ABSTRACT

**ABSTRACT:** With the rapid development of information and network technology, Internet users and emerging applications are growing in strength. In addition, Internet has gradually evolved from a packet-switched network to an integrated service network that hosts mass service information, such as multimedia communications, cloud computing, datacenter, and user interaction etc. Network with single transport function can hardly satisfy the new service mode of dynamically developing and changing demand of users and services It has become urgent need to be addressed in the future Internet design to provide support and protection for the emerging information-based demands. Aiming to meet the urgent need for massive emerging demand and improve the circumstance of lacking in perception and adaption mechanisms, the dissertation presents intensive researches on the demand-aware resource adaptive schemes from the following three perspectives, i.e. how to perceive demands, how to utilize demands, and how is the effect of introducing demands. Main contributions and innovations of the dissertation are as follows:

1. Focusing on how to perceive demands, a general fundamental framework of demand perception in the future Internet architecture is proposed in order to realize flexible embedding, integration and extension for information based massive emerging demand. By arguing the necessity and challenges of demand perception design, the dissertation detailed describes the overall idea and design of the general fundamental framework of demand perception, summarizes the advantages of the framework, and illustrates its basic workflow. Finally, a prototype is built to realize the framework, and presents evaluation based on resource adaptation mechanism taking the service size as an example. Numerical results testify the feasibility and effectiveness of the framework with obvious advantages in several major aspects.
2. Focusing on how to utilize user demands and how is the effect of introducing demands, a user-aware model for describing the process of service selection and a user-aware adaptive transmission scheme is proposed. In order to analyze the influence of introducing user demands, the dissertation takes user demands and preference into consideration, and builds a service selection probabilistic model based on user preference to describe user preference based service decision-making in terms of statistical probability distribution of service selections.

changing rules of service selection, the basic model for describing the process of service selection takes user demands and preference into consideration.

然后通过理论推导得出基于用户偏好需求的服务选择变化规律，并运用蒙特卡洛仿真对服务选择过程进行大量抽样统计，从理论推导和仿真测试两个角度直观展示不同场景下用户偏好需求的引入对服务选择的影响。分析结果表明，用户需求的引入会导致服务选择呈现一定分布的变化规律，有助于服务解析系统的个性化推送决策。为了研究用户需求感知与适配的方法，本论文以用户级别这一用户需求为例，在内容中心网络架构模型下，提出了基于用户级别感知的自适应分级传输机制（UAAT）。本论文详细介绍了UAAT机制中用户级别需求的嵌入与感知方法、用户级别管理器模块设计以及基于用户级别的链路带宽适配方案。仿真结果表明，相比于传统CCN机制而言，UAAT机制能够实现基于用户需求的差异化服务，有效提升链路的带宽利用率，有助于实现服务个性化的定制。

In order to describe and analyze the changing rules of service selection, the basic model for describing the process of service selection takes user demands and preference into consideration. Firstly, the random/preference decision mechanisms are introduced to represent the final decision on the user side. Then, the full/limited candidates pattern is used to discuss its influence on recommendation range. Thirdly, the ranking and matching mechanisms are introduced .Numerical simulation results show that user demand for preference makes the access of services following a certain regularity. Utilizing user demand turns to be practicable and effective for a user to obtain a relative preferred service. UAAT scheme is proposed for adapting to user’s personalized transmission demands in content centric network. UAAT scheme advises a new component – Level Manager (LM) to the original CCN router which can combine user level together with bandwidth and allocate corresponding bandwidth for different level users depending on the interface condition and the incoming Interest level. UAAT scheme introduces a new idea of adaptively distributing network resources based on different user level, especially suitable for content centric multimedia network. The simulation results show that UAAT has implemented a better performance (about 25.6% improvement than the original CCN scheme) in utilizing available bandwidth.

（3）针对服务需求如何利用及其效果如何的问题，研究服务需求优化网络传输的适配方法，提出了期限感知下的低开销数据流传输适配机制（DSLO）。DSLO机制的核心思想是通过对有效期这一服务需求的感知，预判数据中心网络中的无效期限数据流，从而降低网络开销浪费情况，实现需求信息对网络资源的优化管理。本论文首先对DSLO机制的总体设计思路进行详细描述，设计了对应的低开销网络资源适配流程。在此基础上，基于对拥塞窗口变化的合理假设与理论推导，提出以无效概率变量为预判依据的一种发送端驱动的无效期限流预判机制（basic DSLO），并对机制进行详细描述和鲁棒性分析。然后建立无效期限流概率预测模型，提出以无效期限流预测概率变量为预判依据的基于阈值和基于概率的两种交换机驱动的自适应预判机制（adaptive DSLO），并对该机制进行必要的流程描述。最后，在典型数据中心拓扑下针对包括DSLO机制在内的不同期限流传输机制进行仿真实现，仿真结果表明，DSLO机制能够有效减小网络开销，并进一步降低网络中无效期限流的比例包，且在乱序情况下的鲁棒性表现最优。

3. Foucing on “how to utilize the service demands and how is the effect on network optimization”, a deadline-aware scheme with low overhead (DSLO) for datacenter networks is proposed , aiming at reducing the bandwidth overhead incurred by unproductive flows. To achieve this, DSLO introduces a sender-driven strategy for predicting unproductive flows based on the Explicit Congestion Notification mechanism. DSLO also proposes a probabilistic model and two switch-driven adaptive strategies (a threshold-based strategy and a probability-based strategy) to predict unproductive flows more precisely. Simulation results show DSLO’s effectiveness on reducing both the overhead and the ratio of flows whose deadlines are not met compared with some typical mechanisms for scheduling deadline-aware flows.

（4）同样针对服务需求如何利用及其效果如何的问题，研究服务需求提升服务质量的适配方法，提出基于期限需求的数据流调度机制（FBDS）。FBDS机制的核心思想是采用集中式的控制方案，推迟一些数据流至期限时刻完成，通过预留传输空间的方式缓解数据中心网络中高并发流情况下的期限数据流失效问题。本论文依次对FBDS机制的设计目标、通信流程、核心停等调度算法进行具体介绍，并从调度块独立性和停等机制合理性两个方面对停等调度算法进行理论推导验证。最后针对包括FBDS机制在内具有代表性的几类期限流传输机制进行大规模仿真，仿真结果表明，基于调度块的传输和停等机制可以在高并发数据流的情况下，实现基于期限约束的集中式调度决策，进一步提升了数据流的期限满足率。

4. Foucing on “how to utilize the service demands and how is the effect on service quality”, a flow-based deadline scheduling scheme for datacenter networks (FBDS) is proposed, aiming to alleviate the unsatisfactory effect in the condition of high fan-in degree. FBDS makes the unilateral deadline-aware flow transmission with priority transform into a compound centralized single-machine deadline-based flow scheduling decision. In addition, FBDS blocks the flow sets and postpones some flows with extra time until their deadlines to make room for the new arriving flows in order to improve the deadline meeting rate. The simulation results on flow completion time and deadline meeting rate reveal the potential of FBDS in terms of a considerable deadline-sensitive transport protocol for deadline-sensitive interactive services.

**KEYWORDS：**demand perception;resource adaption; user preference; user level; deadline