

信雅达

内容管理平台 (SUNECM)

产品白皮书

Sunyard Enterprise Content Management (SUNECM) White Paper

1 目录

第 1 章	概述	1
1.1	序言	1
1.2	内容简介	1
第 2 章	需求分析	2
2.1	项目背景	2
2.2	内容管理平台功能需求描述	3
2.3	平台在银行 IT 架构中的位置	6
第 3 章	SunECM 平台架构	8
3.1	架构设计原则	8
3.2	SUNECM 平台与其他系统的相互关系	9
3.2.1	基于传统技术的应用系统	9
3.2.2	基于 J2EE 的应用系统	9
3.2.3	基于 Web Services 系统	9
3.2.4	基于开源组件	9
3.3	系统架构概览	10
第 4 章	SunECM 关键技术分析	13
4.1	面向服务的体系架构	13
4.1.1	什么是面向服务的体系架构	13
4.1.2	原有系统的集成	13
4.2	SUNECM 技术特点	14
4.2.1	SUNECM 物理架构	14
4.2.2	业务系统接入架构设计	15
4.2.3	安全与权限设计	18
4.2.4	SUNECM 提供 SOA 的内容服务	19
4.2.5	SUNECM 关键技术	20
4.3	SUNECM 功能概述	22

4.3.1	软件模块图.....	22
4.3.2	软件模块划分.....	23
4.3.3	SUNECM 主要功能描述.....	25
4.4	业务应用接入过程.....	47

第1章 概述

1.1 序言

SUNECM (Sunyard Enterprise Content Management 信雅达企业内容管理平台)是信雅达公司根据其多年的金融领域 IT 架构、内容管理软件开发及实施应用的成功经验,总结出企业内容管理的共性,针对企业对内容管理的项目建设目标,考虑到企业现有的业务关系、系统环境以及未来的扩展需要,以面向服务的架构(SOA)思想为指导,重点推出的企业内容管理软件产品,为管理、共享、重用和检索各种类型的数字内容提供了一个单独的、开放的和综合性的平台。

SUNECM是企业内容管理解决方案的核心,统一管理各类非结构化电子数据。电子数据的存储、查询等都以服务的形式提供给各业务应用系统调用。

本文分别从内容管理服务、系统整合等几个方面进行了说明。包括关键技术,推荐使用的软件产品,同时也提出了实施建议。

1.2 内容简介

本建议书主要包括:

第 2 章简述了企业内容管理平台的需求。

第 3 章分析了与企业内容管理平台相关的关键技术。

第 4 章描述了 SUNECM 企业内容管理平台的详细功能包括与其他系统的关系, 系统架构概览以及系统物理运行架构等。

第2章 需求分析

2.1 项目背景

在当今信息爆炸的时代，信息每天都在以惊人的速度增长。有统计表明，来自交易中的数据每年增长的速度达到 61%，而其他各种相关信息的每年增长率甚至超过 92%。同时，信息的种类也在不断地扩展，越来越多的非结构化信息不断出现，包括企业的各种单证、报表、账单、电子文档、网站的各种元素、图片、传真，以及大量的多媒体的音频、视频信息等等。而信息爆炸的背后，我们面临的巨大挑战是如何有序地存储、管理并利用这些信息，尤其是对于那些大量的、非结构化的信息。如果一个企业无法将这些信息进行有系统地编目和管理，则随之面临的危机是不断流失的生产力。同时，企业和组织也希望对这些信息有更多的控制，保障信息的安全，挖掘信息的价值。

银行中大量涉及单证、凭据、支票、证明、发票、客户外来信息等各种信息，这些信息都是非结构化的大对象，难以用传统的关系型数据库管理，将这些非结构化的信息称为内容（Content），银行中大量的内容信息都是电子影像。

内容按照检索视图可以划分成结构化数据、非结构化数据，结构化数据包括关系型数据库数据，结构化数据文件；非结构化数据包括数字化的纸质文件、照片、传真件、信件、报表，音频、视频文件，电子邮件，网页，软件产品、应用程序等。

对于结构化数据，由于其基于 2 维表或 2 维文件，因此检索比较方便；但是对于非结构化数据，由于其数据是一个对象，如果没有对其进行结构化处理，基于目前的智能情况，无法对其进行快速分类、索引、排序、检索，因此对非结构化数据的管理依赖于对数据的结构化。

目前大部分银行已完成数据集中，该部分数据是结构化数据，对于非结构化数据，比如纸质文件，大部分没有数字化处理，业务处理流程尚处于纸质文件流转，成本较高，风险也无法控制，即使部分纸质文件数字化处理后由于没有一个完整的内容管理规划，文件无法实现共享，没有充分利用数据。

随着信息化管理的不断深入，银行中和大量业务相关的内容信息也急需通过有效的方式进行管理和访问。根据银行的业务系统分析可以查看到目前和内容管理密切相关的业务系统包括：

1. 会计档案系统——大量的报表、凭证影像需要进行有效的管理。
2. 国际结算业务——信用证、托收、保函、福费廷、押汇等国际结算业务系统中涉及的凭证影像。
3. 前台后移业务——包括支付、对公授信类业务中涉及的各种凭证的电子影像等。

支票结算业务——包括支票相关的票据影像等。

4. 事后监督——包括支票相关的票据影像等。
5. 信贷影像——包括多种类影像资料

2.2 内容管理平台功能需求描述

企业内容管理就是以信息共享为目的，面向海量信息处理，信息数字化、结构化处理，分布存储，具有四大功能：内容集成、内容管理、内容智能和内容发布。主要功效和价值主要表现在两方面：对外，帮助企业建立统一的对外协作平台，加强企业与业务合作伙伴、客户的沟通和联系，为客户提供更加个性化的服务；对内，帮助企业打破信息孤岛，理顺业务流程，加强团队之间的协作，使企业拥有的“核心知识”得到充分的利用。根据目前及未来企业对内容管理平台业务的需求。

信息化建设是国民经济信息化建设的重要组成部分。金融行业作为国民经济的支柱产业更是不断地在加快信息化建设的步伐。随着技术进步的加快，互联网的出现和迅速发展，一个全球性的信息社会正在逐步形成，推进工作的自动化、网络化、电子化，已是大势所趋。

随着金融业的发展，许多新的业务也在不断地出现。对于这些新的业务，都会面临着大量的纸张、电子文档等内容信息；在金融行业的竞争中，如果希望成为覆盖面最广的金融服务机构，成为最优秀的银行之一，则需要建立一套新的业务系统和流程。在此系统中会遇到大量的原来基于纸面的客户资料——例如信用卡申请、客户证明材料、传真等等。为了提高业务效率，同时能够提高客户服务水平，需要建立内容管理平台进行此类信息的存储和查询。分析内容管理平台的需求，总结如下项目目标：

1. 简化作业，减少人工分派及递送；
2. 加速文件传递时效，加快作业处理；
3. 减少纸质文档分派过程损害或遗失；
4. 集中文件管理、可同时供多人多部门使用；
5. 便于文件存取、可迅速调阅、增进服务品质；
6. 保存文件方便、节省纸质文档存储管理成本；

而作为内容管理平台，则需要满足以下功能：

● 内容引擎（Content Engine）

包括文档发布和获取校验、版本控制、事件驱动(Event Driven) 架构、安全性校验以及对文件提供存储检索服务；提供文档的使用权限管理、分类、对象间的关系管理支持内容信息的属性检索和全文检索，并能够实现多维检索、全文检索、智能检索；支持海量存储，分类存储，分布式存储。

◆ 文档管理

- 各类文本文件，包括OFFICE、TXT、PDF、HTML文档等；

- 静态影像文件，包括JPG、BMP、TIFF等格式；
- 音频、视频文件，包括WAV、MEPG3、MEGP4等格式；
- 邮件；
- 其他类型文件，包括二进制文件等；
- 版本管理（根据业务需要可采用版本管理，亦可不采版本管理）
- 文件的权限管理（根据业务需要对于不同类型的文件可设置不同的访问权限）
- 文档服务器支持分布式部署，集中式部署
- 支持文档服务器组负载均衡，有效的均衡服务器组内各服务器的负载

● 元数据管理

- ◆ 可以根据自身业务需求进行灵活的元数据定义
- ◆ 支持不同的业务应用采用各自的元数据模型
- ◆ 能够进行增、删、改、查等各种管理操作
- ◆ 根据不同的操作权限来操作元数据管理
- ◆ 支持海量数据的存储
- ◆ 支持业务数据按日/月/年等模式的目录管理
- ◆ 灵活的内容组织形式。可通过索引及文档方式来灵活组织内容，以便于内部用户和外部用户快速组织及使用内容信息。
- ◆ 支持索引下单/复文档结构的元数据模型

● 生命周期管理

提供完善的存储、备份方案及对在线、离线数据的管理策略。记录信息的产生、接收、维护、使用直到最后的处理；

- ◆ 提供内容清晰的创建、迁移途径、演变历史和销毁功能
- ◆ 可以根据数字化信息的使用频率、种类、优先级、及保存年限、以及存储介质的读写速度等因素制定相应的存储策略
- ◆ 查看内容时，从相应的存储介质提取信息

● 数据的一致性和统一性

能够将元数据、大对象数据存储到后端不同系统中，同时保证这些数据存储的一致性及前端调用的统一性

● 灵活方便的用户权限管理机制

具备数据数字权限管理扩展的能力，支持用户/角色级别的用户权限定义

● 对业务的活动内容支撑能力

基于活动内容的思想，利用活动内容基础结构来设计内容引擎和业务流程引擎功能。由内容引擎管理内容的创建或捕获可以自动触发事件，启动或修改

业务流程，或者与其他企业应用程序交互。支持商业流程和内容传递的工作流，配置工作任务和状态，并创建查找索引。

● 灵活的系统扩展能力

- ◆ 提供多层次、分布式的架构能力可支持的企业级的扩展能力
- ◆ 提供垂直和水平的扩展解决方案
- ◆ 支持负载均衡功能
- ◆ 支持不同的操作系统，主流商业数据库支持，以及主流应用服务器

● 提供灵活的业务扩展能力

- ◆ 一次规划、分布实施，建立全行的内容管理平台，方便地进行扩展，以满足未来业务种类、业务量、及业务产生的和类数据内容增加；
- ◆ 能够方便地进行扩展，以满足未来业务种类、业务量、及业务产生的和类数据内容增加
- ◆ 能够支持 B/S 架构及 C/S 架构
- ◆ 能够支持基于 VC、JAVA 语言的业务系统，甚至是未来的 Webservice 系统接入

● 统一的内容接入接口

- ◆ 基于面向服务的方式，向业务接入系统提供统一的访问接口
- ◆ 提供丰富的接口，能够满足不同业务系统的需要，包括各种灵活的组件、模块、控件、API 等输入输出接口
- ◆ 能够保证业务接入的安全性

支持对所有内容管理数据的增、删、改、查等操作，能够支持对不同应用的活动数据和静态历史数据的统一处理；提供各种灵活的组件、模块、控件、API 等各种输入输出接口方案，方便各业务系统调用、载入、展现和利用影像高效地开展业务。支持服务器接口方式和客户端接口方式的相互融合，业务应用在进行接入时可以根据自身要求进行控制。

● 系统监控

- ◆ 提供平台监控功能，包括平台软件及硬件性能参数；
- ◆ 监控除提供桌面监控界面外还提供各类图形形式监控界面；
- ◆ 结合存储管理软件，系统管理模块应能监控存储设备的利用情况；
- ◆ 对于系统告警，模块提供多种手段通知管理员；

由于企业内容管理平台采用混合架构，所以系统在物理上是分散部署，但从逻辑来说是一个整体，因此监控应采取单点监控（可以包含备份监控点），各实际部署点除对硬件设施进行巡检外对平台无需额外监管；

● 存储管理

银行具有大量的内容信息资料。因此内容管理平台必须具备存储大量信息的能力。支持二级存储设备，能够实现不同介质之间的定期转存，以及对介质上的信息

访问定位能力。系统对数据的存储操作简便、可靠，针对不同优先级数据具有不同的存储策略。实现方便灵活的数据存储管理能力。（包括：二级存储管理的能力、迁移、清洗能力等

◆ 存储模式

〈1〉 集中式存储： 集中式存储是存储设备物理储存在单一地点；

〈2〉 分布式存储： 分布式存储是存储设备物理储存在多个地点；

◆ 存储设备

〈1〉 在线存储设备：

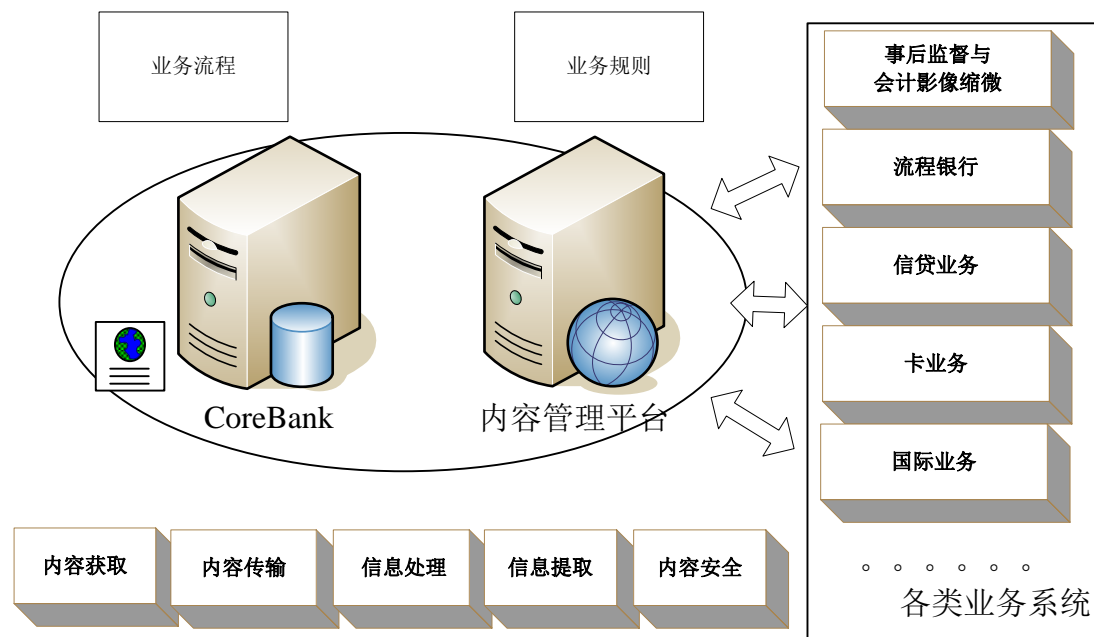
在线（On-Line）存储，又称联机存储，是指设备以永久方式连接在计算机系统中，可以直接访问这些设备。最具代表性的在线存储设备是硬盘，任何时刻均可以快速的存取数据。在线设备具有最高的响应速度，获取保存在硬盘上的数据通常在数秒的水平。在线设备一般存储热数据，适合频繁访问、数据量较小的情况。

〈2〉 离线存储设备：

离线（Off-Line）存储，又称脱机存储，是指设备或介质平时没有装载在计算机系统中，在存取数据时需将设备或介质临时性地装载或连接到计算机系统，当数据访问完成时可以脱开连接。一旦断开连接之后，就可以更换介质，如磁带、软盘或光盘等。离线方式通常在访问速度上要低一些，但是因其可以随时更换存储介质，所以其存储容量上没有实际限制。离线设备因其可以随意装入或卸下存储介质，所以是数据交换的最传统的方式。离线设备具有存储容量无限的特点，通常介质的价格较为低廉，是海量存储的必然选择。但是，因为手工或机械查找和装载介质，其响应速度较低，甚至更长。离线设备一般用于存储冷数据（即不常使用的数据），并且数据量极大的情况。

2.3 平台在银行 IT 架构中的位置

银行内容管理平台构建了银行非结构化数据的存储管理平台，对于银行大量的非结构化数据进行存储管理。银行的结构化数据，如帐务数据、实时流水数据等，存放在核心数据库中，然而结构化数据仅占银行数据内容的 20%，对于大量如票据影像、报表、历史流水、多媒体等非结构化数据，数据库就无能为力了。平台构建的内容存储管理平台与核心数据库共同构建了银行整体数据内容的存储中心。



内容管理平台在银行 IT 架构中的位置

通过存储中心的应用和部署，各种业务系统在进行业务处理时，能够方便地对各类信息数据进行获取、处理、存储和管理操作。业务办理过程中所需要的结构化数据信息内容和信息处理可以在 CoreBank 中实现，非结构化数据的信息检索和信息处理可以在内容管理平台中加以实现，从而实现了银行信息的整合处理。

非结构化数据的容量大大超过结构化数据，差距在一个到两个数量级，因此 IT 架构体系需要重新考虑，如网络带宽、存储空间等。

在业务处理过程中接收或形成的票据通过影像采集形成非结构化数据，经过影像处理和信息提取后，这些数据利用影像安全系统进行加密和防伪处理，通过影像传输系统传输到内容管理平台上，业务办理的后续处理，可以利用这些非结构化数据进行办理，从而形成无纸化的办公和银行业务处理。

第3章

SunECM 平台架构

SUNECM 架构以 SOA 的思想方法为指导, 结合信雅达多年来在金融领域 IT 架构、内容管理软件开发及实施应用以及业务整合的成功经验, 主要着眼点和目标是整个内容管理平台架构的灵活性和适应能力。使得该架构能够适应各合作伙伴在政策法规、组织、业务功能、处理流程、IT 技术以及外系统接口等多方面的变化。同时, 该架构能够支持海量的处理要求。

3.1 架构设计原则

我们的方案将基于以下设计原则:

- 松耦合 (Loosely-Coupled) 整合方式: 系统设计将应用程序定义为不同组件 (或称为服务), 通过这些服务之间定义良好的接口和契约联系起来。接口是采用中立的方式进行定义的, 它应该独立于实现服务的硬件平台、操作系统和编程语言。这使得构建在各种这样的系统中的服务可以以松耦合的整合方式, 并采用一种统一和通用的方法进行交互。

- 适应性 (Flexibility): 由于需要整合的系统相当多并且复杂, 系统设计必须能够方便地适应当前相关系统的不同情况以及未来变化。包括支撑技术、系统接口以及业务需求等方面的变化。同时也能通过流程描述的方式适应面临的改变。本系统应尽可能减少对原系统的改变。

- 扩充性 (Scalability): 能够通过增加系统的资源, 如 CPU、内存、网络、和存储等显著地提升系统的吞吐能力。包括垂直扩展 (在一台服务器中增加更多的资源) 和水平扩展 (使用多台服务器)。

- 可用性 (Availability): 系统通过冗余的方法避免单点故障。同时, 系统应尽量减少计划内的停机。

- 安全性 (Security): 系统的安全性涉及多个方面。在这里我们主要关注安全管理。包括认证 (Authentication)、授权 (Authorization)、审计 (Audit) 和管理 (Administration) 等几个方面。

- 成熟性: 使用的产品都是经过了市场的考验, 并且在全球范围内有广泛的用户。尽量避免采用一些小的厂商开发的、或者自己开发的中间件产品。

- 先进性: 设计方案中采用市场领先并成熟的技术, 使项目具备国内同业领先的地位。便于系统的升级和今后的维护。

- 标准性和开放性: 在本项目建议书中建议的产品, 从网络协议到操作系统, 全部遵循通用的国际或行业标准。系统整体架构

- 充分利用现有资源

- 统筹考虑，长远规划

3.2 SUNECM 平台与其他系统的相互关系

SUNECM 提供内容信息供其他业务系统使用，SUNECM 和其他业务系统的关联为松耦合联结关系，按照 SOA 的设计思想实现 SUNECM 和其他业务系统的内容交换。

基于接口方式的不同，与 SUNECM 相关的系统可以分为以下几类：

3.2.1 基于传统技术的应用系统

基于传统技术(如 C/C++等)的系统，如国际结算单证中心、会计档案、重要客户、信用卡等，将使用 VC 语言作为一个通用的客户端通过 HTTP/SOCKET 与 SUNECM 实现整合。

3.2.2 基于 J2EE 的应用系统

基于 J2EE 的系统，如网上银行、客服中心、个贷等，将使用 JAVA 客户端通过 SOCKET/HTTP 与 SUNECM 实现整合。

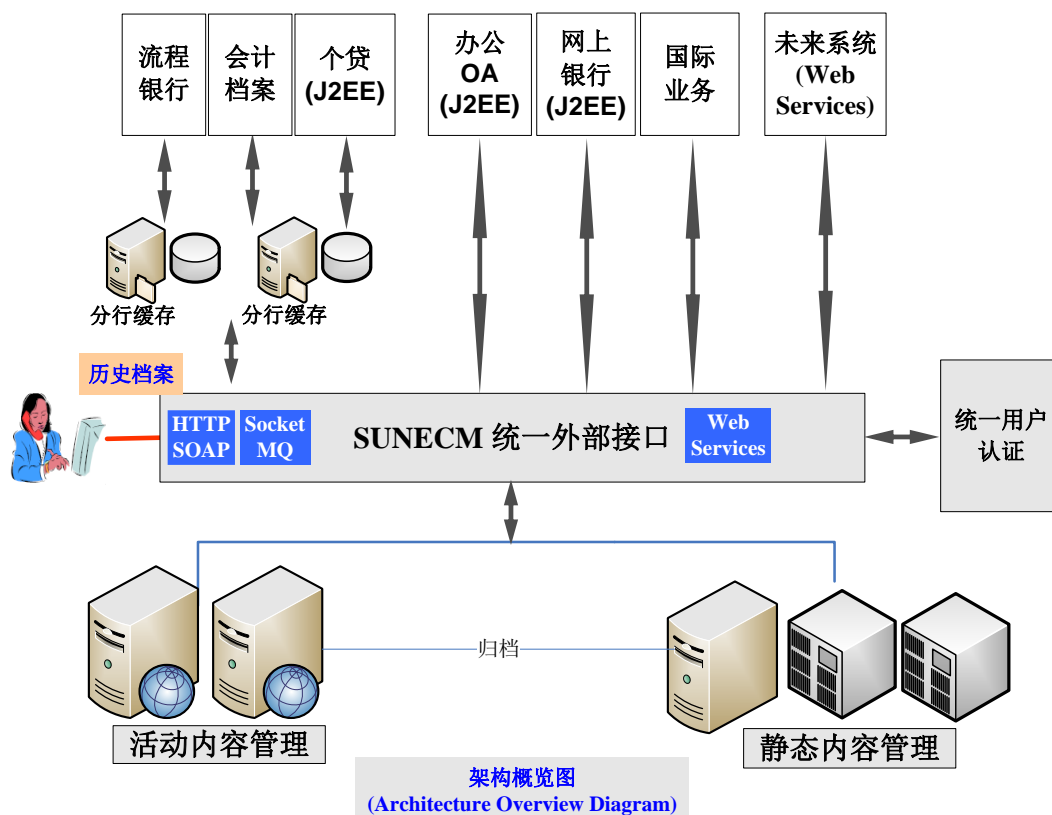
3.2.3 基于 Web Services 系统

基于 Web Services 的未来系统，将使用 SOAP/HTTP(S) 与 SUNECM 实现整合。

3.2.4 基于开源组件

SunECM 的开发使用了 spring、structs2、c3p0、extjs2、log4j、sigar 等开源组件，未使用其他商业组件。

3.3 系统架构概览



SUNECM 与业务系统集成的整体架构图

根据SOA的松耦合设计原则，设计中的企业内容管理平台与企业的各应用系统之间的关系是相互提供服务的关系，平台向各应用系统提供内容检索，内容提取等服务，同时也接受各应用系统向平台输入内容，以及对各应用系统功能的调用。

- 既有总行集中的模式，比如数据仓库；也有总分行分布式的模式，比如个贷。
- 文档和影像文件都能够通过B/S方式获得。
- 现有业务系统采用外部调用方式访问内容管理平台，以减少现有系统的改动量。

内容根据业务需求可以概述为二个部分：

▪ 静态内容

内容在各个业务环节不会被修改，作为原始凭证、内容提供给业务人员进行使用，内容具有只能读、不能修改的特性。银行行业大部分的内容都属于静态内容，如票据影像、单证影像、报表、邮件归档等等。这类内容具有信息量大、保存时间长、访问频率高的特点。

■ 活动内容

处于这种状态的内容数据其内容随着业务的需要，需要进行更改、版本管理等。例如办公应用系统中涉及的办公文档、网站内容中涉及的介绍文字、不同时期内容信息的不同版本等。活动内容在银行行业相对静态内容来讲比较少。

目前银行正在重新组合经营，借助现代信息技术，重新设计银行的管理模式和业务流程，使银行集中核心力量，获得可持续竞争的优势。但是各种业务的不同，集中的程度也不同。中国地大物博，各地区差异较大，某些业务的判断标准也不同（如信贷），考虑到业务的地方特色，某些业务运作方式不能采取完全集中的方式；同时也因为中国地理位置广阔，集中的程度和业务量以及网络的相关度很高，同样也会制约业务集中的程度。

内容管理系统需要管理全行的海量内容数据，目前内容管理系统主要应用于对公会计档案和流程银行的业务，这些业务分布在分行/分中心的缓存上进行业务处理，在业务完成后，上传到总行内容管理中。内容管理系统涉及到多个环节，主要有：采集（如扫描）、处理、安全传输、装载、存储管理、获取各个环节，业务模式也会有多个环节来运行包括：业务处理中心和网点或其他分中心；网络架构包括网点—>分行—>总行，内部的局域网以及多个环节。内容管理系统中的内容通常是传统的关系型数据库信息的几十上百倍，因此其部署方式要充分考虑业务模式和网络环境。

SUNECM 统一外部接口以面向服务为基础采用多种方式与影像应用系统和业务相关系统进行整合，通过标准的协议（如 HTTP、Socket、SOAP、Web Service 等）或者定制的接口实现应用系统与 SUNECM 的集成，通过 Socket 适应目前分行缓存的结构；通过统一的用户认证实现各系统用户账户的同步以支持应用系统的统一认证以及权限认证控制要求。

平台解决的主要需求和功能是：

- 1、银行各类业务在使用以影像及各种文档为基础的内容时，可以通过统一的内容管理平台实施接入；内容管理平台需要根据业务系统的要求提供以下接入方式：
 - 服务端接入方式
 - 服务端+客户端集中接入方式
 - 服务端+客户端分布接入方式
- 2、实现各种业务系统内容的元数据管理，从前端展现到各级缓存都要支持各种业务元数据要求，实现对元数据的各种管理功能；
- 3、内容生命周期管理：涵盖缓存——SUNECM 分支文档服务器——SUNECM 中心文档服务器——离线归档——销毁，能够根据业务应用区管理策略
- 4、支持各种技术接口标准，适应业务接入，包括 Socket、Http、SOAP 等
- 5、支持海量内容存储：

在线方式通过批次内打包、内容压缩，压缩影像增加单个 SUNECM 的存储量

离线方式通过支持业务应用的生命周期管理策略，进行离线归档，使 CM 中的存量数据得到控制

6、缓存的管理和监控

7、实现完整的管理接口功能，包括增、删、改、查等

考虑内容管理系统架构的时候一种处理方式是不考虑外部环境约束，为管理维护方便，部署为总集中方式。网络带宽、业务模式适应内容管理系统，网络带宽不足时扩充网络带宽；但需要投入大量的网络成本。

充分考虑银行内容管理系统的需求、业务模式和网络状况，推荐根据业务过程及模式进行内容管理系统部署（集中或分散）的体系架构。本体系架构充分分析业务各个环节，分析网络流量及业务要求，可以将内容管理系统部署在业务频繁发生内容读取的位置，利用局域网弥补网络带宽的不足。

第4章

SunECM 关键技术分析

SUNECM 需要成为一个统一内容管理平台,支持总行级和总/分行级应用系统的不同要求。这要求它必须具有丰富的功能、足够的灵活性和扩展能力,来适应未来组织架构、业务流程、接口规范以及单个系统实现技术的改变。基于 SOA 思想的架构是其理想的选择。下面我们对 SOA 的思想,以及 SUNECM 平台相关的关键技术进行简要分析。

4.1 面向服务的体系架构

目前,大多数企业都有各种各样的系统、应用程序以及不同时期和技术的体系结构。集成来自多个厂商跨不同平台的产品和应用系统,一直是企业 IT 部门的主要挑战。面向服务的体系结构(Service Oriented Architecture, SOA)为解决这一问题提供了良好的途径。

4.1.1 什么是面向服务的体系架构

面向服务的体系结构是一个组件模型,它将应用程序的不同功能单元(称为服务)通过这些服务之间定义良好的接口和契约联系起来。接口是采用中立的方式进行定义的,它应该独立于实现服务的硬件平台、操作系统和编程语言。这使得构建在各种这样的系统中的服务可以以一种统一和通用的方式进行交互。

这种具有中立的接口定义(没有强制绑定到特定的实现上)的特征称为服务之间的松耦合。松耦合系统的好处有两点,一点是它的灵活性,另一点是,当组成整个应用程序的每个服务的内部结构和实现逐渐地发生改变时,它能够继续存在。而另一方面,紧耦合意味着应用程序的不同组件之间的接口与其功能和结构是紧密相连的,因而当需要对部分或整个应用程序进行某种形式的更改时,它们就显得非常脆弱。

对松耦合的系统的需要来源于业务应用程序需要根据业务的需要变得更加灵活,以适应不断变化的环境,比如经常改变的政策、业务级别、业务重点、合作伙伴关系、行业地位以及其他与业务有关的因素,这些因素甚至会影响业务的性质。我们称能够灵活地适应环境变化的业务为按需应变的(On Demand)业务,在按需应变的业务中,一旦需要,就可以快速地对完成或执行任务的方式进行必要的更改。

4.1.2 原有系统的集成

当基于 SOA 来构建一个企业级的系统架构的时候,一定要注意对原有系统架构中的集成需求进行细致的分析和整理。我们都知道,面向服务的体系结构是当前及未来应用程序系统开发的重点,面向服务的体系结构本质上来说是一种具有特殊性质的体系结构,它由具有互操作性和位置透明的组件集成构建并互连而成。基于 SOA 的企业系统架构通常都是在现有系

统架构投资的基础上发展起来的，我们并不需要彻底重新开发全部的子系统；SOA 可以通过利用当前系统已有的资源（开发人员、软件语言、硬件平台、数据库和应用程序）来重复利用系统中现有的系统和资源。SOA 是一种可适应的、灵活的体系结构类型，基于 SOA 构建的系统架构可以在系统的开发和维护中缩短产品上市时间，因而可以降低企业系统开发的成本和风险。因此，SOA 架构首先考虑的应该是如何重用已有的投资而不是替换遗留系统，因为如果考虑到有限的预算，整体系统替换的成本是十分高昂的。

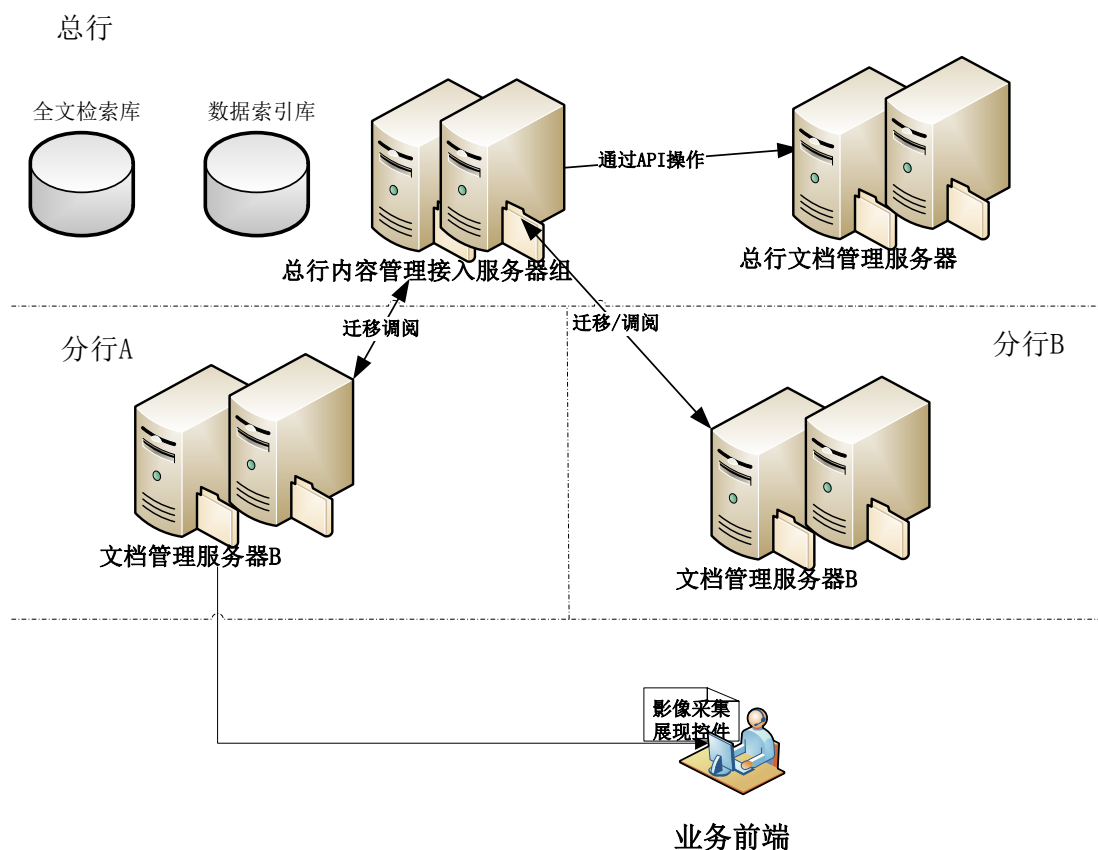
当分析原有系统中的集成需求的时候，不应该只限定为基于组件构建的已有应用程序的集成，真正的集成比这要宽泛得多。在分析和评估一个已有系统体系结构的集成需求时，我们必须考虑一些更加具体的集成的类型，这主要包括以下几个方面：应用程序集成的需求，终端用户界面集成的需求，流程集成的需求以及已有系统信息集成的需求。

4.2 SUNECM 技术特点

4.2.1 SUNECM 物理架构

SUNECM 由客户端、Console Server(索引服务器)、Document Manager Server(文档管理服务器)、Unity Access（统一接入服务器）四大逻辑部分组成。Console Server(索引服务器)管理内容的索引（即结构化数据的管理），Document Manager Server(文档管理服务器)根据 Console Server 的配置要求管理（或收集）非结构化数据，并可以采取分级存储管理方法，可连接磁带库和光盘库等价格较低而性能较好的二级存储设备。SUNECM 系统由一个索引服务器以及一个或多个文档管理服务器组成。索引服务器和文档管理服务器可运行在同一台机器或节点，也可运行在不同机器或节点。客户端、索引服务器和文档管理服务器构成信息传递的三角形架构。客户登录索引服务器，然后查询需要的信息，将指定文档管理服务器信息送达客户，而如果通过统一接入服务器则可以全局查询架构内的文档服务器信息，便于多级，分布部署时的信息查询。

SUNECM 架构非常灵活，可以满足多种行业、不同规模，不同硬件环境、不同区域的需求，同时其体系架构随着业务扩展可方便的扩展而无需修改相关应用。其体系架构可根据业务需要、业务规模、网络环境等条件灵活的部署为集中、集中式索引+分布式存储、联邦式部署等多种架构方式。



SUNECM 部署逻辑图

4.2.2 业务系统接入架构设计

统一内容管理接入平台提供以面向服务为基础采用多种方式与影像应用系统和业务相关系统进行整合，通过标准的协议（如 HTTP、Socket、SOAP、Web Service 等）或者定制的接口实现应用系统与平台的集成。接入平台可以由客户端调用，如影像组件，也可以由业务系统服务端调用。

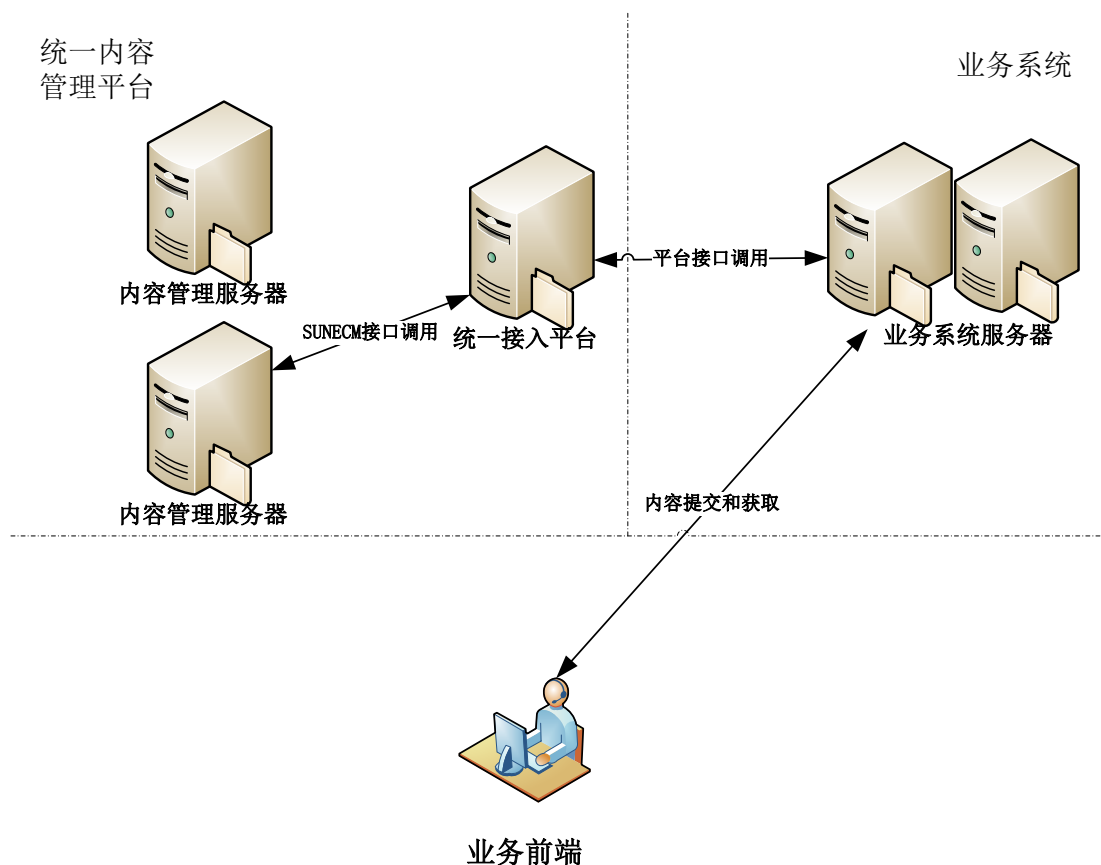
根据业务应用接入的要求，考虑业务系统的部署情况、网络情况以及业务系统架构情况，业务系统接入内容管理平台，有三种方式：

4.2.2.1 服务端集中接入模式

服务端接入模式：内容和元数据全部由统一内容管理平台服务器集群和业务系统服务器存储转发。

业务系统的服务器通过内容管理平台的统一接口，对内容进行操作，内容在客户端的展现处理以及内容到业务系统服务器的过程由业务系统实现，接口仅存在于统一接入平台和业务系统服务器之间，既可以实现系统之间的单笔内容及元数据提交与获取，也可以实现系统之间的批量内容及元数据提交与获取，内容提交和获取的过程全部由服务器之间调用接口实现，过程和权限可控，但业务系统设计较为复杂，如果业务系统已经实现了对内容的基本处理，如展现和传输等，可以较为方便地采用该方式，如 OA 系统、提回系统、会计档案系统等。

业务前端随业务处理过程，获取内容，提交到相应的业务系统服务器，业务系统服务器可以临时存储内容，然后根据规则策略，可以即时，也可以定时调用平台接口将内容及元数据提交到统一接口平台，然后存储到 SUNECD 中进行管理。内容调阅获取也通过业务系统调用平台接口获取内容及元数据，由业务系统完成从服务端到业务前端的传输。



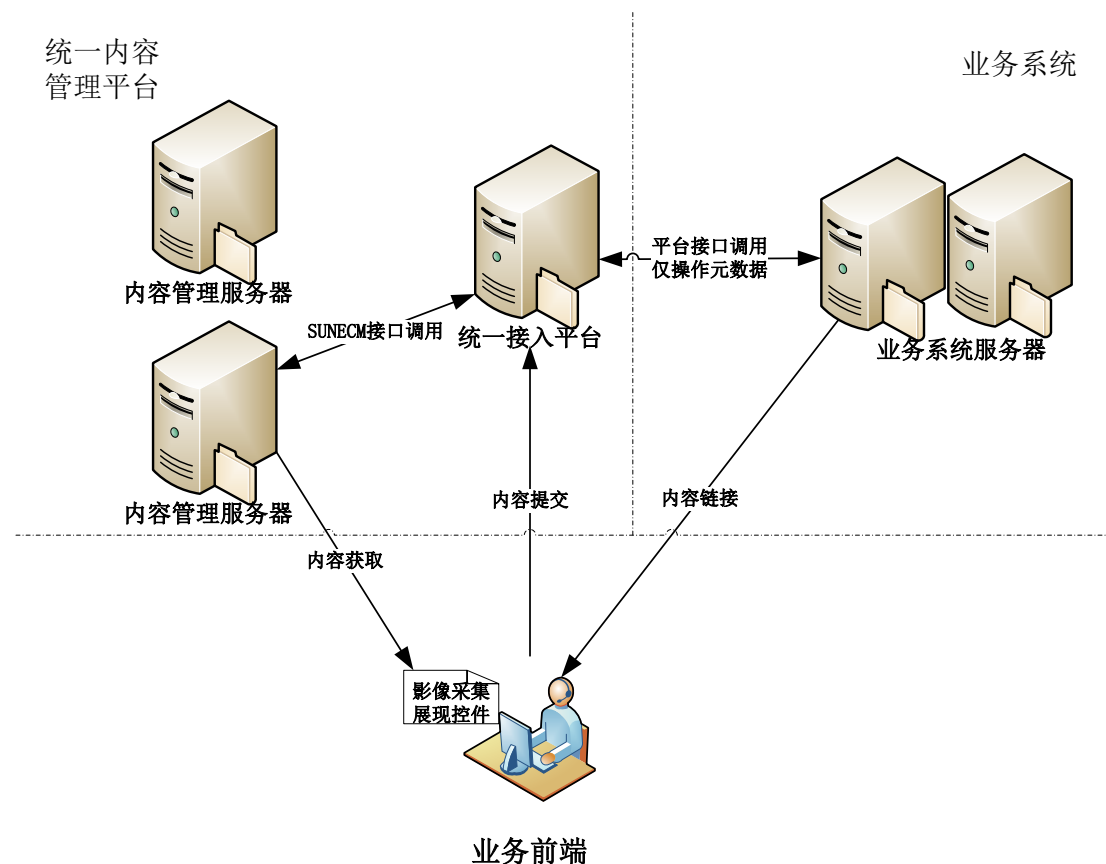
服务端集中接入模式

4.2.2.2 服务端+客户端集中接入方式

服务端+客户端集中接入方式：元数据由业务系统服务器通过内容管理平台服务器集群获取，同时获取内容的链接，内容由客户端通过链接直接向内容管理平台获取。

业务系统的服务器通过内容管理平台的统一接口，对内容的元数据进行操作，并获取内容的链接传递到业务系统前端，内容在客户端的展现等处理可以采用统一 UI 设计的采集展现控件实现，该控件可以嵌入 IE 等各种前端系统。实现了内容元数据访问过程和权限可控，通过控制元数据也可以控制内容的访问，内容本身是大容量数据，可以直接从存储管理端发送到业务前端，提高效率，业务系统接入时可以直接使用现成的前端采集展现控件，可以非常方便地接入，如信贷系统、国结系统等。

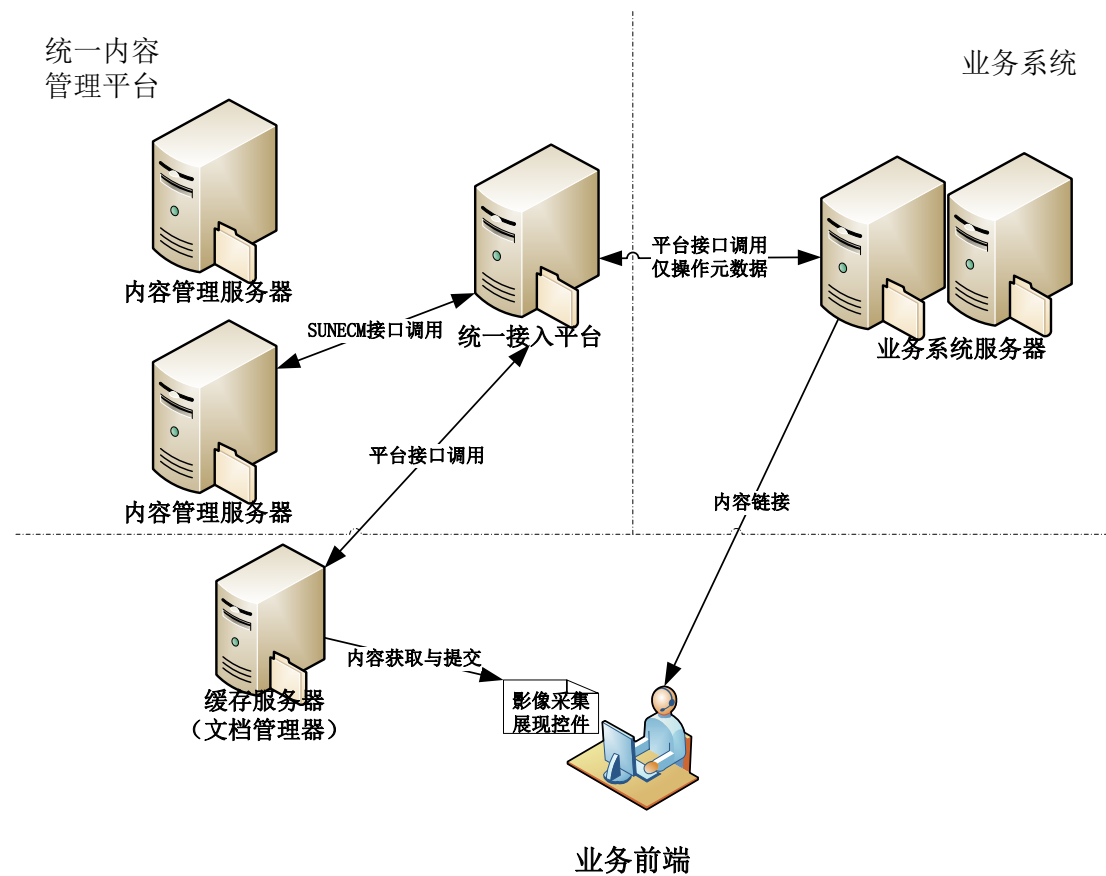
业务前端随业务处理过程，获取内容，提交到统一接入平台，存储到 SUNECM 中进行管理。内容调阅获取时通过业务系统调用平台接口获取内容元数据，获取内容链接，提供到前端采集展现控件，由前端展现控件通过链接直接从存储服务器上获取实际内容。



服务端+客户端集中接入方式

4.2.2.3 服务端+客户端分布接入方式

由于部分业务数据量非常大，全部依赖 SUNECM 和网络带宽，容易影响效率，可以在分支机构部署缓存实现临时存储和就近访问，如账务中心系统等。在集中方式上增加了在缓存上提交和获取内容的功能。应用系统不需要了解影像内容管理平台的部署方式，内容管理前置提出访问需求，内容管理前置通过安全和权限访问控制验证后，直接，实现应用系统完全透明访问内容平台数据。



服务端+客户端分布接入方式

4.2.3 安全与权限设计

● 应用接口操作安全

对内容元数据的访问采用系统服务器直接的接口处理实现，相对客户端直接访问，提高了安全性。

对内容本身的访问采用直接访问可以提高效率，同时由于对内容元数据的访问进行了控制，可以内容链接的访问时间和次数进行控制，同时可以通过技术手段限制指定的客户端访问内容

● 访问权限控制安全

增加权限控制，主要是对应用及组织机构关系中，所属业务模型对象的增、删、改、查权限，应用的权限控制就是对用户所属应用或可以操作的用户进行权限控制，组织机构的权限是指上级机构有对下级机构的相关权限，接口中需要增加用户和相关密码，在用户认证上有两种方式：

一种是真实用户，统一内容管理平台需要调用行方的统一认证平台来验证用户，按照认证平台返回的可以操作的应用以及所属网点，对用户可以操作的权限进行控制，同时必要时统一内容管理平台上根据每一个真实用户控制其权限，增加对响应应用的增、删、改、查权限；

另一种是虚拟用户，每个业务一个用户或多个用户区分权限（增、删、改、查权限），业务在进行相关影像文档操作时，需要提供相应的虚拟用户或业务系统用户登录凭据等来进行处理（业务系统来控制操作的权限，简单实现方式是对显示界面和按钮等用来操作影像的功能进行权限控制，而不需要单独定义内容管理权限）；

如若需要其它类型操作权限，须单笔单独授权，如跨区域和跨业务的查询

4.2.4 SUNECM 提供 SOA 的内容服务

面向服务的体系结构是一个组件模型，它将应用程序的不同功能单元（称为服务）通过这些服务之间定义良好的接口和契约联系起来。作为内容管理的平台，SUNECM 提供了面向 SOA 的完备的 API，用于构建那些需要使用集成的虚拟化内容、内容管理的内容服务应用程序。

● 集成的 API

集成 API 是一种可轻松控制的、面向对象的 Java 应用编程接口，可为您的应用程序提供内容管理和 workflow 功能。应用编程接口和应用编程接口类的实现不会根据内容管理系统提供的、应用编程接口所提交的服务而发生更改。这意味着可以创建支持内容的应用程序或内容管理应用程序，它们可同时使用多个内容管理系统进行联邦操作，可在不同用途的或不同的内容管理系统之间移动。

● Web 服务应用编程接口

Web 服务应用编程接口（Web Services API）支持企业业务应用程序与内容和内容管理服务的集成，可通过 Internet 或 WAN 和防火墙访问这些服务。WSAPI 将一系列非常类似的服务显示为基于 Java 的集成应用编程接口，但使用的是 SOAP 和 WSDL 标准，且可通

过 HTTP、HTTP/TLS 和 SSL 进行访问。WSAPI 具有与 Microsoft .NET 和 Java SOAP 开发工具的互操作性。这些接口是使用 Web Services 描述语言（Web Services Description Language, WSDL）进行说明的。

4.2.5 SUNECM 关键技术

一、安全性

1、数据的访问控制

内容管理系统中使用的软件都有完善的授权机制，可以控制用户只能访问授权的数据，并记录对于非授权数据的访问。系统对非结构化数字信息的访问和使用进行全方位、多层次的许可、控制和监督，可以保证只有经过授权的用户才可以对内容数据进行修改和删除等操作。可以实现内容对象的访问权限控制；这些权限包括对内容的创建、查看、修改、删除以及添加、删除、查看内容注解等等。另外，用户也可以定义自己所需要的权限并添加到系统中。同时，用户可以配置记录相关的日志信息，这些日志信息包括数据的创建、读取、修改以及删除等各种交易日志，以保证各种操作的可控性。对于内容管理系统管理的数据，可以通过定期备份的方式进行保存，一旦发生意外的情况，可以在数据的级别，进行数据的恢复。

2、数据的一致性和完整性

提供对于内容的事务处理机制，来提高数据的一致性和完整性，而这是文件系统无法保证的。

二、高效性

SUNECM 拥有优良的体系架构，当系统从低并发到高并发，从小数据量大数据量，从低并发到高并发的时候，均可以提供高效的访问和优良的性能。

SUNECM 体系结构是基于由一个数据索引服务器、一个或多个文档对象服务器和一个或多个客户程序构成的分布式三角形模型。在这个架构下，内容和索引数据被分开存放，使用不同的存储和管理策略，当对系统进行调优时，可以拥有不同的调优策略以最大程度实现高性能。

数据索引服务器基于关系型数据库，因此，任何针对关系型数据库的优化都可以对 SUNECM 提供性能上的提升。文档对象服务器基于多种应用服务器，管理底层的各种具体内容的存储。SUNECM 提供了很多有效的存储手段，例如，多层存储和多层级目录结构。

为了增加系统的存储性能，支持多层级存储结构。准许用户将不常用的内容文件从快速的设备上迁移到二级存储设备。这样可以有效的利用高速设备的资源，避免高速的盘阵设备因为存储文件数过多而降低系统性能。

同时 SUNECD 还使用多层级目录结构，将大量的内容文件保存在不同的子目录下，使用 Hash 算法，均匀的将大量的文件分散开。提高文件系统目录性能。

三、扩展性

SUNECD 拥有良好的扩展性包括：元数据模型的扩展能力、对象存储的扩展能力、应用的扩展能力，对象格式的扩展能力等方面。正是这些良好的扩展和集成能力，才使得我们能够渐进式的完成整个方案，管理企业产生的各种内容，真正实现企业内容管理平台。

首先，SUNECD 系统拥有强大的元数据的扩展能力。支持数字内容管理界各种流行统一的元数据标准，如 Z39.50、Dublin-Core 等。同时，数据索引服务器所提供的元数据管理中心提供非常灵活的多层次多版本的数据模型，并通过关系型数据库的技术管理数据模型。

从存储上来看，SUNECD 支持基于存储成本、存储策略的多层级存储管理，以支持更海量、更长期的存储需求。支持多个分布式资源管理器，允许用户在需要扩充存储的时候任意增加新的存储服务器，提高系统容量。扩充时，支持对多机种服务器的支持，允许在建立新的数据类型的处理过程中，使用不同的存储服务器，将新应用快速发布到新的服务器上。同时，SUNECD 支持在相同或不同操作系统上，集成多个文档对象服务器，而且可以从一个操作环境到另一个操作环境实现转换。

从应用角度上看，SUNECD 提供了强大的 API 支持，包括了完整的 Java 和 C++ API 以及 Web Service 接口。使得 SUNECD 可以挂接任何应用系统。

SUNECD 在设计初就优先完成了系统操作的各种 API，再在 API 层之上开发产品客户端。也就是说，SUNECD 的管理客户端、用户客户端都是基于 Java 或者 C++ API 基础上开发完成。因此，用户针对 SUNECD 的所有操作，都可以通过 API 完成。同时也确保用户应用可以使用到 SUNECD 的全部功能。这种 API 架构决定了 SUNECD 是一个彻底开放的可扩展的平台，方便的接入和开发各种新应用。

四、可靠性

SUNECD 基于成熟的数据库技术和 HACMP 软件的协助保证数据的高可用性。当用户有多个文档对象服务器时，可以指定在多个文档对象服务器间复制数据。复制可以是单向的也可以是双向的。用户可以使用这个功能，将远程的资源管理器的数据同步到本地，增加访问效率；也可以使用这个功能，将各个分布的远程文档对象服务器的内容集中到中心；也可以使用这个功能实现文档对象服务器的数据备份。

同时 SUNECD 可以通用应用服务器的集群方案来进行负载均衡，对于访问量大的文档对象服务器可以进行单独的集群，当然，数据索引服务器也可以进行集群。

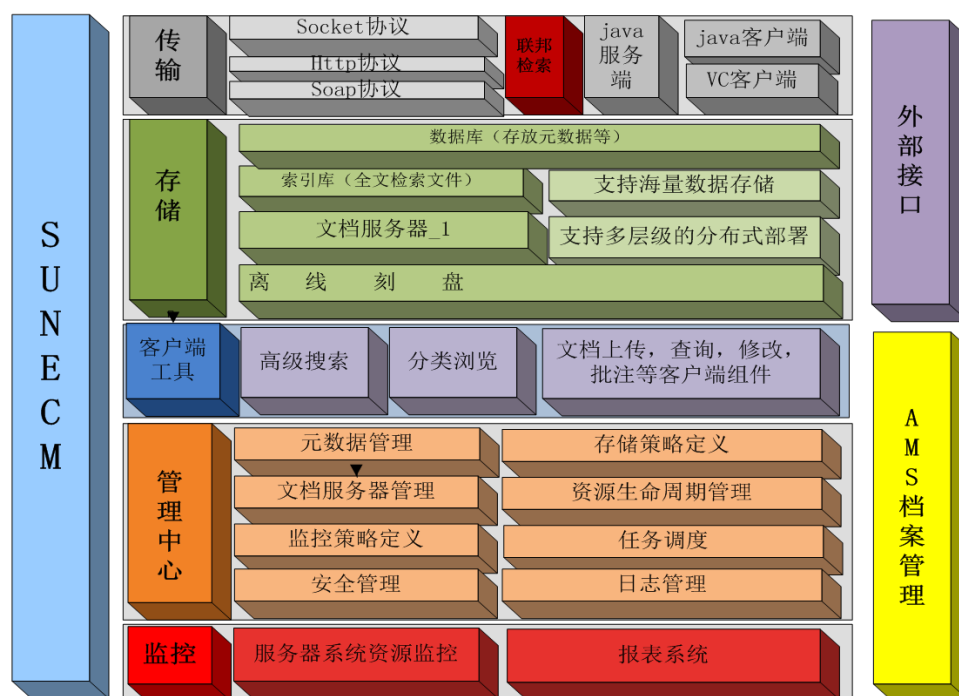
五、易用性

SUNECD 拥有基于 B/S 架构的管理工具，方便客户对服务器进行全面的管理和控制，可以创建和维护元数据模型，创建节点服务器，配置系统的存储迁移策略，管理用户

和权限等任务。提供了资源监控接口，方便管理分布式部署的多个文档对象管理器，提供了编程指南和 API 参考手册大量的样本程序供开发人员熟悉 API 的使用。

4.3 SUNECM 功能概述

4.3.1 软件模块图



SUNECM 软件模块图

4.3.2 软件模块划分

序号	模块
1	SUNECM Console Center(管理中心)
2	SUNECM Document Server (文档服务器)
3	SUNECM Trans Engine (传输引擎)
4	SUNECM Unity Access (监控中心)
5	SUNECM Monitor (监控中心)
6	SUNECM Client (客户端工具)

SUNECM 模块表

➤ SUNECM Console Center (管理中心)

模块功能:

- 元数据模型的配置
- 元数据属性配置
- 文档服务器的管理
- 文档服务器组管理
- 文档服务器负载均衡管理
- 文档服务器的路由策略配置
- 统一接入服务器管理
- 安全管理
- 存储策略管理
- 资源的生命周期定义
- 版本管理
- 监控策略定义
- 任务定义
- 监控管理

- 客户端工具

➤ SUNECD Document Server（文档服务器）

模块功能：

- 非结构化数据存储/查询
- 传输服务
- 迁移服务
- 清理服务
- 资源的生命周期管理
- 非结构化数据的存储
- 支持分布式、集中式的部署
- 目录服务

➤ SUNECD Trans Engine（传输引擎）

模块功能：

- 提供文档的增、删、改、查等接口
- 文档传输过程的安全验证
- 提供 JAVA 、Web service 的客户端
- 提供多种传输方式 http、socket 、soap 等

➤ SUNECD Unity Server（统一接入平台）

模块功能：

- 进行全局式接入的文档增、删、改、查等操作
- 提供全局式分类浏览、高级搜索
- 各种格式内容的展示
- 提供 JAVA 、Web service 的客户端
- 提供多种传输方式 http、socket 、soap 等

➤ SUNECD Monitor（监控平台）

模块功能：

- 文档服务器、数据索引服务器的系统资源的使用情况
- 文档服务器间文档迁移情况
- 文档服务器的文档存储情况

➤ SUNECM Client（客户端组件）

模块功能：

- 进行文档的增、删、改、查、批注等操作
- 提供分类浏览、高级搜索
- 各种格式内容的展示

4.3.3 SUNECM 主要功能描述

4.3.3.1 元数据管理

元数据模型的设计应该是企业内容管理中最有挑战性的技术实现之一。元数据模型的能力直接表现出一个企业内容管理平台适应用户需求的能力。在实施企业内容管理的初始阶段，我们只能了解很少的一部分需求，如果我们设计数据模型时仅基于这些特定的需求，那么这个应用可能只能适应客户当时的需求，而无法适应客户的需求变化。在内容管理中，通常元数据越复杂，内容提升价值的潜力就越大。而丰富元数据的模型不是一蹴而就的，这就要求一个面向客户内容管理的通用数据模型，以适应客户不断变化的需求，提升信息的价值。

SUNECM 中的元数据模型采取面向对象的方式，我们可以把每一个业务类型看作一个对象，将之统称为业务模型对象。业务模型对象主要由二部分组成，一种是属性（即业务类型的索引信息）通常我们可以称之为结构化数据，另一种是文件信息（即业务类型所包含的一些文档），通常我们称之为非结构化数据，在元数据对象间可以定义关联关系。同时对于每个业务模型对象，用户可以选择是否记录历史版本。

➤ SUNECM 的元数据模型：

索引信息				文件信息 1	文件属性
批次 ID	属性 1	属性 2	文件属性

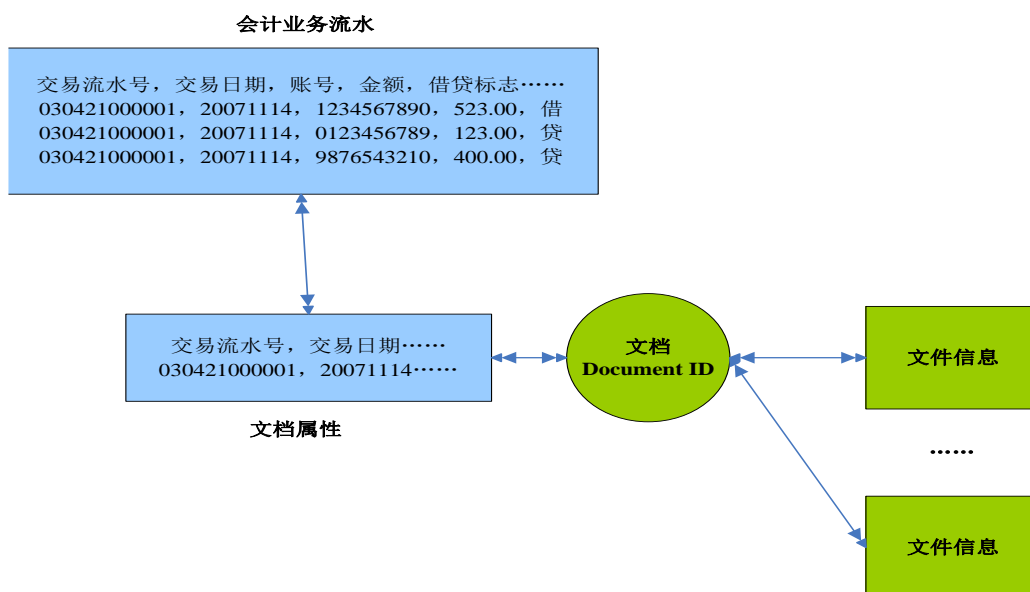
元数据对象模型

在这个模型中，在对一个内容的描述中，支持对复合文档的精确描述，可以对同一文档中不同的文档对象实施不同的处理和控制；该模型也可以灵活地描述内容和内容之间的关系以及内容在使用中的工作状态。而图像化的建模工具以及支持建模的 API 可以帮助方便客户和应用开发商通过不同方式快速建模并进行调整。

根据银行业务资料特点和业务索引构成，我们一般将业务影像的元数据模型分为以下两种：

● 单层索引结构

多张业务影像对应单一的索引信息，如会计档案。会计凭证影像与业务流水号一一对应，通过业务流水号调阅该笔业务的多张业务凭证影像。

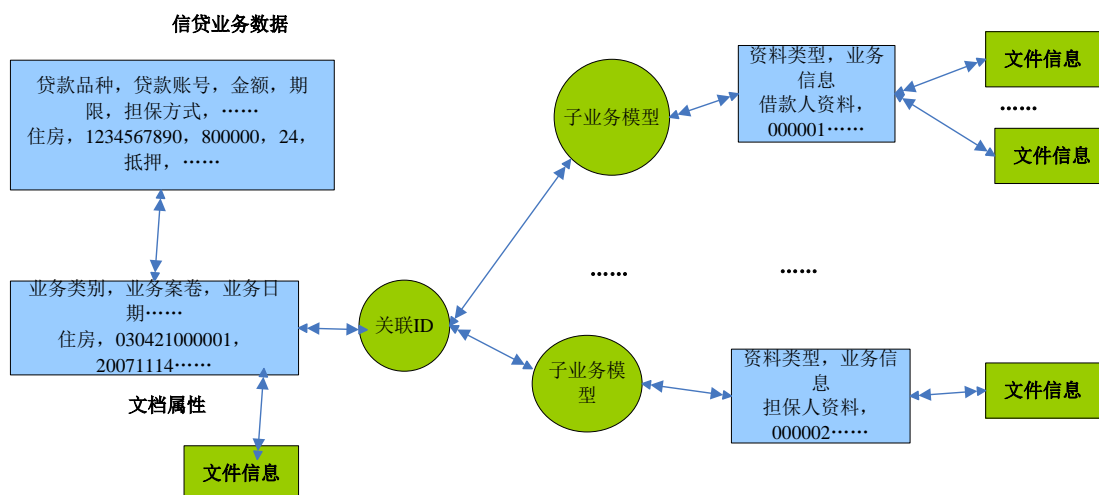


单层索引模型

● 多层索引结构

多张业务影像包含多层次的业务索引信息，如信贷档案。信贷档案对应信贷档案信息，在信贷档案内部需要对业务影像进行细分，按照业务资料类型、业务资料对应业务信息建立多级索引信息。

从索引查询性能等综合考虑，一般建议索引模型层次不宜太多。按照目前的业务资料特点，可采用二层索引模型。



多层索引模型

● 海量数据支持

内容管理系统中必然后存在着大业务量的数据，如果提供海量数据存储及检索的效率，SUNECM 充分的考虑到银行业务使用和海量历史查询的需要，其适应于海量信息保存和查询的自动分表技术，可以很好地解决索引信息持续增加对数据库的压力，消除了当数据库中的记录存储到一定数目后急剧下降的表现，为海量内容信息查询提供稳定的访问能力。不会因为数据库的不断加大，而提高相应的硬件和软件成本，同时降低系统的维护成本。

4.3.3.2 传输引擎

在内容处理过程中，影像等业务资料需要在影像采集录入节点向文档服务器传输，各业务工作节点需要从文档服务获取影像资料，加以显示。传输引擎主要负责在各级行间的文档服务、档案存储的各类资源的传输。

能够实现到 workflow、各种内容管理的支持，内容管理系统可以通过与 workflow 的结合，配置成影像内容存储到内容管理系统后，即启动相关的流程进行业务处理。

传输引擎用于在服务端和客户端主机之间进行文件的传输。它基于 Http、Socket、Web Service 等协议开发，可以保证服务端与客户端之间在大数据传输情况下的可靠性，稳定性。使在网络带宽很低的情况下能够保证数据传输的完整性。

在影像处理过程中，影像等业务资料需要在影像采集录入节点向资源服务器传输，各业务工作节点需要从文档服务器获取影像资料，加以显示。传输引擎主要负责在各级行间的资源服务、档案存储的各类资源的传输。

传输引擎能够达到资源分布式部署的特点，在资源采集的每个节点，传输引擎具有缓存资源的功能，能对本节点资源进行管理，提供简单的资源检索。

4.3.3.2.1 传输引擎的架构特点

传输引擎本身提供的二次开发的接口，客户可以开发丰富的应用和传输引擎集成，从而提升传输引擎的功能。系统有以下特点：

- 提供多种传输协议支持：soap、http、socket、webservice。
- 提供 Java 客户端各类协议 API，方便各应用系统调用。
- 服务端基于 J2EE 架构设计，模块分层设计，模块之间松耦合。
- 基于 XML 设计。
- 高度并发的多线程调度、处理设计机制，保证系统的性能和稳定性。
- 二次开发能力，基本 SOA 的架构设计，可以更好的将二次开发出来的模块集成进入统一接入平台。
- 提供内容的增加、修改、删除、查询、批量查询、批量导入、分页查询、高级搜索等功能丰富的 API，满足不同的应用系统的需求应用
- 传输高效性、稳定性。
- 二次开发应用的开放性。
- 安装简单方便，仅仅需要一个 JRE 环境。
- 是一个纯 Java 应用，跨平台、能够在不同的操作系统运行 (Windows、Linux、Unix<AIX、Solaris、HP-Unix 等>)。
- 采用底层 Socket 通讯机制，提高通讯性能。

为了提高性能，传输引擎平台服务端启动时，可以同时启动多个服务线程，形成服务线程池。这样新的客户端请求可以立即得到处理，避免了建立和中止服务线程的开销。当客户端请求数量巨大超过了预先启动的线程数时，可以通过配置文件允许增加服务线程来处理客户请求。

4.3.3.2.2 传输功能列表

序号	功能	描述
1	断点续传	为保证数据的完整性所提供的断点续传、自动重发
2	流量控制	在网络带宽较小，且需要办理其他业务情况下，传输服务能够在控制影像传输所占用的网络带宽范围内充分利用带宽
3	统一接入管理	所有的操作通过统一接入平台进行操作，减少了操作的复杂性，使得 SunECM 架构部署对操作人员透明化
4	统一认证	可防止未经授权的人员访问和修改传输代理服务中的内容，传输数据经过加密。
5	均衡负载	根据目前传输数量和连接数的大小调整服务器组内的负载，平衡组内服务器性能
6	缓存功能	与资源服务一起提供本地资源缓存功能
7	文件下拉功能	可将上级服务器中最新的业务文件下拉至本地缓存服务器，使得分布的版本一致
8	支持上传策略	可以在传输服务器进行配置，按照业务确定上传方式：实现异步、定时、即时、实时（直接传输到 CM 中）的传输策略
9	策略管理	各级文档服务器依据各类策略自发的进行相关迁移，清理等数据操作

4.3.3.2.3 数据传输的完整性

数据传输的完整性包括两方面的内容，即在网络上传输的数据不能被恶意篡改和数据完整的传输到目的地。采用 MD5 码或其他加密方式进行完整性检验

第一、传输引擎对接入服务的连接进行身份验证，通过合法验证的用户才能登录传输服务，同时登录传输平台的其他用户不能停止或修改传输中的数据，可以保证数据不被恶意篡改和数据传输前后的一致性。

第二、传输引擎对整个系统中的上传文件和下载文件进行状态的监控，记录文件传输中“未传输”、“已传输”两个状态。传输引擎具有出错记录，断点续传，自动重连，多线程上传下载多个文件。通过这些传输机制，可以有效的保证数据不重复传输，确保数据完整传输到目的地。

4.3.3.2.4 数据传输的稳定性

数据传输在很大程度上要受网络带宽约束，在网络带宽不好的情况下，传输连接容易中断，数据包容易丢失，造成文件传输失败。SUNECM 传输服务能够在有限的网络资源的前提下，保证数据包准确完整的发送和接收，以最低的响应时延通过网络而不受时间、用户数变化和突发网络事件因素的影响。系统采用断点续传，自动重连，解决数据传输的稳定性。

4.3.3.2.5 数据传输的安全性

传输引擎采用目前流行的安全认证方式，对数据操作进行严格的审核，同时对传输的数据进行加密后再传输。端到端加密是指信息由发送者加密，然后，作为不可阅读和不可识别的数据通过网络，这些信息一旦到达目的地，将自动重组、解密，成为可读数据。

对用户操作通过令牌的管理进行操作合法性审核；

通过软件加密安全传输控制的方法，在不增加任何硬件设备的情况下，就可以有效的保证影像数据在网络中的传输安全；

同时通过对 URL 的合法性进行一定的加密，以保证此 URL 只会在一定时间内有效，同时 URL 地址并非物理实际存储地址，避免了存储信息外泄。

4.3.3.2.6 缓存传输策略

传输引擎对数据传输提供了策略配置，可以按不同业务处理性质，（如联机处理、历史信息查询等）配置多种传输策略。并且所有的配置都只需在总级的配置界面中修改，自动下发各级文档服务器，无需每次到下级缓存服务器进行修改。



文档服务器的传输策略

4.3.3.2.7 信雅达传输引擎优势

- 自主研发基于 TCP/IP 协议的传输文件和消息交易中间件。
- 提供丰富的二次开发接口，能够开发相应的应用。
- 保证传输数据的完整性、稳定性、安全性。
- 提供资源分布式部署解决方案，完成文档流转和资源共享。
- 开放的接口，方便灵活的与内容管理产品做集成。

系统在多家银行的内容管理系统中得到的大量应用和验证，系统稳定可靠。

4.3.3.3 内容整合

4.3.3.3.1 统一接口服务

内容管理软件是一个基础性的平台，他为多个业务系统提供了业务支撑，所以基于安全性以及通用性来考虑，基于面向服务的先进思想，我们在 SUNECM 的接口封装成统一接口服务，屏蔽所有底层内容访问的细节，包括内容的存取、内容的归档、索引信息与内容之间的同步功能等，业务系统不用直接操作内容管理软件，所有对内容管理软件访问全部能过统一接口服务来进行，然后由统一接口平台来负责操作 SUNECM 内容管理平台。

每个业务应用系统配置独立的索引结构，在 SUNECM 中保存为单独的业务模型对象，业务量比较大的业务可以根据时间（按年或月）来定时创建新的业务模型对象。对于业务量特别大的业务应用，以达到分散 SUNECM 压力，满足业务系统对内容存储和管理期限的要求。

- 多 SUNECD 存储支持

存储到 SUNECD 时，根据业务生成时间和业务类型，通过数据库配置的内容，确定需要存放的 SUNECD 服务器名，在服务期连接池中获取 SUNECD 连接，进行存储操作。

- 多 SUNECD 查询支持

根据数据库配置的某一业务在一定时间内存放的 SUNECD 服务器名，在服务器连接池中获取 SUNECD 连接，进行查询。

多业务类型对象支持

根据业务类型和时间确定存储和查询需要的 SUNECD 中业务模型对象名，一般的处理方式是在确定多 SUNECD 存储支持的同时确定业务模型对象的名字，这样就同时支持多 SUNECD 和多业务模型对象系统架构情况了。

4.3.3.3.2 联邦检索

在整个银行的网络中，分布部署着多个的 SUNECD 服务器，每个 SUNECD 服务器上又存放大量的内容，所以每个 SUNECD 服务器相当于一个信息孤岛，如何在众多的信息孤岛上快速的定位查询内容，加快信息检索的速度，成为每个企业日益关注的话题。

提供一个信息整合的体系架构来实现单一入口的针对多个 SUNECD、同种类或不同种类的访问接口。为此我们提供开放的、易于扩展的体系架构，通过连接器访问多个 SUNECD 内容存储服务器，同时支持基于公共接口的联合连接器来访问任何的内容数据源，并且用户能依据标准接口来定制其他的内容源的连接器，此时，如果查询缓存节点与影像内容所在节点，网络是相通的，查询节点能够直接访问到内容，而如果此网络是星形网络，那么会根据路由间的关系，找到网络相通的上级节点，通过上级节点的路由获取内容。并且用户能依据标准接口来定制其他的内容存储，能整合企业内部的所有内容存储包括已有的内容存储库和未来搭建的内容管理平台。系统具备统一访问、统一展现的能力。

当业务系统无法确定需要检索内容的时间情况下进行查询时，需要通过联邦检索功能向多个 SUNECD 和文档服务器进行查询，系统在接收请求后，根据配置的要求（根据请求类型确定需要检索的 SUNECD 组，不能任意在全部的 SUNECD 中进行查询，否则会极度影响性能），由多线程同时查询多 SUNECD 后，通过整合接口将这些查询结果进行整合，在将整合后的最终结果返回给客户端。

4.3.3.3.3 版本控制

基于不同业务类型的业务差异，可以选择性使用版本控制，采取版本控制的业务，将所有历史版本的影像会在 SUNECD 中保留，并且保证能随时能够调阅历史版本。

4.3.3.4 文档服务器

由于影像采集分散于各地，且采集点通过本级的广域网或者局域网连接起来，并且通过广域网与一级分行或总行相连进行通讯。为了进行业务操作，必定要在网络上传输大量数据，特别是在广域网上进行大数据量的传输对保障传输的稳定性、高效性、可靠性至关重要，因为传输不稳定可靠的话，外在表现出来的就是程序系统的不可靠稳定，从而降低业务处理的效率，降低了用户的满意率。

在采用了本地资源缓存服务后，采集点扫描或者导入电子文档数据，可以先把这些数据存放到本地文档管理器，并且在此本地文档管理器保留一个备份。由传输服务所设定传输服务处理逻辑决定是否实时、或者定时的传送到数据中心。收单点或者本级其它业务处理点需要调阅相应的电子文档数据时，对本地仍存在影像数据可以直接提取，而不需远程访问数据中心提取。在本地文档管理服务器中的备份清除逻辑规则由后台的服务来实现，可采用定期或事情触发方式。

基本功能如下：

- 增加相关管理监控接口
- 提供资源服务（即非结构化数据的存储）
- 通过统一管理服务的配置来进行备份、迁移清理服务
- 提供目录服务功能
- 提供权限控制
- 由于传输引擎提供和总行传输相仿的接口，文档管理器在上传上除了实现其它接口功能外还不需要改动，即和总行传输接口是相同的

4.3.3.4.1 支持对多应用的定义、存储管理的扩展能力

SUNECM 系统的体系结构是基于由一个数据库索引服务器、一个或多个文档管理器以及一个或多个客户程序构成的三角形客户/服务器模型。这种三角架的架构可以支持文档管理器灵活的进行部署，可以多层次的多级部署。

4.3.3.4.2 提供多级存储的管理和定位能力

SUNECM 内置的多层级存储管理技术可以使每个文档服务器都可支持多级存储体系的模式。其内置的存储功能可以根据用户已有的存储介质及存储内容的访问速度、期限等需要制定灵活的存储管理策略，可以提供给用户更多存储介质的选择，从而实现多级存储模式。

4.3.3.4.3 提供内容的生命周期管理能力

SUNECM 内置的存储管理功能提供对存储介质的管理。这些存储介质与系统的磁盘资源一起构成了一个容量巨大的存储池（Storage Pool），为数字化信息提供存储服务，存储管理模块在存储介质中保存和取得数据。其构成的内容管理平台可以使每类数字信息具有清晰的创建、迁移途径、演变历史和销毁功能，极大地简化了用户在内容生命周期管理的维护工作。

用户在实际应用中，可以根据数字化信息的使用频率、种类、优先级、及保存年限、以及存储介质的读写速度等因素，通过内容管理系统为每类数字化信息制定相应的存储策略。内容装入时，根据预先定义的存储策略，将内容装入到相应的存储介质上；查看内容时，从相应的存储介质提取信息，并返回给用户。

4.3.3.4.4 非结构化数据的存储

在内容获取后，通过传输引擎可将客户端的资源文件发送给传输引擎服务端，客户端根据就近原则来判断要连接的文档服务器上的传输引擎服务，在传输引擎接收文件之后会启动相应的服务模块对接收的影像文件（文件夹）按照一定的规则进行有组织的存放。

4.3.3.4.5 备份功能

文档服务和传输引擎结合可实现实时资源镜像备份功能，即可把本地文档服务中的内容数据上传到核心文档服务器中。采用此镜像功能后，各内容数据集中保存，通过各种数据备份设备实现多级备份，而各本地文档服务器只作为本地的一个数据缓冲备份，而不再需要配备那些复杂昂贵的备份设备。即提高数据访问速度，降低了网络负担，又可减少设备投资，降低系统维护难度。

镜像备份还可通过定时服务类来实现对批量文件送往核心文档服务器的动作，资源导入后的资源文件在存入核心文档库前都存放在本地文档管理器中，在利用晚上广域网络空闲时间，定时调度任务会将本地文档管理器中的文件倒到核心文档库中（此时不删除本地文档管理器中的资源文件）。

同时为了数据的安全性，也可以根据不同的策略将资源备份到其他的存储设备中如其他的磁盘、磁带、光盘等介质。

4.3.3.4.6 迁移清理服务

考虑到不同业务的需求及本地资源的存放空间有限及查询效率，不能将资源文件永久地存放在本地并无限地增加，必须制定一套本地资源文件的缓存迁移清理策略，对资源文件的去留能够有很好的判断能力判断是否进行提交上一级文档服务器，或迁移到其他位置，以方便下次查询时可以直接通过本地得到资源文件为衡量标准。因此我们在本地索引引擎中记录最近访问时间、访问次数，以便程序进行有效判断，提高再次访问的本地命中率，减少对核心资源服务器访问的负担。

影像迁移清理服务提供按时间，当资源文件存放一定时间后自动迁移清理的策略。

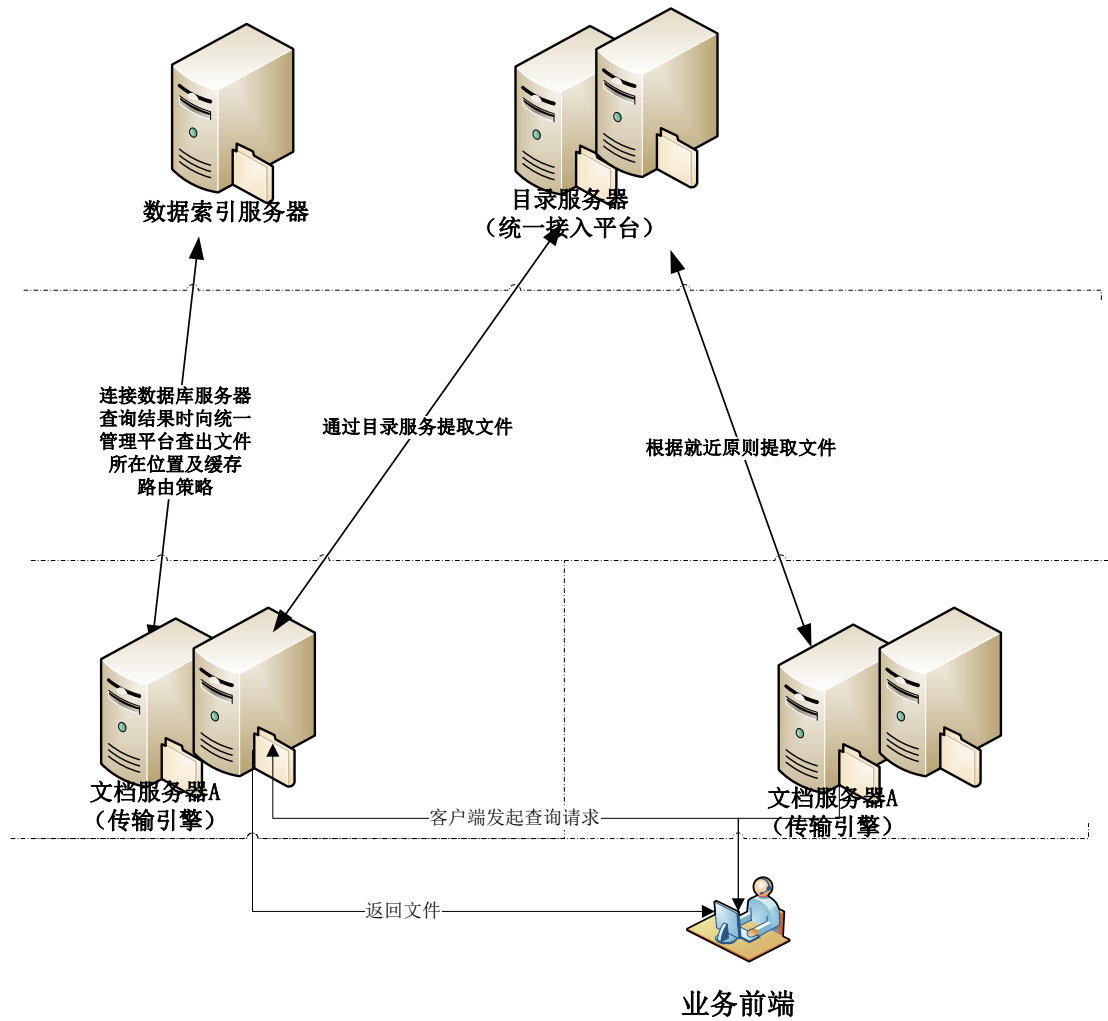
4.3.3.5 统一接入平台

文档服务器作为一个资源服务器，应具备内容存储及内容提取的功能，并能够实现及时的跨文档服务器的内容查询调阅等操作，以实现业务的高效性

在内容提取或展示时，客户端将所提取的资源请求发送到内容管理平台的定位服务中，定位服务将根据所提交的请求条件查询该资源文件在离该客户端最近的副本的资源服务，并产生一个位置指针和查询令牌，客户端在接受到此位置指针和查询令牌后直接从最近的资源服务中提取数据，资源服务在提取数据时将验证此位置指针和查询令牌，并将提取到的资源文件直接从该文档服务器中发送到客户端。

另外，目前，大部分银行采取网络是星形网络，分行与分行间的网络是互相封闭的网络部署环境，所以，当内容存储在其他分行时，缓存会根据统一管理服务所配置的路由策略来进行查找，取到最近的文档服务器上的内容，返回给客户端。

目录服务提取方式如下图：



统一接入平台工作方式图

上图所示统一接入平台流程为：当业务前端向文档服务器 A 发起查询请求，当请求的业务并未在文档服务器 A 时，可以通过查询统一接入平台服务器，统一接入平台服务器查询到文件目前保存在文档服务器 B 上，但是此文档服务器 A 与文档服务器 B 网络是不相通的，这时统一接入平台服务器根据路由策略就近来连接文档服务器 B，作为文档服务器 A 与文档服务器 B 的中继桥梁，将文档服务器 B 中的文件返回给业务前端进行调阅。

4.3.3.6 SUNECM 统一管理平台

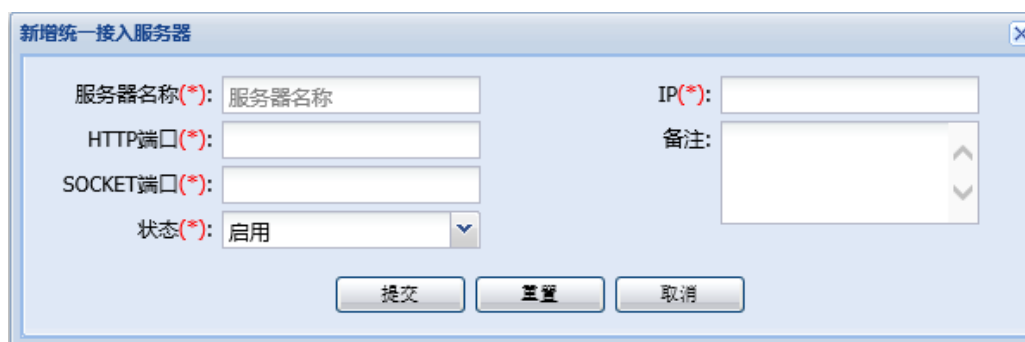
基于整个内容管理系统的分布式，集中式这种多层次的灵活部署方式，为了方便管理，使用资源达到共享的目的，我们设计了统一管理平台集中部署，文档服务器分布式部署的架构，解决了多文档服务器不易管理的问题。所有的文档服务器全部由统一管理平台来进行管理配置，同时统一管理平台又提供了对于文档服务器的监控服务。

4.3.3.6.1 业务模型对象管理

因为整个内容管理平台的部署方式，所以在每个文档服务器上存在着不同的业务类型数据，通过统一管理平台，可以灵活的定制业务模型对象的属性，权限，文件类型属性，所属的文档服务器，以及存储方式等。

4.3.3.6.2 文档服务器策略

文档服务器通过策略定义了文档服务器间的层级关系，如本地文档服务器向上级文档服务器的迁移策略，因为网络环境是一个星形的网络结构，如果在二个文档服务器间的网络情况是不能互相通信，那么一个文档服务器要访问另一个文档服务器上的内容，只有通过统一接入平台进行传递，达到文档服务器间的内容共享。此时这个统一接入平台相当一个连接中转器，通过这个连接中转器，达到整个网络环境上的内容共享。



新增统一接入服务器

服务器名称(*):	<input type="text" value="服务器名称"/>	IP(*):	<input type="text"/>
HTTP端口(*):	<input type="text"/>	备注:	<input type="text"/>
SOCKET端口(*):	<input type="text"/>		
状态(*):	<input type="text" value="启用"/>		

统一接入平台服务器配置页面

根据业务间的差异性及业务量的大小，不同业务对文档服务器的要求是不一样的，有的业务需要将内容临时存放到文档服务器，在非业务高峰期将内容集中，而文档服务器是一个多层级的架构设计，通过这种架构来缓解业务对系统，对网络带宽的压力，这时需要配置文档服务器的迁移等策略，这些需求全部是由文档服务器策略来进行配置的。配置后将策略分发给相应的文档服务器。

文档服务器管理提供了文档服务器的开启终止等任务，让管理员更好的通过页面远程的管理分布在网络内的文档服务器。

新增内容存储服务器

服务器名称(*): 服务器名称

IP(*):

HTTP端口(*):

备注:

SOCKET端口(*):

连接方式(*): 远程直连

状态(*): 启用

提交

重置

取消

文档服务器配置界面

为了文档服务器可以根据业务的变化而弹性的配置，SUNECM 引入了文档服务器组的概念，所有的文档服务器都归属于某一个文档服务器组中，组内的服务器协调工作，支持负载均衡，当业务量增大需要配置新的服务器时，只需在配置界面中为组内增添一个文档服务器即可。同时支持对组内不同服务器配置负载权重，为业务能力提供便利的性能弹性扩展。

新增内容存储服务器组

服务器组信息

服务器组名称(*): 服务器组名称

操作系统(*): Windows

部署方式(*): ECM负载均衡

状态(*): 启用

备注:

下一步

取消

配置服务器

服务器组配置界面

4.3.3.6.3 任务调度

作为一个基于全行的服务性平台，特别是部署的多层级文档服务器，特别需要有一个机制能够满足一些自动定时执行的任务，比如文档服务器的迁移清理策略，报表的跑批，迁移的备份，为此系统提供了任务调度的功能，能够自动执行用户自定义执行的一些任务，系统

具有很强的可扩展性，可以在系统运行一段时间后，系统不停机的情况下，执行一段新的 Java 程序。以完成特定的任务，如备份，导入/导出，格式转换等等。

首先需要将编译好的 class 文件或 BAT 文件放到工程目录下，然后再任务信息中输入到系统中。然后选择执行方式即可。

支持仅一次执行、每日定时执行、每周定时执行、每月定时执行、每年定时执行。

同时 SUNECM 平台内置了 4 种定时任务模式：

	任务名称
1	内容迁移
2	内容清理
3	内容离线
4	内容离线清理

支持的任务

➤ 内容迁移

各级文档服务器业务是不需要即时上传的，但是可能会在规定的时间内将业务数据汇总到上级文档服务器，此时配置内容迁移任务。为了节省网络，避开业务高峰期，给核心应用让出带宽。这部分业务将在规定的时间内上传（一般为夜间网络空闲时）。定时上传可以分多个时段上传，以免影响其他系统的业务。

➤ 内容清理

当内容达到文档服务器，并上传到上级节点以后，文件会被自动清理掉，以节约磁盘空间。

➤ 内容离线

当数据在文档服务器存储一定时间后，存储的内容按照策略转移到离线设备上，以降低在线数据文件系统的存储压力。

文档服务器任务配置界面

4.3.3.6.4 日志管理

作为多层次架构的文档服务器，如果查看分布在每个文档服务器上系统日志，对管理员来说是一件无比头疼的工作，统一管理客户端提供了统一的日志管理及查看功能，管理能够通过访问统一管理平台的页面来定制这些日志的级别，同时也可以随时查看分布在不同机器上的日志文件。方便快捷，而不必登陆到每台机器上去查看日志。同时可以日志管理界面定制需要输出那些日志。

A screenshot of a software dialog box titled "新增日志策略" (Add Log Policy). It contains four input fields: "类路径(*)" (Class Path), "日志级别(*)" (Log Level) with a dropdown arrow, "文件夹名称(*)" (Folder Name), and "日志大小(推荐300M)(*)" (Log Size (Recommended 300M)). At the bottom are three buttons: "提交" (Submit), "重置" (Reset), and "取消" (Cancel).

日志配置界面

4.3.3.7 二次开发接口

无论是自定义现有的封装应用或重新开发解决方案，都需要一系列丰富的、根据相应的开发技术定制的 API 来完成工作。内容管理平台提供了完备的 API，用于构建那些需要使用集成的虚拟化内容、内容管理和 workflow 功能的应用程序。包括集成的 API、虚拟内容管理系统应用编程接口、订阅事件服务应用编程接口、Web 服务应用编程接口等丰富的接口能力。

SUNECM 内容管理平台参考了业界大量的标准，为应用系统的开发测试提供了丰富的面向对象的 API，包括 C++，Java、Web Service 的 API。这些 API 包括了多种。包括内容处理、安全、系统管理等各方面的 API。基于 SOA 的设计思想方便用户快速搭建应用。同时也因为支持标准，使得跨文档服务器的信息检索调用成为可能。联邦检索系统可以根据通用的接口，完成对企业现有数据系统的综合查询调用。

内容管理平台在设计初就完成了系统操作的各种 API，再在 API 层之上开发产品客户端。也就是说，统一接入平台的管理客户端、用户客户端都是基于 Java 或者 C++ API 基础上开发完成。因此，用户针对统一接入平台的所有操作可以通过 API 完成。同时也确保用户应用可以使用到统一接入平台的全部功能。这种 API 架构决定了统一接入平台是一个彻底开放的平台，可以和任何应用系统、工作流引擎整合。

- 基本 SOA 的这种设计风格的架构设计，决定了其具有分布式、跨平台、可

互操作性、松散耦合的特点，可以解决企业信息化过程中不断变化的需求及异构环境集成的难题，所系统具有更好的灵活性及扩展性

- 基于 VC、JAVA 语言的客户端，可以满足多种应用系统，满足基于 JAVA、VB、VC 等语言的应用系统接入
 - 基于 J2EE 架构的服务端，完全可以将整个系统进行跨平台的部署
 - 提供的增加、修改、删除、查询、批量导入、批量上传、批量查询、高级搜索、分类浏览、分页查询等丰富的 API 可以更好的为应用提供了有力的支撑
- 支持多种传输协议及结构化数据基于 XML 格式的传输的可以满足多种业务系统的需求。

4.3.3.8 安全管理

安全管理是系统安全高效运行的基础，包括用户管理，用户组织机构管理，用户角色管理，用户权限管理。

4.3.3.8.1 用户管理

用户的管理有以下几种方式：

- 令牌授权管理；
- 内部管理；

可以使用其中的一种或者同时使用多种用户管理模式。

4.3.3.8.2 令牌授权管理

将平台的用户与令牌进行绑定，既减轻了系统运维人员的负担，也使得最终用户无需记忆新的用户名和密码。同时避免了忘记密码，密码过期等问题所引起的不便和系统维护开销。也避免了新员工加盟和老员工调离时的系统维护负担。

对于特定 IP 的机器授予对应的操作权限，在每次操作前都会核对核发的令牌是否有对应的权限，否则拒绝操作。

令牌授权又分为动态令牌和静态令牌两类。

4.3.3.8.3 内部管理

内部管理用户名密码，需要逐条建立每个用户的信息，并设置其初始密码，为了系统的安全性考虑，对手工管理的密码设置如下约束：

- 第一次登录必须修改密码（首次登陆密码过期）。
- 每3个月更改一次密码（系统强制）（强制密码过期，密码不得重复使用）。
- 密码不得纪录在任何介质上，或者泄露给任何其他人。
- 3次失败 账户冻结 如需解冻，需要提出申请
- 对于停职，辞退，辞职的人员，运维人员在接到人力资源部通知后立即解除其对应应用系统的访问权限
- 系统有导出用户权限报表的功能，每半年对所有能访问核心系统的人员和权限进行查验。
- 密码过期后，禁止使用重复使用同一密码。

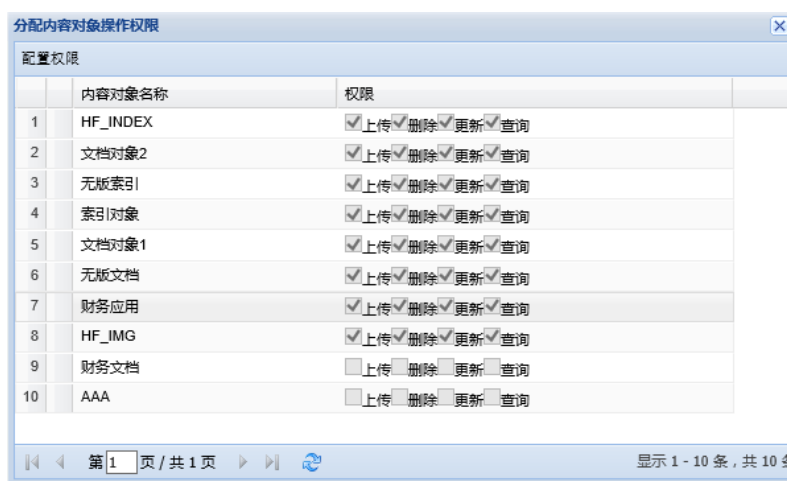
4.3.3.8.4 用户角色与权限

角色是一组权限的集合，有许多用户他们的角色是相同的，因此我们可以定义一些角色，让后批量的授予用户，可以大大方便用户权限的管理

某些情况下用户需要被单独的授予一些特定的权限，这时，用户可以直接获得某种操作的权限。

4.3.3.8.5 业务类型及缓存节点的权限控制

将用户权限控制到业务类型、缓存节点、文件类型，包括文件的增、删、改、查、打印、批注，这样在对接的业务系统中可以更方便的调用。在不同的业务处理环节，操作员能进行操作权限是不同的。



用户文档对象操作权限界面

4.3.3.9 监控系统

作为一个多层级部署的文档服务器，监控每个文档服务器的工作情况，可以更好的为系统运维保驾护航。监控系统提供了统一的页面，对每台文档服务器运维情况进行监控。

4.3.3.9.1 系统资源监控

对于每台文档服务器，监控系统资源的使用情况，方便系统运维，可以跨平台的监控系统 CPU、IO、磁盘使用率、网络使用率、内存使用率，应用服务器（WEB SPHERE、WEBLOGIC、TOMCAT 等）占用系统 CPU、内存。

4.3.3.10 客户端组件

SUNECM 提供了基于 WEB 应用的客户端，实现内容的捕获以及对内容的检索查询等功能，并将业务模型下属的文件类型以多层级的节点目录树的方式进行展示。并提供高级搜索及分类浏览

4.3.3.11 生命周期管理

按照文件从产生到消亡，从动态到静态的客观规律，为满足人们对影像不同阶段的不同访问要求，平台引入并完整实现了文件生命周期管理的概念。

4.3.3.11.1 生命周期阶段和划分标准

按照文件从产生到消亡，从动态到静态的客观规律，为满足人们对影像不同阶段的不同访问要求，平台引入并完整实现了文件生命周期管理的概念。

根据内容从产生到失效的规律，可以将内容的生命周期分为在线，离线的两个阶段。或者划分为缓存、在线、离线三个阶段（视业务期限规律和政策法规而定）。

按照文档的使用特征和使用周期来划分，文档可以分为活动和静态两类。其中，活动文档所附属的业务流程仍未结束，需要对其进行频繁的调阅并对内容进行修改，活动文档主要存储在缓存 SUNECM 中，其中缓存存储管理(文档管理器)的主要是会计流程业务系统这类业务量比较大的文档，SUNECM 中存储管理的主要是国结系统等这类业务量不大的文档；静态

文档相关的业务流程已经结束，其内容较少被调阅，也不需要进行修改。因此，在存储策略上，平台将活动文档保存在 SUNECD 中，而将静态的文档导出 SUNECD，进行脱机归档，以合理利用系统资源，保证内容管理系统足够高的整体性能。

文档生命周期阶段	主要业务流程是否完成	被访问的频率	要求的访问速度	是否还有修改可能	距离业务发生的日期
缓存	未完成	频繁	最快	有	3 个月内 (视业务种类)
在线	未完成	频繁	快	有	1 年内 (视业务种类)
离线	已完成	几乎没有	最少	无	视政策法规要求而定

不同生命周期阶段的判断标准

通过统一的生命周期管理配置，管理缓存阶段、SUNECD 存储阶段以及离线管理阶段的策略，各阶段可以按照统一的生命周期管理策略进行处理。

缓存阶段可以实现生命周期管理，即根据影像生成、处理、调阅、归档和销毁的过程，对影像缓存进行管理。影像缓存生命周期管理可以由 workflow 驱动，在 workflow 中预先将影像复制或迁移到最近的和和工作内容较多的缓存服务器上，流程中不同的客户端可以就近获取影像，以减少网络压力、同时提高响应速度。业务处理完成后，在 workflow 驱动下，将影像传递到 SUNECD 归档。

在线阶段，主要是 SUNECD 文档服务器存储阶段，主要的策略有归档策略、清理策略；

离线阶段，主要是对离线介质，比如说是光盘的管理。

4.3.3.11.2 缓存阶段的管理

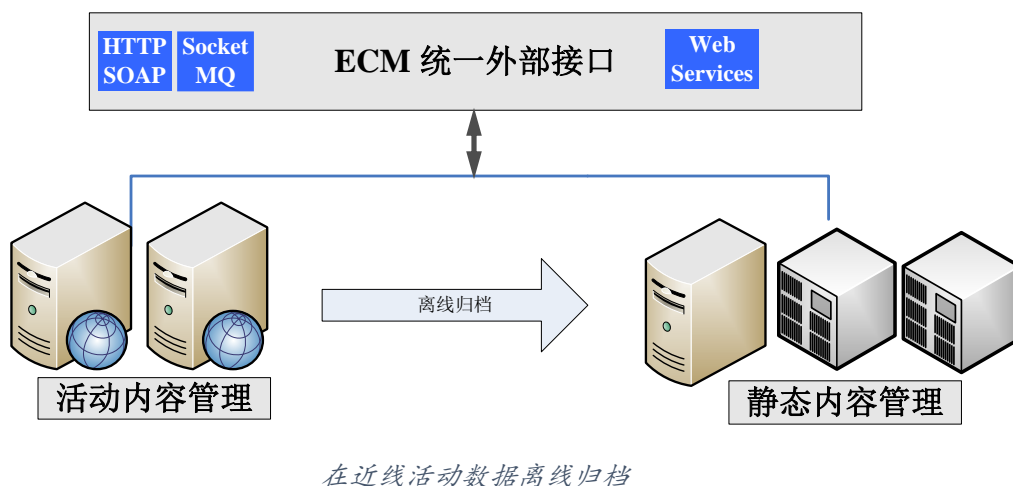
此阶段关于此影像的业务还在办理中，随时有可能对影像进行查询，因此要求平台能迅速响应用户的查询。必须使用最快速的存储设备，并尽快将非活动数据迁移到下级文档服务器。

缓存阶段内容生命周期管理主要是内容迁移归档和内容清理，由文档服务器实现，通过文件服务器的迁移策略的配置，并由任务调度系统来执行。也可以由业务系统或 workflow 系统来触发迁移归档的工作。

4.3.3.11.3 在线阶段的管理

此阶段关于此影像的业务还在办理中，这部分数据主要存储在 SUNECM 的统一文档服务器中，数据包括分支文档服务器归档上来的内容，以及事后监督等系统处理完成后归档的内容，这部分数据在 SUNECM 管理期间不会发生变化，还有各个业务系统在业务处理过程中随着业务处理要求存储到 SUNECM 中的内容，这部分数据随时有可能对影像进行查询和更改，因此要求平台能迅速响应用户的查询，

在线阶段内容生命周期管理主要是内容离线和清理，在 SUNECM 中的数据在业务处理完成后，可以考虑归档到离线系统中，SUNECM 中的数据根据业务在线时间要求进行管理，生命周期管理根据应用及时间配置在 SUNECM 或归档系统中进行查询。

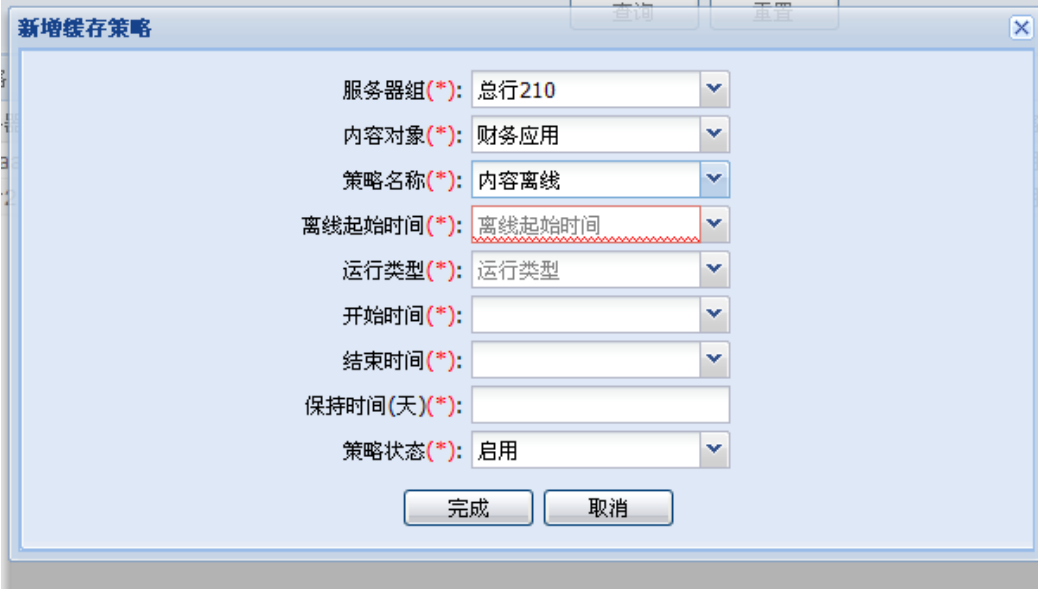


4.3.3.11.4 离线归档过程处理

- 离线归档程序

根据生命周期管理中的缓存策略管理，对不同的业务按照规则进行离线，主要过程如下：SUNECM 系统会在缓存策略中配置的时间启动离线策略，开始对 SUNECM 中某一业务模型下的所有符合条件的批次进行离线，将对应对象的批次及挂载文件导入离线管理系统，完成一笔批次的离线后，将该批次的状态设定为已离线状态。在完成一天的所有批次离线后，将该天的批次数，业务模型名，时间等信息记录入 SUNECM 的离线压缩表。然后再进行下一天的离线。直到完成所有需要离线的数据。

- 形成归档卷后，需要在离线系统数据库中记录卷名、卷路径及其基本信息，包括应用类型、时间区间等，便于离线数据查询和管理；
- 离线归档时在归档数据内容文件的同时归档索引文件及元数据文件，由于内容文件是打包形式存放的，需要通过索引文件及元数据文件的 XML 来解析，以便在查询时能够对打包文件进行处理，并通过 XML 能直接定位到指定的影像文件，提高效率；
- 归档目录结构及管理内容建议：
归档文件的根目录直接在离线服务端的配置文件进行配置。
- 离线策略配置页面如下：



新增缓存策略

服务器组(*)：总行210

内容对象(*)：财务应用

策略名称(*)：内容离线

离线起始时间(*)：离线起始时间

运行类型(*)：运行类型

开始时间(*)：

结束时间(*)：

保持时间(天)(*)：

策略状态(*)：启用

完成 取消

4.3.3.11.5 离线数据清理

对于离线成功的批次，将其记录状态置为离线成功状态，脱机成功的内容由离线数据清理排程调用。将所有离线作业控制表中状态为脱机成功的对象从 SUNCM 系统中永久删除。

对与离线失败的批次, 将其记录状态置为离线失败状态, 转由人工处理。

离线清理配置页面如下:

4.3.3.11.6 离线数据查询

客户端向 SunECMDM 发起查询请求时, 判断批次是否已离线, 已离线的则转发查询请求到离线系统, 解析得到的结果中的 URL, 转换成 ECMDM 的 URL, 然后返回给客户端。

当且仅当查询条件中包含业务发生日期时才到离线查询。

SunECMDM 或统一接入在接收到访问离线文件的 URL 时, 转发到离线服务器:

解析 URL, 如果 URL 中包含离线系统 URL 的 Key, 则将 URL 解析转换成 URLBean, 调用“离线文件访问 API”

4.3.3.11.7 离线阶段的管理

此阶段关于此影像的业务已经办理结束, 影像通过离线系统打包存放到相应目录下, 可以人工去做刻盘或清理操作。

4.4 业务应用接入过程

业务应用要使用统一内容管理接入平台服务, 在需求讨论阶段, 需要明确部

署方式、业务元数据、业务量、网络带宽、档案服务接口（总体情况，详细情况在系统设计阶段实现）、档案离线处理等内容，流程图示如下：

