



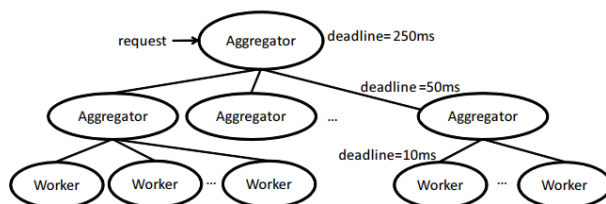
## چکیده

این مقاله در مورد چالش‌هایی است که مراکز داده ابری در مدیریت برنامه‌های متنوع با ترکیب بارهای کاری نیازمند تاخیر پیش‌بینی‌پذیر کوچک با سایر بارهای کاری نیازمند پهنای باند مستمر و بزرگ هستند. در این محیط، پروتکل TCP امروزی کمبود دارد. در اینجا اندازه‌گیری‌هایی از یک خوشه با ۶۰۰۰ سرور ارائه کرده و نقص‌هایی که منجر به تاخیر بالای برنامه‌ها می‌شوند آشکار می‌شوند. به عنوان مثال، جریان‌های "پس‌زمینه" با نیاز به پهنای باند افزایش می‌یابند و صف‌ها در سوئیچ‌ها را پر می‌کنند، و در نتیجه عملکرد ترافیک "پیش‌زمینه" حساس به تاخیر را تحت تأثیر قرار می‌دهند. برای حل این مشکلات، ما پروتکل DCTCP را معرفی شده است، یک پروتکل مانند TCP برای شبکه‌های مراکز داده. DCTCP از ECN در شبکه استفاده می‌کند. ارزیابی DCTCP را با سرعت‌های ۱ و ۱۰ گیگابیت بر ثانیه با استفاده از سوئیچ‌های بافر کم انجام شده و نتایج نشان می‌دهند که DCTCP می‌تواند همان یا حتی بهتر از TCP در انتقال داده با استفاده از ۹۰ درصد کمتر از فضای بافر باشد. علاوه بر این، برخلاف TCP، DCTCP همچنین تحمل شدت بالا به افزایش ناگهانی و تاخیر پایین برای جریان‌های کوتاه ارائه می‌دهد.

## خلاصه

در سال‌های اخیر، تحولات در عرصه محاسبات از طریق ادغام گسترده‌ی IT شرکت‌های بزرگ به مراکز داده و ظهور ارائه‌دهندگان خدمات محاسبات ابری منجر به تمرکز بر طراحی زیرساخت محاسباتی و ذخیره‌سازی با کارایی و دسترسی بالا با استفاده از اجزاء ارزان و عمومی شده است. این روند به شبکه‌های مراکز داده نیز گسترش یافته است و سوئیچ‌های ارزان به عنوان یک اجزاء عمده رواج یافته‌اند. این مقاله به بررسی اجرای واقع‌گرایانه این چشم‌انداز با مرور عملکرد سوئیچ‌های عمومی در مقابل ترافیک واقعی برنامه‌های مراکز داده می‌پردازد. نویسندگان پروتکل Data Center TCP (DCTCP) را معرفی کرده‌اند که از اعلان آشکار آگاهی از ازدحام (ECN) بهره می‌برد تا به این نیازهای برنامه‌ها پاسخ دهد. طراحی DCTCP به دنبال حفظ اشغال کم فضای بافر سوئیچ‌ها و همزمان دستیابی به بهره‌وری بالا برای جریان‌های مختلف کوتاه و بلند در محیط مراکز داده است. مقاله بر امکان پیاده‌سازی DCTCP در سخت‌افزارهای موجود، به ویژه سوئیچ‌های عمومی، تأکید دارد و کارایی آن در دستیابی به توان کامل با فضای بافر کمتر نسبت به TCP را برجسته می‌کند.

در بخش دوم این مقاله، به بررسی چالش‌هایی که پروتکل‌های انتقال در مراکز داده با آنها روبه‌رو هستند، می‌پردازیم. ابتدا یک ساختار معمولی برنامه به نام Partition/Aggregate را معرفی کرده و دلایلی که تاخیر به عنوان یک معیار حیاتی در مراکز داده محسوب می‌شود را شرح می‌دهیم. الگوی Partition/Aggregate که در شکل روبرو نشان داده شده است، پایه‌ای برای



بسیاری از برنامه‌های بزرگ وب است و بر اساس آن، درخواست‌ها از لایه‌های بالاتر برنامه به قطعاتی تقسیم و به کارگران در لایه‌های پایین ارسال می‌شوند. پاسخ‌های این کارگران جمع‌آوری می‌شوند تا یک نتیجه تولید شود. این الگو مبتنی بر نیازهای طراحی بسیاری از برنامه‌های وب

مانند جستجوی وب، ترکیب محتوای شبکه‌های اجتماعی و انتخاب تبلیغات است. در برنامه‌ها زمان واقعی تعاملی مانند این‌ها، تاخیر به عنوان معیار اصلی در نظر گرفته می‌شود و حداکثر تاخیر مجاز کلی توسط مطالعات تأثیر مشتری مشخص می‌شود. ادغام

تاخیرهای معمول اینترنت و تاخیرهای نمایش، بخش "بک‌اند" برنامه را در نظر می‌گیرد که به طور معمول بین ۲۳۰ تا ۳۰۰ میلی‌ثانیه تخصیص داده شده است. این مقاله دو ارائه اصلی دارد. ابتدا ترافیک تولیدی از یک خوشه با ۶۰۰۰ سرور را اندازه‌گیری و تحلیل می‌کند (بخش ۲)، و الگوها و نیازهای برنامه‌ها را استخراج می‌کند. در ادامه، Data Center TCP (DCTCP) را معرفی می‌کند که این نقص‌ها را برطرف می‌کند (بخش ۳). DCTCP از اعلان آشکار آگاهی از سرکوب (ECN) استفاده می‌کند تا به اطلاعات چندبیتی به میزبان‌های پایانی برسد. در نتیجه، DCTCP همان یا بهتر از TCP بهره‌وری را ارائه می‌دهد، در حالی که از ۹۰٪ کمتر از فضای بافر استفاده می‌کند. برخلاف TCP، DCTCP همچنین تحمل شدت بالا و تاخیر کمی برای جریان‌های کوتاه فراهم می‌کند.

طراحی DCTCP از TCP سنتی الهام گرفته و هدف آن دستیابی به تحمل بالای خطا، کاهش تاخیر و افزایش نرخ انتقال با استفاده از سوئیچ‌های ارزان با حافظه کم می‌باشد. اجزاء الگوریتم: ۱. نشانگان ساده در سوئیچ ۲. ECN-Echo در گیرنده ۳. کنترل‌گر در فرستنده.

مزایا: DCTCP به سه مشکل اصلی پاسخ می‌دهد:

- تشکیل صف: DCTCP به سرعت واکنش نشان دار شدن صف را نشان می‌دهد. حجم بیشتری از فضای بافر به‌عنوان فضای اضافی برای جذب micro-bursts گذرا در دسترس است که به حداقل رساندن از دست‌رفتن پرده‌های پرهزینه که می‌تواند به توقف‌ها منجر شود، کمک می‌کند.
- فشار بافر: DCTCP مشکل فشار بافر را حل می‌کند زیرا طول صف یک پورت در وضعیت افتتاح زیاد شده به طور محدود می‌شود. بنابراین، در سوئیچ‌های حافظه مشترک، چندین پورت افتتاح شده منابع بافر را به هم نمی‌زند.
- انکست: DCTCP سناریوی انکست که تعداد زیادی از جریان‌های کوچک و هماهنگ به هم می‌خورند را کاهش می‌دهد. با نشان دار کردن زود هنگام (و مستقیم – بر اساس طول لحظه‌ای صف)، منابع DCTCP در یک یا دو زمان پاسخ تعداد کافی از نشان‌ها را دریافت می‌کنند تا اندازه انکست‌های بعدی را کنترل کنند.

عملکرد DCTCP: آزمایشات مبتنی بر میکروبنچمارک نشان می‌دهند که DCTCP با کارایی حداکثر متناسب با TCP به دست می‌آید. اختلاف اصلی در اینجا مربوط به طول صف در رابط گیرنده است. صف DCTCP در حدود ۲۰ بسته پایدار است، در حالی که صف TCP ۱۰ برابر بزرگتر است و به طور گسترده نوسان می‌کند. تنظیمات DCTCP با مقادیر توصیه شده ( $K=20$ )، حتی برای  $K$  به اندازه ۵ نیز، به بهره‌وری حساس نیست. در سرعت خطوط 10Gbps نیز، DCTCP با  $K=65$  به بهره‌وری مشابه TCP می‌رسد و به اندازه ۶۵ حساس نیست.

مقایسه با RED: در مقایسه با الگوریتم RED، مشاهده می‌شود که RED نوسانات گسترده‌ای در طول صف ایجاد می‌کند و اغلب نیاز به دو برابر حافظه دارد تا به بهره‌وری مشابه DCTCP برسد.

همگرایی: نتایج نشان می‌دهند که DCTCP به سرعت به حداکثر بهره‌وری ممکن متصل می‌شود. همه جریان‌ها به بهره‌وری متناسب با تقسیم حقیقی بهره‌وری می‌رسند. معرفی شده که DCTCP به TCP عادلانه نیست و در صورت نیاز، جریان‌های DCTCP می‌توانند از جریان‌های TCP جدا شوند.

شبکه‌های چندپرت: DCTCP با موفقیت در محیط چندپرت با چندین گره و بطن‌های مختلف عمل کرده و قابلیت مقابله با مسائل چندپرت و RTT متفاوت را نشان می‌دهد.