

Data Acquisition for Real-time Decision-making under Freshness Constraints

*Shaohan Hu, Shuochao Yao, Haiming Jin, Yiran Zhao, Yitao Hu,
Xiaochen Liu, Nooreddin Naghibolhosseini, Shen Li, Akash Kapoor,
William Dron, Lu Su, Amotz Bar-Noy, Pedro Szekely, Ramesh Govindan,
Reginald Hobbs, Tarek Abdelzaher*



- ✦ **Chaotic, dynamic environments**
- ✦ **In response, need to decide...**
 - ✦ **what course of action to take**
 - ✦ **how to carry it out**

**Shelter has
vacancy**

**Comm.
center up**

**Road not
blocked**

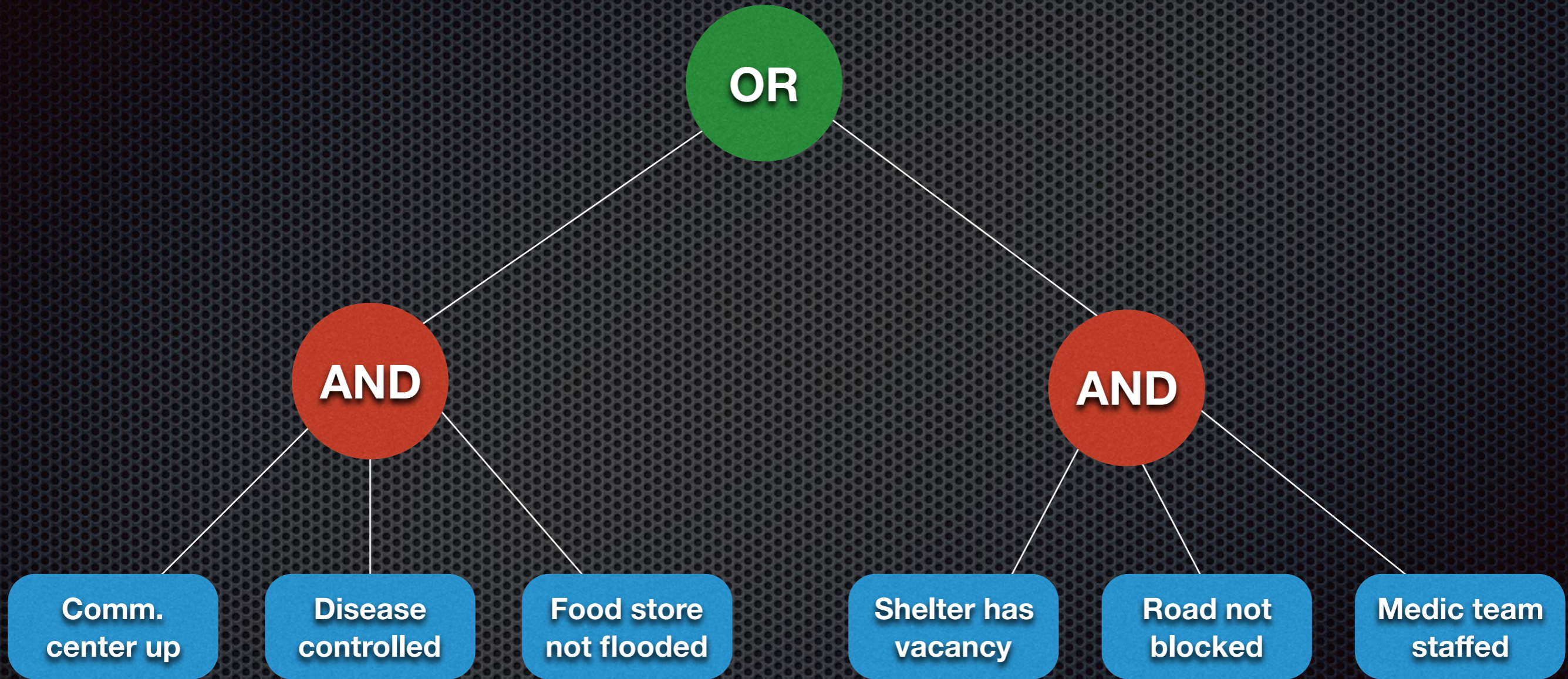
DECISION

**Disease
controlled**

**Medic team
staffed**

**Food store
not flooded**

DECISION



Decision Maker

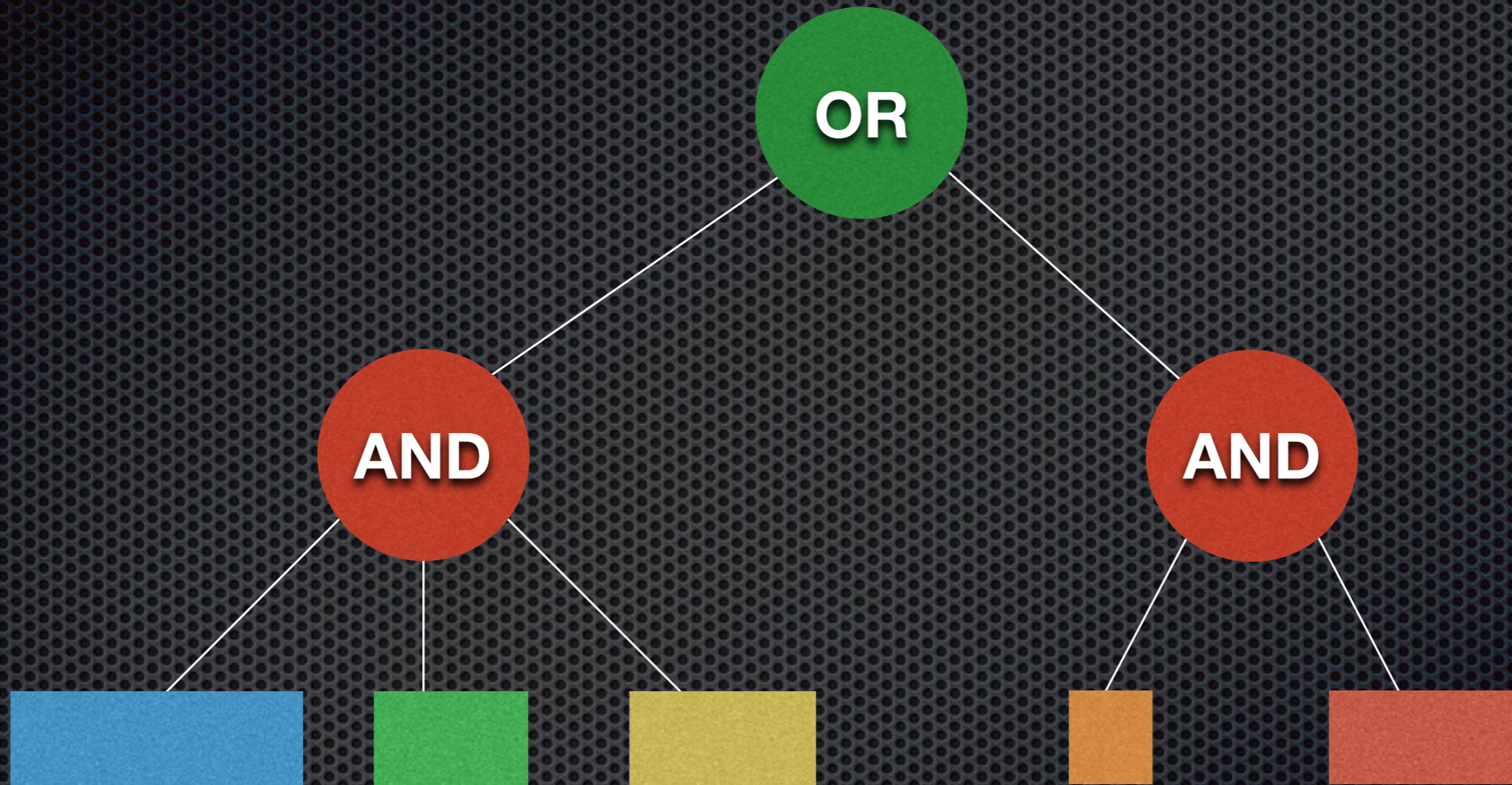




- ✦ **Resource Limitation**
- ✦ **Environment Dynamics**

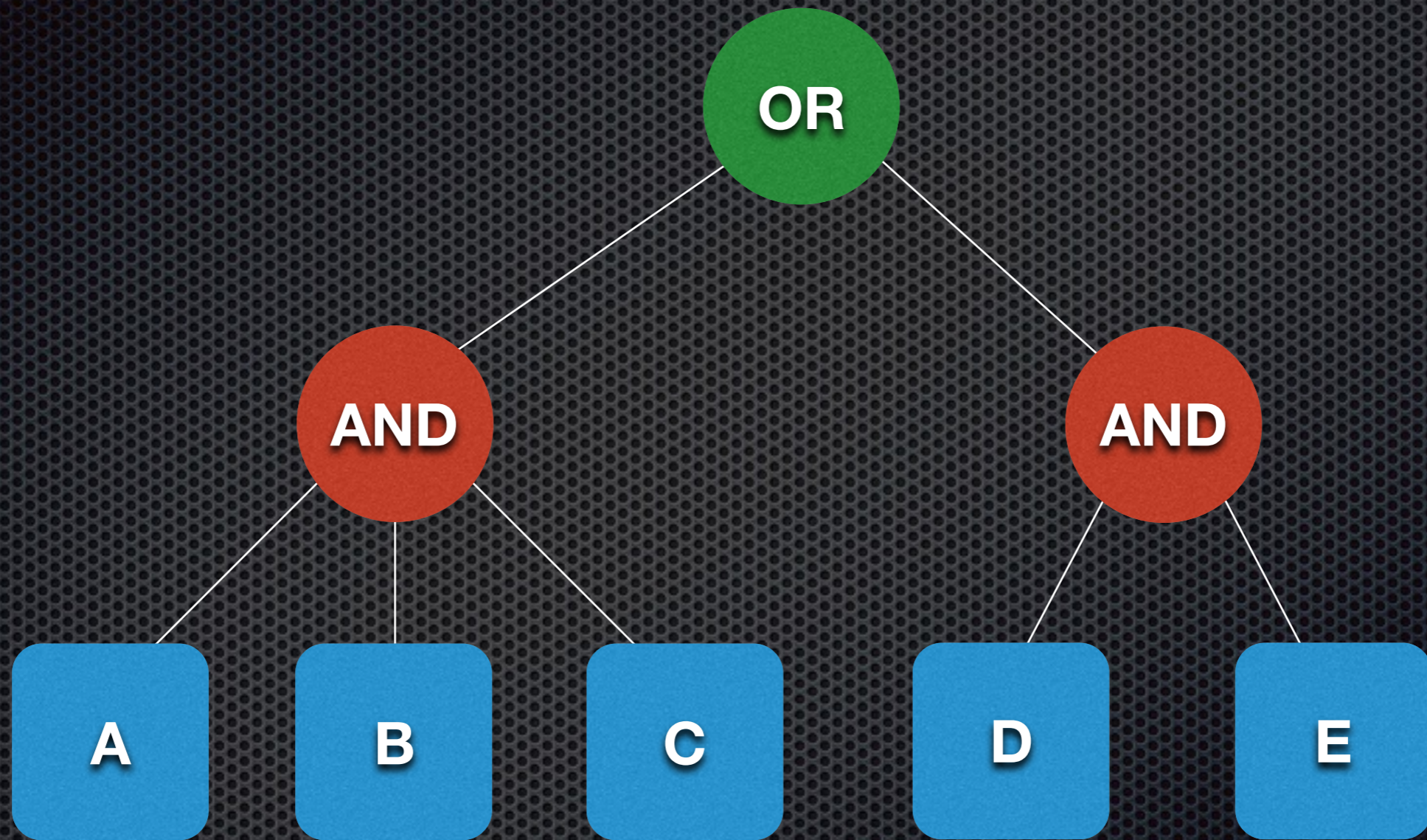
1





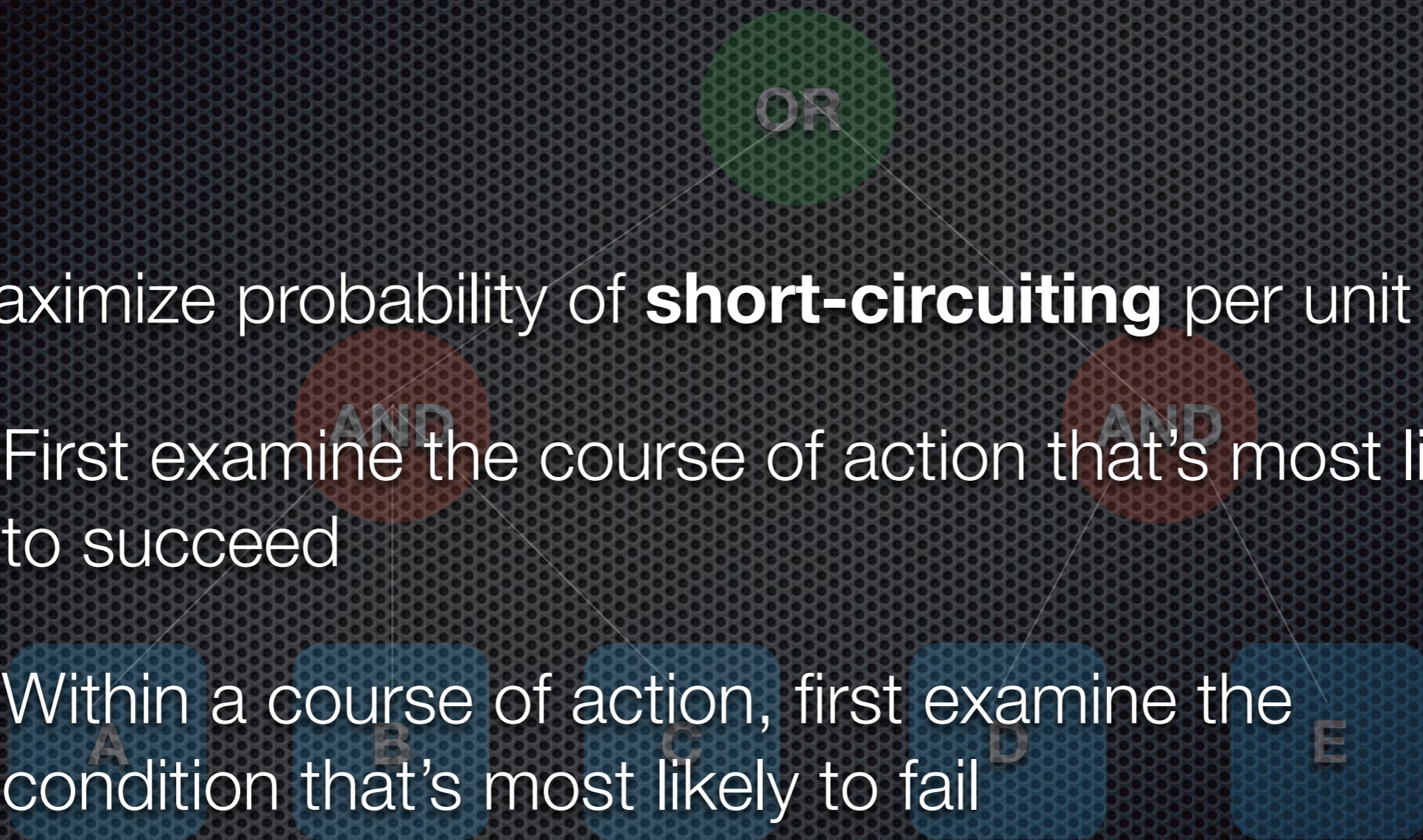


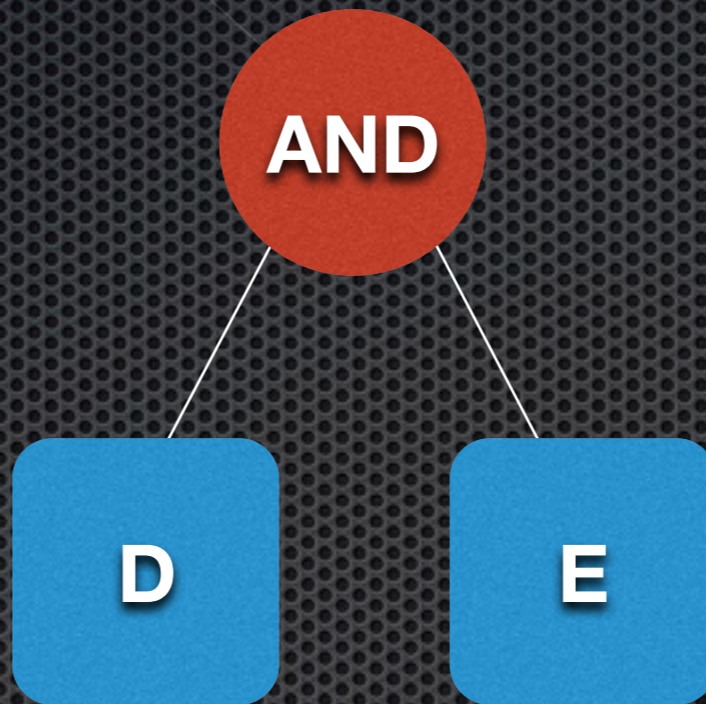
Order?



Intuition

- ✦ Maximize probability of **short-circuiting** per unit cost
 - ✦ First examine the course of action that's most likely to succeed
 - ✦ Within a course of action, first examine the condition that's most likely to fail





C

4

3

P(True)

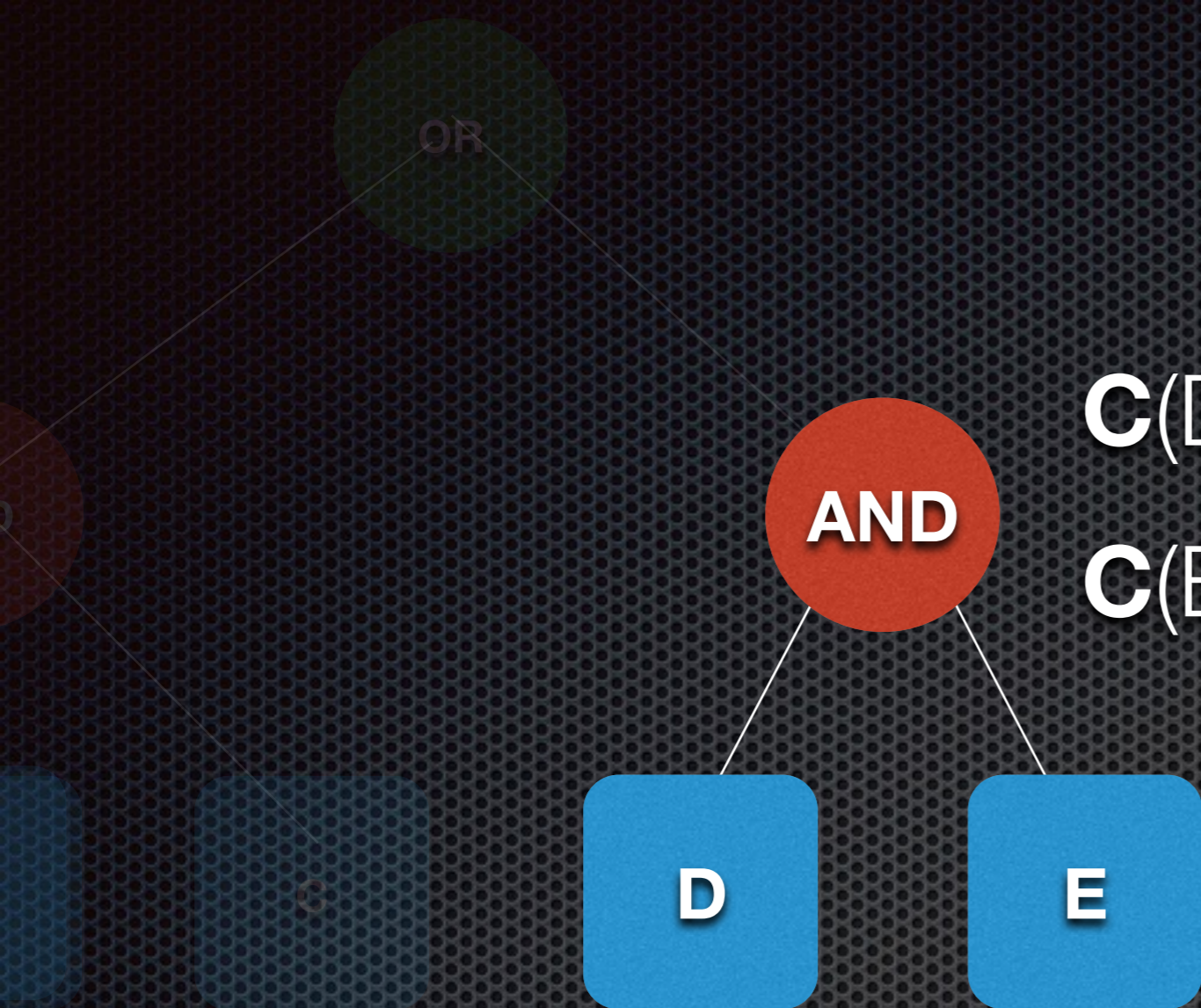
0.3

0.6

R

$(1-0.3)/4$
 $=0.175$

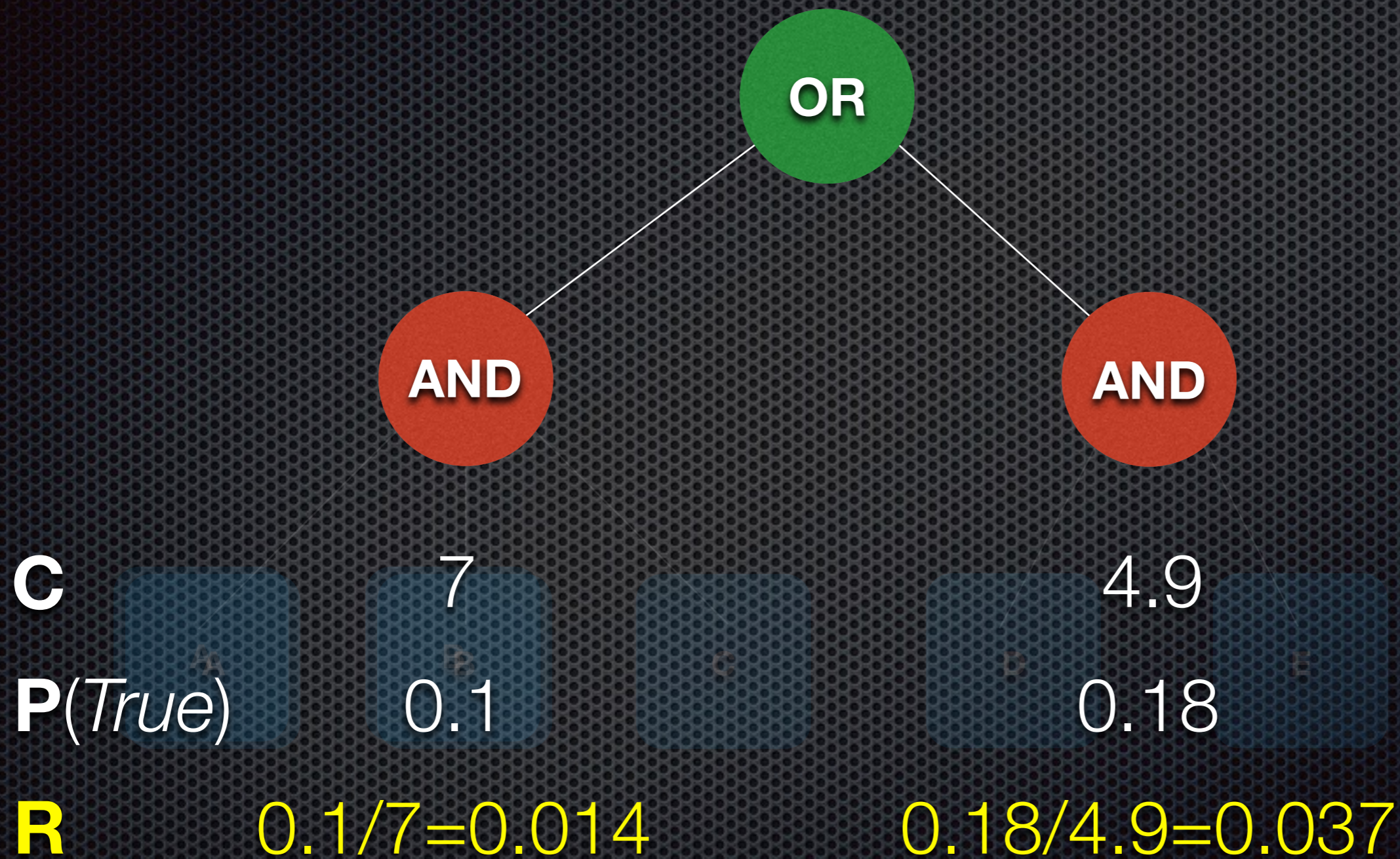
$(1-0.6)/3$
 $=0.133$



$$C(D \rightarrow E) = 4 + 0.3(3) = 4.9$$

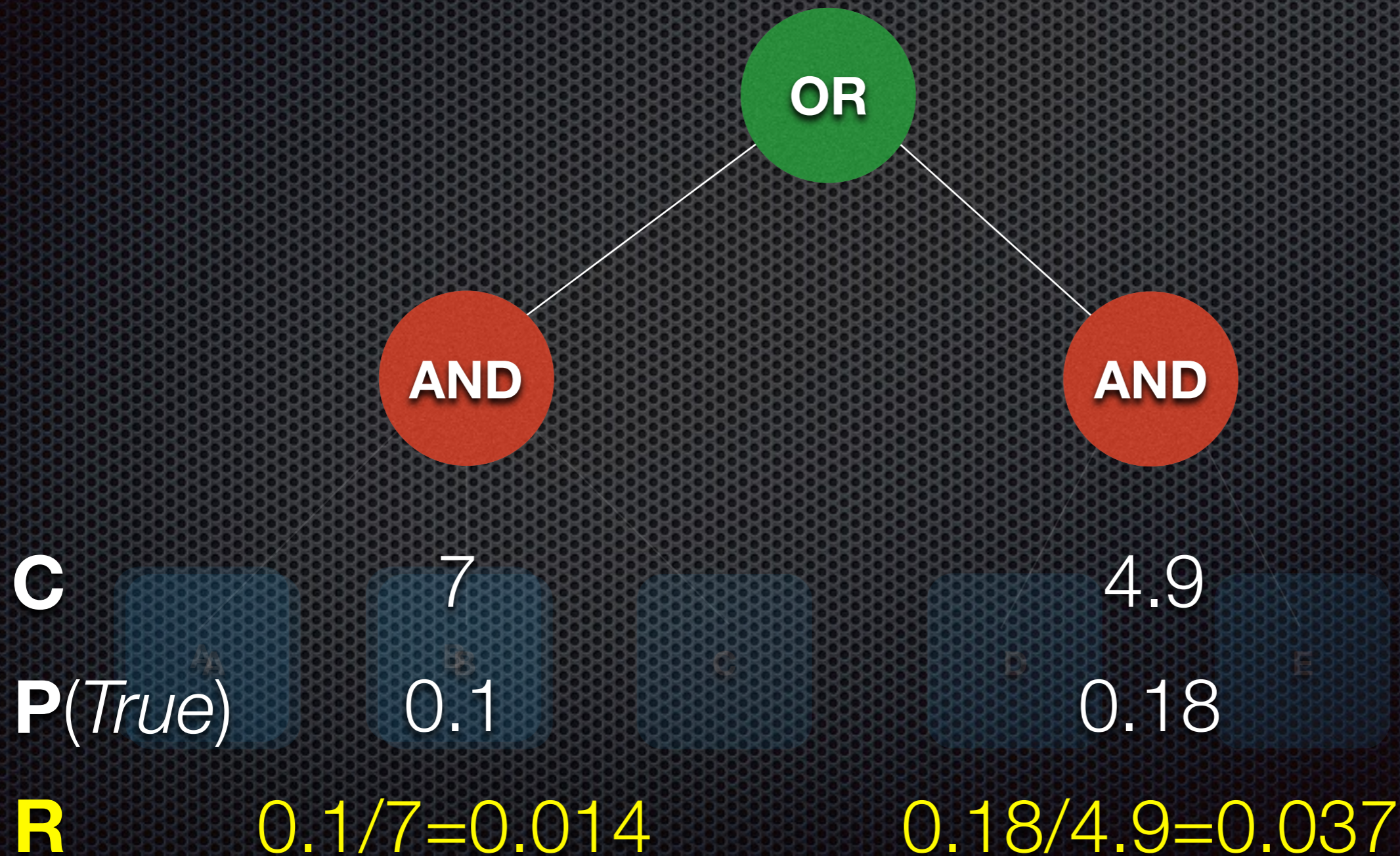
$$C(E \rightarrow D) = 3 + 0.6(4) = 5.4$$

C	4	3
P(True)	0.3	0.6
R	$(1-0.3)/4$ =0.175	$(1-0.6)/3$ =0.133



$$C(L \rightarrow R) = 7 + (1 - 0.1) \times 4.9 = \mathbf{11.41}$$

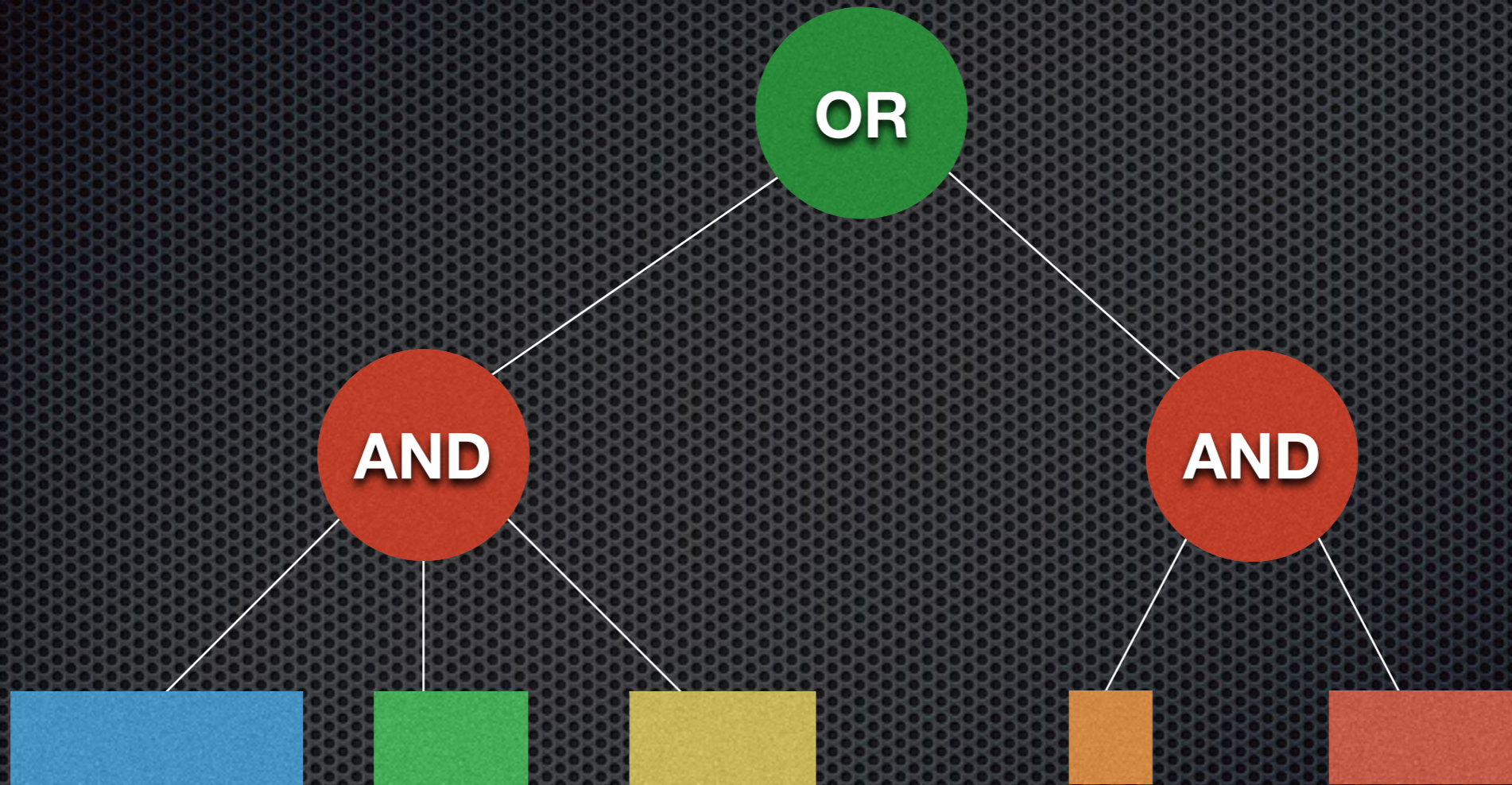
$$C(R \rightarrow L) = 4.9 + (1 - 0.18) \times 7 = \mathbf{10.64}$$

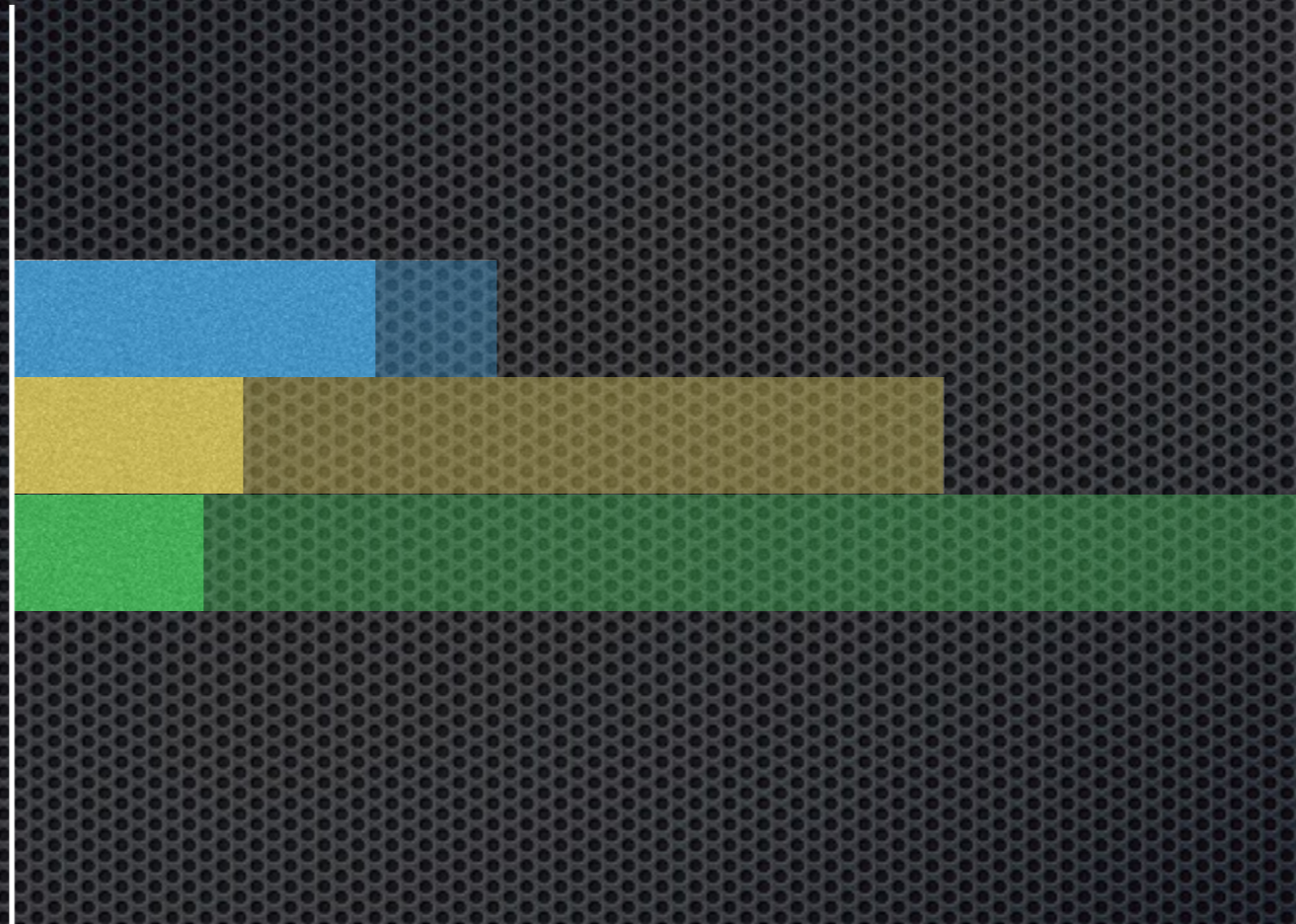


key: $\frac{\text{Short-circuit probability}}{\text{cost}}$

2



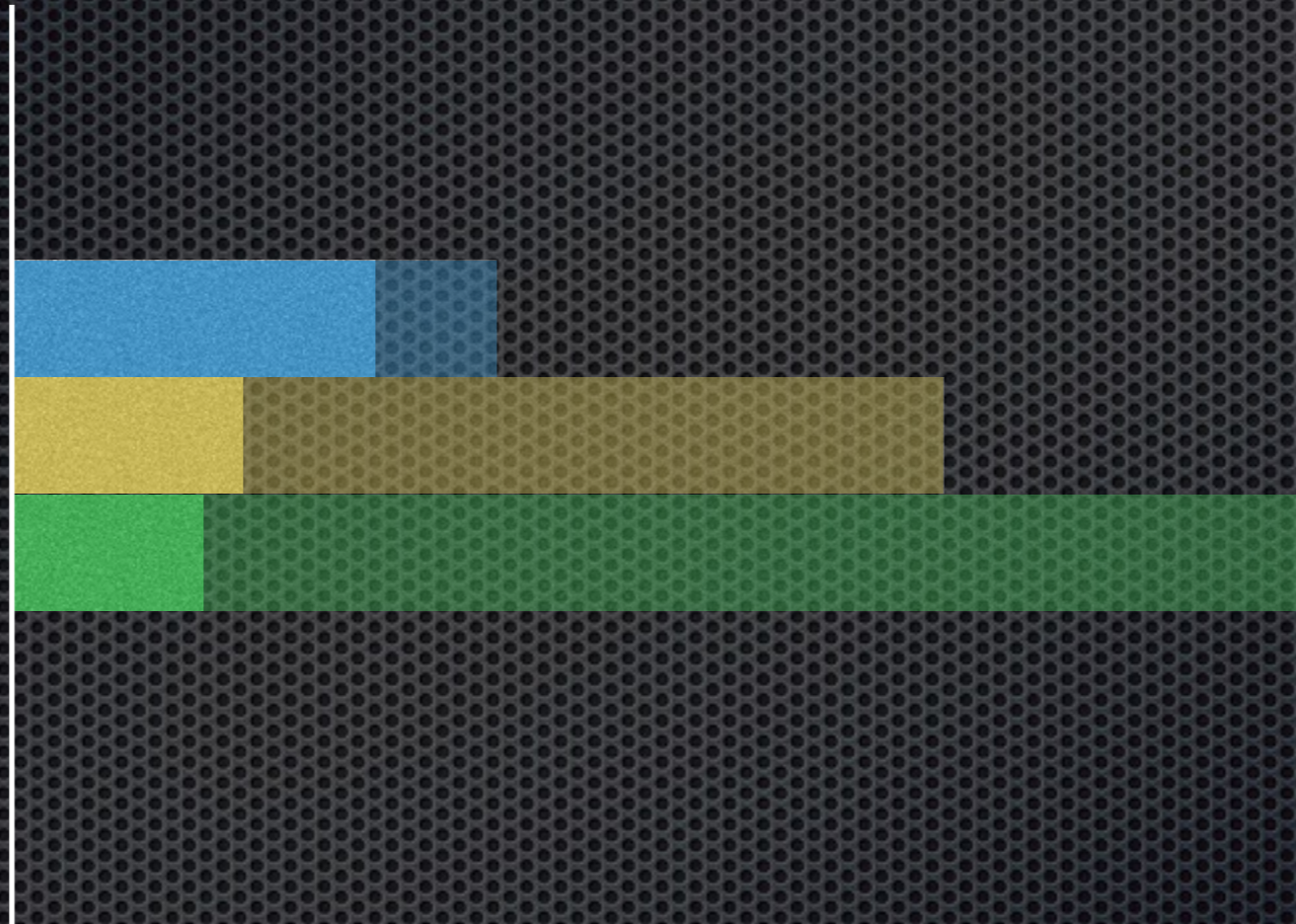


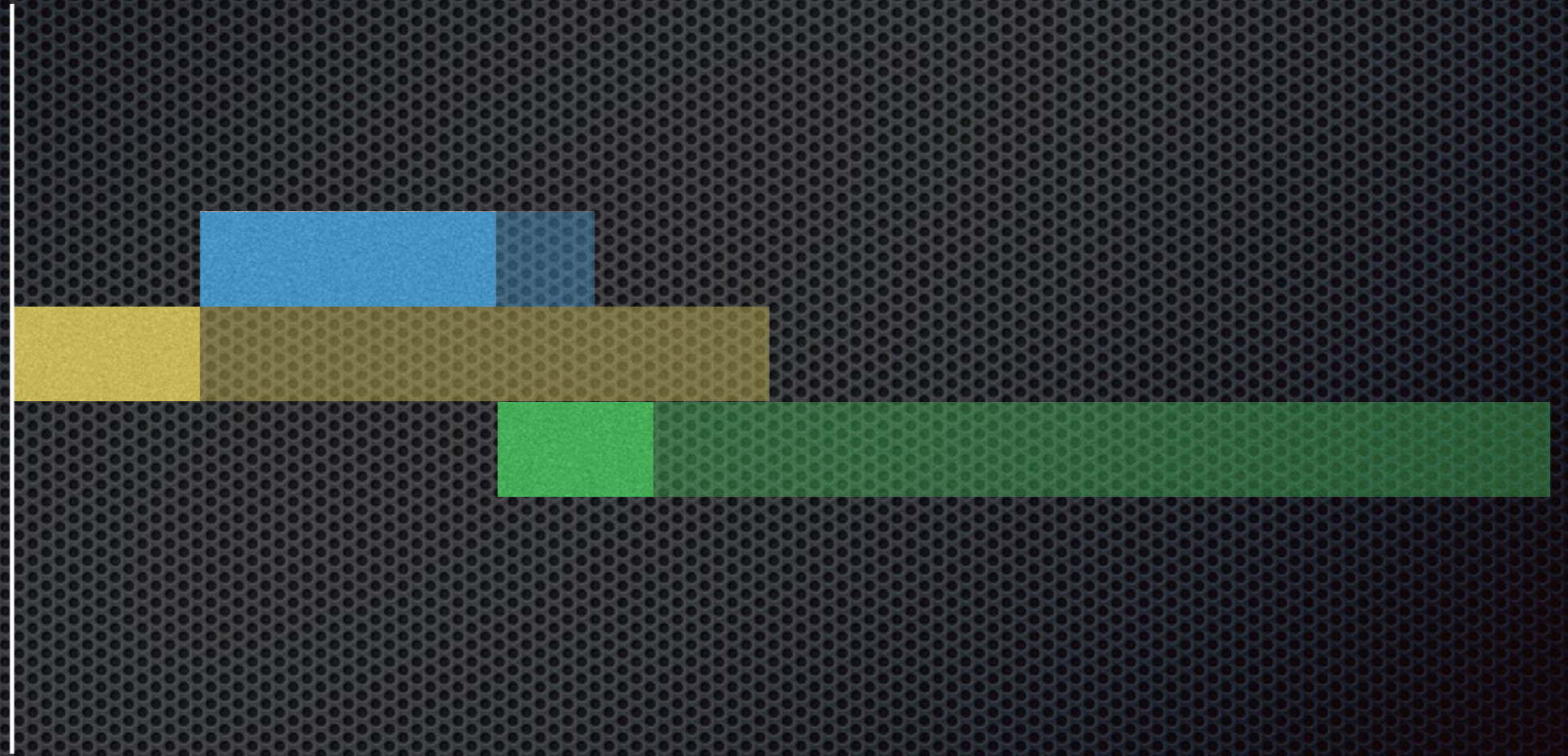


retrieval latency

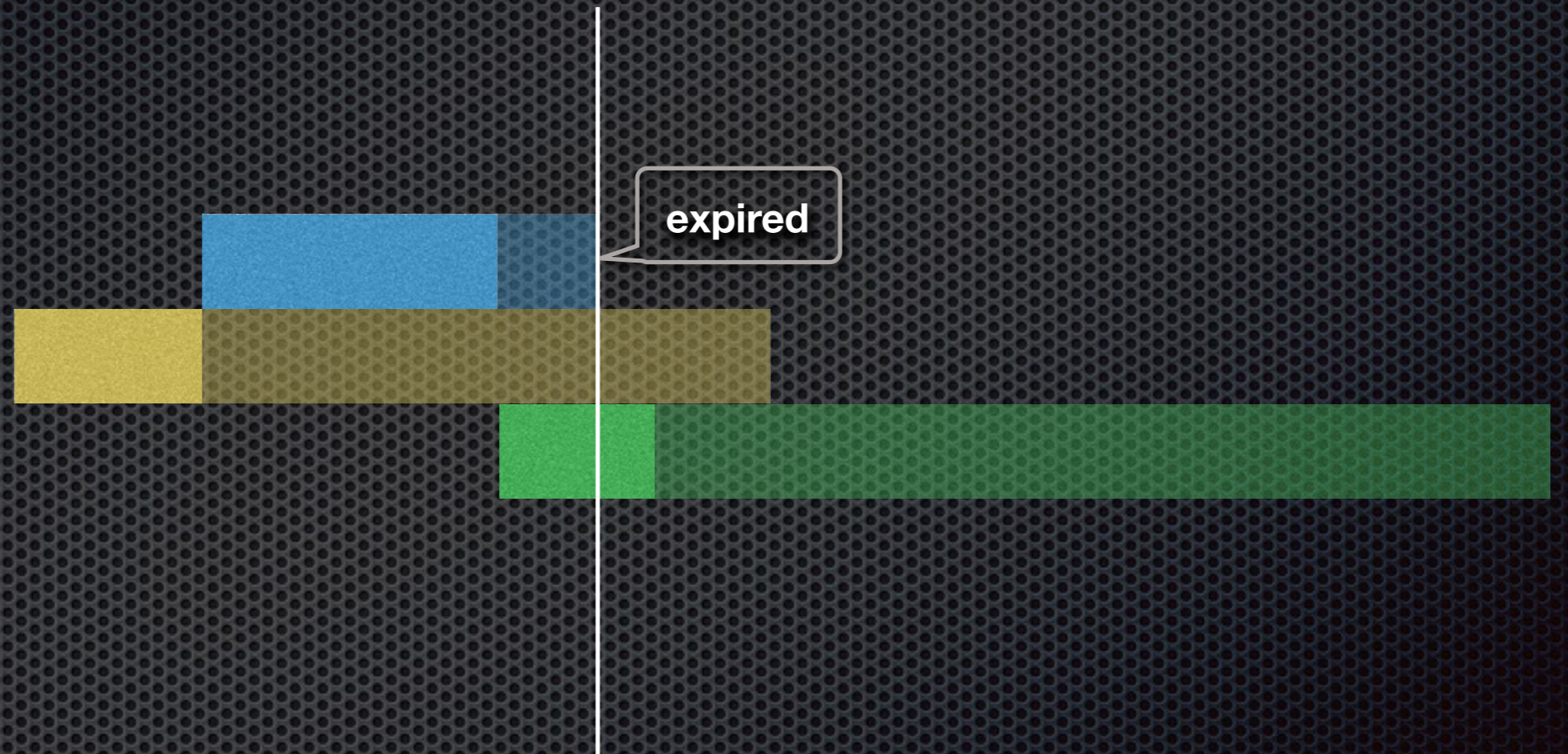


freshness interval

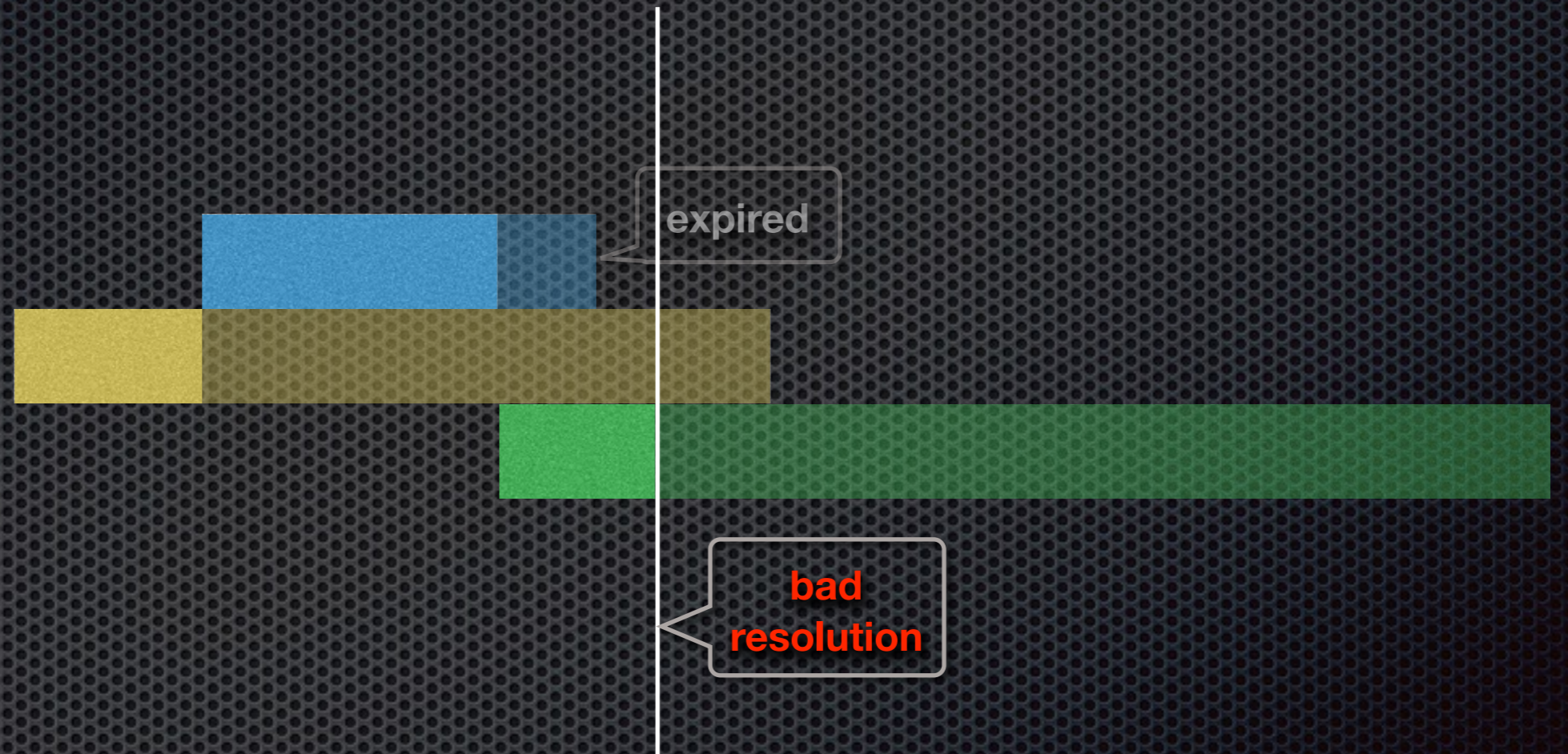




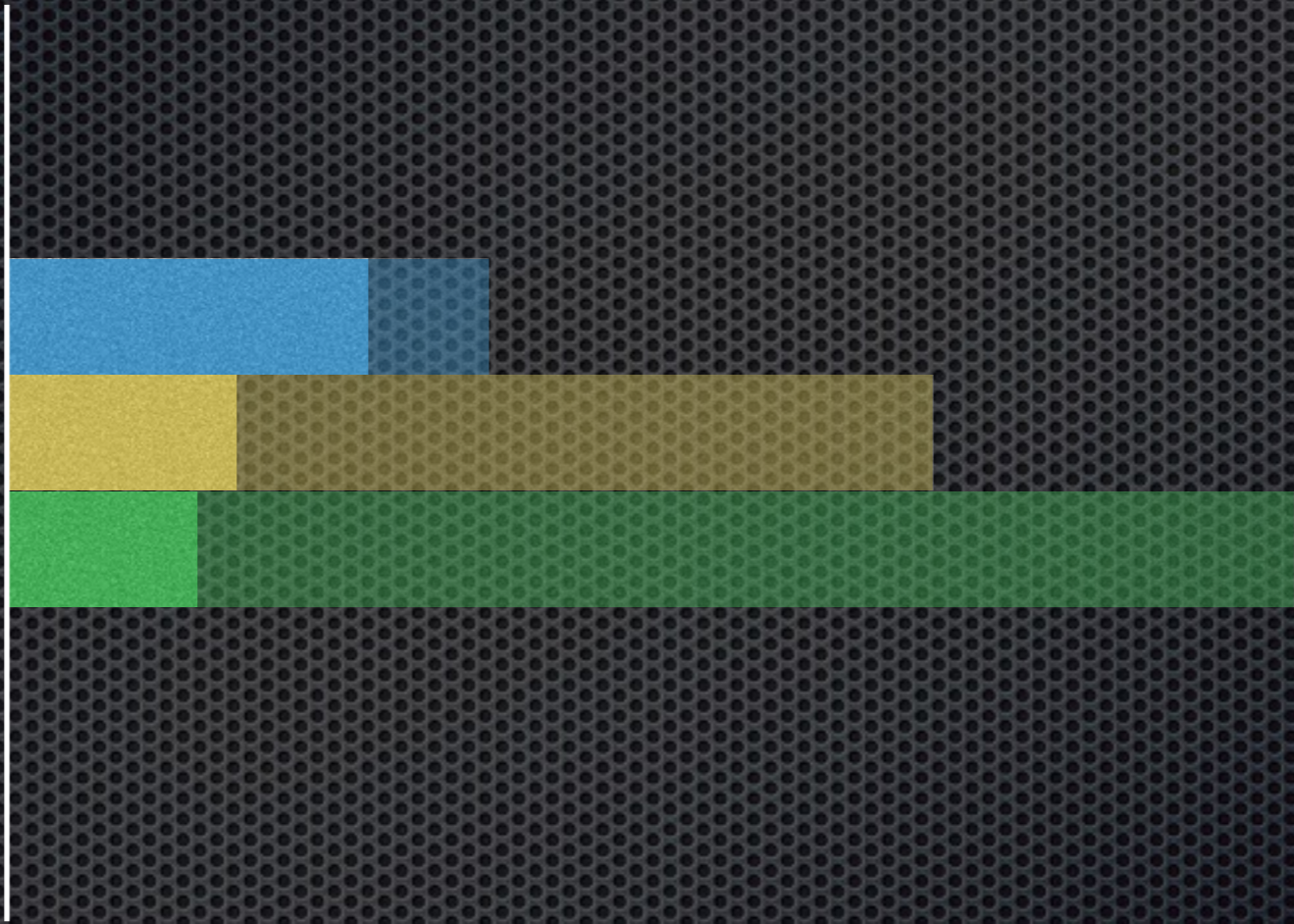
Some random ordering

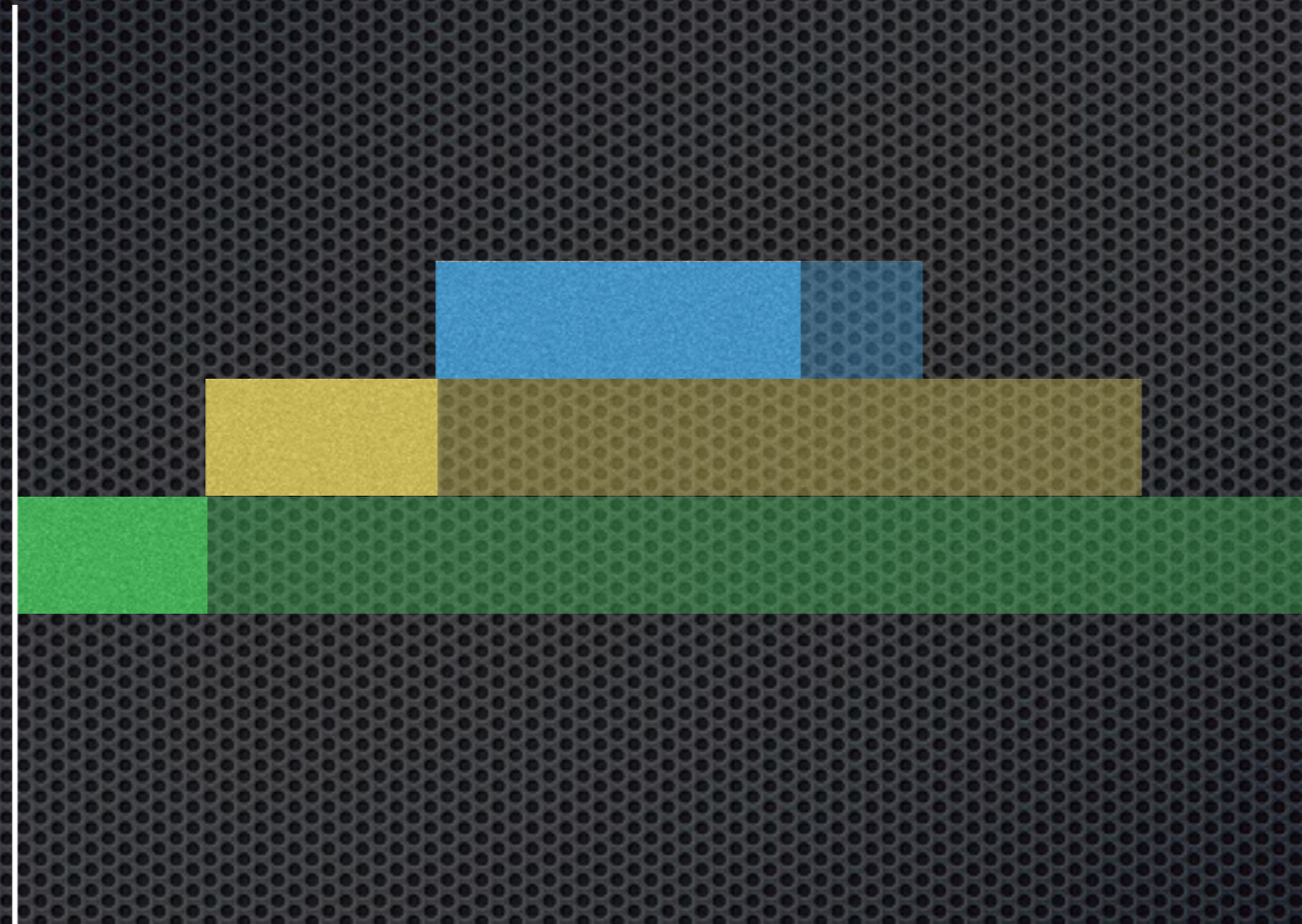


Some random ordering

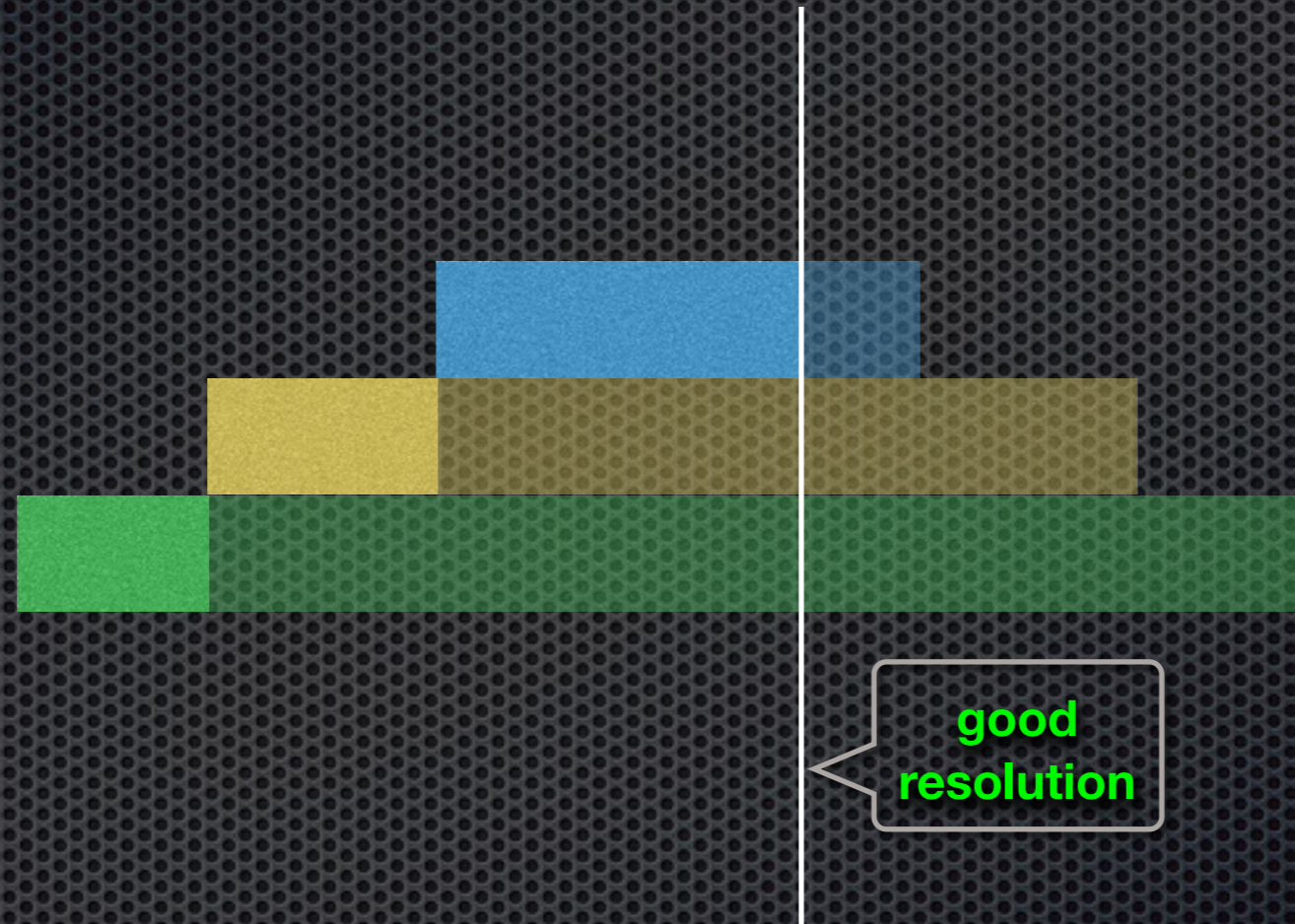


Some random ordering





Latest Deadline First (LDF)



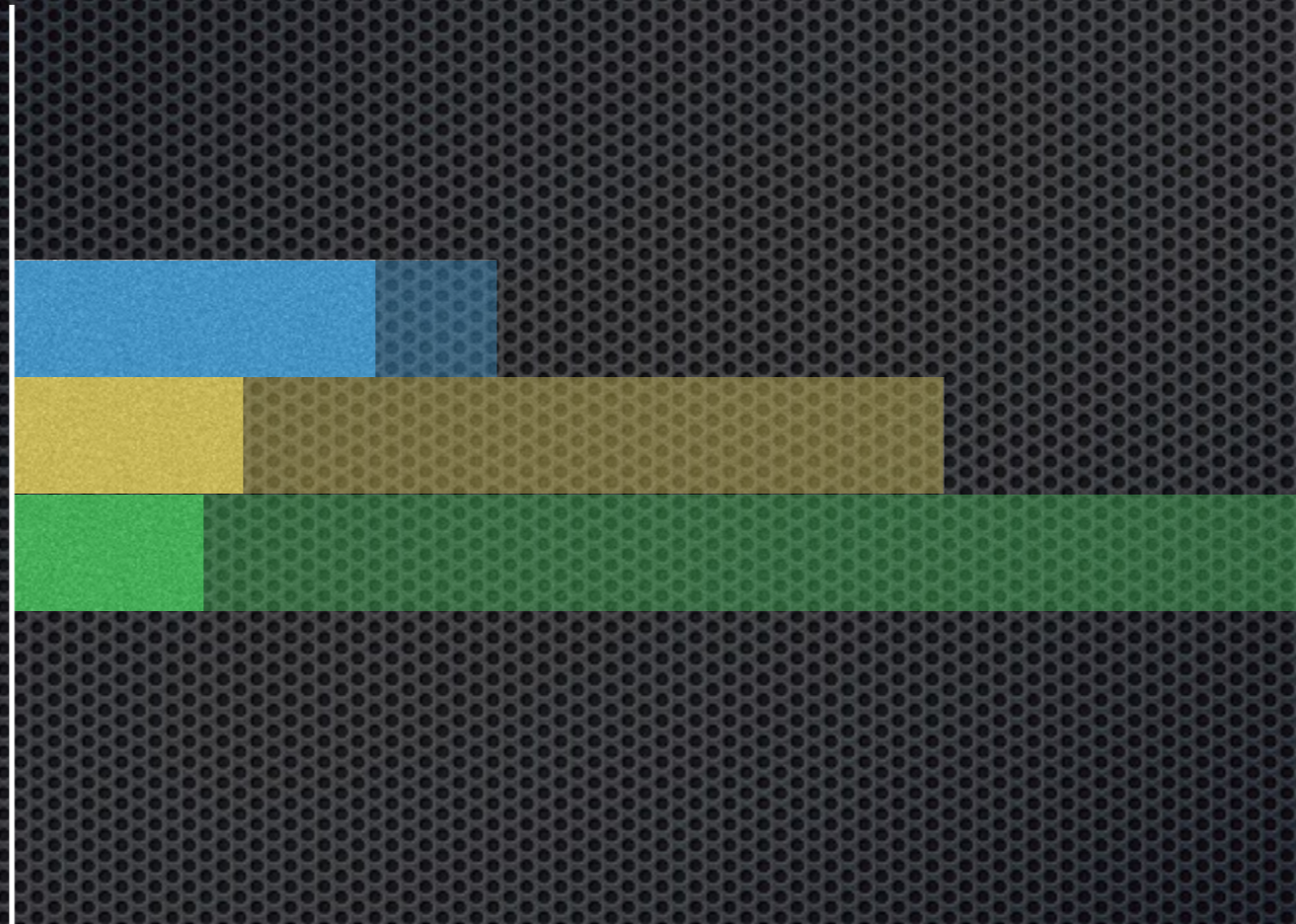
Latest Deadline First (LDF)

LDF - Latest Deadline First

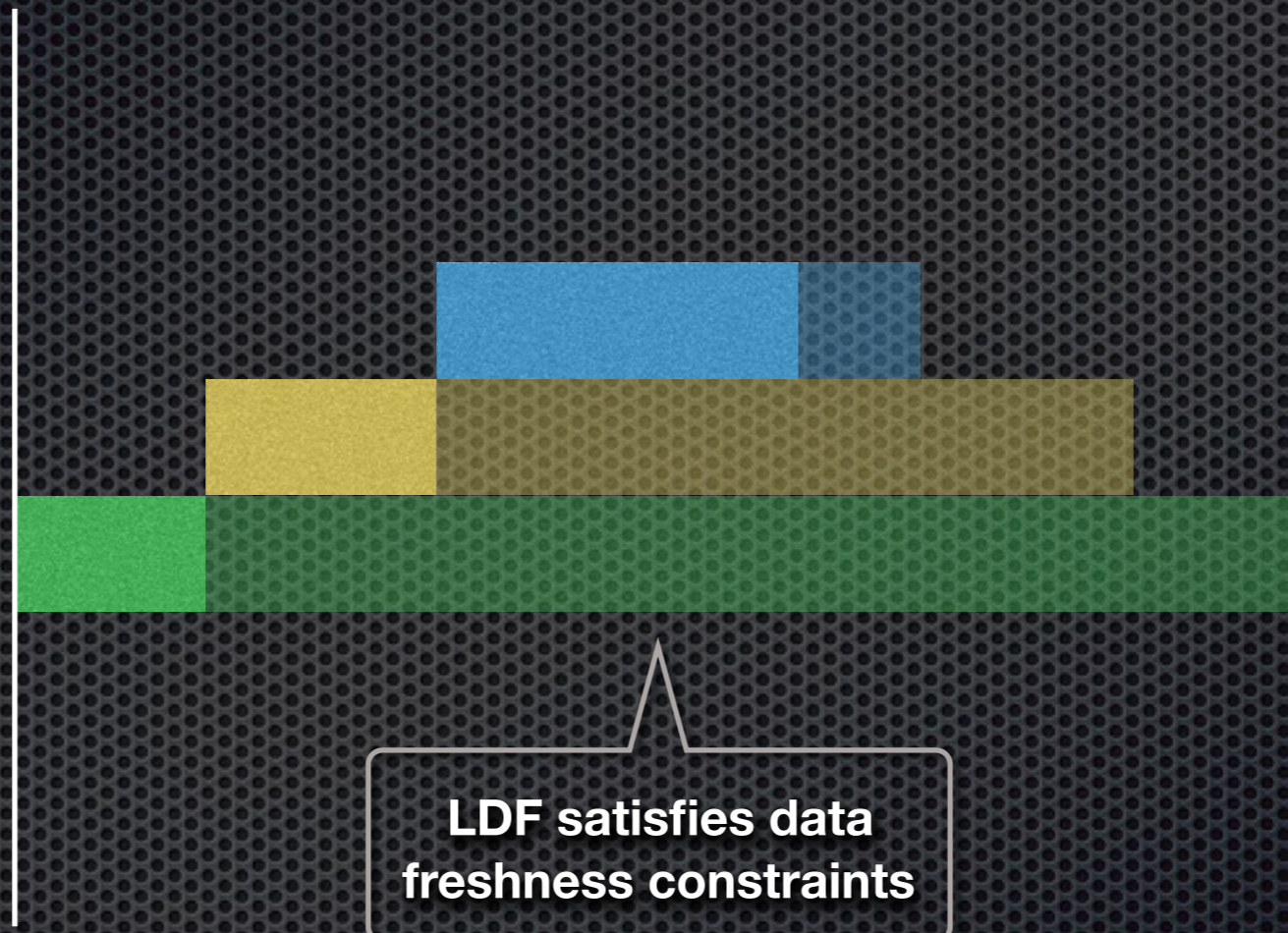
- ✦ Inspired by EDF: data objects with later freshness deadlines are retrieved sooner
- ✦ Optimal: if LDF cannot avoid freshness deadline violation, no sequential order can

LDF, what about Cost?

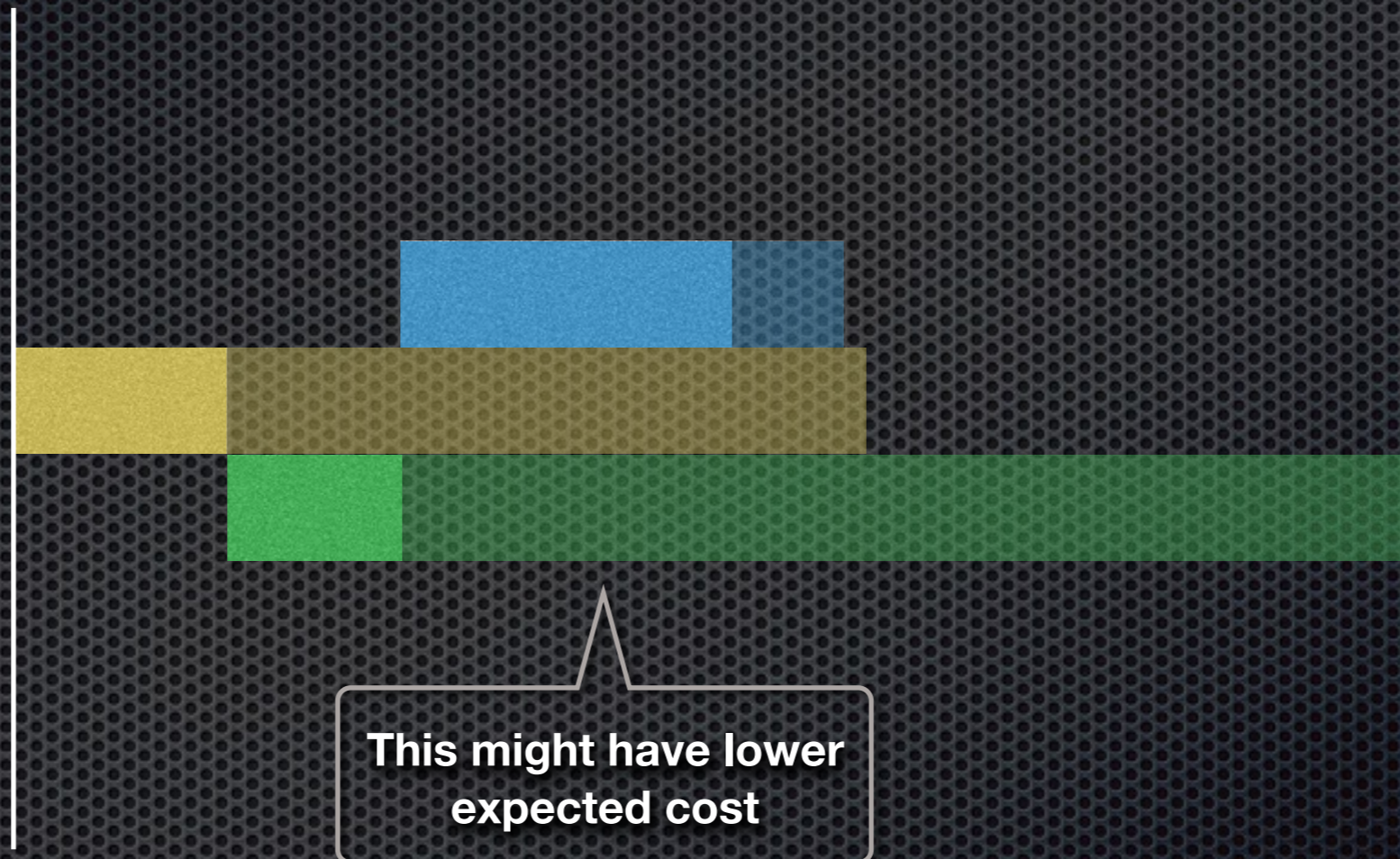
LDF, what about Cost?



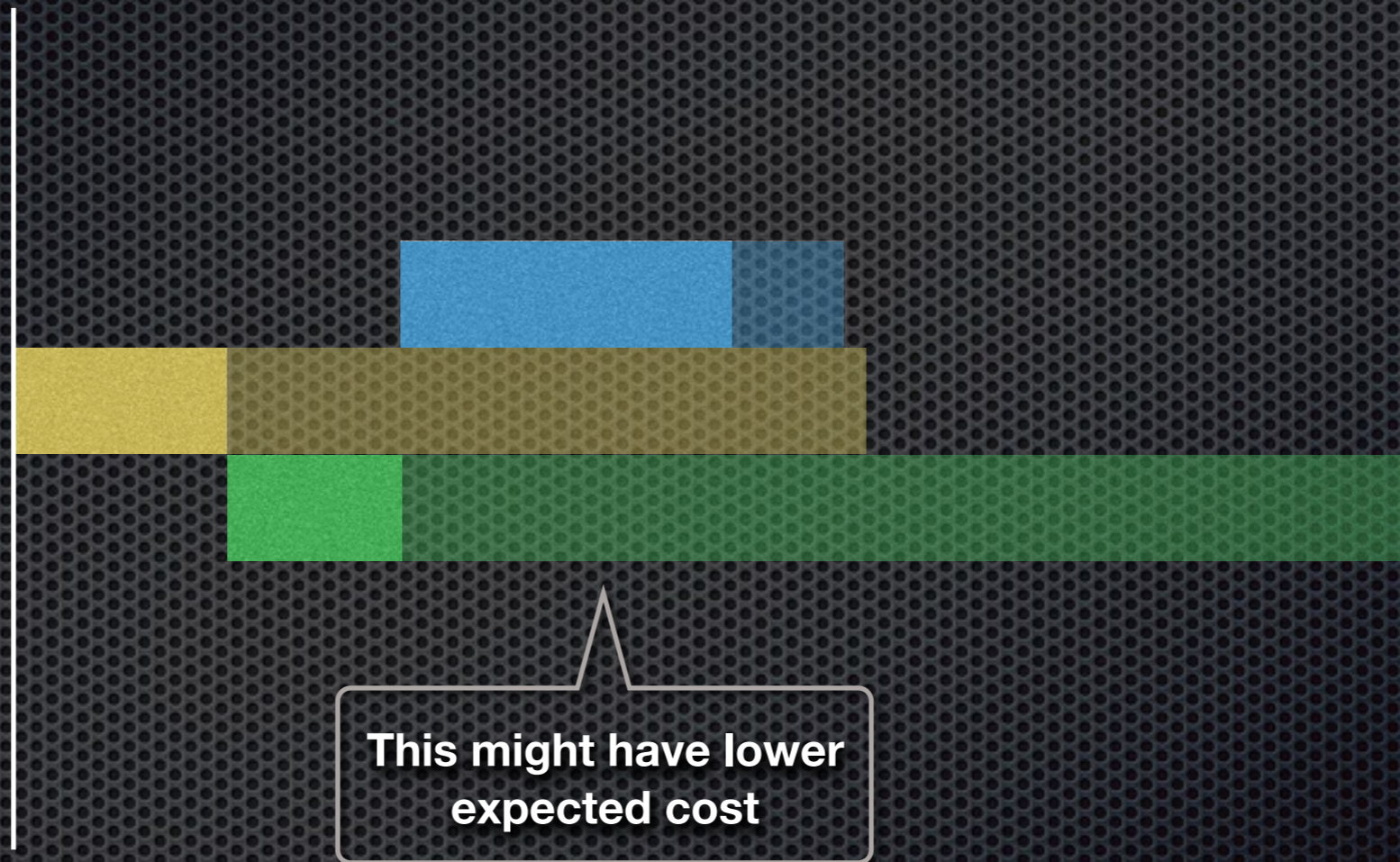
LDF, what about Cost?



LDF, what about Cost?



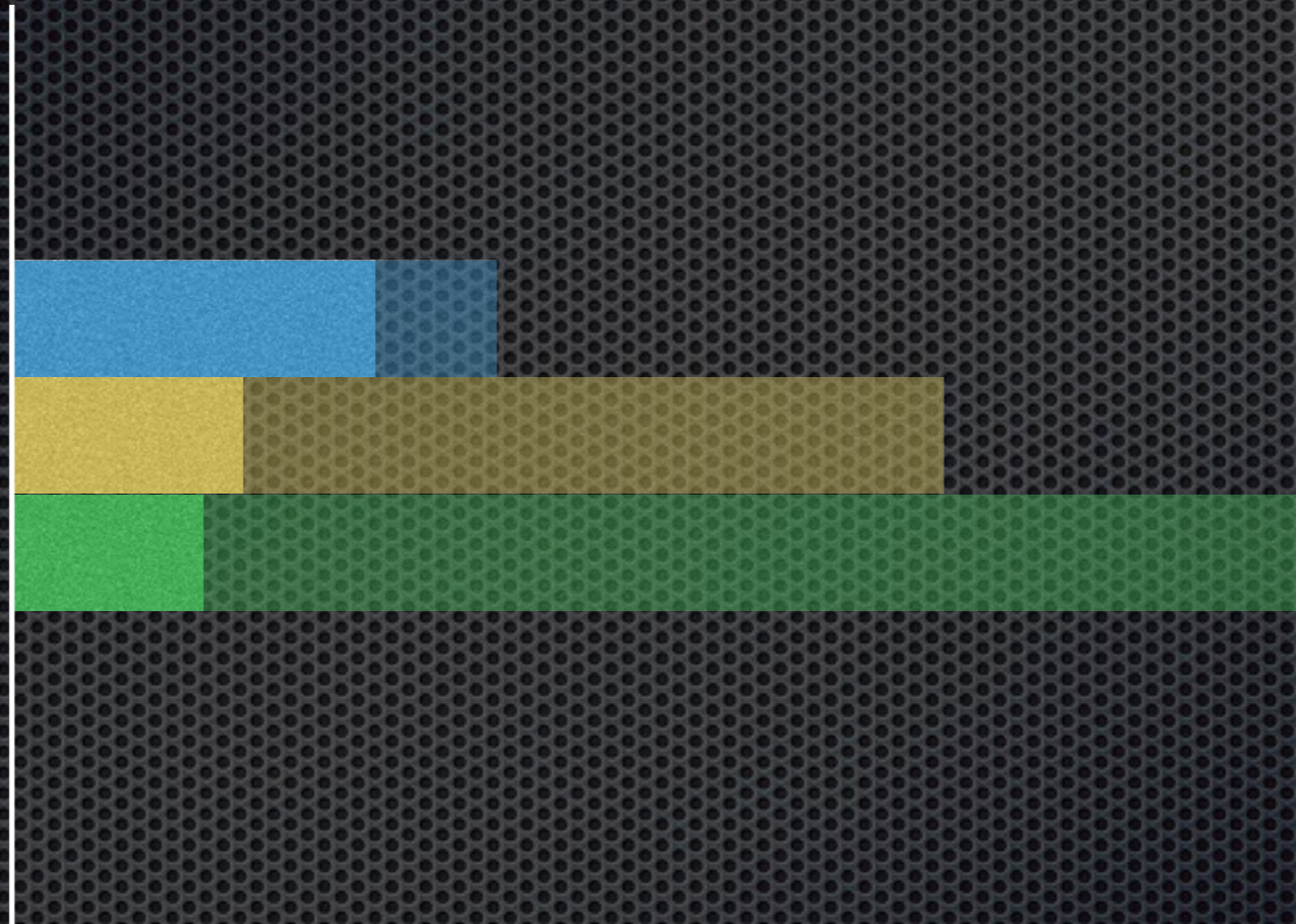
LDF, what about Cost?



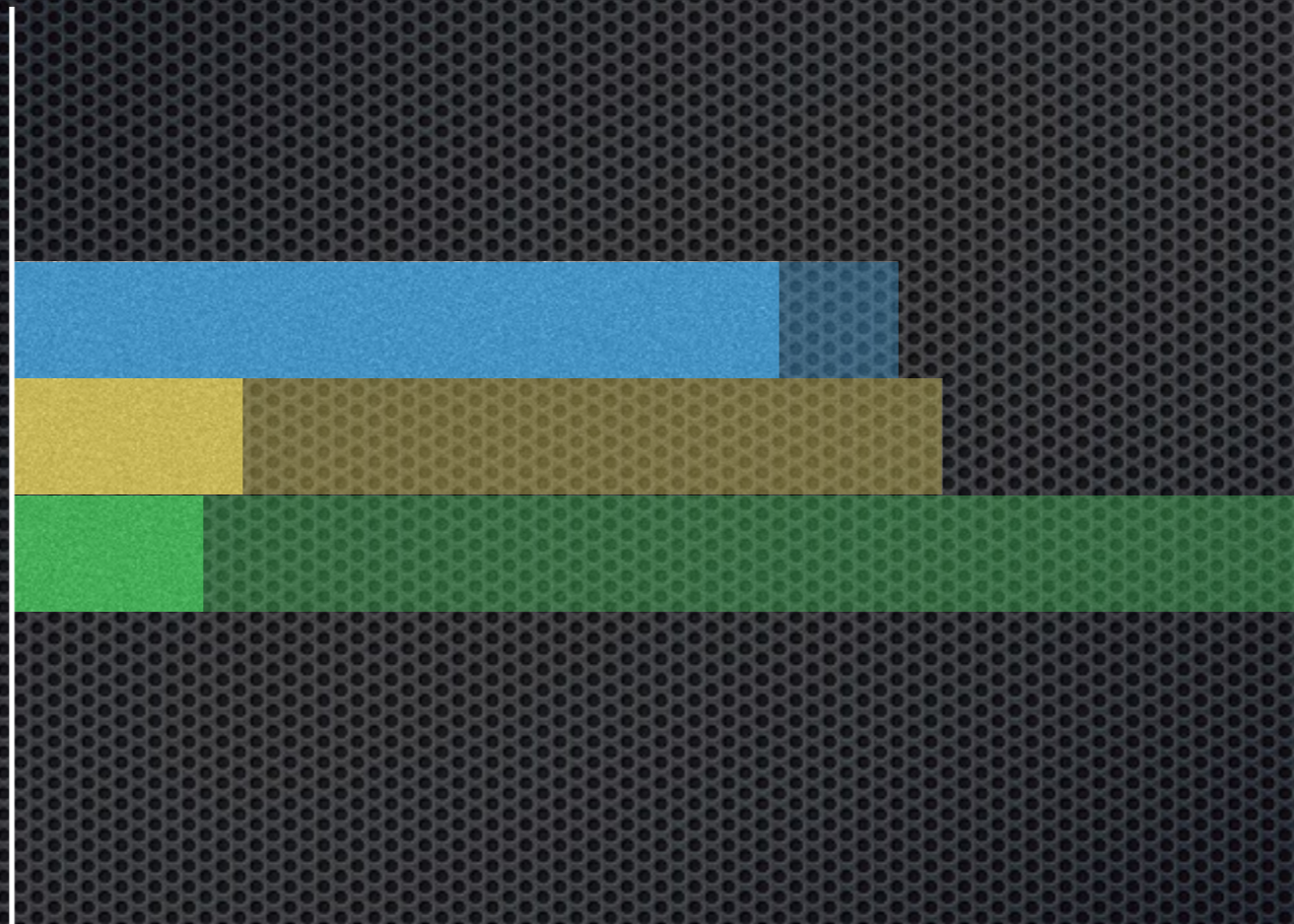
Greedily rearrange LDF order to reduce the expected data retrieval cost

LDF, still Freshness Violations?

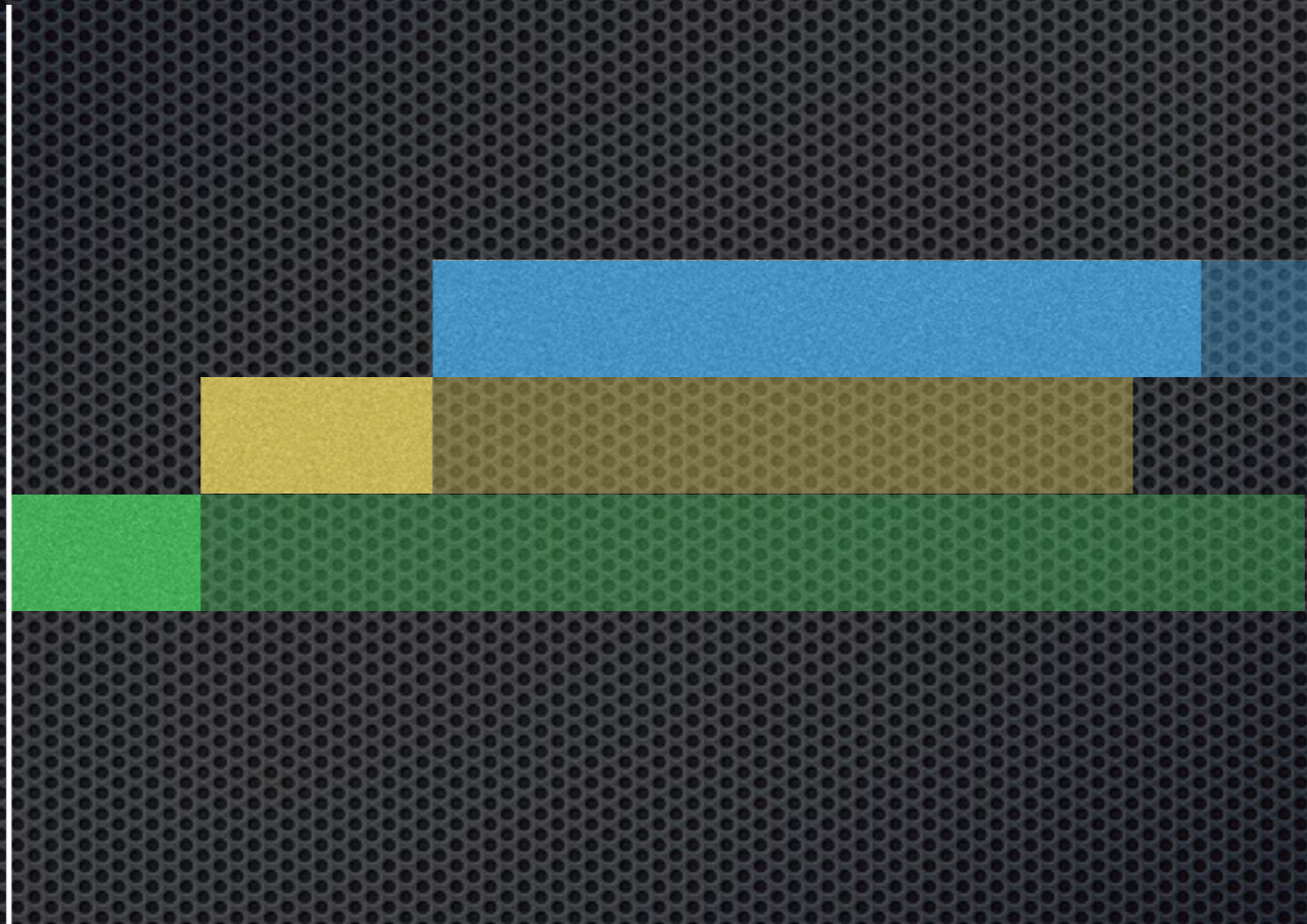
LDF, still Freshness Violations?



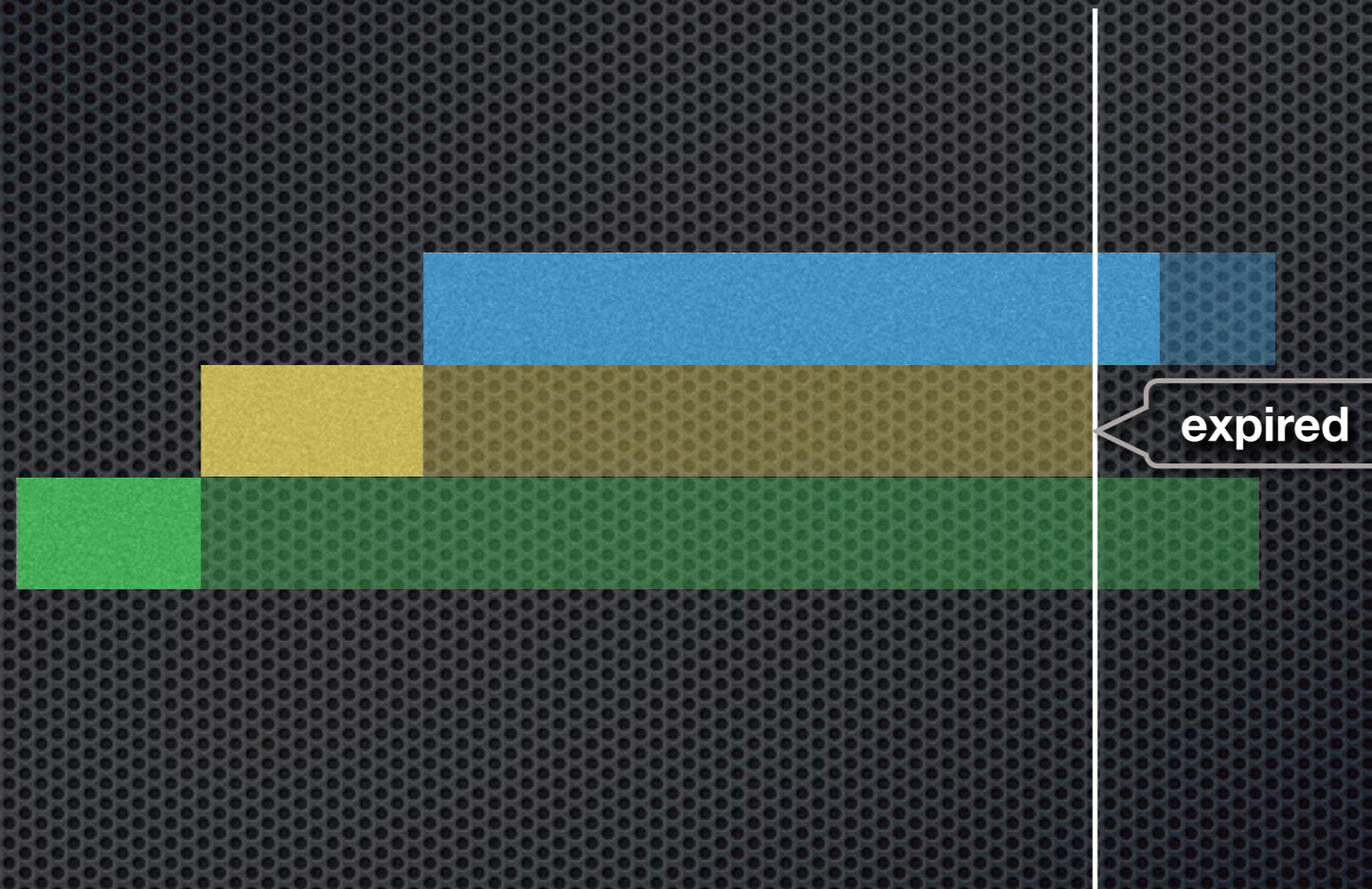
LDF, still Freshness Violations?



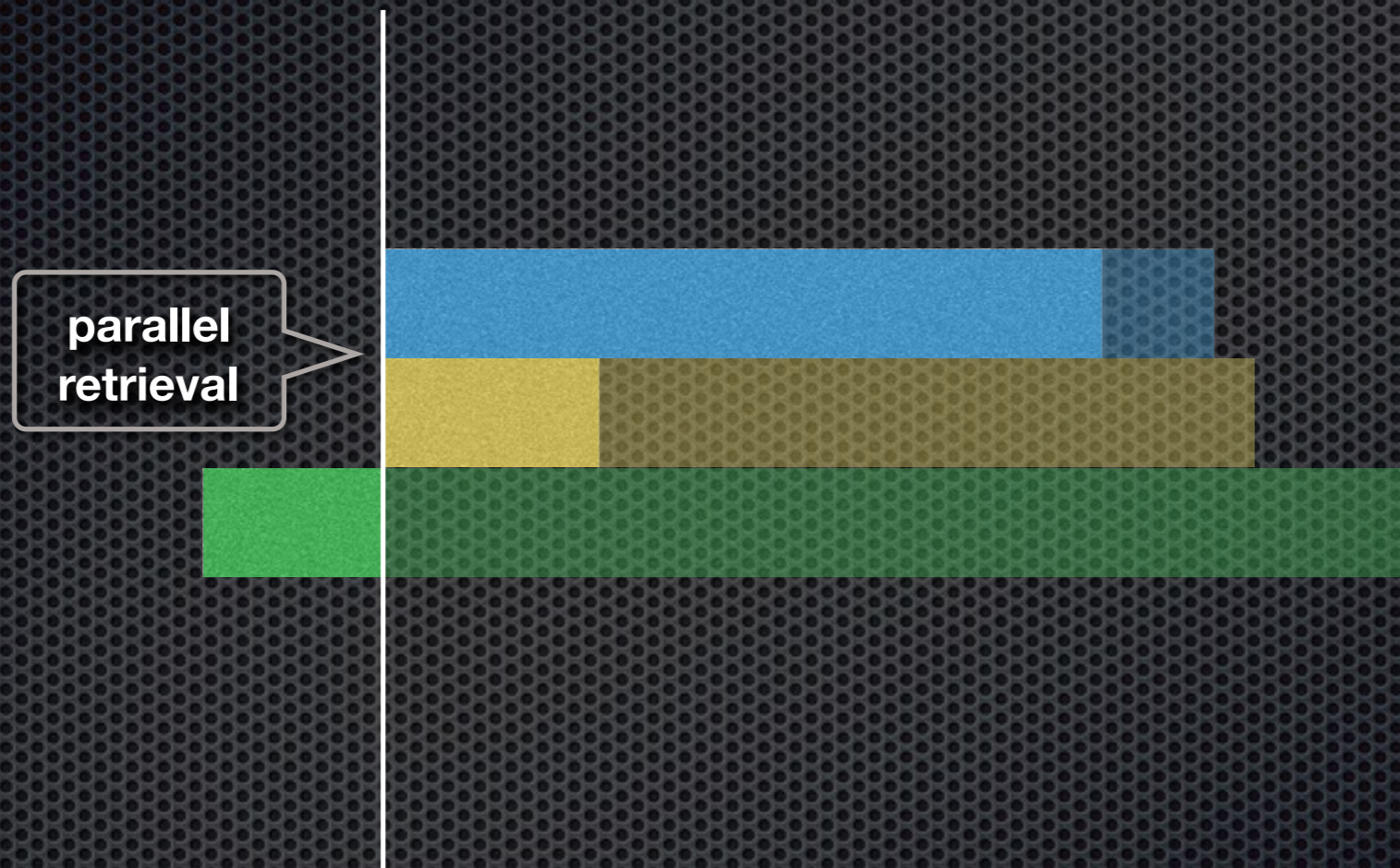
LDF, still Freshness Violations?



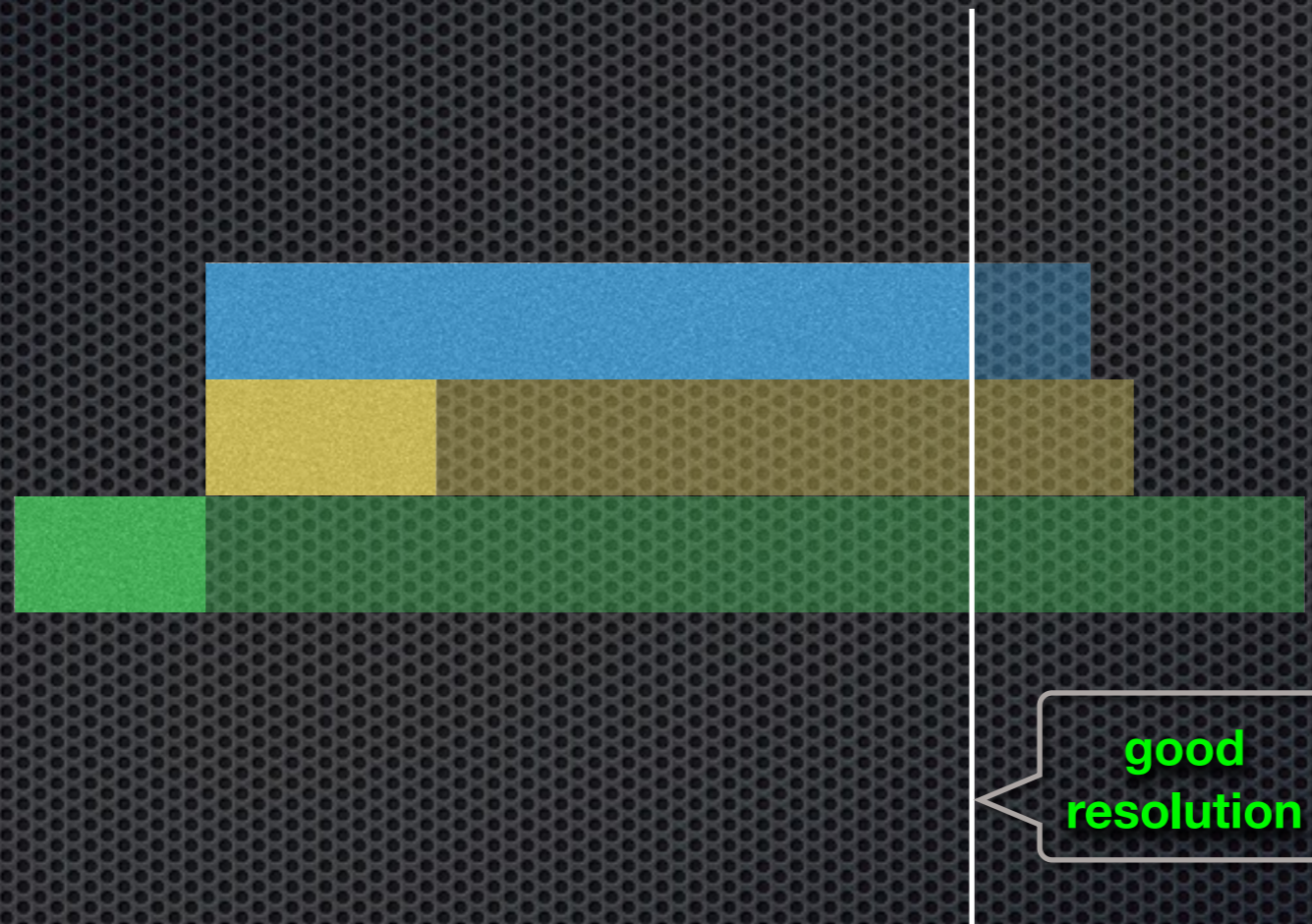
LDF, still Freshness Violations?



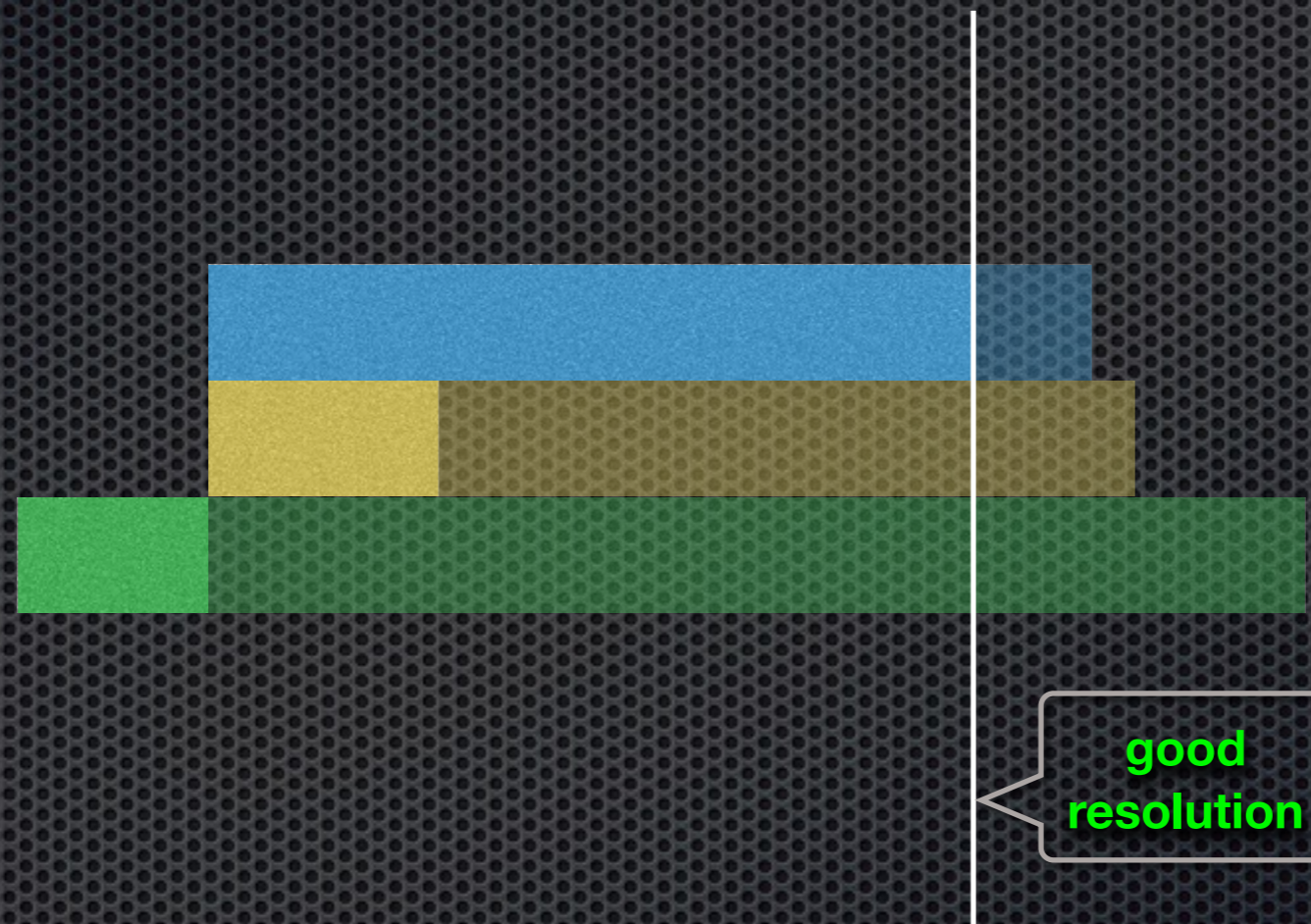
LDF, still Freshness Violations?



LDF, still Freshness Violations?



LDF, still Freshness Violations?



**Increment parallel retrieval level until
freshness constraints are met**

vLDF Data Retrieval

- ✦ Compute LDF order
- ✦ Greedily rearrange LDF order to reduce the expected data retrieval cost
- ✦ Gradually increment parallel retrieval level until freshness constraints are met

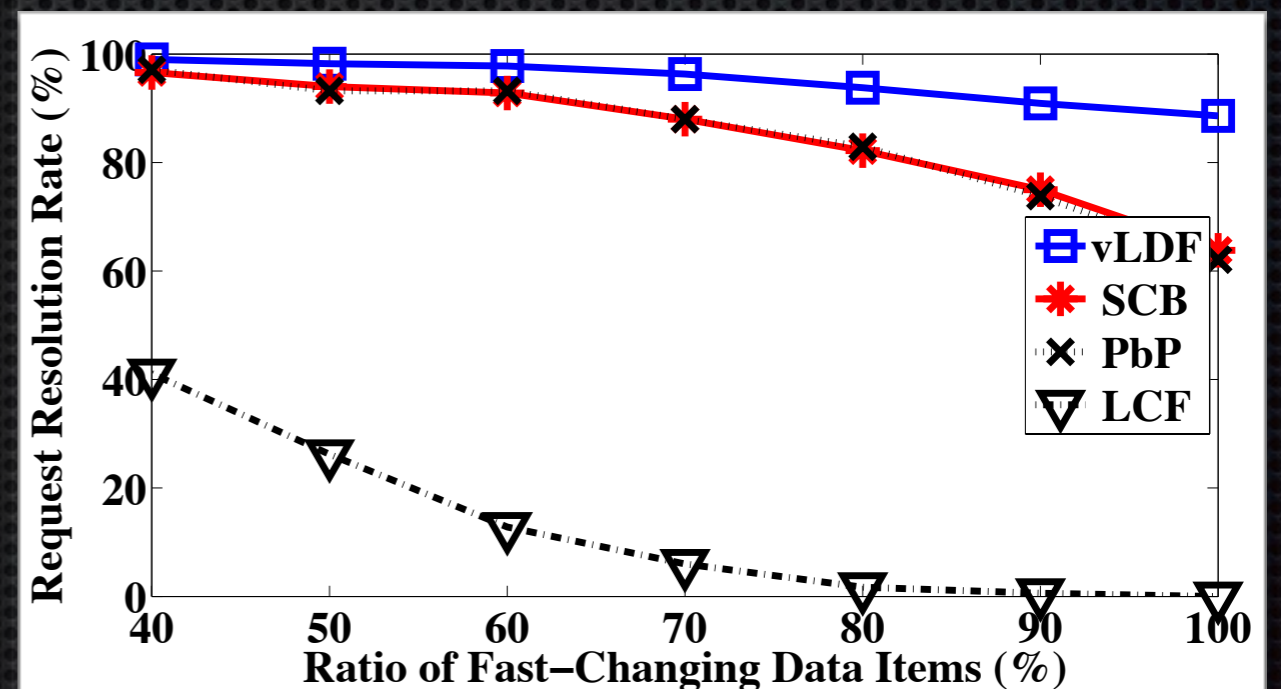
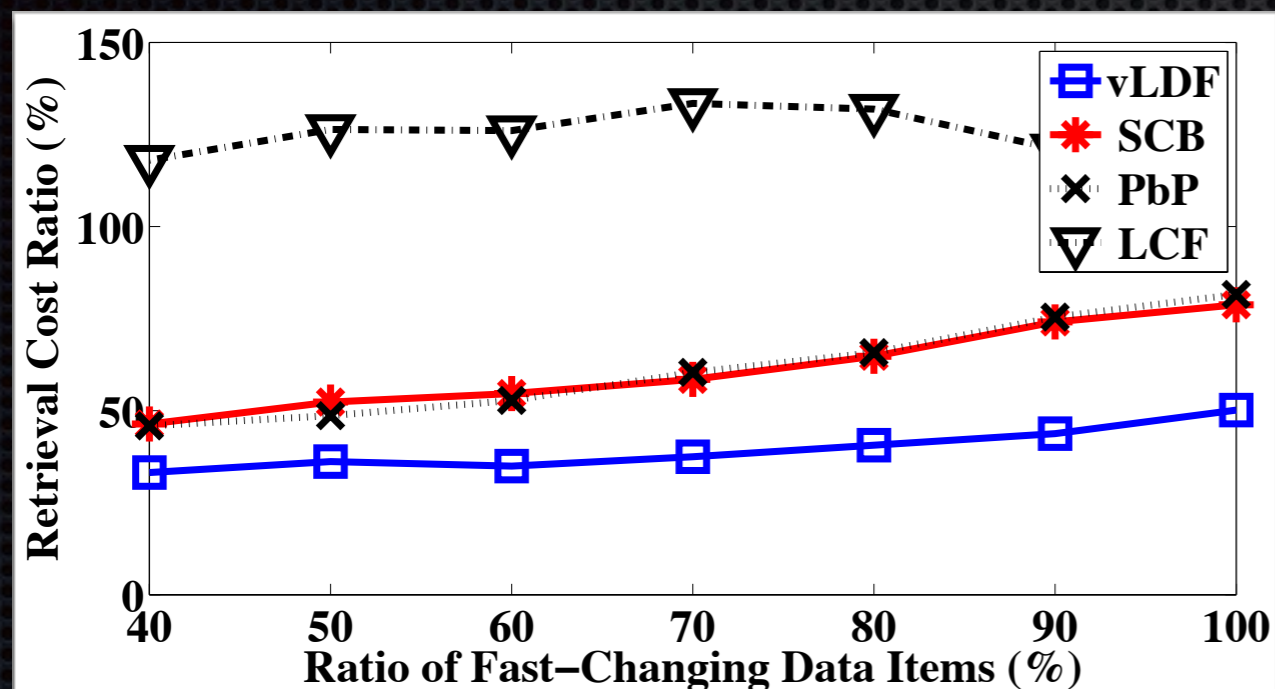
Evaluation

- Simulation experiments
- An application scenario

Simulation Experiments

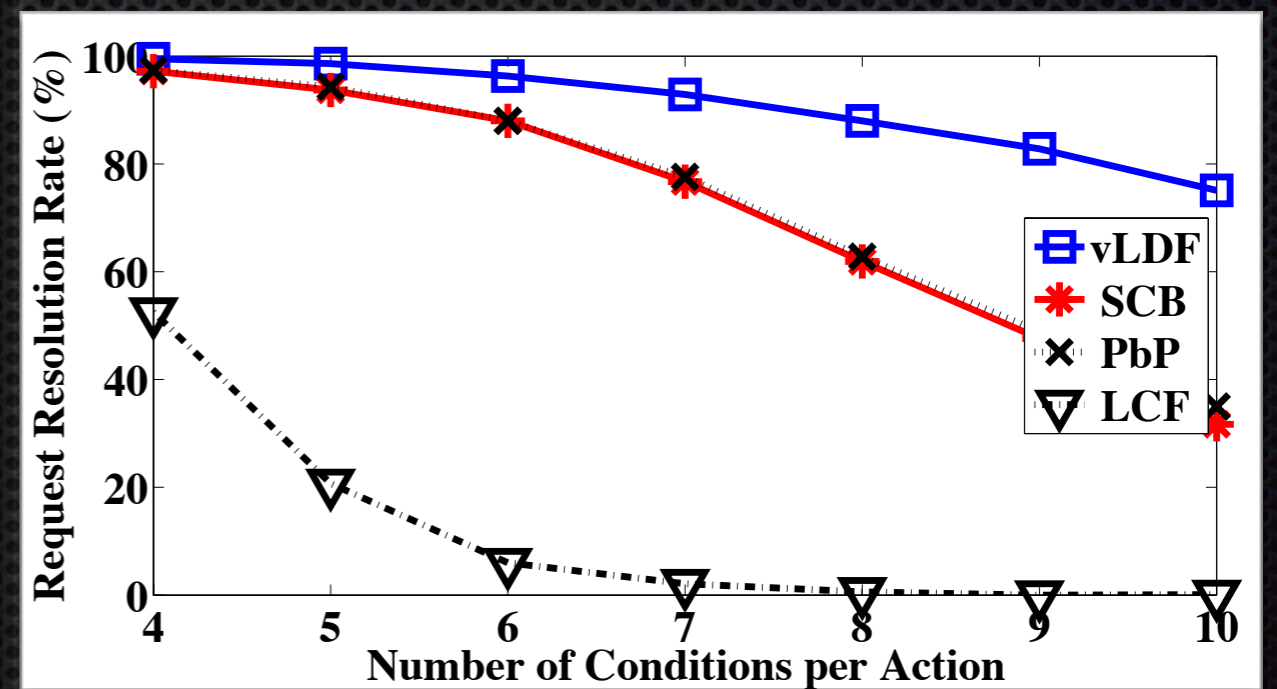
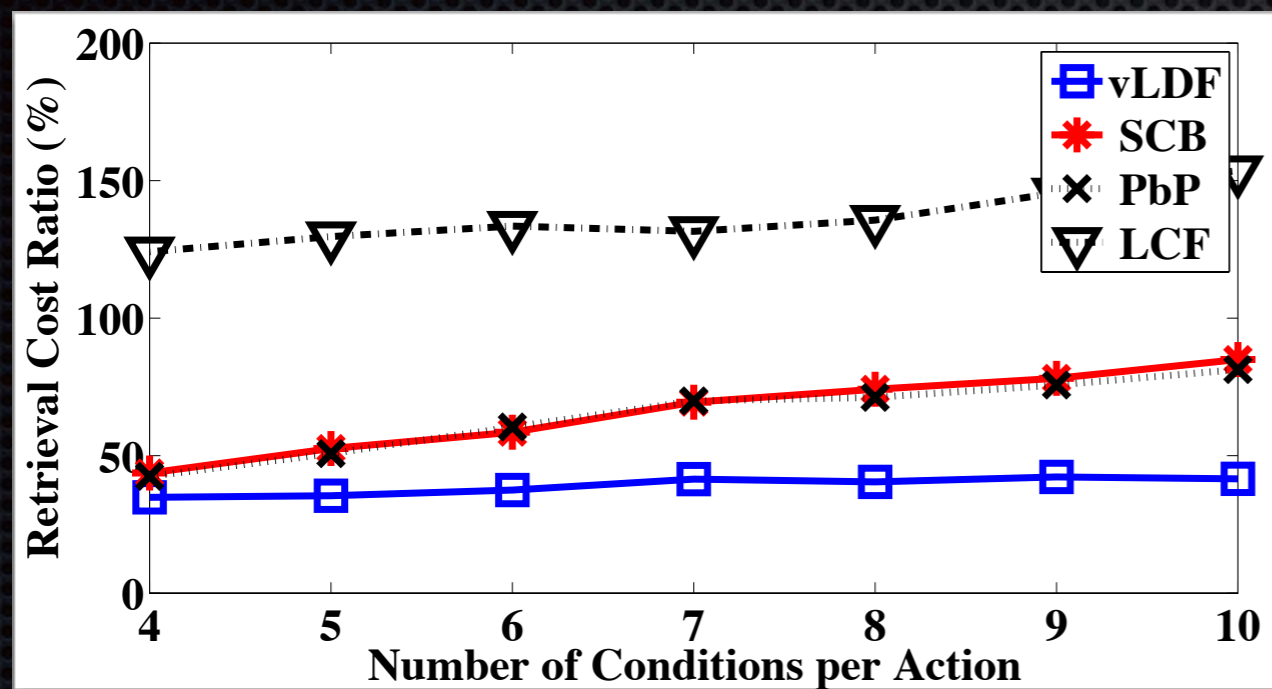
- ✦ Baselines
 - ✦ LCF - Lowest Cost Source First
 - ✦ SCB - Shortcircuit Benefit only
 - ✦ PbP - Probability based Prediction
- ✦ Settings
 - ✦ % fast changing data: 40~100%, default 70%
 - ✦ # Action size: 4~10, default 6
 - ✦ Data object size: 3~5 MB, default 3.45 MB
 - ✦ Network bandwidth: 3.5~6.5 KBps, default 5 KBps
 - ✦ Transmission latency fluctuation: -3~3 min, default 0

Simulation Results



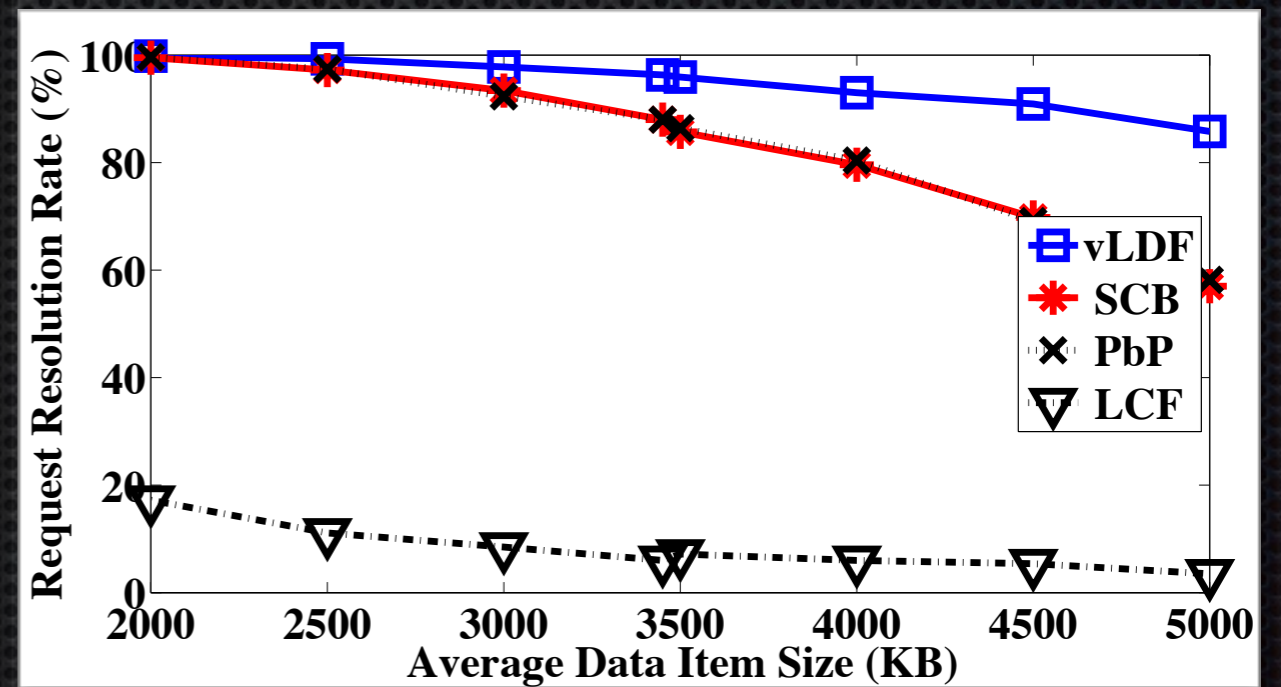
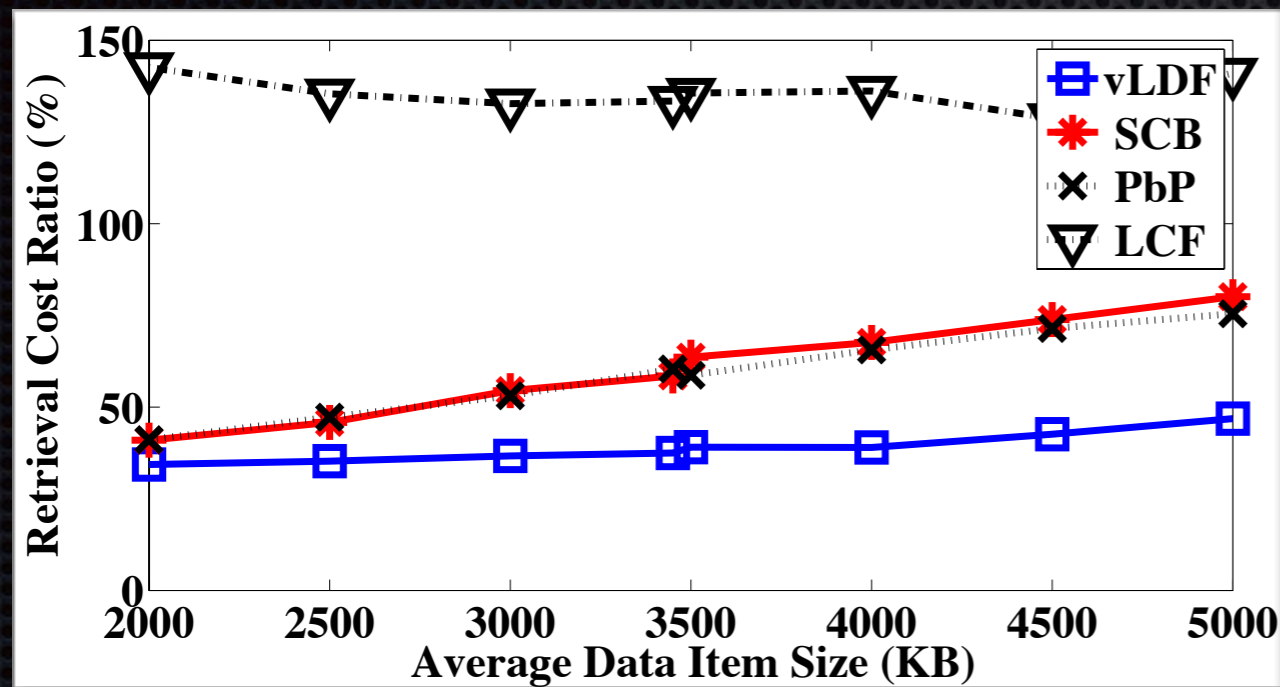
Varying % of fast changing data

Simulation Results



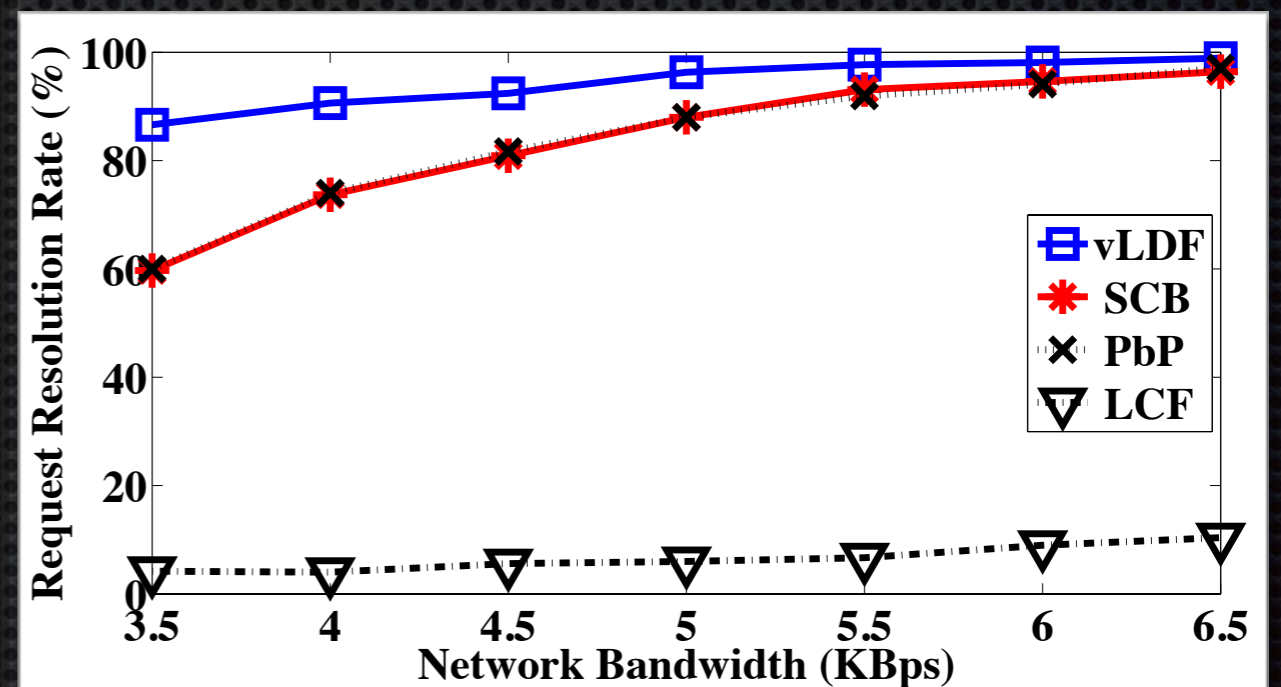
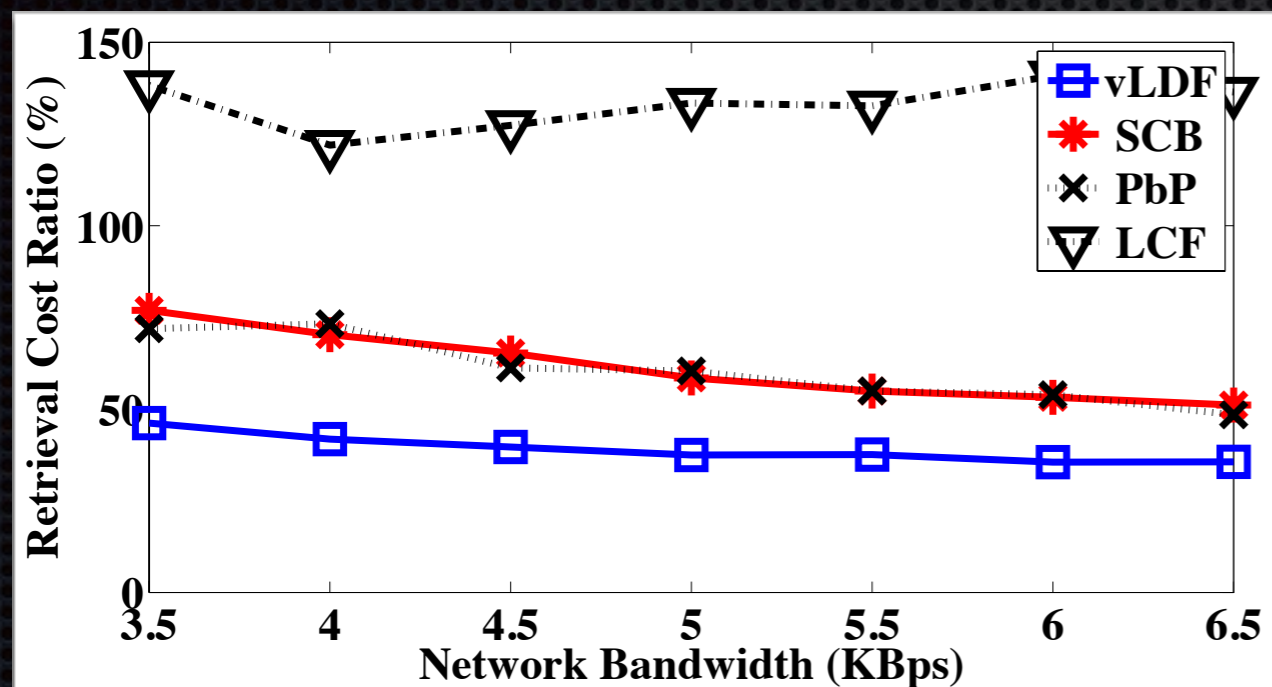
Varying action size

Simulation Results



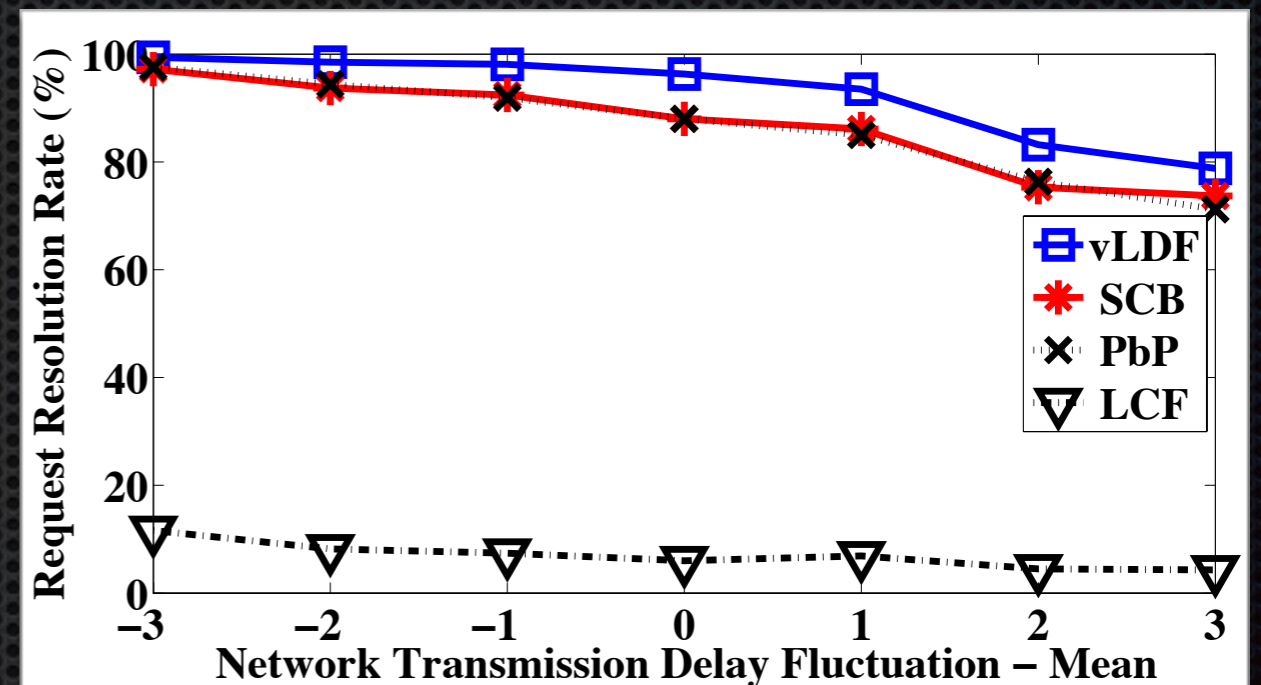
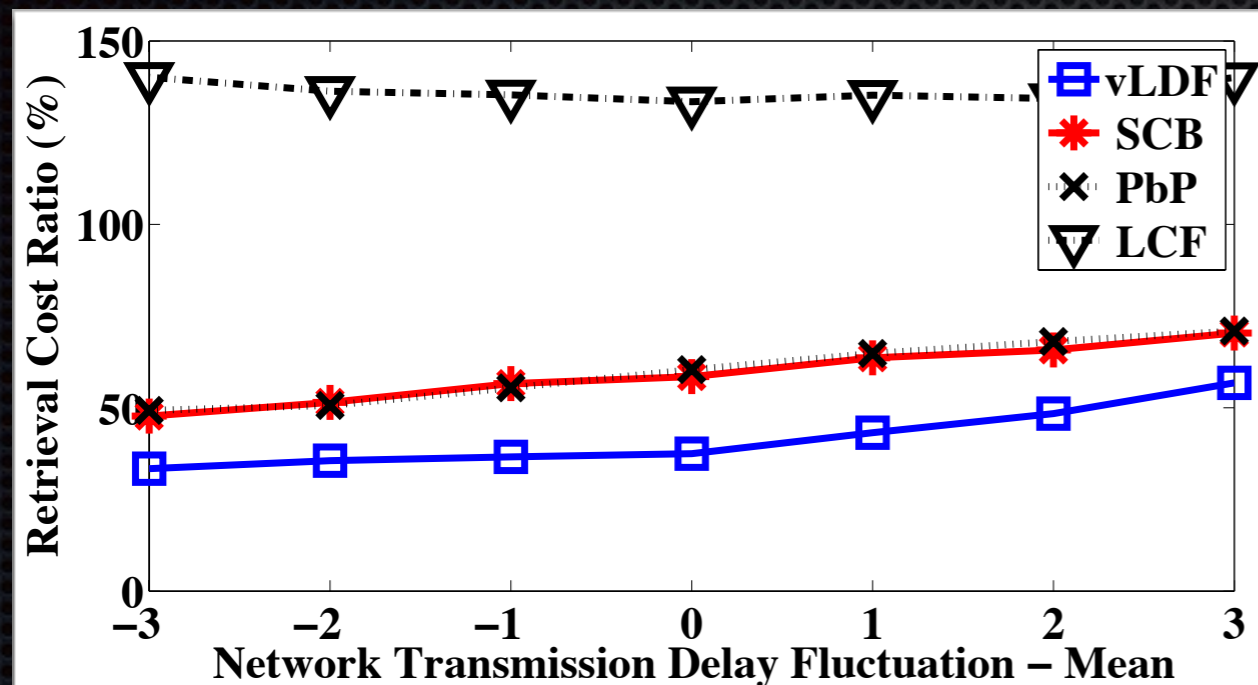
Varying data object size

Simulation Results



Varying network bandwidth

Simulation Results



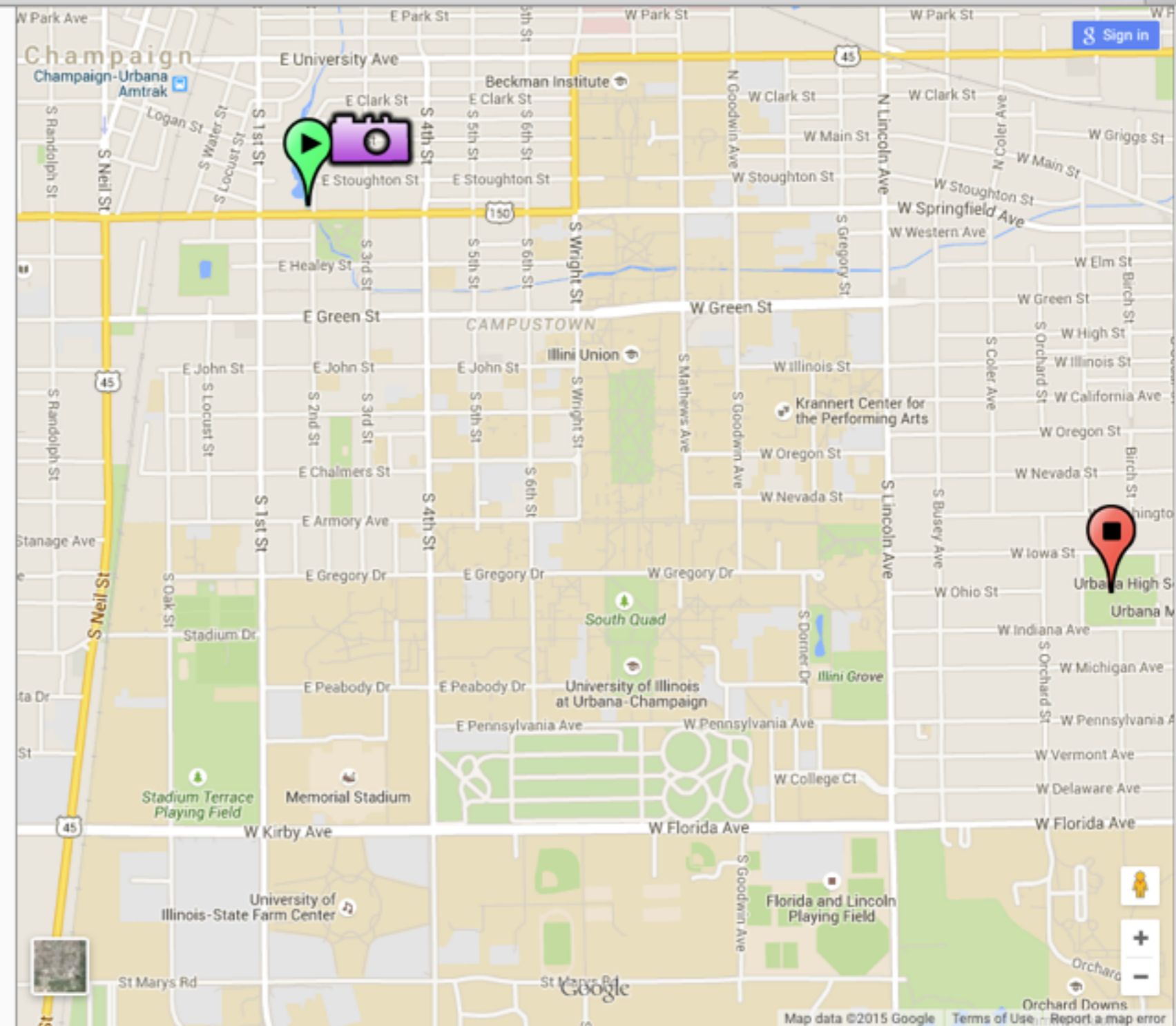
Varying network transmission fluctuation

Application: Route Finding

- ✦ Find routes for $\langle \text{src}, \text{dst} \rangle$ pairs
 - ✦ Each candidate route: **AND** of its segments
 - ✦ Routing result: **OR** of all candidate routes
- ✦ Visual verification for route segment conditions

Q from Helms park to Carle park

- Random Source Selection
- Optimal Source Selection
- Optimal Retrieval Order

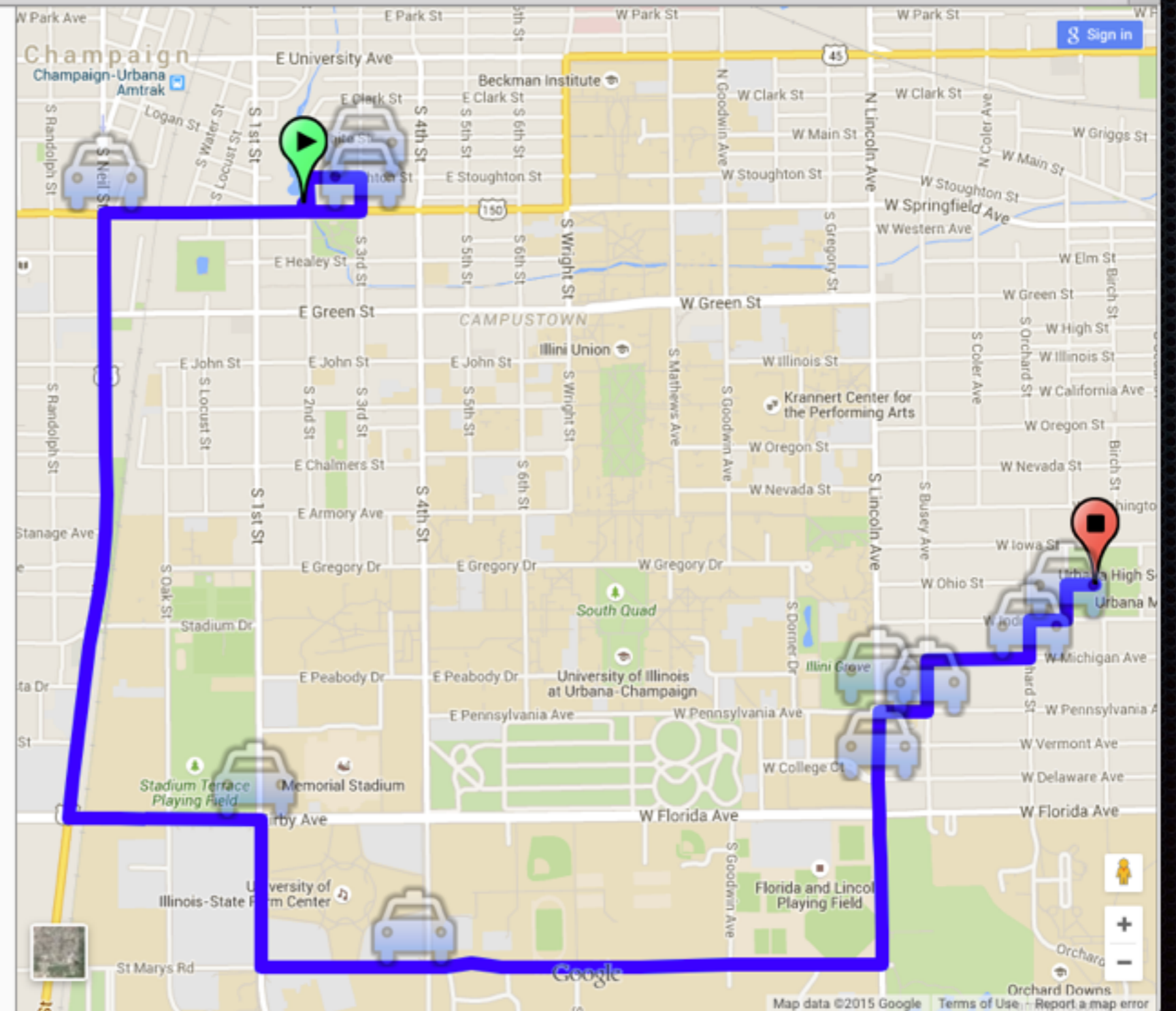


Q from Helms park to Carle park

- Random Source Selection
- Optimal Source Selection
- Optimal Retrieval Order



A good route is found!



Results of 5 Runs

vLDF Cost (KB)	PbP Cost (KB)	vLDF Time (s)	PbP Time (s)
516	685	164	255
343	598	150	206
319	485	160	248
506	1093	165	372
524	1042	175	206

Conclusion

- ✦ Environment dynamics & resource limitations affect real-time decision-making
- ✦ Efficient data acquisition algorithm
- ✦ Promising results through simulations and concrete route finding application scenario

Thanks