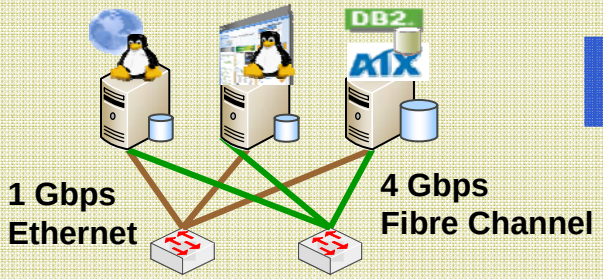
FCoCEE* Enterprise Data Center Use Cases

Dan Eisenhauer,
IBM FC over CEE Development

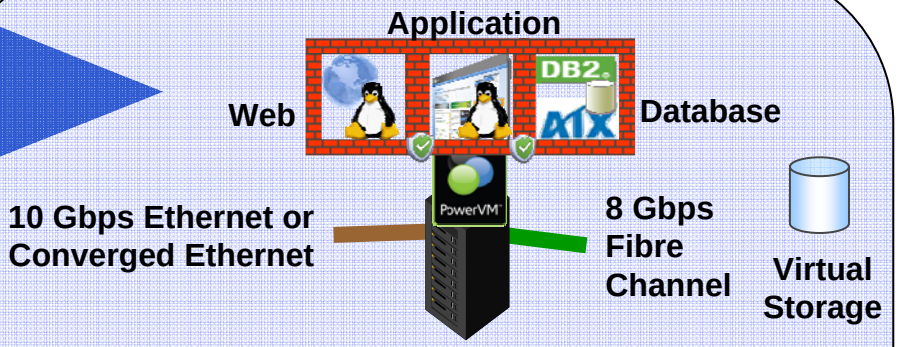
Renato Recio, DE,
IBM Data Center Networking CTO

***Fibre Channel over Convergence Enhanced Ethernet**

The Data Center is Undergoing Transition

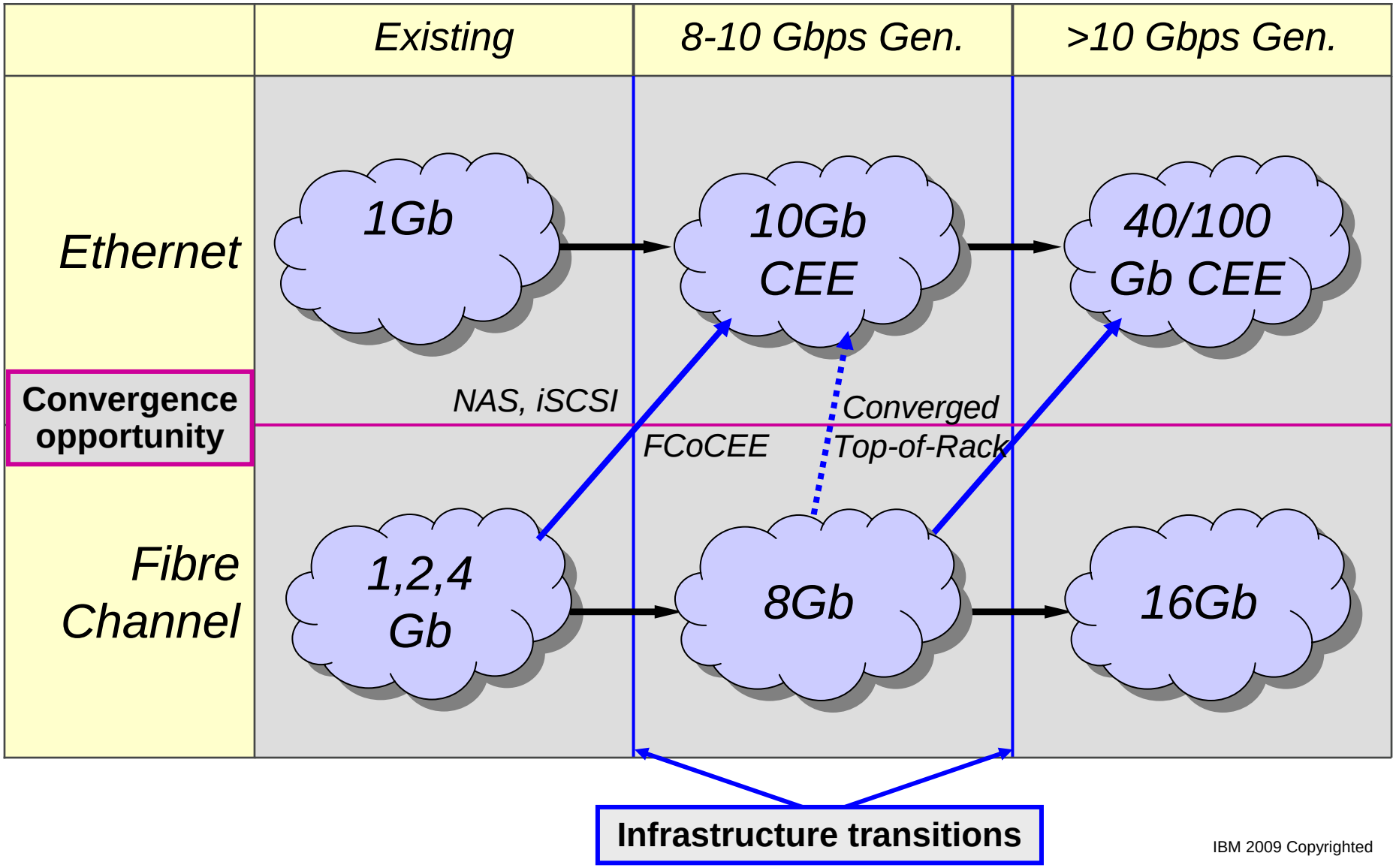


- **From a physical compute model that has:**
  - VMs held captive by low bandwidth links & manual network configuration
  - Few Virtual Machines (VMs) per server
  - Using lower bandwidth links
  - Sprawl of manually intensive, expert based IT management tools
  - Poor image life-cycle management tools
  - Physical appliance sprawl
- **Expense issues with this model:**
  - High capital expenses due to under utilized servers and multiple fabrics
  - High operational expenses due to manual administration of many management tools.

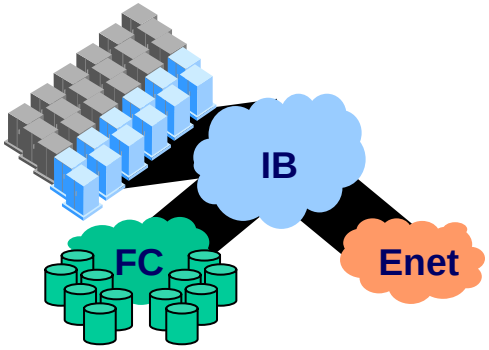


- **To a virtualized, Dynamic Infrastructure model that:**
  - Has simpler, **integrated & virtualized appliances and systems**
  - Has **many** highly utilized servers, each with **many VMs**; using virtual storage, networks & appliances
  - Is enabled by high bandwidth links and migrating to **virtual & converged fabrics**
  - Includes new, "faster time to value" Cloud Building Block (CBB) Multi-Rack acquisitions
- **The value of this new model is:**
  - Lower capital expenses through higher server, storage and network utilization.
  - Lower operational expenses through **automated & integrated management** that optimizes and automates Data Center infrastructure and delivers self configured services to the business.

Data Center Network Evolution

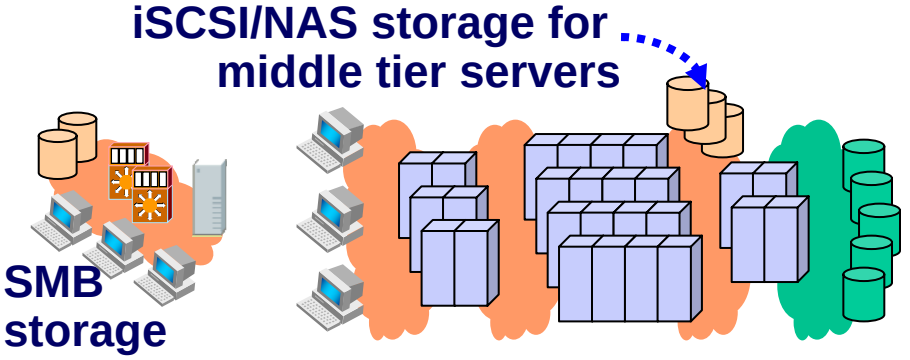


Fabric Convergence Options



High performance clusters

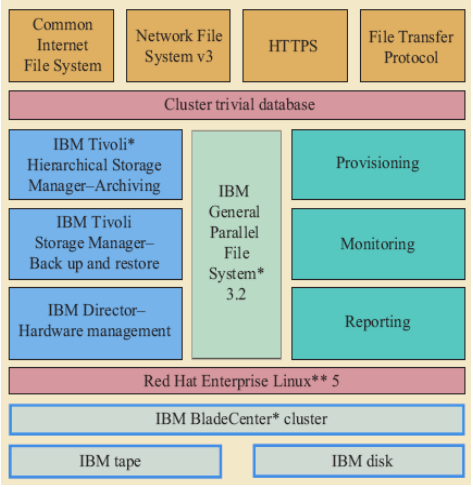
- IB: 2 → 4 → 8 GB/s and 150 → <100 ns
- Lack native storage & use slow gateways



SMB and Mid-tier Storage

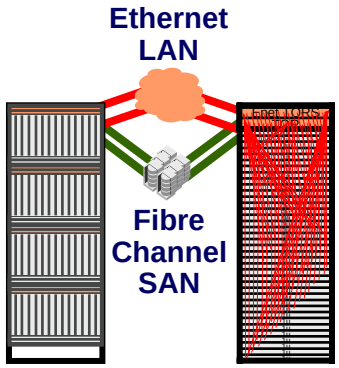
- Low $/Gbps (GigE)
- Growing and maturing, considering 10 GigE

Scale-out File System (SoFS)



- Clustered File System with Global Namespace
- Centrally managed
- Scales horizontally
- Integrated Information Lifecycle Mgt, HA & disaster recovery

Fibre Channel over Convergence Enhanced Ethernet



Converged switch

- Enables 50% CapEx reduction within the chassis (rack or BC).
- Lower power
- Improved RAS
- Plugs into existing DC infrastructure

FCoCEE Refresher

- **FC over Ethernet**
  - Layers FC frames directly over Ethernet (requires jumbo/mini-jumbo)

  - Replacing lowest level of FC with Ethernet

  T11 FC-BB-5 standard (Fibre Channel standards group).

- FC over Ethernet requires **convergence enhancements to Ethernet**:
  - To carry maximum sized FC packets

    Requires Ethernet Jumbo (or mini-jumbo) frames
  - To provide FC like no-drop behavior in face of congestion

    Requires new "**Priority-based Flow Control**" protocol
  - To control traffic interferences:

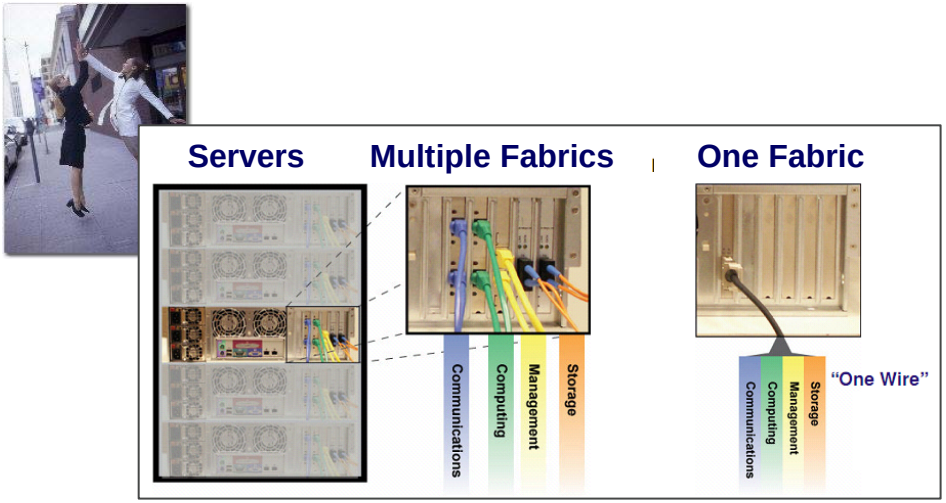
    Requires new "**Enhanced Transmission Selection**" protocol
  - To detect both ends of the wire support the above

    Requires new "**Data Center Bridging eXchange**" protocol

  CEE Authors was formed to facilitate/accelerate definition & adoption of above.

  CEE Authors submitted proposals for above protocols to IEEE 802.1.

Fabric Convergence → Value Proposition vs Barriers



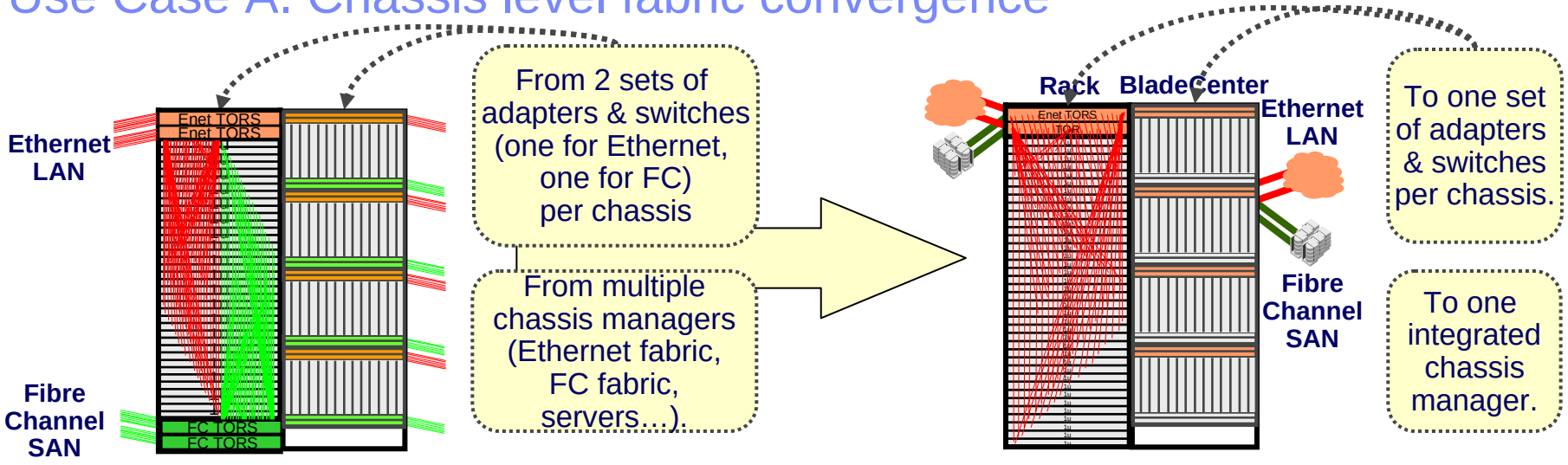
Value Proposition

- Lower Cost, Lower Power
  - Less adapters, cables and switches
- Improved RAS
  - Reduced failure points, time, misconnections, bumping, …
- Simpler Management
  - Single physical fabric to manage.

Barriers

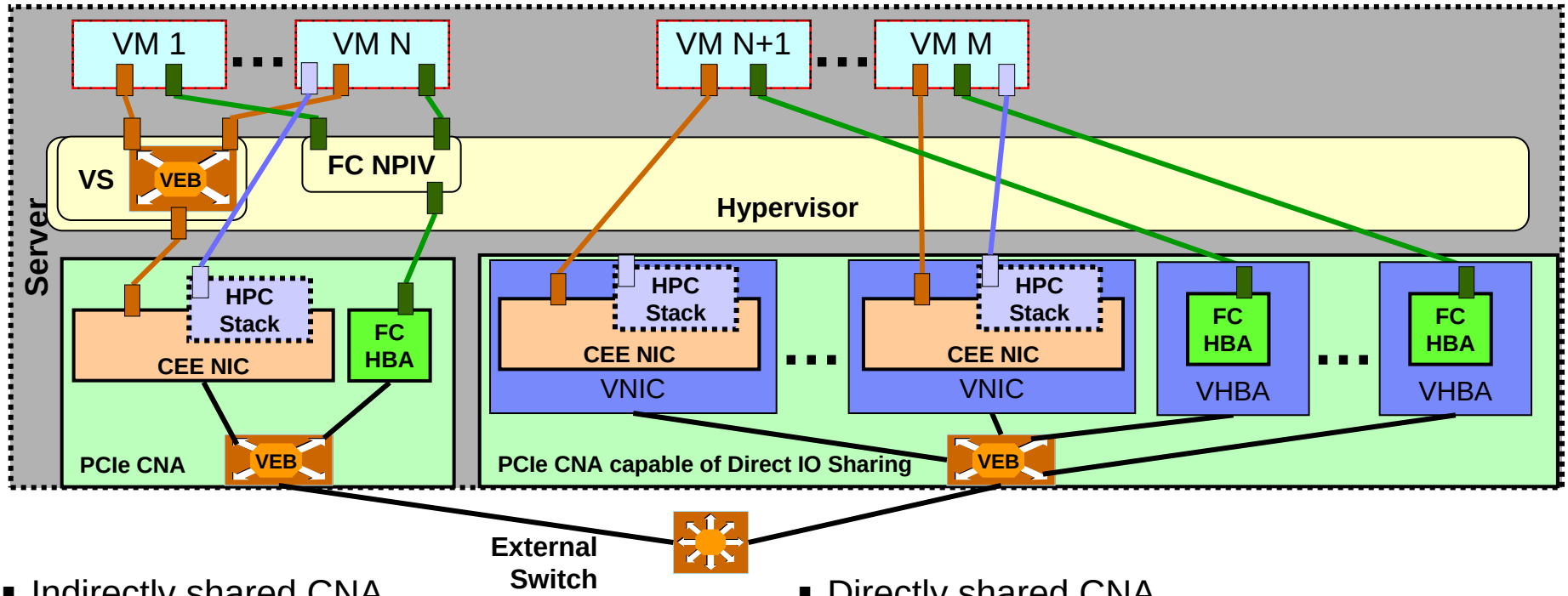
- Customer organizational
  - Server, Network, Storage silos
- Operational management
  - Quality of service control
  - Security and access control
- Maturity
  - Standards, technology, management
  - Resiliency and robustness

Use Case A. Chassis level fabric convergence



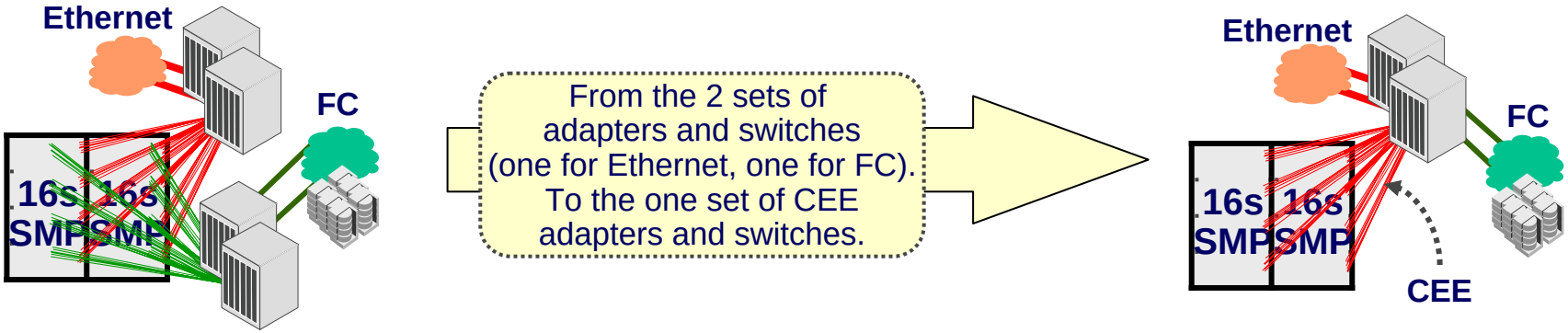
- Today's rack optimized servers use:
  - Fibre Channel (FC) adapters to attach an FC Top-of-Rack (TOR) switch, which connects to the FC Data Center (DC) fabric.
  - Ethernet adapters to attach an Ethernet Top-of-Rack (TOR) switch, which connects to the Ethernet DC fabric.
- Similarly, today's blade servers use:
  - FC adapters to connect to FC DC fabric through an integrated blade FC switch.
  - Ethernet adapters to connect to Ethernet DC fabric through an integrated Ethernet switch.
- Fabric convergence is contained to a single, Blade Server or a rack chassis.
- Within the (rack or blade) chassis:
  - A Converged Network Adapter (CNA) is used to connect to a Fibre Channel over Ethernet (FCoE) Forwarder (FCF) enabled switch.
  - The FCF uses Ethernet and FC to connect to the DC's existing Ethernet and FCfabrics.
  - This approach eliminates FC adapters, cables and switches within the chassis.
- Clients may want to start in dev/test mode (e.g. to define bandwidth allocations needed thru system's lifecycle).

Converged Network Adapter (CNA) Examples



- Indirectly shared CNA
  - Shared through extensions to server's existing virtualization infrastructure.
  - HBA shared through server's software Fibre Channel N_Port Identifier Vitualization mechanism.
  - NIC shared through server's software Virtual Ethernet Bridge (VEB) mechanism.
  - Adapter includes additional CEE capable VEB to share NIC & HBA paths.
- Directly shared CNA
  - Shared through adapter virtualization infrastructure.
  - HBA shared through server's software Fibre Channel N_Port Identifier Vitualization mechanism.
  - NIC shared through adapter VEB mechanism.
  - Adapter includes additional CEE capable VEB to share NIC & HBA paths.

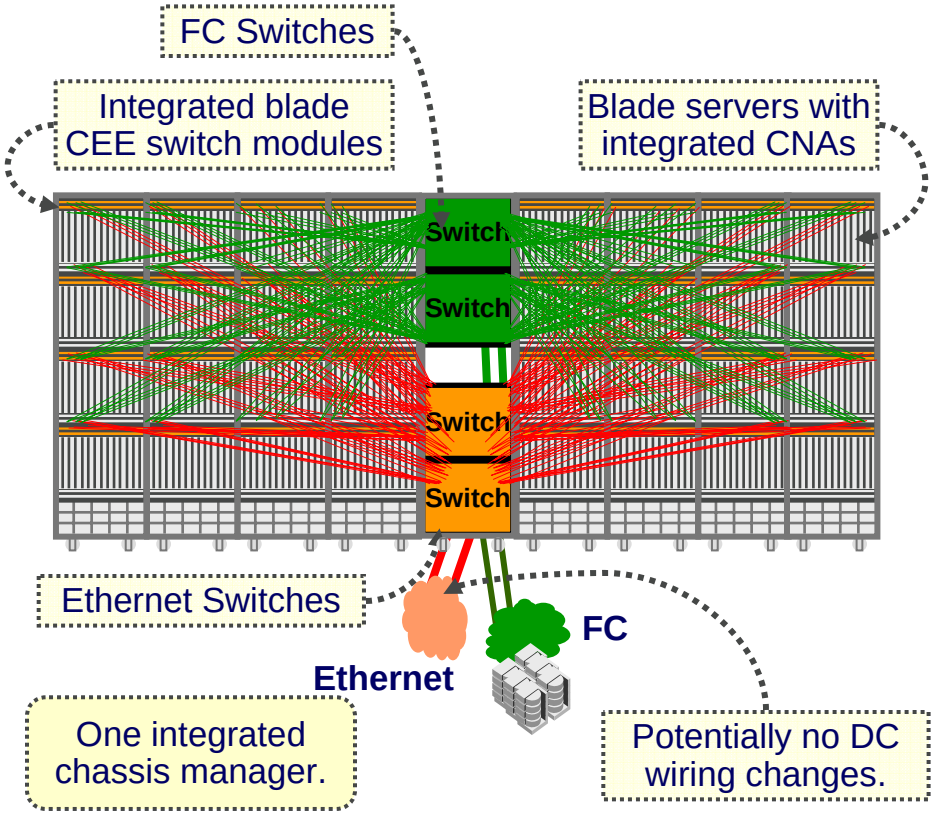
Use Case B. Large SMP level fabric convergence



- Large SMPs use a server virtualization to consolidate many Virtual Machines (VMs).
  - An integrated Virtual Ethernet Bridge (VEB) is used for communication with external systems and between local VMs.
  - The VEB essentially serves as an access layer switch, which eliminates the need for a standalone aggregation switch, such as a TOR switch.
- Large SMPs today use multiple adapters to connect to the data center's existing Ethernet and FC fabrics.
- For Use Case B, the large SMP uses FC over CEE enabled Converged Network Adapters CNAs to attach directly to a converged Modular switch.
  - This converged switch is used to connect into the data center's existing Ethernet and FC infrastructures.
- Clients may want to start in dev/test mode, for same reasons as Use Case A.

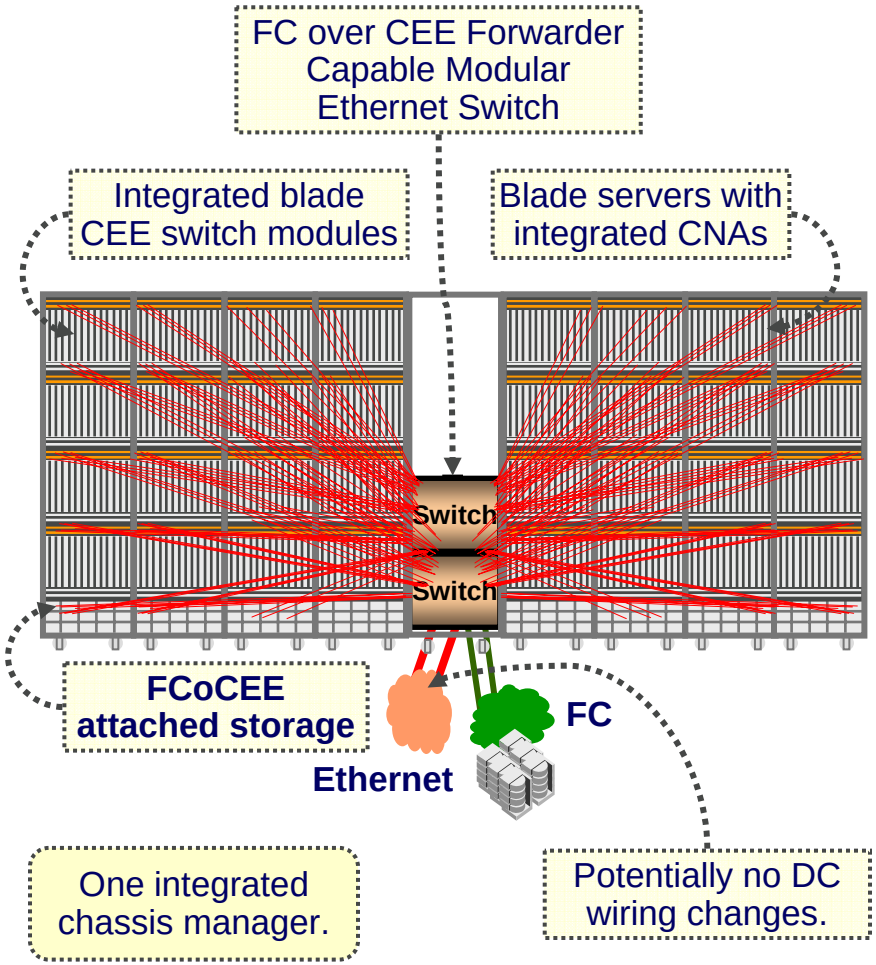
Use Case C. Chassis level Cloud Building Block fabric convergence

Cloud Building Block (CBB) Definition: A DC granular unit of scale, which includes: servers, storage, networking, virtualization infrastructure and associated platform / service management.



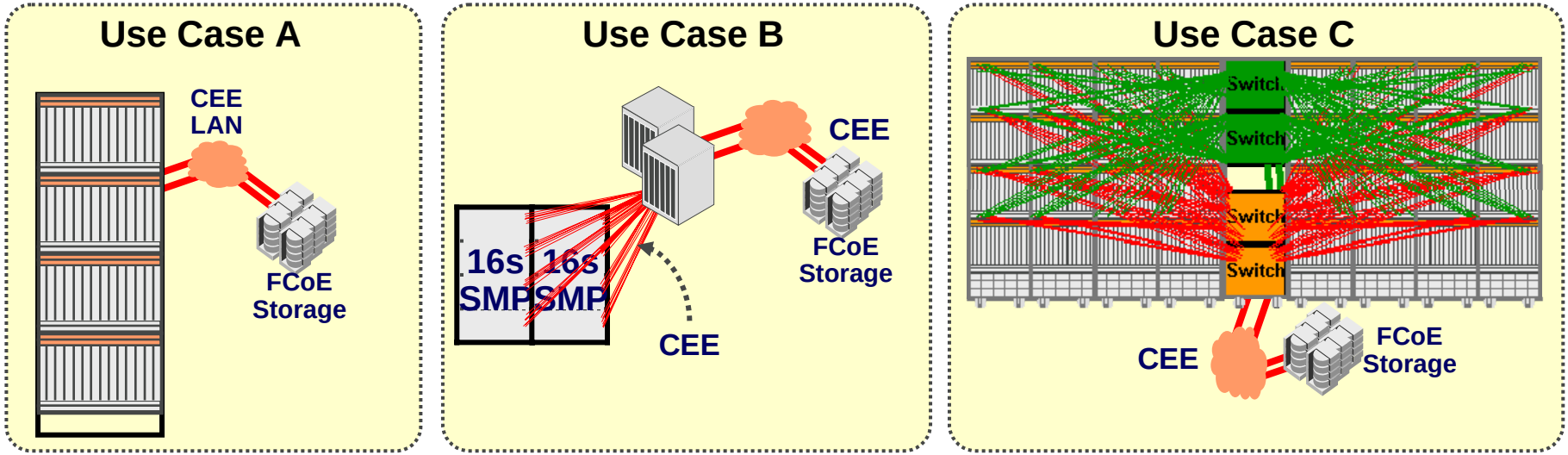
- This use case extends Use Case A & B chassis level convergence to a CBB level.
  - A converged fabric is used within each chassis.
  - But the CBB uses two fabrics: Ethernet and FC.
- This use case provides a significant cost advantage over non-converged CBBs, without converging the modular switches used to connect the CBB to the data center's existing Ethernet and FC infrastructures.
- For the Blade example shown here:
  - Each server uses a dual-ported CNA to connect through to an integrated CEE Blade switch.
  - These integrated CEE switches connect to the CBB's modular Ethernet switches thru Ethernet and the modular FC switches thru FC.
- Use case C has the potential of saving a considerable amount of hardware.
  - For example the 8 rack configuration shown has 4 blade chassis per rack and 14 server blades per chassis.
  - This use case may eliminate 448 FC adapters and 64 FC integrated blade switches per CBB.

Use Case D. Converged Cloud Building Block



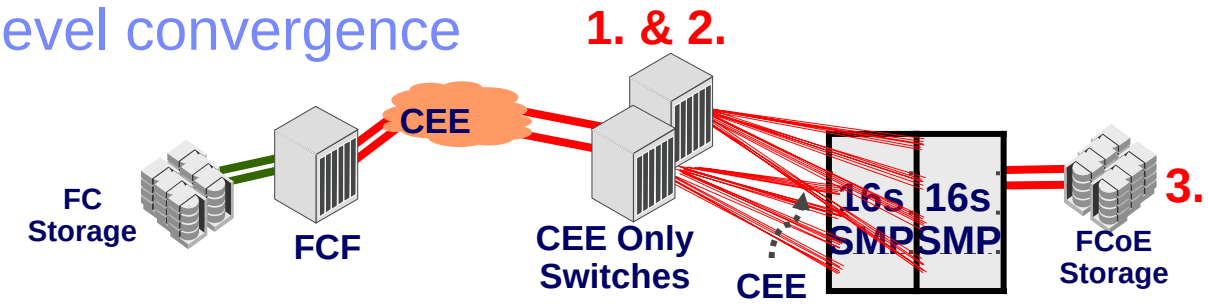
- This use case fully converges the fabric within the CBB, but at the data center level there are still two separate fabrics: Ethernet and FC.
  - Like use case C, this use case eliminates all the adapters & chassis switches.
  - Plus, within the CBB, it also eliminates the modular FC switches.
- This use case provides the full value proposition of FC over Ethernet convergence within the CBB, without having to rip and replace the data center's existing Ethernet and FC infrastructures.
- Like use case C, the DC's fabric management remains the same, but within the CBB an integrated manager can be used to manage the CBB.

Use Case E. Storage attachment to converged fabrics



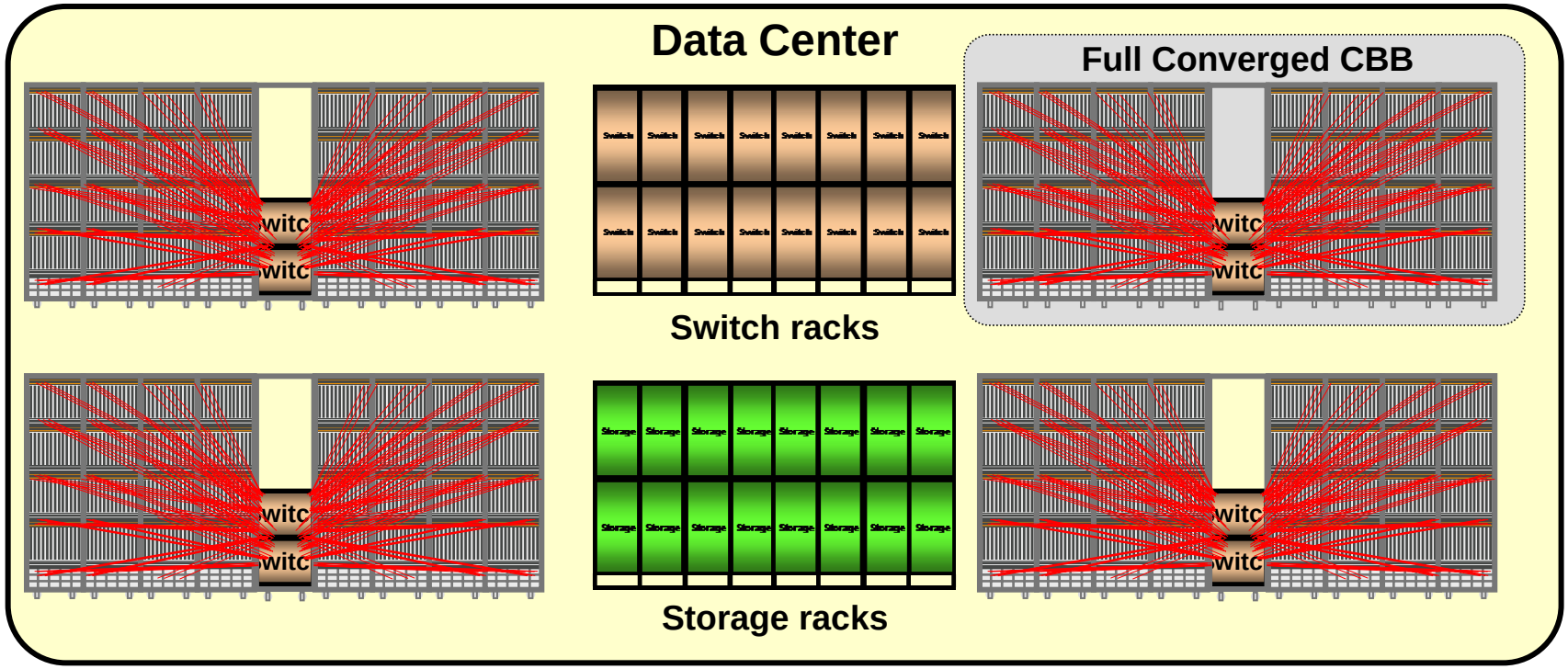
- Each of the uses cases covered so far can take an additional evolutionary step towards full fabric convergence by using FCoCEE enabled storage servers.
  - Each storage server use a FC over CEE capable CNAs to connect into a CEE capable fabric.
- However, migrating from FC attached storage to FCoCEE attached storage doesn't yield the same degree of savings as use cases A-D, because it just replaces an FC adapter with an FCoCEE capable CNA.
- Additionally, in DC environments this use case requires the fabric from the server to the storage to be CEE based (vs plain Ethernet); otherwise the issues covered in "A Case for Convergence Enhanced Ethernet: Requirements and Applications" can surface.
  - Also, as recommended in the above paper, IETF TRILL should be used to simplify active-active configurations and, for larger fabrics, IEEE 802.1Qau Congestion Management should be used.

Use Case F. FICON level convergence



- FICON environments have requirements above most open FC configurations. For example:
  1. FICON requires FC Class 2 for faster error detection.
     - Intermediate “CEE only switches” residing between a server and an FC Forwarder (FCF) would not respond in the event of errors to Class 2 frames as legacy FC switches would.
     - Class 2 frames, just like all other FCoCEE frames, would be treated just as any other Ethernet frame. Any frames that could not be delivered due to congestion or offline destinations would eventually be dropped by the CEE switch, not busied (F_BSY) or rejected (F_RJT) as they would by an FC aware switch.
     - FC-BB-5 makes note of this potential lack of Class 2 functionality, but leaves the solution to managing the supported configurations and simply not allowing intermediate CEE only switches.
  2. FICON requires Link Incident Detection and Reporting.
     - The FC Back Bone 5 specification (FC-BB-5) defined a Link Error Status Block definition for Ethernet port statistics to satisfy the reporting requirement. FC-BB-5 has also addressed detection of lost links by periodic Link Keep Alive and Advertisement messages.
     - However, the timeliness of the detection (the period between Keep Alives can be large) coupled with the above mentioned lack of Class 2 responses this could be problematic for FICON, further prohibiting the use of intermediate “CEE only switches”.
  3. FICON environments require support of direct server to storage configurations (e.g. no switch).
     - FCoE as defined in FC-BB-5 does not support this (i.e. FCoE Forwarder is always required).
     - The FC-BB-6 project proposes to add this capability.

Use Case G. Data Center level convergence



- This use case uses CEE and FCoCEE for all servers and storage.
- Though this use case can be achieved directly, a more prudent approach is to get to data center level convergence by progressing through either use case (C) or (D).
  - This can be achieved by deploying FCoCEE capable switches throughout the data center, but only using them as Ethernet switches initially.
  - As the organizational silo and fabric management issues described in the paper are addressed, the Data Center FCoCEE switches can either use FC line cards to attach FC storage or simply attach FCoCEE capable storage directly.

How does FCoCEE Compare?

		A.1	A.2	B	C	D	E	F	G
What native Fibre Channel equipment is eliminated within each use case installation?	Server Adapters	✓	✓	✓	✓	✓	✓		✓
	Server Cables	✓	✓	✓	✓	✓	✓		✓
	Access Switches	✓	✓	✓		✓	✓		✓
	Cloud Cell Modular Switches				✓	✓	✓		✓
	Storage Adapters						✓		✓
	Storage cables						✓		✓
	System z Adapters							✓	✓
	System z cables							✓	✓
	Data Center Modular Switches								✓
What is the fabric contention scope for each use case?	Adapter to Access	✓	✓	✓	✓	✓	✓	✓	✓
	Access to Modular					✓			✓
	Full DC								✓
Is existing DC infrastructure protected?		✓	✓	✓	✓	✓	✓	✓	

Conclusion

- With the large install base of FC based storage in the enterprise datacenter, FCoCEE offers a fabric convergence solution that aims to protect FC storage investment while providing a consolidated network for clustering, storage and IP/Ethernet traffic.

- As FCoCEE matures & meets Enterprise performance, reliability & quality requirements, we expect it will play well in large enterprises.

- The use cases covered earlier provide an evolutionary model for FCoCEE converged fabrics.
  - They enable clients to obtain the convergence value proposition, through evolutionary steps that help mitigate convergence risks.