

## IBM zHyperlink: Performance Analysis For A New Mainframe I/O Link

Anna Shugol & Martin Packer

IBM

08/11/2017

Session LF



---

## Please note

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion.

Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract.

The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including 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 results similar to those stated here.

---

## Acknowledgements

- Harry Yudenfriend, IBM Fellow, Storage Development, Poughkeepsie
- Peter Sutton, IBM Distinguished Engineer, System Z, Poughkeepsie
- Denis Senin, IBM Storage Expert, Hursley
- “zHyperlinks and z14 new features exploration” zChampion workgroup

---

## Agenda

- Why I/O latency matters
- I/O today
- zHyperlinks: changing I/O paradigm
- Technical practicalities
- zHyperlinks and replication solutions
- zHyperlinks: roadmap
- Performance analysis deep dive

---

# Latency Matters

## Business value of the latency reduction

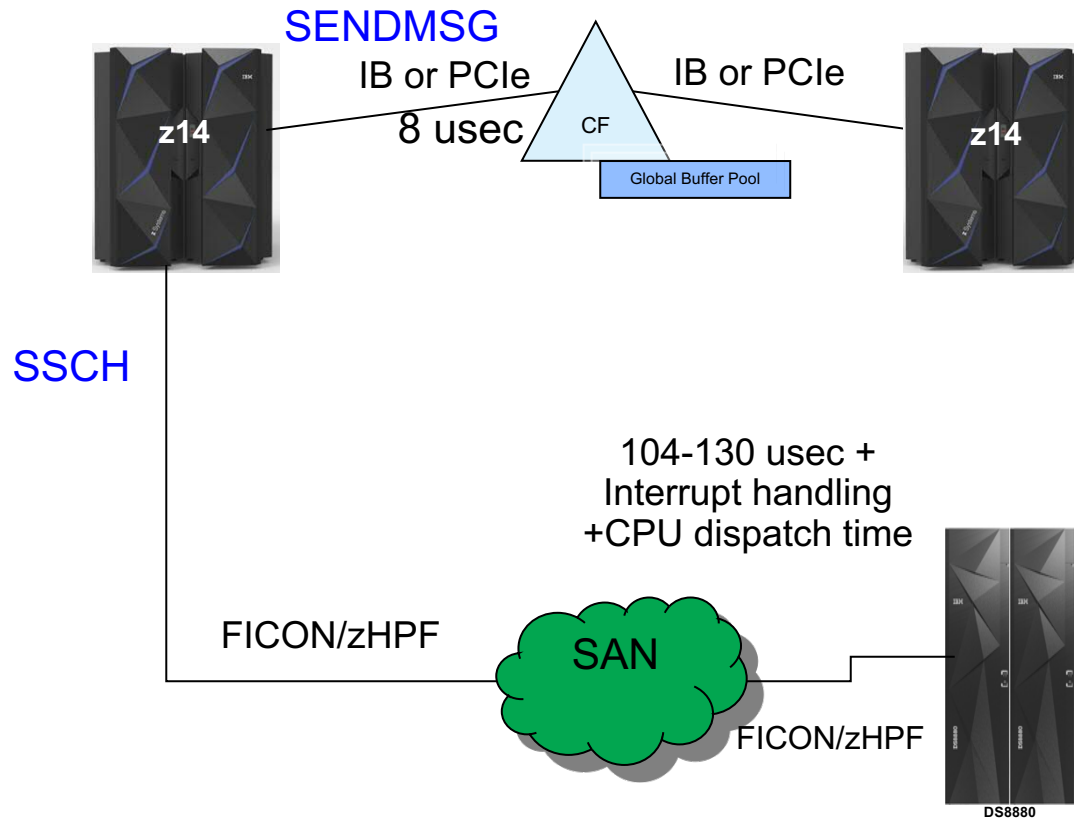
- Avoid Lost Revenue Due to:
  - Abandoning fraud detection
  - Clients switching credit cards
- Grow business opportunities
  - Add additional information on simple transactions
  - Add additional business offers
- I/O sensitive DB2 and VSAM transactions execute faster
  - Better client experience with lower I/O latencies
  - Additional business opportunities for top line growth with more functional applications
- Faster links improve availability
  - Handle work load spikes and recover faster after outages
  - Faster VSAM VERIFY processing for faster application and System restart (e.g. Catalog, HSM)
- Faster DB2 Batch processing with faster DB2 index splits
  - Index split performance is the main bottleneck for high volume INSERTs
- Dramatically improved application and middleware scaling
  - Workload growth without redesigning applications for scale

---

## SAP/DB2 z/OS Transaction elapsed time breakdown

DB2 Server CPU time:	5%
<b>Lock/Latch + Page Latch:</b>	<b>2-4%</b>
<b>I/O service time</b>	<b>60-65%</b>
<b>Dispatcher (CPU) Latency:</b>	<b>20-25%</b>
Network (TCP/IP):	<b>4-6%</b>

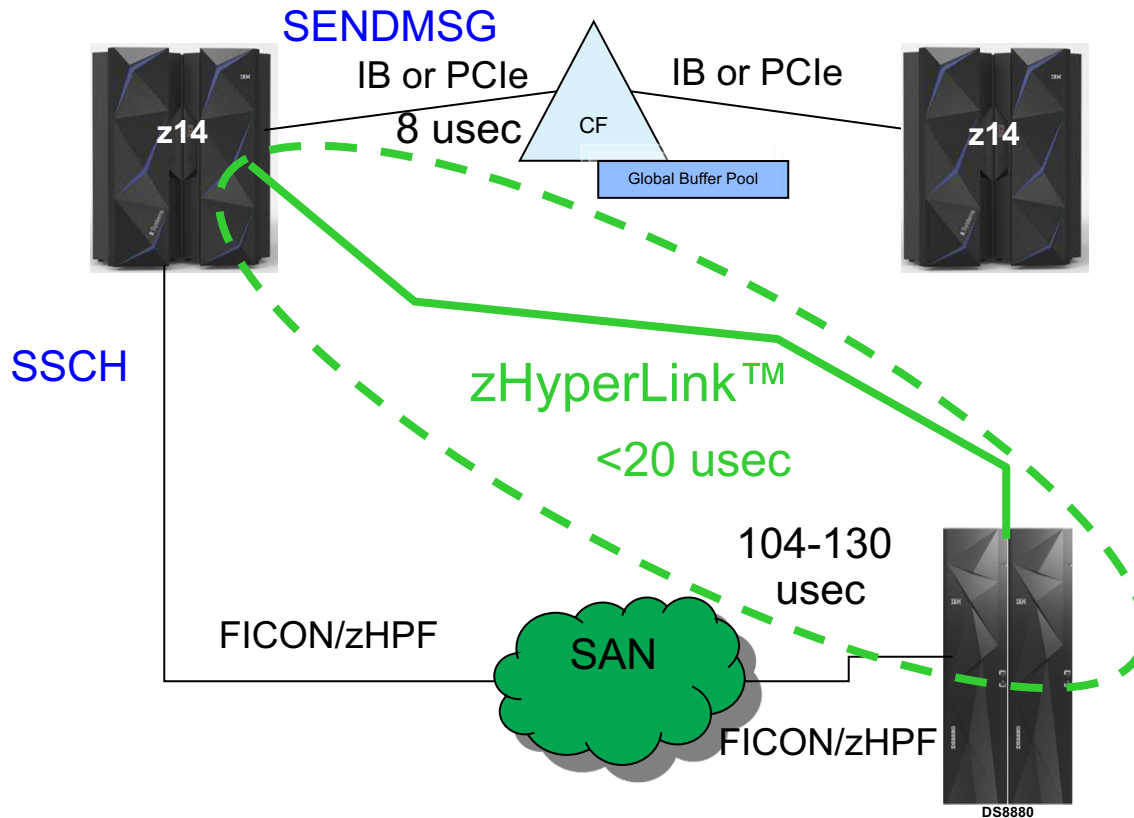
# Today's I/O



Typically over 80%  
cache hit ratio on  
random reads.

100% cache hit on  
writes.

# With zHyperLink™



Typically over 80% cache hit ratio on random reads.

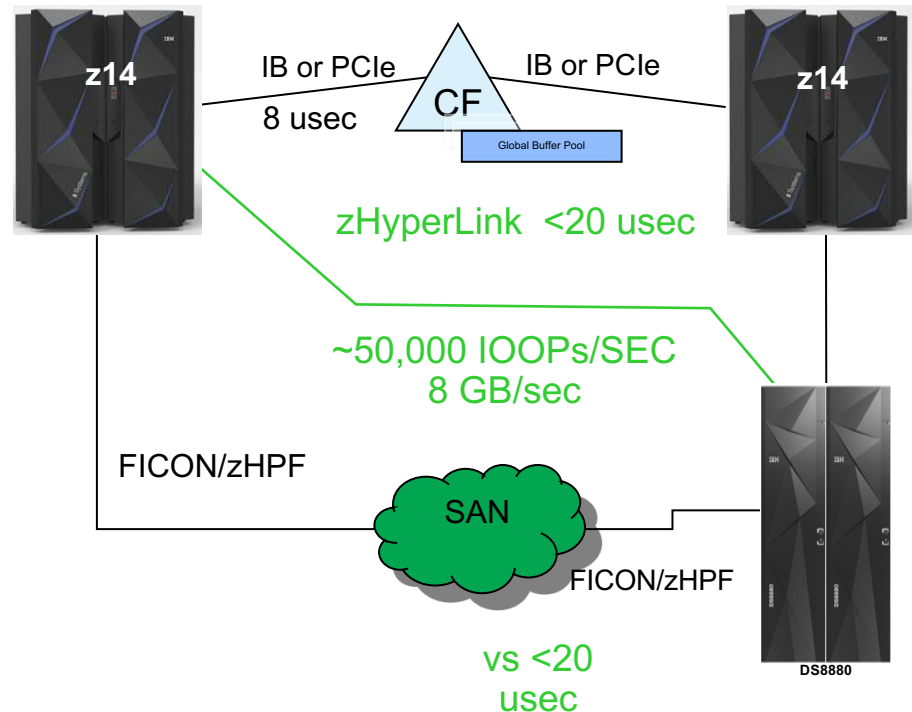
100% cache hit on writes.



# How does IBM zHyperLink™ change the game?

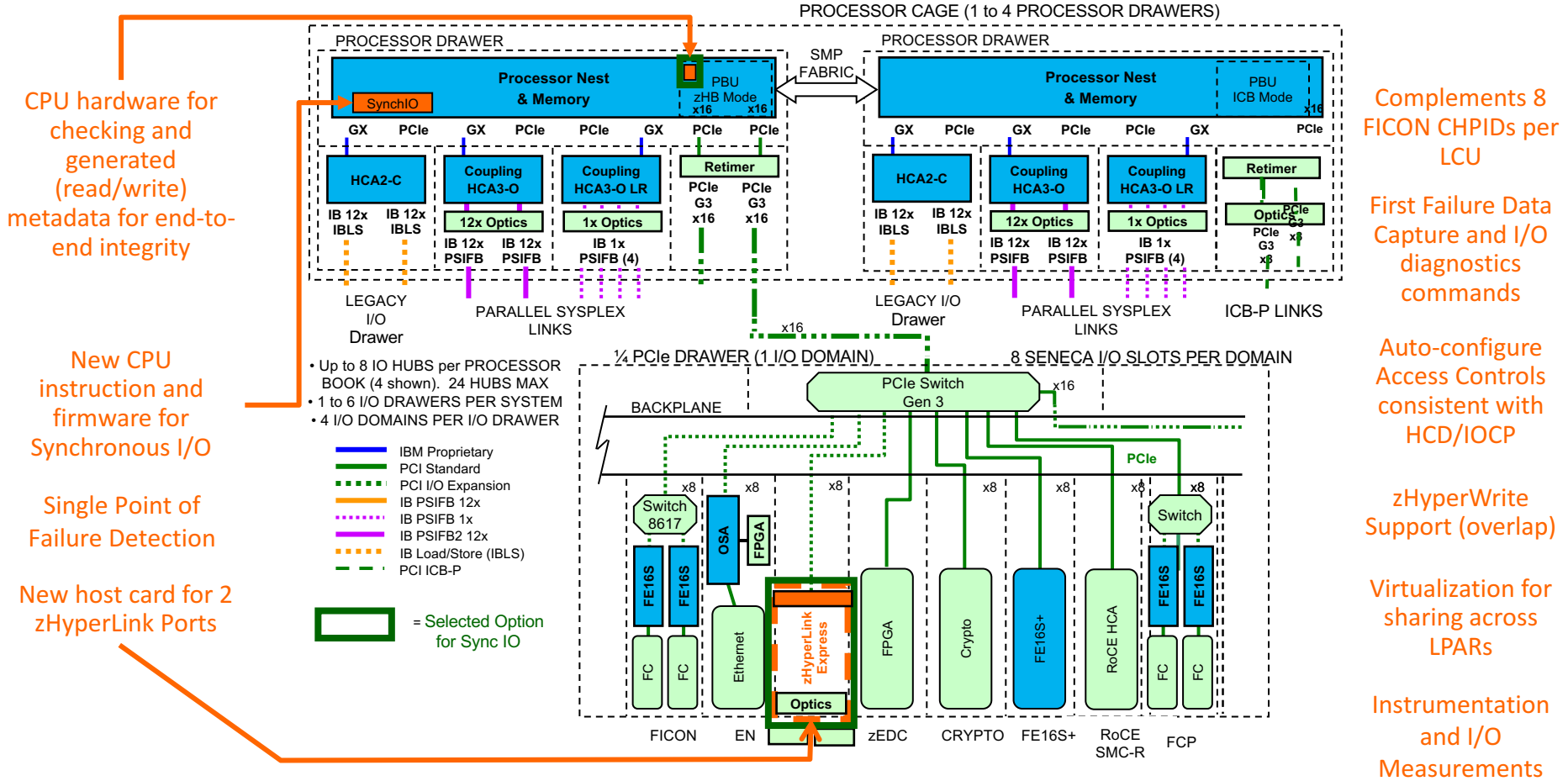
- zHyperLink™ is FAST enough the CPU can just wait for the data
  - No Un-dispatch of the running task
  - No CPU Queueing Delays to resume it
  - No host CPU cache disruption
  - Very small I/O service time
- Operating System and Middleware (e.g. DB2) are changed to keep running over an I/O
- Transparently Gives DB2 apps fundamentally better latency than apps on platforms without zHyperLink
  - Excluding 100% in memory databases

## New industry I/O link for Mainframe



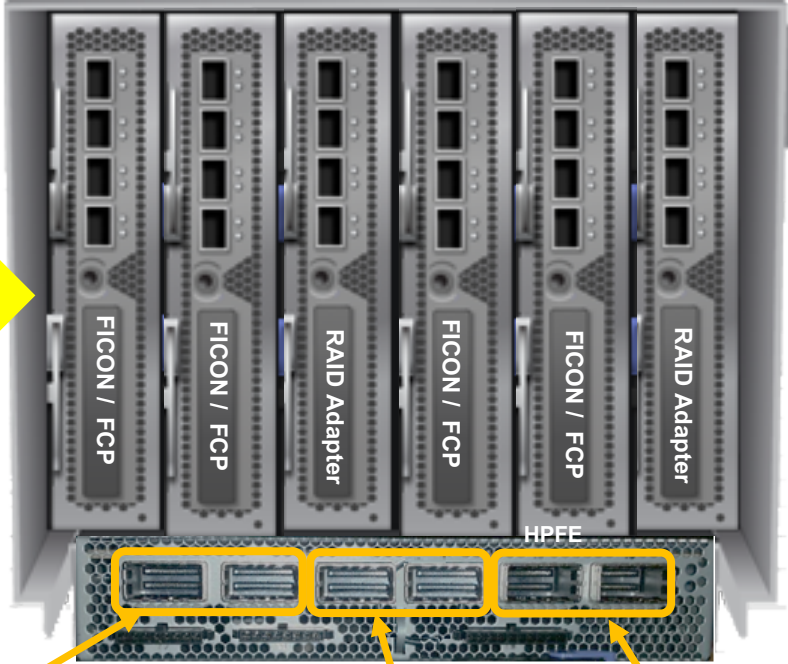
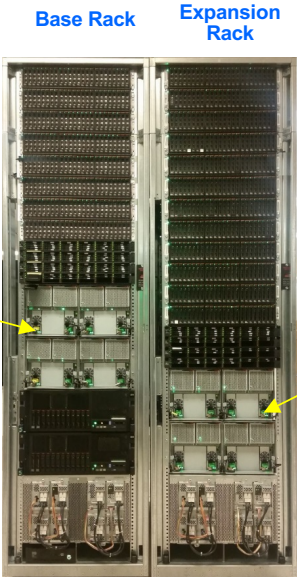
{ 104-130 usec + Interrupt handling + CPU dispatch time + Re-load L1/L2Cache }

# z14 I/O and zHyperLink Express



# DS8880 zHyperLink™ Ports

The DS8880 I/O bay supports up to six external interfaces using a CXP connector type.



DS8880 internal  
PCIe Fabric

zHyperLink ports

HPFE

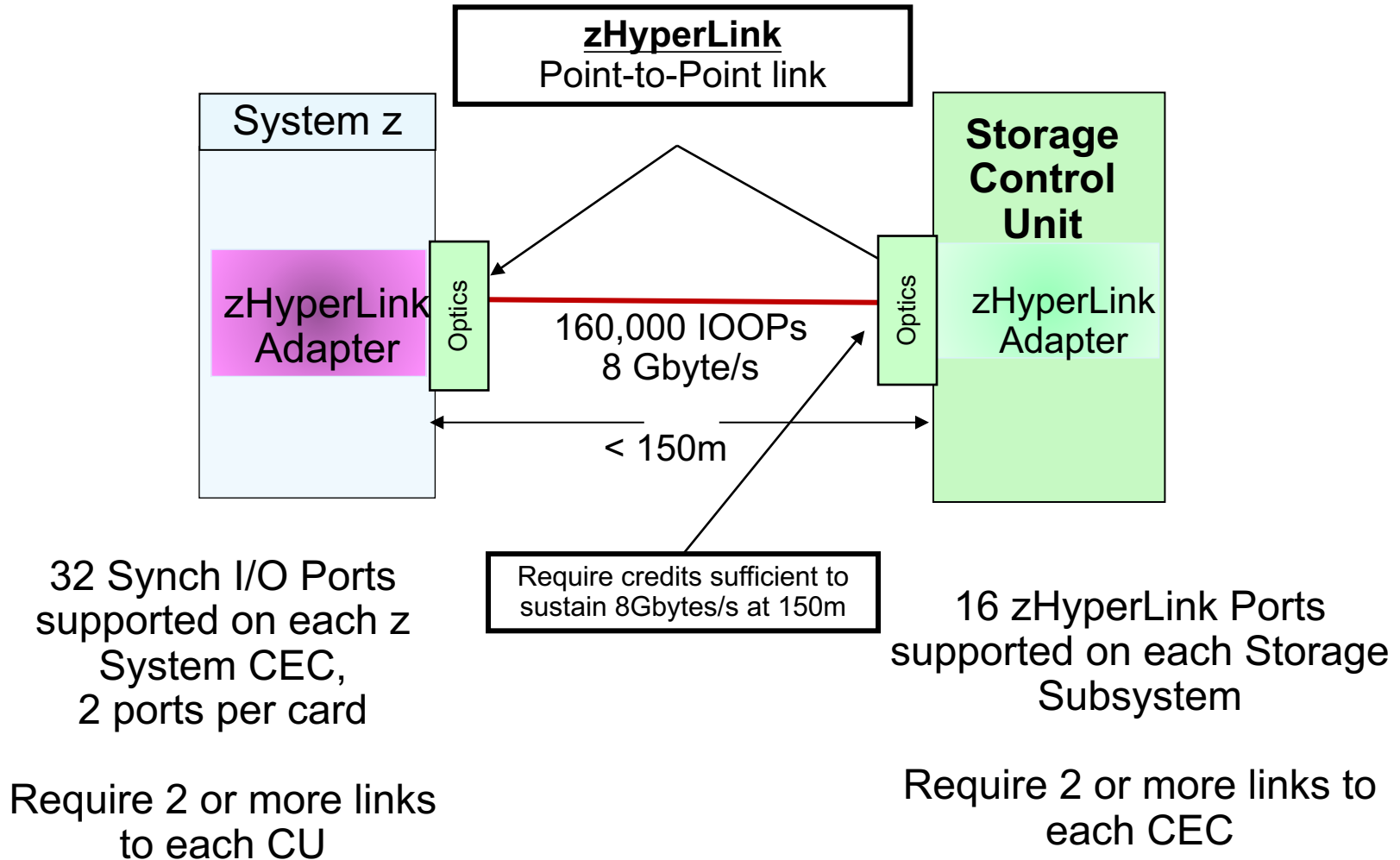
Investment Protection – DS8880 hardware shipping 4Q2016 (models 984, 985, 986 and 988), older DS8880's will be field upgradeable at GA

---

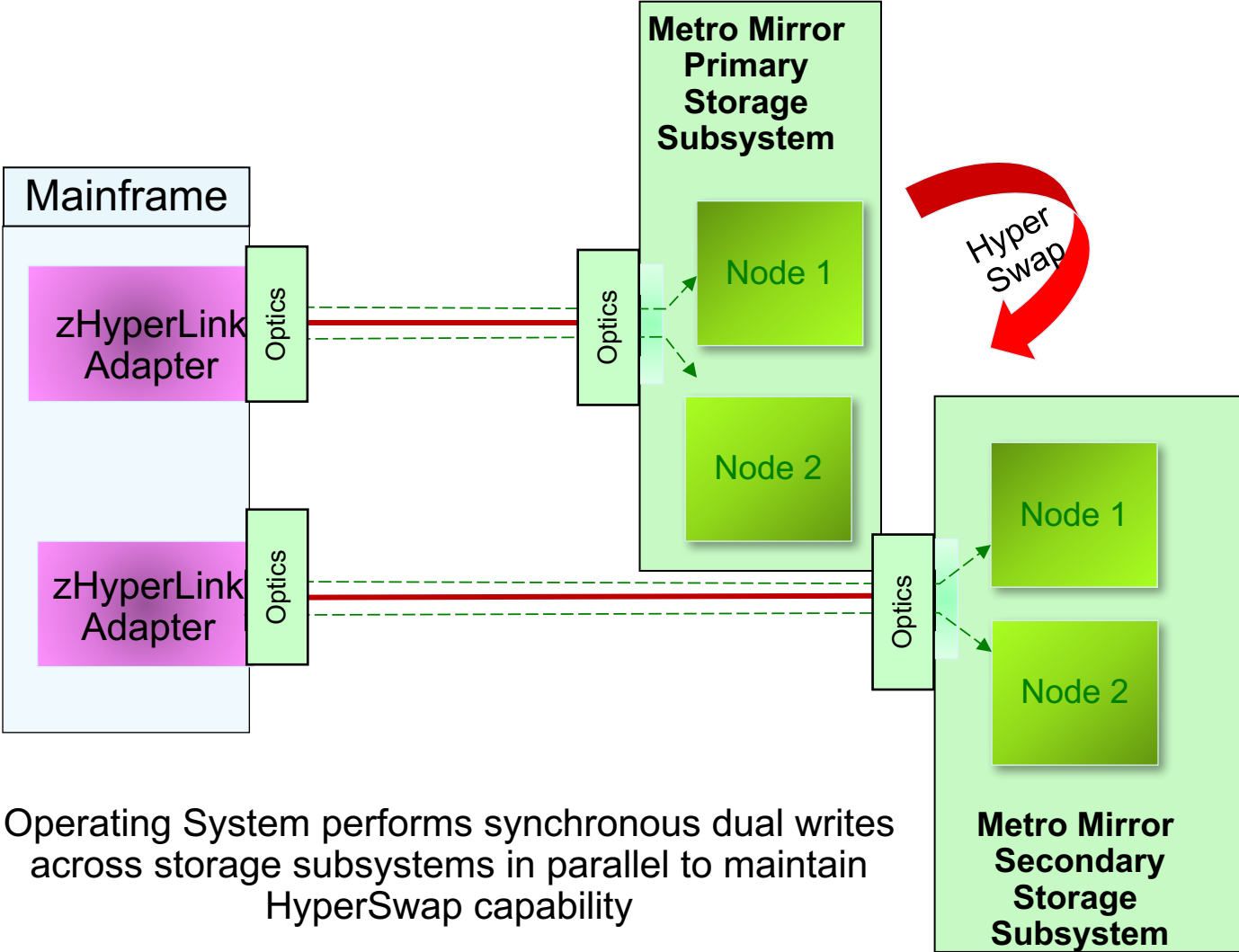
## zHyperlink exploitation: now and roadmap

- Now
  - Db2 sync I/O reads (batch and transactional)
  - IMS, CICS etc exploitation of zHyperlinks (via Db2)
- zHyperlink Statement of Direction (SOD)
  - VSAM Read Support (batch and transactions)
  - IMS, CICS etc exploitation of zHyperlinks (via VSAM)
  - Db2 z/OS log write support

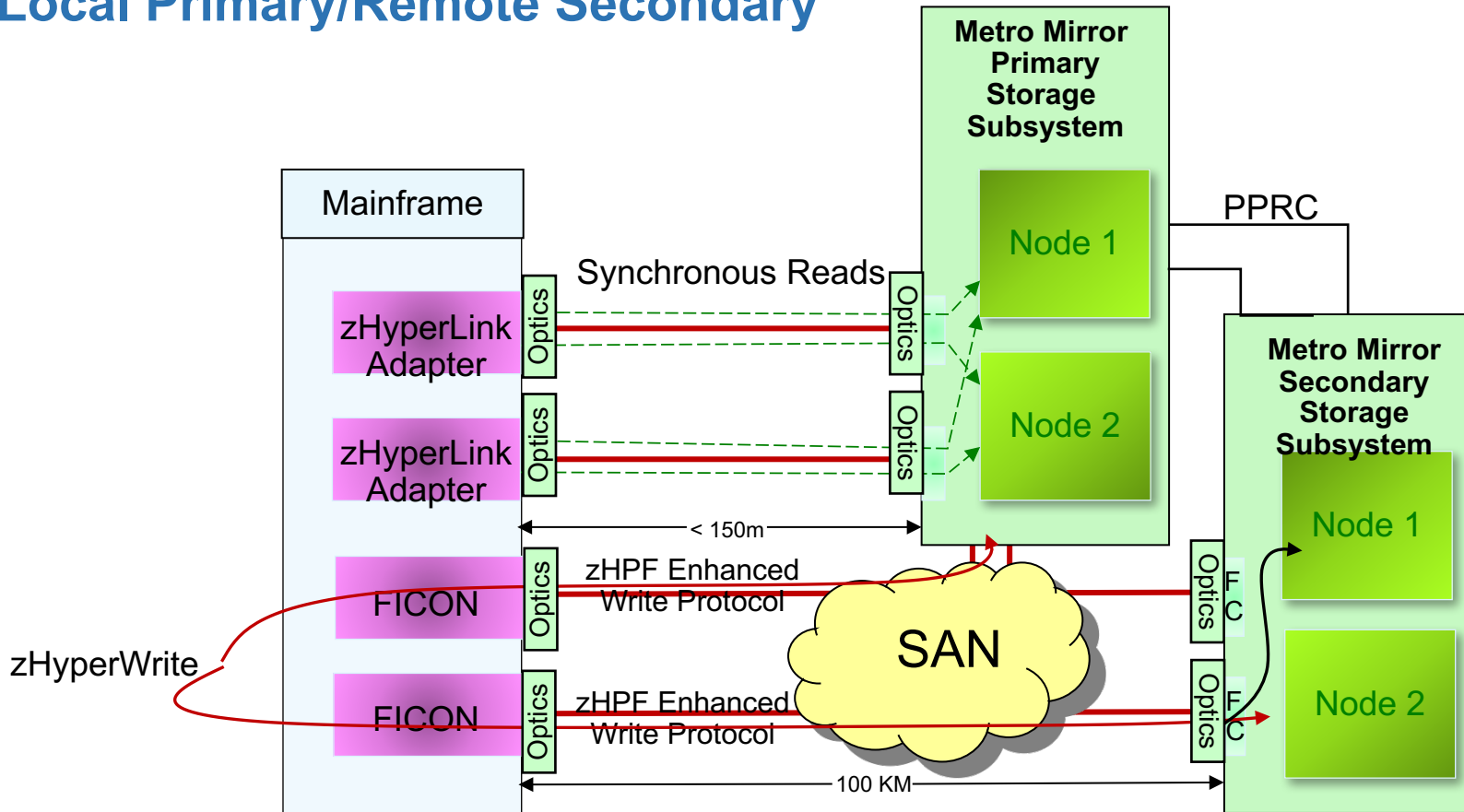
# Physical Connectivity



# Continuous Availability – Synchronous zHyperWrite



# Local Primary/Remote Secondary



Local Primary uses synchronous I/O for reads, zHPF with enhanced write protocols and zHyperWrite for writes at distance

---

## Technical Practicalities

- Z14 GA
- DS8880 R8.2 (model numbers 984, 985, 986 and 988)
- Distance between z14 and DS8k <150m
- z/OS v2.3 + APARs
- Db2 for z/OS v12 +APARs
- Approximate perf estimations: zBNA (dataset activity level)



# zHyperLink Analysis

---

# zHyperLink Analysis

Analysis should be approached from two angles:

- What the effect would be on workloads
- Which infrastructure could benefit
  - > For example, which data sets

There are various levels:

- Disk volume
- Data set
- DB2 application

These can all be rolled up

- For example, at the disk controller level

---

## Disk Volume Perspective

RMF SMF data establishes disk and cache behaviour

- 74-1 Device Activity Report
  - > I/O rates and response times **by system**
- 74-5 Cache Activity Report
  - > Cache behaviour **for all systems sharing**

Summarise at the controller, storage group, or volume level

- For example, DB2 storage group

Consider proportion of disk response time that is Connect

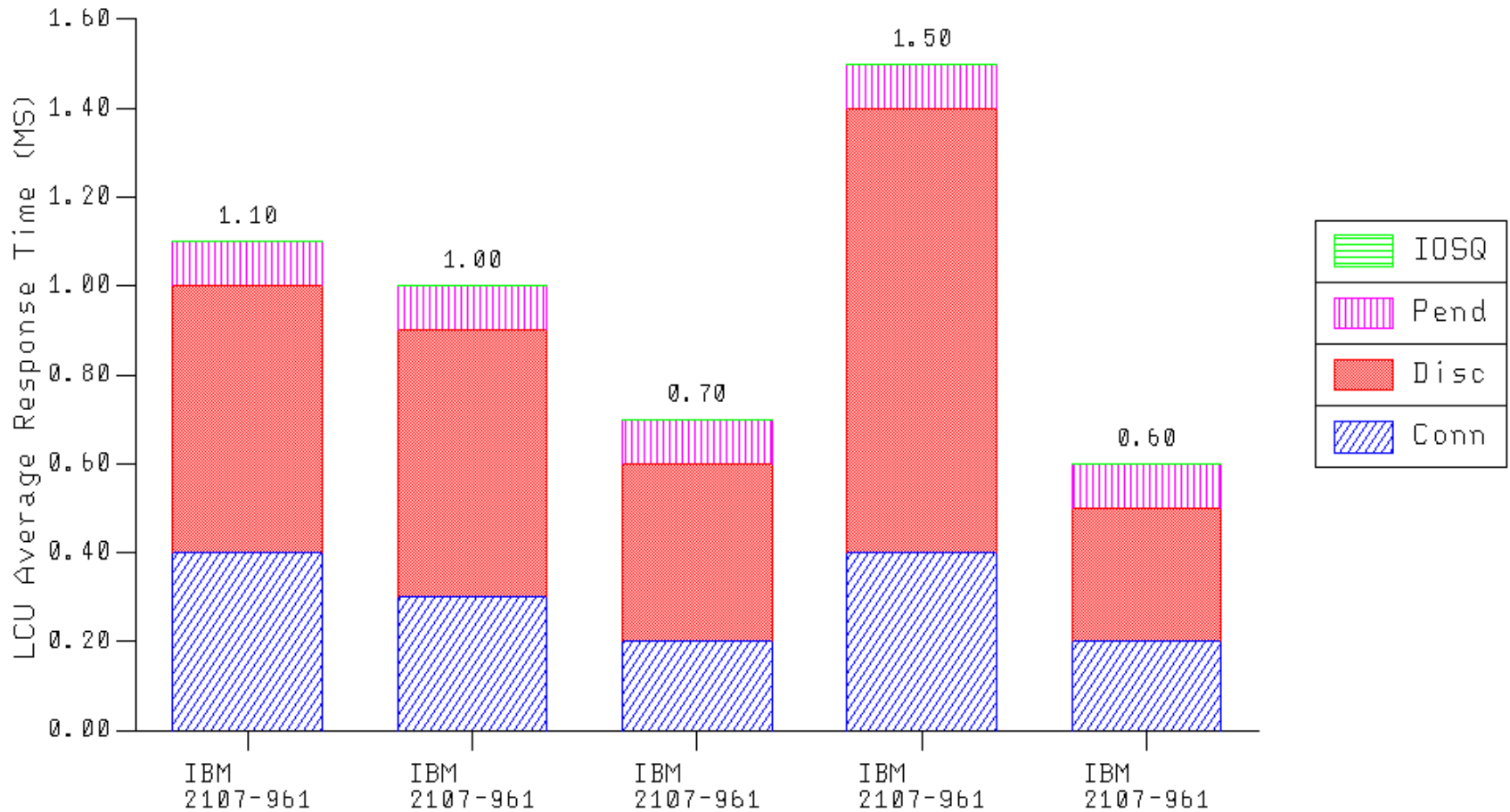
- Optimistic to assume it all disappears with zHyperLink

Depending on your goal, time of day might be useful

- Examples: Batch Window shortening, CICS Transaction Response Time

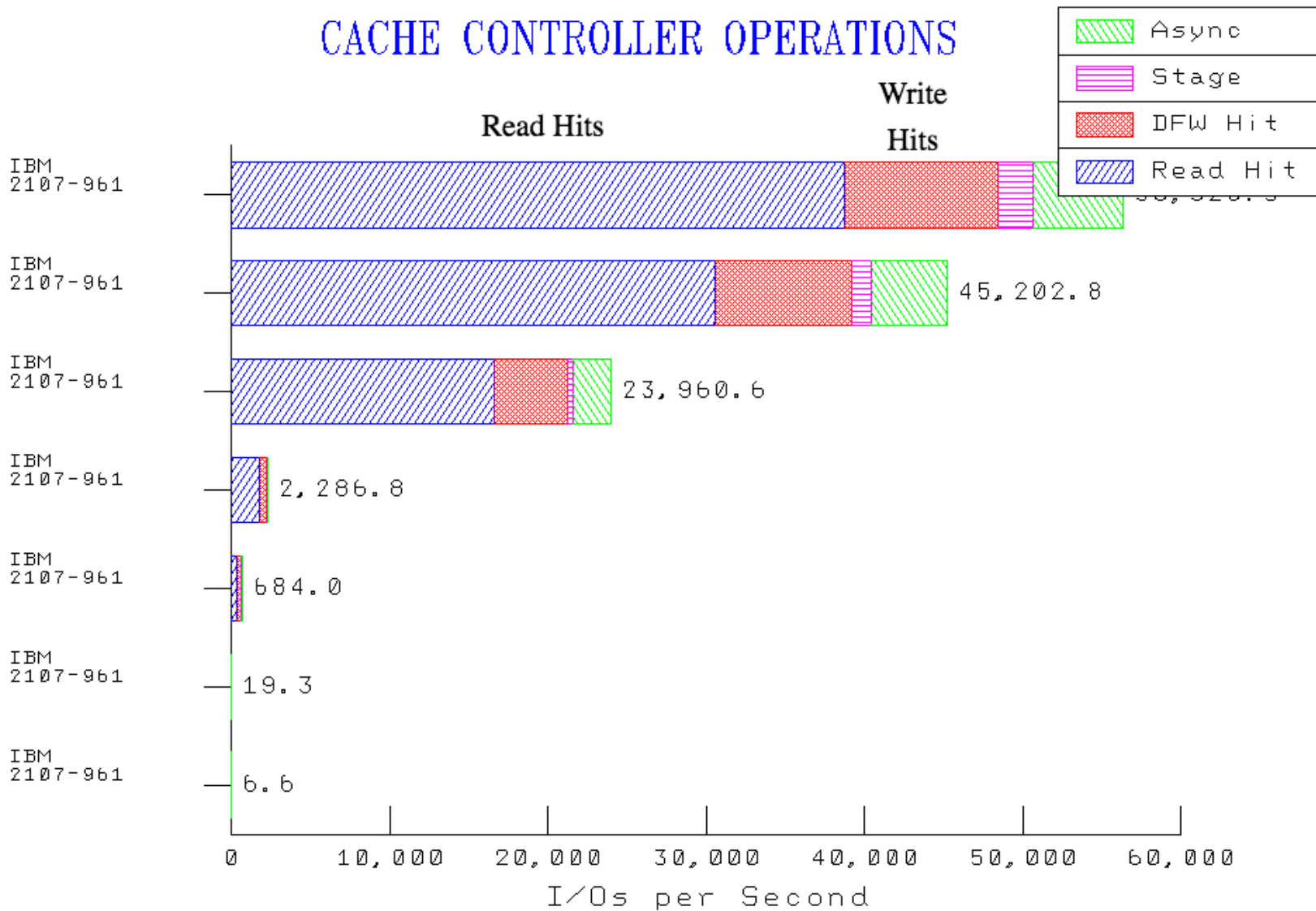
# Device Activity: Examine Connect Time And I/O Rates

## CONTROL UNIT RESPONSE TIME



# Cache Activity: Examine Cache Hit Percentages

## CACHE CONTROLLER OPERATIONS



---

# Data Set Perspective

Data set information is available at two levels

- Storage Class - SMF 42-5
- Data Set - SMF 42-6

Approaching the application perspective

- Data sets usually have some application relatedness

Look for relevant, high I/O rate, good cachiers

Again, response time can be broken down into the 4 buckets

- So Connect Time for e.g. a particular DB2 table space could be interesting

---

# DB2 Application Perspective

DB2 Accounting Trace (SMF 101) crucial

Availability of data varies by installation

SMF 101 breaks application time down into buckets

- Relevant buckets are in the Synchronous I/O Wait category:
  - > Database Read
  - > Log Write - per Statement Of Direction
- It also breaks down by application component:
  - > CICS transaction
  - > Batch job step
  - > ...
- Package-level breaks down further - by program

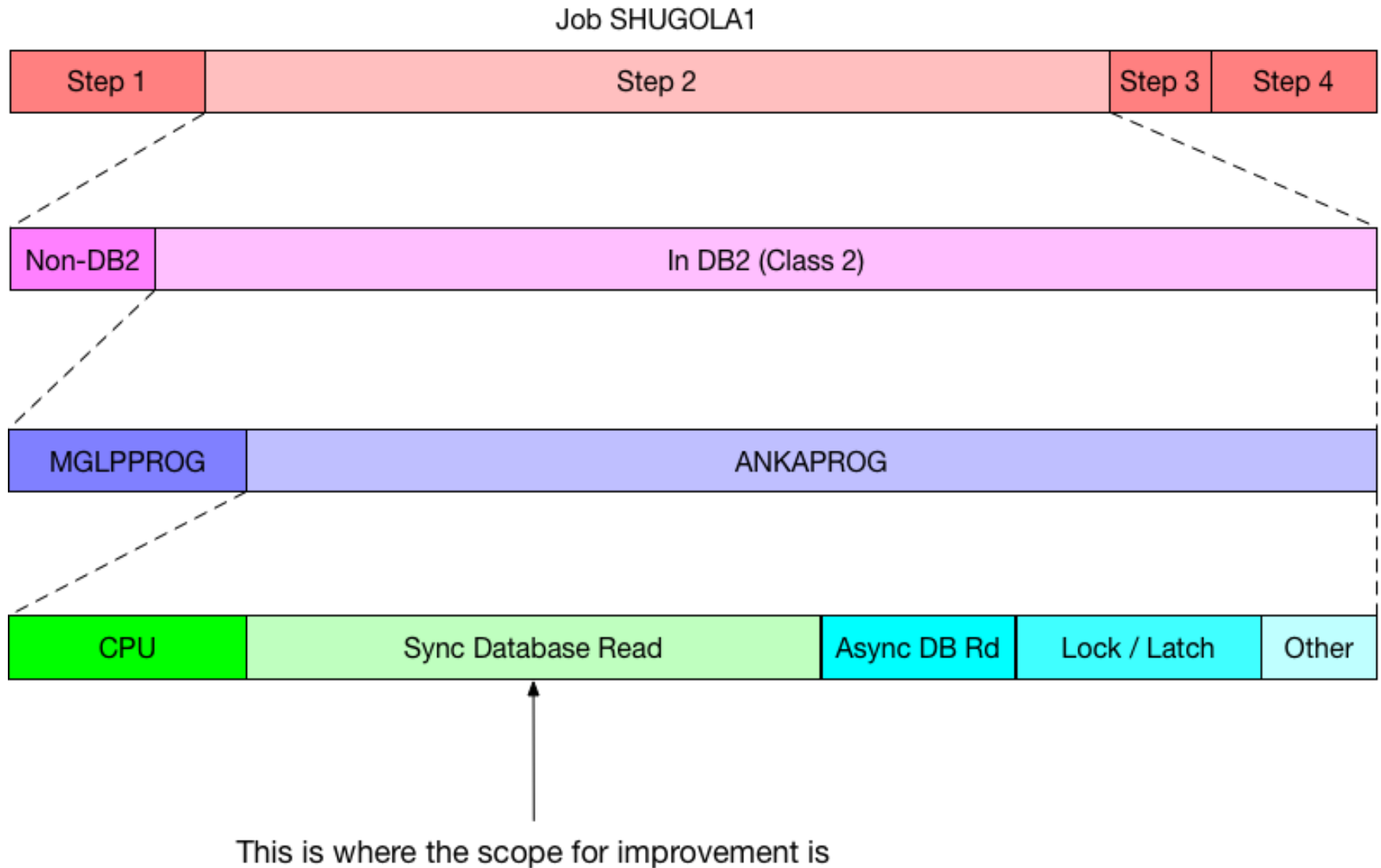
Caution: Not all time in the relevant buckets will disappear

- Almost all would if 100% cache hits

Placing in context is important

- Which applications do you want to improve?
- Could anything else speed them up?

# zHyperLink DB2 Example - A Batch Job





---

## Summary & next steps

- Planning the datacenter infrastructure:
  - Placing z14 and DS8k within 150m distance
  - Verifying HA/DR practices (if HA/DR solutions are used)
- Performance analysis:
  - Preliminary: zBNA
  - Collect 24h data – Db2 accounting traces, disk activity
  - Looking for DB2 sync I/O operations (I/O must be 100% disk cache hit ratio)

---

## Session feedback

Please submit your feedback online at:

<http://conferences.gse.org.uk/2017/feedback/LF>

Paper feedback forms also available from the Chair person

This session is LF

