

高承载力物联网接入方案

针对蒙西电网物联网数字底座工程的技术规范要求，平台需满足≥10万台终端接入、≥5万并发长连接、≥10万条/秒消息吞吐量、数据上报延迟≤5秒、API响应时间≤2秒等硬性性能指标。为解决承载能力，需围绕“**边缘分流 + 负载均衡 + 消息削峰 + 分层存储 + 弹性扩展**”构建高并发高吞吐的专项解决方案。

性能需求与架构对策对应表

性能指标	规范要求	架构实现与技术对策
并发连接数	≥ 5万并发长连接	云边分布式接入 + Netty高并发接入网关集群（支持MQTT/CoAP协议长连接保持与心跳保活）
消息吞吐量	≥ 10万条/秒	分布式消息中间件集群（如 Kafka/RocketMQ）削峰填谷 + 多分区并行消费
数据上报延迟	≤ 5秒（端到端）	内存缓存（Redis）+ 实时计算集群（Flink/Spark Streaming）快速处理告警与流式计算
API 响应时间	≤ 2秒（95% 请求）	Spring Cloud Gateway API 网关 + Redis 热点缓存 + 读写分离
命令下发响应	≤ 10秒	双向异步指令通道 + 边缘节点本地代理，支持指令离线缓存与恢复后补发
整体可用性	≥ 99.9%	K8s 容器化弹性伸缩 + 双区（专网区/互联网区）边缘高可用部署

接入能力专项解决方案设计

1. 海量长连接接入与负载均衡方案（支持 5万+ 并发连接）

- 四层与七层混合负载均衡**：前端采用 SLB / LVS 实现接入层流量分发，将海量终端连接平摊至后端的接入服务节点。
- 高性能协议接入网关**：基于 Netty 框架自研或优化物联网接入网关，支持 MQTT、CoAP、HTTP/HTTPS 等主流协议的拆包、组包与心跳保活。优化 JVM 参数与 Linux 操作系统内核参数（如 `somaxconn` 和文件描述符限制 `nofile` ），确保单节点支持高并发连接。
- 专网与互联网隔离接入**：在专网区和互联网区分别部署边缘接入网关与安全网关。互联网区终端通过 APN 专网与安全网关接入，专网区终端直连专网边缘节点，实现流量分散与物理/逻辑隔离。

2. 高吞吐异步削峰与消息分发方案（支持 10万条/秒 吞吐）

- 消息中间件集群削峰**：接入网关收到终端上报数据后，不做复杂业务逻辑处理，直接推入高吞吐消息中间件集群（分布式 Broker 配置）。通过多 Topic 与多 Partition 设计，解耦接入层与业务逻辑层，防止突发大流量冲垮后端数据库。
- 云边协同异步数据同步**：边缘节点具备数据本地缓存能力。在网络波动或突发峰值时，边缘侧先进行数据暂存与初步清洗，再按优先级异步同步至云端中心节点。

3. 端到端低延迟与分层存储方案（上报 ≤5s / API ≤2s）

- 热数据内存缓存**：采用 Redis 集群存储终端最新状态、在线会话、热点配置及实时告警，保障 API 查询在毫秒级内响应。
- 海量时序数据存储**：针对终端周期上报的遥测、遥信等海量时序数据，采用专业时序数据库（或大数据技术）进行高吞吐写入与降采样存储；业务结构化数据（如设备台账、权限、资产）采用 GaussDB 关系型数据库存储，实现读写分离与高可用集群。
- 实时计算与告警触发**：接入实时计算集群，对上报的设备数据流进行实时规则匹配与异常检测，确保严重告警能在 5 秒内推送到前端看板与外部系统（如智能运检、安全监控平台）。

4. 命令高可靠下发与通道保障（响应 $\leq 10s$ ）

- 双向实时控制通道：**基于 MQTT Pub/Sub 或 WebSocket 建立反向控制通道，下发命令通过消息中间件保证至少送达一次（QoS 1/2）。
- 指令队列与离线补发：**针对离线或弱网终端，命令下发进入待执行队列，由边缘节点代理管理。终端恢复连接后立即触发状态同步与未完成命令下发。

5. 高可用与弹性伸缩机制（保障 99.9% 可用性）

- Kubernetes 容器化动态扩缩容：**依托 Docker 和 K8s 编排能力，监控 CPU、内存及消息队列积压指标。当接入流量突发增长时，自动触发接入网关与微服务节点的 Horizontal Pod Autoscaler (HPA) 弹性扩容。
- 限流熔断与服务降级：**API Gateway 集成限流熔断机制（基于 Resilience4j/Sentinel）。在超出系统承载上限时，优先保障实时告警与核心数据接入，对非核心历史数据查询实施降级或限流。