FIS GT.M™ – Multi-purpose Universal NoSQL Database

K.S. Bhaskar
Development Director, FIS
ks.bhaskar@fisglobal.com
+1 (610) 578-4265

Outline

- M – a Universal NoSQL database
- Using GT.M as a Multi-purpose NoSQL database for analysis, reporting data warehousing, etc.

M – a Universal NoSQL database

- Builds on NoSQL discussion of July 3, 2012
  http://www.osehra.org/sites/default/files/sql_vs_nosql_discussion_-_osehra_awg_7-3-2012.ppt
- Concepts and content inspired by & taken from:
  A Universal NoSQL Engine, Using a Tried and Tested Technology
  by Rob Tweed & George James
  http://www.mgateway.com/docs/universalNoSQL.pdf

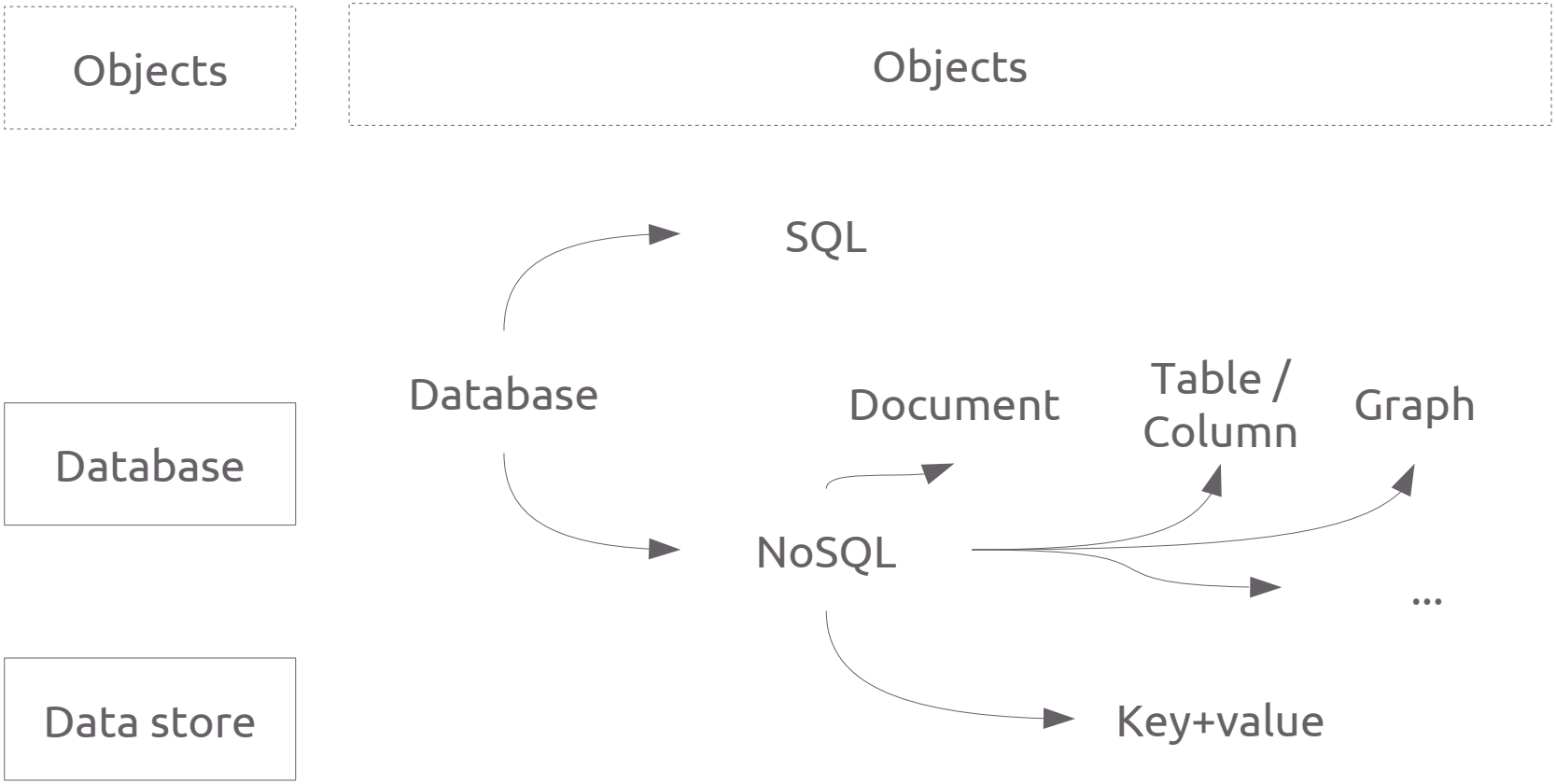
FIS

Traditional View ... Black Box

- Data store – associative memory used for access and retrieval, works for all types of data and association supported
- Database – provides structure and meaning to data in a data store
- Objects – unite code and data
  - (GT.M feature – alias variables for creating containers for objects)

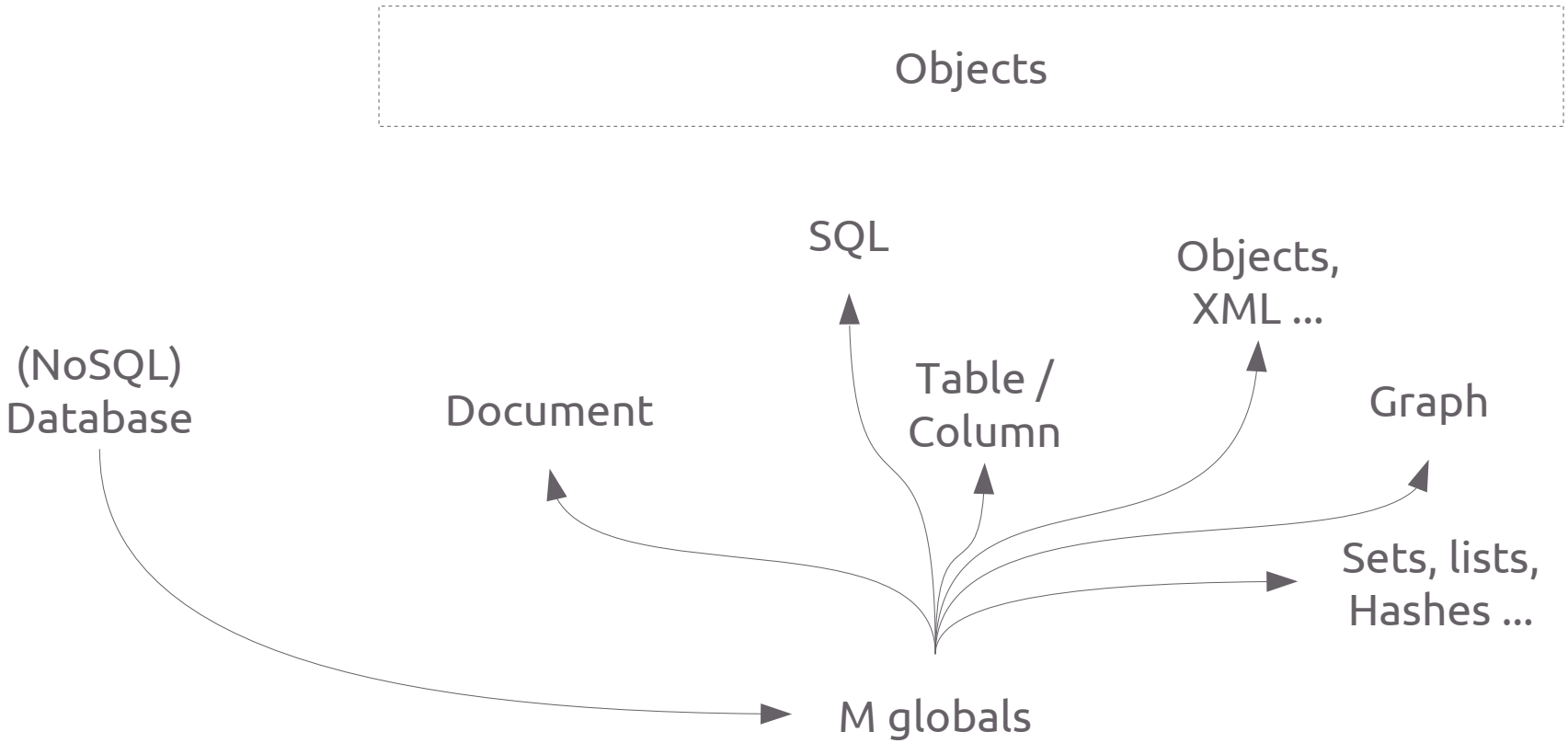
More Contemporary View ... Glass Box

- Distinction between data store and database is blurred



M Universal NoSQL (not only SQL)

- View the same data with the mapping that makes the most sense to the problem domain – you don't need multiple databases



Document Database Example – JSON View

```
{'name':'Rob',
'age':26,
'knows':[
'George',
'John',
'Chris'],
'medications':[
{'drug':'Zolmitripan','dose':'5mg'},
{'drug':'Paracetamol','dose':'500mg'}],
'contact':{
'eMail':'rob@foo.com',
'address':{'street':'112 Beacon Street',
'city':'Boston'},
'telephone':'617-555-1212',
'cell':'617-555-1761'},
'sex':'Male'
}
```

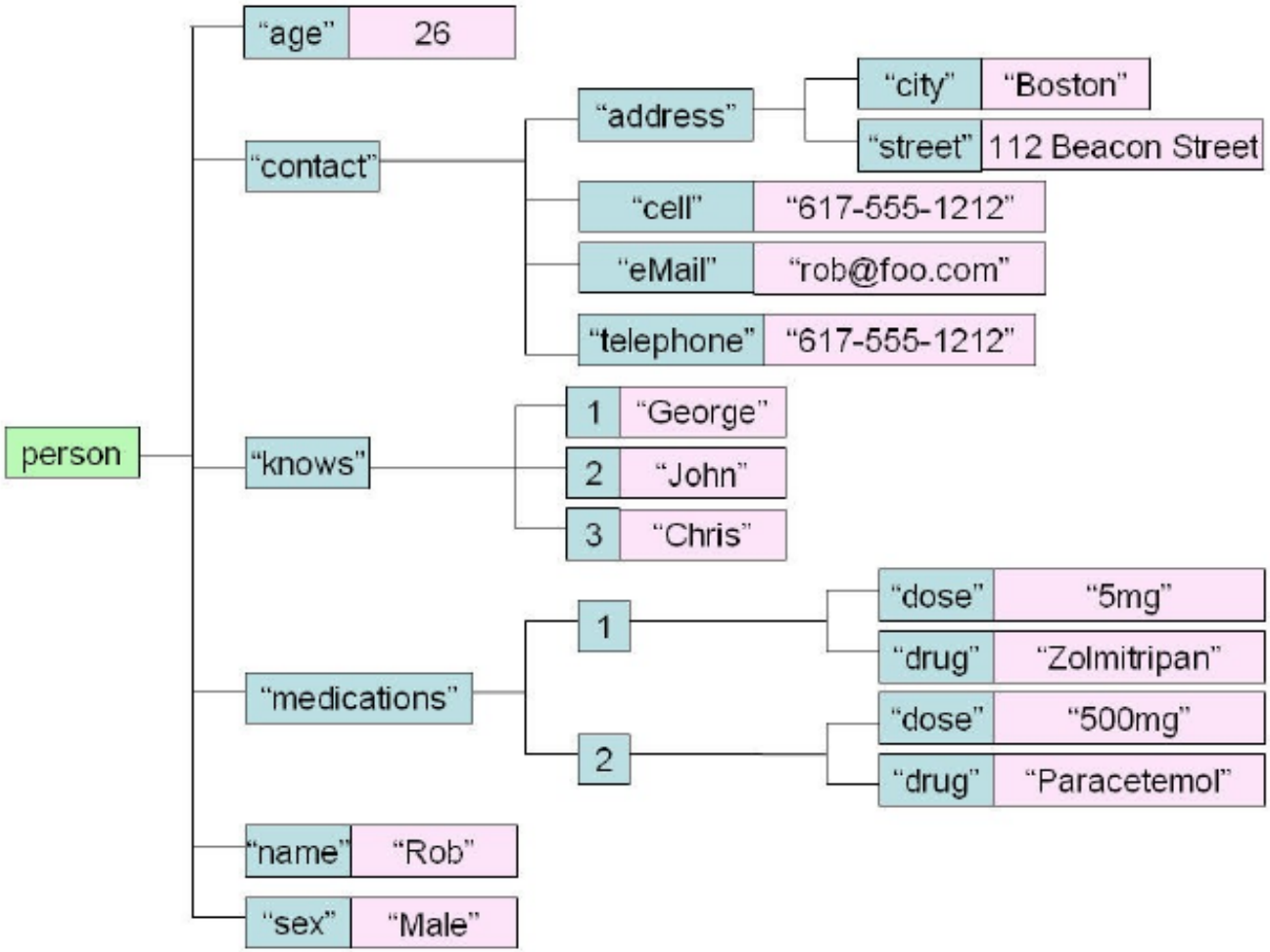
FIS

Document Database Example – Global view

```
person("age")=26
person("contact","address","city")="Boston"
person("contact","address","street")="112 Beacon Street"
person("contact","cell")="617-555-1761"
person("contact","eMail")="rob@foo.com"
person("contact","telephone")="617-555-1212"
person("knows",1)="George"
person("knows",2)="John"
person("knows",3)="Chris"
person("medications",1,"drug")="Zolmitripan"
person("medications",1,"dose")="5mg"
person("medications",2,"drug")="Paracetamol"
person("medications",2,"dose")="500mg"
person("name")="Rob"
person("sex")="Male"
```

FIS

Document Database Example – 1000 Words



Summary – Universal NoSQL

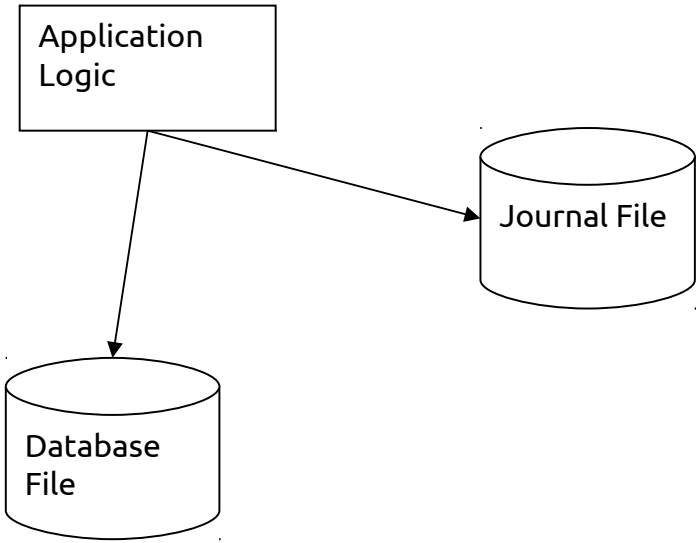
- At the heart of all databases is a data store
- Layered code provides all other interpretations of data, as well as schema management, objects, etc.
  - Black box (proprietary databases)
  - White box (e.g., M/Wire, M/DB:X, FM Projection)
- Transaction processing need not use the same data model as analytic processing
  - Multiple data models coexist as views of the same data
- Common to *all* M implementations!

GT.M Transactional & Analytical Processing

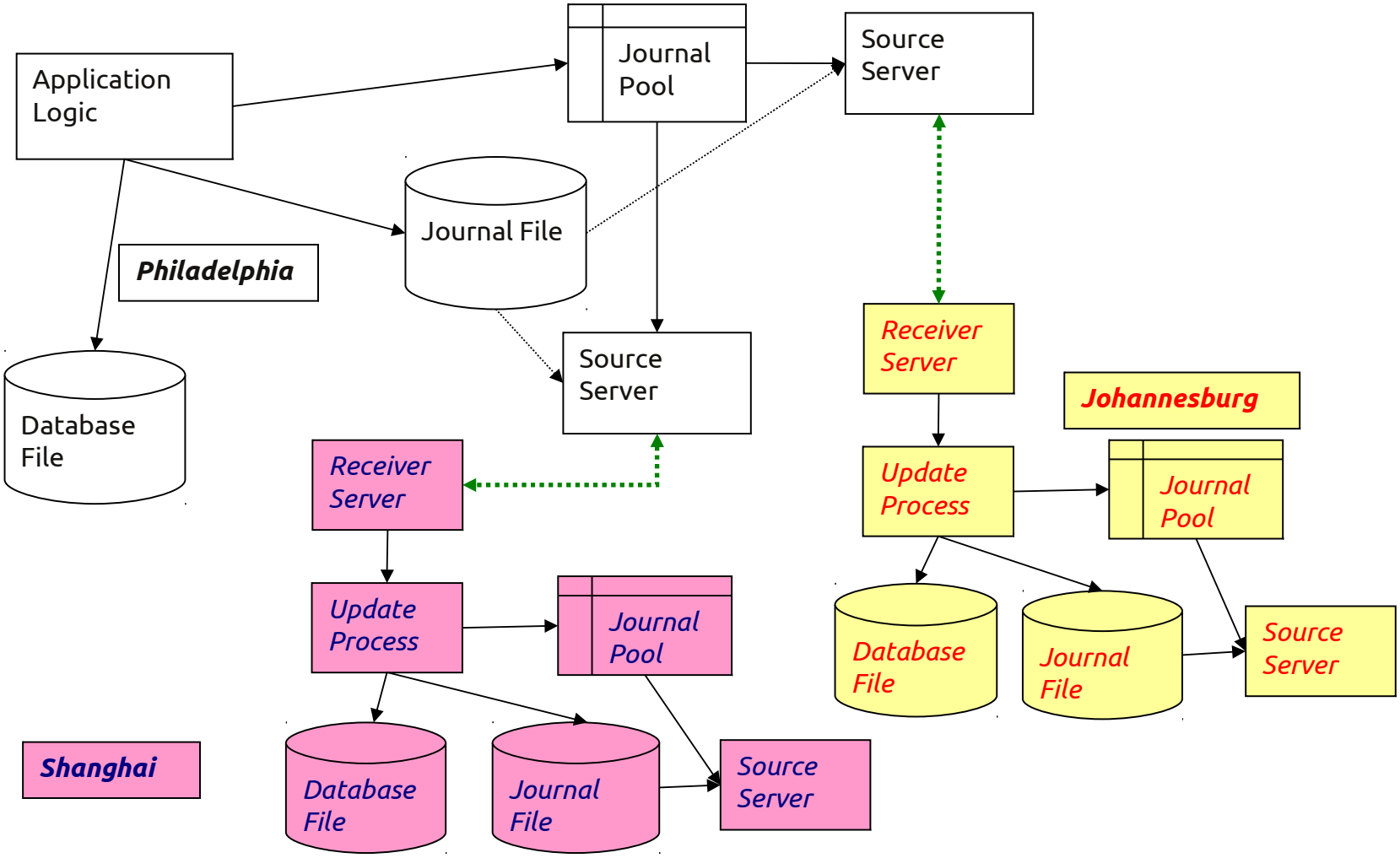
- Common use case – transactional system feeds analytical system
- Traditional implementation – extract, transform, load (ETL)
- "The GT.M way" – real-time replication
  - Originally designed for business continuity ("BC replication") – mature technology, in production since 1999 and regularly enhanced since
  - Now available for real-time feeds for reporting, data warehousing, research, etc. ("SI replication")

FIS

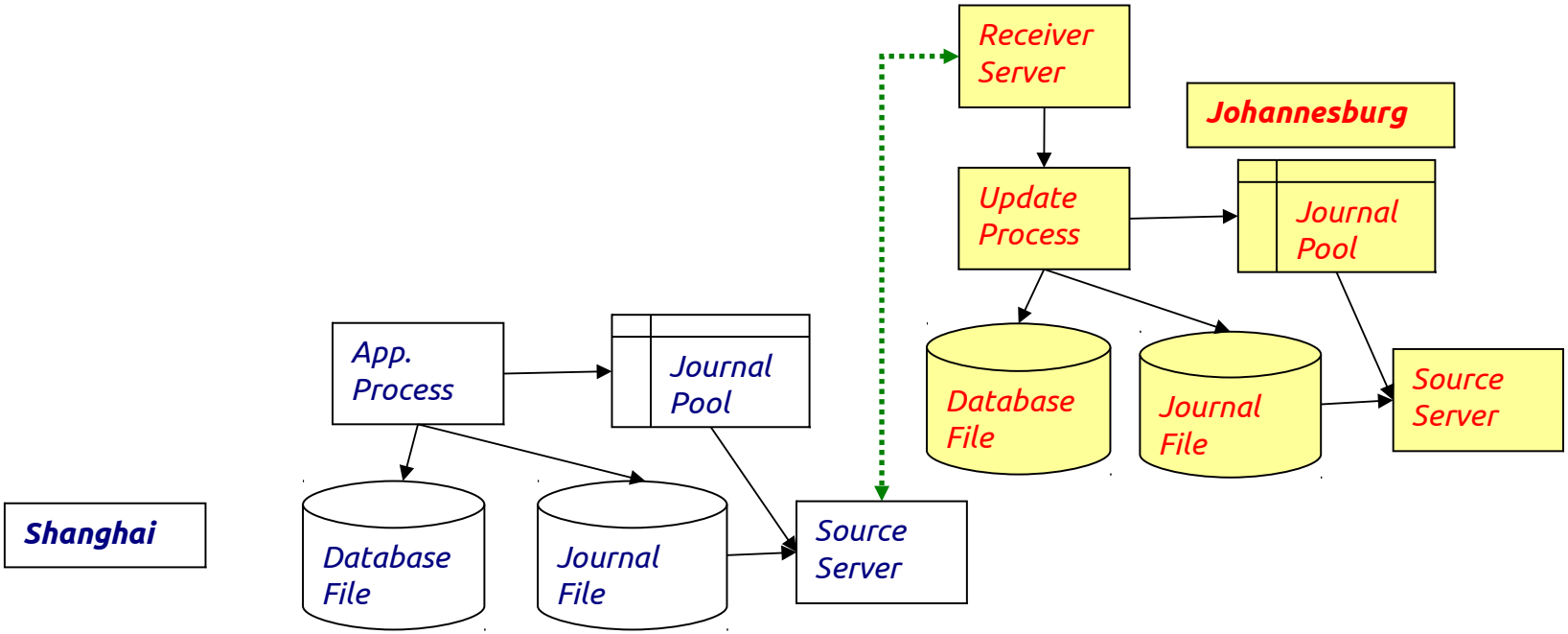
Single Site



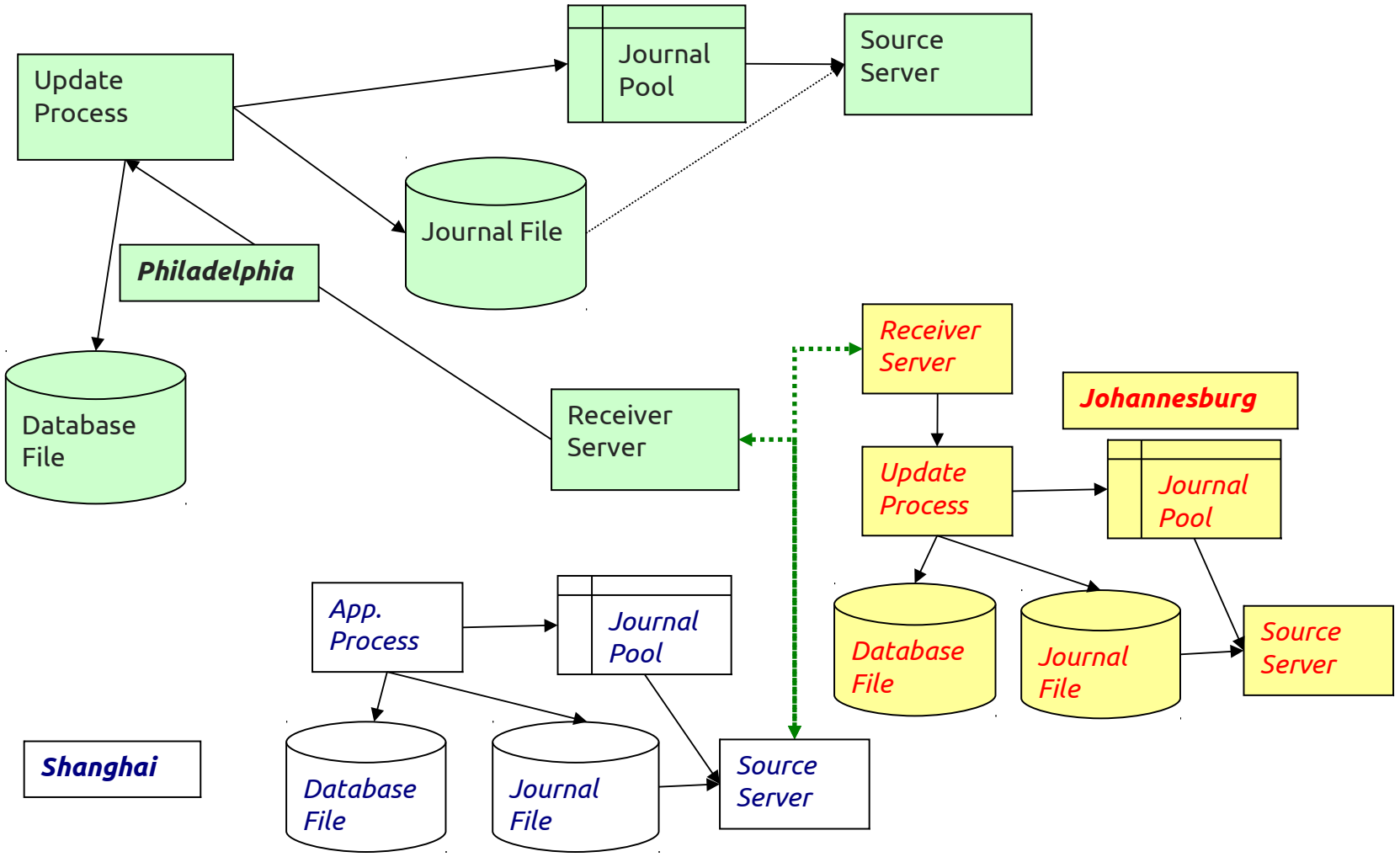
Logical Multi-Site (LMS)



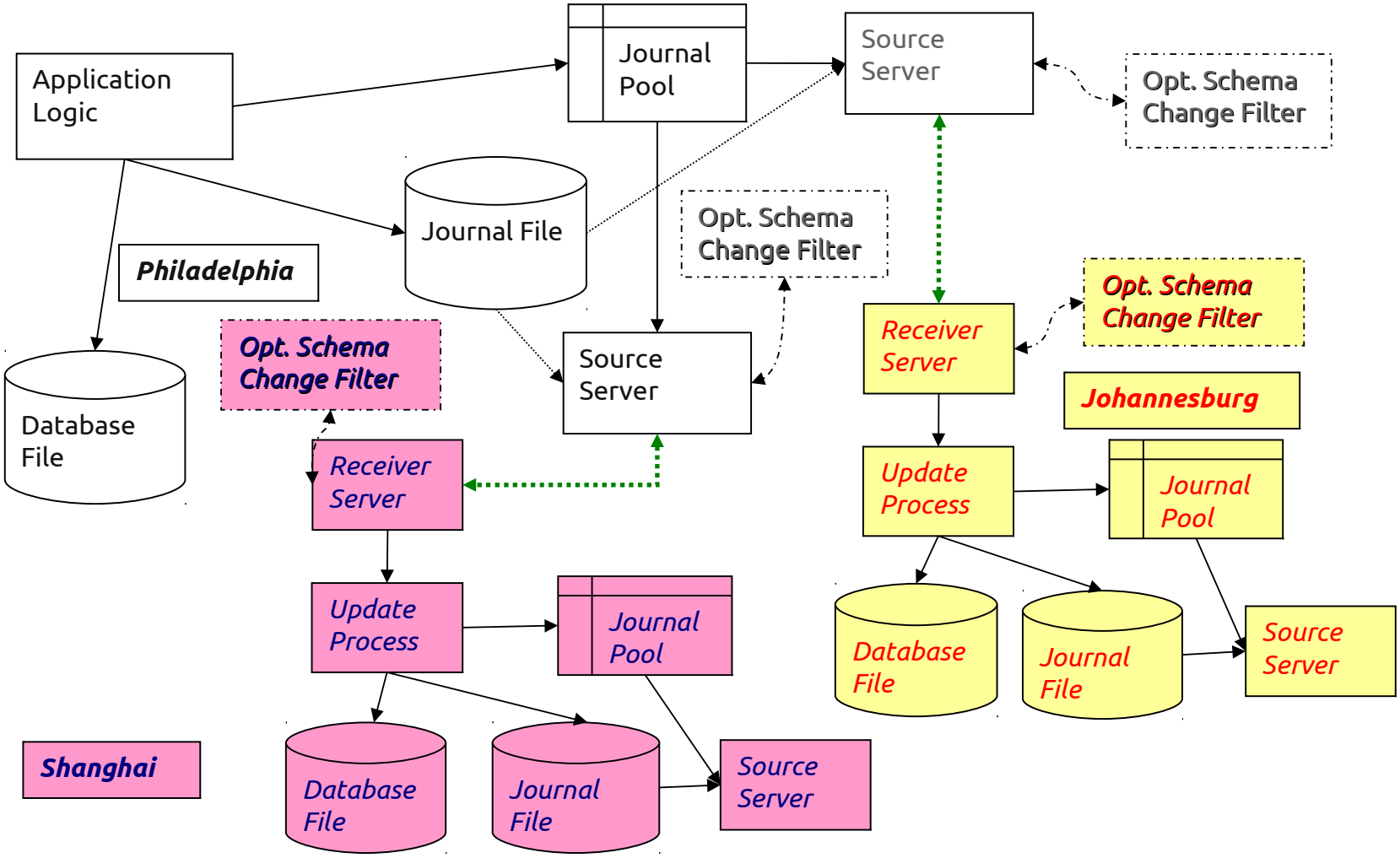
Philadelphia Down



Philadelphia Recovers



Rolling Upgrades with Schema Changes



Rolling Upgrades

Site Ardmore	Site Bryn Mawr
P: Old software	S: Old Software
P: Continue processing	Take down and upgrade
P: Continue processing	S: With old to new filter
S: Continue operation	P: With new to old filter
Take down and upgrade	P: Continue processing
S: New software	P: New software, no filter

BC Replication & the CAP Theorem

- Consistency, Availability, Partition tolerance – pick any two
  - Real world systems are AP (Availability + Partition tolerance)
  - Must restore Consistency after Partition event (“Eventual Consistency”)
- Eventual Consistency requirement
  - Traditional – all nodes (instances) reach the same state
  - Financial application requirement is that all nodes must eventually have *the same path through state space*, not just the same state, with Consistency (as in the ACID property) *at each point* – achieved by giving each transaction a unique serial number and performing rollback / reapply as needed to restore Eventual Consistency while ensuring (ACID) Consistency
    - Requires cooperation from application code

FIS

Unique Serial Numbers for BC Replication

Site Ardmore	Site Bryn Mawr
P: 100	S: 95
Goes down	P: 96 ...
Recovers	P: ...110 ...
S: rollback 95 ... 100 (send to Bryn Mawr for reprocessing)	P: ... 120
	P: 125
S: 125	P: 130

Replication

- 1 primary instance, replicating to
- 16 secondary instances, replicating to
- 256 tertiary instances, replicating to ...

- Any BC instance can in principle take over as a primary instance

(Just because every instance can doesn't mean that any instance should –network considerations guide choice)

BC Replication Limitation

- No locally generated updates allowed on replicating secondary instances
  - Updates have unique serial numbers across all systems
- Required to ensure consistent results from business logic

SI Replication for Analytical Processing (AP)

- Databases for AP need different content from transactional database
  - Additional cross references
  - Statistics
  - Reporting
  - Demographics – house-holding, population analytics
  - Data scrubbing
- Provides an originating primary instance that can receive a replication stream
  - Unlike BC replication, direction not reversible with SI replication
- Works with BC replication to provide business continuity of AP functions
  - Updates tagged with origin

SI & Eventual Consistency ... rollback

Site Ardmore	Site Bryn Mawr	Site Malvern
O: ... A97, A98, A99	R: ... A97	S: ... A97, M37, M38, A98, M39, M40
Goes down	O: ... A97	... A97, M37, M38, A98, M39, M40 (rollback performed)
	O: ... A97, B61, B62	S: ... A97, M37, M38, B61
	O: ... A97, B61, B62, B63, B64	S: ... A97, M37, M38, B61, B62, M39a, M40a, B63

FIS

SI & Eventual Consistency ... no rollback

Site Ardmore	Site Bryn Mawr	Site Malvern
O: ... A97, A98, A99	R: ... A97	S: ... A97, M37, M38, A98, M39, M40
Goes down	O: ... A97, B61, B62	... A97, M37, M38, A98, M39, M40 (no rollback)
	O: ... A97, B61, B62	S: ... A97, M37, M38, A98, M39, M40, B61, B62

FIS

Techniques for Creating AP Databases

- Batch (scheduled) – statistical calculations, e.g, regressions; data mining, e.g., hierarchical clustering
- Real time
  - Replication filters (previously discussed; useful for BC & SI replication)
  - Triggers (useful for single instance, BC & SI replication)

Trigger Example Need

- Global nodes in `^CIF(ACN,1)=NAM|XNAME|...` where `XNAME` is a canonical name (e.g., `"Doe,John"`).
- Cross reference index, `^XALPHA("A",XNAME,ACN)=""`
- Would like to not depend on application programmers to maintain consistency

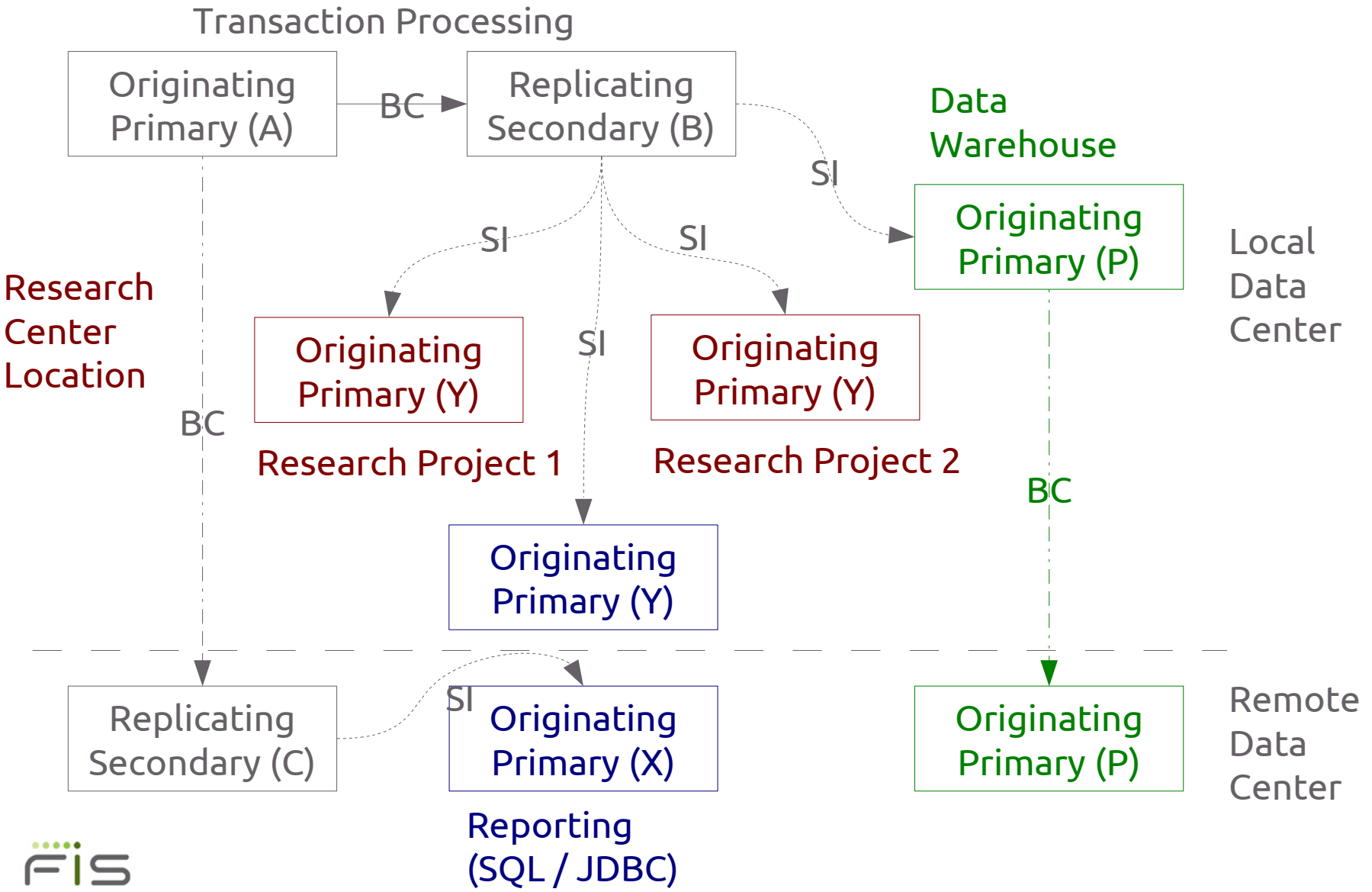
Trigger Example Definition

```
^CIF(acn=:,1) -delim="|" -pieces=2 -commands=SET,KILL -xecute="Do ^XNAMEinCIF"
```

Trigger Example Code

```
XNAMEinCIF ; Triggered Update for XNAME change in ^CIF(acn=:,1)

    ; Get old XNAME - $zchar(254) used in ^XAPLHA for null XNAME in ^ACN
    Set oldxname=$Piece($ZTOLDval,"|",2)
    Set:'$Length(oldxname) oldxname=$zchar(254)

    ; Remove any old xref
    Kill ^XALPHA("A",oldxname,acn); remove any old xref

   ; Create new cross reference if command is Set
    Do:$ZTRIggerop="S"
    . Set xname=$Piece($ZTVALue,"|",2) Set:'$Length(xname) xname=$zchar(254)
    . Set^XALPHA("A",xname,acn)=""

   Quit
```

Sample Configuration



Performance Considerations

- First approximation
  - IO throughput of receiver needs to match that of source
  - CPU and RAM requirements are modest
- Refinements
  - Peak vs. average IO throughput
  - Journaling: two types, with different trade-offs in throughput vs. MTTR
  - Increase / decrease in data volume
  - Local processing needs

Main Tuning Parameters

- Database partitioning
- Block size
- Access Method (BG vs. MM) & dependent parameters
  - Global Buffers (BG only)
  - Journaling type
- Lock space

Questions / Discussion

K.S. Bhaskar
ks.bhaskar@fisglobal.com
+1 (610) 578-4265
http://fis-gtm.com