DTCC

第十四届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA

数智赋能 共筑未来

IT168.com ChinaUnix ITPUB

北京国际会议中心 | 2023/8/16-18

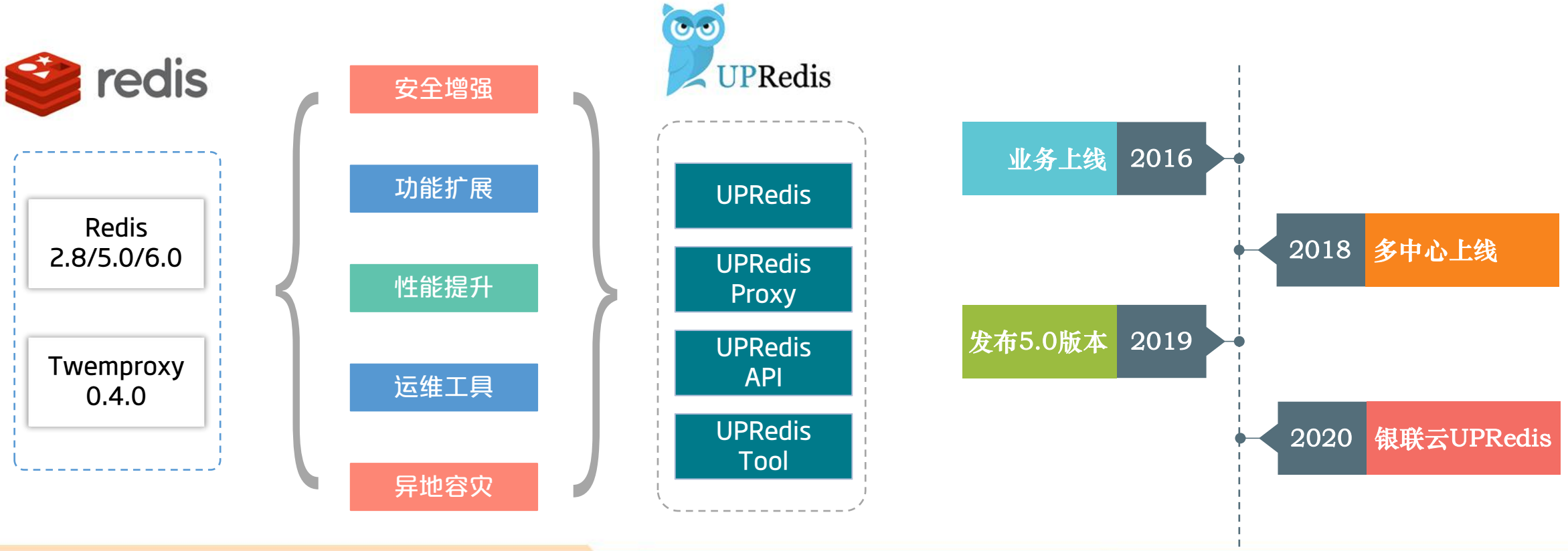
中国银联分布式缓存的探索与发展

中国银联 潘孟琦

目　录

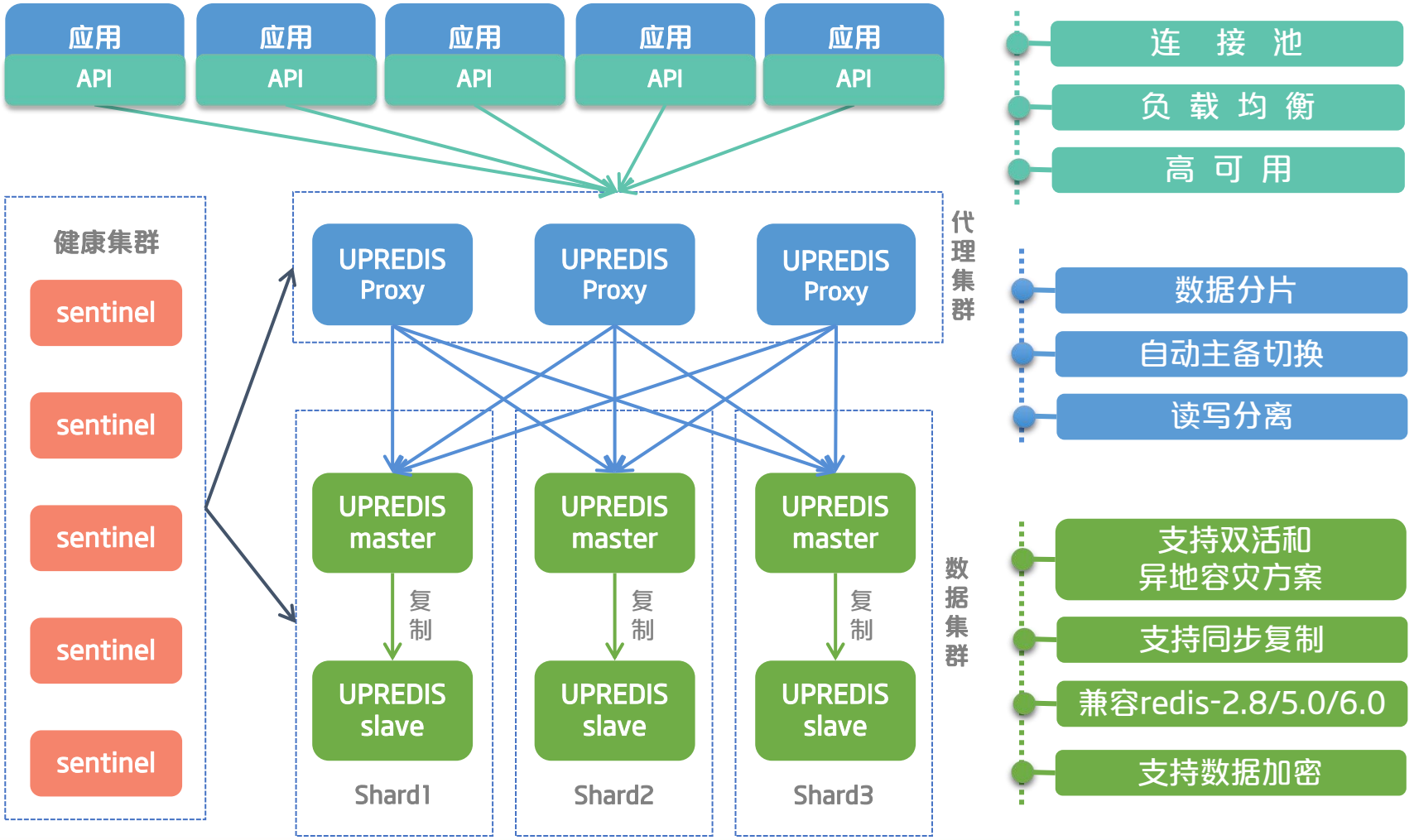
总体概况

UPRedis是中国银联在开源的redis数据库社区版基础上，根据银联业务特点定制开发的KV数据库产品。



UPRedis集群架构

UPRedis
集群架构



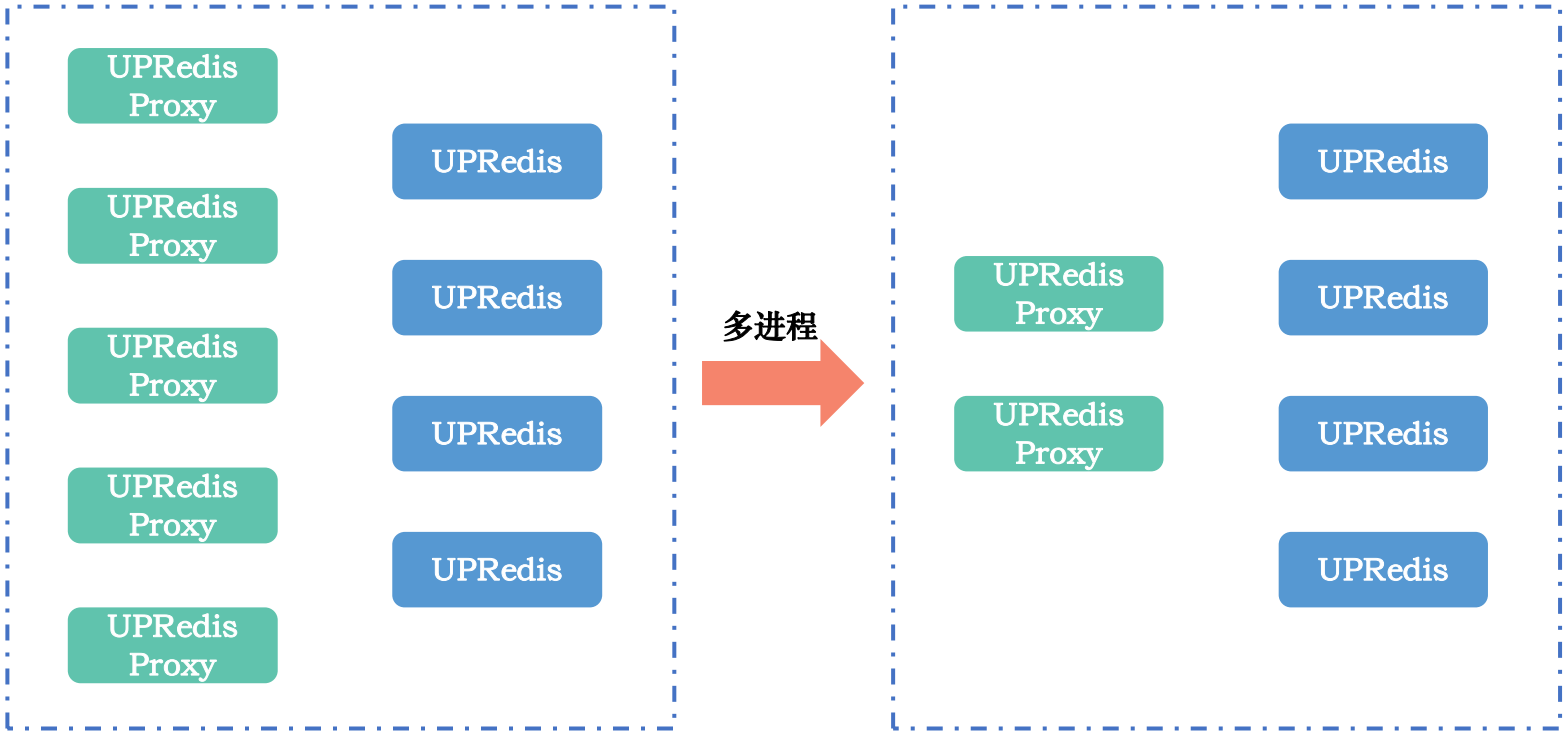
业务场景

- 云闪付营销：消费券抢杀、商品秒杀等活动，特点是并发高，压力大，要保证系统的可用性
- 联机交易：交易限流限额、交易查询等，需保证数据可靠性、一致性
- 其他：二维码信息、会员信息、会话缓存、生物特征信息等

- ✓ 数据可靠性：保证数据不丢失
- ✓ 数据安全：安全合规要求，如身份信息、生物特征等不能明文落盘
- ✓ 性能：业务量逐渐增长，解决性能瓶颈
- ✓ 数据多中心：异地多活趋势，支持数据同步

目 录

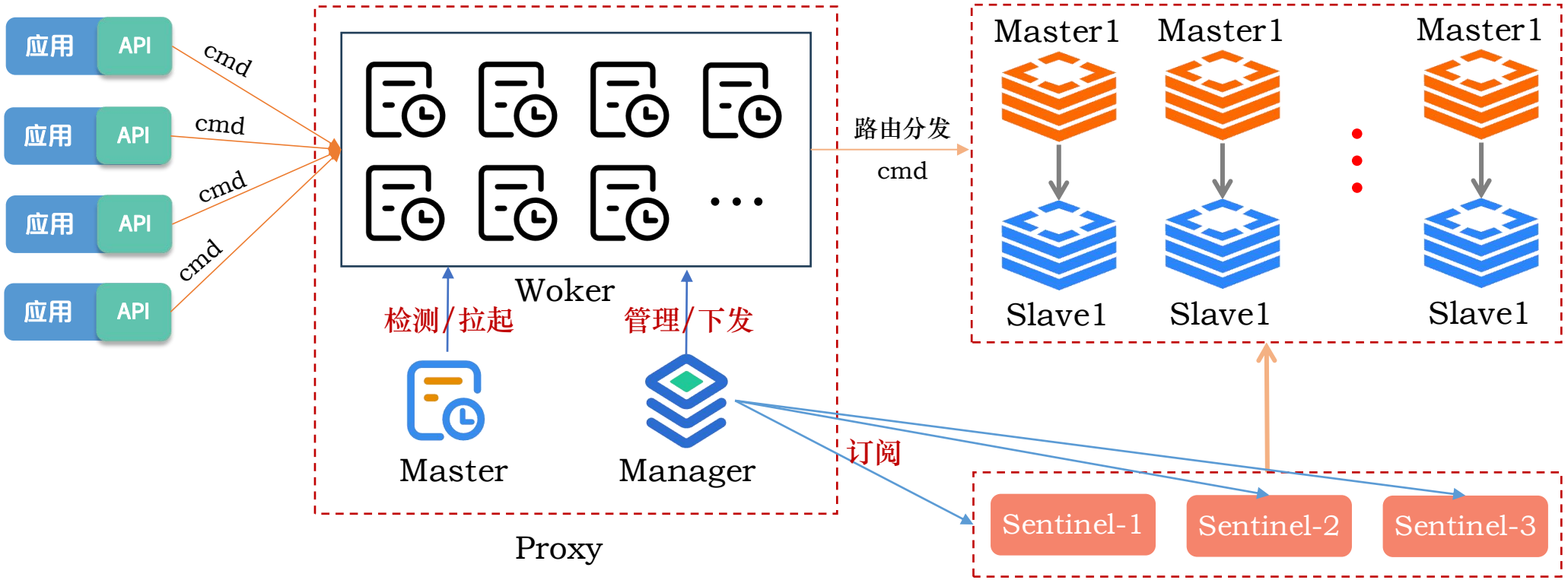
UPRedis-Proxy的探索



UPRedis Proxy单进程性能略低于UPRedis
UPRedis Proxy需要部署比UPRedis更多实例

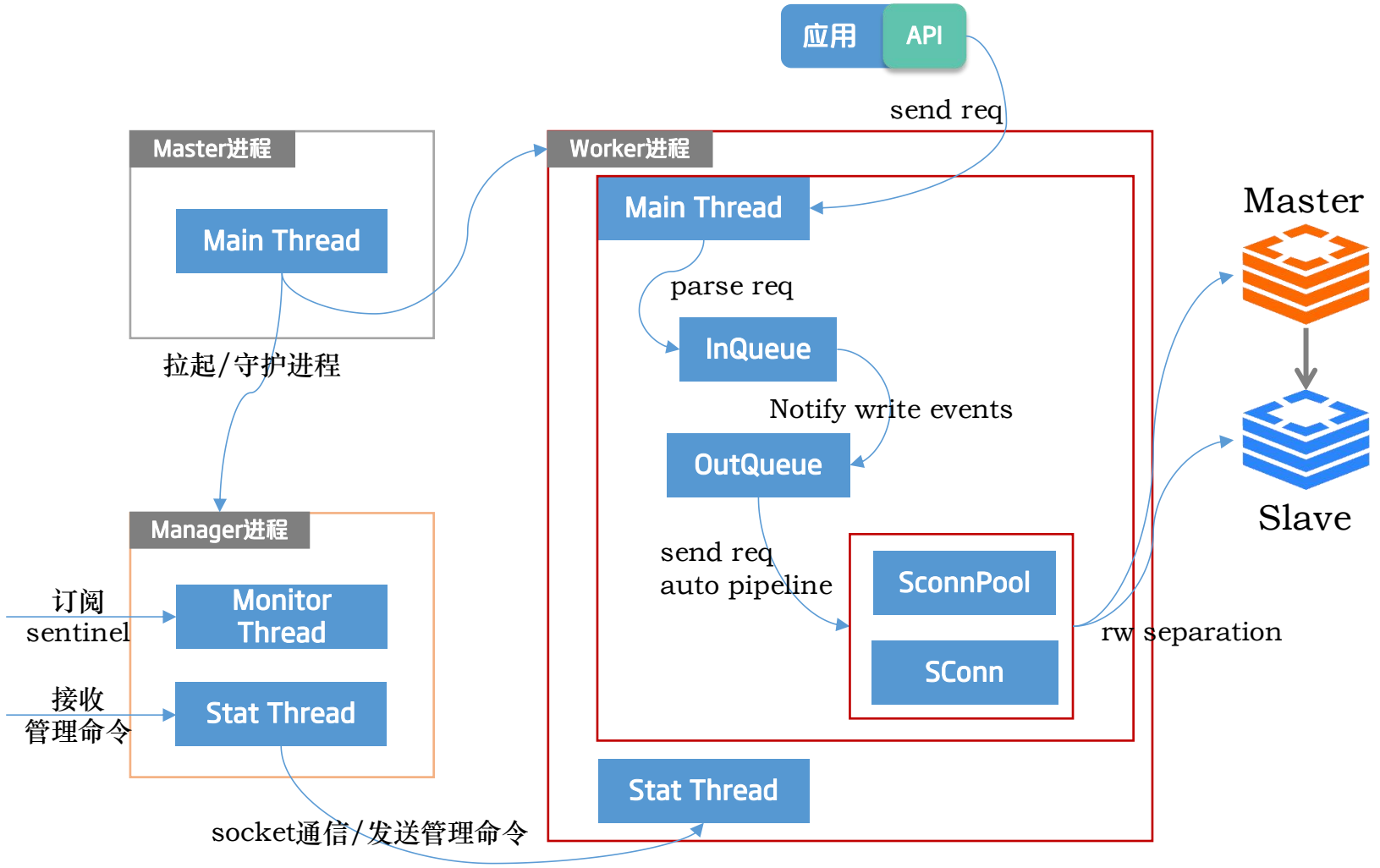
大幅提升UPRedis Proxy单实例性能
减少UPRedis Proxy部署数量，简化运维

UPRedisProxy技术框架



UPRedisProxy技术框架

- Master
  - ✓ main thread: 守护worker和manager进程
- Manager
  - ✓ monitor thread: 监控主从切换
  - ✓ state thread: 用于接收管理命令和与worker进程通信
- Worker
  - ✓ main thread： 处理业务
  - ✓ state thread: 用于与master进程通信，执行管理命令



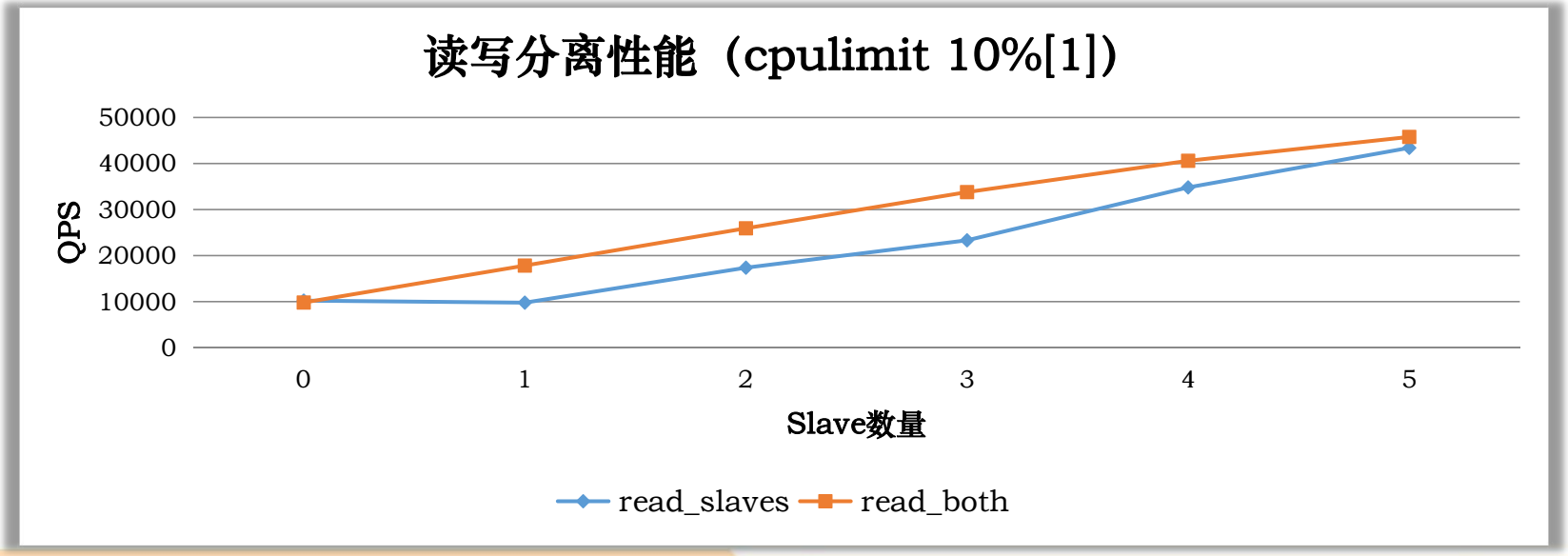
多进程性能提升

环境	命令	UPRedis		UPRedis + Proxy	
		QPS	延迟(us)	QPS	延迟(us)
局域网： 跨机 数据库： 物理机 代理： 虚拟机 数据量： 1500万 数据大小：512字节 并发数： 100 QPS： 10000	SET	110000	235	**505000**	1361
	GET	120000	240	**713000**	1331
	HSET	106000	239	**822000**	1286
	INCR	120000	243	**802000**	1172
	LPUSH	120000	235	**536000**	1372
	LPOP	120000	241	**628000**	1325
	SADD	120000	241	**930000**	1145
	SPOP	125000	312	**1335000**	1155

UPRedis Proxy自动pipeline请求，将Redis单机性能从10万TPS提高到50万TPS

读写分离性能

- 性能调优
  - 直接读写分离并没有线性提升读性能，频繁切换路由对象影响了pipeline效果
  - 读写分离策略修改为批量发送再切换slave，读性能几乎与slave数量成正比
  - 经验证sep_rw_batch为128时，性能和负载均衡效果能达到较好的平衡
- 性能测试
  - 经过模拟测试，显示读写分离性能与slave数量成正比
  - 理论上，读写分离读取性能：1主1从，最高可到100万；1主2从最高可到150万



[1] 由于至少需要6个benchmark进程，8个Proxy进程才能将Redis压到100%，同时还需要使用万兆网卡。而测试环境资源有限，以上测试结果为：cpulimit将redis耗费CPU限制到10%的QPS性能。

Proxy其他优化

命令兼容性

- 兼用上游命令达95%+
- 支持Multi/Exec
- 支持pubsub
- 支持阻塞命令

安全控制

- proxy黑白名单限制
- 支持多用户配置
- 支持命令禁用
- 对接密码管理系统

功能增强

- 读写分离
- 参数动态生效
- 支持管理命令，与业务命令分开

目 录

RDB加密背景

◆ 出于安全合规要求，某些场景redis-server需要对落盘文件（RDB和AOF-BINLOG）进行加密存储。

- 生物特征信息
- 用户身份信息
- 密钥
- ……

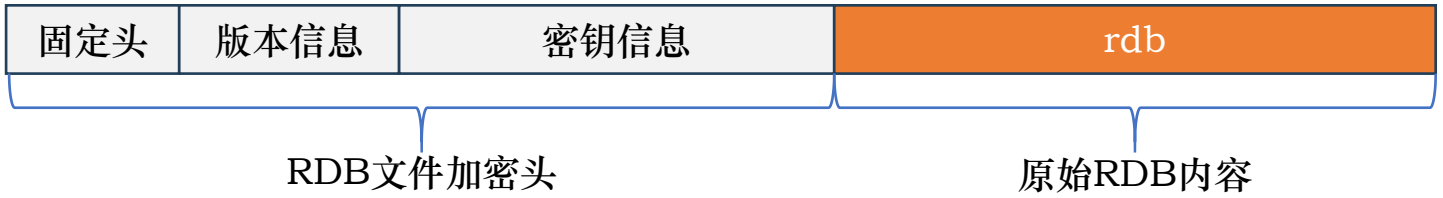
◆ 加密分析

- 加密RDB文件，只在落盘和加载RDB时加密明文
- 主从复制流不进行加密，slave获取全量时为临时的明文RDB
- AOF暂不进行加密，需要考虑性能问题

RDB加密

采用加密算法对RDB文件进行对称加密，加密密钥支持随机生成、文件配置、加密机存储

- master-key全局，初始化时由keyring模块初始化（随机生成，file读取，加密机读取）
- 写入：每次rdb生成都随机生成(subkey,subiv)，用master-key加密得到(encrypt-subkey, encrypt-subiv)存储在RDB文件的encryption-header（应该考虑扩展，版本号等）
- 读取：加载encryption-header使用masterkey解密enc-subkey和enc-subvec，得到(subkey,subiv)，然后解密



加密头：

加密头中存储相关的固定字段、版本号以及加密的密钥字符串信息等，用于解析rdb时解密需要

RDB加密过程

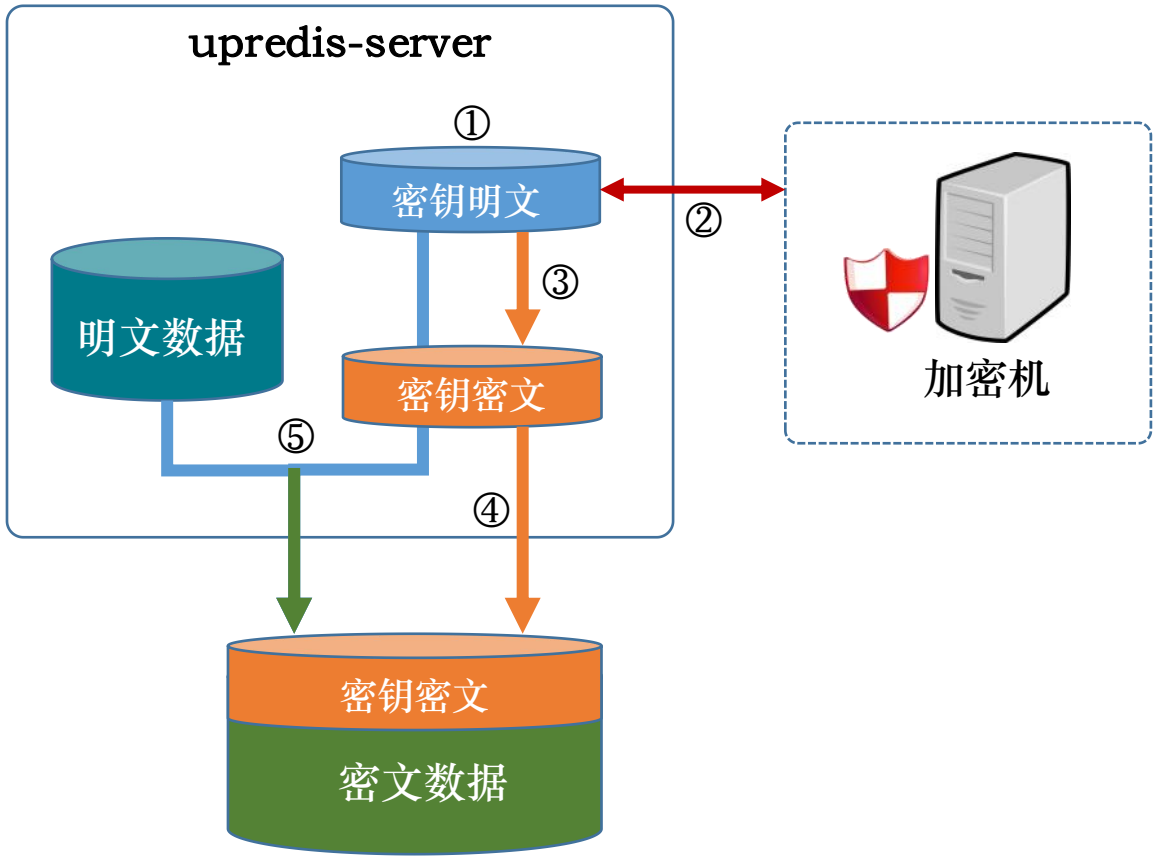
①：随机生成密钥明文

②：使用加密机或密钥文件加密密钥明文

③：生成密钥密文

④：保存密钥密文到RDB文件

⑤：使用密钥明文加密明文数据存储到RDB



RDB解密过程

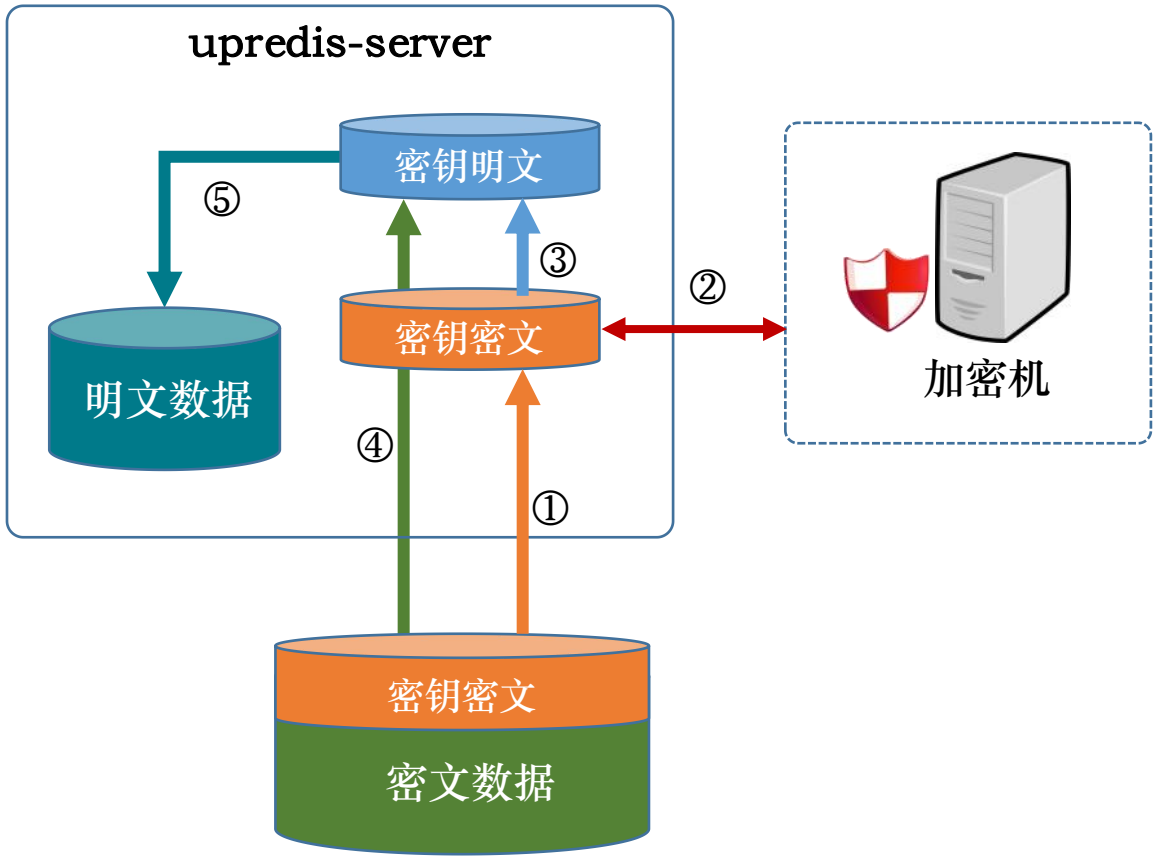
①：从RDB中读取密钥密文

②：使用加密机或密钥文件解密

③：得到密钥明文

④：从RDB读取密文数据

⑤：使用密钥明文解密密文数据得到明文数据



性能

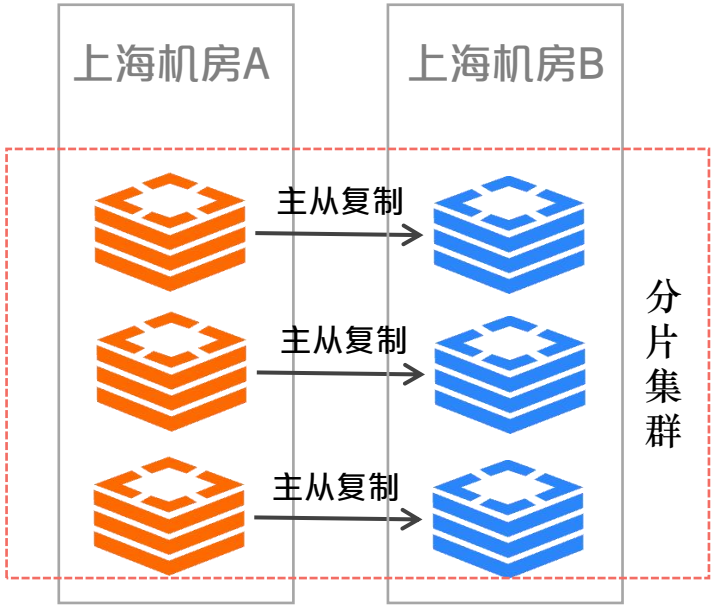
测试了330MB的RDB文件加密解密性能，结果显示文件加解密增加RDB加载、写入时间大约25%。

类型\动作	RDB SAVE (ms)	RDB LOAD (ms)	差值
明文	16	20	+25%
加密	19	24	+26%

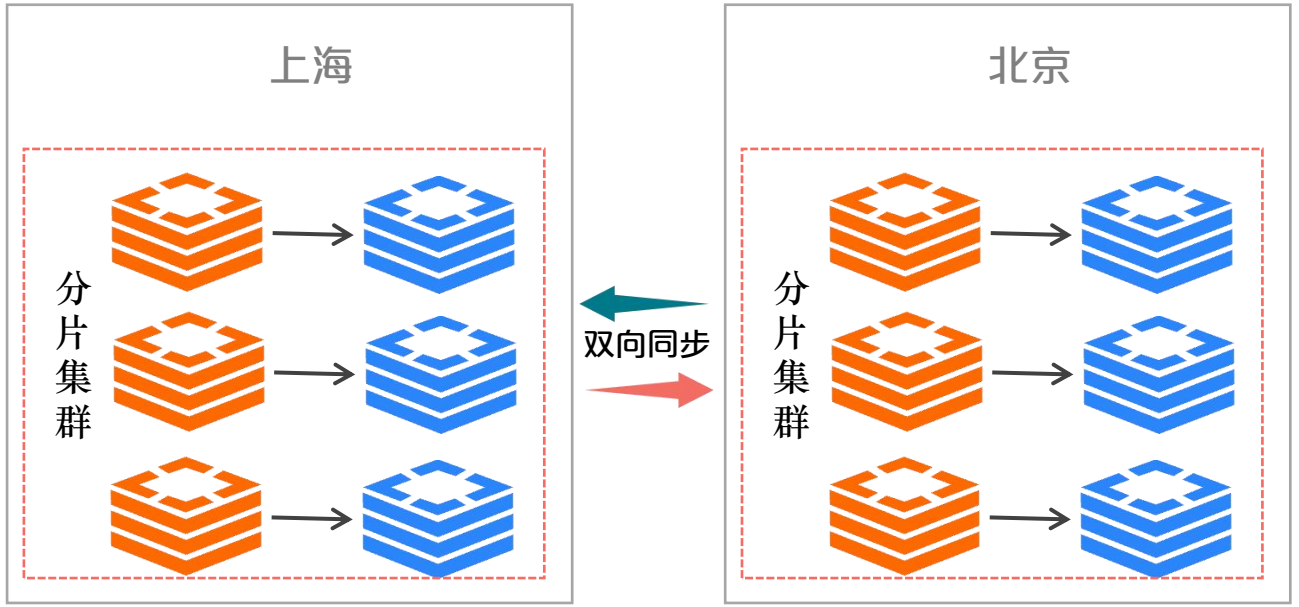
目 录

两种场景

- 同城双机房：采用主备复制，达到机房间的容灾，但只能单向同步
- 异地两中心：两个UPRedis集群进行数据双向同步



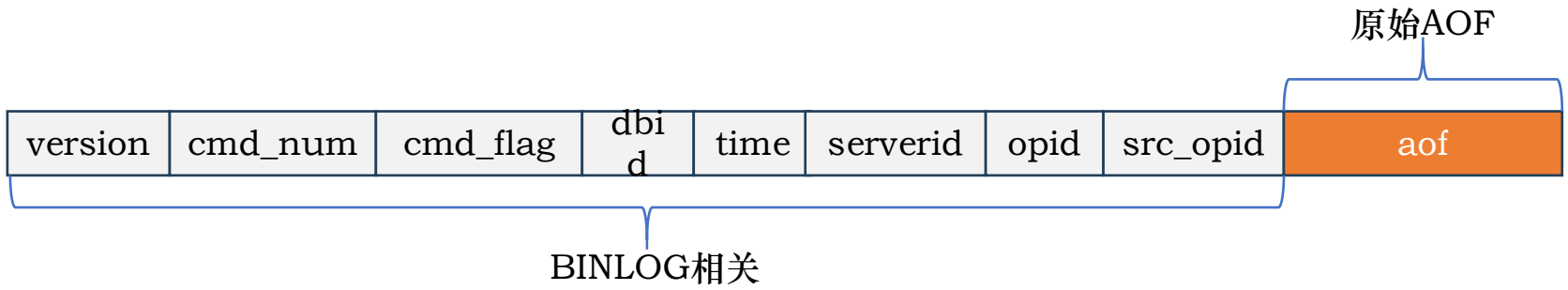
同城复制



异地同步

AOF改造为AOF-BINLOG

- 改造原有个aof格式，在每个命令前添加了一个opinfo命令，记录了命令的opid、serverid、timestamp等信息



例子：

```
*2\r\n$6\r\nopinfo\r\n$32^@ÍZW^D^\r\n*3\r\n$3\r\nset\r\n$1\r\nk\r\n$1\r\nv\r\n
```

- server_id：server标识，用于过滤本中心日志，防止循环复制（双向同步场景）
- opid：一条记录的标识，用于断点续传
- timestamp：时间戳

AOF-BINLOG

新增命令：

- "opget <startopid> [count <count>] [matchdb <db>]* [matchkey <key>]* [matchid <serverid>]* [skipflags delbyexpire|delbyeviction]*" --> "<next_start_opid> <matched_count> <<val1> <val2> ...>"
- "opapply" --> OK/ERR
- "getopidbyaof <aof-filename>" --> <aof-first-opid>
- "purgeaofto <aof-filename>" --> <aof-purged-count>
- "aofflush" --> OK/ERR
- "opRestore <key>" --> OK/ERR

rdb.index

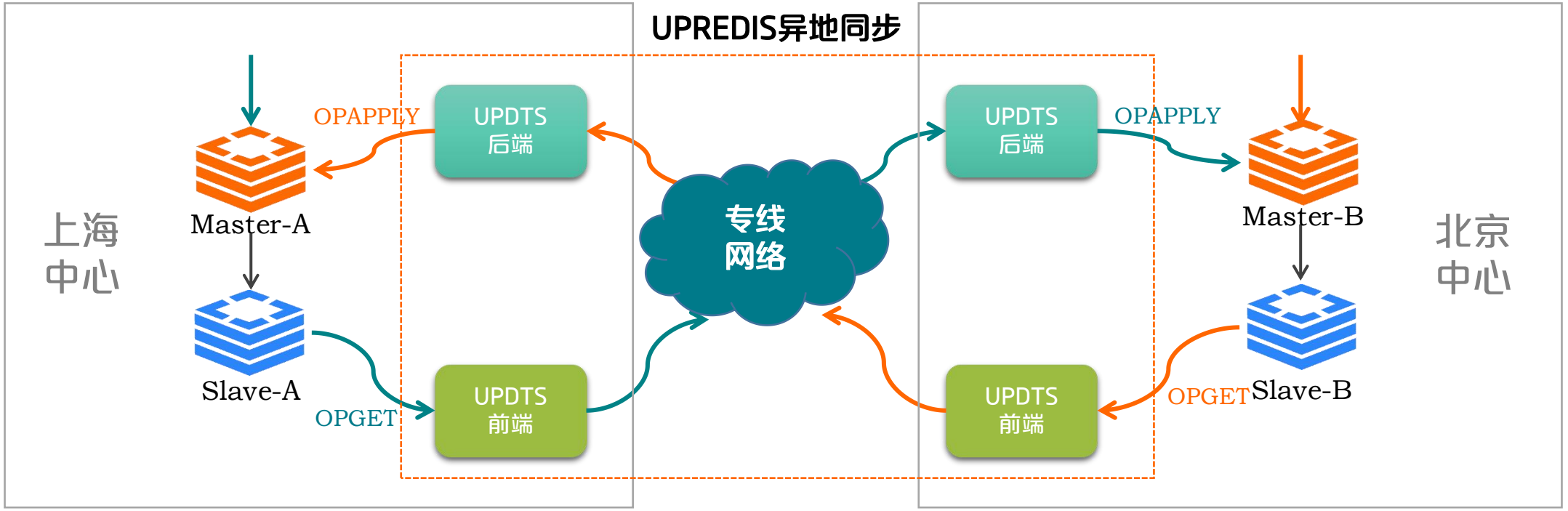
dump.rdb appendonly-inc-1523297820.aof 0 1		
全量文件	接续增量日志文件	opid

aof-inc.index

appendonly-inc-1523297040.aof start_opid appendonly-inc-1523297097.aof start_opid	
aof文件	起始opid

异地容灾架构

- 以SH和BJ两中心为例， UPRedis 1主1备、异地双活架构实现数据同步：
  - ① 通过OPGET命令， UPDTS从SH中心的slave-A获取日志，将数据压缩后传输BJ中心
  - ② BJ中心UPDTS将日志回放到master-B，完成同步
  - ③ 从BJ-->SH的同步过程与SH-->BJ的同步过程类似

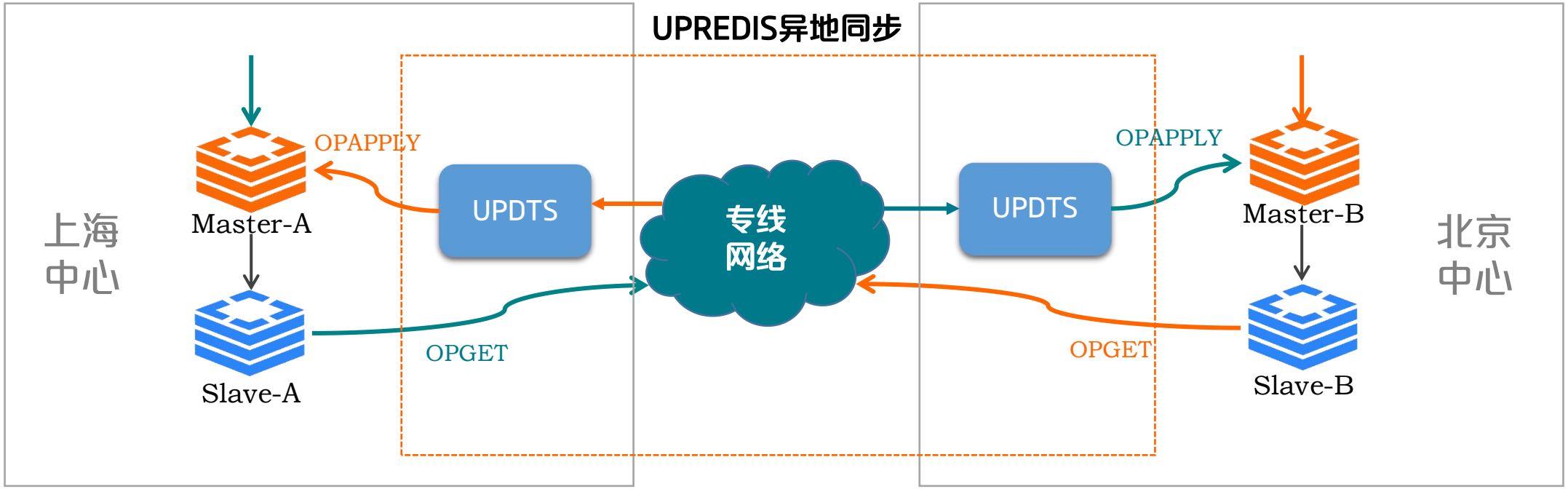


异地容灾架构

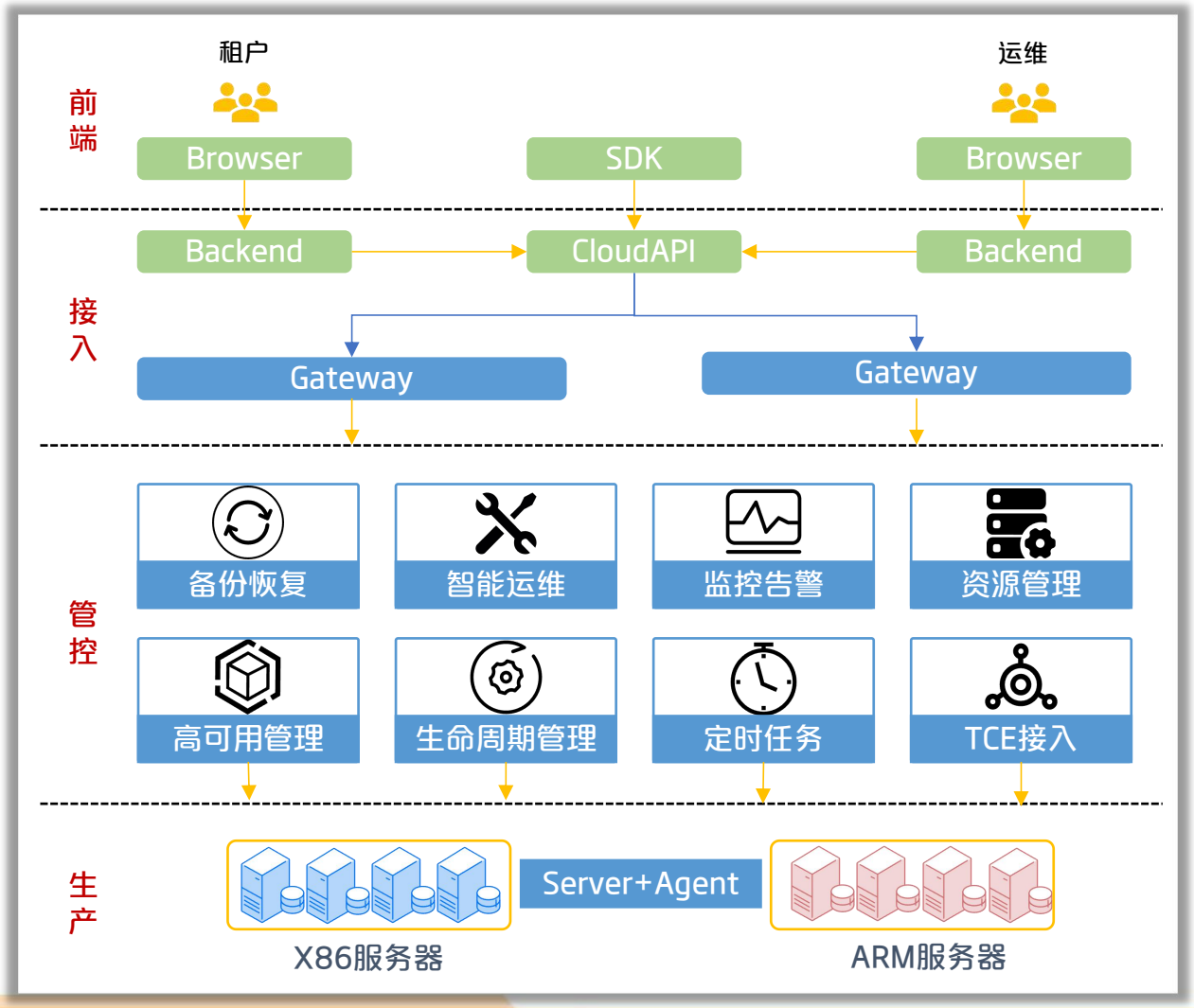
- 另外一种架构（银联云上UPREDIS多中心同步架构）：

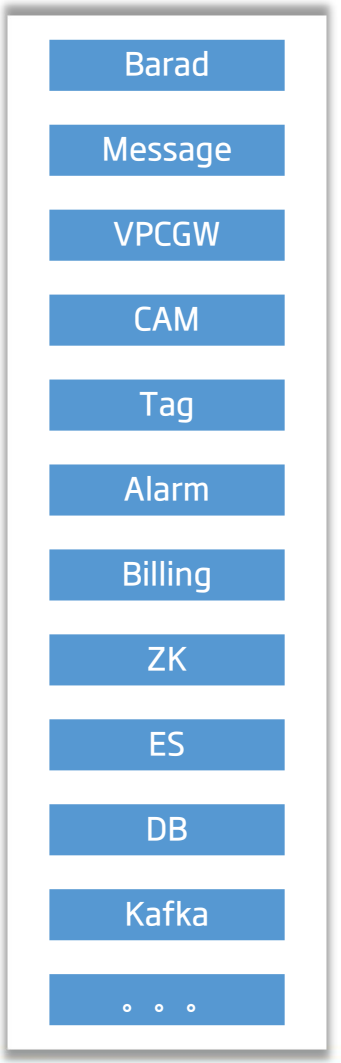
① 通过OPGET命令，BJ中心UPDTS从SH中心的slave-A获取日志，然后回放到自己中心的master-B

② SH中心UPDTS从BJ中心的slave_B获取日志，然后回放到自己中心的master-A

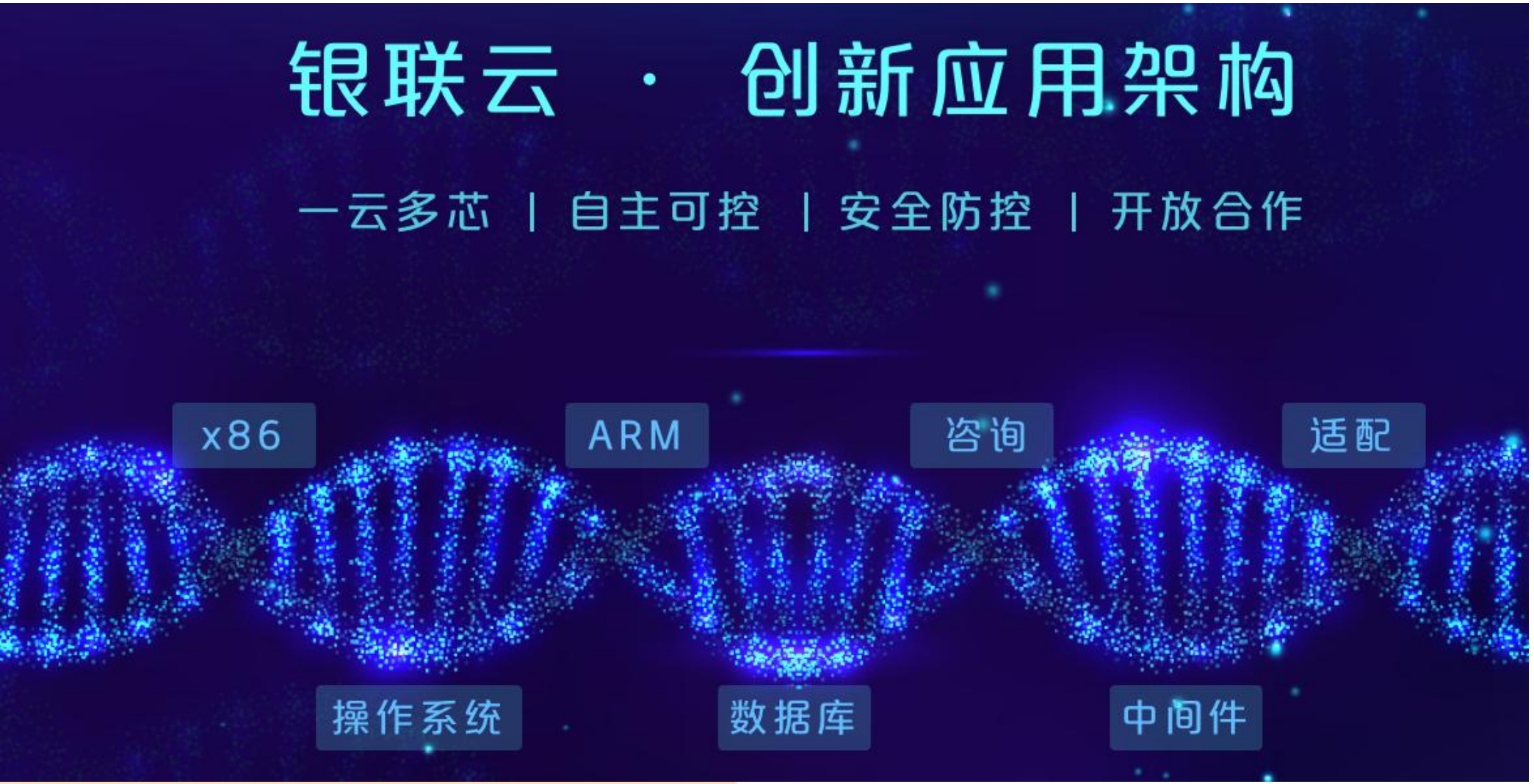


银联云UPRedis的探索





银联云官网：https://yun.unionpay.com/



中国银联

银联云

咨询电话
021- 2063 8828

THANKS