第八届全国大学生信息安全竞赛

暨首届网络安全人才创新创业发展论坛

第八届全国大学生信息安全竞赛

暨首届网络安全人才创新创业发展论坛

数据安全创新解决之道及创新人才培养思考

单　位：天津神舟通用数据技术有限公司

演讲人：顾 云 苏

日　期：2015年8月2日

目录

Contents

目录

Contents

01 数据安全挑战

大数据成为国家和企业的核心资产

2012年瑞士达沃斯论坛上发布的《大数据大影响》报告称，数据已成为一种新的经济资产类别，就像货币或黄金一样。奥巴马政府已把“大数据”上升到国家战略层面，2012年3月，美国宣布投资2亿美元启动“大数据研究和发展计划”，借以增强收集海量数据、分析萃取信息的能力。美国政府认为，大数据是“未来的新石油”，一个国家拥有数据的规模、活性及解释运用的能力将成为综合国力的重要组成部分，**未来对数据的占有和控制甚至将成为继陆权、海权、空权之外国家的另一个核心资产**。

云计算和大数据带来的安全挑战

云安全联盟已经将不正确的数据和网络隔离列为云计算的首要安全威胁

- 大数据平台往往暴露在网络上，大数据的数据量大且相互关联，对于攻击者而言，数据汇集后一次攻击的收益被放大。近年来在互联网上发生的用户账号的信息失窃等连锁反应可以看出，大数据更容易吸引黑客，而且一旦遭受攻击，失窃的数据量也是巨大的。

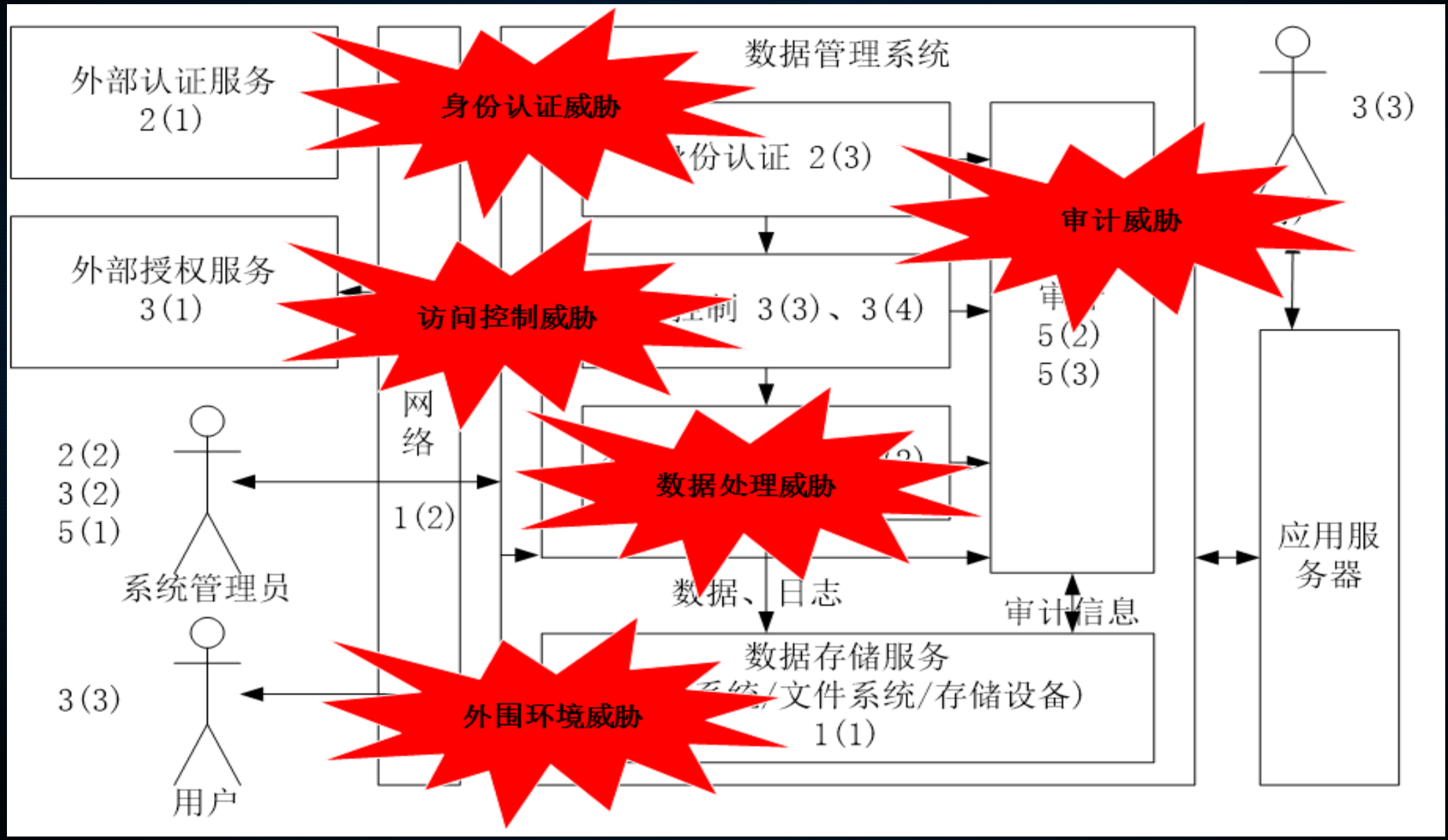
大数据加大隐私泄露风险

- 大量数据的汇集不可避免地加大了用户隐私泄露的风险。一方面，数据集中存储增加了泄露风险；而这些数据不被滥用，也成为人身安全的一部分；另一方面，一些敏感数据的所有权和使用权并没有明确界定，很多基于大数据的分析都未考虑到其中涉及的个体隐私问题。

新出现的大数据存储在安全性上存在很大隐患

- NoSQL数据存储为大数据存储提供了初步解决方案，但是NoSQL数据存储仍存在以下问题：缺少严格访问控制和隐私管理的SQL技术；NoSQL软件存在各种漏洞；三是由于NoSQL服务器软件没有内置足够的安全；API访问权限控制以及密钥生成、存储和管理方面的不足都可能造成数据泄漏

一个数据管理系统面临的安全威胁



目录

Contents

02

解决思路
与关键技术

安全数据访问的技术思路

- **高可信依旧来自源头上的解决方案**：从可信计算的角度在系统设计源头上避免引入安全问题，减少系统安全漏洞，增加隐通道和推理控制的计算复杂度
- **高安全手段在大数据环境中的增强**：采用强认证、强制访问控制、强审计、三权分立、数据传输加密和鉴别、数据库加密等技术，应对运行态安全问题
- **系统融合**：数据安全机制与云环境、操作系统、网络、应用系统的安全机制进一步融合，整体考虑数据安全问题
- **高性能技术保证**：在保证安全性的前提下，依靠精心的结构和算法设计，提高系统的运行效率
- **用大数据的方法设计新的防御手段**：采用大数据分析挖掘的技术手段，识别和定义新的攻击模式

云平台的大数据安全隔离威胁及应对

威胁类型	受害者	攻击者	发生可能性	后果	解决方案
一个租户部署的服务有漏洞或者配置错误	租户自己	其它租户或云外部攻击者	高	数据丢失或服务崩溃	检查代码bug，防火墙及IDS保护，设计多租户数据隔离方案
云基础平台无法防止side-channel攻击	所有租户	其它租户	中	拒绝服务以及有可能数据泄密	可信云技术，多租户数据隔离方案
一个租户的应用程序或内核有bug	租户自己	其它租户或云外部攻击者	中	数据丢失以及服务崩溃	及时更新，多租户数据隔离方案
云服务提供商自身是恶意攻击者	所有租户	云服务提供商雇员	低	所有租户数据及服务被攻破	员工背景审查，数据全程加密

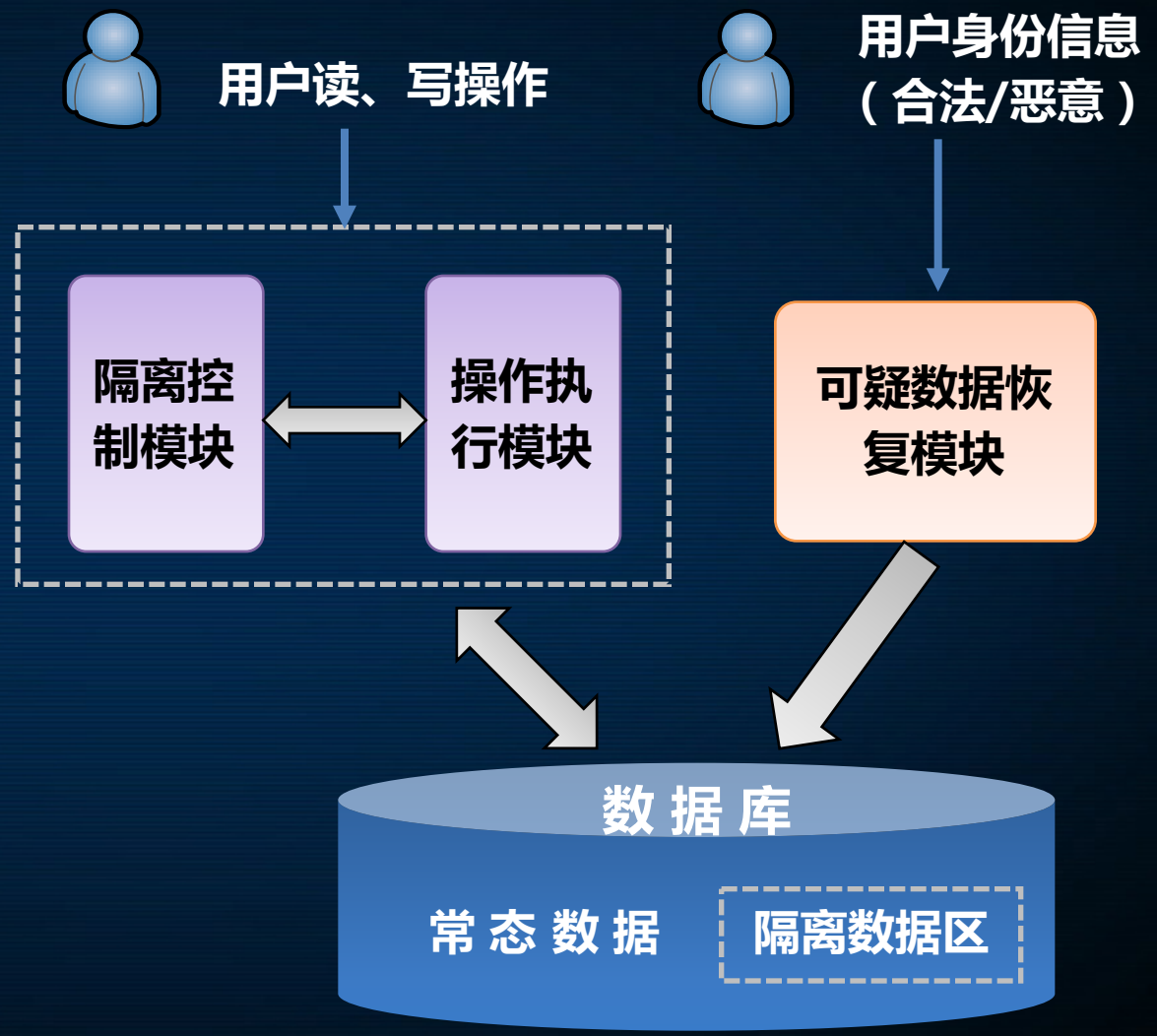
安全隔离系统框架

- 使用标记服务(Labeling Service)对数据进行全程标记
- 在用户处部署一个系统内核级信息流追踪组件
- 追踪所有在租户实例内进程和文件之间的信息流
- 在特权域中设计数据和网络隔离模块用于追踪与阻断；

神通安全数据库集群系统的安全隔离框架

可疑实体的在线隔离和恢复

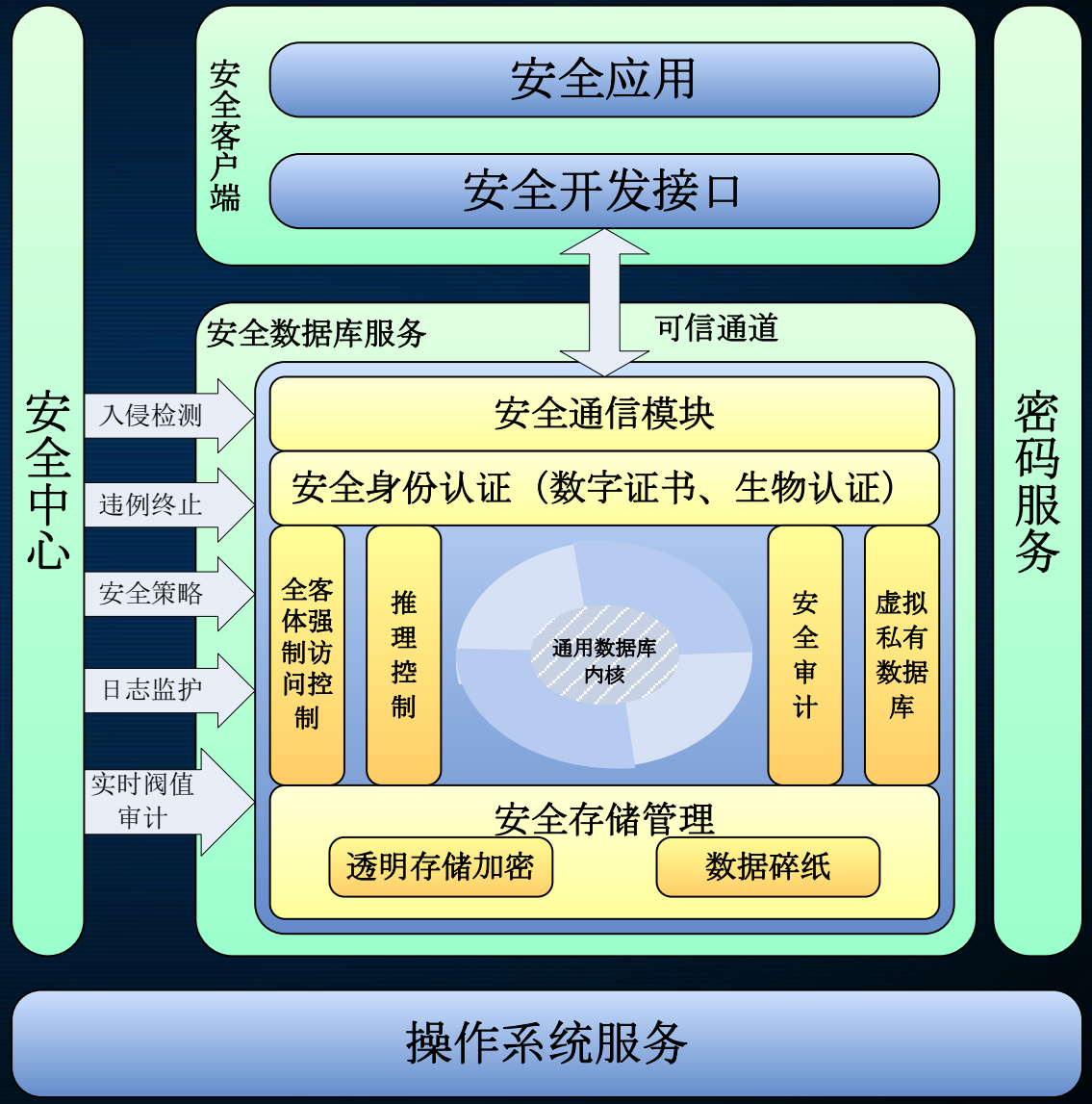
- 设计隔离和操作执行模块统一处理用户读写操作
- 由可疑数据恢复模块根据用户身份信息判断可疑用户和操作
- 设计隔离区存储可疑操作，可疑数据正常用户名不可见
- 当可疑用户确认为恶意用户时，对该用户在可疑期内的可疑数据进行动态实时恢复
- 当可疑用户确认为合法用户时，将该用户创建的隔离态数据到对应的常态数据



数据隔离和恢复操作流程

传统高安全技术手段的云化

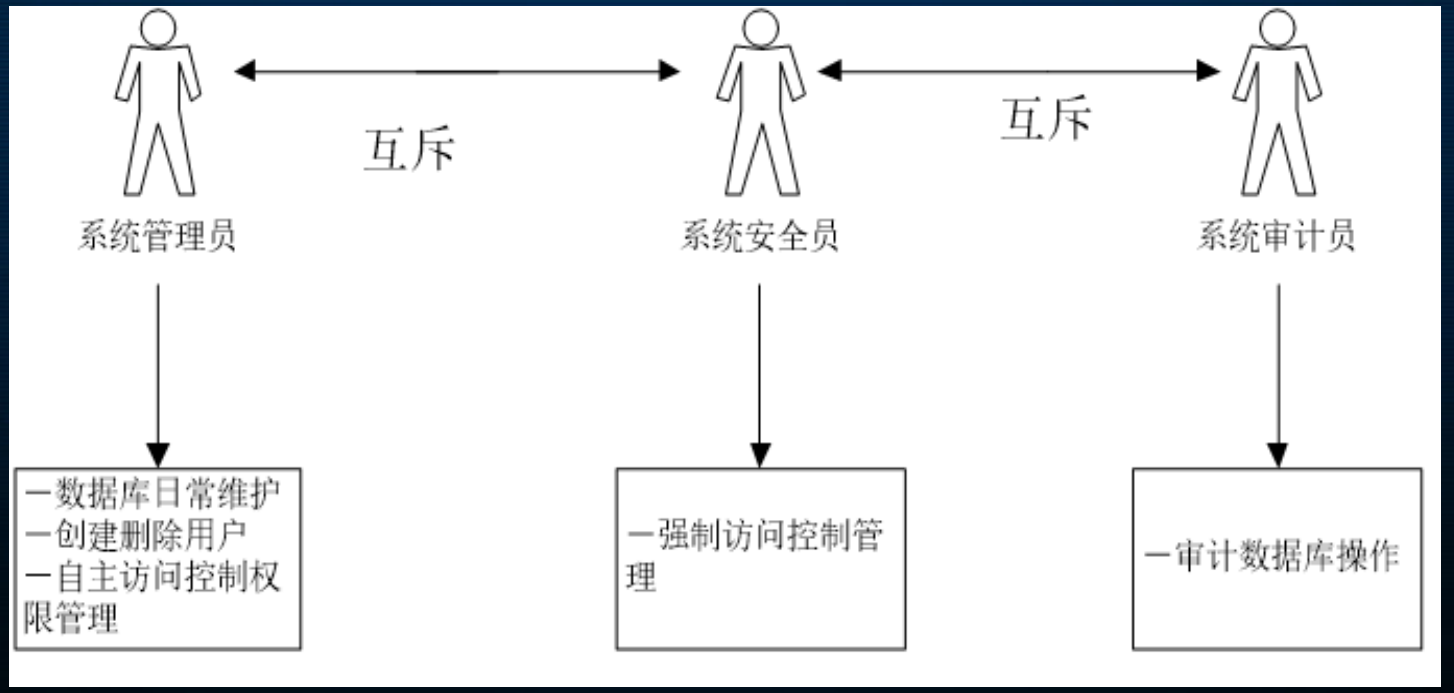
- GBT-20273四级安全等级
  - 全主客体强制访问控制
  - 多种认证方式支持（指纹、
    动态密码、数字证书）
  - 攻击检测和审计防护
  - 可信路径
  - 隐通道分析
  - 推理控制
  - 客体安全重用
  - 虚拟私有数据库
  - 日志监护



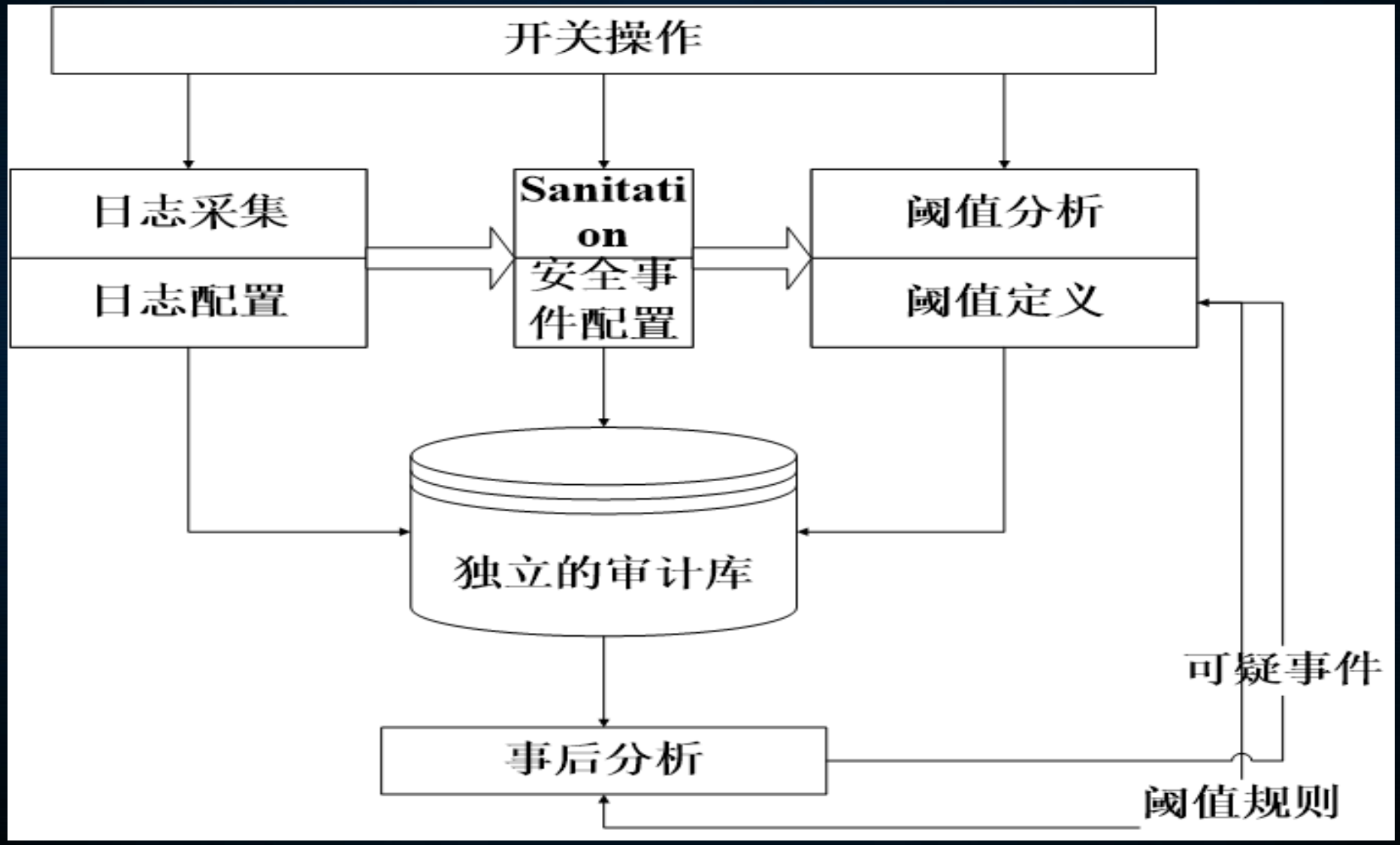
三权分立机制

三权分立的特点在于保证系统中自主访问控制、强制访问控制、审计三者的管理工作相互独立，互不干涉。在大数据环境下，三权分立技术依旧有效，但需要进行如下改进：

- 权限的云化管理
- 强制访问控制的自动化授权
- 对低价值密度和高价值密度数据的不同保护
- 分布式审计



独立的分布式审计


开关操作
日志采集
日志配置
Sanitation
安全事件配置
阈值分析
阈值定义
独立的审计库
事后分析
可疑事件
阈值规则

推理控制的应用

推理规则的挖掘

采用基于带状划分与一致集概念的挖掘方法，根据已有数据分析生成最小非平凡函数依赖集，实现了针对函数依赖关系的自动挖掘功能。

推理通道构造

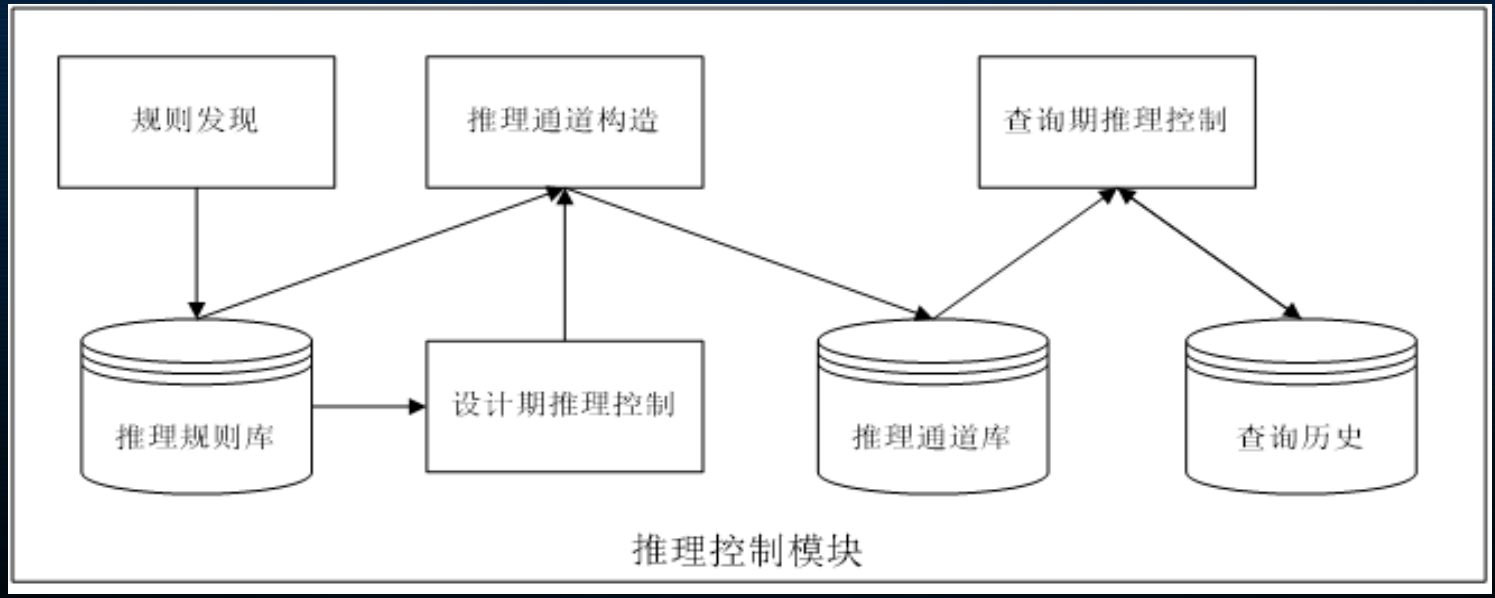
推理通道基于推理规则构造，由于数据依赖关系的性质不同，分别基于函数依赖关系和多值依赖关系构造推理通道，并根据上述两种类型的通道分别进行推理控制。

设计期推理控制

改进现有设计期推理控制算法，对高价值密度和低价值密度的数据进行分别设计。

运行期推理控制

基于Staddon提出的推理通道思想，提出了一种动态组合通道实例的方法，实现了针对数据依赖关系的动态推理控制。



目录

Contents

03

人才需求与培养

大数据安全人才的需求

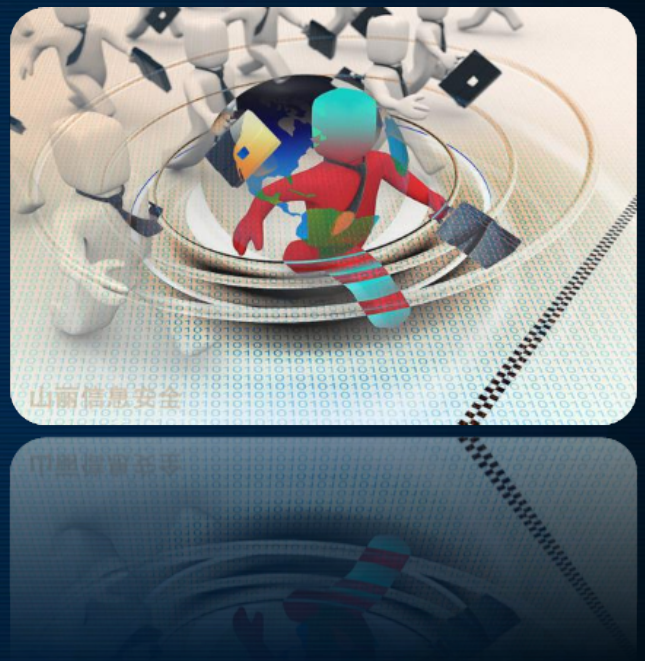
- **大数据安全规划人才**
  - 对大数据有着敏锐的大局观，发现关键问题所在
  - 对安全问题有着深刻理解，洞察数据价值并设计合理的安全措施
  - 能够制定完善的大数据管理和安全操作制度
  - 熟悉可信计算原理和具体技术，了解主要安全技术
  - 了解相关的大数据安全技术
- **大数据安全架构人才**
  - 能够设计完整的大数据系统安全架构，或者针对已有的大数据系统构架，增加关键的安全框架
  - 熟悉和了解可信计算原理和具体技术，了解主要安全技术
  - 熟悉和掌握各类安全软件的特点

安全隔离技术

- **大数据安全设计人员**
  - 熟悉大数据的特点和难度
  - 能够根据安全策略和要求，设计出相应的技术应对方案
  - 了解相关的大数据安全技术
- **大数据安全开发人员**
  - 较好的大数据处理经验，明晰大数据处理的难度
  - 优秀的大数据算法设计和编码能力
  - 熟悉可信计算和各类安全技术
- **大数据安全实施运维人员**
  - 理解大数据安全架构特点，理解安全制度的精髓
  - 理解相应的安全制度
  - 能够做出正确的决策和操作
  - 熟悉和掌握各类安全软件的特点

大数据安全人才的培养

- **先解决大数据人才的培养问题**
  - 我国目前大数据方面的人才仍然很缺乏
  - 大量的人员主要是使用国外开源体系
  - 无法理解核心技术，安全问题自然失控
- **在熟悉大数据的人才中培养大数据安全规划人才**
  - 必须熟悉和了解大数据特点
  - 必须熟悉和了解主流大数据架构体系
- **培养合格的技术设计开发运维人才**
  - 必须掌握和了解大数据的关键技术，掌握大数据平台架构核心
  - 进行安全技能的培训，明确安全在系统中的重要性



目录

Contents

04

总结

总结

- 数据是核心资产，大数据时代对数据的保护面临着很大的新挑战，培养我国的人才刻不容缓
  - “国家信息安全人才培养测评实验室”人才培训认证计划
- 挑战也是机遇
  - 大数据时代对国内外都是新事物，我国基本和国外处于同一条起跑线
  - 大数据安全技术和软件发展的良好机遇

祝您有个美好的一天

謝　　謝

第八届全国大学生信息安全竞赛暨首届网络安全人才创新创业发展论坛