

Distributed cooperation schemes for document lookup in multiple cache servers

Riccardo
Lancellotti
University of Rome
"Tor Vergata"

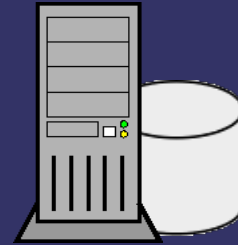
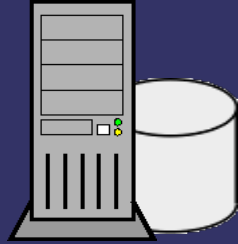
Bruno Ciciani
University of Rome
"La Sapienza"

Michele Colajanni
University of Modena and
Reggio Emilia

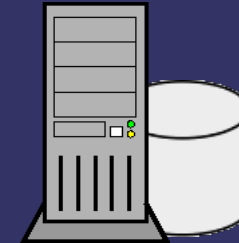
Web caching



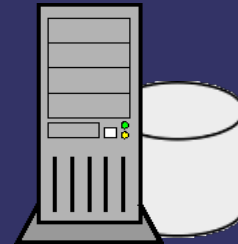
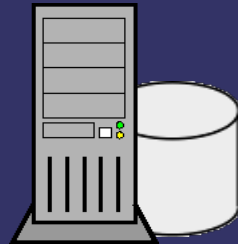
Clients



Cache servers



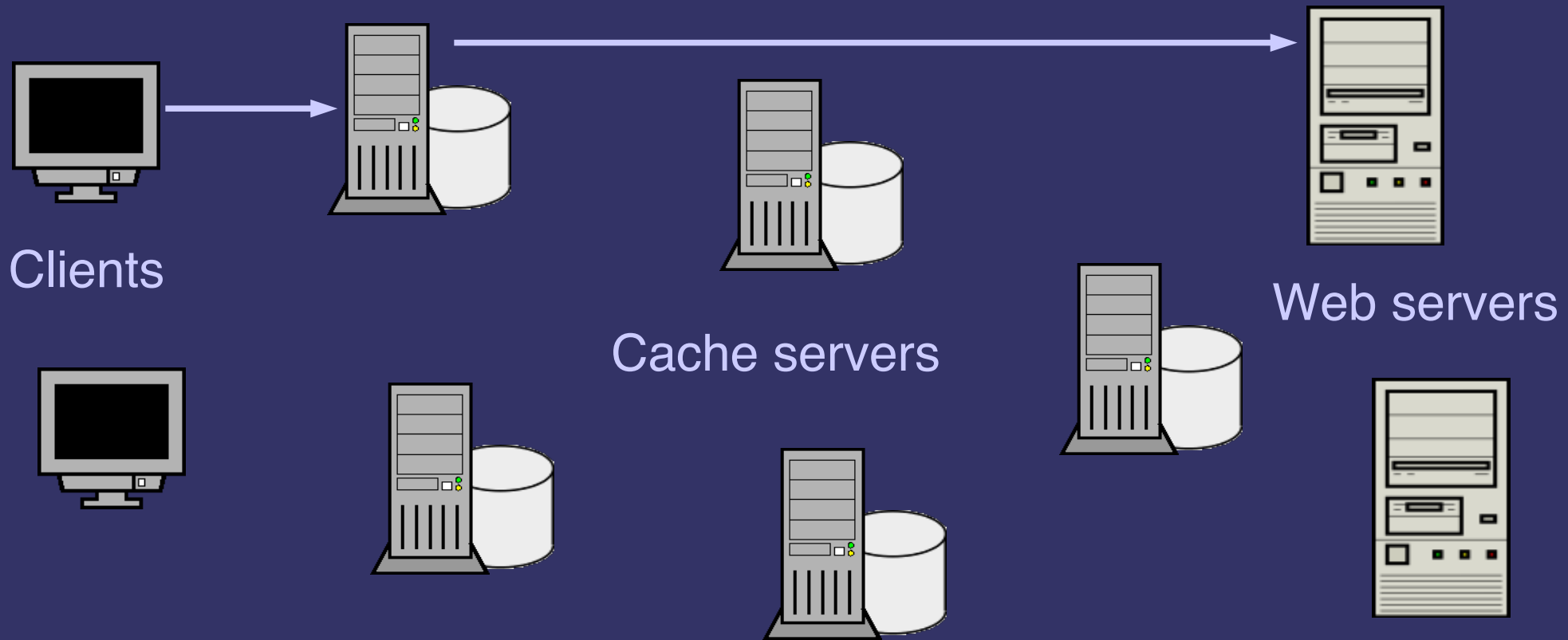
Web servers



With 25 years of Internet experience, we've learned exactly one way to deal with exponential growth: **Caching**.

[Van Jacobson]

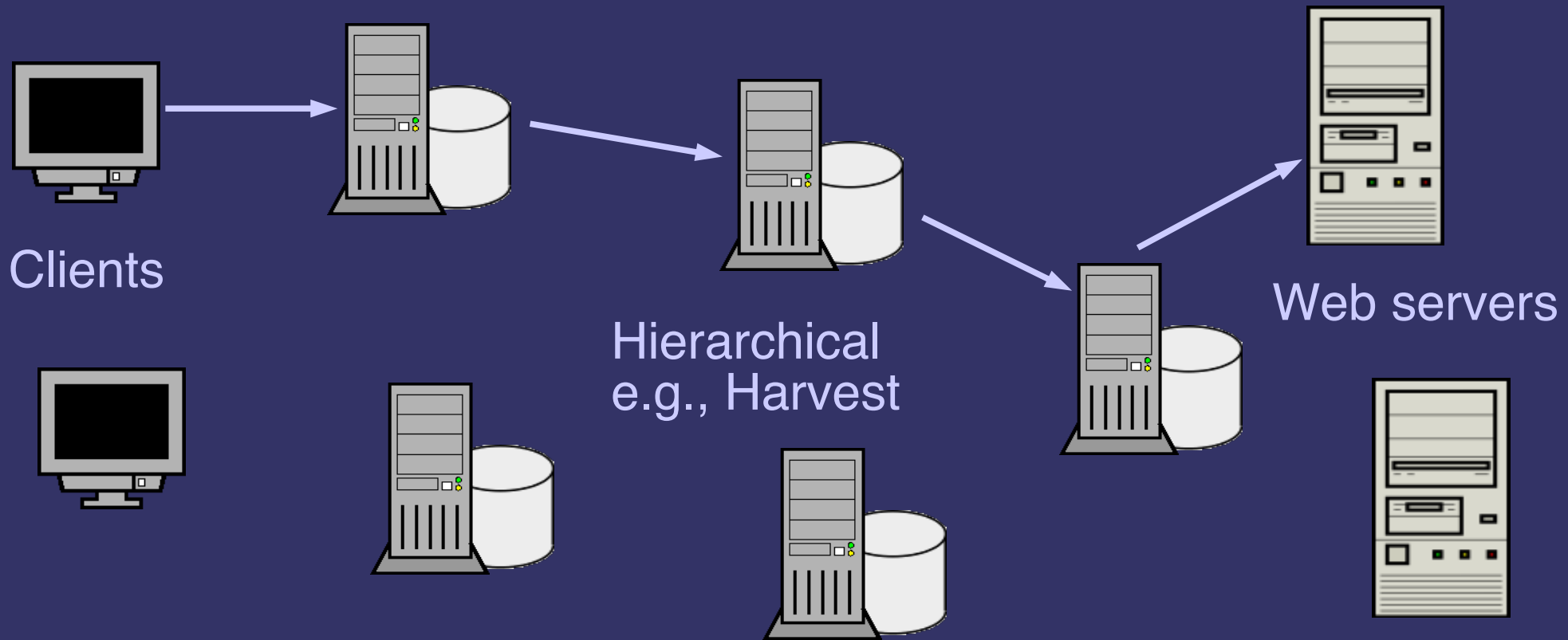
Web caching



With 25 years of Internet experience, we've learned exactly one way to deal with exponential growth: **Caching**.

[Van Jacobson]

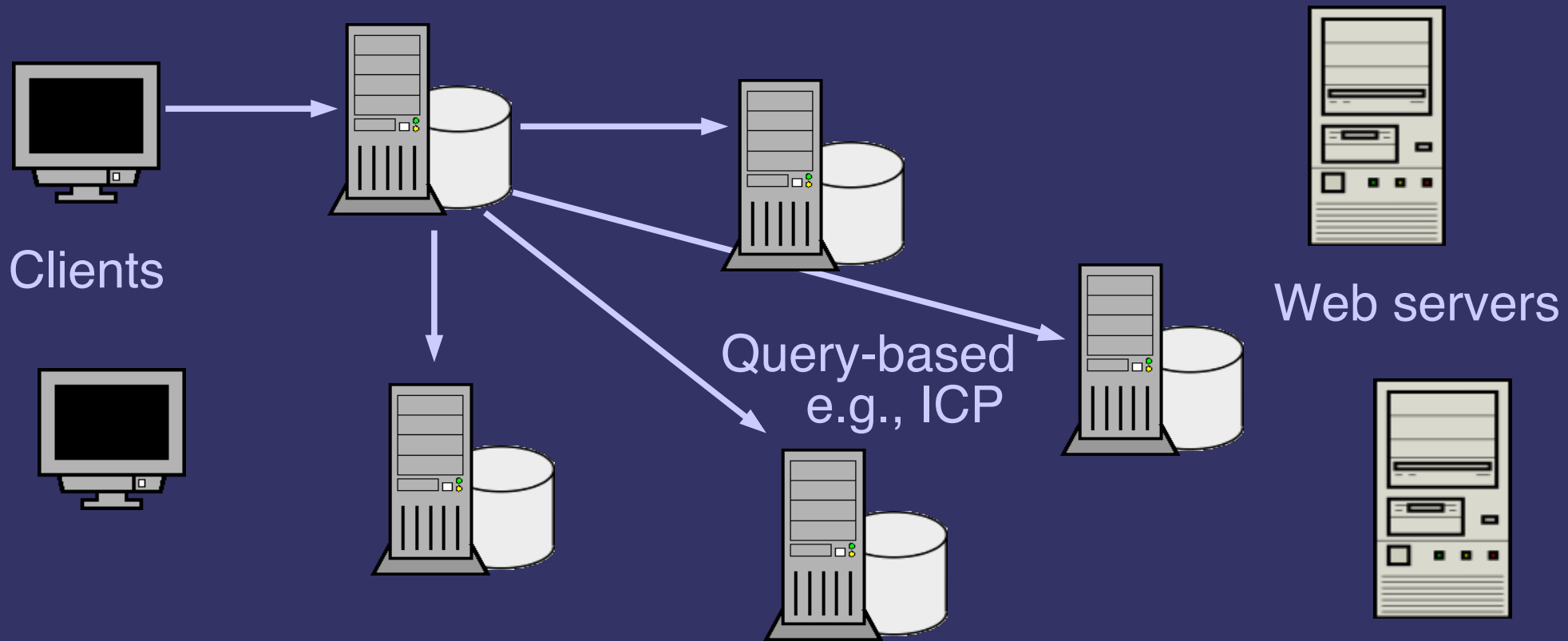
Web caching



With 25 years of Internet experience, we've learned exactly one way to deal with exponential growth: **Caching**.

[Van Jacobson]

Web caching

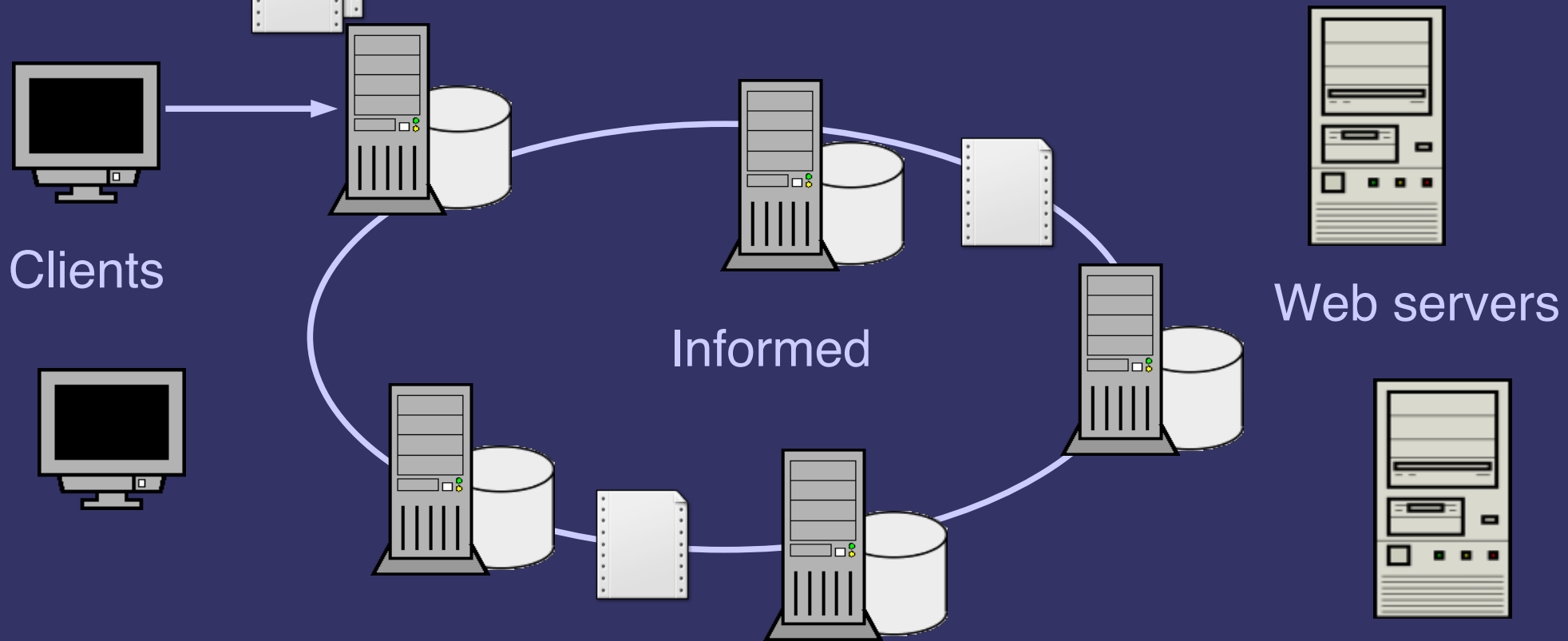


With 25 years of Internet experience, we've learned exactly one way to deal with exponential growth: **Caching**.

[Van Jacobson]

Metadata

Web caching



With 25 years of Internet experience, we've learned exactly one way to deal with exponential growth: **Caching**.

[Van Jacobson]

Limitations of "traditional" cooperation schemes

Hierarchical schemes

Bottleneck

Query-based

High cooperation overhead

Slow response time in case of miss

Summary-based (e.g., Cache Digests)

Lower hit rate

Highly sensitive to workload (client request frequency, object popularity, etc...)

IDEA: Can a mixed scheme compensate the drawbacks?

Two-tier cooperation schemes

Two-step lookup process

Combination of Summary- and Query-based schemes

Work in **heterogeneous** network environments

High/low speed connections

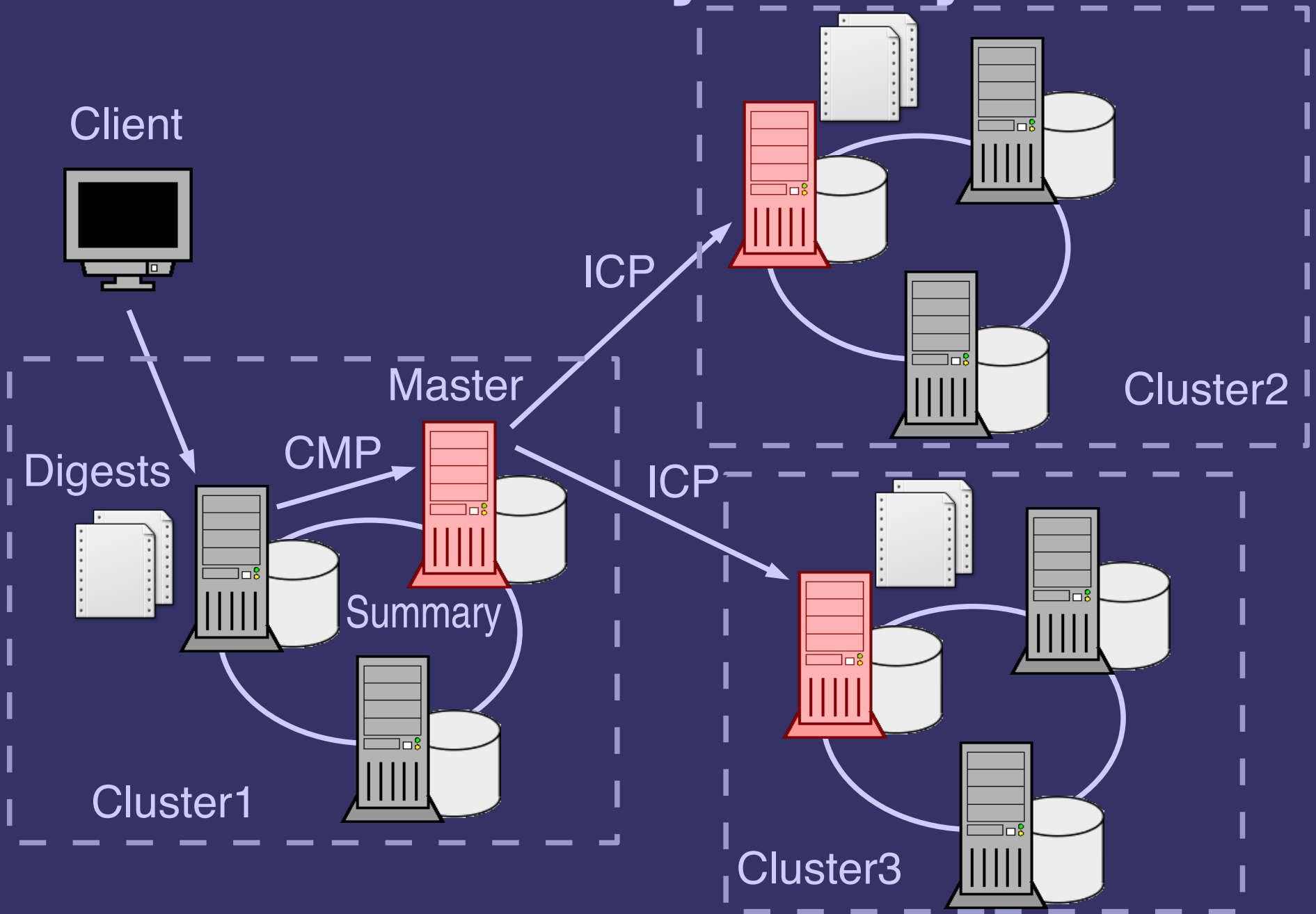
Peering points

Goal: higher scalability

$HR(\text{two-tier}) \gg HR(\text{summary})$

$OH(\text{two-tier}) \ll HR(\text{query})$

Summary-Query



SummaryMaster-Query

Summary-Query:

Twofold role of the *cluster master*

cluster master can become **bottleneck**

"Network-oriented" simulations did not evidence the risk of congestion



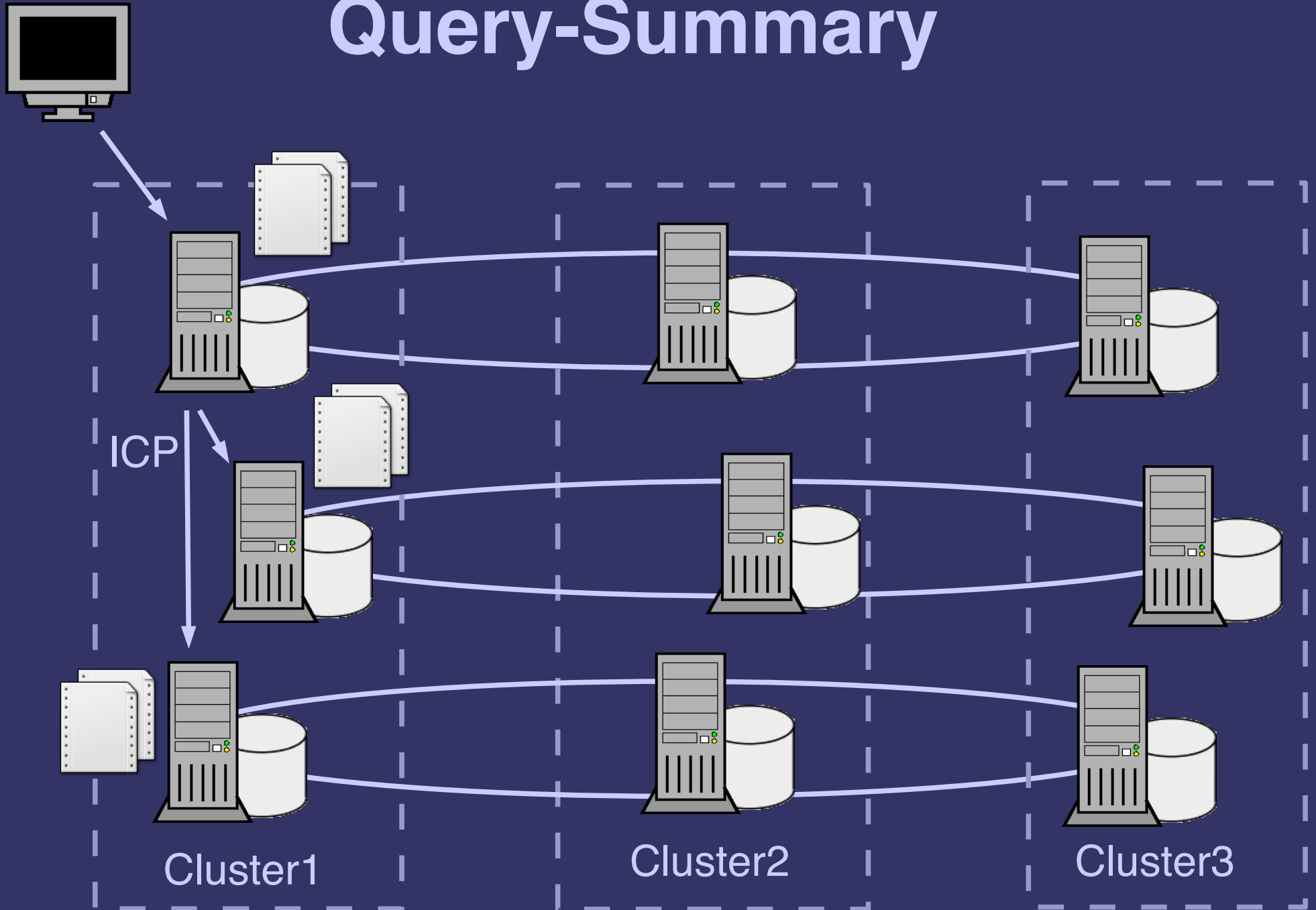
SummaryMaster-Query

Dedicated *cluster masters*

PRO: No risks of congestion

CON: Less cache servers available to clients

Query-Summary



Prototypes

Three versions

Based on Squid 2.4.x

Modification on the lookup mechanism

- ICP (ICP_MISS_POINTER)

- New protocol: Cache-to-Master Protocol (CMP)

- Neighbor management

Experiments

Two network scenarios

Local (two different workloads)

Scalability

Cooperation overheads

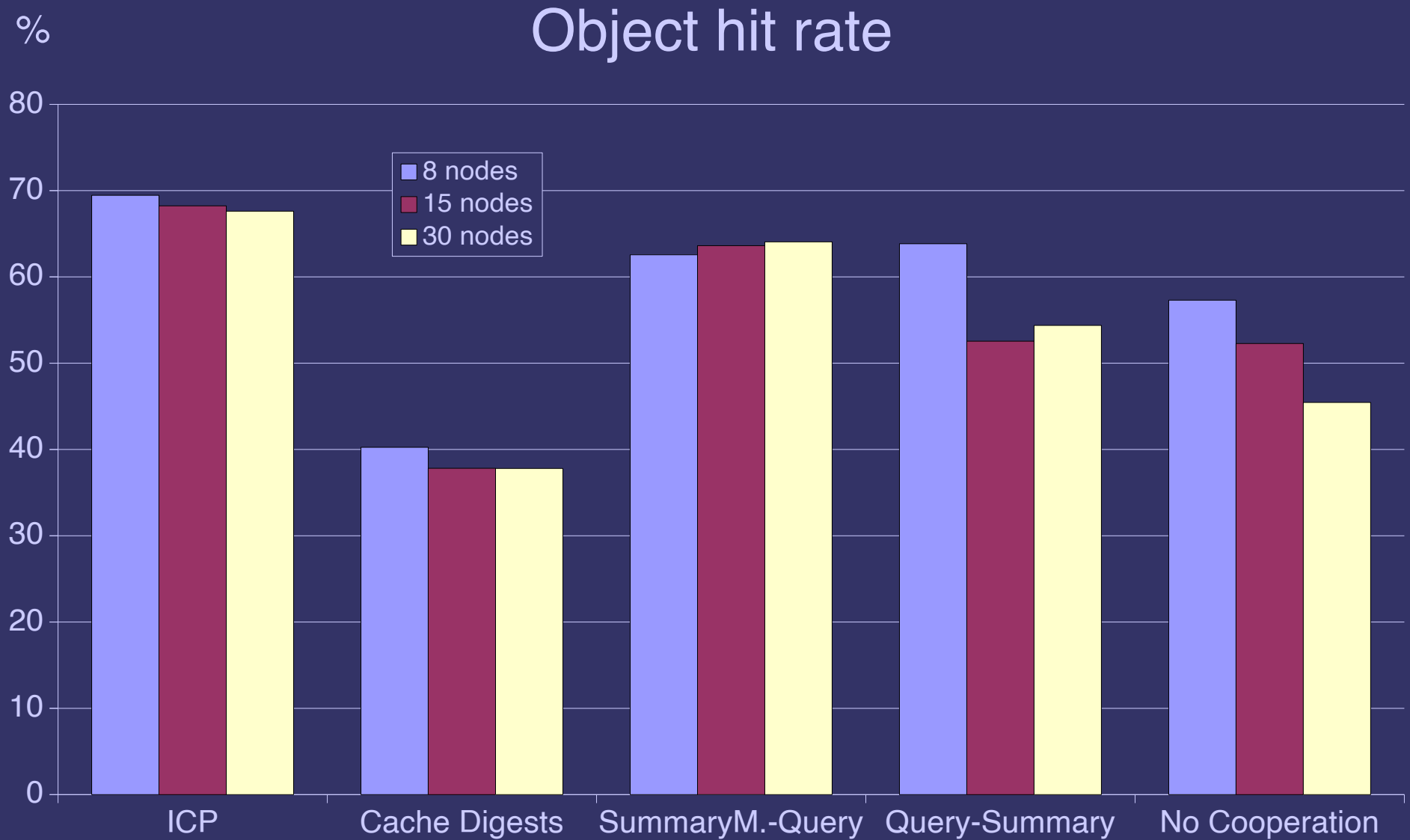
Geographic (two different network status)

User response time

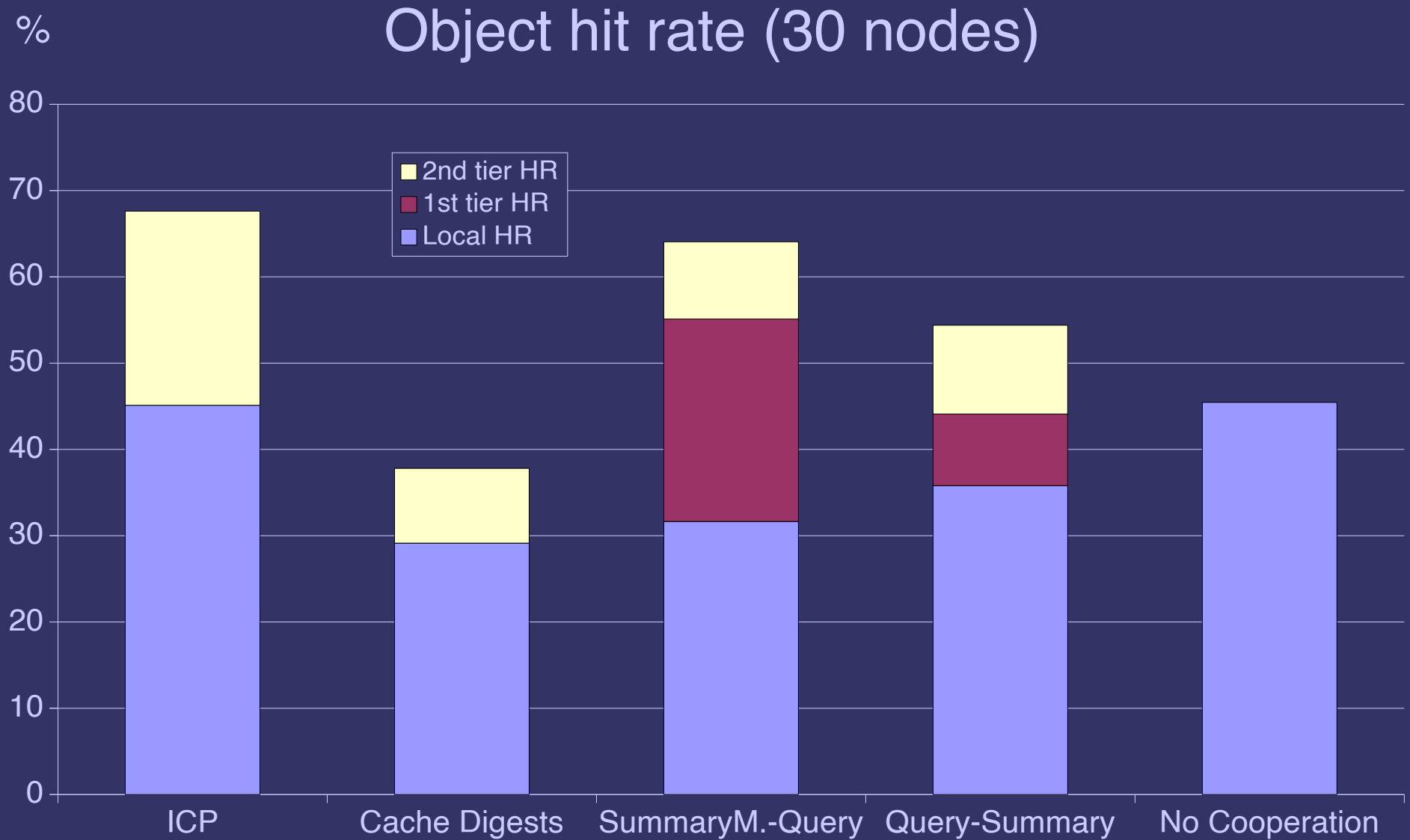
Web Polygraph as workload generator

Workload based on 2nd IRCache cache-off

Scalability test



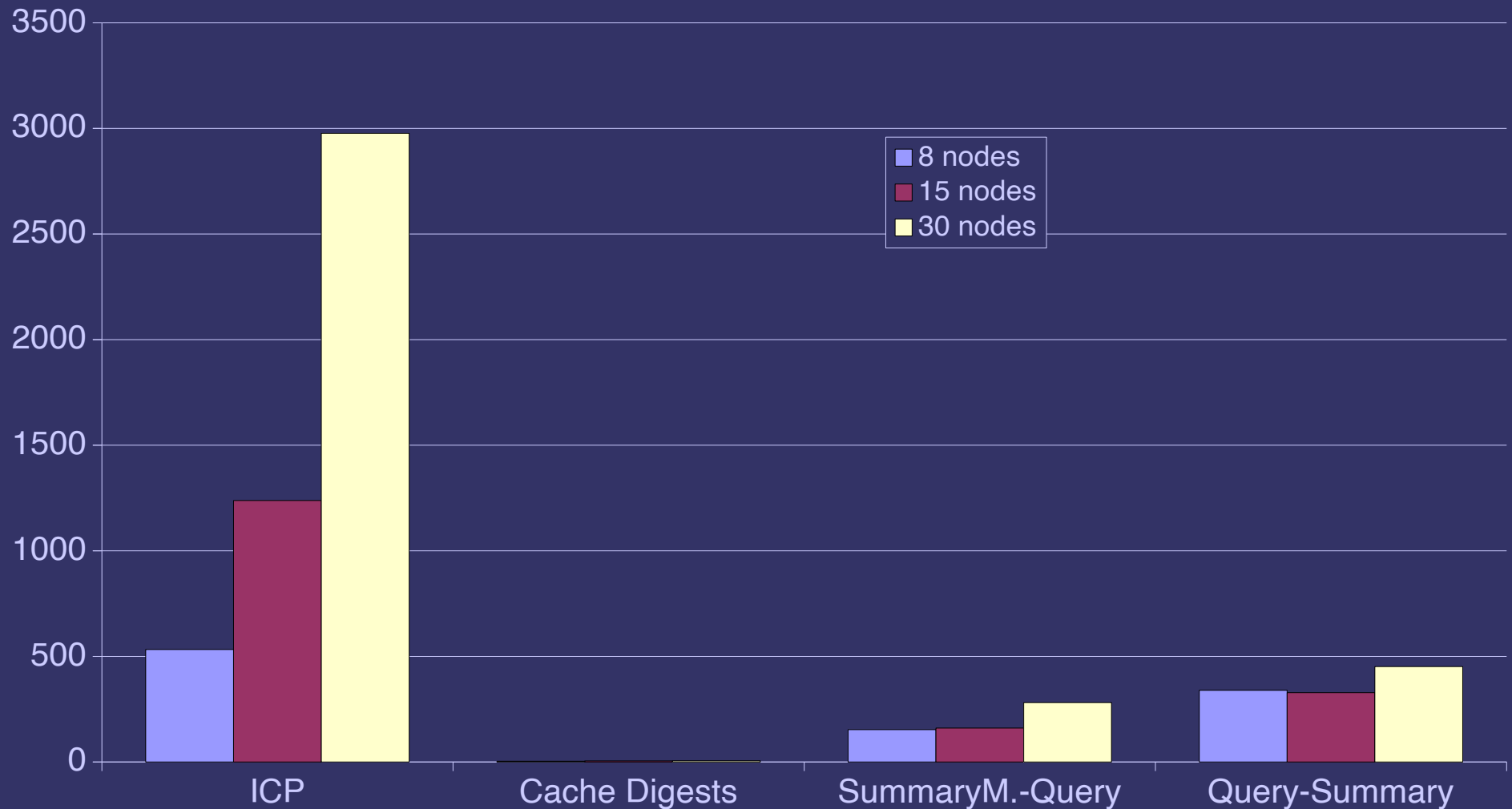
Hit rate contributions



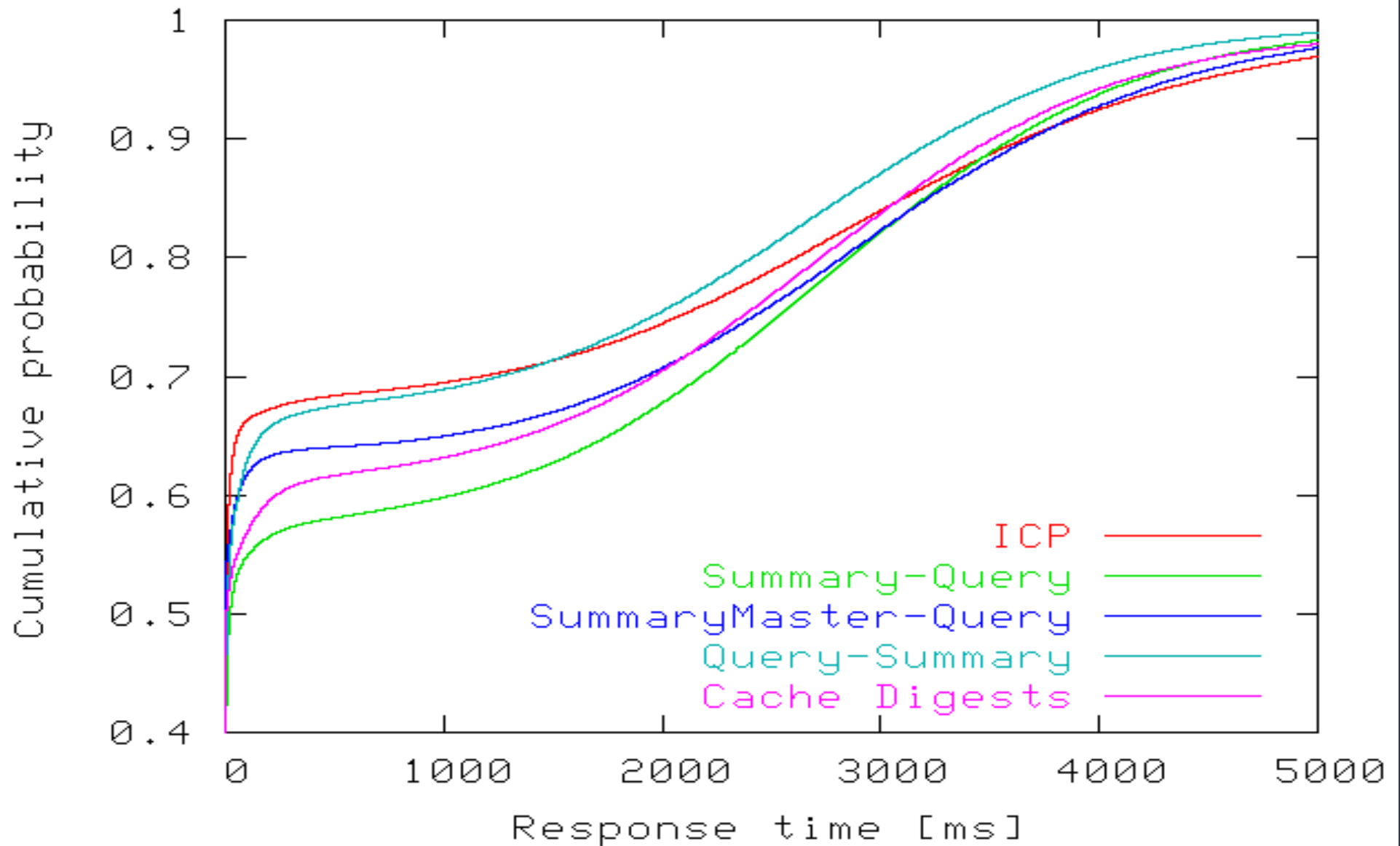
Scalability test

Bytes/req

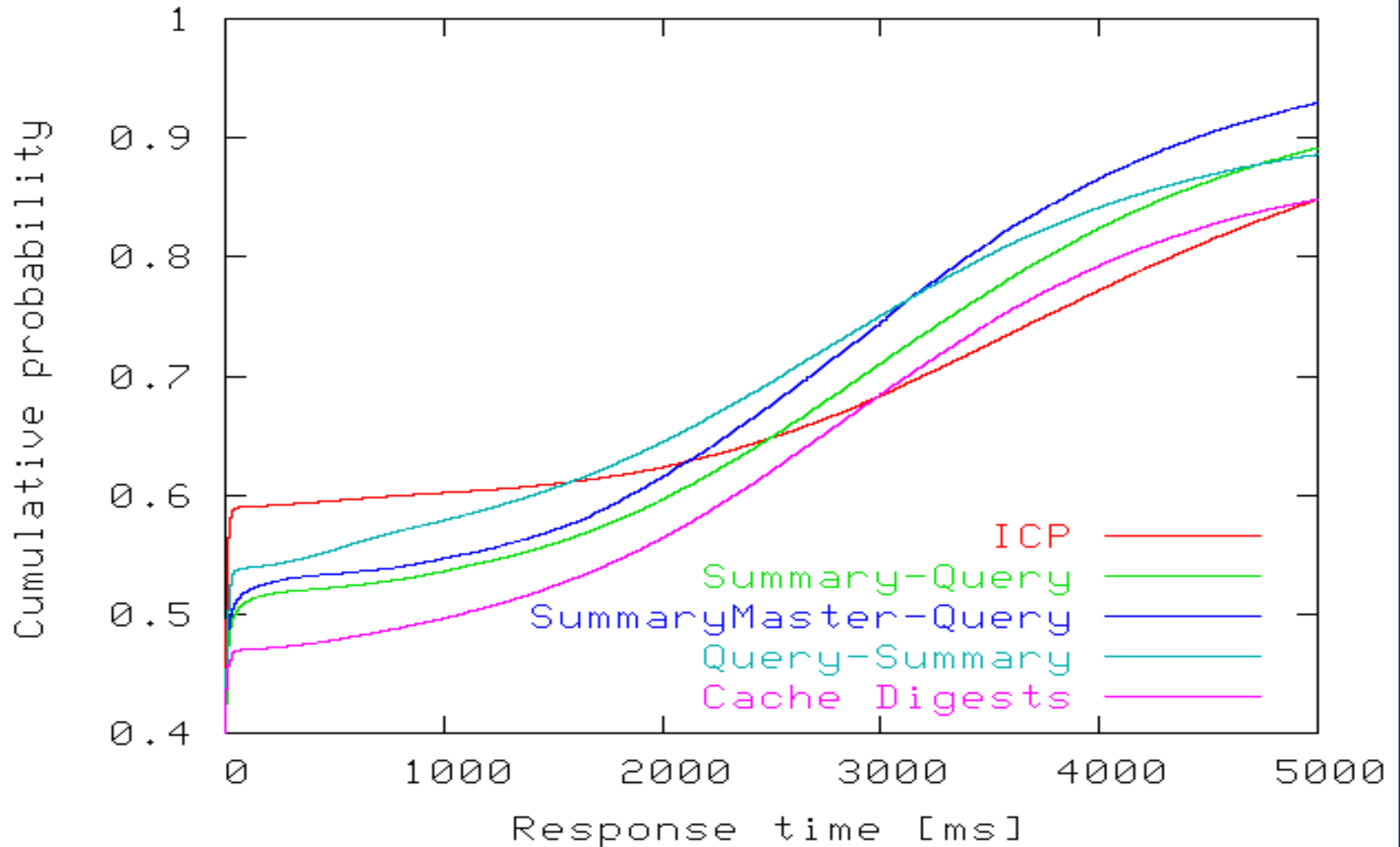
Cooperation overhead



Light network load



Heavy network load



Conclusions

Message for "network-oriented simulations"

Two-tier protocols offer:

- Reduced cooperation overhead (\ll ICP)

- Low penalty in hit rate (HR \gg Cache Digests)

- Viable solution in geographic scenarios

Distributed cooperation schemes for document lookup in multiple cache servers

Riccardo
Lancellotti
University of Rome
"Tor Vergata"

Bruno Ciciani
University of Rome
"La Sapienza"

Michele Colajanni
University of Modena and
Reggio Emilia