

The logo for SDCC 2016, featuring the text "SDCC 2016" in a bold, blue, sans-serif font. The background of the slide includes a light blue grid pattern, a large blue dot pattern on the left side, and several circular icons: an hourglass, a power button, a mail envelope, a gear, and a star.

中国软件开发者大会

SOFTWARE DEVELOPER CONFERENCE CHINA

# 实时系统 Erlang 实践

张虎

zhanghu@yunba.io

# 关于我

- 云巴 创始人，CEO
- 极光推送 创始人，原 CTO
- 微博: @Tiger\_张虎
- 云巴 Yunba.io: 基于 Erlang 大规模实时系统

# 关于云巴

- 云端实时通信
- 物联网 (IoT)
  - MQTT
- 支持 WebSocket
  - [socket.io](http://socket.io)

# 云巴的挑战

- 支持亿级用户
- 线性扩容
- 高可用
  - 无单点失效
  - 多 IDC 容灾
- 软实时
  - 最大延迟 200毫秒

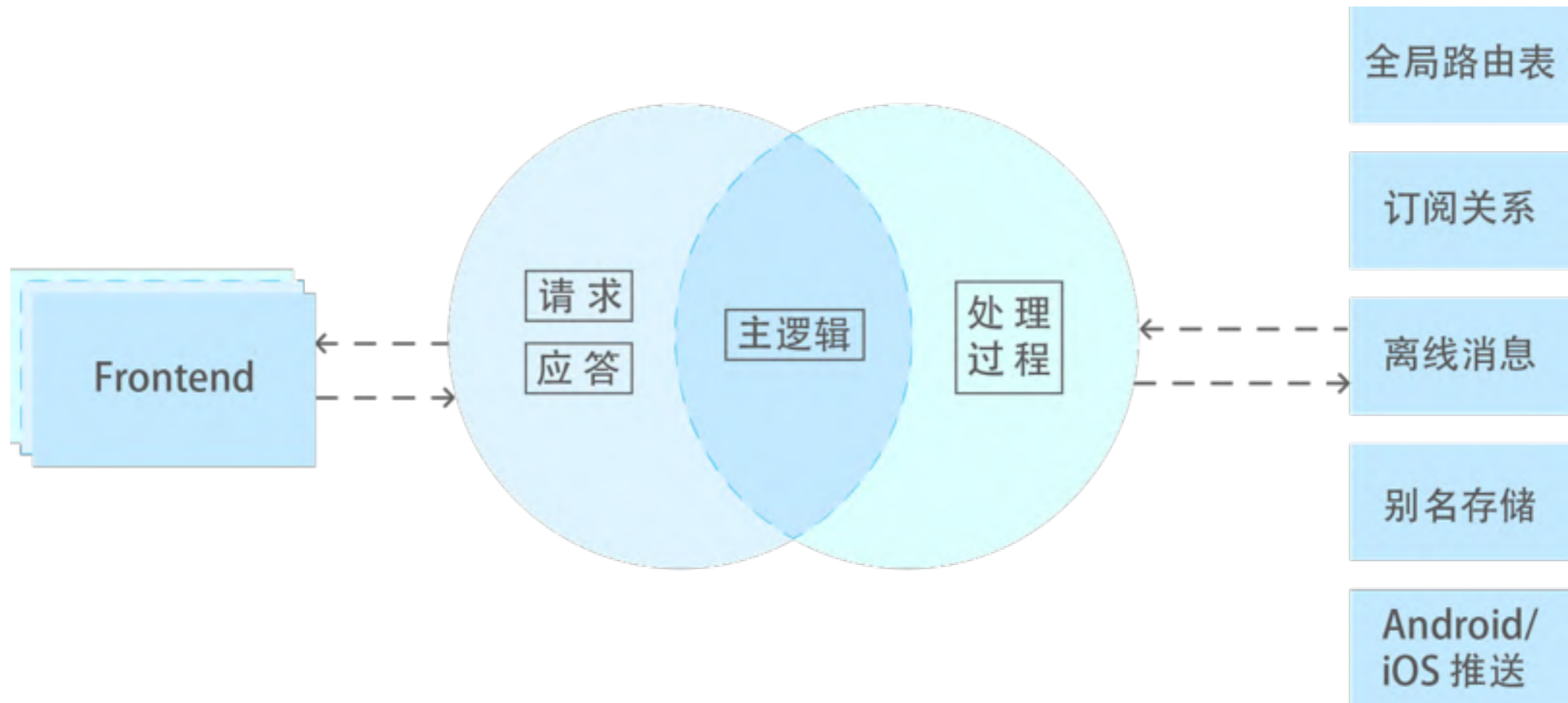
# 可选方案

- 异步 IO
  - Libev/epoll: 编码效率低下
  - Node.js: 并发、压力大时很多问题
- 轻量级进程
  - Erlang
  - Golang
  - Akka

# 为什么选 Erlang

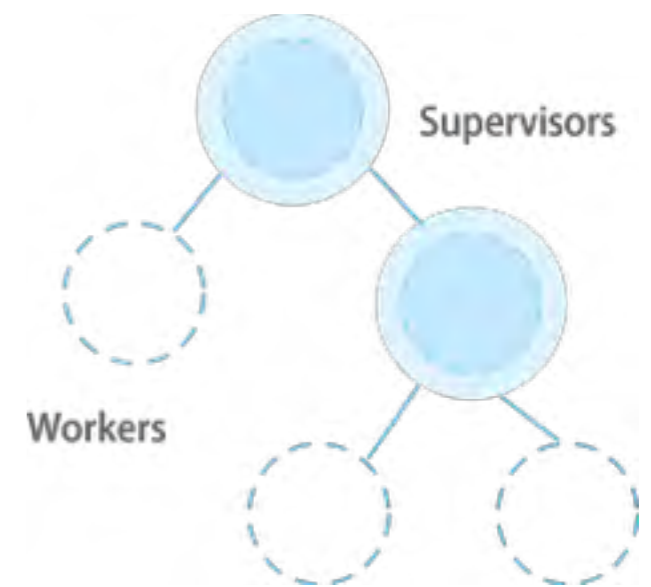
- 公平调度
  - 可抢占, reduction budget
- 支持分布式
  - 远程 RPC
- 进程独立的 GC
  - 有效减少 GC 带来的延迟
- OTP

# 架构



# OTP Applications

- 每一个模块封装一个 Application (Supervisor)
- 每一个子模块有一个 Supervisor
- 业务逻辑由 Worker 实现，对外提供 RPC 接口 (微服务)





# 接入负载均衡模块

- 客户端每次连接前，先请求负载均衡模块分配一个 IP 地址、端口
- 实时监控接入模块负载
- 根据配置的策略，接入点的负载情况分配
- 动态添加、移除接入点
- 每一个接入点用一个 Erlang 进程监控

# 接入模块

- 维持长连接
- 转发请求、返回结果
- MQ -> RPC
- 每一个连接就是一个 Erlang 进程

# 主逻辑模块

- MQ -> RPC 接口
  - 网络减少一跳
  - 降低部署复杂度，内部路由更直接
- Dispatcher worker，分发任务
  - poolboy
- Worker，每个 worker 每次处理一个业务请求

# 大规模 publish

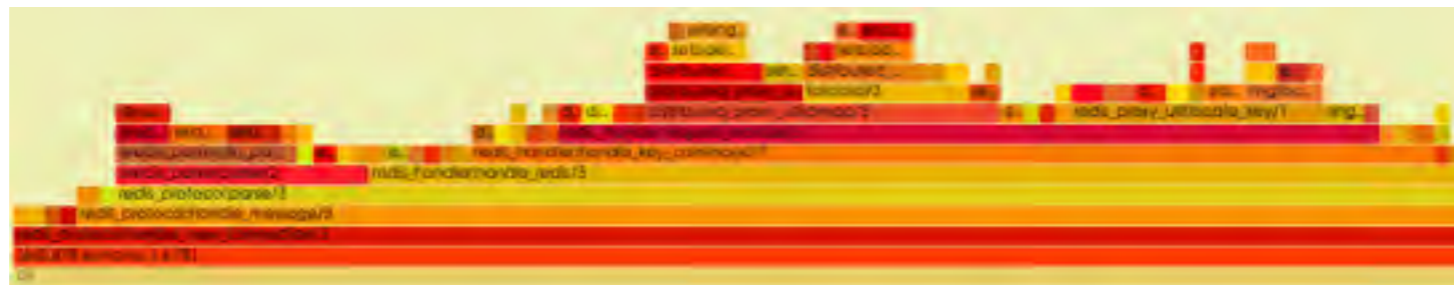
- 目标客户端数 >200
- Worker 把 publish 分片
  - 200个UID每片
- 分片处理任务池
  - 每个分片由一个 Erlang 进程处理

# 离线消息性能问题

- 无线网络抖动，需要离线消息保证送达率
- 一次 publish 有可能大量离线消息
- Write-back cache，延迟写磁盘；丢消息的风险。Couchbase
- 同步写日志，异步回放日志写离线消息；重复发送的风险。Raft

# CPU Profiling

- fprof/eprof
- 火焰图 (<http://www.brendangregg.com/flamegraphs.html>)
- eflame
  - erlang:trace/3 定期采样调用栈
  - 对 BEAM 打 Patch, 减少 trace 的输出
  - <https://github.com/TigerZhang/eflame>



# 监控 / 测量

- 每个节点的系统参数收集绘图
  - 每分钟、每10秒, Munin/influxdb
  - Grafana 绘图
- 每个节点监控
  - CPU, mem, network, disk
  - BEAM (进程数、队列长度, scheduler usage); <https://github.com/ferd/recon>
  - 队列长度 (msgq length) 可以认为是模块 CPU 占用率的指标
- 业务数据绘图
  - Influxdb/grafana

# 性能陷阱 Caveats

- `list_to_atom/1`
- 避免高频率的短连接。加代理



# 为什么 Erlang “难学”

- 函数式语言？
- 几天的适应期
- 工具链不好用？ Rebar? Rebar3?
- 很多人在这放弃

The logo for SDCC 2016, featuring the text "SDCC 2016" in a bold, blue, sans-serif font. The background of the slide includes a large graphic on the left composed of blue dots of varying sizes, and several circular icons scattered across the page: an hourglass, a power button, an envelope, a recycling symbol, a gear, a laptop, and a starburst.

**SDCC 2016**

**中国软件开发者大会**

SOFTWARE DEVELOPER CONFERENCE CHINA

**Q&A**

**谢谢**