Evaluation of Replica Placement and Retrieval Algorithms in Self Organizing CDNs

Jan Coppens, Tim Wauters, Filip De Turck, Bart Dhoedt and Piet Demeester

IFIP/IEEE International Workshop onSelf-Managed Systems & Services (SELFMAN)

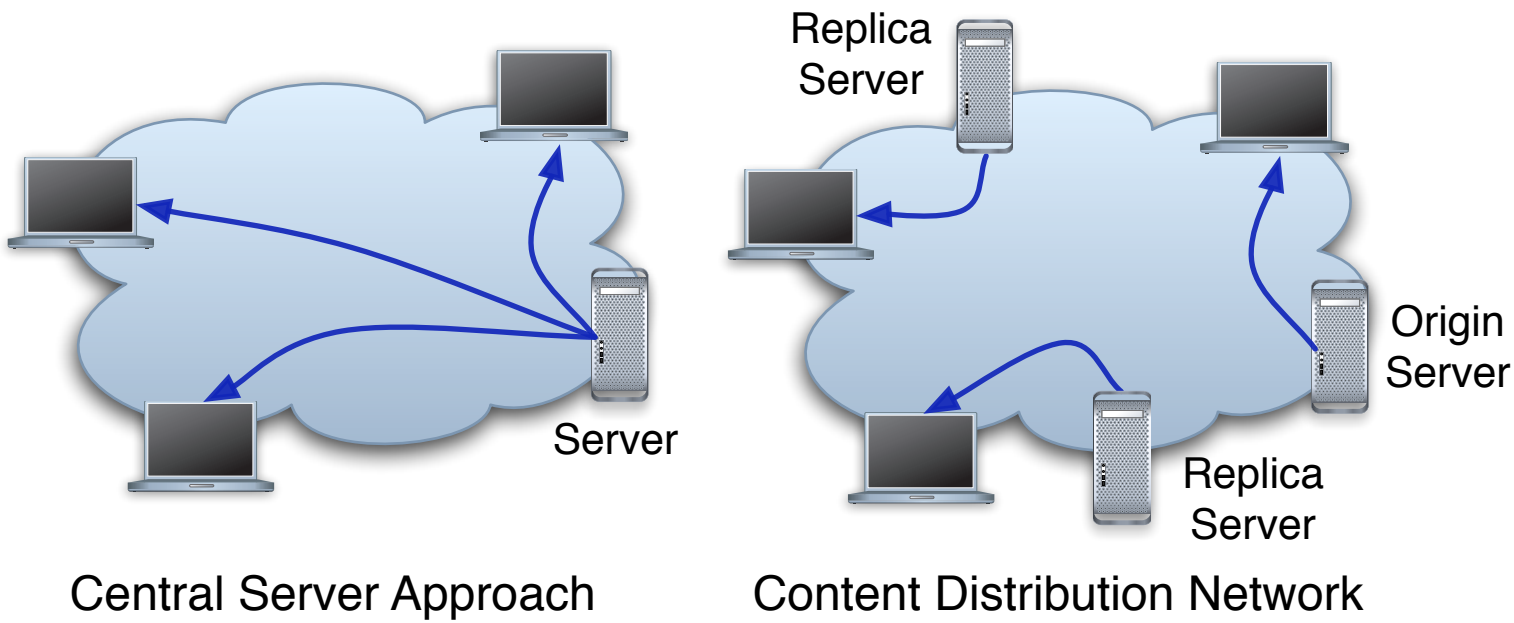
UNIVERSITEIT GENT

Overview

- CDN Architecture
- Replica Placement Algorithms
  - Reference solution by means of an ILP
  - General real-time placement heuristics
  - COCOA heuristic
- Evaluation of the placement algorithms
  - Complexity and scalability
  - Performance analysis of the RPA algorithms
  - Using Traffic Engineering for load balancing
- Conclusion and future work

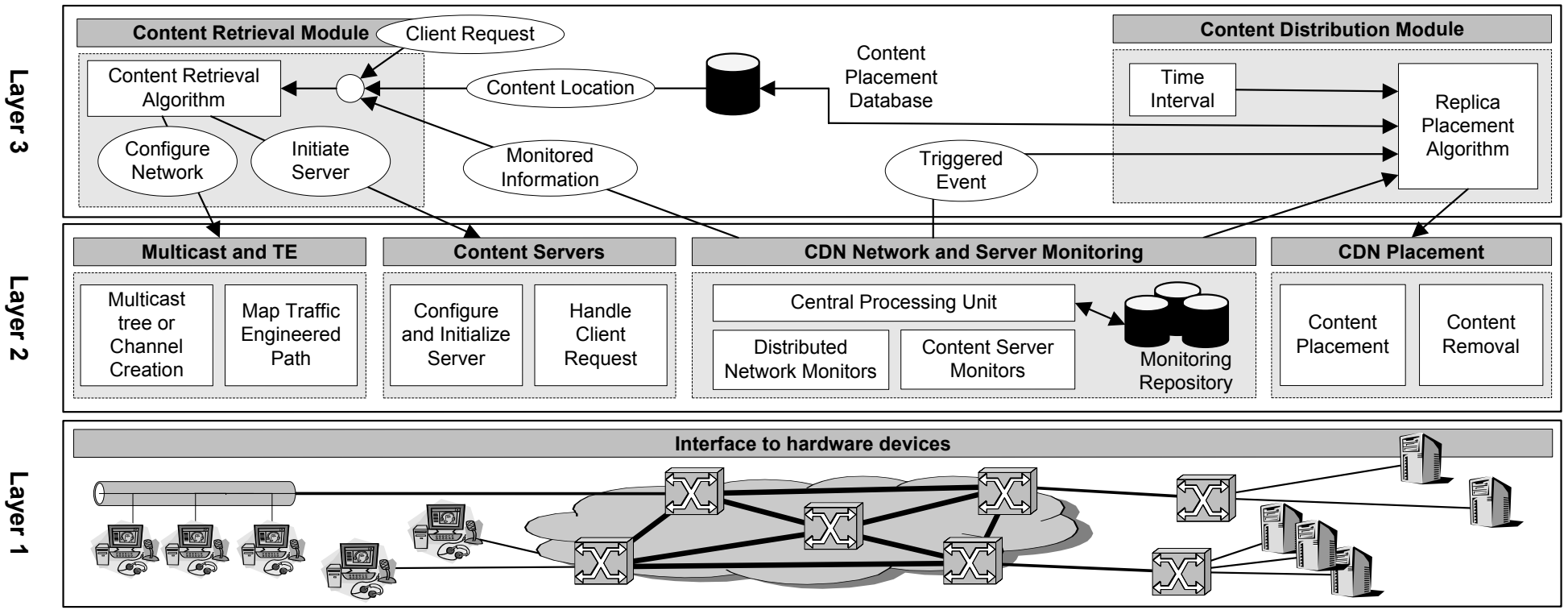
Content Distribution Networks

- Replicate and distribute the content to the edges of the network
- Increase availability and throughput
- Decrease end-to-end delay and packet loss
- Focus on the delivery of video streams



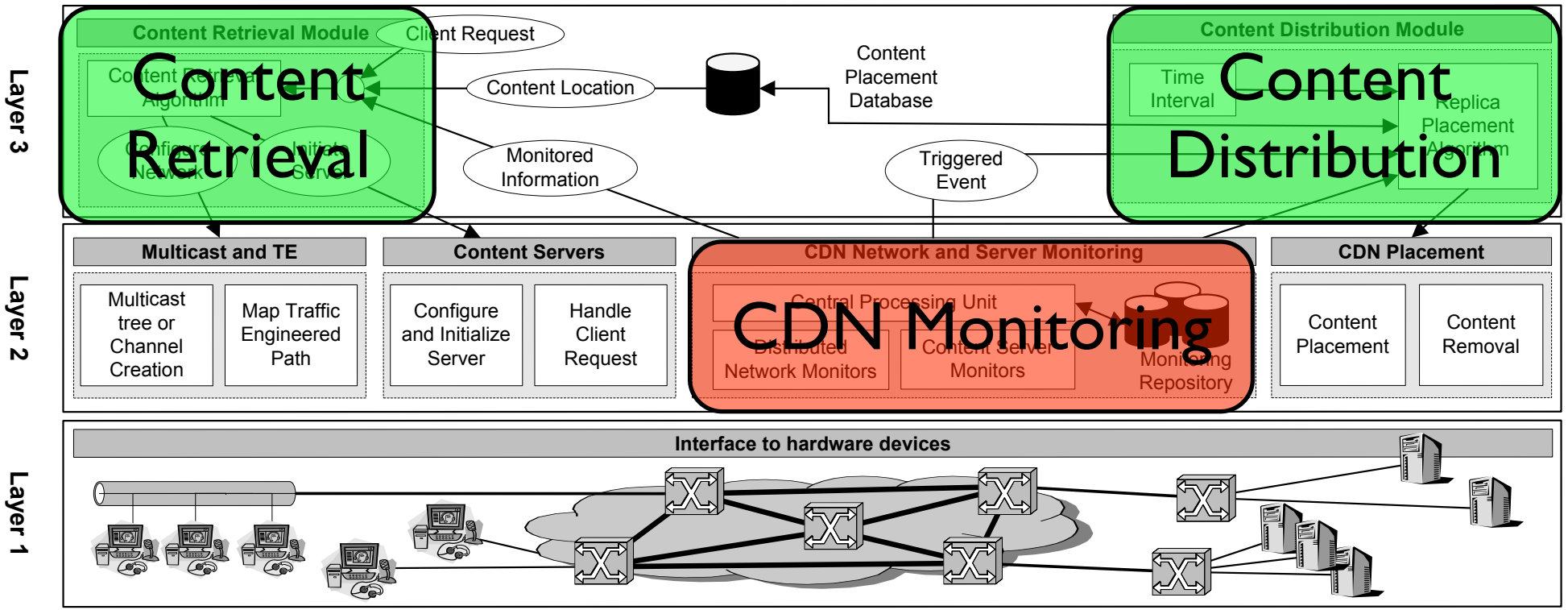
CDN Architecture

- Layered architecture for a content distribution network
- Consists of multiple functional modules



CDN Architecture

- Layered architecture for a content distribution network
- Consists of multiple functional modules



Replica Placement Algorithms

- Reference solution by means of an ILP
  - Determines the optimal placement of a static request distribution
  - Evaluation off-line (NP-complete)
- General real-time RPA heuristics
  - Periodically replaces content in the CDN
  - Evaluation on-line

Replica Placement Heuristics

- Random replica placement
- Popularity algorithms (parallel for each server)
  - Popularity Local (pop-L) - local content popularity
  - Popularity Global (pop-G) - global content popularity
- Greedy algorithms (sequential execution for each content position)
  - Greedy Single (gre-S) - cost of retrieving from origin
  - Greedy Global (gre-G) - cost from other servers
  - Greedy All (gre-A) - cost of all streams in CDN

COCOA RPA

- COCOA: Co-Operative Cost Optimization Algorithm
- Requires the aid of the Content Retrieval module
  - CR module determines the profit of available content or cost of missing content (real-time)
  - The COCOA RPA uses this information to make its placement decision (through monitoring module)
- Hybrid algorithm
  - Centralized content retrieval algorithm (also used with other RPA algorithms, but more intelligent)
  - Distributed replica placement

RPA Complexity

- Compare the computational complexity of the RPA heuristics
- COCOA has the same complexity as popularity local, but uses more accurate information

RPA	Requests	Topology	Process	Complexity
Random	None	None	Distributed	$O(C_s S)$
Pop-L	Local	None	Distributed	$O(C_s SF)$
Pop-G	Global	None	Hybrid	$O(C_s SF)$
Gre-S	Local	Origin	Distributed	$O(C_s SF)$
Gre-G	All	Entire	Centralized	$O(C_s S^3 F)$
Gre-A	All	Entire	Centralized	$O(C_s S^4 F)$
COCOA	CR Module	None	Hybrid	$O(C_s SF)$

RPA Complexity

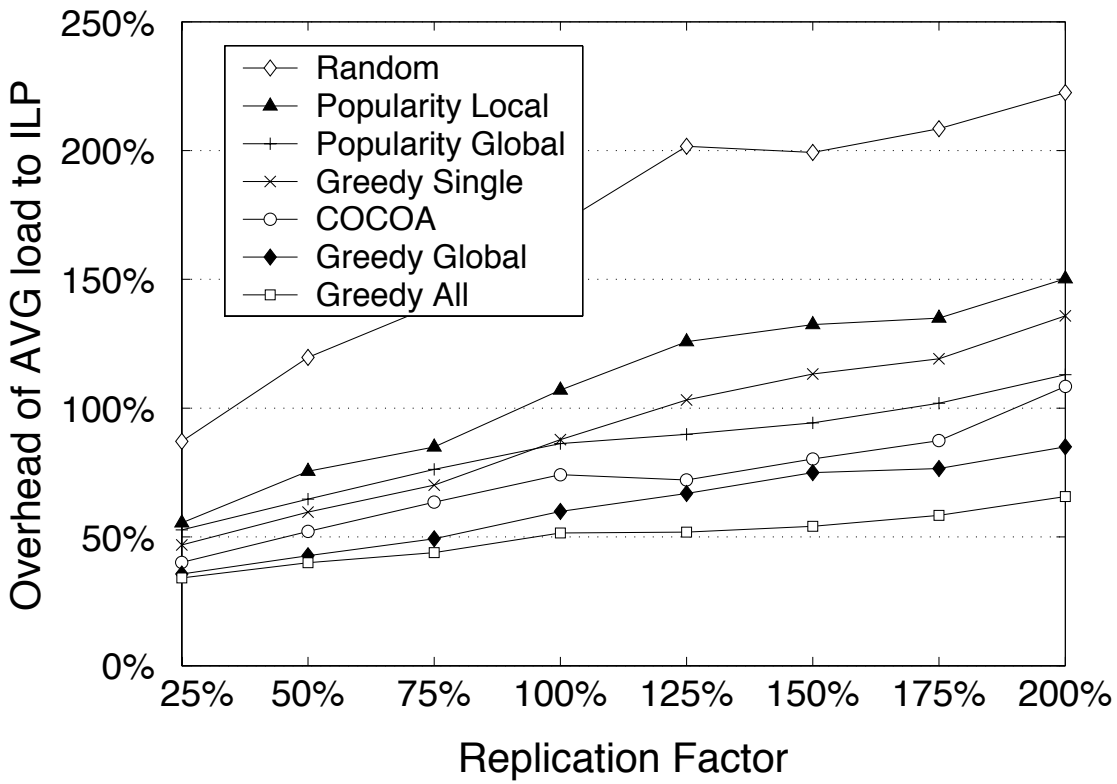
- Compare the computational complexity of the RPA heuristics
- COCOA has the same complexity as popularity local, but uses more accurate information

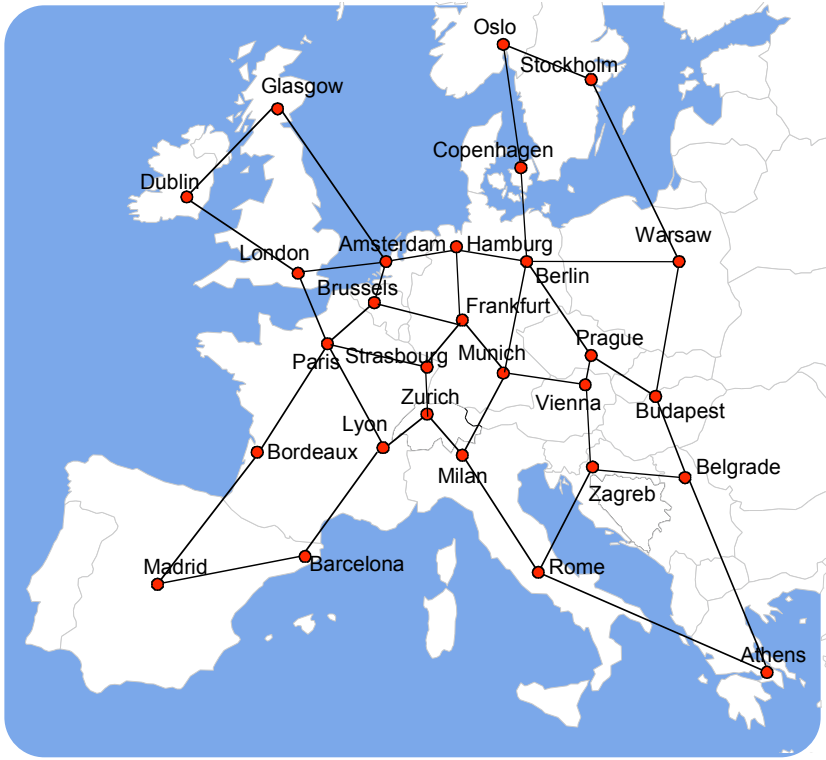
Parallelism

RPA	Requests	Topology		Complexity
Random	None	None	Distributed	$O(C_s)$
Pop-L	Local	None	Distributed	$O(C_sF)$
Pop-G	Global	None	Hybrid	$O(C_sF)$
Gre-S	Local	Origin	Distributed	$O(C_sF)$
Gre-G	All	Entire	Centralized	$O(C_sS^3F)$
Gre-A	All	Entire	Centralized	$O(C_sS^4F)$
COCOA	CR Module	None	Hybrid	$O(C_sF)$

Performance Analysis

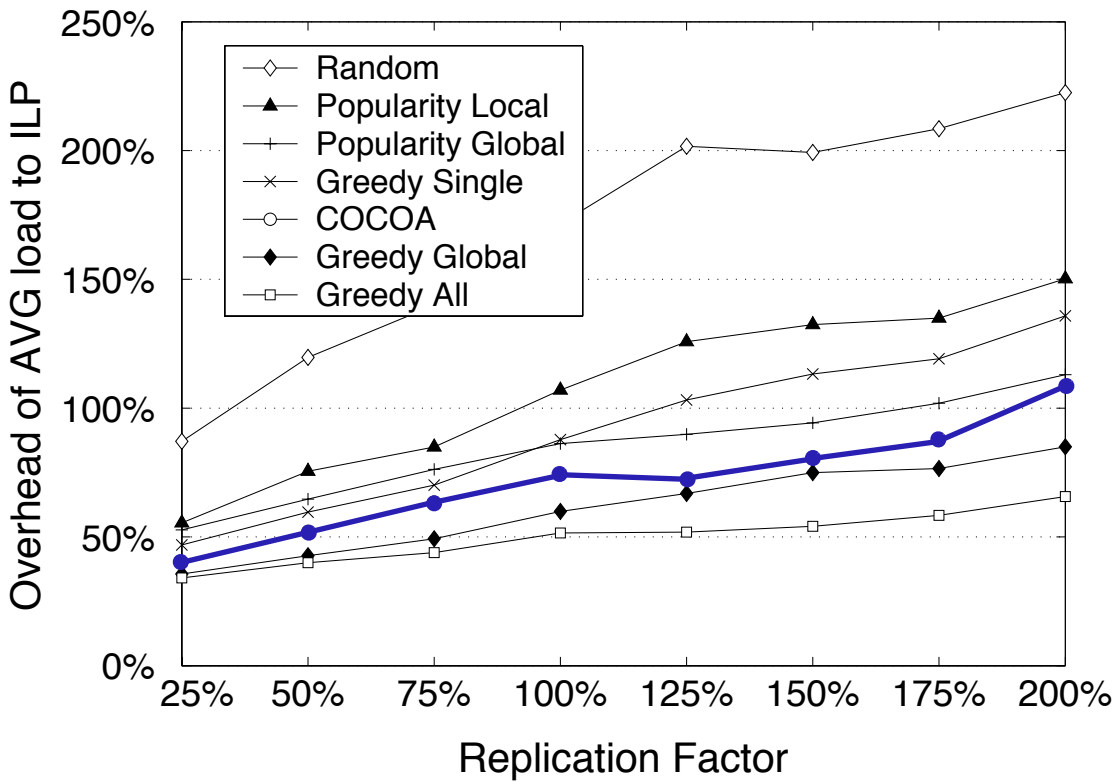
- Overhead of the average load of the algorithms to the ILP solution
- COCOA is better than pop-L and close to the gre-G placement

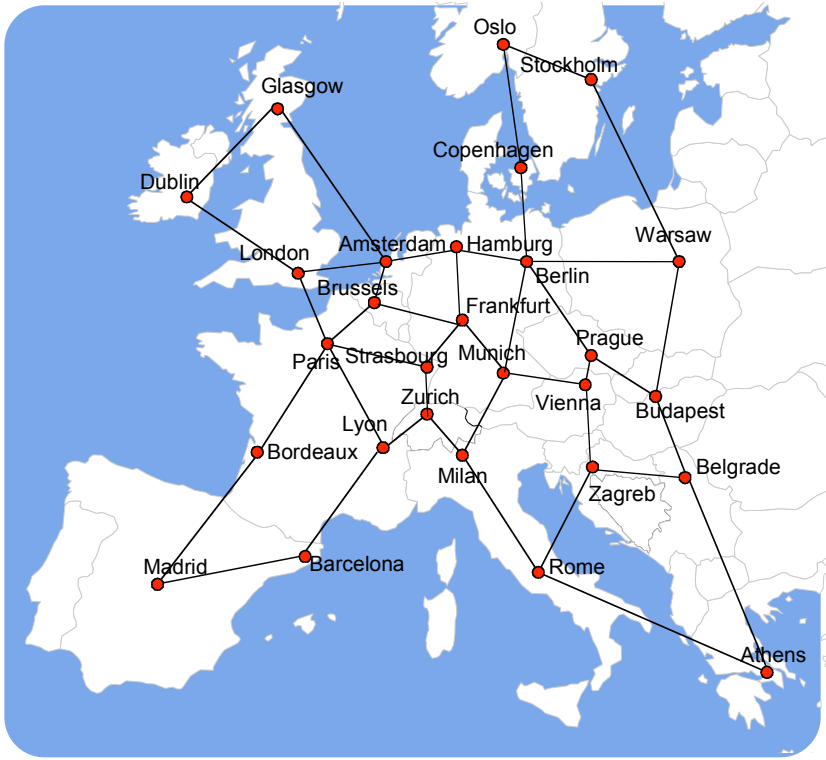




Performance Analysis

- Overhead of the average load of the algorithms to the ILP solution
- COCOA is better than pop-L and close to the gre-G placement

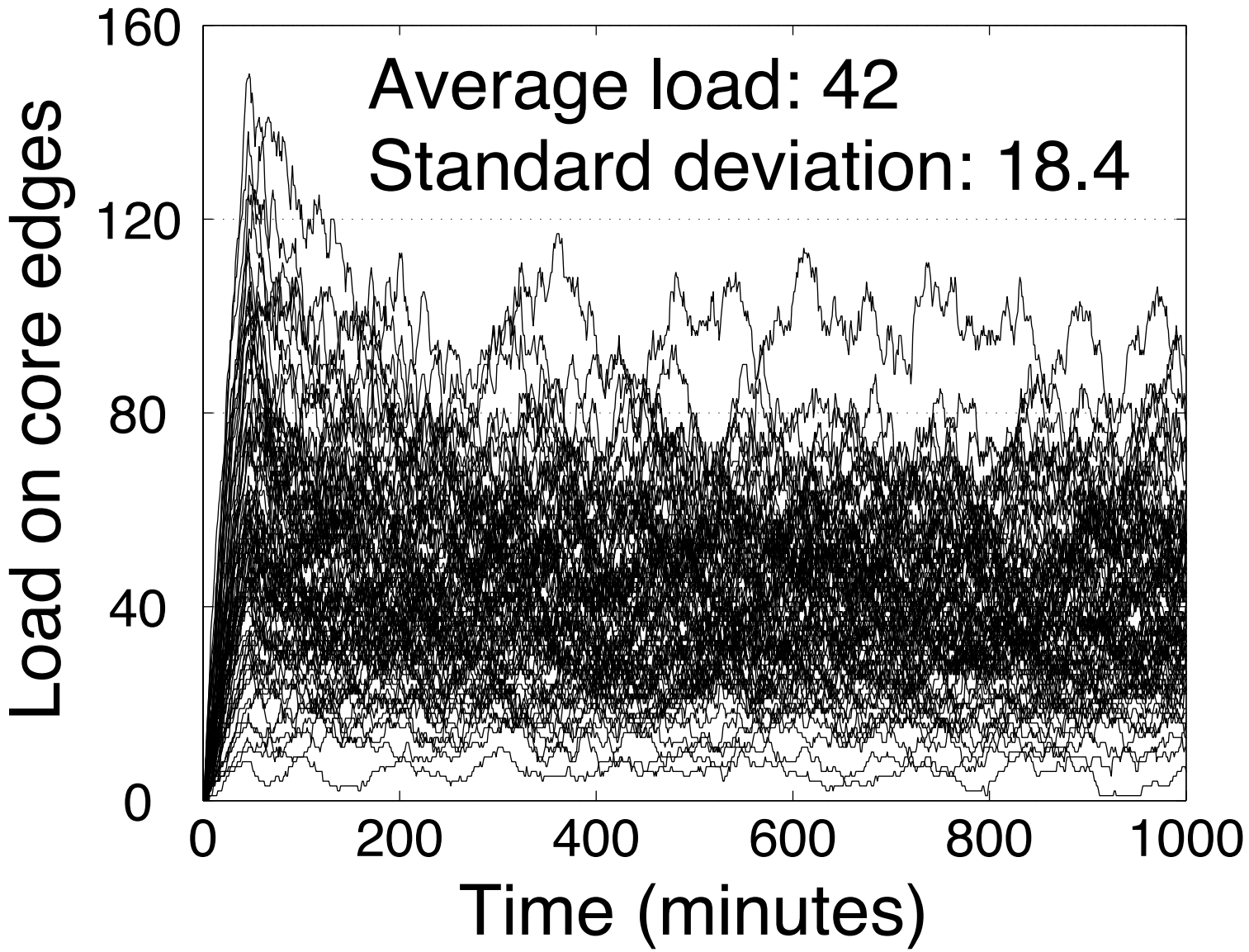




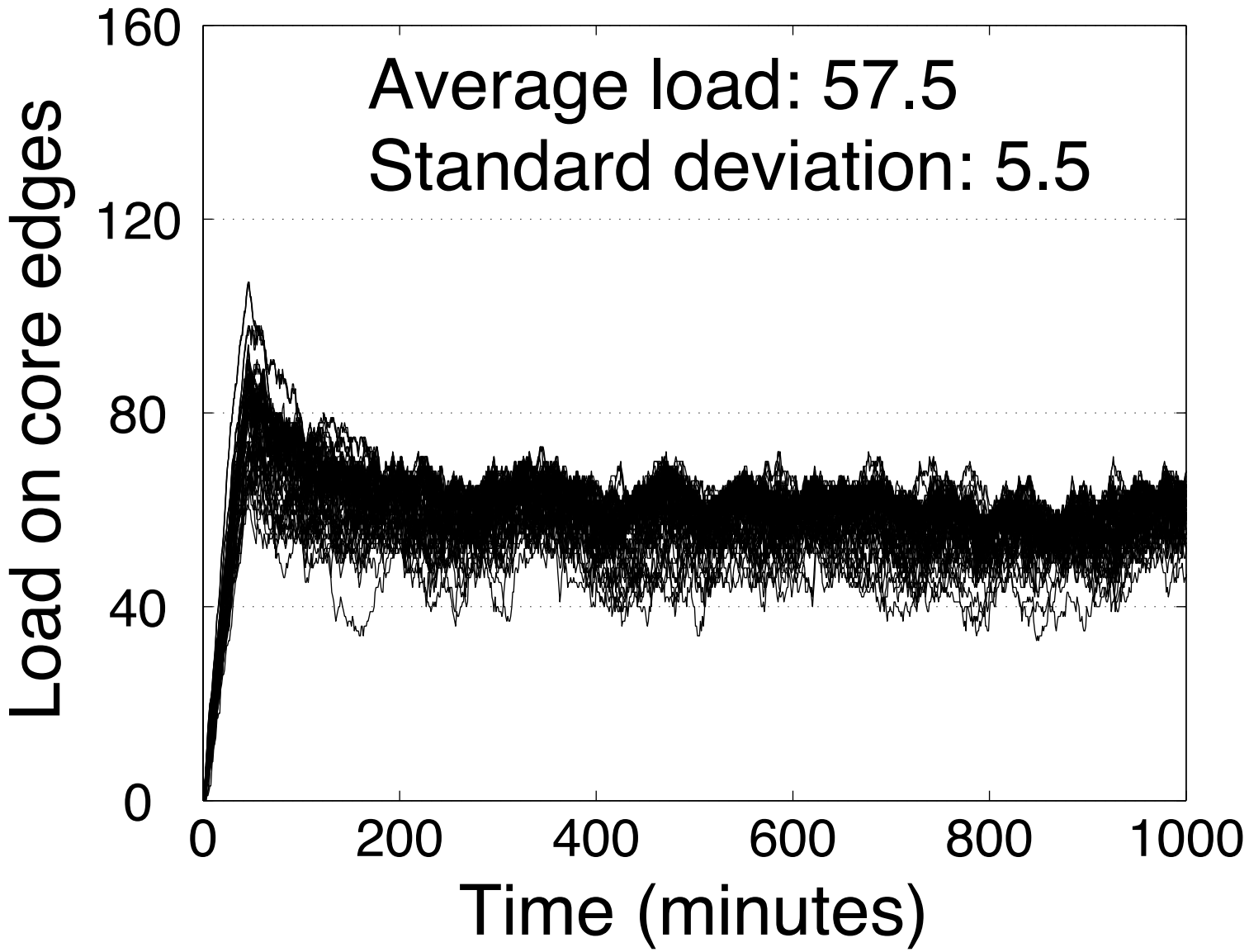
Traffic Engineering for Load Balancing

- Because of asymmetric topology and request distribution, the load is distributed unevenly over the network
- Can cause congestion in certain parts of the network (e.g. during a flash crowd)
- Traffic Engineering can be used to spread the flows over the entire CDN
  - Proactive in order to off-load core edges
  - Reactive in order to route flows round congested bottlenecks

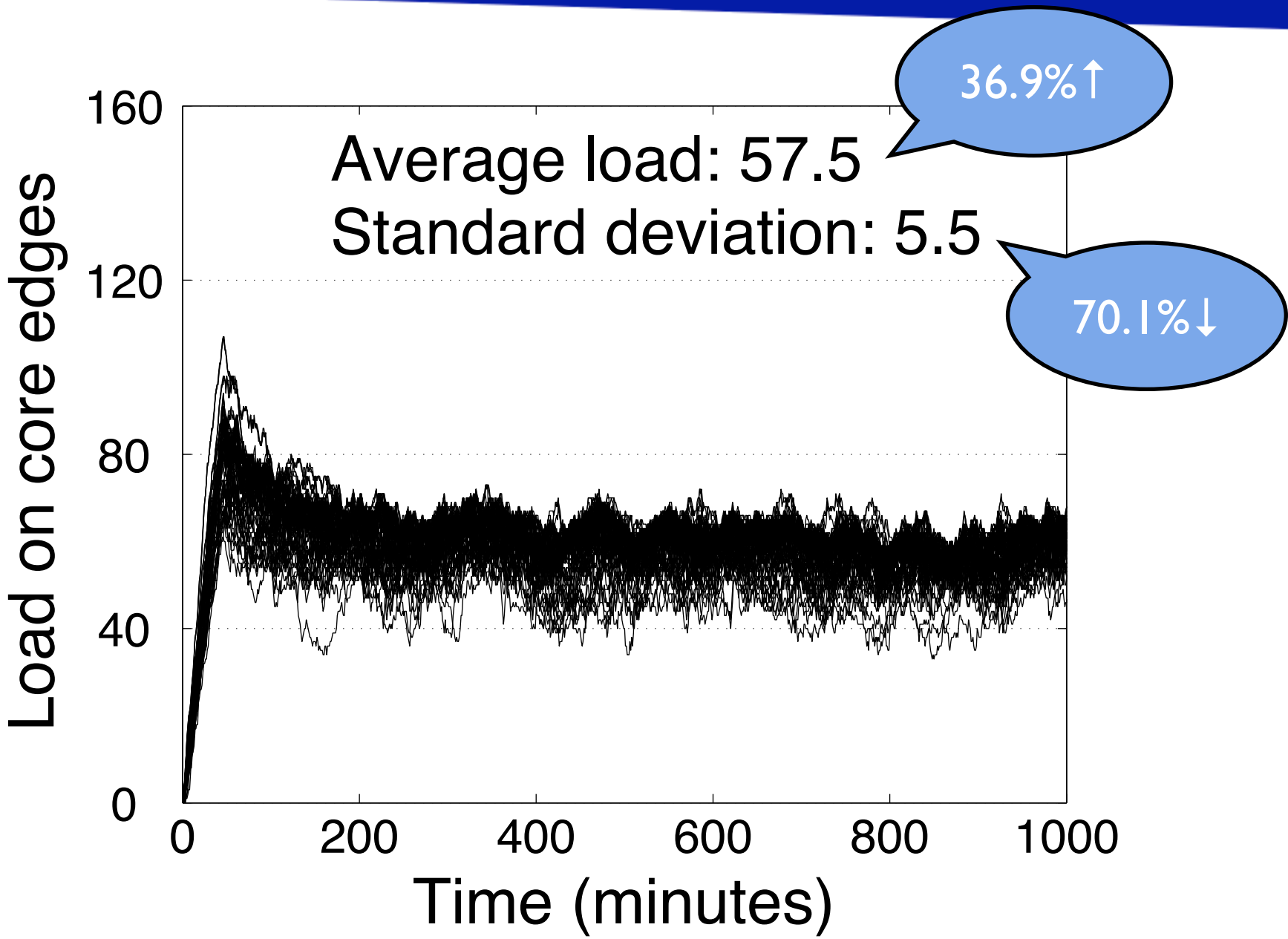
Proactive Traffic Engineering (1)



Proactive Traffic Engineering (1)

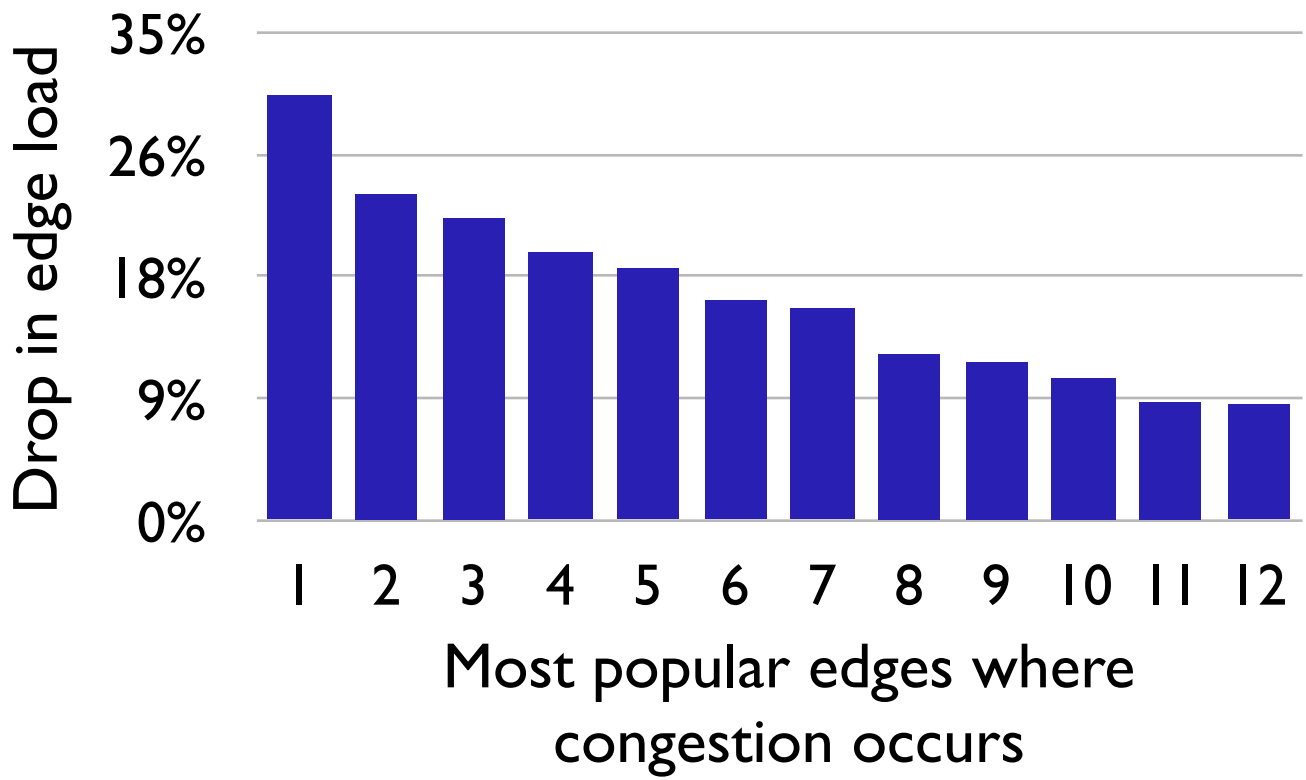


Proactive Traffic Engineering (1)



Proactive Traffic Engineering (2)

- Using Traffic Engineering the average load increases, but the standard deviation drops
- The load on the core edges decreases:



Conclusion and Future Work

- Novel hybrid CDN architecture and COCOA RPA
- COCOA placement algorithm
  - Nearly as scalable as popularity local
  - Close to the performance of greedy global
- Traffic engineering used to off-load the core edges
- Influence of replacements on the load of the network?
- Frequency of RPA executions?
- How can we make content retrieval distributed?