**An Efficient and Adaptive Decentralized File Replication Algorithm in P2P File Sharing Systems**

In peer-to-peer file sharing systems, file replication technology is widely used to reduce hot spots and improve file query efficiency. Most current file replication methods replicate files in all nodes or two end points on a client-server query path. However, these methods either have low effectiveness or come at a cost of high overhead. File replication in server side enhances replica hit rate, hence, lookup efficiency but produces overloaded nodes and cannot significantly reduce query path length. File replication in client side could greatly reduce query path length, but cannot guarantee high replica hit rate to fully utilize replicas. Though replication along query path solves these problems, it comes at a high cost of overhead due to more replicas and produces underutilized replicas. This paper presents an Efficient and Adaptive Decentralized (EAD) file replication algorithm that achieves high query efficiency and high replica utilization at a significantly low cost. EAD enhances the utilization of file replicas by selecting query traffic hubs and frequent requesters as replica nodes, and dynamically adapting to nonuniform and time-varying file popularity and node interest. Unlike current methods, EAD creates and deletes replicas in a decentralized self-adaptive manner while guarantees high replica utilization. Theoretical analysis shows the high performance of EAD. Simulation results demonstrate the efficiency and effectiveness of EAD in comparison with other approaches in both static and dynamic environments. It dramatically reduces the overhead of file replication, and yields significant improvements on the efficiency and effectiveness of file replication in terms of query efficiency, replica hit rate, and overloaded nodes reduction.

# ****Demand Based File Replication and Consistency Mechanism****

File replication and consistency are well known techniques in distributed systems to address the key issues such as scalability, reliability and fault tolerance. For many years, file replication and consistency in distributed environment has been researched to enhance and optimize the availability and reliability of the entire system. An effort has been made in the present work to propose a file popularity based, adaptive, on demand, reliable and comprehensive file replication and consistency mechanism. This paper introduces a group of File Replication Servers (FRS) that is responsible for ensuring the file availability, in order to fulfill the file request. The present work proposes Hybrid file update propagation mechanism so as to minimize the number of updates. Proposed model replicates the file, from one node to the other node, when the total number of request for a particular file reaches a threshold value. Proposed approach is implemented on JAVA platform. Based on the file popularity, model creates and updates the file replica on the server. The proposed Hybrid approach is compared with write invalidate and write update propagation policies. Comparison results show that for balanced load hybrid file update propagation model updates the optimum number of replicas. Whereas, for unbalanced load, hybrid approach outperforms other two approaches by handling all the file requests with minimum number of updates.

# ****Data Replication in P2P Collaborative Systems****

Data replication techniques have been extensively used in distributed systems to achieve, among others, due to system nodes failures: (a) high data availability, (b) system's reliability and (c) scalability. Due to the various characteristics of distributed systems as well as system's and application's requirements, a variety of data replication techniques have been proposed in the distributed computing field. One important distributed computing paradigm is that of P2P systems, which distinguish for their large scale and unreliable nature. The study and application of replication techniques becomes a must in such systems. While it is well understood and easy to achieve replication of immutable information (typically files) in P2P systems, it becomes more challenging to implement data replication techniques of dynamic data under highly dynamic nature of large P2P systems. Indeed, replicating documents that could change over time requires addressing the consistency issues. In this paper we study some data replication techniques for P2P collaborative systems. We identify several contexts and use cases where data replication can greatly support collaboration. We then consider as a case study replication techniques for dynamic documents in the context of a peer-group based P2P system of super-peer architecture. P2P collaborative systems arise in many groupware applications, such as collaborative work in online teams, having requirements for high availability and system reliability (e.g. in disaster management scenarios). We propose a replication system for documents structured as XML files to address the dynamics of the documents at peers and use the super-peer to ensure a satisfactory level of document consistency among peers.

# ****A proactive low-overhead file replication scheme for structured P2P content delivery networks****

File replication is a widely used technique for high performance in peer-to-peer content delivery networks. A file replication technique should be efficient and at the same time facilitates efficient file consistency maintenance. However, most traditional methods do not consider nodes' available capacity and physical location in file replication, leading to high overhead for both file replication and consistency maintenance. This paper presents a proactive low-overhead file replication scheme, namely Plover. By making file replicas among physically close nodes based on nodes' available capacities, Plover not only achieves high efficiency in file replication but also supports low-cost and timely consistency maintenance. It also includes an efficient file query redirection algorithm for load balancing between replica nodes. Theoretical analysis and simulation results demonstrate the effectiveness of Plover in comparison with other file replication schemes. It dramatically reduces the overhead of both file replication and consistency maintenance compared to other schemes. In addition, it yields significant improvements in reduction of overloaded nodes.

# ****S-CLONE: Socially-aware data replication for social networks****

Online social networking has become one of the most important forms of today's communication. While an online social network can be attractive for many socially interesting features, its competitive edge will diminish if it is not able to keep pace with increasing user activities. Deploying more servers is an intuitive way to make the system scale, but for the best performance one needs to determine where best to put the data, whether replication is needed, and, if so, how. This paper is focused on replication; specifically, we propose S-CLONE, a socially-aware data replication scheme which can significantly improve a social network's efficiency by taking into account social relationships of its data. S-CLONE's performance is substantiated in our evaluation study.

# ****GlobeDB: autonomic data replication for web applications****

We present GlobeDB, a system for hosting Web applications that performs autonomic replication of application data. GlobeDB offers data-intensive Web applications the benefits of low access latencies and reduced update traffic. The major distinction in our system compared to existing edge computing infrastructures is that the process of distribution and replication of application data is handled by the system automatically with very little manual administration. We show that significant performance gains can be obtained this way. Performance evaluations with the TPC-W benchmark over an emulated wide-area network show that GlobeDB reduces latencies by a factor of 4 compared to non-replicated systems and reduces update traffic by a factor of 6 compared to fully replicated systems.

# ****A novel data replication and management protocol for mobile computing systems****

Mobile computing has enabled users to seamlessly access databases even when they are on the move. Mobile computing environments require data management approaches that are able to provide complete and highly available access to shared data at any time from any where. In this paper, we propose a novel replicated data protocol for achieving such goal. The proposed scheme replicates data synchronously over stationary sites based on three dimensional grid structure while objects in mobile sites are asynchronously replicated based on commonly visited sites for each user. This combination allows the proposed protocol to operate with less than full connectivity, to easily adapt to changes in group membership and not require all sites to agree to update data objects at any given time, thus giving the technique flexibility in mobile environments. The proposed replication technique is compared with a baseline replication technique and shown to exhibit high availability, fault tolerance and minimal access times of the data and services, which are very important in an environment with low-quality communication links.

# ****Application specific data replication for edge services****

The emerging edge services architecture promises to improve the availability and performance of web services by replicating servers at geographically distributed sites. A key challenge in such systems is data replication and consistency so that edge server code can manipulate shared data without incurring the availability and performance penalties that would be incurred by accessing a traditional centralized database. This paper explores using a distributed object architecture to build an edge service system for an e-commerce application, an online bookstore represented by the TPC-W benchmark. We take advantage of application specific semantics to design distributed objects to manage a specific subset of shared information using simple and effective consistency models. Our experimental results show that by slightly relaxing consistency within individual distributed objects, we can build an edge service system that is highly available and efficient. For example, in one experiment we find that our object-based edge server system provides a factor of five improvement in response time over a traditional centralized cluster architecture and a factor of nine improvement over an edge service system that distributes code but retains a centralized database.

# ****An Efficient Data Replication Scheme for Peer-to-Peer Video Streaming over Wireless-Mesh Community Networks****

Wireless mesh networks (WMNs) are expected to play an important role in the next–generation Internet because of its several advantages, such as self-organization, high capacity, low deployment cost and low maintenance cost. In this paper, we discuss how to provide robust home-to-home on-demand streaming on the basis of WMNs. A key challenge for such a community-based P2P multimedia system is how to efficiently place and replicate video clips across different home devices in the face of frequent online/offline behavior of users. In this paper, we propose an efficient data replication scheme called H2-VIP for home-to-home video streaming. H2-VIP computes the optimal number of replicas of video blocks. Our simulation results show that the proposed scheme can largely reduce the failure rate of home-to-home streaming at the cost of a small amount of extra storage space.

**A Mobile Peer-to-Peer System for Opportunistic Content-Centric Networking**

In this work we present a middleware architecture for a mobile peer-to-peer content distribution system. Our architecture allows wireless content dissemination between mobile nodes without relying on infrastructure support. Contents are exchanged opportunistically when nodes are within communication range. Applications access the service of our platform through a publish/subscribe interface and therefore do not have to deal with low-level opportunistic networking issues or matching and soliciting of contents. Our architecture consists of three key components. A *content structure* that facilitates dividing contents into logical topics and allows for efficient matching of content lookups and downloading under sporadic node connectivity. A *solicitation protocol* that allows nodes to solicit content meta-information in order to discover contents available at a neighboring node and to download content entries disjointedly from different nodes. An *API* that allows applications to access the system services through a publish/subscribe interface. In this work we describe the design and implementation of our architecture. We also discuss potential applications and present evaluation results from profiling of our system.

**SCALAR: Scalable data lookup and replication protocol for mobile ad hoc networks**

Data replication, as an essential service for MANETs, is used to increase data availability by creating local or nearly located copies of frequently used items, reduce communication overhead, achieve fault-tolerance and load balancing. Data replication protocols proposed for MANETs are often prone to scalability problems due to their definitions or underlying routing protocols they are based on. In particular, they exhibit poor performance when the network size is scaled up. However, scalability is an important criterion for several MANET applications. We propose a scalable and reactive data replication approach, named SCALAR, combined with a low-cost data lookup protocol. SCALAR is a virtual backbone based solution, in which the network nodes construct a connected dominating set based on network topology graph. To the best of our knowledge, SCALAR is the first work applying virtual backbone structure to operate a data lookup and replication process in MANETs. Theoretical message-complexity analysis of the proposed protocols is given. Extensive simulations are performed to analyze and compare the behavior of SCALAR, and it is shown to outperform the other solutions in terms of data accessibility, message overhead and query deepness. It is also demonstrated as an efficient solution for high-density, high-load, large-scale mobile ad hoc networks.

# ****Dynamic Data Replication Scheme in the Cloud Computing Environment****

In the cloud computing environment, data replication strategy (DRS) is used to improve data access. Related studies have proposed data replication strategies. The performances of these strategies are closely related to the users' access patterns, and work optimally for a particular data access pattern. However, as the data access patterns become more flexible and unpredictable, it is difficult to manage them with traditional replication strategies. Given this circumstance, this paper proposes an algorithm that detects changes in a user's data access pattern and dynamically applies an optimal replication strategy. The proposed algorithm has the advantage of maintaining an optimal performance by responding to various data access patterns. We tested the proposed algorithm and validated its effectiveness.