Economics of advanced HTTP Caching in eNodeBs

Chris Drechsler, Gerd Windisch

Chair for Communication Networks

Chemnitz University of Technology

Outline

- Introduction
- Improved HTTP Caching Method
- Cost Model
- Results of techno-economic Analysis
- Summary

Introduction

- Key challenge for mobile network operators:
  - tremendous increase in mobile data traffic (dominant protocol: HTTP)

- Solution for HTTP traffic reduction in RAN and core:
  → **Caching in eNodeBs**

- Advantages:
  - no access to GTP-tunnel (S1-interface) required
  - access transport cost savings (compared to centralized caching at S/P-GW)
  - QoS/QoE improvement
- Disadvantages:
  - small population size (at eNodeBs) → low hit rate (caching efficiency)
  - higher cost for distributed cache deployment

Introduction

- Motivation:
  - increase caching efficiency → improved HTTP caching method
  - evalute cost/efficiency tradeoff of the improved caching method in an eNodeB application scenario

Outline

- Introduction
- **Improved HTTP Caching Method**
- Cost Model
- Results of techno-economic analysis
- Summary

Improved HTTP Caching - HTTP Caching Efficiency

- Estimated efficiency potential of HTTP caching:
  - up to 68% HTTP traffic reduction (byte hit rate, BHR)
- Caching efficiency observed today:
  - only 10-20% (byte hit rate)
- Reasons for low caching efficiency:
  - difficult detection of duplicate payloads, example:

    http://s1.videoportal.com/PopularVideo.webm?userid=1111 vs.

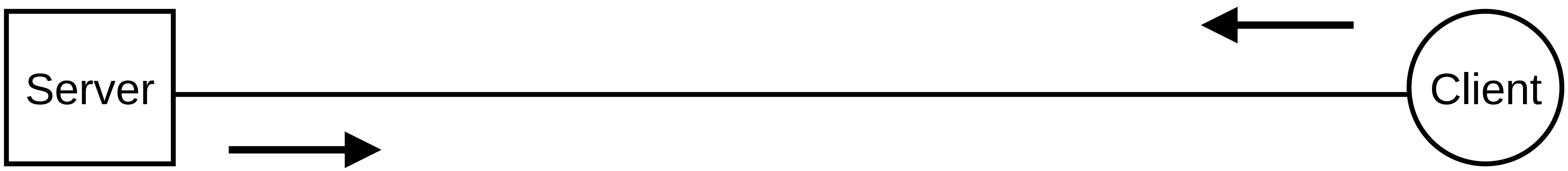
    http://s2.videoportal.com/PopularVideo.webm?userid=2222
  - personalization
  - explicit suppression of caching by content producers
  - too small cache sizes

- → new caching method to improve the caching efficiency

Improved HTTP Caching - Basic Concept

- HTTP header field extension:

```
GET /videos/PopularVideo.webm HTTP/1.1
Host: example.com
```



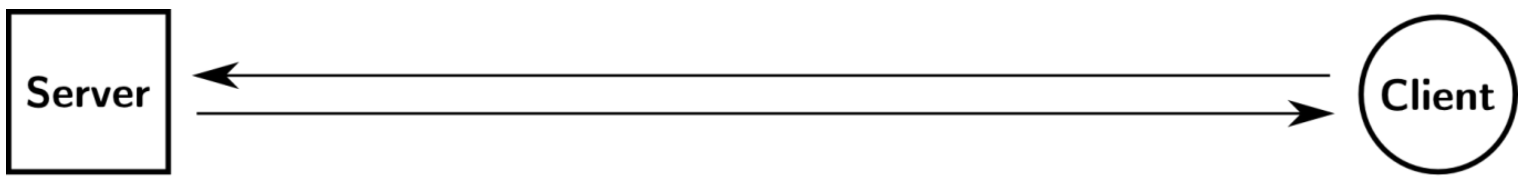

```
HTTP/1.1 200 OK
Date: Fri, 11 Nov 2011 11:11:11 GMT
Cache-NT: sha256=7ab53f24d8c96d1cc87452a6b113 …

<HTTP Body>
```

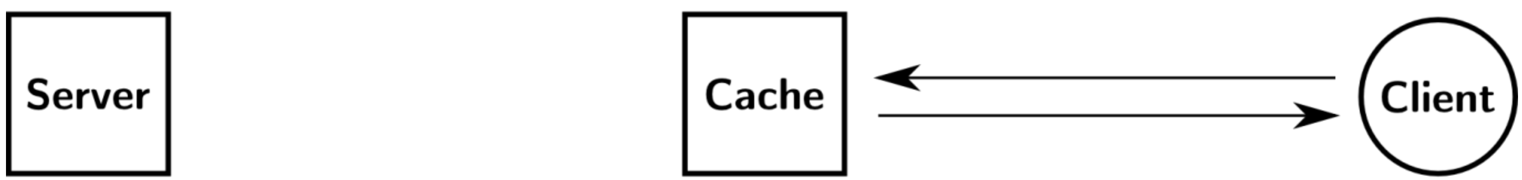
Improved HTTP Caching - Basic Concept

- Modified cache operation:

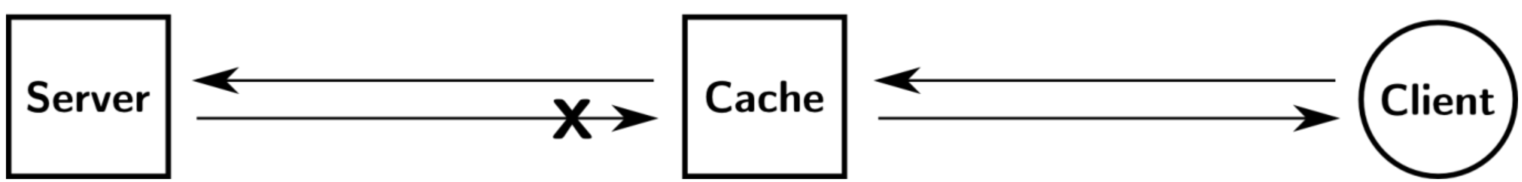
Without Caching:



Traditional Caching (example: cache hit):

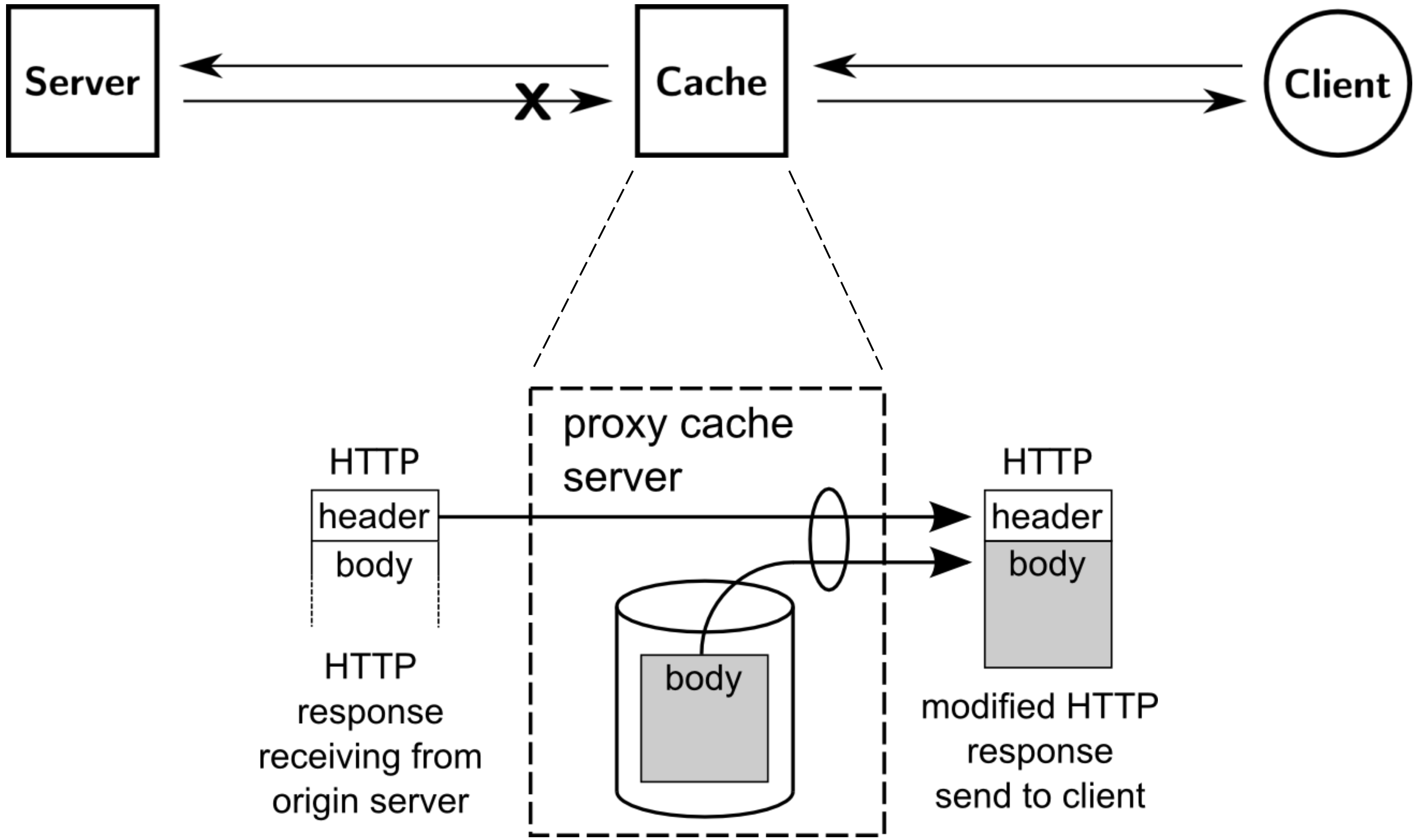


Modified Caching (example: cache hit):



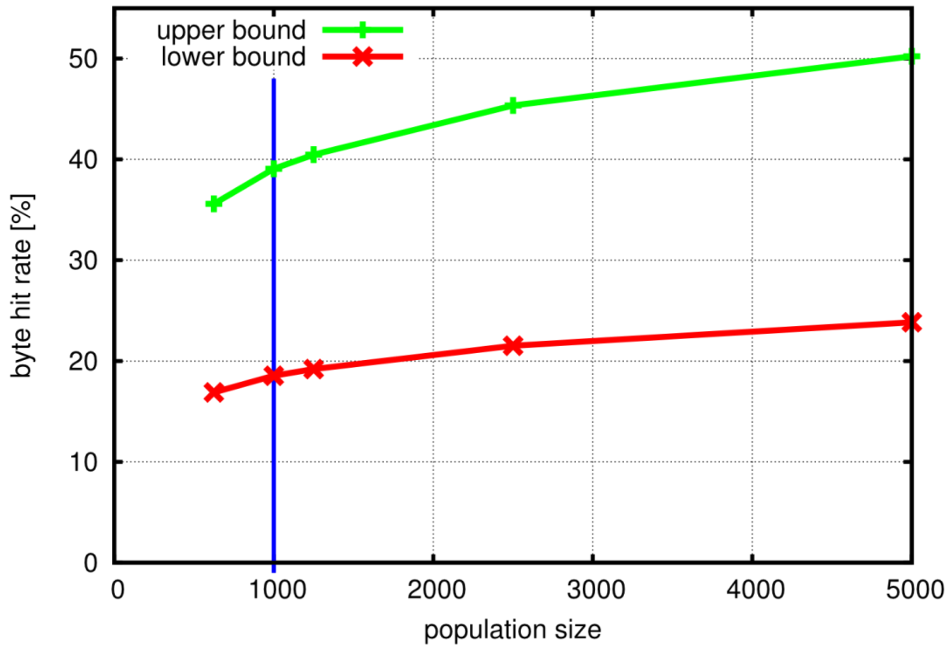
Improved HTTP Caching - Basic Concept

- Modified cache operation:



Improved HTTP Caching - Efficiency Evaluation

- Evaluation: byte hit rate (modified caching method) vs. population size
  - simulation model based on study of Erman et al.
  - BHR upper bound: optimistic HTTP caching scenario - some header fields (e.g. cache-control, cookies) are ignored by the cache
  - BHR lower bound: pessimistic HTTP caching scenario - all header fields are strictly considered by the cache



Outline

- Introduction
- Improved HTTP Caching at eNodeB sites
- **Cost Model**
- Results of techno-economic Analysis
- Summary

Cost Model - Assumptions

- Access network:
  - tree topology
  - microwave links:
    - assumption: discrete capacity steps: link type1 - link type 4
  - multiple root nodes (Point of Connect, POC) - connection to wired backhaul
    - assumption: all POCs of same size/capacity
- eNodeBs:
  - LTE eNodeBs with 3 cell sectors (10 MHz carrier) (NGMN Alliance model)
  - traffic per eNodeB (NGMN Alliance model):
    - busy hour traffic: 73.2 Mbit/s
- HTTP caches:
  - HTTP cache integrated in **each** eNodeB
  - assumption: sufficient cache size (fixed)

Cost Model - Cost Components

- Access transport link cost:
  - variable cost - cost depends on link type i: $C_T^i = 2^{(i-1)E_{cost}} C_T^1$ i=1, ...4
  - cost of link type 1: $C_T^1 = F_T C_C$
  - E_{cost} used as a scaling parameter (cost step width)
- POC cost:
  - fixed cost (all POCs of same size/capacity): $C_{POC} = F_{POC} C_C$
- Cache cost:
  - fixed cost (due to fixed cache size): C_C
- eNodeB cost:
  - not considered because not relevant for techno-economic analysis of caching benefits
- Remark:
  - the cost factors F_T,F_{POC} are used to express the transport link (type 1) cost C_T^1 and the POC cost C_{POC} relative to the cache cost C_C

Cost Model - Total Costs

- Total cost (CAPEX) C of access network with caches in eNodeBs:

$$C = C_C^S + C_T^S + C_{POC}^S$$

∑ cache cost ∑ transport cost ∑ POC cost

where:

$$C_C^S = N_{eNB} C_c$$

$$C_T^S = \sum_{i=1}^{L} N_T^i C_T^i$$

$$C_{POC}^S = N_{POC} C_{POC}$$

N_{eNB} : number of eNodeBs

N_T^i : number of links of type i

N_{POC}: number of POCs

derived from access network dimensioning

Cost Model - Cost Savings (compared to non-Caching)

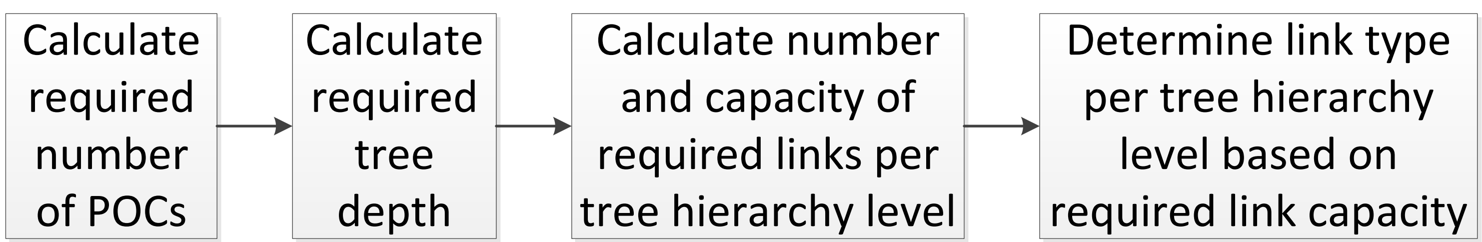
- Cost savings S_C of caching solution compared to non-caching case:

$$S_C = 1 - \frac{C_{with\ Caches}}{C_{without\ Caches}}$$

$$S_C = 1 - \frac{N_{eNB} + \sum N_{T,wC}^{i}\, 2^{(i-1)E_{cost}} F_T + N_{POC,wC} F_{POC}}{\sum N_{T,woC}^{i}\, 2^{(i-1)E_{cost}} F_T + N_{POC,woC} F_{POC}}$$

Cost Model - Access Network Dimensioning Algorithm

- Access network dimensioning algorithm:
  - heuristic algorithm for tree-shaped access networks based on dimensioning guidelines from NGMN Alliance
- Input:
  - number of eNodeBs: N_{eNB} (= 10000)
  - traffic per eNodeB: (73,2 Mbit/s)
  - aggregation ratio from one hierarchy level to the next higher: AR
- Output:
  - $N^{i}_{T,wC}$, $N_{POC,wC}$, $N^{i}_{T,woC}$, $N_{POC,woC}$
- Algorithm main steps:



Outline

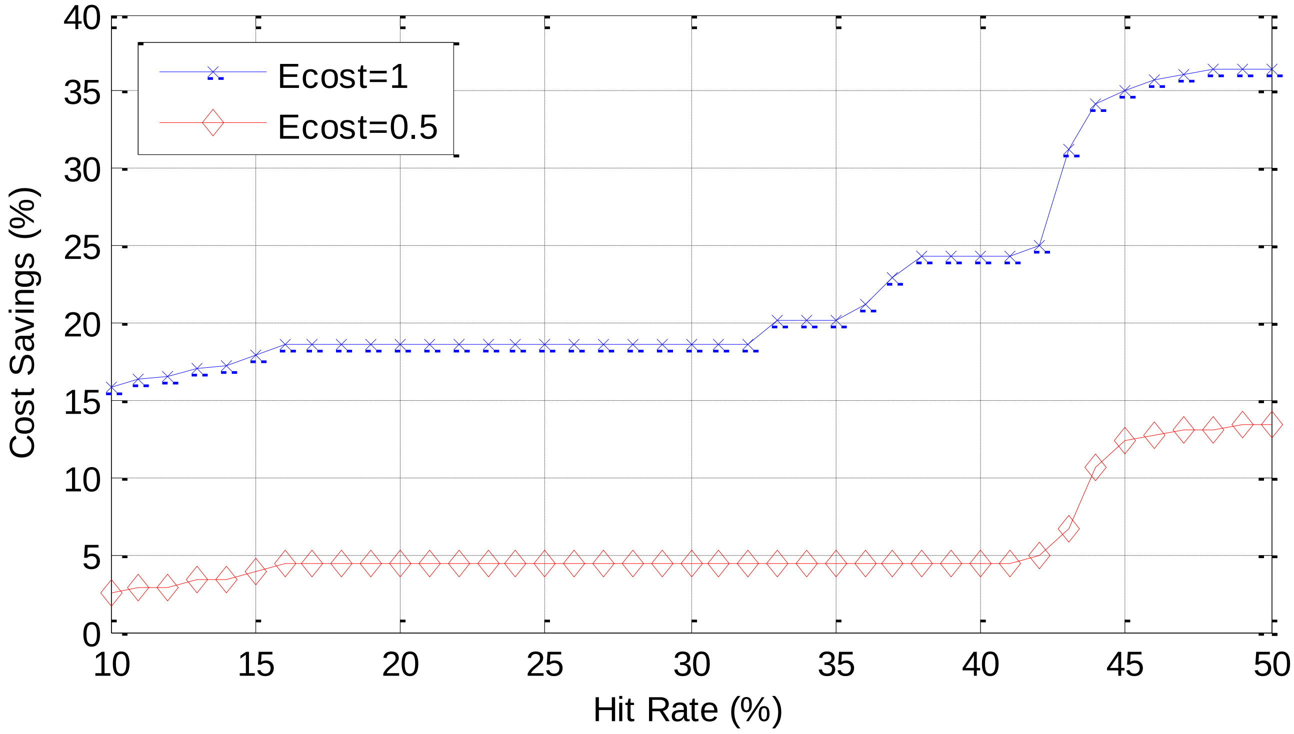
- Introduction
- Improved HTTP Caching at eNodeB sites
- Cost Model
- **Results of techno-economic Analysis**
- Summary

Results - Input Parameter

- Traffic demand per eNodeB (NGMN Alliance model):
  - assumption: 3 sector eNodeB, 10 MHz carrier
  - busy hour demand: 73.2 Mbit/s
- Number of eNodeBs: 10000
- Aggregation ratio from one hierarchy level to the next higher: 4
- Initial cost factor settings:
  - F_T = 10
  - F_{POC} = 100

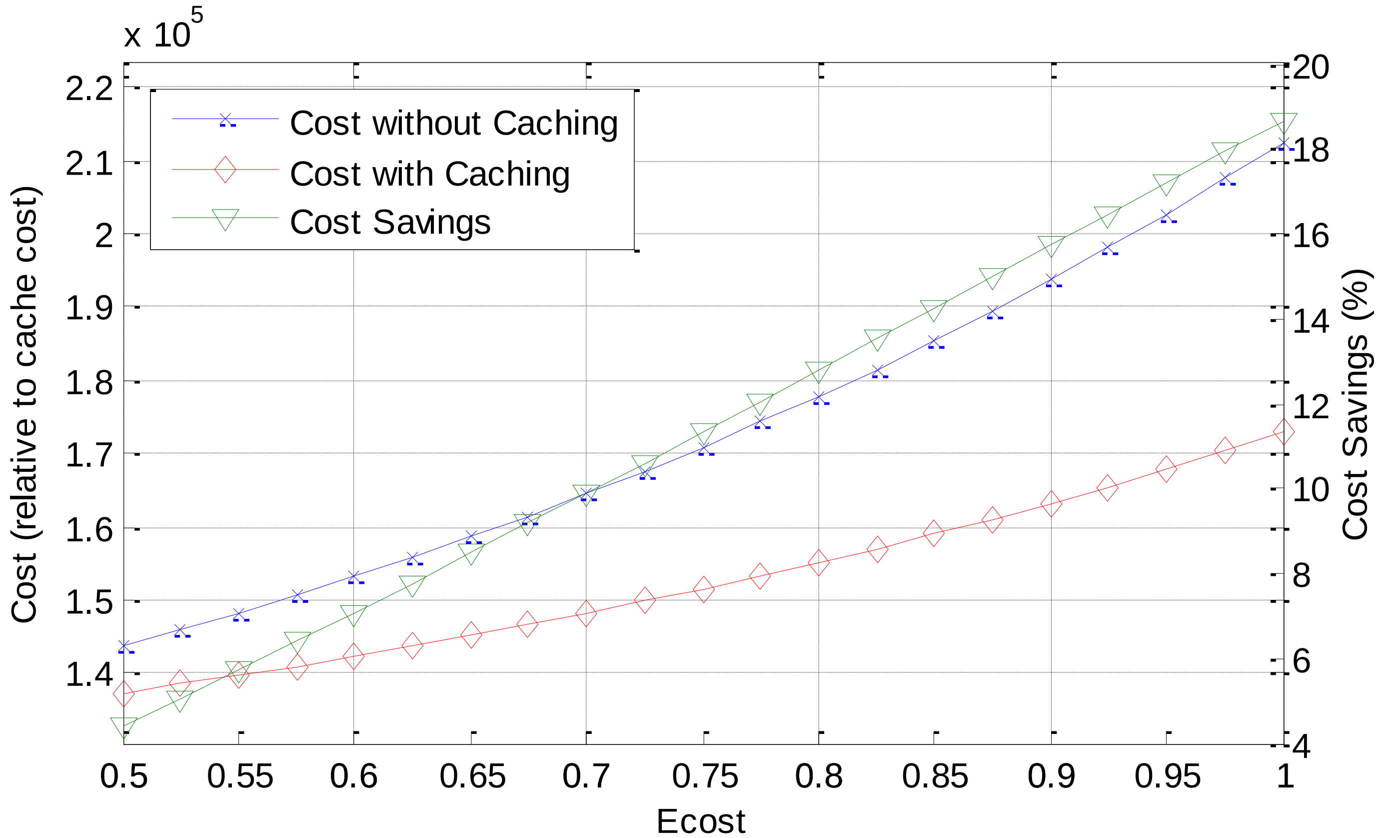
Results - Cost Savings vs. Byte Hit Rate

- Cost savings vs. byte hit rate



Results - Cost / Cost Savings vs. Ecost (BHR 20%)

- Costs and Cost savings vs. Ecost (byte hit rate 20%)



Results - Sensitivity Analysis

- Motivation: evaluation of the influence of the cost parameters on the cost savings
- 4 scenarios:
  - E_{cost} = 0.5, byte hit rate = 20%
  - E_{cost} = 0.5, byte hit rate = 40%
  - E_{cost} = 1, byte hit rate = 20%
  - E_{cost} = 1, byte hit rate = 40%
- Results:

Parameter	Parameter Variation Range	Cost Savings Sensitivity	Scenario
E_{cost}	0.5 to 1	4% - 24%	hit rate = 40%
F_T	-50% to +50%	18% - 24%	E_{cost}=1, hit rate = 40%
F_{POC}	-50% to +50%	3% - 5%	E_{cost}=0.5, hit rate = 20%

Outline

- Introduction
- Improved HTTP Caching at eNodeB sites
- Cost Model
- Results of techno-economic Analysis
- **Summary**

Summary

- Contributions:
  - method to improve the efficiency of HTTP caching
  - analysis of access network cost savings through improved caching in eNodeBs

- Key results:
  - significant cost savings through caching in eNodeBs possible
  - but:
    - modified caching method required (as traditional caching yields to low byte hitrate → not cost efficient)
    - cost savings strongly depend on parameters (F_T, E_{cost}) → careful individual cost analysis required