



**Internet measurement**

**Summary of the paper:**

**Queues don't matter**

**when you can JUMP them!**

**مینا فریدی**

**810100430**



QJUMP رویکردی ساده و سریع است که تداخل شبکه در شبکه های مراکز داده را کنترل می کند. ازدحام ناشی از برنامه های پر توان و باعث ایجاد صف می شود که ترافیک برنامه های حساس به تأخیر را به تأخیر می اندازد و تداخل شبکه زمانی اتفاق می افتد.

برای کاهش تداخل شبکه، QJUMP از تکنیک های الهام گرفته از QoS اینترنت برای برنامه های مرکز داده استفاده می کند. هر برنامه بر حسب به یک کلاس حساسیت اختصاص داده می شود. بسته های سطوح بالاتر در میزبان پایانی با نرخ محدود می شوند، اما پس از ورود به شبکه، می توانند روی بسته های سطوح پایین تر در صف بپرند. در تنظیماتی با تعداد گره ها و سرعت پیوند شناخته شده، QJUMP می تواند سطوح خدمات را از تأخیر کاملاً محدود (اما با نرخ پایین) تا توان عملیاتی نرخ خط (اما با واریانس تأخیر بالا) پشتیبانی کند.

تأخیر دم، به عنوان تأخیر با درصد بالا شناخته می شود، به تأخیرهای بالایی اشاره دارد که مشتریان نسبتاً به ندرت مشاهده می کنند. مواردی مانند: "سرویس من بیشتر در حدود 10 میلی ثانیه پاسخ می دهد، اما گاهی اوقات حدود 100 میلی ثانیه طول می کشد." دلایل زیادی برای تأخیر دم در جهان وجود دارد، از جمله مشاخره، جمع آوری زباله، از دست دادن بسته ها، خرابی میزبان و کارهای عجیب و غریبی که سیستم عامل ها در پس زمینه انجام می دهند.

در این مقاله میزان تاخیر سه برنامه مختلف اندازه گیری شده است:

#### 1. همگام سازی ساعت

همگام سازی دقیق ساعت برای سیستم های توزیع شده مانند Google's Spanner مهم است. PTPd همگام سازی زمانی با دانه بندی میکروثانیه را از یک سرور زمان به ماشین های موجود در یک شبکه محلی ارائه می دهد. در حالت مشترک، فازهای زدن Hadoop باعث ایجاد صف می شود که بسته های همگام سازی PTPd را به تأخیر می اندازد. این باعث می شود که PTPd به طور موقت 200-500  $\mu$ s از هماهنگی خارج شود.

#### 2. نگهداری key value

Memcached یک ذخیره سازی کلیدی-مقدار حافظه محبوب است که توسط فیس بوک و دیگران برای ذخیره اشیاء کوچک برای بازیابی سریع استفاده می شود. با اجرای Hadoop، تأخیر درخواست صدک 99 به میزان  $1.5 \times$  از 779  $\mu$ s به 1196  $\mu$ s کاهش می یابد.

#### 3. جریان داده های تکراری

Naiad چارچوبی برای محاسبات جریان داده توزیع شده است. در محاسبات تکراری، عملکرد Naiad به همگام سازی حالت تأخیر کم بین گره های کارگر بستگی دارد. در یک شبکه بیکار، Naiad حدود 500  $\mu$ s در صدک 99 طول می کشد تا یک همگام سازی مانع چهار طرفه را انجام دهد. با تداخل، این میزان به 1.1 تا 1.5 میلی ثانیه افزایش می یابد که کاهش عملکرد 2 تا 3 برابر است.



آزمایش‌های این مقاله نشان می‌دهد که برنامه‌ها به تداخل شبکه حساس هستند و تداخل شبکه عمدتاً در نتیجه صف‌های سوئیچ مشترک رخ می‌دهد. بنابراین QJUMP با کاهش صف سوئیچ با تداخل شبکه مقابله می‌کند: اساساً، اگر بتوانیم مقدار صف را در شبکه کاهش دهیم، تداخل شبکه را نیز کاهش خواهیم داد. در حالت شدید، اگر بتوان یک کران محدود و کم را در صف قرار داد، آنگاه می‌توان تداخل شبکه را به طور کامل کنترل کرد. این ایده اساس QJUMP را تشکیل می‌دهد.

این مقاله، یک مدل بصری برای قرار دادن چنین محدودیتی در صف در هر توپولوژی شبکه مرکز داده استخراج می‌کند. ابتدا یک مورد سوئیچ را در نظر می‌گیریم، قبل از اینکه مدل را برای پوشش چندین سوئیچ گسترش دهد. سپس محدودیت‌های توان عملیاتی مدل را کاهش داده و واریانس تأخیر در مقابل مبادله توان عملیاتی را کمی کم می‌کنند. در نهایت، توضیح می‌دهند که چگونه ترافیک حساس به تأخیر مجاز است به صف پرش بیش از ترافیک با توان بالا پردازد.

یک استقرار QJUMP به پنج پارامتر برای پیکربندی نیاز دارد:  $n$ ، تعداد میزبان‌ها،  $P$ ، حداکثر اندازه بسته،  $R$ ، نرخ کندترین پیوند لبه،  $\varepsilon$ ، تأخیر پردازش سوئیچ تجمعی لبه به لبه و  $\alpha$ ، کسر فرضی میزبان‌های ارسال کننده همزمان در هر سطح.

یکی از ویژگی‌های منحصر به فرد QJUMP، سطح تأخیر تضمین شده آن است تأخیر محدود، طرح‌های جدید جالبی را برای نرم‌افزار هماهنگ‌سازی مرکز داده‌ها مانند هواپیماهای کنترل SDN، تشخیص سریع شکست و سیستم‌های اجماع توزیع شده امکان‌پذیر می‌سازد.

از آنجایی که QJUMP تحویل قابل اعتمادی را ارائه می‌دهد، هماهنگ کننده می‌تواند پیام‌های خود را با پخش UDP هنگامی که QJUMP فعال است ارسال کند. این بهینه سازی 30٪ بهبود توان عملیاتی را نسبت به TCP و UDP به همراه دارد.

QJUMP همچنین تداخل شبکه تجربه شده توسط مقالات دیگر را حل می‌کند، که یک سیستم توزیع شده برای اجرای برنامه‌های داده موازی جریان داده در آن‌ها پیاده سازی شده است. در یک شبکه شبکه غیرفعال، 90٪ از همگام سازی‌ها بیش از 600 میکرو ثانیه طول نمی‌کشد. با ترافیک تداخلی از Hadoop، این مقدار به 1.2 میلی ثانیه دو برابر می‌شود. با این حال، زمانی که QJUMP فعال است، علیرغم اشتراک گذاری شبکه با Hadoop، توزیع به دقت توزیع پایه بی ادعا را دنبال می‌کند. QJUMP در اینجا 2 تا 5 برابر بهبود تأخیر در سطح برنامه را ارائه می‌دهد.

آزمایش‌های مقاله‌های قبلی این نویسنده نشان داد که تداخل شبکه، عملکرد برنامه را کاهش می‌دهد. اکنون آن آزمایش‌ها را با فعال کردن QJUMP تکرار می‌شوند و نشان داده می‌شود که QJUMP تداخل شبکه را کاهش می‌دهد و در نتیجه عملکرد تقریباً ایده‌آل را به همراه دارد. همچنین نشان می‌دهند که در یک تنظیمات چند کاربرد واقعی، QJUMP هم تداخل شبکه را برطرف می‌کند و هم از سایر سیستم‌های در دسترس بهتر عمل می‌کند.