

文章编号:1671-251X(2020)05-0104-05

DOI:10.13272/j.issn.1671-251x.2019110054

负载均衡算法在智慧矿山软件平台中的应用

郭秀才¹, 张悦¹, 贺耀宜^{2,3}

- (1. 西安科技大学 电气与控制工程学院, 陕西 西安 710054;
2. 中煤科工集团常州研究院有限公司, 江苏 常州 213015;
3. 天地(常州)自动化股份有限公司, 江苏 常州 213015)



扫码移动阅读

摘要:针对现有负载均衡算法在处理智慧矿山系统数据时存在处理速度慢、无法合理利用现有资源完成任务调度等问题,提出一种基于布谷鸟搜索的加权最小连接数(CS-WLC)算法,并将其应用于智慧矿山软件平台解决负载均衡问题。该算法综合考虑后端服务器处理速率、内存容量、磁盘IO速率、网络吞吐量、进程数指标,通过对指标赋予权值计算各后端服务器利用率;根据计算结果,采用布谷鸟搜索算法对后端服务器进行全局寻优,得到一组较优解;考虑连接数及使用频率对后端服务器赋予权值,采用加权最小连接数(WLC)算法在较优解中选取负载较轻的后端服务器处理实时数据存取和用户访问请求。采用分布式融合性监控系统软件平台进行负载均衡测试,结果表明在数据量不断增多的情况下,与WLC算法相比,CS-WLC算法应答时延小、响应连接数多,从而验证了CS-WLC算法具有更好的负载均衡效果。

关键词:智慧矿山;煤矿监控系统;软件平台;负载均衡;布谷鸟搜索算法;加权最小连接数算法
中图分类号:TD67 **文献标志码:**A

Application of load balancing algorithm in smart mine software platform

GUO Xiucai¹, ZHANG Yue¹, HE Yaoyi^{2,3}

- (1. School of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China; 2. CCTEG Changzhou Research Institute, Changzhou 213015, China;
3. Tiandi(Changzhou) Automation Co., Ltd., Changzhou 213015, China)

Abstract: For problems of current load balancing algorithms such as slow processing speed and unreasonable use of existing resources to complete task scheduling in processing data of smart mine system, a cuckoo search based weighted least connection (CS-WLC) algorithm was proposed and applied to smart mine software platform for load balancing. The algorithm synthetically considers backend server indexes including processing speed, RAM, disc IO speed, network throughput and process quantity, and calculates availability of each backend server through assigning weight to each index. Cuckoo search algorithm is used for global optimization of the backend servers according to the calculation results, so as to get an optimal solution collection. The backend servers are assigned weights considering linking number and usage frequency, and weighted least connection (WLC) algorithm is used to select a backend server with lighter load in the optimal solution collection to process real-time data and user access requests. Load balancing test is carried out on software platform of distributed fusion monitoring system. The results show that CS-WLC algorithm has shorter response time-delay and larger linking number than WLC algorithm under the condition of data increase, which verifies CS-WLC algorithm has better load balancing effect.

收稿日期:2019-11-22;**修回日期:**2020-04-08;**责任编辑:**李明。
基金项目:天地(常州)自动化股份有限公司科研项目(2018GY003)。
作者简介:郭秀才(1963—),男,陕西西安人,教授,硕士,主要研究方向为煤矿井下监控监测系统,E-mail:1036385641@qq.com。
引用格式:郭秀才,张悦,贺耀宜. 负载均衡算法在智慧矿山软件平台中的应用[J]. 工矿自动化,2020,46(5):104-107.

GUO Xiucai,ZHANG Yue,HE Yaoyi. Application of load balancing algorithm in smart mine software platform[J]. Industry and Mine Automation,2020,46(5):104-107.

Key words: smart mine; coal mine monitoring system; software platform; load balancing; cuckoo search algorithm; weighted least connection algorithm

0 引言

近年来,随着智慧矿山建设逐步发展,煤矿安全水平得到了极大提升,但智慧矿山系统中的海量数据处理问题随之而来。据统计,目前每座矿山建有30~50个监控子系统^[1-2],包括安全监控系统、带式输送机控制系统、通风机控制系统、变电所监控系统等,且每个监控子系统有1 000~2 000个测点,导致智慧矿山系统数据量大、种类繁多^[3-5]。目前,智慧矿山均基于云计算架构,软件平台运行在云端,不同矿山之间的数据将产生关联,且子系统、测点、用户数量会越来越多。单一服务器无法支撑当前海量数据处理请求,不能满足实际生产管理需求,因此需要多台服务器同时处理并发的数据存取及用户访问请求,通过负载均衡算法^[6-7]将相关请求合理分配给每台服务器。

负载均衡算法包括静态、动态两种。静态负载均衡算法按照固定比例分配任务,如轮询算法、加权轮询算法;动态负载均衡算法根据当前服务器状态分配任务,如最小连接数(Least Connection,LC)算法、加权最小连接数(Weighted Least Connection,WLC)算法^[8-9]。其中WLC算法是目前最常用的动态负载均衡算法。由于煤矿行业的特殊性和智慧矿山系统数据类型的复杂性,现有负载均衡算法在处理智慧矿山系统数据时,存在处理速度慢、无法合理利用现有资源完成任务调度等问题。针对该问题,本文提出了一种基于布谷鸟搜索(Cuckoo Search,CS)^[10]的WLC(CS-WLC)算法,并将其部署在智慧矿山软件平台,通过试验验证了其具有较好的负载均衡效果。

1 负载均衡算法

1.1 WLC 算法

假设智慧矿山系统中有 n 台服务器, H_{S_i} 为第 i ($i=1,2,\cdots,n$) 台服务器 S_i 的权值, C_{S_i} 为 S_i 当前连接数,则系统中总连接数 $C_{\text{sum}} = \sum_{i=1}^n C_{S_i}$ 。当服务器 S_k ($k=1,2,\cdots,n,k \neq i$) 的连接数 C_{S_k} 满足式(1)时,将新的连接请求发送至 S_k 。

$$\frac{C_{S_k}}{C_{\text{sum}} H_{S_k}} = \min \frac{C_{S_i}}{C_{\text{sum}} H_{S_i}}$$

(1)

式中 H_{S_k} 为服务器 S_k 权值。

WLC 算法中各服务器用相应权值表示其性能,在分配新的连接请求时,尽可能使服务器权值与

该服务器性能呈正比。但 WLC 算法仅使用连接数作为性能指标,任务分配的准确度不高,因此需要进一步引入服务器类别、CPU 利用率等指标进行综合考量^[7]。

1.2 CS-WLC 算法

CS 算法是由布谷鸟寻窝方式演变而来。在自然界中,布谷鸟通过随机方式寻找合适的鸟窝;布谷鸟在飞行过程中发现候选鸟窝进行保留并不断更新,最终找到最优鸟窝。鸟窝位置更新公式为^[10]

$$\mathbf{x}_u(t+1) = \mathbf{x}_u(t) + \alpha \otimes \mathbf{L}$$

(2)

式中: $\mathbf{x}_u(t)$ 为第 u 个鸟窝在第 t 代的位置; α 为步长; $\mathbf{L}$ 为随机搜索路径。

当鸟窝位置更新后,取 $[0,1]$ 内服从均匀分布的随机数 r 与被宿主发现的概率 P 比较,若 $r > P$,则更新 $\mathbf{x}_u(t+1)$,反之保留当前的鸟窝位置。

本文在传统 WLC 算法基础上引入 CS 算法,基本思想:针对智慧矿山系统服务器,先通过 CS 算法选取较优解集,再运行 WLC 算法,在较优解集内选取当前连接数最小的服务器,完成负载匹配任务。

2 基于 CS-WLC 算法的负载均衡

智慧矿山软件平台各业务模块根据数据存取和访问请求的不同分布在多台后端服务器上。当煤矿各类监控子系统向智慧矿山软件平台发送数据时,由于后端服务器长时间运行,负载均衡服务器上记录的负载量不能实时准确地反映各后端服务器负载情况,所以需要周期性采集各后端服务器的负载信息,即每隔一定时间 T ,各后端服务器向负载均衡服务器反馈自身 CPU 利用率、内存利用率、磁盘访问率、网络带宽占用率、进程数量占用率,从而保证负载均衡服务器记录数据的准确性。

智慧矿山软件平台整体部署如图 1 所示。各后端服务器承载数据存取及用户访问业务功能。前端

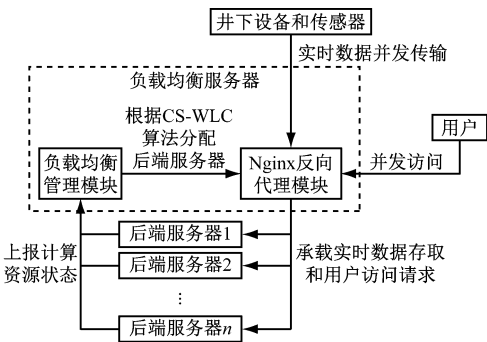


图 1 智慧矿山软件平台整体部署

Fig. 1 Overall deployment of smart mine software platform

负载均衡服务器内置负载均衡管理模块和 Nginx 反向代理模块,前者用于判别后端服务器运行情况,后者根据前者判别结果将实时数据存取和用户访问请求分配给当前负载较轻的后端服务器。

基于 CS-WLC 算法的负载均衡过程如下。

(1) 计算各后端服务器最大处理能力。假设第 i 台后端服务器 S_i 的 CPU 数量为 b_i ,最大处理速率、内存容量、磁盘 IO 速率、网络吞吐量、进程数分别为 E_i, M_i, D_i, N_i, f_i ,其最大处理能力为

$$W_i = K_1 b_i E_i + K_2 M_i + K_3 D_i + K_4 N_i + K_5 f_i \tag{3}$$

式中 K_1-K_5 分别为后端服务器处理速率、内存容量、磁盘 IO 速率、网络吞吐量、进程数对应权值, $\sum_{y=1}^5 K_y = 1$ 。

(2) 计算各后端服务器当前利用率。假设第 i 台后端服务器 S_i 的实时处理速率、内存容量、磁盘 IO 速率、网络吞吐量、进程数分别为 $E'_i, M'_i, D'_i, N'_i, f'_i$,其实时处理能力为

$$W'_i = K_1 b_i E'_i + K_2 M'_i + K_3 D'_i + K_4 N'_i + K_5 f'_i \tag{4}$$

当前后端服务器利用率为

$$G_{S_i} = \frac{W'_i}{W_i} \times 100\% \tag{5}$$

(3) 设置 m 个参考值 $A_1, A_2, \dots, A_m$,并与 G_{S_i} 比较。当 $G_{S_i} \leq A_1$ 时,将 G_{S_i} 放入集合 g_{sel} ;当 $A_1 < G_{S_i} \leq A_2$ 时,将 G_{S_i} 放入集合 g_{se2} 。以此类推,当 $A_{m-1} < G_{S_i} \leq A_m$ 时,将 G_{S_i} 放入集合 g_{sem} 。

(4) 采用 CS 算法对 $g_{sel}, g_{se2}, \dots, g_{sem}$ 进行全局寻优,得到一组较优解,其中每个元素均为各集合中的最优值。

(5) 采用 WLC 算法选取最终后端服务器。在周期 T 内,单一的连接数无法准确反映当前后端服务器的负载接收能力,因此为各后端服务器赋予新的权值,该权值与后端服务器连接数及使用频率呈正比。当新的请求到来时,采用 WLC 算法在较优解中选取当前最优解,并将请求分配给该后端服务器。

假设第 $j(j=1,2,\dots,m)$ 台后端服务器 S_j 权值为 H_{S_j} ,其连接数为 C_{S_j} ,则总的连接数为

$$C'_{sum} = \sum_{j=1}^m C_{S_j} \tag{6}$$

当第 $h(h=1,2,\dots,m, h \neq j)$ 台后端服务器 S_h 的连接数 C_{S_h} 满足式(7)时,将新的请求分配给它。

$$\frac{C_{S_h}}{C'_{sum} H_{S_h}} = \min \frac{C_{S_j}}{C'_{sum} H_{S_j}} \tag{7}$$

式中 H_{S_h} 为后端服务器 S_h 权值。

C'_{sum} 为常数,因此式(7)可简化为

$$\frac{C_{S_h}}{H_{S_h}} = \min \frac{C_{S_j}}{H_{S_j}} \tag{8}$$

3 试验验证

利用某企业“监控类软件平台、框架研究及示范项目”研究成果建立分布式融合性监控系统软件平台,设置 1 台负载均衡服务器、5 台后端服务器,后端服务器运行数据存取及访问服务功能模块。采用煤矿井下实际数据(瓦斯浓度、CO 含量、风速、温度等),对 CS-WLC 算法负载均衡效果进行验证。试验设备见表 1。利用 Nginx 搭建 Web 服务器集群^[11-13],采用 Apache 测试工具 ApacheBench^[14-15]测试软件平台分别采用 WLC 算法和 CS-WLC 算法时的应答时延和响应连接数。设置 CS-WLC 算法中被宿主发现的概率 $P=0.25$,后端服务器处理速率、内存容量、磁盘 IO 速率、网络吞吐量、进程数对应权值 $\{K_1, K_2, K_3, K_4, K_5\} = \{0.3, 0.2, 0.2, 0.2, 0.1\}$,各后端服务器权值为 $\{0.3, 0.2, 0.3, 0.1, 0.1\}$ 。

表 1 试验设备
Table 1 Test devices

设备 编号	CPU 主频/ GHz	内存容 量/GB	系统 类型	设备用途
1	3.2	16	64 位	负载均衡服务器
2	3.2	4	64 位	后端服务器
3	3.2	16	64 位	后端服务器 (含 2 台虚拟机)
4	2.16	8	64 位	后端服务器
5	2.16	8	64 位	后端服务器

试验结果如图 2 所示。可看出在融合性监控系统数据量不断增加的情况下,CS-WLC 算法在数据匹配方面有明显优势:在连接数为 1 000 时,采用 CS-WLC 算法时系统应答时延比采用 WLC 算法时降低了 12.997%,响应连接数提高了 5.61%;在连接数为 1 500 时,采用 CS-WLS 算法时系统应答时延比采用 WLS 算法时降低了 13.829%,响应连接数提高了 9.15%。

4 结语

CS-WLC 算法用于智慧矿山软件平台解决负载均衡问题时,先采用 CS 算法对后端服务器进行全局寻优,得到一组较优解集合,再采用 WLC 算法从该集合中选取负载较轻的后端服务器来处理数据存取及用户访问请求,从而减少了最优后端服务器选取时间,有利于智慧矿山软件平台高效运行。在相同条件下分别将该算法及 WLC 算法运行在分布

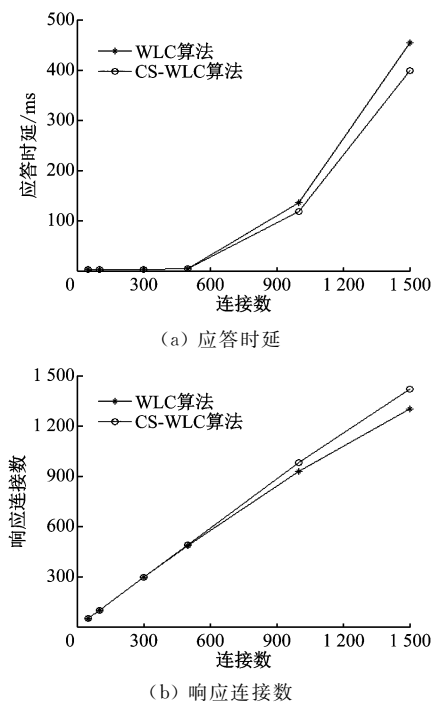


图2 负载均衡试验结果

Fig.2 Load balancing test results

式融合性监控系统软件平台,测试2种算法的应答时延及响应连接数,结果表明CS-WLC算法的应答时延较小、响应连接数较多,从而验证了该算法具有更好的负载均衡效果,可有效提高智慧矿山软件平台的数据处理效率。

参考文献 (References):

[1] 贺耀宜,王海波. 基于物联网的可融合性煤矿监控系统研究[J]. 工矿自动化,2019,45(8):13-18.
HE Yaoyi, WANG Haibo. Research on coal mine fusion monitoring system based on Internet of things [J]. Industry and Mine Automation, 2019, 45 (8): 13-18.

[2] 孙继平,张高敏. 矿井应急通信系统[J]. 工矿自动化,2019,45(8):1-5.
SUN Jiping, ZHANG Gaomin. Mine emergency communication system [J]. Industry and Mine Automation, 2019, 45(8): 1-5.

[3] 裴秋艳. 煤矿海量通风安全数据挖掘算法及智能分析系统研究[D]. 太原:太原理工大学,2017.
PEI Qiuyan. Research on coal mine massive ventilation safety data mining algorithm and intelligent analysis system [D]. Taiyuan: Taiyuan University of Technology, 2017.

[4] 高文. 煤矿监控类系统远程故障分析平台[J]. 工矿自动化,2019,45(2):105-108.
GAO Wen. Remote fault analysis platform of coal mine monitoring system [J]. Industry and Mine Automation, 2019, 45(2): 105-108.

[5] 路娜,郝建军,丁晓明. 矿山透地通信 OFDM 符号定时同步算法[J]. 工矿自动化,2015,41(8):56-59.
LU Yuan, HAO Jianjun, DING Xiaoming. Timing synchronization algorithm of OFDM symbols for mine through-the-earth communication system[J]. Industry and Mine Automation, 2015, 41(8): 56-59.

[6] 王鹏,黄洪琼. 基于神经网络反馈机制的改进型负载均衡算法[J]. 现代计算机,2018(12):3-9.
WANG Peng, HUANG Hongqiong. An improved load balancing algorithm based on neural network feedback mechanism [J]. Modern Computer, 2018 (12): 3-9.

[7] 陈泰安. 一种改进的动态反馈负载均衡算法[J]. 电子设计工程,2013,21(5):171-172.
CHEN Taian. An improved load balancing algorithm based on dynamic feed-back [J]. Electronic Design Engineering, 2013, 21(5): 171-172.

[8] 汪佳文,王书培,徐立波,等. 基于权重轮询负载均衡算法的优化[J]. 计算机系统应用,2018,27(4):138-144.
WANG Jiawen, WANG Shupe, XU Libo, et al. Optimization of load-balancing algorithm based on weight round-robin [J]. Computer Systems & Applications, 2018, 27(4): 138-144.

[9] 高振斌,潘亚辰,华中,等. 改进的基于加权最小连接数的负载均衡算法[J]. 科学技术与工程,2016,16(6):81-85.
GAO Zhenbin, PAN Yachen, HUA Zhong, et al. Improved load balancing algorithm based on weighted least-connections [J]. Science Technology and Engineering, 2016, 16(6): 81-85.

[10] 刘辉,谭延亮. 基于改进的布谷鸟算法在汽车系统电子节点负载均衡中的研究[J]. 科技通报,2015,31(5):179-183.
LIU Hui, TAN Yanliang. Research of electronic node's balanced load in automobile electronic system based on improved cuckoo algorithm[J]. Bulletin of Science and Technology, 2015, 31(5): 179-183.

[11] 汪文君. 基于 Nginx 服务器集群负载均衡方案的研究和改进[J]. 电子世界,2017(2):179-181.
WANG Wenjun. Research and improvement of load balancing scheme based on Nginx server cluster[J]. Electronics World, 2017(2): 179-181.

[12] 李梓杨,于炯,卞琛,等. 基于负载感知的数据流动态负载均衡策略[J]. 计算机应用,2017,37(10):2760-2766.
LI Ziyang, YU Jiong, BIAN Chen, et al. Dynamic data stream load balancing strategy based on load awareness [J]. Journal of Computer Applications, 2017, 37(10): 2760-2766.

[13] 赵梦,李蜀瑜. 云计算环境下基于蚁群优化的任务负载均衡调度算法[J]. 电子设计工程,2016,24(8):

30-33.

ZHAO Meng, LI Shuyu. Load balancing of task scheduling based on ant colony optimization in cloud computing environment [J]. Electronic Design Engineering,2016,24(8): 30-33.

[14] 阮灿华. 基于 SDN 技术的服务器负载均衡设计[J]. 北华大学学报(自然科学版),2018,19(3):405-409.

RUAN Canhua. Design of load balancing of servers based on SDN technology [J]. Journal of Beihua University(Natural Science),2018,19(3):405-409.

[15] 张鹏,李鹏霄,任彦,等. 面向大数据的分布式流处理技术综述[J]. 计算机研究与发展,2014,51(增刊 2): 1-9.

ZHANG Peng, LI Pengxiao, REN Yan, et al. Distributed stream processing and technologies for big data;a review[J]. Journal of Computer Research and Development,2014,51(S2):1-9.