

The Kubernetes Control Plane

...For Busy People Who Like Pictures

Daniel Smith
dbsmith@google.com
github: [lavalamp](#)
twitter: [originalavalamp](#)
SIG API Machinery Co-chair, co-TL
Staff Software Engineer @ Google

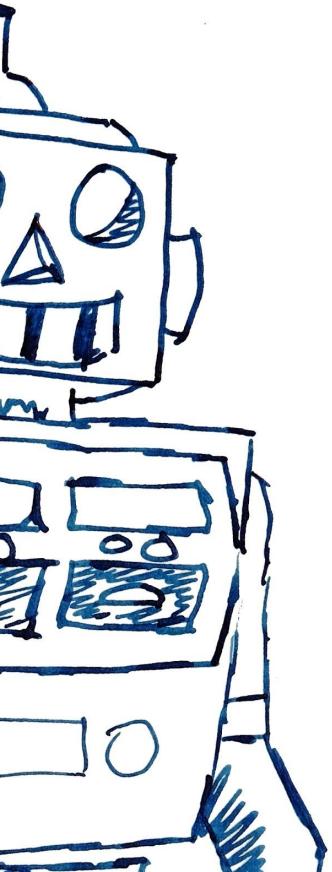
THE KUBERNETES
CONTROL
PLANE

FOR BUSY PEOPLE
WHO LIKE PICTURES

THE KUBERNETES CONTROL PLANE

FOR BUSY PEOPLE

WHO LIKE ^{BAD} PICTURES



THE KUBERNETES CONTROL PLANE

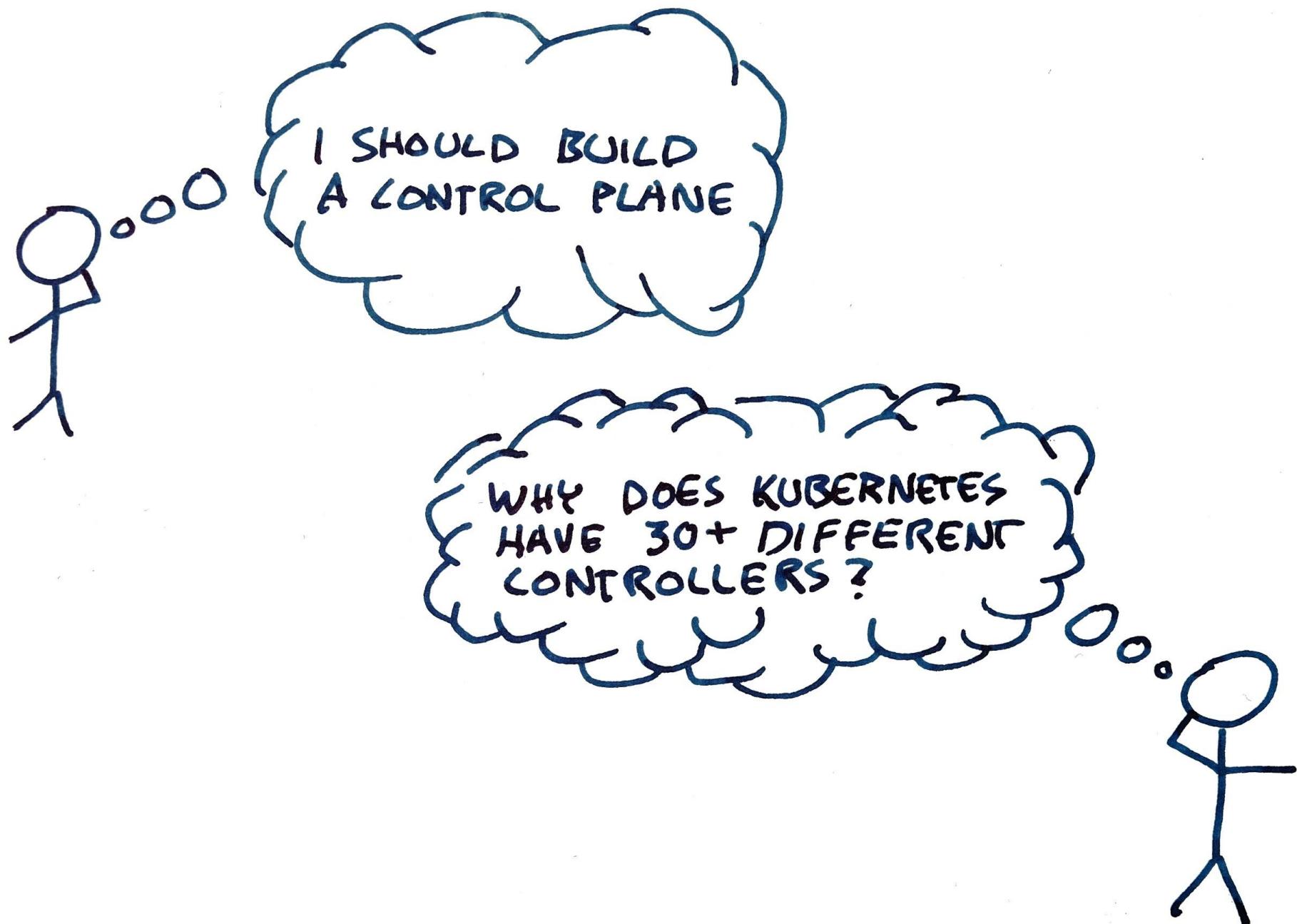
FOR BUSY PEOPLE
WHO LIKE ^{AND} PICTURES



DANIEL SMITH

STAFF SOFTWARE ENGINEER — GOOGLE
LAVALAMP — GITHUB

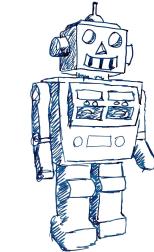
ORIGINALAVALAMP — TWITTER
SIG API MACHINERY
CO-CHAIR * CO-TL



TRILOGY

COPENHAGEN 2018: KUBERNETES-STYLE APIs OF THE FUTURE
SEATTLE 2018: A VISION FOR API MACHINERY
BARCELONA 2019: THE KUBERNETES CONTROL PLANE
→ FOR BUSY PEOPLE WHO LIKE PICTURES

YOU
ARE
HERE



THE KUBERNETES API
IS ABOUT HUMANS AND
MACHINES WORKING TOGETHER.

THE KUBERNETES API
IS ABOUT HUMANS AND
MACHINES WORKING TOGETHER.

... YOU CAN'T DO THAT
WITHOUT SOME MACHINES!!

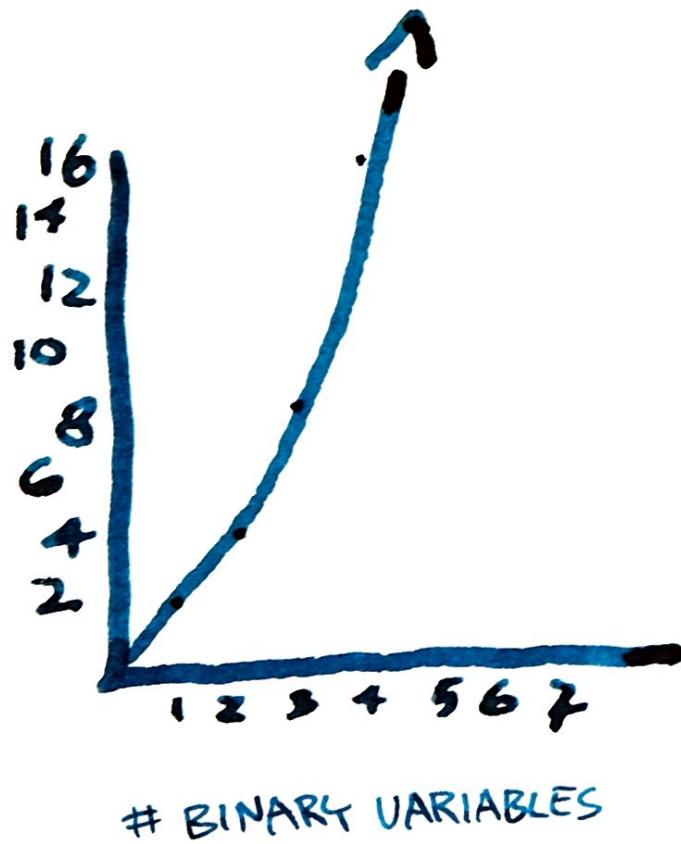
CONTROLLERS: THE  IN THE MACHINE

CONTROL
THEORY!

THE AGE-OLD DEBATE: ~~NATURE VS NURTURE~~
~~STATE MACHINE~~
VS
~~CONTROL LOOP~~

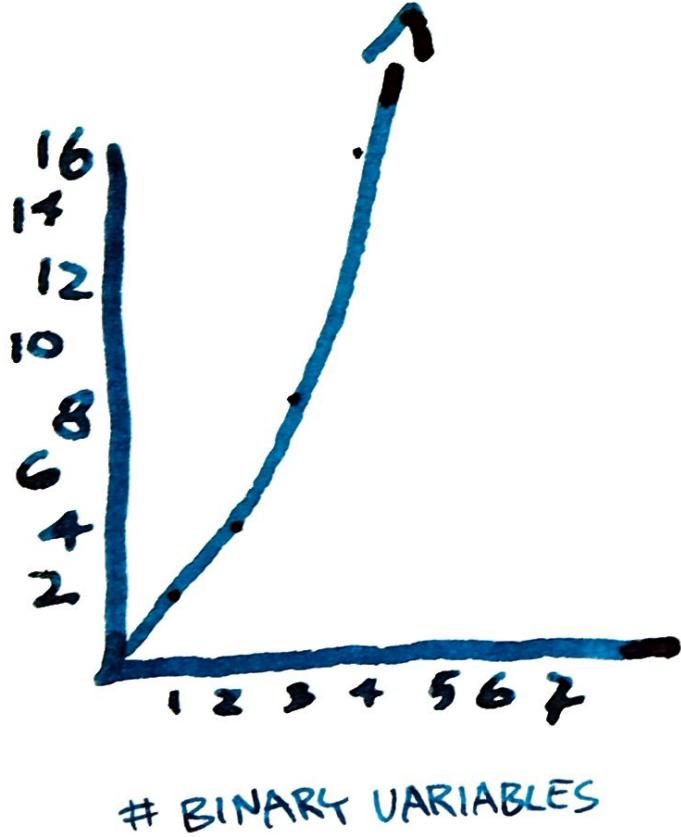
STATE MACHINES

Possible states
you most handle
perfectly



STATE MACHINES

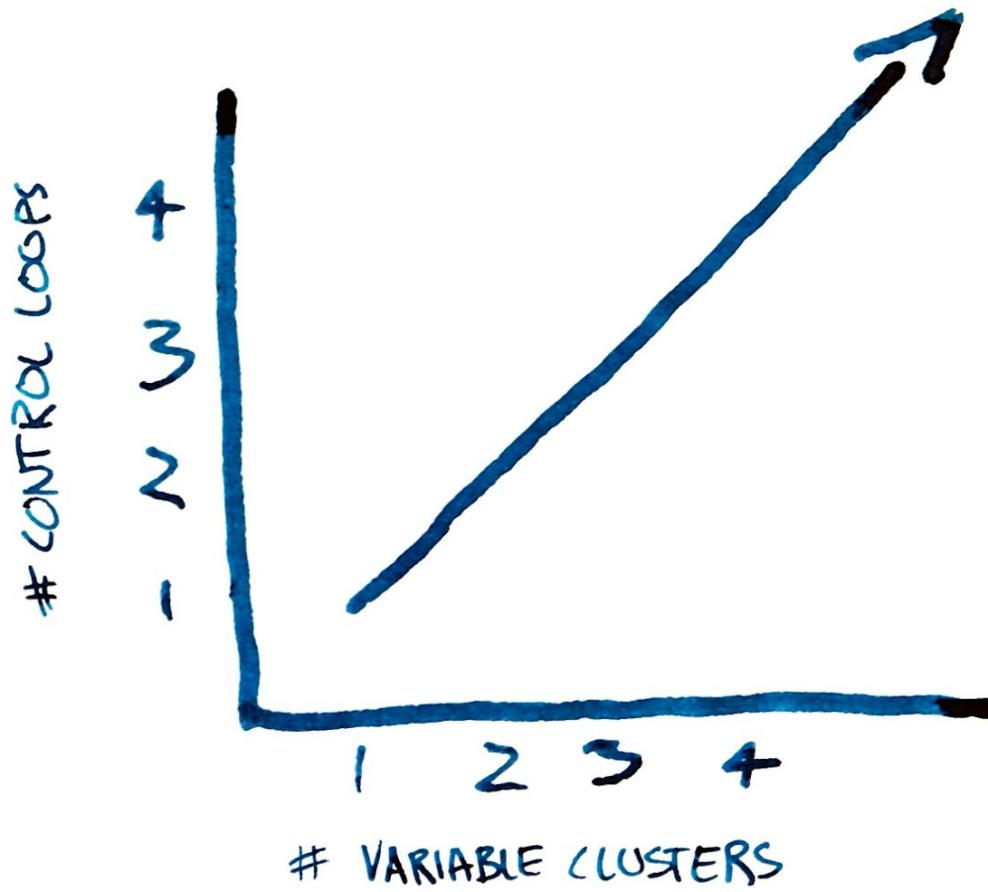
Possible states
you most handle
perfectly



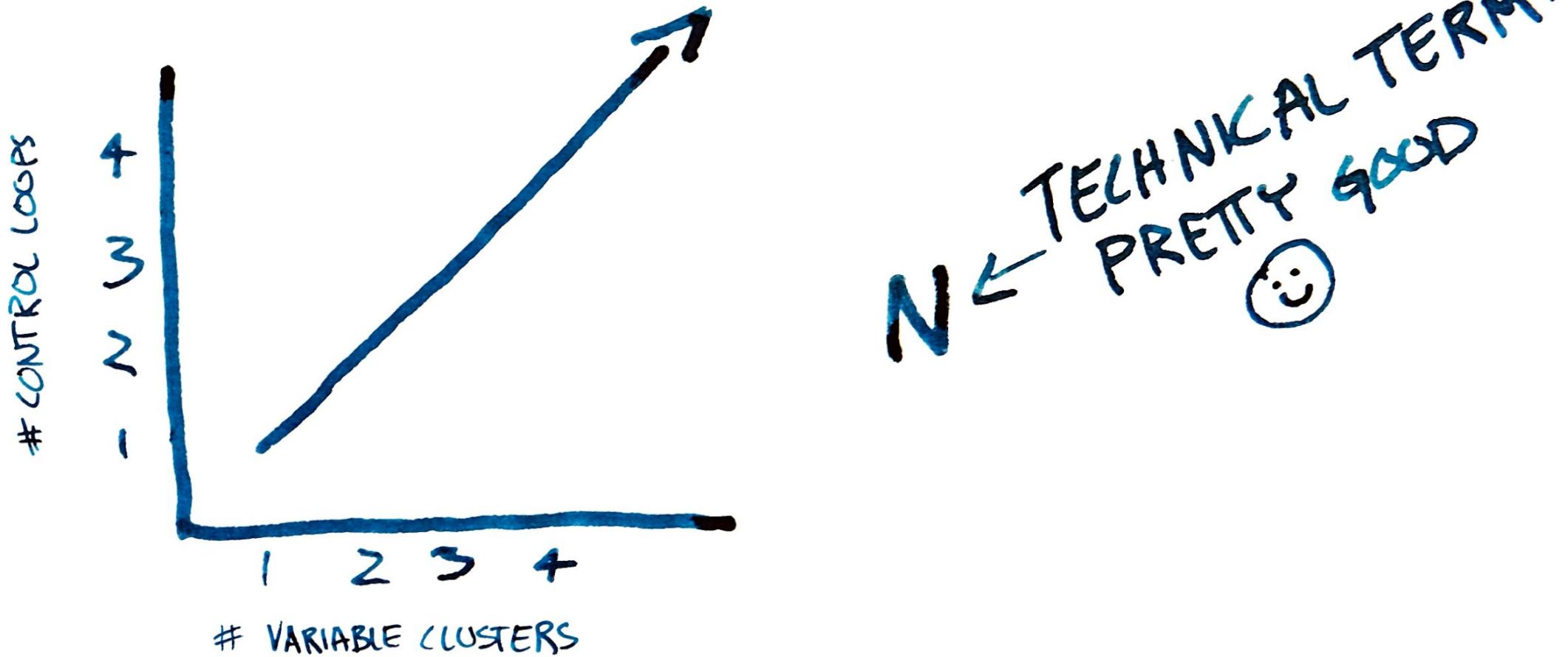
2^n ← TECHNICAL TERM:
NOT GOOD

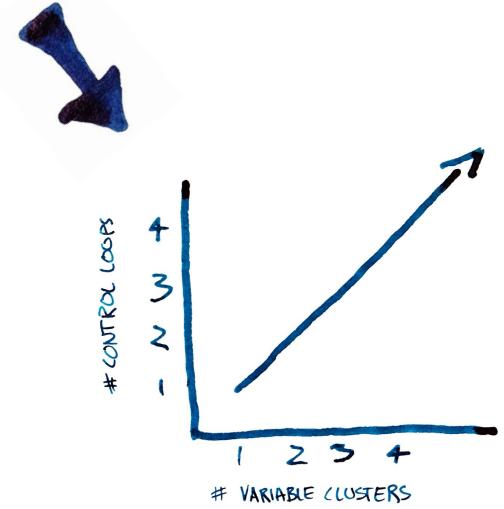
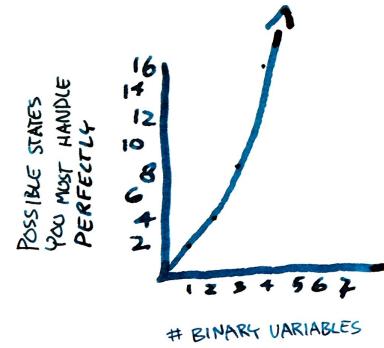
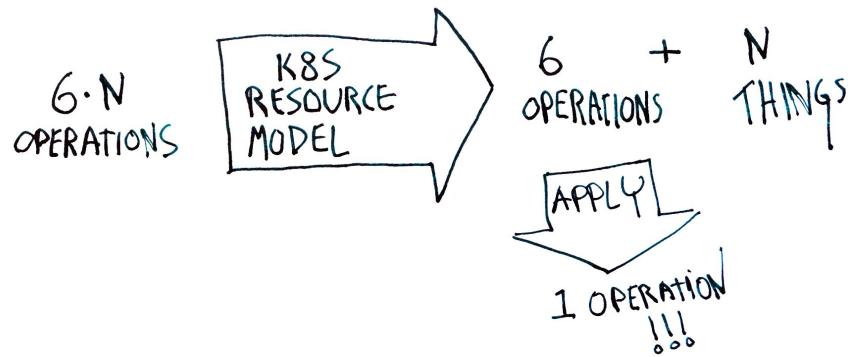
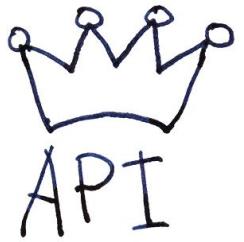


CONTROLLERS



CONTROLLERS





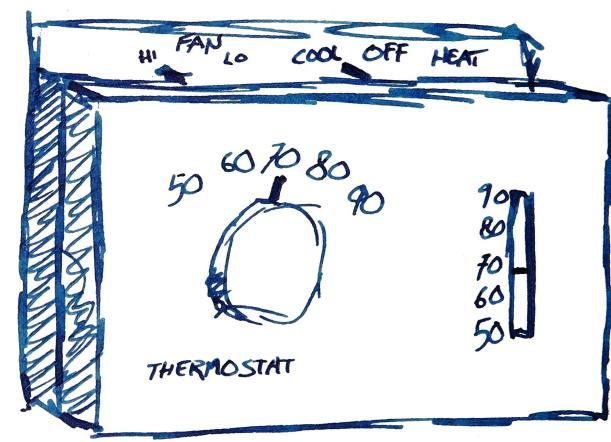
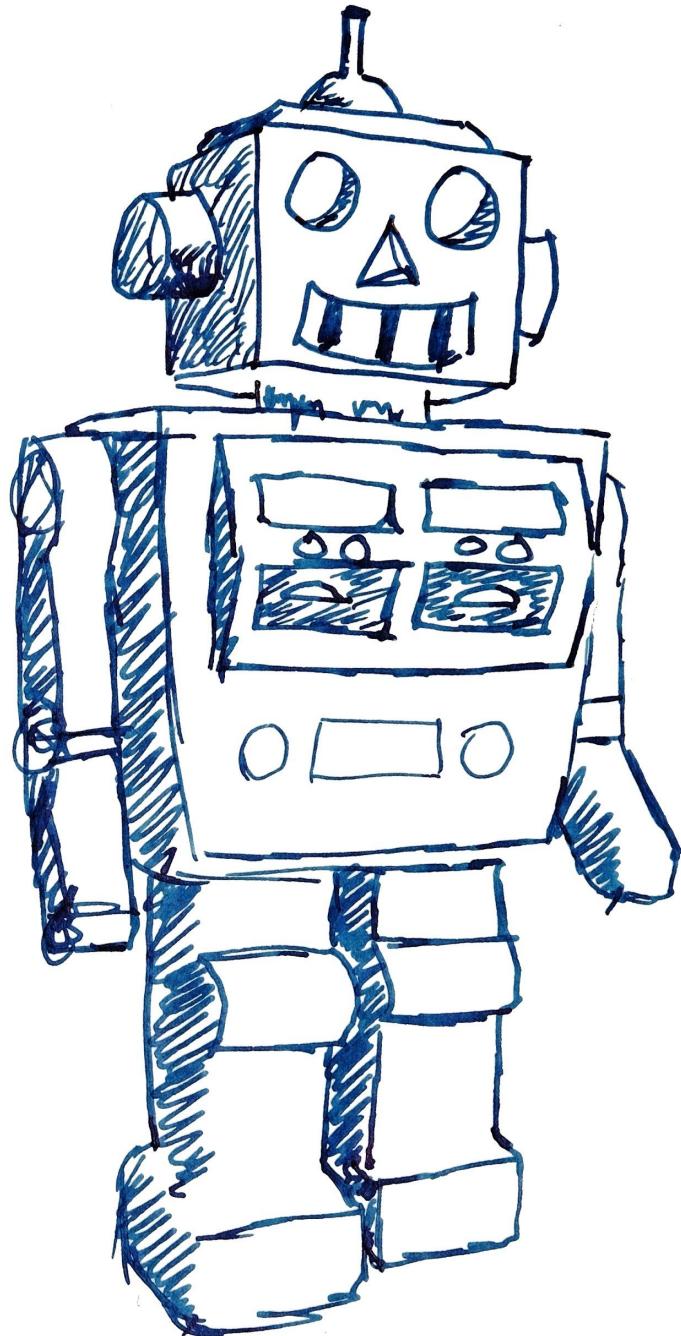
INTEGRATION COMPLEXITY VS IMPLEMENTATION COMPLEXITY

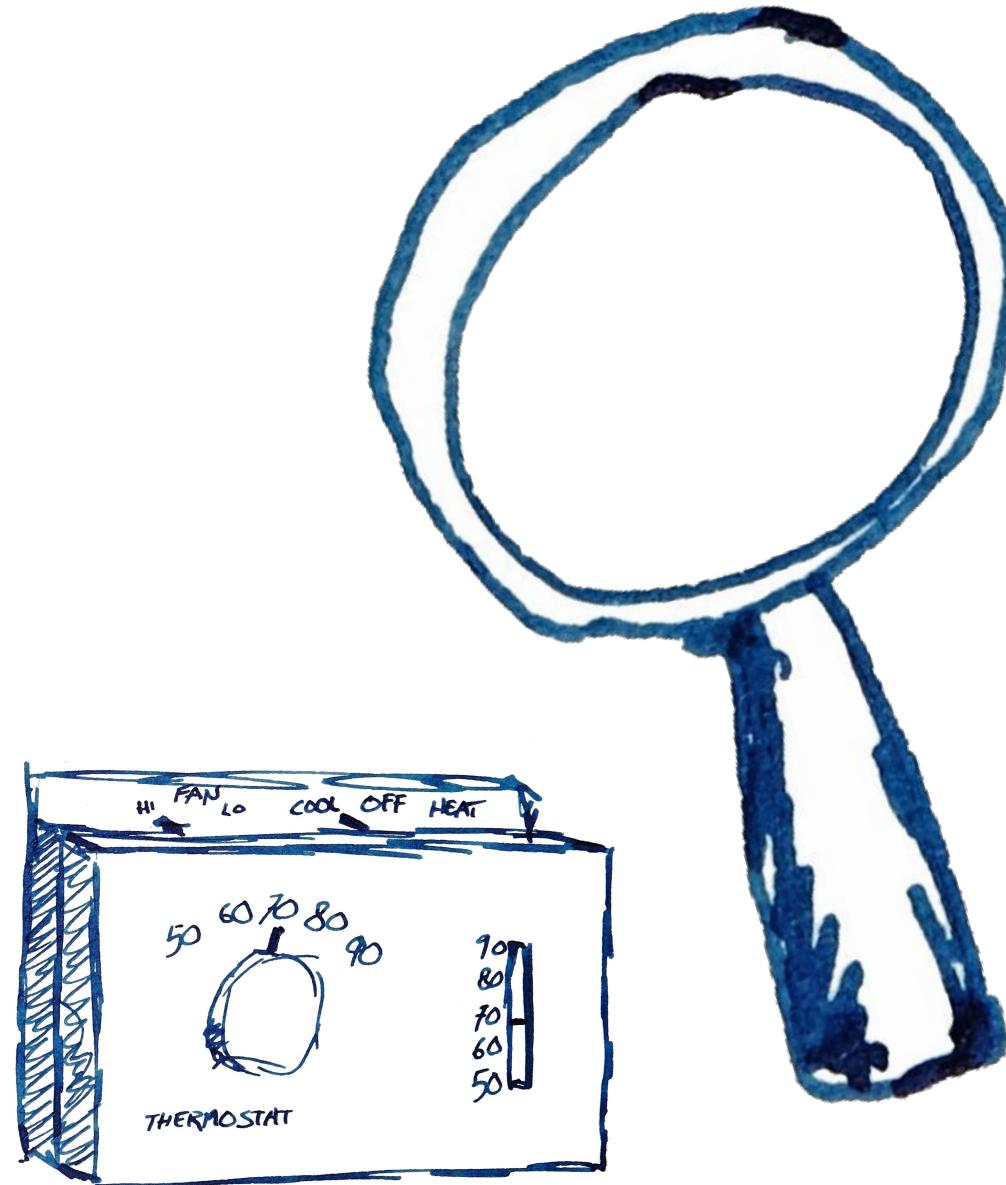
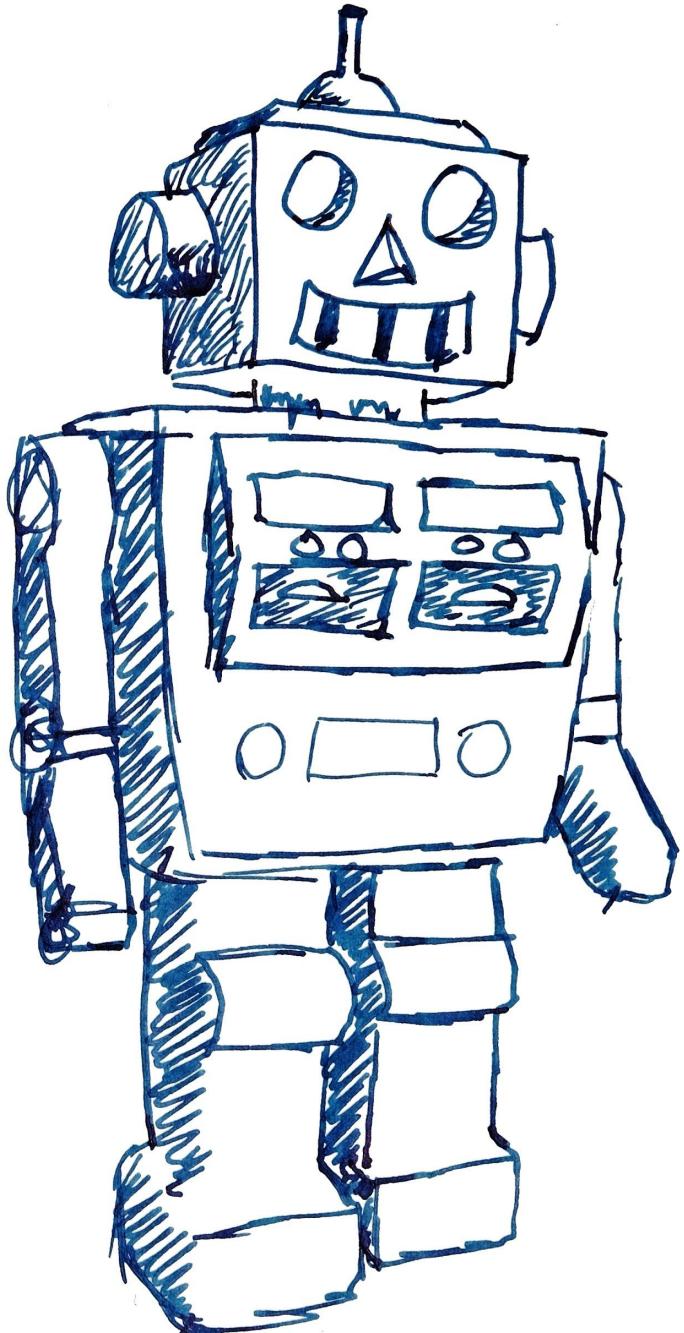
GLOBALLY EASIER
LOCALLY HARDER

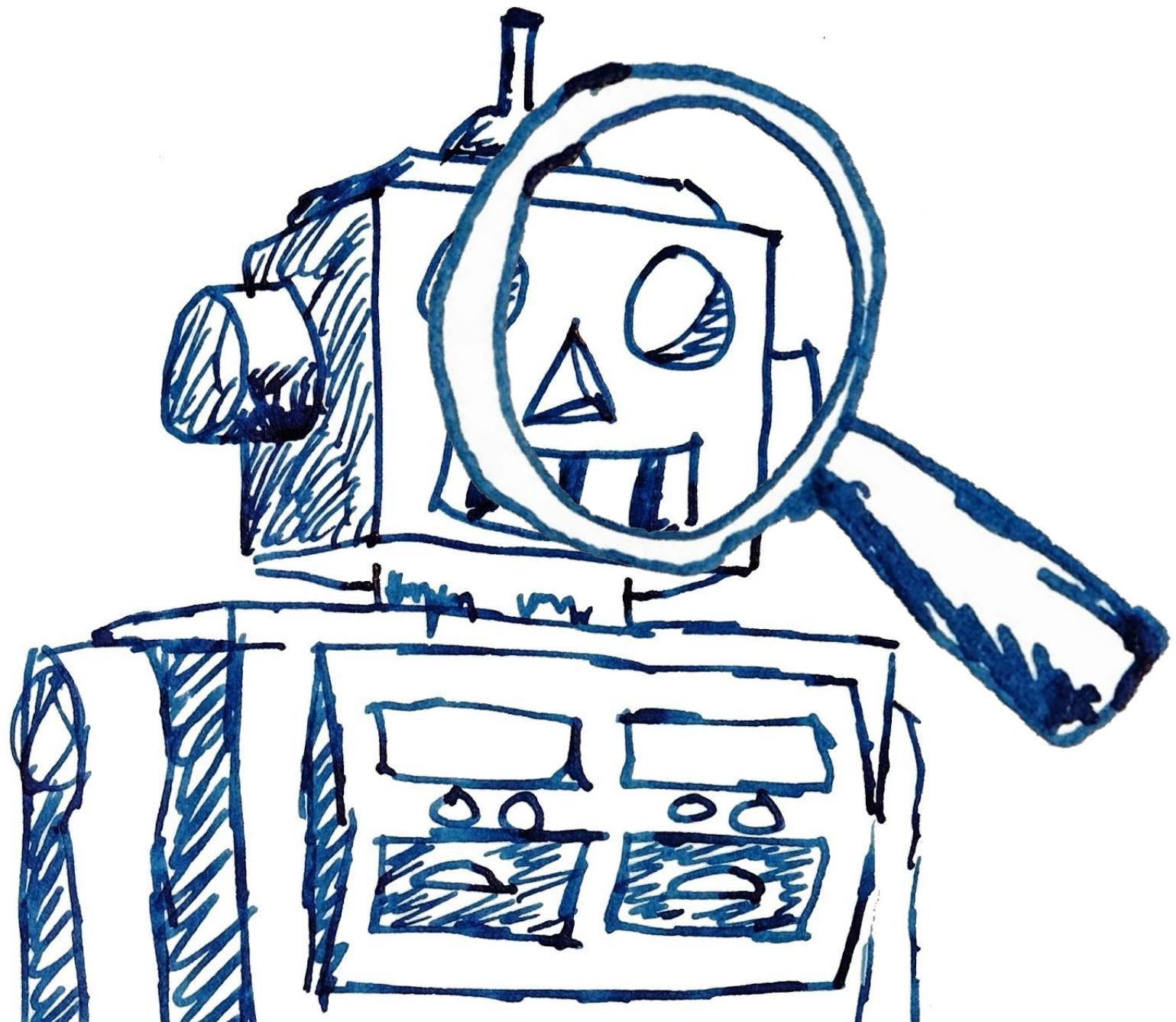
AN IDEAL KRM CONTROLLER

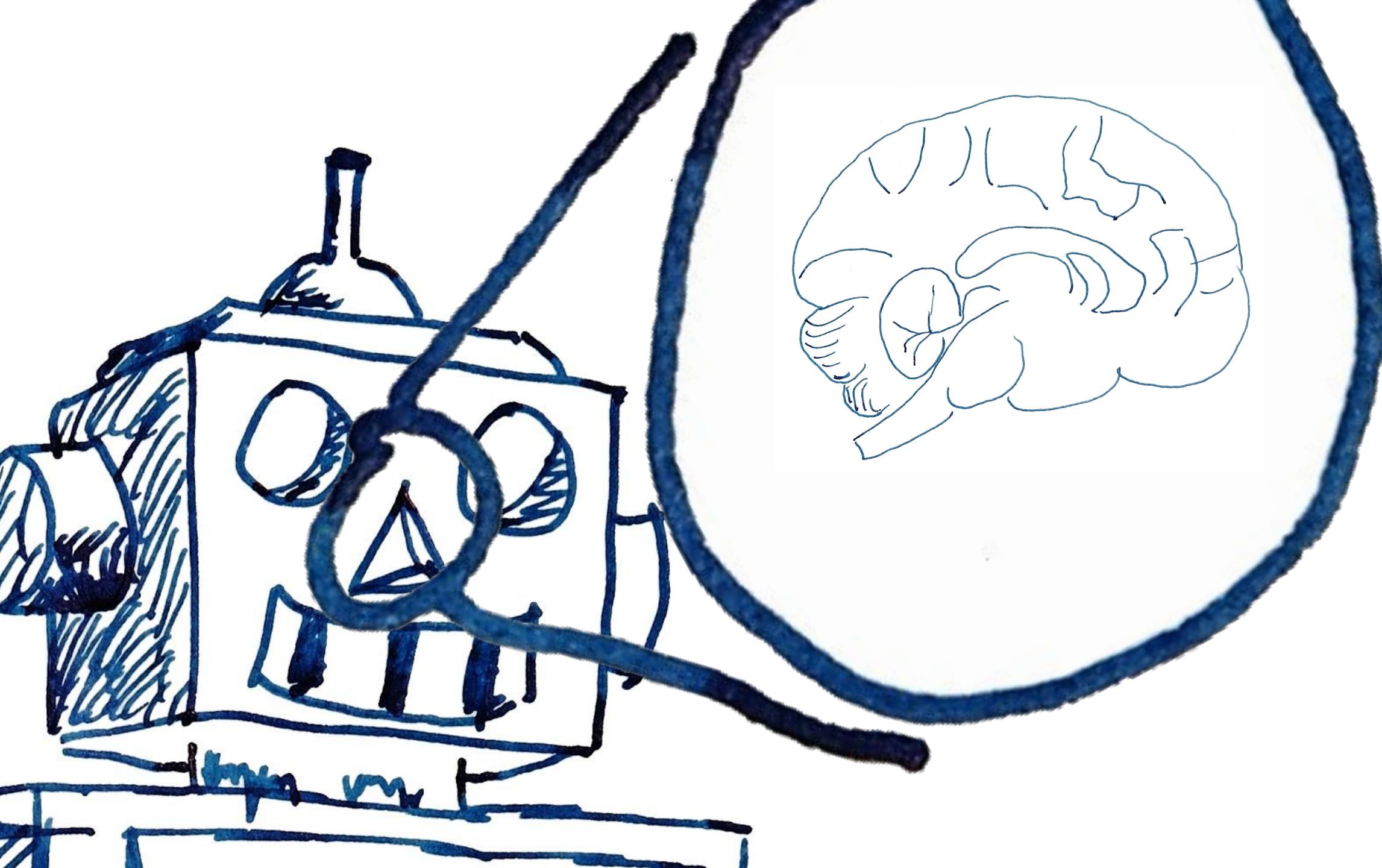
AN IDEAL KRM CONTROLLER SHOULD:

- * ONLY DO ONE THING
- * HAVE AN INPUT SOURCE
- * HAVE A PLACE TO WRITE STATUS
- * HAVE AN OUTPUT LOCATION
- * ANTICIPATE ITS OWN EFFECTS ON THE REST OF THE SYSTEM
- * BREAK THINGS EXACTLY A LITTLE BIT ON FAILURE

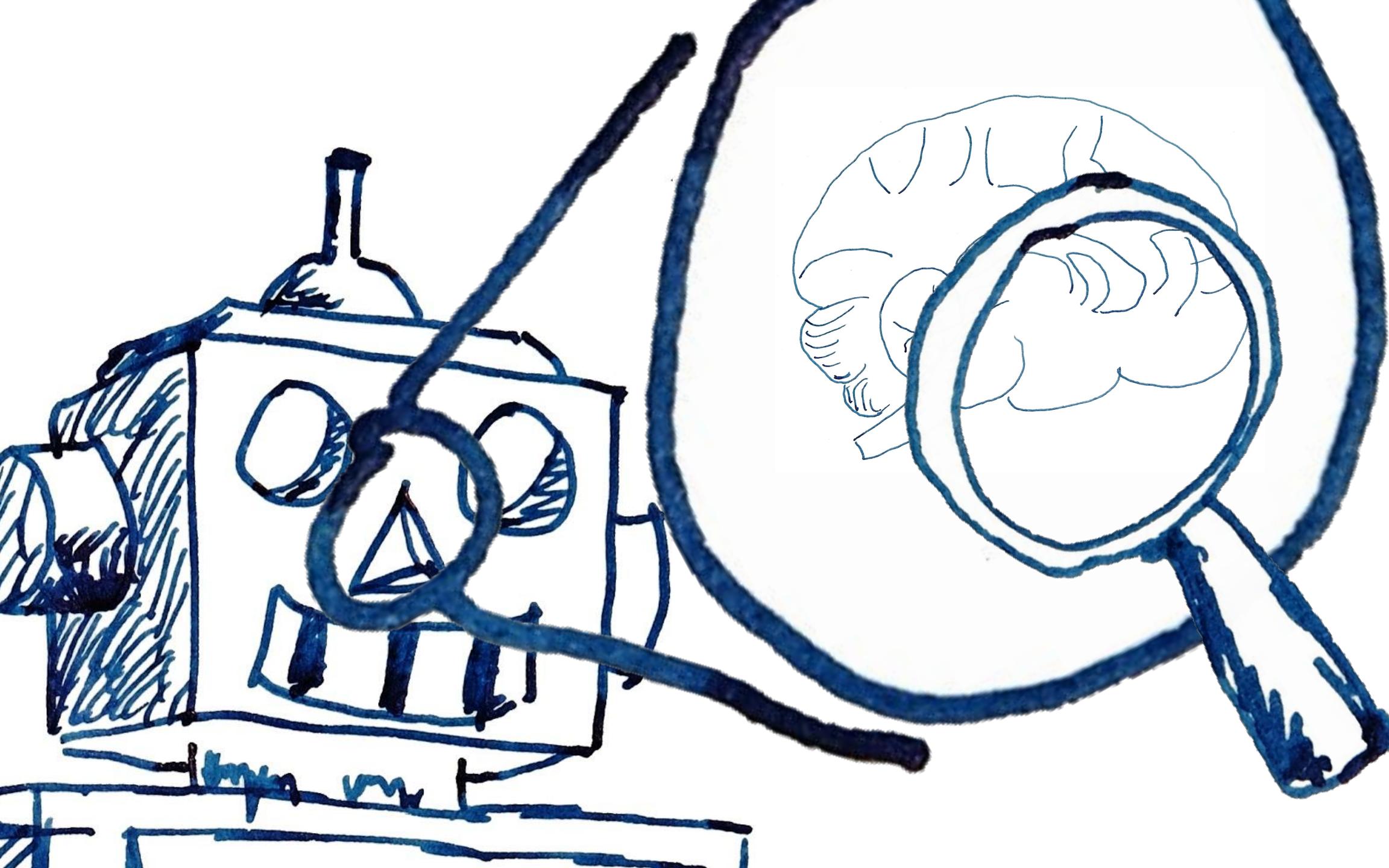


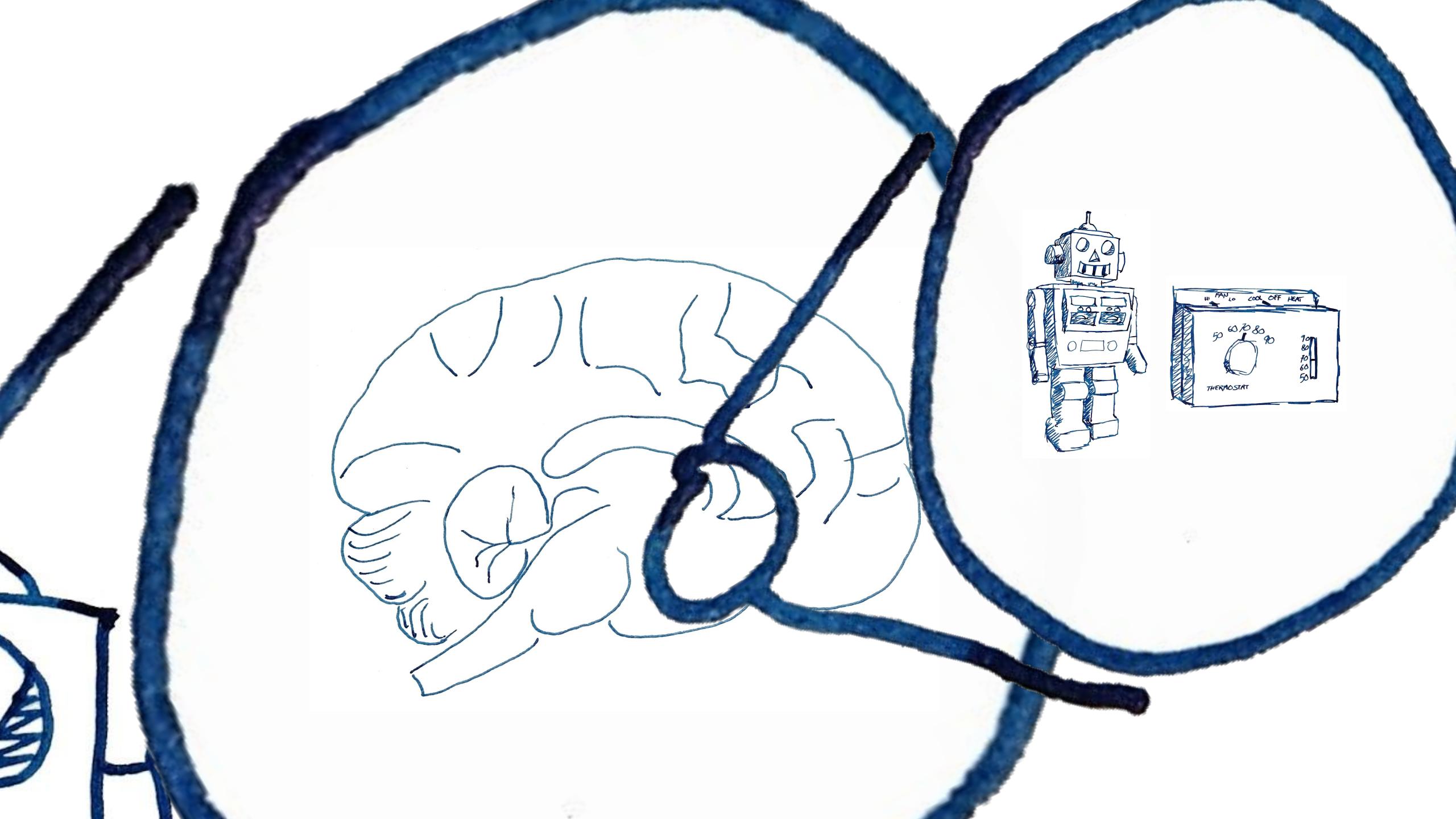


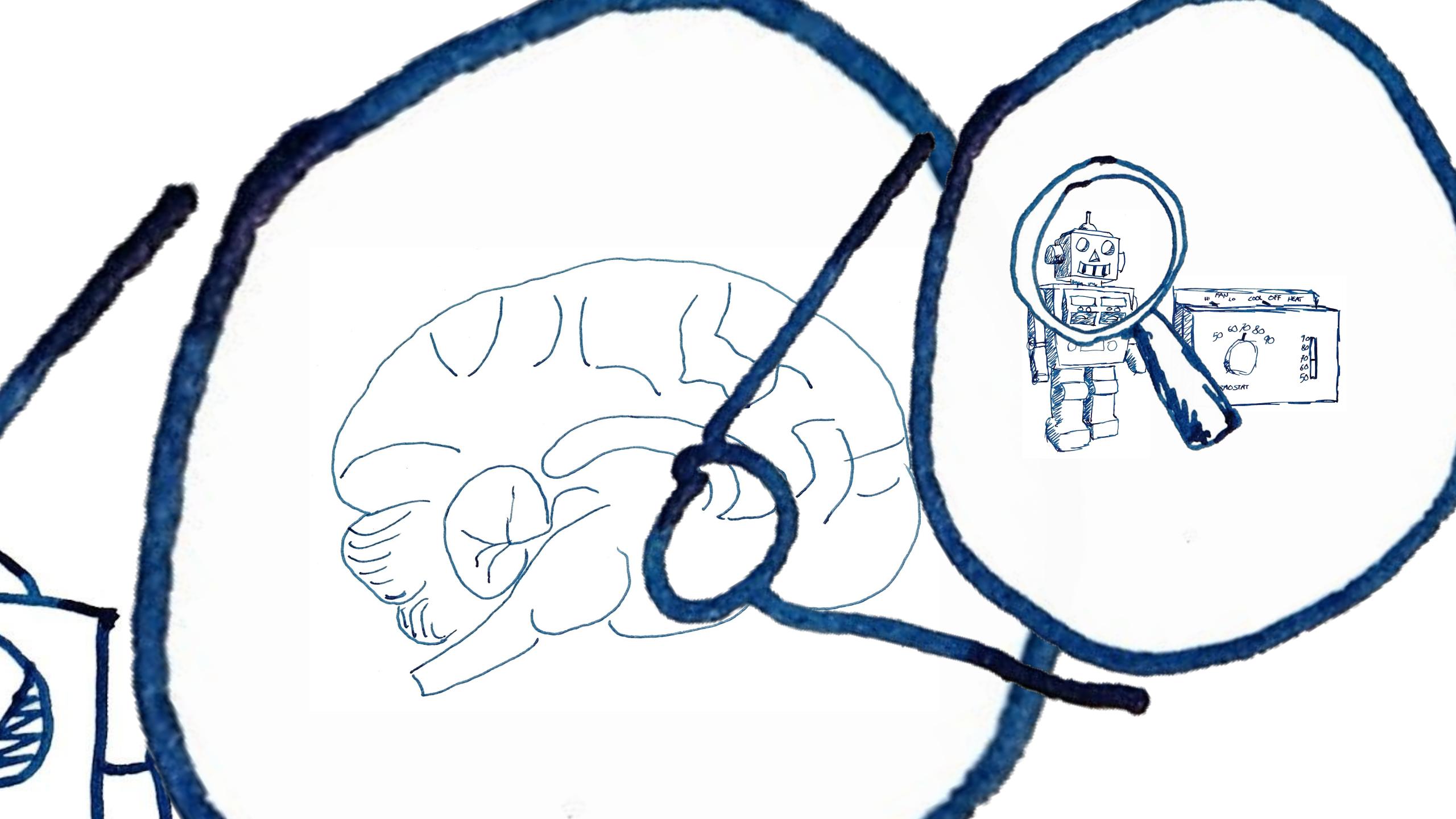


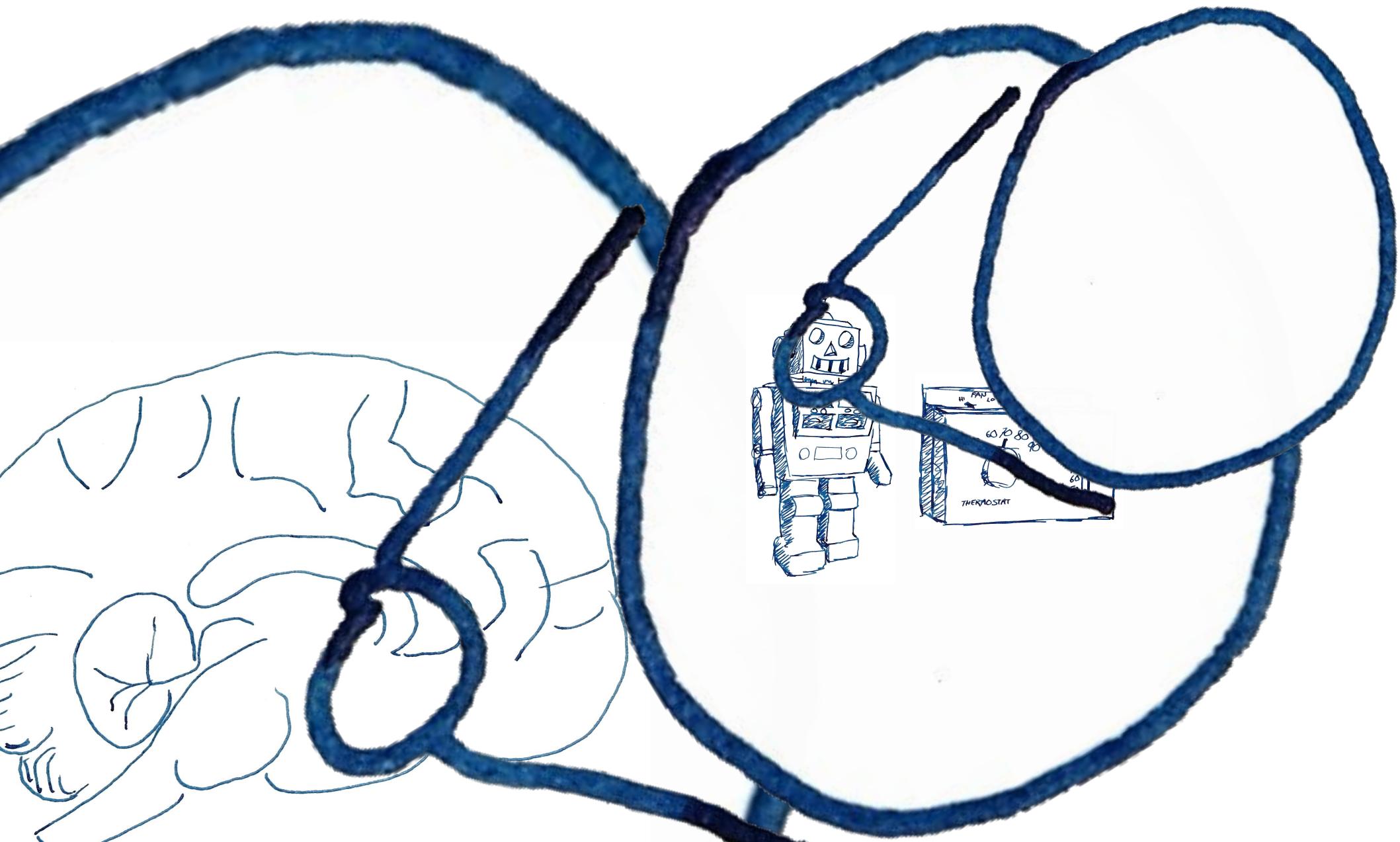


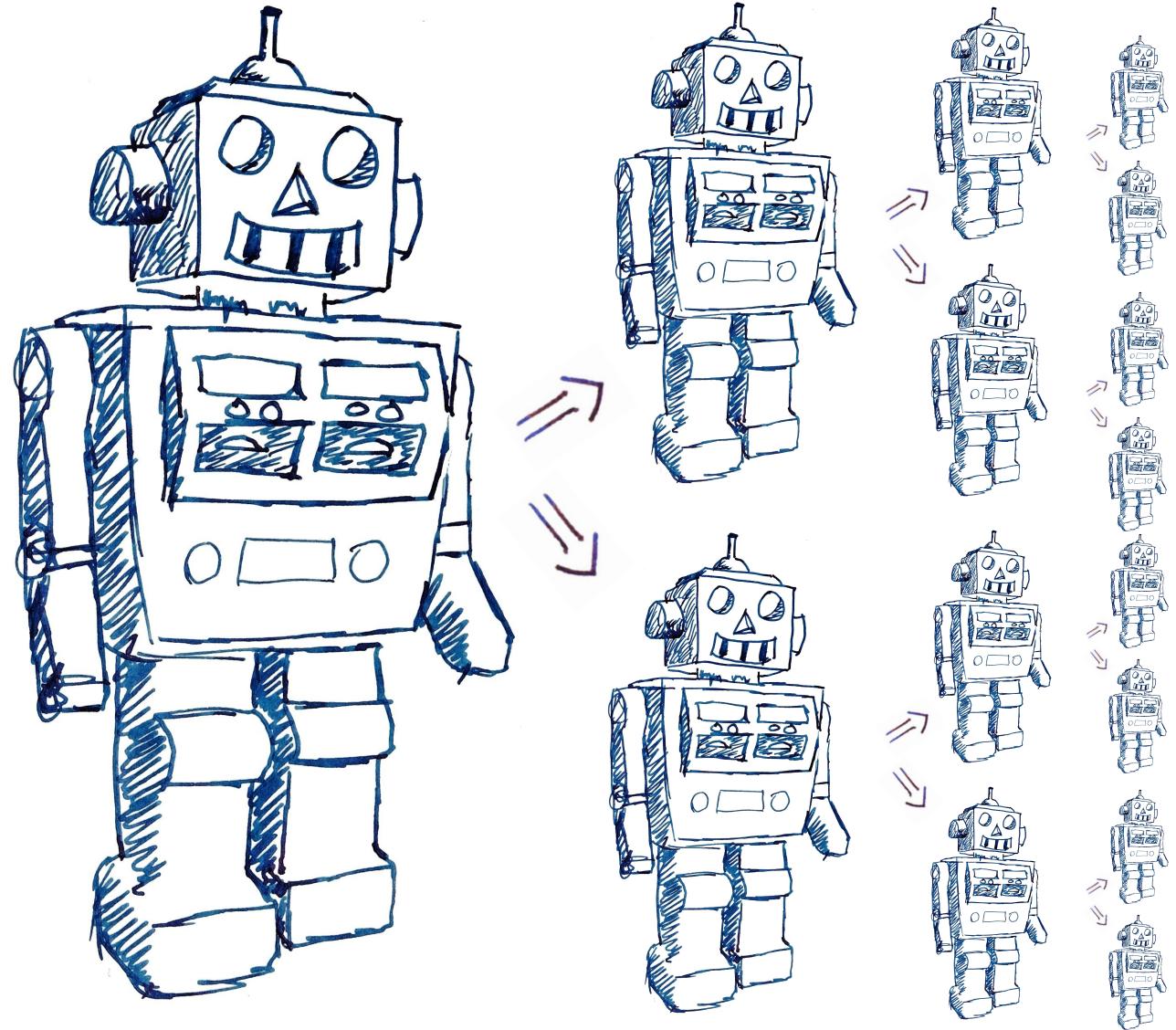
Hyper m











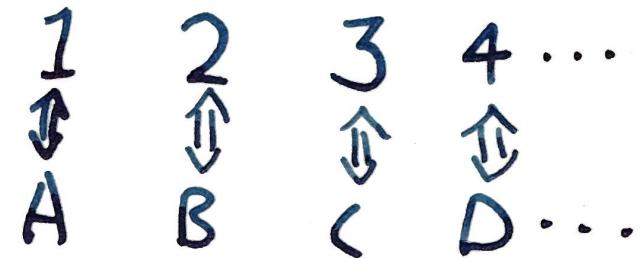
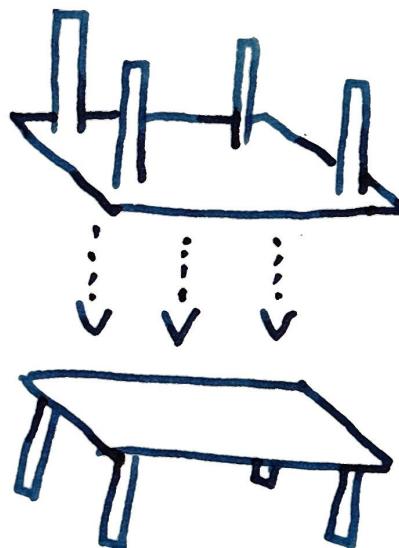
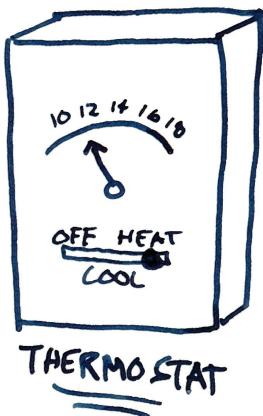
AN IDEAL KRM CONTROLLER SHOULD:

- * ONLY DO ONE THING
- * HAVE AN INPUT SOURCE
- * HAVE A PLACE TO WRITE STATUS
- * HAVE AN OUTPUT LOCATION
- * ANTICIPATE ITS OWN EFFECTS ON THE REST OF THE SYSTEM
- * BREAK THINGS EXACTLY A LITTLE BIT ON FAILURE

CONTROL THEORY
~~THEORY~~
PRACTICE!

CONTROLLER CATEGORIES

- * THE "CLASSIC" CONTROLLERS
- * STANDING QUERY / "TABLE JOIN"
- * IN- OR BIJECTION ENFORCERS

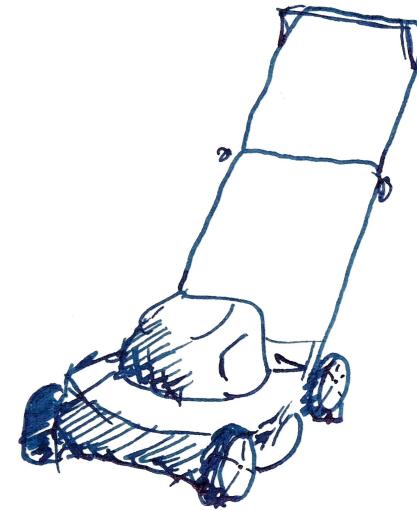
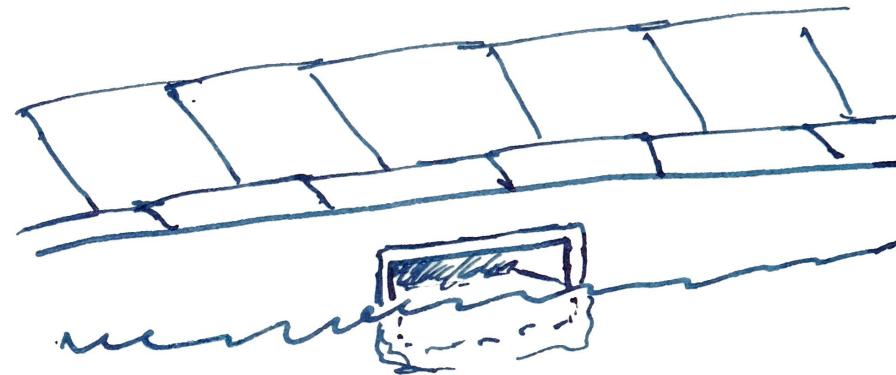
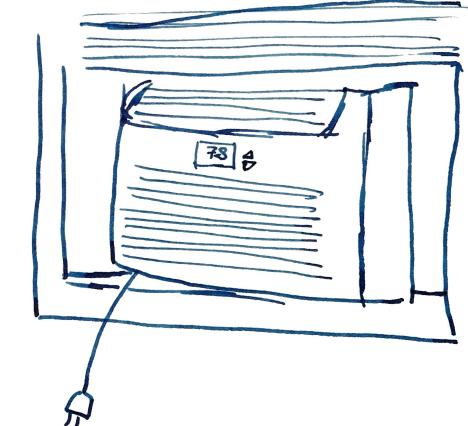


DISCLAIMERS!

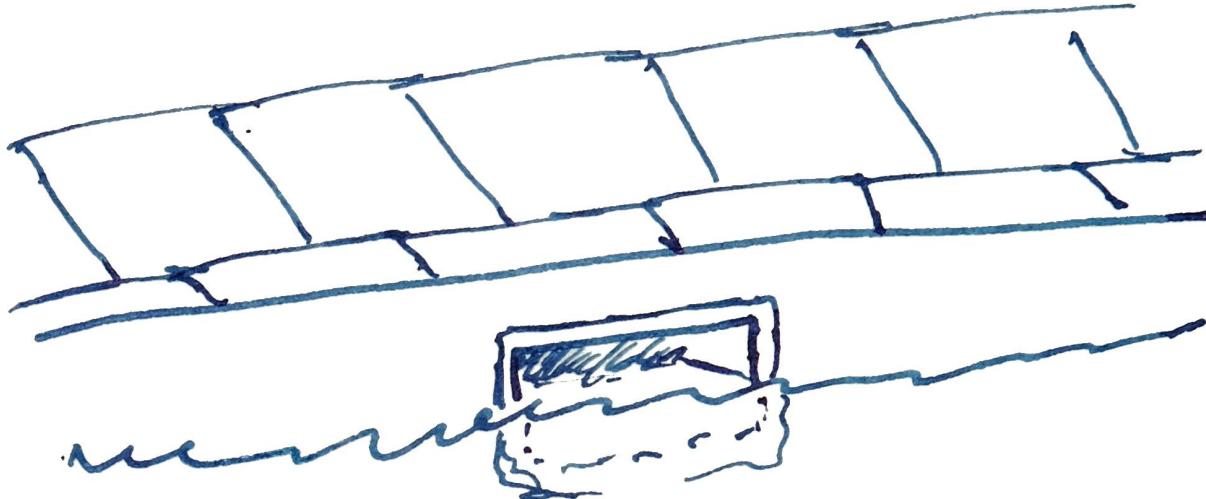
- MANY CONTROLLERS WRITE EVENTS: NOT SHOWN!
- NOT ALL "STATUS" PATHS SHOWN!
- WE'LL GO FAST ON SOME OF THESE!

CLASSIC CONTROLLER

FIRST UP:
GARBAGE COLLECTORS!

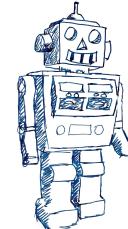


CLASSIC CONTROLLER



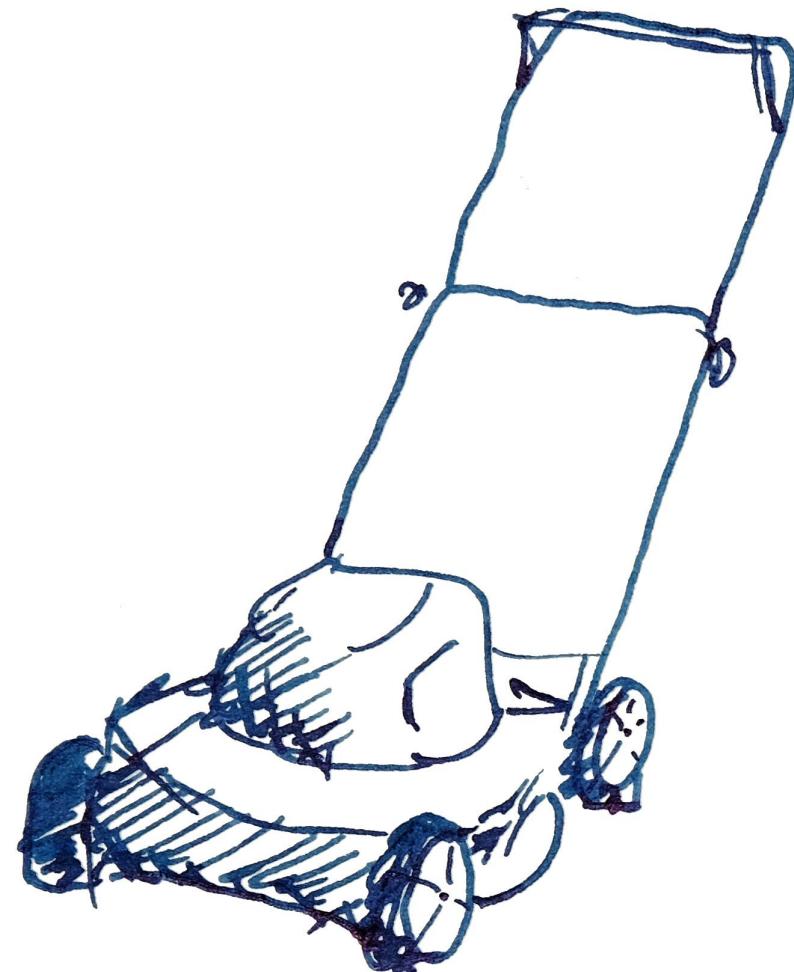
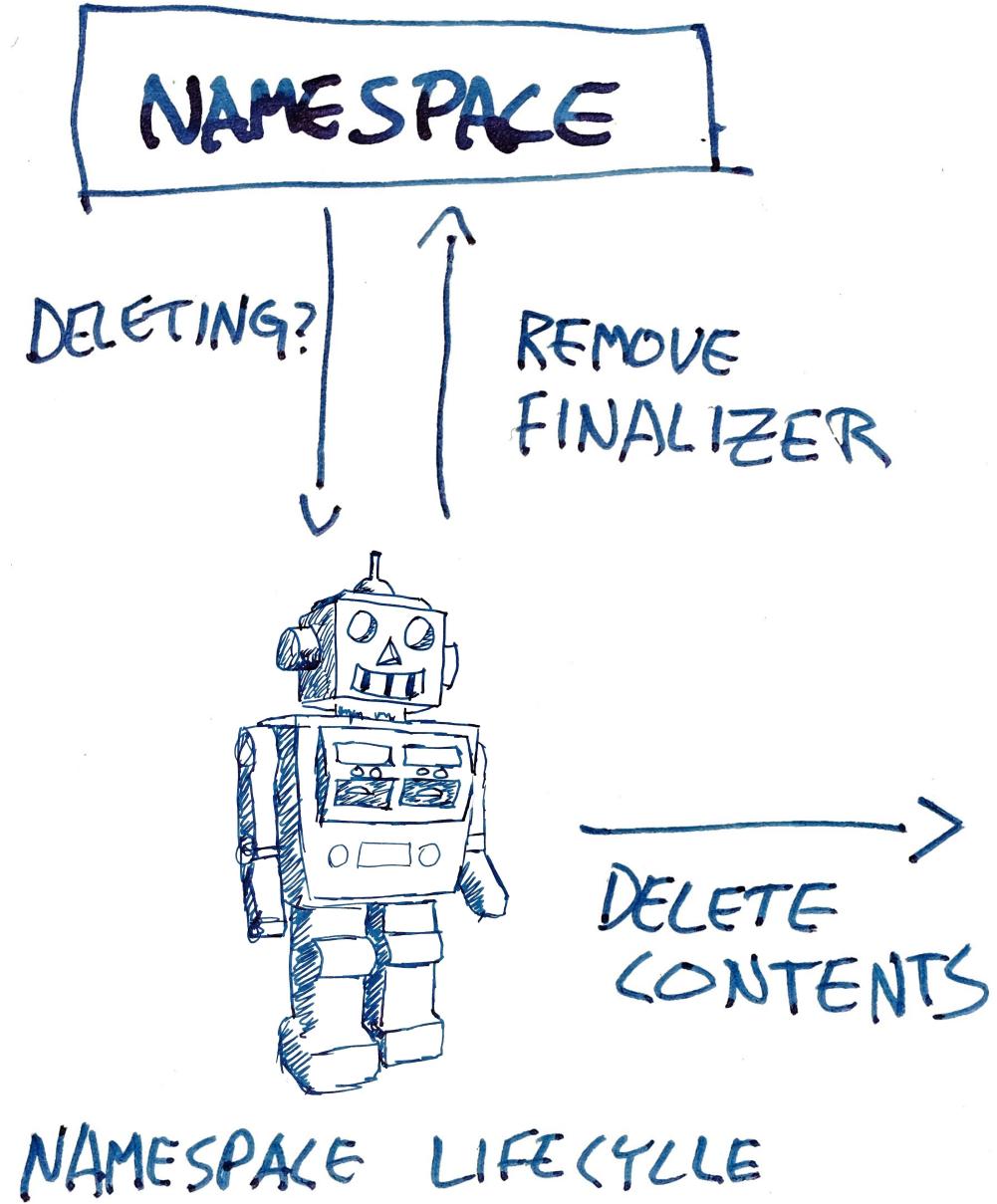
POD

TOO MANY
FINISHED?
DELETE



POD GC

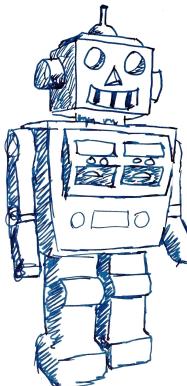
CLASSIC CONTROLLER



CLASSIC CONTROLLER

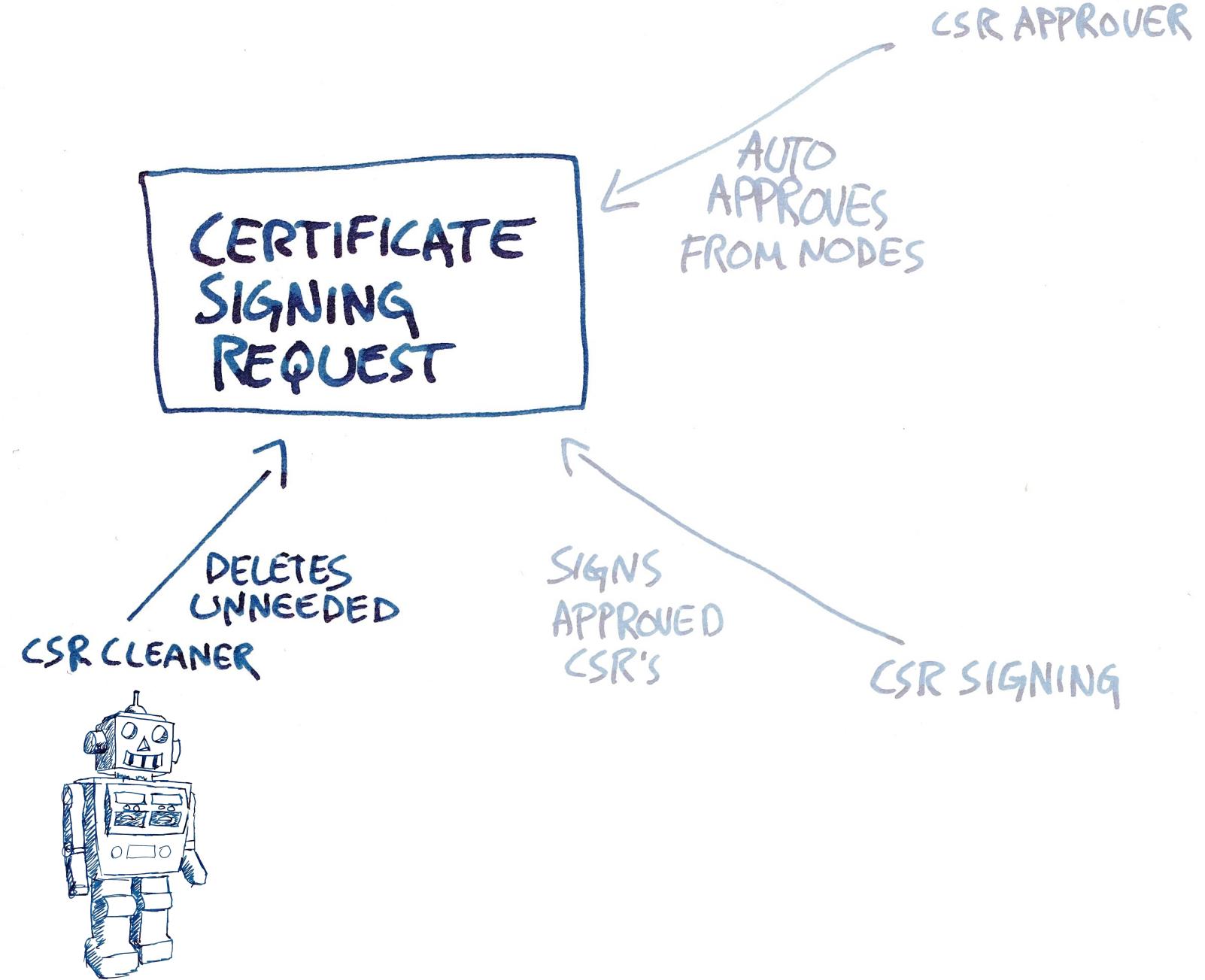


PARENTS
ALL DELETED?
DELETE!



GARBAGE COLLECTOR

CLASSIC CONTROLLER





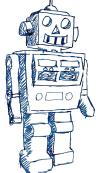
CLASSIC CONTROLLER

PERSISTENT VOLUME

PERSISTENT VOLUME CLAIM

WAIT FOR CLEANUP
REMOVE FINALIZER

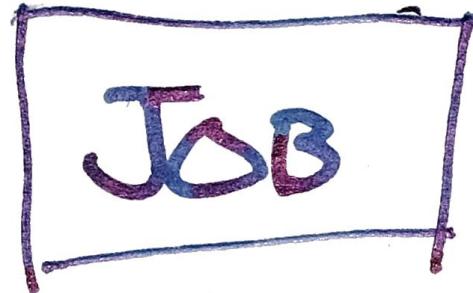
WAIT FOR CLEANUP
REMOVE FINALIZER



PV - PROTECTION



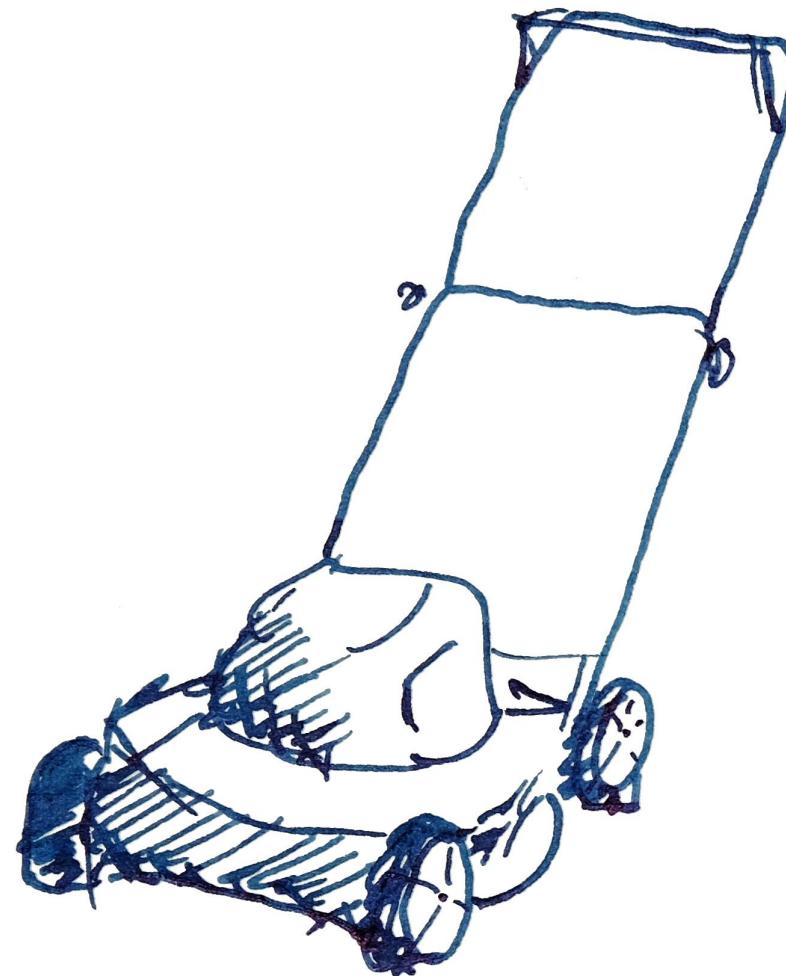
PVC - PROTECTION



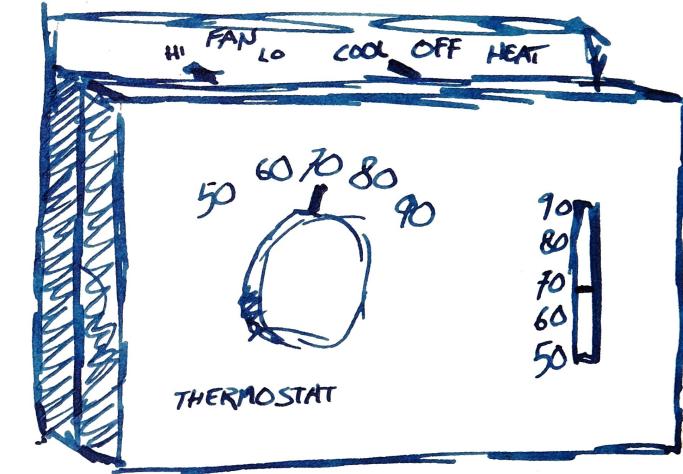
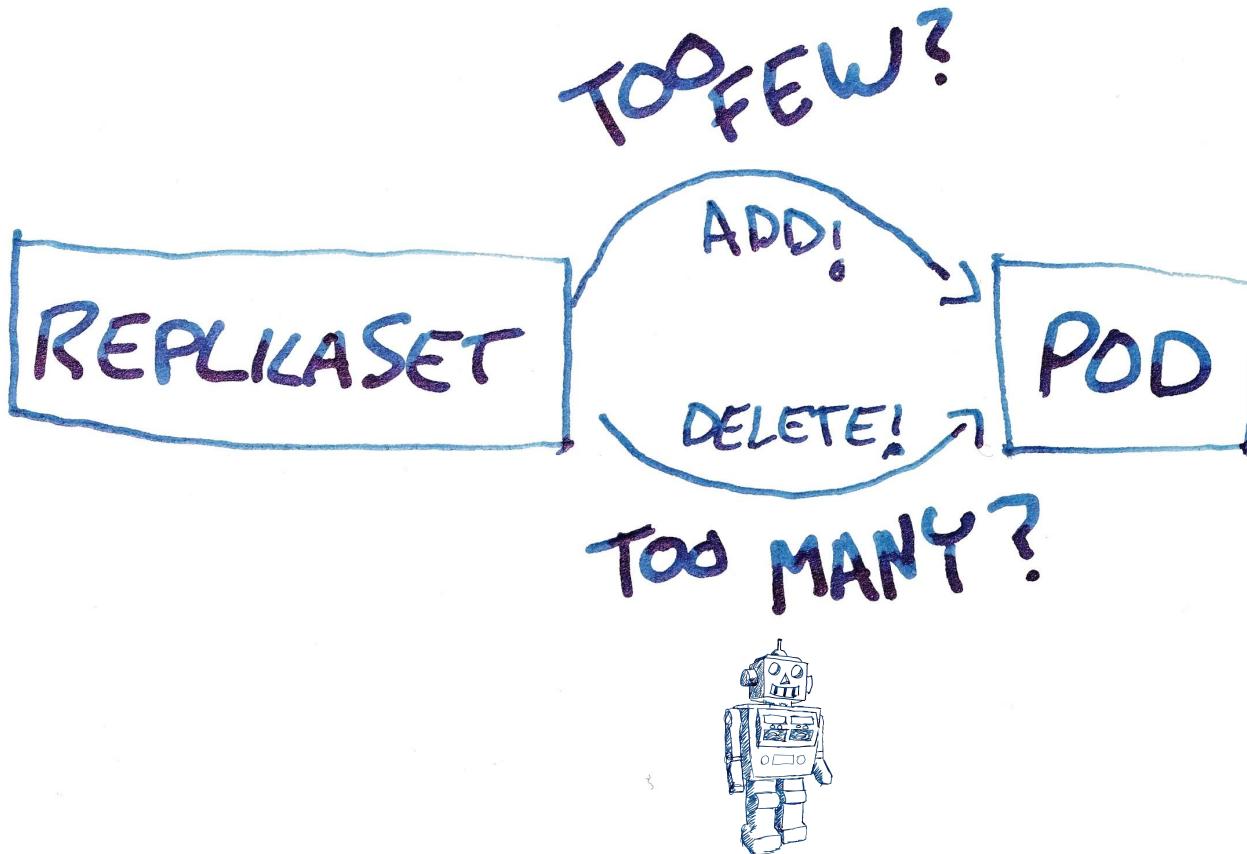
↑
TOO OLD?
DELETE

TTL-AFTER-FINISHED

CLASSIC CONTROLLER

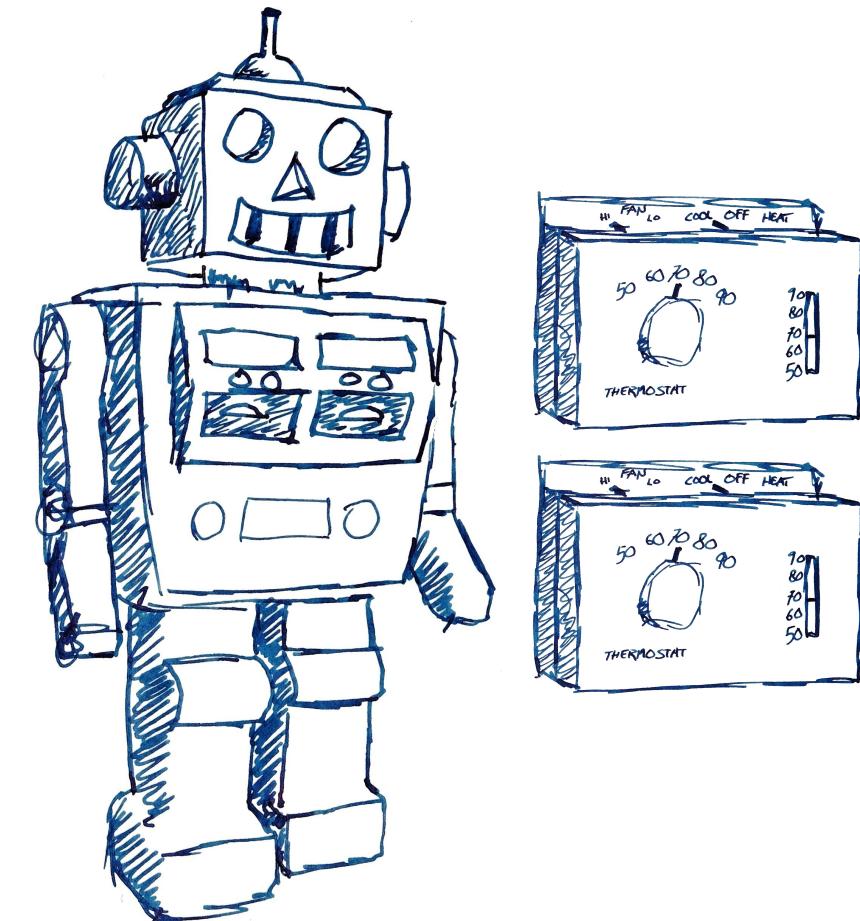
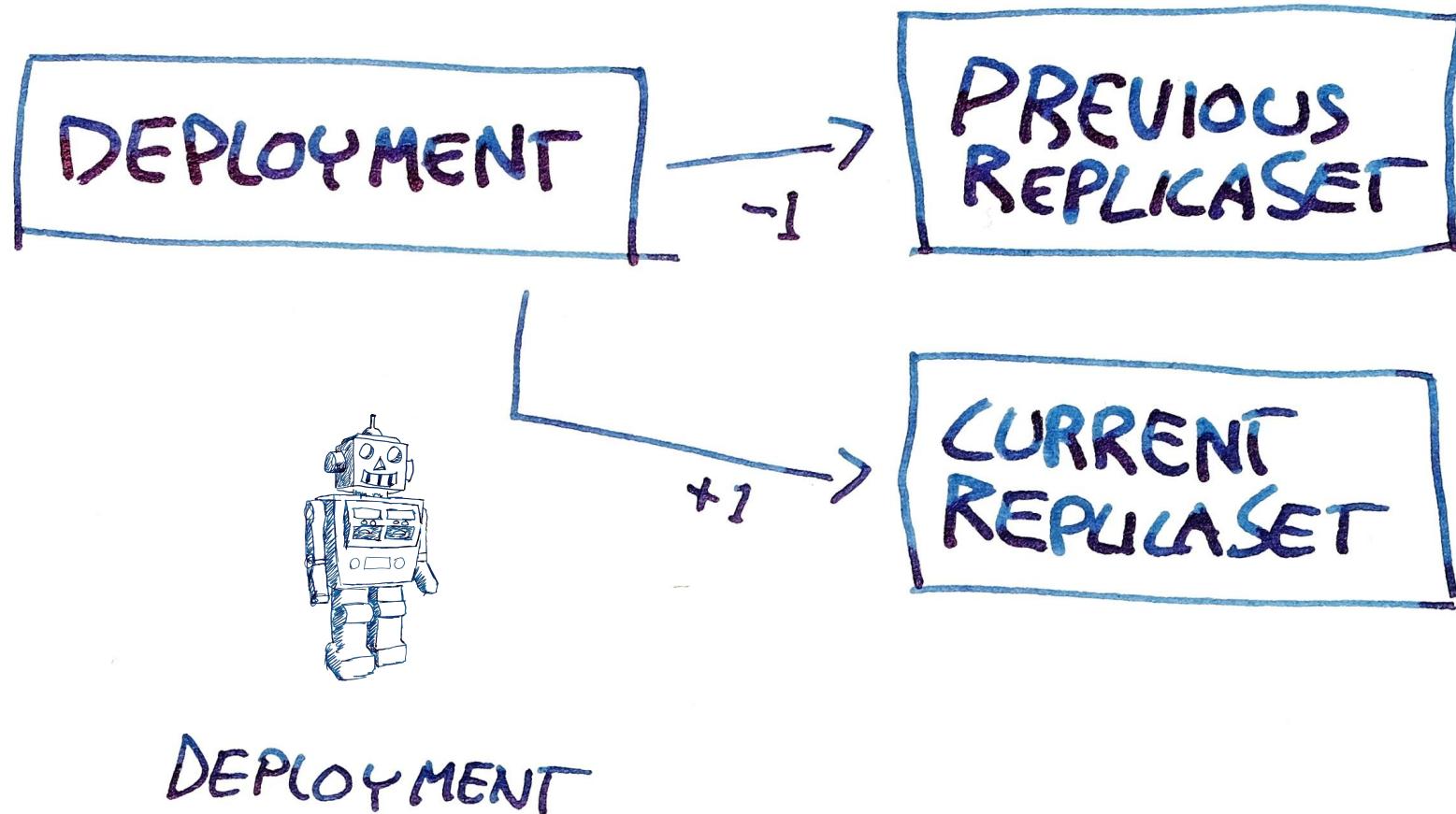


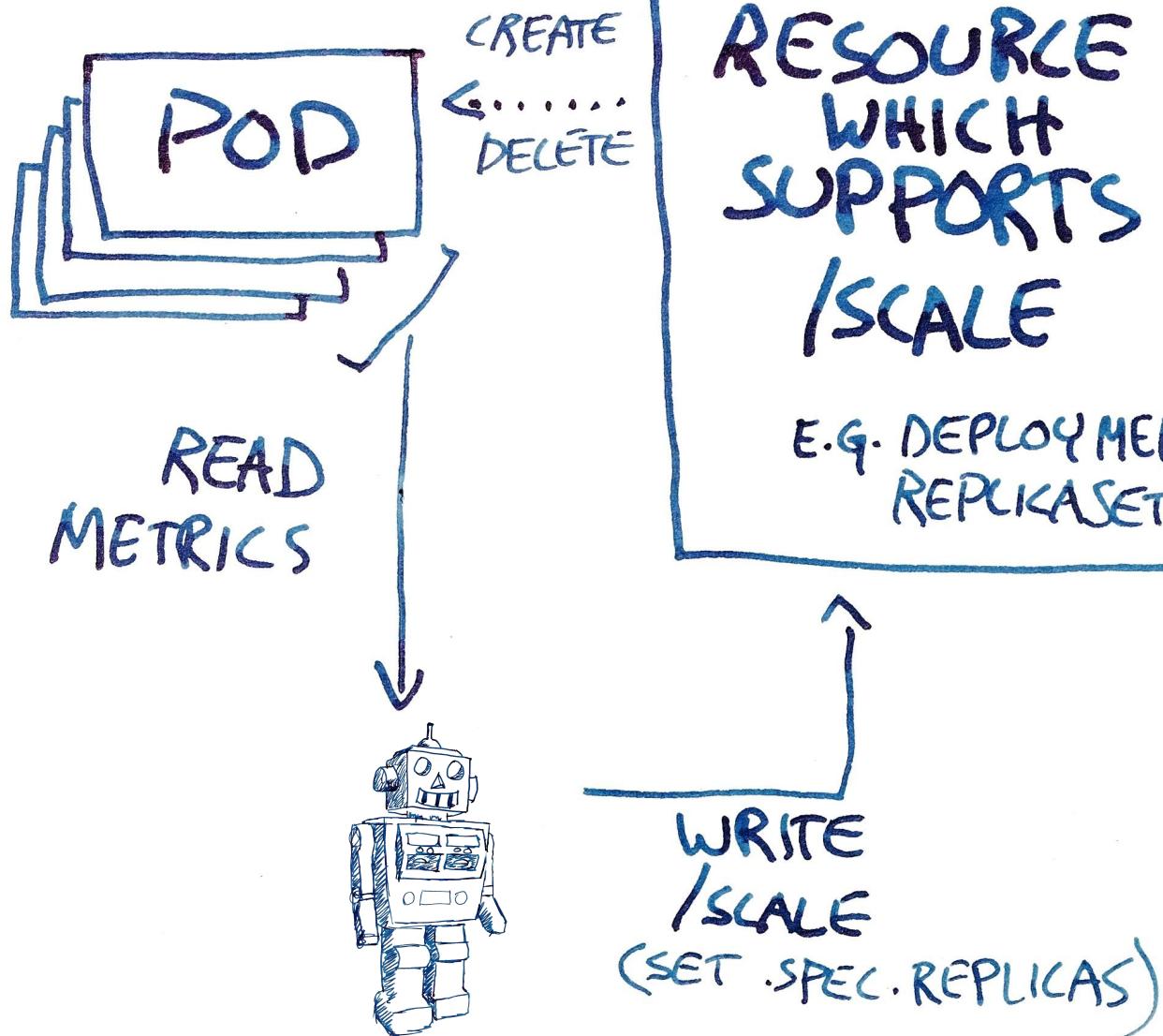
CLASSIC CONTROLLER



REPLICASET

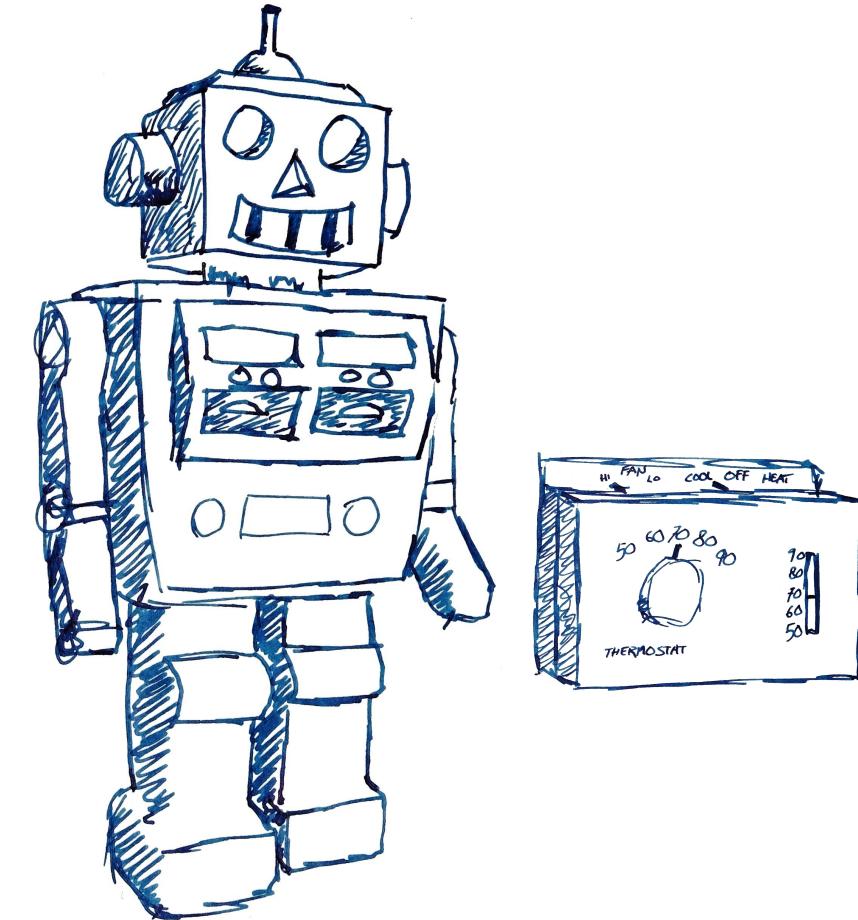
CLASSIC CONTROLLER





HORIZONTAL POD AUTOSCALER

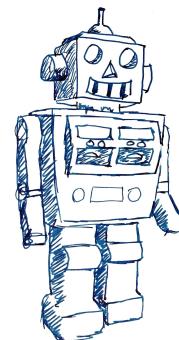
CLASSIC CONTROLLER



CLASSIC CONTROLLER

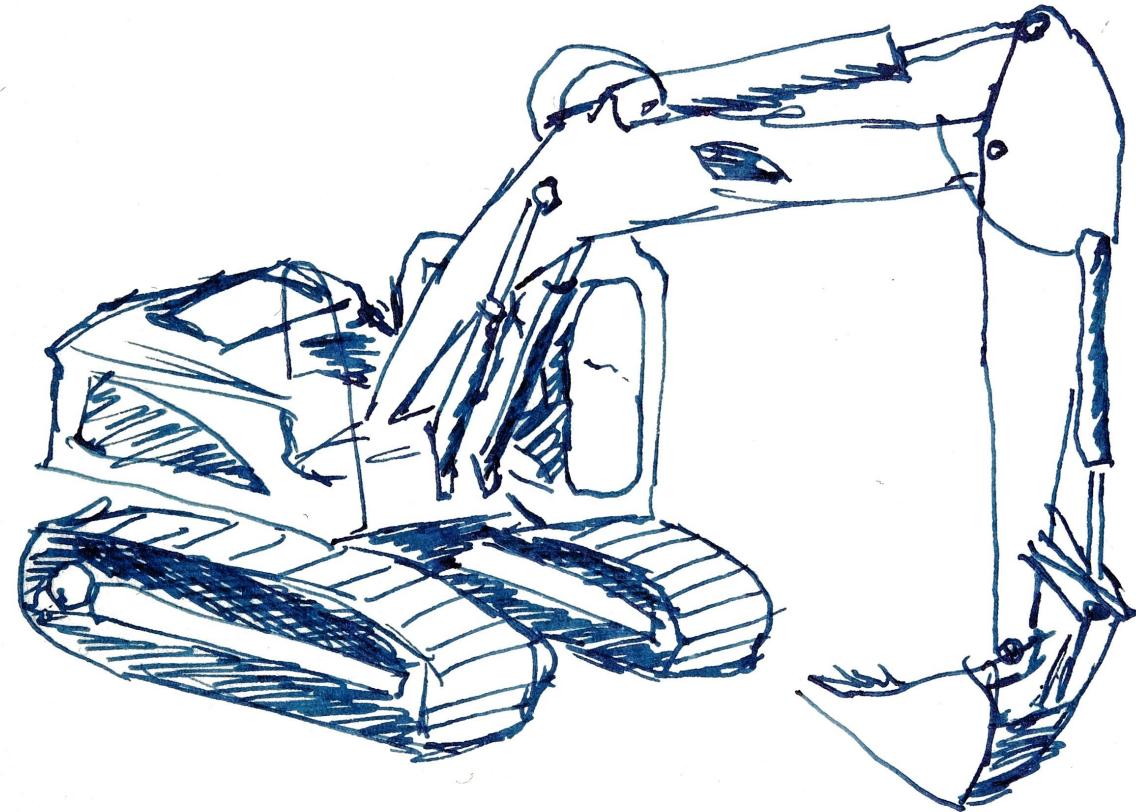


↓
TOO
MANY?



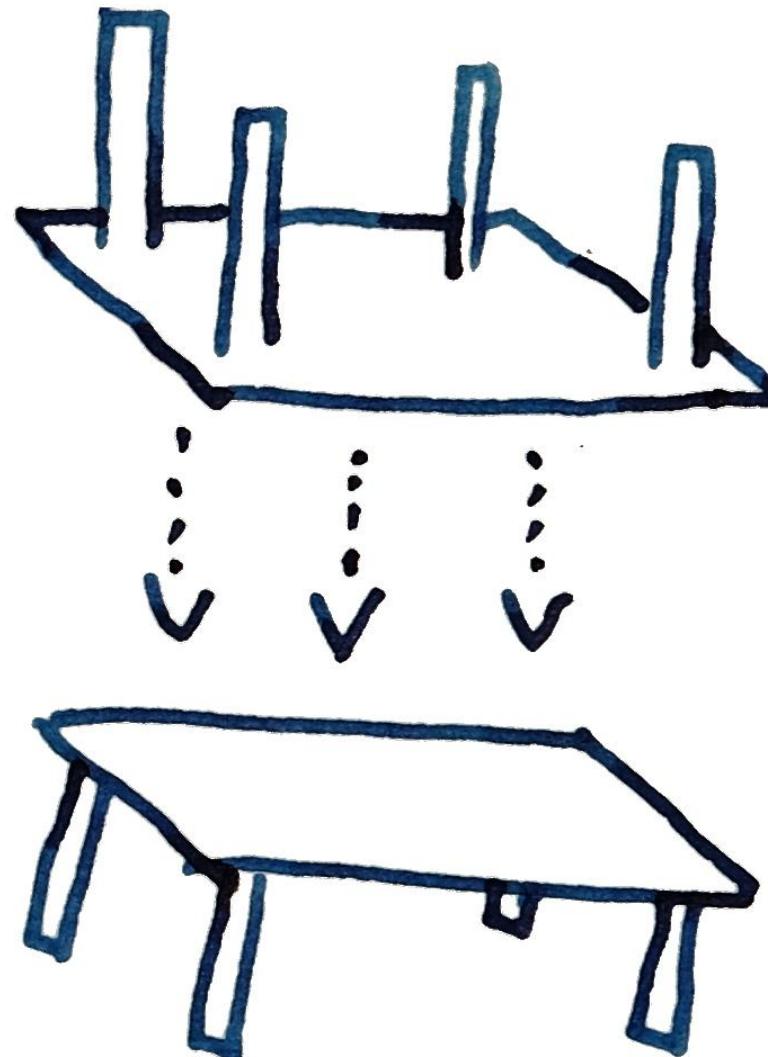
ADD NODES

CLOUD
PROVIDER

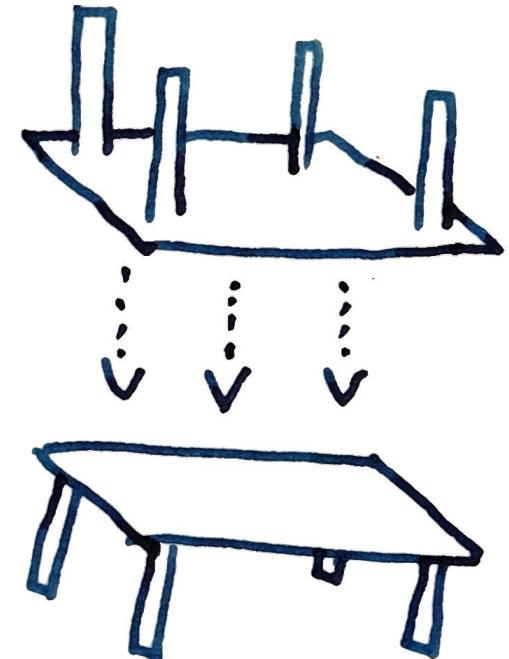
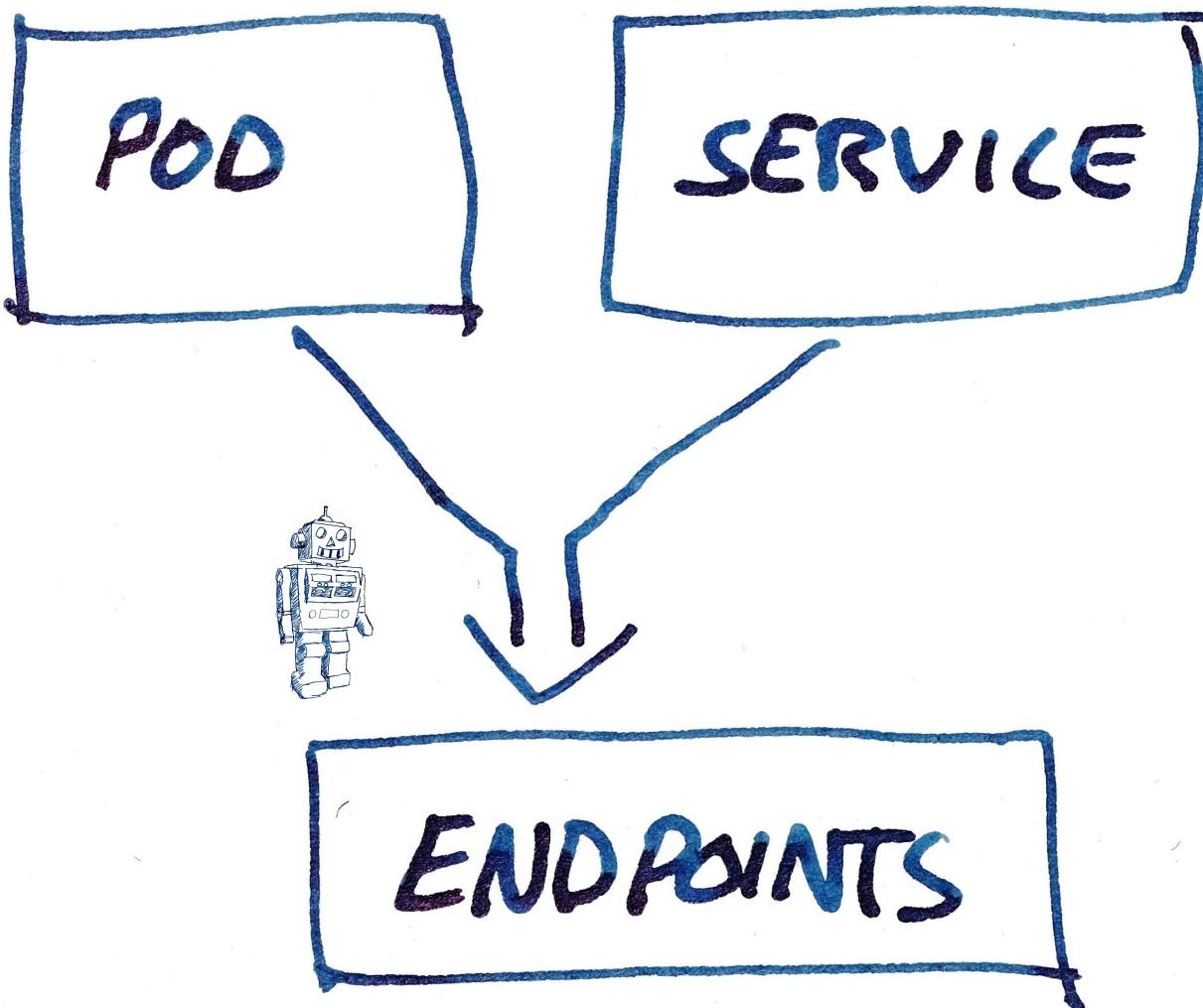


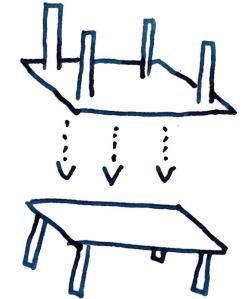
CLUSTERAUTOSCALER

STANDING QUERY TABLE JOIN

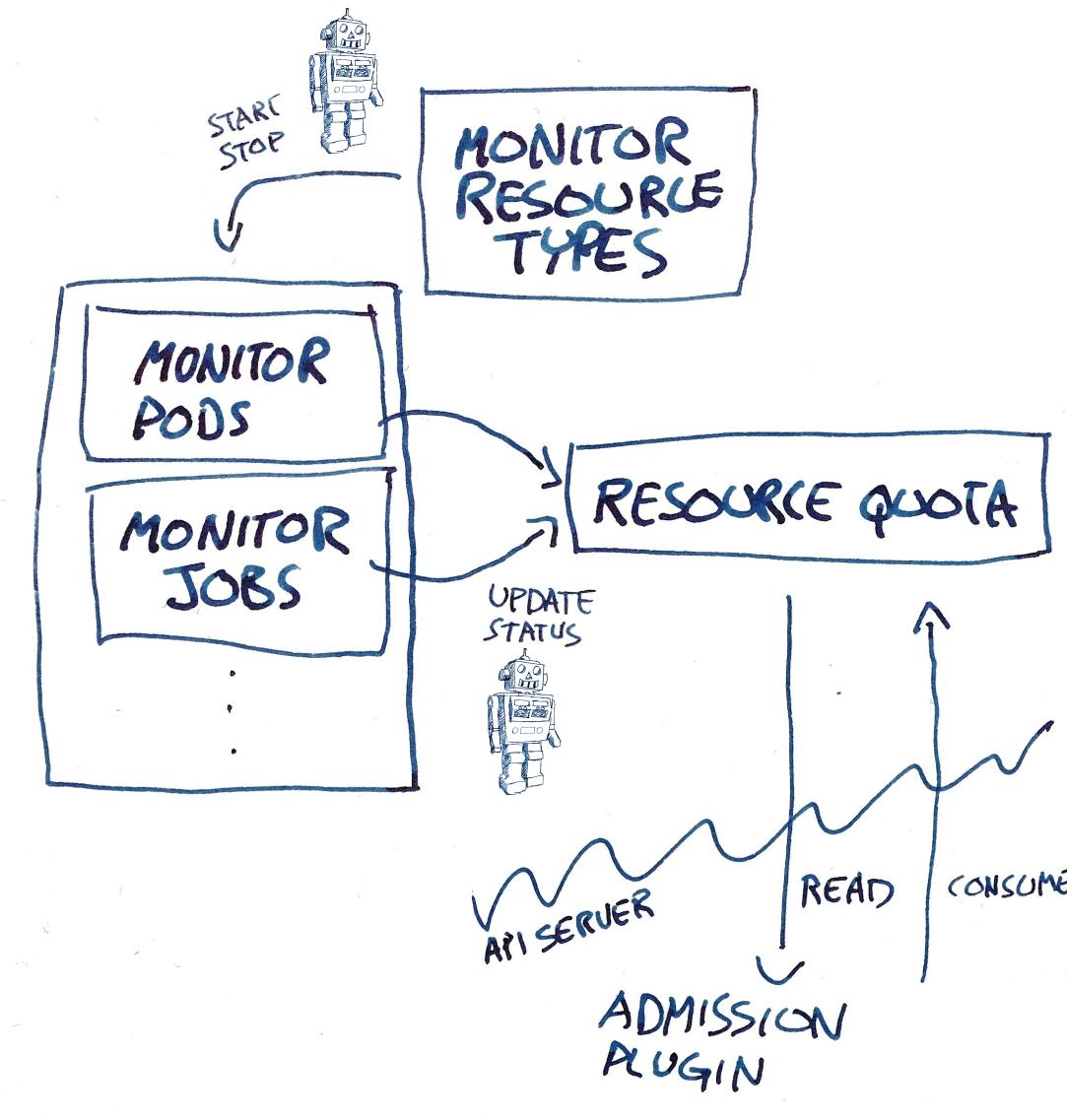
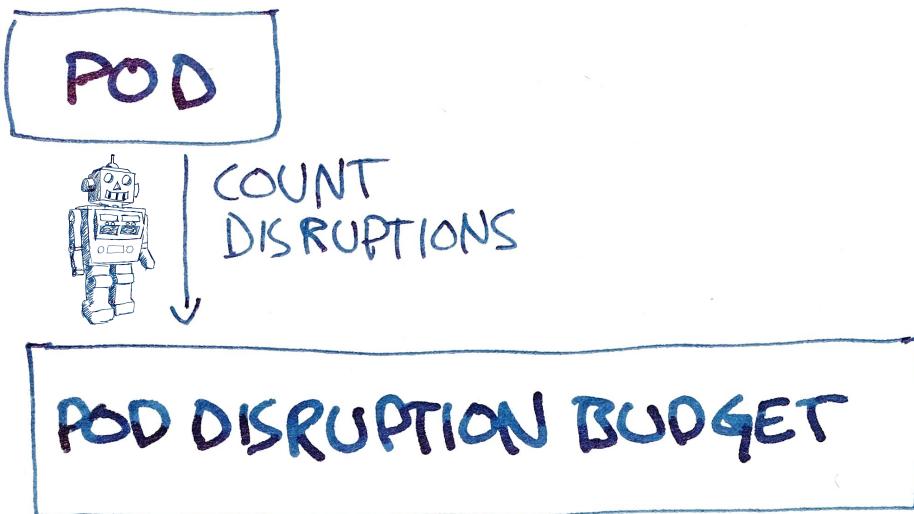


STANDING QUERY TABLE JOIN

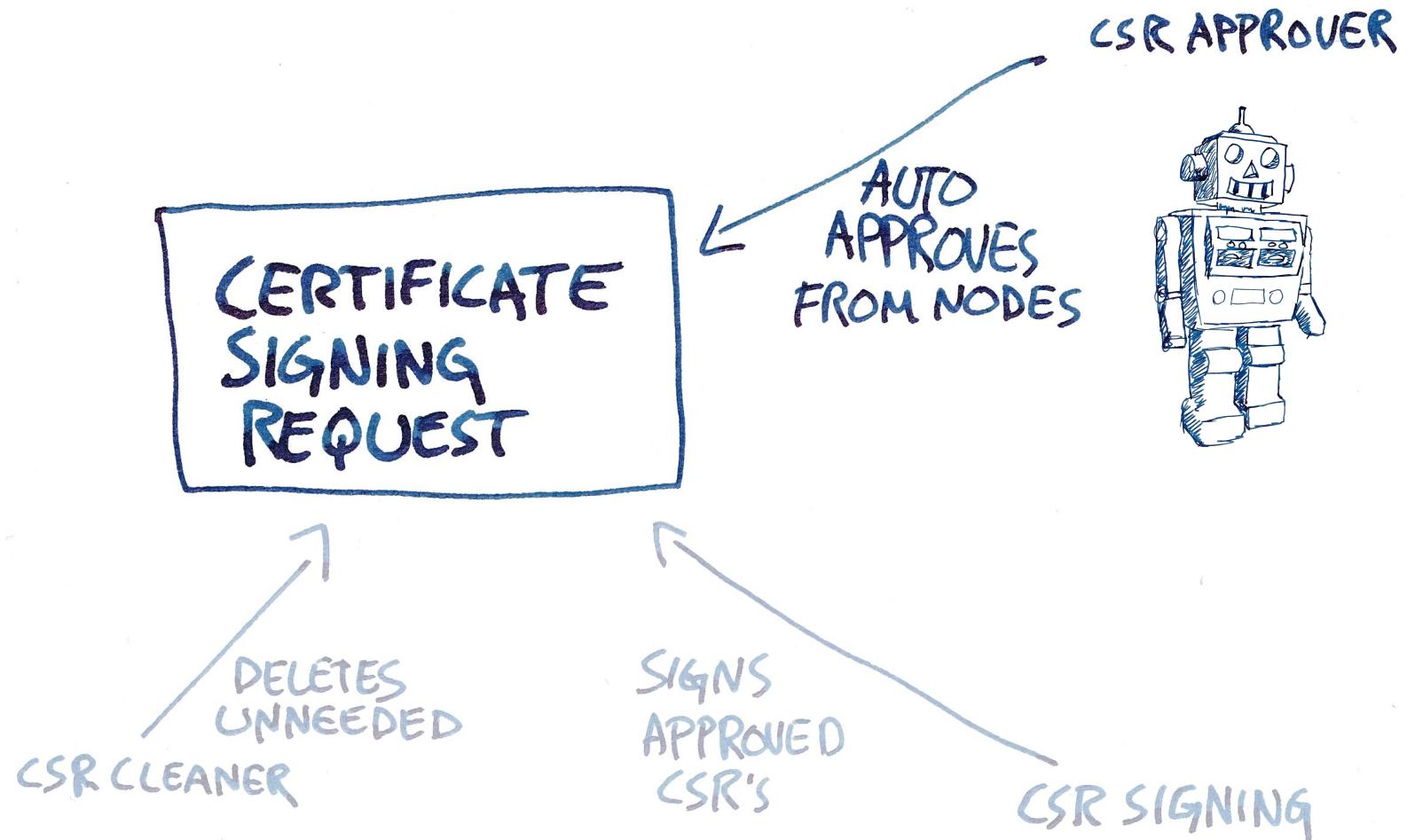
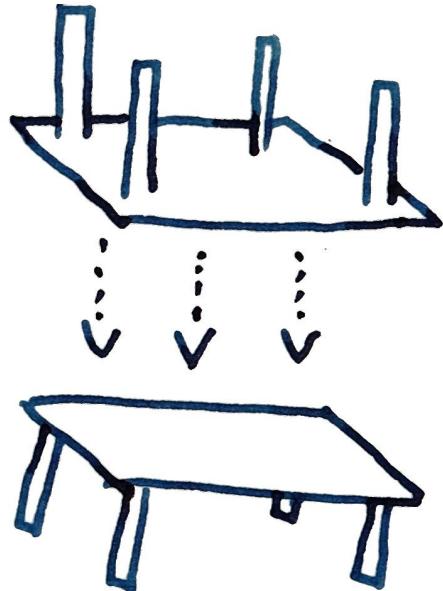




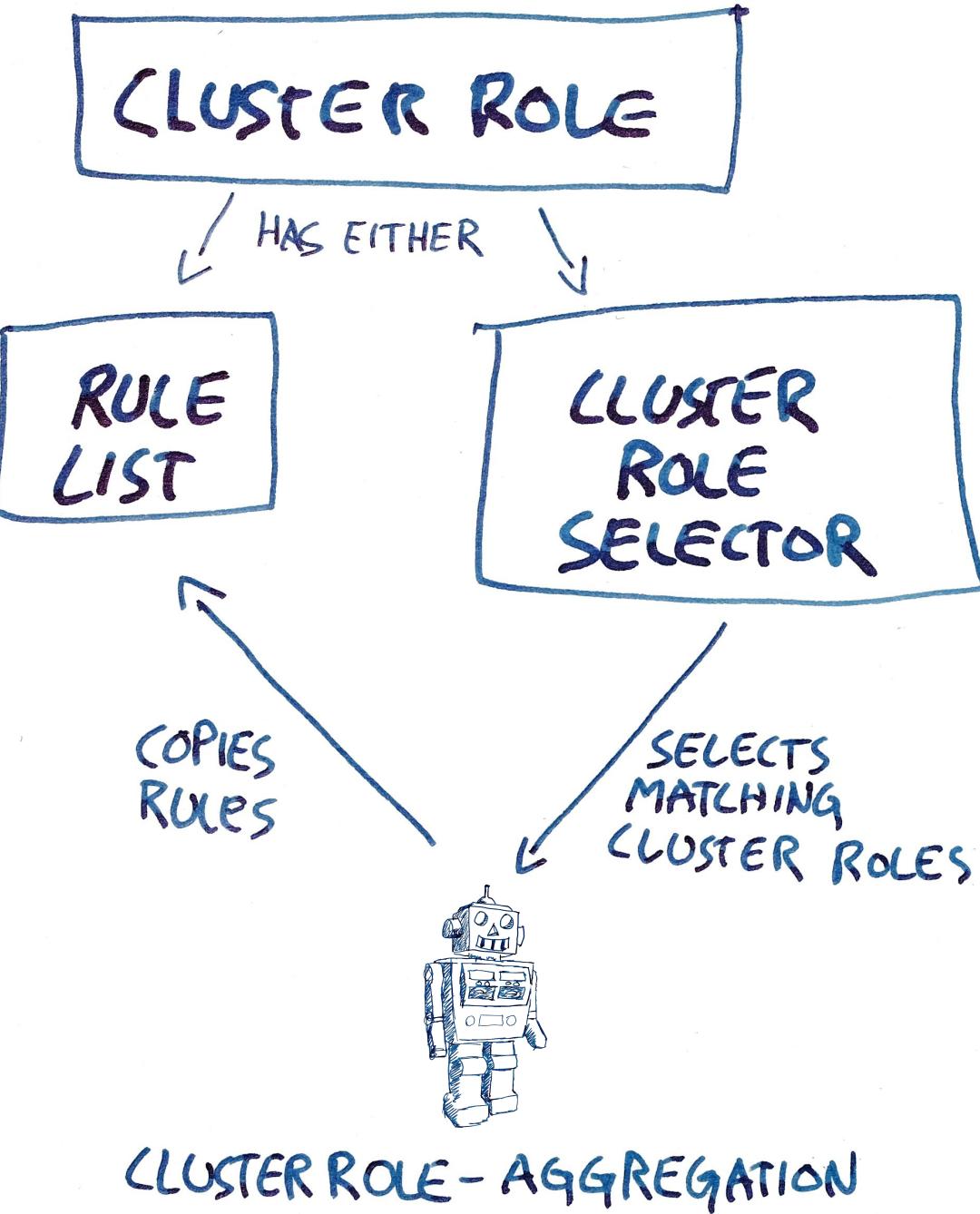
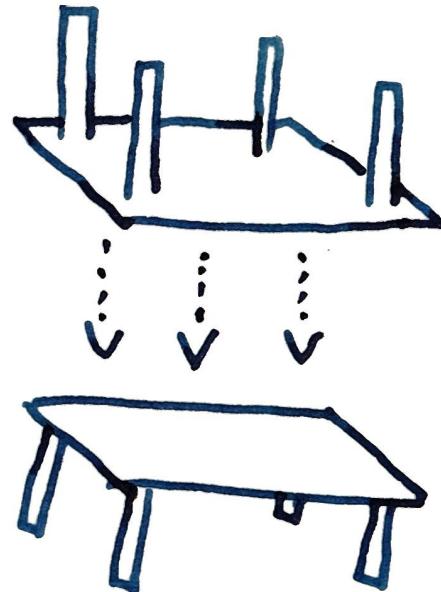
STANDING QUERY TABLE JOIN



STANDING QUERY TABLE JOIN

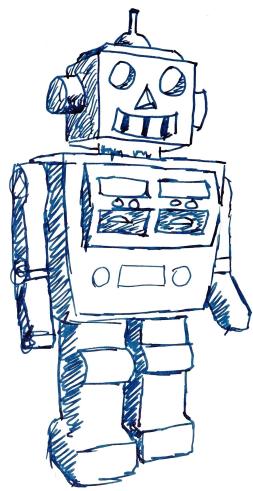


STANDING QUERY TABLE JOIN



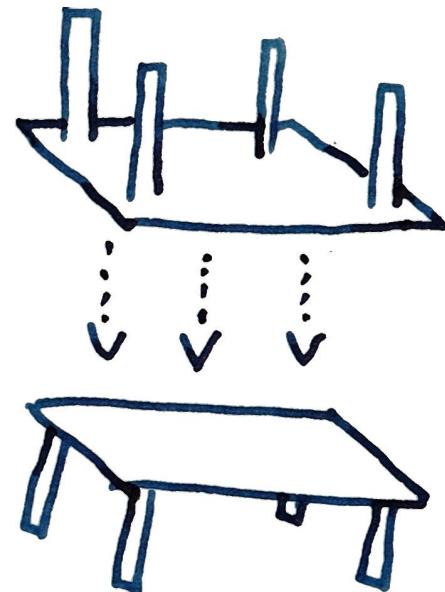
STANDING QUERY TABLE JOIN

NODE ← WRITE
CONDITIONS

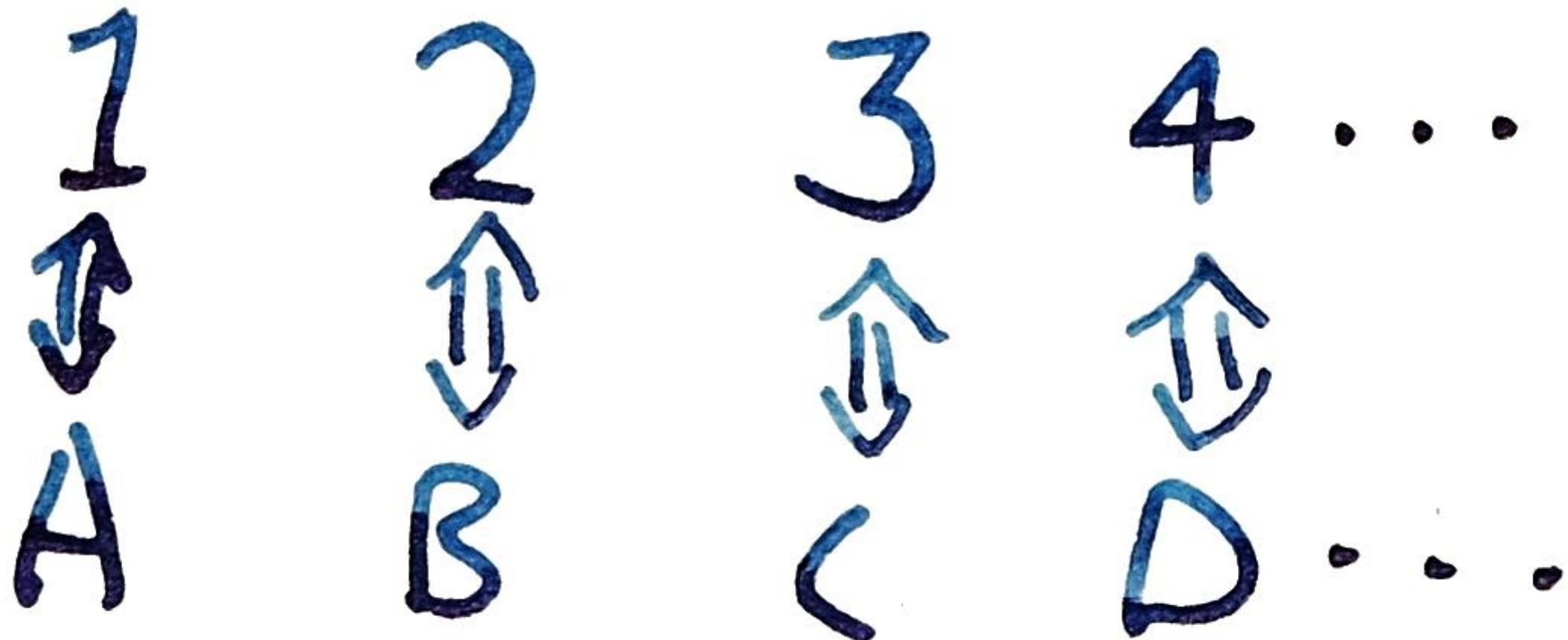


← OBSERVE O.S.

NODE PROBLEM DETECTOR

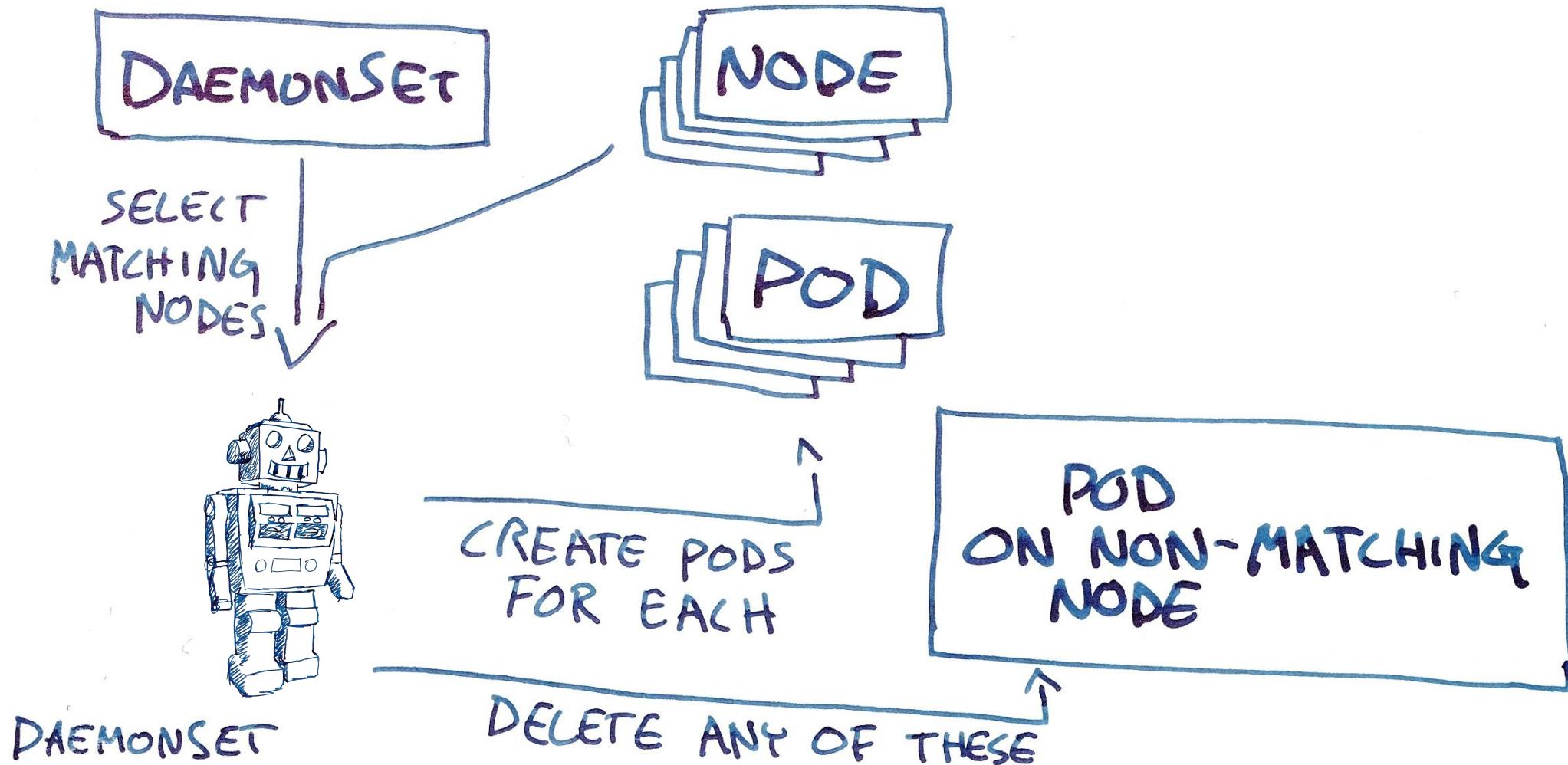


INJECTION ENFORCER



INJECTION ENFORCER

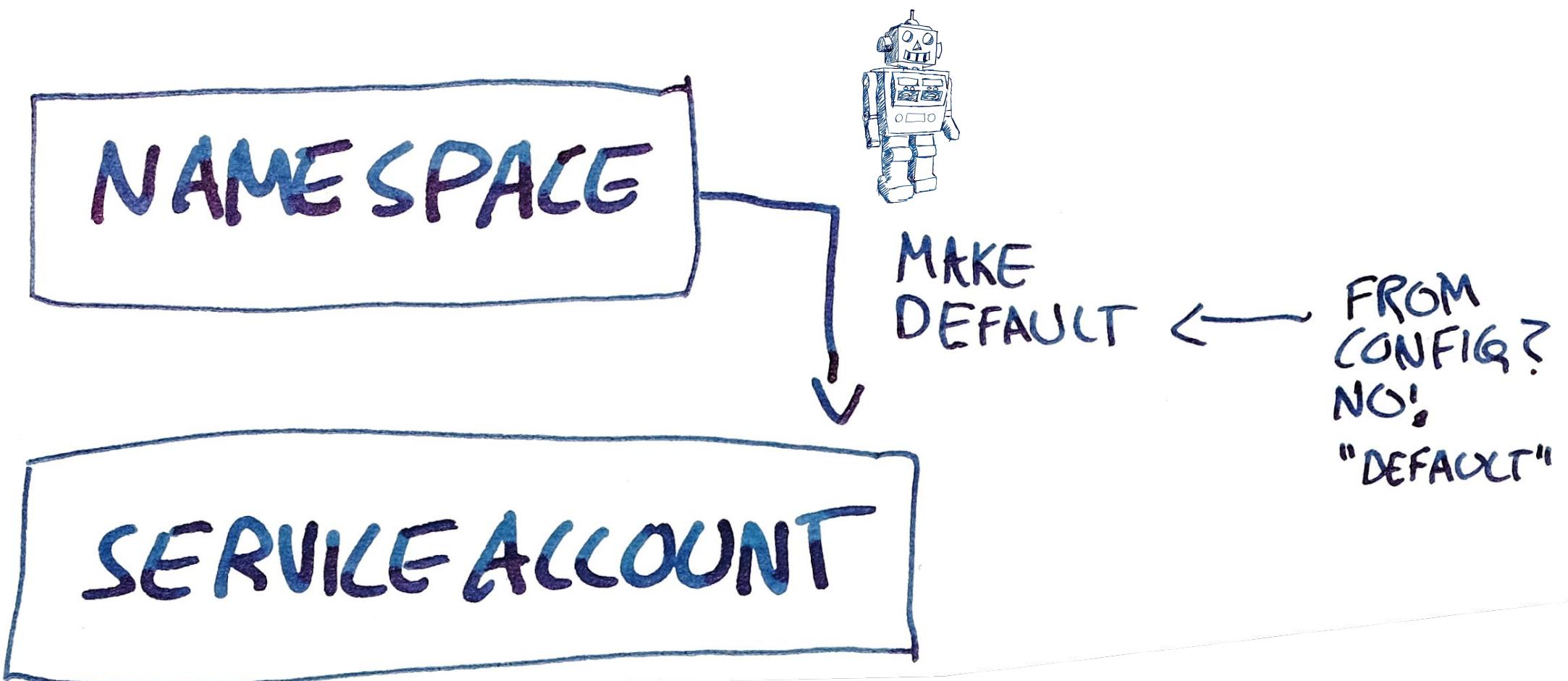
1
2
3
4...
A
B
C
D...



INJECTION ENFORCER

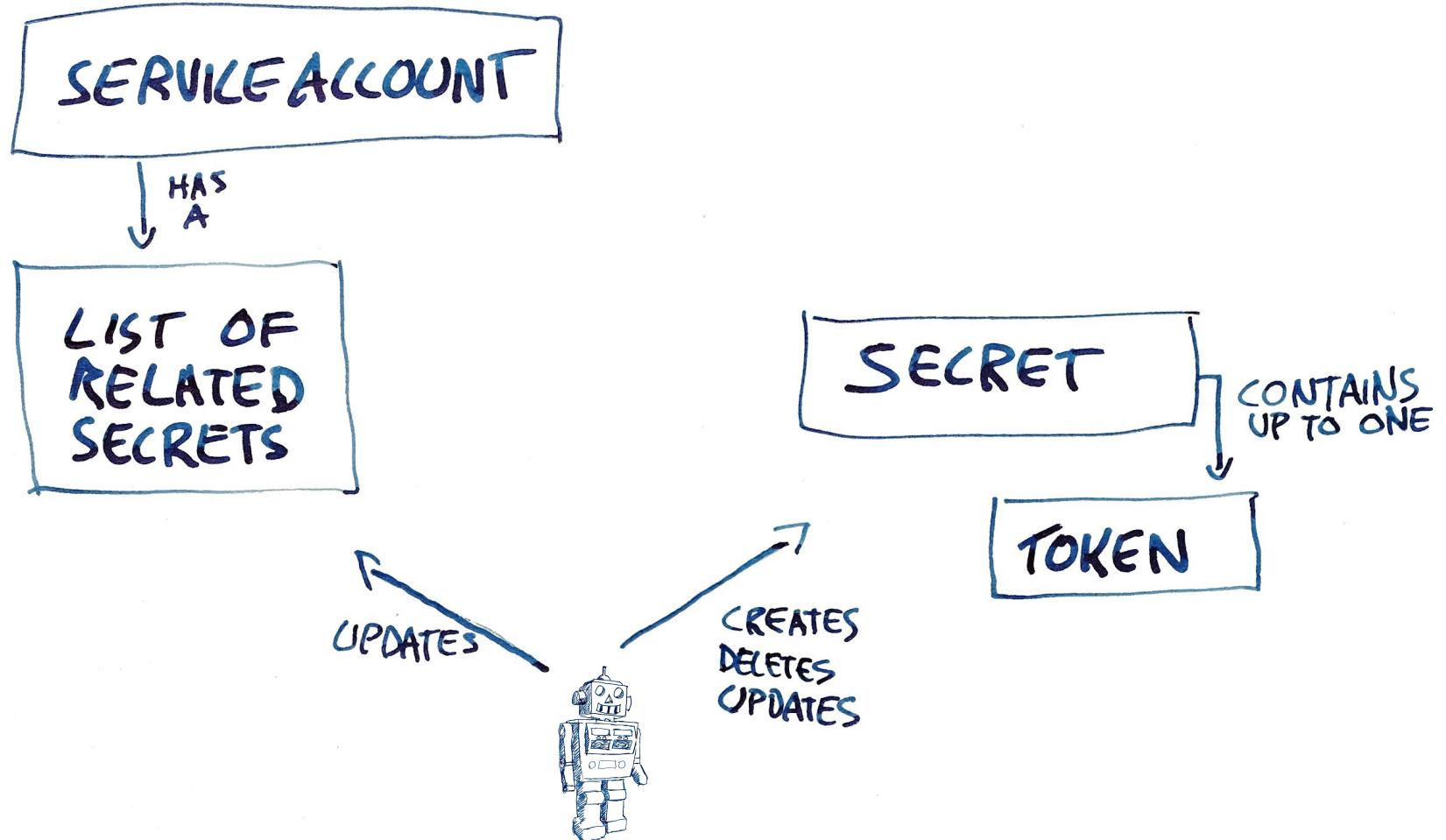
1
A
2
B
3
C
4...
D...

SERVICE Account



1
↓
A
2
↑
B
3
↓
C
4...
↑
D...

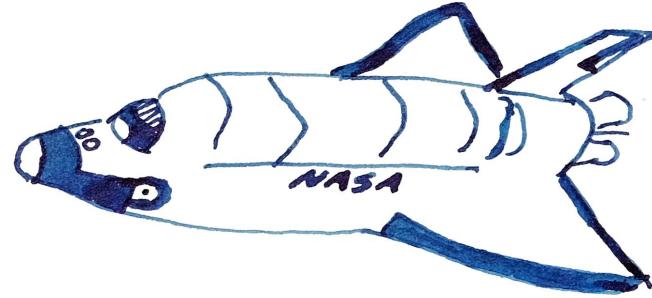
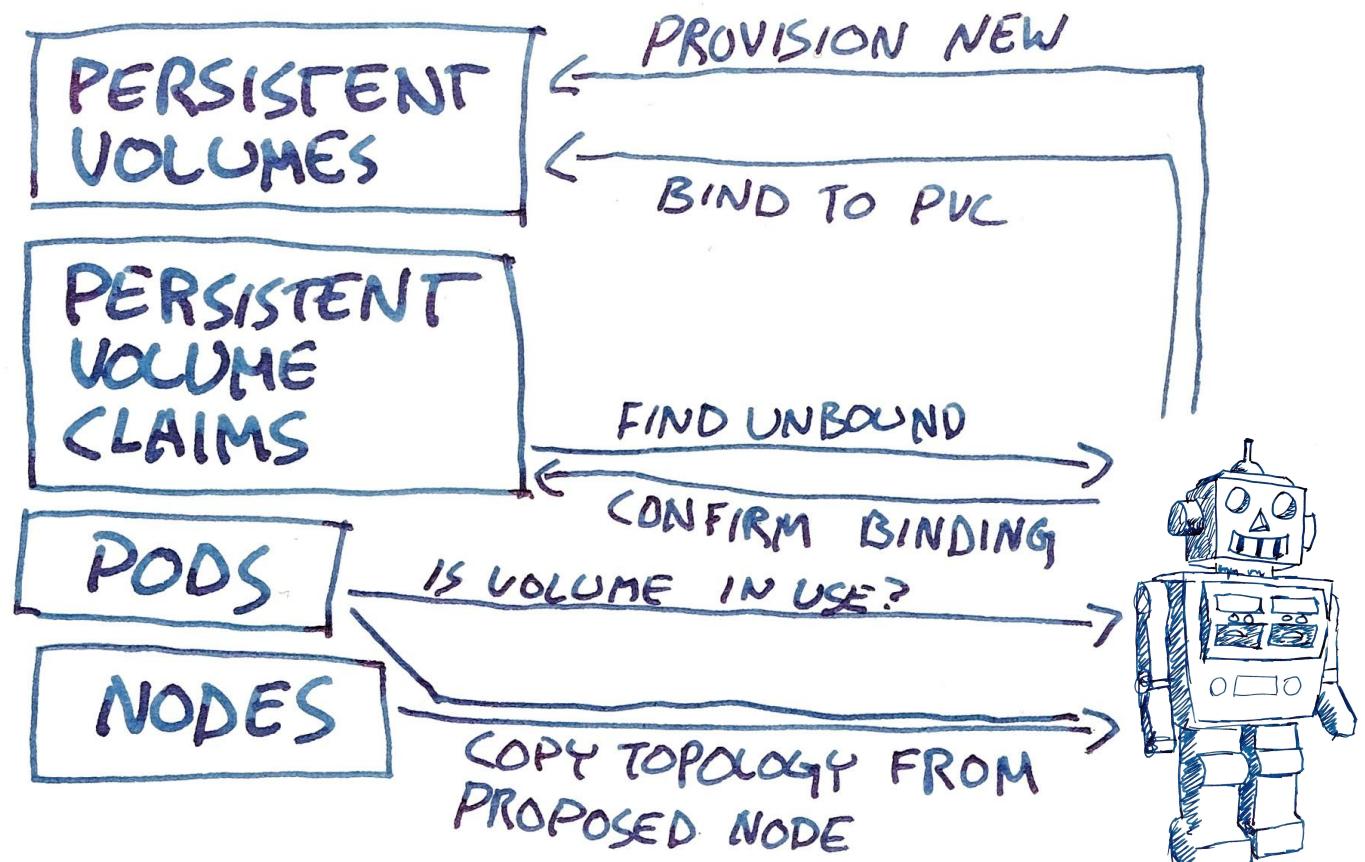
INJECTION ENFORCER



SERVICE ACCOUNT TOKEN CONTROLLER

INJECTION ENFORCER

1 ↴ A
2 ⇄ B
3 ↵ C
4 ... D ...

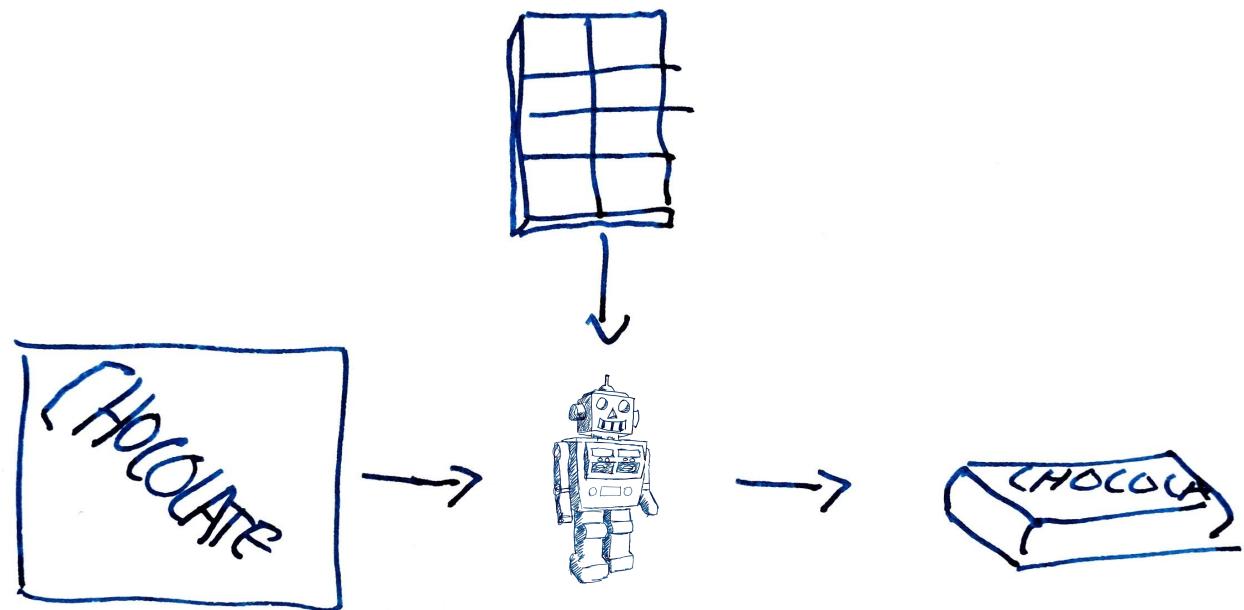


//KEEP THE SPACE SHUTTLE FLYING

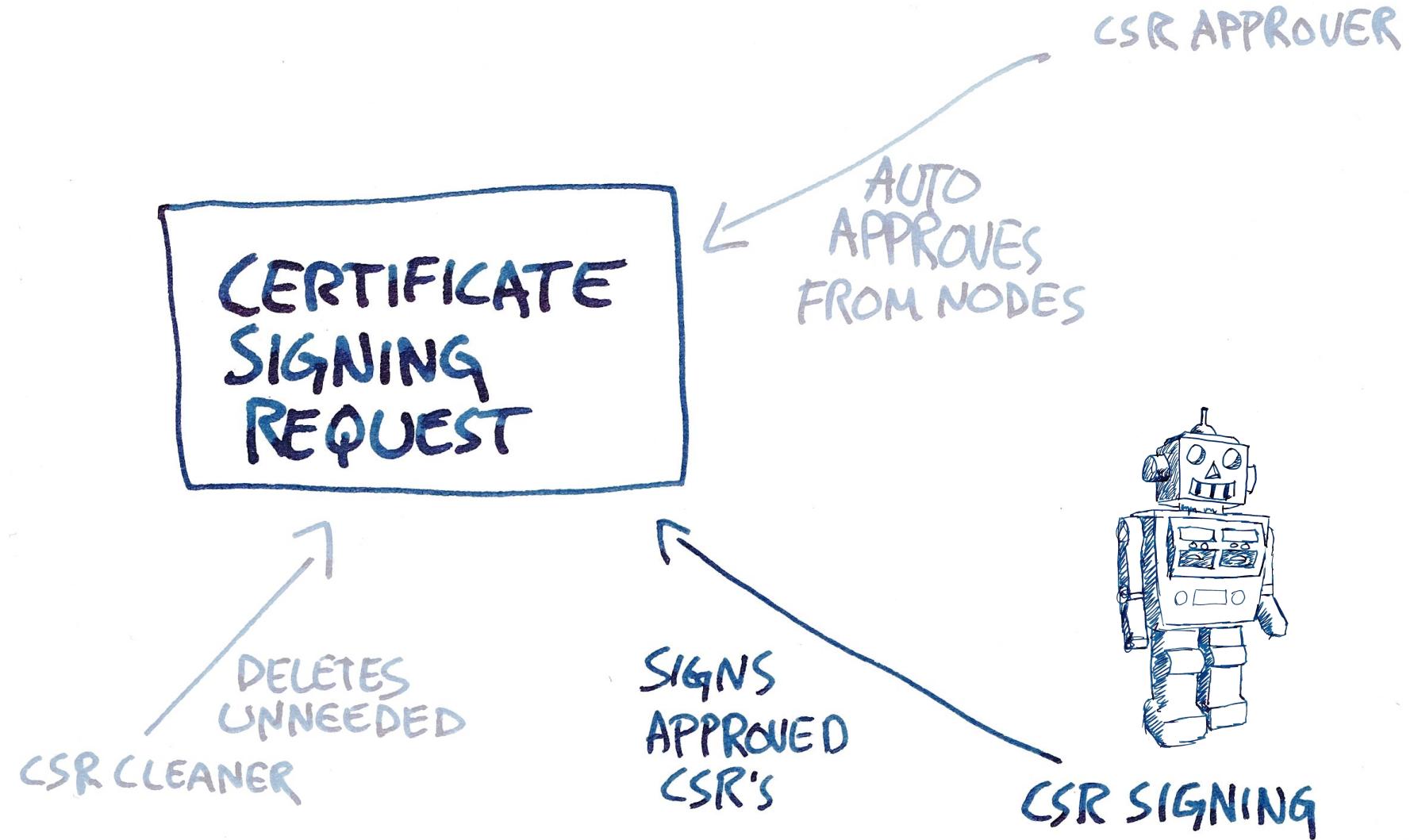
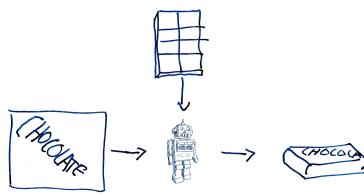
PERSISTENT VOLUME-BINDER

INJECTION ENFORCER

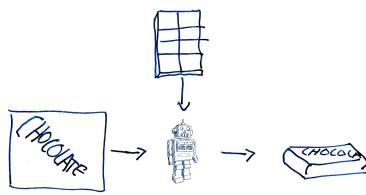
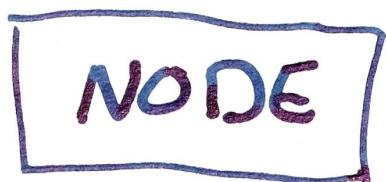
1 ↪ A
2 ↪ B
3 ↪ C
4 ... ↪ D ...



INJECTION ENFORCER



INJECTION ENFORCER



NODE

SET ANNOTATION
"NODE.ALPHA.KUBERNETES.IO/TTL"
BASED ON CLUSTER SIZE

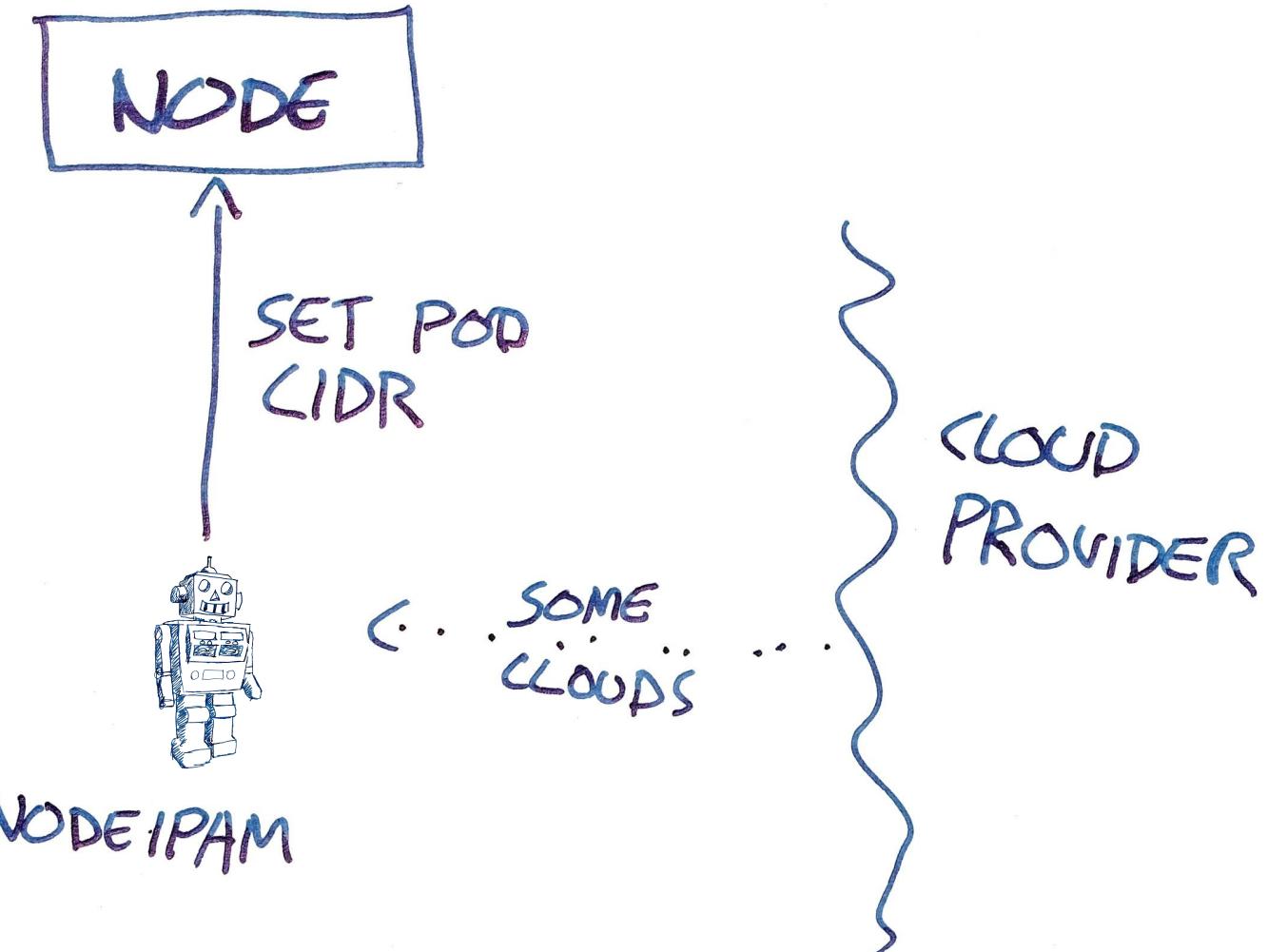
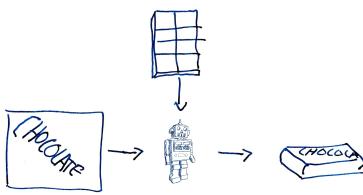


TTL CONTROLLER

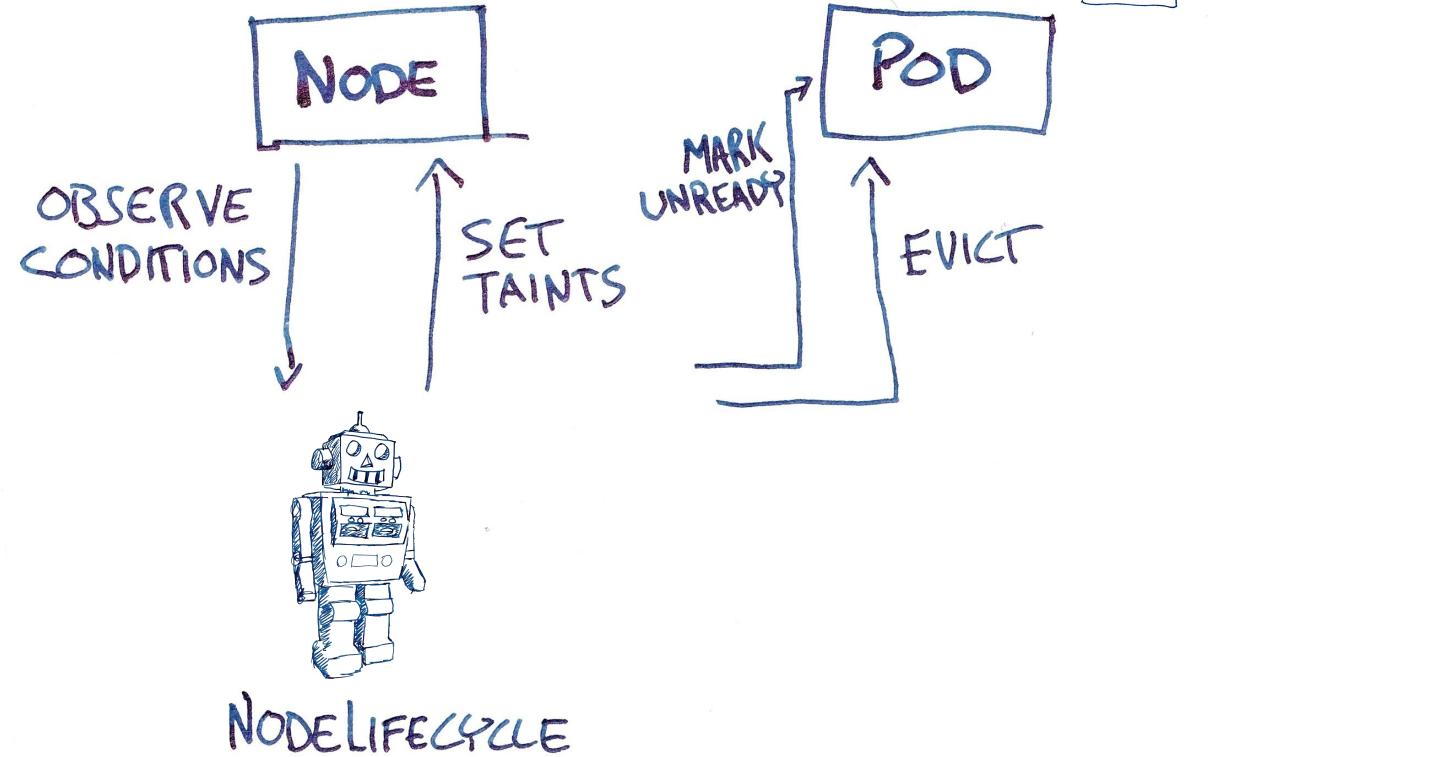


MEANS CACHE TIME
FOR SECRETS, CONFIG MAPS, ...

INJECTION ENFORCER

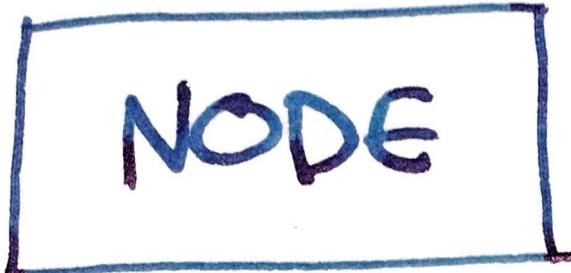
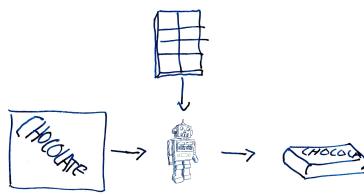


INJECTION ENFORCER

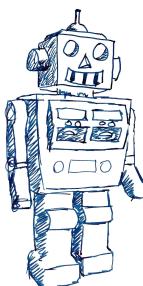


TAKE CHARGE OF THE
K8S RESOURCES IF
SOMETHING HAPPENS TO
KUBELET

INJECTION ENFORCER

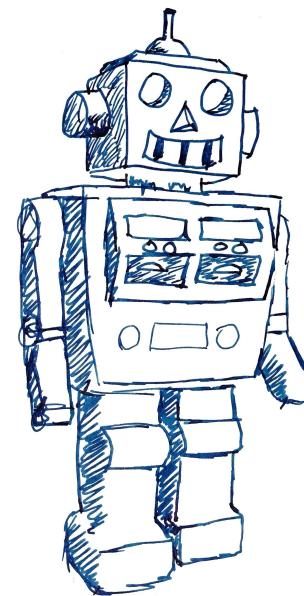
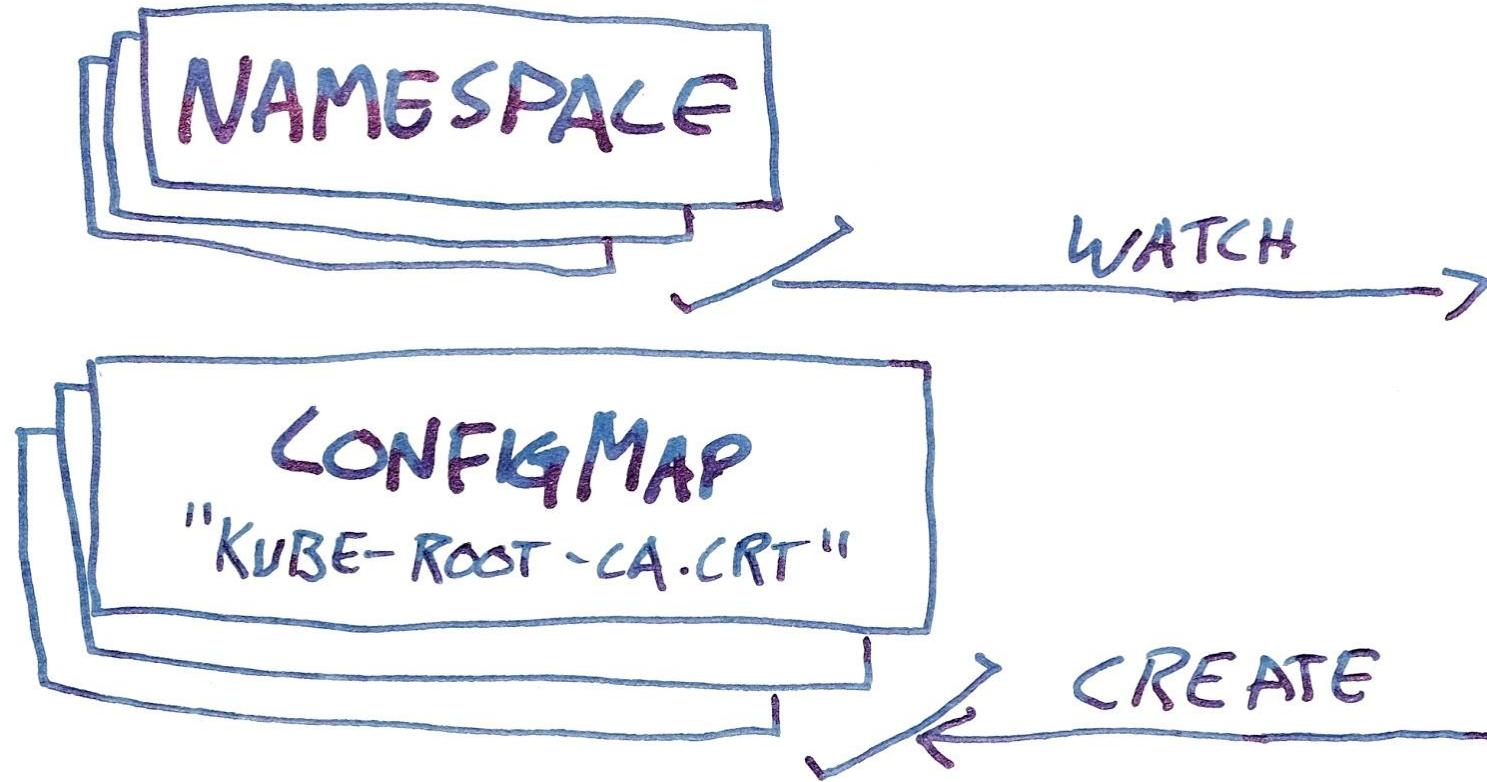


↑ REMOVE "CLOUD" TAINT
ADD CLOUD-SPECIFIC
NODE PROPERTIES
E.G. TOPOLOGY LABELS

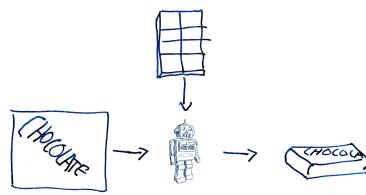


CLOUD-NODE

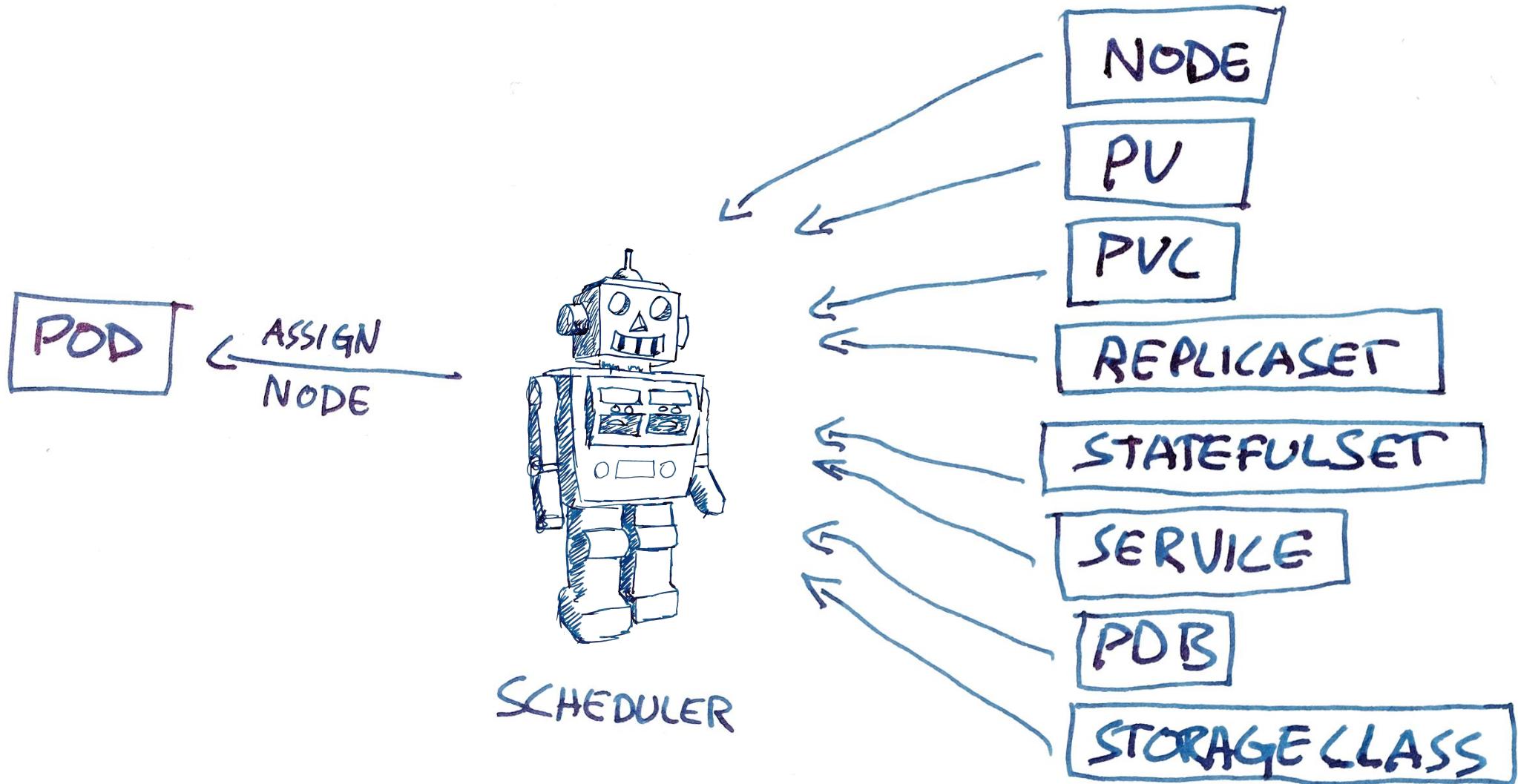
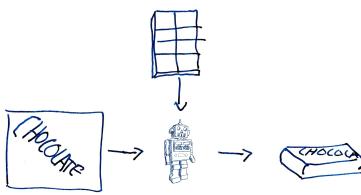
INJECTION ENFORCER



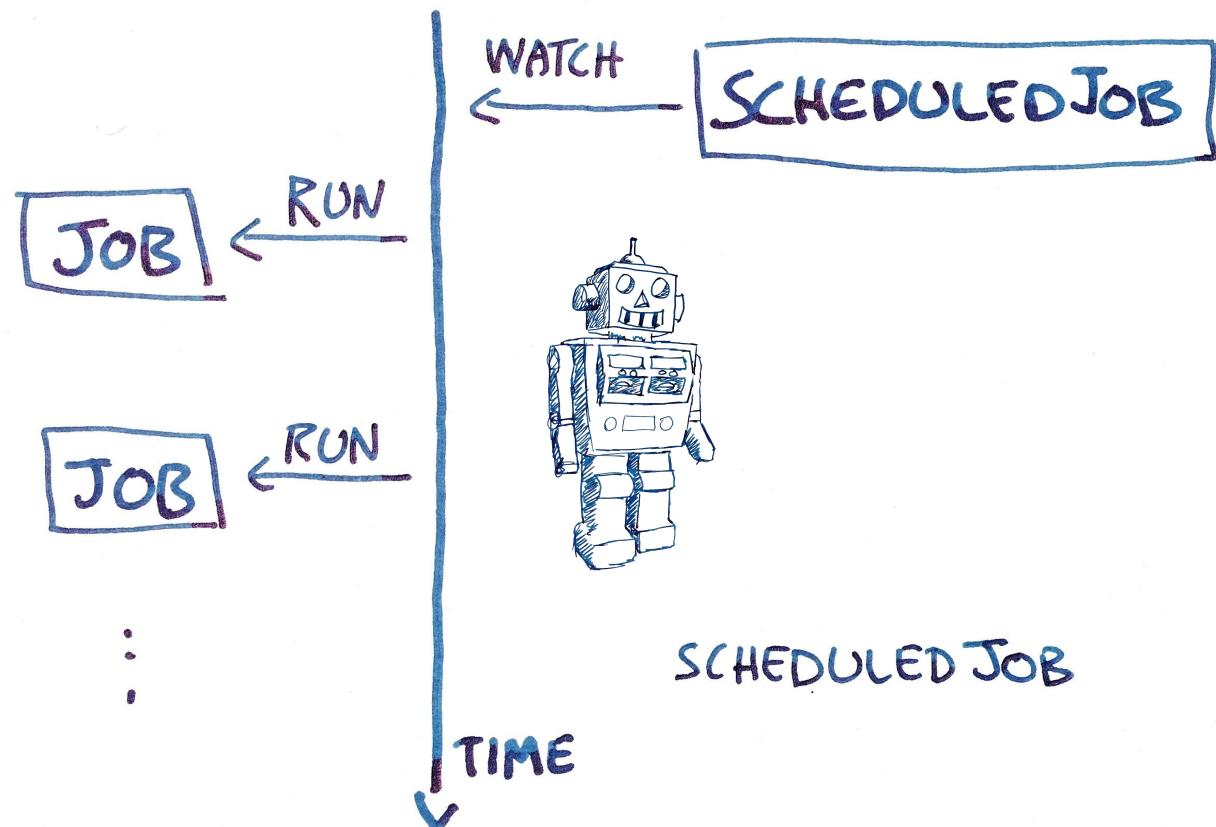
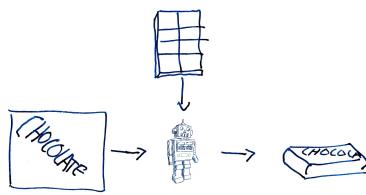
ROOT-CA-CERT - PUBLISHER



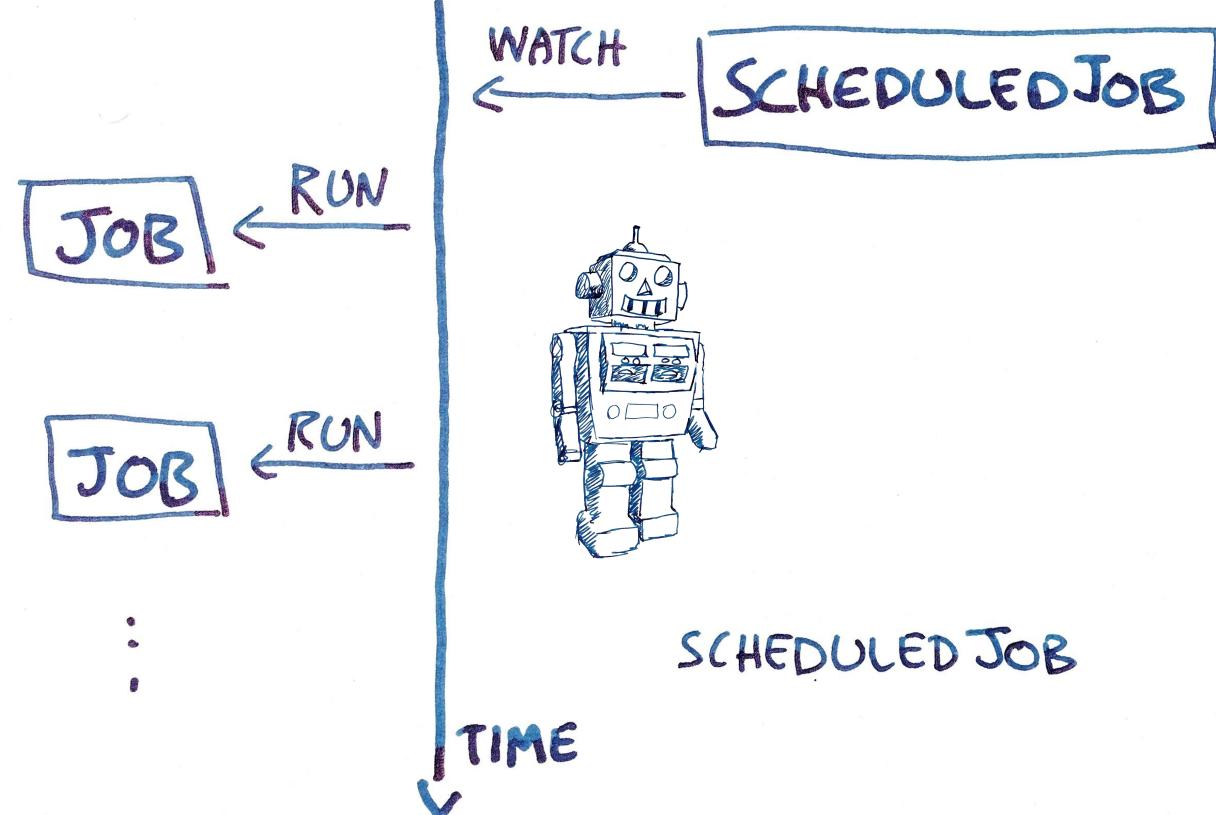
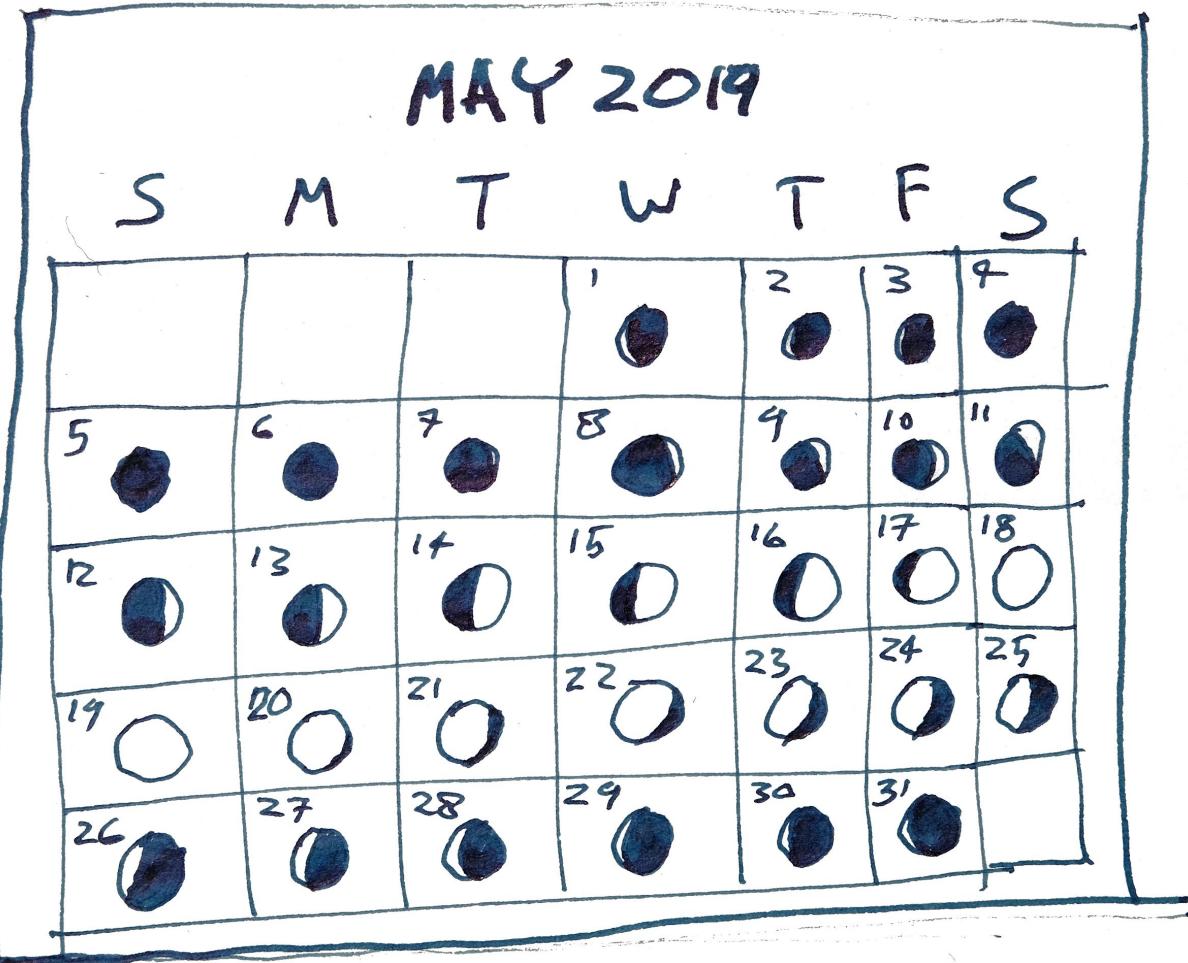
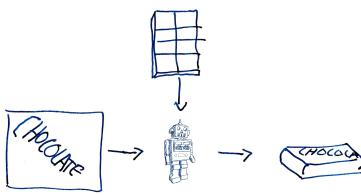
INJECTION ENFORCER



INJECTION ENFORCER



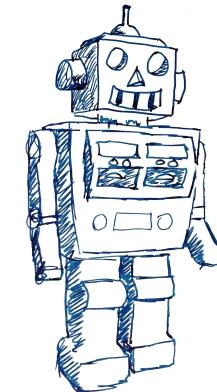
INJECTION ENFORCER



INJECTION ENFORCER

1 ↪ A
2 ↪ B
3 ↪ C
4 ...
D ...

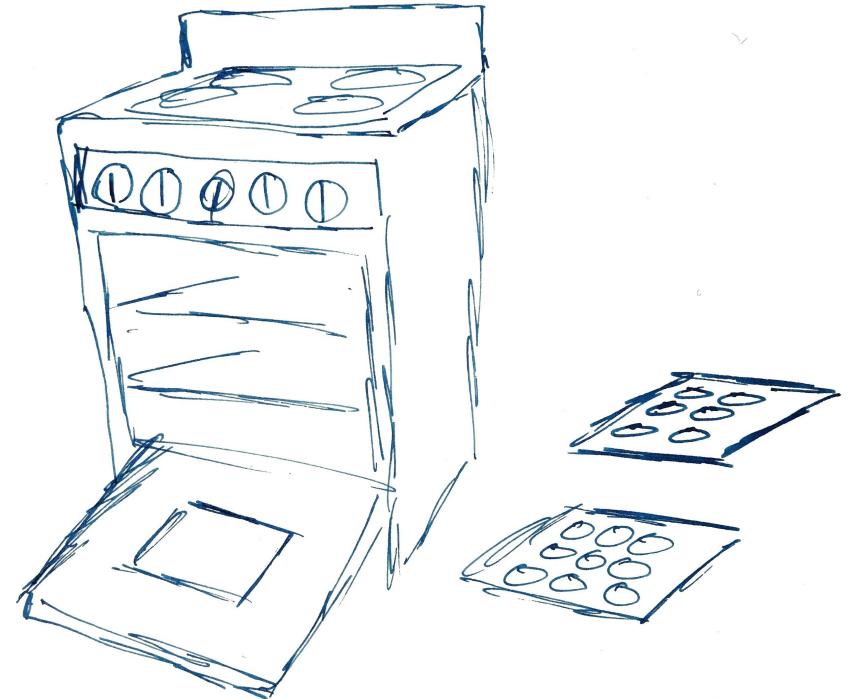
JOB →



N TOTAL
M AT A TIME → POD

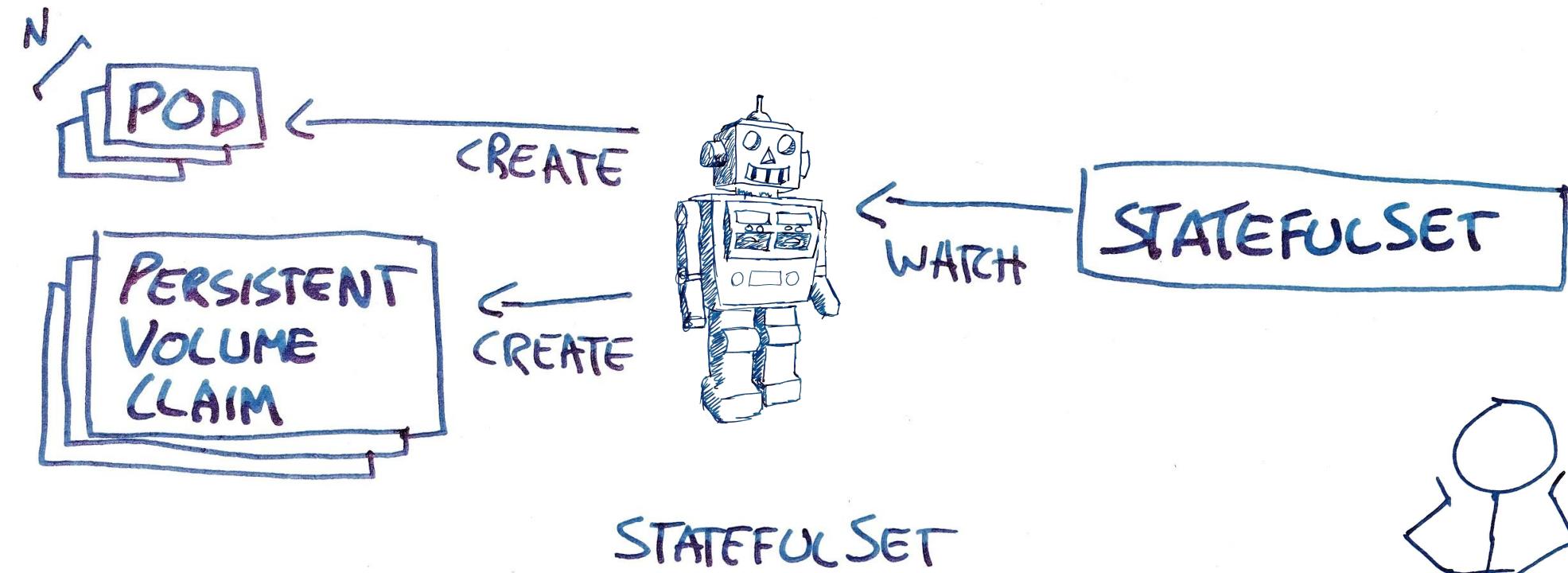
JOB

"COMPLEX INJECTION
WITH A THING"



1
A
2
B
3
C
4...
D...

INJECTION ENFORCER

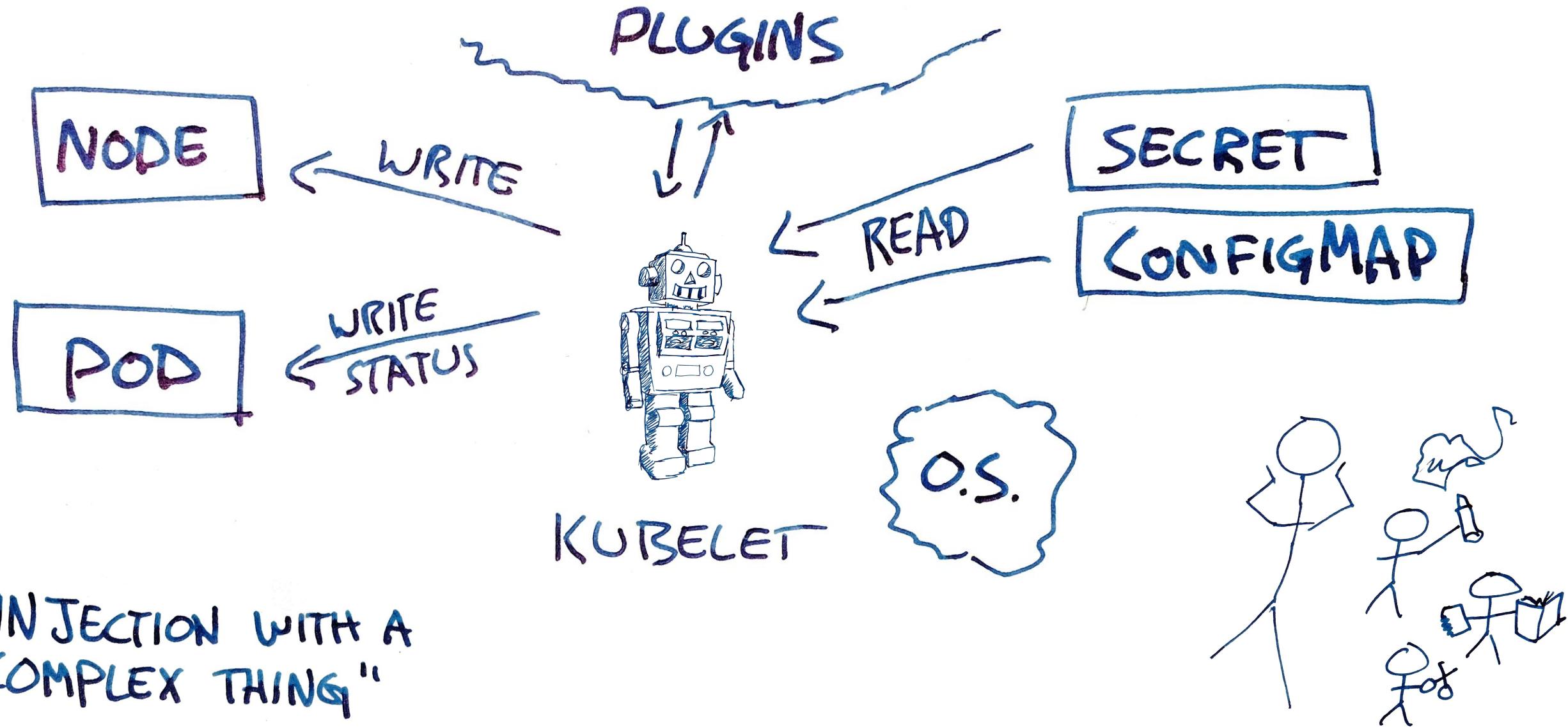


"INJECTION WITH A
COMPLEX THING"

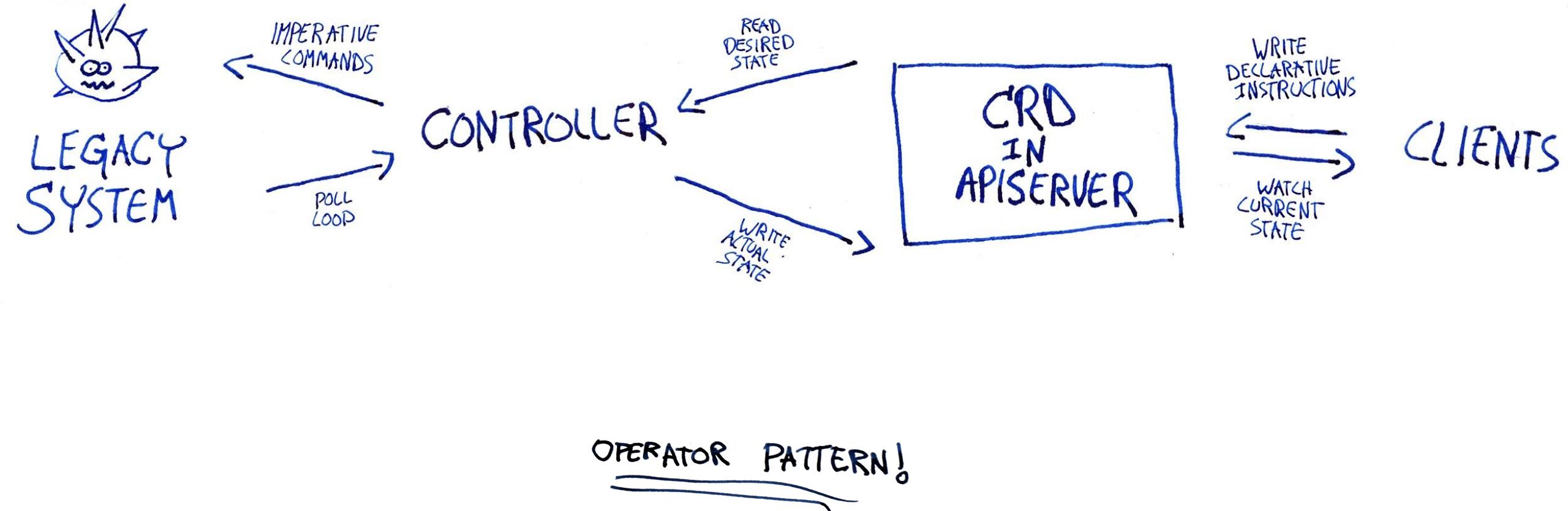


INJECTION ENFORCER

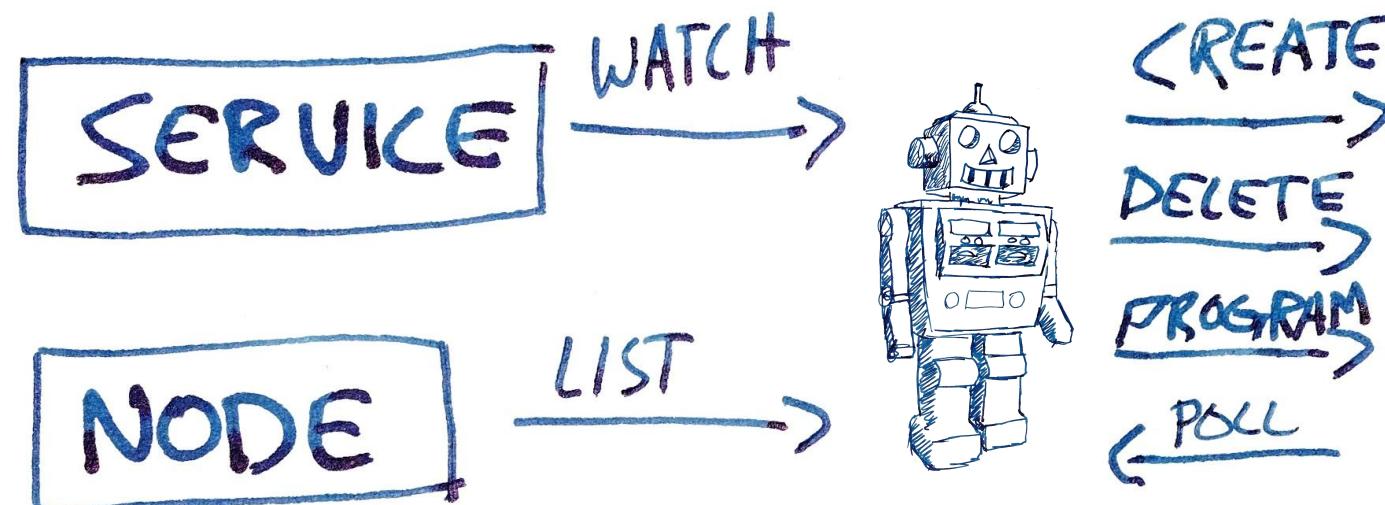
1 ↪ A
2 ↪ B
3 ↪ C
4 ...
D ...



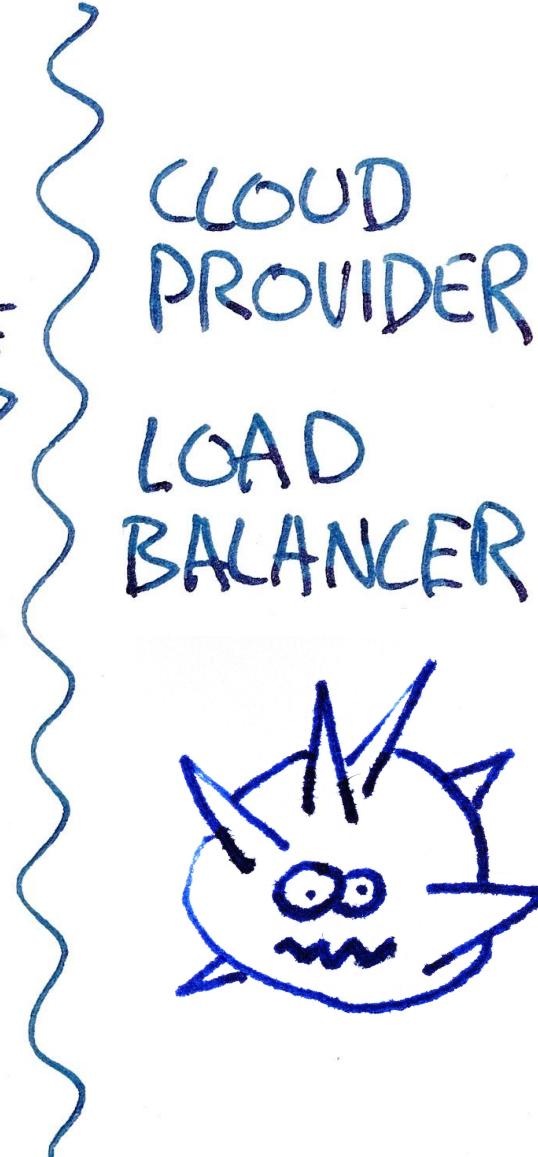
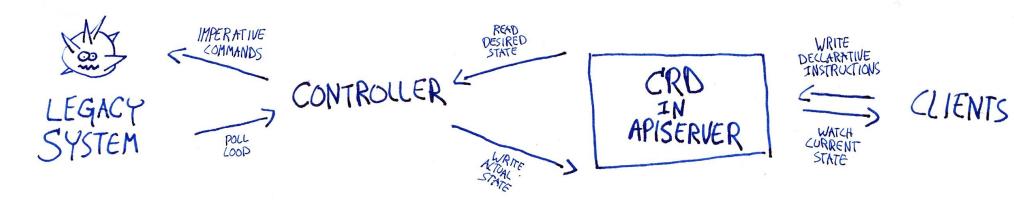
INJECTION ENFORCER



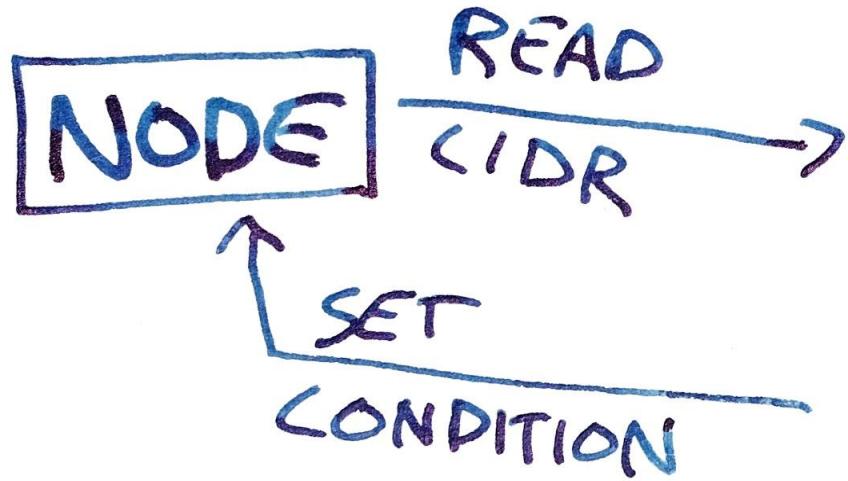
INJECTION ENFORCER



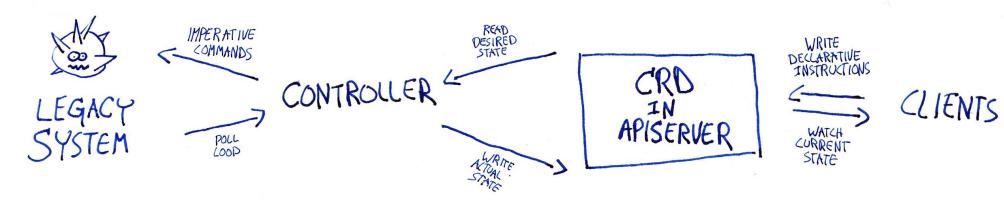
SERVICE



INJECTION ENFORCER



ROUTE

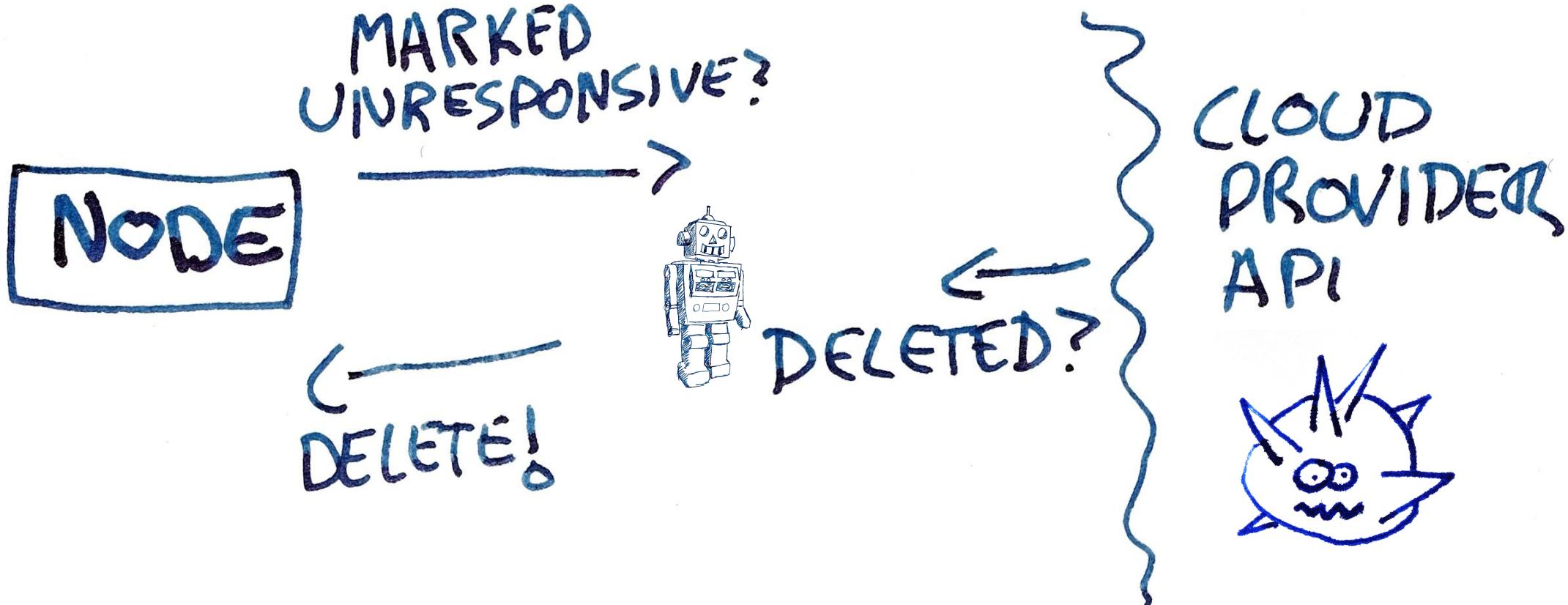
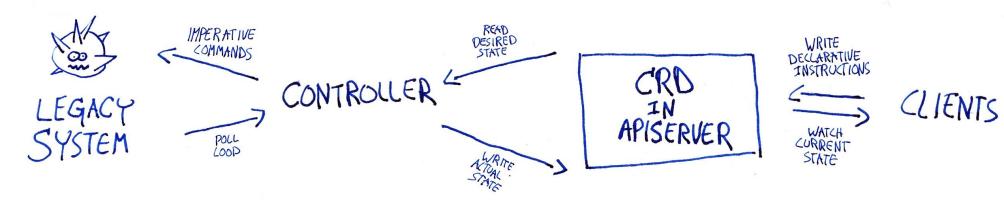


CLOUD PROVIDER
NETWORK FABRIC

MAINTAIN
→
ROUTES



INJECTION ENFORCER



CLOUD-NODE-LIFECYCLE

THAT'S ALL* OF THEM !

READY TO
WRITE
YOUR OWN??

*MOST

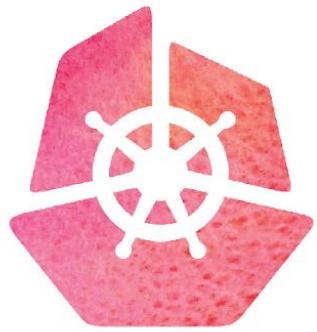
SOUND LIKE A FUN
PROBLEM SPACE?

GET INVOLVED!!

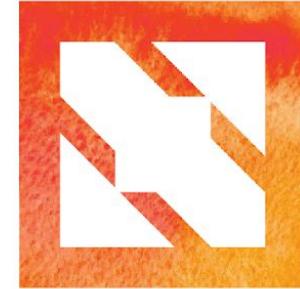
→ SIG API MACHINERY

→ SIG APPS

→ SIG ARCHITECTURE



KubeCon



CloudNativeCon

Europe 2019