Centrality-driven Scalable Service Migration

Panagiotis Pantazopoulos, Merkouris Karaliopoulos, Ioannis Stavrakakis

presented by M. Karaliopoulos

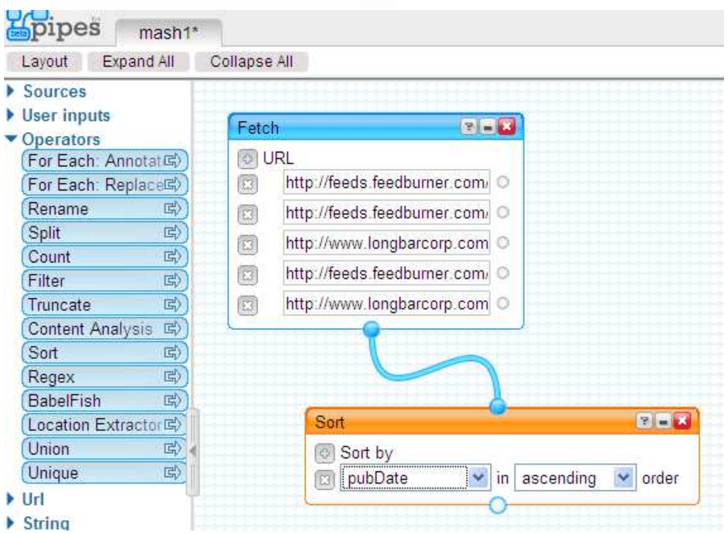
Advanced Networking Research Group
National and Kapodistrian University of Athens
{ppantaz, mkaralio,ioannis@di.uoa.gr}

The changing role(s) of end-users

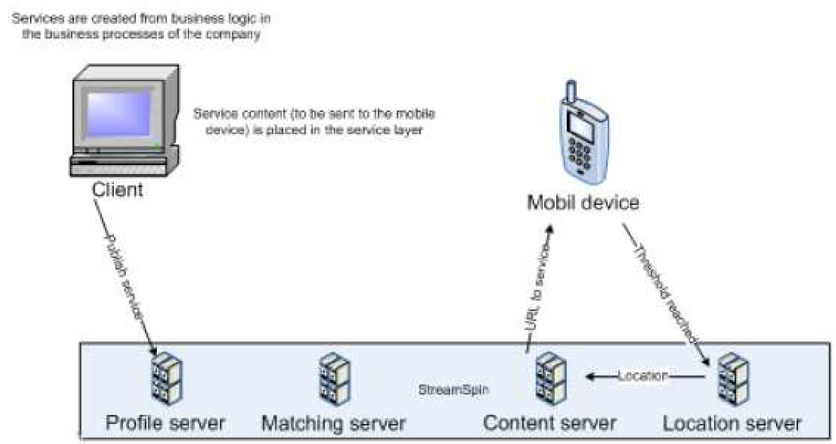
- User Generated Content → end-user as content provider
  - 25% of Google results point to UGC sites (as of 2009)
  - UGC is expected to triple in 2008-2013

- User-centric service creation → end-user as service provider

YAHOO! **Pipes**



Streamspin



Where to place services and content

- Properties/assumptions
  - generated almost anywhere across the network
    - lack of centralized control/coordination
  - many in number, often of local (small-scale) demand (replication: not preferable)

- Objective
  - deploy scalable and distributed mechanisms for “optimally” placing UG Service components
  - <u>in this work</u> : “optimally” ≡ minimize aggregate service access costs

Facility location problem

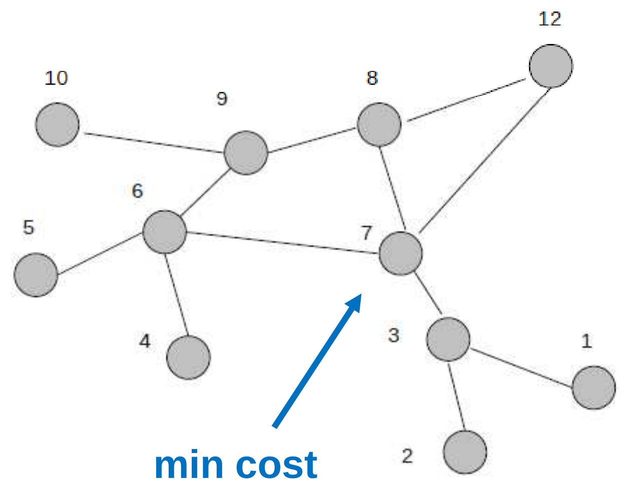
- INPUT
  - V : set of nodes
  - w_n : demand generated by node n
  - $d(x_j,n)$: distance between nodes x_j and n

OUTPUT

F : placement

- k-median problem : open up to k facilities so as to minimize the total service cost

$$Cost(\mathcal{F}) = \sum_{n \in \mathcal{V}} w_n \cdot min_{x_j \in \mathcal{F}} \{d(x_j, n)\}$$



- 1-median: minimize the access cost of a service located at node k

$$Cost(k) = \sum_{n \in \mathcal{V}} w(n) \cdot d(k, n).$$

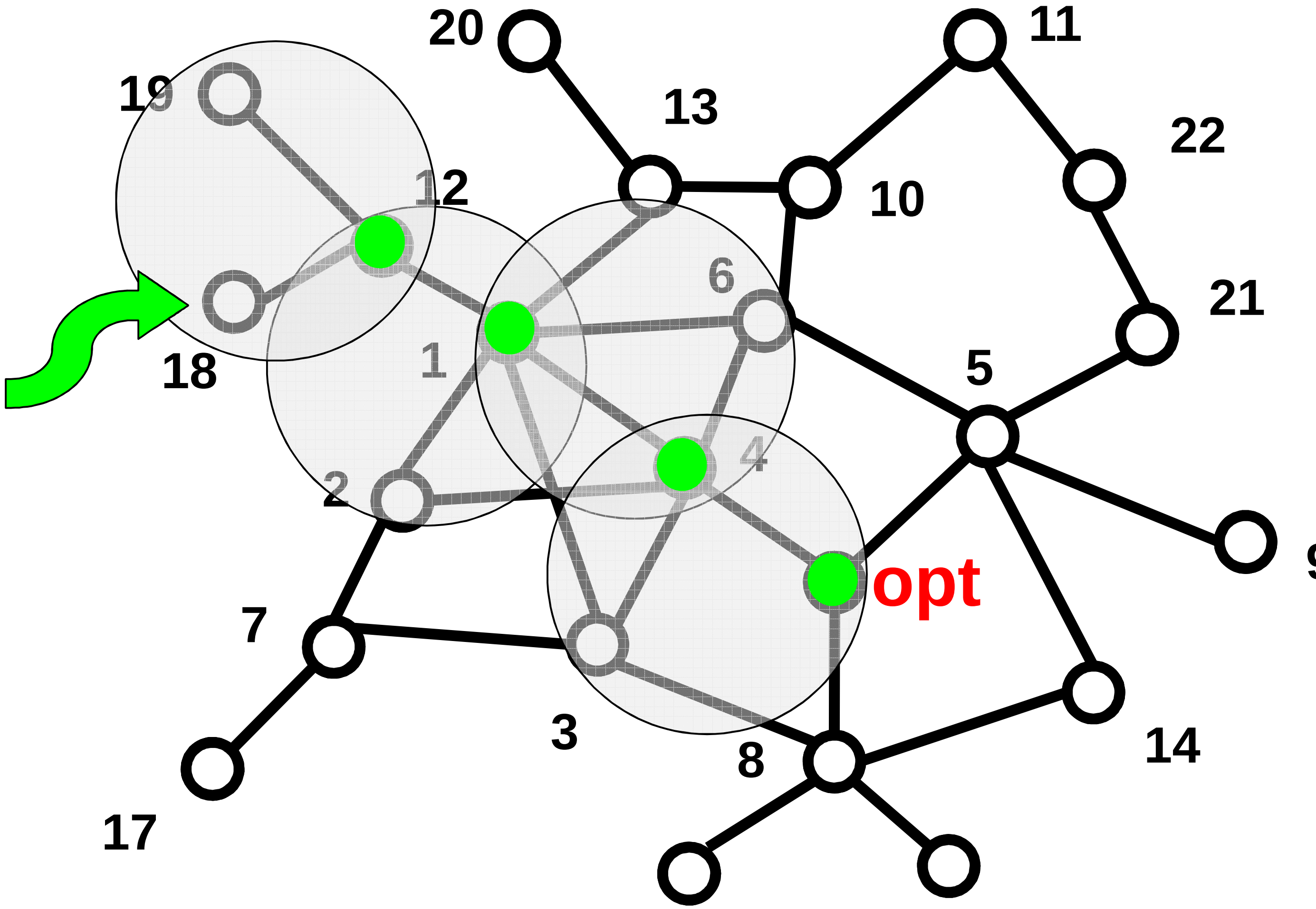
Distributed approaches to facility location

- Centralized solutions – requires a single super-entity that
  - Gathers network wide information
  - Undertakes computations
  - Accounts for demand/topology changes

- Distributed solutions
  - Theoretical work
    - require a certain (albeit small) amount of global knowledge
    - require impractical communication models (client-facility communication in each round)
    - do not always improve over existing heuristic solutions
  - Heuristic solutions
    - Less rigorous but practically implementable

Heuristic local-search approaches

- service migrates towards the *optimum* host (opt) in a finite number of steps

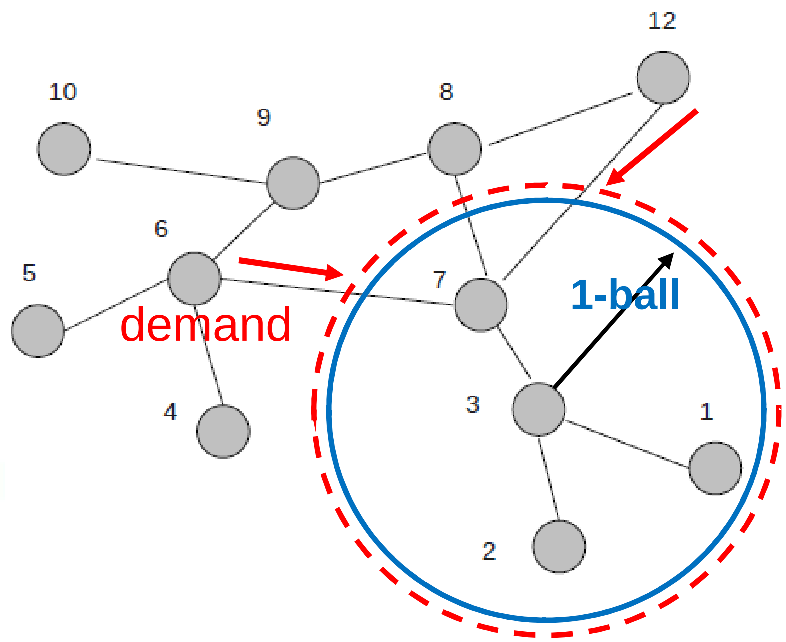


The R-balls heuristic*

- Reduce the original k-median to multiple smaller 1-median problems
  - solved within a limited neighborhood of R-hops around current facility

- Demand generated by outer nodes is *mapped* to the nodes at the outer shell of the R-hop neighborhood

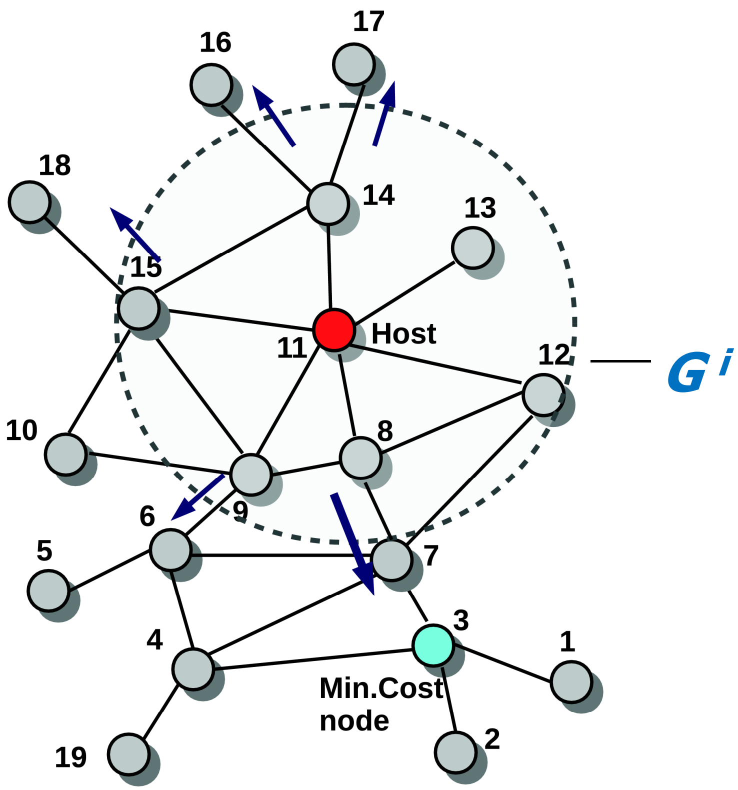
$$Cost(\mathcal{F}) = \sum_{n \in \mathcal{V}} w_n \cdot min_{x_j \in \mathcal{F}} \{d(x_j, n)\}$$



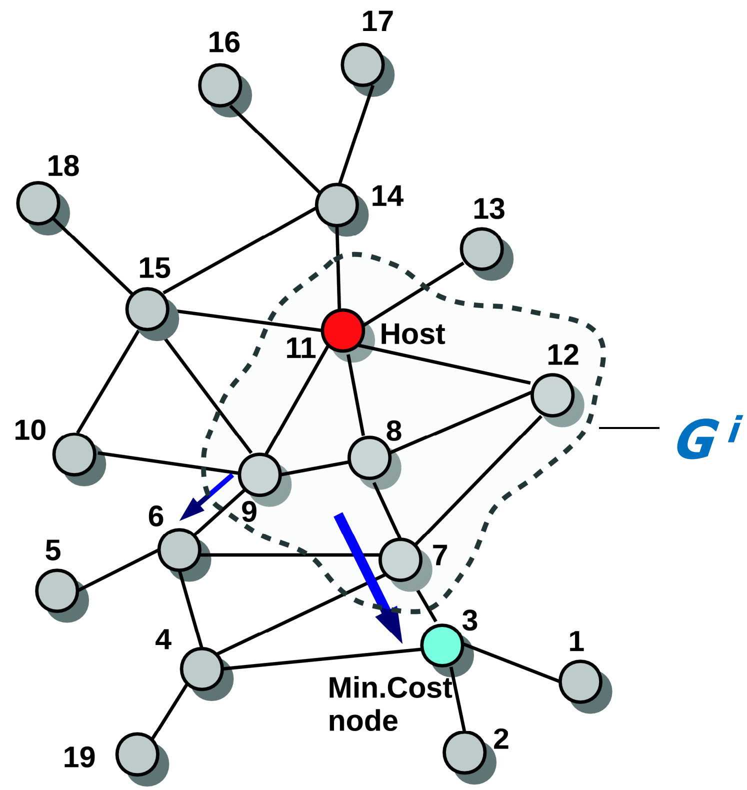
* G. Smaragdakis, N. Laoutaris, K. Oikonomou, I. Stavrakakis, A. Bestavros, *"Distributed Server Migration for Scalable Internet Service Deployment,"* to appear in IEEE/ACM ToN

Making the search "*more informed*": cDSMA

R-ball heuristic



cDSMA



The presentation remainder

- How does cDSMA work
  - Choice of 1-median subgraph G^{i}
  - Demand mapping and solution of the smaller-scale optimization problem

- How well *can* cDSMA perform
  - how close to optimal is the chosen location, how fast is this reached, how complex is this

- How can cDSMA be practically implemented in real networks
  - local approximations for global information

cDSMA : Capturing the topology factor

a measure of the importance of node's u social position : lies on paths linking others

- Betweenness Centrality (u): sums the portions of all pairs' shortest paths in G that pass through node u

$$BC(u) = \sum_{s=1}^{|V|} \sum_{t=1}^{s-1} \frac{\sigma_{st}(u)}{\sigma_{st}}$$

- **Conditional** Betweenness Centrality (*u,t*) : portion of all shortest paths **towards target node t** in G, that pass through node *u*

a measure of the importance of node's u social position : ability to control information flow towards *target* node

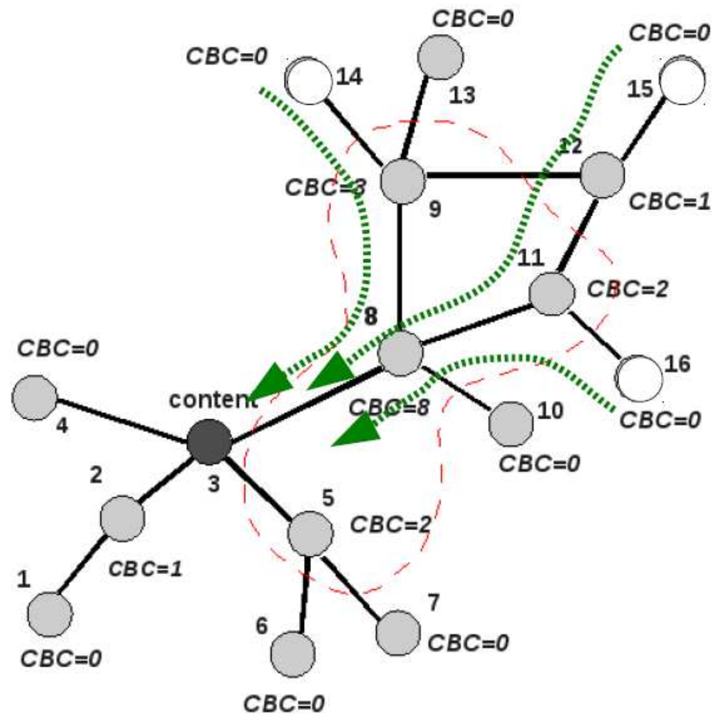
$$CBC(u;t) = \sum_{s \in V, u \neq t} \frac{\sigma_{st}(u)}{\sigma_{st}}$$

cDSMA : Capturing the demand distribution

- a high number of shortest paths through the node *u* (e.g. node 8) does not necessarily mean that equally high demand stems from their sources!

- *weighted* conditional BC

$$wCBC(u;t) = \sum_{s \in V, u \neq t} w(s) \cdot \frac{\sigma_{st}(u)}{\sigma_{st}}.$$



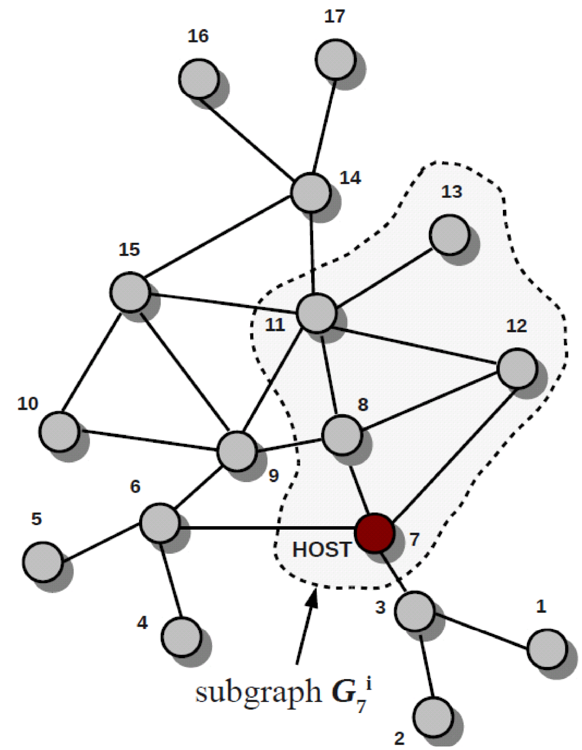
- $wCBC$ assesses to what extent a node can serve as demand concentrator towards a given service location
  - The top a% wCBC-valued nodes are included in the 1-median subgraph

Projecting the "world outside" on the selected nodes

- wCBC metric eases the demand mapping of the G/G^i nodes (world outside), on the selected G^i ones
  - nodes in G^i exhibit an effective demand:

$$w_{eff}(n;Host) = w(n) + w_{map}(n;Host)$$

- $w_{map} \neq 0$ only for the outer nodes of the 1-mediad subgraph
  - demand of node z ϵ G/G^i is "credited" *only to the first* G^i nodes encountered on each shortest path from z towards the host



cDSMA in summary

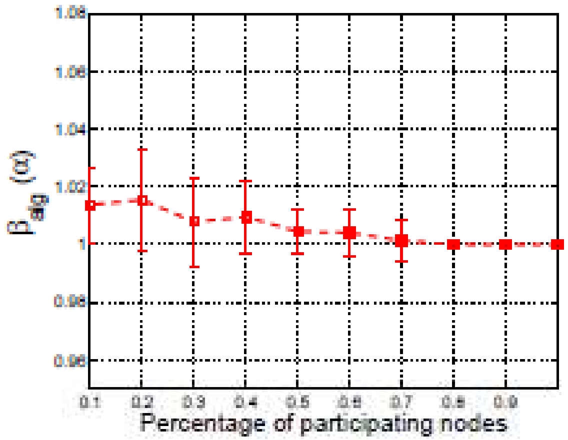
Algorithm 1 cDSMA in *G(V,E)*

choose randomly node s
place SERVICE @ s
$C_{current} \leftarrow \infty$
for all $u \in G$ **do** *compute* $wCBC(u; s)$
$G_s^o \leftarrow \{\alpha\%\ of\ G\ with\ top\ wCBC\ values\} \cup \{s\}$
for all $u \in G_s^o$ **do**
    compute $w_{map}(u; s)$
    $w_{eff}(u; s) \leftarrow w_{map}(u; s) + w(u)$
$Host \leftarrow$ *1-median solution in* G_s^o
$C_{next} \leftarrow C(Host),\ i \leftarrow 1$
while $C_{next} < C_{current}$ **do**
    move SERVICE to Host
    $C_{current} \leftarrow C_{next}$
    for all $u \in G$ **do** *compute* $wCBC(u; Host)$
    $G_{Host}^i \leftarrow \{\alpha\%\ of\ G\ with\ top\ wCBC\ values\} \cup \{Host\}$
    for all $u \in G_{Host}^i$ **do**
        compute $w_{map}(u; Host)$
        $w_{eff}(u; Host) \leftarrow w_{map}(u; Host) + w(u)$
    $NewHost \leftarrow$ *1-median solution in* G_{Host}^i
    $Host \leftarrow NewHost, C_{next} \leftarrow C(NewHost), i \leftarrow i + 1$
end while

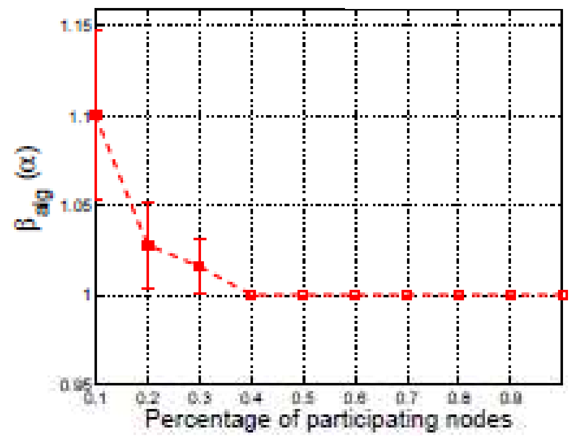
- Convergence in O(N) steps

Evaluation in synthetic topologies

- 2D Grids and B-A like networks, demand model : Zipf(s) distribution
- Performance metric : normalized access cost $\beta_{alg}(\alpha; G, \overline{w}) = E[\ \frac{C_{alg}(\alpha; G, \overline{w})}{C_{opt}(\alpha; G, \overline{w})}\ ]$

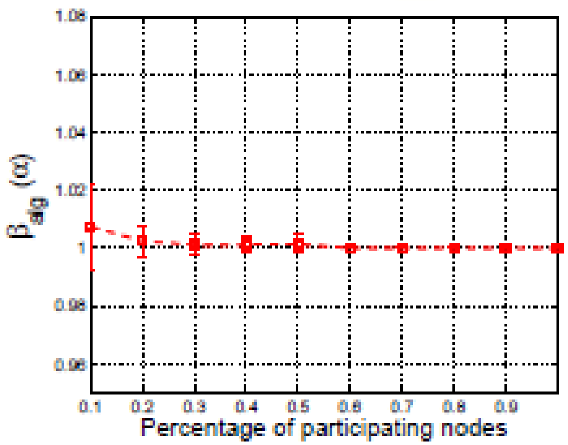


a. B-A (s=0)

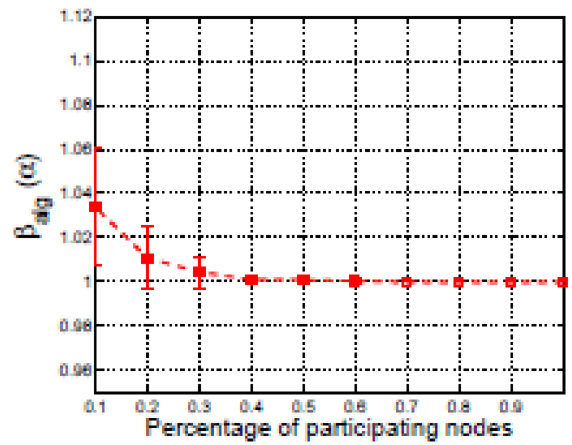


b. Grid (s=0)

100 nodes



c. B-A (s=1)



d. Grid (s=1)

Evaluation of real-world network topologies

- Datasets correspond to different snapshots of 7 ISPs collected by mrinfo multicast tool*

$$\alpha_\epsilon = argmin\ \{\alpha | \beta_{alg}(\alpha) \leq\ (1+\epsilon)\}$$

					s=0		s=1		s=2	
ISP	Dataset id/AS#	mCC nodes	Diameter	<Degree>	$\alpha_{0.025}$	$\lceil \lvert G^s \rvert \rceil$	$\alpha_{0.025}$	$\lceil \lvert G^s \rvert \rceil$	$\alpha_{0.025}$	$\lceil \lvert G^s \rvert \rceil$
type: **Tier-1**										
Global Crossing	36/3549	76	10	3.71	0.047±0.001	4	0.047±0.002	4	0.046±0.001	4
-//-	35/3549	100	9	3.78	0.045±0.002	5	0.045±0.001	5	0.043±0.001	5
NTTC-Gin	33/2914	180	11	3.53	0.024±0.002	5	0.022±0.002	4	0.019±0.002	4
Sprint	23/1239	184	13	3.06	0.019±0.002	4	0.018±0.002	4	0.017±0.002	4
-//-	21/1239	216	12	3.07	0.016±0.002	4	0.016±0.002	4	0.014±0.003	4
Level-3	27/3356	339	24	3.98	0.018±0.002	7	0.017±0.002	6	0.014±0.003	5
-//-	13/3356	378	25	4.49	0.012±0.002	5	0.012±0.002	5	0.011±0.002	5
type: **Transit**										
TDC	46/3292	71	9	3.30	0.033±0.003	3	0.027±0.004	2	0.026±0.003	2
DFN-IPX-Win	41/680	253	14	2.62	0.019±0.003	5	0.015±0.003	4	0.015±0.003	4
JanetUK	40/786	336	14	2.69	0.012±0.003	5	0.012±0.002	5	0.013±0.002	5

Less than half a dozen nodes suffice in almost all cases, even under uniform demand

* J.-J. Pansiot, P. Mérindol, B. Donnet, and O. Bonaventure, "Extracting intra-domain topology from mrinfo probing," in Proc. Passive and Active Measurement Conference (PAM), April 2010.

How much improvement does cDSMA bring?

- Migration hop count metric (h_m) reflects the convergence speed

- Experiment
  - generate asymmetric service demand, Zipf (1)
  - fix set of service generation points at D_{gen} *hops away from optimal location*

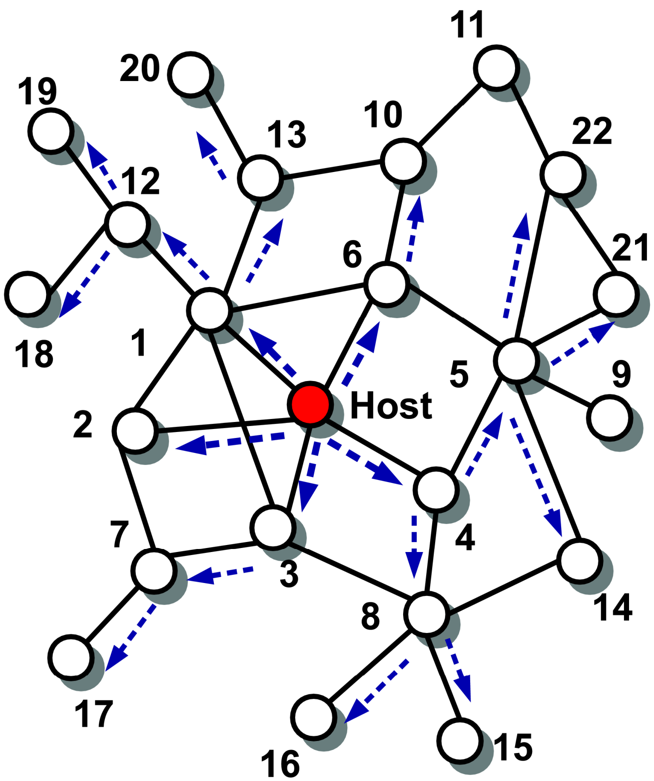
- Compare 2-ball-like (LOM) vs. cDSMA
  - 3% of total number of nodes form the *1-median* subgraph (6-12 nodes)

D_{gen}	Dataset 23				Dataset 33				Dataset 27				Dataset 13			
	LOM		cDSMA		LOM		cDSMA		LOM		cDSMA		LOM		cDSMA	
	h_m	β_{alg}	h_m	$\beta_{alg}(3\%)$	h_m	β_{alg}	h_m	$\beta_{alg}(3\%)$	h_m	β_{alg}	h_m	$\beta_{alg}(3\%)$	h_m	β_{alg}	h_m	$\beta_{alg}(3\%)$
3	1	1.1050	2	1	1	1.0308	2	1	1	1.1109	1	1.0057	1	1.1054	1	1
4	1	1.1275	3	1	1	1.3206	2	1	1	1.2523	1	1.0057	1	1.2312	1	1
5	1	1.1632	2	1	1	1.2800	1	1.2800	2	1.1109	1	1	1	1.0434	2	1
7	1	1.6060	2	1	3	1.0308	1	1.0308	3	1.1763	1	1	1	1.4202	1	1
10	–	–	–	–	–	–	–	–	1	1.7094	2	1	1	1.4604	2	1
13	–	–	–	–	–	–	–	–	2	1.8579	1	1.0057	3	1.6887	1	1.1054

***the "blind" LOM* search for next-best solution either terminates prematurely or, more rarely, gets to the same result more slowly**

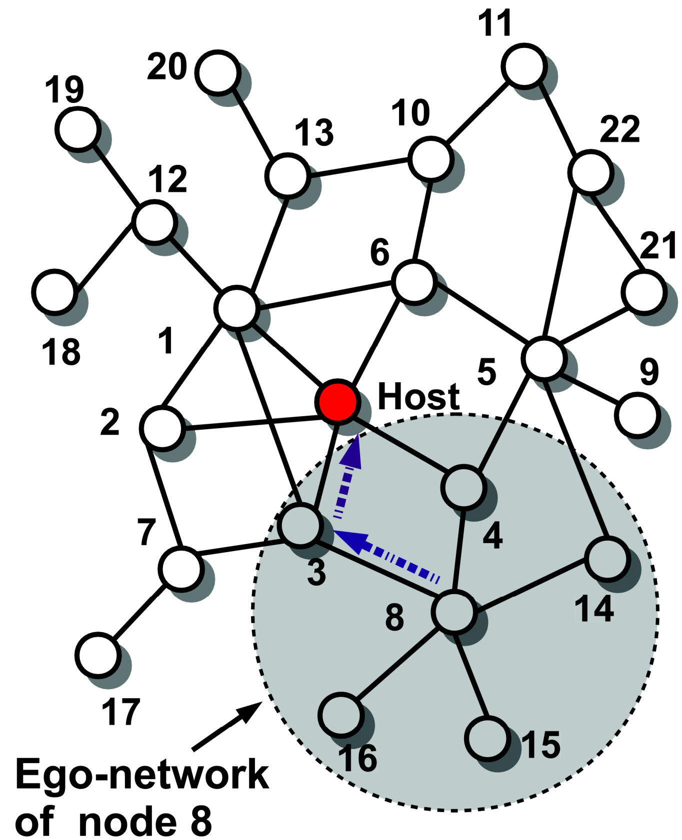
Towards a distributed protocol implementation

Step 1: Service Advertisement



$\mathcal{O}$(E) messages $\mathcal{O}$(D) time
D:diameter

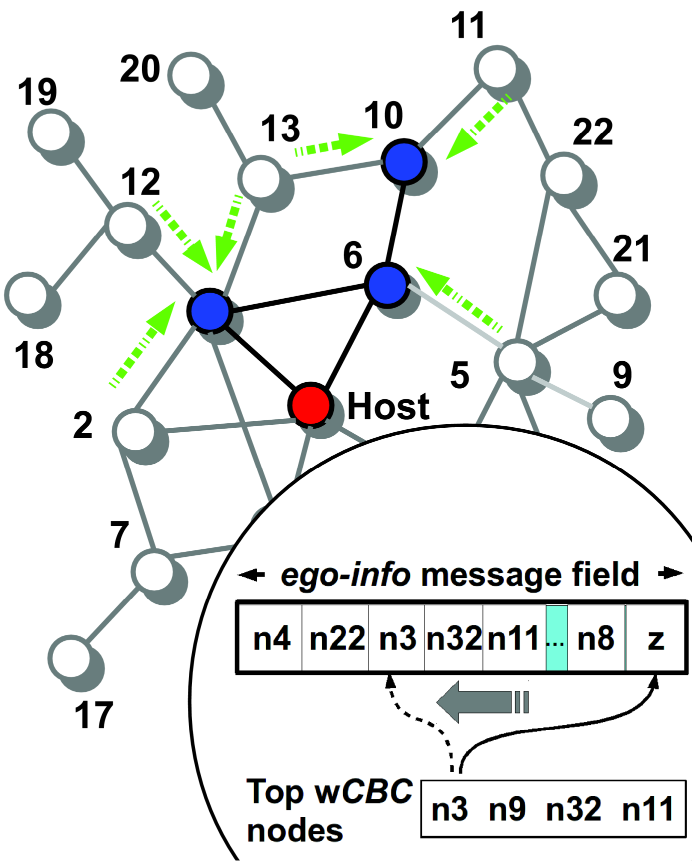
Step 2: Local metrics computation & reporting



$\mathcal{O}(d^3)$ time $\mathcal{O}$(E) messages
d: maximum degree

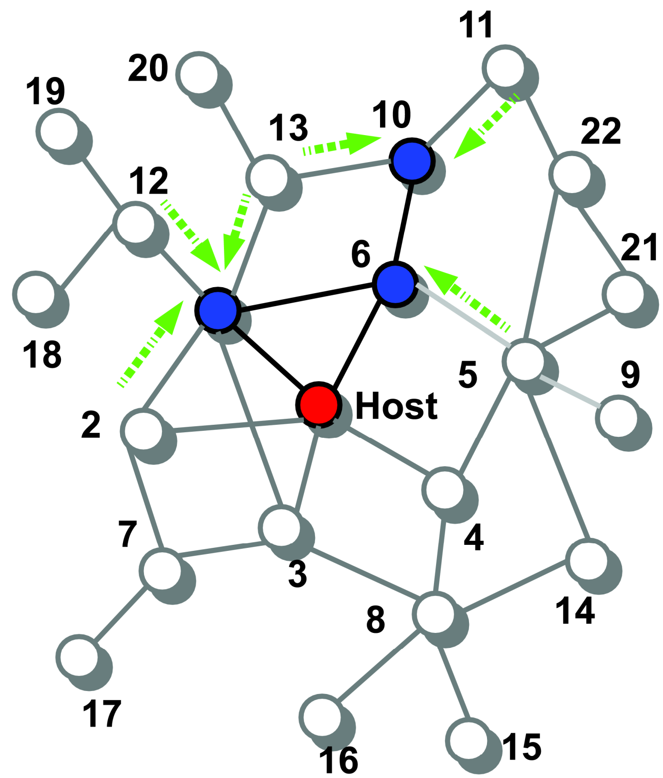
Towards a distributed protocol implementation

Step 3: Host identifies key nodes/maps the demand



O((1-a)VDlog(aV)) time

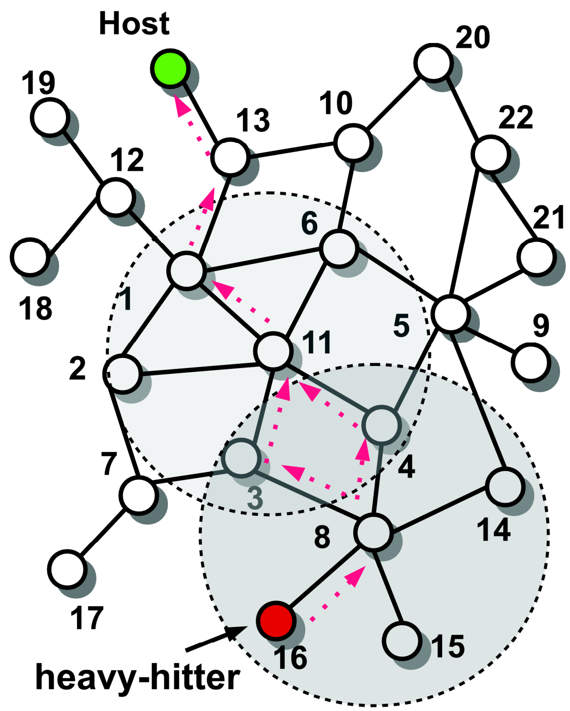
Step 4 : Host solves 1-median after nodes report their pair-wise distances



pair-wise distances among key nodes $O(a^2V^2)$

Implementation caveats and complexity

- Demand-aware assessment on ego-network scale fails to detect distant heavy-hitters
  - Egocentric centrality estimation of '11' does not account for the demand load coming from '16'



- Solution: node-centric passive measurements of the passing-through demand instead of computing it
  - may lose in accuracy when there are multiple shortest paths towards the service host

- Resulting complexity
  - $O(hm(a)a^2V^2)$ *vs.* $O(V^3)$ (brute-force approach)
  - for a=3% and hm(a)≈ 3 lead to cost reduction of one or more orders of magnitude!

Summarizing…

- ✓ We propose a heuristic algorithm (cSDMA) for *scalable* and *distributed* service placement drawing on the Social Network Analysis
  - the service migrates to the (sub)optimal location via a sequence of small-scale optimizations
  - the centrality metric singles out a subset of nodes that can act as demand concentrators and projects on them the attraction forces of the ones left-out

- ❑ The network topology structure spatial demand dynamics affect the accuracy/convergence speed of the algorithm
  - the *higher* the asymmetry (in either of above factors) the *better* the performance
  - ✓ realistic topologies exhibit enough asymmetry to achieve very good accuracy with less than a dozen nodes !

- ✓ A distributed protocol implementation was sketched and its complexity was analyzed
  - Egocentric approximations do not perform satisfactorily under asymmetric demand distributions -- passive measurement based approach under evaluation

Future directions

- Ultimate assessment of practical implementation through simulations
  - How much is lost due to local approximations

- Relaxing the ‘perfect cooperation assumption' :
  - Sensitivity to node churn and selfishness expressions
    - Nodes denying hosting services they are not interested in
  - Decision/game-theoretic dimension
    - e.g., mechanism design for truthful declaration of demand

Centrality-driven Scalable Service Migration

Panagiotis Pantazopoulos, Merkouris Karaliopoulos, Ioannis Stavrakakis

presented by M. Karaliopoulos

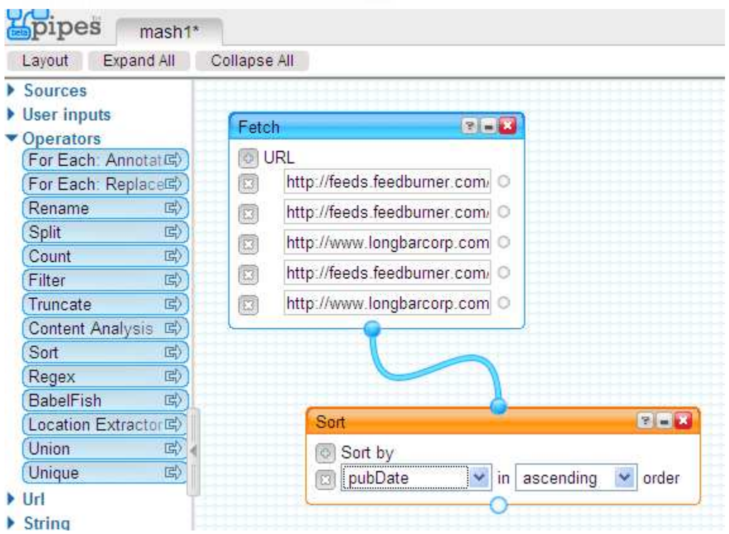
Advanced Networking Research Group
National and Kapodistrian University of Athens
{ppantaz, mkaralio,ioannis@di.uoa.gr}

Backup slides

...and the future : User-Centric Services

- User-centric service creation: engaging end-users to generate/distribute service components
  - end users with no specific knowledge
  - use high-level abstraction
  - graphical tools to provide interfaces for creating simple applications

YAHOO! **Pipes**



Streamspin