

پیاده سازی یک تکنیک حفاظت داده ای نوین مبتنی بر الگوهای ذخیره سازی RAID

محمد رضا میبیدی	تیمور ایزدی	جواد اکبری ترکستانی
دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیر کبیر meybodi@ce.aut.ac.ir	دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی اراک t_izadi@iau-arak.ac.ir	دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی اراک j_akbari@iau-arak.ac.ir

چکیده

اینگونه از سیستم ها علیرغم هزینه های فضایی نسبتا بالایی که بمنظور تهیه ماندهای داده ای پشتیبان از روی عناصر داده ای اصلی دنباله و یا هزینه های زمانی افزوده ای که در طی عملیات احیاء، ذخیره و بازیابی داده ها به سیستم تحمیل میگردد، با این همه، مدلهای قادر به شناسایی دقیق موقعیت و ترمیم خطا های بروز کرده در سطح بلاکهای داده ای و یا احیاء کامل خطا های دیسکی دنباله توزیع شده نبوده و نرخ تحمل پذیری خطا و سطح حفاظت و امنیت داده ای سیستم برای نگهداری داده های بحرانی پایین می باشد [3][2]. علاوه بر این، ساختار ذخیره سازی در این الگو بگونه ایست که عملیتهای دسترسی داده ای غالبا بصورت متمرکز و تحت نظارت کنترلر مرکزی GRC صورت می گیرد و این گونه از پیاده سازی برای سیستم های توزیع شده چندان مناسب نمی باشد. چرا که در سیستم های ذخیره سازی غیر متمرکز جهت توزیع المانهای پردازشی، داده ها، توابع و کنترل بر روی ماشین های مختلف سیستم توزیع شده، حتی الامکان میبایستی هر سیستم در کنترل و اجراء تمامی عملیات داده ای داخلی خود از سایر سیستم ها مستقل باشد (خود مختاری اجراء) [1].

در الگوی پیشنهادی، بمنظور بهبود نرخ تحمل پذیری و افزایش قابلیت اطمینان سیستم در مقابل بروز گونه های مختلف خطا (خصوصا خطاهای دیسکی) از یک تکنیک توزیع ترکیبی نوین برای گسترش بلاکهای داده ای در میان دیسکهای دنباله توزیع شده استفاده می گردد، که این ساختار ذخیره سازی با تلفیق الگوی ردیفی کردن داده ای در سطح بلاکی و نوع خاصی از الگوی انعکاس داده ای که در آن هر دیسک به دو نیمه تقسیم بندی شده، و نسخه های پشتیبان از بلاکهای داده ای اصلی هر دیسک، در بلاکهای متوالی از سایر دیسکهای دنباله (بگونه ای که در تصویر ۱ آمده است) توزیع میگردد، قادر است تا پهنای باند مؤثر سیستم بهنگام عملیات احیاء داده ای، سطح دسترسی پذیری داده ای و کارایی و بازدهی سیستم را بمقدار قابل توجهی افزایش دهد [17]. در ضمن بکار گیری این ساختار توزیع داده ای در سطح دنباله موجب میگردد تا با افزایش نرخ احتمالی قابلیت شناسایی و احیاء خطاهای داده ای در سطح بلاکی و دیسکی

در این مقاله سعی داریم الگوی داده ای نوینی را بمنظور نگهداری داده های بحرانی در سیستم های با نیاز امنیتی بالا بگونه ای ارائه دهیم، که سیستم قادر باشد در مقابل بروز گونه های مختلف خطاهای داده ای، آنها را با دقت بالایی شناسایی نموده و در کمترین زمان ممکن داده های از دست رفته^۱ را احیاء نماید. از این روی، مدل پیشنهادی را بکمک ترکیب دو الگوی ردیفی کردن داده ای و نگارشی خاص از الگوی انعکاس داده ای که از تکنیک منحصرا بفردی برای توزیع بلاکهای داده ای پشتیبان در سطح دنباله دیسکی توزیع شده (RADD)^۲ بهره میبرد، پیاده سازی نموده ایم. الگوی ذخیره داده ای پیشنهادی با توجه به توزیع نسخه های پشتیبان بلاکهای داده ای هر دیسک، در سایر دیسکهای دنباله قادر است، تا حداکثر پهنای باند مؤثر را بهنگام دستیابی داده ها در شرایط امن و یا بهنگام بروز خطا در سطح دنباله و یا احیاء خطا های داده ای فراهم آورد. نکته قابل توجه در مدل پیشنهادی آنست، که ساختار ذخیره سازی آرایه شده نرخ احتمالی فقدان داده ای را بگونه ای کاهش می دهد که فقدان یک دیسک مشروط به فقدان تمامی بلاکهای داده ای از سایر دیسکهای دنباله بوده و این با نرخ احتمالی بسیار ناچیزی امکان پذیر است. بر این اساس الگوی آرایه شده قادر است تا سطح حفاظت داده ای، قابلیت اطمینان، بازگرد پذیری، کارایی و بازدهی سیستم را برای نگهداری داده های بحرانی بمقدار قابل توجهی افزایش خواهد داد.

واژه های کلیدی: تحمل پذیری خطا، فقدان داده ای، افزونگی توزیع شده، ردیفی کردن داده ای، RADD

۱- مقدمه:

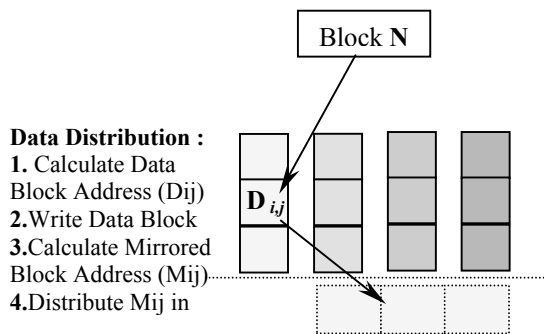
در سیستم های ذخیره داده ای توزیع شده ای که از ساختارهای تکنیک RAID^۳ استفاده مینمایند، غالبا الگوی انعکاس داده ای چندان مورد توجه قرار نمی گیرد، چرا که در

¹ Missed Data

² Redundancy Array of Distributed Disk

³ Redundancy Array of Independent Disk

⁴ Execution Autonomy



تصویر ۲. الگوی توزیع داده ای مدل پیشنهادی

مستقیم به داده های راه دور (درخواست عملیات نوشتن ، بروز سازی و احیاء داده ای) برای کاربر محلی وجود نداشته و درخواستهای دسترسی راه دور^۸ بایستی بروش RPC^۹ و بکمک GRC صورت گیرد^{[6][5]} . که این امر از یک سوی سبب افزایش نرخ ترافیک انتقال پیام (میان Client هاو GRC Server) در سطح شبکه گردیده واز سوی دیگر موجب افزایش ازدحام درخواستها بر روی GRC و بالطبع باعث کاهش بازدهی ، کند شدن و افزایش تداخلات گردیده که با القاء تصویر یک سیستم واحد^{۱۱} (SSI) در سیستم توزیع شده منافات دارد. علاوه براین وجود یک Centralized GRC Server خود یک نقطه شکست^{۱۱} را در سیستم بوجود میآورد که مفاهیمی همچون توزیع شدگی^{۱۲} ، شفافیت^{۱۳} ، مقیاس پذیری^{۱۴} را زیر سؤال میبرد^[9].

بمنظور پیاده سازی این ساختار ذخیره سازی و الگوی داده ای توزیع شده بایستی ویژگیهای خاصی را برای سیستم در نظر گرفت . یکی از موضوعات اساسی در پیاده سازی این الگو بر روی دنباله دیسکی توزیع شده تناسب میان تعداد دیسکهای درون دنباله به تعداد بلاکهای داده ای درون دیسک میباشد. از آنجا که در این ساختار، نیمی از بلاکهای هر دیسک به ذخیره بلاکهای داده ای اصلی تخصیص داده شده و از سوی دیگر بلاکهای داده ای اصلی در هر دیسک از دنباله، در بلاکهای پشتیبان سایر دیسکها توزیع میگردد، بنابر این بایستی رابطه زیر (فرمول ۱) میان تعداد دیسکها و بلاکهای داده ای درون دنباله صدق نماید تا بتوان این الگوی توزیع داده ای را بر روی چنین دنباله دیسکی پیاده سازی کرد^{[8][2][1]}.

$$N_d = 2 \cdot (N_b - 1) \quad \text{فرمول ۱}$$

همانطور که در تصویر ۲ نیز مشاهده می شود ، بمنظور افزایش سطح دسترسی پذیری داده ای (یا پهنای باند مؤثر دسترسی بهنگام انجام عملیات عادی و یا عملیات احیاء) ، کاهش

Disk0	Disk1	Disk2	Disk3
A ₀	A ₁	A ₂	A ₃
B ₀	B ₁	B ₂	B ₃
C ₀	C ₁	C ₂	C ₃
A ₁	C ₀	B ₀	A ₀
B ₂	A ₂	C ₁	B ₁
C ₃	B ₃	A ₃	C ₂

تصویر ۱. ساختار ذخیره سازی مدل پیشنهادی

مدت زمان متوسط فقدان داده ای^۵ MTTL در سیستم افزایش یافته ، و از سوی دیگر با افزایش سطح دستیابی سیستم بهنگام عملیات غیر معمول ، نرخ متوسط احیاء داده ای^۶ MTTRD نیز بهبود خواهد یافت .

نکته قابل توجه دیگر در ساختار توزیع الگوی ارائه شده ، پیاده سازی یک معماری توزیع شده از آن بر روی سیستم های ذخیره داده ای غیرمتمرکزی باشد ، که با گسترش بلاکهای پشتیبان یک دیسک بر روی سایر دیسکهای دنباله، هر دیسک قادر خواهد بود تا بهنگام صدور یک درخواست عملیات ورودی/خروجی منفرد از یک کاربر راه دور و یا عملیات احیاء داده ای بر روی دنباله، بطور مستقل زیر عملیات داده ای را که توسط GRC به آن واگذار گردیده انجام دهد و بصورت محلی به داده های کاربر راه دور دسترسی داشته و بدین ترتیب سطح اجراء همروند عملیاتی داده ای محلی و راه دور بمقدار قابل توجهی افزایش خواهد داشت. در ادامه این مقاله به بحثهای زیر پرداخته ایم . در بخش ۲ ، مدل پیشنهادی خود را در قالب تصاویر، فرمولها و نمودارها تشریح کرده و در بخش ۳ ابتدا مدل را به لحاظ هزینه عملیات احیاء داده ای بررسی نموده و سپس به تحلیل ارزیابی کارایی این مدل در مقابل سایر مدلها مشابه قبلی پرداخته و در بخش بعد نتایج حاصل از شبیه سازی مدل پیشنهادی را در قالب نمودارهایی ارائه و با نتایج تحلیلی بدست آمده از ارزیابی کارایی سیستم در بخش قبلی مقایسه نموده ایم . در بخش ۵ نیز با ارائه یک نتیجه گیری به مقاله خاتمه داده ایم .

۲- مدل پیشنهادی

الگوی انعکاس داده ای پیشنهادی بکمک یک ساختار توزیع داده ای منحصر بفرد بلاکهای داده را بگونه ای در سطح دنباله دیسکی توزیع مینماید که سیستم بیشترین قابلیت اطمینان و تحمل پذیری نسبت به گونه های مختلف خطا را داشته باشد. در اغلب سیستم های ذخیره سازی داده ای توزیع شده ای که از سطوح و ساختارهای مختلف تکنیک RAID استفاده مینمایند ، تمامی عملیاتی داده ای سراسری از طریق کنترلر سراسری RAID^۷ (GRC) صورت میگیرد بگونه ای که امکان دسترسی

⁸ Remote Access

⁹ Remote Procedure Call

¹⁰ Single System Image

¹¹ Failure point

¹² Distribution

¹³ Transparency

¹⁴ Scalability

⁵ Mean Time To Loss Data

⁶ Mean Time To Recover Data

⁷ Global RAID Controller

RAID-1 را بدلیل مشابهت بسیاری که با مدل پیشنهادی دارند مورد تحلیل و ارزیابی قرار داده و آنها را از جنبه های گوناگونی همچون هزینه احیاء بلاکهای داده ای و دیسک در مقابل بروز گونه های مختلف خطا ، نرخ فقدان داده ای سیستم و ارزیابی کارایی و بازدهی سیستم ، مورد مقایسه قرار داده و نقاط ضعف و قوت مدل پیشنهادی را در هر یک از موارد ارزیابی و در برابر سایر مدل های مشابه شناسایی نماییم. برخی از پارامترهای بکاررفته در فرمولهای تحلیلی در جدول ۱ آورده شده است.

جدول ۱. جدول نمادها

نماد	توضیحات
N_d	تعداد دیسکهای دنباله دیسکی توزیع شده
N_b	متوسط تعداد بلاک در هر دیسک
$MTTLD$	متوسط زمان فقدان داده
$MTTRF$	متوسط زمان احیاء خرابی
T_{CM}	هزینه زمانی ارتباطی
T_{RW}	هزینه زمانی عملیات نوشتن راه دور
T_{LR}	هزینه زمانی عملیات خواندن محلی

۳-۱- هزینه عملیات احیاء

در این بخش هزینه عملیات احیاء بلاکهای داده ای دیسکهای دنباله توزیع شده را در اثر بروز خطاهای دیسکی مورد بررسی قرار میدهیم. ویژگی خاص مدل پیشنهادی در مقابل سایر مدل های دیگر آنست که نسخه های پشتیبان بلاکهای داده ای اصلی هر دیسک ، در سطح سایر دیسکهای دنباله توزیع گردیده اند. از این روی بهنگام بروز خرابی در هر یک از دیسکها، حداکثر پهنای باند سایر دیسکهای دنباله بمنظور انجام عملیاتهای احیاء داده ای بطور مؤثر در دسترس خواهد بود. بمنظور احیاء هر بلاک از دیسک مفقود، توسط بلاک پشتیبان نیاز است تا عملیات بازنویسی راه دور صورت گیرد [18][13]. هزینه عملیات احیاء داده ای ، در دو مرحله محاسبه میگردد. نخست، هزینه ارتباطی^{۱۸} عملیات احیاء که مربوط به عملیات انتشار پیام خرابی دیسک مفروض توسط کنترلر سراسری دنباله (GRC) به تمامی سایت های فعال و عملیات انتقال پایه بلاکهای داده ای سایت های راه دور می باشد که بطور موازی با یکدیگر انجام میگردد. هزینه محاسباتی^{۱۹} عملیات هم مربوط به عملیات خواندن محلی بلاک داده ای پشتیبان و بازنویسی آن روی بلاک مفقود راه دور میگردد که عملیات خواندن محلی در تمامی دیسکها بطور موازی و همزمان صورت میگردد اما عملیات بازنویسی بدلیل بکارگیری پروتکل قفل دو مرحله ای^{۲۰} (2PL) بمنظور حفظ پی در پی پذیری^{۲۱} جهت تامین صحت عملیات بازنویسی بلاکها باعث میگردد تا این عملیات بصورت سریال پیاده سازی گردد. در واقع این قسمت

هزینه های زمانی انجام عملیات احیاء دیسک یا بلاکهای داده ای مفقود ، کاهش نرخ احتمال فقدان داده ای و افزایش میزان قابلیت اطمینان و تحمل پذیری نسبت به خطا در مدل پیشنهادی از یک الگوی ذخیره داده ای توزیع شده استفاده خواهد شد. در این الگو هر دیسک به دو نیمه ، که در یکی بلاکهای داده ای اصلی و در دیگری بلاکهای داده ای پشتیبان نگهداری خواهد شد تقسیم میگردد. برای توزیع بلاکهای داده ای اصلی درون دنباله از الگوی ردیفی کردن داده ای^{۱۵} در سطح بلاکی و برای توزیع بلاکهای داده ای پشتیبان در نیمه دیگر، از پیاده سازی الگوی انعکاس داده ای^{۱۶} کمک فرمول زیر استفاده خواهد شد. در این فرمول موقعیت فیزیکی بلاک داده ای پشتیبان پس از ذخیره سازی بلاک داده ای اصلی معادل آن ، و بر اساس موقعیت فیزیکی آن بلاک تعیین خواهد شد.

$$M_{i,j'} = \begin{cases} \left[N_d - (j+1) \right] + \frac{N_b}{2} & ; \quad i \geq N_d - j \\ \left[N_d - (j+1) \right] - \left(1 - \frac{N_b}{2} \right) & ; \quad i < N_d - j \end{cases}$$

$$j' = [[(j+1) \bmod N_d] + i] \bmod N_d$$

فرمول ۲

بمنظور پیاده سازی عملیاتهای دیسکی از قبیل عملیات احیاء دیسک منفرد ، بروز رسانی بلاکهای پشتیبان و یا احیاء بلاکهای داده ای مفقود و با توجه به توزیع داده ای خاص این مدل درون دنباله توزیع شده ، نیاز است تا هر سیستم (که هر سیستم می تواند حاوی یک دیسک یا یک دنباله دیسکی متمرکز باشد) با سایر سیستم ها تبادل اطلاعاتی داشته و این نرخ از انتقال داده یک اتصال از نوع Fully Connected را برای چنین سیستمی میطلبد ، که پیاده سازی آن در یک سیستم توزیع شده با مقیاس پذیری بالا دشوار است. بکارگیری یک بستر ارتباطی با توپولوژی Hypercube H در این ساختار داده ای توزیع شده ، از یک سوی پیاده سازی مدل را امکان پذیر و ساده تر نموده و از سوی دیگر موجب کاهش قابل توجه هزینه های ارتباطی ۱۷ مدل بهنگام عملیاتهای داده ای راه دور نسبت به سایر مدل های ارتباطی خواهد شد [11][10].

۳- مقایسه ارزیابی کارایی مدل پیشنهادی

همانطور که قبلا هم اشاره گردید ، ما در این مقاله تکنیکهای RAID-10, RAID-X , Chained Declustering

¹⁸ Communication cost

¹⁹ Computational cost

²⁰ 2 Phase Locking

²¹ Serializability

¹⁵ Data stripping

¹⁶ Data Mirroring

¹⁷ Communication cost

که این مقدار در مقایسه با سایر مدل‌های مشابه دیگر نرخ حفاظت داده ای را بمقدار قابل توجهی بهبود بخشیده است [16][15].

۳-۳- تحلیل ارزیابی کارایی

با توجه به آنکه هدف از تحلیل و ارزیابی الگوی پیشنهادی، ارزیابی کارایی^{۲۴} مدل میباشد، سعی داریم رفتار مدل را در دراز مدت بکمک زنجیره مارکوف مورد بررسی قرار دهیم. در تحلیل کارایی مدل‌های مورد ارزیابی بکمک زنجیره مارکوف^{۲۵}، بدلیل آنکه عملکرد هر یک از دیسک‌های دنباله کاملاً از یکدیگر مستقل بوده و توزیع خطاهای منتشره در سطح دنباله دیسکی از توزیع پواسون برخوردار است و آنکه پارامتر زمانی فرآیند تصادفی در حالت زمانی پیوسته مورد تحلیل قرار داده می‌شود، بنابر این ارزیابی سیستم را بصورت فرآیند تصادفی مستقل با فضای حالت گسسته برای حالت زمانی پیوسته مورد بررسی مینماییم. با توجه به آنکه وقوع فرآیندهای تصادفی^{۲۶} (بروز خطاهای بلاکی و دیسکی درون دنباله) از یک توزیع پواسون برخوردار است، مدت زمان میان دو رخداد تصادفی متوالی (فاصله زمانی تا بروز خطای بعدی) دارای یک توزیع نمایی با پارامتر توزیع λ و نرخ سرویس دهی (احیاء خطا داده ای) به فرآیند تصادفی نیز دارای توزیعی مشابه، با پارامتر μ میباشد. لازم بذکر است، که هر یک از پارامترهای تصادفی بکار برده شده λ ، μ با مقادیر MTTLD، MTTRF معادل میباشد. علاوه بر این، نیاز است تا تحلیل از نظر توزیع موقعیت خطاهای داده ای در سطح بلاکی نیز مشخص گردد یعنی آنکه خطای بعدی بر روی کدام دیسک از دنباله می‌نشیند. که فاصله مکانی تا خطای بعدی بر روی دنباله نیز توزیعی نمایی دارد [20][19][3].

ارزیابی کارایی مدل پیشنهادی

بمنظور تحلیل الگوی معماری، ساختار ذخیره سازی، نرخ تحمل پذیری و نحوه عملکرد مدل پیشنهادی در قبال بروز دنباله های مختلف خطا و اثبات نتایج تحلیلی بدست آمده از فرمول‌های ارائه شده به ارزیابی کارایی مدل پیشنهادی پرداخته و ابتدا معادلات توازن سراسری جریان را در تحلیل دراز مدت زنجیره

$$\begin{cases} \pi_0 \cdot \lambda_0 = \pi_1 \cdot \mu_1 \\ \pi_0 \cdot \lambda_0 + \pi_2 \cdot \mu_2 = \pi_1 \cdot (\lambda_1 + \mu_1) \\ \pi_1 \cdot \lambda_1 + \pi_3 \cdot \mu_3 = \pi_2 \cdot (\lambda_2 + \mu_2) \\ \vdots \\ \pi_{i+1} \cdot \lambda_k = \pi_k \cdot \mu_k \\ \pi_{n-1} \cdot \lambda_{n-1} = \pi_n \cdot \mu_n \end{cases}$$

از عملیات احیاء به یک گلوگاه کارایی^{۲۲} برای عملیات احیاء تبدیل خواهد شد. که در نتیجه هزینه عملیات احیاء بصورت زیر محاسبه خواهد شد [14][12].

$$T(n) = O(Lg(n).T_{Cm}) + O((n-1).T_{RW} + T_{LR}) \quad \text{فرمول ۳}$$

۳-۲- نرخ فقدان داده ای

نرخ فقدان داده ای که ما آنرا بکمک یک مقدار احتمالی برای از دست دادن یک یا عده ای از بلاکهای داده ای دنباله دیسکی مشخص مینماییم، در مدل پیشنهادی بکمک الگوی خاص توزیع بلاکهای داده ای در سطح دنباله بمقدار قابل توجهی کاهش یافته است. در هر دو تکنیک RAID-1، RAID-10 که نخستین بار توسط Gibson و همکارانش [3] ارائه گردید و تکنیک پایه در آنها الگوی انعکاس داده ای میباشد، نرخ فقدان داده ای برای یک دنباله n دیسکی برابر با $1/n(n-1)$ میباشد که این مقدار در مدل پیشنهادی 'Chained Declustering' در [7] آمده است به مقدار $1/n(n-1)(n-2)$ کاهش یافته است. پس از آن در ساختار ارائه شده توسط Stonebarker (که در واقع ایده های اولیه RADD^{۲۳} را نیز برای نخستین بار در دانشگاه Berkeley ارائه کرد) که بمنظور بهبود انجام عملیات I/O موازی بر روی ساختارهای توزیع شده نیز ارائه گردید، بکمک تکنیک خاصی برای توزیع داده، نرخ فقدان داده ای در بهترین حالت به $1/n(n-1)(n-2)(n-3)$ تقلیل یافت. در [6] Kai Hwang یک معماری توزیع شده ارائه کرد که در آن با استفاده از یک الگوی توزیع داده ای چهارگانه توانست نرخ احتمالی فقدان داده ای را تا $1/n^2$ نرخ بدست آمده در [7] کاهش دهد.

در مدل پیشنهادی، بلاکهای داده ای بگونه ای در سطح دنباله توزیع میگرددند که هر دو دیسک از دنباله در دو بلاک با یکدیگر مشترک میباشدند. یکی از بلاکها، بلاک داده ای اصلی و دیگری نسخه پشتیبان بلاک داده ای دیسک دیگر است. بر این اساس، می توان نشان داد که در بدترین حالت خرابی n دیسک از دنباله، تعداد $(n-1)$ بلاک از دنباله مفقود خواهد گردید، که در نتیجه

$$\frac{2 \cdot [(n-1) + C_2^{n-1}]}{N_d \cdot N_b} \quad \text{فرمول ۴}$$

درصد از کل بلاکهای دنباله دیسکی از دست خواهد رفت. همانطور که در فرمول نیز مشخص است، تنها زمانی بلاکهای داده ای یک دیسک بطور کامل مفقود میگردد که $n = N_d$ باشد. با توجه به این موضوع می توان گفت که نرخ احتمالی فقدان داده ای در این مدل برابر است با:

$$\frac{1}{n \times (n-1) \times (n-2) \times \dots \times 2 \times 1} = \frac{1}{n!}$$

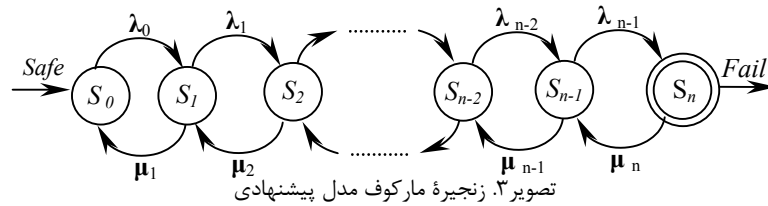
²⁴ Performance evaluation

²⁵ Markovian chain

²⁶ Random process

²² Performance bottleneck

²³ Redundancy Array of Distributed Disk



و کارایی سیستم را بر اساس آن ارزیابی مینماییم بدست خواهد آمد. از آنجاکه در این مدل λ_i و μ_i هر دو از حالت i مستقل هستند، در تمامی حالات آنها را با λ و μ نمایش میدهیم.

$$\sum_{i=0}^n P_{(i)} = 1 \Rightarrow \pi_0 = \left(1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_0 \cdot \lambda_1}{\mu_1 \cdot \mu_2} + \dots + \frac{\lambda_0 \cdot \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_{n-1}}{\mu_1 \cdot \mu_2 \cdot \mu_3 \cdot \dots \cdot \mu_n}\right)^{-1}$$

$$\pi_n = \frac{\left(\frac{\lambda_0 \cdot \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_{n-1}}{\mu_1 \cdot \mu_2 \cdot \mu_3 \cdot \dots \cdot \mu_n}\right)}{\left(1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_0 \cdot \lambda_1}{\mu_1 \cdot \mu_2} + \frac{\lambda_0 \cdot \lambda_1 \cdot \lambda_2}{\mu_1 \cdot \mu_2 \cdot \mu_3} + \dots + \frac{\lambda_0 \cdot \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_{n-1}}{\mu_1 \cdot \mu_2 \cdot \mu_3 \cdot \dots \cdot \mu_n}\right)}$$

$\lambda_i = \lambda, \mu_i = \mu \Rightarrow \pi_n = \frac{\left(\frac{\lambda}{\mu}\right)^n}{\left(1 + \left(\frac{\lambda}{\mu}\right) + \left(\frac{\lambda}{\mu}\right)^2 + \dots + \left(\frac{\lambda}{\mu}\right)^n\right)}$

اگر نسبت λ/μ را بعنوان نرخ بلاکهای در حال احیاء سیستم ، با پارامتر ρ نمایش دهیم . بمنظور حفظ Ergodicity سیستم (زنجیره ای Ergodic بوده که با شروع از هر یک از حالات، به پایداری برسد) بایستی $\rho < 1$ باشد تا ، چرا که به ازاء مقادیر $\rho > 1$ مقدار احتمالی گذار π_0 بسمت صفر میل کرده و موجب عدم پایداری سیستم خواهد شد.

$$P_{(n)} = \frac{\rho^n}{1 + \rho + \rho^2 + \dots + \rho^n} \Rightarrow P_{(n)} = \rho^n \cdot (1 - \rho)$$

با کمی دقت میتوان مشاهده کرد که نتایج تحلیلی بدست آمده از ارزیابی کارایی سیستم برای نرخ احتمالی گذار $P_{(n)}$ ، بخوبی ویژگی منحصر بفرد مدل پیشنهادی را در کاهش چشمگیر احتمال فقدان داده ای ، افزایش نرخ قابلیت احیاء بلاکهای داده ای و سطح تحمل پذیری سیستم نسبت به خطا را اثبات مینماید. یکی از ویژگیهایی که مدل پیشنهادی را برای پیاده سازی بر روی سیستم های توزیع شده برجسته تر میکند ، سیر نزولی نرخ احتمال گذار سیستم به وضعیت نهایی، بازاء افزایش پارامتر n (تعداد دیسکهای دنباله توزیع شده) در سیستم میباشد. نتایج تحلیلی ارایه شده در نمودار ۱ نشان میدهد که نرخ فقدان داده ای در مدل پیشنهادی بازاء مقادیر کمتر از یک، (بازاء نرخ بلاکهای در حال احیاء) در حدود صفر می باشد و میزان کارایی سیستم در این بازه تقریباً ۱۰٪ بوده و در واقع می توان گفت احتمال حضور سیستم در وضعیت نهایی $P_{(n)}$ غیر ممکن است. با افزایش نرخ بروز خطا نسبت به نرخ احیاء داده ای، روند فقدان داده ای سرعت افزایش یافته و در مقابل با کاهش پایداری

مارکوف بکمک ماتریس گذار یک مرحله ای برای حالات مختلف زنجیره مارکوف تصویر ۳ خواهیم نوشت.

همانگونه که در تصویر ۳ هم دیده می شود بدلیل تشابه الگوی رفتاری مدل پیشنهادی بهنگام بروز و احیاء خطا با فرایندهای تصادفی تولد و مرگ^{۲۷} می توان بکمک یک زنجیره مارکوف زمان پیوسته^{۲۸} به تحلیل عملکرد مدل پرداخت. در ماتریس نرخ گذار Q ، نرخ گذار از حالت K به $K+1$ $(Q_{k, k+1})$ برابر است با λ_k و نرخ گذار از حالت K به $K-1$ $(Q_{k, k-1})$ برابر است با μ_k و نرخ k که سیستم از حالت k خارج گردد $(Q_{k, k})$ برابر است با $-(\lambda_k + \mu_k)$. هر گذار سیستم به حالات بالاتر در اثر بروز یک خطا و هر گذار معکوس معادل احیاء خرابی می باشد. نرخ گذار تنها وابسته به وضعیت فعلی سیستم بوده و از مدت زمان درنگ سیستم در یک حالت خاص مستقل است . بر همین اساس، توزیع زمانهای میان بروز خطا و زمان سرویس (احیاء)، توزیع نمایی بوده و احتمال بروز بیش از یک خطا (یا احیاء)، در بازه زمانی کوچک t قابل چشم پوشی است و ما در زنجیره مارکوف از تحلیل این حالات خودداری نموده ایم. با توجه به

$$\pi_1 = \left(\frac{\lambda_0}{\mu_1}\right) \pi_0, \quad \pi_2 = \left(\frac{\lambda_0 \cdot \lambda_1}{\mu_1 \cdot \mu_2}\right) \cdot \pi_0$$

$$\pi_3 = \left(\frac{\lambda_0 \cdot \lambda_1 \cdot \lambda_2}{\mu_1 \cdot \mu_2 \cdot \mu_3}\right) \cdot \pi_0, \quad \pi_{n-1} = \left(\frac{\lambda_0 \cdot \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_{n-2}}{\mu_1 \cdot \mu_2 \cdot \mu_3 \cdot \dots \cdot \mu_{n-1}}\right) \cdot \pi_0$$

$$\pi_n = \left(\frac{\lambda_0 \cdot \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_{n-1}}{\mu_1 \cdot \mu_2 \cdot \mu_3 \cdot \dots \cdot \mu_n}\right) \cdot \pi_0$$

تصویر زنجیره مشخص میگردد که حالت S_n یک حالت نهایی بوده و در اثر بروز خرابی در n امین دیسک دنباله توزیع شده سیستم به حالت Failure وارد شده و هیچ امکانی برای احیاء بلاکهای داده ای مفقود وجود ندارد بر همین اساس مقادیر احتمالی گذار سیستم از هر یک از حالات زنجیره در تحلیل دراز مدت بصورت زیر خواهد بود.

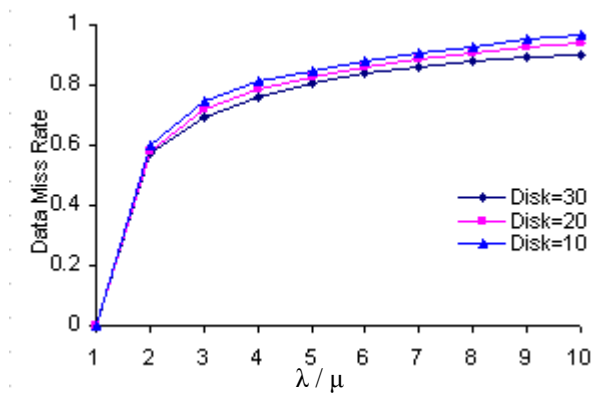
با حل دستگاه معادلات سراسری توازن، احتمال گذار حالات سیستم برحسب احتمال گذار از حالت آغازین بیان میگردد که با قراردادن مقادیر بدست آمده در رابطه تعادلی (مجموع احتمال گذار حالات سیستم برابر با ۱ است) زیر، احتمال گذار تمامی حالات و در نهایت نرخ احتمالی انتقال سیستم به وضعیت Failure (فقدان کامل دیسک) که ما یازدهمی

²⁷ Birth and death process

²⁸ Continuous-time markov chain

²⁹ Transition rate matrix

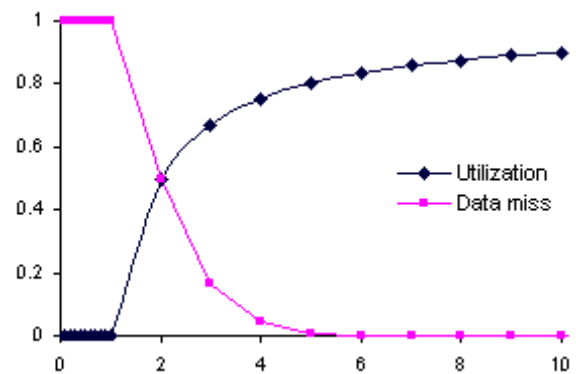
سازی که نتایج آن در نمودار ۲ ارائه گردیده است، ما سطح حفاظت داده ای الگوی پیشنهادی را درمقایسه با دو مدل مشابه دیگر RAID-10, Chained declustering قرار داده ایم و دیده می شود که تکنیک RAID-10 نسبت به دو مدل دیگر، پایین ترین سطح امنیتی را داراست. نرخ احیاء بلاکهای داده ای مفقود بازاء مقادیر کمتر از ۴ روی محور افقی (4?ρ) در دو مدل با هم برابر بوده ولی بازاء مقادیر 4 > ρ، الگوی Chained declustering قادر است بلاکهای داده ای را با نرخ بالاتری احیاء نماید. اما مدل پیشنهادی بازاء تمامی مقادیر، سطح حفاظتی بسیار بالاتری را در سیستم فراهم می آورد. اما تمامی مدلها رفتار مشابهی (روند نزولی) را در مقابل افزایش ρ از خود نشان میدهند. در نمودار ۳ نرخ فقدان داده ای مدل را بازاء تغییرات تعداد دیسک در سطح دنباله مورد بررسی قرار داده و مشاهده می شود که با افزایش تعداد دیسکهای دنباله، همانطور که در بخش ارزیابی کارایی سیستم نیز بدان اشاره گردید، نرخ فقدان داده ای خصوصاً با افزایش میزان بلاکهای در حال احیاء درون دنباله، کاهش چشمگیری خواهد داشت. نتایج بدست آمده از این گام از فرآیند شبیه سازی نشان می دهد که تکنیک توزیع داده ای ارائه شده در این مدل بمنظور پیاده سازی بر روی سیستم های توزیع شده مناسب بوده و الگوی توزیع مدل پیشنهادی از مقیاس پذیری بالایی (برای پیاده سازی در سیستم های مقیاس بالا) برخوردار است.



نمودار ۳. مقایسه نرخ فقدان داده ای بازاء تغییرات دیسک

در انتهای فرآیند شبیه سازی بمنظور اثبات صحت فرمولهای بدست آمده در بخش های قبل، و مقایسه نتایج تحلیلی ارزیابی کارایی سیستم با عملکرد مدل دریک سیستم حقیقی و هم بمنظور تعیین رفتار الگوی پیشنهادی در مقابل تغییرات افزایشی تعداد دیسکهای دنباله توزیع شده، نتایج فرمولاسیون را در کنار نتایج حاصل از شبیه سازی مدل در قالب نمودار ۴ آورده ایم. همانطور که در نمودارهای مرتبط با شبیه سازی و فرمولاسیون الگوی پیشنهادی دیده می شود، میزان کارایی سیستم و قابلیت احیاء بلاکهای داده ای با افزایش تعداد دیسکهای دنباله توزیع شده روند روبه رشدی داشته و از میزان فقدان داده ای سیستم کاسته می شود. بگونه ای که نرخ بازدهی

سیستم از میزان کارایی مدل کاسته خواهد شد. همانطور که در نمودار ۱ نیز دیده می شود، نتایج بدست آمده از شبیه سازی مدل، پیش بینی های قبلی پیرامون نرخ فقدان داده ای و نتایج تحلیلی حاصل از ارزیابی کارایی و فرمولهای ارائه شده در بخشهای قبلی را بخوبی تایید کرده می توان گفت که این همگونی میان نتایج بدست آمده بر صحت مدلسازی سیستم و روابط بدست آمده از ارزیابی کارایی آن اشاره دارد.

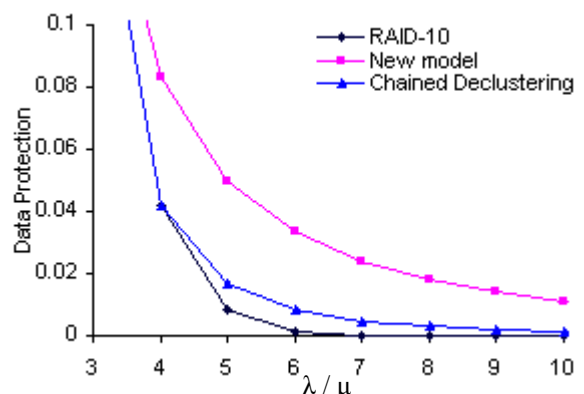


نمودار ۱. ارزیابی کارایی و نرخ فقدان داده ای سیستم

۴- شبیه سازی مدل پیشنهادی

بمنظور حصول اطمینان از نتایج بدست آمده در مراحل مدلسازی تحلیلی و ارزیابی کارایی الگوی ذخیره داده ای پیشنهادی و بررسی و مقایسه نتایج بدست آمده با سایر مدلهای مشابه، فرآیند شبیه سازی را طی مراحل مختلفی انجام داده ایم. و در هر مرحله رفتار مدلها را در مقابل گونه های مختلف خطا آزموده و هرآزمون را بمنظور افزایش دقت فرآیند شبیه سازی تا ۱۰۰۰۰ تکرار انجام داده ایم. ابتدا همانگونه که در نمودار ۱ مشاهده می شود، نرخ فقدان داده ای سیستم و بازدهی مدل بازاء تغییرات (۱۰ تا ۱) پارامتر ρ، مورد بررسی قرار گرفت و نتایج حاصله، صحت پیش بینی های صورت گرفته براساس فرمولهای تحلیلی بدست آمده از بخشهای قبلی را تایید کرد. در بازه ۰ تا ۱ نرخ فقدان داده ای مدل در حد ضفر و بازدهی سیستم ۱۰۰٪ بوده و بازاء مقادیر بالاتر از ۱ نرخ فقدان داده ای روند عودی را در پی خواهد داشت. که بطور معکوس از بازدهی سیستم میکاهد [3][2].

صحت و امنیت داده ای بالاتر، درگام دیگر از فرآیند شبیه



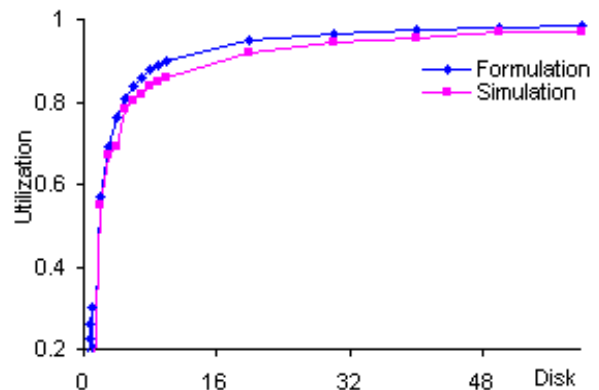
نمودار ۲. مقایسه سطح حفاظت داده ای مدلها مختلف

فقدان داده ای کمتر، قابلیت اطمینان بالاتر، و دسترس پذیری بهتری را برای داده های بحرانی درون سیستم فراهم آورده و در مجموع سطح کارایی و قابلیت اطمینان سیستم را بهبود بخشید.

مراجع

- [1] P. M. Chen, E. K. Lee, G. A. Gibson, R. H. Katz and D. A. Patterson; "RAID: High-Performance, Reliable Secondary Storage", ACM Computing Surveys, Vol.26, No.2, June 1994, pp.145-185.
- [2] J.Akbari and M.Meybodi . " A New Data Mirroring Algorithm to Enhance Reliability and Fault tolerance", The Second International conference on Information and Knowledge Technology, Amirkabir university of technology, May 2005.
- [3] G. Gibson, D. Nagle, K. Amiri, F. Chang, H. Gobioff, E. Riedel, D. Rochberg and J. Zelenka, "A Cost-effective, High-bandwidth Storage Architecture", Proc. of the 8th Conf. on Architectural Support for Programming Languages and Operating Systems, 1998, pp.97-106.
- [4] J.Akbari and A.T.Haghighat . " A New Redundancy Algorithm For Distributed Environment", Operating system & Security Conference-OSSC 2003, Sharif university of technology, Dec 2003, pp.70-81.
- [5] J.Akbari and M.Meybodi . "A New Redundancy Structure for Using in High Reliable Systems", The Second International conference on Information and Knowledge Technology, Amirkabir university of technology, May 2005.
- [6] K. Hwang, H. Jin, E. Chow, C. L. Wang, and Z. Xu , "Designing SSI Clusters with Hierarchical Checkpointing and Single I/O Space". IEEE Concurrency Magazine, March 1999, pp.60-69.
- [7] H. I. Hsiao and D. DeWitt, "Chained Declustering: A New Availability Strategy for Multiprocessor Database Machines", Proc. of 6 th Int'l Data Eng. Conf., 1990, pp.456-465.
- [8] H. Jin and K. Hwang, "Optimal Striping in RAID Architecture", Concurrency: Practice and Experience ,Vol.12, No.10, August 2000, pp.909-916.
- [9] R. P. Martin, A. M. Vahdat, D. E. Culler, T. E. Anderson, "Effects of Communication Latency, Overhead, and Bandwidth in a Cluster Architecture", Proc. of the 24th Annual International Symp. on Computer Architecture, June 1997, pp.85-97.
- [10] P. F. Corbett, D. G. Feitelson, J.-P. Prost, and S. J. Baylor. "Parallel Access to Files in the Vesta File System". Proceedings of Supercomputing'93, 1993, pp.101-111.
- [11] P. Cao, S. B. Lim, S. Venkataraman, and J. Wilkes, "The TickerTAIP Parallel RAID Architecture", ACM Trans. on Computer System, Vol.12, No.3, August 1994, pp.236-269.
- [12] J. H. Howard, M. L. Kazar, S. G. Menees, D. A. Nichols, M. Satyanarayanan, R. N. Sidebotham, and M. J. West, "Scale and Performance in a Distributed File System". ACM Trans. on Computer System, Vol.6, No.1, February 1988, pp.51-81.

سیستم بسمت یک میل مینماید و این مقیاس پذیری سیستم، در مقابل افزایش تعداد دیسکهای درون دنباله، الگوی آرایه شده را برای پیاده سازی بر روی سیستم های توزیع شده مناسب تر خواهد کرد. با توجه به نمودار می توان دید که نحوه عملکرد مدل بازنه تغییرات تعداد دیسکهای درون دنباله و پاسخ سیستم به تغییرات نرخ خطا درون سیستم (روند افزایشی یا کاهششی نرخ p)، در فرایند شبیه سازی با نتایج تحلیلی بدست آمده از فرمولاسیون کاملا منطبق است.



نمودار ۴. مقایسه نتایج ارزیابی تحلیلی و شبیه سازی مدل

۵- نتیجه گیری

با توجه به نتایج بدست آمده از مدل سازی ساختاری و پیاده سازی تحلیلی تکنیک توزیع الگوی پیشنهادی و نتایج حاصل از تحلیل ارزیابی کارایی سیستم و انطباق نتایج حاصله با نتایج شبیه سازی می توان دریافت که مدل پیشنهادی با بکارگیری این ساختار توزیع داده ای خاص در ذخیره سازی بلاکهای داده ای، (خصوصا برای پیاده سازی در سیستم های توزیع شده) قادر است تا با افزایش پهنای باند مؤثر سیستم، نرخ دسترسی به داده ها را خصوصا بهنگام بروز خطا و در حین عملیات احیاء خطا بمقدار قابل توجهی بهبود بخشد. علاوه بر این، ساختار توزیع داده ای بکار رفته در مدل، باعث افزایش سطح بازدهی سیستم و کاهش نرخ فقدان داده ای و در نتیجه افزایش نرخ قابلیت اطمینان و تحمل پذیری سیستم نسبت به خطا خواهد شد. لازم بذکر است که الگوی پیشنهادی با توجه به تمامی ویژگیهای بارز مذکور، علاوه برآنکه نسبت به سایر مدل های مشابه هیچگونه هزینه فضایی و زمانی افزوده ای را به سیستم تحمیل نمی کند، بلکه هر دیسک را قادر می سازد تا عملیات بروز رسانی بلاکهای پشتیبان و احیاء داده ای را بصورت محلی و بدون دخالت کنترلر GRC انجام دهد. مقیاس پذیری سیستم در مقابل تغییرات افزایشی تعداد دیسکهای درون دنباله، الگو را برای پیاده سازی بر روی سیستم های توزیع شده مناسب ترمی کند. در مجموع، با توجه به ارزیابی کارایی، مقایسات تحلیلی و نتایج حاصل از شبیه سازی سیستم می توان گفت، الگوی آرایه شده، قادر است نسبت به بهترین مدل های مشابه قبلی، بازدهی و پهنای باند بیشتر، نرخ

International Parallel and Distributed Processing Symposium , Mexico, May 2000,pp.163-174.

[18] T. Anderson, M. Dahlin, D. Patterson, and R. Wang. "Serverless Network File Systems", ACM Trans. on Computer Systems, Jan. 1996, pp.41-79.

[19] S. Asami, N. Talagala, and D. A. Patterson, "Designing a self-maintaining storage system", Proceedings of 16th IEEE Symposium on Mass Storage Systems, March 1999, pp. 222-233.

[20] R. S. Ho, K. Hwang, and H. Jin, "Design and Analysis of Clusters with Single I/O Space", Proceedings of 20th International Conference on Distributed Computing Systems (ICDCS 2000), April 2000, Taiwan, pp.120-127.

[13] L. F. Cabrera, and D. E. Long, "Using Distributed Disk Striping to Provide High I/O Data Rates", Proceedings of USENIX Computing Systems, Fall 1991, pp.405-433.

[14] R. W. Watson and R. A. Coyne, "Parallel I/O Architecture of the High Performance Storage System", Proc. of the 14th IEEE Symp. on Mass Storage Systems, Sept. 1995, pp.27-44.

[15] T. H. Cormen and D. Kotz, "Integrating Theory and Practice in Parallel File Systems", Proceedings of DAGS '93 Symposium, June 1993, pp. 64-74.

[16] I. Foster, D. Kohr, Jr., R. Krishnaiyer, and J. Mogill, "Remote I/O: Fast Access to Distant Storage". Proc. of the Fifth Workshop on I/O in Parallel and Distributed Systems, November 1997, pp.14-25.

[17] N. Muppalaneni and K. Gopinath, "A Multi-tier RAID Storage System with RAID1 and RAID5", Proceedings of 14th