Powered by UBS Evidence Lab

Global Semiconductors

AI Data Center Power Semis - watts the problem?

Equities

Global

Semiconductors

Francois-Xavier Bouvignies
Analyst
francois.bouvignies@ubs.com
+44-20-7568 7105

Harry Blaiklock, CFA
Analyst
harry.blaiklock@ubs.com
+44-20-7568 5385

Timothy Arcuri
Analyst
timothy.arcuri@ubs.com
+1-415-352 5676

Kenji Yasui
Analyst
kenji.yasui@ubs.com
+81-3-5208 6211

Nicolas Gaudois
Analyst
nicolas.gaudois@ubs.com
+65-6495 5148

Randy Abrams
Analyst
randy.abrams@ubs.com
+886-2-8722 7338

Sunny Lin
Analyst
sunny.lin@ubs.com
+886-2-8722 7346

Shingo Hirata, CFA
Analyst
shingo.hirata@ubs.com
+81-3-5208 6224

A deep dive into the semis powering AI data centers

We update our views on the growing opportunity for analog semis in the powering of AI data centers along side our APAC team, which we first published on in March. We now forecast a TAM of $13bn by 2028E vs. $11bn previously, rising to $15bn in 2030, driven by higher custom ASICs power semis content and higher rack estimates across GPU and custom ASICs. We see execution as the main risk to these estimates given the 51GW of capacity installed implied by our 2027E estimates growing to 105GW in 2030E.

1. Revamped BoM estimates and higher units estimates drives TAM higher

We have expanded our bill of materials estimates to 30+ xPU generations, beyond the six GPU configurations we previously modelled. We have also adjusted component count to align rack configurations with UBS Taiwan Hardware team and refined ASP estimates based on deep dives on the shifting product architecture of each power stage. Our new estimates drive a content per GW range of $100-200m/GW and an average rising from $95m/GW in 2024 to $130m/GW in 2026E and $150-170m/GW beyond.

2. Is there enough available power to support these estimates?

Our estimates imply c51GW of capacity added in 2027, increasing steadily to c105GW in 2030, vs market expectations of c30-50GW annually. Cumulatively our estimates imply 340GW of capacity additions from 2026-2030E vs. <250GW in a 30-50W scenario. This also compares to c120GW of power grid additions in the US, c60-70% of global data center buildout, over the same period. The delta could be bridged with behind-the-meter generation, lower retirements of energy generation facilities and faster and larger buildout of new facilities than expected.

3. 800VDC delay could lead to $500m-1bn reduction in power semis TAM

Reports of a delay to adoption of the 800VDC power sidecar architecture from high volume adoption in 2028 vs. 2027 previously will likely impact the power semis TAM. We assume 60% penetration of power sidecar for Rubin Ultra in 2028E, which if 100% would add $400-500m to the TAM in our estimates. We forecast Rubin Ultra with power sidecar requiring $170m/GW of power semis vs. $119m/GW for integrated power racks, driven largely by battery backup unit, protection and the new high voltage IBC.

4. Power supply is tightening but is unlikely the key bottleneck

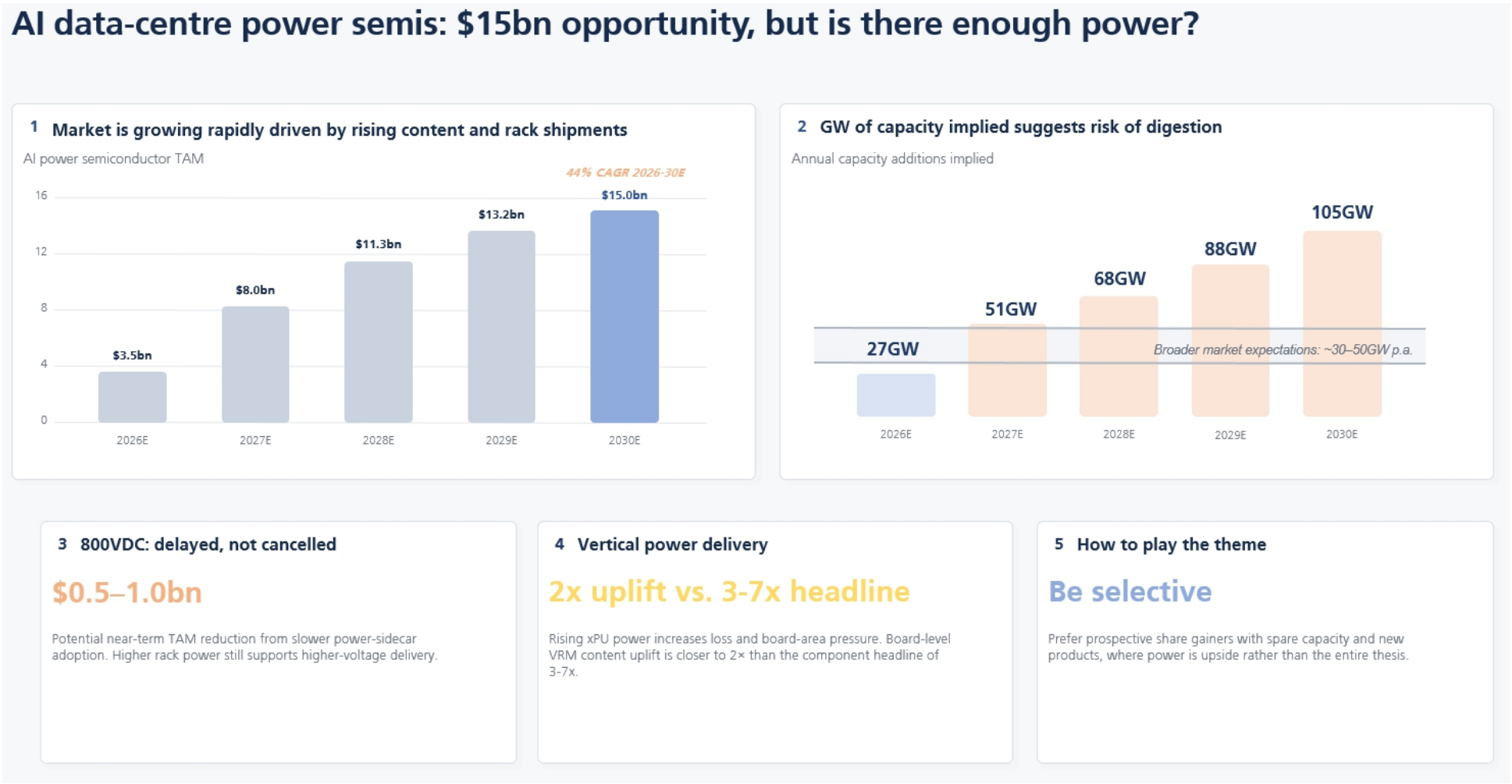
While power semiconductor lead times are rising, we view electrical infrastructure as an earlier bottleneck to AI buildout. Lead times for backup power systems and switchgear extend to ~18 months and beyond 2-3 years for utility scale transformers and on-site gas turbines. Clearly power semis can improve efficiency but against that backdrop, a 5-10% improvement in power delivered to the GPU, while beneficial, is unlikely to move the needle materially. Solid-state transformers offer a longer-term solution, but meaningful adoption remains a post-2028 story at the earliest.

How to play the theme

Given the potential for some level of digestion, we would take a selective approach to playing the theme. We are cautious on market share leaders such as Infineon (Neutral), given high investor expectations and increasing competition. We prefer potential market share winners with spare capacity, bringing new products to markets where they are not currently present such as TI (Buy) with PSUs and IBCs and Analog Devices (Buy) with VRMs. We also prefer stories where power is upside rather than core to the thesis, such as STMicro (Buy) with it's photonics exposure.

This report has been prepared by UBS AG London Branch. **ANALYST CERTIFICATION AND REQUIRED DISCLOSURES, INCLUDING INFORMATION ON THE QUANTITATIVE RESEARCH REVIEW PUBLISHED BY UBS, BEGIN ON PAGE 29.**

Figure 1: AI Data-center power semis: $15bn opportunity, but is there enough power?



Source: UBS

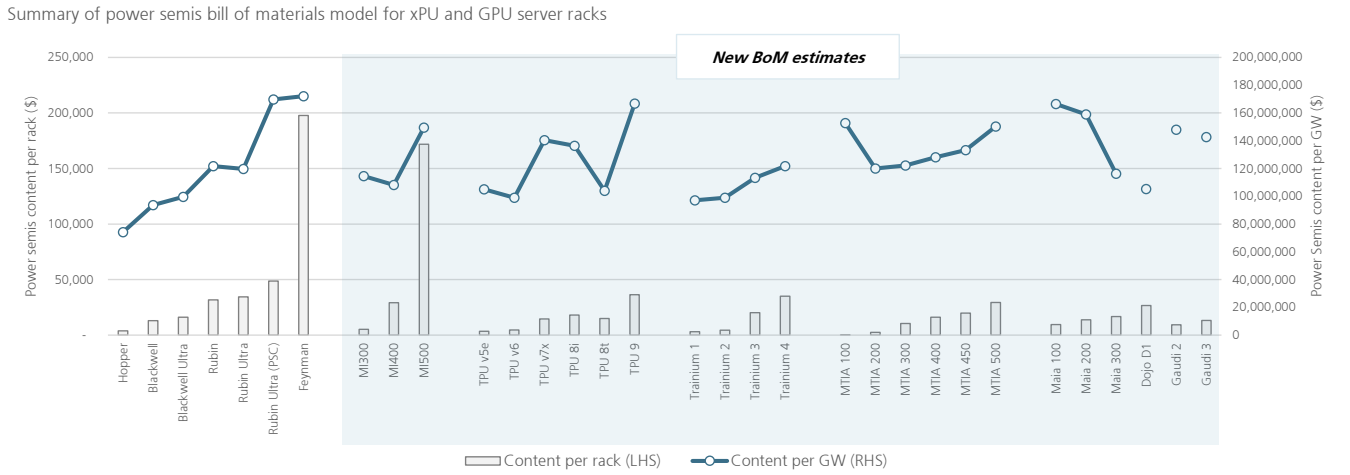
AI DATA CENTER POWER SEMIS: CAN THE WATTS DELIVER THE $15BN OPPORTUNITY

Key Conclusion: AI data-centre power semiconductors are one of the strongest growth opportunities in analog, with the market rising from c$3.5bn in 2026E to around $15bn by 2029-30E. However, the required pace of data-centre capacity additions appears above broader infrastructure forecasts, increasing the risk of a digestion period. We prefer prospective share gainers with excess capacity over incumbent leaders already discounting strong execution.

1. A larger and better-underpinned TAM

Our rebuilt bottom-up model now covers more than 30 GPU and custom-ASIC rack configurations across Nvidia, AMD, Google, Amazon, Meta, Microsoft and other platforms.

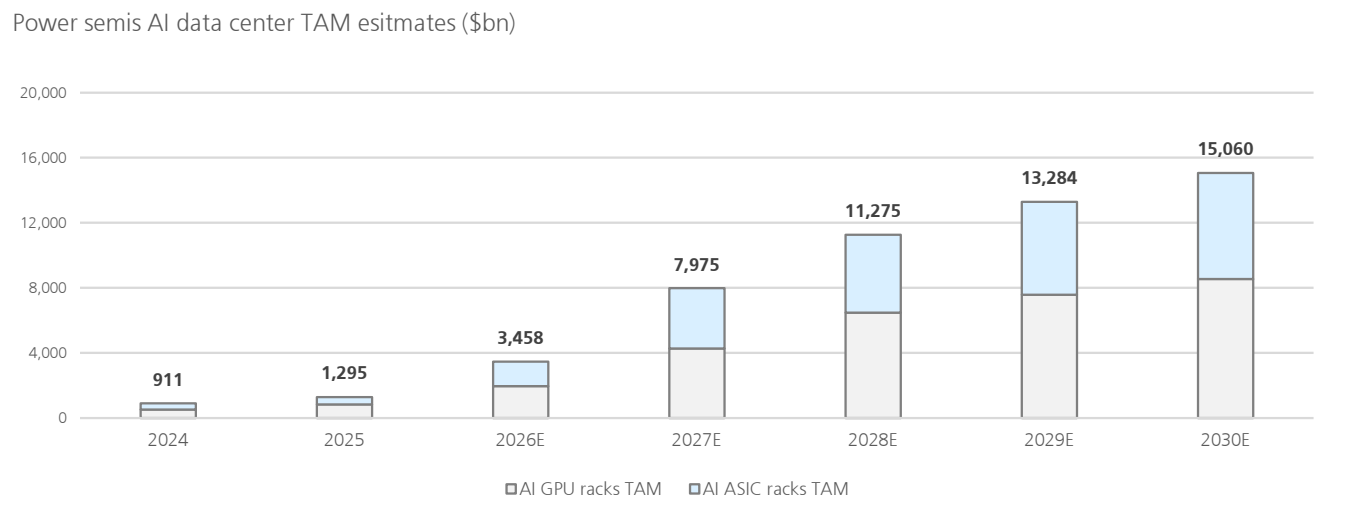
Figure 2: we expand our bill of materials model to include 30+ xPUs



Source: UBS estimates

Higher rack forecasts and refined component assumptions support a power semiconductor TAM of approximately $15bn by 2030E, with custom ASICs contributing about 40%.

Figure 3: we forecast the AI power semis TAM growing to $8bn in 2027E and $15bn in 2030E

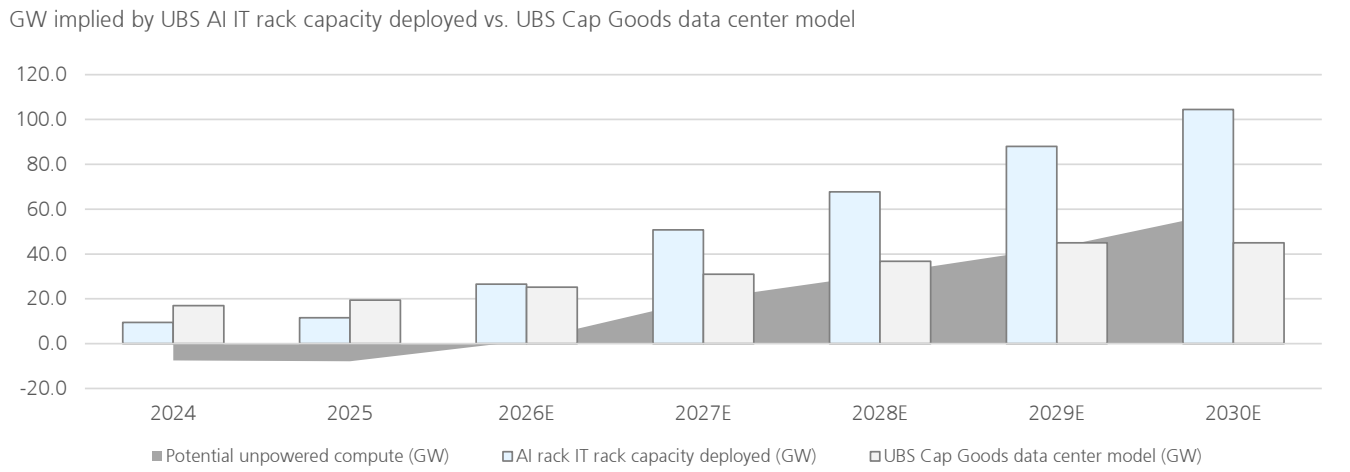


Source: UBS estimates

2. The capacity implied by the model is the question mark

The rack forecasts imply installed compute capacity additions rising from approximately 27GW in 2026E to 51GW in 2027E and 105GW in 2030E, or around 340GW cumulatively during 2026-30E. That is substantially above the approximately 30-50GW annual additions in UBS Cap Goods data center model.

Figure 4: growing gap between compute vs. infrastructure driven capacity addition estimates



Source: UBS estimates

Whilst not all of the compute is being deployed in the US, there does also appear to be a significant gap between UBS US Utilities team's energy supply additions forecasts. The team models roughly 120GW of capacity additions from 2024 to 2030E, though this could be pushed higher through less retirements, faster ramp of projects and behind the meter energy generation.

Figure 5: US energy grid capacity additions

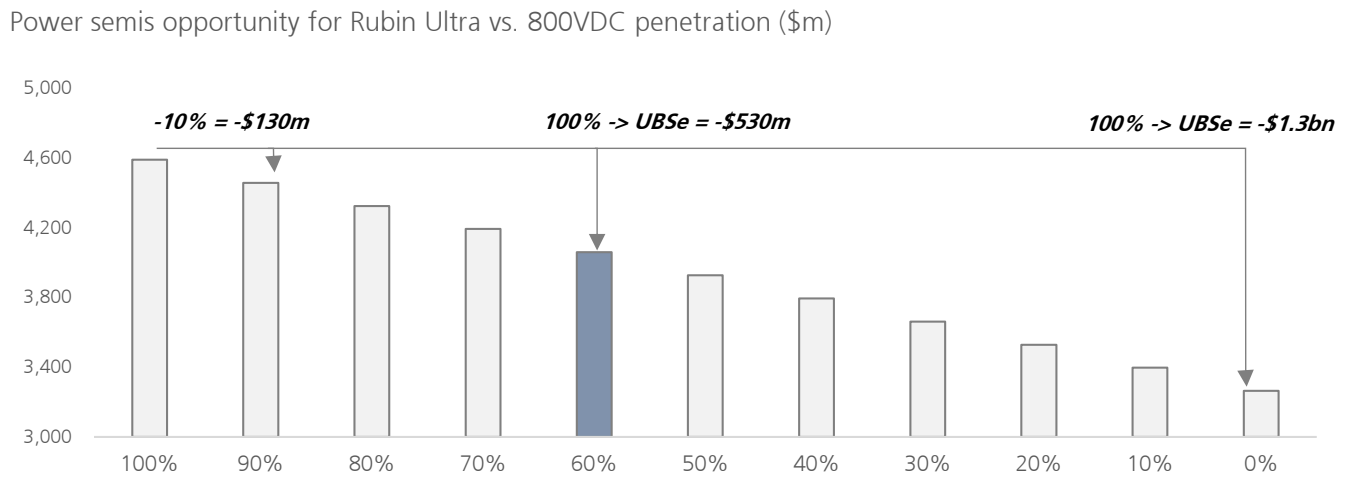
MW	2024A	2025E	2026E	2027E	2028E	2029E	2030E	26-30 Adds
Data Center Peak Demand								
ERCOT (Texas)	4,163	4,618	7,051	11,298	18,521	22,613	28,813	**21,762**
PJM	6,710	7,278	9,507	13,585	18,702	24,634	36,134	**26,627**
Data Center Demand MW	**46,226**	**51,602**	**65,152**	**80,348**	**99,559**	**116,454**	**141,025**	**75,874**
Peak Demand MW	**767,029**	**799,863**	**822,473**	**847,400**	**880,101**	**912,032**	**953,097**	**130,624**
Peak Demand Growth %		4.3%	2.8%	3.0%	3.9%	3.6%	4.5%	
Data Center % Growth		16%	60%	61%	59%	53%	60%	**58%**
Supply	**970,490**	**990,075**	**1,017,822**	**1,024,116**	**1,031,131**	**1,048,836**	**1,092,421**	**74,599**
Supply Growth %		2.0%	2.8%	0.6%	0.7%	1.7%	4.2%	
Reserve Margin	**26.5%**	**23.8%**	**23.8%**	**20.9%**	**17.2%**	**15.0%**	**14.6%**	

Source: UBS estimates

3. 800VDC is timing has a impact on the power semis TAM

A slower transition to 800VDC sidecars could reduce the TAM by approximately $0.5-1.0bn. We estimate Rubin Ultra power semis content is c$153m/GW with a sidecar versus $108m/GW with integrated power supply. Rising rack power still makes higher-voltage delivery increasingly necessary over time, but intermediate 50V sidecars may extend the transition.

Figure 6: we estimate a delay of 800VDC could lead to >$1bn impact on the power semis TAM



Source: UBS

4. Vertical power content uplift is likely closer to 2-3x vs 3-7x discussed with investors

As xPU power rises, lateral power delivery becomes constrained by both losses and board area. Vertical power delivery becomes critical and the headline content uplift of 3-7x vs. lateral discrete solutions is a very attractive story. We estimate the uplift at the total GPU-board level is closer to 2x because fewer devices are needed.

Figure 7: discrete vs. module for GB300 equivalent board level uplift

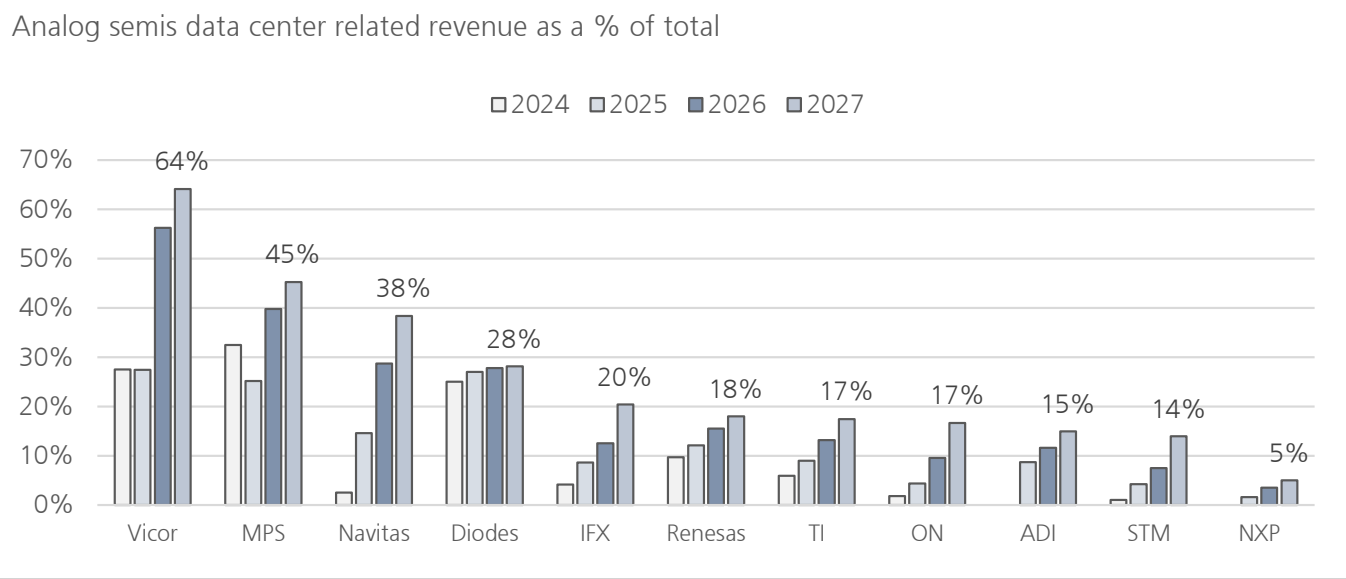
	Discrete	Module	Module v2	Module v3
Current rating (A)	90	140	160	280
Phases	Discrete	Dual	Dual	Quad
Discrete/module ASP ($)	2	6	7	14
Discrete/module #	24	15	14	8
Total VRM BoM per GPU ($)	48	90	98	112

Source: UBS, company data

The lay of the land

As demand for compute and AI data center capacity booms, the segment is becoming increasingly important for analog semis. For the large, diversified companies, data center has grown from <5% revenues before 2024 and is approaching c15% in 2026 and towards 20% by 2027, with most companies expecting it to be a significant growth driver going forward.

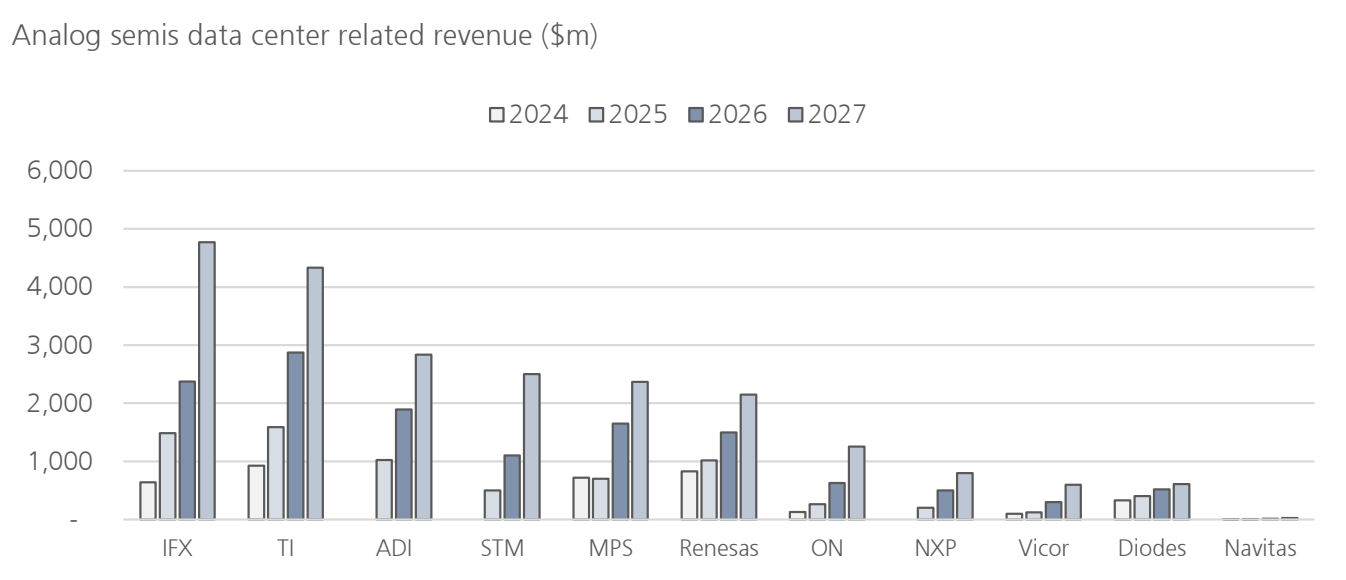
Figure 8: DC revenue approaching 15% of group for diversified analog names



Source: Company data, VA consensus, UBS estimates; Navitas, Vicor, MPI, Diodes not covered

By 2027, we expect Infineon to be the leader in the segment, with TI shortly behind, generating $2.4bn and $2.2bn in revenue respectively. Other players like Analog Devices and Renesas have long had a presence in the segment, while ON Semi and STMicro are now entering the market, with offerings specifically targeted at AI data centers.

Figure 9: Infineon is the leader in the segment, just ahead of TI



Source: Company data, VA consensus, UBS estimates; Navitas, Vicor, MPI, Diodes not covered

1. Updated BoM and TAM models

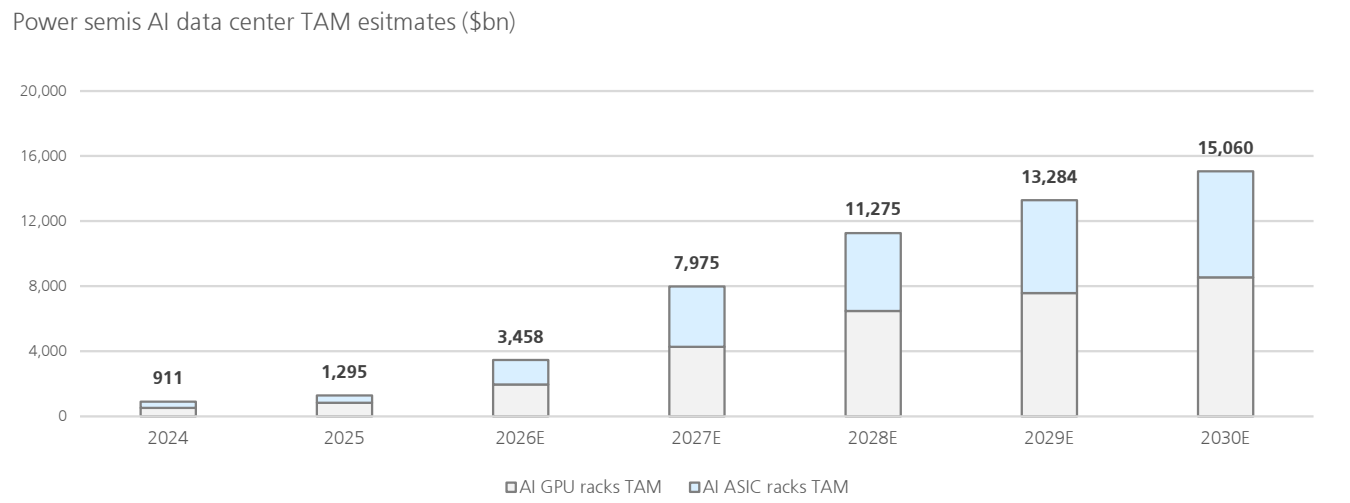
We make several changes to our BoM model and AI data center TAM model, which ultimately drives our 2030E TAM to c$15bn.

Key changes (*prev.*)

- **Expanded BoM model** - we expand our granular BoM model to include AMD, Google, Amazon, Meta, Microsoft, Intel and Tesla server variants, as well as Nvidia (*prev. just Nvidia*)
- **Refined component estimates** - we align component estimates with UBS Taiwan Hardware team's server model. For example we now assume 24 PSU units with a higher power rating for VR200 rack (*prev. PSU units aligned with Infineon estimates for GB200/300*)
- **Refined ASP estimates** - we align ASP estimates with details from UBS Taiwan Hardware server model. For example VR200 PSUs are assumed to be c18kW or c$350 at $22 per kW (*prev. back calculated from PSU unit volumes, rack power + 2N redundancy*)
- **New xPU board ASP estimates** - we now model xPU board at the individual xPU and CPU VRM level, and adjust for discrete, module and vertical architectures (*prev. assumed $/kW at the GPU board level*)

Our new assumptions and inputs drive a TAM of $15bn by 2030E. By 2030 we expect c40% of power semis AI data center revenue to be driven by products used in custom ASIC racks. The majority of this uplift is driven by higher rack forecasts from UBS Taiwan Hardware team as well as higher power semis content estimates for custom ASICs racks.

Figure 10: we forecast the AI power semis TAM growing to $8bn in 2027E and $15bn in 2030E



Source: UBS estimates

UPDATED BOM ESTIMATES

Figure 11: GPU power semis BoM estimates

	Nvidia							AMD		
GPU	Hopper	Blackwell	Blackwell Ultra	Rubin	Rubin Ultra	Rubin Ultra (PSC)	Feynman	MI300	MI400	MI500
Year of introduction	H2'23	H2'24	H2'25	H2'26	H2'27	H2'27	H2'28	Dec-23	H2'26	H2'27
GPU/xPU per rack	32	72	72	72	72	72	144	32	72	256
Die per GPU/xPU	1	2	2	2	2	2	4	8	8	8
Power per die (W)	700	650	700	1,150	1,380	1,380	1,500	94	288	316
Power per GPU/xPU (W)	700	1,300	1,400	2,300	2,600	2,600	6,000	750	2,300	2,530
Power per rack (kW)	45	120	140	227	250	250	1,000	40	235	1,000
Rack architecture	HGX	Oberon	Oberon	Oberon	Oberon	Oberon	TBC	HGX	Helios	Mega Pod
Power delivery	1-phase AC	3-phase AC	3-phase AC	3-phase AC	3-phase AC	Power sidecar	Power sidecar	3-Phase	3-Phase	Power sidecar
IT rack to power rack ratio						2:1	1:1			1:1
# of components per rack										
HV IBC						18	30			18
PSU	24	36	48	24	24	24	42	12	30	45
BBU						14	26			30
PCS						5	10			10
Motherboard	4							4		
GPU/xPU board	32	36	36	36	36	36	72	32	36	128
GPU/xPU per board	1	2	2	2	2	2	2	1	2	2
Power delivery board	32	18	18	18	18	18	72	32	18	64
Smart Network Interface Cards		108	108	108	108	108	288		108	512
Top of rack switches		1	1	1	1				1	
Protection	100	300	300	300	300	600	600	100	300	600
Power semis $ per component										
HV IBC						242	634			845
PSU	66	121	121	403	403	403	990	121	264	990
BBU						234	574			574
PCS						350	861			861
Motherboard	55							55		
GPU/xPU board	50	118	146	419	473	473	1,092	34	419	460
Power delivery board	10	130	168	276	312	312	360	75	230	253
Smart Network Interface Cards		8	8	9	9	10	11		9	10
Top of rack switches		200	200	200	200				200	
Protection	1	3	3	3	3	10	10	1	3	10
$ TAM per rack										
HV IBC						4,348	19,008			15,206
PSU	1,584	4,356	5,808	9,662	9,662	9,662	41,580	1,452	7,920	44,550
BBU						3,269	14,929			17,226
PCS						1,750	8,607			8,607
Motherboard	220							220		
GPU/xPU board	1,613	4,259	5,242	15,070	17,035	17,035	78,624	1,092	15,070	58,939
Power delivery board	320	2,340	3,024	4,968	5,616	5,616	25,920	2,400	4,140	16,192
Smart Network Interface Cards		864	864	972	972	1,080	3,168		972	5,120
Top of rack switches		200	200	200	200				200	
Protection	100	900	900	900	900	6,000	6,000	100	900	6,000
Total $ content per rack	**3,837**	**12,919**	**16,038**	**31,772**	**34,386**	**48,761**	**197,836**	**5,264**	**29,202**	**171,840**
Phase 1	50%	52%	55%	46%	44%	51%	56%	73%	41%	59%
Phase 2	50%	48%	45%	54%	56%	49%	44%	27%	59%	41%
Power semis per GW										
$ content per GW	74,178,133	93,661,300	99,662,229	121,769,339	119,661,888	169,687,556	172,117,109	114,492,000	108,108,051	149,500,659
$ content per W	0.09	0.11	0.11	0.14	0.14	0.20	0.20	0.13	0.12	0.17

Source: Company data, UBS estimates

Figure 12: TPU and Tranium power semis BoM estimates

	Google						Amazon			
GPU	TPU v5e	TPU v6	TPU v7x	TPU 8i	TPU 8t	TPU 9	Trainium 1	Trainium 2	Trainium 3	Trainium 4
Year of introduction	H223	H224	H225	H226	H226	H227	H2'22	H2'24	H2'26	H2'27
GPU/xPU per rack	64	64	64	64	64	64	32	32	72	72
Die per GPU/xPU	1	1	2	2	1	2	1	2	2	2
Power per die (W)	200	300	550	600	800	1,000	350	250	500	800
Power per GPU/xPU (W)	200	300	1,100	1,200	800	2,000	350	500	1,000	1,600
Power per rack (kW)	28	40	90	115	125	190	27	39	155	250
Rack architecture	Torus Cube	Torus Cube	Torus Cube	Torus Cube	Torus Cube	Torus Cube	2D Torus	2D Torus	3D Torus	3D Torus
Power delivery	3-phase AC	3-phase AC	3-phase AC	3-phase AC	3-phase AC	Power sidecar	3-phase AC	3-phase AC	3-phase AC	Power sidecar
IT rack to power rack ratio						2:1				2:1
# of components per rack										
HV IBC						15				18
PSU	12	18	30	18	18	18	12	12	24	30
BBU	8	8	12	16	16	10	8	8	20	14
PCS						5				5
Motherboard										
GPU/xPU board	16	16	16	16	16	16	16	16	36	36
GPU/xPU per board	4	4	4	4	4	4	2	2	2	2
Power delivery board	16	16	16	16	16	16	16	16	18	18
Smart Network Interface Cards	48	48	48	48	48	48	48	48	108	108
Top of rack switches	1	1	1	1	1	1	1	1	1	1
Protection	100	100	300	300	300	600	100	100	300	600
Power semis $ per component										
HV IBC						198				211
PSU	66	66	121	264	264	403	66	121	264	264
BBU	38	38	70	153	153	234	38	70	153	153
PCS						350				230
Motherboard										
GPU/xPU board	73	109	400	437	291	728	32	46	182	291
Power delivery board	24	36	132	144	96	240	42	60	120	192
Smart Network Interface Cards	9	9	9	9	9	9	9	9	9	9
Top of rack switches	200	200	200	200	200	200	200	200	200	200
Protection	1	1	3	3	3	10	1	1	3	10
$ TAM per rack										
HV IBC						2,970				3,802
PSU	792	1,188	3,630	4,752	4,752	7,247	792	1,452	6,336	7,920
BBU	306	306	842	2,450	2,450	2,335	306	561	3,062	2,144
PCS						1,750				
Motherboard										
GPU/xPU board	1,165	1,747	6,406	6,989	4,659	11,648	510	728	6,552	10,483
Power delivery board	384	576	2,112	2,304	1,536	3,840	672	960	2,160	3,456
Smart Network Interface Cards	432	432	432	432	432	432	432	432	972	972
Top of rack switches	200	200	200	200	200	200	200	200	200	200
Protection	100	100	900	900	900	6,000	100	100	900	6,000
Total $ content per rack	**3,379**	**4,549**	**14,523**	**18,027**	**14,929**	**36,422**	**3,012**	**4,433**	**20,182**	**34,976**
Phase 1	44%	46%	45%	53%	59%	50%	59%	67%	57%	50%
Phase 2	56%	54%	55%	47%	41%	50%	41%	33%	43%	50%
Power semis per GW										
$ content per GW	104,991,600	98,950,320	140,384,747	136,376,056	103,906,675	166,773,872	97,048,178	98,899,815	113,281,858	121,718,150
$ content per W	0.12	0.11	0.16	0.16	0.12	0.19	0.11	0.11	0.13	0.14

Source: Company data, UBS estimates

Figure 13: Meta, Microsoft and other custom ASIC power semis BoM estimates

	Meta						Microsoft			Tesla	Intel	
GPU	MTIA 100	MTIA 200	MTIA 300	MTIA 400	MTIA 450	MTIA 500	Maia 100	Maia 200	Maia 300	Dojo D1	Gaudi 2	Gaudi 3
Year of introduction	2023	2024	2026	2026	2027	2027	2024	2026	2026	2027	2027	2027
GPU/xPU per rack	12	72	72	72	72	72	64	64	64	6	64	64
Die per GPU/xPU	1	1	1	2	2	2	1	1	1	25	1	2
Power per die (W)	25	90	800	600	700	850	700	750	1,250	400	600	450
Power per GPU/xPU (W)	25	90	800	1,200	1,400	1,700	700	750	1,250	10,000	600	900
Power per rack (kW)	1	18	75	110	130	170	50	75	125	220	55	80
Rack architecture	MTIA	MTIA	MTIA	MTIA	MTIA	MTIA	Maia	Maia	Maia	Dojo D1	Gaudi	Gaudi
Power delivery	3-phase AC	3-phase AC	3-phase AC	3-phase AC	3-phase AC	Power sidecar	3-phase AC	3-phase AC	3-phase AC	3-phase AC	3-phase AC	3-phase AC
IT rack to power rack ratio						1.5:1						
# of components per rack												
HV IBC						10						
PSU	1	12	24	12	18	20	12	18	18	48	18	24
BBU			4	12	16	16	4	4	16	8	8	8
PCS						7						
Motherboard												
GPU/xPU board	12	36	36	36	36	36	16	16	16	2	8	8
GPU/xPU per board	1	2	2	2	2	2	4	4	4	6	8	8
Power delivery board	6	36	36	36	36	36	36	36	36	12	36	36
Smart Network Interface Cards	36	108	108	108	108	108	48	48	48	6	24	24
Top of rack switches	1	1	1	1	1	1	1	1	1	1	1	1
Protection	100	100	100	300	300	600	100	100	100	100	100	100
Power semis $ per component												
HV IBC						238						
PSU	44	44	121	264	264	264	121	264	121	121	121	121
BBU			70	153	153	153	70	153	70	70	70	70
PCS						230						
Motherboard												
GPU/xPU board	1	8	73	109	127	155	255	273	455	2,730	437	655
Power delivery board	3	11	96	144	168	204	84	90	150	1,200	72	108
Smart Network Interface Cards		9	9	9	9	9	9	9	9	9	9	9
Top of rack switches		200	200	200	200	200	200	200	200	200	200	200
Protection	1	1	1	3	3	10	1	1	1	1	1	1
$ TAM per rack												
HV IBC						2,376						
PSU	44	528	2,904	3,168	4,752	5,280	1,452	4,752	2,178	5,808	2,178	2,904
BBU			281	1,837	2,450	2,450	281	612	1,123	561	561	561
PCS						1,530						
Motherboard												
GPU/xPU board	14	295	2,621	3,931	4,586	5,569	4,077	4,368	7,280	5,460	3,494	5,242
Power delivery board	18	389	3,456	5,184	6,048	7,344	3,024	3,240	5,400	14,400	2,592	3,888
Smart Network Interface Cards	-	972	972	972	972	972	432	432	432	54	216	216
Top of rack switches	-	200	200	200	200	200	200	200	200	200	200	200
Protection	100	100	100	900	900	6,000	100	100	100	100	100	100
Total $ content per rack	**176**	**2,484**	**10,534**	**16,193**	**19,908**	**29,345**	**9,566**	**13,704**	**16,713**	**26,583**	**9,342**	**13,111**
Phase 1	35%	37%	63%	63%	67%	57%	50%	63%	52%	78%	57%	56%
Phase 2	65%	63%	37%	37%	33%	43%	50%	37%	48%	22%	43%	44%
Power semis per GW												
$ content per GW	152,815,500	120,042,600	122,188,832	128,069,062	133,232,603	150,178,247	166,440,048	158,971,968	116,321,645	105,125,422	147,770,924	142,582,560
$ content per W	0.18	0.14	0.14	0.15	0.15	0.17	0.19	0.18	0.13	0.12	0.17	0.16

Source: Company data, UBS estimates

Updated TAM estimates

To update our TAM estimates we now use AI GPU and ASIC unit and rack estimates from UBS Taiwan Hardware team, which are informed by UBS HBM, GPU and custom ASIC models. Rack unit estimates now extend to 2030 vs. 2028 previously. The full model breaks down these rack forecasts by xPU generation.

Figure 14: UBS AI server rack estimates

AI GPU and ASIC Rack Forecast (UBSe)	2024	2025	2026E	2027E	2028E	2029E	2030E
NVIDIA NVL72+	18	27,034	77,506	95,098	82,000	34,429	27,000
NVIDIA Other (HGX/LPU/CPU)	130,744	63,000	66,303	72,941	83,188	100,016	111,868
AMD	6,164	6,466	8,450	13,765	11,338	9,000	9,168
Total AI GPU Racks	136,926	96,500	152,260	181,803	176,526	143,444	148,035
AWS	39,583	43,935	63,026	82,083	54,952	42,361	44,792
Google	34,820	39,672	59,155	133,465	122,266	100,156	105,113
Meta	5,631	6,448	3,465	5,439	7,448	9,340	10,903
Microsoft	474	553	611	832	965	1,192	1,440
Intel	9,804	7,447	4,225	2,869	1,844	859	391
Tesla	1,451	1,025	781	718	986	1,415	1,917
Other ASICs	8,910	12,837	13,094	21,089	31,722	43,590	56,862
Total AI ASIC Racks	100,674	111,918	144,357	246,496	220,182	198,914	221,416
Total Racks	**237,600**	**208,418**	**296,616**	**428,299**	**396,708**	**342,359**	**369,452**
y-o-y growth		*-12%*	*42%*	*44%*	*-7%*	*-14%*	*8%*

Source: UBS estimates

We overlay our BoM estimates from above onto the rack estimates, breaking down power semis content by xPU generation and aggregating at the company level below. We forecast the power semis TAM rising from $3.5bn in 2026E to $15bn in 2030E.

Figure 15: power semis TAM estimates

AI Power Semis TAM forecast ($bn)	2024	2025	2026E	2027E	2028E	2029E	2030E
NVIDIA NVL72+	0	369	1,236	2,705	3,724	4,570	5,342
NVIDIA Other (HGX/LPU/CPU)	490	429	612	1,005	1,666	1,645	1,631
AMD	32	34	104	573	1,100	1,363	1,577
AI GPU racks TAM	522	833	1,952	4,282	6,490	7,578	8,550
AWS	119	137	484	824	1,179	1,457	1,603
Google	108	159	854	2,493	3,045	3,460	3,875
Meta	1	3	17	92	169	249	326
Microsoft	3	4	7	12	15	20	24
Intel	73	55	31	29	12	8	4
Tesla	27	14	14	14	22	32	44
Other	58	91	99	229	343	480	634
AI ASIC racks TAM	390	463	1,506	3,692	4,785	5,706	6,510
Total TAM	**911**	**1,295**	**3,458**	**7,975**	**11,275**	**13,284**	**15,060**
y-o-y growth		*42%*	*167%*	*131%*	*41%*	*18%*	*13%*

Source: UBS estimates

Translating this into GW capacity additions implied

We do think it is important to contextualise these numbers in terms of the GW of capacity additions they imply. The rack forecasts in 2026E imply c.27GW, increasing to 51GW in 2027E and 105GW by 2030E for a cumulative 340GW from 2026E to 2030E.

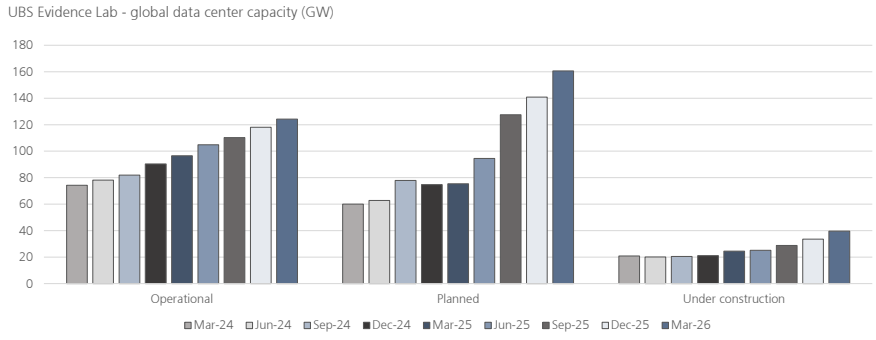
Figure 16: GW capacity installed implied by rack forecasts

GW of Racks Deployed	2024	2025	2026E	2027E	2028E	2029E	2030E
NVIDIA NVL72+	0.0	3.4	10.8	19.8	22.2	27.6	32.4
NVIDIA (HGX/LPU/CPU)	5.9	4.1	5.9	8.0	10.5	14.5	18.1
AMD	0.2	0.3	0.7	3.3	6.0	7.7	8.9
AI GPUs IT rack capacity deployed (GW)	**6.1**	**7.7**	**17.3**	**31.1**	**38.7**	**49.7**	**59.4**
AWS	1.1	1.2	3.5	6.5	10.0	13.0	14.8
Google	1.0	1.2	4.2	9.9	13.9	18.0	20.8
Meta	0.0	0.0	0.1	0.6	1.0	1.4	1.8
Microsoft	0.0	0.0	0.0	0.1	0.1	0.1	0.2
Tesla	0.2	0.1	0.1	0.1	0.1	0.2	0.2
Intel	0.5	0.4	0.3	0.2	0.1	0.1	0.0
Other ASICs	0.6	0.8	1.1	2.3	3.8	5.4	7.3
AI ASIC IT rack capacity deployed (GW)	**3.4**	**3.8**	**9.2**	**19.6**	**29.0**	**38.2**	**45.1**
AI rack IT rack capacity deployed (GW)	**9.5**	**11.5**	**26.5**	**50.8**	**67.7**	**87.9**	**104.47**
Facility GW required at 1.18 PUE	11.2	13.6	31.3	59.9	79.9	103.8	123.3
Facility GW required at 1.30 PUE	12.3	15.0	34.4	66.0	88.0	114.3	135.8

Source: UBS estimates

Below we show data from a UBS Evidence Lab data set which provides a high-level overview of the global data center market. The data set identifies 124GW of operational DC capacity, 161GW of planned capacity and 40GW currently under construction.

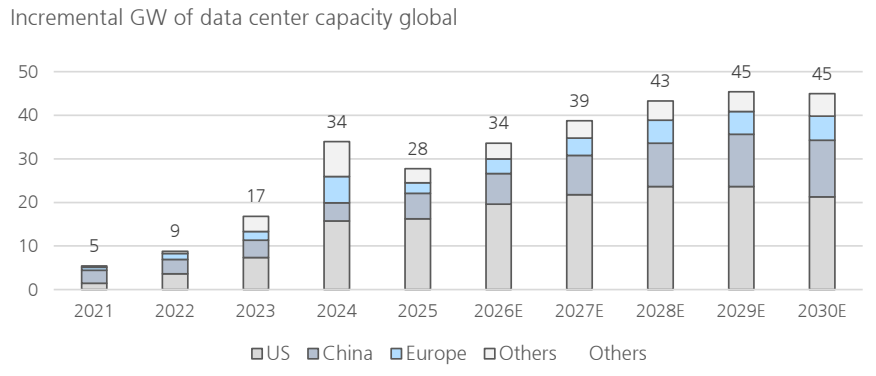
Figure 17: total planned and under construction capacity globally sits at c200GW



Source: UBS Evidence Lab (> Access Dataset)

We also show below UBS Global Capital Goods team Data Center equipment market model, published on most recently here, which implies roughly 30-50GW annual capacity additions from 2027-2030E.

Figure 18: UBS Capital Goods team assumes stabilising GW capacity additions

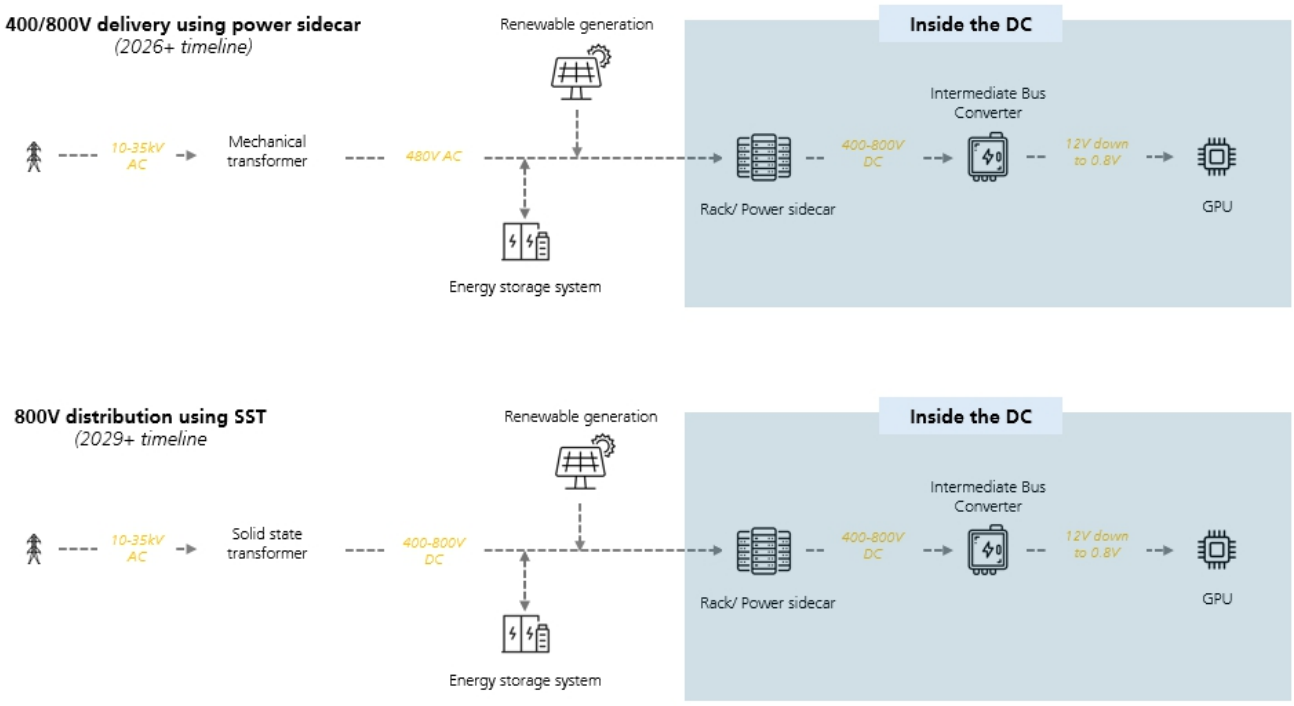


Source: UBS research

2. What is the impact of an 800VDC delay

800VDC power delivery had initially been expected to ramp in 2027 with the Kyber, or power sidecar, rack architecture however there have been reports that this has been pushed to 2028 or later. Some hyperscalers also now appear to be prioritising 50VDC sidecars as an interim step. We set out the differences between 800V power delivery and 800V power distribution below, as the two are often conflated.

Figure 19: key differences between 800VDC power distribution and delivery



Source: UBS, company data

There are several known challenges around the shift to 800VDC and potential reasons for the delay which we highlight below:

- **Safety/certification bottlenecks:** Moving to 800VDC sidecar delivery introduces a more complex high-voltage DC safety regime, which was flagged at this year's OCP APAC Summit. Protection devices and industrial safety hardware still need certification from bodies such as UL and other safety authorities.
- **Ecosystem not yet fully standardised:** Supplier readiness, OCP alignment, protection, connectors, busways and serviceability standards all need to mature before broad deployment.
- **50V power sidecar:** Hyperscalers appear to be exploring intermediate sidecar steps such as 50V power sidecars.

It's important to note that this is very much a delay rather than cancellation of the 800VDC transition. Assuming compute racks continue to scale in terms of power towards and above 1MW racks, 800V power deliver becomes a necessity because of the thermal and copper issues associated with the extremely high current required to move electricity at lower voltages. The industry is likely to focus on interim solutions before 800VDC is more widely adopted.

How do delays impact power semis revenues?

Our TAM numbers imply 60% penetration in 2028E of 800VDC delivered in power sidecars for Rubin Ultra, which delivers $3bn of power semis revenue that year, of which $2bn is power sidecar. Our estimates assume $34k power semis content for an Oberon VR300 rack vs. $49k for a VR300 power sidecar. We run simple scenarios on the impact of power sidecar mix on our VR300 power semis estimates, with further delays driving c20% downside to our estimates and accelerated adoption driving 10-20% upside.

Figure 20: power sidecar mix sensitivities on our VR300 power semis estimates ($m)

				Delta vs. UBSe		
Power sidecar	2027E	2028E	2029E	2027E	2028E	2029E
100%	77	3,413	1,023	17%	13%	10%
90%	75	3,313	993	14%	10%	6%
80%	72	3,212	963	10%	7%	3%
70%	70	3,111	932	7%	3%	0%
60%	68	3,011	902	3%	0%	-3%
50%	66	2,910	872	0%	-3%	-6%
40%	63	2,809	842	-3%	-7%	-10%
30%	61	2,709	812	-7%	-10%	-13%
20%	59	2,608	782	-10%	-13%	-16%
10%	56	2,508	751	-14%	-17%	-19%
0%	54	2,407	721	-17%	-20%	-23%

Source: UBS

If we keep GW of capacity installed constant and use our BoM implied power semis content per GW of $119m/GW for VR300 Oberon and $170m for VR300 power side car we arrived at the below.

Figure 21: power side car mix impact at constant GW ($m)

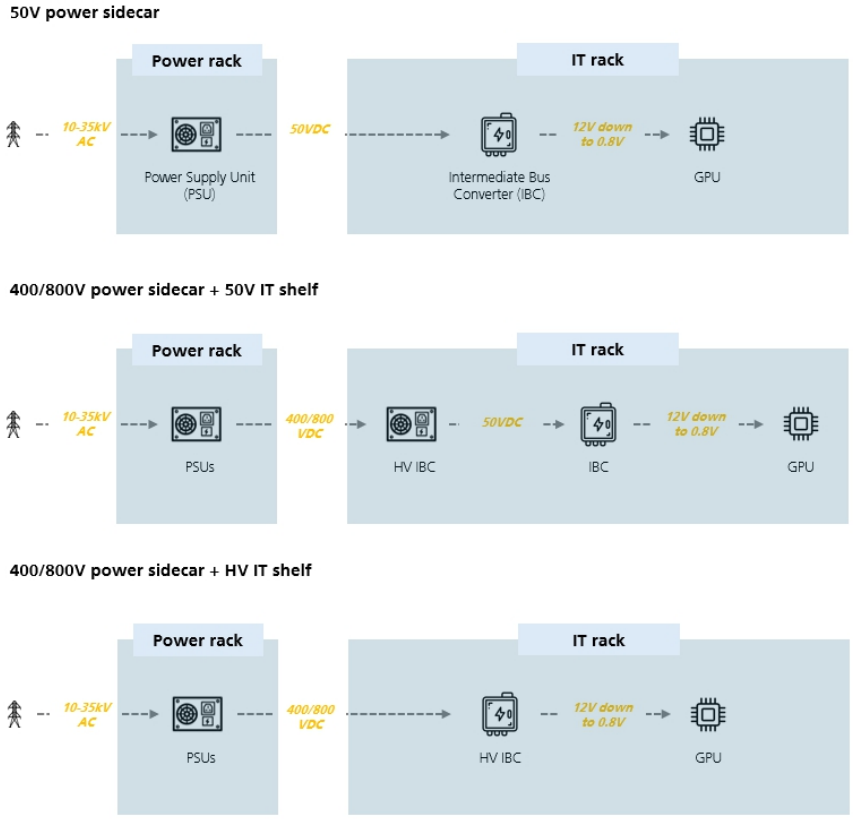
Power sidecar	GW of VR300 shipped				
	15	20	25	30	35
100%	2,545	3,394	4,242	5,091	5,939
90%	2,470	3,294	4,117	4,941	5,764
80%	2,395	3,194	3,992	4,790	5,589
70%	2,320	3,094	3,867	4,640	5,414
60%	2,245	2,994	3,742	4,490	5,239
50%	2,170	2,893	3,617	4,340	5,064
40%	2,095	2,793	3,492	4,190	4,889
30%	2,020	2,693	3,367	4,040	4,713
20%	1,945	2,593	3,242	3,890	4,538
10%	1,870	2,493	3,117	3,740	4,363
0%	1,795	2,393	2,992	3,590	4,188

Source: UBS

What is the likely progression towards 800V sidecars

Power sidecars or power racks can and are being introduced today, even when IT racks have not shifted to high voltage input. Infineon for example, see two potential steps in power delivery architecture before moving to a 400/800V IT rack. The first involves a 50V power rack, where the existing PSUs that convert AC to 50VDC are shifted from the IT rack to the power rack. The second involves using high voltage PSUs in the power rack but keeping some PSUs in the IT rack that step down the high voltage to 50V DC, while the IBC and VRM remain unchanged from current architectures.

Figure 22: three potential steps in the journey to 800V sidecars



Source: UBS, company data

What are the key components?

In terms of the shift to 400/800VDC, the main components that are impacted are the PSUs and the IBC, while the direct power delivery to the GPU remain unchanged. There has been a lot of investor focus on the adoption of SiC and GaN during the transition, but we believe it is more nuanced, it is not as simple as higher voltage = higher SiC or GaN content.

i. PSU

The job of the PSU in the data center is to convert the incoming electricity from alternating current to direct current and to step down the voltage from the typically higher input voltage to typically lower output voltage. From a power semis perspective the two most important parts of the PSU are the power factor correction (PFC) where current is converted and the isolated DC/DC stage, where voltage is stepped down. We summarise each step below in simplified topologies.

a. Power Factor Current AC to DC - limited changes as you move to 800VDC

While PFC architectures continue to evolve, the underlying switch topology remains broadly similar, comprising a slow leg for current steering and a fast leg for high-frequency power conversion. Given that current-generation PSUs already employ SiC extensively in these functions, we see limited scope for a meaningful increase in SiC content per kW from architecture migration alone.

Figure 23: PFC is already a SiC heavy domain and sees little change with 800VDC

	50V	+-400V	800V
Fast leg	650V SiC	650V SiC	650V SiC
Slow leg	650V Si	650V Si	650V Si

Source: UBS

b. DC/DC stage - more SiC but less GaN

The DC/DC stage comprises of two groups of switches either side of a transformer. The first set of switches briefly convert the current back to AC called the primary side, the voltage is stepped down by a transformer, and the second set of switches, the secondary side, convert the current back to DC to be outputted onwards to the IBC. SiC and GaN are already widely used on the primary side, while the second side is typically Si or GaN at 50V. At 400 or 800V the secondary side moves to SiC from GaN.

Figure 24: the DC/DC stage sees more SiC but less GaN at 800VDC

	50V	+-400V	800V
Primary side	650V SiC/GaN	650V SiC/GaN	1.2kV SiC / 650V SiC/GaN
Secondary side	MV Si/GaN	650V SiC/Si	1.2kV SiC

Source: UBS

Clearly through this transition the power requirements of each PSU are increasing from 5.5kW to 12kW and then to 18kW and beyond for, which drives higher content. The key change we see across the whole PSU from the shift to 400/800V is on the secondary DC-DC side, with a shift from a high number e.g. 20+ MV Si/GaN MOSFETs to a lower number of high-voltage SiC MOSFETs.

ii. Intermediate Bus Converter

The IBC sits inside the compute tray and converts the PSU output voltage to an intermediate bus voltage, typically 12V or 6V, before a final conversion stage delivers the ~0.8V required by the GPU. This intermediate step is necessary because a direct conversion from 48V to 0.8V would require an impractically high conversion ratio of ~60:1. Traditional IBCs share a similar architecture to the isolated DC-DC stage of a server PSU, employing primary-side switches, a high-frequency transformer and secondary-side rectification. By contrast, newer interleaved buck and switched-capacitor (xSC) architectures can remove the transformer altogether, shifting power conversion to a larger number of lower-voltage switching devices located closer to the processor.

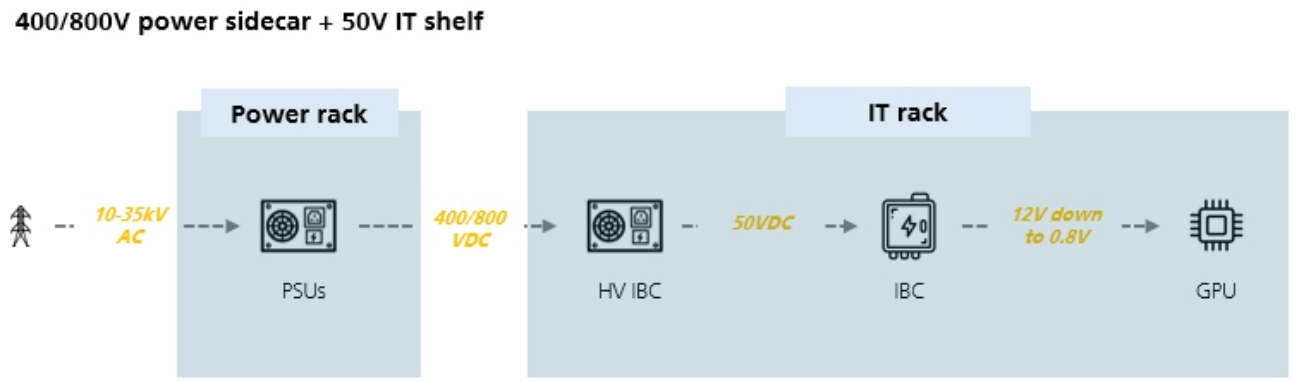
Figure 25: key changes with the shift to 800VDC

	50V -> 12V IBC	800V -> 50V IBC	800V -> 12V IBC
Primary side	MV Si	650V GaN	650V GaN
Secondary side	LV Si	MV GaN	LV Si

Source: UBS

Early 800 VDC architectures will likely remain 50 V inside the IT rack. A new rack-level HV IBC power shelf converts the sidecar's 800 VDC output to 50 V, rather than a straight conversion from 800V to 12V output to be delivered to the VRM. This represents one of the largest uplifts in content/GW.

Figure 26: the HV IBC is a key driver of content upside in initial 800VDC adoption



Source: UBS

Bringing it all together

The shift to 800VDC drives several changes that we have run through above. Bringing it together it's important to look at the several puts and takes across the different products. We show this below, assuming the same power requirement per rack. There is limited change in power semis content

Figure 27: there are several puts and takes with the shift to 800VDC

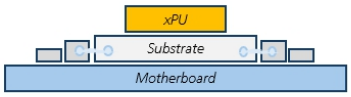
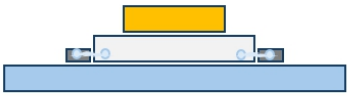
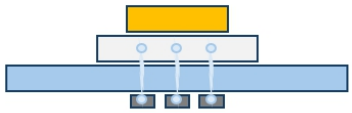
	Conversion	50V IT Rack	50V Sidecar	HV Sidecar	HV IT tray
PSU	400AC -> 50VDC	SiC/Si/GaN	SiC/Si/GaN		
	400AC -> 800VDC			SiC	SiC
IBC	800V -> 50V			GaN/Si	
	800V -> 12V				GaN/Si
	50V -> 12V	Si/GaN	Si/GaN	Si/GaN	
VRM	12V -> 0.8V	Si	Si	Si	Si
Power semis BOM $		**$**	**$**	**$$**	**$$**

Source: UBS

3. Vertical power delivery / Integrated voltage regulators

In this section we explore the second power stage or voltage regulator module (VRM), where voltage is dropped down to the GPU/xPUs core voltage. The VRM has been a key area of investor focus as it represents ~50% of the power semis opportunity and has an architecture driven content uplift story. There are broadly three power delivery architecture approaches currently in use: lateral power delivery achieved with discrete components, which is the lowest cost; lateral delivery achieved with a module which integrates the driver, MOSFETs and inductors into one module; and vertical delivery achieved with a module, where the VRM is placed on the underside of the compute board.

Figure 28: summary of the three main VRM approaches

	Lateral delivery – Discrete	Lateral delivery – Module	Vertical delivery – Module
			
Power delivery	Lateral	Lateral	Vertical
Placement	On motherboard right next to xPU	On motherboard right next to xPU	Flipside of motherboard underneath xPU
Integration	Driver + MOSFETs integrated	Driver + MOSFETs + inductors + capacitors integrated	Driver + MOSFETs + inductors + capacitors integrated
Board space for VRM	c.3000mm2	40% less	Nothing on GPU side of board
Performance	10% of input power lost	8% of input power lost	~1% of input power lost

Source: UBS, company data

Why is the shift to vertical inevitable?

There are two key reasons why the shift to vertical power delivery is inevitable, assuming xPU power continues to scale. Firstly because of power loss which increases as xPU power scales and secondly because of board space constraints, both of which we explore in detail below.

i. Reducing power losses

The final power conversion stage, where 12V is stepped down to ~0.8V for the xPU, is one of the largest sources of power loss in the data center because it carries the highest current anywhere in the power delivery chain. This is because of the relationship between power, current and voltage, which we frame below.

Figure 29: lower voltage needs of xPU drive high current requirements

Power (W)	=	*Voltage (V)*	*x*	*Current (A)*
Worked example				
GB200 GPU power of 1,300W	=	*VRM output voltage of 0.8V*	*x*	*Current of **1,625A***
VR300 GPU power of 2,600W	=	*PSU output voltage of 0.8V*	*x*	*Current of **3,250A***

Source: UBS

Higher current typically drives higher power losses but there are some offsets, which board designers and engineers can implement. To explore this it is important to understand the relationship between power losses and current, which we outline below.

There is limited flex on the current part of the equation, the industry is pushing higher power GPUs and GPUs physically can't handle higher voltage than c0.8V. This puts power distribution network resistance in focus.

Figure 30: power loss scale with current and resistance

Power Loss (W)	=	*Current (A) ²*	x	*Resistance (µΩ)*
Worked example				
GB200 power loss at VRM of c240W	=	*VRM current 1,625A ²*	x	*Power distribution resistance of 90µΩ*
VR300 power loss at VRM of c1,400W	=	*VRM current of 3,250A ²*	x	*Power distribution resistance of 90µΩ*

Source: UBS

There are broadly three ways to reduce resistance: 1) Shorten the current path - resistance increases the longer the electricity has to travel; 2) Increase conductivity of the path - use wider, thicker copper within the power stage, more power vias, often implemented within modules; 3) Optimise conversion - using more power phases, lower-RDS power stages/modules. Below we show how each VRM architecture choice impacts resistance.

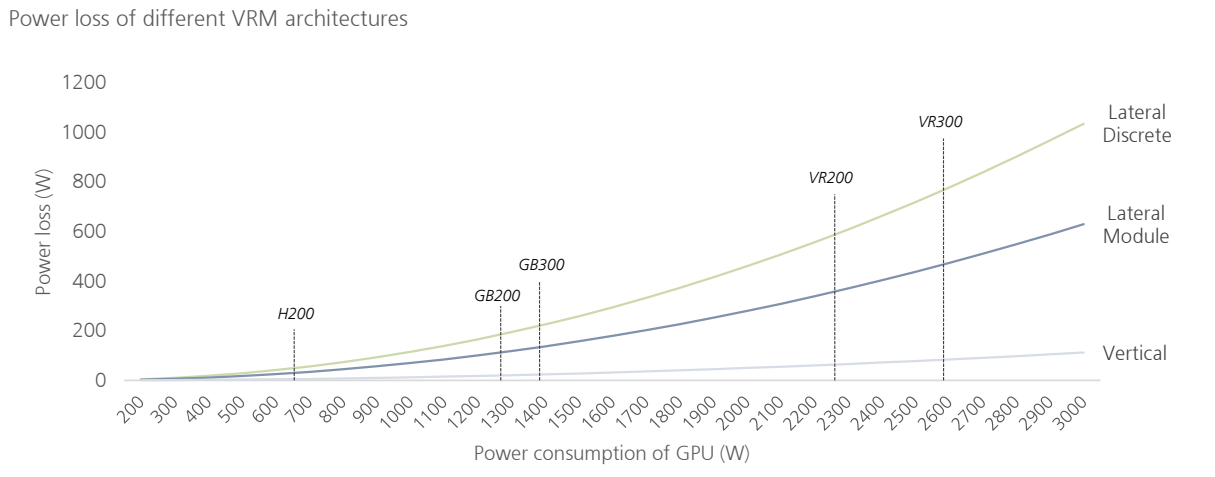
Figure 31: shortening the current path is the most impactful lever to pull to reduce resistance

	Lateral – Discrete	Lateral – Module	Vertical – Module	Substrate Embedded
Shorten current path				
Increase conductivity of path				
Optimise power conversion				
Lumped PDN resistance (µΩ)	*90-140*	*55-80*	*10-15*	*7-10*

Source: UBS, Infineon PDN figures

If we plot this to various levels of output power using the above power loss = A² x µΩ equation, we can visualise real world examples. A H200 using lateral discrete has power losses of c70W or around 10% of GPU power, while a GB300 using discrete has losses of 275W or c20% of GPU power, which is an unsustainable level off loss. This likely drove a shift from a discrete lateral solution to lateral module, which would point to a more manageable 13% of GPU power lost.

Figure 32: power losses scale and make lateral discrete solutions unfeasible

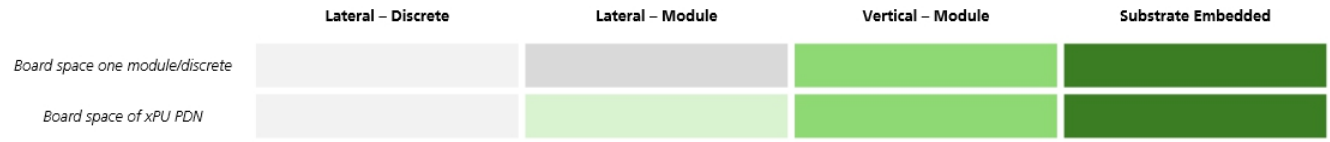


Source: UBS

ii. Optimising board space

The goal of compute board designers is to maximize the amount of compute that can be deployed within a given board and rack footprint. As a result, there is constant pressure to minimize the area occupied by supporting components, including power delivery, networking and cooling hardware, in order to allocate as much space as possible to xPUs and HBM, which are the primary drivers of system performance.

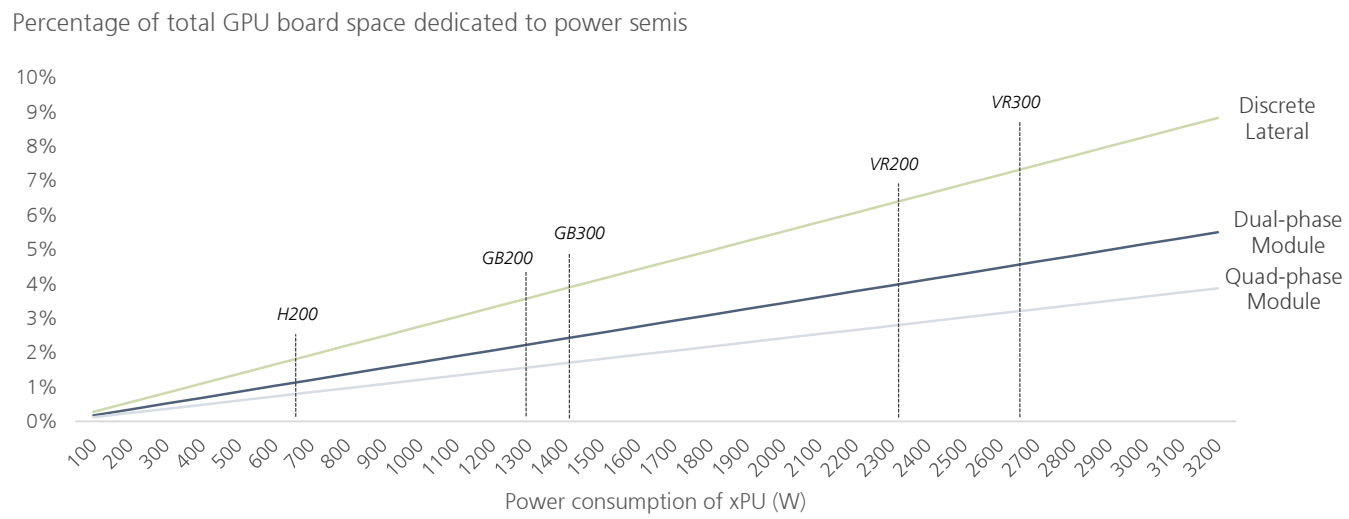
Figure 33: lateral modules reduces board space vs. discrete solutions



Source: UBS

Below we show how VRM architecture impacts board space dedicated to the power delivery network on the GPU board. We assume a GPU board around 66cm², in line with GB200 and calculate what percentage of the board is taken up by final power stage components. Modules drive lower board space in two main ways: 1) Integrating capacitors and inductors into the module; 2) Integrating multiple power stages. Modules tend to be larger than a discrete power stage but much smaller than all the individual components that are now integrated, and they also tend to deliver more power per mm, especially newer modules.

Figure 34: board space savings can occur before the shift to vertical



Source: UBS; we assume discrete represents 45% of board space of components that are integrated into modules e.g. capacitors, inductors

Because of the high current of power delivery to the xPU, components need to be placed in very close proximity, meaning they compete for critical board space. If we assume 30% of all board space is "critical" or close in proximity to the xPU, then we estimate that power components will reach an unsustainable proportion of this space, which will likely necessitate a shift to backside mounted vertical power delivery, freeing up critical board space.

Figure 35: power components will likely reach an unsustainable % of critical board space

	H200	GB200	GB300	VR200	VR300
Discrete	8%	15%	16%	26%	30%
Dual-phase	5%	9%	10%	16%	19%
Quad-phase	4%	7%	7%	12%	13%

Source: UBS; critical assumed at 30% of total board space

Why is everyone not already using vertical VRMs?

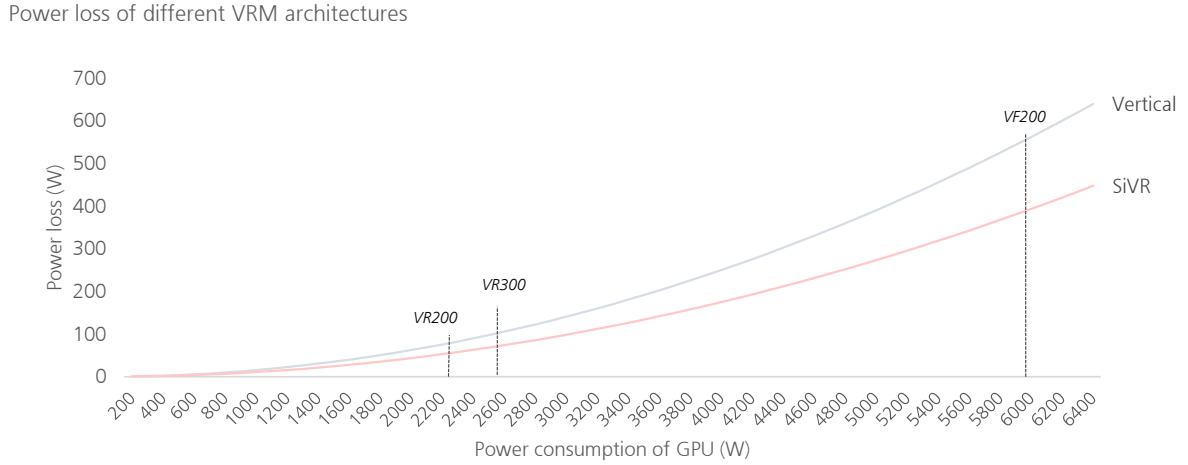
The obvious question here is, if the benefits of vertical architectures are so clear, then why hasn't it been widely adopted. We believe the TPU was early to adopt vertical, likely to ensure its compute board architecture was scalable to significantly higher xPU power, however others have been slower to adopt, with VR200 for example still using lateral. We see three main risks around adoption of vertical power delivery, which have lead to xPU suppliers stretching the capabilities of lateral architectures:

1. **Thermal risk** - Vertical VRMs are concentrated directly beneath or behind the GPU, creating a high heat density hotspot that often requires backside cooling. This can be seen in Google TPUv7 liquid-cooling cold plates on the backside of the PCB
2. **Packaging complexity** - delivering hundreds of amps through the package/ substrate requires new interconnect structures, power vias, copper pillars and advanced packaging technologies.
3. **Packaging yield and cost** - linked to point number two, backside power architectures are significantly more complex to assemble and test than placing VRMs on a PCB. Lateral VRMs use a mature supply chain and design ecosystem, while vertical power requires new module formats, cooling solutions and board designs.

What comes next?

Vertical power delivery also has its limits informed by power losses and board space constraints. The industry is already exploring alternatives and appear to be focussed on substrate embedded power delivery (SiVR), where the power components are embedded into the substrate that the xPU is packaged onto. Substrate embedded is unlikely to be adopted for Rubin or Rubin Ultra as power losses are manageable using vertical solutions. For Feynman at 6000W power per GPU, losses approach an 10% of GPU power, which would cause significant heat losses. A shift to substrate embedded would reduce this by 30%.

Figure 36: power losses can be reduced further by shifting to substrate embedded VRM



Source: UBS

Infineon and Analog Devices have recently made acquisitions in the space, with C2i and Empower respectively, and several module suppliers are also looking at substrate embedded solutions.

Figure 37: many companies are exploring substrate embedded solutions

	IDM		Module		
Company	**Infineon**	**Analog Devices**	**TDK**	**Murata**	**Saras**
Product/relevant acquisition	*C2i*	*Empower*	µPOL/SESUB	iPaS	STILE
Focus	Voltage regulation			Passive components	

Source: Company data, UBS

What is the content uplift from lateral to vertical

The headline content uplift from discrete to vertical VRM has been flagged as between 3-7x, however we believe it is important to delve further into the drivers and the implications of this: i) We see the majority of this uplift as driven by the shift from discrete to a module solution, rather than purely lateral to vertical; ii) A 3-5x uplift at the discrete to module level does not necessarily deliver similar uplift in total VRM content.

i. From discrete to module drives a large uplift

In a discrete lateral architecture, there are multiple power stages with separate inductors and capacitors located next to the processor. When a supplier integrates these inductors and capacitors into a module with multiple power stages, the module supplier is able to capture a larger share of the BoM. The extra board space and improved energy efficiency also allow the module supplier to capture more value. We break down this uplift below.

How we arrive at our numbers

- **Base product:** we use a 90A discrete smart power stage (SPS) at a starting point
- **Base ASP:** we assume an ASP for the discrete lateral SPS of $2, in line with distributor pricing
- **Module uplift:** we assume an uplift of 3x for a 140A dual-phase module, 2x of which is from the extra power stage and the remainder from capturing additional inductor and capacitor content
- Power loss and board space calculated as above

Figure 38: discrete vs. module for GB300 equivalent ASP uplift

	Discrete	Module	Module v2	Module v3
Current rating (A)	90	140	160	280
Phases	Discrete	Dual	Dual	Quad
Discrete/module ASP ($)	2	6	7	14
Power loss (W)	276	168	160	144
% of total power lost	20%	12%	11%	10%
% of critical board space	14%	14%	9%	7%

Source: UBS, company data; critical board space assumed at 30% of total

ii. Uplift at the board level is more moderate

Whilst uplift at the module level is significant, it is important to look at the impact on the GPU board or per GPU basis. With the uplift in current rating from discrete to module, less components are required to deliver the same level of power. Taking a GB300 as an example, each GPU requires 1,400W of power, delivered at 0.8V, which requires a current of 1,750A. If we assume 20% headroom, or 2,100A, then 22 discrete smart power stages are required vs. 14 dual-phase 140A modules or 7 quad-phase 280A modules. Overall this leads to a 2x of total VRM BoM per GPU compares with the 3-7x uplift at the component level.

How we arrive at our numbers

- **GPU power:** we model the below for a GB300 or equivalent GPU at 1,400W, or 1,750A. We model a 20% headroom, requiring 2,100A current capacity from the VRMs
- **Unit count:** we arrive at required units by dividing that required current with the current rating of the discrete/module, which gives us the number of units required

Figure 39: discrete vs. module for GB300 equivalent board level uplift

	Discrete	Module	Module v2	Module v3
Current rating (A)	90	140	160	280
Phases	Discrete	Dual	Dual	Quad
Discrete/module ASP ($)	2	6	7	14
Discrete/module #	24	15	14	8
Total VRM BoM per GPU ($)	48	90	98	112

Source: UBS, company data

How does this change our modelling?

We use the above work to adjust our modelling of the xPU board power semis content, details of which we outline below, as well as the impact on our xPU board BOM estimates.

Assumptions (*prev. assumptions)*

- **Model granularity:** we model xPU board content based on individual VRM ASP assumptions for xPUs (*prev.we modelled at xPU board content level*)
- **xPU content:** we assume xPU power delivery content of $0.35/W for discrete and $0.7/W for module, roughly in line with our modelling above. This number is then multiplied by the number of xPU per xPU board (*prev. xPU assumptions embedded into GPU board)*
- **CPU content:** we assume CPU power delivery content represents c20-25% of the total GPU board content by multiplying xPU content by 1.3x (*prev. CPU assumptions embedded into GPU board)*
- **Discrete vs. module:** we assume a shift to module based power delivery when power losses go above 20%, or around 1,400W xPU power, unless otherwise known e.g. TPU using module

Figure 40: these changes have the largest impact on our xPU board estimates for Feynman

	Hopper	Blackwell	Blackwell Ultra	Rubin	Rubin Ultra	Feynman
Current xPU board BOM $	50	118	127	419	473	1092
Prev. estimates	54	108	126	360	389	648
Change	-7%	10%	1%	16%	22%	69%

Source: UBS estimates

4. Is power a bottleneck?

There has been lots of focus on power semis as a potential bottleneck to the AI infrastructure build out. Whilst this is clearly well founded in growing capacity constraints in parts of the supply chain, we believe there are other areas that will act as a bottleneck before supply is slowed by power semis constraints, such as switch gear, power transformers and gas turbines.

Figure 41: power semis getting tighter but bottlenecks more likely elsewhere

Supply Chain Component	Core Equipment Examples	Typical Lead Time
Power Semiconductors	PMICs, High-Power MOSFETs, SiC and GaN Devices	6 to 12 months
Backup Power Systems	Large-Megawatt Diesel and Gas Generators	10 to 18 months
Switch Gear	MV and LV switchgear	10 to 18 months
Power Transformers	Utility-Scale Substation Transmission Transformers	36 to 42 months
Gas Turbines	Heavy-Duty On-Site Generation Turbines	2-7 years

Source: UBS, company data

a. How capacity constrained are power semis?

When looking at supply constraints for power semis we think it is important to split out stage one and stage two power supplies. Stage one is dominated by high voltage and increasingly wide band gap semis, while stage two is high end silicon. We do not see stage one

i. Stage 1 - PSUs and IBCs

Stage 1 is dominated by high voltage silicon MOSFETs and increasingly wide band gap materials like GaN and SiC. Given the substantial SiC capacity build out to supply traction inverters for xEVs, we do not foresee capacity issues in SiC. 2025 SiC revenues and 2026 guides amongst the leading suppliers sit 16% and 3% below peak revenues in 2023, but 50-60% below 2025E expectations, which a proportion of capacity has already been built out for.

Figure 42: unlikely to see capacity constraints in SiC

	2019	2020	2021	2022	2023	2024E	2025E	2025 act.	2026E
STM	204	290	550	700	1,200	1,700	2,000	500	600
Wolfspeed	110	107	394	560	428	700	1,000	435	465
Rohm	73	94	137	205	320	521	782	250	325
Infineon	70	95	230	315	540	810	1,080	810	972
ON Semi	-	-	120	273	820	1,600	1,900	790	840
Total	**457**	**586**	**1,431**	**2,053**	**3,308**	**5,331**	**6,762**	**2,785**	**3,202**
						vs. 2023 peak		-16%	-3%
						vs. 2025 expectations		-59%	-53%

Source: Company data, UBSe

GaN is already used on the primary side of PSU DC-DC converters and will be needed in the HV IBC / DC-DC stage. While this materially increases unit content, our analysis suggests that demand remains manageable relative to existing and planned capacity. Innoscience alone has disclosed approximately 15 kwpm of capacity, rising toward 20 kwpm. Meanwhile, GaN content in HV IBCs may largely replace content previously located in the PSU, rather than represent wholly incremental demand.

Figure 43: GaN wafer demand

	Assumptions
12kW PSU GaN FETs	96
die size mm2	9
12kW wafer area (mm2)	864
Dpw 200mm	2859
Dpw 300mm	6771
PSU per rack	24
Racks (k)	100
Total PSUs (k)	2400
Required wafer area (mm2 '000)	230400
200mm Wafers (k)	**81**
300mm Wafers (k)	**34**

Source: UBS

There are signs of capacity constraints in high and medium voltage Si MOSFETs that are used in stage 1, however, this could be at least partially eased by increasing adoption of SiC and GaN as higher power modules are used. Alongside this we estimate additional capacity coming online within six to twelve months, where existing MOSFET designs are processed on previously IGBT focussed lines.

ii. Stage 2 - VRMs

Stage 2 is dominated by low to medium voltage silicon MOSFETs and other high end silicon analog components. This is where there is the most focus on capacity constraints. Whilst supply is certainly constrained in 2026E, we estimate a near 2x in capacity coming online in 2027E across the key players in the low to medium voltage space.

Key assumptions

- **Infineon** - 70% of c€1.8bn AI revenue related to medium to low voltage silicon / analog components e.g IBC and VRM and lower voltage silicon components within the PSU
- **Monolithic Power Systems** - we assume 70% of enterprise data consensus estimates are AI related. Uplifted is supported by the company's statement that they will increase capacity from over $4bn in 2026 to $6bn in the near future
- **Renesas -** 30% of AI data center revenues is low to medium voltage silicon related, growing 40% driven largely by higher utilisation rates, currently sitting around 60-70%
- **Vicor -** management expect advanced products excluding royalties of $300m in 2026E and expect to increase group level capacity of $1.5bn up from $1bn

Figure 44: >2x uplift in capacity coming online in 2027E

Capacity ($m)	2026E	2027E	Assumptions
Infineon	1,260	2,860	70% of AI revenue related to medium to low voltage silicon / analog components
MPS	1,155	2,079	70% of enterprise data is AI related. Group capacity increase from $4bn to $6bn
Renesas	449	629	30% of AI revenue related to medium to low voltage power. 40% growth higher utilisation
Vicor	300	800	Advanced products excluding royalties. Stated by mgmt.
Total	**3,164**	**6,368**	***~2x increase in capacity***

Source: UBS, VA cons

b. Can power semis ease constraints elsewhere?

Every conversion stage loses energy and creates heat. SiC and GaN, or vertical power delivery can support architectures with fewer conversion stages and as a result increases energy efficiency, delivering more usable compute per MW of power delivered to the facility. In a simplified scenario we assume roughly 6% total uplift from adopting a vertical VRM, GaN in the IBC and more GaN/SiC in the PSU.

Figure 45: power semis can improve energy efficiency but don't fix the bottleneck

	Energy efficiency uplift
VRM to vertical	2%
GaN based IBC	2%
More GaN/SiC in PSU	2%
Total uplift	**6%**

	Base	More compute	Same compute
Facility level power	130	130	122
IT rack power	100	100	94
GPU power	90	96	90
IT rack power effciency	90%	96%	96%
Facility PUE	1.3	1.3	1.3

Source: UBS

There are clearly many benefits of adopting technologies that can improve energy efficiency such as lower facility level energy costs, lower heat loss and as such cooling needs. That said if we are looking at this purely from a bottleneck perspective, we do not think a 6% improvement eases situations where a 500MW or 1GW+ facility is attempting to connect to the grid or deploy behind the meter energy supplies.

***UBS Evidence Lab** is a sell-side team of experts that work across numerous specialized labs creating insight-ready datasets. The experts turn data into evidence by applying a combination of tools and techniques to harvest, cleanse, and connect billions of data items each month. Since 2014, UBS Research analysts have utilized the expertise of UBS Evidence Lab for insight-ready datasets on companies, sectors, and themes, resulting in the production of thousands of differentiated UBS Research reports. UBS Evidence Lab does not provide investment recommendations or advice but provides insight-ready datasets for further analysis by UBS Research and by clients. All published UBS Evidence Lab content is available via UBS Neo. The amount and type of content available may vary. Please contact your UBS sales representative if you wish to discuss access.*

Global Data Centers Regional Exposure Monitor (>Access Dataset)

This report leverages the following UBS Evidence Lab asset: Global Data Centers Regional Exposure Monitor. UBS Evidence Lab, includes content supplied by S&P Global; Copyright © S&P Global Market Intelligence 2026. All rights reserved. UBS Evidence Lab has performed comprehensive validation, cleansing, and aggregation using geographical attributes and additional data dimensions.



Valuation Method and Risk Statement

Tech investing involves risk. It is difficult for the investment community, UBS included, to project the financial results of tech companies, as their operating models are highly volatile and unpredictable, and they compete in a highly dynamic marketplace. In addition to their low predictability, valuing tech stocks can be challenging because neither traditional nor non-traditional valuation measures have provided much insight into how tech stocks trade.

Analog Devices Inc:
We use NTM P/E methodology to value ADI. Analog is keyed to the macro-economy, especially the Industrial end market which is ADI's largest, and we see this as a risk both to the upside and downside. Analog names are also levered to Electric Vehicle adoption and we see them as being sensitive to EV adoption rates. Specific to ADI, MXIM integration is also a challenge and opportunity.

Infineon Technologies AG:
Infineon is a top 10 global semiconductor company with exposure to auto, industrial and chipcard semiconductors and hence, in spite of its broad portfolio, the company is exposed to fluctuations of the semiconductor cycle. The company is also exposed to overall autos production volumes and the mix towards hybrid and electric vehicles (which have higher semiconductor content). There are also risks of disruptions with new forthcoming technologies. The company is also exposed to fluctuations in the USD:EUR exchange rate with a strong USD positive for the business. We value Infineon using a DCF methodology with a WACC of 9% and a terminal growth rate of 2%.

STMicroelectronics:
Risk factors include GDP growth rates, continuing weakening of US$/EURO exchange rate, technology product life cycles, capacity growth, demand and supply and utilisation rates which contribute to general semiconductor growth. STMicro's biggest medium-term risk is centred on being able to deliver new image sensing products into smartphones. Our valuation for STMicro is based on a DCF methodology with a WACC of 9% and a terminal growth rate of 2%.

Texas Instruments Inc:
TXN: Our valuation is based on an EV/FCF multiple methodology. Texas Instruments faces a variety of risks, including rapidly changing technology, intense competition, high level of capital investment, pricing pressure, and cyclical end-market demand.