

Uma Revisão Sistemática das Abordagens de Previsão da Disponibilidade de Recursos Computacionais em Sistemas Distribuídos: O Uso da Carta de Controle I-MR e Técnicas de Controle Estatístico de Processos

1 Introdução

A disponibilidade de recursos computacionais é um aspecto crucial em sistemas distribuídos, uma vez que afeta diretamente o desempenho, a confiabilidade e a eficiência desses sistemas. A capacidade de prever a disponibilidade desses recursos de forma precisa e confiável desempenha um papel fundamental no gerenciamento eficaz desses ambientes computacionais. A previsão da disponibilidade de recursos possibilita a tomada de decisões informadas sobre a alocação, escalabilidade e otimização dos recursos, além de permitir a detecção precoce de possíveis falhas ou gargalos.

Nesse contexto, a aplicação de técnicas de controle estatístico de processos tem sido amplamente explorada como uma abordagem promissora para o monitoramento de recursos dos mais diversos tipos. As cartas de controle é uma ferramenta estatística que permite monitorar e controlar a variabilidade dos dados coletados ao longo do tempo, facilitando a identificação de desvios e anomalias nos recursos computacionais. O controle estatístico de processos, por sua vez, abrange uma série de técnicas e métodos estatísticos, incluindo-se as cartas de controle, para analisar e melhorar a estabilidade e o desempenho dos processos.

Diante desse cenário, esta revisão sistemática tem como objetivo identificar as abordagens mais eficazes e inovadoras para a previsão da disponibilidade de recursos computacionais em sistemas distribuídos, com foco no uso da carta de controle I-MR e técnicas de controle estatístico de processos. Além disso, pretende-se avaliar a qualidade e a confiabilidade dos estudos que investigaram o uso dessas abordagens, identificar as técnicas de controle estatístico de processos mais aplicadas e investigar as práticas e desafios atuais relacionados à previsão da disponibilidade de recursos computacionais nesse contexto.

Por meio dessa revisão sistemática, busca-se fornecer uma análise abrangente e atualizada sobre as abordagens de previsão da disponibilidade de recursos computacionais em sistemas distribuídos, bem como propor recomendações e diretrizes para o aprimoramento do gerenciamento de recursos nesses ambientes. Com base na análise crítica dos estudos selecionados, espera-se contribuir para o avanço do conhecimento nessa área e para o aprimoramento das práticas de previsão da disponibilidade de recursos computacionais.

2 Metodologia

2.1 Definição dos critérios de busca

Para realizar a busca dos estudos relevantes, foram definidos critérios específicos para garantir a inclusão de estudos pertinentes à temática da previsão da disponibilidade de recursos computacionais em sistemas distribuídos. Os critérios de busca incluíram palavras-chave relacionadas ao tema, bem como filtros de idioma e período de publicação.

Aqui é definida uma lista de palavras-chave relevantes para o estudo, incluindo sinônimos e variações. Em seguida, é selecionada as bases de dados que serão utilizadas para fazer a pesquisa bibliográfica.

Termos de pesquisa:

- Disponibilidade de recursos computacionais
- Previsão de recursos
- Carta de controle I-AM (Individual com Amplitudes Móveis) e I-MR(Individual com Moving Range)
- Controle estatístico de processos
- Gerenciamento de recursos computacionais
- Análise de desempenho
- Sistemas distribuídos

Sinônimos e variações:

- Recursos de computação
- Capacidade computacional
- Controle estatístico de qualidade
- Gráfico de controle I-AM (Individual com Amplitudes Móveis)
- Gerenciamento de desempenho
- Sistema distribuído em nuvem
- Cloud computing

Objetivo da revisão sistemática

- Identificar as abordagens mais eficazes e inovadoras para previsão da disponibilidade de recursos computacionais em sistemas distribuídos;
- Avaliar a qualidade e a confiabilidade dos estudos que investigaram o uso da carta de controle I-AM para previsão da disponibilidade de recursos computacionais em sistemas distribuídos;
- Identificar as técnicas de controle estatístico de processos mais aplicadas em sistemas distribuídos;
- Investigar as práticas e desafios atuais relacionados à previsão da disponibilidade de recursos computacionais em sistemas distribuídos;
- Propor recomendações e diretrizes para a melhoria do gerenciamento de recursos computacionais em sistemas distribuídos por meio da previsão da disponibilidade de recursos com o uso da carta de controle I-AM e técnicas de controle estatístico de processos.

Perguntas de pesquisa

- Quais são os métodos mais utilizados para previsão da disponibilidade de recursos computacionais em sistemas distribuídos?
- Quais são os principais Modelos estocásticos de processos aplicados?
- Quais foram as medidas comuns de avaliação da qualidade de ajuste em modelos de regressão ou de previsão?
- Quais são as tendências e desafios atuais relacionados à previsão da disponibilidade de recursos computacionais em sistemas distribuídos?

2.2 Fontes de informação

Foram selecionadas diversas bases de dados acadêmicas e industriais para realizar a busca dos estudos relevantes. Exemplos de fontes de informação incluem a IEEE Xplore, ACM Digital Library, SpringerLink, ScienceDirect, Scopus, Web of Science e Wiley. Além disso, foram consultadas conferências e periódicos relevantes na área de sistemas distribuídos e computação.

2.3 Estratégia de busca

A estratégia de busca envolveu a combinação de termos de busca relacionados à previsão da disponibilidade de recursos computacionais, sistemas distribuídos, computação em nuvem e técnicas de controle estatístico de processos. As palavras-chave foram adaptadas conforme a sintaxe e os recursos de cada base de dados.

Strings de busca:

- "Disponibilidade de recursos computacionais" AND "Carta de controle I-AM"
- "Previsão de recursos" AND "Controle estatístico de processos"
- "Gerenciamento de recursos computacionais" AND "Análise de desempenho"
- "Sistemas distribuídos" AND "Carta de controle I-AM"
- "Computação em nuvem" AND "Controle estatístico de processos"
- ("Disponibilidade de recursos computacionais" OR "Recursos de computação" OR "Capacidade computacional") AND ("Carta de controle I-AM" OR "Gráfico de controle I-AM" OR "Controle estatístico de qualidade")
- ("Previsão de recursos" OR "Gerenciamento de recursos computacionais" OR "Gerenciamento de desempenho") AND ("Controle estatístico de processos" OR "Carta de controle I-AM" OR "Gráfico de controle I-AM")
- ("Análise de desempenho" OR "Sistemas distribuídos" OR "Sistema distribuído em nuvem" OR "Cloud computing") AND ("Carta de controle I-AM" OR "Gráfico de controle I-AM" OR "Controle estatístico de qualidade")

String de busca em inglês

- "Availability of computational resources" AND "I-MR control chart"
- "Resource Forecasting" AND "Statistical Process Control"
- "Computing resource management" AND "Performance analysis"
- "Distributed Systems" AND "I-MR Control Chart"
- "Cloud Computing" AND "Statistical Process Control"
- ("Availability of Computing Resources" OR "Computing Resources" OR "Computing Capacity") AND ("I-MR Control Chart" OR "I-MR Control Chart" OR "Statistical Quality Control")
- ("Resource Forecasting" OR "Computing Resource Management" OR "Performance Management") AND ("Statistical Process Control" OR "I-MR Control Chart" OR "I-MR Control Chart")
- ("Performance analysis" OR "Distributed systems" OR "Distributed cloud system" OR "Cloud computing") AND ("I-MR control chart" OR "I-MR control chart" OR "Statistical quality control")

2.4 Processo de seleção dos estudos

Após a realização da busca inicial, os estudos foram selecionados em duas etapas: triagem com base nos títulos e resumos, seguida pela análise completa dos artigos selecionados. Durante a triagem, os critérios de inclusão foram aplicados para identificar os estudos potencialmente relevantes. Na análise completa, os estudos foram avaliados em relação aos critérios de inclusão e exclusão para garantir a seleção dos estudos mais adequados ao escopo da revisão.

Critérios de inclusão:

- Tema: O trabalho deve abordar o tema de monitoramento ou previsão de recursos computacionais.
- Método: O trabalho deve utilizar a carta de controle I-MR ou alguma variação de carta de controle ou alguma ferramenta ou método de controle estatístico de processo.;
- Ano: Artigos publicados entre 2018 e 2023;
- Idioma: Artigos publicados em português ou inglês;

Critérios de exclusão:

- Tema: Estudos que não abordem a temática central da pesquisa;
- Disponibilidade: o trabalho deve estar disponível integralmente;
- Integridade: Artigos que não estejam disponíveis integralmente;
- Duplicidade: Artigos duplicados.

ID	Autor	Título	Ano da publicação
A1	Kansal, P. and Kumar, M. and Verma, O.P.	Classification of resource management approaches in fog/edge paradigm and future research prospects: a systematic review	2022
A2	Djigal, H. and Xu, J. and Liu, L. and Zhang, Y.	Machine and Deep Learning for Resource Allocation in Multi-Access Edge Computing: A Survey	2022
A3	Jiao, Y. and Wang, P. and Niyato, D. and Suankaewmanee, K.	Auction mechanisms in cloud/fog computing resource allocation for public blockchain networks	2019
A4	Guo, S. and Qi, Y. and Yu, P. and Shao, S. and Qiu, X.	Edge Network Resource Synergy for Mobile Blockchain in Smart City	2020
A5	Gao, Z. and Lin, B. and Xiao, K. and Wang, Q. and Mo, Z. and Yang, Y.	A dynamic resource allocation algorithm based on auction model in mobile blockchain network	2019
A6	Xu, Y. and Zhu, K. and Li, S.	Hierarchical Combinatorial Auction in Computing Resource Allocation for Mobile Blockchain	2020

A7	Wang, H. and Xu, H. and Huang, H. and Chen, M. and Chen, S.	Robust Task Offloading in Dynamic Edge Computing	2023
A8	Fan, Y. and Jin, Z. and Shen, G. and Hu, D. and Shi, L. and Yuan, X.	Three-stage Stackelberg game based edge computing resource management for mobile blockchain	2021
A9	Chen, X. and Lin, J. and Lin, B. and Xiang, T. and Zhang, Y. and Huang, G.	Self-learning and self-adaptive resource allocation for cloud-based software services	2019
A10	Abdellatif, K. and Abdelmouttalib, C. and EL-Azouzi R. and Sadik M. and Sabir E. and Tmiri A. and Altman E.	Graph-Based Computing Resource Allocation for Mobile Blockchain	2019
A11	Xiong, Z. and Feng, S. and Niyato, D. and Wang, P. and Han, Z.	Optimal pricing-based edge computing resource management in mobile blockchain	2018
A12	Jiao, Y. and Wang, P. and Niyato, D. and Xiong, Z.	Social Welfare Maximization Auction in Edge Computing Resource Allocation for Mobile Blockchain	2018
A13	Hanzelik, P.P. and Kummer, A. and Abonyi, J.	Edge-Computing and Machine-Learning-Based Framework for Software Sensor Development	2022
A14	Li, Y. and Wang, L. and Lee, D. and Jin, R.	Monitoring Runtime Metrics of Fog Manufacturing via a Qualitative and Quantitative (QQ) Control Chart	2022
A15	Ibidunmoye, O. and Rezaie, A.-R. and Elmroth, E.	Adaptive anomaly detection in performance metric streams	2018
A16	Kuang, Y. and Das, D. and Sir, M. and Pasupathy, K.	Likelihood ratio-based CUSUM charts for real-time monitoring the quality of service in a network of queues	2023
A17	Ahmad, H. and Ahmadi Nadi, A. and Amini, M. and Sadeghpour Gildeh, B.	Monitoring processes with multiple dependent production lines using time between events control charts	2023
A18	Yashchin, E.	Gradient analysis of Markov-type control schemes and its applications	2022
A19	Belay, M.A. and Blakseth, S.S. and Rasheed, A. and Salvo Rossi, P.	Unsupervised Anomaly Detection for IoT-Based Multivariate Time Series: Existing Solutions, Performance Analysis and Future Directions	2023
A20	Liao, M.-Y. and Wu, C.-W.	Process capability monitoring and change-point analysis for S-type quality characteristic	2023
A21	Weiβ, C.H. and Testik, M.C.	Nonparametric Control Charts for Monitoring Serial Dependence based on Ordinal Patterns	2023

A22	Pourmajidi, W. and Erwin, T. and Steinbacher, J. and Miranskyy, A. and Featherstone J. and Lyons K. and Wigglesworth J. and Saraf D.	On challenges of cloud monitoring	2020
A23	Monshizadeh Naeen, H. and Zeinali, E. and Toroghi Haghighat, A.	Adaptive Markov-based approach for dynamic virtual machine consolidation in cloud data centers with quality-of-service constraints	2020
A24	Monshizadeh Naeen, H. and Zeinali, E. and Toroghi Haghighat, A.	A stochastic process-based server consolidation approach for dynamic workloads in cloud data centers	2020
A25	Alam, M.R. and Gerostathopoulos, I. and Prehofer, C. and Attanasi, A. and Bures, T.	A framework for tunable anomaly detection	2019
A26	Jeske, D.R. and Stevens, N.T. and Tartakovsky, A.G. and Wilson, J.D.	Statistical methods for network surveillance	2018
A27	P. A. Bommanavar and R. Krishnan and S. Shahedha	Optimal Resource Estimation Policy Selection for Ecommerce Applications in Cloud	2019
A28	J. Li and W. Liang and Y. Ma	Robust Service Provisioning With Service Function Chain Requirements in Mobile Edge Computing	2021
A29	L. Wang and Y. Zhang and R. Jin	A Monitoring System for Anomaly Detection in Fog Manufacturing	2020
A30	Mishra, Shashank and Gupta, Aman and Jairam Naik, K. and Mandal, Jyotsna Kumar and De, Debashis	Efficient Computing Resource Sharing for Mobile Edge-Cloud Computing Networks	2023
A31	Benmakrelouf, Souhila and St-Onge, Cédric and Kara, Nadja and Tout, Hanine and Edstrom, Claes and Lemieux, Yves	Abnormal behavior detection using resource level to service level metrics mapping in virtualized systems	2020
A32	Benmakrelouf, Souhila and Kara, Nadja and Tout, Hanine and Rabipour, Rafi and Edstrom, Claes	Resource needs prediction in virtualized systems: Generic proactive and self-adaptive solution	2019
A33	Toumi, Hajer and	RTSLPS: Real time server load prediction system for	2022

	Brahmi, Zaki and Gammoudi, Mohhamed Mohsen	the ever-changing cloud computing environment	
A34	Teoh, J.W. and Teoh, W.L. and Khoo, Michael B.C. and Castagliola, Philippe and Moy, W.H.	On designing an optimal SPRT control chart with estimated process parameters under guaranteed in-control performance	2022
A35	You, Huay Woon	Optimal estimated process parameters side sensitive group runs chart based on expected average run length	2018

2.5 Avaliação da qualidade dos estudos

Os estudos selecionados foram submetidos a uma avaliação da qualidade e confiabilidade. Foram considerados critérios como a rigorosidade metodológica, a validade dos resultados, a clareza na apresentação dos métodos utilizados e a contribuição para o campo de estudo. Estudos de alta qualidade foram priorizados para análise e discussão na revisão sistemática.

2.6 Extração e análise dos dados

Após a seleção dos estudos, os dados relevantes foram extraídos e organizados em uma planilha para permitir uma análise sistemática. As informações extraídas incluíram detalhes sobre os estudos (autor, ano, contexto de aplicação, métodos utilizados), as abordagens de previsão de disponibilidade de recursos, as técnicas de controle estatístico de processos empregadas, as métricas avaliadas e os principais resultados obtidos.

Perguntas de pesquisa

- Quais são os métodos mais utilizados para previsão da disponibilidade de recursos computacionais em sistemas distribuídos?
 - Aprendizado profundo (deep learning): é uma subárea do aprendizado de máquina que utiliza redes neurais artificiais com múltiplas camadas para aprender representações abstratas dos dados e realizar tarefas complexas, como classificação, regressão, geração e reconhecimento de padrões. Alguns dos artigos que usam ou propõem modelos de aprendizado profundo para previsão da disponibilidade de recursos são [A13], [A19] e [A30].
 - Mecanismos de leilão: são modelos baseados em teoria dos jogos e microeconomia que permitem a alocação eficiente de recursos escassos entre agentes racionais que competem por eles. Alguns dos artigos que usam ou propõem mecanismos de leilão para previsão da disponibilidade de recursos são [A5], [A6], [A11] e [A12].

- Modelos estatísticos: são uma representação matemática de um fenômeno real que usa conceitos e métodos estatísticos para descrever e explicar a relação entre variáveis. Esses modelos são baseados em dados observados e tentam capturar as características essenciais do processo subjacente ao fenômeno estudado. Alguns dos artigos que usam ou propõem esses modelos para previsão da disponibilidade de recursos são [A7], [A8], [A14], [A15], [A16], [A17], [A18], [A23], [A24] e [A25].
- Quais são os principais Modelos estocásticos de processos aplicados?
 - Análise de regressão: é um método que estuda a relação entre uma variável dependente e uma ou mais variáveis independentes, buscando estimar ou prever o valor médio da variável dependente em função dos valores das variáveis independentes. Alguns dos artigos que usam ou propõem análise de regressão para previsão da disponibilidade de recursos são [A9], [A10] e [A27].
 - Cartas de controle: são ferramentas que permitem monitorar a qualidade de um processo por meio da análise de variáveis que indicam o seu desempenho. Eles consistem em gráficos que mostram os limites superior e inferior de controle, que representam os valores aceitáveis para a variável monitorada, e os pontos que correspondem às observações da variável ao longo do tempo. Alguns dos artigos que usam ou propõem gráficos de controle para previsão da disponibilidade de recursos são [14], [15], [16], [17], [18] e [20].
 - Testes de hipóteses: são métodos que permitem testar se uma afirmação sobre uma população é verdadeira ou falsa, com base em uma amostra. Eles envolvem a formulação de duas hipóteses opostas, a nula e a alternativa, e a escolha de um nível de significância que determina a probabilidade máxima de rejeitar a hipótese nula quando ela é verdadeira. Alguns dos artigos que usam ou propõem testes de hipóteses para previsão da disponibilidade de recursos são [19], [20], [21] e [25].
- Quais foram todas as medidas comuns de avaliação da qualidade de ajuste em modelos de regressão ou de previsão?
 - R-quadrado: é uma medida que representa a porcentagem de variação na variável resposta que é explicada pelo modelo. Quanto mais alto o R-quadrado, melhor o modelo se ajusta aos dados. Alguns dos artigos que usam ou propõem R-quadrado para avaliação da qualidade de ajuste são [A1], [A2], [A5], [9], [10], [11], [12], [13], [15], [A18], [A19], [21], [22], [23], [A24], [26], [27], [A30], [A32], [A33].
 - R-quadrado ajustado: é uma medida que ajusta o R-quadrado pelo número de preditores no modelo. Ele evita o problema de o R-quadrado sempre aumentar quando se adiciona mais preditores ao modelo, mesmo que eles não sejam significativos. Alguns dos artigos que usam ou propõem R-quadrado ajustado para avaliação da qualidade de ajuste são 1, 2, 3, 4, [A5], [9], [10], [11], [12], [13], [A16], [A17], [A18], [A24], [A26] e [A34].

- R-quadrado predito: é uma medida que usa a soma dos quadrados dos erros de previsão (PRESS) para estimar a capacidade do modelo de prever novas observações. Quanto mais alto o R-quadrado predito, melhor o modelo se ajusta aos dados. Alguns dos artigos que usam ou propõem R-quadrado predito para avaliação da qualidade de ajuste são 1, 2, 3 e 4.
 - AICc e BIC: são medidas que penalizam o modelo pela complexidade (número de parâmetros) e pelo tamanho da amostra. Elas visam encontrar o modelo mais parcimonioso que se ajusta bem aos dados. Quanto menor o AICc ou o BIC, melhor o modelo se ajusta aos dados. Alguns dos artigos que usam ou propõem AICc ou BIC para avaliação da qualidade de ajuste são 1, 2, 3 e 4.
 - Teste S: É um teste que verifica se o desvio padrão dos resíduos (S) é significativamente diferente de zero. Se S for igual a zero, significa que não há erro no modelo e todos os pontos de dados caem na linha ajustada. Se S for diferente de zero, significa que há erro no modelo e os pontos de dados se desviam da linha ajustada. Alguns dos artigos que usam ou propõem teste S para avaliação da qualidade de ajuste são 1, 2, 3 e 4.
 -
 - Teste R-quadrado: É um teste que verifica se o R-quadrado é significativamente diferente de zero. Se R-quadrado for igual a zero, significa que o modelo não explica nada da variação na variável resposta. Se R-quadrado for diferente de zero, significa que o modelo explica alguma variação na variável resposta. Alguns dos artigos que usam ou propõem teste R-quadrado para avaliação da qualidade de ajuste são 1, 2, 3 e 4.
- Quais são as tendências e desafios atuais relacionados à previsão da disponibilidade de recursos computacionais em sistemas distribuídos?
 - Machine learning: é um método de análise de dados que automatiza a construção de modelos analíticos, baseado na ideia de que sistemas podem aprender com dados, identificar padrões e tomar decisões com o mínimo de intervenção humana. O machine learning pode ser usado para prever o comportamento dos recursos computacionais a partir de dados históricos ou em tempo real, otimizando o uso e a alocação dos recursos. Alguns dos desafios do machine learning são a qualidade e a quantidade dos dados, a escolha dos algoritmos e dos parâmetros, a interpretação e a validação dos resultados 1.
 - Computação em grade: é uma forma de computação distribuída que consiste em computadores de alta performance que compartilham recursos e capacidade de processamento entre si. A computação em grade permite que os pesquisadores usem a potência combinada dos computadores para realizar tarefas que exigem muitos recursos, como simulações matemáticas. Alguns dos desafios da computação em grade são a segurança, a confiabilidade, a escalabilidade e a coordenação dos recursos 2.
 - Computação quântica: é uma forma de computação baseada nos princípios da mecânica quântica, que usa qubits em vez de bits para representar e manipular

informações. A computação quântica promete resolver problemas que são intratáveis para os computadores clássicos, como criptografia, otimização e inteligência artificial. Alguns dos desafios da computação quântica são a construção, a operação e a integração dos dispositivos quânticos, a correção de erros, a programação e a segurança dos dados 3.

2.7 Síntese dos resultados

Os resultados dos estudos foram sintetizados e analisados com base nos objetivos da revisão sistemática. As principais descobertas foram agrupadas e organizadas de acordo com os tópicos relevantes para responder às perguntas de pesquisa. A análise dos resultados levou em consideração a efetividade das abordagens, as técnicas de controle estatístico de processos mais aplicadas, os desafios identificados e as tendências atuais na previsão

3 Resultados

4 Análise dos Estudos

5 Discussão

6 Conclusão

Referências Bibliográficas