2017 Data Investment Symposium

Driving Development and Building a Network of Leaders with a Data Focus

February 22, 2017

CU*ANSWERS
A Credit Union Service Organization

Data Investment Symposium

THE FINAL YEAR TO WRAP UP OUR DIRECTION

- Year 1: **Testing the waters**
  - Introduced the concept of specialized investments for selling, creating, and supporting data warehouse environments for our network
  - Discussed possibilities for a third-party solution
  - Small, vested group of people who would drive the investment
- Year 2: **Serving CUs that stuck with the ideas**
  - Sharing new warehouse designs and building a business for the 10% of CUs who will one day have an aggressive investment
- Year 3: **An "in-the-business" approach**
  - Use what we've learned to go into the data warehouse business for the long term

Data Investment Symposium

WHERE MY HEAD IS TODAY

- After 2+ years of thinking about how the data craze fits into our future, we are ready to launch extended data services and project management for our network
  - We'll start by offering new ideas about where to store data, and what data warehouses to offer
  - We'll consider how CUs who like to share things will use our solutions, and how CUs who like to do it by themselves will need support
- Data warehouses are the foundation of data analytics, and we'll align them with our goals to enhance the data analysis capabilities in our network

FOCUS GROUP INPUT

Round 1: Introduce Yourself

- Tell us your name, credit union name, and your position
- Which of these best describes your role:
  - Writing the plan to maximize data?
  - Encouraging others to maximize data?
  - On the front lines, responsible for finding, presenting, and creatively leveraging the data towards action for success?
- Is this your first Data Investment Symposium?

FOCUS GROUP INPUT

Round 2: Which one are you?

- Where does your credit union fall among our emerging audiences?
  - "My CU has studied the data warehouse provider market and we are getting close to investing or reinvesting in new data strategies and tactics"
  - "My CU has casually been watching the projects at CU*Answers, at the Boot Camps, etc., and we are ready to propose some of these ideas to our business plan authors and leaders"
  - "I'm more interested than my CU is, and I need some help to engage my CU around these ideas"

Today's agenda: Steps for moving forward with our commitment to the data warehouse business

Options for how we change our current data management, understanding that new warehouses give us new options

Considering what it would mean for CU*Answers to buy an Intel data warehouse product and share it with CUs via an online business model

Attacking the low-hanging fruit when it comes to building a data warehouse business: FILExxDW

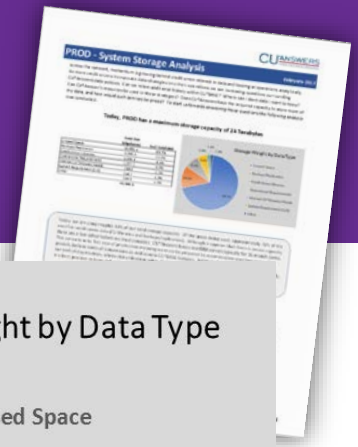
Considering our options related to CU*BASE data

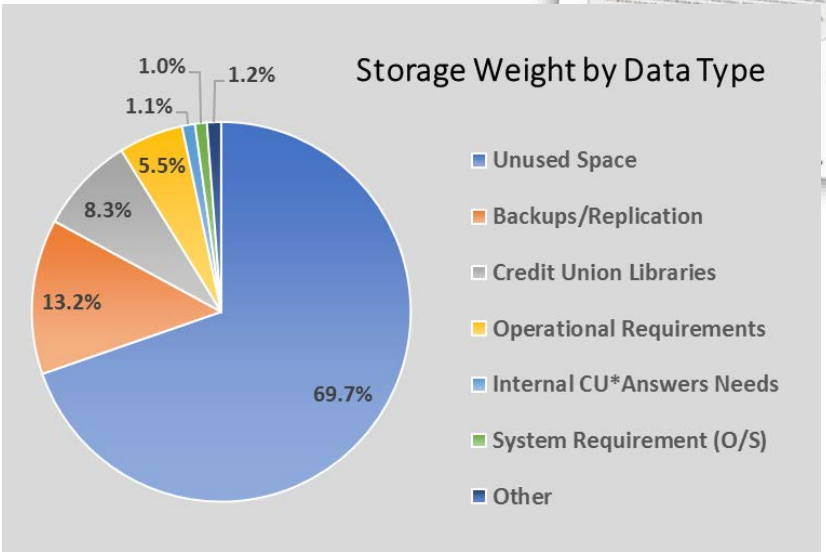
When we think about our current data warehouse (FILExx), what changes could we make so that every CU is better off in 2018?

CU*ANSWERS
A Credit Union Service Organization

2017 Data Investment Symposium

Some facts about our existing data warehouse





	Total Size (Gigabytes)	% of Total Disk
Unused Space	16,931.1	69.7%
Backups/Replication	3,208.4	13.2%
Credit Union Libraries	2,006.6	8.3%
Operational Requirements	1,337.2	5.5%
Internal CU*Answers Needs	269.0	1.1%
System Requirement (O/S)	246.7	1.0%
Other	293.5	1.2%
	24,292.4	

All stats gathered from online CU*Answers clients as of February 7, 2017

Average size of CU file library

7.2 Gb

Largest CU file library

57.5 Gb

Average size of CU EOM library

2.2 Gb

Largest CU EOM library

28.9 Gb

Average size of CU Query library

0.2 Gb

Largest CU Query library

8.4 Gb

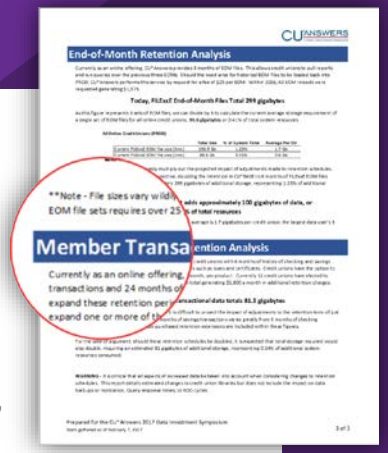
Thinking about the most basic CU*BASE data retention policies

- Before we add a data warehouse where it costs extra for retaining more of your existing CU*BASE data, should we consider just retaining more data in FILExx right now?
- We don't want the new data warehouse to be simply a place to pay for the ability to be a packrat for data you'll never use

Type of Data	Standard Online Retention
Savings/checking history	6 months
Certificate history	24 months
IRA history	24 months
HSA history	24 months
Club account history	24 months
Tax Escrow history	24 months
Loan history	24 months
EOM libraries	3 months
GL trans history	24 months
Check register	24 months
Credit report detail	60 days or life of loan
Credit report summary	6 months
Statements	18 months online
Reports	90 days online

Our new data warehouses need to lift up CU capabilities, not just extend old ideas

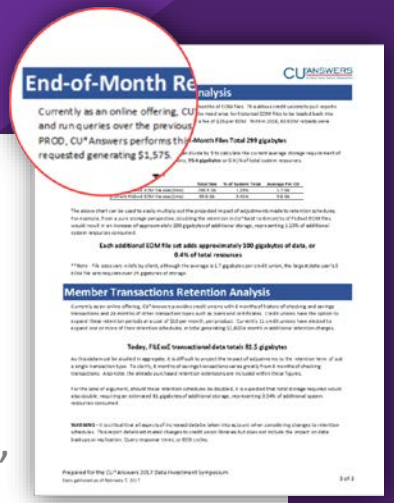
What if we added more member transaction history to our retention schedule?



- We gathered some stats about transaction data stored at our GR data center
  - How did we set the retention policies way back when?
  - Given the size of boxes and the way people think about data, the actual storage used seems rather insignificant – is it?
- If we keep this, will we be extending CU*BASE Inquiry/Phone Op simply for the sake of seeing more? Or is there analysis?
  - Remember, this will lead to other projects, like what you can view in online and mobile banking
- What is a reasonable extension? Double? Quadruple? Shares vs. checking vs. loans, etc.

We're going to do something before June – what ideas do you have?

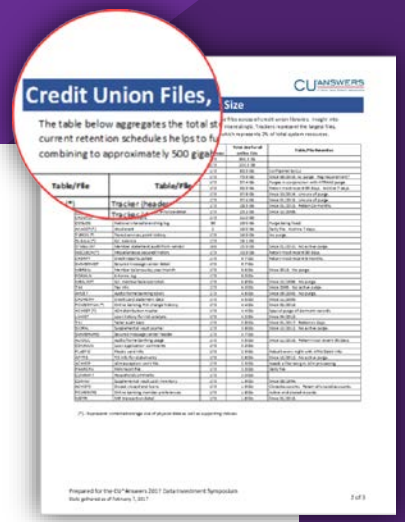
What would it mean to keep more month-end periods available online?



- We gathered some stats about EOM snapshots stored at our GR data center
  - How did we set the retention policies way back when?
  - Given the size of boxes and the way people think about data, the actual storage used seems rather insignificant – is it?
- EOM files are more about researching, but generally related to some kind of member service request – if we doubled to six months, do you think it would change the way you use these files?
- What would be a reasonable change to our policy?

Should we leave it at 3 months and spend the money on trans history?

What about data outside of transaction history and EOM snapshots?



- We gathered some stats about other data stored at our GR data center data such as Trackers, GL history, OFAC scan results, credit bureau reporting history, etc.
  - How did we set the purge procedures way back when?
- What is different about these files compared to transaction history?
- Should we be aggressive and adjust the purge routines to *retain* more data – or *purge* more of it?
  - Duh, who would vote for less? But we should really think about why we *shouldn't* keep this data forever

We're going to do something before June – what ideas do you have?

Christmas in June!

TO BE ANNOUNCED AT THE JUNE LEADERSHIP CONFERENCE?

- We're not restoring old data, but we will be changing the retention schedule to start building new trendlines
- We want to give this a try and see if we get what we all think we're buying: better member service and a more active set of credit union analysts
- Time will tell…retention schedules are meant to be adjusted

Type of Data	Standard Online Re…
Savings/checking history	~~6 months~~ **12 months**
Certificate history	~~24 months~~ **48 months**
IRA history	~~24 months~~ **48 months**
HSA history	~~24 months~~ **48 months**
Club account history	~~24 months~~ **48 months**
Tax Escrow history	~~24 months~~ **48 months**
Loan history	~~24 months~~ **48 months**
EOM libraries	~~3 months~~ **6 months**
GL trans history	24 months
Check register	24 months
Credit report detail	60 days or life of loan
Credit report summary	6 months
Statements	18 months online
Reports	90 days online

Our new data warehouses need to lift up CU capabilities, not just extend old ideas

Should we make room for a new kind of client-architected data in CU*BASE?

- Currently in development, targeted for the 17.05 release:
  - #43454: Custom data fields at the account (MEMBER1-6) level
  - #43455: Custom data fields at the membership (MASTER) level
- Modeled after the 5300 account code database to allow CUs to create their own field names and data types, attached either to a specific sub-account or to the membership itself
  - Dollars $xxx,xxx,xxx.xx
  - Quantities xxx,xxx,xxx
  - Percentages xxx.xx%
  - Date May 07, 2013
  - Yes/no checkbox
  - Free-form text
  - CU-defined list (with lookup)

Let's take a look…

Session 5 CU*BASE GOLD Edition - TEST CREDIT UNION

File Edit Tools Help

Mockup only; subject to change

Custom Account Field Configuration

Field name CRDTGRADE

Display sequence # 00015

Description Credit grade at time of loan approval

Data type

- ○ Dollars
- ○ Date
- ◉ Choose from a List
- ○ Quantity
- ○ Text
- ○ Percentage
- ○ Checkbox

Application type ☐ SH ☐ SD ☐ CD ☐ IR ☐ TX ☑ LN ☑ OC

Pre-defined values section at the bottom appears for the "List" data type

Pre-defined Values

Code	Value
1	Whatever description you want
B	Credit grade 2
EX1	Credit grade 3
TST	Credit grade 4
5	Credit grade 5

Code = 3 characters
Value = 30 characters
(both alphanumeric)

■ Edit ■ Delete ■ View ■ Suspend/Reactivate

Individual values can be suspended without affecting data already entered on a member account

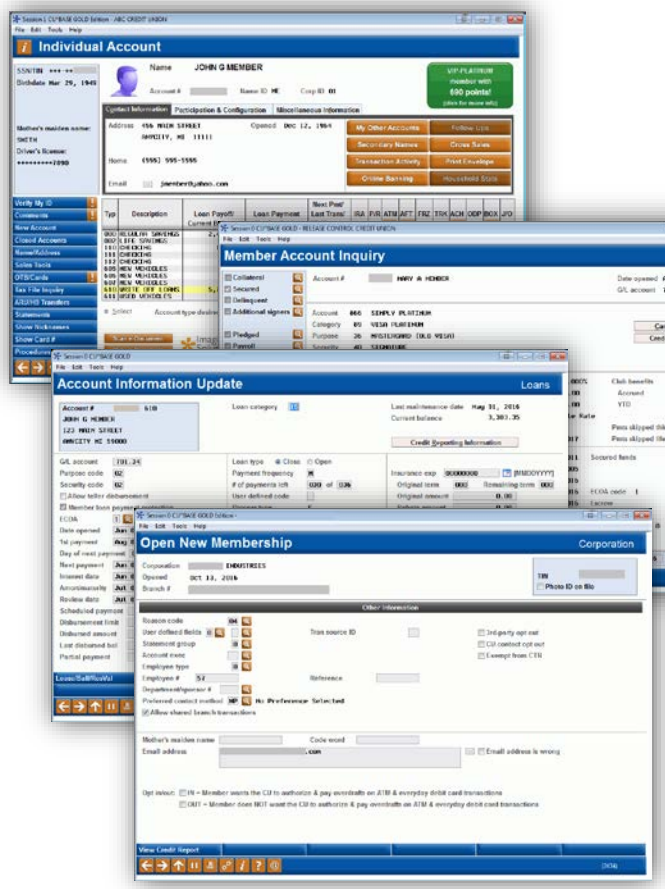
Add Value

WU (3248) 5/08/13

Session 5 CU*BASE GOLD Edition - TEST CREDIT UNION
File Edit Tools Help
This example is for account-level fields for AplTyp LN and OC
Mockup only; subject to change
Custom Account Information
Member 123456789 JOE Q MEMBER
Account 789 SIGNATURE LOAN
Position to custom field
Search for description containing
Custom Field
Description
Content
ORIGLTV
Original LTV (before backend)
$14,234.93
CRDTGRADE
Credit grade at time of loan approval
TST – Credit grade 4
COAPSCORE
Co-applicant score
COAPPTYPE
Co-applicant type
ORIGLNCATG
Original loan category (for write-offs)
LASTREPRICE
Date of last rate reprice
May 07, 2013
LOCOPTOUT
CC automatic LOC increase opt out
Lookups will display your list of configured values
Just like on the 5300 Call Report screen, GOLD will display an appropriate fill-in field based on the data type
Show Filled In Only
WU (3248) 5/08/13

Custom Account and Membership Fields

- Access points for entering or viewing custom data:
  - Custom fields can be maintained via **membership creation/update** and via **account information update**
  - Custom data can be viewed via **Inquiry/Phone Op** main screens and **account inquiry** screens
  - Custom data will also be available for use with Query



A product looking for a name: Randy offered a cash prize to Boot Campers for a workable brand name … ideas, anyone?

FILExxDW

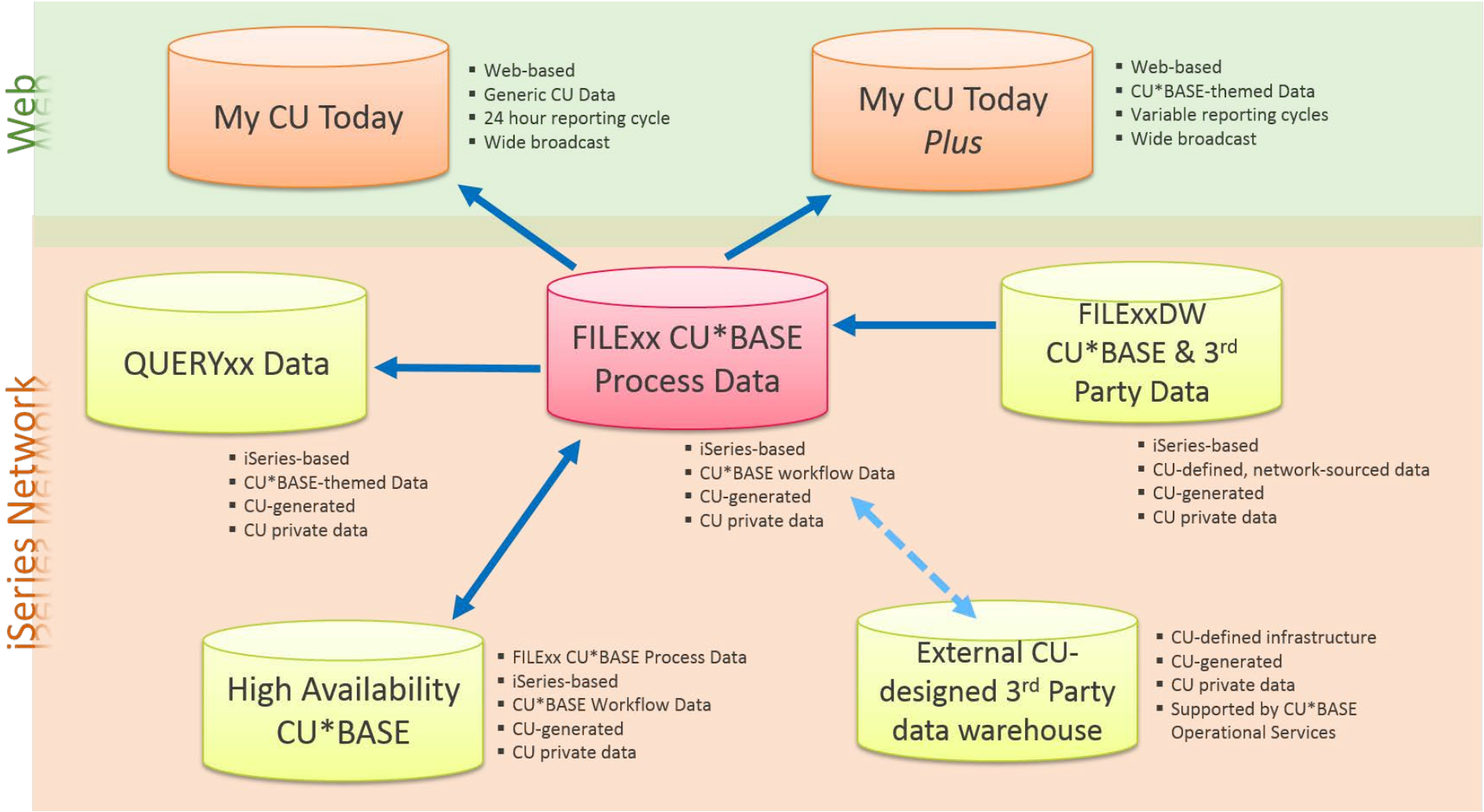
The low-hanging fruit when it comes to building a data warehouse business

By adding a second cylinder to our online environment, we can leverage our data center investments and our primary computing environment

CU*ANSWERS
A Credit Union Service Organization

2017 Data Investment Symposium

From last year's Leadership Conference…



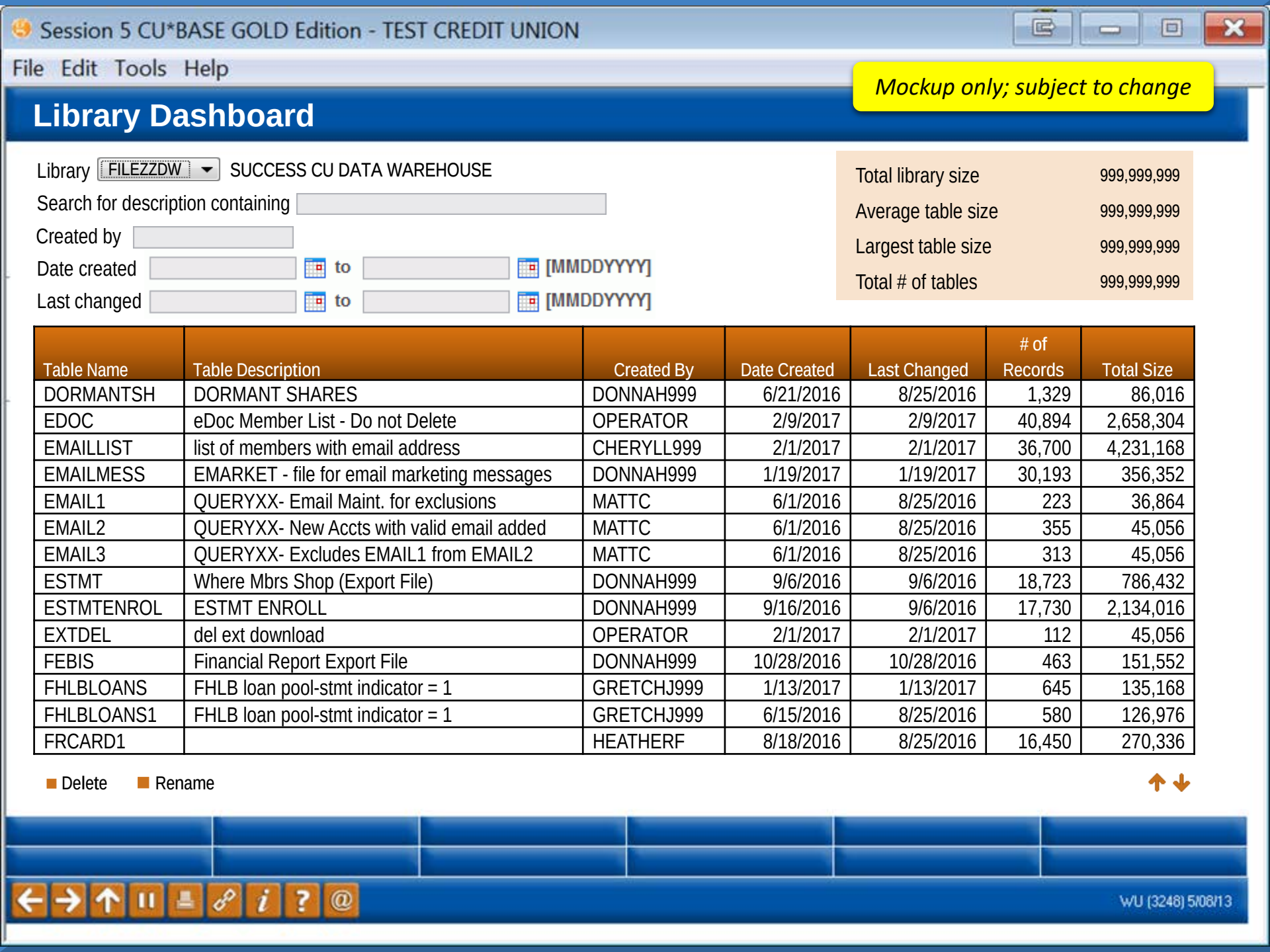
How do you support the investments and work necessary to offer a data warehouse? **Fees!**

- Everyone understands we have to charge for the additional space we'll need to build and maintain
- But it's much more than that…
  - What/how much data is allowed, how to control?
  - How to manage?
  - What about redundancy, high availability?
  - Backups – Frequency, impact on nightly operations, ability to restore?
  - Access – Embedded in CU*BASE vs. kept separated?
  - Encryption – How do we stay safe and compliant? (What if your tables contain PANs?)

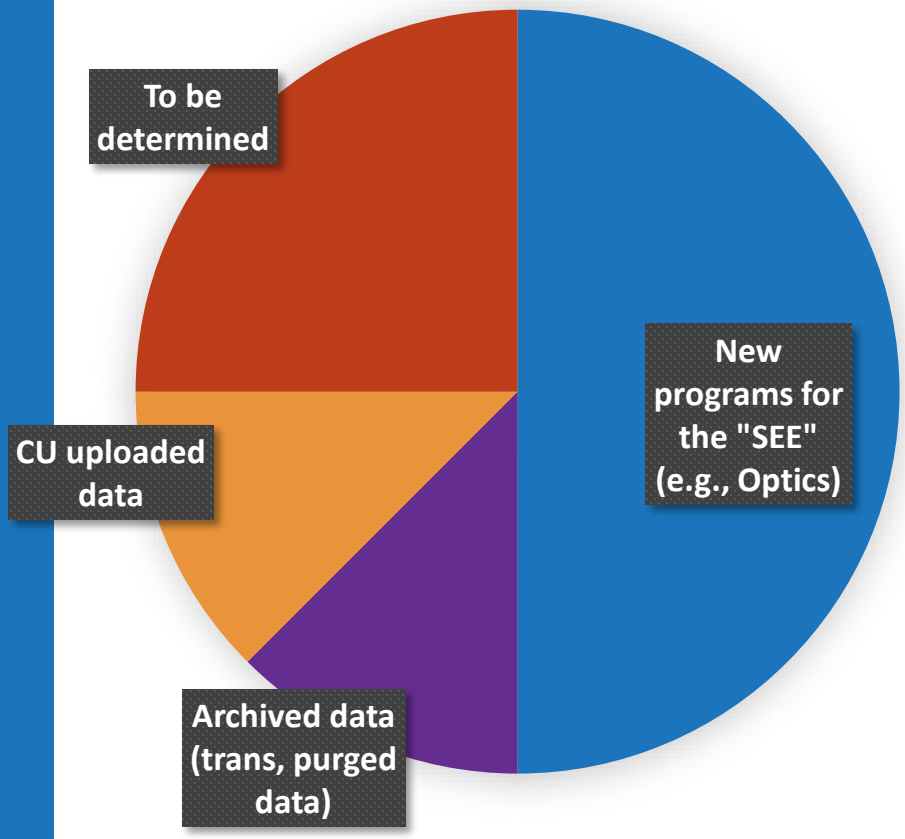
Over the next few months we need to publish procedures and controls and design a revenue model that makes this a long-term offering

For example, if we charge by storage space used…

- …how would you inventory what is in your warehouse?
- …how would you know how much space you're using?
- …how would you delete unused data?
- …how would you keep your files organized?
- We need to build something for the data warehouse user group and the CU's database administrator
- How about this?


Session 5 CU*BASE GOLD Edition - TEST CREDIT UNION
File Edit Tools Help
Mockup only; subject to change
Library Dashboard
Library FILEZZDW SUCCESS CU DATA WAREHOUSE
Search for description containing
Created by
Date created to [MMDDYYYY]
Last changed to [MMDDYYYY]
Total library size 999,999,999
Average table size 999,999,999
Largest table size 999,999,999
Total # of tables 999,999,999
Table Name | Table Description | Created By | Date Created | Last Changed | # of Records | Total Size
DORMANTSH | DORMANT SHARES | DONNAH999 | 6/21/2016 | 8/25/2016 | 1,329 | 86,016
EDOC | eDoc Member List - Do not Delete | OPERATOR | 2/9/2017 | 2/9/2017 | 40,894 | 2,658,304
EMAILLIST | list of members with email address | CHERYLL999 | 2/1/2017 | 2/1/2017 | 36,700 | 4,231,168
EMAILMESS | EMARKET - file for email marketing messages | DONNAH999 | 1/19/2017 | 1/19/2017 | 30,193 | 356,352
EMAIL1 | QUERYXX- Email Maint. for exclusions | MATTC | 6/1/2016 | 8/25/2016 | 223 | 36,864
EMAIL2 | QUERYXX- New Accts with valid email added | MATTC | 6/1/2016 | 8/25/2016 | 355 | 45,056
EMAIL3 | QUERYXX- Excludes EMAIL1 from EMAIL2 | MATTC | 6/1/2016 | 8/25/2016 | 313 | 45,056
ESTMT | Where Mbrs Shop (Export File) | DONNAH999 | 9/6/2016 | 9/6/2016 | 18,723 | 786,432
ESTMTENROL | ESTMT ENROLL | DONNAH999 | 9/16/2016 | 9/6/2016 | 17,730 | 2,134,016
EXTDEL | del ext download | OPERATOR | 2/1/2017 | 2/1/2017 | 112 | 45,056
FEBIS | Financial Report Export File | DONNAH999 | 10/28/2016 | 10/28/2016 | 463 | 151,552
FHLBLOANS | FHLB loan pool-stmt indicator = 1 | GRETCHJ999 | 1/13/2017 | 1/13/2017 | 645 | 135,168
FHLBLOANS1 | FHLB loan pool-stmt indicator = 1 | GRETCHJ999 | 6/15/2016 | 8/25/2016 | 580 | 126,976
FRCARD1 | | HEATHERF | 8/18/2016 | 8/25/2016 | 16,450 | 270,336
Delete Rename
WU (3248) 5/08/13

The easy part is thinking about what CU*Answers will put into FILExxDW



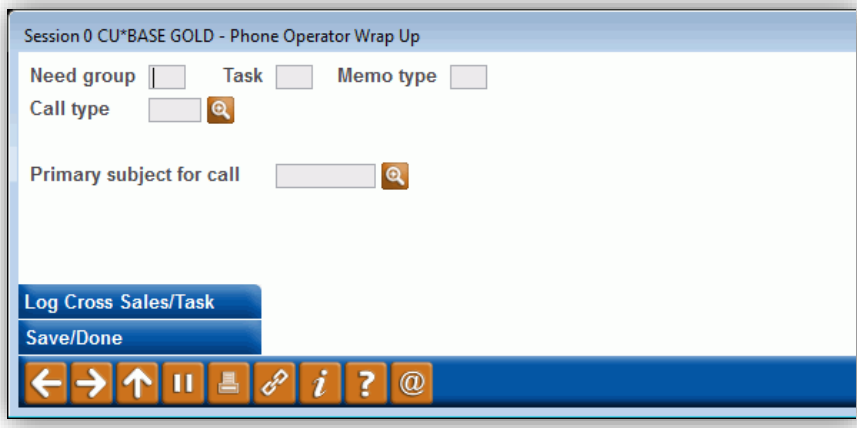
- CU*BASE
  - Phone optics
  - **It's Me 247** optics
  - OTB raw data
  - Bill pay subscribers, payees, activity
  - Purged data/extended retention (trans, Trackers, GL, etc.)
- Snapshots
- Other, TBD
  - Conversion data?
  - ??

It's easy because we are the architects of the data, and we'll set up the rules

SPEAKING OF GETTING DATA INTO CU*BASE…

Let's take a look at a couple of projects we talked about last year

- Phone Op Wrap-up Codes were released in 16.05
  - One-click Trackers were added to the wrap-up window in 16.10



- Coming in 2017-2018:
  - Phone Center Optics (repositioning this project for FILExxDW)
  - **It's Me 247** Optics

The hard part is thinking about how to control the things you'll put into the warehouse

- What kind of controls can we even put into place?
  - We can't really know content or amount of data being uploaded
- How do we prevent someone from attempting to upload so much data it crashed the system (exceeds available DASD space)?
- Research to be done on if and how we can control the size of libraries
- We can create FILExxDW and be ready for every CU, but still need to decide which credit union employees have access

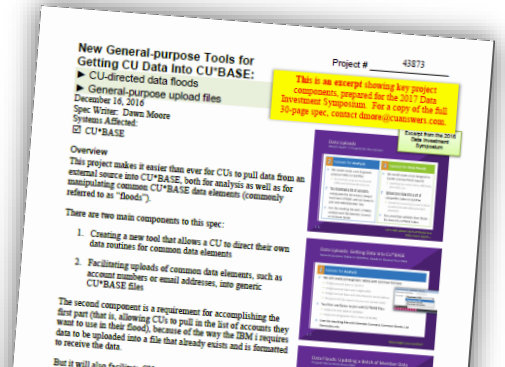
As a customer of the data warehouse, what rules do you think would be reasonable?

SPEAKING OF UPLOADS…

Coming in 2017: CU-directed data floods of common CU*BASE fields

DIY/DIT
Do it yourself
Do it together

- **Run your own data floods for FREE**
  - Built-in audit and exception reports; validation to ensure only clean data is written to member files
  - Values can be selected from a config or even pulled out of your uploaded file
- **Upload any account list for data analysis or floods**
  - New generic tables (complete with *.FDF files needed for the upload process) into which you can pull data
  - Phase 1 includes **26 different flood types**, such as:
    - ANR limits, due diligence flags, ECOA codes, electronic hold groups, freeze codes, delinquency fine codes, loan purpose and security codes, reason codes, opt in/outs, statement style codes...and many more!


New General-purpose Tools for Getting CU Data Into CU*BASE:

Project # 43873

▶ CU-directed data floods

▶ General-purpose upload files

This is an excerpt showing key project components, prepared for the 2017 Data Investment Symposium. For a copy of the full 30-page spec, contact dmoore@cuanswers.com.

Excerpt from the 2016 Data Investment Symposium

December 14, 2016

Spec Writer: Dawn Moore

Systems Affected:

☑ CU*BASE

Overview

This project makes it easier than ever for CUs to pull data from an external source into CU*BASE, both for analysis as well as for manipulating common CU*BASE data elements (commonly referred to as "floods").

There are two main components to this spec:

1. Creating a new tool that allows a CU to direct their own data routines for common data elements
2. Facilitating uploads of common data elements, such as account numbers or email addresses, into generic CU*BASE files

The second component is a requirement for accomplishing the first part (that is, allowing CUs to pull in the list of accounts they want to use in their flood), because of the way the IBM i requires data to be uploaded into a file that already exists and is formatted to receive the data.

Pricing Strategies
BUILDING A REVENUE MODEL

- As customer-owners, you need to wear both hats and help us build something we can all afford and that your CU can earn a return on
- Pricing strategies to consider:
  - **Storage**
    - Amount of data in the warehouse on the last day of the month?
    - Peak usage? (Track usage each day of the billing cycle and bill the highest day)
    - Peak average? (Average out the 3 highest days)
    - Monthly average?
  - **Access** (number of people using the warehouse)
  - **Bandwidth** (amount of data being uploaded per period)
  - **Processing** (for now, just the work to extract purged data over to the warehouse)

FILExxDW and our 2017 Calendar

- **Phase 1 (July 2017):** Come to market with FILExxDW for CU*Answers-generated data only
  - Complete a review of FILExx files
  - Why and how these files might move to FILExxDW
- **Phase 2 (Oct 2017):** Introduce data warehouse uploads to complete the FILExxDW business approach
  - Complete a beta with a CU that has an aggressive plan to use third-party data with CU*BASE data on the same system
- **Phase 3 (Nov 2017):** Introduce our data warehouse business at the CEO Strategies event and launch a long-term effort

Sharing an Intel-based server data warehouse

What would it mean for CU*Answers to buy an Intel data warehouse product and share it with CUs via an online business model?

CU*ANSWERS
A Credit Union Service Organization

2017 Data Investment Symposium

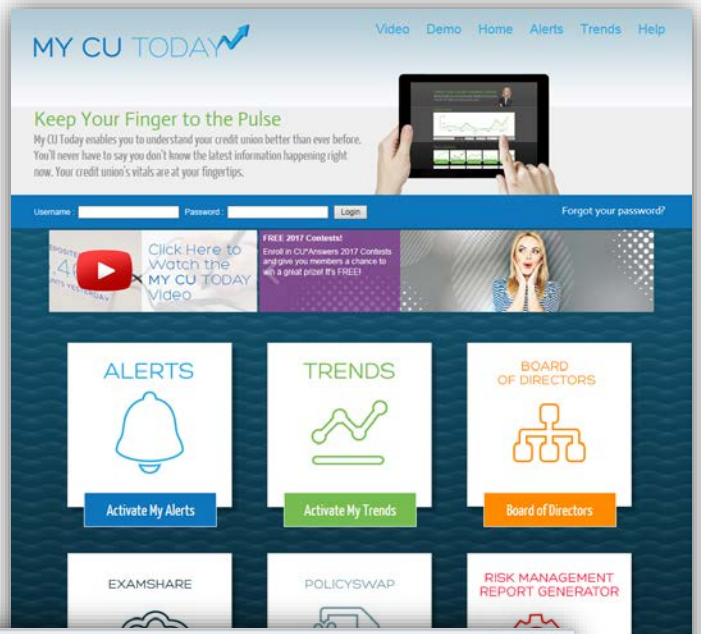
Thinking about Intel-server-based data warehouses

- Once we have an audience that's aggressive about data warehouses and analyzing third-party data, we believe there will be a market for CUs who need CU*Answers to:
  - Provide an Intel server-based warehouse as a shared property, or
  - Be the facilities manager for a credit union's own Intel server-based data warehouse
- At the first Data Investment Symposium we introduced OnApproach… but it was too early
  - During 2018 we'll work with interested CUs to search for a third party data warehouse platform
  - We have a lot of work to do in 2018

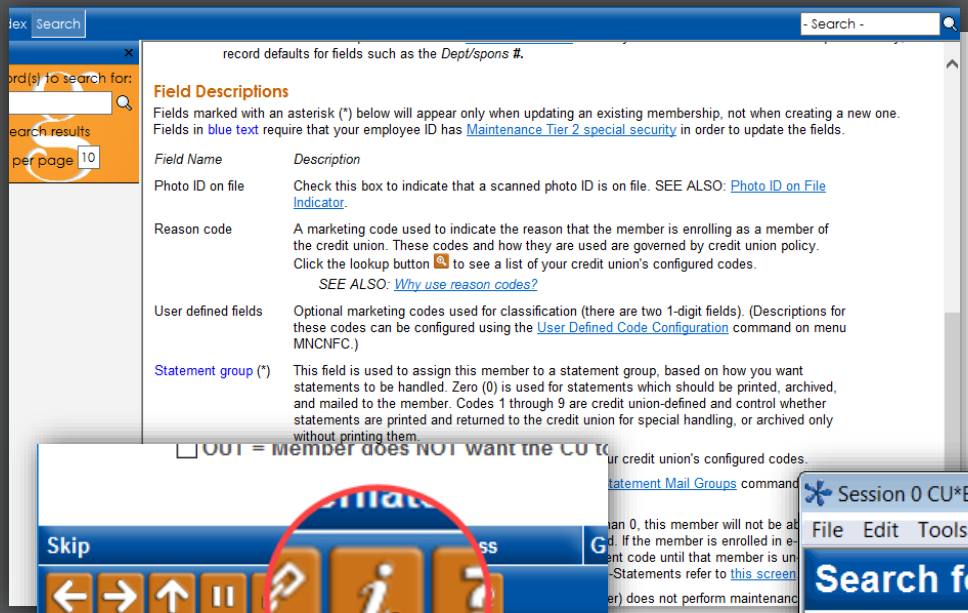
Many individual CUs are doing this work today; we may have a ready-made template by 2018

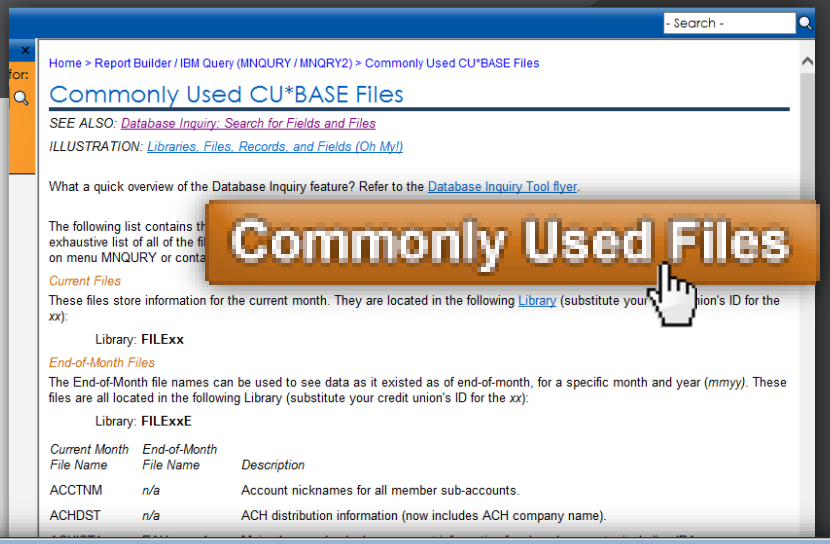
We are further along than you might think

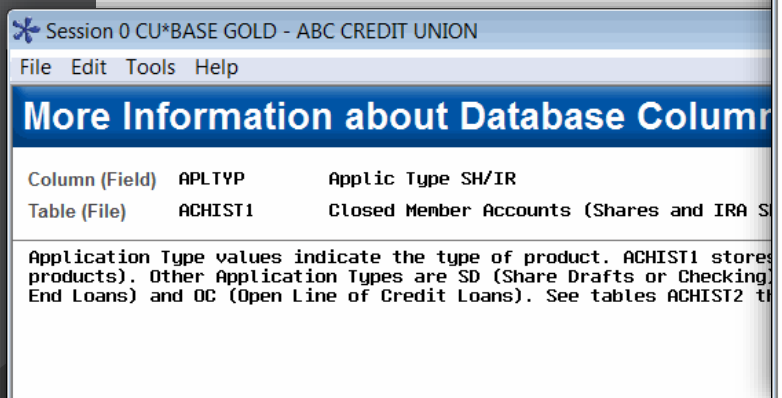
MY CU TODAY



Continuing to teach CUs our tables and how they can work for them







And coming in 17.05:
Enter your own comments into the Database Search Assistant!

Thinking about external data warehouses and the future

- Some stats from our most aggressive data warehouse user to date, Notre Dame CU:
  - Current data warehouse stores 1.25 terabytes
  - CU receives daily uploads from CU*BASE (some add records, some simply replace the prior day's data)
  - Also take some snapshots of daily balances or perform other calculations to record specific data points
  - Backup storage at two sites
  - Only 4 users have authority to log in to the data warehouse
  - Full-time staff dedicated to data warehouse activities

Wrap-Up

Remember, data warehouses are not the point…earning from data analytics and what you see in the data is what will save your future

Thanks for the day!