

ai_work

- [\[REDACTED\]](#)
 - [\[REDACTED\]](#)
 - [\[REDACTED\] 2](#)
 - [\[REDACTED\] web_book_agent 2026-09-18](#)
 - [\[REDACTED\] CZ-697D 2026-09-18](#)
 - [\[REDACTED\] \[REDACTED\] 2026-09-18](#)
 - [ITX68000L\[REDACTED\]](#)
 - [\[REDACTED\] ITX68000L \[REDACTED\] 2026-09-18](#)
 - [ITX68000L\[REDACTED\] \[REDACTED\] LED\[REDACTED\]](#)
 - [\[REDACTED\] ITX68000L\[REDACTED\] LED](#)
 - [\[REDACTED\]](#)
- [LLM\[REDACTED\]](#)
 - [20260916-\[REDACTED\] -ThinkingOFF](#)
 - [20260916-\[REDACTED\] -ThinkingON](#)
 - [20260916-\[REDACTED\] -Haiku](#)
 - [20260918-\[REDACTED\]](#)

??????

□□□□

???

□□ URL: https://zinichi.net/home/cz-697d2_00/
□□ : □□□□□□ CZ-697D□□□□□□
□□ : 3D□□□□ /DIY
□□ : 2025.02.01

□□□□□□□□□□ CZ-697D□□□□□□ CZ-617D
□□□□□□□□□□□□□□□□□□□□□□□□□□□□ CZ-
697D
□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□

?????

□□□□□□□□□□ CZ-697D□□□□□□ CZ-617D
□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□ CZ-697D□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□

????????????????????

X68000□□□□□□□□□□□□□□□□□□□□□□□□□□□□ 1/2
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□ Z
□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□ Z□□□□□□□□□□

XXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXX

CZ-697D

?????????????LED

XXXXXXXXXXXXXXXXXXXXXXXXXXXX LEDXXXXXXXXXXXX

XXXX 15KHzXXXXXXXXXXXXXXXXXXXX OFF
XXXXXXXXXXXXXXXXXXXX LEDXXXXXXXXXXXX

XXXXXXXXXXXX

XXXXXXXXXXXXXXXXXXXXXXXXXXXX LEDXXXX

??CZ??????????

XXXXXXXXXXXX CZXXXXXXXXXXXX 300
XXXXXXXXXXXX

XXXXXXXXXXXXXXXXXXXXXXXXXXXX

XX

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXXXX XXXXXXX USB-TypeC
XXXXXXXXXXXX 5VXXXXXXXXXXXXXXXXXXXX

#CZXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXX CZ

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

XX @XXXX (@zinfyk) January 18, 2025 XXXXX

CZXXXXXXXXXXXX XXXXXXXXXXXXXXX TypeCXXXXXXXXXXXX

XXXX A C k XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

pic.twitter.com/gCgRWulqsJ — XX @XXXX (@zinfyk) January 12, 2025

XXXXXXXXXXXXXXXXXXXXXXXXXXXX

XXXXXXXXXXXX CZXXXXXXXXXXXXXXXXXXXX USB-TypeCXXXX



??2

□□□□□

????web_book_agent 2026-09-18

??

□□□□□□□□□□

□□□□

□□	□
□□□□	2026-09-18

??????? CZ-697D 2026-09-18

Diagram illustrating the structure of a 697D chromosome. The chromosome is shown as a horizontal bar with a centromere in the middle. The left arm is labeled 'CZ-697D' and the right arm is labeled 'CZ-617D'. The centromere is labeled 'CZ-697D'.

?????

CZ-697D  CZ  CZ-617D
 CZ-697D

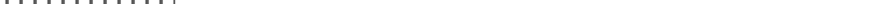

????????????????

X68000  1/2





 $\frac{1}{2}$ 

[illegible]

The diagram shows two horizontal bars. The top bar is a single continuous rectangle divided into 20 equal vertical segments, representing 100%. The bottom bar is a single continuous rectangle divided into 10 equal vertical segments, representing 50%.

????????????????????

XXXXXXXXXX

XXXXXXXXXX

XXXXXXXXXX

XXXXXXXX USB-TypeCXXXXXXXX 5VXXXXXXXXXXXXXXXXXXXX

XX

XXXXXXXXXX

??????????????

XX

項目	仕様	備 考
材質	XXXXXXXXXXXXXXXXXXXX (樹脂) XXXXXXXXXXXXXXXXXXXX (樹脂) XXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXX 材質
色	XXXX LED XXXXXXXXXXXXXXXXXXXX PH XXXXXXXXXXXXXXXXXXXX 材質 (M2-6 2XXXXXX 2)	CZ XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXX
形状	XXXXXXXXXXXXXXXXXXXX	CZ-697D CZ-617DXXXXXXXXXX CZ XXXXXXXXXXXX
CZXXXX	XXXXXXXXXXXXXXXXXXXX (6)XXXXXXXXXX (M4-10 2 M3-15XXXXXX 4)	CZ-697D XXXXXXXXXXXXXXXXXXXX 材質

XX

XX

XXXXXXXXXXXXXXXXXXXXXXXXXXXX 2025 1XXXXXX CZ-697D

XX

XXXXXXXXXXXX

????

XX

XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXX

XXXX

XEiJ+t ? XM6t

Timing diagram for the LED output. The diagram shows four horizontal bars representing different signals over time. The first bar is labeled 'XEij+t' and 'XM6t' with a value of 'X68000'. The second bar is labeled 'X68000'. The third bar is labeled 'LED'. The fourth bar is labeled 'LED'.

??????????????

Windows X68000

LED

SNS

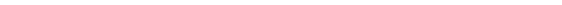


Diagram illustrating the structure of the SNS (Social Network Service) data, showing two rows of boxes representing data points or nodes.

?????

Diagram illustrating a 32-bit bus system with three 32-bit registers (USB, PC, LED) connected to a common bus. Each register is represented by a horizontal row of 32 small squares.

WINDOWS → X68000  **→ USB**  **→ LED**

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
 USB ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐
 LED ☐ ☐ ☐ ☐

???

[illegible]

- 
- PCB 
- 

[illegible]

???????

??????

??????

????

??????????

????????????

Kicad????????

PCB????????????

????????????????????????????????????

????????????????????????????????????

??????

ITX68000L????

LED

????????????????????????????????????

????????????????????????????????

CAD

????????????????????????????????

????????????????????????????????

????????????????????

??

???? ITX68000L?? XM6t

????????????????????????????????

??????

????????????????????????????????

LED

????????????????????????????????

????????????

????

????????

ITX68000L MiniITX PC X68000 PC

140 325 295mm GeForce RTX4070Ti Super X68000 3D

ITX68000L ITX68000L ITX68000L

???????????

ITX68000L HDD X68000 ON/OFF

HDD X68000 PC

HDD

X68000?????ITX68000L?????

X68000 ITX68000L 2

LED

X68000 5.25 2 LED LED

ITX68000L PC X68000

X68000 FDD LED

????????

???????? 2????????

????????????????

????????????????????

LED

??????? XEij+t? XM6t???? X68000

??

???????? LED

????????????????????????????????????

????????

??????? XEij+t / XM6t????????

????????????????????

????????????????????

?? Windows??????? X68000????????????????

LED????????????

????????????????????????????????

??

???? SNS?? ITX68000L??????? LED

????????????????????????????????

????????????????????

????????????????????????????????

????????????????????

???????? SNS????????????????

??????

????????????????????

???????? USB????????????

PC

???????????????? LED????????

WINDOWS → X68000??????? → USB???????? → LED

????????????????

USB????

LED

????????????

????

LED????

X68000

XXXXXXXXXXXXXXXXXXXX

XXX KicadXXXXXXXX PCB
XX
XXXXXXXXXXXXXXXXXXXXXXXXXXXX

XX
XX

XX
XX

XXXXXXXX KicadXX PCBXXXXXXXXXXXXXXXXXXXX

ITX68000L????

LEDXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

CAD
XX
XXXXXXXXXX 140mm
XX
XXXXXXXXXX

CAD

XXXXXXXXXXXX X68000XXXXXXXXXXXXXXXXXXXXXXXXXXXX
XX
XXXXXXXXXXXXXXXXXXXXXXXXXXXX

3D

XXXXXXXX CADXXXXXXXXXXXXXXXXXXXX

??

XXXXXXXXXX

XM6t
XX
XXXXXXXXXXXXXXXXXXXX LED
XX
XX

XXXXXXXXXXXXXXXXXXXXXXXXXXXX

XX
XXXXXXXXXXXX X68000XXXXXXXXXXXXXXXXXXXXXXXXXXXX

XXXXXXXXXX

ITX68000LXXXXXXXXXX

LEDXXXXXXXXXX

????

XXXXXXXXXX

XXXX

[illegible]
$$X_{Eij+t} \square \quad X_{M6t} \square \square \quad X_{68000}$$
[illegible]

LED

LED ITX68000L

????????????????

Windows

X68000

LED

[illegible]

SNS

[illegible][illegible]

????????????????

USB

PC

LED

WINDOWS → X68000

→ USB

→ LED

LED

[illegible]

--	--	--	--	--	--	--	--

[illegible]LED 

??????????

[illegible]

ITX68000L

[illegible]

[illegible]

??????????

Diagram illustrating a 100-bit bus structure. The bus is divided into segments, with labels indicating LED connections at specific points.

??????????

Kicad PCB

ITX68000L????

LED

The diagram illustrates the structure of a Light Emitting Diode (LED). It consists of a p-type semiconductor layer on top of an n-type semiconductor layer. A depletion region is formed at the junction between the two layers, indicated by a dashed line. The p-type layer is labeled 'p' and the n-type layer is labeled 'n'. The depletion region is labeled 'Depletion Region'.

ITX68000L CAD

LLM????

20260916-?????-ThinkingOFF

LLM

gpt-oss:120b Qwen3.8 NVFP4 8bit
Qwen3.8 bf16

- : 2026-09-16 17
- : Claude (Opus 5)
- : workflow/records/20260906_llm_historical_replay_comparison.md
- : workflow/records/llm_eval_20260916/

0. ???

: 3

8bit NVFP4 gpt-oss

					8	4
Qwen3.8 8bit	4	2	4	1	1 1	2
Qwen3.8 NVFP4	4	1	5	1	0	0
gpt-oss:120b	4	0	6	1	0	4

	4	1		
Qwen3.8 8bit	1,053	397.2 Alexa	8,192 7,014	28.6GB 56%
Qwen3.8 NVFP4	367	151.3 Alexa	8,192 7,070	18GB

項目	項目	項目	項目	項目
項目 4	NVFP4	65,536	8,192	項目 3600
項目 4	8bit	MLX 項目	8,192	項目
項目 4	gpt-oss	65,536 → 32,768	8,192 → 16,384 → 20,000	項目

1-4. ?????????????

- thinking: 3項目 off gpt-oss 項目 off 項目
- 項目
 - gpt-oss: 20,000 項目 5,209 16,737 項目 8,192 項目
SKM Alexa 項目 16,384 項目 Alexa 項目
 - NVFP4 8bit: 8,192 項目 7,070 7,014 項目
12,000 項目
- 項目
 - 項目 7,077 15,550 項目
 - gpt-oss 項目 32,768 項目 12,566項目 20,000 項目 65,536
項目 12 15% 項目 Claude Code 項目 2項目
 - MLX 項目 mlx_lm.server項目 8bit 項目
15,550 項目
- 項目
 - gpt-oss 項目 : 項目 66GB項目 12 15%
項目
 - 8bit 項目 : MLX 項目 28.6GB項目 56%項目
- 項目 : 項目 600項目 8bit 項目 397.2
項目

2. ?????

項目 : 項目 Codex 項目 11
項目 4項目 LLM
項目

2-1. ?????

- 項目 2026-09-06項目 11
項目
- 項目 workflow/guidance/codex_studiollm_debug_procedure.md 19項目 7項目
Owner項目 Scope項目 diff項目 Source項目 test
項目 4項目
- 項目 4
項目

mox_law_v3

- airscope_led_initial

- airscope_led_review2

- codex_control_same_turn_retry

- every_golf_serigaya_recovery_retry?????

3-1. ?????

IMPORTANT PASS

- skm_phase5_compact_retry: 2 NUL _text
- alexawebweather_initial_retry: BLOCKER 1 r
- codex_control_same_turn_retry mox_law_v3 airscopecled_review2 every_golf:

Qwen3.8 NVFP4

- mox_law_v2
- mox_law_initial_retry: 2
- alexawebweather_initial_retry: Issue 1 46: This is correct
- alexahandoff_initial: gpt-oss:120b
- 4 0

Qwen3.8 8bit

- mox_law_initial_retry: 3 1
- alexawebweather_initial_retry: 10 1
- skm_phase5_compact_retry: BLOCKER 1 _cleanup_temporary
- codex_control_same_turn_retry: High
- 4

4. Claude ??????????

: 8 0

1? Qwen3.8 8bit

- 8bit 1 1
- 4
- : 4 NVFP4 2.9 2

2? Qwen3.8 NVFP4

- 4 0

- 1 : 1
 1
 1

3? gpt-oss:120b

- 0 4 4
- BLOCKER IMPORTANT PASS 2
- 16,737 4,096

?????????

- 8bit
- NVFP4

5. ?????????Claude ?????????

:

Claude

5-1. ?????????2026-09-16 ?????????5????

-
- bf16 NVFP4 11434 11436
- SHA-256 2
- bf16
-
- 33
- mox_low_v2 2
- 100

5-2. ?????????2026-09-16 ??17?

- codex_control_lab
- 8bit 1 gpt-oss
- Hook
- 8bit 4 900 60 11436 MLX
- gpt-oss 65,536 2

- B : systemb/
194,559 131,072
- thinking ON:
- smartphone_axis_initial B-03
- 3 mox_law_v3 airscope_led_review2
every_golf_serigaya_recovery_retry
- 256GB Mac

10. ????????

: 2

- : workflow/records/llm_eval_20260916/harness/README.md
- : workflow/records/llm_eval_20260916/cleanup_before_after.md

10-1. ????????

11434 Ollama	PID 1167	PID 1167
11435 Alexa	PID 869	PID 869
11436 MLX	bf16 PID 67454	
codex_control_lab	HEAD 97ca15f9	HEAD 09-15
mox_system		

20260916-?????-ThinkingON

thinking OFF 20260916-ThinkingOFF 11 thinking ON

B

Qwen3.8 bf16

- : 2026-09-17
- : Claude (Opus 5)
- thinking OFF : workflow/records/llm_eval_20260916/report/Qwen3.8_OSS260916.html
- : workflow/records/llm_eval_20260916/ answers_thinking_on/ thinking_on_progress.md

0. ???

: gpt-oss thinking ON 11

NVFP4

8bit MLX

0-1. thinking ON ???

						8	11
gpt-oss:120b	5	1	4	0	1	1 mox_law_v2	678 11 18
Qwen3.8 NVFP4	4	1	4	1	1	0	6,079 1 41 19
Qwen3.8 8bit	1	0	0	5	0		3600

0-2. thinking OFF ????

Model	Config OFF → ON	11 Config OFF → ON	Config OFF → ON
gpt-oss:120b	4 6 1 → 5 1 4 1	939 15 39 → 678 11 18	16,737 → 12,549
Qwen3.8 NVFP4	4 1 5 1 → 4 1 4 1 1	547 9 7 → 6,079 1 41 19	7,070 → 31,850
Qwen3.8 8bit	4 2 4 1 → 1 5 5	1,609 26 49 → 1	7,014 → 3,627 1

NVFP4 1 10 OFF 395 6 35 → ON 2,479 41 19
6.3

1. LLM ???

1-1. thinking ON ??????

Model	Context Length	Config	Thinking Config
gpt-oss:120b	131,072	ollama show context length	think: true 4 9 medium
Qwen3.8 NVFP4	262,144	ollama show context length	think: true
Qwen3.8 8bit	262,144 MLX max_position_embeddings	config.json 121 max_position_embeddings	chat_template_kwargs. enable_thinking: true

- gpt-oss 14,746 NVFP4 15,539 8bit 15,550 skm_phase5_compact_retry
- temperature 0 top_p 1 1 3600
- OFF 7
- 4
- 8bit MLX OFF --max-tokens 1024 --chat-template-args {"enable_thinking":false} enable_thinking: true

1-2. ??????

- memory_pressure

 : NVFP4 75% 8bit 61
- 1

2. ??????

thinking OFF

11

thinking OFF

2

- : mox_low_v2 2 alexa_handoff_initial 1 mox_low_initial_retry 3 skm_phase5_compact_retry 1 alexa_weather_initial_retry 1 smartphone_axis_initial
- : mox_low_v3 airscoope_led_initial airscoope_led_review2 codex_control_same_turn_retry every_golf_serigaya_recovery_retry

3. ??????

thinking OFF

1

airscoope_led_review2

3-1. ??????

- :
- : 1
- :
- : 3600 : MLX
-

3-2. 11?????????thinking ON?

		gpt-oss	NVFP4	8bit
mox_low_v2	2	/ 26.3 / 1,725	/ 394.0 / 24,660	8 50

項目	項目	gpt-oss	NVFP4	8bit
alexa_handoff_initial	項目 1項目	項目 / 27.2項目 / 1,448	項目 / 122.0項目 / 6,261	項目
mox_law_initial_retry	項目 3項目	項目 / 46.0項目 / 2,613	項目 / 368.8項目 / 20,688	項目
skm_phase5_compact_retry	項目 1項目	項目 / 141.2項目 / 7,969	項目 / 631.4項目 / 31,850	項目
alexa_weather_initial_retry	項目 1項目	項目 / 208.4項目 / 12,549	項目 / 3600項目	項目
smartphone_axis_initial	項目	項目 / 52.3項目 / 3,659	項目 / 135.0項目 / 7,095	項目 8項目 50項目
mox_law_v3	項目	項目 / 8.4項目 / 542	項目 / 26.1項目 / 1,633	項目
airscope_led_review2	OK	項目 / 50.5項目 / 440	項目 / 26.8項目 / 1,098	項目 / 178.3項目 / 3,627
every_golf_serigaya_recovery_retry	項目	項目 / 43.2項目 / 2,226	項目 / 563.6項目 / 17,674	項目
airscope_led_initial	項目 BLOCKER項目	項目 / 26.3項目 / 1,810	項目 / 59.1項目 / 3,138	項目 8項目 50項目
codex_control_same_turn_retry	項目	項目 / 48.2項目 / 3,294	項目 / 152.3項目 / 8,430	項目 8項目 50項目

項目 / 項目 / 項目

項目 5項目

8bit

3-3. thinking OFF ??????????

項目	項目	OFF	ON	項目
gpt-oss	mox_law_v2	項目	項目	項目 2 項目 項目 項目
gpt-oss	airscope_led_initial	項目	項目	項目 項目 BLOCKER 項目
NVFP4	alexa_weather_initial_retry	項目	項目	3600 項目 項目

項目

OFF 項目 8bit 項目

3-4. ??????????????

gpt-oss:120b

- mox_law_v2:

項目

[]
 []
 • skm_phase5_compact_retry: [] BLOCKER 1[] Studio [] NUL
 [] _read_knowledge [] S_ISREG
 [] NUL
 [] 1
 [] Studio [] 497[]
 report [] FIFO
 [] PASS[]
 • alexa_weather_initial_retry: OFF [] BLOCKER[] r
 [] PASS[]
 • alexa_handoff_initial: [] NON-BLOCKING []
 • mox_law_initial_retry: NO_FINDING []
 • airscopec_led_initial: CONDITIONAL_OK[] BLOCKER
 []

Qwen3.8 NVFP4

• mox_law_v2: [] 2[] 1[] DB
 [] operation_re [] DB [] []
 [] 130[]
 • skm_phase5_compact_retry: FIFO [] 360[] _mapping(studio_value,
 ...) [] studio_entry [] NameError []
 • codex_control_same_turn_retry: [] _text_content
 []
 • alexa_handoff_initial: [] OFF
 [] gpt-oss:120b[]
 • mox_law_initial_retry[] every_golf_serigaya_recovery_retry: NO_FINDING []

Qwen3.8 8bit

• airscopec_led_review2: VERDICT OK[] BLOCKER [] 1
 []

3-5. ????????

[]

- NVFP4 [] mox_law_v2: [] []\n[]
 []
 []
- NVFP4 [] skm_phase5_compact_retry: [] FIFO[]
- NVFP4 [] mox_law_initial_retry: [] 3[]

[] NVFP4 [] 3[] gpt-oss
 [] answers_thinking_on/
 []

- mlx_lm.server [REDACTED] message.reasoning [REDACTED]
server.py 1353 [REDACTED] 1358 [REDACTED] runner.py [REDACTED]

5-2. ???11:18?12:57?: ?????2???????

- 1 [REDACTED] airscopec_led_review2 [REDACTED] 174.1 [REDACTED] 13,608 [REDACTED]
- 2 [REDACTED] mox_law_v3 [REDACTED] RuntimeError: [metal::malloc] Resource
limit (499000) exceeded. [REDACTED] 12:21:43 [REDACTED]
3600 [REDACTED]
- 3 [REDACTED]

5-3. ???13:09?14:09?: 1?????????????

- [REDACTED] 1 [REDACTED] 2 [REDACTED] 1
[REDACTED]
- 1 [REDACTED] airscopec_led_review2 [REDACTED] 178.3 [REDACTED]
- 2 [REDACTED] mox_law_v3 [REDACTED] 13:19 [REDACTED] 13:29
[REDACTED] runner [REDACTED] 13:33:26 [REDACTED]
- [REDACTED] Traceback [REDACTED] 10 [REDACTED] runner.py
[REDACTED]
- [REDACTED]

[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
airscopec_led_initial	13:33:32	13:42:22	[REDACTED] 8 [REDACTED] 50 [REDACTED]
smartphone_axis_initial	13:42:27	13:51:18	[REDACTED] 8 [REDACTED] 51 [REDACTED]
codex_control_same_turn_r etry	13:51:23	14:00:13	[REDACTED] 8 [REDACTED] 50 [REDACTED]
mox_law_v2	14:00:19	14:09:09	[REDACTED] 8 [REDACTED] 50 [REDACTED]

[REDACTED] 10 [REDACTED]
10 [REDACTED]

5-4. ????

- 5 [REDACTED] 4 [REDACTED] NVFP4 [REDACTED] ON [REDACTED] 1
[REDACTED] 8bit [REDACTED] 8 [REDACTED] 50
[REDACTED]
- [REDACTED] stop_q8_thinking_on.flag [REDACTED] 5 [REDACTED] alexa_handoff_initial [REDACTED]
mox_law_initial_retry [REDACTED] every_golf_serigaya_recovery_retry [REDACTED] alexa_weather_initial_retry [REDACTED]
skm_phase5_compact_retry [REDACTED] 14:09:14 [REDACTED] 14:09:36
[REDACTED]

5-5. ??????????????

- [REDACTED] : [REDACTED] Metal
[REDACTED] MLX [REDACTED] allocator.h 72 [REDACTED] 73 [REDACTED]

num_resources_ resource_limit_ 61 62%
8 50
• : mlx_lm
Qwen3.8
• Qwen3.8 Ollama NVFP4

6. Claude

: Claude

- 8bit 262,144 thinking ON
- 10 11:21 MLX 1 2
3600 gpt-oss 12:57 1 36
- 8bit 1 4
- OFF NVFP4
1
- 3 13:12:25 13:33:26 21
3
- 1 2 30 NVFP4 1
52 gpt-oss 58
- /add-dir 2

7.

: thinking OFF 6 thinking ON

8.

: 4,096 600 thinking ON

- 4,096 qwen_shared.py 356
tools/codex-studiollm-debug-client.py 115
116

- gpt-oss ON 11 2 4,096 skm_phase5_compact_retry 7,969 alexa_weather_initial_retry 12,549 208.4 600
- NVFP4 ON 10 7 4,096 31,850 skm_phase5_compact_retry 631.4 alexa_weather_initial_retry 600
- thinking

9.

- 1
- Claude
- 7 4 OFF
- 1
- gpt-oss medium
- NVFP4 3
- 1 1

10.

- Qwen3.8 bf16:
- B : thinking OFF ON
- Qwen3.8 8bit thinking ON: MLX Metal
- gpt-oss low high
- B-03
- 3 8bit OFF mox_low_v3 airscope_led_review2 every_golf_serigaya_recovery_retry
- gpt-oss thinking ON 4,096

11.

:

11-1.

- : workflow/records/llm_eval_20260916/harness/ README.md
- thinking ON 1 11434 11436

- /usr/bin/python3 run_thinking_on_unattended.py --model nvfp4 --context 262144
- /usr/bin/python3 run_thinking_on_unattended.py --model gptoss --context 131072
- /usr/bin/python3 run_thinking_on_unattended.py --model q8 --context 262144 8bit 1
- : /usr/bin/python3 runner.py export --series A --thinking on answers_thinking_on/ OFF
- 8bit thinking ON workflow/records/llm_eval_20260916/stop_q8_thinking_on.flag 8bit thinking ON
- thinking_on_progress.md 8bit thinking_on_mlx_server.log

11-2. 2026-09-17 14

11434 Ollama		
11435 Alexa	PID 869 09-16	PID 869
11436 MLX		
qwen-shared.py	PID 46674 9 15	PID 46674
codex_control_lab	09-15	
mox_system		

20260916-?-Haiku

thinking OFF thinking ON 11 Claude Haiku

LLM

1

- : 2026-09-17
- : Claude (Opus 5)
- : 20260916-ThinkingOFF thinking OFF 20260916-ThinkingON thinking ON
- Haiku : workflow/records/llm_eval_20260916/haiku_answers_20260917.md

0. ?

: Haiku gpt-oss thinking ON 1

BLOCKER 4 3

3

						8	4	4
Claude Haiku	5	1	4	0	1	1	0	3
gpt-oss:120b thinking ON	5	1	4	0	1	1	3	0

	11	1	
Claude Haiku	931 15 31	29.1 169.2	2 49 11
gpt-oss:120b thinking ON	678 11 18	8.4 208.4	11 19 1

Haiku
 1. LLM ???

1. LLM ???

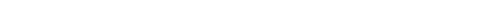
: Mac Anthropic API Claude Code Haiku
 1. LLM ???

1-1. thinking OFF?thinking ON?Haiku ??????

	thinking OFF	thinking ON	Claude Haiku
	Mac	Mac	Anthropic
	runner.py Ollama MLX API		Claude Code model: haiku
			Read
			Read 2
	32,768 65,536 8bit		
	8,192 20,000	–	
thinking	off	on gpt-oss medium	
temperature top_p	0 1	0 1	
	4		4
	1		
	1	1	11
	1	1	1

1-2. Haiku ??????

- Read Web

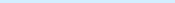
- 
 - 
 -  4
- 
- 

2. ?????

[illegible]

- `00000000` : `mox_law_v2` `2000` `alexa_handoff_initial` `1000` `mox_law_initial_retry` `3000` `skm_phase5_compact_retry` `1000` `alexa_weather_initial_retry` `1000` `smartphone_axis_initial``00000000`
- `00000000000000000000` : `mox_law_v3` `airscope_led_initial` `airscope_led_review2` `codex_control_same_turn_retry` `every_golf_serigaya_recovery_retry`

3. ?????

 :  thinking OFF  thinking ON


3-1. ?????

- [illegible]

3-2. Haiku ?11????????

項目	単位	値	割合	総数	割合
mox_law_v2	1000 回	1000 回	49.9%	36,818	100.0%
alexa_handoff_initial	1000 回	1000 回	83.1%	44,385	100.0%
mox_law_initial_retry	1000 回	1000 回	169.2%	55,616	100.0%

項目	項目	項目	項目	項目	項目
skm_phase5_com pact_retry	項目 1項目	項目	110.1項目	60,380	項目
alexa_weather_in itial_retry	項目 1項目	項目	138.1項目	60,373	項目
smartphone_axis _initial	項目	項目	35.5項目	34,229	項目
mox_law_v3	項目	項目	38.2項目	33,698	項目
airscope_led_revi ew2	OK	項目	29.1項目	32,988	項目
every_golf_seriga ya_recovery_retr y	項目	項目	129.8項目	54,844	項目 PASS項目
airscope_led_initi al	項目 BLOCKER項目	項目	37.6項目	34,308	項目
codex_control_sa me_turn_retry	項目	項目	110.4項目	42,708	項目

3-3. ??????????thinking OFF?thinking ON?Haiku?

項目	項目	gpt-oss OFF	gpt-oss ON	NVFP4 OFF	NVFP4 ON	8bit OFF	8bit ON	Haiku
mox_law_ v2	項目 2項目	項目	項目	項目	項目	項目	項目	項目
alexa_han doff_initial	項目 1項目	項目	項目	項目	項目	項目	項目	項目
mox_law_i nitial_retry	項目 3項目	項目	項目	項目	項目	項目	項目	項目
skm_phas e5_compa ct_retry	項目 1項目	項目	項目	項目	項目	項目	項目	項目
alexa_wea ther_initial _retry	項目 1項目	項目	項目	項目	項目	項目	項目	項目
smartphon e_axis_init ial	項目	項目	項目	項目	項目	項目	項目	項目
mox_law_ v3	項目	項目	項目	項目	項目	項目	項目	項目
airscope_l ed_review 2	OK	項目	項目	項目	項目	項目	項目	項目

		gpt-oss OFF	gpt-oss ON	NVFP4 OFF	NVFP4 ON	8bit OFF	8bit ON	Haiku
every_golf_serigaya_recovery_retry								
airscope_initial	BLOCKER							
codex_control_same_turn_retry								

MLX thinking ON 8bit thinking ON

thinking ON 5

3-4. Haiku

- mox_law_v2:
operation_re operation_re
- codex_control_same_turn_retry: same_turn_retry_initial_prompt_mismatch
106
- alexandra_handoff_initial: PASS Hook
gpt-oss:120b
- mox_law_initial_retry: defect PASS
skm_phase5_compact_retry: PASS FIFO
- alexandra_weather_initial_retry: NO_FINDING 9 stopped.wav
SHA 451
gpt-oss thinking ON
- every_golf_serigaya_recovery_retry: NO_FINDING
- 11 BLOCKER 1

4. Claude

: Haiku 8

gpt-oss?thinking ON

- BLOCKER 1 gpt-oss ON
4 3

- ☐ ☐ 2 ☐ 49 ☐
- ☐ Mac ☐

gpt-oss?thinking ON???????

- ☐ 4 ☐ 3
☐
☐
- ☐
☐
- ☐

???

- ☐ Codex ☐ LLM
☐
- ☐ gpt-oss ☐ thinking ON ☐

5. ??????????????

- ☐ thinking ☐
- ☐ Haiku ☐
- ☐
☐
- ☐
☐
- ☐ 1 ☐
- ☐ thinking OFF ☐ ON ☐ Claude
☐

6. ??????????

- API
☐
☐
- Haiku ☐ ☐ results.jsonl
☐ haiku_answers_20260917.md ☐

7. ????

- API ☐ : Claude Code ☐ 1
☐ model: haiku ☐ 1-2 ☐
duration_ms ☐ subagent_tokens ☐

- API : runner.py ☐ Anthropic API thinking OFF ☐
ON
- harness/cases.json path

8. thinking OFF?thinking ON?Haiku ????

: 3 1 Haiku

[illegible]

0000 00	0000	00	000	000	0000 0000 0000	0000	0000 0000 8 0000	0000 00 4 0000 0000	11 0000 00
Claude Haiku 0000	0000	5	1	4	0	1	0000 1	0	93100 150 31 0000 0000 000 20 4900

- 11 [] thinking ON [] Haiku [] 5 [] 1 []
- [] Haiku [] NVFP4 [] thinking OFF [] 0 []
- [] NVFP4 [] thinking OFF [] gpt-oss [] thinking ON []
- [] 8 []
- [] Haiku [] 1 []
[]

9. ?????????????40%????30%??30%?

Category	Percentage
Open source LLMs	40%
Closed source LLMs	30%
Proprietary LLMs	30%
Claude	30%

9-1. ??

項目	項目名	項目値 40%	項目値 30%	項目値 30%	項目値
1	Qwen3.8 NVFP4 thinking OFF	45	87.5	100	74.3
2	gpt-oss:120b thinking ON	55	87.5	80.7	72.5
3	Claude Haiku 	55	100	58.8	69.6
4	Qwen3.8 8bit thinking OFF	50	91.3	34.0	57.6
5	gpt-oss:120b thinking OFF	40	75.0	58.3	56.0
6	Qwen3.8 NVFP4 thinking ON	45	80.8	9.0	44.9
7	Qwen3.8 8bit thinking ON	10	61.8	0	22.5

$$\frac{1}{10} = \frac{4}{10} \times 0.4 \frac{1}{10} \times 0.3 \frac{1}{10} \times 0.3 \frac{1}{10} \times 100 \frac{1}{10}$$

9-2. ????????

- $\text{smartphone_axis_initial} = \text{smartphone_axis_initial} \times 0.5 \div \text{smartphone_axis_initial} \times 10 \times 100$
- $\text{smartphone_axis_initial} \times 10 \times 100$
- $\text{smartphone_axis_initial} \times 10 \times 100$
- $\text{smartphone_axis_initial} \times 10 \times 100$

9-3. ???????

Diagram illustrating the LLM-Codex architecture. The input sequence is divided into two parts: the first 10 tokens are processed by the LLM, and the remaining 20 tokens are processed by the Codex.

- [] = [] :
[]
 - [] = [] : []
 - [] = 3 [] :
[]
[]
- [] = [] $21 \div [] \times 100$
 - LLM [] 0 [] Haiku [] 11
[] 0.49 [] 2.45 [] API [] [] 1 [] 5 [] 100
[] 490,347 []
• [] : [] Haiku [] API
[] Claude Opus 5 [] 5 [] 25 [] 100 [] 1
[] Haiku []
Haiku [] gpt-oss [] thinking ON [] 3 [] 3 [] 1
[] 0.16 [] 0.82 [] 1
[] Codex
[] 1 []

モデル名	思考モード ON/OFF	思考ステップ数	思考時間 (秒)	推理精度 (Top-1)	推理速度 (tokens/sec)
Claude Haiku v1.0	0	0	7	21	100%
Qwen3.8 8bit thinking OFF	2	0	7	23	91.3%
gpt-oss:120b thinking ON	3	0	7	24	87.5%
Qwen3.8 NVFP4 thinking OFF	0	0	8	24	87.5%
Qwen3.8 NVFP4 thinking ON	1	1	8	26	80.8%

9-6. ????????????

- 1 2 1.8
- 3 1 Haiku 69.6
- gpt-oss thinking ON 67.2 2
- 7 8
- Haiku

10. ?????????????????????

: 11

11

10-1. ?????

		thinking			1	
gpt-oss:120b thinking ON	Ollama 11434	think: true = medium	32,768 25,120 131,072	16,384 12,549	600 208.4	131,072 23%
gpt-oss:120b thinking OFF	Ollama 11434	think: false	32,768 65,536 12 15% Claude Code	20,000 off 16,737	600 284.9	12 15% 65,536
Qwen3.8 NVFP4 thinking OFF	Ollama 11434	think: false	32,768 20,264	12,000 8,192 7,070	600 151.3	18GB

配置	模型	thinking 开关	显存	内存	1.推理 速度	吞吐量 TPS
Qwen3.8 NVFP4 thinking ON	Ollama 11434	think: true	65,536 显存 47,387 262,144 显存	40,000 显存 31,850	600 显存 skm_phase5_c ompact_retry 631.4 alex_a_weather _initial_retry 3600 显存	262,144 显存 75%
Qwen3.8 8bit thinking OFF	MLX mlx_lm.server 11436	enable_thinking: false	显存 显存 15,550 显存 显存	显存 12,000 8,192 显存 显存 7,014 显存 --max-tokens 1024 显存 显存	600 显存 显存 397.2 显存	显存 28.6GB 56%

显存 11显存 5显存 11
显存 显存

10-2. ??

- Ollama
显存
显存
- 显存
显存 OFF 显存 NVFP4 显存 46显存
- 显存 4,096 qwen_shared.py 356显存 4,096
显存
- 显存 600 Qwen3.8 NVFP4 thinking ON
显存
- 11434 Ollama 11436 MLX 8bit
显存
显存

10-3. ????????????

- Qwen3.8 8bit thinking ON : 显存 8 50 MLX [metal::malloc] Metal
Resource limit (499000) exceeded 显存 11显存
显存
- Claude Haiku: 显存 thinking
显存 API 显存
- Qwen3.8 bf16: 显存

20260918-????????

0. ??????????

LLM

- : 60KB 1.2
- 1 25KB 2 54KB 2 28KB 3 36KB 1 3
- 1 129KB 1 102KB 1 NVFP4 12
- 8bit: 5,723 278 36,272 2,062

: ① ②
③1 NG Haiku Luna

1. ?????

6 Opus
Opus
NG 3 3 NG

- 1: codex_control_lab f7c3862 2 25KB
- 2: mox_system f3e69844 hub.py contract.json 54KB
- 1: alexa_cfw00 6e7e0f0 2 129KB
- 2: codex_control_lab e3cb8a4 1 28KB
- 3: mox_system 6a3aa19c 222 36KB
- 1: mox_system 0e1e43b6 7 102KB

Qwen3.8 NVFP4 gpt-oss:120b Qwen3.8 8bit
Claude Haiku

2. ?????

項目	1	2	1	2	3	1
NVFP4 項目	2項目	項目	項目	3項目	2項目	項目
NVFP4 項目	2項目	項目	項目	項目	項目	項目
gpt-oss 項目	2項目	項目	項目	項目	項目	項目
gpt-oss 項目	1項目	項目	項目	項目	項目	項目
8bit 項目	1項目	1項目	項目	項目項目 項目	項目	項目
Claude Haiku 項目	1項目	1項目	項目	項目 1項目 NG 項目	項目 1項目 NG 項目	3項目

項目 22項目項目 11項目 9項目 2項目

3. ???????

Haiku

項目	項目	項目	項目	項目
NVFP4 項目	1	2	152.4	11,438
NVFP4 項目	2	3	498.1	33,355
NVFP4 項目	1	3	2,358.0	109,021
NVFP4 項目	2	3	291.6	20,415
NVFP4 項目	3	2	137.9	9,214
NVFP4 項目	1	3	1,048.9	61,062
NVFP4 項目	1	2	189.1	13,906
NVFP4 項目	2	3	870.8	52,491
gpt-oss 項目	1	2	176.7	11,869
gpt-oss 項目	2	3	539.1	33,348
gpt-oss 項目	1	1	92.1	6,251
gpt-oss 項目	2	3	517.6	32,256
8bit 項目	1	1	278.1	5,723
8bit 項目	2	1	556.2	11,124

8bit	1	3 3	6,499.6	72,554
8bit	2	2	697.3	13,941
Claude Haiku	1	1	95.1	47,510
Claude Haiku	2	1	154.3	62,947
Claude Haiku	1	1	585.9	140,865
Claude Haiku	1	3	728.2	335,862

8bit15,723278136,272

2,0626.37.4

4.

9

NVFP42		hub.py1
NVFP41	353/353	60
NVFP41	64/64	2
NVFP42	110/110	130
gpt-oss2	109/110	SQL1
gpt-oss2		3150
8bit1	353/353	3
8bit2	32/34→	
Claude Haiku1		32,0002

4NVFP411NVFP42

8bit1

5.

NG

-  : NVFP4  1  353/353  60 
 NVFP4  1  64/64  2  NVFP4  2  110/110  30 
-  NG  : gpt-oss  2 1  110/110  1

-  NG


Haiku ☐ Claude Code ☐ API ☐

7. ???????????

項目	単位	単位換算
1GB	1024MB	1024MB
1MB	1024KB	1024KB
1KB	1024B	1024B

-  1  NG  Haiku  Luna  1  3
-  30 
- 
-  1  7  95 
- 

