The Design and Application of Baetyl

LI Leding
Chair of Baetyl TAC

Jun 2022
Computing is everywhere

- Cloud IDC
- CDN
- 5G/MEC
- IoT
- LAN Edge
- WAN Edge
- Thing Edge
- 1~10ms
- 10~100ms
- >100ms
- IoT
- APP
- CDN
- 5G/MEC
Fusion between the Cloud and the Edge

**Edge Side**
- Conn & Storage
- Data Processing
- Synchronization
- Edge Framework
- Edge node Infra

**Cloud Side**
- Devices & HDW
- Applications & Conf
- AI Models & Algo
- Mgmt. Console

Data Uploading:
- Log & Diagnostic
- Training Sample
- Runtime Status
- etc

Knowledge Downloading:
- Application
- ML Model
- Device Control
- etc

**Sources**
- Video
  - Camera & Screen
- Audio
  - Microphone & Speaker
- Data
  - Sensor & Machine
Launched in 2017 as Baidu IntelliEdge, open source in 2018

Join LFEdge umbrella in 2019

Provide Cloud Native environment for all applications and AI services

Unified management for both cloud and edge

Unified security model and data privacy protection

Vendor neutral, Apache License

https://baetyl.io/
Baetyl Architect

- Leverage Kubernetes
- Edge cluster, multiple node
- Open-source management
  - baetyl-cloud
  - with OpenAPI
- Self upgradable
  - Baetyl is running as a Pod
- More integrated services
  - Local KV storage
  - Logging & Tracing
  - Remote metrics report
Open ecosystem, Vendor neutral

- **A Cluster-to-Cluster design**
  - inspired by K8S Federation
  - self-scheduling
  - local load balance & failover

- **No modification** to kubelet
  - no specified K8S vendor requirement
  - no specified K8S version requirement

- **Full K8S semantics in edge instance**
  - Pod, Deployment & Service
  - DaemonSet
  - ConfigMap & Secret
**AI Vision**

**Steps:**
1. Power on
2. Connect camera to POE line
3. Connect AI chipset to USB port

**Workflow:**
1. Load AI model
2. Sampling video as images
3. Do inference, accelerated by AI chipset
4. Save target image
5. Send result to next process node
6. Push target image into cloud
AI auto convert

Report hardware capabilities

Acceleration

Train AI model on cloud

Inference on edge

Match framework with hardware

Convert

<table>
<thead>
<tr>
<th>Source</th>
<th>Device</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any</td>
<td>Intel Movidius</td>
<td>OpenVINO</td>
</tr>
<tr>
<td>PaddlePaddle</td>
<td>CUDA</td>
<td>PaddleLite</td>
</tr>
<tr>
<td>TensorFlow</td>
<td>EdgeTPU</td>
<td>TF-Lite</td>
</tr>
</tbody>
</table>
The AI platform includes:

• Cloud management suite, responsible for batch distribution and configuration update of AI applications

• Local client, supports deployment on the edge and end, and is responsible for AI application start-stop and state management at the edge layer and end-side

Edge service includes:

• When the local client is deployed on the edge node, the cloud delivers the application to the edge layer, the edge layer collects the terminal data of the device layer, implements prediction and inference on the edge node.

• When the local client is deployed on the terminal device, it is applied to the terminal under the cloud, and prediction and inference are realized on the terminal.
Intelligent quality inspection

**Issues**

- Model deployment time is long
- Training data collection is complex
- Spare parts quality inspection cloud inference delay affects quality inspection speed

**Effects**

- Model deployment went from 1 day to 1 minute
- Automatic collection of training data and automatic upload to the cloud
- Edge AI inference realizes edge quality inspection and improves quality inspection efficiency

**Diagram Description**

- Raw images go through the process of Video Sampling, AI Model, Marks & Save, followed by Baetyl, CPU & GPU, and Cameras to reach the Inspection Machine.
- False result goes through the Baetyl Cloud and PaddlePaddle to generate AI result.
- New models and new configurations are generated and saved.
Using Baidu IntelliEdge to achieve vehicle-road collaboration

Over 60% reduction in automatic driving costs to help solve cross-regional road synergy challenges
Global optimization of a wide range of multi-base station nodes
Contact us

https://baetyl.io/
https://github.com/baetyl/baetyl
https://github.com/baetyl/baetyl-cloud

mailto:baetyl@lists.lfedge.org
lfedge.slack.com/#baetyl

Baidu Cloud IoT
https://iot.baidu.com/
Thanks