

مدیریت منابع در شبکه‌های چندرسانه‌ای با استفاده از شبکه‌های نرم‌افزار محور

احمد رضا منتظرالقائم

تماس‌پذیرنده (یا UAS)، اطلاعات مورد نیاز برای ایجاد یک جلسه از طریق سیگنالینگ SIP مبادله می‌شود [۵]. سیگنالینگ SIP از طریق ارسال پیام‌های درخواست و پاسخ به وسیله پروکسی سرورهای SIP انجام می‌شود. مسیر پیام‌های درخواست و پاسخ از مسیر عبور مدیا مستقل است [۶]. سیگنالینگ SIP را به وسیله (۱) تا (۷) در این شکل نشان داده‌ایم. ارتباط (۸) نیز نشان‌دهنده عبور مدیا از سوئیچ‌های شبکه از طریق پروتکل RTP است [۷]. پس پیام‌های سیگنالینگ هم از سوئیچ‌های شبکه عبور می‌کنند و هم از پروکسی‌های SIP. اما پیام‌های مدیا فقط از سوئیچ‌های شبکه عبور می‌کنند. پیام‌ها برای رسیدن به مقصد (تماس‌پذیرنده) نیاز به مسیریابی دارند [۸].

پروکسی‌های SIP وظیفه مسیریابی لایه هفت پیام‌ها و سوئیچ‌های شبکه وظیفه مسیریابی لایه سه پیام‌ها را بر عهده دارند [۹]. عدم مسیریابی مناسب پیام‌ها سبب عدم استفاده درست از منابع پروکسی‌های SIP و همچنین سوئیچ‌های شبکه می‌شود [۵]. از آنجا که دو فاز مطرح‌شده مستقل از هم عمل می‌کنند، مسیریابی و مدیریت منابع نیز به صورت مستقل در دو سطح پروکسی‌های SIP و سوئیچ‌های شبکه انجام می‌گیرد [۱۰]. علاوه بر این که این نوع مسیریابی بین تمام سوئیچ‌ها و پروکسی‌ها توزیع شده است. این باعث عدم یک دید کلی از کل شبکه به عنوان یک شبکه واحد می‌شود که سبب پدیده اضافه‌بار می‌گردد و مستقیماً بر گذردهی شبکه تأثیرگذار است [۱۱].

۲- مفاهیم

در شکل ۲ مشاهده می‌شود که پدیده اضافه‌بار در شبکه‌های چندرسانه‌ای سبب شده که گذردهی شبکه به صفر نزول کند [۴]. پس نیاز به یک مدیریت منابع درست و متمرکز بسیار احساس می‌شود. به عبارت دیگر منابع در شبکه‌های چندرسانه‌ای به دو دسته منابع محاسباتی (پروکسی سرورهای SIP) و منابع شبکه‌ای (سوئیچ‌های شبکه) تقسیم می‌شوند که بایستی به صورت مرکزی مدیریت شوند تا مسیریابی مبتنی بر منابع به درستی انجام گردد.

همان‌طور که در شکل ۳ مشاهده می‌شود، شبکه‌های چندرسانه‌ای تشکیل‌شده از مجموعه‌ای از پروکسی‌های SIP هستند که پیام‌های سیگنالینگ را با استفاده از سرآیند لایه ۷، پرش به پرش تا مقصد مسیریابی می‌کنند [۱۲]. این شبکه، خود بر روی شبکه‌ای از سوئیچ‌های سنتی قرار دارد که وظیفه مسیریابی پیام‌ها توسط سرآیند لایه ۳ را دارند. متأسفانه در این ساختار سنتی، مسیریابی توزیع‌شده و بدون در نظر گرفتن یکپارچه منابع محاسباتی و شبکه‌ای انجام می‌شود [۱۳]، علاوه بر این که پیچیدگی زیادی دارد. همان‌طور که قبلاً ذکر شد، مسیریابی نامناسب باعث پدیده اضافه‌بار می‌شود که عواقب وخیمی برای شبکه دارد [۱۴].

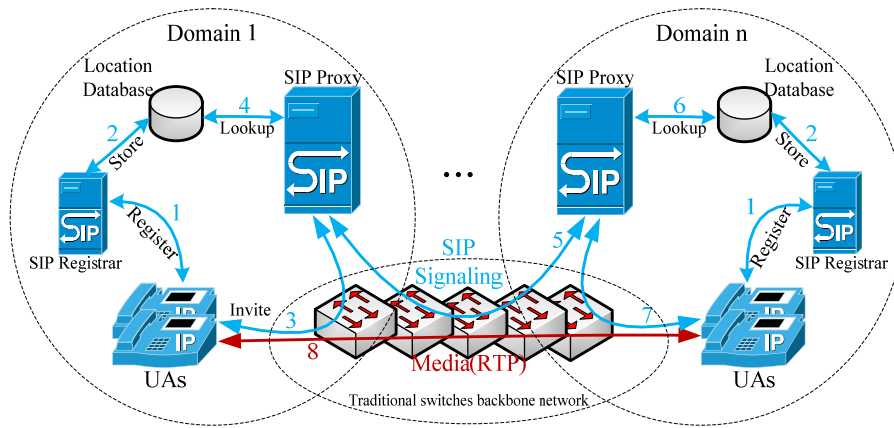
چکیده: امروزه شبکه‌های چندرسانه‌ای بر روی اینترنت به یک جایگزین کم‌هزینه و کارآمد برای PSTN تبدیل شده است. برنامه‌های کاربردی جهت انتقال مالتی‌مدیا بر روی بستر اینترنت روزبه‌روز فراگیرتر شده و به محبوبیت بسیار چشم‌گیری دست پیدا کرده‌اند. این ارتباط از دو فاز تشکیل شده است: فاز سیگنالینگ و فاز تبادل مدیا. فاز سیگنالینگ توسط پروکسی‌های SIP و فاز تبادل مدیا توسط سوئیچ‌های شبکه انجام می‌شود. از مهم‌ترین چالش‌ها در شبکه‌های چندرسانه‌ای، اضافه‌بار شدن پروکسی‌های SIP و سوئیچ‌های شبکه به ترتیب در فازهای سیگنالینگ و مدیا است. وجود این چالش سبب می‌شود که طیف وسیع کاربران شبکه با افت شدید کیفیت سرویس مواجه شوند. ما در این مقاله به مدل‌سازی مسئله مسیریابی در شبکه‌های چندرسانه‌ای جهت مقابله با اضافه‌بار می‌پردازیم. در این راستا یک روش مبتنی بر فناوری شبکه‌های نرم‌افزار محور و بر پایه یک مدل برنامه‌ریزی ریاضی محدد در شبکه‌های چندرسانه‌ای ارائه می‌کنیم. روش پیشنهادی تحت سناریوها و توپولوژی‌های متنوع شبیه‌سازی می‌گردد و نتایج نشان می‌دهند که گذردهی و مصرف منابع، بهبود یافته است.

کلیدواژه: شبکه‌های نرم‌افزار محور، کنترل اضافه‌بار، مدل‌سازی ریاضی، شبکه SIP، مدیریت منابع شبکه.

۱- مقدمه

امروزه علاقه‌مندی به شبکه‌های چندرسانه‌ای بر روی اینترنت، رو به افزایش است [۱]. گستردگی و تنوع سرویس‌های فراهم‌شده توسط شبکه‌های IP سبب روی آوردن فناوری‌های مختلف به یکپارچه و مجتمع سازی انواع شبکه‌ها و گرویدن به شبکه‌های چندرسانه‌ای بر روی اینترنت شده است. ارتباطات در این شبکه دو فاز اصلی دارد: فازهای سیگنالینگ و مدیا. پروتکل SIP [۲] برای سیگنالینگ مکالمات صوتی و تصویری از رشد روزافزونی برخوردار گردیده و به عنوان پروتکل سیگنالینگ پیشنهادی برای شبکه‌های نسل آینده در نظر گرفته شده است [۳]. همچنین روزبه‌روز فناوری‌های بی‌سیم نسل چهارم و پنجم از سوی شرکت‌های مخابراتی بیشتری مورد استفاده قرار می‌گیرد و در نتیجه در آینده‌ای نه‌چندان دور تقریباً تمامی تلفن‌ها و دستگاه‌های موبایل مجبور خواهند بود که برای جلسات چندرسانه‌ای خود، از پروتکل SIP به عنوان پروتکل سیگنالینگ پشتیبانی نمایند [۴]. پس از فاز سیگنالینگ، فاز مدیا برای تبادل مالتی‌مدیا است. شکل ۱ یک پیکربندی معمول SIP را نشان می‌دهد. پیش از ایجاد یک جلسه بین تماس‌گیرنده (یا UAC) و

این مقاله در تاریخ ۲۱ خرداد ماه ۱۳۹۹ دریافت و در تاریخ ۲۰ مهر ماه ۱۴۰۰ بازنگری شد.
احمد رضا منتظرالقائم، دانشکده مهندسی کامپیوتر، دانشگاه اصفهان، اصفهان، ایران،
(email: a.montazerolghaem@comp.ui.ac.ir)



شکل ۱: سیگنالینگ و مدیا جهت برقراری مالتی مدیا.

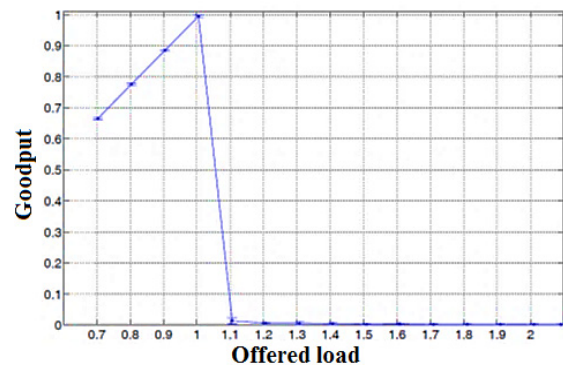
ساختار کلی شبکه را نگهداری می کند (صفحه کنترل). صفحه برنامه های کاربردی نیز شامل مجموعه ای از برنامه های کاربردی مانند مسیریابی، فایروال، تعادل بار، نظارت و غیره است. پروتکل ارتباطی بین صفحات، یک سری Open APIs استاندارد از جمله پروتکل OpenFlow است. در صورت وجود چنین رابط هایی، کنترل کننده قادر به برنامه ریزی دستگاه های ناهمگون شبکه به صورت پویا می باشد.

مشارکت های اصلی این مقاله در ادامه آمده است:

- پیش گیری از وقوع اضافه بار و توزیع بار در شبکه SIP از طریق مدیریت یکپارچه منابع
 - طراحی یک روش متمرکز مبتنی بر شبکه های نرم افزار محور جهت مسیریابی در دو سطح سوئیچ ها و پروکسی ها
 - مدل سازی ریاضی مسئله مسیریابی در شبکه های SIP در دو فاز سیگنالینگ و مدیا
 - شبیه سازی روش پیشنهادی در یک بستر آزمون واقعی و ارزیابی عملکرد تحت سناریوهای متنوع
- پس در این راستا اهداف ذیل کسب می شوند:
- ۱) افزایش گذردهی کل شبکه SIP
 - ۲) کاهش تأخیر برقراری تماس
 - ۳) مسیریابی در مسیرهای پروکسی ها و سوئیچ ها جهت توزیع بار

۳- کارهای پیشنهادی

عمده مقالات در این حوزه در تلاش برای حل مشکل اضافه بار از طریق باطراحی سیستم کنترلی اضافه بار سنتی هستند. طراحی مناسب شبکه و توزیع مناسب بار بین پروکسی های SIP یکی از عوامل جلوگیری از کارافتادن سرویس در شبکه های چندرسانه ای است [۱۸] تا [۲۳]. یک رویکرد برای مقابله با اضافه بار، توزیع ترافیک ورودی جدید بین سرورهای SIP بر اساس ظرفیت در دسترس آنها با استفاده از یک متعادل ساز بار است [۲۴] تا [۲۶]. بنابراین احتمال این که اضافه بار در سرور مشخصی اتفاق بیفتد، کاهش می یابد. اما عملکرد این روش ها وابسته به کارایی متعادل ساز بار است، چرا که تمام ترافیک سیگنالینگ شبکه SIP از آن عبور می کند. در نتیجه در اضافه بار شدید امکان وقوع اضافه بار در خود متعادل ساز بار نیز وجود دارد و باید با روش هایی ظرفیت آن را نیز بالا برد. در [۲۵] سه الگوریتم توزیع بار به نام های $CJSQ^1$ ، $TJSQ^2$ و



شکل ۲: افت گذردهی پروکسی SIP در مواجهه با اضافه بار [۴].

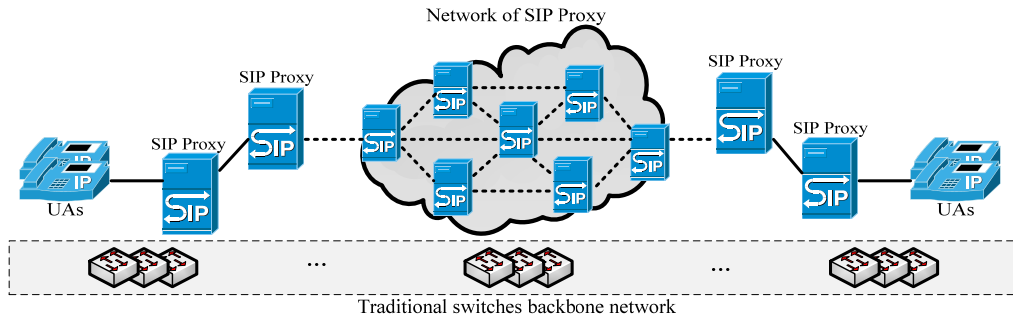
در این ساختار هر پروکسی SIP مستقلاً یک واحد کنترل اضافه بار شامل تابع کنترل بار، ناظر و بکارانداز دارد [۱۵]. پردازنده SIP هم وظیفه بررسی پیام های SIP و مسیریابی آنها را بر عهده دارد. در این ساختار، شناسایی و مقابله با اضافه بار با همکاری همه پروکسی ها اتفاق می افتد و آمار بار ترافیکی به صورت محلی بین پروکسی ها دست به دست می شود [۱۶]. عدم وجود یک کنترل کننده مرکزی در این ساختار باعث عدم آگاهی از توپولوژی و منابع کل شبکه می گردد [۱۷].

ما در این مقاله یک روش متمرکز کنترل اضافه بار را برای شبکه SIP توسط مفهوم شبکه های نرم افزار محور پیشنهاد می دهیم. این روش به مدیریت کلیه منابع محاسباتی و شبکه ای به صورت یکپارچه می پردازد. شبکه های نرم افزار محور یک معماری جدید در شبکه هستند و با استفاده از معماری کل شبکه و اجزای آن به صورت یک شبکه مجازی واحد دیده می شوند که با استفاده از نرم افزارها و API های طراحی شده قابل کنترل می گردند.

پس نیاز به یک پروتکل ارتباطی بین نرم افزارها و سخت افزارها مشهود است که پروتکل OpenFlow مطرح گردید. هر جزیی از شبکه که این پروتکل را پشتیبانی کند می تواند در شبکه نرم افزار محور قرار گیرد. معماری شبکه های نرم افزار محور دارای ویژگی های ارتقای پیکربندی، بهبود کارایی، پویایی، مدیریت یکپارچه منابع و توان، مدیریت ترافیک، قابلیت برنامه ریزی، کاهش هزینه ها و دسترس پذیری بالا است.

شکل ۴ نمایی از معماری شبکه های نرم افزار محور را نمایش می دهد که از سه قسمت اصلی صفحه کنترل، صفحه داده و صفحه برنامه های کاربردی تشکیل شده است. صفحه داده شامل دستگاه های شبکه از جمله دستگاه های مسیریابی، سوئیچ و ... است به طوری که فاقد بخش کنترلی و یا نرم افزاری جهت تصمیم گیری های خودکار می باشند. بخش هوشمند شبکه در کنترل کننده های نرم افزاری شبکه های نرم افزار محور قرار دارد که

1. Call-Join-Shortest-Queue
2. Transaction-Join-Shortest-Queue



شکل ۳: سیگنالینگ SIP در شبکه‌ای از پروکسی‌ها و سوئیچ‌ها.

در TLWL یک تماس جدید به سروری که کمترین کار را دارد، مسیره می‌شود. در اینجا کار (یا به عبارت دیگر بار) بر مبنای تخمین‌های نسبی از هزینه‌های تراکنش می‌باشد. به عبارت دیگر در این روش، شمارشگرها مجموع وزن‌دار تراکنش‌های تخصیص‌یافته به هر سرور را مشخص می‌کنند و تماس‌های جدید به سرورهایی که پایین‌ترین شمارشگر را دارند اختصاص می‌یابند. به عنوان مثال اگر یک سرور در حال پردازش یک تراکنش INVITE (با هزینه نسبی ۱/۷۵) و BYE (با هزینه نسبی ۱/۱۰) باشد، آن گاه بار آن سرور برابر ۲/۷۵ است.

در [۲۶] نیز یک متعادل‌ساز بار برای SIP با استفاده از تابع آمیزش^۲ ارائه شده که بر مبنای دریافت‌کننده تماس، درخواست‌ها به سرورها مسیره می‌شوند. در [۲۷] یک مدل افزونگی برای توزیع‌کننده‌های بار در شبکه SIP پیشنهاد شده و با استفاده از OpenAIS [۲۸] سلامت پروکسی‌های پایین‌دست، نظارت و خرابی یک پروکسی به سرعت کشف و پروکسی دیگری جایگزین آن می‌شود.

در [۲۹] ساختاری تحت عنوان Snocer برای توزیع بار در شبکه‌های SIP معرفی شده که در آن کاربر در صورت دسترسی به توزیع‌کننده بار به آن وصل می‌شود و در غیر این صورت از طریق DNS و یا به طور مستقیم به یکی از پروکسی‌های موجود در شبکه وصل می‌گردد.

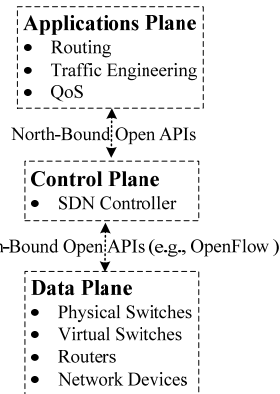
در [۲۸] و [۲۹] از سرورهای پشتیبان به عنوان افزونگی استفاده شده است. در [۳۰] انواع افزونگی به چهار دسته مبتنی بر مشتری، مبتنی بر DNS، مبتنی بر IP و مبتنی بر تکرار پایگاه داده، طبقه‌بندی گردیده و مورد مطالعه قرار گرفته است. همچنین انواع روش‌های تقسیم و توزیع بار با استفاده از DNS، شناسه SIP، پروکسی‌های با آدرس یکسان و مترجم آدرس شبکه، معرفی و کارایی آنها روی یک معماری دوسطحی متشکل از پروکسی‌های SIP ارزیابی شده است.

به عنوان نتیجه‌گیری این بخش می‌توان به این نکته اشاره کرد که تقریباً تمامی رویکردهای پیشین این حوزه توزیع‌شده هستند و همین توزیع‌شدگی باعث پیچیدگی، کاهش کارایی و همچنین عدم مدیریت یکپارچه و متمرکز منابع می‌شود. البته ذکر این نکته ضروری است که تحقیقات پژوهشی در این زمینه هنوز به شدت ادامه دارد.

۴- رویکرد پیشنهادی

۴-۱ مدل سیستم

در حالت کلی، شبکه SIP یک مجموعه n تایی از پروکسی‌های SIP به همراه یک مجموعه z تایی از سوئیچ‌ها با منابع پردازشی و حافظه‌ای محدود می‌باشد. هر پروکسی یا سوئیچ منابع خود را جهت برقراری جلسه بین کاربران محلی خود و یا کاربران دامنه‌های مختلف مصرف می‌کند. در

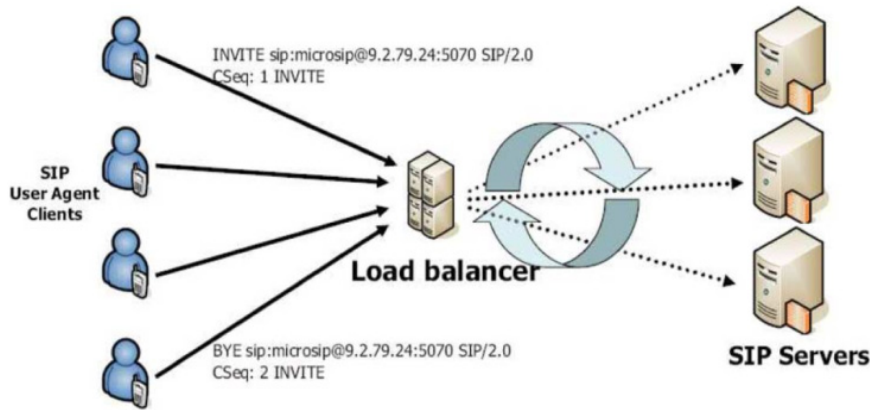


شکل ۴: معماری لایه‌ای در شبکه‌های نرم‌افزارمحور.

^۱ TLWL معرفی شده است. CJSQ بر مبنای تعداد جلسات تخصیص‌یافته به یک سرور، میزان کاری را که بر عهده آن سرور می‌باشد برآورد می‌کند. مطابق شکل ۵، متعادل‌ساز بار برای هر سرور یک شمارشگر جهت شمارش تعداد تماس‌های تخصیص‌یافته دارد. زمانی که یک درخواست INVITE جدید (که متناظر با یک تماس جدید است) دریافت می‌شود، این تقاضا به سروری که کمترین شمارشگر را دارد اختصاص و شمارشگر آن سرور یک عدد افزایش می‌یابد. زمانی که متعادل‌ساز بار برای یک درخواست BYE مربوط به این تماس، یک پاسخ OK ۲۰۰ دریافت کند، متوجه می‌شود که پردازش آن تماس توسط این سرور به اتمام رسیده و در نتیجه شمارشگر این سرور را یک عدد کاهش می‌دهد.

یکی از محدودیت‌های این روش این است که تعداد تماس‌های اختصاص‌یافته به یک سرور همواره سنجه دقیقی از بار آن سرور نیست. ممکن است بین تراکنش‌های یک تماس، زمان‌های توقف (بی‌باری) طولانی‌ای وجود داشته باشد. علاوه بر این، تماس‌های مختلف، ممکن است از تعداد متفاوتی تراکنش تشکیل شده باشند.

در TJSQ، متعادل‌ساز بار تماس جدید را به جای سرور با کمترین تماس‌های فعال، به سروری که کمترین تعداد تراکنش‌های فعال را دارد، مسیره می‌کند. در واقع شمارشگرها، تعداد تراکنش‌های تخصیص‌یافته به هر سرور را مشخص می‌کنند و تماس‌های جدید به سرورهایی که پایین‌ترین شمارشگر را دارند اختصاص می‌یابند. یکی از محدودیت‌های این روش این است که در آن تمامی تراکنش‌ها به طور برابر وزن‌دهی می‌شوند. در پروتکل SIP، تراکنش‌های INVITE از تراکنش‌های BYE پرهزینه‌تر هستند، زیرا ماشین حالت تراکنش INVITE نسبت به ماشین حالت تراکنش‌های BYE پیچیده‌تر است. در [۲۵] با یک سری آزمایش نشان داده شده که تراکنش‌های INVITE در حدود ۷۵٪ بیشتر از تراکنش‌های BYE هزینه دارند.



شکل ۵: توزیع بار بین سرورهای SIP توسط متعادل‌ساز بار [۲۵].

$$\max \gamma \left(\frac{\sum_{i=1}^n \sum_{j=1}^n C^{ij}}{\sum_{i=1}^n \sum_{j=1}^n C^{ij}} \right) + \rho \left(\frac{\sum_{i=1}^z \sum_{j=1}^z D^{ij}}{\sum_{i=1}^z \sum_{j=1}^z \mathbb{D}^{ij}} \right) - \phi \left(\frac{\sum_{g=1}^n P_g^\phi}{\sum_{g=1}^n P_g^\phi} + \frac{\sum_{g=1}^n m_g^\phi}{\sum_{g=1}^n M_g^\phi} \right) - \delta \left(\frac{\sum_{f=1}^z p_f^\theta}{\sum_{f=1}^z P_f^\theta} + \frac{\sum_{f=1}^z m_f^\theta}{\sum_{f=1}^z M_f^\theta} \right) \quad (1)$$

Subject to:

Proxies Constraints:

$$C^{ij} \leq \mathbb{C}^{ij}, \quad \forall i, j \in R \quad (2)$$

$$\sum_{k=1}^n L_{kv} X_{kv}^{ij} = \sum_{u=1}^n L_{vu} X_{vu}^{ij}, \quad \forall i, j, v \in R, i \neq v, j \neq v \quad (3)$$

$$\sum_{k=1}^n L_{kv} X_{kv}^{iv} = C^{iv}, \quad \forall i, v \in R, i \neq v \quad (4)$$

$$\sum_{u=1}^n L_{vu} X_{vu}^{vj} = C^{vj}, \quad \forall v, j \in R, j \neq v \quad (5)$$

$$X_{kv}^{ii} = 0, \quad \forall i, k, v \in R \quad (6)$$

$$X_{ki}^{jj} = 0, \quad \forall i, j, k \in R \quad (7)$$

$$\alpha_v C^{vv} + \alpha_v \left(\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n L_{vk} X_{vk}^{ij} + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n L_{kv} X_{kv}^{ij} \right) \leq P_v^\phi, \quad \forall v \in R \quad (8)$$

$$\beta_v C^{vv} + \beta_v \left(\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n L_{vk} X_{vk}^{ij} + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n L_{kv} X_{kv}^{ij} \right) \leq m_v^\phi, \quad \forall v \in R \quad (9)$$

$$P_v^\phi \leq P_v^\theta, \quad \forall v \in R \quad (10)$$

$$m_v^\phi \leq M_v^\theta, \quad \forall v \in R \quad (11)$$

Switches Constraints:

$$D^{ij} \leq \mathbb{D}^{ij}, \quad \forall i, j \in S \quad (12)$$

$$\sum_{k=1}^z L_{kv} X_{kv}^{ij} = \sum_{u=1}^z L_{vu} X_{vu}^{ij}, \quad \forall i, j, v \in S, i \neq v, j \neq v \quad (13)$$

$$\sum_{k=1}^z L_{kv} X_{kv}^{iv} = D^{iv}, \quad \forall i, v \in S, i \neq v \quad (14)$$

این مدل سیستم، فرض شده که ماتریس باینری و متقارن L_{kl} توپولوژی شبکه SIP را در بر دارد. در این ماتریس اگر L_{kl} برابر یک باشد به معنی وجود لینک فیزیکی بین نود k و l است و در غیر این صورت، عدم وجود لینک فیزیکی بین دو نود را نشان می‌دهد. قطر اصلی این ماتریس صفر می‌باشد.

همچنین ماتریس مربعی \mathbb{C}^{ij} نشان‌دهنده تعداد درخواست تماس از کاربران پروکسی i به کاربران پروکسی j می‌باشد. در این صورت \mathbb{C}^{ii} تعداد درخواست‌های محلی پروکسی i و \mathbb{C}^{ij} تعداد درخواست‌های دامنه‌ای پروکسی i می‌باشد. فرض می‌کنیم مقدار بهینه تعداد تماس‌های پذیرفته‌شده از پروکسی i به پروکسی j ، C^{ij} باشد به طوری که $C^{ij} \leq \mathbb{C}^{ij}$ است.

همچنین \mathbb{D}^{ij} نشان‌دهنده تعداد درخواست‌های سیگنالینگ از سوئیچ i به سوئیچ j می‌باشد و بهینه آن D^{ij} است. تعداد درخواست‌های مدیا بین دو سوئیچ i و j نیز \mathbb{D}^{ij} و بهینه آن D^{ij} می‌باشد. همچنین X_{kv}^{ij} نیز نشان‌دهنده تعداد درخواست‌های پذیرفته‌شده از مبدأ i به مقصد j است که از دو پروکسی مجاور k و v می‌گذرد ($X_{kv}^{ij} \leq \mathbb{C}^{ij}$). بنابراین تعداد کل درخواست‌های پذیرفته‌شده از i به j با C^{ij} مشخص می‌گردد به طوری که X_{kv}^{ij} نحوه توزیع آن بین مسیرهای موجود بین i و j را نشان می‌دهد. ما در این مقاله X_{kv}^{ij} را جریان مبدأ i به مقصد j که از مسیر k و v می‌گذرد، تعریف می‌کنیم.

هر پروکسی یا سوئیچ برای انجام عملیات خود متکی به میزان باقیمانده منابع خود یعنی پردازنده و حافظه می‌باشد. مقدار باقیمانده پردازنده و حافظه پروکسی یا سوئیچ v را به ترتیب با P_v^θ ، M_v^θ و P_v^ϕ نشان می‌دهیم.

کنترل‌کننده وظیفه مدیریت این منابع را بر عهده دارد. در ابتدای هر بازه زمانی τ ، هر پروکسی و سوئیچ تعداد درخواست‌های جدید خود را به همراه مقادیر P_i ، M_i از طریق پروتکل OpenFlow به کنترل‌کننده ارسال می‌کنند (جمع‌آوری اطلاعات در ابتدای زمان τ). کنترل‌کننده با توجه به ماتریس L_{kl} و اطلاعات دریافتی و با استفاده از مدل برنامه‌ریزی خطی معرفی‌شده در زیر، مقادیر بهینه پارامترهای C^{ij} ، D^{ij} ، X_{kl}^{ij} ، P_v^ϕ ، m_v^ϕ ، P_v^θ و m_v^θ را محاسبه (محاسبات) می‌کند و سپس به پروکسی‌ها جهت اعمال، اطلاع می‌دهد (اطلاع‌رسانی).

۴-۲ فاز سیگنالینگ

در ادامه یک مدل خطی چندهدفه برای مسئله مسیریابی پیام‌های سیگنالینگ با توجه به محدودیت‌های مطرح‌شده ارائه می‌دهیم

هستند. از آنجا که طرفین (۱۸) و (۱۹) هم‌جنس نیستند، از ضرایب ω و ψ برای نرمالایز کردن این معادلات استفاده شده است. نهایتاً مدل فوق بدون وقوع اضافه‌بار به دنبال افزایش مجموع گذردهی شبکه از طریق تخصیص بهینه منابع می‌باشد. مجموع گذردهی را می‌توان به صورت مجموع جریان‌هایی که از شبکه عبور می‌کند، در نظر گرفت. پس با توجه به منابع، کنترل‌کننده سعی در حداکثر نمودن پذیرش تماس‌ها دارد به نحوی که بتواند آنها را بین پروکسی‌ها و سوئیچ‌های شبکه توزیع کند.

۴-۳ فاز مدیا

فاز بعدی کنترل‌کننده، مسیریابی مدیا در بین سوئیچ‌های شبکه می‌باشد و بدین جهت مدل خطی زیر پیشنهاد شده است. هدف، افزایش گذردهی مدیا در شبکه سوئیچ‌ها با حداقل مصرف منابع می‌باشد

$$\max \zeta \left(\frac{\sum_{i=1}^{\bar{z}} \sum_{j=1}^{\bar{z}} \bar{D}^{ij}}{\sum_{i=1}^{\bar{z}} \sum_{j=1}^{\bar{z}} \bar{\mathbb{D}}^{ij}} \right) - \zeta \left(\frac{\sum_{f=1}^{\bar{z}} \bar{P}_f^{\theta}}{\sum_{f=1}^{\bar{z}} \bar{P}_f^{\theta}} + \frac{\sum_{f=1}^{\bar{z}} \bar{m}_f^{\theta}}{\sum_{f=1}^{\bar{z}} \bar{M}_f^{\theta}} \right) \quad (22)$$

Subject to:

Switches Constraints:

$$\bar{D}^{ij} \leq \bar{\mathbb{D}}^{ij}, \quad \forall i, j \in S \quad (23)$$

$$\sum_{s=1}^{\bar{z}} L_{sl} Y_{sl}^{ij} = \sum_{e=1}^{\bar{z}} L_{le} Y_{le}^{ij}, \quad \forall i, j, l \in S, i \neq l, j \neq l \quad (24)$$

$$\sum_{s=1}^{\bar{z}} L_{sl} Y_{sl}^{il} = \bar{D}^{il}, \quad \forall i, l \in S, i \neq l \quad (25)$$

$$\sum_{e=1}^{\bar{z}} L_{le} Y_{le}^{ij} = \bar{D}^{ij}, \quad \forall l, j \in S, j \neq l \quad (26)$$

$$Y_{sl}^{ii} = 0, \quad \forall i, k, l \in S \quad (27)$$

$$Y_{si}^{ij} = 0, \quad \forall i, j, s \in S \quad (28)$$

$$\xi \left(\sum_{i=1}^{\bar{z}} \sum_{j=1}^{\bar{z}} \sum_{s=1}^{\bar{z}} L_{ls} Y_{ls}^{ij} + \sum_{i=1}^{\bar{z}} \sum_{j=1}^{\bar{z}} \sum_{s=1}^{\bar{z}} L_{sl} Y_{sl}^{ij} \right) \leq \bar{P}_l^{\theta}, \quad \forall l \in S \quad (29)$$

$$\eta \left(\sum_{i=1}^{\bar{z}} \sum_{j=1}^{\bar{z}} \sum_{s=1}^{\bar{z}} L_{ls} Y_{ls}^{ij} + \sum_{i=1}^{\bar{z}} \sum_{j=1}^{\bar{z}} \sum_{s=1}^{\bar{z}} L_{sl} Y_{sl}^{ij} \right) \leq \bar{m}_l^{\theta}, \quad \forall l \in S \quad (30)$$

$$\bar{p}_l^{\theta} \leq \bar{P}_l^{\theta}, \quad \forall l \in S \quad (31)$$

$$\bar{m}_l^{\theta} \leq \bar{M}_l^{\theta}, \quad \forall l \in S \quad (32)$$

$$\text{Variables: } \bar{D}^{ij}, Y_{sl}^{ij}, \bar{p}_l^{\theta}, \bar{m}_l^{\theta} \geq 0, \quad \forall i, j, s, l$$

در (۲۲) ضرایب ξ و η موازنه بین گذردهی مدیا و منابع ایجاد می‌کنند. اگر ξ از η بیشتر باشد آن گاه وزن بیشتر با افزایش گذردهی است اما اگر η از ξ بیشتر باشد آن گاه کاهش مصرف منابع، وزن بیشتری دارد. قید (۲۳) حداکثر ظرفیت سوئیچ‌ها برای تبادل مدیا را مشخص می‌کند. قیود (۲۴) تا (۲۸) برای مسیریابی مدیا با توجه به ترافیک مدیا (Y_{le}^{ij}) است. قیود (۲۹) تا (۳۲) نیز منابع را به مسیریابی مدیا تخصیص می‌دهند. ضرایب ξ و η جهت نرمالایز کردن (۲۹) و (۳۰) استفاده می‌شوند.

۵- ارزیابی کارایی

به منظور ارزیابی متد پیشنهادی، ۳ سناریوی کم‌بار، بار متوسط و پر بار

$$\sum_{u=1}^{\bar{z}} L_{vu} X_{vu}^{vj} = D^{vj}, \quad \forall v, j \in S, j \neq v \quad (15)$$

$$X_{kv}^{ii} = 0, \quad \forall i, k, v \in S \quad (16)$$

$$X_{ki}^{ij} = 0, \quad \forall i, j, k \in S \quad (17)$$

$$\omega \left(\sum_{i=1}^{\bar{z}} \sum_{j=1}^{\bar{z}} \sum_{k=1}^{\bar{z}} L_{vk} X_{vk}^{ij} + \sum_{i=1}^{\bar{z}} \sum_{j=1}^{\bar{z}} \sum_{k=1}^{\bar{z}} L_{kv} X_{kv}^{ij} \right) \leq P_v^{\theta}, \quad \forall v \in S \quad (18)$$

$$\psi \left(\sum_{i=1}^{\bar{z}} \sum_{j=1}^{\bar{z}} \sum_{k=1}^{\bar{z}} L_{vk} X_{vk}^{ij} + \sum_{i=1}^{\bar{z}} \sum_{j=1}^{\bar{z}} \sum_{k=1}^{\bar{z}} L_{kv} X_{kv}^{ij} \right) \leq m_v^{\theta}, \quad \forall v \in S \quad (19)$$

$$P_v^{\theta} \leq P_v^{\theta}, \quad \forall v \in S \quad (20)$$

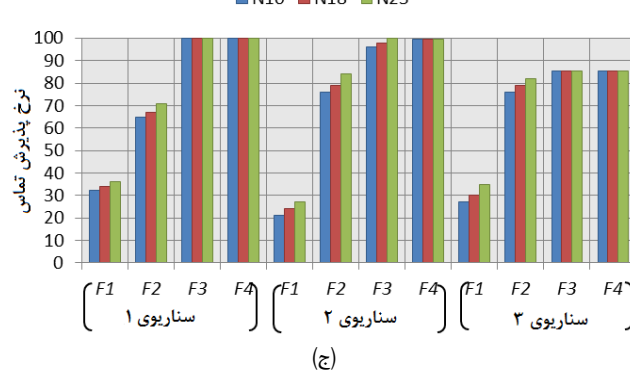
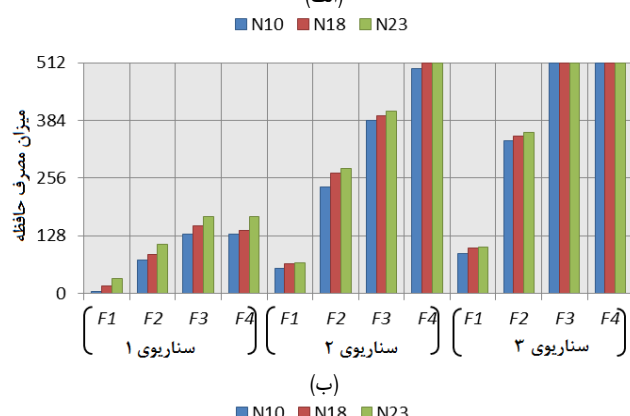
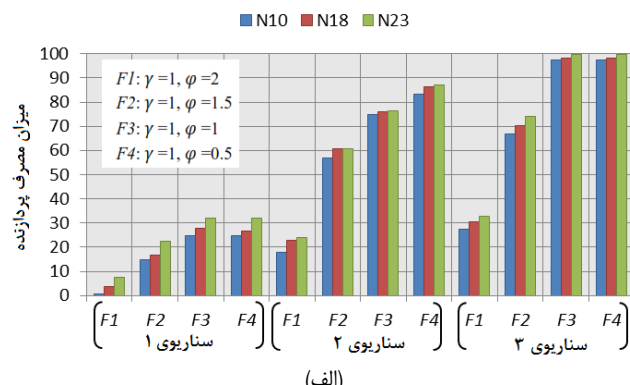
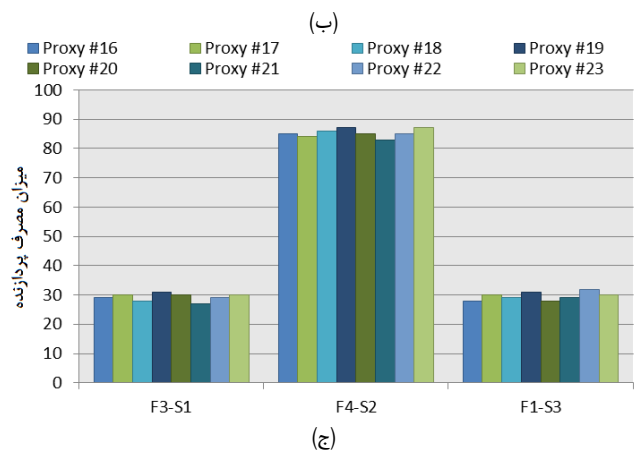
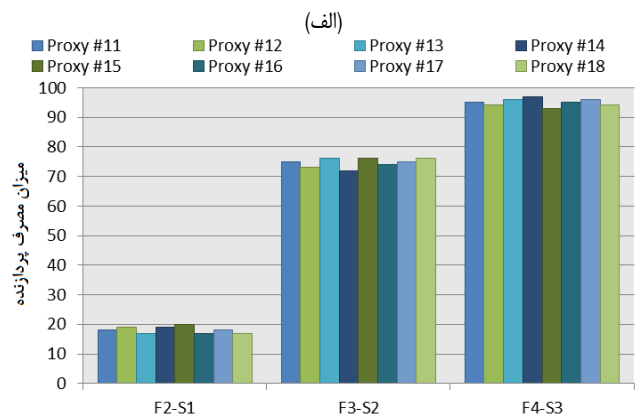
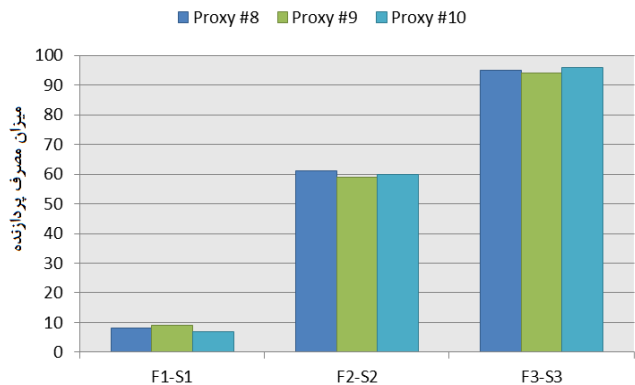
$$m_v^{\theta} \leq M_v^{\theta}, \quad \forall v \in S \quad (21)$$

$$\text{Variables: } C^{ij}, D^{ij}, X_{kl}^{ij}, P_v^{\theta}, m_v^{\theta}, p_v^{\theta}, m_v^{\theta} \geq 0, \quad \forall i, j, k, v$$

در مدل فوق، هدف، ماکسیم کردن گذردهی پروکسی‌ها و سوئیچ‌ها و همچنین کاهش مصرف منابع آنهاست. C^{ij}/\mathbb{C}^{ij} تعداد درخواست‌های پذیرفته به کل درخواست‌های پروکسی و D^{ij}/\mathbb{D}^{ij} تعداد درخواست‌های پذیرفته به کل درخواست‌های سوئیچ را نشان می‌دهد. ترم‌های سوم و چهارم در تابع هدف به ترتیب نشان‌دهنده منابع مصرفی (پردازنده و حافظه) پروکسی‌ها و سوئیچ‌ها هستند که به دنبال کاهش آنها هستیم. ضرایب γ, ρ, φ و δ درجه تأثیر گذردهی یا منابع در تابع هدف را نشان می‌دهد. با این ضرایب می‌توان یک موازنه بین گذردهی و منابع ایجاد کرد، به این معنی که اگر ضرایب مربوط به گذردهی بیشتر باشد (ρ و γ)، آن گاه تابع هدف با صرف منابع بیشتر، به دنبال ماکسیم کردن گذردهی است. برعکس اگر ضرایب مربوط به منابع بیشتر باشد (δ و φ)، آن گاه تابع هدف به دنبال کاهش مصرف منابع است.

محدودیت‌ها به ۲ دسته کلی محدودیت‌های پروکسی‌ها و سوئیچ‌ها تقسیم شده است. قید (۲) نشان می‌دهد که حداکثر می‌توان به تعداد درخواست‌ها، پذیرش داشت ($C^{ij} \leq \mathbb{C}^{ij}$). قید (۳) در هر پروکسی بین جریان‌های ورودی و خروجی توازن برقرار می‌کند، یعنی مجموع جریان‌های ورودی و خروجی به پروکسی v که از مبدأ i به مقصد j می‌باشند، باید برابر باشند. قید (۴) مجموع جریان‌های ورودی به پروکسی v از مبدأ i که از پروکسی‌های بین مسیر k می‌گذرد را ملزم می‌کند که برابر با C^{iv} باشد. قید بعدی نیز مجموع جریان‌های خروجی از سرور v به مقصد j را بین سرورهای هم‌جوار سرور v توزیع می‌کند. قید (۶) امکان ایجاد جریانی با مبدأ و مقصد یکسان را نخواهد داد. قید (۷) نیز از ایجاد دورها در مسیر جلوگیری می‌نماید. قیود (۸) تا (۱۱) قیود تخصیص منابع پروکسی‌ها هستند. قید (۸) نشان می‌دهد که پردازنده پروکسی‌ها با ضرایب α_1 و α_2 صرف رسیدگی کردن تماس‌های محلی و خارج دامنه‌ای می‌شود. قید (۹) نیز نشان می‌دهد که حافظه پروکسی‌ها با ضرایب β_1 و β_2 صرف رسیدگی کردن تماس‌های محلی و خارج دامنه‌ای می‌شود.

حدود بالای منابع پردازشی و حافظه‌ای مصرفی بهینه در مدل بالا P_v و m_v هستند که بنا بر قیود (۱۰) و (۱۱) حداکثر می‌توانند P_v و M_v باشند. قید (۱۲) پذیرش درخواست‌ها توسط سوئیچ‌ها را محدود به ظرفیت آنها می‌کند. قیود (۱۳) تا (۱۷) جهت مسیریابی ترافیک سیگنالینگ در سوئیچ‌ها می‌باشد. قیود (۱۸) تا (۲۱) نیز جهت مدیریت منابع سوئیچ‌ها



شکل ۶: پردازنده و حافظه مصرفی به همراه نرخ پذیرش مربوط به فاز سیگنالینگ در پروکسی‌ها، (الف) میزان مصرف پردازنده پروکسی‌ها (p_i^e) ، (ب) میزان مصرف حافظه پروکسی‌ها (m_i^e) و (ج) نرخ پذیرش تماس‌ها توسط پروکسی‌ها (C^{ij}) .

شکل ۷: پردازنده مصرفی پروکسی‌ها در سه توپولوژی در فاز سیگنالینگ، (الف) مصرف پردازنده پروکسی‌ها در توپولوژی N_{10} ، (ب) مصرف پردازنده پروکسی‌ها در توپولوژی N_{18} و (ج) مصرف پردازنده پروکسی‌ها در توپولوژی N_{23} .

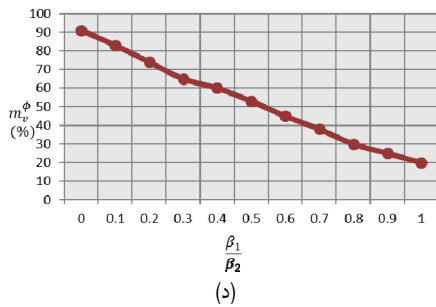
رشد و درصد پذیرش تماس‌ها نیز افزایش می‌یابد، چرا که نسبت ضرایب γ و ρ به یکدیگر نشان‌دهنده اهمیت پذیرش تماس‌ها یا حفظ منابع می‌باشد. در F^1 حفظ منابع نسبت به F^4 بسیار پراهمیت‌تر بوده و حتی باعث می‌گردد درصدی از تماس‌ها بلاک شوند. برعکس در حالت F^4 پذیرش هرچه بیشتر تماس‌ها مهم است، حتی اگر به قیمت صرف منابع بیشتر باشد. می‌توان با ایجاد یک توازن بین این ضرایب هم درصد پذیرش تماس بالایی داشت و هم از منابع به صورت مناسبی استفاده کرد. با حرکت از سناریوی ۱ (کم‌بار) به سناریوی ۳ (پربار)، منابع مصرفی رشد می‌یابند اما ممکن است بار ورودی به شبکه آن قدر زیاد شود که حتی با صرف تمام منابع شبکه نیز نتوان به تمام تماس‌های ورودی پاسخگو بود (سناریوی ۳ حالت F^3 و F^4).

شکل ۷ مصرف پردازنده پروکسی‌ها در سه توپولوژی را نشان می‌دهد. در این شکل‌ها ترکیب‌های مختلفی از حالت F و سناریو در نظر گرفته شده است (مثلاً F^1S^1 یا F^4S^2). همان طور که مشاهده می‌شود، مستقل از توپولوژی و سناریو، بار بین پروکسی‌ها، بالانس توزیع شده است.

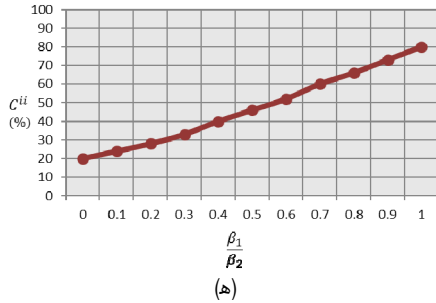
در نرم‌افزار Matlab استفاده شده است. سناریوها $(\mathbb{C}^{ij}, \mathbb{D}^{ij}, \bar{\mathbb{D}}^{ij})$ تصادفی با توزیع نرمال هستند. همچنین هر سناریو با حالت‌های متفاوت γ و ρ در فاز سیگنالینگ (که با F^1 تا F^4 نمایش می‌دهیم) مورد بررسی قرار گرفته است. مقادیر P_v ، M_v و E_v را برای تمامی پروکسی‌ها و سوئیچ‌ها در ابتدا به ترتیب ۱۰۰، ۵۱۲ و ۳۵۰ در نظر گرفتیم. همچنین توپولوژی‌های N_{10} ، N_{18} و N_{23} نشان‌دهنده تعداد نودهای شبکه می‌باشند (به ترتیب ۱۰، ۱۸ و ۲۳). در ادامه نتایج ارائه می‌گردد.

۵-۱ آزمایش اول: فاز سیگنالینگ

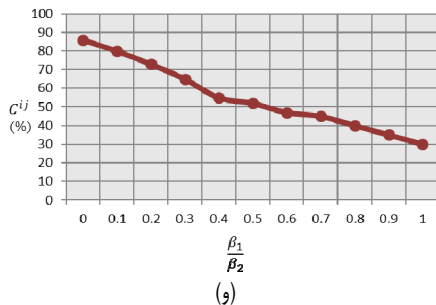
در این آزمایش به بررسی عملکرد پروکسی‌ها و سوئیچ‌ها در فاز سیگنالینگ می‌پردازیم. در آزمون‌های این بخش، فرض گردیده است که $\alpha_1 = 0.02168$ ، $\alpha_r = 0.07169$ و $\beta_1 = 0.02329$ و همچنین $\beta_r = 0.08012$ ، $\omega = 0.06345$ و $\psi = 0.04635$ است. پروکسی: در شکل‌های ۶ مقادیر بهینه p_v^e ، m_v^e و C^{ij} برای حالات مختلف F در سه سناریو و سه توپولوژی نشان داده شده است. در هر سه سناریو و در هر سه توپولوژی با تغییر حالت از F^1 تا F^4 ، منابع مصرفی



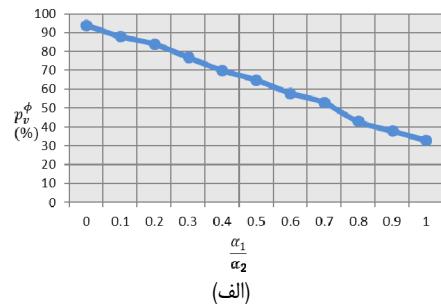
(د)



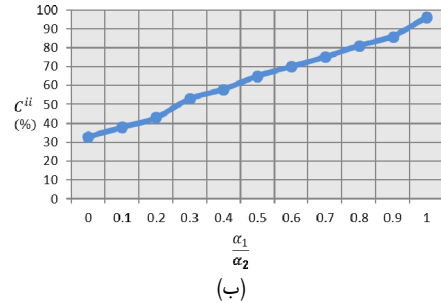
(ه)



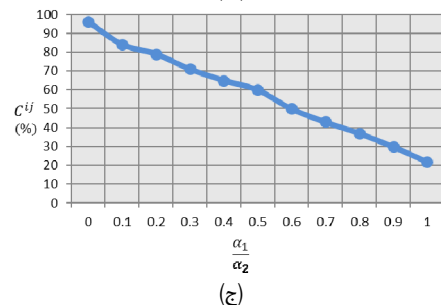
(و)



(الف)



(ب)



(ج)

شکل ۸: تأثیر α و β بر m_r^ϕ ، p_r^ϕ ، C^{ii} و C^{ij} پروکسی‌ها، (الف) تأثیر α_r/α بر پردازنده، (ب) تأثیر α_r/α بر تماس‌های محلی، (ج) تأثیر α_r/α بر تماس‌های خارج دامنه‌ای، (د) تأثیر β_r/β بر حافظه، (ه) تأثیر β_r/β بر تماس‌های محلی و (و) تأثیر β_r/β بر تماس‌های خارج دامنه‌ای.

گسترده‌ای از سوئیچ‌های OpenFlow را به دلیل داشتن دید کلی مدیریت کند.

پردازنده و حافظه مصرفی سوئیچ‌ها در فاز سیگنالینگ برای توپولوژی‌های N_{10} و N_{18} به ترتیب در جداول ۱ و ۲ آورده شده است. همان طور که در این جدول‌ها مشاهده می‌شود، منابع مصرفی سوئیچ‌ها بسیار نزدیک به هم هستند. به عنوان مثال در سناریوی ۱ و در حالت G^4 ، پردازنده مصرفی هر هفت سوئیچ حدوداً ۲۲٪ است که این روند برای حافظه مصرفی هم صادق می‌باشد. این نشان می‌دهد که بار به صورت عادلانه و یکنواخت بین سوئیچ‌ها توزیع شده است. نتایج برای توپولوژی N_{23} نیز از همین الگو پیروی می‌کند و از آوردن مجدد آن صرف نظر کردیم. فاز بعدی، فاز مدیا است.

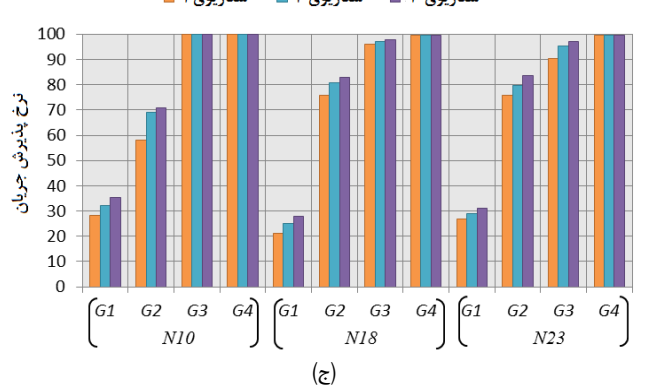
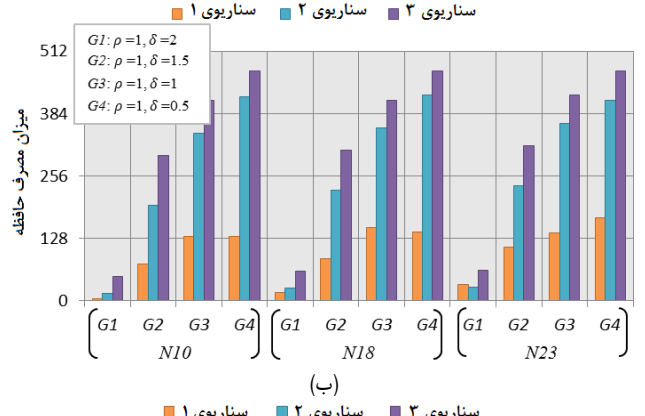
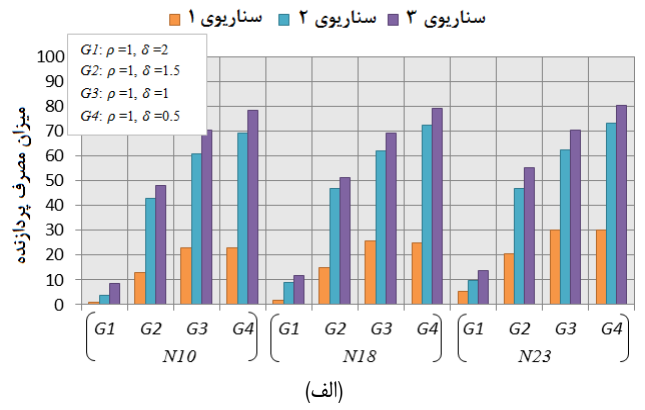
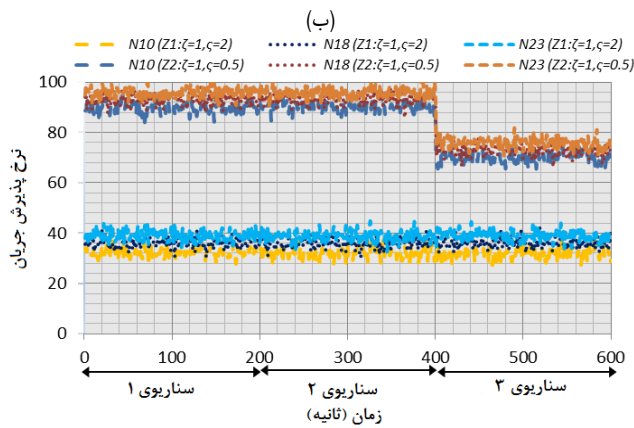
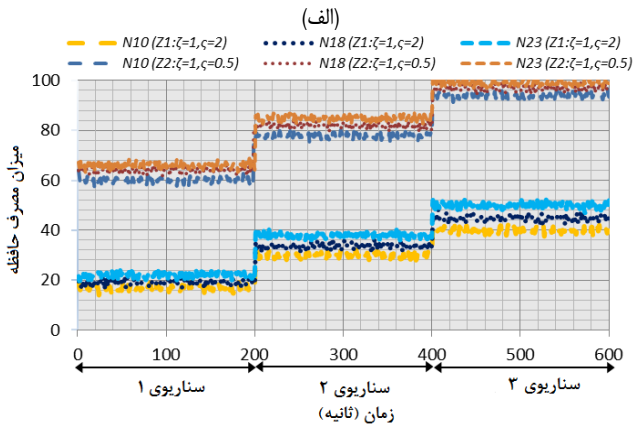
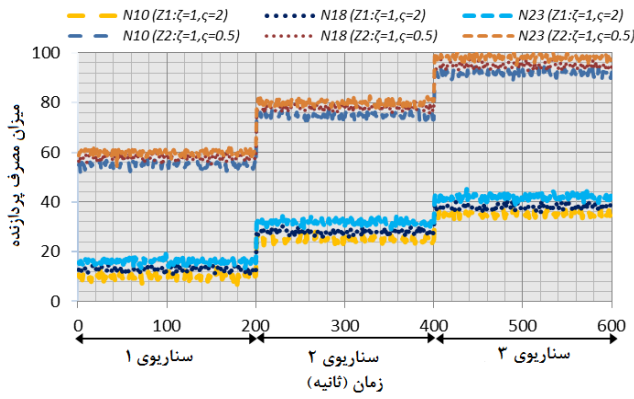
۵-۲ آزمایش دوم: فاز مدیا

در این فاز سوئیچ‌ها درگیر هستند و پروکسی‌ها نقشی ندارند. نتایج این فاز برای سه سناریوی کم‌بار، بار متوسط و پر بار در شکل ۱۰ آورده شده است. Z تابعی از k و ζ است. در Z_1 مقدار ζ (حفظ منابع) و در Z_2 مقدار k (نرخ پذیرش) بیشتر است. مقادیر بهینه \bar{p}_i^θ ، \bar{m}_i^θ و \bar{D}^j در این شکل برای مدت ۶۰۰ ثانیه آورده شده است. در ۲۰۰ ثانیه اول، سناریوی کم‌بار اجرا شده و سپس در ۲۰۰ ثانیه دوم و سوم به ترتیب سناریوهای بار متوسط و پر بار اجرا گشته است. همان طور که مشاهده می‌شود با افزایش بار در گذر زمان، منابع مصرفی (\bar{m}_i^θ و \bar{p}_i^θ) نیز به جهت ثابت نگه داشتن نرخ پذیرش جریان، افزایش می‌یابند. این افزایش با Z_2 بیشتر است به طوری که در سناریوی سوم با مصرف کل منابع (شکل ۱۰-الف و ۱۰-ب)، نرخ پذیرشی بهتر از ۷۸٪ به دست نمی‌آید.

زیرا کنترلر از منابع باقیمانده پروکسی‌ها آگاه بوده و با توجه به آن در مورد توزیع بار و مسیر ترافیک تصمیم‌گیری می‌کند. همین روند برای حافظه نیز صادق است و از ارائه نتایج آن صرف نظر می‌کنیم.

شکل‌های ۸ به بررسی اثر پارامترهای α و β بر منابع پردازشی، حافظه‌ای، تماس‌های محلی و تماس‌های خارج دامنه‌ای می‌پردازد. همان طور که در شکل‌های ۸-الف تا ۸-ج مشخص است، با افزایش α_r نسبت به α ، تماس‌های محلی پذیرفته‌شده افزایش و تماس‌های خارج دامنه‌ای کاهش می‌یابد. همچنین پردازنده مصرفی نیز کاهش می‌یابد. دلیل آن این است که در (۸)، α_r ضریب C^{ii} است و همچنین رسیدگی کردن به تماس‌های محلی منابع کمتری را نیاز دارد. از آنجا که در معادله تخصیص حافظه (۹)، β_r ضریب C^{ii} است پس افزایش β_r باعث افزایش تماس‌های محلی پذیرفته‌شده می‌شود (شکل ۸-ه). افزایش تماس‌های محلی نسبت به تماس‌های خارج دامنه‌ای به حافظه کمتری نیاز دارد (شکل ۸-د).

سوئیچ: در شکل‌های ۹ مقادیر بهینه p_r^θ ، m_r^θ و D^j برای حالات مختلف تابع G و همچنین سه سناریو و سه توپولوژی نشان داده شده است. G یک تابع از ضرایب ρ و δ است که وزن "پذیرش جریان‌ها" و "منابع سوئیچ‌ها" در تابع هدف (۱) هستند. همان طور که در این شکل‌ها مشاهده می‌شود، در هر سه سناریو با حرکت از G_1 تا G_4 منابع مصرفی رشد و درصد پذیرش تماس‌ها نیز افزایش می‌یابد، چون ضریب ρ در G_4 بیشتر می‌شود. همچنین با حرکت از سناریوی ۱ (کم‌بار) به سناریوی ۳ (پر بار)، منابع مصرفی رشد می‌یابند. اما نتایج مستقل از توپولوژی است و با پیچیده‌شدن توپولوژی عملکرد کنترلر تحت تأثیر قرار نمی‌گیرد. این گویای این است که شبکه نرم‌افزارمحور می‌تواند شبکه



شکل ۹: پردازنده و حافظه مصرفی و نرخ پذیرش در فاز سیگنالینگ سوئیچها، (الف) میزان مصرف پردازنده سوئیچها (P_v^0) ، (ب) میزان مصرف حافظه سوئیچها (m_v^0) و (ج) نرخ پذیرش جریان سیگنالینگ توسط سوئیچها (D^j) .

شکل ۱۰: پردازنده و حافظه مصرفی سوئیچها و نرخ پذیرش در فاز مدیا، (الف) میزان مصرف پردازنده سوئیچها (P_v^0) ، (ب) میزان مصرف حافظه سوئیچها (m_v^0) و (ج) نرخ پذیرش جریان مدیا توسط سوئیچها (D^j) .

در نتیجه نرخ پذیرش تماس توسط سوئیچها بالاست. با این حال همان گونه که در شکل ۱۲-ب و ۱۲-ج مشخص است، میانگین پردازنده و حافظه مصرفی در حالت τ برابر ۸ ثانیه، کمتر از سایر موارد است. دلیل این امر این است که هرچه τ کمتر باشد، سربار کنترلر نیز بیشتر است، چرا که تصمیمات و حل مدل ریاضی در بازه‌های زمانی کمتری انجام می‌شود، پس مصرف منابع در این حالت بیشتر می‌باشد. تصمیم‌گیری در مورد τ و بازه‌های زمانی حل مدل توسط کنترلر از این جهت حایز اهمیت است که مستقیماً روی مصرف منابع و افزایش یا کاهش پیچیدگی روش پیشنهادی تأثیرگذار است. پس باید با توجه به منابع و نرخ پذیرش تماس‌های مد نظر، یک τ مناسب را در نظر گرفت.

۴-۵ آزمایش چهارم: تأثیر خرابی نودها

علاوه بر ازدحام آنی، یکی دیگر از دلایلی که ممکن است سرورهای SIP را با اضافه‌بار روبه‌رو کند، خراب‌شدن ناگهانی اجزای شبکه و کاهش

۳-۵ آزمایش سوم: تأثیر زمان اجرا

در این بخش به تأثیر τ بر کارایی روش پیشنهادی می‌پردازیم. بدین منظور ما نرخ پذیرش تماس و منابع مصرفی را با مقادیر مختلف τ (۲، ۴، ۶ و ۸ ثانیه) به دست می‌آوریم. همان گونه که در شکل ۱۱-الف مشاهده می‌شود، زمانی که τ کم است (۲ ثانیه)، زمان انتظار برای تماس‌های جدید کم می‌باشد و فاصله زمانی تصمیم‌گیری کنترلر نیز کوتاه است. در نتیجه C^{ij} ، D^{ij} و \bar{D}^{ij} دقیق‌تر هستند که این باعث افزایش نرخ پذیرش تماس می‌شود. با این حال همان گونه که در شکل ۱۱-ب و ۱۱-ج نشان داده شده است، میانگین پردازنده و حافظه مصرفی در حالت τ برابر ۲ ثانیه، بیشتر از سایر موارد است.

در شکل‌های ۱۲ نیز به تأثیر τ بر کارایی سوئیچها در روش پیشنهادی می‌پردازیم. بدین منظور ما نرخ پذیرش تماس و منابع مصرفی سوئیچها را با مقادیر مختلف τ (۲، ۴، ۶ و ۸ ثانیه) به دست می‌آوریم. همان گونه که در شکل ۱۲-الف مشاهده می‌شود، زمانی که τ کم است (۲ ثانیه)، زمان انتظار در سوئیچها برای تماس‌های جدید کم می‌باشد که

جدول ۱: منابع مصرفی سوئیچ‌ها در توپولوژی N10. (%)

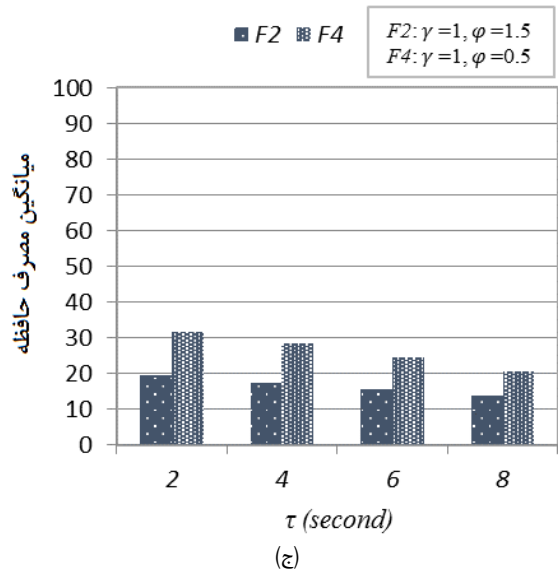
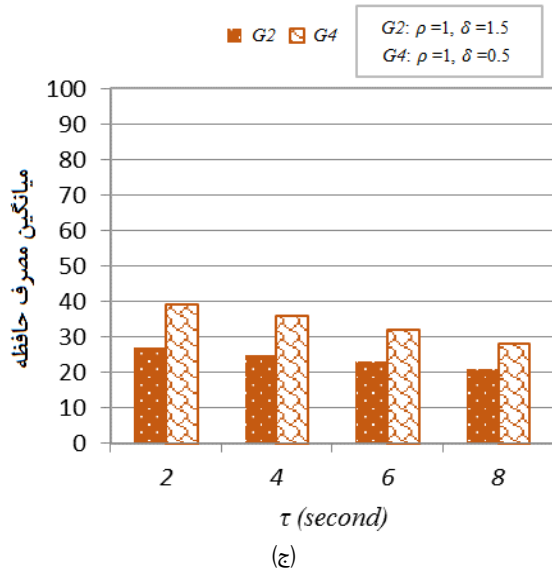
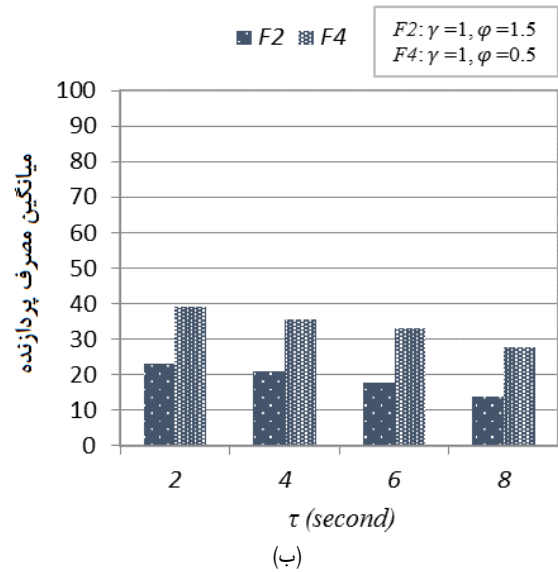
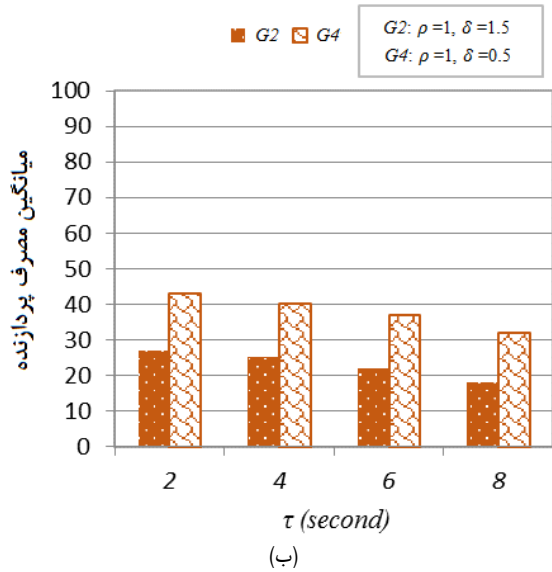
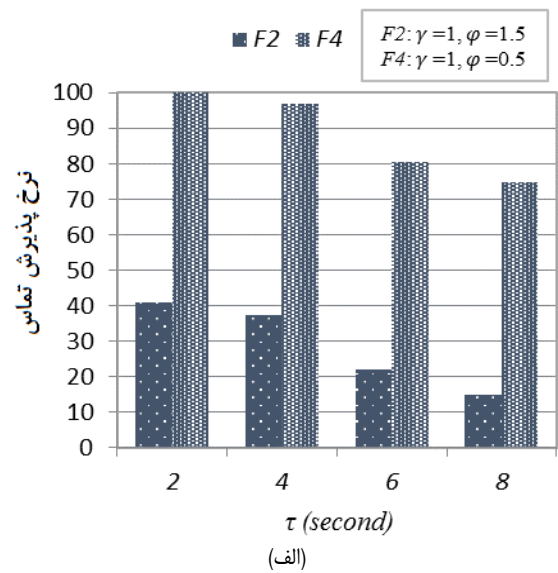
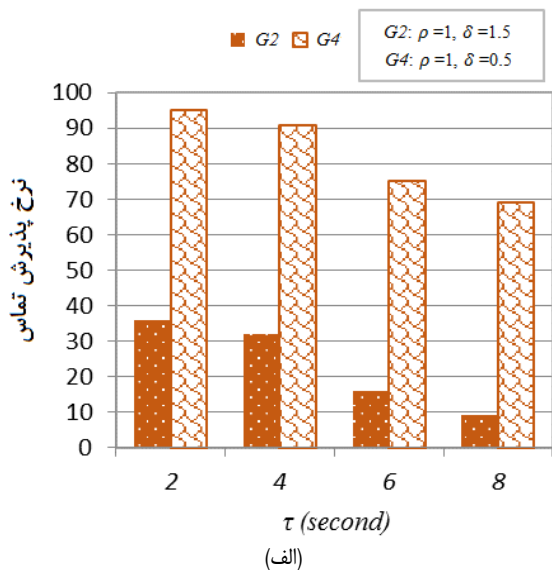
	سناریوی ۱				سناریوی ۲				سناریوی ۳			
	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4
۱ سوئیچ ۱ پردازنده	۲	۱۲	۲۲	۲۲	۵	۴۲	۶۰	۷۰	۸	۴۷	۷۰	۷۸
۲ سوئیچ ۲ پردازنده	۳	۱۱	۲۳	۲۳	۵	۴۱	۶۲	۶۹	۷	۴۶	۷۱	۷۹
۳ سوئیچ ۳ پردازنده	۲	۱۱	۲۲	۲۲	۶	۴۰	۶۲	۶۹	۸	۴۶	۷۰	۷۷
۴ سوئیچ ۴ پردازنده	۳	۱۳	۲۱	۲۲	۴	۴۳	۶۳	۶۸	۷	۴۵	۷۲	۷۶
۵ سوئیچ ۵ پردازنده	۲	۱۱	۲۲	۲۲	۵	۴۴	۶۰	۷۰	۷	۴۸	۷۰	۷۵
۶ سوئیچ ۶ پردازنده	۳	۱۰	۲۱	۲۳	۶	۴۲	۶۱	۶۹	۶	۴۶	۷۱	۷۷
۷ سوئیچ ۷ پردازنده	۲	۱۲	۲۰	۲۲	۴	۴۱	۶۲	۷۰	۸	۴۵	۷۳	۷۴
۱ سوئیچ ۱ حافظه	۶	۶۴	۱۲۸	۲۸	۱۵	۱۹۶	۳۴۴	۴۱۸	۴۸	۲۹۸	۴۱۲	۴۷۲
۲ سوئیچ ۲ حافظه	۷	۶۴	۱۲۸	۱۲۶	۱۴	۱۹۹	۳۴۳	۴۱۷	۴۷	۲۹۸	۴۱۰	۴۷۱
۳ سوئیچ ۳ حافظه	۵	۶۳	۱۲۷	۱۲۶	۱۶	۱۹۸	۳۴۲	۴۱۶	۴۶	۲۹۷	۴۱۱	۴۷۰
۴ سوئیچ ۴ حافظه	۶	۶۲	۱۲۵	۱۲۸	۱۴	۲۰۰	۳۴۴	۴۱۸	۴۸	۲۹۶	۴۱۳	۴۷۲
۵ سوئیچ ۵ حافظه	۷	۶۶	۱۲۸	۱۲۵	۱۵	۱۹۷	۳۴۲	۴۱۷	۴۷	۲۹۵	۴۱۱	۴۷۳
۶ سوئیچ ۶ حافظه	۷	۶۱	۱۲۵	۱۲۸	۱۶	۱۹۶	۳۴۴	۴۲۰	۴۶	۲۹۸	۴۱۰	۴۷۲
۷ سوئیچ ۷ حافظه	۵	۶۲	۱۲۶	۱۲۴	۱۵	۱۹۶	۳۴۳	۴۱۸	۴۸	۲۹۷	۴۱۱	۴۷۰

جدول ۲: منابع مصرفی سوئیچ‌ها در توپولوژی N18. (%)

	سناریوی ۱				سناریوی ۲				سناریوی ۳			
	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4
۱ سوئیچ ۱ پردازنده	۲	۱۴	۲۴	۲۵	۸	۴۶	۶۱	۷۲	۱۱	۵۱	۶۹	۷۹
۲ سوئیچ ۲ پردازنده	۲	۱۳	۲۳	۲۵	۷	۴۵	۶۰	۷۲	۱۰	۵۰	۶۹	۷۸
۳ سوئیچ ۳ پردازنده	۳	۱۲	۲۴	۲۴	۸	۴۵	۶۲	۷۱	۱۲	۵۲	۶۸	۷۷
۴ سوئیچ ۴ پردازنده	۴	۱۴	۲۲	۲۶	۶	۴۷	۶۱	۷۰	۱۱	۵۱	۶۷	۷۹
۵ سوئیچ ۵ پردازنده	۳	۱۳	۲۳	۲۷	۷	۴۴	۶۲	۷۰	۱۰	۵۱	۶۹	۷۷
۶ سوئیچ ۶ پردازنده	۲	۱۴	۲۴	۲۵	۸	۴۶	۶۰	۷۱	۱۲	۵۲	۶۹	۷۸
۷ سوئیچ ۷ پردازنده	۳	۱۲	۲۲	۲۴	۸	۴۶	۶۰	۷۲	۱۱	۵۳	۶۷	۷۷
۸ سوئیچ ۸ پردازنده	۴	۱۳	۲۴	۲۴	۷	۴۷	۶۲	۷۲	۱۲	۵۲	۶۶	۷۸
۹ سوئیچ ۹ پردازنده	۳	۱۴	۲۳	۲۵	۸	۴۷	۶۱	۷۲	۱۳	۵۱	۶۸	۷۹
۱۰ سوئیچ ۱۰ پردازنده	۲	۱۴	۲۲	۲۵	۷	۴۶	۶۲	۷۲	۱۲	۵۳	۶۷	۷۹
۱ سوئیچ ۱ حافظه	۱۶	۸۵	۱۵۰	۱۶۰	۲۵	۲۲۶	۳۵۴	۴۲۲	۵۹	۳۰۸	۴۱۲	۴۷۲
۲ سوئیچ ۲ حافظه	۱۵	۸۳	۱۴۹	۱۶۱	۲۴	۲۲۴	۳۵۳	۴۲۱	۵۸	۳۰۸	۴۱۲	۴۷۰
۳ سوئیچ ۳ حافظه	۱۴	۸۴	۱۵۲	۱۶۰	۲۳	۲۲۵	۳۵۴	۴۲۰	۵۷	۳۰۷	۴۱۱	۴۷۱
۴ سوئیچ ۴ حافظه	۱۶	۸۵	۱۵۱	۱۶۲	۲۵	۲۲۶	۳۵۵	۴۱۸	۵۹	۳۰۶	۴۱۰	۴۷۰
۵ سوئیچ ۵ حافظه	۱۵	۸۶	۱۵۲	۱۶۳	۲۴	۲۲۷	۳۵۴	۴۱۹	۵۸	۳۰۸	۴۱۲	۴۶۹
۶ سوئیچ ۶ حافظه	۱۴	۸۵	۱۵۳	۱۶۰	۲۳	۲۲۴	۳۵۵	۴۲۳	۵۷	۳۰۸	۴۱۱	۴۷۳
۷ سوئیچ ۷ حافظه	۱۶	۸۴	۱۵۲	۱۶۱	۲۵	۲۲۴	۳۵۶	۴۲۲	۵۹	۳۰۷	۴۱۲	۴۷۲
۸ سوئیچ ۸ حافظه	۱۴	۸۵	۱۵۱	۱۶۲	۲۴	۲۲۶	۳۵۴	۴۲۲	۵۸	۳۰۶	۴۱۰	۴۷۱
۹ سوئیچ ۹ حافظه	۱۵	۸۶	۱۵۰	۱۶۳	۲۳	۲۲۵	۳۵۵	۴۲۱	۵۷	۳۰۶	۴۱۰	۴۷۰
۱۰ سوئیچ ۱۰ حافظه	۱۶	۸۴	۱۴۹	۱۶۰	۲۴	۲۲۴	۳۵۴	۴۲۰	۵۸	۳۰۷	۴۱۱	۴۶۹

پروکسی‌ها در این ۲۰۰ ثانیه تقریباً ۶۱٪ و ۵۹٪ می‌باشد (شکل ۱۳-ب و ۱۳-ج). با خراب شدن یک سرور در ثانیه ۲۰۰ مسیری از مسیرهای شبکه حذف می‌گردد که تنها کنترلر از آن آگاه است. در این حالت، کنترلر به دلیل حالت F4 قصد دارد بدون دچار شدن به عواقب اضافه بار از بیشتر ظرفیت سرورهای موجود در مسیرهای جایگزین برای مقابله با افت شدید نرخ سرویس استفاده کند. به همین خاطر میانگین مصرف پردازنده و حافظه افزایش می‌یابد (شکل ۱۳-ب و ۱۳-ج منحنی با برچسب "With controller" از ثانیه ۲۰۰ تا ۴۰۰). در صورت عدم وجود کنترلر، تعداد درخواست‌های سرویس داده شده در ۲۰۰ ثانیه دوم تقریباً به ۶۳۶

ناگهانی ظرفیت می‌باشد. این خرابی ممکن است تحمیل بار سرور خراب شده روی سایرین در یک مجموعه سرور باشد. برای آزمایش چنین وضعیتی، ۱۰۰۰ تماس (بار ارائه شده) در یک بازه زمانی ۶۰۰ ثانیه‌ای به توپولوژی N10 تزریق شد. در ثانیه ۲۰۰، نودی از شبکه دچار خرابی شده و در ثانیه ۴۰۰ مجدداً به کار می‌افتد. همان گونه که در شکل ۱۳-الف مشاهده می‌شود، با کنترلر تا قبل از ثانیه ۲۰۰ به طور میانگین ۹۵۵ درخواست سرویس داده می‌شوند. میانگین مصرف پردازنده و حافظه برای



شکل ۱۲: نرخ پذیرش، پردازنده و حافظه مصرفی سوئیچ‌ها با تغییر τ ، (الف) نرخ پذیرش تماس توسط سوئیچ‌ها (D^H)، (ب) میزان مصرف پردازنده سوئیچ‌ها (p_s^H) و (ج) میزان مصرف حافظه سوئیچ‌ها (m_s^H).

شکل ۱۱: نرخ پذیرش، پردازنده و حافظه مصرفی پروکسی‌ها با تغییر τ ، (الف) نرخ پذیرش تماس توسط پروکسی‌ها (C^H)، (ب) میزان مصرف پردازنده پروکسی‌ها (p_p^H) و (ج) میزان مصرف حافظه پروکسی‌ها (m_p^H).

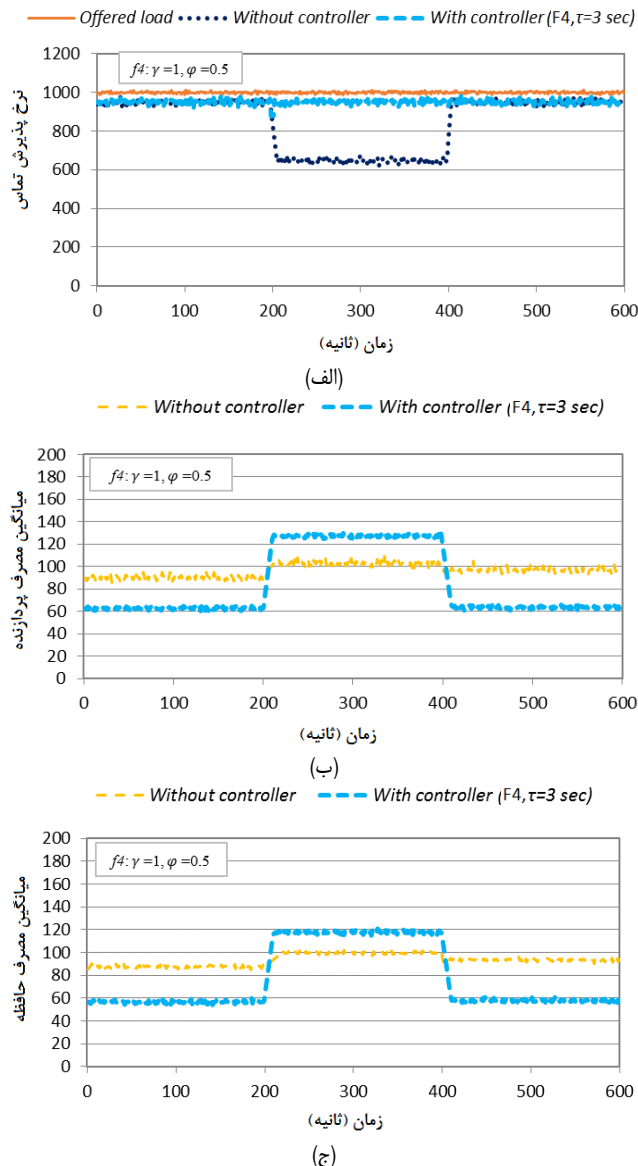
با بازگشت مجدد پروکسی سرور خراب‌شده به مسیر در ثانیه ۴۰۰ و عبور از وضعیت اضافه‌بار، نرخ پذیرش تماس در هر دو حالت برابر می‌شود اما به دلیل مدیریت هوشمندانه منابع توسط کنترلر، میانگین مصرف

افت می‌یابد و میانگین نرخ پذیرش تماس برابر ۶۳/۵٪ می‌شود (شکل ۱۳- الف منحنی با برچسب "Without controller" از ثانیه ۲۰۰ تا ۴۰۰).

گذردهی و مصرف منابع، در هر دو فاز سیگنالینگ و مدیا بود. برای کارهای آتی به تعمیم و توسعه مدل‌های ریاضی برای تخصیص بهینه منابع برای همه تجهیزات شبکه به صورت عمومی خواهیم پرداخت. همچنین چارچوب پیشنهادی را مجهز به فناوری مجازی‌سازی توابع شبکه خواهیم کرد. در این صورت منابع به صورت مجازی و در صورت تقاضا در دسترس خواهند بود. همچنین در نظر داریم متد پیشنهادی را در زیرسیستم چندرسانه‌ای^۱ (IMS) با استفاده از Open IMS پیاده‌سازی کنیم.

مراجع

- [1] J. Rosenberg, et al., *SIP: Session Initiation Protocol*, IETF RFC 3261, Jun. 2002.
- [2] J. Rosenberg, et al., Requirements for Management of Overload in the Session Initiation Protocol, RFC5390, Dec. 2008.
- [3] E. M. Nahum, J. Tracey, and C. P. Wright, "Evaluating SIP server performance," *ACM SIGMETRICS Performance Evaluation Review*, vol. 35, no. 1, pp. 349-350, Jun. 2007.
- [4] M. Ohta, "Performance comparisons of transport protocols for session initiation protocol signaling," in *Proc. 4th Int. Telecommunication Networking Workshop on QoS in Multiservice IP Networks*, pp. 148-153, Venezia, Italy, 13-15 Feb. 2008.
- [5] E. Noel and C. Johnson, "Initial simulation results that analyze SIP based VoIP networks under overload," in *Proc. 20th int. Teletraffic Conf. on Managing traffic performance in converged networks*, pp. 54-64, Ottawa, Canada, 17-21 Jun. 2007.
- [6] M. Ohta, "Simulation study of SIP signaling in an overload condition," in *Proc. Communication Internet, and Information Technology*, pp. 321-326, Cote d'Azur, France, 26-29 Aug. 2004.
- [7] V. K. Gurbani and R. Jain, "Transport protocol considerations for session initiation protocol networks," *Bell Labs Technical J.*, vol. 9, no. 1, pp. 83-97, 2004.
- [8] V. Pascual, *SIP Server Overload Problem Statement*, http://siprouter.org/wiki/media/tbd/overload_control_problem_statement_vpascualv10.pdf, 2009.
- [9] V. Hilt and I. Widjaja, "Controlling overload in networks of SIP servers," in *Proc. IEEE Int. Conf. on at the Network Protocols*, pp. 83-93, Orlando, FL, USA, 19-22 Oct. 2008.
- [10] P. McGregor, R. Kaczmarek, V. Mosley, D. Dease, and P. Adams, "National security/emergency preparedness and the next-generation network," *IEEE Communications Magazine*, vol. 44, no. 5, pp. 133-143, May 2006.
- [11] V. Hilt, Design Considerations for Session Initiation Protocol (SIP) Overload Control, RFC6357, Aug. 2011.
- [12] I. Widjaja, V. Hilt, and H. Schulzrinne, *Session Initiation Protocol (SIP) Overload Control*, IETF, <http://tools.ietf.org/html/draft-hilt-sipping-overload-02>, 2008.
- [13] C. Shen and H. Schulzrinne, "On TCP-based SIP server overload control," in *Proc. The Principles, Systems and Applications of IP Telecommunications*, pp. 71-83, Munich Germany, 2-4 Aug. 2010.
- [14] H. Schulzrinne, S. Narayanan, J. Lennox, and M. Doyle, *SIP Stone-Benchmarking SIP Server Performance*, Columbia University, 2002.
- [15] H. Lindholm, T. Vahakangas, and K. Raatikainen, "A control plane benchmark for telecommunications signalling applications," *Sort*, vol. 20, no. 5, pp. 100-125, May 2007.
- [16] M. Cortes, J. R. Ensor, and J. O. Esteban, "On SIP performance," *Bell Labs Technical J.*, vol. 9, no. 3, pp. 155-172, Nov. 2004.
- [17] S. Wanke, M. Scharf, S. Kiesel, and S. Wahl, "Measurement of the SIP parsing performance in the SIP express router," in *Proc. 13th open European Summer School and IFIP TC6.6 Conf. on Dependable and Adaptable Networks and Services*, pp. 103-110, Enschede, The Netherlands, 18-20 Jul. 2007.
- [18] S. S. Gokhale and J. Lu, "Signaling performance of SIP based VoIP: a measurement-based approach," in *Proc. IEEE Global Telecommunications Conf.*, pp. 761-765, St. Louis, MO, USA, 28-Nov.-3 Dec. 2005.
- [19] C. Chi, D. Wang, R. Hao, and W. Zhou, "Performance evaluation of SIP servers," in *Proc. 3rd Int. Conf. on the Communications and Networking in China*, pp. 674-679, Hangzhou, China, 25-27 Aug. 2008.



شکل ۱۳: کارایی در طول زمان و در صورت خرابی یک نود در ثانیه ۲۰۰، (الف) نرخ پذیرش تماس، (ب) میزان مصرف پردازنده و (ج) میانگین حافظه مصرفی.

پردازنده و حافظه کمتر است. یعنی کنترلر به نرخ پذیرش مناسب با مصرف منابع کمتر دست می‌یابد و این به دلیل آگاهی جامع از کل شبکه توسط کنترلر است. از ثانیه ۴۰۰ به بعد مجدداً تمام پروکسی سرورها می‌توانند به نرخ سرویس قبل از اضافه‌بار بازگردند.

۶- جمع‌بندی و کارهای آتی

اضافه‌بار پروکسی‌های SIP از مشکلات مهم شبکه‌های چندرسانه‌ای مبتنی بر اینترنت است که به دلیل مسیریابی نامناسب درخواست‌های تماس در لایه کاربرد، باعث اشباع منابع و در نتیجه افت شدید کارایی می‌شود. علاوه بر این، در فاز مدیا نیز عدم مسیریابی متمرکز باعث افت کارایی سوئیچ‌ها و گذردهی شبکه می‌شود. در این مقاله، یک روش مبتنی بر فناوری شبکه‌های نرم‌افزارمحور برای ارتقای شبکه‌های SIP ارائه شد تا بتوان بر این مشکلات فایز آمد. در این راستا، در لایه کنترلر یک کنترلر OpenFlow بر پایه مدل‌های برنامه‌ریزی خطی طراحی شد و قیود این مدل‌ها، مصرف منابع و مسیریابی هستند. برای ارزیابی کارایی روش پیشنهادی به شبیه‌سازی آن پرداخته شد و آزمایش‌های متنوع در قالب چند سناریو و در سه توپولوژی انجام گردید. نتایج نشان‌دهنده بهبود

- [29] G. Kambourakis, D. Geneiatakis, T. Dagiuklas, C. Lambrinouidakis, and S. Gritzalis, "Towards effective SIP load balancing," in *Proc. 3rd VoIP Security Workshop*, 6 pp., Berlin, Germany, 15-17 Jun. 2006.
- [30] S. Montagna and M. Pignolo, "Performance evaluation of load control techniques in sip signaling servers," in *Proc. 3rd Int. Conf. on the Systems*, pp. 51-56, Cancun, Mexico, 13-18 Apr. 2008.
- [20] D. Pesch, M. I. Pous, and G. Foster, "Performance evaluation of SIP-based multimedia services in UMTS," *Computer Networks*, vol. 49, no. 3, pp. 385-403, Oct. 2005.
- [21] M. Homayouni, et al., "Configuration of a sip signaling network: an experimental analysis," in *Proc. 5th Int. Joint Conf. on INC, IMS and IDC*, pp. 76-81, Seoul, South Korea, 25-27 Aug. 2009.
- [22] K. K. Ram, I. C. Fedeli, A. L. Cox, and S. Rixner, "Explaining the impact of network transport protocols on SIP proxy performance," in *Proc. IEEE Int. Symp. on the Performance Analysis of Systems and Software*, pp. 75-84, Austin, TX, USA, 20-22 Apr. 2008.
- [23] X. Wenfeng, et al., "A survey on software-defined networking," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 27-51, Firstquarter 2015.
- [24] A. R. Montazerolghaem, S. K. Shekofteh, M. H. Yaghmaee, and M. Naghibzadeh, "A load scheduler for SIP proxy servers: design, implementation and evaluation of a history weighted window approach," *International J. of Communication Systems*, vol. 30, no. 3, Article ID: e2980, May 2015.
- [25] J. Hongbo, et al., "Design, implementation, and performance of a load balancer for SIP server clusters," *IEEE/ACM Trans. on Networking*, vol. 20, no. 4, pp. 1190-1202, Aug. 2012.
- [26] K. Singh and H. Schulzrinne, "Failover, load sharing and server architecture in SIP telephony," *Computer Communications*, vol. 30, no. 5, pp. 927-942, Mar. 2007.
- [27] W. M. Wu, K. Wang, R. H. Jan, and C. Y. Huang, "A fast failure detection and failover scheme for SIP high availability networks," in *Proc. 13th Pacific Rim Int. Symp. on the Dependable Computing*, pp. 187-190, Melbourne, Australia, 17-19 Dec. 2007.
- [28] Y. J. Cheng, K. Wang, R. H. Jan, C. Chen, and C. Y. Huang, "Efficient failover and load balancing for dependable SIP proxy servers," in *Proc. IEEE Symp. on the Computers and Communications*, pp. 1530-1546, Marrakech, 6-9 Jul. 2008.

احمدرضا منتظرالقائم تحصیلات خود را در مقاطع کارشناسی مهندسی فناوری اطلاعات و کارشناسی ارشد مهندسی کامپیوتر به ترتیب در سال‌های ۱۳۸۹ و ۱۳۹۲ به پایان رسانده است. او همچنین دکتری سیستم‌های نرم افزاری خود را در سال ۱۳۹۶ با درجه عالی از دانشگاه فردوسی مشهد اخذ نموده است و هم اکنون استادیار دانشکده مهندسی کامپیوتر دانشگاه اصفهان می‌باشد. ایشان در سال ۱۳۹۳ موفق به دریافت جایزه تحصیلی بنیاد ملی نخبگان شده و پژوهشگر برتر سال ۱۳۹۵ دانشکده مهندسی دانشگاه فردوسی مشهد بوده است. نام‌برده قبل از پیوستن به دانشگاه اصفهان، در سال‌های ۱۳۸۹ الی ۱۳۹۷ مدیر کیفی آزمایشگاه تایید نمونه تجهیزات شبکه ویپ دانشگاه فردوسی مشهد بوده است. از دیگر سوابق وی می‌توان به شرکت در پروژه توسعه زیرساخت ابر مجازی و هوشمند SAVI دانشگاه تورنتو کانادا و تحقیق به عنوان پژوهشگر پسادکتری بر روی مدیریت انرژی شبکه‌های ارتباطی نرم‌افزار محور در دانشگاه فردوسی مشهد اشاره کرد. همچنین ایشان در طرح‌های پژوهشی و پروژه‌های متعددی از جمله پیاده سازی ویپ شرکت ایران خودرو خراسان، طراحی و پیاده سازی معماری دیتاستر شرکت توزیع برق مشهد و همچنین بررسی، مطالعه، راه اندازی و استقرار سیستم ویپ استانداری خراسان رضوی مشارکت داشته‌اند. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: شبکه‌های کامپیوتری، شبکه‌های نرم‌افزار محور، مجازی سازی توابع شبکه، شبکه‌های چندرسانه‌ای، شبکه‌های سلولی و اینترنت اشیا.