1+ slices:
  - Loss resilient : FEC with different dedicated slice.
  - Higher throughput : 1+ connections as GridFTP, or split 1 connection data.
- User who pays premium service can take into account pros. and cons.

Conclusion

- Pricing and differentiation.
  - Reflect how much share of "bottleneck"
  - "Pay more, get better" through Internet
- RSR for traffic demand control
- RSR with multiple network slice instead of TDM

- Just show my idea, and possibly include significant flaws