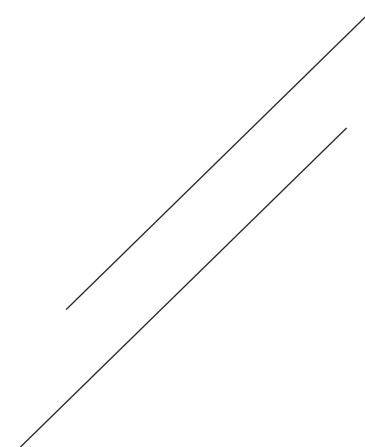




04

NTT



►NTT/INTT

- ✓ NTT is a variant of the Fast Fourier Transform (FFT) based on a finite field.
- ✓ It transforms polynomial multiplication into “pointwise multiplication,” accelerating polynomial multiplication operations.
- ✓ Cooley-Tukey decomposition is used for NTT and Gentleman-Sande decomposition for INTT in a butterfly architecture.
- ✓ It adopts a Radix-2 Multi-path Delay Commutator (MDC) FFT, using 8 Butterfly Units (BU) to process data with $N = 256$.

► NTT - Mathematical Derivation

- ✓ The Number Theoretic Transform (NTT) is defined as:

$$\hat{a}_j = \sum_{i=0}^{n-1} \psi^{2ij} a_i \mod q$$

Note:

$$\begin{aligned}\psi^{k+2n} &= \psi^k \\ \psi^{k+n} &= -\psi^k\end{aligned}$$

- ✓ Using the Cooley-Tukey decomposition:

$$\begin{aligned}\hat{a}_j &= \sum_{i=0}^{n/2-1} \psi^{4ij+2i} a_{2i} + \sum_{i=0}^{n/2-1} \psi^{4ij+2j+2i+1} a_{2i+1} \mod q \\ &= \sum_{i=0}^{n/2-1} \psi^{4ij+2i} a_{2i} + \psi^{2j+1} \sum_{i=0}^{n/2-1} \psi^{4ij+2i} a_{2i+1} \mod q\end{aligned}$$

- ✓ Define :

$$A_j = \sum_{i=0}^{n/2-1} \psi^{4ij+2i} a_{2i}, \quad B_j = \sum_{i=0}^{n/2-1} \psi^{4ij+2i} a_{2i+1}$$

- ✓ Thus, the transformed coefficients are:

$$\begin{aligned}\hat{a}_j &= A_j + \psi^{2j+1} B_j \mod q \\ \hat{a}_{j+n/2} &= A_j - \psi^{2j+1} B_j \mod q\end{aligned}$$

► INTT - Mathematical Derivation

- ✓ The Inverse Number Theoretic Transform (INTT) is given by:

$$\mathbf{a}_i = \sum_{j=0}^{n-1} \psi^{-(2i+1)j} \hat{a}_j \mod q$$

Note:

$$\begin{aligned} \psi^{k+2n} &= \psi^k \\ \psi^{k+n} &= -\psi^k \end{aligned}$$

- ✓ Using the Gentleman-Sande decomposition:

$$\begin{aligned} \mathbf{a}_i &= \left[\sum_{j=0}^{n/2-1} \psi^{-(2i+1)j} \hat{a}_j + \sum_{j=0}^{n/2-1} \psi^{-(2i+1)(j+n/2)} \hat{a}_{j+n/2} \right] \mod q \\ &= \psi^{-i} \left[\sum_{j=0}^{n/2-1} \psi^{-2ij} \hat{a}_j + \sum_{j=0}^{n/2-1} \psi^{-2i(j+n/2)} \hat{a}_{j+n/2} \right] \mod q \end{aligned}$$

- ✓ Define :

$$A_i = \sum_{j=0}^{n/2-1} \hat{a}_j \psi^{-4ij}, \quad B_i = \sum_{j=0}^{n/2-1} \hat{a}_{j+n/2} \psi^{-4ij}$$

- ✓ Thus, the inverse transform coefficients are:

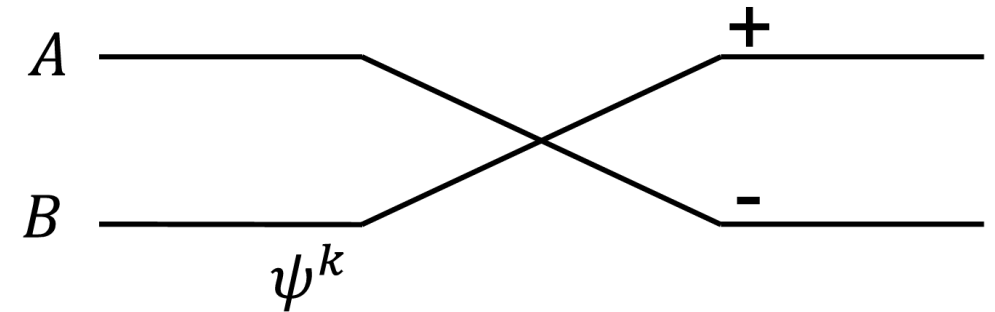
$$\begin{aligned} \mathbf{a}_{2i} &= (A_i + B_i) \psi^{-2i} \mod q \\ \mathbf{a}_{2i+1} &= (A_i - B_i) \psi^{-2i} \mod q \end{aligned}$$

► NTT/INTT - Butterfly diagram

✓ Fundamental in FFT/NTT computations, is given by:

$$A = A + \psi^k B$$

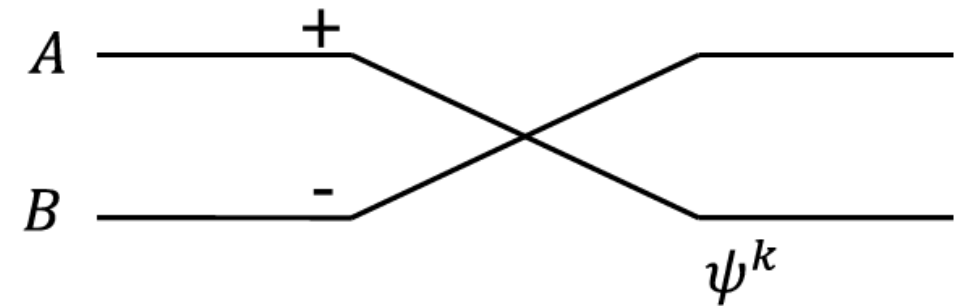
$$B = A - \psi^k B$$



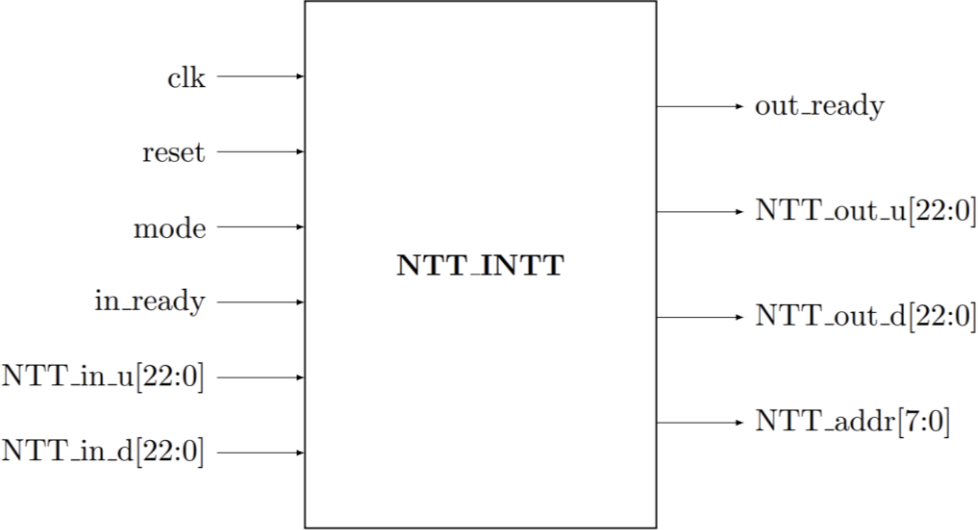
✓ Fundamental in FFT/NTT computations, is given by:

$$A = A + B$$

$$B = (A - B)\psi^{-k}$$



►NTT_INTT - Module Pin Function Definition

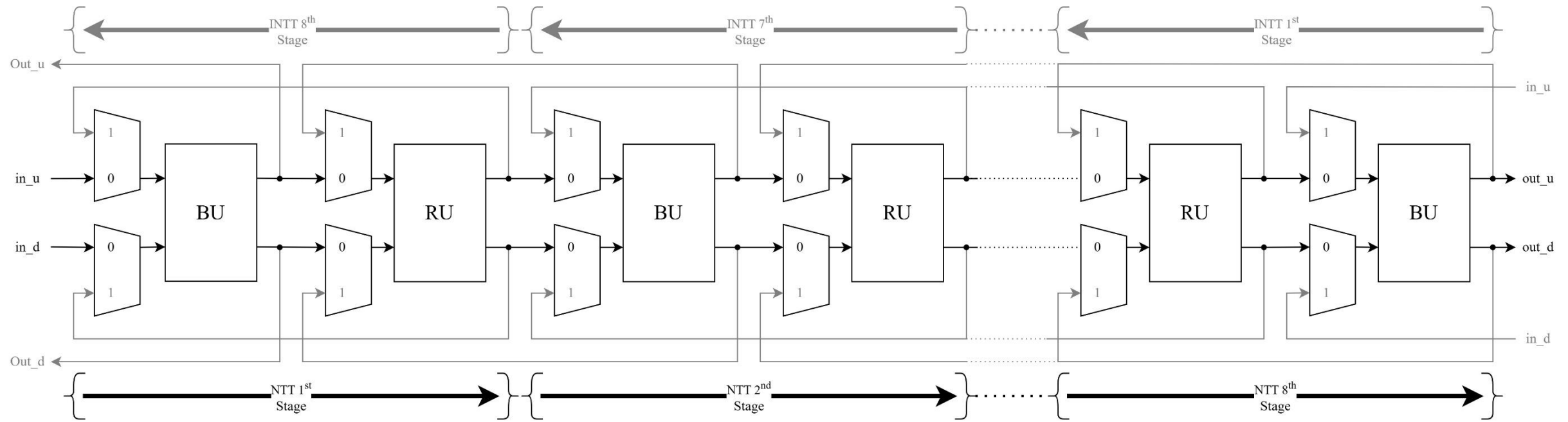


Pin name	I/O	Bit width	Function
clk	I	1	系統時脈
reset	I	1	系統同步重置信號
mode	I	1	0 : NTT mode / 1 : INTT mode
in_ready	I	1	由待轉換的MEM輸入資料準備好的指示信號
NTT_in_u	I	23	由待轉換的MEM輸入data
NTT_in_d	I	23	由待轉換的MEM輸入data
out_ready	O	1	轉換完成資料開始輸出的指示信號
NTT_out_u	O	23	轉換完成資料輸出data至指定MEM
NTT_out_d	O	23	轉換完成資料輸出data至指定MEM
NTT_addr	O	8	指定輸出的MEM位址

►NTT_INTT - Module Task Assignment Table

Pin name	I/ O	Bit length	Function
clk	I	1	系統時脈
reset	I	1	系統同步重置信號
mode	I	1	0 : NTT mode / 1 : INTT mode
in_ready	I	1	由待轉換的MEM輸入資料準備好的指示信號
NTT_in_u	I	23	由待轉換的MEM輸入data
NTT_in_d	I	23	由待轉換的MEM輸入data
out_ready	O	1	轉換完成資料開始輸出的指示信號
NTT_out_u	O	23	轉換完成資料輸出data至指定MEM
NTT_out_d	O	23	轉換完成資料輸出data至指定MEM
NTT_addr	O	8	指定輸出的MEM位址

► NTT_INTT – Block Diagram



► NTT – Timing Diagram

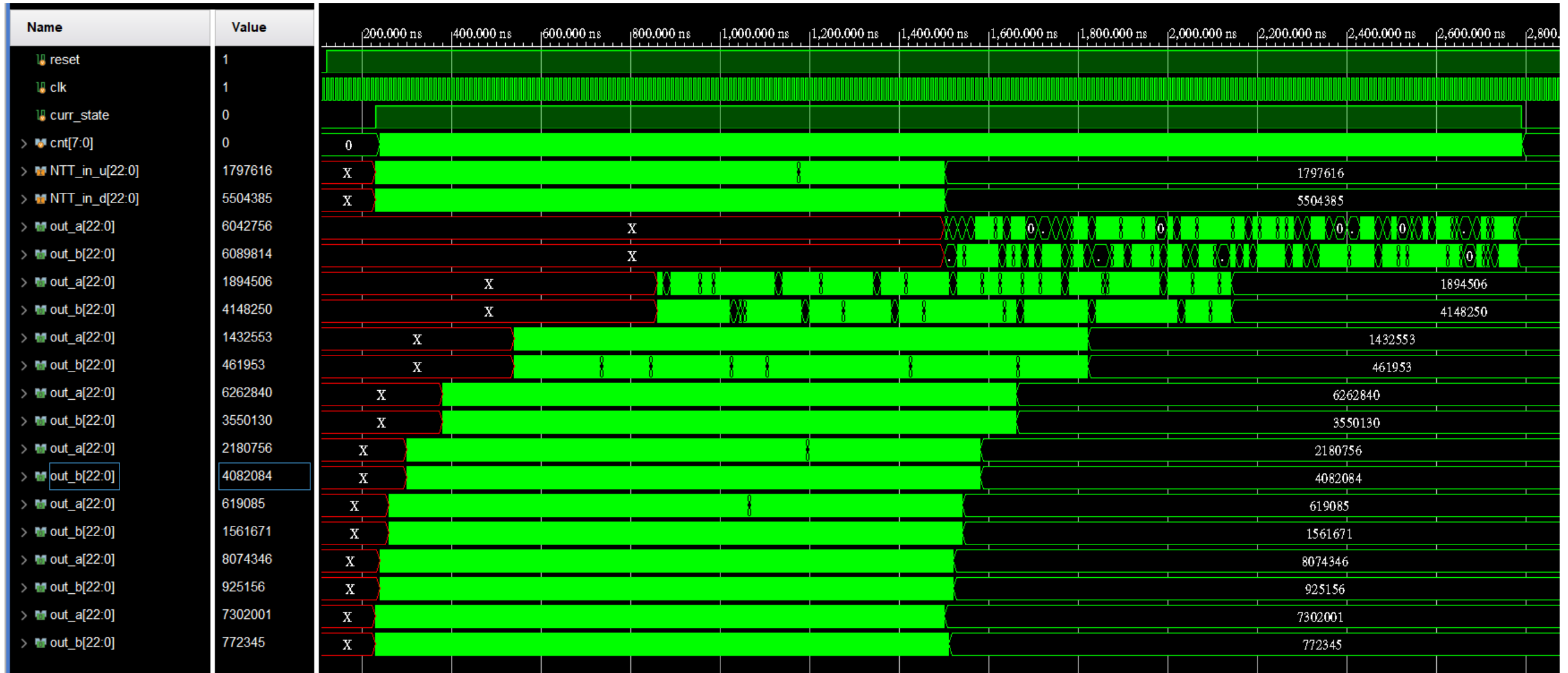


►NTT - – Timing Diagram

> out_b[22:0]	0	0														
> out_a[22:0]	0	3292360	5639478	5157967	1713881	3162035	7669833	6641394	4903989	0						
> out_b[22:0]	0	6442904	5232688	8144313	8208803	5357188	3774935	5558247	1994723	0						
> out_a[22:0]	0	6189896	906783	5793821	3339230	5215824	832905	6436092	4288883	5223614	6917442	584788	6170123	0		
> out_b[22:0]	0	6911461	1654786	5202208	2925445	1368896	2065634	3879842	7519296	7662194	3547934	7373421	1867066	0		
> out_a[22:0]	0	958422	6629204	1153275	2424097	7067290	2409715	1575198	4222356	6608599	2905824	1439604	3095744	7483179	5524427	0
> out_b[22:0]	0	8031462	7824196	2846100	7769886	6755632	899857	476033	5823871	4509610	1225444	627207	2358723	7841209	1571441	0
> out_a[22:0]	0	5719810	7847079	3957586	614766	7118893	7254363	3656372	2956723	5031747	5940663	6967820	3793081	2168270	7132662	1797616
> out_b[22:0]	0	332082	2450182	3724921	1691784	6953724	6880217	1474475	193673	4300736	7276535	2051400	7466544	7466561	7833696	5504385

5157967	8144313	5793821	5202208	1153275	2846100	3957586	3724921
1713881	8208803	3339230	2925445	2424097	7769886	614766	1691784
3162035	5357188	5215824	1368896	7067290	6755632	7118893	6953724
7669833	3774935	832905	2065634	2409715	899857	7254363	6880217
6641394	5558247	6436092	3879842	1575198	476033	3656372	1474475
4903989	1994723	4288883	7519296	4222356	5823871	2956723	193673
0	0	5223614	7662194	6608599	4509610	5031747	4300736
0	0	6917442	3547934	2905824	1225444	5940663	7276535
0	0	534788	7373421	1439604	627207	6967820	2051400
0	0	6170123	1867066	3095744	2358723	3793081	7466544
0	0	0	0	7483179	7841209	2168270	7466561
0	0	0	0	5524427	1571441	7132662	7833696
0	0	0	0	0	0	1797616	5504385
0	0	0	0	0	0	0	0

► INTT – Timing Diagram



► INTT – Timing Diagram

> out_a[22:0]	256	X		8380161	256		512		256		8380161	512	
> out_b[22:0]	8380161	X			8380161		0	512	8380161	512	8379905	256	
> out_a[22:0]	3678519	...	1023635	1023763	3678263	3678519		256	1023891	3678519	1023763	7356654	4702282
> out_b[22:0]	128	256	4702154	0	3678647	8380161	0	3678647	256	128	7357038	4701770	7356782
> out_a[22:0]	8324729	...	2239637	6473191	206200	1438946	6029468	2557405	912259	8324729	5226122	1121050	6665068
> out_b[22:0]	4066362	...	3416375	3678455	4978493	3954858	5711636	7744753	5435169	4066362	2351013	7744753	2557213
> out_a[22:0]	8356669	...	3270576	5145730	489870	719475	4704154	5279623	6142372	8356669	2709753	371237	8138681
> out_b[22:0]	4780255	...	2701048	3968203	5510760	4949176	3521803	1797269	4179009	4780255	2846760	5216195	1358765
> out_a[22:0]	7753974	...	5677629	3844460	2458359	3173158	5853236	3957719	1110448	7753974	5412541	4907194	
> out_b[22:0]	6213124	...	8187643	7190754	5830377	4953084	2136765	2050770	3933271	6213124	3475634	7369790	
> out_a[22:0]	619085	...	7742068	5330631	1260730	8339836	6491585	7007505	619085				
> out_b[22:0]	1561671	...	2845650	5710244	955419	5058229	2563891	7577367	1561671				

121	0	0	1023635	7356782	4258239	7992318	5864397	4518333
122	0	0	8380289	7356526	4010802	523968	607845	2328505
123	0	0	7356526	1023635	3940535	387971	3023220	4851779
124	0	0	1023763	256	8103694	7468350	2822564	2451780
125	0	0	1023635	4702154	2239637	3416375	3270576	2701048
126	0	0	1023763	0	6473191	3678455	5145730	3968203
127	8380161	8380161	3678263	3678647	206200	4978493	489870	5510760
128	256	8380161	3678519	8380161	1438946	3954858	719475	4949176
129	256	8380161	3678519	0	6029468	5711636	4704154	3521803
130	512	0	256	3678647	2557405	7744753	5279623	1797269
131	512	512	1023891	256	912259	5435169	6142372	4179009

► Modular Reduction

- ✓ For a 46-bit value s , modular reduction is performed recursively by exploiting the relation :

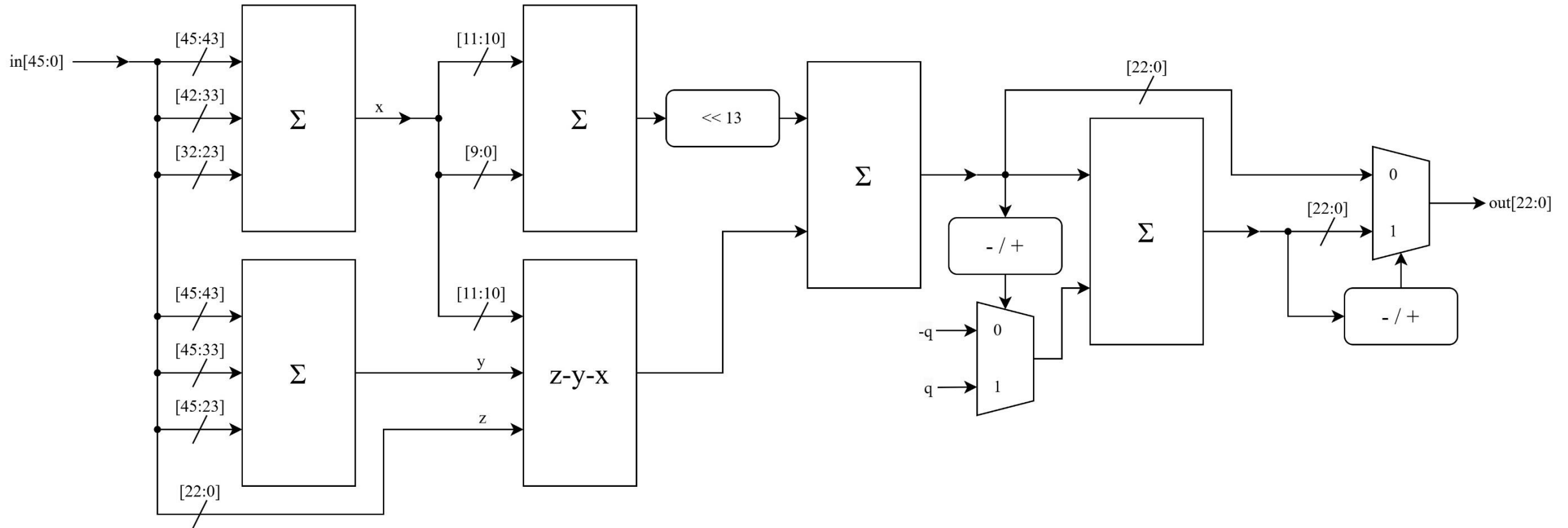
$$2^{23} \equiv 2^{13} - 1 \text{ mod } q$$

- ✓ The reduction ensures that the result falls within the interval $(-q, 2q)$ allowing for adjustments by adding q if negative or subtracting q if positive.
- ✓ After performing necessary additions or subtractions with q of the reduced result, the final output is determined by selecting the non-negative value.
- ✓ Each BU module contains one, and there are eight BU modules, requiring a total of eight modular reduction modules.

► Modular Reduction - Mathematical Derivation

$$\begin{aligned} s[45 : 0] &\equiv 2^{23}s[45 : 23] + s[22 : 0] \equiv 2^{13}s[45 : 23] - s[45 : 23] + s[22 : 0] \\ &\equiv 2^{23}s[45 : 33] + 2^{13}s[32 : 23] - s[45 : 23] + z \\ &\equiv 2^{13} (s[45 : 33] + s[32 : 23]) - (s[45 : 33] + s[45 : 23]) + z \\ &\equiv 2^{23}s[45 : 43] + 2^{13} (s[42 : 33] + s[32 : 23]) - (s[45 : 33] + s[45 : 23]) + z \\ &\equiv 2^{13} (s[45 : 43] + s[42 : 33] + s[32 : 23]) - (s[45 : 43] + s[45 : 33] + s[45 : 23]) + z \\ &\equiv 2^{13}x - y + z \equiv 2^{23}x[11 : 10] + 2^{13}x[9 : 0] - y + z \\ &\equiv 2^{13} (x[11 : 10] + x[9 : 0]) - (y + x[11 : 10]) + z \pmod{q} \end{aligned}$$

► Modular Reduction – Block Diagram



► Modular Reduction – Timing Diagram

Name	Value	10.000 ns	15.000 ns	20.000 ns	25.000 ns	30.000 ns	35.000 ns	40.000 ns	45.000 ns	50.000 ns	55.000 ns
> in[45:0]	98765432	0	70368744177663	123456789	8380416	987654321					
> out[22:0]	7145532	0	49144	6130951	8380416	7145532					
> in[45:0]	00003ade	000000000000	3fffffff	000075bed15	000007fe000	00003ade68b1					
> out[22:0]	6d083c	000000	00bff8	5d8d07	7fe000	6d083c					
> x[11:0]	075	000	805	00e	000	075					
> y[23:0]	000075	000000	802005	00000e	000000	000075					
> z[22:0]	5e68b1	000000	7ffff	5bed15	7fe000	5e68b1					
> d[11:0]	075	000	007	00e	000	075					
> e[26:0]	05e683c	0000000	7ffaff8	05bed07	07fe000	05e683c					
> f[23:0]	6d083c	000000	00bff8	5d8d07	7fe000	6d083c					
> adjust[23:0]	801fff			801fff							
> g[23:0]	ed283b	801fff	80dff7	ddad06	ffff	ed283b					



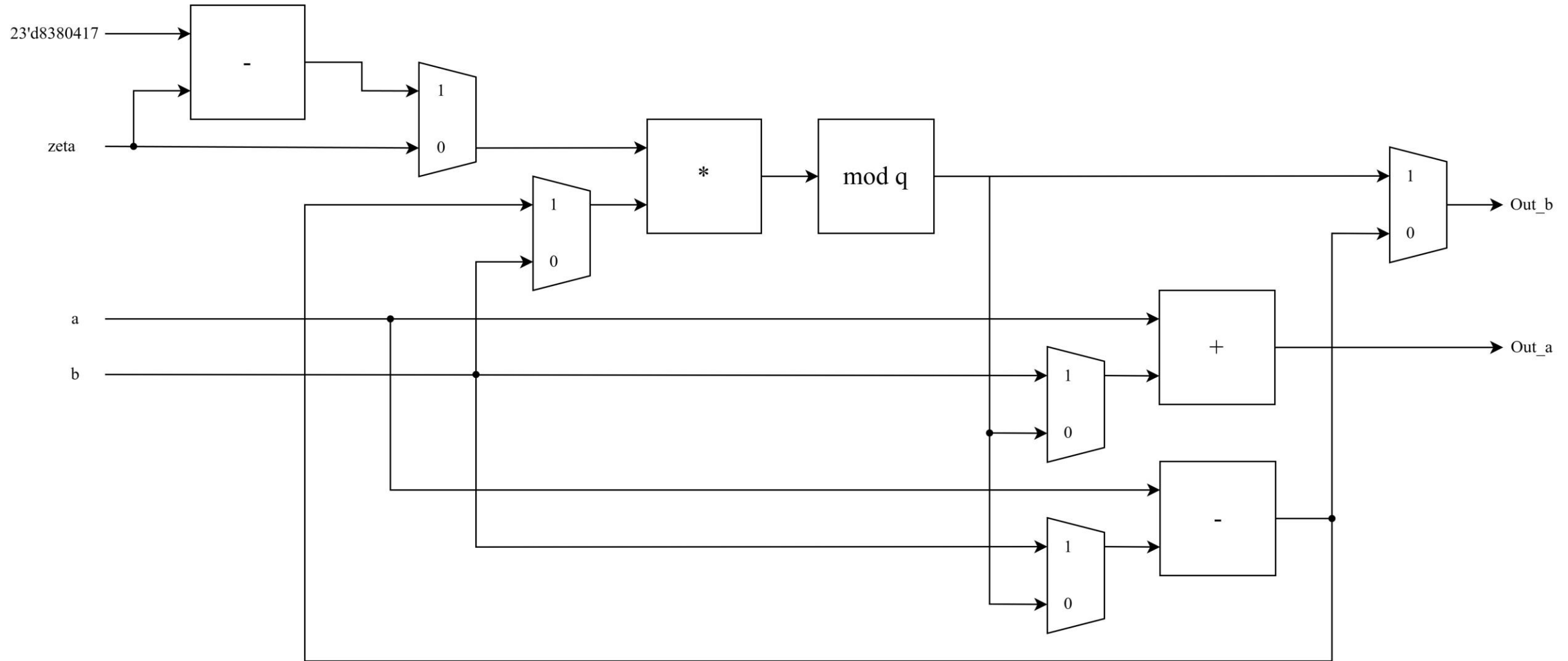
- ✓ The Butterfly Unit is formed based on the symmetry and parity properties in NTT and INTT computations.
- ✓ For a pair of input numbers a and b , together with a corresponding twiddle factor w , the butterfly operation in a finite field (modulo p) proceeds as follows:

$$x=(a+w\times b)\bmod q$$

$$y=(a-w\times b)\bmod q$$

- ✓ The structures of NTT and INTT are similar, but the twiddle factors used in INTT are the modular inverses of those in NTT. A normalization factor must also be applied at the end.
- ✓ Eight BU modules are used in the NTT/INTT of the thesis.

► BU – Block Diagram



- ✓ To match the butterfly structure, the output of each stage's BU is reordered accordingly.
- ✓ There are a total of 7 RU_i , where $1 \leq i \leq 7$, in our implemented NTT and INTT.
- ✓ In NTT/INTT, the MEM depth of each RU_i Stage is different:

$$\text{NTT : MEM_Depth_i} = 2 \times ((8-i)-1)$$

$$\text{INTT : MEM_Depth_i} = 2 \times (i-1)$$

► RUi – Block Diagram

