## Formula Optimization

A guide on identifying and optimizing Anaplan formulas


## Mark Warren

Manager - Operational Excellence Group

## /Anaplan

$G 6$
3 easy steps to follow if you want to improve your model performance"

## Golden Rules

Calculate once and
refer many times

- Avoid repetition!
- Why calculate the same thing in multiple places when we can avoid it
- Use System modules learn how to DISCO

Calculate at lowest cell count possible

- Why calculate more than we need to?
- Calculate over just the dimension the formula applies to
- Break up formulas to calculate at the right cell count for each part

Test as you go!

- Spotting poor performance early makes it easier to address
- Test different ways of achieving the same goal
- Test and prove optimizations


## Calculate once and refer many times

## Repetition

- Where it occurs
- Generally find it within the same module
- Between similar modules and functional areas
- Simple common elements like list items and dates
- Conditional checks are a common area
- Why is it bad?
- Because it is a calculation we can avoid
- They add up, increasing overall calculation time
- Can lead to more recalculation


## Calculate once and refer many times

Common example of repeated functions

- ITEM(X)
- NAME(ITEM(X))
- CODE(ITEM(X))
- PARENT(ITEM(X))
- START()
- END()
- ITEM(Time) = TIME.'Current Period'
- CURRENTPERIODSTART() < START()
- Boolean checks in IF statements


## Calculate once and refer many times



## Calculate once and refer many times

Use a Time Management Module


Cell Count $=105$


VS


|  | rormura |
| :---: | :---: |
| Call Data Forecast |  |
| Override Call Flag | IF Time Management.Future Period? 1 HEN 0 ELSE IF Call Ce |
| Override Avg Handle Time Flag | IF SIARI() <=CURRENIPERIUUSTART() THEN 0 ELSE IF C |
| Distance from Actuals | 'Current Period'.Actuals Through - START() + 1 |
| - |  |
| Total Actual Calls | IF Time Management.Future Period? HEN 0 ELSE Import Ca |
| 1_Max_Calls | \|FY_Calls <> OTHENTELSE0 |
| X_All_Calls | CUMULATE(1) |
| X_Max_Calls | IF Y_Calls <> 0 THEN CUMULATE('1_Max_Calls') ELSE 0 |
| X_Total_Calls | X_Max_Calls |
| X_Sqr_Calls | POWER(X_Max_Calls, 2) |
| Y_Calls | Total Actual Calls |
| XY_Calls | X_Max_Calls ${ }^{\text {P Y_Calls }}$ |
| m_Calls | (X_Max_Calls[SELECT: TIME.All Periods] * XY_Calls[SELECT |
| c_Calls | (Y_Calls[SELECT: TIME.All Periods] - m_Calls * X _Total_Calls] |
| E(Y) - Straight Line_Calls | m_Calls * X_All_Calls + c_Calls |
| E(Y) - Moving Average_Calls | IF START() <= CURRENTPERIODSTART() THEN Y_Calls ELs |
| E(Y) - Exponential Smoothing_Calls_Formula | IF CUMULATE( 1 ) $=1$ THEN Y_Calls ELSE Trend Line Method |
| E(Y) - Exponential Smoothing_Calls | IF START() <= CURRENTPERIODSTART() +1 AND END() $>1$ |
| Total Inbound | IF START() <= CURRENTPERIODSTART() THEN Total Actual |
| -- |  |
| Total Actual Avg Handle Time | IF Time Management.Future Period? ${ }^{1}$ HEN 0 ELSE Import Ca |
| 1_Max_Avg Handle Time | IFY_Avg Hanaie nime es ( IRENTEESE 0 |
| X_All_Avg Handle Time | CUMULATE(1) |
| X_Max_Avg Handle Time | IF Y_Avg Handle Time <> 0 THEN CUMULATE(1_Max_Avg H |

CPX $7 /$ /.naplan

## Calculate at lowest cell count possible

A common example is adding text together to form a unique code

In this case we are adding a Company code to Product code with an underscore separator

|  | Formula | Parent | Is Summary | Format | Applies To | Time Scale |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| REP01 Data Summary |  |  |  |  | Company, Product | Month |
| Code | \| $\operatorname{CODE}$ (ITEM(Company)) \& "_" \& CODE(ITEM(Product)) |  |  | Text | - | Month |

There are two text additions here (two \&'s) done at a combined cell count of 128,895,624

## Calculate at lowest cell count possible

- Product has 3531 items
- Company has 1014 items

So we do one text addition at the lowest cell count possible with no timescale


The main formula is now

|  | Formula | Parent | Is Summary | Format | Applies To | Time Scale |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| REP01 Data Summary |  |  |  |  | Company, Product | Month |
| Code | \|'SYS05 Company Details'. Code to Use \& 'SYS06 Product Details'.Code |  |  | Text | - | Month |

CPX $7 /$ /.naplan

## Calculate at lowest cell count possible

We can take it a step further here...
The formula does not apply to time

| Jan 14 | Feb 14 | Mar 14 | Apr 14 | May 14 | Jun 14 |
| :--- | :---: | :---: | :---: | :---: | :---: |
| $840 \_1274$ | $840 \_1274$ | $840 \_1274$ | $840 \_1274$ | $840 \_1274$ | $840 \_1274$ |
| $840 \_5337$ | $840 \_$- 3337 | $840 \_5337$ | $840 \_5337$ | $840 \_5337$ | $840 \_5337$ |
| $840 \_6990$ | $840 \_6990$ | $840 \_6990$ | $840 \_6990$ | $840 \_6990$ | $840 \_6990$ |
| $840 \_4184$ | $840 \_4184$ | $840 \_4184$ | $840 \_4184$ | $840 \_4184$ | $840 \_4184$ |
| $840 \_1501$ | $840 \_1501$ | $840 \_1501$ | $840 \_1501$ | $840 \_1501$ | $840 \_1501$ |
| $840 \_1580$ | $840 \_1580$ | $840 \_1580$ | $840 \_1580$ | $840 \_1580$ | $840 \_1580$ |
| $840 \_8034$ | $840 \_8034$ | $840 \_8034$ | $840 \_8034$ | $840 \_8034$ | $840 \_8034$ |
| $840 \_1900$ | $840 \_1900$ | $840 \_1900$ | $840 \_1900$ | $840 \_1900$ | $840 \_1900$ |

Doing the calculation without a timescale reduces cell count to $3,580,434$

|  | Formula | Parent | Is Summary | Format | Applies To | Time Scale |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| REP01 Data Summary |  |  |  |  | Company, Product | Month |
| Code Optimised | 罧 'SYS05 Company Details'. Code to Use \& 'SYS06 Product Details'. Code |  |  | Text | - | Not Applicable |
| Code | Code Optimised |  |  | Text | - | Month |

CPX $7 /$ /.naplan

## Calculate at lowest cell count possible

- The first example calculation took 5.68 sec
- (two text additions at $128,895,624$ cells)
- The last example took $0.52 \mathbf{~ s e c}$
- (one text addition at $3,580,434$ cells)
- $91 \%$ reduction in duration


## Demo

1. Identifying large formulas
2. How to edit large formulas
3. Splitting up formulas to reduce repetition and complexity
4. Testing the optimizations
5. Performance analysis

## CPX $7 /$ /.naplan

## Demo

## - Taking a complex formula (small part shown here 3,504 characters of 273,668 )

(IF ISNOTBLANK(LEFT(String, FIND(""", String, 1) - 1)) THEN LEFT(String, FIND("_", String, 1) - 1) ELSE String) \& (IF LENGTH(MID(String, FIND(" ", String, 1) + 1, IF FIND("", String, LENGTH(LEFT(String, FIND("", String, 1) + 1)) - -LENGTH(LEFT (String, FIND("n, String, 1)) < 0 THEN 1000 ELS



 LENGTH(MID(String, FIND("", String, 1) + 1, IF FIND("n, String, LENGTH(LEFT(String, FIND("-", String, 1) + 1))) - LENGTH(LEFT(String, FIND(" ", String, 1)) < 0 THEN 1000 ELSE FIND(" ", String, LENḠTH(LEFT(String, FIND("_", String, 1) + 1 ) )) - LENGTH(LEFT(String, FIND("_", String, 1)))))= 0 THEN BLANK ELSE " " \& MID(String, FIND ("", String, 1) + 1, IF FIND("n, String, LENGTH(LEFT(String, FIND("_", String, 1) + 1))) LENGTH(LEFT(String, FIND(" ", String, 1))) < 0 THEN 1000 ELSE FIND(" ", String, LENGTH(LEFT(String, FIND(" ", String, 1) + 1))) LENGTH(LEFT(String, FIND("-", String, 1))) - 1)) + 2, IF FIND(" "", String, [ENGTH ((IF ISNOTBLANK(LEFT(String, F'IND("n", String, 1) - 1)) THEN LEFT(String, FIND(" "", String, 1) - 1) ELSE String) \& IF LENGTH(MID(String, FIND("", String, 1) + 1, IF FIND("", String, LENGTH(LEFT(String,
 1) + 1) ) - LENGTH(LEFT(String, FIND(" ", String, 1) ))) $=0$ THEN BLANK ELSE " " \& MID(String, FIND("" ", String, 1) + 1, IF FIND("_", String, LENGTH(LEFT(String, FIND(" $\quad$ ", String, 1) + 1))) - LENGTH(LEFT(String, FIND("_", String, 1))) < 0 THEN 1000 ELSE FIND (" $\_$", String,


 LENGTH(LEFT(String, FIND("-", String, 1) + 1)) - LENGTH(LEFT(String, FIND("-", String, 1)) < 0 THEN 1000 ELSE FIND (" " String, String, 1) + 1, IF FIND(" ", String, LENGTH(LEFT(String, FIND (" ", String, 1) + 1 1)) - LENGTH(LEFT(String, FIND(" ", String, 1)) < 0 THEN 1000 ELSSE FIND (" " , String, LENGTH(LEFT (String, FIND("", String, 1) + 1)) - LENGTH(LEFT (String, FIND ("", String, 1))) - 1) ) + 2) <= 0 THEN 1000000 ELSE FIND (" LENGTH(MID(String, FIND(" "", String, 1) + 1, IF FIND(""", String, ZENGTH(LEFT(String, FIND("", String, 1) + + 1) )) - LENGTH(LEFT(String, FIND(" " ", String, 1) ) < 0 THEN 1000 ELSE FIND ("n", String, LENGTH (LEFT(String, FIND("_", String, 1) + 1)) ) - LENGTH(LEFT(String, FIND("_", String, 1))))) ='0 THEN BLANK ELSE " " \& MID(String, FIND("_", String, 1) + 1, IF FIND("_-, Str...

## Demo

- Finding and highlighting repeated parts in the text editor
(IF ISNOTBLANK (LEFT(String, FIND("_", String, 1) - 1)) THEN LEFT(String, FIND("_", String, 1) - 1) ELSE String) \& (IF LE $\operatorname{LENGTH}(\operatorname{LEFT}(S t r i n g, ~ F I N D(" . ", ~ S t r i n g, ~ 1) ~+~ 1))) ~-~ L E N G T H(L E F T(S t r i n g, ~ F I N D(" . ", ~ S t r i n g, ~ 1))) ~<~ 0 ~ T H E N ~ 1000 ~ E L S E ~ F I N D(" ~ " ~ ' ~$ $\operatorname{LENGTH}\left(\operatorname{LEFT}\left(S t r i n g, \operatorname{FIND}\left("{ }^{-}\right.\right.\right.$", String, 1))) < 0 THEN 1000 ELSE FIND("_", String, LENGTH(LEFT(String, FIND("_", String, 1) FIND("_", String, LENGTH((IF ISNOTBLANK(LEFT(String, FIND("_", String, 1) - 1)) THEN LEFT(String, FIND("_", String, 1) FIND("_", String, LENGTH(LEFT(String, FIND("_", String, 1) + 1))) - LENGTH(LEFT(String, FIND("_", String, 1))) < 0 THEN $\operatorname{LENGTH}\left(\operatorname{LEFT}\left(S t r i n g, ~ F I N D\left(" \_", S t r i n g, ~ 1\right)\right)\right)$ < 0 THEN 1000 ELSE FIND("_", String, LENGTH(LEFT(String, FIND("_", String, 1) THEN 1000000 ELSE FIND("_", String, LENGTH ((IF ISNOTBLANK (LEFT(String, FIND("_", String, 1) - 1)) THEN LEFT(String, FINL String, 1) + 1, IF FIND("_", String, LENGTH(LEFT(String, FIND("_", String, 1) + 1))) - LENGTH(LEFT(String, FIND("_", Str String, 1) + 1) )) - LENGTH(LEFT(String, FIND("_", String, 1))) < 0 THEN 1000 ELSE FIND("_", String, LENGTH(LEFT(String,
1)) ) - 1) ) + 2) ) = BLANK THEN BLANK ELSE " " \& MID(String, LENGTH((IF ISNOTBLANK(LEFT(String, FIND("_", String, 1) - 1)) MID (String, FIND("_", String, 1) + 1, IF FIND("_", String, LENGTH(LEFT(String, FIND("_", String, 1) + 1))) - LENGTH(LEFT
- Using the Find or Count function od the editor to look for part with most repetitions
- Replacing that part with a reference to a new line item named $A$

| 4356 results found for 'FIND("_", String, 1)' | Finding with Option | se Insensitive | .* | Aa |  | 亓 | $\times$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| FIND("_", String, 1) | 4356 found | Find | Find All |  |  |  |  |
| A |  | Replace | Replace All |  |  |  |  |

## Demo

- Repeating until all repetition is removed...

| Optimised Formula | C \& D \& \& H \& K \& M | $\square$ | Text |
| :---: | :---: | :---: | :---: |
| Substitute Function | SUBSTITUTE(String, "_-", " ") | $\square$ | Text |
| A | FIND("_", String, 1) | $\square$ | Number |
| B | FIND("_", String, LENGTH(LEFT(String, A + 1))) - LENGTH(LEFT(String, A )) | $\square$ | Number |
| C | IF ISNOTBLANK(LEFT(String, A - 1)) THEN LEFT(String, A - 1) ELSE String |  | Text |
| D | IF LENGTH(MID(String, A +1 , IF B < 0 THEN 1000 ELSE B)) $=0$ THEN BLANK E |  | Text |
| E | LENGTH(C \& D \& MID(String, A +1, IF B $<0$ THEN 1000 ELSE B - 1 ) +2 | $\square$ | Number |
| F | MID(String, E, IF FIND("_", String, E) - E <= 0 THEN 1000000 ELSE FIND("_", St\| |  | Text |
| G | MID(String, A + 1, IF B < 0 THEN 1000 ELSE B - 1) |  | Text |
| H | IF F = BLANK THEN BLANK ELSE " " \& F |  | Text |
| 1 | $\operatorname{LENGTH}(\mathrm{C}$ \& D \& G \& H) +2 | $\square$ | Number |
| J |  |  | Text |
| K | IF $\mathrm{J}=$ BLANK THEN BLANK ELSE " " \& J |  | Text |
| L | MID(String, LENGTH(C \& D \& G \& H \& K) $+2,10000$ ) |  | Text |
| M | IF L = BLANK THEN BLANK ELSE " \% \& L |  | Text |

## Demo

- Calculation analysis - comparison of calculation times before and after
- Test by comparing formulas and how long did it take before and after

Calculations by Line Item

|  | Module | Line Item ₹ | Total calculation time (ms) | \% of total calculation time |
| :---: | :---: | :---: | :---: | :---: |
| 1 | Complexity | Original Formula | 1,170, 317.36 | 96.93 |
| 2 | Complexity | B | 5,580.09 | 0.46 |
| 3 | Complexity | F | 5,160.04 | 0.43 |
| 4 | Complexity | String | 5,079.23 | 0.42 |
| 5 | Complexity | A | 4,738.79 | 0.39 |
| 6 | Complexity | G | 3,483.94 | 0.29 |
| 7 | Complexity | D | 3,392.21 | 0.28 |
| 8 | Complexity | Substitute Function | 3,289.74 | 0.27 |
| 9 | Complexity | C | 2,866.46 | 0.24 |
| 10 | Complexity | L | 1,414.55 | 0.12 |
| 11 | Complexity | H | 616.94 | 0.05 |
| 12 | Complexity | Optimised Formula | 462.33 | 0.04 |

## Key takeaways

- Large performance gains can be made
- Using a text editor can aid in finding repetition
- The gain in model size can be worth it for the improved performance
- Difficult to write an optimized formula - testing and optimising after build is key


## Complexity

## Complexity

- Where it occurs
- Multiple functions in one line item
- Nested IF statements
- Copy and pasted or Excel formula generation
- Why is it bad?
- Large calculations
- Frequent recalculation
- Difficult logic to understand \& maintain


## Complexity

User makes a cell change


Products.Purchase Price[SUM: Region] + Marketing.Total Costs[SUM: Region] / Sales.Total Sales[LOOKUP: Region] * (Currency Modifier.USD / 1000)

In this scenario all four functions will recalculate if any reference changes

Calculation Time: $30,320 \mathrm{~ms}+29,780 \mathrm{~ms}+21,320 \mathrm{~ms}+4560 \mathrm{~ms}=85,980 \mathrm{~ms}$

## Complexity

Functions split out to their own line items
A = Products.Purchase Price[SUM: Region]
B = Marketing.Total Costs[SUM: Region]
C = Sales.Total Sales[LOOKUP: Region]
D = (Currency Modifier.USD / 1000)
If $B$, Marketing Costs, is the only reference affected by the changes then only this calculates A+B/C*D

Calculation Time: $1 \mathrm{~ms}+29,780 \mathrm{~ms}+1 \mathrm{~ms}+1 \mathrm{~ms}=29,783 \mathrm{~ms}$
Giving us a $65 \%$ reduction in calculation time!

## Complexity

- Export the line items from settings

- Create a column in the export to check formula length using: =LEN(B3)
- Anything longer than 500 to 1000 characters is worth looking at
- Combine with cell count to work out which line items to focus on first


CPX $7 /$ /.naplan

## Text Concatenation

## Modification Performance



## Memory Use



## CPX $7 /$ /naplan

## Finding text concatenation



## Excel formula for \& count: <br> =LEN(B3)-LEN(SUBSTITUTE(B3,"\&",""))

- Combine with formula length and cell count to decide which line items to focus on first


## Demo

1. Identifying text concatenation formulas
2. Splitting formulas to calculate at lowest cell count possible
3. Testing the optimization
4. Performance analysis

## CPX $7 /$ /.naplan

## Demo

- Breaking down formula (at each applies to)
- "CODE1:" \& CODE(ITEM('500')) \& " " \& "CODE2:" \& CODE(ITEM('200')) \& "_" \& "DATE:" \& NAME(ITEM(Time))
$\approx$ Original Formula "CODE1:" \& CODE(ITEM( $\left.\left.500^{\prime}\right)\right)$ \& "_" \& "CODE2:" \& CODE(ITEM(200')) \& "_" \& "DATE:" \& NAME(ITEM(Time))

|  |  | Formula | Parent | Is Summary | Format | Applies To | Time : |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Text Concatenation |  |  |  |  |  | 500, 200 | Month |
| Original Formula |  | -CODE1:" \& CODE(ITEM(500)) \& "-" \& "CODE2:" \& CODE(ITEM(200) ) "-" \& |  |  | Text | - | Month |
| Optimised Formula |  | 'Text 1' \& 'Text 2' |  |  | Text | - | Month |
| Text 1 | 田 | 'CODE1:" \& CODE(ITEM(500')) |  |  | Text | 500 | Not Appl |
| Text 2 | 回 | "_CODE2:" \& CODE(ITEM( $200{ }^{\prime}$ ) \& \& 'Text 3' |  | - | Text | 200 | Month |
| Text 3 |  | "_DATE:" \& NAME(ITEM(Time)) |  |  | Text |  | Month |

- New line items with lowest cell counts possible

| Cell Count |
| ---: |
| $2,402,912$ |
| $1,200,000$ |
| $1,200,000$ |
| 500 |
| 2,400 |
| 12 |

## Demo

- Test optimisation - compare original and optimised
- Filter out any items that don't match - should be none



## Demo

- Performance analysis

Calculations by Line Item

|  | Module $=$ | Line ltem $=$ | Total calculation time $(\mathrm{ms})=$ | $\%$ of total calculation time $=$ |
| :--- | :--- | :--- | ---: | ---: |
| 1 | Text Concatenation | Original Formula | $22,253.84$ | 90.29 |
| 2 | Text Concatenation | Optimised Formula | $2,170.20$ | 8.81 |
| 3 | Text Concatenation | Text 3 | 120.56 | 0.49 |
| 4 | Text Concatenation | Text 1 | 56.31 | 0.23 |
| 5 | Text Concatenation | Text 2 | 45.60 | 0.19 |

CPX $7 /$ /.naplan

## Key takeaways

- Text formula performance is poor
- Combine as many text parts as possible
- Split apart by each 'applies to’
- Use System modules where possible

CPX $7 /$ /.naplan

## Session takeaway

- Following the golden rules will lead to improvements
- Use Excel to analyse blueprints
- Learn to use the text editor to break down formulas

Key factors in reducing formula complexity

- Calculate once and refer many times
- Calculate at lowest cell count possible


## Don't forget

- Test as you go!


## Further reading

- https://community.anaplan.com/t5/Best-Practices/Reduce-Calculations-for-Better-Performance/ta-p/33667
- https://community.anaplan.com/t5/Best-Practices/Formula-Optimization-in-Anaplan/tap/41663
- https://community.anaplan.com/t5/Best-Practices/Formula-Structure-for-Performance/tap/33177
- PLANS - https://community.anaplan.com/t5/Best-Practices/PLANS-This-Is-How-We-Model/ta-p/33530
- DISCO - https://community.anaplan.com/t5/Best-Practices/Best-Practices-for-Module-Design/ta-p/35993
- Planual - https://community.anaplan.com/t5/Best-Practices/The-Planual/ta-p/49773
- My text editor https://atom.io/


## Exercise 1

- An empty model with the complex formula
- Try to recreate the demo by finding repeated elements in a text editor and populating the model
- Remember to add in a line item to compare the results

|  | Formula | Parent | Is Summary | Format |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Complexity |  |  |  |  | 1000, 500 |
| String | "ABC_DEF_GHI_JKL_MNO" |  | $\square$ | Text | - |
| Original Formula | (IF ISNOTBLANK(LEFT(String, FIND('_-", String, 1) - 1)) THEN LEFT(String, FINT |  | $\square$ | Text | - |
| ---- |  |  | $\square$ | No Data | - |
| Optimised Formula |  |  | $\square$ | Text | - |
| A |  |  | $\square$ | Number | - |
| B |  |  | $\square$ | Number | - |
| C |  |  |  | Text | - |
| D |  |  |  | Text | - |
| E |  |  | $\square$ | Number | - |
| F |  |  |  | Text | - |
| G |  |  | $\square$ | Text | - |
| H |  |  |  | Text | - |
| 1 |  |  | $\square$ | Number | - |
| J |  |  |  | Text | - |
| K |  |  |  | Text | - |
| L |  |  |  | Text | - |
| M |  |  |  | Text | - |

CPX $7 /$ /.naplan

## Exercise 2

- Add the line items needed to separate out the text concatenation
- Remember to change dimensions to get lowest cell counts
- Use a line item to compare the two formulas

|  | Formula | Parent | Is Summary | Format |  | Applies To |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Text Concatenation |  |  |  |  | 500, 200 |  |
| Original Formula |  |  | - | Text | - |  |
| Optimised Formula |  |  |  | Text | - |  |

## Thank you!

CONNECTED
PLANNING XPERIENCE
1 A.naplan

