第05课：区块链...

大家好，我是丹华。在接下来几节，我们将分别介绍构成区块链技术的四大支柱：共识机制、密码学、分布式账本技术和智能合约。

5.1 区块链技术概述

区块链（Blockchain）是一系列现有成熟技术的有机组合，它对业务账本进行分布式的有效记录，并且提供完善的底层语言脚本以支持多样的业务逻辑。在典型的区块链系统中，数据以区块（block）为单位产生和存储，并按照时间顺序连成链式（chain）数据结构。所有节点共同参与区块链系统的数据验证、存储和维护。新区块的创建通常需得到全网多数（数量取决于不同的共识机制）节点的确认，并向各节点广播实现全网同步，之后不能更改或删除。

区块链的技术本质是一种去中心化、面向业务、跨主体、健壮与安全的分布式状态机，具有存储数据、共享数据、分布式、防篡改与保护隐私、智能合约等核心特征。基于这些特征，可以在不同参与者之间部署节点（具体可以采用公有链、联盟链或私有链模式），用区块链技术搭建一张社会化的业务流程与数据共享网络，从而以技术手段来解决跨主体之间存在的信任问题。

区块链技术是一个集成了多项技术成果的综合性技术集合，主要解决了交易的信任和安全问题。一般认为，其中有四项核心技术：共识机制、密码学、分布式账本技术和智能合约。当然，未来随着技术的演进，可能会有其他的技术被添加进来。

本节我们先介绍区块链技术的第一个支柱：共识机制。

5.2 为什么需要共识机制?

武侠小说里，汇聚在一起的各路英雄意见各不相同，群龙无首，大家听谁的呢？公认的办法是打，一轮轮打下来，谁武功天下第一，大家都听他的。这就是江湖上的共识机制，靠实力说话。

类似地，区块链是一个去中心化的电脑网络，各节点间如何就交易和数据状态达成一致性的认识，是网络要解决的首要问题。区块链通过一套共识机制来使得各节点之间形成普遍认可。所谓共识，是指多方参与的节点在预设规则下，通过多节点网络的交互，对数据、行为或流程达成一致的过程。共识机制是指定义共识过程的算法、协议和规则。

这一点在去中心化体系中至关重要，而且与中心化的体系完全不同。中心化的业务体系中，都有一个可信的中心，它就是权威，所有其他参与者必须参照它的账本。比如政府机关、银行、支付宝。

因此，区块链的共识机制应具有民主自治的特点，包括少数服从多数以及人人平等。其中少数服从多数不一定指节点个数，也可以是计算能力、股权数或者其他特征量。人人平等是指所有节点都有权优先提出共识结果、直接被其他节点认同后并最后有可能成为最终的全网共识结果。

更进一步地，共识协议用于在分布式系统中实现可用性与一致性，其核心指标包括共识协议的强壮性（容错、容恶意节点的能力）、高效性（收敛速度，也即系统达成一致性或"稳态"的速度）及安全性（协议抽象理论模型的安全边界）。代表性协议包括 POW 共识、POS 共识、PBFT 共识及混合共识等。

5.3 共识机制的分类

按节点属性，共识算法分为两类，可信节点间的共识算法与不可信（无信任或任意）节点间的共识算法。前者已经被深入研究且广泛应用。不可信节点间的共识算法，是在区块链出现以后才得到广泛研究和应用的。

根据应用场景的不同，共识算法又可以分为两大类，一类算法适用于公有链，包括以工作量证明 PoW（Proof of Work），权益证明 PoS（Proof of Stake）等算法。

另一类适用于联盟链或私有链，以实用拜占庭容错算法 PBFT（Practical Byzantine Fault Tolerance）及其变种算法为代表。提示：关于公有链、联盟链和私有链的概念，因本节容量不够，将移至下一节介绍。

5.4 工作量证明 POW 算法

工作量证明 POW 算法是比特币和以太坊采用的共识算法，该算法于1998年由 W. Dai（戴伟）在 B-money 的设计中提出。

比特币的工作量证明是寻找满足特定难度值的区块头哈希（参考第二节比特币中介绍的挖矿机制），并引入了经济激励，设计了随着区块生成而持续滚动的无限次投票流程：

- 任何人都可以生成一个包含交易的新区块(增加账本数据)并广播；
- 必须完成足够难度的大量随机计算才能获得区块记账权；
- 其他人如果同意该区块纳入账本，则将该区块的哈希作为自己构造的区块数据的一部分，以对该区块进行"确认"；
- 对某个区块的"确认"也包含了对该区块前序所有区块的"确认"。

也有一类共识机制的思路是设计内存消耗型算法，比如 Ethereum 基于 Dagger-Hashimoto 的 Ethash，Zcash 基于广义生日悖论问题的 Equihash 等。这类算法在计算时需要占用大量内存，而内存作为成熟产品优化空间小，设计专用 ASIC 芯片的成本优势不大。

5.5 权益证明 PoS（Proof of Stake）算法

权益证明 PoS（Proof of Stake）算法最早由 Sunny King 在2012年的点点币 PPC（Peer To Peer Coin）系统中首先实现。PoS 及其变种算法可以解决 PoW 算法一直诟病的浪费算力、运行成本高问题，但其本身尚未经过足够验证。PoS 协议下，节点获得区块创建权的概率取决于该节点在系统中所占有的权益比例的大小。简单说，PoS 机制中，你拥有的币越多，你就能够获得更多的区块奖励。以太坊目前正在计划以 PoS+PoW 机制来代替 PoW。

PoS 一般需要用户时刻在线，这对应用带来了很大挑战。为了解决这个问题，衍生出了 DPoS（Delegated Proof of Stake）共识，其核心思想是从先从全网节点中选出部分节点，保证这些节点的有效性，然后在该子节点集合内进行 PoS 共识。

5.6 PBFT 算法

BFT（Byzantine Fault-Tolerant）算法于20世纪80年代开始被研究，旨在解决所谓拜占庭将军问题。拜占庭将军问题（Byzantine failures）是指，拜占庭帝国军队的将军们必须全体一致，以决定是否攻击某一支敌军。问题是这些将军在地理上是分隔开来的，并且将军中存在叛徒。BFT 类算法中最著名的是 PBFT 算法。

PBFT 算法最早由卡斯特罗（Miguel Castro）和利斯科夫（Barbara Liskov）在1999年提出，该算法运行效率更高。假设系统中共有 N 个节点，那么 PBFT 算法可以容忍存在不多于1/3比例的恶意节点。

BFT 类共识随着参与共识节点的增加，通信开销会急剧上升，达成共识的速度则快速下降，难以支撑上万节点规模的分布式系统，一般 PBFT 共识系统中节点数很少超过100个。此外，节点参与共识首先要获得投票权，因此要为节点的加入和退出过程设计额外的机制，增加了协议复杂度和实现难度。

PBFT 的优点是收敛速度快、节省资源、具有理论上的安全界（理论上允许不超过1/3的恶意节点存在）。

5.7 共识算法小结

无论是 PoW 算法还是 PoS 算法，其核心思想都是通过经济激励来鼓励节点对系统的贡献和付出。为了鼓励更多节点参与网络共识，公有链通常会发放代币（token）以激励节点的付出。

联盟链一般没有代币激励机制，联盟链的节点通常更愿意获得可信数据、获得业务合作关系等。通常，联盟链的参与节点数较少，节点信任度相对较高，因此 PBFT 类算法正好适用于联盟链或私链的应用场景。

长远来看，共识算法也在快速进化中，未来可能出现更好的共识算法机制。

下一节，我们将介绍区块链技术的另一个重要构成：密码学。