Static Resource Allocation

Mark Towers

September 9, 2019

1 Problem Statement

In prior research into resource pricing for cloud based resource allocation, the jobs considered have had fixed resource requirements that could then be requested from a cloud provider. However this could result in different job's fighting over a single resource instead of sharing the resource fairly due to the lack of knowledge of each other. In this research we consider jobs that have requirements that are constrainted by a deadline allowing the cloud services the ability to dynamically allocate resources depended how the usage of it's resources by other jobs and allowing resources to be shared more effectively. We have also developed a greedy algorithm to that finds near-optimals of problem cases to maximise the social welfare where there it total knowledge. However in real life, two problems arise of servers wished to be fairly payed for the work it does and jobs may be want to reveals their job utility. To solve these problems we have created an iterative auctions that we compare to VCG auctions.

2 Problem Case

2.1 Variable

There are J jobs, indexed with $j = 1, \ldots, J$ and I servers, indexed with $i = 1, \ldots, I$.

- $x_{i,j} \in \{0, 1\}$ indicates whether the job j was done on server i
- $s_{i,j}$ the rate that the program is loaded at (MB/s)
- $w_{i,j}$ the rate of computation (TFlop/s)
- $r_{i,j}$ the rate that the result's data is sent back (MB/s)

2.2 Constants

Server - i

- Maximum storage - S_i (MB)

- Maximum computation capacity - W_i (TFlop/s)
- Maximum communication bandwidth - R_i (MB/s)

Job - j

- Required storage - s_j (MB)
- Required computation capacity - w_j (TFlop)
- Required data for results - r_j (MB)
- Utility - U_j ($)
- Deadline - D_j (s)

2.3 Optimisation

$$max \sum_{j=1}^{J} U_j x_{i,j} \qquad \forall i = 1, \ldots, I \tag{1}$$

2.4 Constraints

Job to server allocation

$$\sum_{i=1}^{I} x_{i,j} \leq 1 \qquad \forall j = 1, \ldots, J \tag{2}$$

$$x_{i,j} \in \{0, 1\} \qquad \forall i = 1, \ldots, I; j = 1, \ldots, J \tag{3}$$

Server resource available

$$\sum_{j=1}^{J} s_j x_{i,j} \leq S_i \qquad \forall i = 1, \ldots, I \tag{4}$$

$$\sum_{j=1}^{J} w_{i,j} x_{i,j} \leq W_i \qquad \forall i = 1, \ldots, I \tag{5}$$

$$\sum_{j=1}^{J} (r_{i,j} + s_{i,j}) x_{i,j} \leq R_i \qquad \forall i = 1, \ldots, I \tag{6}$$

Process completed within deadline

$$\frac{S_j}{s_{i,j}} + \frac{W_j}{w_{i,j}} + \frac{R_j}{r_{ij}} \leq D_j \qquad \forall i = 1, \ldots, I; j = 1, \ldots, J \tag{7}$$

Resource usage

$$0 \leq s_{i,j} \qquad \forall i = 1, \ldots, I; j = 1, \ldots, J \tag{8}$$

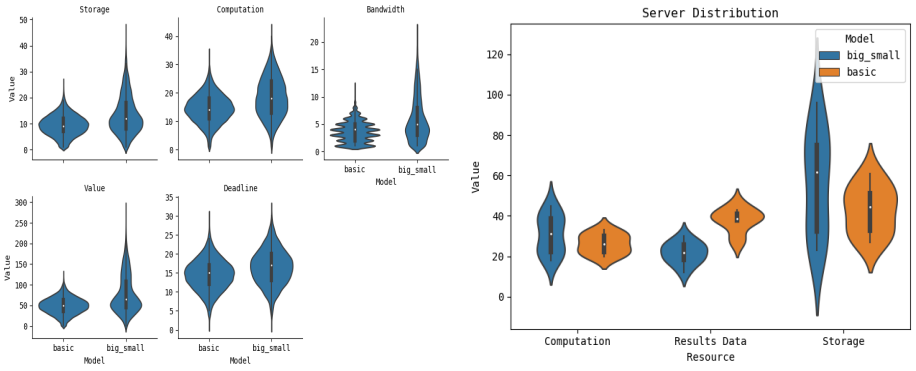
$$0 \leq w_{i,j} \qquad \forall i = 1, \ldots, I; j = 1, \ldots, J \tag{9}$$

$$0 \leq r_{i,j} \qquad \forall i = 1, \ldots, I; j = 1, \ldots, J \tag{10}$$

2.5 Problem Case Explanation

- Equation 1 is the objective function that maximises the sum of the job utility for jobs completed.
- Equation 2 and 3 enforce that a job is only done on a single server.
- Equations, 4 to 6, ensures that the server resource used are within the maximum resources available.
- Equation 7 enforces that the job will be completed within the deadline on only the servers that is a job runs on.
- Equations, 8 to 10, ensures that resource speeds are within a valid range of greater than 0

2.6 Example cases



(a) Example jobs (b) Example Servers

Figure 1: Example jobs and servers

2.7 Model creation

To create each of the jobs and servers we choose a mean and standard deviation for each attribute that is then used to generate a random number from a normal

distribution. We normalise the value generated to make it an integer and check that the value is greater than zero.

Job Attribute	Mean	Std	Server Attribute	Mean	Std
Storage	10	4	Storage	45	10
Compute	15	5	Compute	30	5
Results data	5	4	Bandwidth	40	5
Utility	50	20			
Deadline	15	4			

3 Greedy Algorithm

The problem case laid out above is a knapsacking variation which has been shown to be NP-Hard to solve optimality. Therefore while a solution can be found through integer programming, it is extremely slow to do and become impossible to solve for problem cases with more than 15 jobs and 3 servers on a regular computers.
This means that to calculate the optimal solution, with full knowledge, then a greedy algorithm must be developed. To do this, I have created two greedy algorithm that are compared and explained below.

3.1 Greedy implementation

The first algorithm is a three stage process where the jobs are sorted on how good they are based on a metric like the ratio of utility to resource's required. Then For each jobs in order of value, a server is selected for the jobs to run on based on the another metric like which server has the least resources available. Resources are then allocated to a job based on a third policy that calulcates the values for each possible allocation like the percentage of the available resource that the allocated resources would use. The maximum of this is then chosen and allocated for the job on the server with the process repeated till no jobs can be allocated to any server.
This allows for a large amount of experimentation due to the number of permutations that exists for different policies of the different stages however any errors are propagated through the system more easily and the job value doesnt change has the server's resources are allocated off.
Below is the algorithm used to generate the solution written in python with the github containing a large collection of policies for each stage.

```
job_values = sorted(((job, value_density.evaluate(job)) for job
    ↪ in jobs), key=lambda jv: jv[1], reverse=True)
```

```
# Loop through all of the job in order of values
for job, _ in job_values:
    # Allocate the server using the allocation policy function
    allocated_server = server_selection_policy.select(job, servers
        ↪ )

    # If an optimal server is found then calculate the resource
        ↪ allocation policy
    if allocated_server:
        value, (s, w, r) = resource_allocation_policy.allocate(job
            ↪ , allocated_server)
        job.allocate(s, w, r, allocated_server)
        allocated_server.allocate_job(s, w, r, job)
```

3.2 Matrix Greedy implementation

The second algorithm is significantly simpler as it only uses a single policy and hopes to prevent many of the possible problems found with the first algorithm. I believe that it will be easier to prove any approximations of the algorithm with this version due to having a single policy compared to required three policies of the last version.
This algorithm works by thinking about the resource allocation first instead of last and so for each job and server the best resource allocation is found using a metric like the product of the job utility and the percentage of the server available resources after allocation. This is then used to generate a matrix of jobs and servers allocation with the value being the value of the best resource allocation for the job and server pair. From this matrix then the max value is found of this with the job then being allocated to the server and the job removed from the matrix. This is iteratively done updating each time till all of the jobs are allocated or none of the jobs can be allocated.
Due to this simplification with only a single policy, I believe that this should make the algorithm better however as the policy must include the job utility and some function of the server available resources and the resources allocation this is less simple to find good algorithms. Therefore this is something I am still experimenting with and to prove if any approximations are possible.

```
def allocate_resources(job: Job, server: Server, value_policy:
    ↪ MatrixPolicy):
    # Calculates the best resource allocation by loop over all of
        ↪ the possible allocations that satisfy the deadline
        ↪ constraint
    return max(((value_policy.evaluate(job, server, s, w, r), s, w
        ↪ , r)
                for s in range(1, server.available_bandwidth + 1)
                for w in range(1, server.available_computation + 1)
```

```
                for r in range(1, server.available_bandwidth - s +
                    ↪ 1)
                if job.required_storage * w * r + s * job.
                    ↪ required_computation * r +
                s * w * job.required_results_data <= job.deadline *
                    ↪  s * w * r), key=lambda x: x[0])


def matrix_greedy(unallocated_jobs: List[Job], servers: List[
    ↪ Server], value_policy: MatrixPolicy):
    while unallocated_jobs:
        value_matrix = []
        # Loop through all of the jobs and servers to calculate
            ↪ their best resource allocation
        for job in unallocated_jobs:
            for server in servers:
                if server.can_run(job):
                    value, s, w, r = allocate_resources(job, server
                        ↪ , value_policy)
                    value_matrix.append((value, job, server, s, w,
                        ↪ r))

        if value_matrix:
            # Finds the maximum value within the matrix and
                ↪ allocate to job and server
            value, job, server, s, w, r = max(value_matrix, key=
                ↪ lambda x: x[0])
            job.allocate(s, w, r, server)
            server.allocate_job(job)
            unallocated_jobs.remove(job)
        else:
            # No jobs can be allocated to servers so stop
            break
```

3.3 Approximation

Both of these greedy algorithm are near optimal approximation algorithm with a lower bound of at least n/m of the optimal solution.

3.4 Greedy algorithms results

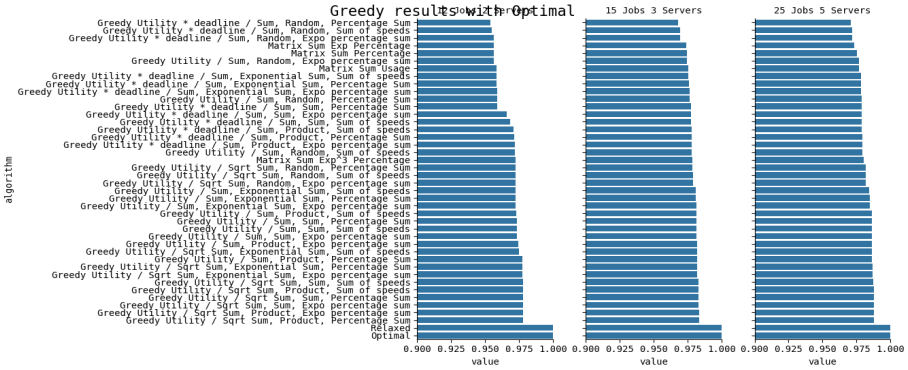


Figure 2: Greedy Results

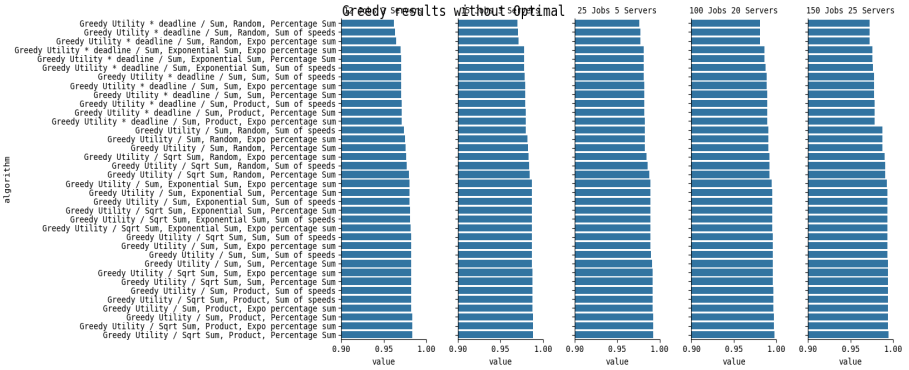


Figure 3: Greedy Results with the optimal algorithms

4 Auctions

graphicx

While the greedy algorithm allows us to find near optimals for problems where we have perfect information about the problem case like the job utility, this may be private information that the job owner doesnt want to reveal also server would want to get paid for that work that they do.

We propose a iterative auctions that occurs before any processing is done and allows the job owner doesnt have to reveal it's true utility. In normal VCG auctions, then the auctioneer will solve the problem to maximise the social welfare (total utility of allocated jobs) and must find the optimal solution. Then for each job and server, the job or server is removed from the problem case and

the optimal is solved for that as well with the server's revenue it is paid and the job's cost being the social welfare of the optimal solution minus the social welfare of the optimal without the job or server. This is extremely slow and unscalable because of this so is not used often however has the properties that it is individually rational, truthful bidding and incentive compatible.

4.1 Proposed Iterative Auction

Our iterative auctions uses the idea of VCG so that when a job asks to run on a server then the server will calculate current revenue of the jobs running minus the revenue if the new job must be running on the server plus a price increase factor. This is then the price for the job to run on the server as the new allocation with the job would be to a greater revenue for the server than is currently allocated.

4.2 Iterative Auction Properties

Auctions mainly consider four different properties that are important: Economic Efficiency, Individual Rational, Incentive Compatibility and Budget Balance. We believe that our iterative auction is Budget Balanced and Individual rational and possibly economic efficient however not incentive compatible.

4.2.1 Proving budget balance for auction

Budget balance is that the sum of revenue and the sum of prices is zero, as our algorithm doesnt require an auctioneer then budget balance is true. The algorithm can be run with a centralised auctioneer or in a decentralised manner such that jobs don't need to be aware of each others existence with a job only directly communicating with the server.

4.2.2 Proving individual rational

The utility of all partitions are non-negative therefore the sum of all server revenue $\geq$ the sum of revenue without jobs. So the revenue of a job must be non-negative.

4.2.3 Economic efficient and Incentive compatible

Our algorithm is not economically efficient or incentive compatible due a server only considering itself in choosing its price and is unaware of the price being charged by other servers. Because of this then the allocation of jobs to server may not be optimal resulting in a suboptimal economic efficient however in over 50

4.3 Iterative Auction algorithm

```
unallocated_jobs = jobs
while len(unallocated_jobs):
    # Select a job, can be at random
    job = unallocated_jobs[0]

    # Calculate the minimum job price on all of the servers
    job_price, allocation_info = min((evaluate_job_price(job,
        ↪ server, epsilon=epsilon)
                                for server in servers), key=
                                    ↪ lambda bid: bid[0])
        # Check if the job can pay the minimum price
        if job_price <= job.utility:
            # Uses the allocation info to create the new
                ↪ allocation on the selected server
            allocate_job(job_price, job, allocation_info,
                ↪ unallocated_jobs)
        else:
            # Remove job as the job can be run ever at a price
                ↪ lower than the job's value utility
            unallocated_jobs.remove(job)
```

5 Iterative Auction Results

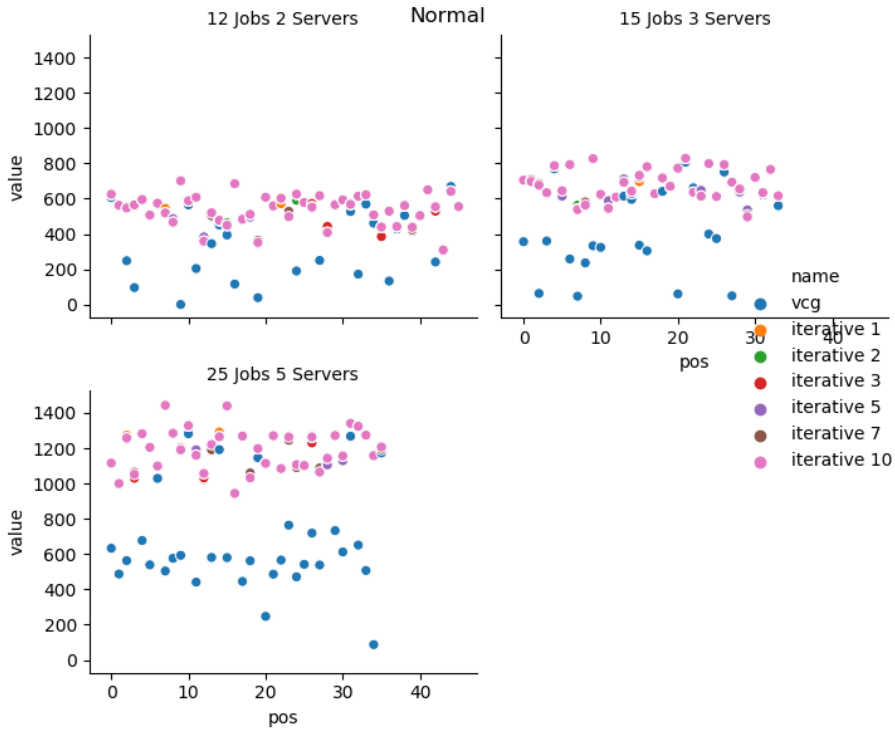


Figure 4: Multiple Price auction results

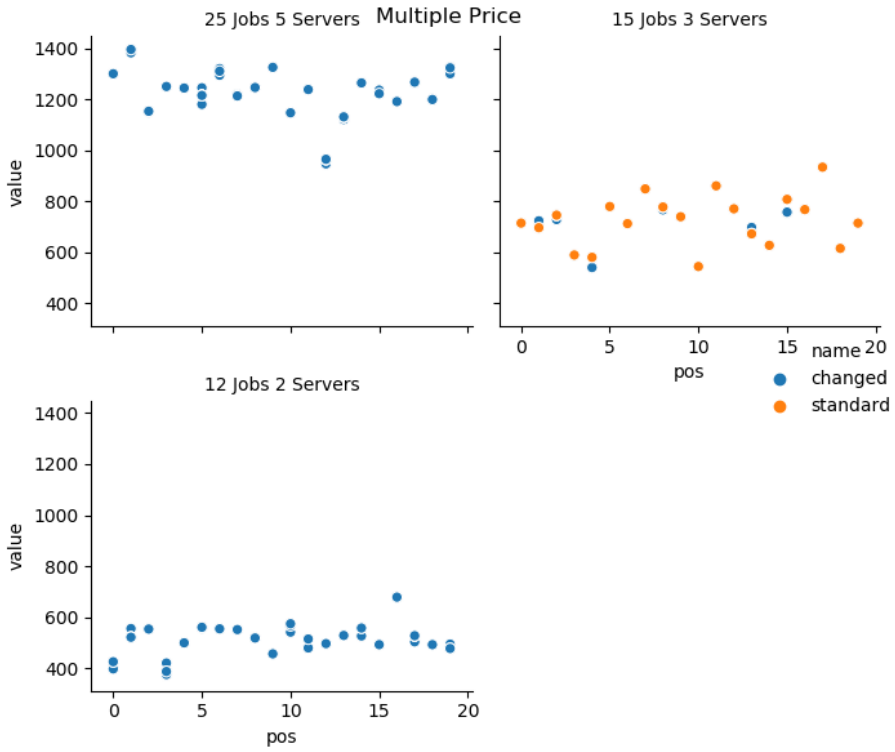


Figure 5: Multiple Price auction results

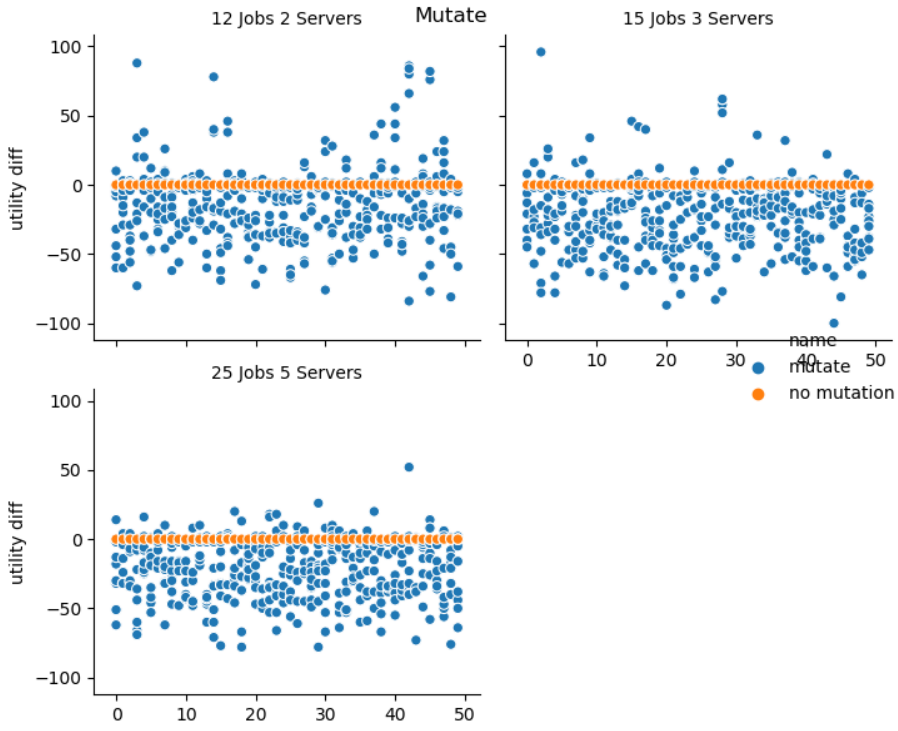


Figure 6: Mutated Jobs and servers in auctions