Application-Network Integration: Possibilities, Challenges, and Possible Next Steps

Y. Richard Yang

IETF 106
November 20, 2019
Singapore

Application-Aware (AA) Networks Can Have Diverse AA Capabilities, Requiring Different Support

Example Capability	Possible Support & Assumption
Treat each packet the same (aka not AA)	
Aware at app-level granularity	
Create different networks/slices (e.g., voice vs data networks)	IP, access; SDN; scheduling
Identify packets by ports (e.g., ACL)	Packet header port; scheduling
Aware of sub-app granularity	
Scheduling each packet according to app-level deadline (e.g., fastpass'14)	Custom packet header; scheduling
Distinguish application-level structures (e.g., I frame vs P frame)	
Fancy: Co-flow scheduling (e.g., VARYS'14, AALO'15)	Network state; scheduling
Aware of cross-app/protocol dependency	Packet header; net state; scheduling
Fancy: identify full dependency (e.g., applicationlevel dependency such as DNS->handshake->…)	Network state; scheduling
…	

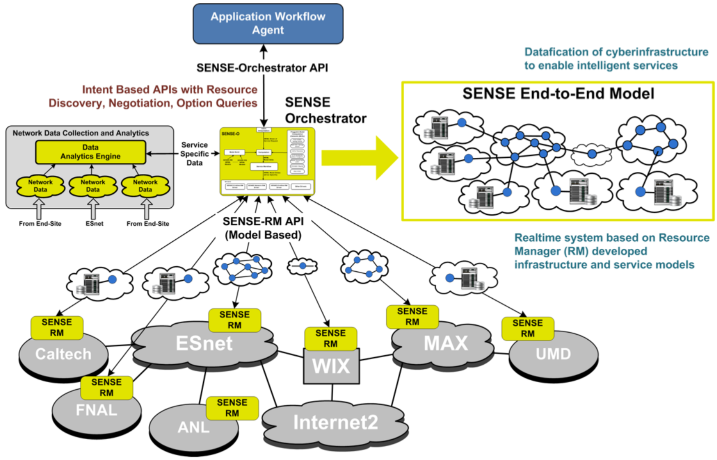
Network-Aware (NA) Applications Can Have Diverse NA Capabilities, Requiring Different Network Information/Support

Example Capability	Support & Assumption
Transfer time selection	Network state in time; can delay
Server direction	Path properties from client to potential servers; has multiple servers
Rate adaptation	None/None/ECN/INT
CC, reacting to loss/delay/ECN bit/INT (e.g., HPCC'19)	
Adaptive streaming	
Lower-than-best-effort (e.g., LEDBAT)	
Multi-path TCP	
…	

Despite Broad Capabilities, A Simple Gap Example

- Collaborative, distributed, exa-scale data sciences [1]
  - Applications: LHC, LIGO, LSST, EHT …
  - Services
    - Time-Block-Maximum Bandwidth
      - Application asks for a specific time block and would like to know (or provision) the maximum bandwidth available for a specific time period.
    - Bandwidth-Sliding-Window
      - Application asks for a specific bandwidth and duration and provides an acceptable time window. For example, a request may be for 40 Gbps for a 10-hour time window, sometime in the next 3 days.
    - Time-Bandwidth-Product (TBP)
      - Application asks for “8 hours of transfer at 10Gbps” representing a TBP of 36 TBytes. The user also specifies an acceptable time window, and other options such as “prefer the highest bandwidth rate available”, or the lowest.
  - Protocol outcomes
    - Immediate provision
    - What is Possible?
    - Negotiation



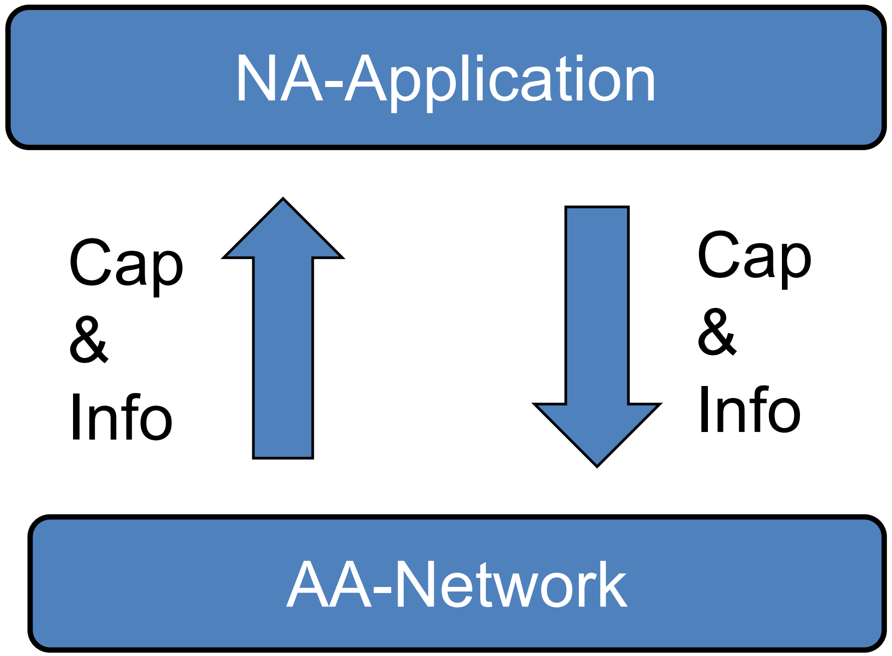


[1] http://sense.es.net/services

Basic Challenge: Architecture

- Applications and networks can be designed with different objectives
  - Application: optimizes application's utility
  - Network: optimizes network's utility, enforces fairness, …

- The end-to-end principle which mostly argues for the minimization of AA-networking



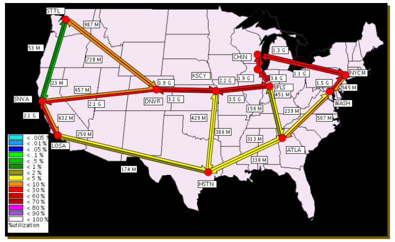
A Simple Example

- Application objective: optimize total throughput
  - Using a fluid model*, we can derive that: optimizing throughput ⇒ maximizing up/down link capacity usage

- Network objective: minimize maximum link utilization (MLU)
  - b_e: background traffic volume on link e
  - c_e: capacity of link e
  - $I_e(i,j)$ = 1 if link e is on the route from i to j
  - t^k : a traffic demand matrix $\{t^k_{ij}\}$ for each pair of nodes (i,j)

$$\max \sum_i \sum_{j\neq i} t_{ij}$$
$$s.t. \forall i, \sum_{j\neq i} t_{ij} \leq u_i \ ,$$
$$\forall i, \sum_{j\neq i} t_{ji} \leq d_i \ ,$$
$$\forall i \neq j, t_{ij} \geq 0$$

$$\min \max_{e\in E} \left(b_e + \sum_k \sum_{i\neq j} t^k_{ij} I_e(i,j)\right) / c_e$$

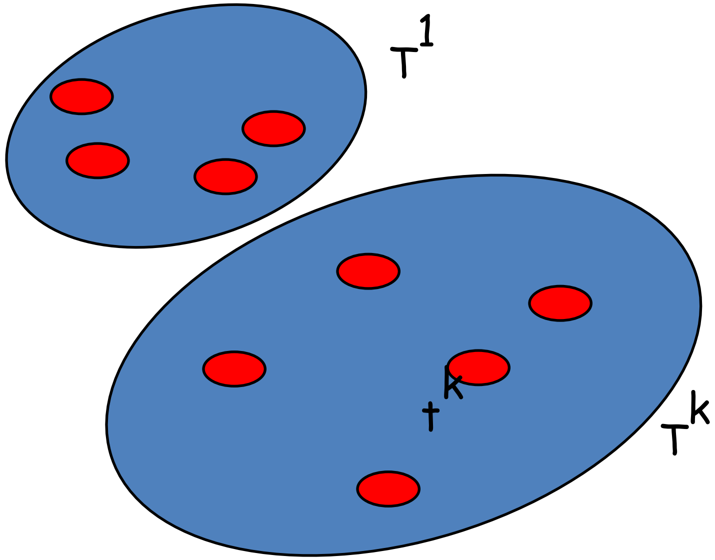


A Simple Example: System Formulation

- Combine the objectives of network and applications

$$\min_{e\in E} \max (b_e + \sum_k \sum_{i\neq j} t_{ij}^k I_e(i,j)) / c_e$$

s.t., for any k,



$$\max \sum_i \sum_{j\neq i} t_{ij}^k$$

$$s.t. \forall i, \sum_{j\neq i} t_{ij}^k \leq u_i^k,$$

$$\forall i, \sum_{j\neq i} t_{ji}^k \leq d_i^k,$$

$$\forall i \neq j, t_{ij}^k \geq 0$$

Why Hard: Constraints Couple Net/App

$$\min_{\forall k: t^k \in T^k} \max_{e \in E} \Big(b_e + \sum_k \sum_{i \neq j} t_{ij}^k I_e(i,j)\Big) / c_e$$

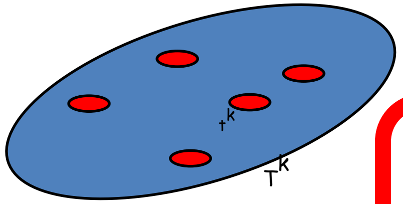
$$\min_{\forall k: t^k \in T^k} \alpha$$

$$s.t. \quad \forall e: b_e + \sum_k \sum_{i \neq j} t_{ij}^k I_e(i,j) \le \alpha c_e$$

Constraints couple network/applications together!

Solution Architecture: Decouple Network and Application

$$\min_{\forall k: t^k \in T^k} \max_{e \in E} \Big(b_e + \sum_k \sum_{i \neq j} t_{ij}^k I_e(i,j)\Big) / c_e$$



Introduce p_e to decouple the constraints

$$\min_{\forall k: t^k \in T^k} \alpha$$

$$s.t. \quad \forall e: b_e + \sum_k \sum_{i \neq j} t_{ij}^k I_e(i,j) \le \alpha c_e \qquad p_e$$

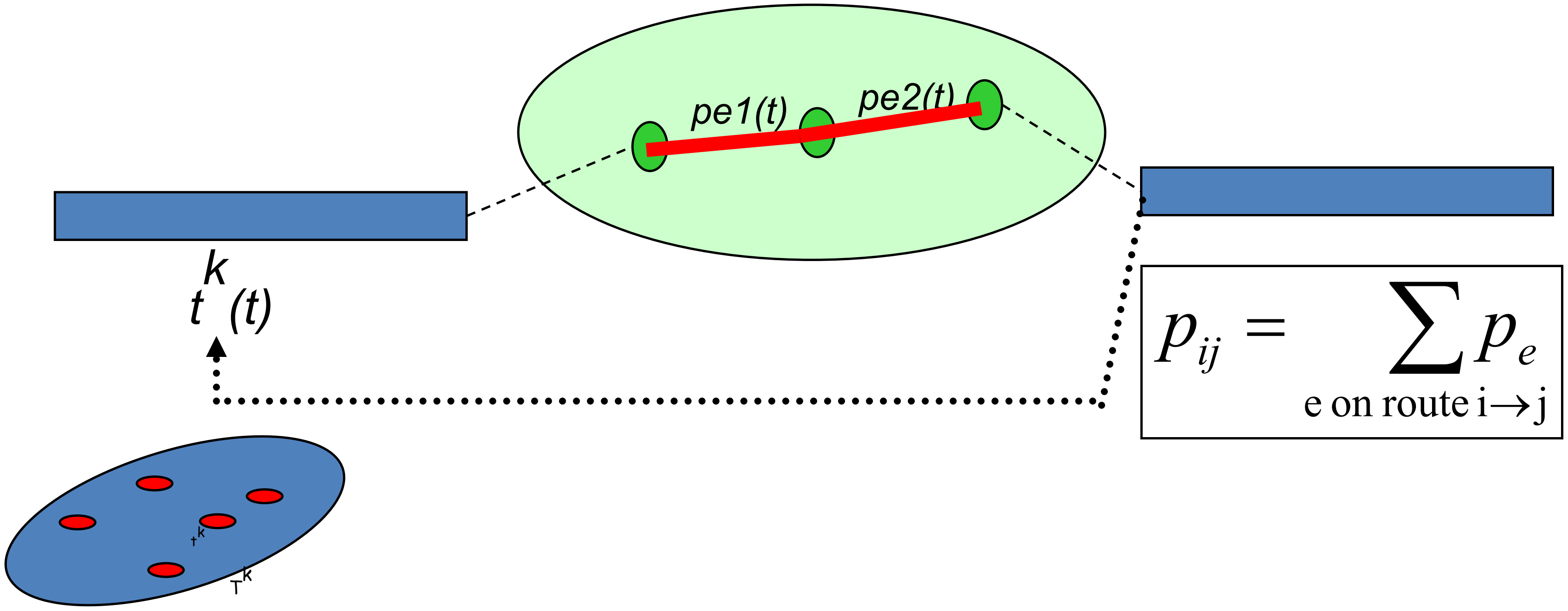
- With dual variable p_e (≥ 0) for the inequality of each link e, and dual finite condition $\sum_e p_e c_e = 1$

p_{ij} is the sum of p_e along the path from node i to node j

$$D(\{p_e\}) = \min_{\forall k: t^k \in T^k} \sum_e p_e \Big(b_e + \sum_k t_e^k\Big) = \sum_e p_e b_e + \sum_k \min_{t^k \in T^k} \sum_{i \neq j} p_{ij} t_{ij}^k$$

Bigger Picture

- The interface between applications and network is the dual variables $\{p_{ij}\}$



$$p_{ij} = \sum_{e \text{ on route } i \to j} p_e$$

From Simple Example to Real Design

- At a high level, one may indeed consider both ALTO (estimation) and (per-packet) INT info as p_{ij}
- Multiple steps to make progress
  - A systematic study of existing NAA capabilities and their support requirements
  - A systematic study of existing AAN capabilities and their support requirements
  - Systematic design of network and application abstraction spaces