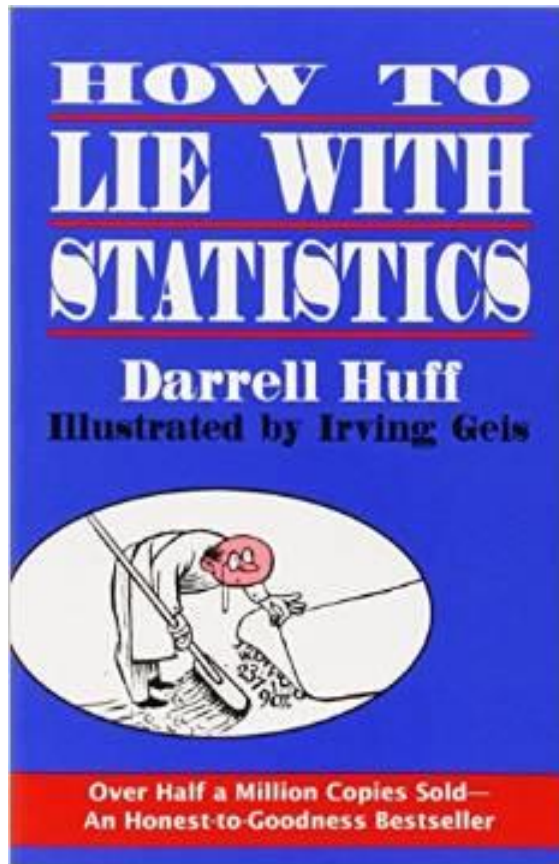


How NOT To Lie With Graphics and Tables – Don't Convey the Wrong Message!.

SHARE session 17370

Greg Caliri
BMC Software, Inc.
Lexington MA, USA

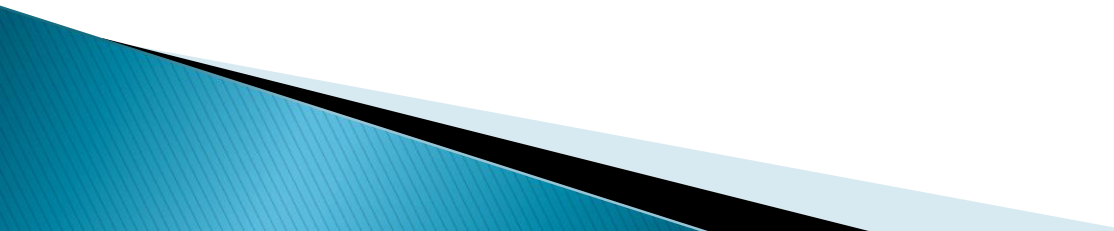
Reference....



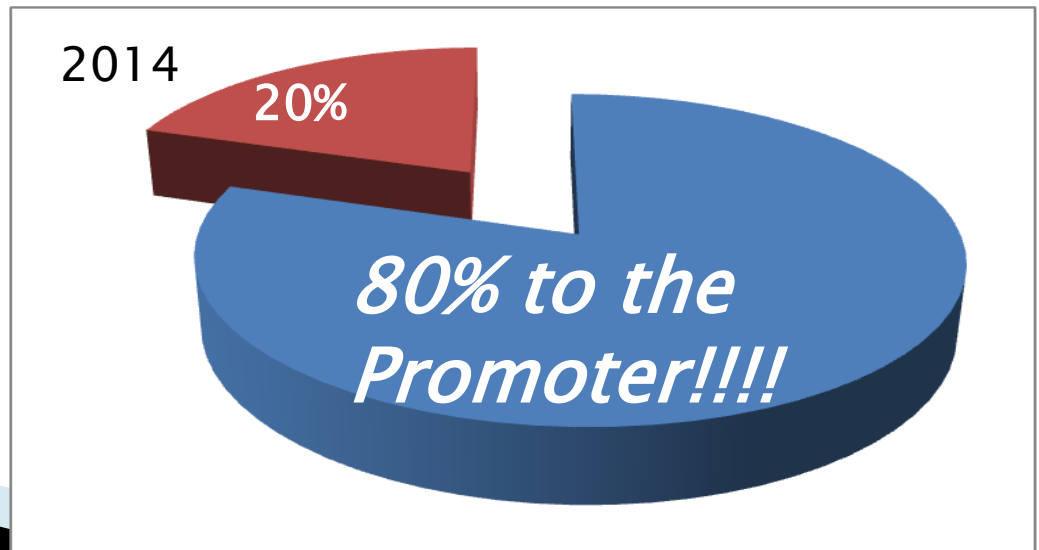
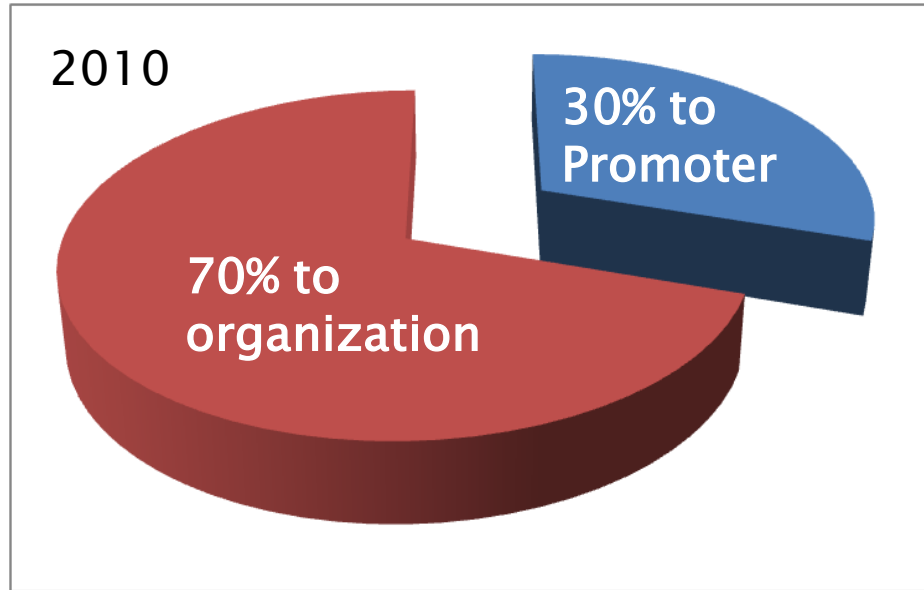
Darrell Huff (1913–2001)

“There is terror in numbers..”

Disclaimer and objective...

- ▶ This is INTRO stuff...
 - ▶ Sure, you can lie, but you don't WANT to lie
 - ▶ And you don't want to inadvertently "lie"
 - ▶ An example of a deliberate lie with statistics...
- 

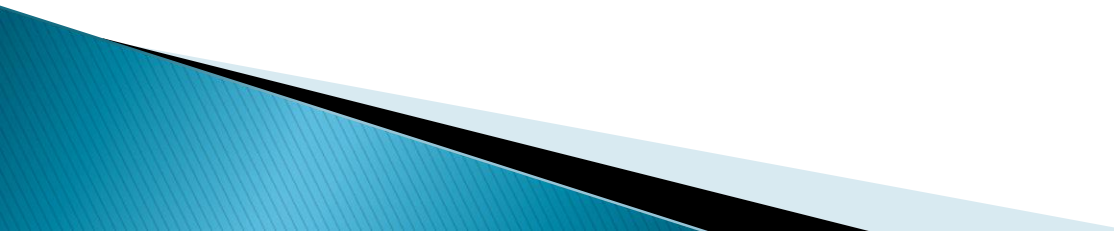
Misrepresentation and Embellishment



Now what if ...

- ▶ 2010 revenue was \$2 million...
 - 80% to organization = \$1.6 million
- ▶ 2014 revenue was \$40 million...???
 - 20% to organization = \$8 million
 - **A “Fair share” argument might be made – politically, but also there is an organizational uptick of over 300 percent**
 - **Also note that I embellished the 80 percent pie slice with an exaggerated description**

Mistakes happen ...

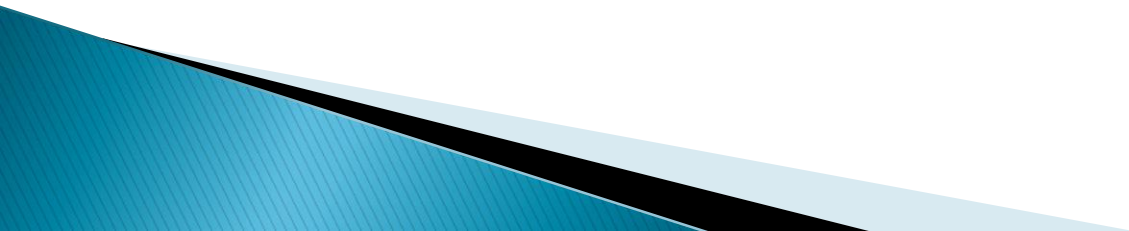
- ▶ Executive suite reporting – you can't pull a trick like that
 - ▶ You AVOID doing it accidentally
 - ▶ Easy to trip up with z/OS charts, numbers, etc.
 - ▶ Other examples follow
- 

In performance and capacity reporting ---

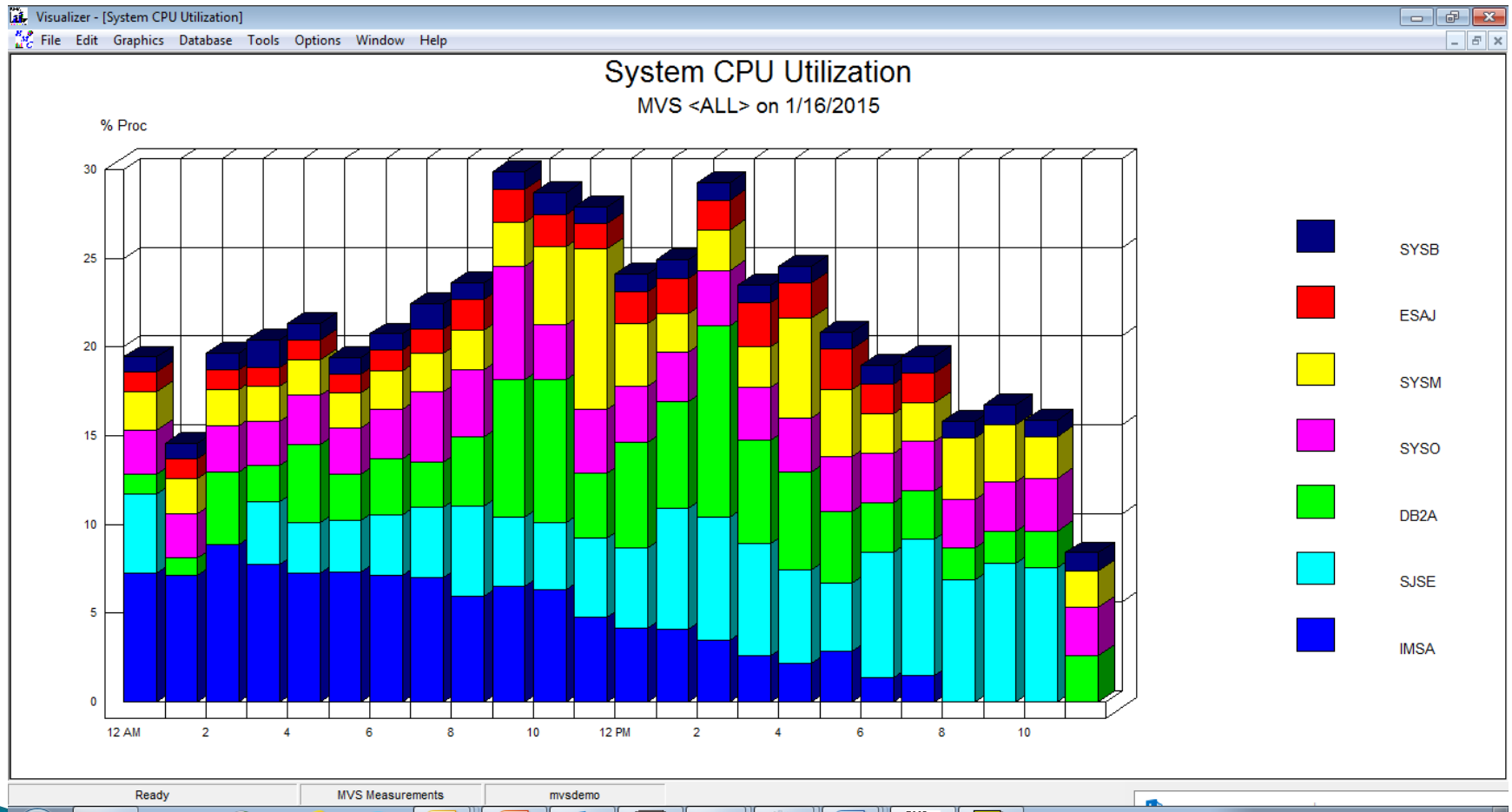
- ▶ Generally, don't use pie charts
- ▶ They show shares, not necessarily numbers for decision-making or trending

The “Gee Whiz” graph

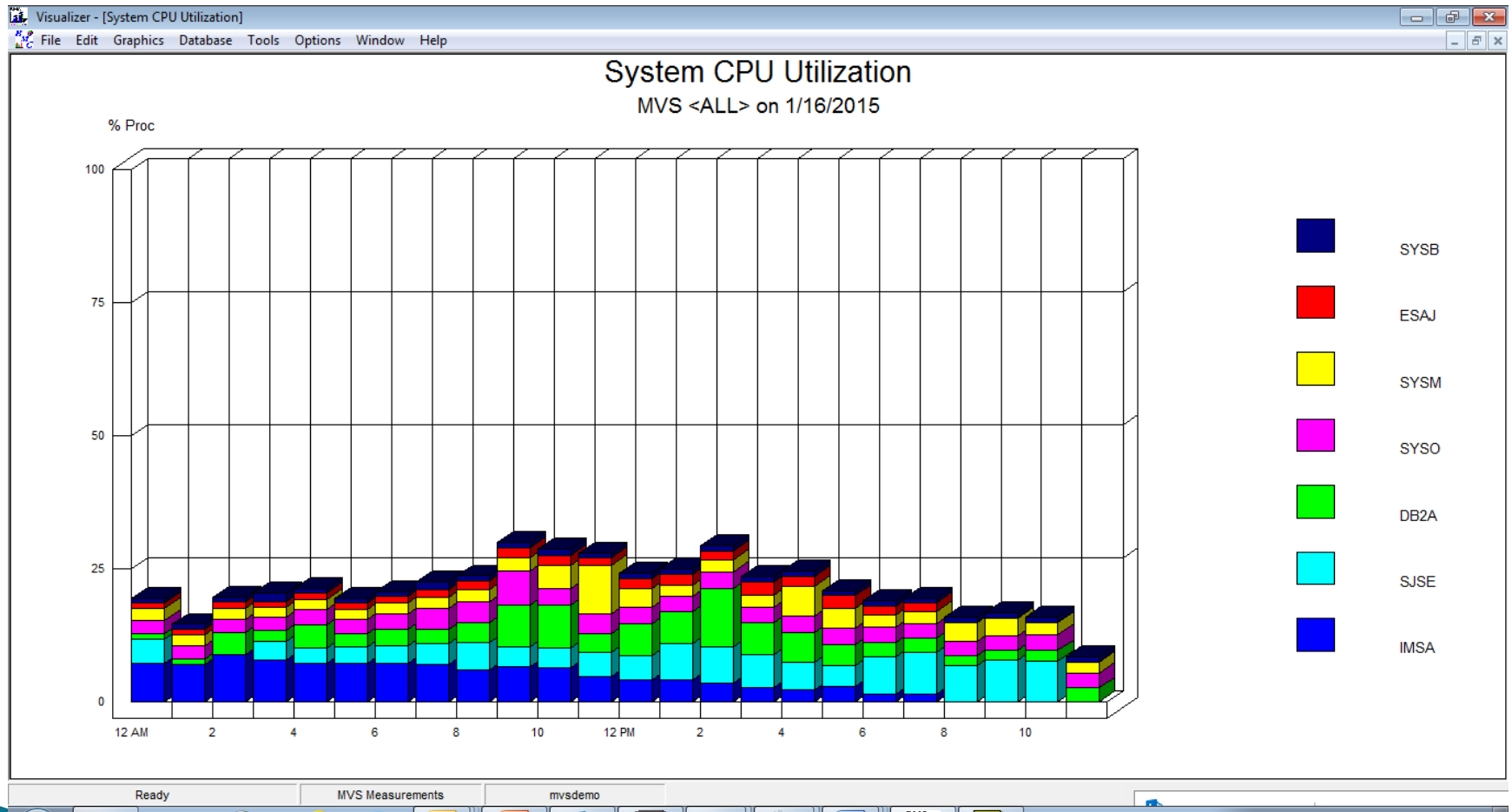
- ▶ Most common method of confusion
- ▶ Easiest way to lie with a graph
- ▶ But you don't to mislead anyone ...



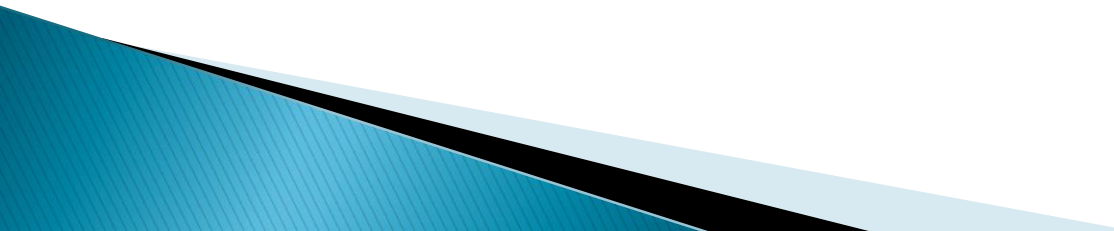
What is wrong here?



What is really happening...



The differences?

- ▶ Graph y-axis set to make a “good fit” – standard deviation
 - ▶ But – for consistency, “fix” the Y-axis to capacity (MIPs, % of processor, MSUs)
 - ▶ Same data, different presentation – and the presentation can be everything
- 

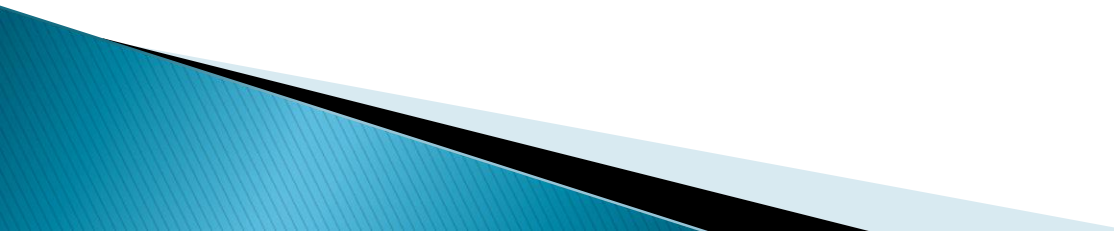
Highly critical

- ▶ Users do comparisons
- ▶ Scaling variations can confuse the reader
- ▶ It will generate questions

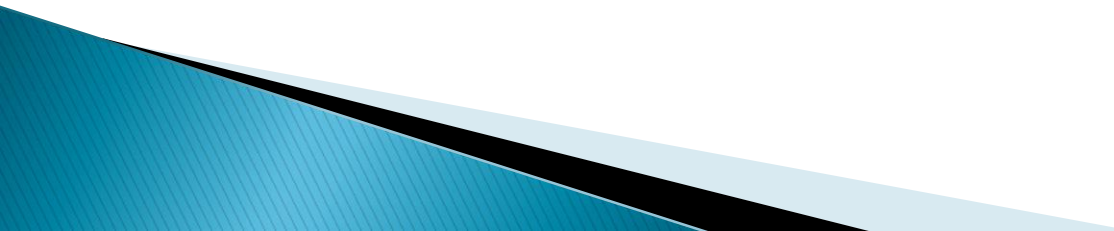
AVOID CONFUSION!

IF YOU GET ASKED FOR EXPLANATIONS OF THE
REPORT ITSELF – YOU'VE CONFUSED THINGS!

Let's take a breath here

- ▶ Mainframe Information from SMF and RMF and other sources.....it's an ocean in z/OS
 - ▶ 50+ years, metrics for everything – most of which you'll never need ...
 - ▶ Unlike other platforms – YOU must report on what's **relevant**....
- 

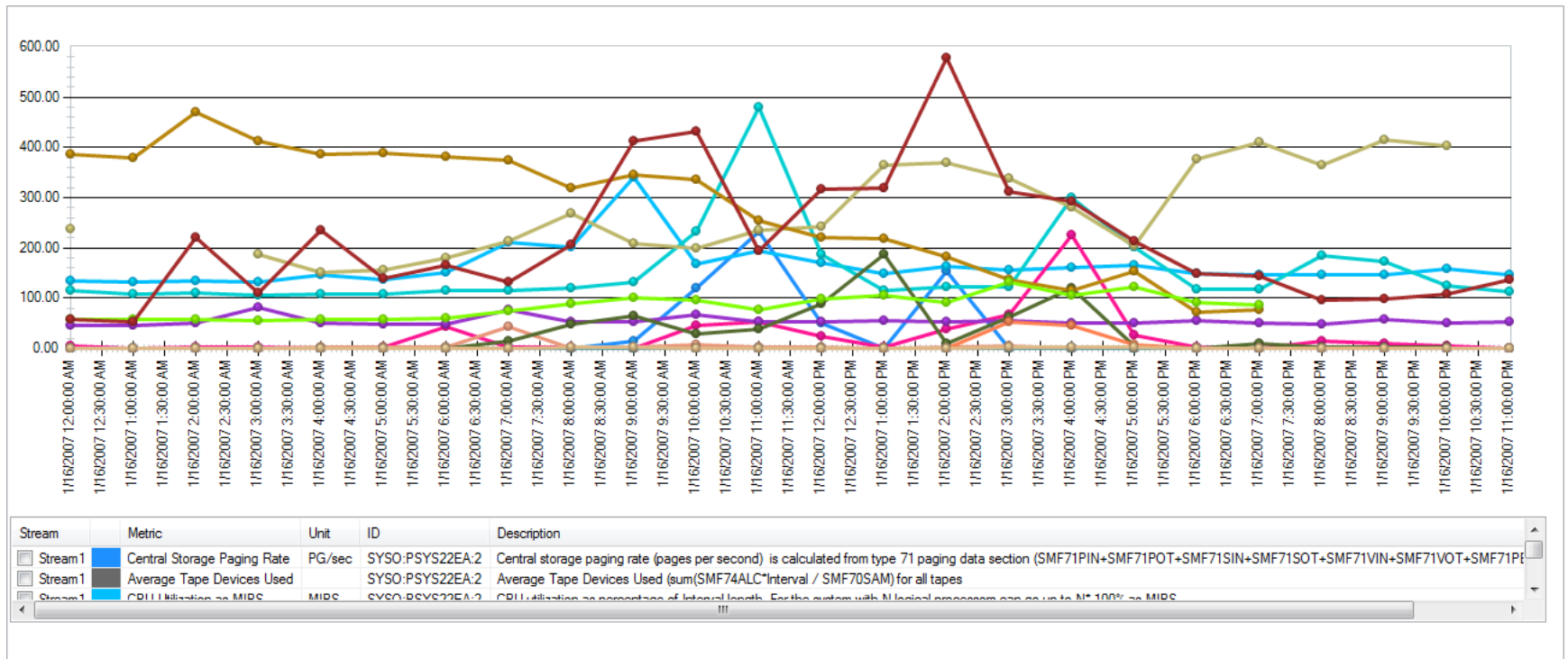
... before going back

- ▶ Metrics chosen and reported
 - Units of Work, organized along lines of business
 - Units the audience can relate to
 - So it's easy to report “irrelevant stuff”
 - Back on topic...
- 

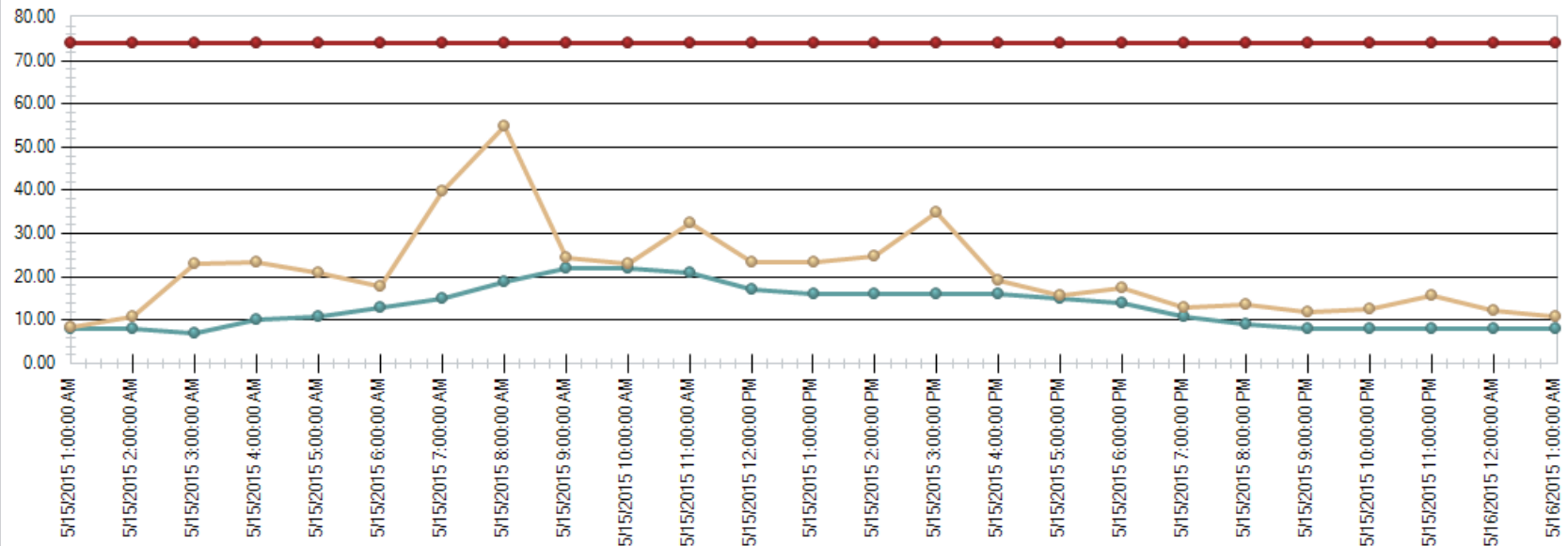
Including too much information on a graph

- ▶ It ISN'T better to report too much
- ▶ Get to the point
- ▶ Examples Here comes a bad one

Too much and irrelevant..attempt at correlation

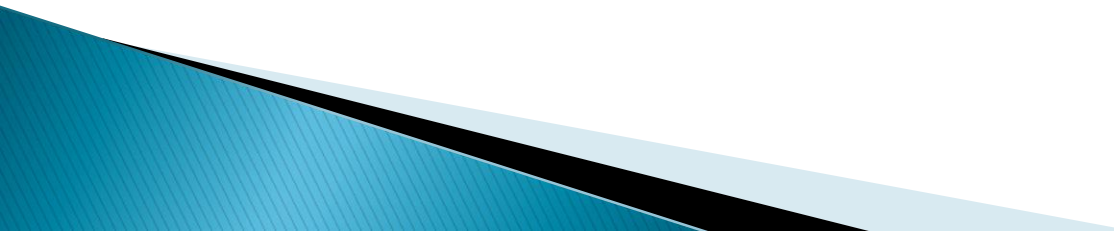


Correlative, right amount to report

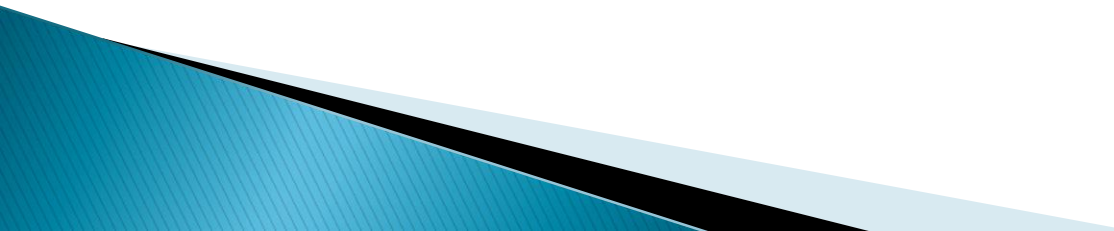


Stream	Metric	Unit	ID	Description
☐ Stream1	Host 4hours Average MSU	MSU/hour	SJSB:PSYSF037:2	Four hour rolling average MSU is compared with LPAR Defined Capacity when LPAR Capping is enabled. It is also used to calculate Vari
☐ Stream1	CPU Utilization as MSU/hour	MSU/hour	SJSB:PSYSF037:2	CPU utilization as percentage of Interval length. For the system with N logical processors can go up to N* 100% as MSU/hour
☐ Stream1	Host Defined Capacity	MSU/hour	SJSB:PSYSF037:2	Host OS image Defined capacity in Million Service units Used per hour (SMF70WLA)

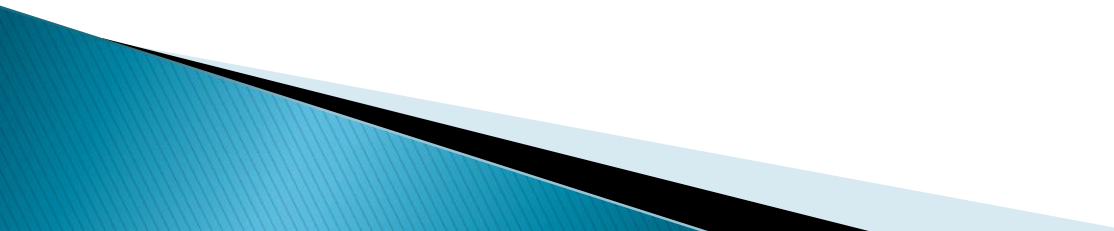
Correlations...

- ▶ OK to do them if the metrics are related to each other
 - ▶ NOT OK – to associate non-related metrics
 - ▶ Sometimes I/O, DB2 Calls, Paging, CPU spikes are related
 - ▶ Non-related activities confuse – and MISLEAD
- 

Correlations

- ▶ Separate coincidences from related events
 - ▶ Avoid the “Stork Correlation”
 - ▶ AFC/NFC Super Bowl correlation vs. Dow Jones Industrial Average
- 

“Harum Scarum” – data out of context

- ▶ Analysis of anything with pure numbers can be very dangerous – tell the wrong story
 - ▶ Large population = relatively small percentage of population (Nielsen ratings)
 - ▶ Small population = relatively large percentage sampling (New Hampshire Presidential primary)
- 

Oh my! We missed 100 percent...! Every hour had something wrong!

BMC Performance Predictor for Mainframes

File View Tools Utilities Help

Tools List

- Performance Explorer
- Report Explorer
- Graphlet Explorer
- Graph Explorer
- Scenario Explorer

Report Explorer

Friday, May 15, 2015 at 01:00 [i]
Daily Report Set (Daily Rep
Daily Health Check Su

Health Check - Daily Health Check Summary Report

Report Description: Daily Baseline performance health check summary, including availability of performance data collection, and summary of performance health check for the running workloads

Baseline: 2015/05/15; 01:00 - 2015/05/16; 01:00 - Duration: 60

Creation Time: 08/08/2015, 13:57

Creation Location: By the user account GCALIRI from ADPROD/W-GCALIRI-77 (172.28.181.213)

Overall Performance Status Summary

Report Description: Daily Baseline data availability and overall performance status chart

Overall Performance Status of the Baseline data of date: 2015/05/15; 01:00 - 2015/05/16; 01:00

Interval	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	00:00	01:00
Health Check Indicator	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	

Back to Baseline Performance Data Status Chart

No Health Check reports are available for the last 5 days.

Click [here](#) for other Daily Health Check reports.

General Status

Status	Source	Data Source	Timestamp
Informational	MVP System Service	z/OS	01:56:24 PM on August 08, 2015
Interactive Evaluation Finished.			
Warning	Report Publisher	z/OS	01:56:24 PM on August 08, 2015
Report Publishing - Web portal publishing is disabled. No reports have been published			
Informational	Application Evaluator	z/OS	01:56:24 PM on August 08, 2015

BMC Performance Predictor for Mainframes

Details of misses? All the important ones “hit”.

BMC Performance Predictor for Mainframes

File View Tools Utilities Help

Tools List

- Performance Explorer
- Report Explorer
- Graphlet Explorer
- Graph Explorer
- Scenario Explorer

Report Explorer

Workload Activity Summary

Report Description: Summary of overall running workload activity, including number of systems running on.

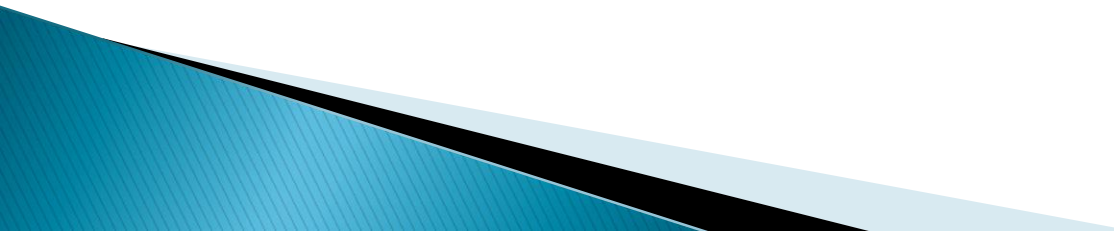
Click on the workload name to show the activity history

Workload Name*	Number Of Logical Systems Running On	Performance Status	Number Of Occurrences	Number Of Occurrences With Bad Performance Index	Percentage Of Occurrences With Bad Performance Index
BATLOW	1	↑	5	0	0.00%
BATNRM	1	↑	24	0	0.00%
BATNRM_CICSUTL	1	↑	3	0	0.00%
BATNRM_OMVS	1	↑	1	0	0.00%
BATPROD	1	↑	25	0	0.00%
BATPROD_DB2UTL	1	↑	3	0	0.00%
DB2TASKS	1	↓	1	1	100.00%
DB2TASKS_DB2_DEB6	1	↓	24	11	45.83%
DB2TASKS_DB2_DEE7	1	↓	24	11	45.83%
OMVSNRM_OMVS	1	↑	25	0	0.00%
STCHI	1	↓	25	21	84.00%
STCLOW	1	↑	25	0	0.00%
STCNRM	1	↑	25	0	0.00%
STCNRM_DB2_DEB6	1	↑	24	0	0.00%
STCNRM_DB2_DEE7	1	↑	24	0	0.00%
STCNRM_HTTP	1	↑	24	0	0.00%
STCPAS	1	↓	25	19	76.00%
STCPROD	1	↓	25	1	4.00%

General Status

BMC Performance Predictor for Mainframes

Present this as =

- ▶ NOT = “Failures in every interval”
 - ▶ BUT = “Six areas – major or minor ones – need addressing”
 - ▶ And analyze and report impact on the business
 - ▶ Also note time of day Development batch at 3 am? One job misses?
- 

Outliers

- ▶ Historical shift right now
- ▶ At one time – you'd drop 'em
- ▶ Examples....

Simplistic but

Two wild
outliers

Use to properly calculate
average response time for
the day

APPLXBNK	tx count	avg rpt	tx > 5 sec	tx > 10	tx > 1000	outliers	factor (tx count * rpt)
9:00 AM	22516	0.23	34	4	0		5178.68
10:00 AM	33534	0.26	36	13	0		8718.84
11:00 AM	66221	0.33	55	22	1	1123	21852.93
12:00 PM	62134	0.31	78	21	0		19261.54
1:00 PM	44986	0.89	45	33	1	3215	40037.54
2:00 PM	57810	0.3	77	29	0		17343
3:00 PM	65007	0.33	49	19	0		21452.31
4:00 PM	58711	0.31	58	15	0		18200.41
5:00 PM	49867	0.39	35	8	0		19448.13
tot tx	460786						171493.38
rpt day	0.372176						
norm !	0.362761						

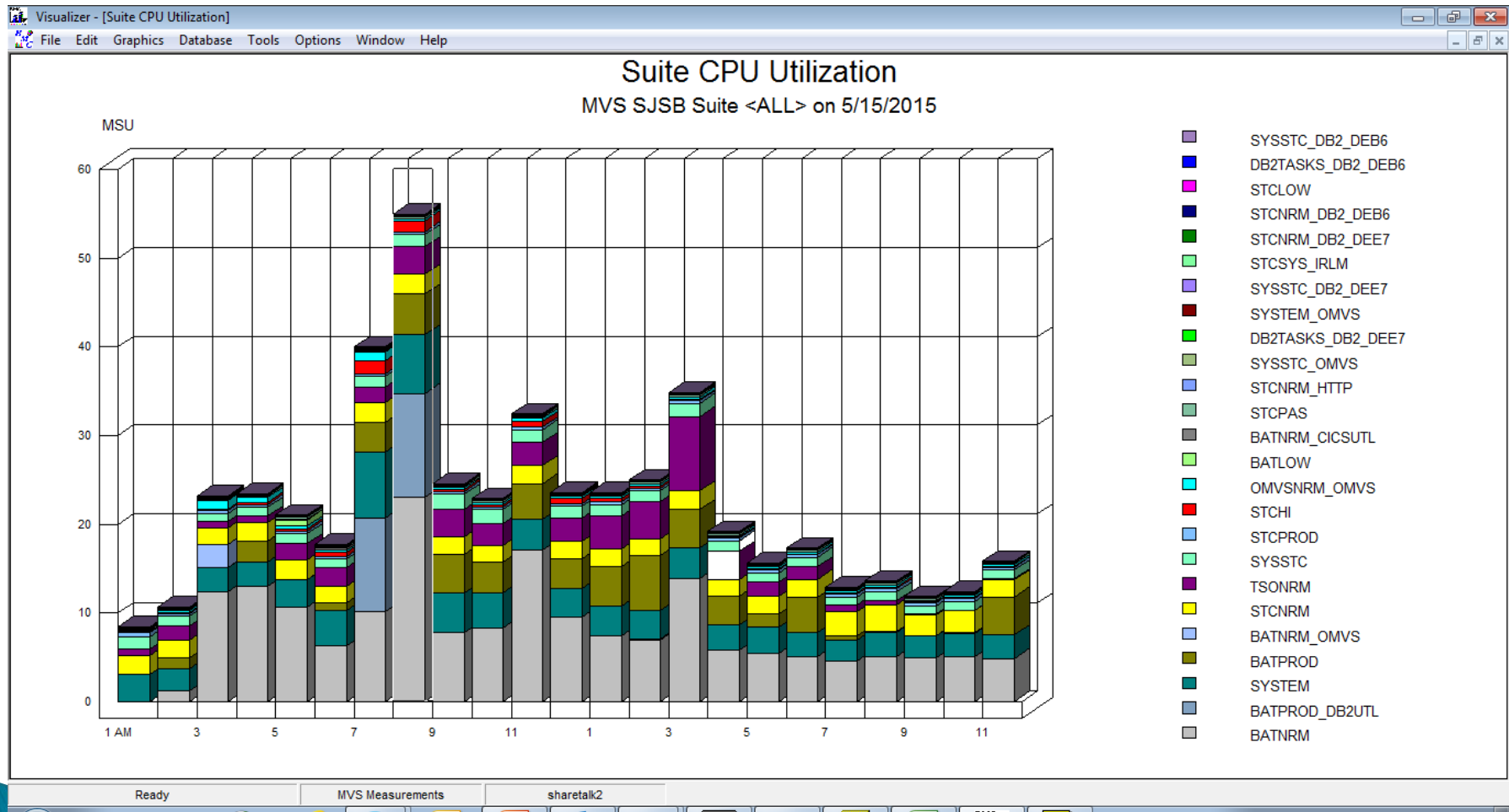
Pulling the one outlier recalcs to .81s,
so there are further problems, possible
looping tx impacting others

Whither monthly license charges?

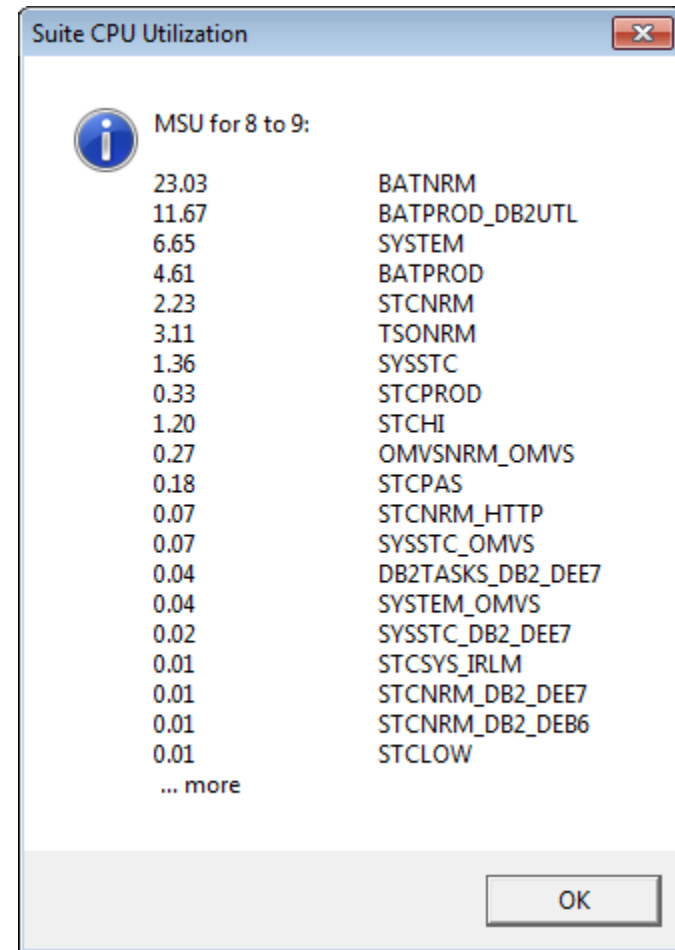
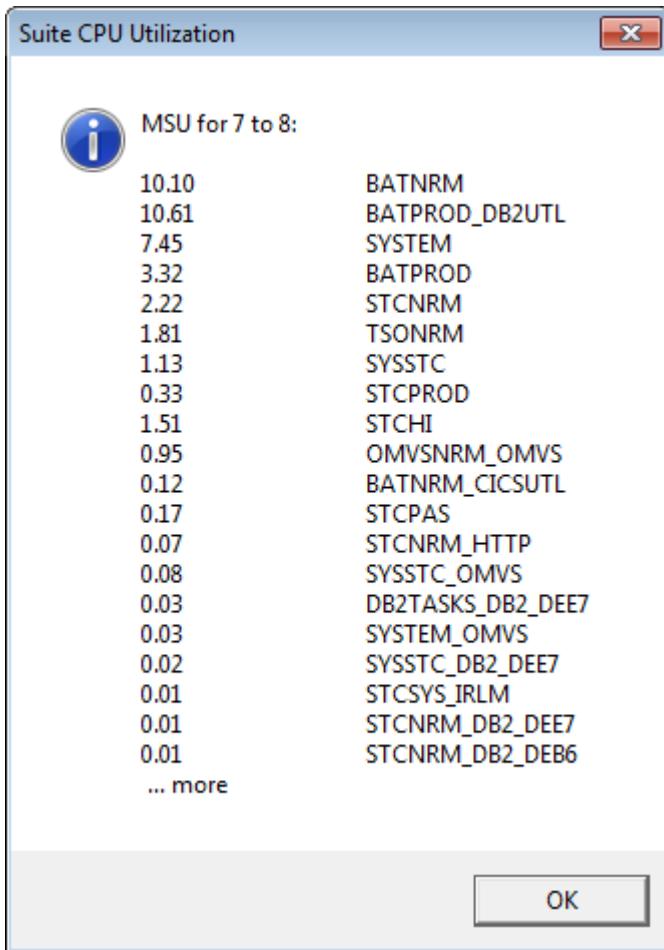
- ▶ With MLC licensing

- Outlying CPU Usage – requiring capping / time shifting of work, now is more critical
- Cannot be ignored, and it must be handled –
-

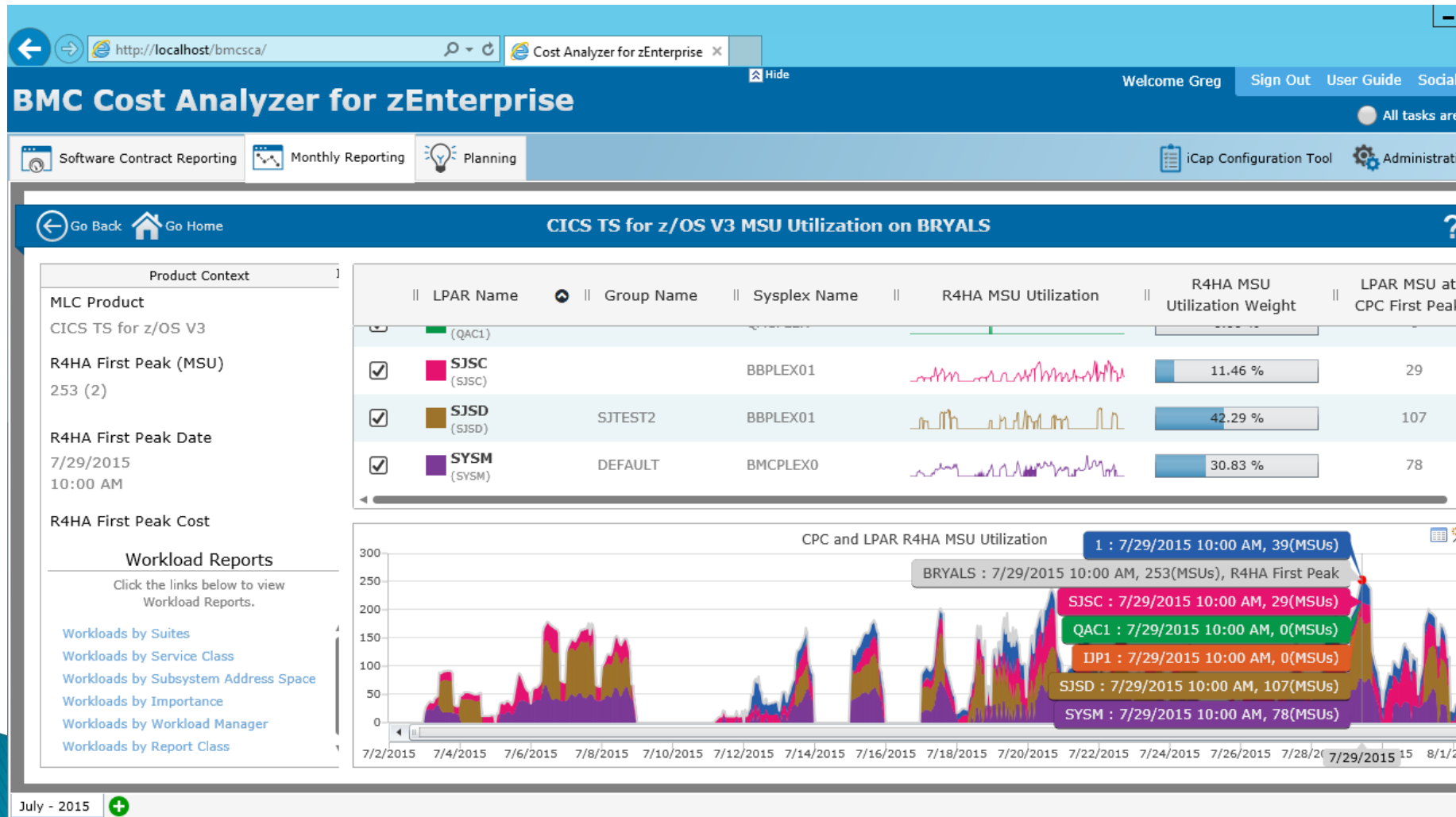
There's a spike I might want to adjust – you might call it an outlier



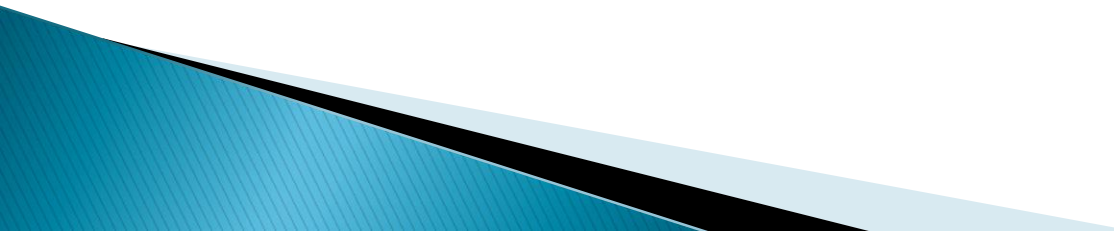
Numbers behind it...



MSU/MLC – look for those outliers!



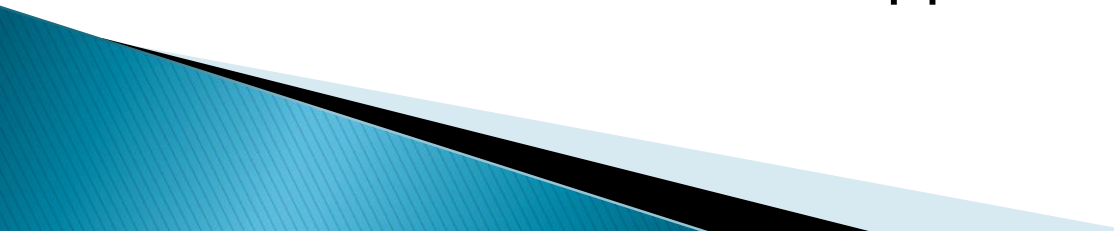
Epilog

- ▶ Back to (another) deep breath
 - ▶ New analysts to mainframe – you need to organize
 - Identify units of work, organize those units along lines of business
 - Establish accountability !
- 

Epilog

- ▶ Understand what is going on within the system
 - Your own graphs can “throw” you
 - Outliers may have low or high impact, know them
 - Be able to explain everything possible

Epilog

- ▶ Know your audience!
 - Executives don't care about paging or I/O rates, but the bottom line (\$\$\$) – know their lingo, speak in their terms
 - Users want to know their own successes/failures
 - Techs want to know those system intricacies
 - Consider a modular approach
- 

Paper on modular approach.

- ▶ Gregory V. Caliri:



Performance reporting in the 21st century – changes in scope and direction. [Int. CMG Conference 2006](#)

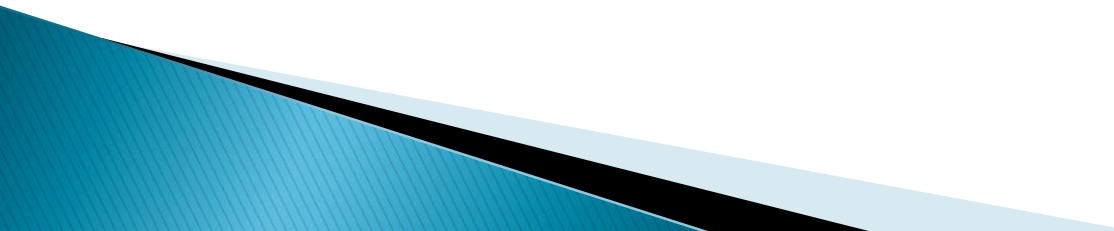
- ▶ Greg_caliri@bmc.com

- ▶ 781 257 3120

Epilog

- ▶ Know your metrics!
 - Know what you'll be explaining
 - Practice practice, and more practice!
 - Take suggestions but
 - Don't get taken down side roads

Epilog

- ▶ If you are asked “what does this mean?” over and over again – next time make it simpler
 - ▶ Don't present something you can't explain
 - ▶ Don't present irrelevant info
 - ▶ Do explain everything that **is** relevant
- 

Thank you!

*“I didn’t know my
parents could
dress so cool”*

*– Jill Caliri
Patruno*

