Personalized Chunk Framework for High Performance Personalized Web

K S Shailesh, IGNOU, New Delhi, India

Suresh Pachigolla Venkata, IGNOU, New Delhi, India

ABSTRACT

Dividing the web site page content or web portal page into logical chunks is one of the prominent methods for better management of web site content and for improving web site's performance. While this works well for public web page scenarios, personalized pages have challenges with dynamic data, data caching, privacy and security concerns which pose challenges in creating and caching content chunks. Web portals has huge dependence on personalized data. In this paper the authors have introduced a novel concept called "personalized content chunk" and "personalized content spot" that can be used for segregating and efficiently managing the personalized web scenarios. The authors' experiments show that performance can be improved by 30% due to the personalized content chunk framework.

KEYWORDS

Content Chunk, Personalization, Personalized Content Chunk, Personalized Web Acceleration, Web Content, Web Performance, Web Performance Optimization

INTRODUCTION

Most of the modern-day websites are personalized wherein the web site content and functionality will be rendered based on user preferences and other context parameters such as access device, language and such. Users expect Internet-facing applications to be responsive, interactive with optimal performance (Galletta et al., 2004). The key success criteria for an Internet-facing web site would be its usability and performance (Schmiedl et al., 2009). Web architecture should also satisfy the SLAs and quality attributes such as extensibility, performance, scalability and flexibility (Shivakumar, 2014).

In Internet-facing Information portals and public web site scenarios, the web site content is mostly stored in content management system (CMS) and we can apply the content chunking or content fragmenting concept. A web page content can be logically divided into modular fragments (Griffin et al., 2005). Content chunks are independent content pieces and represent a semantic entity (Challenger et al., 2005).

On a typical web page, the marquee image, right-hand content boxes can be created as a content chunk. Once the page content is segregated into logical content fragments, each of the individual content chunks can be managed. Content chunks can be independently authored, updated, published and cached. Chunks can also be tagged with metadata for easier identification. Chunks can also be cached individually with unique cache keys and this can improve the page performance.

In case of personalized web scenario, the page content would depend on many implicitly and explicitly specified parameters such as user preference, user profile attributes, user device, locale,

security roles, permissions, location, time and such. Web sites would push targeted content, relevant ads and effective recommendations using personalized web. Personalized web mostly contains dynamic values specific for a user group or for an individual user. Due to its dynamic nature, managing the content on a personalized web page is different than that of a static web page. We cannot statically author a personalized content chunk as the content author would not know of the dynamic values beforehand. Caching personalized content has its own set of challenges: caching a variant of content for each user would quickly become unmanageable. Global caching of personalized content would also violate privacy and security of individual users.

In this paper, we introduce a novel concept called "Personalized chunk" and "Personalized content spot" which brings the best of both worlds: flexibility and reusability of content chunk and security/privacy adherence of personalization scenarios.

Based on our experiments we were able to achieve about 30% performance improvement through usage of personalized chunk and personalized content spot.

PAPER ORGANIZATION

In the remaining portions of the introduction section we will look at state of the art methods along with related work and the significance of the work. We will discuss the complete details of the personalized content chunk framework, process steps, caching design in the "Method" section. In "Results" section we will look at the benchmarked results of performance numbers at various user load and content metrics. Finally, we will discuss the significance of results, explanation of the main findings, threats to validity and future scope of improvements in "discussion" section.

Literature Review and Related work

Content chunk is a predominant theme in this paper and we will be extending the concept of the chunk to personalized scenarios. Content chunk was proposed for reducing network traffic (Zhang et al., 2015) and for reducing download time in mobile device (Jang et al., 2014). Content chunk would improve the performance in CDN scenarios (Zhu et al., 2016). The concept of chunking can also be used to remove duplicates (Griffin et al. 2005) and can be used for device specific rendering and appropriate positioning (Davis et al. 2006). Keyword based chunk detection is another technique proposed by Brodie et al. 2004. Chunking concept is used to differentiate static chunks with dynamic chunks (Khaing & Thein 2005) and are used for prefetching (Maghoul et al. 2008).

For identifying and creating content chunks, Ramaswamy et al., 2004 proposes caching and reusability as the key criteria and Christos et al., 2004 and Ioannis et al., 2004 propose user driven personalization and page's internal structure for creating fragments and Chan & Woo (1999) proposes content similarity in a site for identifying content chunks. Challenger et al. (2005) and Challenger et al. (2000) propose an update mechanism for dynamic content update of content chunk and fragment based publishing.

Dynamic content caching (Challenger et al., 1999) and active cache (Cao et al., 1998) can be used for caching dynamic content chunks.

Content chunks are used for performance improvement (Datta et al. (2002) and for improving quality of service in dynamic content (Mohapatra & Chen (2001). Chunks are also used to improve learning on mobile devices

As far as content caching goes, web graph technique as discussed in Mohapatra & Chen (2001) can be used for chunk caching (Bruck et al. 2012 and Nagler et al., 2007).

10 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/personalized-chunk-framework-for-high-performance-personalized-web/183651

Related Content

A Survey on Privacy Preservation in Location-Based Mobile Business: Research Directions
Ahmed Alouiand Okba Kazar (2021). *International Journal of Web Portals (pp. 20-39).*
www.irma-international.org/article/a-survey-on-privacy-preservation-in-location-based-mobile-business/271396

Analysis of Accessibility Initiatives Applied to the Web
Joselice Ferreira Lima, Gustavo Miranda Caran, Luiz Fernando R. Molinaroand Daniela Favarro Garrossini (2012). *International Journal of Web Portals (pp. 48-58).*
www.irma-international.org/article/analysis-accessibility-initiatives-applied-web/75531

A Mobile Portal for Academe
Hans Lehmann, Stefan Bergerand Ulrich Remus (2007). *Encyclopedia of Portal Technologies and Applications (pp. 577-582).*
www.irma-international.org/chapter/mobile-portal-academe/17932

Web Site Portals in Local Authorities
Robert Laurini (2007). *Encyclopedia of Portal Technologies and Applications (pp. 1169-1176).*
www.irma-international.org/chapter/web-site-portals-local-authorities/18026

Building a Campus Portal-A Strategy that Succeeded
Anne Y. Bishop (2003). *Designing Portals: Opportunities and Challenges (pp. 186-200).*
www.irma-international.org/chapter/building-campus-portal-strategy-succeeded/8225