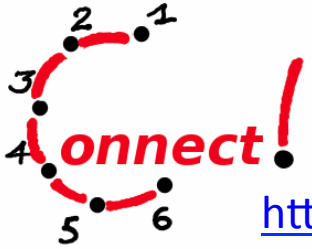
A dive into CCN* caching performance

Dario Rossi
Giuseppe Rossini
Raffaele Chiocchetti

dario.rossi@enst.fr
http://www.enst.fr/~drossi



http://www.anr-connect.org/

*ccn =
+ chunk-based icn
+ interest aggregation
+ data backtrace interests path

Agenda

(Perils of) Naïve multi-path

Parameter space exploration

The catalog issue

Main message of the dive:

Better hit the right spot!

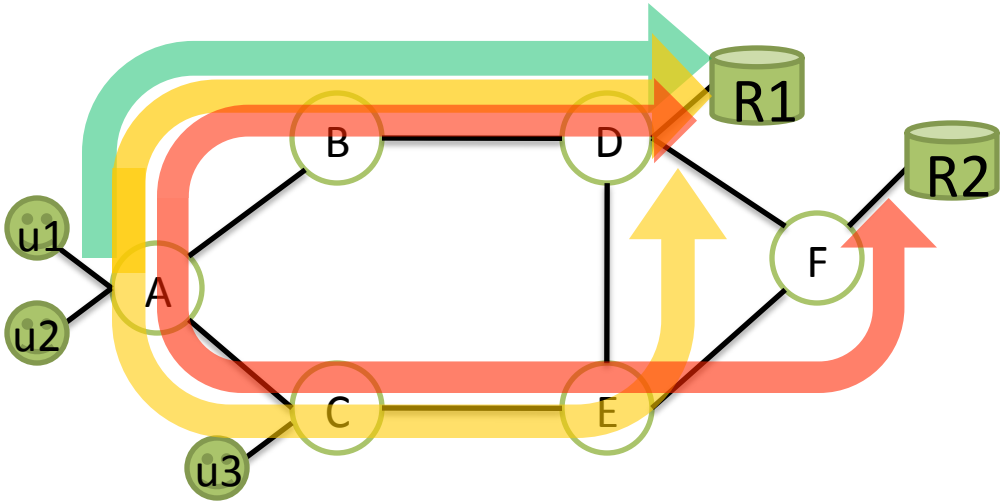
http://www.telecom-paristech.fr/~drossi/ccnSim/

(Perils of) Naïve multi-path (1/2)

- Assumptions
  - One (R=1) or more (R>1) repositories store persistent object replicas
  - Shortest path toward closest repository known
  - Alternate optimal (shortest most disjoint from primary) path toward closest repository (possibly) known
  - In case R>1, shortest paths toward all repositories (possibly) known
- Interest forwarding policies
  - Closest repository (for all chunks)
  - Closest repository (over multiple alternate paths; in parallel for all chunks, or uniformly at random per chunk, or round robin per chunk,…)
  - All repositories (in parallel, random, round robin,...)

Observations:

- Shortest path (u1,R1) may *miss cached* copies from previous (u3,R*) requests
- Uncoordinated decisions on multiple paths further *reduce aggregation*
- >1 requests for the same chunk over >1 paths yield to *multiple downloads*



Note: color code *not* associated with result quality, etc.

(Perils of) naïve multi-path (2/2)

Where we stand

- *Exploit content replicas knowledge*
  - Assume full FIB knowledge
  - May be unrealisitc, but provides baseline of what we could achieve if we had that knowledge
- *Limits of exploitation, network*
  - Multiple interests for the same content toward multiple repo (R>1) reduce aggregation efficiency [2]
  - Multiple interests for same content over independent paths to a single repo (R=1) generate cache collision & eviction on multiple caches [1]
  - Different interests over multiple paths does not successfully exploit ubiquitous caching (limits of en-route caching, with reduced aggregation, due to independent per-chunk decisions)
- *Limits of exploitation, user*
  - Different interests over multiple paths less efficient in terms of delay

Where we're heading

- *Explore cache neighborhood*
  - Ongoing work with Alcatel Lucent Bell Labs (Giovanna & Diego)
- *Aim*
  - Find *closest cached copy*
  - Interest dissemination as a means to enforce *implicit cache coordination*
- *Means*
  - Greedy policies (due to line of speed requirement)
  - Joint dissemination/caching policies investigation
  - A scent of cache coordination (at longer time scales)

Parameter space exploration

- Wide boundaries explored
  - Cache *decision* policies (LCE, LCD, RND) in [1,2]
  - Cache *replacement* (LRU, FIFO, RND, Bias) in [1,2]
  - Topologies (GEANT, Tiger2, Abilene, DTelekom, Level3) [1]
  - Catalog popularity MZipf($\alpha \in [0.5,2.5]$,$q \in [0,50]$) in [1]
  - Cache over catalog ratio [10^{-5}, 10^{-1}] in [1]
  - Locality of user requests [2]
  - Cache sizing according to topological information [3]

- Impact rank (high to low; quantitative figures in the papers)
  - Cache over catalog ratio & MZipf settings *-> hence next slide*
  - Interest forwarding *-> hence prev slide*
  - Decision, replacement, request locality, topology
  - Heterogeneous cache sizing

The catalog issue

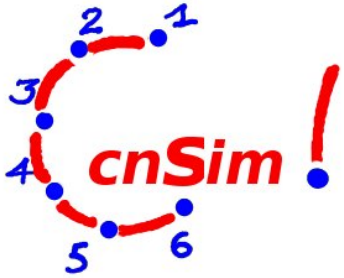
- Many work, many scenarios:
  - Necessary to agree on a reference baseline scenario to promote cross-comparison
    - A scenario that should be tried by everybody – leaving freedom to test any other scenario a
  - Otherwise 10 years from now, ICN community will issue call for papers on
    - Special Issue of Elsevier Ad Hoc Networks on "SCEnarios for ad hoc Network Evaluation Studies (SCENES)" to be published on 2012
  - Right now we have, at best an *heterogenity* of settings
    - E.g. of bizarre settings (very small caches, large cache/catalog ratios, large alphas…) in [1]
- The issue:
  - Parameters not under control, but hard to get in practice (probably a matter of research per se)
  - Many opinions, few based on up-to-date publicly available data
    - Ghodsi et al. at HotNets'11 dismiss caching as unfeasible, but based on 10 year old Web caching studies (adopting infinite cache size)… they mention KaZaa but not video (and according to Cisco visual index they should)!
- One direction to get beyond the issue:
  - Promote *open competitions* on standard reference scenarios (requiring open-source implementation for verification?)
  - Competitions are commonplace in other communities (KDD, Sortbenchmark.org, etc.)
  - IRTF good forum to steer this (risks: TCP benchmark suite ? RFC2544 not common practice !), also faster than most conferences/journal lifecycles

Advertisement

[1] D. Rossi, G. Rossini, *Caching performance of content centric networks under multi-path routing (and more)* . Technical report, Telecom ParisTech, 2011.

[2] G. Rossini, D. Rossi, *A dive into the caching performance of Content Centric Networking* . Technical report, Telecom ParisTech, 2011.

[3] D. Rossi and G. Rossini, *On sizing CCN content stores by exploiting topological information* . In NOMEN Worshop at IEEE Infocom'12, , Orlando, FL, March 25-30 2012.



ccnSim

- scalable : 10GB caches, 1PB catalog, >30 nodes networks
- open source: C++ with Omnet++ framework
- (finally) available http://www.telecom-paristech.fr/~drossi/ccnSim/
- (best effort) support at ccnsim@lincs.fr