

**Dynamo:**  
Amazon's Highly Available Key-value Store

Giuseppe DeCandia, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati,  
Avinash Lakshman, Alex Pilchin, Swaminathan Sivasubramanian, Peter Vosshall  
and Werner Vogels

Amazon.com

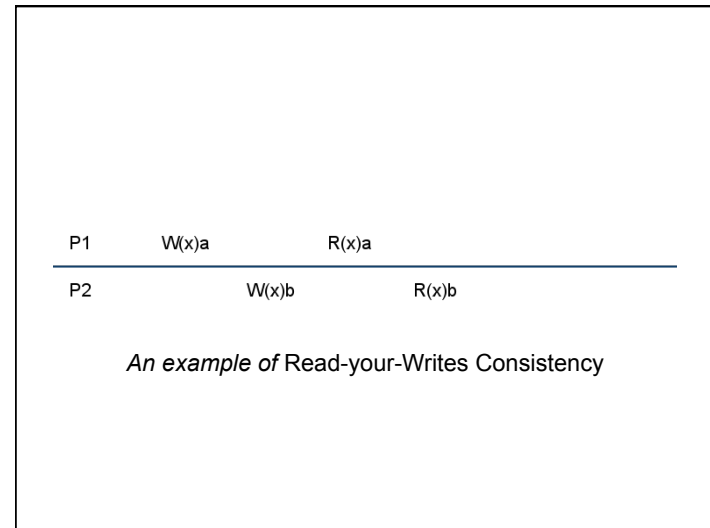
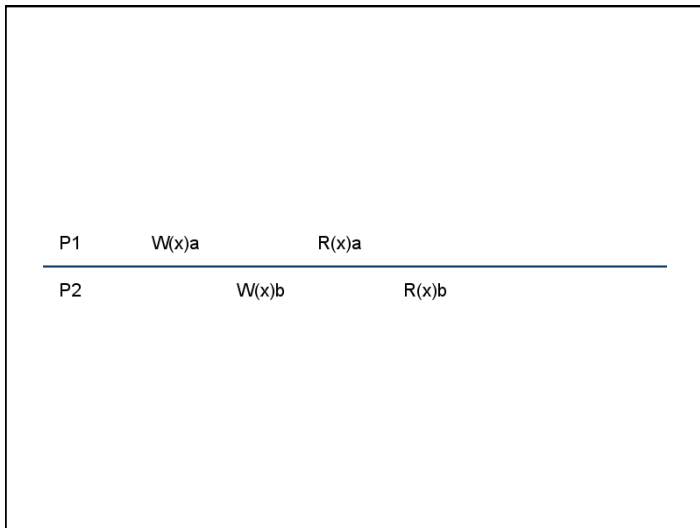
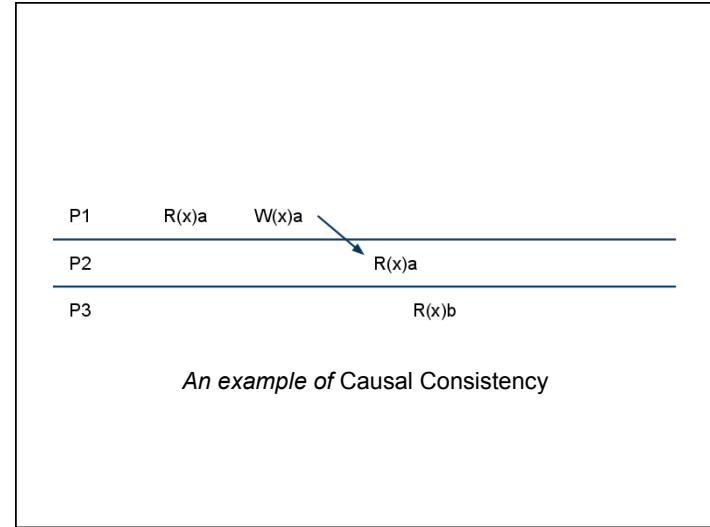
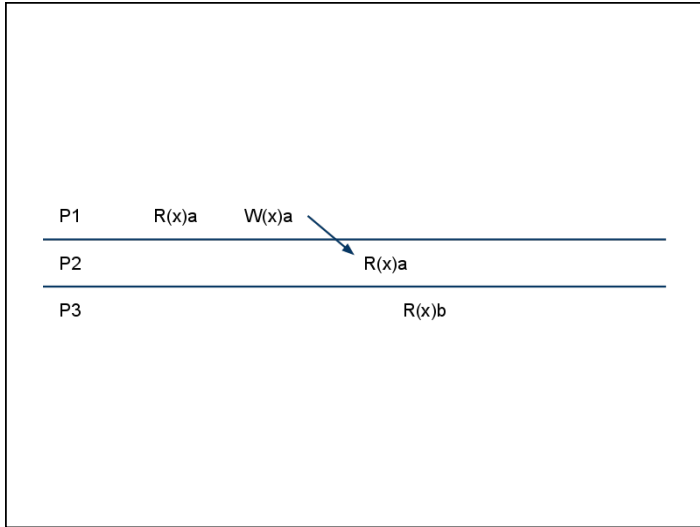
Prologue: Eventual Consistency

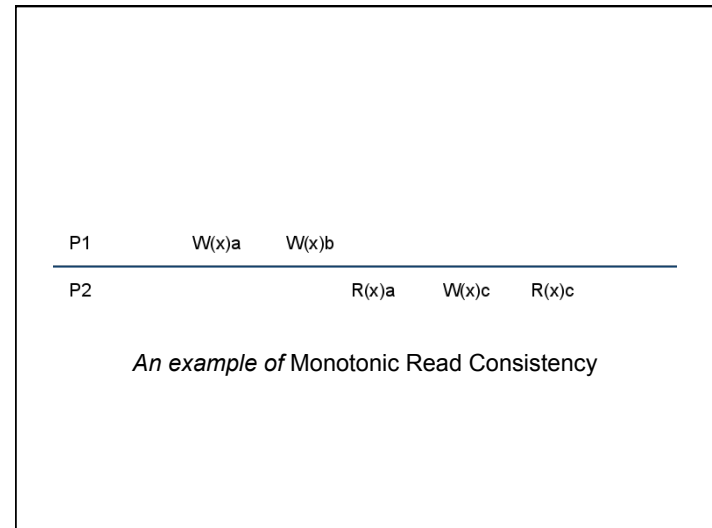
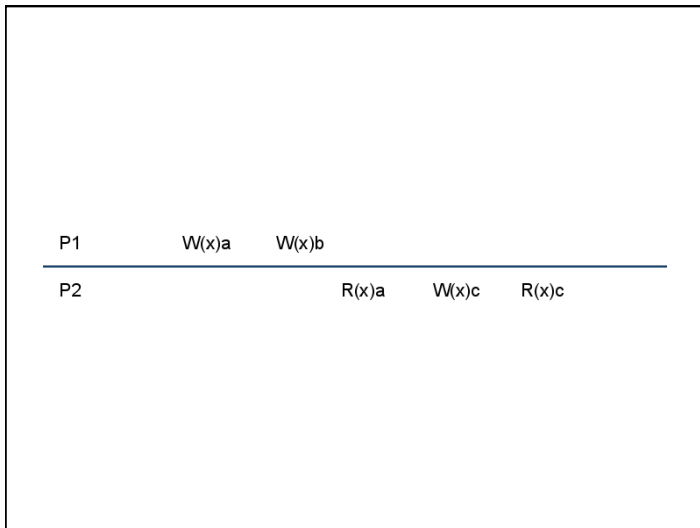
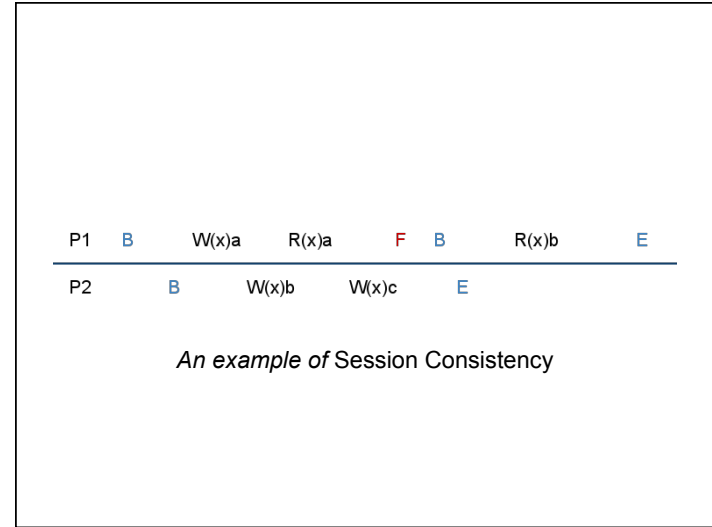
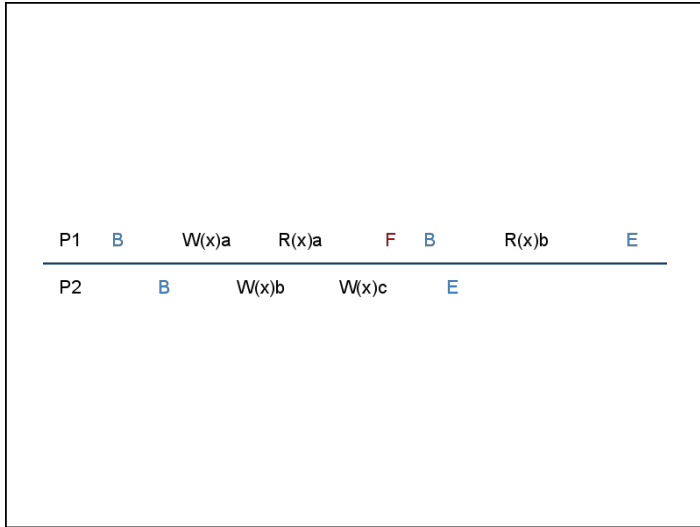
Prologue: Eventual Consistency

- Werner Vogel. Eventually Consistent. CAMC, 52(1): 40-44. 2009

Prologue: Eventual Consistency

- Werner Vogel. Eventually Consistent. CAMC, 52(1): 40-44. 2009
- Notation....
  - $R(x)a$  -- Read from shared variable  $x$ ; the result is  $a$
  - $W(x)a$  -- Write value  $a$  to shared variable  $x$
  - $P$  -- A [client] process or thread
  - $B$  -- Session beginning
  - $F$  -- Session failure
  - $E$  -- Session end





$$W + R > N$$

$$W + R > N$$

$$W + R \leq N$$

## Dynamo

- Prologue: Eventual Consistency
- Act I: 107 / 104
- Act II: 99.9%
- Act III: Related Work
- Entr'Act: Taxonomy of Correctness
- Act IV: System Architecture
- Act VI: Lessons Learned
- Act VII: Conclusion

# 107

**107** Customers *at peak times*

**107** Customers *at peak times*  
**104**

**107** Customers *at peak times*  
**104** Nodes *in data centers around the world*

**107** Customers *at peak times*  
**104** Nodes *in data centers around the world*  
**1000** Customers *per Node*

Act II: 99.9%

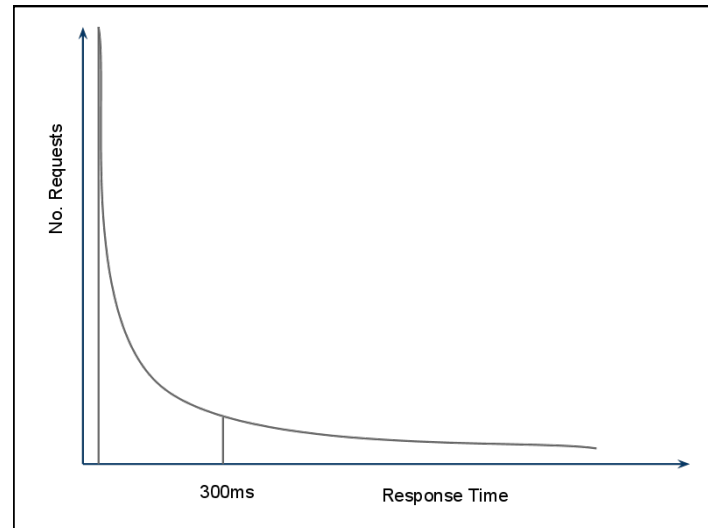
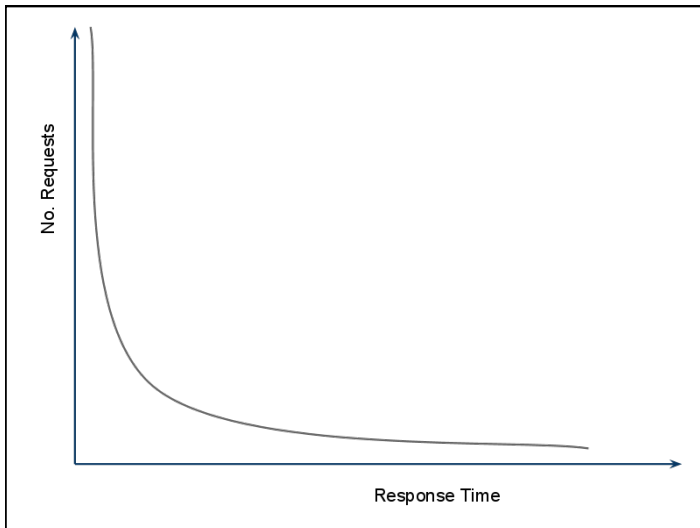
Key	Value
Act II	99.9%

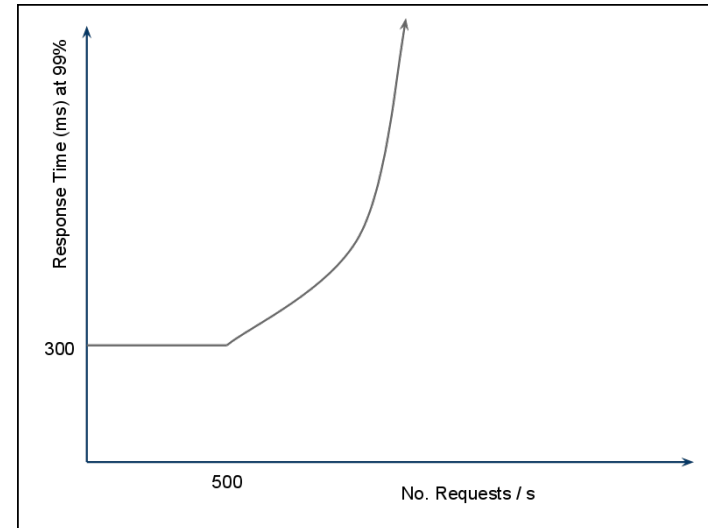
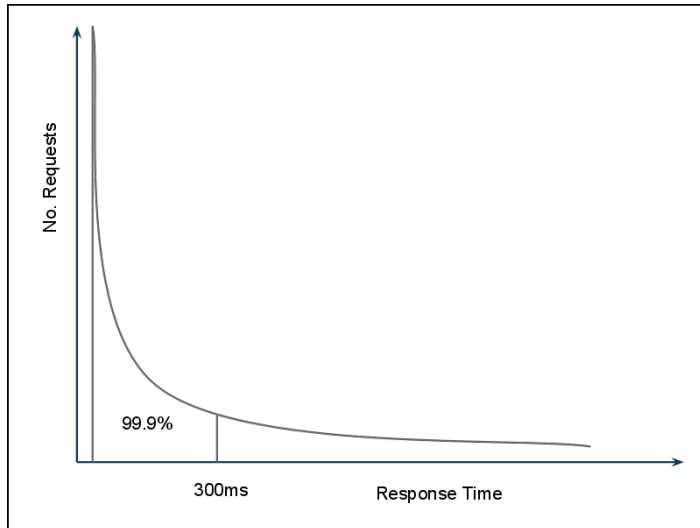
Key	Value
Act II	99.9%
	....

Key	Value
Act II	99.9%
	....
123456	abcdefg

Key	Value
Act II	99.9%
	....
123456	abcdefg

Key	Value
Act II	99.9%
	....
123456	zyxwvu





$QW > QR$

$QW > QR$   
*-> always writeable!*



**QW > QR**

*-> always writeable!*

Data Center      Application

**QW > QR**

*-> always writeable!*

Data Center      Application  
*last write wins*

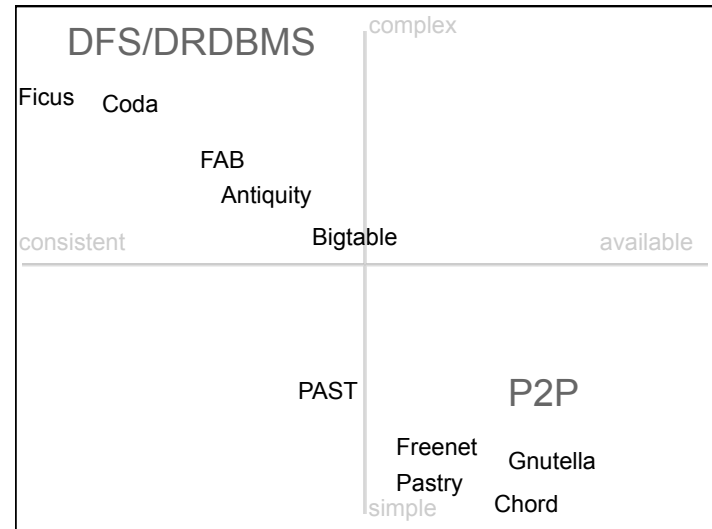
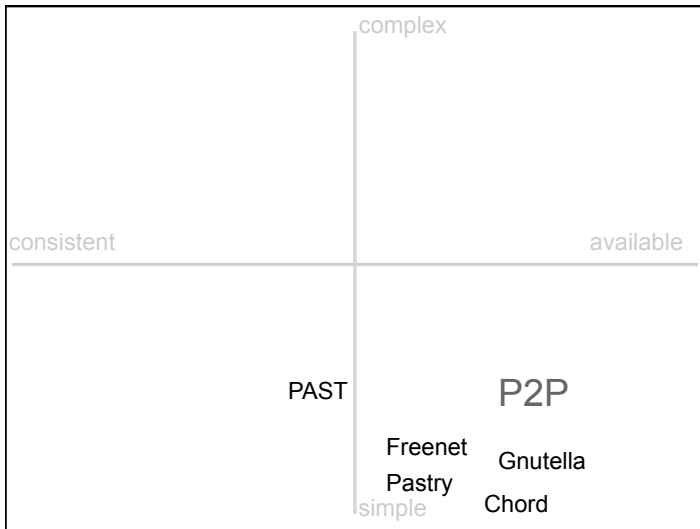
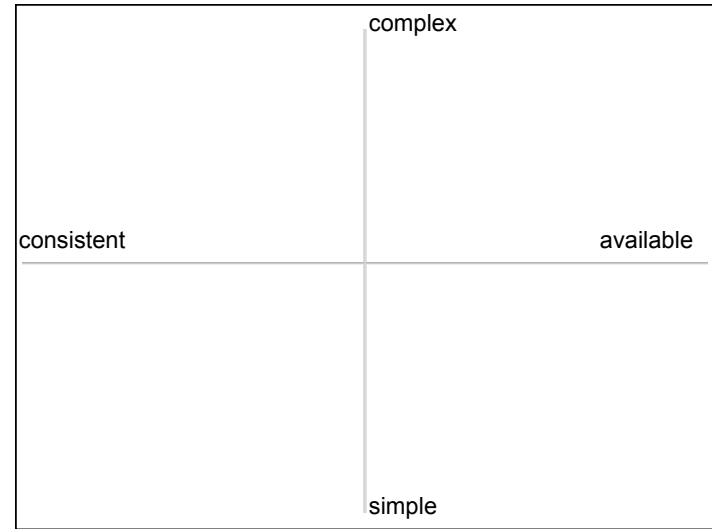
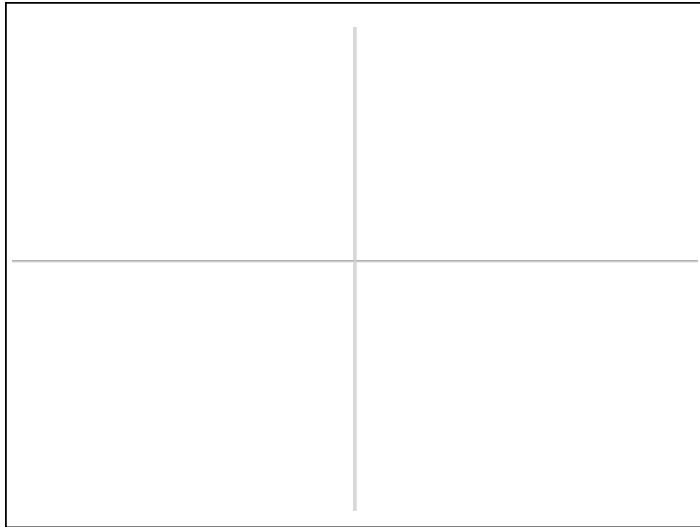
**QW > QR**

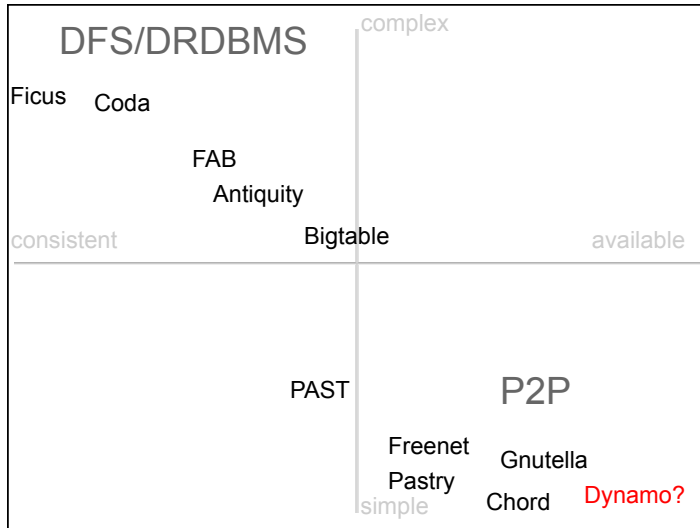
*-> always writeable!*

Data Center      Application  
*last write wins*      *merge?*

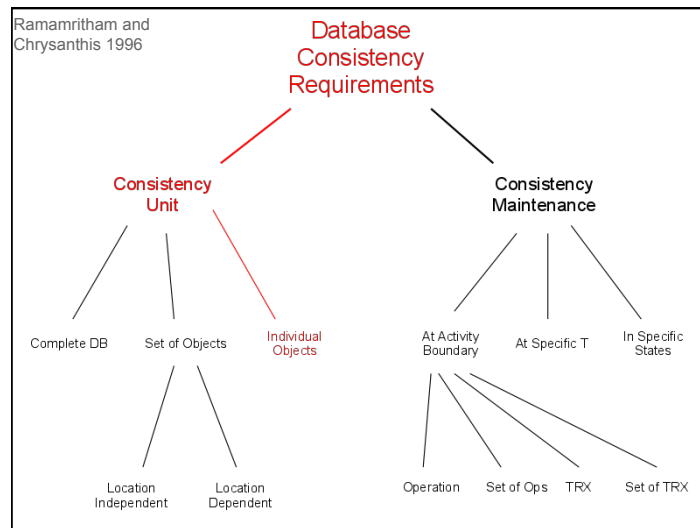
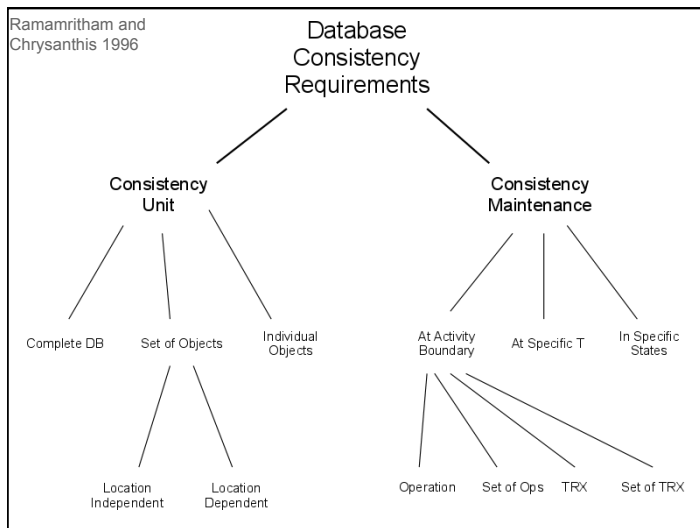
Dynamo

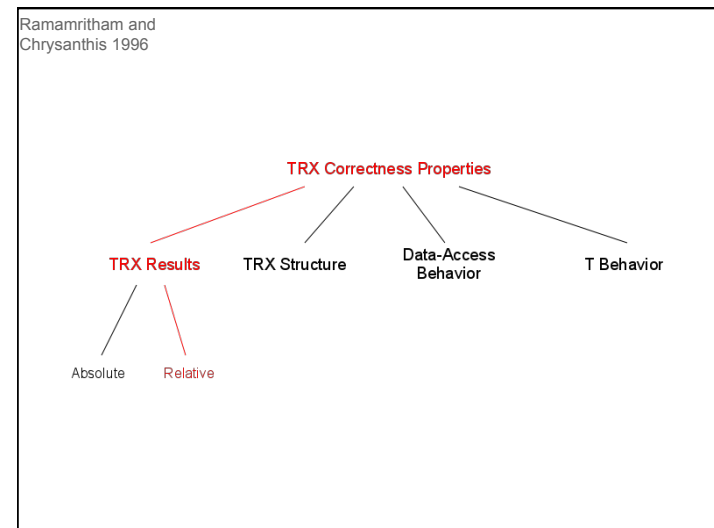
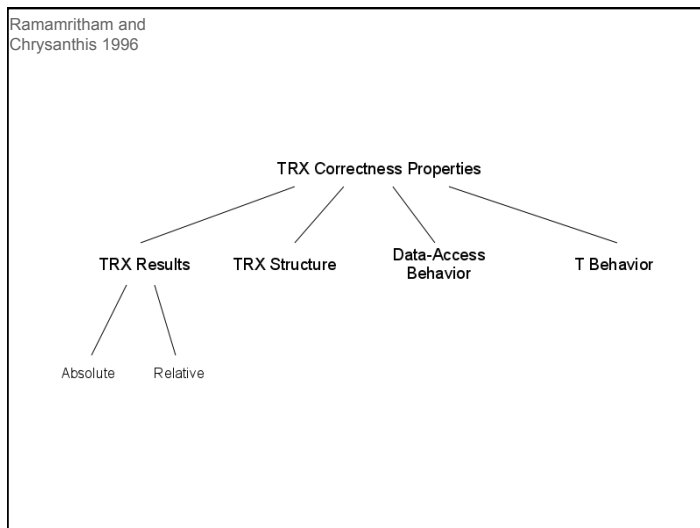
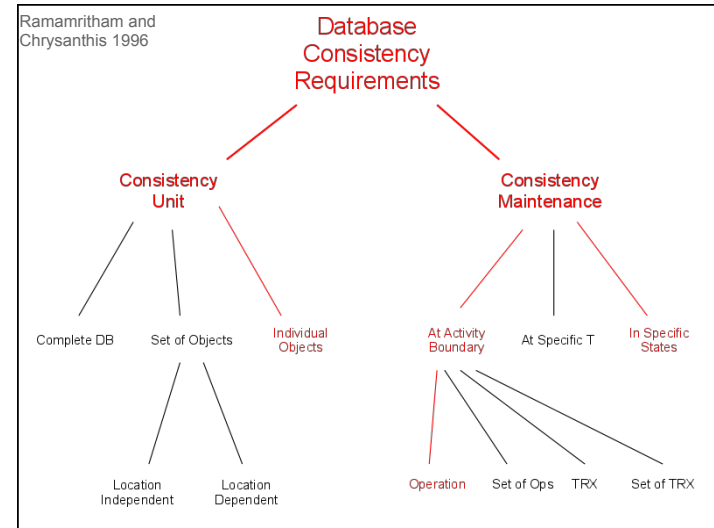
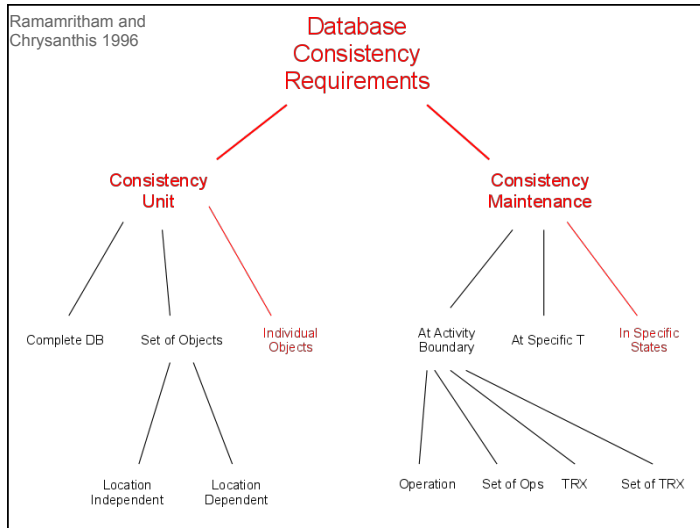
- Prologue: Eventual Consistency
- Act I: 107 / 104
- Act II: 99.9%
- Act III: Related Work
- Entr'Act: Taxonomy of Correctness
- Act IV: System Architecture
- Act VI: Lessons Learned
- Act VII: Conclusion

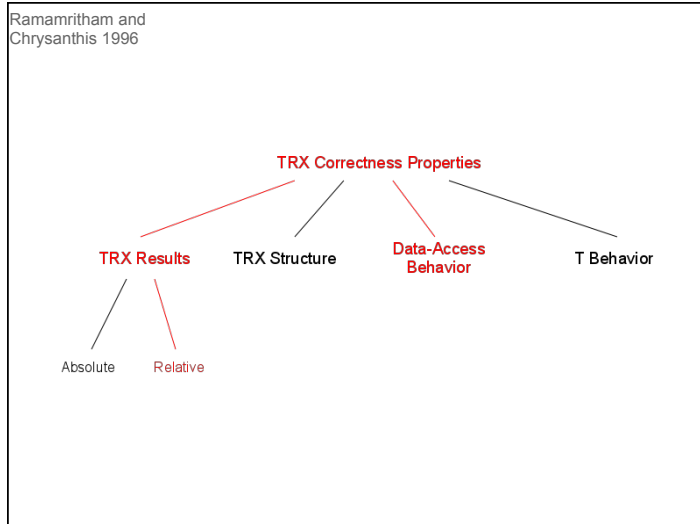




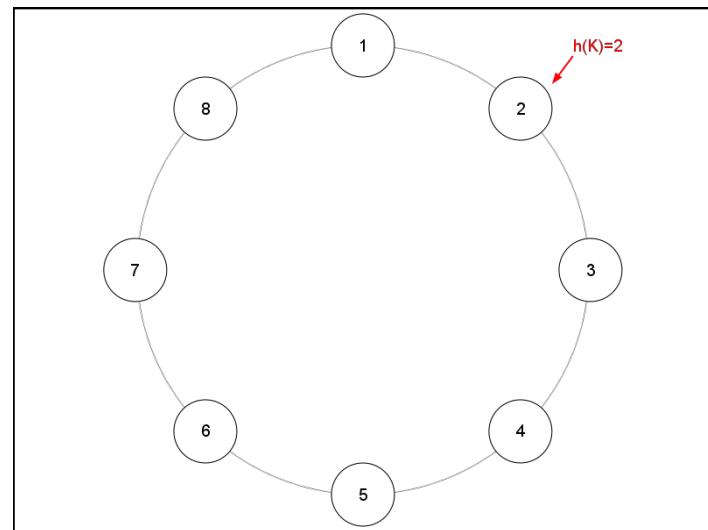
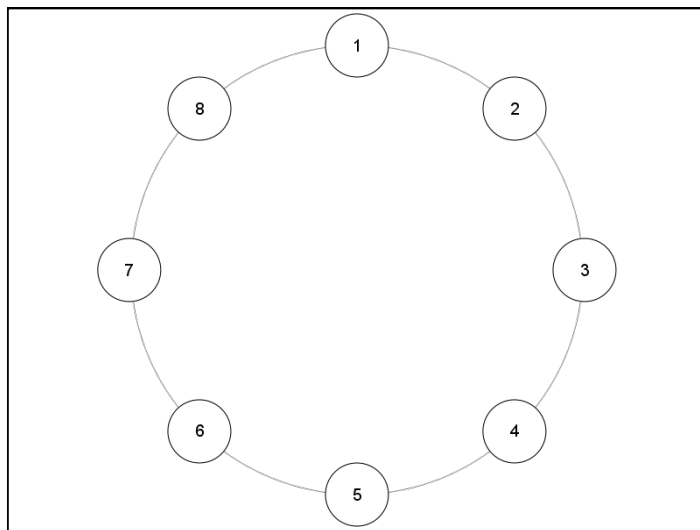
- ### Dynamo
- Prologue: Eventual Consistency
  - Act I: 107 / 104
  - Act II: 99.9%
  - Act III: Related Work
  - **Entr'Act: Taxonomy of Correctness**
  - Act IV: System Architecture
  - Act V: Implementation
  - Act VI: Lessons Learned
  - Act VII: Conclusion

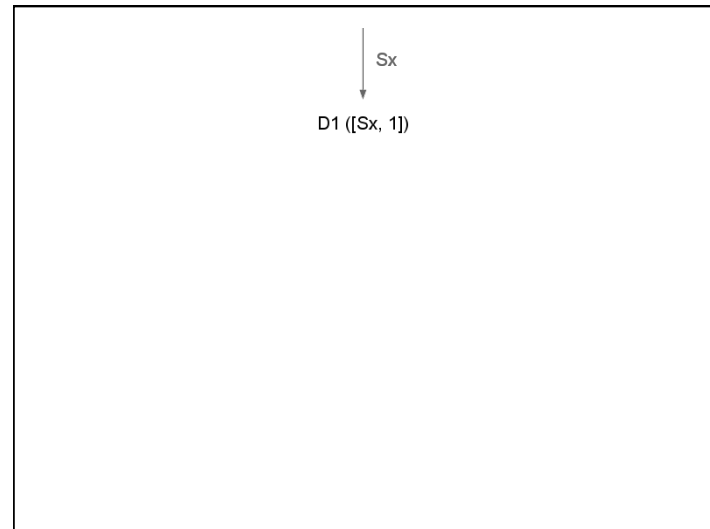
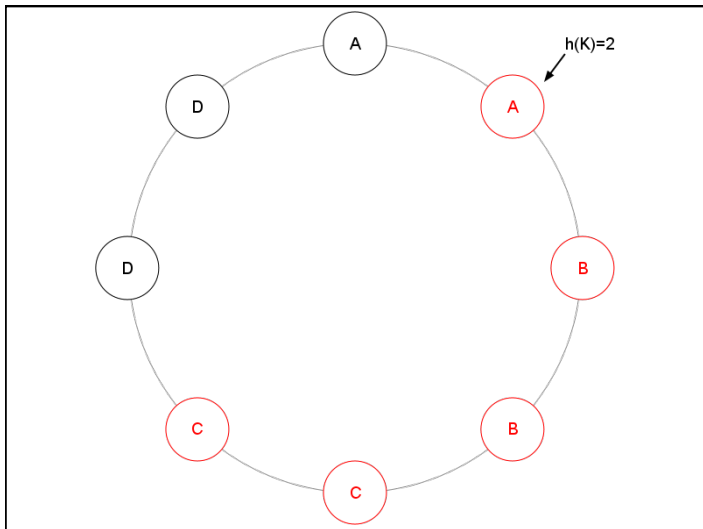
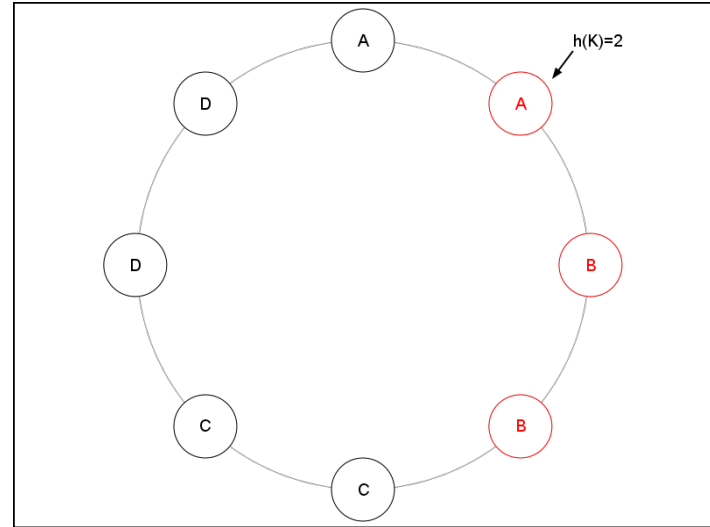
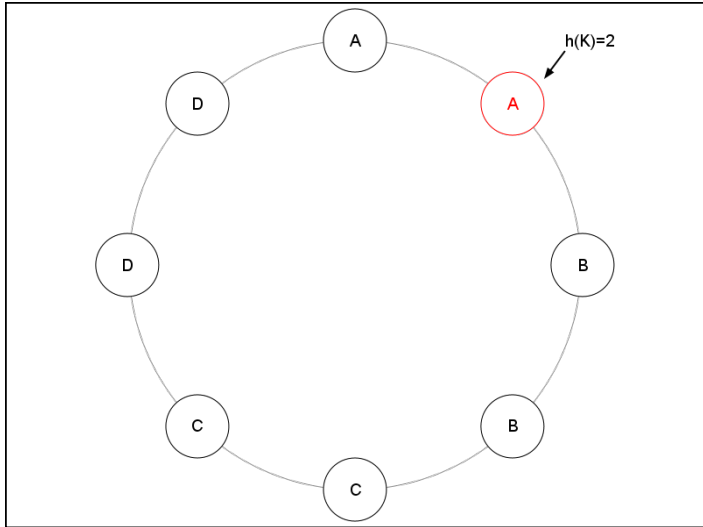


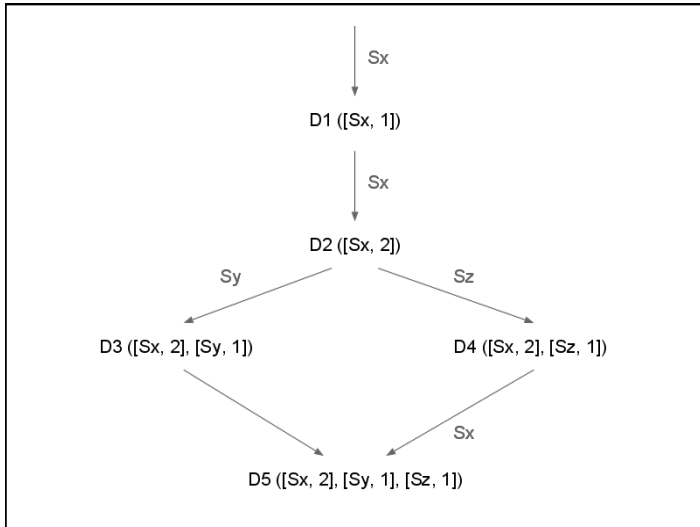
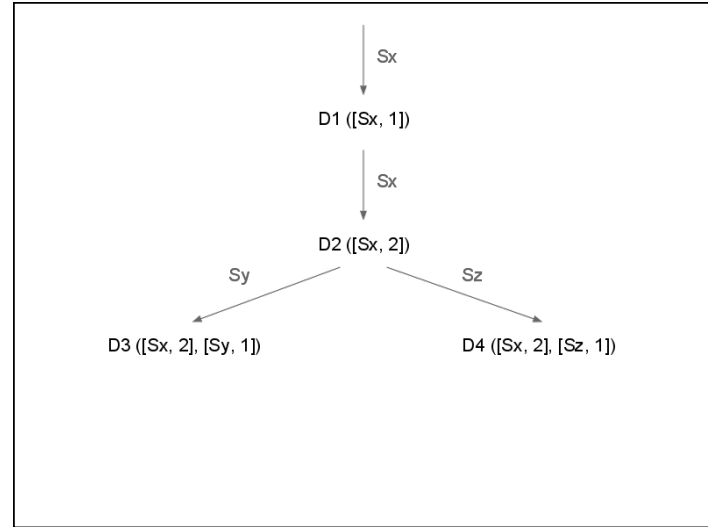
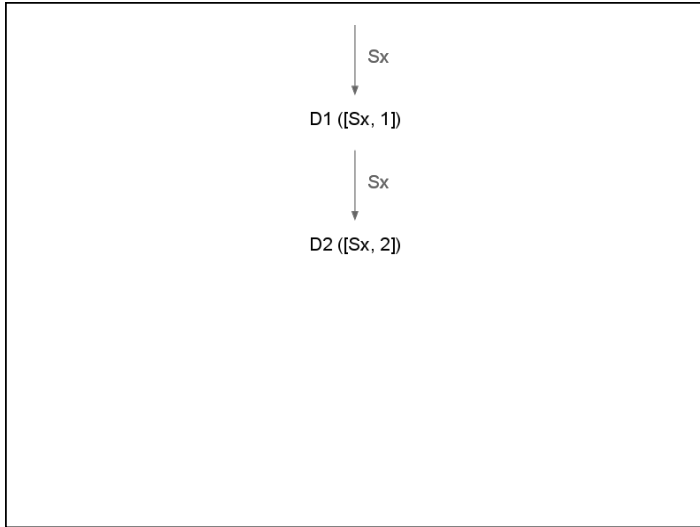




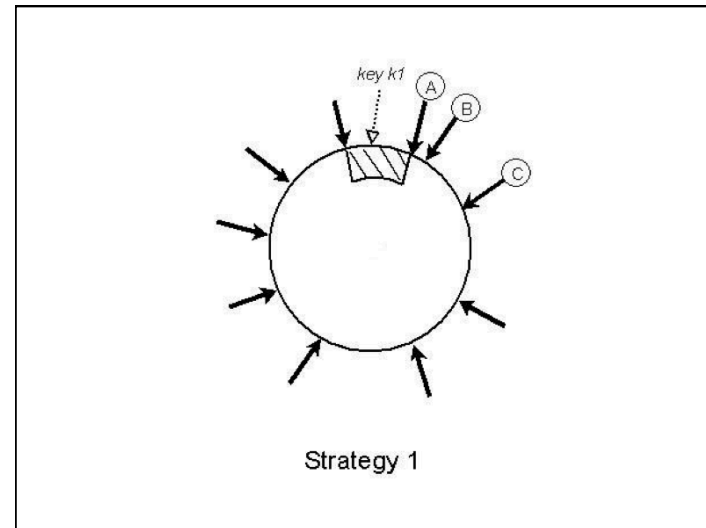
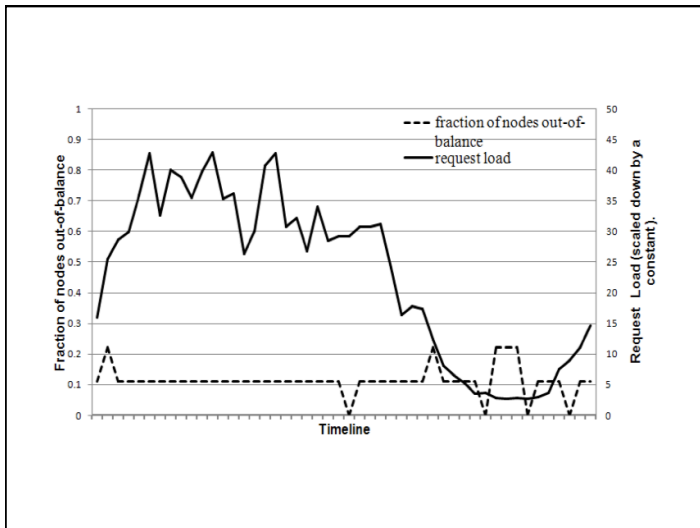
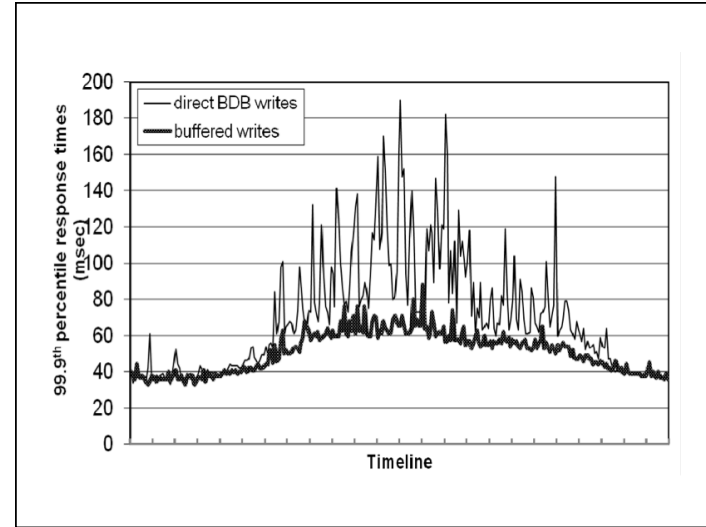
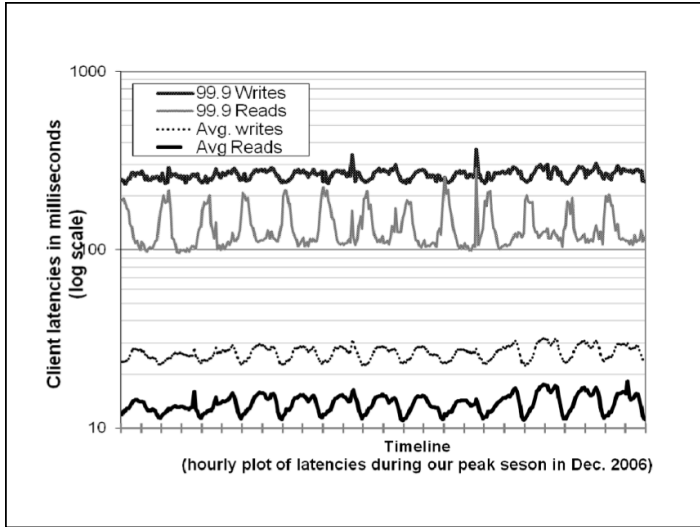
- ### Dynamo
- Prologue: Eventual Consistency
  - Act I: 107 / 104
  - Act II: 99.9%
  - Act III: Related Work
  - Entr'Act: Taxonomy of Correctness
  - **Act IV: System Architecture**
  - Act VI: Lessons Learned
  - Act VII: Conclusion



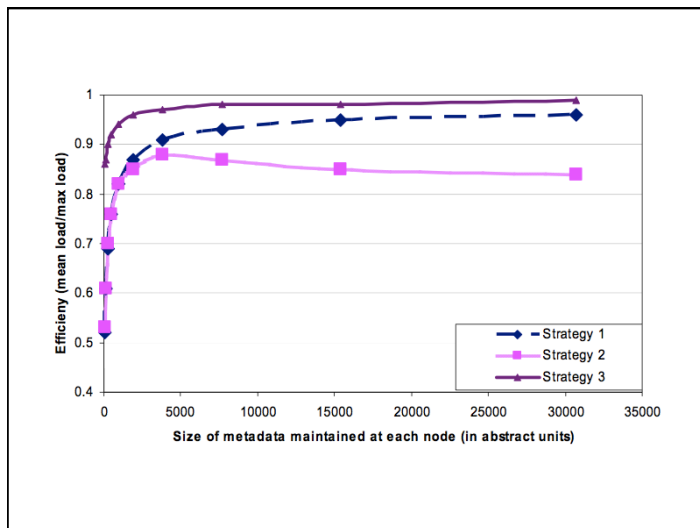
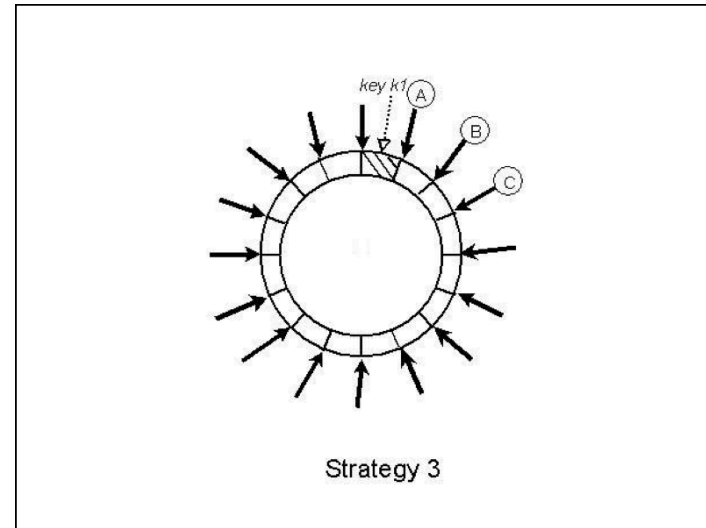
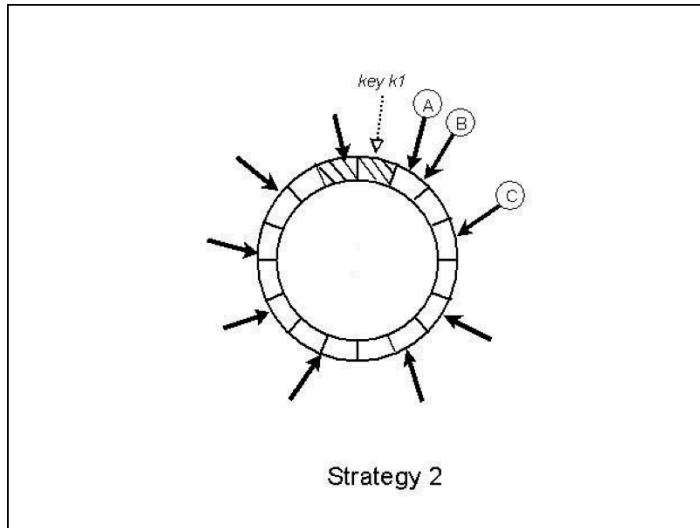




- ### Dynamo
- Prologue: Eventual Consistency
  - Act I: 107 / 104
  - Act II: 99.9%
  - Act III: Related Work
  - Entr'Act: Taxonomy of Correctness
  - Act IV: System Architecture
  - Act VI: Lessons Learned
  - Act VII: Conclusion







	99.9th read latency (ms)	99.9th write latency (ms)	avg. read latency (ms)	avg. write latency (ms)
server-driven	68.9	68.5	3.9	4.02
client-driven	30.4	30.4	1.55	1.9

## Dynamo

- Prologue: Eventual Consistency
- Act I: 107 / 104
- Act II: 99.9%
- Act III: Related Work
- Entr'Act: Taxonomy of Correctness
- Act IV: System Architecture
- Act VI: Lessons Learned
- **Act VII: Conclusion**

