Trustworthiness of Workflows in Cloud Computing

David M. Nicol

Franklin W. Woeltge Professor of ECE

Director, ITI

Trust in Cloud Computing

Trust has different attributes

- Performance attributes
  - Meets latency requirements, meets bandwidth requirements
- Reliability attributes
  - Cloud computation will not fail over some mission time with more than a specified probability
- Availability attributes
  - Cloud is available for use for some specified fraction of time over some specified period
- Privacy
  - Cloud will not divulge personal data or allow inference
- Security
  - Cloud will maintain Confidentiality, Integrity, Availability of data

Trust in Cloud Computing

There is a notion of 'expectation', e.g.

- Service provider meets SLA w.r.t performance
- Service provider practices industry standard norms with respect to
  - Security
  - Data sharing
- Service provider takes extra-ordinary pre-cautions with respect to separation

Expectations will vary with consumer

Consumer might choose service based on *perception* of how well the provider meets expectations

On what basis might that perception be quantified?

What is trust?

- Many definitions …
- Trust is defined as 3 elements:
  - **Expectation**, trustor expects a specific thing from trustee
  - **Belief** in that expectation
  - **Willingness to take risk** for that belief.

- Belief is based on evidence about trustee on
  - **Consistency (C)** -> integrity
  - **Intension (I)** -> goodwill
  - **Ability (A),** or competence

(CIA triad of trust evidence)

User Expectation on reliability?

- SLA might specify service available 99.9% time *per year*
  - User might require
    - Specific bound on recovery time after failure
    - 99% available *every week*
- Note the gap between SLA specification and user needs / expectations

User Expectation on cloud privacy?

- Trust in privacy protection in cloud computing
- [Westin 1967]: **Privacy** is "the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others**".**
- **Domain of Expectation**: how a cloud service provider handles users' data
  - Data collection
  - Data usage
  - Data guarding
  - Data situation informing
  - Data dissemination
  - Data termination and disposal.

Trust Attributes are Interdependent

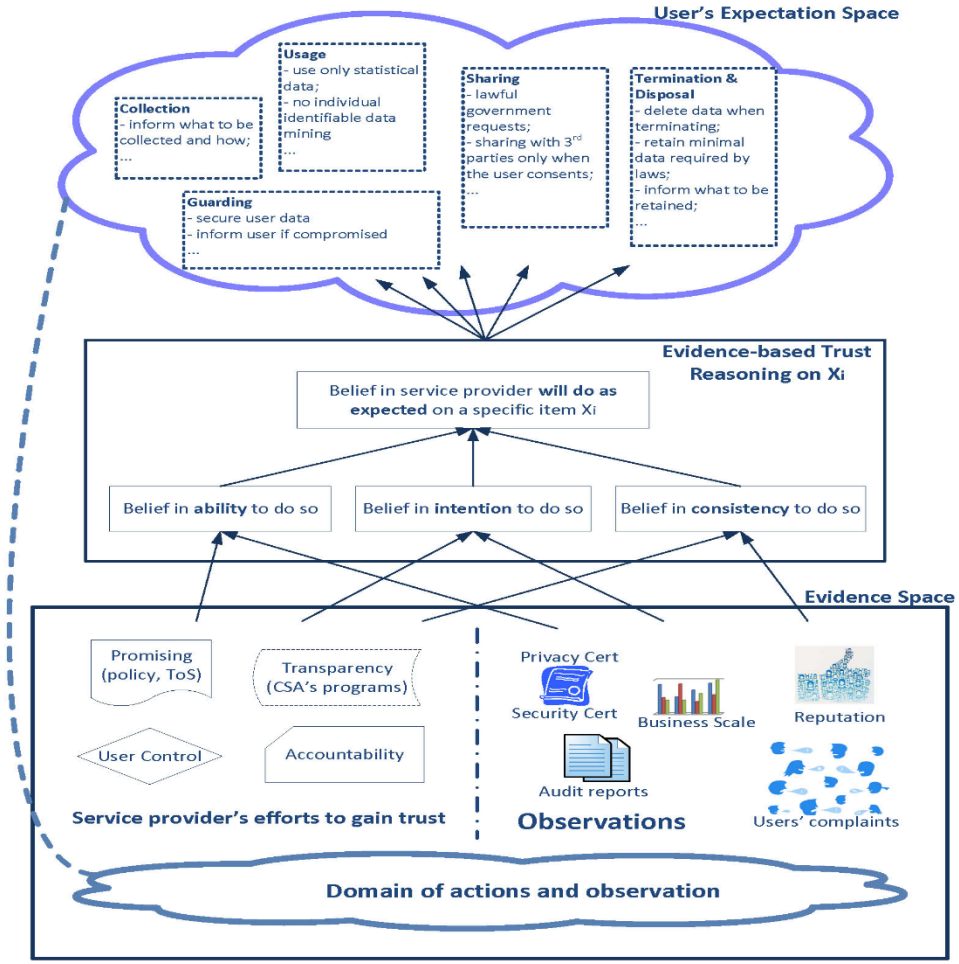
- Security laxity can lead to disruption, which affects everything
- Smallest latency may involve short-cuts that expose data to
  - Lost of confidentiality and/or integrity
- Meeting expectations w.r.t. integrity (using crypto) affects throughput
- Reliability affects availability

We want to be able to make quantitative assessment of trust attributes in the face of

- Lack of knowledge / uncertainty
- Inter-relationships between trust attributes

Inferring Belief from Evidence

Might quantify belief by triple [p{true},p{false},p{unknown}]



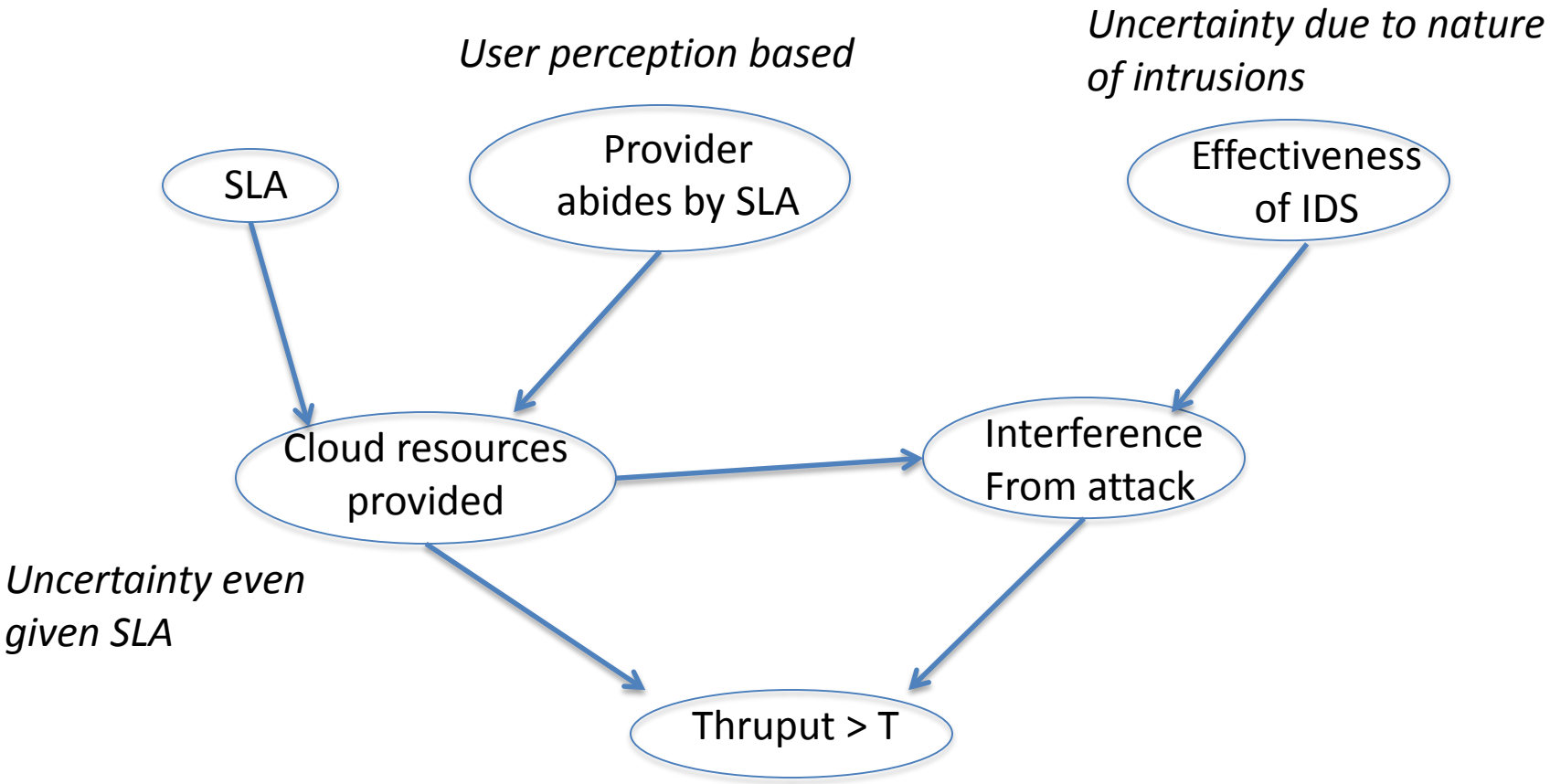
Mathematical Tools

"Belief Network" formalizes one way of computing belief "probabilities"

- Nodes represent Random variables
- Edges represent dependencies
- Belief network forms a directed acyclic graph
  - A node's distribution is specified conditioned on knowledge of its parents' values
  - For us nodes represent probabilities of particular assertions related to trust being true
  - Very general framework

Example

- Throughput depends on many related trust attributes



The Truly Hard Part

Good News:

- Conditional probability functions in BN are very expressive and have the potential to allow us to *model* both uncertainty, and inter-dependency of attributes
- *E.g.. Pr{workflow has throughput T | file system is encrypted }*

Best News: It's a jobs program for researchers!:

- On what scientific / statistical basis can we derive these models from knowledge of software used and empirical measurements?
- How can we sensibly mix objective measurements with subjective quantifications
- How much probability weight can we place on certainty when there unknowns or uncertainty in model parameters, or in model relationships?

Summary

Trust in cloud-based workflows has interesting characteristics

- Subjectivity of provider meeting user expectations
- Uncertainty due to lack of knowledge or uncertainty associated with measurements
- Relationships between trust attributes that have to be modeled, with its own uncertainty

We exploring variants of belief networks, with special emphasis on the the scientific issues of developing models of trust attribute inter-relationships