

***Modeling Paradigms & Tools for Autonomy Management in
Complex Networked Systems:
A case for Probabilistic Model-Checking based approach***

***Kaliappa Ravindran (Professor) City University of New
York (CUNY -- Graduate Center)***

Graduate student participants:

Gregory Javens (Ph.D.)

Gwang Yu Ding, M.S. [currently employed in China]

July 2026

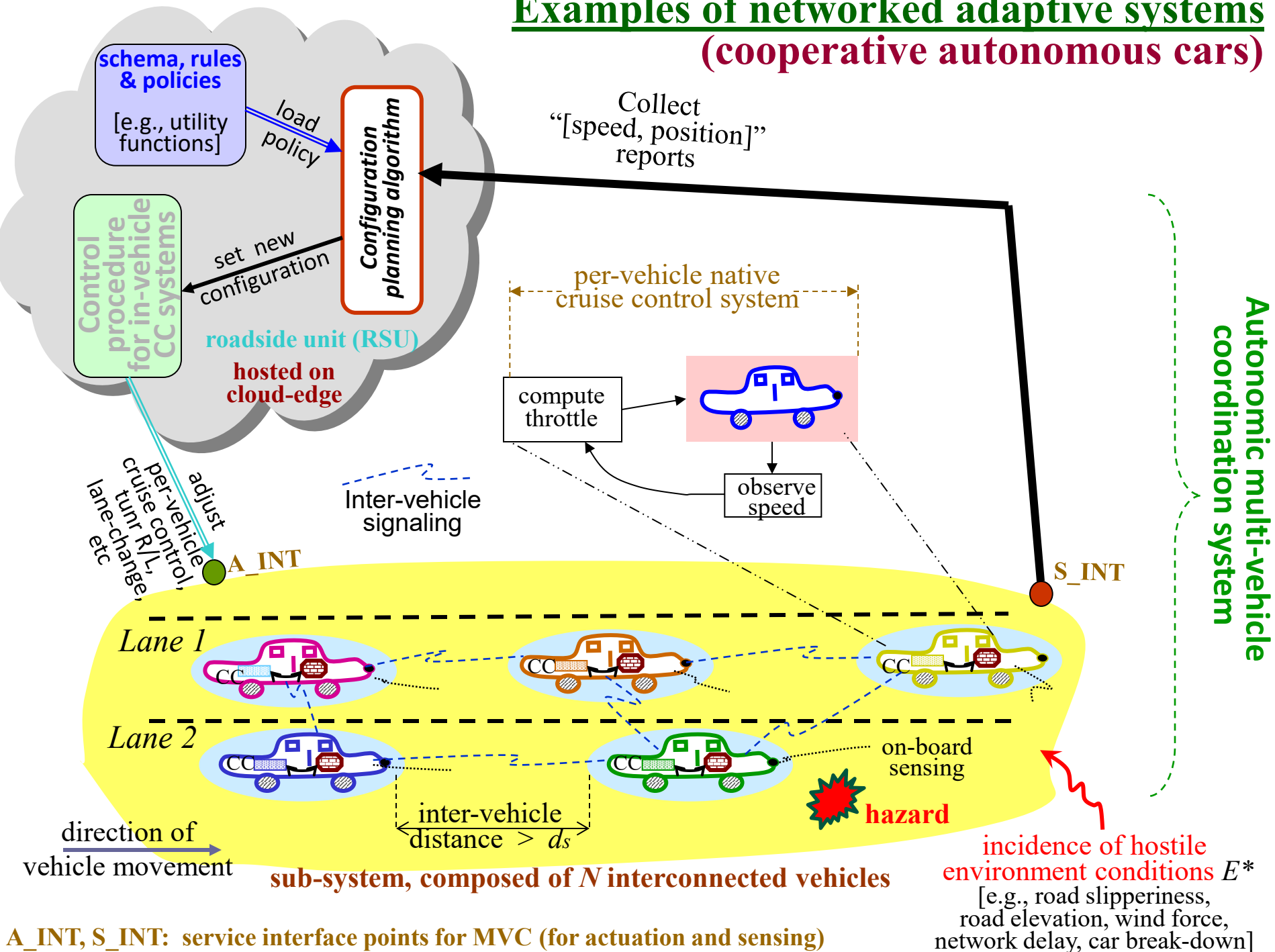
Salient points of our research

1. Self-managing systems with cyber-physical components
2. Algorithm verification *versus* system assessment
 (non-functional properties:
e.g., stability, convergence, agility, adaptivity, etc)
3. Software engineering view of adaptive networked systems
4. Model-checker tools for management of adaptation processes
(Advantage: *Reusable, Reconfigurable, Higher quality of adaptation*
Need for rapid prototyping is exacerbated as systems become complex)
5. *System-level Case Studies*:
 - i) Resilient supply-chain Management for Product Distributions
 - ii) UAV-based surveillance to Enforce Secure Cyber-fencing
 - iii) Collaborative driving autonomous cars in intelligent highways
6. Discussion of alternate modeling tools:
 - 6.1 Probabilistic Model-checkers (PROMELA*, PRISM, STORM, PAT, . .)
 - 6.2 Operational models (closed-form mathematical expressions, . .)
 - 6.3 Miniature prototypes of real-world systems

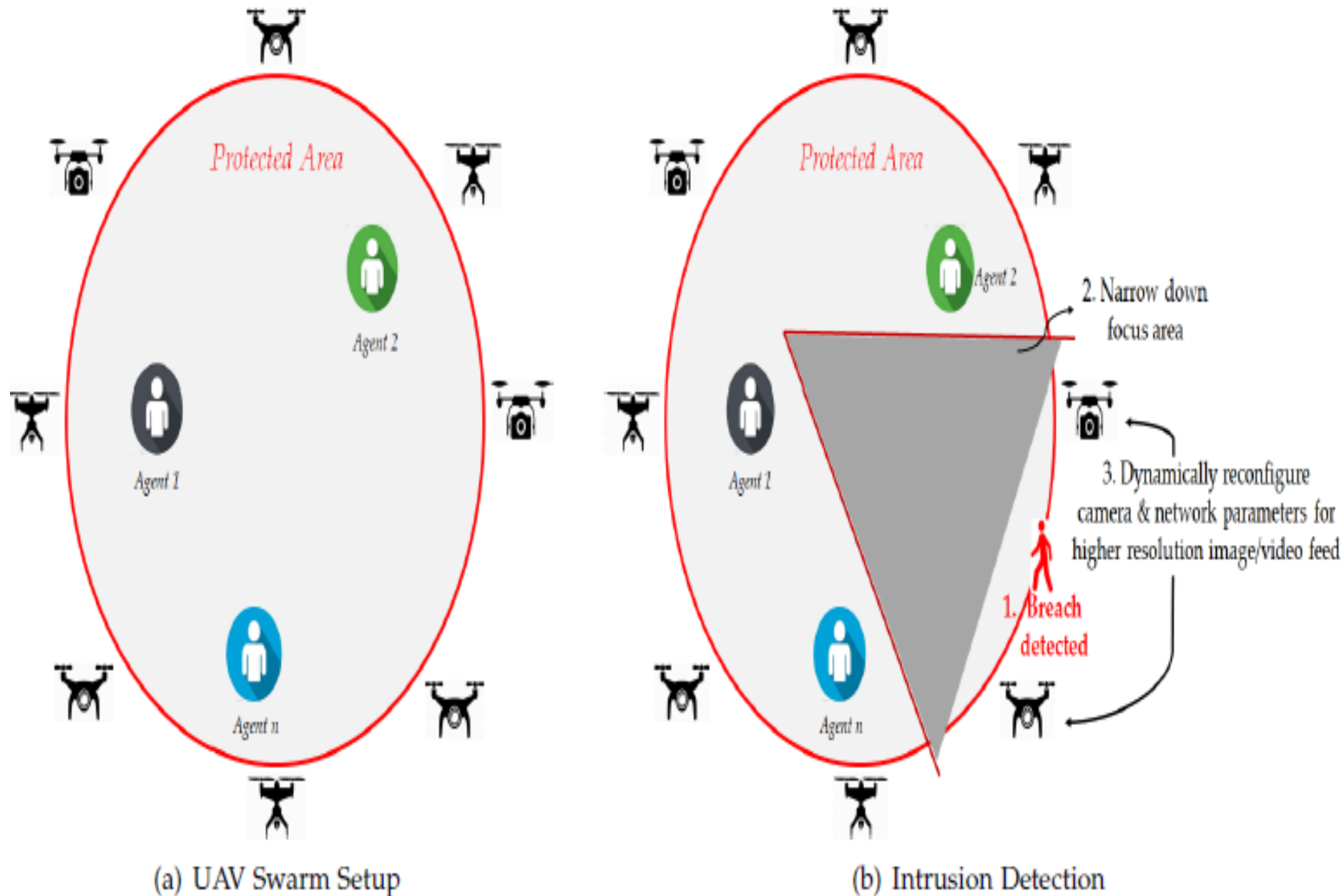
Fundamental concepts of complex network systems

- How complex is the handling of “complexity” ??
- **Proactive** planning to handle faults (instead of **reacting**)
[balancing between **pessimistic** vs **optimistic** approaches]

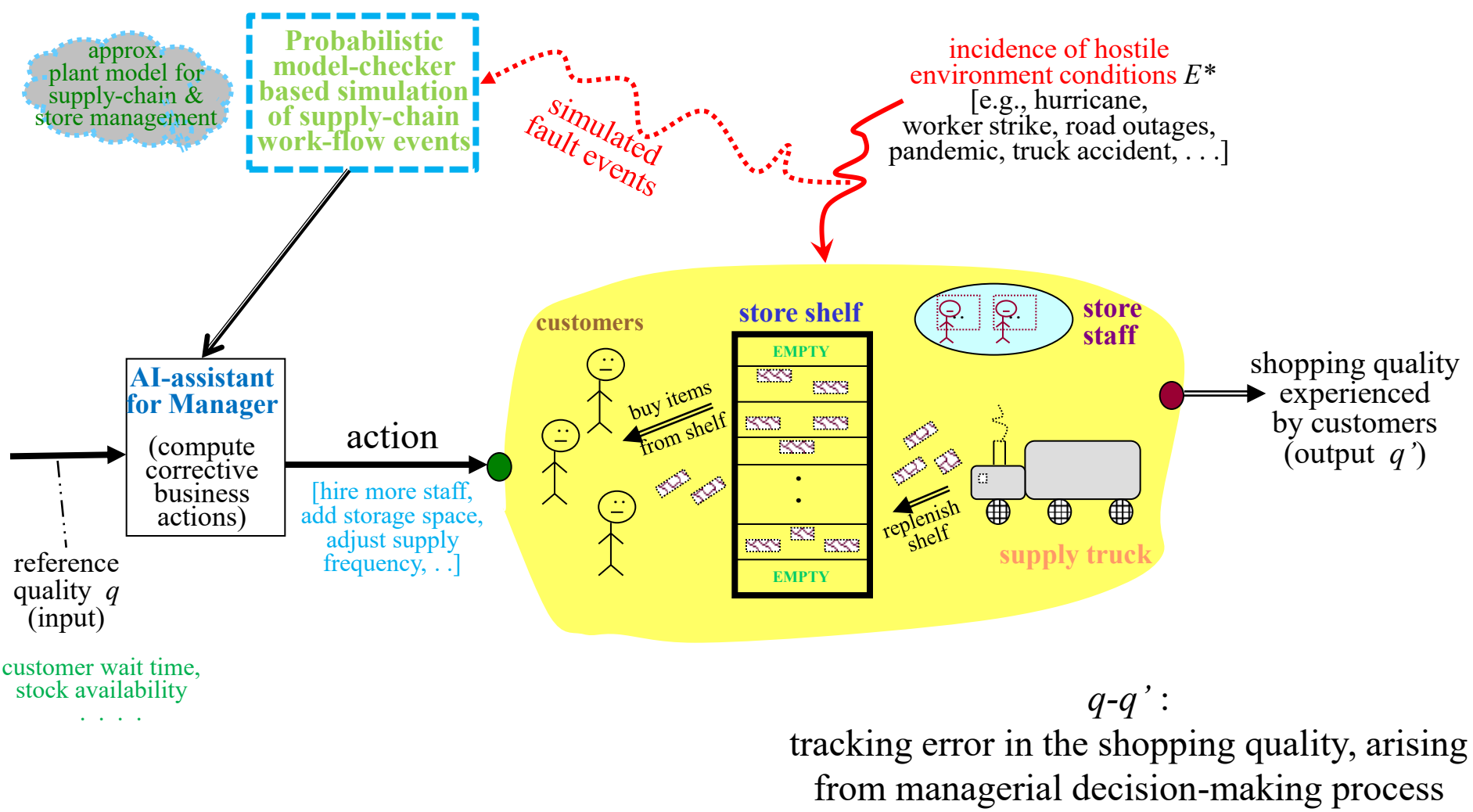
Examples of networked adaptive systems (cooperative autonomous cars)



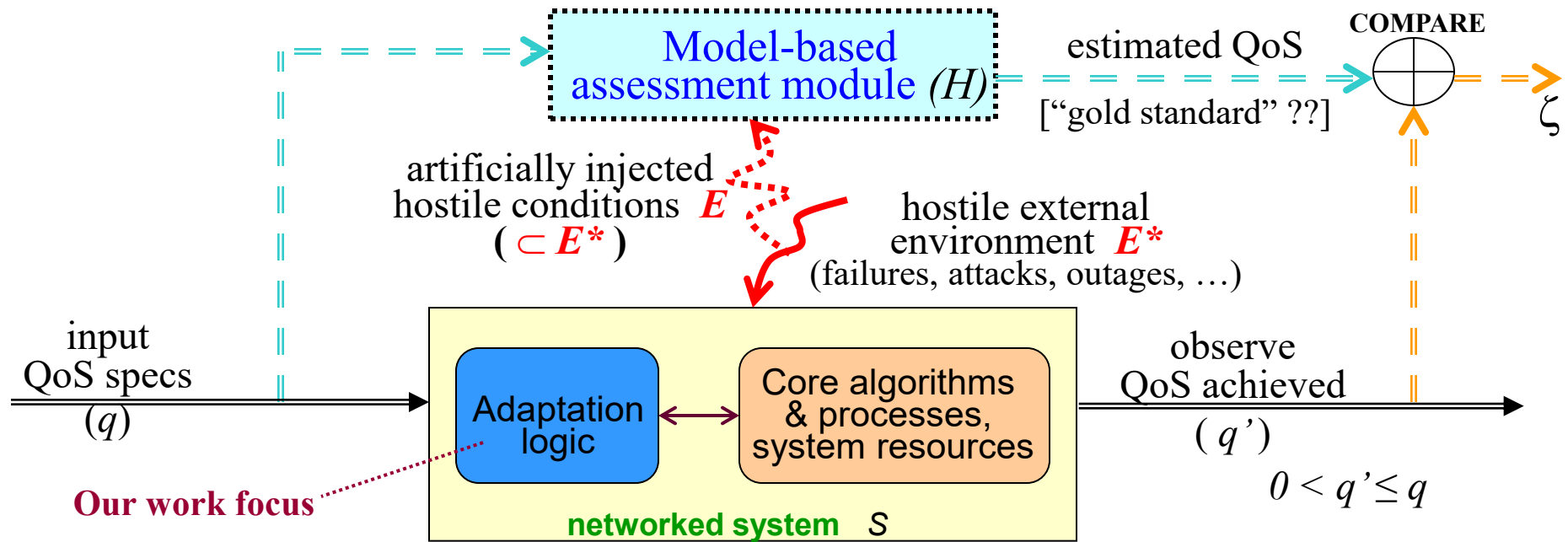
Example 2: *Cyber-fencing of Metro-cities by multiple UAVs or drones*



Example 3: Fault-resilient supply-chain network for product distribution



Quantitative view of designing of resilient systems (digital-twin)



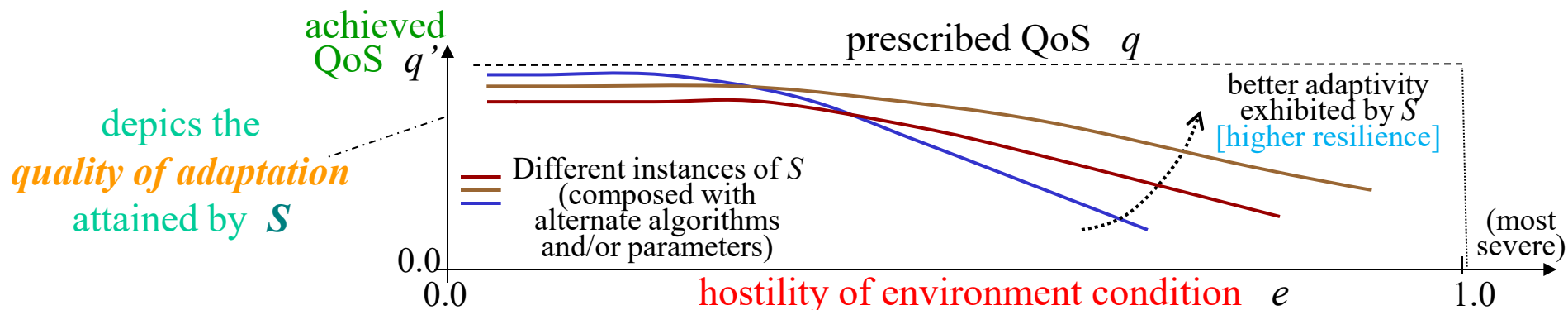
$\Psi = 1.0 \rightarrow$ 100% trustworthiness bestowed on S for meeting the QoS specs

$= 0.0 \rightarrow$ not trustworthy at all

A coarse measure of the "goodness" of system S under harsh condition $e \in E^*$:

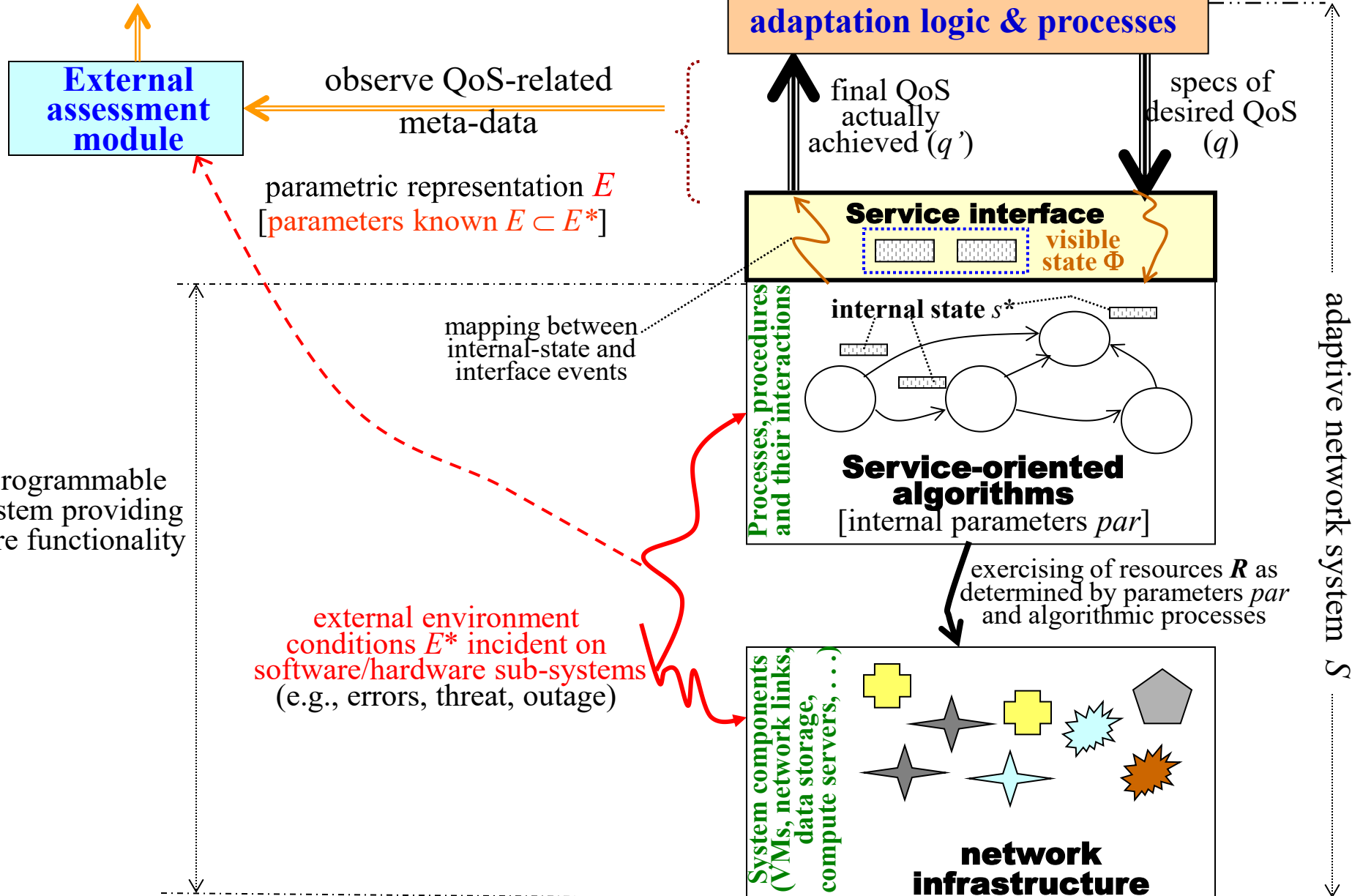
$$\Psi_{(S,q,e)} = \left[1 - \frac{(q-q')}{q} \right]$$

$$0.0 < \Psi_{(S,q,e)} \leq 1.0$$



Software Engineering view of self-managing network systems

reason about how good is the system S in meeting QoS specs



Motto behind resilience assessment

A system that is good
but is not *verifiably good* is
not good enough !!

[e.g., quality rating of restaurants, hotels, taxi services, . . .]

3-star, 4-star, etc

In contrast, **system auditing** reasons about
why an actual QoS attained is less than a desired QoS specs
[say, less resources, node outages, laxity of algorithms, . . .]

Uncertainties emanating from external environment

QoS specs q , algorithm parameters par ,
system resource allocation R are usually controllable inputs

In contrast, environment parameters $e \in E^*$ are often uncontrollable and/or unobservable, but they do impact the system-level performance (e.g., component failures, network traffic fluctuations, attacks, etc)

environment parameter space:

$$E^* = \underbrace{E(yk)} \cup \underbrace{E(nk)} \cup \underbrace{E(ck)} \cup E(uu)$$

parameters that the
designer knows about
(known “knowns”)

parameters that the
designer does not
currently know about
(knowable “unknowns”)

parameters that the
designer can never
know about
(known “impossibilities”)

What about unknown “unknowns” $E(uu)$?? $\rightarrow$ (late) Hon. D. Rumsfeld, US defense secretary

Algorithm design decisions face these uncertainties --- so, designer makes
certain assumptions about the environment

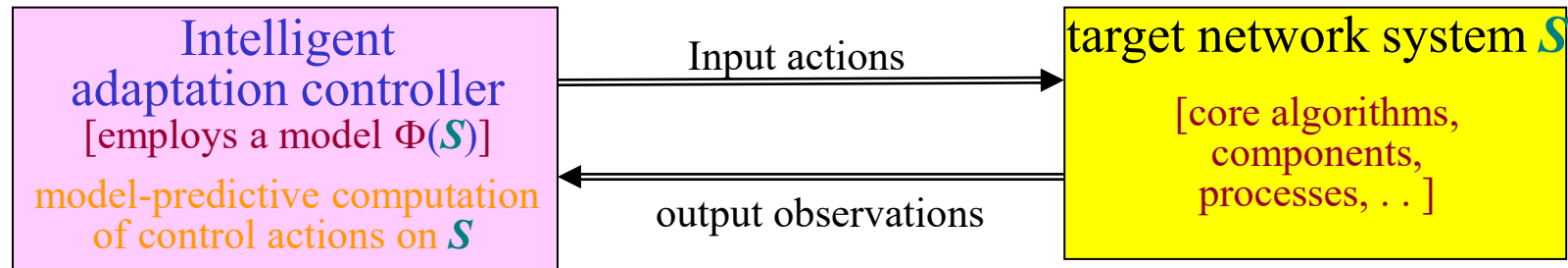
[e.g., at most 3 UAVs in a 10-vehicle surveillance team will fail during operations].

When assumptions get violated, say, due to severe attacks,
algorithms fall short of what they are designed to achieve

$\rightarrow$ Evaluate how good an algorithm performs under strenuous environments

Categorization of modeling paradigms

Model $\Phi(S)$ predicts the future behavior of a networked system S



1. Operational models of S
Quickly computable, but low accuracy of prediction
(due to the simplifying assumptions of **piecewise-linearity, function-separability, . . .**)
2. Close-to-real software prototyping of S **(to mimic its run-time behavior)**
High accuracy of prediction, but building system prototype incurs lot of development efforts, and long turn-around time
3. Model-checker based discrete-event simulation of S
Offers a middle ground: Reasonable accuracy of prediction, Rapid & flexible composition of alternate system-level functions, model reusability & verifiability, low model-development costs

Metaphorical statement of Prof. Judea Pearl (UCLA) on Modeling Accuracy

[well-known for works on Game-Theory & Epistemic reasoning]

The best model of a “cat” is the “cat” itself !!

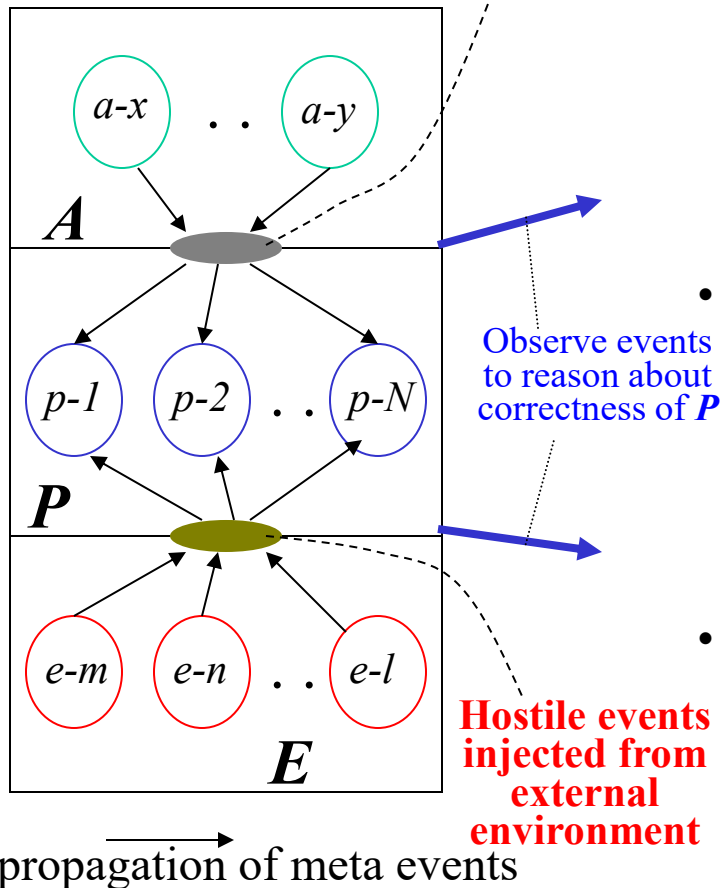
Any computationally-tractable “modeling paradigm” is bound to have inaccuracies (due to the simplifying assumptions made about the system)
[it is just a question how inaccurate one paradigm is relative to others]

- ➔ **What are the sources of inaccuracies in a modeling paradigm**
- ➔ **What is the “gold-standard” for modeling accuracy ??**

Event-based view of *Distributed algorithm* in a networked system of things

Algorithm P : governs the interactions between two or more processes through a set of rules to realize a service-layer property G

goal G at
service interface



- externally visible interface to P , seen by application A , that prescribes *safety*, *liveness* conditions to be met by P

safety $\rightarrow$ “bad thing” never happens during execution of P

liveness $\rightarrow$ “good thing” thing eventually happens

- Algorithm processes of P propagate meta-events & actions to counter the effects of hostile environment E (e.g., ‘packet retransmit’ action to counter ‘packet loss’)

- operating environment E that attempts to disrupt the pursuit of P towards goal G (e.g., ‘packet loss’ during ‘data transfer’ over a network)

What does the correctness of a distributed algorithm mean ??

A distributed algorithm P implements a set of functions, often running on different machines on a network, to meet the goal G set by applications $\{A\}$, in the presence of perturbations emanating from the operating environment E

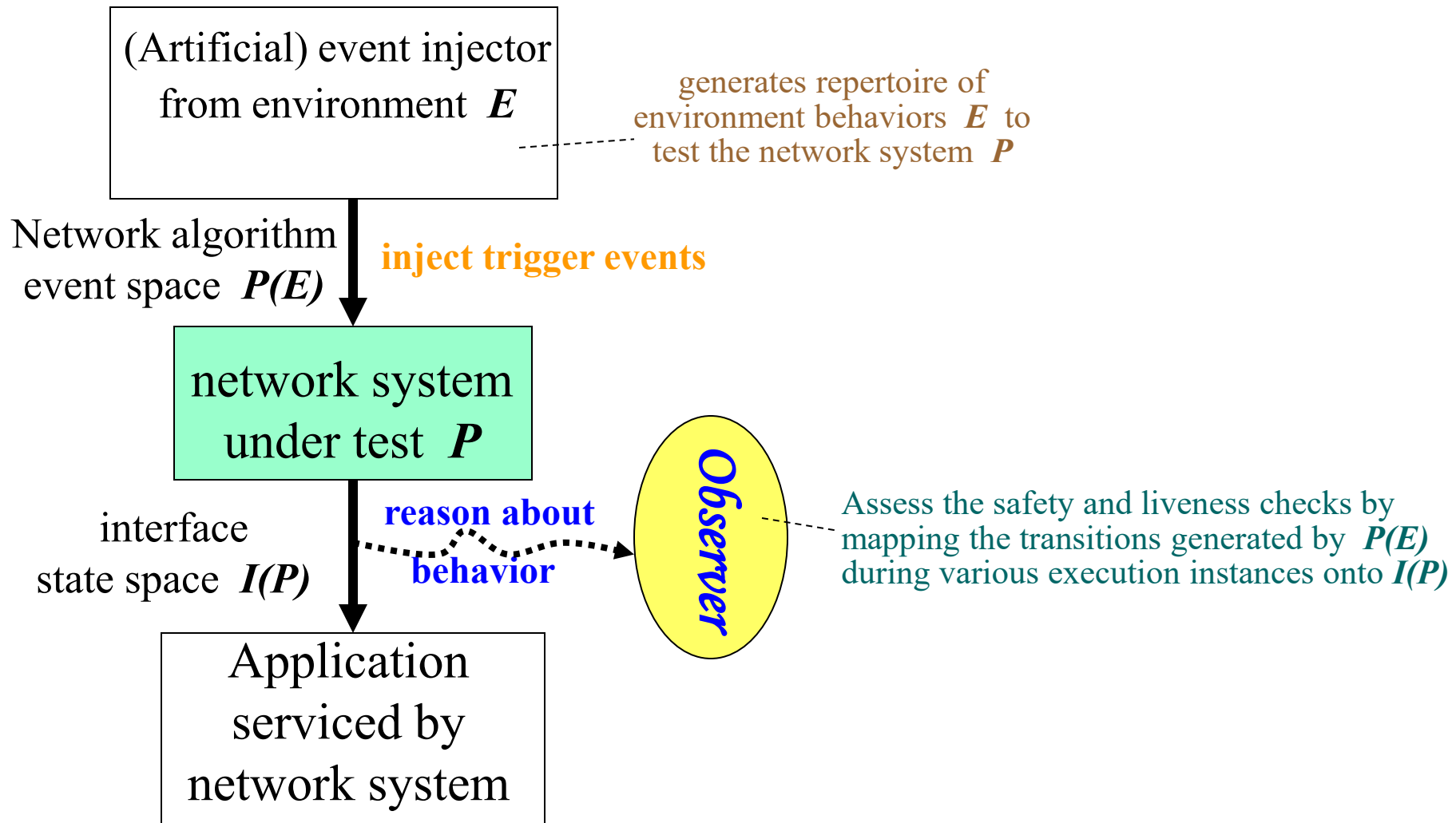
E.g., ‘Uber taxi driver’ interacts with multiple entities/data (booking site, weather report, road transport agency, etc) for on-time passenger drop-off

CLAIM:

$P(E)$ meets G under certain assumptions about E

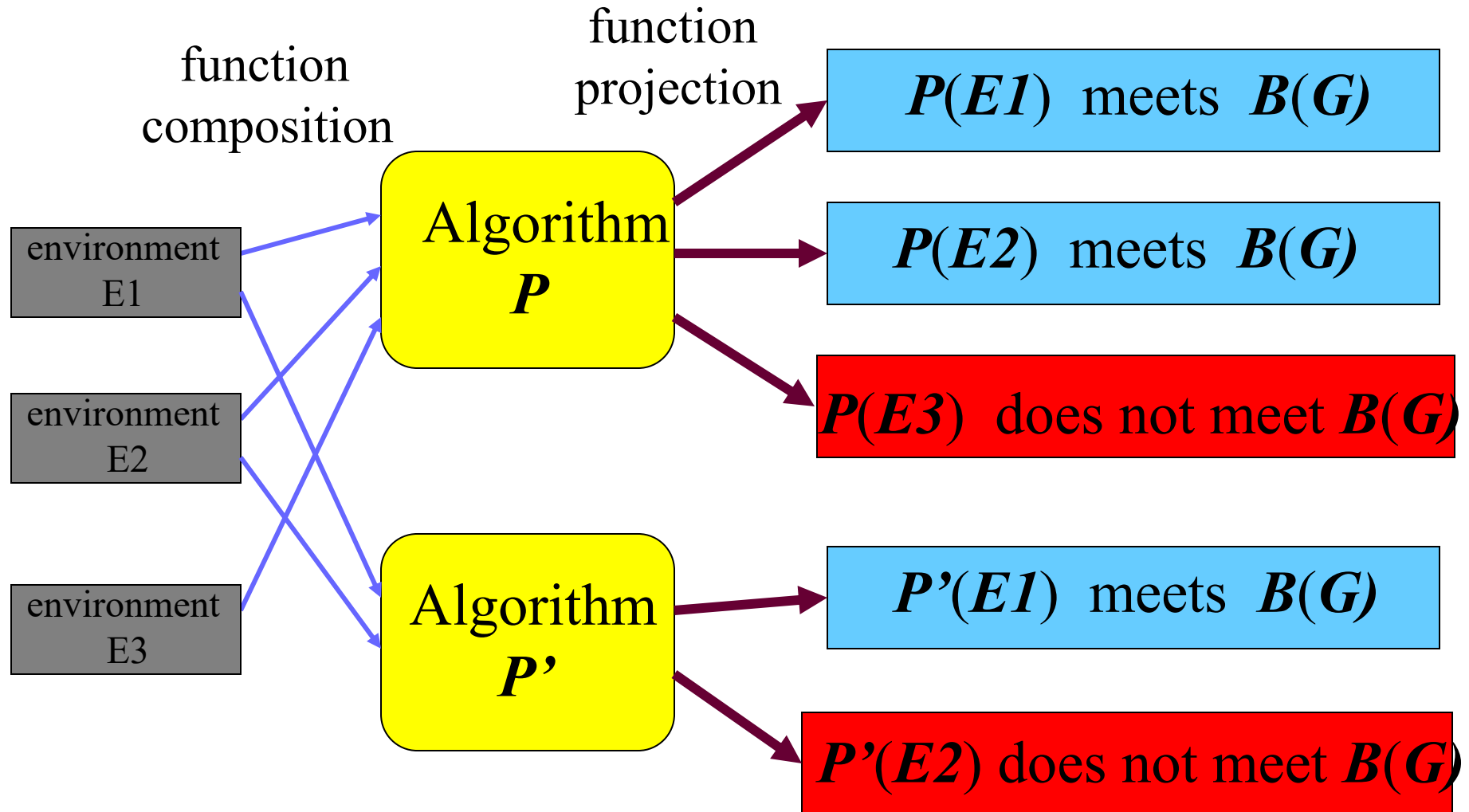
Algorithm correctness refers to a verification of the above claim, i.e., checking through some means as to whether P does in fact meet the goals G when subject to the disturbances emanating from E

Testing modules for fault-tolerance & resilience



Employ *state transition projection relations* for automating the safety and liveness checks from the execution traces of state transition sequences generated by $P(E)$ at run-time

‘function-based’ view of algorithm verification



P' is less resilient than P (so switch from P' to P if/when needed)

$E3$ is more hostile than $E2$; $E2$ is more hostile than $E1$.

$B(G)$: Assertions on the service interface state depicting goal G

Quick tour of PROMELA language for model-checking

Our extensions to PROMELA (Protocol Modeling Meta-Language)
[originally designed at AT&T Bell-Labs, in mid-90's (to analyze telecom systems)]

Key conceptual elements in formulating model $\Phi(S)$ of a network system S :

- Cause-and-effect relationship between actions/events
- State-based encapsulation of past and future events/actions

- PROMELA software overlays to assign state-transition probabilities and SPIN trace-analysis (with home-grown extensions to explicitly control the entry of PROMELA threads for the execution of *Guard* statement blocks)
- Rapid composition of algorithm functions into PROMELA processes
- Discrete-event simulation of message flows & process states and explicit causality tracking (use randomly-generated events to trigger the simulation)
- Consideration of other probabilistic modeling languages/tools: such as PRISM (Oxford Univ), STORM (Tech Univ. of Aachen), FSQ (Carnegie-Mellon Univ), and PAT (Natl. Univ. of Singapore)

Sample PROMELA-modeling of store supply-chain

PROMELA-proc customer(x)

```
spont(x) → send(purchase rqst, ch_shelf)
receive(x, ch_shelf) → {if (x == product)
    my_state := HAPPY;
    pay at check-out counter;
if (x == OUT_OF_STOCK)
    my_state := ¬ HAPPY;
    send(COMPLAINT, ch_shelf);}
```

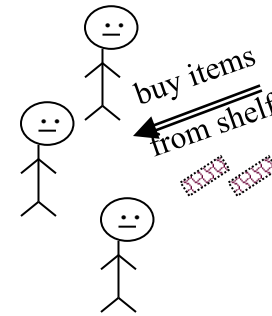
:

PROMELA-proc store_shelf()

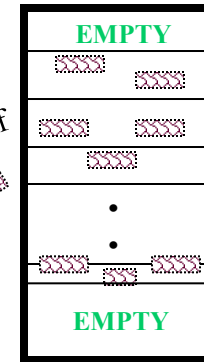
```
receive(supply, ch_truck) → {shelf_full →
    send(NO_SPACE, ch_truck);
¬shelf_full →
    if (inventory + supply ≥ MAX_SIZE)
        shelf_full := true;
        inventory := MAX_SIZE;
    else
        inventory += supply;}
```

:

customers

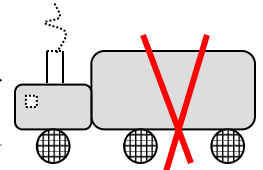


store shelf



MAX_SIZE

truck getting into accident



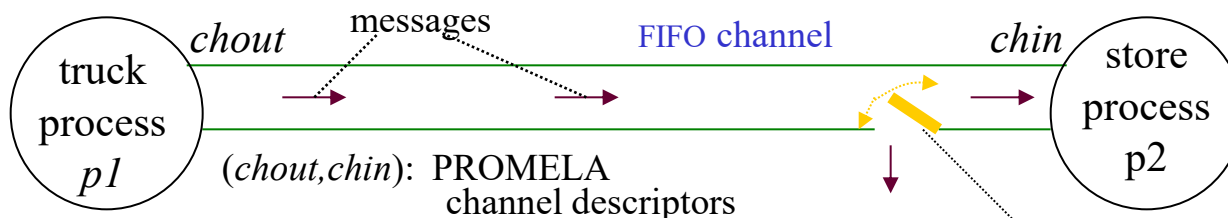
supply truck

How do we simulate this
“accident” in computer model ??

How about a back-up supply truck ??

PROMELA’s built-in language constructs

1. “message-channel” with non-lossy FIFO behavior;
2. Guarded statements, with non-deterministic statement selection.



“non-deterministic leak” of messages
(to simulate “road accidents”)

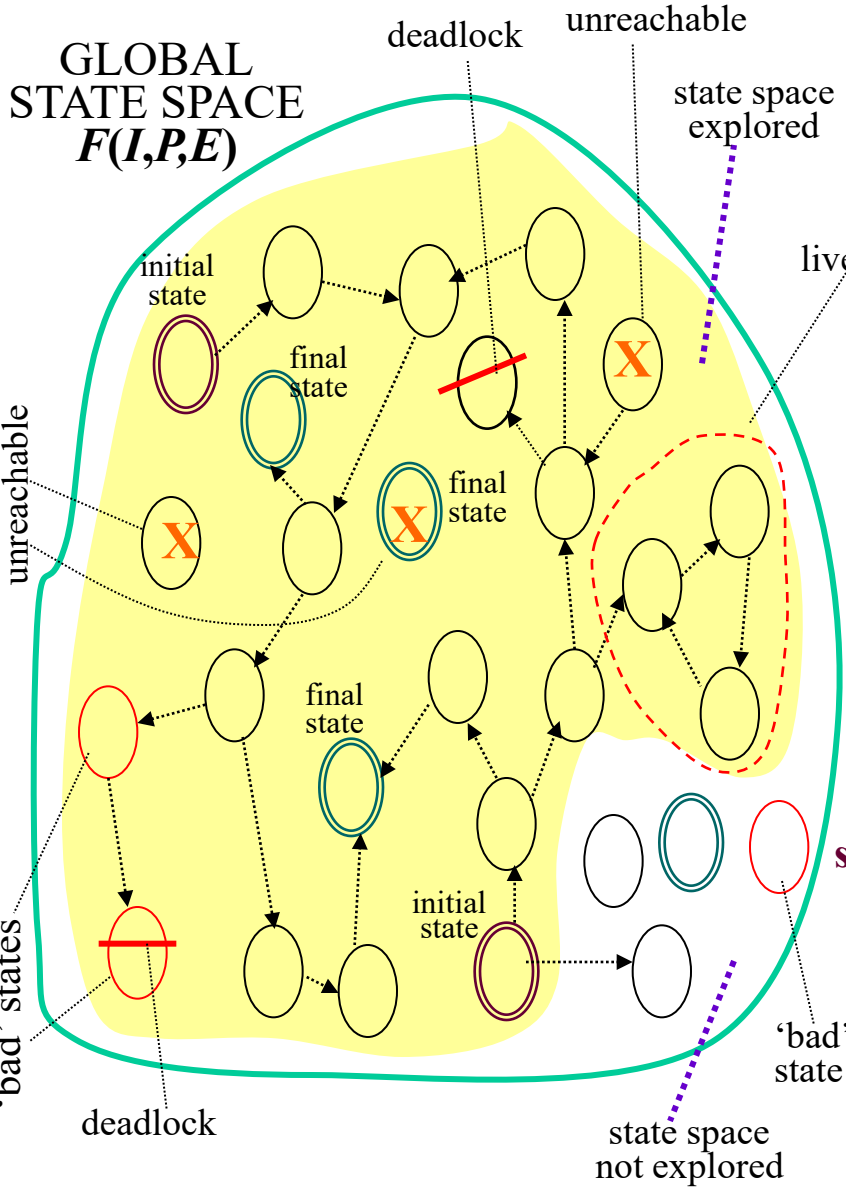
```
send(x, chout) → {receive(x, chin);
    pass on x to process p2;};
```

When the guard “send(x, chout)” evaluates to TRUE,
the statement blocks to execute is non-deterministically
selected by the PROMELA’s run-time system SPIN

spontaneous trigger
to send message x

```
send(x, chout) → receive(x, chin);
spont(x) → send(x, chout);
```

standard PROMELA based
automated algorithm analysis & synthesis



coverage = 85.7%
quality of analysis = 89%

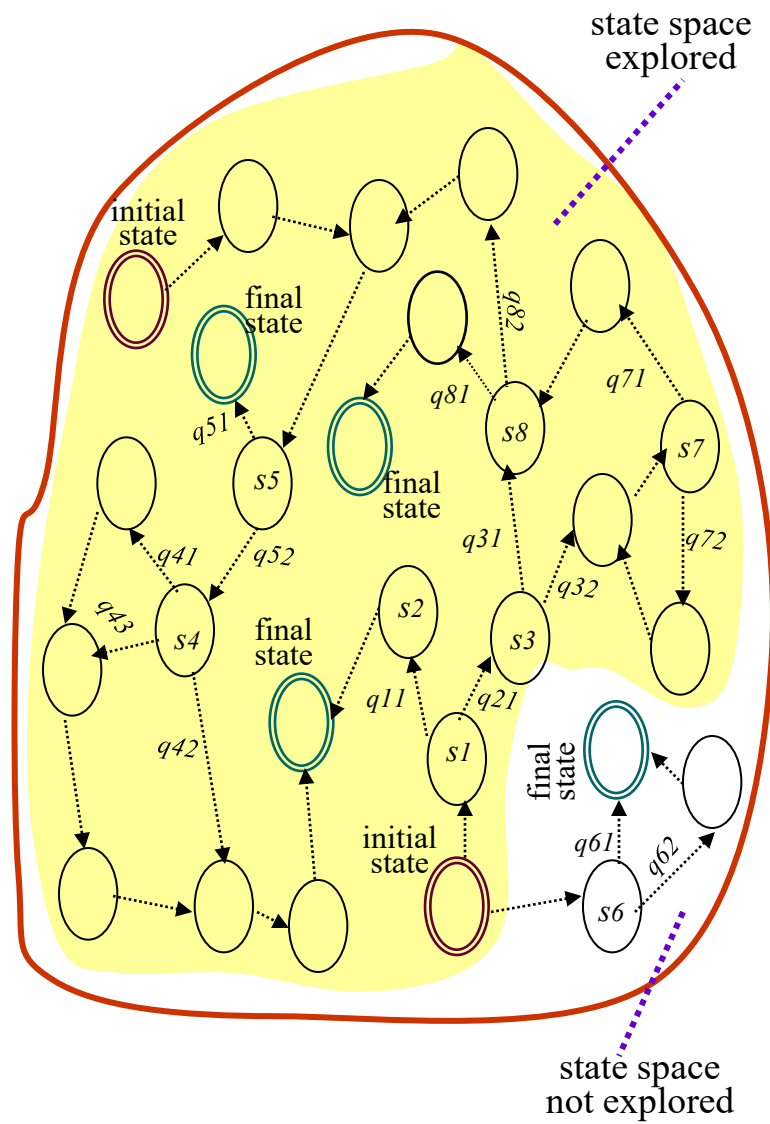
.....> state traversal

Our extended-PROMELA with
probabilistic control of state-transitions

[similar in scope to
PRISM (Oxford), PAT (NUS), . .]

First step
fix FSM
design errors

Second step
convert FSM
to Markovian
state-transition
system



q..: state-transition probability

Case-study of terrain surveillance system **by multi-drones with on-board cameras**

Accuracy of object detection depends on
[camera angle θ , distance r]
relative to the reference point

Represented as **confusion probability** Q^K
that depicts camera's detection capability
over a set of K objects ($K \geq 2$)
[**confusion probability** Q^K
varies with (r, θ)]

Algorithms to analyze UAV camera data for data-fusion
[distributed consensus by voting]

Model-based simulation:
UAV dynamics, resource needs, ...
[battery energy, network bandwidth]
RoS / Gazebo based simulator software

Incentive-based models of sensor software attacks / faults
[use of Economic Theory principles]

Event occurrences posted by software modules

Discrete-event simulation of composite system behavior under various action plans and control decisions
[Model-checker software tools (e.g., PRISM) and/or operational analysis tools]

UAV camera control
[resolution, orientation, ...]

UAV position control

UAV deployment control
[add new UAVs, remove current UAVs, reposition UAVs, ...]

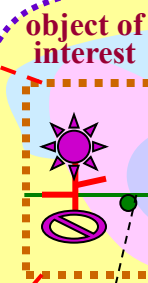
Software-cybernetics interface

Sample images of similar-looking objects in the sky:
plane, bird, baloon, helicopter, ...



Edge-detection computed by object-recognition algorithm

[distance, view-angle] vs coverage quality
fuzzy-data from camera images



spatial reference point in object

coverage map3

coverage map2

coverage map1

veh-3

$r3$

$\theta3$

coverage map2

$r2$

$\theta2$

coverage map1

$r1$

$\theta1$

veh-2

veh-1

malicious

Drones have electronically steered camera-beams

Geographic terrain monitored by drones

CYBER-FENCE

With some probability, a plane may be mis-detected as a bird (or vice-versa) by the underlying "data/image comparison" algorithm
→ **false-positives and false-negatives**

Consensus algorithm as building-block for replicated data-sources

[software prototype runs in CCNY lab (Ack: Support from US Air Force, 2015-18)]

Functional replication of data collecting processes and/or data paths is the key to mask failures, with an appropriate mechanism to manage the replication:

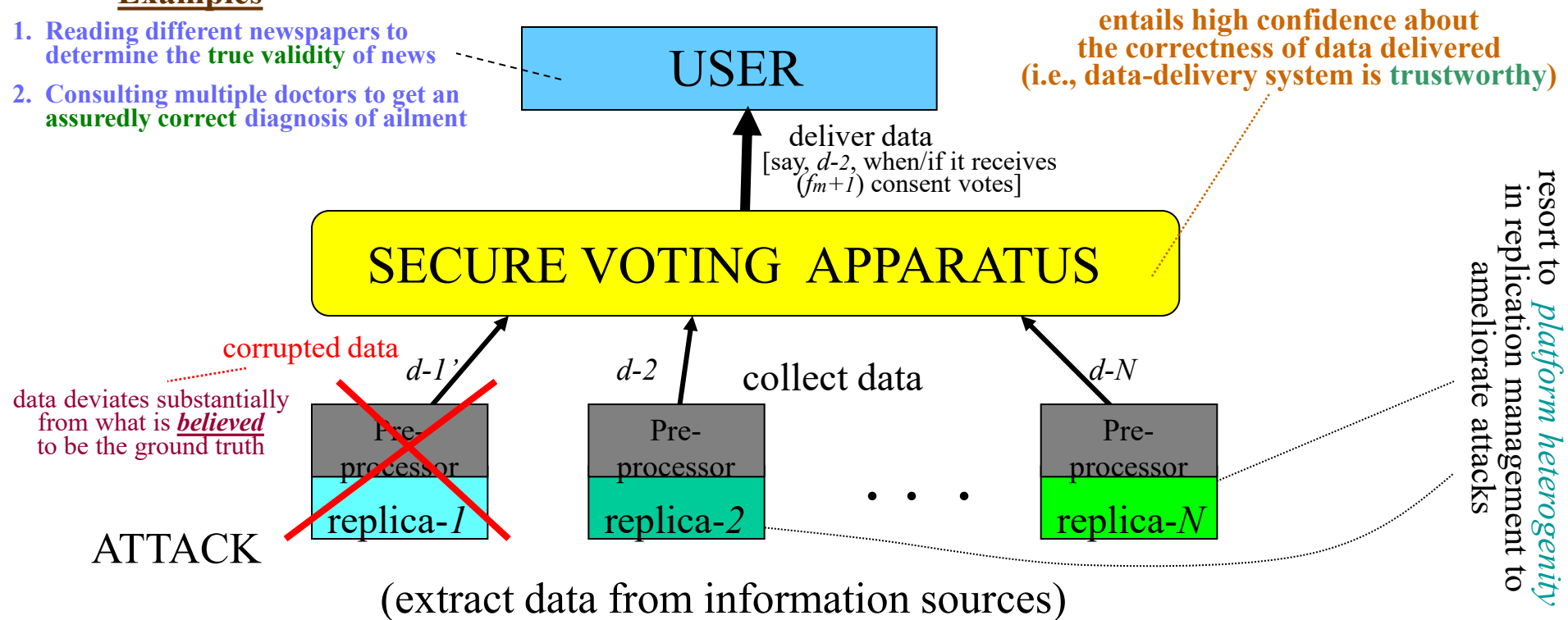
MAJORITY VOTING

$[f_m$: max. # of faulty replicas; $1 \leq f_m < \lceil N/2 \rceil$

f_m : # of server/device failures that the information system is designed to tolerate
[this is an assumption built into the algorithm about the extent of fault occurrences]

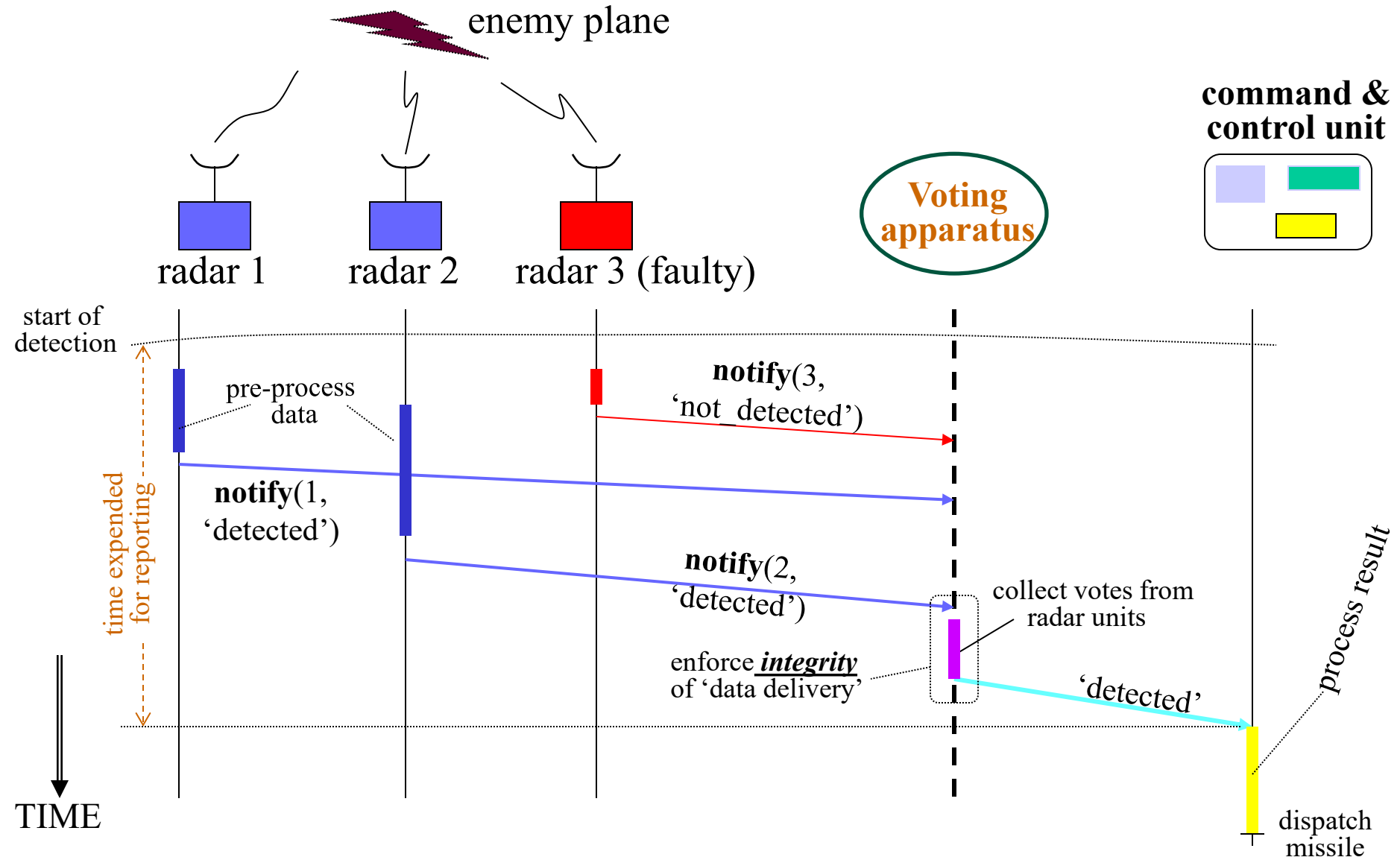
Examples

1. Reading different newspapers to determine the **true validity** of news
2. Consulting multiple doctors to get an **assuredly correct** diagnosis of ailment



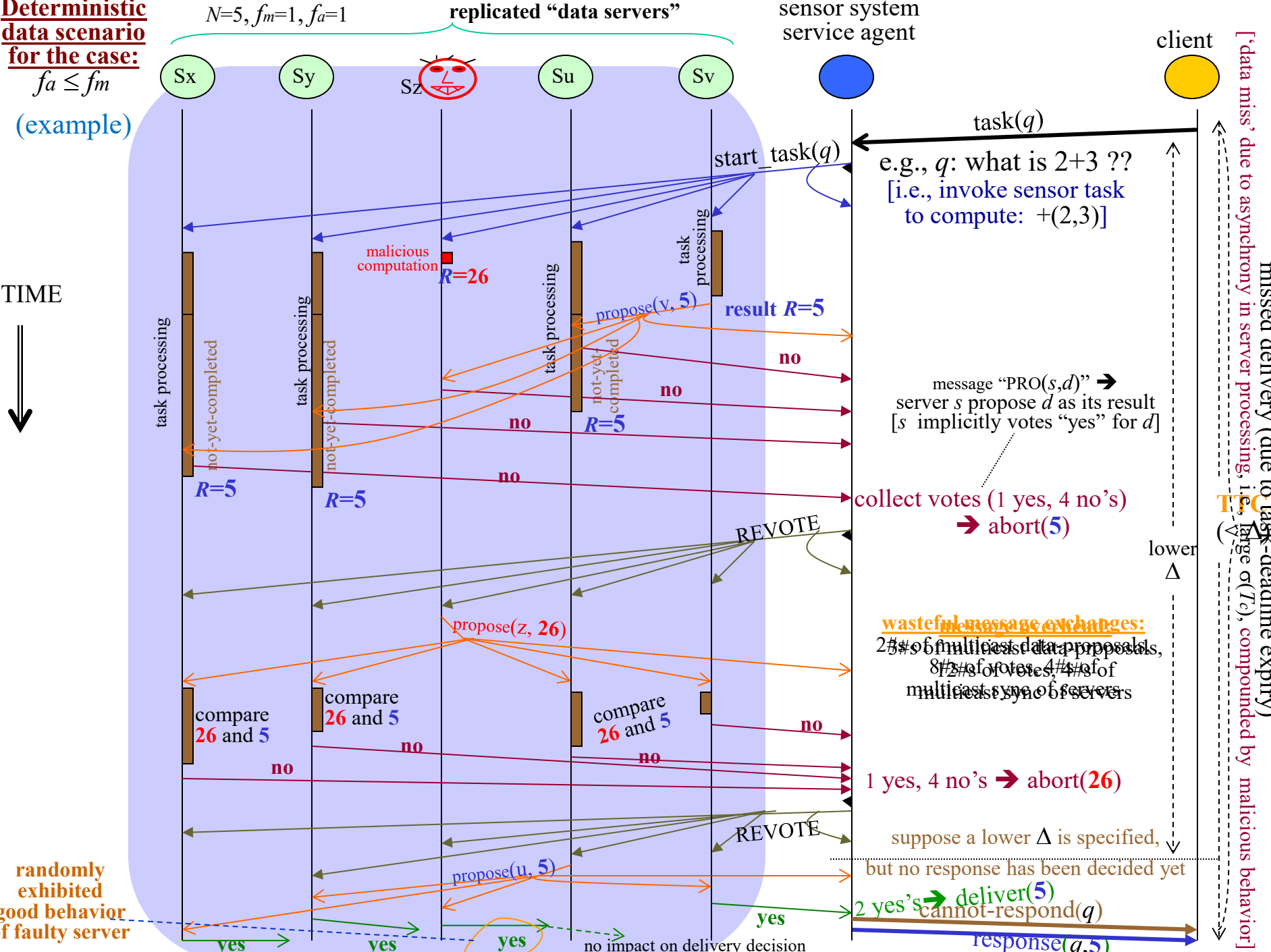
A sample application setting

[to detect enemy planes]



Deterministic data scenario for the case:
 $f_a \leq f_m$
 (example)

TIME
 ↓



Deterministic scenario for a case:
 $f_a > f_m$

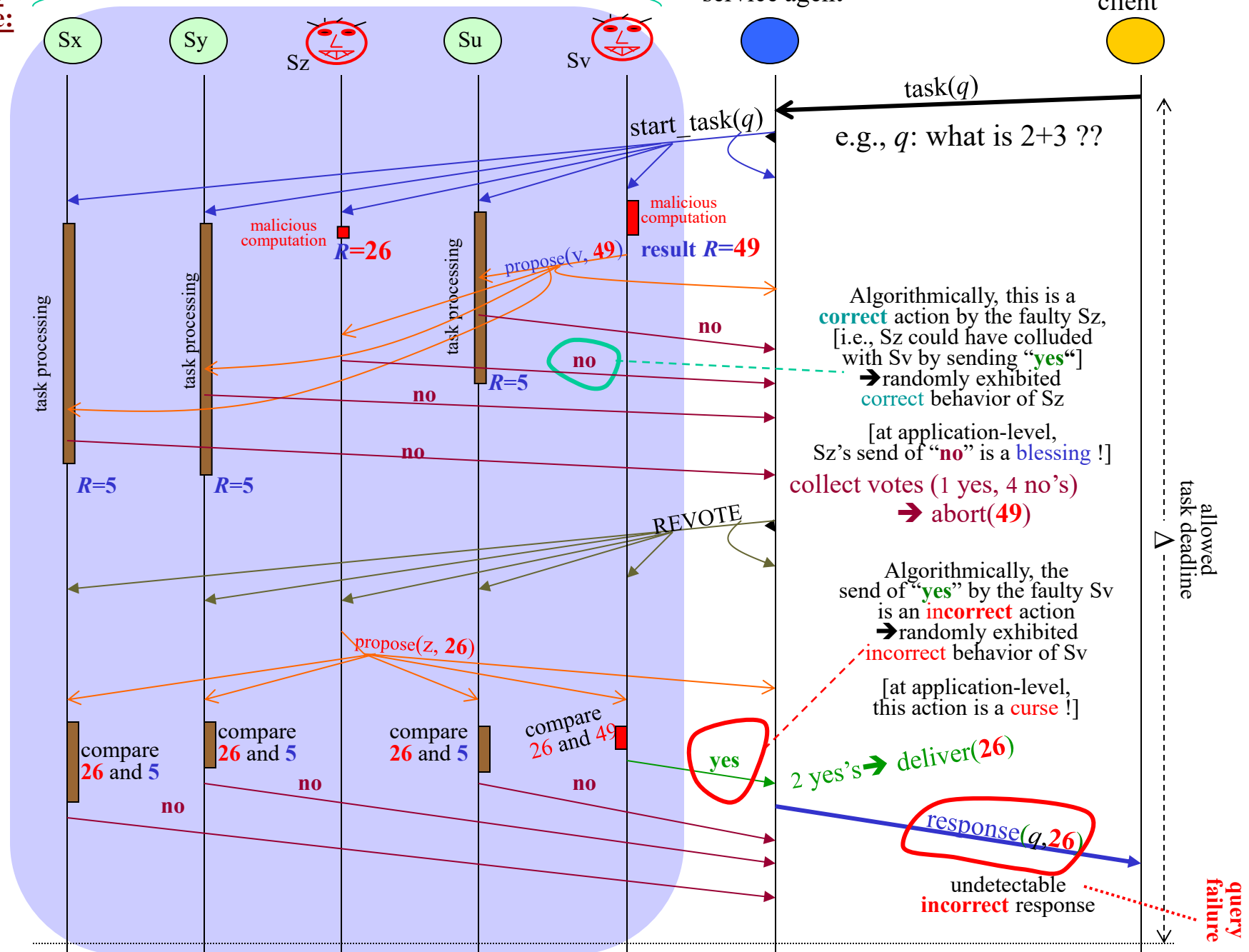
$N=5, f_m=1, f_a=2$

replicated "data servers"

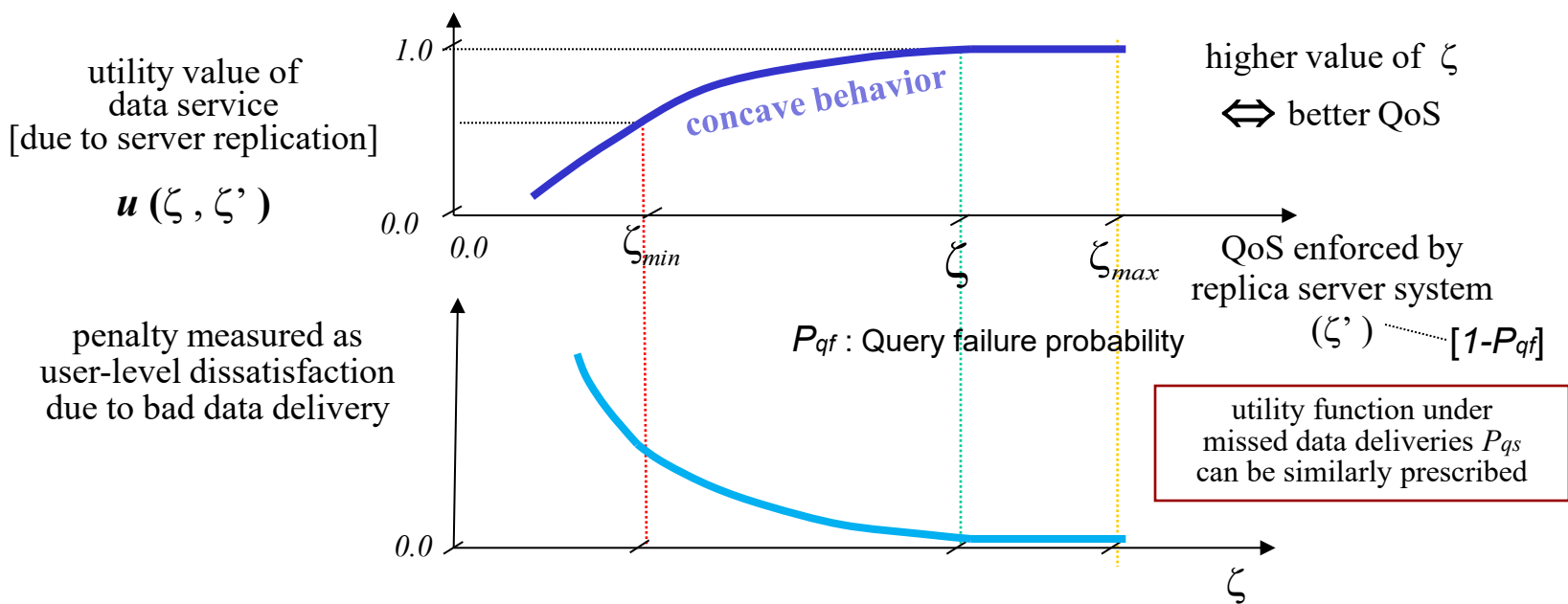
sensor system service agent

client

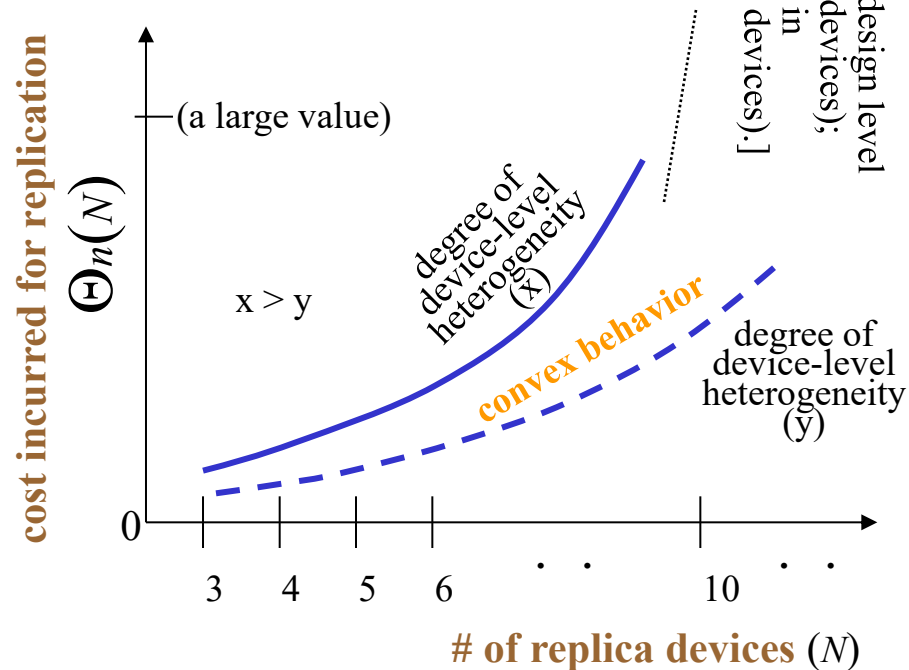
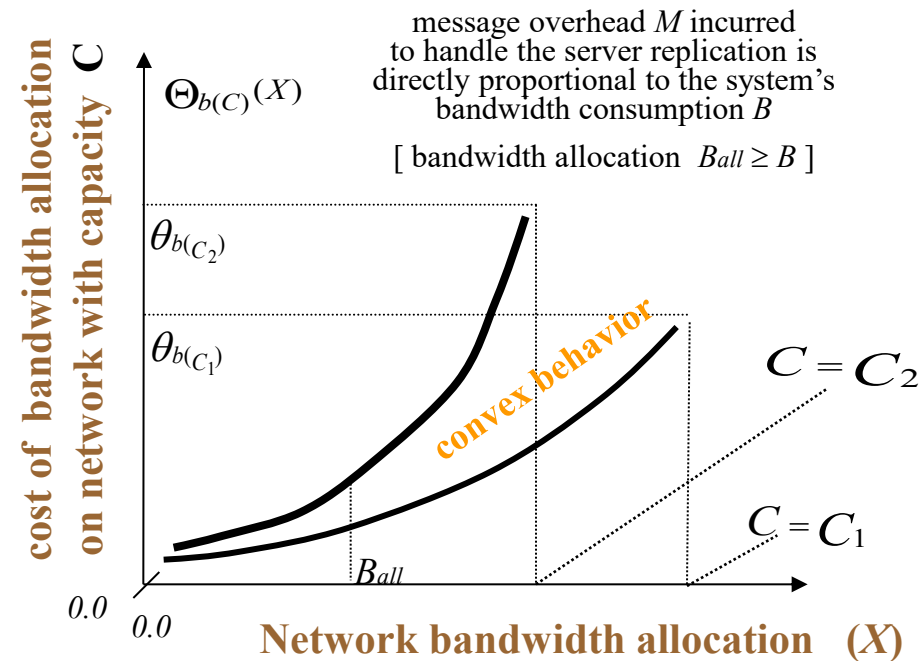
TIME



Cost vs QoS tradeoff in a replicated data-server system



- ['cost' is determined by the amount of:
1. computational efforts expended at system design level (to replicate devices);
 2. deployment/maintenance efforts expended in physical world (to install multiple devices).]

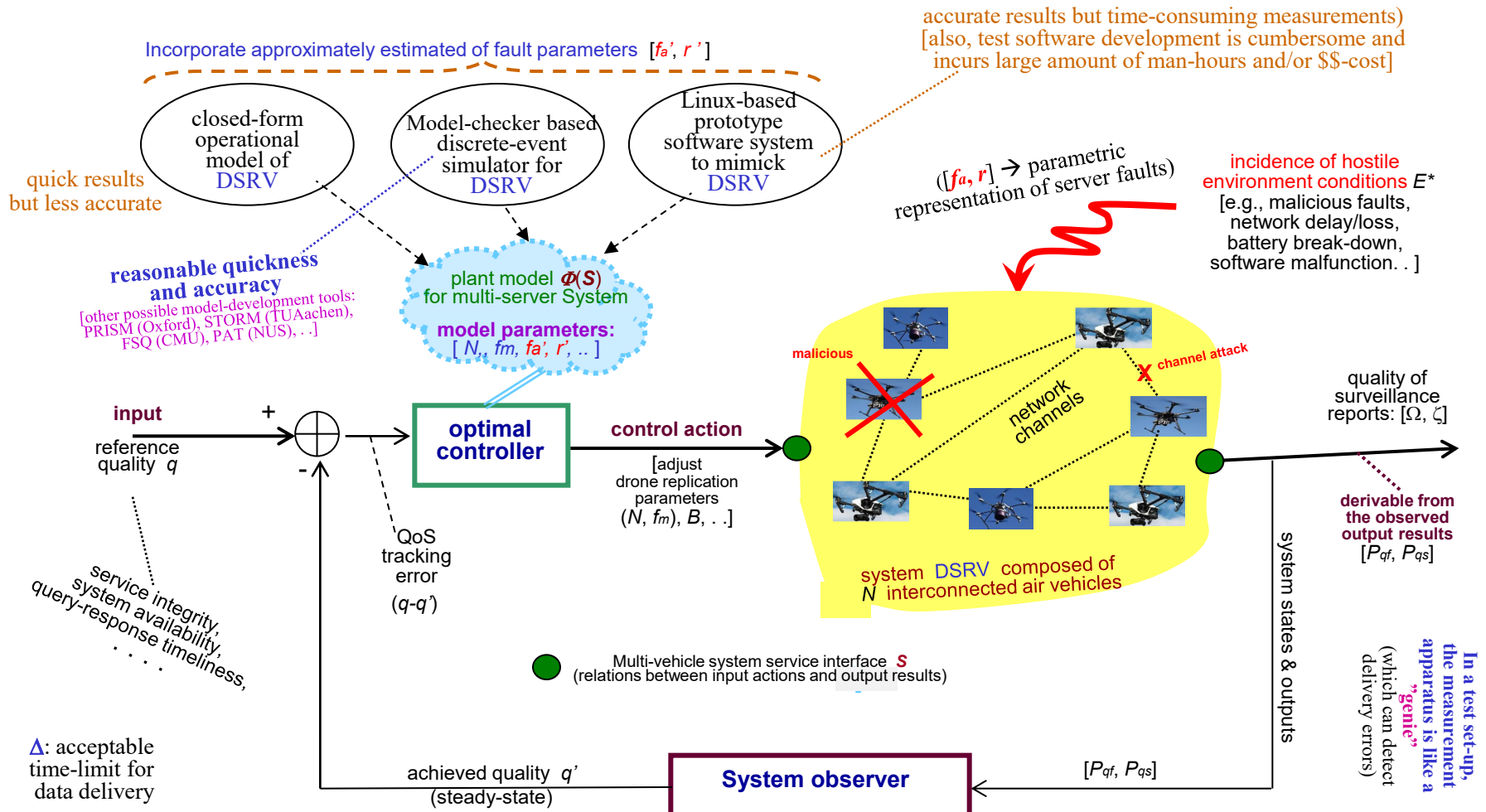


System modeling for: autonomic data-server replication management

Control schema of multi-server query processing system (use AI & ML tools for risk analysis)

Controller solves a constrained optimization problem (to decide on the **control actions** on DSRV system):

Minimize penalty arising from data-service “integrity failures”, timeliness violations, and availability degradations --- **subject to** bandwidth and battery-energy consumption and replication costs



Modeling Approach-I:

**Employ Operational Models of the
Networked System under study**

How does an operational model of replica-server system look like ?

system-level **trust** metric: Ω
 (captures the frequency of service integrity violation)
 [derivable from the QoS parameter $(1-P_{qf})$]

$[f_a, r]$ are the parameters to be inferred from the observed system output Ω

Query failure probability P_{qf}

$$P_{qf} = \sum_{j=f_m+1}^{f_a} r^j \times (1-r)^{f_a-j} \times 1 - \frac{\binom{N-j}{f_m+1}}{\binom{N}{f_m+1}} \quad (1)$$

$$\text{for } f_m < f_a \leq (N - (f_m + 1)) \quad (2)$$

$$\sum_{j=f_m+1}^{N-(f_m+1)} r^j \times (1-r)^{f_a-j} \times 1 - \frac{\binom{N-j}{f_m+1}}{\binom{N}{f_m+1}} + \quad (3)$$

$$\sum_{i=(N-f_m)}^{f_a} r^i \times (1-r)^{f_a-i} \quad (4)$$

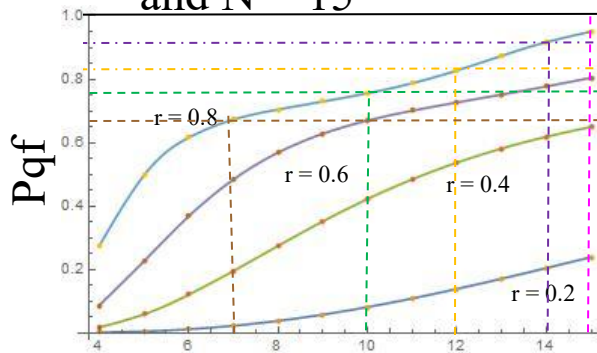
$$\text{for } f_a > (N - (f_m + 1)). \quad (5)$$

Probability Query Failure vs Number of Faulty servers

N: # of server replicas;
 r: failure-severity of a faulty server
 (i.e., probability that a faulty server exhibits an actual failure:
 $0.0 < r \leq 1.0$)

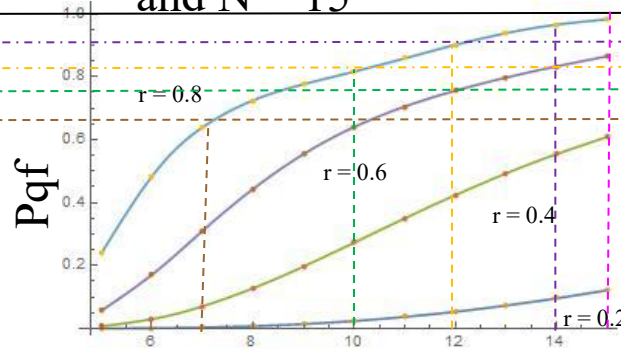
Max. # of faulty servers assumed in the algorithm

Pqf while **fm** = 3
 and N = 15



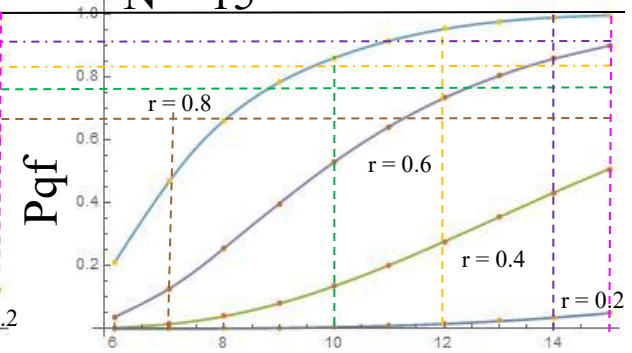
Number of actual Faulty servers (fa)

Pqf while **fm** = 4
 and N = 15



Number of actual Faulty servers (fa)

Pqf while **fm** = 5 and
 N = 15



Number of actual Faulty servers (fa)

Sample set of experimental results, as validated by mathematical model*

Actual $[N, fm]$ unknown to test designer

Claimed: $[N, fm] = [9, 3]$

Observed L' : 4.0

Model: threshold: 0.155

$[fa, r]$ known to test designer

$[fa=2, r=0.5]$		
N	fm	L'
6	2	3.27702
7	2	2.95272
8	2	2.70147
	3	4.13931
9	2	2.50412
	3	3.83729
10	2	2.34707
	3	3.57997
	4	5.09098

	Actual	Model: threshold: 0.14	
	Pqf	Pqf	Diff
$[N=9, fm=3, fa'=4, r'=0.8]$	0.25336	0.39334	0.13998
$[N=8, fm=3, fa'=4, r'=0.8]$	0.2615	0.40374	0.14224

Step 2: Compute Pqf for the two cases to disambiguate between them;

Test designer introduces known faults $[fa', r']$ to a test query in the actual system

Step 1: Explore $[N, fm]$ parameter space and compute L' ;

Identify L' that is (are) within the threshold of the observed L'

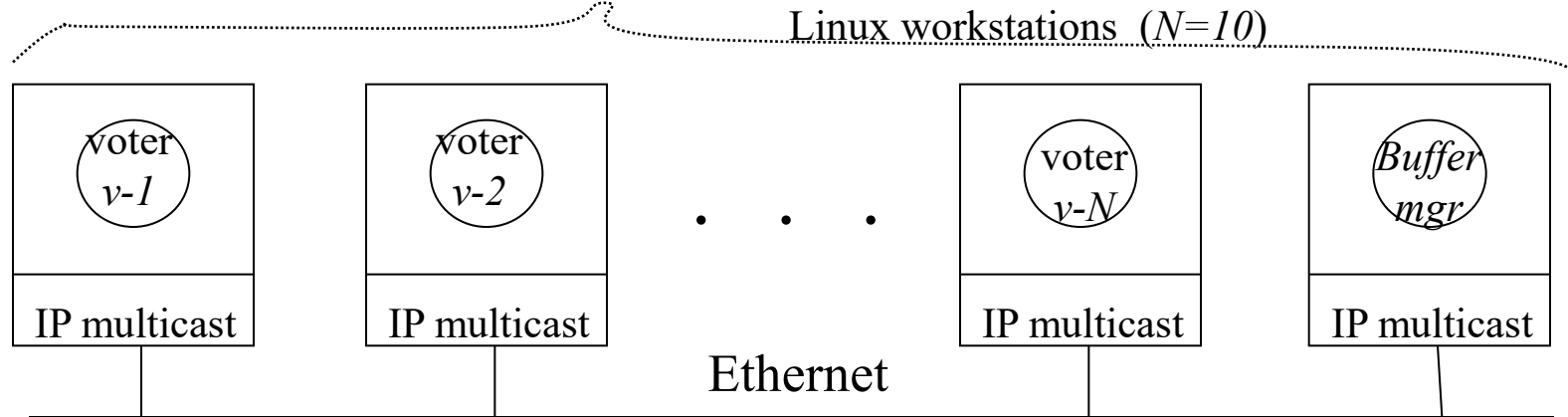
- IEEE INFCOM '26 (Ravi and etal)
- IEEE/IFIPAnnet'19 (Arun, Ravi)

Modeling Approach-II:

**Employ Software Prototypes of
Networked System under study**

Prototype implementation of replica-server software system at CUNY

[completed by a team of 4 Ph.D. and 3 M.S. students, over a 2-year period]



Scope of prototyping efforts

** Functionality checks and protocol validation by ‘stress testing’

Subject the protocol to rigorous failure scenarios to ascertain its ‘correct data delivery’ capability and identify any design errors therein

** Performability checks on protocol

Determine the maximum achievable performance by testing the protocol under ‘normal case’ behaviors in the environment (i.e., message loss is very low, $\sigma(Tc)$ is small, and $fm \ll N/2$)

^^ Experiments with random injection of faulty voter behaviors, control of voter asynchrony by random spread out of the ‘data ready’ time of servers

^^ IP multicast transport, for the purpose of performability studies

Sample experimental results on replica system performance

$N=10$;

of consent votes needed = 6;

$\sigma(T_c)=50\text{ mec}$; $\mu(T_c)=50\text{ msec}$;

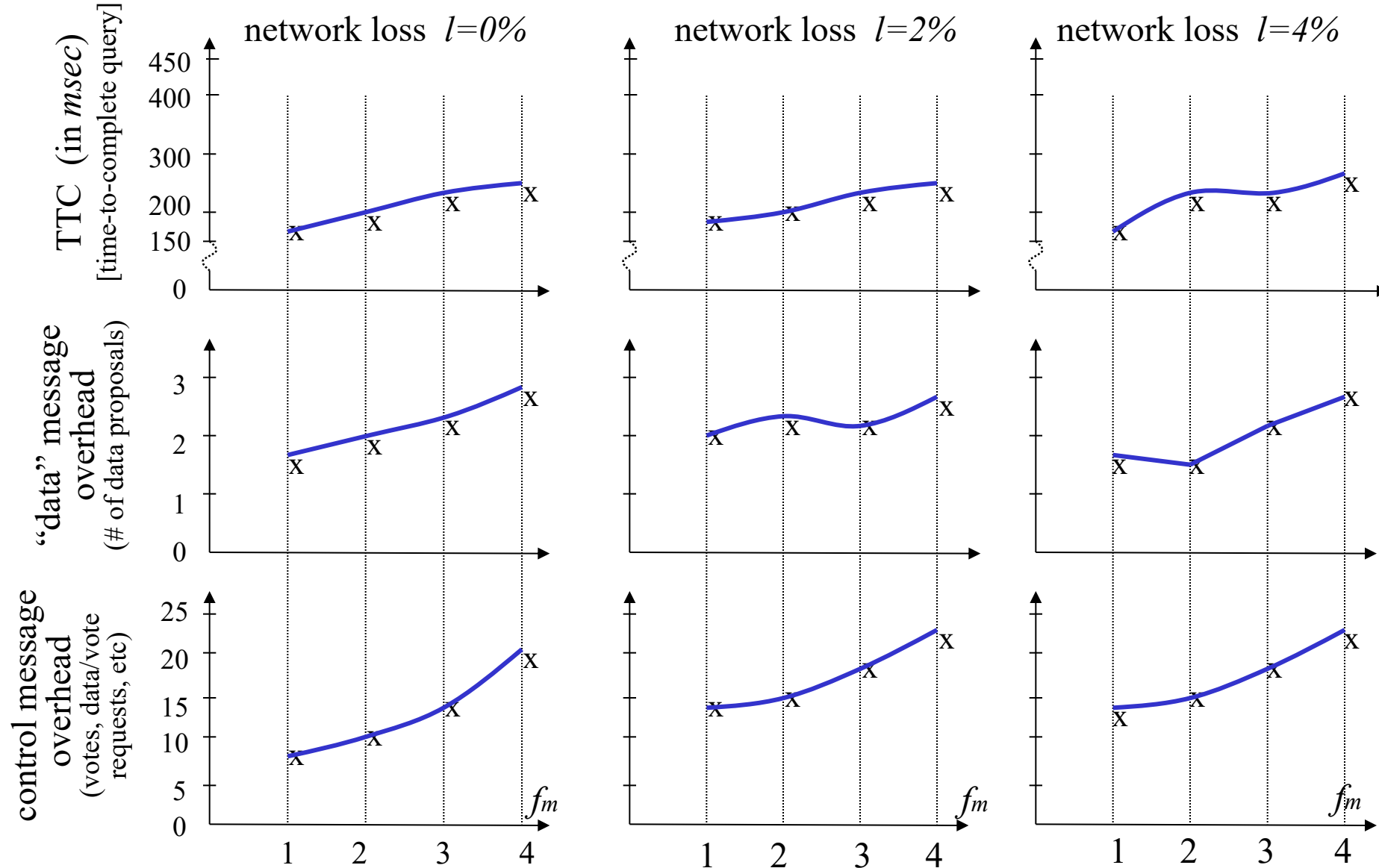
data size = 30 kbytes

control message size = 50 bytes

f_m : algorithmic assumption on # of faulty servers

f_a : actual # of faulty servers

$f_a = f_m$ in our experiments



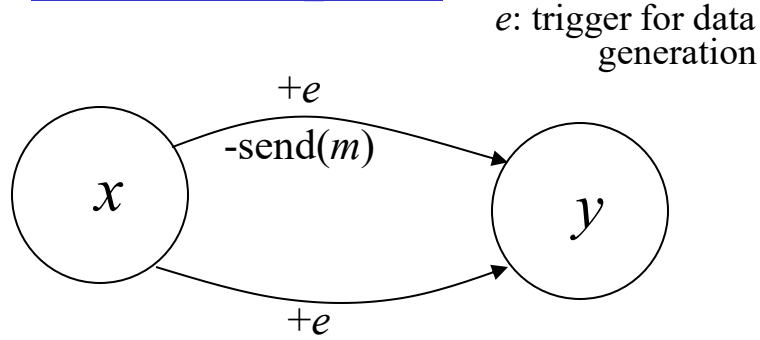
Modeling Approach-III:

**Employ Probabilistic model-checkers of
Networked System under study**

**[illustration of model-checking with
extended-PROMELA (with software-overlays)]**

FAULT: "SEND-OMISSION of process"

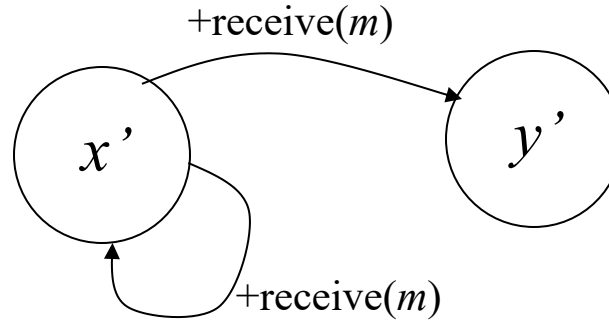
FSM of 'network sender'



state = x && event(e) → {send(m); state := y;};
 state = x && event(e) → state := y;

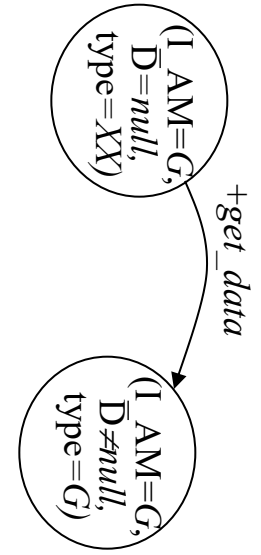
FAULT: RECEIVE-OMISSION of process

FSM of 'network receiver'

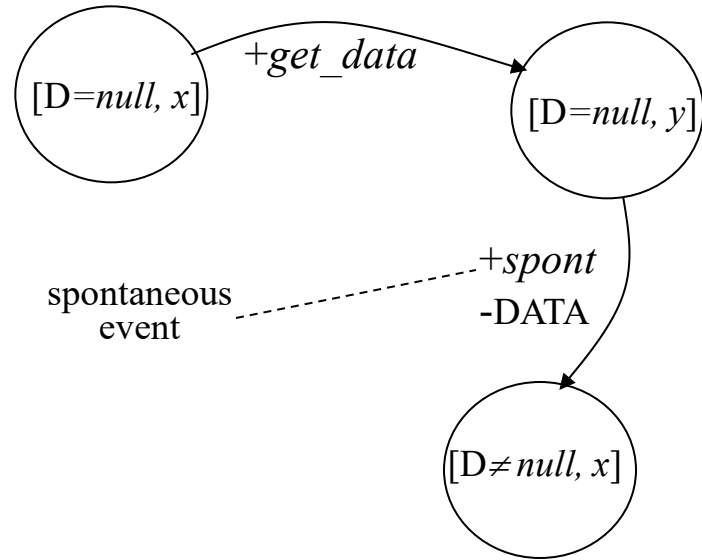


state = x' && receive(m) → state := y';
 state = x' && receive(m) → ;

FAULT: "data corruption"



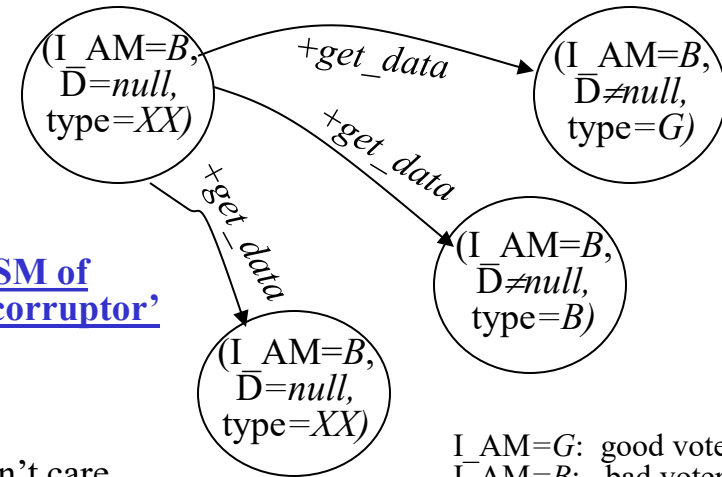
FAULT: "asynchronous data generation"



FSM of 'asynchronous data generator'

[D = null, state=x] && event(get_data) → {state := y;};
 [D = null, state=y] && event(spont) → {D ≠ null; state := x;};

FSM of 'data corruptor'



xx: don't care

I AM=G: good voter
 I AM=B: bad voter

(I_AM=xx, D = null) && event(get_data) → {D ≠ null; type := G;};
 (I_AM=B, D = null) && event(get_data) → {D ≠ null; type := B;};
 (I_AM=B, D = null) && event(get_data) → ;

Capturing faulty behaviors in PROMELA-model of replicated servers

N : # of deployed servers; f_m : max. # of servers *assumed* as faulty; f_a : # of *actual* faulty servers;
 r : faulty-severity of a faulty server; K : # of deployable servers (say, # of cloud-leased servers).
($f_m < f_a \leq N$; $0.0 < r \leq 1.0$; $N \leq K$; $1 \leq f_m < \lceil N/2 \rceil$)

Initial set up of hostile environment

```
for (i := 1; i ≤ fa; i++)  
    server_type[i] := BAD; /* mark server as faulty */  
    server_severity[i] := r;  
if (fa < N)  
    for (i := fa+1; i ≤ N; i++)  
        server_type[i] := GOOD; /* mark server as non-faulty */
```

Issue: How do we set up the probabilistic behavior of a faulty server Z (say, $0.0 < r < 1.0$) ??
[i.e., Z proposes a corrupted data with probability r , and proposes a non-corrupted data with probability $(1-r)$]

Basic set up for each server $j \mid j=1,2,\dots,N$ in standard-PROMELA

‘data trigger’ conditions for server j
are not shown in guard expressions

Non-deterministic guard
(expressed in PROMELA)

→ $r = 0.5$ is the only possible assignment

(server_type[j] == **GOOD**):
{generate data x;
 x.label := **NC**;} /* mark data as ‘not corrupted’ */

(server_type[j] == **BAD**):
{generate data x;
 x.label := **C**;} /* mark data as ‘corrupted’ */;

(server_type[j] == **BAD**):
{generate data x;
 x.label := **NC**;}

Infusion of probabilistic behaviors in replica server systems

“meta-label” set up for each server $j \mid j=1,2,\dots,N$

(for **observational** purposes)

guards, expressed in PROMELA

$(\text{server_type}[j] == \text{GOOD}) \wedge (\text{dtrig}[j] == \text{TRUE}):$

{generate data x ;
 $x.\text{label} := \text{NC}$; /* mark data as ‘not corrupted’ */
 $\text{dtrig}[j] := \text{FALSE};$ }

$(\text{server_type}[j] == \text{BAD}) \wedge (\text{dtrig}[j] == \text{TRUE}):$

{generate random # $R \in (0,1]$;

$(R \leq r):$

{generate data x ;

$x.\text{label} := \text{C}$; /* mark data as ‘corrupted’ */
 $\text{dtrig}[j] := \text{FALSE};$ }

sub-guards

$(\text{server_type}[j] == \text{BAD}) \wedge (\text{dtrig}[j] == \text{TRUE}):$

{generate random # $R \in (0,1]$;

$(r < R \leq 1.0):$

{generate data x ;

$x.\text{label} := \text{NC}$;
 $\text{dtrig}[j] := \text{FALSE};$ }

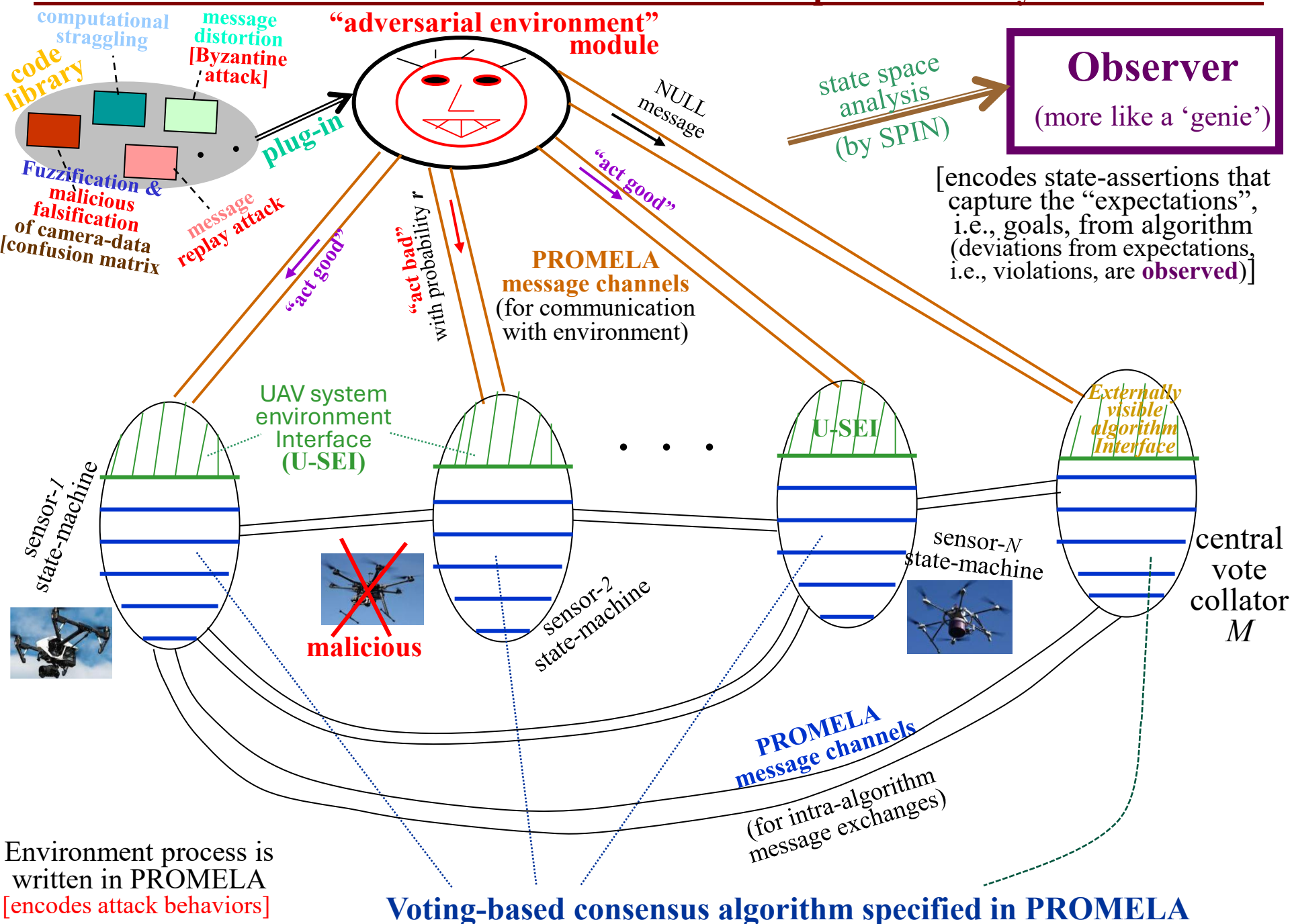
$\text{dtrig}[j]$: external trigger event for
server j to generate data
[e.g., end of server computation,
malicious action, etc]

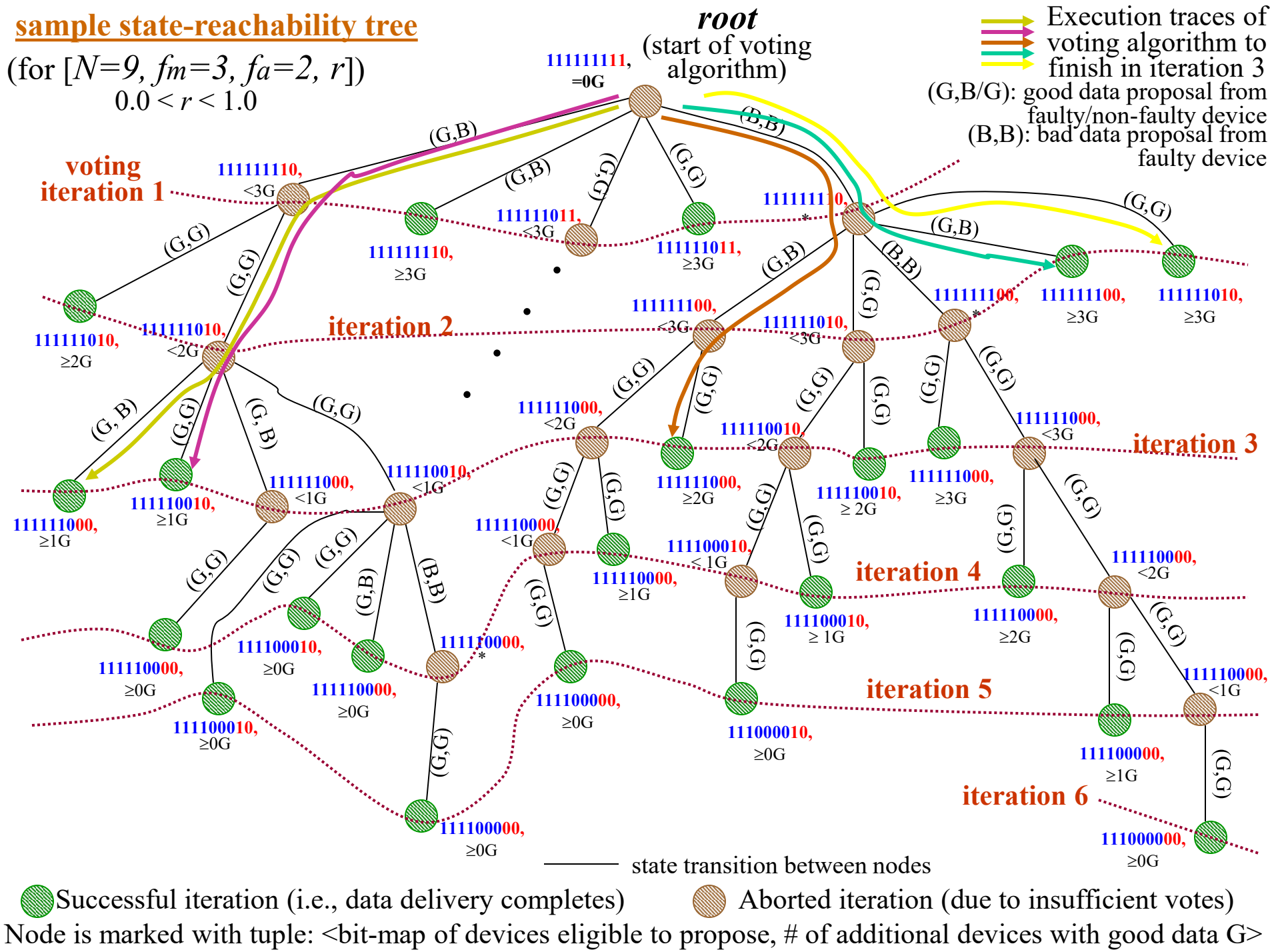
Transactional query failure probability:

$$P_{qf} = \frac{\text{\# of deliveries with NC label}}{\text{Total \# of deliveries (i.e., both NC and C)}}$$

counted by SPIN from state-traces

PROMELA-based simulation structure to assess replica-server system behaviors



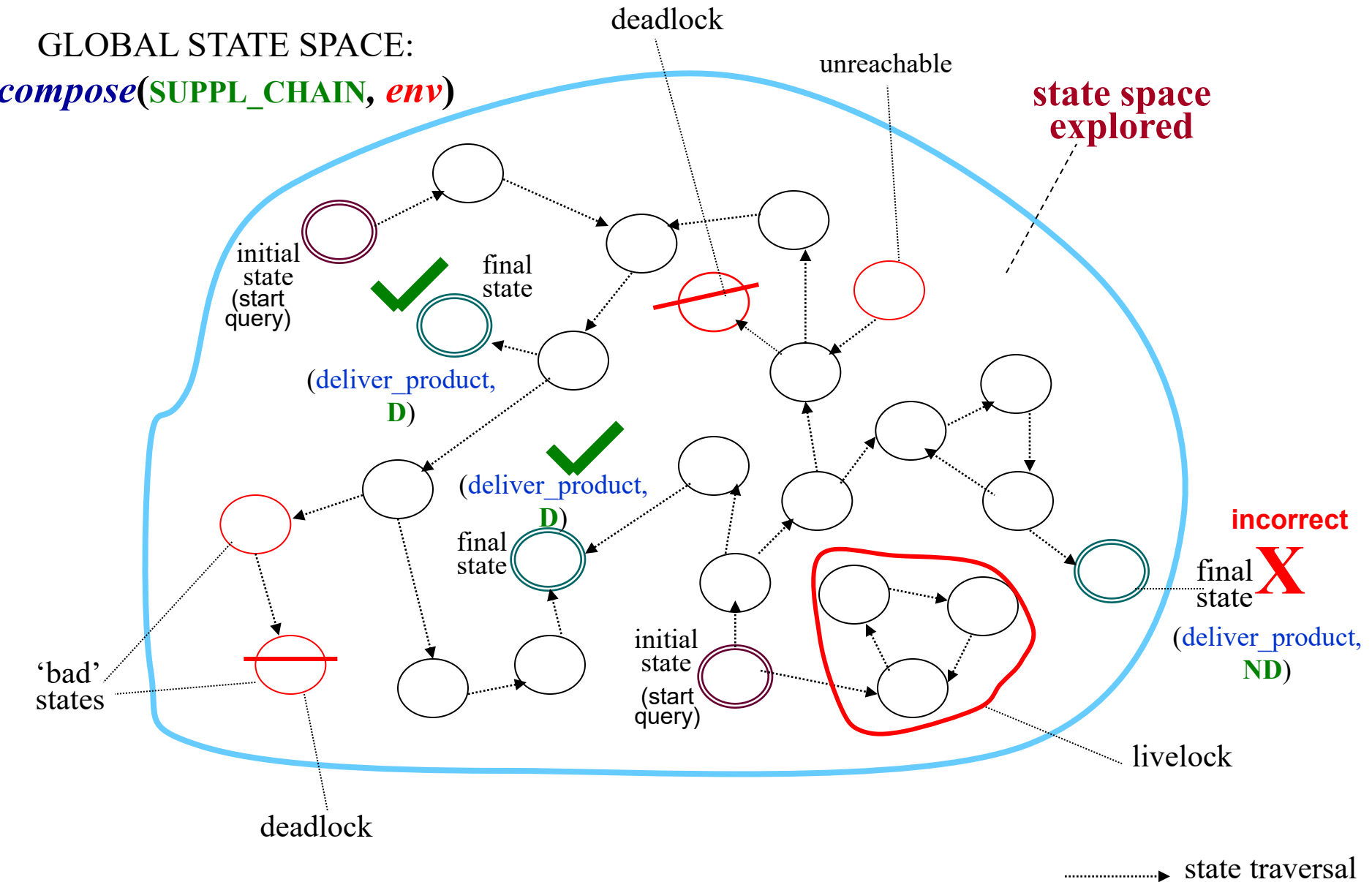


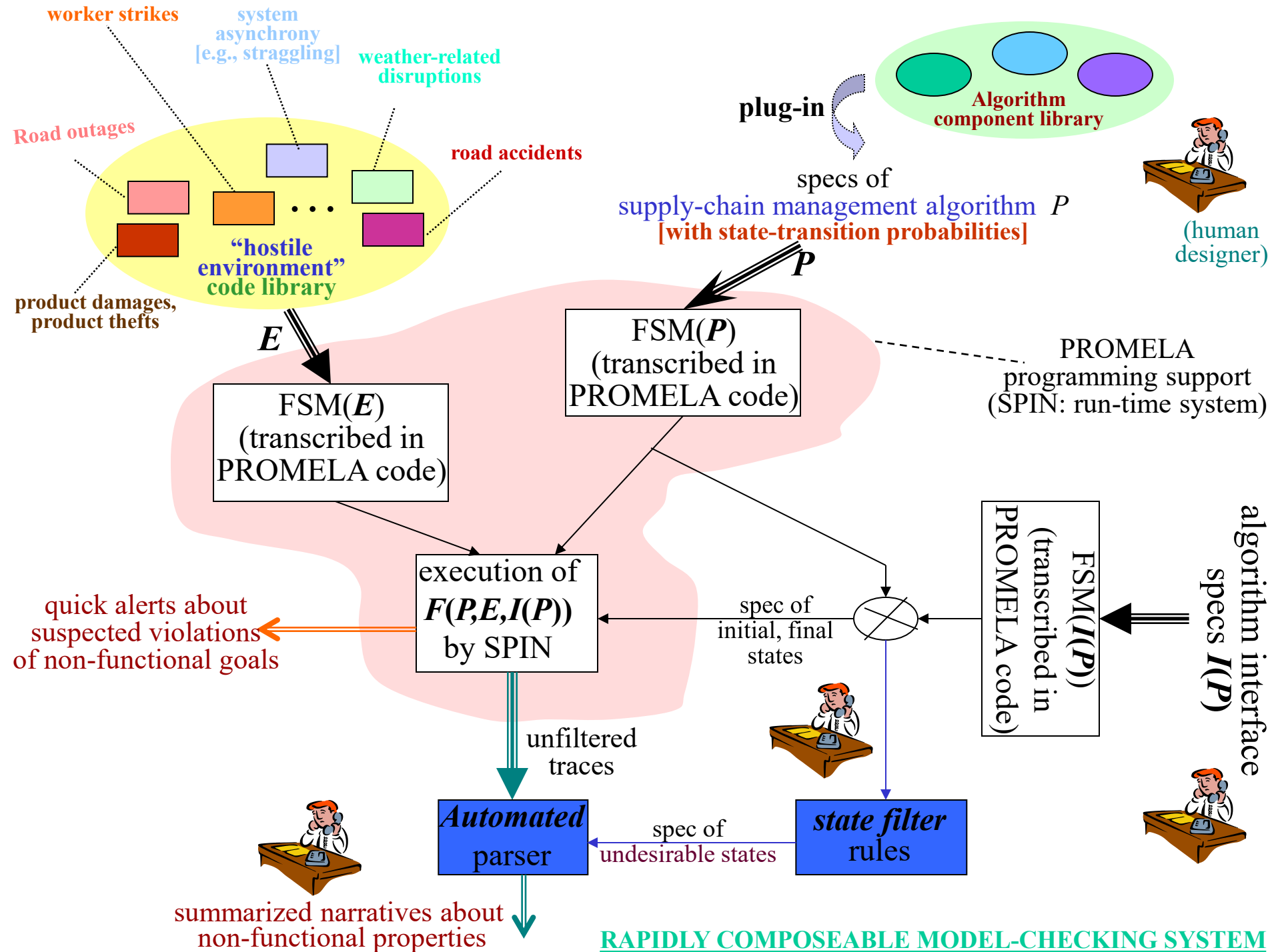
Quick ideas on probabilistic model-checking for resilient supply-chain networks

Use of meta-labels **D** and **ND** to track good and bad product deliveries during supply-chain simulated execution

GLOBAL STATE SPACE:

```
compose(SUPPL_CHAIN, env)
```





Summary of our works on probabilistic model-checking

- PROMELA software overlays to assign state-transition probabilities and SPIN trace-analysis
- Rapid composition of algorithm functions into PROMELA processes
- Exploration of other model checkers: PRISM (Oxford), PAT (NUS), STORM (Tech Univ Aachen), . . .
- Extraction of probabilistic measures of system adaptation processes
- Machine-intelligence and Markov decision tools for system analysis in the face of uncertain events
- Autonomic methods for system-level reconfigurations (such as switching between pessimistic and optimistic algorithms)