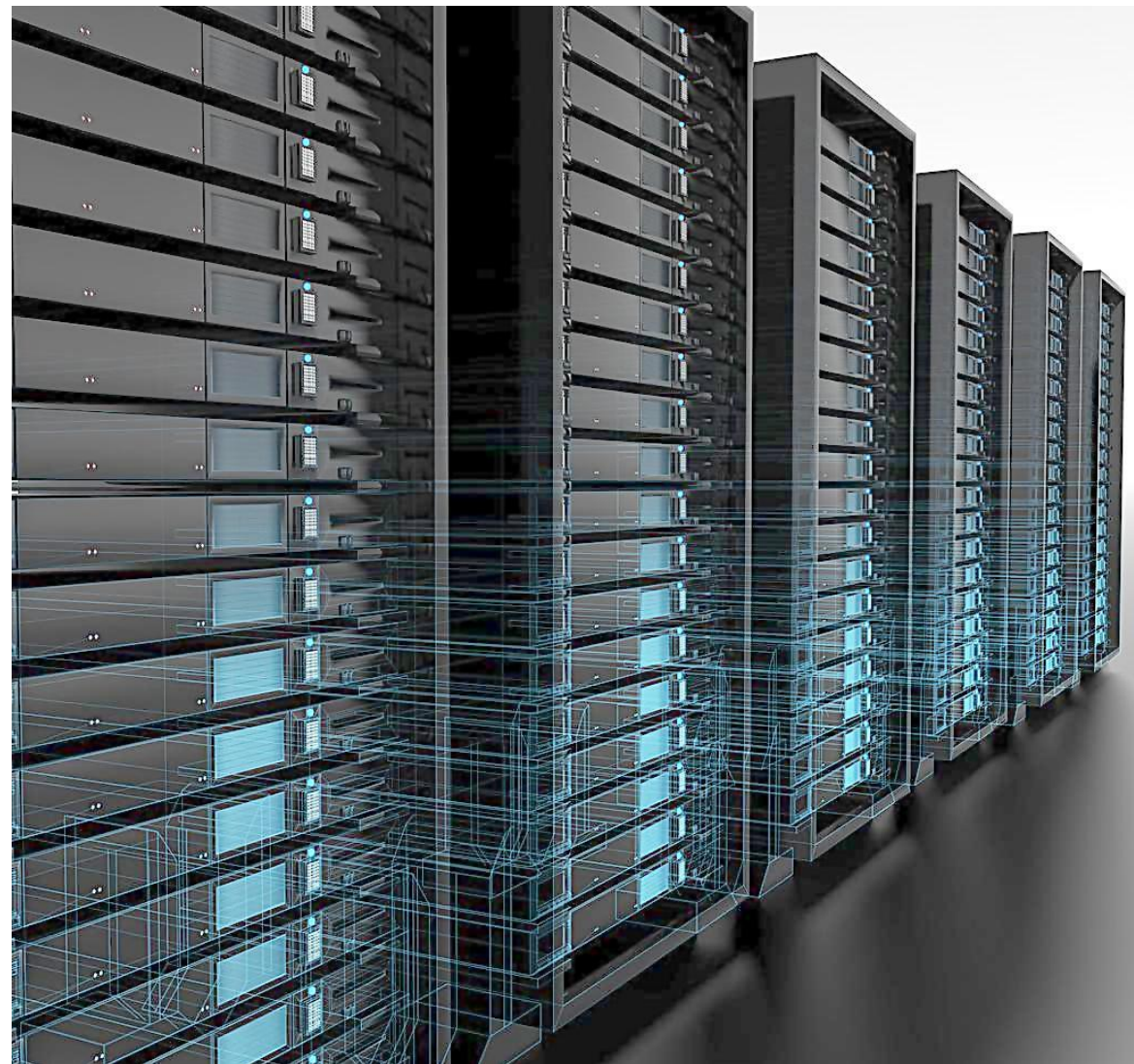


2017级新生培训讲座



DataBase

一、基本概念

1、数据（Data）

2、数据库（DB）

3、数据库管理系统（DBMS）：位于用户和操作系统之间的一层数据管理软件。主要功能包括：

数据定义功能

数据操纵功能

数据库的运行管理

数据库的建立和维护功能

4、数据库系统（DBS）：计算机中引入数据库后的系统



二、数据管理与数据处理

1、数据管理：

对数据收集、整理、组织、存储、维护、检索、传送等

对象

操作

目标：在适当的时候以适当的形式给适当的人提供适当的数据。

2、数据处理：对数据进行加工、计算、提炼，
从而产生新的有效数据的过程

数据

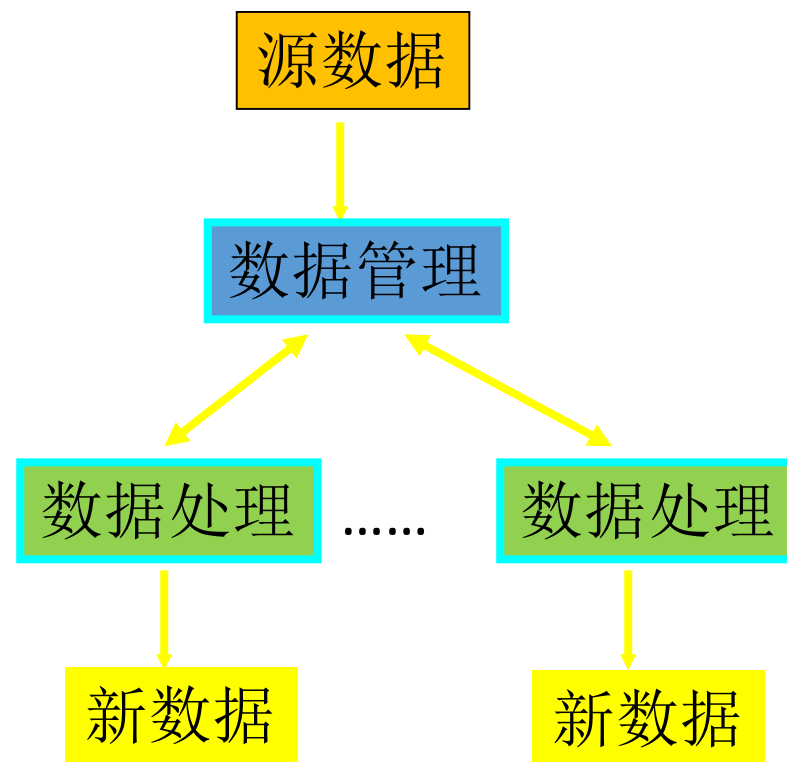


信息

3、管理与处理的关系：

- 管理是处理的基础
- 处理为管理服务

管理和处理又可看成一个问题两个阶段，故可以统一起来，其中心是管理

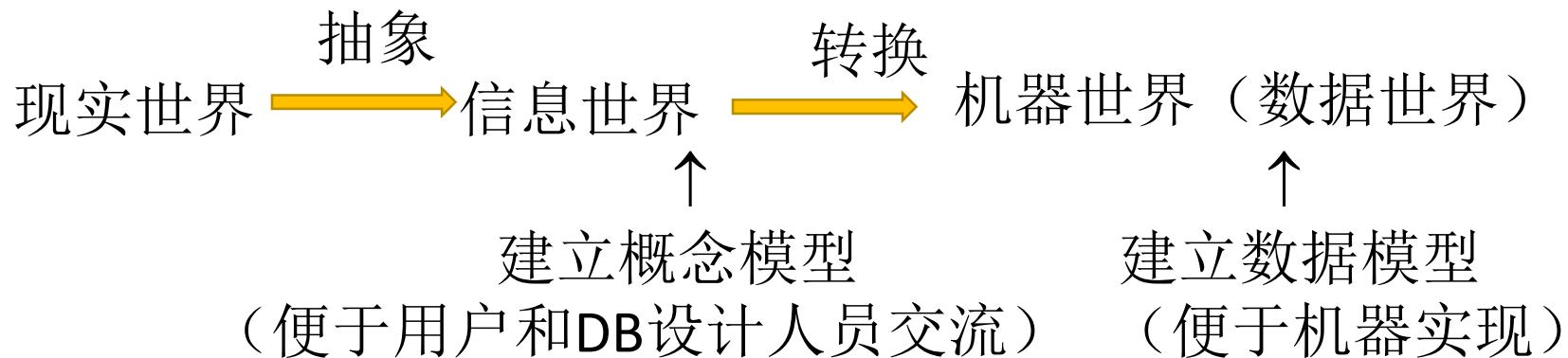


4、数据管理的发展阶段

- ✧ 人工管理阶段（50年代中期以前）
- ✧ 文件系统阶段（50年代中期至60年代后期）
- ✧ 数据库系统阶段（60年代后期以后）

三、数据模型

数据处理的抽象过程（涉及三个领域）



1.概念模型(信息模型)

把现实世界中的客观对象抽象成的某种信息结构，主要用于数据库设计。

独立于具体的计算机系统

独立于具体的DBMS支持的数据模型

2.数据模型

形式化描述数据、
数据之间的联系
以及数据操作
和有关的语义
约束规则的方法

如何表示
实体及联系

(难点是表示联系)

根据现实世界实体间联系的特征
用四种不同的方法进行抽象

层次模型

网状模型

关系模型

面向对象模型

数据结构

数据操作

完整性约束

如何实现
查、增、删、改

如何保证数据的
约束条件得到满足

(因此, 是按照数据结构的
类型来命名数据模型)

四、 SQL概述

1.SQL的特点

- 综合统一
- 高度非过程化
- 面向集合的操作方式
- 以同一种语法结构提供两种使用方法
- 语言简洁， 易学易用

2.SQL的种类

DDL

DML

DCL

SQL 语言的动词

SQL 功 能	动 词
数 据 定 义	CREATE, DROP, ALTER
数 据 查 询	SELECT
数 据 操 纵	INSERT, UPDATE DELETE
数 据 控 制	GRANT, REVOKE

五、 数据库设计范式

- 范式是符合某一种级别的关系模式的集合。
- 关系数据库中的关系必须满足一定的要求。满足不同程度要求的为不同范式。
- 范式的种类：

第一范式(1NF)

第二范式(2NF)

第三范式(3NF)

BC范式(BCNF)

第四范式(4NF)

第五范式(5NF)

第一范式 (1NF)

编号	品名	进货		销售		备注
		数量	单价	数量	单价	



编号	品名	进货数量	进货单价	销售数量	销售单价	备注

符合1NF的设计

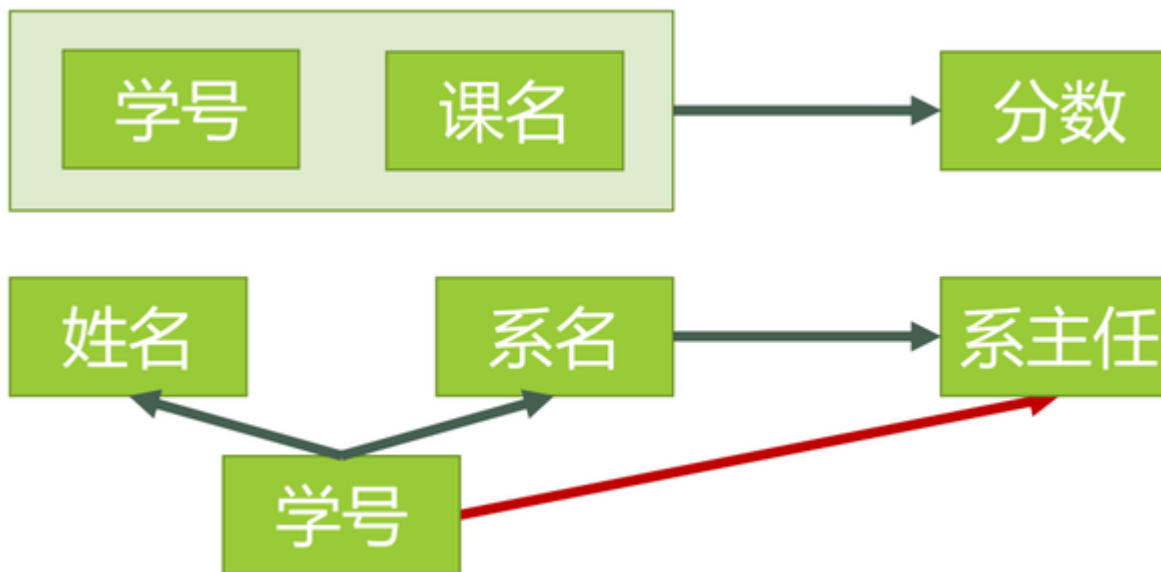
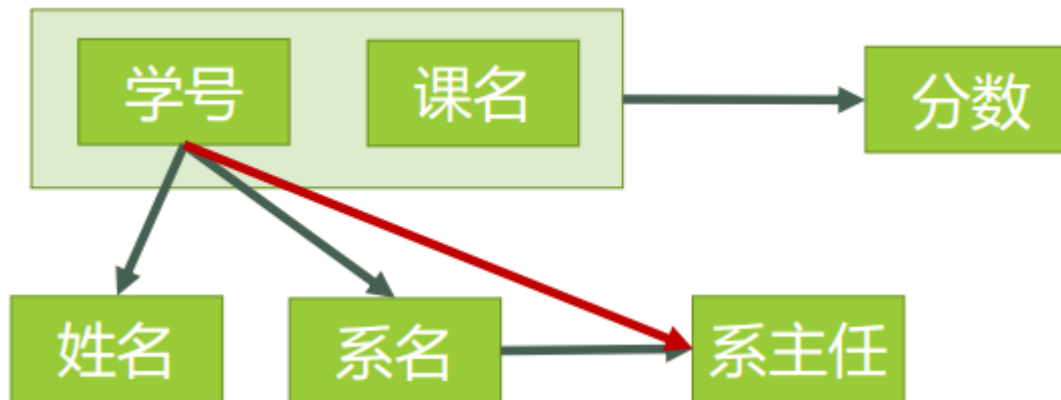
学号	姓名	系名	系主任	课名	分数
1022211101	李小明	经济系	王强	高等数学	95
1022211101	李小明	经济系	王强	大学英语	87
1022211101	李小明	经济系	王强	普通化学	76
1022211102	张莉莉	经济系	王强	高等数学	72
1022211102	张莉莉	经济系	王强	大学英语	98
1022211102	张莉莉	经济系	王强	计算机基础	88
1022511101	高芳芳	法律系	刘玲	高等数学	82
1022511101	高芳芳	法律系	刘玲	法学基础	82

第二范式 (2NF)

第二范式 (2NF) 是在第一范式 (1NF) 的基础上建立起来的，即满足第二范式 (2NF) 必须先满足第一范式 (1NF)。

$$X \xrightarrow{F} Y$$

$$X \xrightarrow{P} Y$$



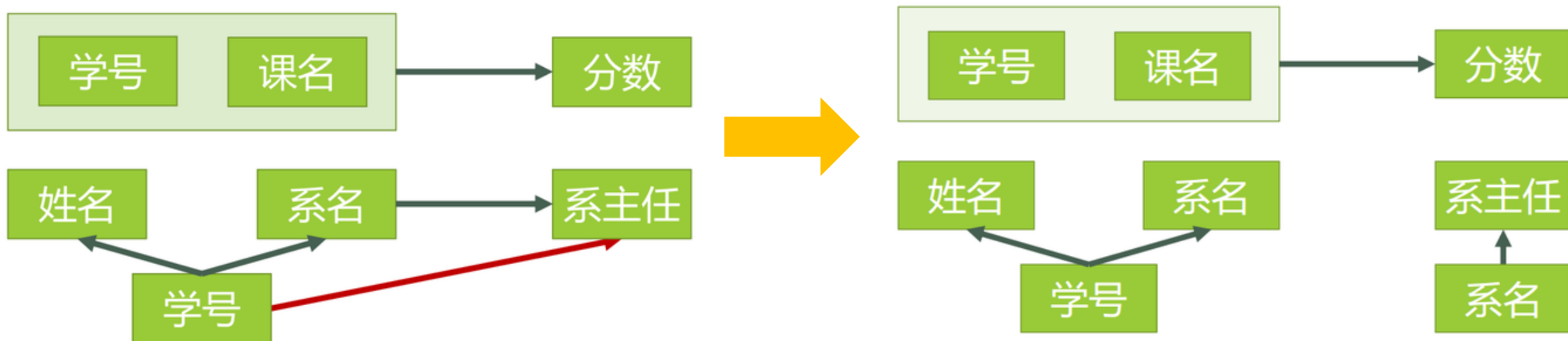
模式分解以后新的数据，符合2NF

学号	课名	分数
1022211101	高等数学	95
1022211101	大学英语	87
1022211101	普通化学	76
1022211102	高等数学	72
1022211102	大学英语	98
1022211102	计算机基础	88
1022511101	高等数学	82
1022511101	法学基础	82

学号	姓名	系名	系主任
1022211101	李小明	经济系	王强
1022211102	张莉莉	经济系	王强
1022511101	高芳芳	法律系	刘玲

第三范式 (3NF)

仅仅符合2NF的要求，很多情况下还是不够的，我们还需要将符合2NF要求的数据表改进为符合3NF的要求



再次模式分解以后新的数据，符合3NF

学号	课名	分数
1022211101	高等数学	95
1022211101	大学英语	87
1022211101	普通化学	76
1022211102	高等数学	72
1022211102	大学英语	98
1022211102	计算机基础	88
1022511101	高等数学	82
1022511101	法学基础	82

学号	姓名	系名
1022211101	李小明	经济系
1022211102	张莉莉	经济系
1022511101	高芳芳	法律系

系名	系主任
经济系	王强
经济系	王强
法律系	刘玲

CASE

- 1.某公司有若干个仓库；
- 2.每个仓库只能有一名管理员，一名管理员只能在一个仓库中工作；
- 3.一个仓库中可以存放多种物品，一种物品也可以存放在不同的仓库中。每种物品在每个仓库中都有对应的数量。

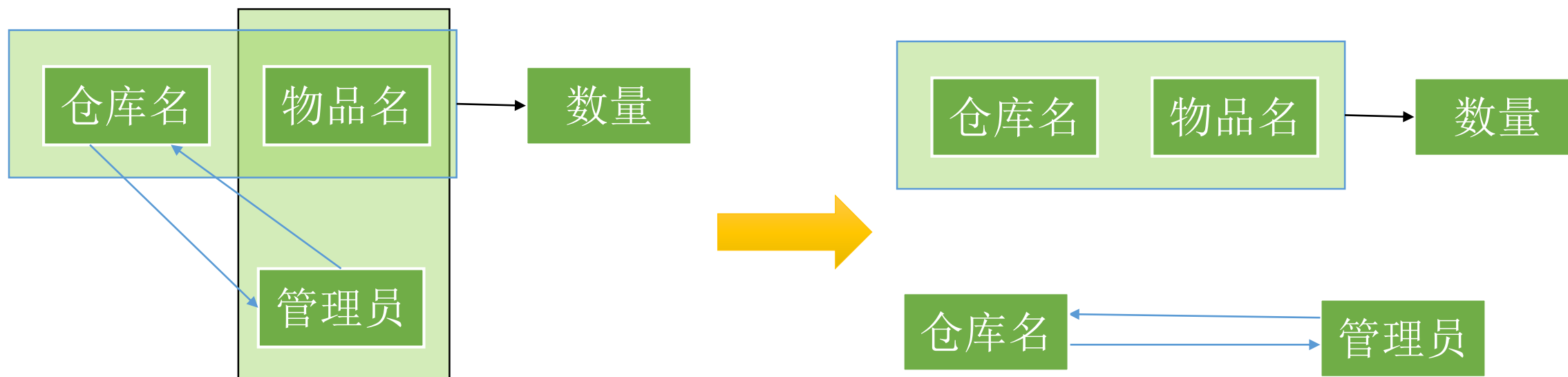
关系模式已经属于了 3NF，那么
这个关系模式是否存在问题呢

?

仓库名	管理员	物品名	数量
上海仓	张三	iPhone 5s	30
上海仓	张三	iPad Air	40
北京仓	李四	iPhone 5s	50
北京仓	李四	iPad Mini	60

BCNF范式

在满足第二第三范式的情况下，主属性内部也不能部分或传递依赖。**判断方法：**箭头左边的必须是候选码，不是候选码的就不是BC范式。



函数依赖与规范化

函数依赖有三种情况：1.完全函数依赖、2.部分函数依赖、3.传递函数依赖

$x=a, y=2a+3$ y 存在对 x 的完全函数依赖

$x=a+b, y=b$ y 存在对 x 的部分函数依赖

$x=a, y=x, z=y$ z 存在对 x 的传递函数依赖。

主键：用于唯一标识一行的键

候选成为主键

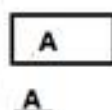
候选键：能够标识每一行的键

主属性：包含在候选键中的属性

非主属性：剩下的属性（列名）

单键：包含一个属性的键

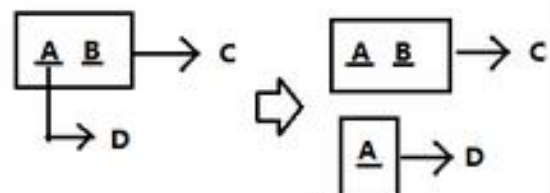
复合键：包含多个属性的键



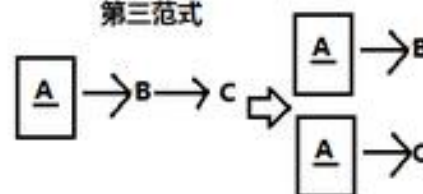
矩形表示主键
下划线表示主属性

能创建出来的表都符合第一范式

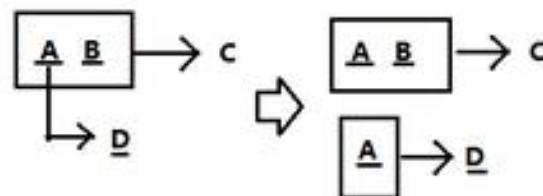
第二范式



第三范式



BCNF范式



第四范式

学号	姓名	兄弟	专业
1	小王	小白	英语
1	小王	小红	英语
1	小王	小蓝	英语
2	张三	张四	数学
2	张三	张五	数学

学号	兄弟
1	小白
1	小红
1	小蓝
2	张四
2	张五

第五范式

A	B	C
a1	b1	c1
a2	b1	c2
a1	b2	c1
a2	b2	c2

A	B
a1	b1
a2	b2

B	C
b1	c1
b2	c2

C	A
c1	a1
c2	a2

- 关系模式规范化的基本步骤



六、 DBA （Data Base administrator）

DBA的角色定义

开发型DBA

- 数据库安装。
- 数据库架构设计（架构和建模）
- 代码开发（存储过程，SQL）

运维型DBA

- 数据库日常监控
- 故障处理
- 性能优化
- 数据备份，容灾
- 数据库安全规划

DBA的职责：

- ① 决定数据库的内容和逻辑结构、存储结构
- ② 确定数据的安全性要求和完整性约束条件
- ③ 监控数据库的使用和运行，维护数据库
- ④ 决定数据库的存储结构和存储策略
- ⑤ 负责数据库的改进和重组重构

DBA的守则

在自己的责任范围内

- 让数据库设计更合理，预防设计导致的性能或安全隐患。
- 数据更安全。
- 数据库性能更优。
- 数据库日常管理更合理。
- 故障发现，处理及时。

数据库的架构设计

数据库架构

- 分布**or**单库

实例的冗余

- RAC or single**

数据的安全和容灾

- DG or streams or RMAN**

空间的考虑，存储的规划

- ASM+SAN**

软件的生命周期和业务（数据）增长的预测。

数据库的运维

数据库的监控

- 定制
- 开源软件+脚本
- OEM+grid control
- 第三方软件

数据的管理及安全

- 备份策略
- 数据的保留，删除策略
- 分区，压缩，只读表空间..

故障处理

- 对数据库的深入理解
- Oraclesupport(metalink)
- 官方文档docs.oracle.com
- asktom.oracle.com
- www.itpub.net

性能优化

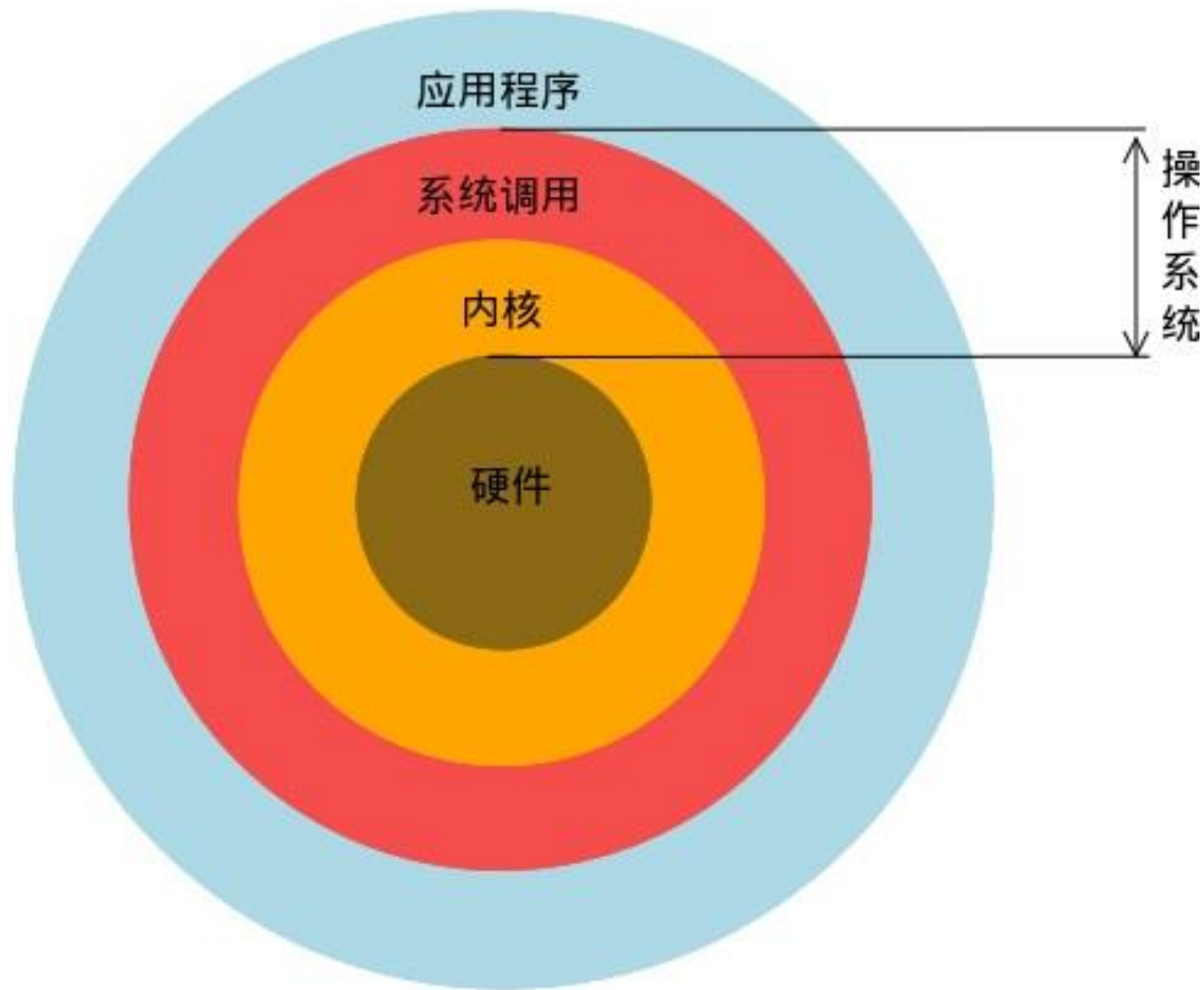
- 对业务流程的深入理解
- 用户感知为导向的优化思路
- 性能基线的建立
 - cpu, i/o, sessions.....
- 定期的AWR报告分析。

数据库的安全

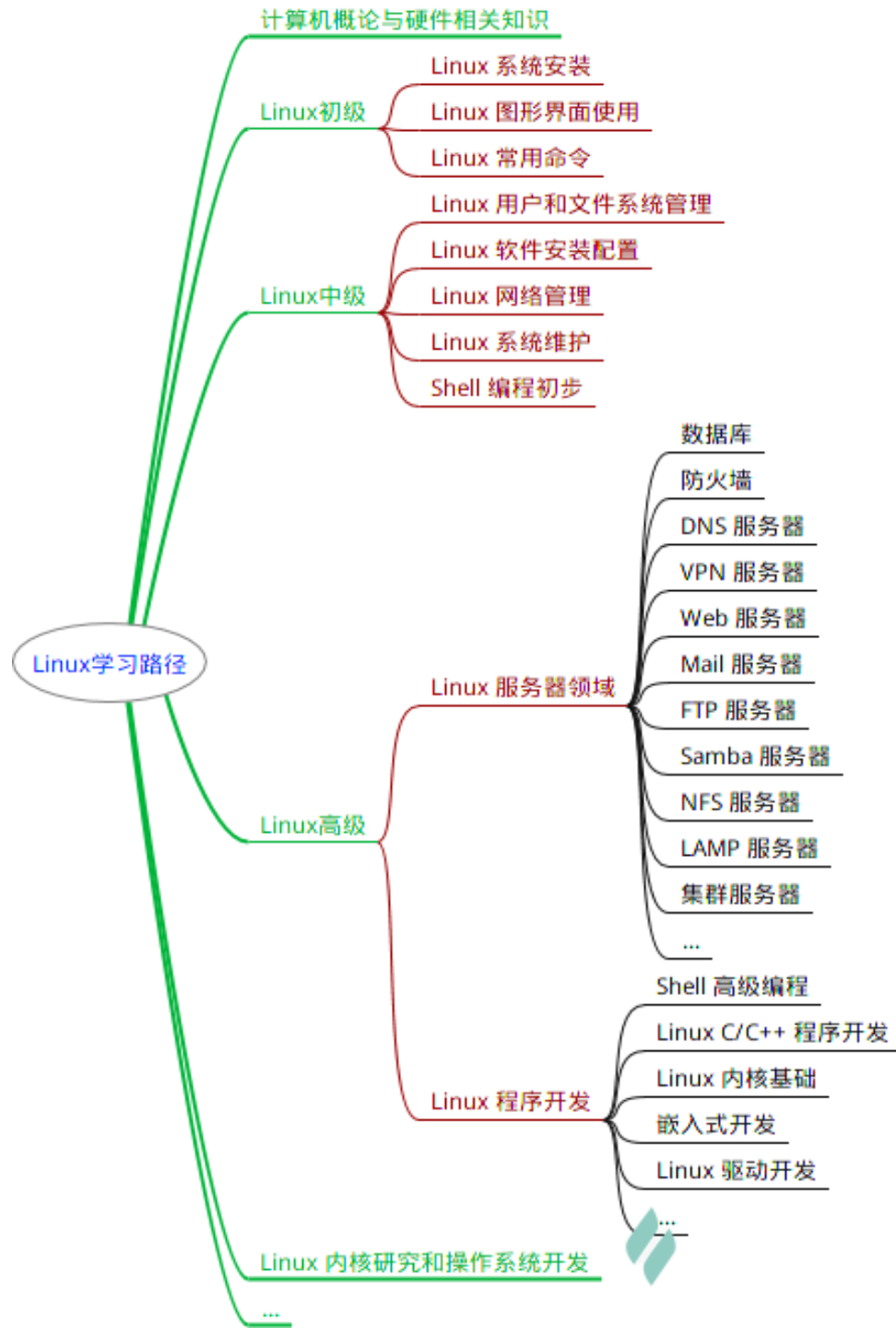
- 口令管理策略
 - 精细化的授权机制。
- DBVAULT, AUDITVALUT
- Oracle的细粒度审计（FGA）
- 数据加密
- 操作系统口令管理

Linux 运维

Linux 也就是系统调用和内核那两层，当然直观的来看，我们使用的操作系统还包含一些在其上运行的应用程序，比如文本编辑器，浏览器，电子邮件。



Linux学习路线图



学习Linux的经验

1.学习心态

- 明确目的
- 面对现实
- 是学习 Linux 操作系统本身还是某一个Linux发行版

2、从基础入手，切勿眼高手低

熟记命令的前提下，使用命令行界面往往要较使用图形用户界面的操作速度要快，这也是你以后玩转Linux的基本条件之一。

3、多总结，勤思考，多记笔记

4、及时充电，提高自学能力

Virtualization

虚拟化是为一些组件创建基于软件的或虚拟（而不是物理）表现形式的过程。虚拟化可以应用于应用、服务器、存储和网络，它是一种可以为所有规模的企业降低 IT 开销，同时提高效率和敏捷性的最有效方式。

虚拟化的类型



服务器虚拟化



网络虚拟化



桌面虚拟化



软件定义的存储

虚拟化的优势

虚拟化可以提高 IT 部门的敏捷性、灵活性和可扩展性，同时大幅节约成本。工作负载的部署速度更快、性能和可用性得到提升、运维实现自动化，所有这一切不仅简化了 IT 管理，还降低了拥有成本和运维成本。其他优势包括：

- 降低资金成本和运维成本。
- 最大限度地减少或消除停机。
- 提高 IT 部门的工作效率、效益、敏捷性和响应能力。
- 更快地调配应用和资源。
- 支持业务连续性与灾难恢复。
- 简化数据中心管理。
- 构建真正的软件定义的数据中心

虚拟机的主要特性

虚拟机具有以下特征，这些特征可提供多项优势。

分区

- 可在一台物理机上运行多个操作系统
- 可在虚拟机之间分配系统资源

隔离

- 可在硬件级别进行故障和安全隔离
- 可利用高级资源控制功能保持性能

封装

- 可将虚拟机的完整状态保存到文件中
- 移动和复制虚拟机就像移动和复制文件一样轻松

独立于硬件

- 可将任意虚拟机调配或迁移到任意物理服务器

虚拟化管理有哪些功能



统一管理

提供虚拟化环境的统一视图，集成 vSphere 并且可延展到整个软件定义的数据中心。



可控的自动化

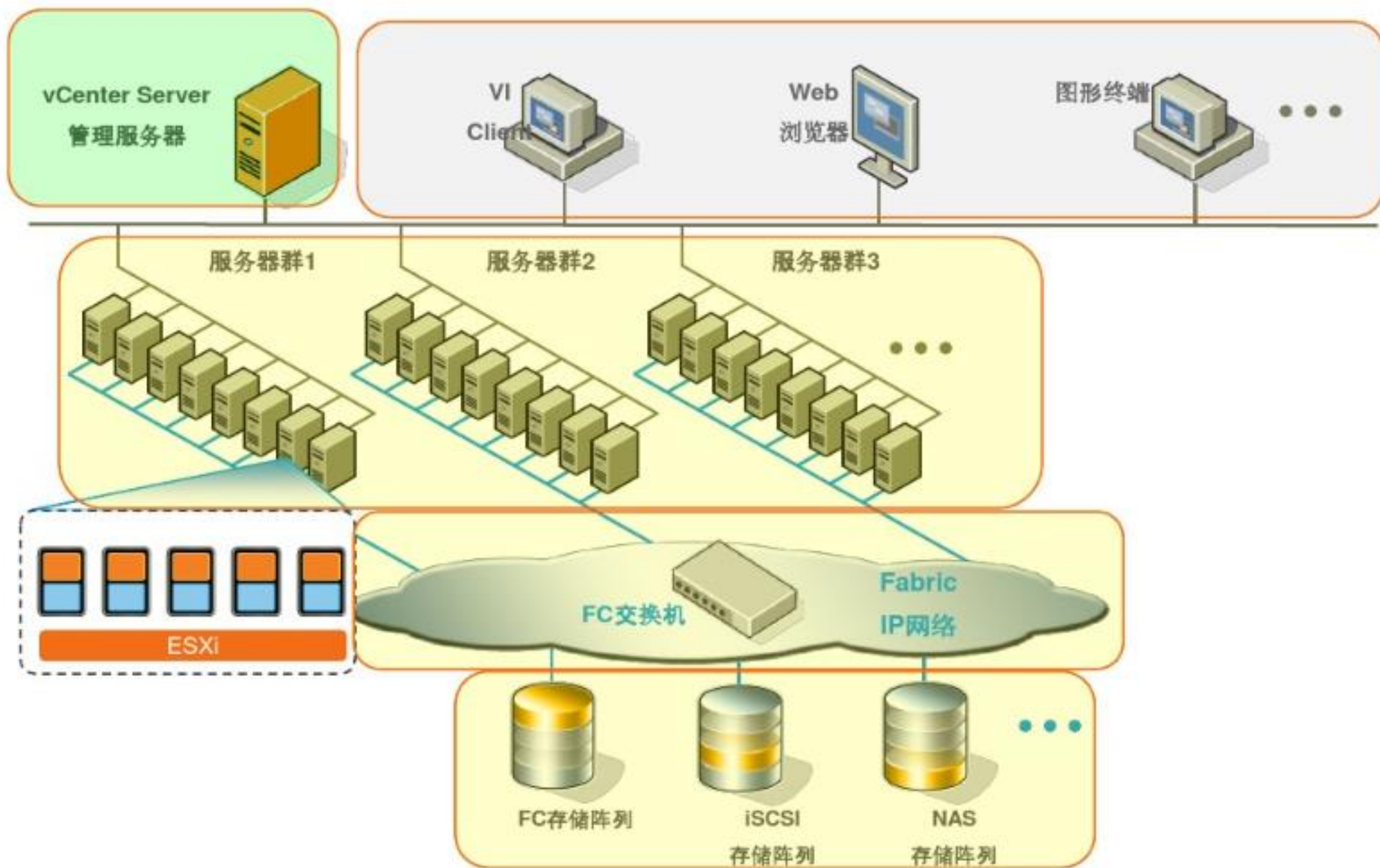
通过安全地自动执行各类管理任务，其中包括工作负载安置、重新均衡以支持基础架构并提升应用性能、优化资源以及强制执行 vSphere 安全强化准则，为战略项目腾出更多时间。



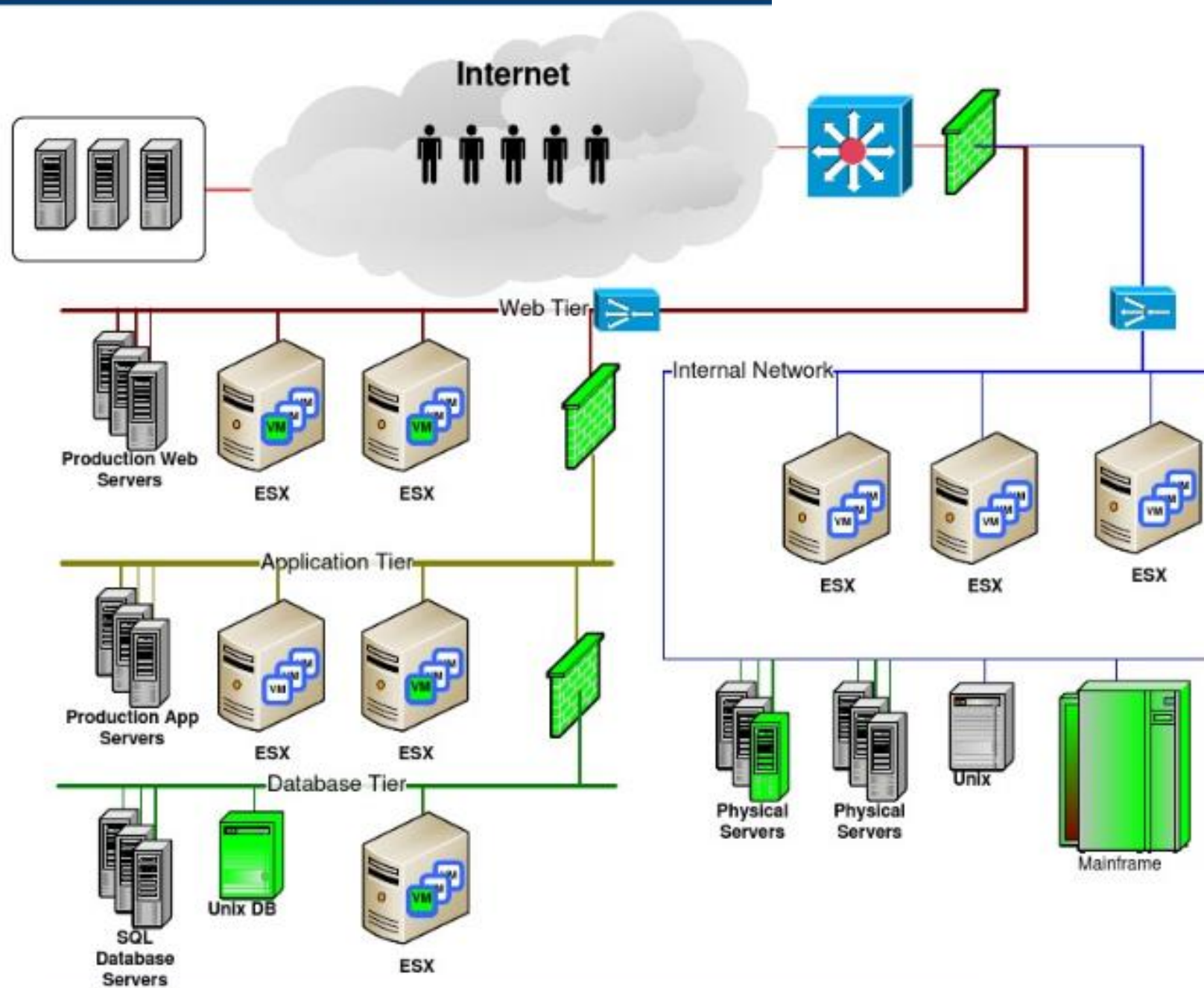
智能运维

自学式预测性分析功能会针对您的特定环境做出响应，并为您提供有意义的洞察信息和引导式修复。利用虚拟化环境中所有类型的 IT 数据（包括实时日志分析）来消除盲点。

VMware vSphere虚拟数据中心物理拓扑结构



虚拟化与传统数据中心环境的无缝集成



云基础架构平台



Thanks

FAQ