第六讲 数据安全

白杨 alicepub@163.com

第六讲 数据安全

① 数据安全概述

② 数据安全与传统安全的关系

③ 数据安全要素

④ 数据安全能力成熟度模型

⑤ 数据安全防护架构

⑥ 数据安全防护关键技术

数据安全概述

数据爆炸增长

数据被增列为新的关键生产要素

加快推动数据要素流通

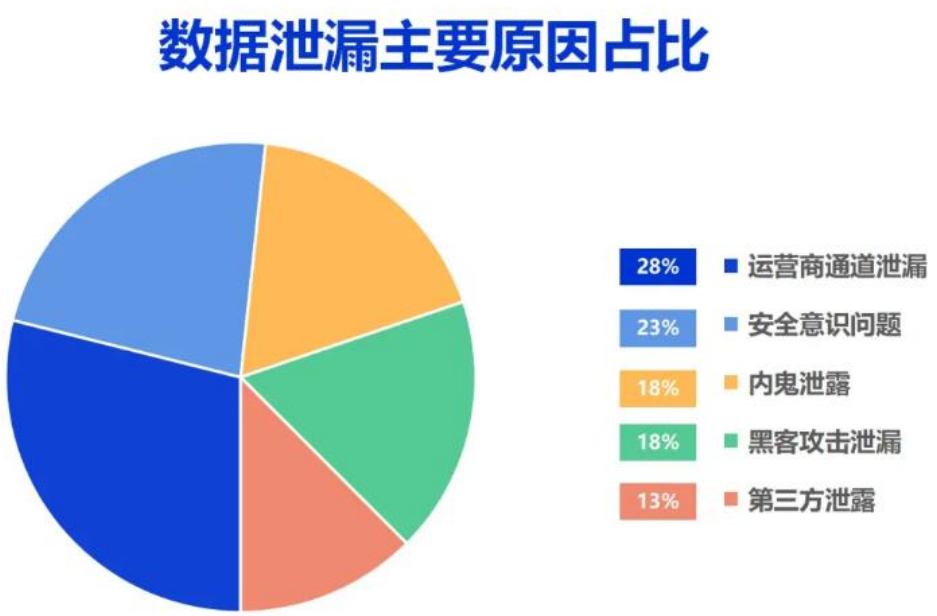
统筹推进数字中国、数字经济、数字社会规划和建设

数据安全概述

随着科技的发展， 数据遭到泄漏越来越普遍。2011-2016年间，多家电商平台大量用户数据外泄，2016年美国大选中希拉里的“邮件门”事件和韩国总统“闺蜜门”事件等比比皆是。据科学调查，70%的数据泄漏由内部威胁造成，每次经济损失达千万美元。

数据安全概述

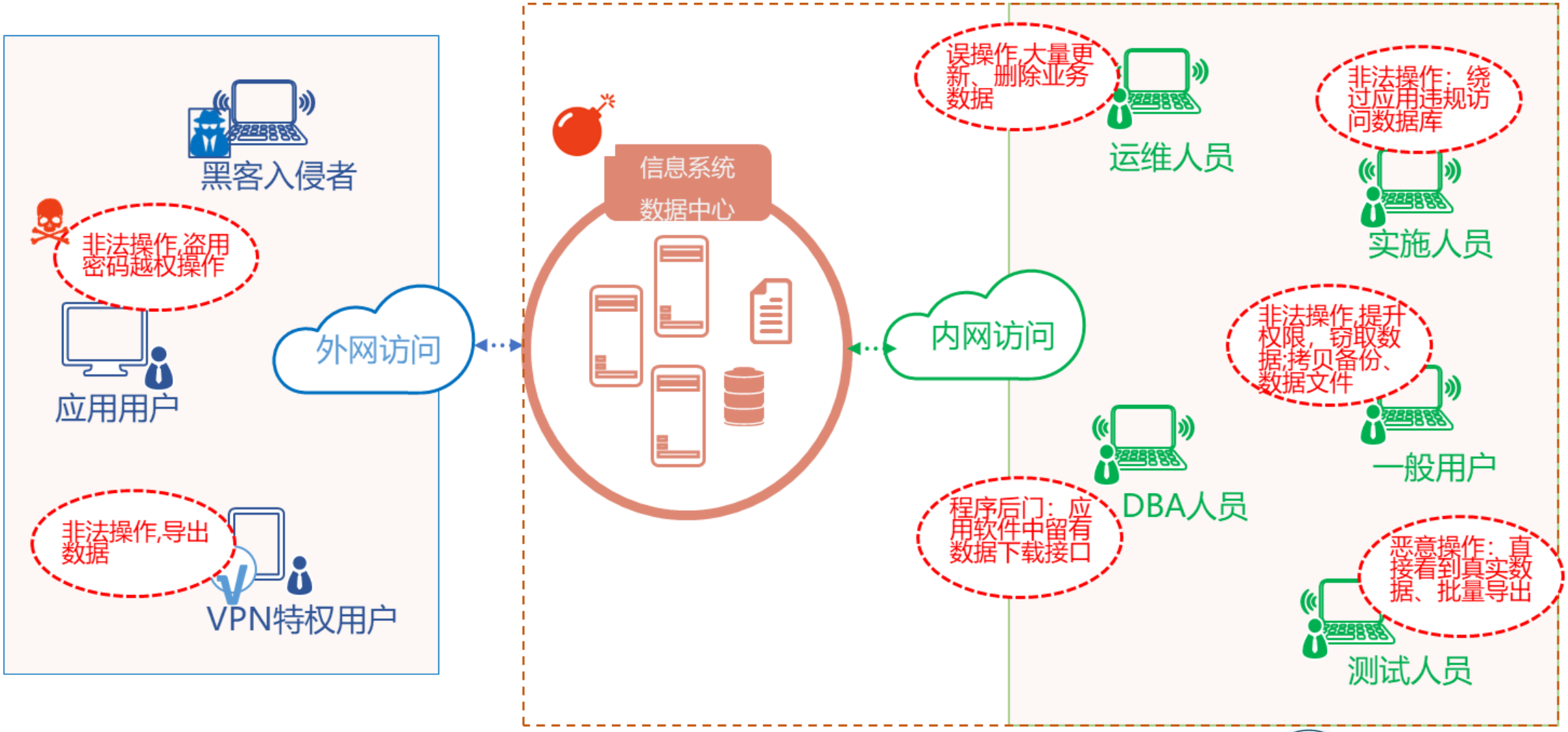
据统计，导致数据泄漏的主要原因：**黑客网络攻击、木马、病毒窃取、设备丢失或被盗、使用管理**不当等。也就是说数据从**数据采集、存储、传输、处理、交换、销毁数据全生命周期**管理过程中都会遇到各种威胁。



2022年数据泄露的主要原因及占比

2022年数据泄露事件月度数量趋势

数据安全概述



数据安全风险

数据安全概述

通过立法来加强数据安全重视：美国立法情况

- 特点：美国互联网监管体系主要包括：立法、司法和行政三大领域和联邦与州两个层次；涉及较为全面，既有针对互联网宏观整体规范，也有围观具体规定，其中包括行业进入规则、电话通信规则、**数据保护规则、消费者保护规则、版权保护规则、反欺诈与误传法规等方面**，多达130多部。

《联邦计算机系统保护法案》1977

《伪造网络信息存取手段及计算机欺诈与滥用法》1981

《电子通信隐私法》1986 《公共网络安全法》1997

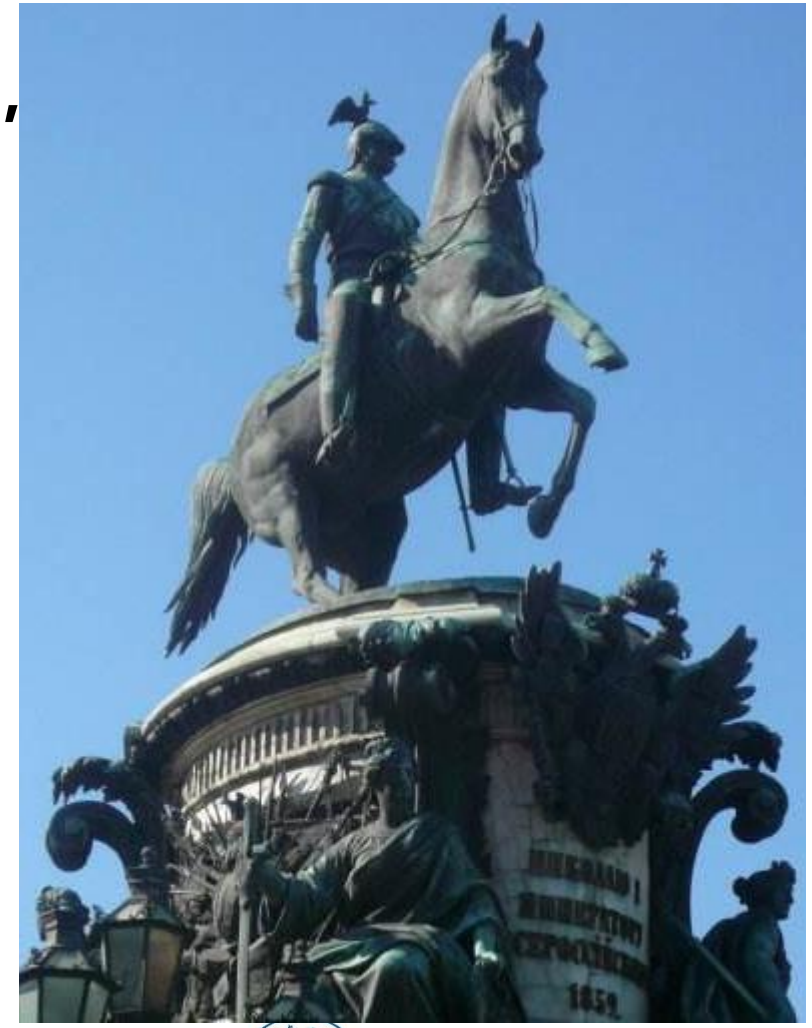
《联邦互联网隐私保护暂行条例》1997

《电子政务法》、《儿童在线隐私法》等

数据安全概述

通过立法来加强数据安全重视：俄罗斯立法情况

- 特点：网络信息安全法制开始于20世纪90年代，到21世纪，俄罗斯形成了比较完善的网络信息安全法制体系
- 《联邦信息、信息化和信息保护法》1997
- 《俄罗斯国家安全构想》1997 （加强个人信息保护）
- 《国家信息安全学说》2000 （由普京批准发布）
- 《电子数字签名法》2001
- 《保护青少年免受对其健康和发展有害信息干扰法》2001
- 《信息保护设备认证法》《电子公文法》《国际信息交易法》《电子合同法》等



数据安全概述

通过立法来加强数据安全重视：欧洲立法情况

- 特点：欧盟在网络安全体系建设方面成效显著。欧盟网络安全体系主要包含三大部分，**一是立法，二是战略，三是实践。**立法体系包含决议、指令、建议、条例等，战略体系包含长期战略与短期战略，实践则包含机构建设、培训、合作演练等多项内容。
- 《有关数据库法律保护的指令》1992 德国
- 《多媒体法》《数字签名法》《信息和通信服务规范》《著作权》《网络服务提供者责任法》1997
- 《关于计算机犯罪的协定》《欧盟电子签名指令》1999
- 《网络刑事公约》2001 《联邦数据保护法》2002….《一般数据保护条例》2018

数据安全概述

通过立法来加强数据安全重视：一般数据保护条例

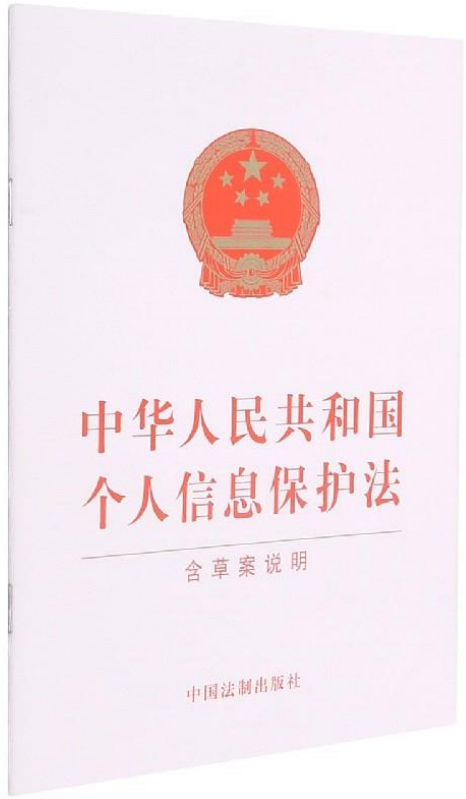
一般数据保护条例(General Data Protection Regulation，简称GDPR)，适用于**数据控制者或者数据处理者**在欧盟境内有分支；或者向欧盟境内的**数据主体提供商品服务**，或者**监控欧盟境内数据主体的行为**。

关注：

- 个人隐私保护权利（含儿童）
- 数据主体信息的获取与删除
- 第三方合作方的合规遵从
- 信息泄露后的应急处置与用户通告
- 数据保护官（具备法律和实践知识）
- ……

数据安全概述

通过立法来加强数据安全重视：国内立法



第六讲 数据安全

① 数据安全概述

② 数据安全与传统安全的关系

③ 数据安全要素

④ 数据安全能力成熟度模型

⑤ 数据安全防护架构

⑥ 数据安全防护关键技术

数据安全与传统安全的关系

计算机系统安全

为数据处理系统建立和采用的技术和管理的安全保护，保护计算机硬件、软件和数据不因偶然和恶意的原因遭到破坏、更改和泄露。

计算机网络安全

通过采用各种技术和管理措施，使网络系统正常运行，从而确保网络数据的可用性、完整性和机密性。

数据安全，是指通过采取必要措施，确保数据处于有效保护和合法利用的状态，以及具备保障持续安全状态的能力。–《中华人民共和国数据安全法》

第六讲 数据安全

① 数据安全概述

② 数据安全与传统安全的关系

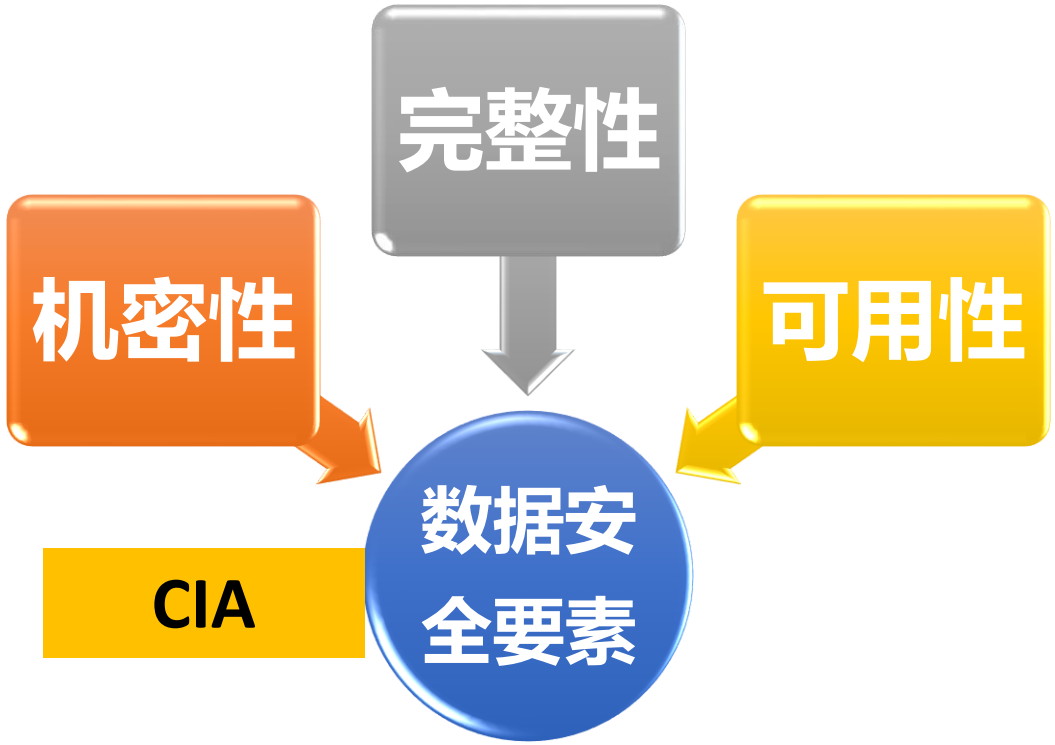
③ 数据安全要素

④ 数据安全能力成熟度模型

⑤ 数据安全防护架构

⑥ 数据安全防护关键技术

数据安全要素



特性含义

数据机密性，是指个人或团体的信息不为其他不应获得者获得，即**具有一定保密程度的数据只能让有权读到或更改的人读到和更改**。**数据完整性**指在传输、存储信息或数据的过程中，确保信息或数据不**被未授权的篡改或在篡改后能够被迅速发现**。数据**可用性**，使用者在浏览的过程中不会产生压力或感到挫折，对于该数据的合法拥有和使用者，在他们**需要这些数据的任何时候，都应该保障他们能够及时得到所需要的数据**。

第六讲 数据安全

① 数据安全概述

② 数据安全与传统安全的关系

③ 数据安全要素

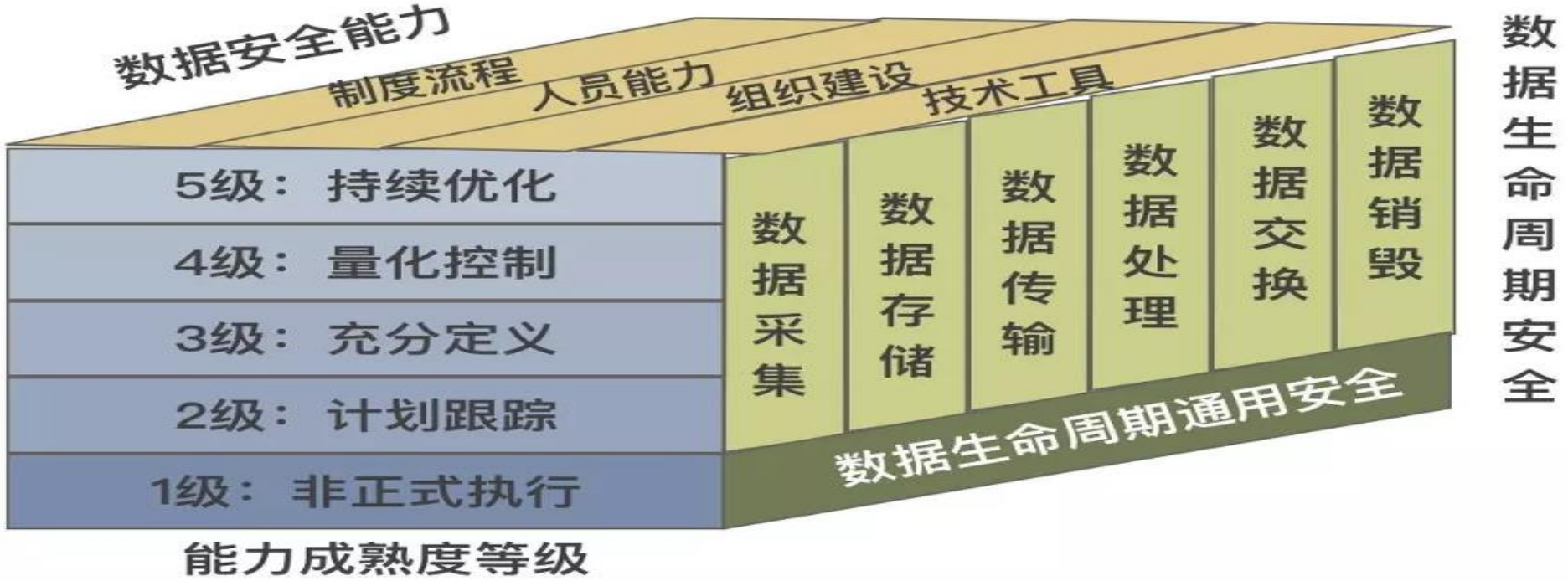
④ 数据安全能力成熟度模型

⑤ 数据安全防护架构

⑥ 数据安全防护关键技术

数据安全能力成熟度模型

数据安全能力成熟度模型（Data Security capability Maturity Model，DSMM）的缩写，中文名为数据安全能力成熟度模型。是以2019-08-30 发布，2020-03-01 实施的GB/T 37988-2019《信息安全技术数据安全能力成熟度模型》为依据的数据安全保护体系。



第六讲 数据安全

① 数据安全概述

② 数据安全与传统安全的关系

③ 数据安全要素

④ 数据安全能力成熟度模型

⑤ 数据安全防护架构

⑥ 数据安全防护关键技术

数据安全防护架构

数据运维管理安全

敏感数据监控

数据安全风险评估

数据溯源追踪

数据安全态势分析

数据安全治理

数据采集安全
- 数据源可信
- 数据接入控制
- 数据内容安全合规

数据传输安全
- 数据防篡改
- 数据防窃取
- 数据防伪造

数据存储安全
- 数据存储合规
- 数据容灾备份
- 数据库/文件安全

数据处理安全
- 数据数据脱敏
- 数据分析安全
- 数据导入导出安全

数据交换安全
- 数据共享安全
- 数据发布安全
- 数据接口安全

数据销毁安全
- 剩余信息保护
- 数据销毁合规

数据安全分级防护

数据分级分类标识与管理

数据加密

数据脱敏

身份鉴别与数据访问控制

数据水印

数据安全基础服务

数据安全标准规范

第六讲 数据安全

① 数据安全概述

② 数据安全与传统安全的关系

③ 数据安全要素

④ 数据安全能力成熟度模型

⑤ 数据安全防护架构

⑥ 数据安全防护关键技术

数据安全防护关键技术

（1）数据加密

加密就是对明文（可读懂的信息）进行翻译，使用不同的算法对明文以代码形式（密码）实施加密。此过程的逆过程称为解密，即将该编码信息转化为明文的过程。

一般我们在保存或传输过程会将数据本身先进行加密。**加密的基本功能**包括:

1. 防止不速之客查看机密的数据文件;
2. 防止机密数据被泄露或篡改;
3. 防止特权用户(如系统管理员)查看私人数据文件;
4. 使入侵者不能轻易地查找一个系统的文件。

数据安全防护关键技术

（1）数据加密

对称加密：加、 解密用同一个密钥，速度快，但密钥保存要格外注意。常用对称加密算法有DES、3DES、AES、IDEA等。安全级别较高的是AES（高级加密标准）。

非对称加密：加、解密需由一对密钥共同完成：公钥和私钥。若是公钥加密，必须由私钥解密，反之亦然。 这里提醒一下：私钥是私有的，不能公开，公钥可以告知他人。在应用时，公钥加密---〉私钥解密，是为了实现数据的机密性；而私钥加密----〉公钥机密，是为了操作的不可否认性（数字签名）。常用的非对称加密算法有RSA和DSA。

Hash(哈希)算法一般用在需要认证的环境下的身份确认或不考虑数据的还原的加密。因为Hash是一种单向散列算法，只能由一种状态变为另一种状态而不可逆。常用的是MD5算法和SHA算法。

数据安全防护关键技术

(2) 数据脱敏

姓名	社保卡号码	住址
张三	12345678	科华北路1号
李四	87654321	科华北路2号

结构化数据

文本数据

多媒体数据

数据全生命周期安全公有安全防护技术，有效防止敏感信息泄露，实现数据内容安全防护

- 自然语言处理
- 数据挖掘
- 机器学习

数据安全防护关键技术

（2）数据脱敏-方法

仿真：是根据敏感数据的原始内容生成符合原始数据编码和校验规则的新数据，使用相同含义的数据替换原有的敏感数据，例如姓名脱敏后仍然为有意义的姓名，住址脱敏后仍然为住址。仿真算法能够保证脱敏后数据的业务属性和关联关系，从而具备较好的可用性。

数据替换：用某种规律字符对敏感内容进行替换，从而破坏数据的可读性，并不保留原有语义和格式，例如特殊字符、随机字符、固定值字符等。

加密：通过加密算法（包括国密算法）进行加密。例如Hash（密码算法）算法是指对于完整的数据进行Hash加密，使数据不可读。

数据截取：数据截取术是指对原始数据选取部分内容进行截断。

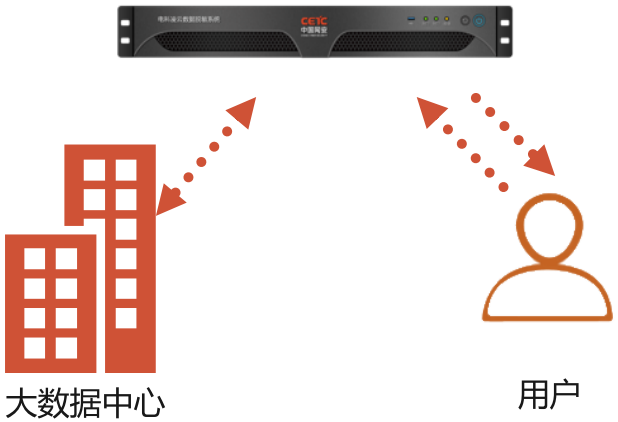
数据混淆：混淆算法是将敏感数据的内容进行无规则打乱，从而在隐藏敏感数据的同时能够保持原始数据的组成方式。

数据安全防护关键技术

(2) 数据脱敏-模式

根据数据开放情况，根据不同服务场景提供不同场景下的脱敏方式，进而自动识别敏感数据并作脱敏处理，确保个人隐私信息或涉密信息不会被滥用。

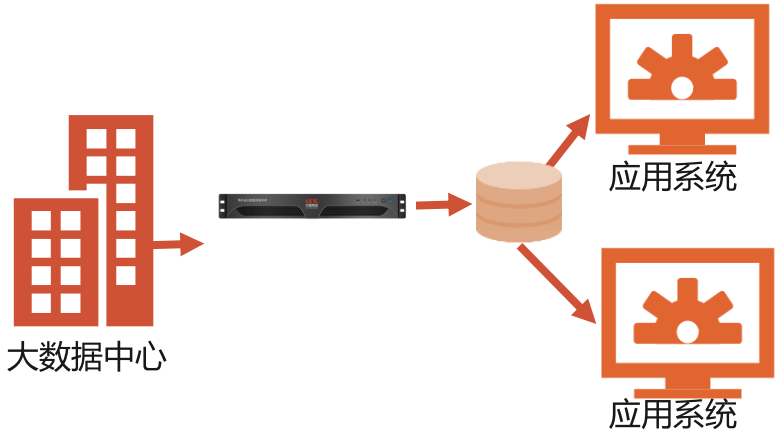
动态脱敏



实时脱敏

在线实时接收用户请求，并根据用户配置的脱敏策略，对数据进行脱敏处理，实时返回脱敏结果。

静态脱敏



非实时脱敏

在生产环境中抽取原始数据，对原始数据进行不同级别的脱敏配置，一次性将不同安全级别的原始数据全部脱敏输出到指定位置，然后授予不同用户访问不同安全级别数据的权限。

数据安全防护关键技术

（3）数据防泄漏

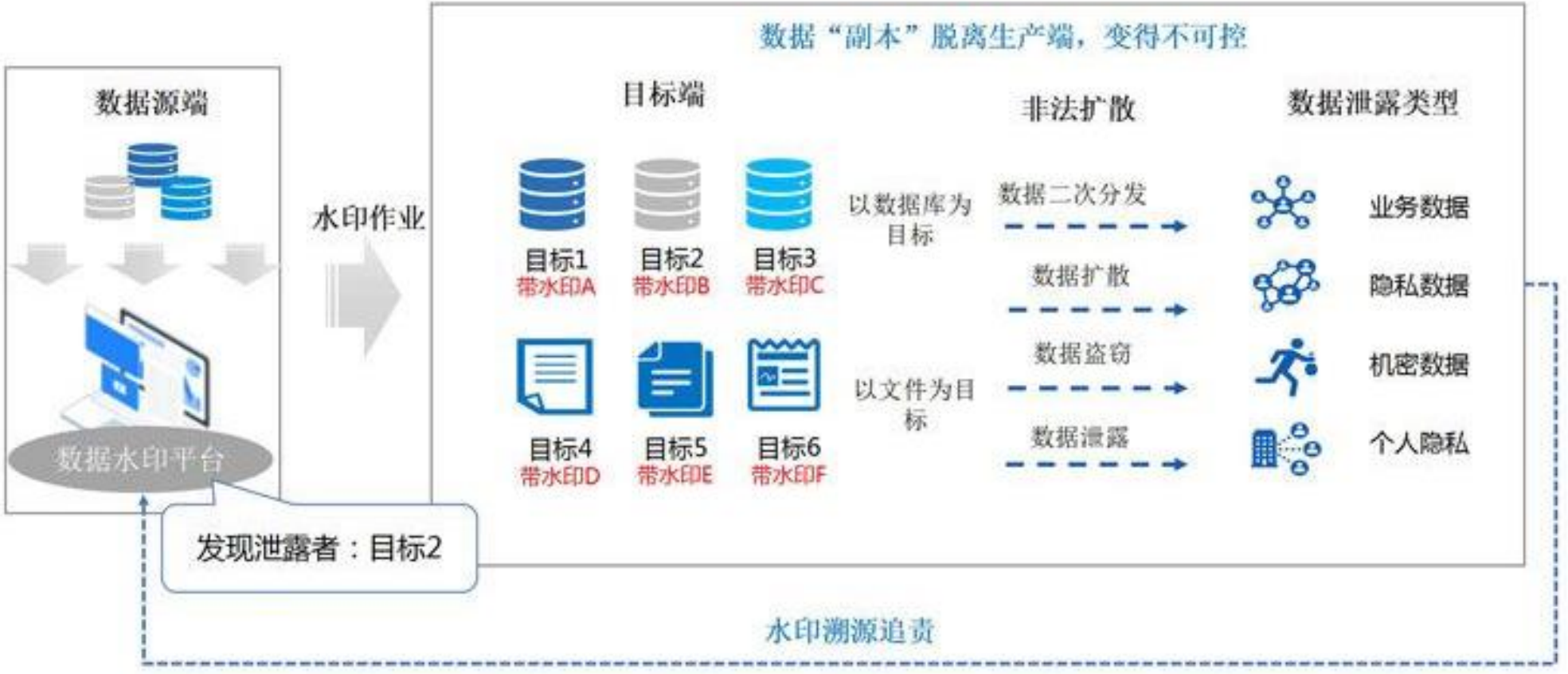
Data Leakage(Loss) Prevention数据泄露防护，简写为DLP，就是通过内容识别达到到对数据的防控。防护范围主要包括网络防护和终端防护。网络防护主要以审计、控制为主，终端防护除审计与控制能力外，还应包含传统的主机控制能力、加密和权限控制能力。可以通过识别如 关键字、正则表达式、文档指纹、确切数据源（数据库指纹）等。

基本来说，DLP其实就是一个综合体。最终实现的效果是智能发现、智能加密、智能管控、智能审计，就是一整套的数据泄露防护方案。从另一个角度保证数据机密。

数据安全防护关键技术

(4) 数据水印

针对数据文件中的敏感数据进行高级别仿真水印标识的数据安全产品，广泛应用于内外部数据共享交换等场景，解决数据扩散后泄漏主体不明确、无法追溯等难题。



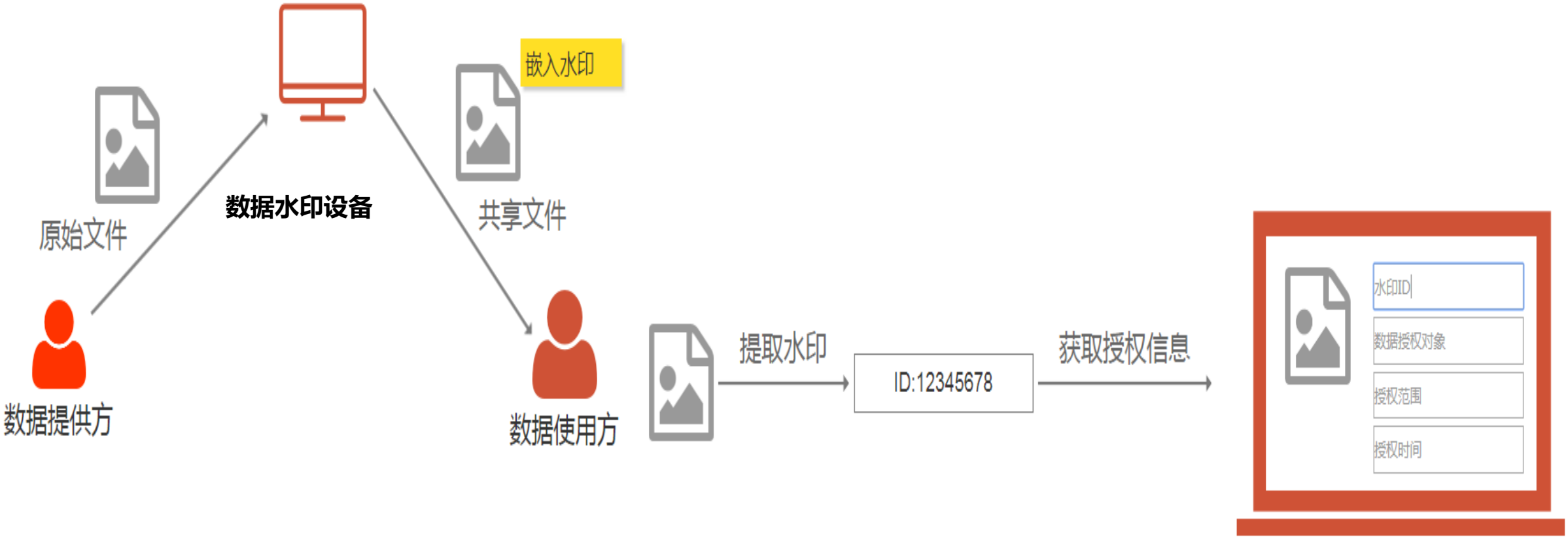
（4）数据水印

数据水印具有高隐蔽性、高易用性、高管理融合性等特点。数据水印可对通过系统外发数据行为进行流程化管理，在不影响数据使用的前提下，自动生成水印并添加数据标记，可以对泄露数据源的追溯，避免了内部人员外发数据过程泄露无法对事件定责追溯，提高了数据传递的安全性和可追溯能力。

水印的类别分为**鲁邦水印和脆弱水印**，**脆弱水印可用于数据篡改验证；鲁棒水印被用在非结构化数据的溯源追踪**。根据功能划分分为**水印添加功能，水印提取功能**。

数据安全防护关键技术

(4) 数据水印



（5）数据容灾备份

数据备份是指为防止系统出现操作失误或系统故障导致数据丢失，而将全部或部分数据集合从应用主机的硬盘或阵列复制到其它的存储介质的过程。

传统的数据备份主要是采用内置或外置的磁带机进行冷备份。但是这种方式只能防止操作失误等人为故障，而且其恢复时间也很长。随着技术的不断发展，数据的海量增加，不少的企业开始采用网络备份。网络备份一般通过专业的数据存储管理软件结合相应的硬件和存储设备来实现。

（5）数据容灾备份-存储方式

1.定期磁带

远程磁带库、光盘库备份。即将数据传送到远程备份中心制作完整的备份磁带或光盘。远程关键数据+磁带备份。采用磁带备份数据，生产机实时向备份机发送关键数据。

2.数据库备份

这种方式就是在与主数据库所在生产机相分离的备份机上建立主数据库的一个拷贝。

由于传统的数据存储方式过于简单化，过于集中管理而造成了大量数据的堆积。这样一个公司或企业要使用大量的数据就需要大量的存储数据的介质，而导致服务器的回应下降乃至崩溃。这样，分布式数据库技术在构建企业级应用程序中更为广泛流行，是因为分布式数据库存储方式给企业带来了很多的方便以及好处。

（5）数据容灾备份-存储方式

3.网络数据

这种方式是对生产系统的数据库数据和所需跟踪的重要目标文件的更新进行监控与跟踪，并将更新日志实时通过网络传送到备份系统，备份系统则根据日志对磁盘进行更新。

4.远程镜像

通过高速光纤通道线路和磁盘控制技术将镜像磁盘延伸到远离生产机的地方，镜像磁盘数据与主磁盘数据完全一致，更新方式为同步或异步。

目前常见的云备份方式可以看作是网络备份和远程镜像的一种应用，具有广阔的应用前景。

（5）数据容灾备份-本地备份

什么是本地备份?

本地备份也叫手工备份，是每台服务器有自己的本地备份设备，这是最简单的备份方案。

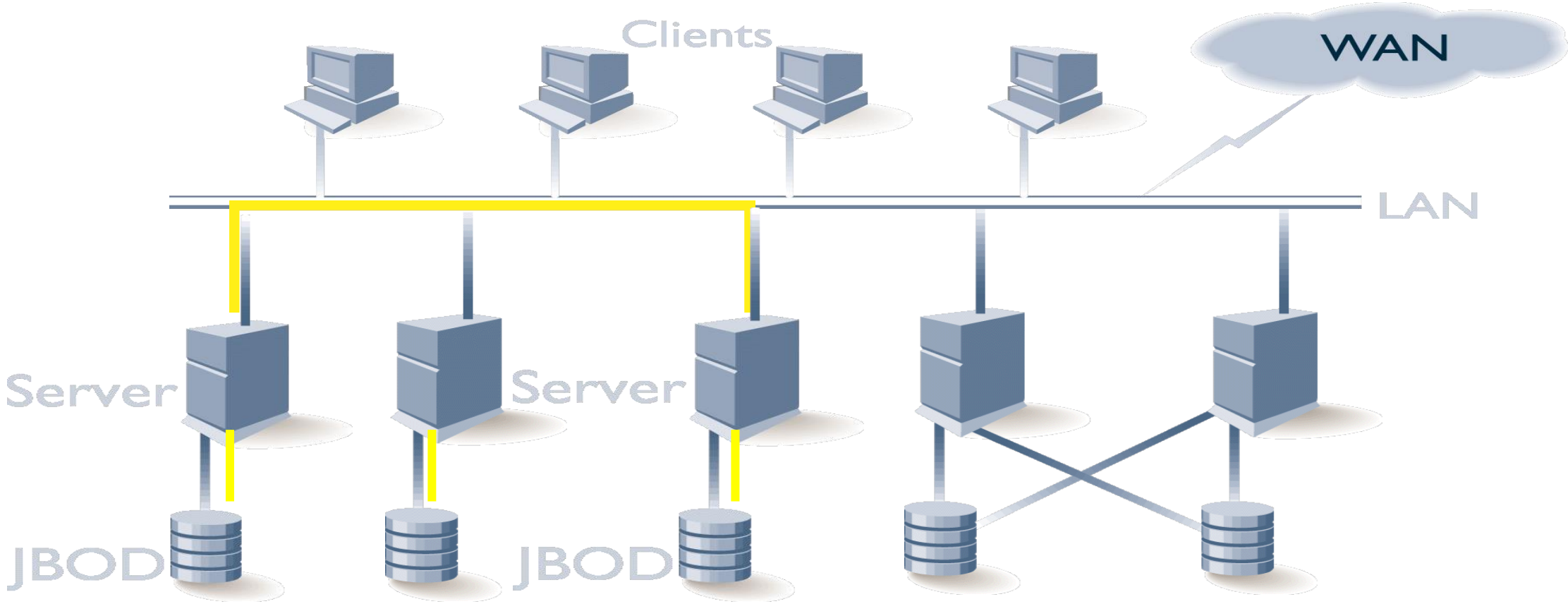
由于备份设备连接到每台服务器上，所以每台服务器不得不单独管理备份进程。这种备份方案不仅增加了硬件投资，还增加了管理的费用。

本地备份的问题:

- **手动，用户每次都要自己配置备份任务**
- **无规律，没有策略**
- **缺乏对备份数据的管理**

数据安全防护关键技术

(5) 数据容灾备份-本地备份



（5）数据容灾备份-网络备份

什么叫网络备份?

网络备份也叫做LAN备份，是一种流行的备份解决方案。通常，带有备份设备的备份服务器被放置在网络中。备份服务器负责整个系统的备份，它管理整个网络的备份策略、备份媒体和备份目标。所有的备份数据必须通过本地局域网进行传输。

网络备份的优点

- 实现了大容量自动化、集中式备份
- 备份过程有策略管理，无需管理员介入
- 网络内所有需要备份的服务器可共享一台备份设备

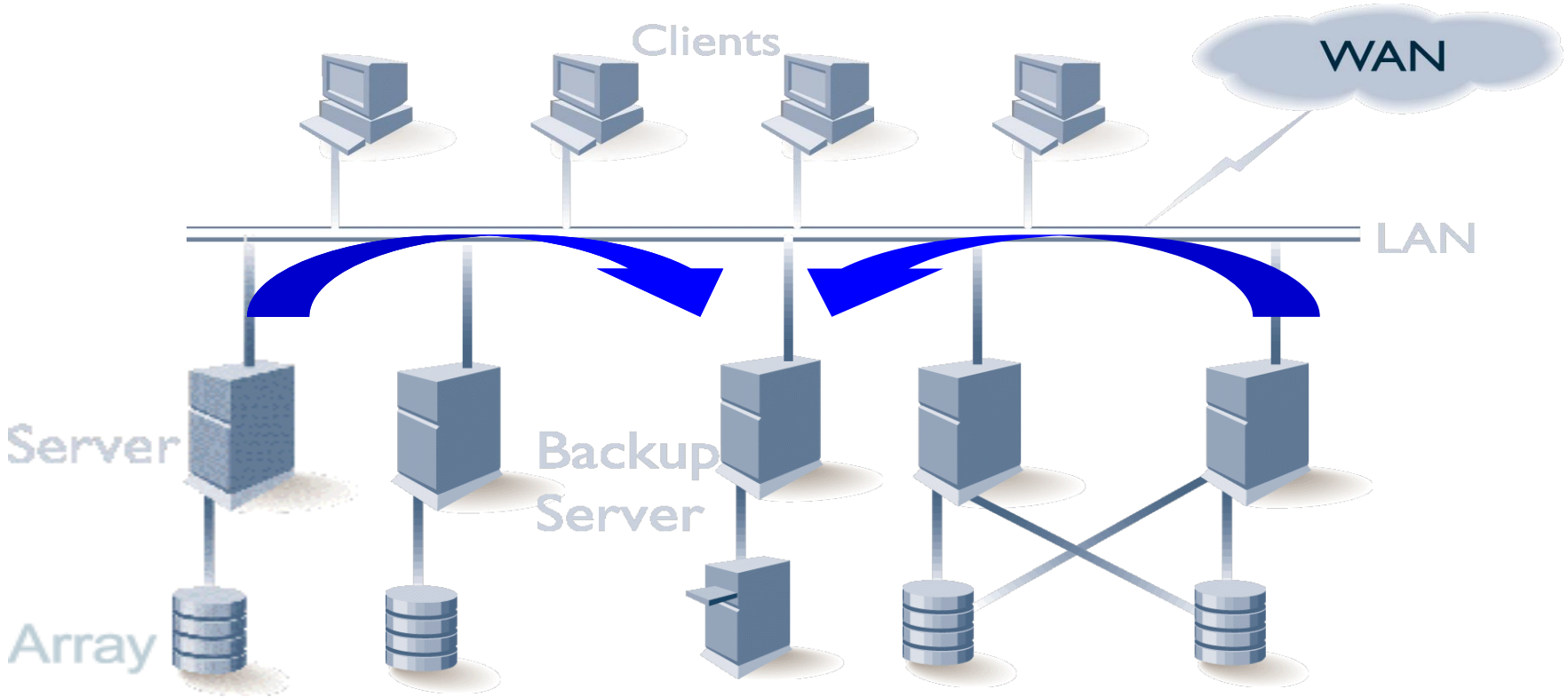
（5）数据容灾备份-网络备份

网络备份的问题

- **这种基于LAN的备份解决方案将强制备份数据通过LAN进行传输，因此在备份过程中网络就会超负荷。**
- **这不仅会导致备份性能下降，还会使备份时间更长。**

数据安全防护关键技术

(5) 数据容灾备份-网络备份



数据安全防护关键技术

（5）数据容灾备份- LAN-free

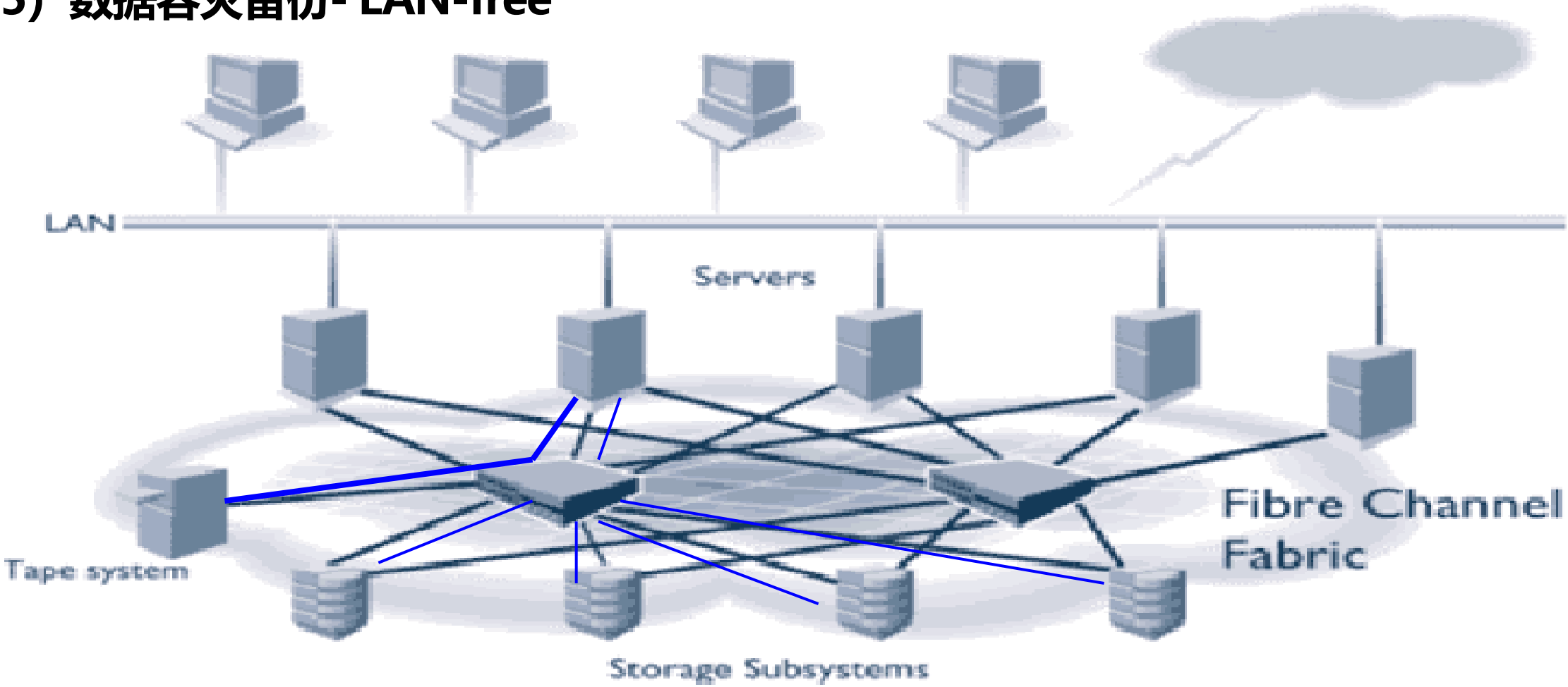
LAN-Free备份由于数据通过LAN传播，当需要备份的数据量较大，备份时间窗口紧张时，网络容易发生堵塞。在SAN环境下，可采用存储网络的LAN-Free备份，需要备份的服务器通过SAN连接到磁带机上，在LAN-Free备份客户端软件的触发下，读取需要备份的数据，通过SAN备份到共享的磁带机。

LAN Free备份的优点

- 提高备份速度，减少备份及恢复窗口
- 优化备份设备的使用
- 降低备份服务器负担
- 消除对业务网络（LAN）的影响

数据安全防护关键技术

(5) 数据容灾备份- LAN-free



（5）数据容灾备份-Server Less

什么叫Server Less备份?

Server less备份是备份技术中最近的技术，它可以在LAN Free备份的基础上节省有价值的服务器资源（CPU、内存等）。一些Server less备份设备放在服务器和存储子系统之间，这些设备负责备份数据的全部责任，它从存储阵列向磁带设备直接发送数据。

Server Less备份的优点

- **实现不影响应用的备份**
- **极大的减少服务器负担**

(6) 数据恢复

硬盘保存文件时是按簇保存在硬盘中的，而保存在哪些簇中则记录在文件分配表里。硬盘文件删除时，并非把所有内容全部清零，而是在文件分配表里把保存该文件位置的簇标记为未使用，文件分配表标记为未使用的簇，以后就可以将文件直接写入该簇，在重新写入之前，上次删除文件的内容实际上依然在该簇中，所以，只要找到该簇，就可以恢复文件内容。这也是为什么误删文件后不要再往该硬盘写入数据的原因。只有在相同簇中写入新文件以后，文件才会被彻底破坏。

同时从物理角度来看，特别是了解硬盘的结构以后，大家会发现，当我们保存数据的时候，盘片会变得凸凹不平从而保存数据。我们删除了文件的时候，并没有把所有的凸凹不平的介质抹平，而是把它的地址抹去，让操作系统找不到这个文件，从而认为它已经消失，进而在这个地方写数据，把原来的凸凹不平的数据信息覆盖掉。

(6) 数据恢复

数据恢复的原理：如果数据没被覆盖，我们就可以**用软件，通过操作系统的寻址和编址方式，重新找到那些没被覆盖的数据并组成一个文件，如果几个小地方被覆盖，可以用差错效验位来纠正。**当然，如果覆盖已满，基本就没办法恢复了。

数据恢复种类：

- 逻辑故障数据恢复，逻辑故障是指与文件系统有关的故障。常见的逻辑故障有：无法进入操作系统、文件无法读取、文件无法被关联的应用程序打开、文件丢失、分区丢失、乱码显示等。逻辑故障造成的数据丢失，大部分情况是可以通过专用数据恢复软件找回的。

数据安全防护关键技术

(6) 数据恢复

- 硬件故障数据恢复，硬盘一般由电路板、固件、磁头、盘片、电机等电子器件/软件/机械三部分组成，而任一组件都可能发生故障。比如雷击、高压、高温等造成的电路故障；高温、振动碰撞等造成的机械故障、存储介质老化造成的物理坏磁道扇区故障和意外丢失损坏的固件BIOS信息等。
- 磁盘阵列RAID数据恢复
  - 硬盘数据恢复，硬件故障的数据恢复的步骤是先诊断，找到故障点。先修复硬件故障，然后再修复其他软故障，最终将数据成功恢复。
  - U盘数据恢复，U盘损坏或出现电路板故障、磁头偏移、盘片划伤等情况时，可采用开体更换、加载、定位等方法进行数据修复。

数据安全防护关键技术

（7）数据安全销毁

数据销毁是通过建立针对数据内容的清除、净化机制，实现对数据的有效销毁，防止因对存储介质中的数据内容进行恶意恢复而导致的数据泄漏风险。

数据销毁有两个目的，一是合规要求，国家法律法规要求重要数据不被泄漏;另外就是组织本身的业务发展或管理需要。日常工作过程中，用户往往采取删除、硬盘格式化、文件粉碎等方法销毁数据，但是这些方法并不是完全安全。

主流的数据销毁技术，主要有**数据删除、数据清除、物理销毁**。

（7）数据安全销毁-数据删除

删除磁盘数据的常规方法主要有：删除和格式化

- 删除数据最便捷的方法，如：经常采用的“Delete”的系统删除命令。实际上并没有真正的将数据从硬盘上删除，只是将文件的索引删除而已，让操作系统和使用者认为文件已经删除，又可以把腾出空间存储新的数据。

- “格式化”有许多不同的含义：物理的或低级格式化，操作系统的格式化，快速格式化，分区格式化等等 …大多数情况下，普通用户采用的格式化不会影响到硬盘上的数据。格式化仅仅是为操作系统创建一个全新的空的文件索引，将所有的扇区标记为“未使用”的状态，让操作系统认为硬盘上没有文件。格式化后的硬盘数据也是能够恢复。

数据安全防护关键技术

（7）数据安全销毁-数据清除

数据清除，由于磁盘可以重复使用，前面的数据被后面的数据覆写后，前面的数据被还原的可能性就大大降低了，随着被覆写次数的增多，能够被还原的可能性就趋于0，但相应的时间支出也就越多。密级要求的高低对应着不同的标准，低密级要求的就是一次性将磁盘全部覆写；高密级要求则须进行多次多规则覆写。

有相应标准对数据清除做规范和要求：

- 美国国防部 DOD 的 5220.22M 标准;
- 北约 NATO 的标准

参考标准实质就是多次覆写的标准，规定了覆盖数据的次数，覆盖数据的形式。

（7）数据安全销毁-数据清除

复写原理（Overwriting），使用预先定义的格式——无意义、无规律的信息来覆盖硬盘上原先存储的数据。这是销毁数据的既有效又可操作的方法。如果数据被“成功”的完全覆写，即使只覆写一次，也可以认为数据是不可恢复的。

硬盘上的数据都是以二进制的“1”和“0”形式存储的。完全覆写后也就无法知道原先的数据是“1”还是“0”了，也就达到了清除数据的目的。

覆写的方法：根据覆写时的具体顺序，软件覆写分为逐位覆写、跳位覆写、随机覆写等模式，根据时间、密级的不同要求，可组合使用上述模式，可靠的专业清除软件应同时支持多种模式。

覆写的局限性：既然，一次完全的覆写就可以彻底清除数据，那标准为什么还要规定须多次覆写呢？因为磁信号泄露了数据的历史痕迹，可以通过特殊的专业设备来识别痕迹进而恢复被覆盖的数据。

（7）数据安全销毁-物理销毁

a. 消磁

使用专门的消磁设备来磁化磁盘表面的磁介质。它产生电磁场来磁化磁盘。而磁盘类磁性介质，一旦磁性结构被破坏，数据也就不复存在了。

a. 化学腐蚀

化学腐蚀的方法把记录涉密数据的物理载体完全破坏掉，从而从根本上解决数据泄露问题的硬盘销毁方式。

a. 物理破坏

常见的是盘片划损、硬盘回炉、外力破损等方法。

(8) 数据库审计

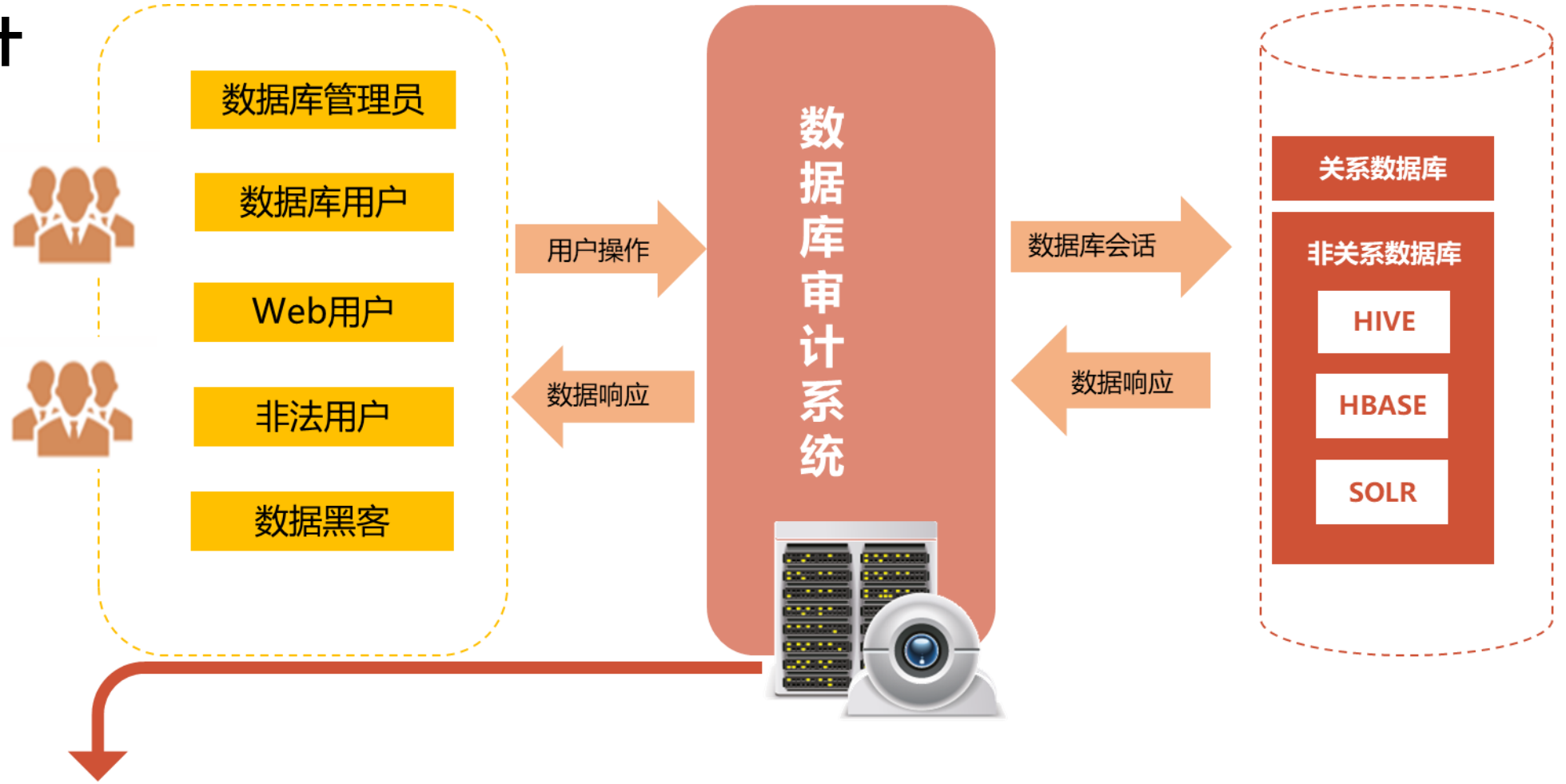
以安全事件为中心，以全面审计和精确审计为基础，实时记录网络上的数据库活动，对数据库操作进行细粒度审计的合规性管理，对数据库遭受到的风险行为进行实时告警。它通过对用户访问数据库行为的记录、分析和汇报，来帮助用户事后生成合规报告、事故追根溯源，同时通过大数据搜索技术提供高效查询审计报告，定位事件原因，以便日后查询、分析、过滤，实现加强内外部数据库网络行为的监控与审计，提高数据资产安全。

探测数据库的访问操作行为，通过深度语义级解析及风险模型匹配判断访问行为是否存在安全威胁，同时对访问操作行为做一个完整的记录；对违反安全规则的事件发生后，能有效的追查责任和分析原因，必要时还可以为惩罚恶意攻击行为提供必要的证据。

数据安全防护关键技术

（8）数据库审计

记录数据库的各种访问操作，智能分析和监控访问者的各种操作，进行实时威胁预警，并对事件进行统计分析记录，有效支持电子取证。



THE END