

Unity

Accelerating DNN Training Through Joint Optimization
of Algebraic Transformations and Parallelization

Colin Unger ^{*,1}

Sina Lin ⁶

Vinay Ramakrishnaiah ⁴

Jamaludin Mohd-Yusof ⁴

Jongsoo Park ³

Zhihao Jia ^{*,2,3}

Mandeep Baines ³

Nirmal Prajapati ⁴

Xi Luo ⁷

Misha Smelyanskiy ³

Wei Wu ^{4,5}

Carlos Efrain Quintero Narvaez ³

Pat McCormick ⁴

Dheevatsa Mudigere ³

Alex Aiken ¹



1



2



3



4



5



6



7

Unity

Accelerating DNN Training Through Joint Optimization
of Algebraic Transformations and Parallelization

Unity

Accelerating DNN Training Through Joint Optimization
of Algebraic Transformations and Parallelization

1.

2.

Unity

Accelerating DNN Training Through Joint Optimization
of Algebraic Transformations and Parallelization

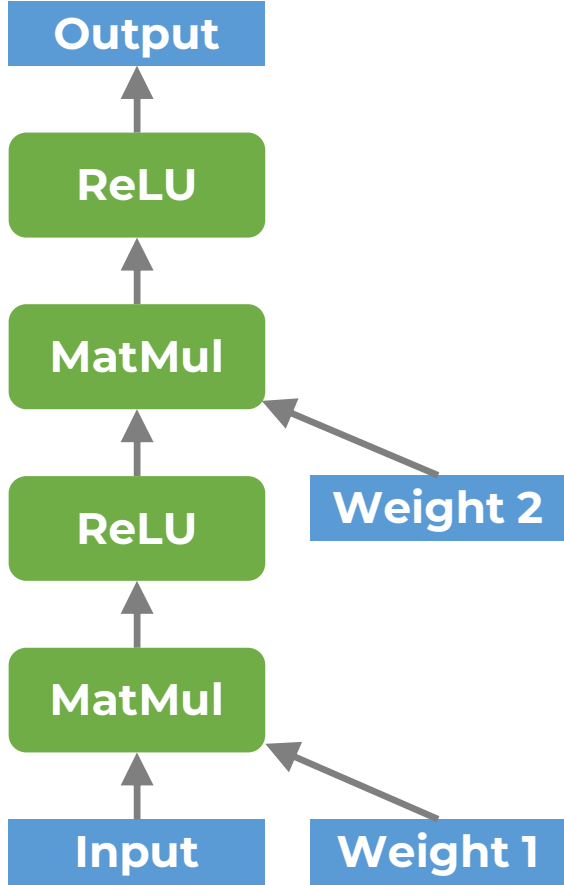
1. Algebraic Transformations
- 2.

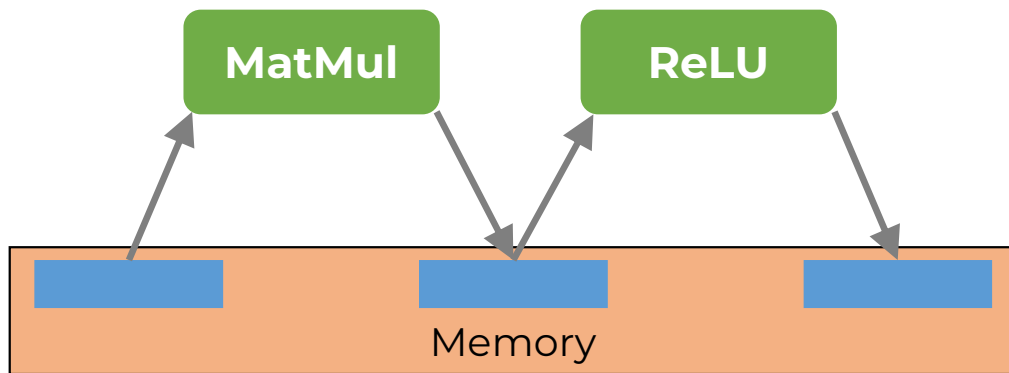
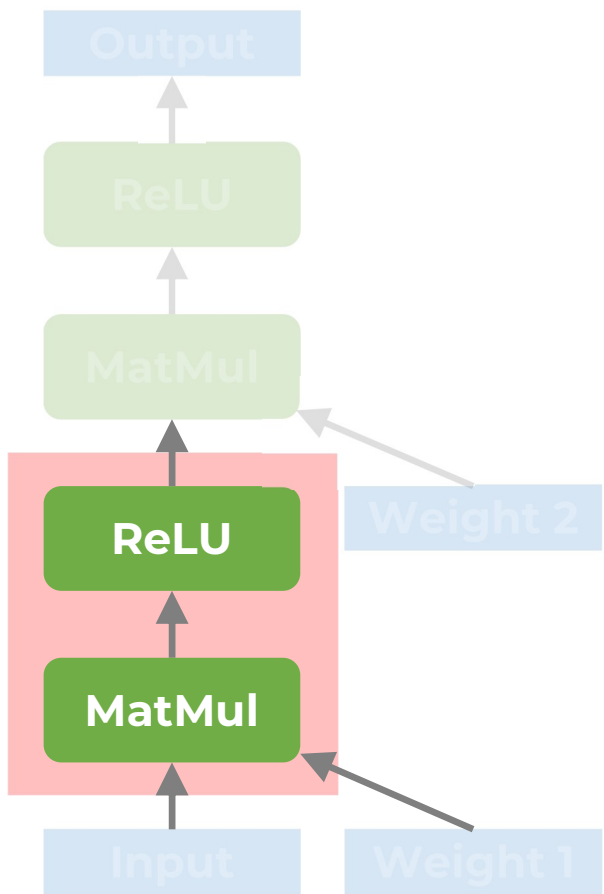
Unity

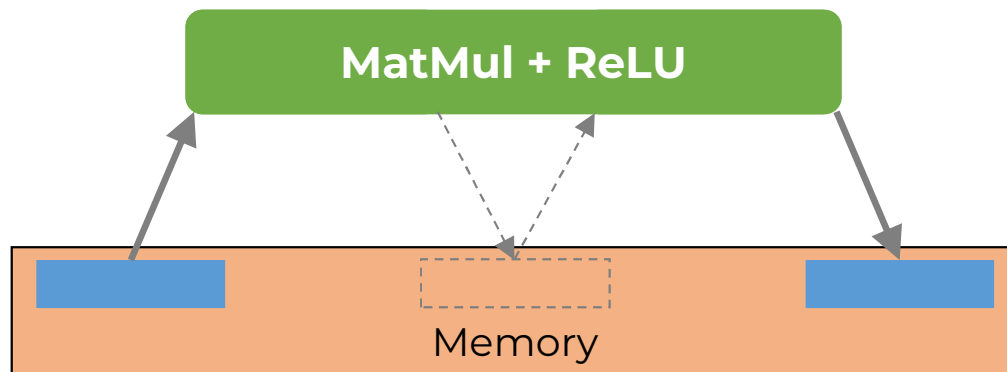
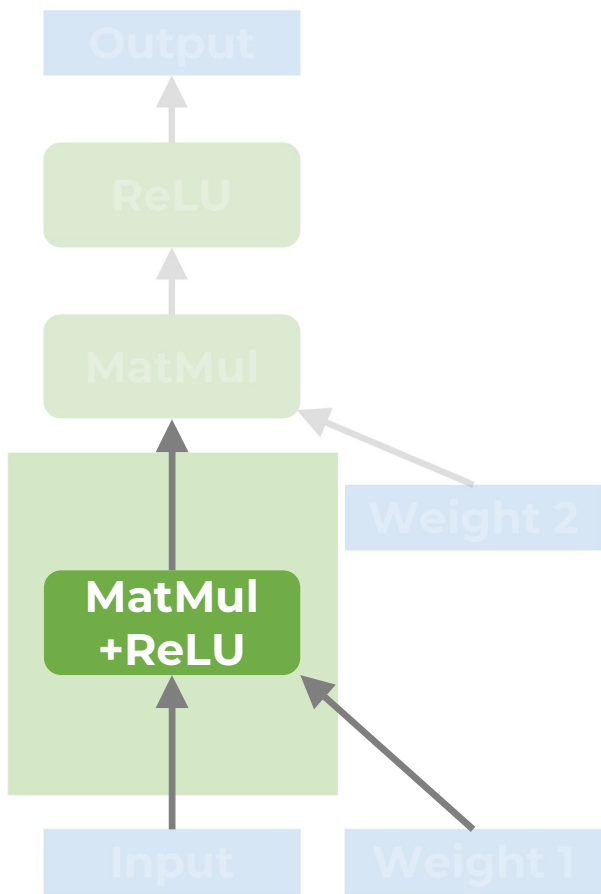
Accelerating DNN Training Through Joint Optimization
of Algebraic Transformations and Parallelization

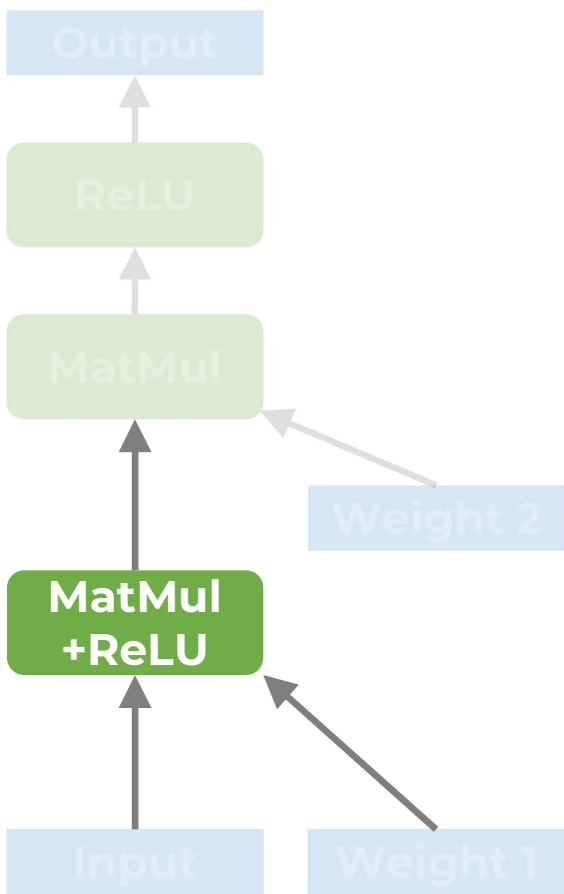
1. Algebraic Transformations
2. Parallelization

1. Algebraic Transformations

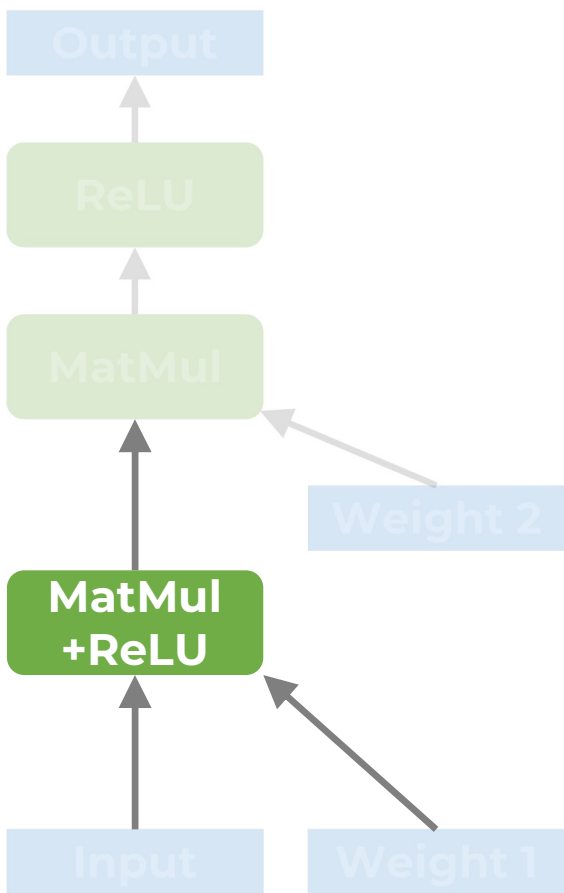




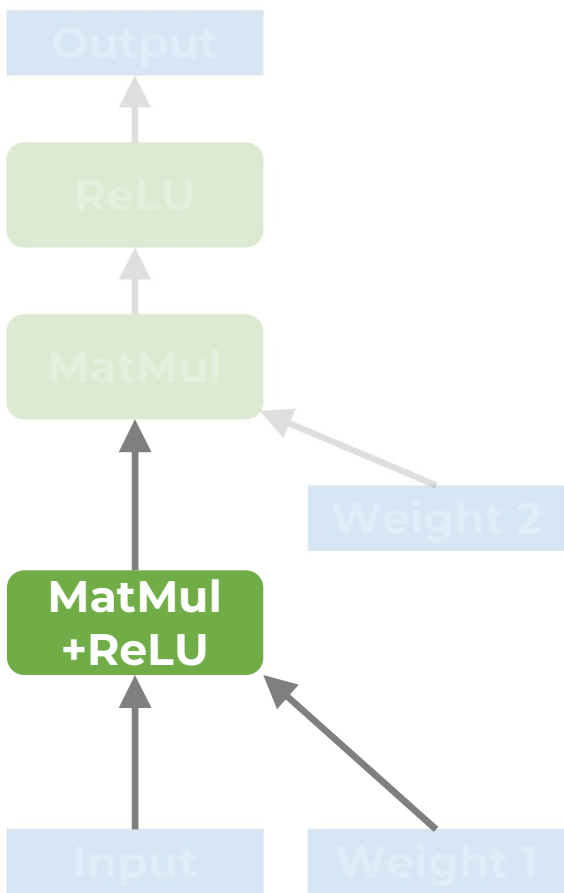




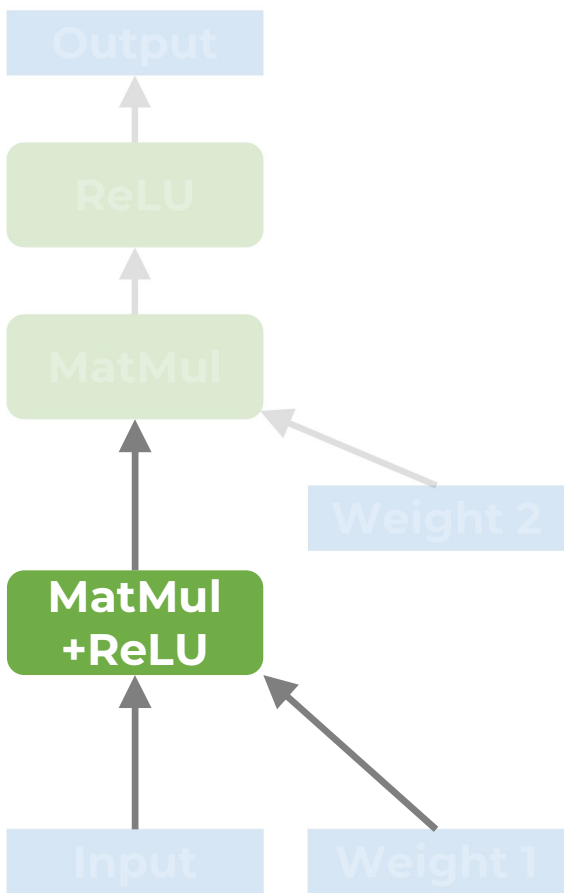
Operator Fusion



Operator Fusion
Operator Splitting



Operator Fusion
Operator Splitting
Operator Reordering



Operator Fusion
Operator Splitting
Operator Reordering
...

Operator Fusion
Operator Splitting
Operator Reordering
...

Operator Fusion
Operator Splitting
Operator Reordering
...

“Algebraic Transformations”

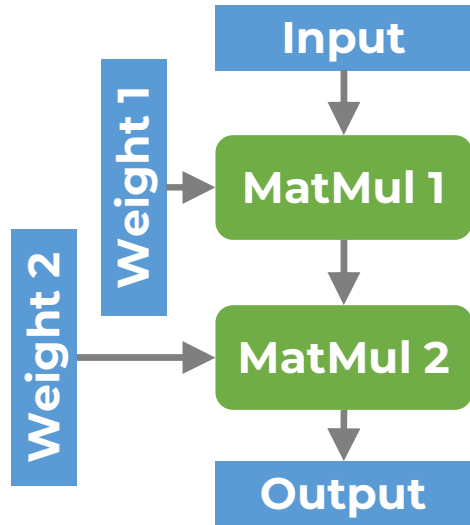
Operator Fusion
Operator Splitting
Operator Reordering
...

1. Algebraic Transformations

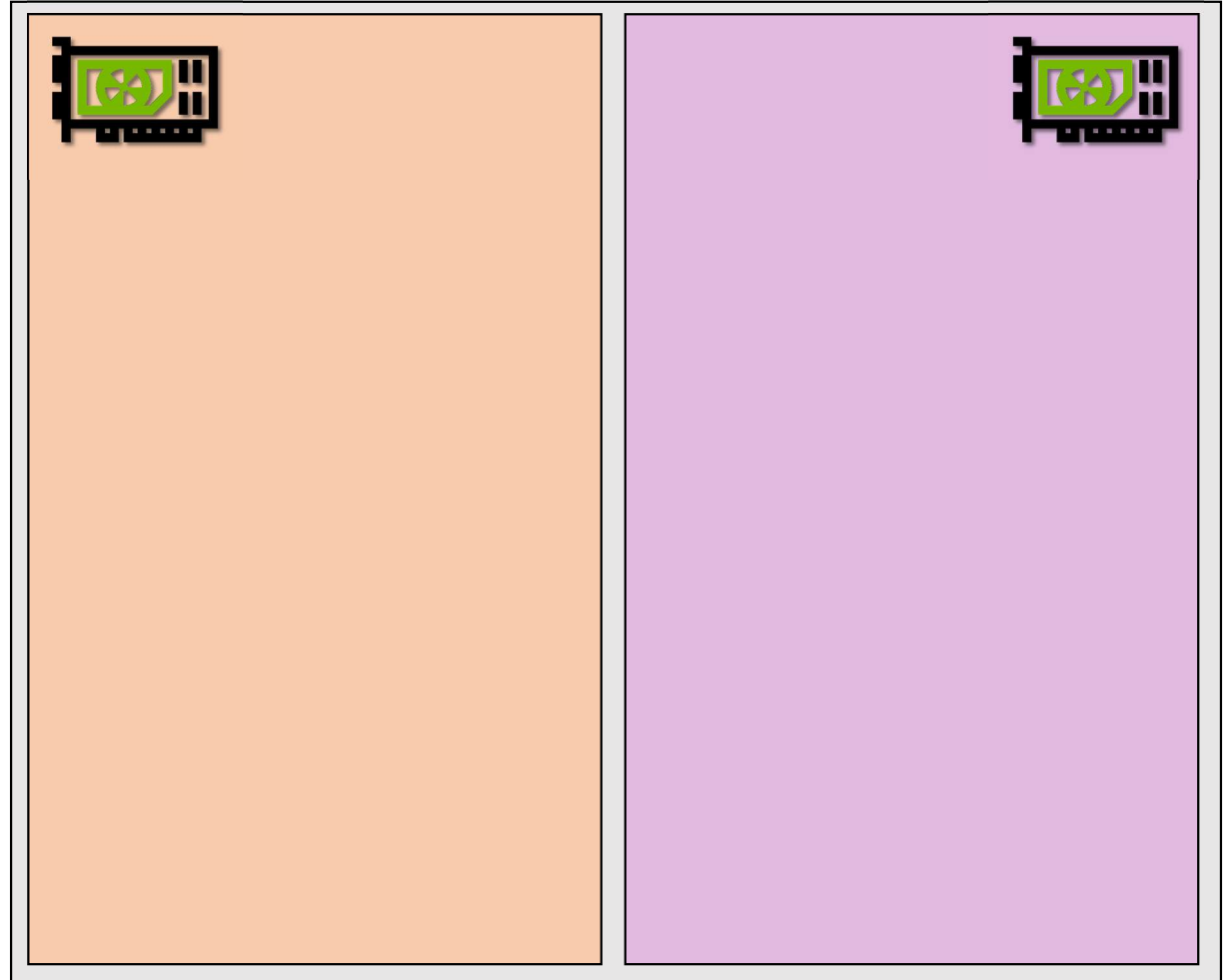
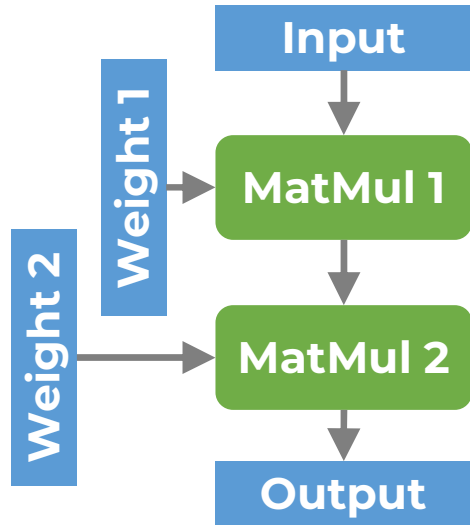
1. Algebraic Transformations
2. Parallelization

Data Parallel

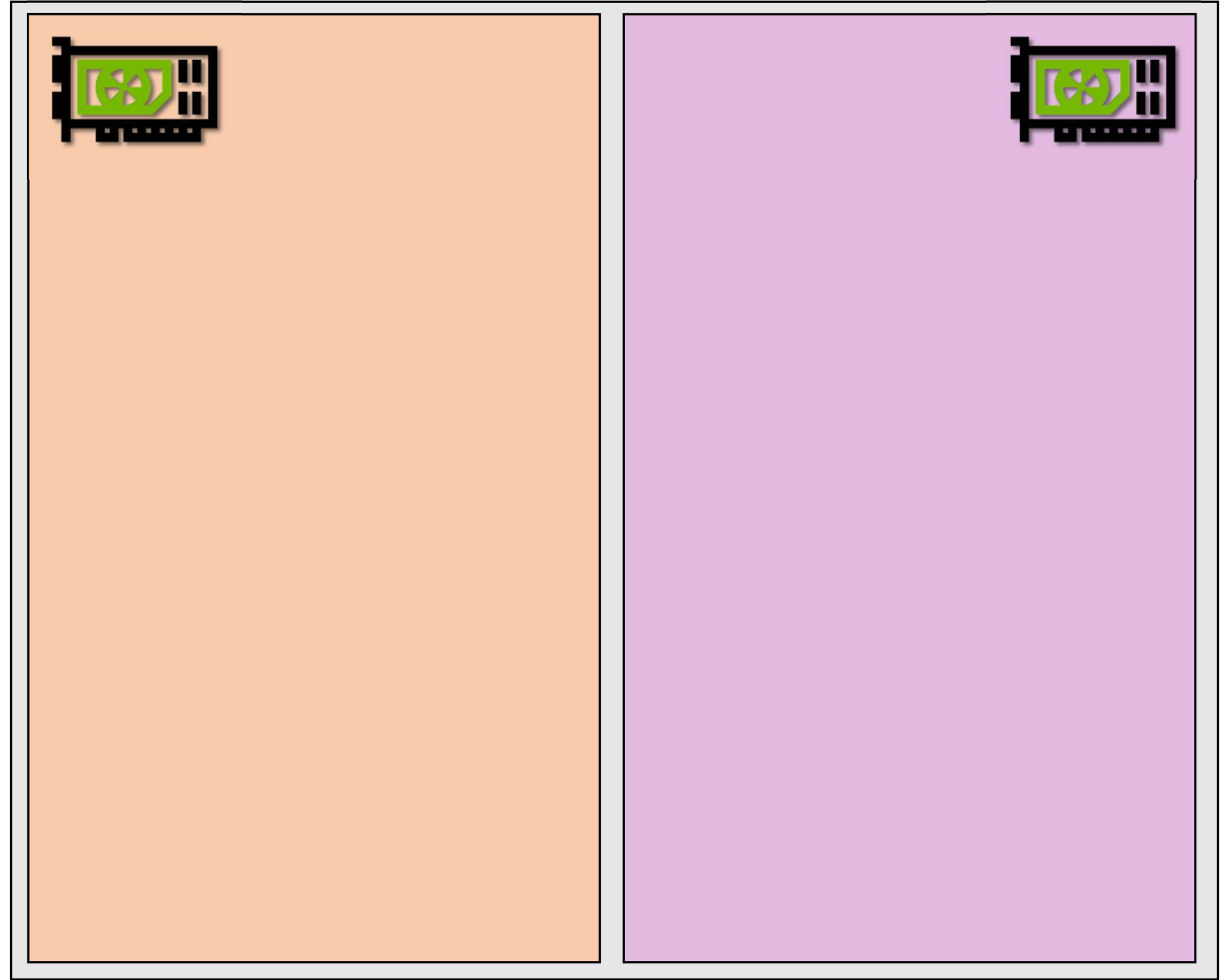
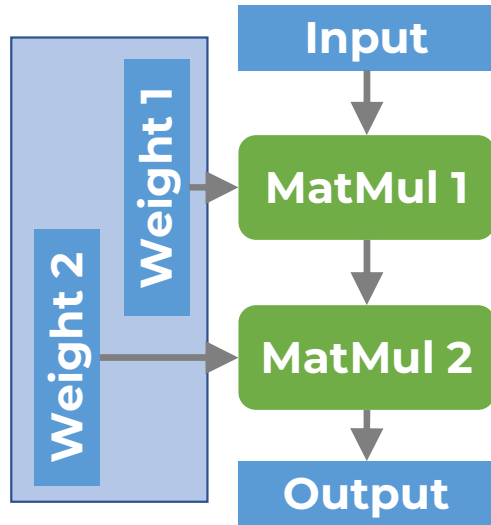
Data Parallel



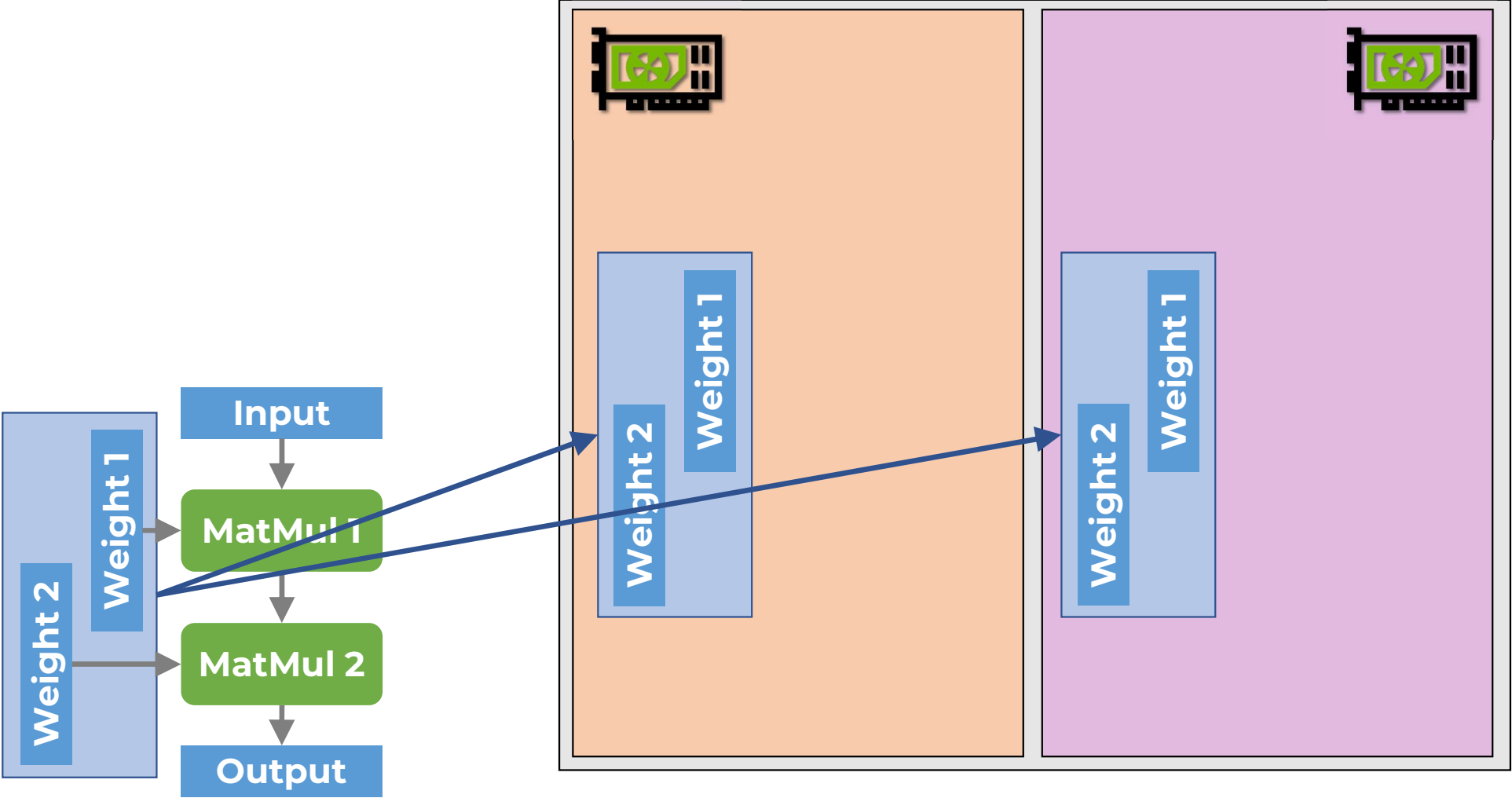
Data Parallel



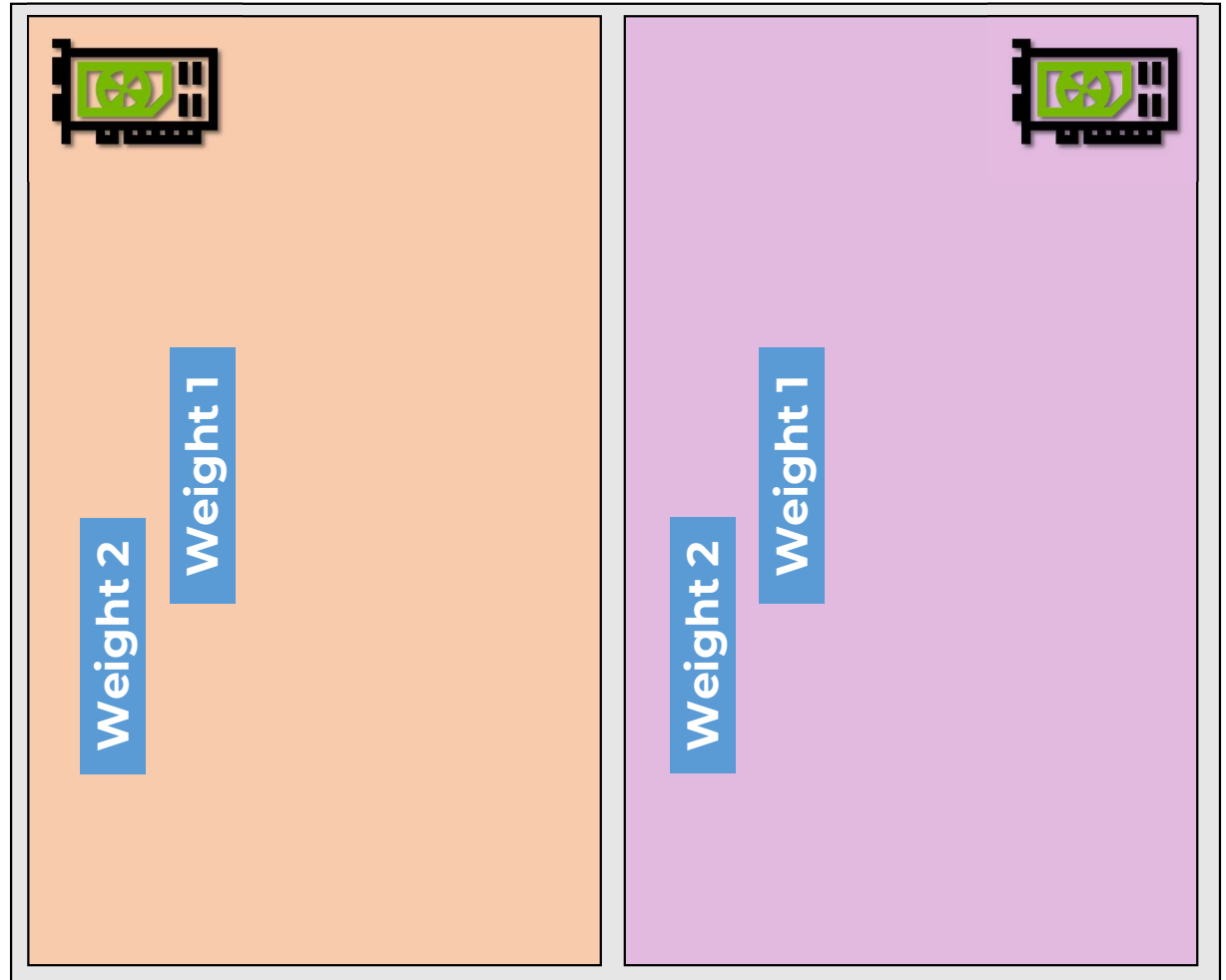
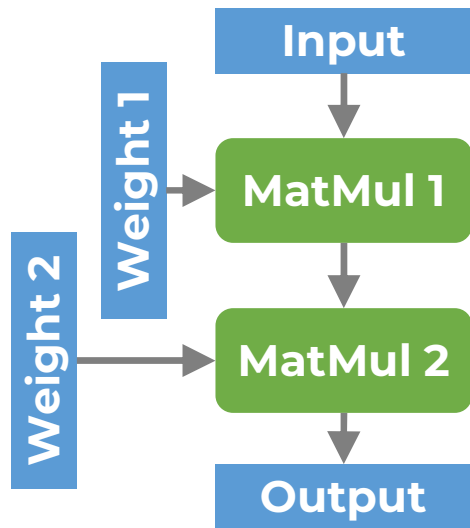
Data Parallel



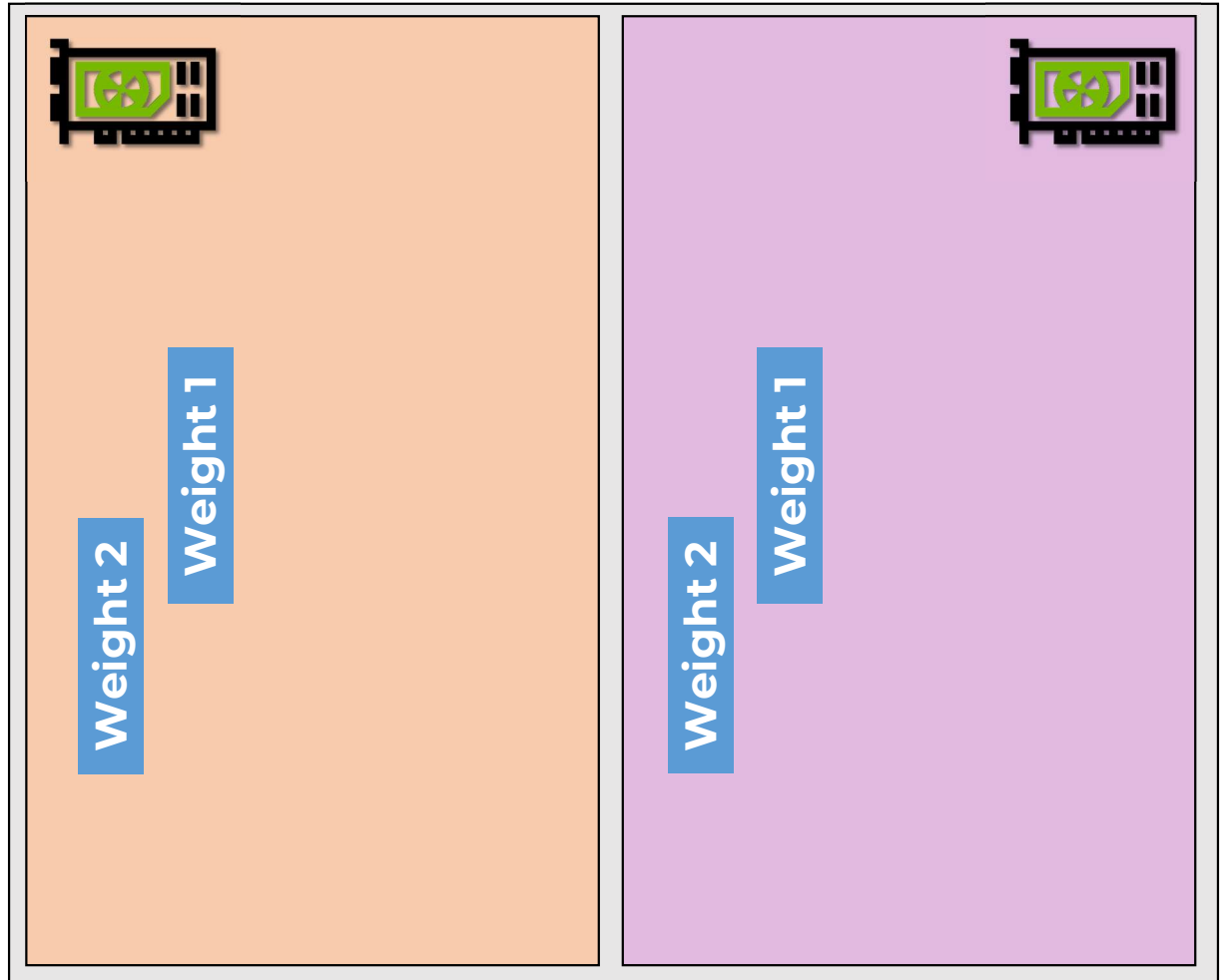
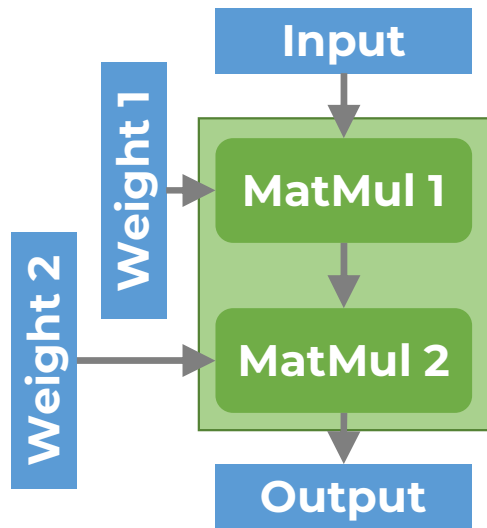
Data Parallel



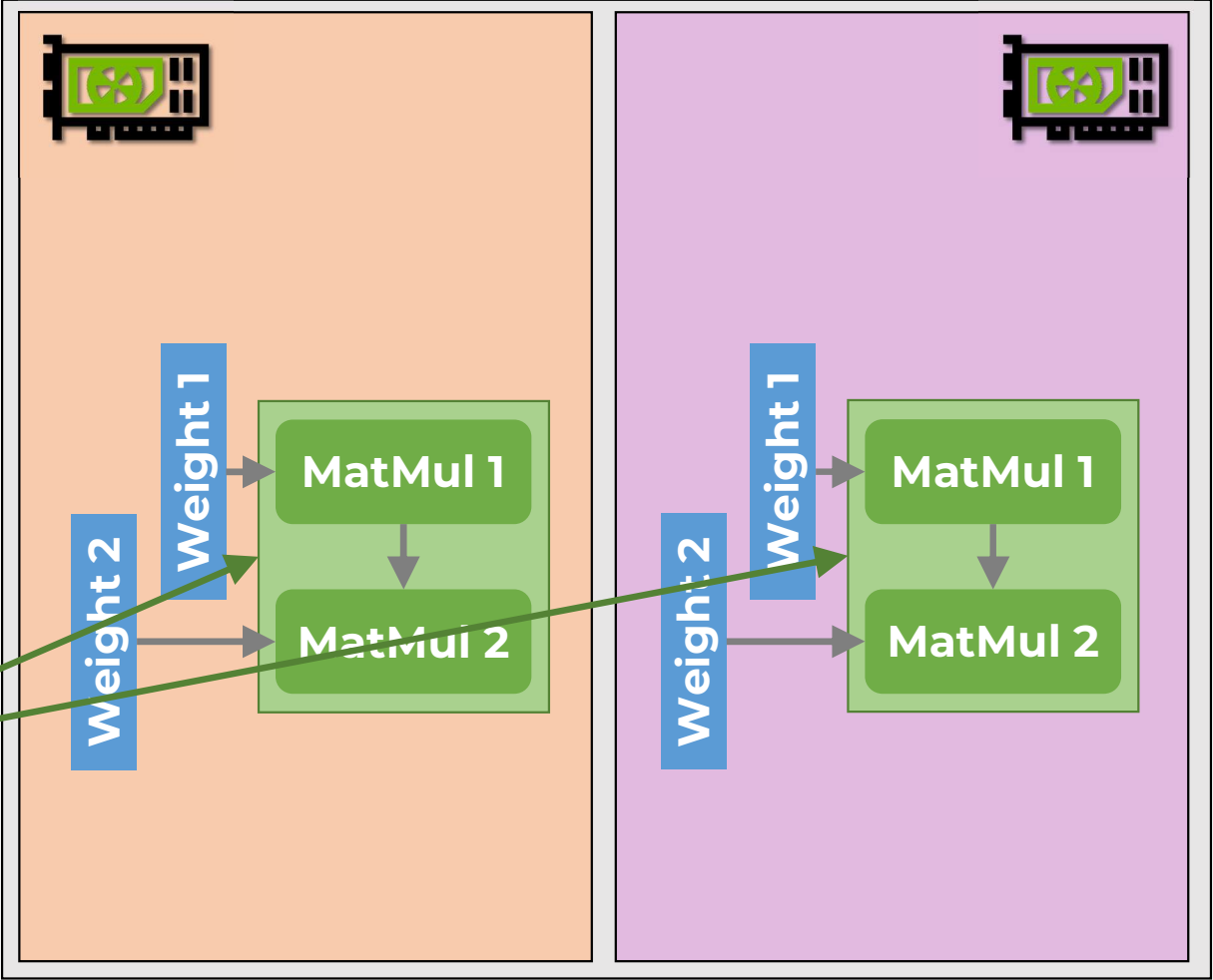
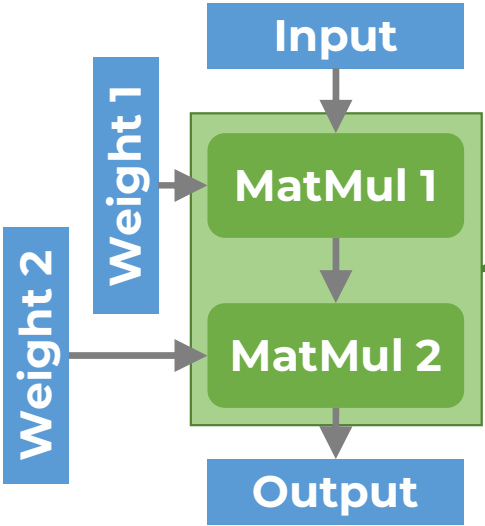
Data Parallel



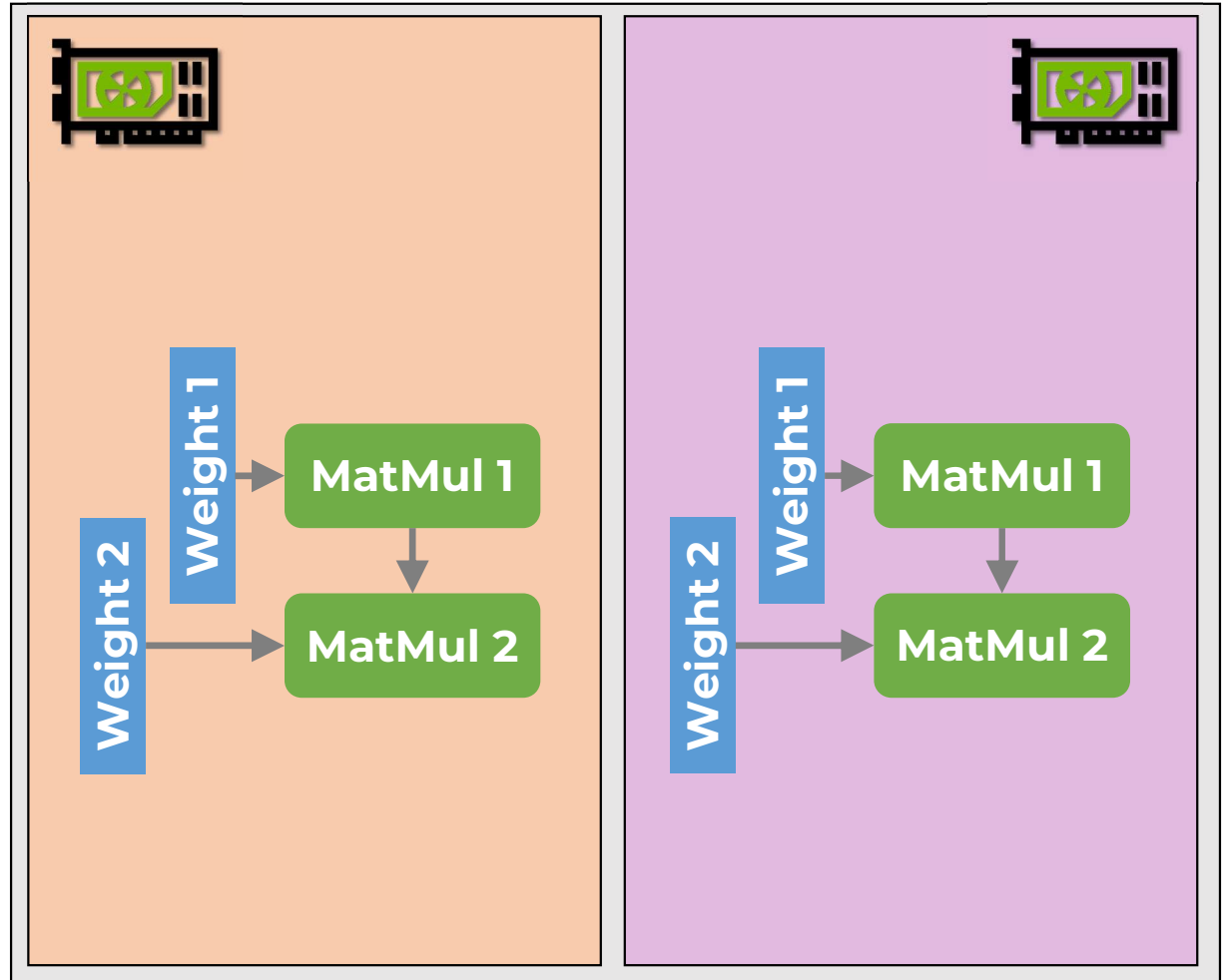
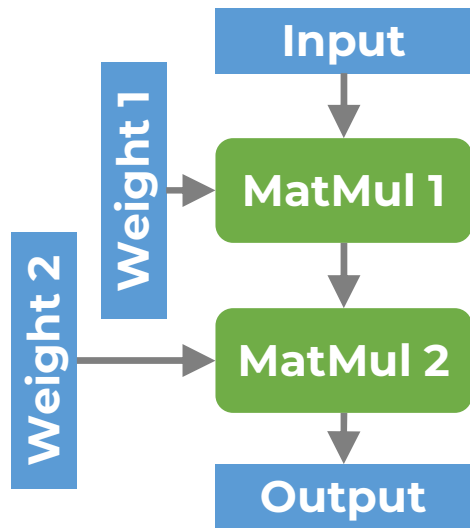
Data Parallel



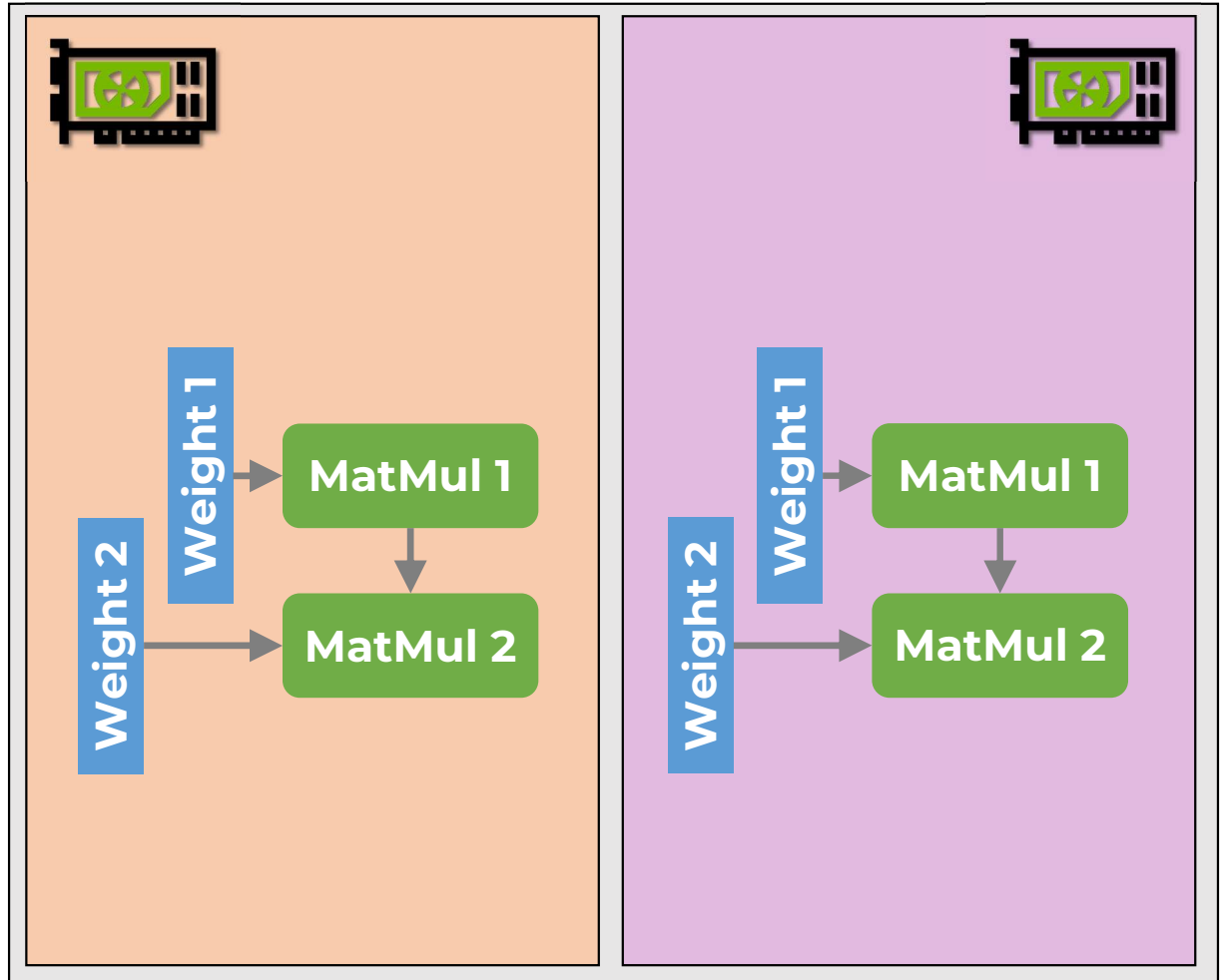
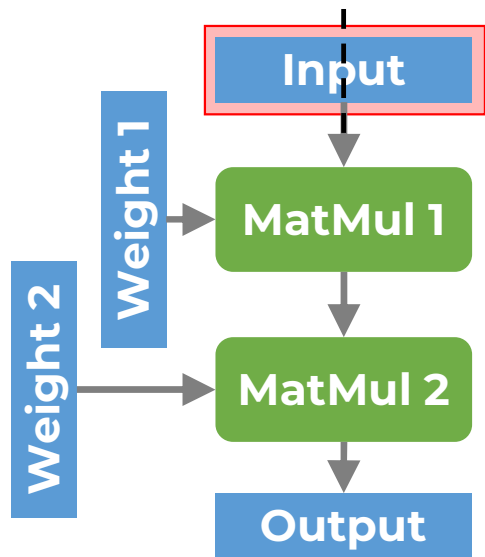
Data Parallel



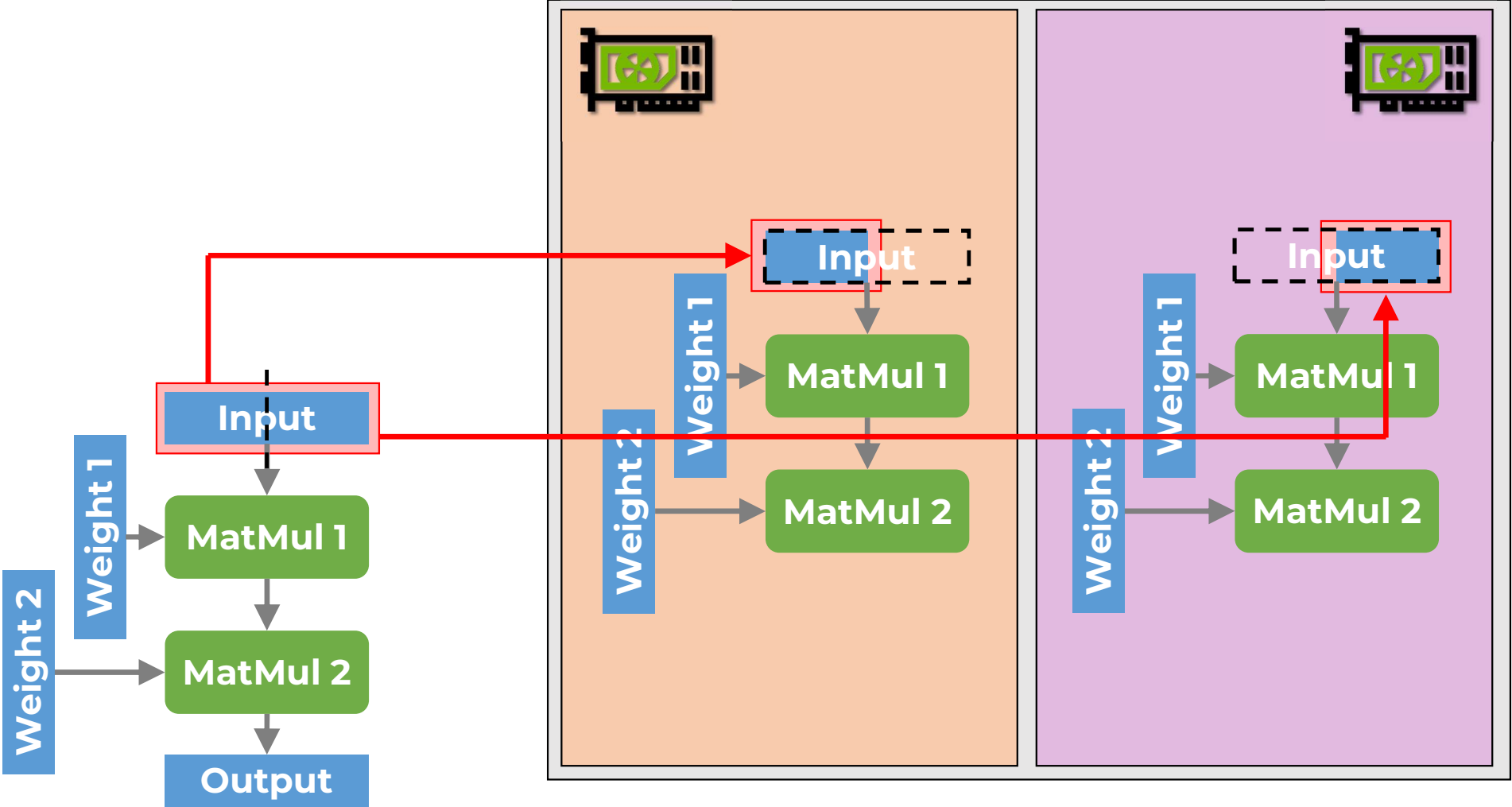
Data Parallel



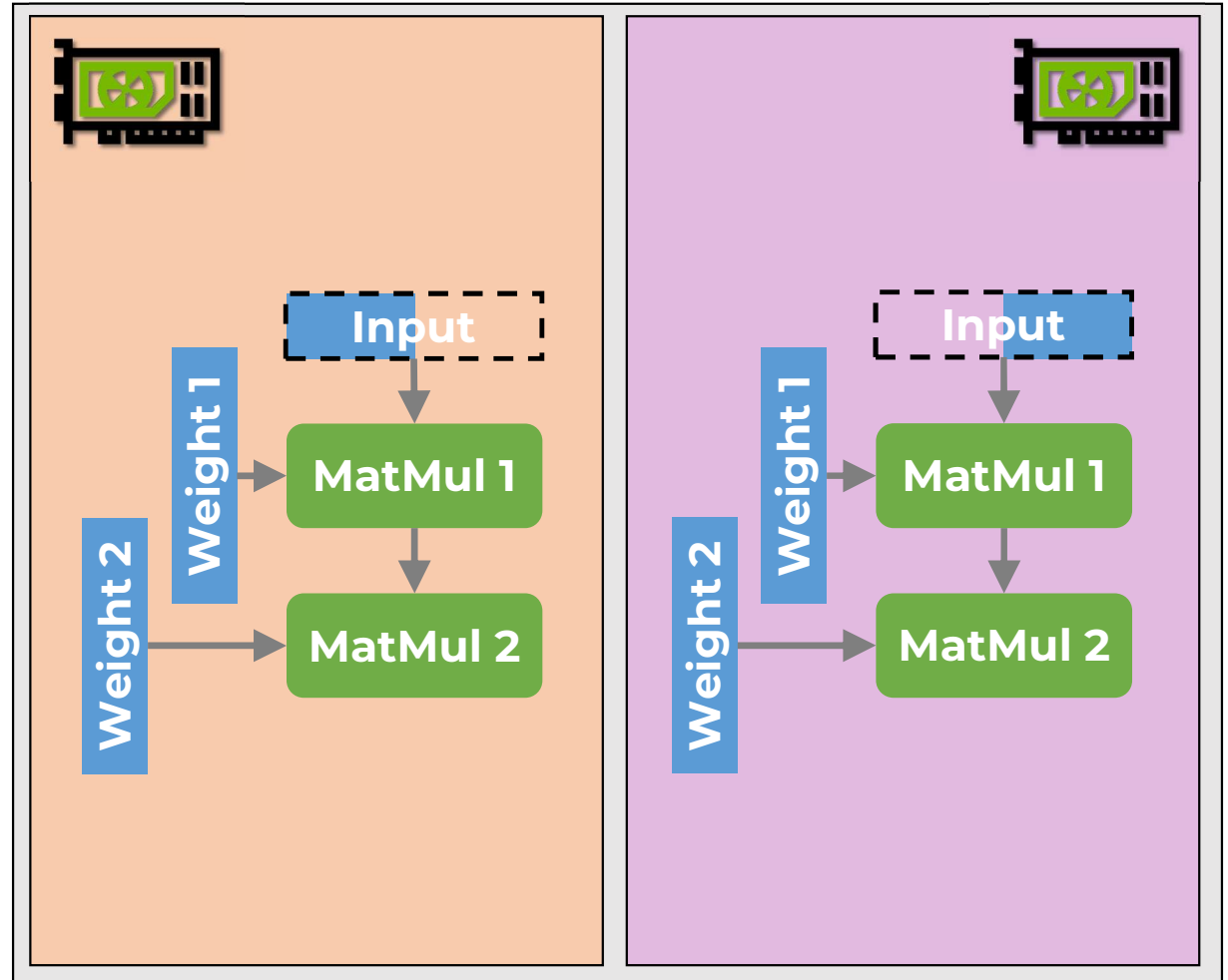
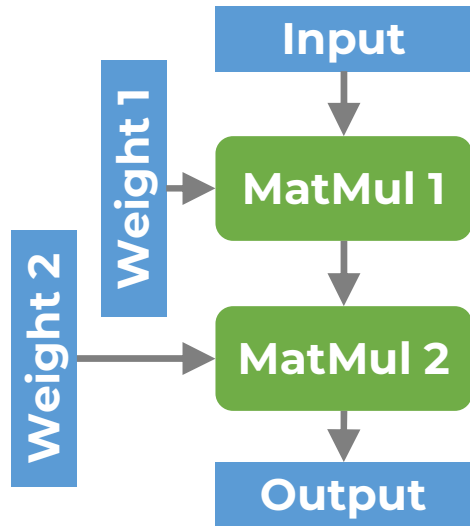
Data Parallel



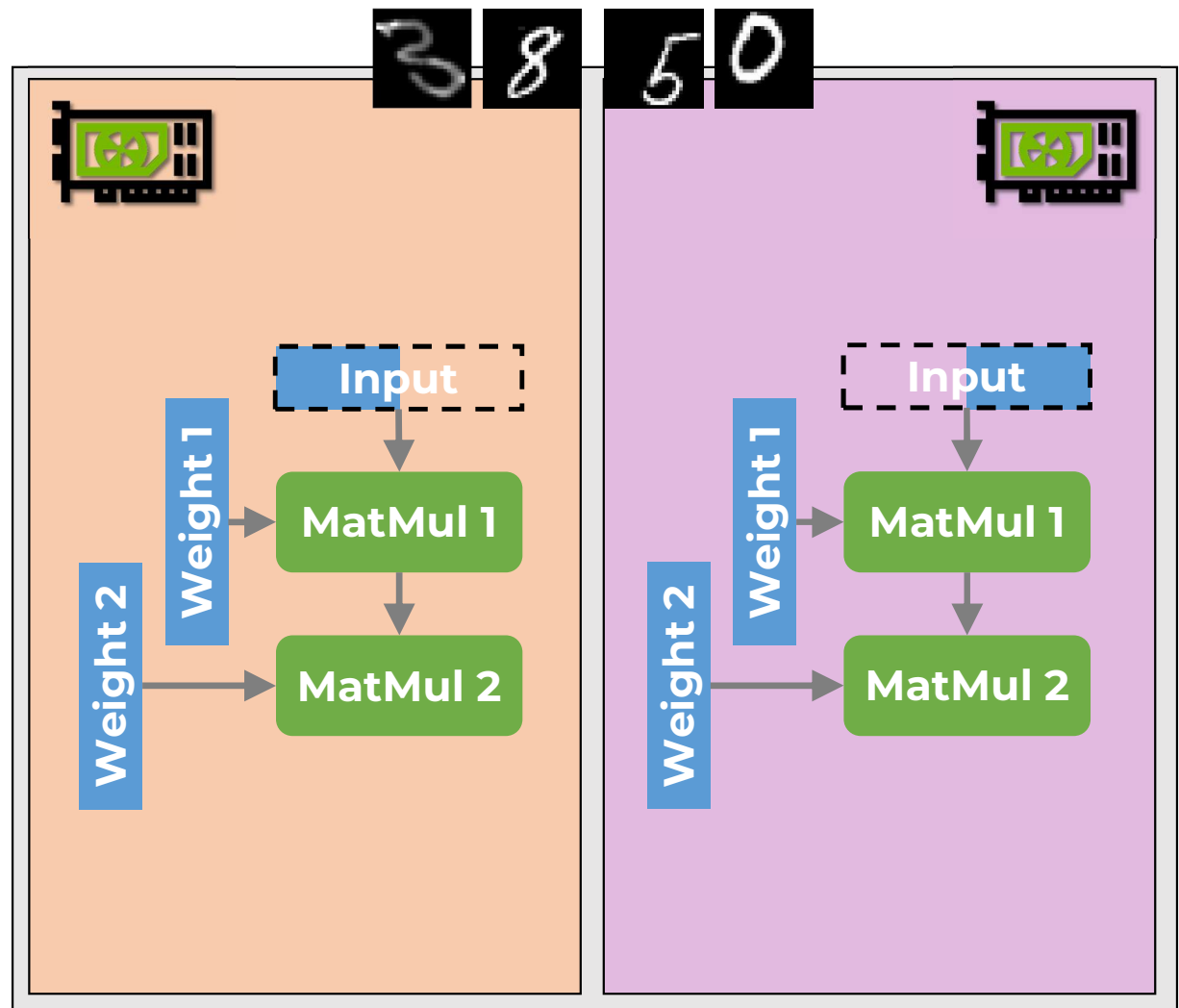
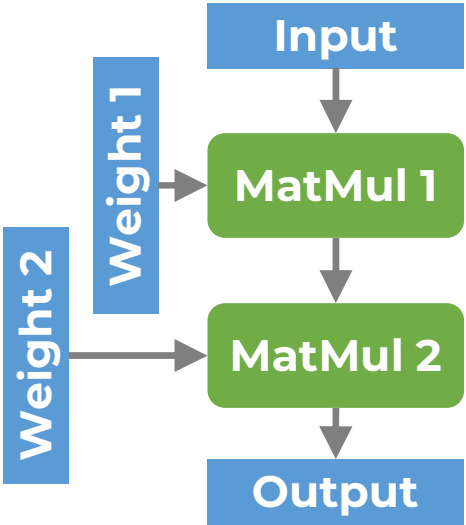
Data Parallel



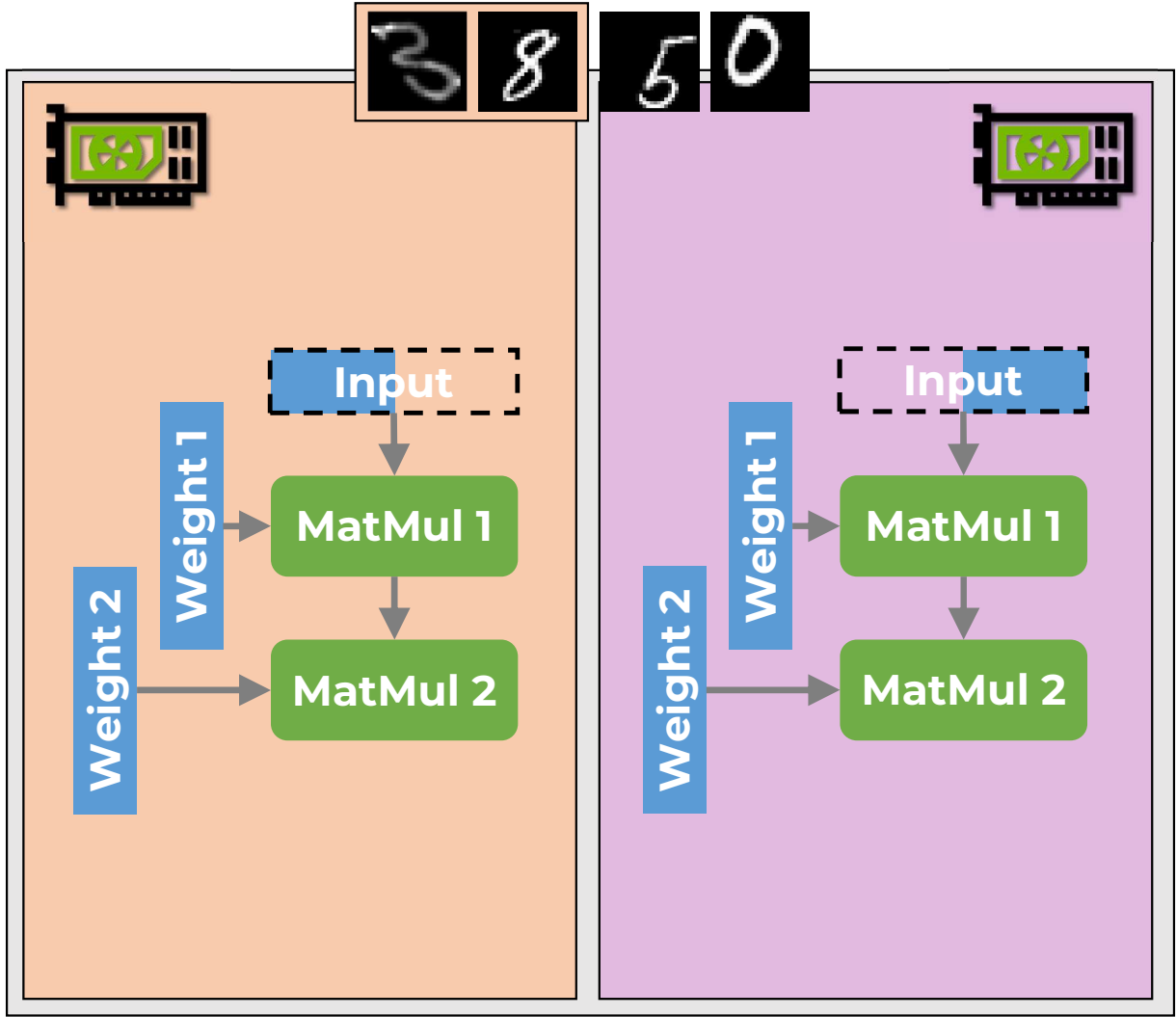
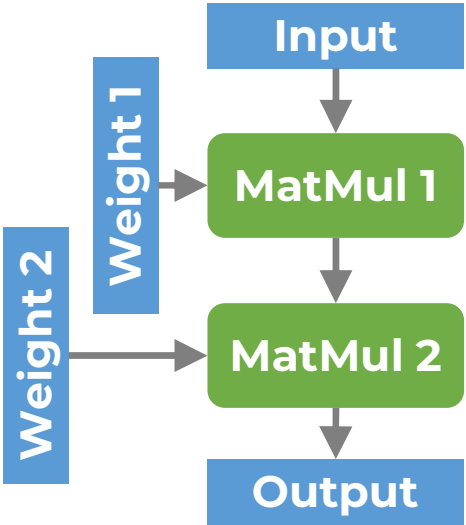
Data Parallel



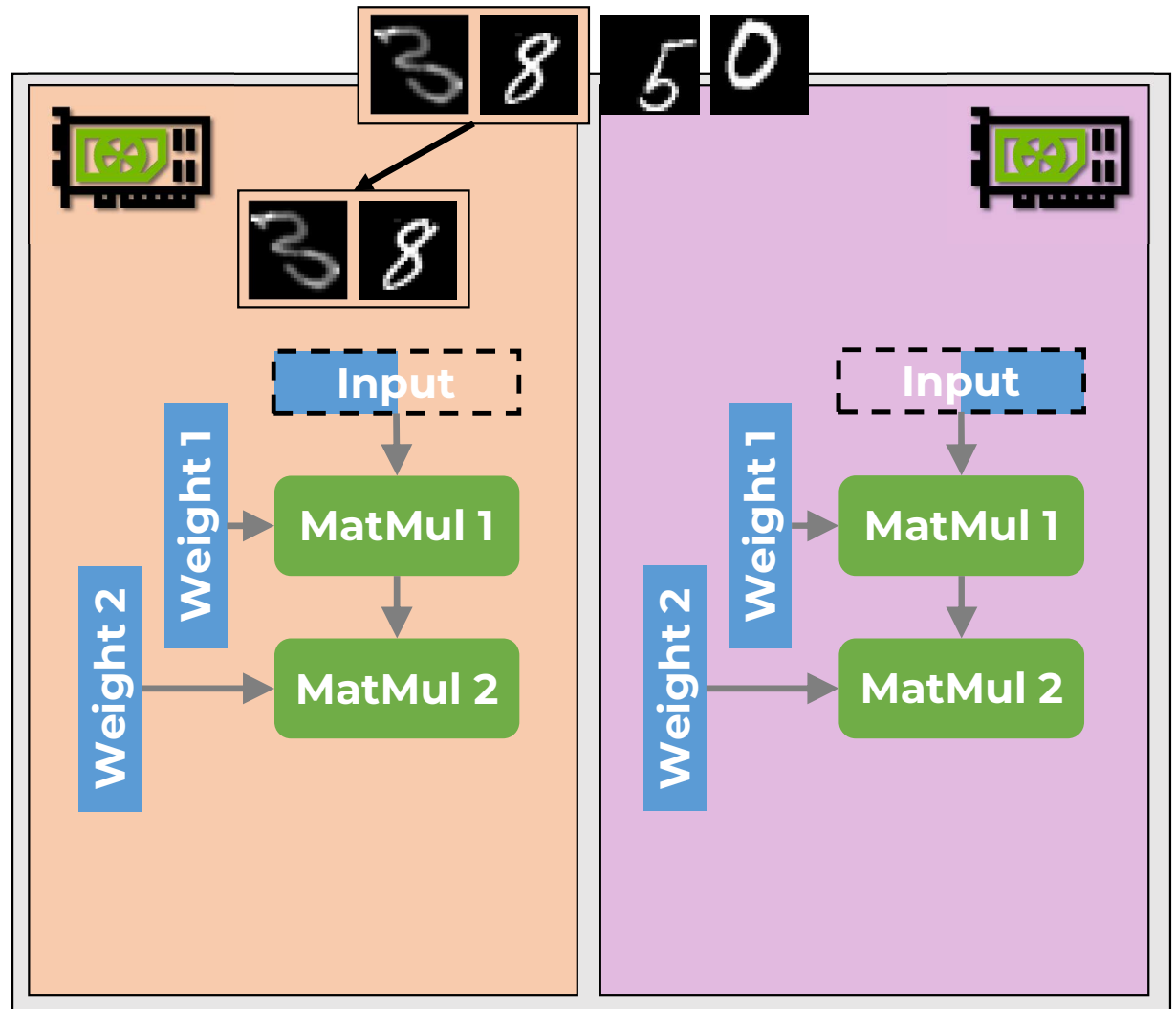
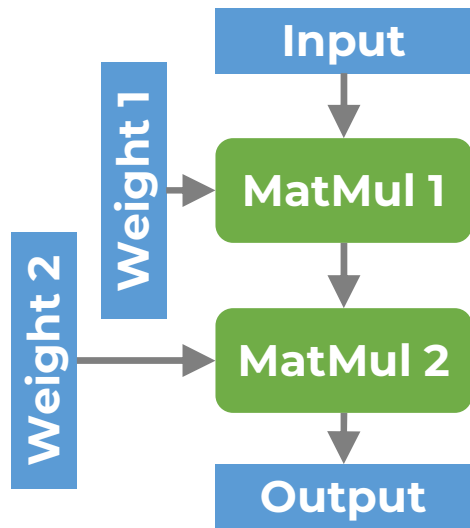
Data Parallel



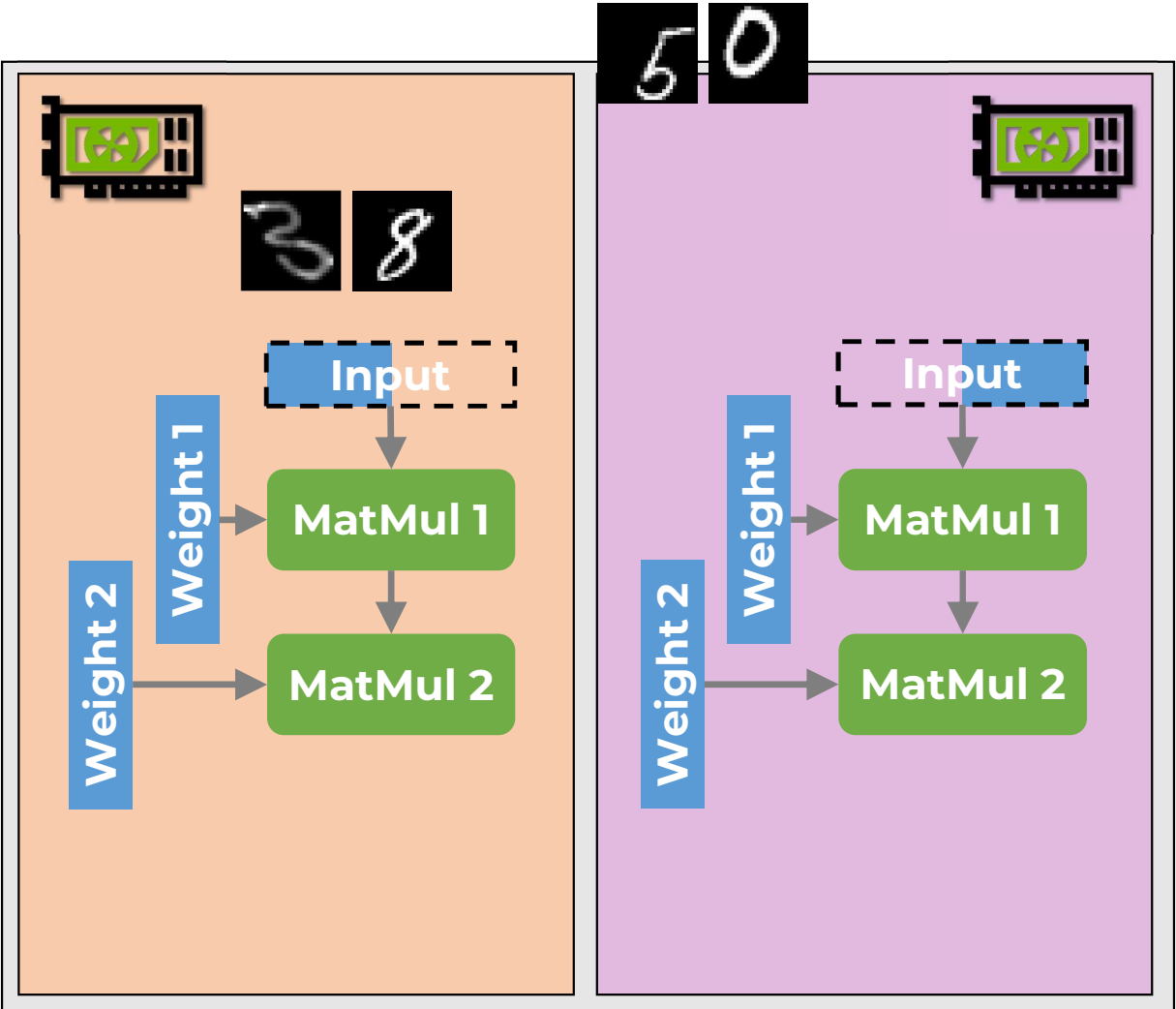
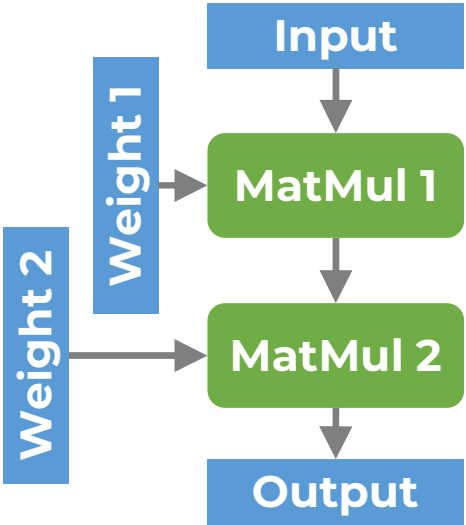
Data Parallel



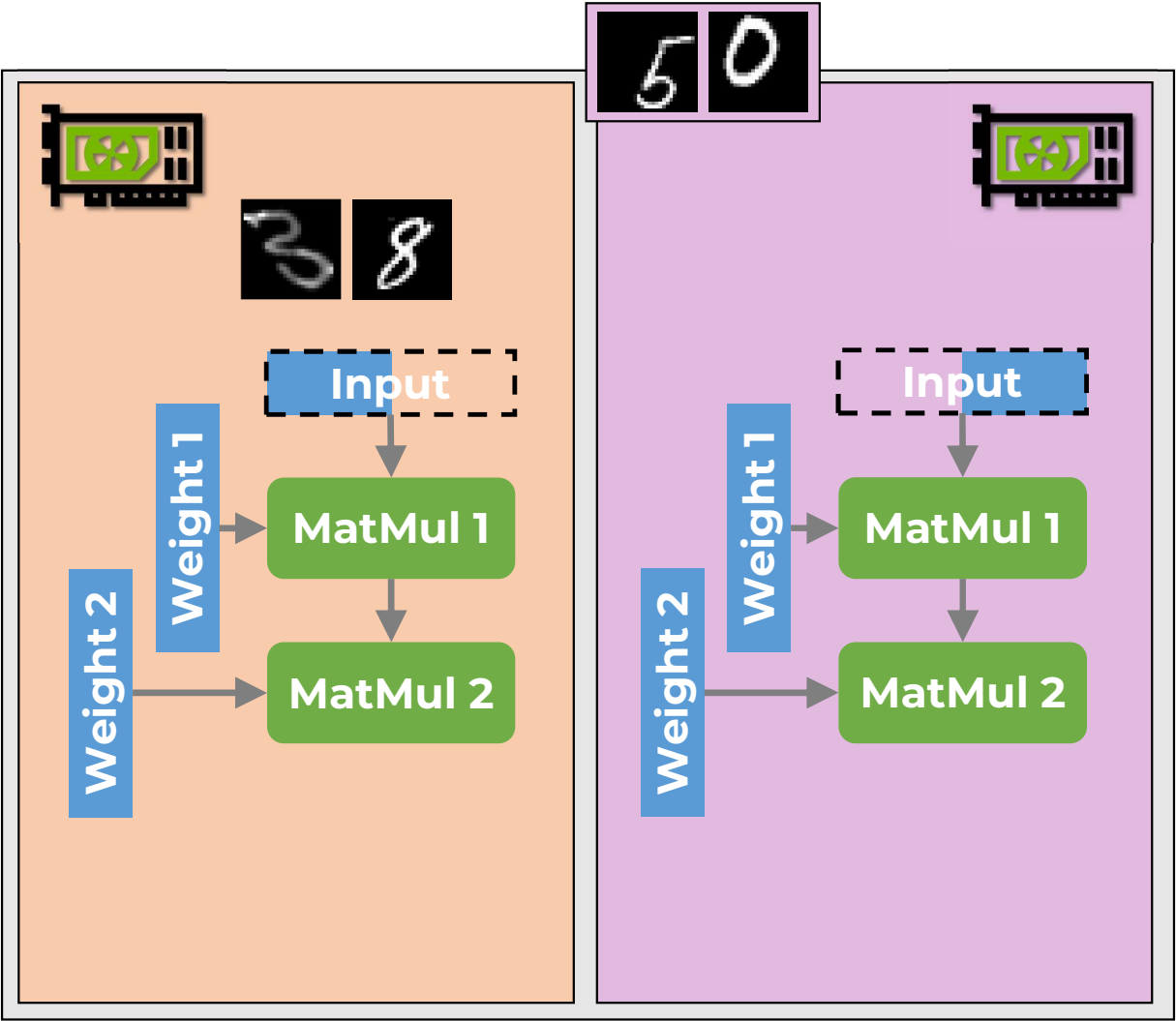
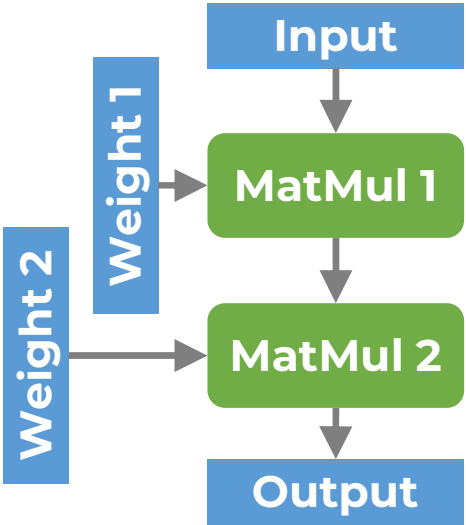
Data Parallel



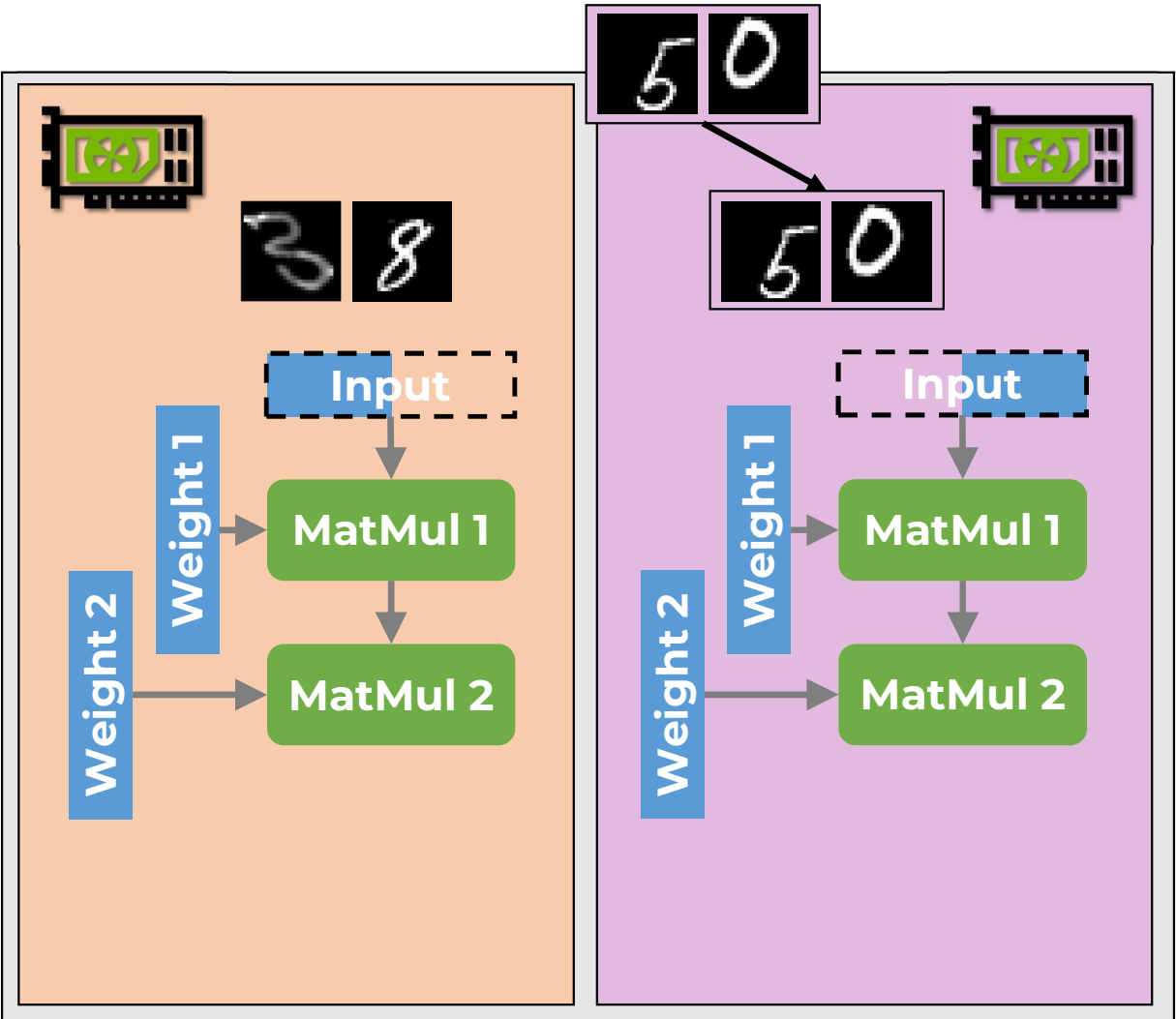
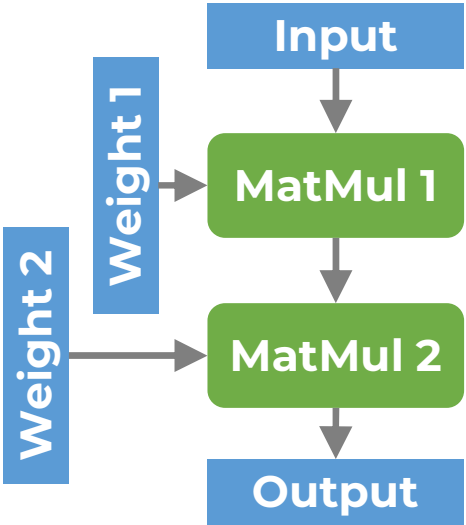
Data Parallel



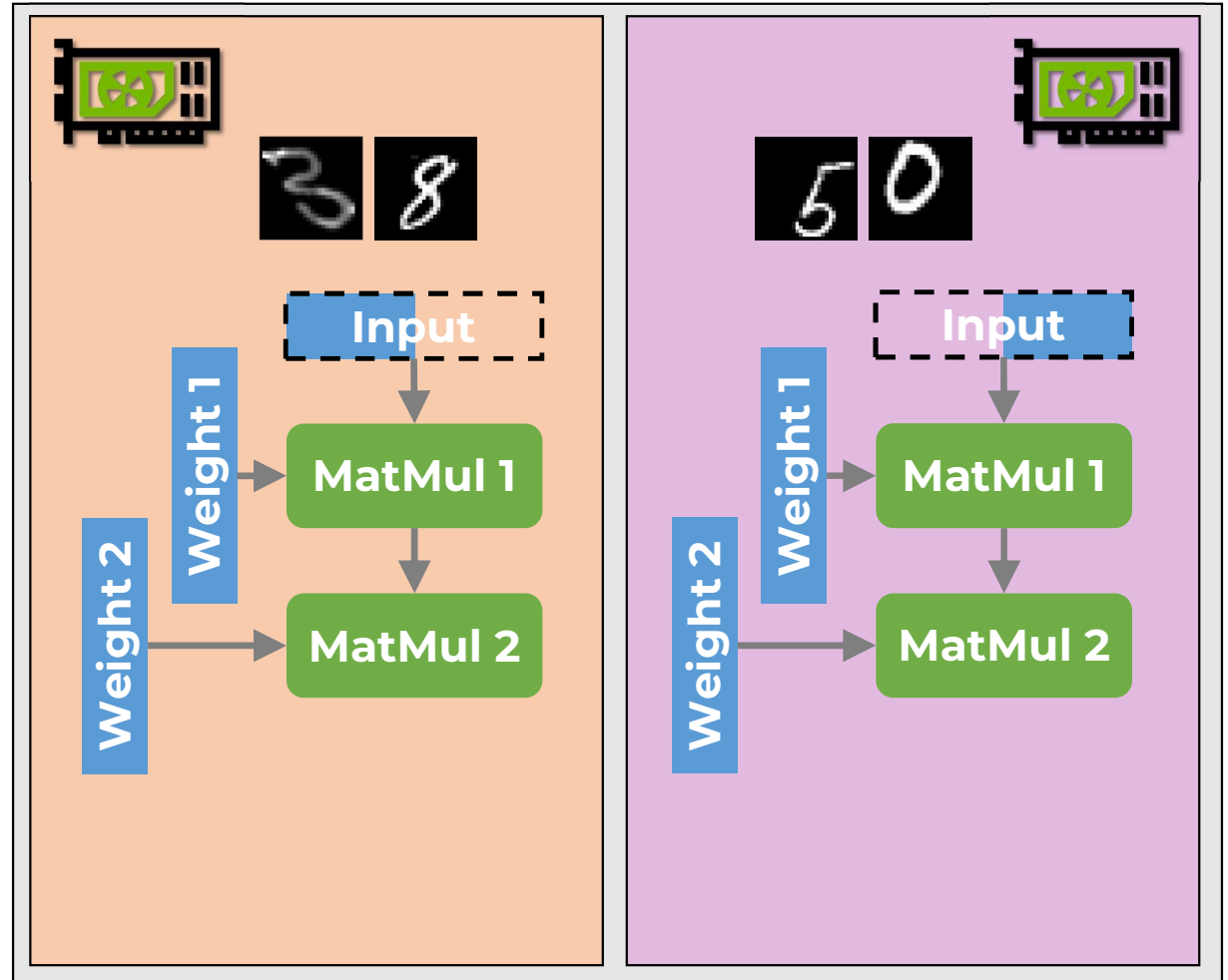
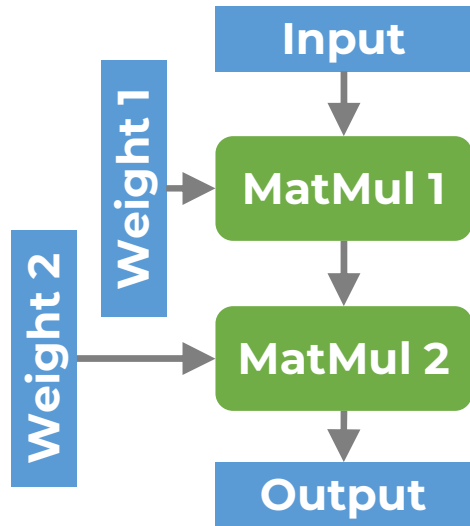
Data Parallel



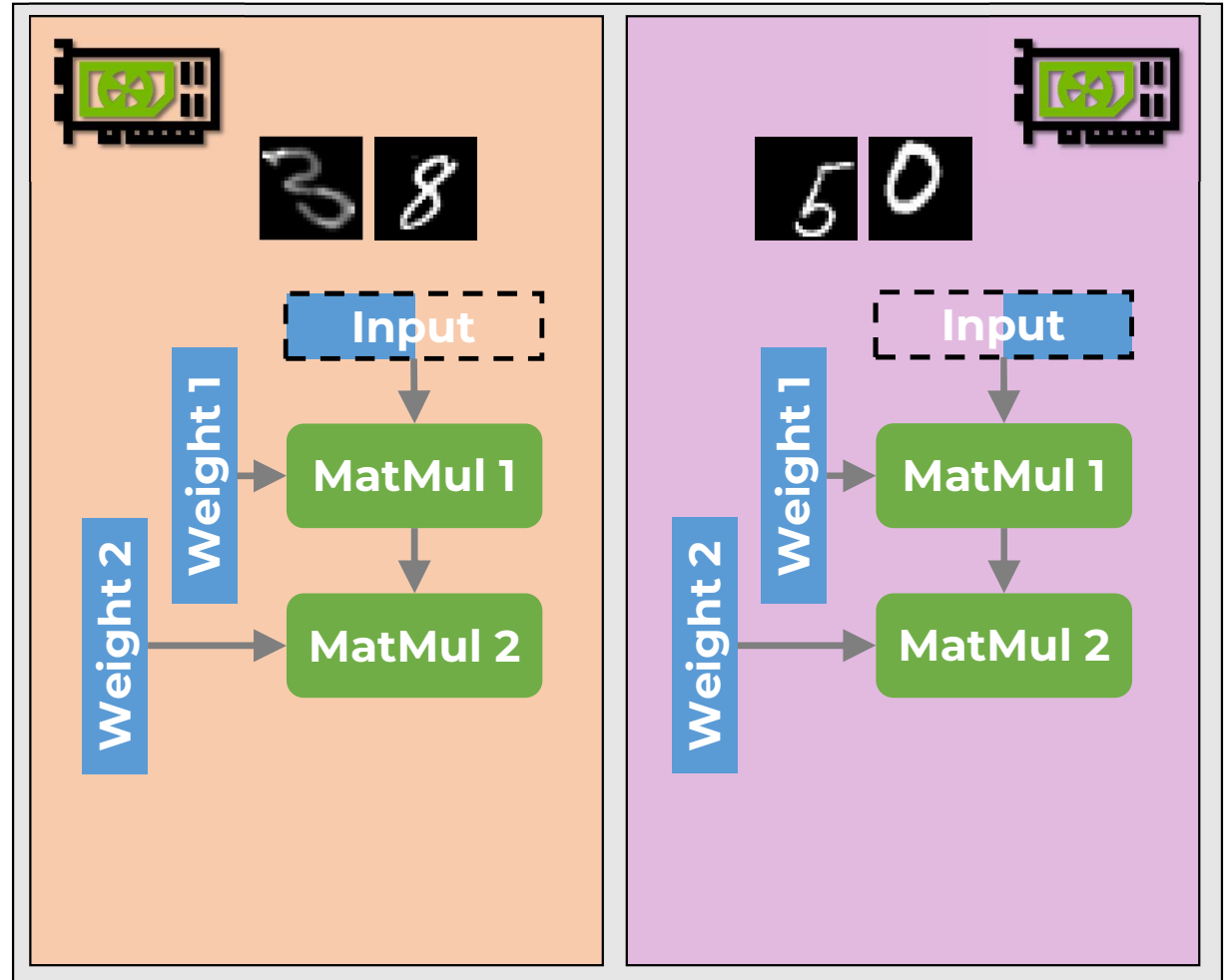
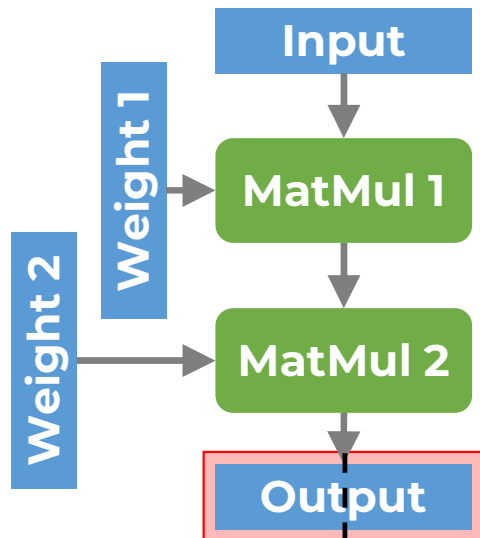
Data Parallel



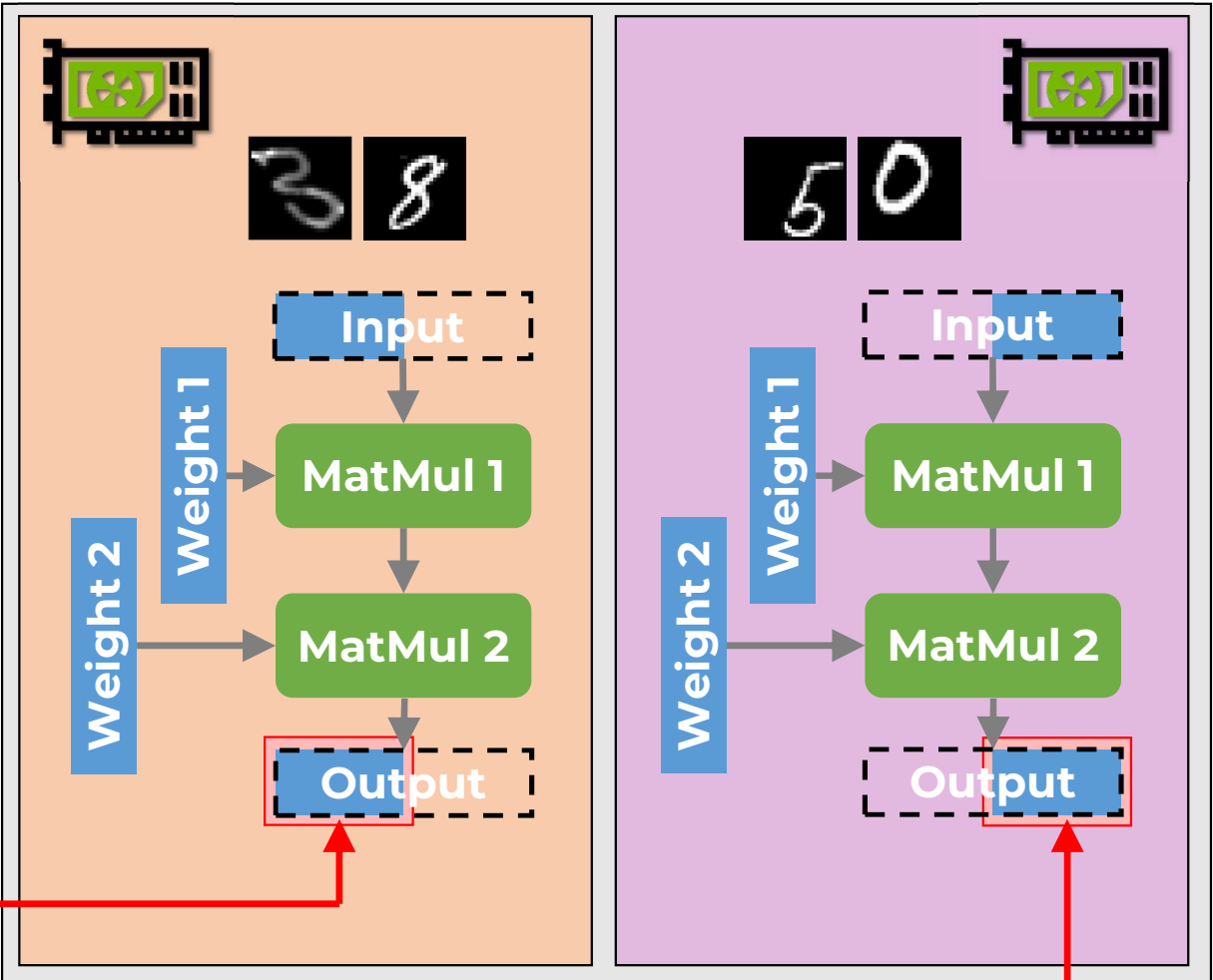
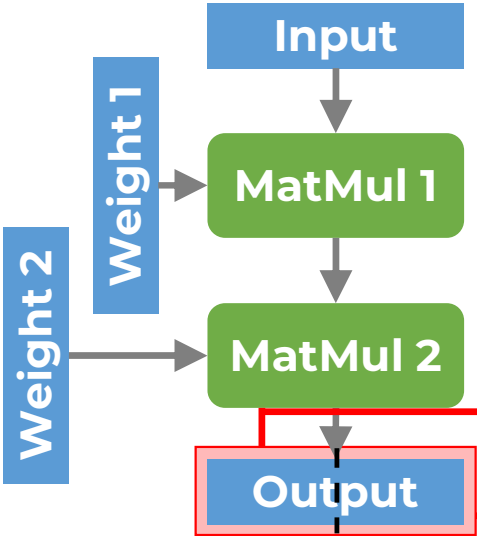
Data Parallel



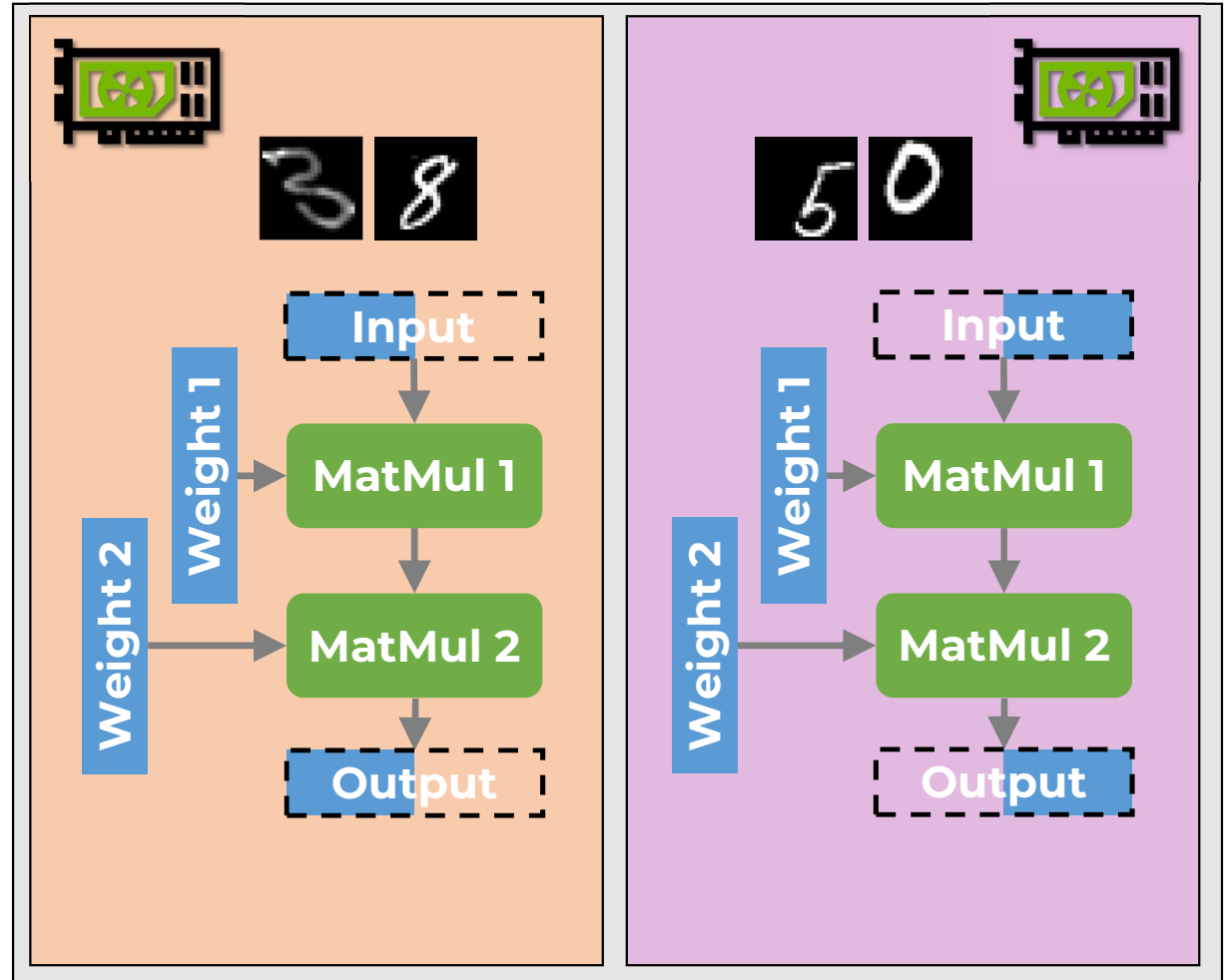
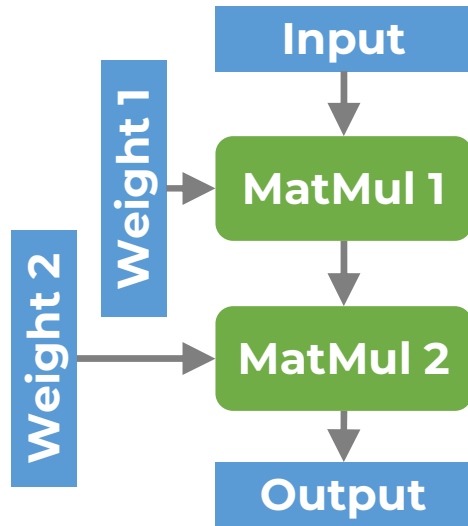
Data Parallel



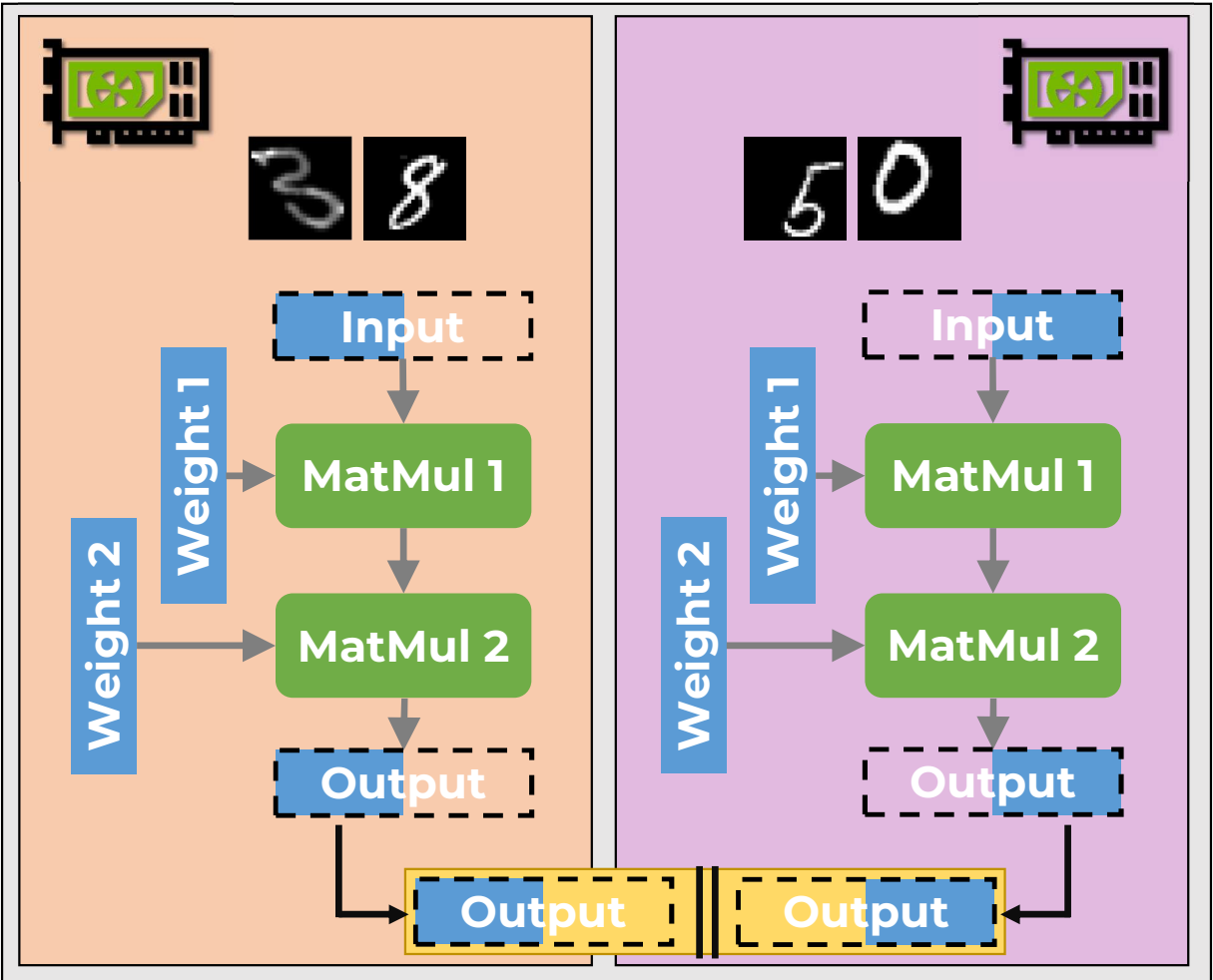
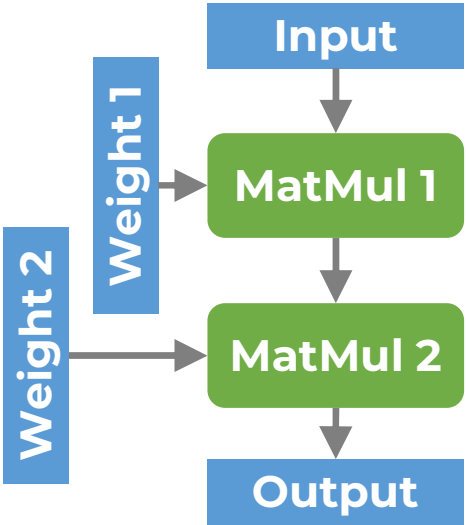
Data Parallel



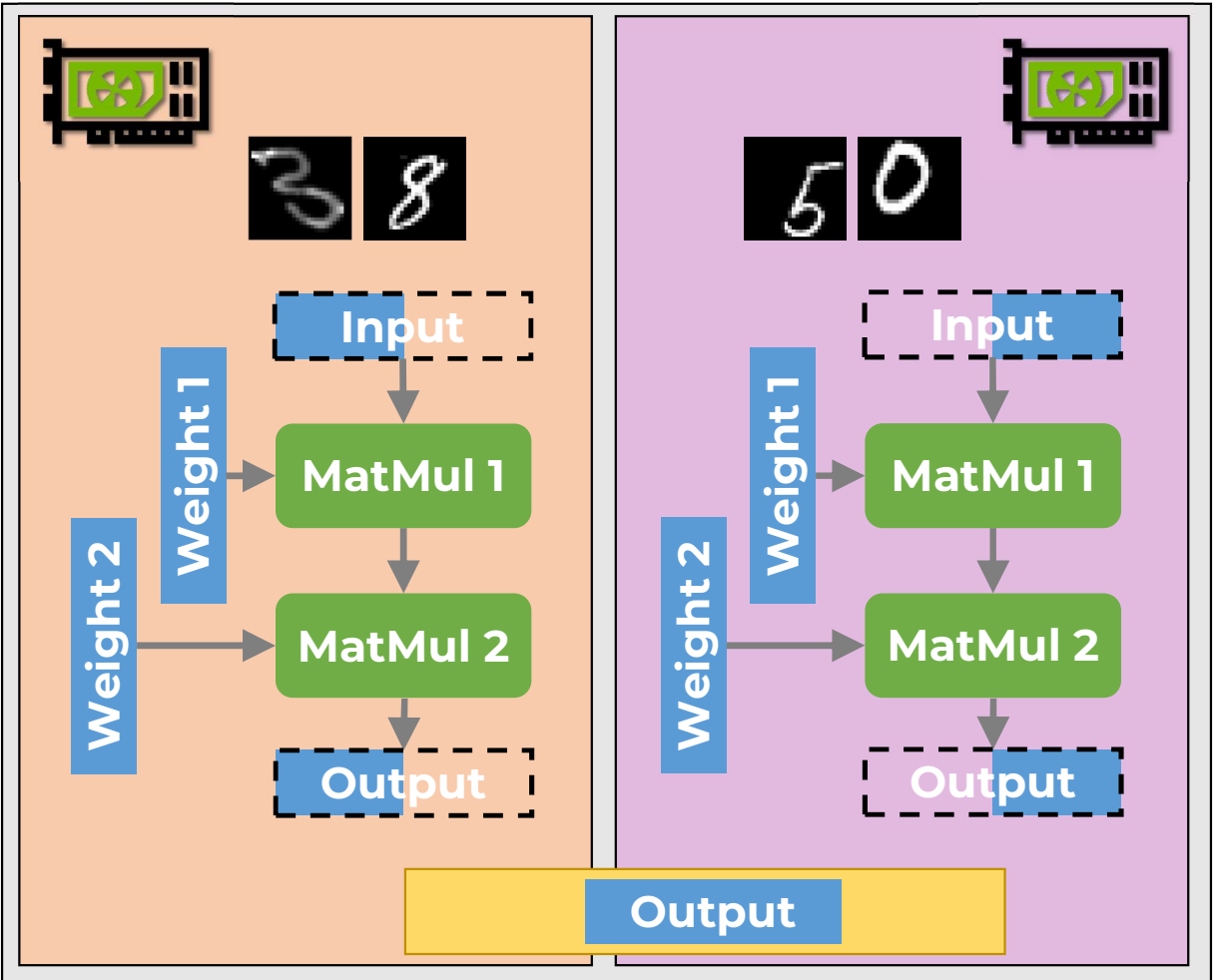
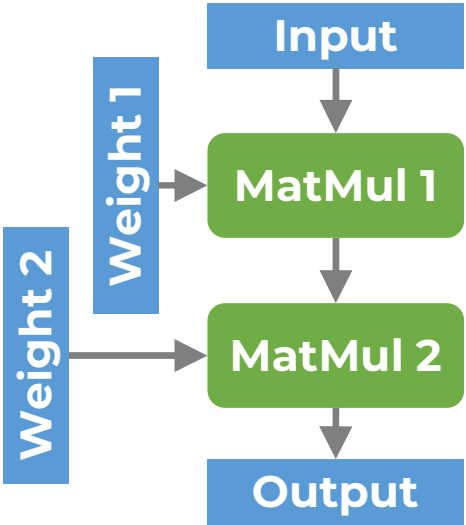
Data Parallel



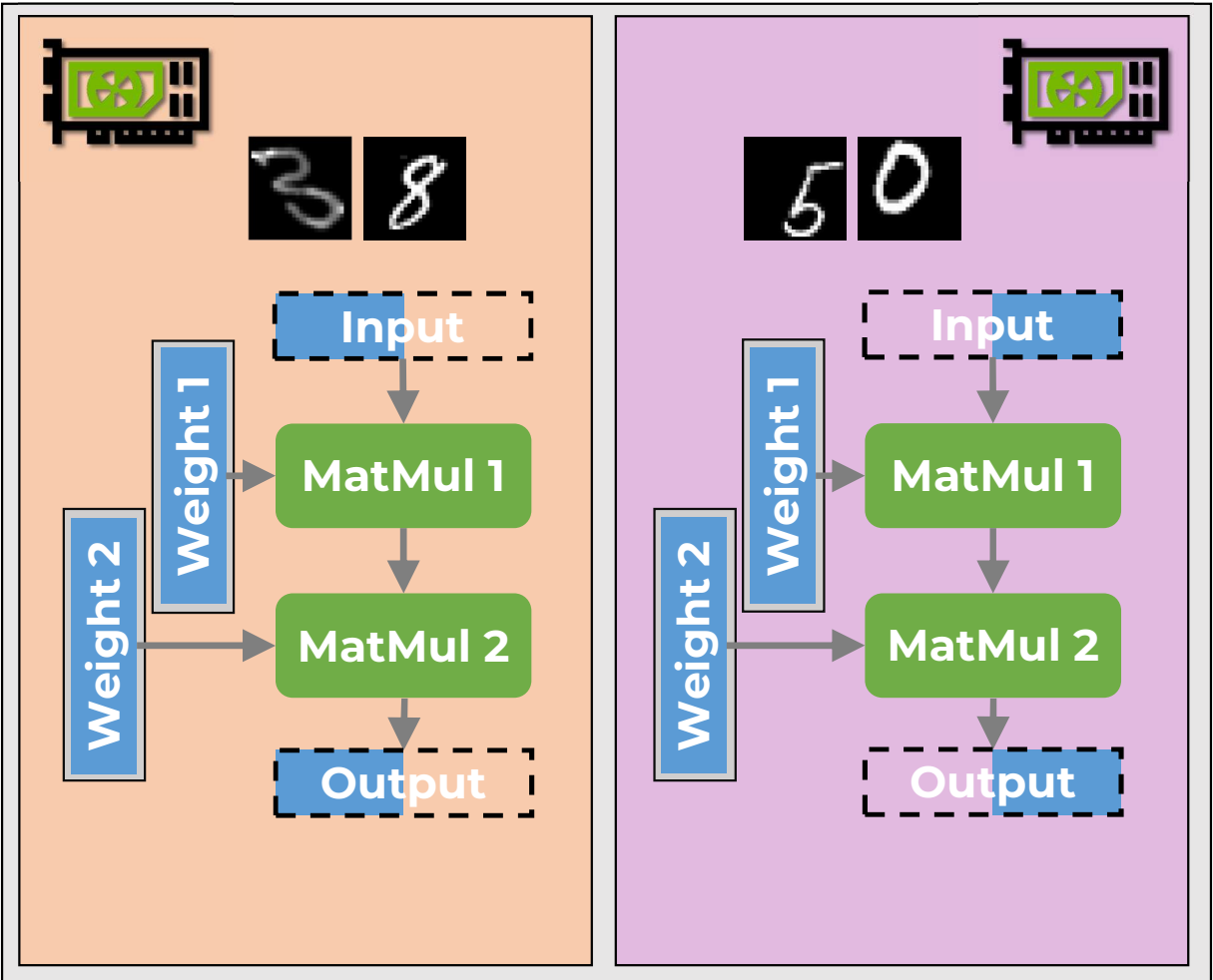
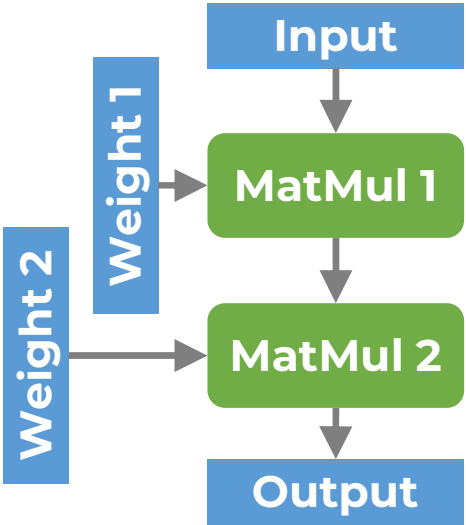
Data Parallel



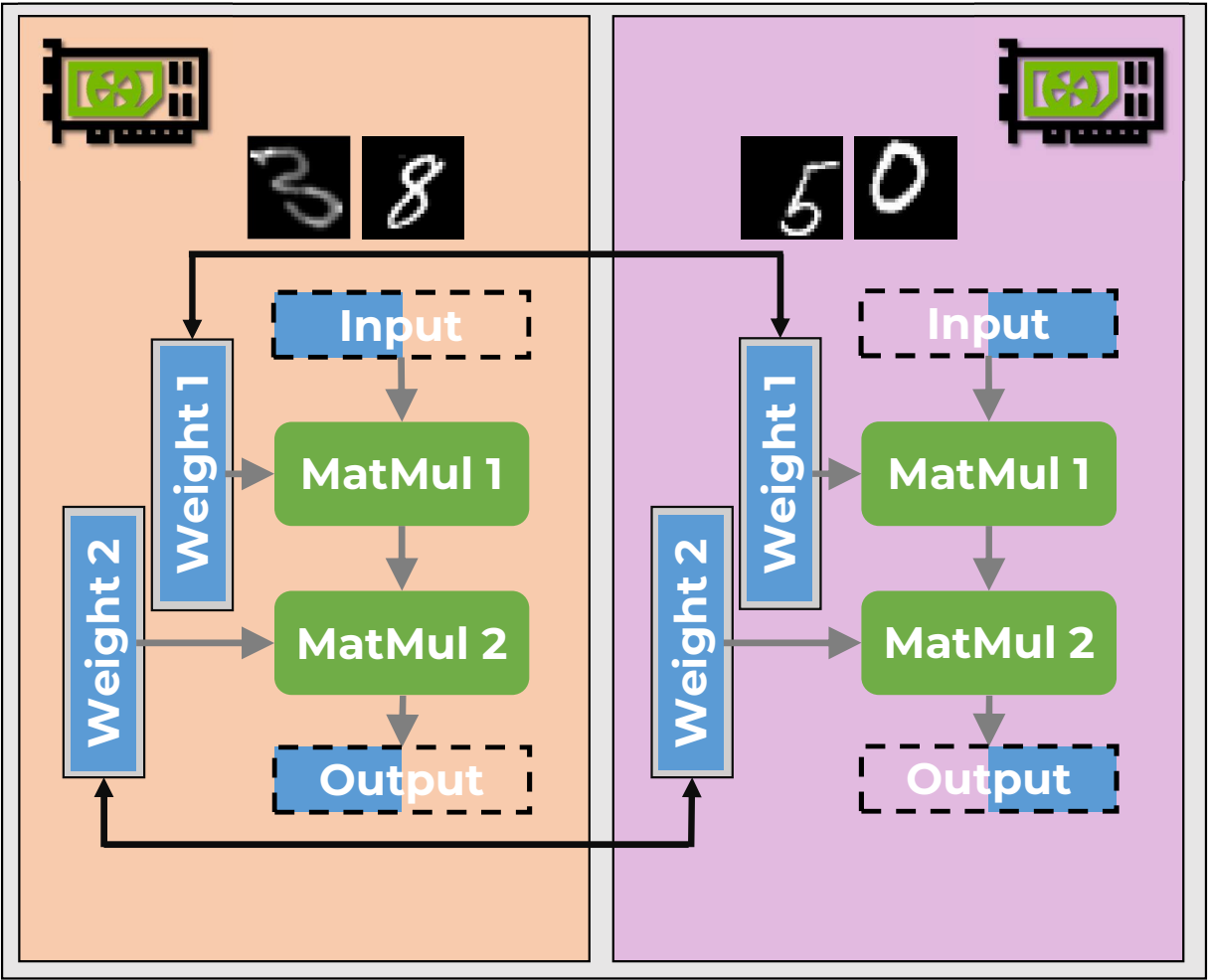
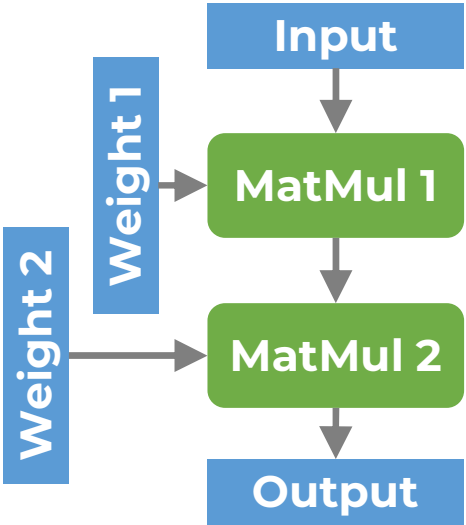
Data Parallel



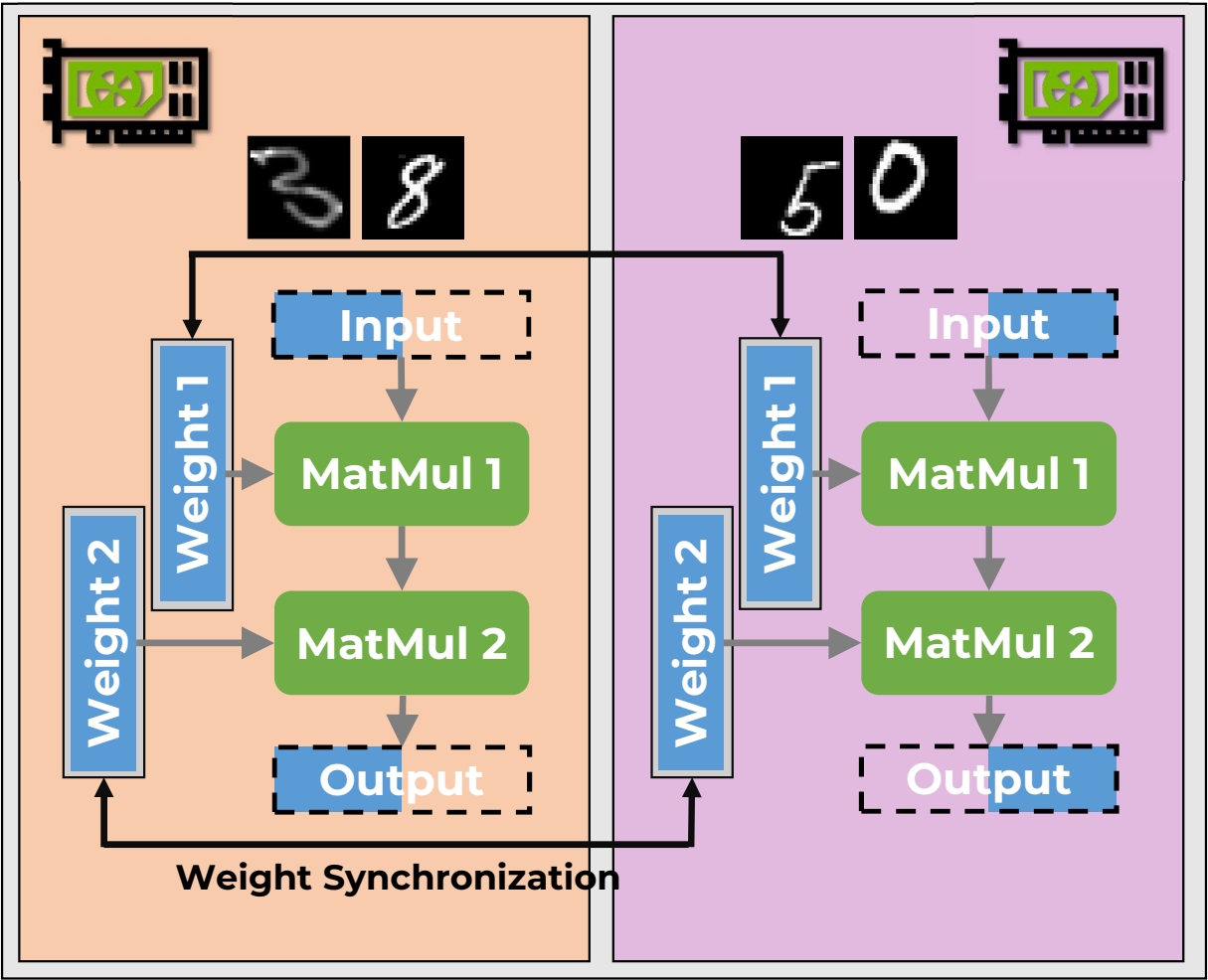
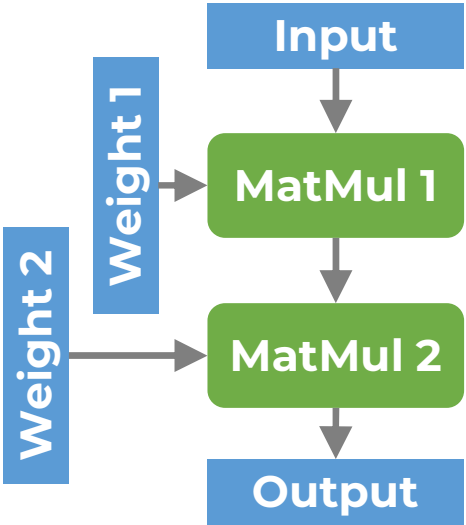
Data Parallel



Data Parallel



Data Parallel



Data Parallel

Data Parallel

Model Parallel

Data Parallel
Model Parallel
Attribute Parallel

Data Parallel
Model Parallel
Attribute Parallel
Reduction Parallel

Data Parallel
Model Parallel
Attribute Parallel
Reduction Parallel
Parameter Parallel

Data Parallel
Model Parallel
Attribute Parallel
Reduction Parallel
Parameter Parallel
Pipeline Parallel

Data Parallel
Model Parallel
Attribute Parallel
Reduction Parallel
Parameter Parallel
Pipeline Parallel

...

Data Parallel
Model Parallel
Attribute Parallel
Reduction Parallel
Parameter Parallel
Pipeline Parallel
...

Parallelization

Data Parallel
Model Parallel
Attribute Parallel
Reduction Parallel
Parameter Parallel
Pipeline Parallel
...

2^{|Parallelization|}

Data Parallel
Model Parallel
Attribute Parallel
Reduction Parallel
Parameter Parallel
Pipeline Parallel
...

2 | Parallelization |

Algebraic Transformations

Operator Fusion
Operator Splitting
Operator Reordering
...

Data Parallel
Model Parallel
Attribute Parallel
Reduction Parallel
Parameter Parallel
Pipeline Parallel
...

2 | Parallelization |

2 | Algebraic Transformations |

Operator Fusion
Operator Splitting
Operator Reordering
...

Auto-Parallelization

FlexFlow [MLSys 19]

Tofu [EuroSys 19]

PipeDream [SOSP 19]

automap [arXiv 19]

Whale [arXiv 21]

Alpa [OSDI 22]

...

Auto-Parallelization

FlexFlow [MLSys 19]

Tofu [EuroSys 19]

PipeDream [SOSP 19]

automap [arXiv 19]

Whale [arXiv 21]

Alpa [OSDI 22]

...

Algebraic Optimizers

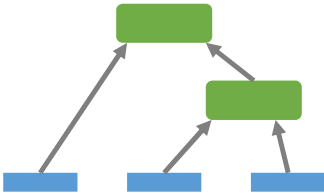
MetaFlow [MLSys 19]

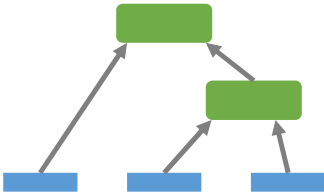
TASO [SOSP 19]

PET [OSDI 21]

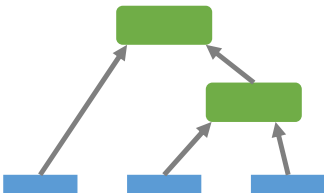
Tensat [MLSys 21]

...



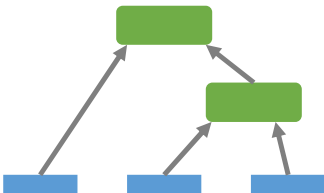


**Auto-
Parallelization**



**Auto-
Parallelization**

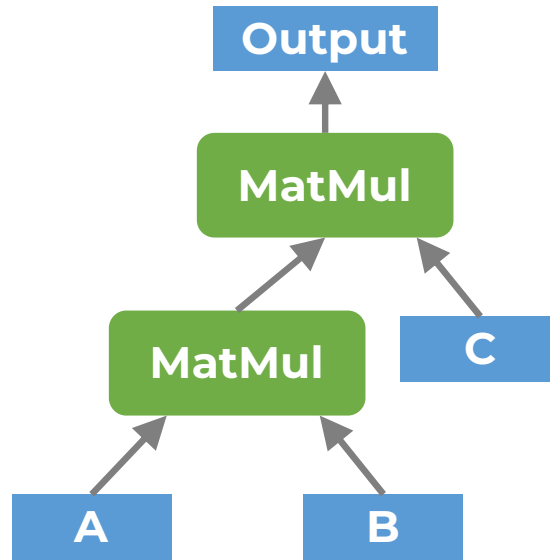
**Algebraic
Optimizer**

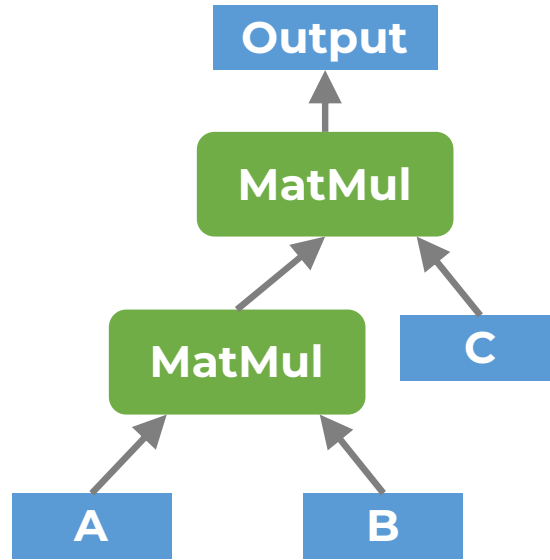


**Auto-
Parallelization**

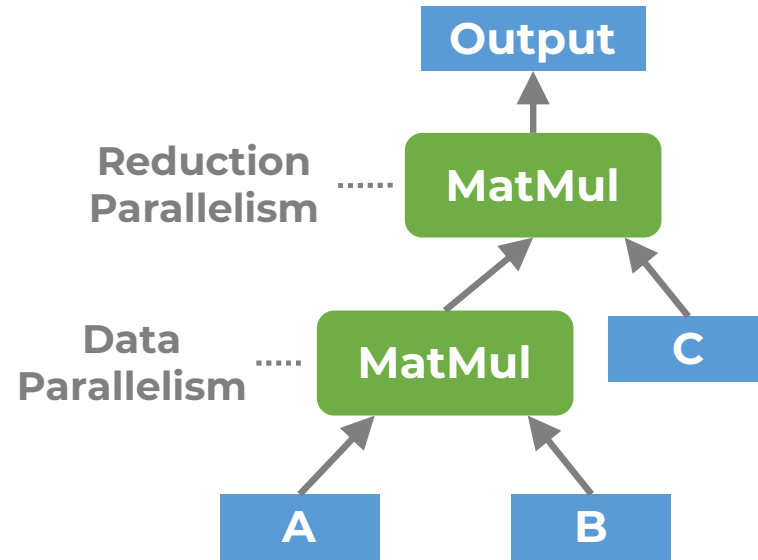
**Algebraic
Optimizer**

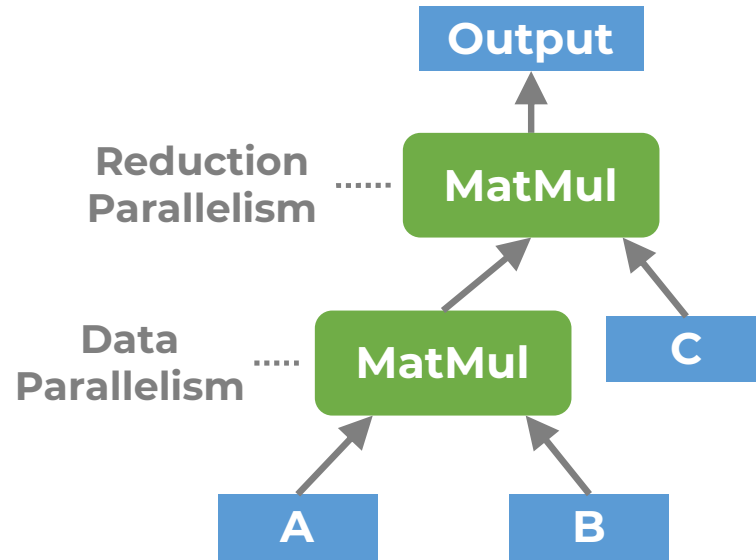
?



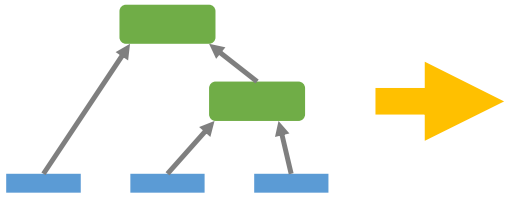


“computation graph”

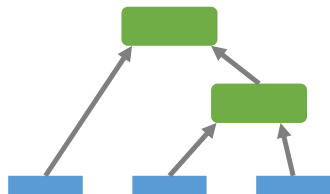




“annotated computation graph”





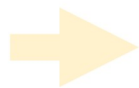
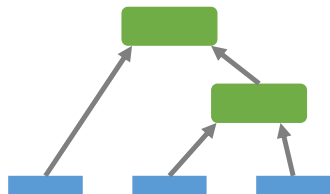


**Auto-
Parallelization**

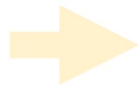


**Algebraic
Optimizer**

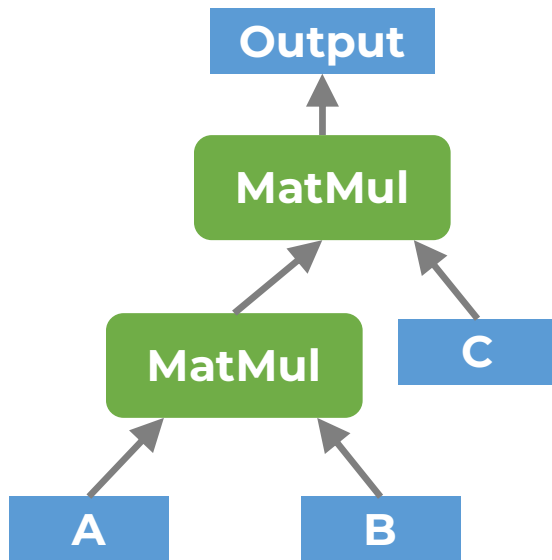
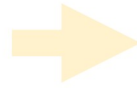


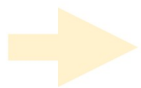
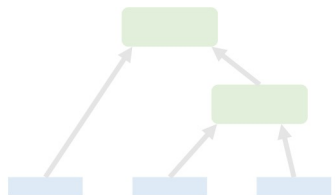


Auto-Parallelization

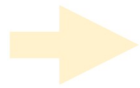


Algebraic Optimizer

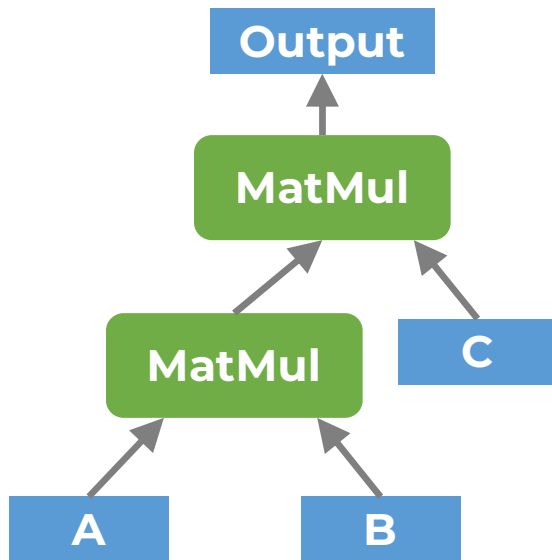
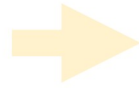


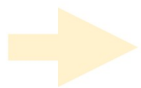
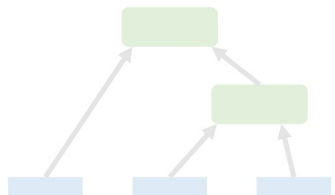


Auto-Parallelization

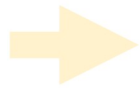


Algebraic Optimizer

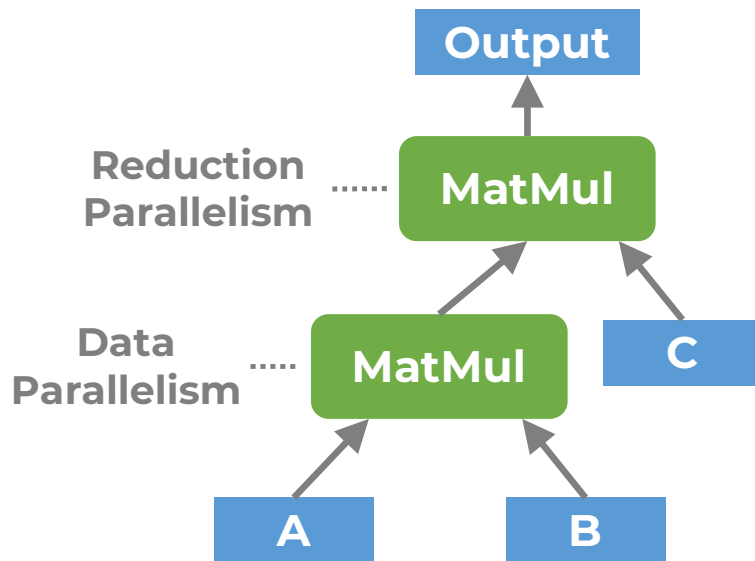
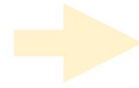


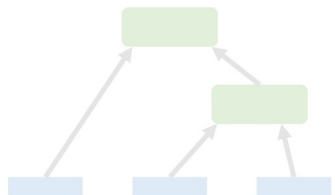


Auto-Parallelization

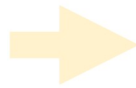


Algebraic Optimizer

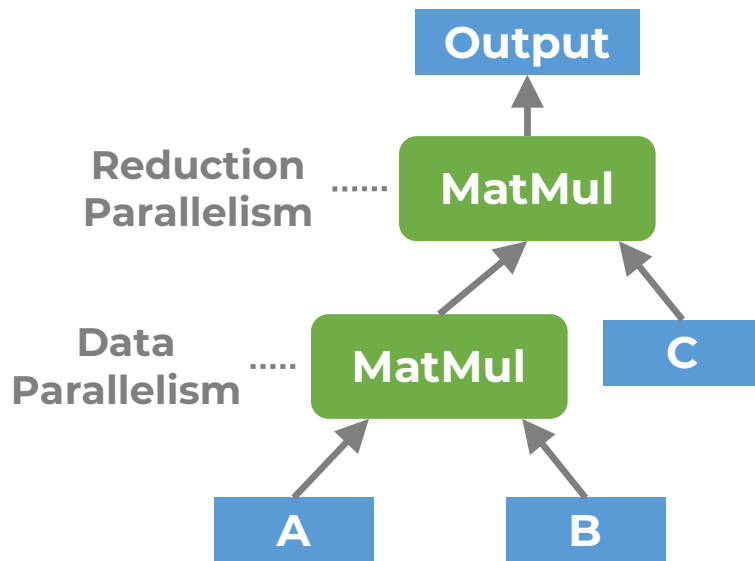
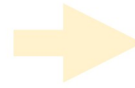


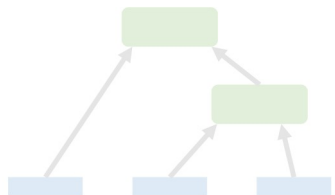


Auto-Parallelization

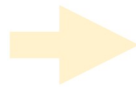


Algebraic Optimizer

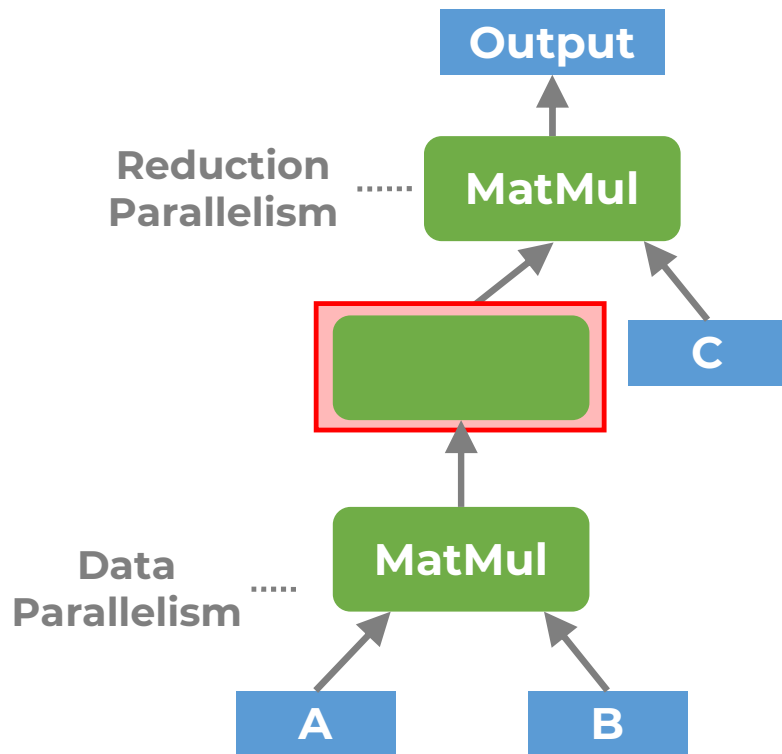
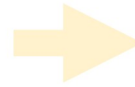


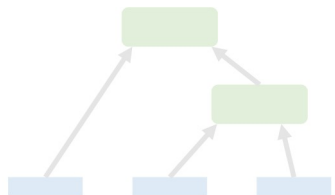


Auto-Parallelization

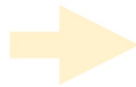


Algebraic Optimizer

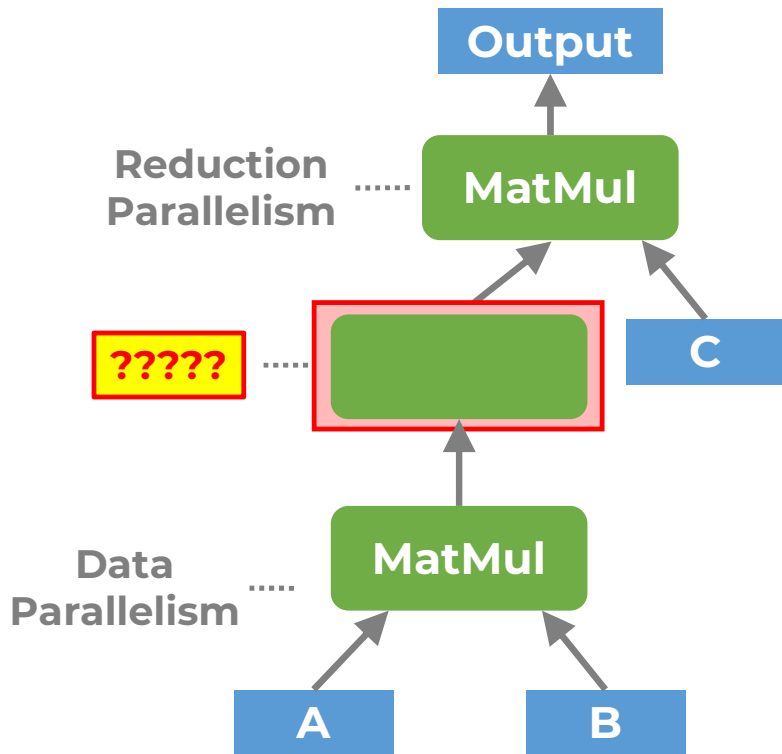
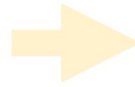


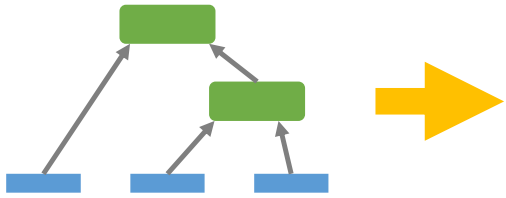


Auto-Parallelization

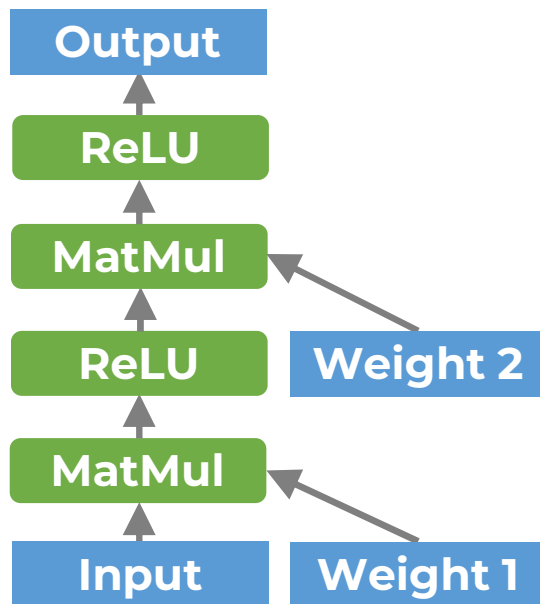


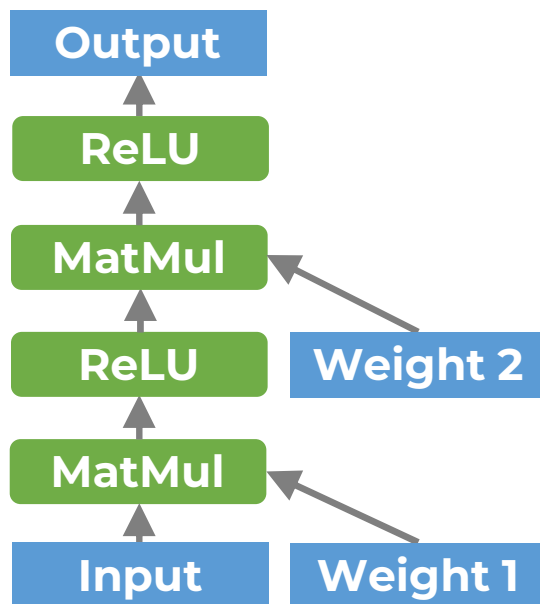
Algebraic Optimizer

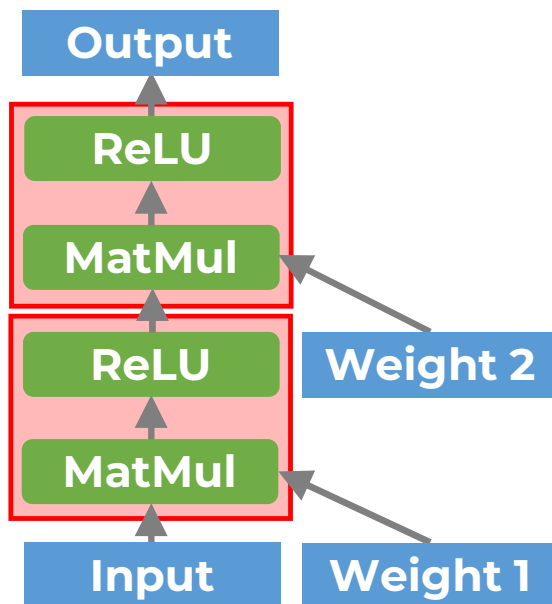


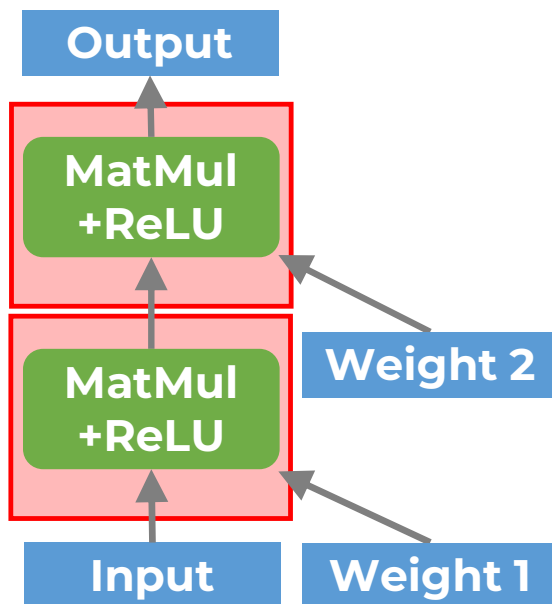


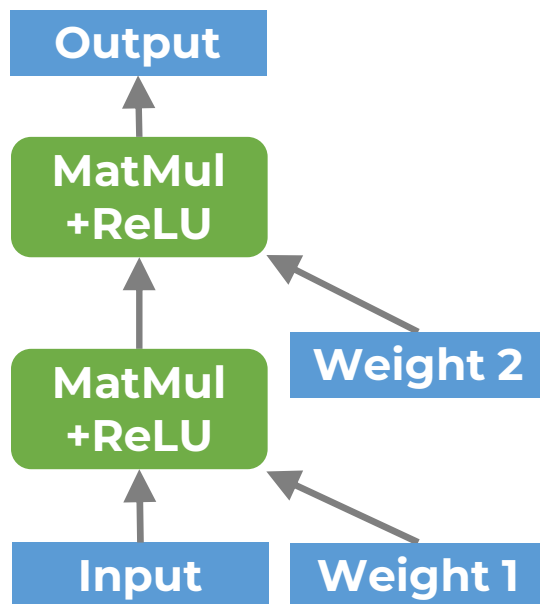


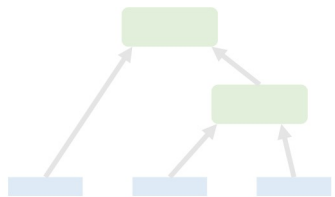




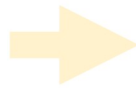




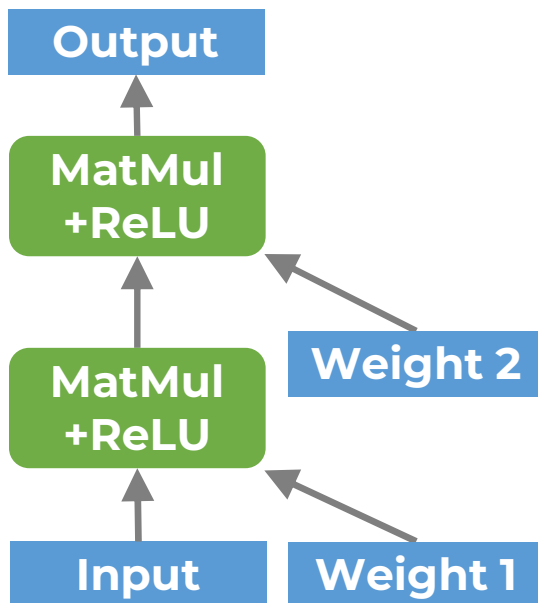
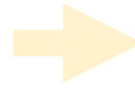


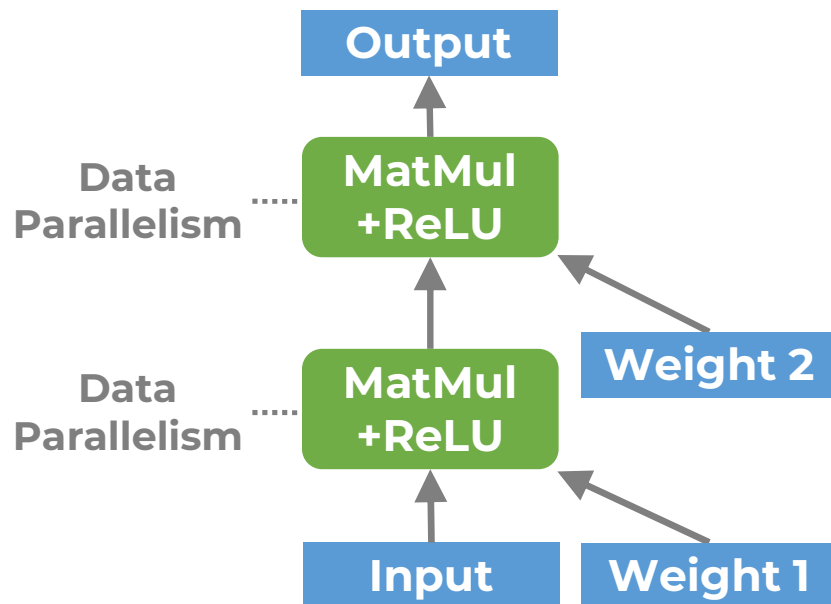


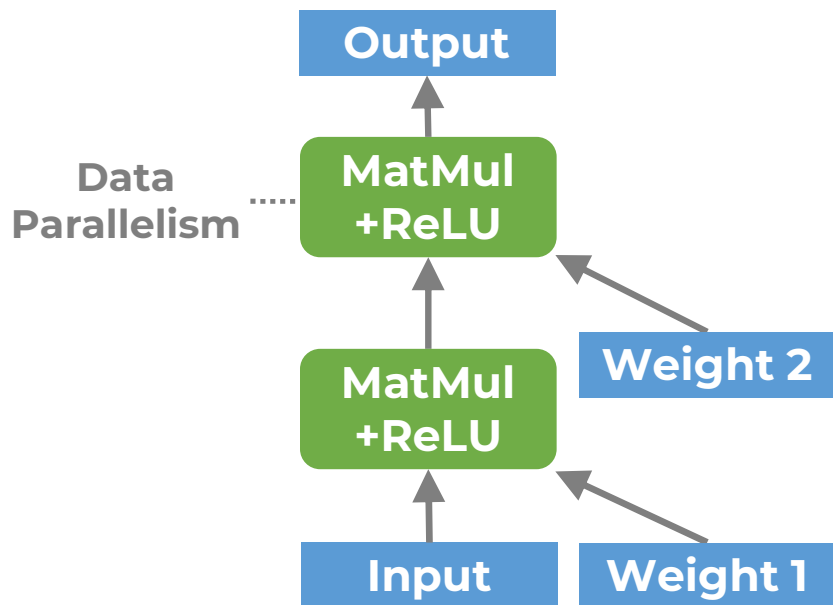
Algebraic
Optimizer

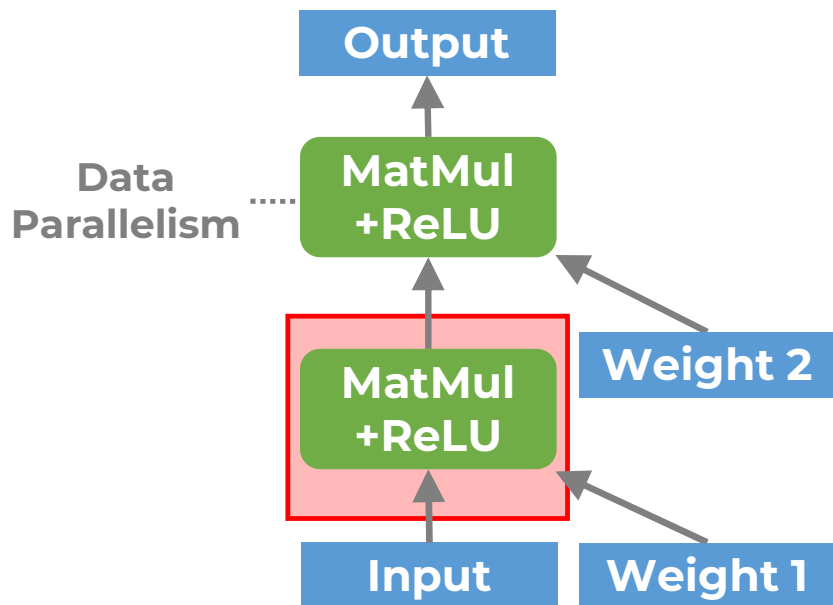


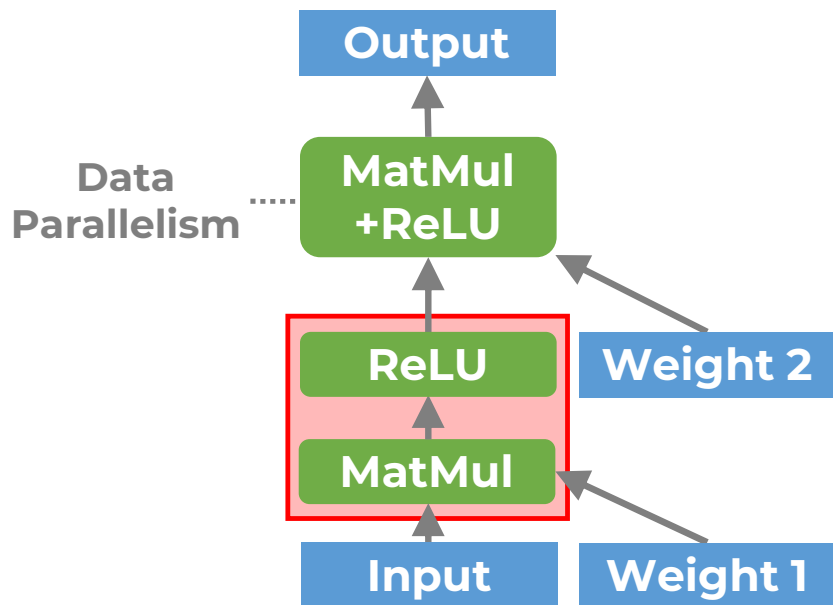
Auto-
Parallelization

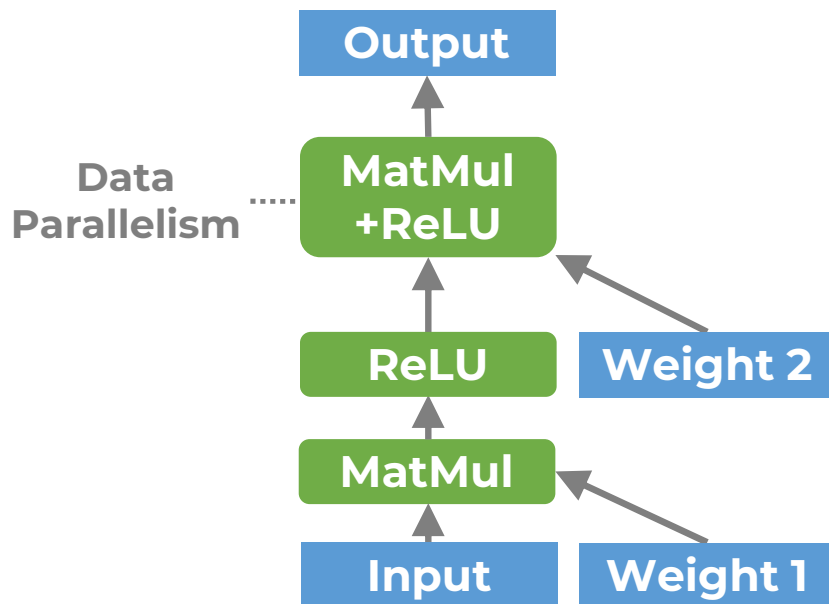
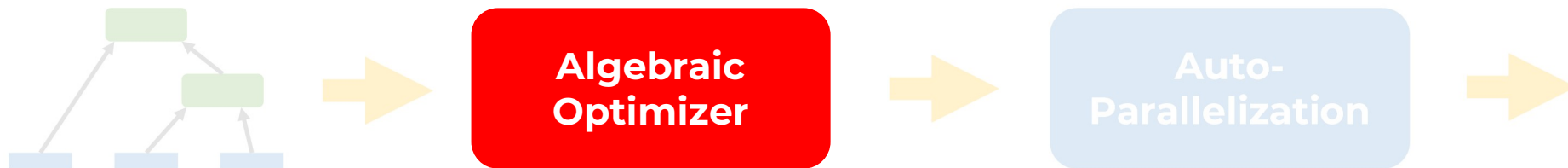


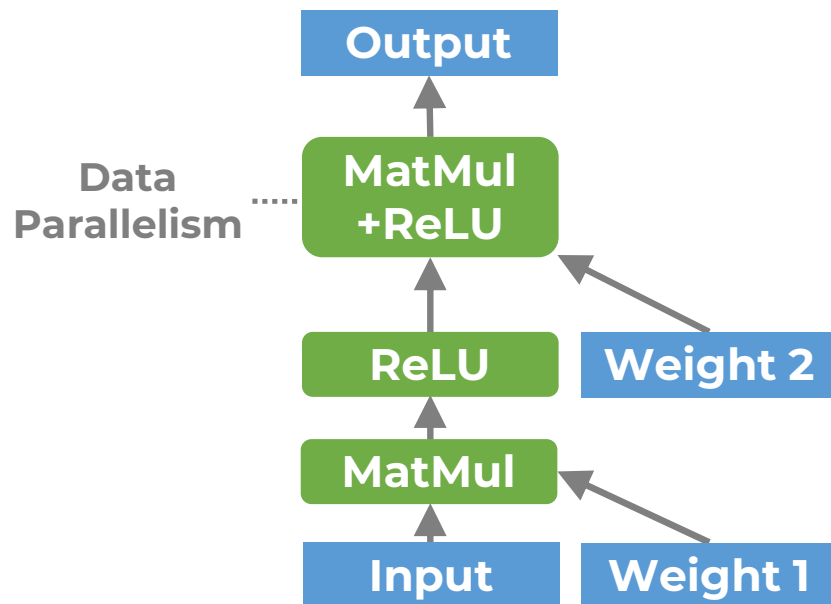


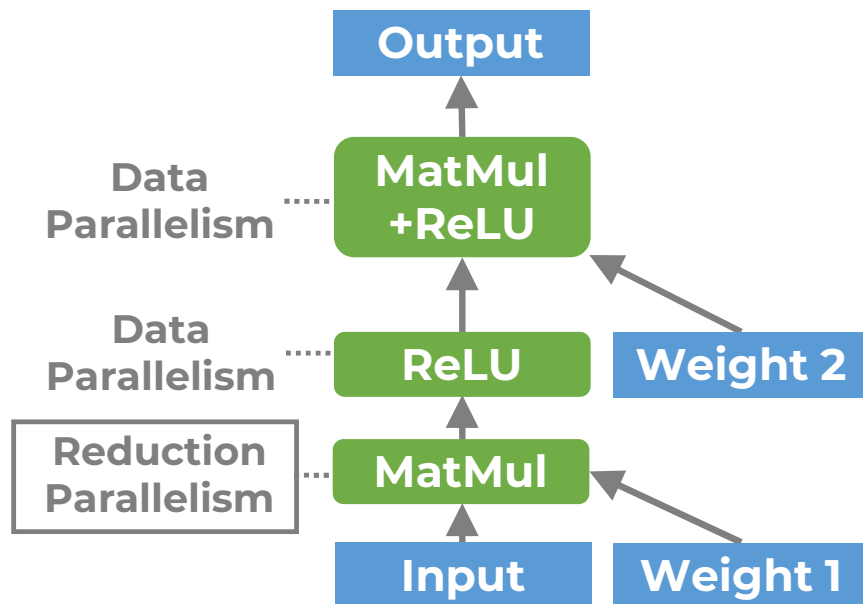


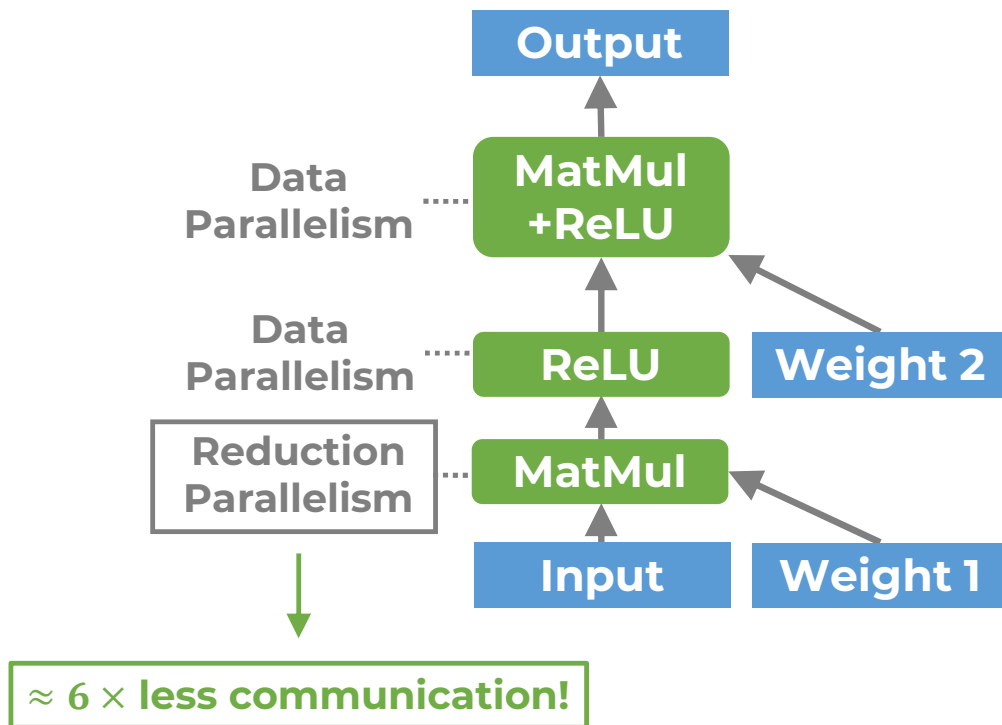
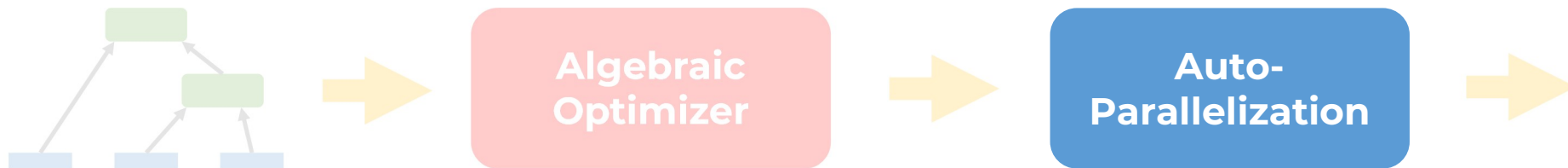




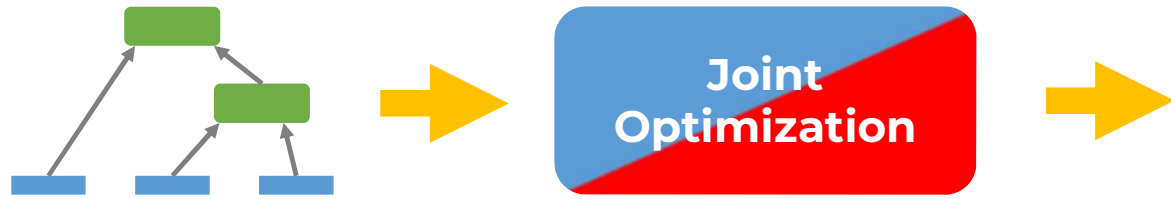






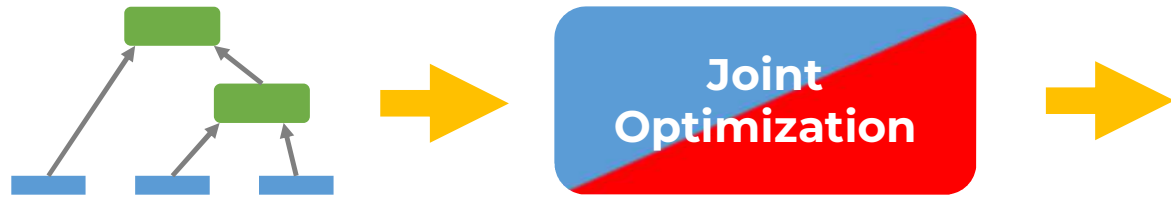






1.

2.



1. Representation

2.



-
1. Representation
 2. Scalability

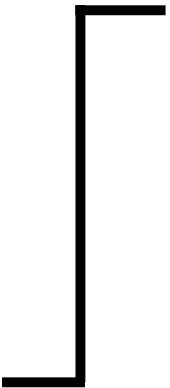
Unity

Unity — Representation

Representation

Parallel Computation
Graph (PCG)

Unity

A diagram consisting of a horizontal line extending from the right side of the word 'Unity', followed by a vertical line segment that turns right to point towards the text 'Parallel Computation Graph (PCG)'.

Representation

Parallel Computation
Graph (PCG)

Unity

Scalability

Representation

Parallel Computation
Graph (PCG)

Unity

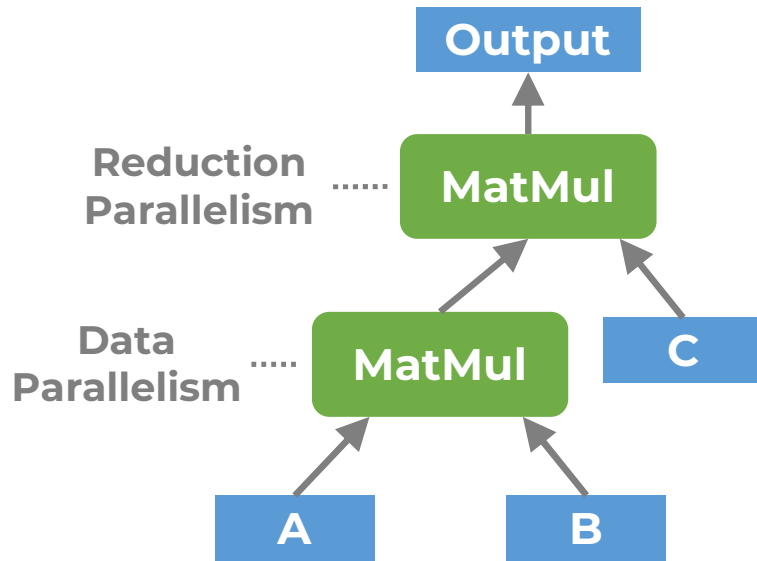
Scalability

Hierarchical Search
Algorithm

Representation

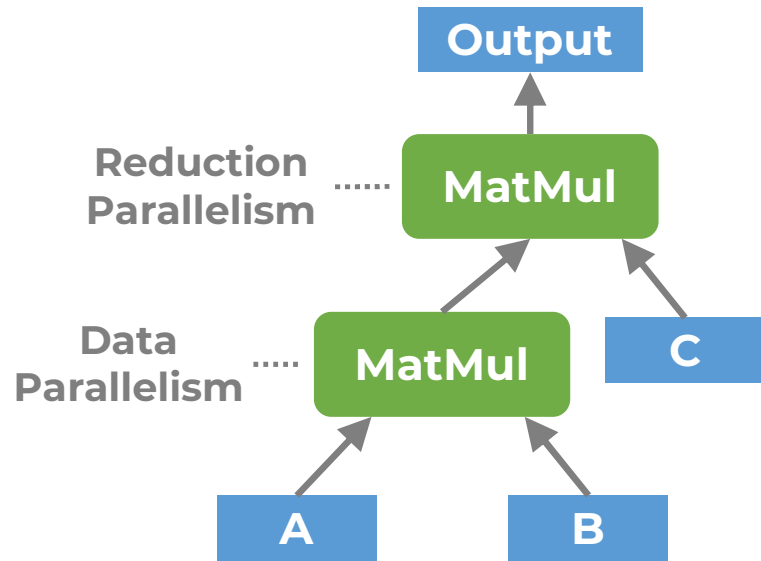
Parallel Computation Graph (PCG)

Parallel Computation Graph (PCG)

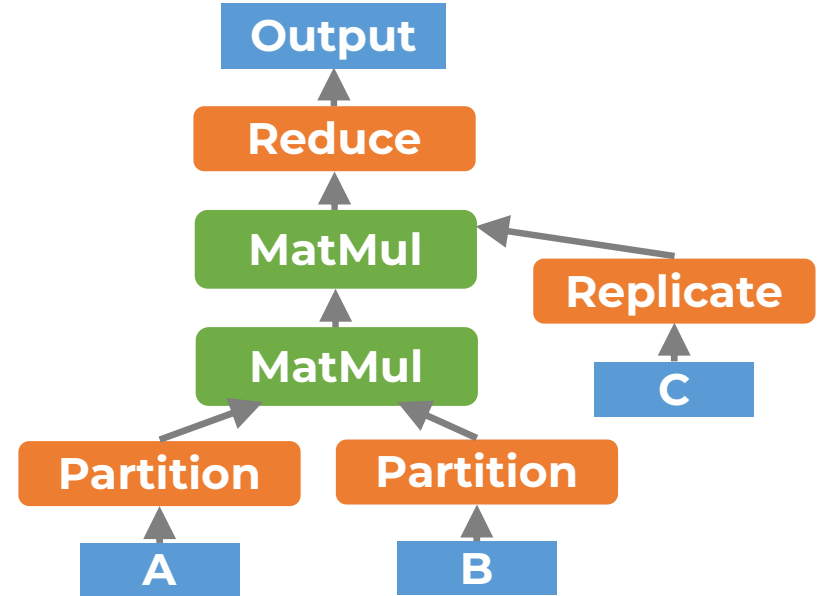


annotated computation graph

Parallel Computation Graph (PCG)

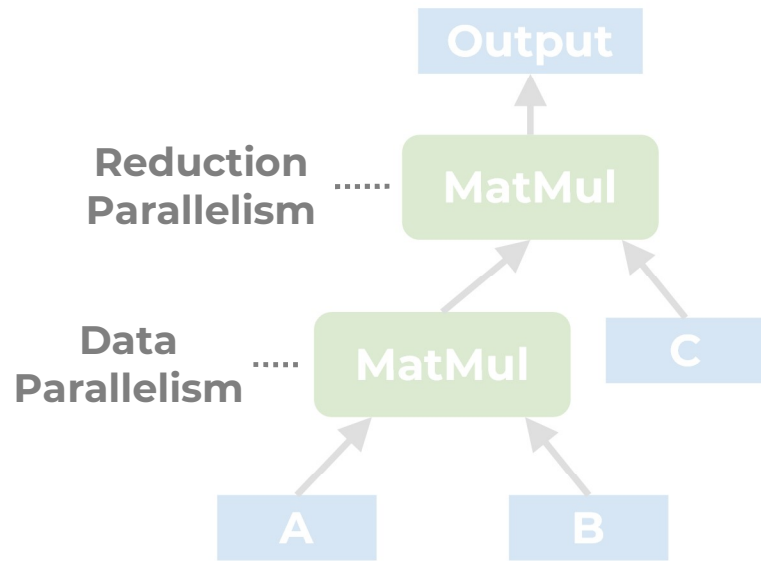


annotated computation graph

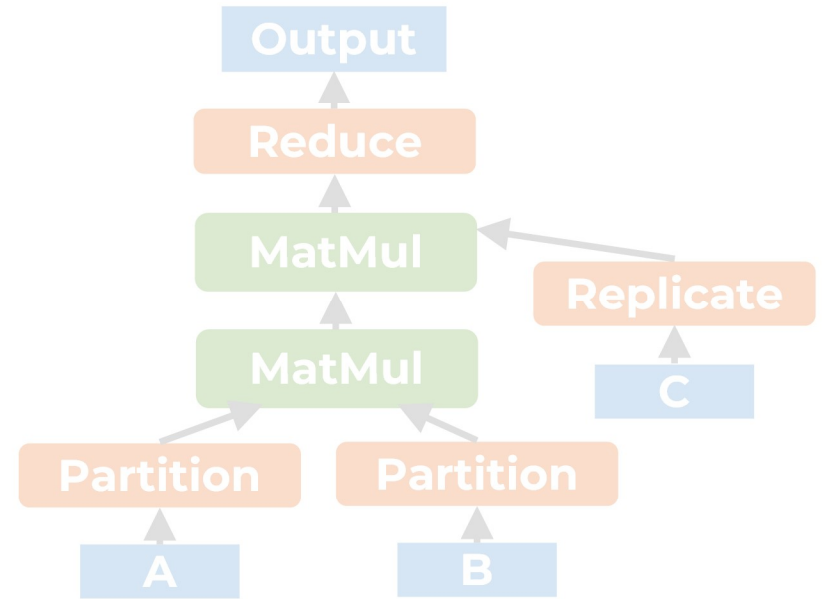
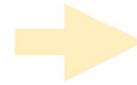


parallel computation graph (PCG)

Parallel Computation Graph (PCG)

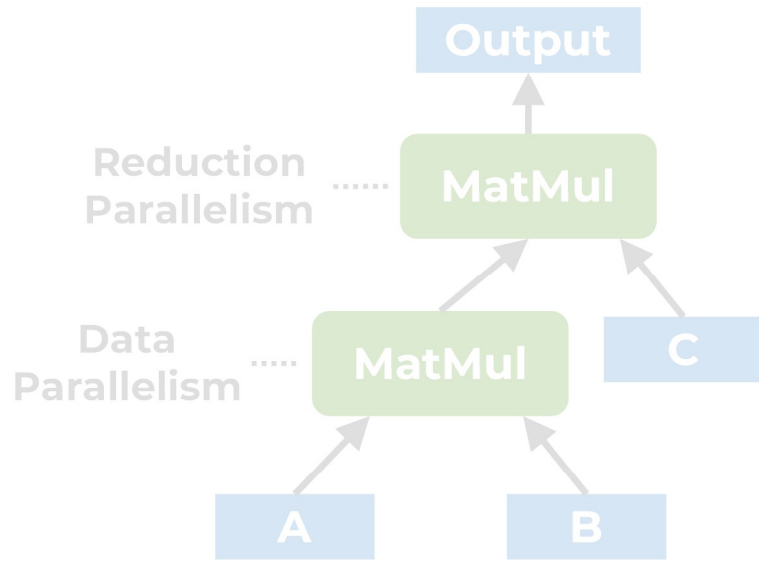


annotated computation graph

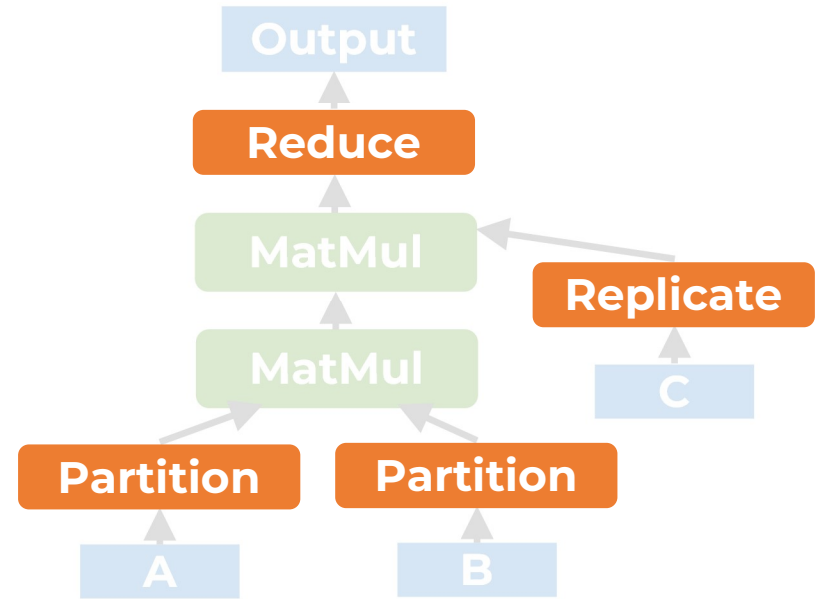
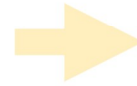


parallel computation graph (PCG)

Parallel Computation Graph (PCG)



annotated computation graph



parallel computation graph (PCG)

Partition

Combine

Replicate

Reduce

Pipeline

Batch

Partition



Combine

Replicate

Reduce

Pipeline

Batch



Partition

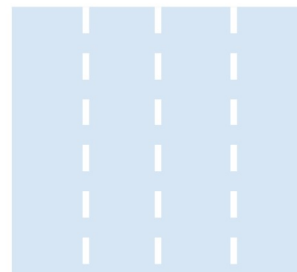
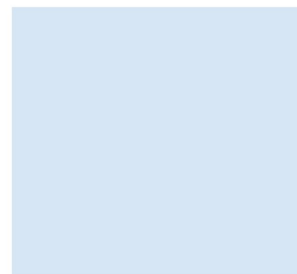
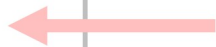
Combine

Replicate

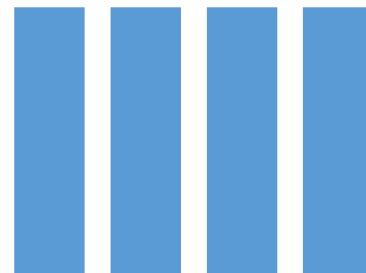
Reduce

Pipeline

Batch



Split



Partition



Combine

Replicate

Reduce

Pipeline

Batch



Partition

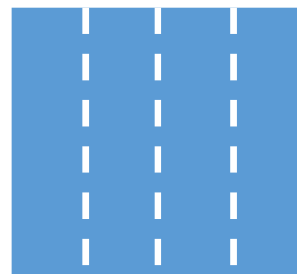
Combine

Replicate

Reduce

Pipeline

Batch



Partition

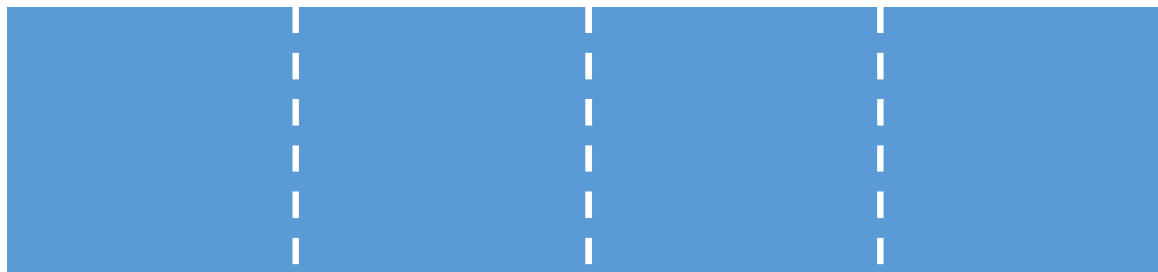
Combine

Replicate

Reduce

Pipeline

Batch



Partition

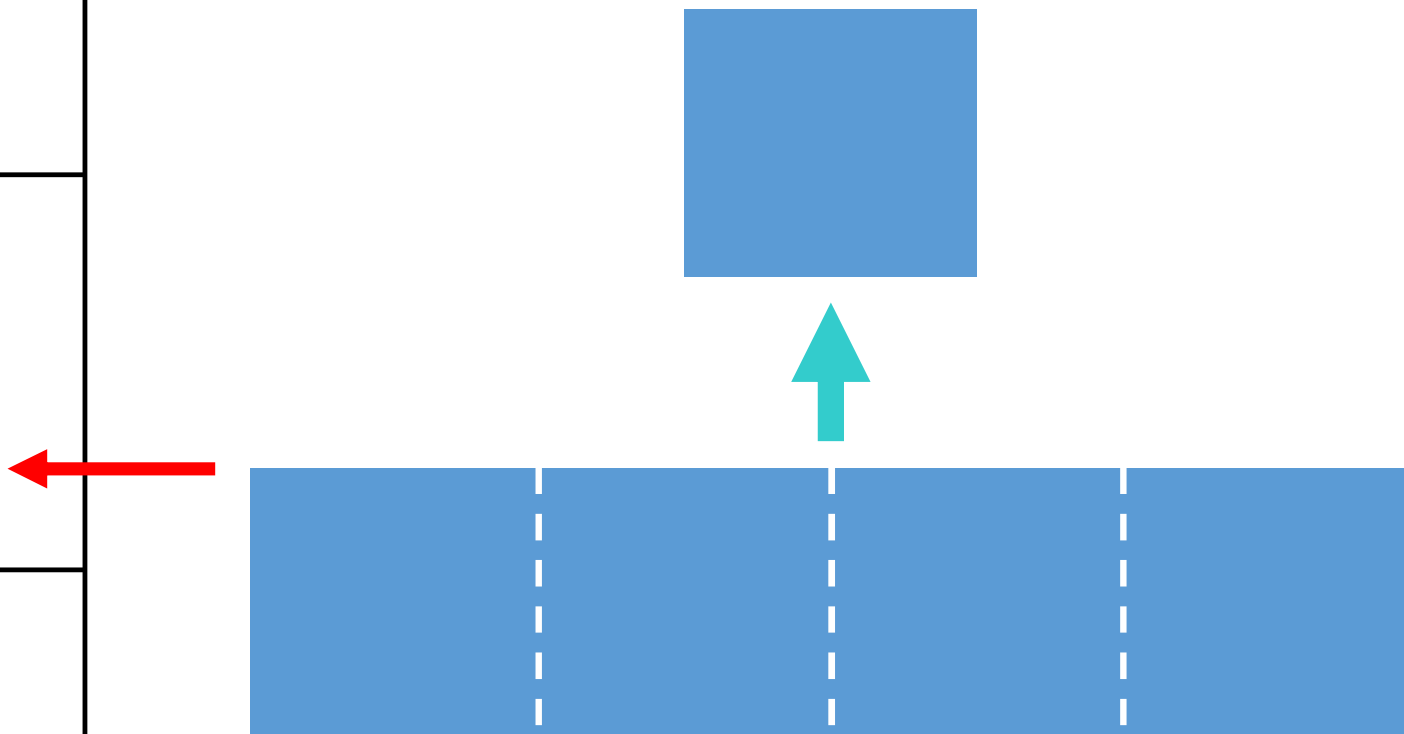
Combine

Replicate

Reduce

Pipeline

Batch



Partition

Combine

Replicate

Reduce

Pipeline

Batch



Partition

Combine

Replicate

Reduce

Pipeline

Batch



Partition

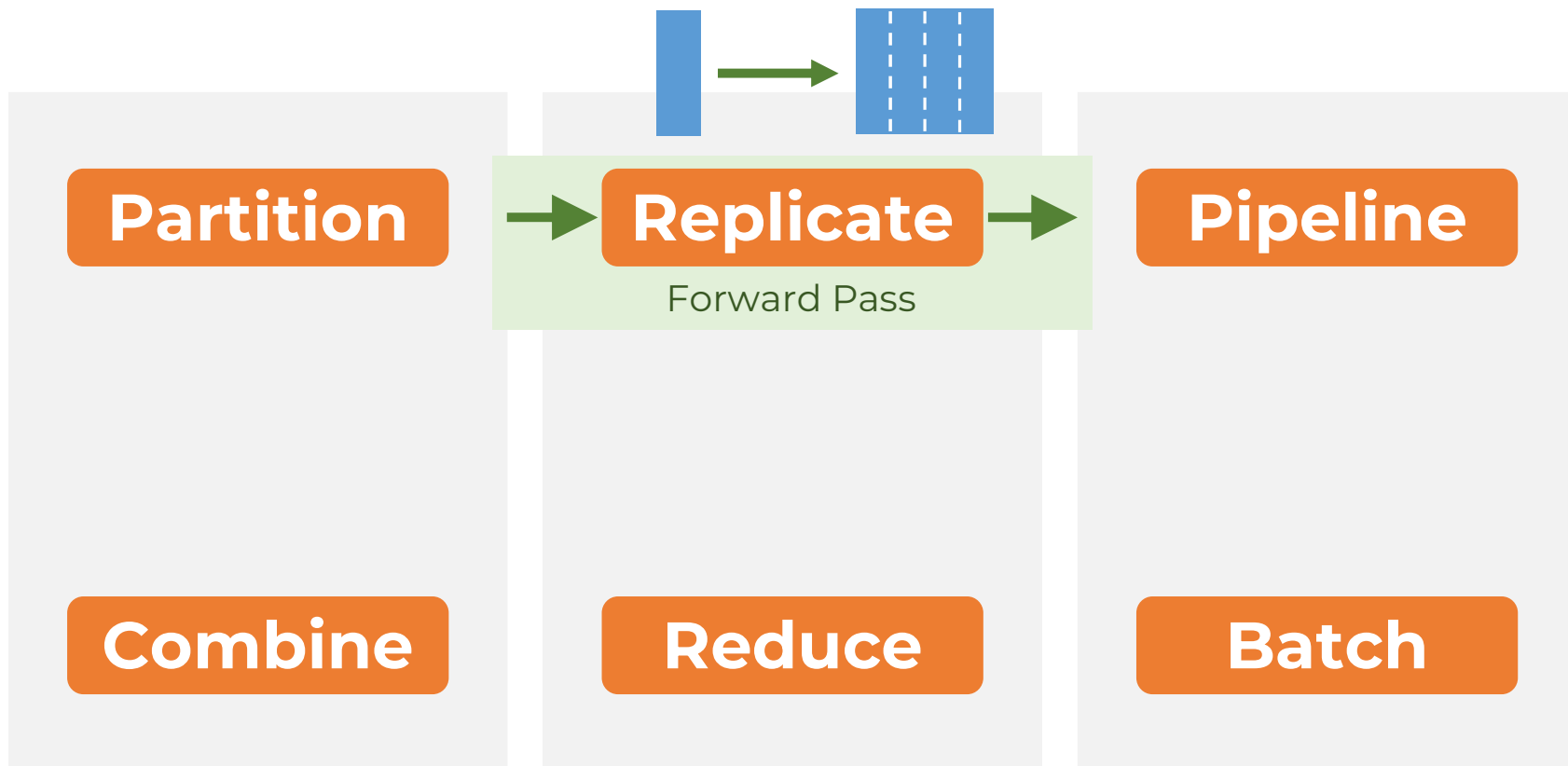
Combine

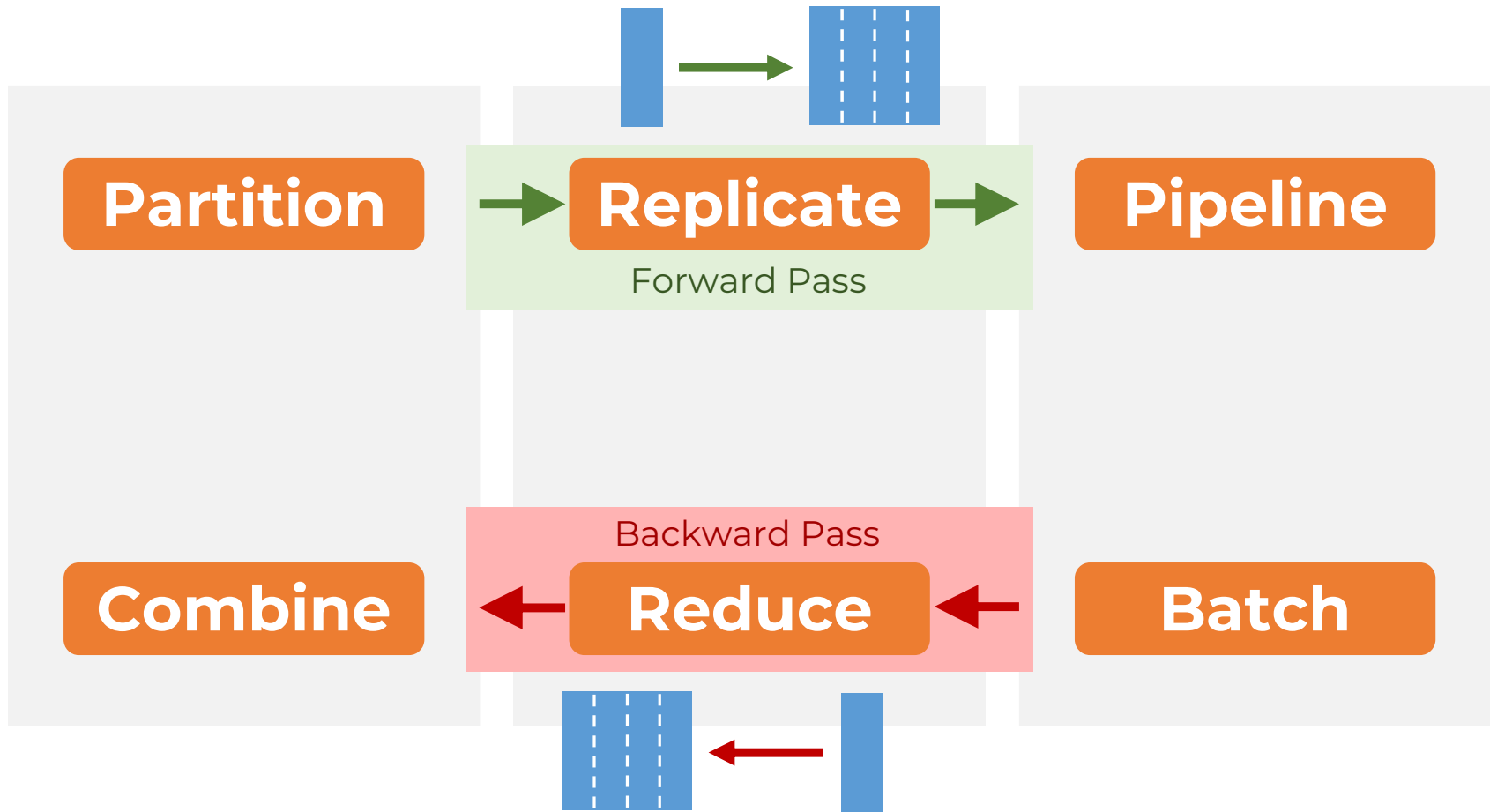
Replicate

Reduce

Pipeline

Batch





Partition

Combine

Replicate

Reduce

Pipeline

Batch

Partition

Replicate

Pipeline

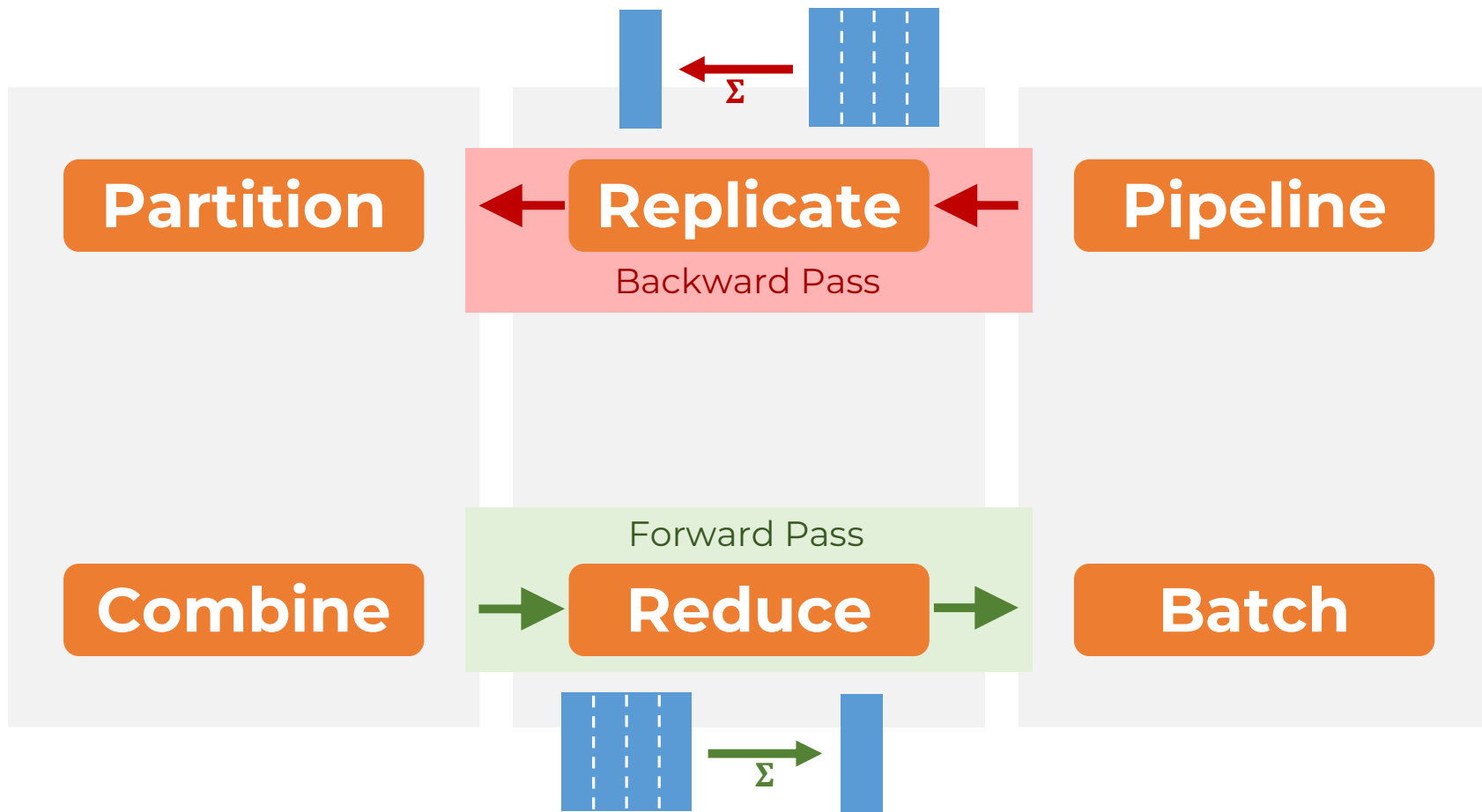
Combine

Forward Pass

Reduce

Batch





Partition

Combine

Replicate

Reduce

Pipeline

Batch

Partition



Combine

Replicate

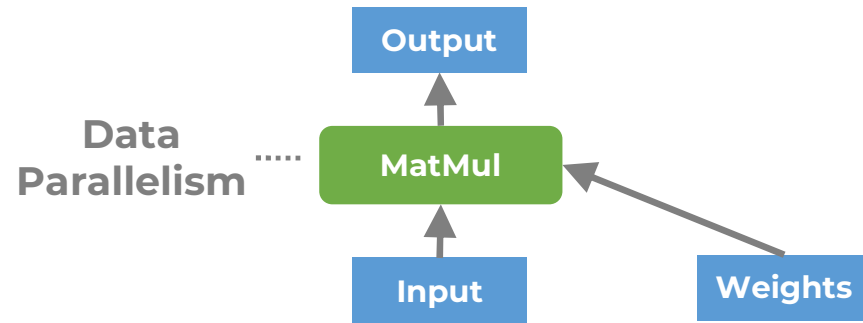


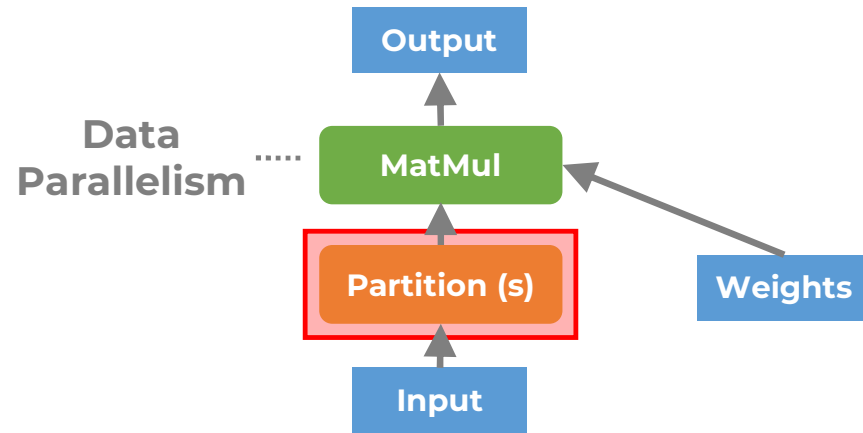
Reduce

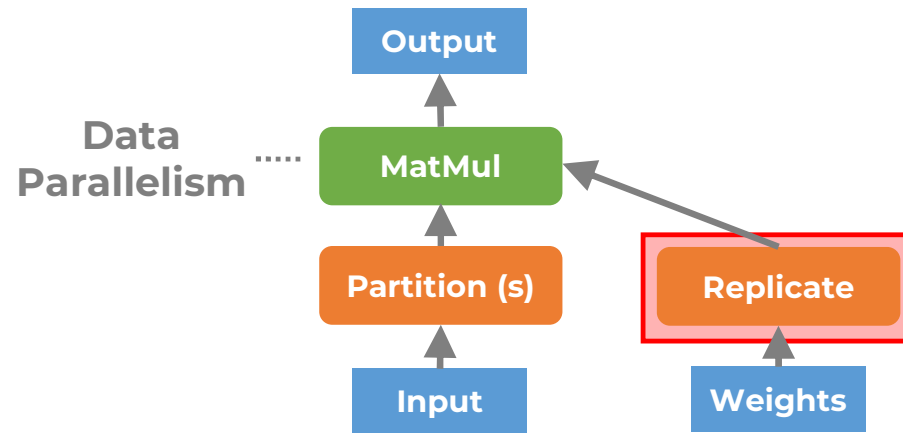
Pipeline

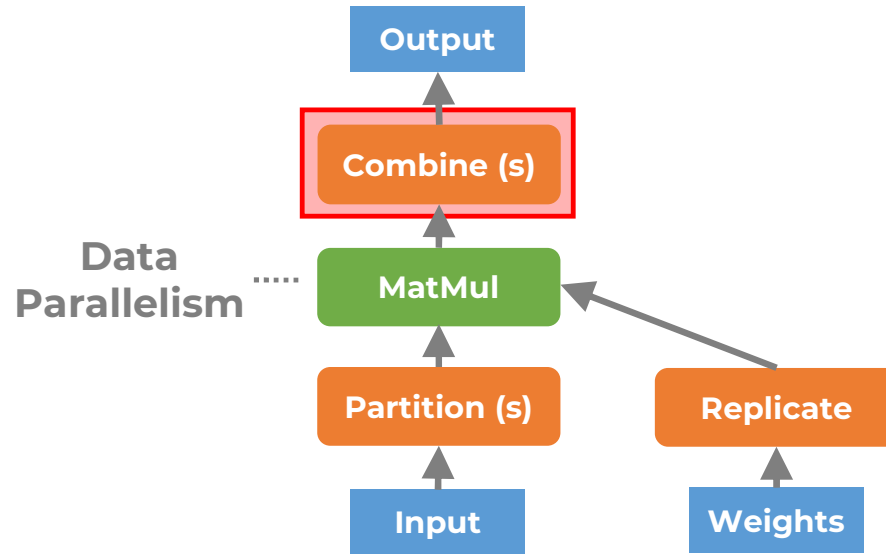


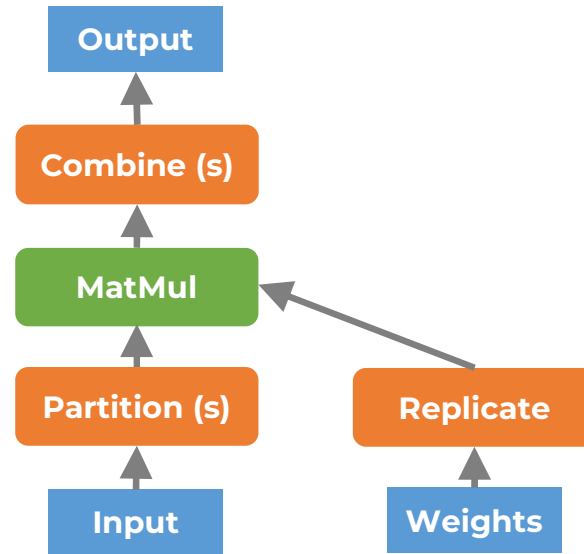
Batch

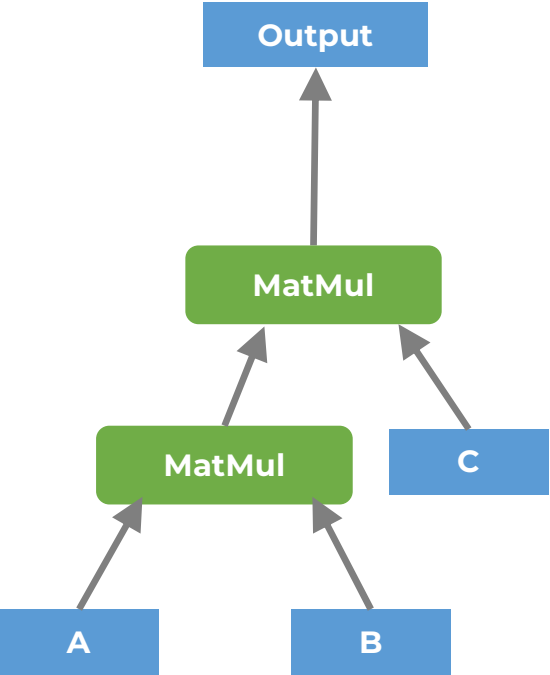


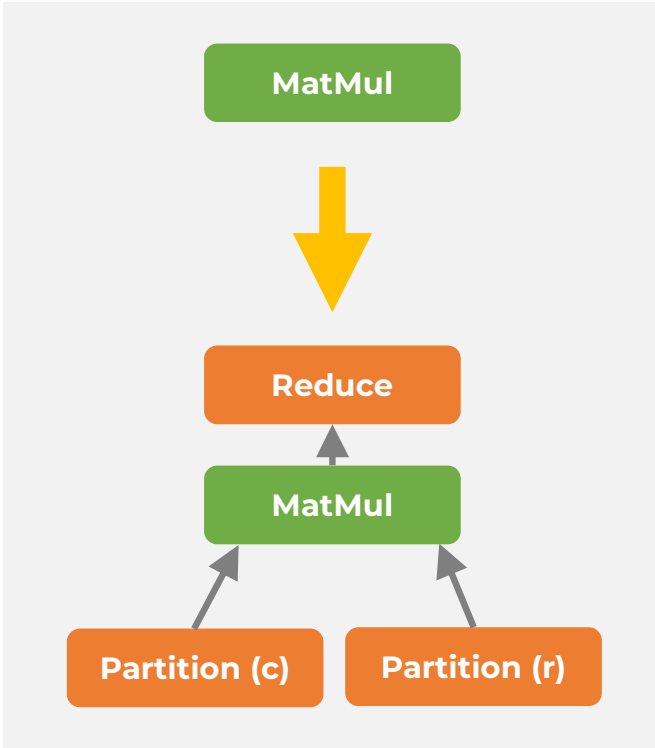
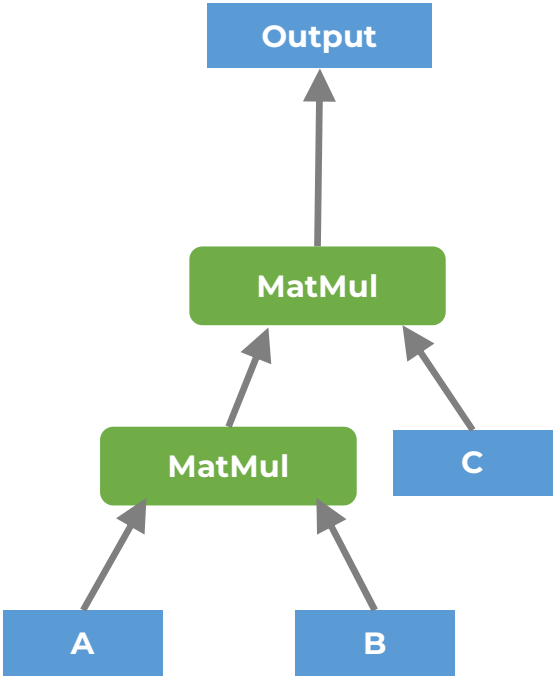


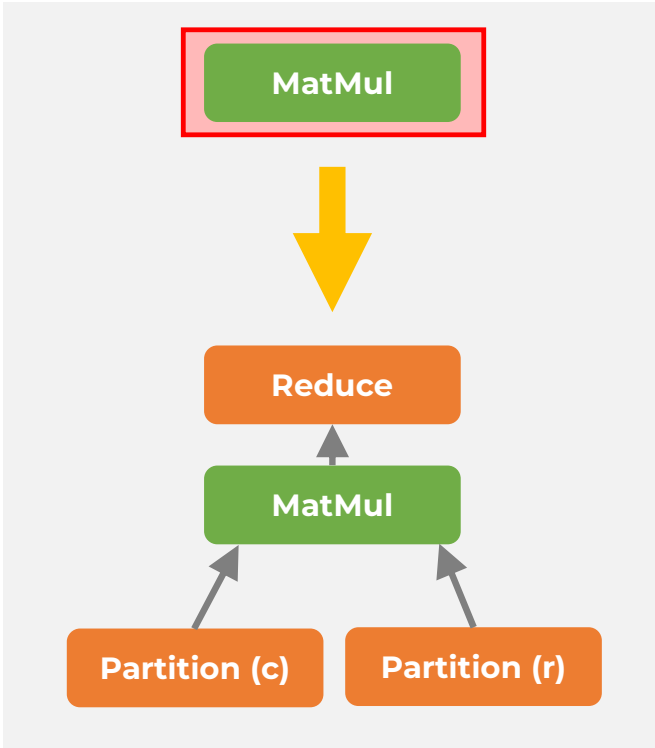
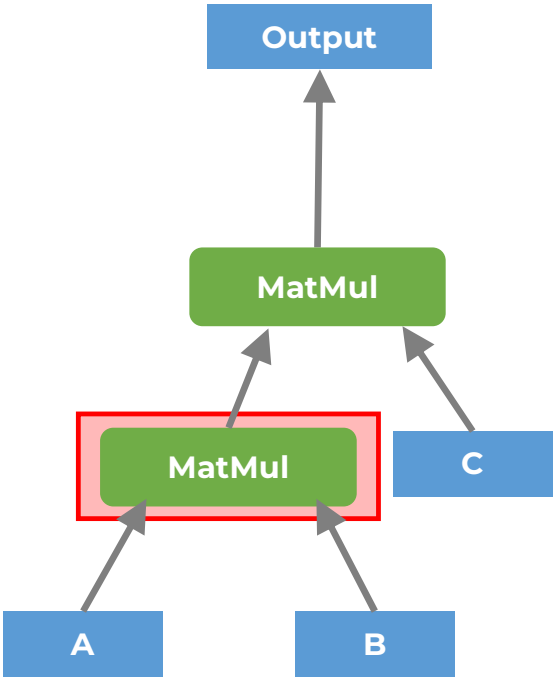


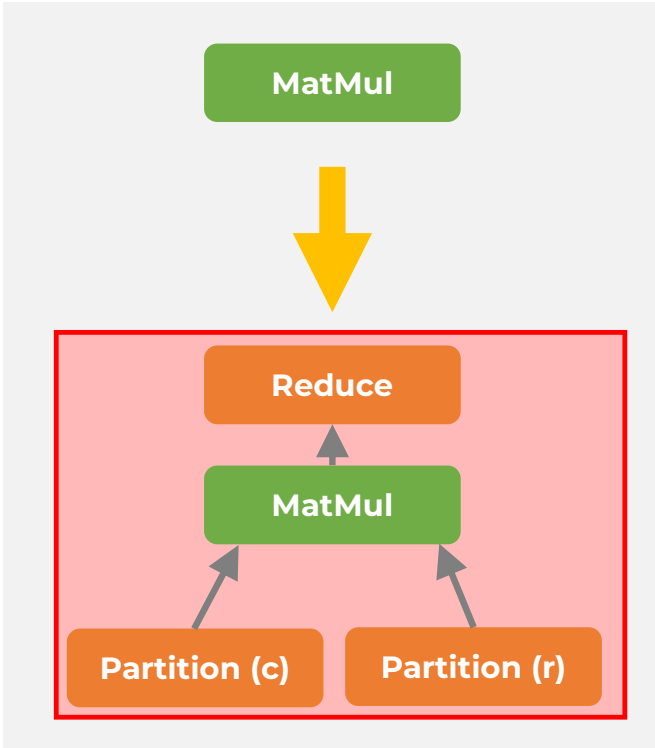
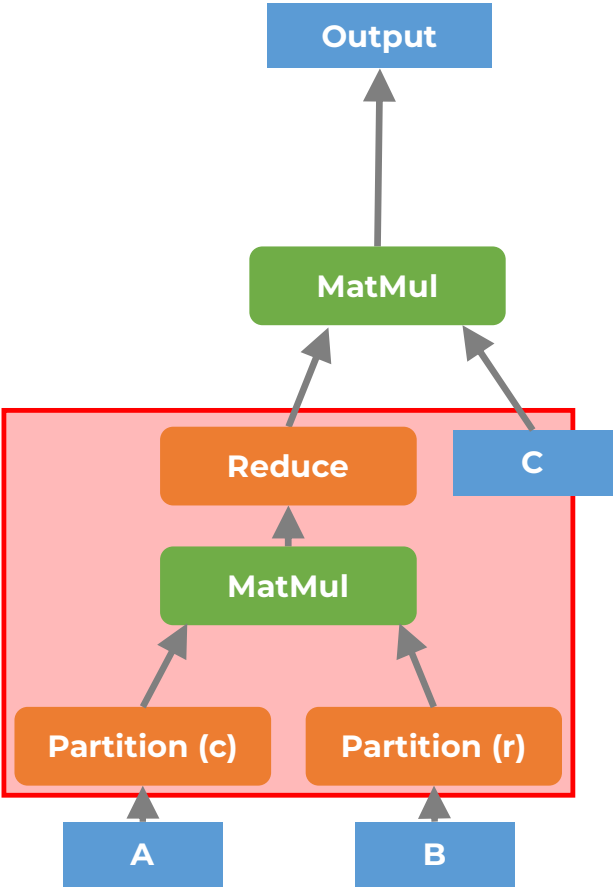


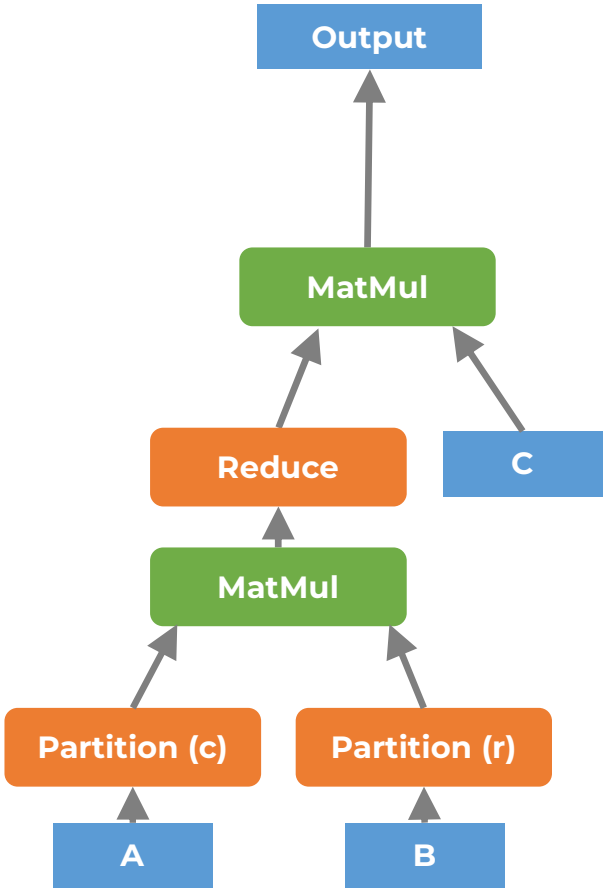


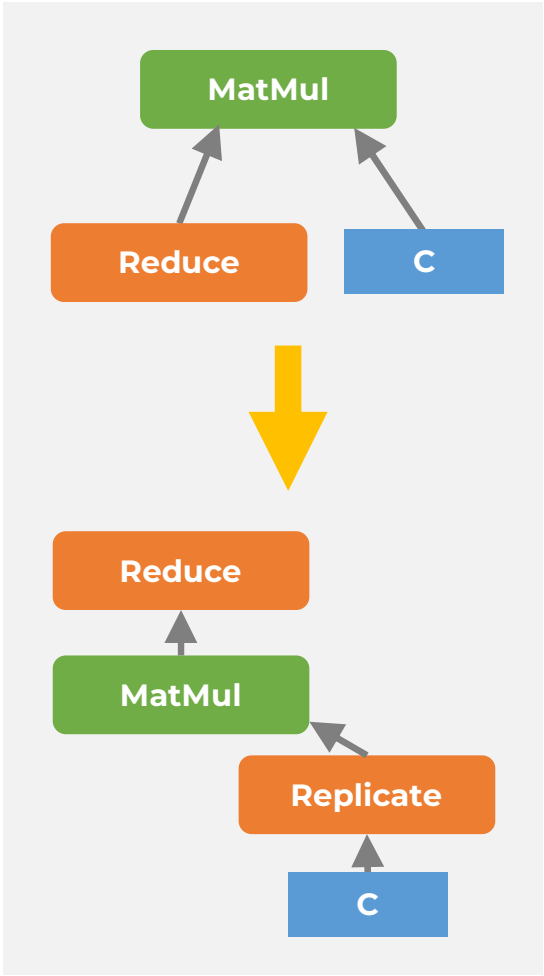
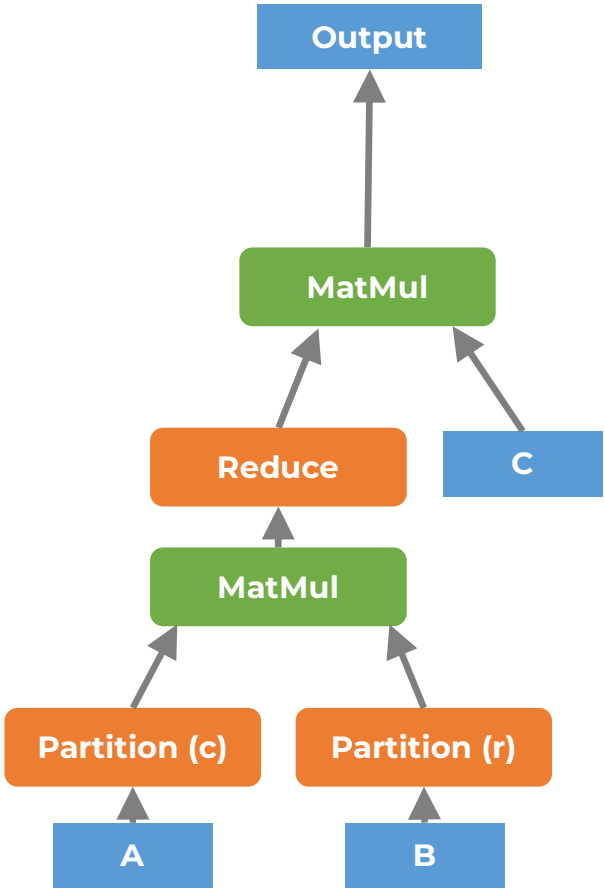


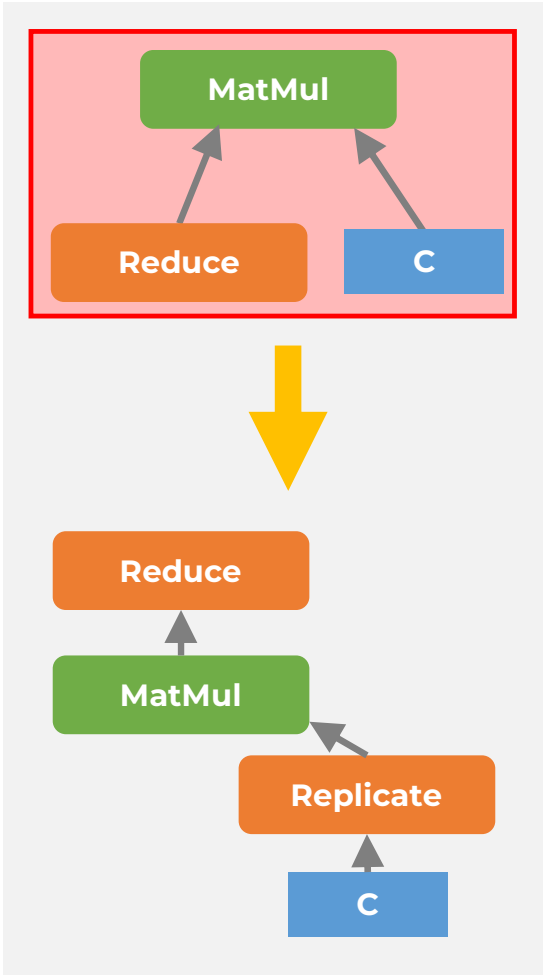
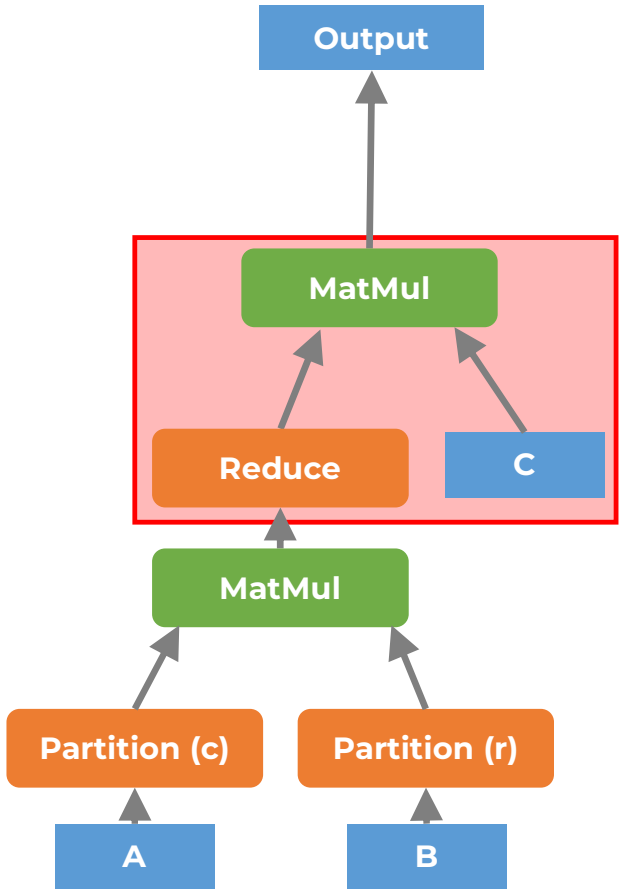


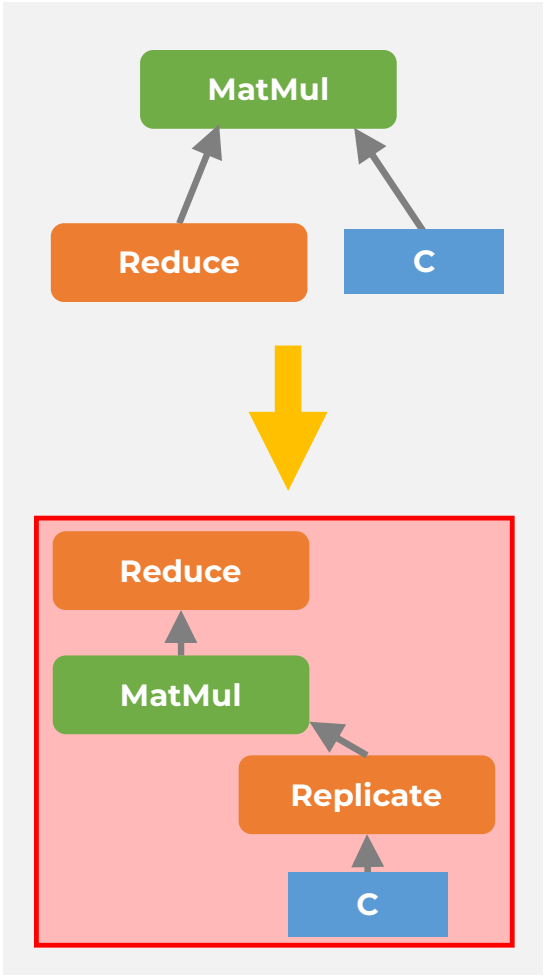
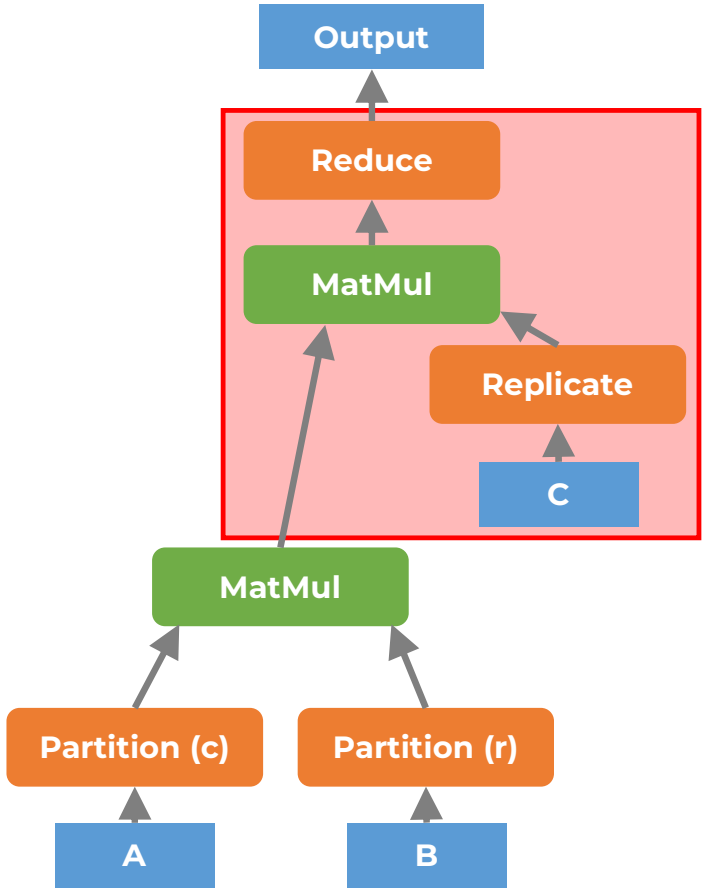


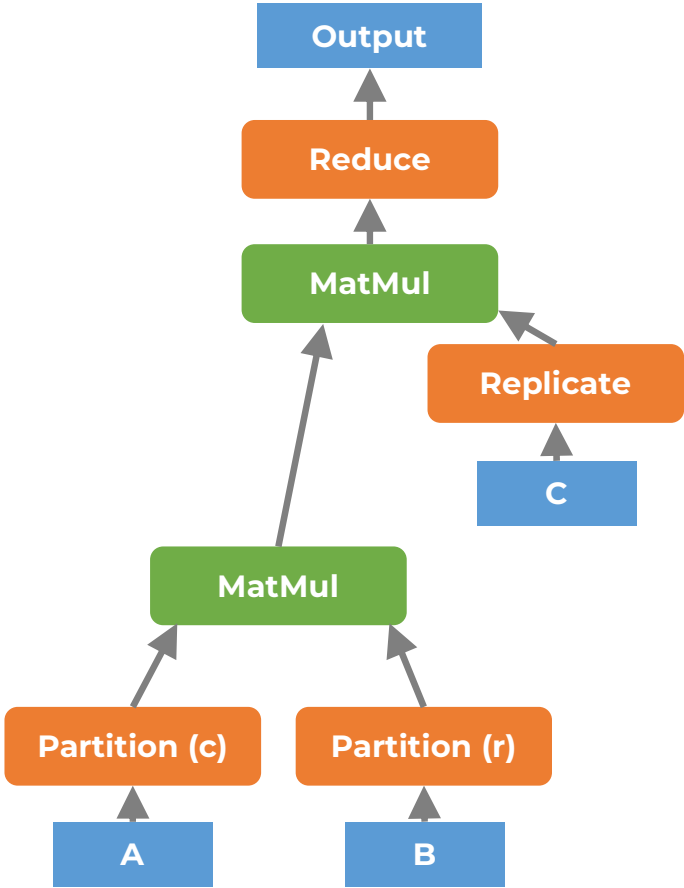












Separation of concerns

Separation of concerns

Automatically generate substitutions

Separation of concerns

Automatically generate substitutions

New operators

Separation of concerns

Automatically generate substitutions

New operators

New forms of parallelism

Separation of concerns

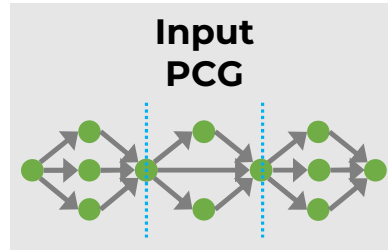
Explicitly represents communication

Separation of concerns

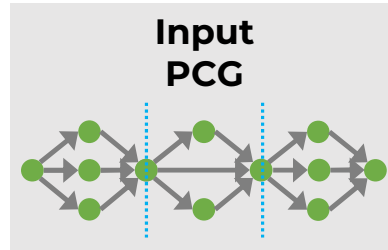
Explicitly represents communication

Concise

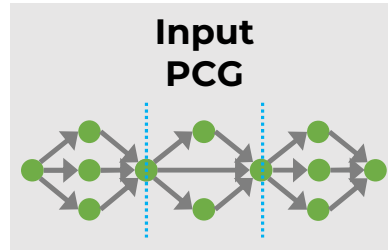
Hierarchical Search Algorithm



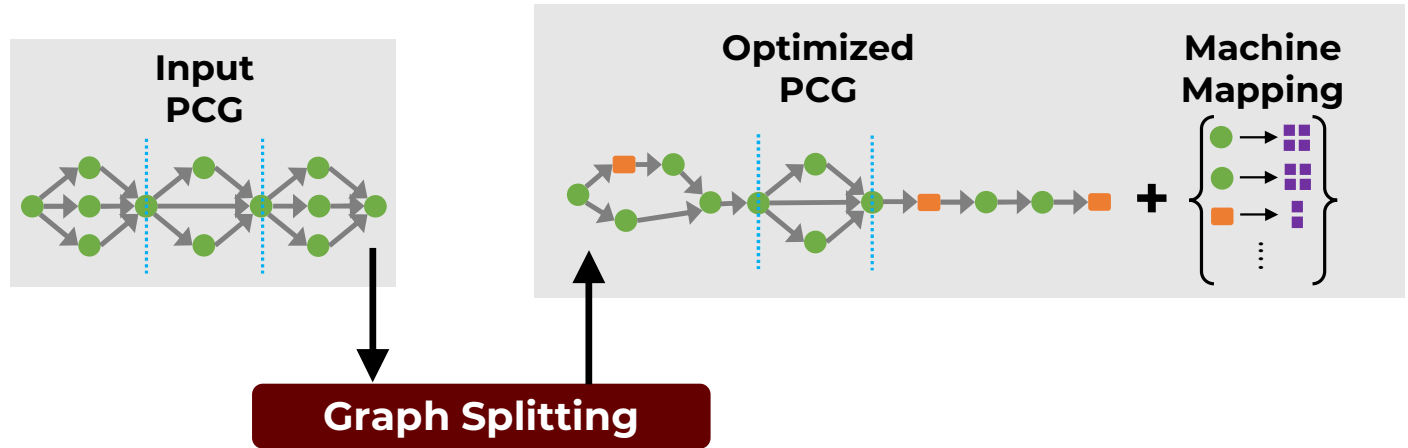
Hierarchical Search Algorithm



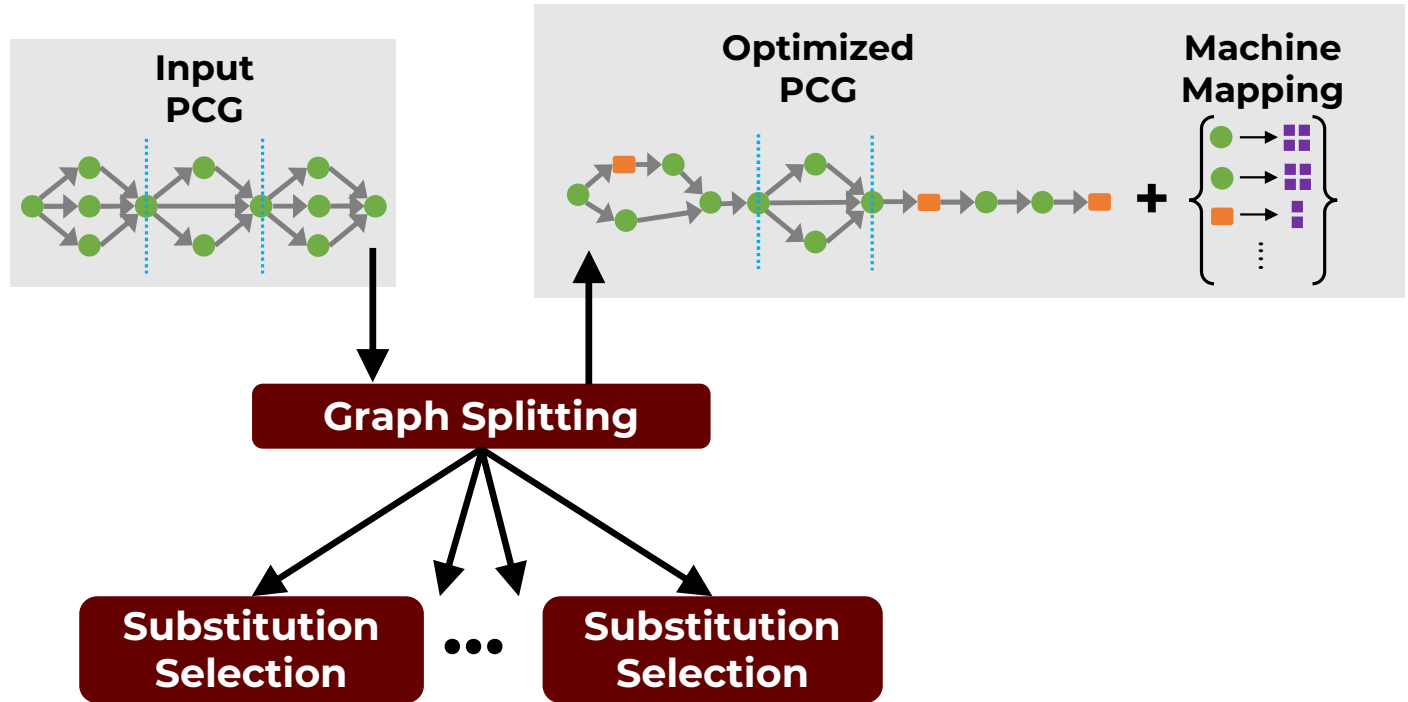
Hierarchical Search Algorithm



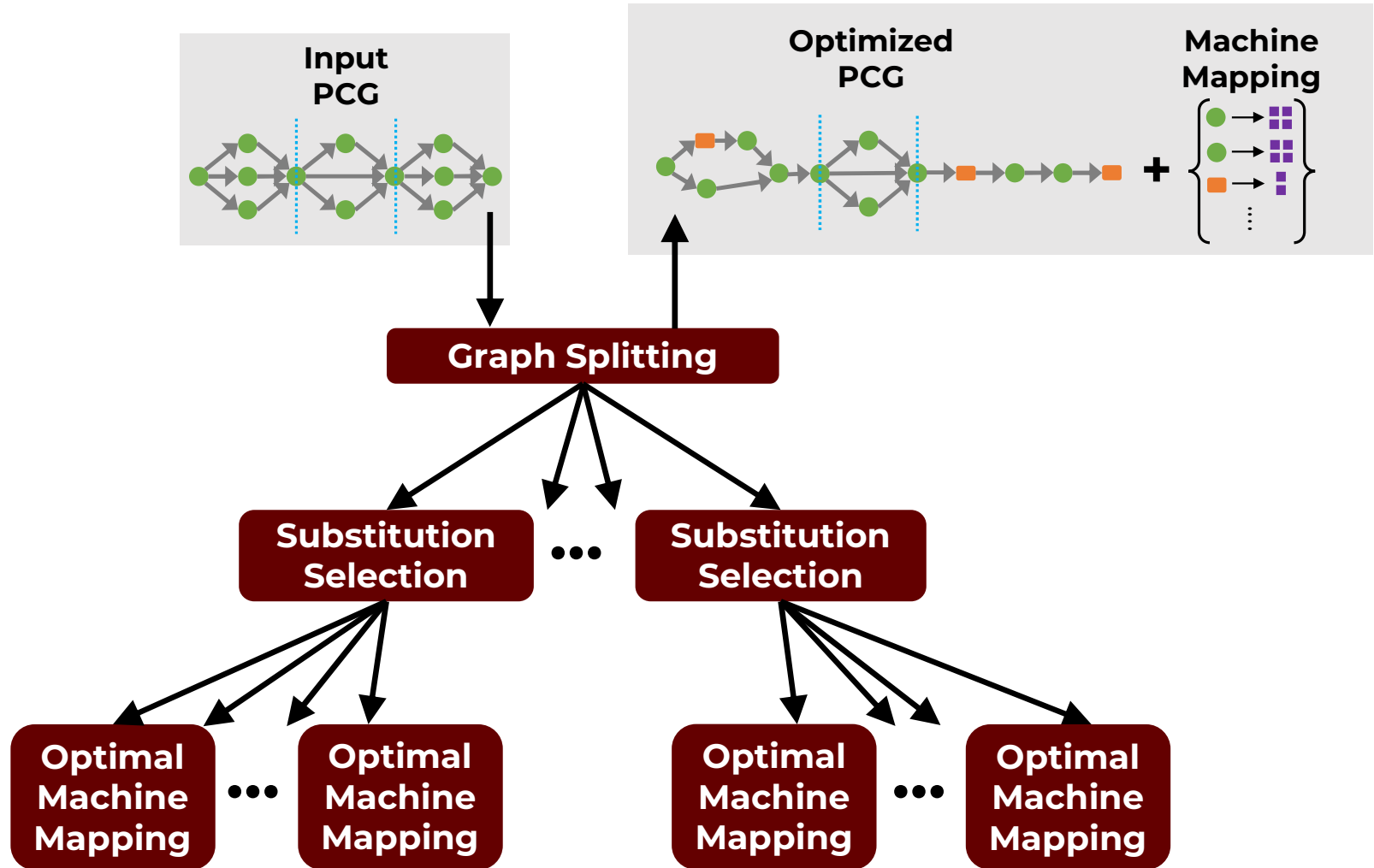
Hierarchical Search Algorithm



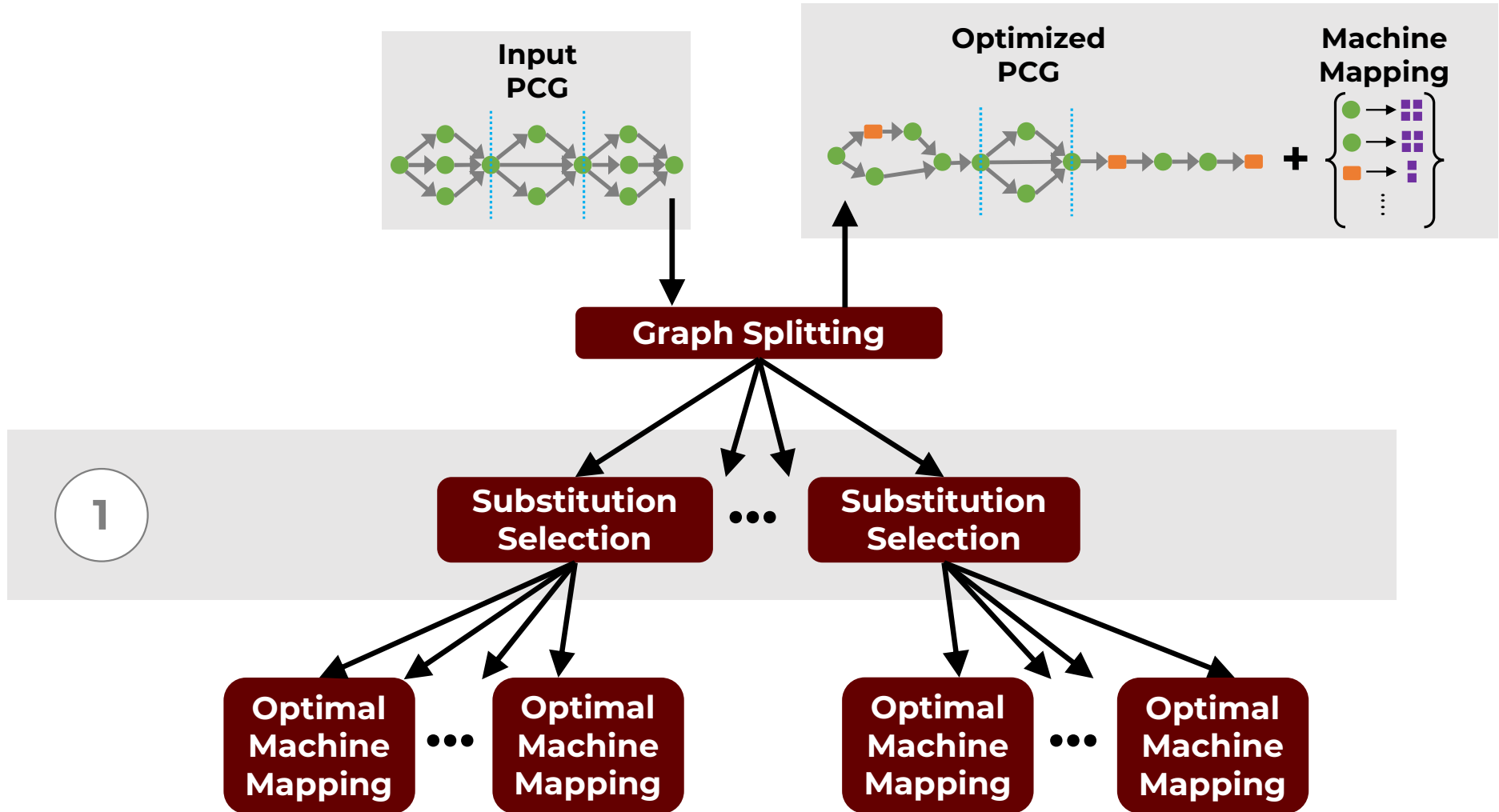
Hierarchical Search Algorithm



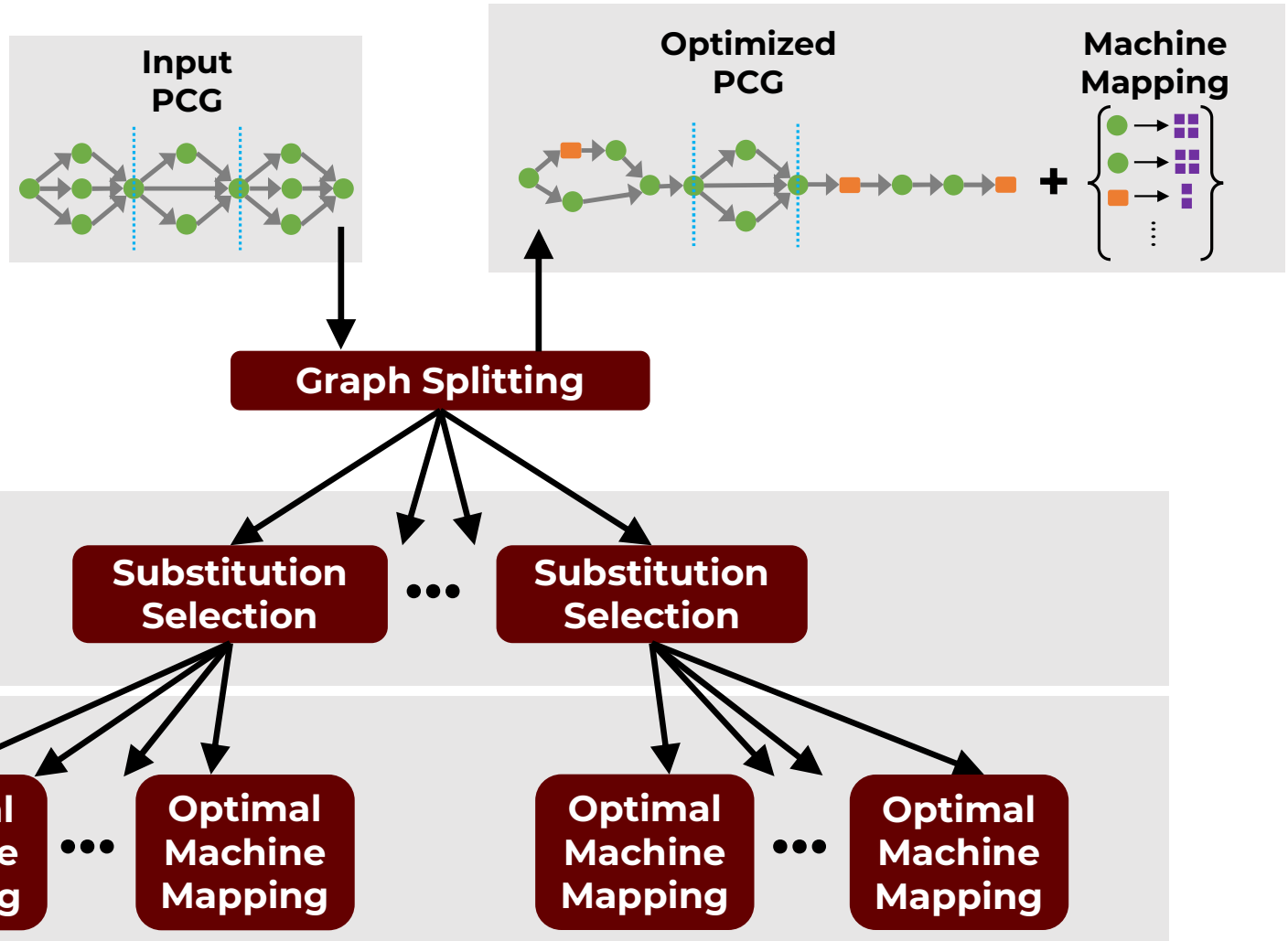
Hierarchical Search Algorithm



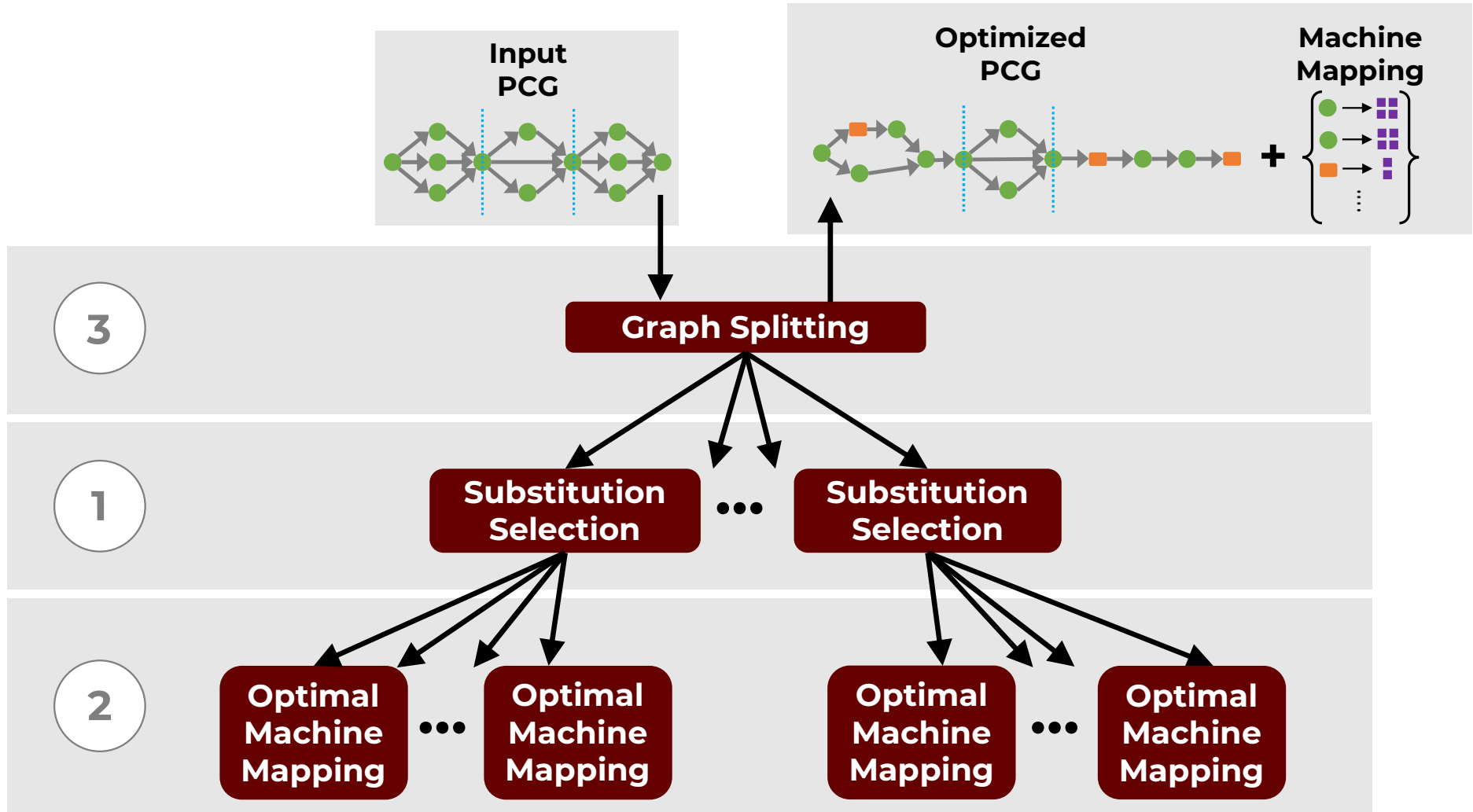
Hierarchical Search Algorithm



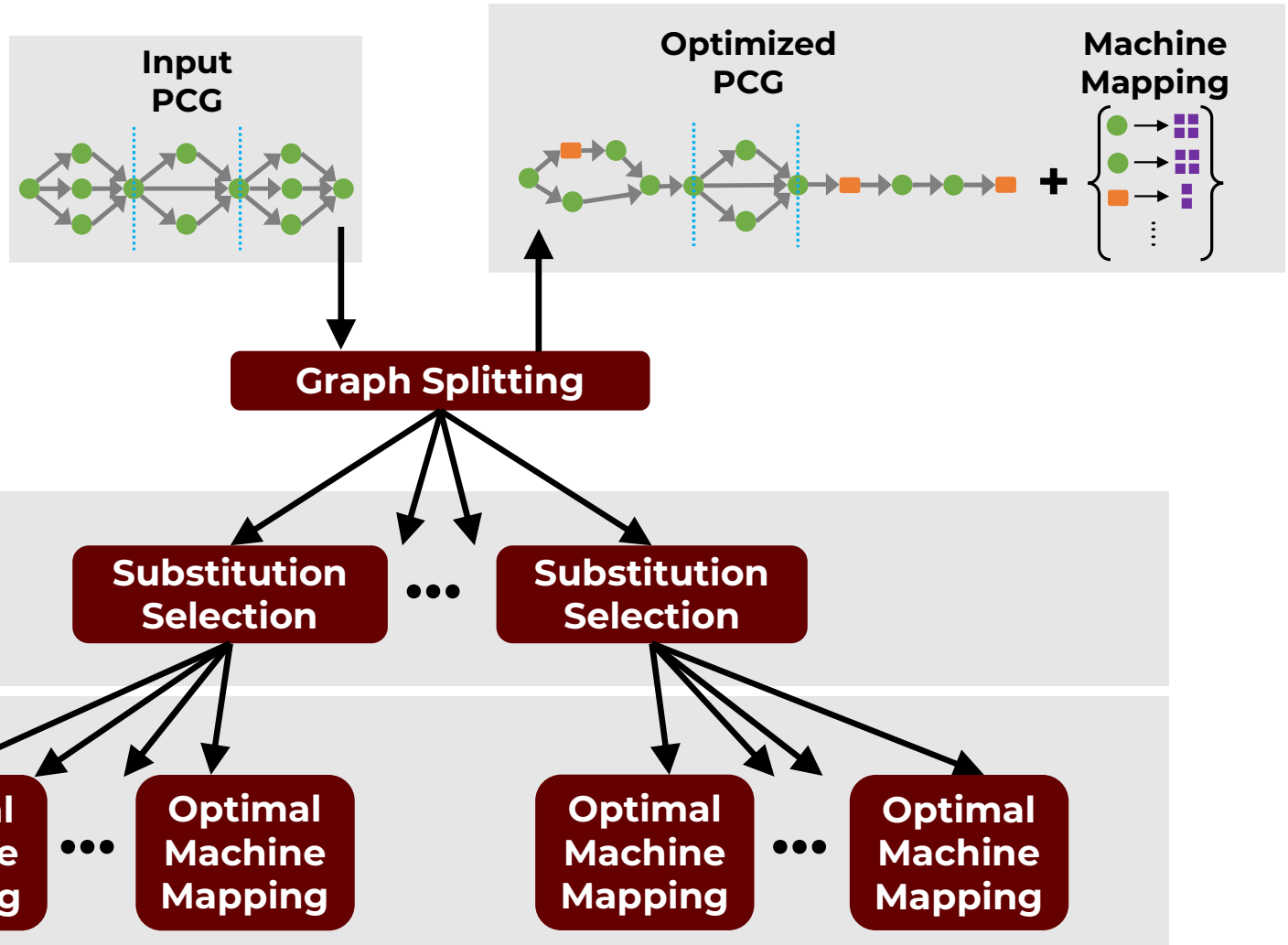
Hierarchical Search Algorithm

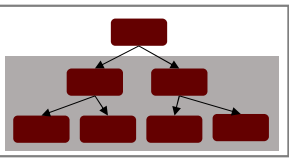


Hierarchical Search Algorithm

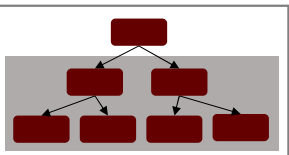


Hierarchical Search Algorithm



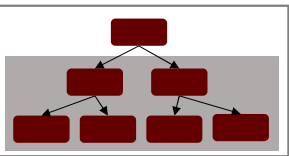


Hierarchical Search Algorithm



Hierarchical Search Algorithm

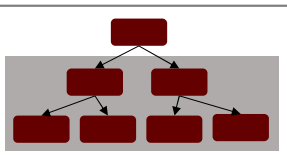
Algebraic Transformation



Hierarchical Search Algorithm

Algebraic Transformation

Parallelism Type

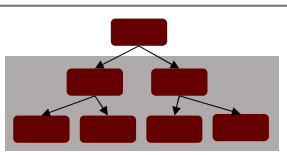


Hierarchical Search Algorithm

Algebraic Transformation

Parallelism Type

Parallelism Degree



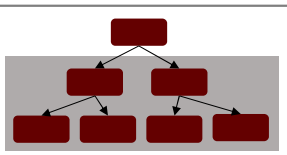
Hierarchical Search Algorithm

Algebraic Transformation

Parallelism Type

Parallelism Degree

Device Mapping



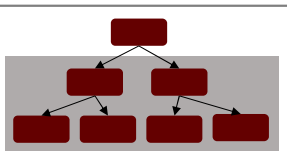
Hierarchical Search Algorithm

Algebraic Transformation

Parallelism Type

Parallelism Degree

Device Mapping



Hierarchical Search Algorithm

Algebraic Transformation

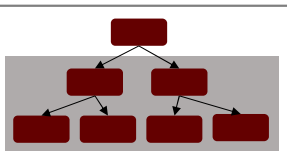
Backtracking
Search

Parallelism Type



Parallelism Degree

Device Mapping



Hierarchical Search Algorithm

Algebraic Transformation

Backtracking
Search

Parallelism Type



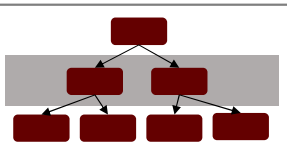
Parallelism Degree

Device Mapping



Dynamic
Programming





Hierarchical Search Algorithm

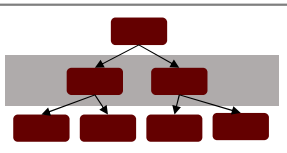
Algebraic Transformation

Parallelism Type

Parallelism Degree

Backtracking
Search

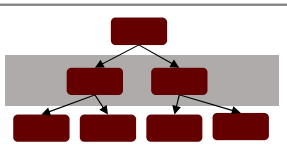




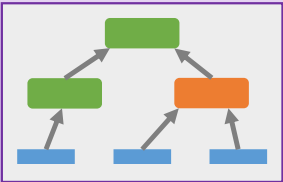
Hierarchical Search Algorithm

Backtracking
Search



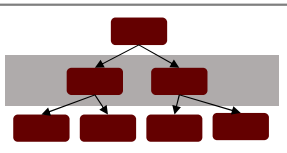


Hierarchical Search Algorithm

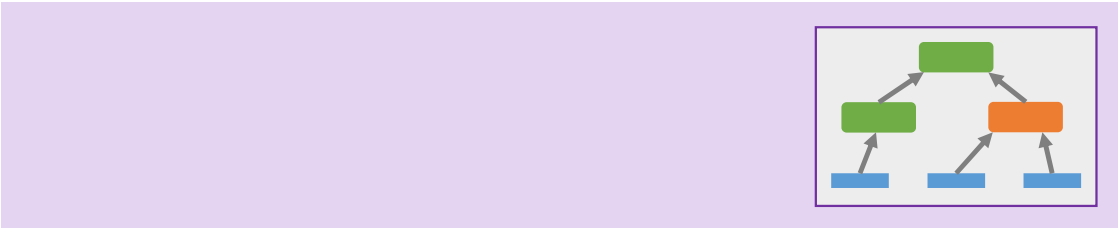


Backtracking Search





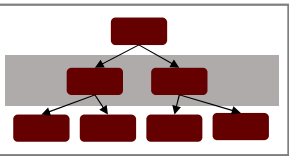
Hierarchical Search Algorithm



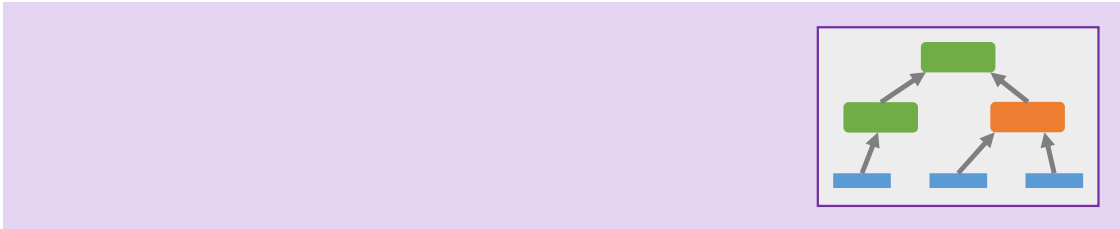
→
decreasing cost

Backtracking Search



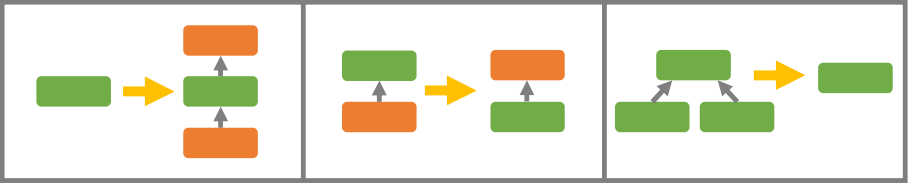


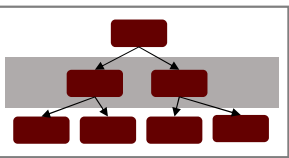
Hierarchical Search Algorithm



→
decreasing cost

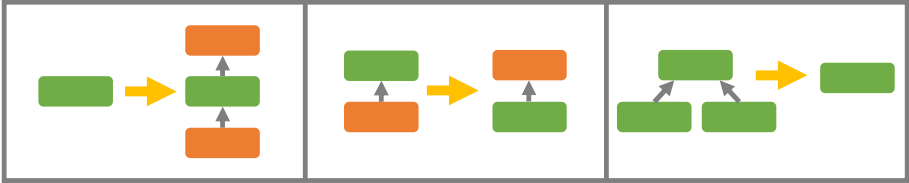
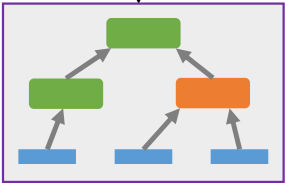
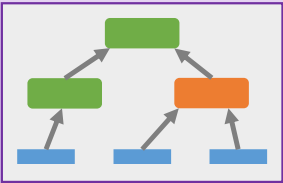
Backtracking Search

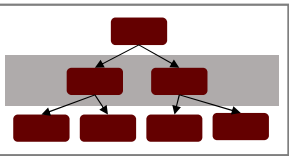




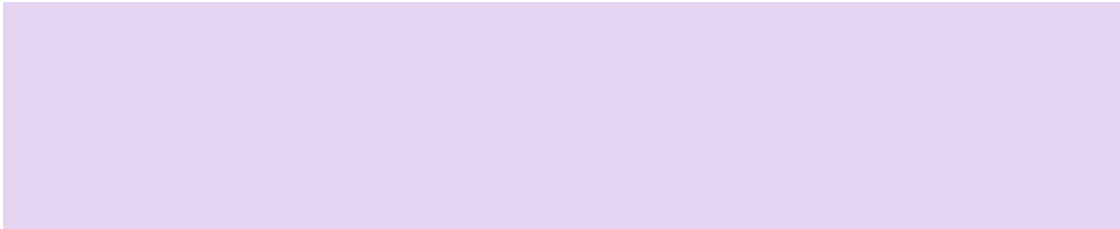
Hierarchical Search Algorithm

Backtracking Search

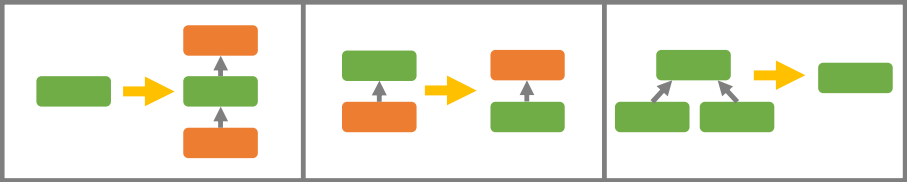
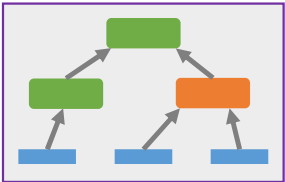


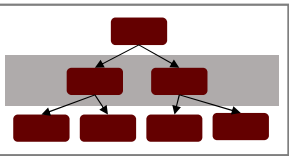


Hierarchical Search Algorithm

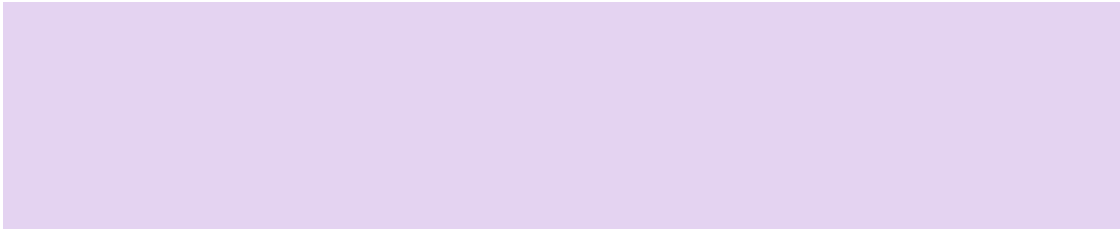


Backtracking Search

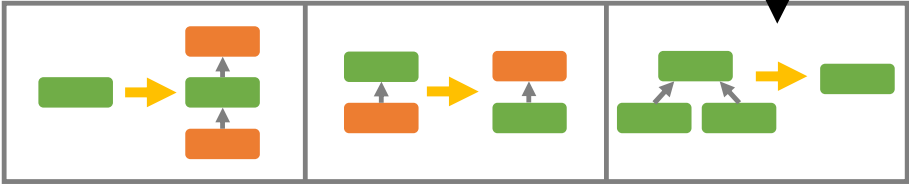
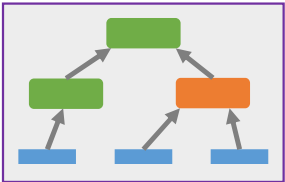


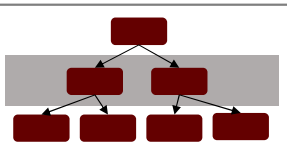


Hierarchical Search Algorithm

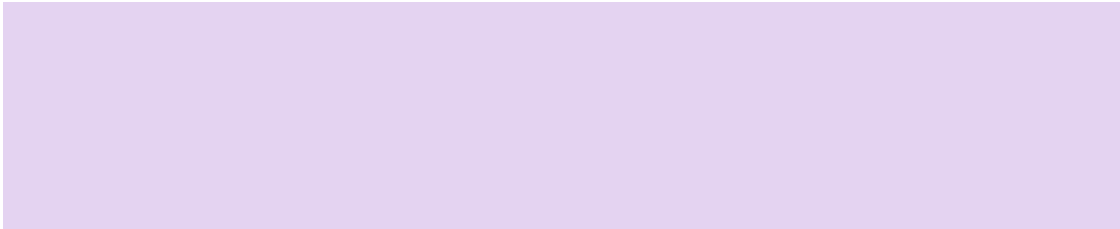


Backtracking Search

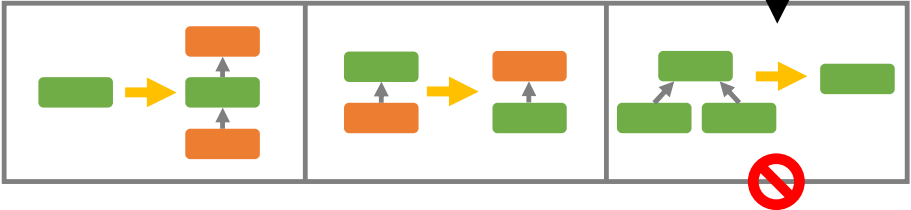
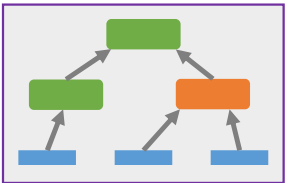


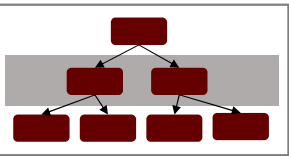


Hierarchical Search Algorithm

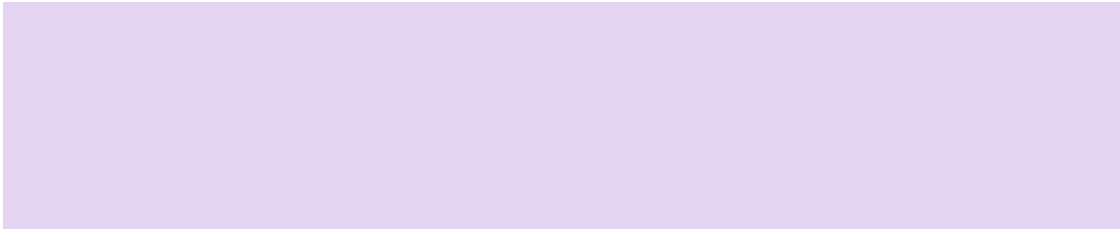


Backtracking Search

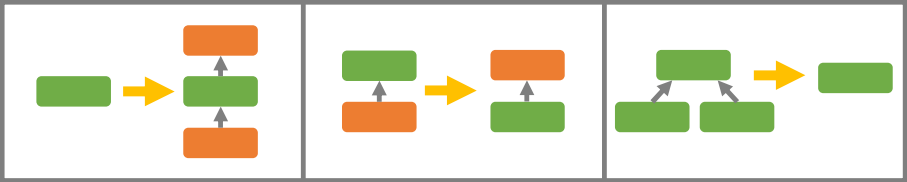
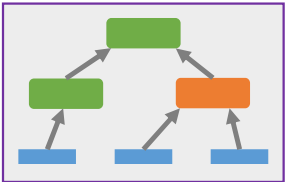


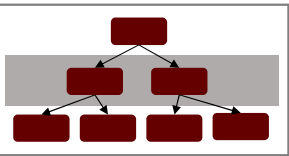


Hierarchical Search Algorithm

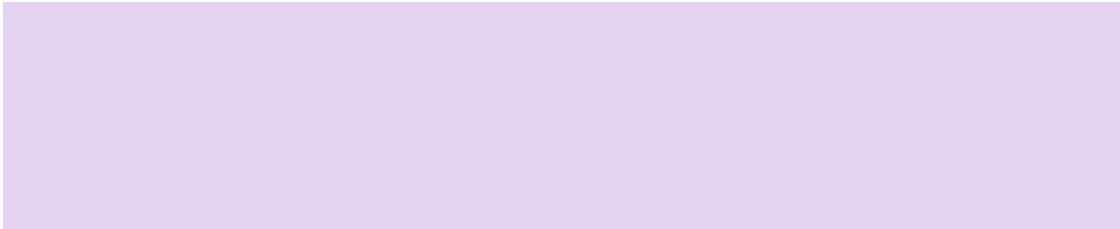


Backtracking Search

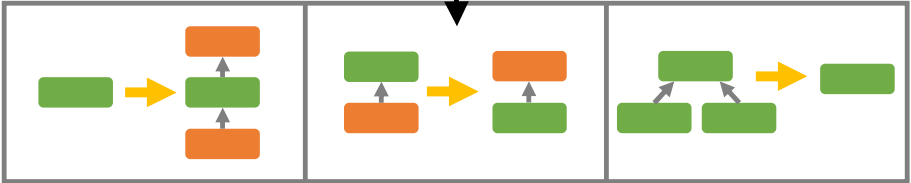
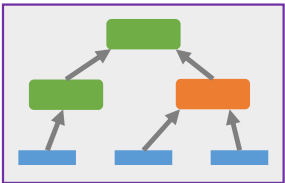


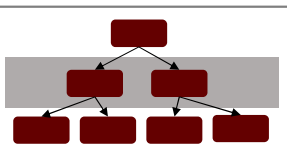


Hierarchical Search Algorithm

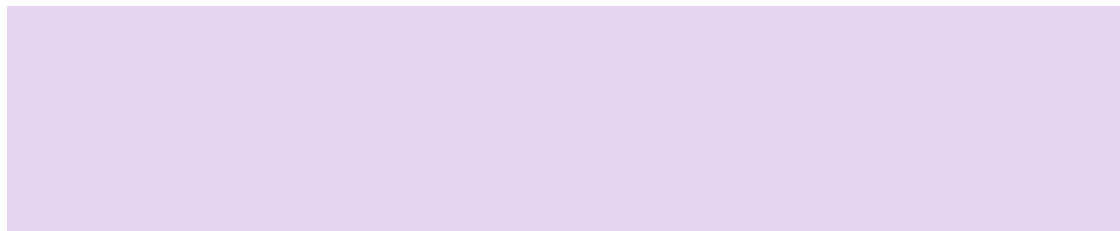


Backtracking Search

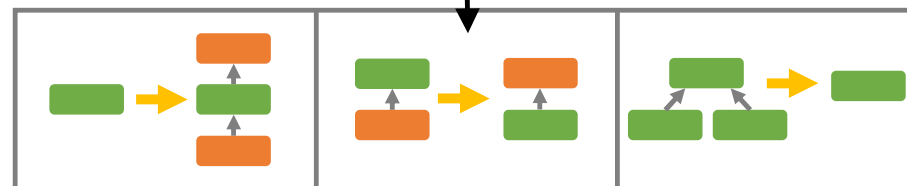
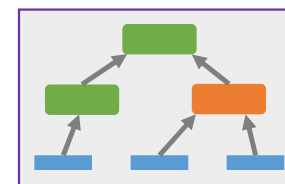




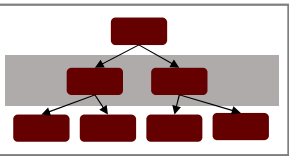
Hierarchical Search Algorithm



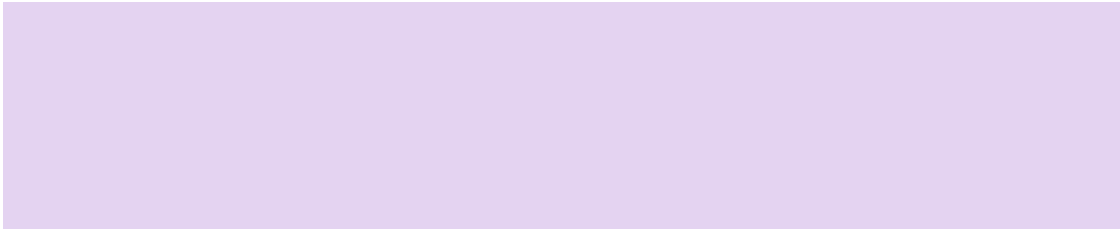
Backtracking Search



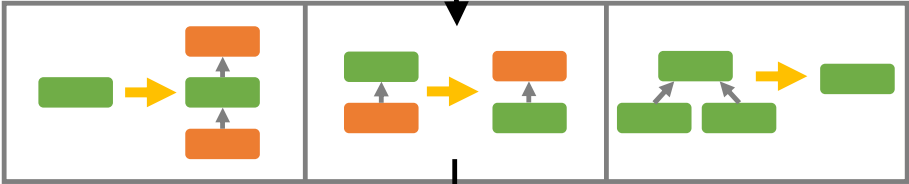
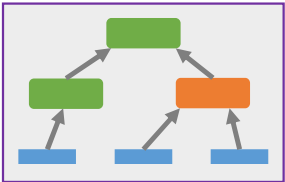
cost > threshold × best



Hierarchical Search Algorithm

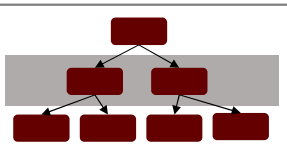


Backtracking Search



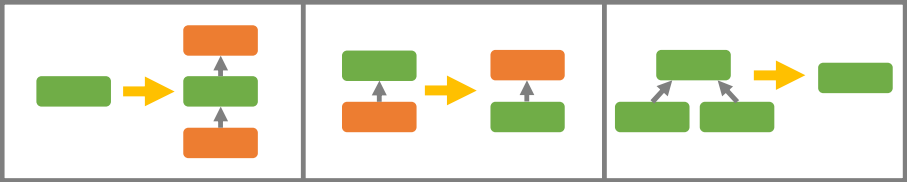
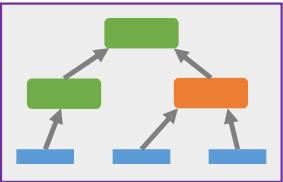
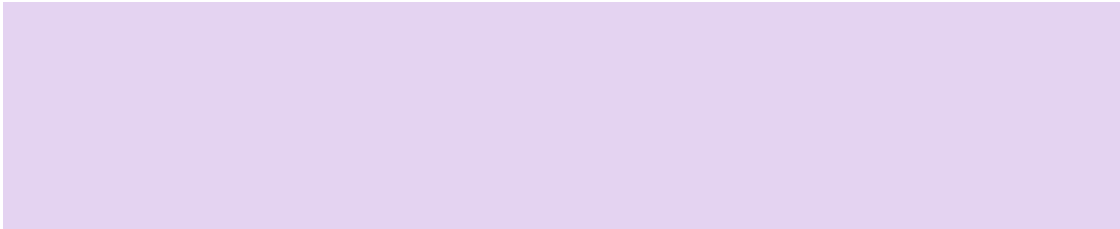
cost > threshold × best



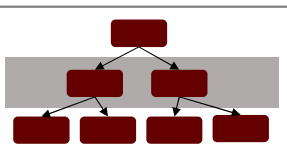


Hierarchical Search Algorithm

Backtracking Search

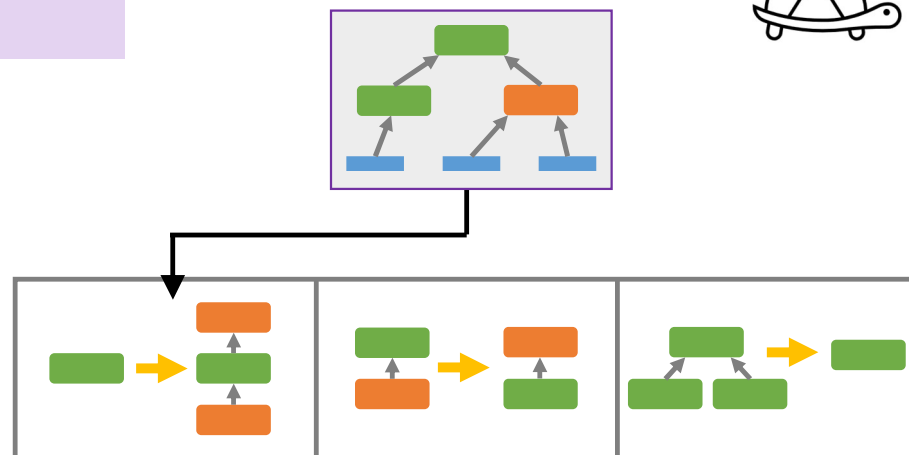
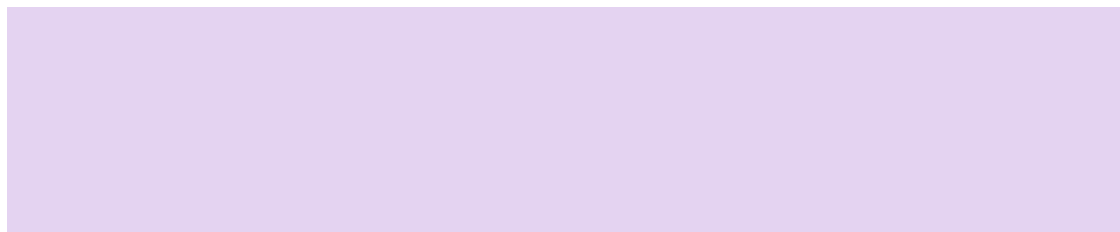


cost > threshold × best

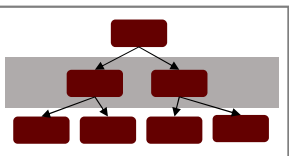


Hierarchical Search Algorithm

Backtracking Search

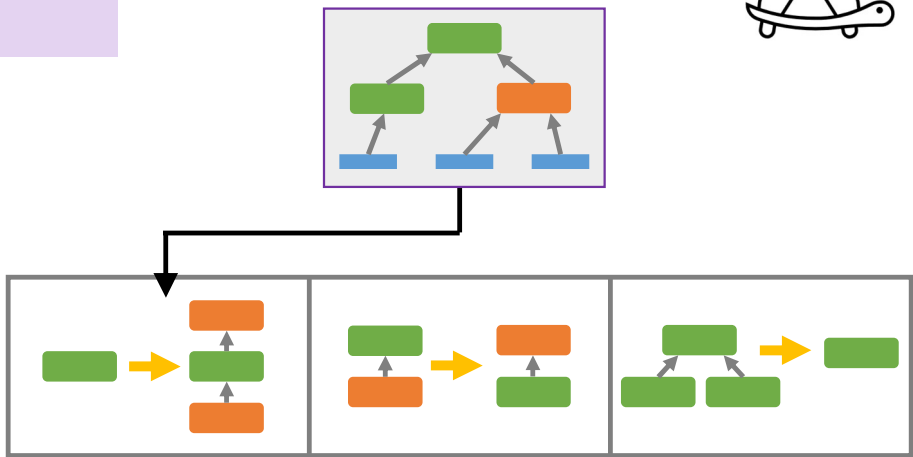
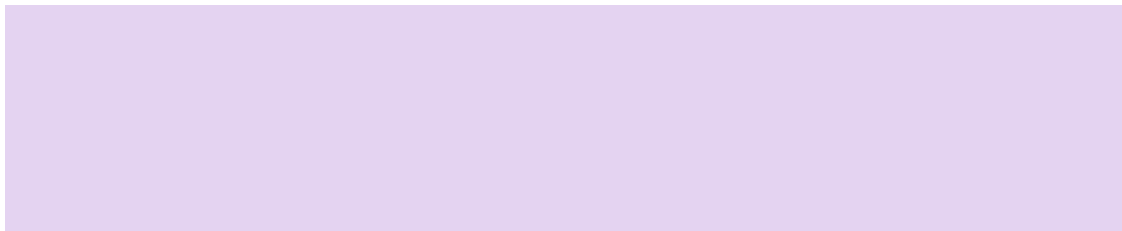


cost > threshold × best

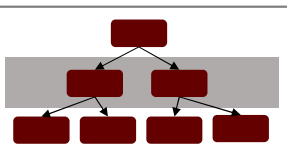


Hierarchical Search Algorithm

Backtracking Search

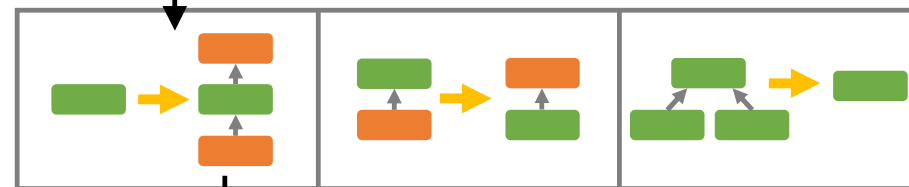
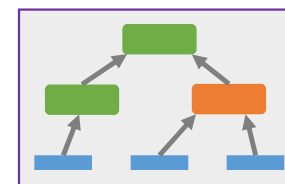
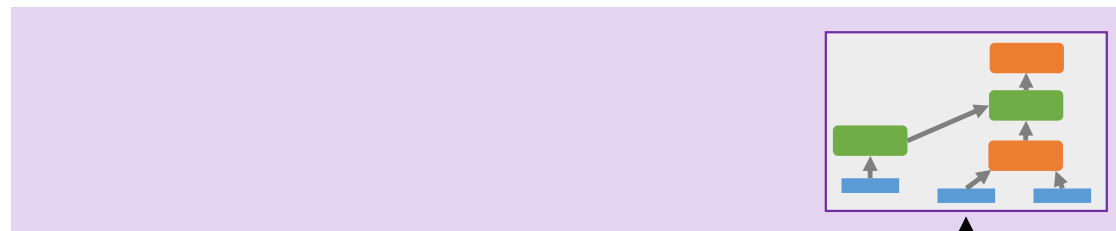


$$\text{cost} < \text{threshold} \times \text{best}$$

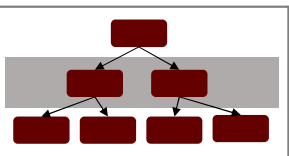


Hierarchical Search Algorithm

Backtracking Search

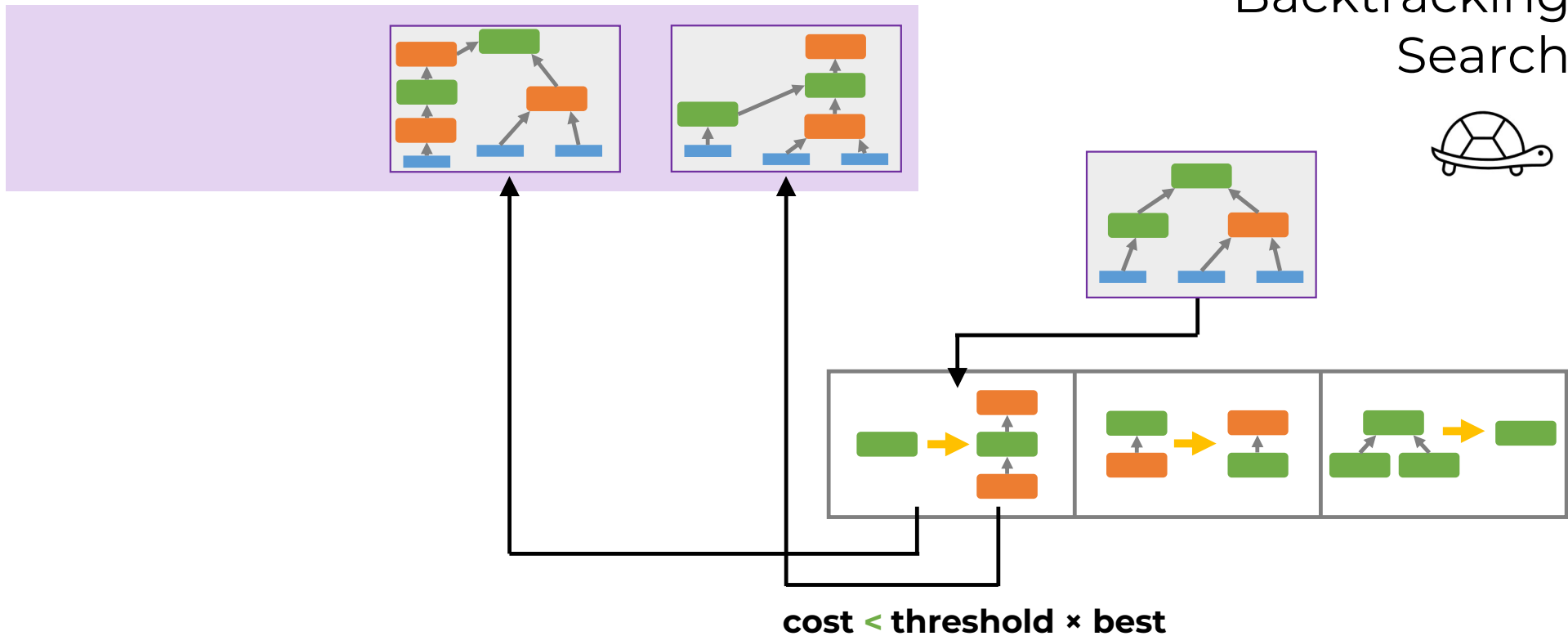


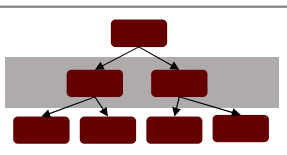
cost < **threshold** × **best**



Hierarchical Search Algorithm

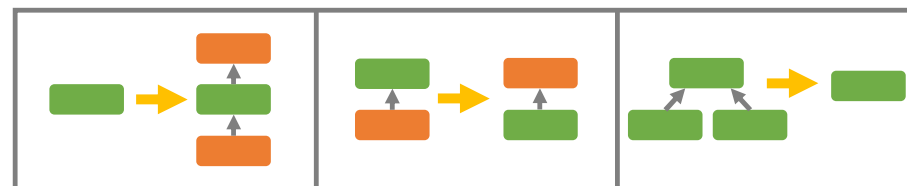
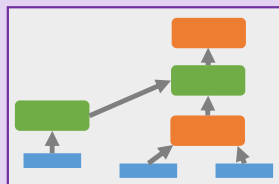
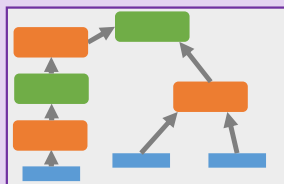
Backtracking Search



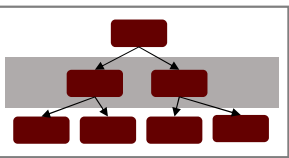


Hierarchical Search Algorithm

Backtracking Search

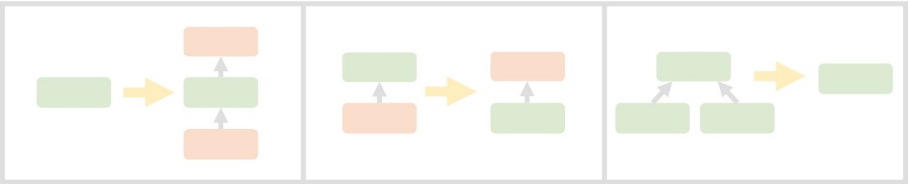
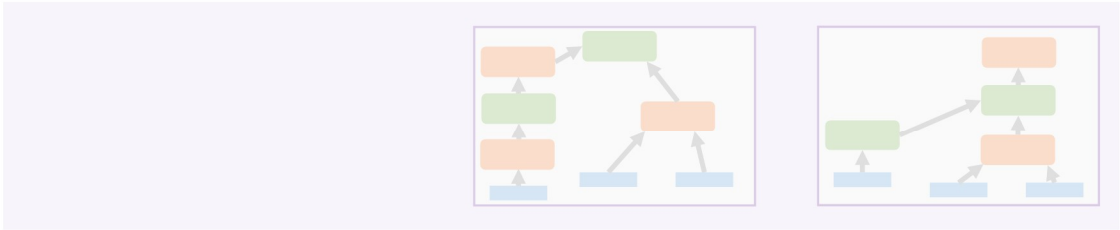


cost < threshold × best

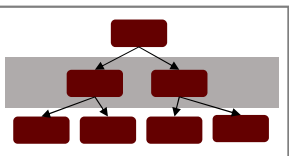


Hierarchical Search Algorithm

Backtracking Search

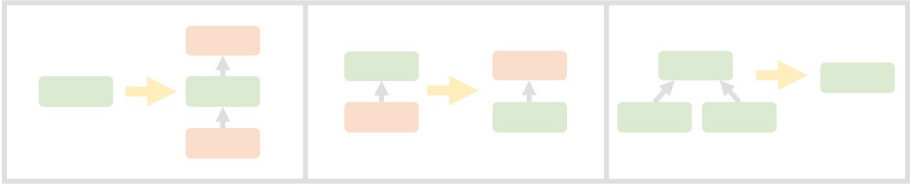
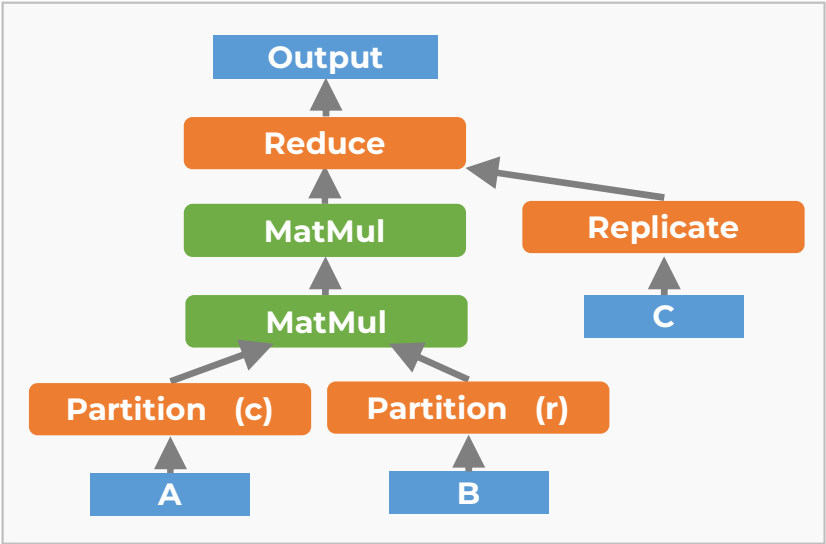
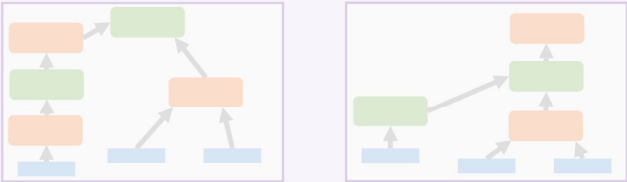


$$\text{cost} < \text{threshold} \times \text{best}$$

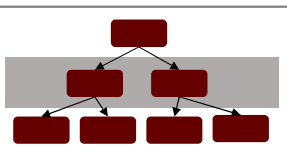


Hierarchical Search Algorithm

Backtracking Search

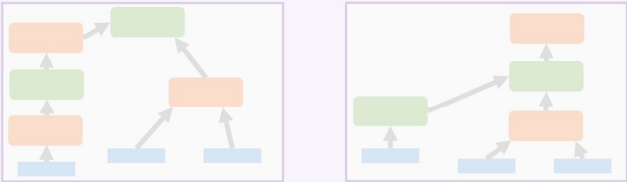


$$\text{cost} < \text{threshold} \times \text{best}$$

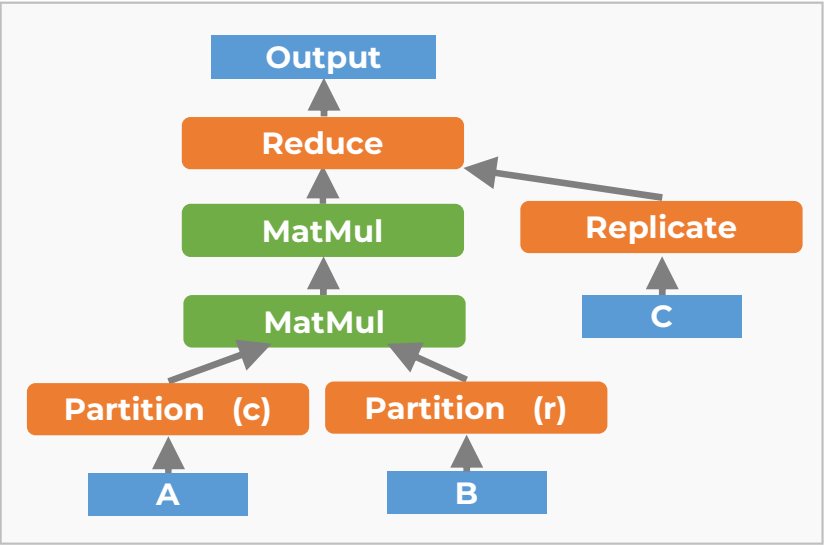


Hierarchical Search Algorithm

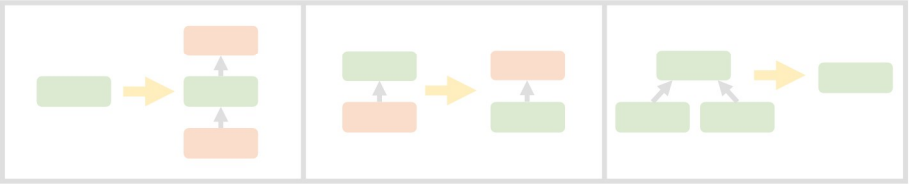
Backtracking Search



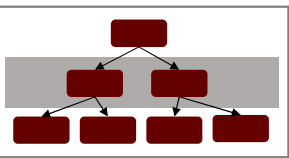
cost (



)

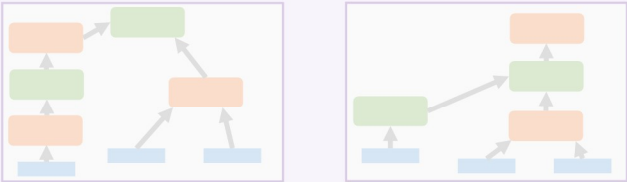


$$\text{cost} < \text{threshold} \times \text{best}$$

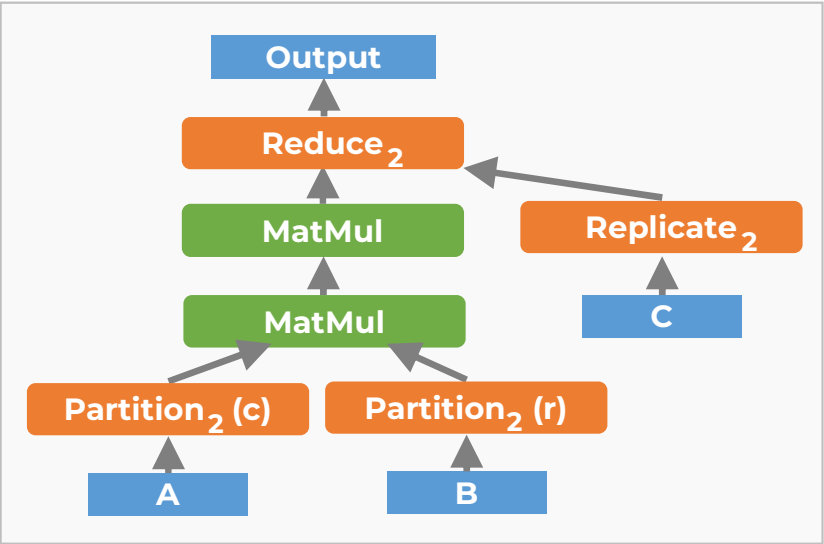


Hierarchical Search Algorithm

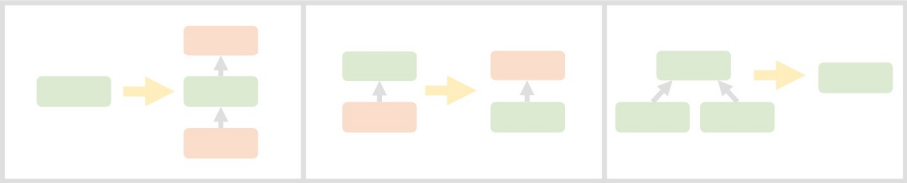
Backtracking Search



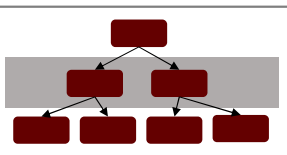
cost (



)



cost < threshold * best

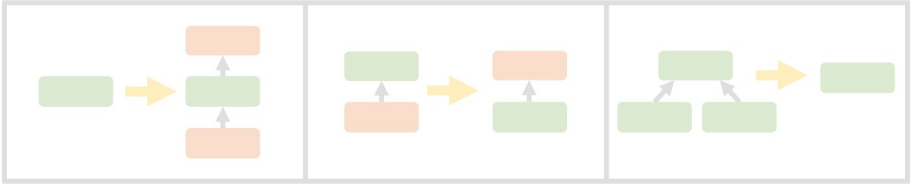
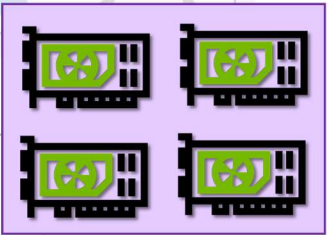
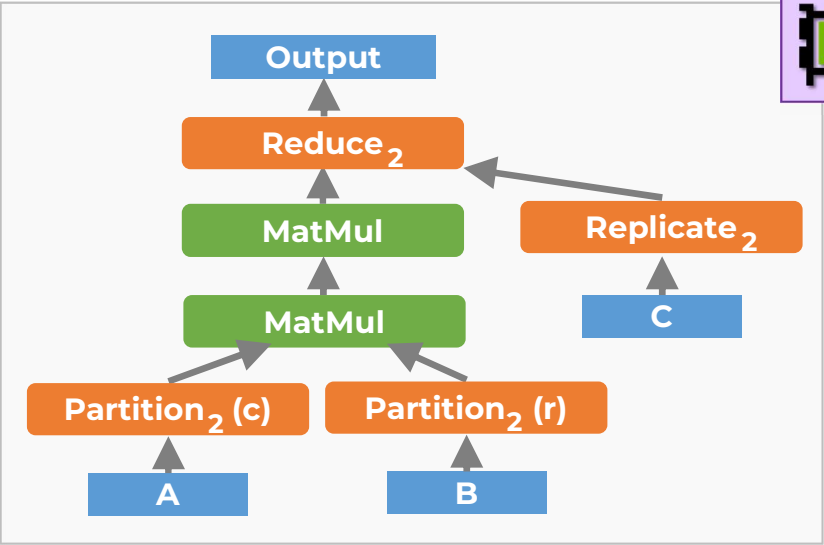


Hierarchical Search Algorithm

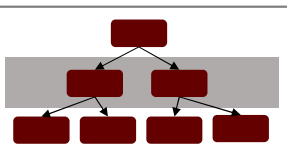
Backtracking Search



cost (

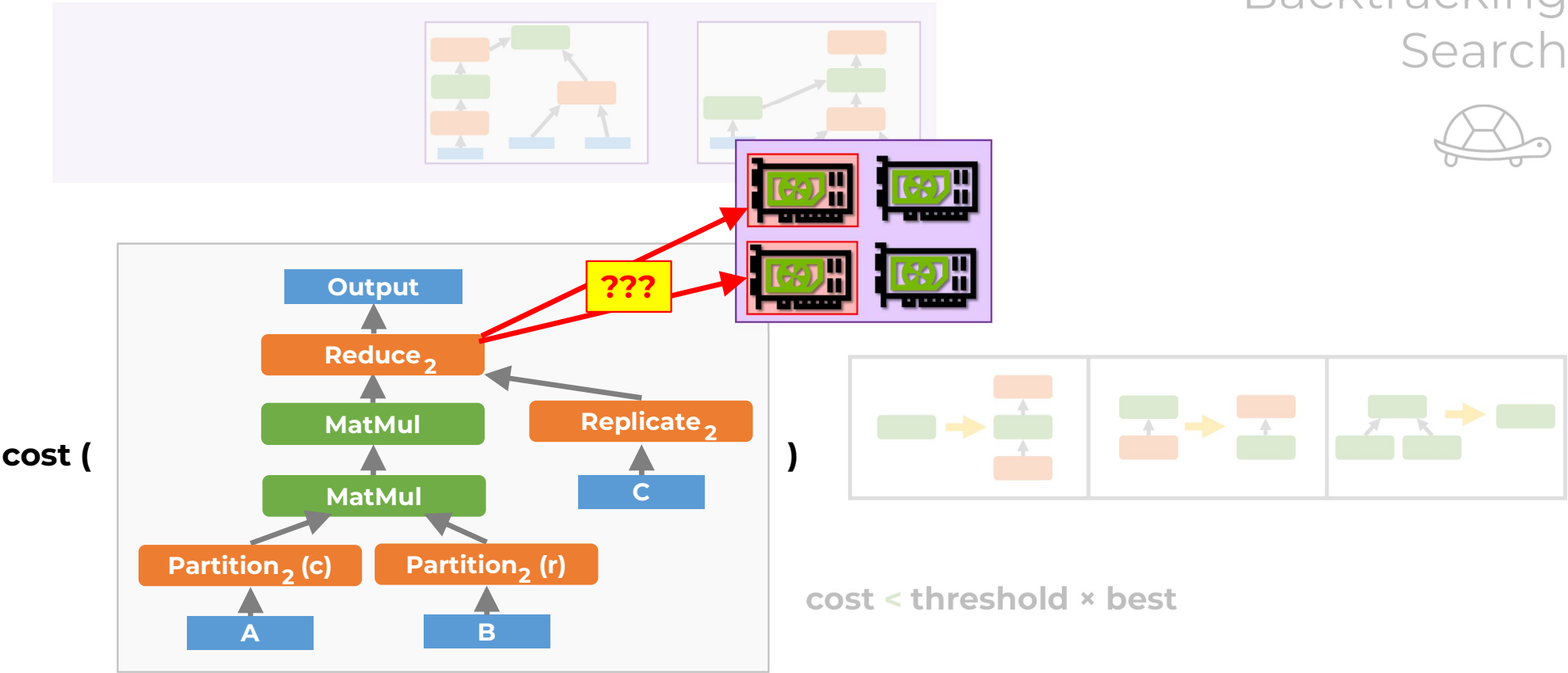


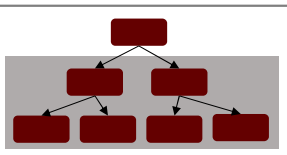
cost < threshold * best



Hierarchical Search Algorithm

Backtracking Search





Hierarchical Search Algorithm

Algebraic Transformation

Backtracking
Search

Parallelism Type



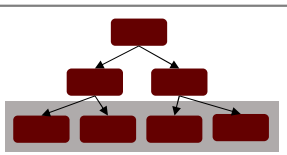
Parallelism Degree

Device Mapping



Dynamic
Programming





Hierarchical Search Algorithm



Dynamic
Programming

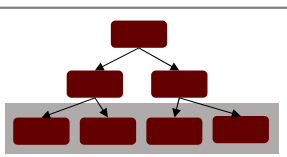
2

2

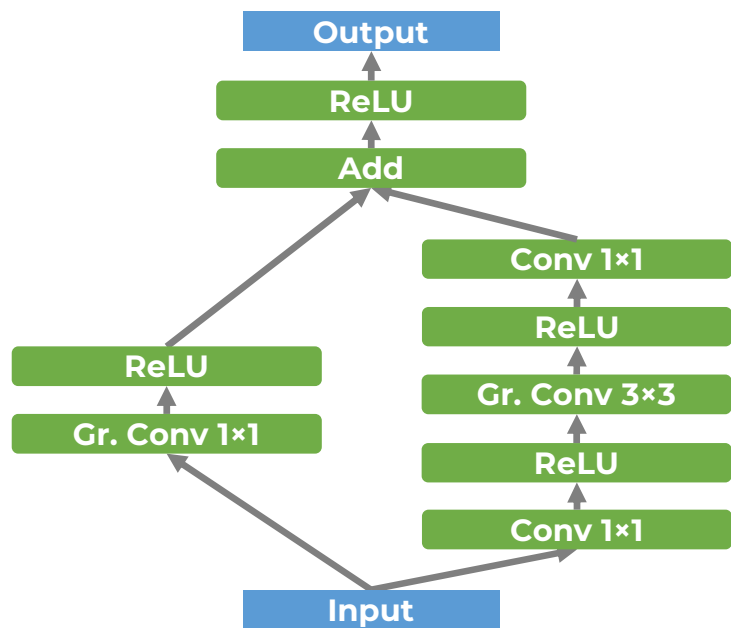
/

3

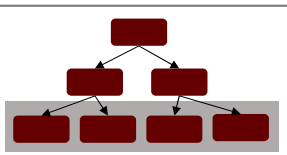
8



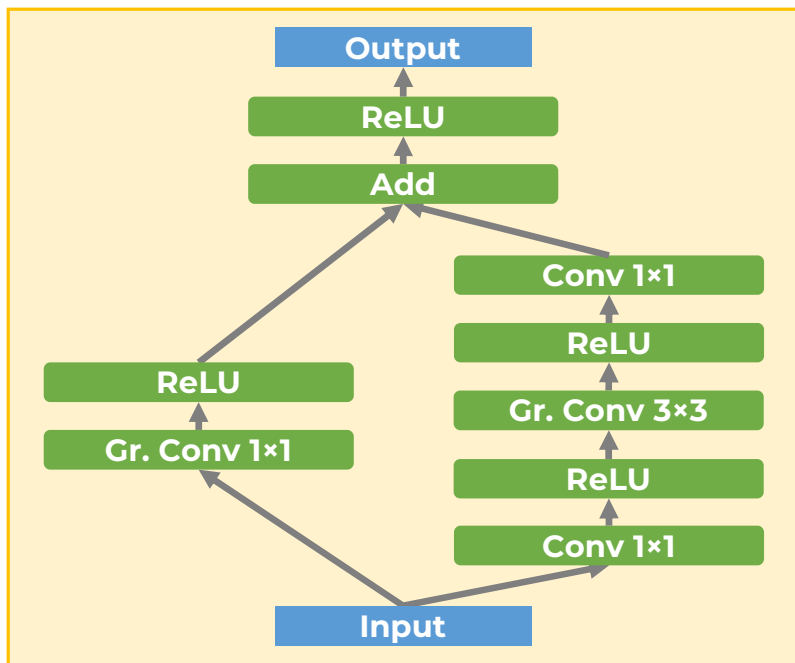
Hierarchical Search Algorithm

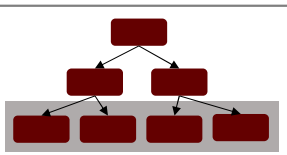


Dynamic
Programming

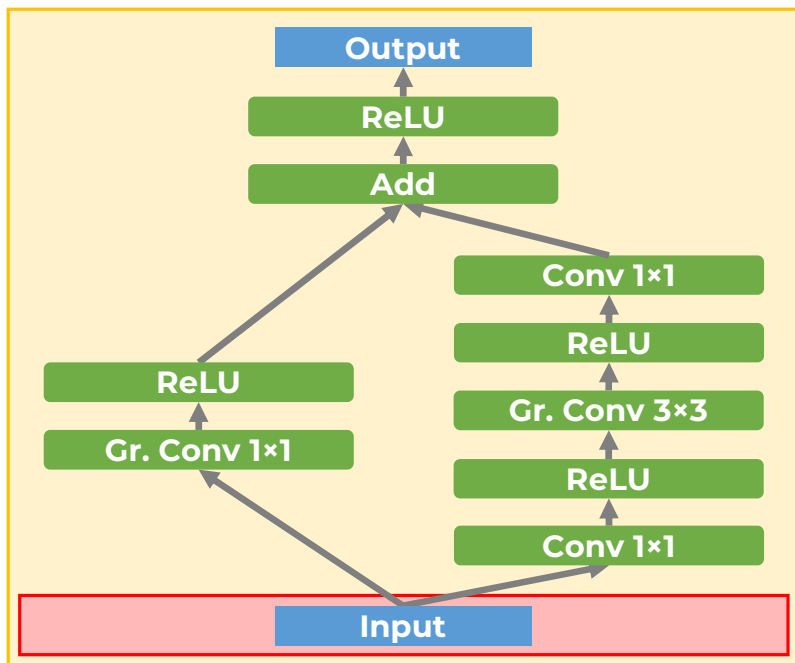


Hierarchical Search Algorithm

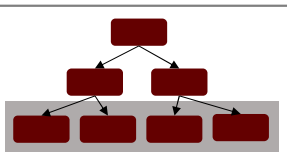




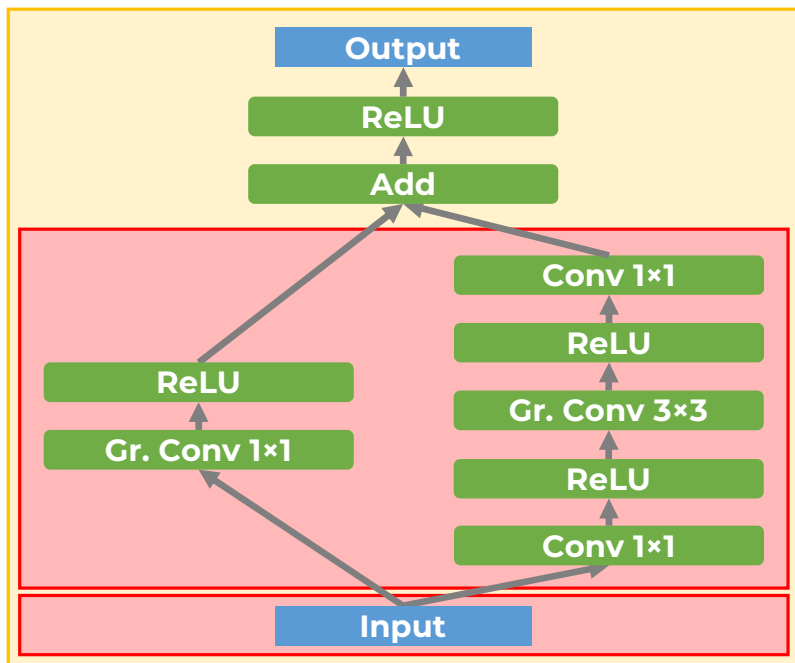
Hierarchical Search Algorithm



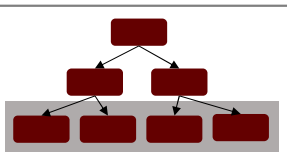
Dynamic
Programming



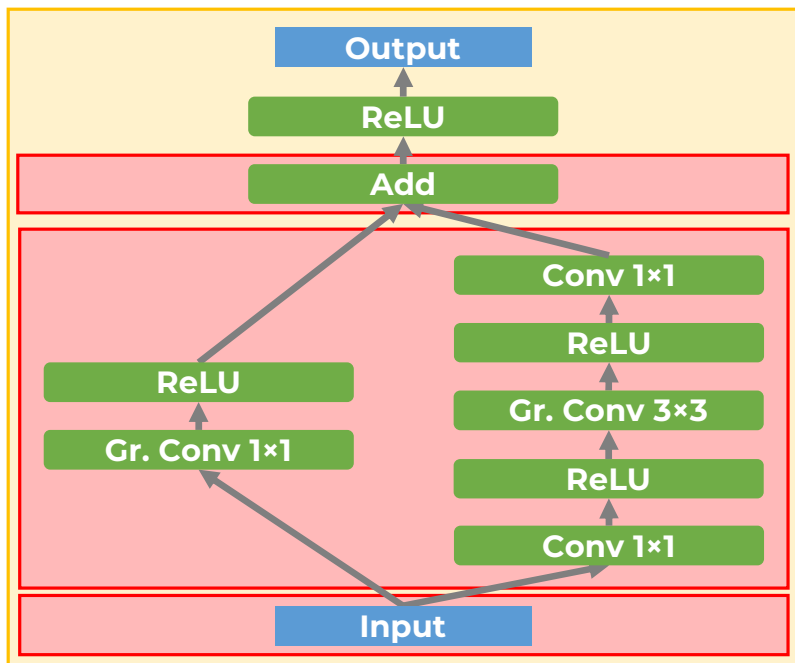
Hierarchical Search Algorithm



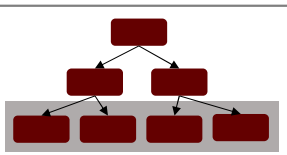
Dynamic
Programming



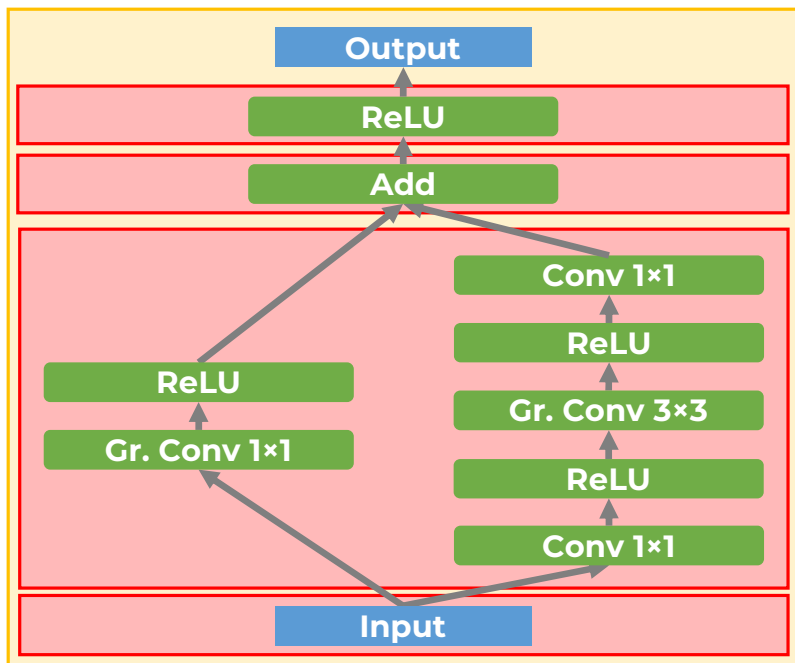
Hierarchical Search Algorithm



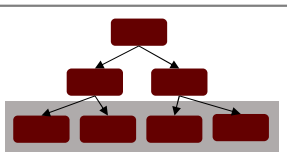
Dynamic
Programming



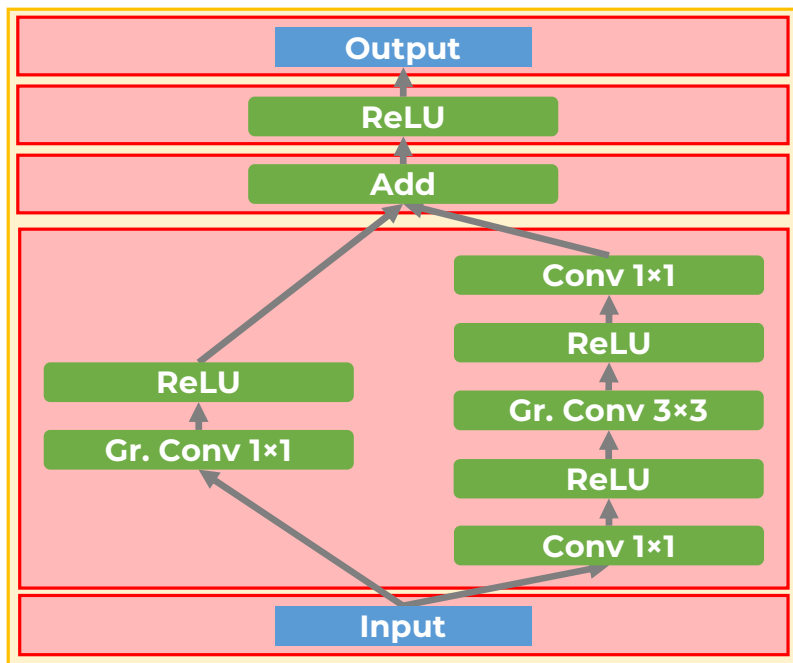
Hierarchical Search Algorithm



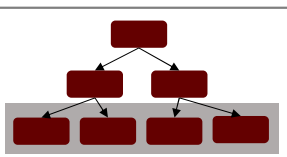
Dynamic
Programming



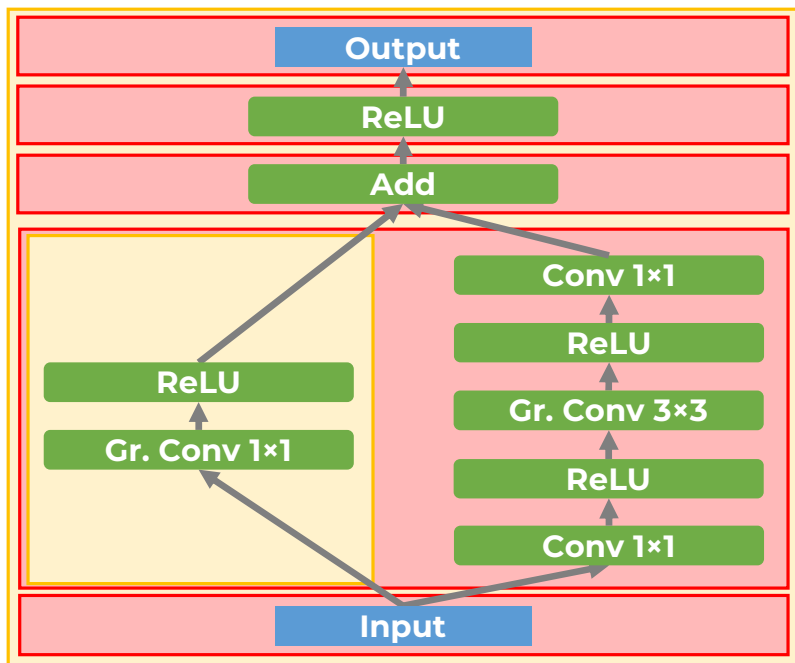
Hierarchical Search Algorithm



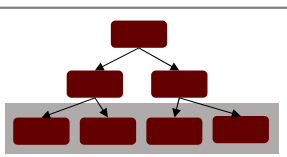
Dynamic
Programming



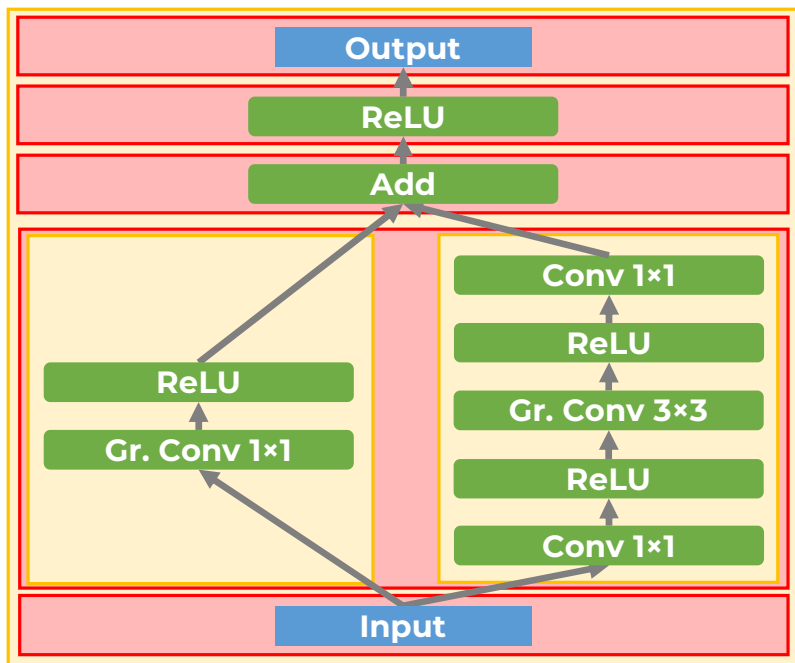
Hierarchical Search Algorithm



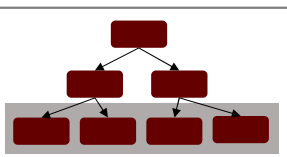
Dynamic
Programming



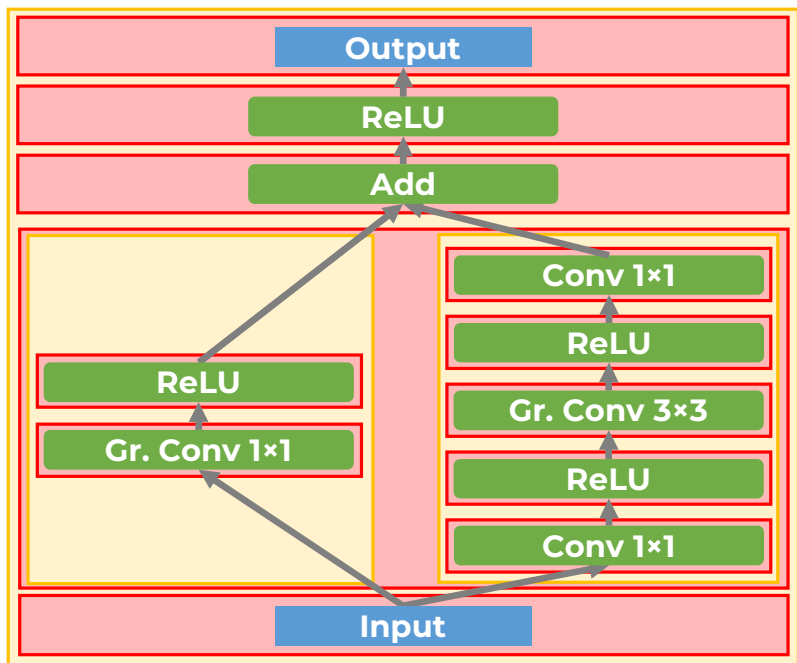
Hierarchical Search Algorithm



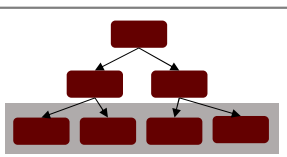
Dynamic
Programming



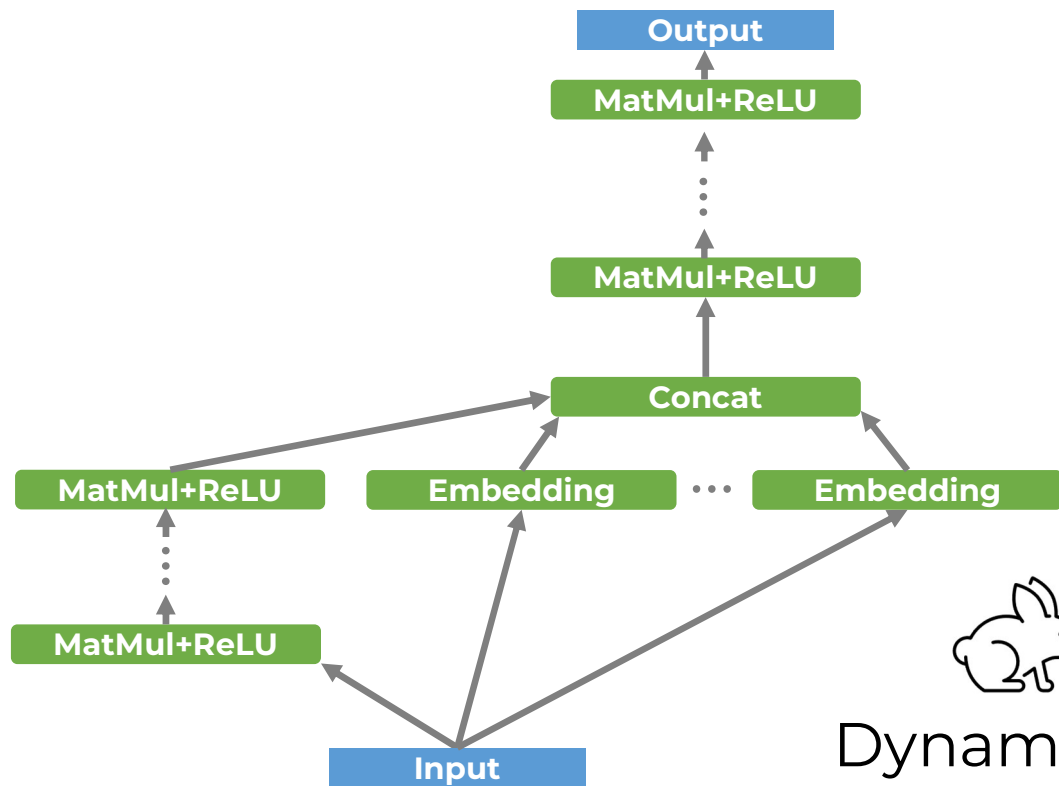
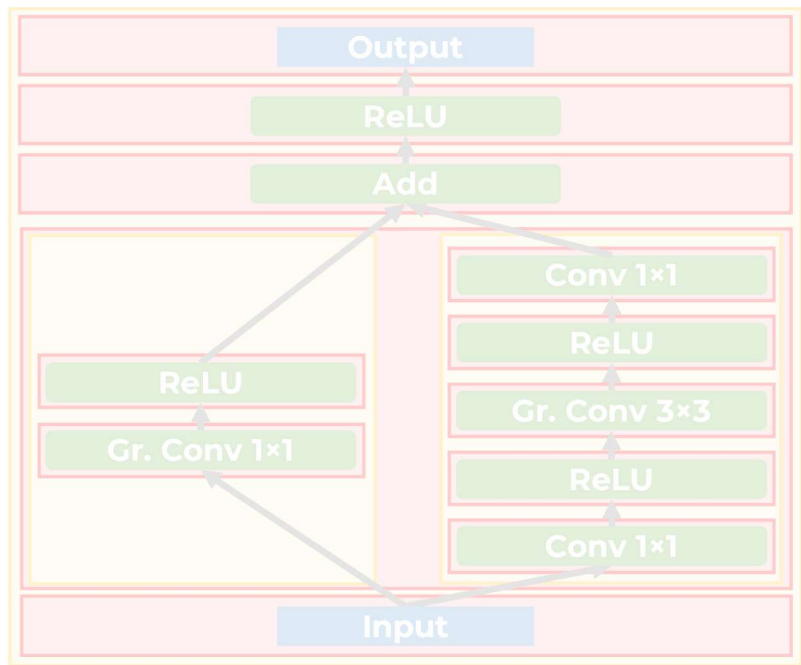
Hierarchical Search Algorithm



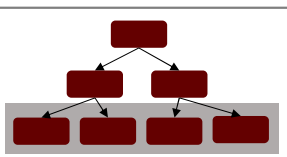
Dynamic
Programming



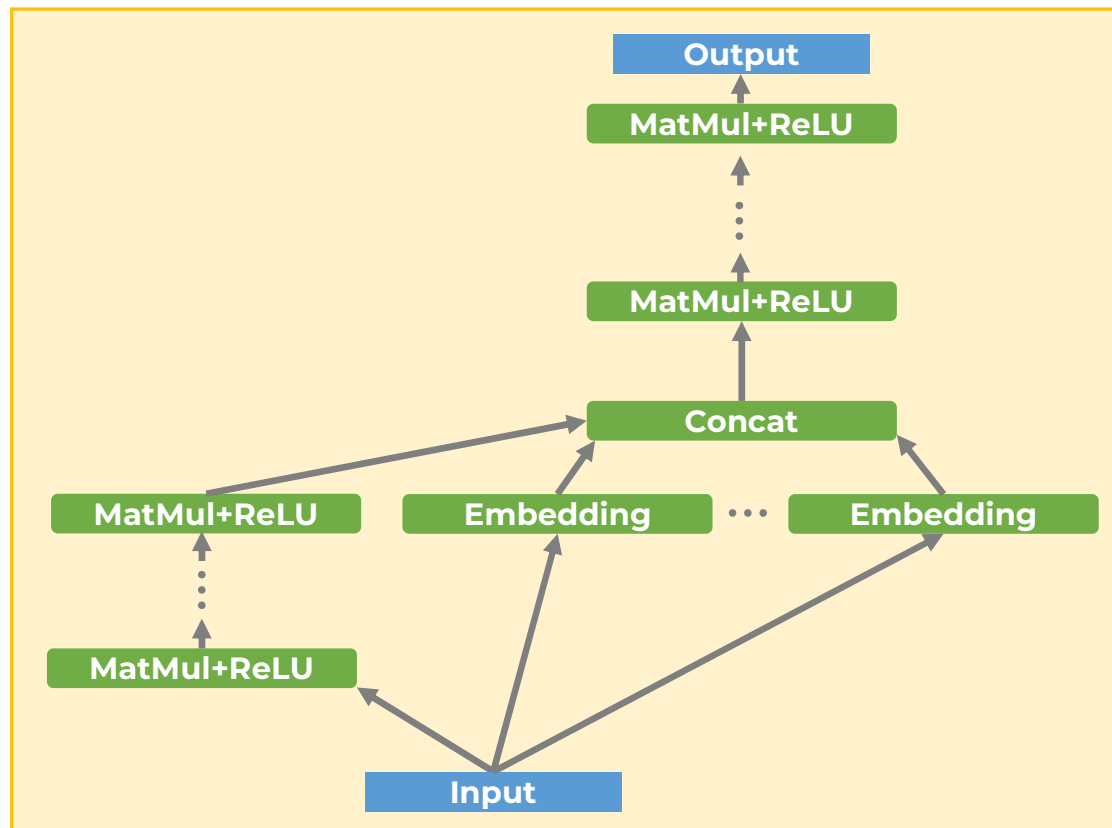
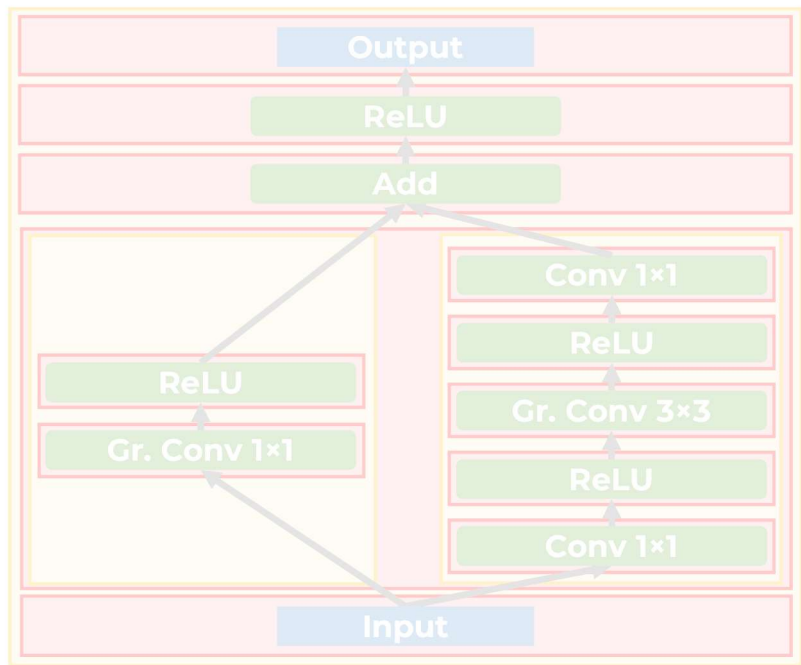
Hierarchical Search Algorithm



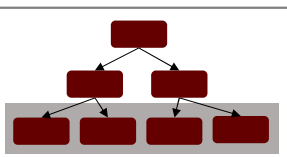
Dynamic
Programming



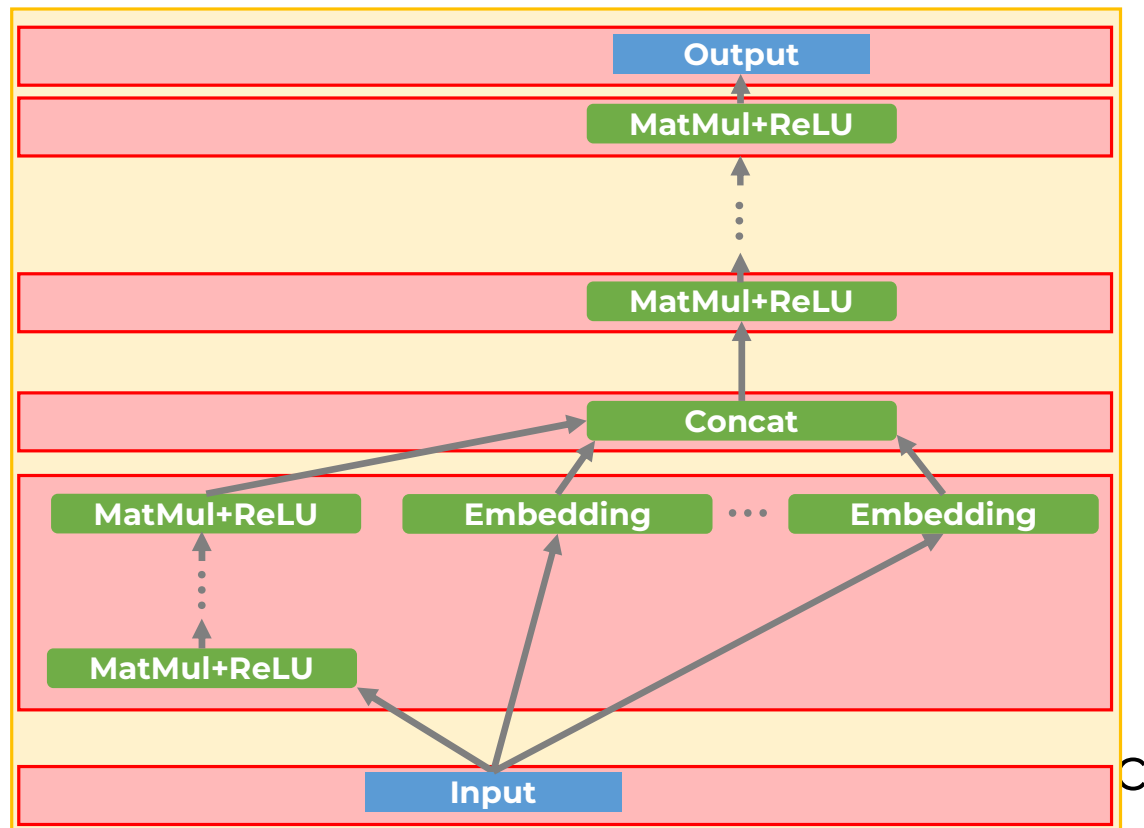
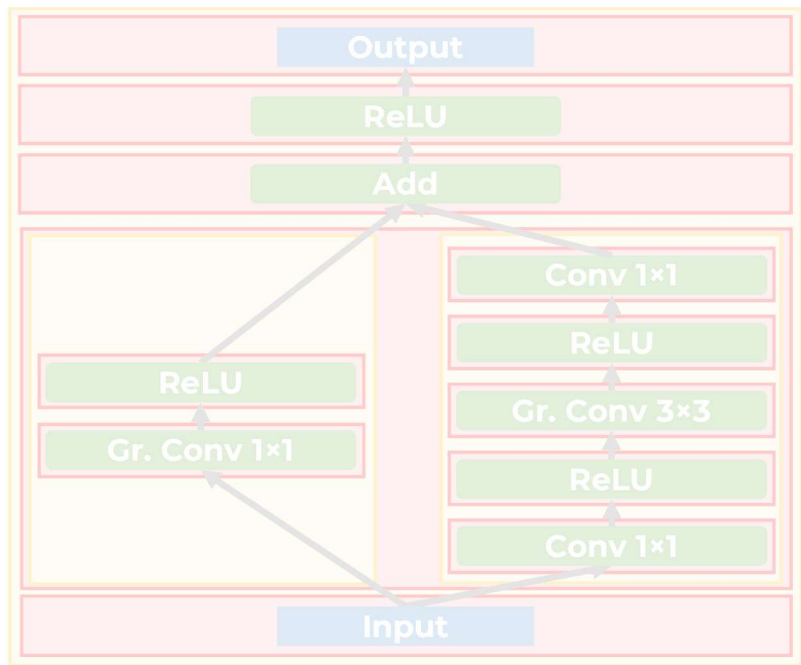
Hierarchical Search Algorithm



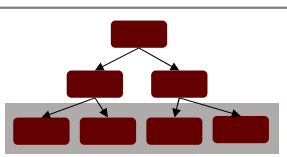
Programming



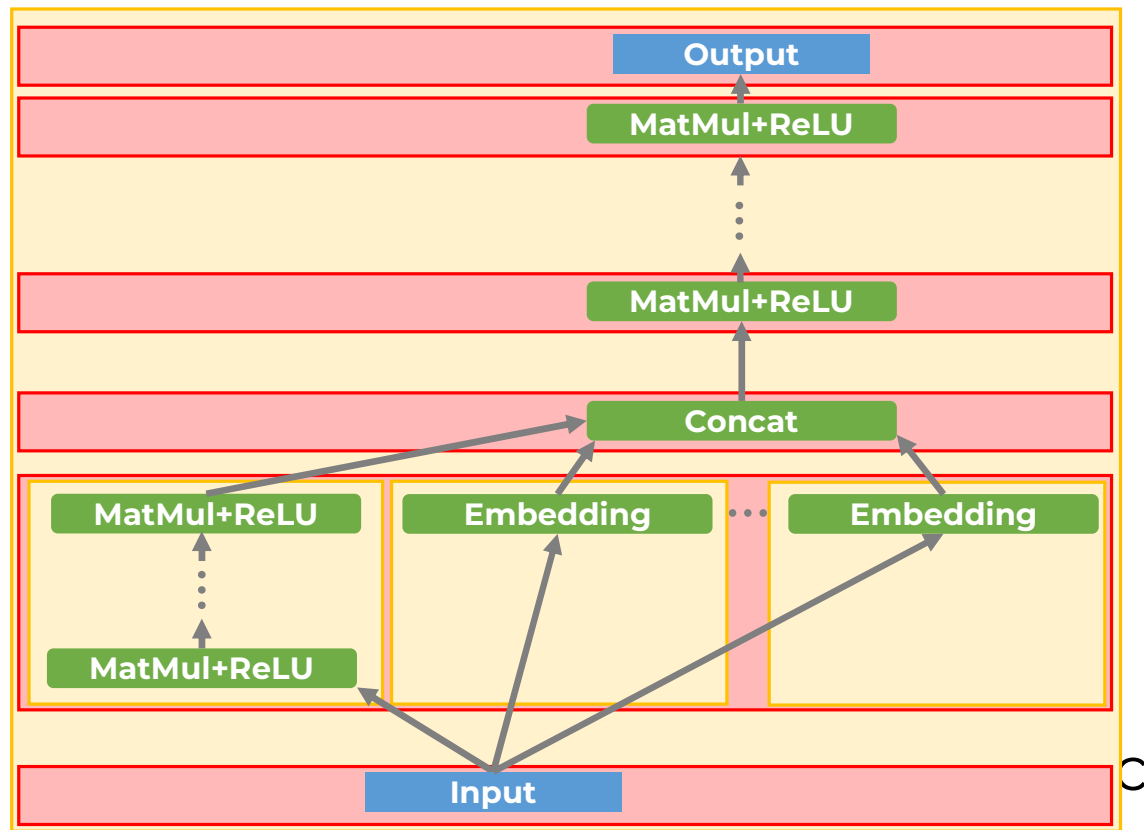
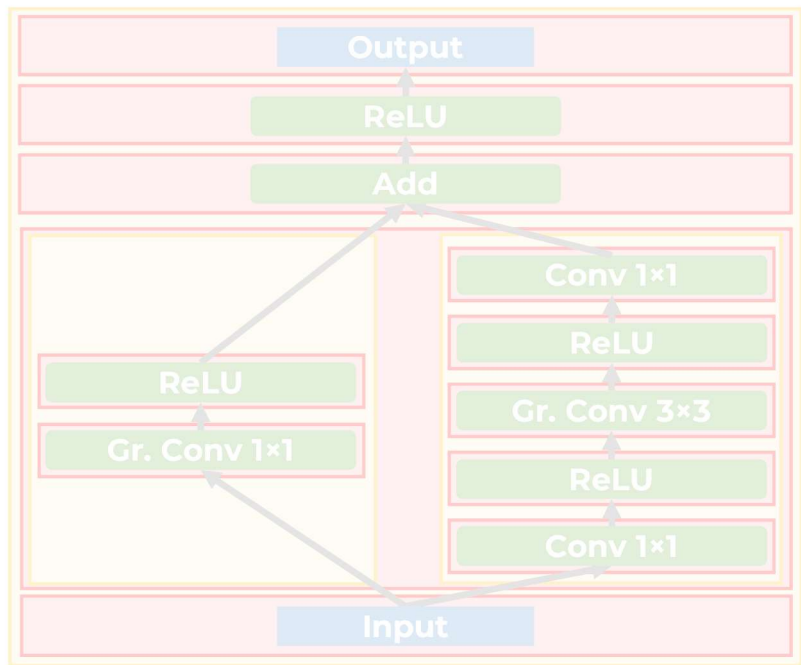
Hierarchical Search Algorithm



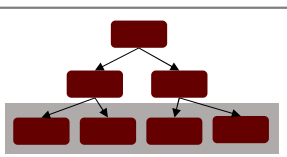
Programming



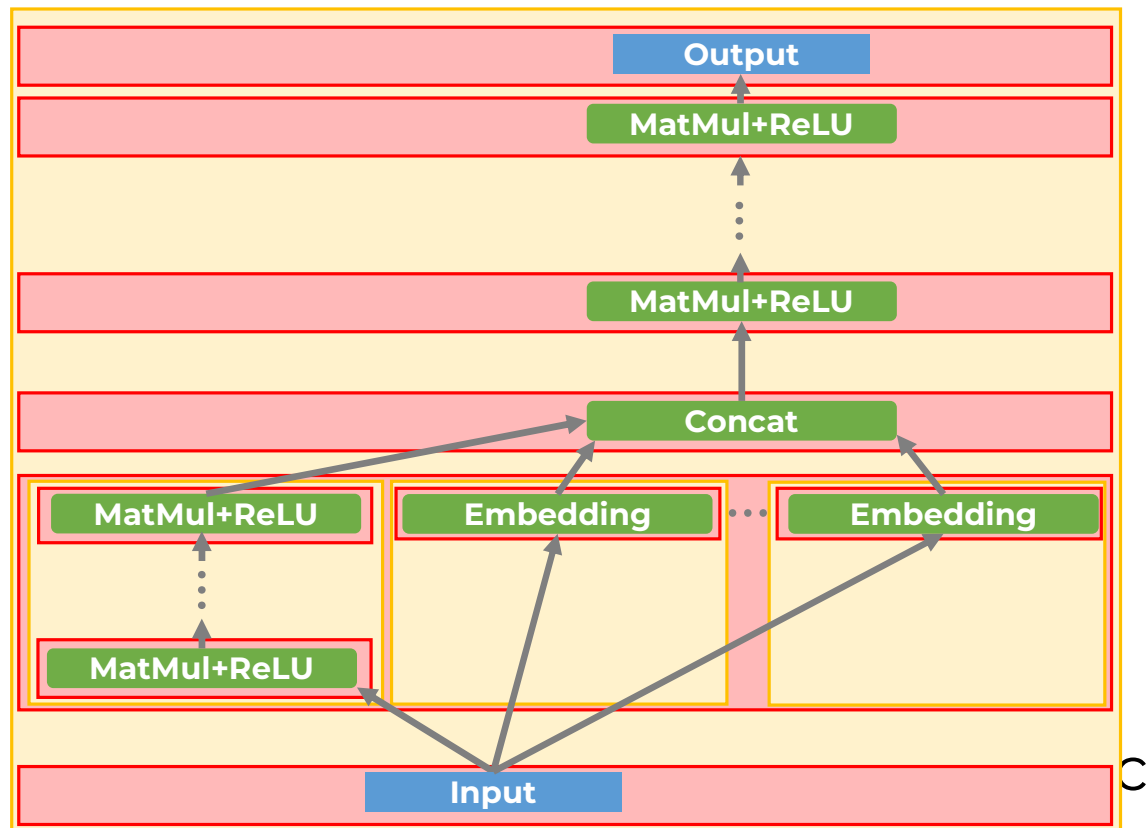
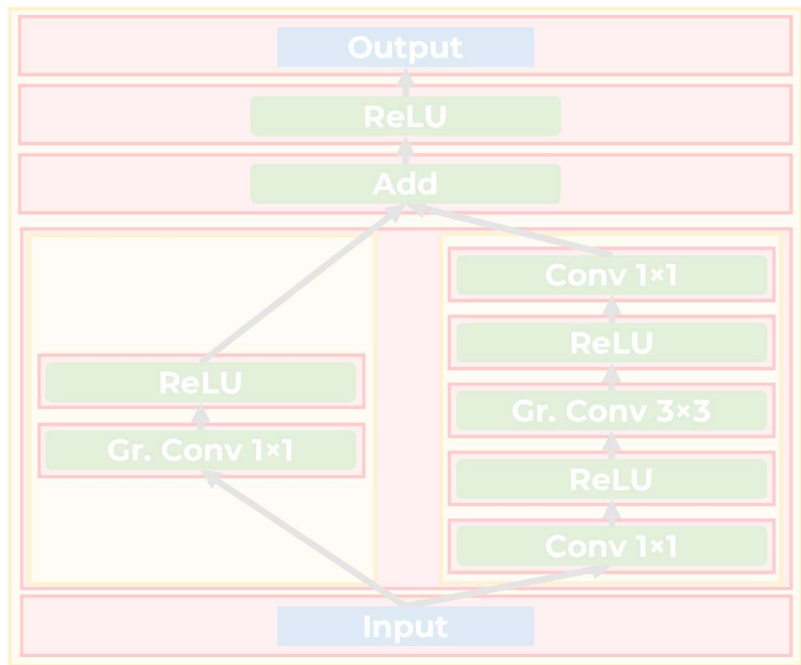
Hierarchical Search Algorithm



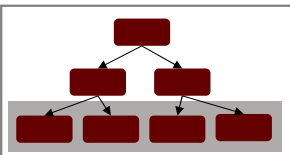
Programming



Hierarchical Search Algorithm



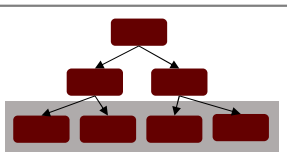
Programming



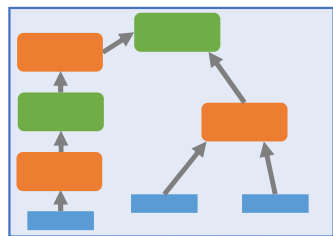
Hierarchical Search Algorithm



Dynamic
Programming



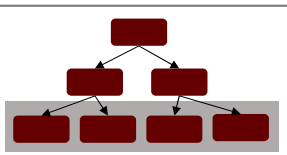
Hierarchical Search Algorithm



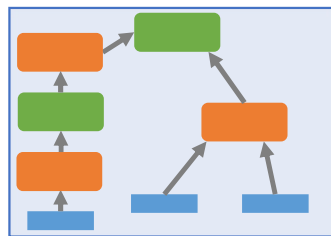
G



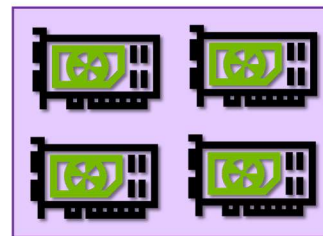
Dynamic
Programming



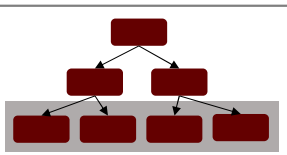
Hierarchical Search Algorithm



G D



Dynamic
Programming



Hierarchical Search Algorithm

$\text{OptCost}(G, D)$



Dynamic
Programming

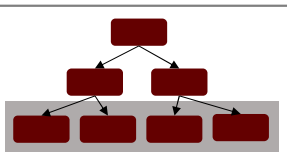
2

3

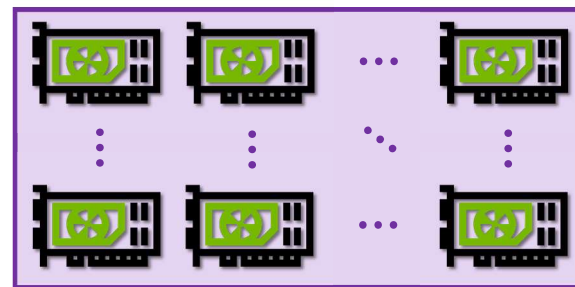
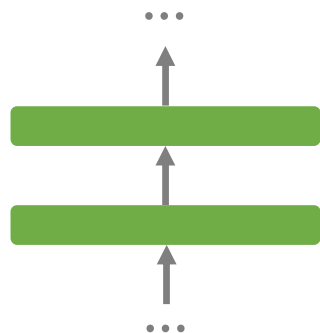
/

3

8



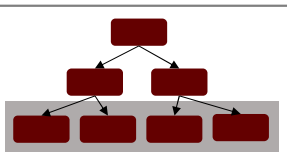
Hierarchical Search Algorithm



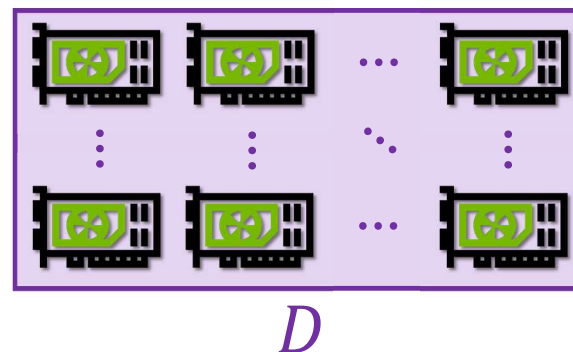
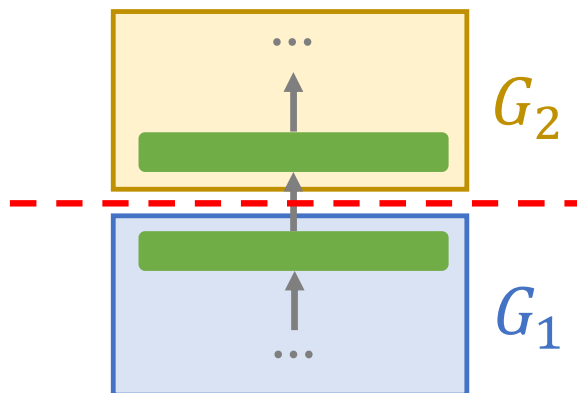
D

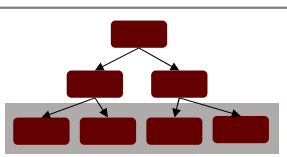


Dynamic
Programming

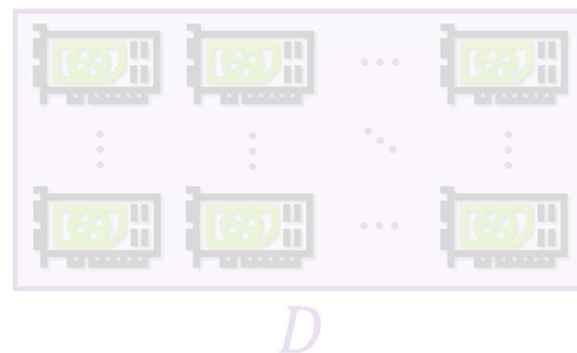
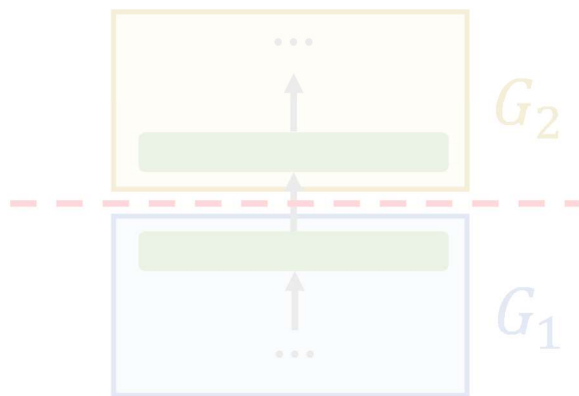


Hierarchical Search Algorithm



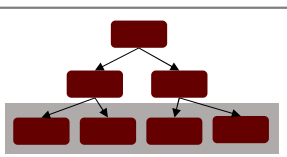


Hierarchical Search Algorithm

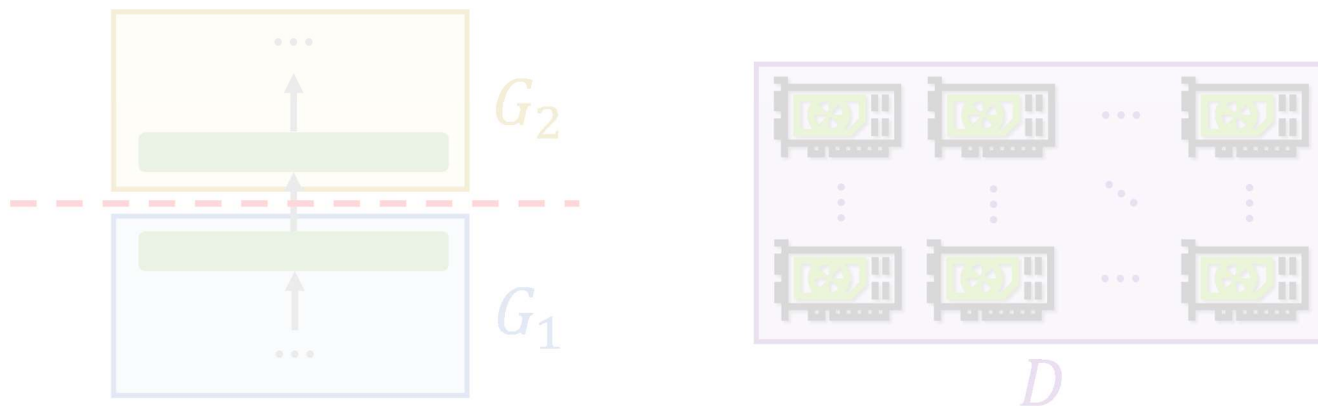


$$\text{OptCost}(G_1 \rightarrow G_2, D)$$



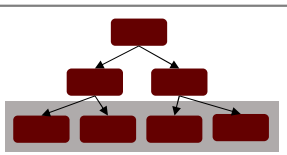


Hierarchical Search Algorithm

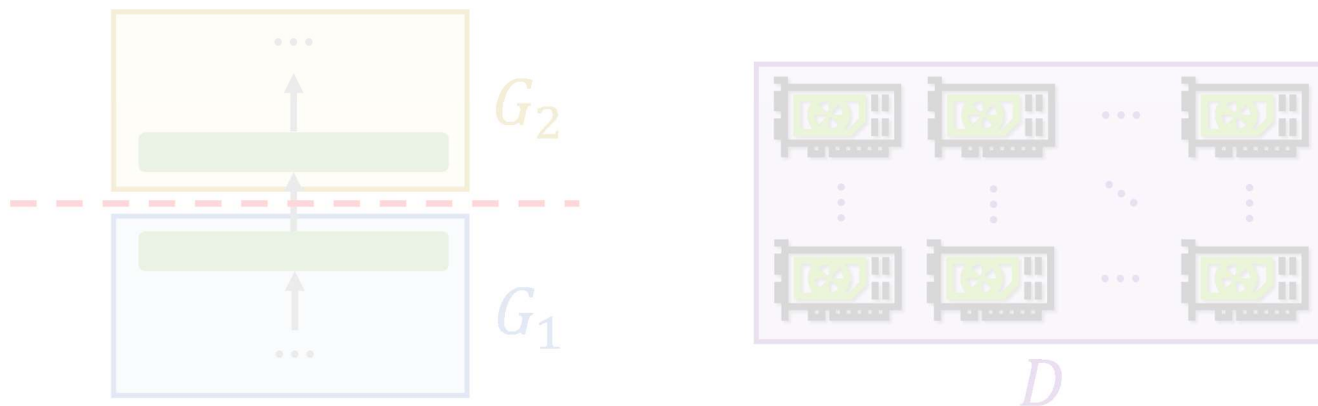


$$\text{OptCost}(G_1 \rightarrow G_2, D) = \text{OptCost}(G_1, \quad)$$



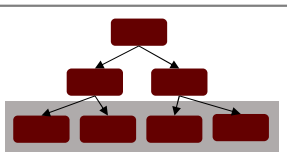


Hierarchical Search Algorithm

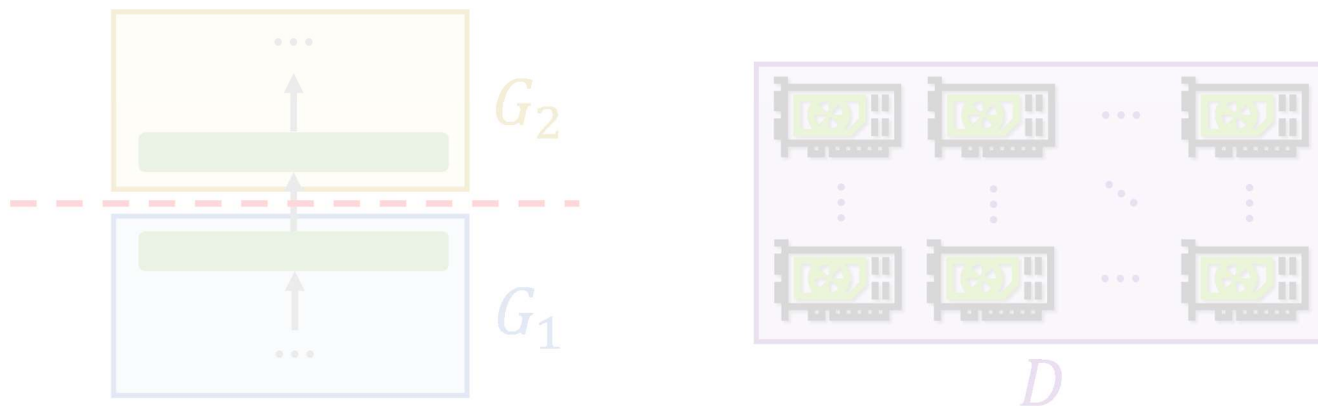


$$\text{OptCost}(G_1 \rightarrow G_2, D) = \text{OptCost}(G_1, \quad) + \text{OptCost}(G_2, \quad)$$



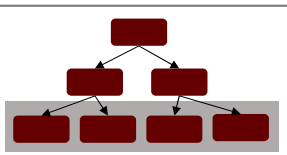


Hierarchical Search Algorithm

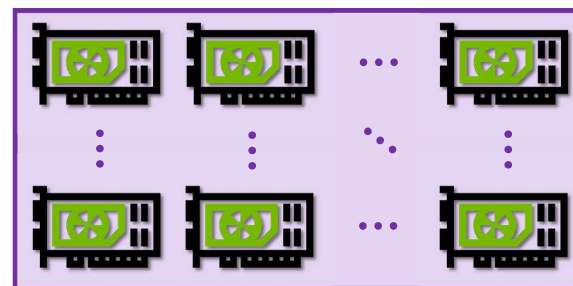
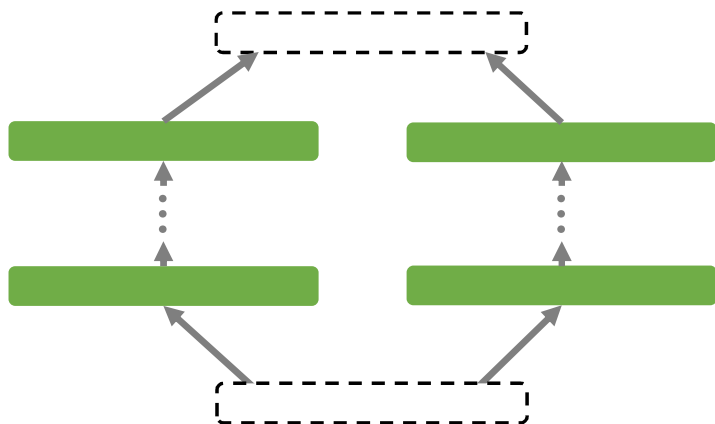


$$\text{OptCost}(G_1 \rightarrow G_2, D) = \text{OptCost}(G_1, D) + \text{OptCost}(G_2, D)$$





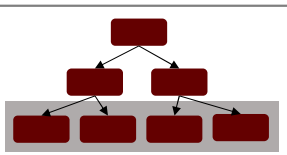
Hierarchical Search Algorithm



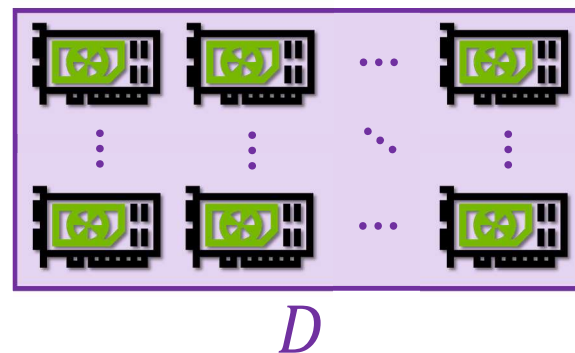
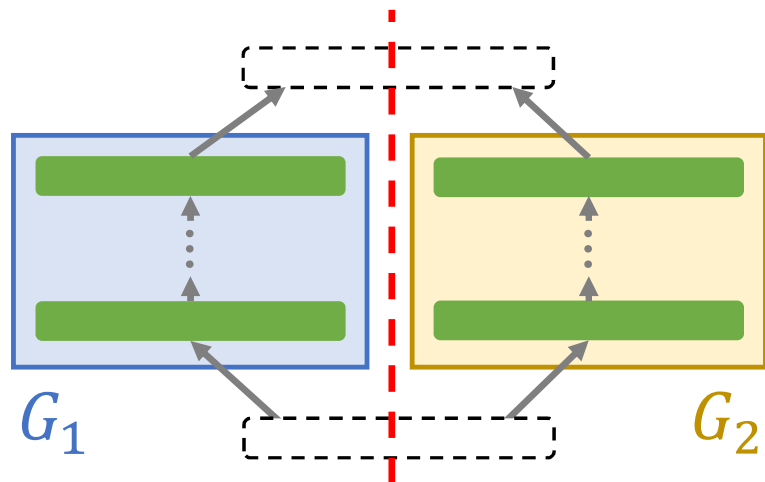
D

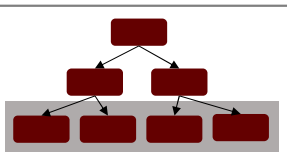


Dynamic
Programming

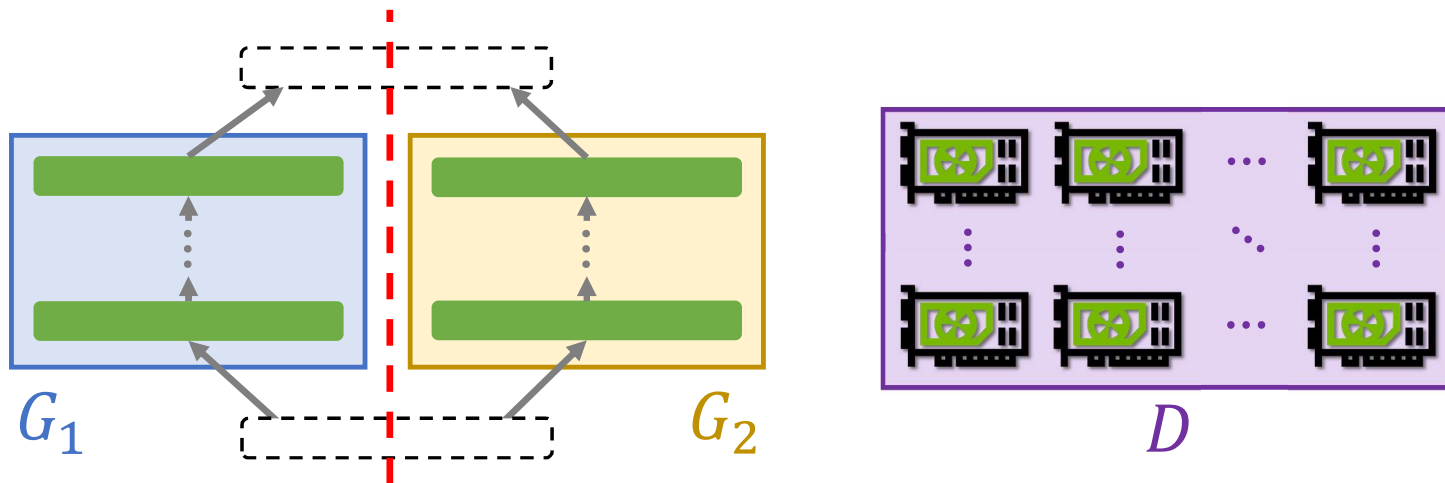


Hierarchical Search Algorithm



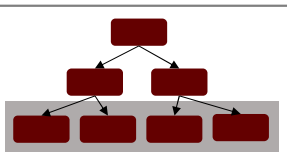


Hierarchical Search Algorithm

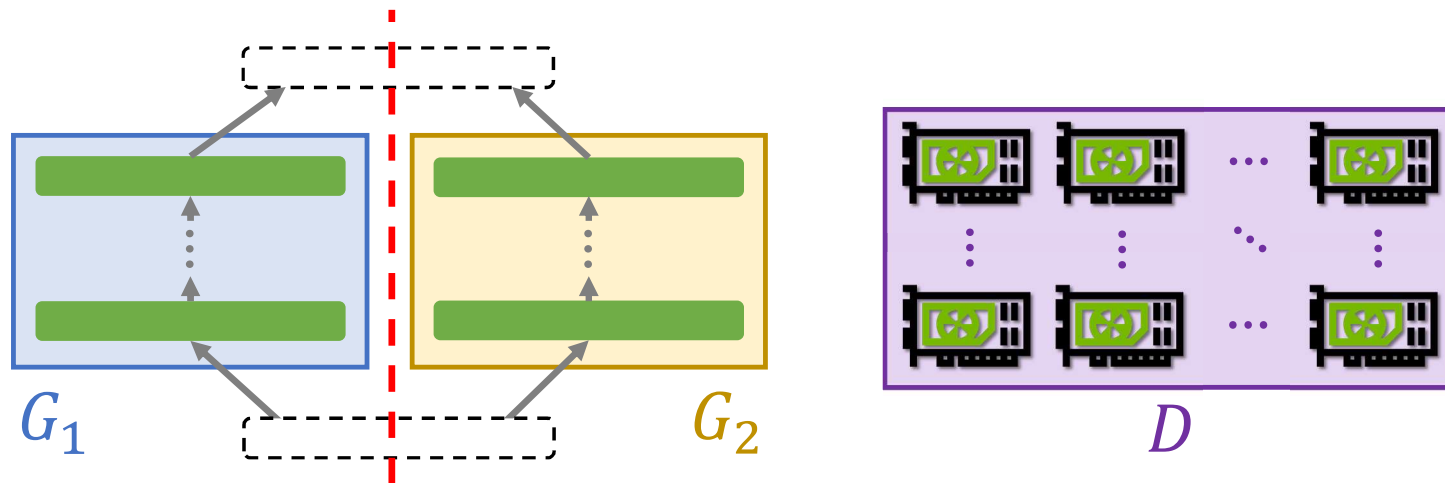


$$\text{OptCost}(G_1 \uparrow\uparrow G_2, D)$$

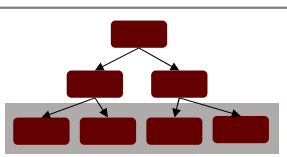




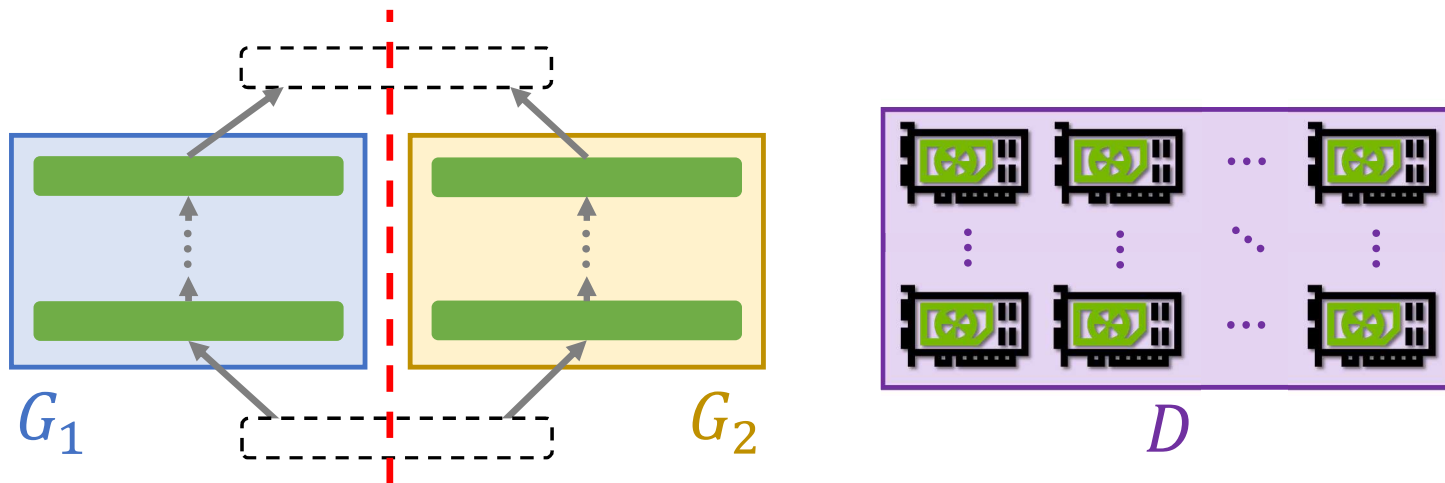
Hierarchical Search Algorithm



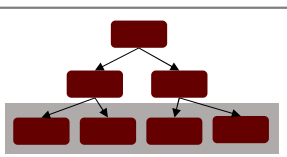
$$\text{OptCost}(G_1 \uparrow\uparrow G_2, D) = \min \left\{ \right.$$



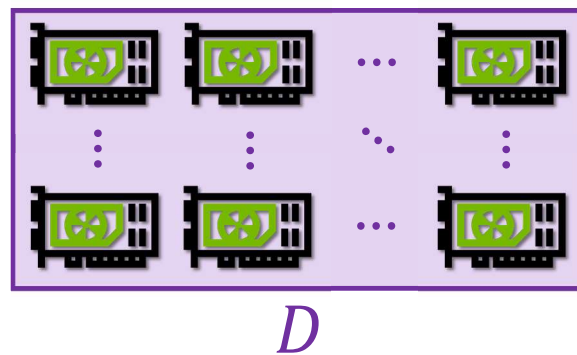
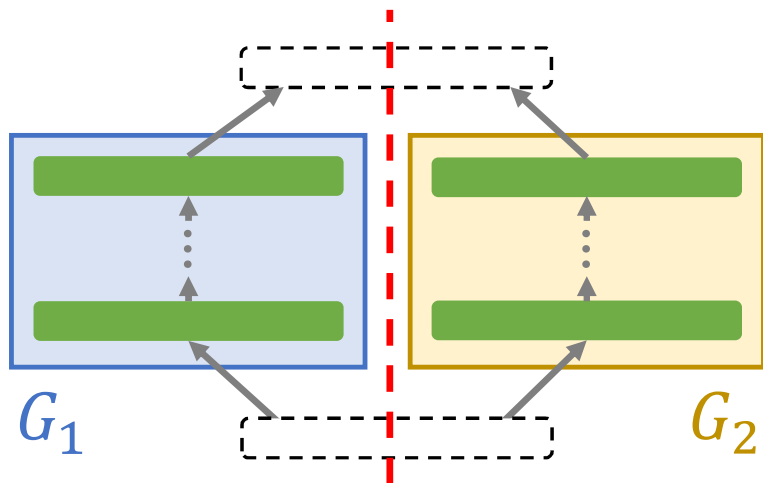
Hierarchical Search Algorithm



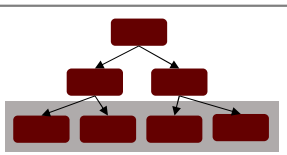
$$\begin{aligned}
 &\rightarrow G_1 \longrightarrow G_2 \rightarrow \\
 &\rightarrow G_2 \longrightarrow G_1 \rightarrow \text{OptCost}(G_1 \uparrow\uparrow G_2, D) = \min \left\{ \begin{array}{l} \\ \\ \\ \end{array} \right.
 \end{aligned}$$



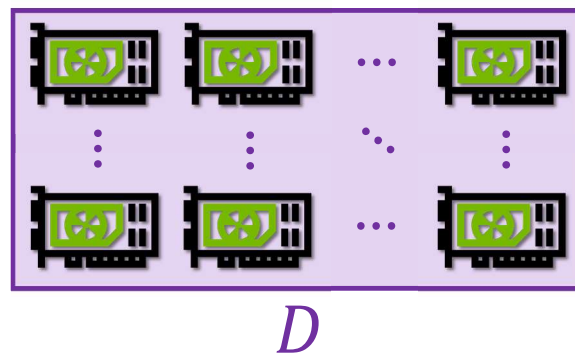
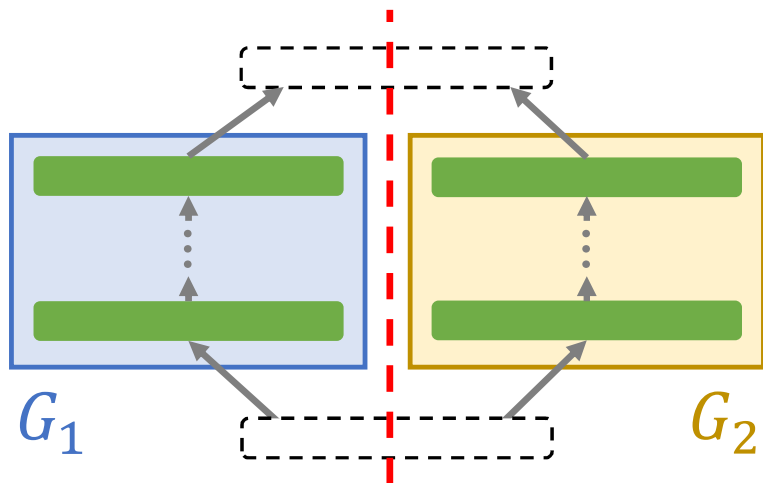
Hierarchical Search Algorithm



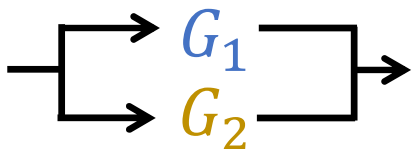
$$\begin{aligned}
 &\rightarrow G_1 \longrightarrow G_2 \rightarrow \\
 &\rightarrow G_2 \longrightarrow G_1 \rightarrow \text{OptCost}(G_1 \uparrow\uparrow G_2, D) = \min \left\{ \begin{aligned} &\text{OptCost}(G_1, D) \\ &+ \text{OptCost}(G_2, D) \end{aligned} \right.
 \end{aligned}$$

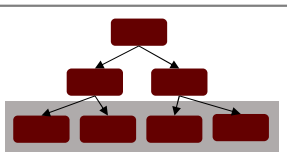


Hierarchical Search Algorithm

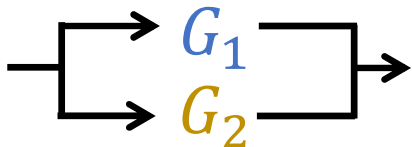
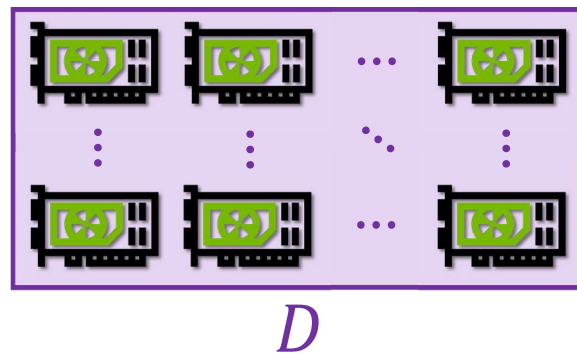
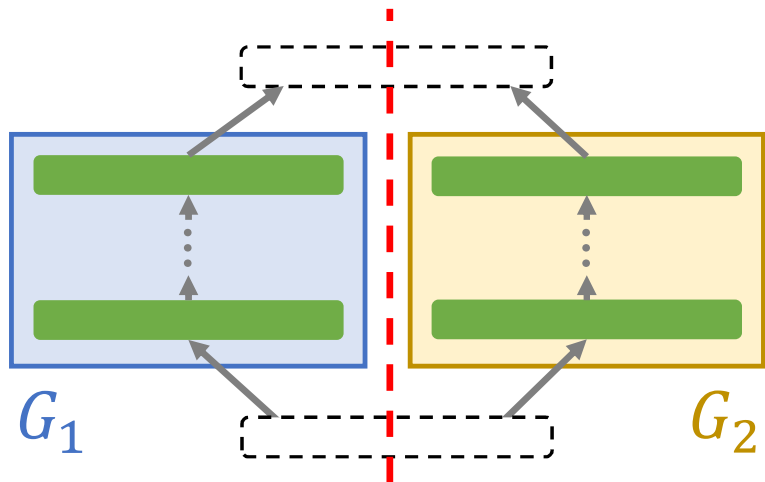


$$\text{OptCost}(G_1 \uparrow\uparrow G_2, D) = \min \left\{ \begin{array}{l} \text{OptCost}(G_1, D) \\ + \text{OptCost}(G_2, D) \end{array} \right.$$

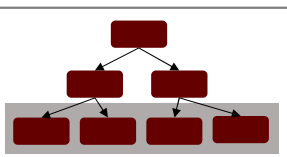




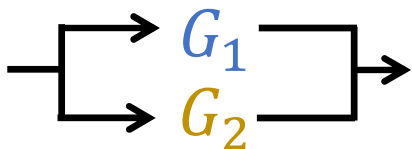
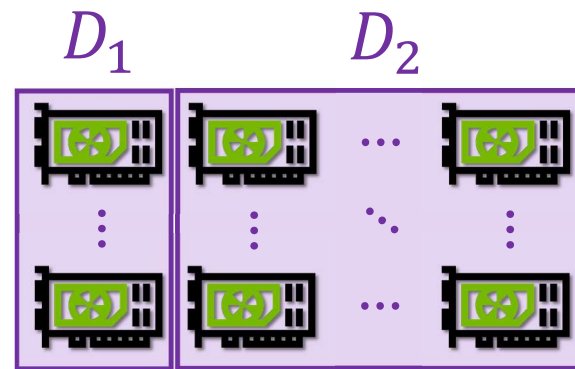
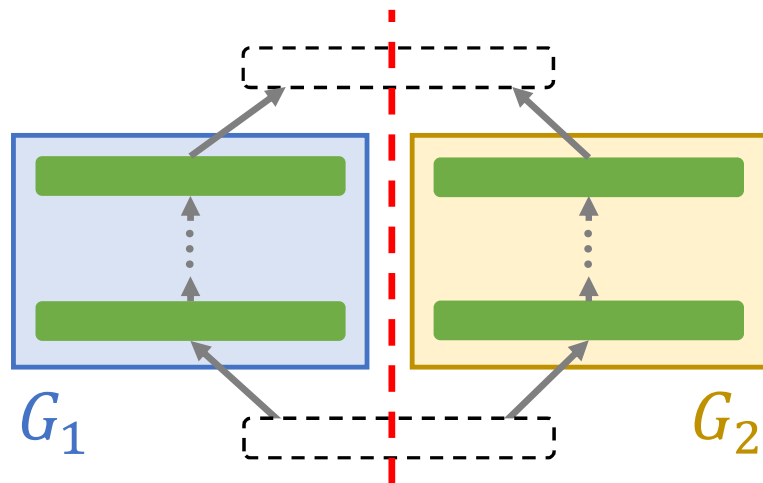
Hierarchical Search Algorithm



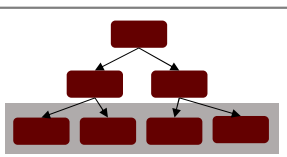
$$\text{OptCost}(G_1 \uparrow\uparrow G_2, D) = \min \left[\begin{array}{l} \text{OptCost}(G_1, D) \\ + \text{OptCost}(G_2, D) \\ \max[\text{OptCost}(G_1, D_1), \\ \text{OptCost}(G_2, D_2)] \end{array} \right]$$



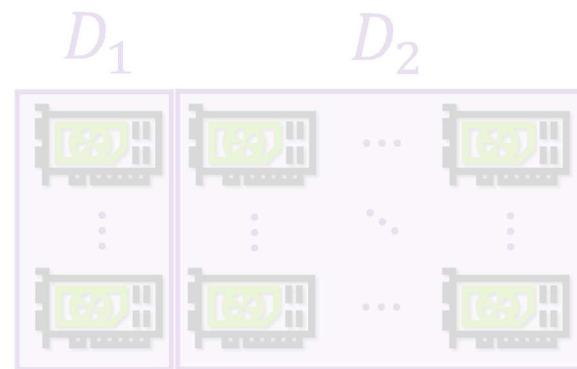
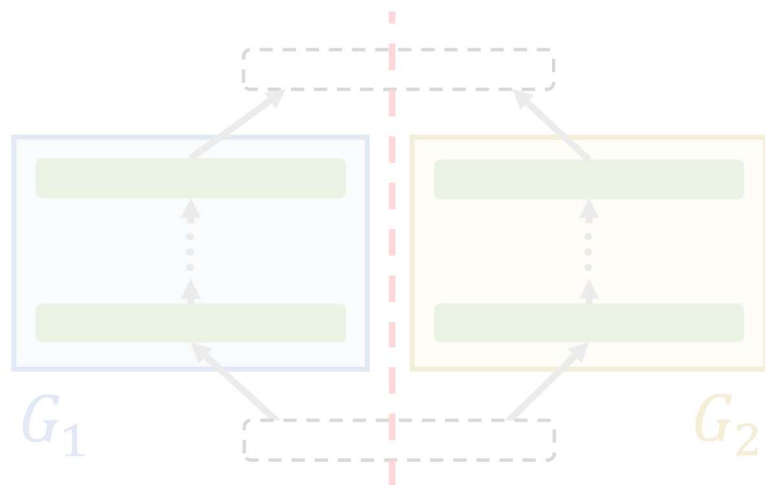
Hierarchical Search Algorithm



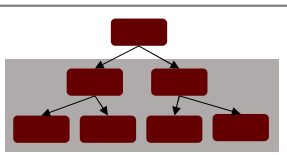
$$\text{OptCost}(G_1 \uparrow\uparrow G_2, D) = \min \left\{ \begin{array}{l} \text{OptCost}(G_1, D) \\ \quad + \text{OptCost}(G_2, D) \\ \max[\text{OptCost}(G_1, D_1), \\ \quad \text{OptCost}(G_2, D_2)] \end{array} \right.$$



Hierarchical Search Algorithm



$$\text{OptCost}(G_1 \uparrow\uparrow G_2, D) = \min \begin{cases} \text{OptCost}(G_1, D) \\ \quad + \text{OptCost}(G_2, D) \\ \max[\text{OptCost}(G_1, D_1), \\ \quad \text{OptCost}(G_2, D_2)] \end{cases}$$



Hierarchical Search Algorithm

Algebraic Transformation

Backtracking
Search

Parallelism Type



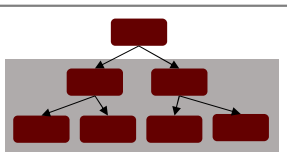
Parallelism Degree

Device Mapping



Dynamic
Programming

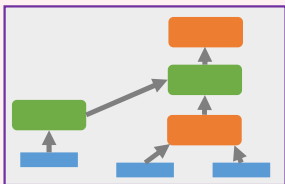




Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



Parallelism Type



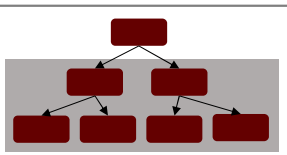
Parallelism Degree

Device Mapping



Dynamic Programming

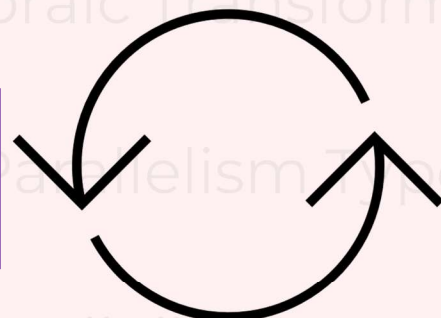
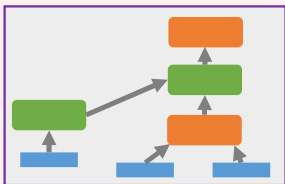




Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



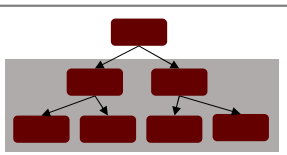
Parallelism Degree

Device Mapping



Dynamic Programming

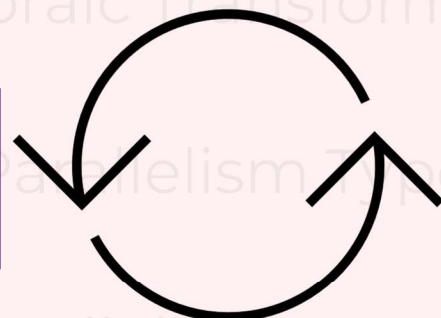
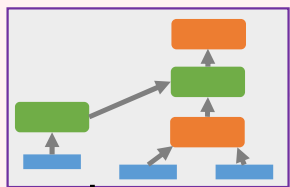




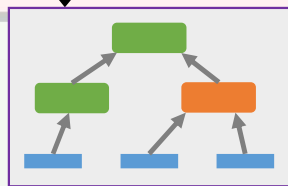
Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



Parallelism Degree

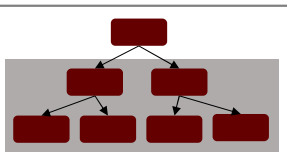


Space Mapping

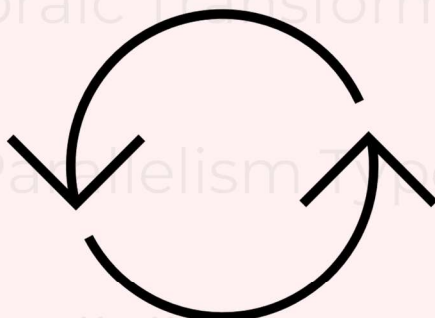
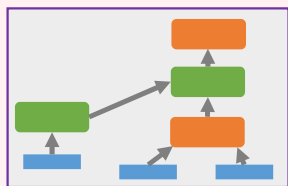


Dynamic Programming





Hierarchical Search Algorithm

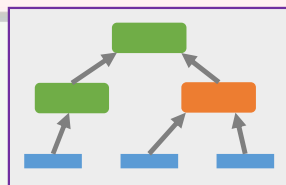


Algebraic Transformation

Backtracking Search



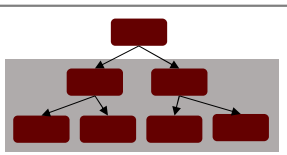
Parallelism Degree



ce Mapping



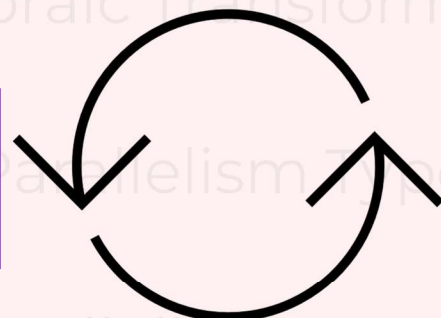
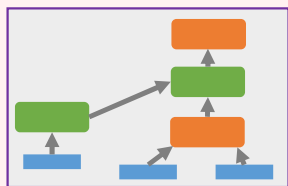
Dynamic Programming



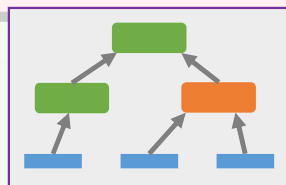
Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



Parallelism Degree

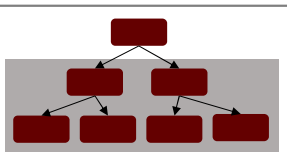


Space Mapping



Dynamic Programming

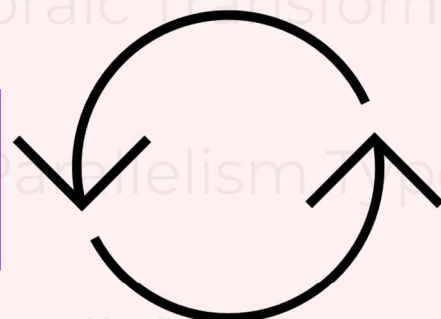
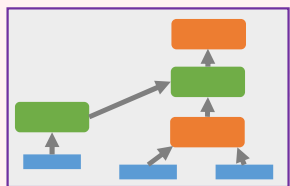




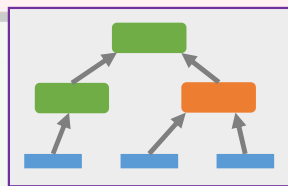
Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



Parallelism Degree

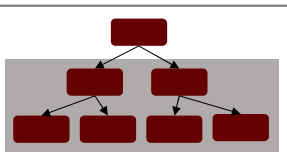


$\text{OptCost}(G, D)$



Dynamic Programming

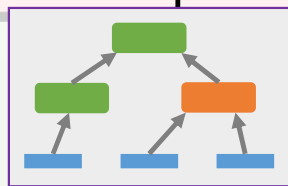
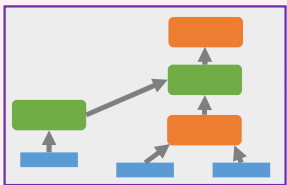
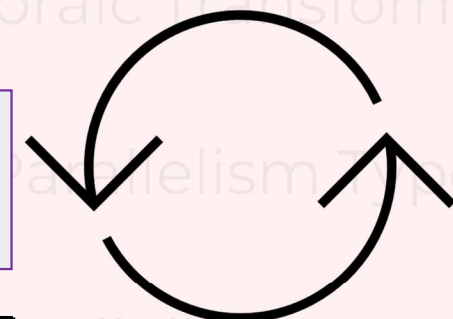




Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



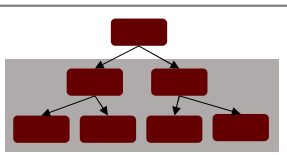
$\text{OptCost}(G, D)$

Cost Matrix



Dynamic Programming

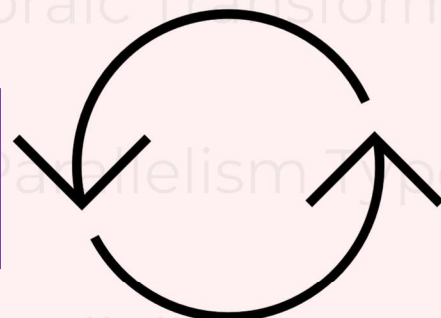
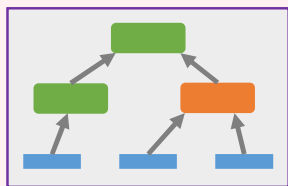




Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



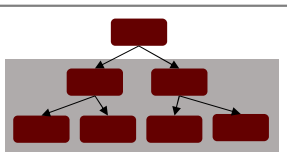
Parallelism Degree

Device Mapping



Dynamic Programming

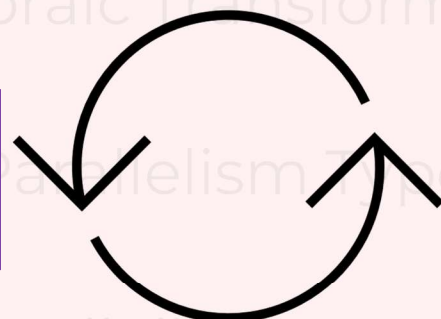
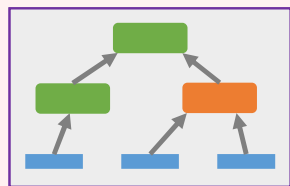




Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



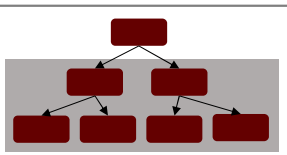
Parallelism Degree

Device Mapping



Dynamic Programming

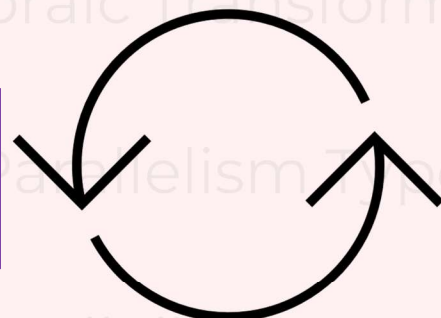
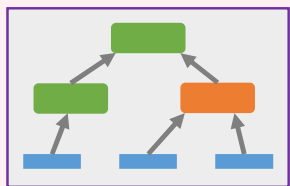




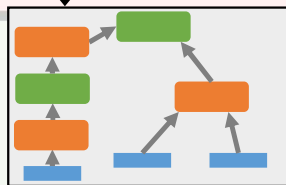
Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



Parallelism Degree

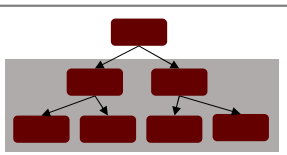


Space Mapping



Dynamic Programming

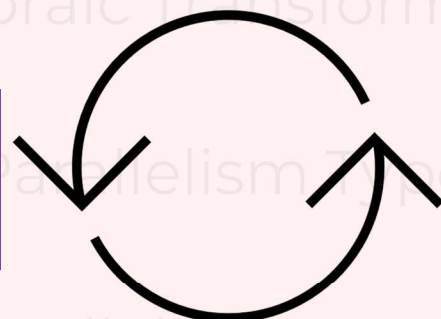
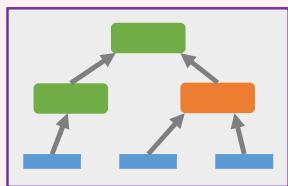




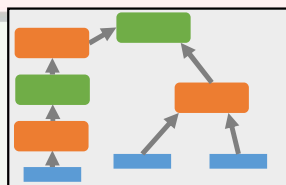
Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



Parallelism Degree

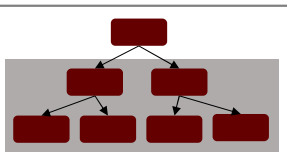


Space Mapping



Dynamic Programming

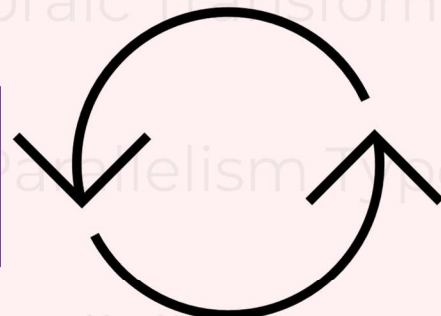
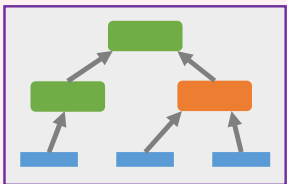




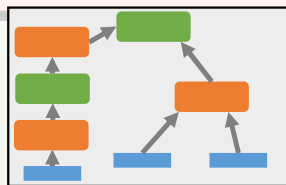
Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



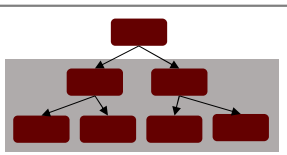
Parallelism Degree



Space Mapping

Dynamic Programming

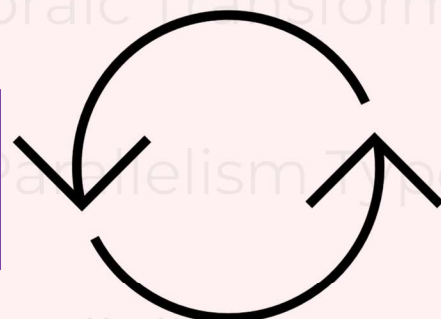
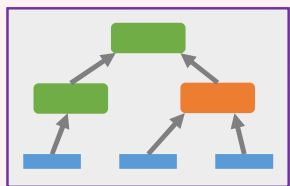




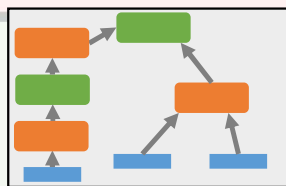
Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



Parallelism Degree

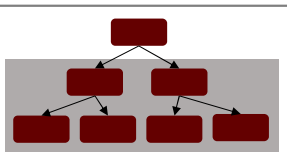


$\text{OptCost}(G, D)$



Dynamic Programming

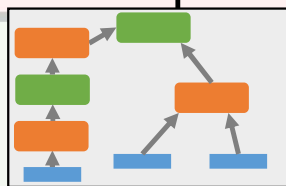
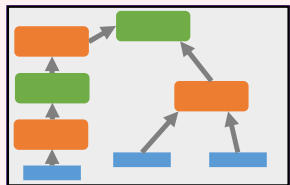
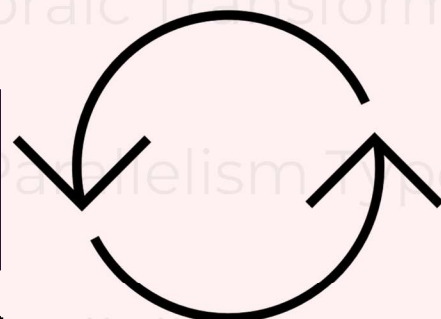




Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



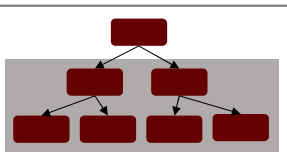
$\text{OptCost}(G, D)$

Cost Matrix



Dynamic Programming

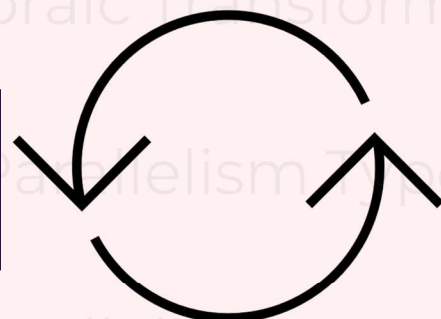
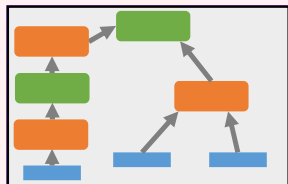




Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search



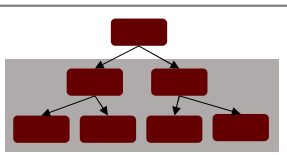
Parallelism Degree

Device Mapping



Dynamic Programming

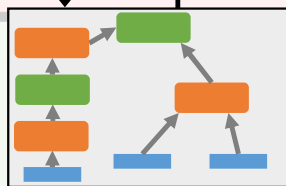
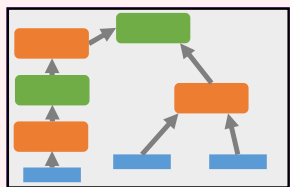
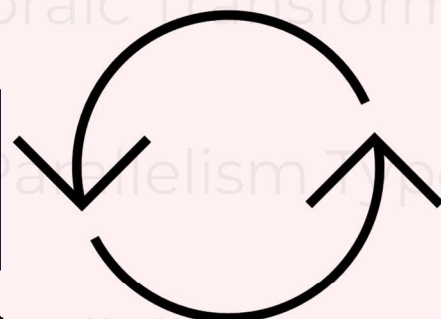




Hierarchical Search Algorithm

Algebraic Transformation

Backtracking Search

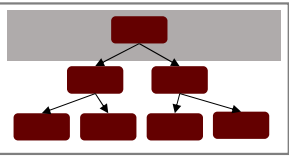


$\text{OptCost}(G, D)$

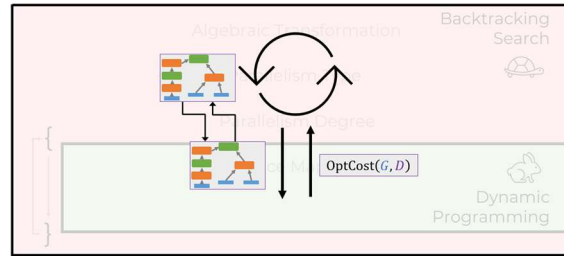


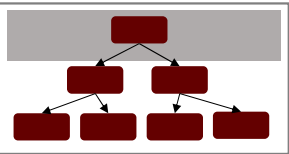
Dynamic Programming





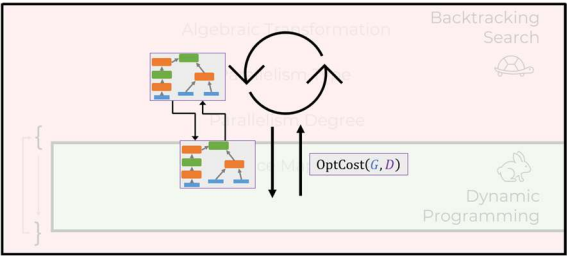
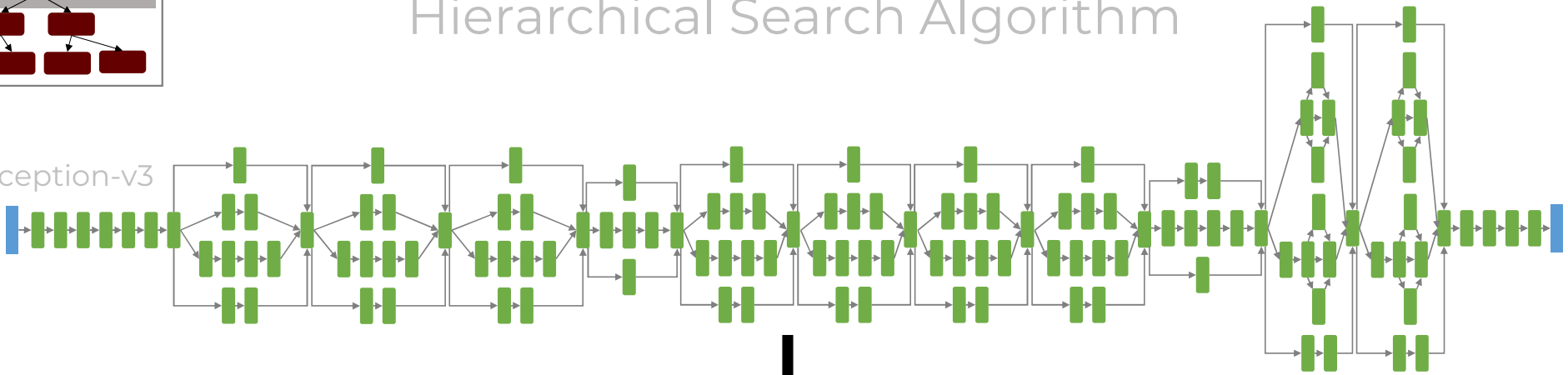
Hierarchical Search Algorithm





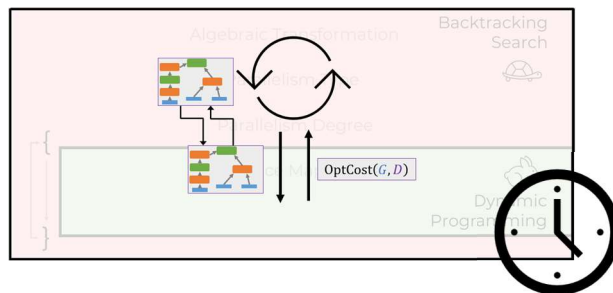
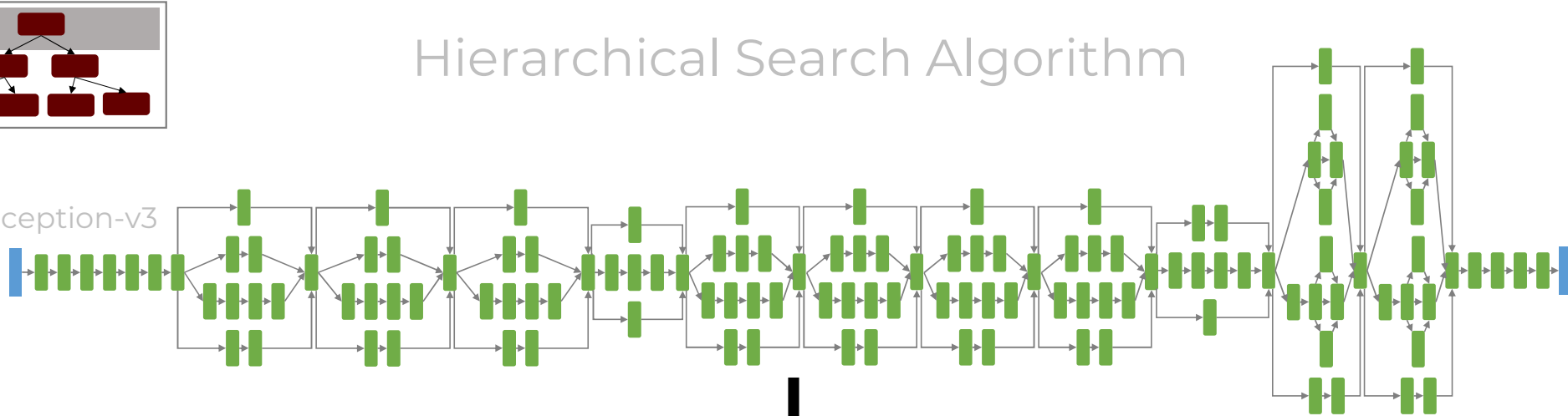
Hierarchical Search Algorithm

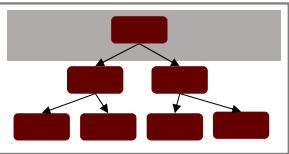
Inception-v3



Hierarchical Search Algorithm

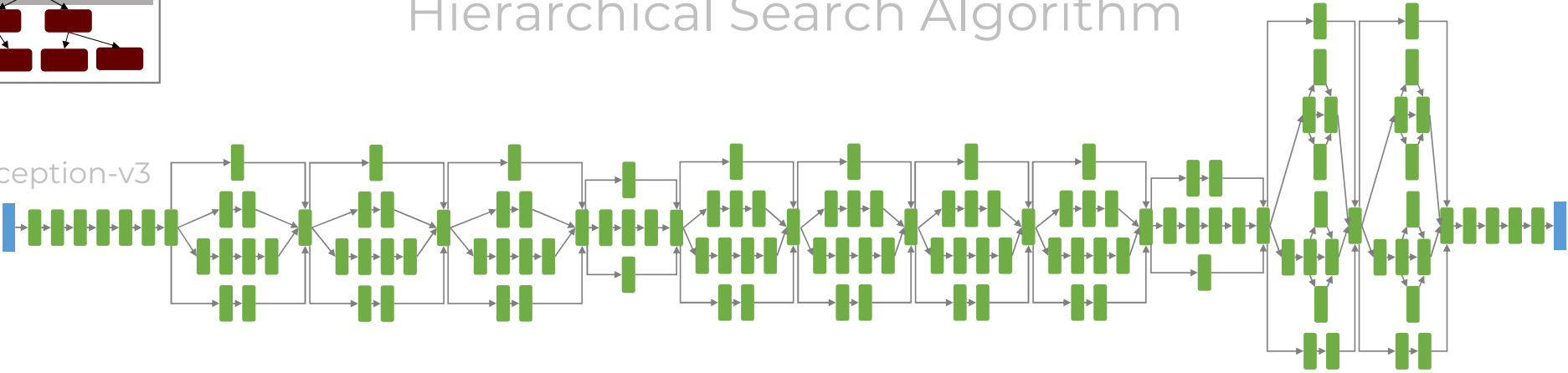
Inception-v3





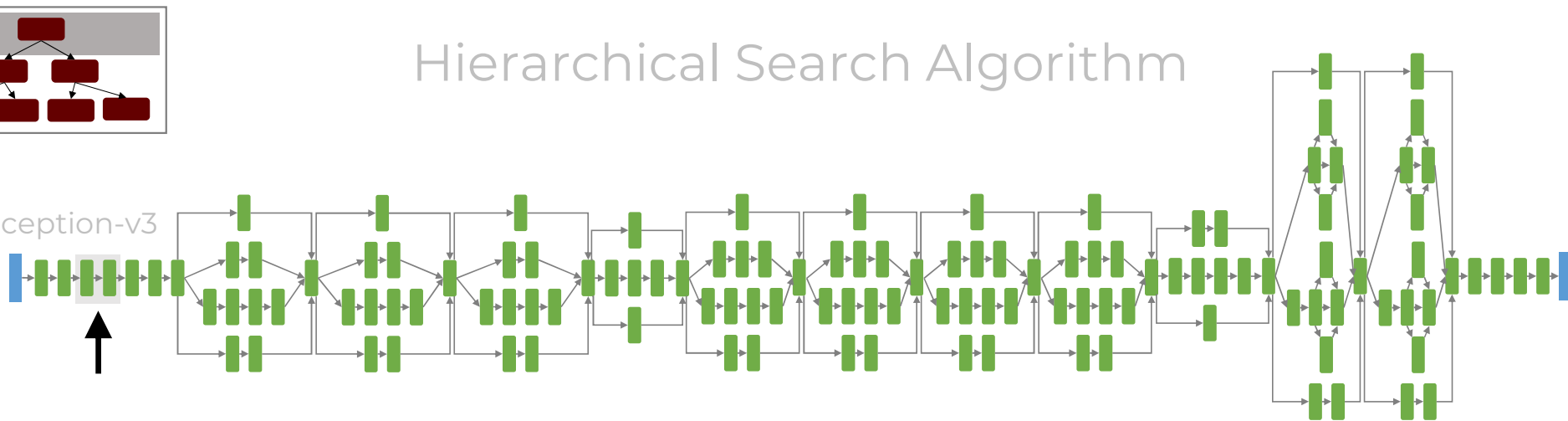
Hierarchical Search Algorithm

Inception-v3



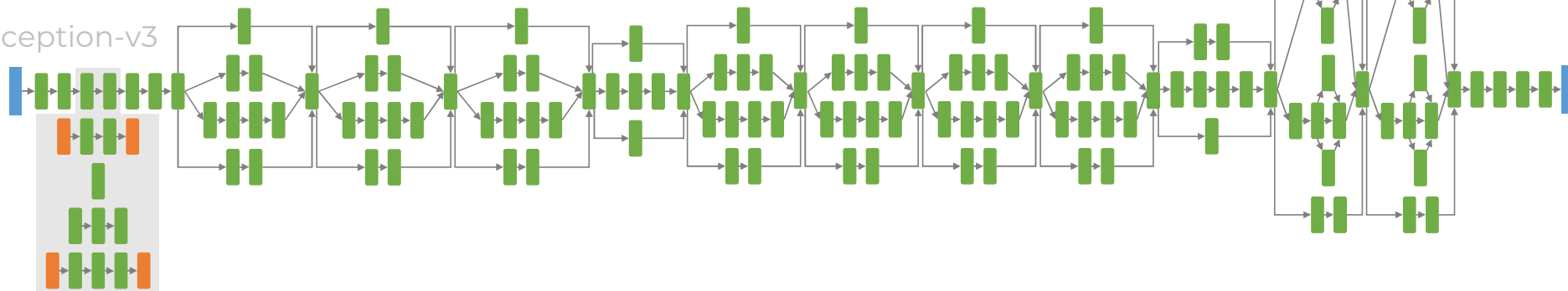
Hierarchical Search Algorithm

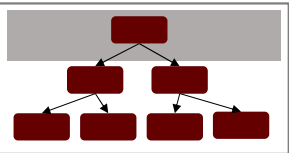
Inception-v3



Hierarchical Search Algorithm

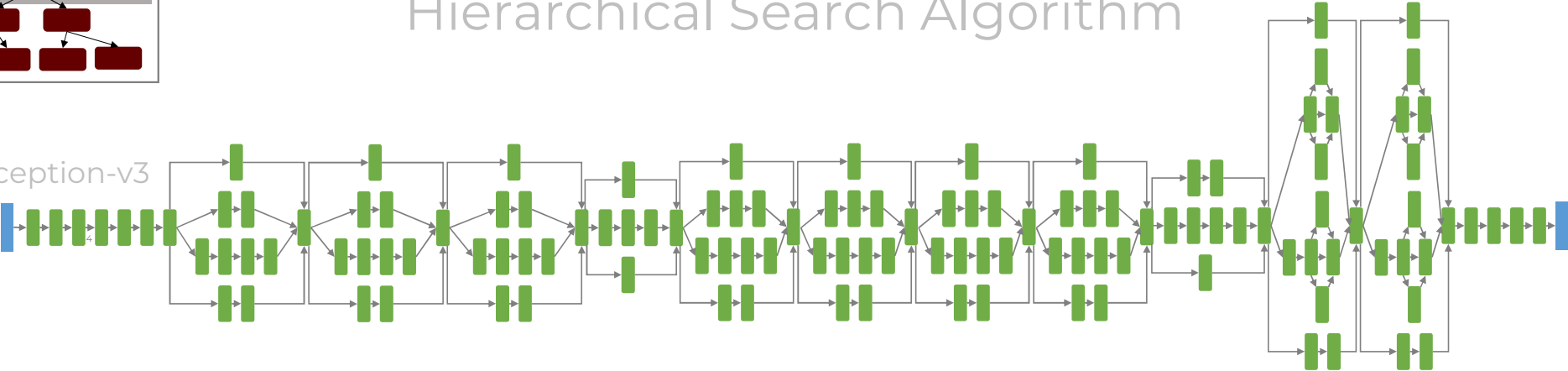
Inception-v3

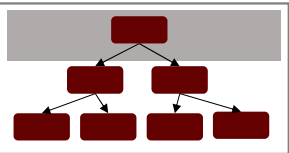




Hierarchical Search Algorithm

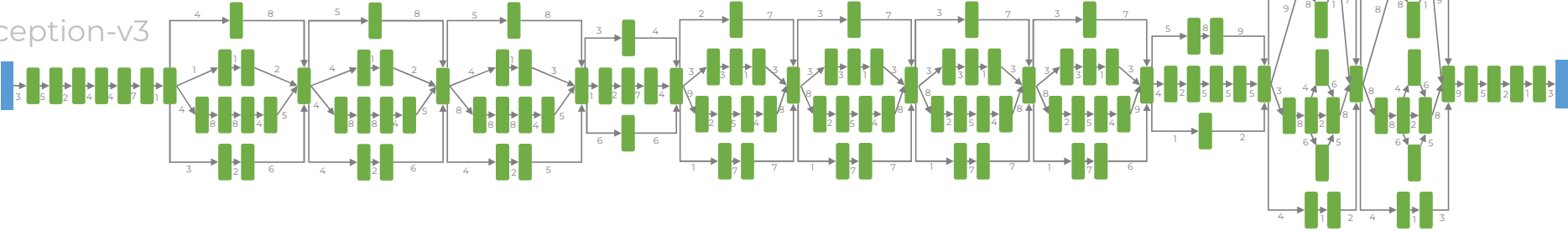
Inception-v3

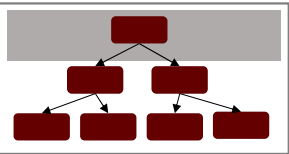




Hierarchical Search Algorithm

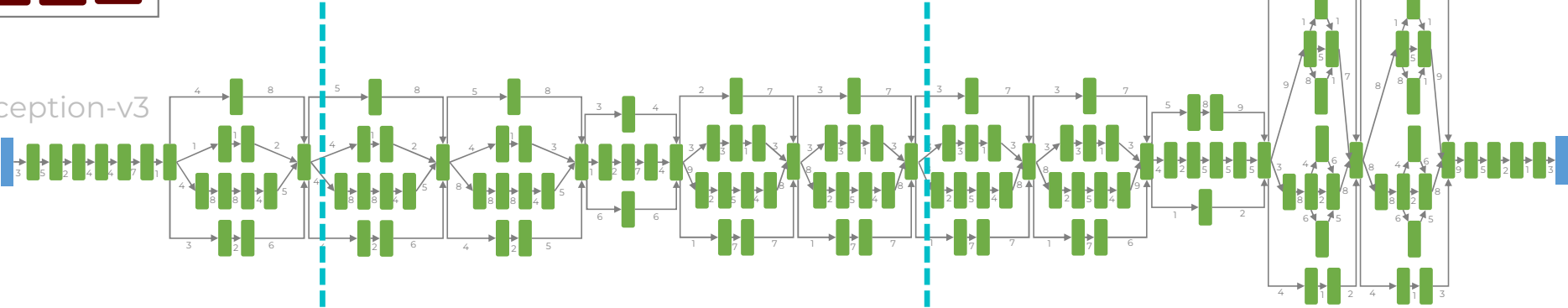
Inception-v3



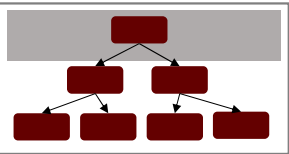


Hierarchical Search Algorithm

Inception-v3

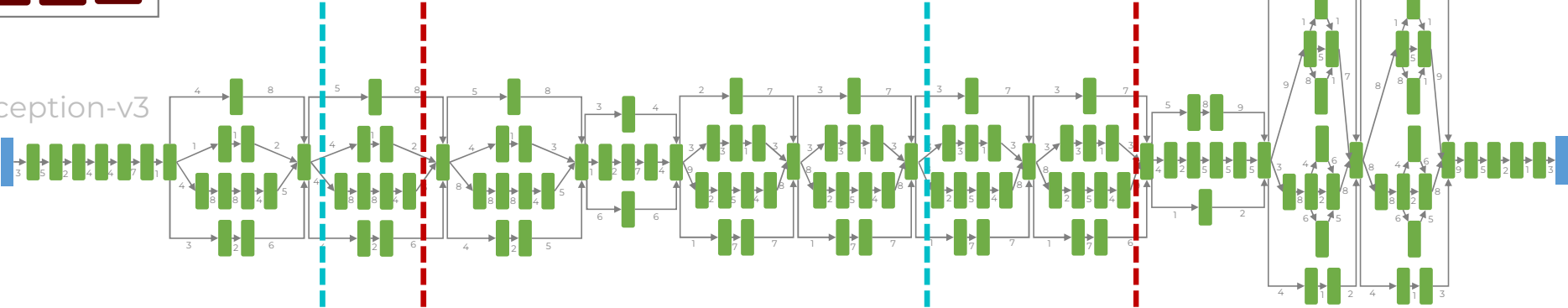


32



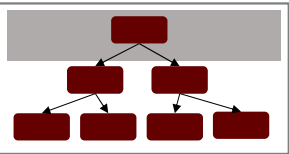
Hierarchical Search Algorithm

Inception-v3



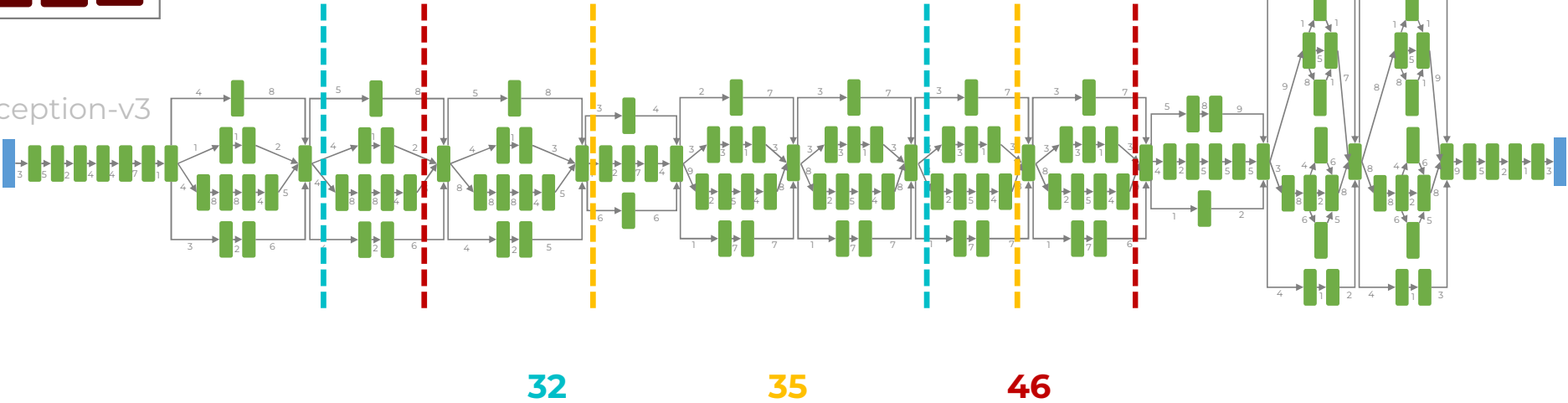
32

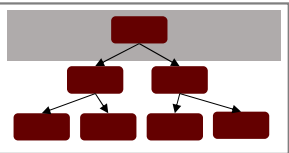
46



Hierarchical Search Algorithm

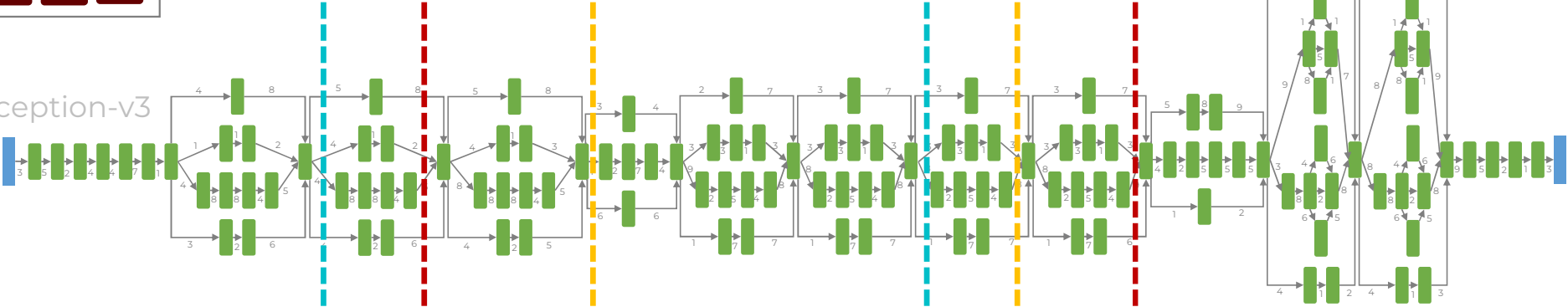
Inception-v3



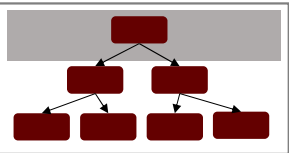


Hierarchical Search Algorithm

Inception-v3

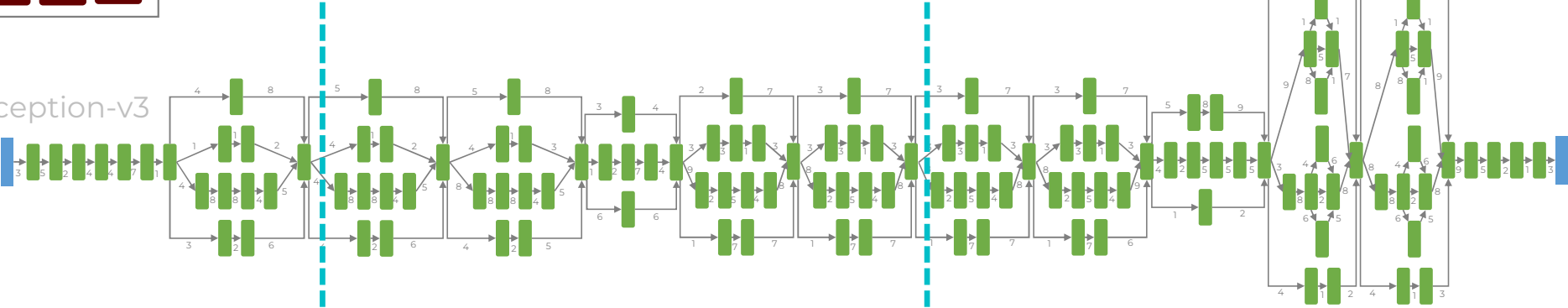


32 < 35 < 46

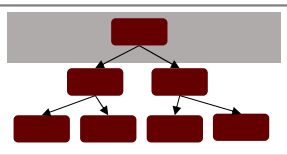


Hierarchical Search Algorithm

Inception-v3

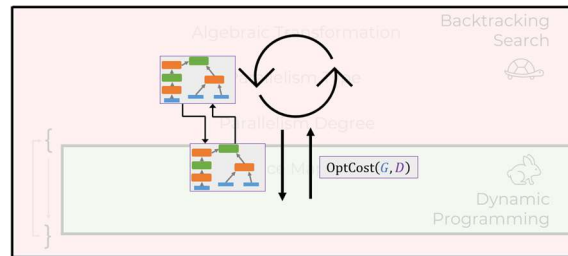
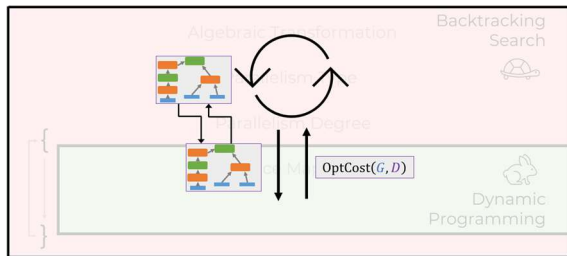
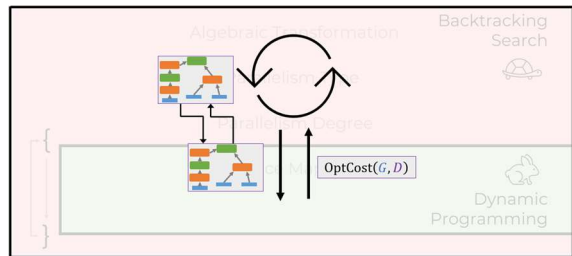
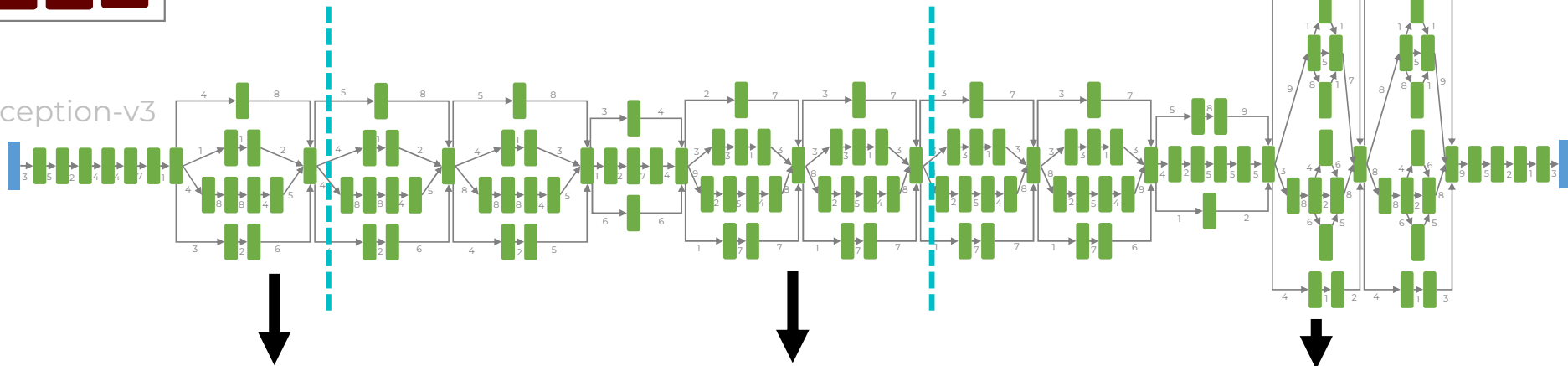


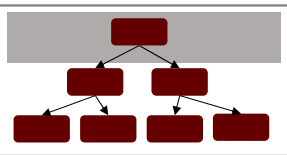
32



Hierarchical Search Algorithm

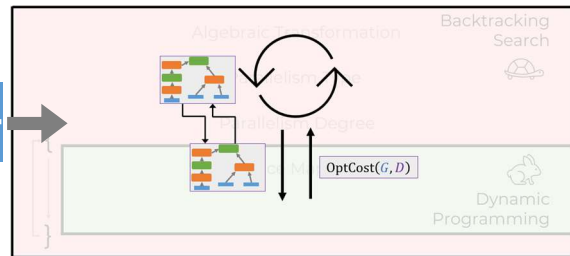
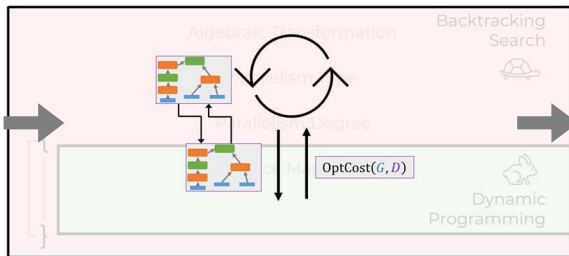
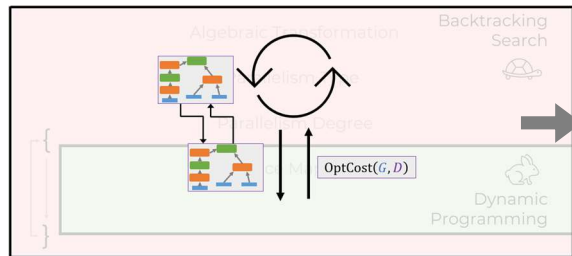
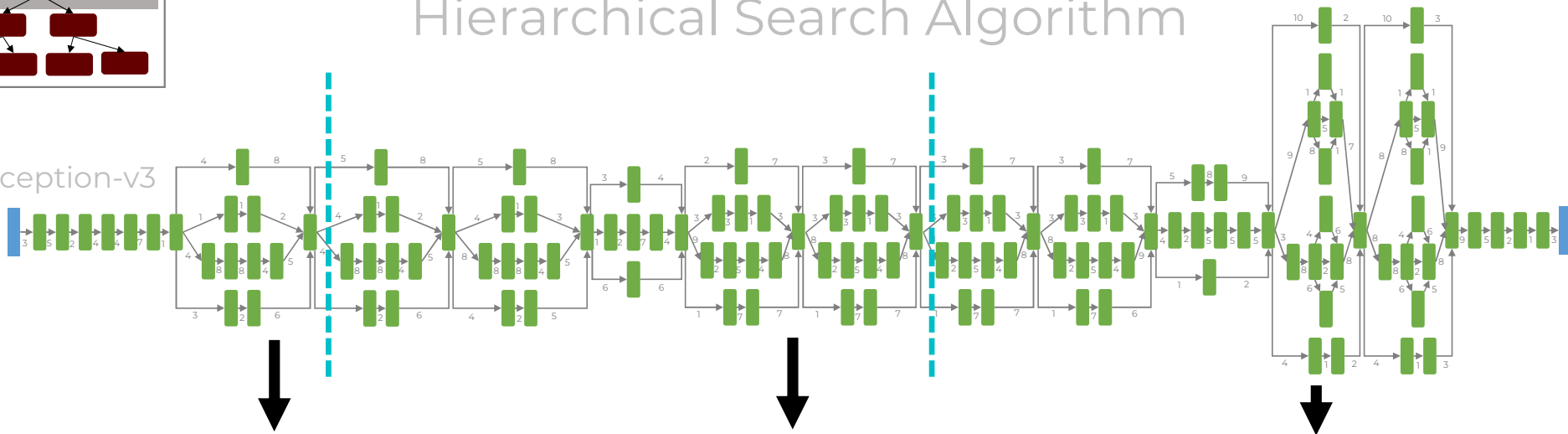
Inception-v3



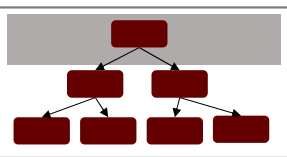


Hierarchical Search Algorithm

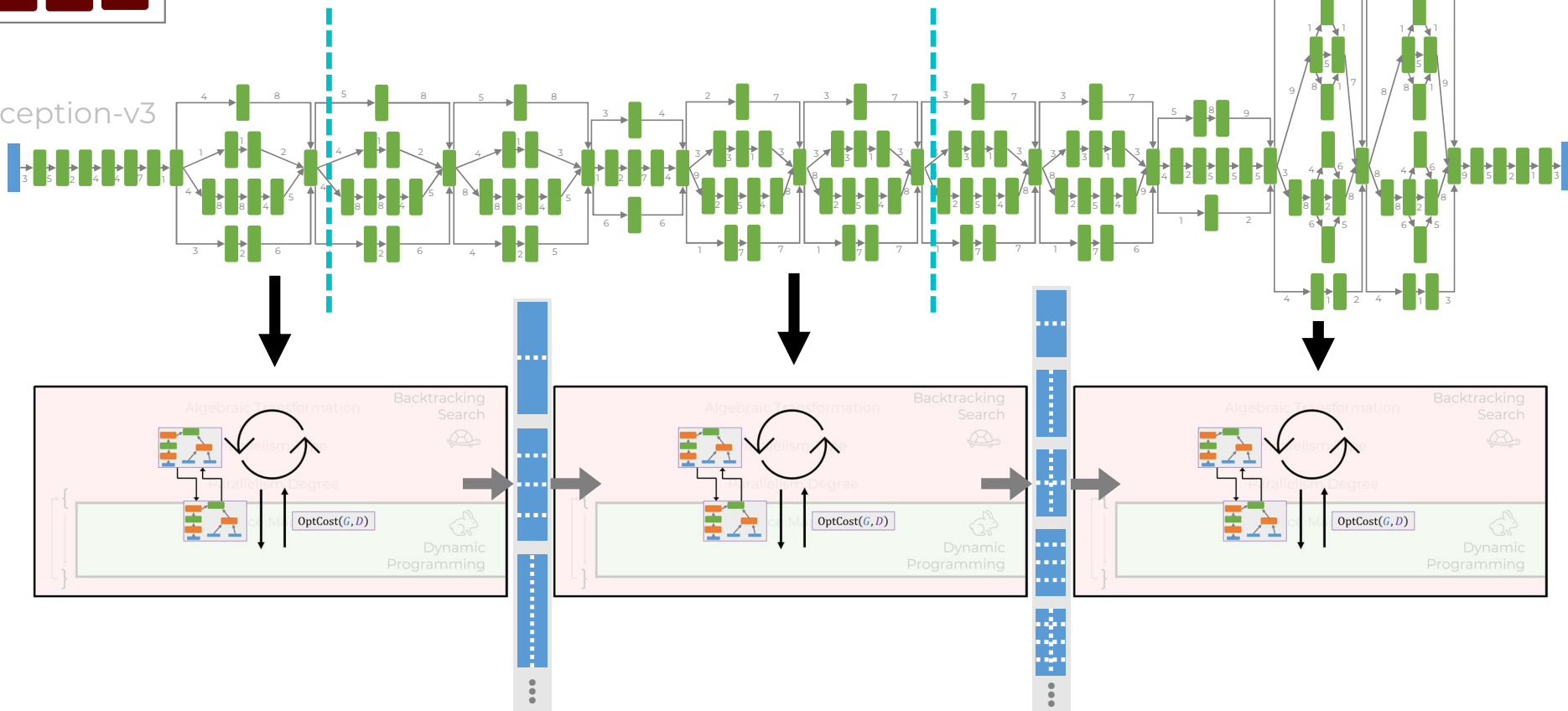
Inception-v3

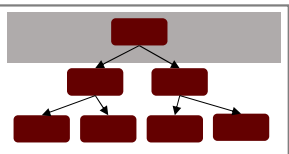


Hierarchical Search Algorithm



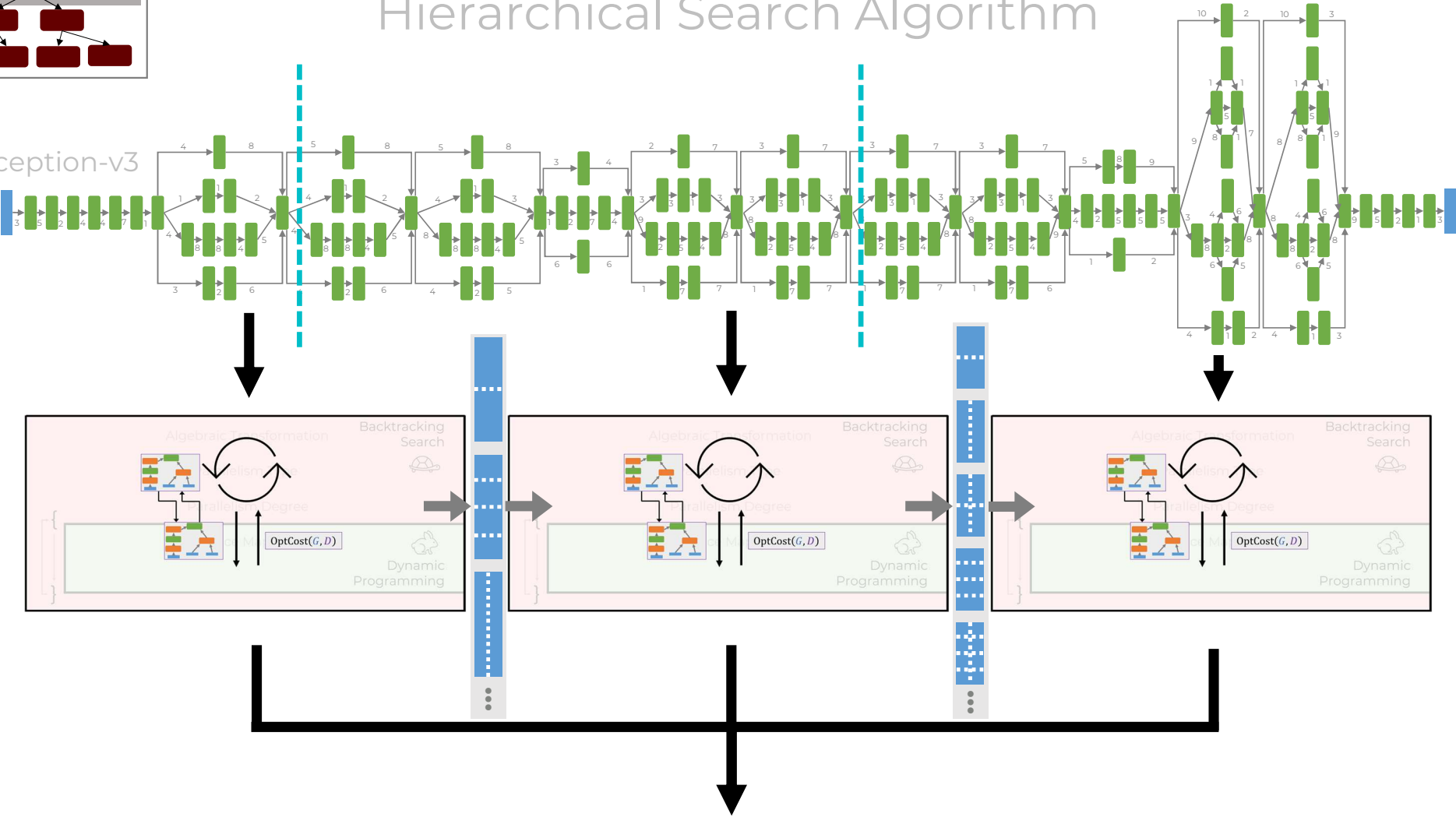
Inception-v3



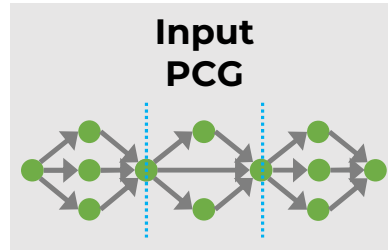


Hierarchical Search Algorithm

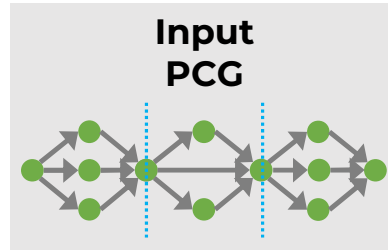
Inception-v3



Hierarchical Search Algorithm

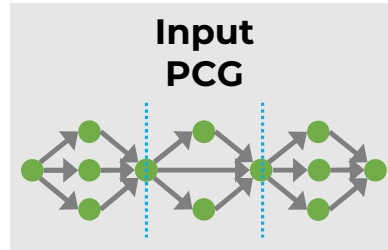


Hierarchical Search Algorithm



**Optimized
PCG**

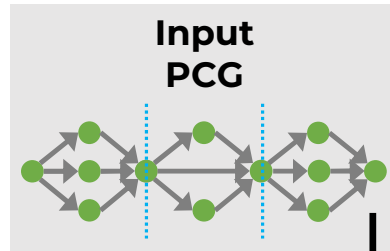
Hierarchical Search Algorithm



**Optimized
PCG**

**Machine
Mapping**

Hierarchical Search Algorithm

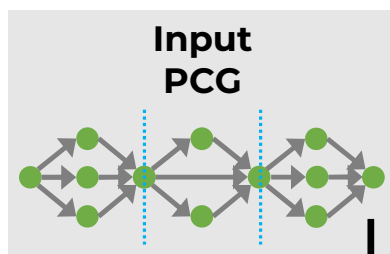


**Optimized
PCG**

**Machine
Mapping**

Graph Splitting

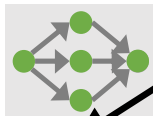
Hierarchical Search Algorithm



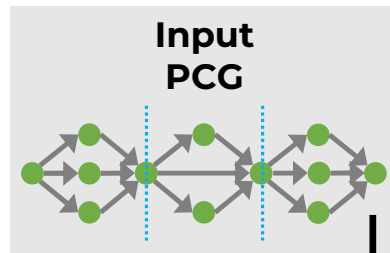
**Optimized
PCG**

**Machine
Mapping**

Graph Splitting



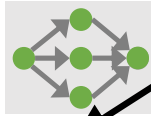
Hierarchical Search Algorithm



**Optimized
PCG**

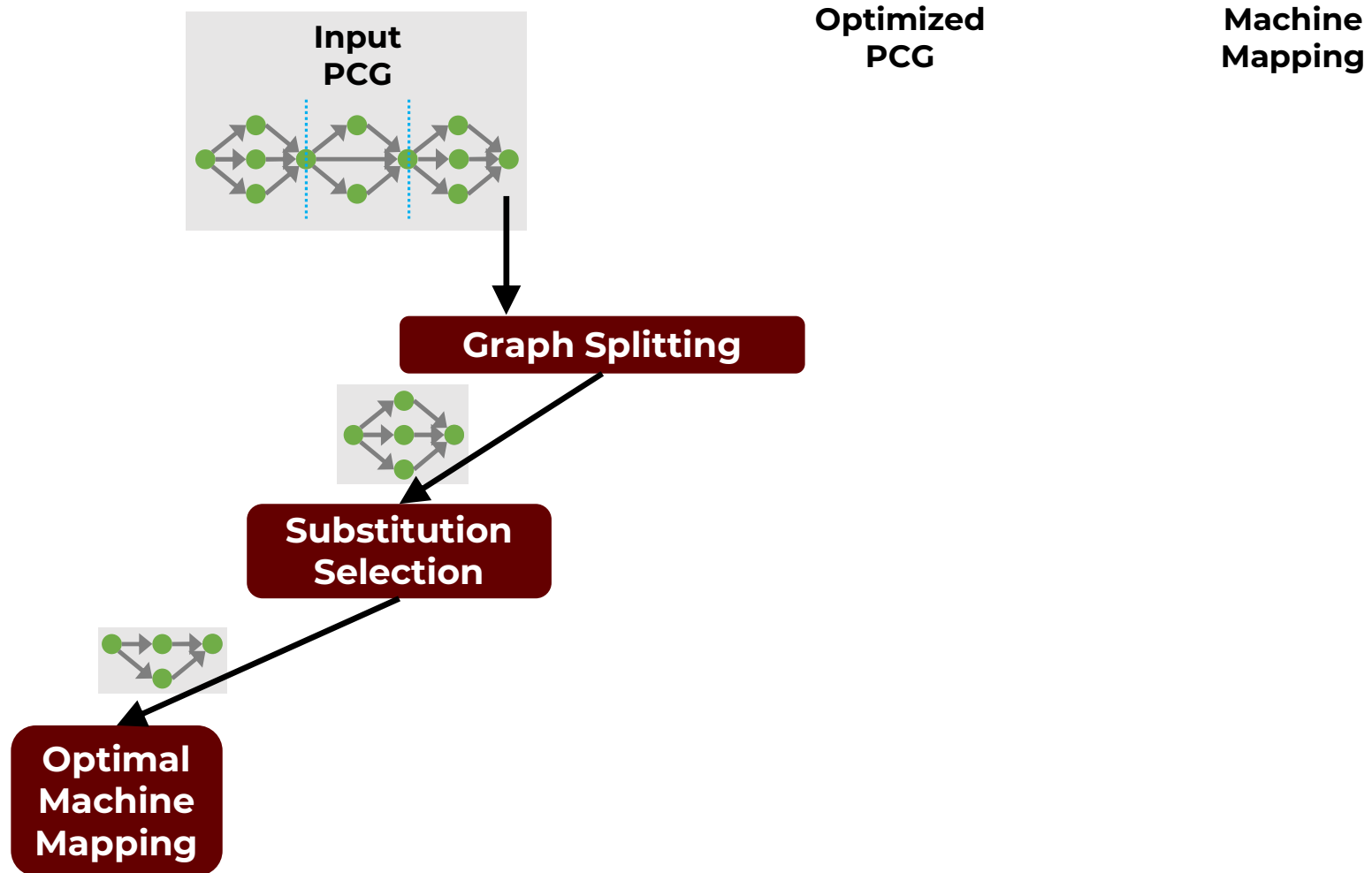
**Machine
Mapping**

Graph Splitting

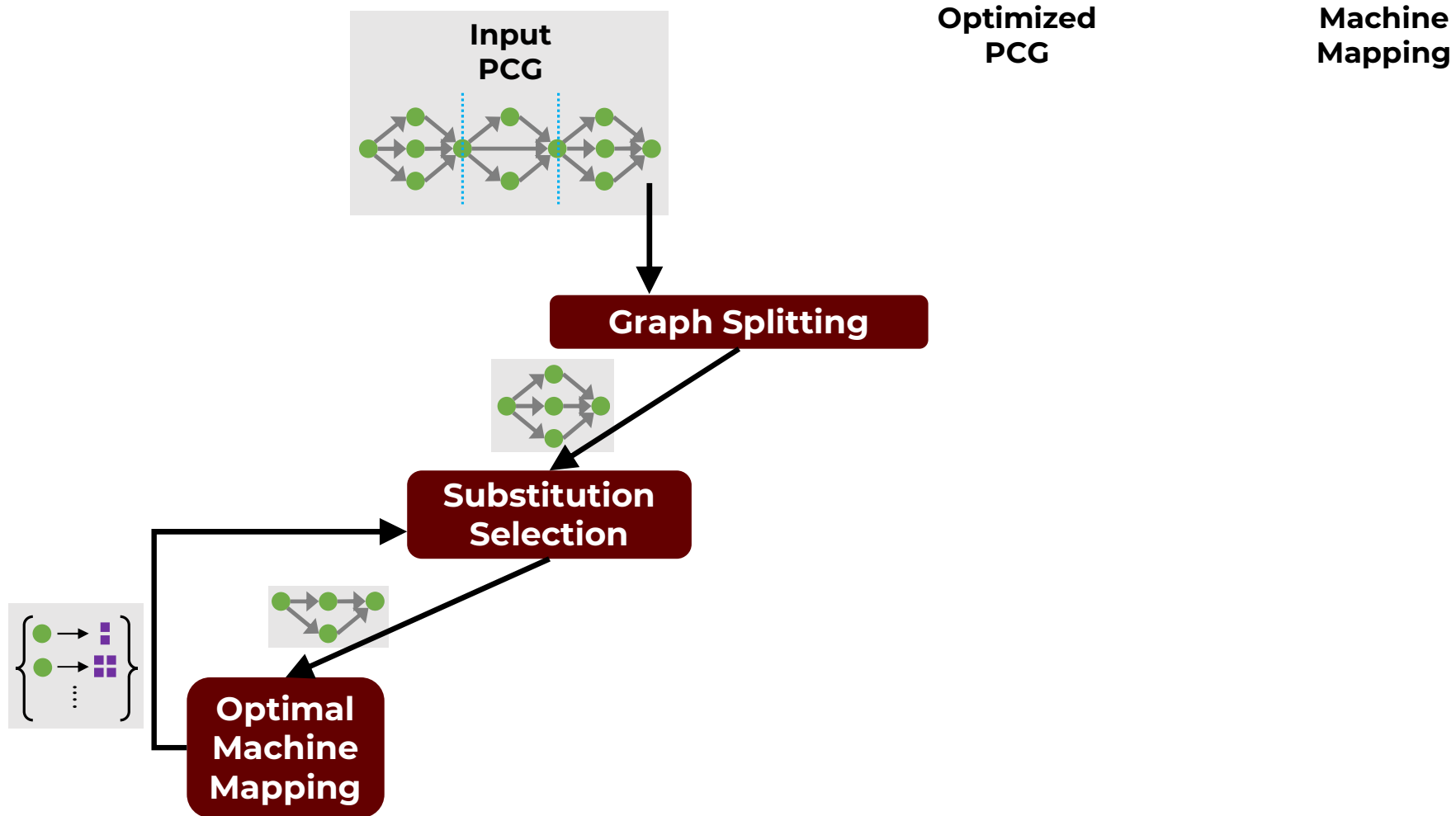


**Substitution
Selection**

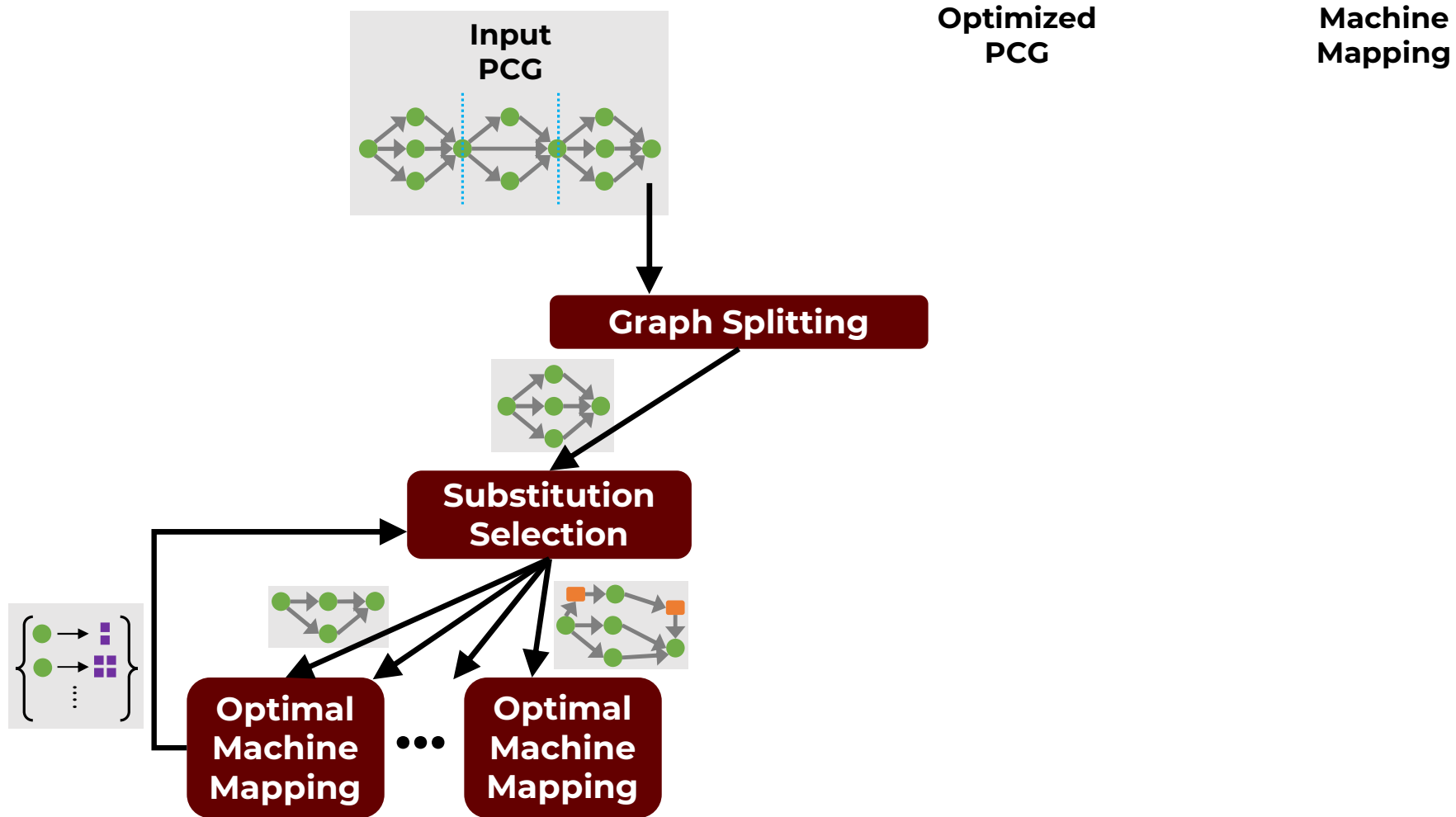
Hierarchical Search Algorithm



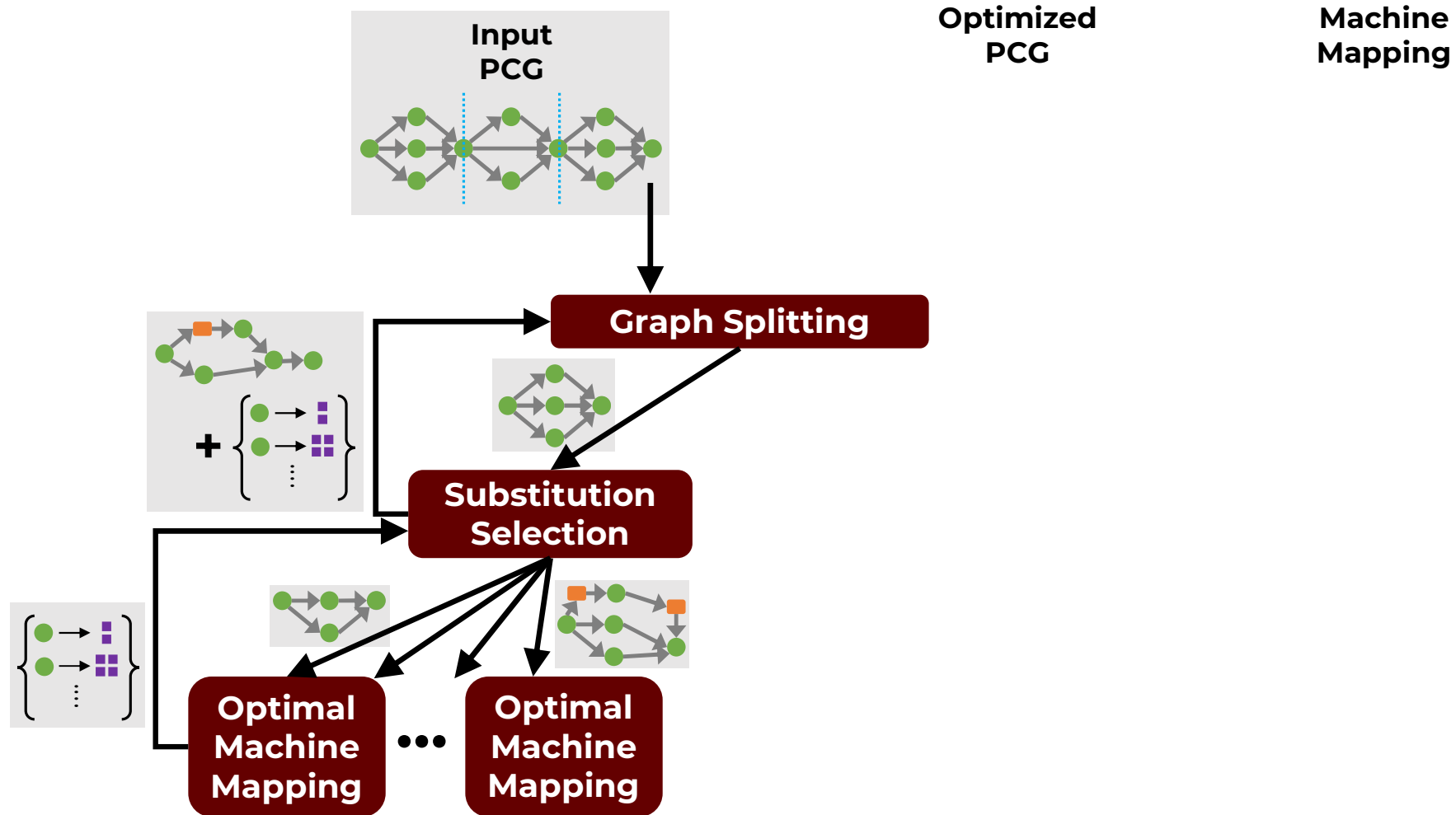
Hierarchical Search Algorithm



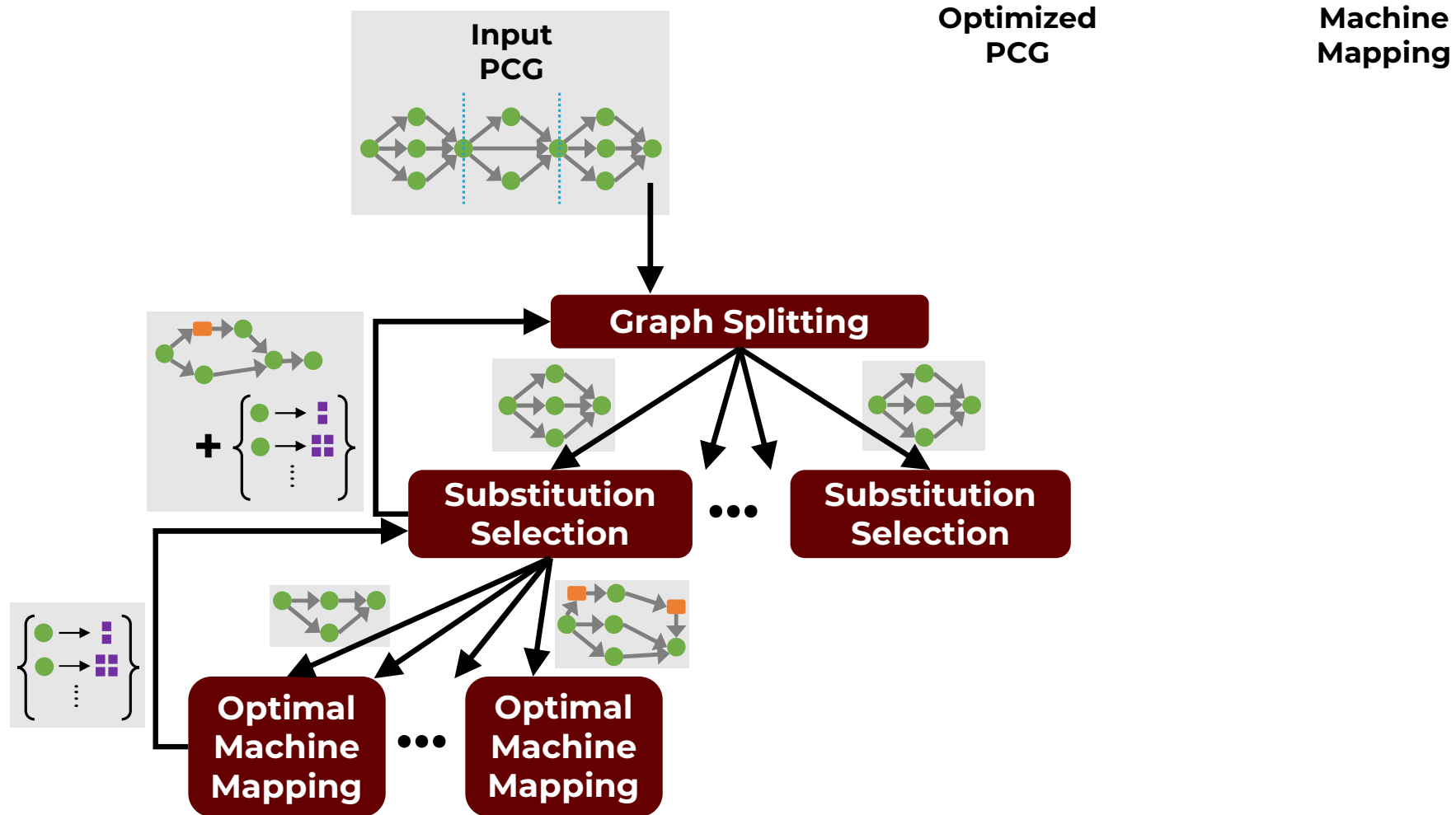
Hierarchical Search Algorithm



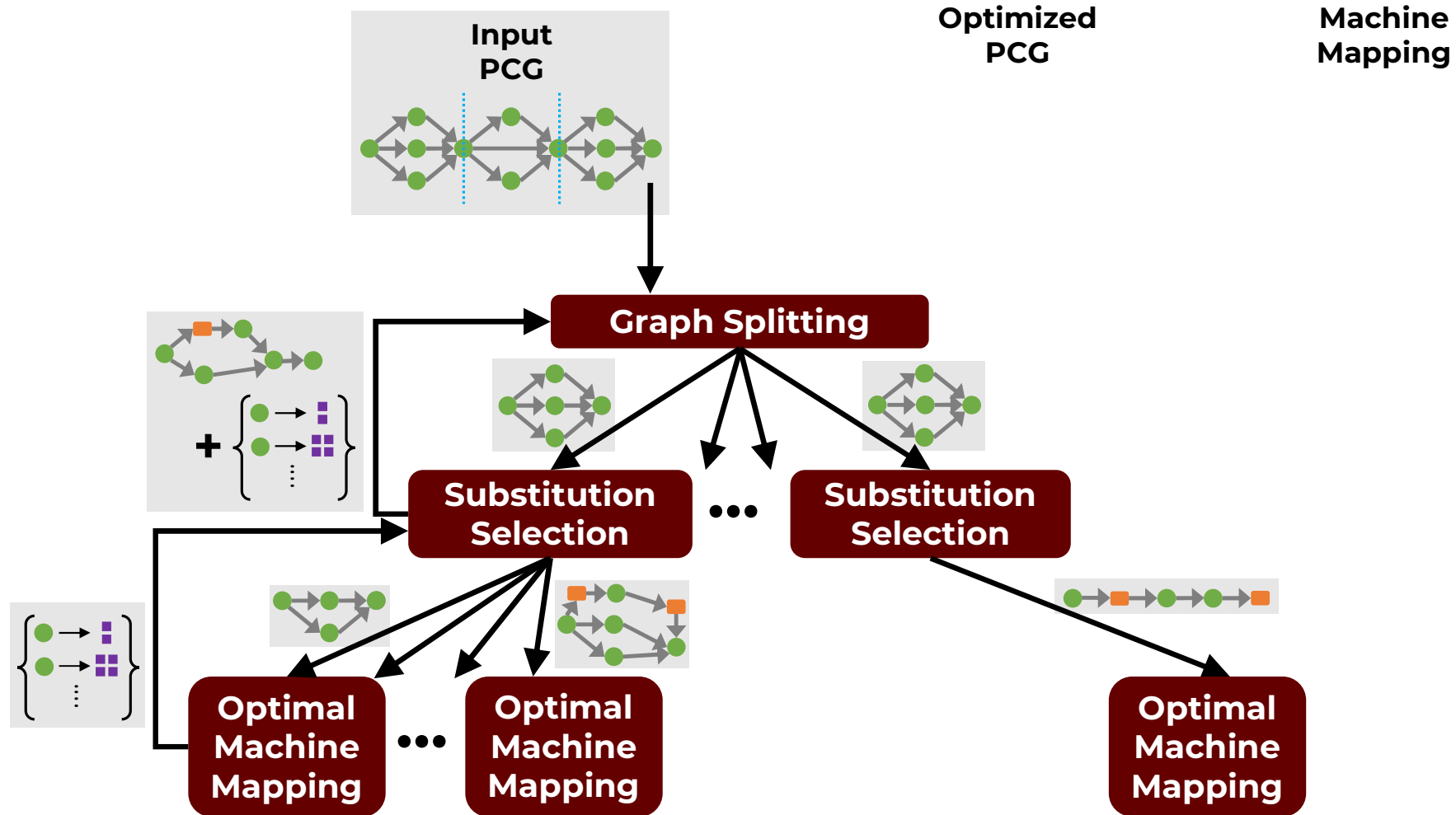
Hierarchical Search Algorithm



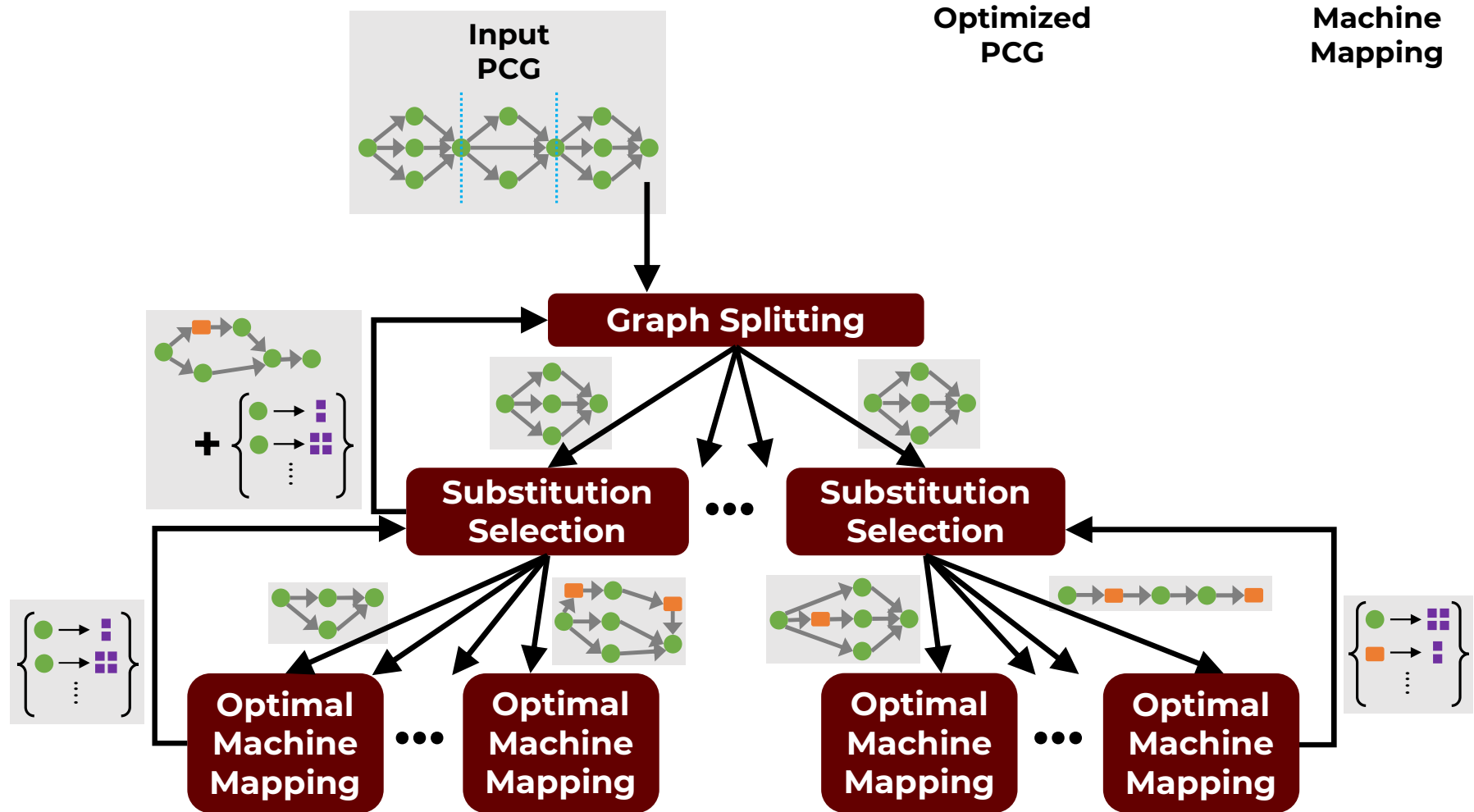
Hierarchical Search Algorithm



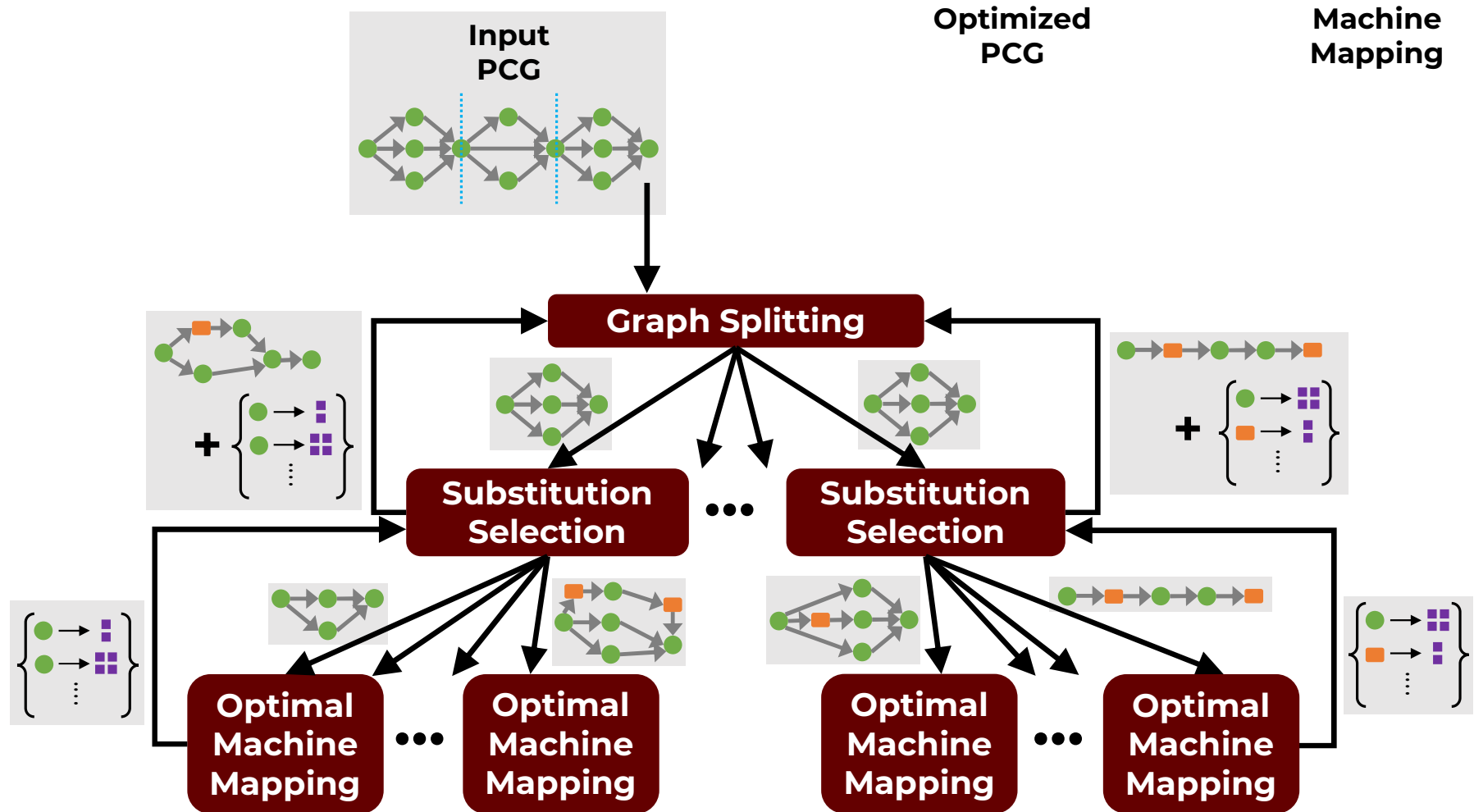
Hierarchical Search Algorithm



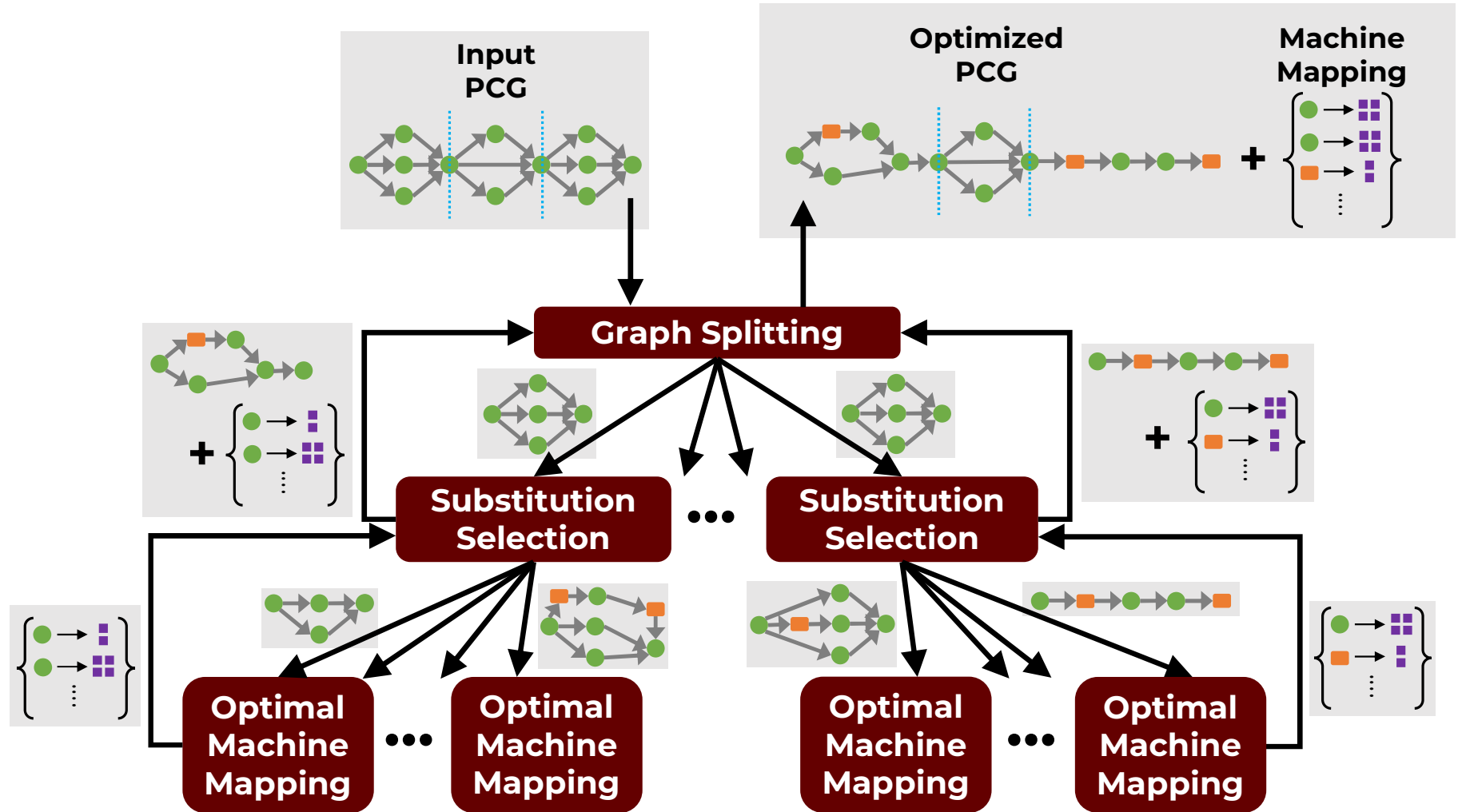
Hierarchical Search Algorithm



Hierarchical Search Algorithm



Hierarchical Search Algorithm



Evaluation

Models

BERT-Large

(Language Modeling)

Candle-UNO

(Precision Medicine)

MLP

(Regression)

DLRM

XDL

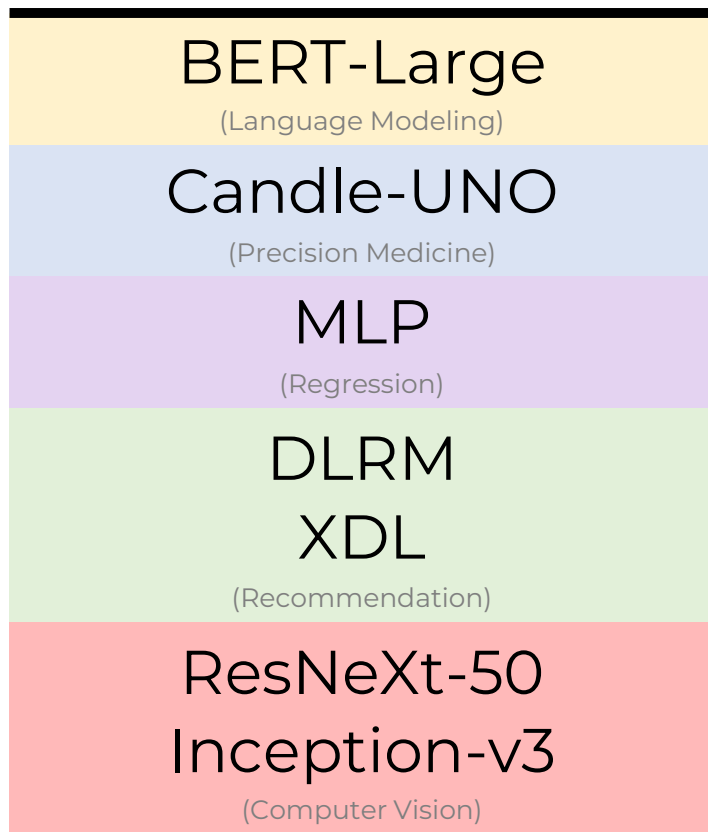
(Recommendation)

ResNeXt-50

Inception-v3

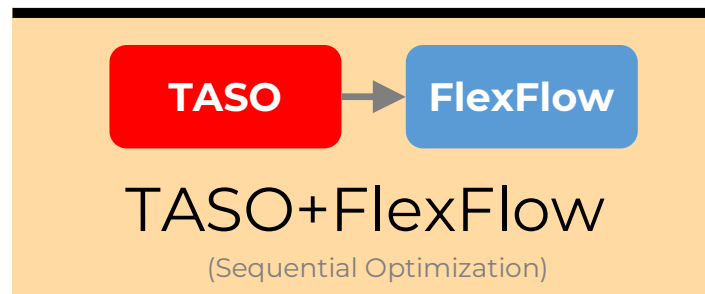
(Computer Vision)

Models

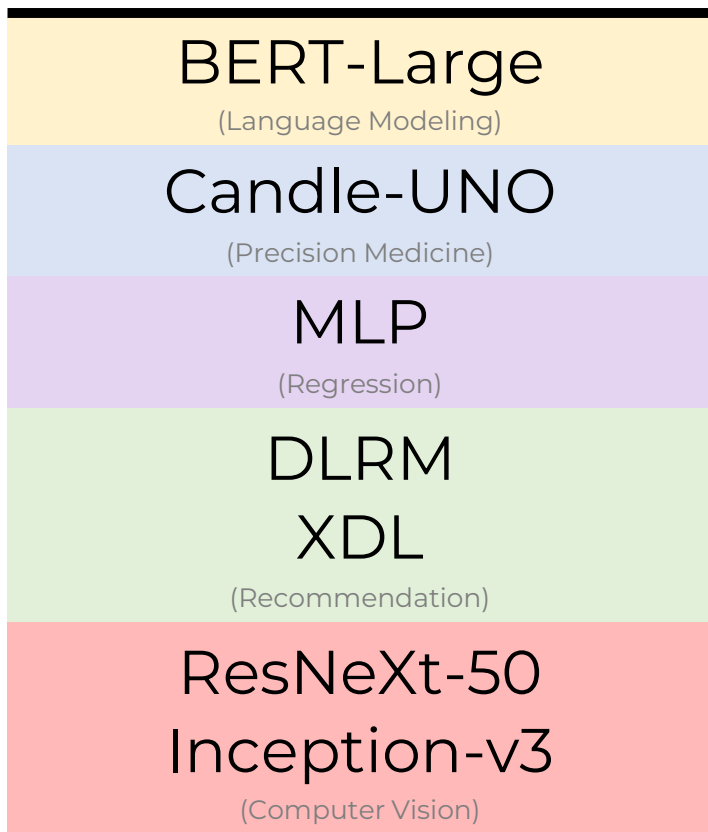


×

Baselines

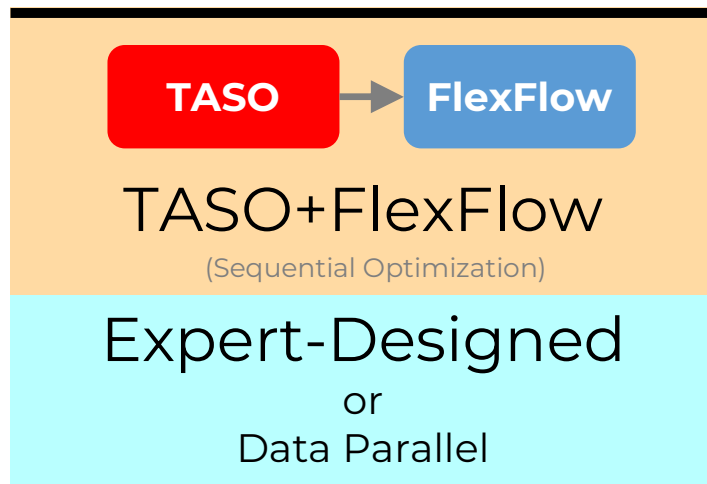


Models

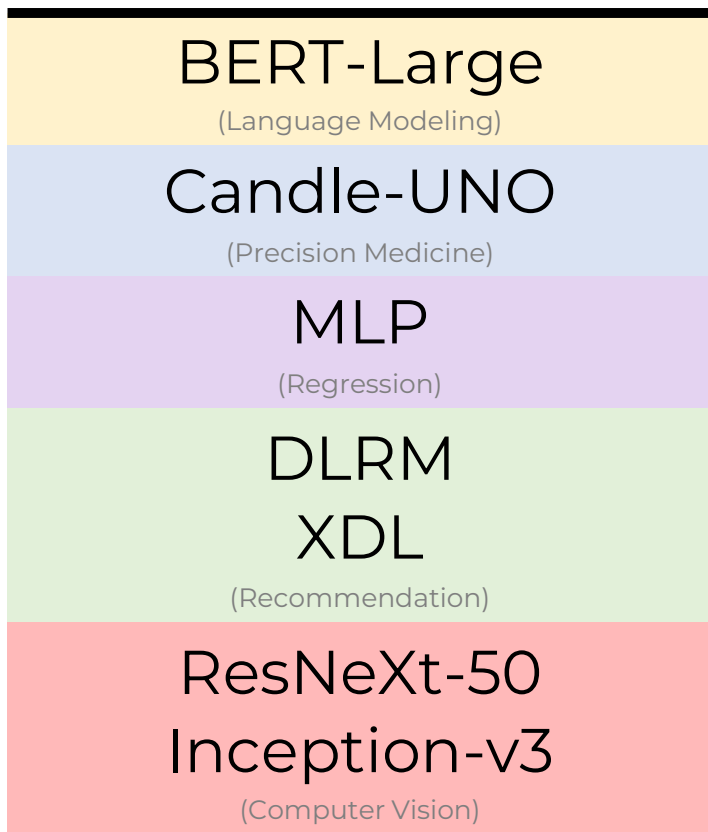


×

Baselines

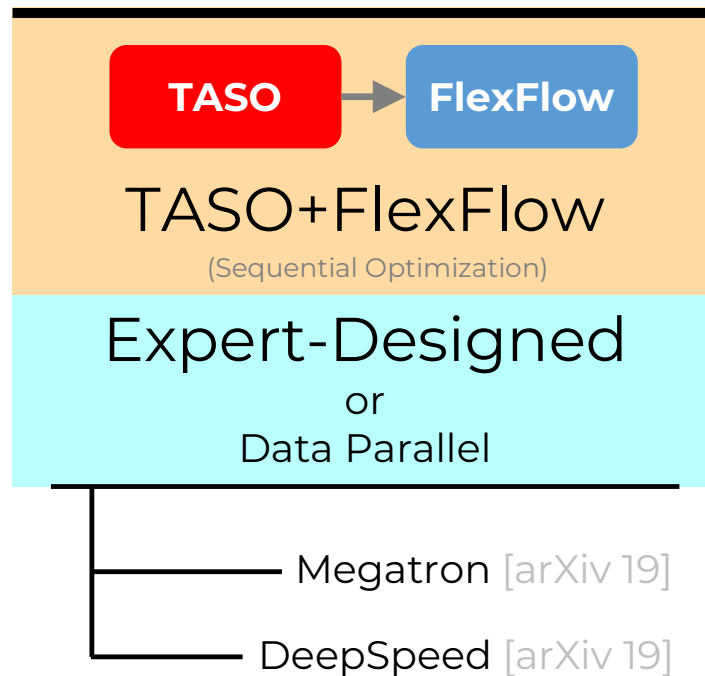


Models

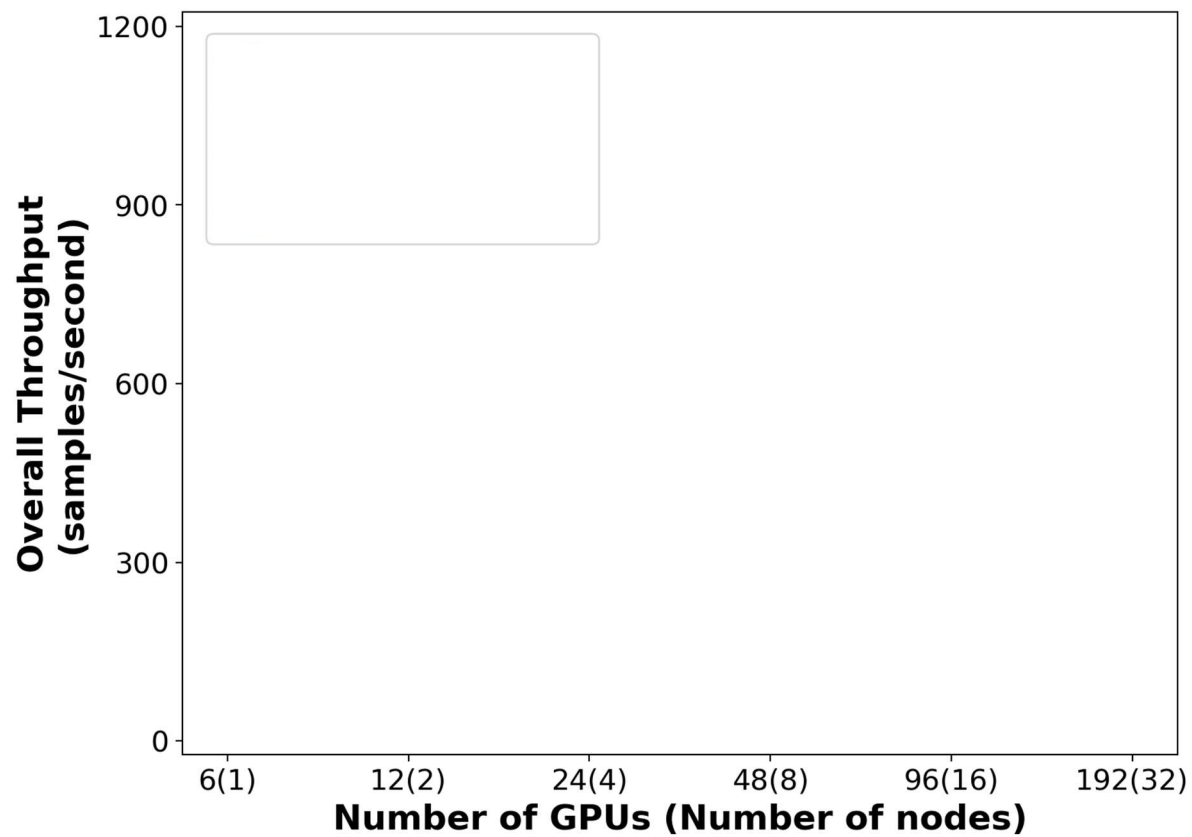


×

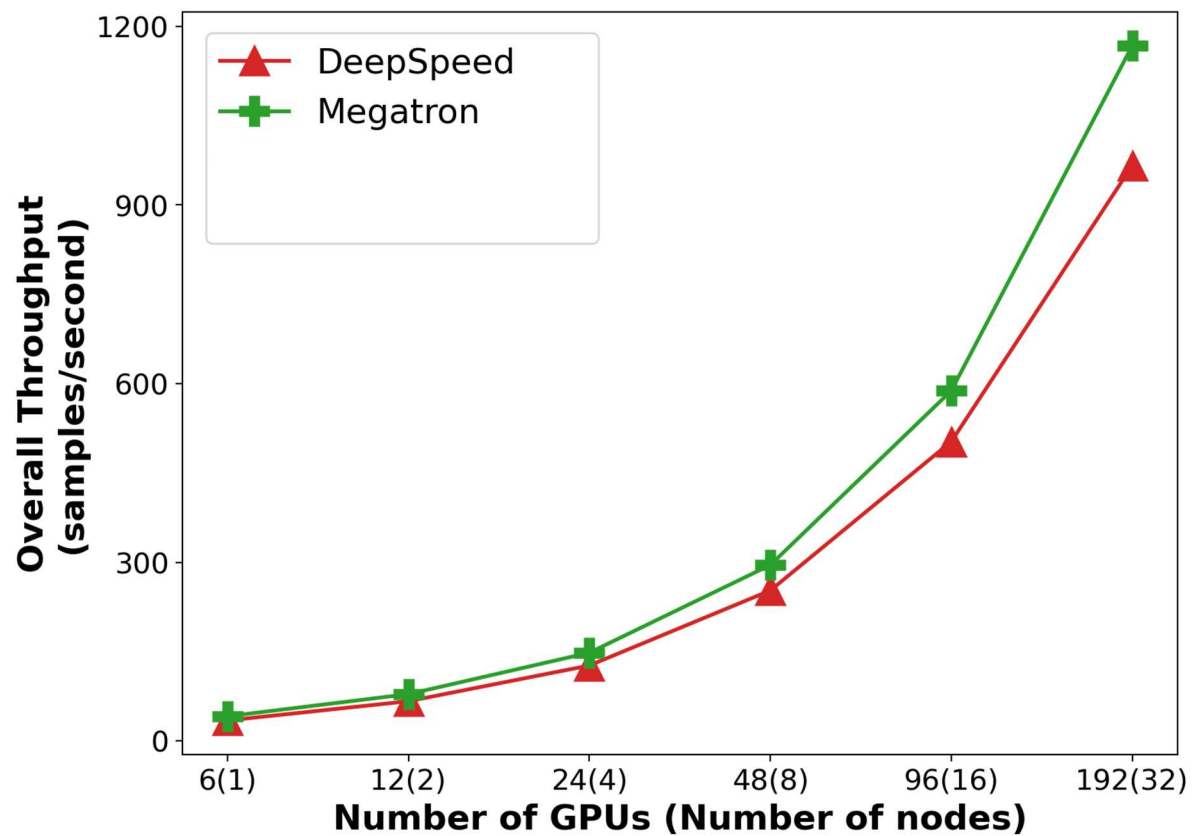
Baselines



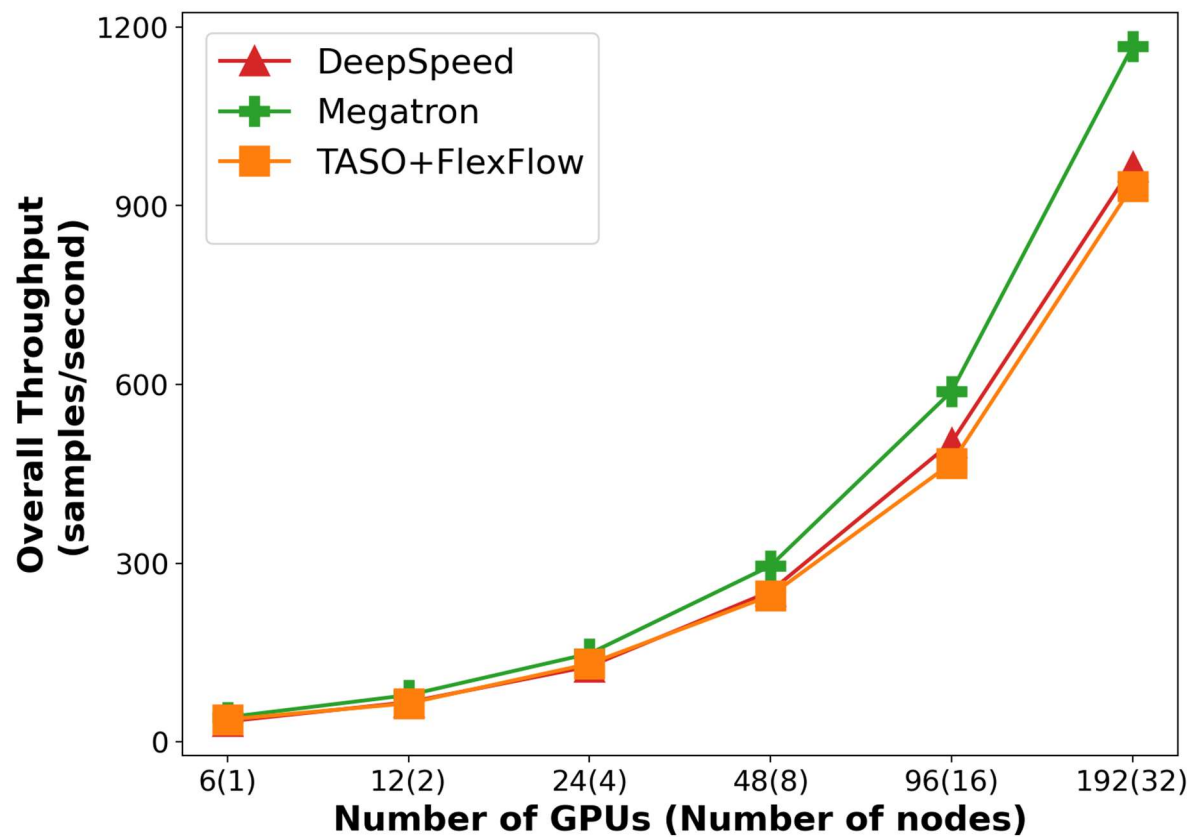
BERT-Large



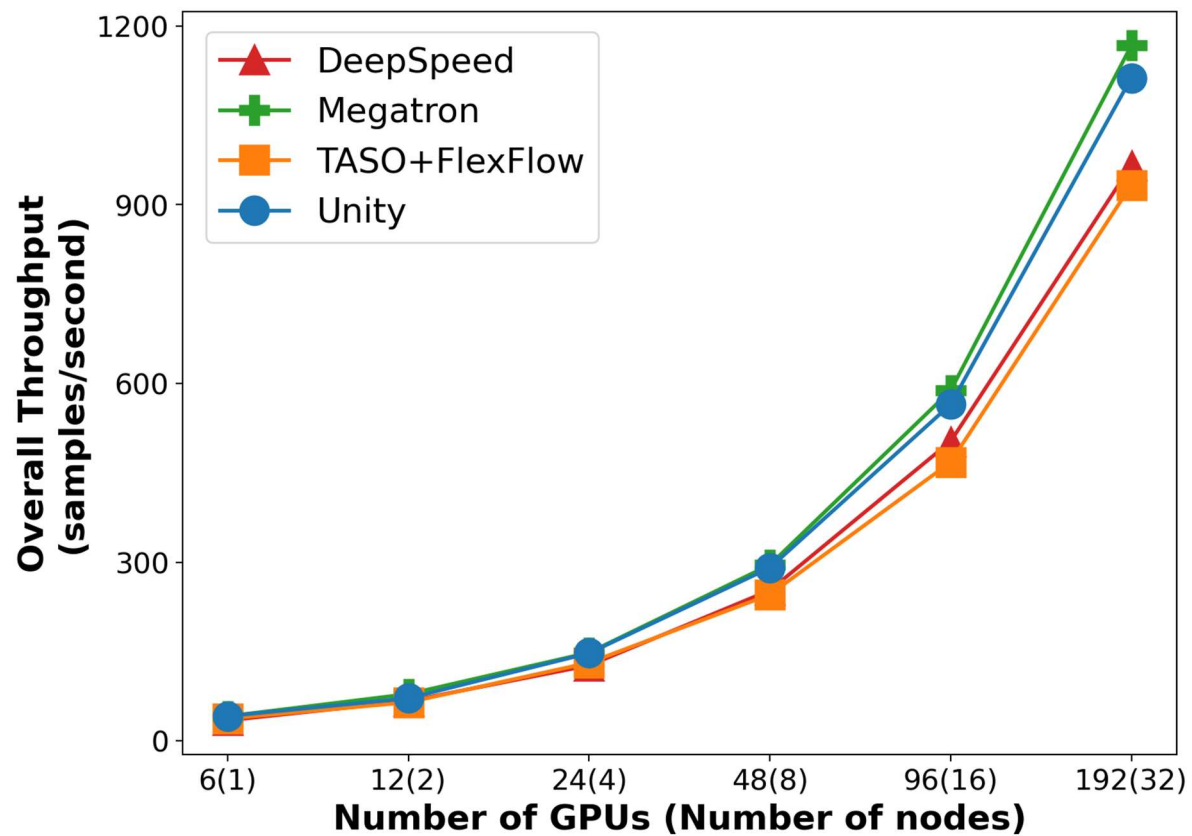
BERT-Large



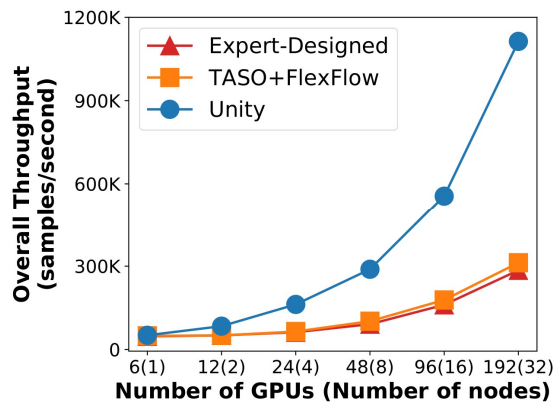
BERT-Large



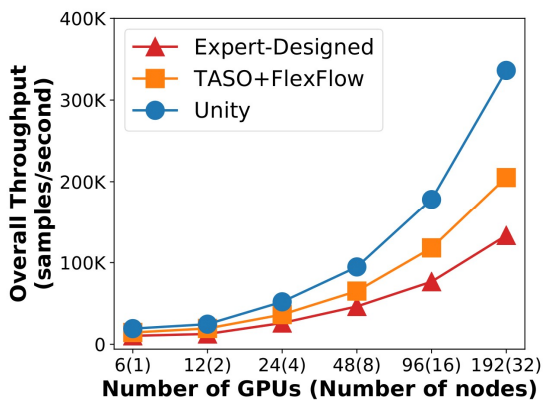
BERT-Large



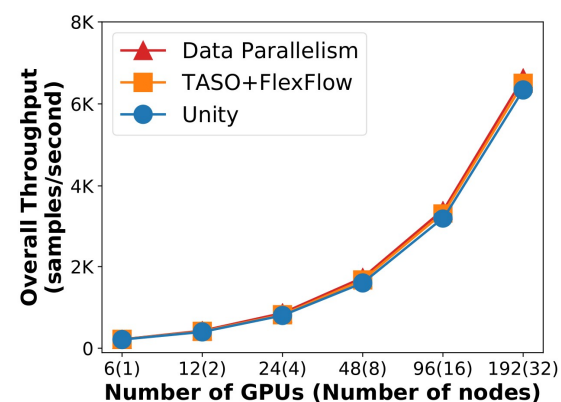
DLRM



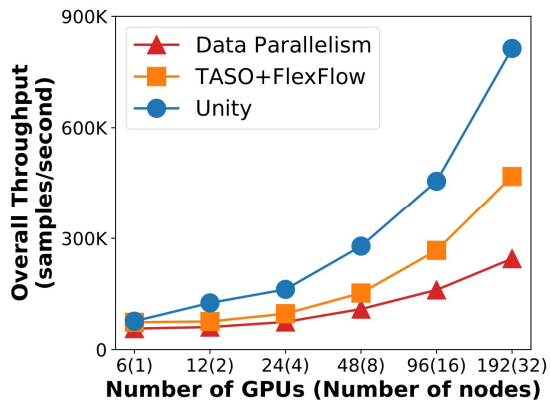
CANDLE-Uno



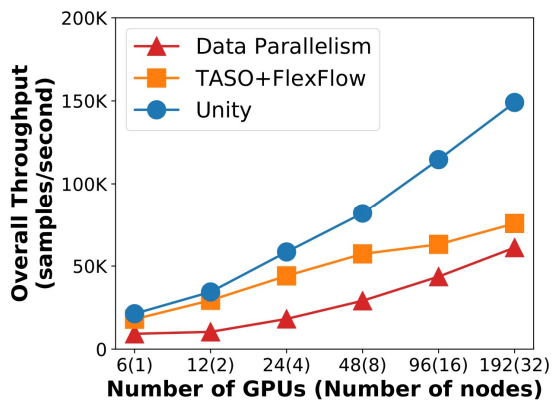
ResNeXt-50



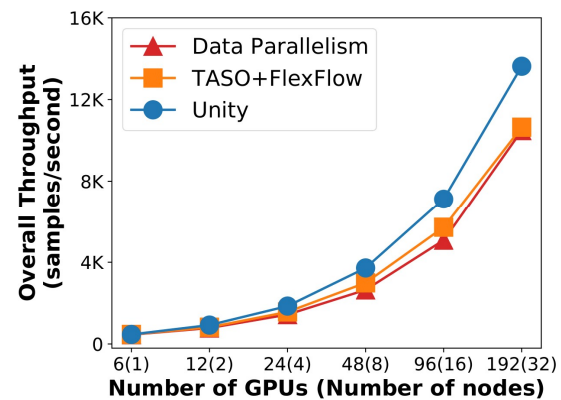
XDL



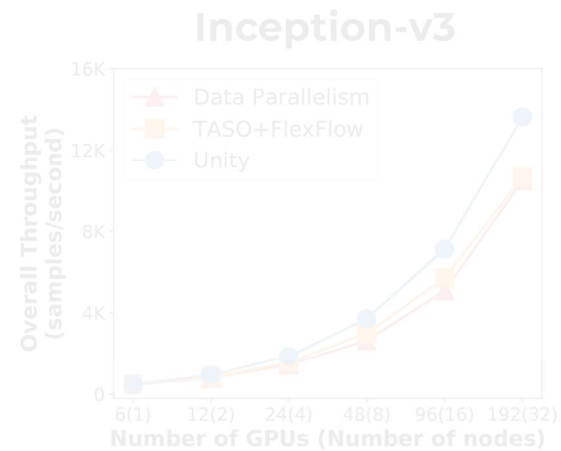
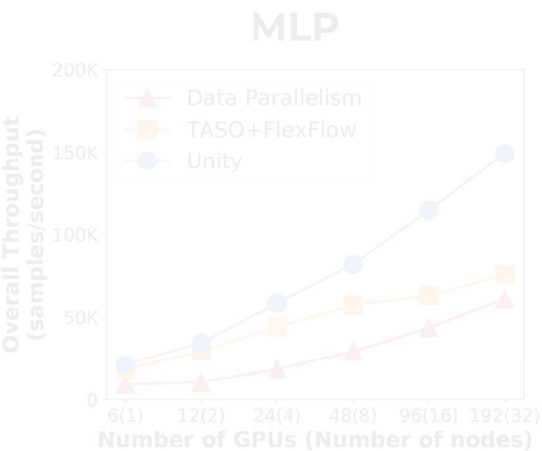
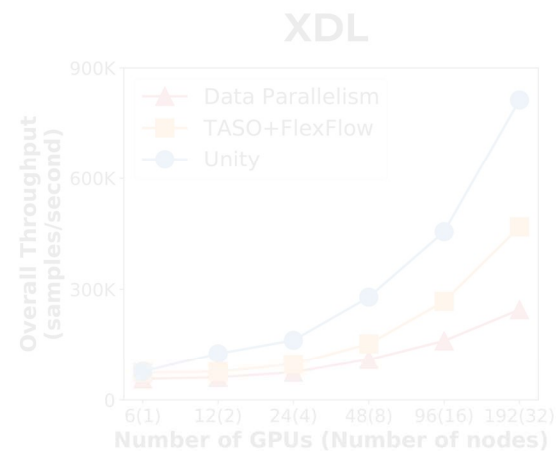
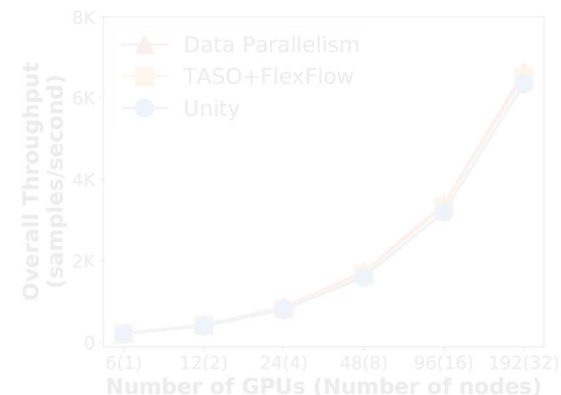
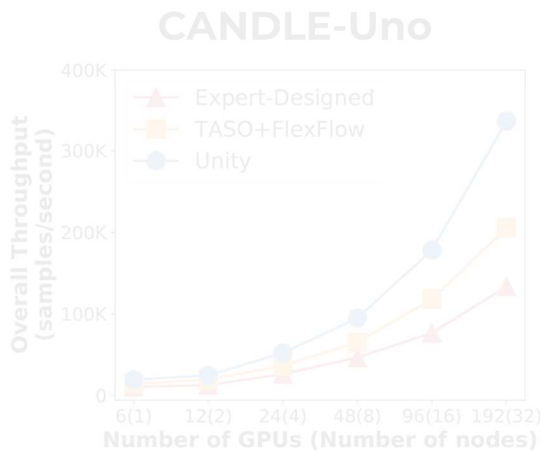
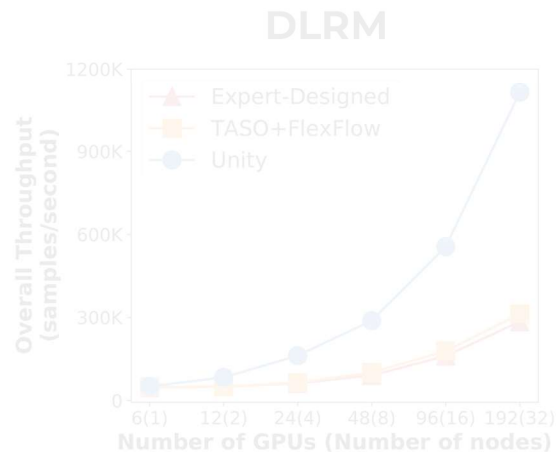
MLP



Inception-v3

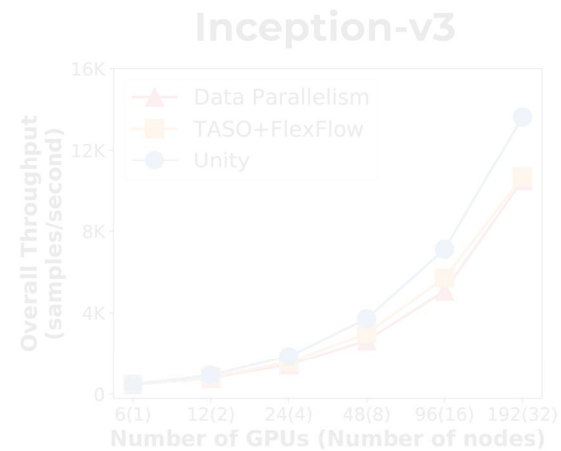
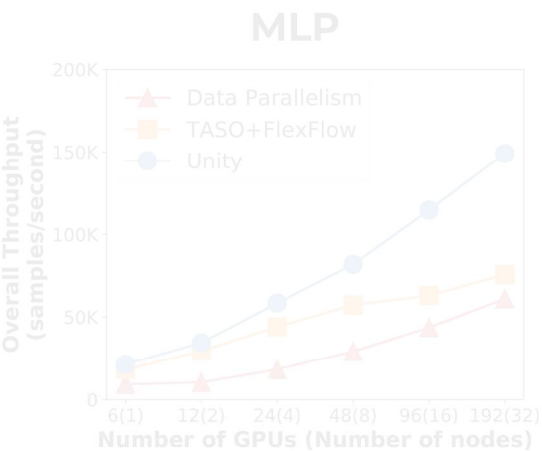
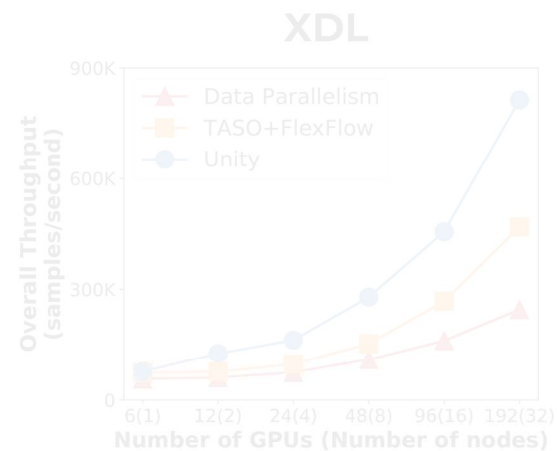
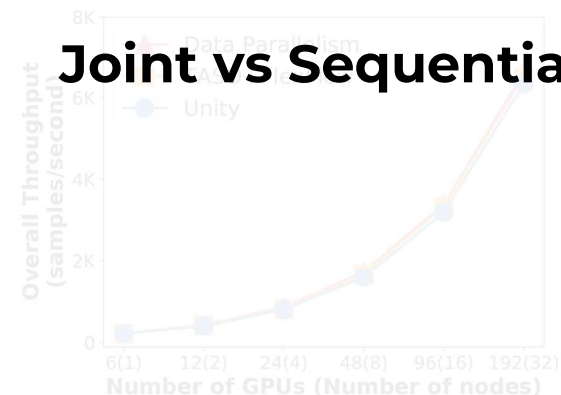
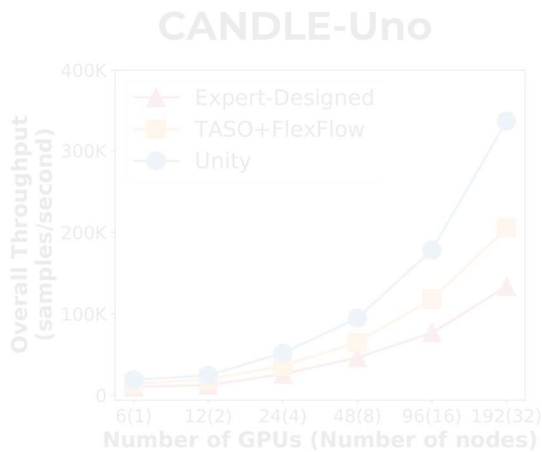
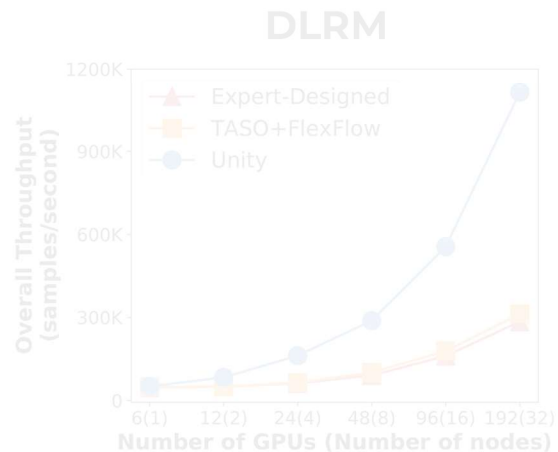


Forms of Parallelism



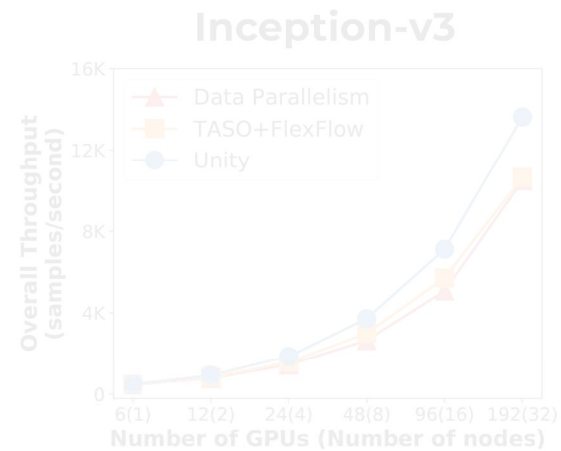
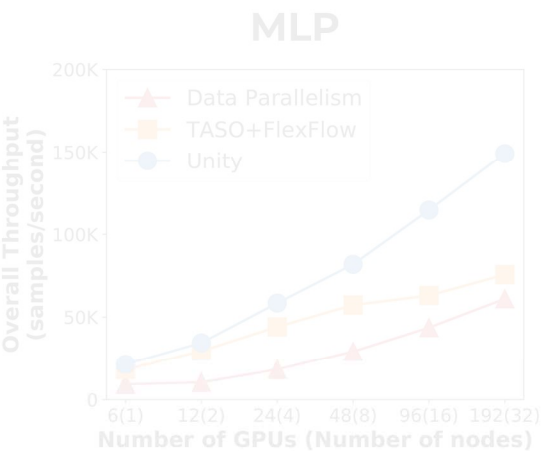
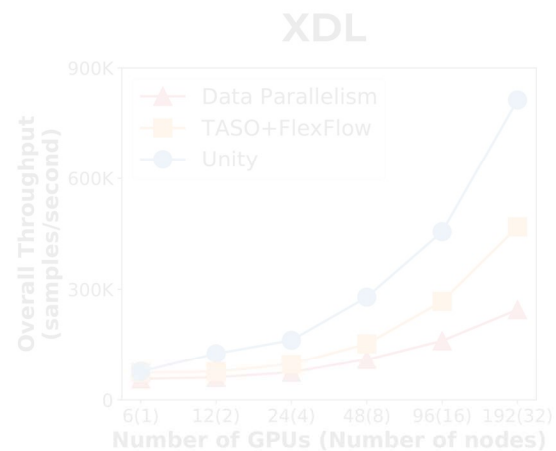
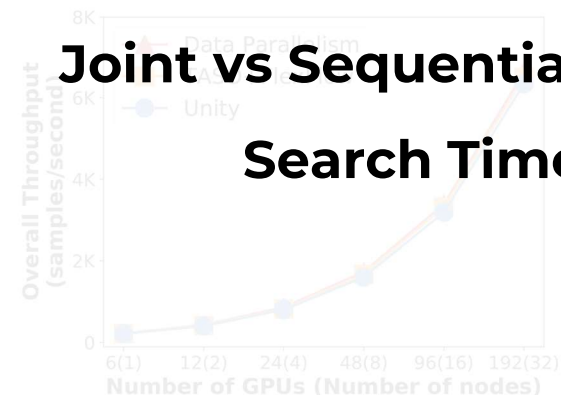
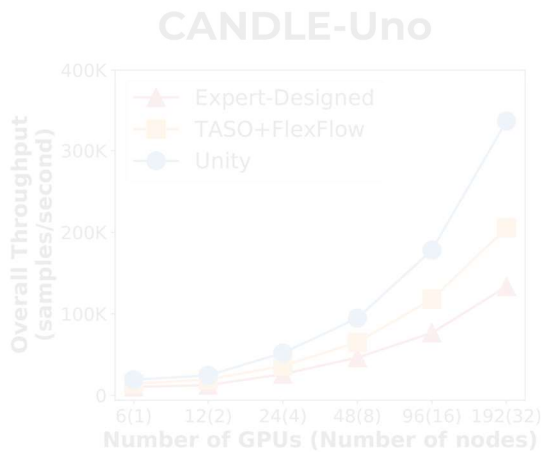
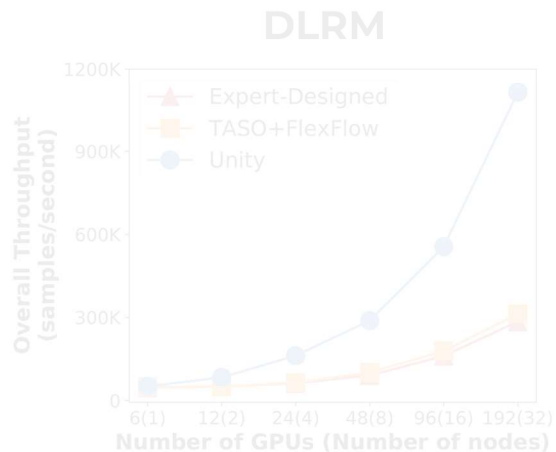
Forms of Parallelism

Joint vs Sequential



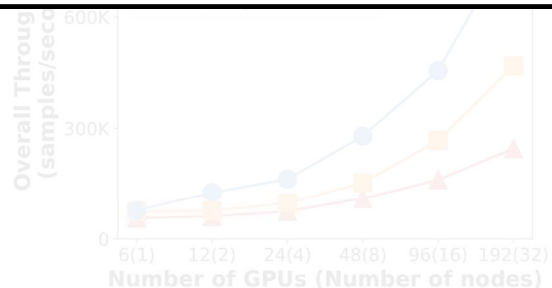
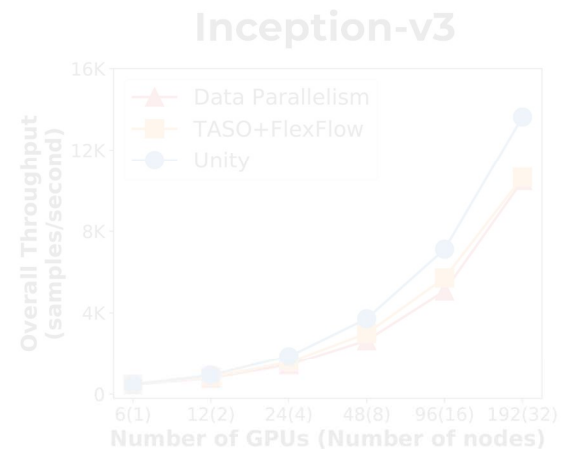
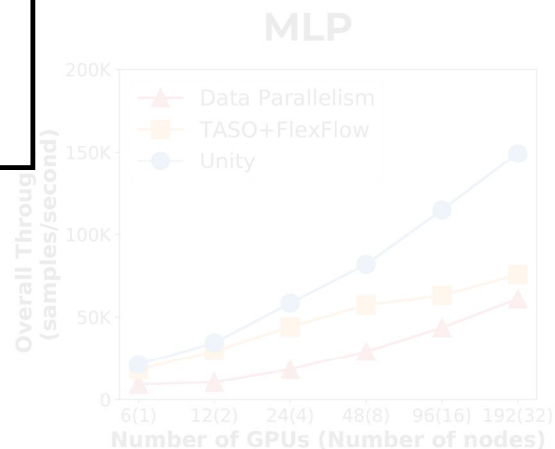
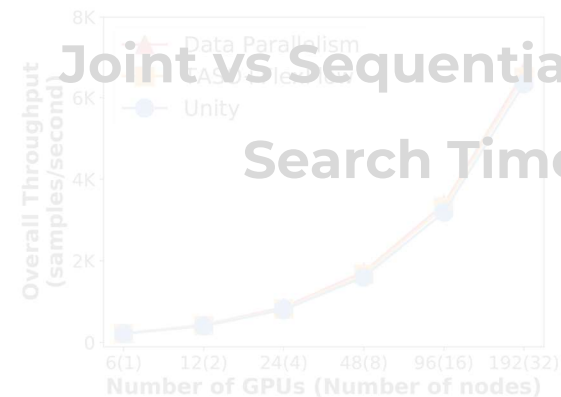
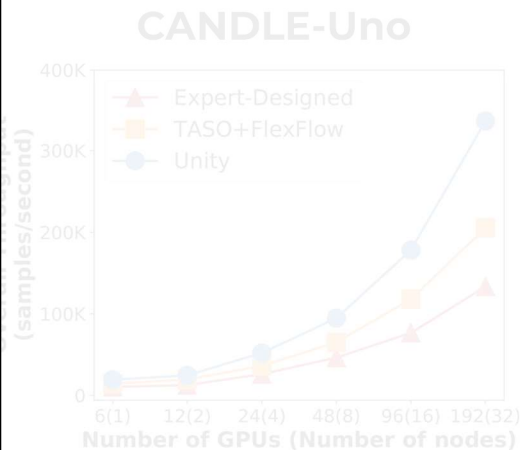
Forms of Parallelism

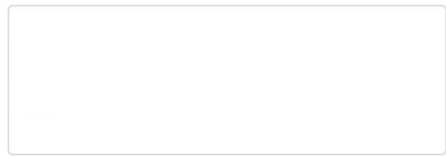
Joint vs Sequential Search Time



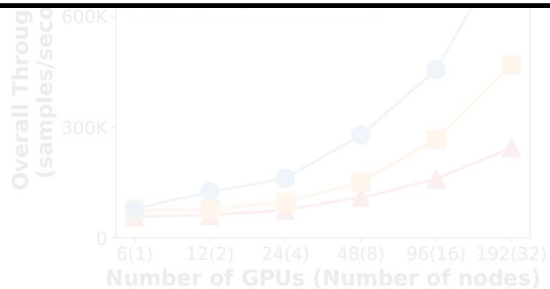
Forms of Parallelism

Joint vs Sequential Search Time

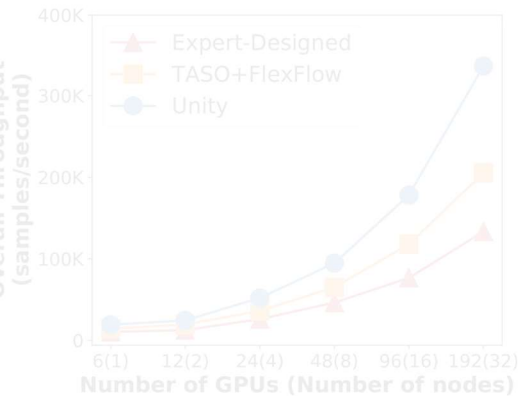




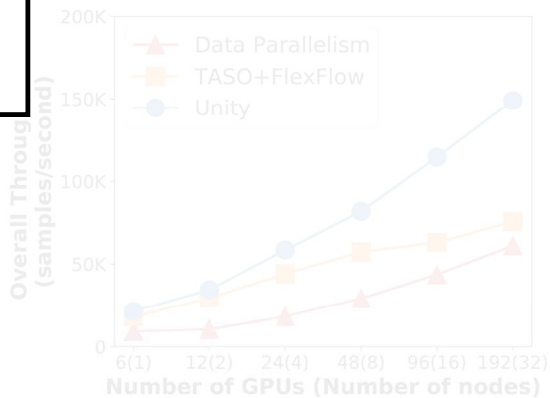
BERT-Large



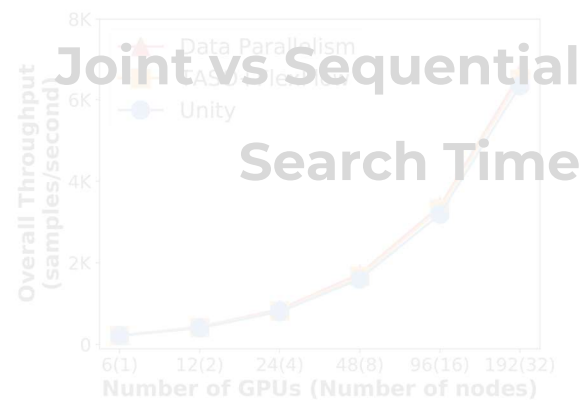
CANDLE-Uno



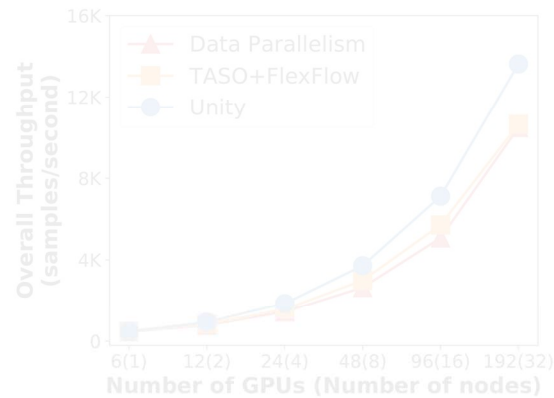
MLP

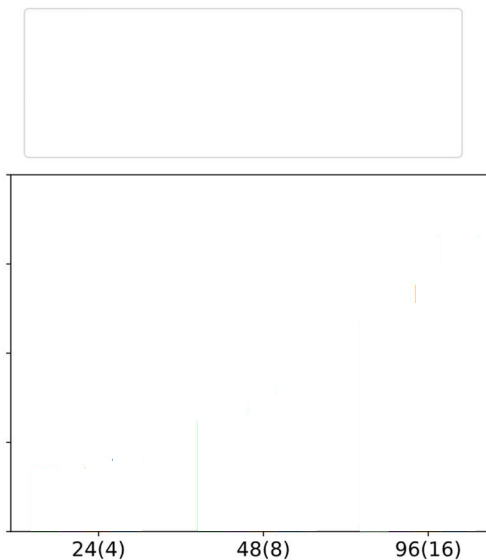


Forms of Parallelism



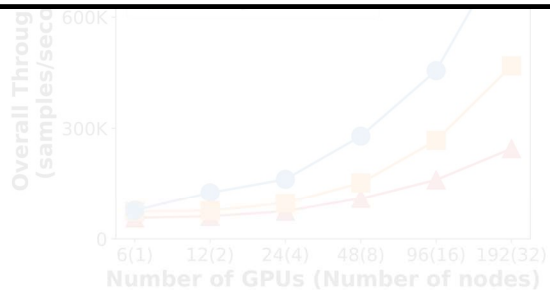
Inception-v3



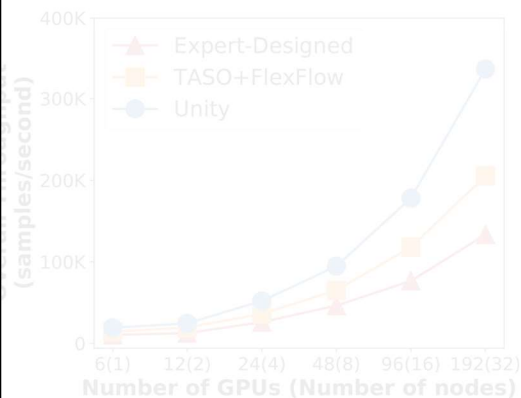


Number of GPUs (Number of nodes)

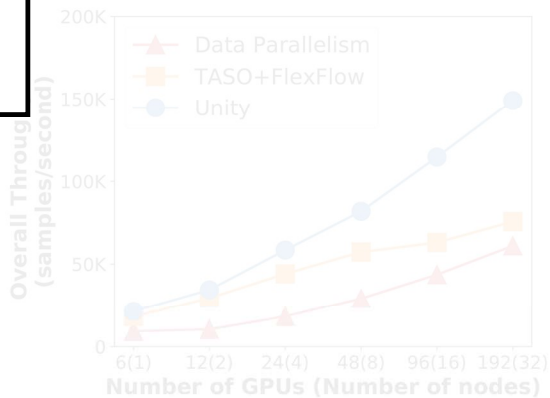
BERT-Large



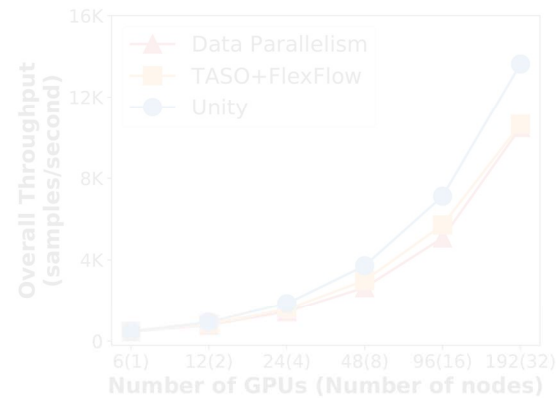
CANDLE-Uno



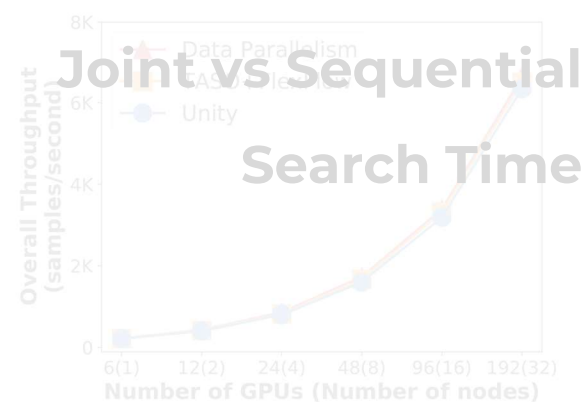
MLP

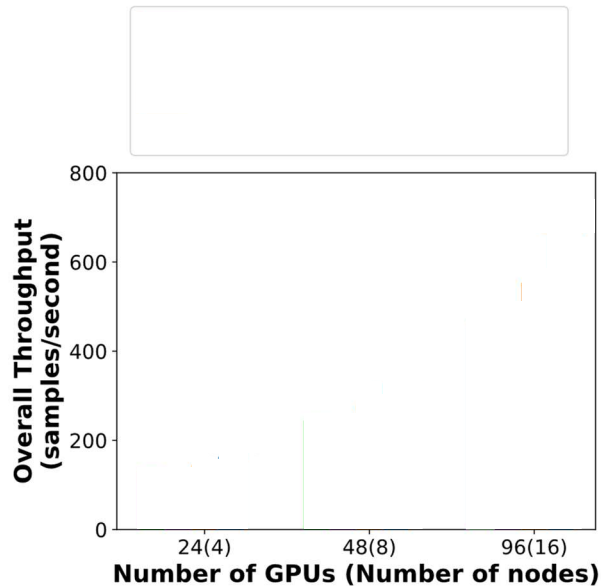


Inception-v3

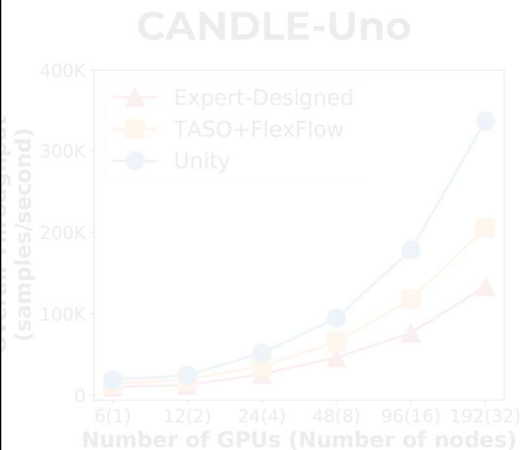


Forms of Parallelism

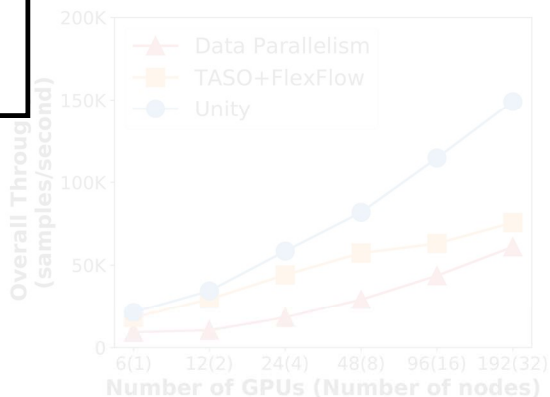




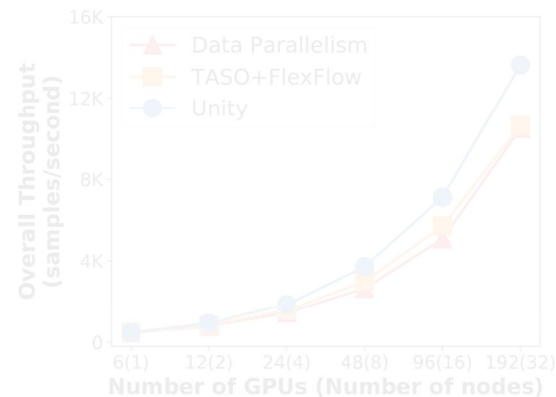
BERT-Large



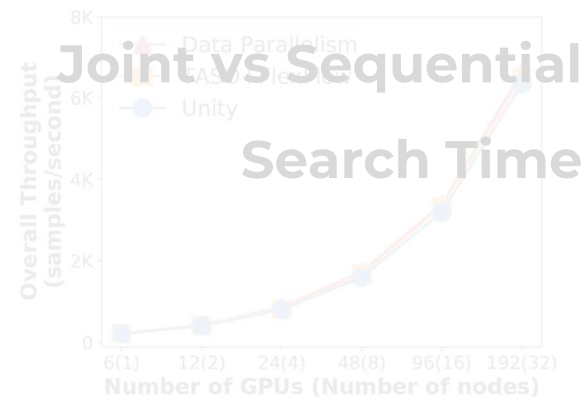
MLP



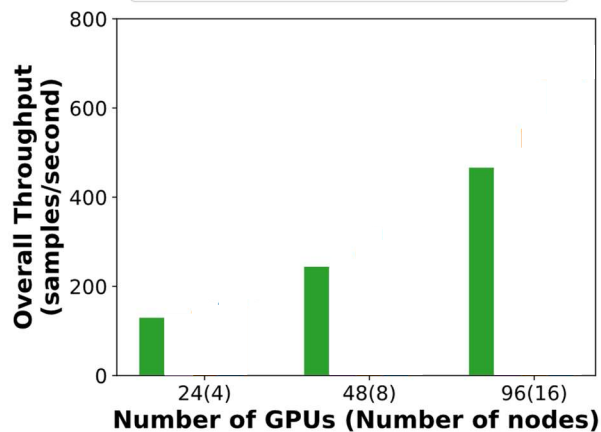
Inception-v3



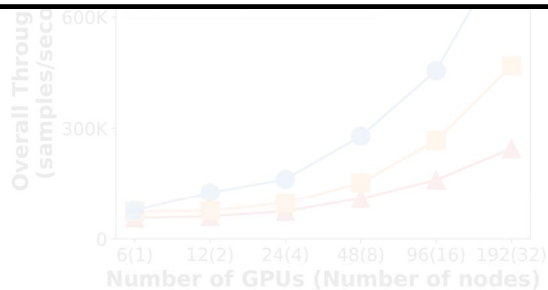
Forms of Parallelism



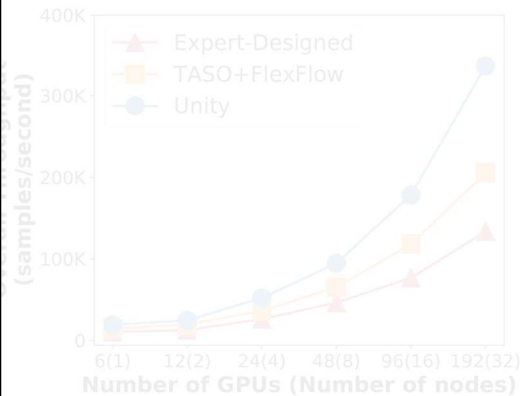
■ Data & Model Parallelism



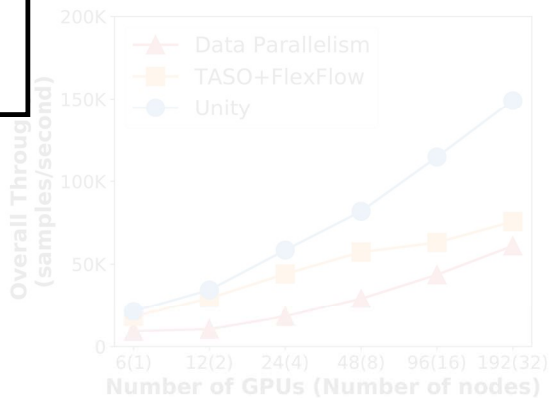
BERT-Large



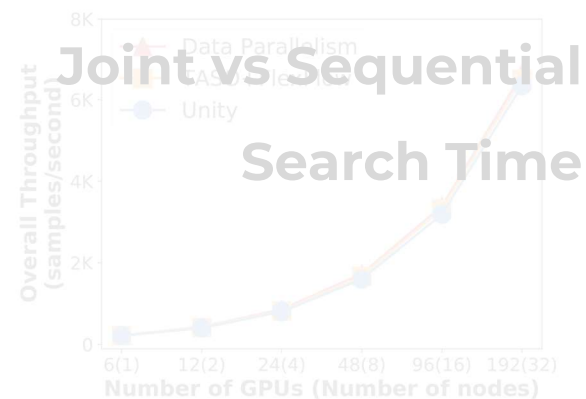
CANDLE-Uono



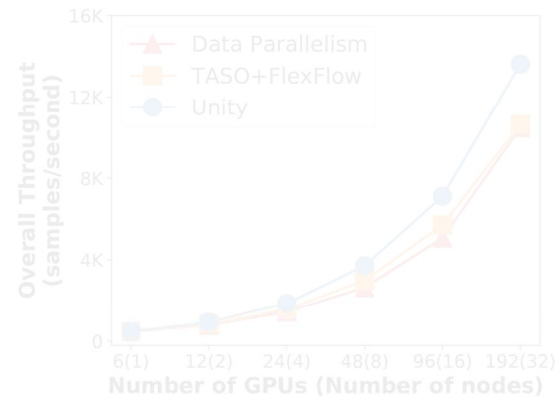
MLP

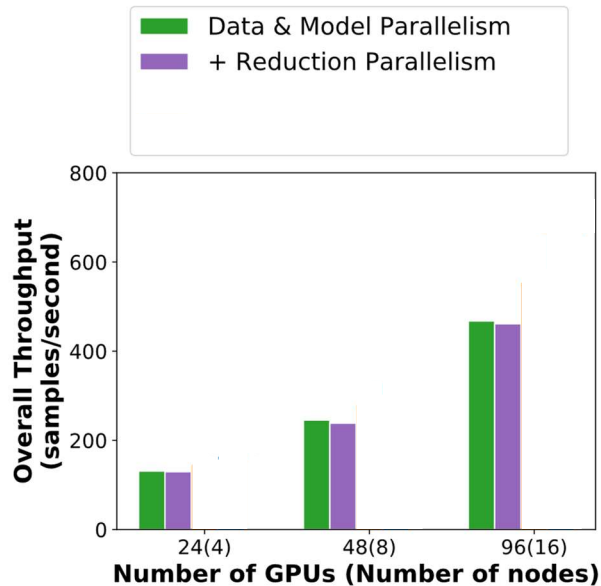


Forms of Parallelism

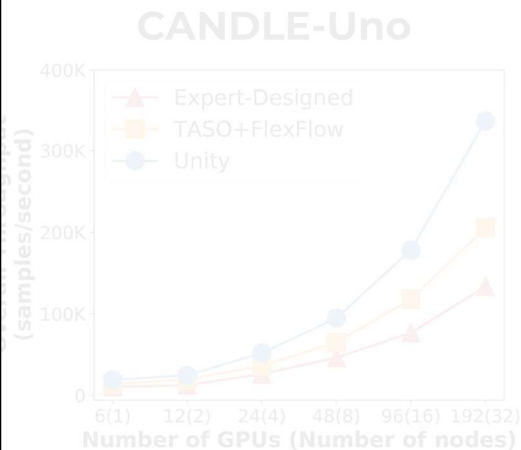


Inception-v3

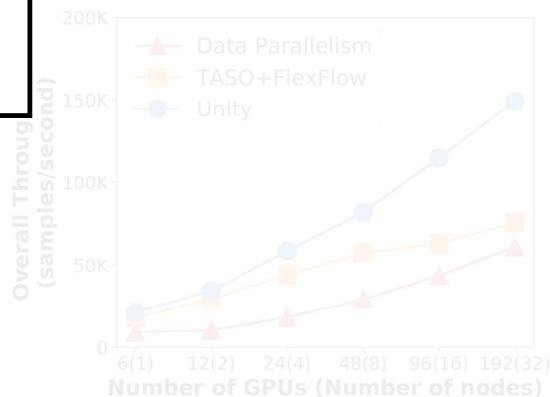




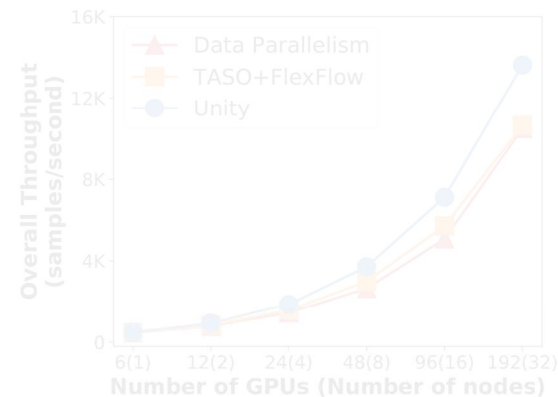
BERT-Large



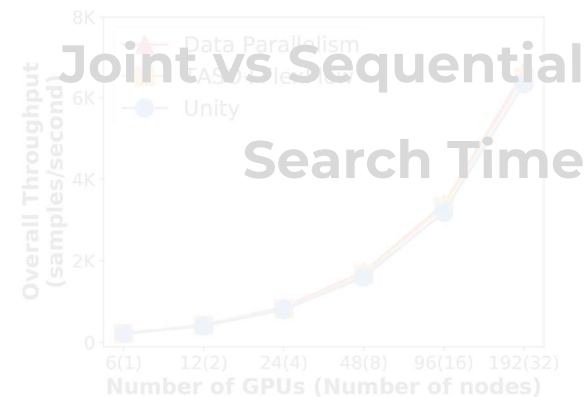
MLP



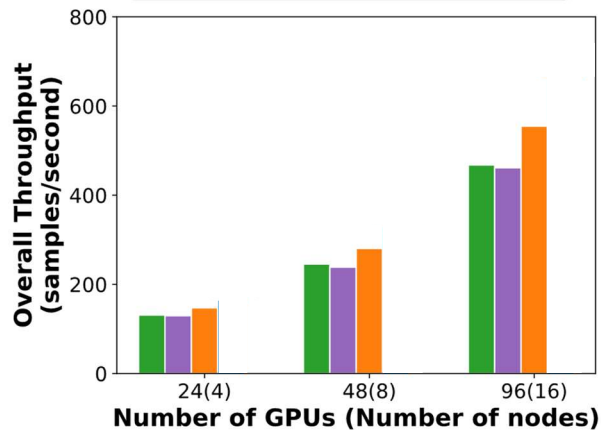
Inception-v3



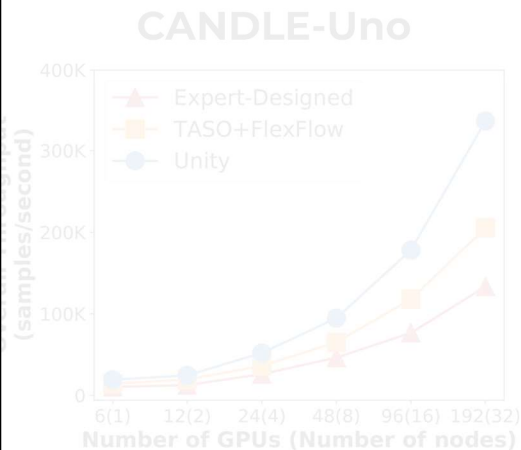
Forms of Parallelism



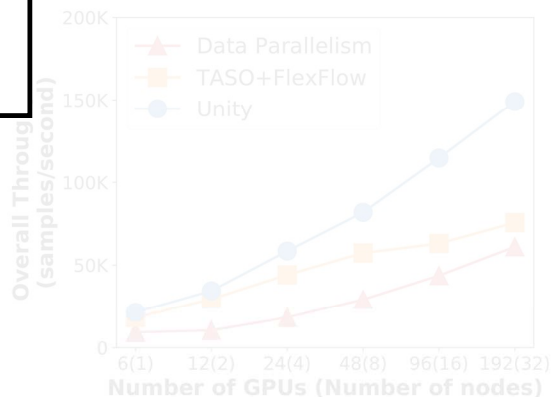
- Data & Model Parallelism
- + Reduction Parallelism
- + Attention-Head Parallelism



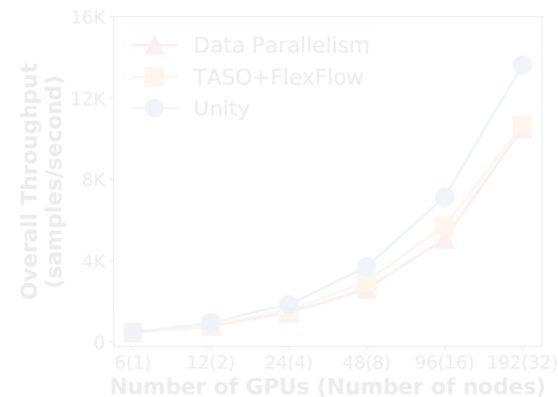
BERT-Large



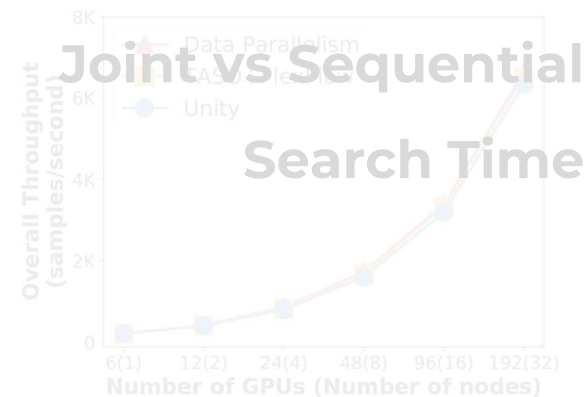
MLP

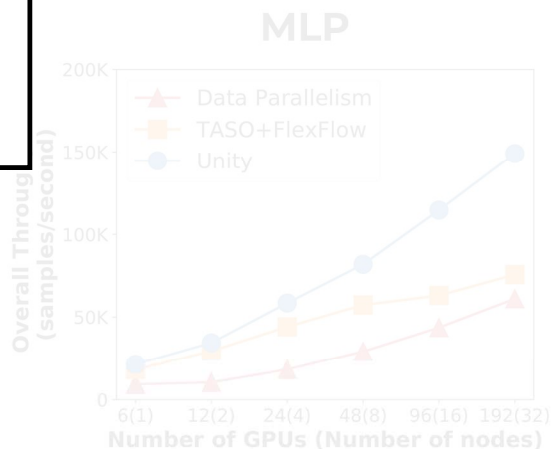
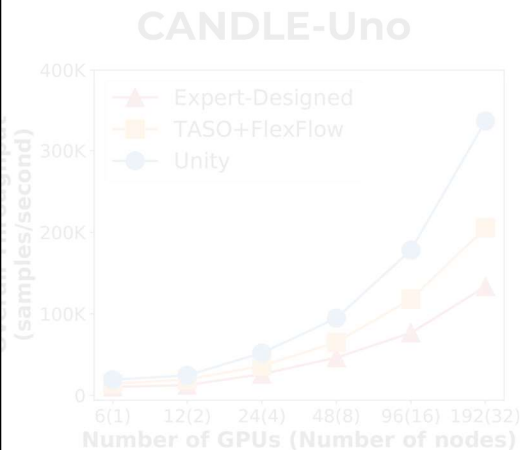
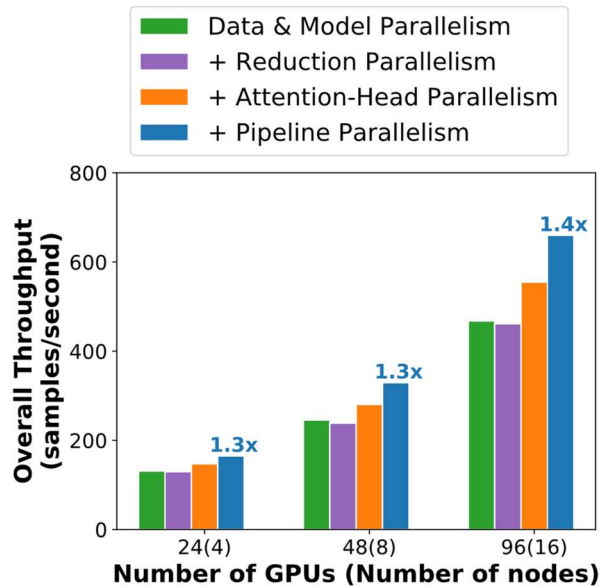


Inception-v3

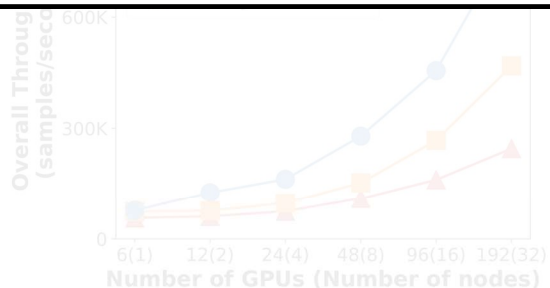
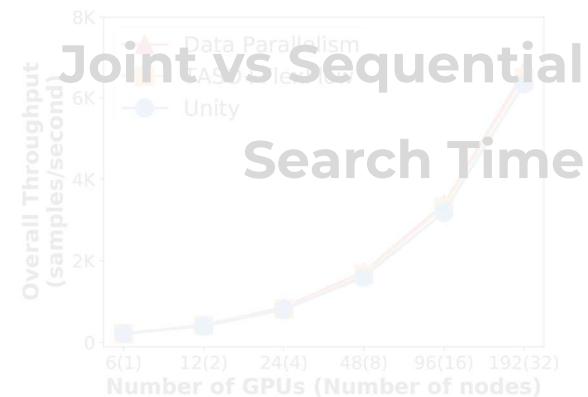


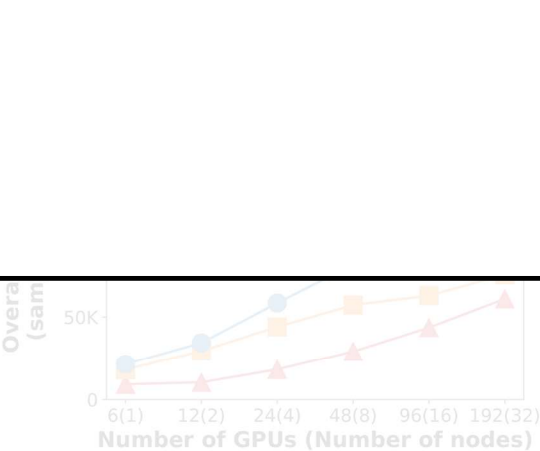
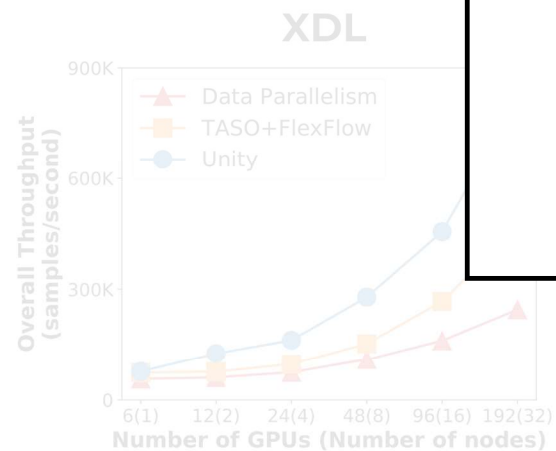
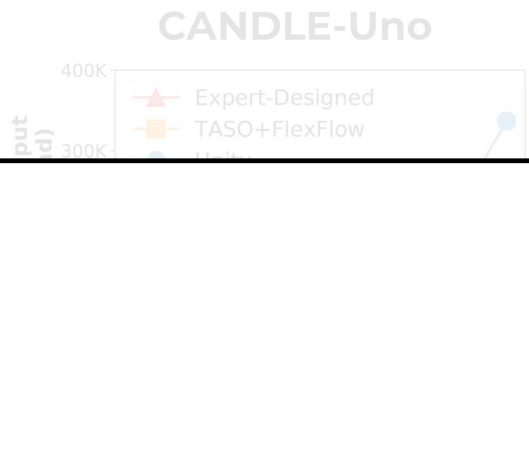
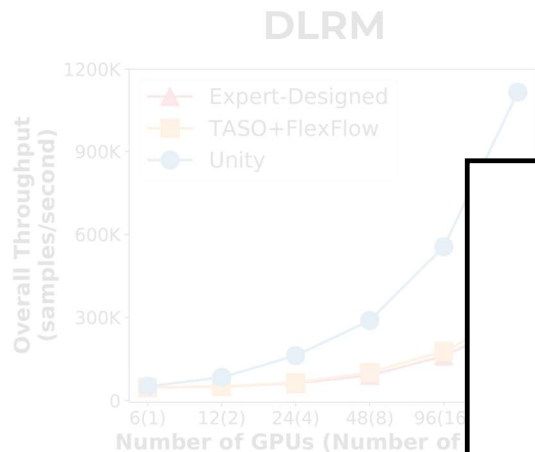
Forms of Parallelism



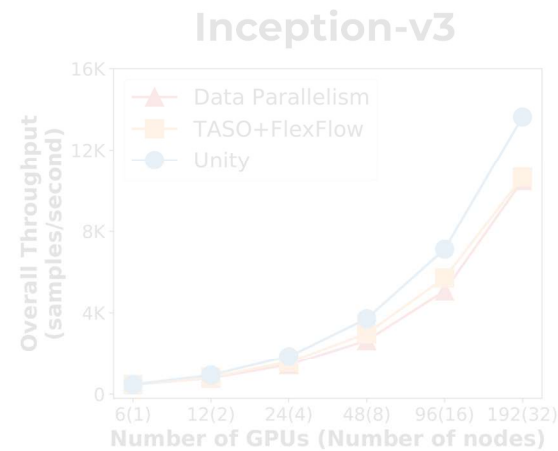
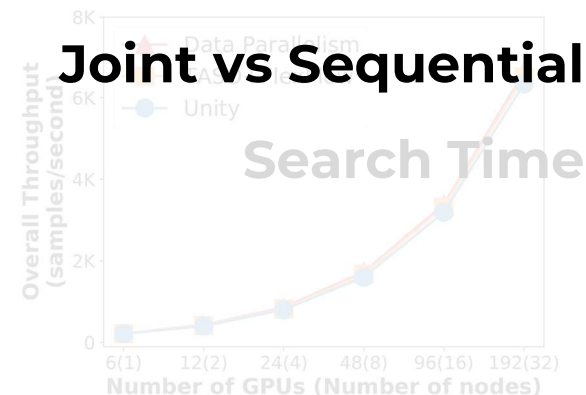


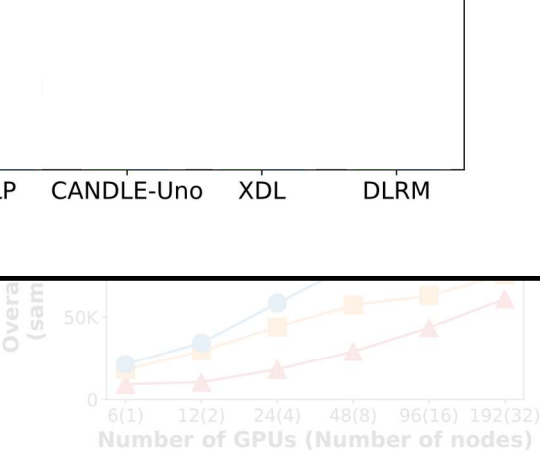
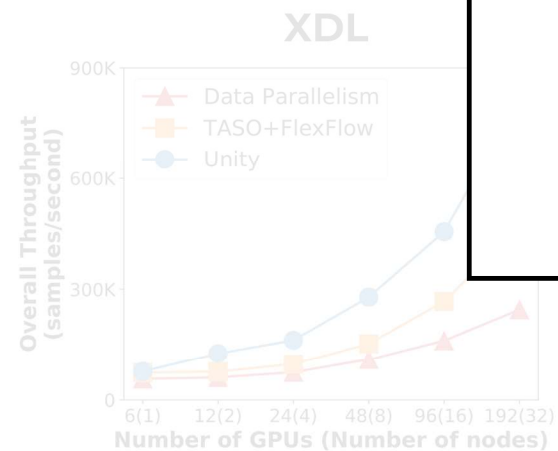
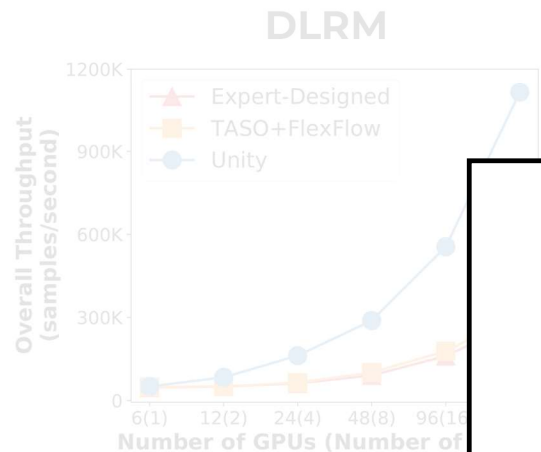
Forms of Parallelism



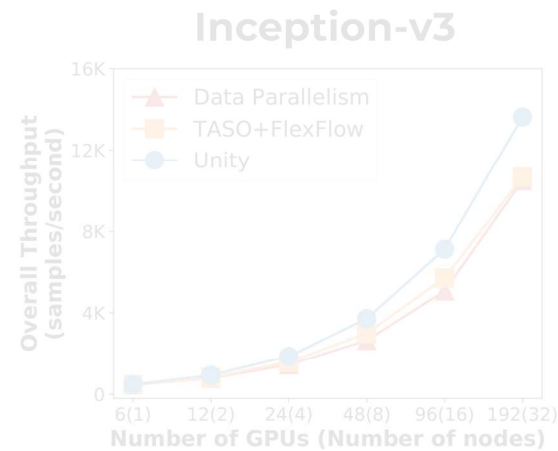
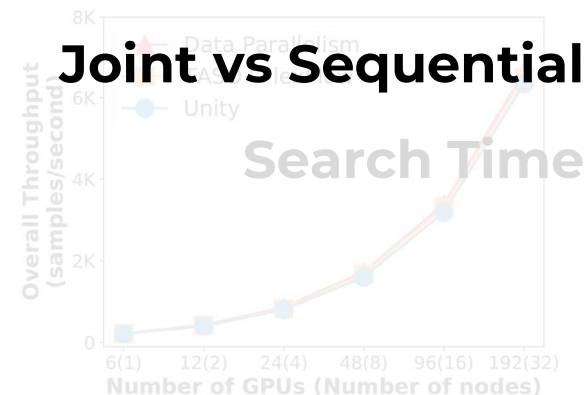


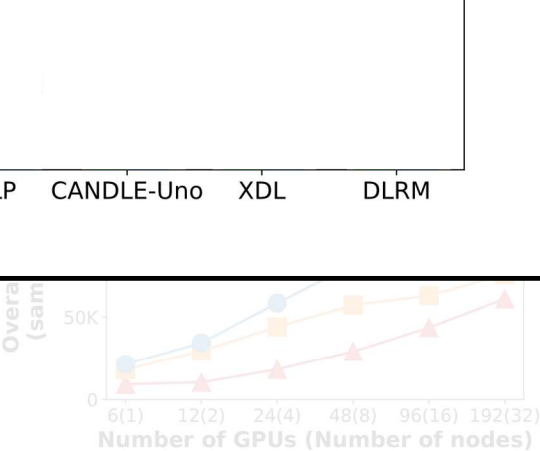
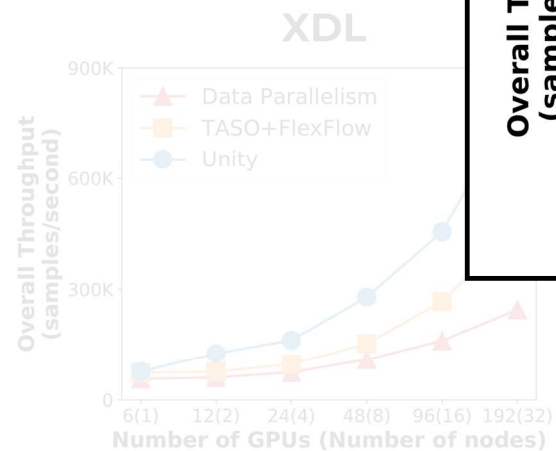
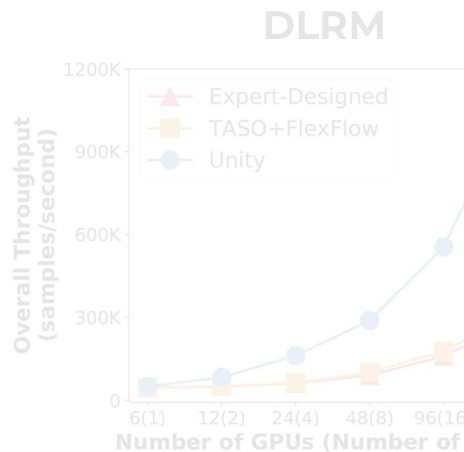
Forms of Parallelism



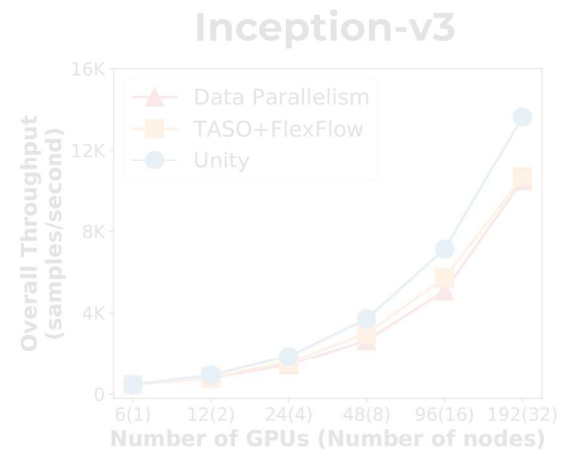
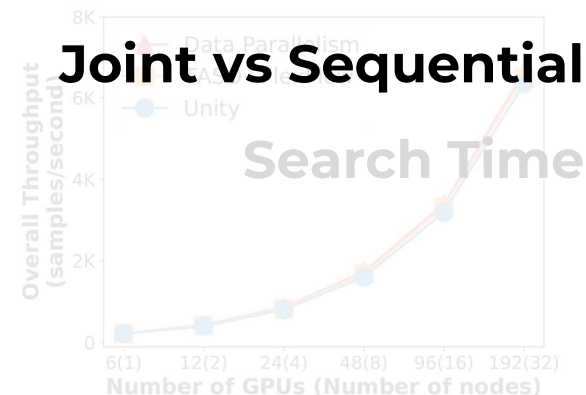


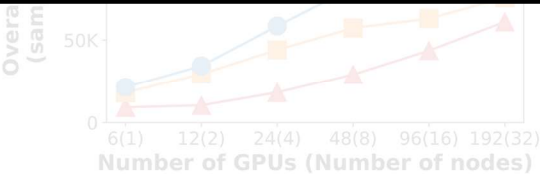
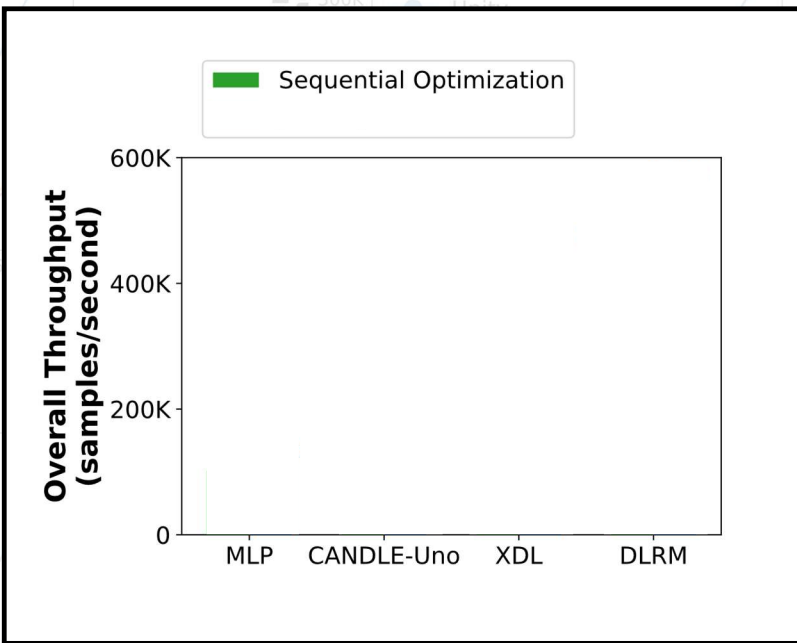
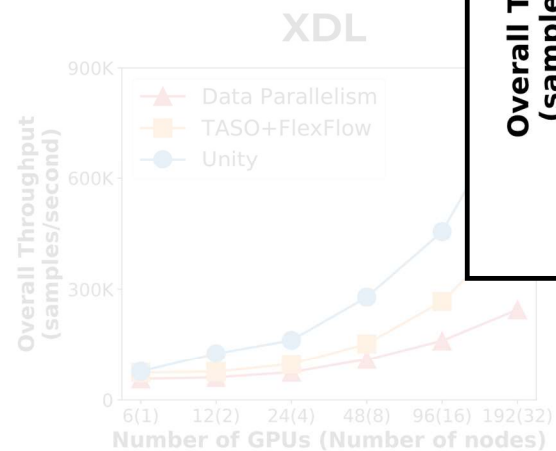
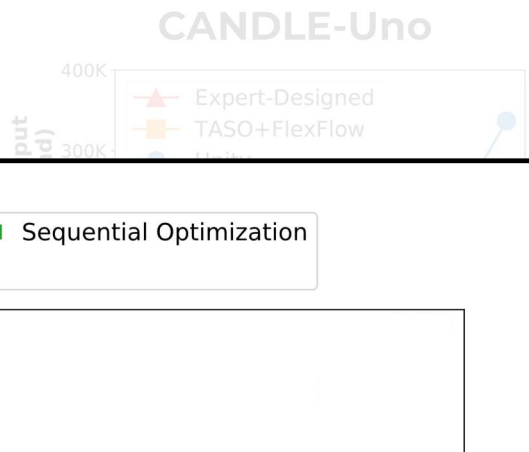
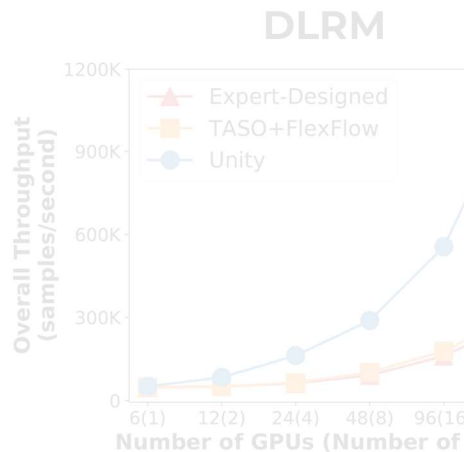
Forms of Parallelism



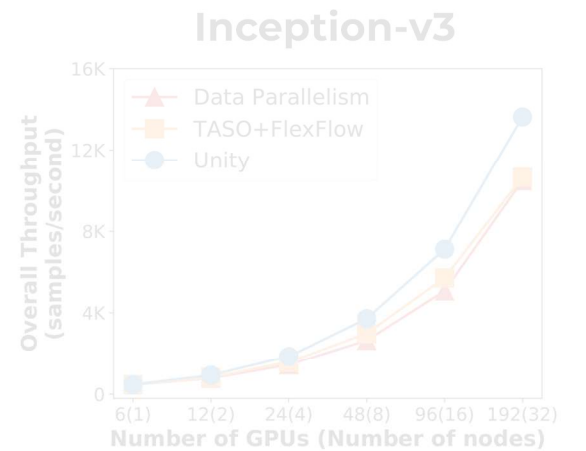
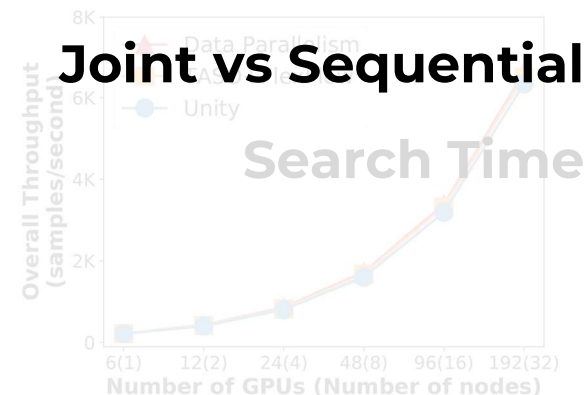


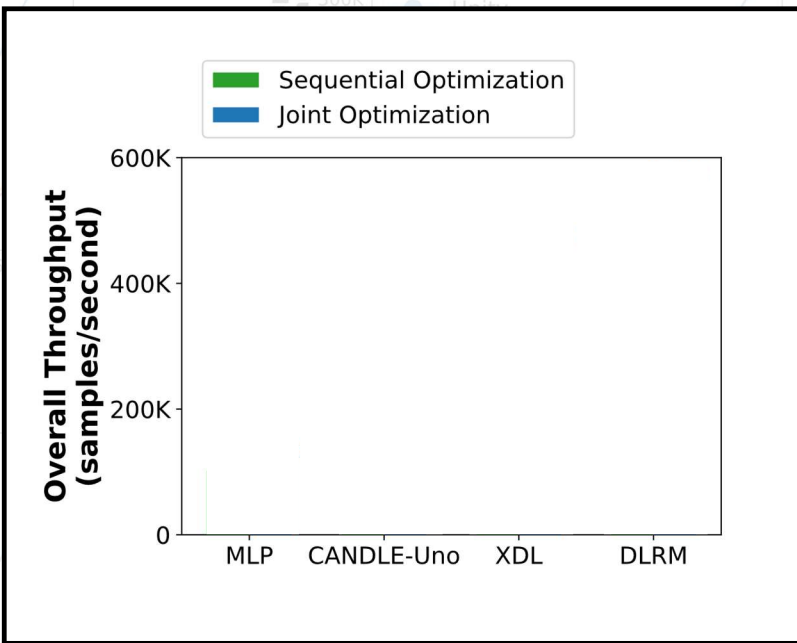
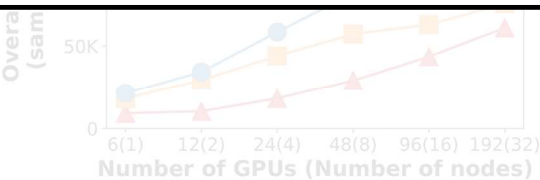
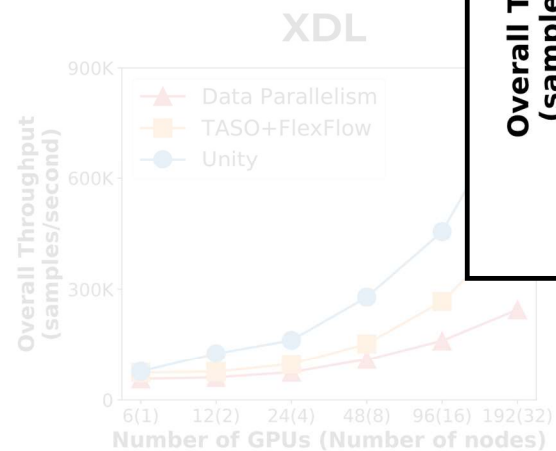
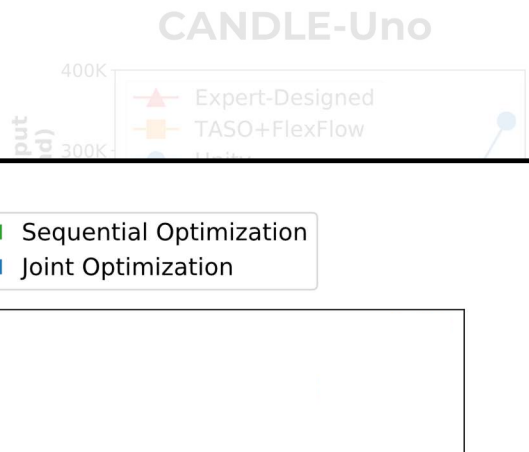
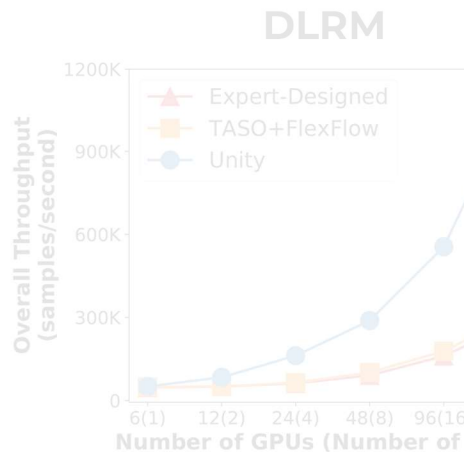
Forms of Parallelism



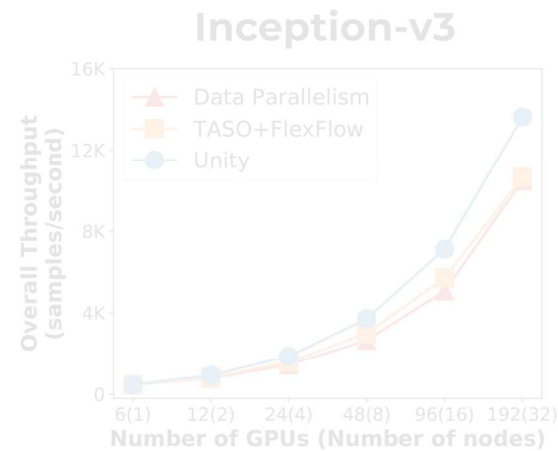
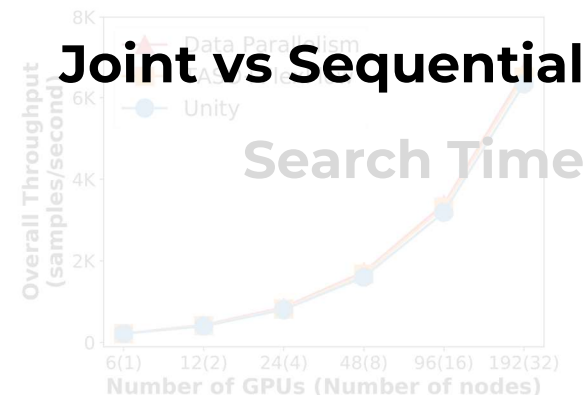


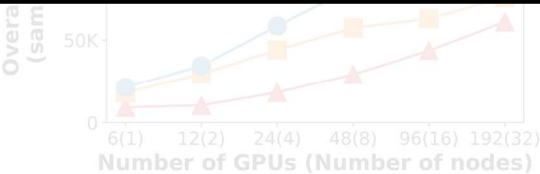
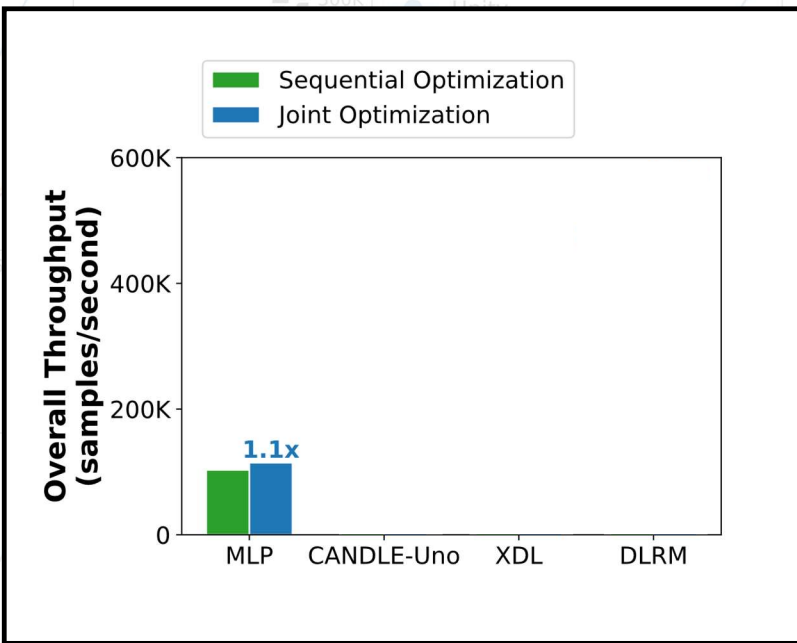
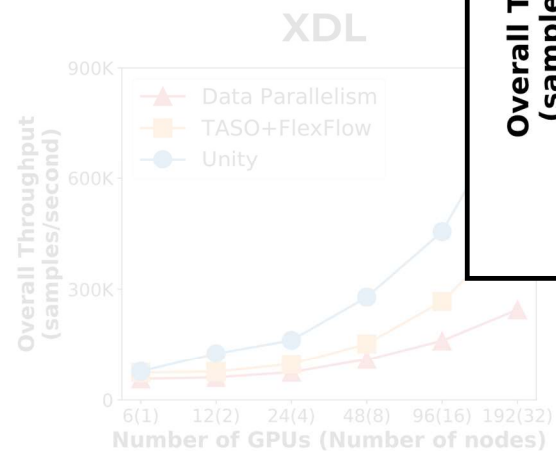
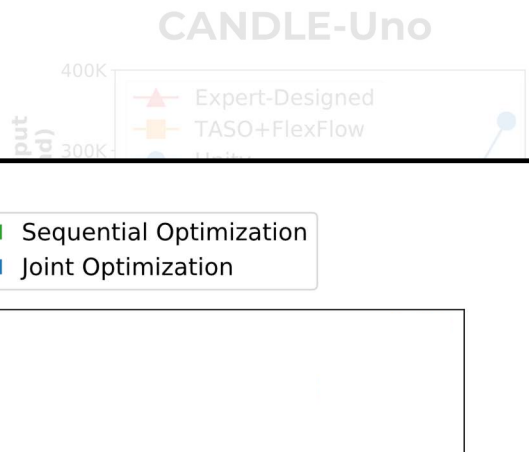
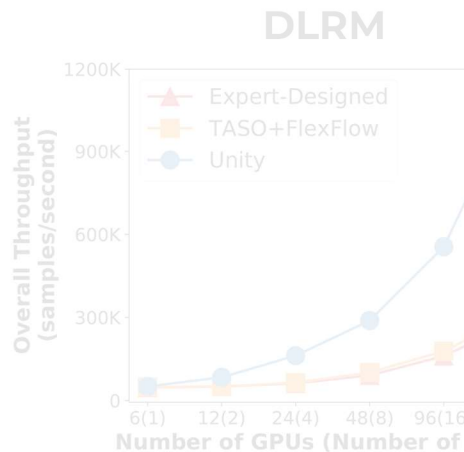
Forms of Parallelism



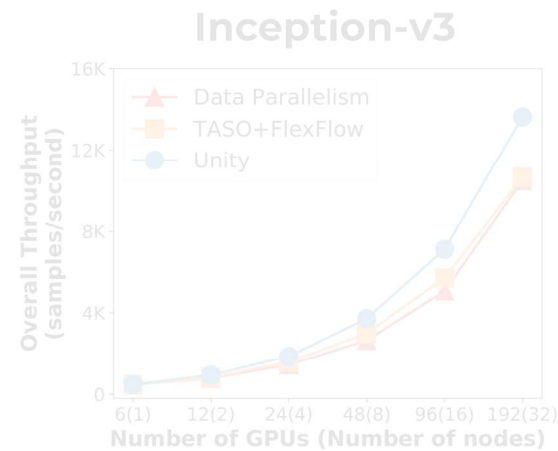
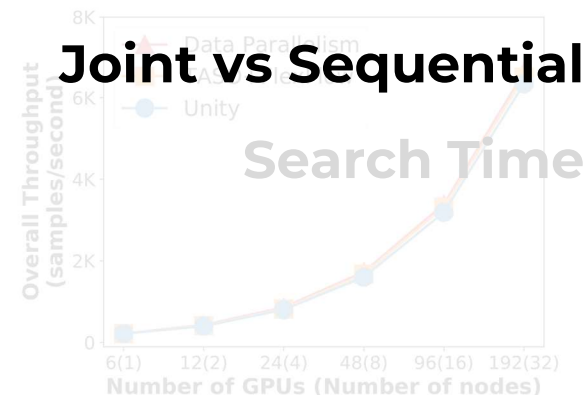


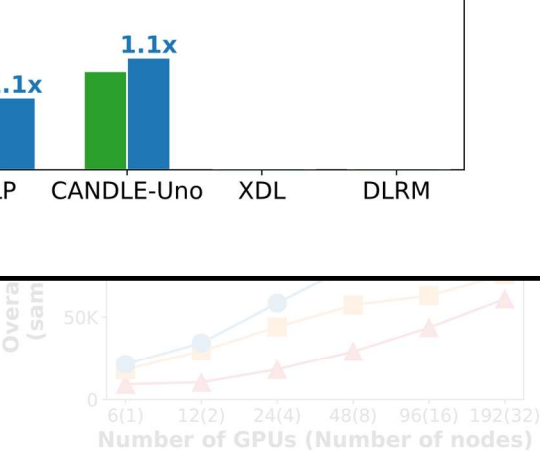
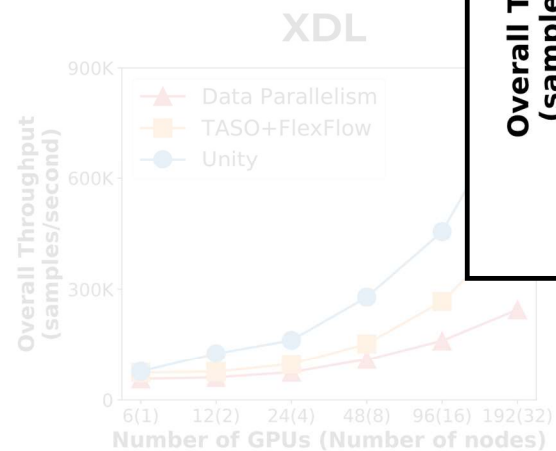
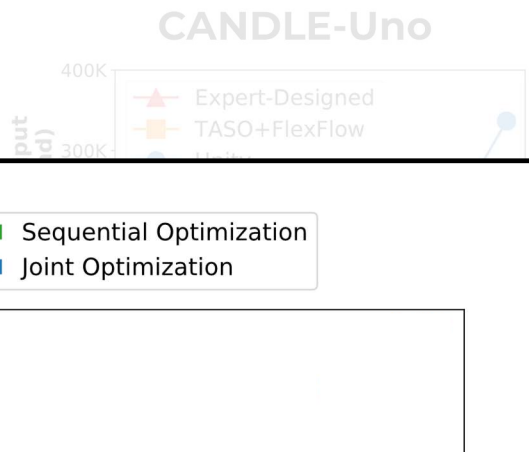
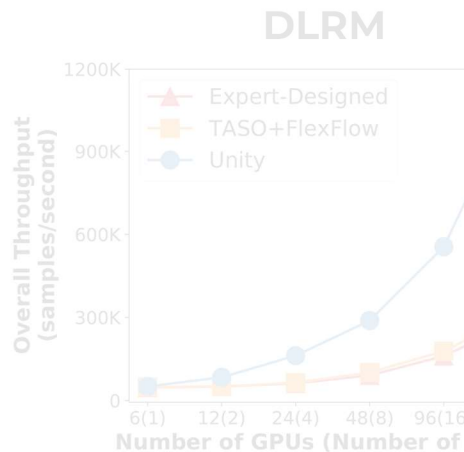
Forms of Parallelism



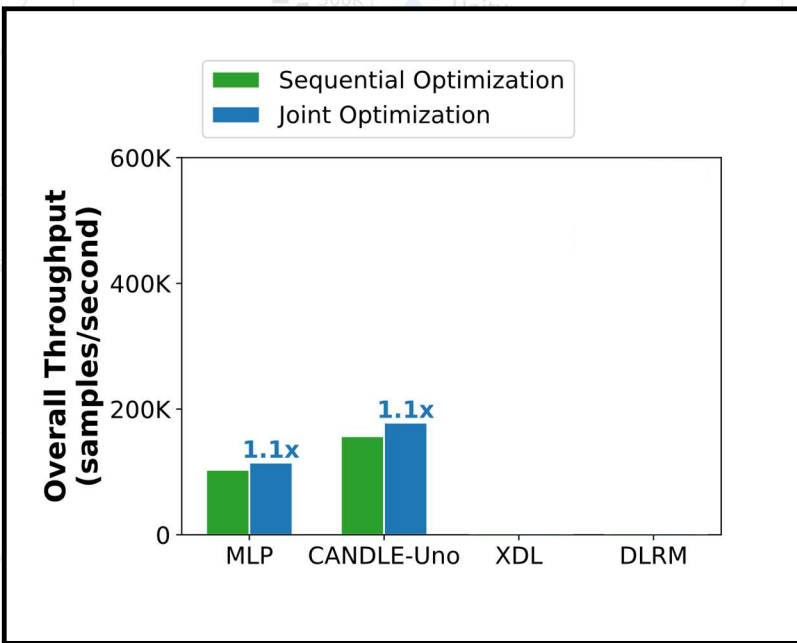
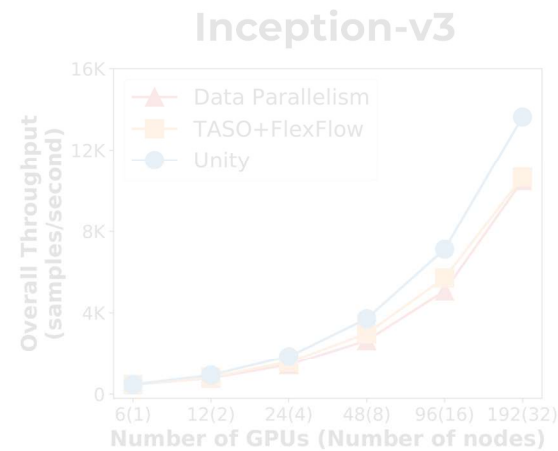
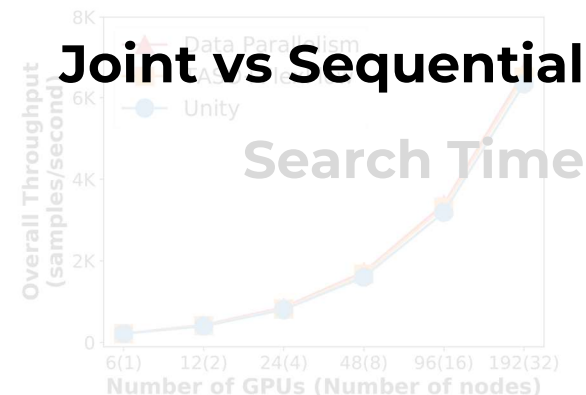


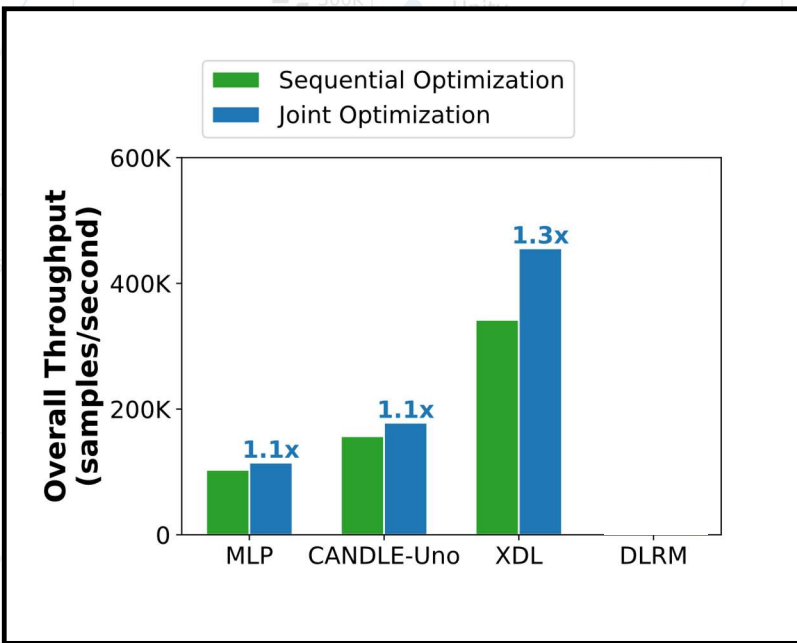
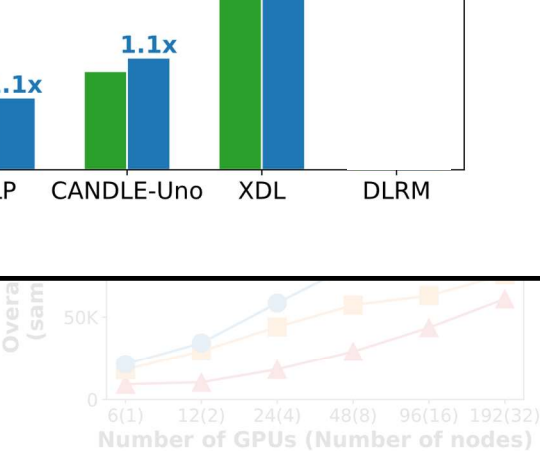
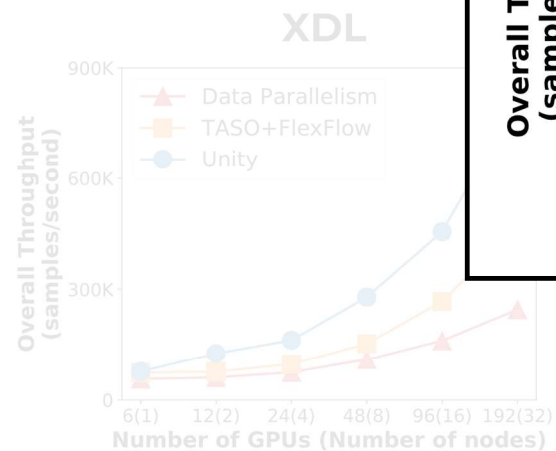
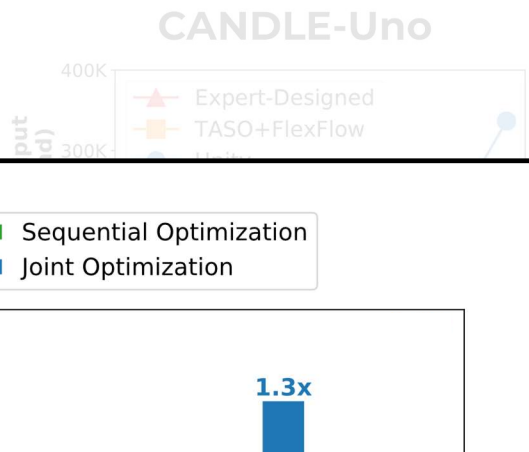
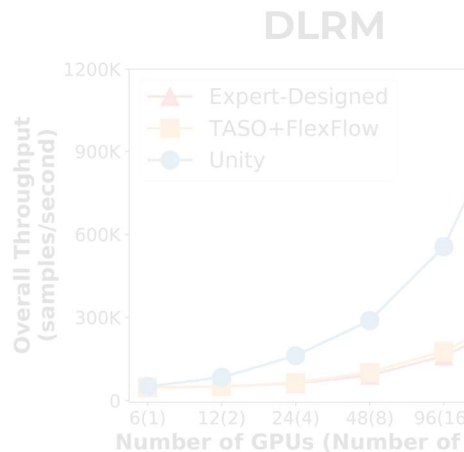
Forms of Parallelism



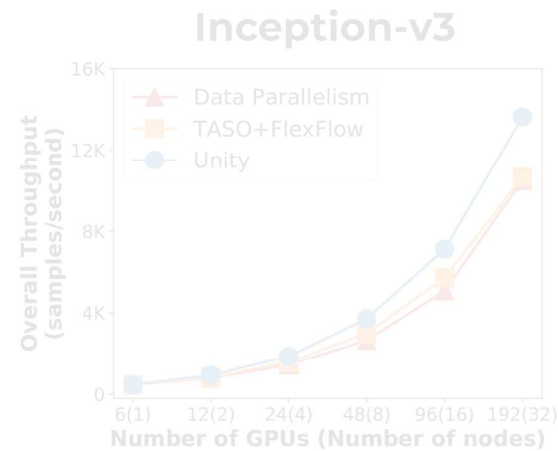
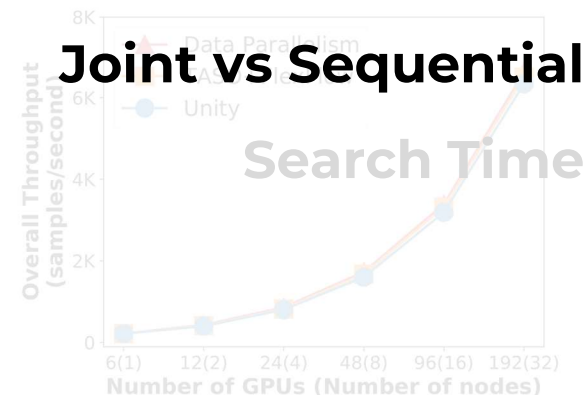


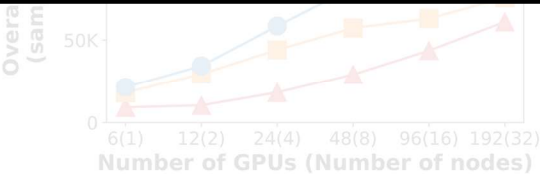
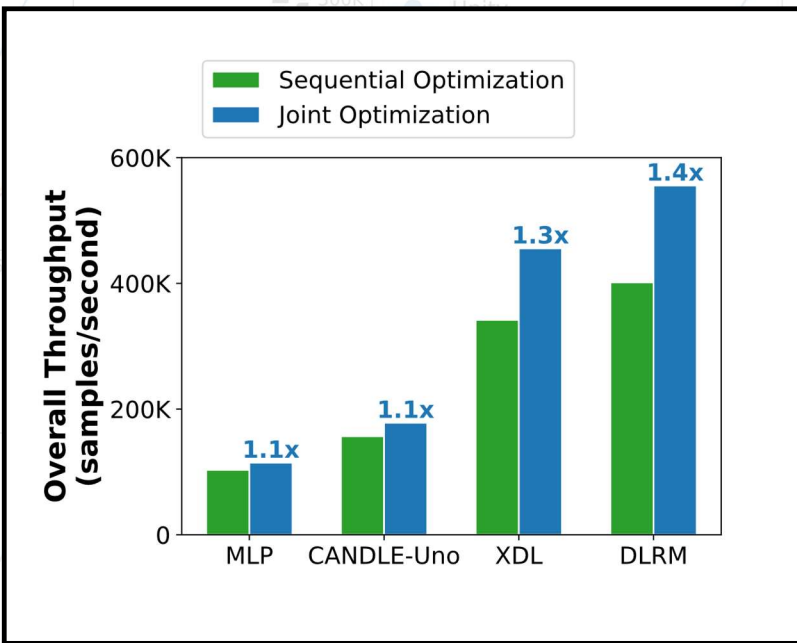
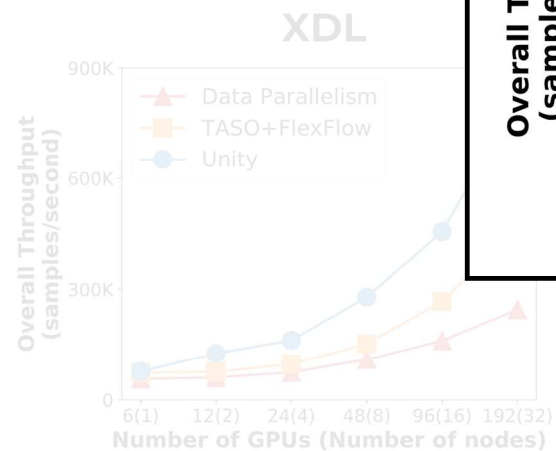
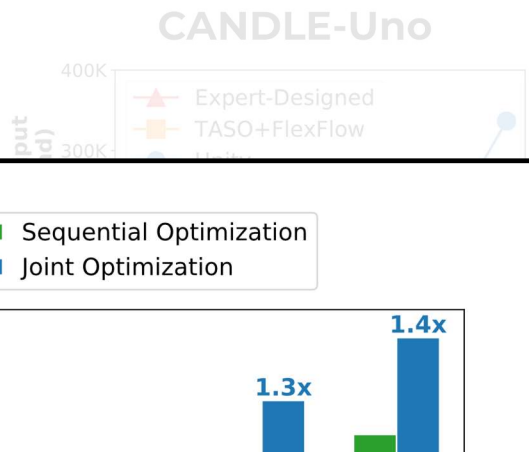
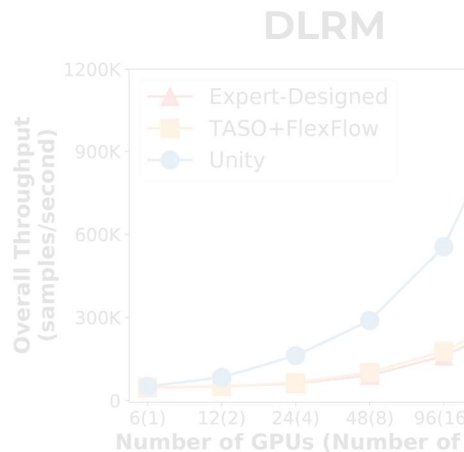
Forms of Parallelism



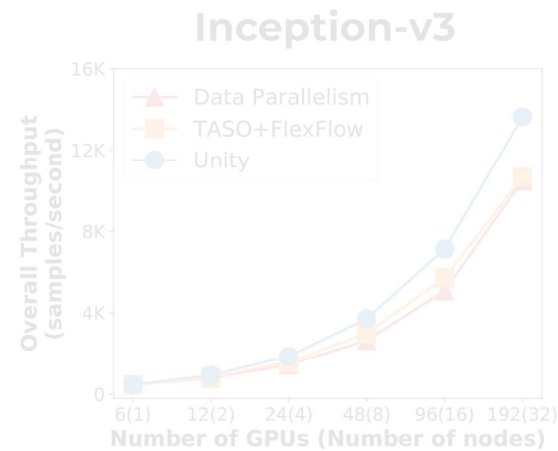
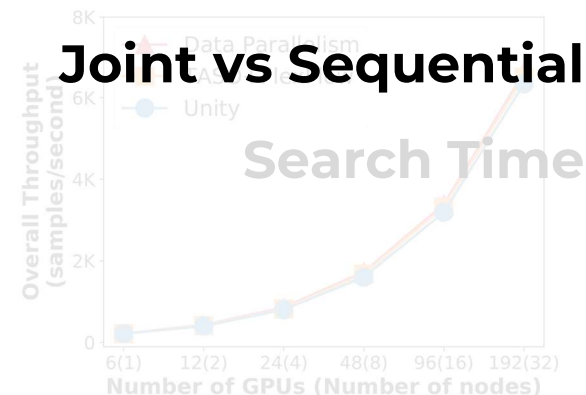


Forms of Parallelism

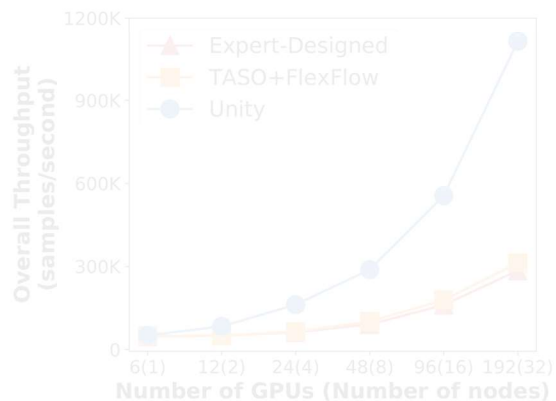




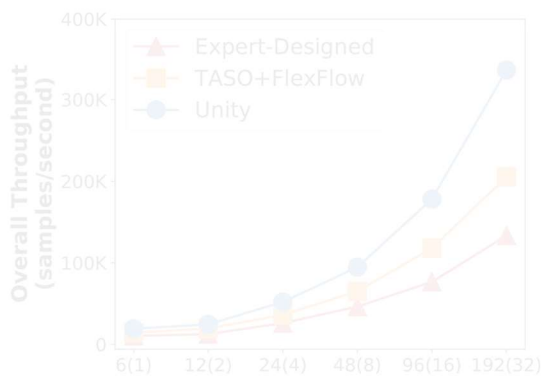
Forms of Parallelism



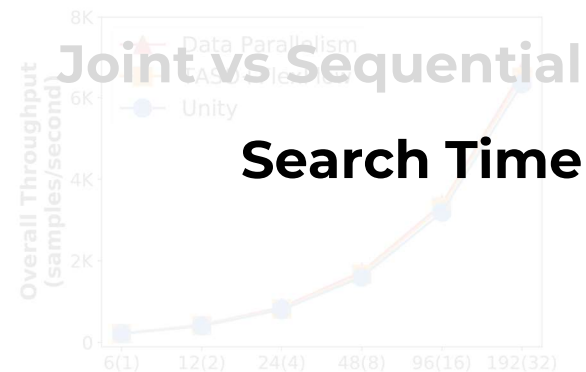
DLRM



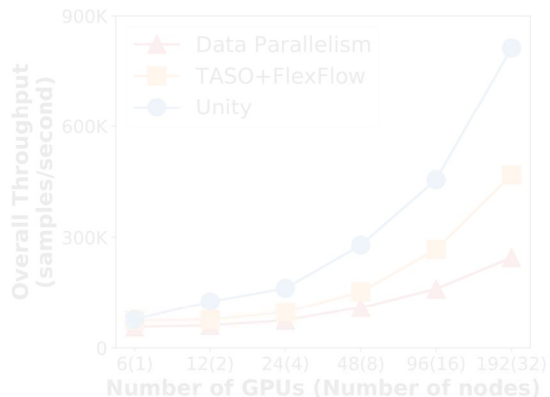
CANDLE-Uno



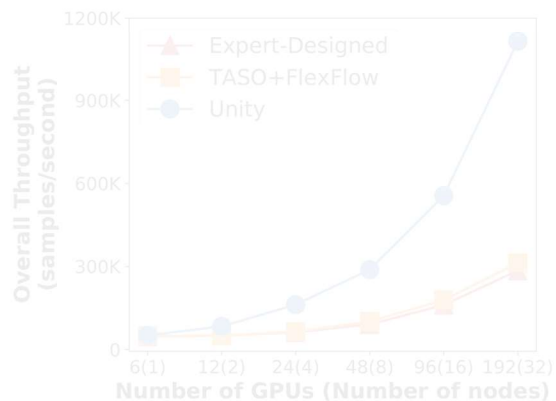
Forms of Parallelism



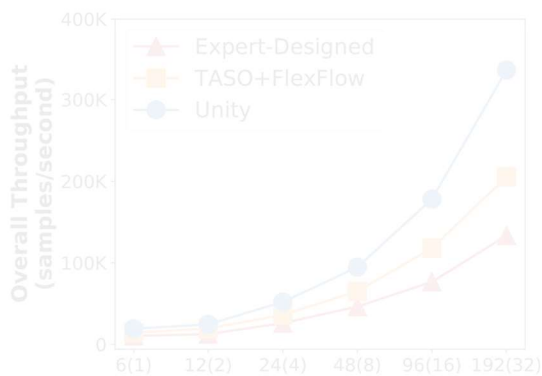
XDL



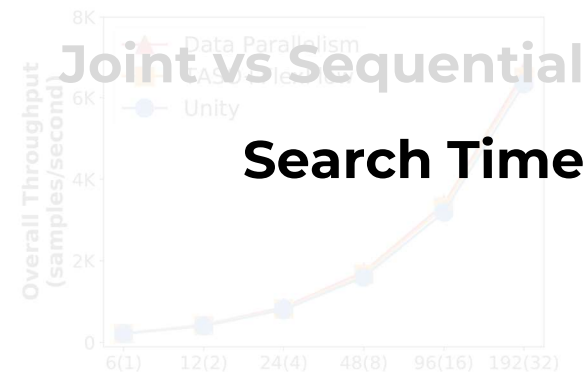
DLRM



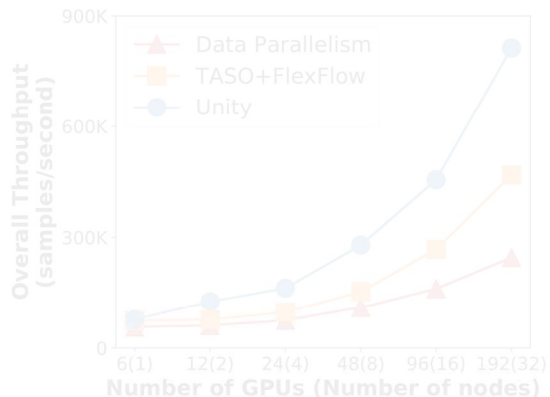
CANDLE-Uno



Forms of Parallelism

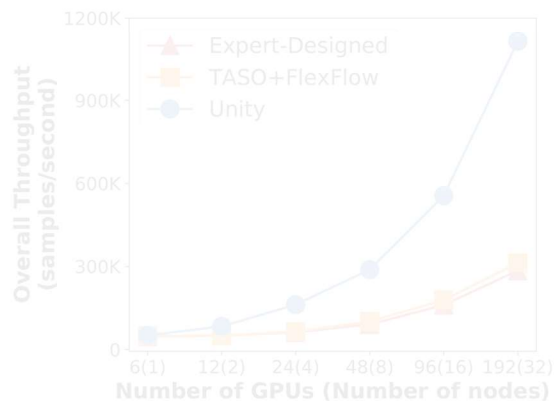


XDL

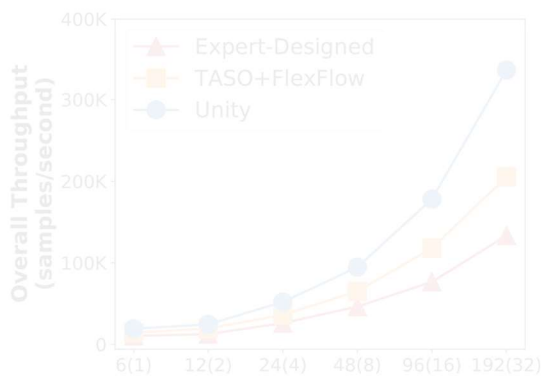


Search Time
< 20 min

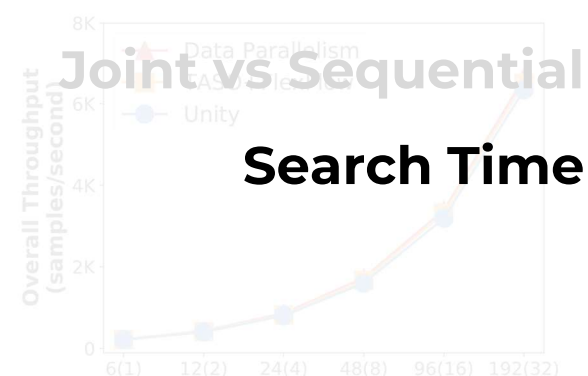
DLRM



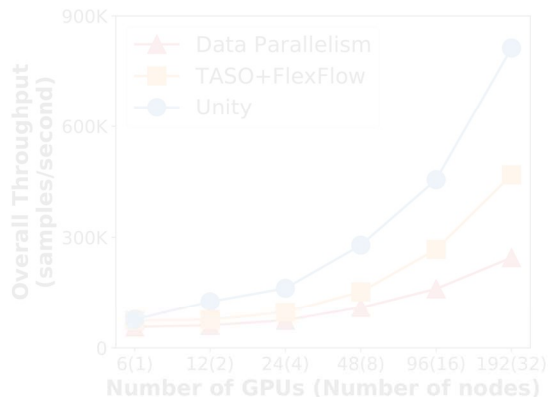
CANDLE-Uno



Forms of Parallelism



XDL

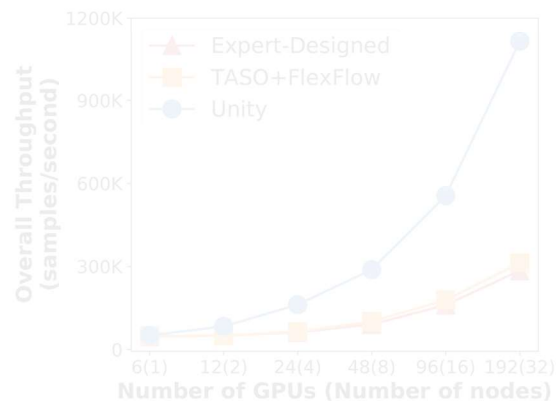


Search Time
< 20 min

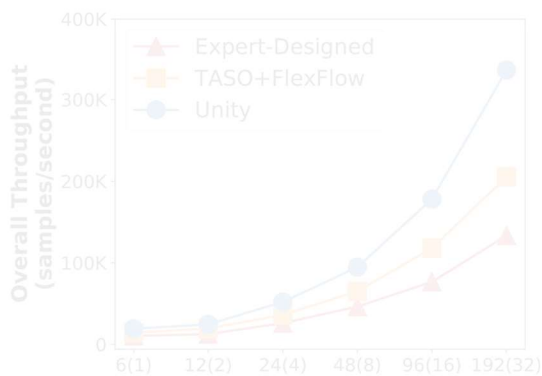


Training Time
hours or days

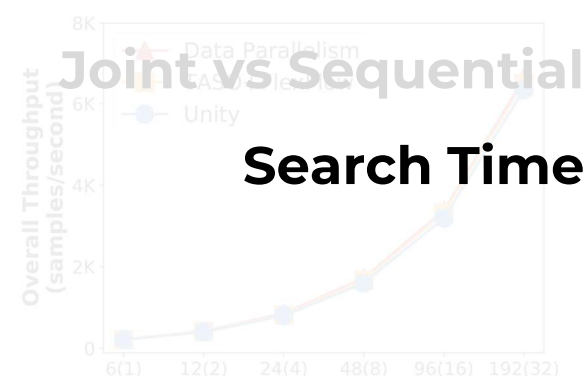
DLRM



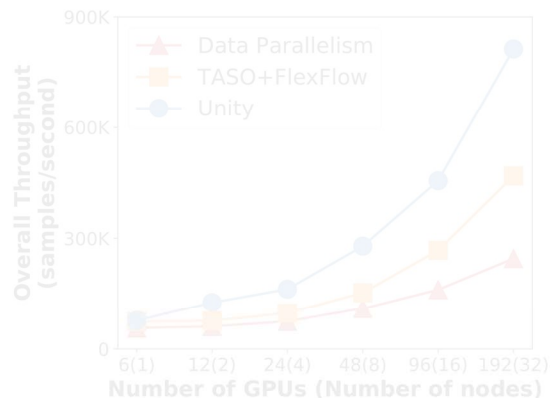
CANDLE-Uno



Forms of Parallelism



XDL

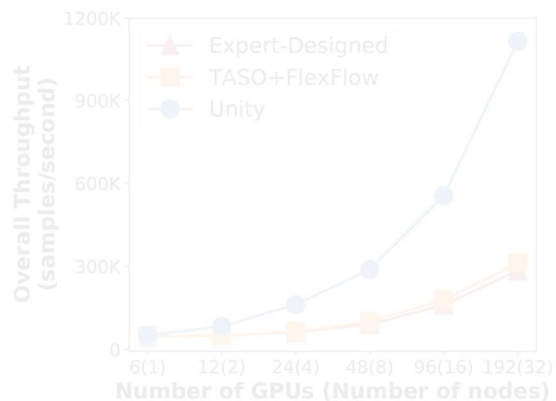


Search Time
< 20 min

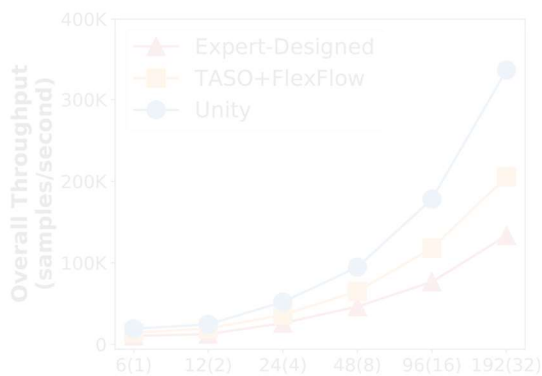


Training Time
hours or days

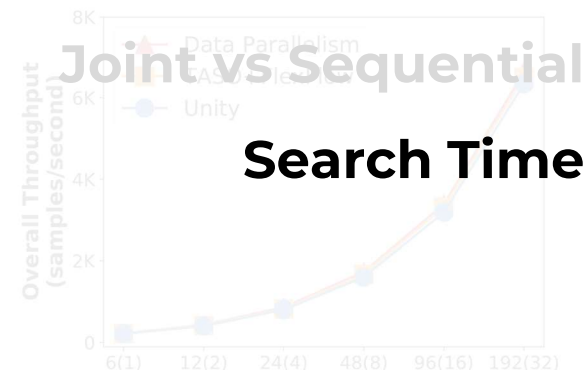
DLRM



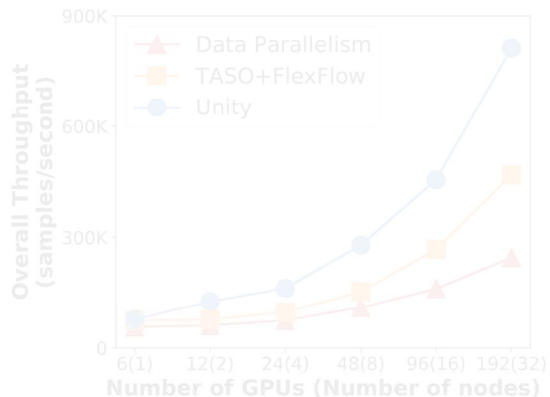
CANDLE-Uno



Forms of Parallelism



XDL



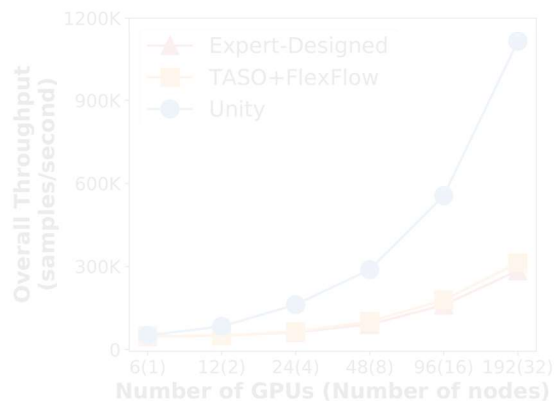
Search Time
< 20 min



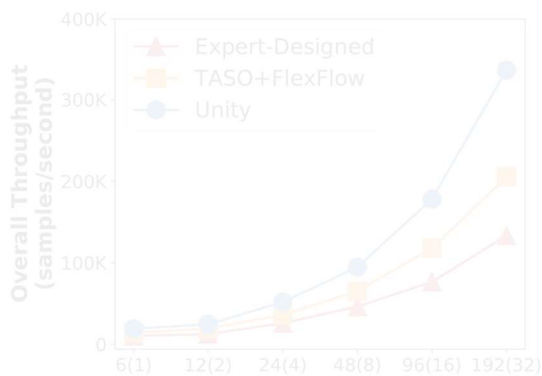
Training Time
hours or days

Naive

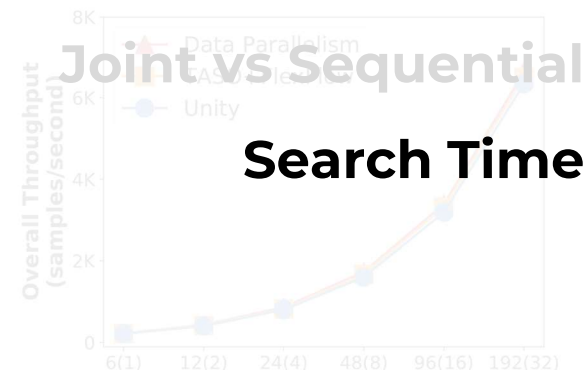
DLRM



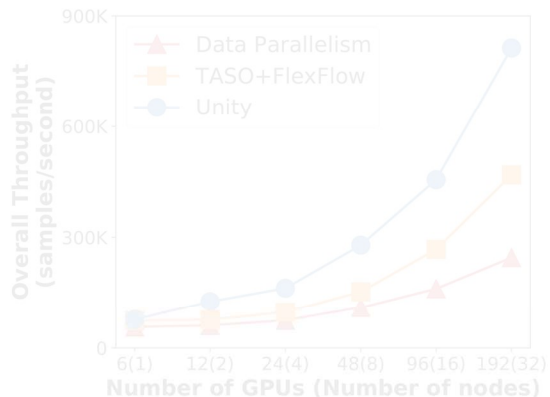
CANDLE-Uno



Forms of Parallelism



XDL



Search Time
< 20 min

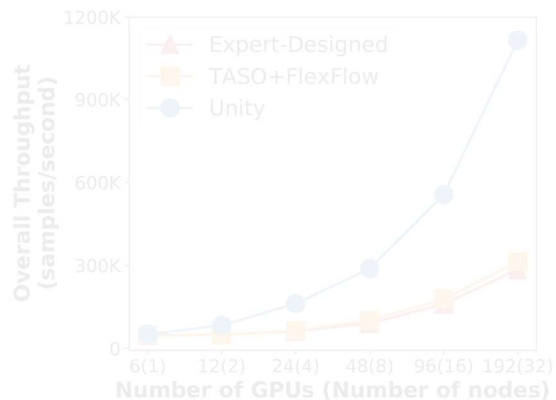


Training Time
hours or days

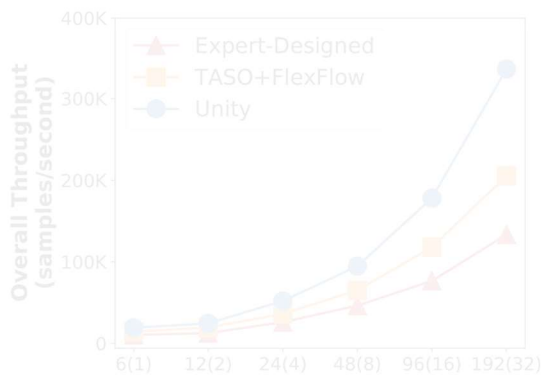
w/o Cache+Split

Naive

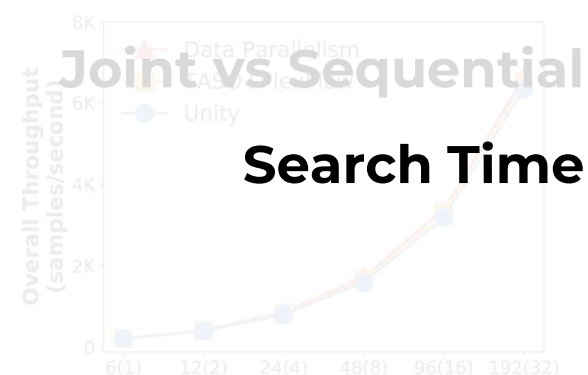
DLRM



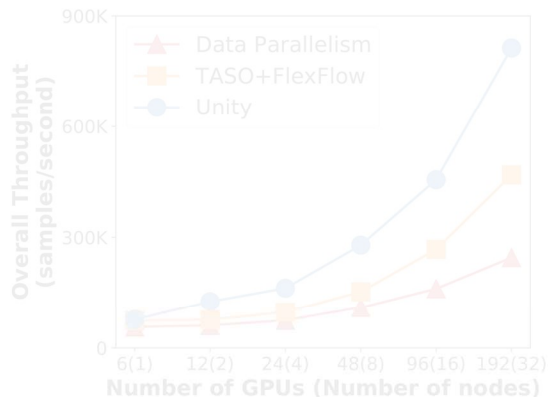
CANDLE-Uno



Forms of Parallelism



XDL



Search Time
< 20 min



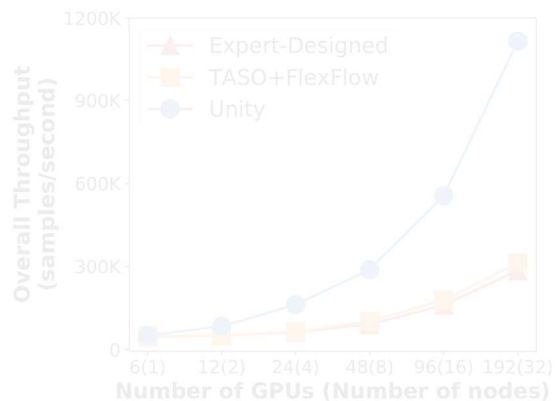
Training Time
hours or days

w/o Split

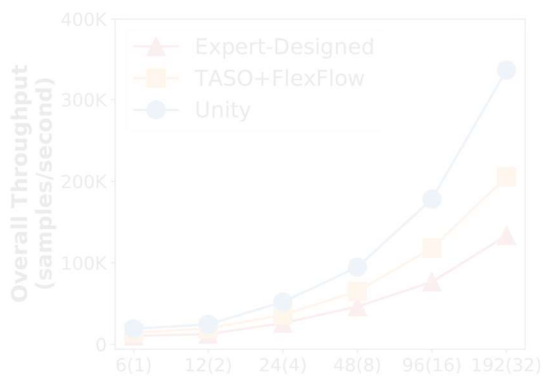
w/o Cache+Split

Naive

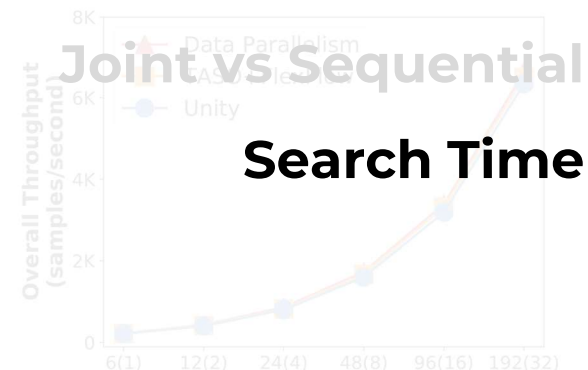
DLRM



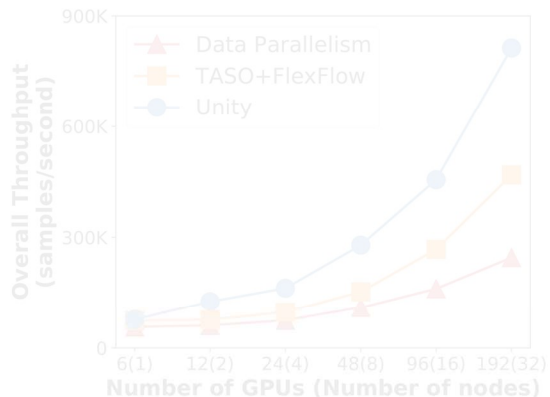
CANDLE-Uno



Forms of Parallelism



XDL



Search Time
< 20 min



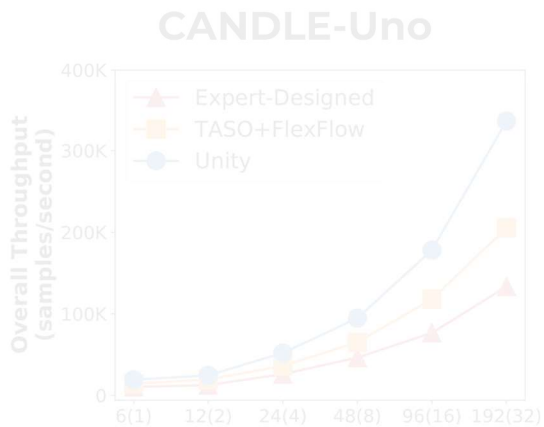
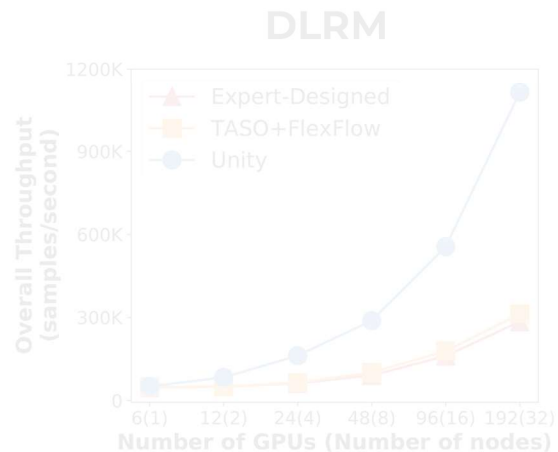
Training Time
hours or days

All

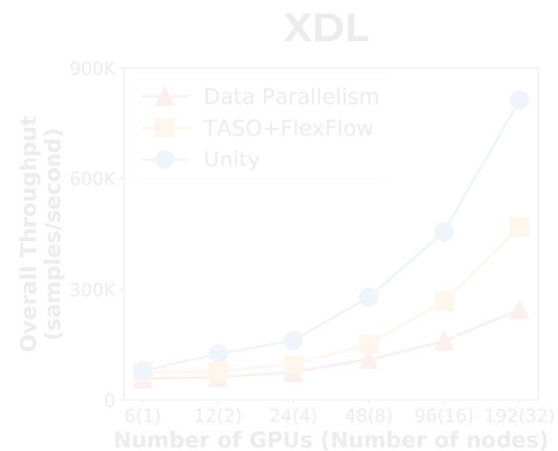
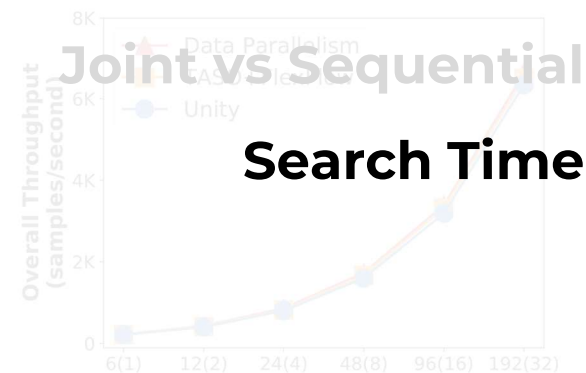
w/o Split

w/o Cache+Split

Naive



Forms of Parallelism

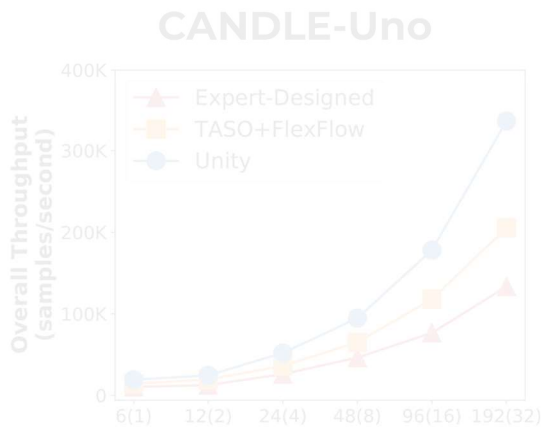
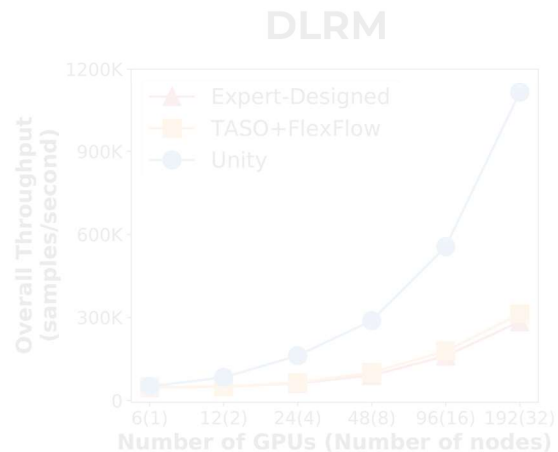


Search Time
< 20 min

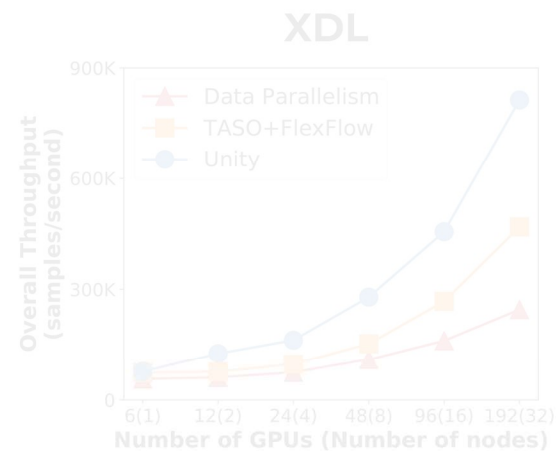
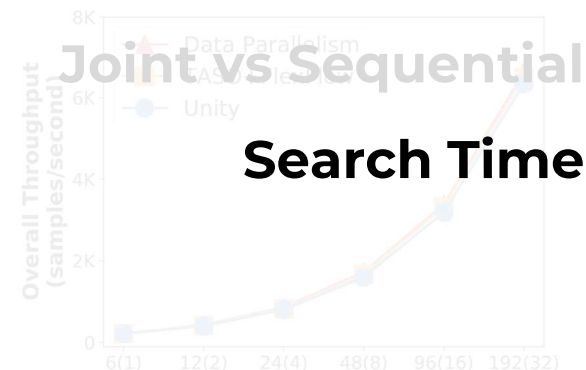


Training Time
hours or days

	All	w/o Split	w/o Cache+Split	Naive
6 GPUs (1 nodes)				
12 GPUs (2 nodes)				
24 GPUs (4 nodes)				
48 GPUs (8 nodes)				



Forms of Parallelism

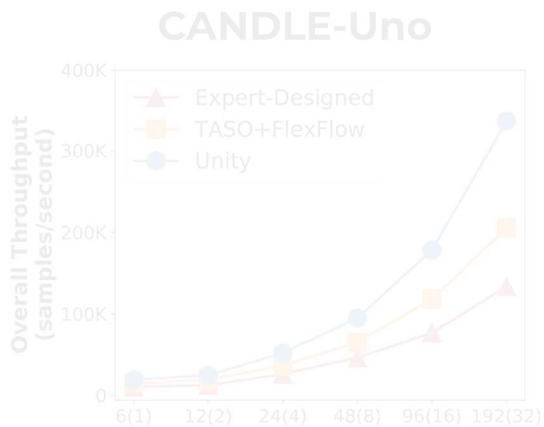
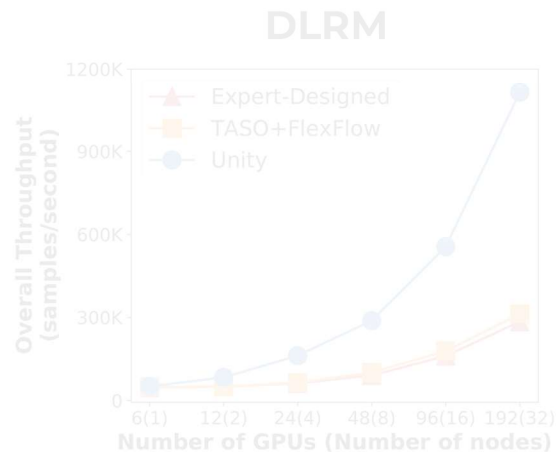


Search Time
< 20 min

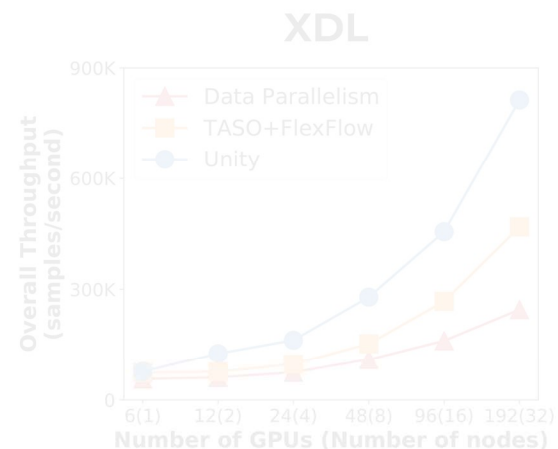
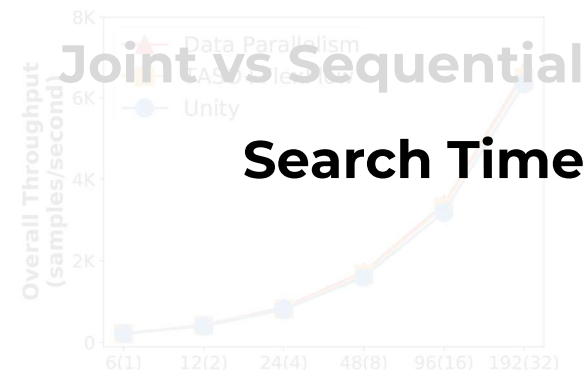


Training Time
hours or days

	All	w/o Split	w/o Cache+Split	Naive
Time				
6 GPUs (1 nodes)	57s			
12 GPUs (2 nodes)				
24 GPUs (4 nodes)				
48 GPUs (8 nodes)				



Forms of Parallelism

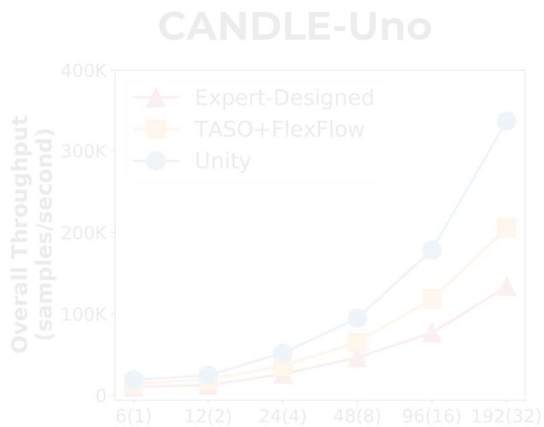
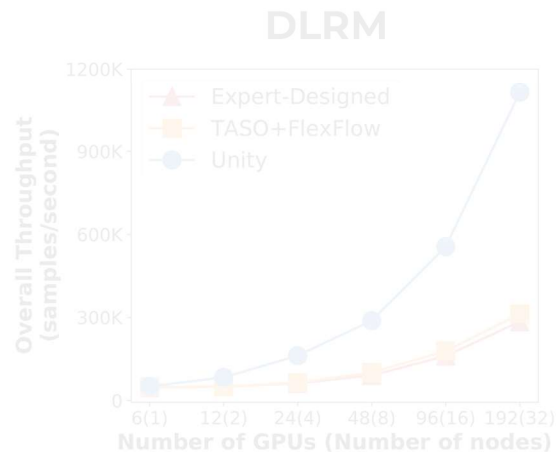


Search Time
< 20 min

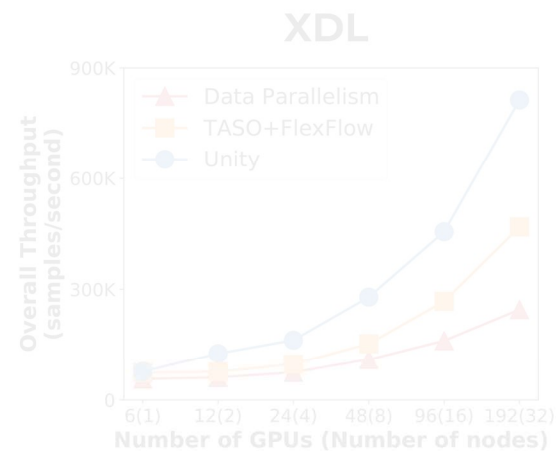
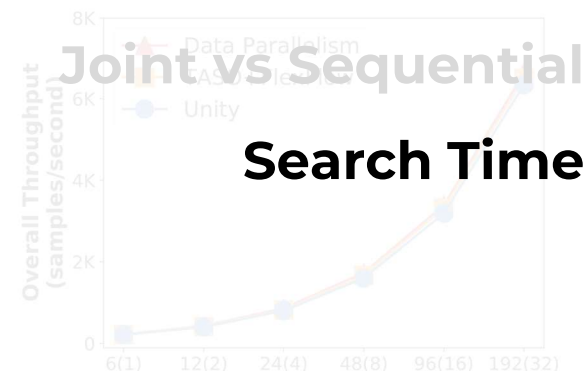


Training Time
hours or days

	All	w/o Split	w/o Cache+Split	Naive
	Time Scaled			
6 GPUs (1 nodes)	57s	1×		
12 GPUs (2 nodes)				
24 GPUs (4 nodes)				
48 GPUs (8 nodes)				



Forms of Parallelism

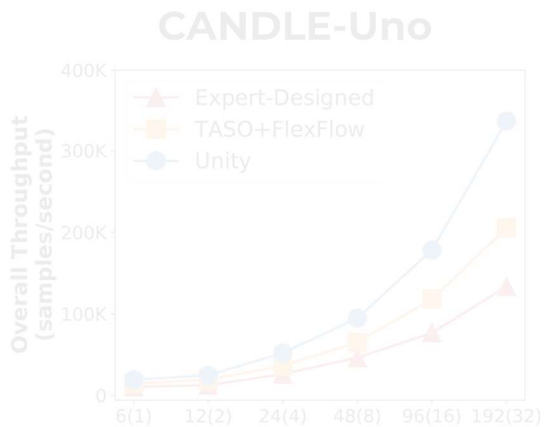
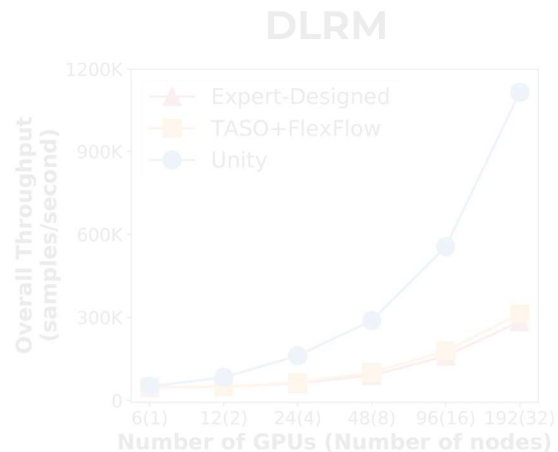


Search Time
< 20 min

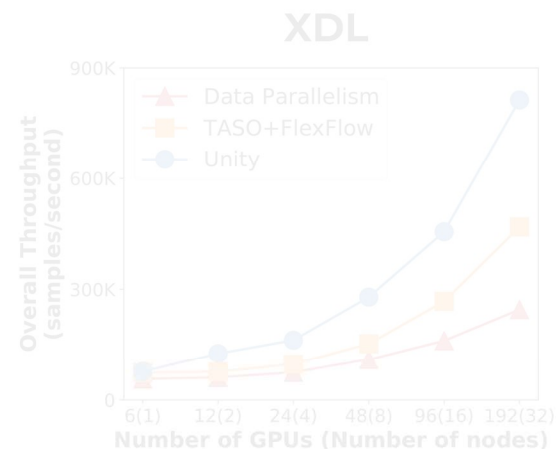
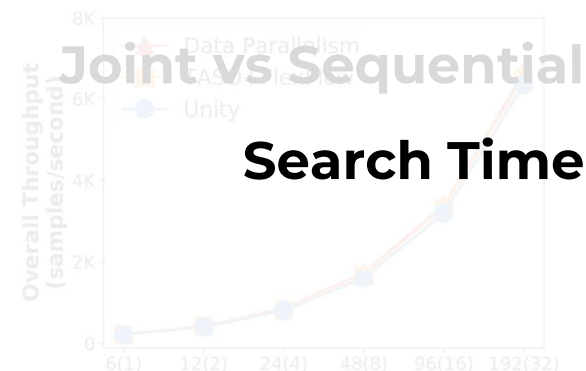


Training Time
hours or days

	All		w/o Split	w/o Cache+Split	Naive	
	Time	Scaled			Time	Scaled
6 GPUs (1 nodes)	57s	1×			> 1h	n/a
12 GPUs (2 nodes)					> 1h	n/a
24 GPUs (4 nodes)					> 1h	n/a
48 GPUs (8 nodes)					> 1h	n/a



Forms of Parallelism

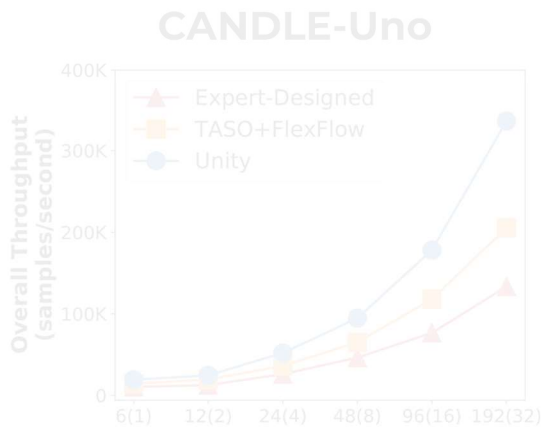
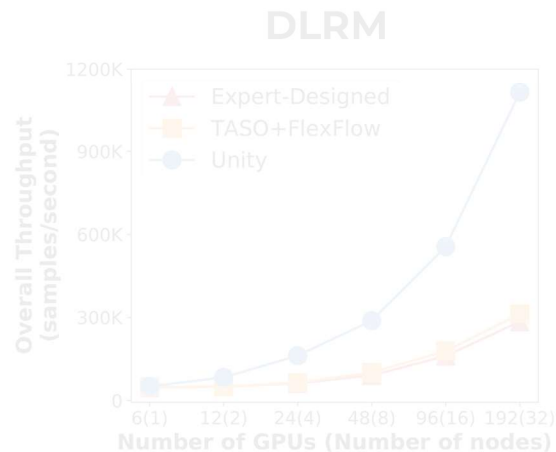


Search Time
< 20 min

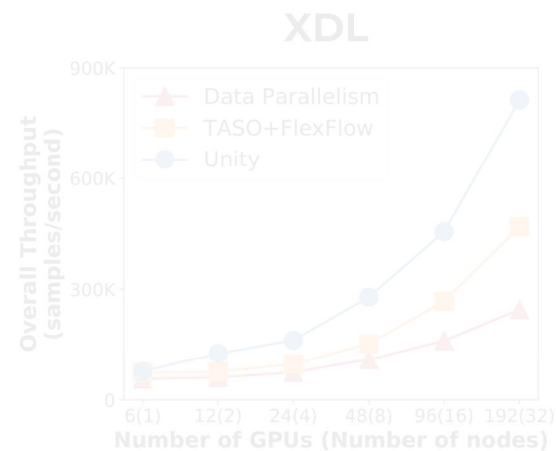
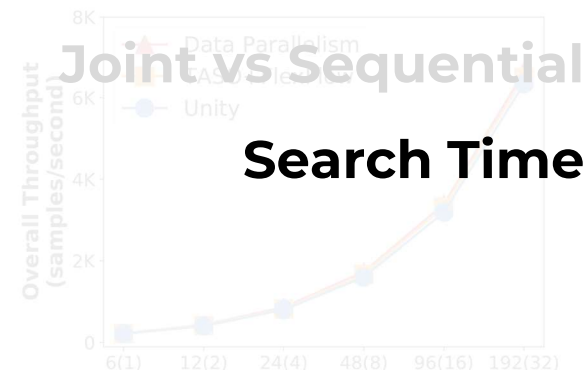


Training Time
hours or days

	All		w/o Split		w/o Cache+Split		Naive	
	Time	Scaled	Time	Scaled	Time	Scaled	Time	Scaled
6 GPUs (1 nodes)	57s	1×			37m 01s	38.5×	> 1h	n/a
12 GPUs (2 nodes)					> 1h	n/a	> 1h	n/a
24 GPUs (4 nodes)					> 1h	n/a	> 1h	n/a
48 GPUs (8 nodes)					> 1h	n/a	> 1h	n/a



Forms of Parallelism

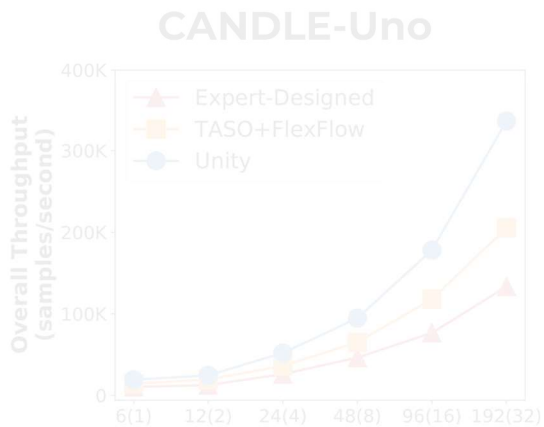
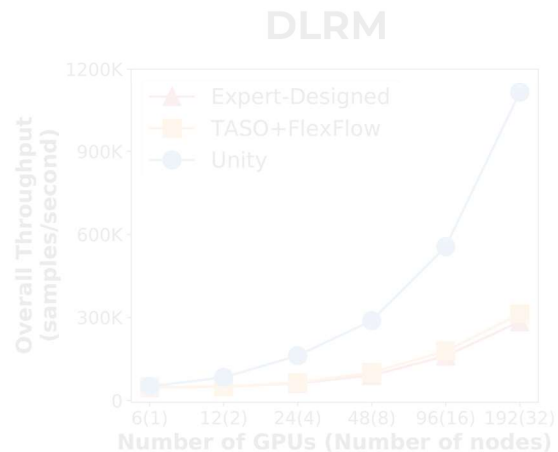


Search Time
< 20 min

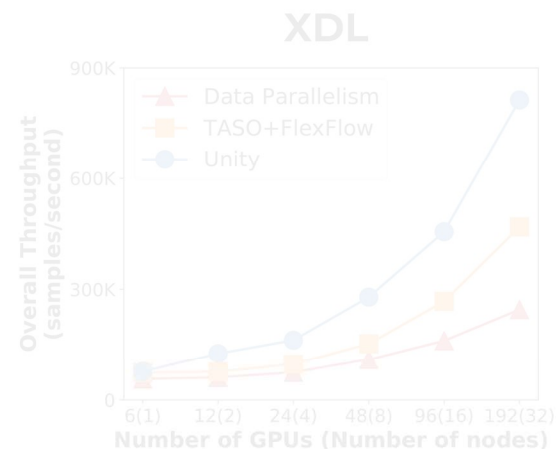
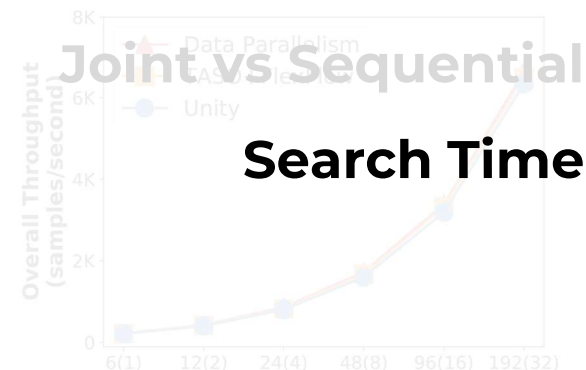


Training Time
hours or days

	All		w/o Split		w/o Cache+Split		Naive	
	Time	Scaled	Time	Scaled	Time	Scaled	Time	Scaled
6 GPUs (1 nodes)	57s	1×	4m 01s	4.3×	37m 01s	38.5×	> 1h	n/a
12 GPUs (2 nodes)			11m 15s	16.8×	> 1h	n/a	> 1h	n/a
24 GPUs (4 nodes)			> 1h	n/a	> 1h	n/a	> 1h	n/a
48 GPUs (8 nodes)			> 1h	n/a	> 1h	n/a	> 1h	n/a



Forms of Parallelism



Search Time
< 20 min



Training Time
hours or days

	All		w/o Split		w/o Cache+Split		Naive	
	Time	Scaled	Time	Scaled	Time	Scaled	Time	Scaled
6 GPUs (1 nodes)	57s	1×	4m 01s	4.3×	37m 01s	38.5×	> 1h	n/a
12 GPUs (2 nodes)	1m 47s	1.9×	11m 15s	16.8×	> 1h	n/a	> 1h	n/a
24 GPUs (4 nodes)	3m 00s	3.1×	> 1h	n/a	> 1h	n/a	> 1h	n/a
48 GPUs (8 nodes)	5m 55s	6.1×	> 1h	n/a	> 1h	n/a	> 1h	n/a

1

Joint Optimization

1

Joint
Optimization

2

Unified Graph
Representation
(PCG)

1

Joint
Optimization

2

Unified Graph
Representation
(PCG)

3

Hierarchical
Search
Algorithm

<https://github.com/flexflow/FlexFlow>

kadinzhang Added tests for Linear operator in align/linear (#264) 488423a 19 days ago 1,339 commits

.circleci	[CircleCI] update script	12 months ago
align	Added tests for Linear operator in align/linear (#264)	19 days ago
bootcamp_demo	[python] make num_samples a property (#121)	14 months ago
cmake	[CMake] retrieve legion installation path for spack	9 months ago
conda	[Conda] fix conda	17 months ago
config	[Alignment] Add embedding alignment (#255)	3 months ago
deps	[Legion] version update	8 months ago
docker	Adding Dockerfile	2 years ago

A distributed deep learning framework that supports flexible parallelization strategies.

 Readme
 Apache-2.0 license
 Code of conduct
 458 stars
 21 watching
 100 forks

Releases

 Release 21.09 (September 30th ...)
 Latest



Keras



PyTorch



ONNX

python

python

[python] fix embedding for pytorch 1.1

Merging the public branch to the master branch (#74)

27 days ago

16 months ago

Publish your first package



CMU



<https://github.com/flexflow/FlexFlow>

Questions?

