CS 744: DRF

Shivaram Venkataraman

Fall 2019

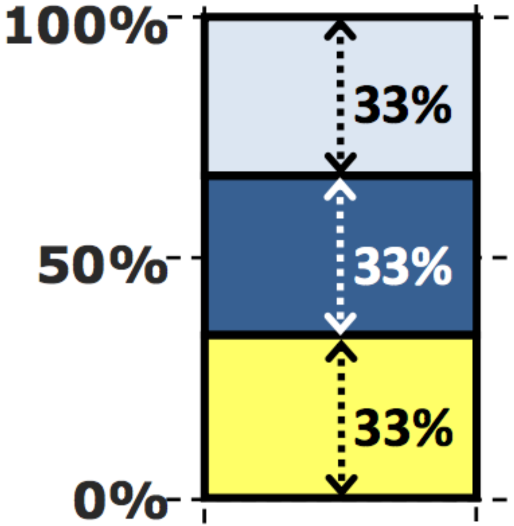
ADMINISTRIVIA

- Assignment 1 details
- Assignment 2 out tonight
- Project groups

SETTING: FAIR SHARING

CPU sharing
Network sharing

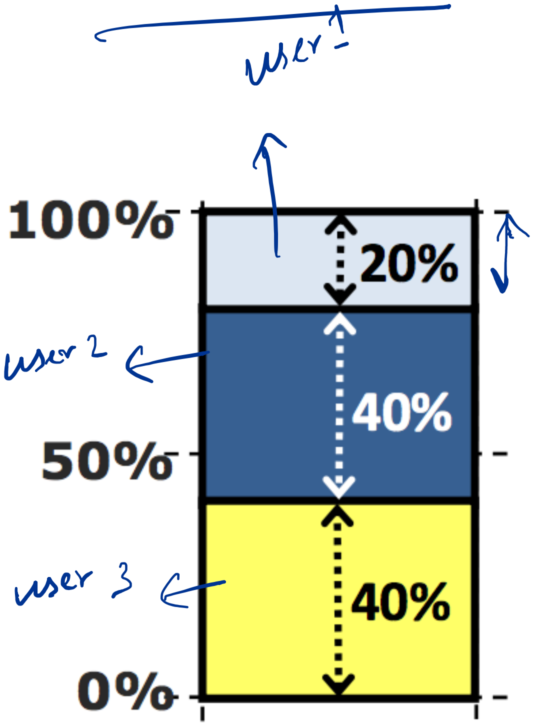
Equal Share



Max-Min Share

Maximize the allocation for most poorly treated users

Maximize the minimum



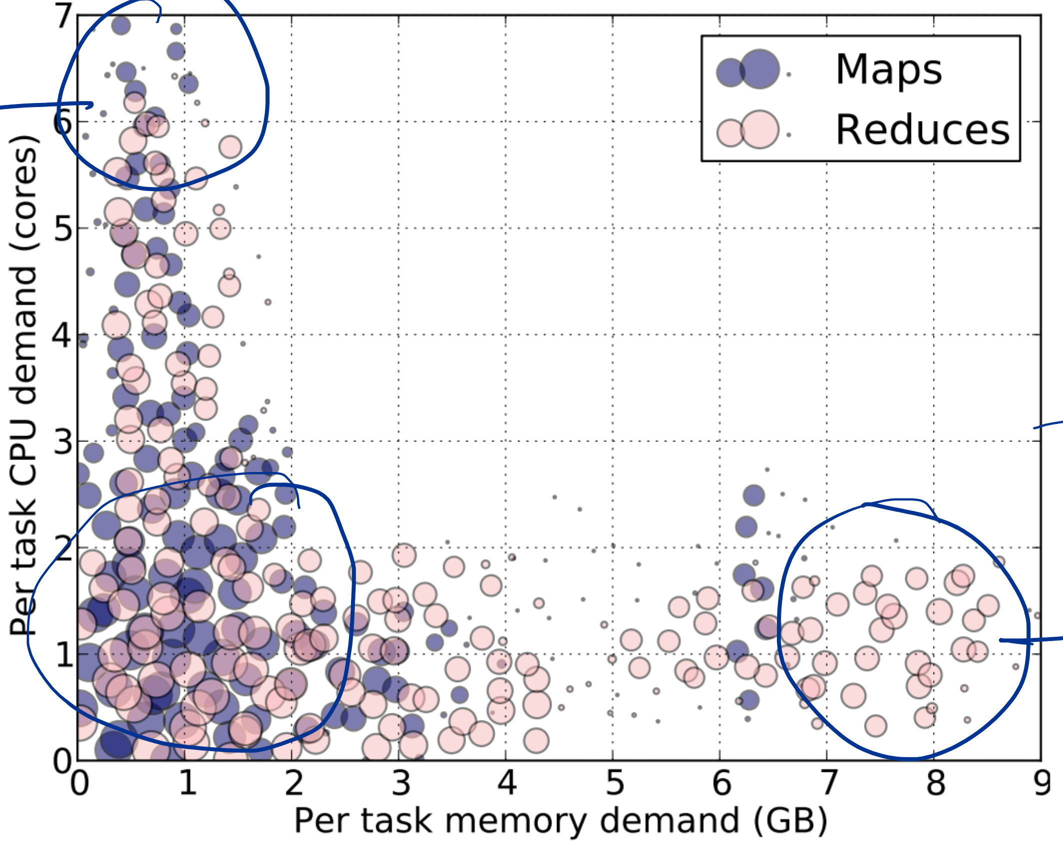
MOTIVATION: MULTI RESOURCES

CPU intensive

4CPU 4CPU

16GB 16GB

~2GB, 2CPU



<7CPU, 1GB>

<2CPU, 7GB>

Classic Hadoop sched.

memory intensive

4 map - slots

2 reduce - slots

DRF: MODEL

Users have a demand vector

<2, 3, 1> means user’s task needs 2 R1, 3 R2, 1 R3

Resources given in multiples of demand vector

i.e., users might get <4,6,2> $\simeq$ 2 tasks with <2, 3, 1>

PROPERTIES

Sharing Incentive

Each user should at least get 1/n cluster assuming n users

Strategy Proof

No user should be able to get more resources by lying

Pareto Efficiency

No user can get more without reducing another user's alloc.

Envy free

No user would prefer the allocation of another user

PROPERTIES

Sharing Incentive

User is no worse off than a cluster with

1/n resources

Strategy Proof

User should not benefit by lying about demands

Pareto Efficiency

Not possible to increase one user without decreasing another

Envy free

User should not desire the allocation of another user

DRF: APPROACH

Dominant Resource

Resource user has the biggest share of → wrt total resources available

Total: <10 CPU, 4 GB>

User 1: <1 CPU, 1 GB>

Dominant resource is memory

$\frac{1}{10}$ CPU $\boxed{\frac{1}{4} \text{ memory}}$ DR → Dominant share

Dominant Share

Fraction of the dominant resource user is allocated

E.g., for User 1 this is 25% or 1/4

DRF: APPROACH

Equalize the dominant share of users

Total: <9 CPU, 18 GB>

User1: <1 CPU, 4 GB>
dom res: mem

User2: <3 CPU, 1 GB>
dom res: CPU

Total <9 CPU, 14 GB>

User	Allocation	Dominant Share
3 tasks User1	<0 CPU, 0 GB> <1 CPU, 4 GB> <2 CPU, 8 GB> <3 CPU, 12 GB>	0 2/9 4/9 2/3
2 tasks User2	<0 CPU, 0 GB> <3 CPU, 1 GB> <6 CPU, 2 GB>	0 1/3 2/3

DRF: APPROACH

Total: <9 CPU, 18 GB>

User1: <1 CPU, 4 GB> per task

<3 CPU, 12 GB> for 3 tasks

dom res: mem

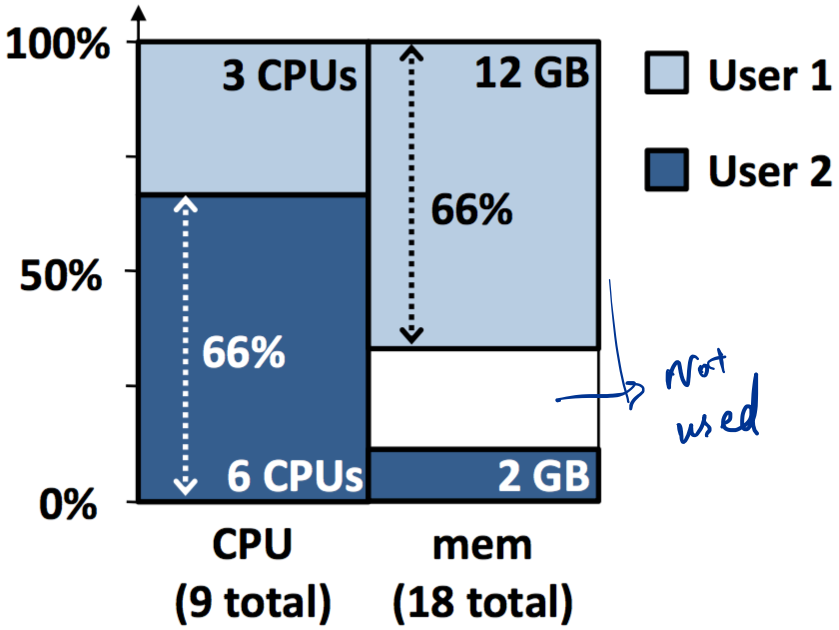
dom share: 12/18 = 2/3

User2: <3 CPU, 1 GB>

<6 GPU, 2 GB> for 2 tasks

dom res: CPU

dom share: 6/9 = 2/3



DRF ALGORITHM

Whenever there are available resources:

Schedule a task to the user with smallest dominant share

DRF ALGORITHM

Algorithm 1 DRF pseudo-code

$R = \langle r_1, \cdots, r_m \rangle$ ▷ total resource capacities
$C = \langle c_1, \cdots, c_m \rangle$ ▷ consumed resources, initially 0
$s_i \ (i = 1..n)$ ▷ user i's dominant shares, initially 0
$U_i = \langle u_{i,1}, \cdots, u_{i,m} \rangle \ (i = 1..n)$ ▷ resources given to user i, initially 0

pick user i with lowest dominant share s_i
$D_i \leftarrow$ demand of user i's next task
if $C + D_i \leq R$ **then**
  $C = C + D_i$ ▷ update consumed vector
  $U_i = U_i + D_i$ ▷ update i's allocation vector
  $s_i = \max_{j=1}^{m}\{u_{i,j}/r_j\}$
else
  return ▷ the cluster is full
end if

Fault tolerance
⇒ Tasks finishing
Update alloc vector, dominant share

choose which user to allocate

if total alloc is within cluster capacity

COMPARISON: ASSET FAIRNESS

Asset Fairness: Equalize each user's sum of resource shares

35 CPU
35 GB

Violates Sharing Incentive

U1: < 50% of CPU, memory

Consider total of 70 CPUs, 70 GB RAM
U1 needs <2 CPU, 2 GB RAM> per task = 4 units
U2 needs <1 CPU, 2 GB RAM> per task = 3 units

Asset Fair Allocation:
U1: 4, .. 12 60
U2: 3, 6 ... 12 ... 60
Total: 10 24 120

U1 <30, 30>
15 task

U2 <20, 40>
20 tasks

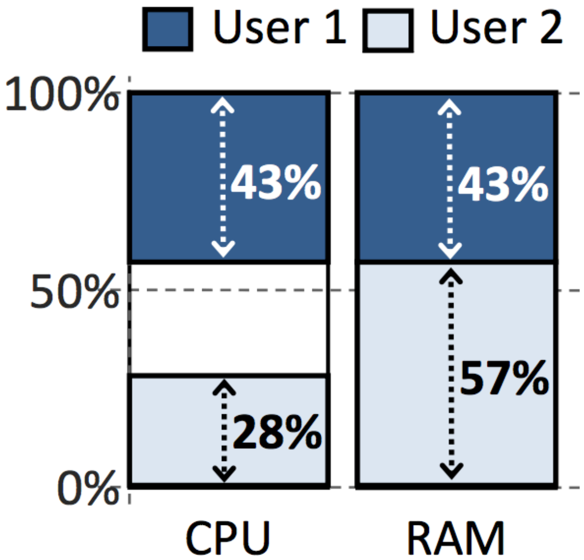
COMPARISON: ASSET FAIRNESS

Asset Fairness: Equalize each user's sum of resource shares

Violates Sharing Incentive

Consider total of 70 CPUs, 70 GB RAM
U1 needs <2 CPU, 2 GB RAM> per task
U2 needs <1 CPU, 2 GB RAM> per task

Asset Fair Allocation:
U1: 15 tasks: 30 CPU, 30 GB (Sum = 60)
U2: 20 tasks: 20 CPU, 40 GB (Sum = 60)



COMPARISON: CEEI

CEEI: Competitive Equilibrium from Equal Incomes

- Each user receives initially 1/n of every resource,
- Subsequently, each user can trade resources with other users in a perfectly competitive market
- Computed by maximizing product of utilities across users

COMPARISON: CEEI

Total: <9 CPU, 18 GB>   User1: <1 CPU, 4 GB>   User2: <3 CPU, 1 GB>

x tasks U1 , y tasks U2

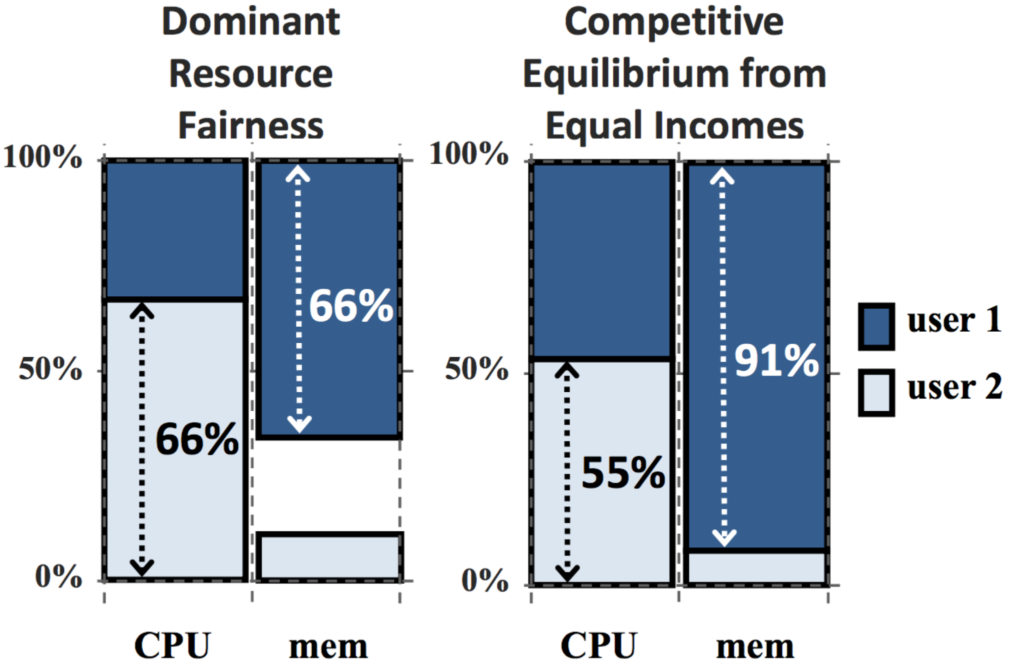
$$\max\,(x \cdot y)$$

subject to

$$\begin{aligned} x + 3y &\leq 9 \\ 4x + y &\leq 18 \end{aligned}$$

$y = 1.62$

$x = 4.05$



CEEI: STRATEGY PROOFNESS

max $x \cdot y$
$x + 3y <= 9$
$4x + 2y <= 18$

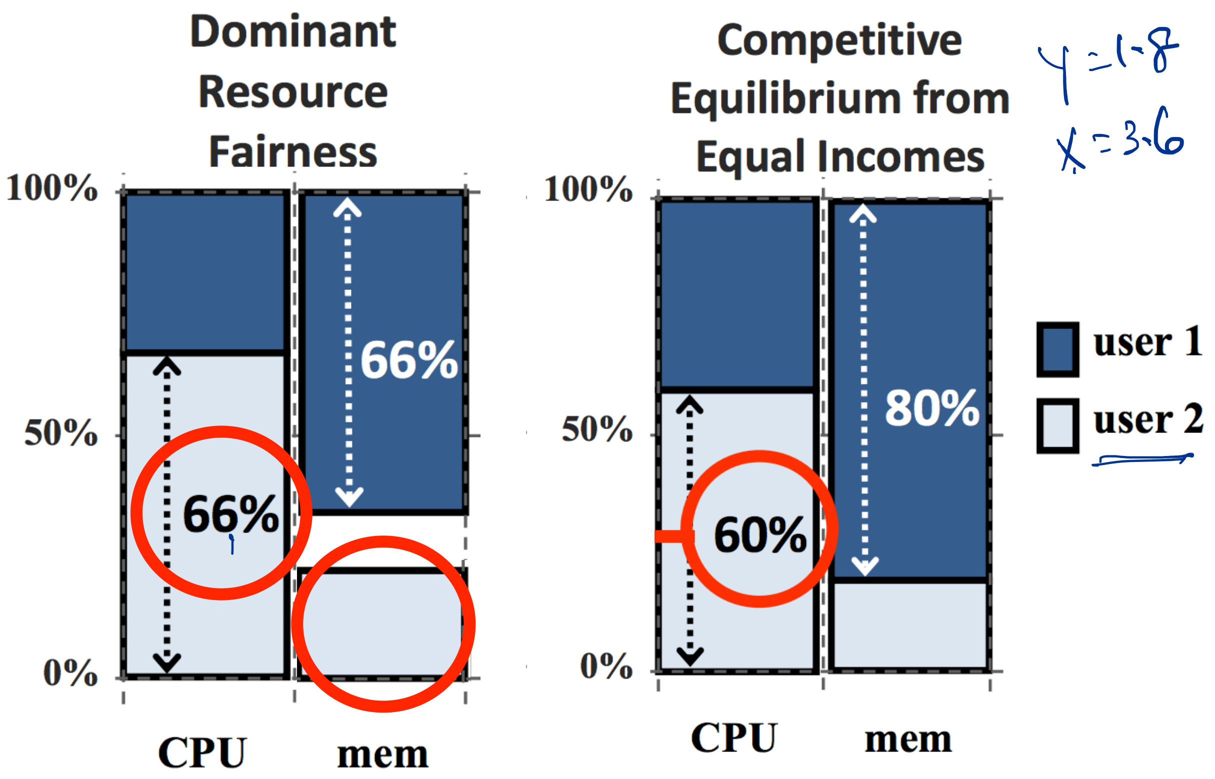
$y = 1.8$
$x = 3.6$

Total: <9 CPU, 18 GB>

User2 Before:
DRF: 66% CPU, 16% mem
CEEI: 55% CPU, 9% mem

Total: <9 CPU, 18 GB>

User1: <1 CPU, 4 GB>
User2: <3 CPU, 2 GB>



COMPARISON

Property	Allocation Policy Asset	CEEI	DRF
Sharing Incentive	X	✓	✓
Strategy-proofness	✓	X	✓
Envy-freeness	✓	✓	✓
Pareto efficiency	✓	✓	✓
Single Resource Fairness	✓	✓	✓
Bottleneck Fairness		✓	✓
Population Monotonicity	✓		✓
Resource Monotonicity			

Table 2: Properties of Asset Fairness, CEEI and DRF.

DISCUSSION

https://forms.gle/s9nm7Gr1uz8Xsn3s5

Consider a system with 100 units of CPU, 50 units of memory and 200 units of disk. Consider three users with the following requirements

Alice (4 CPU, 1 memory, 1 disk)

Bob (1 CPU, 4 memory and 4 disk)

Carol (1 CPU, 2 memory and 16 disk)

share

List the dominant resource as defined in DRF for Alice, Bob and Carol

Alice : CPU

Bob : Memory

Carol : Disk

What would be the final task allocation in the given cluster for Alice, Bob and Carol ?

A: <4, 1, 1>
B: <1, 4, 4>
C: <1, 2, 16>
T: <100, 50, 200>

Total

C	M	D
10	8	22
⋮	⋮	⋮
60	48	132

		D S
A:	~~1~~ 2	2/25 =
B:	1	2/25
C:	1	2/25
	⋮	
A:	12	6/25
B:	6	6/25
C:	6	6/25

Carol → 12, 6, 7

Alice → 13, 6, 6

14, 6, 6

What could be one workload / cluster scenario where DRF implemented on Mesos will NOT be optimal?

1. Heterogeneous task lengths

2. Fragmentation = Packing ?

3. Placement Preferences

NEXT STEPS

Next Week: Machine Learning

Assignment 2 out tonight!

Course projects: Office hours