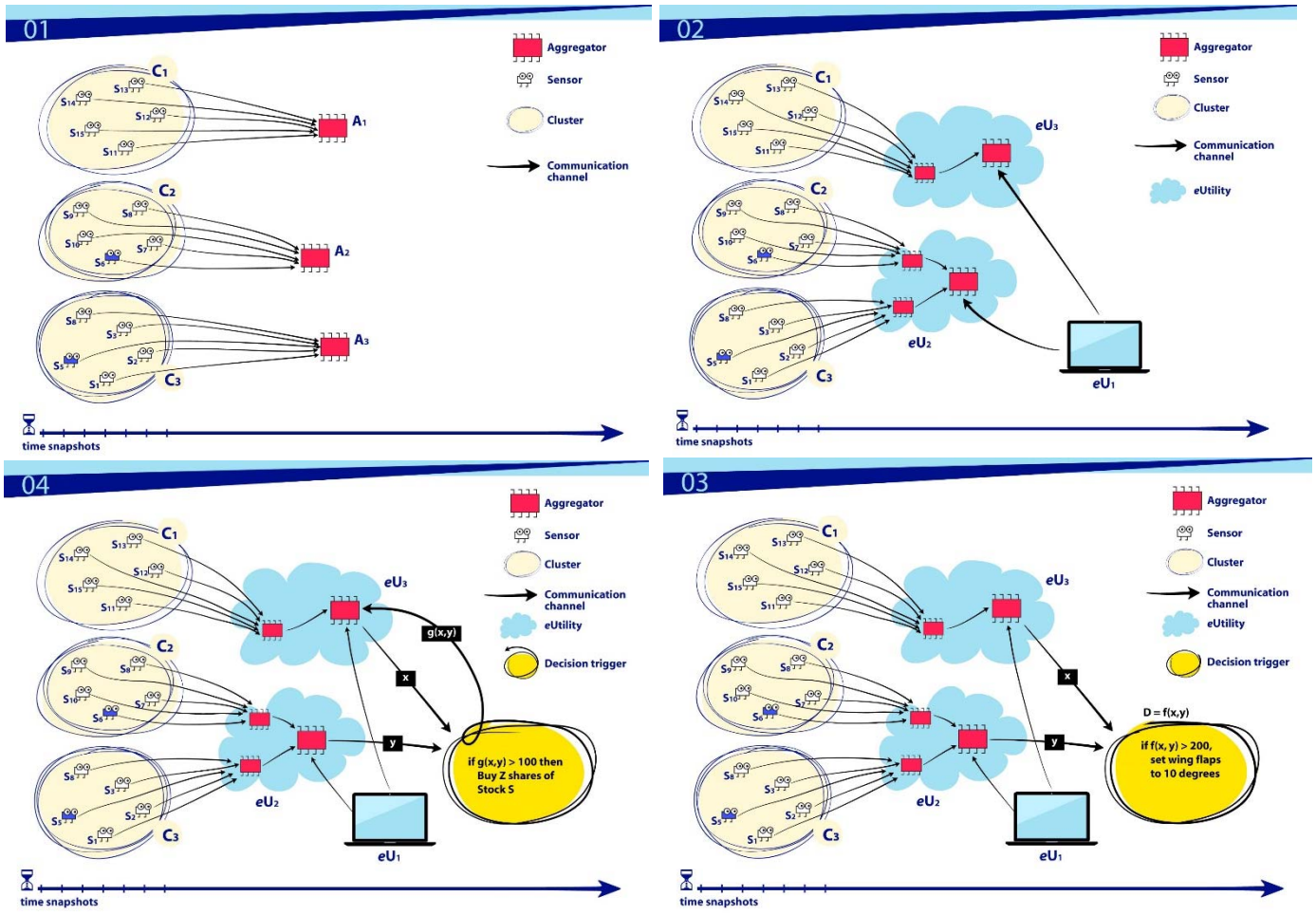
Networks of ‘Things’

(Demystifying IoT*)*



J. Voas

Computer Scientist

National Institute of Standards and Technology

18 Months Ago We Asked

What is IoT*?*

The Reality

No universally-accepted and actionable definition exists to the question, “What is IoT?”

Opening Statement

‘This technology’ employs a mixture of *sensing*, *communication*, *computation, actuation*.

We step back from the acronym IoT

Network of Things (NoT)

IoT *and* NoT

"We use two acronyms, IoT and NoT (Network of Things), extensively and interchangeably—the relationship between NoT and IoT is subtle. IoT is an instantiation of a NoT, more specifically, IoT has its 'things' tethered to the Internet. A different type of NoT could be a Local Area Network (LAN), with none of its 'things' connected to the Internet. Social media networks, sensor networks, and the Industrial Internet are all variants of NoTs. This differentiation in terminology provides ease in separating out use cases from varying vertical and quality domains (e.g., transportation, medical, financial, agricultural, safety-critical, security-critical, performance-critical, high assurance, to name a few). That is useful since there is no singular IoT, and it is meaningless to speak of comparing one IoT to another."

Draft NISTIR 8063

Primitives and Elements of Internet of Things (IoT) Trustworthiness

Jeffrey Voas
Computer Security Division
Information Technology Laboratory

February 2016

U.S. Department of Commerce
Penny Pritzker, Secretary

National Institute of Standards and Technology
Willie May, Under Secretary of Commerce for Standards and Technology and Director

DRAFT NIST Special Publication 800-183

Networks of Things

Jeffrey Voas
Computer Security Division
Information Technology Laboratory

This publication is available free of charge from:
http://dx.doi.org/10.6028/NIST.SP.800-183

June 2016

U.S. Department of Commerce
Penny Pritzker, Secretary

National Institute of Standards and Technology
Willie May, Under Secretary of Commerce for Standards and Technology and Director

Primitives

1. **Sensor** A *sensor* is an electronic utility that measures physical properties such as temperature, acceleration, weight, sound, location, presence, identity, etc. All sensors employ mechanical, electrical, chemical, optical, or other effects at an interface to a controlled process or open environment
2. **Aggregator** An *aggregator* is a software implementation based on mathematical function(s) that transforms groups of *raw* data into *intermediate, aggregated* data. Raw data can come from any source. Aggregators address 'big' data.
3. **Communication channel** A *communication channel* is a medium by which data is transmitted (e.g., physical via USB, wireless, wired, verbal, etc.).
4. ***e*Utility** An *e*Utility (external utility*)* is a software or hardware product or service.
5. **Decision trigger** A *decision* trigger creates the final result(s) needed to satisfy the purpose, specification, and requirements of a specific NoT.

For Each Primitive

Basic properties, assumptions, recommendations, and general statements about Primitive x include:

Sensor (10 of 29)

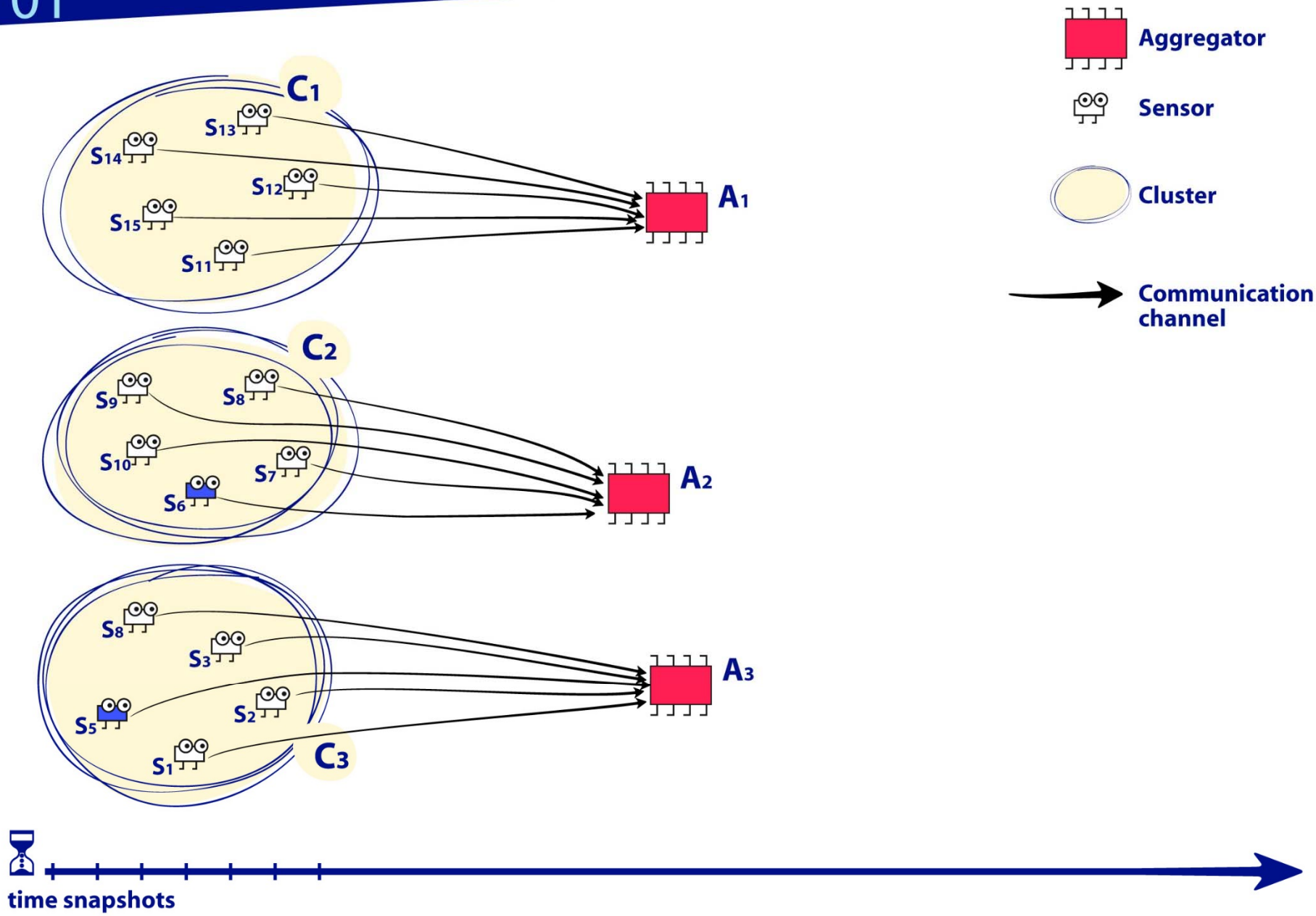
1. Sensors are physical; some may have an Internet access capability.
2. A sensor may also transmit device identification information, such as via RFID
3. Sensors may be heterogeneous, from different manufacturers, and collect data, with varying levels of data integrity.
4. Sensors may be associated with fixed geographic locations or may be mobile.
5. Sensors may have an owner(s) who will have control of the data their sensors collect, who is allowed to access it, and when.
6. Sensors will have pedigree – geographic locations of origin and manufacturers. Pedigree may be unknown, and suspect. This has ties to Supply Chain Risk Management (SCRM).
7. Sensors may be cheap, disposable, and susceptible to wear-out over time.
8. There will differentials in sensor security, safety, and reliability, e.g., between consumer grade, military grade, industrial grade, etc.
9. Sensors may return no data, totally flawed data, partially flawed data, or correct/acceptable data. Sensors may fail completely or intermittently. They may lose sensitivity or calibration.
10. Security is a concern for sensors if they or their data is tampered with, stolen, deleted, dropped, or transmitted insecurely so it can be accessed by unauthorized parties. Building security into specific sensors may or may not be cost effective.

Aggregator (6 of 11)

1. Intermediate, aggregated data may suffer from some level of *information loss*. Proper care in the aggregation process should be given to significant digits, rounding, averaging, and other arithmetic operations to avoid unnecessary loss of precision.
2. Aggregators are: (1) executed at a specific time and for a fixed time interval, or (2) event-driven.
3. Aggregators may be acquired off-the-shelf. Note that aggregators may be non-existent and will need to be home-grown. This may create a problem for huge volumes of data within a NoT.
4. Security is a concern for aggregators (malware or general defects) and for the sensitivity of their aggregated data. Further, aggregators could be attacked, e.g., by denying them the ability to operate/execute or by feeding them bogus data.
5. Reliability is a concern for aggregators (general defects).
6. Aggregators have two actors for consolidating large volumes of data into lesser amounts: Clusters and Weights. This is the only primitive with actors.

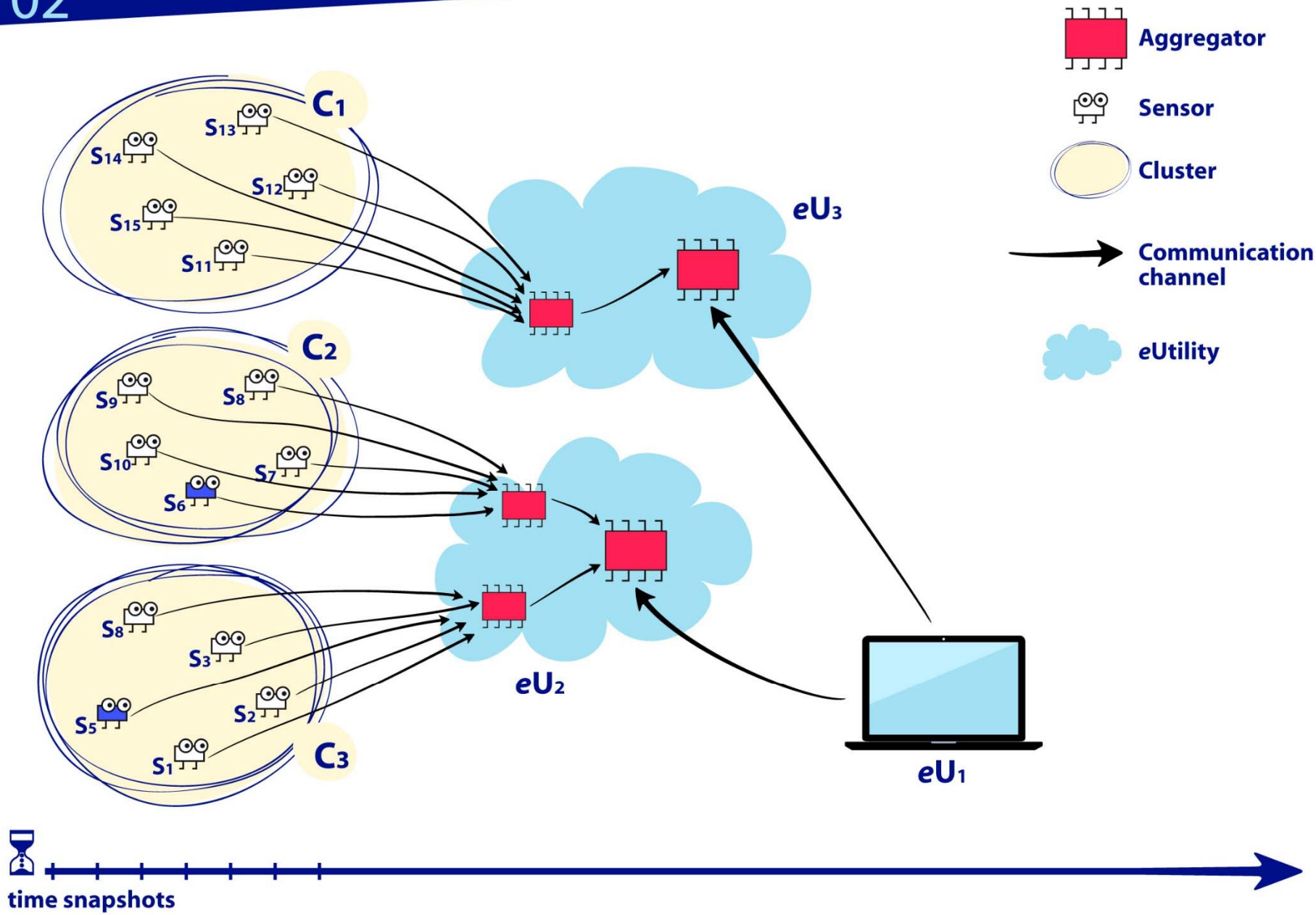
Communication Channel (7 of 12)

1. Communication channels move data between computing, sensing, and actuation.
2. Since data is the “blood” of a NoT, communication channels are the “veins” and “arteries”, as data moves to and from intermediate events at different snapshots in time.
3. Communication channels will have a physical or virtual aspect to them, or both. Protocols and associated implementations provide a virtual dimension, cables provide a physical dimension.
4. Communication channel dataflow may be unidirectional or bi-directional. There are a number of conditions where an aggregator might query more advanced sensors, or potentially recalibrate them in some way (e.g., request more observations per time interval).
5. No standardized communication channel protocol is assumed; a specific NoT may have multiple communication protocols between different entities.
6. Communication channels are prone to disturbances and interruptions.
7. *Redundancy* can improve communication channel reliability. There may be more than one distinct communication channel between a computing primitive and a sensing primitive.


Aggregator
Sensor
Cluster
Communication channel
C1
S13
S14
S12
S15
S11
A1
C2
S9
S8
S10
S7
S6
A2
S8
S3
S5
S2
S1
C3
A3
time snapshots

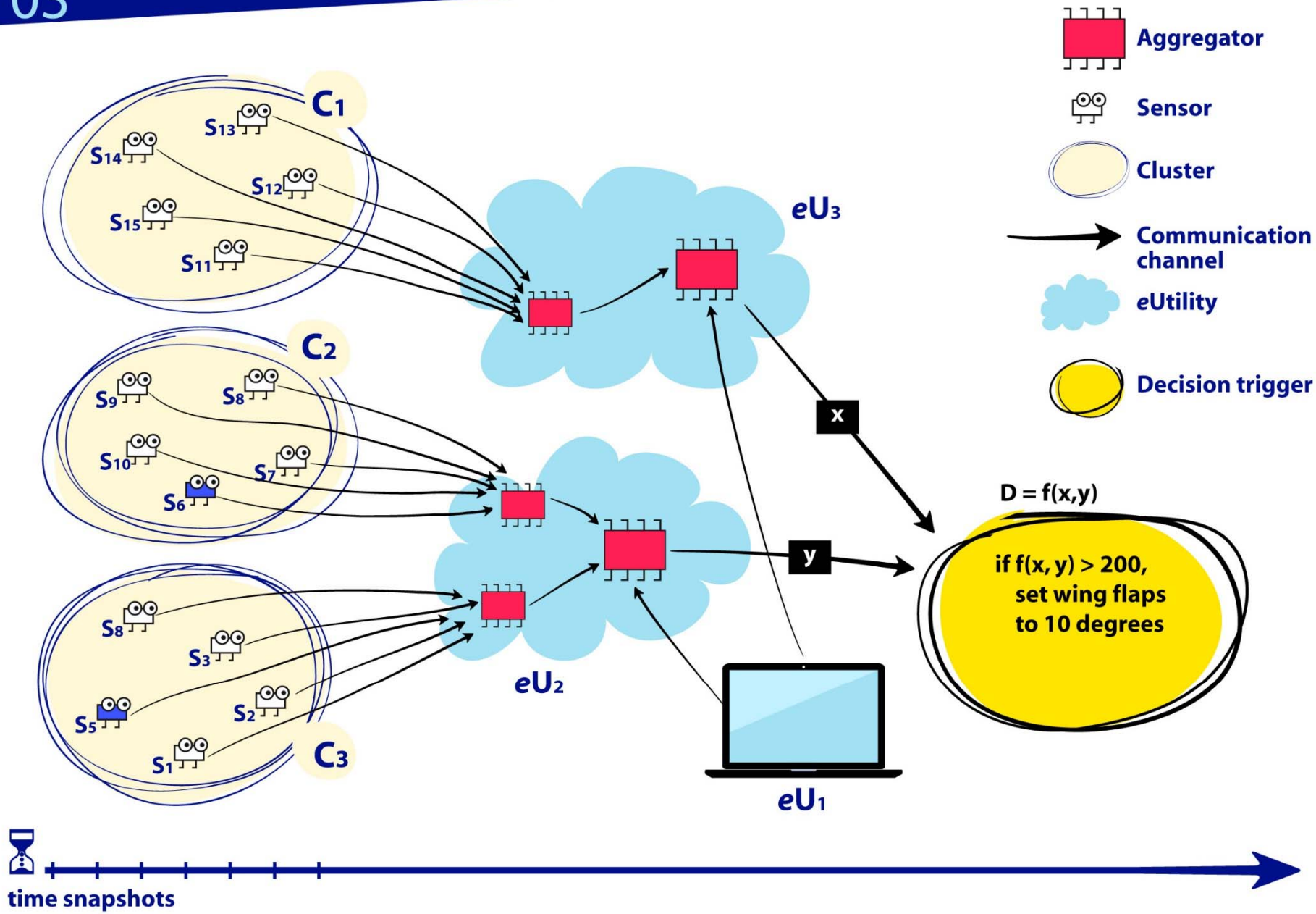
*e*Utility (5 of 9)

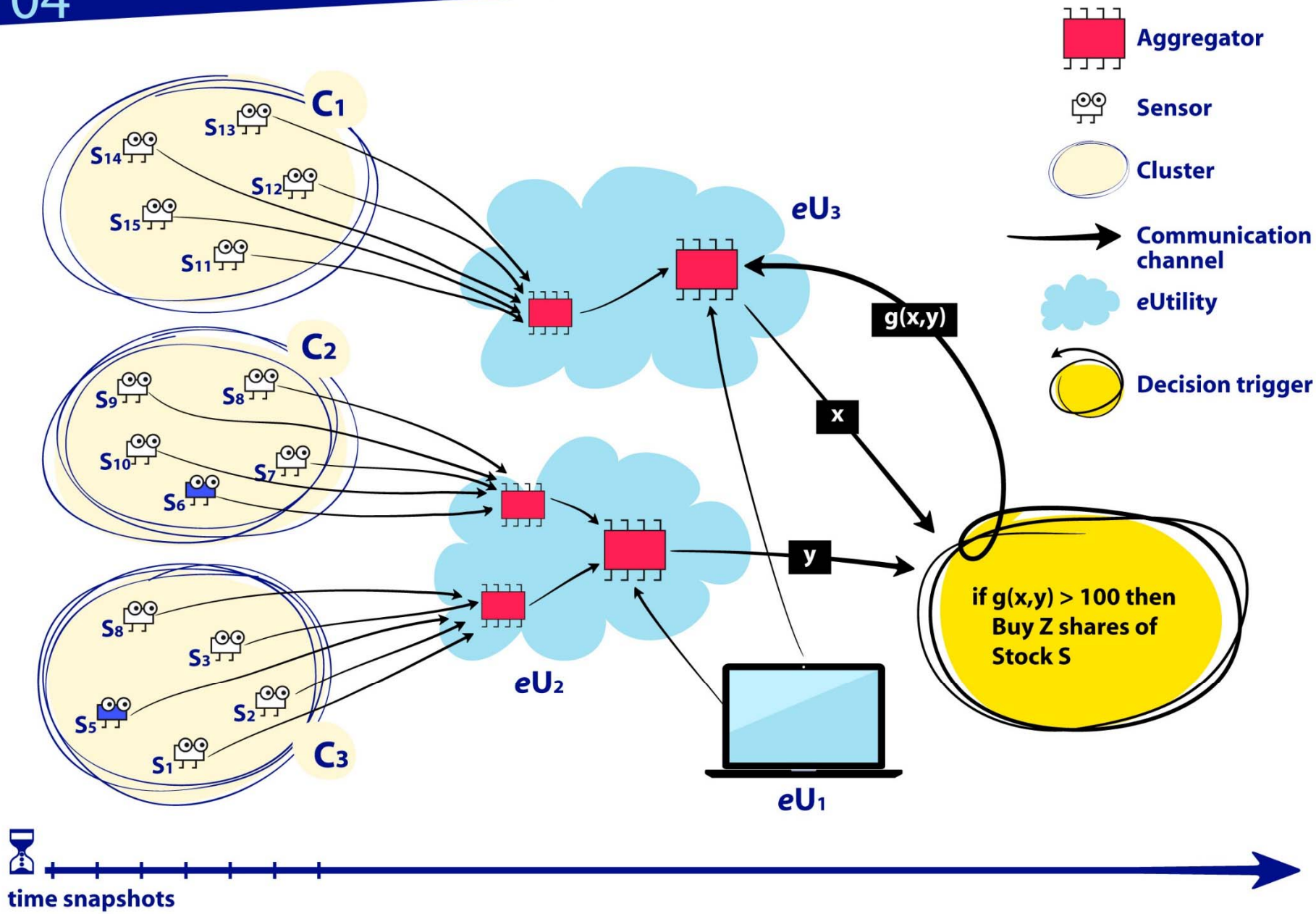
1. *e*Utilities execute processes or feed data into the overall workflow of a NoT.
2. *e*Utilities will likely be acquired off-the-shelf from 3rd parties.
3. *e*Utilities include databases, mobile devices, misc. software or hardware systems, clouds, computers, CPUs, actuators, etc. The *e*Utility primitive can be subdivided.
4. *e*Utilities, such as clouds, provide computing power that aggregators may not have.
5. A human may be viewed as a *e*Utility. A human is sometimes referred to as a 'thing' in public IoT discourse.


Aggregator
Sensor
Cluster
Communication channel
eUtility
C1
C2
C3
S13
S14
S12
S15
S11
S9
S8
S10
S7
S6
S8
S3
S5
S2
S1
eU3
eU2
eU1
time snapshots

Decision trigger (9 of 19)

1. A decision trigger is a conditional expression that triggers an action. A decision trigger's outputs control actuators and transactions. Decision triggers abstractly define the end purpose of a NoT.
2. A decision trigger should have a corresponding virtual implementation.
3. A decision trigger may have a unique owner.
4. Decision triggers may be acquired off-the-shelf or homegrown.
5. Decision triggers are executed at specific times and may execute continuously as new data becomes available.
6. It is fair to consider a decision trigger as an *if-then* rule.
7. Failure to execute decision triggers at time t_x may occur due to tardy data collection, inhibited sensors or *e*Utilities, inhibited communication channels, low performance aggregators, and a variety of other subsystem failure modes
8. Economics and costs play a role in the quality of the decision trigger's output.
9. There may be intermediate decision triggers at any point in a NoT's workflow.


Aggregator
Sensor
Cluster
Communication channel
eUtility
Decision trigger
C1
S13
S14
S12
S15
S11
eU3
C2
S9
S8
S10
S7
S6
x
D = f(x,y)
y
if f(x, y) > 200,
set wing flaps
to 10 degrees
S8
S3
S5
S2
S1
C3
eU2
eU1
time snapshots


Aggregator
Sensor
Cluster
Communication channel
eUtility
Decision trigger
C1
C2
C3
S13
S14
S12
S15
S11
S9
S8
S10
S7
S6
S8
S3
S5
S2
S1
eU3
eU2
eU1
g(x,y)
x
y
if g(x,y) > 100 then
Buy Z shares of
Stock S
time snapshots

Elements

1. **Environment** – The universe that all primitives in a specific NoT operate in; this is essentially the *operational profile* of a NoT. The environment is particularly important to the sensor and aggregator primitives since it offers context to them. An analogy is the various weather profiles that an aircraft operates in or a particular factory setting that a NoT operates in. This will likely be difficult to correctly define.
2. **Cost** – The expenses, in terms of time and money, that a specific NoT incurs in terms of the non-mitigated reliability and security risks; additionally, the costs associated with each of the primitive components needed to build and operate a NoT. Cost is an estimation or prediction that can be measured or approximated. Cost drives the design decisions in building a NoT.
3. **Geographic location –** Physical place where a sensor or *e*Utility operates in or was manufactured, e.g., using RFID. Manufacturing location is a *supply chain* trust issue. Note that the operating location may change over time. Note that a sensor's or *e*Utility's geographic location along with communication channel reliability and data security may affect the dataflow throughout a NoT's workflow in a timely manner. Geographic location determinations may sometimes not be possible.

4. **Owner** - Person or Organization that owns a particular sensor, communication channel, aggregator, decision trigger, or *e*Utility. There can be multiple owners for any of these five. Note that owners may have nefarious intentions that affect overall trust. Note further that owners may remain anonymous. Note that there is also a role for an **operator**; for simplicity, we roll up that role into the owner element.
5. **Device_ID** – A unique identifier for a particular sensor, communication channel, aggregator, decision trigger, or *e*Utility. Further, a Device_ID may be the only sensor data transmitted. This will typically originate from the manufacturer of the entity, but it could be modified or forged. This can be accomplished using RFID for physical primitives. RFID readers that work on the same protocol as the inlay may be distributed at key points throughout a NoT. Readers activate the tag causing it to broadcast radio waves within bandwidths reserved for RFID usage by individual governments internationally. These radio waves transmit identifiers or codes that reference unique information associated with the item to which the RFID inlay is attached, and in this case, the item would be a primitive.
6. **Snapshot** – an instant in time

Special Element: Snapshot

a) Because a NoT is a distributed system, different events, data transfers, and computations occur at different snapshots.

b) Snapshots may be aligned to a clock synchronized within their own network [NIST 2015]. A global clock may be too burdensome for sensor networks that operate in the wild. Others, however, argue in favor of a global clock [Li 2004]. This publication does not endorse either scheme at the time of this writing.

c) NoTs may affect business performance – sensing, communicating, and computing can speed-up or slow-down a NoT's workflow and therefore affect the "perceived" performance of the environment it operates in or controls.

d) Snapshots maybe tampered with, making it unclear when events actually occurred, not by changing time (which is not possible), but by changing the recorded time at which an event in the workflow is generated, or computation is performed, e.g., sticking in a **delay()** function call.

e) Malicious latency to induce delays, are possible and will affect when decision triggers are able to execute.

f) Reliability and performance of a NoT may be highly based on d) and e).

Trustworthiness

Primitive or Element	Attribute	Pedigree Risk?	Reliability Risk?	Security Risk?
Sensor	Physical	Y	Y	Y
Aggregator	Virtual	Y	Y	Y
Communication channel	Virtual and/or Physical	Y	Y	Y
eUtility	Virtual or Physical	Y	Y	Y
Decision trigger	Virtual	Y	Y	Y
Geographic location	Physical (possibly unknown)	N/A	Y	Y
Owner	Physical (possibly unknown)	?	N/A	?
Environment	Virtual or Physical (possibly unknown)	N/A	Y	Y
Cost	Partially known	N/A	?	?
Device_ID	Virtual	Y	Y	Y
Snapshot	Natural phenomenon	N/A	Y	?

Summary

1. IoT is a brand or catalogue of technologies – not a singular technology
2. Discussing NoTs makes more sense than discussing IoT.
3. Primitives and elements offer "a science."
4. NoTs can be defined, measured, and compared.
5. Can I compare your IoT to my IoT? No!
6. The goal is to someday measure Trust of a NoT:

   ***Trust* in some NoT *A*, at some snapshot *X*, is a function of NoT *A*'s assets ϵ {sensor(s), aggregator(s), communication channel(s), *e*Utility(s), decision trigger(s)} with respect to the members ϵ {geographic location, owner, environment, snapshot, cost, Device_IDs}, for each asset in the first set, when applicable.**

7. Feedback from people that have seriously considered this is that there is some level of *elegance* is this *simpl*e 5 + 6 part model that is useful to better understand "What is IoT?"