Cohort Analytics and Cloud Usage: Security Review

Greg Heileman **Terry Babbitt** **Chaouki Abdallah**

Application Development Team
Academic Affairs
University of New Mexico

Cohort Analytics – Overview

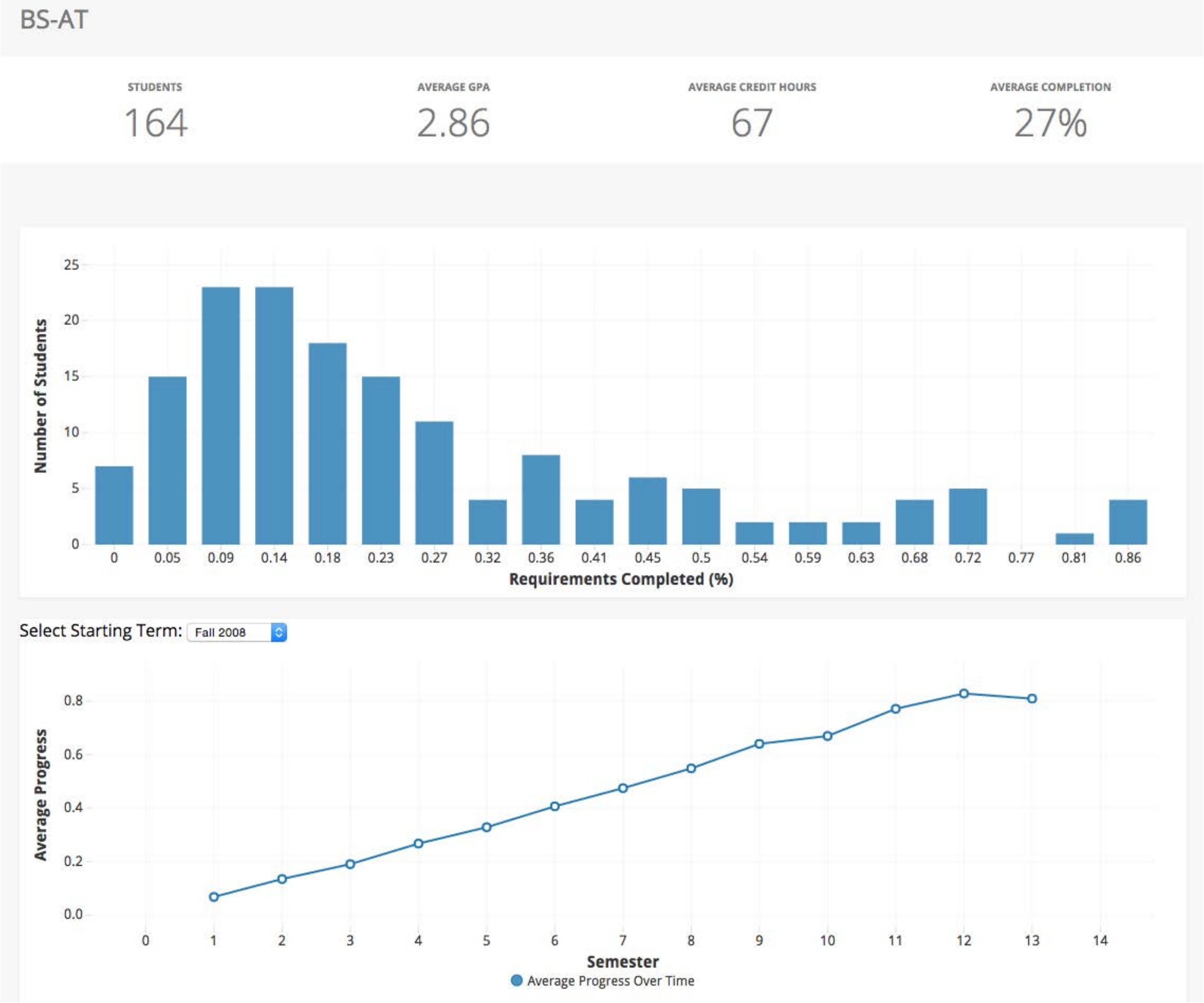
We have developed a cohort analytics application that will dramatically improve our student success capabilities.

This application will enable:

- Advisors, chairs, deans and administrators to track the progress of relevant student cohorts relative to academic progress.
- Earlier insights into various metrics the regents, president, provost have asked us to track. E.g., accurately project the number of students who will graduate in four years (tuition free final semester).
- The ability to set and track program- and college-level success targets.
- More accurate graduation rate projections (years in advance, rather than months in advance of required reporting).

Target Date for Release: August 7, 2015

Cohort Analytics Dashboard

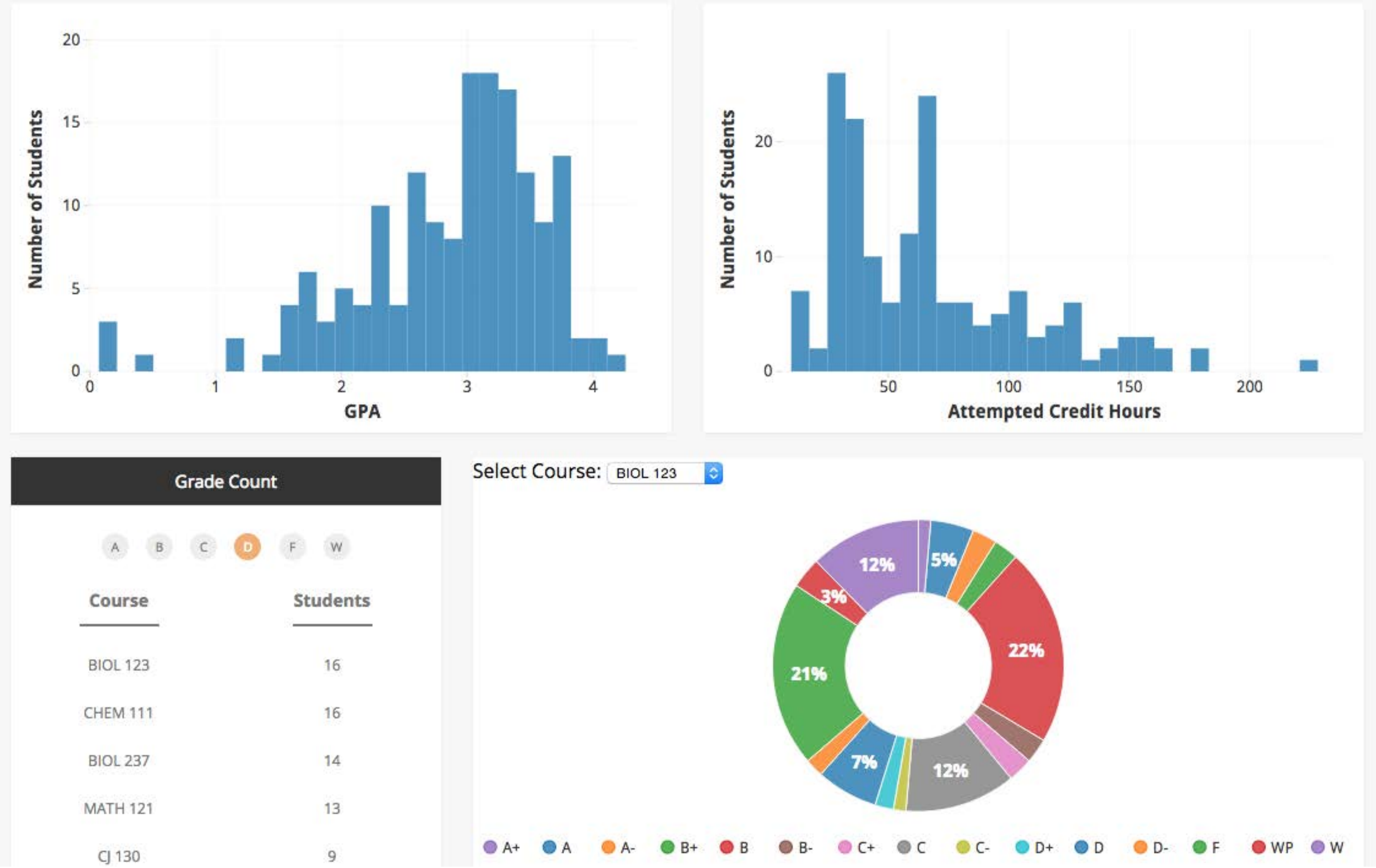


BS-AT

STUDENTS	AVERAGE GPA	AVERAGE CREDIT HOURS	AVERAGE COMPLETION
164	2.86	67	27%

Select Starting Term: Fall 2008

Cohort Analytics Dashboard



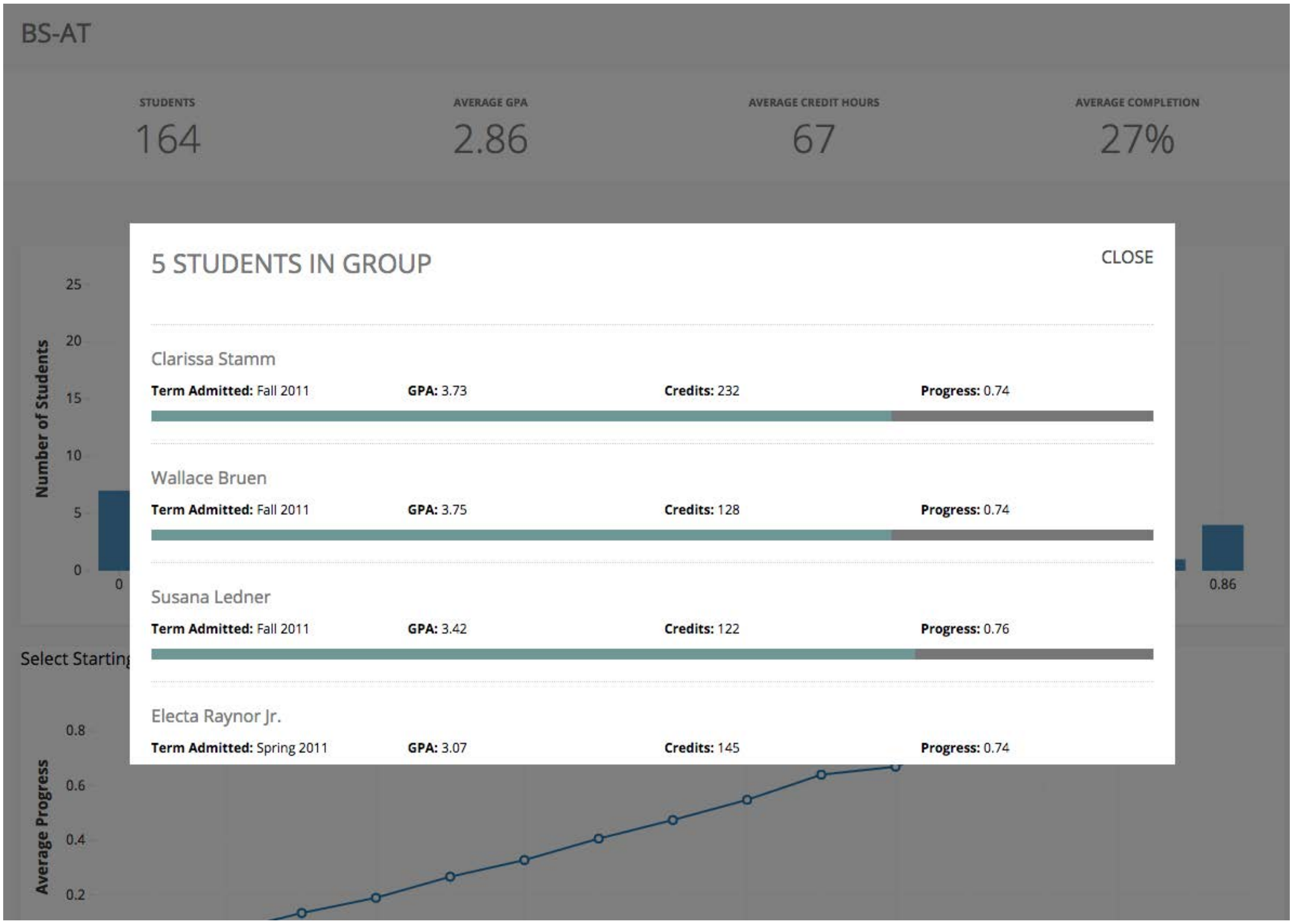
Grade Count

A B C D F W

Course	Students
BIOL 123	16
CHEM 111	16
BIOL 237	14
MATH 121	13
CJ 130	9

Select Course: BIOL 123

Cohort Analytics Dashboard (fake names)



BS-AT

STUDENTS 164

AVERAGE GPA 2.86

AVERAGE CREDIT HOURS 67

AVERAGE COMPLETION 27%

5 STUDENTS IN GROUP

CLOSE

Clarissa Stamm

Term Admitted: Fall 2011 **GPA:** 3.73 **Credits:** 232 **Progress:** 0.74

Wallace Bruen

Term Admitted: Fall 2011 **GPA:** 3.75 **Credits:** 128 **Progress:** 0.74

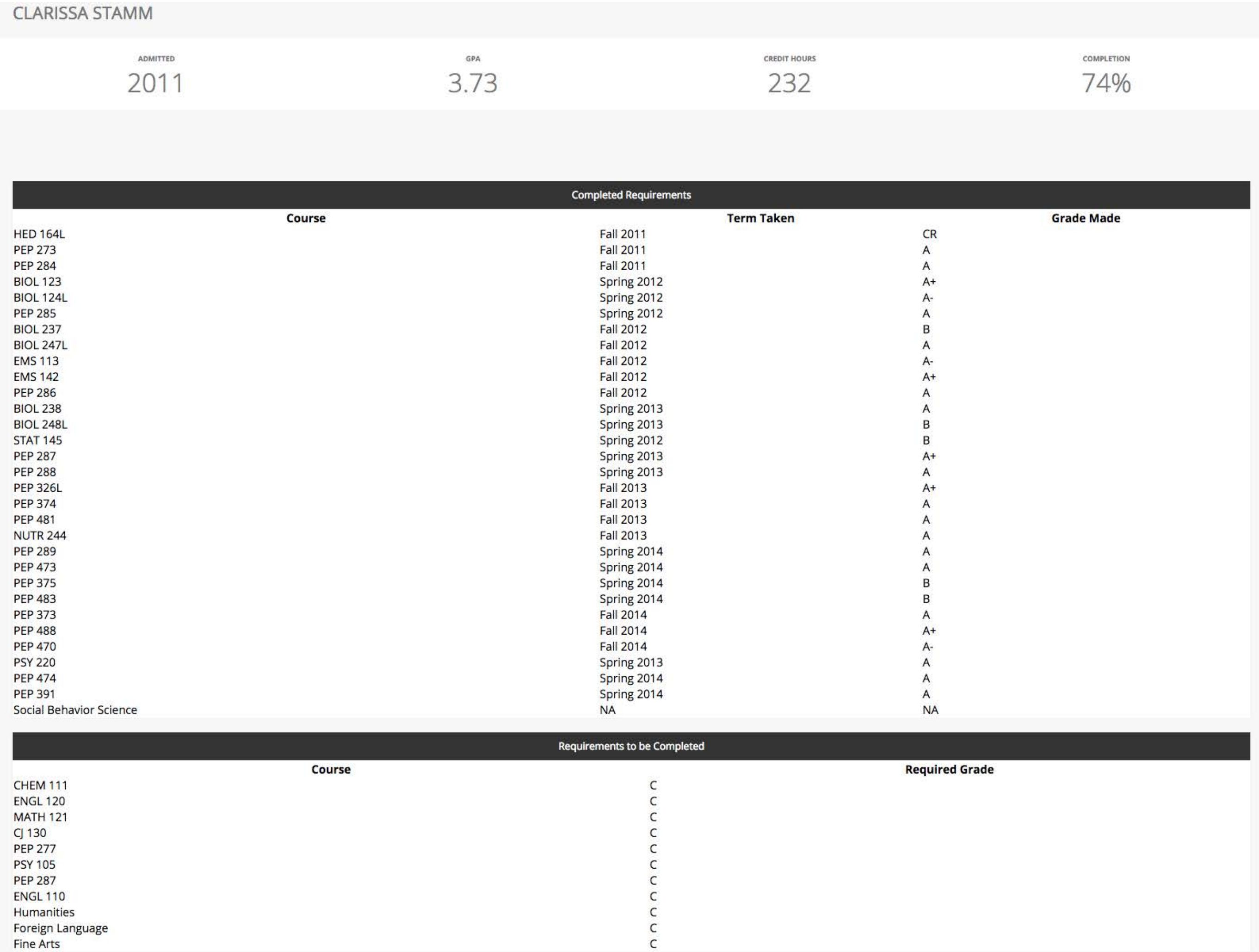
Susana Ledner

Term Admitted: Fall 2011 **GPA:** 3.42 **Credits:** 122 **Progress:** 0.76

Electa Raynor Jr.

Term Admitted: Spring 2011 **GPA:** 3.07 **Credits:** 145 **Progress:** 0.74

Cohort Analytics Dashboard (fake name)



CLARISSA STAMM

ADMITTED	GPA	CREDIT HOURS	COMPLETION
2011	3.73	232	74%

Completed Requirements

Course	Term Taken	Grade Made
HED 164L	Fall 2011	CR
PEP 273	Fall 2011	A
PEP 284	Fall 2011	A
BIOL 123	Spring 2012	A+
BIOL 124L	Spring 2012	A-
PEP 285	Spring 2012	A
BIOL 237	Fall 2012	B
BIOL 247L	Fall 2012	A
EMS 113	Fall 2012	A-
EMS 142	Fall 2012	A+
PEP 286	Fall 2012	A
BIOL 238	Spring 2013	A
BIOL 248L	Spring 2013	B
STAT 145	Spring 2012	B
PEP 287	Spring 2013	A+
PEP 288	Spring 2013	A
PEP 326L	Fall 2013	A+
PEP 374	Fall 2013	A
PEP 481	Fall 2013	A
NUTR 244	Fall 2013	A
PEP 289	Spring 2014	A
PEP 473	Spring 2014	A
PEP 375	Spring 2014	B
PEP 483	Spring 2014	B
PEP 373	Fall 2014	A
PEP 488	Fall 2014	A+
PEP 470	Fall 2014	A-
PSY 220	Spring 2013	A
PEP 474	Spring 2014	A
PEP 391	Spring 2014	A
Social Behavior Science	NA	NA

Requirements to be Completed

Course	Required Grade
CHEM 111	C
ENGL 120	C
MATH 121	C
CJ 130	C
PEP 277	C
PSY 105	C
PEP 287	C
ENGL 110	C
Humanities	C
Foreign Language	C
Fine Arts	C

Cohort Analytics – Components

The application involves the integration of a number of information systems:

- Student Data Mart – student progress data (FERPA applies).
- Degree Requirements and Degree Plans databases.
- Reasoning Engine – reasons over the aforementioned data stores.
- CAS Authentication and Authorization (whitelist until BAR roles are made available).
- Analytics and Interactive Dashboard Framework.

Note: the system involves moving student data to Amazon Web Services.

Security Profile & Controls[1]

UNM Data Classification:

1. Data owners: students
2. Data steward: Enrollment Management (Terry Babbitt), Custodians: AA Application Development Team and end users
3. Information system identification: see slide 10
4. Data categorization: student data – academic performance and other student attributes (e.g., ethnicity, gender, HS attended, etc.)
5. Privacy requirements: FERPA
6. Data Classification: see next slide
7. UNM Information Security Safeguards guidance: not available (see attached document and the following slides for our security control selection analysis)

[1] UNM, Office of the CIO. (2008), Information Technology Standards, [Online]. Available: http://cio.unm.edu/standards/docs/DataClassificationStandard041608r.pdf (visited on 06/25/2015).

UNM Data Classification Worksheet[2]

Data Owner: students

Information System (application name): Cohort Analytics

Specific Pieces of Data: student data – academic performance and other student attributes (e.g., ethnicity, gender, HS attended, etc.)

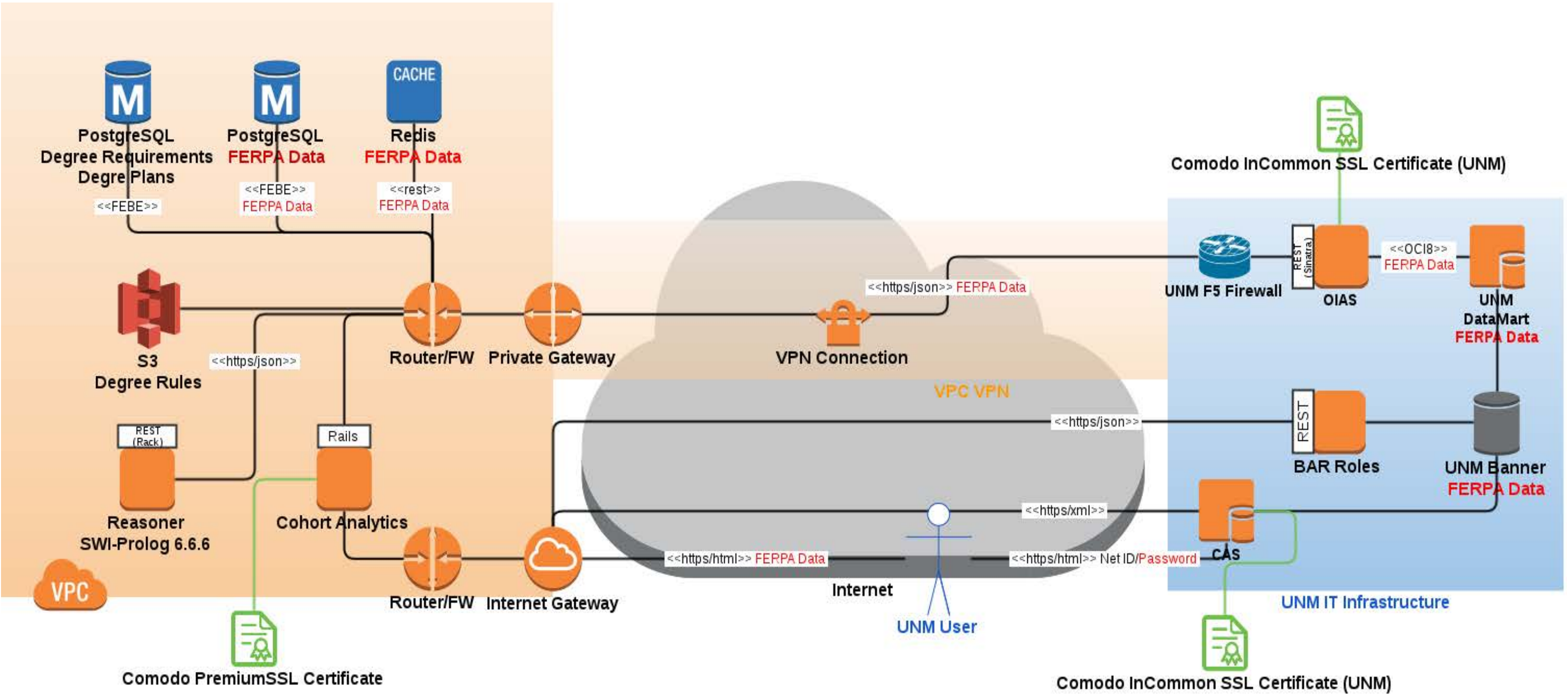
Data Classification: FIPS SC for cohort analytics data =
$\{$(Confidentiality, Moderate), (Integrity, Low), (Availability, Low)$\}$
Note: Moderate confidentiality requirements will be supported through controls including encryption at rest and in transit, via encryption standards described below.

Rationale: see attached document.

The combination of the above data classification, and the appropriate controls given this classification, seem to imply UNM's "C Class."

[2] UNM, Office of the CIO. (2008), Information Technology Standards, [Online]. Available: http://cio.unm.edu/standards/docs/DataClassificationStandard041608r.pdf (visited on 06/25/2015).

Cohort Analytics – Technical Components



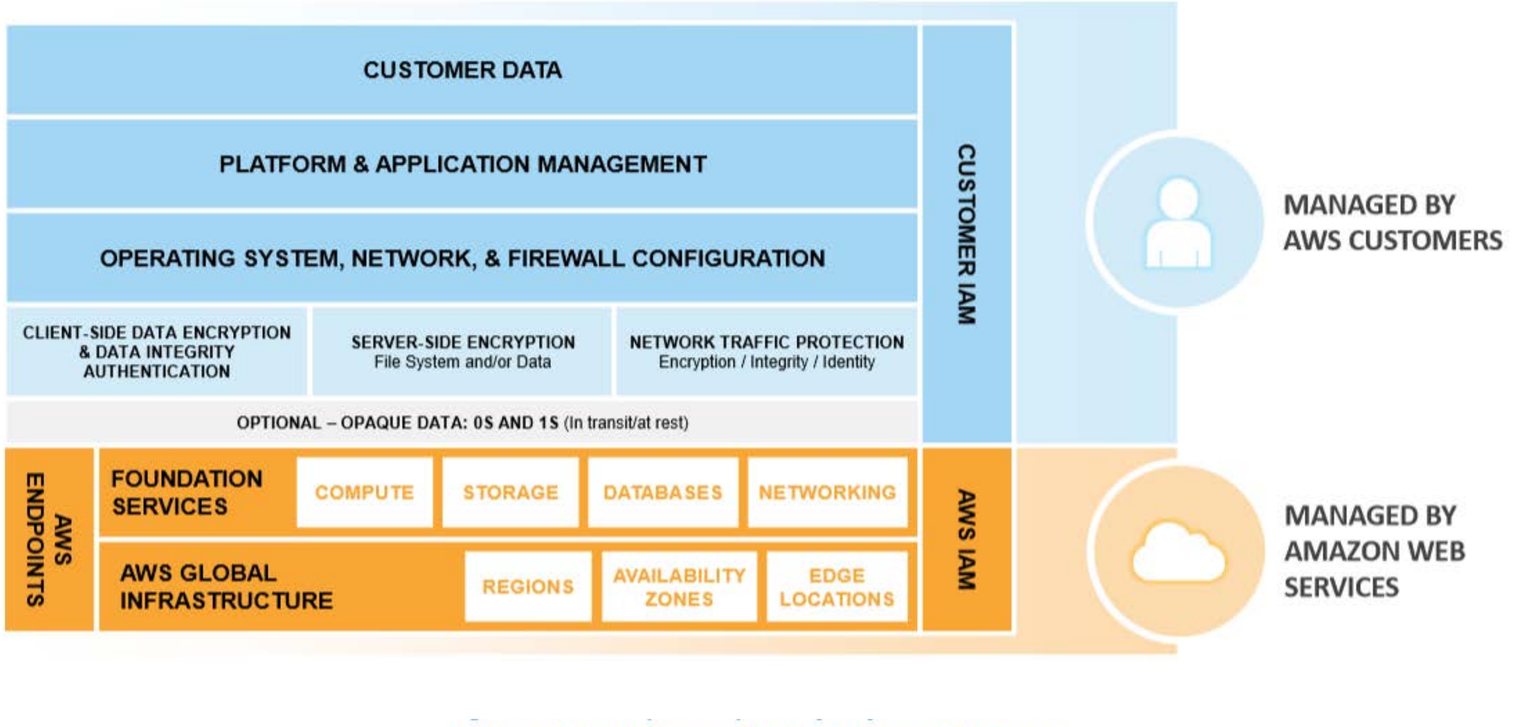
Cohort Analytics – Technical Components

Notes:

- Until the Banner Authorization Role (BAR) can be worked out, we will use UNM's CAS system for user authentication, and we will maintain a whitelist on the AWS side for user authorization. Whitelist entires must have UNM FERPA training, and if this is satisfied will include:
  - UNM President and Provost Office administrators.
  - Deans, Chairs and Program-level administrators
  - Academic Advisors
  - Others with a justified business need.

- For the required encrypted connections between these users and the Cohort Analytics system running on AWS, Academic Affairs will obtain a Premium SSL Certificate from Comodo.

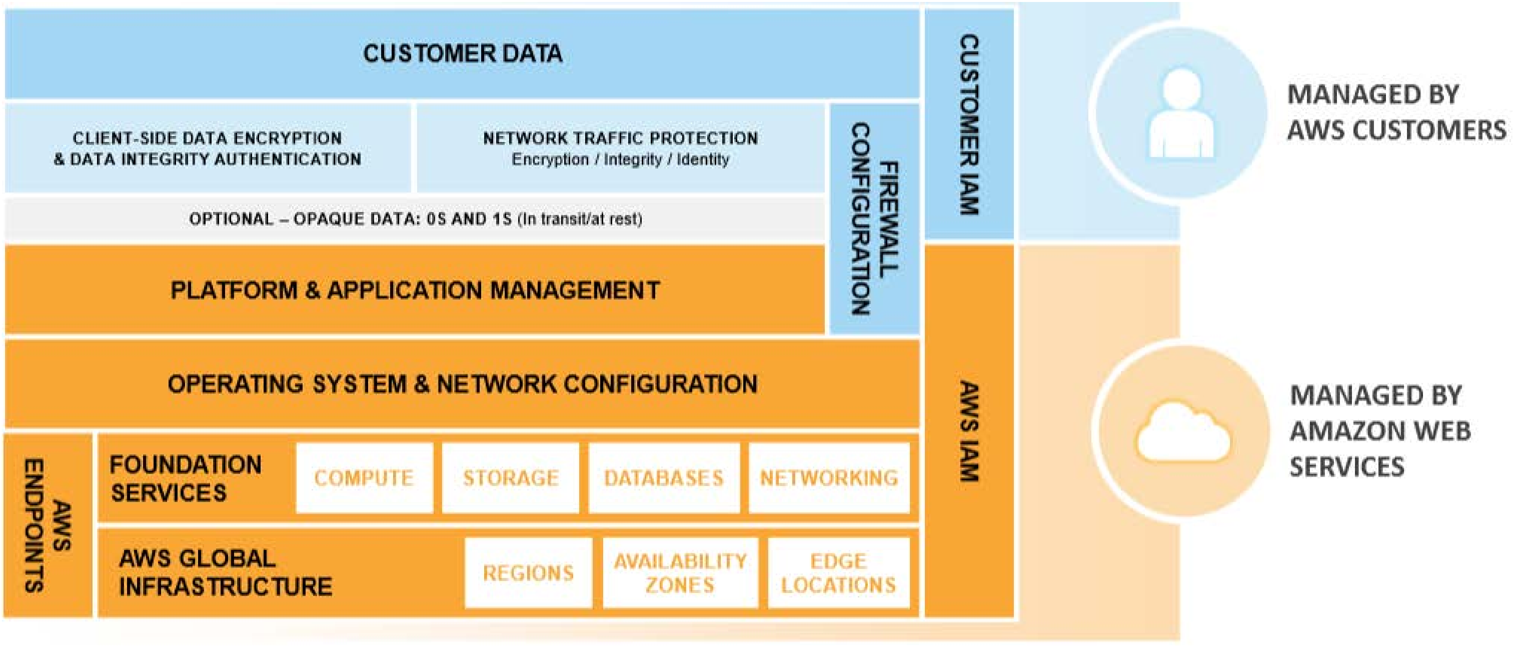
Responsibilities - Infrastructure


CUSTOMER DATA
PLATFORM & APPLICATION MANAGEMENT
OPERATING SYSTEM, NETWORK, & FIREWALL CONFIGURATION
CLIENT-SIDE DATA ENCRYPTION & DATA INTEGRITY AUTHENTICATION
SERVER-SIDE ENCRYPTION File System and/or Data
NETWORK TRAFFIC PROTECTION Encryption / Integrity / Identity
OPTIONAL – OPAQUE DATA: 0S AND 1S (In transit/at rest)
CUSTOMER IAM
MANAGED BY AWS CUSTOMERS
AWS ENDPOINTS
FOUNDATION SERVICES
COMPUTE
STORAGE
DATABASES
NETWORKING
AWS GLOBAL INFRASTRUCTURE
REGIONS
AVAILABILITY ZONES
EDGE LOCATIONS
AWS IAM
MANAGED BY AMAZON WEB SERVICES

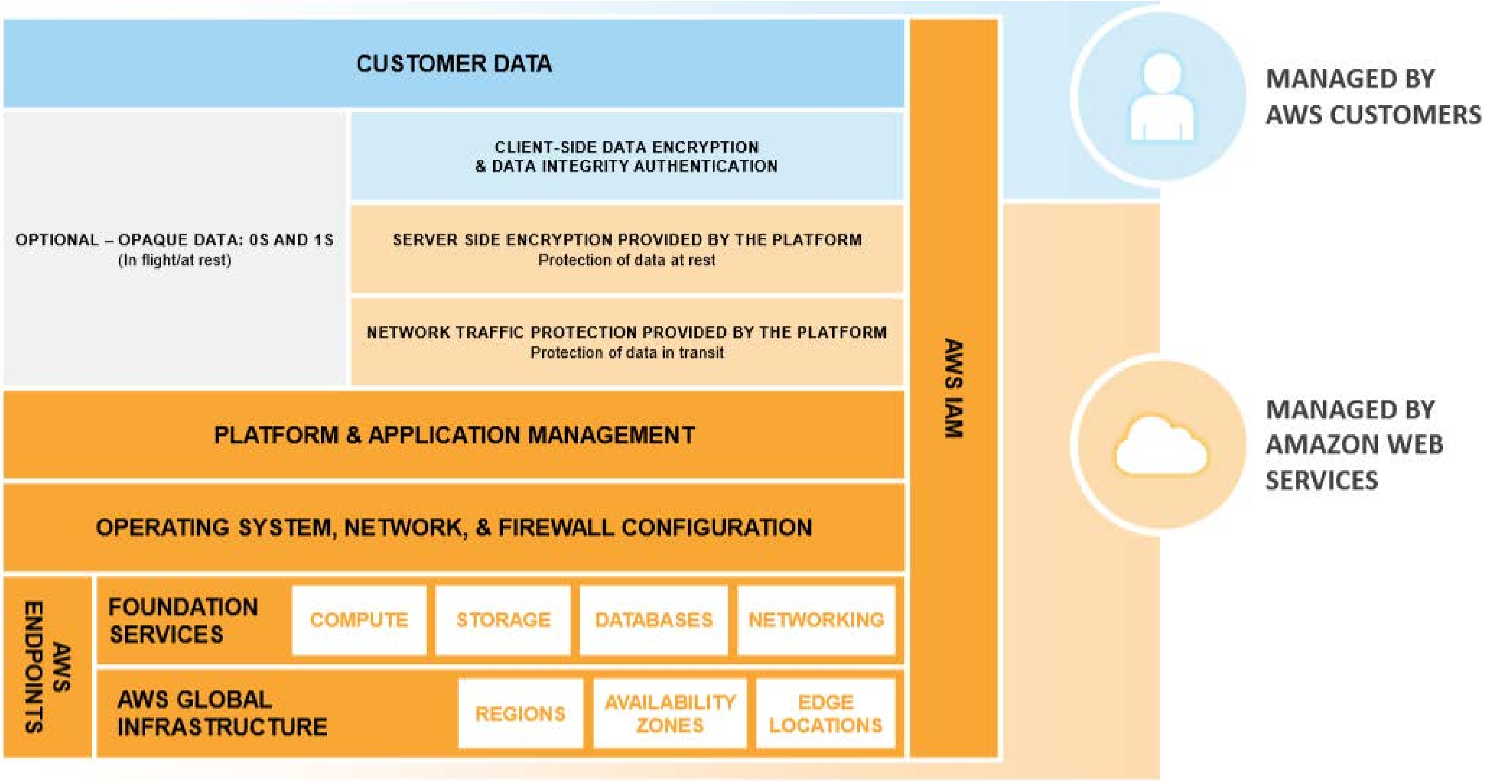
Extracted from content provided by Amazon Web Services[3].

[3] Amazon Web Services, "FERPA Compliance on AWS", , Amazon Web Services, Inc., Tech. Rep., May 2015.

Responsibilities - Containers


CUSTOMER DATA
CLIENT-SIDE DATA ENCRYPTION & DATA INTEGRITY AUTHENTICATION
NETWORK TRAFFIC PROTECTION
Encryption / Integrity / Identity
OPTIONAL – OPAQUE DATA: 0S AND 1S (In transit/at rest)
FIREWALL CONFIGURATION
CUSTOMER IAM
PLATFORM & APPLICATION MANAGEMENT
OPERATING SYSTEM & NETWORK CONFIGURATION
AWS ENDPOINTS
FOUNDATION SERVICES
COMPUTE
STORAGE
DATABASES
NETWORKING
AWS GLOBAL INFRASTRUCTURE
REGIONS
AVAILABILITY ZONES
EDGE LOCATIONS
AWS IAM
MANAGED BY AWS CUSTOMERS
MANAGED BY AMAZON WEB SERVICES

Responsibilities - Abstract Services



CUSTOMER DATA

OPTIONAL – OPAQUE DATA: 0S AND 1S
(In flight/at rest)

CLIENT-SIDE DATA ENCRYPTION
& DATA INTEGRITY AUTHENTICATION

SERVER SIDE ENCRYPTION PROVIDED BY THE PLATFORM
Protection of data at rest

NETWORK TRAFFIC PROTECTION PROVIDED BY THE PLATFORM
Protection of data in transit

PLATFORM & APPLICATION MANAGEMENT

OPERATING SYSTEM, NETWORK, & FIREWALL CONFIGURATION

AWS ENDPOINTS

FOUNDATION SERVICES: COMPUTE, STORAGE, DATABASES, NETWORKING

AWS GLOBAL INFRASTRUCTURE: REGIONS, AVAILABILITY ZONES, EDGE LOCATIONS

AWS IAM

MANAGED BY AWS CUSTOMERS

MANAGED BY AMAZON WEB SERVICES

Our Responsibilities - AWS

OS, Network, FW Configuration:

- Elastic Compute Cloud (EC2) VMs run SELinux/Redhat, UFW
- We don't manage UNM VMs or Firewalls
- We manage host firewalls and maintain the VPC

Platform & Application Management:

- Ruby/Rails — Runs on EC2; manually patched when required via the *bundler* and *gem* utilities
- Redis — Runs on Amazon RDS and Elasticache
- Prolog — Patched via operating system utilities

Student Data:

- Encrypted[4] at rest on AWS side and in motion (HTTPS or equivalent)

[4] K. Beer and R. Holland, "Encrypting Data at Rest", Amazon Web Services, Inc., Tech. Rep., Nov. 2014.

Identity Management, Accounts, and Keys

UNM and Local Identity Management

- Local accounts on EC2 and amazon are managed using UNM password policies (strong passwords with six month rotation)
- Application access is authorized via local whitelists and CAS authentication to UNM.
- We only allow administrative access via `sudo`.
- We use Amazon IAM as much as possible.

Amazon Identity Management

- Initially SSH access to running systems.
- Migration to multi-factor authentication (e.g. Google Authenticator).[5]
- Amazon key management for key storage.

[5] Amazon Web Services. (2015), Multi-factor Authentication, [Online]. Available: http://aws.amazon.com/iam/details/mfa/ (visited on 06/25/2015).

Security Monitoring

CloudWatch

- Syslog, performance, communication, etc.
- Early indicator that VMs have been compromised:
  - Higher usage
  - New VM creation
  - Very large instance creation (great for mining bitcoin, for example).

CloudTrail

- Compliance monitoring, user activity tracking, API access.
- Good for initial intrusion detection:
  - New API access
  - Excessive API access

We would like to collaborate with UNM IT on:

- Monitoring, management, continuity
- Security auditing

Bibliography

- UNM, Office of the CIO. (2008), Information Technology Standards, [Online]. Available: http://cio.unm.edu/standards/docs/DataClassificationStandard041608r.pdf (visited on 06/25/2015).

- Amazon Web Services, “FERPA Compliance on AWS”, Amazon Web Services, Inc., Tech. Rep., May 2015.

- K. Beer and R. Holland, “Encrypting Data at Rest”, Amazon Web Services, Inc., Tech. Rep., Nov. 2014.

- Amazon Web Services. (2015), Multi-factor Authentication, [Online]. Available: http://aws.amazon.com/iam/details/mfa/ (visited on 06/25/2015).