IBM z Systems

Performance Topics

Robert Blackburn Ph.D.

3/2015

IBM

Disclaimer

- Any reference to future plans are for planning purposes only. IBM reserves the right to change those plans at its discretion. Any reliance on such a disclosure is solely at your own risk. IBM makes no commitment to provide additional information in the future.

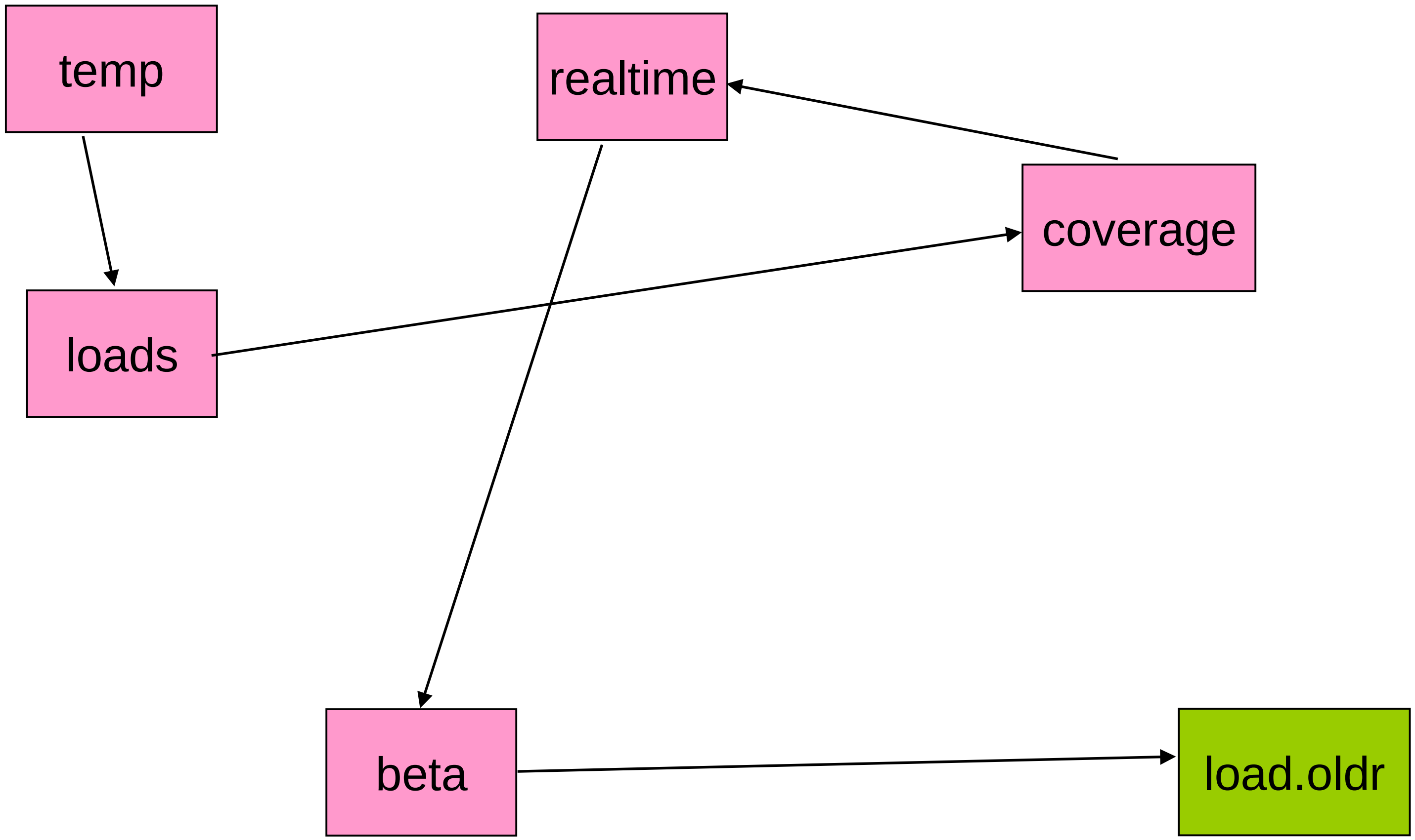
© 2015 IBM Corporation

Logical record cache(LRC)

- Not persistent – Files must go to DASD
- Assume single CEC
- Example that captures essence of the effect
  - From TPF File System
  - Directories
    - /temp/loads/coverage/realtime/beta/load.oldr
  - really only interested in 'load.oldr'
- If put in VFA
  - access all 6 each time to get to 'load.oldr'

© 2015 IBM Corporation

File System Directory Access



/temp/loads/coverage/realtime/beta/load.oldr

© 2015 IBM Corporation

LRC - Loosely Coupled

- Working assumptions
- Number CECs = 8
- Total record access cost = 2 x 1 = 2
  - Accesses to get record = 2 (index + data)
  - Let DASD cost = 1
- LRC read = .5
  - rough estimate
  - just the data and in memory
- Castout cost = 3.5
  - rough estimate
  - one CEC only, others minimal cost, dup W

© 2015 IBM Corporation

Calculating LRC gains

- Assume R/W = 10
  - So cycle is RRRRRRRRRR then W then invalidate all occupied CECs
  - Key number is (R/W)/(# CECs) = 10/8 = 1.25
- Let E(O) = expected number of occupied CECs
- P(CEC not occupied) = exp(-1.25) = .287
  - Not obvious - accept for now
- E(O) = 8 x P(occupied) =8 (1-.287) = 5.7
- Without LRC – cost per cycle
  - 2 x 10 + 2 = 22
- With LRC – cost per cycle
  - Miss + LRC hits + castout
  - 5.7(2) + (10-5.7)(.5) + 3.5 = 17.1

© 2015 IBM Corporation

Similar calculations with R/W = 5

- Without LRC – cost per cycle
  - 2 x 5 + 2 = 12
- With LRC – cost per cycle
  - Miss + LRC hits + castout
  - 3.72(2) + (5-3.72)(.5) + 3.5 = 11.6
- So rough parity now

© 2015 IBM Corporation

Executive summary - LRC

- LRC can be beneficial with modest R/W ratios
- We used 2 = total record access cost (index+data)
  - Better results if this increases
    - e.g. the 6 in file system just discussed

© 2015 IBM Corporation

High utilization and low priority work

- Generally shutdown caused by insufficient Tape or DASD/CU power
  - IOBs, ECBs queue
- Interesting case of CPU at .999 busy and bouncing in and out of ECB caused shutdown
- Small % of msgs were huge consumers of MIPS
  - These were put on Defer list
- As CPU utilization increased less time was available for Defer processing
- Defer ECBs totaled over 500 pushing system into shutdown

© 2015 IBM Corporation

High utilization and low priority work

- So when in shutdown processing largely Defer work
  - Said another way - transactional work is waiting outside the system while the CPU sits at .999 busy
- TPF basically won't shutdown because of CPU utilization driven by interactive work
  - CPU at ~ .99 the queue will be 20 or so per CP
  - In this case 3W MP so noise level increment
- Clearly there will be a huge queue OUTSIDE of TPF

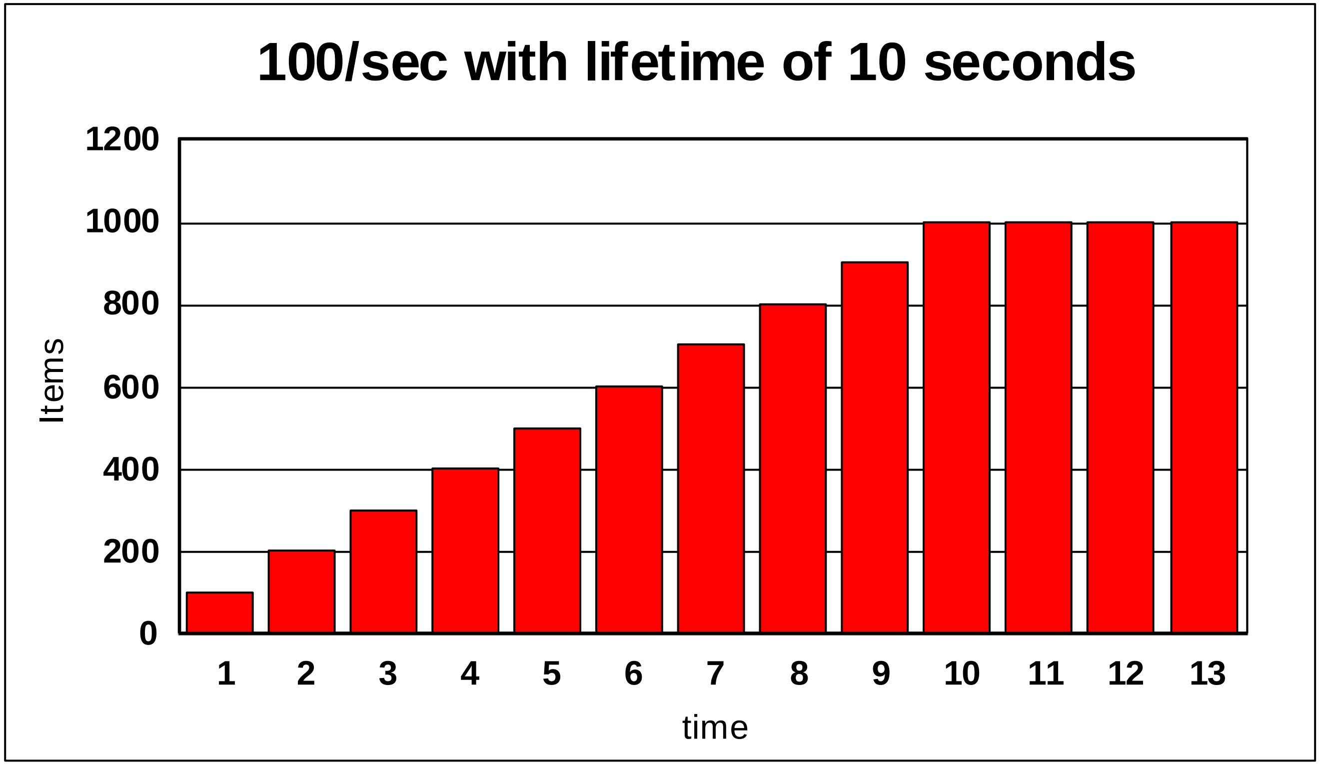
IBM
© 2015 IBM Corporation

Executive summary – high utilization

- Solutions to CPU overload problem
  - Adjustable MIPS
    - - e.g. 507 -> 607
    - no TPF IPL
  - Potentially shed some work
    - via Lodic and depart

© 2015 IBM Corporation

Little’s Theorem



ST Pools – example description

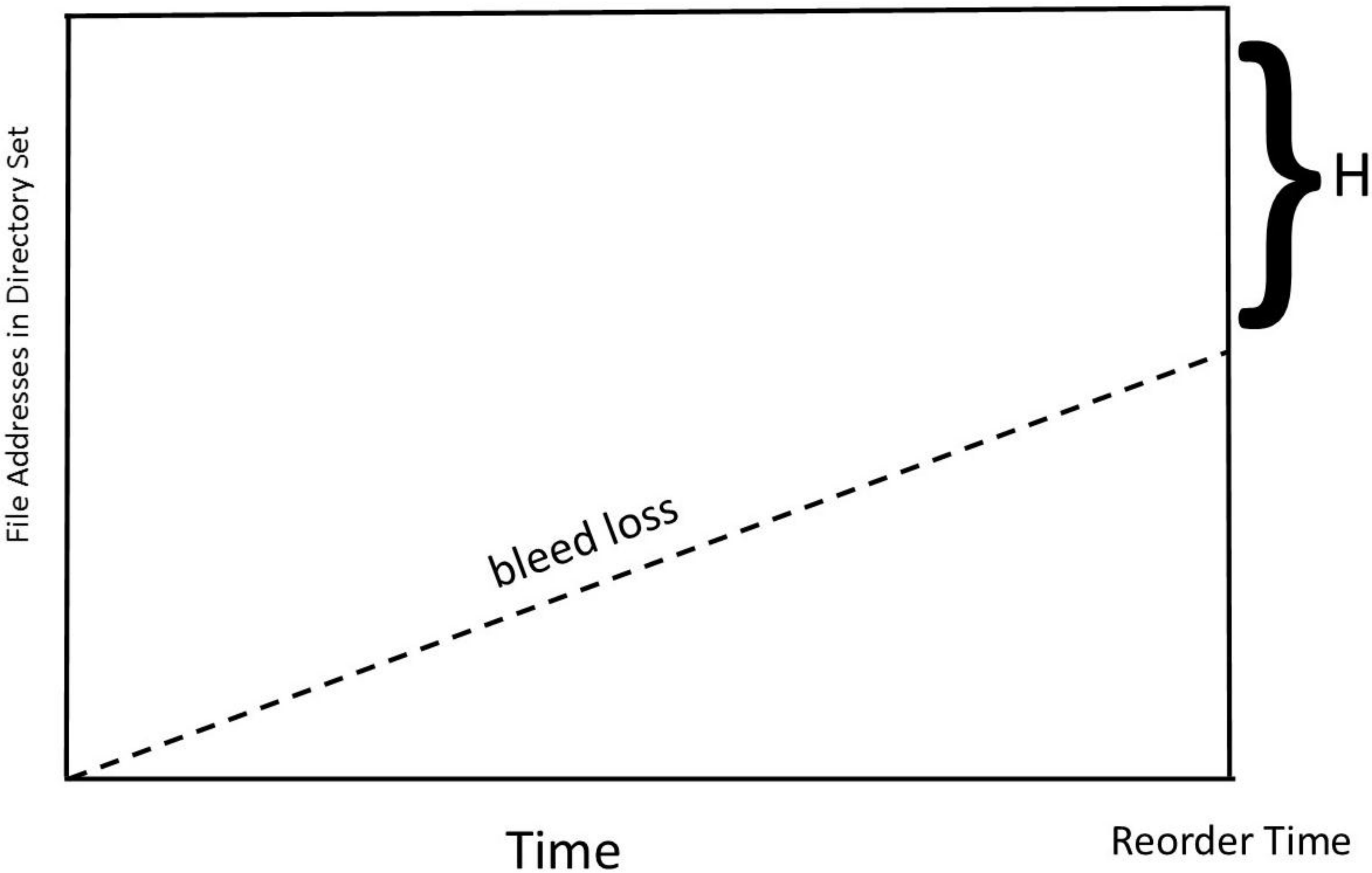
- Terms
  - Reorder is go thru 1 directory set
  - Recycle is thru all 20 directory sets
- 80 directories => 80 x 8000 = 640K file addresses
  - Assume reorder at 100% full – or .9 N = 640K
- 15 minute reorder time
- Get rate = 6000 /sec
- Bleed % = 5%
  - Thus bleed rate = .05(6000) = 300/sec
- First pass time through all 80 directories
  - 640K / 6K = 107 seconds

© 2015 IBM Corporation

ST Pools

- Bleed rate consumption after 15 minutes
  - 300/sec x 900 sec = 270K
- Since we reorder after 15 minutes there must be a resident set H such that
  - 640K = H +270K
  - so H= 370K
- By Little's Theorem
  - W = 370K/6000(.95) = 65 seconds
  - ST pool average lifetime

© 2015 IBM Corporation

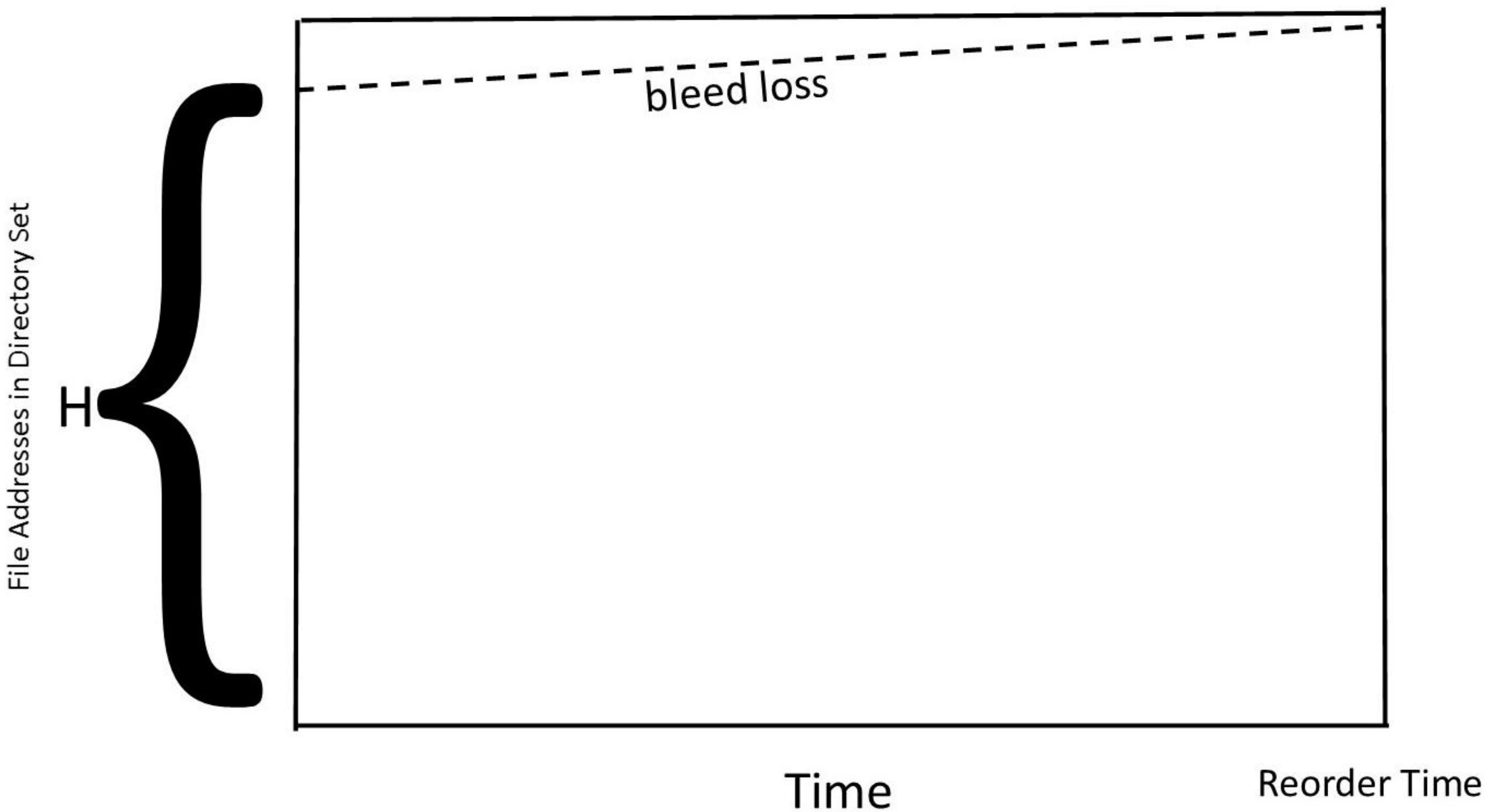


© 2015 IBM Corporation

What happens if Get rate changes

- Increase get rate by 30% and keep characteristics of bleed and wait time
  - 6000 x 1.3 = 7800/sec
- H= .95 x 7800/sec x 65 sec = 482K
- Bleed rate = 300/sec x 1.3 = 390/sec
- Reorder time = (640K -482K) / 390 = 405 sec(6.75 min)
- So new rate =1.3 but reorder time divided by 2.22
  - 15 /2.22 =6.75
- if pool reorder time was pure bleed(H=0)
  - 15/1.3 = 11.5 minutes
- Small increments and observe reorder time you are fine

© 2015 IBM Corporation



bleed loss
File Addresses in Directory Set
H
Time
Reorder Time

Executive summary –ST pools

- You are fine if
  - Small increments of ST rate
  - observe reorder time
- Largest concern is with sudden ST rate increase
  - Proper planning will resolve this

IBM

© 2015 IBM Corporation

Fundamental Limits on Single CPU Speed

"In terms of size [of transistor] you can see that we're approaching the size of atoms which is a fundamental barrier, ..."

Gordon Moore, April 2005*

Modest increases in single-thread engine performance

Smaller increase than zEC12 / z196

Growth rate continues to decline, but **no** decrease in single-thread performance

Faster growth of SMP (n-Way) needed to sustain box growth

Hardware innovations to enable consistent performance on large SMP

Focus on constraint relief across software stack

* Techworld, Operating Systems and Servers News, 13 April 2005

© 2015 IBM Corporation

TPF customers well positioned for large MPs

- Worked with several customer’s performance data
  - Observed very small MIPS loss per CP in the 6 to 12 CP range
- Minimal ‘hot’ cache lines bouncing around
  - TPF scheduler very efficient at dispatching work
  - ECB Heap and private area(381,1K,4K) tend to stay on one CP for their lifetime
- CPUMF is very useful to catch any issues as we grow to 30W and beyond
- ITRRs go up to 30W for z13

© 2015 IBM Corporation

Predicting TPF performance after application changes

- Before see results on production system
- Possible future topic and want your input
- Use of owners
- At least at next TUG
- If sufficient interest then another webinar
- Don’t need to implement all at once
- Guess
  - Top 30 applications consume 50% of MIPS
- How to verify our predictions
- Is this the right stuff to focus on?

IBM

© 2015 IBM Corporation

Trademarks

- IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml.
- *(Include any special attribution statements as required – see Trademark guidelines on https://w3-03.ibm.com/chq/legal/lis.nsf/lawdoc/5A84050DEC58FE31852576850074BB32?OpenDocument#Developing%20the%20Special%20Non-IBM%20Tr)*

Notes

- Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.
- All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.
- This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.
- All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.
- Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.
- Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.
- This presentation and the claims outlined in it were reviewed for compliance with US law. Adaptations of these claims for use in other geographies must be reviewed by the local country counsel for compliance with local laws.

© 2015 IBM Corporation