Towards a Noah’s Ark for the Upcoming Data Deluge

Alessandro Bassi    Spyros Denazis
Pierpaolo Giacomin

Hitachi Sophia Antipolis Lab, 06560 Valbonne, France
Alessandro.Bassi@Hitachi-eu.com

VLDB 2007
1. Intro

2. State of the art

3. INCA: The Concepts
   - Introducing the Intelligent Network Caching Architecture
   - The Architecture

4. INCA: The Implementation
   - The INCA Transport Protocol: HITP
   - The INCA Control Protocol: HICP
   - How to retrieve the stored data - the Middleware Layer

5. Outro

Outline

Towards a Noah’s Ark for the Upcoming Data Deluge
Outline

1. Intro
2. State of the art
3. INCA: The Concepts
   - Introducing the Intelligent Network Caching Architecture
   - The Architecture
4. INCA: The Implementation
   - The INCA Transport Protocol: HITP
   - The INCA Control Protocol: HICP
   - How to retrieve the stored data - the Middleware Layer
5. Outro
Introducing the Exaflood and his Impressive Numbers

- **Scientific World**
  - Large Hadron Collider (LHC) (15 Petabytes/year from Spring 2008)
Introducing the Exaflood and his Impressive Numbers

- **Scientific World**
  - Large Hadron Collider (LHC) (15 Petabytes/year from Spring 2008)

- **Financial Institutions Global Data caching requirements**
  - Each single large Enterprise (Handle 5 Exabytes/year from 2015)
  - Enterprise Data (Creation of 200 Exabytes/year from 2015)
Introducing the Exaflood and his Impressive Numbers

- **Scientific World**
  - Large Hadron Collider (LHC) (15 Petabytes/year from Spring 2008)

- **Financial Institutions Global Data caching requirements**
  - Each single large Entreprise (Handle 5 Exabytes/year from 2015)
  - Entreprise Data (Creation of 200 Exabytes/year from 2015)

- **The Consumer World (Holidays Pictures, Video)**
The technical Goals for a sustainable data management:

Sustainable Data Management?

- User
- Data
- Data Exchange
- Data Provider

Towards a Noah's Ark for the Upcoming Data Deluge
The technical Goals for a sustainable data management:

Sustainable Data Management!

- User
- Data
- Data Exchange
- Data Provider
The technical Goals for a sustainable data management:

Scalability

- User
- Data
- Data Exchange
- Data Provider
The technical Goals for a sustainable data management:

- Compatibility
  - Traditional Network Node
  - User
  - Data
  - Data Exchange
  - Data Provider
The technical Goals for a sustainable data management:

- Security Management

- Traditional Network Node
- User
- Data
- Data Exchange
- Data Provider
- Security Management
The technical Goals for a sustainable data management:

Properties Management

- Traditional Network Node
- User
- Data
- Data Exchange
- Data Provider
- Properties Management
- Security Management

Towards a Noah’s Ark for the Upcoming Data Deluge
The technical Goals for a sustainable data management:

- Genericity
- Generic Interface
- User
- Applications
- Data
- Data Exchange
- Data Provider
The technical Goals for a sustainable data management:

Quality of Experience

- Generic Interface
- User
- Happy User
- Data
- Data Exchange
- Data Provider
Outline

1. Intro
2. State of the art
3. INCA: The Concepts
   - Introducing the Intelligent Network Caching Architecture
   - The Architecture
4. INCA: The Implementation
   - The INCA Transport Protocol: HITP
   - The INCA Control Protocol: HICP
   - How to retrieve the stored data - the Middleware Layer
5. Outro

Towards a Noah’s Ark for the Upcoming Data Deluge
State of the art

- Akamai and Tamanoir as Hierarchical Webcache
State of the art

- Akamai and Tamanoir as Hierarchical Webcache ⇒ HTTP and Websites
State of the art

- Akamai and Tamanoir as Hierarchical Webcache ⇒ HTTP and Websites
- Storage Resource Broker (San Diego Supercomputing Center Middleware)
State of the art

- Akamai and Tamanoir as Hierarchical Webcache ⇒ HTTP and Websites
- Storage Resource Broker (San Diego Supercomputing Center Middleware) ⇒ Scalability because of Metadata Catalog

Towards a Noah’s Ark for the Upcoming Data Deluge
State of the art

- Akamai and Tamanoir as Hierarchical Webcache ⇒ HTTP and Websites
- Storage Resource Broker (San Diego Supercomputing Center Middleware) ⇒ Scalability because of Metadata Catalog
- Internet Backplane Protocol (IBP: Logistical Storage)
State of the art

- Akamai and Tamanoir as Hierarchical Webcache ⇒ HTTP and Websites
- Storage Resource Broker (San Diego Supercomputing Center Middleware) ⇒ Scalability because of Metadata Catalog
Outline

1. Intro
2. State of the art
3. INCA: The Concepts
   - Introducing the Intelligent Network Caching Architecture
   - The Architecture
4. INCA: The Implementation
   - The INCA Transport Protocol: HITP
   - The INCA Control Protocol: HICP
   - How to retrieve the stored data - the Middleware Layer
5. Outro
Introducing the Intelligent Network Caching Architecture

The Architecture

Bringing a copernican revolution to network and storage

INCA is Tight Coupling between Networks and Storage.
Exposition of Storage Capacity
Exposition of Storage Capacity
Introducing the Intelligent Network Caching Architecture

**The Architecture**

**INCA Layers**

- **INTERFACE** – API towards application and services
- **DATA MANAGEMENT LAYER** – Middleware to optimise the positioning of data in the network.
- **STORAGE PROTOCOL LAYER** – (Standardised) protocols to transport data and control operations on the HW.
- **STORAGE CONTROL LAYER** – SW (or HW) that controls racks of disks, optimizing performance and security.
How the data should be stored - Storage Protocol Layer

INCA addons

- HITP (Storage Transport Protocol)

Towards a Noah’s Ark for the Upcoming Data Deluge
How the data should be stored - Storage Protocol Layer

INCA addons

- HITP (Storage Transport Protocol)
- HICP (Storage Control Protocol, it controls HITP)
How the data should be stored - Storage Protocol Layer

INCA addons

- HITP (Storage Transport Protocol)
- HICP (Storage Control Protocol, it controls HITP)
- INCA Middleware (The INCA Intelligence, it speaks HICP)
Outline

1. Intro
2. State of the art
3. INCA: The Concepts
   - Introducing the Intelligent Network Caching Architecture
   - The Architecture
4. INCA: The Implementation
   - The INCA Transport Protocol: HITP
   - The INCA Control Protocol: HICP
   - How to retrieve the stored data - the Middleware Layer
5. Outro
Introducing the implementation

Restrictions

The current Internet as a real-case test-bed.
Introducing the implementation

Restrictions
The current Internet as a real-case test-bed.

HITP Aim
To transport huge volumes (Terabytes) of data in an efficient way.
Introducing the implementation

Restrictions
The current Internet as a real-case test-bed.

HITP Aim
To transport huge volumes (Terabytes) of data in an efficient way.

HICP Aim
To control HITP w/o being too intrusive.
HITP Key Properties

- Minimize Overhead
- Direct Storage Access/Addressing
- Deal with unordered packets (Connectionless??)
- Easily Proxable (Connectionless!!!)
- Deal with many peers (Connectionless!!)

HITP Today Constraints
- Over IP to be routable by INCA-Unaware nodes (a necessary evil)
- Simple header to fast analyzation

Towards a Noah’s Ark for the Upcoming Data Deluge
HITP Key Properties

- Minimize Overhead
HITP Key Properties

- Minimize Overhead
- Direct Storage Access/Addressing
HITP Key Properties

- Minimize Overhead
- Direct Storage Access/Addressing
- Deal with unordered packets (Connectionless??)
HITP Key Properties

- Minimize Overhead
- Direct Storage Access/Addressing
- Deal with unordered packets (Connectionless??)
- Easily Proxable (Connectionless?!)
HITP Key Properties

- Minimize Overhead
- Direct Storage Access/Addressing
- Deal with unordered packets (Connectionless??)
- Easily Proxable (Connectionless?!)  
- Deal with many peers (Connectionless!!)
HITP Key Properties

- Minimize Overhead
- Direct Storage Access/Addressing
- Deal with unordered packets (Connectionless??)
- Easily Proxable (Connectionless?!
- Deal with many peers (Connectionless!!)

HITP Today Constraints
HITP Key Properties

- Minimize Overhead
- Direct Storage Access/Addressing
- Deal with unordered packets (Connectionless??)
- Easily Proxable (Connectionless?!) 
- Deal with many peers (Connectionless!!)

HITP Today Constraints

- Over IP to be routable by INCA-Unaware nodes (a necessary evil)
HITP Key Properties

- Minimize Overhead
- Direct Storage Access/Addressing
- Deal with unordered packets (Connectionless??)
- Easily Proxable (Connectionless?!)  
- Deal with many peers (Connectionless!!)

HITP Today Constraints

- Over IP to be routable by INCA-Unaware nodes (a necessary evil)
- Simple header to fast analyzation
The Channels

Input Channel (ichannel) - identified by to resource ID (dst port)
Output Channel (ochannel) - identified by remote resource ID and remote IP address (dst ip, dst port)
The Channels

- Input Channel (ichannel) - identified by to resource ID (dst port)
The Channels

- Input Channel (ichannel) - identified by resource ID (dst port)
- Output Channel (ochannel) - identified by remote resource ID and remote IP address (dst ip, dst port)
Input Channel Properties

- View of a Writable Storage Medium (sync/async)
- In charge for Acknowledgement
- In charge for Packet Dropping
- In charge for Packet Filtering
- Data Integrity Aware

Towards a Noah’s Ark for the Upcoming Data Deluge
Input Channel Properties

- View of a Writable Storage Medium (sync/async)
Input Channel Properties

- View of a Writable Storage Medium (sync/async)
- In charge for Acknowledgement
Input Channel Properties

- View of a Writable Storage Medium (sync/async)
- In charge for Acknowledgement
- In charge for Packet Dropping
Input Channel Properties

- View of a Writable Storage Medium (sync/async)
- In charge for Acknowledgement
- In charge for Packet Dropping
- In charge for Packet Filtering

Towards a Noah’s Ark for the Upcoming Data Deluge
Input Channel Properties

- View of a Writable Storage Medium (sync/async)
- In charge for Acknowledgement
- In charge for Packet Dropping
- In charge for Packet Filtering
- Data Integrity Aware
Output Channel Properties
Output Channel Properties

- View of a Readable Storage Medium (sync/async)
Output Channel Properties

- View of a Readable Storage Medium (sync/async)
- In charge for Avoiding Network Congestion
Output Channel Properties

- View of a Readable Storage Medium (sync/async)
- In charge for Avoiding Network Congestion
- In charge for Traffic Shaping

Towards a Noah’s Ark for the Upcoming Data Deluge
HICP Key Needs
HICP Key Needs

- Reliability
HICP Key Needs

- Reliability
- Conciseness and Completeness
HICP Key Needs

- Reliability
- Conciseness and Completeness
- Security (Mostly Integrity)
Middleware Mission

- Network Coding;
Middleware Mission

- Network Coding;
- Logistical placement for data chunks;
Middleware Mission

- Network Coding;
- Logistical placement for data chunks;
- Deal with specific data and user needs;
Middleware Definition
Middleware Definition

The Problem:

The Problem: Optimal Replica Placement for storage and content distribution.

The Issues: Avoid Centralization, Performances, Balancing.

The Solution: LCN and VCN.
Middleware Definition

- The Problem: Optimal Replica Placement for storage and content distribution
Middleware Definition

• The Problem: Optimal Replica Placement for storage and content distribution
• The Issues:
Middleware Definition

- The Problem: Optimal Replica Placement for storage and content distribution
- The Issues: Avoid Centralization,
Middleware Definition

- The Problem: Optimal Replica Placement for storage and content distribution
- The Issues: Avoid Centralization, Performances,
 Middleware Definition

- The Problem: Optimal Replica Placement for storage and content distribution
- The Issues: Avoid Centralization, Performances, Balancing

Towards a Noah’s Ark for the Upcoming Data Deluge
Middleware Definition

- The Problem: Optimal Replica Placement for storage and content distribution
- The Issues: Avoid Centralization, Performances, Balancing
- The Solution:
Middleware Definition

- The Problem: Optimal Replica Placement for storage and content distribution
- The Issues: Avoid Centralization, Performances, Balancing
- The Solution: LCAN and VCAN
Outline

1. Intro
2. State of the art
3. INCA: The Concepts
   - Introducing the Intelligent Network Caching Architecture
   - The Architecture
4. INCA: The Implementation
   - The INCA Transport Protocol: HITP
   - The INCA Control Protocol: HICP
   - How to retrieve the stored data - the Middleware Layer
5. Outro

Towards a Noah’s Ark for the Upcoming Data Deluge
Future works: Short term planning

- Providing benchmark on real-case scenario
Future works: Short term planning

- Providing benchmark on real-case scenario
- An Ad-Hoc middleware module
Future works: Short term planning

- Providing benchmark on real-case scenario
- An Ad-Hoc middleware module
- MultimediaN applications
Future works: Short term planning

- Providing benchmark on real-case scenario
- An Ad-Hoc middleware module
- MultimediaN applications
- Developing further the protocols, bringing them to the most appropriate standardisation bodies (IETF, OGF).
Future works: Long term planning

- Extensive long-term technical todo list and road map
Future works: Long term planning

- Extensive long-term technical todo list and road map
- Cooperation with other entities, both academic and industrial
Future works: Long term planning

- Extensive long-term technical todo list and road map
- Cooperation with other entities, both academic and industrial
  and of course...
Future works: Long term planning

- Extensive long-term technical todo list and road map
- Cooperation with other entities, both academic and industrial
  and of course...
- World Domination!
Acknowledgement

We would like to thank:

Vincent Franceschini, from Hitachi Data Systems, for his deep insight on SAN/NAS, and for his enlightening suggestions,
Athanasios Christakidis and Nikolaos Efthymiopoulos, from University of Patras, for their precious suggestions and collaboration,
Frank Seinstra, MultimediaN, Vrije Universiteit,
Yi Zhu

And you maybe.

Towards a Noah’s Ark for the Upcoming Data Deluge
Acknowledgement

We would like to thank:

- Vincent Franceschini, from Hitachi Data Systems, for his deep insight on SAN/NAS, and for his enlighting suggestions,
Acknowledgement

We would like to thank:

- Vincent Franceschini, from Hitachi Data Systems, for his deep insight on SAN/NAS, and for his enlighting suggestions,
- Athanasios Christakidis and Nikolaos Efthymiopoulos, from University of Patras, for their precious suggestions and collaboration,
Acknowledgement

We would like to thank:

- Vincent Franceschini, from Hitachi Data Systems, for his deep insight on SAN/NAS, and for his enlighting suggestions,
- Athanasios Christakidis and Nikolaos Efthymiopoulos, from University of Patras, for their precious suggestions and collaboration,
- Frank Seinstra, MultimediaN, Vrije Universiteit, Yi Zhu
Acknowledgement

We would like to thank:

- Vincent Franceschini, from Hitachi Data Systems, for his deep insight on SAN/NAS, and for his enlighting suggestions,

- Athanasios Christakidis and Nikolaos Efthymiopoulos, from University of Patras, for their precious suggestions and collaboration,

- Frank Seinstra, MultimediaN, Vrije Universiteit, Yi Zhu

- And you
Acknowledgement

We would like to thank:

- Vincent Franceschini, from Hitachi Data Systems, for his deep insight on SAN/NAS, and for his enlighting suggestions,
- Athanasios Christakidis and Nikolaos Efthymiopoulos, from University of Patras, for their precious suggestions and collaboration,
- Frank Seinstra, MultimediaN, Vrije Universiteit, Yi Zhu
- And you maybe.