Connecting people with the
world's greatest travel possibilities.

Data Management at Sabre

Glenn Harper
Chief Infrastructure Architect

Janine Stevens
Enterprise Data Architect

May 2009

Sabre Data in Motion

32,000

Transactions – **every second**

1 million

Transactions **every minute**

500 million+

Transactions **every day**

65 million+

Transactions ever day
via **Web Services**

1

Number one provider of
Travel Technology
products and services
On the Planet

100,000+

Agency Computers
accessing Sabre travel
content globally

24 hours

Every day – **7 days** a week

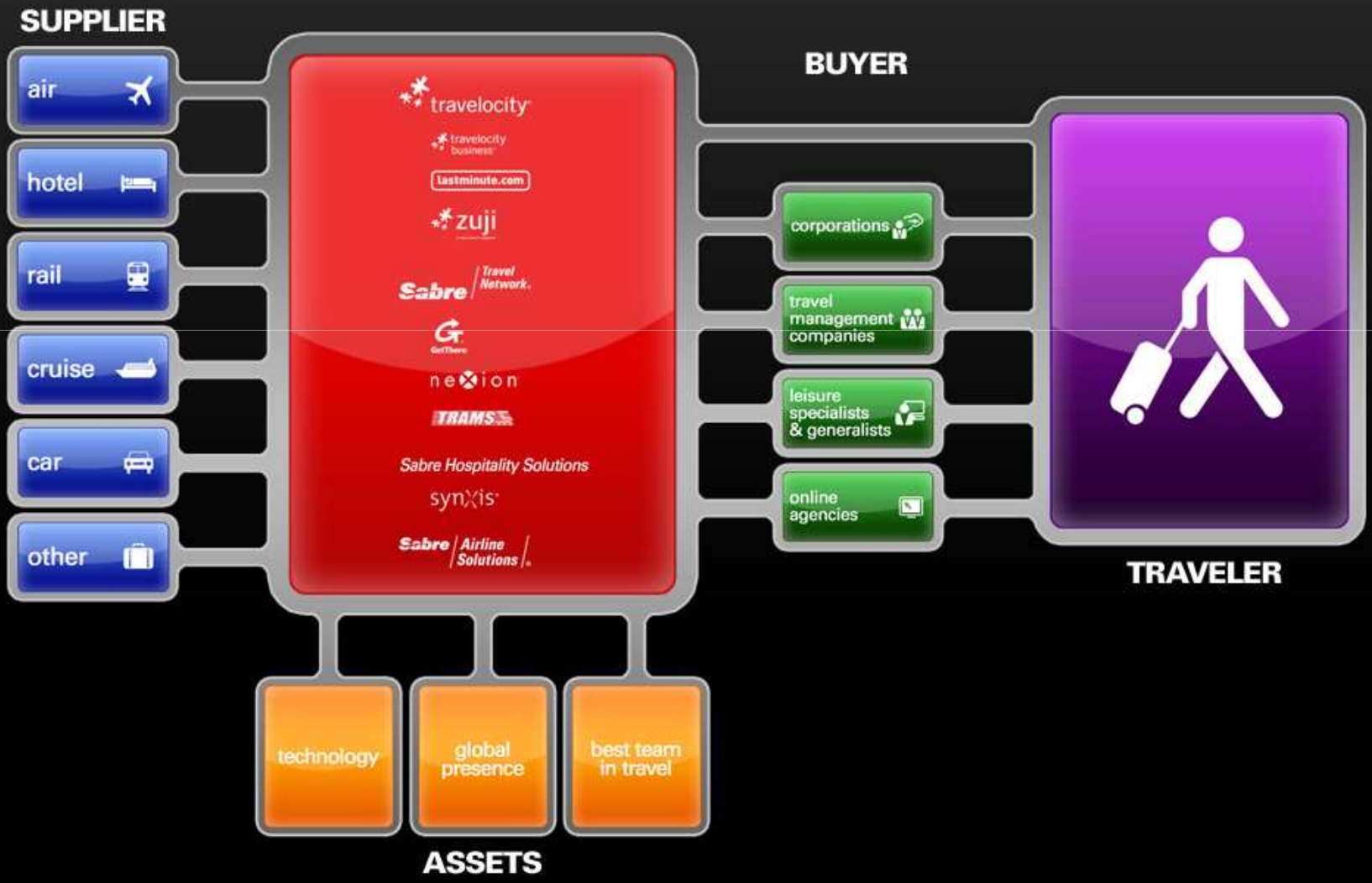
It's always peak hour
somewhere in the world

Sabre Holdings

Our Key Brands

Sabre Holdings

SABRE HOLDINGS



Sabre Worldwide

- More than 9,000 employees in 59 countries
- Headquarters in Southlake TX

Top 10 Office Locations

1. Dallas/Fort Worth, Texas
2. Montevideo, Uruguay
3. London, England
4. Krakow, Poland
5. Woking, England
6. San Antonio, Texas
7. Bangalore, India
8. Plains, Pennsylvania
9. Sydney, Australia
10. Singapore

Sabre History

1960

- **American Airlines and IBM introduced the first passenger reservation system. Semi-Automated Business Research Environment, SABRE in Briarcliff Manor, NY.**

1970

- 1972 - The Sabre system is moved to a new state of the art data center in Tulsa, OK.
- 1976 - The first Travel Agency implements automation using Sabre

1980

- 1980 - Introduction of easySabre allows consumer using personal computers to access Sabre online for airline, hotel and car rentals.
- 1986 - Sabre Airline Solution release the industry's first revenue management system
- 1988 - Sabre Airline Solutions begins providing software, consulting and system management services to other airlines .

1990

- 1992 - Sabre introduces Sabre Airflite flight scheduling system
- 1994 - Sabre and SNCF install RESARAIL rail reservation and distribution system for TGV network
- 1996 - Sabre launches Travelocity.com
- 1998 - Forms a joint venture with ABACUS International to establish SabreSonic
- 1998 - Introduce Best Fare Finder pricing .

Sabre History

2000

- Sabre spins off from American Airlines parent company.
- Purchase Get There an online corporate booking tool.
- Introduced Sabre eVoya Webtop next generation agency tech tools.

2001

- Purchase travel distribution business in Pacific alliance with Quantas, Air New Zealand and Anssett Airlines.
- Sabre Airline Solutions introduces Sabre Aerodynamic Traveler passenger processing solution designed to expedite check-in. (Use of curbside check-in, roving agents, self-serve kiosks).

2002

- Travelocity acquires Site 59.

2003

- Travelocity introduces TotalTrip new packaging capabilities.

2004

- Sabre launches MySabre a new Web based agent booking portal.
- Travelocity launches it's French consumer website 'Odysia'.

2005

- Sabre acquires, SynXis, IgoUgo and Lastminute.com.
- Travelocity introduces Zuji leading online travel agency in the Asia Pacific.

2007

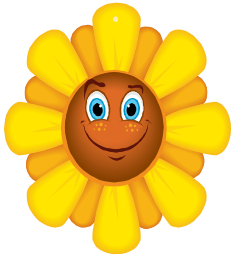
- Sabre is acquired by Silver Lake and TPG.
- Sabre acquires E-site Marketing which will be a division of SynXis.

Innovation

- **In 2006, *Information Week* named SABRE seventh on its list of the "Greatest Software Ever Written"**

- **Sabre Holdings invests more in research and development than any other travel commerce company.**
  - Travel Studio is the hot bed of travel commerce technology innovation.
  - Sabre Airline Solution's Airline Yield Management program is the gold standard used by major airlines to price their fares

But lest you think it is all roses and sunshine…..



……Let's talk Data Management!

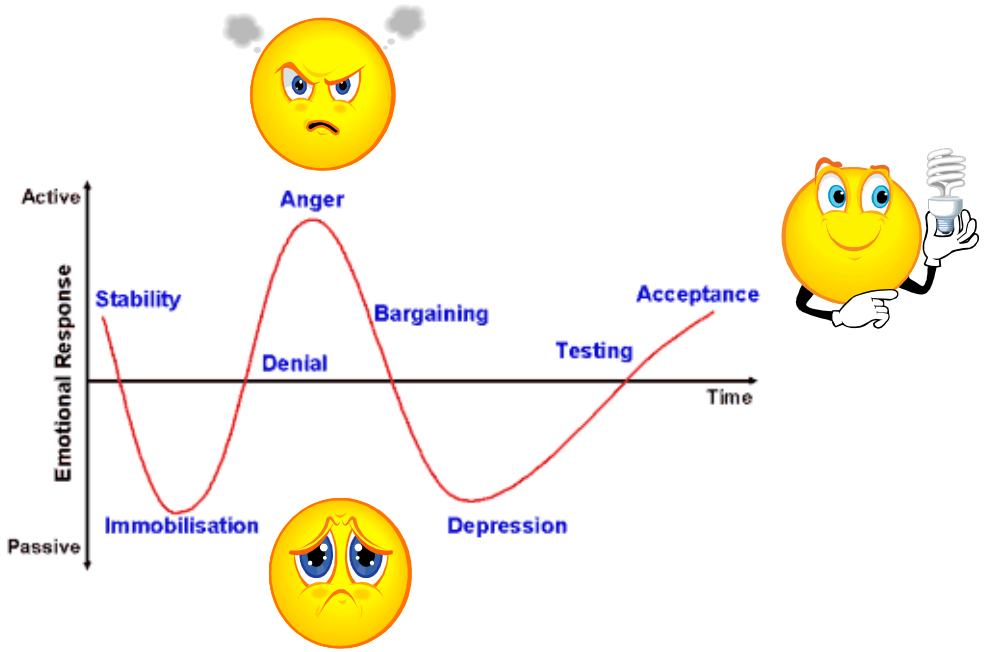
Innovation means change

- **Sabre created a centralized data management team to manage the data for the organization.**
  - It was such a good idea that we centralized and then de-centralized at least 3 times over the last 10 years.
- **We learned that having a centralized team to manage all the models did not work in our environment.**
- **Challenges encountered:**
  - It was clearly us (central team) vs. them (project team) culture
  - Agile development was introduced and all but killed data modeling
  - The project teams wanted more from their data modelers
  - More and more teams were remote and a centralized time zone was difficult to manage
  - Perceived business value was diminished

Change cycle model on the “grief cycle”

- **Sometimes working with teams can be much like the grief cycle.**
- **When people have limited or no input into change, they may feel out of control and exhibit attributes of the grief cycle.**
- **As we drive technology change, the way it is decided and communicated will effect the way teams respond.**

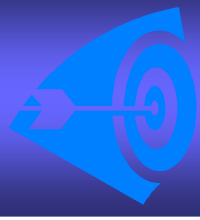
- Shock stage
- Denial stage
- Anger stage
- Bargaining stage
- Depression stage
- Testing stage
- Acceptance stage



Data Management

- **Preparing for the Shift to happen**
  - Known concerns/issues that needed to be addressed to keep a discipline in data modeling.
    - Ensure that expertise in data modeling would not be lost
    - Make certain that the knowledge of the case tool would be sustained
    - Address communication challenges due to globalization and remote users
- **Introduced an internal Data Modeler User Group**
  - We meet monthly to communicate ideas, concepts or discussions on modeling, tips and tricks using the ER tool, problems etc.
  - We introduced Master Series topics regarding data modeling best practices or specific solutions to real problems.
  - We Created training materials for both class instruction and self study for using the case tool and for foundational Data Modeling concepts

Data Management

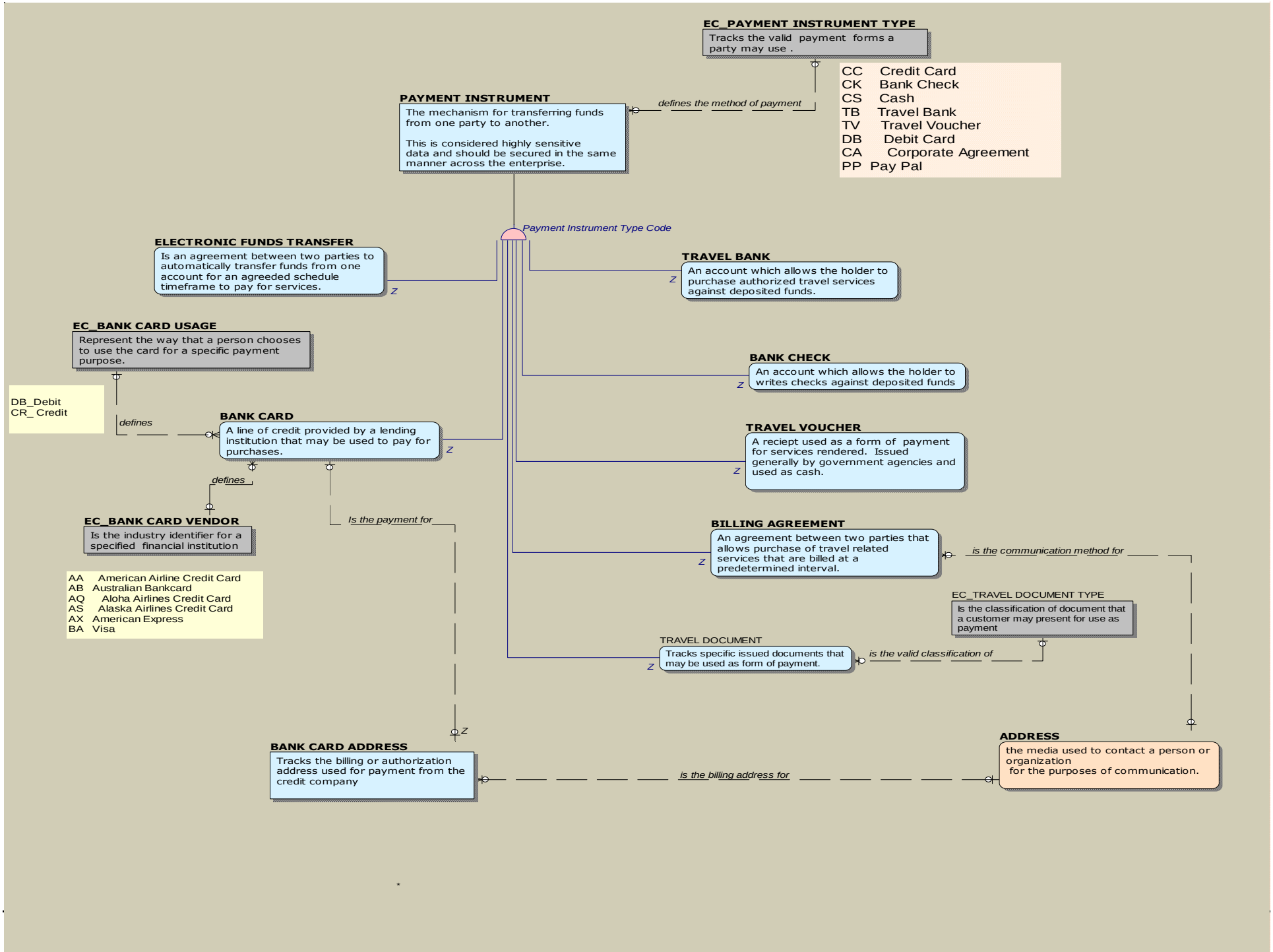


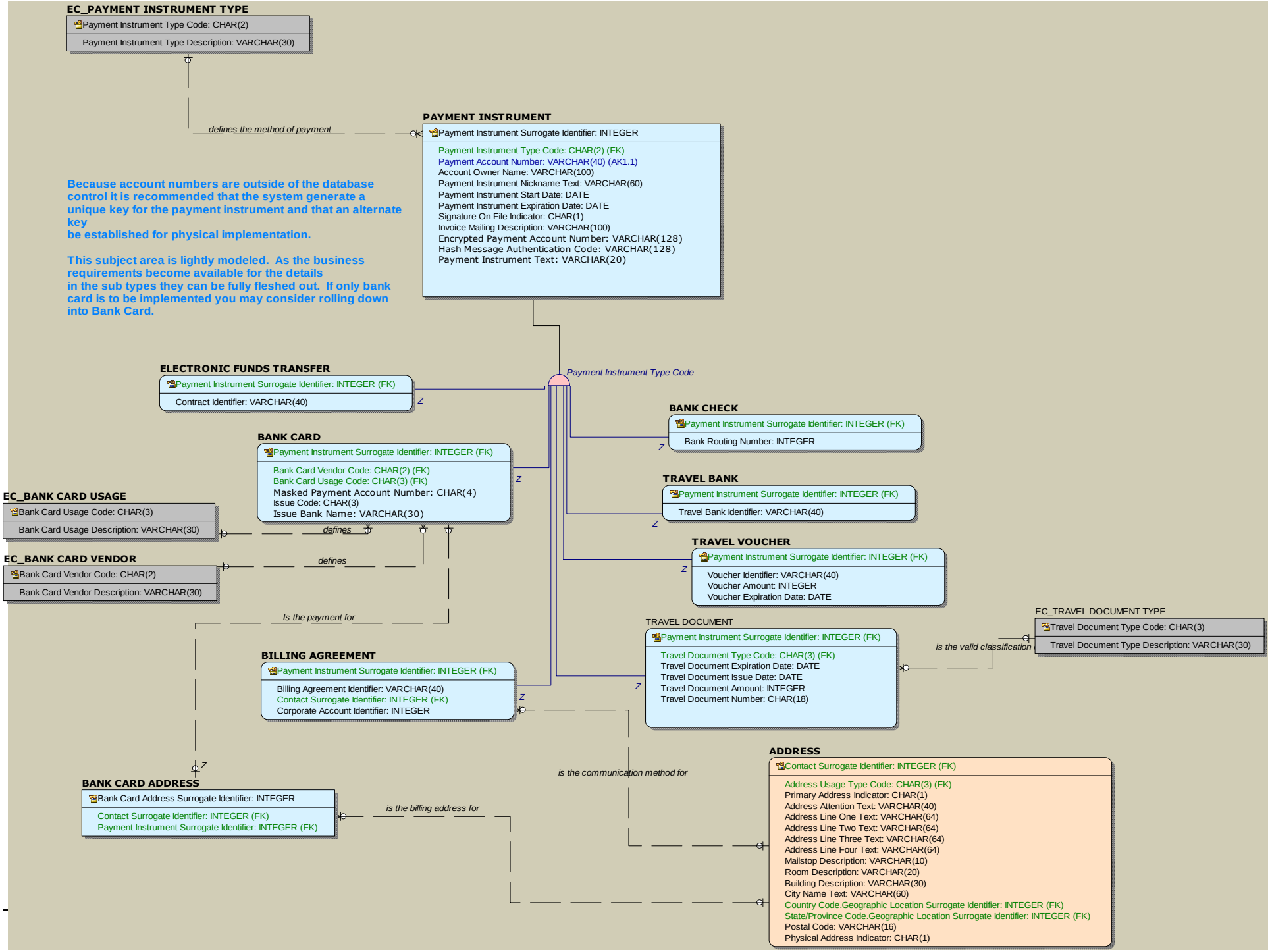
- **Refocused the Enterprise Data Architect to concentrate on providing corporate data standards, enterprise models, and other data management best practices.**
  - Created a matrixed team of Lead Data Architects from each of the business units to work with the CTO.
    - Executive Leadership team sponsored
    - As a cohesive team we have introduced:
      - Naming standards
      - Enterprise model components and patterns
      - Process guidelines
      - Best Practices
      - Modeling guidelines
      - Enterprise Model checklist
- **At one point in time we had ERwin, Embarcadero, Power Designer, Visio, MS Word, MS Access, Rational Rose and note paper used to model data.**
  - We evaluated the various case tools and standardized to ERwin as the tool we would employ for all our data models.

Data Management

- **Introduced Enterprise Model Components**
  - Normalized logical models that covered reusable areas of interest:
    - Payment Instrument
    - Contact
    - Language
    - Geography
    - Remarks
    - Loyalty Programs

- **Implemented standards that required all ERwin to also utilize Model Manager.**
  - Introduced Model Manager to everyone to give visibility to the data models and enable the synchronization and reuse of enterprise components to the teams


EC_PAYMENT INSTRUMENT TYPE
Tracks the valid payment forms a party may use .
CC Credit Card
CK Bank Check
CS Cash
TB Travel Bank
TV Travel Voucher
DB Debit Card
CA Corporate Agreement
PP Pay Pal
PAYMENT INSTRUMENT
The mechanism for transferring funds from one party to another.
This is considered highly sensitive data and should be secured in the same manner across the enterprise.
defines the method of payment
Payment Instrument Type Code
ELECTRONIC FUNDS TRANSFER
Is an agreement between two parties to automatically transfer funds from one account for an agreeded schedule timeframe to pay for services.
TRAVEL BANK
An account which allows the holder to purchase authorized travel services against deposited funds.
EC_BANK CARD USAGE
Represent the way that a person chooses to use the card for a specific payment purpose.
DB_Debit
CR_ Credit
defines
BANK CARD
A line of credit provided by a lending institution that may be used to pay for purchases.
BANK CHECK
An account which allows the holder to writes checks against deposited funds
TRAVEL VOUCHER
A reciept used as a form of payment for services rendered. Issued generally by government agencies and used as cash.
defines
EC_BANK CARD VENDOR
Is the industry identifier for a specified financial institution
AA American Airline Credit Card
AB Australian Bankcard
AQ Aloha Airlines Credit Card
AS Alaska Airlines Credit Card
AX American Express
BA Visa
Is the payment for
BILLING AGREEMENT
An agreement between two parties that allows purchase of travel related services that are billed at a predetermined interval.
is the communication method for
EC_TRAVEL DOCUMENT TYPE
Is the classification of document that a customer may present for use as payment
TRAVEL DOCUMENT
Tracks specific issued documents that may be used as form of payment.
is the valid classification of
BANK CARD ADDRESS
Tracks the billing or authorization address used for payment from the credit company
is the billing address for
ADDRESS
the media used to contact a person or organization for the purposes of communication.


EC_PAYMENT INSTRUMENT TYPE
Payment Instrument Type Code: CHAR(2)
Payment Instrument Type Description: VARCHAR(30)
defines the method of payment
PAYMENT INSTRUMENT
Payment Instrument Surrogate Identifier: INTEGER
Payment Instrument Type Code: CHAR(2) (FK)
Payment Account Number: VARCHAR(40) (AK1.1)
Account Owner Name: VARCHAR(100)
Payment Instrument Nickname Text: VARCHAR(60)
Payment Instrument Start Date: DATE
Payment Instrument Expiration Date: DATE
Signature On File Indicator: CHAR(1)
Invoice Mailing Description: VARCHAR(100)
Encrypted Payment Account Number: VARCHAR(128)
Hash Message Authentication Code: VARCHAR(128)
Payment Instrument Text: VARCHAR(20)
Because account numbers are outside of the database control it is recommended that the system generate a unique key for the payment instrument and that an alternate key
be established for physical implementation.
This subject area is lightly modeled. As the business requirements become available for the details
in the sub types they can be fully fleshed out. If only bank card is to be implemented you may consider rolling down into Bank Card.
Payment Instrument Type Code
ELECTRONIC FUNDS TRANSFER
Payment Instrument Surrogate Identifier: INTEGER (FK)
Contract Identifier: VARCHAR(40)
BANK CHECK
Payment Instrument Surrogate Identifier: INTEGER (FK)
Bank Routing Number: INTEGER
BANK CARD
Payment Instrument Surrogate Identifier: INTEGER (FK)
Bank Card Vendor Code: CHAR(2) (FK)
Bank Card Usage Code: CHAR(3) (FK)
Masked Payment Account Number: CHAR(4)
Issue Code: CHAR(3)
Issue Bank Name: VARCHAR(30)
TRAVEL BANK
Payment Instrument Surrogate Identifier: INTEGER (FK)
Travel Bank Identifier: VARCHAR(40)
EC_BANK CARD USAGE
Bank Card Usage Code: CHAR(3)
Bank Card Usage Description: VARCHAR(30)
defines
EC_BANK CARD VENDOR
Bank Card Vendor Code: CHAR(2)
Bank Card Vendor Description: VARCHAR(30)
defines
TRAVEL VOUCHER
Payment Instrument Surrogate Identifier: INTEGER (FK)
Voucher Identifier: VARCHAR(40)
Voucher Amount: INTEGER
Voucher Expiration Date: DATE
Is the payment for
TRAVEL DOCUMENT
Payment Instrument Surrogate Identifier: INTEGER (FK)
Travel Document Type Code: CHAR(3) (FK)
Travel Document Expiration Date: DATE
Travel Document Issue Date: DATE
Travel Document Amount: INTEGER
Travel Document Number: CHAR(18)
is the valid classification
EC_TRAVEL DOCUMENT TYPE
Travel Document Type Code: CHAR(3)
Travel Document Type Description: VARCHAR(30)
BILLING AGREEMENT
Payment Instrument Surrogate Identifier: INTEGER (FK)
Billing Agreement Identifier: VARCHAR(40)
Contact Surrogate Identifier: INTEGER (FK)
Corporate Account Identifier: INTEGER
is the communication method for
BANK CARD ADDRESS
Bank Card Address Surrogate Identifier: INTEGER
Contact Surrogate Identifier: INTEGER (FK)
Payment Instrument Surrogate Identifier: INTEGER (FK)
is the billing address for
ADDRESS
Contact Surrogate Identifier: INTEGER (FK)
Address Usage Type Code: CHAR(3) (FK)
Primary Address Indicator: CHAR(1)
Address Attention Text: VARCHAR(40)
Address Line One Text: VARCHAR(64)
Address Line Two Text: VARCHAR(64)
Address Line Three Text: VARCHAR(64)
Address Line Four Text: VARCHAR(64)
Mailstop Description: VARCHAR(10)
Room Description: VARCHAR(20)
Building Description: VARCHAR(30)
City Name Text: VARCHAR(60)
Country Code.Geographic Location Surrogate Identifier: INTEGER (FK)
State/Province Code.Geographic Location Surrogate Identifier: INTEGER (FK)
Postal Code: VARCHAR(16)
Physical Address Indicator: CHAR(1)

Data Security

$$PCI = \sum_{\infty}^{now} Y2K \times SarbOx$$

$$\text{Perfect Security} = \sqrt{-1}$$

Data Identification – you can't secure what you can't find.

- **PCI standards were mandated and we looked at the scope of what needed to be delivered and realized we had a wonderful tangible opportunity to illustrate the importance of data models.**
  - Provide value to the organization as the knowledge workers around data
- **Adding governance (with a little g) for PII**
  - As a company we choose to comply with local and national and international laws to ensure our customers data is secure
- **To enable better visibility we introduced a standard that requires all databases that store, move or process credit card data to have:**
  - A managed data model using the standard tool
  - A managed data model stored in a central repository.

Data Security – you can't secure what you can't find.

- **The core team introduced a process to tag the data elements within a model as PCI, or PII data using the UDP (User Defined Property) feature of ERwin.**

- **Data identification is part of the ongoing data modeling lifecycle.**

- **Global Reporting**
  - Hundreds of databases were reversed engineered
  - Thousands of servers that may contain PCI and PII were reviewed
  - Standard reports templates developed
  - Monthly compliance reports are generated and published to the Executive Leadership Team.

Data Identification – you can't secure what you can't find.

- Data identification introduced and standardized
  - The security classification will be established for certain attributes at the logical level. An attribute that is PCI or PII must be at least one of the four items listed below:
  -

UDP Name	UDP Type	UDP Description
Classword	TEXT	Is the classification of attributes into categories with common characteristics
Source System Identifier	TEXT	Is the technical identifier used for a system of record for this piece of data
Business Term	TEXT	Is the term that a business unit may use when talking about this data attribute
Source System Name	TEXT	Is the name of the system of record for this data attribute
PCI Security Classification	LIST	Is the Payment Card Industry security classification for a data attribute
PII Security Classification	LIST	Is the Personally Identifiable Information security classification for a data attribute
SPI Security Classification	LIST	Is the Sensitive Personal Information security classification for a data attribute
EU Security classification	LIST	Is the Sensitive Information security classification for a data attribute

- Tip! Tag the data using Data Browser for quick search and update

Data Identification – you can't secure what you can't find.

- **When the work is completed one or more of the following is added as appropriate to the findings as a Model Level UDP.**

Logical Model UDP Name	Type	Description
Secured PCI	Date	Reviewed by (insert name) and PCI Compliant
Secured PII	Date	Reviewed by (insert name) and PII Compliant
No PCI	Date	Reviewed by (insert name) no PCI data stored
No PII	Date	Reviewed by (insert name) no PII data stored
Not Production	Date	Not a production model no PCI/PII identification made. (insert name)

Physical Model UDP Name	Type	Description
PCI Encrypted	Date	All PCI data encrypted work approved by (insert name)
PCI Obfuscated	Date	All PCI data obscured work approved by (insert name)

Data Identification – you can't secure what you can't find.

- **Global reports are run against the model repository to communicate with the security team identification of PCI exposure or resolution.**

Library		Model	Class	UDP
AS Air Operation		**Aerodynamic Traveler**	Attribute	Attribute.Logical.Business
				Attribute.Logical.EU Security Class
				Attribute.Logical.PCI Security Class
				Attribute.Logical.PII Security Class
				Attribute.Logical.SPI Security Class
			Model	Model.Logical.Secured PCI
		SS Check-In	Attribute	Attribute.Logical.Business
				Attribute.Logical.EU Security Class
				Attribute.Logical.PCI Security Class
				Attribute.Logical.PII Security Class
				Attribute.Logical.SPI Security Class
			Model	Model.Logical.Secured PCI

- **Reports are run by the data modelers for each model to indicate which data is tagged so that the project teams are aware of where the data is so they can take the necessary steps to protect it by encryption, or removal.**

Sabre Holdings

Next Steps

- **PII is next and its impact**
  - PII is obviously more pervasive than Payment Data and therefore poses a more significant impact to the teams to handle the data.
  - As the keepers of the data we must keep and identify the data at the most atomic level. The combination of that data and how it is handled is done by the applications and systems that use the data.
- **What's Next**
  - Mapping databases to messages need to know the data flows between tiers and systems
  - Keeping the identification of sensitive data across model types and into the database.
  - Better ways to search database and data models with the magnitude of databases
  - Search across models in Model Manager ad-hoc queries
  - Better reporting

Challenges

Challenge

- Data modeling in an Agile development environment
  - Rapidly changing requirements
  - Refactoring the database structure after the code is done

Solution

- Use more surrogate keys to lessen the impact
- Use model templates to promote reuse and agility
- Know the data and share information
- Become a valuable resource

Challenges

Challenge

- How to achieve governance in a matrix organ

Solution

- Creation of Core Teams, representatives from every business unit
- Obtain Executive buy-in
- Make sure that individuals on the Core Team have the authority to make decisions for their business unit as it pertains to data management.

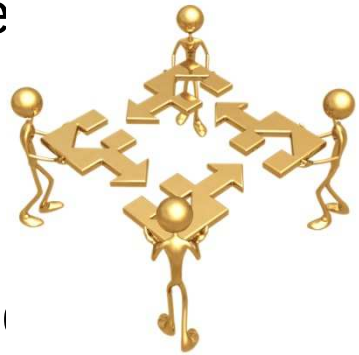
Challenges

Challenge

- Staying effectively connected with remote/global te

Solution

- Use Remote desktop capabilities for model assistance/desktop takeover
- Utilize VM for specific performance issues with Mo Manager
- Train using Webex or similar tool
- Ensure you have good Conference phones/headsets
- English is not necessarily someone's first language ..SLOW DOWN
- Use a graphic to illustrate or demonstrate the point.



Challenges

Challenge

- Model Driven Development
  - Conceptual models
  - Business Process mapping
  - Logical driving Physical models
- Will save time - Prove it!

Solution?

Use the case tools to push this idea?

Challenges

Challenge

- Finding Models and knowing which ones are go consistent and follow standards
- 50 Libraries in Model Manager
- 182 data models in Model Manager
- Model size is an average of:
  - 267 Entities
  - 2797 Attributes

Solutions ??

Challenges

Challenge

- How to apply Governance in a Matrixed organization
  - Enforce good modeling techniques

Solution ?

- Allow new government and regulatory mandates to enforce governance.

Summary

- Matrixed teams work as long as senior management in each business unit recognize the value
- Using regulatory mandates only helps facilitate the data management effort
- Standards are important
  - Small steps to achieve big results
- Communication is critical for success
- Leverage the knowledge of the data practitioners by supplying information to the teams.

THAT'S ALL FOLKS!