نام : دانیال بنکداری

پروژه درس : قابلیت اطمینان و تحلیل ریسک

موضوع : قابلیت اطمینان در مراکز داده

استاد : دکتر عباس رجبی

فهرست

مقدمه

در دهه‌های اخیر با کاهش قیمت تجهیزات کامپیوتری و امکان دسترسی آسان و فراگیر به اینترنت پرسرعت و همچنین وابستگی روز افزون کسب و کارها به فناوری اطلاعات، تقاضا برای خدمات دیتاسنتر بسیار افزایش یافته است. در حال حاضر، دیتاسنترها نقش کلیدی و بسیاری مهمی را در حوزه اقتصاد دیجیتال ایفا می‌کنند و پشتیبانی طیف عظیمی از فعالیت‌های مهم دولت‌ها، جامعه و کسب و کارها را بر عهده دارند. بنابراین پابدار بودن دیتاسنتر ها در طول زمان ضروری است این به معنی است که همیشه در دسترس باشند به این منظور دیتاسنترها باید این قابلیت را داشته باشند که در صورت خرابی سریعا به حالت اولیه بازگشته و نیاز شبکه را برطرف کنند در این گزارش ما به موضوع افزایش قابلیت اطمینان دیتاسنترها میپردازیم

دیتاسنتر

دیتاسنتر (Data Center) یا مرکز داده به مکانی گفته می‌شود که گروه عظیمی از سرورهای کامپیوتری و تجهیزات شبکه با استفاده از امکانات زیرساختی و ارتباطی برای رایانش و میزبانی مجموعه بزرگی از داده‌ها، گرد هم آمده باشند. به بیان دیگر، دیتاسنتر محل استقرار تعداد زیادی از سرورهای کامپیوتری است که در کنار یکدیگر و بدون وقفه، امور مربوط به رایانش، ذخیره‌سازی و انتقال داده‌ها را انجام می‌دهند. ساختمان اغلب این مراکز، دارای سیستم‌های امنیتی پیشرفته، سیستم تهویه، اطفاء حریق و سیستم توزیع برق است که به سامانه برق اضطراری (UPS و دیزل ژنراتور) مجهز شده‌اند. پیاده‌سازی یک دیتاسنتر، عموما بر پایه شبکه عظیمی از منابع پردازشی و ذخیره‌سازی صورت می‌پذیرد که با کمک یک زیرساخت ارتباطی قدرتمند، امکان ارایه سرویس‌های اینترانتی (محلی و داخل سازمانی) یا اینترنتی انتقال داده را در مقیاس‌های کوچک و بزرگ فراهم می‌کند.

اهمیت دیتاسنترها برای سازمان‌ها و کسب و کارها

در دهه‌های اخیر با کاهش قیمت تجهیزات کامپیوتری و امکان دسترسی آسان و فراگیر به اینترنت پرسرعت و همچنین وابستگی روز افزون کسب و کارها به فناوری اطلاعات، تقاضا برای خدمات دیتاسنتر بسیار افزایش یافته است. در حال حاضر، دیتاسنترها نقش کلیدی و بسیاری مهمی را در حوزه اقتصاد دیجیتال ایفا می‌کنند و پشتیبانی طیف عظیمی از فعالیت‌های مهم دولت‌ها، جامعه و کسب و کارها را بر عهده دارند.

حجم فعالیت بعضی از سرویس‌های اینترنتی تا حدی وسیع است که در برخی مواقع حتی تا ۱۰ سرور کامپیوتری نیز جوابگوی نیازهای آنها نیستند. در چنین شرایطی، تنها شبکه عظیمی از صدها و حتی هزاران سرور کامپیوتری قادر خواهند بود تا از عهده تامین نیازهای این سرویس‌ها برای ذخیره‌سازی، رایانش و انتقال داده‌ها بربیایند. از دیدگاه سازمانی در فناوری اطلاعات، دیتاسنترها بستر لازم را برای پیاده‌سازی و اجرای بدون وقفه برنامه‌های کاربردی سازمان‌ها و فعالیت آنها فراهم می‌کنند.

مدیریت ارتباط با مشتری (CRM)، برنامه‌ریزی منابع سازمانی (ERP)، پست الکترونیک، ذخیره‌سازی، تهیه نسخه پشتیبان و اشتراک فایل‌ها، کلان داده، هوش مصنوعی و یادگیری ماشینی، خدمات ارتباطی درون سازمانی، انتقال، پردازش و رایانش بدون وقفه داده‌های عظیم، از جمله رایج‌ترین کاربردهای دیتاسنتر در سازمانها دانشگاه‌ها و شرکت‌های گوناگون می‌باشند. شرکت‌های بزرگی مانند گوگل، مایکروسافت و آمازون، با تاسیس مراکز داده در نقاط مختلف جهان همواره در تلاشند تا زیرساخت‌های لازم را برای پاسخگویی به نیازهای متفاوت جامعه بین‌المللی در مناطق مختلف دنیا فراهم کنند.

پیشینه

بلایای طبیعی غیرمنتظره مانند طوفان و گردبادها ممکن است منجر به قطع برق در سیستم توزیع شوند و منجر به تلفات قابل توجهی برای مشتریان شوند طبق آمار ایالات متحده بر اساس خدمات تامین برق، قطعی برق ناشی از بلایای طبیعی از سال ۲۰۰۲ سالانه بین ۱۸ تا ۳۳ میلیارد دلار هزینه داشته است با توجه به آمار میانگین بار قدرت در IDC (اینترنت دیتا سنتر) های گوگل در سال ۲۰۱۱ تقریباً ۲۶۰ مگاوات است که بیشتر از تقاضای برق شهر سالت لیک است با توجه به وابستگی شدید به برق، قطعی غیر منتظره برق دلیل اصلی قطعی IDC است

اگر چه آمادگی قبل از فاجعه سیستم های توزیع قدرت در سال های اخیر به سرعت توسعه یافته اند قطعی برق همچنان در بسیاری از موارد اجتناب ناپذیر است در چنین شرایطی، یک طرح بازیابی کارآمد می تواند به سرعت بارهای بحرانی را در سیستم توزیع برق بازیابی کند و خسارات ناشی از بلایای طبیعی را تا حد زیادی کاهش دهد

افزایش کارایی و قابلیت اطمینان مرکز داده یکی از چالش های فنی حفظ کیفیت خدمات برای کاربران نهایی در عملیات مرکز داده است مصرف انرژی مدل‌های اجزای مرکز داده برای حصول اطمینان از طراحی بهینه امکانات داخلی و محدود کردن مصرف انرژی مرکز داده بسیار مهم هستند مدلسازی قابلیت اطمینان مرکز داده همچنین مهم است زیرا رضایت کاربر نهایی به در دسترس بودن خدمات مرکز داده بستگی دارد.

نحوه عملکرد دیتاسنتر

مراکز داده اینترنتی را می توان به دو دسته تقسیم کرد: IDC های مالک وIDC های هم مکان. در مقایسه باIDCهایی که توسط مالک اداره می شوند و مالک آنها کنترل کامل بر روی آنها دارد IDC های هم لوکیشن

اساساً کانتینر هستند که سرورهای تحت مالکیت (یا اجاره ای) هستند که توسط چندین مستأجر اداره می شود. بنابراین، اپراتورIDC های colocation هیچ دسترسی به بار کاری در سرورها ندارند و تغییر عملکرد دشوار است

وظیفه اصلی IDC استفاده از منابع سرور برای تکمیل بارهای کاری ارسال شده از سرورهای فرانت اند است حجم کار معمولاً از دو بخش تشکیل شده است: تعداد مورد نیاز واحدهای منبع سرور و کیفیت نیاز خدمات (QoS)

ظرفیت منابع سرور درIDC ها توسط منبع تغذیهIDC ها تعیین می شود . از این رو، مدل سازی از تقاضای برق درIDC ها برای تجزیه و تحلیل ضروری است. قدرت تقاضا IDC ها از دو بخش تشکیل شده است: تقاضای تسهیلات فناوری اطلاعات و تقاضای سیستم های خنک کننده مصرف برق متناسب با تعداد واحد منبع سرور فرض شده است

برای جلوگیری از فروپاشی کل IDC به دلیل وقفه ناگهانی در عرضه برق منبع تغذیه آماده به کار مانند ژنراتورها پشتیبان معمولاً درIDC ها مجهز می شوند. اندازه مورد نیاز به منابع قدرت آماده به کار در مرحله برنامه ریزی تعیین می شود و باید نیاز برق طراحی شده را پوشش دهد با این حال، تقاضای برق در IDCها ثابت نیست ارتقاء تجهیزات آی تی یک چرخه سه ساله دارند ، وIDC ها ممکن است نصب دستگاه های قدرتمندتر در طول عمر خود داشته باشند IDC ها تمایل دارند واحدهای انرژی تجدیدپذیر سازگار با محیط زیست را با ژنراتورهای دیزلی سنتی جایگزین شوند که خروجی غیر قابل کنترل است و ممکن است در ارائه انرژی کافی شکست بخورد بنابراین، منابع موجود قدرت آماده به کار درIDC ها ممکن است پس از قطع منبع تغذیه از طرح عملیاتی اصلی خود پشتیبانی نکنند و در نتیجه مهاجرت و لغو بار کاری ضروری است.

بازیابی پس از رویداد سیستم های توزیع برق

آمار نشان می دهد که بیش از ۹۰ درصد از قطعی های سیستم برق در سطح توزیع رخ می دهد. سیستم های توزیع برق معمولاً به صورت مش طراحی می شوند اما به شکل شعاعی عمل می کنند که به این معنی است که هر گونه خطای جزء در شبکه ممکن است منجر به جداسازی مناطق آسیب دیده شود برای برخی از مناطق با خطوط متصل به شبکه اصلی، اپراتور سیستم می تواند شبکه را با تغییر وضعیت سوئیچ ها و پیکربندی مجدد بازیابی کند در مناطق آسیب دیده بدون اتصال فیزیکی به شبکه اصلی، خاموشی تا پایان ادامه خواهد داشت تا زمانی که قطعه خراب تعمیر می شود. بنابراین، پیکربندی مجدد شبکه ها و تعمیر قطعات خراب دو مورد اصلی ابزار برای بازیابی توزیع برق هستند یک روش رایج برای اندازه‌گیری سطح بحرانی بار، محاسبه وزن اولویت بار برای همه بارها در شبکه‌های برق است.

ظاهراً، با اعمال تابع هدف، نیازی به تعریف این نیست که کدام بار مهم است که در فرآیند ترمیم اولویت اصلی را دارد: مدل بهینه‌سازی طرح بازیابی را پیدا می‌کند که تلفات کلی را به حداقل می‌رساند. با این حال، پس از وقوع بلایای طبیعی، ابتدا باید برخی از زیرساخت‌های حیاتی برای ارائه خدمات اضطراری مانند بیمارستان‌ها و ایستگاه‌های آتش نشانی بازیابی شوند.

از آنجایی که اهمیت IDC های حیاتی به دلیل نقش ویژه ای است که در اقتصاد محلی ایفا می کنند، سودمندی IDC ها باید به روش های اقتصادی مانند دلار آمریکا محاسبه شود. بنابراین، وزن اولویت بار نیز باید با واحدهای مشابه نشان داده شود.

همانطور که در بالا بحث شد، هدف از بازیابی سیستم توزیع برق با IDCهای حیاتی، به حداقل رساندن تلفات ابزار IDCها و سایر بارها در طول فرآیند بازیابی است. از دست دادن ابزار IDC ها توسط تصمیمات عملیات IDC تعیین می شود که تحت تأثیر طرح های بازیابی سیستم توزیع برق قرار می گیرد.می توان یک نتیجه مهم گرفت که

ارزش منبع تغذیه IDCها در IDCها و بازه های زمانی مختلف متفاوت است، زیرا یک بار کاری محاسباتی فقط می تواند قبل از پایان ضرب الاجل در IDC که در آن قرار دارد تکمیل شود. تغییر مقدار منبع تغذیه در بین IDCها ممکن است واضح نباشد، اما تفاوت مقدار منبع تغذیه بین دو شکاف زمانی مختلف می تواند بسیار زیاد باشد زیرا منبع تغذیه بازیابی شده پس از مهلت بار کاری محاسباتی برای حجم کاری بی ارزش است.

با توسعه خدمات و برنامه های کاربردی مبتنی بر ابر، ارائه دهندگان خدمات ابری تجاری مانند گوگل، فیس بوک یا آمازون اکنون مراکز داده عظیم توزیع شده جغرافیایی را مستقر می کنند. طبق تحقیقات انجام شده توسط شرکت بین المللی داده (IDC)، انتظار می رود تقاضای جهانی برای انتقال داده و خدمات دیجیتال دو برابر شود و به 4.2 زتابایت در سال، معادل ۴۲۰۰۰ اگزابایت تا سال ۲۰۲۲ برسد تعداد مراکز داده در سطح جهان در حال افزایش است تا این ترافیک داده به سرعت در حال رشد را مدیریت کند، در حالی که تقاضای انرژی مراکز داده نیز در حال افزایش است. مراکز داده ایالات متحده حدود ۳۰۰ میلیون ترابایت داده را مدیریت کردند که در سال ۲۰۱۶ حدود ۸۳ میلیارد کیلووات ساعت در سال مصرف می کرد، بنابراین ۲۷۷ کیلووات ساعت در هر ترابایت با ردپای کربن تقریباً ۳۵ کیلوگرم CO_2 در هر ترابایت داده مصرف می کرد.

در گزارشی اشاره شده است که تعداد سرورها در مراکز داده طی سال‌های ۲۰۱۰ تا ۲۰۱۸ به دلیل افزایش تقاضا برای بارهای کاری محاسباتی ۳۰ درصد افزایش یافته است. با افزایش تعداد سرورها، تعداد نمونه‌های محاسباتی از جمله ماشین‌های مجازی که بر روی سخت‌افزار فیزیکی کار می‌کنند تا ۵۵۰ درصد افزایش یافت، ترافیک داده‌ها ۱۱ برابر شد و ظرفیت ذخیره‌سازی نصب‌شده در طول مدت مشابه ۲۶ برابر افزایش یافت بنابراین، تقاضای انرژی جهانی مراکز داده از ۱۹۴ تراوات ساعت به ۲۰۵ تراوات ساعت طی سال‌های ۲۰۱۰ تا ۲۰۱۸ افزایش یافت علاوه بر این، مراکز داده به دلیل تقاضای رو به رشد انرژی که تا سال ۲۰۳۰ تا ۷۲۰ میلیون تن پیش بینی شده است، به طور غیر مستقیم بر انتشار CO_2 تأثیر می گذارد

در حال حاضر، شرکت های پیشرو در تجارت فناوری اطلاعات و ارتباطات (ICT) اکنون مراکز داده جدید خود را در مناطق با عرض جغرافیایی بالا در منطقه قطب شمال می سازند تا از مزایای طبیعی از جمله امکانات تولید انرژی های تجدید پذیر، هوای سرد و رطوبت مناسب بهره مند شوند. گوگل در سال ۲۰۱۱ یک مرکز داده در Hamnia فنلاند ساخته است تا از آب سرد دریا از خلیج فنلاند و انرژی باد در خشکی استفاده کند. در حالی که فیس بوک در سال ۲۰۱۳ به سوئد و در سال ۲۰۱۶ به ایرلند نقل مکان کرده است به دلیل مزایای طبیعی در عملیات مرکز داده وجود دارد

این شرکت ها از مزایای طبیعی برای کاهش مصرف انرژی مراکز داده استفاده می کنند و از این رو به طور غیرمستقیم مشارکت خود را در انتشار CO2 کاهش می دهند. دو مرحله اصلی از نوآوری مرکز داده برای مقابله با چالش های بهره وری انرژی وجود دارد. در مرحله اول، اپراتورهای مرکز داده بر بهبود کارایی تجهیزات فناوری اطلاعات (IT) و تأسیسات خنک‌کننده مرکز داده طی سال‌های ۲۰۰۷ تا ۲۰۱۴ تأکید کرده‌اند

در مرحله دوم، اپراتورهای بزرگ مرکز داده بر روی تهیه انرژی های تجدیدپذیر (به عنوان مثال، باد، خورشید) برای تامین برق برای عملیات مرکز داده به جای منابع برق سنتی تمرکز کرده اند مدل های مصرف انرژی و قابلیت اطمینان مرکز داده برای ارائه راه حل هایی برای این دو چالش عملیاتی در مراکز داده مورد نیاز است. مدل‌های مصرف انرژی می‌توانند به پیش‌بینی پیامدهای تصمیمات عملیاتی کمک کنند، که منجر به مدیریت و کنترل مؤثرتر بر سیستم می‌شود علاوه بر این، مدل‌سازی قابلیت اطمینان بخش‌های بار منفرد مرکز داده و ارزیابی قابلیت اطمینان مرکز داده در کل برای جلوگیری از وقفه‌های ناخواسته در خدمات و اطمینان از SLA (توافقنامه در سطح خدمات) متعهد مهم است

در برخی موارد، مدل ارزیابی قابلیت اطمینان نیز مدل‌های مصرف انرژی دستگاه‌ها را در مقاطع بار می‌طلبد در این راستا، یک مدل مناسب مصرف انرژی یا رویکرد مدل‌سازی صرفاً به معنای دقت مدل نیست، در حالی که

رویکرد مدل‌سازی مصرف انرژی مرکز داده اغلب به کاربردهای مدل‌های انرژی یا توان مصرفی قطعات بستگی دارد.

قابلیت اطمینان

مراکز داده باید از نظر محیطی کنترل شده و مجهز به دستگاه های تهویه کننده قدرت باشند تا از عملکرد قابل اعتماد بارهای IT از جمله سرورها و دستگاه های شبکه اطمینان حاصل شود. اپراتورهای مرکز داده تمام اقدامات ممکن را برای جلوگیری از آسیب عمدی یا تصادفی به تجهیزات موجود در مرکز داده انجام می دهند، به طوری که بخش های بارگذاری می توانند درجه بالایی از قابلیت اطمینان در عملکرد را تضمین کنند. طبق تعریف، قابلیت اطمینان احتمالی است که یک دستگاه یا سیستم عملکرد خود را به اندازه کافی تحت شرایط عملیاتی خاص برای یک دوره زمانی مورد نظر انجام دهد

در اینجا درجه اعتماد بر موفقیت بر اساس تجربه گذشته قرار می گیرد، که به عنوان احتمال موفقیت برای یک سیستم ماموریت گرا مانند یک مرکز داده در این مورد کمیت می شود. این تعریف قابلیت اطمینان تنها وضعیت عملیاتی جزء یا سیستم را بدون هیچ وقفه ای در نظر می گیرد در همین حال، احتمال یافتن جزء یا سیستم در حالت عملیاتی به عنوان "در دسترس بودن" شناخته می شود که به عنوان یک شاخص قابلیت اطمینان برای یک سیستم قابل تعمیر استفاده می شود در این حالت، اجزای موجود در بخش‌های بار مرکز داده قابل تعمیر هستند که شامل فرآیند جایگزینی نیز می‌شود، بنابراین شاخص در دسترس بودن به طور گسترده در مدل‌سازی قابلیت اطمینان مرکز داده استفاده می‌شود

صنعت مرکز داده به «طبقه‌بندی‌های لایه‌ای» که توسط مؤسسه Uptime به عنوان مقیاس گرادیان بر اساس پیکربندی‌ها و الزامات مرکز داده، از کمترین (سطح ۱) تا معتبرترین (سطح ۴) معرفی شده است، تکیه کرده است

مؤسسه Uptime این چهار سطح از مراکز داده را تعریف می کند که خطر تأثیر سرویس (به عنوان مثال، در دسترس نبودن و خرابی) را به دلیل فعالیت های مدیریت خدمات و خرابی های برنامه ریزی نشده مشخص می کند

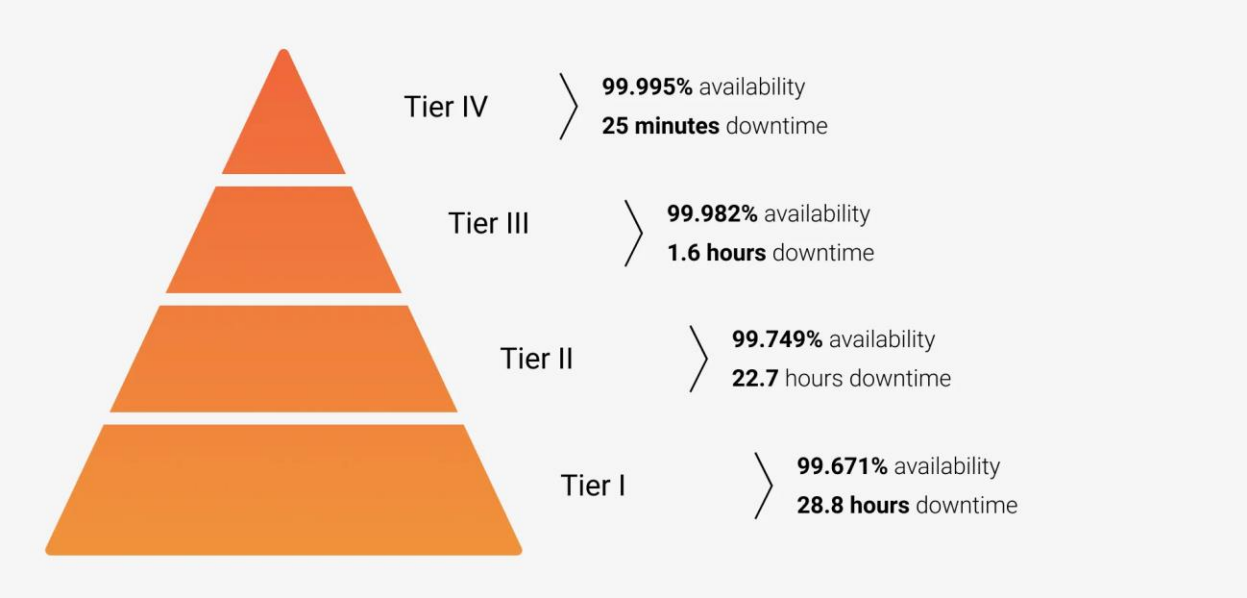
طبقه بندی ردیف مراکز داده

هدف اصلی طبقه‌بندی لایه‌ای مراکز داده، ایجاد دستورالعملی از توپولوژی طراحی است که سطوح مطلوبی از در دسترس بودن را مطابق با پرونده کسب وکار مالک، که توسط مؤسسه Uptime معرفی شده است، ارائه می‌کند سطح مرکز داده با در دسترس بودن IPCS از جمله ابزار و منبع تولید کننده پشتیبان تعیین می شود مؤسسه Uptime در تحقیقات استانداردسازی طراحی مرکز داده و تشریح افزونگی سیستم های منبع تغذیه زیربنایی خود پیشگام است.

طبق سیستم طبقه‌بندی مؤسسه Uptime، زیرساخت داخلی مراکز داده طی حداقل چهار مرحله متمایز در ۴۰ سال گذشته تکامل یافته است که برای مدل‌سازی قابلیت اطمینان استفاده می‌شود و به عنوان " Tiers of Data Center" شناخته می‌شود تا آوریل ۲۰۱۳، مؤسسه Uptime ۲۳۶ گواهینامه برای ساخت مراکز داده در سراسر جهان بر اساس طبقه بندی لایه اعطا کرده بود

این ترکیبی از رویکرد طبقه‌بندی کمی و کیفی است ترکیب این دو رویکرد توسط مؤسسه Uptime برای صدور گواهینامه ردیف استفاده می‌شود، با این حال، رویکرد ارزیابی قابلیت اطمینان به موارد تجاری مالک مرکز داده بستگی دارد. سیستم طبقه‌بندی ردیف، مراکز داده را بر اساس توانایی آنها برای اجازه تعمیر و نگهداری و مقاومت در برابر خرابی در سیستم منبع تغذیه ارزیابی می‌کند. Tier I (کمترین قابل اعتماد) تا Tier IV (مطمئن ترین) بسته به اجزای اضافی در مسیر منبع تغذیه موازی به بخش های بار بحرانی تعریف می شوند. افزونگی در مسیر

منبع تغذیه نه تنها می تواند در دسترس بودن مرکز داده را بهبود بخشد. در دسترس بودن ممکن است به دلیل خرابی های حالت رایج کاهش یابد، که به داده های آماری برای تحقیقات بیشتر نیاز دارد



شکل ۱ : طبقه بندی دیتاسنترها

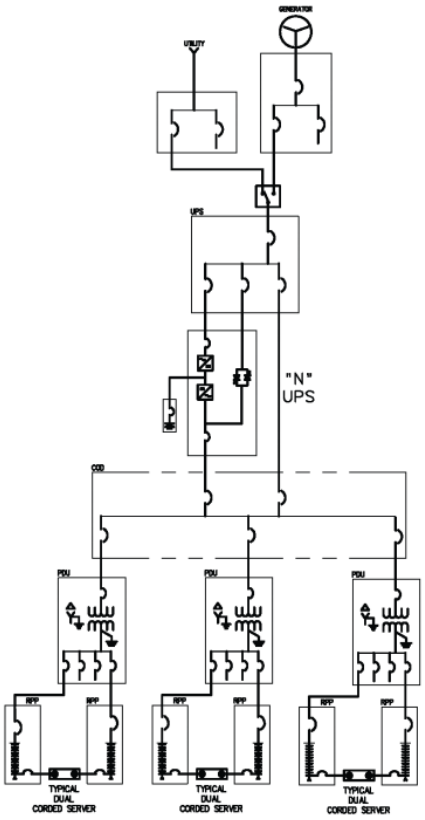
عوامل موثر مدلسازی قابلیت اطمینان در مراکز داده

مهمترین عامل برای ارزیابی قابلیت اطمینان مرکز داده، خرابی اجزای سیستم است. آرنو و همکاران مثالی را تدوین کرده است "اگر UPS در سیستم منبع تغذیه از کار بیفتد و تمام بارهای متصل به مرکز داده انرژی خود را از دست بدهند، بدیهی است که یک "شکست" خواهد بود. ؟ آیا این یک "شکست" برای مرکز داده است؟"

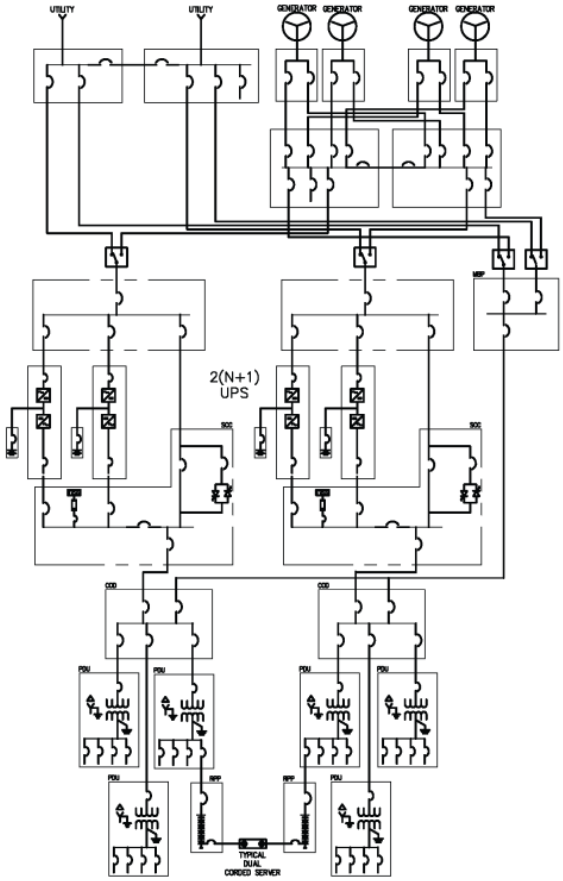
با توجه به تعریف خرابی ارائه شده در فصل ۸ کتاب طلایی IEEE، استاندارد ۴۹۳-۲۰۰۷ «شکست از دست دادن توان یک واحد توزیع برق (یا پانل توزیع UPS در مورد مرکز داده) است. بنابراین از دست دادن کل یک یو پی اس بر ماموریت کلی فناوری متصل تأثیر می گذارد که طبق تعریف، خرابی مرکز داده است. با این حال، اگر یک قطع کننده مدار قطع شود و رک های متصل برق را از دست بدهند، به عنوان خرابی مرکز داده در نظر گرفته

نمی شود، بلکه سرورهای موجود در رک ها (Rack) از کار افتاده یا برای کار در دسترس نیستند. بنابراین، اولین گام هر تحلیل قابلیت اطمینان، تعریف «وضعیت شکست» سیستم مورد مطالعه است. توضیح مشابهی در مورد "خطا و شکست" برای یک سیستم ابری ارائه شده است، جایی که اصطلاح "شکست" برای خطاهای کشنده در سیستم استفاده می شود که غیرقابل جبران هستند و به طور فاجعه باری بر عملکرد سیستم تأثیر می گذارند.

با این حال، "خطا" عملکرد سیستم را کاهش می دهد (یعنی تاخیر، کاهش پرتاب) زیرا خطاها می توانند به طور خودکار حل شوند و سیستم می تواند به حالت اولیه بازگردد علاوه بر این، تعریف خرابی ذکر شده در کتاب طلایی IEEE، استاندارد ۴۹۳-۲۰۰۷ با طبقه‌بندی ردیف مرکز داده در تضاد است، که نشان می‌دهد خرابی با حالت عملکرد ضعیف برای تعریف مراکز داده ضروری است. خرابی PSU های (Power distribution unit) سطح رک برای ارزیابی کفایت منابع محاسباتی در نظر گرفته شده است



شکل ۲ : سیستم برق N با کمترین قابلیت اطمینان



شکل ۳ : سیستم برق 2(N+1) با بیشترین قابلیت اطمینان

شاخص های قابلیت اطمینان و معیارهای مورد استفاده برای مدلسازی قابلیت اطمینان در مراکز داده

۱) شاخص های قابلیت اطمینان برای بارها و خدمات آن

شاخص‌هایی که برای بخش بار فناوری اطلاعات استفاده می‌شوند را می‌توان به دو گروه (۱) شاخص‌های مربوط به عملکرد و خدمات فناوری اطلاعات و (۲) شاخص‌های مربوط به آمادگی بخش بار فناوری اطلاعات در مرکز

داده طبقه‌بندی کرد. QoS یک شاخص کلیدی برای ارزیابی عملکرد مرکز داده است، که همچنین شامل شاخص های کیفیت کلیدی(KQI) و شاخص های کلیدی عملکرد (KPI) برای خدمات فناوری اطلاعات ارائه شده توسط مرکز داده است

این شاخص ها برای نظارت بر خدمات فناوری اطلاعات و مدیریت ظرفیت محاسباتی در مرکز داده استفاده می شود. شاخص‌های «قابلیت اطمینان خدمات» و «در دسترس بودن خدمات» برای حفظ SLA با مشتری یا کاربر مرکز داده استفاده می‌شوند. به عبارت دیگر، در دسترس بودن یا قابل اطمینان بودن سرویس، آمادگی یک سیستم مرکز داده برای ارائه خدمات فناوری اطلاعات وعده داده شده به کاربر را مشخص می کند. معمولاً به آمادگی یک سیستم "بالا" بودن گفته می شود. شاخص «در دسترس بودن خدمات» برای پرداختن به قابلیت اطمینان بارهای فناوری اطلاعات یا سرورها در سطح رک استفاده می شود. محققان همچنین «احتمال قطع» سرور را به عنوان یک شاخص قابلیت اطمینان نشان داده‌اند که با افزایش تلفات برق در IPCS در مرکز داده تغییر می‌کند.

$$A_{Service} = \frac{t_{up}}{t_{up} + t_{down}}$$

که در آن، AService در دسترس بودن خدمات است. t up و t down به ترتیب زمان های uptime و downtime سیستم هستند. جدا از احتمال قطع، "قابلیت اطمینان خدمات" نیز مورد تاکید قرار گرفته است زیرا احتمال انجام درخواست های خدمات بدون تاخیر با این شاخص مشخص می شود. خدمات مهم جلسه محور در مراکز داده، هم احتمال شروع موفقیت آمیز یک جلسه با سرویس به نام «دسترسی» و هم احتمال اینکه یک جلسه سرویس را با QoS وعده داده شده تا پایان جلسه به نام «قابلیت نگهداری» ارائه دهد، اندازه گیری می

کند. از این نظر، نقص در هر میلیون عملیات (DPM) شاخصی است که عملیات شکست خورده را در هر میلیون عملیات اندازه گیری می کند تا قابلیت اطمینان سیستم را ارزیابی کند

$$R_{DPM} = \frac{O_f}{O_a} \times 1,000,000$$

$$r_{Service} = 100\% - \frac{R_{DPM}}{1,000,000}$$

که در آن، RDPM، Of، و Oa به ترتیب نقص در هر میلیون عملیات، تعداد عملیات ناموفق، و عملیات تلاش شده هستند. قابلیت اطمینان سرویس توسط rService نشان داده می شود. شاخص دیگری به نام «تأخیر سرویس» با اهمیت برای ارزیابی قابلیت اطمینان سیستم به ویژه برای مراکز داده لبه و اینترنت در ذکر شده است. تأخیر تراکنش مستقیماً بر کیفیت تجربه کاربران نهایی تأثیر می گذارد. افزایش ۵۰۰ میلی ثانیه ای در تأخیر سرویس باعث کاهش ۲۰ درصدی ترافیک برای Google.com می شود و افزایش ۱۰۰ میلی ثانیه ای در تأخیر سرویس باعث کاهش ۱ درصدی فروش برای Amazon.com می شود. شاخص در دسترس بودن سرویس با در نظر گرفتن میانگین سطح بار CPU، در نتیجه بارهای کاری محاسباتی، و عملکرد خطر CPU، با یک شاخص جدید "در دسترس بودن ماشین وابسته به بار" است. شاخص‌های قابلیت اطمینان وابسته به بار مشابه با نام‌های «عملکرد متوسط» و «متوسط در دسترس بودن تحویل‌شده» تعریف شده‌اند.

اصول اولیه QoS و شاخص های قابلیت اطمینان خدمات مشابه هستند. بیشتر بر اساس شاخص های در دسترس بودن خدمات و عملکرد سیستم فناوری اطلاعات است. شاخص‌های عملکرد سیستم فناوری اطلاعات به روش‌های مختلفی مدل‌سازی می‌شوند به غیر از شاخص‌های قابلیت اطمینان مبتنی بر QoS، شاخص‌های دیگری مانند میانگین زمان بین شکست (MTBF)، میانگین زمان تعمیر (MTTR)، در دسترس بودن، قابلیت اطمینان وجود

دارد که در مدل‌سازی قابلیت اطمینان برای اجزای فیزیکی بخش بار فناوری اطلاعات استفاده می‌شود یک شاخص قابلیت اطمینان مشابه به نام "احتمال از دست دادن حجم کار (LOWP)" بر اساس احتمال قطع شدن سرور در سطح رک تعریف شده است. خطر قطع شدن سرور به دلیل خطاهای الکتریکی و افت ولتاژ ناشی از آن نیز تعریف شده است. علاوه بر این، شاخص های SLA-aware مبتنی بر عملکرد بار فناوری اطلاعات نیز برای راه حل های مبتنی بر نرم افزار در مراکز داده استفاده می شود. شاخص‌های آگاه از SLA به عنوان مثال، کاهش عملکرد به دلیل مهاجرت (PDM)، نقض تجمیع سطح سرویس (SLAV) برای ارزیابی عملکرد بارهای فناوری اطلاعات با بارهای کاری تلفیقی در سیستم ابری اعمال می‌شوند

۲) شاخص های قابلیت اطمینان برای بخش خنک کننده

شاخص در دسترس بودن عملیاتی به عنوان احتمال اینکه سیستم در حالت عملیاتی مورد نظر قرار گیرد تعریف شده است و به صورت ریاضی تابعی از میزان خرابی سیستم بیان می شودλsys و نرخ تعمیر μsys است. یکی دیگر از شاخص های قابلیت اطمینان به نام در دسترس بودن عملکردی Af نیز بر اساس دمای اتاق سرور پیش بینی شده و شرایط کاری سرورها در استفاده می شود. در دسترس بودن عملکرد کلی سیستم خنک کننده مراکز داده عمدتاً توسط در دسترس بودن عملیاتی، چگالی گرما، ویژگی های انتقال حرارت دمای اتاق، زمان راه اندازی سیستم خنک کننده و زمان تعمیر خرابی سیستم خنک کننده تعیین می شود. تحلیل مبتنی بر شرایط عملکردی مشابهی برای سیستم تهویه مطبوع مرکز داده بر اساس ظرفیت منبع تغذیه تهویه مطبوع انجام شده است

$$A_O(\infty) = \frac{1}{\frac{\lambda_{sys}}{\mu_{sys}} + 1}$$

$$A_f(\infty) = A_o(\infty) \times (1 - p_{us}) + (1 - A_o(\infty)) \times p_t$$

که در آن AO و Af در دسترس بودن عملیاتی و عملکردی سیستم خنک کننده هستند. میزان خرابی و میزان تعمیر سیستم خنک کننده به ترتیب λsys و μsys می باشد. p us احتمال دمای اتاق خارج از محدوده مورد نظر زمانی است که سیستم در حال کار است. pt احتمال مقدار مورد نظر دمای اتاق در هنگام از کار افتادن سیستم خنک کننده است محققان بر قابلیت اطمینان سیستم خنک‌کننده تأکید کردند، زیرا قابلیت اطمینان به تحمل خطا و قابلیت اطمینان مرتبط است. شاخص های اهمیت قابلیت اطمینان (Ii) و اهمیت قابلیت اطمینان-هزینه (Ci) استفاده می شود.

$$I_i = R_s(U_i, \mathbf{p}^i) - R_s(D_i, \mathbf{p}^i)$$
$$M_i = I_i \times (1 - \frac{C_i}{C_{sys}})$$

که در آن، Ii اهمیت قابلیت اطمینان جزء i است. pi بردار قابلیت اطمینان مؤلفه را با حذف مؤلفه i نشان می دهد. Di و Ui به ترتیب نشان دهنده خرابی و وضعیت بالا مؤلفه i هستند. Ci هزینه اکتساب جزء i و Csys هزینه اکتساب سیستم است. جدای از شاخص های ذکر شده، شاخص های معمولی مانند در دسترس بودن بر اساس MTBF، MTTR، میزان خرابی و تعمیر به طور گسترده برای مدل سازی قابلیت اطمینان بخش بار خنک کننده مرکز داده استفاده می شود ذکر این نکته ضروری است که تحقیقات در مورد قابلیت اطمینان سیستم خنک کننده مرکز داده هنوز به اندازه کافی مورد توجه قرار نگرفته است، در حالی که زیرساخت های خنک کننده برای ساختمان های تجاری در دهه گذشته به شدت مورد توجه محققان قرار گرفته است. تحقیق در مورد زیرساخت های خنک کننده مطمئن مرکز داده بسیار مورد نیاز است زیرا حساسیت دمایی سالن سرور مرکز داده باید با سایر امکانات ساختمان مقایسه شود

رویکردهای تحلیلی برای ارزیابی قابلیت اطمینان

کاربردهای رویکردهای تحلیلی مانند نمودارهای بلوکی قابلیت اطمینان (RBD) و تجزیه و تحلیل درخت خطا در مدل‌سازی قابلیت اطمینان مرکز داده به دلیل سادگی و نیاز کمتر به ظرفیت محاسباتی بسیار رایج هستند. یکی از اولین چنین تحقیقاتی در سال ۱۹۸۸ منتشر شد ، که در آن نویسندگان در دسترس نبودن سیستم منبع تغذیه توزیع شده یک اتاق کنترل مخابراتی را با سیستم توزیع متمرکز نیرو تجزیه و تحلیل و مقایسه کردند. یک رویکرد تحلیلی مشابه وجود دارد که در آن قابلیت اطمینان سیستم توزیع معمولی جریان متناوب (AC) با سیستم توزیع برق جریان مستقیم (DC) در مراکز داده با استفاده از RBD مقایسه می‌شود. خرابی سیستم توزیع برق تنها بدون در نظر گرفتن خرابی بارهای فناوری اطلاعات در نظر گرفته شده است، در حالی که در دسترس بودن IPCS (سیستم تهویه داخلی) با در نظر گرفتن احتمال شکست بارهای فناوری اطلاعات از جمله PSU است. بسته به سطح ولتاژ در IPCS، قابلیت اطمینان ساختارهای مختلف IPCS با استفاده از RBD ارزیابی می‌شود. مدل‌سازی قابلیت اطمینان زیرساخت‌های منابع محاسباتی (بخش بار فناوری اطلاعات) مراکز داده با استفاده از مدل RBD انجام شده است. نوع مشابهی از تجزیه و تحلیل در مقالات دیگر ارائه شده است، که در آن محققان از نمودارهای جهت دار و غیر جهت دار با استفاده از حداقل مجموعه های برش استفاده کرده اند. رویکرد تحلیلی همچنین برای ارزیابی قابلیت اطمینان توپولوژی های شبکه مرکز داده با استفاده از مفهوم تئوری مجموعه برش و بهینه سازی تخصیص منابع برای شبکه های قابل اعتماد اعمال می شود. رویکرد تحلیلی یعنی RBD، شبکه پتری تصادفی و مدل جریان انرژی برای ارزیابی قابلیت اطمینان IPCS استفاده می شود. یک مدل RBD توسعه یافته پیشنهاد شده است که می تواند وابستگی قابلیت اطمینان اجزای IPCS به قابلیت اطمینان کلی IPCS را در نظر بگیرد. مدل پیشنهادی پویا RBD نامیده می‌شود، که بیشتر با مدل شبکه پتری رنگی مقایسه می‌شود تا تحلیل خصوصیات رفتاری را انجام دهد که صحت مدل پیشنهادی را برای قابلیت اطمینان IPCS تأیید می‌کند، همانطور که توضیح داده شد. تکنیک تجزیه و تحلیل درخت خطا برای تخمین میزان شکست، MTBF، MTTR و قابلیت اطمینان

توپولوژی های مختلف UPS استفاده می شود.RBD همچنین برای تجزیه و تحلیل قابلیت اطمینان سیستم خنک کننده مرکز داده استفاده می شود. در دسترس بودن یک سیستم خنک کننده با آب با استفاده از حداکثر زمان توقف مجاز در مدل پیشنهادی RBD ارزیابی می شود، در حالی که مدل RBD و شبکه پتری تصادفی برای کمی سازی اثرات پایداری، هزینه ها و قابلیت اطمینان زیرساخت خنک کننده مرکز داده استفاده می شود.

زیرساخت‌های مراکز داده نسبت به ساختمان‌ها و صنایع معمولی حیاتی‌تر هستند. داده های آماری خرابی مؤلفه مرکز داده برای تحقیقات بیشتر برای بهبود قابلیت اطمینان مناسب در کاربرد مرکز داده مورد نیاز است. یک مجموعه داده در دسترس عموم وجود دارد که زمان خرابی و تعمیر سرورها را منتشر می‌کند در حالی که داده‌های خرابی و تعمیر سایر مؤلفه‌ها (به عنوان مثال، PDU، PSU، دستگاه‌های خنک‌کننده) بخشی از هیچ مجموعه‌ای در دسترس عموم نیستند. داده ها. تمایل اپراتور مرکز داده به حفظ محرمانه بودن و محرمانه بودن اطلاعات داخلی مراکز داده، دلایل اصلی فقدان چنین مجموعه های داده ای است

رویکردهای مبتنی بر شبیه سازی در ارزیابی قابلیت اطمینان

در کنار مدل‌های تحلیلی، رویکردهای مدل‌سازی احتمالی نیز برای ارزیابی قابلیت اطمینان مرکز داده رایج هستند. مدل‌های فضای حالت شامل مدل مارکوف و زنجیره مارکوف مونت کارلو (MCMC) برای مدل‌سازی قابلیت اطمینان سیستم‌های مقیاس بزرگ و قابل تعمیر استفاده می‌شوند، بنابراین استفاده از مدل‌های مارکوف اخیراً برای مدل‌سازی قابلیت اطمینان مراکز داده رایج شده است برای اجتناب از مدل فضای حالت غیر خطی متغیر زمانی در مدل مارکوف، میزان خرابی و تعمیر اجزای سیستم های مورد مطالعه ثابت فرض شده است. اگر اثر پیری با در نظر گرفتن نرخ شکست ثابت نادیده گرفته شود، نرخ خرابی و تعمیر می‌تواند برای یک جزء ثابت باشد بنابراین، مدل‌های پایایی مبتنی بر شبیه‌سازی برای ارزیابی قابلیت اطمینان مراکز داده امروزه به طور گسترده در تحقیقات مورد استفاده قرار می‌گیرند.

مونت کارلو یکی از پرکاربردترین رویکردهای مبتنی بر شبیه‌سازی برای مدل‌سازی قابلیت اطمینان مرکز داده است. رویکرد شبیه‌سازی مونت کارلو بیشتر برای تولید خرابی وابسته به زمان و رویدادهای تعمیر اجزای سیستم با استفاده از تابع توزیع احتمال و مشاهده عملکرد کلی سیستم بر اساس داده‌های تصادفی استفاده می‌شود روش شبیه‌سازی مونت کارلو نیز برای مدل‌سازی قابلیت اطمینان اجزایی که در مراکز داده استفاده می‌شوند، یعنی UPS سیستم شبکه نوری استفاده می‌شود. در رویکردهای مبتنی بر شبیه‌سازی، مدل خرابی جزء سیستم مهم است زیرا نتیجه شبیه‌سازی شده قابلیت اطمینان کلی می‌تواند بسته به حالت خرابی متفاوت باشد، به‌ویژه برای کاربردهای با قابلیت اطمینان بالا مانند مرکز داده. به عنوان مثال، در دسترس بودن مرکز داده Tier IV باید دارای ۵ تا ۶ عدد ۹ باشد، که به این معنی است که رویدادهای شکست بسیار کمی در میلیون ها رویداد تصادفی مشاهده می شود. بنابراین، دقت در بررسی حالت خرابی و مدل‌سازی خرابی اجزا در رویکردهای مبتنی بر شبیه‌سازی برای مدل‌سازی قابلیت اطمینان مراکز داده مهم است.

جدا از تعداد نمونه و حالت خرابی اجزا، توابع توزیع احتمال خرابی و رویدادهای تعمیر قطعات در سیستم مورد مطالعه نیز نقش مهمی در رویکردهای مبتنی بر شبیه‌سازی در صورت مدل‌سازی قابلیت اطمینان دارند. توابع توزیع احتمال و کاربردهای توابع توزیع برای مدل‌سازی قابلیت اطمینان سرورها در مرکز داده تجزیه و تحلیل می‌شوند. تابع توزیع خرابی و زمان تعمیر دستگاه های شبکه و سایر اجزای سرور یعنی هارد دیسک، حافظه و کارت های شبکه بیشتر برای مدل سازی قابلیت اطمینان سیستم کلی استفاده می شود. علاوه بر مونت کارلو، شبکه‌های پتری تصادفی و زنجیره مارکوف مونت کارلو (MCMC) نیز برای مدل‌سازی قابلیت اطمینان مرکز داده استفاده می‌شوند.

وابستگی بخش ها و زیرسیستم های بارگذاری مرکز داده

قابلیت اطمینان یک سیستم به عنوان توانایی سیستم برای ارائه خدماتی که به طور موجه قابل اعتماد است تعریف می شود. از طرف دیگر، ارائه معیار برای تصمیم گیری در مورد قابل اعتماد بودن سرویس، قابلیت اطمینان یک سیستم است به عنوان مثال، وابستگی سیستم A به سیستم B نشان دهنده میزانی است که قابلیت اطمینان سیستم A تحت تاثیر سیستم B است (یا خواهد بود) ،قابلیت نگهداری و غیره، زیرا در دسترس بودن خدمات مرکز داده به تداوم خدمات ارائه شده توسط اجزای سیستم های فرعی بستگی دارد

وابستگی به بخش بار خنک کننده

نشان داده شده است که عمر باتری در IPCS به دلیل افزایش دمای عملیاتی به میزان ۱۰ درجه سانتیگراد ۵۰٪ کاهش می یابد. در حالی که عناصر غیرفعال در سرورها مانند خازن ها زمان عمر را تا ۵۰ درصد برای افزایش ۱۰ درجه سانتی گراد در دما کاهش می دهند همچنین به این نتیجه منجر شده که افزایش دمای سالن داده بازده انرژی را بهبود می بخشد اما بر قابلیت اطمینان سرورها وPSU ها در IPCS تأثیر می گذارد. مصرف برق بارهای خنک کننده به آرایش سرورها در سالن داده بستگی دارد، از این رو آرایش متراکم سرور باعث مصرف انرژی بالا توسط بارهای خنک کننده می شود. علاوه بر این، تأخیر شبکه و ذخیره سازی به دلیل بارهای خنک کننده بیش از حد و سرورهای استفاده نشده یا بیکار بیشتر افزایش می یابد، که همچنین بر قابلیت اطمینان کلی استراتژی های قرار دادن مرکز داده تأثیر می گذارد

نتیجه گیری و توصیه ها

به عنوان ستون فقرات تحولات فناوری اطلاعات و ارتباطات امروزی (ICT)، بهره وری انرژی و قابلیت اطمینان بیشتر مراکز داده برای اطمینان از عملکرد مرکز داده مورد نیاز است. جنبه‌های مدل‌سازی مصرف انرژی و مدل‌سازی قابلیت اطمینان مراکز داده بررسی شد. با تجزیه و تحلیل مدل‌های مصرف انرژی بخش‌های بار مرکز داده به پر کردن شکاف‌های تحقیقاتی مربوط به مدل‌سازی مصرف انرژی مرکز داده کمک می‌کند، که کاربرد مدل‌ها را در تحقیقات بیشتر آسان‌تر می‌کند مشخص شده است که مدل‌های مصرف انرژی اجزای مرکز داده اغلب برای مدل‌های قابلیت اطمینان مرکز داده ضروری هستند، اگرچه مدل‌های مصرف انرژی نیز کاربردهای دیگری برای مدیریت انرژی مرکز داده دارند.

بر اساس بررسی مدل‌های مصرف انرژی مؤلفه‌های مرکز داده، تأکید بیشتری بر در دسترس بودن پارامترها و متغیرهای مدل مصرف انرژی نسبت به دقت برای کاربرد در تحقیق وجود دارد. دقت بالاتر چنین مدل هایی اغلب کاربرد را پیچیده می کند و نمی تواند کمک زیادی به بهبود روش پیشنهادی کند.

علاوه بر این، فقدان تحقیق در مورد مدل سازی مصرف انرژی تجهیزات سیستم تهویه داخلی توان (IPCS) در این بررسی شناسایی شده است. کل مصرف برق IPCS می تواند تا ۱۰ درصد از کل تقاضای مرکز داده را افزایش دهد، که همچنین می تواند باعث قطع و مشکلات اطمینان در مراکز داده شود. این بررسی همچنین به نشان دادن رابطه بین مصرف برق و قابلیت اطمینان مرکز داده کمک می کند و به این نتیجه می رسد که تحقیقات بیشتری برای کاهش مصرف برق به ویژه در بخش IPCS به عنوان یک توصیه انجام شود.

جنبه‌های مدل‌سازی قابلیت اطمینان مرکز داده بررسی شد که نیاز به کد استاندارد برای عملکرد مرکز داده به همراه طبقه‌بندی ردیف موجود را نشان می‌دهد که به عنوان توصیه ذکر شده است. این تجزیه و تحلیل همچنین به نشان دادن پیشرفته‌ترین روش‌های مدل‌سازی قابلیت اطمینان مبتنی بر شبیه‌سازی و تحلیلی کمک می‌کند

که می‌تواند به محققان آینده در انتخاب مدل‌های مناسب بر اساس کاربرد کمک کند. تجزیه و تحلیل نیاز به خرابی آماری و داده های تعمیر اجزای مرکز داده را نشان می دهد که به ندرت به دلیل عدم تمایل اپراتور برای اشتراک گذاری در دسترس است. بنابراین توصیه می‌شود اطلاعات آماری خرابی و تعمیر قطعه منتشر شود تا بتوان از آن برای تحقیقات بیشتر استفاده کرد. همچنین توصیه برای تمرکز بیشتر بر بهبود تجزیه و تحلیل قابلیت اطمینان بخش خنک کننده و تجزیه و تحلیل وابستگی قابلیت اطمینان کلی مرکز داده به سایر بخش های بار با جزئیات بیشتر است.

مراجع

[1] Ahmed, K., Bollen, M. and Alvarez, M., 2021. A Review of Data Centers Energy Consumption and Reliability Modeling. *IEEE Access*, 9, pp.152536-152563.

[2] Arno, B., Friedl, A., Gross, P. and Schuerger, R., 2010. Reliability of example data center designs selected by tier classification. *2010 IEEE Industrial and Commercial Power Systems Technical Conference - Conference Record*,.

[3] Liu, Y., Lei, S. and Hou, Y., 2019. Restoration of Power Distribution Systems With Multiple Data Centers as Critical Loads. *IEEE Transactions on Smart Grid*, 10(5), pp.5294-5307.

[4] Xu, Y., Liu, C., Wang, Z., Mo, K., Schneider, K., Tuffner, F. and Ton, D., 2019. DGs for Service Restoration to Critical Loads in a Secondary Network. *IEEE Transactions on Smart Grid*, 10(1), pp.435-447.