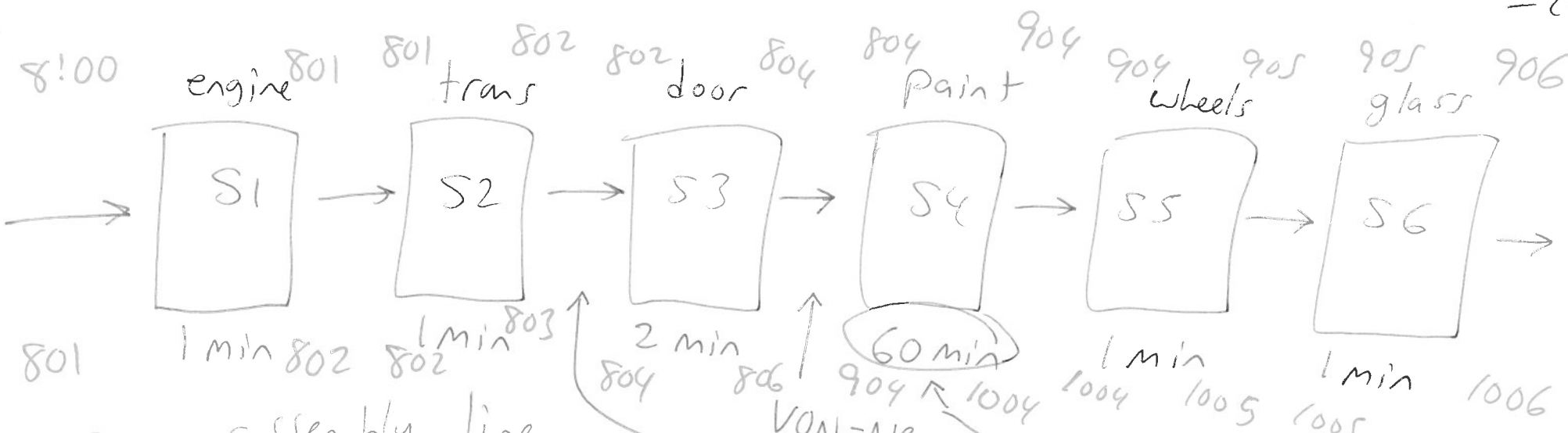
Pipelines: ≈ assembly line

multiple hardware working in parallel (and series) to fetch, decode, and execute instructions in order to maximize bandwidth.

eg MIPS

Bandwidth: # of instructions that are completed in a unit of time

Latency: the total time for one instruction to pass through the pipeline from start to finish.



Car assembly line
6 stages

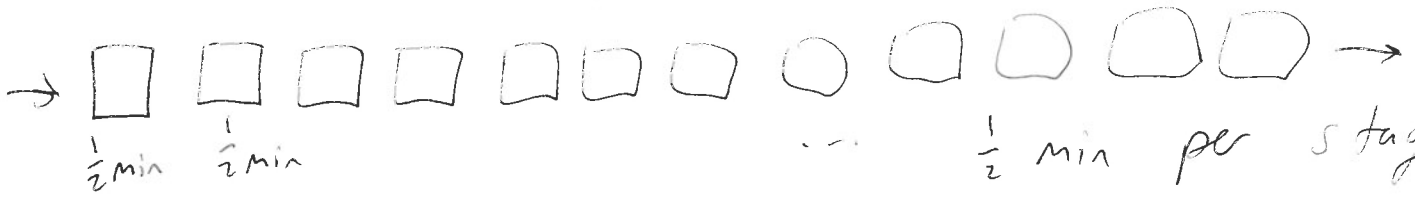
Latency: 66 minutes / car

Bandwidth: 1 car / 60 mins

Fix bottlenecks: superscalar architecture:
put extra at slowest stages:

Latency: $\frac{6 \text{ min}}{\text{car}}$

BW: $\frac{1 \text{ car}}{\text{min}}$



ie doubled the pipeline length

Latency: 6 min/car

Bw: 2 car/min

} doubled the Bw but same latency

we like deep pipelines

Disneyland

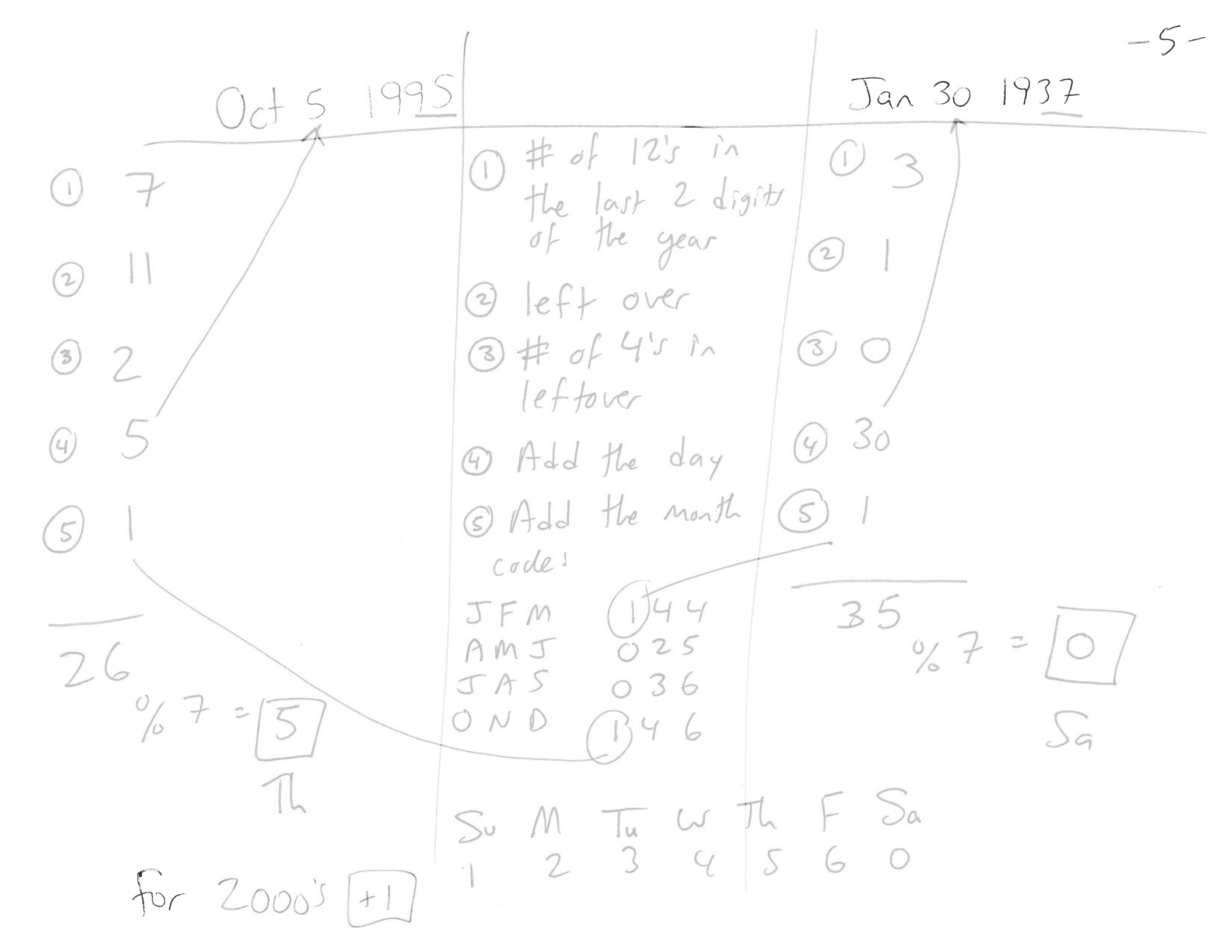
It's a small world — 15 min — short line

Space Mountain — 3 min — long line

Latency: want long ride short line

Bw: more riders per hour

No line long ride } lots of hardware


Oct 5 1995
Jan 30 1937
① # of 12's in the last 2 digits of the year
② left over
③ # of 4's in leftover
④ Add the day
⑤ Add the month code:
JFM 144
AMJ 025
JAS 036
OND 146
① 7
② 11
③ 2
④ 5
⑤ 1
26
%7 = 5
Th
① 3
② 1
③ 0
④ 30
⑤ 1
35
%7 = 0
Sa
Su M Tu W Th F Sa
1 2 3 4 5 6 0
for 2000's +1

April 1 1984 Sunday
Dec 25 1950 Monday

this type of question is on quiz (bonus)

Chapter 2 of textbook

pipeline: multiple HW

purpose: FDE instructions in parallel to maximize BW

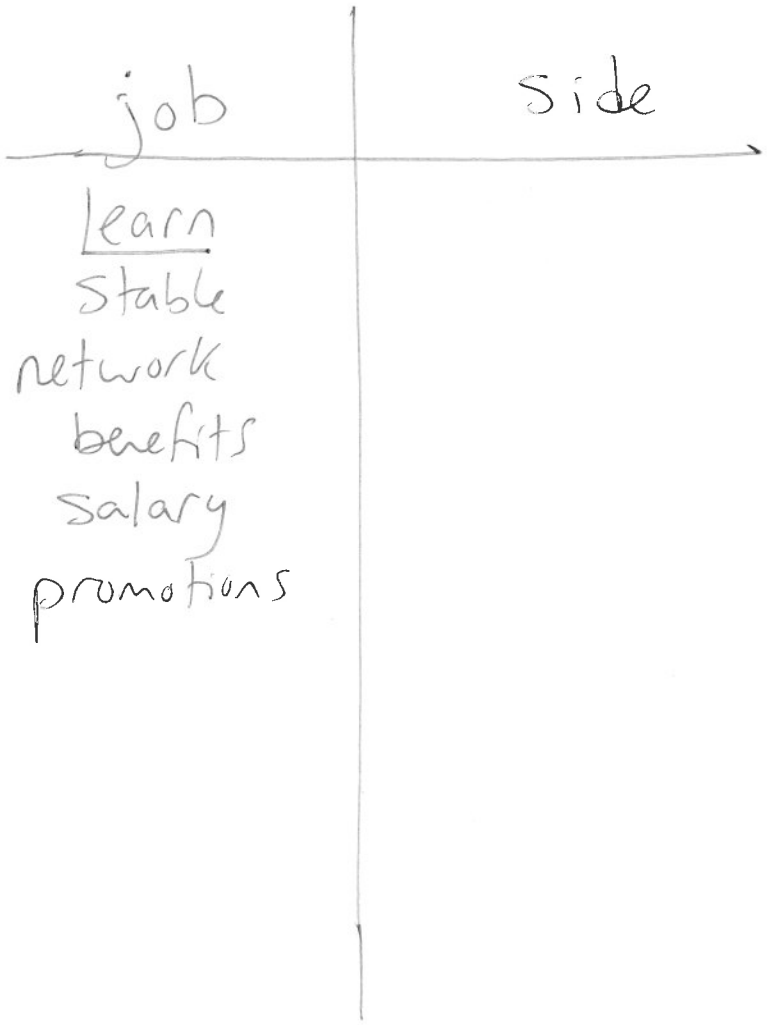
superscalar arch: @ bottlenecks

BW: instructional BW: prefetch, FDE maximize MIPS

Latency: minimize latency

Netflix: want high BW

~~low~~ high latency: long load time then ok


job
side
Learn
Stable
network
benefits
salary
promotions