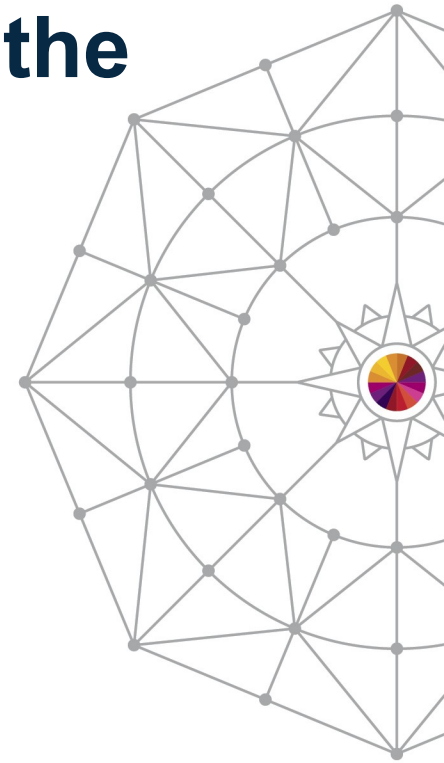
15950: Cross-server Sizing in the Linux (and zLinux) World

Monte Bauman
IBM Columbus
mbauman@us.ibm.com

Monday, August 04, 2014
15950

http:/www.SHARE.org



#SHAREorg

SHARE in Pittsburgh 2014

Topics

- **The Challenge**
- The (Statistical) Remedy
- The Real (Messy) World
- Conclusions (and Plausible Next Steps)

The Challenge… The Opportunity…

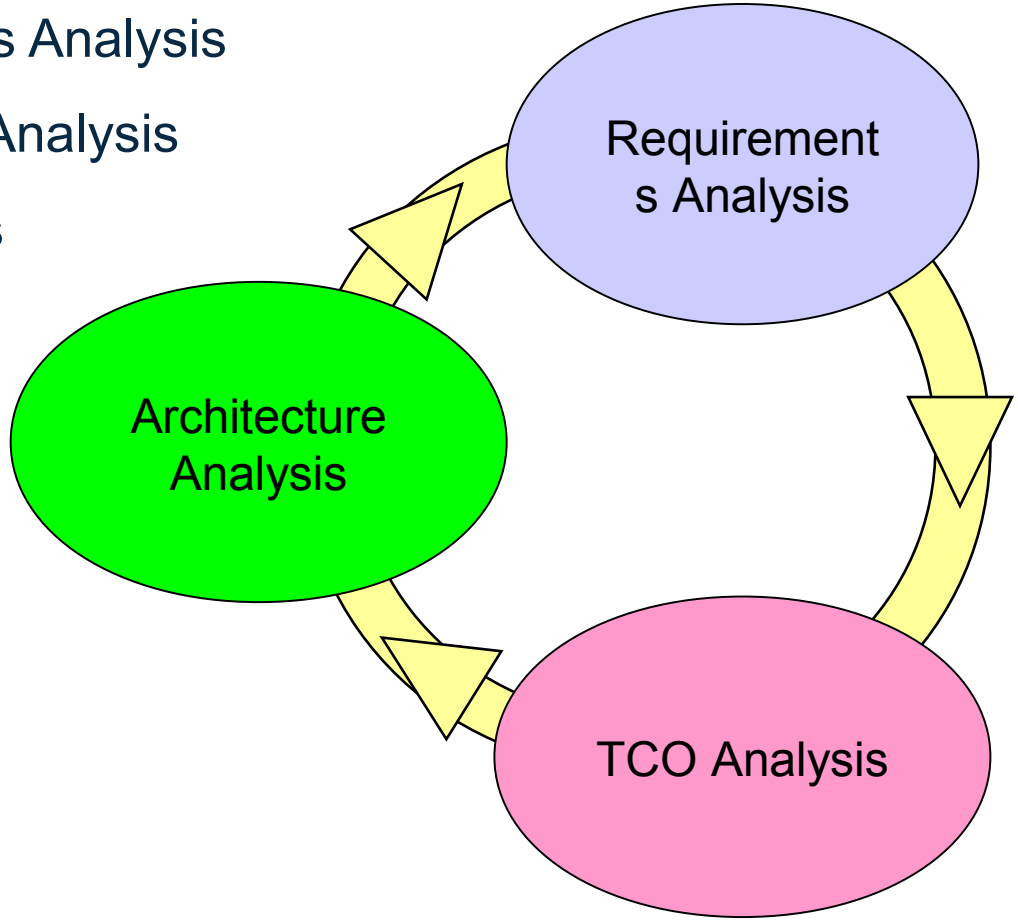
- Robust and pervasive “Middleware” has made for a software world full of “portability”
  - e.g. Java (aka WebSphere, Tomcat, etc.)
  - e.g. Oracle
  - e.g. BPM ODM WMB DB2 etc.
- Allowing IT shops to run those apps where it makes the most sense
  - aka “IT Optimization”

IT Optimization

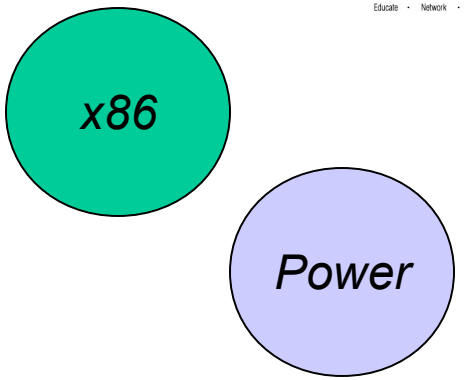
- Run workloads where it makes the most “sense”…
  - Where it costs the least “cents”! … ?
  - Where it risks the least “cents”! … ?
  - Where it most quickly starts making “cents”! … ?

IT Optimization Analysis

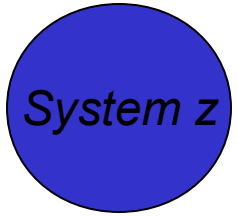
- How do we make sure work is running in the right place?
  - Requirements Analysis
  - Architecture Analysis
  - Cost Analysis



Analysis w/o Paralysis



- Doing all that analysis is a LOT of work
  - Takes a lot of time
    - Costs a lot of money
      - *And when we do it we're still not sure it's right, so …*
      - *Let's just do things like we've always done things*
- Or…



  - You find patterns and built methodolgies
    - You build tools and automate
      - *Harvest (encode) experiences (measurements) of others*

End of Section

Topics

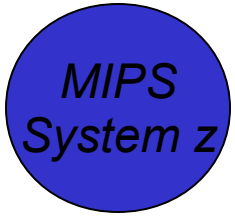
- The Challenge
- **The (Statistical) Remedy**
- The Real (Messy) World
- Conclusions (and Plausible Next Steps)

Inches and Centimeters

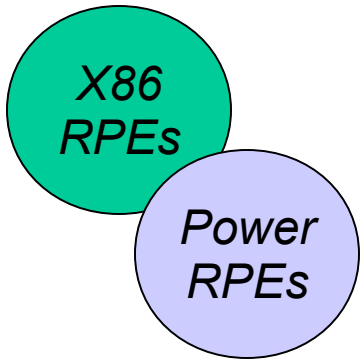
- How many centimeters are there in one inch?
- What if you are on the west coast?
- What if you are on an island?
- What if you are on a reality TV show?

- So what's this got to do with "Cross Server Sizing?" …

MIPS and RPEs



- How many MIPS are there in an RPE?
- How many RPEs are there in a MIP?
- What's this you say?!?!?! … "It Depends" … Ugh (not again)



- BTW
  - What's an RPE?
  - Gartner (formerly Ideas International)
  - Relative Performance Equivalent (aka MIPS for distributed servers)
  - License their data (enjoy their Ts&Cs) … enjoy their arbitration!

The Gartner Group RPE Table

Dynamic Link to RACEv Ideas Table

Workload Factors

- IBM Patent / Joe Temple (IBM DE (not retired))
- MIPS = RPEs / WLF
  - WLF = Workload Factor
  - WLF = a constant
    - derived from workload movement observations
      - *distributed to z, or z to distributed (the more the better)*
  - WLF = a variable
    - Observations show different workload types map from MIPS to RPEs differently (i.e. the WLF will differ by workload type)
      - *aka … CPU-intensive workloads map differently then I/O-intensive workloads map differently then memory-nest intensive workloads … etc.*

The Middleware-Based Workload Factor Table

Dynamic Link to RACEv Workload Factor Table

End of Section

Topics

- The Challenge
- The (Statistical) Remedy
- **The Real (Messy) World**
- Conclusions (and Plausible Next Steps)

MIPS = RPEs / WLF

- So now we can cross-server size, right?
- Well …
  - The world is a little messy
  - Workloads are collections of servers
    - Servers are discrete (sometimes) / virtual (most times)
    - Servers (and the workloads being run) are connected
      - *Sometimes logically / sometimes virtually*
      - *Sometimes physically locally*
      - *Sometimes physically over looooong distances*
    - Sometime “connections” effect the response time budget…

Real (Messy) Sizing Methods (the Cookbook)

- Step 1 – Understand the architecture
- Step 2 – Get inventory data
- Step 3 – Get measurement data
- Step 4 – Peak-based measurements analysis
- Step 5 – Map net peaks onto candidate target servers
- Step 6 – Sensitivity analysis

Step 1 – Understand the Architecture

- Understand the architecture
  - Code
  - Containers
  - Platforms
  - Connectors
  - Connections

Architecture Analysis Workbook

Dynamic Link to RACEa Architecture Analyzer

Step 2 – Get Inventory Data

- Get inventory data
  - Physical server attributes (make model chips cores GHz)
  - Virtual server attributes (allocation)
  - Hypervisor
  - Operating system
  - Middleware (and other software)
  - Workload characterization and priority
    - Interactive production high priority
    - Head-less non-production low priority
    - Or something somewhere in between

Step 3 – Get Measurement Data

- Get measurement data
  - 15 minute intervals (preferred / others can be dealt with)
  - VMstat SAR or similar reporting output
  - Over 5 or 7 day period
  - Capturing normal workload peaks

Step 4 – Peak-Based Measurements Analysis

- Peak-based Measurements Analysis
  - The collective peak
    - Find when the entire collection peaks together
      - *That's what the target server needs to accomodate*
  - Time period exclusions
    - Remove backups and virus scan measurement intervals
  - Anomaly removal
    - Percentile filtering
  - Peak elongation
    - Allow peaks for low/medium … or headless processes … to "potentially" elongate on the target

Step 5 – Map Peaks onto Candidate Target Servers

- Map net peaks onto candidate target servers
  - Allow for target server's hypervisor overhead
  - Allow for hypervisor's inefficiency if managing mixed workloads
  - Allocate resources for topology variable overhead
    - Co-location (or lack of co-location) effects between app tiers
      - *TCP/IP stack processing*
      - *SSL processing*
      - *Data marshalling and serialization processing*
  - Plan for the target server's planned Saturation Design Point
    - required headroom
      - *Based upon size of target and number of virtual servers on target*

Inventory & Peak Analysis Workbook

Dynamic Link to RACEi Inventory and Peak Analyzer

Step 6 – Sensitivity Analysis

- A house of cards…
  - Is only a bad thing if you don't know it is "a house of card"
- The methodology is sound and repeatable
  - Based upon numerous observations and experiences
- But you are not (necessarily) like everyone else
  - What factors (which estimates made) are "suspect"?
  - Change them and see…
    - re-drive the process
      - *That's why we encoded it*
        - *Bound the analysis (best and worst case)*

End of Section

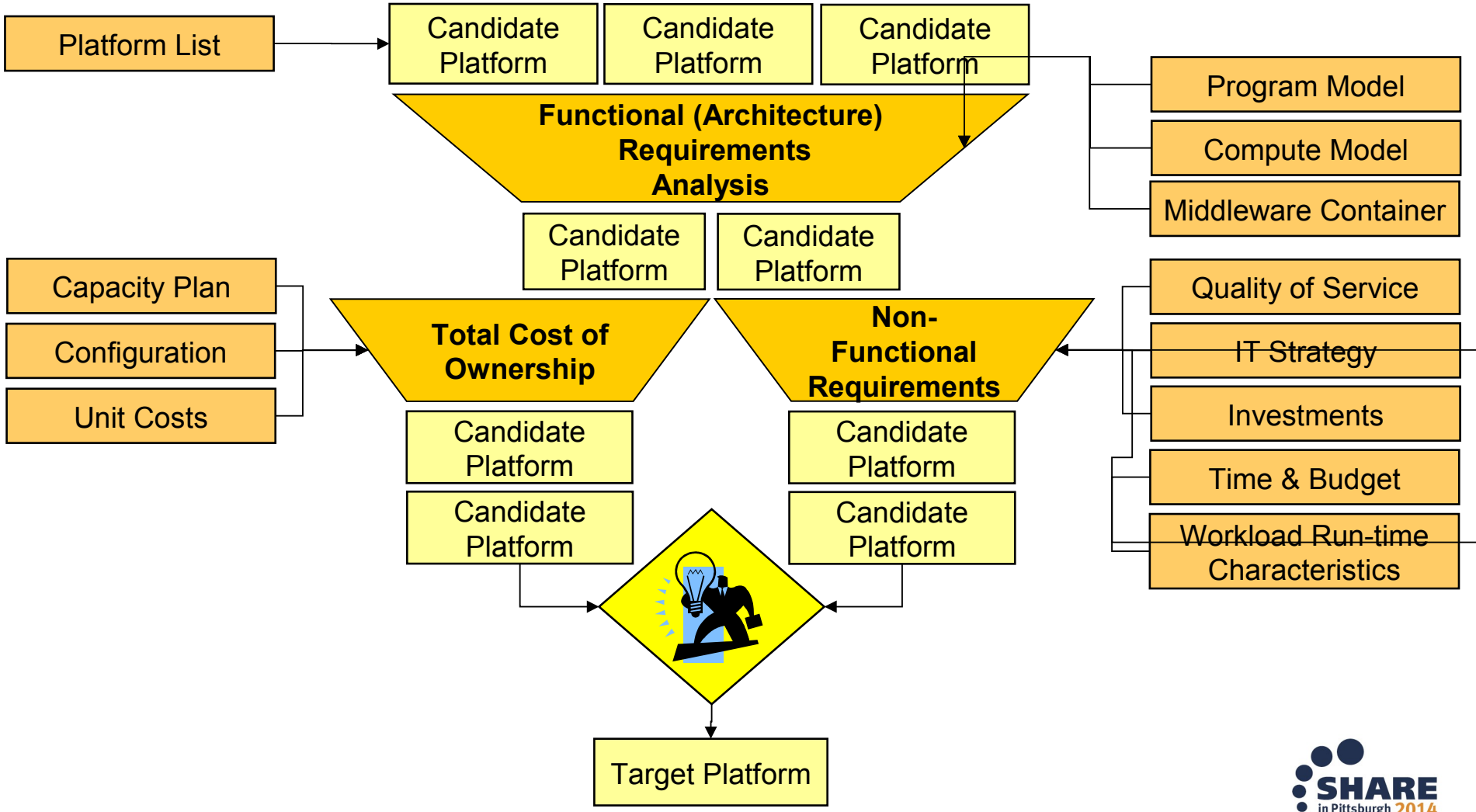
Topics

- The Challenge
- The (Statistical) Remedy
- The Real (Messy) World
- **Conclusions (and Plausible Next Steps)**

Conclusions

- Tools and Methods
- Break the Analysis Paralysis Log Jam
- Establish Technical Equity
- To Compare Platforming Alternatives
  - for “IT Optimization” Purposes
    - Run the right work
      - *At the right time*
        - *On the right platform*
        - *For the right reason(s)*

IT Optimization … Enabled



And then it was the end…

- Final comments?
- Questions?
- Requests?

- Thank You

- Monte Bauman … mbauman@us.ibm.com