

HP Capstone

Big Data Analytics

Prepared by:
Nic Desilets,
James Stallkamp,
Nathaniel Whitlock

Introduction

The following slides will be an overview of the individual experiences of each of the members of our senior design group. They be organized by individual in the order of; Nic Desilets, James Stallkamp, and Nathaniel Whitlock.

Nic's Experience

■ Weeks 0-2

- Introduction to the class and what it entails.
- Introduced to some of this year's capstone projects and their sponsors.
- Read over the complete list of available projects.
- Was primarily interested in web related projects.

■ Week 3

- Got on board with HP Big Data Analytics.
- Initial meeting went over project details and expectations.
- Discussed a little bit about Oracle Database 12c.
- Core problem: 72-core, 3TB server not being fully utilized.
- Established a regular weekly meeting time.
- Began problem statement document.

■ Week 4

- Established weekly meeting time with TA.
- Tried to get Oracle Database 12c installed prior to the next HP meeting.
- Although the install was successful, the database itself was not correctly configured.
- Spent a fair amount of time installing and setting up Oracle 12c with help from our clients.
- Began discussing the dynamic performance views in the database.
- Goal for next meeting was to begin designing the toolkit.

■ Week 5

- Spent some time this week revising our problem statement document.
- Removed some technical jargon from it.
- Had a better idea of what our target audience would be.
- Started working on the requirements document.
- Reinstalled Oracle 12c and got it configured properly.
- Got set up with HP's test server and database.

■ Week 6

- Struggled a bit with the requirements document.
- The requirements for the document did not quite fit our project.
- Was pretty busy during this week due to midterms.
- Setup a Slack channel for our group to help improve communications.
- Email is fine but not quite suited for instant messaging.
- Started reading Explain the Explain Plan document.

■ Week 7

- Read over the Explain the Explain Plan document.
- Got a little more insight into the internals of the database.
- We were given a script to generate a set of data that is similar to the real data set.
- Talked a little bit more about the dynamic performance views.
- Discussed different IDEs, toolkits, and partitioning types.
- Goal before next meeting was to continue familiarizing ourselves with Oracle 12c (V\$, X\$).

■ Week 8

- A lot of time was spent on the tech. review early on in the week.
- Ensure that assignment requirements were appropriately met, finalizing.
- This week's meeting with our client was one of the most productive.
- Wasn't tied down with worrying about another paper.
- Was able to start formulating queries to query different V\$ tables.
- Constructed sequence of queries to obtain database times for each parallel executor.

■ Week 9

- Struggled with getting a start on the design document.
- While other papers were difficult, this one seemed the most difficult.
- Seemed way overkill for our toolkit that we're designing.
- We continued to make more progress on obtaining metrics information during our weekly meeting.
- Main goal at this point was to continue working on the design document.

■ Week 10

- This week was primarily working on and completing the design document.
- Definitely one of the more difficult papers, but we got it done.
- Got introduced to tracing during our weekly meeting.
- Verbose logging of what exactly the database is doing.
- Helps determine how the optimizer is coming up with its execution plans.
- Can be used in conjunction with dynamic performance views.
- Main goal was to compile our blog reports into our progress report document and presentation.

James' Experience

■ Weeks 0-3

- During the first three weeks we were introduced to the structure of our capstone project.
- We heard project proposals from clients and voted on the ones we preferred to work with.
- I was assigned to work with HP's big data project.
- In week three we had the initial meeting with our client, which went well.
- We spent the meeting getting a basic understanding of the project. Which is that HP has a very powerful server that seems to be under utilizing its resources.
- Personally I have concerns because I have very little experience with databases but this presents a good opportunity to learn.

■ Week 4

- This week the rest of my group began to setup the software packages on their personal laptops
- I don't own a laptop so I took notes and repeated the process on my desktop.
- I discussed with our mentors how important it would be to have a laptop, and I will try to find one to use or borrow.
- Began revision on Problem Statement based on feedback
- Discussed method of retrieving metrics from dynamic performance views
- Missed TA meeting. I completely forgot about and missed the TA meeting. Moving forward I will be better about attending meetings.

■ Week 5

- I confirmed I have a laptop to borrow, and I will be driving to Portland this weekend to pick it up from my parents.
- I have successfully set up SQL developer on my personal desktop and will do so on the laptop next week.

■ Week 6

- Spent most of the time working on the requirements document.
- As a group we had issues coordinating for this document and largely worked on it in individual segments.
- I have the weakest skills with Latch and Nathaniel and Nic are doing most of the formatting. This is something I will need to work on.
- For next weeks meeting we all will be reading Explain the Explain plan which will explain how Oracle databases execute and optimize queries.

■ Week 7

- For the meeting this week we all read Explain the Explain plan, then discussed it.
- The document does into detail on how a sql query is executed as well as optimized.
- We learned about the various internal tables on Oracle SQL databases, and which contain useful performance data.
- I am still having issues with my personal machine the VM that I set up keeps crashing or corrupting the image when I try to resume a session.
- Still having issues with my virtual box image not loading when I try to resume, will plan on talking to Andy next meeting to fix this.

■ Week 8

- Spent time this week working on my section of the tech review.
- My section to research is which IDE which we will be using.
- Our mentors have introduced us to SQL developer which is what we will probably use but I will research SQL plus as well as TOAD as alternatives

■ Week 9

- Spent most of our time this week working on the Design document.
- This document proved to be very difficult to write because our project is researched based.
- Missed Weekly HP meeting because I was sick.
- After having continued trouble with my virtual machine I have decided that when I am at HP I will use the test server they have provided and at home I will use the server on my Desktop.
- For some reason I can not seem to maintain the integrity of the image in the virtual machine on my laptop work

■ Week 10

- The weekly meeting was focused on understanding how tracing works.
- We started the time line for our progress report but did not accomplish much.

Nathaniel's Experience

■ Weeks 0-3

■ Prior to Group Assignment

- Learning about structure of our course
- Read about available projects
- Spoke with client about getting involved in their project
- Voted on projects of interest and assigned to group

■ Post Group Assignment

- Initial meeting with client
- Learned about the scope of our problem
- Learned about some Oracle specific database concepts
- Scheduled remained of meeting times for the term
- Began working on Problem Statement document

■ Week 4

- Getting software installed on person computers
 - 1 Issues with package dependencies
 - 2 Network limitations due to HP firewall
 - 3 Configuration issues with Oracle environment and listener
- Began revision on Problem Statement based on feedback
- Discussed method of retrieving metrics from dynamic performance views

■ Week 5

- Swapped from personal computer use to work laptop
- Finished installing required software on work laptop
- Dealt with network issues from within virtual machine
- Began working on the Requirements Document

■ Week 6

- Majority of time spent on finishing the Requirements Document
- Hard to find time to work in concert on the Requirement Document
- Began executing queries against dynamic performance views on personal computer
- Ran into major issues forcing formatting on L^AT_EX class file
 - T.A. helped point me in the right direction
 - **Lesson learned:** Don't try to force formatting changes on a class file
- Began reading Oracle white paper

■ Week 7

- Discussed the Explain Plan white paper
- **Issues encountered:**
 - Forgot password to personal 12c environment
- Provided PL/SQL script to generate data
- Emphasis on learning more about dynamic performance views
- Began the partition of work for the Technology Review, my areas of focus included:
 - Partitioning
 - Parallelism

■ Week 8

- The majority of efforts were spent on finalizing Technology Review
- Unsure about initial feedback from document
- Began initial efforts on outlining the Design Document
- No real progress on toolkit development
- Meeting with clients becoming increasingly focused on development

■ Week 9

- Further pushing into pre-toolkit development process
- Majority of efforts placed on finishing portion of design document
- News from client about Oracle convention

■ Week 10

- Began focusing on the topic of tracing
- Experimented with Oracle native tools for trace file analysis
- Began focusing on creating a time line of progress over the term

Conclusion

- Majority of the term was spent on writing documentation
- We have hope that this will guide us in development
- Issues with software installation
- Benefit of our experienced mentors
- Plan moving forward