

VERSION (修订版本)

Version	Data	Updata
V1	2025/01/01	initial release(Beta)
V2	2025/02/28 12:19	Over CPU/GPU/Cache
V3	2025/03/01 7:19	Over NPU/ISP/BP
V4	2025/03/01 11:39	Small Debug

DATA

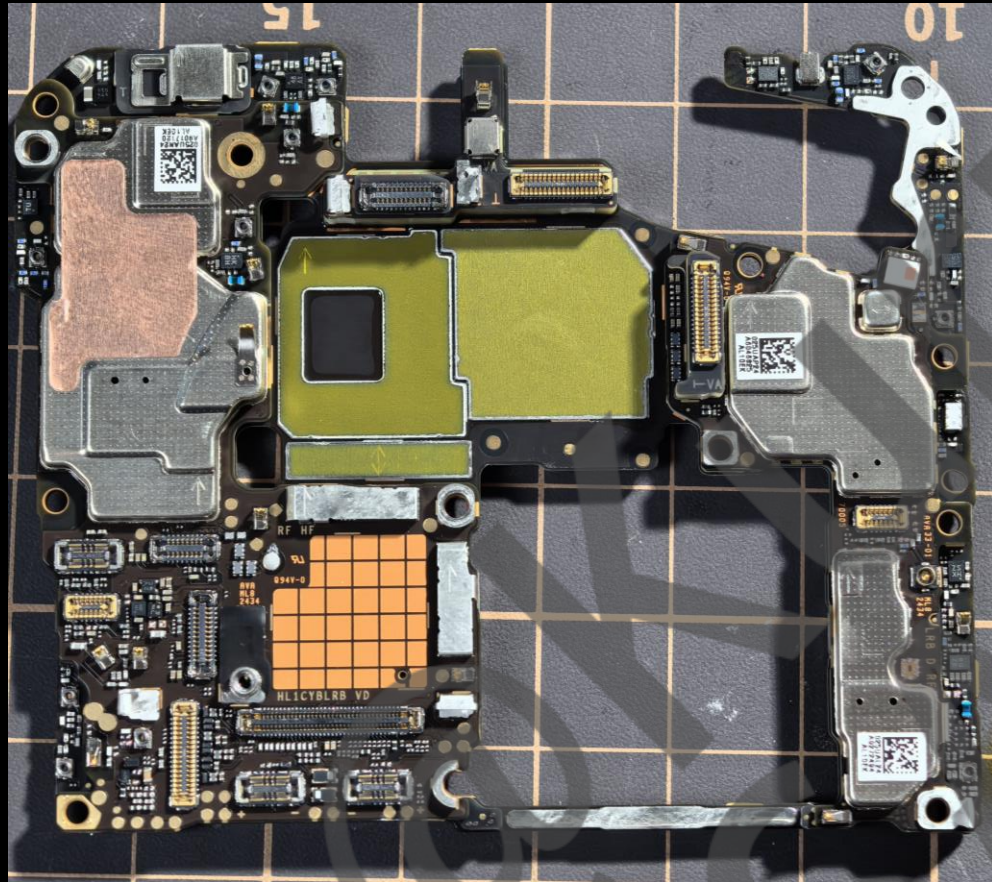
118页

Bilibili工房文件

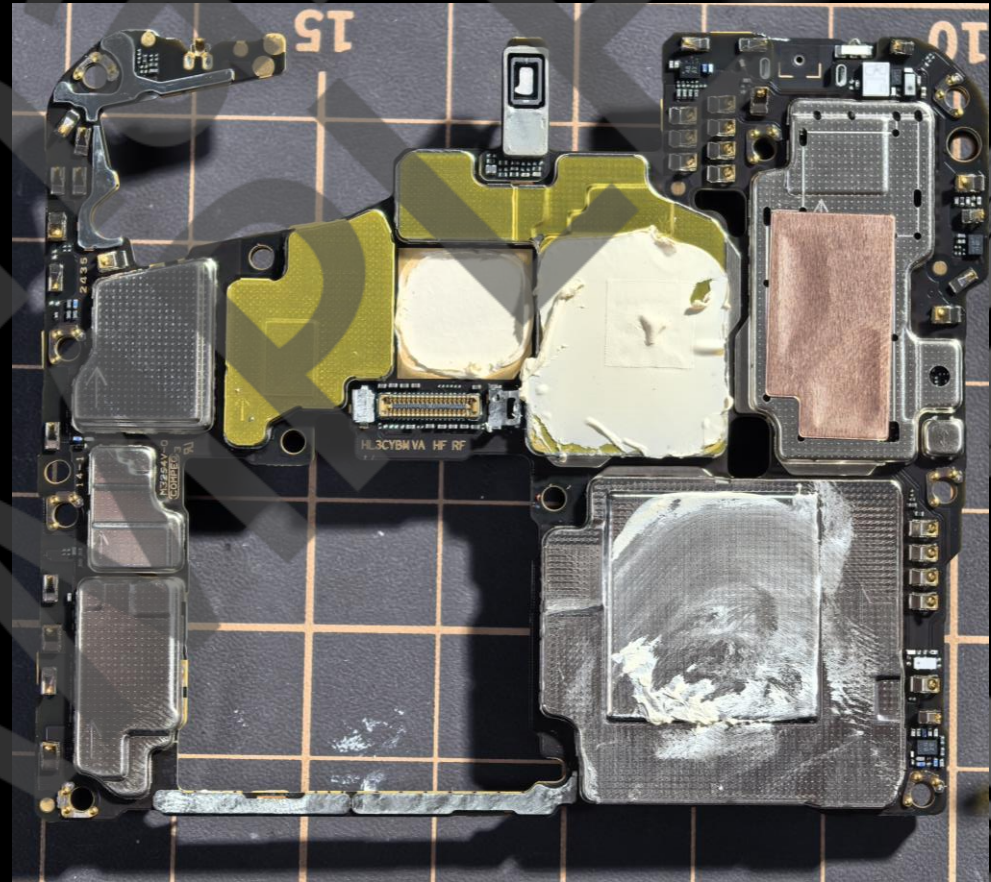


MOTHER BOARD

MOTHERBOARD SHOT

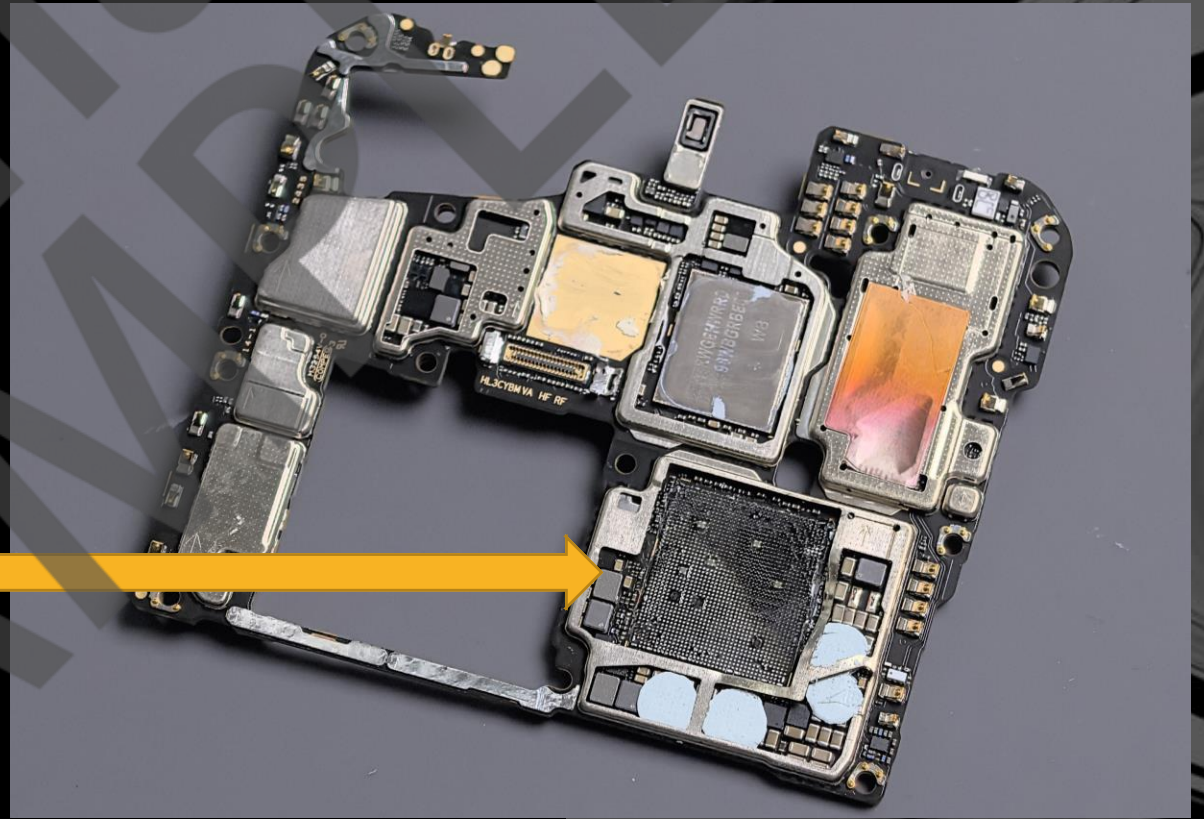
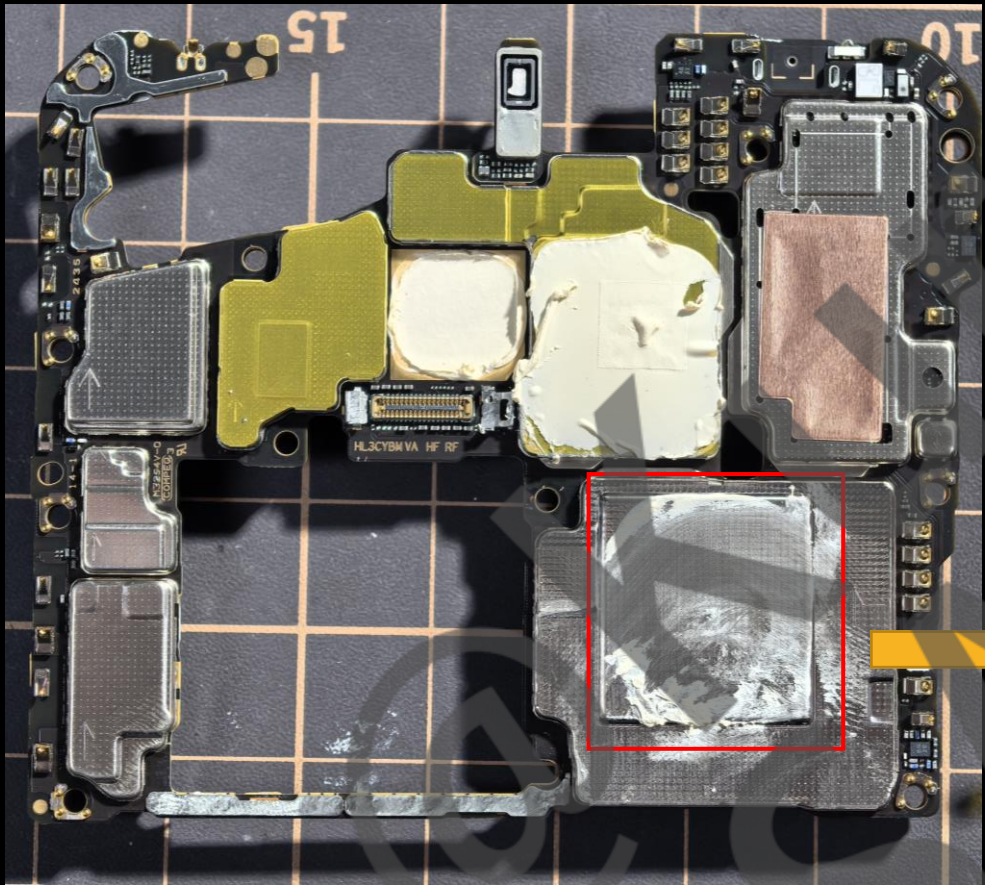


Front shot

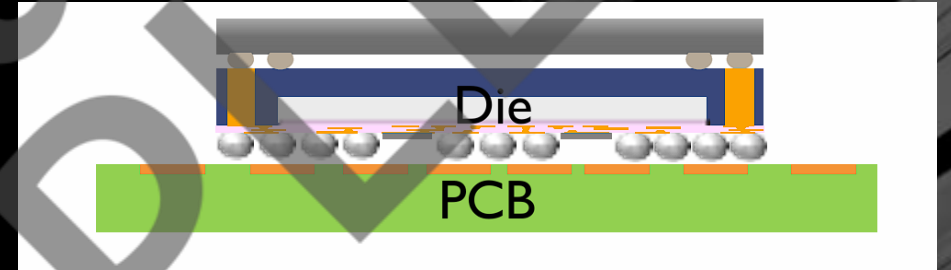
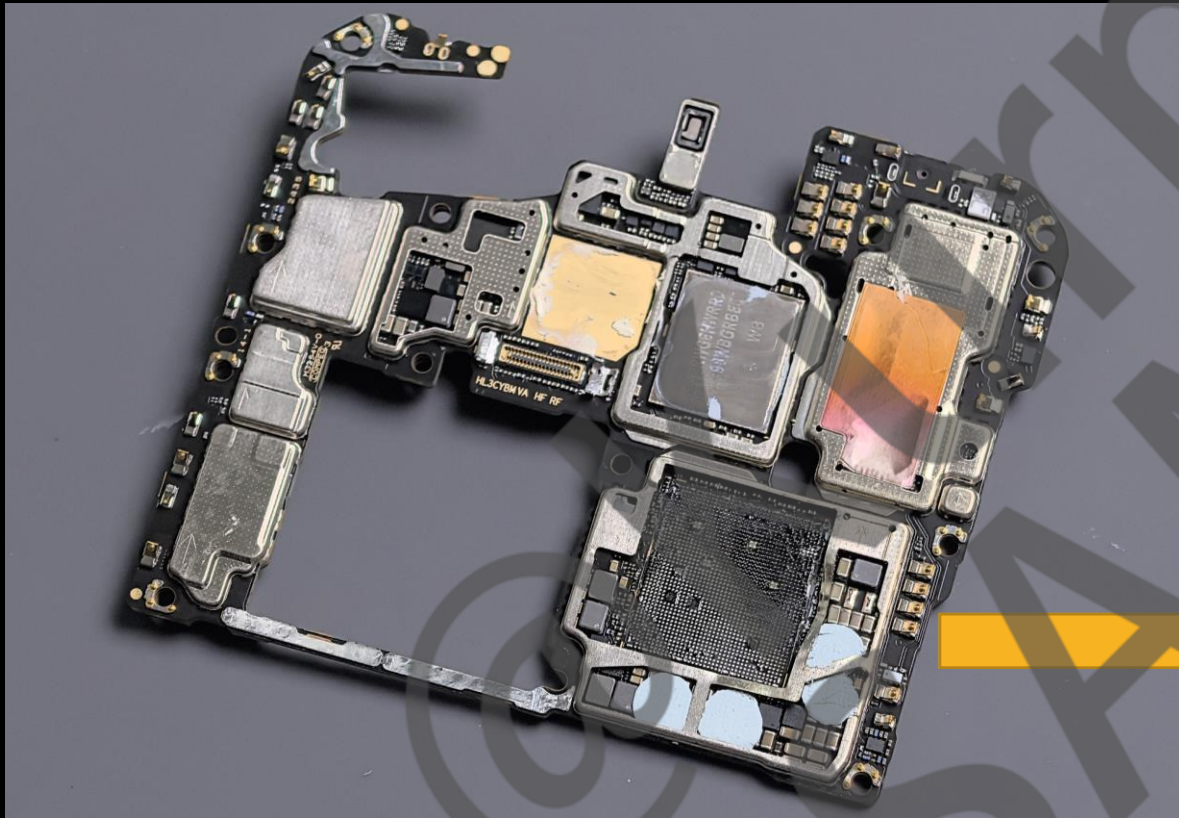


Back shot

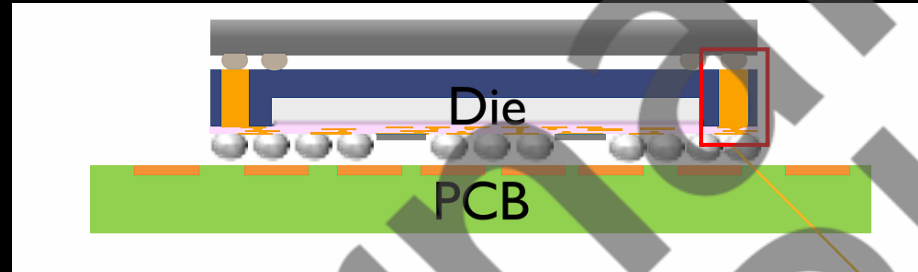
MOTHERBOARD SHOT



MOTHERBOARD SHOT



PACKAGE SHOT

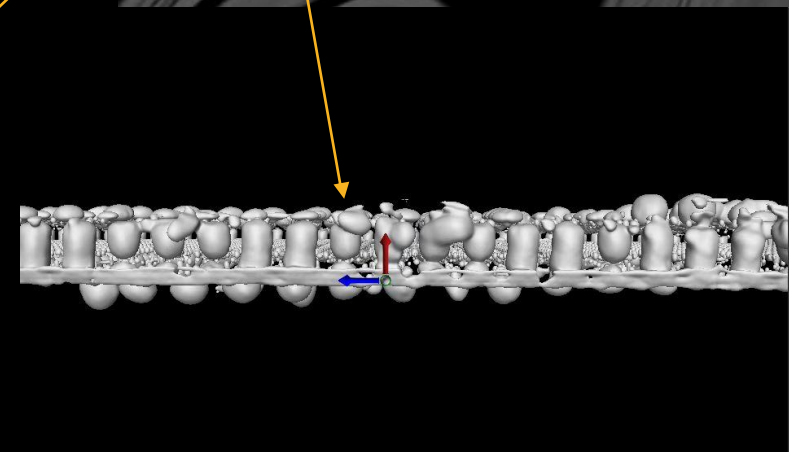


Top layer

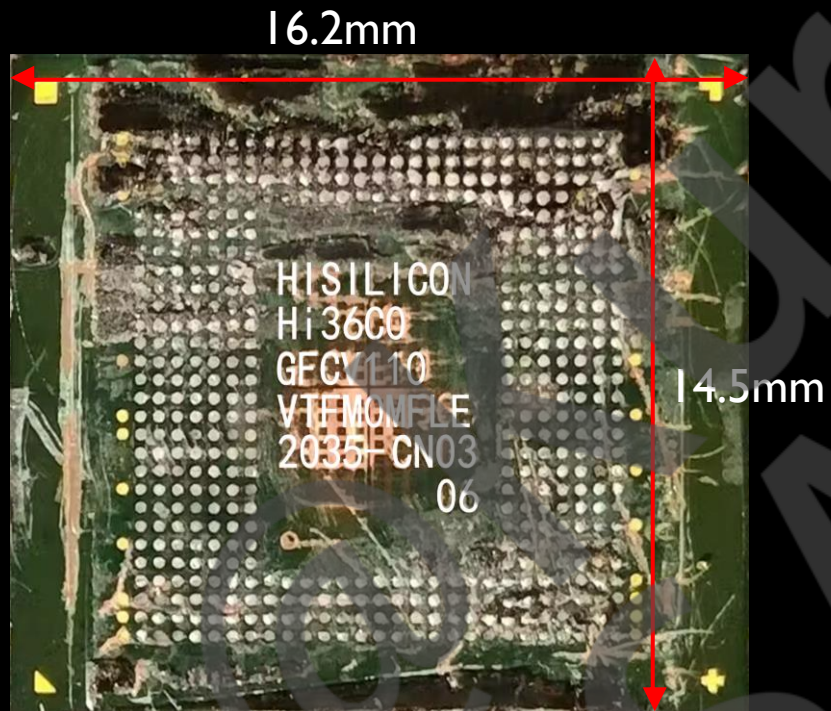


Middle layer

Filp Chip
Package on Package
(FC POP)

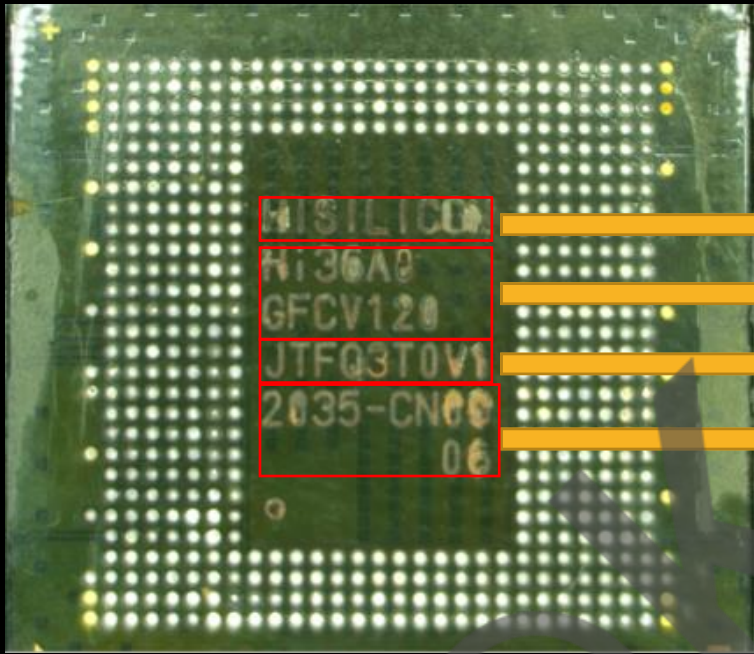


MOTHERBOARD SHOT



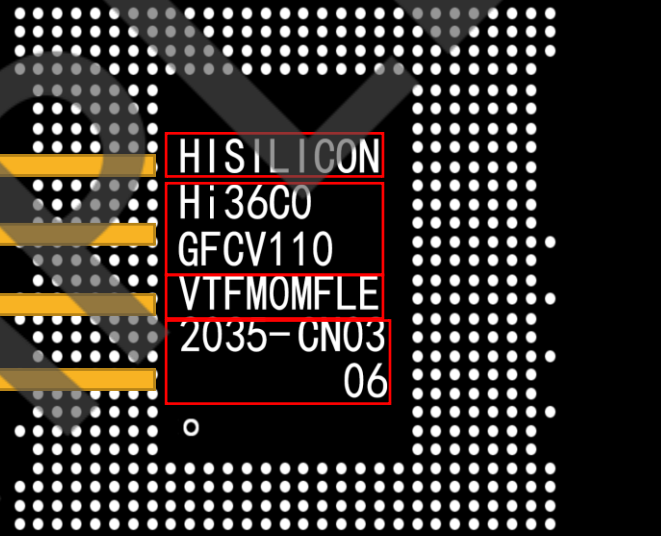
Package size: 234.9mm²
Package Length: 16.2mm
Package Width: 14.5mm

PACKAGE SHOT



Kirin9000s Topmark

Hisilicon Kirin
Product code
Traceability code
Fixed code



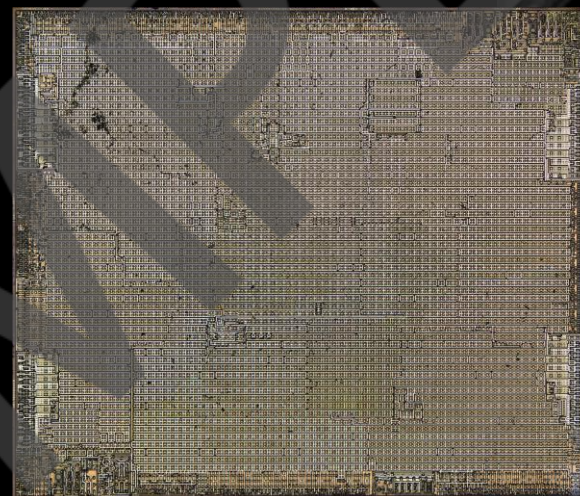
Kirin 9020 Topmark

PACKAGE SHOT



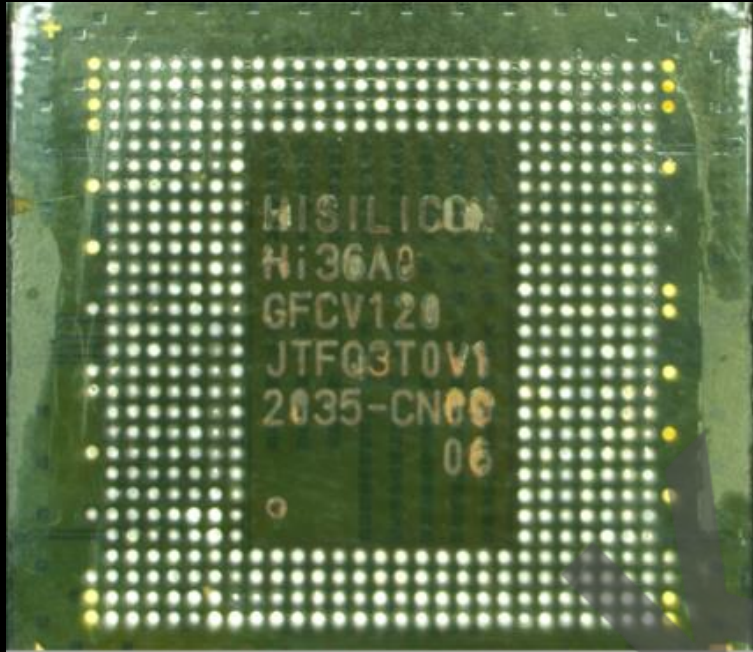
Kirin9020 Package

Decaped



Kirin9020 Die photo

PACKAGE SHOT



Kirin9000s Topmark

Mate60series

Half step



Kirin9010 Topmark

Pura70series

Next Gen

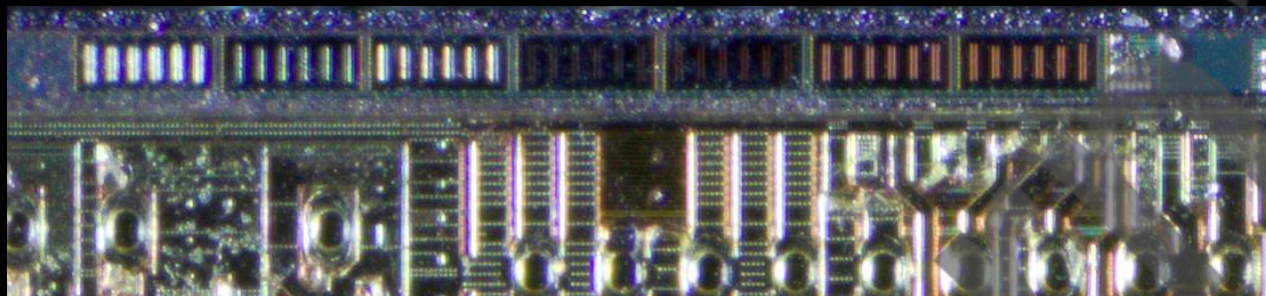


Kirin 9020 Topmark

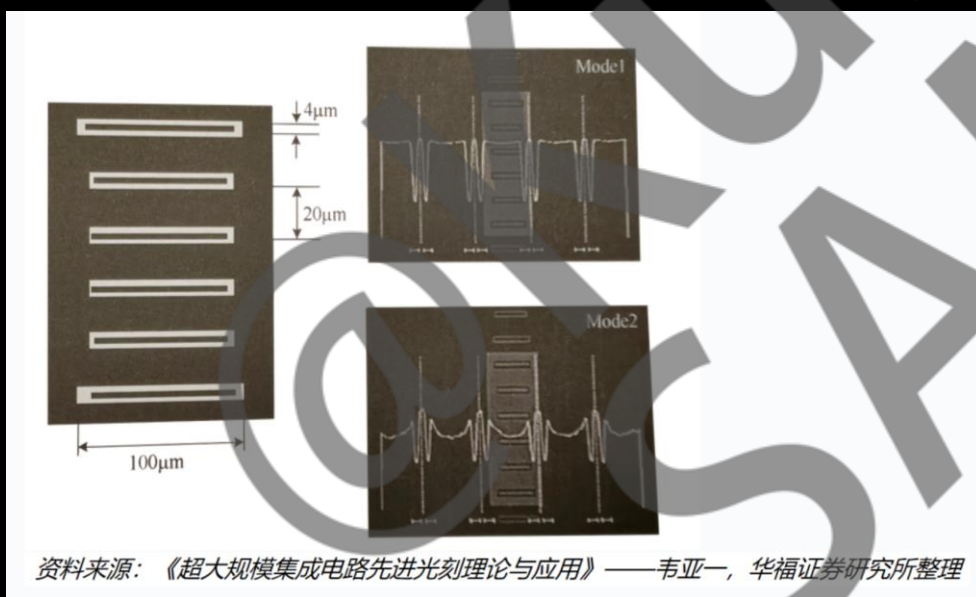
Mate70 series

DIE MARK ANALYZE

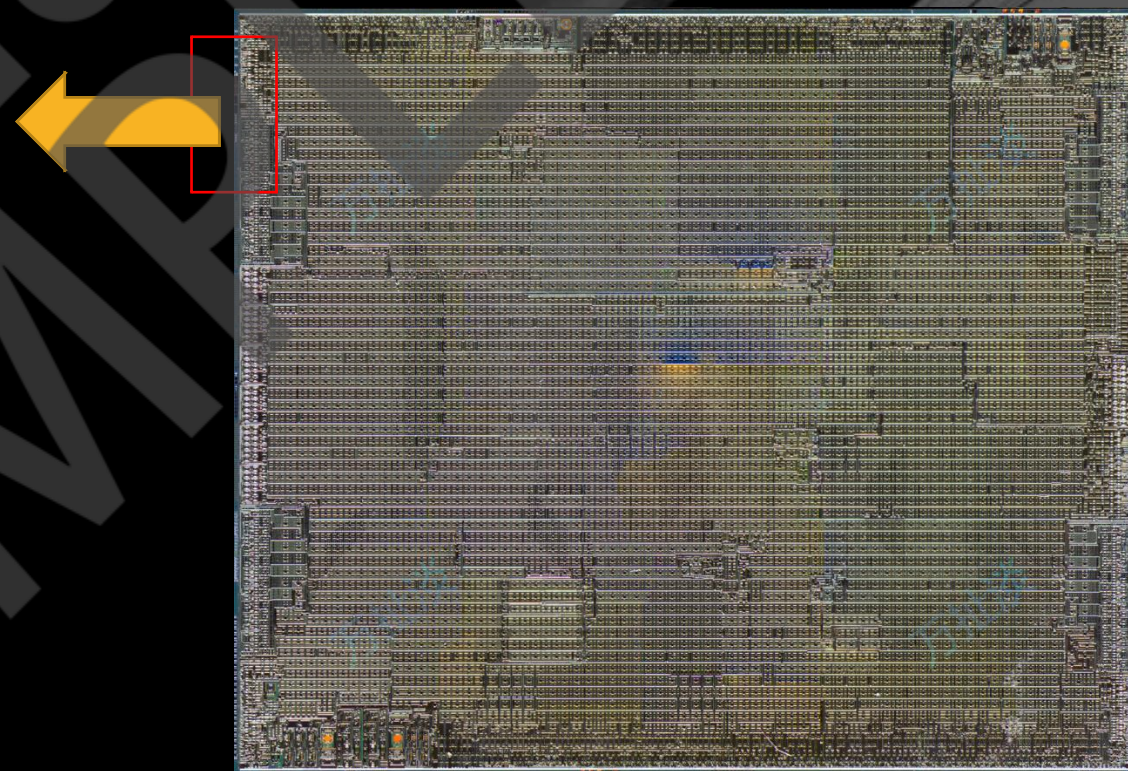
PACKAGE SHOT



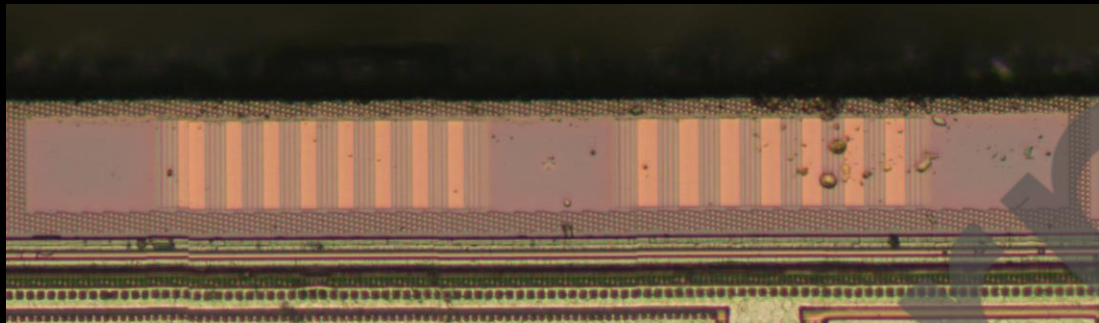
Canon Alignment Mark



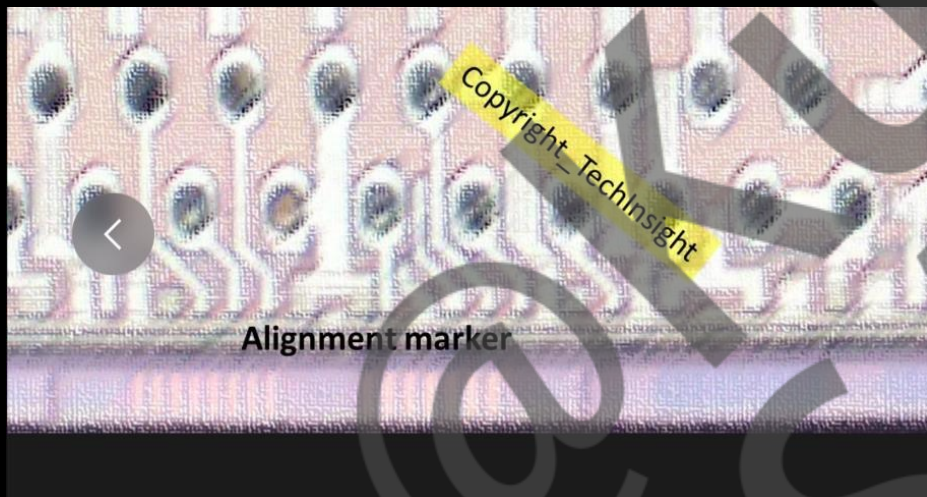
资料来源:《超大规模集成电路先进光刻理论与应用》——韦亚一, 华福证券研究所整理



PACKAGE SHOT

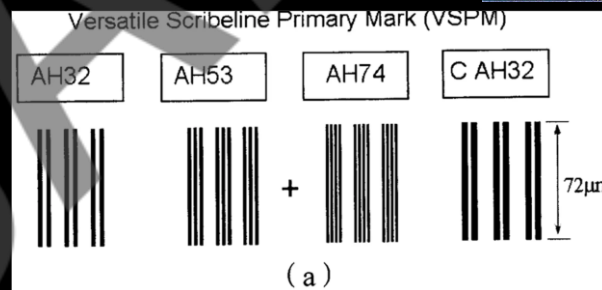
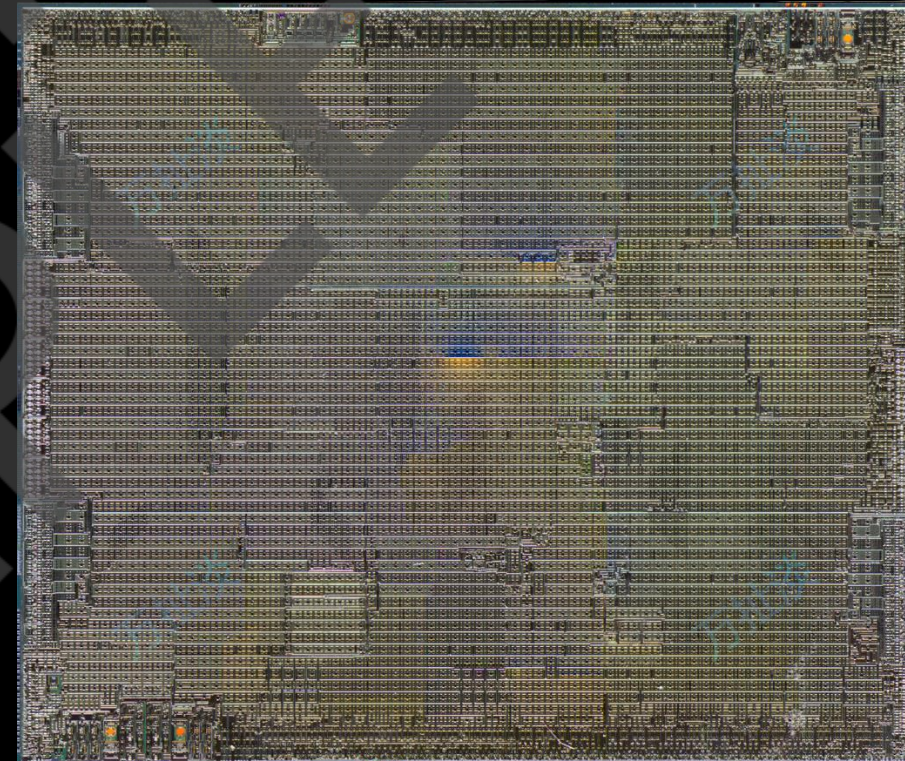


ASML VSPM AH74



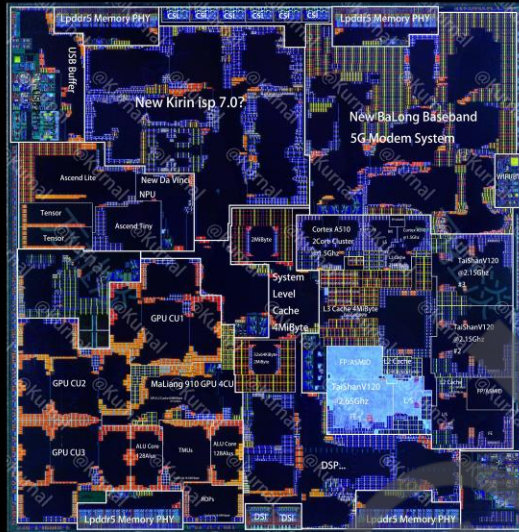
Alignment marker

By @Techinsights



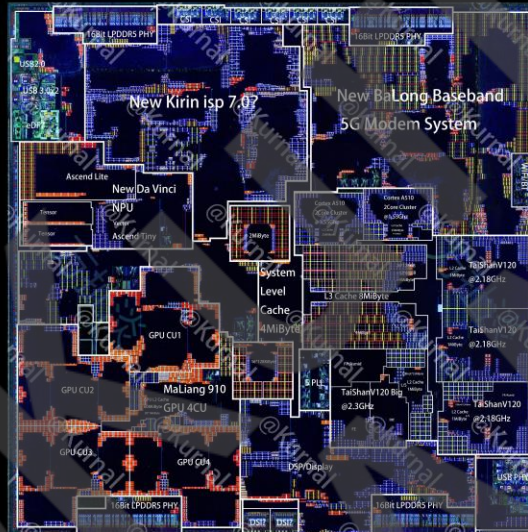
PACKAGE SHOT

Kirin 9000s (10.39x10.65)
| 10.65mm²



10mm

Kirin 9010 (10.74*10.88)
| 16.85mm²

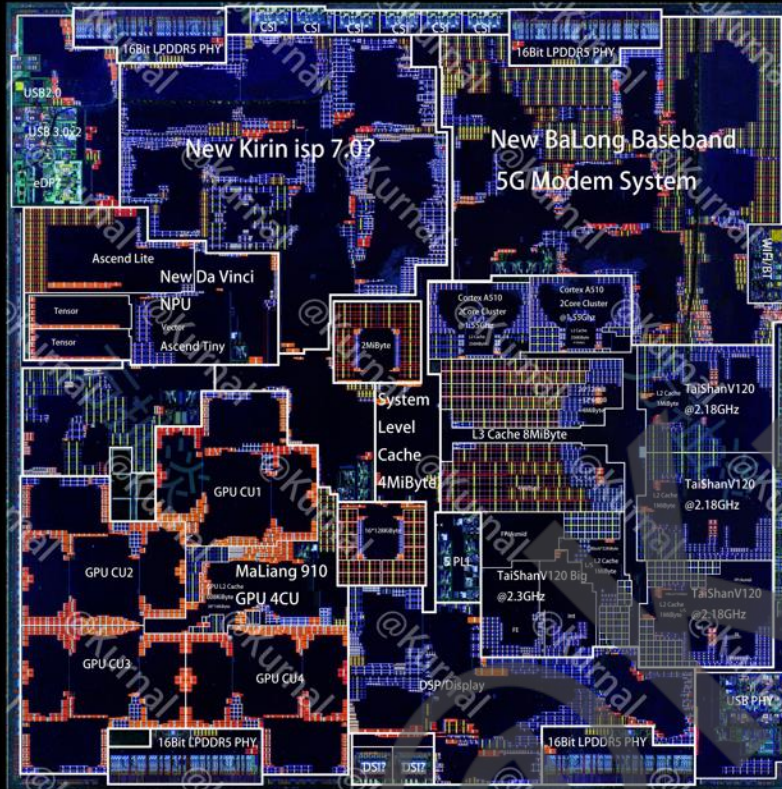


Kirin 9020 (12.89x10.9)
| 40.5mm²



Up 20.24% area

PACKAGE SHOT



Kirin 9010 | 16.85mm²



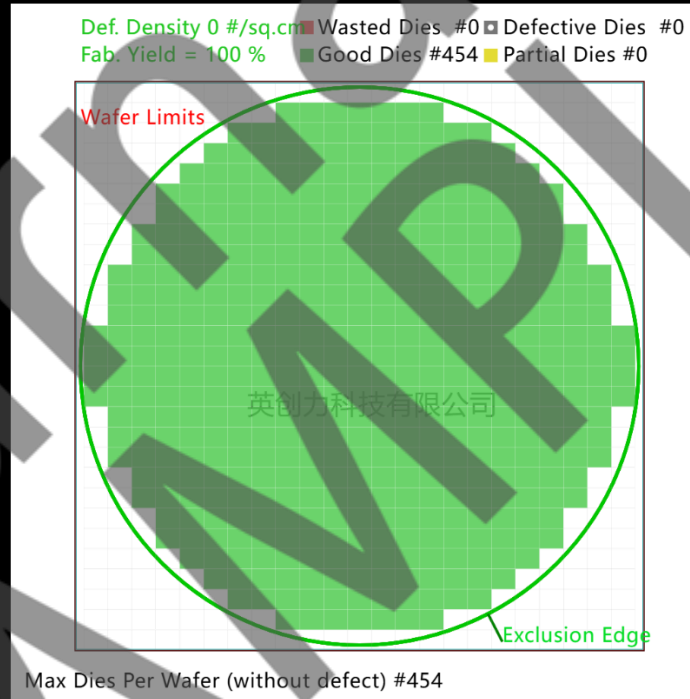
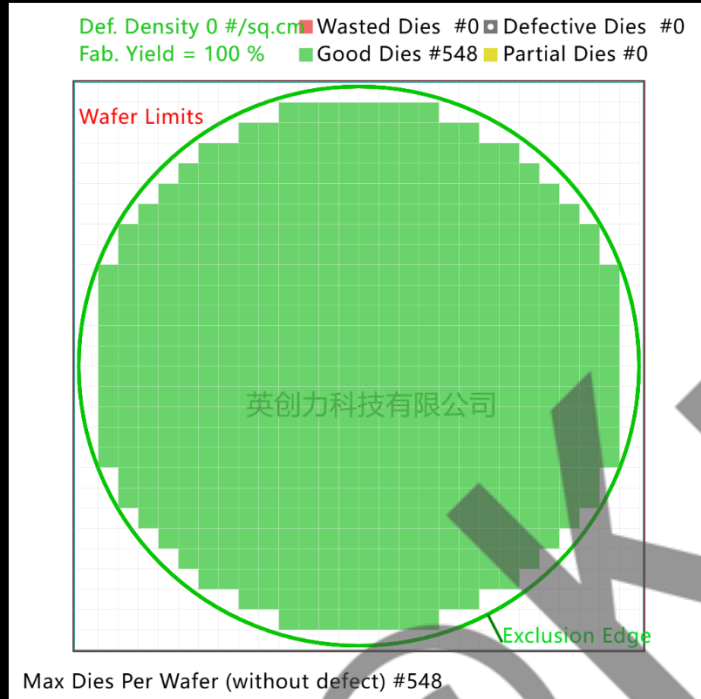
Kirin 9020 | 40.5mm²

Up 20.24% area

The process technology is the same as SMIC N+2

CHIP PRICE ESTIMATE

CHIP PRICE ESTIMATE



$$DPW = \frac{\pi \times \left(\frac{d}{2}\right)^2}{W \times H} - \frac{\pi \times d}{\sqrt{2} \times W \times H}$$

π: 圆周率
 d: 晶圆直径
 W: 晶粒长度
 H: 晶粒宽度

Because the die size is **larger**
 So that
 In DPW data
 Each wafer can cut die number
 decreased by **94/17.15%**

Kirin 9010
 Die per wafer nb :

116.85mm²
548Die

Kirin 9020
 Die per wafer nb :

140.5mm²
454Die

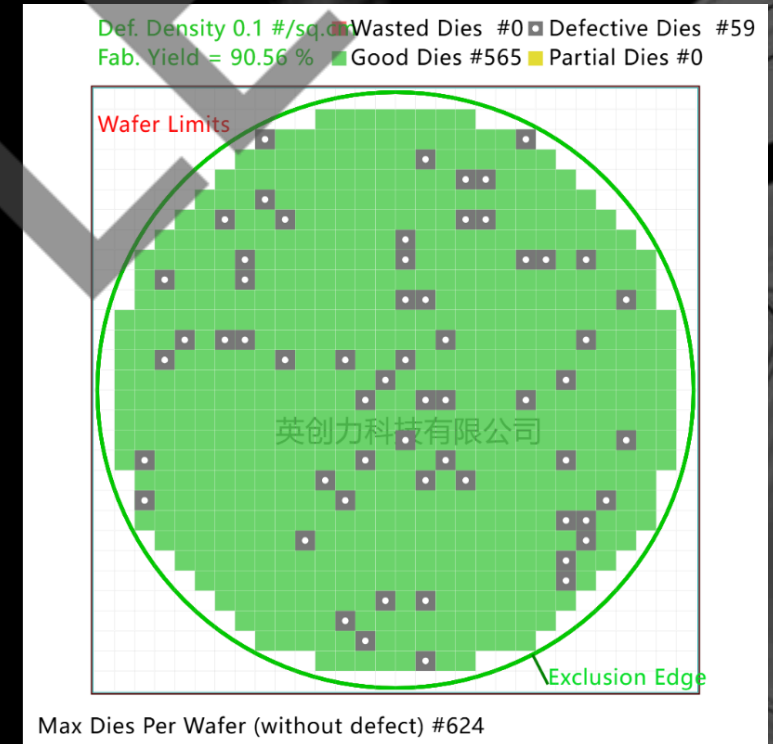
CHIP PRICE ESTIMATE

How to know the product price?

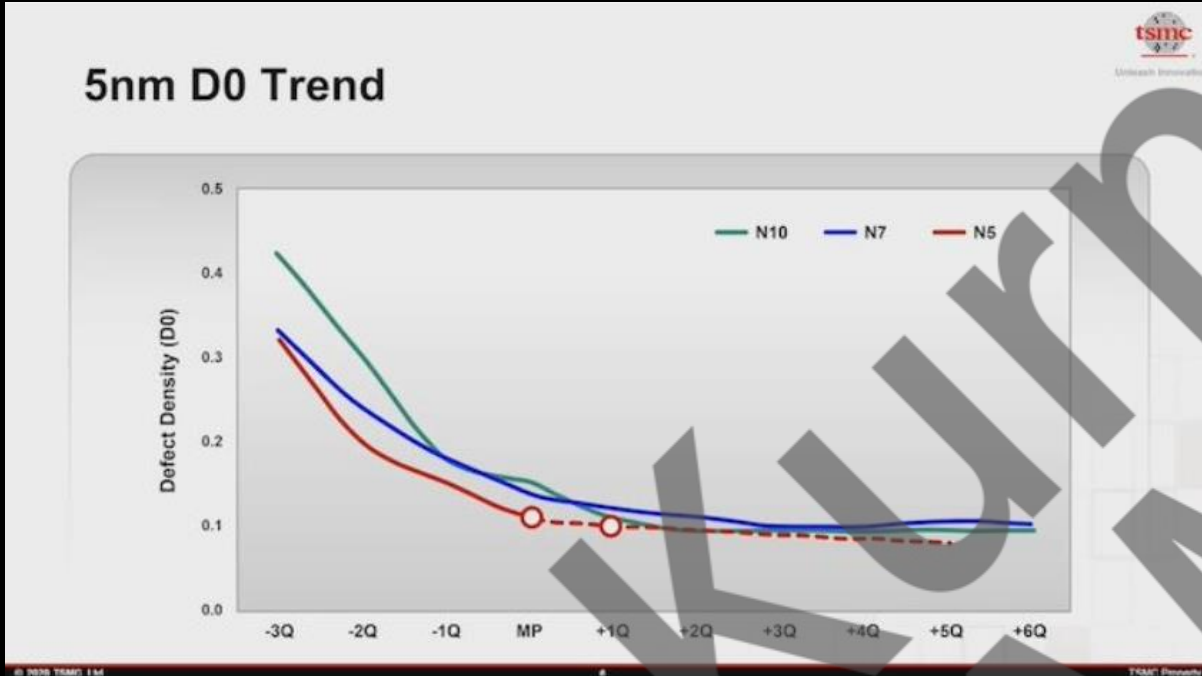
We need to know the **yield**

In this analyze data, we **just** compute **Manufacturing yield**
No Package yield

We need to know the **D0** (used Murphy's) and **Wafer Price**
Then we can know the **Price/die** and **yield**



CHIP PRICE ESTIMATE(D0)



Estimate Price and yield we can use **Level data**

In this photo shows that
When the process enters mass production
We can roughly estimate **D0=0.1**



High Yield: **D0=0.06**
Middle Yield: **D0=0.1**
Low Yield: **D0=0.15**
Ultra Low Yield: **D0=0.2**



CHIP PRICE ESTIMATE(WAFER PRICE)

SMIC N+2 Price/wafer is nearly TSMC N7 Price/wafer

Line	Node (nm)	90	65	40	28	20	16/12	10	7	5
1	Mass production year and quarter ²²⁰	2004 Q4	2006 Q4	2009 Q1	2011 Q4	2014 Q3	2015 Q3	2017 Q2	2018 Q3	2020 Q1
2	Capital investment per wafer processed per year	\$4,649	\$5,456	\$6,404	\$8,144	\$10,356	\$11,220	\$13,169	\$14,267	\$16,746
3	Net capital depreciation at start of 2020 (25.29% / year)	65%	65%	65%	65%	65%	65%	55.1%	35.4%	0.0%
4	Undepreciated capital per wafer processed per year (remaining value at start of 2020)	\$1,627	\$1,910	\$2,241	\$2,850	\$3,625	\$3,927	\$5,907	\$9,213	\$16,746
5	Capital consumed per wafer processed in 2020	\$411	\$483	\$567	\$721	\$917	\$993	\$1,494	\$2,330	\$4,235
6	Other costs and markup per wafer	\$1,293	\$1,454	\$1,707	\$2,171	\$2,760	\$2,990	\$4,498	\$7,016	\$12,753
7	Foundry sale price per wafer	\$1,650	\$1,937	\$2,274	\$2,891	\$3,677	\$3,984	\$5,992	\$9,346	\$16,988
8	Foundry sale price per chip	\$2,433	\$1,428	\$713	\$453	\$399	\$331	\$274	\$233	\$238

N7 Price by 2020 was nearly 10000\$/Wafer



CHIP PRICE ESTIMATE)

Yield data

Yield	High yield	Middle yield	Low yield	Ultra low yield
9020	92.10%	87.24%	81.58%	76.34%
9010	93.39%	89.26%	84.40%	79.85%
9000S	93.73%	89.80%	85.16%	80.80%

Die Price in 7000\$ per Wafer

Price	High yield	Middle yield	Low yield	Ultra low yield
9020	16.741	17.673	18.9	20.2
9010	13.685	14.314	15.129	16
9000S	12.0702	13.258	13.979	14.737

Die per Wafer nb

	DPW
9020	454
9010	548
9000S	588

Good Die per Wafer data

GDPW	High yield	Middle yield	Low yield	Ultra low yield
9020	418.13	396.07	370.37	346.58
9010	511.78	489.14	462.51	437.58
9000S	551.13	528.02	500.74	475.1

Die Price in 10000\$ per Wafer

Price	High yield	Middle yield	Low yield	Ultra low yield
9020	23.916	25.248	27	28.855
9010	19.54	20.44	21.621	22.853
9000S	18.145	18.938	19.971	21.052

High Yield: **D0=0.06**
 Middle Yield: **D0=0.1**
 Low Yield: **D0=0.15**
 Ultra Low Yield: **D0=0.2**

Die Price in 13000\$ per Wafer

Price	High yield	Middle yield	Low yield	Ultra low yield
9020	31.09	32.823	35.1	37.512
9010	25.402	26.578	28.107	29.709
9000S	23.589	24.62	25.962	27.368

CPU ANALYZE

CPU ANALYZE

In this Kirin 9020 Chip
It was **3 CPU Cluster**
Design **1+3+4**

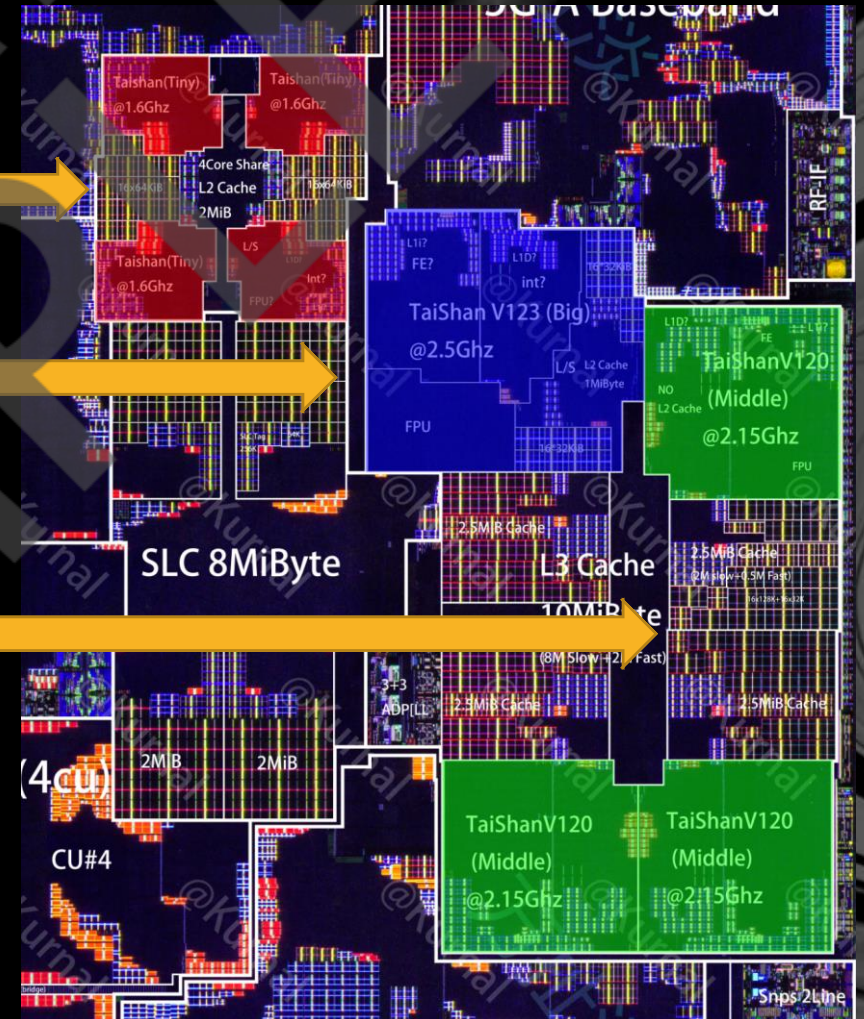
- 1 Biggest (Have SMT)
- 3 Middle (Have SMT)
- 4 Little/tiny (No SMT)

All Taishan Core, No ARM IP

(NEW) Taishan Tiny

(NEW) Biggest Taishan

Middle Taishan V120



CPU ANALYZE-ALL

CPU Cluster Only



Kirin9000s CPU Cluster

Kirin9010 CPU Cluster

Kirin9020 CPU Cluster

They CPU all 1 Biggest 3 Middle 4 Little/tiny

CPU ANALYZE-ALL

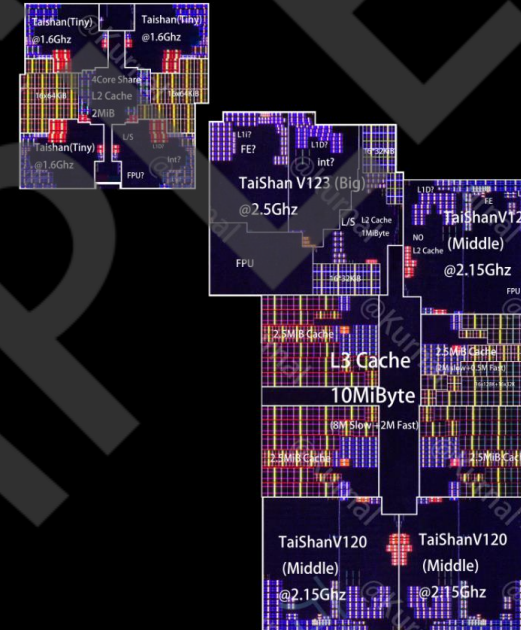
CPU Cluster Only



Kirin9000s CPU Cluster



Kirin9010 CPU Cluster



Kirin9020 CPU Cluster

CPU ANALYZE-ALL

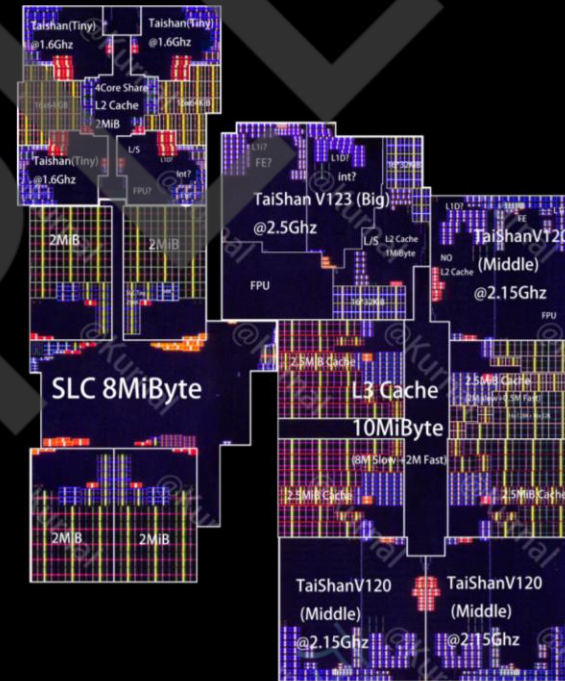
CPU Cluster+SLC



Kirin9000s CPU Cluster



Kirin9010 CPU Cluster

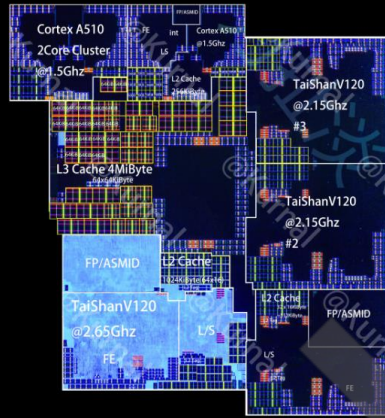


Kirin9020 CPU Cluster

Similar

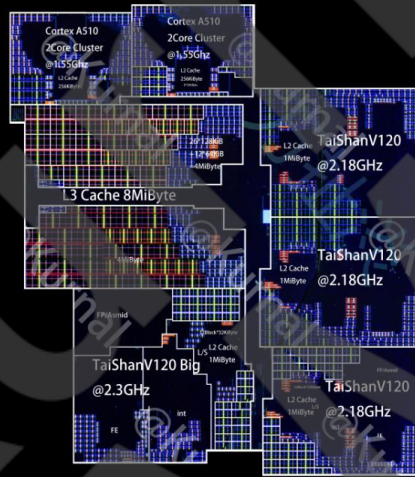
Next Gen

CPU ANALYZE-ALL



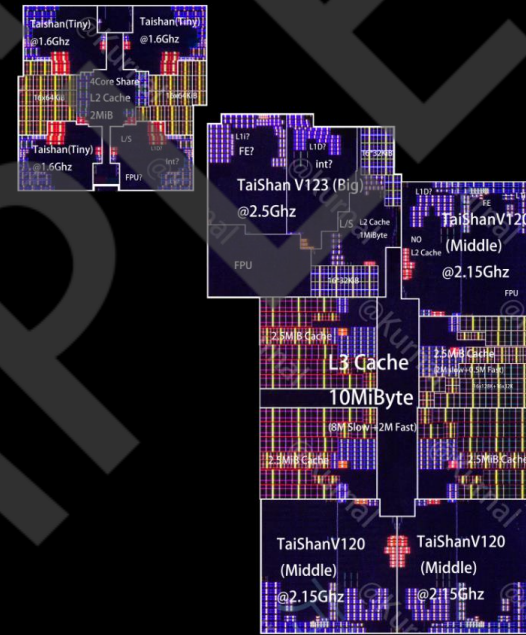
Kirin9000s CPU Cluster

Size: 17.113



Kirin9010 CPU Cluster

Size: 21.199



Kirin9020 CPU Cluster

Size: 18.5+4.345=22.845

Up 23.88%

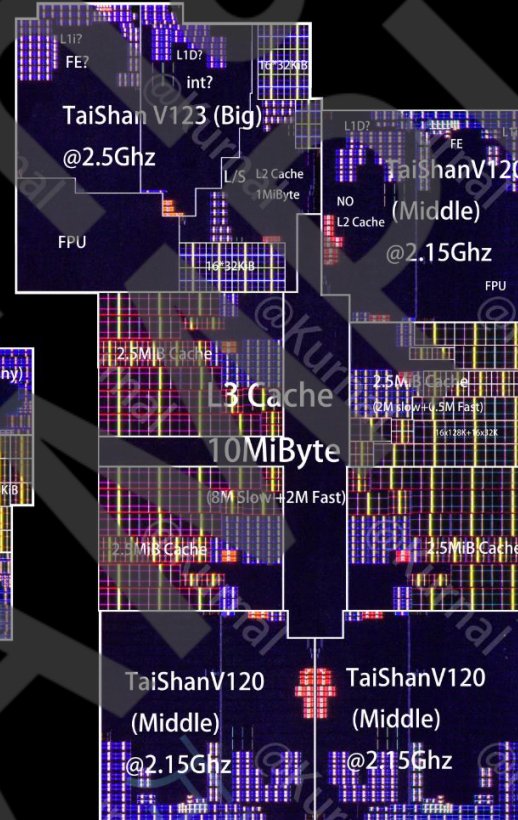
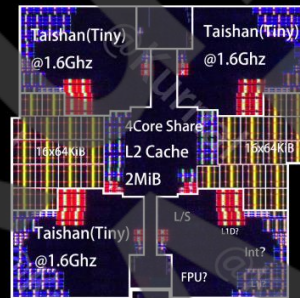
Up 7.76%

CPU ANALYZE-ALL

This time the CPU cluster is divided into a **dual** cluster
Named **Big Cluster / Little Cluster**

Big Cluster area: **18.5**
Little Cluster area: **4.345**

Little Cluster



Big Cluster



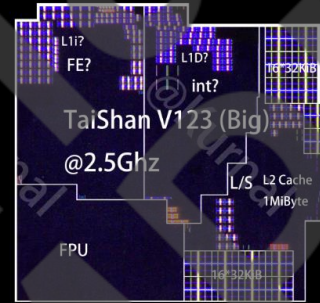
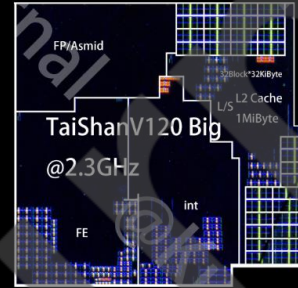
CPU ANALYZE-ALL

Kirin9000s

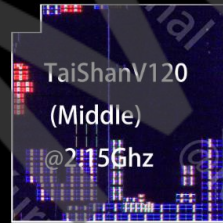
Kirin9010

Kirin9020

Big Core



Middle Core

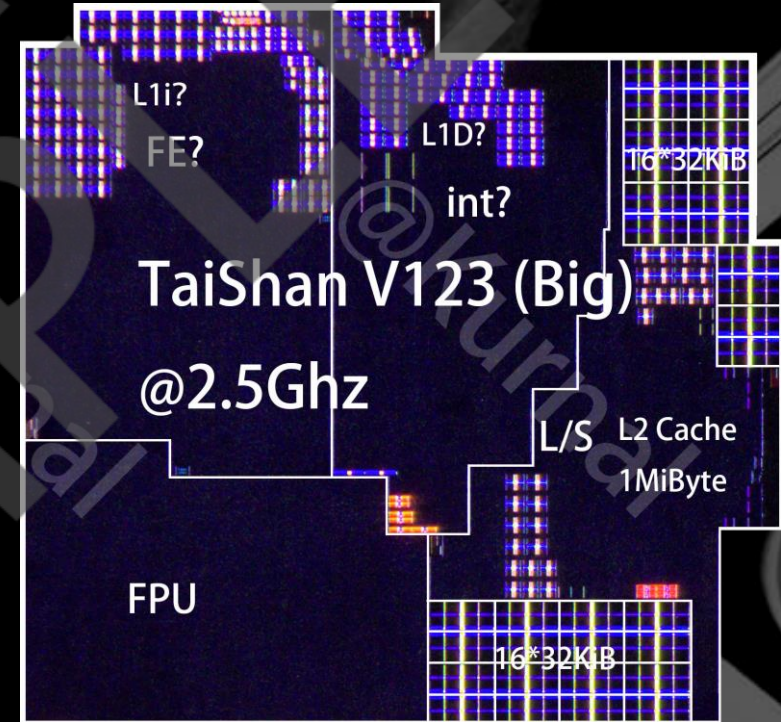
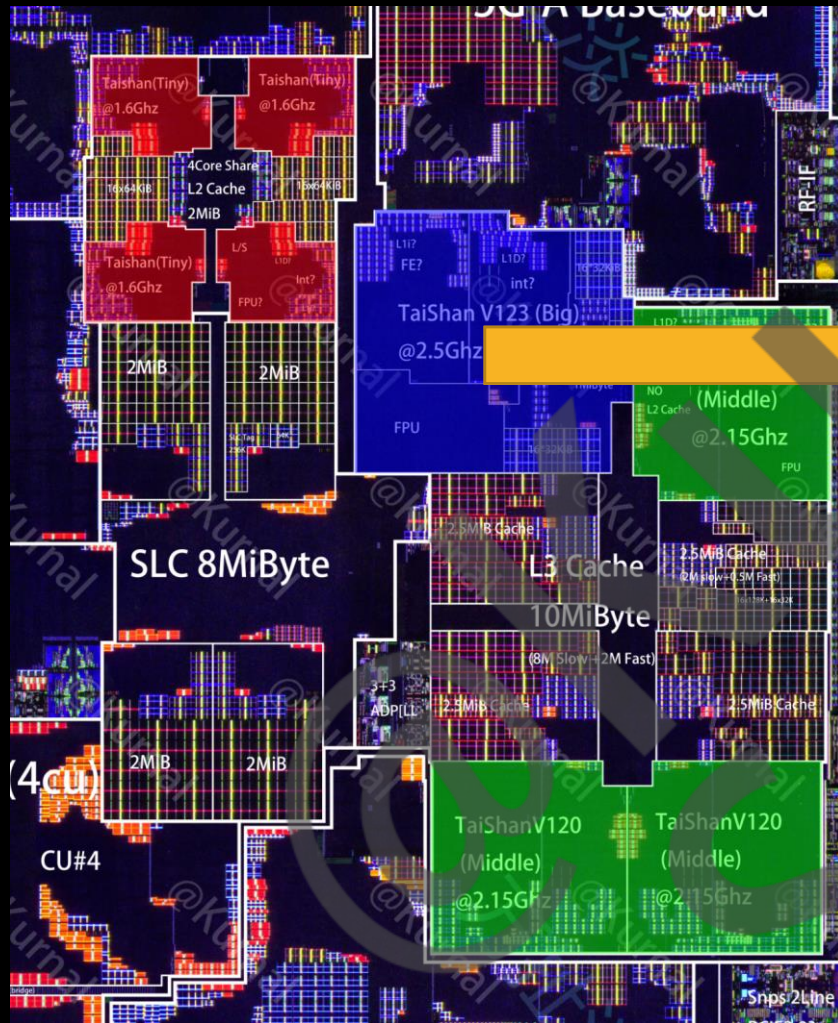


Tiny Core



CPU-BIGGEST CORE

CPU ANALYZE-BIGGEST



New Biggest CPU Core

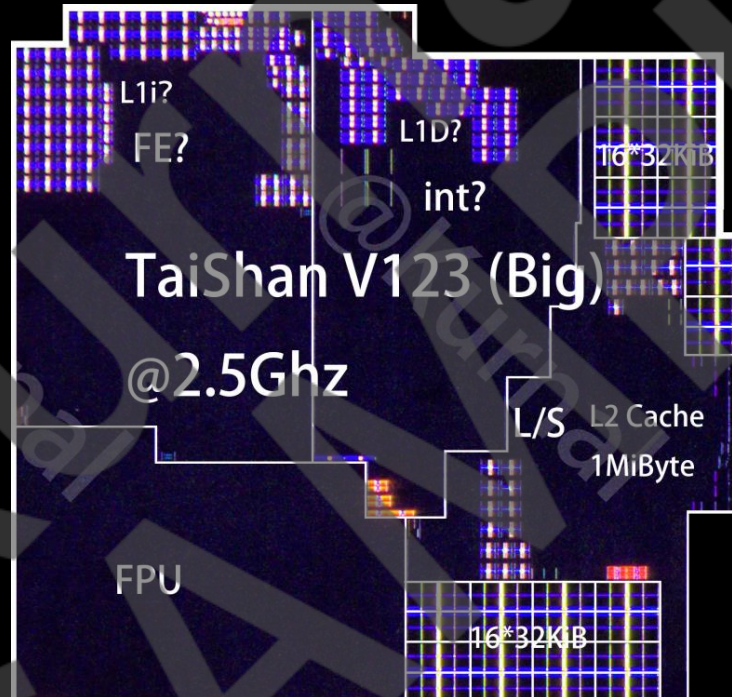
CPU ANALYZE-BIGGEST-AREA

New Biggest CPU Core

Taishan V123
@2.5Ghz

L2 Cache 1MiB

Size: 2.349mm²



CPU ANALYZE-BIGGEST-AREA

Kirin 9000s

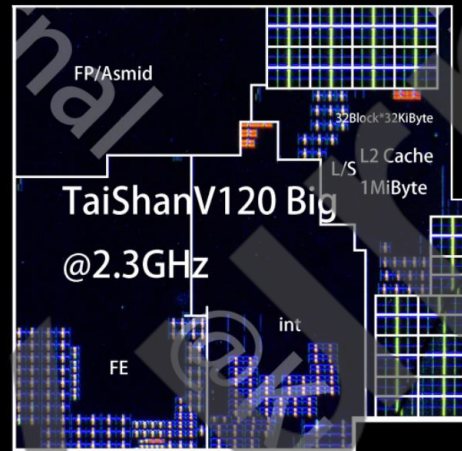


Taishan V120
@2.65Ghz

L2 Cache 1MiB

Size: 3.357mm²

Kirin 9010

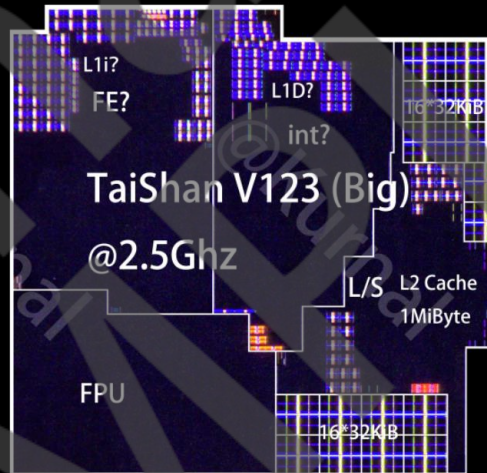


Taishan V121
@2.3Ghz

L2 Cache 1MiB

Size: 4.034mm²

Kirin 9020



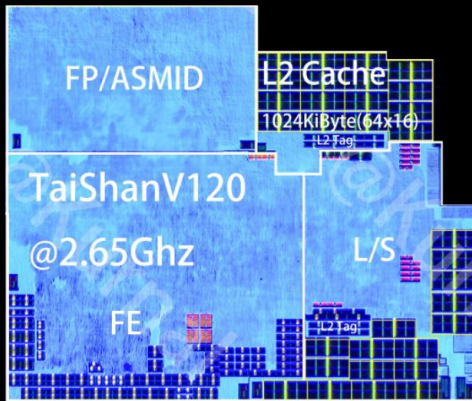
Taishan V123
@2.5Ghz

L2 Cache 1MiB

Size: 4.315mm²

CPU ANALYZE-BIGGEST-AREA

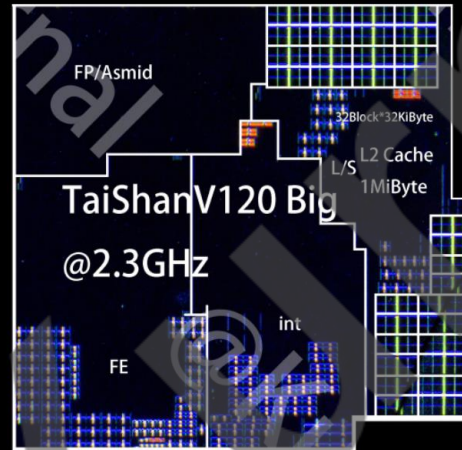
Kirin 9000s



Taishan V120
@2.65Ghz

If No L2 Cache
Size: 2.791mm²

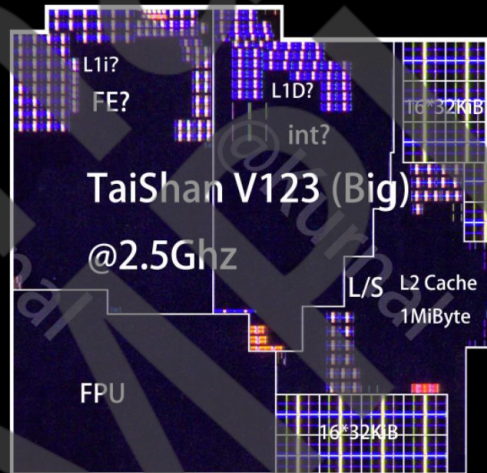
Kirin 9010



Taishan V121
@2.3Ghz

If NoL2 Cache
Size: 3.48mm²

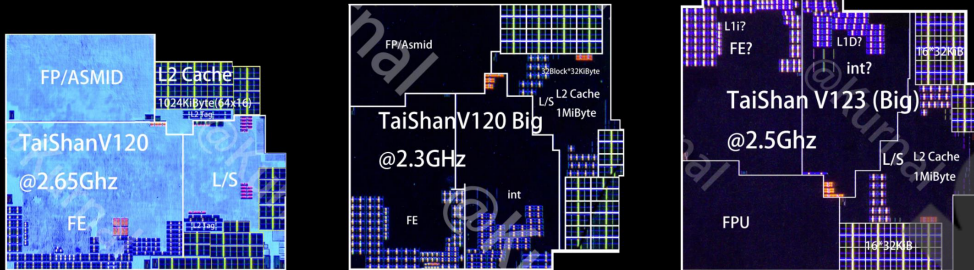
Kirin 9020



Taishan V123
@2.5Ghz

If NoL2 Cache
Size: 3.743mm²

CPU ANALYZE-BIGGEST-AREA



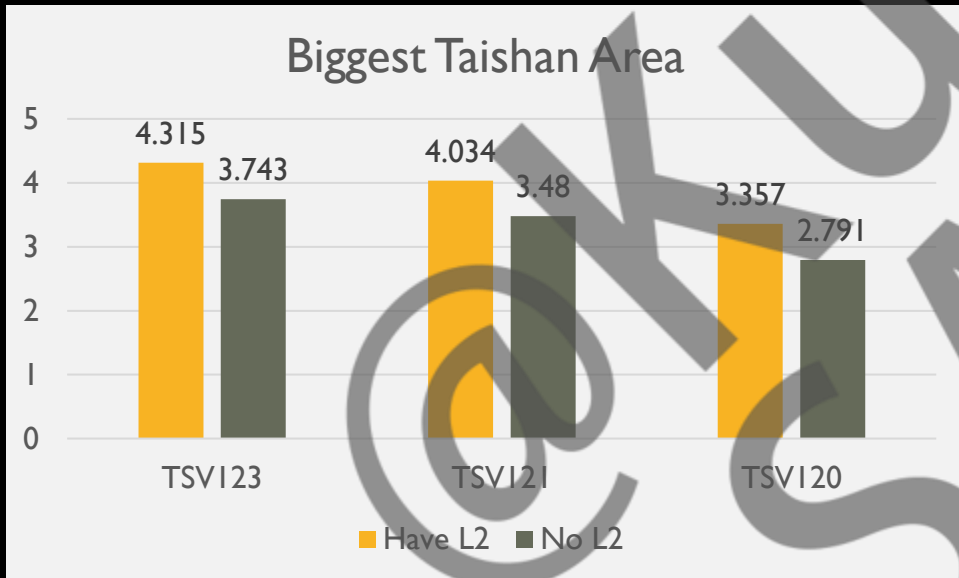
	TSV123	TSV121	TSV120
Have L2	4.315	4.034	3.357
No L2	3.743	3.48	2.791
Only L2	0.572	0.554	0.566

New Taishan Core have the same L2 Cache size and area compare previous generations

TSV 123 vs TSV121

The area is larger by **6.96%**

No L2 area is larger by **7.56%**



CPU ANALYZE-BIGGEST-PERF

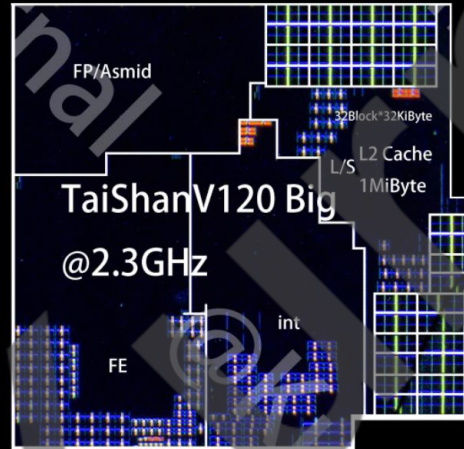
Kirin 9000s



Taishan V120
@2.65Ghz

GB6 Perf: 1299
IPC: 490.2/Ghz

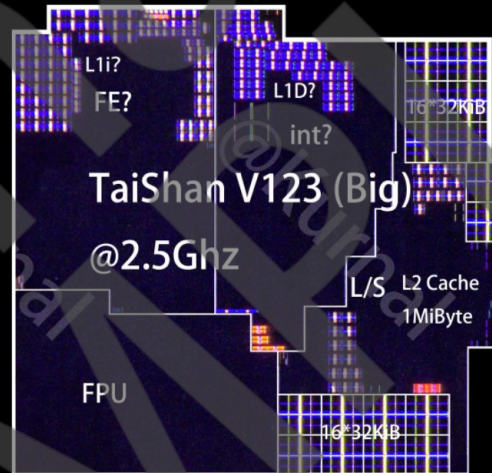
Kirin 9010



Taishan V121
@2.3Ghz

GB6 Perf: 1458
IPC: 633.9/Ghz

Kirin 9020



Taishan V123
@2.5Ghz

GB6 Perf: 1616
IPC: 646.4/Ghz

CPU ANALYZE-BIGGEST-PERF

Taishan V120
@2.65Ghz

GB6 Perf: 1299
IPC: 490.2/Ghz

Taishan V121
@2.3Ghz

GB6 Perf: 1458
IPC: 633.9/Ghz

Taishan V123
@2.5Ghz

GB6 Perf: 1616
IPC: 646.4/Ghz

Up Perf: 12.24%
Up ipc: 29.31%

Up Perf: 10.83%
Up ipc: 1.97%

Up Perf: 24.4%
Up ipc: 31.86%

CPU ANALYZE-BIGGEST-PERF

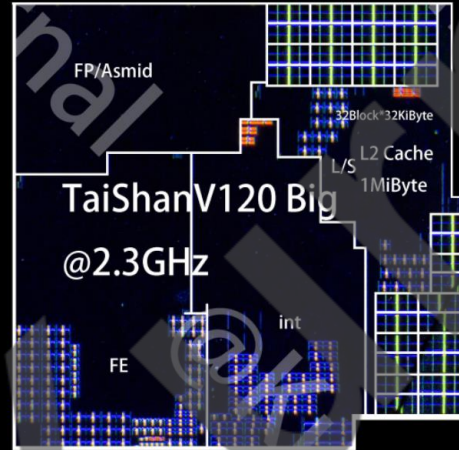
Kirin 9000s



Taishan V120
@2.65Ghz

Size/IPC: 146.023/mm²

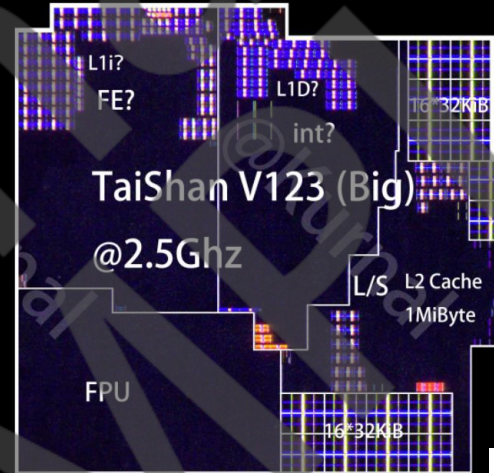
Kirin 9010



Taishan V121
@2.3Ghz

Size/IPC: 157.139/mm²

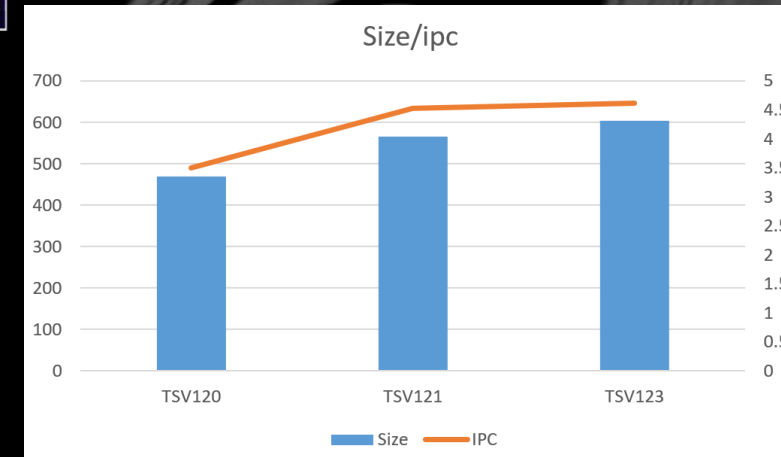
Kirin 9020



Taishan V123
@2.5Ghz

Size/IPC: 149.8/mm²

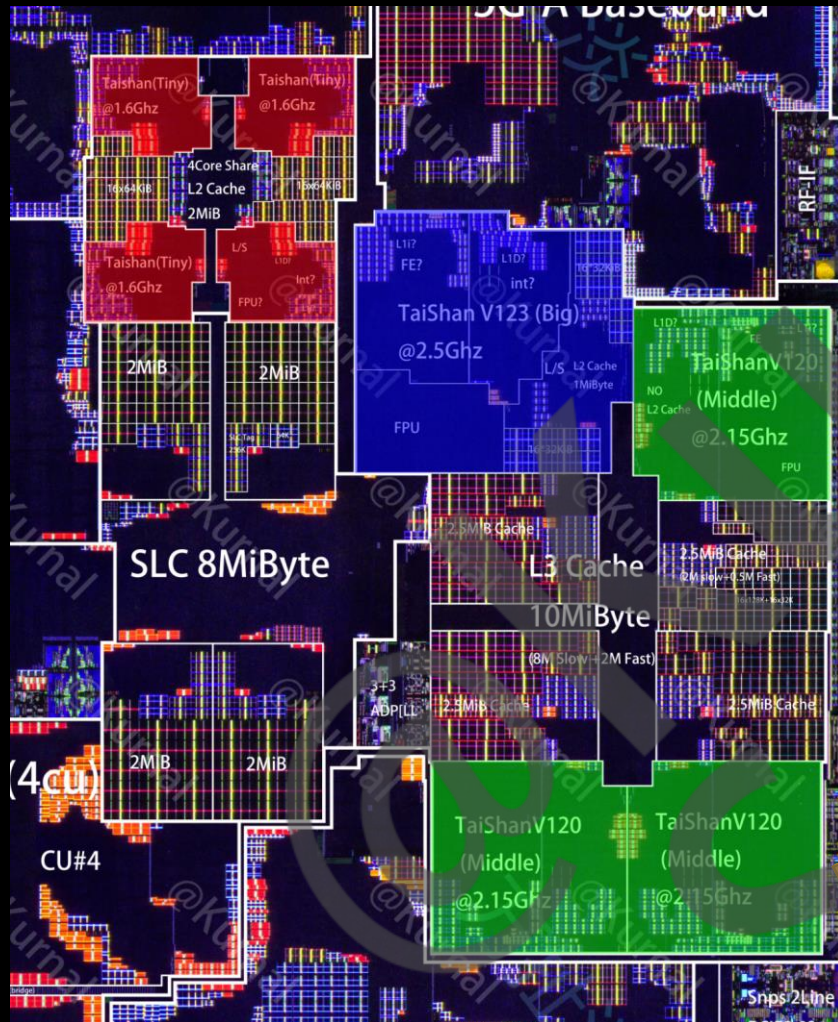
	Size	IPC	Size/IPC
TSV120	3.357	490.2	146.023
TSV121	4.034	633.9	157.139
TSV123	4.315	646.4	149.8





CPU-MIDDLE CORE

CPU ANALYZE-MIDDLE



Middle CPU Core

CPU ANALYZE-MIDDLE

New Middle CPU Core

Taishan V120 Middle
@2.15Ghz

No L2 Cache

Size: 2.33mm²



CPU ANALYZE-MIDDLE

Kirin 9000s



Taishan V120
@2.15Ghz

L2 Cache 512KiB

Size: 2.259mm²

Kirin 9010



Taishan V120
@2.18Ghz

L2 Cache 1MiB

Size: 2.652mm²

Kirin 9020



Taishan **New** V120
@2.15Ghz

No L2 Cache

Size: 2.33mm²

CPU ANALYZE-MIDDLE

Kirin 9000s



Taishan V120
@2.15Ghz

If No L2 Cache
Size: 1.971mm²

Kirin 9010



Taishan V120
@2.18Ghz

If No L2 Cache
Size: 2.063mm²

Kirin 9020



Taishan **New** V120
@2.15Ghz

No L2 Cache
Size: 2.33mm²

CPU ANALYZE-BIGGEST-AREA

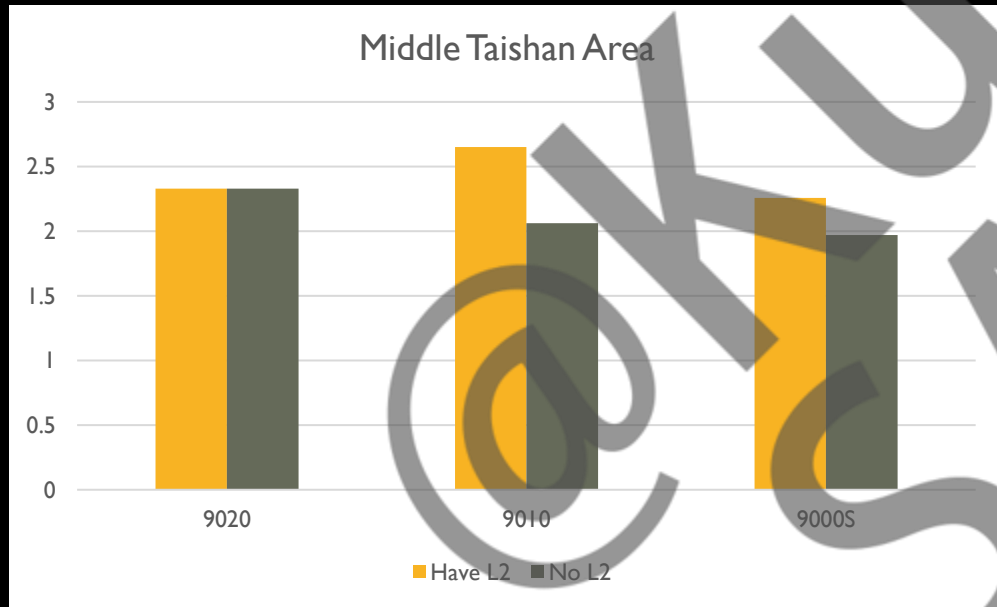


	9020	9010	9000S
Middle	9020	9010	9000S
Have L2	2.33	2.652	2.259
No L2	2.33	2.063	1.971
Only L2	0	0.589	0.288

In 9020 Middle Core
They have **not L2 Cache**

So the area is smaller **12.14%/0.322mm²**

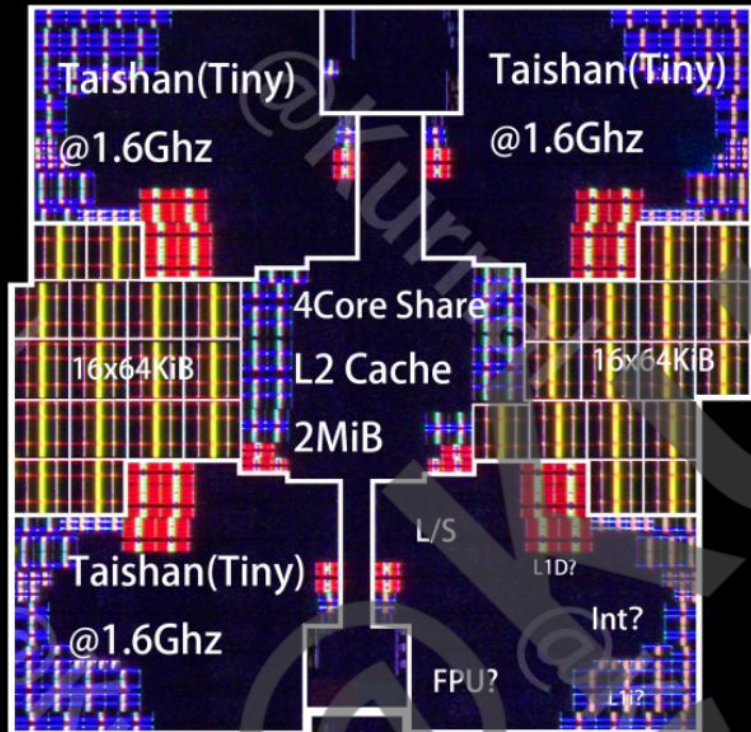
But when the data in NoL2 Size
The area is bigger **12.94%/0.267mm²**





CPU-TINY CORE

CPU-TINY CORE



Taishan Tiny
@1.6Ghz

Taishan Tiny
@1.6Ghz

Taishan Tiny
@1.6Ghz

Taishan Tiny
@1.6Ghz

share L2 Cache 2MiB

In Little Cluster
4 Tiny Core shared the L2 Cache 2MiB

CPU-TINY CORE

New Tiny CPU Core

Taishan Tiny Core
@1.6Ghz

4 Core Share L2 2MiB

Size: 0.652mm²



CPU-TINY CORE

Kirin 9000s



Arm A510x2
@1.5Ghz

2Core Shard 256KiB L2
Core Per L2: 128KiB

Size/Core: 0.689 mm²

Kirin 9010

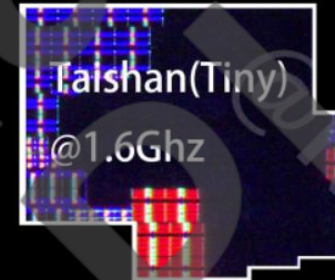


Arm A510x2
@1.55Ghz

2Core Shard 256KiB L2
Core Per L2: 128KiB

Size/Core: 0.698 mm²

Kirin 9020



Taishan **New** Tiny
@1.6Ghz

4Core Share 2MiB L2
Core Per L2: 512KiB

Size/Core: 0.652mm²
Cluster/Core: 1.086mm²

CPU-TINY CORE

Kirin 9000s



Arm A510x2
@1.5Ghz

If No L2 Cache
Size/Core: 0.591mm²

But A510x2 share the FP unit

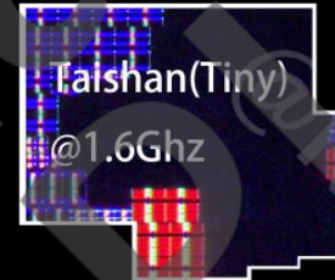
Kirin 9010



Arm A510x2
@1.55Ghz

If No L2 Cache
Size/Core: 0.605 mm²

Kirin 9020

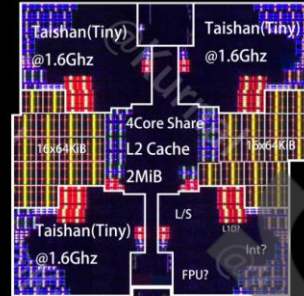


Taishan **New** Tiny
@1.6Ghz

If No L2 Cache
Size/Core: 0.652mm²
Cluster/Core: 0.855mm²

Exclusive the FP unit

CPU-TINY CORE



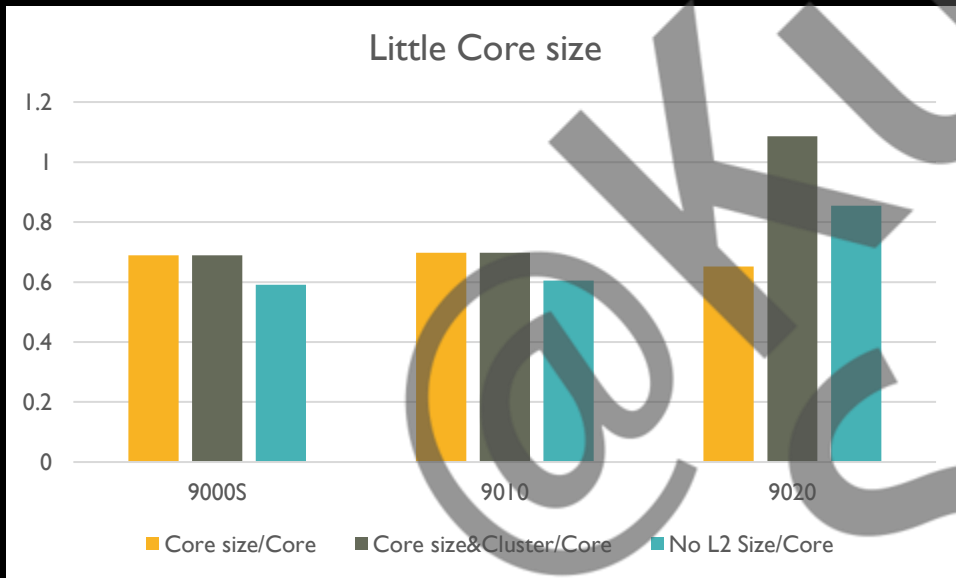
	9000S	9010	9020
Core size/Core	0.689	0.698	0.652
Core size & Cluster/Core	0.689	0.698	1.086
No L2 Size/Core	0.591	0.605	0.855

In 9020Tiny Cluster
Use the New Taishan IP , Taishan Tiny
Per Core L2 Cache is 512K
4 times than A510 (128K Per Core)

Area per : Cluster/Core
TaishanTiny is bigger Arm A510
No L2 Data is bigger ArmA510

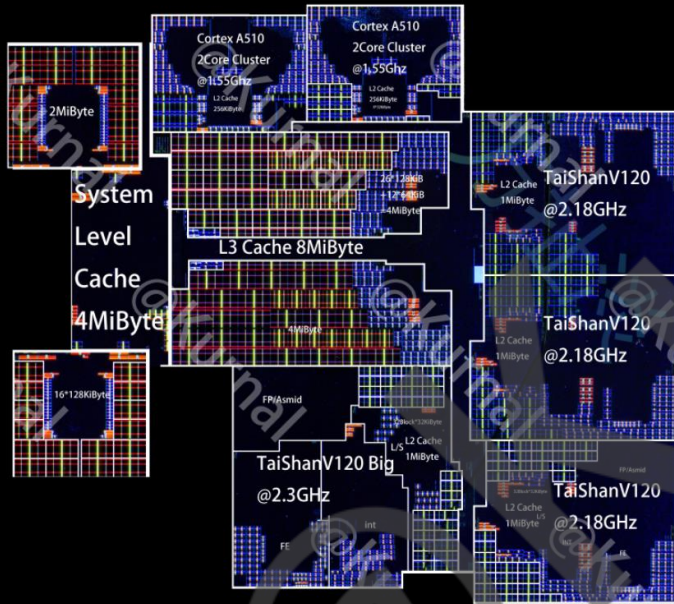
55.59 %

41.32%

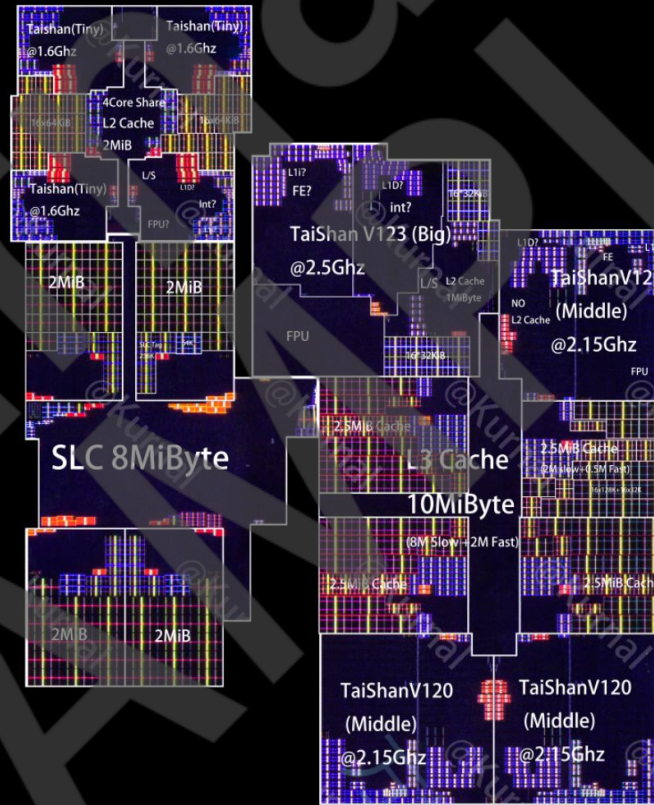


CPU ANALYZE VS

CPU ANALYZE VS-LAYOUT

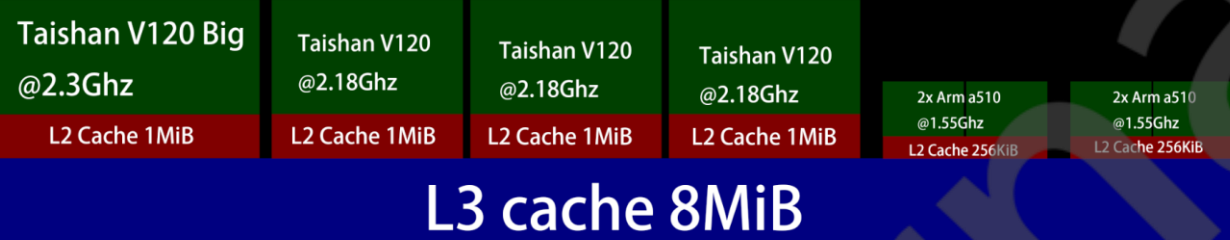


Kirin 9010



Kirin 9020

CPU ANALYZE VS-LAYOUT



Kirin 9010

System Level Cache 4MiB



Kirin 9020

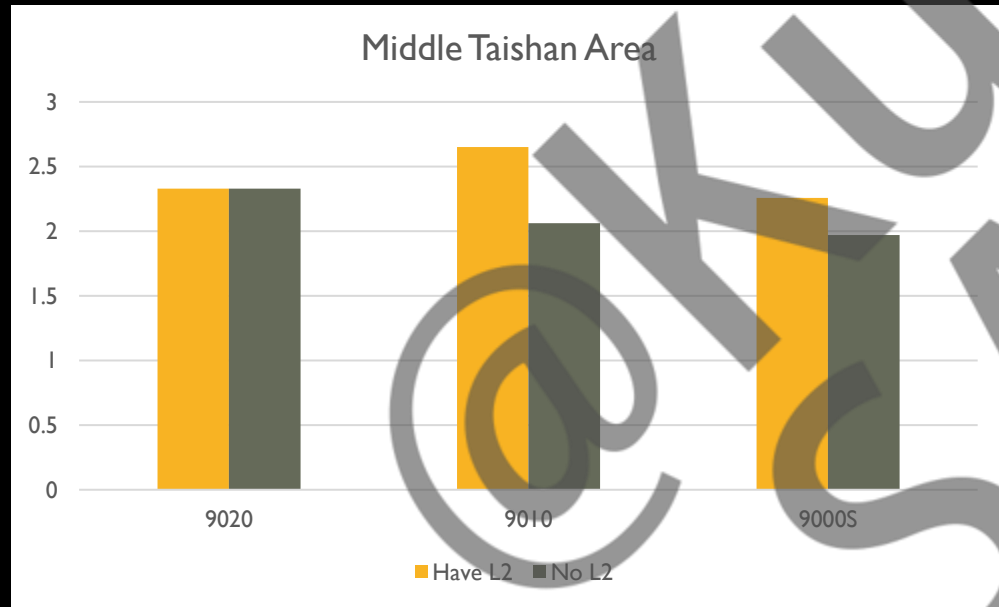
System Level Cache 8MiB

L3 CACHE

L3 CACHE



Middle	9020	9010	9000S
Have L2	2.33	2.652	2.259
No L2	2.33	2.063	1.971
Only L2	0	0.589	0.288



In 9020 Middle Core
They have **not L2 Cache**

So **why** doesn't the middle core use L2 Cache?

L3 CACHE

Taishan V120 Big @2.3Ghz	Taishan V120 @2.18Ghz	Taishan V120 @2.18Ghz	Taishan V120 @2.18Ghz	2x Arm a510 @1.55Ghz	2x Arm a510 @1.55Ghz
L2 Cache 1MiB	L2 Cache 1MiB	L2 Cache 1MiB	L2 Cache 1MiB	L2 Cache 256KiB	L2 Cache 256KiB

L3 cache 8MiB

System Level Cache 4MiB

$$L3+L2=12.5MiB$$

Kirin 9010

8 Core per 8MiB L3

Taishan V123 @2.5Ghz	Taishan V120 @2.15Ghz	Taishan V120 @2.15Ghz	Taishan V120 @2.15Ghz	Taishan Tiny @1.6Ghz	Taishan Tiny @1.6Ghz	Taishan Tiny @1.6Ghz	Taishan Tiny @1.6Ghz
L2 Cache 1MiB							

L3 cache 10MiB

share L2 Cache 2MiB

System Level Cache 8MiB

Just More than 0.5MiB

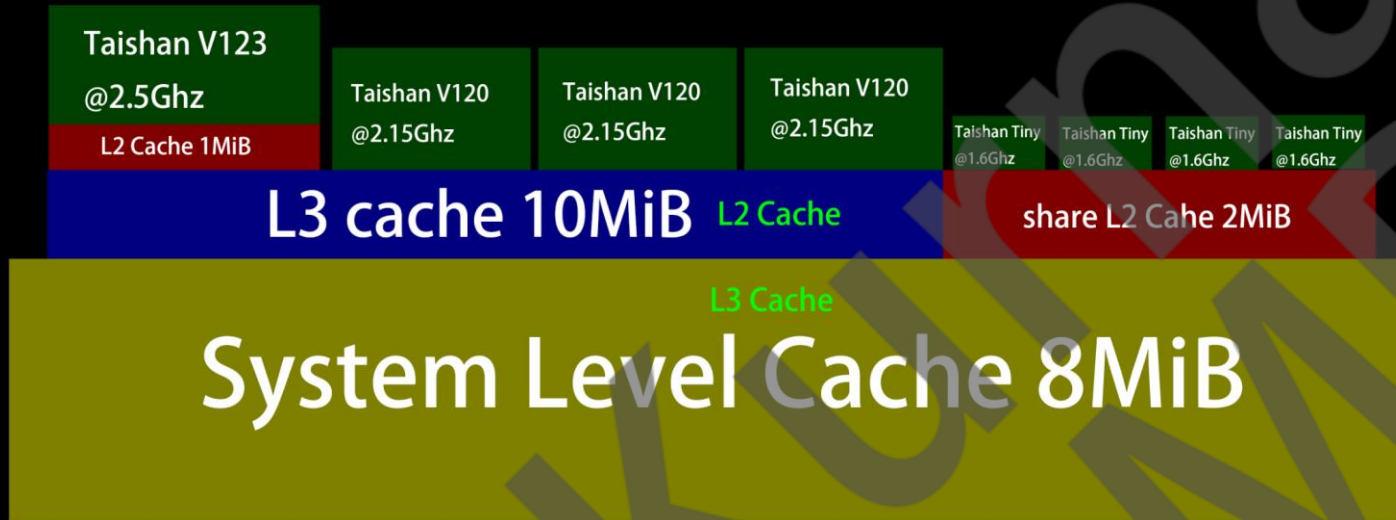
$$L3+L2=13MiB$$

Kirin 9020

4Core per 10MiB L3

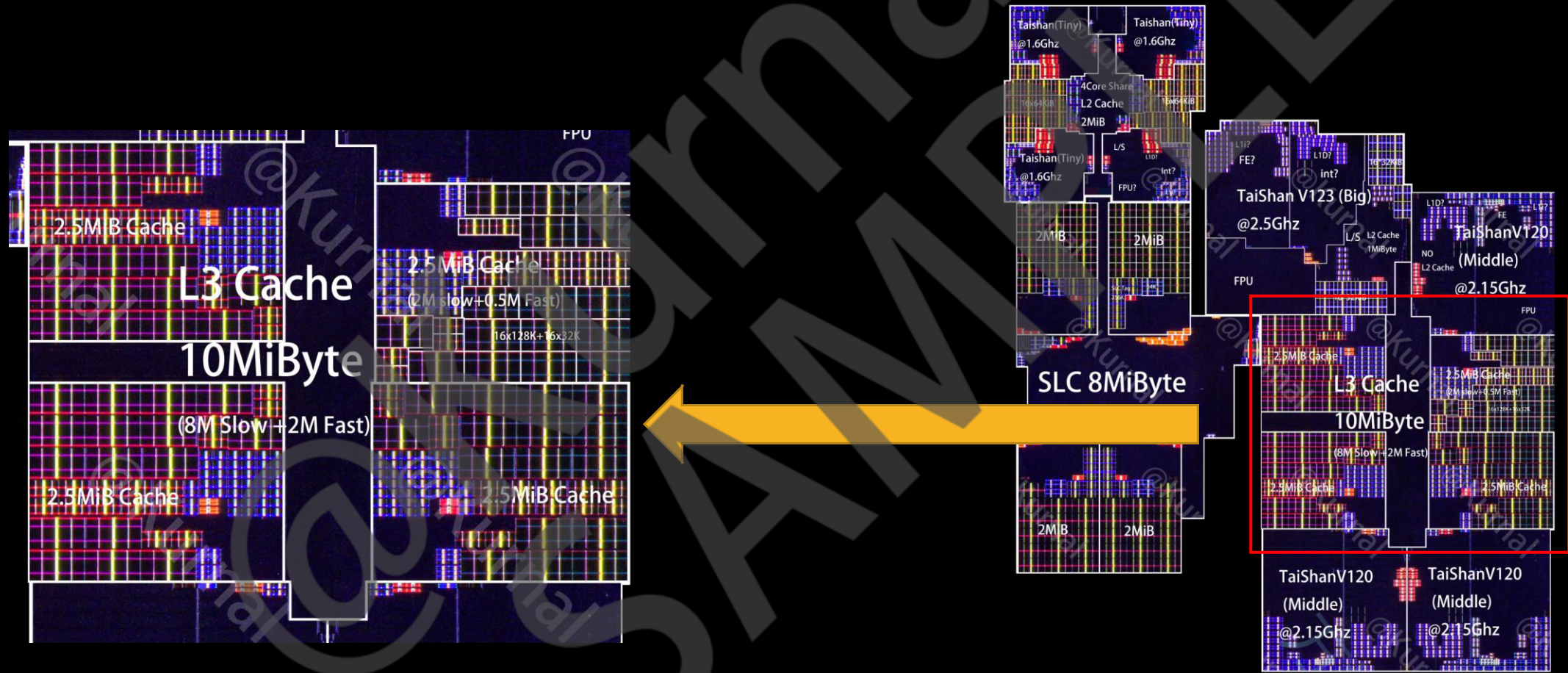
4Core per 2MiB L2

L3 CACHE

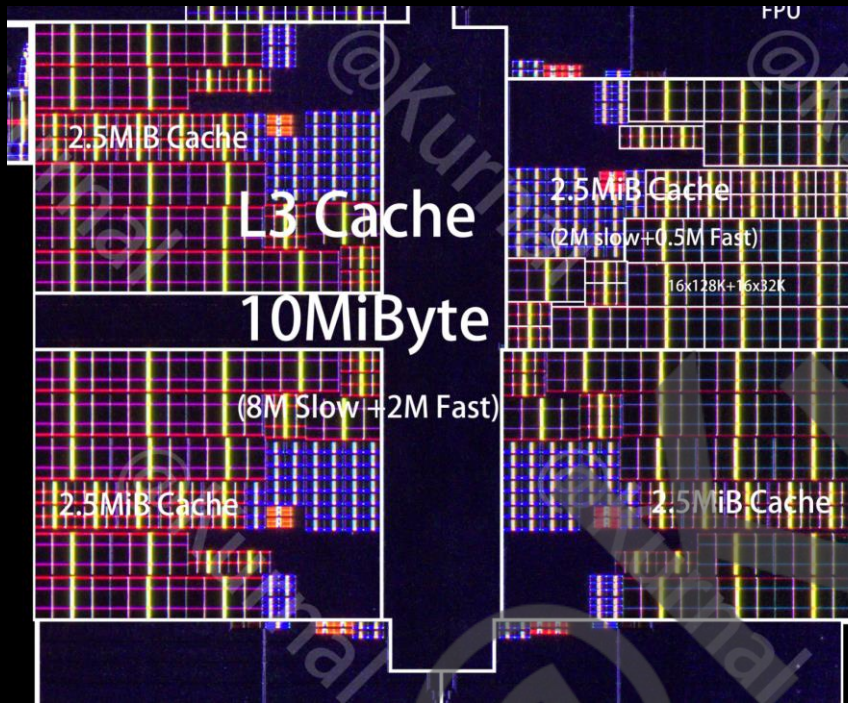


Middle Core maybe use the Level 3 Cache like one-time L2
Because the High Perf Compute need **cache consistency**

L3 CACHE



L3 CACHE

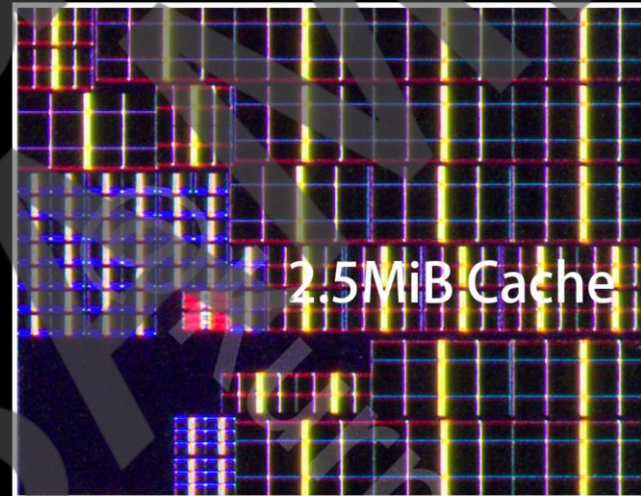


In Big cluster L3 Cache

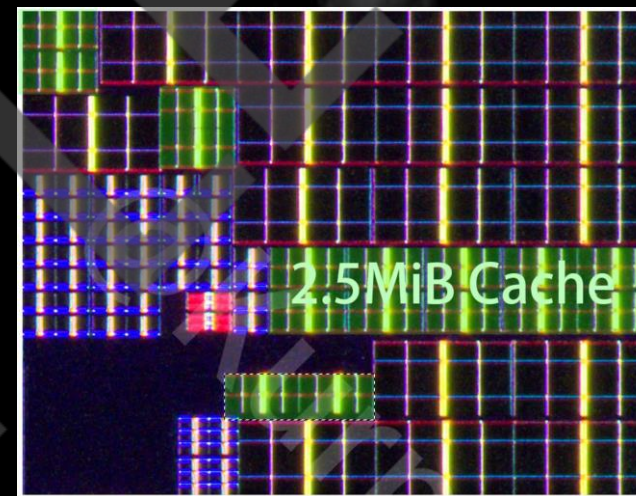
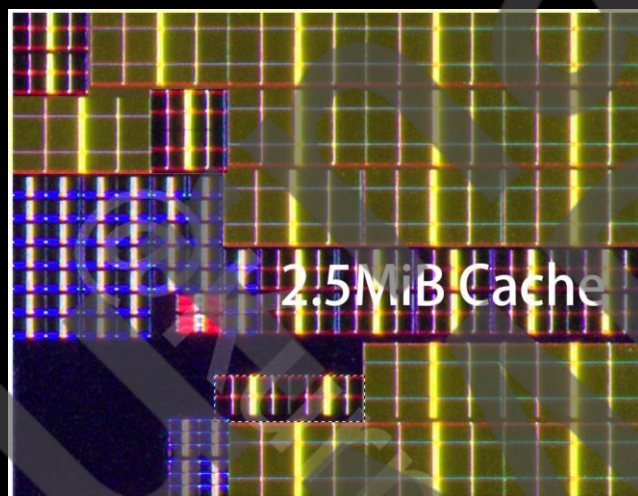
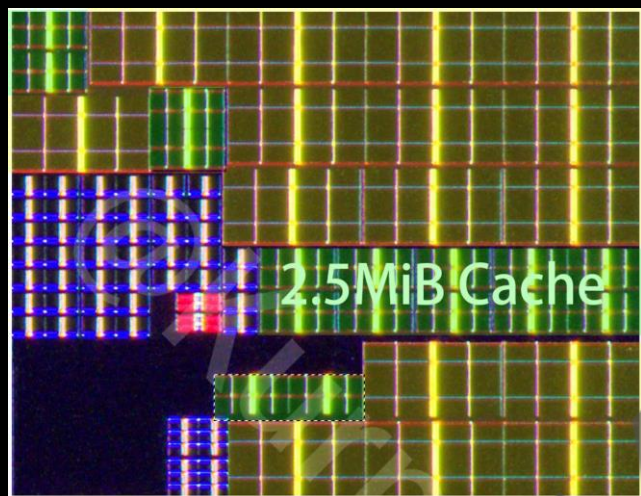
L3 Cache divided into 4 clusters
Each Cluster have 2.5MiB Cache

4 Cluster L3 Cache area **5.644**

1 Cluster L3 Cache area **1.411**



L3 CACHE

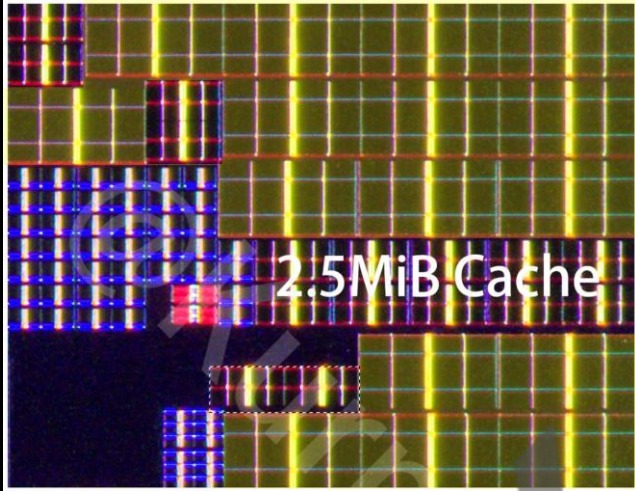


Slow Cache

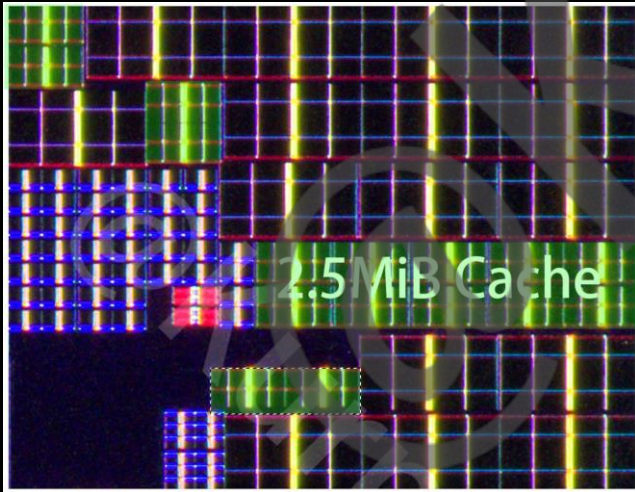
Fast Cache

In L3 Cache Cluster
They have 2 kinds of Cache

L3 CACHE

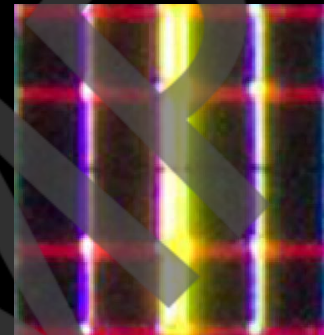


Slow Cache 2MiB



Fast Cache 0.5MiB

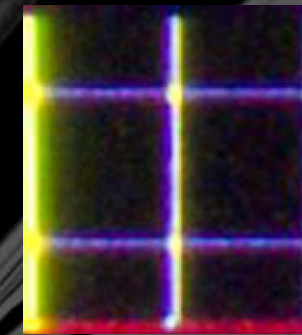
Both 64KiB/Block



Fast Cache Block

0.0275mm²/64KiB

2.27MiB/mm²



Slow Cache Block

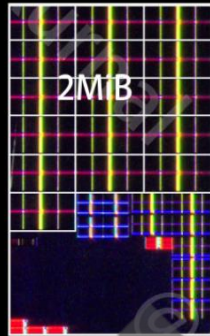
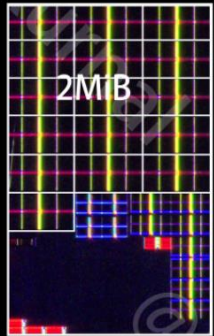
0.024mm²/64KiB

2.6MiB/mm²

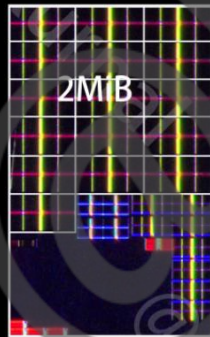
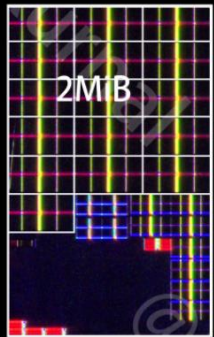
Cache Block ip desity

SYSTEM LEVEL CACHE

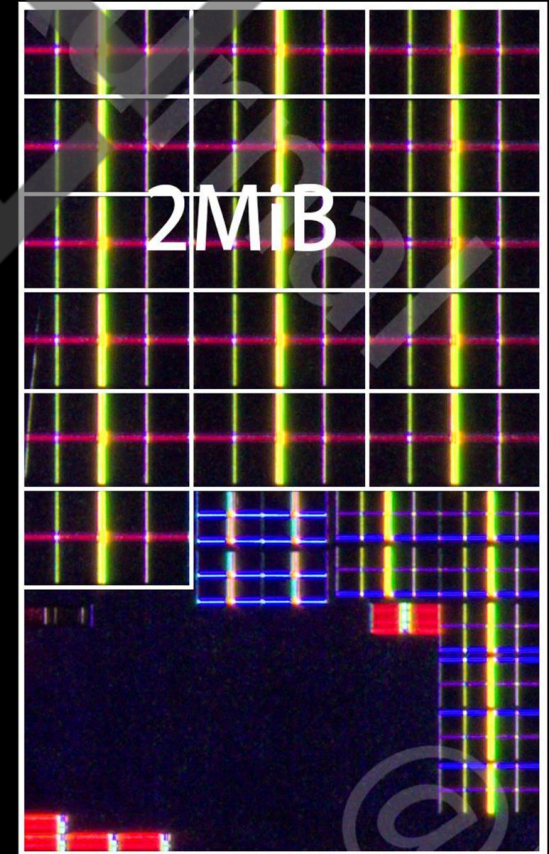
SYSTEM LEVEL CACHE



System Level Cache have 4 Cluster Cache
Each Cluster have 2MiB Cache

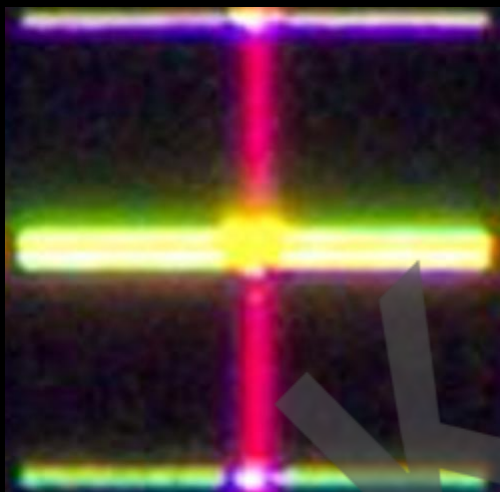


SLC Total area: 5.06mm²
SLC Cluster area: 1.265mm²



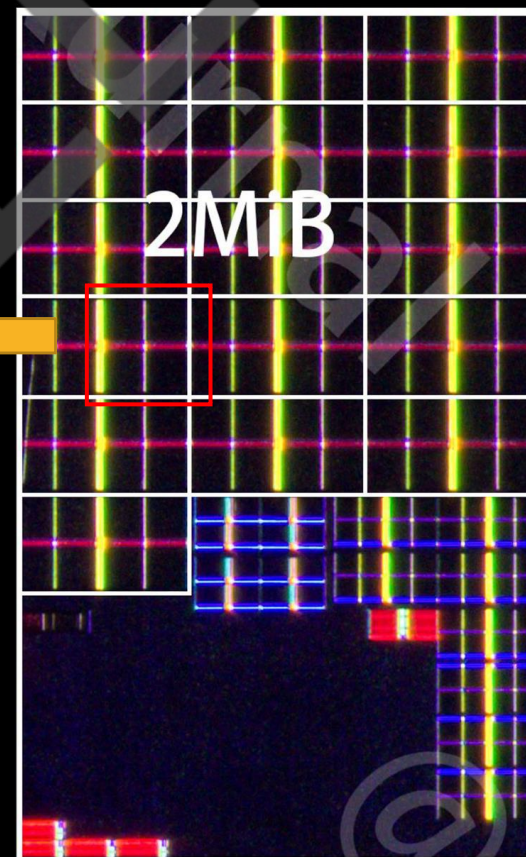
SYSTEM LEVEL CACHE

SLC Cache Block IP



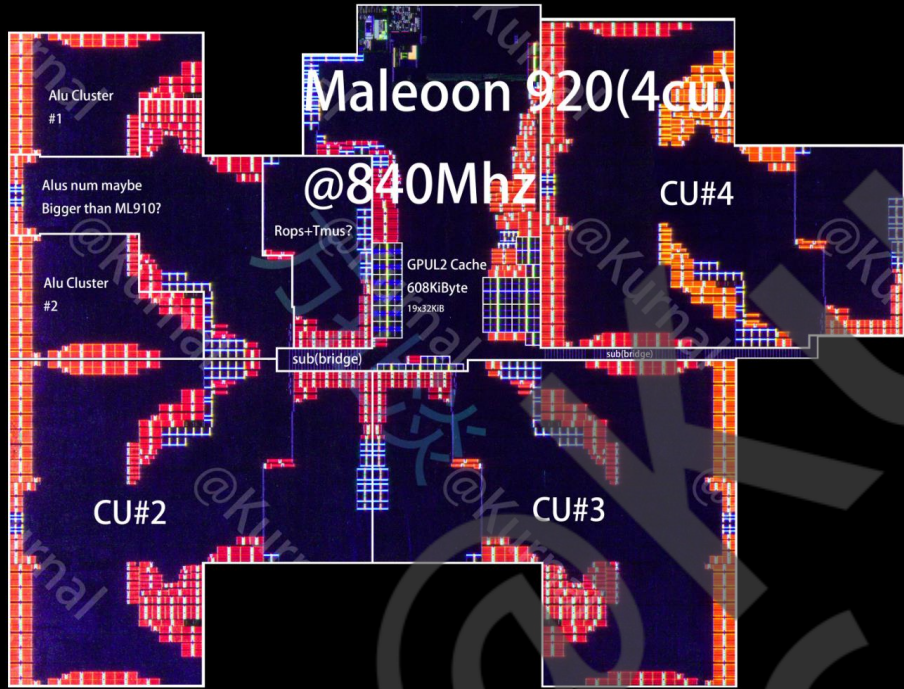
0.0235mm²/64KiB

2.66MiB/mm²



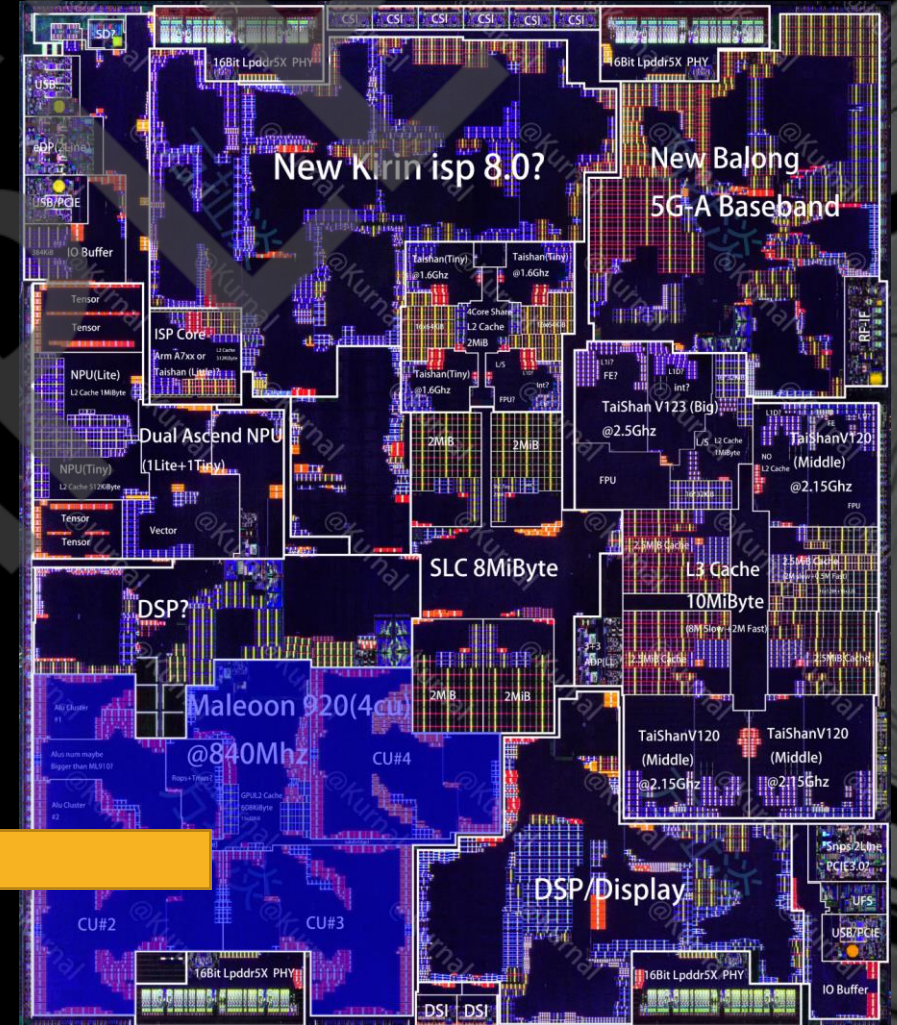
GPU ANALYZE

GPU ANALYZE

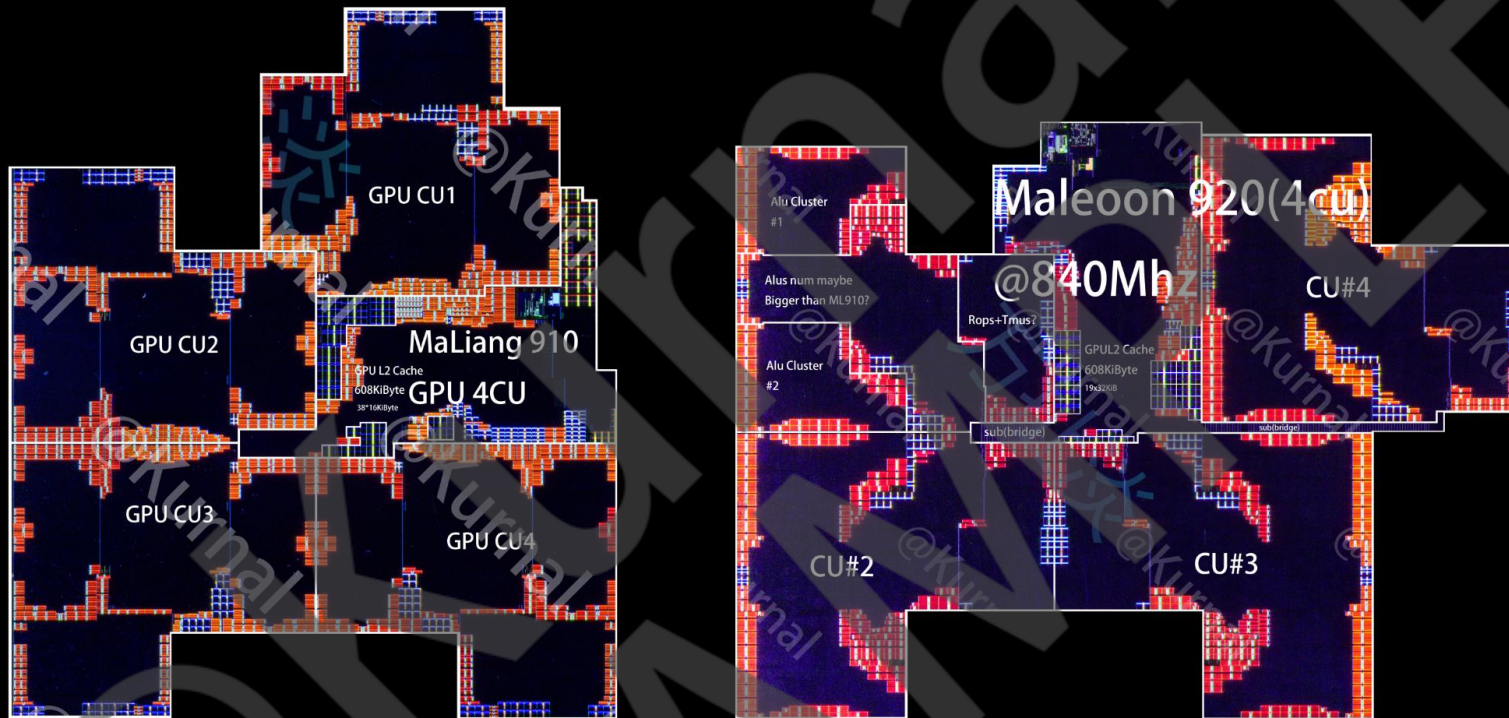


Next Gen GPU

Maleoon 920
@840Mhz



GPU ANALYZE



Kirin9000S/9010
Maleoon 910

Kirin9020
Maleoon 920

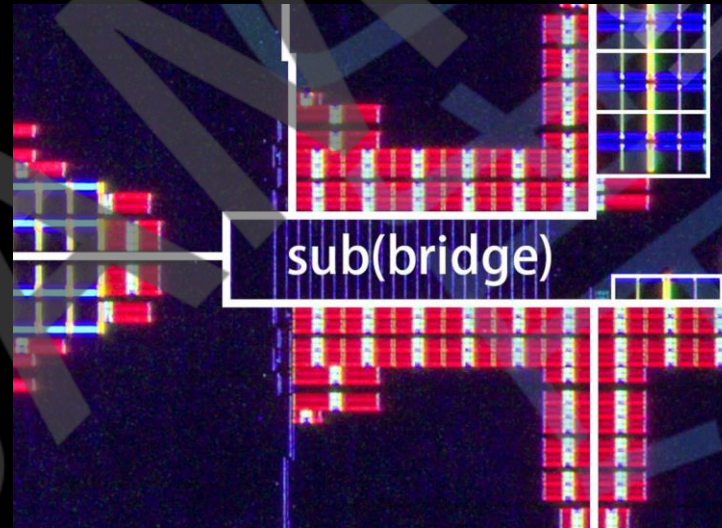
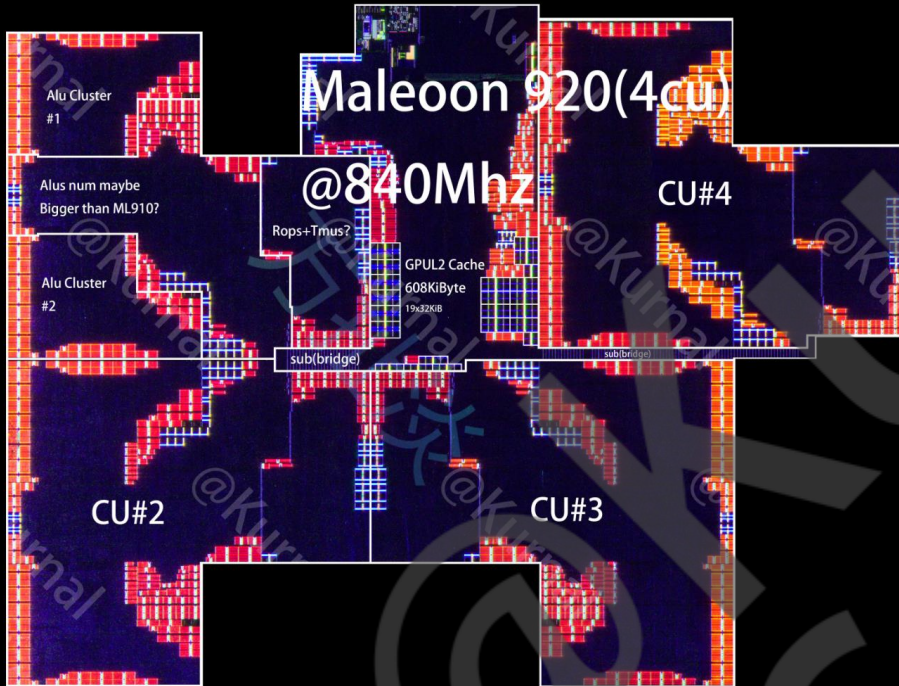
Small update?

GPU ANALYZE

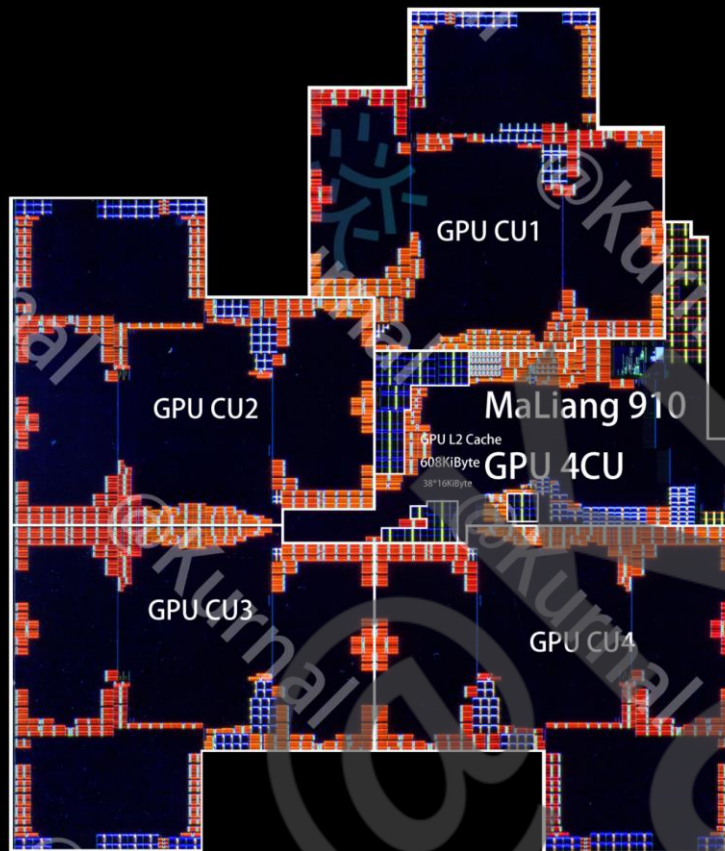
Maleoon 920 have 4 GPU CU

GPU L2 : 512KiB
GPU Size: 19.356mm²
GPU Size Per CU: 4.05mm²
GPU area per die size: 13.78%

Cu to Cu have the Metal layer Bridge



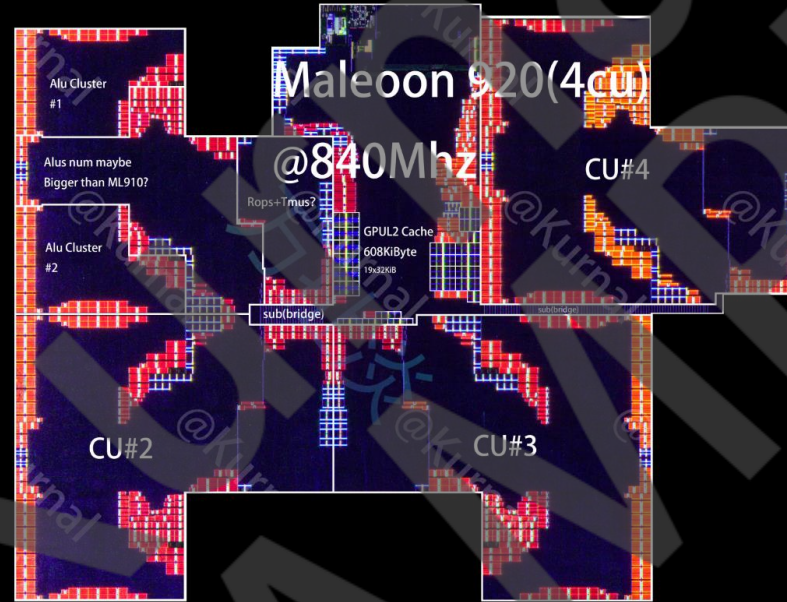
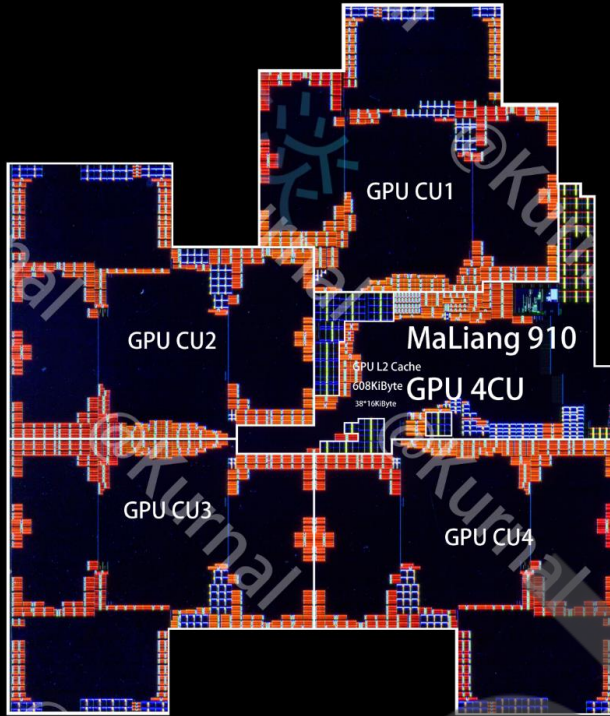
GPU ANALYZE



Maleoon 910 have 4 GPU CU

GPU L2 : 512KiB
GPU Size: 18.404mm²
GPU Size Per CU: 3.957mm²
GPU area per diesize: 15.75%

GPU ANALYZE



About ML910/ML920

The GPU L2 Size is same as **512KiB**
The GPU Cu number is same as **4 cu**

Size have a little bit different

GPU Size is bigger than nearly **1mm² / 7.36%**
CU Size is bigger than **0.1mm² / 2.35%**

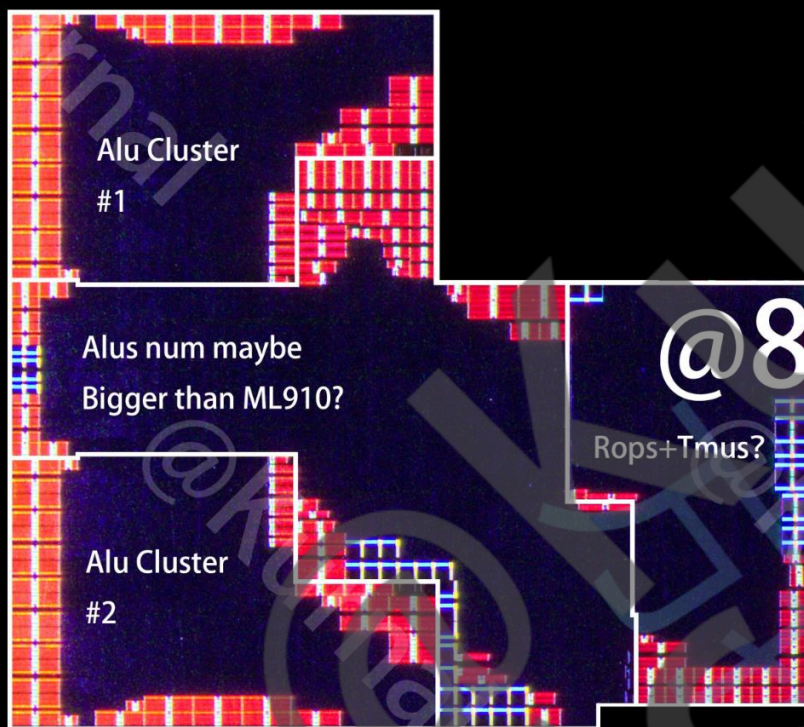
Maleoon 910 have **4 GPU CU**

GPU L2 : **512KiB**
GPU Size: **18.404mm²**
GPU Size Per CU: **3.957mm²**

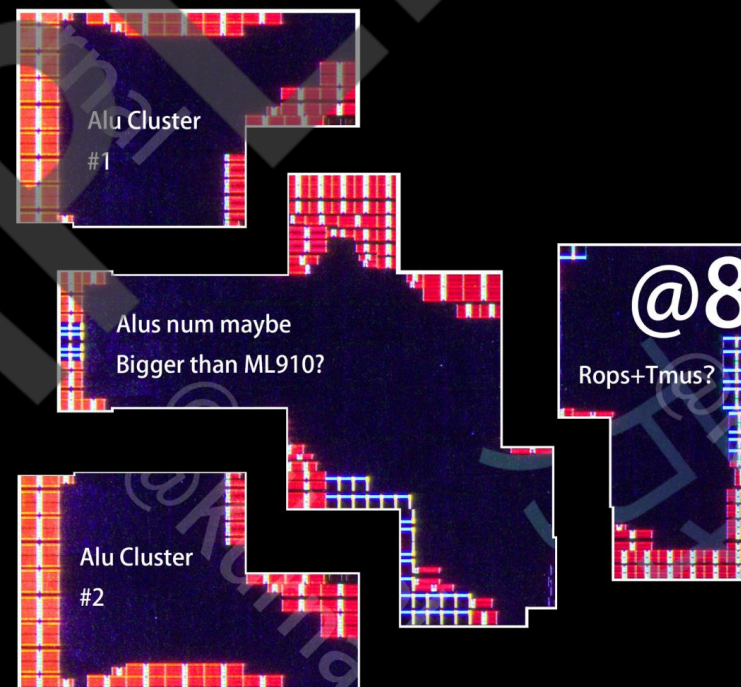
Maleoon 920 have **4 GPU CU**

GPU L2 : **512KiB**
GPU Size: **19.356mm²**
GPU Size Per CU: **4.05mm²**

GPU ANALYZE-MALEOON 920



Maleoon 920



Maleoon 920

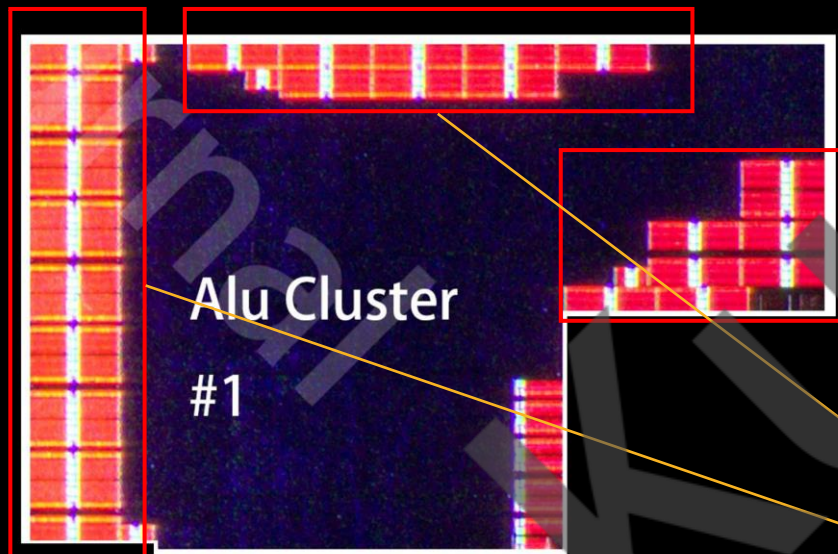
GPU ANALYZE-MALEOON 920



Each GPU Cu have 2 Alu Cluster



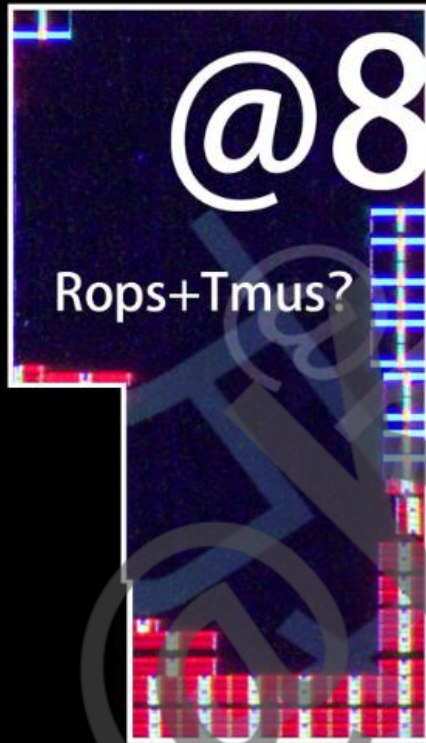
GPU ANALYZE-MALEOON 920



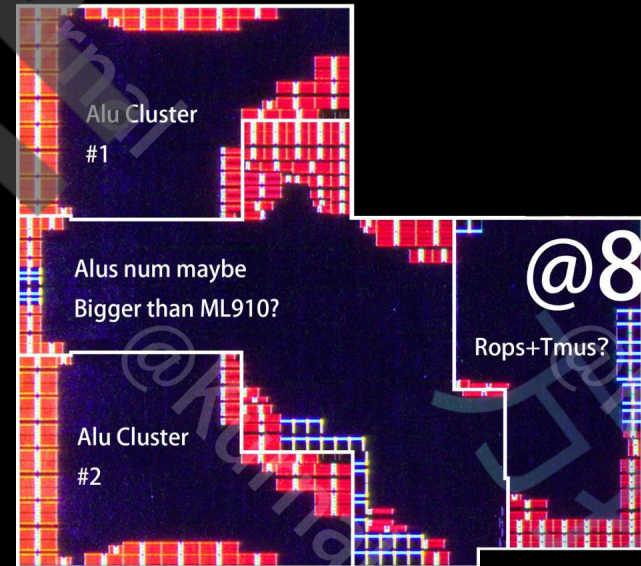
Alu Cluster size: 0.854mm²

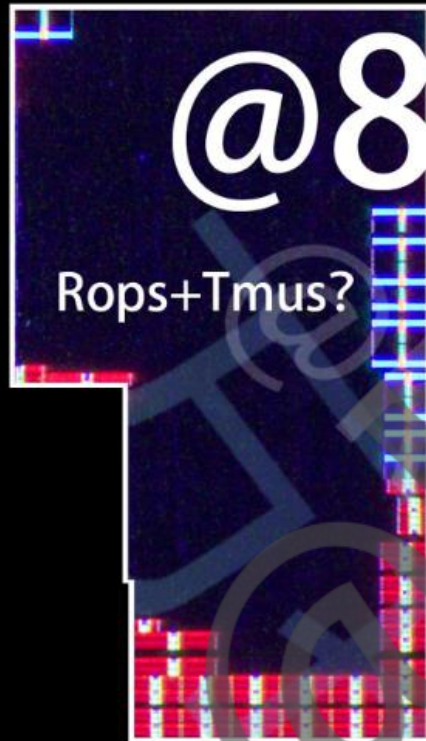
Every Alu cluster have
16 Block VGPR Cache





Each GPU Cu have
| RB/Shader Engine





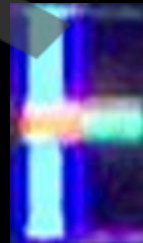
RB/Shader Engine Size: 0.758mm²

Every RB/SE have

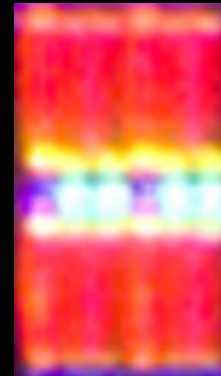
4Block Cache#1



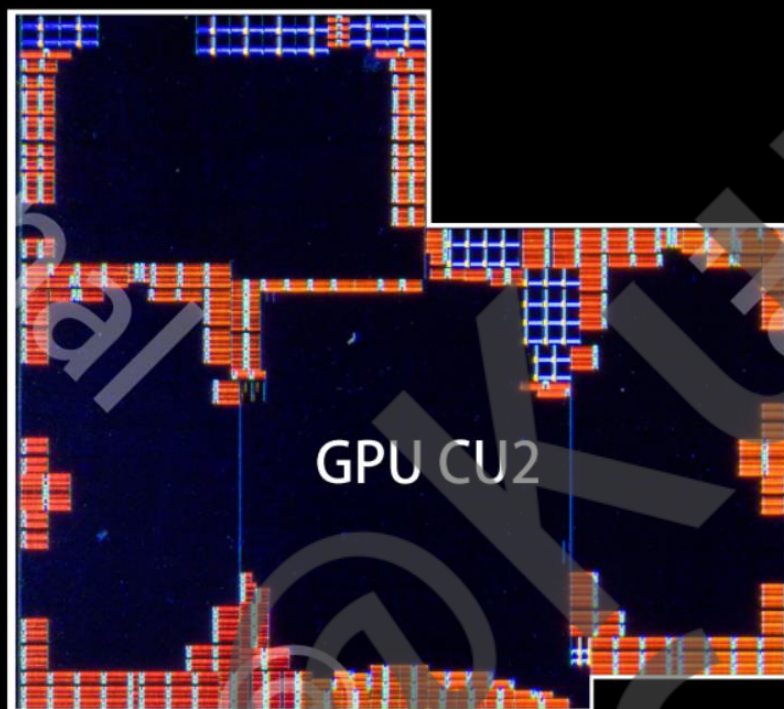
5Block Cache#2



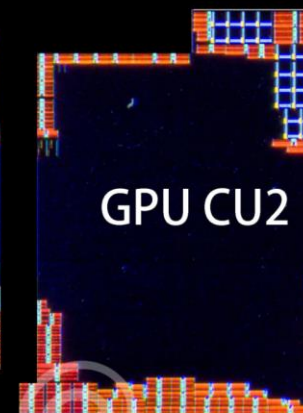
16Block Cache#3



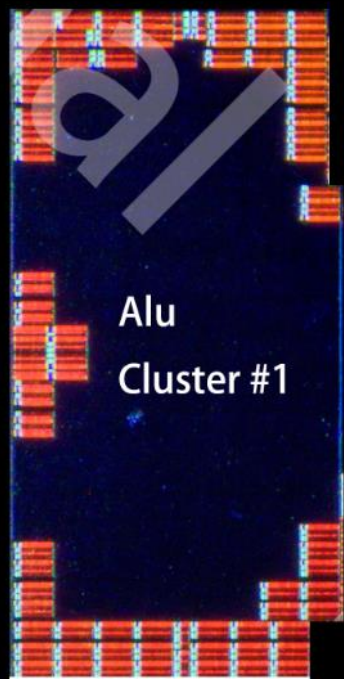
GPU ANALYZE-MALEOON 910



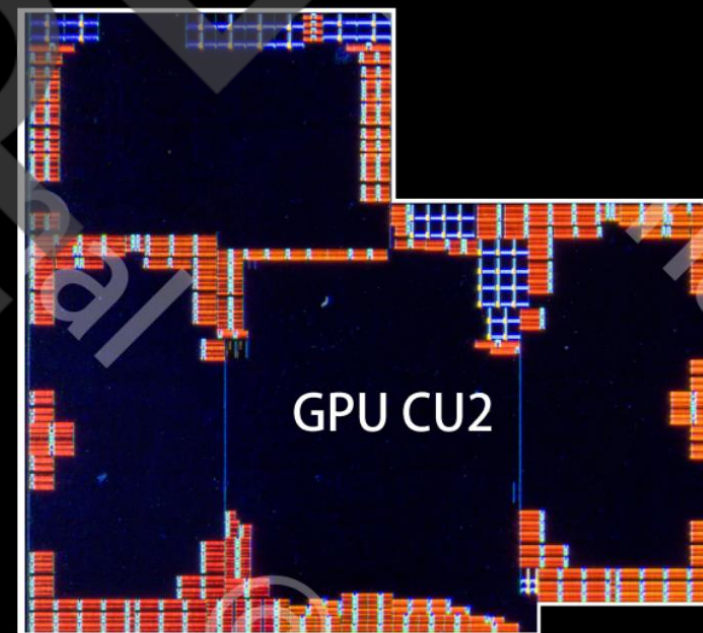
Maleoon 910



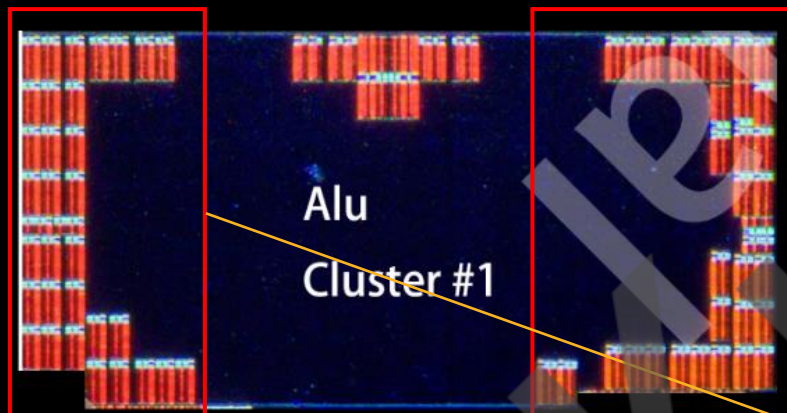
Maleoon 910



Each GPU Cu have 2 Alu Cluster

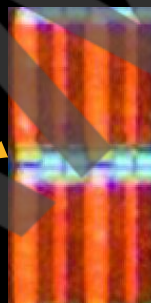


GPU ANALYZE-MALEOON 910

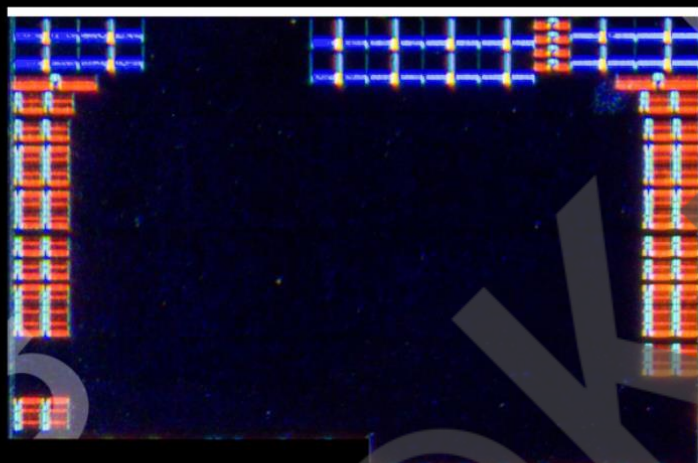


Alu Cluster size: **0.831mm²**

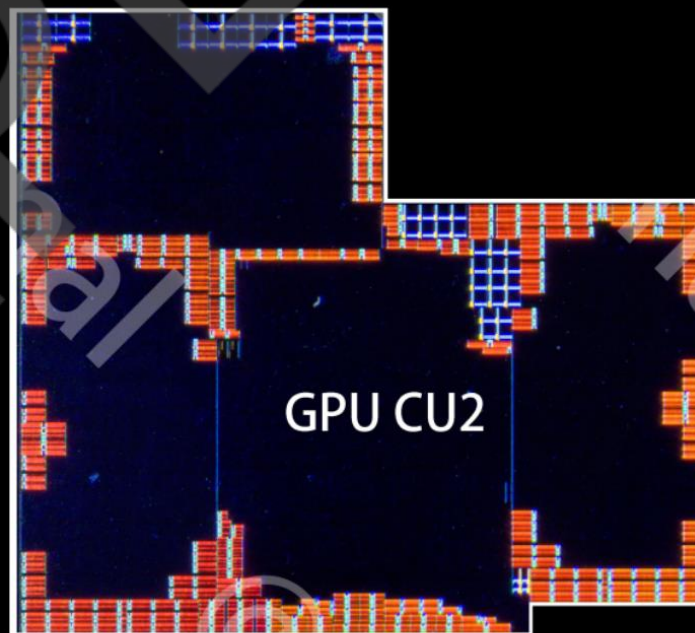
Every Alu cluster have **16 Block** VGPR Cache



GPU ANALYZE-MALEOON 910



Each GPU Cu have
| RB/Shader Engine



RB/Shader Engine Size: 0.897mm²

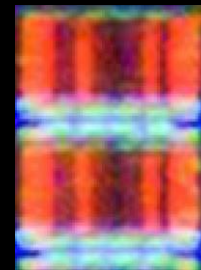
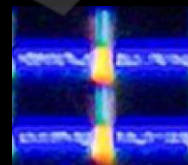


Every RB/SE have

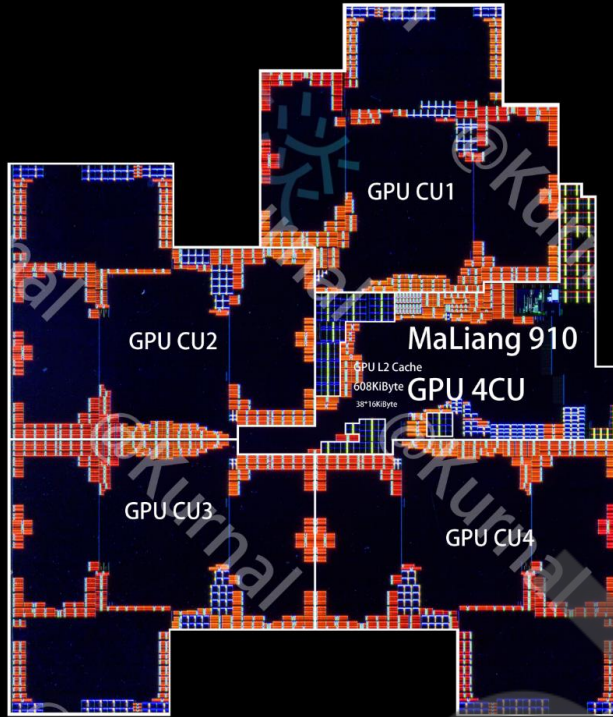
4Block Cache#1

4Block Cache#2

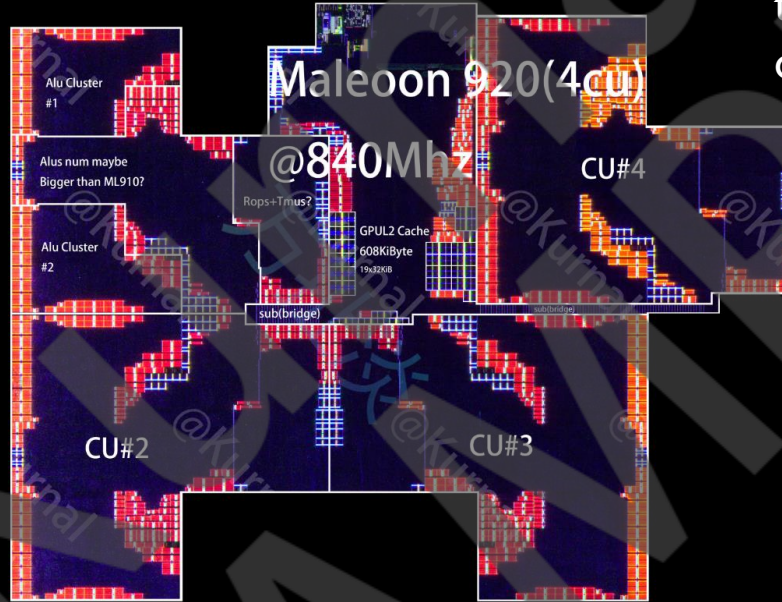
12Block Cache#3



GPU ANALYZE-910VS920



Maleoon 910



Maleoon 910

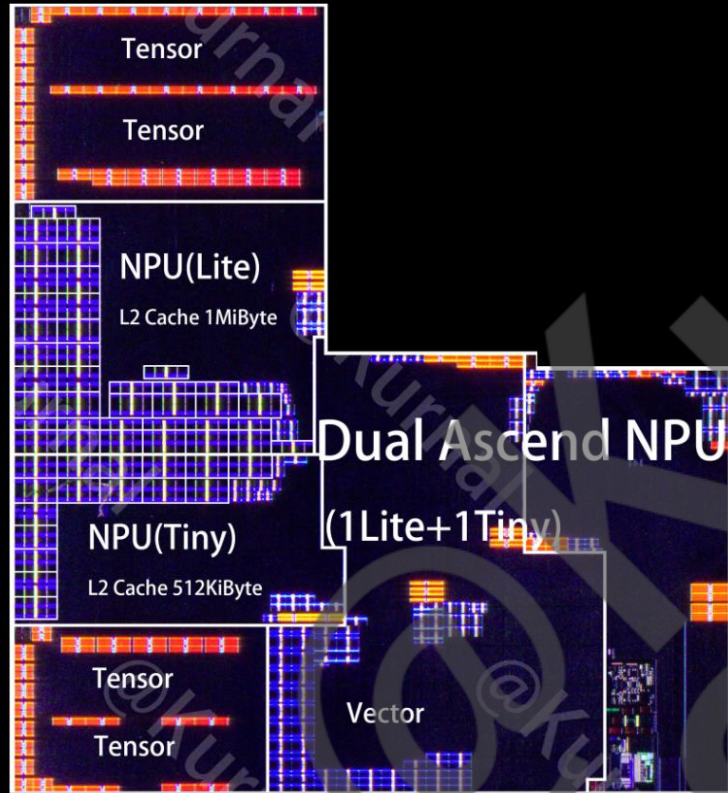
Half Gen GPU

In fact, the iteration and architecture improvement of this GPU is likely to be a small improvement.

	ML910	ML920
GPU Size	18.404	19.356
GPU CU Size	3.957	4.05
GPU Alu cluster size	0.831	0.854
GPU CU RB/SE size	0.897	0.758

NPU ANALYZE

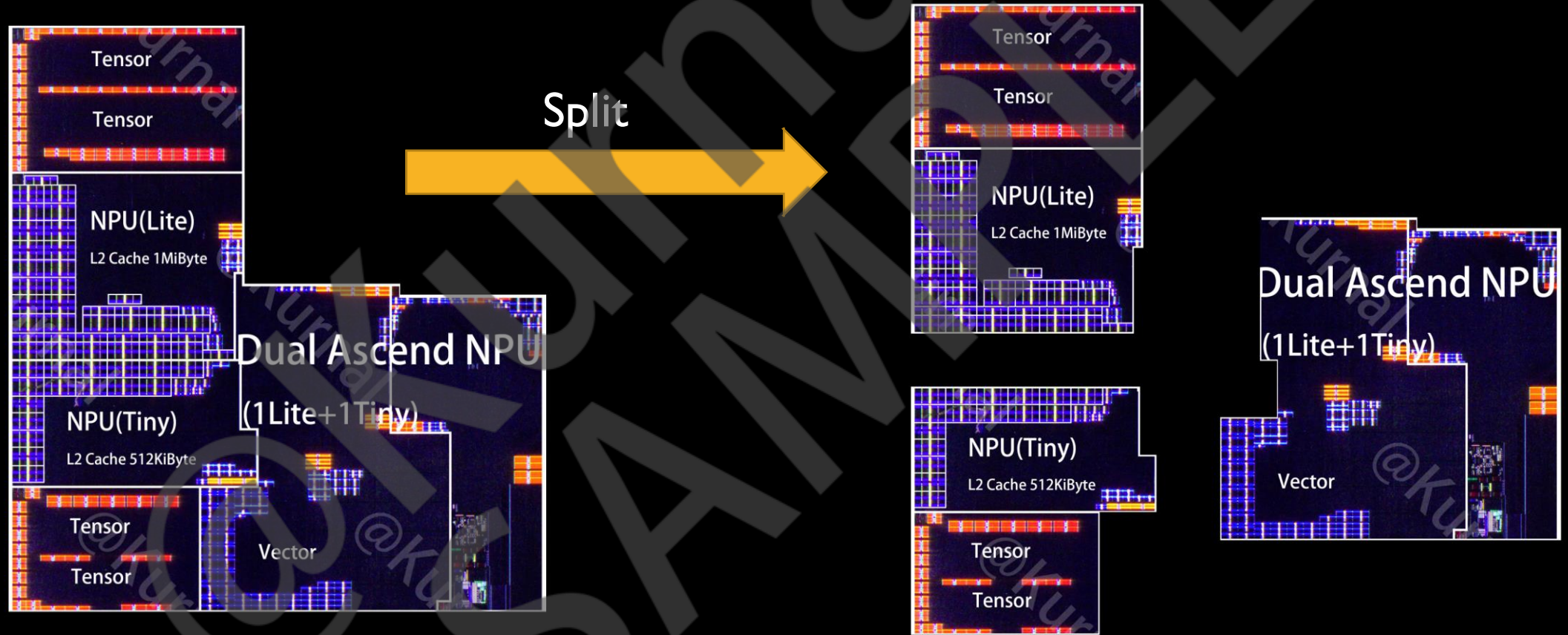
NPU ANALYZE



New Ascend NPU

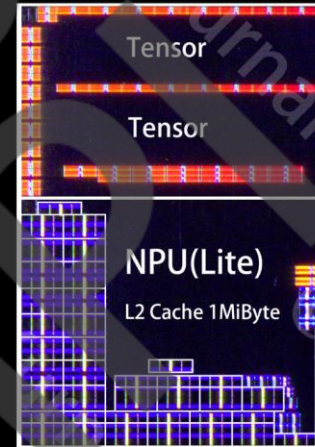


NPU ANALYZE

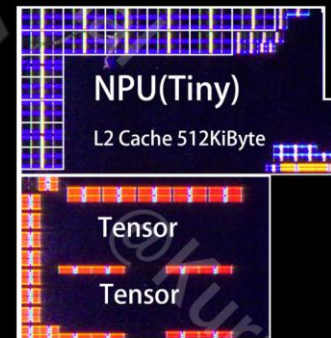


NPU ANALYZE

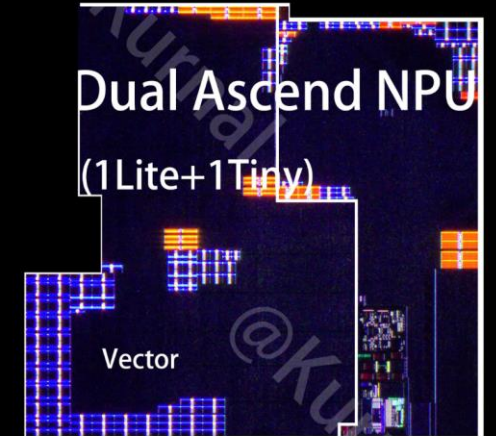
Ascend NPU **Lite** Cluster



Ascend NPU **Tiny** Cluster

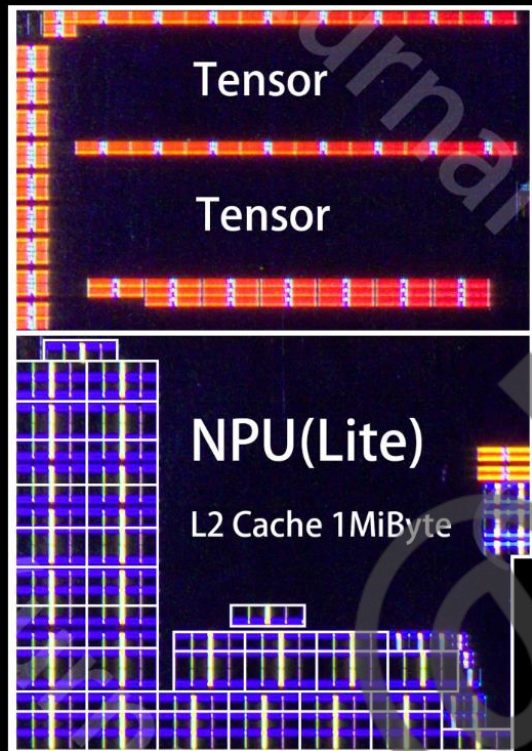


Dual Ascend NPU
(1Lite+1Tiny)



NPU ANALYZE

Ascend NPU Lite Cluster



Have **Dual** Tensor Unit

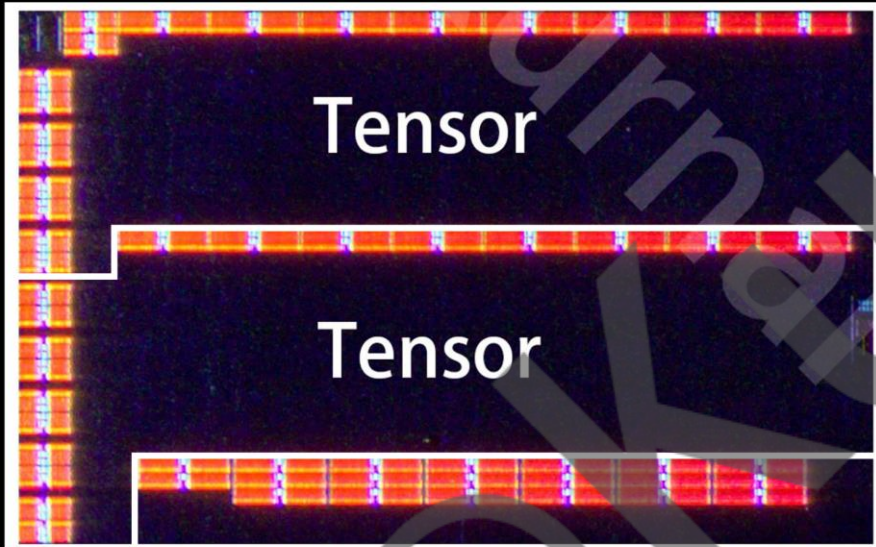


Have only Ascend Lite L2 Cache **1MiB**



Ascend NPU Lite Total Area: **2.62mm²**
Ascend NPU Lite Tensor Area: **1.168mm²**
Ascend NPU Lite L2 Area: **1.452mm²**

NPU ANALYZE

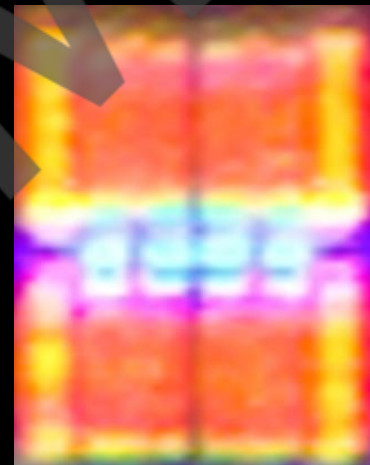


In tensor unit
Have 3 Sram ip

NPU Sram ip#1



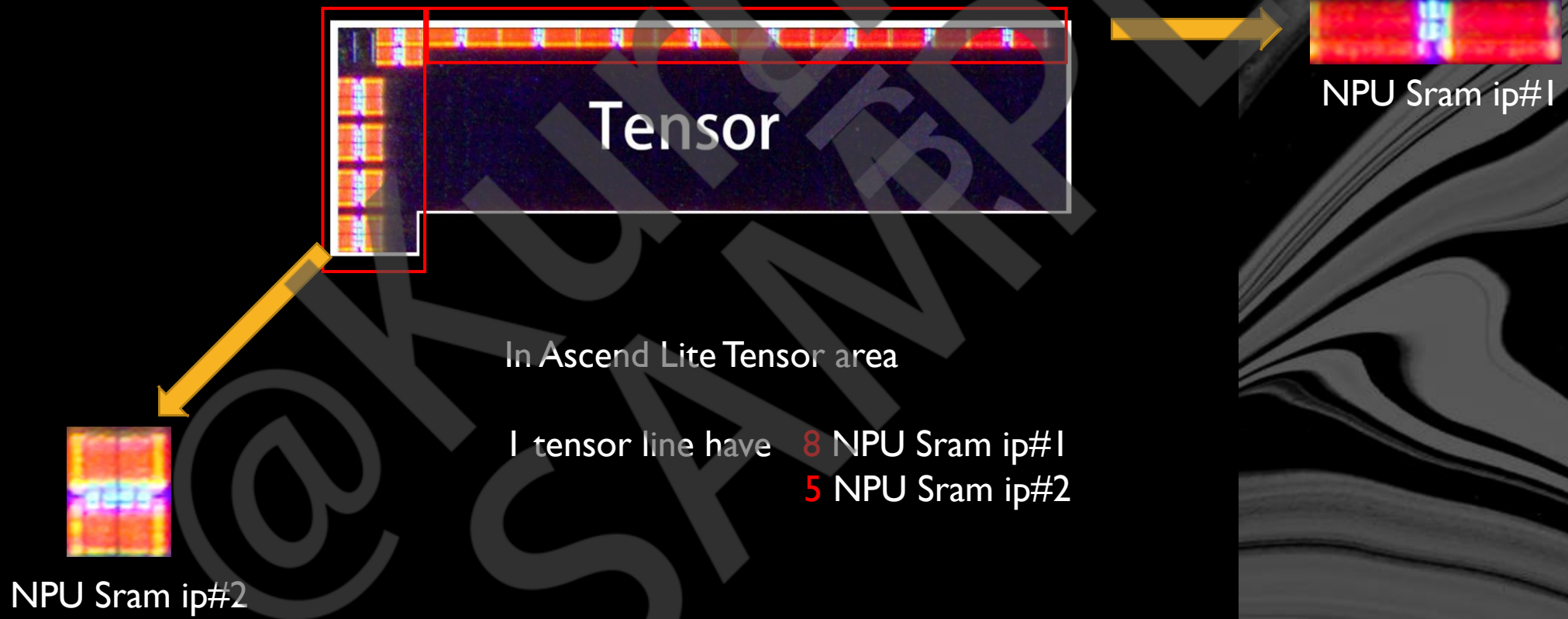
NPU Sram ip#2



NPU Sram ip#3



NPU ANALYZE



NPU ANALYZE



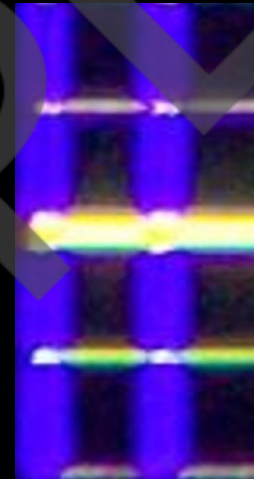
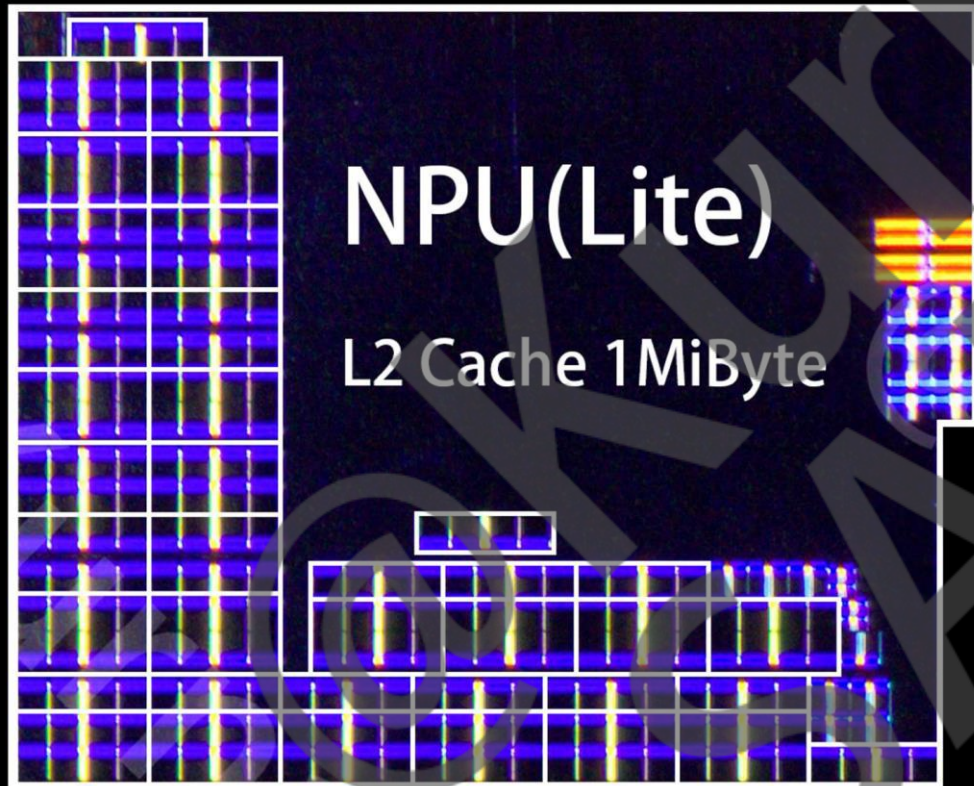
In Ascend Lite Tensor area

The tensor control have 20 NPU Sram ip#3

NPU ANALYZE

Have only Ascend Lite L2 Cache 1MiB

Ascend L2 block IP

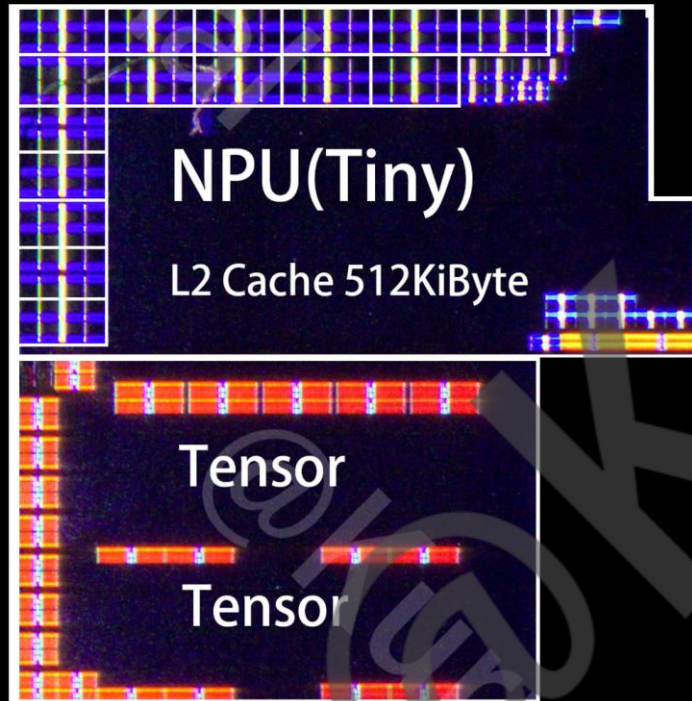


0.020mm²/64KiB

3.125MiB/mm²

NPU ANALYZE

Ascend NPU Tiny Total Area: 1.837mm²
Ascend NPU Tiny Tensor Area: 0.814mm²
Ascend NPU Tiny L2 Area: 1.023mm²



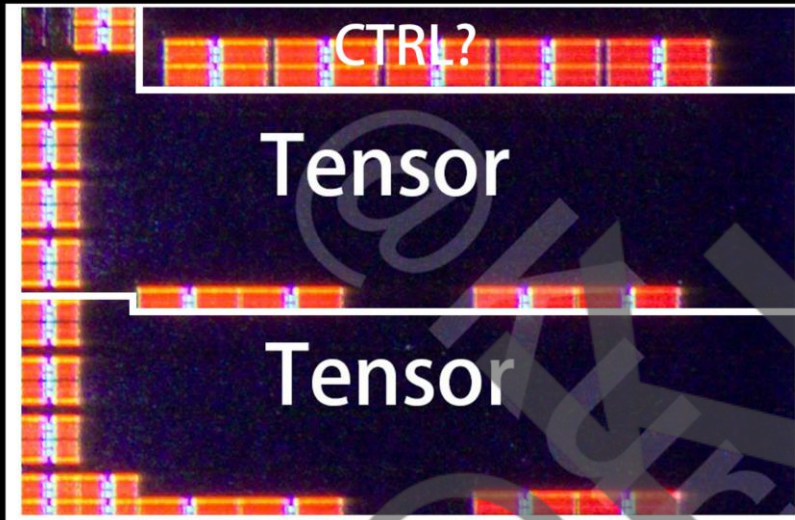
Have only Ascend Tiny L2 Cache 512KiB



Have Dual Tensor Unit



NPU ANALYZE

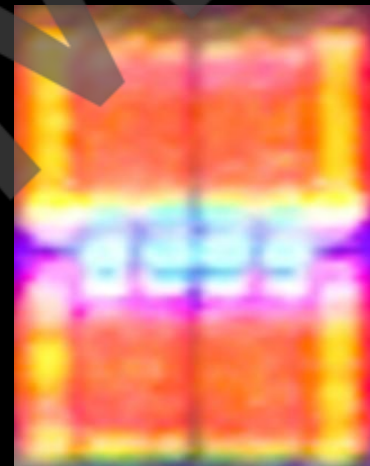


In tensor unit
Have 3 Sram ip

NPU Sram ip#1



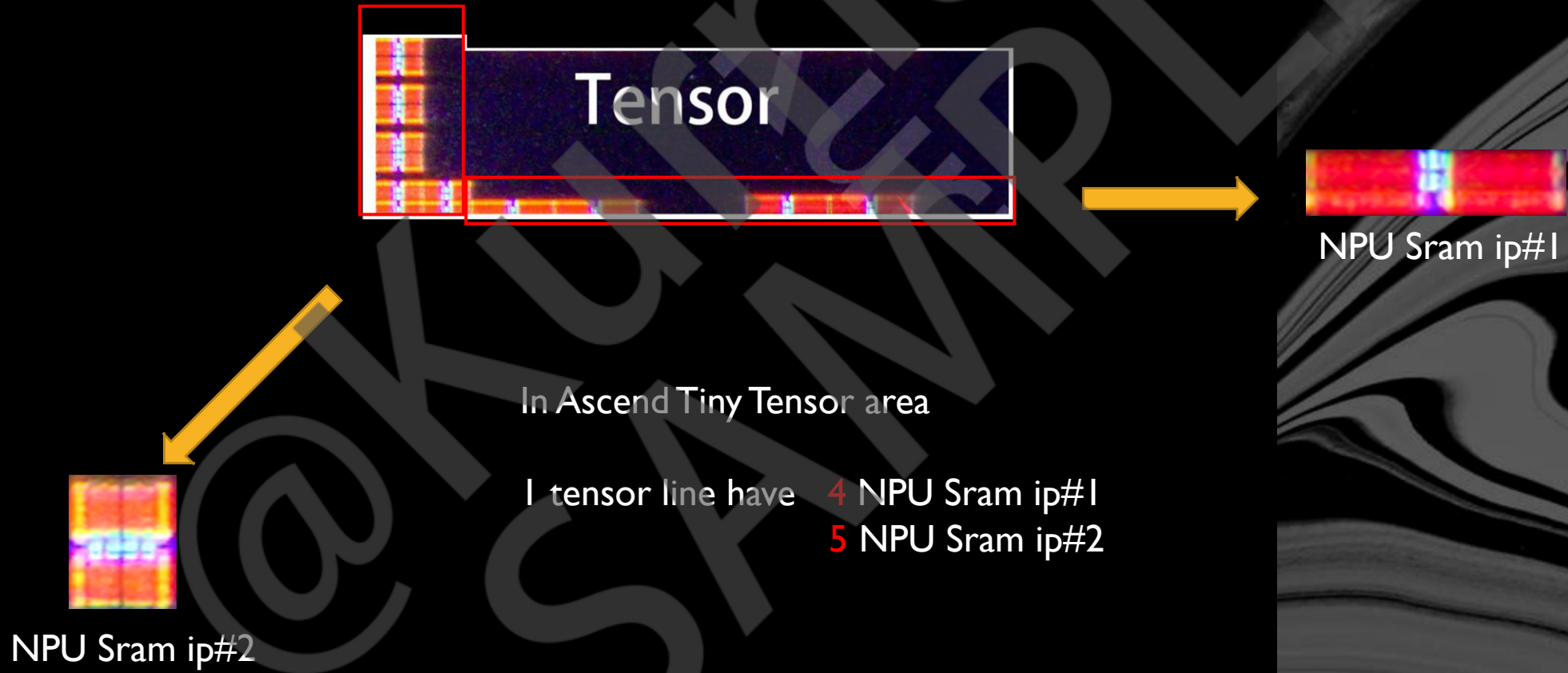
NPU Sram ip#2



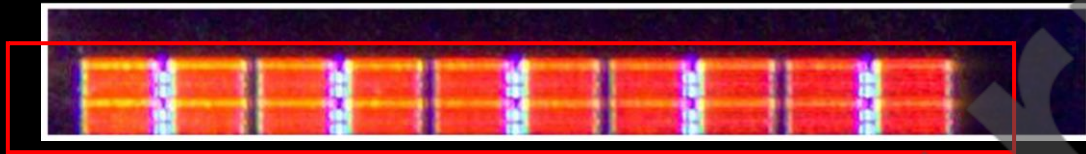
NPU Sram ip#3



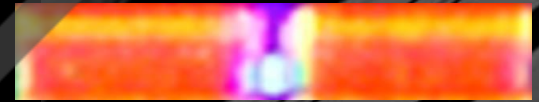
NPU ANALYZE



NPU ANALYZE



NPU Sram ip#3



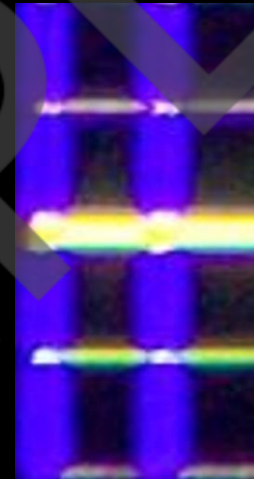
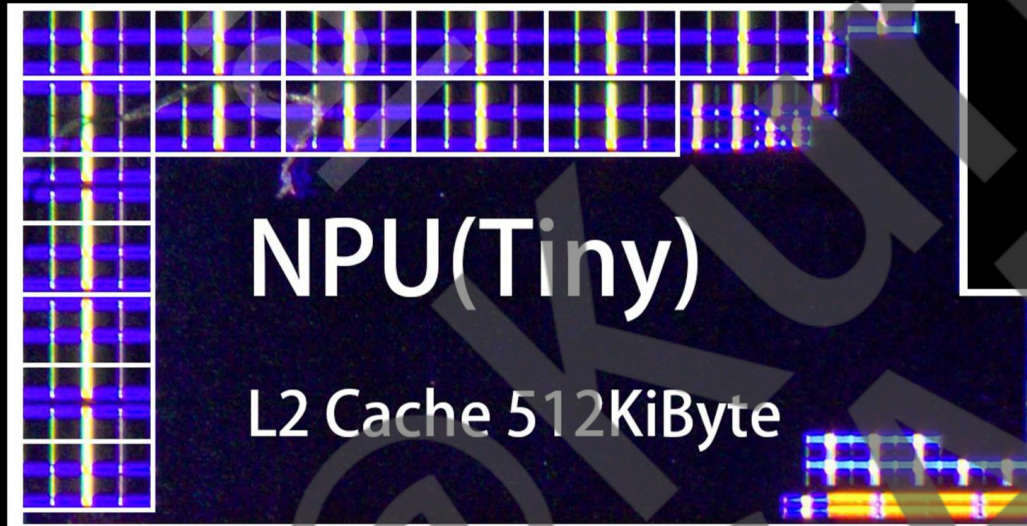
In Ascend Tiny Control area

The tensor control have 10 NPU Sram ip#3

NPU ANALYZE

Have only Ascend Tiny L2 Cache **512KiB**

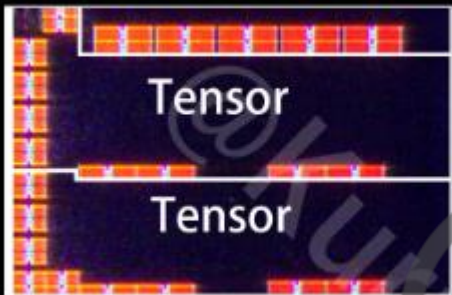
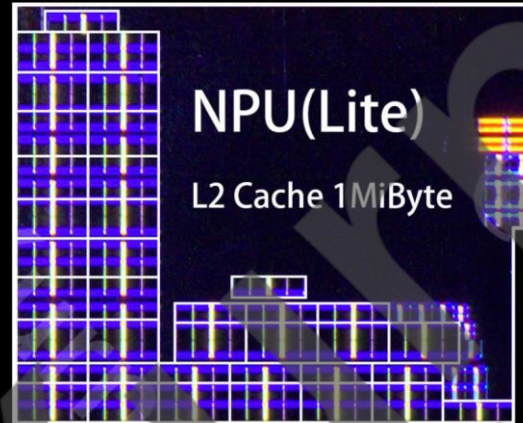
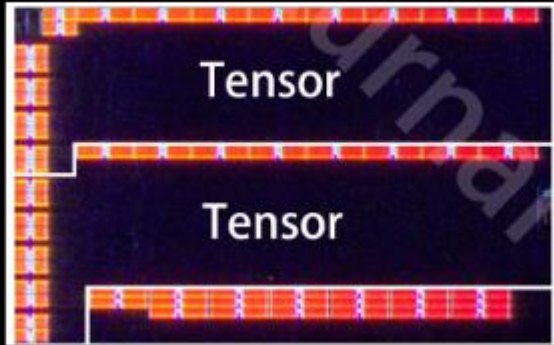
Ascend L2 block IP



0.020mm²/64KiB

3.125MiB/mm²

NPU ANALYZE



The data is half

Tensor	Sram ip1	Sram ip2	Sram ip3
Ascend Lite	8	5	20
Ascend Tiny	4	5	10

Area	Total	Tensor	L2
Ascend Lite	2.62	1.168	1.452
Ascend Tiny	1.837	0.814	1.023

So that, I think

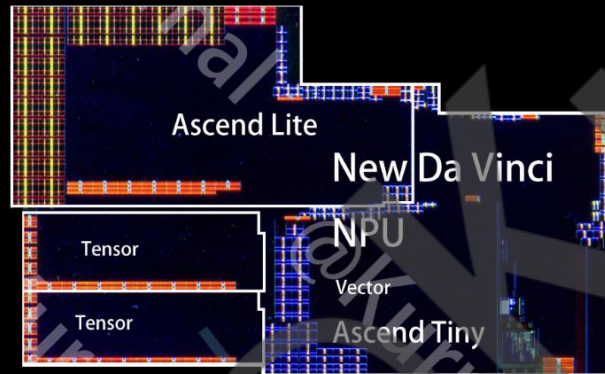
The **Ascend Tiny** use the **same Arch** with **Ascend Lite**

And

The **Perf of Ascend Tiny** is **Half** the **Ascend Lite**

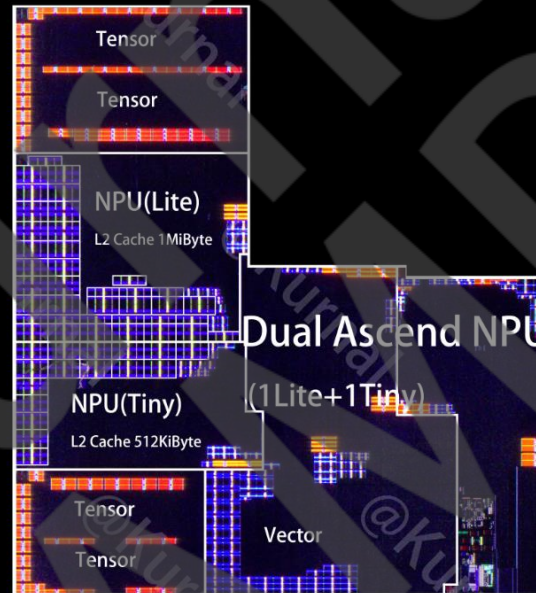
NPU ANALYZE VS

Kirin9000s/9010 NPU



NPU area: **6.528mm²**
NPU area Percent **5.59%**

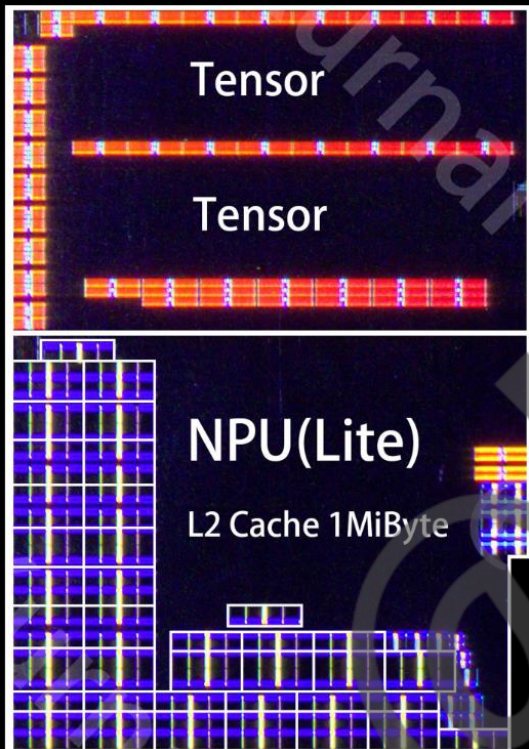
Kirin9020 NPU



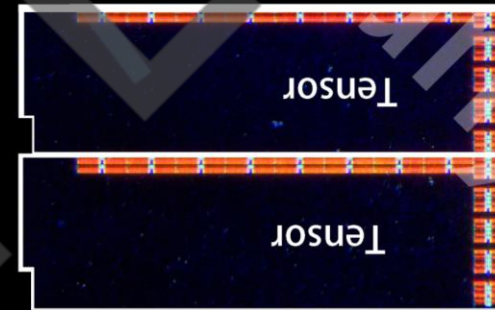
NPU area: **7.932mm²**
NPU area Percent **5.65%**

NPU ANALYZE

Kirin 9020 NPU Lite



Kirin 9010 NPU Lite



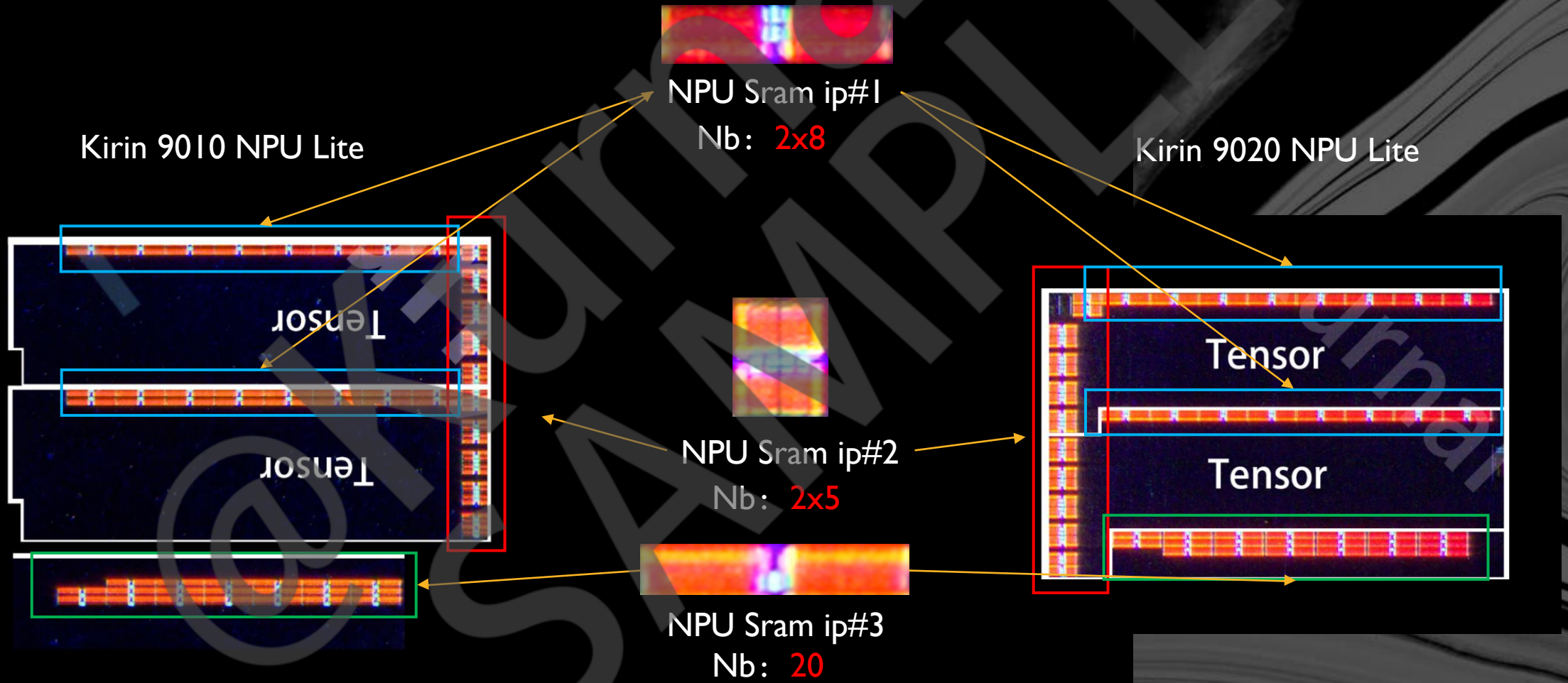
Have **Dual** Tensor Unit



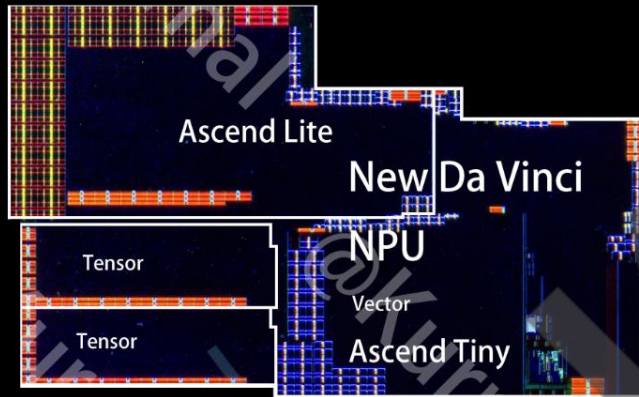
Have only Ascend Lite L2 Cache **1MiB**



NPU ANALYZE



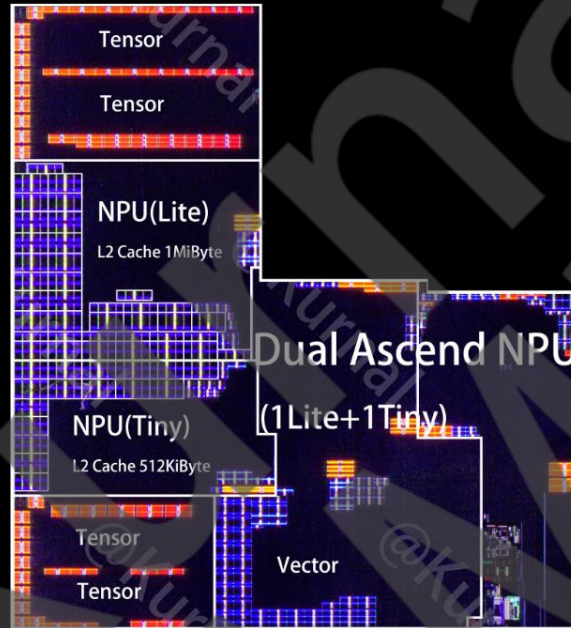
NPU ANALYZE



Kirin9000s/9010 NPU

NPU area: 6.528mm²

| Ascend NPU Lite



Kirin9020 NPU

NPU area: 7.932mm²

| Ascend NPU Lite

+

| Ascend NPU Tiny

The Kirin 9020 shares the **same NPU IP** (Ascend NPU Lite) as the 9000s/9010

but adds an Ascend NPU Tiny core

forming a **1+1 dual-core NPU architecture**

with the Tiny delivering **half** the Lite's scale

ISP ANALYZE

ISP ANALYZE



Isp size bigger than **5mm²/36.88%**
For isp percent is bigger than **1.61%**

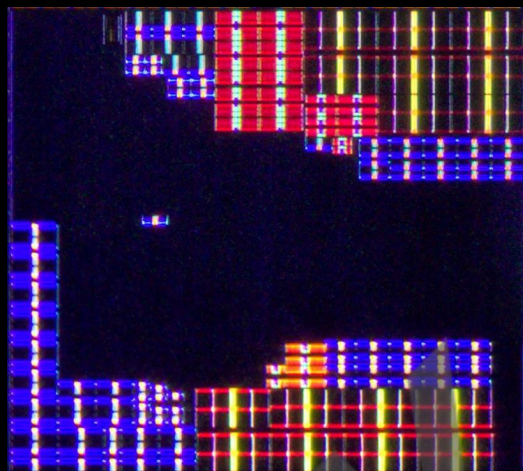
Kirin 9000S/9010 ISP IP (OLD)

New Kirin isp 7.0 Size **13.546mm²**
ISP size / Diesize **11.59%**

Kirin 9020 ISP IP (NEW)

New Kirin isp 8.0 Size **18.542mm²**
ISP size / Diesize **13.2%**

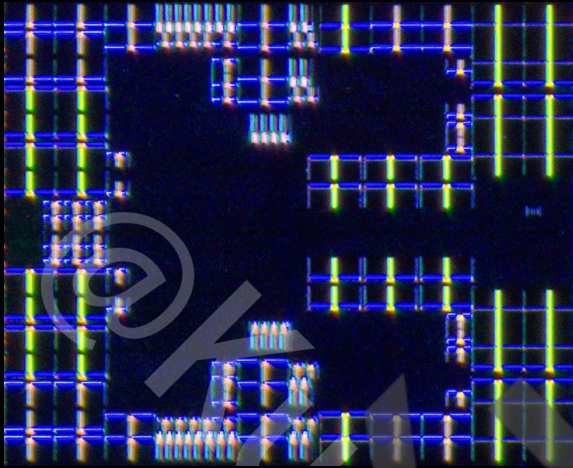
ISP ANALYZE



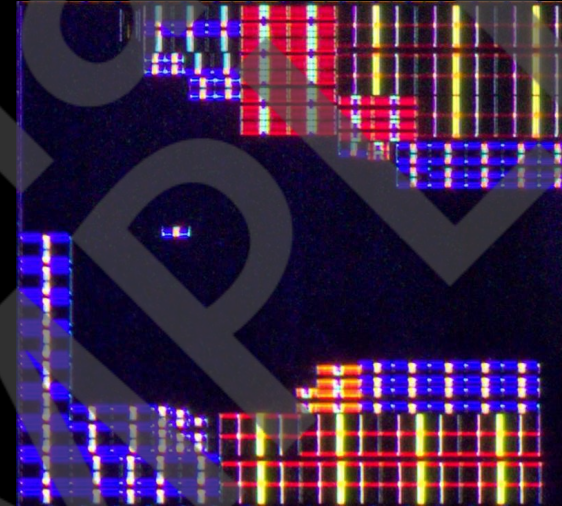
New ISP CPU IP



ISP ANALYZE



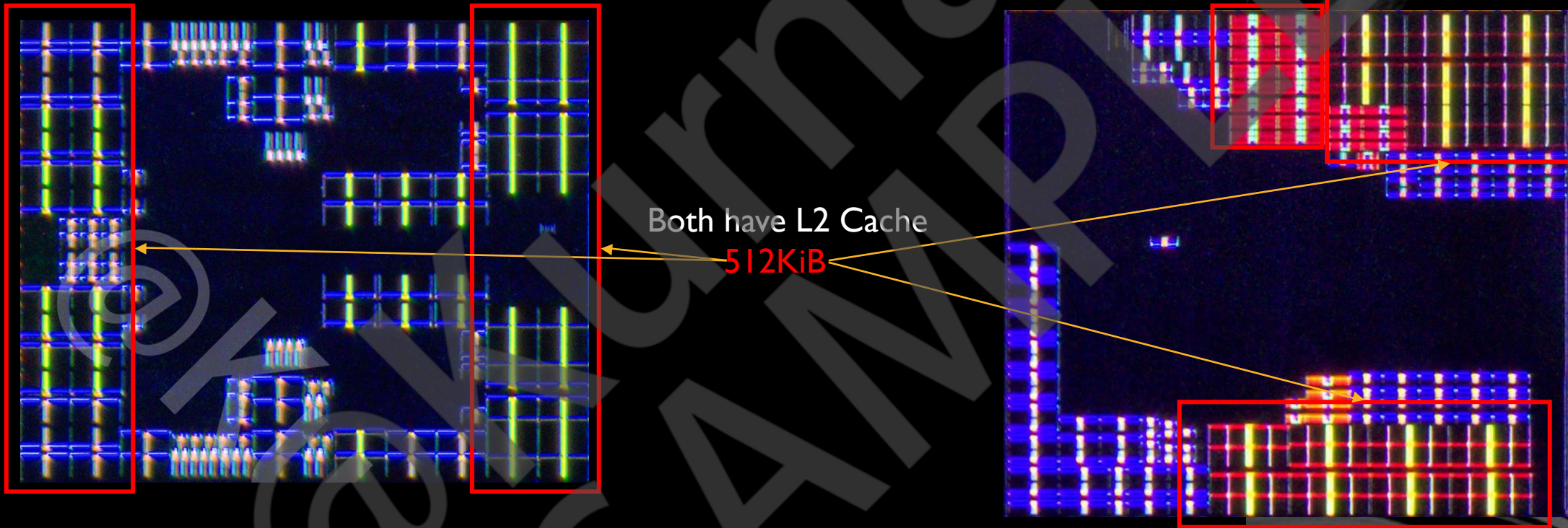
Kirin ISP 7.0 Size **0.982mm²**



Kirin ISP 8.0 Size **1.15mm²**

ISP CPU Size is bigger than **17.11%**

ISP ANALYZE



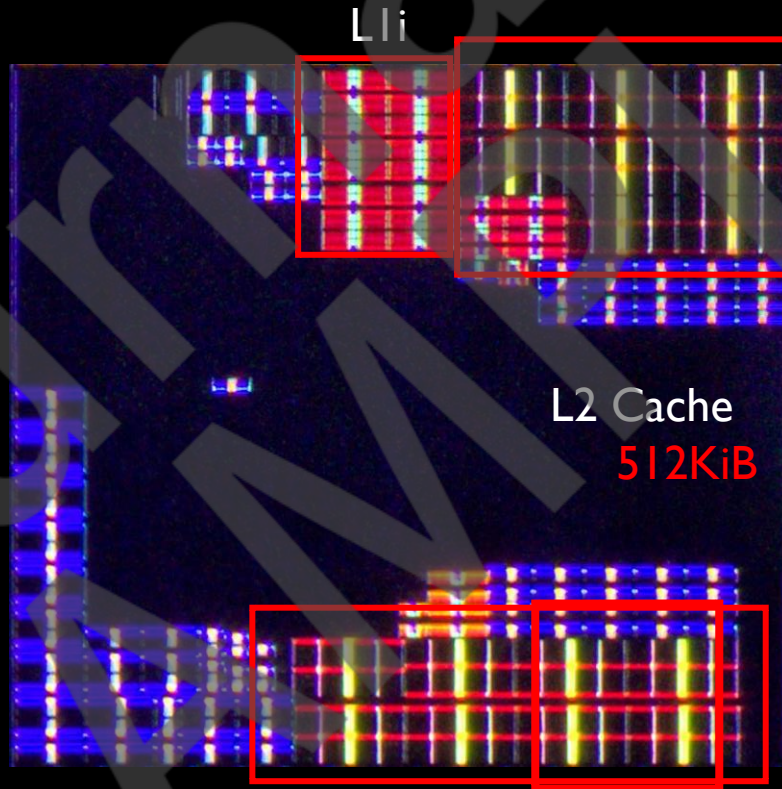
ISP ANALYZE

Taishan CPU Series



New Taishan Core

LID

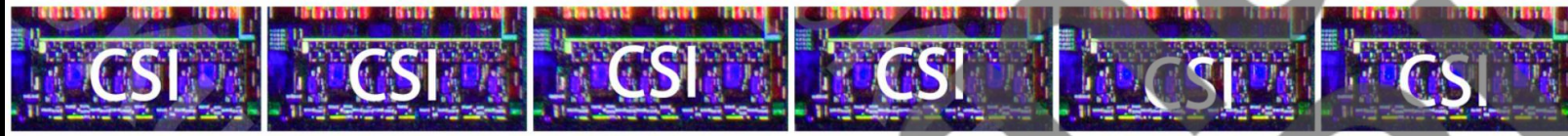


For just see
This core have 2 Level cache (L1/L2)

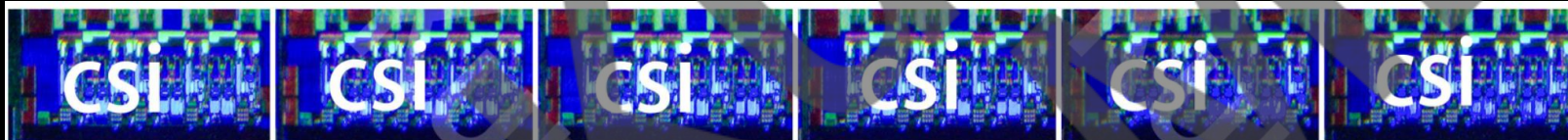
It is a new CPU
Nearly like Arm A7xx series

If it is ture a New CPU IP
Maybe it is the **Taishan Little?**

ISP ANALYZE



In Kirin9020
have **6Line** Camera Serial Interface

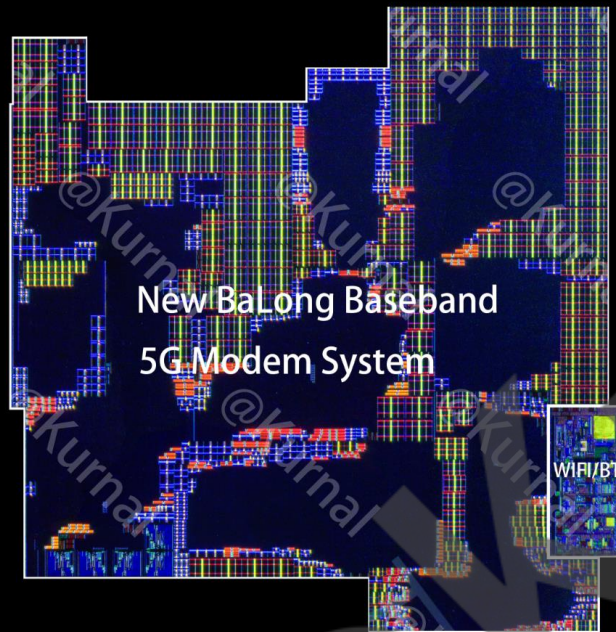


In Kirin9010
have **6Line** Camera Serial Interface

The CSI IP is **same**
And the CSI nb is same **6Line**

5G BASEBAND ANALYZE

5G BASEBAND ANALYZE



BP size Smaller than **2.78mm²/16.7%**
The proportion of BP decreased by **4.38%**
Nearly smaller than **30.74%**

Kirin 9000S/9010 Balong Baseband IP (OLD)

Balong Baseband Size **16.654mm²**
BP size / Diesize **14.25%**

Kirin 9020 Balong Baseband IP (NEW)

New Balong Baseband Size **13.873mm²**
BP size / Diesize **9.87%**

END

The author of this report is @Kurnal
Copyright @Kurnal

