

mini01

Runtime Agent

http://127.0.0.1:8090

Backends

Backend	Kind	Model	Runtime	Status	Control
text-main	text_generation	n/a	n/a	off	n/a
embedding-main	embedding	n/a	n/a	off	n/a
whisper-main	transcription	whisper-v3-turbo-FM	lemnade	running	on
image-sd-turbo	image_generation	SD-Turbo	lemnade	running	on
image-sd-1-5	image_generation	stablediffusion-1.5-sd-turbo	sd-turbo-backend	running	on
describe-gemini	image_description	gemini-1.5-4k-FM	lemnade	running	on

gpu01

Runtime Agent

http://localhost:8180

Backends

Backend	Kind	Model	Runtime	Status	Control
text-gpu	text_generation	n/a	bertaccp	running	on
embedding-gpu	embedding	n/a	bertaccp	running	on

Inference Broker

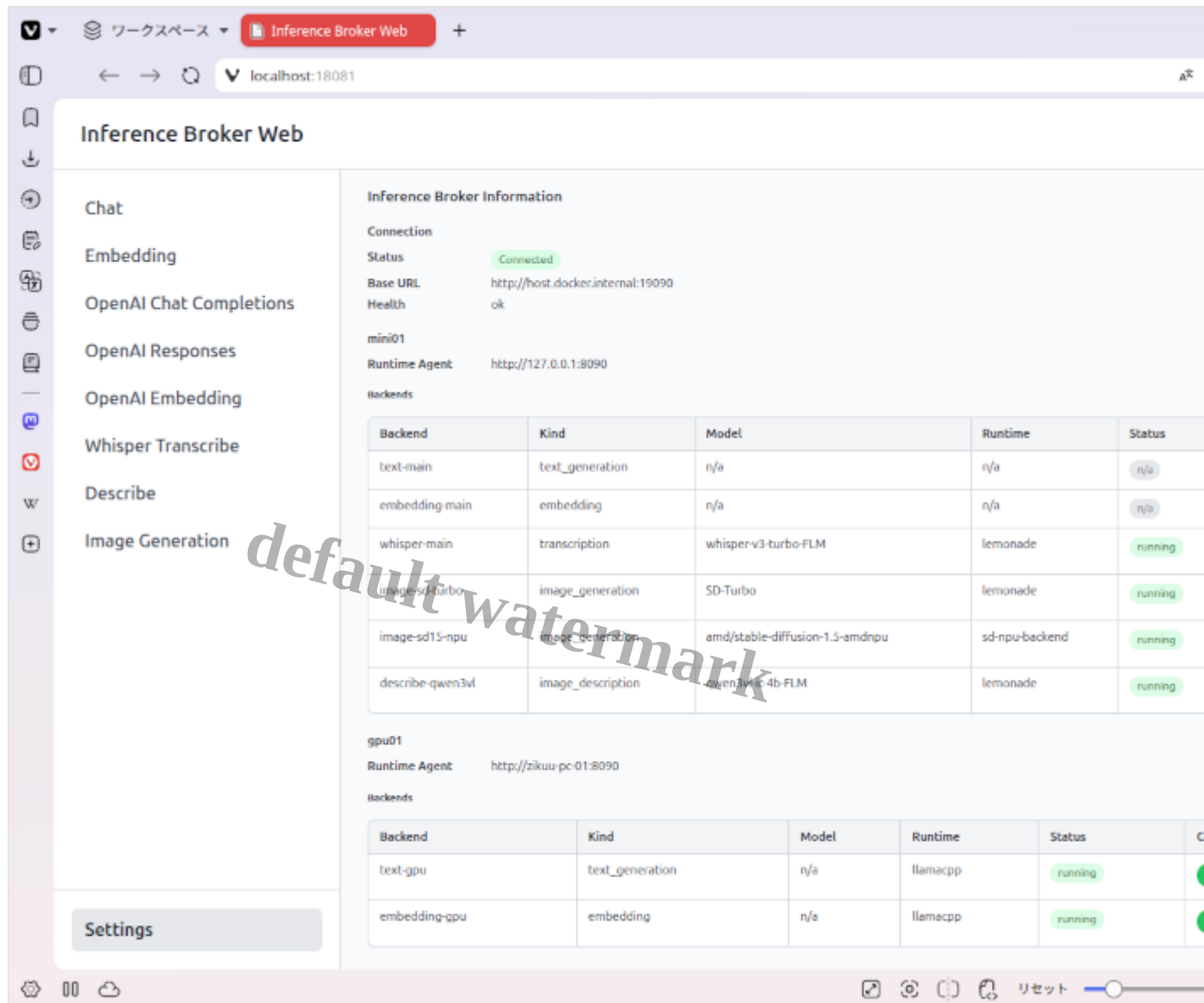
Description

LLM Runtime Agent Inference Broker

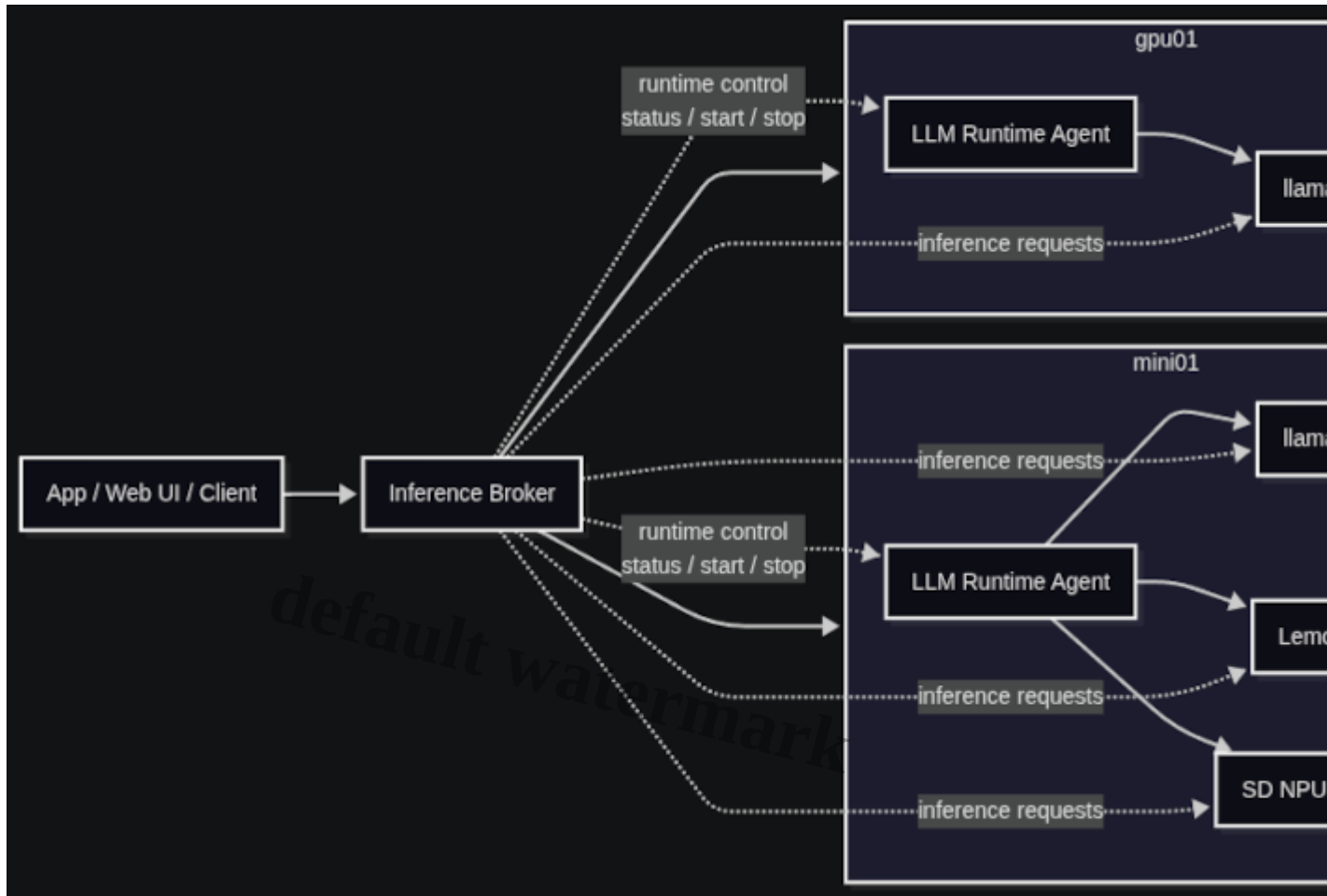
mini01 PC • gpu01 GPU

Backends

LLM Runtime Agent



ç•¼æ??ç?¹ã•šã•ã?å•ç³»æ?è«?è³?æ°•ã??ã½¿ã•£ã•ã??ã?©ã?¼ã?«ã?
ã??ã?ã??ã?ªã?ã?ã?¹ã??è² èå??æ?£ã•®æ©?è?½ã•ã?ã??ã?¼ã•ã??ã?ã?•ä¿?å¿?ã?
æ?è«?è³?æ°•ã?ªã³ã??ã?©ã•ã?ã?¼ã?ã•šã•ã??ç?¶æ ?ã•«ã•ã?ã??ã?¼ã?ã?ã??
ã?ã??ã?Inference Brokerã•ã»?ã•®ã³ã³ã?ã?¼ã?ã³ã??ã•®é?é?£ã??èjã?å³ã•šã?ã??

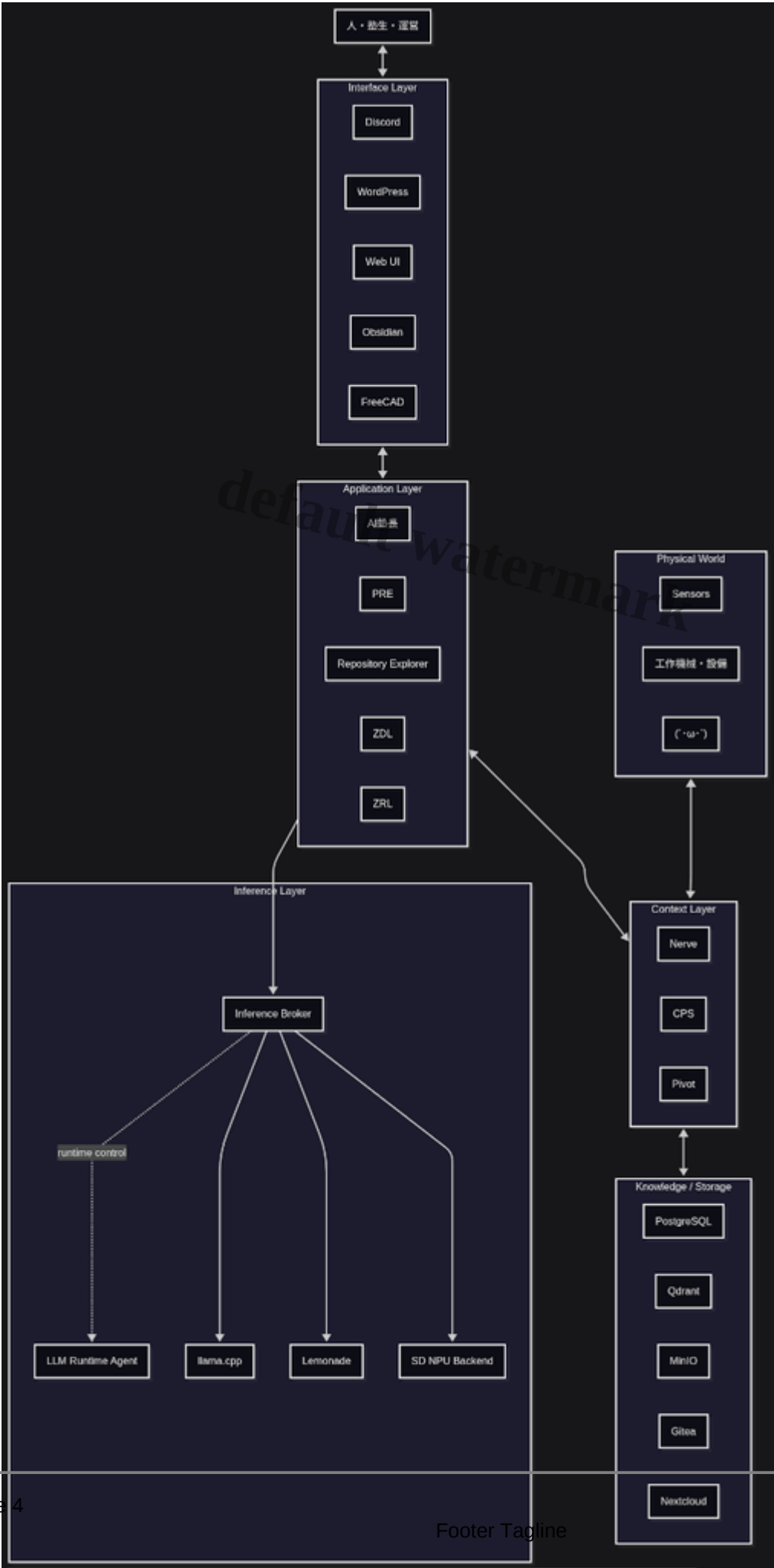


ã•ã•®1á¹´ã»ã•©ã•®é??ç?°ã•šã?•ZIKUUã•®ã?•ã?¹ã??ã?ã•¯ã?ã•ªã??ã•šã•ã•é?•å±•ç??ã•«ã•ªã??ã•¾ã•?ã•?ã??

- $\ddot{a}^{00}$
- Interface $\mathring{a} \pm \mathfrak{a}$
- Application $\mathring{a} \pm \mathfrak{a}$
- $\mathfrak{c} ? \textcircled{\mathfrak{c}} \cdot ? \ddot{a} \cdot ? \mathfrak{c} ? ?$
- Context $\mathring{a} \pm \mathfrak{a}$
- Inference $\mathring{a} \pm \mathfrak{a}$
- $\mathfrak{c} ? \text{¥} \mathfrak{e} ? \ddot{a} \cdot \ddot{e} \cdot ? \mathfrak{e} ? ?$

ã·?æ??æ©?ç??ã·«æ?¥ç¶?ã·?ã·?æ§?é? ã·§ã·?ã·??

å ¨ä½?ã??å?³ã«ã·?ã??ã¨ã·?ã??ã·æ??ã·?ã·§ã·?ã??



Category

1. æ?¥ã? ã•®æ´»å??

Tags

1. AI
2. Inference Broker
3. LLM
4. LLM Runtime Agent
5. ä,?é??å??å,?
6. å±±æ¢ç??

Date Created

2026å¹8æ??29æ?¥

Author

kazuo-tsubaki

default watermark