Information Protection in Content-centric Networks

Christopher C. Lamb

Department of Electrical and Computer Engineering
University of New Mexico

November 6, 2012

THE UNIVERSITY of NEW MEXICO

Outline

Original Goals

Contribution of Work

The contribution of this work is a quantitative analysis of policy-centric overlay network options, associated taxonomies of use, and prototypical technology proofs-of-concept.

- *Network Control Options* — This includes various types networks and associated strengths and weaknesses addressing centralized and decentralized models.
- *Taxonomies of Use* — Depending on the specific usage management requirements and context, different overlays have different applicability; this work will provide guidance on suitability; it will eventually lead to how to manage data flow within SDN-capable infrastructure.
- *Prototypical Technologies* — Examples and proofs-of-concept will be required to appropriately analyze various architectural alternatives.

Meeting the Goals

Network Control Options

I have developed and analysed multiple types of overlay systems, both centralized (hierarchical) and non-centralized (non-hierarchical), with differing topologies and integrated content-centric control.

Taxonomies of Use

I have established an verified a taxonomy of usage management and applied that within the network providing mechanisms extendable to SDN use.

Prototypical Technologies

Prototype information-centric networks are running between the Rackspace and Amazon clouds.

Impact and Originality

- Information-centric architectures common in future internet designs
- Significant work with respect to name/object binding, overall topologies, approaches
- No significant work yet on exploiting information-centricity for enhanced security
- They have significant new capabilities inherent in approach that allow for better information security

Additional Contributions

This work, as well as providing alternatives analysis with respect to information-centric security with respect to architectures and approaches, also demonstrates the first implementation of granular context-sensitive security functionality embedded in an information-centric network.

Publications

Accepted Conference Papers:

C.C. Lamb and G.L. Heileman. *Overlay architectures enabling cloud computing for multi-level security environments*. In Services (SERVICES), 2012 IEEE Eighth World Congress on, pages 116-124, june 2012.

Christopher Charles Lamb, Pramod A. Jamkhedkar, Mathew P. Bohnsack, Viswanath Nandina, and Gregory L. Heileman. *A domain specific language for usage management*. In Proceedings of the 11th annual ACM workshop on Digital rights management, DRM '11, pages 51-62, New York, NY, USA, 2011. ACM.

Christopher C. Lamb, Pramod A. Jamkhedkar, Gregory L. Heileman, and Chaouki T. Abdallah. *Managed control of composite cloud systems*. In System of Systems Engineering (SoSE), 2011 6th International Conference on, pages 167-172, june 2011.

P.A. Jamkhedkar, C.C. Lamb, and G.L. Heileman. *Usage management in cloud computing*. In Cloud Computing (CLOUD), 2011 IEEE International Conference on, pages 525-532, july 2011.

Pramod A. Jamkhedkar, Gregory L. Heileman, and Chris C. Lamb. *An interoperable usage management framework*. In Proceedings of the tenth annual ACM workshop on Digital rights management, DRM '10, pages 73-88, New York, NY, USA, 2010. ACM.

Publications

Accepted Journal Submissions:

Christopher C. Lamb and Gregory L. Heileman, "Content-centric Information Protection in Cloud Computing", *International Journal of Cloud Computing and Services Science* vol. 1, no. 5, December 2012.

Accepted Book Chapters:

Pramod A. Jamkhedkar, Christopher C. Lamb, and Gregory L. Heileman, *Digital Rights Management: Technology, Standards and Applications*, Auerbach Publications, 2013.

Pending Journal Submissions:

Christopher C. Lamb and Gregory L. Heileman, "Dynamic Context-sensitive Information Protection", *IEEE Internet Computing - Dynamic Collective Work*.

Results Overview

Overall evaluation of impact against strategy:

- Encryption most likely to be used...
- ...Rerouting likely the best compromise (but expensive)
- Hierarchical and non-hierarchical networks had similar performance
- No clear leading strategy under all conditions

Property	Redaction	Rerouting	Encryption
Confidentiality	3	2	1
Integrity	0	1	3
Availability	0	1	2

Strategy Impact by Attribute (3 is best, 0 is worst)

What does this mean? How did we get it?

Methodology

Confidentiality, Integrity characteristics based on approach.

- **Redaction**, by removing information, by definition destroys integrity while guaranteeing confidentiality; unavailable information that is cannot be leaked
- **Rerouting** removes information from a context damaging integrity that can possibly be repaired later, potentially increasing confidentiality by rendering that information unavailable
- **Encryption** minimizes integrity impacts be keeping ciphered data with original context at the expense of possible interception and cryptanalysis exposure

Availability is based on performance.

- **Performance** is measured via end-to-end time of transmittal

Redaction

Redaction: Removing content that is not approved for transmission over a given link or consumption by a given agent from a larger context of suitable content.

- Strongest confidentiality
- Destroys integrity
- Mixed impact on availability

Fast and easy to implement

Property	Redaction	Rerouting	Encryption
Confidentiality	3	2	1
Integrity	0	1	3
Availability	0	1	2

Rerouting

Rerouting: Removing content that is not approved for transmission over a given link and rerouting that content to its destination through secondary means (e.g. SMTP).

- Confidentiality dependent on secondary links
- Integrity compromised temporarily and perhaps permanently
- Availability dependent on secondary links

Undependable, expensive, good information control

Property	Redaction	Rerouting	Encryption
Confidentiality	3	2	1
Integrity	0	1	3
Availability	0	1	2

Encryption

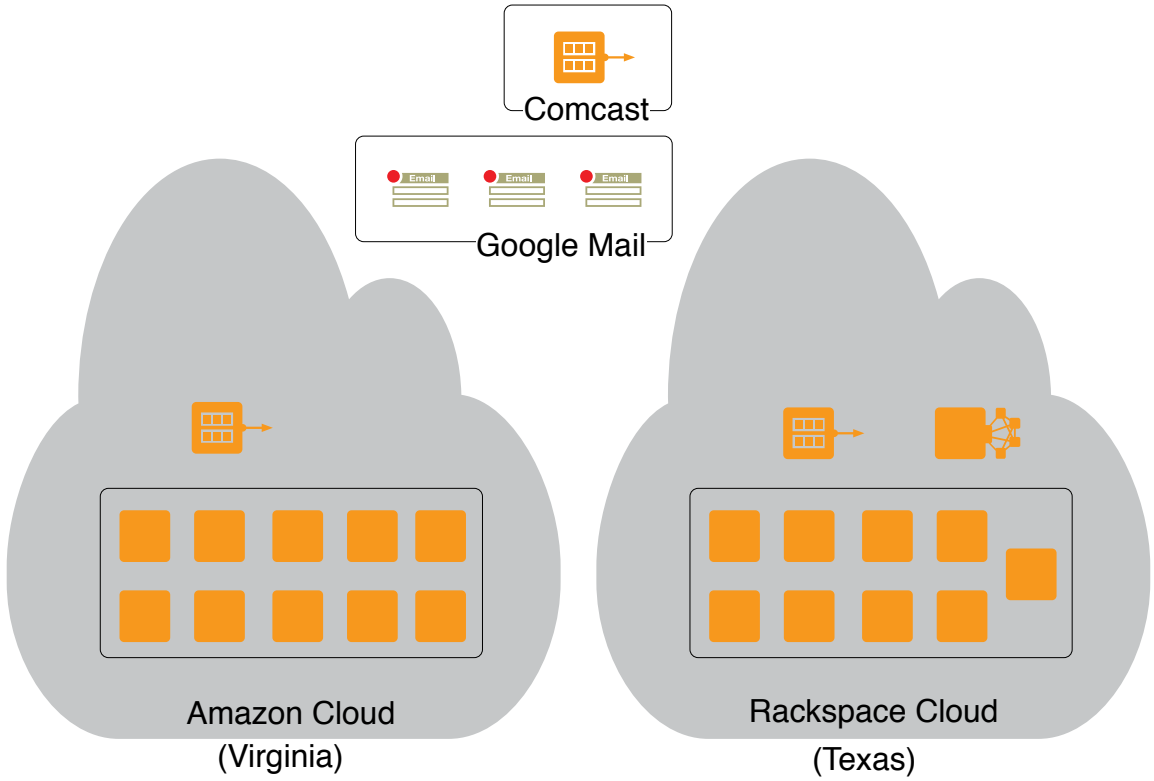
Encryption: Enciphering content within larger documents, deciphering enciphered sections when suitable by defined policy and when content needs to be re-evaluated.

- Confidentiality questionable over time
- Integrity compromised temporarily and perhaps permanently
- Availability dependent on secondary links

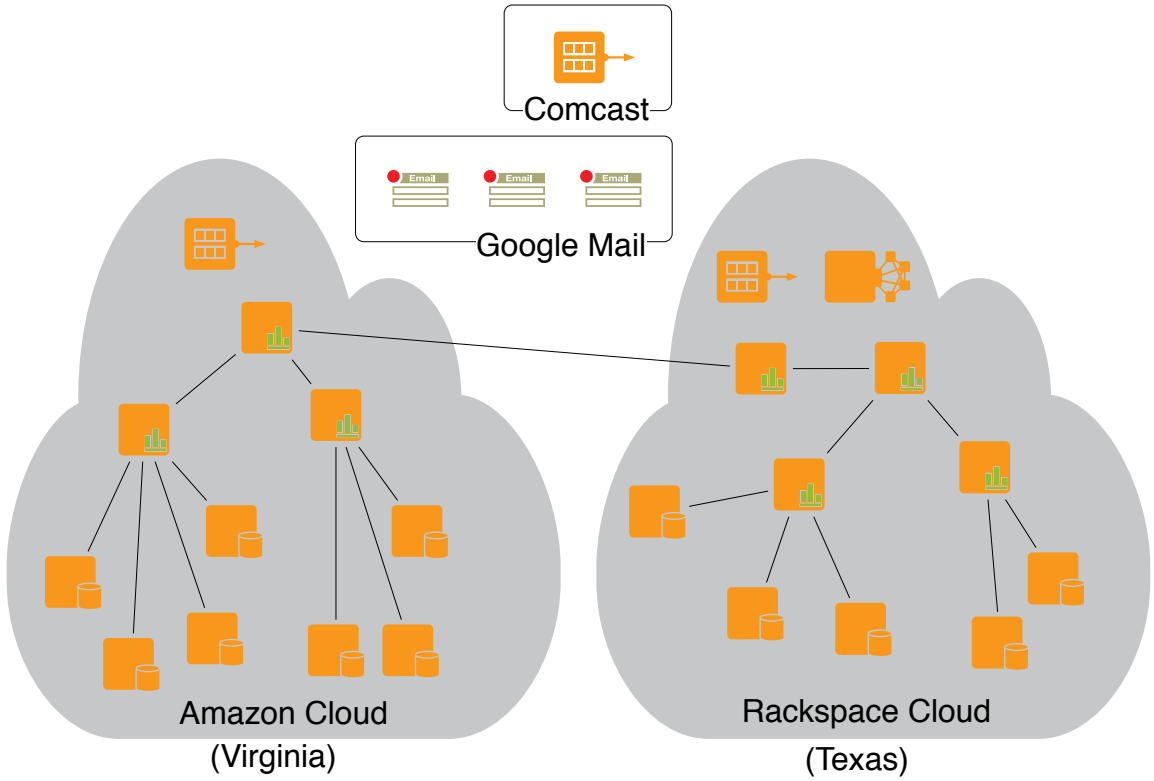
Reasonably secure, simple and performant

Property	Redaction	Rerouting	Encryption
Confidentiality	3	2	1
Integrity	0	1	3
Availability	0	1	2

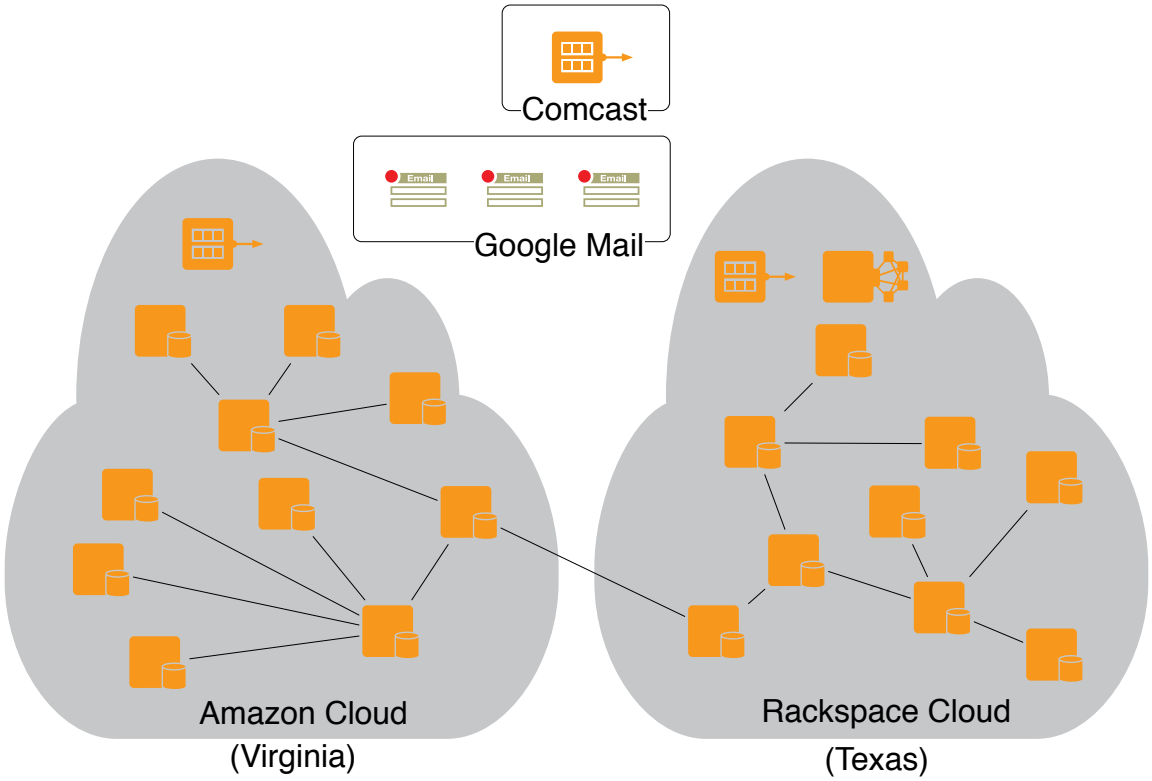
Physical Topology



Hierarchical Topology



Non-Hierarchical Topology



Hierarchical Effects

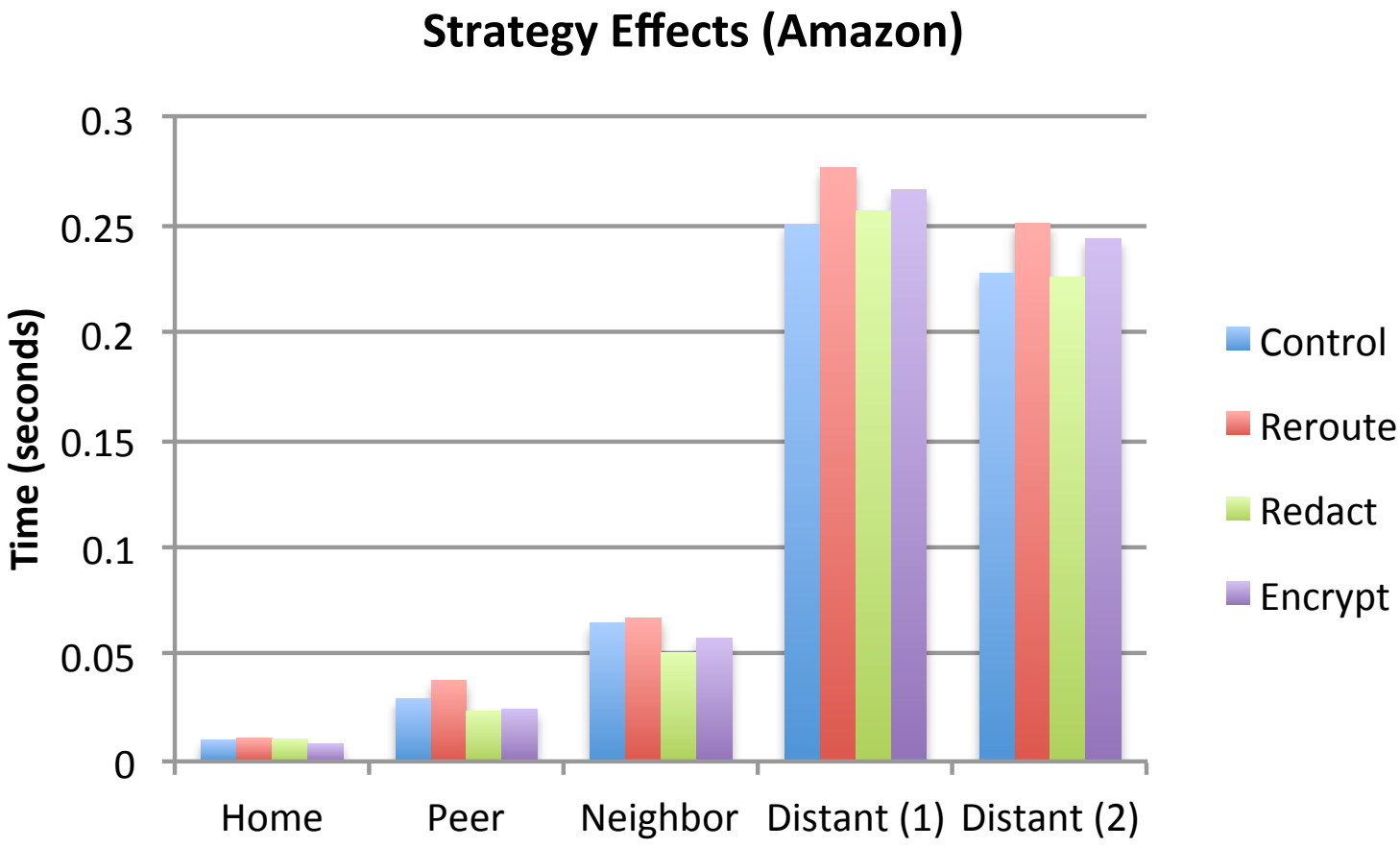


Figure: Hierarchical Results from Amazon

Hierarchical Effects

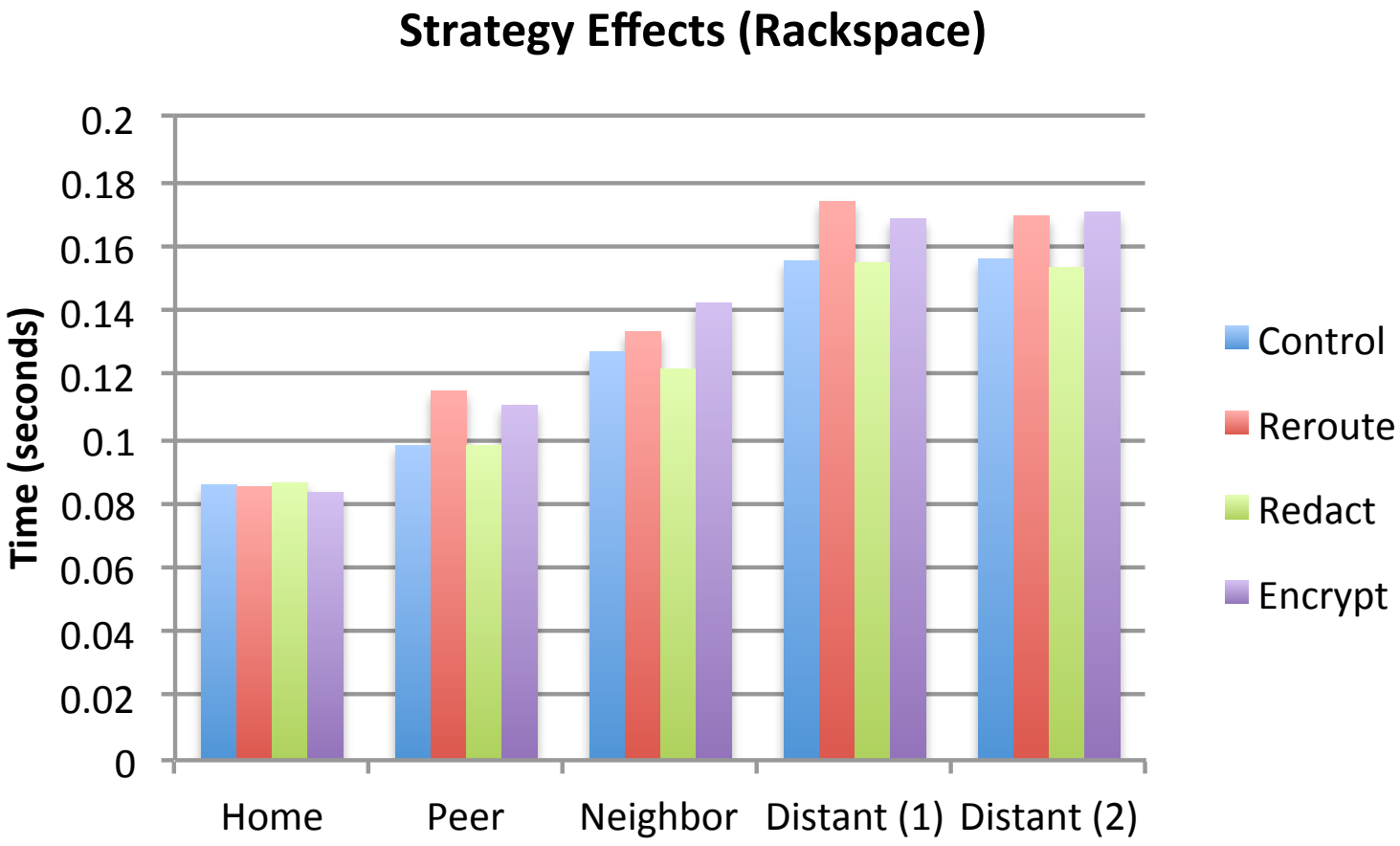


Figure: Hierarchical Results from Rackspace

Hierarchical Effects

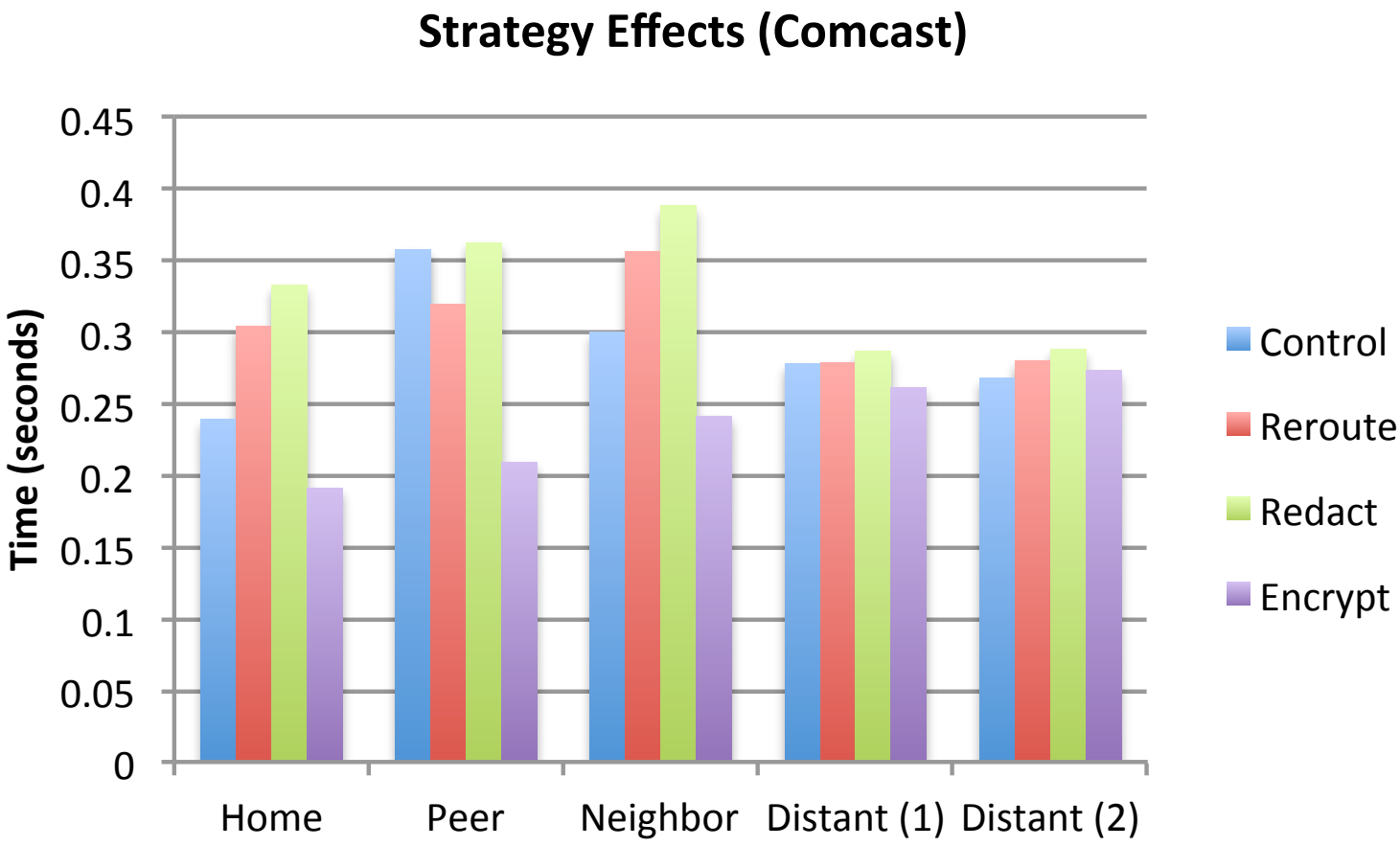


Figure: Hierarchical Results from Comcast

Hierarchical Analysis

Caching is important.

Network effects have a tremendous impact on performance. Using in-node content caching helps take advantage of query locality.

Less naive routing wouldn't hurt either.

A query from the amazon test node to rackspace nodes travels to rackspace, to amazon, to rackspace, and then back to amazon. Better node location services or caching could eliminate this back-and-forth.

Infrastructure costs are high.

Rerouting content has some significant conceptual advantages, but establishing reliable secondary infrastructure can be expensive and difficult.

Encryption is a realistic compromise.

Encryption seemed performant and didn't sacrifice integrity and availability for confidentiality gains.

Non-Hierarchical Effects

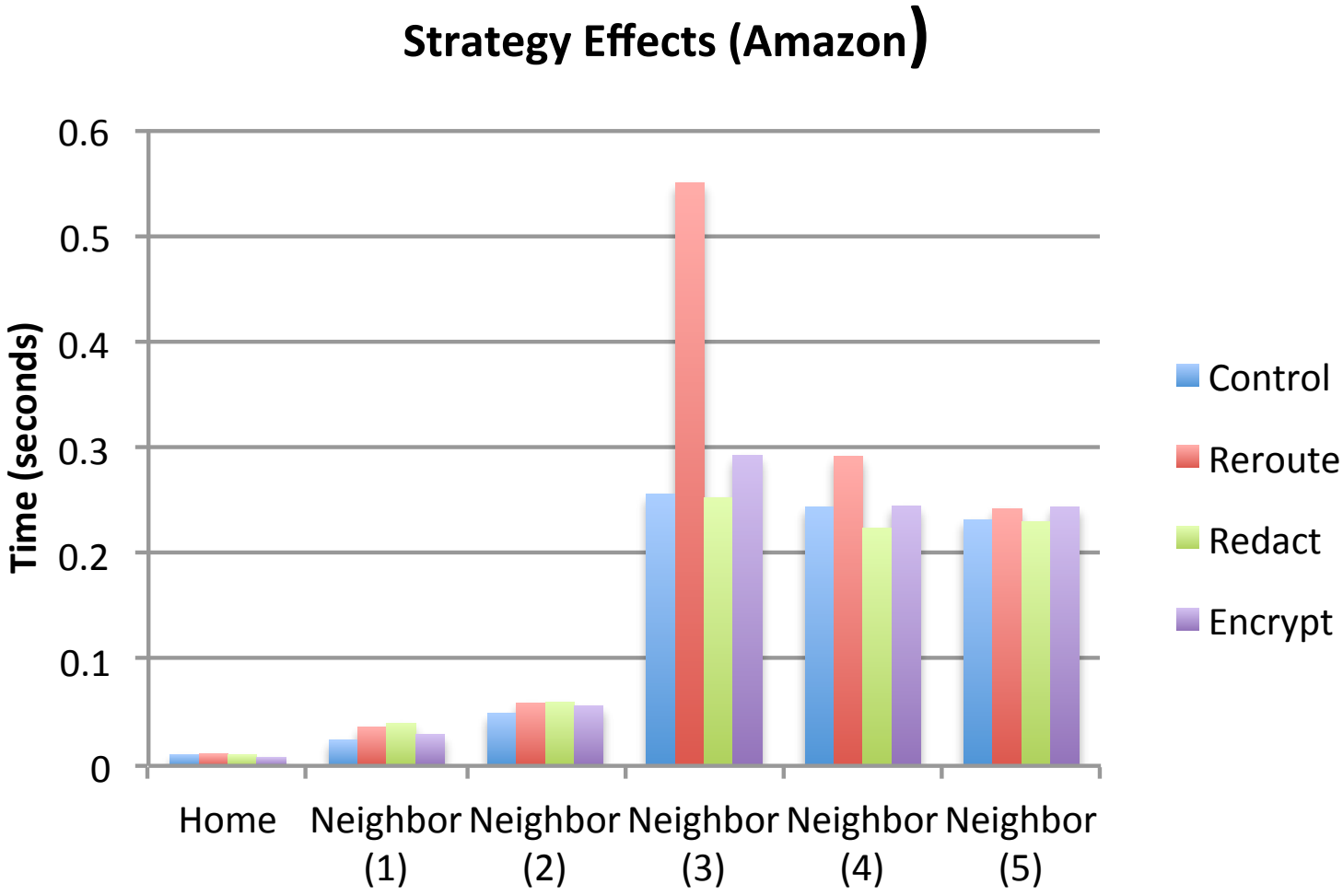


Figure: Non-Hierarchical Results from Amazon

Non-Hierarchical Effects

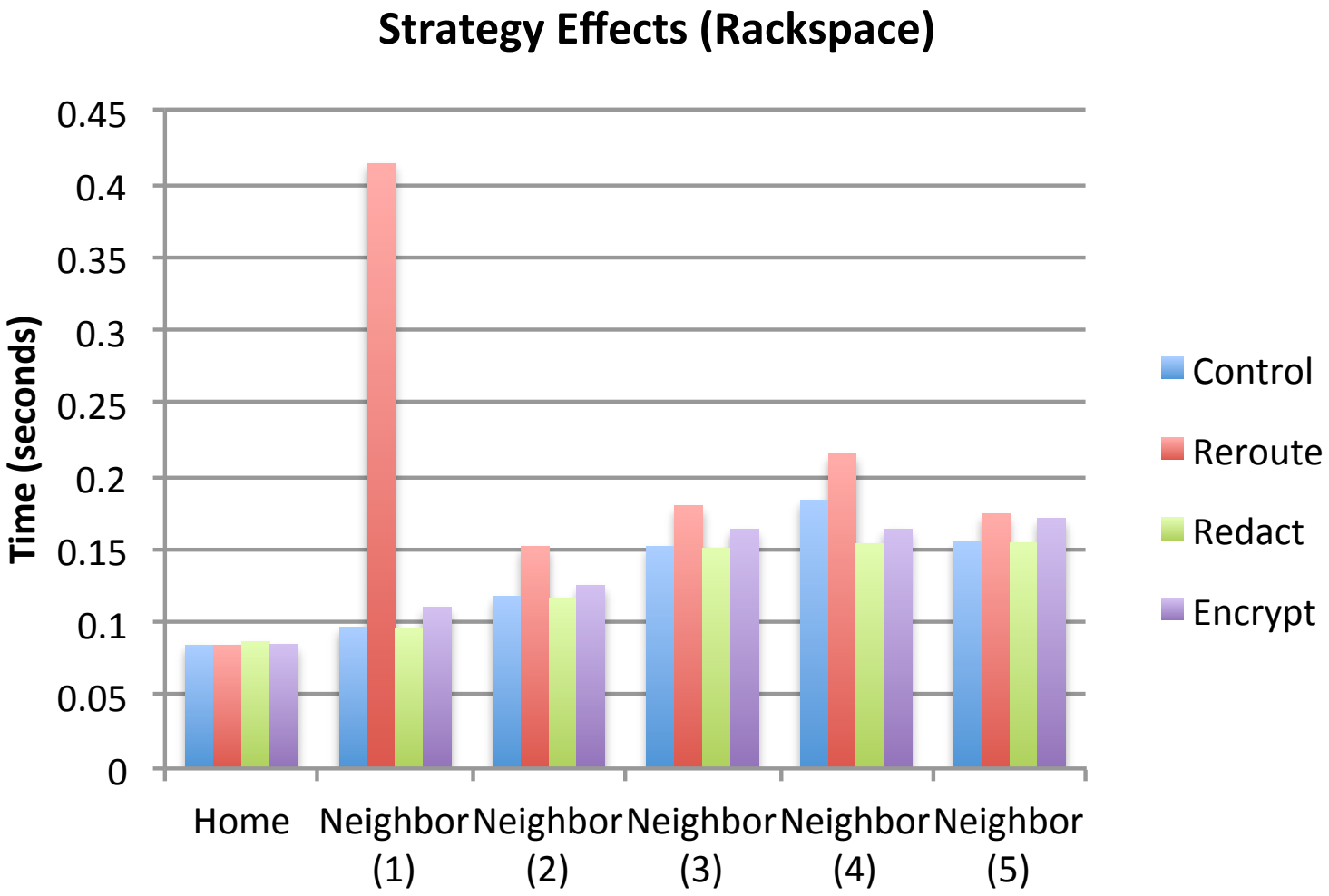


Figure: Non-Hierarchical Results from Rackspace

Non-Hierarchical Effects

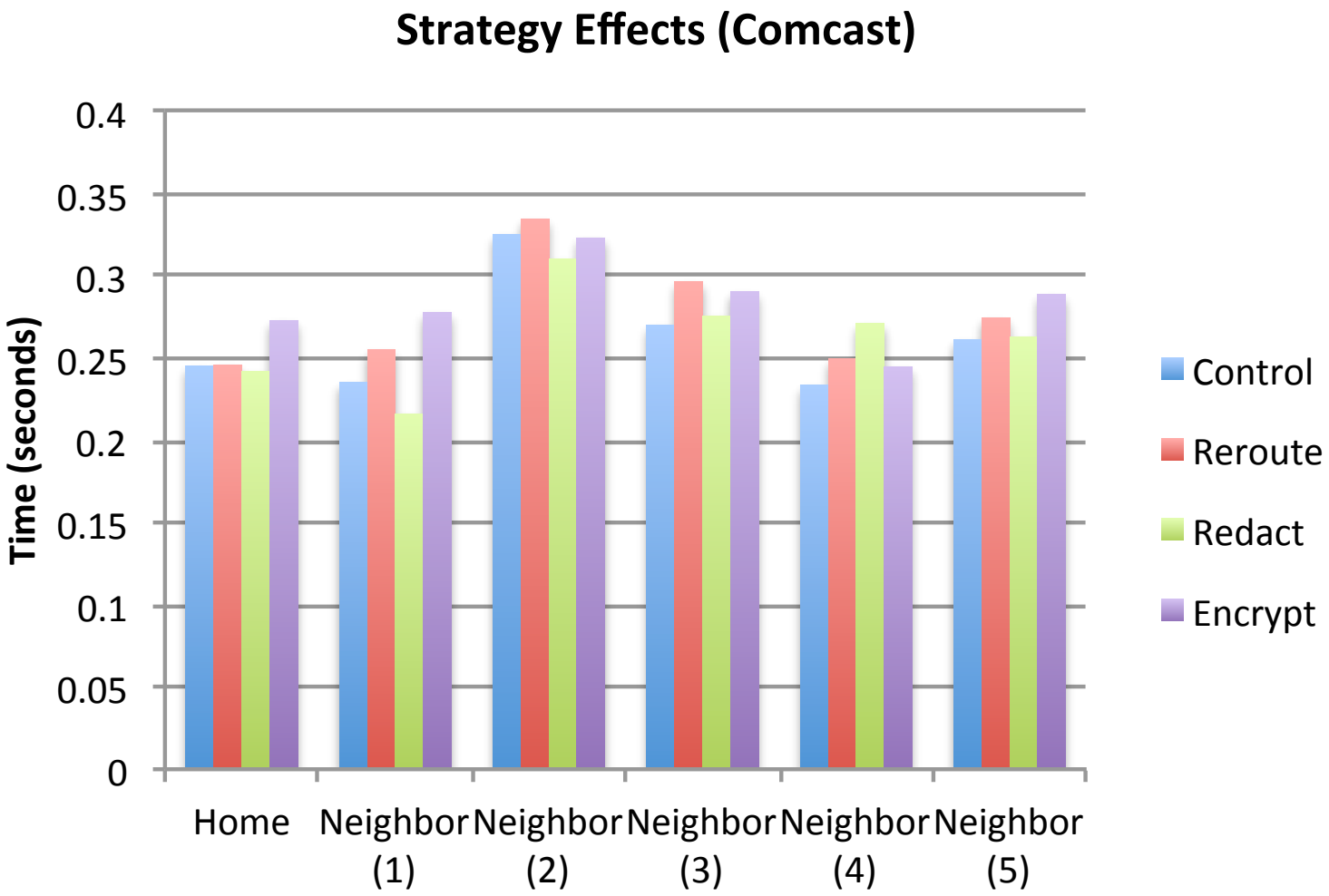


Figure: Non-Hierarchical Results from Comcast

Non-Hierarchical Analysis

Similar performance to hierarchical topologies.

We see similar performance profiles for hierarchical and non-hierarchical networks. Conclusions from hierarchical networks with respect to routing and caching hold here as well. Non-hierarchical networks do have additional content processing demands.

Secondary infrastructure effects more prominent.

In these tests, we had significant infrastructural issues with respect to rerouting with our provider. Again, secondary infrastructure can be problematic.

Non-hierarchical and hierarchical have different strengths.

Hierarchical networks may more effectively resist random failures due to functional centralization. Non-hierarchical networks may better maintain functionality in the face of directed attacks. This will likely depend on topology as well as hierarchical characteristics. *More research is needed to support these conclusions.*

Network-Free Evaluation

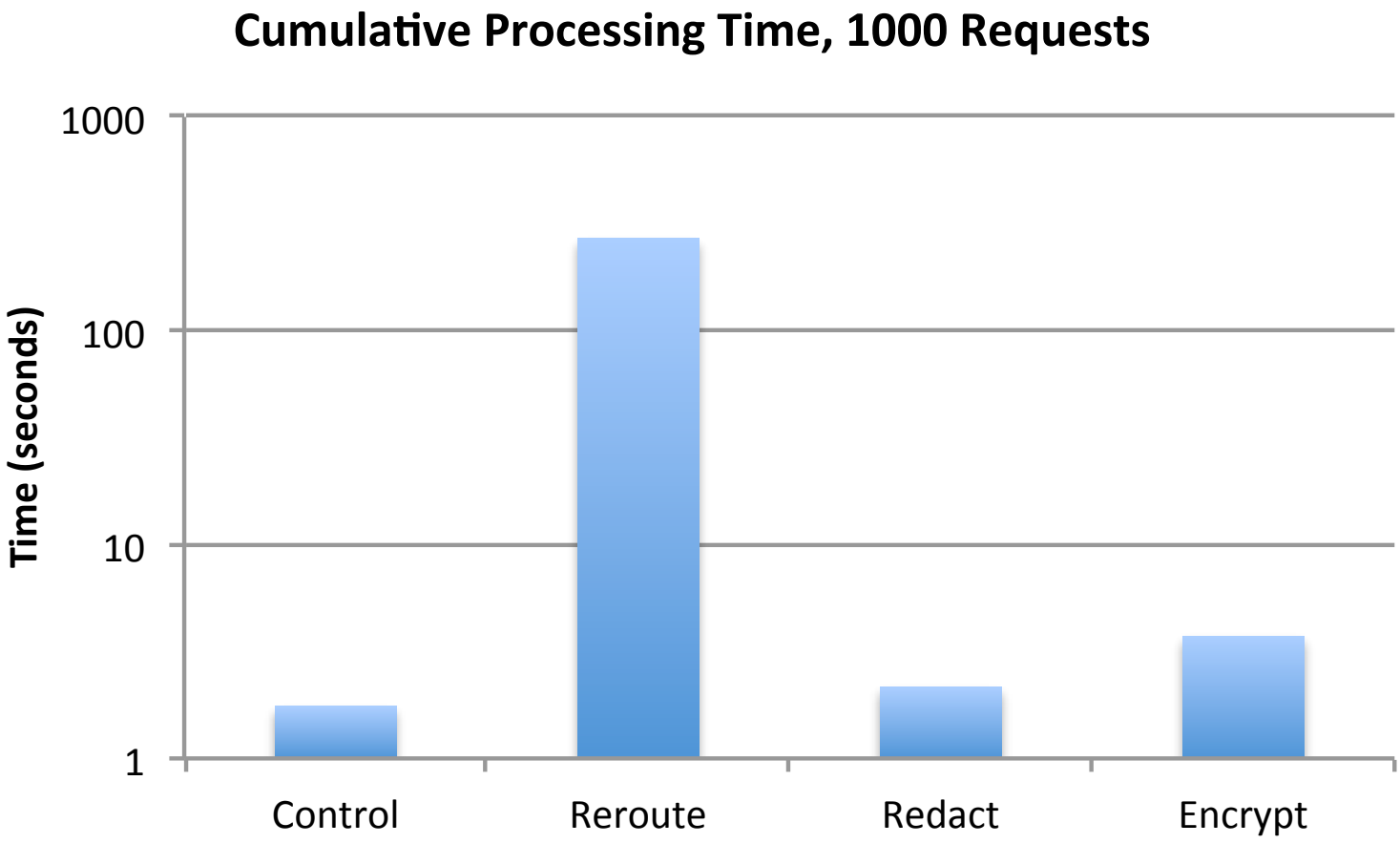


Figure: Results from Requests to a Singe Node

Network-Free Analysis

Doing nothing is cheapest.

As expected, the control group has the best performance. We are not processing content in the control example, rather we just pass content directly through the node.

Redaction is next.

Computationally, redaction is the simplest strategy, and is just slightly more expensive than the control set, but it destroys integrity.

Encryption is just slightly more expensive than redaction.

We do use symmetric ciphers here, so this is very fast. Public key encryption would be more expensive.

Rerouting is by far the most expensive.

Note that this is a logarithmic scale graph! Setting up secondary channel communication is expensive and time consuming.

Conclusions

What did we learn?

- Redaction destroys integrity
- Rerouting is expensive
- Encryption can be broken

Surprisingly, rerouting is not a realistic option!

- Rerouting may provide the best compromise between confidentiality and integrity, but an a potentially high availability, and likely financial, cost. Maintaining dual infrastructures, especially if the second is highly secure, will more than double infrastructure spend.
- Encryption may be eventually broken, but is difficult to decipher short-tem, is easier to implement, and is likely much less expensive than rerouting.

What's next?

This work is funded for an additional six months, and likely for an additional three years of continuing funding. Goals currently include:

- **Productionization** of the current system for eventual deployment in operational systems. This will involve extending the current infrastructure to handle additional types of data and potentially different and lower-level protocols.
- **Public Key Infrastructure** to provide non-repudiation as well as confidentiality to protected information. This will likely involve some combination of symmetric and asymmetric cryptography (e.g. TLS, IPsec).
- **Software Defined Networking** to enable real-time network configuration in response to changing contexts.
- **Publish/Subscribe** approaches to registering for content of interest, and associated decoupling of request and response network paths.

Questions? Comments?