CS856 – Web Data Management

Proxy-Based Acceleration of Dynamically Generated Content on the World Wide Web: An Approach and Implementation

A. Datta, K. Dutta, H. Thomas, D. VanderMeer, Suresha, K. Ramamritham

Discussion presented by
Mohammed Abouzour

School of Computer Science
University of Waterloo

Outline

- Problem and background information
- Existing approaches
- Proposed approach
- Q & A

Problem and Background Information

- Sections 1 and 2 of paper
- Did a good job at introducing the problem
- Detailed information about the different bottlenecks associated with serving dynamically generated web pages (layout and content)

Existing Approaches

- Section 3 of paper
- Back-end caching approach works well in a dynamically generated content environment
- Proxy-based approaches work “somewhat”

Proposed Approach

- Setup cost
- Overhead of cache invalidation manager
- Performance depends on granularity of cacheable fragments

Proposed Approach (Cont.)

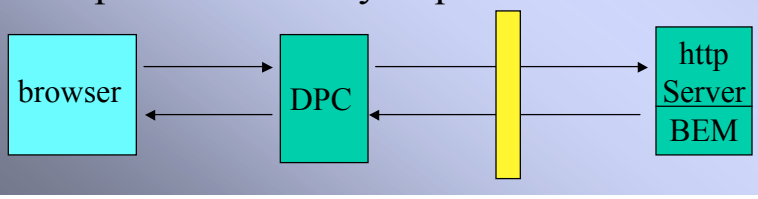
- Is the DPC a central point of failure/bottleneck?
- Scalability
- Proxy-based acceleration in a server farm environment

Setup Cost

- The author suggest building the site from the ground up with a new design structure: what is the cost involved in migrating existing sites to this architecture?
- *Tagging* process: what kind of APIs? XML?
- DPC has to be able to decode replies coming back from the original web site

Cache Invalidation Manager Overhead

- Every single request has to be completely interpreted and executed at the original Web Server
- What is the overhead of checking the availability and validity of cacheable components of every request?

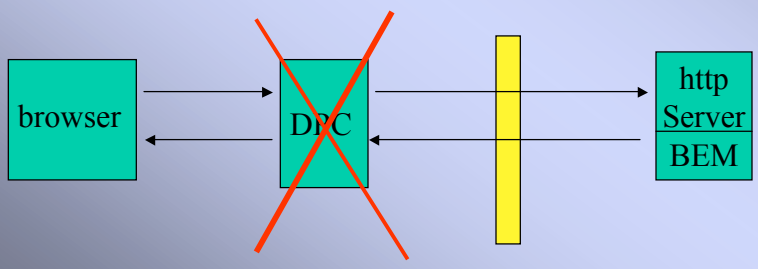


Performance depends on granularity of cacheable fragments

- Performance results depend largely on deciding which fragments are cacheable: might need retuning? (Fragment size and fragment id/key in DPC)
- In existing back-end approaches this is done automatically at finer granularity levels and provides better reuse of server resources
- Requires user intervention -> maintenance cost?

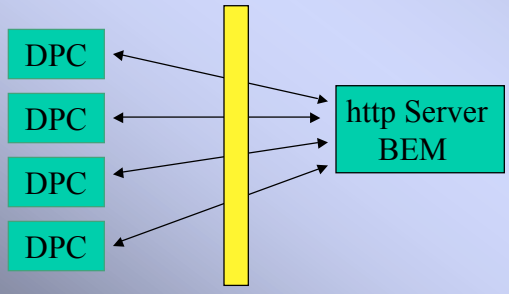
Dynamic Proxy Cache

- Central point of failure?
- A “new” performance bottleneck?



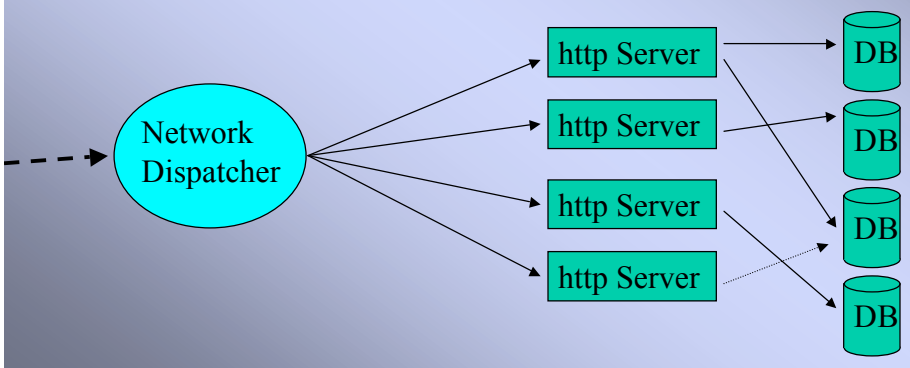
Scalability

- How scalable is the BEM component to handle multiple DPCs in a distributed reverse proxy environment?



Classical http Server Farm

- How well will does this solution apply in a http server farm setup?



Q & A