

عنوان فارسی مقاله :

تضمین عملکرد برای سیستم های نهایی- وب سرور:
شیوه کنترل نظری

عنوان انگلیسی مقاله :

Performance Guarantees for Web Server End-Systems:
A Control-Theoretical Approach



توجه !

این فایل تنها قسمتی از ترجمه میباشد. برای تهیه مقاله ترجمه شده کامل با فرمت ورد (قابل ویرایش) همراه با نسخه انگلیسی مقاله، [اینجا](#) کلیک نمایید.

8 Conclusions and Future Work

In this paper, we demonstrated the application of control theory to Internet server performance control. We presented a QoS-management architecture that relies on adapting the delivered content to control server utilization. Unlike contemporary non-adaptive servers, and unlike servers that implement “binary” — accept or reject — admission control, content adaptation enables a server to provide a smooth range of client degradation, thereby handling overload gracefully. We proposed in Section 4 the design and implementation of a utilization control loop that regulates the extent of degradation (in the service level and number of clients) so as to satisfy a pre-specified utilization bound in the presence of variable server load while virtually eliminating connection errors. In Section 4.4 we have shown how utilization control may be used to satisfy individual time constraints. We demonstrated several extensions to this mechanism that provide performance isolation, service differentiation, excess-capacity sharing, and QoS guarantees. The mechanisms described in this paper are largely independent of workload assumptions, and can be easily applied to different platforms by appropriately tuning a small set of parameters using well-founded analytic techniques. The architecture can be implemented in a middleware layer transparently to existing server and browser code thereby facilitating its deployment.



8. نتایج و کار آتی

در این مقاله، در مورد کاربرد تئوری کنترل برای کنترل عملکرد سرور اینترنتی توضیح دادیم. در اینجا معماری مدیریت QoS را مطرح کردیم که بر انطباق محتوای تحویل داده شده جهت کنترل بهره برداری از سرور متکی می باشد. برخلاف سرورهای غیر انطباقی و برخلاف سرورهای پیاده کننده کنترل پذیرش دودویی - پذیرش یا رد، انطباق محتوا به سرو امکان تنزل رتبه کاربر و در نتیجه جابجایی حجم کار را می دهد. در بخش 4 طراحی و پیاده سازی حلقه کنترل بهره برداری را پیشنهاد کردیم که میزان تنزل رتبه را تنظیم کرده (در سطح خدمات یا سرویس و تعداد کاربران) و بدین طریق حدبهره برداری از پیش تعیین شده در حضور بار متغیر سرور را تعیین وبالاخره خطاهای اتصال را به حداقل می رساند. در بخش 4.4 روش استفاده از کنترل بهره برداری برای تامین محدودیت های زمانی فردی را نشان داده ایم. سپس این مکانیزم را به طرق مختلفی توسعه دادیم که تفکیک عملکرد، تمایز خدمات، به اشتراک گذاشتن ظرفیت اضافی و تضمین QoS فراهم می نماید. مکانیزم های تشریح شده در این مقاله عمدتاً مستقل از فرضیات حجم کار بوده و با تنظیم مناسب مجموعه کوچکی از پارامترها با استفاده از تکنیک های تحلیلی درست بنیان، به راحتی می توان از آنها در پلتفرم های مختلف استفاده نمود. معماری را می توان در لایه میان افزار سرور موجود و کد جستجوگر پیاده و بدین طریق روند استقرارش را تسهیل نمود.

توجه!

این فایل تنها قسمتی از ترجمه میباشد. برای تهیه مقاله ترجمه شده کامل با فرمت

ورد (قابل ویرایش) همراه با نسخه انگلیسی مقاله، [اینجا](#) کلیک نمایید.

همچنین برای مشاهده سایر مقالات این رشته [اینجا](#) کلیک نمایید.