China Blockchain Conference

区块链—形式化表示与体系结构演进

（浙江·杭州 中国区块链技术大会）

斯雪明

复旦大学 教授

中国计算机学会区块链专委会 主任

目录

01

传统信息系统的数学模型与特点

信息系统定义

信息系统是一个人造系统，它由人、硬件、软件和数据资源组成，目的是及时、正确的收集、加工存储、传递和提供信息，实现组织中各项活动的管理、调节和控制

信息系统形式化表示

一个信息系统 S 可以表示为： $S=\{U, A, V, f, g\}$

- $U=\{x_1, x_2, \ldots, x_n\}$ 称为论域,代表信息系统的不同模块(网络、数据…)
- $A=\{a_1,a_2,\ldots,a_m\}$是属性集
- $V=\cup_{a\in A}V_a$,V_a称为属性 a 的值域
- $f_i: x_i \rightarrow V$ 是信息函数

$$x_i \rightarrow \left(a_{11}^i, a_{12}^i, \cdots, a_{1m_i}^i\right)$$

- $g_i(t): f_i \rightarrow f_i$ 是单域状态转换算子

$$\left(x_i \rightarrow \left(a_{11}^i, a_{12}^i, \cdots, a_{1m_i}^i\right)\right) \rightarrow \left(x_i \rightarrow \left(a_{t1}^i, a_{t2}^i, \cdots, a_{tm_i}^i\right)\right)$$

- $g(t): x_i \rightarrow x_i$是单域状态转换函数

$$x_i \rightarrow \begin{bmatrix} a_{11}^i & a_{12}^i & \cdots & a_{1m_i}^i \\ a_{21}^i & a_{22}^i & \cdots & a_{2m_i}^i \\ \vdots & \vdots & \ddots & \vdots \\ a_{t1}^i & a_{t2}^i & \cdots & a_{t,m_i}^i \end{bmatrix}$$

信息系统形式化表示

- $G(t):U\rightarrow U$是系统状态转换函数

$$U\{x_1,x_2,\cdots,x_n\}\rightarrow\begin{bmatrix} f_1(x_1) & f_2(x_2) & \cdots & f_{m_1}(x_{m_1}) \\ g_1(f_1(x_1)) & g_2(f_2(x_2)) & \cdots & g_{m_1}(f_{m_1}(x_{m_1})) \\ g_1(g_1(f_1(x_1))) & g_2(g_2(f_2(x_2))) & \cdots & g_{m_1}(g_{m_1}(f_{m_1}(x_{m_1}))) \\ \vdots & \vdots & \ddots & \vdots \\ (g_1)^t(f_1(x_1)) & (g_2)^t(f_2(x_2)) & \cdots & (g_{m_1})^t(f_{m_1}(x_{m_1})) \end{bmatrix}$$

传统信息系统的静态性

考虑传统信息系统的某一论域时，属性取值在时间轴上**稳定不变**

例如对域 i 的某个属性值是不变的

$$g(x_i) \rightarrow \begin{bmatrix} a_{11}^i & a_{12}^i & \cdots & a_{1m_i}^i \\ a_{21}^i & a_{22}^i & \cdots & a_{2m_i}^i \\ \vdots & \vdots & \ddots & \vdots \\ a_{t1}^i & a_{t2}^i & \cdots & a_{t,m_i}^i \end{bmatrix}$$

比如第一个属性是**静态**的，则有 $a_{11}^i = \cdots = a_{t1}^i$ 成立

传统信息系统的相似性

对于两个子系统$S=\{U, A, V, f, g\}$ 和$S'=\{U, A, V, f, g\}$

定义S和 S' 之间的**相似度**:

$$d(S,S') = \sum_t \frac{\sum_{i=1}^{\max_{j=1}^{n}(m_j)} d\left(\left(g_i^S\right)^t\left(f_i^S\left(x_i^S\right)\right),\left(g_i^S\right)^t\left(f_i^{S'}\left(x_i^{S'}\right)\right)\right)}{n \cdot t}$$

$d(S,S')$越小表示两个系统越相似，对于传统的计算机信息系统，这两个系统的$d(S,S')$值往往很小

传统信息系统的中心化

传统信息系统具有**中心化**的特点，中心（特权用户）**独自**可以进行数据的**增、删、改、存**等操作，也可以执行服务的**发布**与**撤回**等。而普通用户与特权用户的权限有很大的差别。中心化导致系统的**机密性**、**完整性**、**可用性**容易受到**内部人**或**外部黑客**的破坏。

传统信息系统的特点

静态性、相似性、中心化

传统信息系统各维度通常是**静态的**，数据通常**无冗余**或**少冗余**，各信息系统之间通常是**相似的**，具有**中心化**的特点，操作权限和数据的增删改权限总是**集中的**

02

区块链系统的形式化表示

区块链产生过程

发起交易 → 传播交易 → 验证交易 → 添加区块

- **发起交易**，该交易可能涉及Token，智能合约，记录或其他信息
- 在节点的帮助下，该交易被**广播**到P2P网络中
- 借助已知的算法，网络中节点**验证**交易的合法性和用户状态
- 交易一旦完成，新的**区块**被添加到**区块链**上，从此永久**不可篡改**

区块链形式化表示

消息单元

Bool M($parameter_1$,$parameter_2$,...,$parameter_n$,sig)

environment

共识

打包

Ver(M)=True?→Block{}+M

Bool PROVE(...*parameter*)

Bool CON(PROVE,Block{...*parameter*,M_1,...,M_k})

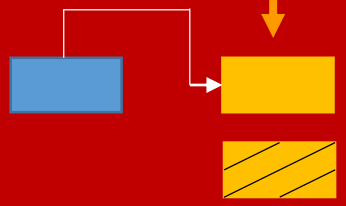
同步

Ver(CON)=False?

Ver(CON)=True?→Reset PROVE

存储

数据存储



状态存储

$state_r\{a_1(s_1),...,a_m(s_m)\} \xrightarrow{Block} state\{a_1(s_1'),...,a_m(s_m'),?\}$

STATE{}+state

以比特币为代表的单链非许可链

- **比特币：**
  - 消息单元以普通转账交易为主(交易双方地址、输入输出金额)
  - 消息打包前，需要工作量证明(PoW)取得记账权
  - 算力决定了进行PoW的能力
  - PoW的验证与区块的传播异步，可能存在分叉
  - 状态以UTXO为主，传递模式
- **以太坊**
  - 智能合约
  - 世界状态

比特币形式化表示

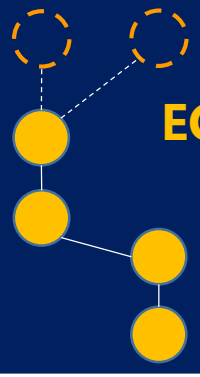
消息单元

Bool M(add_{sen},add_{rec},$value_{sen}$,$value_{rec}$,*ECDSA*)

environment

共识

打包



ECDSA(M)=True?→Block{}+M

PoW值：H(ctr,G(s,x))

同步

H(ctr,G(s,x))>=D?

H(ctr,G(s,x))<D?→下一轮打包操作

存储

数据存储

状态存储

$UTXO_r(utxo_1,\ldots,utxo_m) \longrightarrow UTXO_{r+1}(utxo_1,\ldots,utxo_m,utxo_{m+1})$

以fabric为代表的单链许可链

- CA(Certification Authority)作为所有节点加入系统的认证
- 普通客户端通过部署链码(chaincode)、调用链码发起交易
- 通过背书策略申请peer节点(Endorse node)的背书签名
- Order节点验证客户端节点的交易和背书消息，转发给peer节点(Commit node)，通过Kafka(Zookeeper)对交易进行排序
- Commit node通过CouchDB(或者LevelDB)存储区块信息和实时状态以及世界状态

fabric形式化表示

消息单元

Bool M(type,data,payload{channel_header{type,version,timestamp,channel_id,tx_id,extension,...}}*ECDSA*)

environment **CA**

共识

打包

Very(M,endorse nodes,endorse policy)=True?
→Block{}+M

Block(previous_hash,data_hash,number,{M_1,...,M_k},ECDSA)

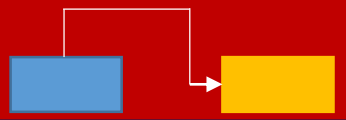
同步

Very(Block,ECDSA)=False?

Very(Block,ECDSA)=True?→下一轮打包操作

存储

数据存储 { }←Block



状态存储

$state_r\{a_1(s_1),...,a_m(s_m)\} \xrightarrow{Block} state\{a_1(s_1'),...,a_m(s_m'),?\}$

STATE{}+state

以IOTA为代表的有向无环图

- **有向无环图**中(Directed Acyclic Graph,DAG)无传统区块概念
- 交易**并发**进入系统
- 交易产生，同时对已有交易进行验证
- 消息产生后**直接进入**Tangle系统
- 存在**中心化**倾向

IOTA形式化表示

消息单元

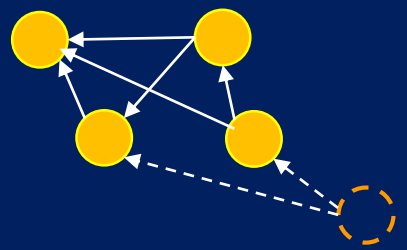
Bool M(version,parent_hash$_1$,parent_hash$_2$,messages,Pow_proof,*sig*)

environment

共识

打包

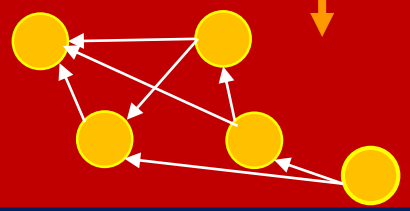
Very(M)=True?→Tangle+M



同步

存储

数据存储



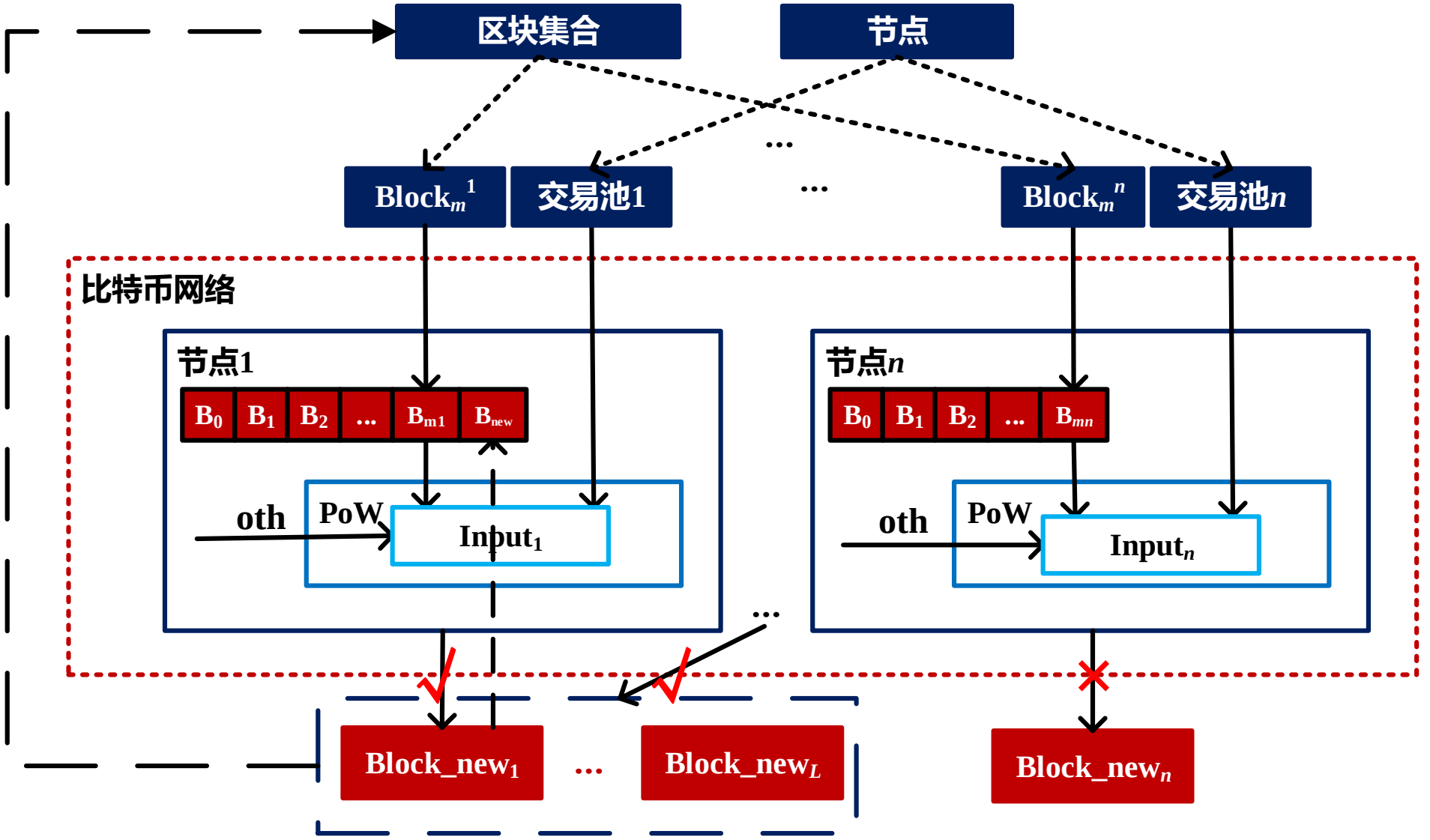
状态存储

$$state_r\{a_1(s_1),...,a_m(s_m)\} \xrightarrow{M} state\{a_1(s_1'),...,a_m(s_m'),?\}$$

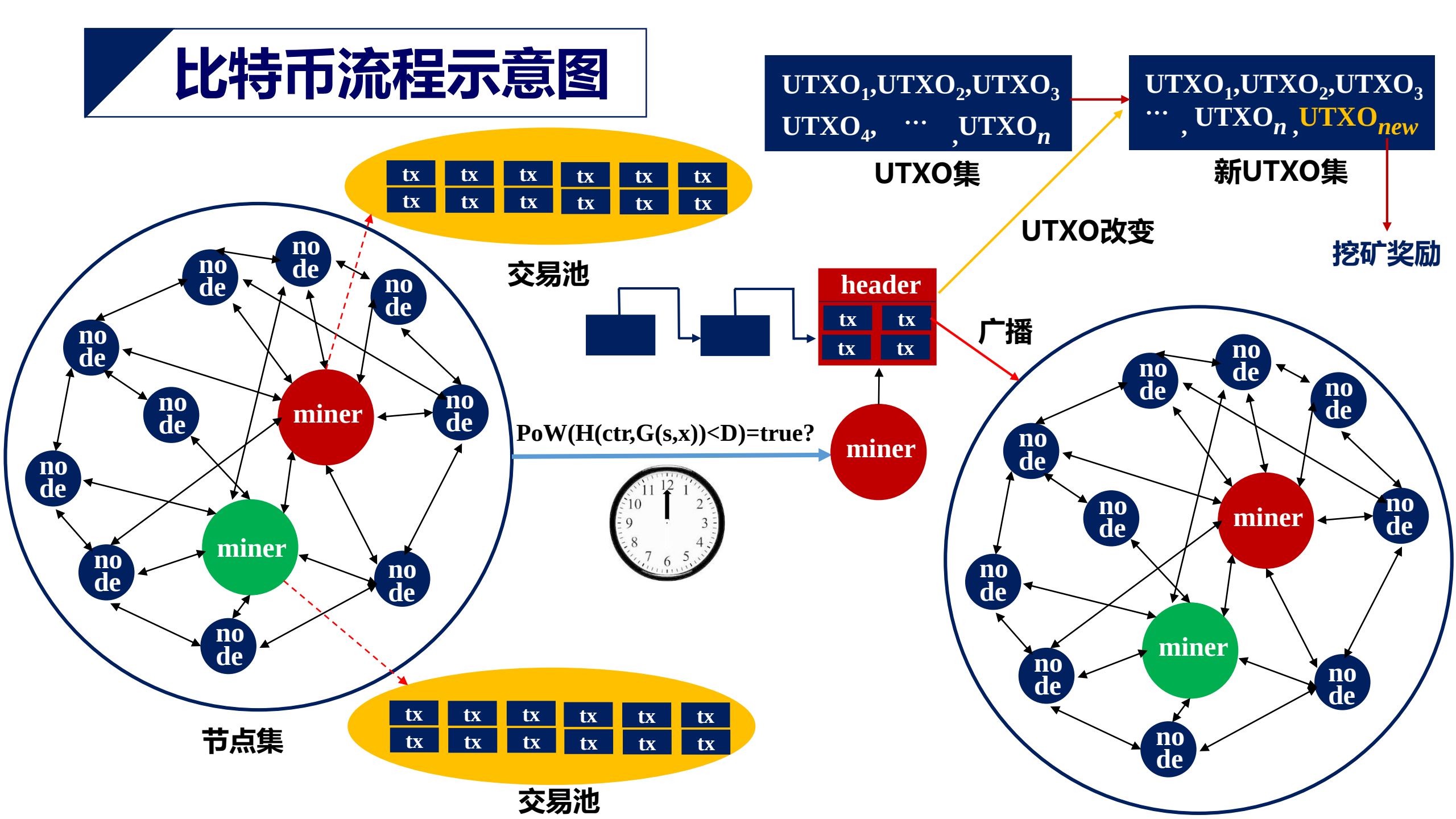
03

典型区块链体系结构

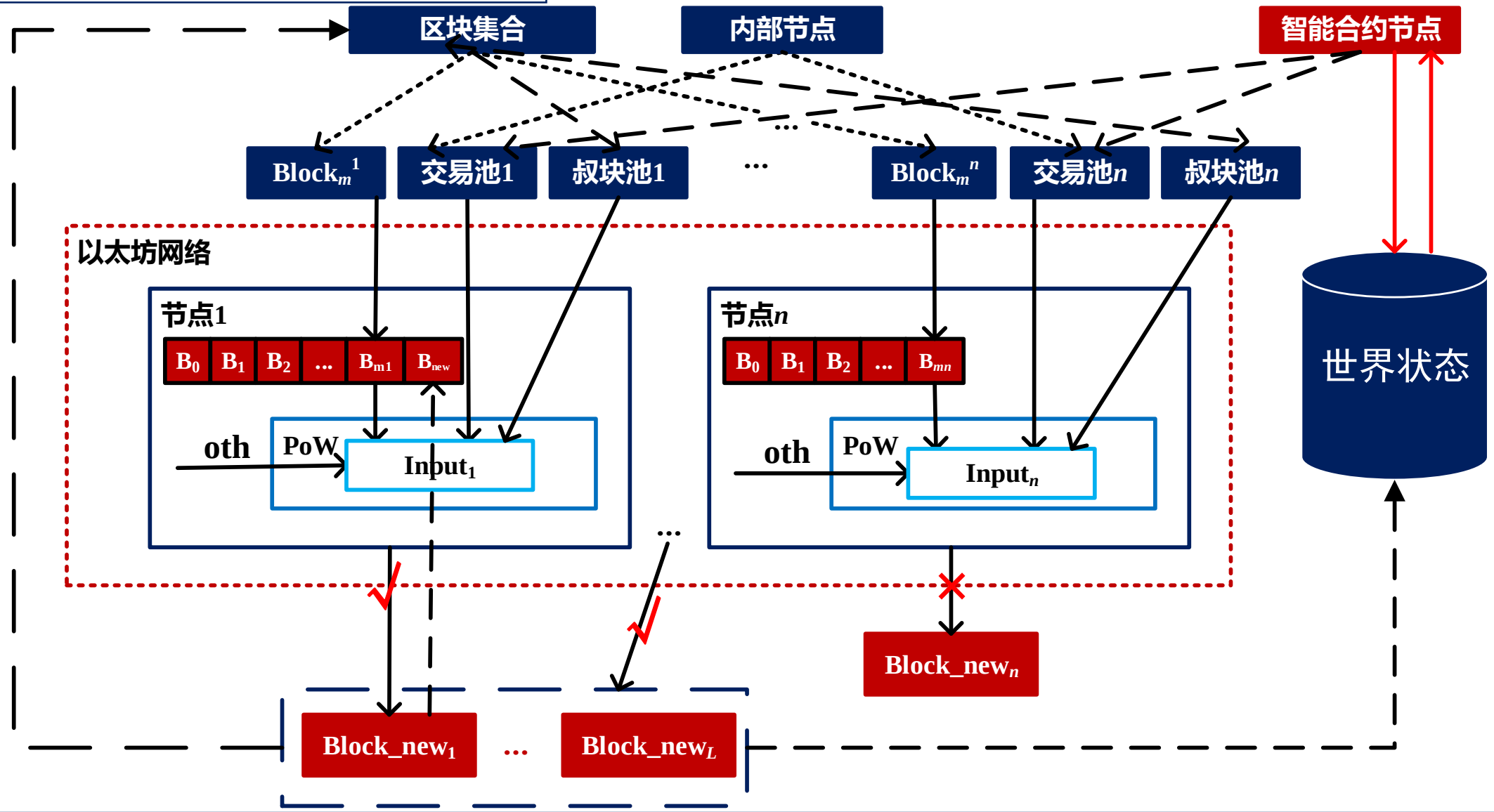
比特币体系结构


区块集合
节点
Block_m^1
交易池1
Block_m^n
交易池n
比特币网络
节点1
B0 B1 B2 ... Bm1 Bnew
oth
PoW
Input1
节点n
B0 B1 B2 ... Bmn
oth
PoW
Inputn
Block_new1
Block_newL
Block_newn

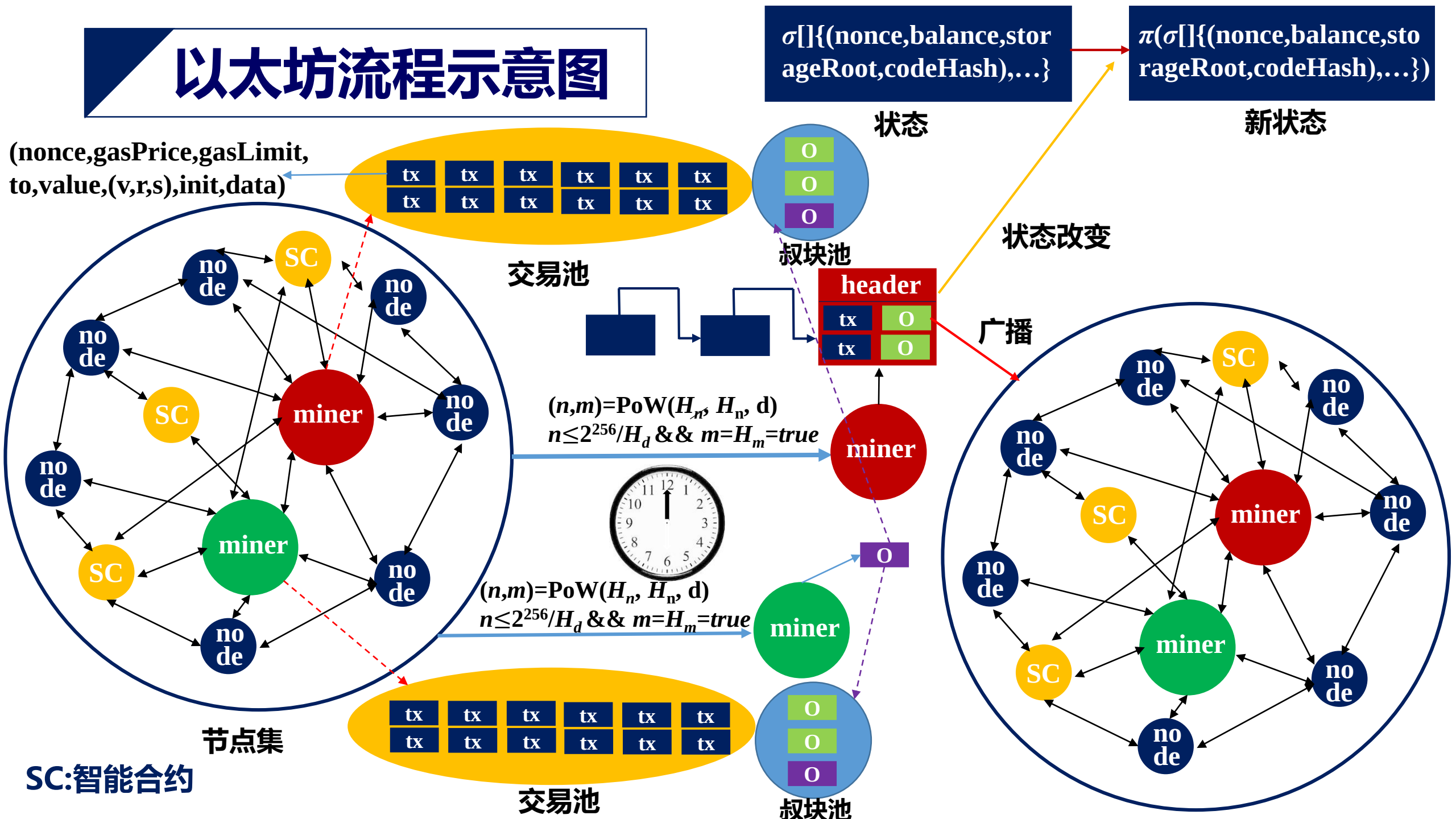
比特币流程示意图


UTXO1,UTXO2,UTXO3
UTXO4, … ,UTXOn
UTXO集
UTXO1,UTXO2,UTXO3
… , UTXOn , UTXOnew
新UTXO集
UTXO改变
挖矿奖励
tx
交易池
header
广播
node
miner
PoW(H(ctr,G(s,x))<D)=true?
节点集
交易池

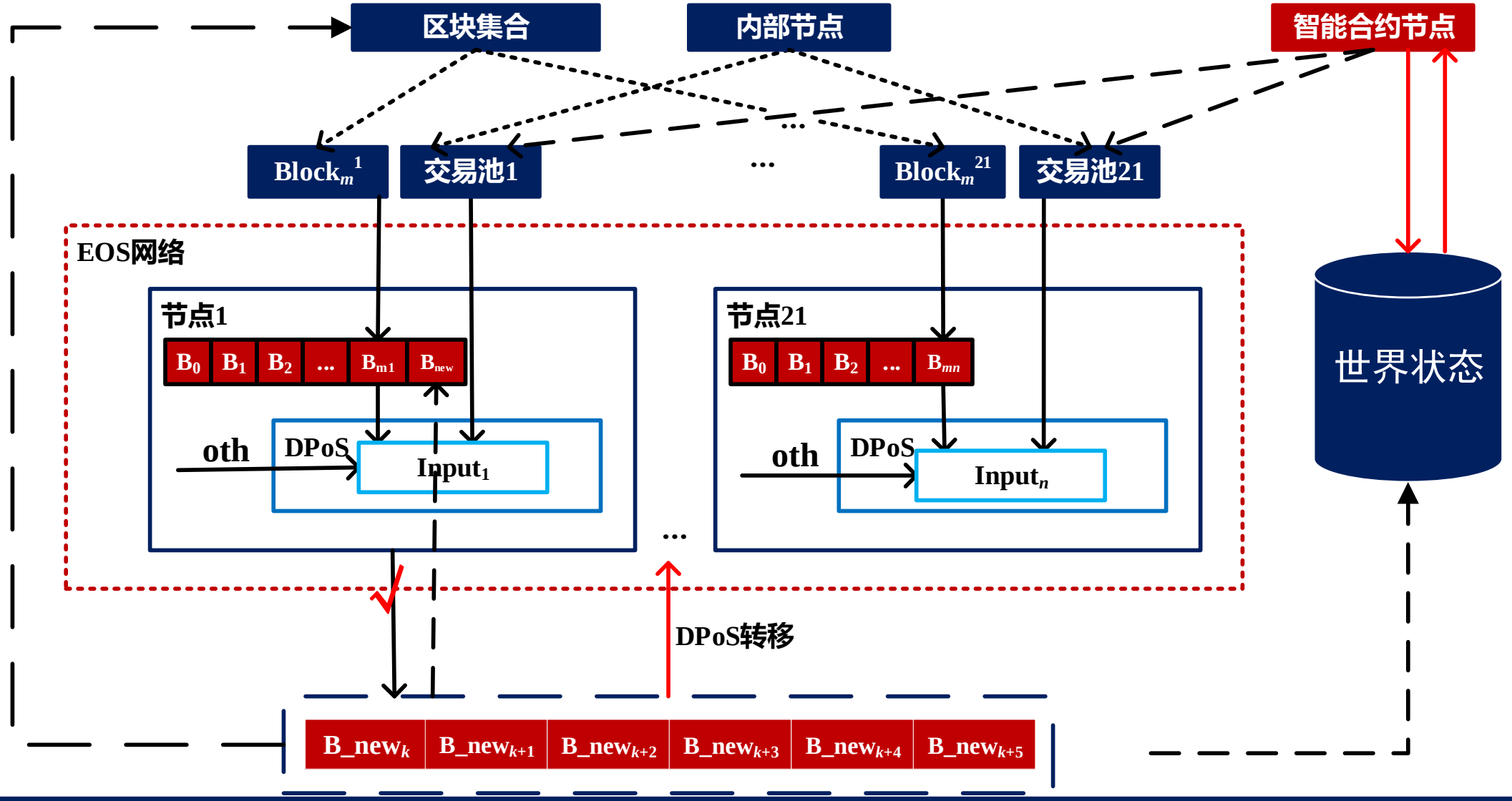
以太坊体系结构


区块集合
内部节点
智能合约节点
Block_m^1
交易池1
叔块池1
...
Block_m^n
交易池n
叔块池n
以太坊网络
节点1
B0 B1 B2 ... Bm1 Bnew
oth
PoW
Input1
节点n
B0 B1 B2 ... Bmn
oth
PoW
Inputn
世界状态
Block_new1
...
Block_newL
Block_newn

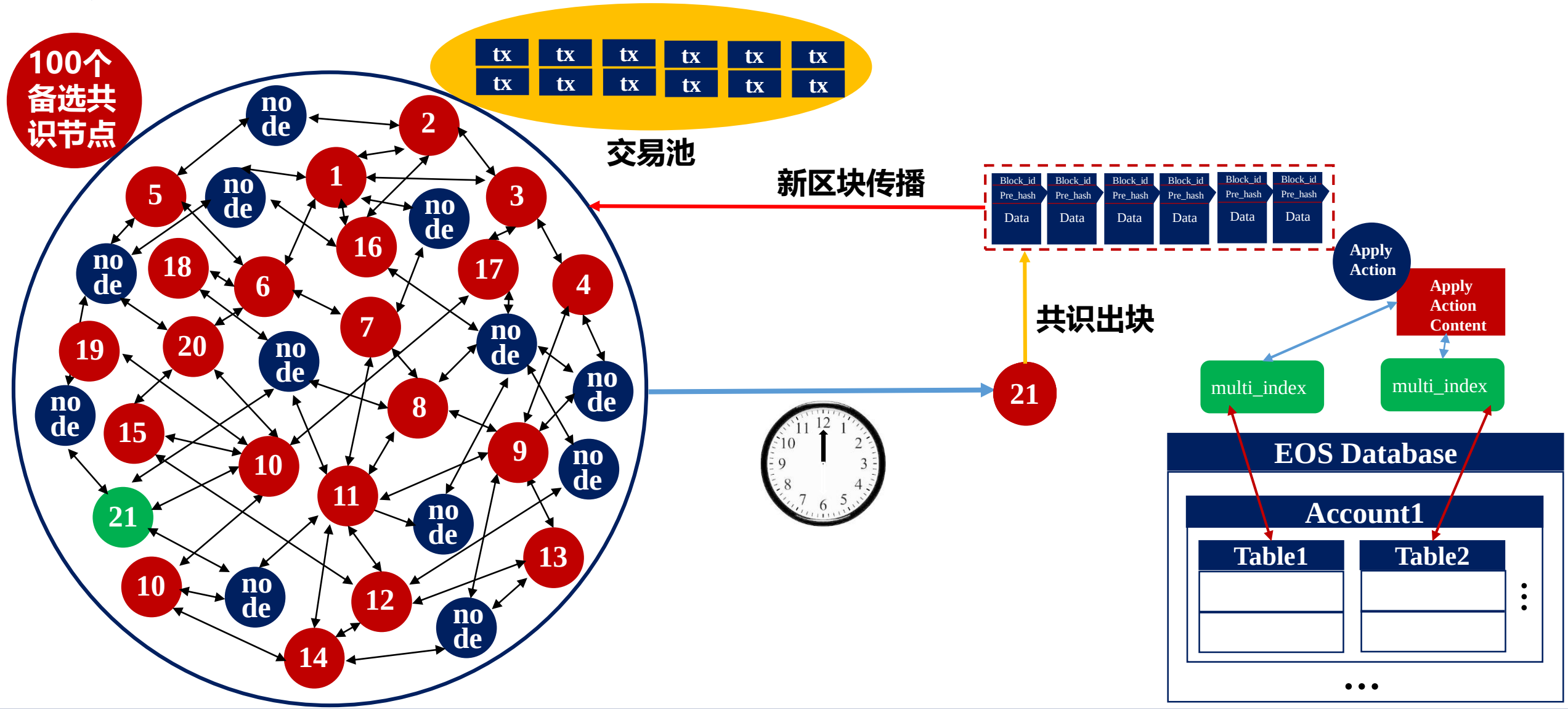
以太坊流程示意图


σ[]{(nonce,balance,storageRoot,codeHash),…}
π(σ[]{(nonce,balance,storageRoot,codeHash),…})
状态
新状态
(nonce,gasPrice,gasLimit,
to,value,(v,r,s),init,data)
tx
交易池
叔块池
状态改变
header
广播
(n,m)=PoW(H_n, H_n, d)
n≤2^256/H_d && m=H_m=true
miner
SC
node
节点集
SC:智能合约

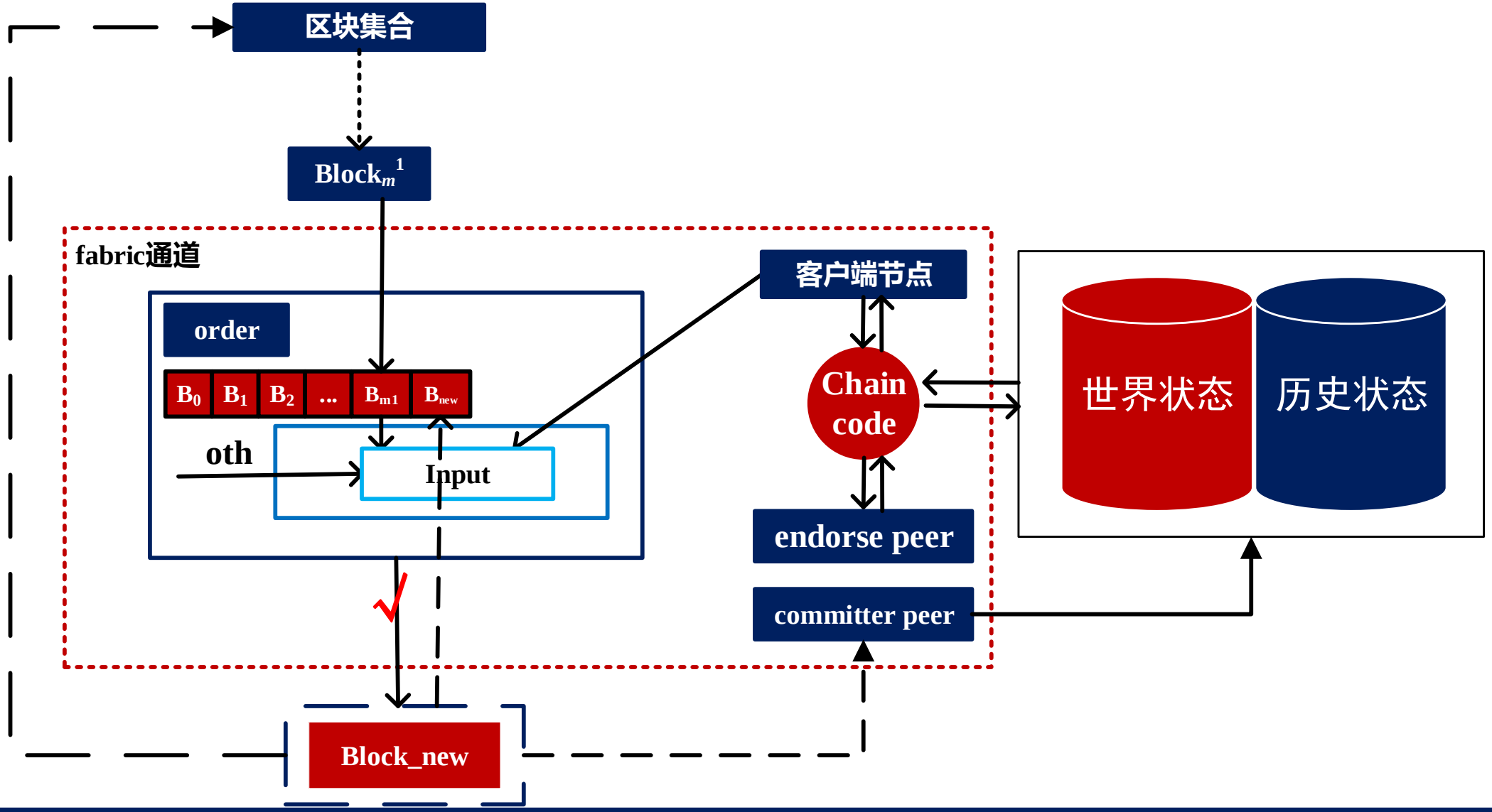
EOS体系结构


区块集合
内部节点
智能合约节点
Block$_m^1$
交易池1
...
Block$_m^{21}$
交易池21
EOS网络
节点1
B$_0$ B$_1$ B$_2$ ... B$_{m1}$ B$_{new}$
oth
DPoS
Input$_1$
节点21
B$_0$ B$_1$ B$_2$ ... B$_{mn}$
oth
DPoS
Input$_n$
世界状态
DPoS转移
B_new$_k$ B_new$_{k+1}$ B_new$_{k+2}$ B_new$_{k+3}$ B_new$_{k+4}$ B_new$_{k+5}$

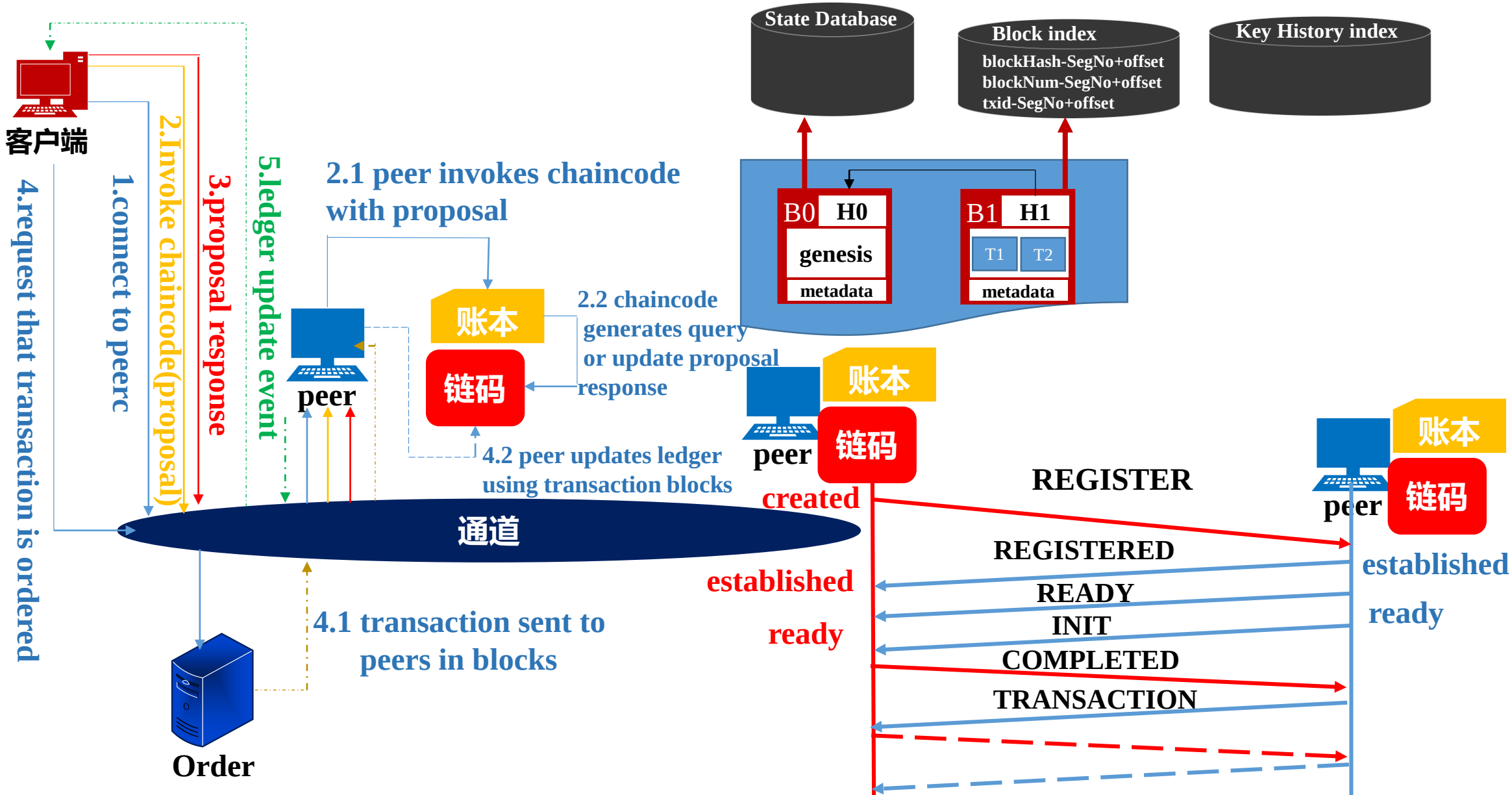
EOS流程示意图


100个备选共识节点
tx tx tx tx tx tx
tx tx tx tx tx tx
交易池
新区块传播
Block_id Pre_hash Data
共识出块
Apply Action
Apply Action Content
multi_index
multi_index
EOS Database
Account1
Table1
Table2

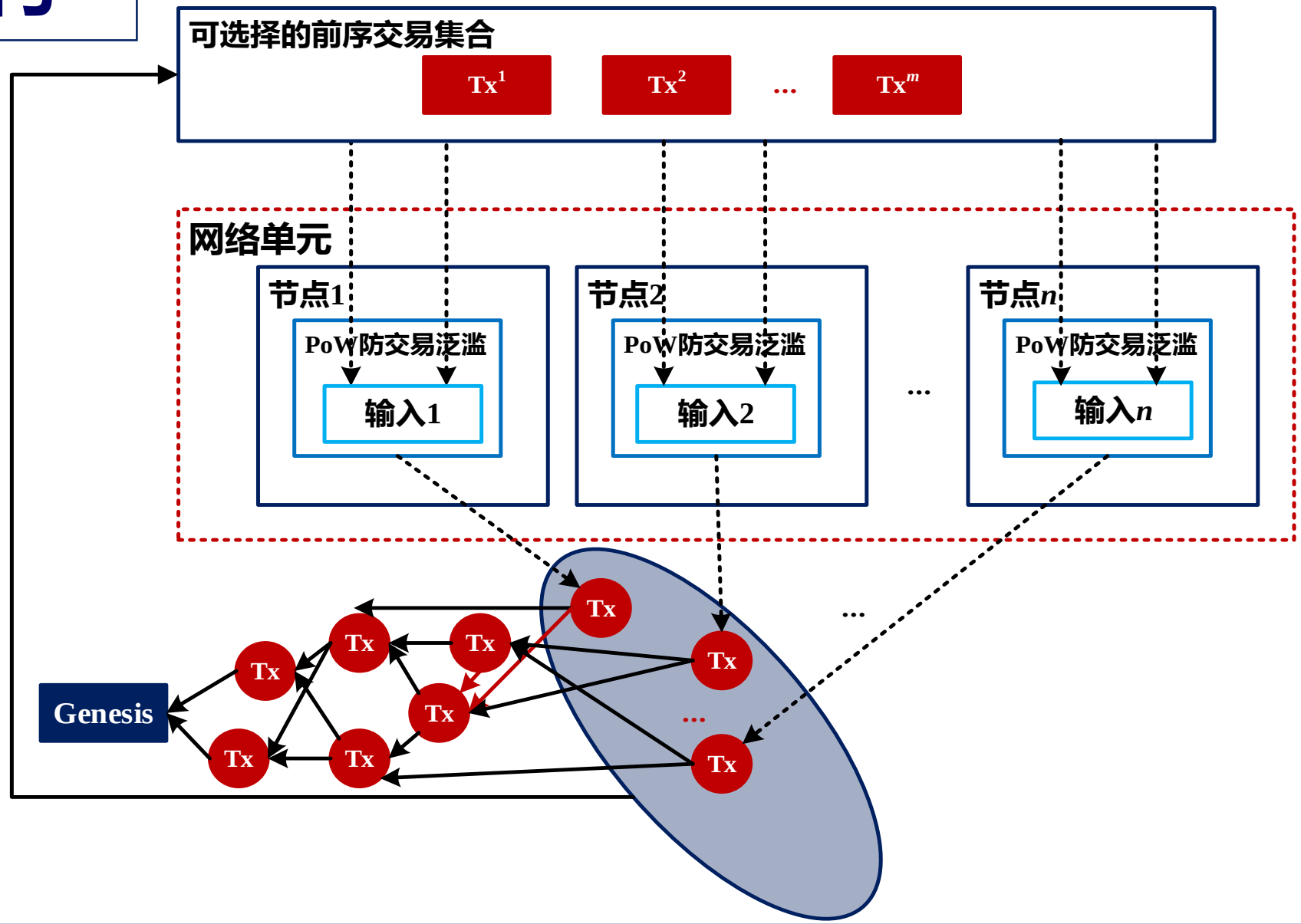
Fabric1.2体系结构



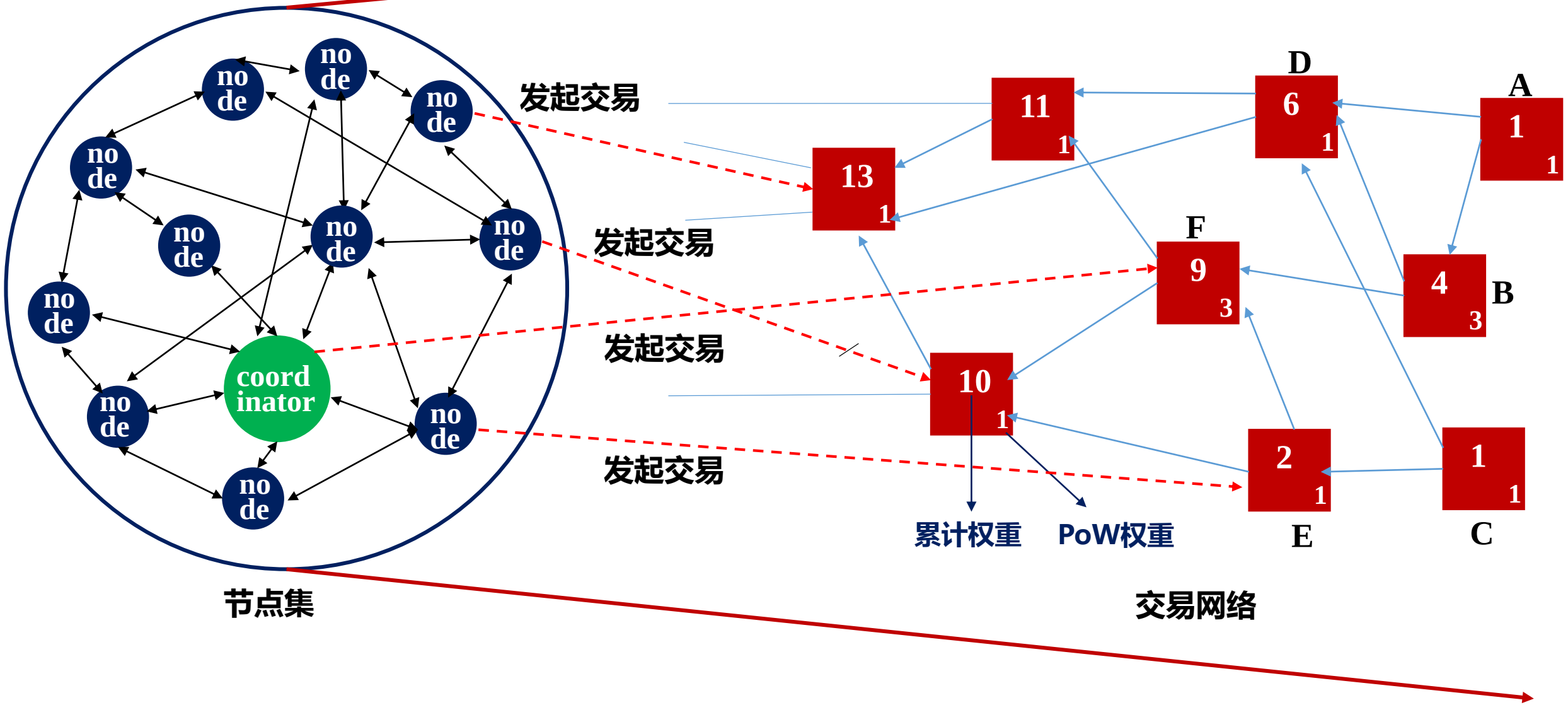
Fabric1.2流程示意图



IOTA体系结构


可选择的前序交易集合
Tx1
Tx2
...
Txm
网络单元
节点1
PoW防交易泛滥
输入1
节点2
PoW防交易泛滥
输入2
...
节点n
PoW防交易泛滥
输入n
Genesis
Tx
Tx
Tx
Tx
Tx
Tx
Tx
Tx
...
...

IOTA流程示意图


node
coordinator
发起交易
发起交易
发起交易
发起交易
11
1
13
1
6
1
D
A
1
1
F
9
3
4
3
B
10
1
2
1
1
1
累计权重
PoW权重
E
C
节点集
交易网络

区块链系统的特点

静态性、多中心、分布式、冗余、高度同构

目前的区块链信息系统通常也是**静态的**，具有**多中心**的特点，数据**分布式**存储和共享，在许可链中，各节点数据存在一定的**冗余**，在非许可链中，**高度冗余**，各节点运行的系统高度**同构**。

04

区块链体系结构可能的演进方向

区块链体结构演进方向

静态性
相似性
中心化

传统信息系统特点

静态性
多中心
分布式
冗余
高度同构

典型区块链系统特点

动态性
多中心
分布式
适度冗余
异构性
群体性

理想区块链系统特点

区块链体结构演进方向

如何构建理想的区块链体系结构?

我们从自然界得到启示

自然界的启示一:拟态现象

- **拟态现象**是一种生物在形态、行为等特征上模拟另一种生物或环境，从而使一方或双方**受益**的生态适应现象
- 典型代表：**拟态章鱼**

拟态伪装大师—拟态章鱼

本征功能不变条件下， 以不确定的色彩、纹理、形状和行为变化给攻击者***造成认知困境***

自然界的启示二：群飞和群游现象

- **群飞**又叫聚集式飞行，鸟群能够在快速飞行的过程中整齐划一地转向，其目的是能够免遭老鹰等捕食者的袭击
- 典型代表：椋鸟
- **群游**是一种鱼群的集体行为，鱼群能够通过对周围环境变化做出迅速感知和统一变向，不仅可以有效抵御捕食者，还能够提高捕食效率和游动速度

群飞现象

群游现象

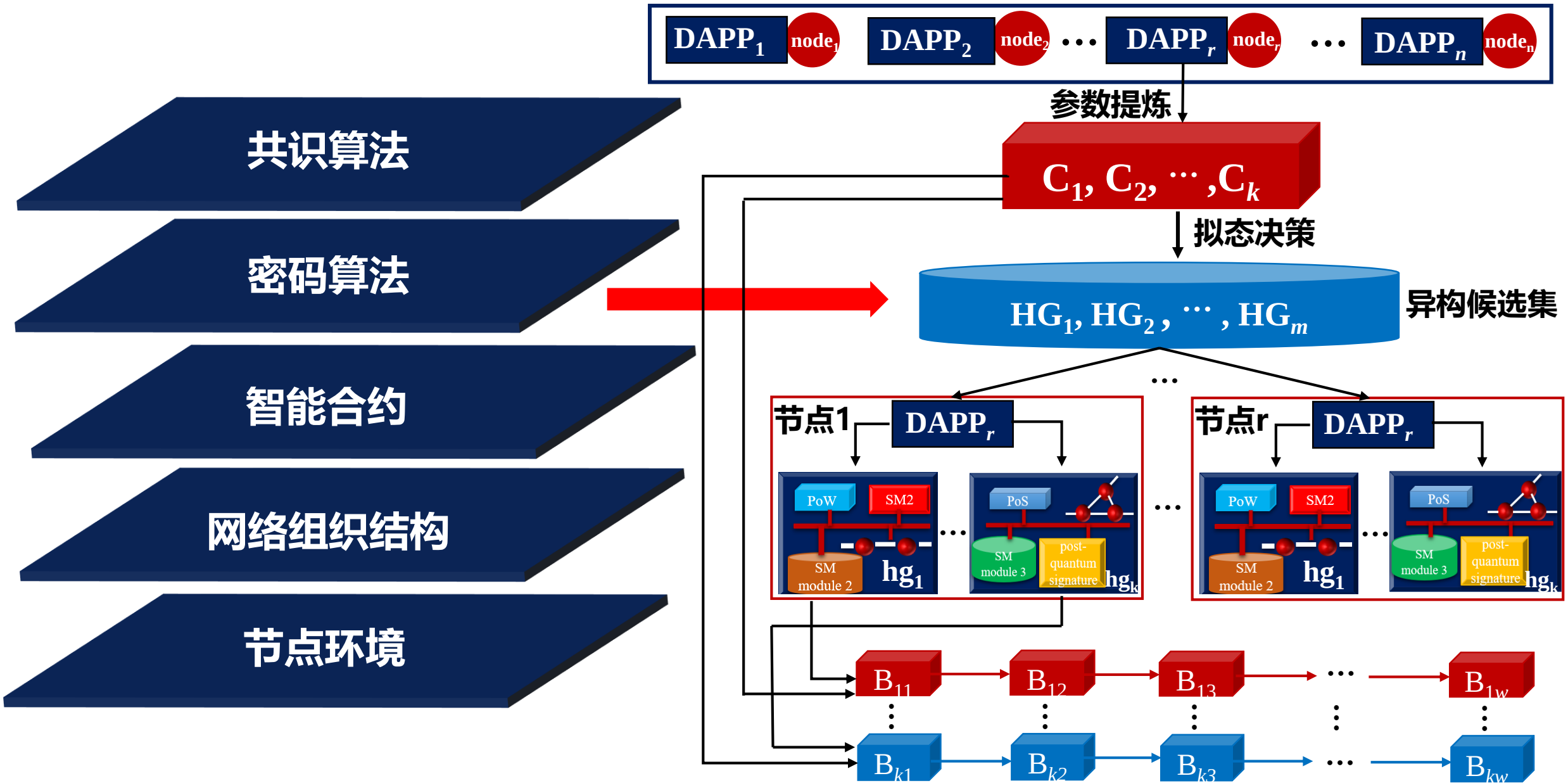
区块链体结构演进方向

根据上述生物学中的拟态现象和群飞、群游现象，我们提出了拟态群体异构体系结构，简称MCH体系结构

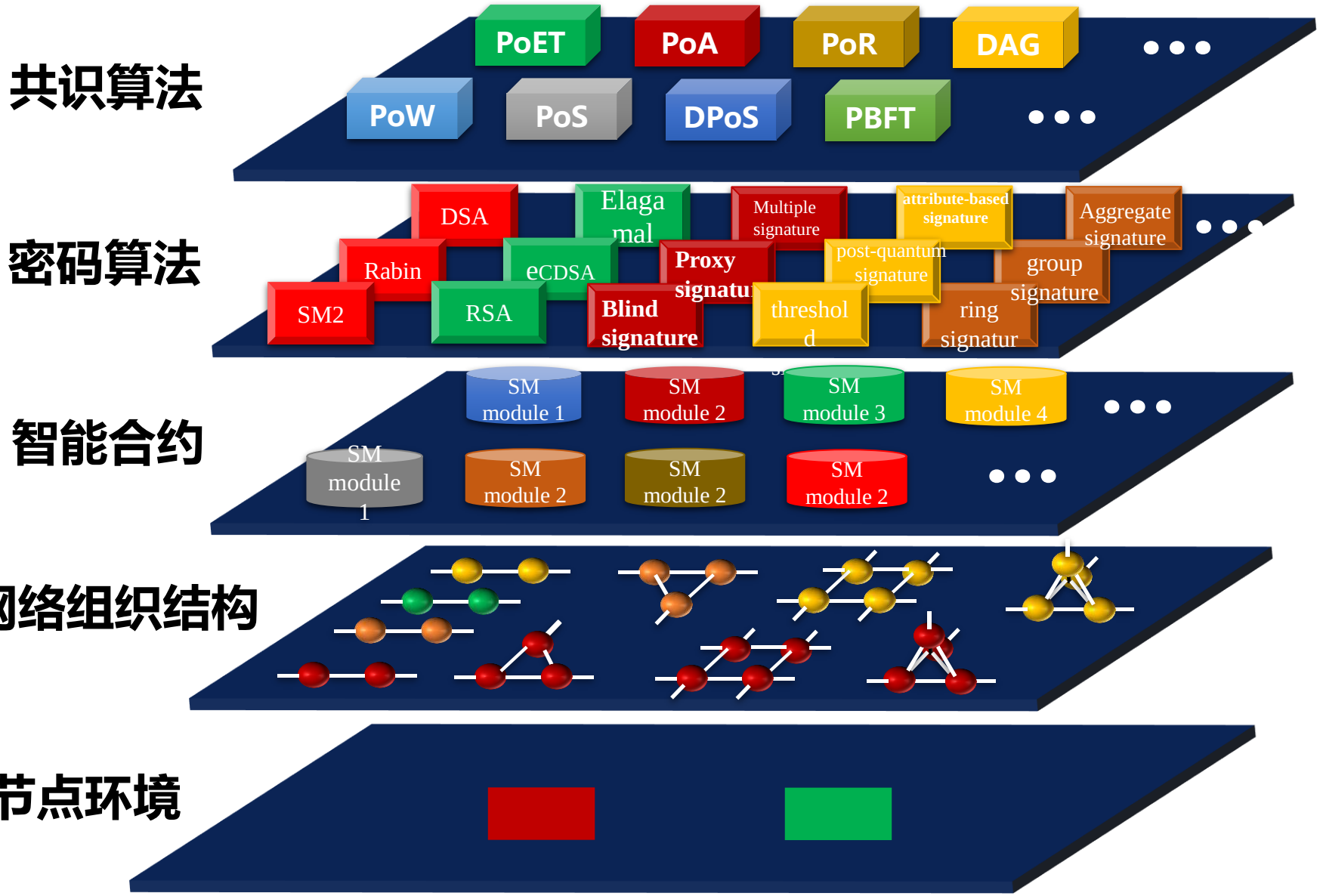
- 区块链系统中的部件可以根据应用场景动态调整变化
- 不同部件的不同实现方式，完成统一的目的
- 系统的部件异构，构建最适合应用的区块链系统

针对特定的应用场景，选择合适的异构部件，构建最优的体系结构，达到理想的应用目标

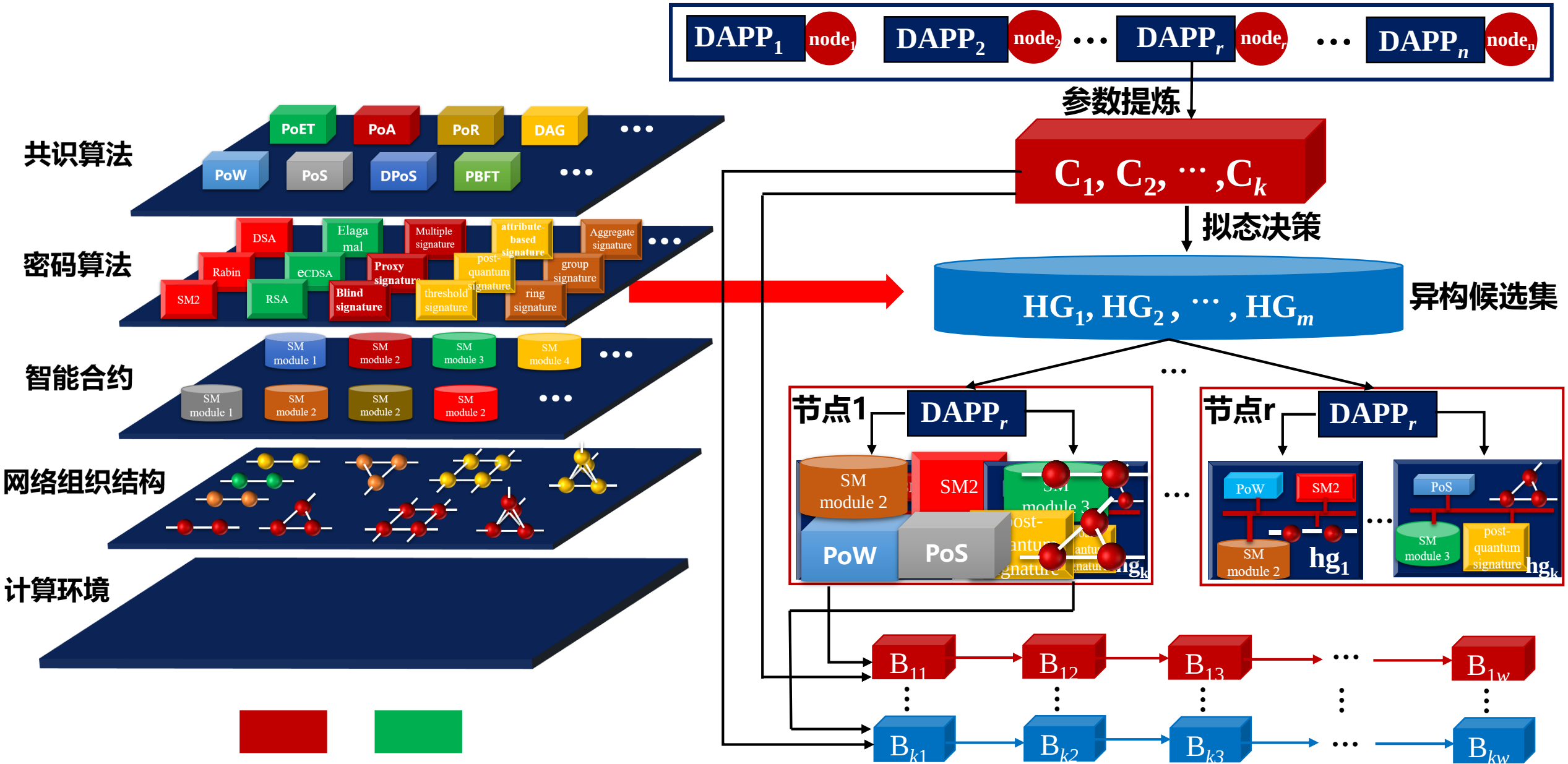
拟态群体异构体系结构


共识算法
密码算法
智能合约
网络组织结构
节点环境
DAPP1 node1
DAPP2 node2
DAPPr noder
DAPPn noden
参数提炼
C1, C2, ⋯ ,Ck
拟态决策
HG1, HG2 , ⋯ , HGm
异构候选集
节点1
DAPPr
PoW
SM2
SM module 2
hg1
PoS
SM module 3
post-quantum signature
hgk
节点r
DAPPr
PoW
SM2
SM module 2
hg1
PoS
SM module 3
post-quantum signature
hgk
B11
B12
B13
B1w
Bk1
Bk2
Bk3
Bkw

拟态群体异构体系结构主要组成



拟态群体异构体系结构动态演示



区块链体结构演进方向

长期以来，信息系统在不断地演进，**体系结构创新**是信息系统发展的永恒之路。我们从信息系统的角度来考察区块链的体系结构演进；首先对传统信息系统进行了数学建模，指出出了传统信息系统的特点；然后对经典区块链系统进行了研究，给出了区块链系统的形式化表示和体系结构模型，提炼了经典区块链系统的特点；最后，我们根据理想区块链系统的特点，从自然界的**拟态、群飞、群游**现象中得到启发，提出了**拟态群体异构**体系结构。

谢 谢