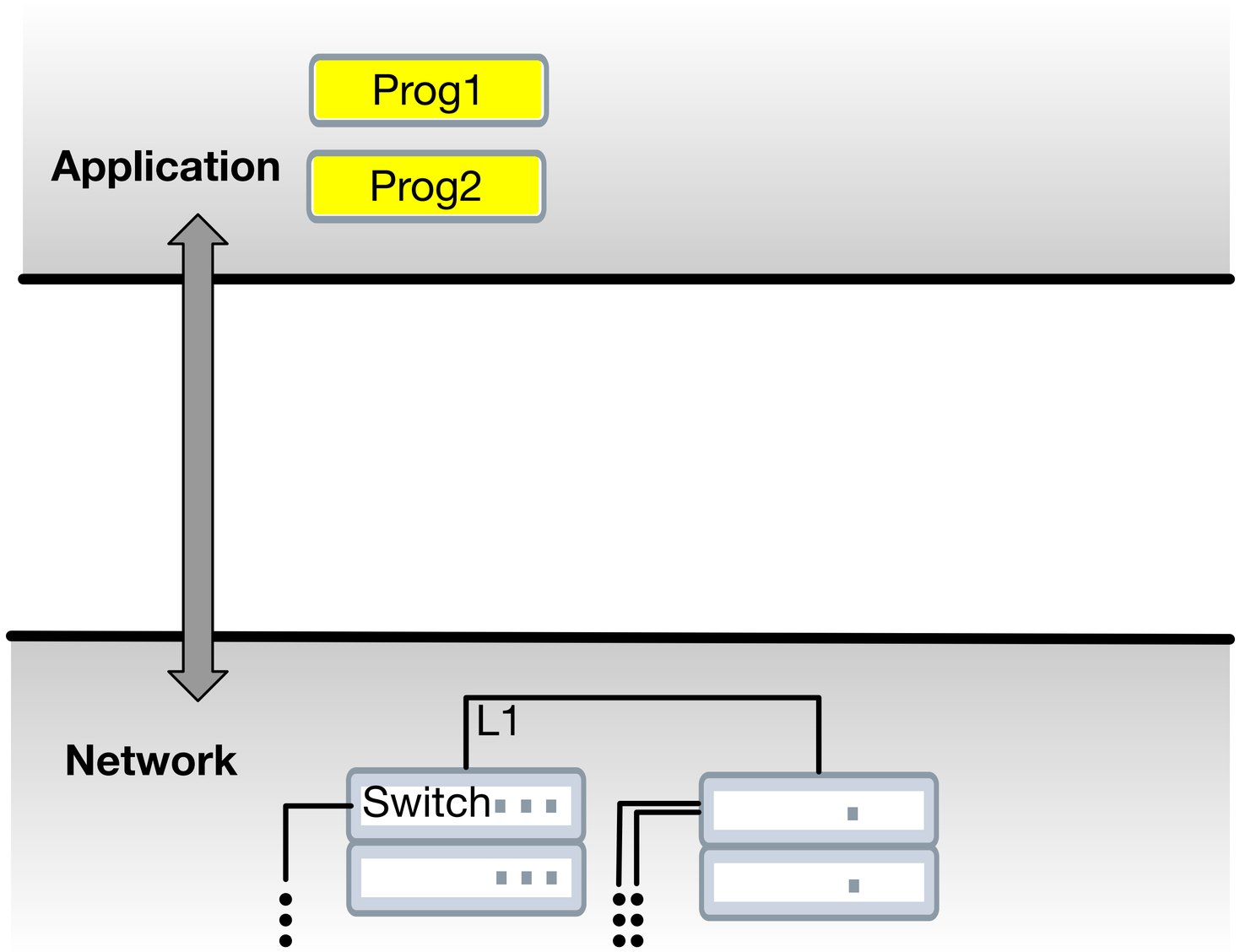
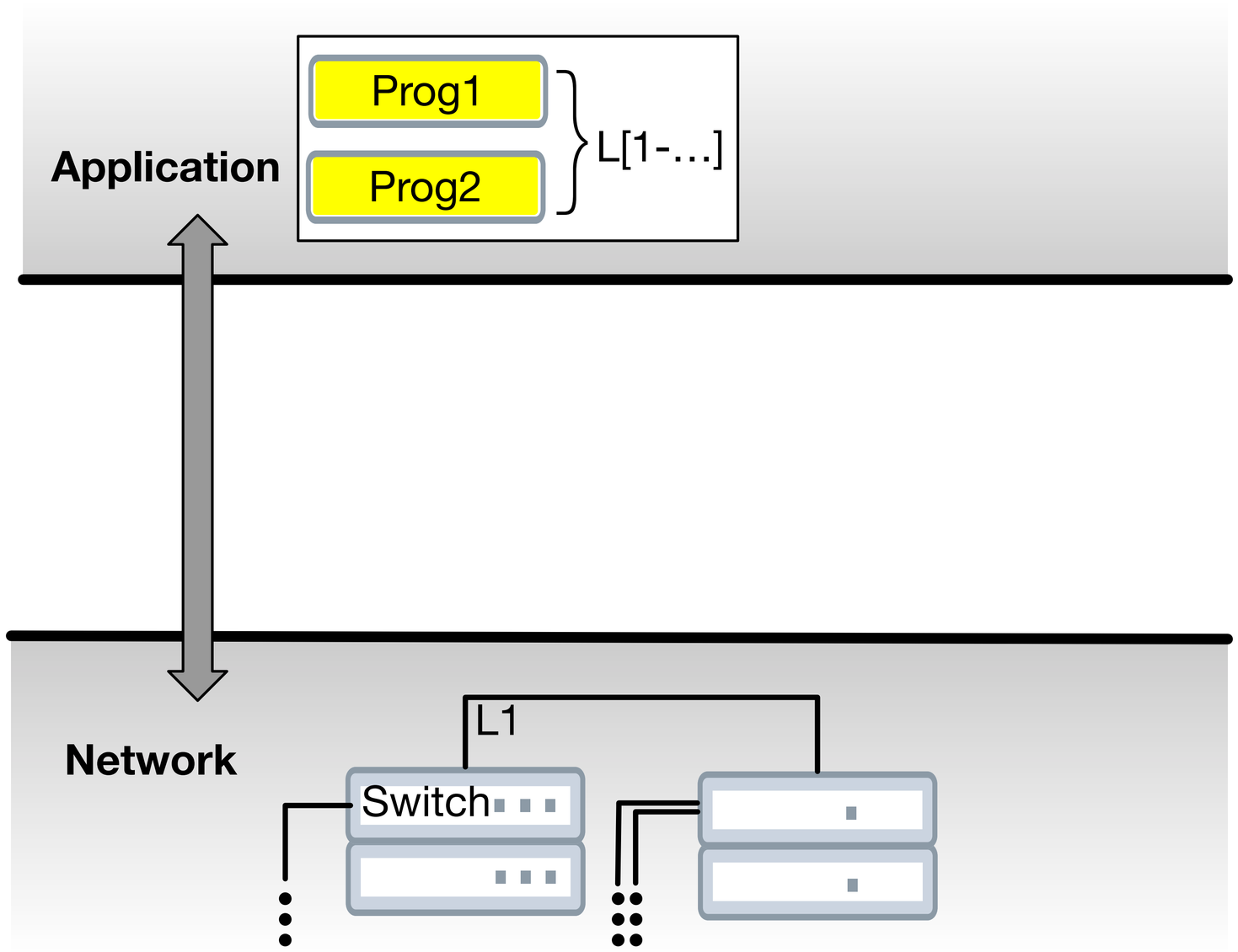
Leveraging In-Network Application Awareness

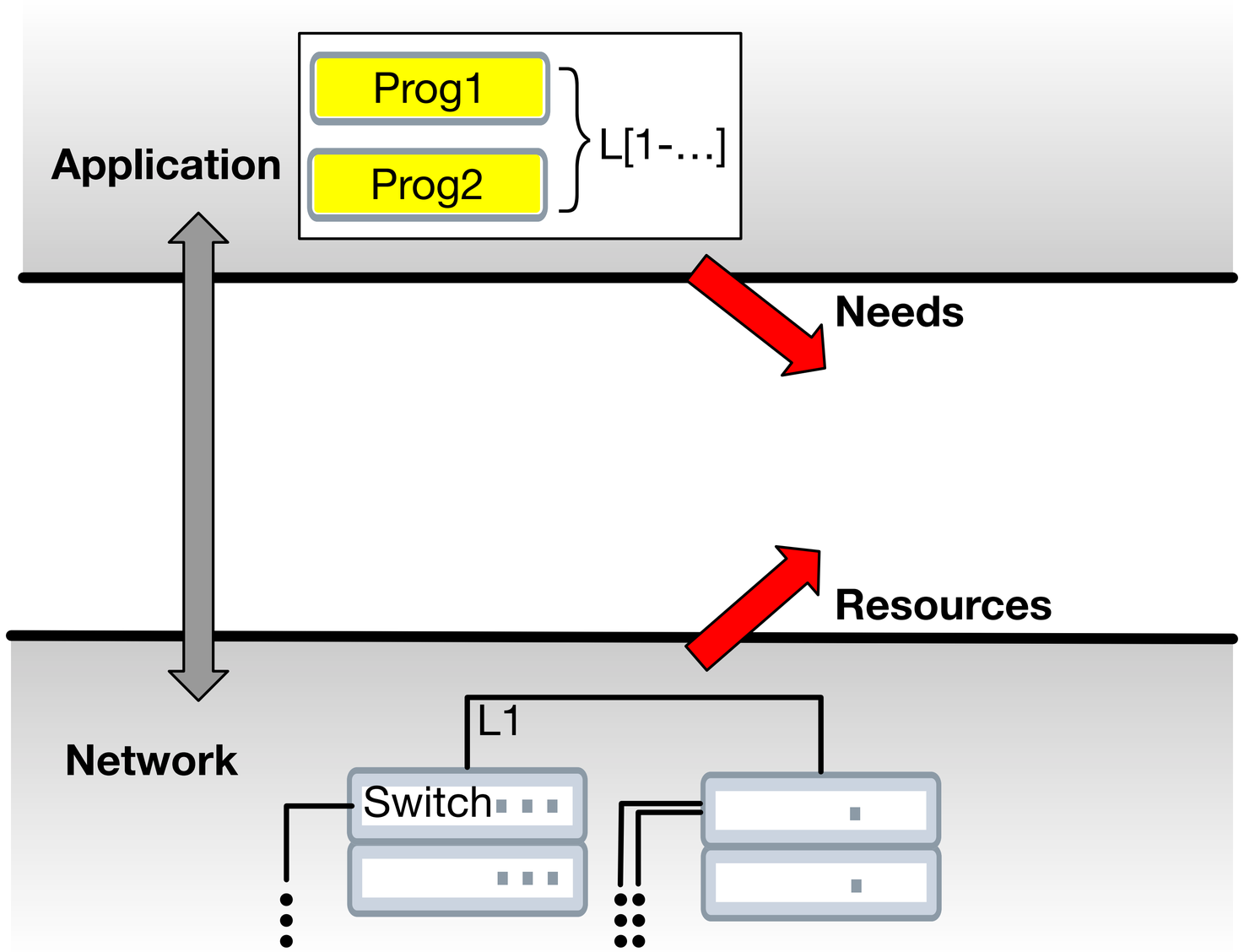
Nik Sultana
University of Pennsylvania

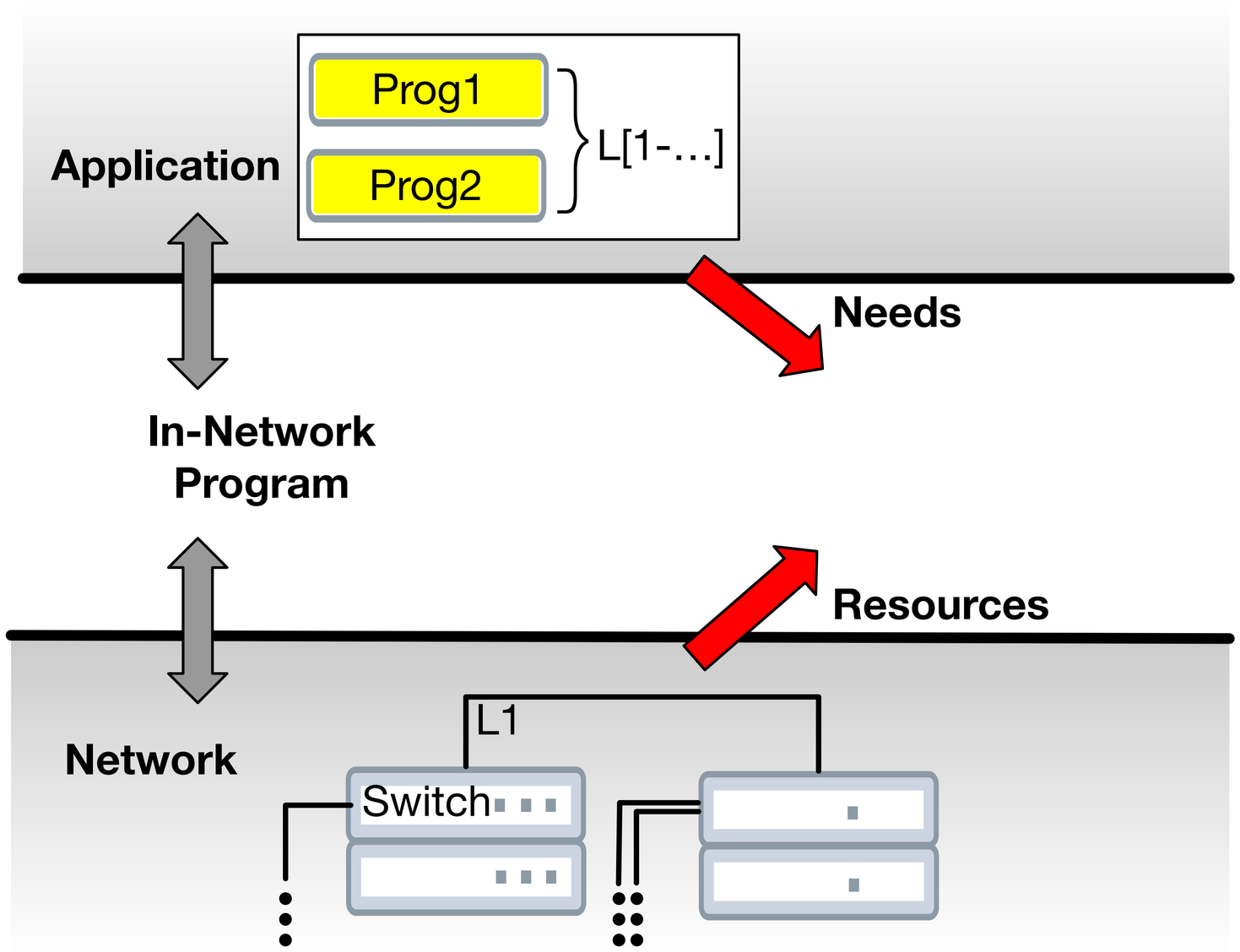
https://www.seas.upenn.edu/~nsultana

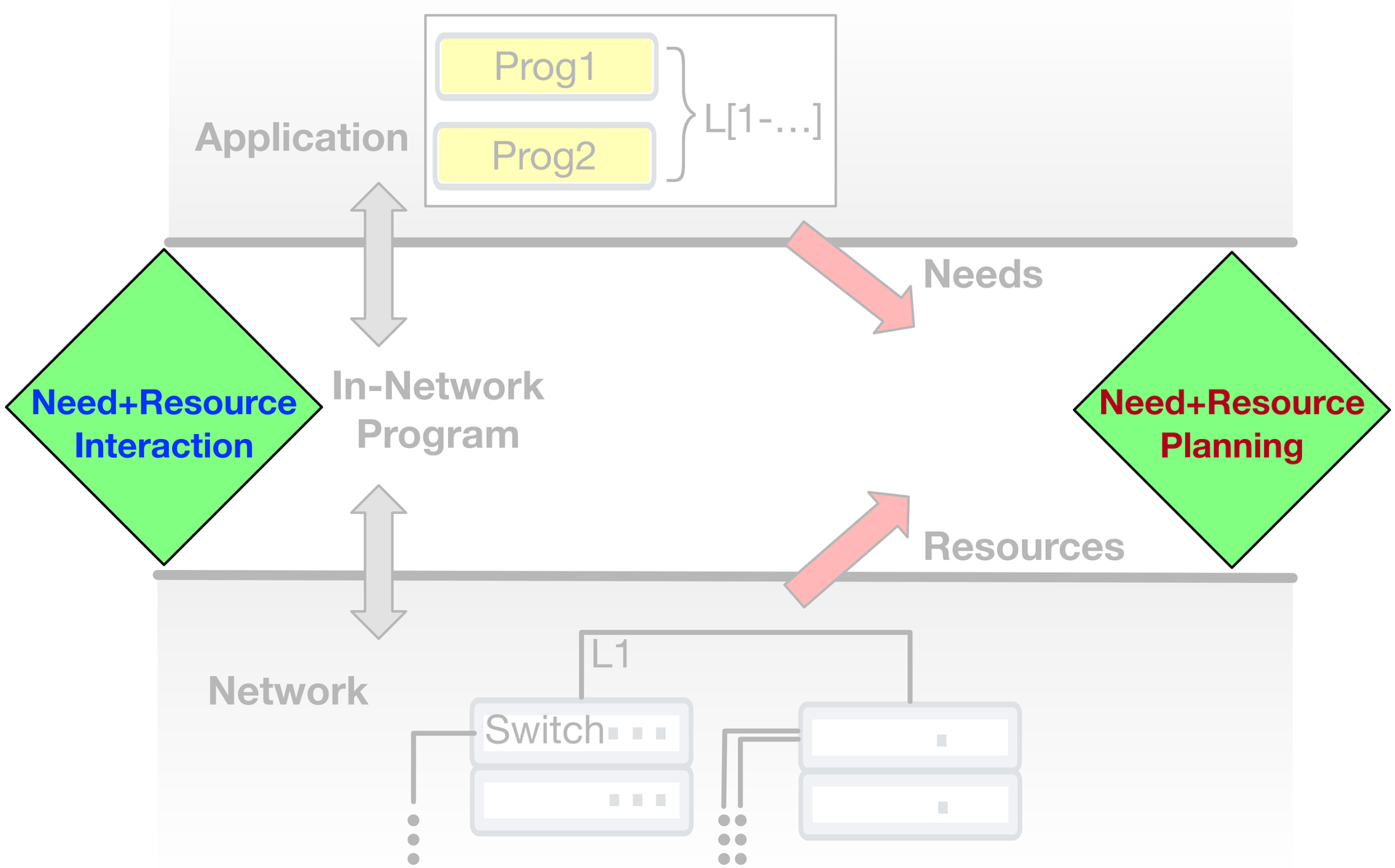
NAI'21

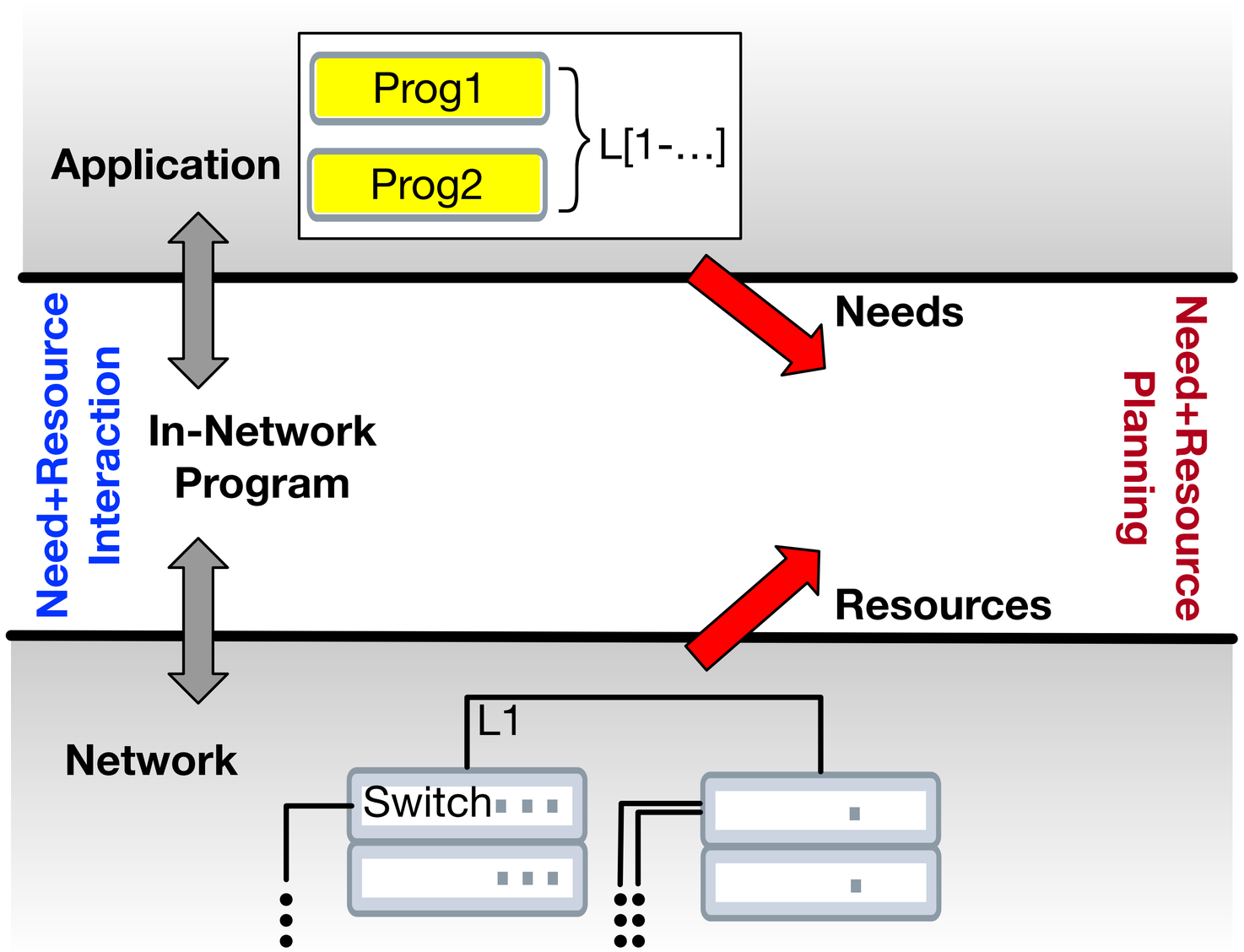

Prog1
Application
Prog2
L1
Network
Switch

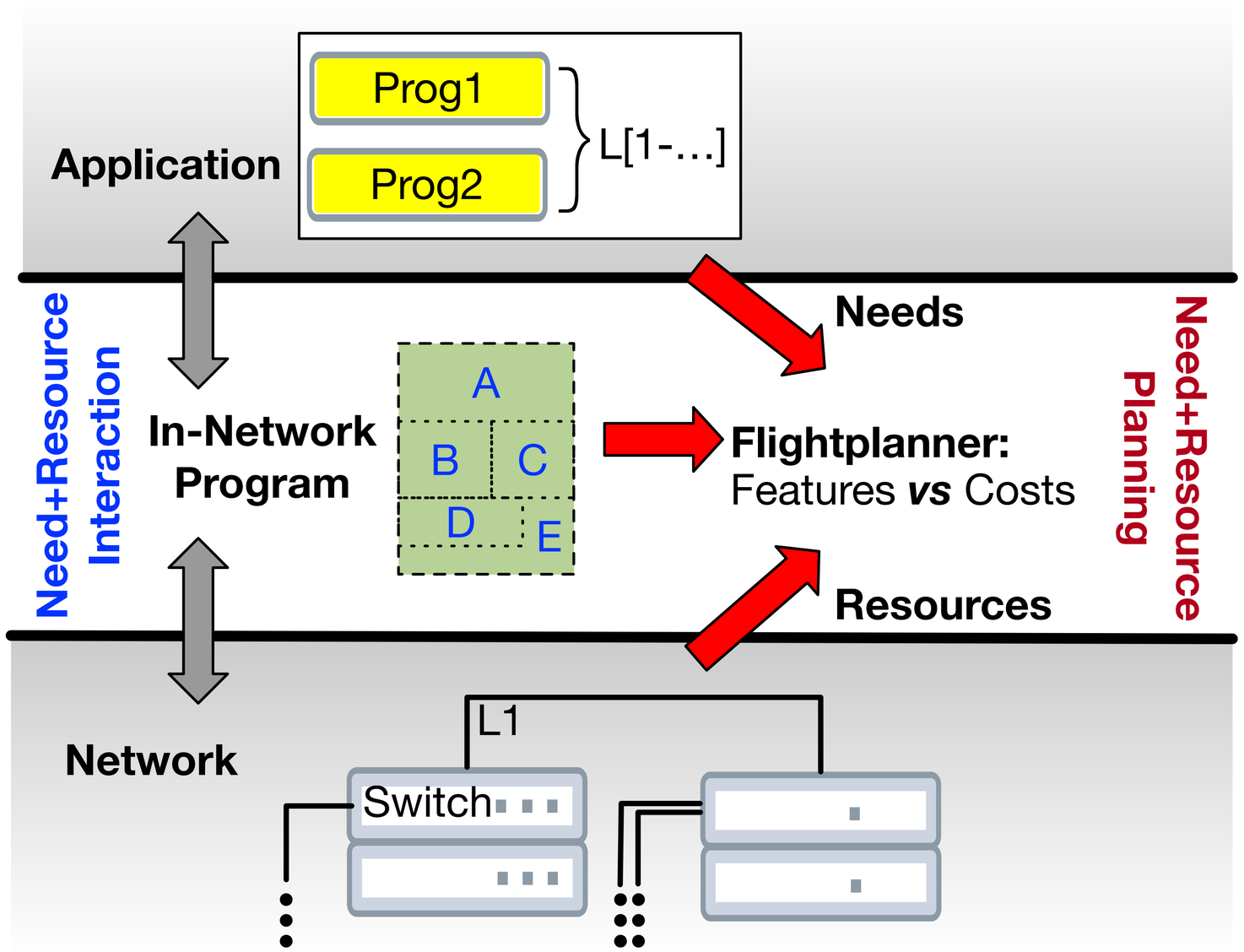

Application
Prog1
Prog2
L[1-…]
Network
L1
Switch


Application
Prog1
Prog2
L[1-…]
Needs
Resources
Network
L1
Switch


Application
Prog1
Prog2
L[1-…]
Needs
In-Network
Program
Resources
Network
L1
Switch


Application
Prog1
Prog2
L[1-…]
Needs
Need+Resource
Interaction
In-Network
Program
Need+Resource
Planning
Resources
Network
L1
Switch


Application
Prog1
Prog2
L[1-…]
Needs
Need+Resource Interaction
In-Network Program
Need+Resource Planning
Resources
Network
L1
Switch


Application
Prog1
Prog2
L[1-…]
Needs
Need+Resource Interaction
In-Network Program
A
B
C
D
E
Flightplanner:
Features *vs* Costs
Need+Resource Planning
Resources
Network
L1
Switch

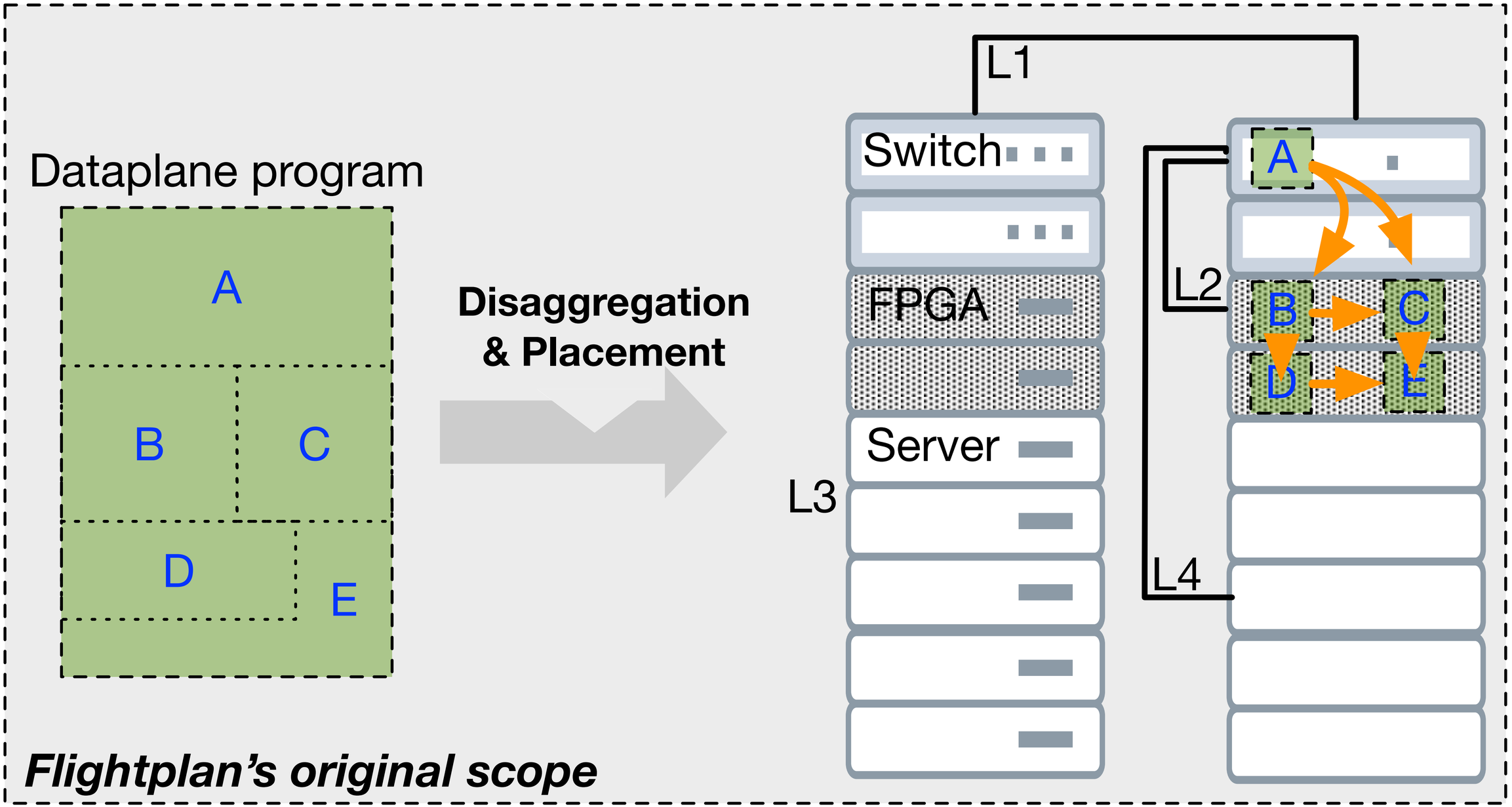
Flightplan: Dataplane Disaggregation and Placement for P4 programs

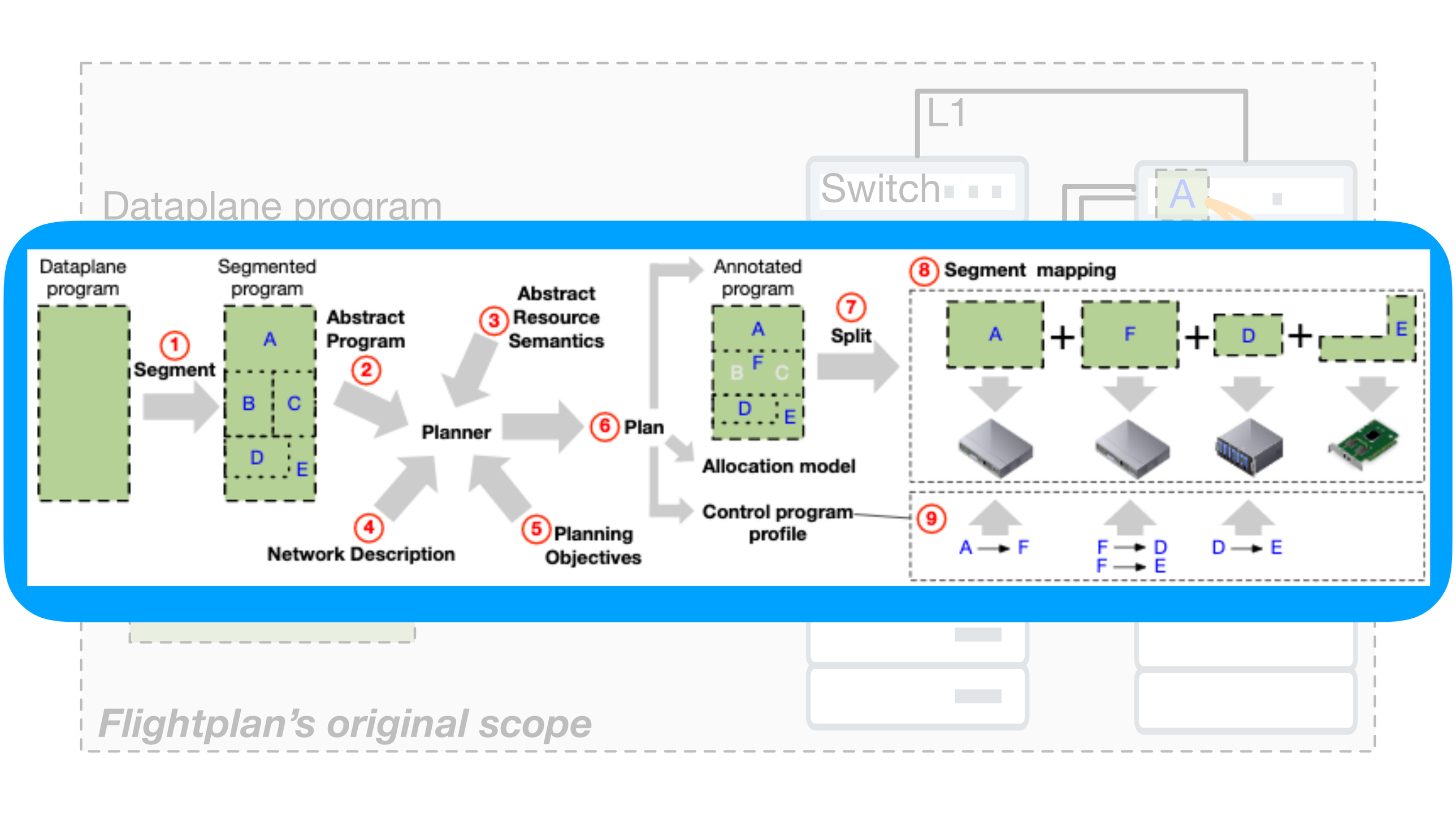
Nik Sultana John Sonchack Hans Giesen Isaac Pedisich Zhaoyang Han
Nishanth Shyamkumar Shivani Burad André DeHon Boon Thau Loo

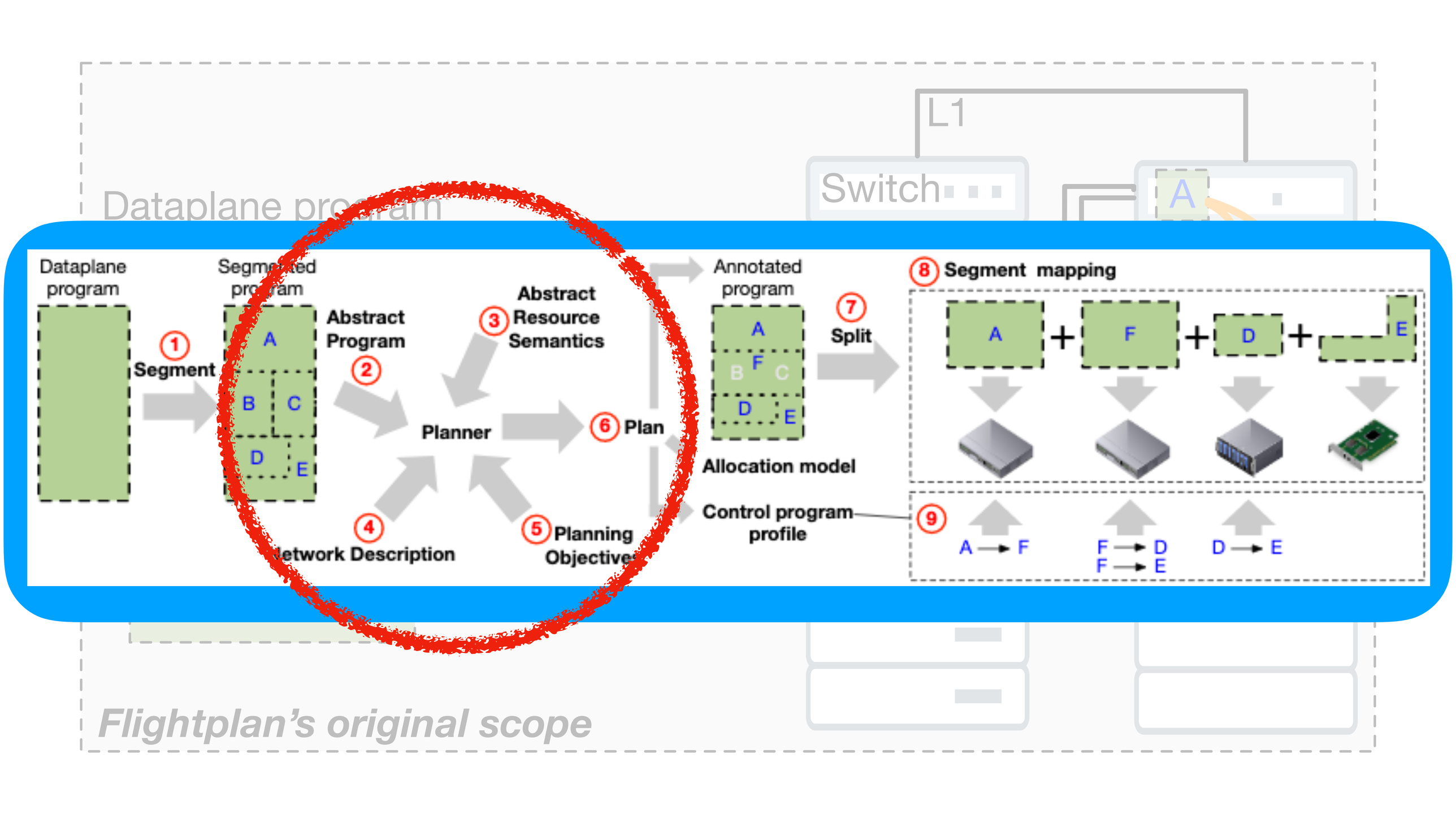
University of Pennsylvania

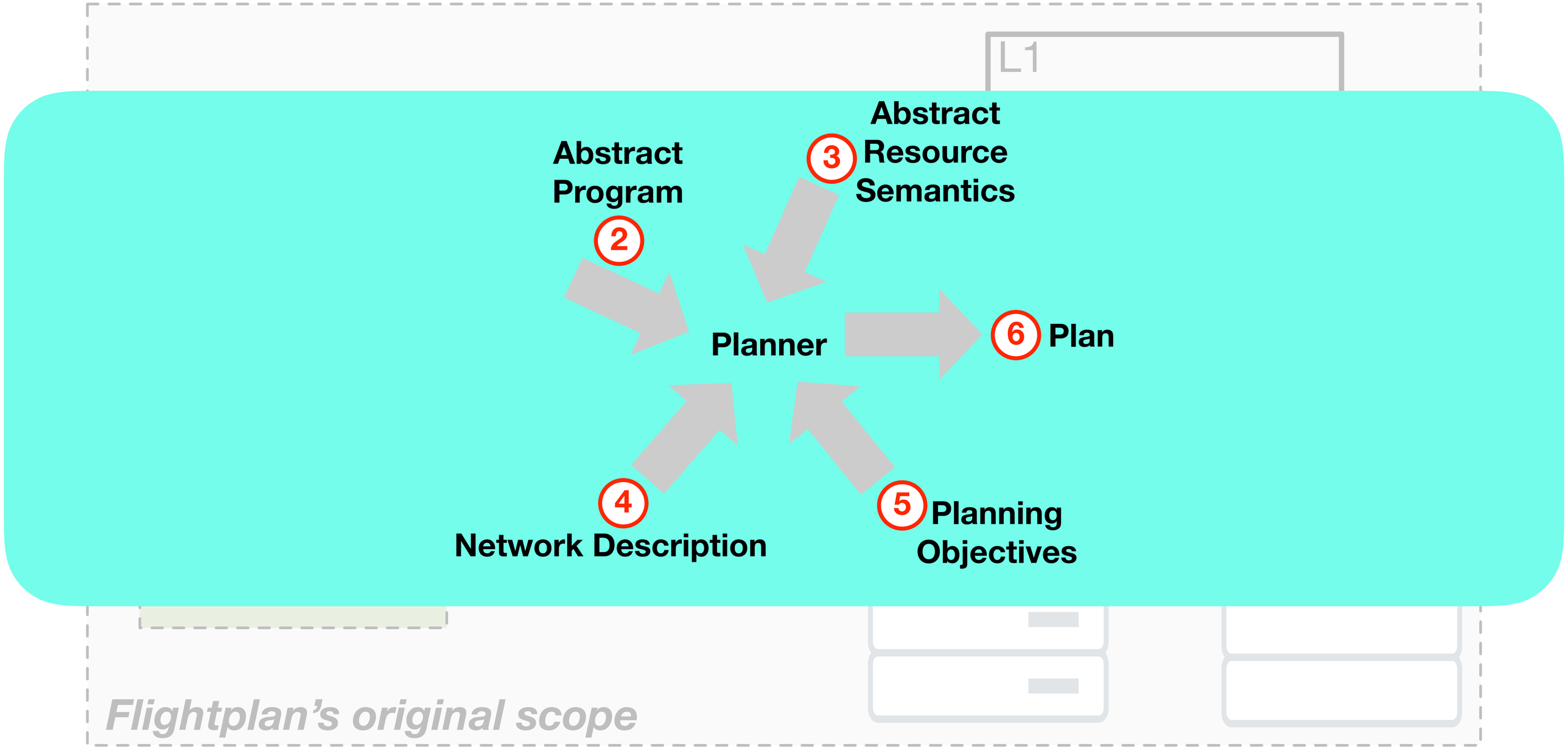
NSDI'21

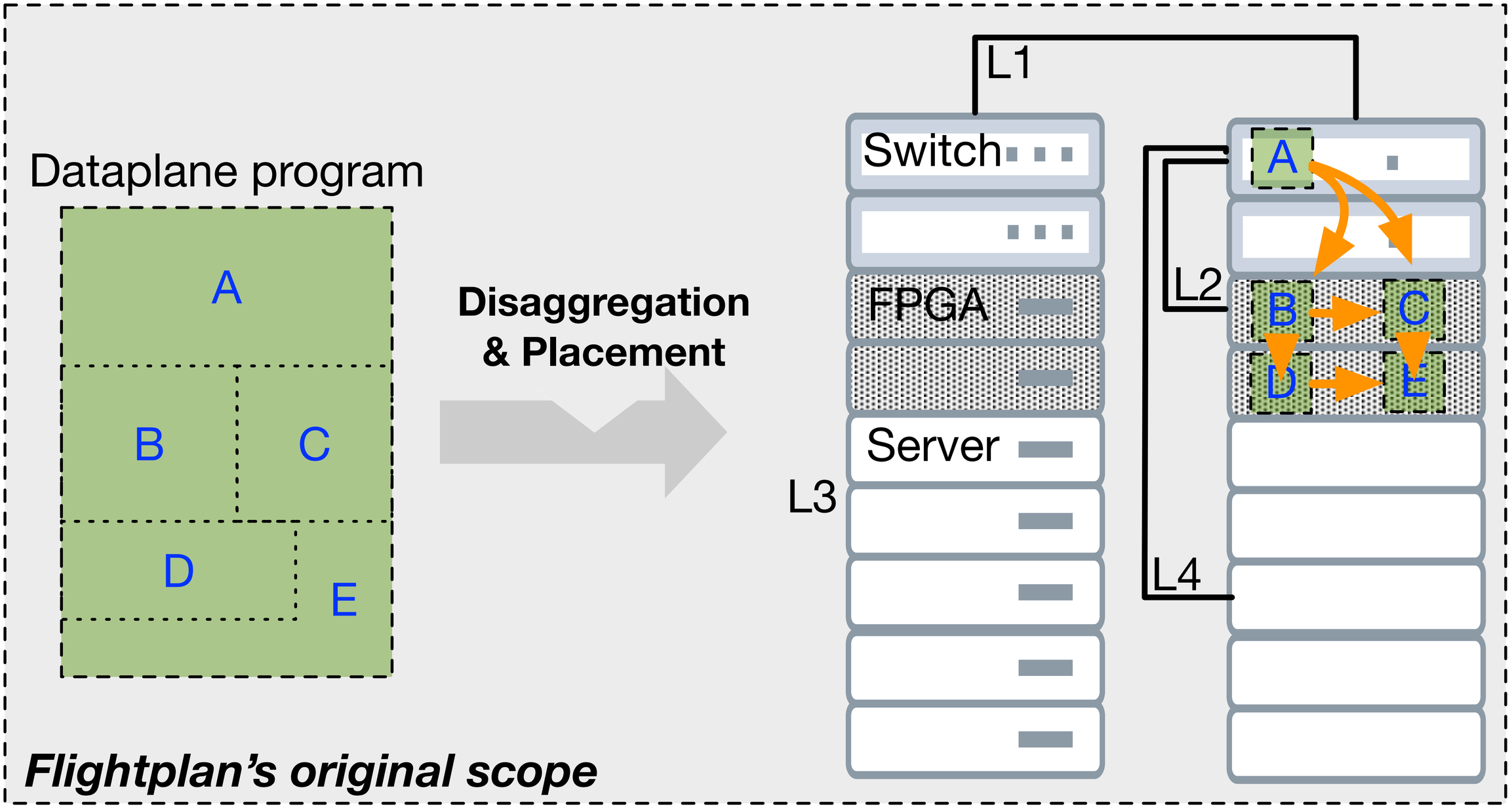
https://flightplan.cis.upenn.edu/

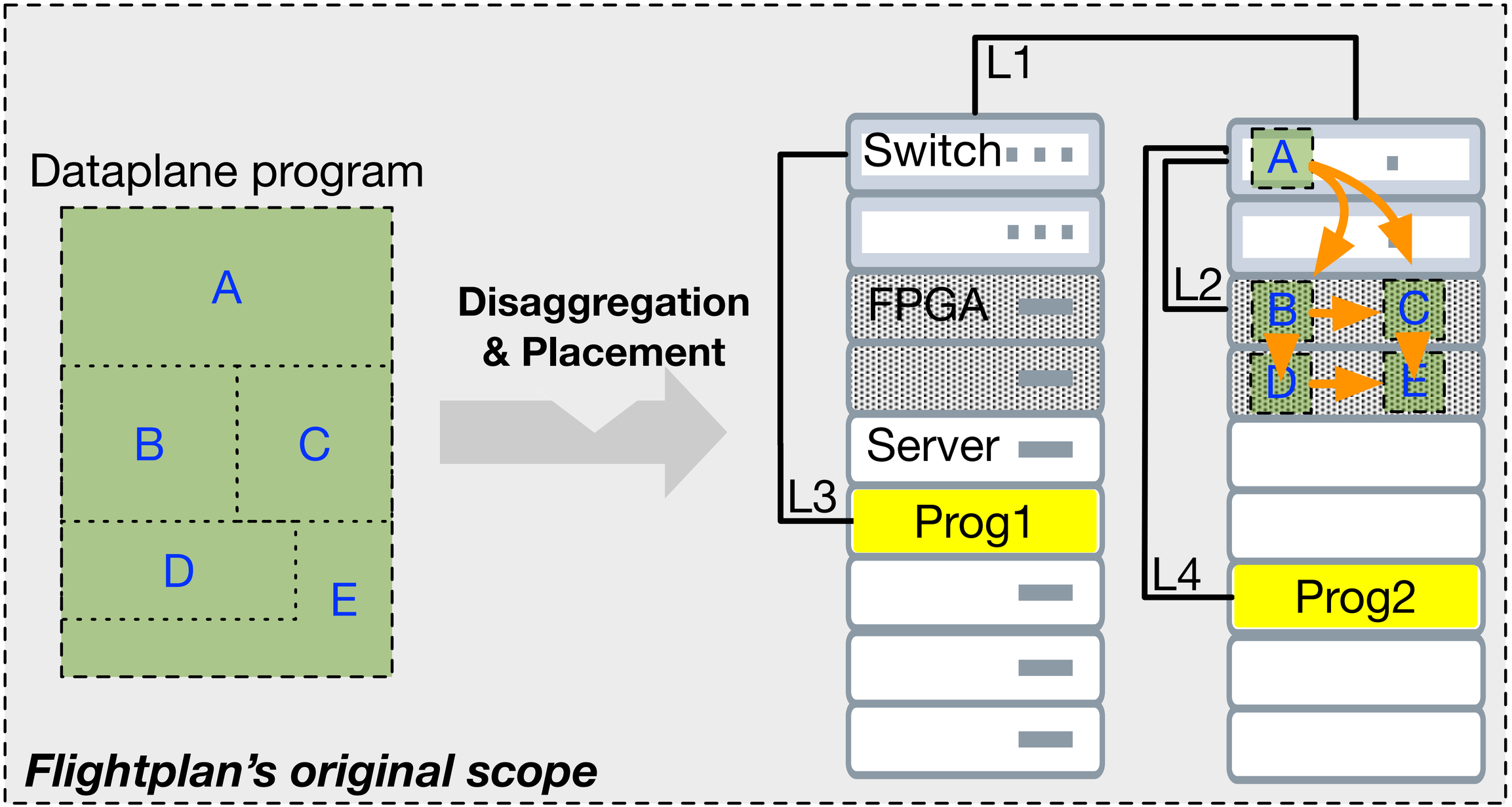

Dataplane program
A
B
C
D
E
Disaggregation
& Placement
L1
Switch
FPGA
Server
L3
L2
L4
A
B
C
D
E
Flightplan's original scope

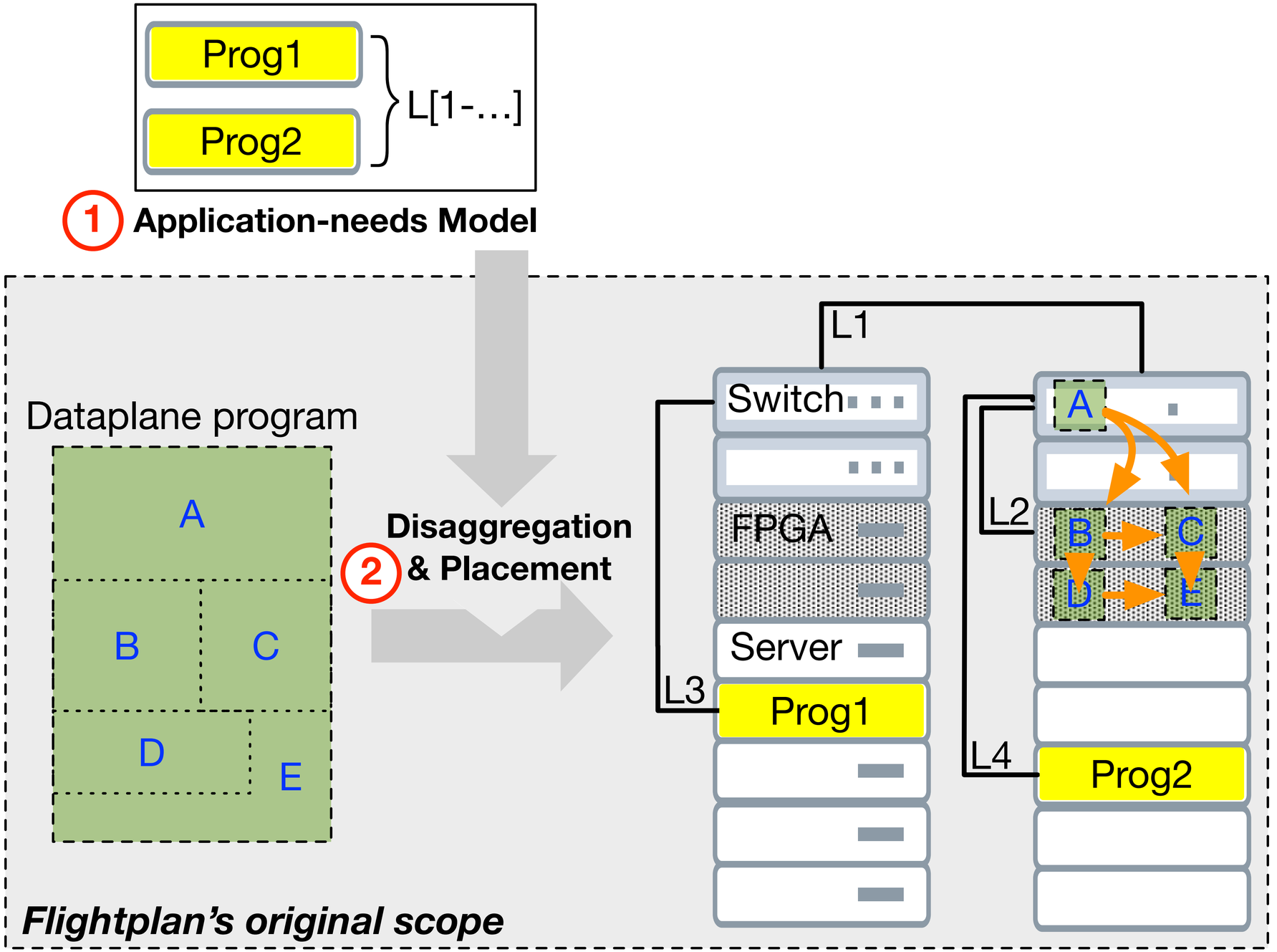

Dataplane program
L1
Switch
A
Dataplane program
Segmented program
Abstract Program
Abstract Resource Semantics
Annotated program
Split
8 Segment mapping
1 Segment
2
3
7
A
B
C
D
E
F
Planner
6 Plan
Allocation model
Control program profile
4 Network Description
5 Planning Objectives
9
A → F
F → D
F → E
D → E
Flightplan's original scope

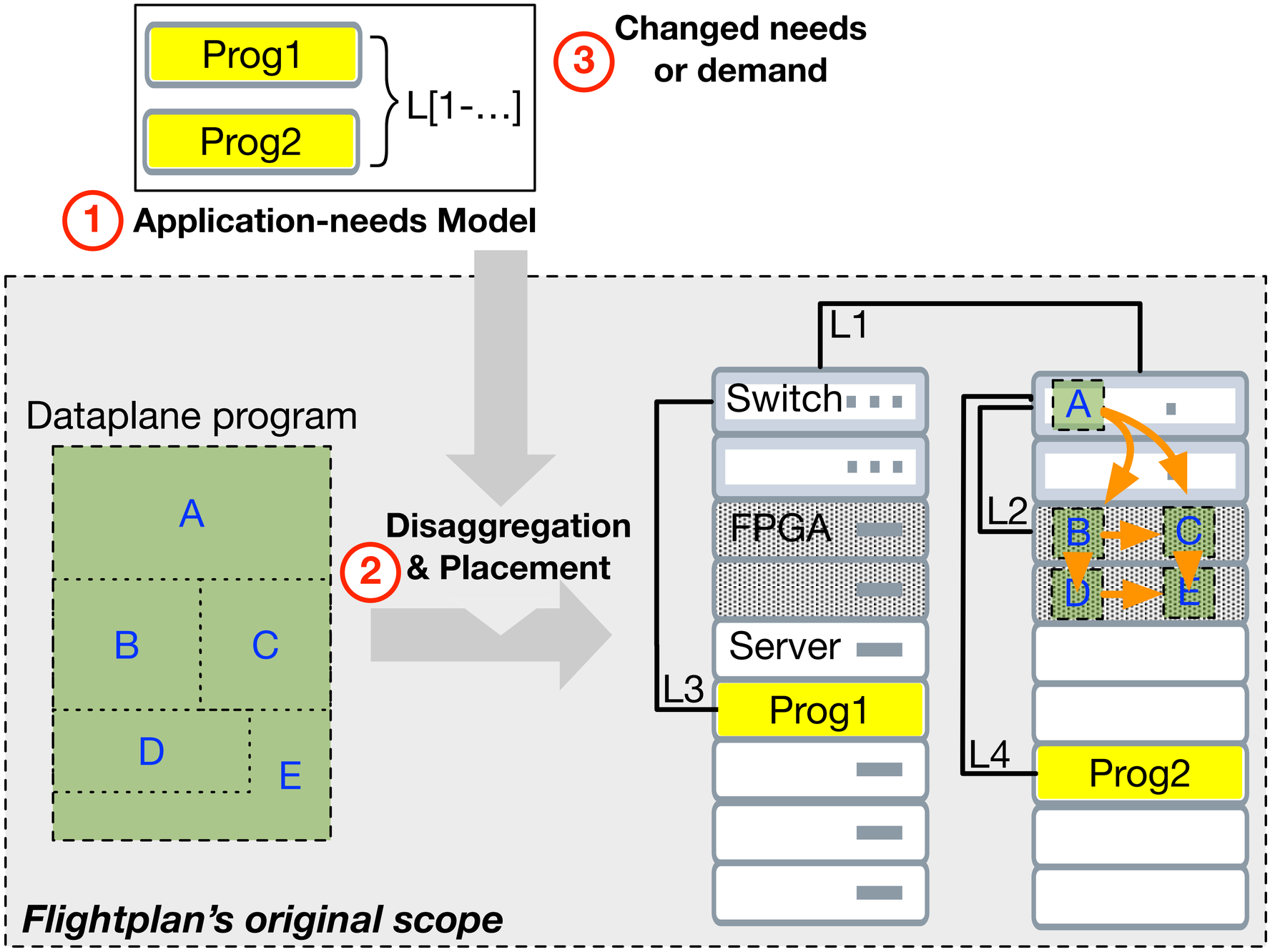

Dataplane program
L1
Switch
A
Dataplane program
Segmented program
1 Segment
A
B
C
D
E
Abstract Program
2
3 Abstract Resource Semantics
Planner
6 Plan
4 Network Description
5 Planning Objectives
Annotated program
A
B
F
C
D
E
Allocation model
Control program profile
7 Split
8 Segment mapping
A
+
F
+
D
+
E
9
A → F
F → D
F → E
D → E
Flightplan's original scope

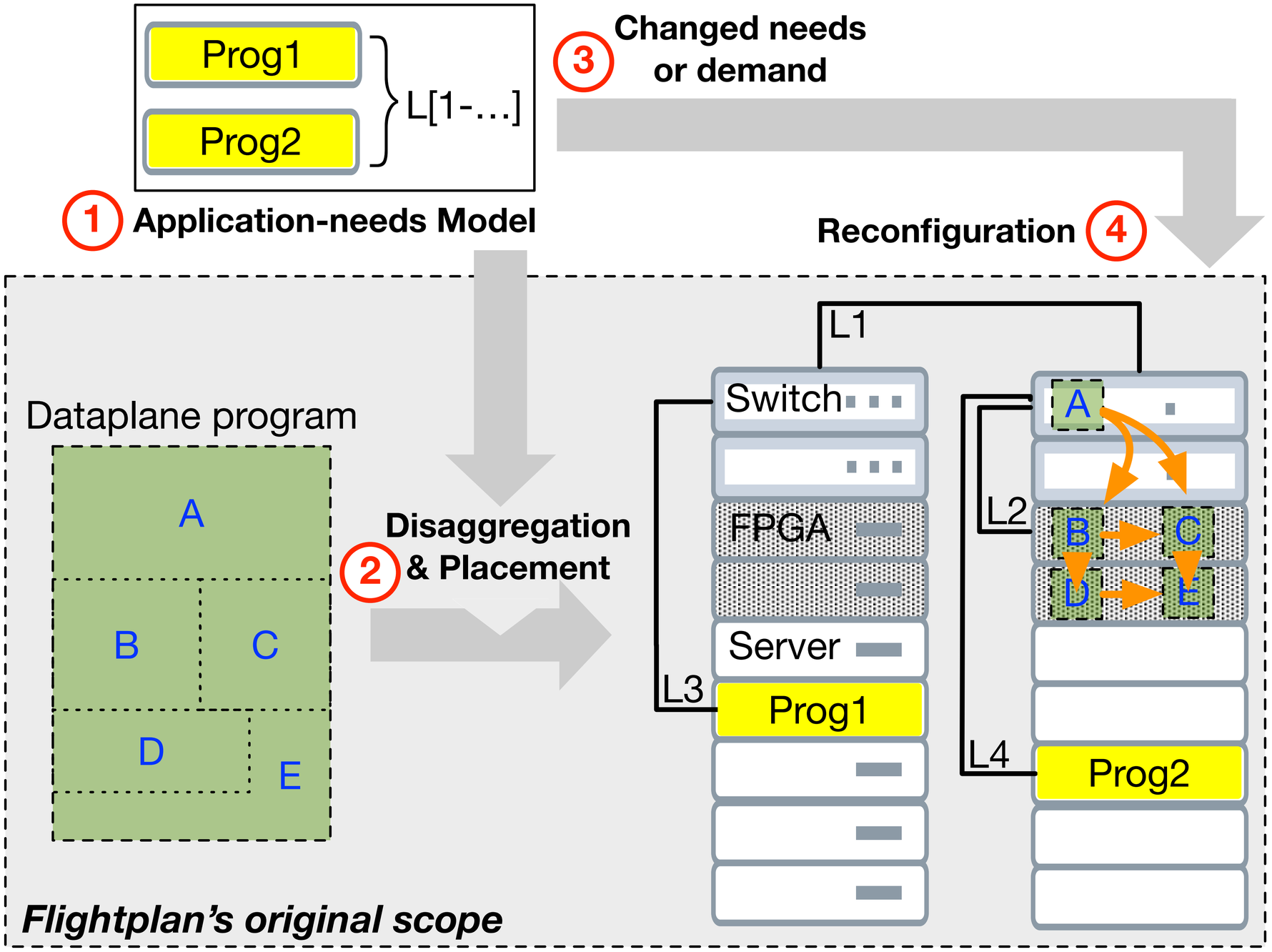

L1
Abstract
Program
2
3
Abstract
Resource
Semantics
Planner
6
Plan
4
Network Description
5
Planning
Objectives
Flightplan's original scope


Dataplane program
A
B
C
D
E
Disaggregation
& Placement
L1
Switch
FPGA
Server
L3
L2
L4
A
B
C
D
E
Flightplan's original scope


Dataplane program
A
B
C
D
E
Disaggregation
& Placement
L1
Switch
FPGA
Server
L3
Prog1
A
L2
B
C
D
E
L4
Prog2
Flightplan's original scope


Prog1
Prog2
L[1-…]
1 Application-needs Model
Dataplane program
A
B
C
D
E
2 Disaggregation & Placement
L1
Switch
FPGA
Server
Prog1
L3
A
L2
B
C
D
E
L4
Prog2
Flightplan's original scope


Prog1
Prog2
L[1-…]
3
Changed needs
or demand
1
Application-needs Model
Dataplane program
A
B
C
D
E
2
Disaggregation
& Placement
L1
Switch
FPGA
Server
L3
Prog1
A
L2
B
C
D
E
L4
Prog2
Flightplan's original scope


Prog1
Prog2
L[1-…]
1 Application-needs Model
3 Changed needs or demand
Reconfiguration 4
Dataplane program
A
B
C
D
E
2 Disaggregation & Placement
L1
Switch
FPGA
Server
L3
Prog1
L2
L4
Prog2
Flightplan's original scope

$$\text{Xeon2450-1}:$$
$$\pi_{\text{Requires}} = \left\{\text{Rate} \leq 10^{10}\right\}, \qquad \pi_{\text{Provides}} = \{\text{CPU}\},$$
$$\text{Ports} = \left\{\ 1 \mapsto \left\{\pi_{\text{Requires}} = \left\{\text{Rate} \leq 10^{10}\right\},\ \pi_{\text{Provides}} = \{\}\right\}\ \right\}$$

Prog1 mode:**M1** :

$\pi_{\text{Requires}} = \{\text{Rate} \leq 1^{10}\}$, $\pi_{\text{Provides}} = \{$ Traffic $\}$,

Ports = {(Xeon2450-1, 1)}, Peers = {Prog2},

OnPath = { InNetProg1 }

Prog1 mode:M1 :

$$\pi_{\text{Requires}} = \{\text{Rate} \leq 1^{10}\}, \quad \pi_{\text{Provides}} = \{\text{Traffic}\},$$
$$\text{Ports} = \{(\text{Xeon2450-1}, 1)\}, \quad \text{Peers} = \{\text{Prog2}\},$$
$$\text{OnPath} = \{\text{InNetProg1}\}$$

Prog1 = mcdClient
Prog2 = mcdServer
Traffic = UDP(200)
M1 = LowActivity (lower π_{Requires})

Prog1 mode:M1 :

$$\pi_{\text{Requires}} = \{\text{Rate} \leq 1^{10}\}, \quad \pi_{\text{Provides}} = \{\text{Traffic}\},$$
$$\text{Ports} = \{(\text{Xeon2450-1}, 1)\}, \quad \text{Peers} = \{\text{Prog2}\},$$
$$\text{OnPath} = \{\text{InNetProg1}\}$$

Prog1 = mcdClient
Prog2 = mcdServer
Traffic = UDP(200)
M1 = LowActivity (lower π_{Requires})
InNetProg1 = (None)

Prog1 mode:M1 :

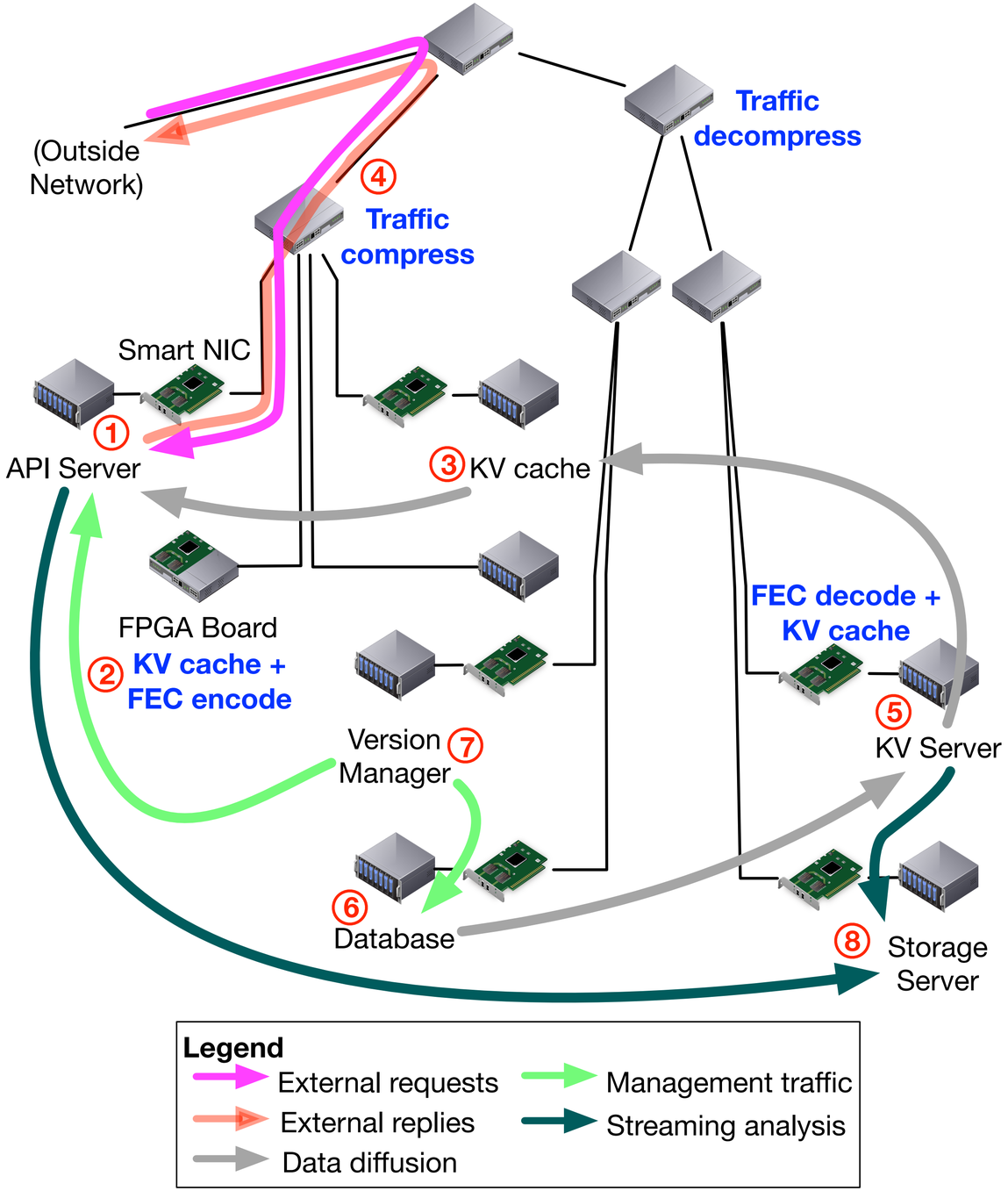
$$\pi_{\text{Requires}} = \{\text{Rate} \leq 1^{10}\}, \quad \pi_{\text{Provides}} = \{\text{Traffic}\},$$

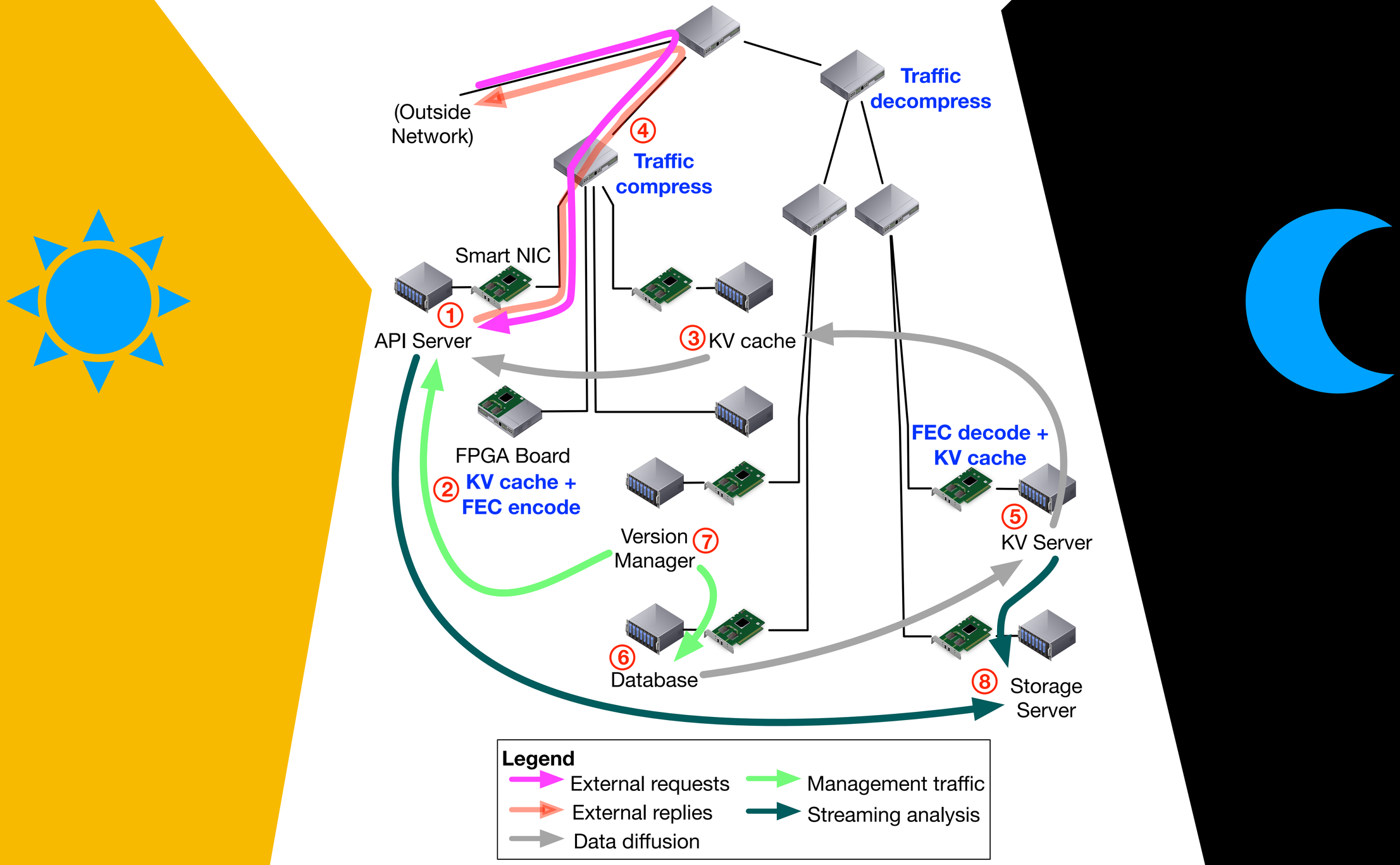
$$\text{Ports} = \{(\text{Xeon2450-1}, 1)\}, \quad \text{Peers} = \{\text{Prog2}\},$$

$$\text{OnPath} = \{\text{InNetProg1}\}$$

Prog1 = mcdClient
Prog2 = mcdServer
Traffic = UDP(200)
M1 = LowActivity
InNetProg1 = (None)

Prog1 = mcdClient
Prog2 = mcdServer
Traffic = UDP(200)
M1 = HighActivity (higher π_{Requires})
InNetProg1 = MCD_Cache


(Outside Network)
Traffic decompress
4
Traffic compress
Smart NIC
1
API Server
3 KV cache
FPGA Board
2 KV cache + FEC encode
FEC decode + KV cache
5
KV Server
Version Manager
7
6
Database
8
Storage Server
Legend
External requests
External replies
Data diffusion
Management traffic
Streaming analysis

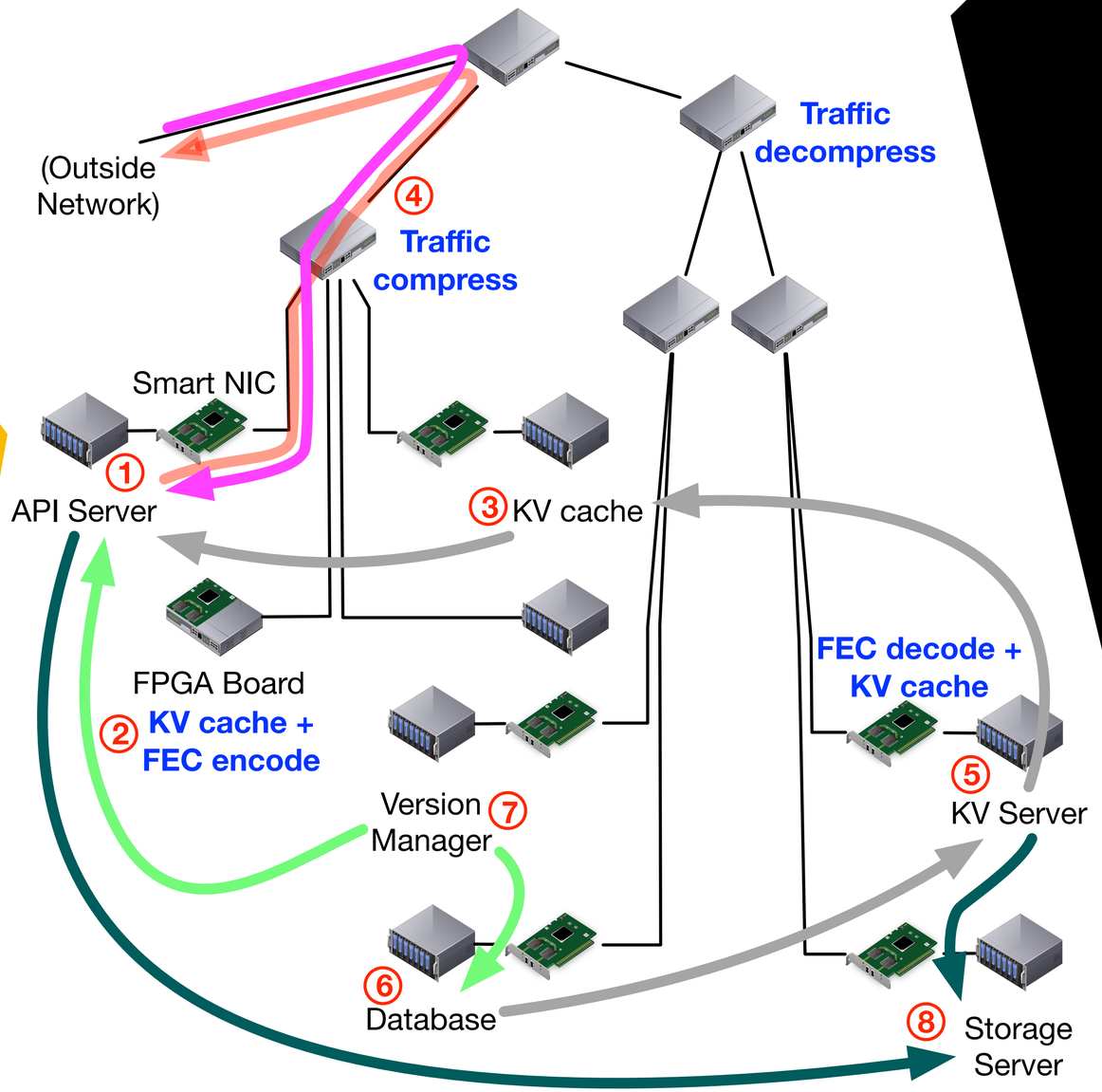

(Outside Network)
Traffic decompress
④
Traffic compress
Smart NIC
①
API Server
③ KV cache
FPGA Board
② KV cache + FEC encode
FEC decode + KV cache
⑤
KV Server
Version Manager
⑦
⑥
Database
⑧
Storage Server
Legend
External requests
External replies
Data diffusion
Management traffic
Streaming analysis

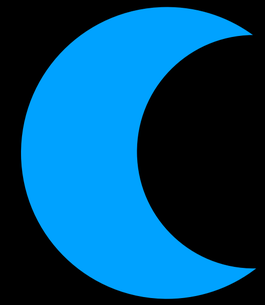
Prog1 mode:M1 :

$$\pi_{\text{Requires}} = \{\text{Rate} \leq 1^{10}\}, \quad \pi_{\text{Provides}} = \{\text{Traffic}\},$$
$$\text{Ports} = \{(\text{Xeon2450-1}, 1)\}, \quad \text{Peers} = \{\text{Prog2}\},$$
$$\text{OnPath} = \{\text{InNetProg1}\}$$

Prog1 = mcdClient
Prog2 = mcdServer
Traffic = UDP(200)
M1 = HighActivity, NightTime
InNetProg1 = MCD_Cache

HTTP
Memcached
rsync





Leveraging In-Network Application Awareness

Nik Sultana
University of Pennsylvania

https://www.seas.upenn.edu/~nsultana

NAI'21