**数据交换平台**

**需求说明书**

**2015年11月**

目录

**[1](#_Toc434575049)****[项目概述](#_Toc434575049)** [2](#_Toc434575049)

**[1.1](#_Toc434575050)****[项目建设背景](#_Toc434575050)** [2](#_Toc434575050)

**[1.2](#_Toc434575051)****[项目建设目标及总体要求](#_Toc434575051)** [2](#_Toc434575051)

**[1.3](#_Toc434575052)****[项目建设原则](#_Toc434575052)** [3](#_Toc434575052)

**[2](#_Toc434575053)****[技术要求](#_Toc434575053)** [4](#_Toc434575053)

**[2.1](#_Toc434575054)****[技术总体要求](#_Toc434575054)** [4](#_Toc434575054)

**[2.2](#_Toc434575055)****[系统架构要求](#_Toc434575055)** [4](#_Toc434575055)

**[2.3](#_Toc434575056)****[接口规范要求](#_Toc434575056)** [5](#_Toc434575056)

**[2.4](#_Toc434575057)****[系统风险控制及业务持续性计划要求](#_Toc434575057)** [5](#_Toc434575057)

**[2.5](#_Toc434575058)****[系统人机交互要求](#_Toc434575058)** [5](#_Toc434575058)

**[2.6](#_Toc434575059)****[系统安全和访问控制要求](#_Toc434575059)** [5](#_Toc434575059)

**[2.7](#_Toc434575060)****[系统性能要求](#_Toc434575060)** [6](#_Toc434575060)

**[3](#_Toc434575061)****[业务需求](#_Toc434575061)** [6](#_Toc434575061)

**[3.1](#_Toc434575064)****[设计原则](#_Toc434575064)** [6](#_Toc434575064)

**[3.2](#_Toc434575065)****[总体要求及项目实施范围](#_Toc434575065)** [7](#_Toc434575065)

**[3.3](#_Toc434575066)****[行权系统](#_Toc434575066)** [7](#_Toc434575066)

**[3.4](#_Toc434575067)****[电子政务网站群](#_Toc434575067)** [8](#_Toc434575067)

**[3.5](#_Toc434575068)****[宏观经济数据](#_Toc434575068)** [10](#_Toc434575068)

**[3.6](#_Toc434575069)****[IT基础设施核心设备](#_Toc434575069)** [11](#_Toc434575069)

1. **项目概述**
   1. **项目建设背景**

当前，凉山彝族自治州人民政府电子政务建设快速发展，部署在电子政务外网上的各类应用系统也日渐增多，但是还缺乏一个统一的实时数据收集、储存、分析、应用的平台，数据交换平台因此提上建设的日程。随着互联网、云计算、物联网等信息技术的迅猛发展，大量数据的收集、储存、分析、处理及其应用变得更加方便，政府或公众的决策行为将逐渐基于数据和分析而做出，而非像以前基于经验和直觉。

通过数据交换平台的建设，首先改变政府信息系统的数据收集管理模式，将模糊掉各部门系统间的边界，大幅消减信息孤岛现象，使数据共享成为可能，从而提高政府各机构的协同办公效率和为民办事效率，提升政府社会治理能力和公共服务能力。

* 1. **项目建设目标及总体要求**

当前，凉山彝族自治州人民政府电子政务建设快速发展，部署在电子政务外网上的各类应用系统也日渐增多，但是还缺乏一个统一的实时数据收集、储存、分析、应用的平台，数据交换平台因此提上建设的日程。随着互联网、云计算、物联网等信息技术的迅猛发展，大量数据的收集、储存、分析、处理及其应用变得更加方便，政府或公众的决策行为将逐渐基于数据和分析而做出，而非像以前基于经验和直觉。

通过数据交换平台的建设，首先改变政府信息系统的数据收集管理模式，将模糊掉各部门系统间的边界，大幅消减信息孤岛现象，使数据共享成为可能，从而提高政府各机构的协同办公效率和为民办事效率，提升政府社会治理能力和公共服务能力。

通过数据交换平台的建设，收集整合职能系统数据。在无需各应用系统开发厂商提供接口和不改变职能系统数据的情况下，对职能部门所收集的信息进行共享采集和比对，通过分析，可以发现监管漏洞，提高执法水平。

通过数据交换平台的建设，可以加强政务数据的获取、组织、分析、决策，通过云计算技术实现从数据层面对政务信息资源的统一管理，依据法律法规和各部门的需求进行政务资源的开发和利用，可以提高设备资源利用率、避免重复建设、降低维护成本，有效地将数据资源转换为数据资产。

通过数据交换平台的建设，收集整合各类数据，通过展现分析，将进一步提高决策的效率，提高政府决策的科学性和精准性，提高政府预测预警能力以及应急响应能力，节约决策的成本。

当前，凉山彝族自治州人民政府电子政务建设已完成部分建设内容，如基于数据交换平台中间件的行权系统等，通过上述信息化建设工作，已初步形成一个全州各级单位应用的、采用数据交换平台中间件交互数据的信息共享平台，达到了建设的目标，但当前数据交换平台还不能解决数据的规范化、数据的有效性判定、数据的分类采集、数据的交互关联分析、电子政务数据的有效利用、乃至最终数据被用于决策分析等等方方面面的问题，而解决上述问题正是本州电子政务数据交换平台建设的最根本的需求。

电子政务数据交换平台建设需要落在实处，见到实效，能最终解决问题。从现有实际条件出发，有效利用原有的信息化建设成果，即充分利用原有的信息共享平台，在不改变原有的东方通数据交换平台技术架构的基础上，分三步完成电子政务数据交换平台的信息化建设工作：

第一、 为解决面向整体电子政务数据而不仅仅是行权平台的数据采集、数据集成的发展需要，进一步深化原有的数据共享体系架构的功能，通过在原东方通数据交换平台上的数据采集与集成二次开发，使原数据共享平台升级为能满足电子政务数据交换平台建设需要的数据整合平台的能力；

第二、 通过多级部署、集成实施、数据交换平台前置机开发和整合规范数据，使纵向和横向的、符合要求的电子政务数据流动起来，最终汇集到数据交换平台核心数据库；

第三、 通过平台软件的管理及数据分析扩展功能开发，以多种多样的展现方式使汇集的电子政务数据变为对电子政务业务处理、决策分析的有力支撑。

* 1. **项目建设原则**

本项目的建设遵循统一的标准，统一规划来构建电子政务数据交换平台，在总体规划与指导下，在规范上与电子政务相关建设项目保持一致。

电子政务数据交换平台的建设应遵循统一的标准，数据交换平台是基于电子政务系统的应用型平台，虽然各相关职能部门的应用不同、数据不同，但应根据实际制订数据采集标准、数据存储标准、数据利用标准、信息共享标准，实现资源的集约化管理。

电子政务数据交换平台的建设是一项长期而复杂的工作，而且随着业务的快速发展，政府对信息资源利用的需要和要求也将不断提高。所以数据交换平台项目工程的建设，必须坚持可持续发展的原则，应充分考虑后期应用的接入和其他应用的数据利用方式，实现通过可持续的配置方式来扩大数据交换平台的应用范围。。

1. **技术要求**
   1. **技术总体要求**

为了使电子政务数据交换平台既适应当前凉山州的需求，又面向未来发展的需要，系统应采用“一体化建设”原则，即统筹规划，统一建设的原则，保证整个系统顺利连通，并高效、稳定地运行。在进行系统方案的设计中，实现如下要求：

* 高可扩展性；

数据采集点可扩展:数据采集部分能够根据采集数据量的要求，通过添加服务器和配置的方式，无需编程实现水平扩展。

数据存储可扩展：数据存储部分能够根据数据量的要求，通过添加服务器和配置的方式，无需编程实现水平扩展。

分析/处理部分可扩展：数据分析部分能够根据应用的要求，通过添加服务器和配置的方式，无需编程实现水平扩展。.

* 高性能

支持流数据采集：能够实时采集业务日志、流水和各种软件运行状况的数据；业务数据采集速度不低于50笔/秒，每种软件运行状况数据不低于1笔/秒。

细节数据存储：存储每笔业务活动记录（包括查询）。

实时动态展示：业务统计数据实时变化延时不超过1分钟，软件系统统计数据变化延时不超过1分钟。

* 高可用

系统可以通过集群等方式实现分布式部署，避免单点故障。

* 1. **系统架构要求**
* 按照国际标准、国家标准或行业标准，利用先进的软件设计方法论、设计模型和数据模型，进行符合工业标准和金融行业规范的系统开发。
* 系统开发遵循组件化、模块化、参数化设计原则，保持软件系统架构的易于改造和扩展，满足新业务功能的不断扩充，不影响应用系统的各种原有功能。
* 支持应用服务器集群模式，实现服务动态分配、负载均衡。
* 系统设计方案须能充分集成或兼容凉山州信息中心现有网络环境；
* 系统须具有良好的扩展性，采用组件化技术，参数化和插件式设计，伸缩性好，能良好支撑功能扩展；
* 系统须健壮高效，具有快速的响应速度和良好的并发支持能力；
* 系统须具有较好的可维护性，须提交所有源码，须结构清晰，并包括齐全文档
  1. **接口规范要求**
* 不能对现有生产系统造成影响，在现有系统基础上做数据采集，存储。支持多种数据采集方式，如旁路监听、文件、数据库访问。
* 对未来新上系统可以进行方便配置监控，而不需要较大的改造量。
  1. **系统风险控制及业务持续性计划要求**
* 在系统建设全过程加强对各类科技风险的防范，严格按照电子政务安全管控的要求进行系统建设全过程各环节的有效控制。
* 系统应运行稳定，提供连续可靠的服务，具有多级安全保密、权限控管机制。
  1. **系统人机交互要求**
* 系统应具有友好的用户操作和管理维护界面，操作简洁、高效。
* 结构清晰，模块化、参数化程度高，可灵活设置，方便维护和管理。
* 采用纯B/S架构，用户无需安装客户端，通过标准浏览器就能完成对系统的访问。
* 系统支持对不同业务、应用、运行环境（含主机、网络）集中监控和统一视图管理等功能。
  1. **系统安全和访问控制要求**
* 数据交换平台本身能够提供详细有效的系统运行和用户使用日志，便于对故障、事件和错误等进行分析和定位，方便事件处理和解决。
* 提供系统资源使用情况实时监控报告，系统出现异常预警报告等功能。
* 数据交换平台须保证软件本身有足够的健壮性和成熟度。
  1. **系统性能要求**
* 系统整体性能支持凉山州信息中心主要应用纳入信息交换平台进行数据整合并进行监控及数据分析。

1. **业务需求**
3. 1. **设计原则**

* 实用性原则；

数据交换平台体系框架的设计将适应系统运行管理体制和人员的实际情况，满足现有的人工监控的内容要求，所建议的技术方案和提供的产品是成熟的，具备方便的人机界面、易于使用。

* 可用性原则；

数据交换平台的部署不应对原有的系统结构、安全策略等方面做较大修改和调整，对原有系统性能影响最小化，不能对生产系统自身的运行造成不良影响，不能干扰系统的正常运行；尽量少的占用消耗原系统的资源、网络资源，被监控服务器的资源整体消耗不超过5%。

* 健壮性原则；

数据交换平台具有较强的免维护能力，能够长时间稳定运行，自身维护要求简单，具有快速恢复功能。

* 开放性和扩展性原则；

数据交换平台架构设计应遵循业界标准，并提供开放、灵活的信息交互及管理接口，实现灵活的数据交换功能；提供灵活的部署方式，在客户化、管理策略、事件关联、报警方式、报表生成、信息展示、管理流程等方面可以按实际需求进行定制，并支持用户的二次开发；监控管理软件具有较强的扩展性，能够在包括管理范围、管理功能、管理对象数量等方面提供灵活、多样的扩展能力；可以适应生产系统新业务、新技术的要求，适应于监控系统未来发展的需要，并能与其他基于业界标准的软件进行集成；

* 展示形式多样性原则；

数据交换平台能够对收集的数据进行分析处理，生成技术、管理维护等层面的相关报表、视图等，根据不同级别用户的实际需求，提供灵活、多样的展示形式，并提供良好、易操作的自定义功能。

* 1. **总体要求及项目实施范围**

数据交换平台目的是为了满足凉山州信息中心系统所运行维护主要业务系统的数据采集、监控及分析的需求，将分布在不同应用及业务系统中的不同来源及不同格式的数据，通过大数据采集引擎，进行采集，并统一存储至大数据中心库，以供今后对政务网大数据分析提供基础，并根据业务需求，提供实时数据的分析及业务状态监控。本次实施范围主要包括以下几个方面：

* 行权系统的基础数据采集，行权系统的运行状态监控。
* 电子政务网站群数据采集，网站群的运行状态监控。
* 经济信息数据库数据采集及展示。
* 电子政务IT基础设施核心设备运行数据采集及状态监控。
  1. **行权系统**
     + 行权数据聚合

目前，行权系统采用分布式部署方式，由州、县分别管理。其数据分布在17个州、市/县各自不的行权系统中。本期项目需将分布在数据库的数据进行抽取，聚合到大数据平台，给大数据分析处理提供数据基础。由于行权系统未提供对外的数据接口，其系统内数据变化无法实时捕获，因此，本次行权系统的数据的抽取方式将通过读取底层数据库方式获取。为避免对数据抽取对行权数据库性能造成影响，将应用访问低谷时段按一定时间间隔（如每天/多天）取一次数据，将数据同步至大数据交换平台。为了保证行权数据库的安全，由行权数据库创建一个只读权限用户，信息中心定时通过指定的脚本由只读权限用户，将数据从数据库拉出并存放在特定目录的相应的数据文件中（建议格式为CSV），由大数据交换平台读取相关数据文件并加载进大数据平台。需要抽取至大数据交换平台的数据，在与信息中心确定相应的数据库表后，将明确范围补充至项目相关文档中。

* + - 行权数据展示
      * + 按行权大类进行展示。根据行政职权目录划分，对每个大类的行政权力运行结果进行展示。
        + 按地域（州/县）展示不同行政区域行政权力运行结果。结合地图展示行政权力的运行结果。州/市/县的数据在数据表中的位置需信息中心提供说明。
        + 支持按用户角色（部门）进行展示。分角色及行政区域对展示进行权限控制。角色分类：监察局、法制办及平台管理单位。区域：州本级/市（县）。支持用户对行权记录进行查询。
        + 展示报表形式可定制。支持常用的展示形式（直方图，饼图，曲线图等等）。
    - 行权系统的用户登录类型统计

为对用户使用情况进行监控，大数据交换平台将从用户登录日志中提取用户登录行为的相关数据，并进行统计分析，提供报表呈现。行权系统用户登录日志由于未公开接口，详细实施细则将补充至项目相关文档中。

* + - 各地行权平台服务可用性监控

提供各州/县的行权服务状态监控并提供告警服务。大数据交换平台将通过主动探测方式，根据规定的时间间隔，主动发起测试，测试行权平台的可用性。在探测到行权平台不能正常提供服务，在大数据交换平台将提供告警，并将相应的告警事件归档。

* 1. **电子政务网站群**

电子政务网站群由于其门户外接了其它部分非由州信息中心管理的部门及县/市网站，其流量无法通过流量镜像获取，因此，本期网站群的监控范围仅限于我们能获取到流量的网站。网站群数据采用交换机流量镜像方式获取至大数据交换平台。网站群监控包括如下内容：

* + - 用户行为监控；

实时了解用户访问电子政务网站群的行为。捕捉用户请求并跟踪其通过所以响应，收集、处理并显示用户行为的细节数据。

具体实现以下功能：

错误页面跟踪，对返回码为404,500等出错页面进行统计跟踪；

用户终端类型，对用户访问电子政务网站群的终端进行统计；

受访页面统计，用户访问网站所浏览的页面统计；

来路页面，用户通过页面浏览电子政务网站群统计

地区分布，根据用户IP统计访问电子政务网站群的地区分布，并区分内外网用户（内网IP地址范围及相关部门的对照表需信息中心提供）

IP/PV,一天之内独立IP数，相同IP数被计数一次；页面浏览量，用户每次刷新被计算一次。

重复访问率，同一IP，不在同一天内访问同一页面的访问量/总访问量

* + - 电子政务网站群页面监控

关键词搜索频率，用户搜索关键词的频率

热门关键词统计

二级域名访问统计（需信息中心提供二级域名对照表）

频道访问统计（需信息中心提供频道名称对照表）

热点页面统计

根据考核指标，依据上级的要求，检查部分指定栏目的更新频率（仅限于进行于信息中心的网站，检查规则基于文件的更新时间进行确认，需要信息中心提供相应的文件与网站栏目的对应关系及创建相应具有访问相应目录读权限的用户）。

* + - 电子政务网站群运行状态监控

实现对电子政务网站群系统软件运行监视。由于网站群内容提供都是通过静态资源提供。其性能主要取决于Apache web服务器的处理能力及网络传输速度。因此性能及后台服务器监控着重于以下几方面：

动态页面响应性能监控。主要是搜索页面监控。

服务器资源监控（CPU，内存，I/O）等。

网站可用性监控。（需中心提供相应的需监控可用性的网站对应地址及测试可用性的入口）。

告警规则设置：

支持下列业务规则定义的监控策略

1. 单指标阈值规则
2. 单指标统计规则，为某个时间段内的统计指标（如累加、平均等）设定阈值
3. 多指标+统计复合指标，支持与/或/非逻辑表达式定义规则

实时告警功能，包括：

1. 屏幕显示实时报警；
2. 发送电子邮件；（需邮件服务器）
3. 短信通知。（需短信网关支持）
   1. **宏观经济数据**

凉山州宏观经济数据目前主要来源于州统计局定时提供的内部小册子，其数据来源相对固定的，数据格式是规范的，准备展示的内容也是确定的。由于数据内部关联规则复杂，因此，数据将由信息中心相关人员进行预处理，处理后的数据再行导入大数据交换平台。

* + - 宏观经济数据入库

由宏观经济数据库相关的人员将整理相关的经济指标并以excel文件提供，大数据交换平台提供工具自动进行excel文件相关数据导入。 为管理方便，excel文件的名字建议命名规则为Macroeconomic\_indicators\_XX。XX为文件编号。文件编号对应的经济指标需双方进行约定。

* + - 宏观经济数据展示：
* 以地图的方式展示全州及各个市县的主要经济指标（全州地区生产总值、全州非公有制经济增加值、物价等）。为了提供更好的展示效果，建议主要经济指标数不超过5个。（主要经济指标可以更换）。
* 以柱状图、线图、饼图等展示全州各个市县的经济指标的排名、经济指标的发展趋势，以及同比和环比的增长情况。
* 以柱状图、线图、饼图等展示全省各市州的GDP、非公有制经济增加值等的排名情况，提供凉山州在全省范围内的一个发展情况快速参考。
* 对于每个展示页面支持提供一个注释说明框，注释框内容由提供者提交经济指标的时候指定，以便于了解展示内容相关的其它信息。
  1. **IT基础设施核心设备**

IT基础设施包含支撑IT系统的所有设备，所涉及软件、硬件、网络、安全及机房等各个层面。其所涉及到设备数量、型号、版本、接口类型等非常庞杂。其涉及到业务支撑系统，运营支撑及管理支撑几个IT域。本项目并不定位于实现一个完全的IT管控系统，而是做为一个数据整合平台，将几个主要业务的相关的前后台数据，通过大数据的手段，对散布于IT各个层面的数据进行整合、关联并进行运行状况展示，并为今后展开大数据分析提供基础。

按照凉山州信息中心目前设备的管控机制划分，本次监控实施范围共分为如下7个域：

1. 州全网核心网络

核心网络是整个网络骨干，由州本级与17个市、县的核心路由器构成。

1. 州本级核心网络

由州本级核心路由器、城域网汇聚路由器、城域网汇聚交换机及出、入口防火墙构成。

1. 电子政务外网公共服务区

由州本级核心路由器、入口防火墙及虚拟交换机一台构成。

1. 互联网接入区

由州本级核心路由器、F5负载均衡器、主防火墙、备用防火墙及虚拟交换机构成。

1. 交付服务器

物理服务器一台

1. 核心业务服务器

运行核心业务的物理服务器一台（VCenter）及核心业务VM（行权系统，门户网站，数据库等）

1. 公用存储设备

存储阵列

以下为各个域的监控需求：

* + - 核心网络状态监控。

核心网络涵盖了数据中心与省、市/县及对外提供互联网访问所有网络流量。对核心网络状态的监控，可以帮助数据中心及时掌握全网的总体运行情况。

核心网络包括州两台核心路由器（华为NE40E-8，IP：59.213.208.53，

59.213.208.57）及17市/县的核心交换机（华为的不同型号，共17台）。设备状态通过SNMP协议采集。状态包括如下几个方面：接口名称，接口状态，接口收到的字节数，接口发送的字节数，接口收到的错误包，接口发送的错误包。

监控要求：通过网络拓扑逻辑示意图，结合实时采集的数据，动态展示网络运行状态，并对网络异常提供及时告警。

* + - 州本级核心网络状态监控。

州本级核心网络设备涵盖了州数据中心负责运维的核心设备。对州本级核心网络状态的监控，可以帮助数据中心及时掌握数据中心运维的网络主干的总体运行情况。

州本级核心网络包括州两台核心路由器（华为NE40E-8），3台城域网汇聚路由器（华为ma5200g，IP：10.21.2.198，10.21.0.45，10.21.0.73）及2台汇聚交换机（华为S8508一台，华为S7802一台）。设备状态通过SNMP协议采集。状态包括如下几个方面：接口名称，接口状态，接口收到的字节数，接口发送的字节数，接口收到的错误包，接口发送的错误包。

监控要求：通过网络拓扑逻辑示意图，结合实时采集的数据，动态展示网络运行状态，并对网络异常提供及时告警。

* + - 电子政务外网公共服务区状态监控

电子政务外网公共服务区核心设备包括州两台核心路由器（华为NE40E-8），入口防火墙2台（华为USG5530，IP: 59.213.208.78,59.213.208.121）及虚拟交换机一台（H3C s12510，IP：10.21.1.1）构成。设备状态通过SNMP协议采集。状态包括如下几个方面：接口名称，接口状态，接口收到的字节数，接口发送的字节数，接口收到的错误包，接口发送的错误包。防火墙需采集syslog中安全告警信息。

监控要求：结合实时采集的数据，动态展示网络设备的运行状态，并提供及时网络异常告警。针对syslog中采集的安全告警及时发出告警提示。

* + - 互联网接入区状态监控

互联网接入区核心设备包括州两台核心路由器（华为NE40E-8），入口防火墙2台（华为USG5530，IP: 59.213.208.74，59.213.208.105），出口防火墙1台（华为USG5530,IP:59.213.208.54），备用防火墙（网御星云 UTM2000，IP：59.213.208.58），链路负载均衡器（F5 BIG-IP-4000S，IP： 59.213.209.1）及虚拟交换机一台（H3C s12510，IP：192.168.2.10）构成。设备状态通过SNMP协议采集。状态包括如下几个方面：接口名称，接口状态，接口收到的字节数，接口发送的字节数，接口收到的错误包，接口发送的错误包。防火墙需采集syslog中安全告警信息。

监控要求：结合实时采集的数据，动态展示网络设备的运行状态，并提供及时网络异常告警。针对syslog中采集的安全告警及时发出告警提示。

* + - 交付服务器物理主机运行状态监控

本期将对政务云平台交付主机服务器（HyperV，IP：10.21.12.200）进行监控，监控信息将通过SNMP获取。主要包括CPU、内存、网卡、HBA卡、电源、风扇、主板运行状况。

* + - 关键业务运行状态监控

本期将选择政务云平台主机服务器（Vmvare vCenter，IP：10.21.21.200）进行监控，监控信息将通过SNMP获取。主要包括CPU、内存、网卡、HBA卡、电源、风扇、主板运行状况。

由于物理主机上运行的VM众多，本期将选择部分重要服务器进行监控，如网站群服务器（IP：192.168.12.24，192.168.12.25，192.168.12.26），行权系统部分服务器（IP：59.213.213.68，59.213.213.69，59.213.213.70）及数据库服务器（IP：192.168.12.28），监控指标：CPU、内存、I/O等。

* + - 提供主存储设备的状态监控

存储设备监控通过捕获syslog中告警的方式，监控主存储设备阵列中磁盘开销情况、坏盘告警等。存储设备包括两台S5600T及存储虚拟化网关VIS6600。本期至少完成对其中一台存储设备的监控以提供后续扩展样例。