ICN caching feasibility and trade-offs

Diego Perino, Bell Labs, Alcatel-Lucent

ICNRG side meeting @ IETF-83 in Paris, March 28th 2012

Joint work with G. Carofiglio and M. Varvello (Bell Labs, Alcatel-Lucent), M. Gallo and L. Muscariello (Orange Labs)

Storage in Content Centric Networking

ICN provides a unique pervasive storage infrastructure allowing efficient and flexible utilization of network resources (storage, bandwidth, processing)

- Packet-level management at high speed
- Multiple applications sharing the same resources
- Joint proactive content placement and reactive caching
- Close relation with transport protocols
- Caching architecture related to name-based routing

In-network storage can reduce network bandwidth utilization, improve user Quality of Experience, decrease server resource requirements, enhance service availability

Feasibility - Cache Design

Packet Store where Data packets are actually contained

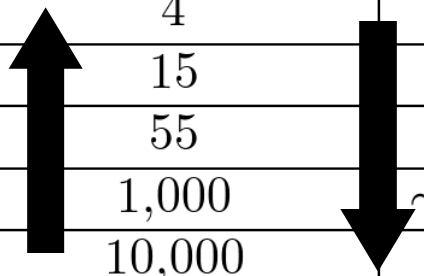
- Design guidelines
  - Sustain the Data packet arrival/departure rates
  - Maximize the number of entries

Index table which keeps track of Data packets in the packet store

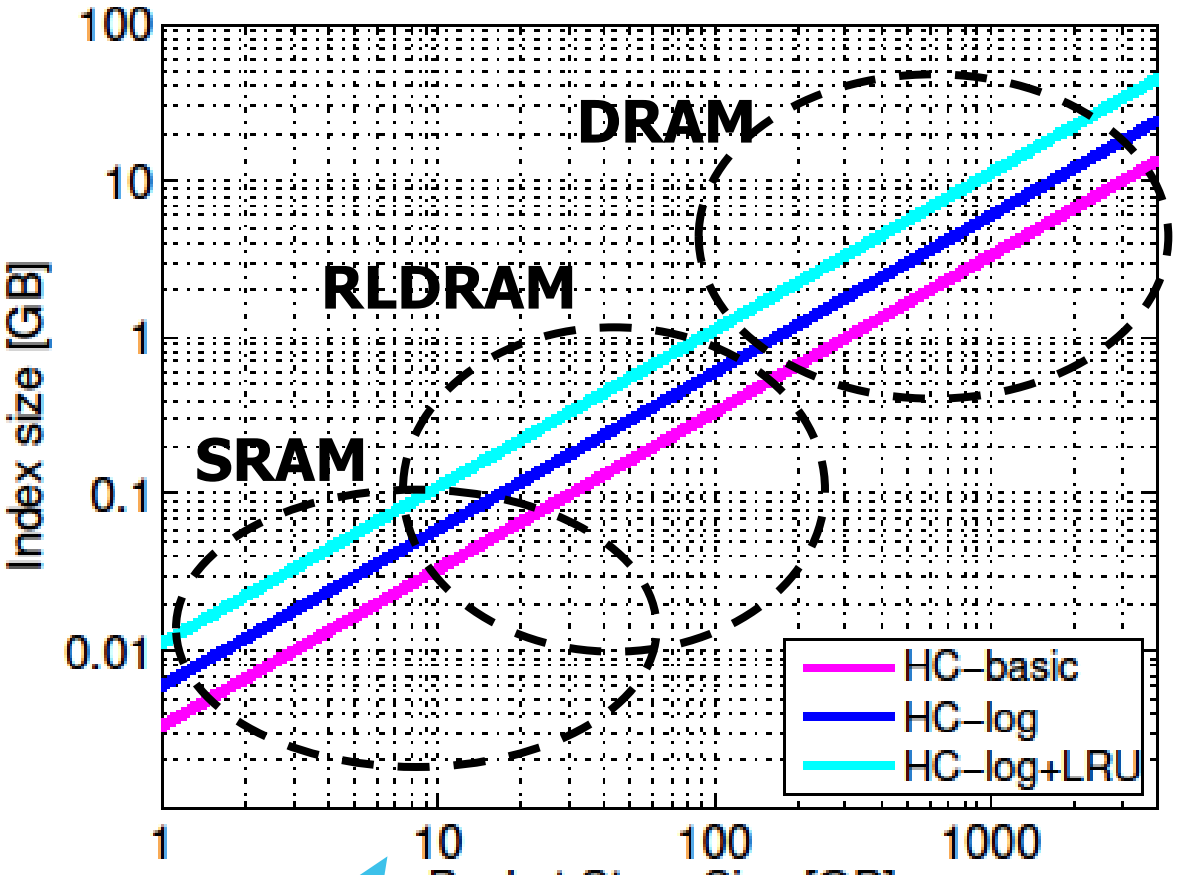
- Design guidelines
  - Minimize the number of bit per entry
  - Sustain the Request arrival rate

Max size / Access time trade-off

Technology	Access time [ns]	Max. size
SRAM	4	~210 Mb
RLDRAM	15	~2 Gb
DRAM	55	~10 GB
High-speed SSD	1,000	~10-100 TB
SSD	10,000	~1-10 TB



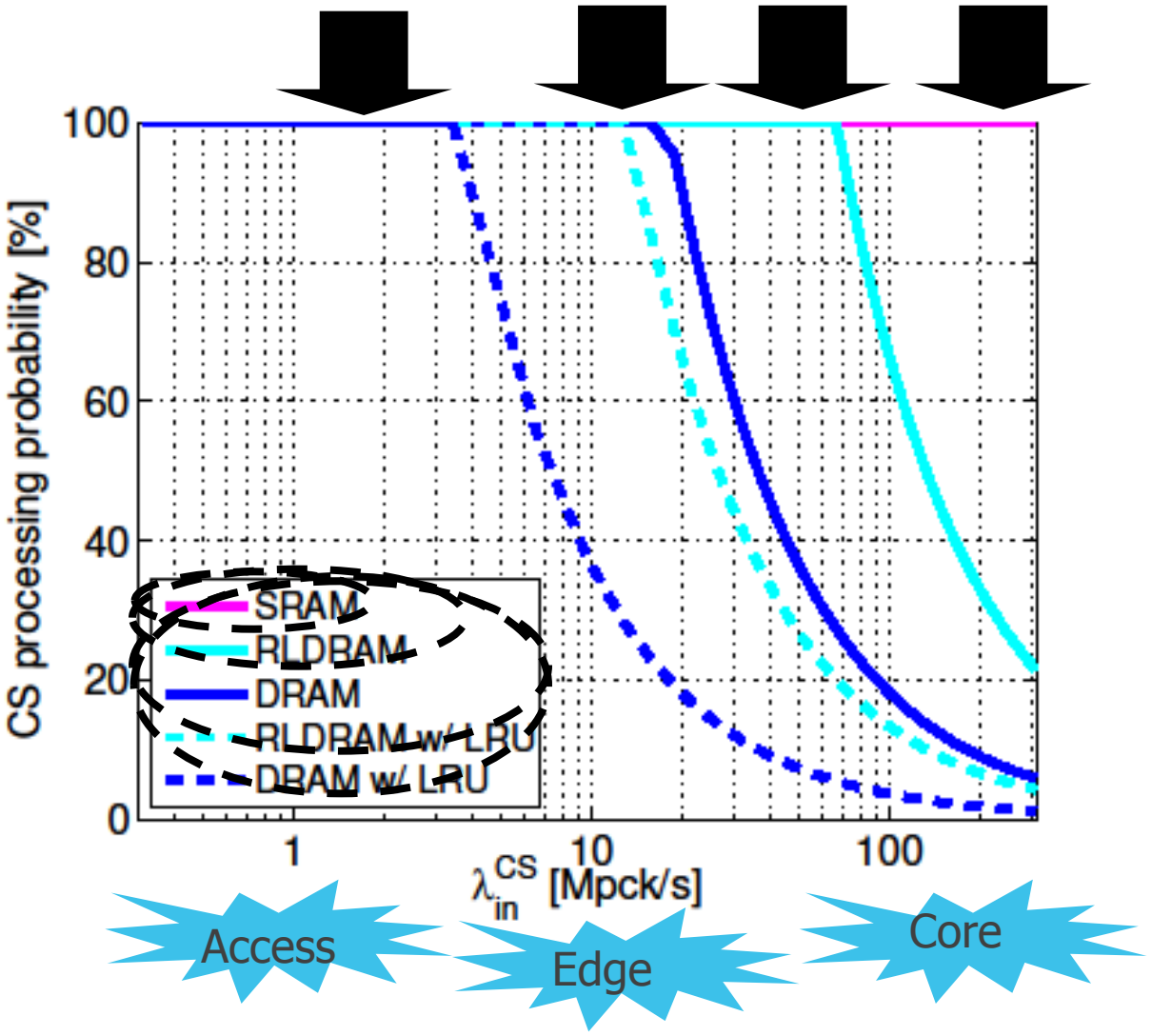
Feasibility - Cache Evaluation



DRAM

SSD

Feasibility - Cache Evaluation (cont' d)



40 Bytes Request packets and 1500 bytes Data packets, line-speed in the range 100 Mbps - 100 Gbps

Storage management techniques - Design

- **Simple cache replacement policies and coordination techniques**
  - LRU, Random, FIFO, Probabilistic caching, modulo caching, …
- **Cache sharing among applications**
  - Different applications share a common cache, and the replacement/coordination policy is applied over packets of all applications
  - Storage Partitioning (SP): each application is statically assigned a fraction of the cache, and the replacement/coordination policy is independently applied on every partition
- **Dynamic Storage management**
  - Packets are characterized by a given Time-to-Live depending on the application
  - Priority Storage Management (PSM)
  - Weighted-Fair Storage Management (WFSM)

Storage management techniques - Some results

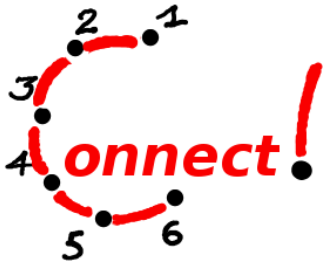
	p_{hit} LRU [%]	p_{hit} SP [%]
HTTP Web (A1)	43.0	47.6
Live Streaming (A2)	26.5	39.7
Streaming/UGC (A3)	17.6	14.9
Unknown (A4)	12.5	11.8
All content items	17.7	17.1

Packet-level observations collected at a BRAS – 1 GB cache size

- Static partitioning preserves application performance and can provide guarantee to their users. However, it may result in higher transport costs for a provider, and poor resources utilization when content has finite lifetime (TTL)
- Dynamic management can better utilize resources while controlling application performance and only slightly increasing provider costs

References & Acknowledgements

- D. Perino and M. Varvello, **A reality check for content-centric networking**, in ACM SIGCOMM workshop on information-centric networking, Toronto, Canada, 2011
- G. Carofiglio, V. Gehlen, D. Perino, **Experimental evaluation of storage management in Content-Centric Networking,** in Proc. of in Next Generation Networking and Internet symposium, IEEE ICC 2011, Kyoto, Japan
- G. Carofiglio, M. Gallo, L. Muscariello, D.Perino**, Evaluating per-application storage management in Content-Centric Networking,** under submission
- G. Carofiglio, M. Gallo, L. Muscariello, D.Perino**, Modeling data transfer in content-centric networking,** in Proc. of 23rd International Teletraffic Congress, ITC23 San Francisco, CA, USA, 2011



Work partially funded by the French national research agency (ANR), CONNECT project. http://www.anr-connect.org/

www.alcatel-lucent.com