

# 面向多级调度管理的云搜索体系架构研究

孙大雁<sup>1</sup>, 丁杰<sup>2</sup>, 彭林<sup>2</sup>, 施建华<sup>3</sup>

(1.江苏省电力公司,江苏南京 210024;2.中国电力科学研究院信息通信研究所,江苏南京 211106;

3.南京南瑞集团公司信息与通信技术分公司,江苏南京 211106)

**摘要:**提出了一种面向大电网多级调度管理一体化的云计算搜索引擎体系架构设计方法。该方法引入云计算理念,在大电网多级调度管理一体化IT基础设施资源充分整合的基础上,通过调度数据网络将分布于各调度中心的业务数据资源进行有效组织、布局和管理,按照“需则可用”的总体思路,为调度管理等业务应用提供高效的数据共享、交互、检索等各类服务,实现“按需共享”目标,并支持数据的定制化访问和可视化展现。文中着重分析了大电网多级调度管理一体化云搜索引擎的体系架构、技术体系和应用整体设计,为基于成熟、稳定、可靠的开源搜索引擎平台研发系统原型提供理论依据。

**关键词:**智能电网;调度管理;调控一体化;云计算;搜索引擎

**中图分类号:** TM734

**文献标志码:** A

**文章编号:** 1009-0665(2013)06-0005-03

随着特高压交直流电网建设、区域电网互联以及风电等间歇性新能源的接入,电网运行特性愈加复杂,迫切需要对传统的管理模式进行创新,实现管理模式由粗放型向集约化转变,促进各级调度的一体化运行,建立调控一体化技术体系,为国家电网公司的“大运行”体系提供技术支撑。为有效提高生产管理的集约化支撑能力、大电网的协调控制能力以及多级调控间的协同处理能力,需进一步解决电网调控一体化信息检索和共享方式。然而调度的同质化管理要求对数据和信息获取提出全新要求。

近年来互联网上快速发展的搜索引擎技术提供了快速信息检索的有效方法,但这些搜索引擎主要针对网页信息,算法具有一定的局限性<sup>[1-3]</sup>,无法适应国网调度与分控中心一体化运行、与省网调度同质化管理的业务要求,难以支撑调度管理的纵向贯通体系。针对上述问题,本文结合电力行业智能信息处理技术的应用实践<sup>[4,5]</sup>,在现有智能化搜索的研究进展基础上<sup>[6,7]</sup>,以现有调度管理信息检索需求为导向,提出基于云计算支撑技术的搜索引擎体系结构及应用部署模式。该搜索框架关注于高电力信息搜索的查准率、查全率和查找速度,充分利用云计算技术优点<sup>[8,9]</sup>,实现节点分布式管理、资源动态伸缩、服务高可用性功能,为用户提供个性化、智能化的电力资源搜索服务,支撑多级调度机构数据的纵向贯通,为研发性能良好的电力资源搜索引擎系统提供依据。

## 1 搜索引擎问题选型

随着 D5000 的建设,调度信息的规范性越来越高,使调度领域的信息搜索逐渐具备了技术可行性。调度管理应用代表了调度领域对信息搜索的主要需

求,可以基于 OMS 对调度信息搜索做典型分析。调度管理对信息搜索的需求,主要体现在扩大信息搜索范围、多级搜索集成等方面:一是扩大信息搜索范围,目前在 OMS 中已实现了基于文字索引的关键词搜索,支持对调度管理(3区)范围内中的各类文档资料索引提取及搜索功能。云计算环境下的搜索引擎需要扩大搜索范围,横向覆盖 1区、2区、3区的关系库和文件,实现调度中心内部的全面搜索;二是多级搜索集成,需要在关键词搜索的基础上,研究跨调度中心的搜索引擎纵向集成,从而保证搜索的灵活性和全面性。以上需求,在 D5000 调度管理类应用的建设中已有了不同程度的实现,但随着应用功能的深化,目前基于文件索引的信息搜索已不能完全满足需求,迫切需要扩大搜索覆盖面,实现纵向多级调度之间的分布式搜索。

## 2 云搜索体系架构设计

在智能电网调度技术支持系统的已有成果基础上,进行了探索性的研究和分析,结合云计算的理念并充分借鉴了电力行业研究成果<sup>[10,11]</sup>,从 IT 角度构建了调度管理层面的云搜索体系架构,对部署形式及其所提供的按需服务进行了探索性的研究和分析,形成了资源和管理适度集中的部署方案,充分发挥云计算作为调度生产管理重要技术支撑手段的作用。

### 2.1 部署架构

在云搜索整体架构中,搜索引擎分两级部署(国网省一级,地县一级),两级搜索通过远程服务调用进行数据交换。调度机构使用云桌面系统,厂站也通过云桌面实现接入,形成“两级部署,多级应用”的部署架构。如图 1 所示。

在云搜索引擎的部署模式中,全部功能都集中部署在调度中心,通过和数据中心相类似的方式为用户

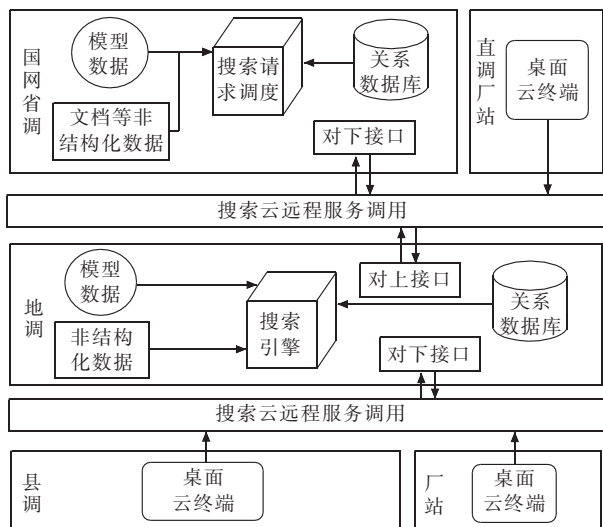


图1 云搜索引擎部署架构

提供服务, 改进了传统的数据中心之间的数据交换方式, 上下级搜索引擎之间通过远程服务调用实现了更快速灵活的集成, 而且只传递搜索结果, 并通过排序减少数据流量, 实现广域规模下的分布式部署架构。因此调度中心的搜索节点部署基本相同, 二级搜索引擎具有向上、向下的服务接口, 一级搜索引擎仅有向下的服务接口。各级调度机构和厂站节点在部署上完全相同, 通过桌面终端, 获得搜索服务并应用。

## 2.2 服务架构

为满足调度领域的搜索需求, 云搜索引擎需要提供基于结构化数据、非结构化以及下级搜索引擎等数据源, 覆盖调度领域各类数据的信息搜索服务。对于搜索服务的使用者, 由于搜索引擎的分布式部署架构屏蔽了上下级搜索引擎之间的关系, 而且国调、省调所部署的搜索引擎在功能上是类似的, 因此讨论搜索服务功能只用基于单搜索节点。搜索服务可分为两大类: 人机服务和公共服务。人机服务是直接向用户提供的搜索服务, 可集成到云搜索引擎的整体门户中; 公共服务是向各级搜索模块提供的公共服务, 可通过标准的SOA 服务接口被其他模块调用。两大类服务的最终目标都是为了支持应用模块, 如图2所示。

## 2.3 技术架构

以云搜索引擎的服务架构和调度管理应用需求为依据, 从搜索技术和云计算技术两个方面进行技术选型, 构建云计算环境下搜索引擎的服务平台, 支持上层业务应用。整体技术架构如图3所示。

电力搜索层通过将 IEC 61970, D5000 等模型中的数据结构固化到搜索引擎框架, 提供面向调度管理(早会报、设备查询、日志查询等)的电力搜索服务。在检索结果反馈上, 搜索云基于 B/S 架构, 采用富客户端技术进行信息展现。在具体的信息展现上, 主要关注于两方面目标: 一是丰富的用户界面; 二是智能的数据模型。

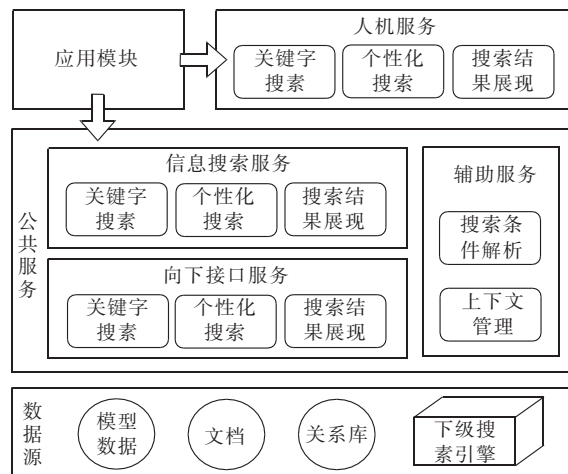


图2 搜索引擎服务架构

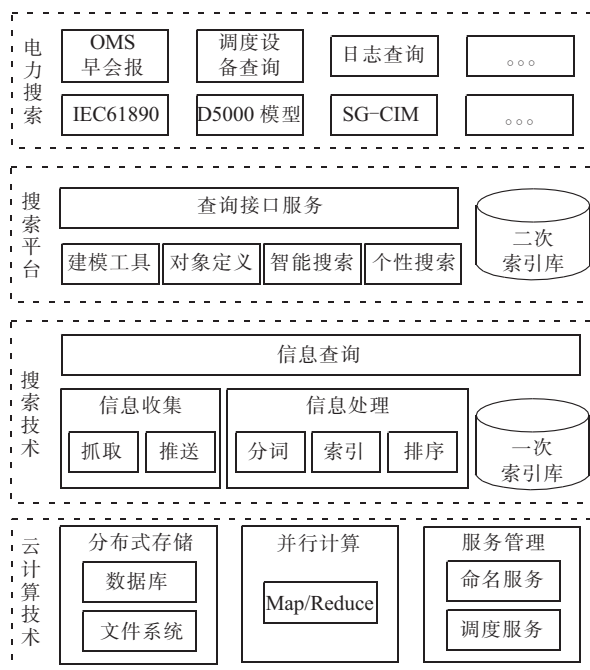


图3 搜索引擎技术架构

对此通过富客户端技术集成桌面应用的交互性和传统网络应用的部署灵活性, 进行信息展现建模和成本分析, 基于 Web 服务驱动实现面向服务的展现层模型。

搜索平台层提供搜索服务的操作接口, 包括信息发布与提交、信息索引和信息查询等。信息的信息发布与提交采用了主动与被动相结合的方式: 一方面通过抓取器轮询文件目录和数据库, 在发现数据变更的情况下对变更数据进行处理; 另一方面针对业务系统提交的信息进行主动索引建立。信息的索引与检索是沿用开源搜索引擎框架, 主要在开源框架的基础上增加中文分词和专属电力行业的索引字段。搜索云底层技术关注于各类数据的索引和存储, 技术选型上主要融合了搜索技术和云计算技术, 主要关注于分布式存储的高可用性和可扩展性, 借鉴 HDFS 的技术架构和实现机理进行文件备份和元数据管理, 并在云存储的基础上构建分布式文件系统来进行索引文件的高效存储。

### 3 搜索引擎应用整体流程

搜索引擎的整体应用流程主要包含4个阶段:索引抽取、索引存储、数据查询、用户访问。信息搜索请求原则上都必须经过4个阶段流转。由于索引抽取是开销最大的阶段,在索引存储阶段引入缓存机制,使绝大多数搜索请求只用经历后3个阶段就可以快速得到结果。基于应用流程,分析构成搜索应用的模块视图,建立对应的应用架构,形成包含4个步骤的主流程、6个主要模块的应用总体设计,如图4所示。

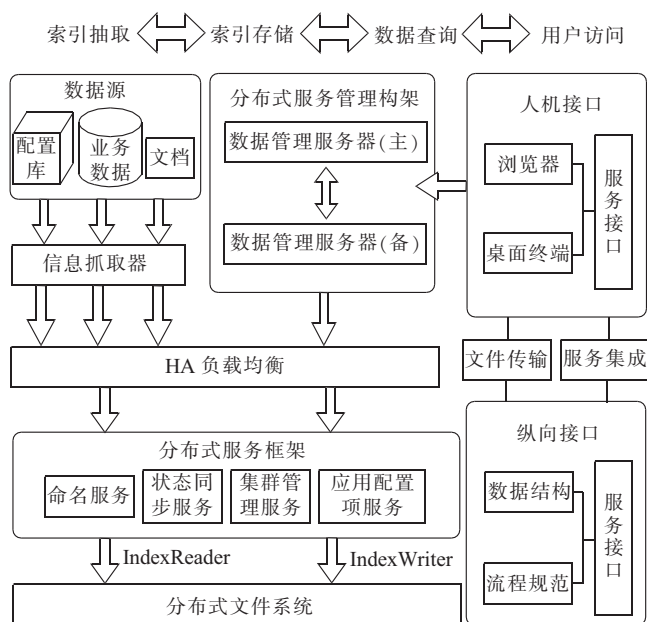


图4 搜索引擎应用流程

### 4 结束语

本文针对调度管理应用对信息搜索的主要需求,探讨了面向多级调度管理的云搜索引擎体系架构。通过借鉴云计算、信息检索等技术的关键特征,研究了面向调度机构的搜索部署架构、服务架构、应用架构,分

析调度信息搜索的应用模式,在扩大信息搜索范围、多级搜索集成、高性能搜索、个性化搜索等方面进行可行性分析和应用设想,更好地支撑调度业务一体化运作。

#### 参考文献:

- [1] 张卫丰,徐宝文. Web 搜索引擎框架研究[J]. 计算机研究与发展, 2005, 11(3):376-378.
- [2] 刘畅. 综合搜索引擎与垂直搜索引擎的比较研究[J]. 情报科学, 2007(1):97-102.
- [3] 印鉴,陈忆群,张钢. 搜索引擎技术研究与发展[J]. 计算机工程, 2005, 31(14):54-56.
- [4] 崔晓祥,李娟. 基于支持向量机回归的电力系统负荷建模[J]. 江苏电机工程, 2012, 31(3):37-42.
- [5] 尹飞,李新家,祝永晋,等. 单亲遗传算法在电力计量器具配送问题中的研究[J]. 江苏电机工程, 2013, 32(1):43-46.
- [6] 赵海波,黄勇理,陈胜,等. 基于电厂运行数据库的专业智能搜索引擎[J]. 电力自动化设备, 2003, 23(8):25-28.
- [7] 袁林. 基于搜索引擎的电力调度知识管理系统[J]. 电力信息化, 2008, 06(7):115-118.
- [8] WEISS A. Computing in the Cloud[J]. ACM Networker, 2007, 11(4):18-25.
- [9] 陈康,郑纬民. 云计算:系统实例与研究现状[J]. 软件学报, 2009, 20(5):1337-1348.
- [10] 赵俊华,文福拴,薛禹胜,等. 云计算:构建未来电力系统的核心计算平台[J]. 电力系统自动化, 2010, 34(15):1-8.
- [11] 王德文,宋亚奇,朱永利. 基于云计算的智能电网信息平台[J]. 电力系统自动化, 2010, 34(22):7-12.

#### 作者简介:

- 孙大雁(1970),男,江苏无锡人,高级工程师,研究方向为电力系统运行及管理;
- 丁杰(1983),男,江苏盐城人,工程师,研究方向为电力系统自动化、智能系统、模式识别;
- 彭林(1976),男,四川巴县人,高级工程师,研究方向为电力系统自动化;
- 施建华(1977),男,安徽宁国人,工程师,研究方向为电力系统自动化。

## Research on Search Engine Architecture for Multi-Level Scheduling Management Based on Cloud Computing

SUN Dayan<sup>1</sup>, DING Jie<sup>2</sup>, PENG Lin<sup>2</sup>, SHI Jianhua<sup>3</sup>

(1. Jiangsu Electric Power Company, Nanjing 210024, China;

2. Information & Communication Department, China Electric Power Research Institute, Nanjing 211106, China;

3. Information & Communication Department, NARI Group Corporation, Nanjing 211106, China)

**Abstract:** This paper presents a method to design the architecture of cloud computing based search engine for the large-scale power network's unified multi-level scheduling management. With the integration of IT infrastructure resources of the large-scale power network's unified multi-level scheduling management system, the data resources from various dispatching centers are effectively organized, dispatched and managed. This makes it possible to provide services of data sharing, interaction and retrieval for power network scheduling and management. Also, it is possible to custom data visiting and to visualize the data. The analysis of integrated cloud search engine architecture for the multi-level scheduling management is particularly stressed to provide a theoretical basis for overall design of search engine prototype system.

**Key words:** smart grid; dispatching management; dispatch control integration; cloud computing; search engine