

An Integrated Biological Screening System Using Automated MS and Analytical Informatics

Eric Stauffer
Eli Lilly and Company

Chemistry 696D Guest Lecture

The Challenge

- 150,000+ ligands
- 4 targets
- Pool ligands (following rules)
- Look for binding

Goals

- Never be the rate limiting factor
- Reduce required human interaction
- Simplify sample genealogy
- Maximize pooled sample efficiency
- Platform for reporting tools
- Anticipate change

Compound Pooling

- Based on chemical formula
- Maximize number of ligands that can be distinguished by mass spec
- Eliminate (mask out) 'bad actors'
- Mindful of DMSO concentration
 - Above 1% unacceptable

Manual process

Follow pooling instructions, locate 1 plate in 4,700, aspirate each well from that plate, dispense into "pooling" plate, repeat for subsequent 375,999 wells

At 45 seconds per well, it would take over 300 days of 15-hour days to *pool* the plates.

What's the Solution??

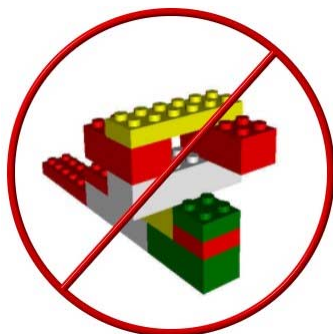
automation.

The Solution

- **Hardware Selection**
- Database Selection
- Compound Pooling
- Robotic Liquid Handling
- Sample Acquisition and Processing

Hardware Selection

- Liquid handlers
- Plate stackers
- Barcode Readers
- Mass Specs
- PCs
- Networking infrastructure



NOT created in the Lego mindset.

The Solution

- Hardware Selection
- **Database Selection**
- Compound Pooling
- Robotic Liquid Handling
- Sample Acquisition and Processing

Data Tracking

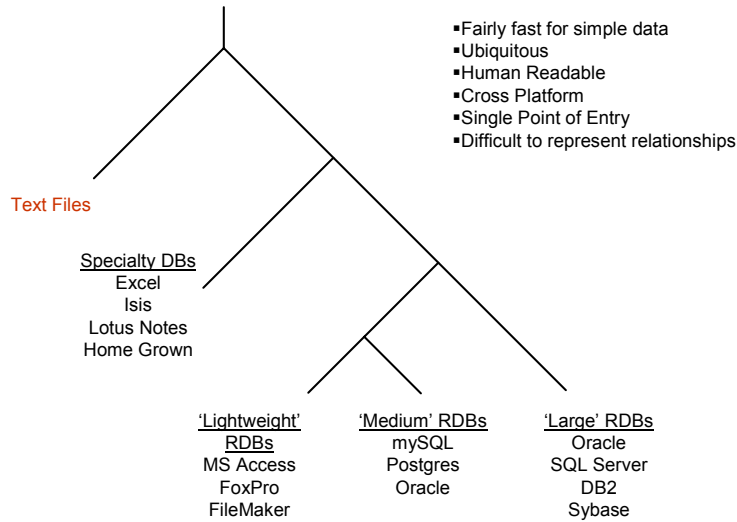
How do we represent all this data?

- Data is not X,Y pairs
- Tracking ancestors and descendents of real world objects
- What has happened to any object at any point in time

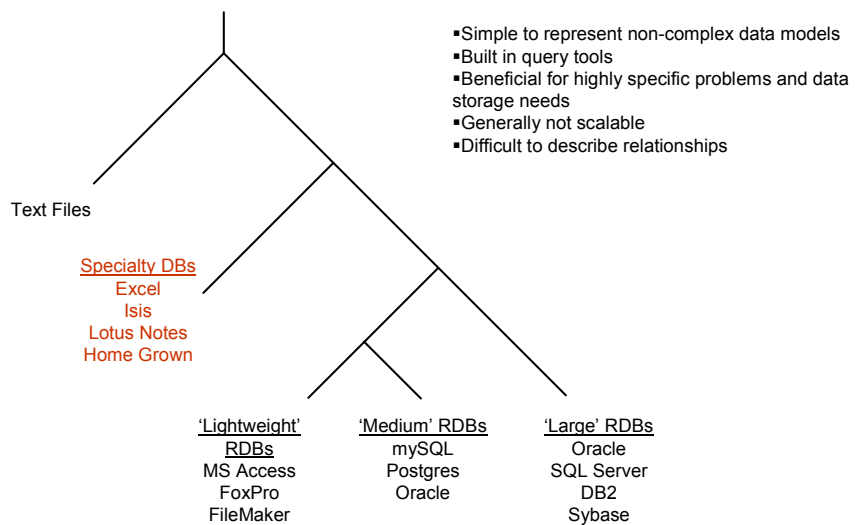
Selecting Tracking Database

- Fast
- Reliable
- Foundation upon which to represent real world associations in a data model (relationships)
- Support considerations
- Network Issues

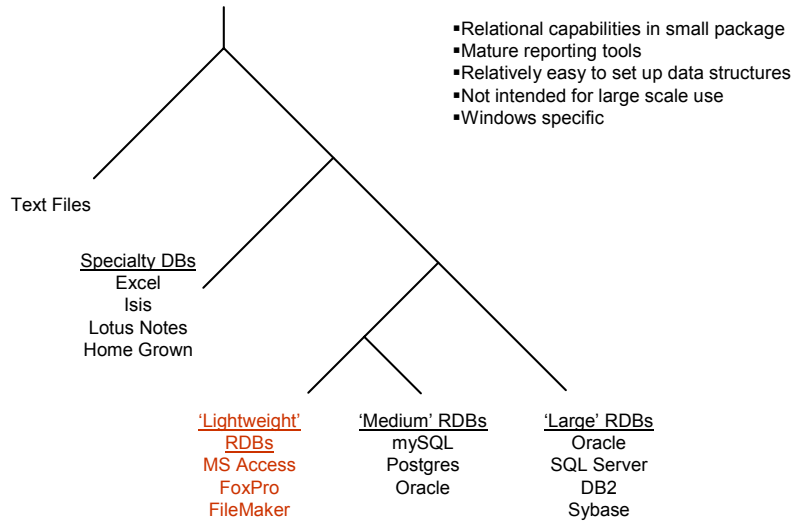
Text Files



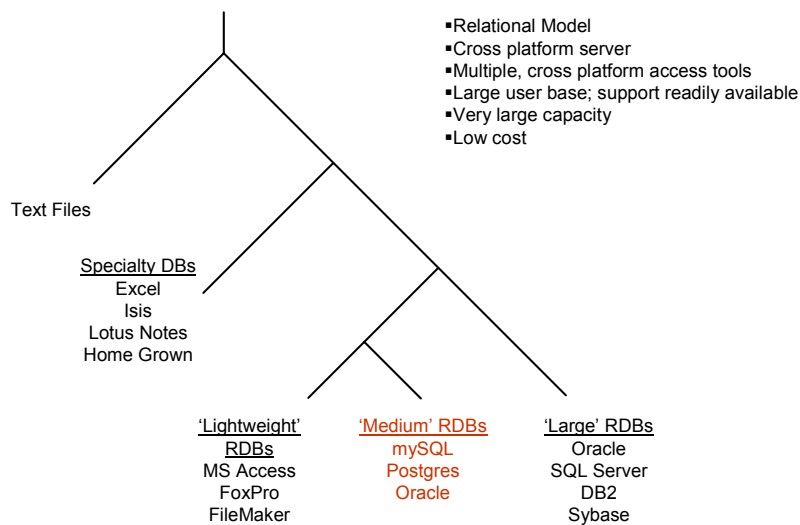
Specialty Databases



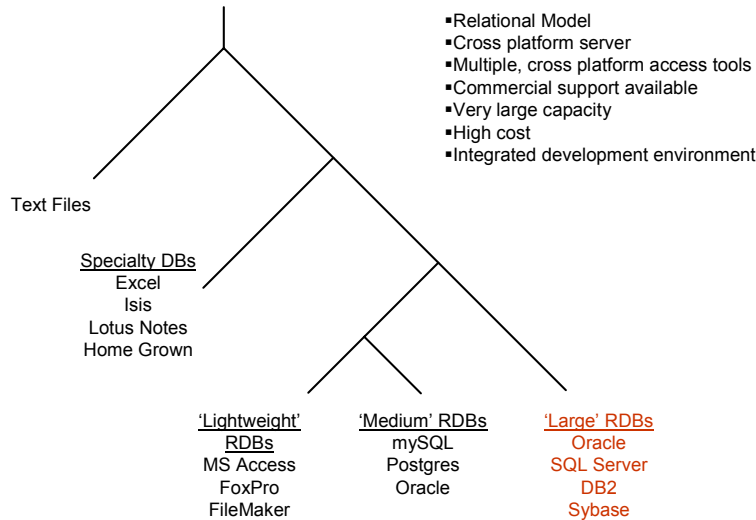
Lightweight Databases



Medium Databases



Large Databases



Contributing Factors

- Ability to represent real world objects in the database
- Fast and dependable
 - 24 x 7 real-time database access
- Multiple clients from various platforms
- Scalable

Data Modeling

Two primary methods to represent our data

- Transactional
- Process Driven

Transactional

- Optimal for logging actions resulting from dynamic processes
 - Checkbook register
 - Real-time data acquisition
 - Point of Sale software

TX ID	Plate ID	Column	Action	Liquid	Volume
1	00098	1	Aspirate	Mobile Phase	12
2	00098	1	Dispense	Sample	150
3	00098	1	Aspirate	Sample	100
4	00098	1	Dispense	Buffer	25
5	00163	1	Aspirate	Mobile Phase	15
6	00163	1	Dispense	Target	125
7	00163	1	Dispense	Buffer	100

Process Driven Model

- Requires defined processes
- Can be burdensome to add new steps
- Reporting on data very straightforward
- Very fast

TX ID	Step ID	Plate ID	Column	Action	Volume	Done
1	1	00098	1	Aspirate	15	6/5/02 23:12
2	2	00098	1	Dispense	100	6/5/02 23:15
3	3	00098	1	Aspirate	125	6/5/02 23:30
4	4	00098	1	Dispense	35	
5	1	00163	1	Aspirate	50	
6	2	00163	1	Dispense	100	
7	3	00163	1	Dispense	25	

The Solution

- Hardware Selection
- Database Selection
- Compound Pooling
- Robotic Liquid Handling
- Sample Acquisition and Processing

Compound Pooling Data

- From what and where?
 - Source Plate ID
 - Source Plate Column
 - Mask

- To where?
 - Destination 'Group'
 - Destination Column

- When was it done?

source_plate_id	source_column	dest_group	dest_column	tip_mask	batch	completed
A000377380	7	3	7	11111111	49	2002-04-23 07:42:57
A000346923	2	1	2	11111111	48	2002-04-22 10:10:57
A000346923	3	1	3	11111111	48	2002-04-22 10:10:57
A000346923	4	1	4	11111111	48	2002-04-22 10:10:57
A000346923	5	1	5	11111111	48	2002-04-22 10:10:57
A000346923	6	1	6	11111111	48	2002-04-22 10:10:57
A000346923	7	1	7	11111111	48	2002-04-22 10:10:57
A000346923	8	1	8	11111111	48	2002-04-22 10:10:57
A000346923	9	1	9	11111111	48	2002-04-22 10:10:57
A000346923	10	1	10	11111111	48	2002-04-22 10:10:57
A000346923	11	1	11	11111111	48	2002-04-22 10:10:57
A000331619	2	2	2	11111111	48	2002-04-22 10:10:57
A000331619	3	2	3	11111111	48	2002-04-22 10:10:57
A000331619	4	2	4	11111111	48	2002-04-22 10:10:57
A000331619	5	2	5	11111111	48	2002-04-22 10:10:57
A000331619	6	2	6	11111111	48	2002-04-22 10:10:57
A000331619	7	2	7	11111111	48	2002-04-22 10:10:57
A000331619	8	2	8	11111111	48	2002-04-22 10:10:57
A000331619	9	2	9	11111111	48	2002-04-22 10:10:57
A000331619	10	2	10	11111111	48	2002-04-22 10:10:57

Pooling Steps

- Prefill Plate
- Read Source Plate
- “Validate” correct group
- Generate automation script
- Pool compound

‘Prefilling’

- Pooling rules eliminate guarantee of 10:1 pooling
- Must account for varying volume

Prefilling

Database query for Group 1 yields

Column	Total
8	9
12	8

Makeup Volume

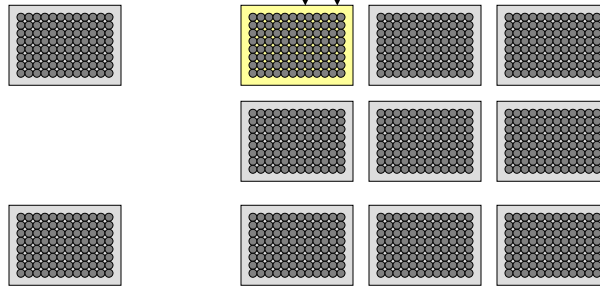


Plate	Column	Destination Group	Destination Column	Mask
00238	1	A1	4	11111111
00238	2	C2	1	11001111

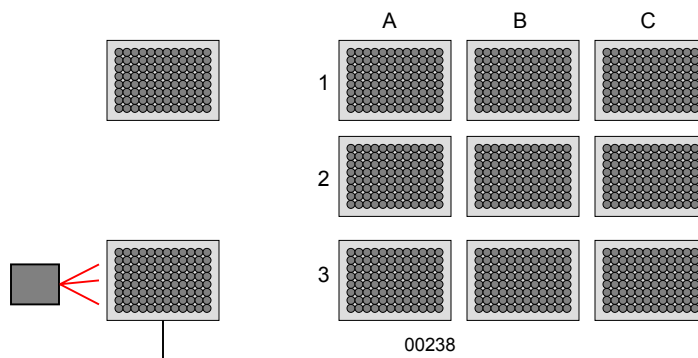
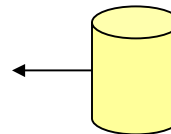


Plate	Column	Destination Group	Destination Column	Mask
00238	1	A1	4	11111111
00238	2	C2	1	11001111

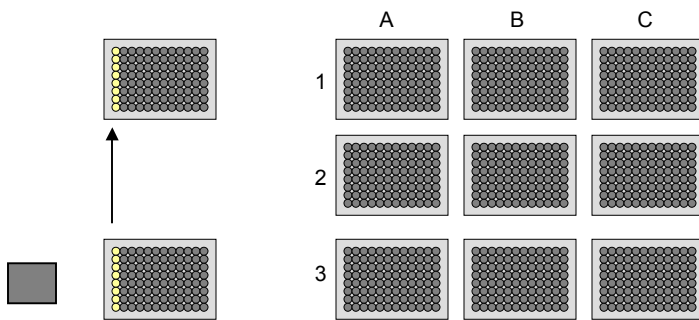


Plate	Column	Destination Group	Destination Column	Mask
00238	1	A1	4	11111111
00238	2	C2	1	11001111

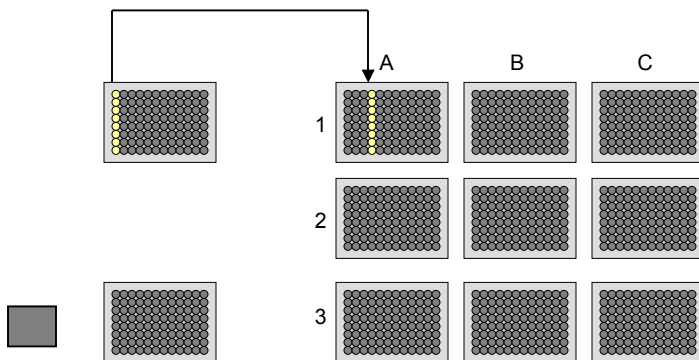


Plate	Column	Destination Group	Destination Column	Mask
00238	1	A1	4	11111111
00238	2	C2	1	11001111

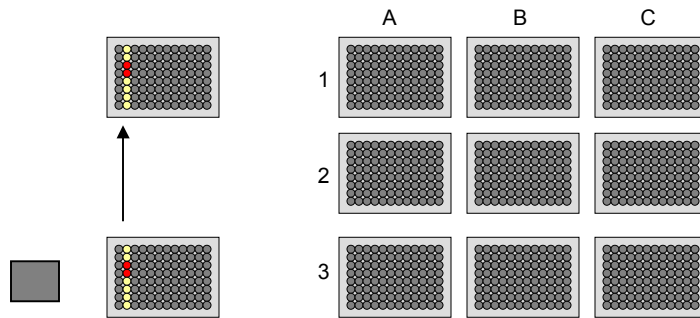
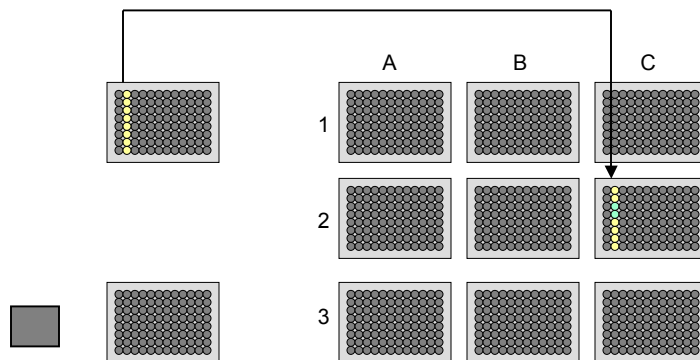


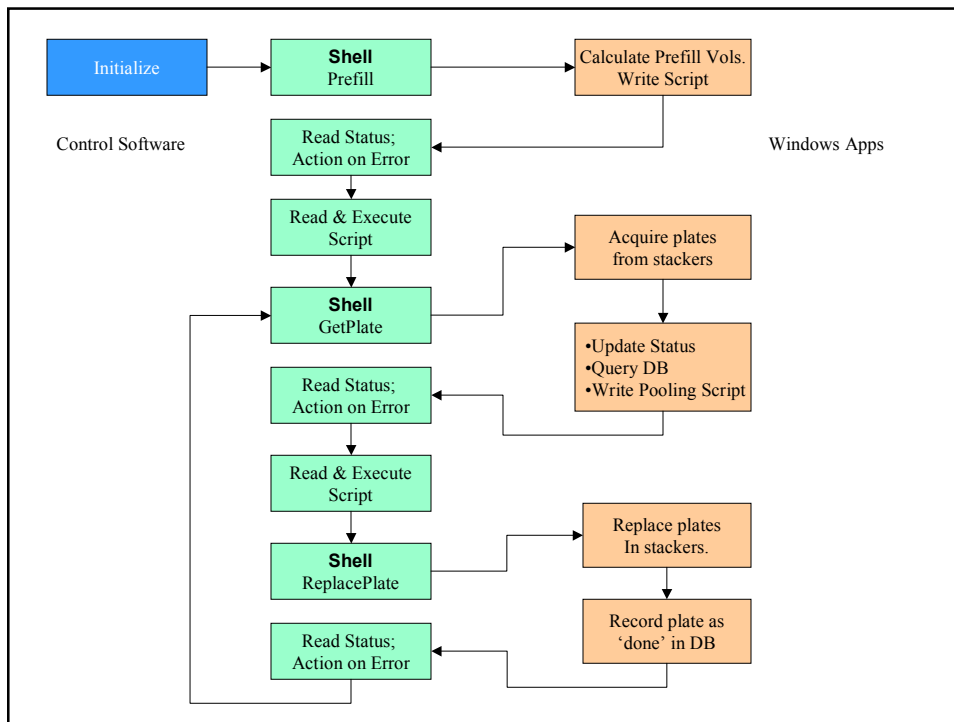
Plate	Column	Destination Group	Destination Column	Mask
00238	1	A1	4	11111111
00238	2	C2	1	11001111



The Pooling Problem

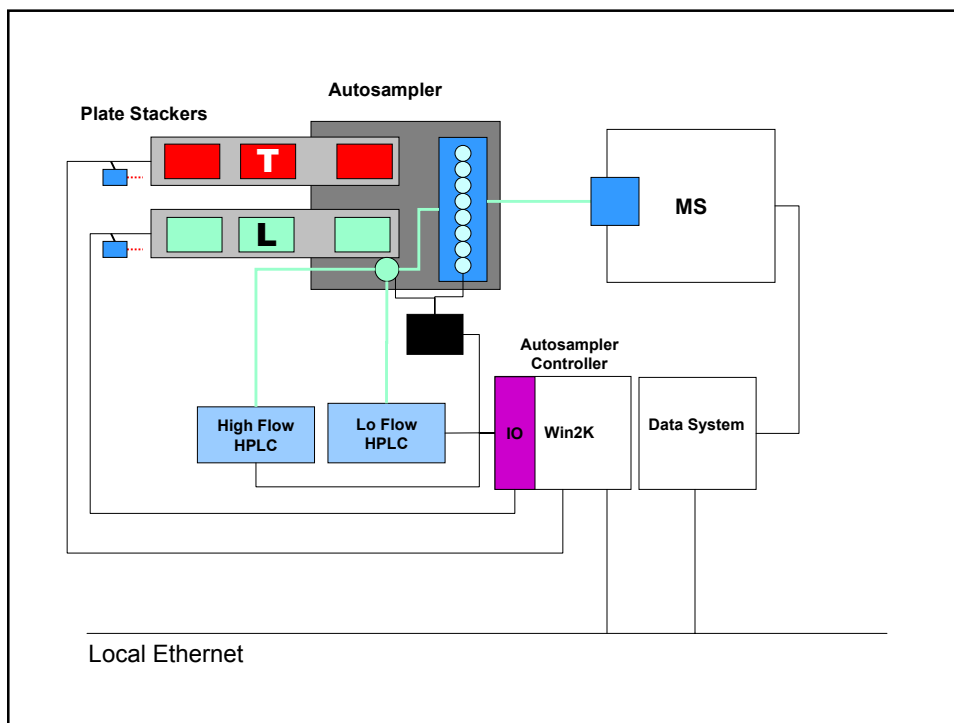
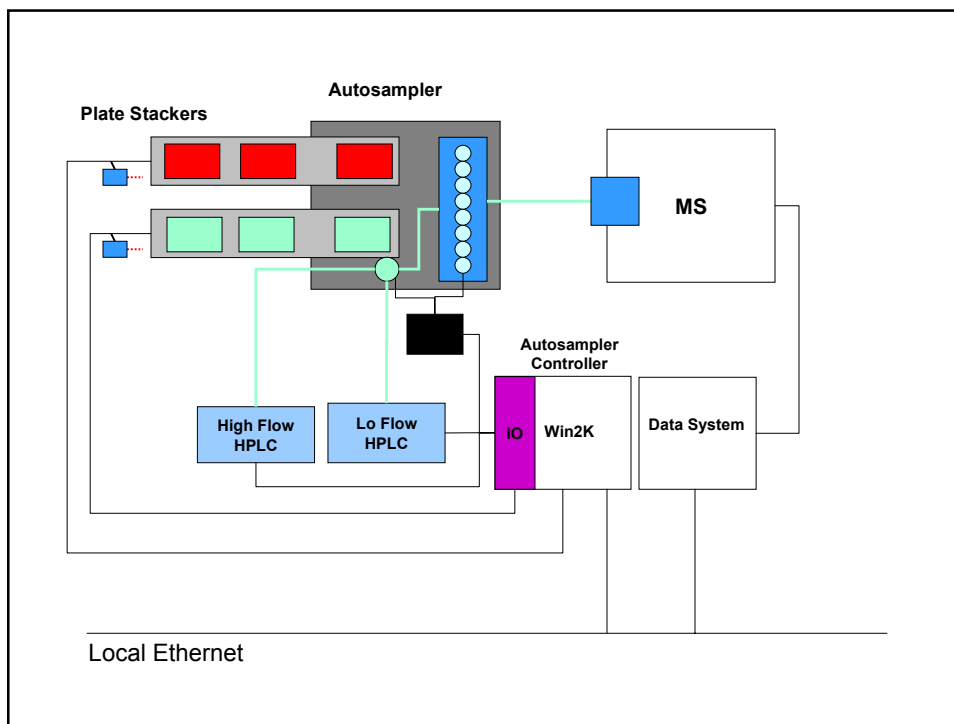
- Goal of pooling 10:1
- 100 source plates @ 12 columns per plate
- Each column represents
 - 3 Aspirations
 - 3 Dispenses
 - 2 mixes
- 9600 distinct activities exclusive of washes and prefill steps
- Need a liquid handling robot that can talk to the database

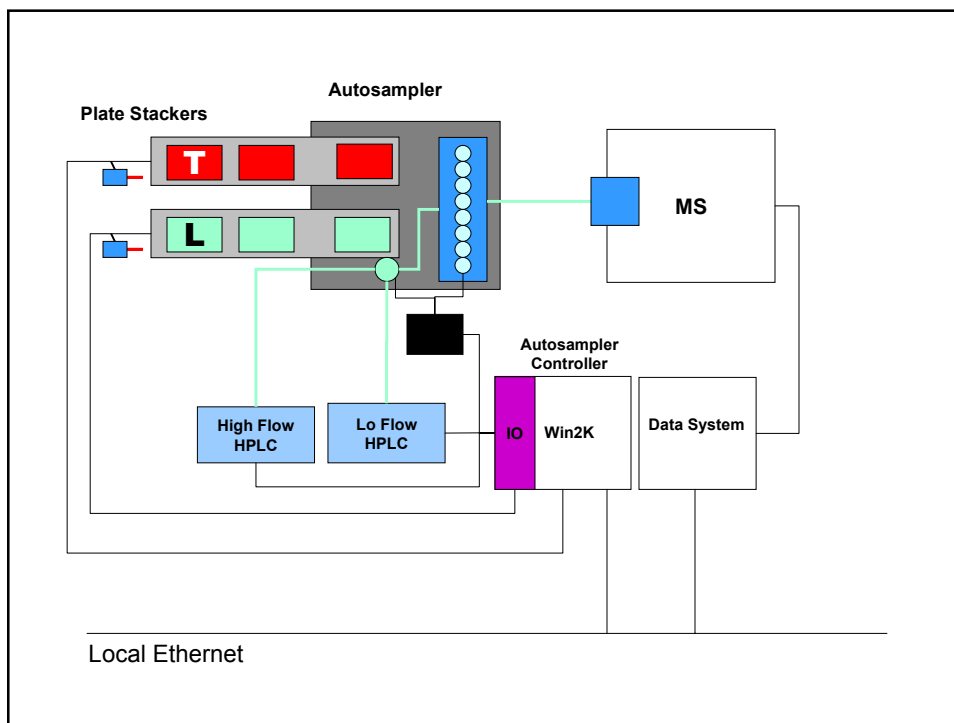
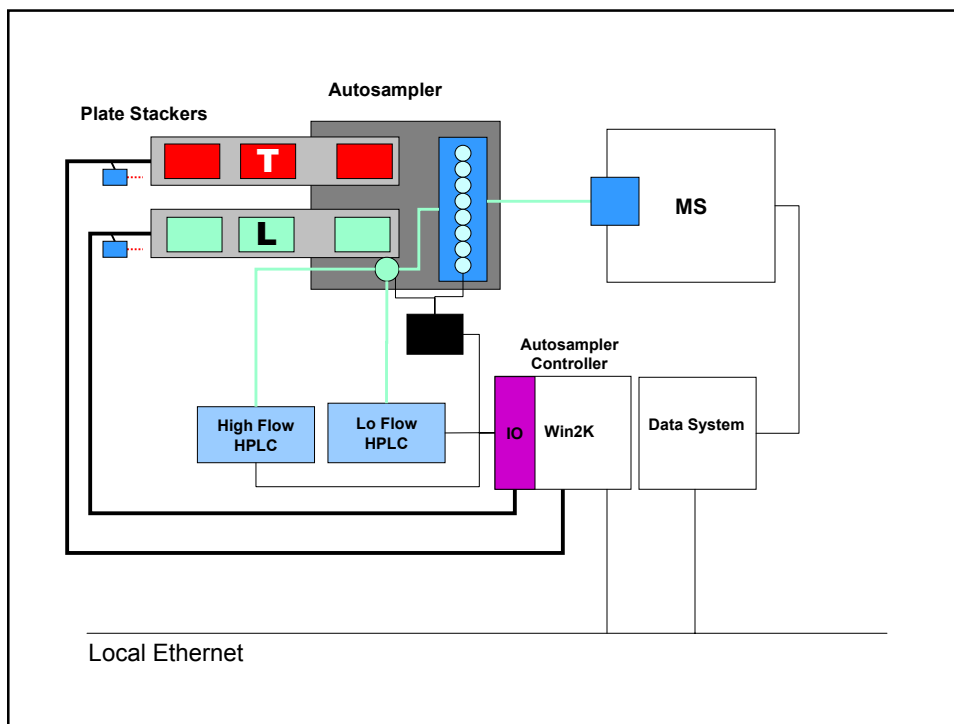
Take advantage of
software that
communicates with the
outside world

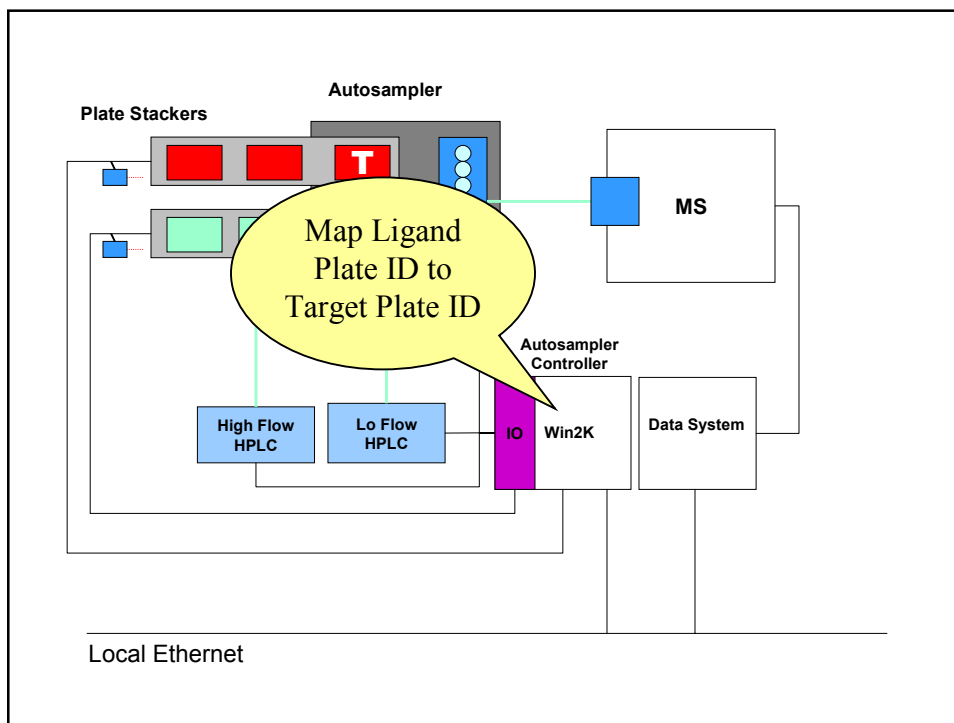
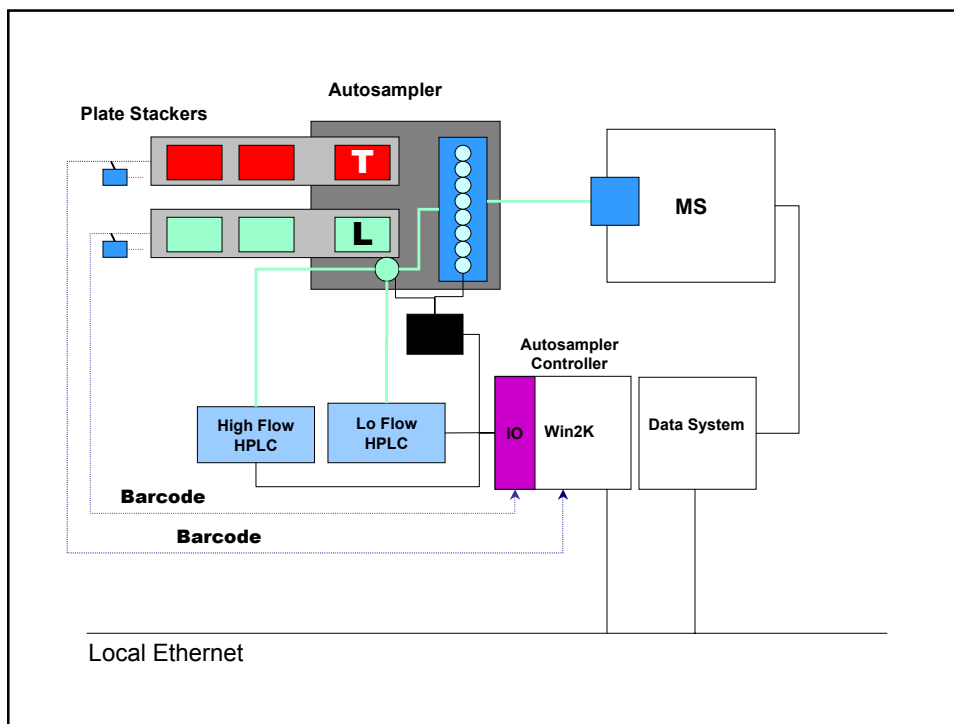


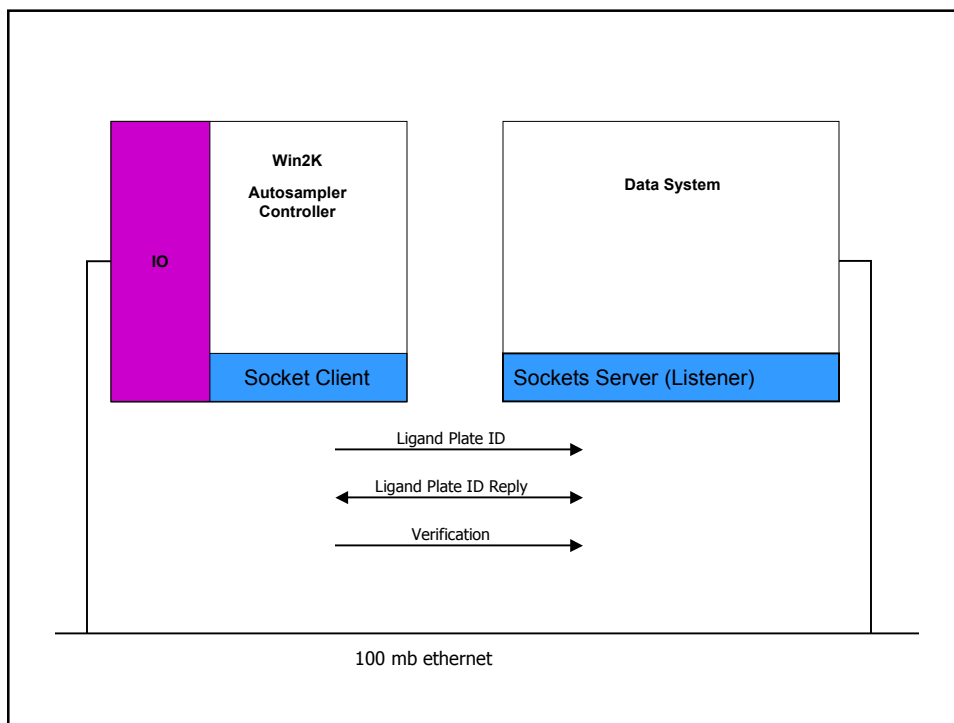
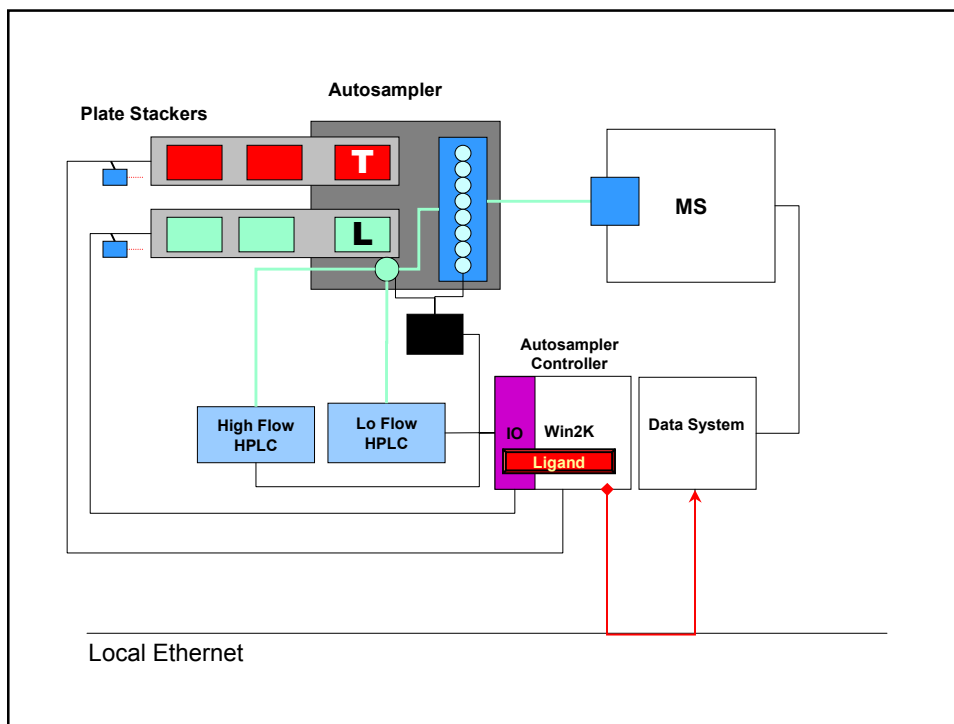
The Solution

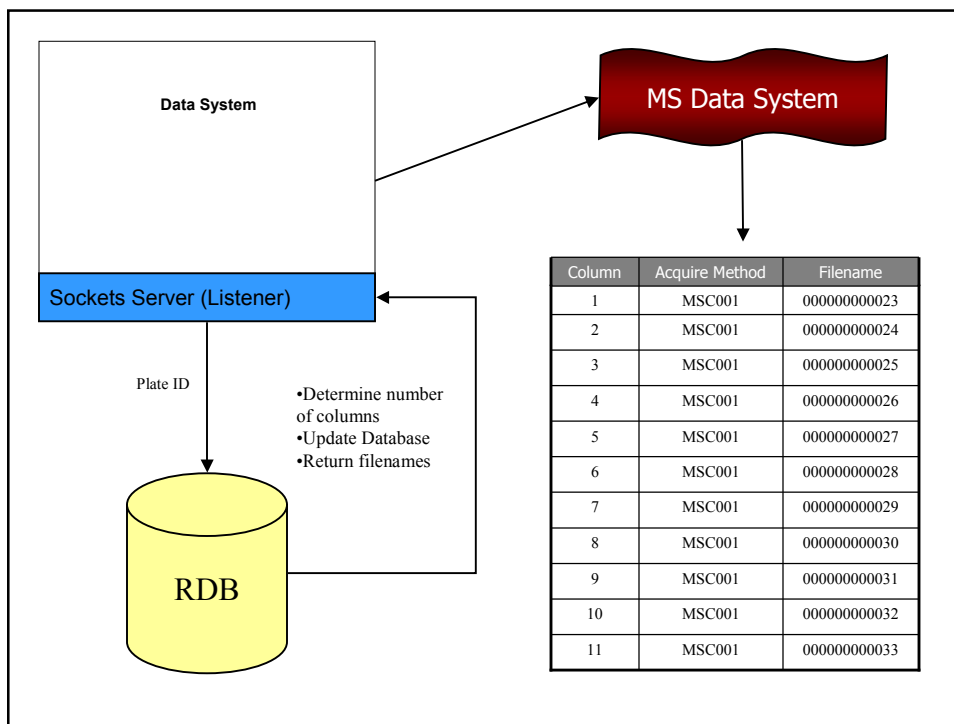
- Hardware Selection
- Database Selection
- Compound Pooling
- Robotic Liquid Handling
- **Sample Acquisition and Processing**











Acquisition Method

- Control Software Specific
- Controls various instrument acquisition parameters.
- *Primary automation concern is in post-acquisition program defined in the acquisition method*

Column	Acquire Method	Filename
1	MSC001	000000000023
2	MSC001	000000000024
3	MSC001	000000000025
4	MSC001	000000000026

MSC001 Says:

1. Collect sample with these instrument parameters
2. Execute the program foo.exe and
3. pass in the filename as a command line parameter.

Power lies in the ability to call a program after each acquisition and pass in an identifier as a command line

What we can now do...

- Validate acquisition quality; stop mass spec acquisition if quality test fails
- Update database marking sample as 'acquired'
- Generate email to sample submitter
- Add acquired sample to 'post-processing' queue
- Query eBay bid status for Dave Matthews tickets

Conclusion

- The database must serve the scientific process
- Software must communicate with the outside world