Network Optimization via Path Selection in a Heterogeneous Cloud-Based Internet of Things

Eric Sturzinger

Group Meeting

19AUG 2016

UCDAVIS

Outline

- Purpose/Motivation
- Problem Statement
- Functional Based Application Requirements
- Performance & Cost Application Requirements
- Topology/Flow Scenarios
- ILP Formulation
- Dynamic Simulation
- Extending over WSNs
- Conclusion

UCDAVIS
UNIVERSITY OF CALIFORNIA

Purpose/Motivation

- 2019: Cisco predicts > 500 zettabytes of Internet traffic/year
  - 1 zettabyte = 10^{21} bytes = 1 billion TB = 1 trillion GB
- 2020: > 50 billion connected devices, cloud consumption will increase 4x by 2020 to 8.6 zettabytes, enterprises plan to transfer 56% of applications and data to cloud
- M2M Devices: 30% of all connected devices today to 46% in 2020, but only ~ 2-4 % of traffic
- Will IoT simply require more physical layer core capacity? Or will traffic nature and increased heterogeneity require more robust traffic engineering and/or policy based/constraint-based routing?

http://siliconangle.com/blog/2015/10/28/cisco-predicts-internet-of-things-will-generate-500-zettabytes-of-traffic-by-2019/

http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html

UCDAVIS UNIVERSITY OF CALIFORNIA

Problem Statement

1st Priority: Optimize performance metrics of a set of traffic flows w/ heterogeneous application & traffic profiles via path selection within a core layer internetwork.

Solution form: Unique path for each node pair & application profile, intended to generate diverse set of SLAs for each node pair

2nd Priority: Given above path solutions by node pair, minimize operational cost to MAN service provider while maintaining end to end performance requirements. Simulate various application profile mixtures and loads; determine at what loads performance is substantially degraded. Repeat for various link capacities (higher costs).

3rd Priority: Inject a multi-homed redundant 4G/5G connection from specific sensors/sink node to specific nodes in access/metro layer

UCDAVIS UNIVERSITY OF CALIFORNIA

Functional Based Application Requirements

- Each application requires various tiers of data processing capability prior to delivery to final destination
- Applications may require storage at DC
- Interactive applications: location of the client (and thus the path metrics) will affect the computation location
- Traffic may be aggregated from multiple sources for a single event

Storage Required	Processing Required	Cause	Source	Destination
None	None	Primary Event	Single Sensor	Single Sensor
Data Center	Tier 1- Sink/GW	Secondary Event	Multiple Sensors	Multiple Sensors
	Tier 2	Client Demand	DC	DC
	Tier 3 - DC	Periodic to DC	Client	Client
		Periodic to Client		

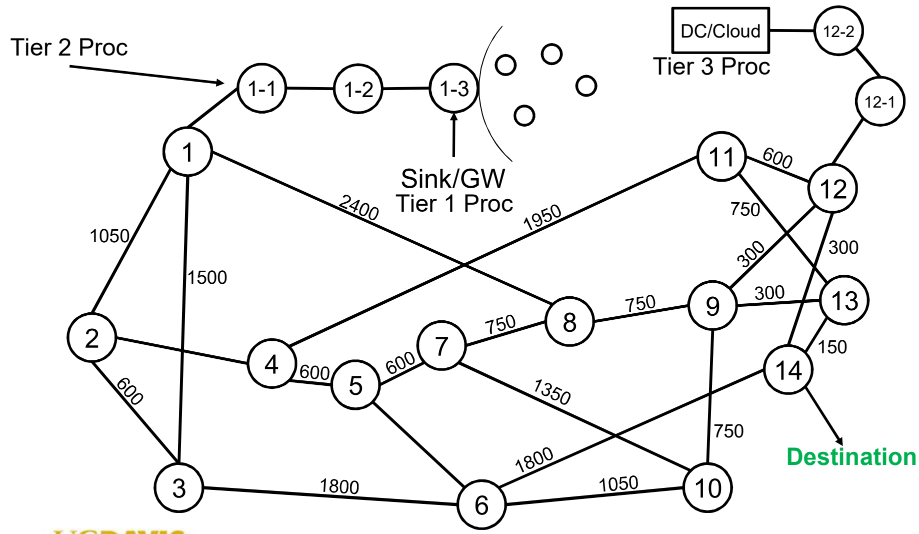
Application Performance/Cost Requirements

Latency	Bandwidth	Reliability	Jitter
Uni-directional: < 50 ms	Min 50 Mbps per link	Prob. Delivery: > 99.9%	Latency +/- 10%
Bi-directional: < 100ms		Path of Least Congestion	Latency +/- 20%

Upstream to Internet	Downstream from Internet	Storage	Processing (Tier I, II, III)
X dollars per Mbps	X dollars per Mbps	X $ per Mbit	I: X dollars per Mbps (high)
			II: X dollars per Mbps (moderate)
			III: X dollars per Mbps (low)

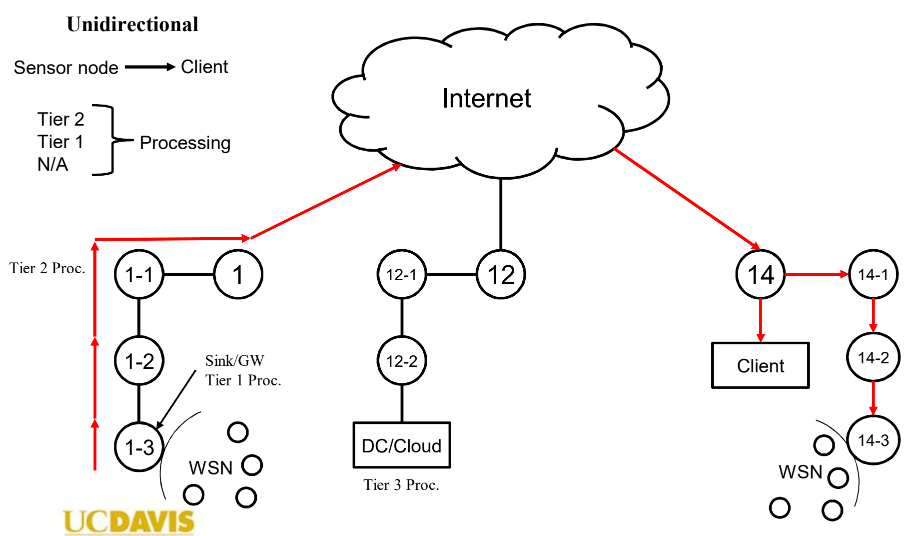
UCDAVIS
UNIVERSITY OF CALIFORNIA

NSFnet Topology



UCDAVIS
UNIVERSITY OF CALIFORNIA

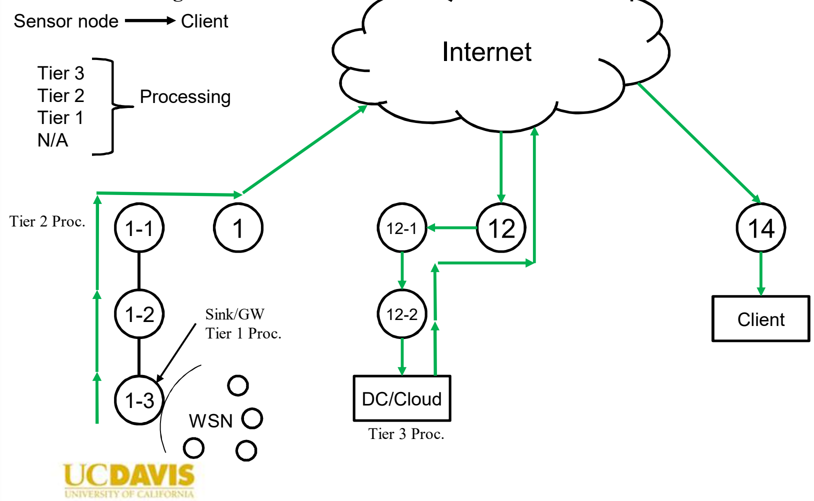
Flow Scenarios



UCDAVIS
UNIVERSITY OF CALIFORNIA

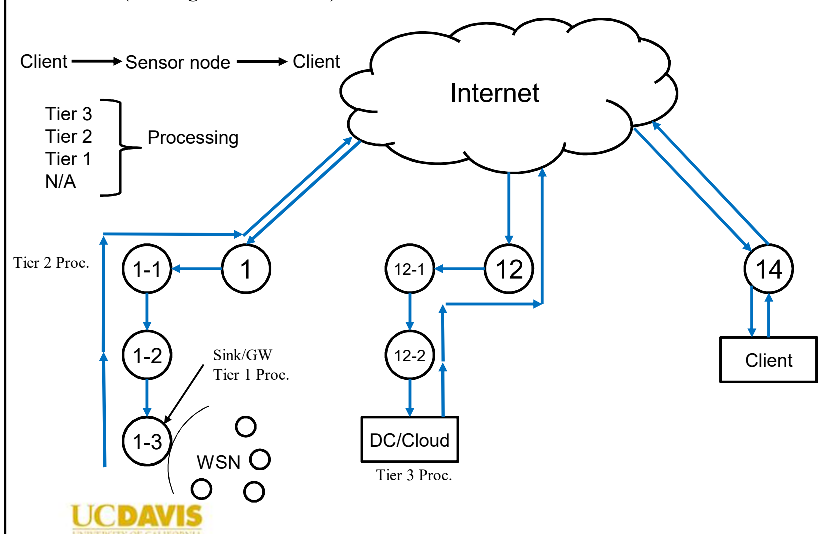
Flow Scenarios (cont.)

Unidirectional – w/ Storage and/or Tier 3 Processing

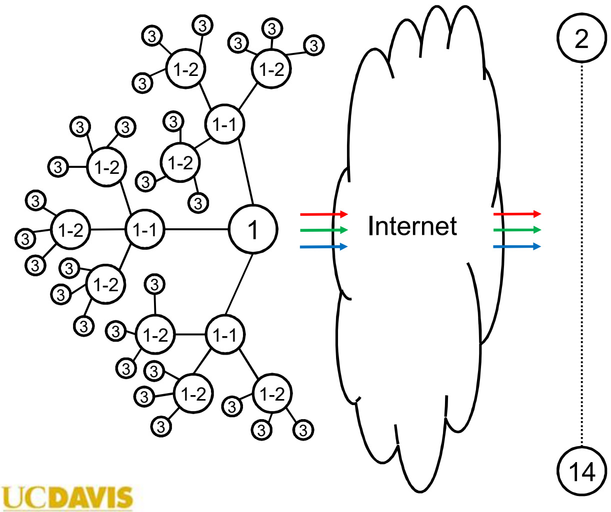


Flow Scenarios (cont.)

Interactive (Through DC or direct)

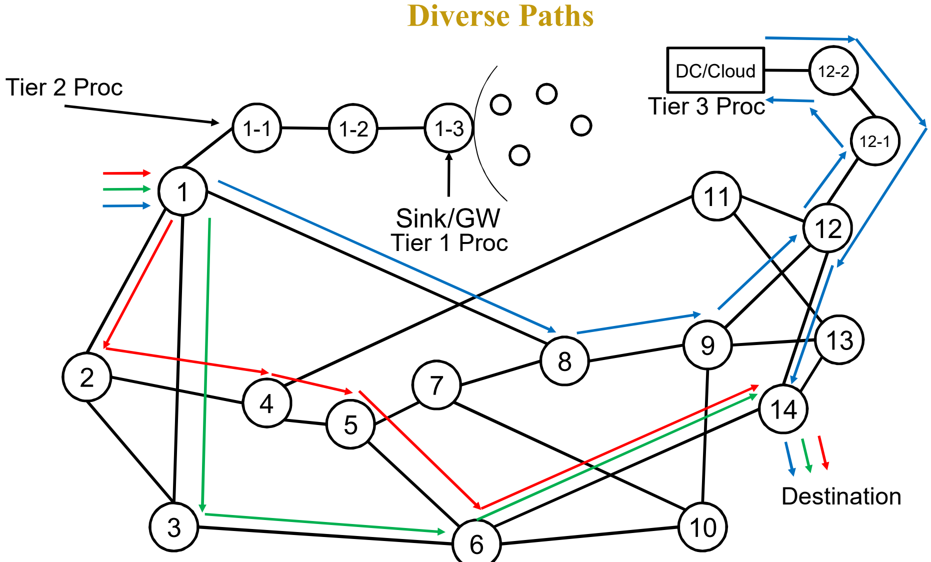


Flow Scenario Visualization



UCDAVIS
UNIVERSITY OF CALIFORNIA

Diverse Paths



UCDAVIS
UNIVERSITY OF CALIFORNIA

Mathematical Formulation

Variables:

$v_{c,a}$: Offered traffic of application profile **a**, between node pair **c**.

$r_{c,k,a}$: Traffic of application profile **a**, routed over the $\mathbf{k}^{th}$ admissible path between node pair **c**.

$S_{a,i} = 1$, if traffic of application profile **a** requires storage at DC, located behind node **i**

$d_{t,i,j}$: Transmission delay at node **i** on link **i j**

$d_{p,i,j}$: Propagation delay on link **i j**

UCDAVIS UNIVERSITY OF CALIFORNIA

Mathematical Formulation (cont.)

Variables:

$d_{n,i}$: Processing delay at node **i**

$K_{c,a}$: Set of all admissible paths between node pair **c** for appl **a**

$k_{c,a}$: $\mathbf{k}^{th}$ admissible path between node pair **c** for appl **a**

$h_{c,k,a}$: Hop count of $\mathbf{k}^{th}$ admissible path for node pair **c**, appl **a**

$P^{c}_{a,t} = 1$, if traffic of application profile **a** between node pair **c** is processed at tier **t**

UCDAVIS UNIVERSITY OF CALIFORNIA

Mathematical Formulation (cont.)

Variables:

$\kappa_{a,t,c}$: Tier **t** processing cost per unit of traffic of application profile **a** between node pair **c**

β_a: Parameter that defines relationship between computational power required and traffic of application profile **a**

ν_a: Cost per unit of traffic of application profile **a** for data center storage

UCDAVIS
UNIVERSITY OF CALIFORNIA

Mathematical Formulation (cont.)

Objective Function:

$$\max \prod_{c=1}^{N(N-1)} \prod_{a=1}^{A} \frac{Perf_a^c}{Cost_a^c}$$

$Perf_a^c$ and $Cost_a^c$ are Performance and Cost functions of **A** application profiles between each node pair **c**.

Subject to:

$$\sum_k r_{c,k,a} = v_{c,a} \ \forall (c,a)$$ Solenoidality

$$\sum_a \sum_{r_{c,k \in R_{i,j}}} r_{c,k,a} \le F_{i,j} \quad \forall (i,j)$$ Capacity

UCDAVIS
UNIVERSITY OF CALIFORNIA

Mathematical Formulation (cont.)

$$\sum_{t=0}^{3} P^c_{a,t} = 1 \;\forall (a,c)$$ Processing

$$i \in k_{c,a} \text{ if } S_{a,i} = 1 \;\forall (a,c)$$ Storage

UCDAVIS UNIVERSITY OF CALIFORNIA

Mathematical Formulation (cont.)

Performance Functions:

Total Latency (uni-directional):

$$Perf^c_1 = \sum_{(i,j)\in k_c} d_{p,i,j} + \sum_{(i,j)\in k_c} d_{t,i,j} + \sum_{(i)\in k_c} d_{n,i}$$

Total Latency (bi-directional/interactive):

$$Perf^c_2 = 2(\sum_{(i,j)\in k_c} d_{p,i,j} + \sum_{(i,j)\in k_c} d_{t,i,j} + \sum_{(i)\in k_c} d_{n,i})$$

Min Throughput Link x Total Latency

$$Perf^c_3 = (min \sum_{r_{c,k\in R_{i,j}}} r_{c,k,3})(\sum_{(i,j)\in k_c} d_{p,i,j} + \sum_{(i,j)\in k_c} d_{t,i,j} + \sum_{(i)\in k_c} d_{n,i})$$

UCDAVIS UNIVERSITY OF CALIFORNIA

Mathematical Formulation (cont.)

Cost Function:

$$Cost_a^c = \alpha_{u,a}^{s_c} v_a^c + \alpha_{d,a}^{d_c} v_a^c + v_{c,a} \sum_{t=0}^{3} \kappa_{a,t}^c P_{a,t}^c + \nu_a S_{a,i} v_{c,a}$$

Upstream, Downstream, Processing, Storage

$\alpha_{u,a}^{d_c}$ = Cost per unit of upstream traffic of application profile **a** from source of node pair **c**

$\alpha_{d,a}^{d_c}$ = Cost per unit of downstream traffic of application profile **a** to destination of node pair **c**

UCDAVIS
UNIVERSITY OF CALIFORNIA

Mathematical Explanation

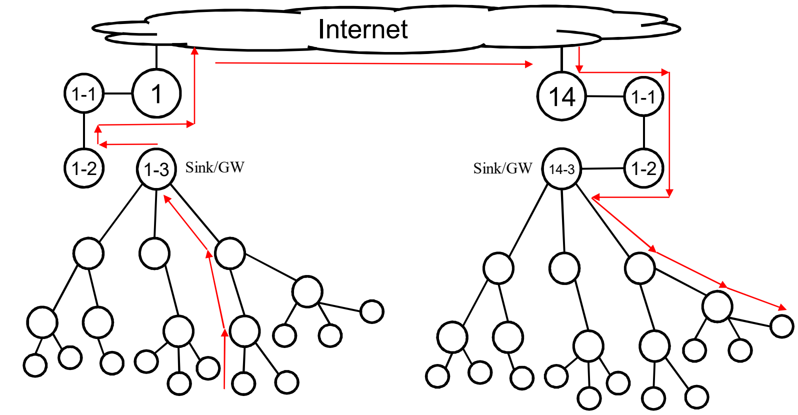
Inputs: Performance functions based on the application profiles
Cost functions
Single/multiple cloud storage/DC locations
Multiple data processing locations
Topology, profile proportions at each source node (1/4, ¼, ½)
Link Capacities

Objective function: Maximize product of performance/cost ratios via ideal paths of heterogeneous application traffic flows across all possible node pairs.

Outputs: For each node pair and application profile:
Designated path through core nodes
Intermediate processing location, if necessary

UCDAVIS
UNIVERSITY OF CALIFORNIA

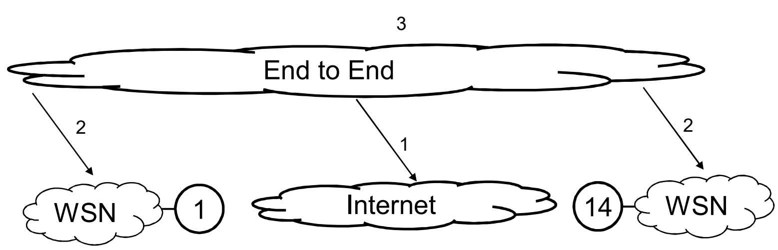
WSN/Core Integration



Logical (tree) topology in a WSN computed based on RPL

Optimized path selection per application profile can be based on the core network ILP solution

WSN/Core Integration



Future Work

ILP Solution: Ideal paths that maximize performance/cost ratio for all node pairs and application profiles

Formulate similar ILP for internal WSN path selection; determine feasibility of integration

Compute integrated (complete end-to-end) best path solution

Dynamic simulation: Given ILP solutions, simulate dynamic traffic to determine time varying delay (jitter) and other performance metrics in order to formulate strategy for further optimization.

UCDAVIS
UNIVERSITY OF CALIFORNIA

References

1. http://siliconangle.com/blog/2015/10/28/cisco-predicts-internet-of-things-will-generate-500-zettabytes-of-traffic-by-2019/
2. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html
3. Parwekar, P., 2011, September. From internet of things towards cloud of things. In *Computer and Communication Technology (ICCCT), 2011 2nd International Conference on* (pp. 329-333). IEEE.
4. 2. Bonomi, F., Milito, R., Natarajan, P. and Zhu, J., 2014. Fog computing: A platform for internet of things and analytics. In *Big Data and Internet of Things: A Roadmap for Smart Environments* (pp. 169-186). Springer International Publishing.
5. 3. Borges, L.M., Velez, F.J. and Lebres, A.S., 2014. Survey on the characterization and classification of wireless sensor network applications.*IEEE Communications Surveys & Tutorials*, *16*(4), pp.1860-1890.
6. 4. Botta, A., Pescapé, A. and Ventre, G., 2008. Quality of service statistics over heterogeneous networks: Analysis and applications. *European Journal of Operational Research*, *191*(3), pp.1075-1088.
7. 5. Fox, G.C., Kamburugamuve, S. and Hartman, R.D., 2012, May. Architecture and measured characteristics of a cloud based internet of things. In*Collaboration Technologies and Systems (CTS), 2012 International Conference on* (pp. 6-12). IEEE.

UCDAVIS
UNIVERSITY OF CALIFORNIA