
Efficient Compression of Multi-view Video Exploiting Inter-view Dependencies Based on H.264/MPEG4-AVC

Dr.-Ing. Aljoscha Smolic

3DV and FVV as Functionalities

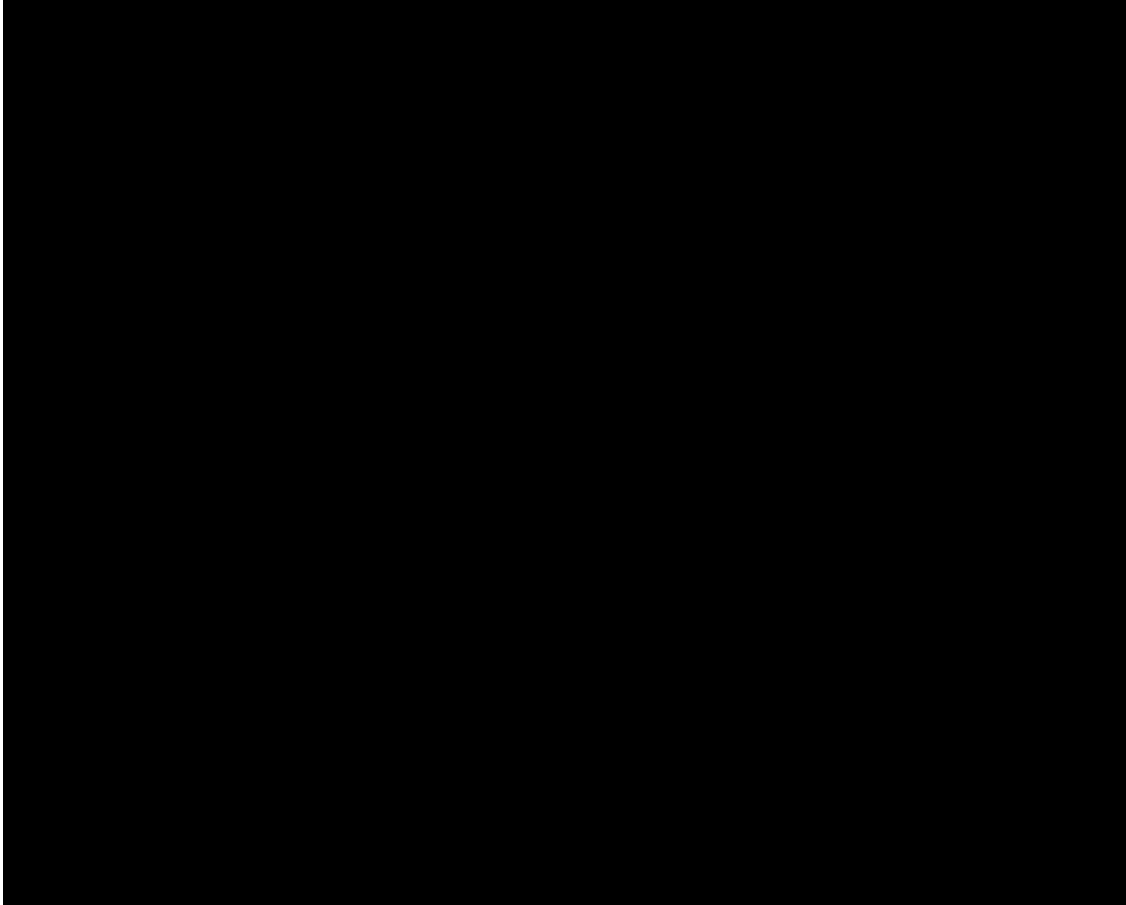
- 3D Video (3DV, also known as stereo) and Free Viewpoint Video (FVV) describe **functionalities**
- Expand the user's sensation beyond what is offered by classical 2D video
- They **do not exclude each other**, can be combined in a single system

FVV Acquisition

- Courtesy of:
Stephan Wuermlin
et al.,
ETH Zuerich,
Switzerland

[VMV_ETH\isfvv_dooyoung_input.wmv](#)

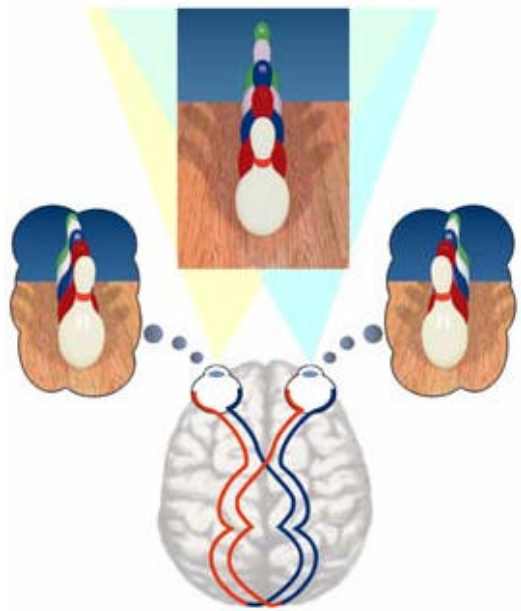
Free Viewpoint Video



- **Same functionality as CG objects**
 - free navigation, can be viewed from any viewpoint/direction
 - Integration into complete scenes (virtual/augmented/real)
- **But:** depict appearance, motion, deformation of **real world objects**

3DV / Stereo Video

- Generation of a 3D depth impression from separate views for each eye



More Details

- **3D Video and Free Viewpoint Video – Technologies, Applications and MPEG Standards**

Wed, 15:30

- **A Flexible 3DTV System for Different Multi-baseline Geometries**

Wed, 10:30

- **3D-TV: Primed for Success?**

Special Session I, II, III

- **Analysis of Hierarchical B Pictures and MCTF**

Wed, 10:50

Multi-view Video Coding (MVC)

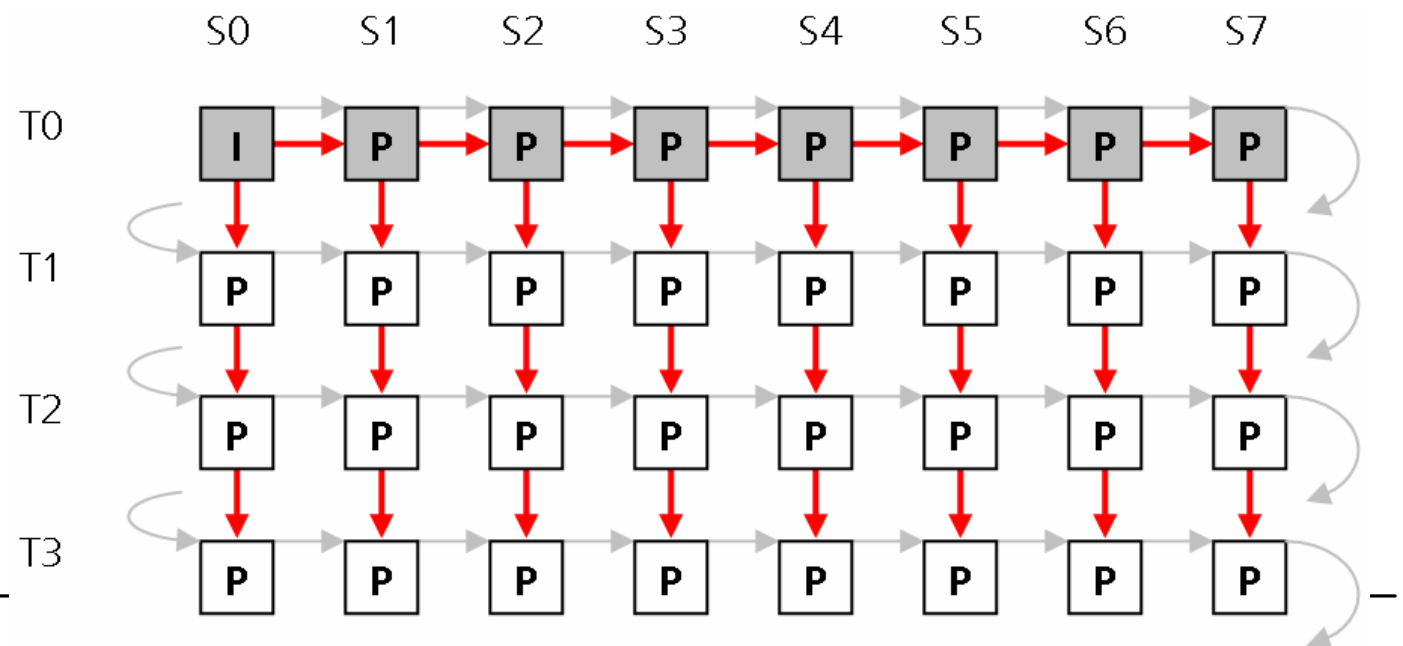
- FVV and 3DV representations require transmission of **multiple synchronized video signals** that show the same scenery from different viewpoints
- **Huge amount of data** that need to be compressed efficiently
- Contains a large amount of **inter-view statistical dependencies**



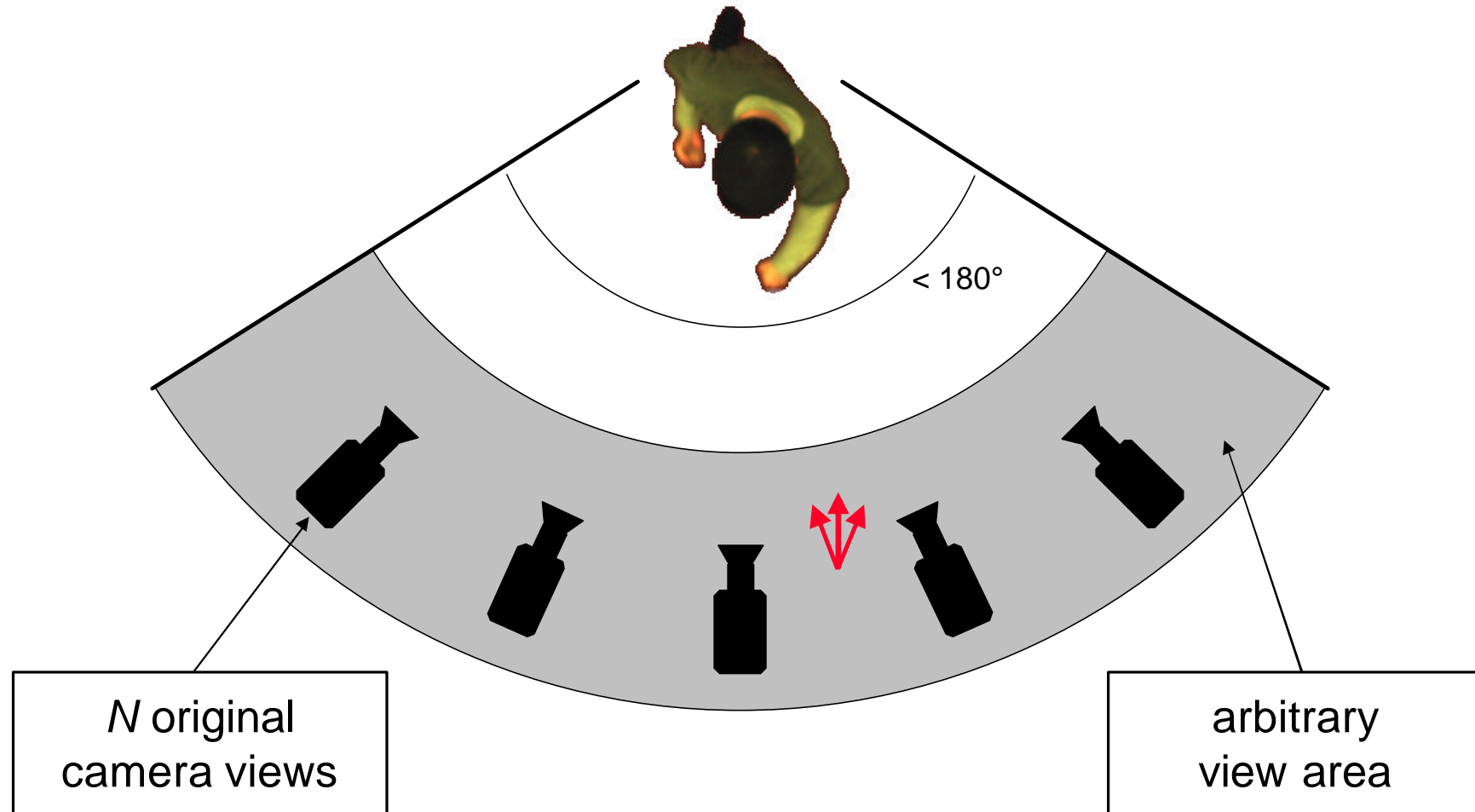
We would like to thank the Image Based Realities Group of Microsoft Research for providing the Breakdancers and Ballroom data sets.

Multi-view Video Coding (MVC)

- Evaluations have shown that specific MVC exploiting inter-view statistical dependencies provides significant coding gain
- Only for dense camera settings, not for dome-type arrangements



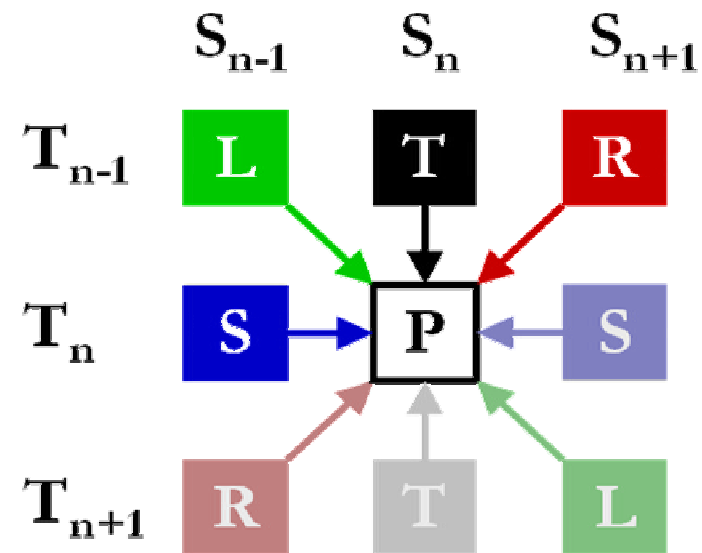
Dense Camera Settings

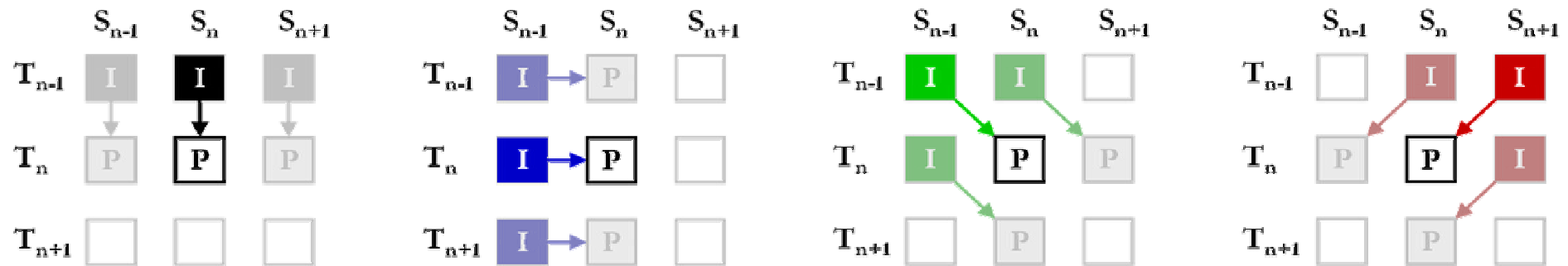


Multi-view Prediction

1st order neighbors

- **T** (temporal)
- **S** (inter-view)
- **L/R** (combined)

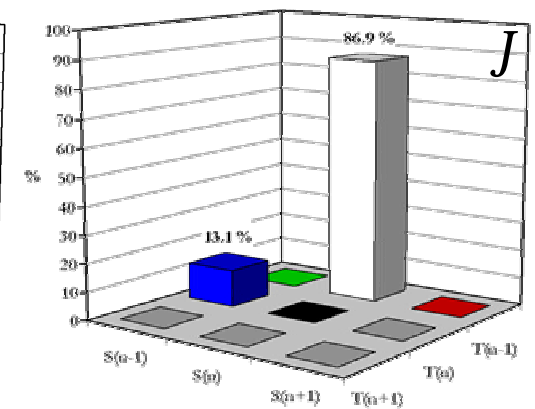
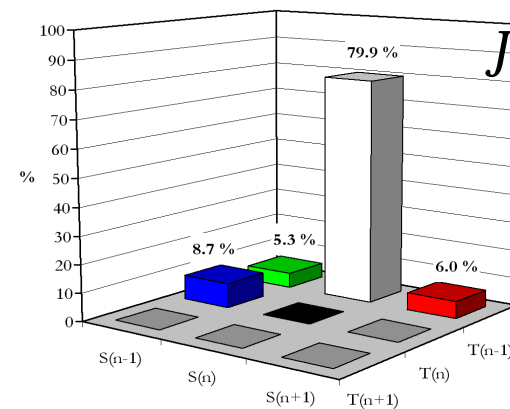




T better than S better than R/L

Influencing factors

Frame rate, camera distance, complexity of Content (motion, spatial detail)



MVC Standard

- MPEG decided to launch a new standard based on H.264/MPEG4-AVC
- MPEG-4 Part 10, Amd. 4

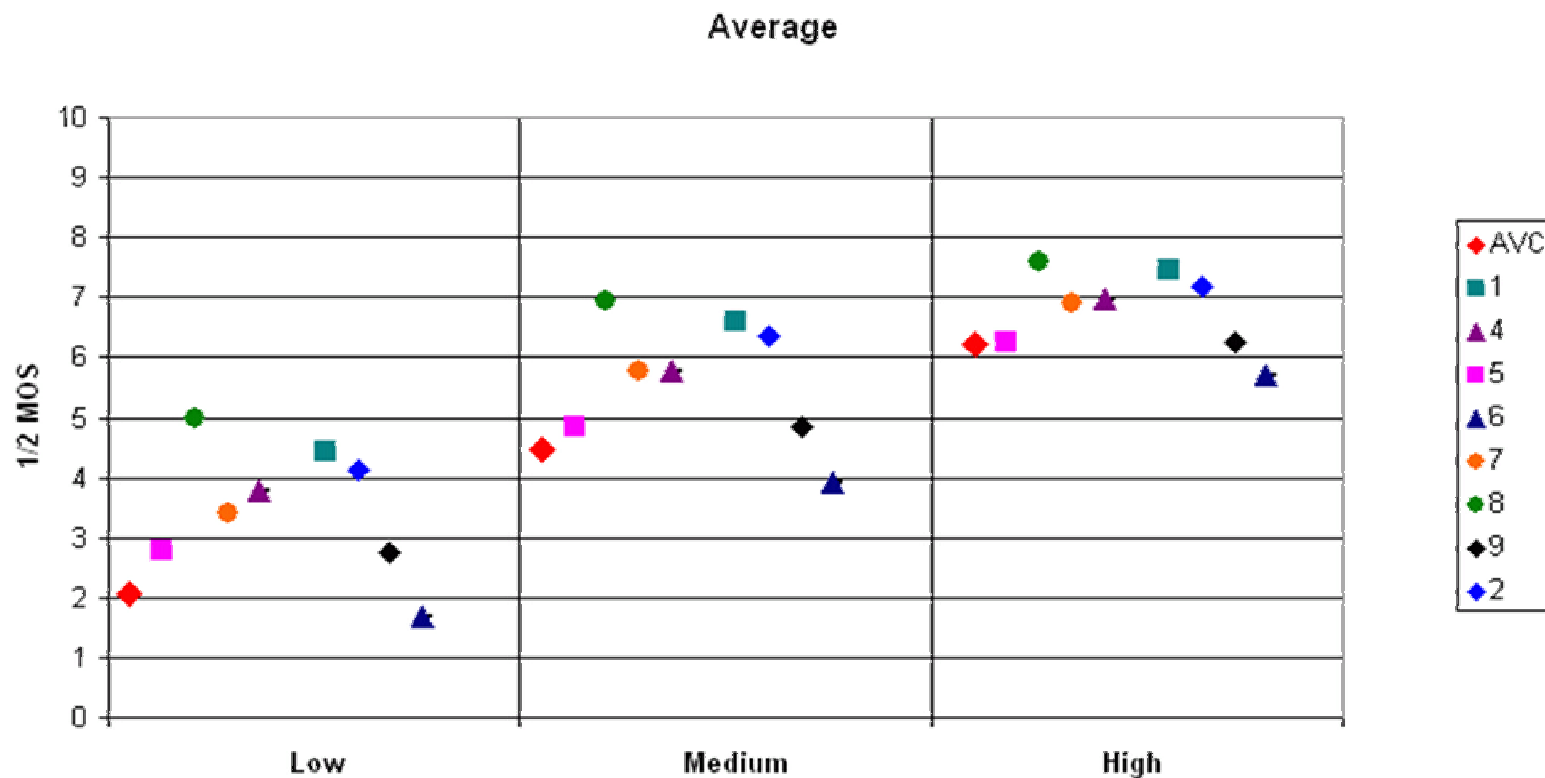
<input type="checkbox"/> Evaluation of Call for Proposals	January 2006
<input type="checkbox"/> WD	July 2006
<input type="checkbox"/> PDAM	January 2007
<input type="checkbox"/> FPDAM	July 2007
<input type="checkbox"/> FDIS	January 2008

Test Data & Conditions

- 8 MV data sets with 5-16 sequences each
- Different frame rates, image resolutions, camera distances, camera setups
- Representative test set for MV applications
- 3 bitrates, low, medium, high quality
- Over 200 sequences to be generated

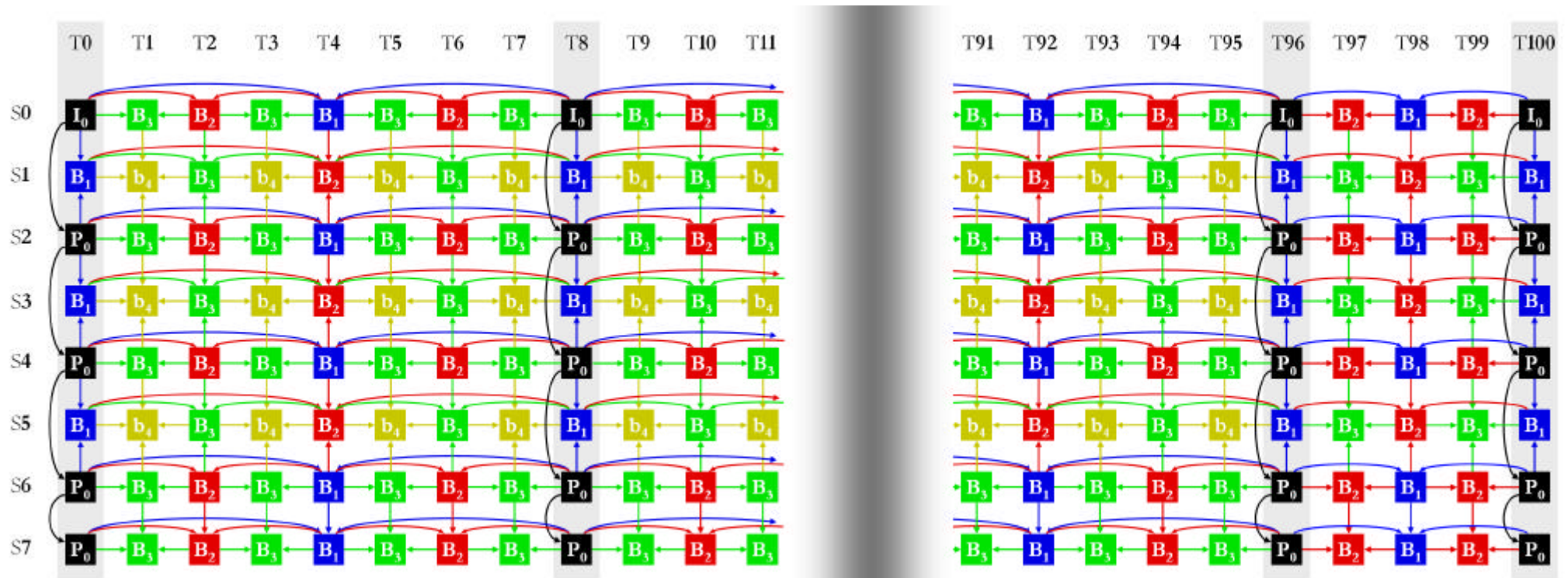
MVC Call for Proposals

- Current reference model is based on a proposal made by FhG-HHI
- Proved best performance out of 8 proposals in response to a CfP in formal subjective tests



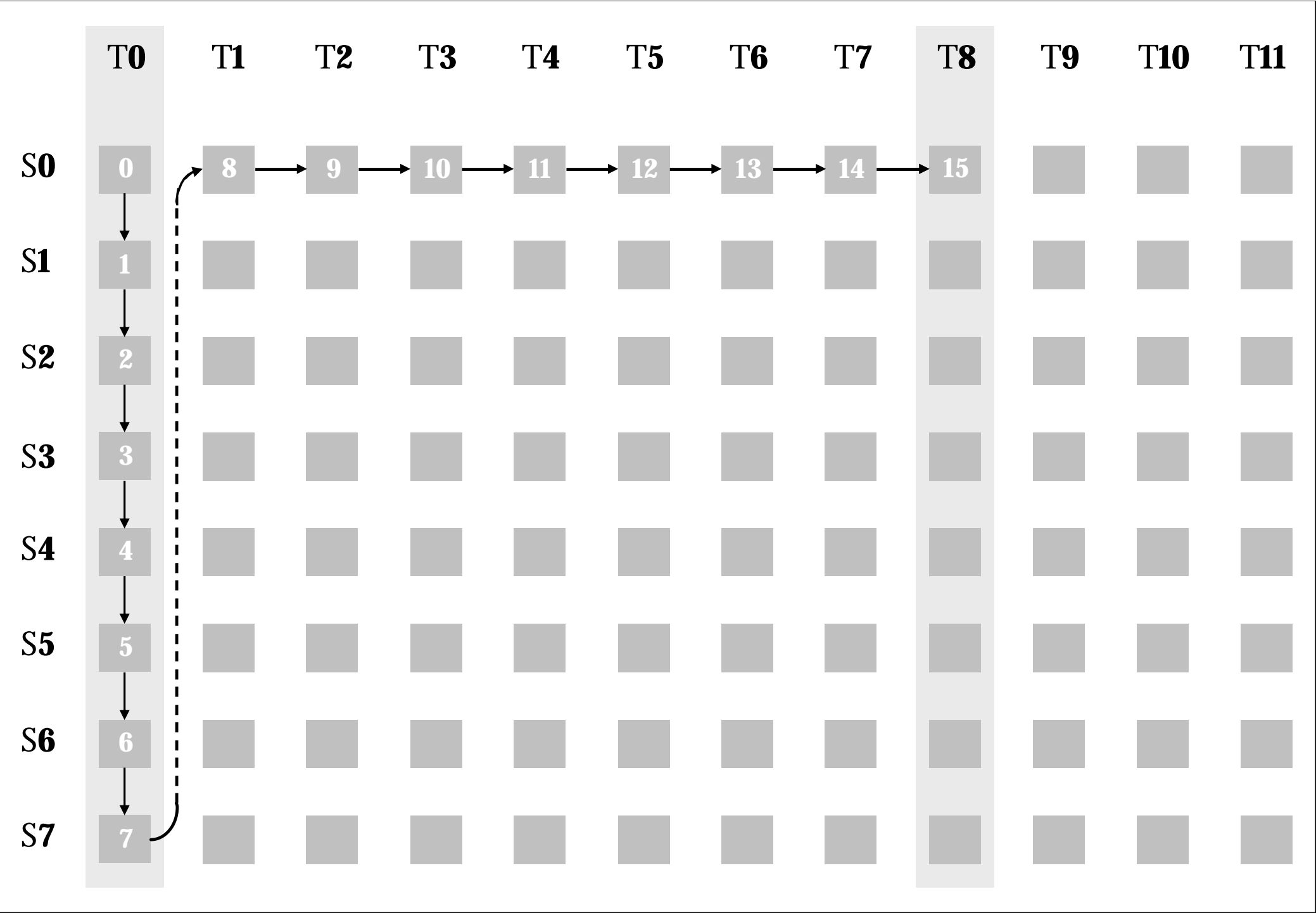
MVC Reference Model

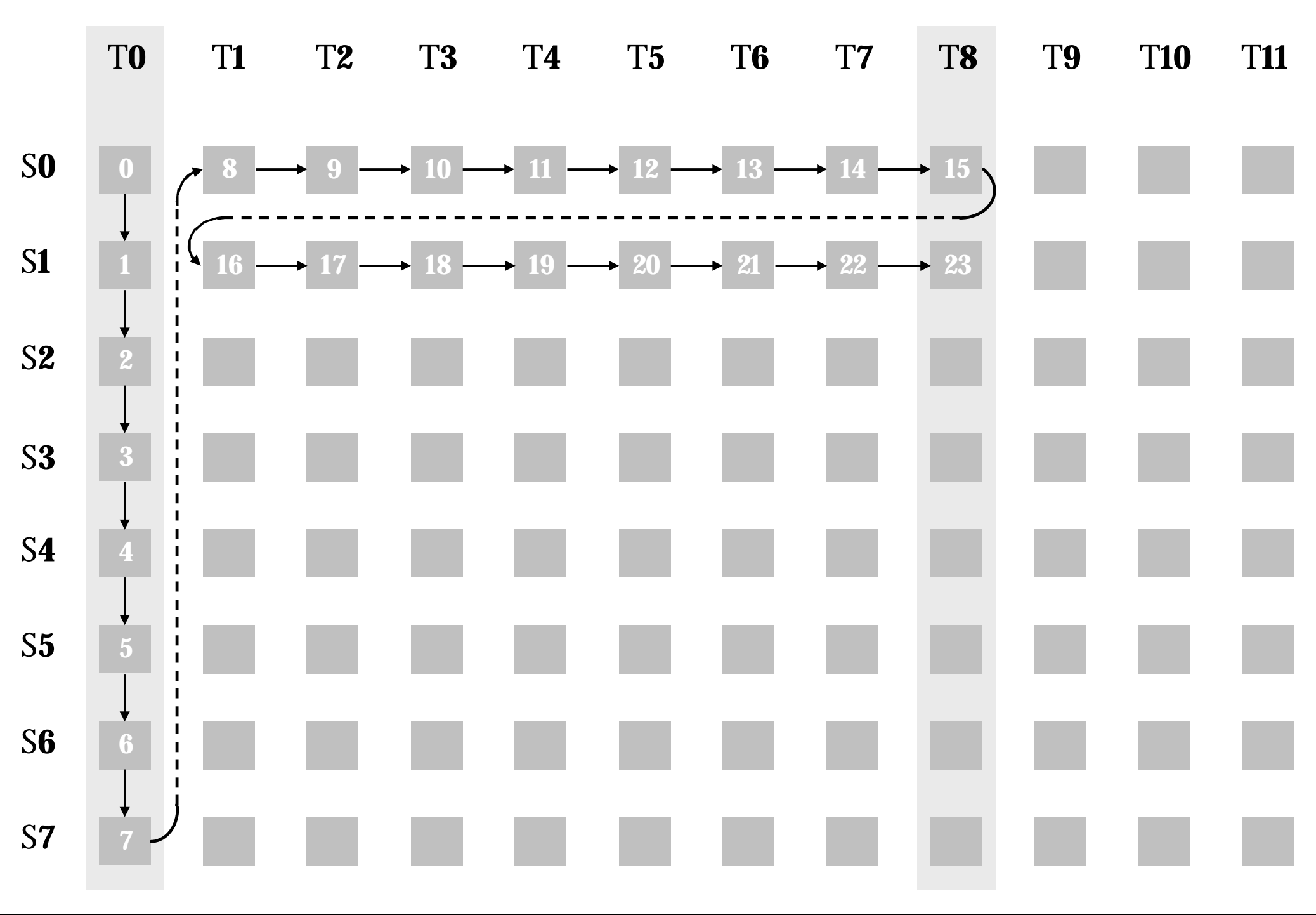
- Fully compatible to H.264/MPEG4-AVC
- Reorganization of input images into a single stream prior to encoding
- Uses hierarchical B-pictures combined in temporal and inter-view dimension

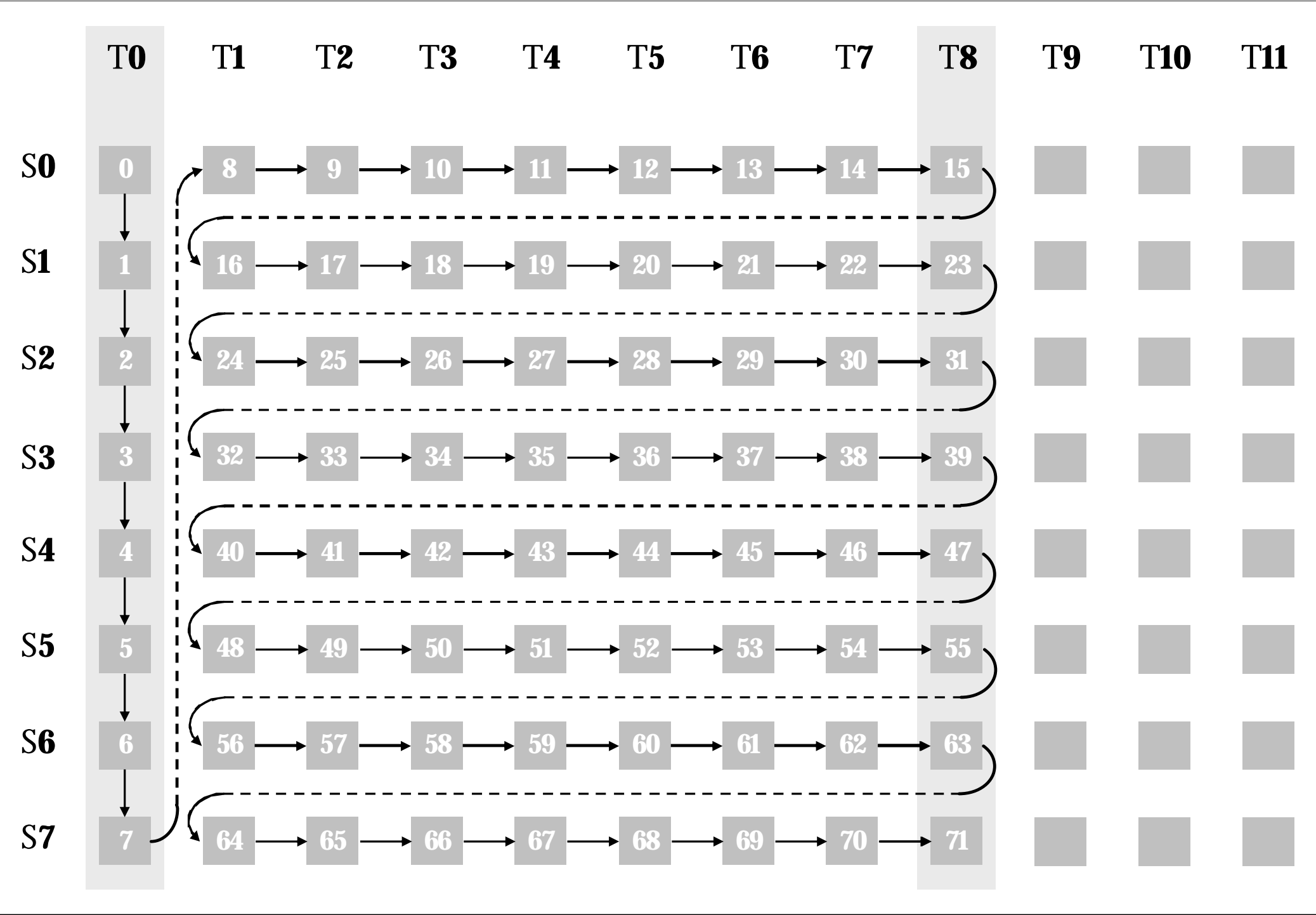


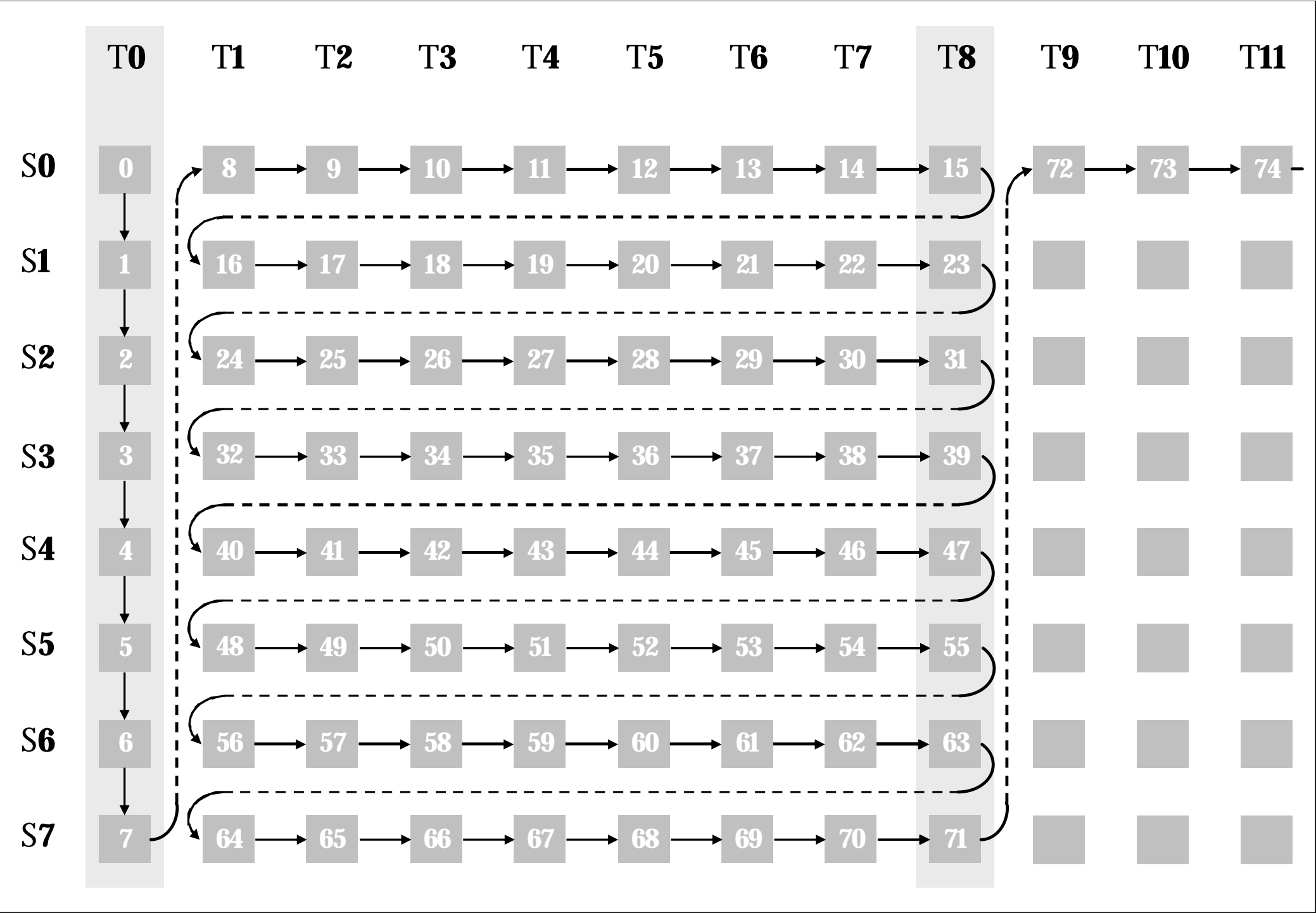
[illegible]

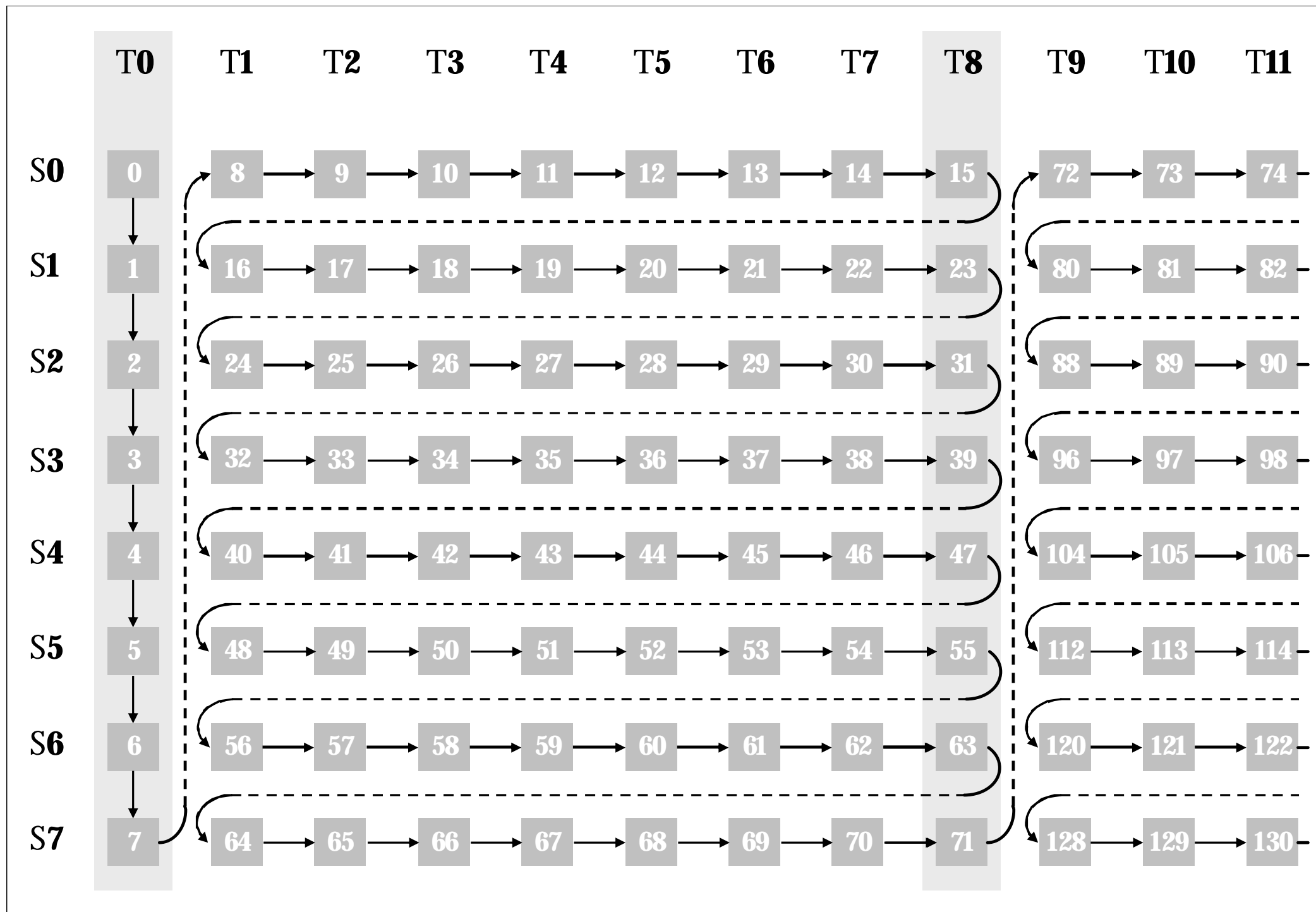
[illegible]











	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
S0	I ₀	8	9	10	11	12	13	14	15	72	73	74
S1	1	16	17	18	19	20	21	22	23	80	81	82
S2	2	24	25	26	27	28	29	30	31	88	89	90
S3	3	32	33	34	35	36	37	38	39	96	97	98
S4	4	40	41	42	43	44	45	46	47	104	105	106
S5	5	48	49	50	51	52	53	54	55	112	113	114
S6	6	56	57	58	59	60	61	62	63	120	121	122
S7	7	64	65	66	67	68	69	70	71	128	129	130

	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
S0	I ₀	8	9	10	11	12	13	14	15	72	73	74
S1	1	16	17	18	19	20	21	22	23	80	81	82
S2	P ₀	24	25	26	27	28	29	30	31	88	89	90
S3	3	32	33	34	35	36	37	38	39	96	97	98
S4	4	40	41	42	43	44	45	46	47	104	105	106
S5	5	48	49	50	51	52	53	54	55	112	113	114
S6	6	56	57	58	59	60	61	62	63	120	121	122
S7	7	64	65	66	67	68	69	70	71	128	129	130

	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
S0	I ₀	8	9	10	11	12	13	14	15	72	73	74
S1	B ₁	16	17	18	19	20	21	22	23	80	81	82
S2	P ₀	24	25	26	27	28	29	30	31	88	89	90
S3	3	32	33	34	35	36	37	38	39	96	97	98
S4	4	40	41	42	43	44	45	46	47	104	105	106
S5	5	48	49	50	51	52	53	54	55	112	113	114
S6	6	56	57	58	59	60	61	62	63	120	121	122
S7	7	64	65	66	67	68	69	70	71	128	129	130

	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
S0	I ₀	8	9	10	11	12	13	14	15	72	73	74
S1	B ₁	16	17	18	19	20	21	22	23	80	81	82
S2	P ₀	24	25	26	27	28	29	30	31	88	89	90
S3	3	32	33	34	35	36	37	38	39	96	97	98
S4	P ₀	40	41	42	43	44	45	46	47	104	105	106
S5	5	48	49	50	51	52	53	54	55	112	113	114
S6	6	56	57	58	59	60	61	62	63	120	121	122
S7	7	64	65	66	67	68	69	70	71	128	129	130

Diagram illustrating a grid structure with rows S0 through S7 and columns T0 through T11. The grid contains numerical values and specific labels.

Key elements and connections:

- Row S0, Column T0: I₀ (black box)
- Row S1, Column T0: B₁ (blue box)
- Row S2, Column T0: P₀ (black box)
- Row S4, Column T0: P₀ (black box)
- Arrows: A curved arrow connects S0 to S2, and a straight arrow connects S1 to S2. A straight arrow points from I₀ to B₁, and another points from B₁ to P₀.

	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
S0	I ₀	8	9	10	11	12	13	14	15	72	73	74
S1	B ₁	16	17	18	19	20	21	22	23	80	81	82
S2	P ₀	24	25	26	27	28	29	30	31	88	89	90
S3	B ₁	32	33	34	35	36	37	38	39	96	97	98
S4	P ₀	40	41	42	43	44	45	46	47	104	105	106
S5	5	48	49	50	51	52	53	54	55	112	113	114
S6	6	56	57	58	59	60	61	62	63	120	121	122
S7	7	64	65	66	67	68	69	70	71	128	129	130

Diagram illustrating a data structure layout with rows S0 through S7 and columns T0 through T11. The layout is divided into two main vertical sections by a light gray background.

Left Section (Columns T0, T1, T2, T3, T4, T5, T6, T7):

- Row S0: T0 contains I₀ (black box).
- Row S1: T0 contains B₁ (blue box).
- Row S2: T0 contains P₀ (black box).
- Row S3: T0 contains B₁ (blue box).
- Row S4: T0 contains P₀ (black box).
- Row S5: T0 contains 5 (gray box).
- Row S6: T0 contains 6 (gray box).
- Row S7: T0 contains 7 (gray box).

Right Section (Columns T8, T9, T10, T11):

- Row S0: T8 contains 15 (gray box).
- Row S1: T8 contains 23 (gray box).
- Row S2: T8 contains 31 (gray box).
- Row S3: T8 contains 39 (gray box).
- Row S4: T8 contains 47 (gray box).
- Row S5: T8 contains 55 (gray box).
- Row S6: T8 contains 63 (gray box).
- Row S7: T8 contains 71 (gray box).

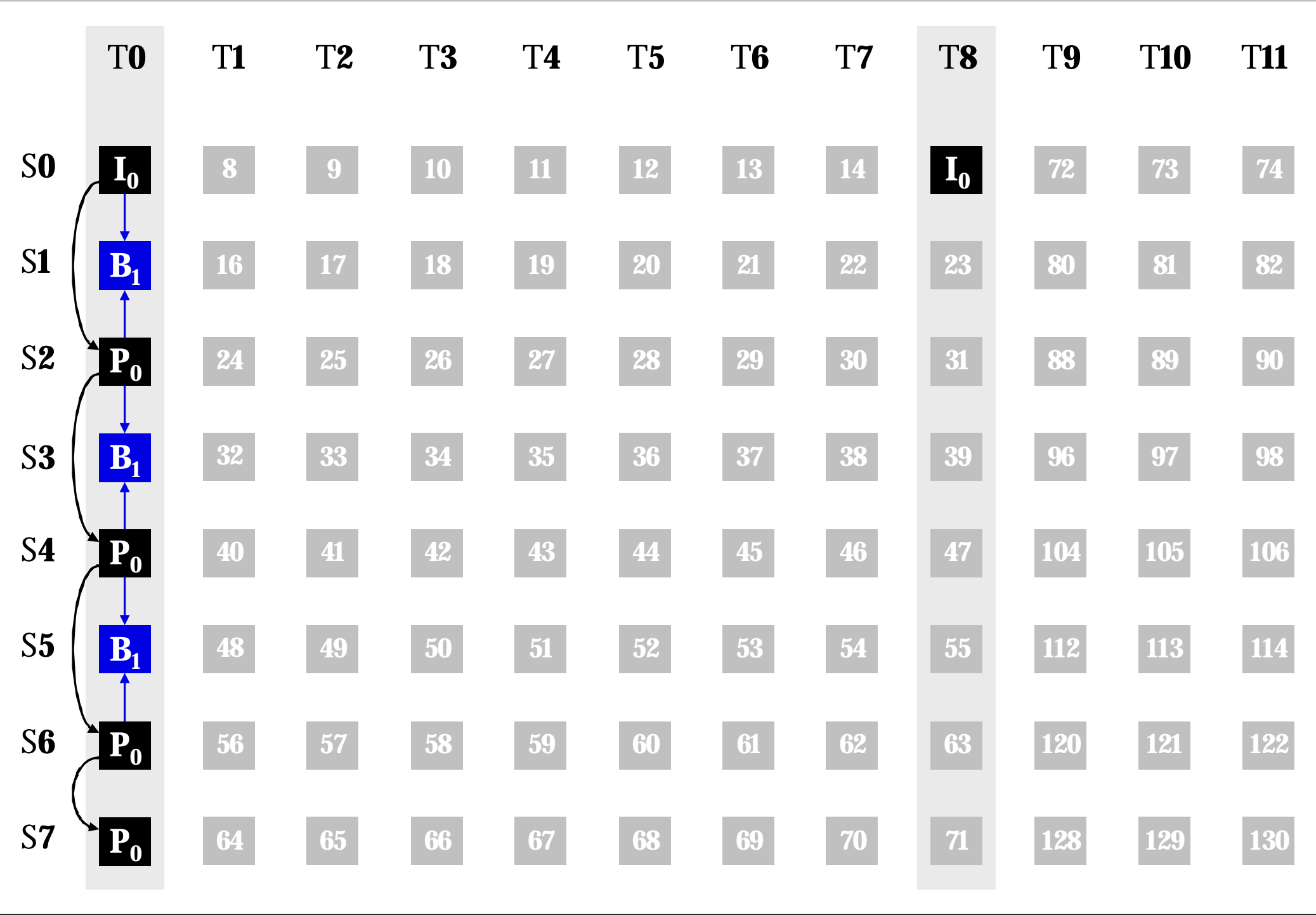
Connections:

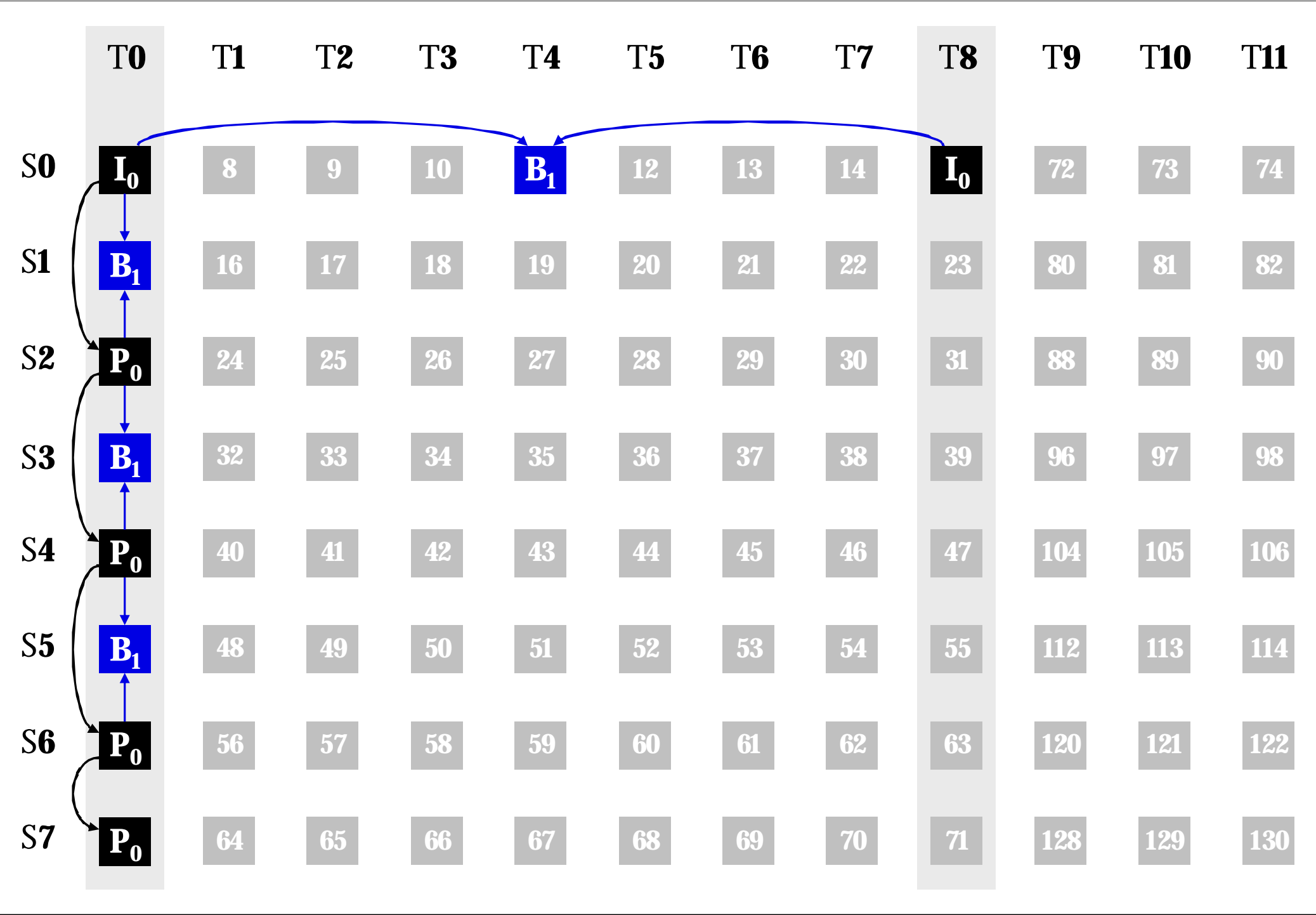
- A curved arrow connects S0 to S2.
- A curved arrow connects S3 to S4.
- Vertical arrows connect S0 to S1 and S2 to S3.
- Vertical arrows connect S1 to S2 and S3 to S4.

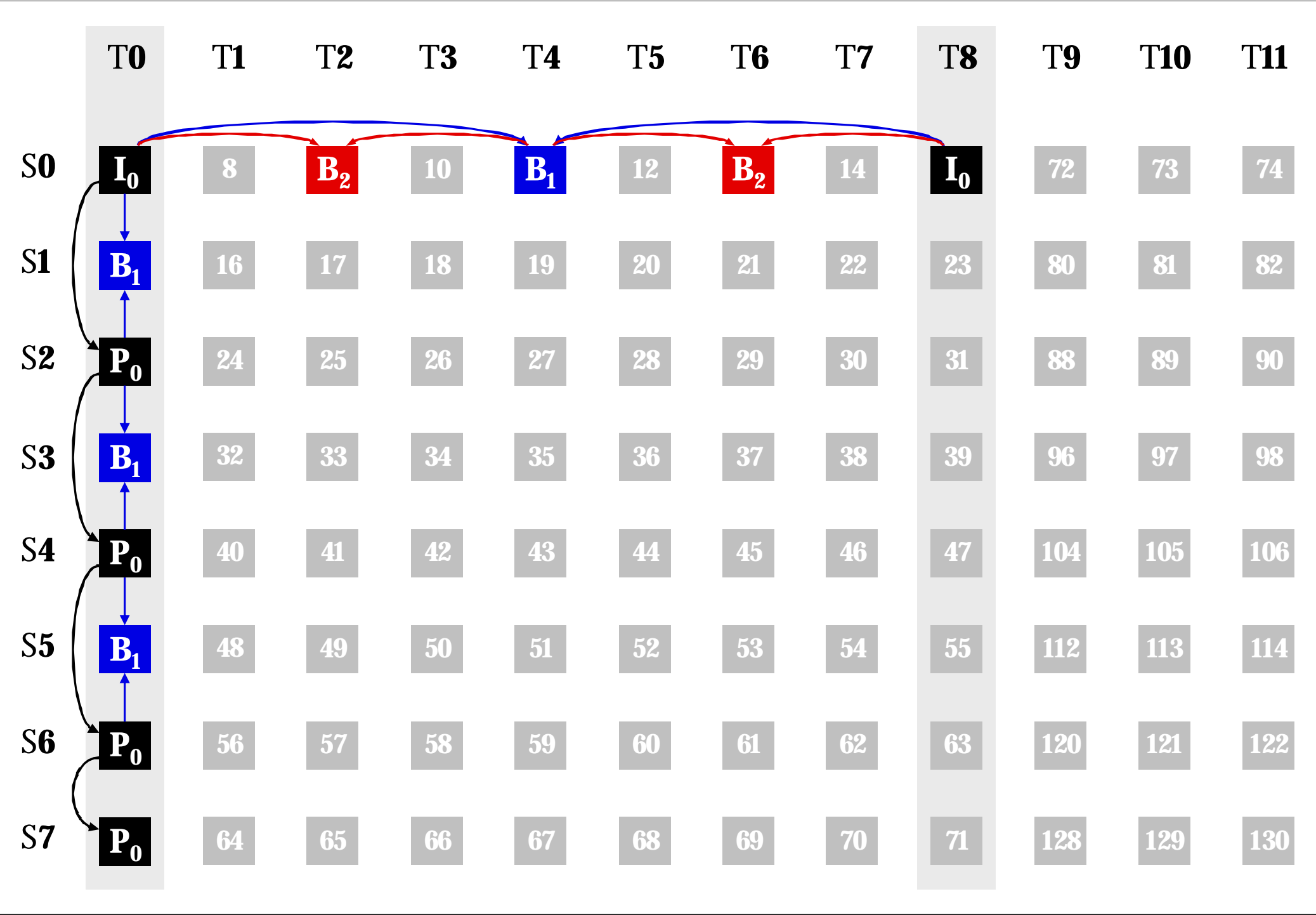
	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
S0	I ₀	8	9	10	11	12	13	14	15	72	73	74
S1	B ₁	16	17	18	19	20	21	22	23	80	81	82
S2	P ₀	24	25	26	27	28	29	30	31	88	89	90
S3	B ₁	32	33	34	35	36	37	38	39	96	97	98
S4	P ₀	40	41	42	43	44	45	46	47	104	105	106
S5	5	48	49	50	51	52	53	54	55	112	113	114
S6	P ₀	56	57	58	59	60	61	62	63	120	121	122
S7	7	64	65	66	67	68	69	70	71	128	129	130

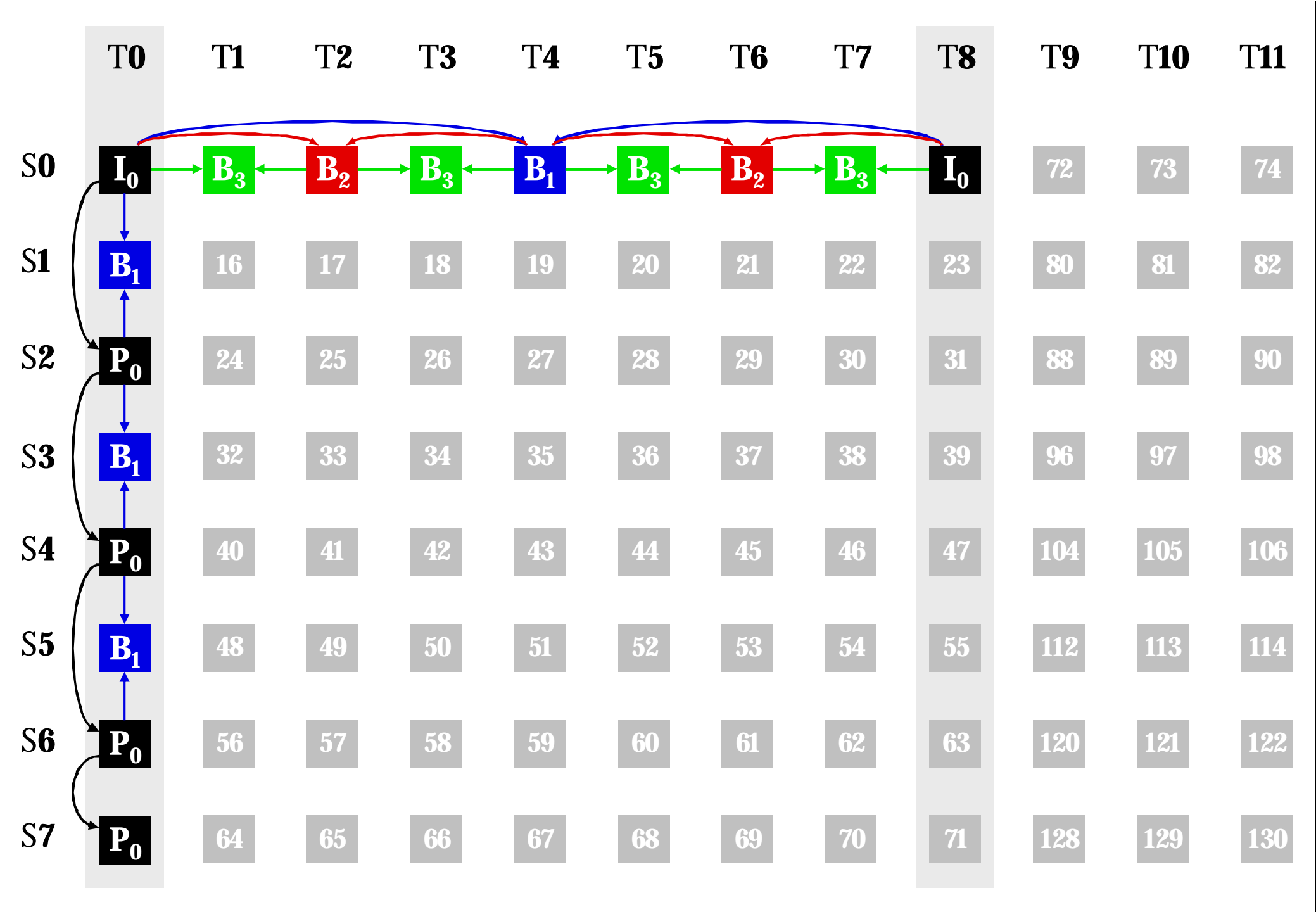
	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
S0	I ₀	8	9	10	11	12	13	14	15	72	73	74
S1	B ₁	16	17	18	19	20	21	22	23	80	81	82
S2	P ₀	24	25	26	27	28	29	30	31	88	89	90
S3	B ₁	32	33	34	35	36	37	38	39	96	97	98
S4	P ₀	40	41	42	43	44	45	46	47	104	105	106
S5	B ₁	48	49	50	51	52	53	54	55	112	113	114
S6	P ₀	56	57	58	59	60	61	62	63	120	121	122
S7	7	64	65	66	67	68	69	70	71	128	129	130

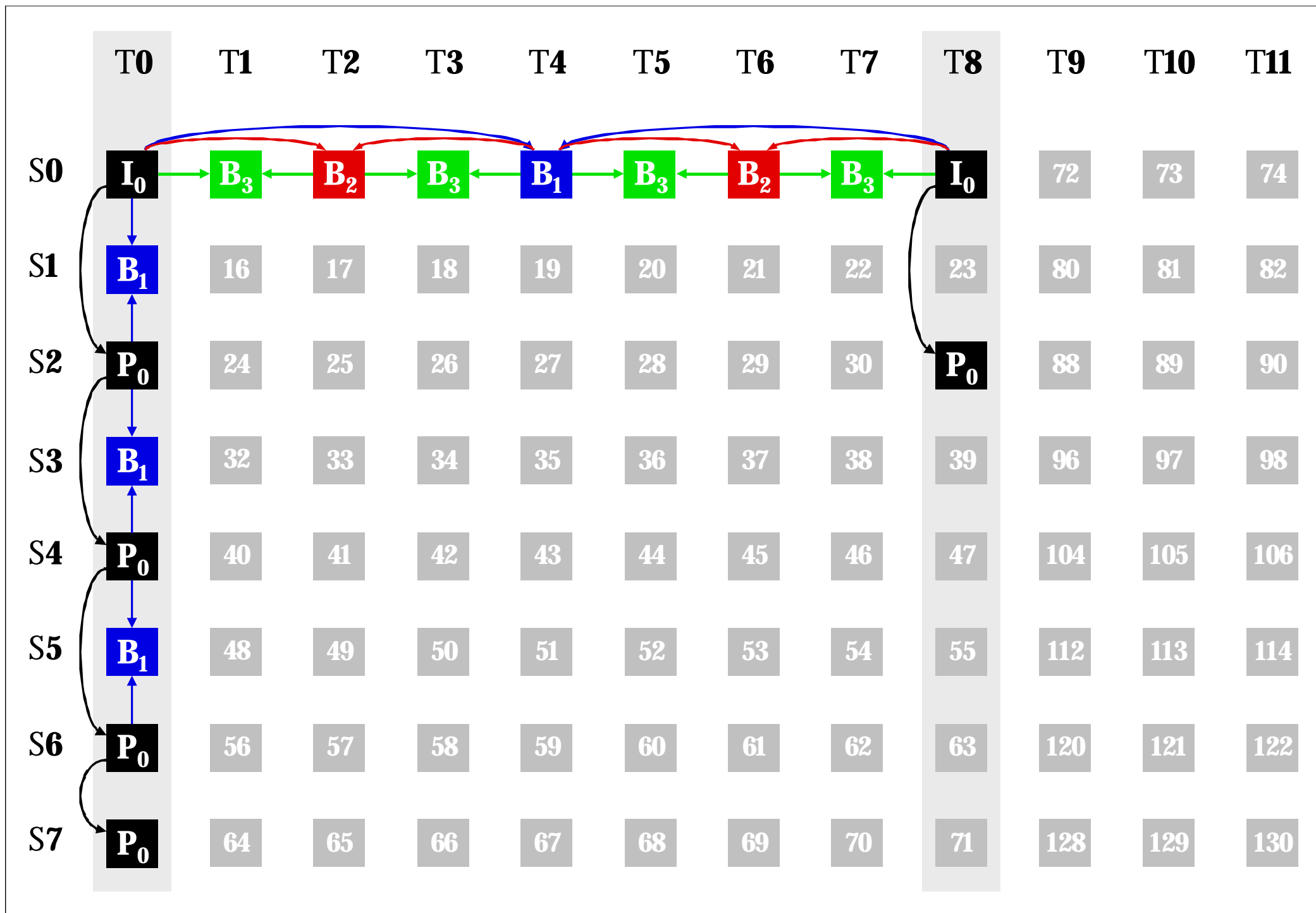
	T0	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
S0	I ₀	8	9	10	11	12	13	14	15	72	73	74
S1	B ₁	16	17	18	19	20	21	22	23	80	81	82
S2	P ₀	24	25	26	27	28	29	30	31	88	89	90
S3	B ₁	32	33	34	35	36	37	38	39	96	97	98
S4	P ₀	40	41	42	43	44	45	46	47	104	105	106
S5	B ₁	48	49	50	51	52	53	54	55	112	113	114
S6	P ₀	56	57	58	59	60	61	62	63	120	121	122
S7	P ₀	64	65	66	67	68	69	70	71	128	129	130

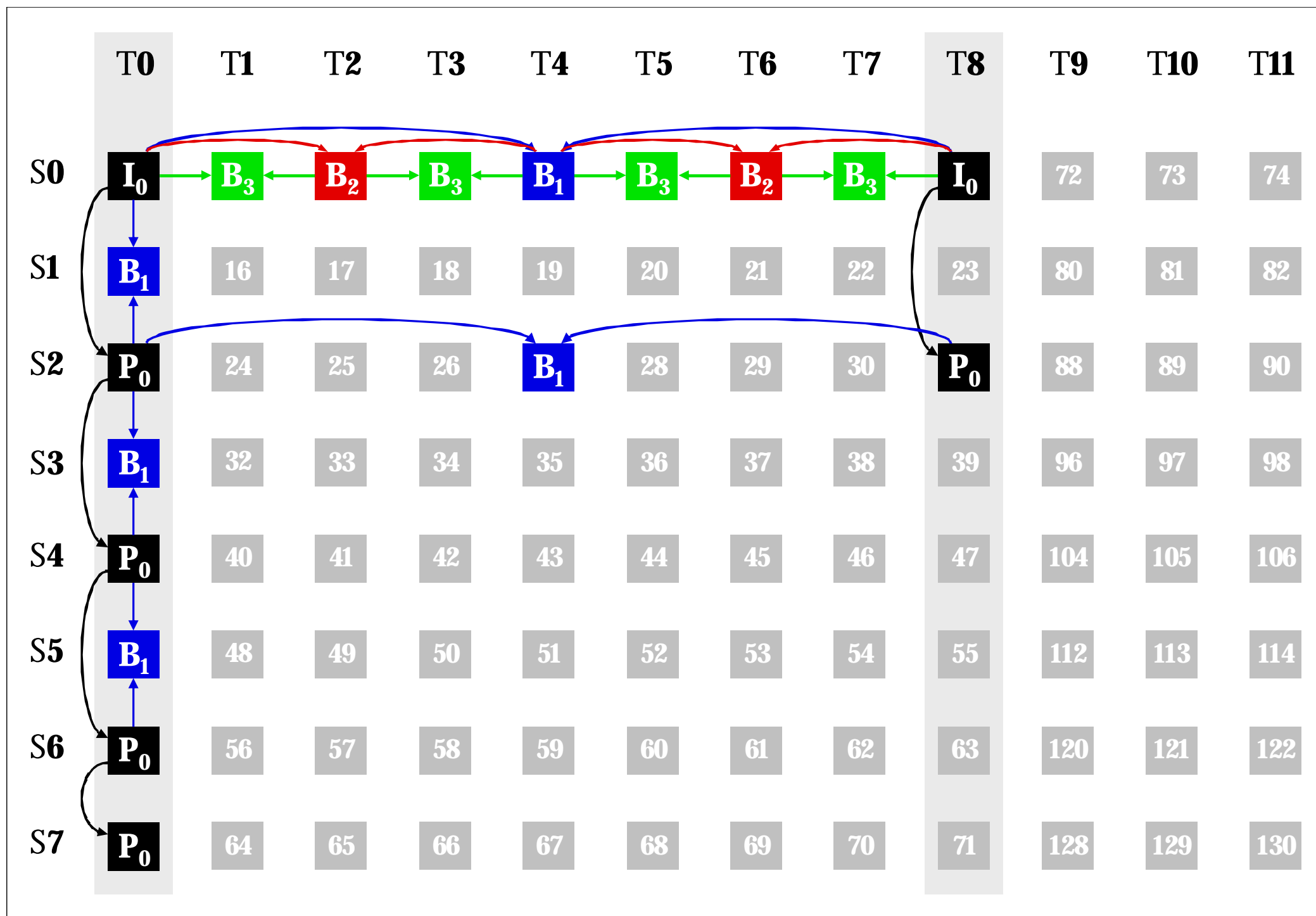


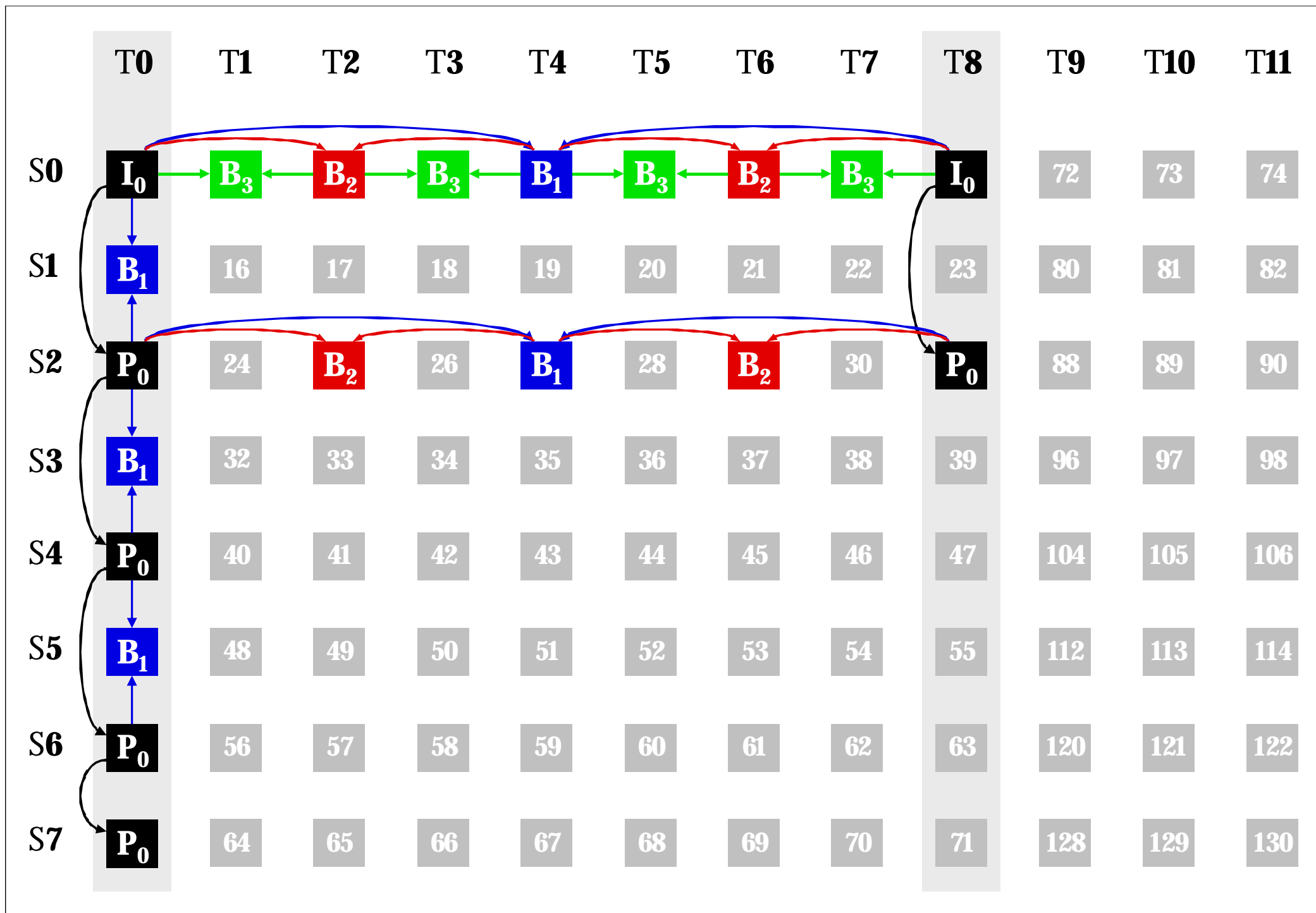


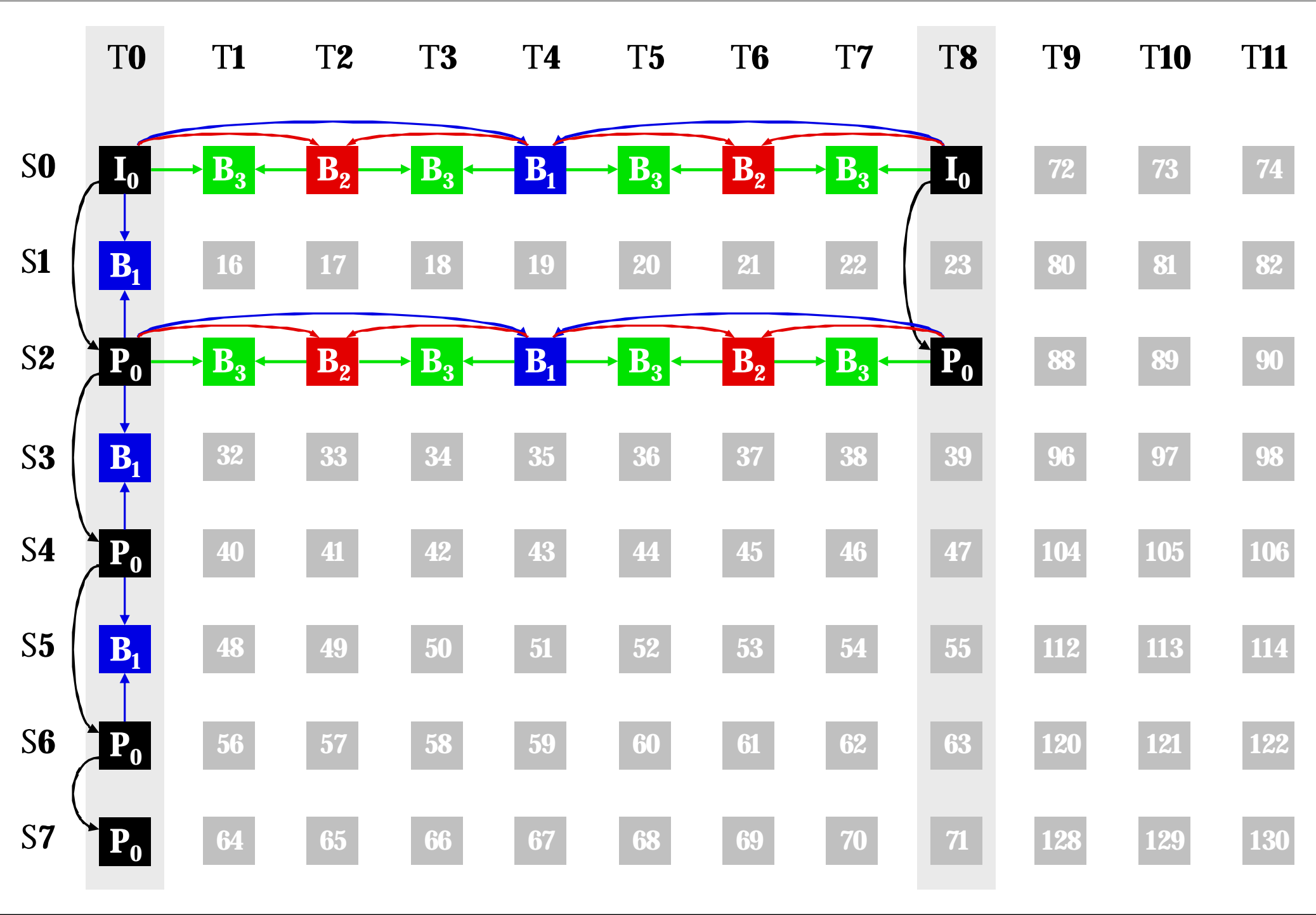


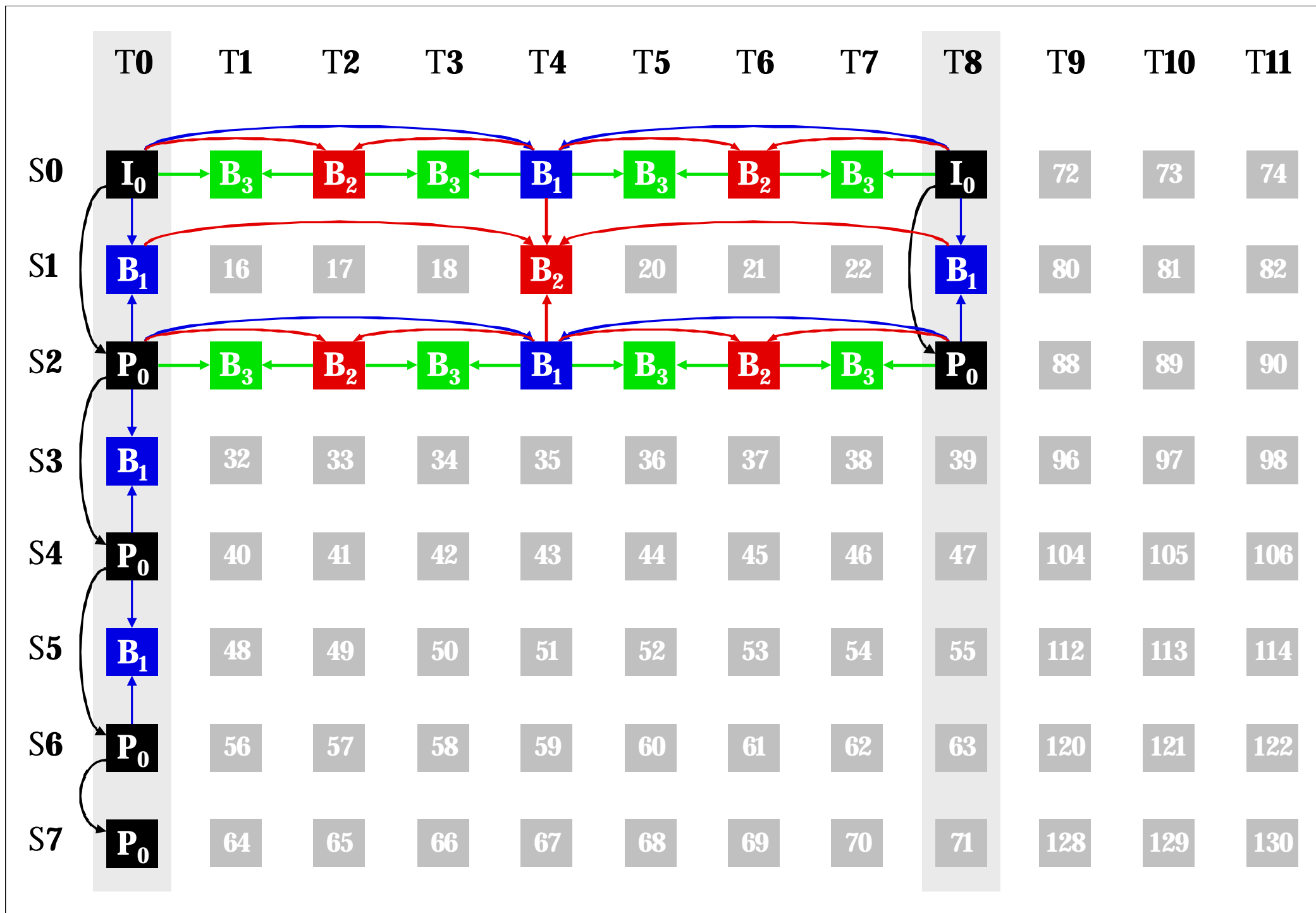


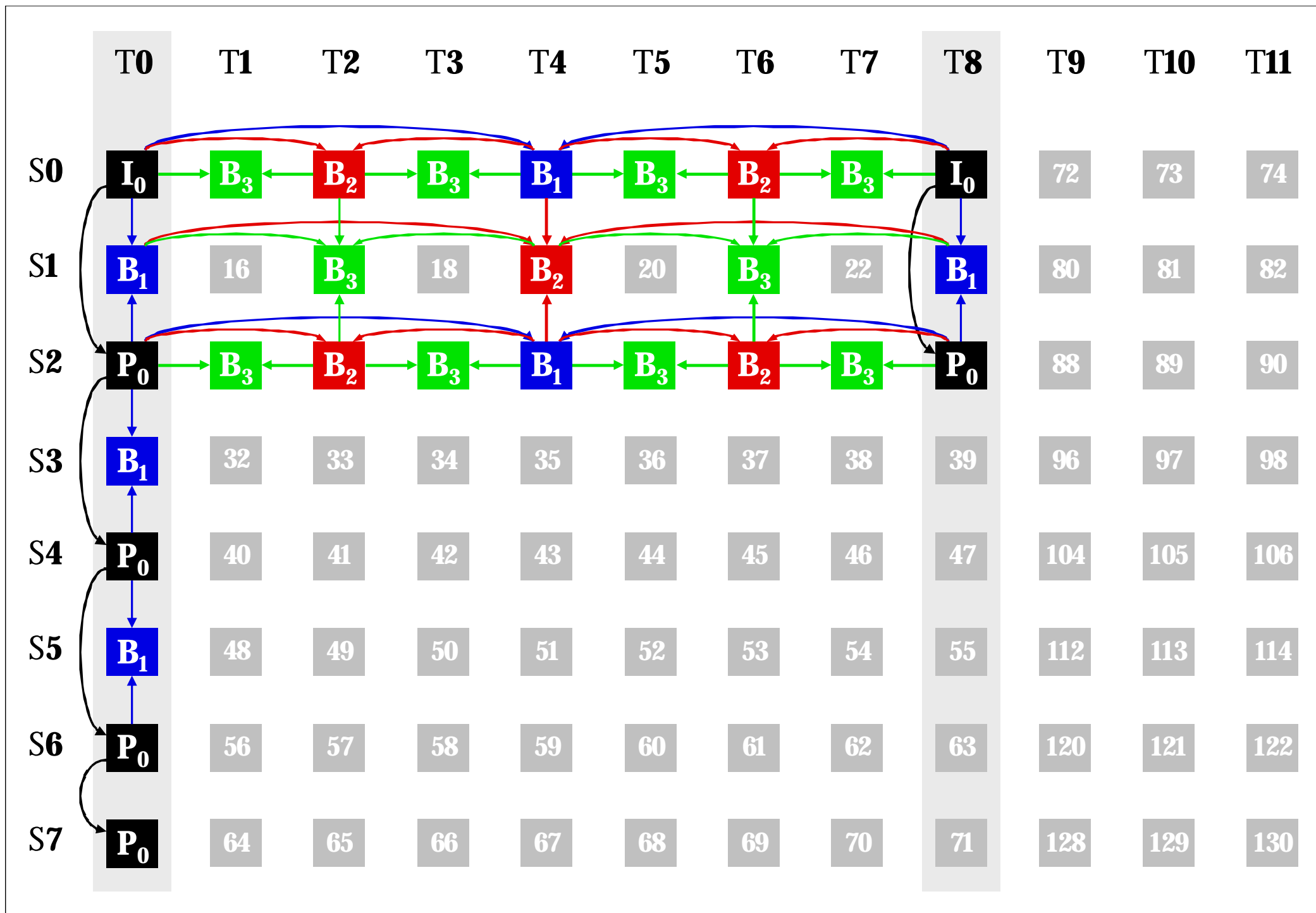


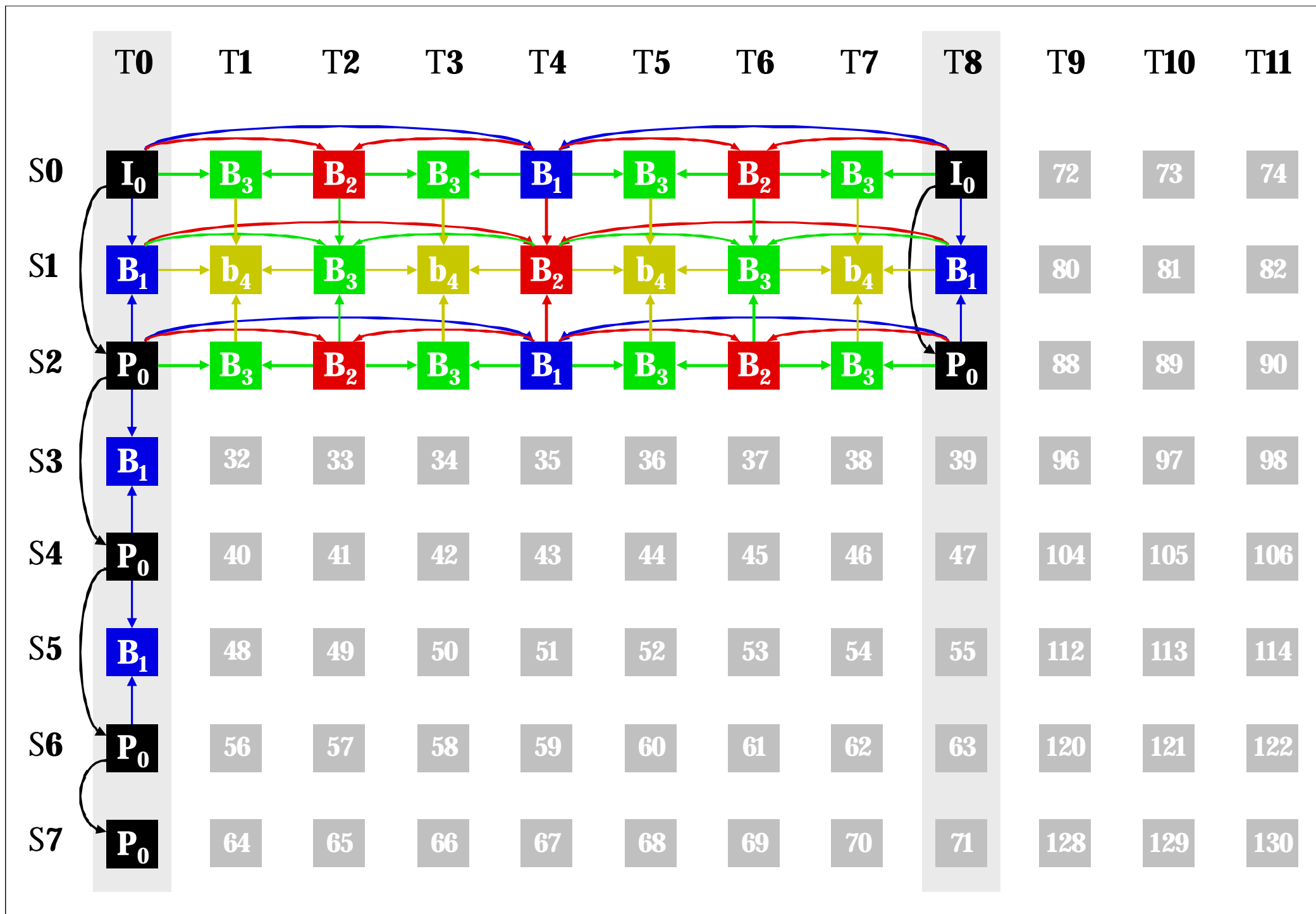


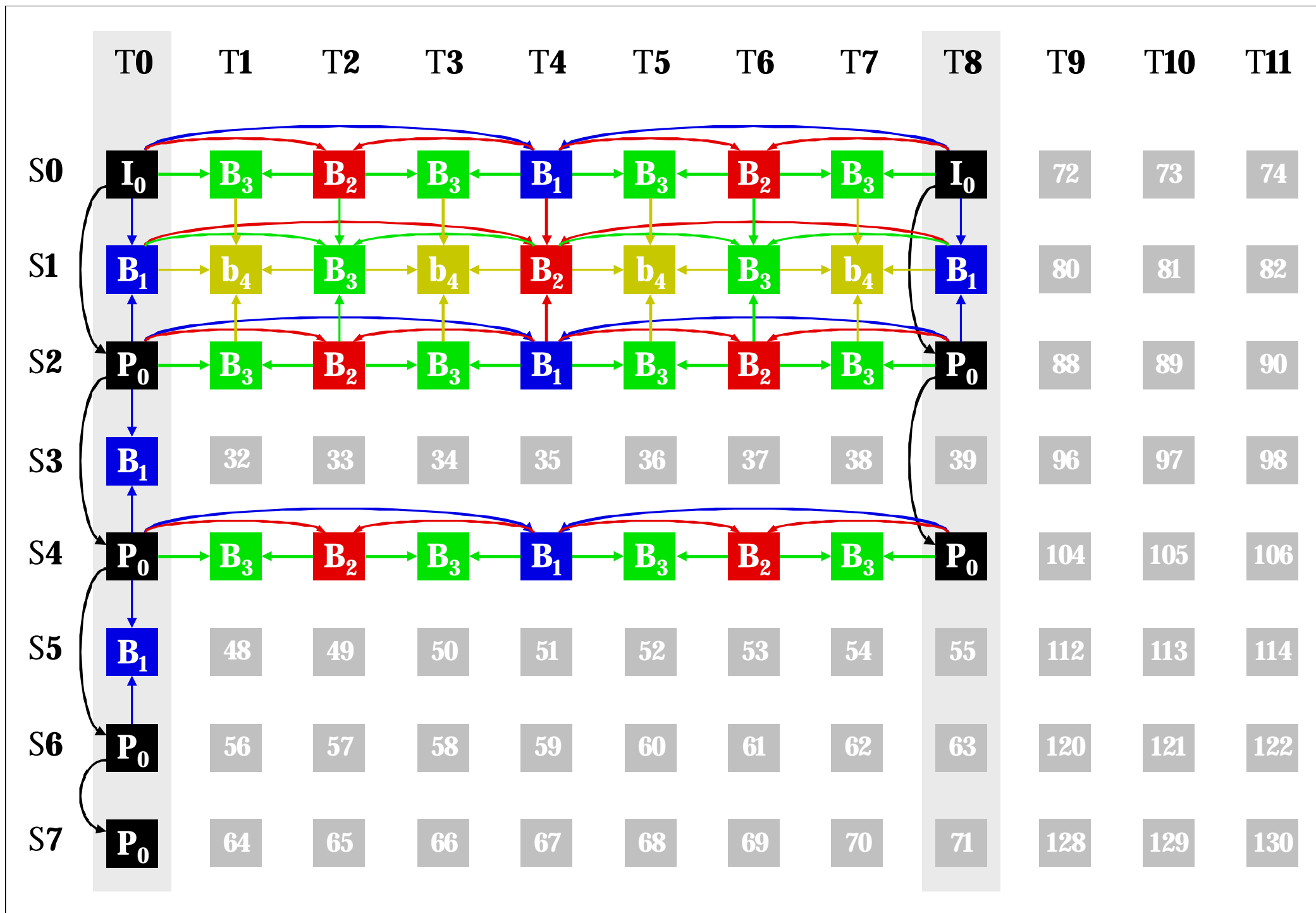


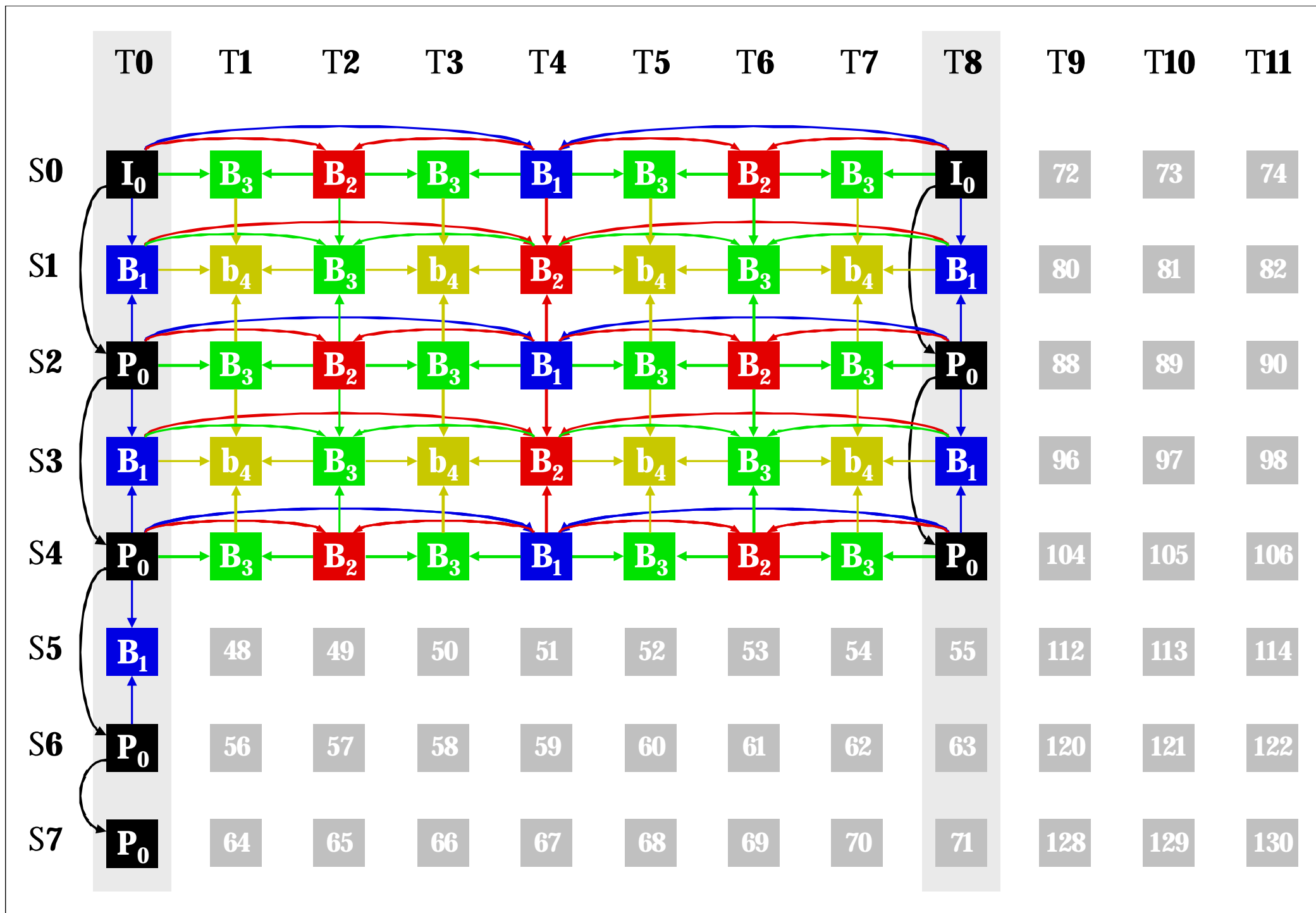


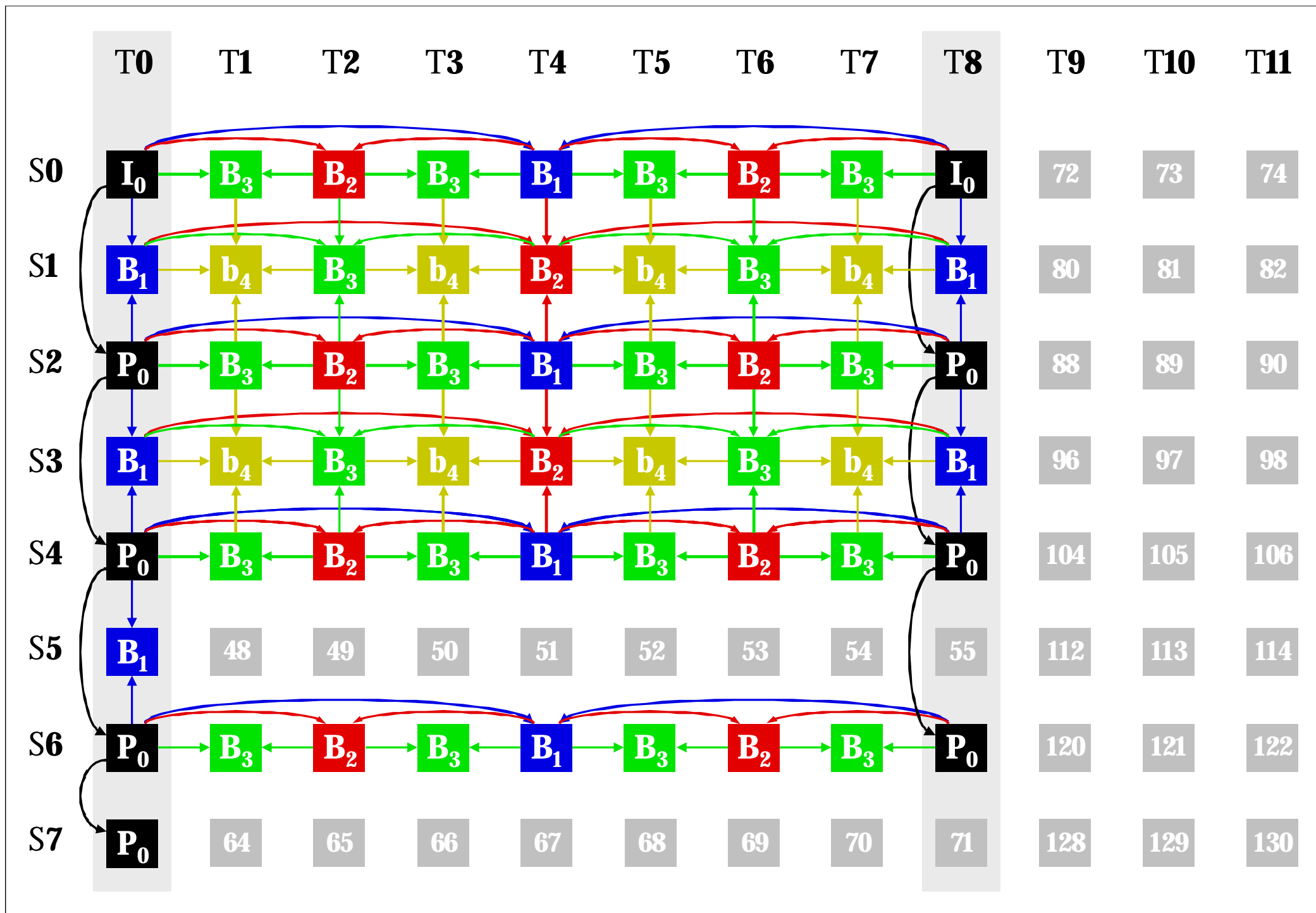


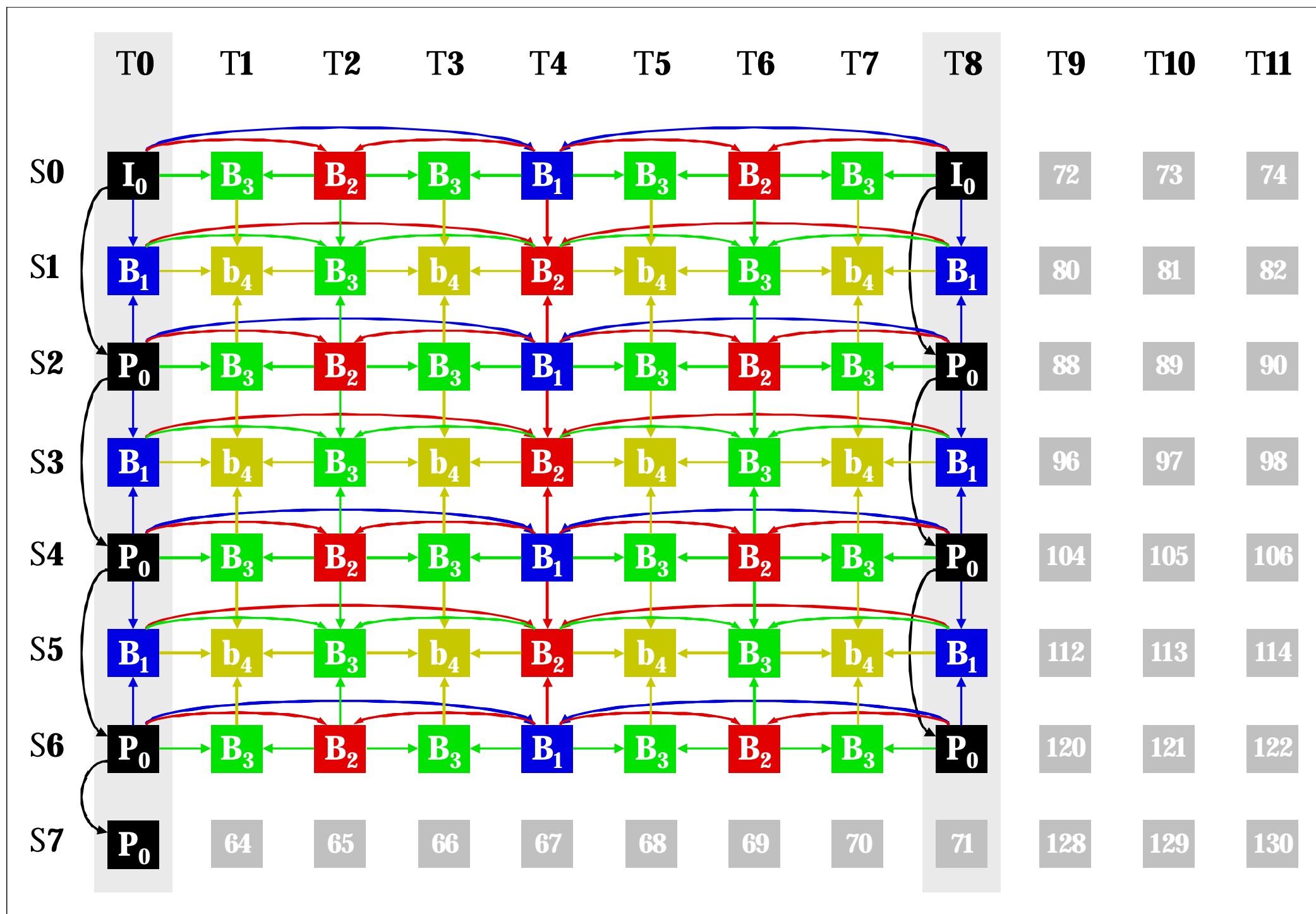


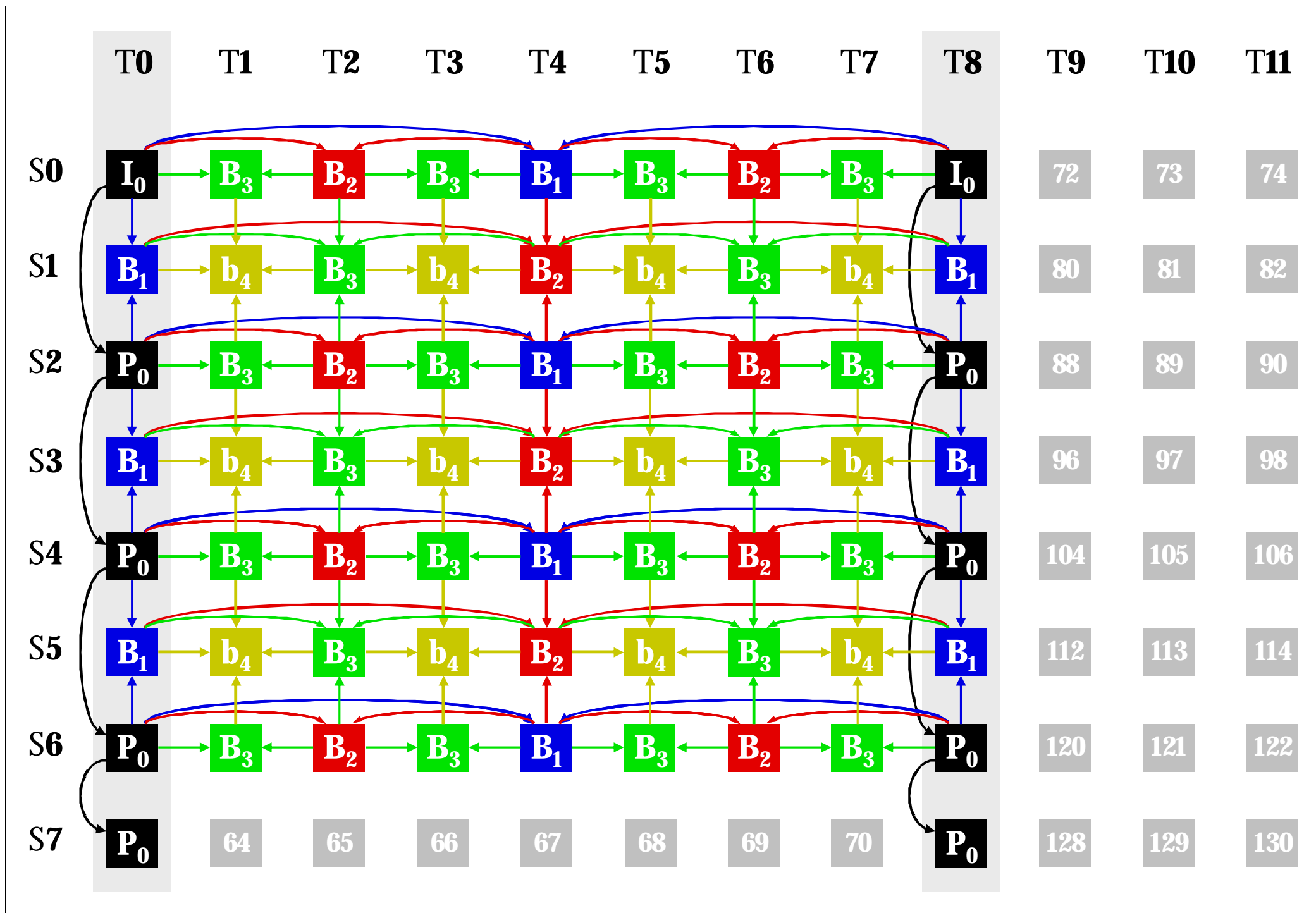


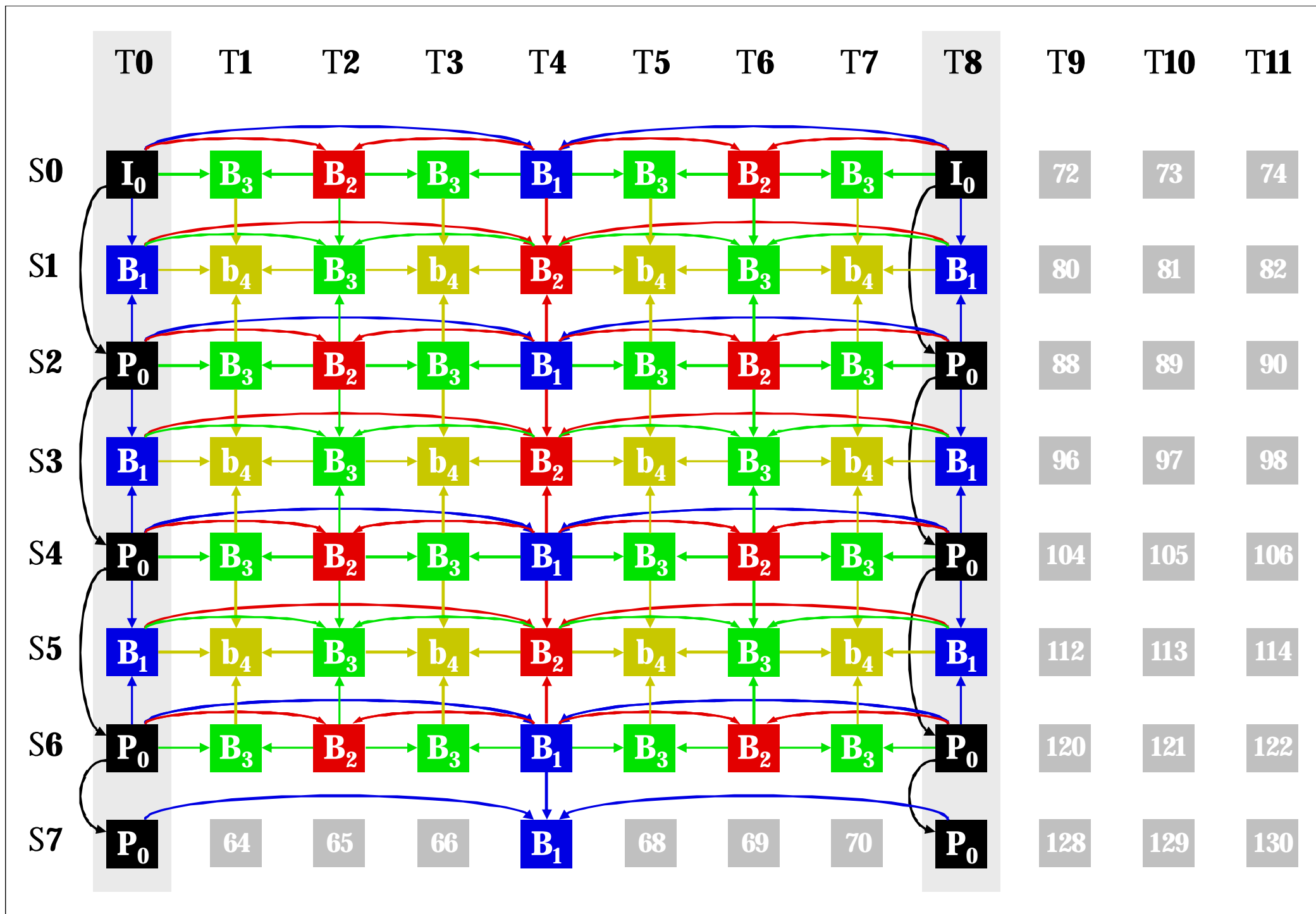


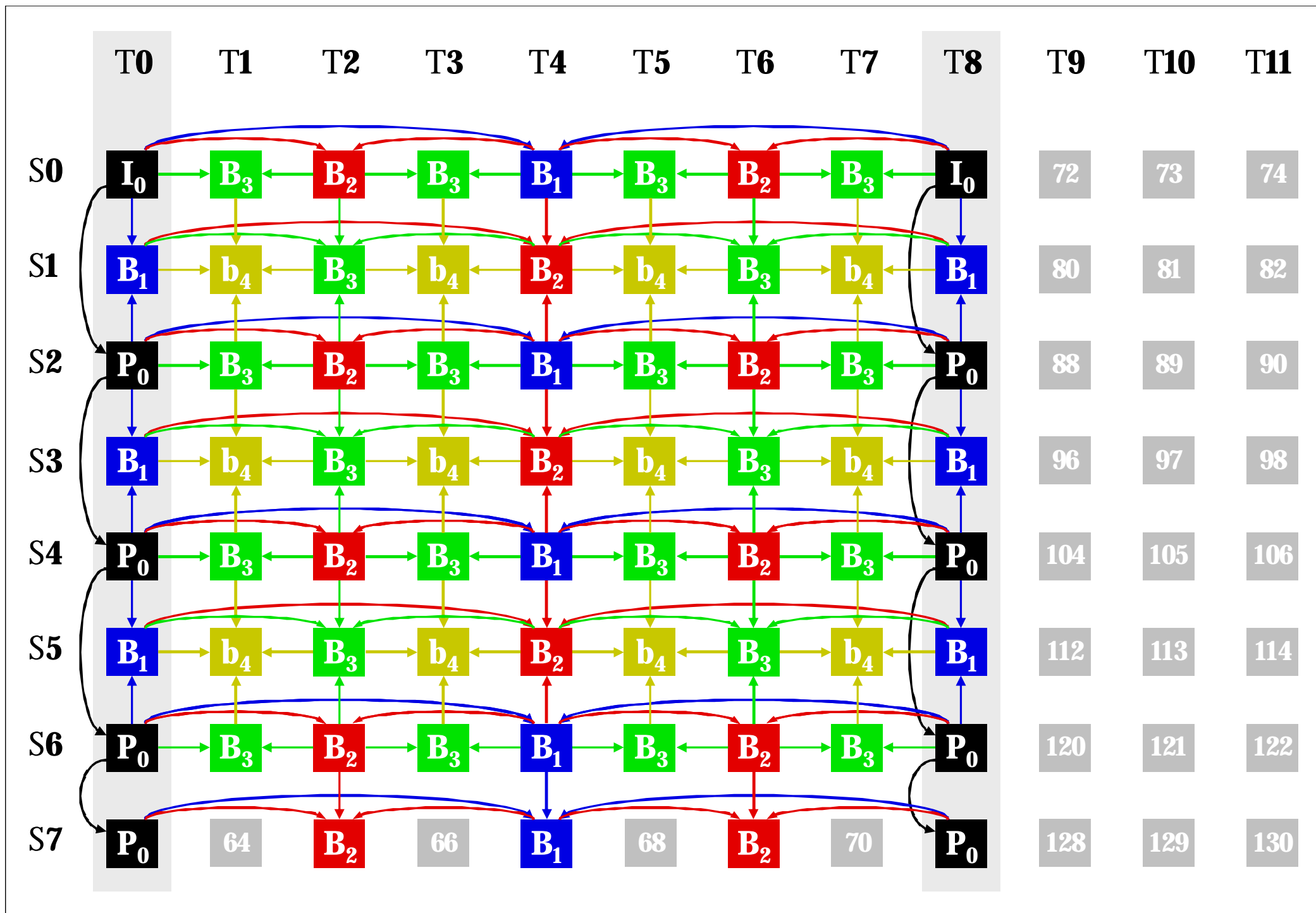


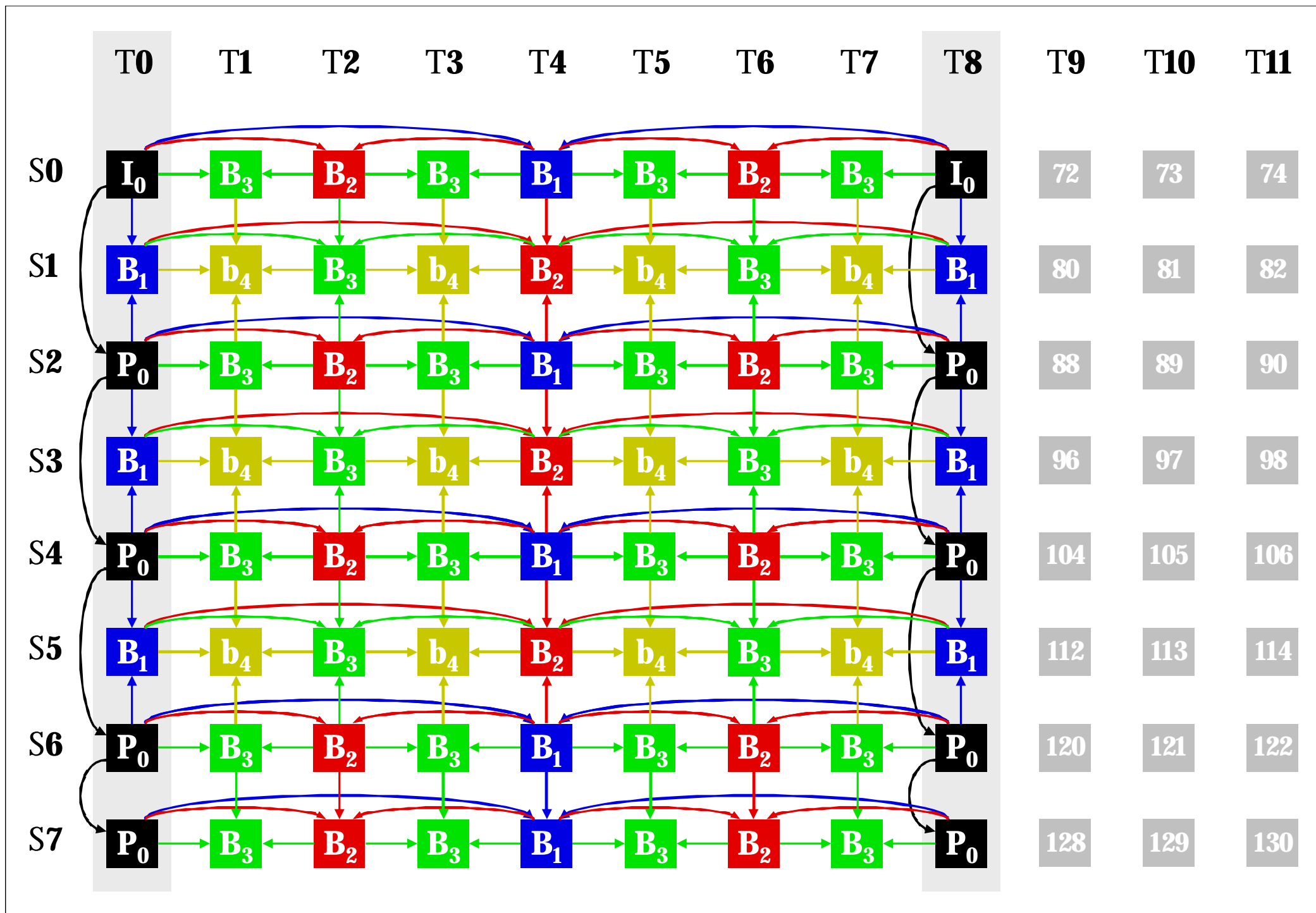


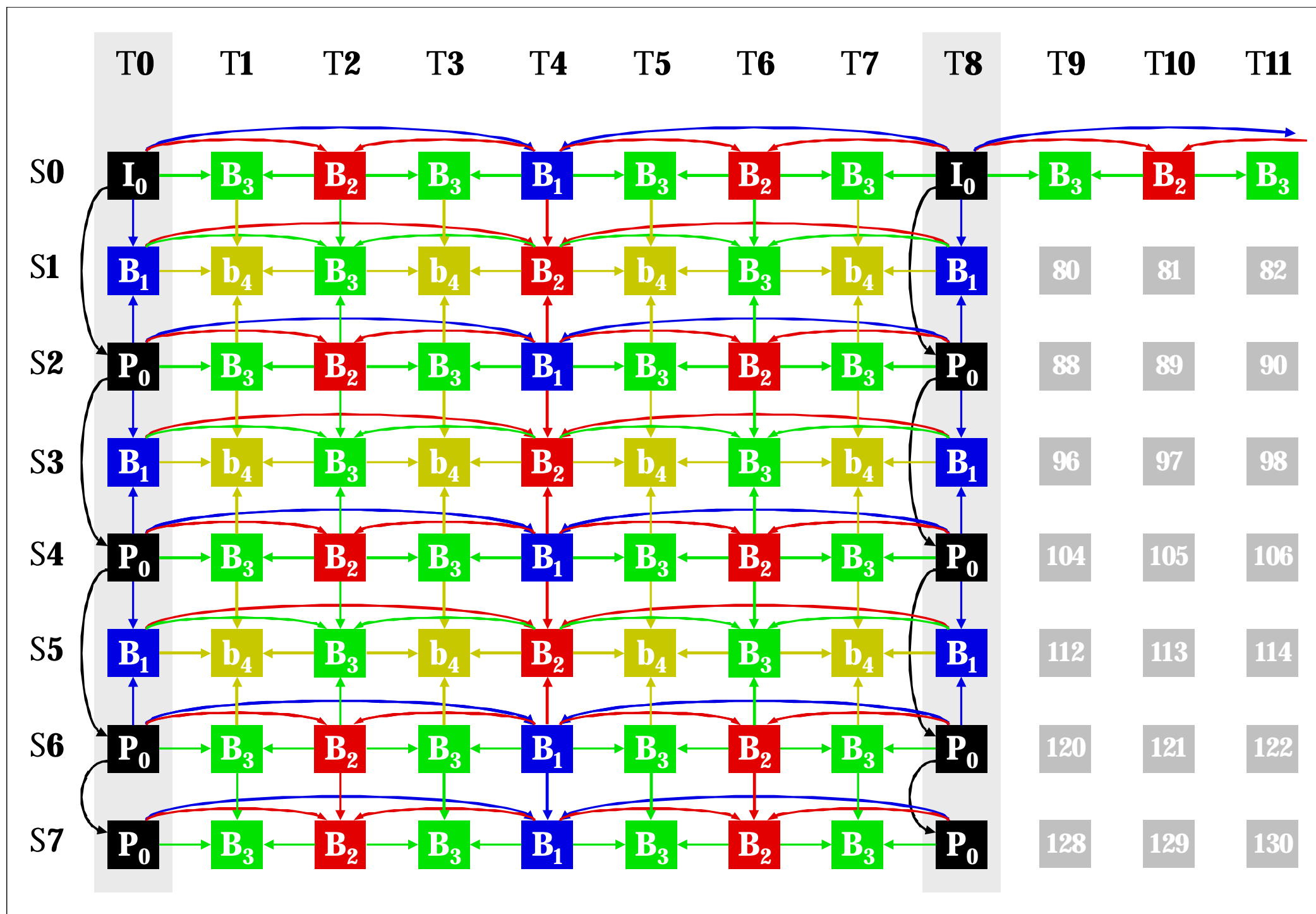


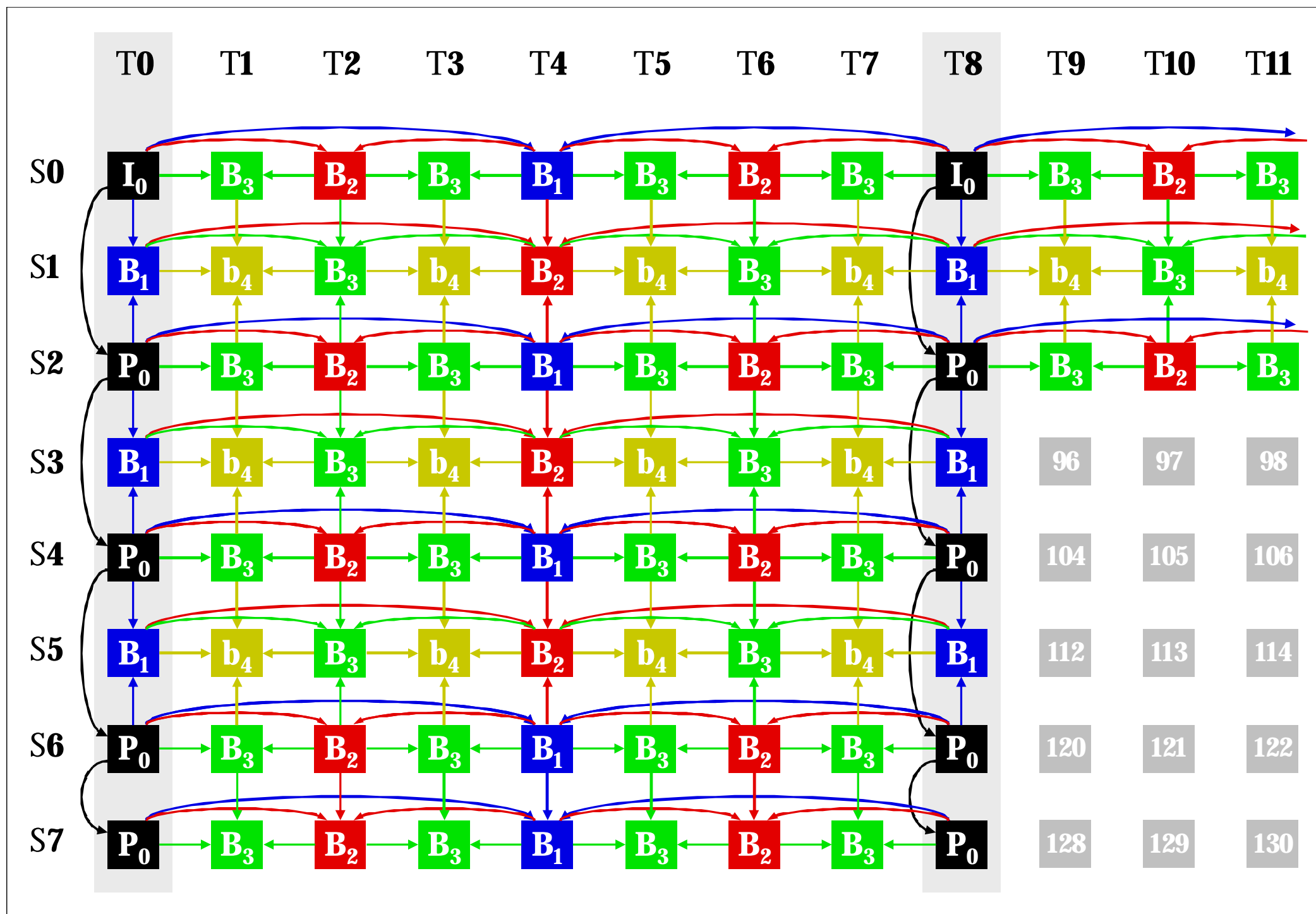


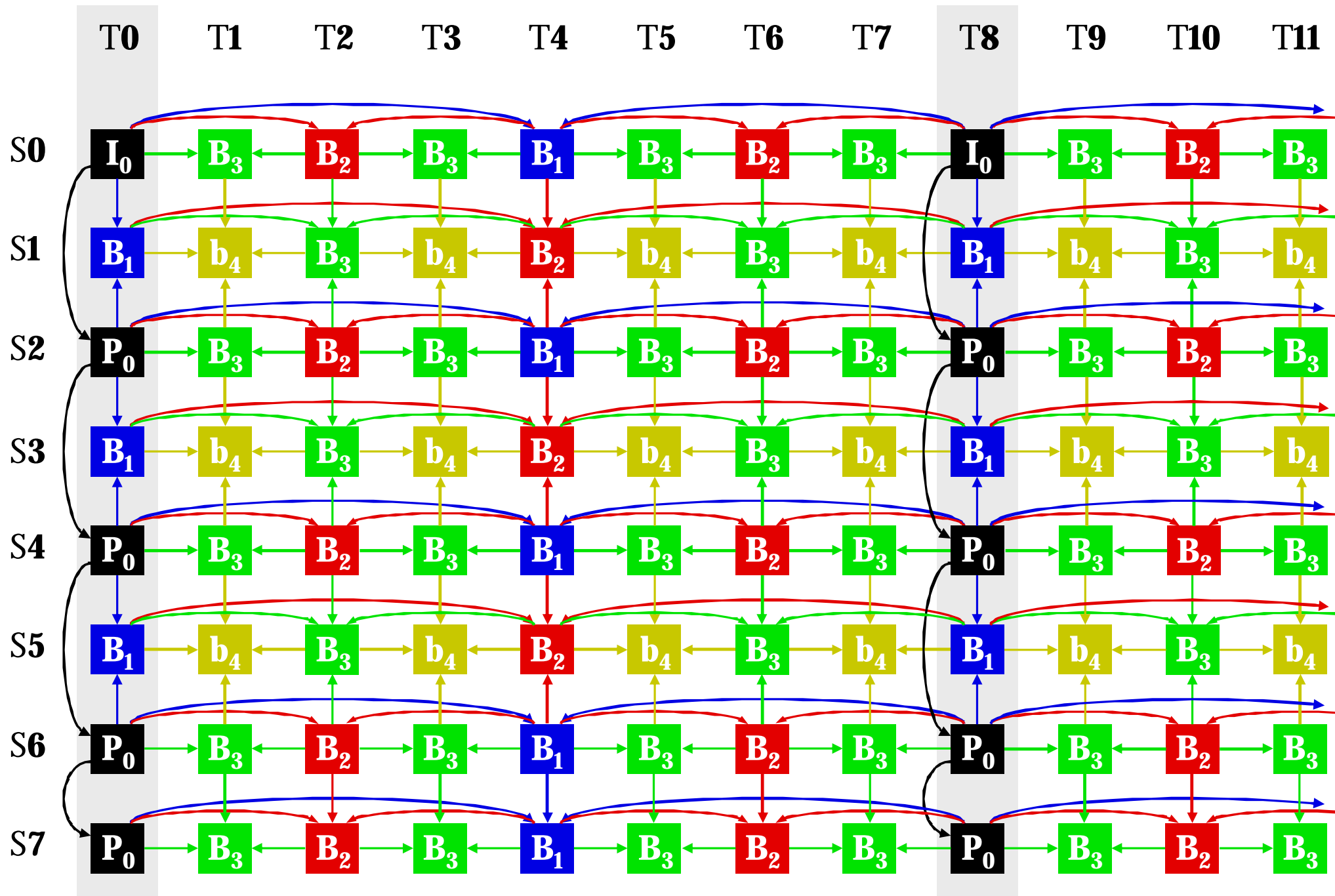












Summary & Conclusion

- Fully H.264/MPEG4-AVC compatible bitstream
- Decoder needs to de-interleave the decoded pictures into MV video streams
- Sophisticated temporal+inter-view prediction structures can be implemented based on hierarchical B-picture syntax
- Performs best among all MVC proposals so far



MVC Core Experiments

Specific tools under investigation in Core Experiments:

- **Illumination compensation**
- **View interpolation based on depth**
 - a) **calculated at decoder (+ correction data encoding)**
 - b) **sending depth**
- **Disparity vector prediction**

MVC Core Experiments

High-level syntax for:

- **Improved random access**
- **Low delay**
- **Memory optimization**

Thank you!

smolic@hhi.de

<http://ip.hhi.de/~smolic/>