



Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

Regional HMIS
Data Warehouse



Regional Collaboration in the San Francisco and Monterey Bay Area

Ray Allen
Jen Hoey
Community Technology Alliance
Santa Clara County
California



11 County Region of Northern California

Fall 2008 Conference

Atlanta, Georgia

September 24, 2008



Population =
7,617,857

Geographic Area =
10,691 (sq. miles)

Equivalent in size to
the state of
Maryland – bigger
than New Jersey
and Massachusetts

P.I.T. count 2007:
35,902 (undercount)



Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

Who is involved?

- 11 Counties of the San Francisco and Monterey Bay area
- 11 Separate and independent Continuums of Care
- Local Foundation (Charles and Helen Schwab Foundation) brought 11 Continuums of Care together for initial conversation. Foundation interested in getting a regional picture of homelessness
- Each of 11 Continuums represented by County Homelessness Coordinator or person overseeing Continuum
- Foundation agreed to fund facilitation of group of 11 Continuums to discuss possible collaboration around regional data
- Foundation funded first 18 months of work



BACHIC

Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

- **Bay Area Counties Homeless Information Collaborative**
- **Mission:** To better enable policy makers, service agencies, and funders to understand and service the needs of the homeless within the community
- **Goals:**
 - Obtain unduplicated regional count of homeless persons
 - Identify prevalence of cross-county chronic homelessness
 - Understand client movement across continuum boundaries
 - Analyze service usage across continuums
 - Inform funders about effectiveness of sponsored programs in the region
 - Leverage HMIS learning and expertise across multiple communities; increase success factors, reduce risk factors



Initial Challenges

- **All Continuums working on HMIS but in different stages of implementation**
- **No agreement on sharing of client level data across jurisdictions**
- **Many different HMIS software applications in use across the Region so needed a way to pool information from local HMIS's**
- **Some counties locally host systems while others host with vendor**
- **Agreed Data Warehouse concept.**
- **Agreed unsigned “Guiding Principles” NOT “MOU”**



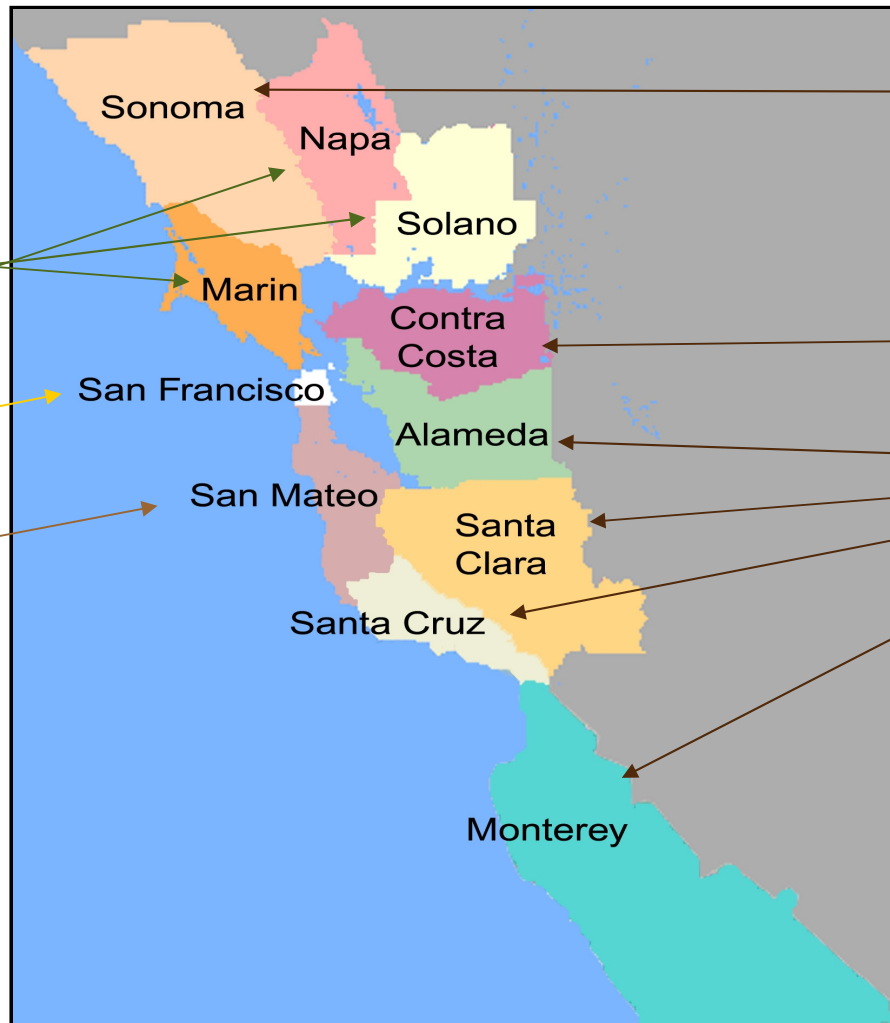
HMIS Implementations by County

Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

- Legacy System
- Homegrown
- Deloitte



ServicePoint
- All locally hosted except for Contra Costa, Monterey and Alameda



Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

RHINO Data Collection

- **Regional Homeless Information Network (RHINO)**
- Consider what data elements needed to be collected in order to meet goals
- BACHIC group agreed to the collection of
 - All Universal Data Elements
 - All Program-Specific Data Elements
- What each county has agreed to forward RHINO
 - All Universal & Program Elements **except** for Protected Personal Information (PPI)
 - Exceptions: Year of Birth, (age range) Program Entry & Exit Dates (length of time in system) and ZIP Code (where client came from)



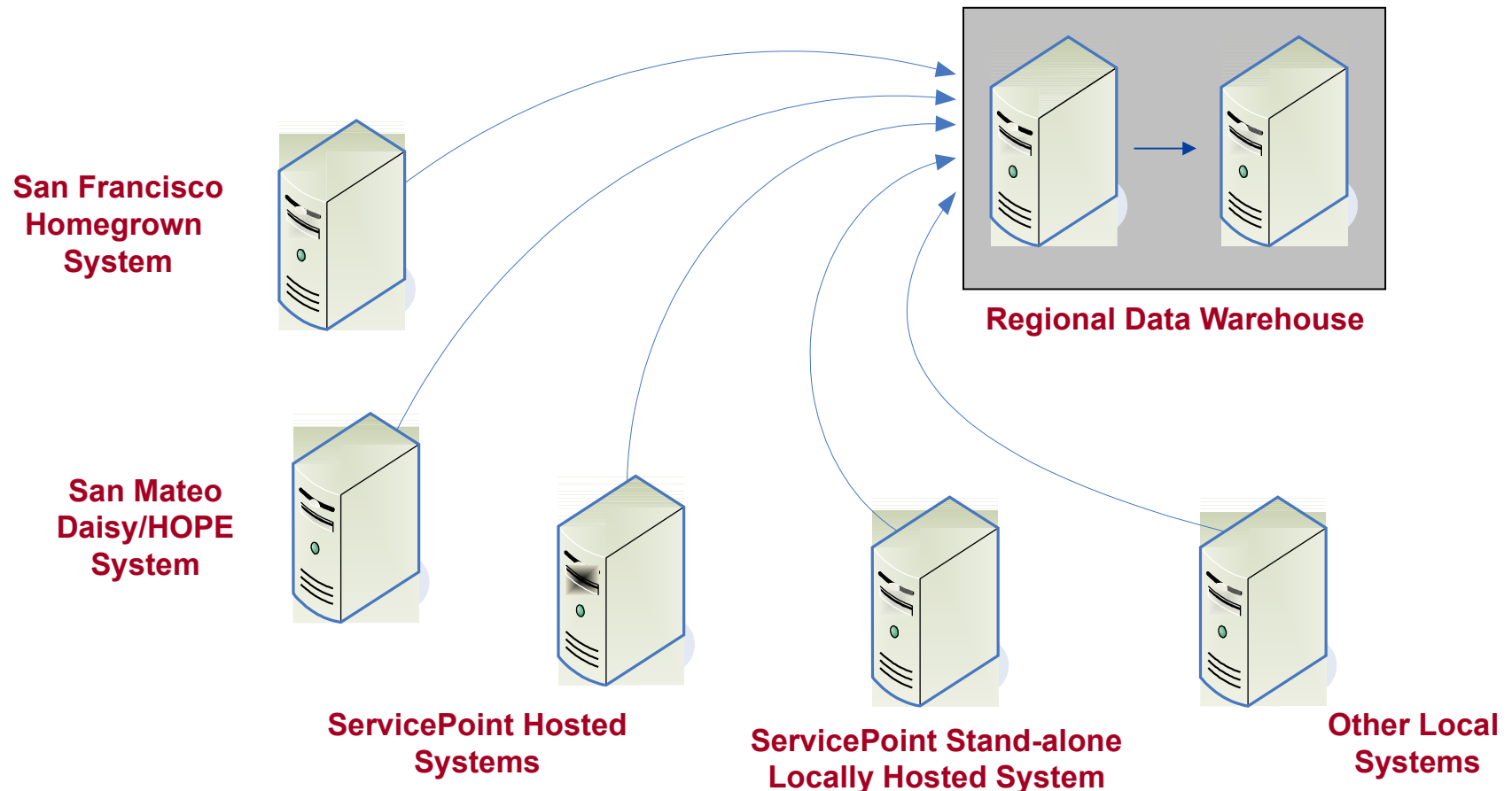
HMIS Data and RHINO Overview

Fall 2008 Conference

Atlanta, Georgia

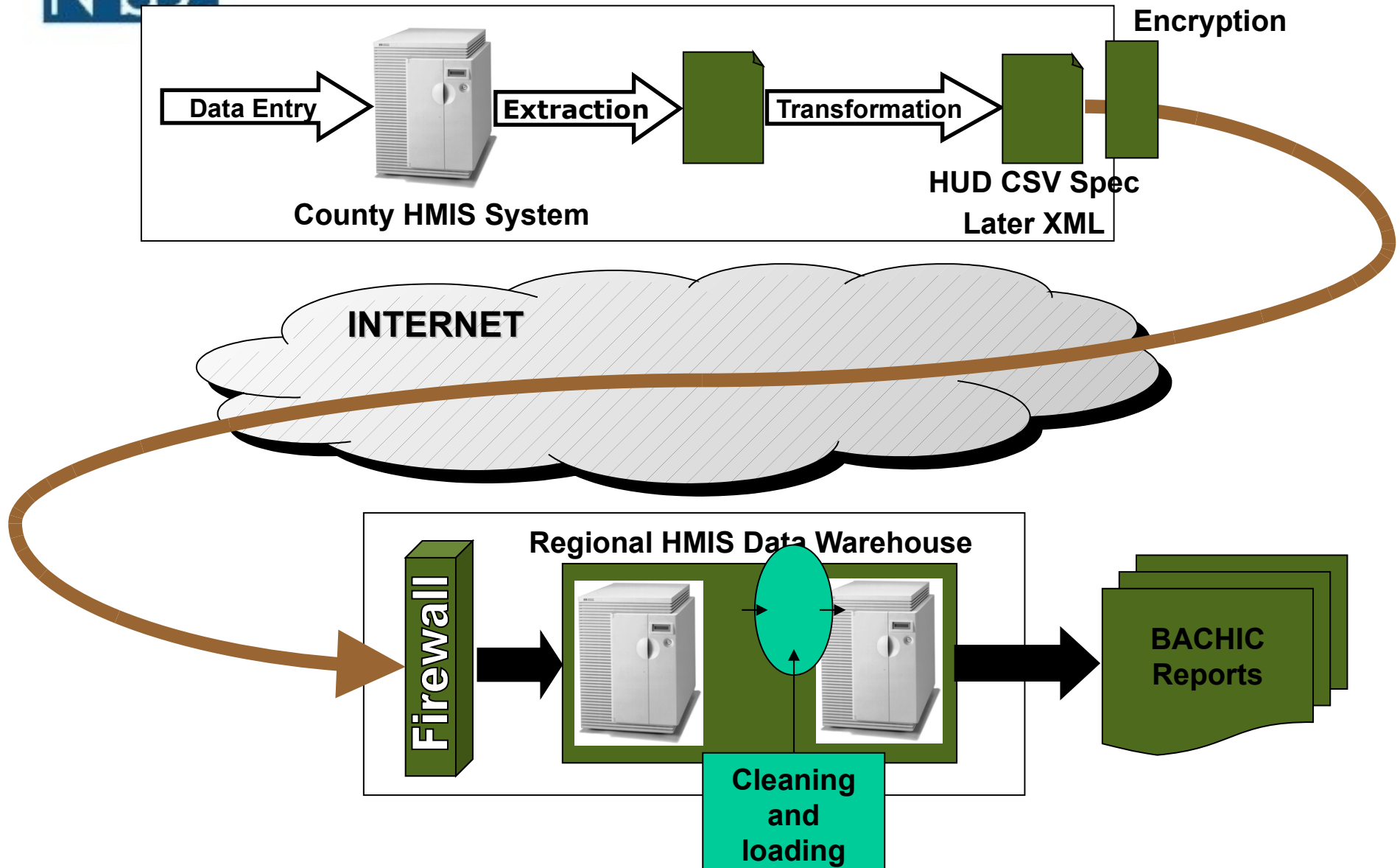
September 24, 2008

Multiple diverse HMIS systems, in different stages of implementations, with different data formats and database systems, provide data to RHINO in a common CSV file format based almost entirely on the HUD CSV File Format specification. Possible future use of XML.





RHINO Design





Data Warehouse System

Regional HMIS Data Warehouse

Firewall



Stage Server



Warehouse Server



DAT72 Backup

- Firewall only allows access from specified CoC computers (using Static IP)
- All access through Firewall restricted to Stage Server
- Stage Server acts as second Firewall
- Data moved from Stage Server to Warehouse Server for processing
- Data backed up regularly onto DAT tapes
- All systems housed in secure hosting facility with 24 hour security, fire protection, video surveillance and biometric secured access



Mid-Project Challenges

- **Wide variation in technical expertise within counties for exporting data**
- **Minimize counties' efforts**
 - **Especially ongoing duties/ obligations**
- **XML vs. CSV – HUD Standards say “must be able to export in CSV at a minimum”**
- **Tight control on values for CSV export similar to constraints provided by XML**
- **Future move to XML**
- **De-duplication of clients between counties**
- **Creation of Regional Unique Identifier**
- **Confidentiality and Security of client data**
- **Reporting**
 - **Flexibility so as not to limit future reporting choices**
- **Work flow**
 - **Processes and procedures for resolving exceptions**



CoC RHINO Data Transformation

- The RHINO CSV specification is the HUD HMIS CSV specification with just a few additions. (Thanks are due to Eric Jahn, Alexandria Consulting, and Brian Sokol, Abt Associates, for writing the underlying HMIS CSV specification!)
- CSV was considered easier by the BACHIC members, and the CSV coding will be largely reused if or when RHINO switches to XML.
- RHINO provided database tables to the CoC's for data validation to the RHINO CSV specification, which is similar to the validation within the XML specification.



Regional Unique Identifier

Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

- Required for de-duplication of customers within 11 counties.
- Uses partial information from personal identifiable data elements.
- Uses a hash algorithm to encrypt ID.
- Letters converted to numbers then scrambled.
- Can not be reverse engineered without the Hash algorithm - since ID does not include complete personal identifiable information, breaking algorithm will not produce personally identifiable information.



Other Challenges

- **No signed binding “Agreements”**
- **Participation dependant on Counties budgets and staff availability for meetings**
- **Agree data dump regularity** – monthly, but challenging due to individual county CoC staffing issues
- **Agree protocols for “allowable values” and “exceptions”**
- **Withdrawal of Foundation funding**
- **Lockheed Martin/HUD funding for “Best Practice” project**
- **Production of Best Practice “White Papers”**
- **No on-going funding for analysis of data**



Reports and Data Ownership

Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

- Each county “owns” its own data
- No client level data shared between counties
- No individual county will have access to Data Warehouse
- Reports will be generated by Warehouse staff
- Only “regional” reports will be generated – local reports to be generated by local HMIS
- Only aggregate reports to be generated – no client level reports
- Definition of terms in Reports



RHINO Travel Event Definition

Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

- **A Travel Event is simply the occurrence of a client receiving services in one CoC, and then later receiving services in another CoC.**
- **We are defining a Travel Event as when a client has entered one CoC (A), and then at a later time appears in another CoC (B), where B is the next time sequential occurrence of the client in a CoC different from A. For example, if a client receives services in A, then B, then back at A, and then in C, the client has three travel events: A-B, B-A, A-C.**
- **A Travel Event is somewhat distinguished from a migration which implies that the client was tending to reside in one area, and is now tending to reside in another area. This type of migration analysis is also possible with a regional data warehouse.**



Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

Travel Event Data Analysis

Travel Event analysis will initially focus on unduplicated regional reporting of travel statistics and service utilization by people who have traveled to receive the services.

- How wide spread is homeless travel across CoC's?
- How often are homeless traveling to different CoC's?
- Where are homeless people coming from, and where are they going?
- Are they drawn by services, by housing opportunity, by other factors?



Clients Traveling to Multiple CoC's

Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

- Description**

Count of clients who have received services
In at most *N-many* counties.

- Data Elements**

Program Start Date, Program End Date,
Service Start Date, Service End Date,
Last Permanent Zip Code

- Counties**

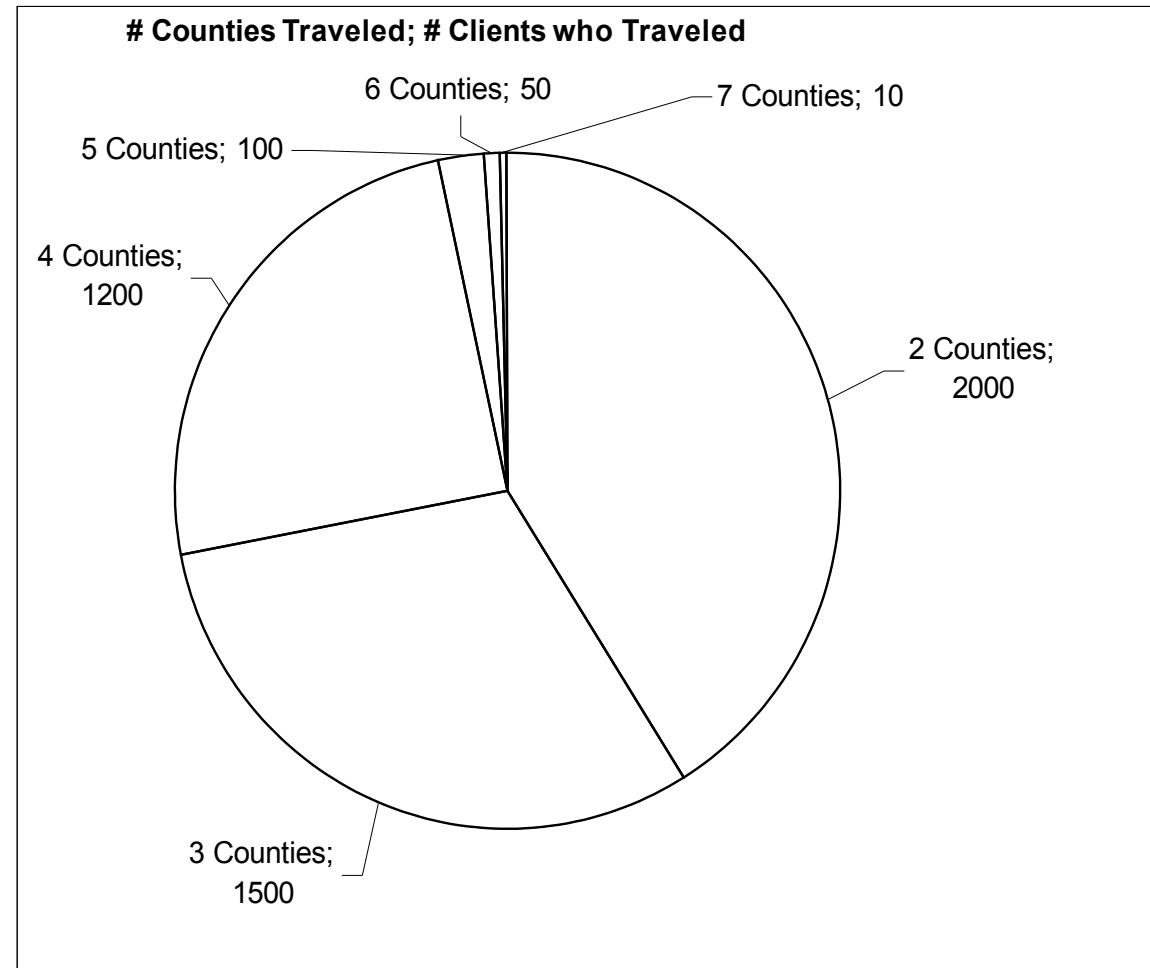
All Counties

- Date Range**

2007

- Data Statistics & Data Limitations (...)**

# Counties Traveled	# Clients Traveled
2 Counties	2000
3 Counties	1500
4 Counties	1200
5 Counties	100
6 Counties	50
7 Counties	10





Future Reporting & Data Analysis Challenges

Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

- **Under the agreed Policies and Procedures, BACHIC must approve all report formats and data releases. Each CoC has a right to opt its data out. Achieving consensus will remain a challenge.**
 - Each county owns its own data
 - What if a county's "weakness" is exposed by data?
 - Does each county have veto power?
 - Withdrawal of any of the counties
 - Deletion of data from the Warehouse
- **CoC participation in BACHIC & RHINO is voluntary, and BACHIC & RHINO will discuss data quality issues with CoC's, including**
 - Program participation rates
 - Completeness of Universal and Program HUD Data Elements
 - General accuracy of data element values by cross checking related data elements.
- **The need for project funding remains a major challenge for RHINO, not only for data analysis and reporting, but also to sustain the data warehouse.**



Contacts

- Community Technology Alliance (RHINO lead agency):
 - 115 East Gish Road, Suite 222, San Jose, CA 95112
 - (408) 437-8800
 - www.ctagroup.org
- Ray Allen, Executive Director, Community Technology Alliance
(408) 437-9170 ray@ctagroup.org
- Jen Hoey, Community Technology Alliance
(408) 451-0731 jen@ctagroup.org



Aggregate Reports Partial List

Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

Aggregate reports will summarize data at the regional level, and will allow for drill down analysis by a variety of groupings, including CoC specific reports. The following two pages are a partial list ...

1. Total client population by Age Groups
2. Total client population by Race
3. Total client population by Ethnicity
4. Adult client population by Gender
5. Adult client population by Income Sources
6. Adult client population by Income Group
7. Adult client population by Disabilities
8. Adult client population by Prior Residence
9. Adult client population by Last Permanent Zip
11. Adult client population by Prior Residence Length of Stay
Adult client population by Homeless Reasons
12. Adult client population by Health Conditions
13. Adult client population by Domestic Violence
14. Adult client population by Homeless Reasons
15. Adult client population by Destination, Subsidies
16. Adult client population by Reason for Leaving
17. Adult client population by Employment
18. Adult client population by Education
19. Adult client population by Veteran Information



Aggregate Reports Partial List cont...

Fall 2008 Conference

Atlanta, Georgia

September 24, 2008

20. Child client population by Child School Information
21. Total client population by Single, Family, Adults, Children
22. Total client population by Family Income Groups
23. Total client population by Family Size
24. Total client population by Program Type
25. Total client population by Service Type
26. Count of Programs by Program Type
27. Count of Services by Service Type
28. Adult client population by HUD Homeless
29. Adult client population by HUD Homeless
30. Adult client population by HUD Chronically Homeless
31. Adult client population by Chronically Homeless (*Inferred from data.*)
32. Total client population by Repeated Homeless Episodes
33. Types of homelessness, by First Time, 2-3 Times, Chronic, Long Term
34. Adult client population by duration of homeless (*Inferred from data.*)