Why Study Large Scale Computing (aka Mainframes)
What is “Large Scale Computing”?

• Developing “really big” applications

• Focus on highly available, transaction processing systems

• Using the “mainframe” as a server
Capabilities of a Mainframe

• Large-scale transaction processing (thousands of transactions per second)

• Support thousands of users and application programs

• Simultaneously accessing resources

• Terabytes of information in databases

• Large-bandwidth communications
An Example

A Large System in a large Company
JP Morgan Chase Overview

By The Numbers…

• Net Revenue: $99B
• Net Income: $14.4B
• Assets: $1.2T
• Headcount: 174,360
• Technology & Communication Expense: $3.5+B
JP Morgan Chase Overview

**JP Morgan:** Wholesale banking

- Relationships with over 99% of the Fortune 1000 companies
- $3.2 trillion in wire transfers daily

**Chase:** Consumer & commercial banking

- 2,600+ branches (& growing via acquisition)
- 110+ million credit cards in circulation
The JP Morgan Chase “Merger” Tree

Major Mergers and Acquisitions of JPMorgan Chase and Bank One
Integration Challenges in Financial Services

Key Challenges

• Application integration challenges inherent from M&A efforts
• Variety of applications and platforms across business units
• Business unit silos not conducive to customer integration

Implications on Business & Technology

• Need to maintain a holistic view of the customer relationship
• Ensure a consistent understanding of the customer – across all businesses
• Distribute information within and across business units in an easy, reliable and cost effective manner.
Technologies used within Banking

Technologies used

✓ PCs
✓ Distributed (aka Unix)
✓ Large Scale Computing Platforms (aka Mainframes)

Why mainframes?

– Large-scale transaction processing (thousands of transactions per second)
– Support thousands of users and application programs
– Simultaneously accessing resources
– Terabytes of information in databases
– Large-bandwidth communications
Thoughts on Outsourcing

- Need to control strategic systems & processes

- JP Morgan Chase has “insourced” more than “outsourced” over the past several years.

- JP Morgan Chase has offices in over 50 countries – we live and work in a global marketplace

- Most companies outsource within the United States

“The issue boils down to how much of a company's IT should be outsourced. If you're J.P. Morgan, the answer is: almost none of it.”

Business Week (Oct, 2004)
Why Insourse?

Become the Low Cost Provider / Drive Down TCO

• Reduce TCO through consolidation of existing internal applications
• Streamline/re-engineer business processes

Support proprietary initiatives

• Rapid deployment - control prioritization of development efforts
• End-to-end ownership enables protection of proprietary info

Control Our Architecture

• Improve service levels (predictability, simplicity, future growth) by streamlining end-to-end systems and processes
An example of a Large System
The Credit Card Core Processing System

What does a “core processing system” do?
– Authorizations (your card “swipe”) - Credit & Fraud analysis
– Statements - transaction history, rewards
– Customer & Account System of record
– Interact with Visa & Mastercard

The Size of a “Large” Chase System
– 68K Source Code Components
– 10K+ COBOL Executables
– 12K+ batch jobs
– 50TB of data (5,000 Production Database Areas)
– 100M Accounts
– 15M POS decisions a day (peak of 400/sec, each taking 100MS)
The Key Characteristics of a Card Processing System

Architecture

– IMS & DB2 Databases
– Separate Customer and Account structure
– Option (business) driven
– Highly available, 24x7
– Information accessible via: Native 3270, MQ and XML
– Workflow management to streamline customer service
– Parallel nightly batch processing
– Real-time credit & risk evaluation (in less than 200 milliseconds)
Data Model for a core processing system
## Data for a core processing system

<table>
<thead>
<tr>
<th>Entity Name</th>
<th>Entity Definition / Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCOUNT HISTORY</td>
<td>Account history root segment</td>
</tr>
<tr>
<td>CYCLE HISTORY</td>
<td>This is current cycle information that is rolled into history. A new history segment gets created for each new cycle.</td>
</tr>
<tr>
<td>HISTORY DETAILS</td>
<td>History Details of the three types of “Actions” or “Maintenance” - Promise to Pay, Credit Limit and Dispute, performed on an account (e.g. raise Credit Limit, Accept a promise to pay or put a transaction into dispute). A new segment is created whenever t</td>
</tr>
<tr>
<td>BALANCE HISTORY</td>
<td>This is the history from the balance segment. At Cycle, a copy of the balance segment is rolled into this segment, creating a segment for each transaction category (TCAT).</td>
</tr>
<tr>
<td>TERMS BALANCE</td>
<td>This is history of the balance history rolled up to the terms balance (TBAL) level. The multiple TCATs can roll up into one or more TBALs.</td>
</tr>
<tr>
<td>TRANSACTION LEVEL PROCESSING</td>
<td>This is history from the transaction level processing. This is for money that is either Deferred or Same as Cash. An example of “Deferred” is no interest, no payment for 6 months. “Same as Cash” is payment only, no interest.</td>
</tr>
<tr>
<td>ACCOUNT MASTER</td>
<td>Account Master root segment. This is the largest IMS segment for account. This segment stores over 80 Account Level fields that provide information about the status of the account. For example, Past Due, Activity today, Fraud block, etc.</td>
</tr>
<tr>
<td>ACCOUNT TO CUSTOMER RELATIONSHIP</td>
<td>The purpose of this segment is to link the Card Number to the Customer ID. The segment contains Card Level information, one for every customer, regardless of whether or not the card numbers are the same or different.</td>
</tr>
<tr>
<td>ACTIVE ACCOUNT INFORMATION</td>
<td>This contains active account information. It provides a picture of the total balance on the account. It contains past due, over the limit information and many date fields such as date of last purchase, date of last cash, date of last debit, date of last.</td>
</tr>
<tr>
<td>ACCOUNT ACTIVE DISPUTES</td>
<td>This is the active dispute segment. There is one segment for each transaction in dispute.</td>
</tr>
<tr>
<td>BALANCE SEGMENT INFORMATION</td>
<td>This is the balance segment. There is one segment for each TCAT that contains money (Examples: Fee Money, Purchase Money and Cash).</td>
</tr>
<tr>
<td>TRANSACTION LEVEL PROCESSING</td>
<td>This segment records information on “Same as Cash” or “Deferred” transactions. There is a segment for each “Same as Cash” or “Deferred” transactions.</td>
</tr>
<tr>
<td>COLLECTION INFORMATION</td>
<td>This segment is redefined in three ways to contain the three types of collection information: Past Due, Over the Limit and Promise to Pay.</td>
</tr>
<tr>
<td>CHARGEOFF SEGMENT</td>
<td>This segment is generated when an account charges off. It contains information about the chargeoff. There is no occurrence prior to an account charging off.</td>
</tr>
<tr>
<td>BALANCE INFORMATION</td>
<td>This segment contains the balance information at the time of the account charge off.</td>
</tr>
<tr>
<td>CARD NUMBER SEGMENT</td>
<td>This segment contains Fraud and Warning Bulletin information. There is one segment for every card number. This is the parent segment for Warning Bulletin Information and Security Report Information.</td>
</tr>
<tr>
<td>Entity Name</td>
<td>Entity Definition / Description</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>WARNING BULLETIN INFORMATION</td>
<td>This is Warning Bulletin Information. This is information about blocking the account at the association level.</td>
</tr>
<tr>
<td>SECURITY REPORT INFORMATION</td>
<td>This is security report information and it can only occur once. It contains information collected from a customer for</td>
</tr>
<tr>
<td></td>
<td>incidents such as lost / stolen card.</td>
</tr>
<tr>
<td>ACTIVE HISTORY INFORMATION</td>
<td>This is active history segment.</td>
</tr>
<tr>
<td>CRITICAL CUSTOMER DATA</td>
<td>Critical Customer Information. This segment was created for performance reasons to enable storing Customer Information</td>
</tr>
<tr>
<td></td>
<td>in the Account Database so that batch processes do not have to access the Customer Database.</td>
</tr>
<tr>
<td>ACCOUNT MASTER EXTENSION</td>
<td>Account Master Extension root segment. It contains account data that is stored separately to facilitate easy access. The</td>
</tr>
<tr>
<td></td>
<td>segment was created for performance reasons because the Account Master Database already contains a lot of data.</td>
</tr>
<tr>
<td>PRODUCT CHANGE</td>
<td>This segment is used when the Customer trades from one product to another product (e.g. Master Card to Visa Card to AAA</td>
</tr>
<tr>
<td></td>
<td>card to Disney Card.) It records a snapshot of many account level fields at the time of the product change.</td>
</tr>
<tr>
<td>WARNING BULLETIN HISTORY</td>
<td>This is an extension record that contains additional data related to the Warning Bulletin contained.</td>
</tr>
<tr>
<td>INCENTIVE PROGRAM HISTORY</td>
<td>Consumer Rebate/Reward segment, one occurrence at Cycle. This contains a history of incentive programs such as</td>
</tr>
<tr>
<td></td>
<td>Consumer Rebate and Consumer Reward.</td>
</tr>
<tr>
<td>EXT. SECURITY FRAUD REPORT INFORMATION</td>
<td>This is additional Fraud Security Information.</td>
</tr>
<tr>
<td>PRIVACY ACT DATA</td>
<td>This is where privacy and opt-out information is stored.</td>
</tr>
<tr>
<td>AUTHORIZATION BASE</td>
<td>This is the authorization root segment. It contains authorization information such as total authorization performed on the</td>
</tr>
<tr>
<td></td>
<td>account (number approved today, number declined today etc.).</td>
</tr>
<tr>
<td>AUTHORIZATION DETAILS</td>
<td>This is a child segment that is created for every authorization. There is one created for each authorization. The purpose</td>
</tr>
<tr>
<td></td>
<td>is to withhold money from the available credit until the transaction posts. After the transaction posts, this segment is</td>
</tr>
<tr>
<td></td>
<td>deleted.</td>
</tr>
<tr>
<td>AUTHORIZATION HISTORY BASE</td>
<td>This is the authorization history root segment. It contains authorization information such as total authorization performed</td>
</tr>
<tr>
<td></td>
<td>on the account (number approved today, number declined today etc.).</td>
</tr>
<tr>
<td>ACCOUNT CROSS REFERENCE</td>
<td>This segment ties together the keys of several databases so that if one key is known, the rest can be looked up. It ties</td>
</tr>
<tr>
<td></td>
<td>the application number, customer number and rewards number together.</td>
</tr>
<tr>
<td>ACCOUNT TRANSACTION INFORMATION</td>
<td>This is the root segment for the account transaction database.</td>
</tr>
<tr>
<td>TRANSACTION CYCLE INFORMATION</td>
<td>This segments contains information that delineates the beginning of the transaction cycle. The key is &quot;Date Statement</td>
</tr>
<tr>
<td></td>
<td>Begin&quot;.</td>
</tr>
<tr>
<td>TRANSACTION INFORMATION</td>
<td>This segment contains the transactions for the cycle.</td>
</tr>
<tr>
<td>TRANSACTION EXTRA INFORMATION</td>
<td>This is a child segment of the Transaction Information segment.</td>
</tr>
<tr>
<td>ADDITIONAL TRANSACTION DETAIL</td>
<td>Additional transaction detail segment containing SKU type information on retail transactions, including Department ID,</td>
</tr>
<tr>
<td></td>
<td>SKU level description, item quantity, item price and SKU or tracking number.</td>
</tr>
</tbody>
</table>
### Entity Name | Entity Definition / Description
--- | ---
**CUSTOMER MASTER** | Customer database root segment. There is one segment for each customer on an account. The keys are Client Number and Customer ID. Customer Name, SSN, Zip and Phone numbers are also stored and serve as secondary keys for lookup.

**NAME LOOKUP** | This is the Customer Cross-reference segment used for searching (using the ICAL screen) on Customer information (search by Customer Name) to locate specific accounts. This is also regarded as a secondary index.

**SSN LOOKUP** | This is the Customer Cross-reference segment used for searching (using the ICAL screen) on Customer information (search by SSN) to locate specific accounts. This is also regarded as a secondary index.

**ZIP LOOKUP** | This is the Customer Cross-reference database used for searching (using the ICAL screen) on Customer information (search by Zip) to locate specific accounts. This is also regarded as a secondary index.

**PHONE LOOKUP** | This is the Customer Cross-reference database used for searching (using the ICAL screen) on Customer information (search by Phone number) to locate specific accounts. This is also regarded as a secondary index.

**CUSTOMER ADDRESS** | This is the Customer Address segment.

**CUSTOMER TO ACCOUNT** | This establishes the relationship between the customer and specific account numbers. Since the account number is part of the key, there is a different segment for each account number.

**CUSTOMER TO CUSTOMER RELATIONSHIP** | Establishes the relationship between the primary Customer and additional Customers. For example, Father, Daughter etc.

**CUSTOMER NOTES** | This is a Customer notes segment.

**CUSTOMER NOTES TEXT** | This is a child of the Customer Notes segment and it contains the details of the notes.

**PERSONAL INFORMATION** | This segment contains Customer Specific personal information such as hobbies, marital status, checking and savings account numbers, widowed etc.

**EMPLOYMENT INFORMATION** | This segment contains employment information.

**CUSTOM DATA** | This is Customer custom data segment.

**SKIP TRACE INFORMATION** | This is Collections skip trace segment. This is used to share information with other Credit Card Companies for Collection or recovery purposes with regards to delinquent accounts.

**EVENT HEADER** | Event root segment.

**EVENT DETAILS** | This data is used for Batch Event Reporting.

**EVENT DATA** | This segment contains data for on-line viewing of events, memos and memory messages. There are rules that determine how long the information is kept, based on the event type.

**SCORE** | This is the scores root segment.

**SCORE DATA** | This is the scores segment that contains detail information for each score type. The scores within this segment can be used within risk calculations.

**BUREAU SCORE** | This contains supplemental information for a bureau score. It contains some limited bureau information, such as number of disputed trade lines, trade lines and public records. A trade line is an entry in the credit bureau record for a customer.
<table>
<thead>
<tr>
<th>Entity Name</th>
<th>Entity Definition / Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PURCHASE ANALYSIS</td>
<td>Purchase analysis root segment.</td>
</tr>
<tr>
<td>ACCOUNT PURCHASES</td>
<td>Account level purchase analysis details for general purchases and cash only, life to date, cycle to date and prior year.</td>
</tr>
<tr>
<td>PURCHASES BY MERCHANT</td>
<td>Purchase analysis details by transaction summary code which is selected in the purchase analysis option. For example car rental, medical expenses. Each category will generate one of these segments.</td>
</tr>
<tr>
<td>APPLICATION ENTRY</td>
<td>This is the Application Entry Root Segment. This segment is the first IMS segment to be created when an application gets into new account analysis. It contains a number of application information that eventually goes to the account master segment</td>
</tr>
<tr>
<td>APPLICATION CUSTOMER INFORMATION</td>
<td>This is the Application Customer Information segment.</td>
</tr>
<tr>
<td>APPLICATION AUTO DECISION ACTION</td>
<td>Application Auto Decision Action segment. This segment is used for New account analysis internal processing, for example queuing, generating correspondence etc.</td>
</tr>
<tr>
<td>APPLICATION DATA ELEMENT VALUE DATA</td>
<td>Application Data Element Value Data segment. This is a client defined data element segment. The client defined data element aid the application of business rules, for example assigning scores for having a checking account.</td>
</tr>
<tr>
<td>APPLICATION MEMOS</td>
<td>This segment stores Application Memos. These are the memos generated during the processing of the application.</td>
</tr>
<tr>
<td>APPLICATION STATUS DATA</td>
<td>This is the application status data segment that stores the status of the application.</td>
</tr>
<tr>
<td>MESSAGE DATA</td>
<td>This is the application Message Data segment.</td>
</tr>
<tr>
<td>APPLICATION ALERT HITS</td>
<td>This is the Alert Hit Data segment. The alert file is a fraud database. If an application comes in that has an entry in the alert file (e.g. phone number, address), we will get an alert hit.</td>
</tr>
<tr>
<td>ALERT INFORMATION</td>
<td>This is the alert file root segment. This stores the data in the alert file.</td>
</tr>
<tr>
<td>ALERT ACTIVITY INFORMATION</td>
<td>This is an individual alert file segment either captured by the system or entered by an operator. It contains information about individual fraudulent names, addresses, phone numbers, social security number etc.</td>
</tr>
<tr>
<td>PIN MAILER KEYS</td>
<td>Embossing file segment for PIN mailer key information for individual records on the embossing database.</td>
</tr>
<tr>
<td>STAGE OF CARD REQUEST</td>
<td>Additional card request processing information on the embossing database.</td>
</tr>
<tr>
<td>EMBOSsing DATABASE</td>
<td>This is the embossing database root segment. Any request for cards, PIN, mailer etc are stored in child segments to this segment.</td>
</tr>
<tr>
<td>PIN MAILER</td>
<td>This is the PIN / PIN mailer segment. This stores information related to the PIN.</td>
</tr>
<tr>
<td>CARD MAILER</td>
<td>Embossing database segment containing card mailer information.</td>
</tr>
<tr>
<td>CMS CROSS-REFERENCE INFORMATION</td>
<td>Embossing database cross-reference record. This cross-references the individual Application Number to the embossing file transmission ID.</td>
</tr>
<tr>
<td>CARD INFORMATION</td>
<td>Embossing database segment containing card, formatting, embossing and magstripe encoding information.</td>
</tr>
<tr>
<td>INQUIRY LOGGING</td>
<td>Inquiry database logging root segment. Every operator ID defined will have an inquiry logging segment.</td>
</tr>
<tr>
<td>INQUIRY LOG DETAILS</td>
<td>This segment is created whenever an operator looks at an account level green screen. It contains the timestamp, operator screen looked at, what account, etc.</td>
</tr>
</tbody>
</table>
Course Details
Viability of the Mainframe

Who uses mainframes?
– Most Fortune 1000 companies use a mainframe environment
– 60% of all data available on the Internet is stored on mainframes
– There are more CICS transactions processed than Web pages served

Mainframe Resources:
– A growing issue for mainframe enterprises is the shrinking set of skilled staff to maintain the environment - Gartner 2004/2005.
– Wanted: 20,000 IBM Mainframe Experts - InformationWeek, Aug 2005

Financial Impact to IBM:
Mainframe-related sales likely contribute around 25 percent of IBM’s revenue and, according to some, close to 45 percent of their operating profit.
Factors contributing to mainframe use

- RAS -- *reliability, availability, serviceability*
- Security
- Scalability
- Centralized control
- Workload management
- Partitioning / virtualization
- Continuing compatibility
- Evolving architecture
- Application enablement, complexity, variety
- Potential for thousands of users
# The World Depends on Mainframes

Mainframes are a part of everyday life. When you:

<table>
<thead>
<tr>
<th>Activity</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turn on a light</td>
<td>Get a business loan</td>
</tr>
<tr>
<td>Make a telephone call</td>
<td>Process accounting records</td>
</tr>
<tr>
<td>Use your ATM card</td>
<td>Control inventories</td>
</tr>
<tr>
<td>Put money in a bank</td>
<td>Process payroll</td>
</tr>
<tr>
<td>Rent a car</td>
<td>Update personnel records</td>
</tr>
<tr>
<td>Purchase life insurance</td>
<td>Control an assembly line</td>
</tr>
<tr>
<td>Travel</td>
<td>Control a railroad</td>
</tr>
<tr>
<td>Send a package</td>
<td>Use corporate data bases</td>
</tr>
<tr>
<td>Track in-transit packages</td>
<td>Run a government agency</td>
</tr>
<tr>
<td>Trade stocks</td>
<td>Conduct international business/banking</td>
</tr>
</tbody>
</table>

And more...  

... you are likely using a mainframe!
### What Would You Title This Chart?

<table>
<thead>
<tr>
<th></th>
<th>IBM zSeries Z/OS</th>
<th>Fujitsu Primepower Solaris</th>
<th>Sun 15000 Solaris</th>
<th>HP Superdome HP-UX</th>
<th>IBM p690 AIX5L</th>
<th>Dell Intel SHV W2K</th>
<th>Unisys ES7000 W2K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor Technology</td>
<td>☀</td>
<td>☀</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
</tr>
<tr>
<td>System Performance</td>
<td>☀</td>
<td>☀</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
</tr>
<tr>
<td>Parallel Clustering</td>
<td>☀</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
</tr>
<tr>
<td>Unplanned Downtime (S)</td>
<td>☀</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
</tr>
<tr>
<td>Unplanned Downtime (C)</td>
<td>☀</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
</tr>
<tr>
<td>Disaster Tolerance</td>
<td>☀</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
<td>☁</td>
</tr>
</tbody>
</table>

**Based on September 2003 Updated Model**

- **Best**
- **Worst**
Course Objectives

• Understand key concepts required for large scale application development - which is independent of “mainframe” vs “distributed computing”

• Describe programs from an “Application Architecture” perspective

• Gain some real world insight into the application of “Computer Science” concepts
Course Objectives

Be able to have an intelligent conversation with respect to Large Scale Computing

• Understand the differences when creating applications to run on a PC, traditional web-based applications and “really big” applications.

• Become “fluent” in the TLAs/Buzzwords

• Understand Hype verses reality
Course Topics

1. System Architecture overview
2. Introduction to COBOL and JCL
3. Data: VSAM and Sequential Files
4. Application Architecture of transaction processing systems
5. Databases: Relational and hierarchical databases
6. Using the mainframe as a “server” (MQ, HTTP, J2EE)
7. Transaction Processing – CICS
8. High Availability - Parallel Sysplex
Why Learn Cobol

• Most mainframe programs are written in Cobol

• Most applications are large
  (ex. 10,000,000 lines of code)

• Cobol can be very efficient for certain types of applications
Course Format

• **Lectures** - Review key concepts - includes both applied (ex. compile/link a program) and foundational topics (ex. security)

• **Labs** - Focus on getting familiar with the mainframe environment

• **Online** - helping to create/use our “Wiki” (largescalecomputing.wikispaces.com)
Our Wiki

• Areas of the Wiki
  • Course info (ex. reading assignments, schedule, …)
  • General mainframe computing info (ex. terms & abbreviations,

• One contributes to the Wiki by editing (improving) the web pages, and as appropriate, contributing to the discussion on specific web pages

• It is expected that students will be an active reader / contributor to our web site!
Welcome to the Large Scale Computing Wiki home page.

This website has information on CSCI388 as well as general topics on mainframe computing.

These general topics contain useful information on large scale (mainframe) computing that will be used (and updated) within the course, but the information is not specific to our class. Some of the general topics include:

- **Buzzwords** - A quick look up of acronyms often used in mainframe computing
- **Useful References** and tutorials - Articles and books (typically in PDF format)

**Class announcements:**

Added how to install the 3270 terminal emulator (for using on a home PC)

**Adding Content to this website**

Almost all pages in this website are editable (click on the edit button in the upper right part of the web page). Everyone is encouraged to add content to the website (some of the assignments will require you to update some of these web pages). Feel free to use this test page to learn how to use this Wiki space (there is also help if needed).

Note that in addition to updating/creating web pages, you can (and should) be reviewing the "discussion" tab each of the pages (where you can post comments and reply to others that have posted comments).
Course Grading

• 5 % Lab 1
• 10% Lab 2
• 20% Lab 3
• 20% Exams (two, combined they count 20%)
• 5% Lab presentation
• 10% Wiki update assignment
• 20% Final Exam
• 10% Participation (Wiki & Classroom)
Data Sets
What is a data set?

- **z/OS files are called data sets.** Before you can write data into them, space for data sets must be reserved on disk. The user is involved in specifying the amount of space as well as the formatting of it.

  - A data set is a collection of logically related data records stored on one disk storage volume or a set of volumes.

  - To use a data set, you first *allocate* it.

- A data set can be:
  - a source program
  - a library of macros
  - a file of data records used by a processing program.
Types of Datasets
Partitioned Data Sets (PDS)
Why use a PDS
Naming Conventions

TSO Data Set Name

MFM3984.PAYROLL.COBOL

PAY1
PAY2
PAY3
PAY4
PAY5
Naming Conventions - HLQ

High Level Qualifier
Usually indicates the catalog where the data set is stored

TSO Data Set Name

MFM3984.PAYROLL.COBO
Naming Conventions - User Qualifier

User Determined Qualifier
Designates the type of data or application that is served by the data set
Naming Conventions - Type of Data

- **ASM**: Assembler language source code
- **CLIST**: CLIST procedure
- **CNTL**: JCL job stream used for batch job facility
- **COBOL**: COBOL source code
- **DATA**: Uppercase text data
- **EXEC**: REXX procedure
- **FORT**: FORTRAN source code
- **LOAD**: Executable program module
- **OBJ**: Object module
- **PLI**: PL/I source code
- **TEXT**: Upper and lowercase text data
PDS Members
How data sets are named

• Data set naming convention
  – Unique name (**Maximum 44 characters**)
  – Maximum of 22 name segments: level qualifier
    • The first name in the left: high level qualifier (HLQ)
    • The last name in the right: low level qualifier (LLQ)
    • Level qualifiers are separated by ‘.’
  – Each level qualifier:
    • From 1 up to 8 characters
    • The first must be alphabetical (A-Z) or special (@ # $)
    • The 7 remaining: alphabetical, national, numeric (0-9) or hyphen (-)
    • Upper case only
  – Example: **MYID.JCL.FILE2**  HLQ: MYID  3 qualifiers
Data set concepts

- On a mainframe system, files are referred to as *data sets*.
- Each data set has a name that consists of one or more *qualifiers* separated by periods, like MM01.TEST.COBOL.
- Under ISPF, a data set name consists of a project-id, which is usually your TSO user-id; a library name or group name that can be any name you want; and a type that indicates the kind of data that’s stored in the data set.
- You typically store program development data like source, object, and load files in *partitioned data sets*.
- A partitioned data set can contain one or more *members* that are functionally the same as sequential files.
- Before you can use a new or existing data set, you must *allocate* it.
How data is stored in a z/OS system

• Data is stored on a direct access storage device (DASD), magnetic tape volume, or optical media.

• You can store and retrieve records either directly or sequentially.

• You use DASD volumes for storing data and executable programs, including the operating system itself, and for temporary working storage.

• You can use one DASD volume for many different data sets, and reallocate or reuse space on the volume.
DASD: Use and terminology

- *Direct Access Storage Device* (DASD) is another name for a disk drive.

- DASD volumes are used for storing data and executable programs.

- Data sets in a z/OS system are organized on DASD volumes.
EBCDIC

- The IBM S/360 through to the latest zSeries machines use the Extended Binary Coded Decimal Interchange character set for most purposes.
- This was developed before ASCII and is also an 8 bit character set.
- z/OS Web Server stores ASCII data as most browsers run on PCs which expect ASCII data.
- UNICODE is used for JAVA on the latest machines.
Catalogs
How a catalog is used

- A catalog associates a data set with the volume on which the data set is located.

- Locating a data set requires:
  - Data set name
  - Volume name
  - Unit (volume device type)

- Typical z/OS system includes a master catalog and numerous user catalogs.
Catalog Structure

Master Catalog
Data Set-SYS1.A1 or HLQs (alias) IBMUSER...USER

SYSTEM.MASTER.CATALOG

Master Catalog

User Catalog
Data Set with HLQ=IBMUSER

IBMUSER.A1 IBMUSER.A3
volume (wrk002) unit (3390)

Catalog Structure

User Catalog
Data Set with HLQ=USER

USERCAT.IBM

IBMUSER.IBM
Data Set with HLQ=IBMUSER

volume (wrk001) unit (3390)

USERCAT.COMPANY

User Catalog
Data Set with HLQ=USER

USER.TAPE.A1
volume (012345) unit (tape)
z/OS UNIX file systems

- z/OS UNIX System Services (z/OS UNIX) allows z/OS to access UNIX files.
- A z/OS UNIX file system is hierarchical and byte-oriented.
- Files in the UNIX file system are sequential files and are accessed as byte streams.
- UNIX files and traditional z/OS data sets can reside on the same DASD volume.
UNIX file system structure
TSO
How do we interact with z/OS?

**TSO/E**

- Allows users to logon to z/OS and use a limited set of basic commands. This is sometimes called using TSO in its *native mode*.

**ISPF**

- Provides a menu system for accessing many of the most commonly used z/OS functions.
TSO overview

• Acronym for Time Sharing Option/Extensions (TSO/E)

• Allows users to create an interactive session with z/OS

• Provides a single-user logon capability and a basic command prompt interface to z/OS

• Most users work with TSO through its menu-driven interface, Interactive System Productivity Facility (ISPF)
TSO overview (continued)

- In a z/OS system, each user gets a user ID and a password authorized for TSO logon.

- During TSO logon, the system displays the TSO logon screen on the user’s 3270 display device or TN3270 emulator.

- z/OS system programmers modify the layout and text of the TSO logon panel to better suit the needs of the system’s users.
### TSO/E logon screen

Enter LOGON parameters below:  

- **Userid**: ZPROF  
- **Password**:  
- **Procedure**: IKJACCNT  
- **Acct Nmbr**: ACCNT#  
- **Size**: 860000  
- **Perform**:  
- **Command**:  

Enter an 'S' before each option desired below:  

- Nomail  
- Nonotice  
- Reconnect  
- OIDcard  

<table>
<thead>
<tr>
<th>Key</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF1/PF13</td>
<td>Help</td>
</tr>
<tr>
<td>PF3/PF15</td>
<td>Logoff</td>
</tr>
<tr>
<td>PA1</td>
<td>Attention</td>
</tr>
<tr>
<td>PA2</td>
<td>Reshow</td>
</tr>
</tbody>
</table>

You may request specific help information by entering a '?' in any entry field.
Lab 1

• Hints
  – Return does not equal “CNRL”
  – If needed, review the tutorial in our Wiki