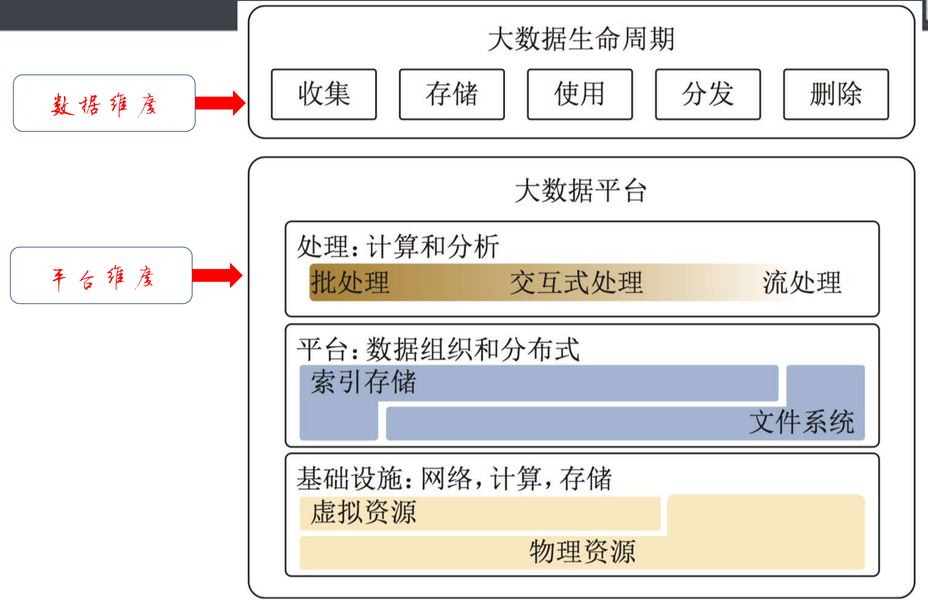

# 大数据安全标准及平台安全保护体系

演讲人：四川大学 陈兴蜀

# 大数据基本概念

## ■NIST大数据定义

- Variety：数据多样性，如不同领域、不同类型
- Velocity：数据增长速度，处理速度
- Volume：数据规模
- Variability：多变性，即规模、速率等特征易变



# 大数据相关法律法规

- 2012年《全国人大常委会加强网络信息保护的决定》；
- 2013年工信部发的《电信和互联网用户个人信息保护的规定》；
- 2014年《消费者权益保护法》里有关“个人信息保护”的规定；
- 2015年国务院发的《促进大数据发展行动纲要》里有“在发展大数据的时候要保障个人隐私和信息安全”；
- 2016年《“十三五”规划》里也提出了“实施国家大数据战略”，大数据变成一个国家的发展战略是很重要的一件事情，全国各地都在成立大数据管理局，说明国家非常重视大数据的安全和大数据的发展；
- 2016年《网络安全法》里对数据安全提的最多；最近网信办的《国家网络空间安全战略》里面也提到了“数据的跨境流动、个人信息的保护和大数据的发展”

| 序号 | 法律法规和部门规章 | 发布/生效时间 | 备注 |
|---|---|---|---|
| **一、美国** | | | |
| 1 | 《隐私盾协议》（替代《安全港协议》） | 2016 年发布 | 通用法律 |
| 2 | 《加州在线隐私保护法案》 | 2014 年生效 | 州法律 |
| 3 | 《联邦隐私法案》 | 2014 年发布 | 通用法律 |
| 4 | 《数字问责和透明法案》（FFATA） | 2014 年发布 | 部门规章 |
| 5 | 《数字政府战略》 | 2012 年发布 | 通用法律 |
| 6 | 《开放政府指令》 | 2009 年发布 | 通用法律 |
| 7 | 《加州安全违约告知法律》 | 2002 年生效 | 州法律 |
| 8 | 《金融服务现代化法案》（GLBA） | 1999 年发布 | 部门规章 |
| 9 | 《健康保险携带和责任法案》（HIPAA） | 1996 年发布 | 部门规章 |
| 10 | 《联邦贸易委员会法案》（FTCAct） | 1914 年发布 | 部门规章 |
| **二、欧盟** | | | |
| 1 | 《通用数据保护规则》（GDPR） | 2016 年发布 | 通用法律 |
| 2 | 《欧盟数据留存指令》 | 2006 年发布 | 通用法律 |
| 3 | 《隐私与电子通讯指令》 | 2002 年发布 | 通用法律 |
| 4 | 《欧盟数据保护指令》 | 1995 年发布 | 通用法律 |
| **三、澳大利亚** | | | |
| 1 | 《电信法案》 | 1997 年发布 | 部门规章 |
| 2 | 《联邦隐私法案》 | 1988 年发布 | 通用法律 |
| **四、俄罗斯** | | | |
| 1 | 俄罗斯联邦法律第 152-FZ 条中 2006 年个人数据相关内容（PersonalDataProtectionAct，个人数据保护法案） | 2015 年发布 | 通用法律 |
| 2 | 俄罗斯联邦法律第 149－FZ 条 2006 年信息、信息技术和数据保护相关内容（DataProtectionAct，数据保护法案） | 2006 年发布 | 通用法律 |
| 3 | 《斯特拉斯堡公约》 | 2005 年发布 | 通用法律 |
| **五、新加坡** | | | |
| 1 | 《个人数据保护法令》(PDPA) | 2012 年发布 | 通用法律 |

# 大数据相关标准

•ISO/IEC 隐私保护相关标准

## 已发布

- 29100:2011 《隐私保护框架》
- 29101:2013 《隐私保护体系结构框架》
- 29190:2015《隐私保护能力评估模型》
- 29191:2012《部分匿名、部分不可链接鉴别要求》
- 27018:2014《PII处理者在公有云中保护PII的实践指南》

## 即将发布

- 29134 《隐私影响评估指南》
- 29151 《可识别个人信息保护实践指南》

## 工作草案

- 29184 《在线隐私通知和准许指南》
- 27550 《隐私保护工程》
- 27551《对ISO/IEC 27001在隐私保护管理方面的增强要求》

# 大数据相关标准

## •ISO/IEC 大数据工作组（WG9）

### 工作草案

- 20546 《大数据概述和词汇》
- 20547 《大数据参考架构》
  - 20547-1：框架和应用过程
  - 20547-2：用例和导出需求
  - 20547-3：参考架构
  - 20547-4：安全与隐私保护
  - 20547-5：标准路线图

## •ITU-T：国际电信联盟 电信标准化部门

- Y.3600： 基于云计算的要求和能力
  Big data – Cloud computing based requirements and capabilities
- 大数据交换框架和要求
  Big data exchange framework and requirements
- 大数据即服务的功能架构
  Functional architecture of Big data as a Service
- 大数据驱动的网络要求
  Requirements of big data-driven networking
- 基于大数据驱动的网络DPI框架
  Framework of big data driven networking based on DPI
- 大数据环境下深度包检测机制
  Mechanism of deep packet inspection applied in networking big data context
- 大数据 – 数据溯源的要求
  Big data - Requirements for data provenance

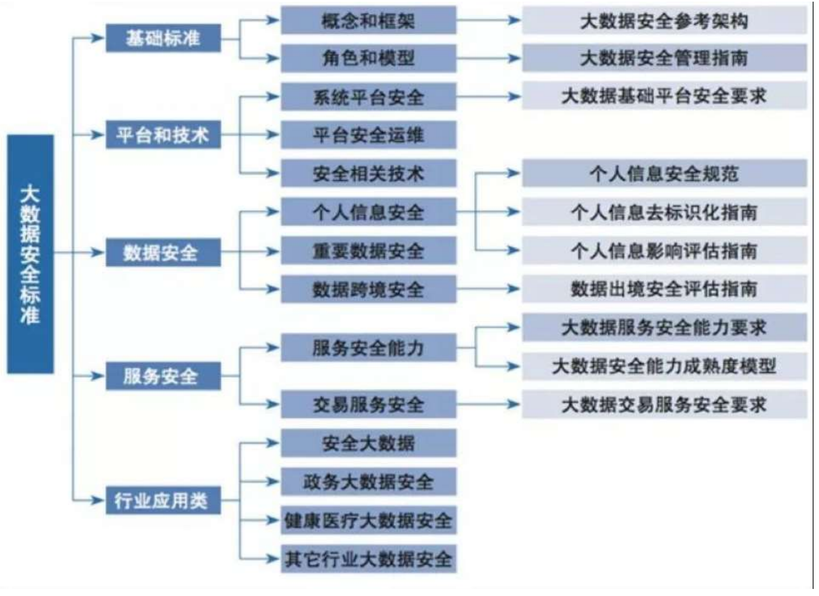
# 大数据相关标准

- NIST：美国国家标准与技术研究院
  - 2015年出版SP 1500 《NIST大数据互操作框架》系列标准

- SP 1500-1：卷一：定义
  NIST Big Data Interoperability Framework: Volume 1, Definitions
- SP 1500-2：卷二：大数据分类法
  NIST Big Data Interoperability Framework: Volume 2, Big data Taxonomies
- SP 1500-3：用例和一般要求
  NIST Big Data Interoperability Framework: Volume 3, Cases and General Requirements
- SP 1500-4：安全和隐私保护
  NIST Big Data Interoperability Framework: Volume 4, Security and Privacy
- SP 1500-5：架构调研白皮书
  NIST Big Data Interoperability Framework: Volume 5, Architectures White paper Survey
- SP 1500-6：参考架构
  NIST Big Data Interoperability Framework: Volume 6, Reference Architecture
- SP 1500-7：标准路线图
  NIST Big Data Interoperability Framework: Volume 7, Standards Roadmap
- SP 1500-8：大数据参考架构接口(草案)
- SP 1500-9：大数据采用情况及技术现代化（草案）

# 大数据相关标准

- 我国大数据安全标准体系



## ■ 标准范围

■本标准为组织的大数据安全管理提供指导，本标准提出了大数据安全管理基本原则，从大数据安全需求、数据分类分级、大数据活动的安全要求、评估大数据安全风险等方面，指导组织针对大数据的特点开展数据保护的管理工作。

■本标准适用于所有控制和处理大数据的组织，包括企业、事业单位、政府部门等等，也适用于第三方机构对组织的数据安全管理能力进行评估。

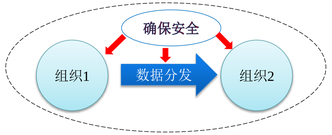
## ■ 编制思路

■范围限制为拥有和使用大数据的组织

- ■欧盟数据保护法案 -> 一般数据保护条例
- ■欧盟-美国隐私盾
- ■加拿大《个人信息保护和电子文档方案（PIPEDA）》

■突出大数据安全管理的特殊内容。

- ■大数据安全需求特殊考虑内容
- ■大数据活动及要求
- ■风险评估特殊内容等

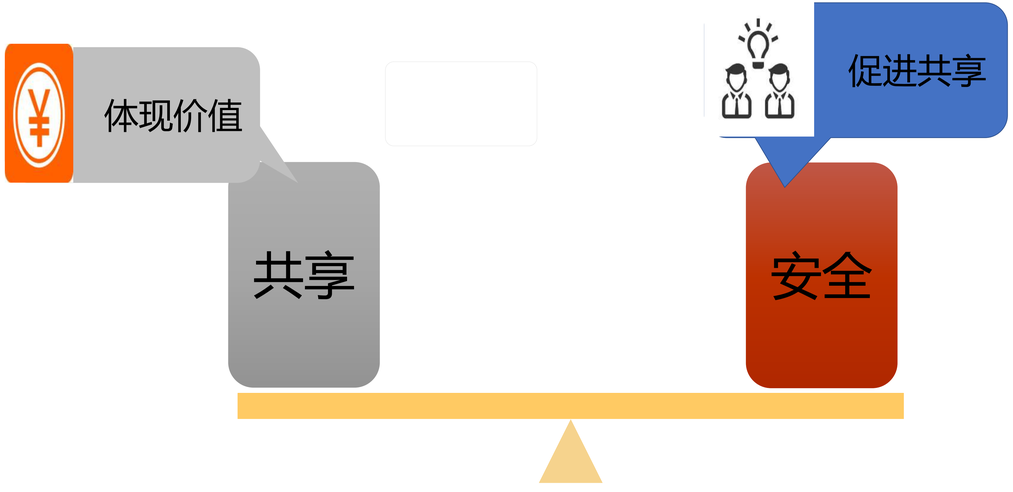


## ■ 安全目标

- ➢满足个人信息保护和数据保护的法律法规、标准等要求；
- ➢满足业务相关方的数据保护要求；
- ➢通过技术和管理手段，保证自身控制和管理的数据安全风险可控。

## ■主要内容

- ➢明确数据安全需求；
- ➢数据分类分级；
- ➢明确大数据活动安全要求；
- ➢评估大数据安全风险。


体现价值
共享
促进共享
安全

# 大数据平台的安全需求

- ■边界保护
- ■细粒度访问控制
- ■数据保护
- ■审计与溯源

目标：
- ◆ 进不来！
- ◆ 拿不走！
- ◆ 看不懂！
- ◆ 可追溯！

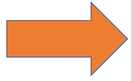
# 大数据平台的安全需求

**•边界保护**

•细粒度访问控制

•数据保护

•审计与溯源

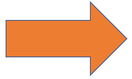


•暴露大数据平台细节，增加被攻击的风险

•用户账号管理不集中，增加管理成本

# 大数据平台的安全需求

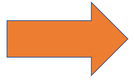
- 边界保护
- **细粒度访问控制**
- 数据保护
- 审计与溯源



- 主体集合构成复杂
- 多源异构数据导致访问控制策略中客体描述困难
- 访问控制场景复杂多变
- 开源大数据软件访问控制能力参差不齐

# 大数据平台的安全需求

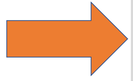
- 边界保护
- 细粒度访问控制
- **数据保护**
- 审计与溯源



- 不同敏感程度数据共同存储导致敏感信息泄露
- 数据关联分析导致敏感信息/隐私泄露
- 数据发布、交易等导致敏感信息/隐私泄露

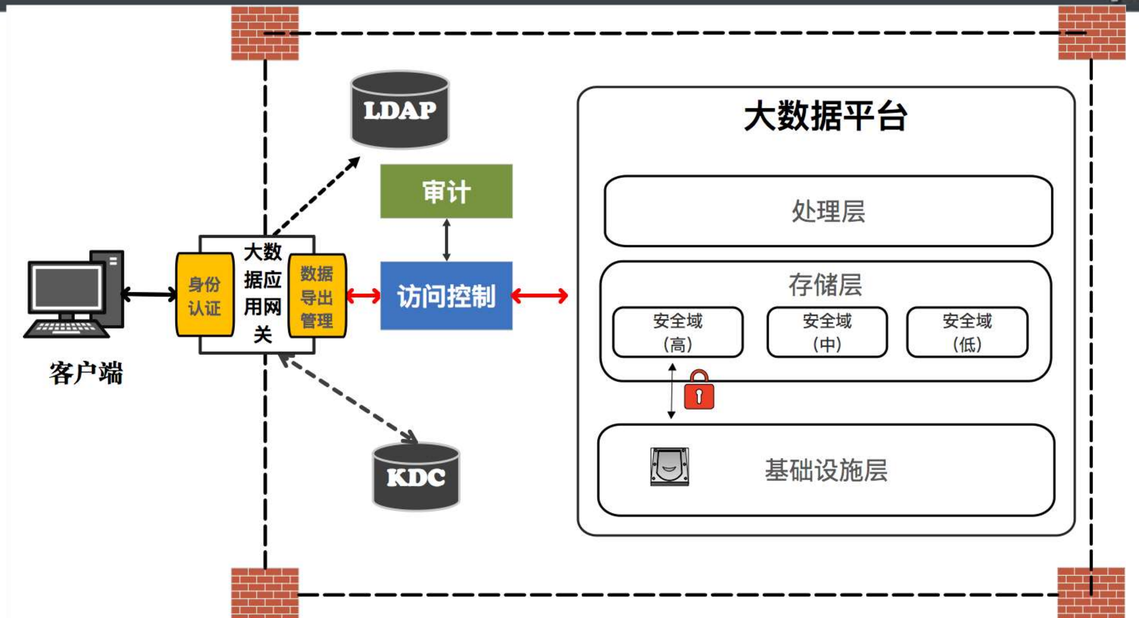
# 大数据平台的安全需求

- 边界保护
- 细粒度访问控制
- 数据保护
- **审计与溯源**



- 无法及时发现非法数据访问行为
- 大数据组件众多，增加审计与溯源难度

# 大数据平台安全解决思路



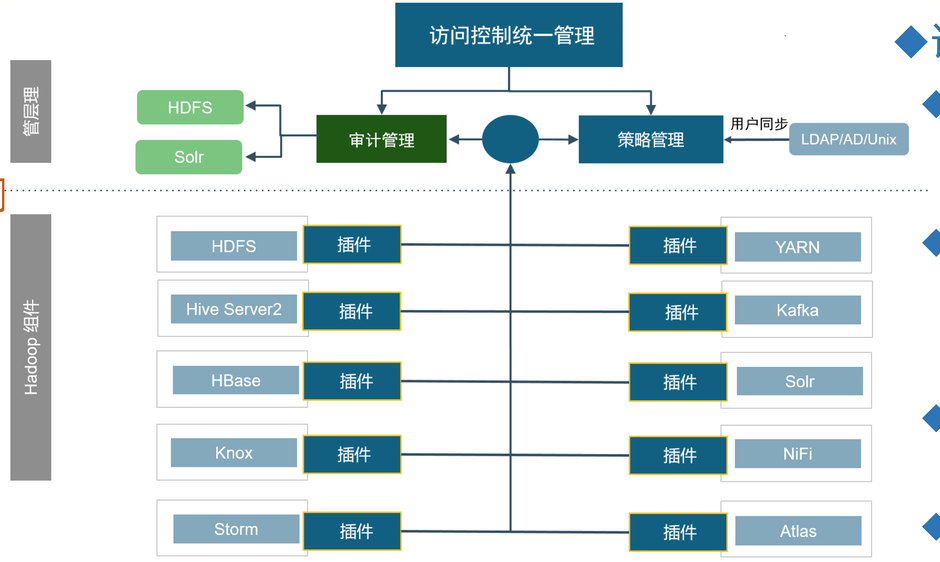
# 细粒度访问控制

■统一访问控制

■访问控制策略统一管理

■解决不同大数据服务组件访问控制能力参差不齐的问题

■提供全面数据访问日志




◆访问控制模型

◆世系数据相关的访问控制模型

- 将数据客体的状态作为访问控制判定依据，满足客体在不同状态下具有不同的访问权限。

◆半/非结构化数据的访问控制模型

◆针对图数据，基于关系的访问控制模型能更加直观地描述了访问控制需求。

◆针对文本数据，基于内容的访问控制模型能对文本数据客体进行细粒度描述，避免繁重的授权管理。

◆风险访问控制模型

- 通过评估访问行为风险，动态调整访问权限，及时发现数据窃取行为。

◆多种访问控制模型结合使用

# 数据保护：脱敏和隐私保护

■无损数据脱敏算法，保证脱敏后数据类型、数据关系、数据统计特性的一致性

■ 面向数据发布的隐私保护

■k-anonymity、l-diversity，t-closeness、差分隐私

■ 面向数据挖掘的隐私保护

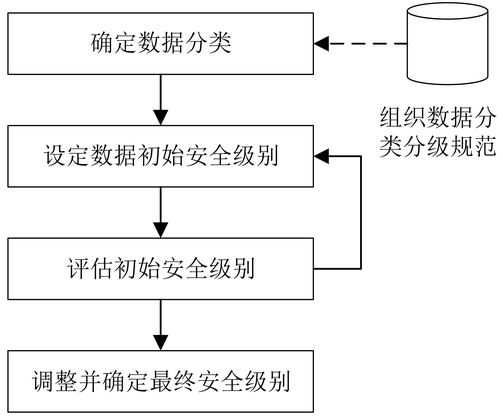
■数据扰动

■密文计算：同态加密、可搜索加密

■隐私保护的数据聚类、分类算法

# 数据保护：基于安全域的数据隔离保护

- 安全域构建依据：数据安全级别
- 访问控制模型：BLP模型
- 不同域采用不同安全措施
  - 高敏感级别数据加密存储



# 边界保护

- 通过使用大数据应用网关实现
  - 用户首先通过应用网关认证与授权，获得大数据服务的使用权限
  - 应用网关使用代理身份执行大数据操作，如访问数据，提交作业等
  - 应用网关对返回结果进行安全风险评估，满足要求则放行
- 优势
  - 隐藏大数据集群信息，降低安全风险
  - 结合网络层网关等其他安全措施，降低数据泄露风险
  - 用户集中管理
  - 提升数据共享安全


# THANKS